On some classes of optimal control problems governed by elliptic PDEs

Alexis de Villeroché

This internship was realized under the supervision of G. Buttazzo at the Università di Pisa from April to July 2022.

1 Introduction

We consider optimal control problems for a general integral cost functional, governed by an elliptic PDE over a bounded domain Ω of $\mathbf{R}^d$.

Statement of the problem - For m in a given class $\mathcal{M}$ of admissible controls that will be specified later, we define the state u_m as the unique solution in $W_0^{1,p}(\Omega)$ of the state equation

$$\begin{cases} -\Delta_p u + m\,|u|^{p-2}u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1}$$

where $p > 1$ and f is some fixed non negative function in $W^{-1,p}(\Omega)$.

Given u_m, we define the cost function that we wish to optimize over the class $\mathcal{M}$

$$J(m) = \int_\Omega j(x, u_m, \nabla u_m)\,dx$$

where j is a general integrand whose properties will also be discussed later. We consider two types of optimization problems: either the minimization and the maximization of J. These problems read as

$$\min_{m\in\mathcal{M}} J(m) \qquad \text{and} \qquad \max_{m\in\mathcal{M}} J(m).$$

The goal is to study the properties of the optimizer m^*.

Classes of admissible controls - We consider two different cases for the class $\mathcal{M}$ of controls m.

In the most general case, we only want to assume m to be non negative and that the state equation is well defined in the weak sense, namely,

$$\int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v\,dx + \int_\Omega |u|^{p-2}u\,v\,m\,dx = \int_\Omega f\,v\,dx.$$

for all v in the right set of test functions. If we want the class to be as large as possible, we need to relax it to a suitable class of Borel measures. This class is known to be the class of *capacitary measures*, that naturally depends on the exponent p.

After working on an example which shows that the relaxation to the larger class of capacitary measures is actually necessary, we consider the optimization problems to some restricted classes of controls:

$$\mathcal{M} = \left\{ m \in L^\infty(\Omega) \ : \ \alpha \leq m \leq \beta, \ \int_\Omega m \, dx = V \right\}$$

where α and β are two non negative constants, and $\alpha|\Omega| \leq V \leq \beta|\Omega|$.

This class of control appears naturally in several questions of distribution of resources, the mass constraint accounts for the fact there is only a finite amount of resources available, and the bounds account for environmental properties; the upper bound could be a point of saturation for resources and the lower bound a minimum requirement.

Properties of the optimizer - As explained before, we want to study the properties of the optimizer m^* of our problem.
When we consider the relaxed problem to p-capacitary measures, the first relevant property is whether m^* is of finite mass or not.
Then, the questions of regularity arise: is m^* a function or a general measure, and how much regularity does it have.

When we work with bounded controls, we will try to understand under which assumptions the optimizer is of a *bang-bang* type.

Definition 1. *We say that $m \in L^\infty(\Omega)$ such that $\alpha \leq m \leq \beta$ has the bang-bang property if it is of the form*

$$m = (\beta - \alpha)\mathbf{1}_E + \alpha,$$

where E is a Borel subset of Ω.

We will then study the properties of the set E. We will show that in the general case, E is p-quasi open (see Definition 2 below) and we will look for the cases where E is open, we will complete the study by trying to determine whether E is of finite perimeter. Then, the further step would consist in studying the regularity of the set E; this is usually a very delicate issue that we will not consider here.

Definition 2. *p-quasi open set*
A set E is said to be p-quasi open if it is a preimage of an open subset of Ω by an element of $W^{1,p}(\Omega)$.

When $p > d$, since Sobolev functions are Hölder continuous, p-quasi open set simply reduce to open sets.

2 Relaxed problem and p-capacitary measures

The goal of this section is to illustrate the relaxed framework of p-capacitary measures, that are the most natural objects to consider in our optimization problems. In the following we use the notion of p-capacity of a set E:

$$Cap_p(E) = \inf\Big\{ \int |\nabla u|^p + |u|^p \, dx \ : \ u \in \mathcal{U}_E \Big\},$$

where $\mathcal{U}_E$ is the set of all functions u of the Sobolev space $H^1(\mathbf{R}^d)$ such that $u \geq 1$ almost everywhere in a neighborhood of E.
From this we can first define the notion of being p-quasi continuous

Definition 3. *A function f on Ω is said to be p-quasi continuous if for every $\epsilon > 0$ there exists a continuous function f_ϵ such that*

$$Cap_p\big(\big\{x \in \Omega, \ f(x) \neq f_\epsilon(x)\big\}\big) < \epsilon$$

We also define the p-capacitary measures which are the Borel measures absolutely continuous with respect to the p-capacity

Definition 4. *p-capacitary measure*
A Borel measure μ on Ω, taking its values in $[0, +\infty]$, is said to be p-capacitary if for every Borel set $E \subset \Omega$

$$\begin{cases} Cap_p(E) = 0 \Rightarrow \mu(E) = 0 \\ \mu(E) = \inf\left\{\mu(A) \ : \ A \supset E, \ A \ p\text{-quasi open}\right\}. \end{cases}$$

Again, when $p > d$, due to the fact that all nonempty sets have a strictly positive capacity, p-capacitary measures simply reduce to Borel measures taking their values in $[0, +\infty]$. On the contrary, when $p \leq d$, points and more generally $d-2$ dimensional sets have capacity zero; as a consequence, Dirac masses or more generally measures concentrated on $d-2$ dimensional sets, are not p-capacitary measures.

The important result is that any Sobolev function in $W^{1,p}(\Omega)$ has a p-quasi continuous representative, uniquely defined up to a zero p-capacity set. Then as explained in [1] we can make sense of a weak solution to (1) u_μ for μ a p-capacitary measure such that

$$\int_\Omega |\nabla u_\mu|^{p-2} \nabla u_\mu \cdot \nabla v \, dx + \int_\Omega |u_\mu|^{p-2} u_\mu \, v \, d\mu = \int_\Omega f \, v \, dx.$$

for all test function v in $W^{1,p}_{\mu,0}(\Omega) = \Big\{ v \in W^{1,p}_0(\Omega), \int_\Omega |v|^p \, d\mu < \infty \Big\}$. We will then relax the initial problem by considering weak solution to (1) in this sense.

3 Existence

In this section we look at the existence of optimal solutions, in the different cases, for our problems. We begin by stating two general properties that do not depend on the class $\mathcal{M}$ of controls that we consider. For the sake of clearness we will only work with the min problem here even though all the results are of course adaptable for the max problem

The first property just states that the weak max principle holds.

Proposition 1. *For a given control m in $\mathcal{M}$, the state u_m satisfies the weak max principle and is non negative.*

Proof -The state u_m can be characterized as the solution of the variational problem

$$\min_{u \in W_0^{1,p}(\Omega)} F_m(u),$$

with

$$F_m(u) = \int_\Omega \left(\frac{1}{p}|\nabla u|^p + \frac{1}{p} m|u|^p - fu \right).$$

Consider for $u \in W_0^{1,p}(\Omega)$, $\tilde{u} = \max\{u, 0\} \in W_0^{1,p}(\Omega)$, then $|\nabla \tilde{u}| \leq |\nabla u|$, $|\tilde{u}| \leq |u|$ and $\tilde{u} \geq u$, so, since the function f and the control m are non negative, $\tilde{u}$ is a better candidate than u to the problem, hence u_m must be non negative. □

The second property gives a partial order on the set of states; it was proven in [1] Proposition 3.3 and we only state here the first part of the property.

Proposition 2. *The map $m \mapsto u_m$ is decreasing, namely,*

$$u_{m_1} \leq u_{m_2} \qquad \textit{whenever } m_1 \geq m_2.$$

Proof - This proof is directly taken from [1]. For simplicity of notations, here we write $u_i = u_{m_i}$ and the corresponding functional $F_i = F_{m_i}$. As for the max principle, the proof uses the variational formulation of the state equation.
If we prove that $F_2(u_1 \wedge u_2) \leq F_2(u_2)$, by minimality and uniqueness, we would have $u_1 \wedge u_2 = u_2$ concluding the proof. Noticing that

$$F_i(u_1 \wedge u_2) + F_i(u_1 \vee u_2) = F_i(u_2) + F_i(u_1),$$

it is equivalent to prove $F_2(u_1) \leq F_2(u_1 \vee u_2)$. Rewriting F_2 in term of F_1, it is equivalent to showing

$$F_1(u_1) \leq F_1(u_1 \vee u_2) + \int_\Omega \frac{1}{p} \left(|u_1 \vee u_2|^p - |u_1|^p\right)(m_2 - m_1).$$

This last inequality is true since u_1 minimizes F_1 and the integral is non negative by the hypothesis on m_1 and m_2. □

Remark - In particular from these two properties, we deduce that for all $m \in \mathcal{M}$, u_m must satisfy the inequalities $0 \leq u_m \leq u_0$ with u_0 the solution in $W_0^{1,p}(\Omega)$ of

$$-\Delta_p u = f.$$

When $\mathcal{M}$ is the set of p-capacitary measures, we prove that the set of solutions u_m is compact for the weak $W^{1,p}$ convergence so we only need mild assumptions on the integrand of the cost j, which only needs to provide the lower semicontinuity to the integral functional. It is then enough to require that $j(x,s,z)$ be measurable in x, lower semicontinuous in (s,z) and convex in z.

Let's write the state equation,

$$-\Delta_p u_m + m|u_m|^{p-2}u_m = f.$$

Since by Proposition 2, the term $m|u_m|^{p-2}u_m \geq 0$, it implies

$$-\Delta_p u_m \leq f.$$

Now, considering

$$E = \left\{u \in W_0^{1,p}(\Omega) \ : \ -\Delta_p u \leq f,\ 0 \leq u \leq u_0\right\},$$

we have the following compactness result.

Theorem 1. *If $p > d$ and $f \in L^1(\Omega)$, or if $p \leq d$ and $f \in L^{p^*/(p^*-1)}(\Omega)$ with $p^* = pd/(d-p)$ (any $p^* < +\infty$ if $p = d$), then the set E is compact for the weak $W_0^{1,p}(\Omega)$ convergence.*

Proof - If $u \in E$ the p-Laplacian condition implies

$$\int_\Omega |\nabla u|^p\,dx \leq \int_\Omega f u\,dx,$$

then applying Hölder inequality, if $p < d$,

$$\int_\Omega |\nabla u|^p \leq \left(\int_\Omega f^{\frac{p^*}{p^*-1}}\right)^{\frac{p^*-1}{p^*}} \left(\int_\Omega |u|^{p^*}\right)^{\frac{1}{p^*}} \leq \left(\int_\Omega f^{\frac{p^*}{p^*-1}}\right)^{\frac{p^*-1}{p^*}} \left(\int_\Omega |u_0|^{p^*}\right)^{\frac{1}{p^*}},$$

here, the fact that u and u_0 are in $L^{p^*}(\Omega)$ is a consequence of the Sobolev embedding theorem. The second inequality is true because $0 \leq u \leq u_0$.

If $p > d$, u and u_0 are Holder continuous on Ω and since they are null on the boundary, they are both continuous on $\bar{\Omega}$, then they are necessarily bounded on Ω with $\|u\|_\infty \leq \|u_0\|_\infty$, then we obtain the inequality

$$\int_\Omega |\nabla u|^p\,dx \leq \|u_0\|_\infty \int_\Omega f\,dx$$

Then, in both cases, since E is closed, it is compact for the weak $W_0^{1,p}(\Omega)$ convergence. $\square$

We deduce form this result that up to extraction, all sequence of states (u_m) must converge in E. But actually, if one takes $u \in E$,

$$\mu_u = \begin{cases} \dfrac{f + \Delta_p u}{|u|^{p-2}u} & \text{if } u > 0, \\ +\infty & \text{if } u = 0, \end{cases}$$

is a p-capacitary measure. Then, if (m_n) is an optimizing sequence, up to extraction (u_{m_n}) converges weakly to u and we can construct μ_u a p-capacitary measure such that $u = u_{\mu_u}$. So we have proven that the set of states is compact for the weak convergence and we only need the asssumptions described at the beginning on j to have existence.

Remark - This approach of the problem is really interesting because we forget about the structure of the control and only consider the set of states as a subset of $W_0^{1,p}(\Omega)$ with explicit constraints. The problem can then be viewed as standard variational problem and, more importantly, solved as such.

This result is even better in the case $p = 2$ because then by linearity, the condition on the Laplacian becomes

$$\Delta(u_m - u_0) \geq 0.$$

So, the variational problem that we consider can be viewed as a problem on the set of sub-harmonic functions v null on the boundary and bounded below by $-u_0$. More precisely, setting $v = u_m - u_0$ we have that our problem reduces to the optimization of the functional

$$\tilde{J}(v) = \int_\Omega j\big(x, v + u_0(x), \nabla v + \nabla u_0(x)\big)\, dx$$

on the class $\big\{v\ :\ \Delta v \geq 0,\ -u_0 \leq v \leq 0\big\}$. In particular, in the one-dimensional case, the class above is the class of convex functions $\big\{v \text{ convex}\ :\ -u_0 \leq v \leq 0\big\}$.

One may also notice that to prove this existence result, we have proven that the class of capacitary measures is compact for the γ_p convergence introduced in [1] (Definition 3.4) namely a sequence (m_k) of p-capacitary measures is said to γ_p converge to m if the sequence of associated states (u_{m_k}) converges strongly in L^p to u_m.

Now, when we work with the class of bounded control functions $\alpha \leq m \leq \beta$, we can actually have the exact same approach the Proposition 2 imposes that the admissible states live in

$$\tilde{E} = \big\{u \in W_0^{1,p}(\Omega)\ :\ f - \beta|u|^{p-2}u - \beta - \Delta_p u \leq f - \alpha|u|^{p-2}u,\ u_\beta \leq u \leq u_\alpha\big\}.$$

Since $\tilde{E}$ is a closed subset of E, it is of course also compact for the weak $W_0^{1,p}(\Omega)$ convergence. And if one takes an element u of $\tilde{E}$, it is bounded from below by u_β which is positive on Ω. Then, as before, considering

$$\mu_u = \frac{f + \Delta_p u}{|u|^{p-2}u},$$

the bounds on the p-Laplacian of u immediately imply $\alpha \leq \mu_u \leq \beta$. Then, once again, the set of admissible states is compact for the weak $W_0^{1,p}(\Omega)$ convergence and we need the same assumptions on the integrand j to have existence of an optimizer.

4 Dimension 1

4.1 Example $f = 1$

As a first example, we will work in dimension 1 with p = 2 optimizing over the set of 2-capacitary measures that coincide with the Borel measures in this case. For the sake of simplicity, we choose $f = 1$. In this setting, the state equation is the following

$$\begin{cases} -u'' + m\,u = 1, & \text{in }]-1,1[, \\ u(-1) = u(1) = 0. \end{cases} \tag{2}$$

Taking c a constant, we define the functional J for a given m

$$J(m) = \int_{-1}^{1} |u_m - c|^2 .$$

And we want to solve

$$\min_m J(m).$$

We will use the convex formulation introduced in the previous section to prove the following theorem

Theorem 2. *The functional J admits a unique minimizer m^*. And the dependence on c is such that*
- if $c \leq 0$, $m^* = +\infty$,
- if $0 < c < \frac{1}{4}$,

$$m^* = \frac{1}{c}\mathbf{1}_{]-\gamma,\gamma[} + \frac{\sqrt{c}}{2}\left(\delta_{-\gamma} + \delta_{\gamma}\right),$$

where $\gamma = 1 - 2\sqrt{c}$,
- if $\frac{1}{4} \leq c < \frac{5}{12}$,

$$m^* = 3\left(\frac{5}{12} - c\right)\delta_0,$$

- if $c \geq \frac{5}{12}$, $m^* = 0$.

Proof

The first point is trivial, if $c \leq 0$, the minimizing state is of course $u_{m^*} = 0$ and $m^* = +\infty$. Remember that by Proposition 2, we know that $0 \leq u_m \leq u_0$, with

$$u_0(x) = -\frac{x^2 - 1}{2}$$

by direct calculation. The last point of the theorem is just a direct consequence of this inequality when $c \geq \frac{1}{2}$ since $\sup u_0 = \frac{1}{2}$.

The main part of the theorem is to prove the other points ($0 < c < 1/2$), to do so we will rewrite the problem in the convex formulation. For a given control m, the function $\phi = u_m - u_0$ is convex, satisfying $-u_0 \leq \phi \leq 0$ and $\phi(-1) = \phi(1) = 0$. And as explained for a given convex function ψ satisfying these properties, we can construct back an admissible control m,

$$m = \begin{cases} \dfrac{\psi''}{\psi + u_0} & \text{if } \psi \neq -u_0, \\ +\infty & \text{if } \psi = -u_0. \end{cases}$$

Then our problem rewrites equivalently as

$$\min_{\phi} J(\phi),$$

with

$$J(\phi) = \int_{-1}^{1} \left| \phi(x) - \left(c + \frac{x^2 - 1}{2} \right) \right|^2,$$

where the functions ϕ are taken convex satisfying $\phi(-1) = \phi(1) = 0$ and $\phi \geq -u_0$. Note that in this setting the fact that $\phi \leq 0$ is contained in the convexity and its value on the boundary.

Uniqueness- We can note that the functional is strictly convex defined on a convex set so the minimizer is unique. It only remains to prove that the solution is of the expected form.

Our first step is to restrain the problem to two types of symmetric functions,

$$v_y(x) = \begin{cases} a(y)\,(1+x) & \text{for } x \in [-1, -y] \\ \dfrac{x^2 - 1}{2} + c & \text{for } x \in [-y, y] \\ a(y)\,(1-x) & \text{for } x \in [y, 1] \end{cases}$$

here, for a given y, $a(y)$ is explicitly given by the continuity of v_y and we choose y chosen such that v_y stays convex, and,

$$w_h(x) = \begin{cases} h\,(1+x) & \text{for } x \in [-1, 0] \\ h\,(1-x) & \text{for } x \in [0, 1] \end{cases}$$

with $c - \dfrac{1}{2} \leq h \leq 0$.

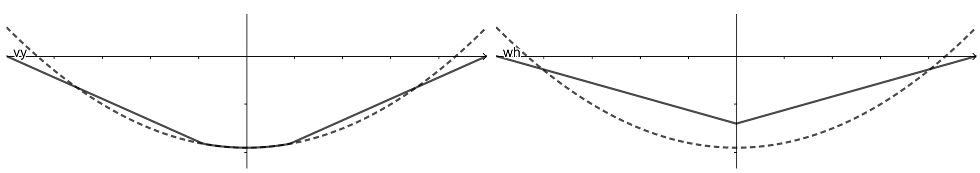


Figure 1: example of functions v_y and w_h

Then, we will derive the different conditions en c that give the expression of the minimizer

Reduction of the problem - Consider a generic admissible function ϕ, if we can find a better candidate of either of the previous forms on each halves of the space, namely ψ_r and ψ_l such that

$$\int_0^1 \left|\phi(x) - \left(\frac{x^2-1}{2} + c\right)\right|^2 dx - \int_0^1 \left|\psi_r(x) - \left(\frac{x^2-1}{2} + c\right)\right|^2 dx \geq 0$$
$$\int_{-1}^0 \left|\phi(x) - \left(\frac{x^2-1}{2} + c\right)\right|^2 dx - \int_{-1}^0 \left|\psi_l(x) - \left(\frac{x^2-1}{2} + c\right)\right|^2 dx \geq 0$$

Then

$$J(\phi) \geq \min\{J(\psi_l), J(\psi_r)\}$$

By symmetry of the problem we only need to prove the existence of ψ_r.

In the following we need to define two constants $\alpha = 1 - \sqrt{2c}$ and $\beta = 1 - 2c$. The constant α is the smaller root of the polynomial

$$X^2 - 2\,X + 1 - 2c,$$

and is such that

$$v(x) = v_\alpha(x) = \begin{cases} -\alpha(x+1) & \text{for } x \in [-1, -\alpha] \\ \dfrac{x^2-1}{2} + c & \text{for } x \in [-\alpha, \alpha] \\ \alpha(x-1) & \text{for } x \in [\alpha, 1] \end{cases}$$

is C^1.
The constant β solves

$$\frac{\beta^2-1}{2} + c = \left(\frac{1}{2} - c\right)(\beta - 1),$$

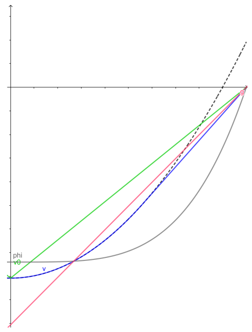
it is the x-coordinate of the right intersection point between v_0 and $-u_0$.

We will denote by $J_r(\phi) = \displaystyle\int_0^1 \left|\phi(x) - \left(\frac{x^2-1}{2} + c\right)\right|^2 dx$ the cost of ϕ on $[0,1]$. Take ϕ an admissible convex function, and define x_0 such that

$$\phi(x_0) = \frac{{x_0}^2 - 1}{2} + c$$

and $\forall x > x_0$, $\phi(x) < \dfrac{x^2-1}{2} + c$.
If there is no such x_0 or $x_0 \leq \alpha$, then $\phi(\alpha) \leq \dfrac{\alpha^2-1}{2} + c$, and by convexity of ϕ, $\forall x \in [\alpha, 1]$,

$$\begin{aligned} \phi(x) &\leq \frac{\phi(\alpha)}{1-\alpha}(1-x) \\ &\leq \alpha\,(1-x) \\ &\leq v(x) \end{aligned}$$



by definition of α and v, then $J_r(\phi) \geq J_r(v)$.

If $x_0 \geq \alpha$, by convexity of ϕ,

$$\begin{aligned} \forall x \geq x_0, \quad \phi(x) &\leq \frac{\phi(x_0)}{1-x_0}(1-x) \\ \forall x < x_0 \quad \phi(x) &\geq \frac{\phi(x_0)}{1-x_0}(1-x). \end{aligned}$$

By convexity of $x \mapsto \dfrac{x^2-1}{2} + c$,

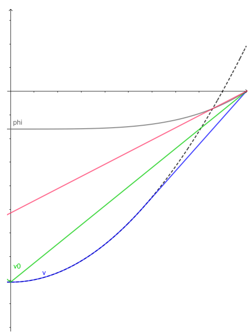
$$\forall x \geq x_0, \quad \frac{\phi(x_0)}{1-x_0}(1-x) \leq \frac{x^2-1}{2} + c,$$

if $x_0 \geq \beta$, the convexity also imposes

$$\forall x \leq x_0, \quad \frac{\phi(x_0)}{1-x_0}(1-x) \geq \frac{x^2-1}{2} + c,$$

then, we deduce $J_r(\phi) \geq J_r(w)$, with

$$w(x) = \begin{cases} \dfrac{\phi(x_0)}{1-x_0}(1+x) & \text{for } x \in [-1,0] \\ \dfrac{\phi(x_0)}{1-x_0}(1-x) & \text{for } x \in [0,1] \end{cases}$$

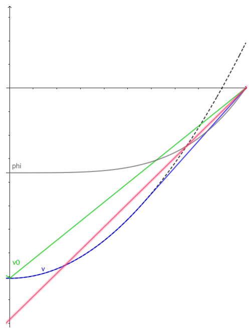


If $x_0 < \beta$, there exists a positive $y \neq x_0$ such that

$$\frac{\phi(x_0)}{1-x_0}(1-y) = \frac{y^2-1}{2} + c,$$

then the convexity imposes

$$\forall x \in [y, x_0], \quad \frac{\phi(x_0)}{1-x_0}(1-x) \geq \frac{x^2-1}{2} + c,$$



then $J_r(\phi) \geq J_r(w)$, with

$$w(x) = \begin{cases} \frac{\phi(x_0)}{1-x_0}(1-x) & \text{for } x \in [-1, -y] \\ \frac{x^2-1}{2} + c & \text{for } x \in [-y, y] \\ \frac{\phi(x_0)}{1-x_0}(1-x) & \text{for } x \in [y, 1]. \end{cases}$$

It remains to prove that w is convex, to do so we just need to check that $y \leq -\frac{\phi(x_0)}{1-x_0}$.

$$\begin{aligned} -\frac{\phi(x_0)}{1-x_0} &= \frac{\phi(x_0) - \frac{\phi(x_0)}{1-x_0}(1-y)}{x_0 - y} \\ &= \frac{\frac{{x_0}^2-1}{2} + c - \left(\frac{y^2-1}{2} + c\right)}{x_0 - y} && \text{by definition of } x_0 \text{ and y} \\ &= \frac{x_0 + y}{2} \\ &\geq y && \text{since } x_0 \geq y. \end{aligned}$$

We have now proven that we can restrain our search for a minimizer of the forms v_y of w_h.

Computation of the explicit minimizer - Define for $h \in [c - 1/2, 0]$

$$F(h) = J_r(w_h) = \frac{1}{2} J(w_h).$$

Then we can directly compute

$$\begin{aligned} F(h) &= h^2 \int_0^1 (1-x)^2 dx - h \int_0^1 (1-x)(x^2 + 2c - 1) dx + \int_0^1 \left(\frac{x^2-1}{2} + c\right)^2 dx \\ &= \frac{1}{3} h^2 - \left(c - \frac{5}{12}\right) h + D. \end{aligned}$$

The minimum of this polynomial is obtained for $h^* = \frac{3}{2}\left(c - \frac{5}{12}\right)$ we deduce that there are several cases,
if $c < 1/4$,, $h^* < c - 1/2$ and the best candidate is $w_{c-1/2} = v_0$,
if $1/4 \leq c \leq 5/12$, h^* is inside the domain and the best candidate is w_{h^*},
and if $5/12 < c$, h^* is positive and the best candidate is $w_0 = 0$.

Now, we need to compare these possible candidates with the functions of the form v_y. By definition,

$$v_y(x) = \begin{cases} a(y)\,(1+x) & \text{for } x \in [-1, -y] \\ \dfrac{x^2-1}{2} + c & \text{for } x \in [-y, y] \\ a(y)\,(1-x) & \text{for } x \in [y, 1] \end{cases}$$

Where y is chosen so that v_y is convex, which imposes the two following conditions,

$$\begin{aligned} y &\leq -a(y) \\ \frac{y^2-1}{2} + c &= a(y)\,(1-y), \end{aligned}$$

which put together imply

$$y^2 - 2y + 1 - 2c \geq 0.$$

Since α is the smaller root of this polynomial, the two conditions rewrite

$$\begin{aligned} 0 \leq y \leq \alpha \\ a(y) = \frac{y^2 - 1 + 2c}{2\,(1-y)}. \end{aligned}$$

As previously, we define

$$\tilde{F}(y) = J_d(v_y) = \frac{1}{2} J(v_y),$$

and by definition,

$$\tilde{F}(y) = \int_y^1 \left| a(y)(1-x) - \left(\frac{x^2-1}{2} + c\right)\right|^2 dx.$$

We can derive $\tilde{F}$,

$$\tilde{F'}(y) = -\left| a(y)(1-y) - \left(\frac{y^2-1}{2} + c\right)\right|^2 + 2a'(y) \int_y^1 (1-x) \left(a(y)(1-x) - \left(\frac{x^2-1}{2} + c\right)\right).$$

By definition of a, the first term is null and

$$a'(y) = \frac{1}{1-y}(y + a(y)).$$

Then,

$$\begin{aligned} \tilde{F}'(y) &= 2\,a'(y) \left(\frac{1}{3} a(y)\,(1-y)^3 - \frac{1}{4}\,(2c-1)\,(1-y)^2 - \frac{1}{6}\,(1-y^3) + \frac{1}{8}\,(1-y^4)\right) \\ &= \frac{1}{12}\,a'(y)\,(1-y)\,((1-y)^3 - 4c(1-y)) \\ &= \frac{-1}{24}\,(y^2 - 2y + 1 - 2c)\,(y - (1 + 2\sqrt{c}))\,(y - (1 - 2\sqrt{c})). \end{aligned}$$

the first factor has two roots α and the other being larger than 1 so it is non negative for $y \in [0, \alpha]$ and the second factor stays negative so this derivative is of the sign of $(y - (1 - 2\sqrt{c}))$ and there are two cases, If $c < \dfrac{1}{4}$ then the best candidate is $v_{1-2\sqrt{c}}$ and if $c \geq \dfrac{1}{4}$, the candidate becomes v_0.

Finally, regrouping the two sets of results, we have found the explicit minimizer depending on the value of c.

It remains to compute the associated optimal m^* in the different cases,
- if $c < \dfrac{1}{4}$, introducing $\gamma = 1 - 2\sqrt{c}$,

$$m^* = \frac{w_\gamma''}{w_\gamma - \frac{x^2-1}{2}},$$

then we can write

$$m^* = \mathbf{1}_{]-\gamma,\gamma[}\frac{1}{c} + \frac{\sqrt{c}}{2}\left(\delta_{-\gamma} + \delta_\gamma\right)$$

- if $\dfrac{1}{4} \leq c < \dfrac{5}{12}$,

$$m^* = 3\left(\frac{5}{12} - c\right)\delta_0.$$

- if $c \geq \dfrac{5}{12}$, $m^* = 0$. □

4.2 Generalisation to generic function f

The previous result is actually exactly the same for a generic symmetric non negative L^2 function f. The proof works exactly the same way, first the reduction then the computation of the explicit minimizer, the main difference is in this computation we cannot express the explicit solution but we can show its existence and how to find it.

The precise setting is the following, consider for $f \in L^2(]-1,1[)$ non negative, the state equation

$$\begin{cases} -u'' + m\,u = f, & \text{in }]-1,1[, \\ u(-1) = u(1) = 0, \end{cases}$$

and for c a real number, consider the functional

$$J_f(m) = \int_{-1}^{1} |u_m - c|^2.$$

Then, the following theorem holds

Theorem 3. *If f is symmetric, the functional J_f admits a unique minimizer m^*. And the dependence on c is such that*

- *if* $c \leq 0$, $m^* = +\infty$,
- *if* $0 < c < \gamma - \lambda$,

$$m^* = \frac{f}{c}\mathbf{1}_{]-y^*,y^*[} + C(y^*)\left(\delta_{-y^*} + \delta_{y^*}\right),$$

where y^* *is the smallest solution on* $[0,1]$ *of*

$$\int_y^1 (1-x)\, f(x)\, dx - \int_y^1 (1-x)^3\, f(x)\, dx = c$$

and,

$$C(y^*) = \int_{y^*}^1 \frac{1-x}{1-y^*}\, f(x)\, dx,$$

- *if* $\gamma - \lambda \leq c < \gamma - \dfrac{\lambda}{3}$,

$$m^* = 3\left(\left(\gamma - \frac{\lambda}{3}\right) - c\right)\delta_0,$$

- *if* $c \geq \gamma - \dfrac{\lambda}{3}$, $m^* = 0$.

With,

$$\gamma = \int_0^1 (1-x)\, f(x)\, dx \quad \textit{and} \quad \lambda = \int_0^1 (1-x)^3\, f(x)\, dx.$$

Proof - the proof is basically the same as the previous one, we just need to change the expression of u_0. One can check that

$$u_0(x) = (1-x)\int_0^x f(y)\, dy + \int_x^1 (1-y)\, f(y)\, dy.$$

From this expression, appears $\gamma = \sup u_0 = u_0(0)$, it will play the role of $\frac{1}{2}$ in the previous demonstration. With this different expression, we must redefine the functions v_y accordingly, namely

$$v_y(x) = \begin{cases} a(y)\,(1+x) & \text{for } x \in [-1,-y] \\ c - u_0(x) & \text{for } x \in [-y,y] \\ a(y)\,(1-x) & \text{for } x \in [y,1] \end{cases}$$

with $a(y) = \dfrac{c - u_0(y)}{1-y}$. As before, one can define α as the smaller positive solution of $a(y) = u_0'(y)$, there is a solution because since f is L^2, u_0 is H^2 which imply C^1 in dimension 1. As before, v_y will be well defined if $y \leq \alpha$.

With this new setting, using the convex formulation, one can make the same reduction as before with the same cases.

Then, as previously on can define, for $c - \gamma \leq h \leq 0$,

$$F(h) = \frac{1}{3}h^2 - 2\int_0^1 (1-x)(c - u_0(x))dx + D(u_0)$$

With the minimality of the function F obtained for

$$h^* = 3\int_0^1 (1-x)(c-u_0(x))dx,$$

which after two integration by parts can be expressed as

$$h^* = \frac{3}{2}\left(c - \left(\gamma - \frac{\lambda}{3}\right)\right).$$

Again as previously, one can define, for $0 \le y \le \alpha$, $\tilde{F}(y)$, after much computation, we find that the derivative $\tilde{F}'$ of $\tilde{F}$ satisfies

$$\tilde{F}'(y) = -\frac{1}{3}a'(y)(1-y)^2(c-h(y)).$$

with

$$h(y) = \int_y^1 (1-x)f(x)\,dx - \frac{1}{(1-y)^2}\int_y^1 (1-x)^3 f(x)\,dx.$$

One can actually check that, h is decreasing, $h(0) = \gamma - \lambda$ and, $h(\alpha) = c - \frac{1}{(1-\alpha)^2}\int_\alpha^1 (1-x)^3 f(x)\,dx$. Then, since $a'(y)$ is negative, if $c \ge \gamma - \lambda$, the minimality of the function $\tilde{F}$ is obtained for $y^* = 0$, conversely it is obtained for y^* satisfying $0 < y^* < \alpha$ and $h(y^*) = c$.
Finally, putting all the pieces together one can finish the proof.

□

Remarks The general remark that should be made about this result is that in the generic case, except when c in non positive, the total mass stays finite and except when it hits 0 the minimizer is never a function, So clearly, there is no hope of finding *bang-bang* minimizer to this problem.

The obvious question is how do we go from here to non symmetric functions f. The complication is that we will not be able to reduce the problem to symmetric functions but rather to functions that are a mix of the two forms w_h and v_y on either side of the minimal point of $c - u_0$. Yet, we can expect there will still be the same type range on c that appear. For small positive c there should be two dirac measures one on each side of the minimal point with $\frac{f}{c}$ in between , then probably that a new range will appear when one of the dirac arrives first at the minimal point and finally the two dirac combine and then when c gets above $\sup u_0$ the solution will of course be 0.

This scheme of proof can probably also be adapted to functionals of the form

$$J(m) = \int_{-1}^1 |u_m - g|^2,$$

with $g \in H^2$ symmetric and such that $g - u_0$ stays convex and is not non negative. The reduction should still hold because it only uses the fact $c - u_0$ is convex and

C^1 and then there should be conditions on g'' that will describe the shape of the minimizer.

Finally, this proof is not transferable to higher dimension because the equivalent variational problem would be with subharmonic functions instead of convex and the reduction as it is would not work especially if the domain Ω is too pathological.

5 Energy functional and Bang-Bang property

In this section we go back to the problem in multiple dimensions and we consider the energy functional for the *p-laplacian* problem

$$J(m) = \frac{1-p}{p} \int_\Omega f u_m,$$

Where u_m solves the variational formulation of the state equation

$$\min_{u \in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + m\frac{|u|^p}{p} \right).$$

Here, we consider the class of controls m for $0 \leq \alpha \leq \beta$,

$$\mathcal{M} = \left\{ m \in L^\infty(\Omega) \middle| \alpha \leq m \leq \beta, \text{ and } \int_\Omega m = V \right\},$$

with V chosen so that the set is not empty.
In order to simplify this class, we drop the mass constraint by adding a Lagrange multiplier in the functional

$$J(m) = \frac{1-p}{p} \int_\Omega f u_m - \lambda \int_\Omega m.$$

In order to have a more general result, we replace this multiplier by an integral cost c which is not necessarily linear. And the functional then becomes

$$J(m) = \frac{1-p}{p} \int_\Omega f u_m - \int_\Omega c(x,m).$$

We are interested in both the minimum problem and the maximum problem, namely

$$\min / \max_{\alpha \leq m \leq \beta} J(m).$$

with $0 \leq \alpha \leq \beta$. Since at the minimum u_m, we have

$$\frac{1-p}{p} \int_\Omega f u_m = \int_\Omega \left(\frac{|\nabla u_m|^p}{p} - fu + m\frac{|u_m|^p}{p} \right),$$

we can separate both variables in our problem which rewrites

$$\min / \max_{\alpha \leq m \leq \beta} \min_{u \in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + m\frac{|u|^p}{p} - c(x,m) \right).$$

Now, in the case where we consider the minimum problem we can exchange the two min operators and we are left with

$$\min_{u\in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + \min_{\alpha\le m\le\beta} \left\{ m\frac{|u|^p}{p} - c(x,m) \right\} \right).$$

And in the case of the maximum problem, if $m \mapsto c(x,m)$ is convex we can make the same operation and get

$$\min_{u\in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + \max_{\alpha\le m\le\beta} \left\{ m\frac{|u|^p}{p} - c(x,m) \right\} \right),$$

note that this convexity hypothesis still includes the linear case we started with. This exchange of the max and min operator can be justified by Sion's minimax theorem (see [2]).

There are a few questions that we want to answer. We want to know under which hypothesis on c do we get the *bang-bang* property for the optimizer m^*. Then we want to determine under which conditions the optimal set E is open.

5.1 Hypothesis on c to have the *bang-bang* property

We wish to establish a criteria on c such that the optimal m is *bang-bang* for all $u \in W_0^{1,p}(\Omega)$. We will only do the calculation for the minimisation problem since they are very similar to the maximisation problem.

We want to find a sufficient and necessary condition on c such that for all non negative real number u the minimizer of $h_u(m) = \frac{u^p}{p} m_c(m)$ on $\alpha \le m \le \beta$ is in $\{\alpha, \beta\}$.

Firstly, let's note that $h_u(\alpha) \le h_u(\beta)$ if and only if

$$\frac{u^p}{p} \ge \frac{c(\beta)-c(\alpha)}{\beta-\alpha},$$

and in that case, we want for all m in $[\alpha,\beta]$, $h_u(\alpha) \le h_u(m)$, which translates as

$$\frac{u^p}{p} \ge \frac{c(m)-c(\alpha)}{m-\alpha},$$

so in particular if $c(\beta)-c(\alpha) > 0$ taking

$$\frac{u^p}{p} = \frac{c(\beta)-c(\alpha)}{\beta-\alpha}.$$

We have that c must satisfy for all m

$$\frac{c(\beta)-c(\alpha)}{\beta-\alpha} \ge \frac{c(m)-c(\alpha)}{m-\alpha}.$$

And if $c(\beta) - c(\alpha) \leq 0$ taking $u = 0$ we have that c must satisfy for all m

$$\frac{c(m) - c(\alpha)}{m - \alpha} \leq 0.$$

Which rewrites as for all m,

$$c(m) \leq \max\left\{c(\alpha), \frac{c(\beta) - c(\alpha)}{\beta - \alpha}(m - \alpha) + c(\alpha)\right\}.$$

Now considering the case when $c(\beta) - c(\alpha) > 0$, where

$$\frac{u^p}{p} < \frac{c(\beta) - c(\alpha)}{\beta - \alpha},$$

we want to have for all m, $h_u(m) \geq h_u(\beta)$ which rewrites as

$$\frac{u^p}{p} \leq \frac{c(\beta) - c(m)}{\beta - m}.$$

then in particular, taking the equality for the condition on u, it implies that for all m

$$\frac{c(\beta) - c(\alpha)}{\beta - \alpha} \leq \frac{c(\beta) - c(m)}{\beta - m},$$

which rewrites as

$$c(m) \leq \frac{c(\beta) - c(\alpha)}{\beta - \alpha}(m - \alpha) + c(\alpha)$$

We have now proven that if for all u the minimizer of h_u lies in $\{\alpha, \beta\}$ then, c must satisfy for all m

$$c(m) \leq \max\left\{c(\alpha), \frac{c(\beta) - c(\alpha)}{\beta - \alpha}(m - \alpha) + c(\alpha)\right\}.$$

We will now prove the converse, we assume that c satisfies this condition.
If $c(\alpha) \geq c(\beta)$, since for all m, $c(m) \leq c(\alpha)$ we immediately get $h_u(m) \geq h_u(\alpha)$ for all u and m.
If $c(\alpha) < c(\beta)$, the condition is equivalent to both formulation

$$\forall m, \ \frac{c(\beta) - c(\alpha)}{\beta - \alpha} \geq \frac{c(m) - c(\alpha)}{m - \alpha} \quad \forall m \ \frac{c(\beta) - c(\alpha)}{\beta - \alpha} \leq \frac{c(\beta) - c(m)}{\beta - m},$$

Then, depending on if u is larger or smaller than $\frac{c(\beta) - c(\alpha)}{\beta - \alpha}$ the minimiser is α or β. We have now proven the following theorem

Theorem 4. *The minimizer of*

$$m \mapsto \frac{u^p}{p}m - c(m)$$

in $[\alpha, \beta]$ lies in $\{\alpha, \beta\}$ for all non negative u if and only if c satisfies

$$\forall\, m, \qquad c(m) \leq \max\left\{c(\alpha), \frac{c(\beta)-c(\alpha)}{\beta-\alpha}(m-\alpha)+c(\alpha)\right\}.$$

And the minimizer is α if

$$\frac{u^p}{p} \geq \frac{c(\beta)-c(\alpha)}{\beta-\alpha},$$

and β otherwise.

In a similar fashion one can derive the criteria for the maximization problem the proof works exactly the same inverting the inequalities

Theorem 5. *The maximizer of*

$$m \mapsto \frac{u^p}{p} m - c(m)$$

in $[\alpha, \beta]$ lies in $\{\alpha, \beta\}$ for all non negative u if and only if c satisfies

$$\forall\, m, \qquad c(m) \geq \min\left\{c(\beta), \frac{c(\beta)-c(\alpha)}{\beta-\alpha}(m-\alpha)+c(\alpha)\right\}.$$

And the maximizer is α if

$$\frac{u^p}{p} \geq \frac{c(\beta)-c(\alpha)}{\beta-\alpha},$$

and β otherwise.

Then from both of these criteria we get that the optimum function m^* is *bang-bang* if and only if the cost satisfies the criteria almost everywhere on Ω. And then for a given $u \in W_0^{1,p}(\Omega)$ the optimal function m^* is given by

$$m^* = (\beta - \alpha)\, \mathbf{1}_E + \alpha$$

where, if we consider the minimum problem,

$$E = \left\{ x \in \Omega \,\middle|\, \frac{|u(x)|^p}{p} < \frac{c(x,\beta)-c(x,\alpha)}{\beta-\alpha} \right\},$$

and for the maximum problem

$$E = \left\{ x \in \Omega \,\middle|\, \frac{|u(x)|^p}{p} > \frac{c(x,\beta)-c(x,\alpha)}{\beta-\alpha} \right\}.$$

In both cases the optimum value is then

$$\begin{aligned} h_{|u|}(m^*) &= ((\beta-\alpha)\,\mathbf{1}_E + \alpha)\,\frac{|u|^p}{p} - c\,(x, (\beta-\alpha)\,\mathbf{1}_E + \alpha) \\ &= ((\beta-\alpha)\,\mathbf{1}_E + \alpha)\,\frac{|u|^p}{p} - (\mathbf{1}_E(c(x,\beta) - c(\alpha)) + c(x,\alpha)) \\ &= (\beta-\alpha)\,\mathbf{1}_E\,\left(\frac{|u(x)|^p}{p} - \frac{c(x,\beta)-c(x,\alpha)}{\beta-\alpha}\right) + \alpha\,\frac{|u(x)|^p}{p} - c(x,\alpha). \end{aligned}$$

What is interesting in this formulation, is that not only did we get rid of the dependence on m as we wished but the only relevant property of c is the slope of its cord from α to β, since the set E only depends on this slope, and we see here that we can consider that $c(x, \alpha) = 0$ because it will not participate in the minimisation over u. Then, we have actually shown even though we tried to generalize the problem, the only relevant cost on m is the linear cost.
Finally we have reduced the general class of problems to two problems depending on if we take the max or the min

$$\min_{u \in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + \alpha \frac{|u|^p}{p} \pm (\beta - \alpha) \left(\frac{|u|^p}{p} - \lambda \right)_\pm \right),$$

where λ is an integrable not non positive function in Ω. In this formulation the set that interests us is hidden in the positive or negative part and that's why we add this assumption on λ otherwise we would have $E = \Omega$ or $E = \emptyset$. Note that E is necessarily *p-quasi open*.

5.2 Open minimizing sets

We want to prove that under reasonable assumptions, the optimal solution u of each problem

$$\min_{u \in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + \alpha \frac{|u|^p}{p} \pm (\beta - \alpha) \left(\frac{|u|^p}{p} - \lambda \right)_\pm \right),$$

is continuous which will imply that the optimal set that we found is an open subset of Ω The first essential assumption is that λ is continuous otherwise this approach would not be relevant.
In order to show the continuity we will actually prove that the solution is $C^{0,\alpha}$. In the case where $p > d$, the result is trivial and directly follows from the Sobolev embedding theorem. If we assume $p \leq d$, we need to use a theorem proven by Giaquinta and Giusti [3] that we summarize below for the sake of completeness.

Theorem 6. *Let $\tilde{u}$ be a solution of the problem*

$$min \left\{ \int_\Omega h(x, u, \nabla u) dx \ : \ u \in W_0^{1,p}(\Omega) \right\}$$

where the integrand h satisfies the condition

$$c \left(|z|^p - b(x)|s|^\gamma - g(x) \right) \leq h(x, s, z) \leq C \left(|z|^p + b(x)|s|^\gamma + g(x) \right)$$

for all x, s, z, where $p < 1$, $0 \leq c \leq C$, $p \leq \gamma \leq p^*$, $b \in L^q_{loc}(\Omega)$, $g \in L^\sigma_{loc}(\Omega)$, with $\sigma > d/p$ and $q > p^*/(p^* - \gamma)$.
Then, $\tilde{u}$ is locally Hôlder continuous in Ω.

In our case,

$$h(x, s, z) = \frac{|z|^p}{p} - fs + \alpha \frac{|s|^p}{p} \pm (\beta - \alpha) \left(\frac{|s|^p}{p} - \lambda \right)_\pm,$$

which satisfies

$$\frac{1}{p}\left(|z|^p - \tilde{h}(x,s)\right) \le h(x,s,z) \le \frac{1}{p}\left(|z|^p \tilde{h}(x,s)\right)$$

with

$$\tilde{h}(x,s) = \beta|s|^p + \frac{d-p}{d}|s|^{p^*} + (\beta-\alpha)|\lambda| + \frac{d(p-1)+p}{d} f^{\frac{p^*}{p^*-1}}$$

and noticing that for all non negative x, $x^p < x^{p^*} + 1$, actually it implies

$$\tilde{h}(x,s) \le \left(\beta + \frac{d-p}{d}\right)|s|^{p^*} + \beta + (\beta-\alpha)|\lambda| + \frac{d(p-1)+p}{d} f^{\frac{p^*}{p^*-1}}.$$

Since we supposed λ to be continuous, we just need to have $f^{\frac{p^*}{p^*-1}} \in L^\sigma_{loc}(\Omega)$ with $\sigma > d/p$, so we need f in L^q with $q > \dfrac{d^2}{d(p-1)+p}$.

This result is summarized in the following theorem

Theorem 7. *If $p \le d$ and $f \in L^q$ with $q > \dfrac{d^2}{d(p-1)+p}$, then the solution to each of the problems*

$$\min_{u \in W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{|\nabla u|^p}{p} - fu + \alpha\frac{|u|^p}{p} \pm (\beta-\alpha)\left(\frac{|u|^p}{p} - \lambda\right)_\pm\right),$$

is locally Holder continuous

As a direct consequence the optimal set solution of the original min or max problem is open.

5.3 Dimension 1

To have an example we want to compute the optimal set in dimension 1, for simplicity, we will solve the problem with $f = 1$, $\alpha = 0$, $\beta = 1$, $p = 2$ and λ constant.

Minimization -We first solve the problem originated form the minimization problem, it reads

$$\min_{u \in H^1(\Omega)} \int_{-1}^{1} \left(\frac{u'^2}{2} - u + \mathbf{1}_{E(u)}\left(\frac{u^2}{2} - \lambda\right)\right),$$

with $E(u) = \{u < \sqrt{2\lambda}\}$ and λ a non negative real number. The solution u^* to the problem must satisfy the Euler Lagrange equation

$$\begin{cases} -u'' + u = 1 & \text{if } u < \sqrt{2\lambda} \\ -u'' = 1 & \text{else} \\ u(-1) = u(1) = 0. \end{cases} \tag{3}$$

This equation imposes that if $\bar{u}$ is a solution in H^1 then $\bar{u}''$ is in L^2 which imply that we look for a solution in $C^1(]-1,1[)$. We know that $E(u)$ is an open subset

of $[-1,1]$, so it is a union of open intervals. Also since -1 and 1 are zeros of u, either E includes at least two intervals, one of the form $[-1,a[$ and the other of the form $]b,1]$ with $a<b$, such that $u(a)=u(b)=\sqrt{2\lambda}$ and for all $x\in[-1,a[\cup]b,1]$, $u(x)<\sqrt{2\lambda}$ or if there is no such a and b then $E=[-1,1]$.
We will try to construct solutions of Euler Lagrange where $E\neq[-1,1]$. Take $a<b$, we want to find a solution $w_{a,b}$ such that $[-1,a[$ and $]b,1]$ are included in E and $w_{a,b}(a)=w_{a,b}(b)=\sqrt{2\lambda}$, by symmetry of the problem we can suppose that $a\leq 0$. On $[-1,a[$, since it is a solution to (3), $w_{a,b}$ must satisfy

$$w_{a,b}(x)=(\sqrt{2\lambda}-1)\frac{sh(1+x)}{sh(1+a)}-\frac{sh(a-x)}{sh(1+a)}+1,$$

and on $]b,1]$, it must satisfy

$$w_{a,b}(x)=(\sqrt{2\lambda}-1)\frac{sh(1-x)}{sh(1-b)}-\frac{sh(x-b)}{sh(1-b)}+1.$$

The condition $[-1,a[\cup]b,1]\subset E\left(w_{a,b}\right)$ is equivalent to $w'_{a,b}(a^-)<0$ and $w'_{a,b}(b^+)<0$. If $\sqrt{2\lambda}\geq 1$ this condition is trivial and if $\sqrt{2\lambda}<1$ this condition is equivalent to

$$ch(1+a)<\frac{1}{1-\sqrt{2\lambda}}\quad\text{and}\quad ch(1-b)<\frac{1}{1-\sqrt{2\lambda}},$$

which rewrites for $A(\lambda)=argch\left(\frac{1}{1-\sqrt{2\lambda}}\right)$,

$$a<1-A(\lambda)\quad\text{and}\quad b>1-A(\lambda).$$

Between a and b, either $w_{a,b}$ is always larger than $\sqrt{2\lambda}$ and it satisfies on $[a,b]$

$$w_{a,b}(x)=-\frac{x^2}{2}+\frac{a+b}{2}x-\frac{ab}{2}+\sqrt{2\lambda},$$

or there is at least an other interval $]c,d[$ such that $w_{a,b}(c)=w_{a,b}(d)=\sqrt{2\lambda}$ and $w_{a,b}(x)<\sqrt{2\lambda}$ in between. If there is such an interval, necessarily, on $]c,d[$, $w_{a,b}$ must satisfy

$$w_{a,b}(x)=(\sqrt{2\lambda}-1)\frac{ch\left(x-\frac{c+d}{2}\right)}{ch\left(\frac{d-c}{2}\right)}+1.$$

Here we see that if $\sqrt{2\lambda}<1$, such an interval cannot exist and $E\left(w_{a,b}\right)=[-1,a[\cup]b,1]$. We will tackle this case first,

If $\sqrt{2\lambda}<1$, necessarily, $w_{a,b}$ is of the form

$$w_{a,b}(x)=\begin{cases}(\sqrt{2\lambda}-1)\dfrac{sh(1+x)}{sh(1+a)}-\dfrac{sh(a-x)}{sh(1+a)}+1 & \text{if } x\in[-1,a[\\ -\dfrac{x^2}{2}+\dfrac{a+b}{2}x-\dfrac{ab}{2}+\sqrt{2\lambda} & \text{if } x\in[a,b]\\ (\sqrt{2\lambda}-1)\dfrac{sh(1-x)}{sh(1-b)}-\dfrac{sh(x-b)}{sh(1-b)}+1 & \text{if } x\in]b.1]\end{cases}$$

The C^1 conditions is equivalent to

$$\frac{(\sqrt{2\lambda}-1)ch(1+a)+1}{sh(1+a)} = \frac{(\sqrt{2\lambda}-1)ch(1-b)+1}{sh(1-b)} = \frac{b-a}{2}.$$

Defining h the function on $[1-A(\lambda),1]$ such that

$$h(b) = \frac{(\sqrt{2\lambda}-1)ch(1-b)+1}{sh(1-b)},$$

this function is strictly increasing, so for a fixed a the only b satisfying the first part of the C^1 condition is $b=-a$. On can check that actually, the function $\tilde{h} : b \mapsto h(b)-b$ is also increasing and goes to ∞ in 1, and $\tilde{h}(1-A(\lambda)) = A(\lambda)-1$. In conclusion, if $1 > \sqrt{2\lambda} \geq 1 - \dfrac{1}{ch(1)}$, there is no such solution and the only solution is

$$u^* = 1 - \frac{ch(x)}{ch(1)},$$

and the optimal m^* to the original minimization problem is $m^* = 1$.
If $\sqrt{2\lambda} < 1 - \dfrac{1}{ch(1)}$, the only solution is $w_{-\beta.,\beta}$, for β solving

$$\frac{(\sqrt{2\lambda}-1)ch(1-\beta)+1}{sh(1-\beta)} = \beta,$$

And the optimal set E^* is then $[-1,-\beta[\cup]\beta,1]$.

If $\sqrt{2\lambda} \geq 1$, $w_{a,b}$ could behave in different ways between a ad b but we know that in any case, $w'_{a,b}(a^+) \leq \dfrac{b-a}{2} \leq \dfrac{1-a}{2}$. Then the C^1 condition imposes

$$\frac{1}{sh(1+a)}\Big((\sqrt{2\lambda}-1)ch(1+a)+1\Big) = w'_{a,b}(a^+) \leq \frac{1-a}{2}.$$

Consider the function g in $-1 \leq a \leq 0$ defined by

$$g(a) = \frac{1}{sh(1+a)}\Big((\sqrt{2\lambda}-1)ch(1+a)+1\Big) + \frac{a}{2}$$

it is decreasing and

$$g(0) = \frac{1}{sh(1)}\Big((\sqrt{2\lambda}-1)ch(1)+1\Big) \geq \frac{1}{sh(1)} > \frac{1}{2}.$$

Then this condition cannot be satisfied so there is no admissible $w_{a,b}$ if $\sqrt{2\lambda} \geq 1$ and the optimal m^* to the original minimization problem is $m^* = 1$.

The results are summarized in the following theorem

Theorem 8. *If* $\sqrt{2\lambda} < 1 - \frac{1}{ch(1)}$*, the optimal set is* $E^* = [-1, -\beta[\cup]\beta, 1]$*, for* $0 < \beta < 1$ *solving*

$$(\sqrt{2\lambda} - 1)ch(1-\beta) + 1 = \beta\, sh(1-\beta)$$

If $\sqrt{2\lambda} \geq 1 - \frac{1}{ch(1)}$*, the optimal set is* $E^* = [-1.1]$.

Maximization - We can now take a look at the related problem originated from the maximization problem.

$$\min_{u \in H^1(\Omega)} \int_{-1}^{1} \left(\frac{u'^2}{2} - u + \mathbf{1}_{E(u)} \left(\frac{u^2}{2} - \lambda \right) \right),$$

with $E(u) = \{u > \sqrt{2\lambda}\}$ and λ a non negative real number. Once again we can work with the Euler Lagrange equation

$$\begin{cases} -u'' + u = 1 & \text{if } u > \sqrt{2\lambda} \\ -u'' = 1 & \text{else} \\ u(-1) = u(1) = 0. \end{cases} \tag{4}$$

This time the possible solutions are a bit simpler, either the solution is

$$v^* = \frac{1 - x^2}{2},$$

Or, if $\sqrt{2\lambda} < 1$, it could be of the form

$$v_{a,b} = \begin{cases} \dfrac{(x+1)(a-x)}{2} + \dfrac{\sqrt{2\lambda}}{1+a}(x+1) & \text{if } x \in [-1, a] \\ 1 + (\sqrt{2\lambda} - 1)\dfrac{ch\left(x - \frac{a+b}{2}\right)}{ch\left(\frac{b-a}{2}\right)} & \text{if }]a, b[\\ \dfrac{(1-x)(x-b)}{2} + \dfrac{\sqrt{2\lambda}}{1-b}(1-x) & \text{if } x \in [b, 1] \end{cases}$$

The C^1 condition is equivalent to $a = -b$ and b solves

$$(\sqrt{2\lambda} - 1)\frac{sh(b)}{ch(b)} = -\frac{\sqrt{2\lambda}}{1-b} + \frac{1-b}{2}$$

and the validity condition being

$$\sqrt{2\lambda} > \frac{(1-b)^2}{2},$$

to have a solution, we need $\sqrt{2\lambda} < \frac{1}{2}$. So, if $\sqrt{2\lambda} \geq \frac{1}{2}$, the optimal set E^* is $[-1, 1]$ and, if $\sqrt{2\lambda} < \frac{1}{2}$, the optimal set is $[-b^*.b^*]$, with b^* the only solution to the C^1 condition.

References

[1] F. Paolo Maiale G. Buttazzo and B. Velichkov. “Shape optimization problems in control form”. In: (2021).

[2] F. H. Clarke. “Multiple integrals of lipchitz functions in the calculus of variation”. In: (1977).

[3] M. Giaquinta and E. Giusti. “Quasi-minima”. In: (1984).