Westfälische
Wilhelms-Universität
Münster

Vector-Valued Multi-Bang Control for Linearised Elasticity

Master Thesis
submitted in fulfilment of the requirements for the degree of
Master of Science

Institute for Computational and Applied Mathematics
University of Münster, Germany

Supervisor:
Prof. Dr. Benedikt Wirth

Submitted by:
Carla Maria Tameling

Münster, September 2015

Abstract

In optimal control theory there are often problems where the control takes on only values from a discrete set of given points. These problems are called multi-bang control problems. They are challenging since the penalty functional that achieves that the control takes on only values from the discrete set is neither convex nor lower semi-continuous. Replacing this penalty functional by its convex regularization yields that we can derive a primal-dual optimality system. It can be shown that the primal-dual optimality system has a unique solution and under certain condition coincides with the solution of the original problem. The Moreau-Yosida approximation can be applied to the optimality system. The regularized system is amenable to a solution with a semismooth Newton method.

Eidesstattliche Erklärung

Hiermit versichere ich, *Carla Tameling*, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Gedanklich, inhaltlich oder wörtlich Übernommenes habe ich durch Angabe von Herkunft und Text oder Anmerkung belegt bzw. kenntlich gemacht. Dies gilt in gleicher Weise für Bilder, Tabellen, Zeichnungen und Skizzen, die nicht von mir selbst erstellt wurden.

Alle auf der CD beigefügten Programme sind von mir selbst programmiert worden.

Münster, 04.09.2015

Carla Tameling

Ich erkläre mich mit einem Abgleich der Arbeit mit anderen Texten zwecks Auffindung von Übereinstimmungen sowie mit einer zu diesem Zweck vorzunehmenden Speicherung der Arbeit in eine Datenbank einverstanden.

Münster, 04.09.2015

Carla Tameling

Contents

1. Introduction

If one works in the field of optimal control theory, one can find problems where the control variable takes on only two different values. One of the standard introduction problems is the drag racing pilot who needs to cover a given distance in minimal time. To achieve this goal he needs to use maximum acceleration and after some meters maximum breaking, in order to be able to stop at the finishing line.

This kind of optimal control problem is called bang-bang control problem. Christian Clason and Karl Kunisch introduced a multi-bang control in their paper 'Multi-Bang Control of Elliptic Systems' [7]. According to this paper a multi-bang control is a control that takes on only values from a discrete set of given points. The two authors considered in their paper the following problem:

$$\begin{cases} \min\limits_{u,y} \frac{1}{2} \|y - z\| + \frac{\alpha}{2} \|u\|_{L^2}^2 + \beta \int\limits_{\Omega} \prod\limits_{i=1}^{d} |u(x) - u_i|_0 \,\mathrm{d}x, \\ \text{s. t. } Ay = u,\ u_1 \leq u(x) \leq u_d \text{ for almost every } x \in \Omega, \end{cases} \tag{1.1}$$

for given $\alpha, \beta > 0$, real numbers $u_1 < ... < u_d, d \geq 2$ and a target $z \in L^2(\Omega)$. They assumed that V is a Hilbert space and that $A\colon V \to V^*$ is an isomorphism. Moreover, the embeddings $V \hookrightarrow L^2(\Omega) \hookrightarrow V^*$ should be continuous, compact and dense. Furthermore, $|\cdot|_0$ is defined as follows:

$$|t|_0 := \begin{cases} 0 & \text{if } t = 0, \\ 1 & \text{if } t \neq 0. \end{cases}$$

They have proven that they can find a unique solution if they take the convex relaxation of $\frac{\alpha}{2} \|u\|_{L^2}^2 + \beta \int\limits_{\Omega} \prod\limits_{i=1}^{d} |u(x) - u_i|_0 \,\mathrm{d}x$ and that this solution is continuously dependent of the target z. Additionally, they were able to show that one can find conditions under which the solution of the problem with the relaxation is the solution of (1.1).

This work aims to prove that problem (1.1) still has a unique solution in two dimensions, i.e. we consider vector-valued functions y and u instead of scalar functions. It is too difficult to prove this in generality in two dimensions since $\mathbb{R}^2$ is not ordered. Therefore, we want to prove it for a given set of eight control states and for $\beta = \infty$. Moreover, we take the indicator function instead of $|\cdot|_0$. During the thesis we find some points at which we can give more general statements. Furthermore, we look only on a special constraint. We have decided to take the linearised elasticity equation as the constraint because in this equation the two components of the vector field are linked.

Consequently, we are not able to use the theorems from Clason and Kunisch in each component. We have not found a real application of this model yet. We could make up some artificial problem with an area that is full of Piezo elements. These elements convert a force into electric voltage or vice versa. This in not a problem someone has to deal with in reality.

This thesis is organized in the following way. The next chapter deals with the linearised elasticity equation. We start with the problem that is to be modelled and derive from this the linearised elasticity equation. Afterwards, we prove in 2.1 that the linearised elasticity equation has a unique solution. In chapter 3 we give a short introduction to optimal control theory. Chapter 4 contains all definitions and theorems from convex analysis that we will need during the rest of the thesis. A collection of parts from monotone operator theory that is going to be used can be found in chapter 5.
The mainpart of this thesis starts in chapter 6. We are going to introduce the considered multi-bang control problem. In 6.1 we deal with the optimality system for our problem. Existence of a unique solution and continuous dependence on the target are proven in 6.2. The last section of this chapter deals with the structure of the solution.
Chapter 7 addresses the numerical calculation of a solution. Therefore, we need to introduce the Moreau-Yosida regularization of our system, since we can apply a semismooth Newton method only on the regularized system. In the next section we define a semismooth Newton method and prove that it converges locally superlinear. Last but not least we want to give some numerical examples.
In the last chapter we want to summarize all findings and give some ideas for further research topics.

2. Linearised Elasticity Equation - an Example for a Vector-Valued PDE

This chapter is based on Braess' study [2]. We consider a very flat solid body $\Omega \subset \mathbb{R}^2$. One side of the body is attached to a wall. The question to be model is how the body is deformed when we apply a force to it (cf. figure 2.1).

We assume that the body is made out of an elastic material that is isotropic, i.e. the elasticity is the same in every direction, and homogeneous, this means that the elasticity is the same everywhere in the material. We also assume that we only have small strains, e.g. a very stiff material. This is the requirement for Hooke's linear law. The resulting PDE for this problem is the Lamé equation

$$\begin{aligned} -2\mu \operatorname{div} \epsilon(u) - \lambda \operatorname{grad} \operatorname{div} u = f \text{ in } \Omega, \\ u = 0 \text{ on } \Gamma_0, \\ \sigma(u) \cdot \nu = g \text{ on } \Gamma_1, \end{aligned} \tag{2.1}$$

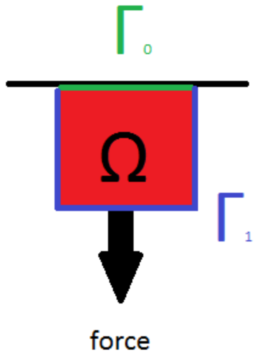


Figure 2.1.: Model of the deformation of a flat body. The black arrow indicates the direction of the force.

where ϵ is the symmetrized gradient that is a linear approximation of the strain. σ represents the stress tensor. ϵ and σ are tensor fields and u is a vector field. This is the translation. The fields are linked through the kinematic principle

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{2.2}$$

and Hooke's law

$$\sigma = \lambda tr(\epsilon) \operatorname{Id} + 2\mu\epsilon. \tag{2.3}$$

f and g are also vector fields. f is the uniformly distributed volume load and g the uniformly distributed linear load. By multiplication with a test function v and integration over Ω we get

$$\begin{aligned} &-2\mu \int_\Omega \operatorname{div} \epsilon(u) \cdot v \, \mathrm{d}x - \lambda \int_\Omega \operatorname{grad} \operatorname{div} u \cdot v \, \mathrm{d}x = \int_\Omega f \cdot v \, \mathrm{d}x \\ \overset{Green}{\Leftrightarrow}\; &2\mu \int_\Omega \epsilon(u) : \nabla v \, \mathrm{d}x + \lambda \int_\Omega \operatorname{div} u \operatorname{div} v \, \mathrm{d}x - 2\mu \int_{\partial\Omega} \epsilon(u)\nu \cdot v \, \mathrm{d}\sigma \\ &-\lambda \int_{\partial\Omega} \operatorname{div}(u) \operatorname{Id} \nu \cdot v \, \mathrm{d}\sigma = \int_\Omega f \cdot v \, \mathrm{d}x. \end{aligned} \tag{2.4}$$

Here $\epsilon(u) : \epsilon(v) := tr(\epsilon(u)^T \epsilon(v))$. Since $A : B = A^T : B = A : B^T$ for $A, B \in \mathbb{R}^{n\times n}$, A symmetric, it holds $\epsilon(u) : \nabla v = \epsilon(u) : \frac{\nabla v + (\nabla v)^T}{2} = \epsilon(u) : \epsilon(v)$. Therefore, (2.4) is equivalent to

$$\begin{aligned}2\mu \int_\Omega \epsilon(u) : \epsilon(v)\,\mathrm{d}x + \lambda \int_\Omega \operatorname{div} u \operatorname{div} v\,\mathrm{d}x - 2\mu \int_{\partial\Omega} \epsilon(u)\nu \cdot v\,\mathrm{d}\sigma \\ -\lambda \int_{\partial\Omega} \operatorname{div}(u)\,\mathrm{Id}\,\nu \cdot v\,\mathrm{d}\sigma = \int_\Omega f \cdot v\,\mathrm{d}x.\end{aligned} \tag{2.5}$$

We have to define the space for our test functions:

$$H^1_\Gamma := \left\{ v \in H^1(\Omega)^2 \colon v(x) = 0 \text{ for } x \in \Gamma_0 \right\}.$$

Furthermore, we can use the definition of σ (2.3) and receive the variational formulation:

Variational formulation of the Lamé equation 2.1. *Find $u \in H^1_\Gamma$ such that*

$$2\mu(\epsilon(u), \epsilon(v))_2 + \lambda(\operatorname{div} u, \operatorname{div} v)_2 = (f, v)_2 + \int_{\Gamma_1} g \cdot v\,\mathrm{d}x \tag{2.6}$$

holds for all $v \in H^1_\Gamma$ and $(\cdot,\cdot)_2$ denotes the corresponding L^2-inner product.

2.1. Existence and Uniqueness of a Solution of the Linearised Elasticity Equation

In the following we are going to prove that (2.6) has a unique solution for all appropriate f and g. Since the left-hand side of (2.6) is a bilinear form and the right-hand side a linear functional in the dual space $(H^1_\Gamma)^*$, we are going to use the lemma of Lax-Milgram for this proof.

Theorem 2.2 (Lemma of Lax-Milgram)**.** *Let V be a real Hilbert space, $a \colon V \times V \to \mathbb{R}$ a continuous and coercive bilinear form and $l \in V^*$. Then there exists a unique solution $u \in V$ of*

$$a(u, v) = l(v) \qquad \forall v \in V.$$

The continuity of $a(u, v) = 2\mu(\epsilon(u), \epsilon(v))_2 + \lambda(\operatorname{div} u, \operatorname{div} v)_2$ follows directly with the triangle inequality and the Cauchy-Schwarz inequality. We need a few more efforts to verify that our bilinear form is coercive.

Theorem 2.3 (1. Korn's inequality)**.** *Let Ω be an open and bounded subset of $\mathbb{R}^n$ with a piecewise smooth boundary. Then it exists a constant $c = c(\Omega) > 0$ such that*

$$\int_\Omega \epsilon(v) : \epsilon(v)\,\mathrm{d}x + \|v\|^2_{L^2} \geq c\,\|v\|^2_{H^1} \;\; \forall v \in H^1(\Omega)^n.$$

For $v \in H^1_\Gamma$ we can simplify this inequality and get Korn's second inequality. We prove only the simplified version, because we need this inequality to confirm that our bilinear form is coercive.

Theorem 2.4 (2. Korn's inequality)**.** *Let $\Omega \subset \mathbb{R}^2$ be an open and bounded subset with a piecewise smooth boundary. Assume that $\Gamma_0 \subset \partial\Omega$ has a positive 1-dimensional measure. Then there exists a constant $\tilde{c} > 0$ that depends on Ω and Γ_0 such that*

$$\int_\Omega \epsilon(v) : \epsilon(v)\,\mathrm{d}x \geq \tilde{c}\,\|v\|_{H^1}^2 \quad \forall v \in H^1_\Gamma(\Omega).$$

Proof. Assume that the inequality is not true. In this case we can find a sequence $(v_n) \in H^1_\Gamma(\Omega)$ such that

$$\|\epsilon(v_n)\|_{L^2}^2 := \int_\Omega \epsilon(v) : \epsilon(v)\,\mathrm{d}x \leq \frac{1}{n} \text{ and } |v_n|_{H^1} = 1.$$

The assumption on Γ_0 and Friedrich's inequality [see Braess [2] Chap. 2.1 for the inequality and the proof] yield that $\|v_n\|_{H^1} \leq c_1$ for all n and a reasonable constant $c_1 > 0$. Due to the fact that $H^1(\Omega)$ is a compact subspace of $L^2(\Omega)$ there exists a convergent subsequence regarding the $\|.\|_{L^2}$-norm of (v_n). With theorem 2.3 we get for this subsequence $c\,\|v_{n_k} - v_{n_l}\|_{H^1}^2 \leq \|\epsilon(v_{n_k} - v_{n_l})\|_{L^2}^2 + \|v_{n_k} - v_{n_l}\|_{L^2}^2 \leq 2\,\|\epsilon(v_{n_k})\|_{L^2}^2 + 2\,\|\epsilon(v_{n_l})\|_{L^2}^2 + \|v_{n_k} - v_{n_l}\|_{L^2}^2 \leq \frac{2}{n_k} + \frac{2}{n_l} + \|v_{n_k} - v_{n_l}\|_{L^2}^2$. Therefore, our subsequence is a Cauchy-sequence in $H^1(\Omega)$ and converges with regard to the $\|.\|_{H^1}$-norm to a u_0 since $H^1(\Omega)$ is a Banach space. It holds for u_0 that $\|\epsilon(u_0)\|_{L^2} = \lim_{k\to\infty} \|\epsilon(v_{n_k})\|_{L^2} = 0$ and $|u_0|_{H^1} = \lim_{k\to\infty} |v_{n_k}|_{H^1} = 1$. It follows from $\|\epsilon(u_0)\|_{L^2} = 0$ that $u_0(x) = Ax + b$. In this case $b \in \mathbb{R}^2$ and $A \in \mathbb{R}^{2\times 2}$ is a skew-symmetric matrix. Since we have zero boundary conditions on Γ_0 we get $u_0 = 0$. This is a contradiction to $|u_0|_{H^1} = 1$. $\square$

With Korn's second inequality we get an upper bound for the first term of our bilinear form and since the second term is positive we get

$$a(u,u) = 2\mu(\epsilon(u), \epsilon(u))_2 + \lambda(\operatorname{div} u, \operatorname{div} u)_2 \geq 2\mu\tilde{c}\,\|u\|_{H^1}^2.$$

This last estimation confirms the coercivity of the bilinear form and Lax-Milgram 2.2 guarantees that there is a unique solution for the variational formulation of the Lamé equation for all given controls f and g.

3. Optimal Control Theory

In this chapter we want to give a short introduction to optimal control theory. It follows with its ideas and definitions chapter 1 from Hinze, Pinnau, Ulbrich and Ulbrich [11].
We often have the problem that we want to minimize a functional with a certain constraint. These constrains are very often partial differential equations (PDEs). An example from physics is the stationary heating of a solid object $\Omega \subset \mathbb{R}^3$. We apply a temperature distribution $u(x)$ - the control - to the boundary $\partial\Omega$ of the object. Our aim is to find a temperature distribution $y\colon \Omega \to \mathbb{R}$ - the state - of the object that is close to a given temperature distribution $y_0\colon \Omega \to \mathbb{R}$. We need to consider the fact that we cannot heat the whole object immediately. The temperature distribution $y(x)$ inside the object can be modelled with the Laplace equation

$$-\Delta y(x) = 0,\ x \in \Omega \tag{3.1}$$

and Robin-Boundary-Conditions

$$\kappa \frac{\partial y}{\partial \nu} = \beta(x)(u(x) - y(x)),\ x \in \partial\Omega. \tag{3.2}$$

Here $\kappa > 0$ is the heat conduction coefficient of the object's material and $\beta\colon \partial\Omega \to (0, \infty)$ is a positive function that models the coefficient for the heat exchange with the exterior. Furthermore, the control is pointwise bounded $a(x) \leq u(x) \leq b(x)$, for all $x \in \partial\Omega$ according to a bounded heating capacity. Since we want $y(x)$ to be close to $y_0(x)$ and we have to take care of the temperature distribution at the boundary, we have to solve the following problem:

$$\begin{aligned} \min_{y,u} J(t,u) &:= \frac{1}{2} \left\| y(x) - y_0(x) \right\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \left\| u(x) \right\|^2_{L^2(\partial\Omega)}, \\ \text{subject to } & -\Delta y = 0 \text{ on } \Omega, \\ & \frac{\partial y}{\partial \nu} = \frac{\beta}{\kappa}(u - y) \text{ on } \partial\Omega, \\ & a \leq u \leq b \text{ on } \partial\Omega. \end{aligned} \tag{3.3}$$

We can reformulate the constraints. The second one is equivalent to

$$\frac{\partial y}{\partial \nu} - \frac{\beta}{\kappa}(u - y) = 0 \text{ on } \partial\Omega$$

and the third constrained can be split into the following two constraints:

$$u - a \geq 0 \text{ and } b - u \geq 0 \text{ on } \partial\Omega.$$

We can now introduce two functionals e and c, where e summarizes all equality constraints and c collects all inequality constraints:

$$e(y,u) = \begin{pmatrix} -\Delta y \\ \dfrac{\partial y}{\partial \nu} - \dfrac{\beta}{\kappa}(u-y) \end{pmatrix}, \; c(y,u) = \begin{pmatrix} u-a \\ b-u \end{pmatrix}.$$

With this functionals we can rewrite our minimization problem (3.3):

$$\min_{y\in Y, u\in U} J(y,u) \qquad \text{s.t. } e(y,u)=0, \; c(y,u)\in \mathcal{N}, \tag{3.4}$$

where Y and U are appropriate Banach spaces that contain functions $y\colon \Omega \to \mathbb{R}$ and $u\colon \Omega \to \mathbb{R}$ respectively and $\mathcal{N} = \{(v_1,v_2)\in U\times U, v_i \le 0, i=1,2\}$. On this level of abstraction we can give the definition of an optimal control problem.

Definition 3.1. *An optimal control problem is an optimization problem of the following form:*

$$\begin{aligned} &\min_{y\in Y, u\in U} J(y,u), \\ \textit{subject to } &e(y,u)=0, \qquad\qquad \textit{(state equation)} \\ &c(y,u)\in \mathcal{K}, \end{aligned} \tag{3.5}$$

where $J\colon Y\times U\to\mathbb{R}$ is called objective function, $e\colon Y\times U\to Z$ is the equality constraint operator, $c\colon Y\times U\to X$ is the inequality constraint operator and $\mathcal{K}\subset X$ is a closed convex cone. U,Y,X and Z are real Banach spaces.
y is called state variable and u is the control variable. A central aspect of optimal control theory is that the state equation is assumed to have a (unique) solution y for all given controls u.

In chapter 6 we are going to introduce the optimal control problem that is going to be considered. Our problem has an elliptic partial differential operator as the state equation. We already know that our state equation, the Lamé equation, has a unique solution for every given right-hand side [see chapter 2].

4. Convex Analysis

In this chapter we want to list several definitions and theorems from convex analysis. These are taken from Bauschke and Combettes [1], Schirotzek [13] and Ekeland and Temam [8].
We start with the definition of the Fenchel conjugate and some additional properties of this.

Definition 4.1 (Fenchel conjugate)**.** *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{F}\colon \mathcal{H} \to \bar{\mathbb{R}}$.*

$$\mathcal{F}^*(y) = \sup_x \langle x, y\rangle - \mathcal{F}(x), \tag{4.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product, is called Fenchel conjugate of $\mathcal{F}$ at y.

Theorem 4.2. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{F}\colon \mathcal{H} \to [-\infty, +\infty]$. Then the Fenchel conjugate $\mathcal{F}^*$ is convex and lower semi-continuous.*

Proof. Assume that $\mathcal{F} \not\equiv +\infty$, otherwise we are finished since the constant function is lower semi-continuous and convex. It follows directly from the definition of $\mathcal{F}^*$ that it yields

$$\mathcal{F}^*(u) = \sup_{x \in \operatorname{dom} \mathcal{F}} \langle x, u\rangle - \mathcal{F}(x) = \sup_{(x,\xi) \in \operatorname{epi} \mathcal{F}} \langle x, u\rangle - \xi \tag{4.2}$$

for $u \in \mathcal{H}$ and $\operatorname{epi} \mathcal{F} = \{(u,t) \in \mathcal{H} \times \mathbb{R} \,|\, \mathcal{F}(u) \leq t\}$ is the epigraph of $\mathcal{F}$. We can show that $\mathcal{F}_{(x,\xi)}\colon \mathcal{H} \to \bar{\mathbb{R}},\ u \mapsto (\langle x, u\rangle - \xi)_{(x,\xi)\in \operatorname{epi} \mathcal{F}}$ is lower semi-continuous and convex:
We start with the proof for the convexity of $\mathcal{F}_{(x,\xi)}$. Let (u_1, t_1) and (u_2, t_2) be elements of $\operatorname{epi} \mathcal{F}_{(x,\xi)}$ and take $\lambda \in [0,1]$, then

$$\begin{aligned}
\mathcal{F}_{(x,\xi)}(\lambda u_1 + (1-\lambda)u_2) &= \langle x, \lambda u_1 + (1-\lambda)u_2\rangle - \xi \\
&= \lambda(\langle x, u_1\rangle - \xi) + (1-\lambda)(\langle x, u_2\rangle - \xi) \\
&= \lambda\, \mathcal{F}_{(x,\xi)}(u_1) + (1-\lambda)\, \mathcal{F}_{(x,\xi)}(u_2) \\
&\leq \lambda t_1 + (1-\lambda)t_2.
\end{aligned}$$

Hence, $\lambda(u_1, t_1) + (1-\lambda)(u_1, t_1) \in \operatorname{epi} \mathcal{F}$ and therefore $\mathcal{F}_{(x,\xi)}$ is convex.
A function is lower semi-continuous if and only if its epigraph is closed. For a proof see Lemma 1.24 in [1]. Therefore, we have to show that $\operatorname{epi} \mathcal{F}_{(x,\xi)}$ is closed.
Let (u_a, t_a) be a sequence in $\operatorname{epi} \mathcal{F}_{(x,\xi)}$ that converges to $(u,t) \in \mathcal{H} \times \mathbb{R}$, then

$$\mathcal{F}_{(x,\xi)}(u) = \langle x, u\rangle - \xi \overset{(*)}{=} \lim \langle x, u_a\rangle - \xi \leq \lim t_a = t.$$

$(*)$ holds since the inner product is continuous in both arguments.
This inequality shows that $(u,t) \in \operatorname{epi} \mathcal{F}_{(x,\xi)}$ and therefore $\mathcal{F}_{(x,\xi)}$ is lower semi-continuous.

The supremum of the family $(\mathcal{F}_{(x,\xi)})_{(x,\xi)}$ of lower semi-continuous and convex functions is lower semi-continuous and convex, since it holds

$$\operatorname{epi}(\sup_{(x,\xi)\in \operatorname{epi}\mathcal{F}} \mathcal{F}_{(x,\xi)}) = \bigcap_{(x,\xi)\in \operatorname{epi}\mathcal{F}} \operatorname{epi} \mathcal{F}_{(x,\xi)} \,. \tag{4.3}$$

Because of the facts that the epigraph of each $\mathcal{F}_{(x,\xi)}$ is convex as we have shown above and that the intersection of convex sets is convex, the epigraph of $\sup_{(x,\xi)\in \operatorname{epi}\mathcal{F}} \mathcal{F}_{(x,\xi)}$ is convex and therefore the supremum is convex, too. It remains to prove that (4.3) holds true. Let $(u,t) \in \mathcal{H} \times \mathbb{R}$, then

$$\begin{aligned}
&(u,t) \in \operatorname{epi}(\sup_{(x,\xi)\in \operatorname{epi}\mathcal{F}} \mathcal{F}_{(x,\xi)}) \\
\Leftrightarrow &\sup_{(x,\xi)\in \operatorname{epi}\mathcal{F}} \mathcal{F}_{(x,\xi)}(u) \leq t \\
\Leftrightarrow &\mathcal{F}_{(x,\xi)}(u) \leq t \ \forall (x,\xi) \in \operatorname{epi}\mathcal{F} \\
\Leftrightarrow &(u,t) \in \operatorname{epi}\mathcal{F}_{(x,\xi)} \ \ \forall (x,\xi) \in \operatorname{epi}\mathcal{F} \\
\Leftrightarrow &(u,t) \in \bigcap_{(x,\xi)\in \operatorname{epi}\mathcal{F}} \operatorname{epi} \mathcal{F}_{(x,\xi)} \,.
\end{aligned}$$

The next step is to show the lower semi-continuity of the supremum.
Combining the fact that lower semi-continuity is equivalent to a closed epigraph with (4.3) we get that the supremum of a family of lower semi-continuous functions is lower semi-continuous. Thus, the assumption is proven.

□

Now we consider the definition of the subdifferential and the interplay between the Fenchel conjugate and the subdifferential.

Definition 4.3 (Subdifferential)**.** *Let $\mathcal{H}$ be a real Hilbert space and $\mathcal{F}\colon \mathcal{H} \to (-\infty, +\infty]$ be a proper functional, i.e. $-\infty \notin \mathcal{F}(\mathcal{H})$ and $\operatorname{dom}\mathcal{F} \neq \emptyset$. The set-valued operator*

$$\partial \mathcal{F}\colon \mathcal{H} \to 2^{\mathcal{H}}, \ x \mapsto \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H}) \ \mathcal{F}(y) \geq \mathcal{F}(x) + \langle y - x, u\rangle\}\,, \tag{4.4}$$

is called subdifferential of $\mathcal{F}$. $\mathcal{F}$ is subdifferentiable at x if $\partial \mathcal{F}(x) \neq \emptyset$ and the subgradients of $\mathcal{F}$ at x are the elements of $\partial \mathcal{F}(x)$.

Theorem 4.4. *Let $\mathcal{H}$ be a real Hilbert space, $\mathcal{F}\colon \mathcal{H} \to \overline{\mathbb{R}}$ be proper, convex and lower semi-continuous, then it holds*

$$x \in \operatorname{dom}\mathcal{F}, y \in \partial \mathcal{F}(x) \quad \Leftrightarrow \quad y \in \operatorname{dom}\mathcal{F}^*, x \in \partial \mathcal{F}^*(y). \tag{4.5}$$

Proof. By the definition of the subgradient we get

$$\begin{aligned}
&y \in \partial \mathcal{F}(x) \\
&\Leftrightarrow \mathcal{F}(v) \geq \mathcal{F}(x) + \langle v - x, y \rangle \ \ \forall v \in \mathcal{H} \\
&\Leftrightarrow \langle v, y \rangle - \mathcal{F}(v) \leq \langle x, y \rangle - \mathcal{F}(x) \ \forall v \in \mathcal{H} \\
&\Leftrightarrow \mathcal{F}^*(y) + \mathcal{F}(x) \leq \langle x, y \rangle \,.
\end{aligned}$$

The Young-inequality reads $\langle x, y\rangle \leq \mathcal{F}(x) + \mathcal{F}^*(y) \ \forall x \in \operatorname{dom} \mathcal{F}, y \in \mathcal{H}$. We deduce

$$\langle x, y \rangle = \mathcal{F}(x) + \mathcal{F}^*(y) \Leftrightarrow y \in \partial \mathcal{F}(x). \tag{4.6}$$

From this it follows that $y \in \operatorname{dom} \mathcal{F}^*$. Furthermore, for each $v \in \mathcal{H}$ we get using the Young-inequality and (4.6) $\langle v - y, x\rangle \leq \mathcal{F}(x) + \mathcal{F}^*(\bar{v}) - (\mathcal{F}(x) + \mathcal{F}^*(y) = \mathcal{F}^*(v) - \mathcal{F}^*(y)$. Consequently, we have $x \in \partial \mathcal{F}^*(y)$.
It remains to show that $y \in \operatorname{dom} \mathcal{F}^*, x \in \partial \mathcal{F}^*(y) \Rightarrow x \in \operatorname{dom} \mathcal{F}, y \in \partial \mathcal{F}(x)$. Since $\mathcal{F}^*$ is proper, convex and lower semi-continuous as we proved in 4.2, we get from (4.6) applied on $\mathcal{F}^*$ instead of $\mathcal{F}$, $\langle y, x\rangle = \mathcal{F}^*(y) + \mathcal{F}^{**}(x)$. As a result, we have $x \in \operatorname{dom} \mathcal{F}^{**}$. Because of the fact that $\mathcal{F}$ is proper, convex and lower semi-continuous, it holds $\mathcal{F}^{**} = \mathcal{F}$. We obtain $y \in \partial \mathcal{F}(x)$ by applying (4.6) again. □

We are going to introduce the normal integrand and see how we can calculate the Fenchel conjugate of an integral of a normal integrand.

Definition 4.5 (Normal integrand)**.** *Let B be a Borel subset of $\mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$ open, $f\colon \Omega \times B \to \overline{\mathbb{R}}$. The function f is called normal integrand if the two following conditions hold:*
a) for almost all $x \in \Omega$, $f(x, .)$ is lower semi-continuous on B,
b) there exists a Borel function, i.e. the inverse image of every closed set is a Borel set, $\bar{f}\colon \Omega \times B \to \overline{\mathbb{R}}$ such that $\bar{f}(x, .) = f(x, .)$ for almost every $x \in \Omega$.

Theorem 4.6. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, g a non-negative normal integrand of $\Omega \times \mathbb{R}^m$. The Fenchel conjugate of the functional $\mathcal{F}\colon L^p(\Omega, \mathbb{R}^m) \to \overline{\mathbb{R}}$, $\ u \mapsto \int_\Omega f(x, p(x))$ is under the assumption that there exists a $p_0 \in L^\infty(\Omega, \mathbb{R}^m)$ such that $\mathcal{F}(p_0) < +\infty$ given by*

$$\mathcal{F}^*(p) = \int_\Omega f^*(x, p(x)) \,\mathrm{d}x, \tag{4.7}$$

where $f^*\colon \Omega \times \mathbb{R}^m \to \overline{\mathbb{R}}$, $\ (x, y) \mapsto \sup_{v \in \mathbb{R}^m} y \cdot v - g(x, v)$ is the Fenchel conjugate of f.

Proof. Let us recall that the dual space of $L^q(\Omega, \mathbb{R}^m)$ is given by $L^{q'}(\Omega, \mathbb{R}^m)$ for $1 \leq q \leq \infty$, where $\frac{1}{q} + \frac{1}{q'} = 1$. We start the proof with fixing a p in $L^{q'}(\Omega, \mathbb{R}^m)$. We set

$$\Phi(x) = \sup_{\xi \in \mathbb{R}^m} \xi \cdot p(x) - f(x, \xi),$$

$$\Phi_n(x) = \max_{|\xi| \leq n} \xi \cdot p(x) - f(x, \xi).$$

The Φ_n's are clearly increasing and it is obvious that $\Phi_n(x) \overset{n\to\infty}{\longrightarrow} \Phi(x)$ for all x in Ω. Moreover, it holds for all $n \geq \|p_0\|_\infty$

$$\Phi_n(x) \geq \underbrace{p_0(x)p(x) - f(x, p_0(x))}_{:=\Psi(x,p_0(x))}.$$

It holds

$$\begin{aligned}\int_\Omega \Psi(x, p_0(x))\,\mathrm{d}x &= \int_\Omega p_0(x)p(x) - f(x, p_0(x))\,\mathrm{d}x \\ &\leq \|p_0\|_{L^\infty} \int_\Omega |p(x)|\,\mathrm{d}x + \int_\Omega f(x, p_0(x))\,\mathrm{d}x < \infty,\end{aligned}$$

since $p_0 \in L^\infty(\Omega, \mathbb{R}^m), p \in L^{q'}(\Omega, \mathbb{R}^m)$ and f is a normal integrand. Hence, Ψ is integrable over Ω and according to the measurable selection theorem there exists for all $n \in \mathbb{N}$ a measurable function $\bar{p}_n \colon \Omega \to \mathbb{R}^m$ such that $\|\bar{p}_n\|_\infty \leq n$ and it holds:

$$\Psi(x, \bar{p}_n(x)) = \bar{p}_n(x)p(x) - f(x, \bar{p}_n(x)) = \Phi_n(x).$$

It follows that Φ_n is measurable for all n and therefore that Φ is measurable and hence

$$\int_\Omega \Phi(x)\,\mathrm{d}x = \sup_{n\in\mathbb{N}} \int_\Omega \Phi_n(x)\,\mathrm{d}x = \sup_{n\in\mathbb{N}} \left| \int_\Omega \bar{p}_n(x)p(x) - f(x, \bar{p}_n(x))\,\mathrm{d}x \right|.$$

As for all $n \in \mathbb{N}$ it is $\bar{p}_n \in L^\infty(\Omega, \mathbb{R}^m) \subset L^q(\Omega, \mathbb{R}^m)$ we get

$$\int_\Omega \Phi(x)\,\mathrm{d}x \leq \sup_{u\in L^p} \left| \int_\Omega u(x)p(x) - f(x, u(x))\,\mathrm{d}x \right| = \mathcal{F}^*(p).$$

It remains to prove that $\mathcal{F}^*(p) \leq \int_\Omega \Phi(x)\,\mathrm{d}x$. By definition of Φ it holds for all $u \in L^q(\Omega, \mathbb{R}^m)$

$$u(x)p(x) - f(x, u(x)) \leq \Phi(x)$$

and therefore

$$\int_\Omega u(x)p(x) - f(x, u(x))\,\mathrm{d}x \leq \int_\Omega \Phi(x)\,\mathrm{d}x.$$

Since this inequality is true for all $u \in L^q(\Omega, \mathbb{R}^m)$, it remains true for the supremum:

$$\mathcal{F}^*(p) \leq \int_\Omega \Phi(x)\,\mathrm{d}x.$$

Because we have bounded $\mathcal{F}^*(p)$ from below and from above by $\int_\Omega \Phi(x)\,\mathrm{d}x$ the two terms must be equal. Looking again at the definition of $\Phi(x)$ we find that it is none other than $f^*(x, p(x))$. □

We are going to present a generalised differential for locally Lipschitz continuous functions that are not convex.

Definition 4.7 (Clarke directional derivative and subdifferential)**.** *Let $\mathcal{H}$ be a real Hilbert space, $D \subset \mathcal{H}$ be open, $\bar{x} \in D$ and $\mathcal{F}\colon D \to \mathbb{R}$. For $y \in \mathcal{H}$*

$$\mathcal{F}^{\circ}(\bar{x}; y) := \limsup_{\substack{t \downarrow 0 \\ x \to \bar{x}}} \frac{1}{\tau}(\mathcal{F}(x + \tau y) - \mathcal{F}(x)) \tag{4.8}$$

is called Clarke directional derivative of $\mathcal{F}$ at $\bar{x}$ in the direction y.
If $\mathcal{F}$ is locally Lipschitz continuous around $\bar{x}$, then the Clarke subdifferential of $\mathcal{F}$ at $\bar{x}$ is given by

$$\partial_{\circ} \mathcal{F}(\bar{x}) := \{x^* \in \mathcal{H} \mid \langle x^*, y\rangle \leq \mathcal{F}^{\circ}(\bar{x}; y) \forall y \in \mathcal{H}\}. \tag{4.9}$$

The next theorem yields that the Clarke subdifferential equals the convex subdifferential if the function is locally Lipschitz continuous and convex.

Theorem 4.8. *If $D \subset \mathcal{H}$ is convex and $\mathcal{F}$ is convex and locally Lipschitz continuous around $\bar{x} \in \mathcal{H}$, then it holds*

$$\partial \mathcal{F}(\bar{x}) = \partial_{\circ} \mathcal{F}(\bar{x}).$$

Proof. Let $\delta > 0$, then by the definition of the limit superior and by the fact that $\mathcal{F}$ is convex we get

$$\begin{aligned} \mathcal{F}^{\circ}(\bar{x}; y) &= \inf_{\epsilon \in (0, \epsilon_0)} \sup_{\substack{\tau \in (0, \epsilon) \\ x \in B(\bar{x}, \delta\epsilon)}} \frac{1}{\tau}(\mathcal{F}(x + \tau y) - \mathcal{F}(x)) \\ &= \inf_{\epsilon \in (0, \epsilon_0)} \sup_{x \in B(\bar{x}, \delta\epsilon)} \frac{1}{\epsilon}(\mathcal{F}(x + \epsilon y) - \mathcal{F}(x)). \end{aligned} \tag{4.10}$$

Let L be the Lipschitz constant of $\mathcal{F}$ around $\bar{x}$, then we obtain with the help of the triangle inequality

$$\begin{aligned} &\frac{1}{\epsilon} \left| \mathcal{F}(x + \epsilon y) - \mathcal{F}(x) - (\mathcal{F}(\bar{x} + \epsilon y) - \mathcal{F}(\bar{x})) \right| \\ \leq &\frac{1}{\epsilon}(|\mathcal{F}(x + \epsilon y) - \mathcal{F}(\bar{x} + \epsilon y)| + |\mathcal{F}(\bar{x}) - \mathcal{F}(x)|) \\ \leq &\frac{2}{\epsilon} L \|x - \bar{x}\| \leq 2\delta L \qquad\qquad \text{if } \|x - \bar{x}\| < \delta\epsilon \text{ for } \epsilon \text{ small enough.} \end{aligned}$$

The directional G-derivative is defined by $\mathcal{F}_G(\bar{x}; y) = \lim_{\tau \downarrow 0} \frac{1}{\tau}(\mathcal{F}(\bar{x} + \tau y) - \mathcal{F}(\bar{x}))$. The last estimate provides an estimation of (4.10) by the directional G-derivative:

$$\mathcal{F}^{\circ}(\bar{x}; y) \leq \inf_{\epsilon \in (0, \epsilon_0)} \frac{1}{\epsilon}(\mathcal{F}(\bar{x} + \epsilon y) - \mathcal{F}(\bar{x})) + 2\delta L = \mathcal{F}_G(\bar{x}; y) + 2\delta L.$$

$\delta > 0$ is arbitrary, hence we have $\mathcal{F}^{\circ}(\bar{x}; y) \leq \mathcal{F}_G(\bar{x}; y)$. It follows directly from the definitions of the two directional derivatives that the reversed inequality holds. Thus, we have $\mathcal{F}^{\circ}(\bar{x}; y) = \mathcal{F}_G(\bar{x}; y)$ and the claim follows with the fact that for a convex $\mathcal{F}$ the subdifferential can be written as $\partial \mathcal{F}(\bar{x}) = \{u \in \mathcal{H} \mid \langle u, y\rangle \leq \mathcal{F}_G(\bar{x}; y) \forall y \in \mathcal{H}\}$ [see Schirotzek [13], Prop. 4.1.6 for a proof]. ☐

In a finite dimensional space as $\mathbb{R}^n$ Rademacher's theorem yields that a locally Lipschitz continuous function is almost everywhere differentiable. This provides the following characterisation of Clarke's subdifferential.

Theorem 4.9. *If $\mathcal{F}\colon\ \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz continuous around $\bar{x}$, let $\Omega_{\mathcal{F}} \subset \mathbb{R}^n$ denote the null set, where $\mathcal{F}$ is not differentiable, and let $S \subset \mathbb{R}^n$ be also a null set, then we have*

$$\partial_\circ \mathcal{F}(\bar{x}) = co\left\{ \lim_{k\to\infty} \mathcal{F}'(x_k) | x_k \to \bar{x}, x_k \notin \Omega_{\mathcal{F}} \cup S \right\}. \tag{4.11}$$

See Clarke [4] theorem 2.5.1 p. 63 for a proof of this theorem.
At this point we have collected all definitions and theorems from convex analysis that we are going to use in the sequel. Our next step is to list all definitions and theorems from monotone operator theory that we need in chapter 6 and 7.

5. Monotone Operators

For the proof of the existence of a solution for our multi-bang control problem in chapter 6 we need some basics from monotone operator theory. The theory of this section follows some parts of chapters 20, 21 and 24 of Bauschke and Combettes [1]. Whenever we have a Hilbert space $\langle \cdot, \cdot \rangle$ denotes the inner product.

Definition 5.1. *Let $\mathcal{H}$ be a real Hilbert space and $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. A subset U of $\mathcal{H} \times \mathcal{H}$ is monotone if*

$$\forall (x, u), (y, v) \in U \;\; \langle x - y, u - v \rangle \geq 0.$$

The operator A is monotone if the graph of A, $\operatorname{graph} A = \{(x, u) \in \mathcal{H} \times \mathcal{H} \,|\, u \in Ax\}$, is a monotone set.

Furthermore, we need to define when a monotone operator is maximally monotone.

Definition 5.2. *Let $\mathcal{H}$ be a real Hilbert space and $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator. A is called maximally monotone if it holds for each $(x, u) \in \mathcal{H} \times \mathcal{H}$ that $(x, u) \in \operatorname{graph} A$ if and only if*

$$\langle x - y, u - v \rangle \geq 0 \;\forall (y, v) \in \operatorname{graph} A.$$

Our next step is to prove that the subdifferential, defined in 4.3, of a proper lower semi-continuous and convex function is maximally monotone.

Theorem 5.3. *Let $\mathcal{H}$ be a real Hilbert space and $\mathcal{F}\colon \mathcal{H} \to (-\infty, +\infty]$ be a proper lower semi-continuous and convex functional. Then the subdifferential $\partial \mathcal{F}$ as defined in 4.3 is maximally monotone.*

For the proof of this theorem we need Minty's theorem. This theorem is recited from Bauschke and Combettes [1, Thm. 21.1].

Theorem 5.4 (Minty's theorem)**.** *Let $\mathcal{H}$ be a real Hilbert space and $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator. A is maximally monotone if and only if*

$$\operatorname{ran}(\operatorname{Id} + A) = \mathcal{H},$$

where $\operatorname{ran}(\cdot)$ *is the range of the operator.*

We do not want to prove this theorem, since this is beyond the limits of this thesis. With Minty's theorem we are able to prove 5.3.

Proof. The first step is to verify that $\partial\mathcal{F}$ is a monotone operator. We take two elements $(x,u),(y,v)$ in graph$(\partial\mathcal{F})$. Then by the definition of the subdifferential 4.3 it holds

$$\mathcal{F}(y) \geq \mathcal{F}(x) + \langle y-x, u\rangle$$

and

$$\mathcal{F}(x) \geq \mathcal{F}(y) + \langle x-y, v\rangle .$$

Adding these two inequalities yields

$$\langle x-y, u-v\rangle \geq 0.$$

Hence, by definition 5.1 $\partial\mathcal{F}$ is monotone.
The next step is to show that $\operatorname{ran}(\mathrm{Id}+\partial\mathcal{F}) = \mathcal{H}$. It is clear that $\operatorname{ran}(\mathrm{Id}+\partial\mathcal{F}) = \operatorname{dom}(\mathrm{Id}+\partial\mathcal{F})^{-1}$. We use at this point without a proof that $\operatorname{prox}_{\mathcal{F}} = (\mathrm{Id}+\partial\mathcal{F})^{-1}$, where $\operatorname{prox}_{\mathcal{F}}$ is the proximal mapping of $\mathcal{F}$ [see 7.4 for the definition of the proximal mapping and 7.7 for the proof of this statement]. It follows from the definition of the proximity operator that $\operatorname{dom} Prox_{\mathcal{F}} = \mathcal{H}$ for a proper lower semi-continuous and convex functional $\mathcal{F}$. Thus, we can conclude that $\operatorname{ran}(\mathrm{Id}+\partial\mathcal{F}) = \mathcal{H}$. Since we verified all requirements of Minty's theorem it follows from 5.4, that $\partial\mathcal{F}$ is maximally monotone for every proper lower semi-continuous and convex functional $\mathcal{F}$. □

We want to state a condition under which a maximally monotone operator is surjective. For this we need the Rockafellar-Veselý theorem. It is recited from Bauschke and Combettes [1, Thm. 21.15], they also give a proof of this theorem.

Theorem 5.5 (Rockafellar-Veselý)**.** *Let $\mathcal{H}$ be a real Hilbert space, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone and let $x \in \mathcal{H}$. Then A is locally bounded at x if and only if $x \notin boundary(\operatorname{dom} A)$.*

Corollary 5.6. *Let $\mathcal{H}$ be a real Hilbert space and let A, that maps from $\mathcal{H}$ to $2^{\mathcal{H}}$, be maximally monotone. A is surjective, if the following holds:*

$$\lim_{\|x\|\to\infty} \inf \|Ax\| = \infty. \tag{5.1}$$

Proof. Our first step is to show that A is surjective if and only if A^{-1} is locally bounded everywhere on $\mathcal{H}$. Assume that A^{-1} is locally bounded everywhere on $\mathcal{H}$. By the Rockafellar-Veselý theorem we get that for all $x \in \mathcal{H}$, $x \notin boundary(\operatorname{dom} A^{-1})$. Thus, $\operatorname{dom} A^{-1} = \mathcal{H}$ and therefore $\operatorname{ran} A = \mathcal{H}$. This yields that A is surjective.
Assume now that A is surjective. Then $\operatorname{dom} A^{-1} = \mathcal{H}$ and it yields that
$x \notin boundary(\operatorname{dom} A^{-1})$. Again by Rockafellar-Veselý A^{-1} is locally bounded everywhere on $\mathcal{H}$.
Consequently, we assume that there exists a $u \in \mathcal{H}$ such that A^{-1} is not locally bounded at u. Then there exists a sequence $(x_n, u_n) \in \operatorname{graph} A$, such that $u_n \to u$ and $\|x_n\| \to \infty$. Ergo, $\infty = \lim_{\|x_n\|\to\infty} \inf \|Ax_n\| \leq \lim \|u_n\| = \|u\|$. This is a contradiction to the definition of u. As a deduction we get that A^{-1} must be locally bounded everywhere on $\mathcal{H}$ and thus, that A is surjective. □

In addition we need to think about sums and compositions of maximally monotone operators.

Theorem 5.7. *Let A and B be maximally monotone operators from a real Hilbert space $\mathcal{H}$ to $2^{\mathcal{H}}$. $A+B$ is maximally monotone if*

$$\operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) = \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B),$$

where $\operatorname{cone}(\cdot)$ *is the conical hull of the set, i.e. the smallest cone in $\mathcal{H}$ that contains the set.*

Proof. The Fitzpatrick function of a monotone operator A from $\mathcal{H}$ to $2^{\mathcal{H}}$ is given by

$$F_A\colon \mathcal{H}\times\mathcal{H} \to [-\infty,+\infty], \quad (x,u) \mapsto \sup_{(y,v)\in \operatorname{graph} A} (\langle y,u\rangle + \langle x,v\rangle - \langle y,v\rangle).$$

We define the projection onto the first argument $Q_1\colon \mathcal{H}\times\mathcal{H} \to \mathcal{H},\ (x,u)\mapsto x$. For a maximally monotone operator A it holds

$$\operatorname{dom} A \subset Q_1(\operatorname{dom} F_A).$$

To verify this take $x \in \operatorname{dom} A$, then there exists a $u \in \mathcal{H}$ such that $(x,u) \in \operatorname{graph} A$. By the definition of the Fitzpatrick function it follows $F_A(x,u) = \sup_{y\in\mathcal{H}}(\langle y, Ax\rangle + \langle x, Ay\rangle - \langle y, Ay\rangle) = \sup_{y\in\mathcal{H}}(\langle x, Ax\rangle - \langle x-y, Ax-Ay\rangle)$. The supremum is $\langle x,u\rangle$ since $\langle x-y, Ax-Ay\rangle \geq 0$. The positivity is given because A is maximally monotone. This implies that $x \in Q_1(\operatorname{dom} F_A)$. Furthermore, it holds

$$Q_1(\operatorname{dom} F_A) \subset \overline{\operatorname{dom}}\ A.$$

Then we get

$$\begin{aligned}
\operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) &\subset \operatorname{cone}(Q_1 \operatorname{dom} F_A - Q_1 \operatorname{dom} F_B)\\
&\subset \overline{\operatorname{span}}(Q_1(\operatorname{dom} F_A) - Q_1(\operatorname{dom} F_B))\\
&\subset \overline{\operatorname{span}}(\overline{\operatorname{dom}}\ A - \overline{\operatorname{dom}}\ B)\\
&= \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B)\\
&= \operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B).
\end{aligned}$$

It follows from this that $0 \in \operatorname{sri}(Q_1(\operatorname{dom} F_A) - Q_1(\operatorname{dom} F_B))$, where sri denotes the strong relative interior. Hence, we get that $A+B$ is maximally monotone. For a proof of this implication see Bauschke and Combettes [1, Thm. 24.2]. □

There are three conditions which imply the condition of theorem 5.7.

Corollary 5.8. *Let $\mathcal{H}$ be a real Hilbert space, $A, B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone. If one of the following holds*

i) $\operatorname{dom} B = \mathcal{H}$, *ii)* $\operatorname{dom} A \cap \operatorname{int}(\operatorname{dom} B) \neq \emptyset$,

iii) $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$,

then $A + B$ *is maximally monotone.*

Proof. Clearly (i) implies (ii). Since $\operatorname{dom} A \cap \operatorname{int}(\operatorname{dom} B)$ is not empty, there exists at least one element such that $x \in \operatorname{dom} A$ and $x \in \operatorname{int} \operatorname{dom} B$. Hence, $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$ because $\operatorname{int}(\operatorname{dom} B)$ is open. It follows from (iii) that $\operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) = \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B)$. Therefore, by theorem 5.7 $A + B$ is maximally monotone. $\square$

The next theorem of this section deals with the composition of a maximally monotone operator and a linear operator. We recite this theorem from Bauschke and Combettes [1, Thm. 24.5].

Theorem 5.9. *Let* $\mathcal{H}$ *and* $\mathcal{K}$ *be real Hilbert spaces,* $A\colon \mathcal{K} \to 2^{\mathcal{K}}$ *be a maximally monotone operator and let* L *be a bounded linear operator from* $\mathcal{H}$ *to* $\mathcal{K}$ *with domain* $\mathcal{H}$ *such that* $\operatorname{cone}(\operatorname{ran} L - \operatorname{dom} A) = \overline{\operatorname{span}}(\operatorname{ran} L - \operatorname{dom} A)$. *Then* L^*AL *is maximally monotone, where* L^* *is the adjoint operator.*

Proof. The Hilbert direct sum of two Hilbert spaces is given by

$$\mathcal{H} \oplus \mathcal{K} := \left\{ (x_h, x_k) \in \mathcal{H} \times \mathcal{K} \mid \|x_h\|^2_{\mathcal{H}} + \|x_k\|^2_{\mathcal{K}} < +\infty \right\}.$$

$\mathcal{H} \oplus \mathcal{K}$ is a Hilbert space and the inner product is given by

$$\langle (x_h, x_k), (y_h, y_k) \rangle = \langle x_h, y_h \rangle_{\mathcal{H}} + \langle x_k, y_k \rangle_{\mathcal{K}}.$$

We are going to write $\mathcal{H} \times \mathcal{K}$ instead of $\mathcal{H} \oplus \mathcal{K}$.
Let $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $A\colon \mathcal{K} \to 2^{\mathcal{K}}$ be maximally monotone. Then $B \times A\colon \mathcal{H} \times \mathcal{K} \to 2^{\mathcal{H} \times \mathcal{K}}, (x, y) \mapsto Bx \times Ay$ is clearly maximally monotone by the definition of the inner product for the Hilbert direct sum.
Set $B = N_{\operatorname{graph} L}$ the normal cone operator of $\operatorname{graph} L$, that is defined by

$$N_{\operatorname{graph} L}(x_h, x_k) = \begin{cases} \{(u_h, u_k) \in \mathcal{H} \times \mathcal{K} \mid \sup \langle \operatorname{graph} L - (x_h, x_k), (u_h, u_k) \rangle \leq 0\} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } (x_h, x_k) \in \operatorname{graph} L, \\ \emptyset \qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise.} \end{cases}$$

From the definition of the normal cone operator follows directly that $N_{\operatorname{graph} L}$ is maximally monotone. Further we set

$$C\colon \mathcal{H} \times \mathcal{K} \to 2^{\mathcal{H} \times \mathcal{K}}, (x, y) \mapsto \{0\} \times Ay.$$

C is also maximally monotone by the statement from above about the product of two maximally monotone operators, since the zero-operator is clearly maximally monotone. We have $\operatorname{dom} B = \operatorname{graph} L$ and $\operatorname{dom} C = \mathcal{H} \times \operatorname{dom} A$ and thus $\operatorname{dom}(B + C) \subset \operatorname{graph} L$.

Claim:
For all $(x,u) \in \mathcal{H} \times \mathcal{H}$ and for all $v \in \mathcal{K}$ it holds

$$(u,v) \in (B+C)(x,Lx) \Leftrightarrow u + L^*v \in L^*(A(Lx)). \tag{5.2}$$

Take $(x,u) \in \mathcal{H} \times \mathcal{H}$ and $v \in \mathcal{K}$. Then $(B+C)(x,Lx) = N_{\operatorname{graph} L}(x,Lx) + (\{0\} \times A(Lx))$. Firstly, we take a closer look on

$$\begin{aligned} N_{\operatorname{graph} L}(x,Lx) &= \left\{ (u_h,u_k) \in \mathcal{H} \times \mathcal{K} \middle| \sup_{y \in \mathcal{H}} \langle (y,Ly) - (x,Lx), (u_h,u_k) \rangle \le 0 \right\} \\ &= \left\{ (u_h,u_k) \in \mathcal{H} \times \mathcal{K} \middle| \sup_{y \in \mathcal{H}} \langle y - x, u_h + L^*u_k \rangle \le 0 \right\}_{\mathcal{H}} \\ &= \{(L^*w, -w) | w \in \mathcal{K}\}. \end{aligned}$$

Then we have

$$\begin{aligned} (u,v) \in (B+C)(x,Lx) &\Leftrightarrow (u,v) \in N_{\operatorname{graph} L}(x,Lx) + (\{0\} \times A(Lx)) \\ &\Leftrightarrow (u,v) \in \{(L^*w,-w)|w \in \mathcal{K}\} + (\{0\} \times A(Lx)) \\ &\Leftrightarrow \exists w \in \mathcal{K} \text{ s.t. } u = L^*w \text{ and } v + w \in A(Lx) \\ &\Leftrightarrow u + L^*v \in L^*(A(Lx)). \end{aligned}$$

This proves (5.2). As we already mentioned above it is $\operatorname{dom} B - \operatorname{dom} C = \operatorname{graph} L - (\mathcal{H} \times \operatorname{dom} A) = \mathcal{H} \times (\operatorname{ran} L - \operatorname{dom} A)$ and since $\operatorname{cone}(\operatorname{ran} L - \operatorname{dom} A) = \overline{\operatorname{span}}(\operatorname{ran} L - \operatorname{dom} A)$ it follows $\operatorname{cone}(\operatorname{dom} B - \operatorname{dom} C) = \overline{\operatorname{span}}(\operatorname{dom} B - \operatorname{dom} C)$. Ergo, we can deduce with theorem 5.7 that B + C is maximally monotone. Take (z,w) such that

$$\forall x \in \mathcal{H} \quad \inf \langle x - z, L^*(A(Lx)) - w \rangle \ge 0. \tag{5.3}$$

Additionally choose $(x,u) \in \mathcal{H} \times \mathcal{H}, v \in \mathcal{K}$ such that $((x,Lx),(u,v)) \in \operatorname{graph}(B+C)$. It follows with (5.2) that $u + L^*v \in L^*(A(Lx))$ and we get from (5.3)

$$\begin{aligned} 0 &\le \langle x - z, u + L^*v - w \rangle \\ &= \langle x - z, u - w \rangle + \langle Lx - Lz, v - 0 \rangle \\ &= \langle (x,Lx) - (z,Lz), (u,v) - (w,0) \rangle. \end{aligned}$$

We can deduce by the maximal monotonicity of $B+C$ that $((z,Lz),(w,0)) \in \operatorname{graph}(B+C)$ and with (5.2) it follows $w \in L^*(A(Lz))$, i.e. $(z,w) \in \operatorname{graph}(L^*(A(Lx)))$ and therefore L^*AL is maximally monotone. □

With this theorem we have all the knowledge that we need to prove that our multi-bang problem in chapter 6 has a solution.
We want to list additionally a lemma from Brezis, Crandall and Pazy [3] that we are going to need for a proof in chapter 7.

Theorem 5.10. *Let $\mathcal{H}$ be a Hilbert space and $B \subset \mathcal{H} \times \mathcal{H}$ be maximally monotone. If (u_n, v_n) is a sequence in B where each component converges weakly, i.e. $u_n \rightharpoonup u, v_n \rightharpoonup v$ and it holds*

$$\limsup_{n,m\to\infty} \langle u_n - u_m, v_n - v_m \rangle \leq 0. \tag{5.4}$$

Then $(u, v) \in B$.

Proof. We have

$$\lim_{n,m\to\infty} \langle u_n - u_m, v_n - v_m \rangle = 0$$

since B is monotone.
Let (u_{n_i}, v_{n_i}) be a subsequence of (u_n, v_n) such that $\langle u_{n_i}, v_{n_i} \rangle \to L$. The first fact together with (5.4) yield

$$\begin{aligned} 0 &= \lim_{n_i\to\infty} \left(\lim_{n_k\to\infty} \langle u_{n_i} - u_{n_k}, v_{n_i} - v_{n_k} \rangle \right) \\ &= \lim_{n_i\to\infty} \langle u_{n_i}, v_{n_i} \rangle - \langle u_{n_i}, v \rangle - \langle u, v_{n_i} \rangle + L \\ &= 2L - 2 \langle u, v \rangle . \end{aligned}$$

Thus $L = \langle u, v \rangle$ and $\langle u_n, v_n \rangle \to \langle u, v \rangle$. From this we can deduce that
$\langle x - u, y - v \rangle \geq 0 \; \forall \; (x, y) \in B$. This implies that $(u, v) \in B$ by the maximal monotonicity of B. $\square$

6. Multi-Bang Control Problem

This chapter is based on the ideas of Clason and Kunisch [7].
As mentioned in the introduction a bang-bang control is a control variable that attains its control constrains u_1 and u_2 almost everywhere. In analogy to this we want to introduce a multi-bang control. This is as already stated before a control that takes on almost everywhere values from a discrete set of given control states u_i [cf. Clason and Kunisch [7], Introduction]. Therefore, we want to find a solution for the following optimal control problem:

$$\begin{cases} \min\limits_{u,y \in L^2(\Omega,\mathbb{R}^2)} \dfrac{\alpha}{2} \|y - z\|_{L^2}^2 + \|u\|_{L^2}^2 + \displaystyle\int_\Omega \delta_B(u(x))\,\mathrm{d}x, \\ \text{s.t. } Ay = u, \end{cases} \tag{6.1}$$

for a given open and bounded subset Ω of $\mathbb{R}^2$ and a given target $z \in L^2(\Omega, \mathbb{R}^2)$.

$$\begin{aligned} B &= \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \end{pmatrix} \right\} \\ &:= \{u_1, u_2, \cdots, u_8\}. \end{aligned}$$

Here δ_B is the indicator function of B given by

$$\delta_B(t) = \begin{cases} 0 & \text{if } t \in B, \\ \infty & \text{otherwise.} \end{cases}$$

$A\colon H^1_\Gamma \to (H^1_\Gamma)^*, Ay = u$ is the state equation and is assumed to be the linearised elasticity equation with homogeneous Neumann conditions on Γ_1

$$\begin{aligned} -2\mu \operatorname{div} \epsilon(y) - \lambda \operatorname{grad} \operatorname{div} y &= u \text{ in } \Omega, \\ y &= 0 \text{ on } \Gamma_0, \\ \nabla y \cdot n &= 0 \text{ on } \Gamma_1, \end{aligned} \tag{6.2}$$

where ∇y is the deformation gradient

$$\nabla y = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{pmatrix}$$

and Id is the identity on $\mathbb{R}^{2\times 2}$. We recall the definition of the symmetrized gradient $\epsilon(y) = \frac{1}{2}(\nabla y + \nabla y^T)$. Γ_0 and Γ_1 are parts of the boundary of Ω as defined in figure 2.1. We think back to the definition of H^1_Γ that is given in section 2.1:

$$H^1_\Gamma := \left\{ v \in H^1(\Omega)^2 \colon v(x) = 0 \text{ for } x \in \Gamma_0 \right\}.$$

In this section we also verified that our state equation has a unique solution for all given controls u, such that we can write $y = A^{-1}u$.
We are able to split our cost functional in two parts. One part is the distance between our state and the target

$$\mathcal{F}\colon L^2(\Omega, \mathbb{R}^2) \to \mathbb{R}, \quad u \mapsto \frac{\alpha}{2} \left\| A^{-1}u - z \right\|_{L^2}^2 .$$

This functional is easy to deal with, since it is Fréchet differentiable. The other part takes care of the form of our control u:

$$\tilde{\mathcal{G}}\colon L^2(\Omega, \mathbb{R}^2) \to \overline{\mathbb{R}}, \quad u \mapsto \|u\|_{L^2}^2 + \int_\Omega \delta_B(u(x))\,\mathrm{d}x.$$

$\tilde{\mathcal{G}}$ is challenging since it is neither lower semi-continuous nor convex. Thus, we are going to look at the convex relaxation $\mathcal{G}$ of $\tilde{\mathcal{G}}$, which is given by

$$\mathcal{G}\colon L^2(\Omega, \mathbb{R}^2) \to \overline{\mathbb{R}}, \quad u \mapsto \int_\Omega \max(\|u\|_\infty, 1)\,\mathrm{d}x + \delta_U(u),$$

where δ_U denotes the indicator function of the admissible set

$$U := \left\{ u \in L^2(\Omega) : \|u\|_\infty \leq 2 \right\}.$$

As we are going to see later we can find conditions under which the control of the problem with $\mathcal{G}$ instead of $\tilde{\mathcal{G}}$ also fulfils the requirement that it takes only values from B.
Hence, we are going to analyse the following optimal control problem:

$$\begin{cases} \displaystyle\min_{u,y \in L^2(\Omega,\mathbb{R}^2)} \frac{\alpha}{2} \|y - z\|_{L^2}^2 + \int_\Omega \max(\|u\|_\infty, 1)\,\mathrm{d}x, \\ \text{s.t. } Ay = u, \ \|u\|_\infty \leq 2. \end{cases} \tag{6.3}$$

With the given definitions of $\mathcal{F}$ and $\mathcal{G}$ we can rewrite our system as

$$\min_u \mathcal{F}(u) + \mathcal{G}(u). \tag{6.4}$$

With the definitions of the Fenchel conjugate 4.1 and the subdifferential 4.3 we can state the primal-dual optimality system.

Theorem 6.1. *The necessary optimality conditions for* (6.4) *are:*

$$\begin{cases} \bar{p} = -\mathcal{F}'(\bar{u}), \\ \bar{u} \in \partial \mathcal{G}^*(\bar{p}). \end{cases} \tag{6.5}$$

Proof. The necessary optimality condition is $0 \in \mathcal{F}'(\bar{u}) + \partial \mathcal{G}(\bar{u})$. We set $\bar{p} = -\mathcal{F}'(\bar{u})$ since $\mathcal{F}$ is Fréchet differentiable. As a result we get the optimality system

$$\begin{cases} \bar{p} & = -\mathcal{F}'(\bar{u}), \\ \bar{p} & \in \partial \mathcal{G}(\bar{u}). \end{cases}$$

By theorem 4.4 we get that the second term is equivalent to $\bar{u} \in \partial \mathcal{G}^*(\bar{p})$. □

$\partial \mathcal{G}^*$ is convex and lower semi-continuous according to theorem 4.2. Hence, the primal-dual optimality system is well-defined. Therefore, we can claim the formal optimality system for (6.3).

6.1. Formal Optimality System

We remember our functions:

$$\begin{aligned} \mathcal{F}\colon L^2(\Omega) \to \mathbb{R}, \qquad & u \mapsto \frac{\alpha}{2} \left\| A^{-1}u - z \right\|_{L^2}^2, \\ \mathcal{G}\colon L^2(\Omega) \to \overline{\mathbb{R}}, \qquad & u \mapsto \int_\Omega \max(\|u\|_\infty, 1)\,\mathrm{d}x + \delta_U(u). \end{aligned} \tag{6.6}$$

In consonance with (6.5) we have to calculate the Fréchet derivative of $\mathcal{F}$ and the subdifferential of the Fenchel conjugate of $\mathcal{G}$.

6.1.1. Fenchel Conjugate and Subdifferential

By definition 4.1 it is

$$\mathcal{G}^*(p) = \sup_{u \in L^2(\Omega, \mathbb{R}^2)} \langle u, p \rangle_{L^2} - \mathcal{G}(u).$$

Since $\mathcal{G}$ is the integral of the function

$$g\colon \mathbb{R}^2 \to \overline{\mathbb{R}}, \quad v \mapsto \max(\max(|v_1|, |v_2|), 1) + \delta_{\|v\|_\infty \leq 2}(v)$$

we can compute the Fenchel conjugate pointwise as stated in theorem 4.6:

$$\mathcal{G}^*(p) = \int_\Omega g^*(p(x))\,\mathrm{d}x.$$

We need to confirm the requirements of 4.6. g is clearly non-negative. Ergo, it remains to verify that g is a normal integrand. The function is convex and therefore it is lower semi-continuous. Furthermore, g is a Borel function itself since the inverse image of every closed subset of $\overline{\mathbb{R}}$ is a Borel set. Consequently, g is a normal integrand as stated in 4.5. Thus, we can apply theorem 4.6 and have to calculate the Fenchel conjugate of g.

For this we assume that the supremum for given q is attained at $\bar{v}$. We have to distinguish the following two cases:

(i) $0 \leq \|\bar{v}\|_\infty \leq 1$. Then $g(\bar{v}) = 1$ and therefore

$$g_1^*(q) = \|q\|_{\ell^1} - 1.$$

(ii) $1 \leq \|\bar{v}\|_\infty \leq 2$. Then $g(\bar{v}) = \|\bar{v}\|_\infty$ and hence

$$g^*(q) = \sup_{1 \leq |v_1|, |v_2| \leq 2} q_1 \cdot v_1 + q_2 \cdot v_2 - \max(|v_1|, |v_2|).$$

We can assume without restriction of generality that $|v_1| \geq |v_2|$. It follows that

$$\begin{aligned} g_2^*(q) &= |v_1| (\|q\|_{\ell^1} - 1) \\ &= \begin{cases} \|q\|_{\ell^1} - 1 & \text{if } \|q\|_{\ell^1} \leq 1, \\ 2(\|q\|_{\ell^1} - 1) & \text{if } \|q\|_{\ell^1} \geq 1. \end{cases} \end{aligned}$$

We have to decide which of these cases gives the biggest value according to q, since $g^*(q) = \max\{g_1^*(q), g_{21}^*(q), g_{22}^*(q)\}$. For $\|q\|_{\ell^1} \leq 1$ it is $g_1^*(q) = g_{21}^*(q)$ and for $\|q\|_{\ell^1} \geq 1$ the maximal value is $g_{22}^*(q)$. From this it follows:

$$g^*(q) = \begin{cases} \|q\|_{\ell^1} - 1 & \text{if } \|q\|_{\ell^1} \leq 1, \\ 2(\|q\|_{\ell^1} - 1) & \text{if } \|q\|_{\ell^1} \geq 1. \end{cases} \tag{6.7}$$

We are going to verify that the Fenchel conjugate g^* is locally Lipschitz continuous, so that we can use theorem 4.8 to calculate the subdifferential.

It is clear that g^* is Lipschitz continuous for $\|q\|_{\ell^1} < 1$ and for $\|q\|_{\ell^1} > 1$. Because of this we have to look only at a neighbourhood U of $\|q\|_{\ell^1} = 1$. Let $q_1, q_2 \in U$ such that $\|q_1\|_{\ell^1} \leq 1$ and $\|q_2\|_{\ell^1} \geq 1$. We get

$$\begin{aligned} |g^*(q_1) - g^*(q_2)| &= |\, \|q_1\|_{\ell^1} + 1 - 2\|q_2\|_{\ell^1} \,| \\ &= 2\|q_2\|_{\ell^1} - \|q_1\|_{\ell^1} - 1 \\ &\leq 2|\, \|q_2\|_{\ell^1} - \|q_1\|_{\ell^1} \,| \leq 2\|q_1 - q_2\|_{\ell^1}. \end{aligned}$$

Hence, g^* is locally Lipschitz continuous and we can calculate the Clarke subdifferential instead of the convex subdifferential according to theorem 4.8. In addition g is a function on $\mathbb{R}^2$. Thus, we can use theorem 4.9 to derive the Clarke subdifferential.
We start with the case $\|q\|_{\ell^1} < 1$. We have to take care of four different cases. The gradient for $q_1 \neq 0$ and $q_2 \neq 0$ is

$$\nabla g^*(q) = \begin{pmatrix} \operatorname{sgn}(q_1) \\ \operatorname{sgn}(q_2) \end{pmatrix}.$$

If $q_1 = 0$ and $q_2 \in (-1, 1), q_2 \neq 0$, Clarke's generalized gradient is by definition (4.11) the convex hull of $(-1, \operatorname{sgn}(q_2))^T$ and $(1, \operatorname{sgn}(q_2))^T$:

$$\nabla g^*(q) = \begin{pmatrix} -1 \\ \operatorname{sgn}(q_2) \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \ \lambda \in [0, 1].$$

Analogously we get for $q_1 \in (-1,1), q_1 \neq 0$ and $q_2 = 0$

$$\nabla g^*(q) = \begin{pmatrix} \operatorname{sgn}(q_1) \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \; \lambda \in [0,1] \, .$$

The convex hull of the last two cases is Clarke's generalized gradient for $q_1 = q_2 = 0$

$$\nabla g^*(q) = [-1,1] \times [-1,1].$$

For the second case we have to discriminate three further cases to calculate the Clarke's generalized gradient.

$$\nabla g^*(q) = \begin{pmatrix} 2\operatorname{sgn}(q_1) \\ 2\operatorname{sgn}(q_2) \end{pmatrix}, \text{ if } q_1 \neq 0 \text{ and } q_2 \neq 0.$$

By theorem 4.9 the Clarke's generalized gradient is given by the convex hull of both gradients if one component of q is zero:

$$\nabla g^*(q) = \begin{pmatrix} -2 \\ 2\operatorname{sgn}(q_2) \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \; \lambda \in [0,1] \quad \text{if } q_1 = 0 \text{ and } q_2 \in (-\infty,-1) \cup (1,\infty),$$

$$\nabla g^*(q) = \begin{pmatrix} 2\operatorname{sgn}(q_1) \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \; \lambda \in [0,1] \quad \text{if } q_1 \in (-\infty,-1) \cup (1,\infty) \text{ and } q_2 = 0.$$

The subdifferential is by 4.8 and 4.9

$$
\partial g^*(q) = \overline{co}(\{\nabla g_1^*(q)\} \cup \{\nabla g_2^*(q)\})
$$
$$
= \begin{cases}
\begin{pmatrix} \operatorname{sgn}(q_1) \\ \operatorname{sgn}(q_2) \end{pmatrix} & \text{if } \|q\|_{\ell^1} < 1 \text{ and } q_1, q_2 \neq 0, \\
\begin{pmatrix} -1 \\ \operatorname{sgn}(q_2) \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \lambda \in [0,1] & \text{if } q_1 = 0 \text{ and } |q_2| \in (0,1), \\
\begin{pmatrix} \operatorname{sgn}(q_1) \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \lambda \in [0,1] & \text{if } |q_1| \in (0,1) \text{ and } q_2 = 0, \\
[-1,1] \times [-1,1] & \text{if } q_1 = q_2 = 0, \\
\begin{pmatrix} 2\operatorname{sgn}(q_1) \\ 2\operatorname{sgn}(q_2) \end{pmatrix} & \text{if } \|q\|_{\ell^1} > 1 \text{ and } q_1, q_2 \neq 0, \\
\begin{pmatrix} -2 \\ 2\operatorname{sgn}(q_2) \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \lambda \in [0,1] & \text{if } q_1 = 0 \text{ and } |q_2| \in (1,\infty), \\
\begin{pmatrix} 2\operatorname{sgn}(q_1) \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \lambda \in [0,1] & \text{if } |q_1| \in (1,\infty) \text{ and } q_2 = 0, \\
(2-\lambda) \begin{pmatrix} \operatorname{sgn}(q_1) \\ \operatorname{sgn}(q_2) \end{pmatrix}, \lambda \in [0,1] & \text{if } \|q\|_{\ell^1} = 1 \text{ and } q_1, q_2 \neq 0, \\
\begin{pmatrix} (2-\delta)\mu \\ (2-\delta)\operatorname{sgn}(q_2) \end{pmatrix}, \delta \in [0,1], \mu \in [-1,1] & \text{if } q_1 = 0 \text{ and } |q_2| = 1, \\
\begin{pmatrix} (2-\delta)\operatorname{sgn}(q_1) \\ (2-\delta)\mu \end{pmatrix}, \delta \in [0,1], \mu \in [-1,1] & \text{if } |q_1| = 1 \text{ and } q_2 = 0.
\end{cases} \tag{6.8}
$$

The different cases of the subdifferential are visualised in figure 6.1.

6.1.2. Fréchet Derivative

We already computed $\partial\, \mathcal{G}^*$, so it remains according to (6.5) to compute the Fréchet derivative of $\mathcal{F}(u) = \frac{\alpha}{2} \left\| A^{-1}u - z \right\|_{L^2}^2$. We can eliminate the state variable y since we have shown in section 2.1 that there exists a solution of the linearised elasticity equation (6.2) for every given control u. In line with Clason [5] the Gateaux derivative is given by

$$
\begin{aligned}
\frac{d}{dt}\mathcal{F}(u + t\phi)\big|_{t=0} &= \frac{d}{dt}\left(\frac{\alpha}{2}\left\|A^{-1}(u+t\phi) - z\right\|_{L^2}^2\right)\big|_{t=0} \\
&= \frac{d}{dt}(\frac{\alpha}{2}\int_\Omega |A^{-1}(u+t\phi) - z|^2 \,\mathrm{d}x)\big|_{t=0} \\
&= \frac{\alpha}{2}\int_\Omega 2A^{-1}uA^{-1}\phi + 2t(A^{-1}\phi)^2 - 2A^{-1}\phi z \,\mathrm{d}x\big|_{t=0} \\
&= \alpha \left\langle A^{-1}u - z, A^{-1}\phi \right\rangle_{L^2} \\
&= \left\langle \alpha A^{-*}(A^{-1}u - z), \phi \right\rangle_{L^2}.
\end{aligned}
$$

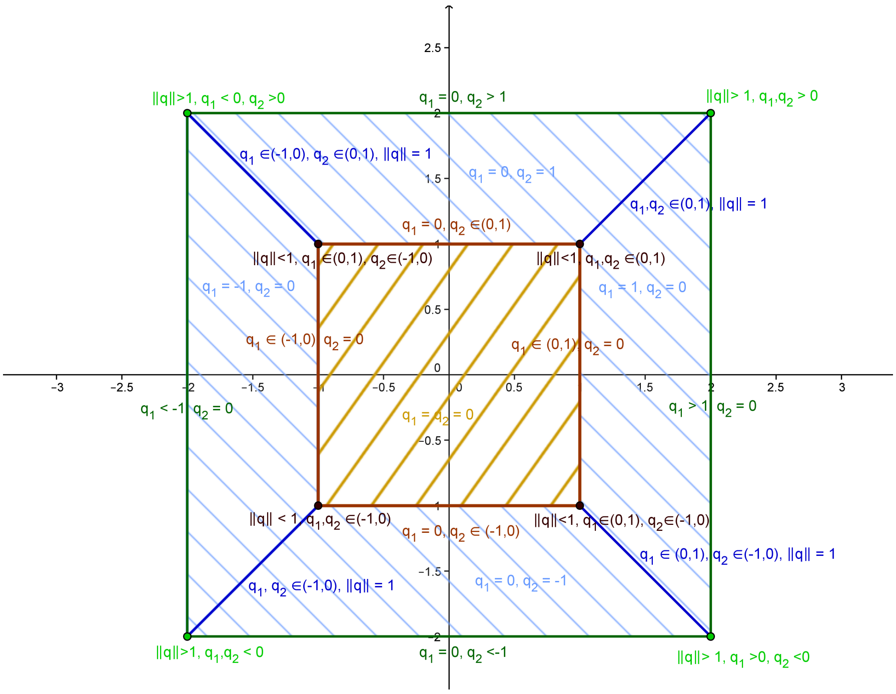


Figure 6.1.: The subdifferential ∂g^*.

Thus, the Gateaux derivative of our functional reads

$$D\,\mathcal{F}(u) = \alpha A^{-*}(A^{-1}u - z).$$

It holds additionally

$$\begin{aligned}
&\frac{|\,\mathcal{F}(u+\phi) - \mathcal{F}(u) - \langle D\,\mathcal{F}(u), \phi\rangle\,|}{\|\phi\|_{L^2}} \\
&= \frac{|\alpha\left\langle A^{-1}u - z, A^{-1}\phi\right\rangle + \frac{\alpha}{2}\left\langle A^{-1}\phi, A^{-1}\phi\right\rangle - \alpha\left\langle A^{-1}u - z, A^{-1}\phi\right\rangle|}{\|\phi\|_{L^2}} \\
&= \frac{\alpha}{2}\left|\left\langle A^{-*}(A^{-1}\phi), \frac{\phi}{\|\phi\|_{L^2}}\right\rangle\right| \\
&\overset{Cauchy-Schwarz}{\leq} \frac{1}{2}\left\|A^{-*}(A^{-1}\phi)\right\|_{L^2} \xrightarrow{\|\phi\|_{L^2}\to 0} 0
\end{aligned}$$

By section 2.2 of Clason's lecture notes [5] $\mathcal{F}$ is Fréchet differentiable and the Fréchet derivative is

$$\mathcal{F}'(u) = \alpha A^{-*}(A^{-1}u - z). \tag{6.9}$$

6.2. Existence and Stability of the Solution

In this section we are going to prove that the formal optimality system (6.5) has a unique solution $(\bar{u}, \bar{p})$. First we introduce sets Q_i's, S, P_{ij}'s and L_{ij}'s to give the subdifferential a more compact expression.

$$Q_i = \begin{cases} \{q \in \mathbb{R}^2 \,|q_1, q_2 > 0 \wedge q_2 + q_2 < 1\} & \text{i} = 1, \\ \{q \in \mathbb{R}^2 \,|q_1 > 0, q_2 < 0 \wedge q_1 - q_2 < 1\} & \text{i} = 2, \\ \{q \in \mathbb{R}^2 \,|q_1 < 0, q_2 > 0 \wedge -q_1 + q_2 < 1\} & \text{i} = 3, \\ \{q \in \mathbb{R}^2 \,|q_1, q_2 < 0 \wedge -q_1 - q_2 < 1\} & \text{i} = 4, \\ \{q \in \mathbb{R}^2 \,|q_1, q_2 > 0 \wedge q_1 + q_2 > 1\} & \text{i} = 5, \\ \{q \in \mathbb{R}^2 \,|q_1 > 0, q_2 < 0 \wedge q_1 - q_2 > 1\} & \text{i} = 6, \\ \{q \in \mathbb{R}^2 \,|q_1 < 0, q_2 > 0 \wedge -q_1 + q_2 > 1\} & \text{i} = 7, \\ \{q \in \mathbb{R}^2 \,|q_1, q_2 < 0 \wedge -q_1 - q_2 > 1\} & \text{i} = 8, \end{cases} \tag{6.10}$$

and

$$S = \bigcap_{i=1}^{4} \overline{Q_i}. \tag{6.11}$$

The P_{ij}'s are defined as follows:

$$P_{ij} = \begin{cases} \bigcap_{k=i}^{j} \left(\overline{Q_k} \cap \overline{Q_{k+4}}\right) & \text{i=1,3, j = i+1}, \\ \bigcap_{k=0}^{3} \overline{Q_{i+2k}} & \text{i=1,2, j = i+2}. \end{cases} \tag{6.12}$$

From here we agree that $P_{ij} = P_{ji}$ since each combination of indices exists only once. This agreement makes the definition of the L_{ij}'s easier:

$$L_{ij} = \begin{cases} \left(\overline{Q_i} \cap \overline{Q_j}\right) \setminus (S \cup P_{ij}) & \text{i = 1,3, j = i + 1 and i = 1,2, j= i+2}, \\ \left(\overline{Q_i} \cap \overline{Q_j}\right) \setminus P_{i-4j-4} & \text{i = 5,7, j = i + 1 and i = 5,6, j= i+2}, \\ \left(\overline{Q_i} \cap \overline{Q_j}\right) \setminus (P_{ij_1} \cup P_{ij_2}) & \text{i = 1,...,4, j = i + 4, j}_1 = 2, \text{j}_2 = 3 \text{ for i} \\ & = 1,4 \text{ and j}_1 = 1, \text{j}_2 = 4 \text{ for i = 2,3}. \end{cases} \tag{6.13}$$

In figure 6.2 the different sets are illustrated.

We can rewrite (6.5) using the Fréchet derivative of $\mathcal{F}$ as

$$\begin{aligned} -\bar{p} &= \alpha A^{-*}(A^{-1}u - z) \\ \bar{u} &\in \partial\, \mathcal{G}^*(\bar{p}), \end{aligned} \tag{6.14}$$

where the subdifferential $\partial\, \mathcal{G}^*$ is given by theorem 4.6 pointwise almost everywhere as

$$\partial\, \mathcal{G}^*(p)(x) = \partial g^*(p(x)) = \begin{cases} u_i & \text{if } p(x) \in Q_i, 1 \le i \le 8, \\ \overline{co}\,\{u_i, u_j\} & \text{if } p(x) \in L_{ij}, 1 \le \text{i} \le 7, \\ & \qquad 2 \le \text{j} \le 8, \\ \overline{co}\,\{u_1, u_2, u_3, u_4\} & \text{if } p(x) = S, \\ \overline{co}\,\{u_i, u_j, u_{i+4}, u_{j+4}\} & \text{if } p(x) = P_{ij}, 1 \le \text{i} \le 3, \\ & \qquad 2 \le \text{j} \le 4. \end{cases} \tag{6.15}$$

Theorem 6.2. *There exists a unique solution* $(\bar{u}, \bar{p})$ *to* (6.14) *in* $L^2(\Omega, \mathbb{R}^2) \times L^2(\Omega, \mathbb{R}^2)$.

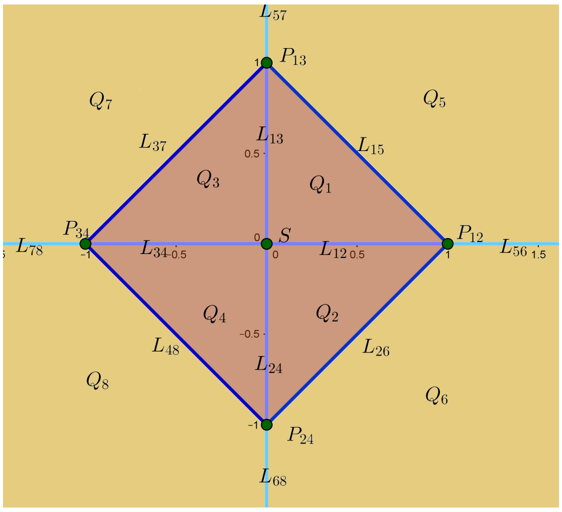


Figure 6.2.: Domains for the definition of ∂g^*.

Proof. We can eliminate $\bar{p}$ from (6.14) by inserting the first equation into the second:

$$\bar{u} \in \partial \mathcal{G}^*(\alpha A^{-*}(z - A^{-1}\bar{u})).$$

Also we aim to eliminate $\bar{u}$ from this equation. So we introduce $\bar{y} = z - A^{-1}\bar{u}$ and get the reduced optimality condition

$$z \in \bar{y} + A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\bar{y}). \tag{6.16}$$

We want to show that $\mathrm{Id} + A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\cdot)$ is maximally monotone, since we can deduce from this that our operator is surjective.
We have proven in theorem 4.2 that $\mathcal{G}^*$ is lower semi-continuous and convex. Therefore, theorem 5.3 yields that $\partial \mathcal{G}^*$ is a maximally monotone operator. Now we have to verify that $A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\cdot)$ fulfils the hypotheses of theorem 5.9 so that we can conclude that $A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\cdot)$ is maximally monotone.
A and therefore A^* are isomorphisms. It follows that A^{-*} is a bounded operator. Since we have shown in section 2.1 that the linearised elasticity equation has a solution for all $u \in (H^1_\Gamma)^*$ the domain of A^{-*} is $(H^1_\Gamma)^*$. It holds that $L^2(\Omega, \mathbb{R}^2)$ is compactly embedded in $(H^1_\Gamma)^*$. It remains to prove that $\mathrm{cone}(\mathrm{ran}\, A^{-*} - \mathrm{dom}\, \partial \mathcal{G}^*) = \overline{\mathrm{span}}(\mathrm{ran}\, A^{-*} - \mathrm{dom}\, \partial \mathcal{G}^*)$.

$$\mathrm{cone}(\mathrm{ran}\, A^{-*} - \mathrm{dom}\, \partial \mathcal{G}^*) = \bigcup_{\lambda > 0} \lambda(\mathrm{ran}\, A^{-*} - \mathrm{dom}\, \partial \mathcal{G}^*) = L^2(\Omega, \mathbb{R}^2),$$

since $\mathrm{ran}\, A^{-*} = H^1_\Gamma \hookrightarrow L^2(\Omega, \mathbb{R}^2)$ because of the fact that A^{-*} is an isomorphism and $\mathrm{dom}\, \partial \mathcal{G}^* = L^2(\Omega, \mathbb{R}^2)$. Clearly $\overline{\mathrm{span}}(\mathrm{ran}\, A^{-*} - \mathrm{dom}\, \partial \mathcal{G}^*) = L^2(\Omega, \mathbb{R}^2)$ and thus, $A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\cdot)$ is maximally monotone by theorem 5.9. Obviously Id is maximally monotone and $\mathrm{dom}\, \mathrm{Id} = L^2(\Omega, \mathbb{R}^2)$. Because of this it holds that $\mathrm{dom}(A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\cdot)) \cap \mathrm{int}(\mathrm{dom}\, \mathrm{Id}) \neq \emptyset$ and therefore we get from theorem 5.8 that $\mathrm{Id} + A^{-1}\partial \mathcal{G}^*(\alpha A^{-*}\cdot)$ is maximally monotone.

As mentioned above $A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\cdot)$ is maximally monotone and it follows by definition 5.1 that

$$\left\langle v, \partial A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}v) - A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}0)\right\rangle_{L^2} \geq 0$$

for all $v \in L^2(\Omega,\mathbb{R}^2)$.

$$\begin{aligned}\left\langle y + A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}y), y\right\rangle_{L^2} &= \|y\|_{L^2}^2 + \left\langle A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}y), y\right\rangle_{L^2} \\ &\geq \|y\|_{L^2}^2 + \left\langle A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}0), y\right\rangle_{L^2} \\ &\overset{\|y\|_{L^2}\to\infty}{\to} \infty\end{aligned}$$

Thus, $\mathrm{Id} + \partial A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\cdot)$ is coercive and with theorem 5.6 the operator is surjectiv. So for all $z \in L^2(\Omega,\mathbb{R}^2)$ there exists a $\bar{y} \in L^2(\Omega,\mathbb{R}^2)$ such that (6.16) is satisfied. We can reformulate (6.16) for the solution $\bar{y}$ as follows:

$$A(z-\bar{y}) \in \partial\mathcal{G}^*(\alpha A^{-*}\bar{y}).$$

So the solution $(\bar{u},\bar{p})$ can be defined by

$$\begin{aligned}\bar{u} &:= A(z-\bar{y}) \in \partial\mathcal{G}^*(\bar{p}), \\ \bar{p} &:= \alpha A^{-*}\bar{y} = \alpha A^{-*}(z - A^{-1}\bar{u}).\end{aligned}$$

We can deduce from this that $\bar{p}$ is even an element of H^1_Γ.
Our last step is to prove the uniqueness of the solution. Assume that $\bar{y}_1, \bar{y}_2 \in L^2(\Omega,\mathbb{R}^2)$ are two solutions of (6.16). Then

$$z \in \bar{y}_1 + A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_1) \text{ and } z \in \bar{y}_2 + A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_2).$$

Subtracting these equations and making use of the fact that $\mathrm{Id} + A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\cdot)$ is maximally monotone yields

$$\begin{aligned}&0 \in \bar{y}_1 + A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_1) - \bar{y}_2 - A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_2) \\ \Rightarrow\ &0 \in \left\langle \bar{y}_1 + A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_1) - \bar{y}_2 - A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_2), \bar{y}_1 - \bar{y}_2\right\rangle_{L^2} \\ &\quad = \|\bar{y}_1 - \bar{y}_2\|_{L^2}^2 + \underbrace{\left\langle A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_1) - A^{-1}\partial\mathcal{G}^*(\alpha A^{-*}\bar{y}_2), \bar{y}_1 - \bar{y}_2\right\rangle_{L^2}}_{\geq 0 \text{ by monotonicity of } \partial\mathcal{G}^*} \\ &\quad \geq \|\bar{y}_1 - \bar{y}_2\|_{L^2}^2\,,\end{aligned}$$

and therefore $\bar{y}_1 = \bar{y}_2$. Now we have to verify that it follows that the pair $(\bar{u},\bar{p})$ is unique. It holds $\bar{u} = A(z-\bar{y})$ and $\bar{p} = A^{-*}\bar{y}$, A is an isomorphism and $\bar{y}$ is unique. Hence, the pair $(\bar{u},\bar{p})$ is also unique. □

Later we need the solution of $\bar{u}$ in order to remember that for a pair $(\bar{u},\bar{p})$ that satisfies $\bar{u} \in \partial\mathcal{G}^*(\bar{p})$ we have pointwise for almost all x

$$\bar{u}(x) \in \begin{cases} u_i & \text{if } p(x) \in Q_i, 1 \leq i \leq 8, \\ \overline{co}\,\{u_i, u_j\} & \text{if } p(x) \in L_{ij}, 1 \leq i \leq 7, 2 \leq j \leq 8, \\ \overline{co}\,\{u_1,u_2,u_3,u_4\} & \text{if } p(x) = S, \\ \overline{co}\,\{u_i,u_j,u_{i+4},u_{j+4}\} & \text{if } p(x) = P_{ij}, 1 \leq i \leq 3, 2 \leq j \leq 4. \end{cases} \tag{6.17}$$

We are going to show that our solution triple $(\bar{u}, \bar{y}, \bar{p})$ depends continuously on our target function $z \in L^2(\Omega, \mathbb{R}^2)$.

Theorem 6.3. *Let $z \in L^2(\Omega, \mathbb{R}^2)$ and $(\bar{u}_z, \bar{y}_z, \bar{p}_z)$ be the corresponding solution to (6.14). The following three statements hold:*
a) There exists a constant $K > 0$ such that

$$\|\bar{y}_{z_1} - \bar{y}_{z_2}\|_{L^2} + \|\bar{p}_{z_1} - \bar{p}_{z_2}\|_{H^1} \leq K \|z_1 - z_2\|_{L^2}$$

for all $z_1, z_2 \in L^2(\Omega, \mathbb{R}^2)$.

b) For a sequence $(z_n)_n \in L^2(\Omega, \mathbb{R}^2)$ with $z_n \to z$ it holds $(u_{z_n}, y_{z_n}) \rightharpoonup (u_z, y_z)$ weakly in $(H^1_\Gamma)^* \times H^1_\Gamma$ and $p_{z_n} \to p_z$ strongly in H^1_Γ.

c) For $z \in L^2(\Omega, \mathbb{R}^2)$, let W be a compact subset of $\bigcup_{i=1}^8 Q_i$. A is an isomorphism from $H^2(\Omega) \cap H^1_\Gamma(\Omega)$ to $L^2(\Omega, \mathbb{R}^2)$ and $\Omega \subset \mathbb{R}^2$ thus we can find a neighbourhood $U_z \subset L^2(\Omega, \mathbb{R}^2)$ and a constant K_W such that

$$\|u_{\tilde{z}} - u_z\|_{H^2(\Omega_W)} \leq K_W \|\tilde{z} - z\|_{L^2} \, \forall \tilde{z} \in U(z),$$

where $\Omega_W = \{x | \bar{p}(x) \in W\}$.

Proof. a) For given $z_1, z_2 \in L^2(\Omega, \mathbb{R}^2)$ let y_1, y_2 be the solution to (6.16) respectively. Then it holds

$$z_1 - z_2 \in y_1 - y_2 + A^{-1}(\partial \mathcal{G}^*(\alpha A^{-*} y_1)) - A^{-1}(\partial \mathcal{G}^*(\alpha A^{-*} y_2)).$$

From this we get

$$\begin{aligned}
&\langle z_1 - z_2, y_1 - y_2 \rangle_{L^2} \\
=& \left\langle y_1 - y_2 + A^{-1}(\partial \mathcal{G}^*(\alpha A^{-*} y_1)) - A^{-1}(\partial \mathcal{G}^*(\alpha A^{-*} y_2)), y_1 - y_2 \right\rangle_{L^2} \\
=& \|y_1 - y_2\|^2_{L^2} \\
&+ \underbrace{\left\langle (\partial \mathcal{G}^*(\alpha A^{-*} y_1)) - (\partial \mathcal{G}^*(\alpha A^{-*} y_2)), A^{-*} y_1 - A^{-*} y_2 \right\rangle_{L^2}}_{\geq 0, \text{ since } \partial \mathcal{G}^* \text{ is monotone}} \\
\geq& \|y_1 - y_2\|^2_{L^2}.
\end{aligned}$$

So we have

$$\|y_1 - y_2\|^2_{L^2} \leq \langle z_1 - z_2, y_1 - y_2 \rangle_{L^2} \leq \|z_1 - z_2\|_{L^2} \|y_1 - y_2\|_{L^2}.$$

The last estimation holds by Cauchy-Schwarz's inequality.

$$\Rightarrow \|y_1 - y_2\|_{L^2} \leq \|z_1 - z_2\|_{L^2}$$

We can obtain from (6.14) and from the definition of y that $y = A^*p$.

$$\begin{aligned}
\|z_1 - z_2\|_{L^2} &\geq \frac{1}{2}\|y_1 - y_2\|_{L^2} + \frac{1}{2}\|y_1 - y_2\|_{L^2} \\
&= \frac{1}{2}\|y_1 - y_2\|_{L^2} + \frac{1}{2}\|A^*(p_1 - p_2)\|_{L^2} \\
&\geq \frac{1}{2}\|y_1 - y_2\|_{L^2} + \frac{1}{2}\|A^*\|\,\|p_1 - p_2\|_{H^1_\Gamma} \\
&\geq \underbrace{\min\left(\frac{1}{2}, \frac{1}{2}\|A^*\|\right)}_{:=\frac{1}{K}}\left(\|y_1 - y_2\|_{L^2} + \|p_1 - p_2\|_{H^1_\Gamma}\right)
\end{aligned}$$

The claim of a) follows.

b) For this part of the proof we set $(\bar{u}_{z_n}, \bar{y}_{z_n}, \bar{p}_{z_n}) = (u_n, y_n, p_n)$ to simplify the notation. The first part of this theorem yields that $(p_n)_{n\in\mathbb{N}}$ is bounded in H^1_Γ. By the Rellich-Kondrachov compactness theorem [see [9], Thm. 5.1] H^1_Γ is compactly embedded in $L^2(\Omega, \mathbb{R}^2)$. So $(p_n)_{n\in\mathbb{N}}$ is precompact in $L^2(\Omega, \mathbb{R}^2)$ and hence there exists a subsequence p_{n_k} such that $p_{n_k} \to \bar{p} := \bar{p}_z$ almost everywhere in Ω. Since $\|u_n\|_\infty \leq 2$ the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega, \mathbb{R}^2)$ and because of this there again exists a subsequence (u_{n_k}) such that $u_{n_k} \rightharpoonup \tilde{u}$ weakly for some $\tilde{u} \in L^2(\Omega, \mathbb{R}^2)$. The Q_i's are open and as shown above $p_n \to p$ almost everywhere, so it follows from (6.17) that

$$u_n \to \bar{u} \text{ almost everywhere on } \left\{x \in \Omega | p(x) \in \bigcup_{i=1}^{8} Q_i\right\}.$$

Since $u_n \rightharpoonup \tilde{u}$ and u_n is constant on Q_i and these sets are open, we conclude that $\tilde{u} = \bar{u}$.
It remains to consider the sets S, P_{ij}, L_{ij} as defined in (6.11), (6.12) and (6.13). As shown above $u_n \rightharpoonup \tilde{u}$ weakly for some $\tilde{u} \in L^2(\Omega, \mathbb{R}^2)$ independent of $p(x)$. Using Mazur's theorem [see Yosida [15], Theorem V.1.2 for a proof] we obtain that for every $\epsilon > 0$ there exist coefficients $\gamma_j^n \geq 0$ with $\sum_{j=1}^{l_n} \gamma_j^n = 1$ and indices $n_j \in n, n+1, ...$ such that

$$\tilde{u}_n := \sum_{j=1}^{l_n} \gamma_j^n u_{n_k} \to \tilde{u} \text{ strongly.}$$

Hence we can take another subsequence $\tilde{u}_{n_k}$ of $\tilde{u}_n$ such that $\tilde{u}_{n_k} \to \tilde{u}$ almost everywhere. Now our aim is to verify that $\tilde{u}$ is in $\partial\mathcal{G}^*(p)$ on the union of the L_{ij}'s, P_{ij}'s and S. All $\tilde{u}_n$'s are in the closed convex hull of some u_i's by (6.17). Therefore, $\tilde{u}$ must also be an element of this closed convex hull. This yields that $\tilde{u} \in \partial\mathcal{G}^*(p)$ for almost all $x \in \Omega$. Hence, $(\tilde{u}, y, p)$ satisfies (6.14). Since we have proven uniqueness of the solution of (6.14) in the previous theorem this implies that $\tilde{u} = u$.

c) Let the sets Q_i belong to the solution $(\bar{u}_z, \bar{y}_z, \bar{p}_z)$ of (6.14). Since we can restrict A to an isomorphism form $H^2(\Omega) \cap H^1_\Gamma(\Omega)$ to $L^2(\Omega, \mathbb{R}^2)$ and our Ω is a subset of $\mathbb{R}^2$

we get from part b) of this theorem that $p_{\tilde{z}} \to p_z$ as $\tilde{z} \to z$. Because of this there exists a neighbourhood U_z of z such that $\{x \in \Omega | \bar{p}_z \in W \cap Q_i\} \subset \{x \in \Omega | p_{\tilde{z}} \in Q_i\}$ for $1 \leq i \leq 8$. It follows that

$$\bar{u}_z = \bar{u}_{\tilde{z}} = u_i \text{ on } \{x \in \Omega | \bar{p}_z \in W \cap Q_i\}.$$

Therefore, we know that the distance between $\bar{u}_z$ and $\bar{u}_{\tilde{z}}$ is zero for all $x \in \Omega_W$. So the claim follows.

□

6.3. Structure of the Solution

In this section we want to deal with the structure of our solution $\bar{u}$. We recall that our aim was to get a multi-bang control, i.e. $\bar{u} = u_i$ almost everywhere. So we want to find conditions under which we can fulfil this.
We observe that we can divide Ω in 9 different sets:

$$\begin{aligned}\Omega &= \bigcup_{i=1}^{8} \{x \in \Omega | \bar{u}(x) = u_i\} \cup \{x \in \Omega | \bar{u}(x) \notin \{u_1, ..., u_8\}\} \\ &=: \bigcup_{i=1}^{8} \mathcal{A}_i \cup \mathcal{S}.\end{aligned}$$

$\mathcal{A} := \bigcup_{i=1}^{8} \mathcal{A}_i$ is called the *multi-bang arc* and $\mathcal{S}$ the *singular arc.*

Clason and Kunisch [7, Prop. 2.3] were able to characterize the structure of the solution for scalar functions as follows:

$$\Omega = \bigcup_{i=1}^{8} \{x \in \Omega | \bar{u}(x) = u_i\} \cup \{x \in \Omega | \bar{y}(x) = z(x)\},$$

under the assumption that A is a second order elliptic partial differential operator of the form

$$Ay = -\sum_{i,j=1}^{n} \partial_{x_i}(a_{i,j}\partial_{x_j} y) + \sum_{i=1}^{n} \partial_{x_i}(b_i y)$$

with $a_{i,j} \in W^{1,\infty}(\Omega), b_i \in L^{\infty}(\Omega)$ and that A fulfils

$$A^{-*}(L^2(\Omega)) \subset W^{2,1}(\Omega).$$

This is not possible for vector-valued functions as the following example shows.
The distance between the state y and the target z is given by the first equation of the optimality system (6.14) as

$$\frac{1}{\alpha} A^* p = y - z. \tag{6.18}$$

Our aim is to construct an example were $y - z \neq 0$ on the singular arc. We can construct a triple (u, y, p) that fulfils the optimality system (6.14) and from this we can deduce a target z which does not fulfil the condition of Calson and Kunisch.
We take $\Omega = [0,1]^2$ and choose $A = -\Delta$ with homogeneous Dirichlet boundary conditions on $\Gamma_0 = \{0\} \times [0,1] \cup \{1\} \times [0,1]$ and inhomogeneous Dirichlet conditions on the other part of $\partial\Omega$. We know that the Laplace operator is self-adjoint. Therefore, we need according to (6.18) a p with $-\Delta p \neq 0$. Our choice is

$$p(x,y) = \begin{pmatrix} -x^2 + x \\ 0 \end{pmatrix}.$$

This p has homogeneous boundary conditions on Γ_0 and $p(x,0) = p(x,1) = -x^2 + x$ on $\partial\Omega \setminus \Gamma_0$.
We observe that $p(x) \in L_{12} = \left\{ q \in \mathbb{R}^2 \,| q \in (0,1) \times \{0\} \right\}$ for all $x \in \Omega$ and therefore Ω is the singular arc.
From the definition of L_{12} (6.8) we get that u must be in the following set if we want to fulfil the second equation the optimality system (6.14):

$$u(x,y) \in \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \lambda \in [0,1].$$

We choose $u(x,y) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. We calculate the solution of the PDE numerically. A linear finite element approach is used. Ω is uniformly triangulated with $N_h = 256 \times 256$ nodes. The code for the mass and the stiffness matrix can be found in appendix A.
Furthermore, we choose $\alpha = 100$.
Figure 6.3 shows $\|z - y\|_2$ at each point of the triangulation. As one can observe in figure 6.3 that $y(x) \neq z(x)$ for all

$$p(x) \in L_{12} = \{x \in \Omega | p(x) \in (0,1) \times 0\} = \Omega.$$

So the characterisation of the singular arc given by Calson and Kunisch is not applicable for vector-valued functions.

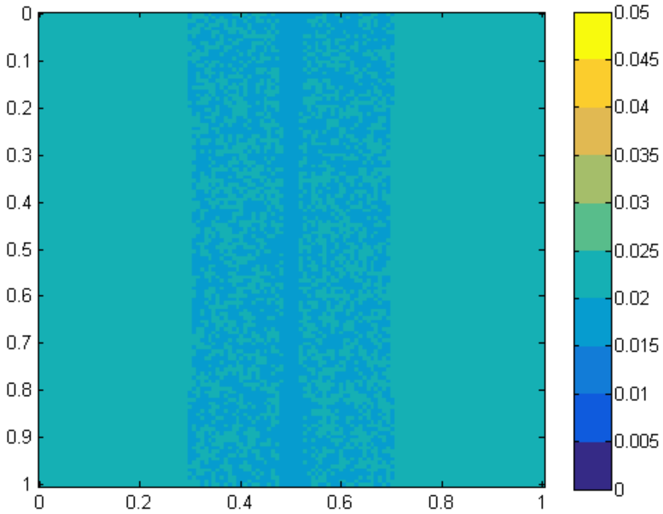


Figure 6.3.: Norm of the residual of the state y and the target z

7. Numerics

We want to compute a solution for our primal-dual optimality system (6.5). This can be written with the introduction of the optimal state $\bar{y} = A^{-1}u$ as follows

$$\begin{cases} A\bar{y} = \bar{u}, \\ \dfrac{1}{\alpha} A^* \bar{p} = z - \bar{y}, \\ \bar{u} \in \partial \mathcal{G}^*(\bar{p}). \end{cases} \tag{7.1}$$

The subdifferential is set-valued and this causes problems for the numerical computation of a solution. Consequently, we are going to regularize the subdifferential. We use the Moreau-Yosida regularization since this regularization is single-valued as we will see later.

7.1. Moreau-Yosida Regularization

At first we want to give the definition of the infimal convolution. From this we can derive the definition of the Moreau envelope. We need this definition to define the Moreau-Yosida regularization. The following definitions are taken form Bauschke and Combettes [1].

Definition 7.1 (Infimal convolution)**.** *Let $\mathcal{H}$ be a real Hilbert space, $f, g\colon \mathcal{H} \to (-\infty, +\infty]$. The infimal convolution of f and g is given by*

$$f \Box g\colon \mathcal{H} \to \overline{\mathbb{R}}, \quad x \mapsto \inf_{y \in \mathcal{H}} (f(y) + g(x-y)). \tag{7.2}$$

$f \Box g$ is exact at $x \in \mathcal{H}$ if the infimum is attained, i.e. $(f \Box g)(x) = \min_{y \in \mathcal{H}} f(y) + g(x-y)$.

The Moreau envelope of a function f is the infimal convolution of f with a function g that depends on a parameter γ.

Definition 7.2 (Moreau envelope)**.** *Again let $\mathcal{H}$ be a real Hilbert space, f a function from $\mathcal{H}$ to $(-\infty, +\infty]$ and $\gamma \in (0, \infty)$.*

$$f_\gamma = f \Box \left(\frac{1}{2\gamma} \|\cdot\|^2 \right), \ \textit{i.e.}\ f_\gamma(x) = \inf_{y \in \mathcal{H}} (f(y) + \frac{1}{2\gamma} \|x - y\|^2) \tag{7.3}$$

is the Moreau envelope of f of parameter γ.

We want to mention several properties of the Moreau envelope.

Theorem 7.3. *Let f be a proper convex function from a Hilbert space $\mathcal{H}$ to $(-\infty, +\infty]$. The Moreau envelope of f is convex, real-valued and exact.*

Proof. By assumption f is convex and $\frac{1}{2\gamma}\|\cdot\|^2$ is supercoercive, hence convex. Thus, we get that the Moreau envelope is convex [for a proof that the infimal convolution of two convex functions is convex see Bauschke and Combettes [1] Prop. 12.11]. The supercoercivity yields also that $f\Box\left(\frac{1}{2\gamma}\|\cdot\|^2\right)$ is exact [see again [1] Prop. 12.14 for a more detailed proof]. The Moreau envelope is real-valued because of the fact that f_γ is bounded from above on every ball in $\mathcal{H}$ [see [1] Prop. 12.9]. □

Our next step is to define the proximal mapping of f and list the interplay between the proximal mapping and the Moreau envelope.

Definition 7.4 (Proximal mapping)**.** *Let f be a proper lower semi-continuous and convex function from $\mathcal{H}$ to $(-\infty, +\infty]$. $\operatorname{prox}_f(x)$ is the unique point in $\mathcal{H}$ for which the minimum in the definition of f_1 is attained, i.e.*

$$f_1(x) = \min_{y\in\mathcal{H}}\left(f(y) + \frac{1}{2}\|x-y\|^2\right) := f(\operatorname{prox}_f(x)) + \frac{1}{2}\|prox_f(x) - x\|^2 .$$

$\operatorname{prox}_f\colon \mathcal{H} \to \mathcal{H}$ *is called proximal mapping.*

Theorem 7.5. *We can write the Moreau envelope with the definition of the proximal mapping in the following way:*

$$f_\gamma(x) = f(prox_{\gamma f}(x)) + \frac{1}{2\gamma}\left\|\operatorname{prox}_{\gamma f}(x) - x\right\|^2 . \tag{7.4}$$

Proof. The first step is to verify that $(\gamma f)_1 = \gamma(f_\gamma)$.

$$(\gamma f)_1 = \min_{y\in\mathcal{H}}(\gamma f(y) + \frac{1}{2}\|x-y\|^2) = \gamma(\min_{y\in\mathcal{H}} f(y) + \frac{1}{2\gamma}\|x-y\|^2) = \gamma f_\gamma$$

From this we get in a second step

$$f_\gamma = \frac{1}{\gamma}(\gamma f)_1 = \frac{1}{\gamma}(\gamma f(\operatorname{prox}_{\gamma f}(x)) + \frac{1}{2}\left\|x - \operatorname{prox}_{\gamma f}(x)\right\|^2 .$$

As a result, we get our claim. □

We want to mention that we can also write the proximal mapping in the following way

$$\operatorname{prox}_{\gamma f}(x) = \arg\min_{y\in\mathcal{H}} f(y) + \frac{1}{2\gamma}\|x-y\|^2 . \tag{7.5}$$

Since we now know about the Moreau envelope of a function we want to define the Yosida approximation of a set-valued operator.

Definition 7.6 (Yosida approximation). *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator on a Hilbert space $\mathcal{H}$ and let $\gamma \in (0, \infty)$. The resolvent of A is given by*

$$J_A = (\mathrm{Id} + A)^{-1}. \tag{7.6}$$

$$A_\gamma = \frac{1}{\gamma}(\mathrm{Id} - J_{\gamma A}) \tag{7.7}$$

is called Yosida approximation of A of index γ.

Our next theorem shows the connection between the resolvent of ∂f and the proximal mapping.

Theorem 7.7. *Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to (-\infty, +\infty]$ be proper and convex. Then proximal mapping of γf equals the resolvent of $\gamma \partial f$, i.e.*

$$\mathrm{prox}_{\gamma f} = J_{\gamma \partial f} = (\mathrm{Id} + \gamma \partial f)^{-1}. \tag{7.8}$$

Proof. Let $y \in \mathcal{H}$ be an arbitrary point and set $p = \mathrm{prox}_{\gamma f}(x)$. For all $\alpha \in (0, 1)$ we define $p_\alpha = \alpha y + (1 - \alpha)p$. From the definition of the proximal mapping we get

$$\begin{aligned}
\gamma f(p) &\leq \gamma f(p_\alpha) + \frac{1}{2} \|x - p_\alpha\|^2 - \frac{1}{2} \|x - p\|^2 \\
&\leq \alpha(\gamma f)(y) + (1 - \alpha)(\gamma f)(p) - \alpha \langle x, y\rangle + \alpha \langle x, p\rangle + \frac{\alpha^2}{2} \langle y, y\rangle \\
&\qquad + \alpha \langle y, p\rangle - \alpha^2 \langle y, p\rangle - \alpha \langle p, p\rangle + \frac{\alpha^2}{2} \langle p, p\rangle \\
&= \alpha(\gamma f)(y) + (1 - \alpha)(\gamma f)(p) - \alpha \langle x - p, y - p\rangle + \frac{\alpha^2}{2} \langle y - p, y - p\rangle .
\end{aligned}$$

The second inequality yields because of the convexity of f. We obtain from this inequality through dividing by α

$$\langle y - p, x - p\rangle + \gamma f(p) \leq \gamma f(y) + \frac{\alpha}{2} \|y - p\|^2 .$$

For $\alpha \to 0$ we get $\langle y - p, x - p\rangle + \gamma f(p) \leq \gamma f(y)$ and since $y \in \mathcal{H}$ was arbitrary we achieve with the definition of the subdifferential 4.3 that

$$x - p \in \gamma \partial f(p) \Leftrightarrow p \in (\mathrm{Id} + \gamma \partial f)^{-1}(x).$$

For the other inclusion let $p = (\mathrm{Id} + \gamma \partial f)^{-1}(x)$. This is obviously equivalent to $x - p \in \gamma \partial f(p)$. Again with the definition of the subdifferential 4.3 it holds

$$\begin{aligned}
&\langle y - p, x - p\rangle + \gamma f(p) \leq \gamma f(y) \forall y \in \mathcal{H} \\
\Leftrightarrow &\gamma f(p) + \frac{1}{2} \|x - p\|^2 \leq \gamma f(y) + \frac{1}{2} \|x - p\|^2 + \langle y - p, p - y\rangle + \frac{1}{2} \|p - y\|^2 \forall y \in \mathcal{H} \\
\Leftrightarrow &\gamma f(p) + \frac{1}{2} \|x - p\|^2 \leq \gamma f(y) + \frac{1}{2} \|x - y\|^2 \forall y \in \mathcal{H} .
\end{aligned}$$

This yields that the minimum is attained at p and therefore we have $p = \mathrm{prox}_{\gamma f}(x)$. Finally we have $\mathrm{prox}_{\gamma f} = (\mathrm{Id} + \gamma \partial f)^{-1}$. □

Now we are able to define the Moreau-Yosida regularization.

Definition 7.8 (Moreau-Yosida regularization). *Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to (-\infty, +\infty)$ be a proper convex function. Then*

$$(\partial f)_\gamma(p) = \frac{1}{\gamma}\left(p - \operatorname{prox}_{\gamma f}(p)\right) \tag{7.9}$$

is the Moreau-Yosida regularization of ∂f.

In the following theorem we want to list several properties of $\operatorname{prox}_{\gamma f}$ and $(\partial f)_\gamma$. This theorem is recited from Clason, Ito and Kunisch [6, Prop. 2.3].

Theorem 7.9. *Let f be a proper convex function from a Hilbert space $\mathcal{H}$ to $(-\infty, +\infty)$. Then*
a) $(\partial f)_\gamma = (f_\gamma)'$,

b) $(\partial f)_\gamma$ *is single-valued, maximally monotone and Lipschitz continuous with Lipschitz constant* $\frac{1}{\gamma}$,

c) $\|(\partial f)_\gamma(x)\| \leq \inf_{q\in\partial f(v)} \|q\|_H$ *for all* $x \in \mathcal{H}$,

d) $f(\operatorname{prox}_{\gamma f}(x)) \leq f_\gamma(x) \leq f(x)$ *for all* $\gamma > 0$ *and* $v \in \mathcal{H}$.

Proof. a) Let $x, y \in \mathcal{H}$ with $x \neq y$. Then set $p = \operatorname{prox}_{\gamma f}(x)$ and $q = \operatorname{prox}_{\gamma f}(y)$. By theorem 7.5 we get

$$\begin{aligned}
f_\gamma(y) - f_\gamma(x) = {} & \underbrace{f(q) - f(p)}_{\geq \frac{1}{\gamma}\langle q-p, x-p\rangle \text{ by Thm. 7.7}} + \frac{1}{2\gamma}\left(\|y-q\|^2 - \|x-p\|^2\right) \\
\geq {} & \frac{1}{2\gamma}\left(\|y-q\|^2 - \|x-p\|^2 + 2\langle q-p, x-p\rangle\right) \\
= {} & \frac{1}{2\gamma}\Big(\|y-q\|^2 - 2\langle y-q, x-p\rangle + \|x-p\|^2 \\
& - 2\|x-p\|^2 + 2\langle y-q, x-p\rangle + 2\langle q-p, x-p\rangle\Big) \\
= {} & \frac{1}{2\gamma}\left(\|y-q-x+p\|^2 + 2\langle y-x, y-p\rangle\right) \\
\geq {} & \frac{1}{\gamma}\langle y-x, y-p\rangle .
\end{aligned}$$

Analogous we get $f_\gamma(x) - f_\gamma(y) \geq \frac{1}{\gamma}\langle x-y, y-q\rangle$. Adding these two inequalities yields

$$\begin{aligned}
0 &\leq \frac{1}{\gamma}\left(\langle y-x, y-q\rangle - \langle y-x, x-p\rangle\right) \\
&= \frac{1}{\gamma}\left(\langle y-x, (y-q)-(x-p)\rangle\right).
\end{aligned}$$

An operator $T\colon D \to \mathcal{H}$, where D is a subset of a Hilbert space $\mathcal{H}$, is called firmly nonexpansive if $\forall\, x, y \in D \;\; \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$. The proximal mapping is firmly nonexpansive [see Bauschke and Combettes [1], Prop. 12.27 for a proof]. Thus, we get that

$$
\begin{aligned}
0 &\leq \frac{1}{\gamma}\left(\langle y - x, (y - q) - (x - p)\rangle\right) \\
&\leq \frac{1}{\gamma}\Big(\|y - x\|^2 - \|q - p\|^2 - \underbrace{\langle q - p, (y - q) - (x - p)\rangle}_{\geq 0,\ \mathrm{prox}_{\gamma f} \text{ firmly nonexpansive}} \Big) \\
&\leq \frac{1}{\gamma} \|y - x\|^2 .
\end{aligned}
$$

Therefore, $\lim_{y \to x} \frac{f_\gamma(y) - f_\gamma(x) - \langle y - x, \gamma^{-1}(x - p)\rangle}{\|y - x\|} = 0$ and therefore f_γ is Fréchet differentiable with derivative $\frac{1}{\gamma}(\mathrm{Id} - \mathrm{prox}_{\gamma f})$. The Fréchet derivative equals the Moreau-Yosida approximation of ∂f.

b) The Lipschitz continuity follows directly from the fact that $\mathrm{prox}_{\gamma f}$ and therefore $\mathrm{Id} - \mathrm{prox}_{\gamma f}$ are firmly nonexpansive and thus both operators are Lipschitz continuous with constant 1. Hence, $\frac{1}{\gamma}(\mathrm{Id} - \mathrm{prox}_{\gamma f})$ is Lipschitz continuous with constant $\frac{1}{\gamma}$. Since we verified in theorem 7.7 that $\mathrm{prox}_{\gamma f} = (\mathrm{Id} + \gamma \partial f)^{-1}$, $(\mathrm{Id} + \gamma\partial f)^{-1}$ and also $\mathrm{Id} - (\mathrm{Id} + \gamma \partial f)^{-1}$ are firmly nonexpansive. This yields the maximally monotonicity of those two operators. From the maximally monotonicity of $\mathrm{Id} - (\mathrm{Id} + \gamma\partial f)^{-1}$ we can deduce that $(\partial f)_\gamma$ is γ-coercive and thus maximally monotone. The Moreau-Yosida regularization is single-valued because the proximal mapping is unique.

c) Let $y = (\partial f)_\gamma(x)$ and $q \in \partial f(x)$. Then

$$
\begin{aligned}
y = (\partial f)_\gamma(x) &\Leftrightarrow y = \frac{1}{\gamma}\left(\mathrm{Id} - (\mathrm{Id} + \gamma \partial f)^{-1}\right)(x) \\
&\Leftrightarrow (\mathrm{Id} + \gamma\partial f)^{-1}(x) = x - \gamma y \\
&\Leftrightarrow y \in \partial f(x - \gamma y) \\
&\Leftrightarrow (x - \gamma y, y) \in \operatorname{graph} \partial f.
\end{aligned}
$$

We have proven in theorem 5.3 that ∂f is maximally monotone and accordingly we get

$$
0 \leq \frac{1}{\gamma}\langle x - (x - \gamma y), q - y\rangle = \langle y, q - y\rangle \leq \|q\|\,\|y\| - \|y\|^2 ,
$$

where the last inequality holds due to Cauchy-Schwarz. Thus, we have

$$
\|(\partial f)_\gamma(x)\| \leq \|q\| \;\; \forall q \in \partial f(x).
$$

Consequently, this is also true for the infimum.

d) From the definition of the proximal mapping it is clear that $f(\mathrm{prox}_{\gamma f}(x)) \leq f_\gamma(x)$. We can derive the second inequality directly from the definition of the Moreau envelope.

$\square$

Now we are able to define the regularized system. For brevity we set $H_\gamma := (\partial \mathcal{G}^*)_\gamma$.

$$\begin{aligned} Ay_\gamma &= u_\gamma \\ \frac{1}{\alpha} A^* p_\gamma &= z - y_\gamma \\ u_\gamma &= H_\gamma(p_\gamma) \end{aligned} \tag{7.10}$$

Since H_γ is maximally monotone and the other prerequisites are not changed we can deduce that (7.10) has a unique solution $(u_\gamma, y_\gamma, p_\gamma)$ with the same arguments as in the proof of theorem 6.2. We need to think about the convergence of our regularised system. The poof of this theorem is in line with Clason, Ito and Kunisch [6].

Theorem 7.10. *The sequence $\{(u_\gamma, y_\gamma, p_\gamma)\}_{\gamma>0}$ converges weakly to the solution $(\bar{u}, \bar{y}, \bar{p})$ of* (7.1) *as $\gamma \to 0$.*

Proof. Our first step is to prove that we can find a subsequence of $\{(u_\gamma, y_\gamma, p_\gamma)\}_{\gamma>0}$ that converges weakly as $\gamma \to 0$ to the solution $(\bar{u}, \bar{y}, \bar{p})$. Therefore, observe that we can bound $(\mathcal{G}_\gamma^*)^*(0)$ in the following way:

$$(\mathcal{G}_\gamma^*)^*(0) = \sup_{p \in L^2(\Omega,\mathbb{R}^2)} -\mathcal{G}_\gamma^*(p) = \inf_{p \in L^2(\Omega,\mathbb{R}^2)} \mathcal{G}_\gamma^*(p) \le \inf_{p \in L^2(\Omega,\mathbb{R}^2)} \mathcal{G}^*(p).$$

Here we have used the definition of the Fenchel conjugate 4.1 and theorem 7.9 c). Since u_γ is optimal for any given $\gamma > 0$ we get

$$\mathcal{F}(u_\gamma) \le \mathcal{F}(u_\gamma) + (\mathcal{G}_\gamma^*)^*(u_\gamma) \le \mathcal{F}(0) + (\mathcal{G}_\gamma^*)^*(0) \le \mathcal{F}(0) + \inf_{p \in L^2(\Omega,\mathbb{R}^2)} \mathcal{G}^*(p)$$

according to the facts that $(\mathcal{G}_\gamma^*)^*$ is non-negative and the estimation above. This yields that $\{\mathcal{F}(u_\gamma)\}_{\gamma>0}$ is bounded. Now we are going to show that the boundedness of $\{\mathcal{F}(u_\gamma)\}_{\gamma>0}$ implies boundedness of $\{\mathcal{F}'(u_\gamma)\}_{\gamma>0}$.

$$\left\|\mathcal{F}'(u_\gamma)\right\|_{L^2} = \left\|\alpha A^{-*}(A^{-1}u_\gamma - z)\right\|_{L^2} \le \alpha \left\|A^{-*}\right\| \left\|A^{-1}u_\gamma - z\right\|_{L^2} \le C,$$

because A^{-*} is an isomorphism and therefore its operator norm is bounded. The $\|\cdot\|_{L^2}$-norm is bounded since $\{\mathcal{F}(u_\gamma)\}_{\gamma>0}$ is bounded.
It follows from the boundedness of $\{\mathcal{F}'(u_\gamma)\}_{\gamma>0}$ that

$$\{p_\gamma\}_{\gamma>0} = \left\{\mathcal{F}'(u_\gamma)\right\}_{\gamma>0}$$

is bounded. Our next goal is to verify that $\{u_\gamma\}_\gamma$ is also bounded. We have

$$\|u_\gamma\|_{L^2} = \|H_\gamma(p_\gamma)\|_{L^2} \overset{7.9}{\le} \inf_{q \in \partial \mathcal{G}^*(p_\gamma)} \|q\|_{L^2} \le \widetilde{C}.$$

The last inequality holds true since our $\partial \mathcal{G}^*$ is bounded. $\{y_\gamma\}_{\gamma>0} = \left\{A^{-1}u_\gamma\right\}_{\gamma>0}$ is also bounded because of the fact that A^{-1} is an isomorphism.

Thus $\{(u_\gamma, y_\gamma, p_\gamma)\}_{\gamma>0}$ is bounded and hence contains a weakly convergent subsequence. [See Bauschke and Combettes [1], Lemma 2.37 for a proof of this fact.] Let $u_{\gamma_n} \rightharpoonup \hat{u}, y_{\gamma_n} \rightharpoonup \hat{y}$ and $p_{\gamma_n} \rightharpoonup \hat{p}$.
We want to prove that the graph of $\mathcal{F}'$ is weakly closed: Let (u_n, g_n) be a sequence in $\operatorname{graph} \mathcal{F}'$ that converges weakly to (u, g). We get that

$$g_n = \alpha A^{-*}(A^{-1}u_n - z).$$

As $u_n \rightharpoonup u$

$$\alpha A^{-*}(A^{-1}u_n - z) \rightharpoonup \alpha A^{-*}(A^{-1}u - z) = g.$$

As a result (u, g) is an element of $\operatorname{graph} \mathcal{F}'$ and thus the claim follows.
Accordingly to the weak closedness of $\operatorname{graph} \mathcal{F}'$ we get

$$\hat{p} = -\mathcal{F}'(\hat{u}).$$

Using that $\mathcal{F}'$ is monotone by theorem 5.3 and the relations of (7.10) yields

$$\langle H_{\gamma_1}(p_{\gamma_1}) - H_{\gamma_2}(p_{\gamma_2}), p_{\gamma_1} - p_{\gamma_2}\rangle = -\left\langle u_{\gamma_1} - u_{\gamma_2}, \mathcal{F}'(u_{\gamma_1}) - \mathcal{F}'(u_{\gamma_2})\right\rangle \leq 0$$

for any $\gamma_1, \gamma_2 > 0$. This implies that

$$\limsup_{n,m\to\infty} \langle H_{\gamma_n}(p_{\gamma_n}) - H_{\gamma_m}(p_{\gamma_m}), p_{\gamma_n} - p_{\gamma_m}\rangle \leq 0.$$

The maximal monotonicity of H_γ and the fact that $H_\gamma(p_\gamma) \in \partial\,\mathcal{G}^*(p_\gamma)$ give that we can apply theorem 5.10 to the graph of $\partial\,\mathcal{G}^*$. According to this theorem $(\hat{p}, \hat{u}) \in \operatorname{graph} \partial\,\mathcal{G}^*$. Since A is an isomorphism it holds $A\hat{y} = \hat{u}$. Altogether $(\hat{u}, \hat{y}, \hat{p})$ is a solution of (7.1).
Now we aim to verify that the whole sequence converges to the solution $(\bar{u}, \bar{y}, \bar{p})$:
Assume that there exists a subsequence $(\tilde{u}_{\gamma_n}, \tilde{y}_{\gamma_n}, \tilde{p}_{\gamma_n})$ of $(u_\gamma, y_\gamma, p_\gamma)$ that converges weakly to $(\tilde{u}, \tilde{y}, \tilde{p}) \neq (\hat{u}, \hat{y}, \hat{p})$. With the same arguments as before we get that $(\tilde{u}, \tilde{y}, \tilde{p})$ is also a solution of (7.1). System (7.1) has a unique solution as mentioned before and therefore $(\tilde{u}, \tilde{y}, \tilde{p}) = (\hat{u}, \hat{y}, \hat{p})$. This is a contradiction to our assumption and this implies that $(u_\gamma, y_\gamma, p_\gamma) \rightharpoonup (\bar{u}, \bar{y}, \bar{p})$. □

7.1.1. Proximal Mapping of g^*

We need to derive the proximal mapping of γg^*. Therefore, we use that $\operatorname{prox}_{\gamma g^*}(v) = (\operatorname{Id} + \gamma\partial g^*)^{-1}(v)$ and that the following equivalence holds

$$w := (\operatorname{Id} + \gamma\partial g^*)^{-1}(v) \Leftrightarrow v \in (\operatorname{Id} + \gamma\partial g^*)(w) = w + \gamma\partial g^*(w). \tag{7.11}$$

To calculate the proximal mapping we go along with the cases in the subdifferential.

i) $w \in Q_1$: In this case we have $v = w + \gamma \begin{pmatrix}1\\1\end{pmatrix}$. Solving this relation for w we get $w = v - \gamma \begin{pmatrix}1\\1\end{pmatrix}$. The definition of Q_1 yields that $w_1, w_2 > 0$ and $w_1 + w_2 < 1$. Using all conditions we get that

$$v_1, v_2 > \gamma \text{ and } v_1 + v_2 < 1 + 2\gamma.$$

For the Moreau-Yosida regularization we have

$$(\partial g^*)_\gamma(p) = \frac{1}{\gamma}\left(p - \operatorname{prox}_{\gamma g^*}(p)\right) = \frac{1}{\gamma}\left(p - p + \gamma \begin{pmatrix}1\\1\end{pmatrix}\right) = \begin{pmatrix}1\\1\end{pmatrix}.$$

We can calculate the Q_i^γ's for $i = 2, 3, 4$ in the same way. For all these Q_i's the proximal mapping is $p - \gamma u_i$ and therefore the Moreau-Yosida regularization equals u_i. The sets for the definition of $(\partial g^*)_\gamma$ are the Q_i's shifted with γu_i.

ii) $w \in Q_5$: In this case we have $v = w + \gamma \begin{pmatrix}1\\1\end{pmatrix}$. Solving this equation for w and using the definition of Q_5, i.e. $w_1, w_2 > 0$ and $w_1 + w_2 > 1$, we get

$$v_1, v_2 > 2\gamma \text{ and } v_1 + v_2 > 1 + 4\gamma.$$

The Moreau-Yosida regularization is again u_i since the subdifferential of g^* is single-valued on all Q_i's. For $i = 5, 6, 7, 8$ Q_i^γ equals Q_i shifted with γu_i.

iii) $w \in S$: In this case $w_1 = w_2 = 0$ and from (7.11) we get that $v_1 \in [w_1 - \gamma, w_1 + \gamma]$ and $v_2 \in [w_2 - \gamma, w_2 + \gamma]$. Hence, we have

$$v \in [-\gamma, \gamma] \times [-\gamma, \gamma].$$

The Moreau-Yosida regularization reads the following:

$$(\partial g^*)_\gamma(p) = \frac{1}{\gamma}\left(\begin{pmatrix}p_1\\p_2\end{pmatrix} - \begin{pmatrix}0\\0\end{pmatrix}\right) = \frac{1}{\gamma}p.$$

iv) $w \in P_{12}$: In this case $w = \begin{pmatrix}1\\0\end{pmatrix}$ and we have $v \in w + \gamma \begin{pmatrix}(2-\delta)\\(2-\delta)\mu\end{pmatrix}$ for $\delta \in [0,1]$ and $\mu \in [-1, 1]$. This yields $v_1 = 1 + \gamma(2-\delta)$ and $v_2 = (v_1 - 1)\mu$. Altogether we have

$$v_1 \in [1+\gamma, 1+2\gamma] \text{ and } v_2 \in [1 - v_1, v_1 - 1].$$

For the Moreau-Yosida regularization we can insert our point P_{12}

$$(\partial g^*)_\gamma(p) = \frac{1}{\gamma}\left(p - \begin{pmatrix}1\\0\end{pmatrix}\right).$$

We can proceed in the same way for the other P_{ij}'s.

v) $w \in L_{12}$: In this case $v \in w + \gamma\left(\begin{pmatrix}1\\-1\end{pmatrix} + \lambda\begin{pmatrix}0\\2\end{pmatrix}\right)$ for $\lambda \in [0,1]$. Additionally, we have $0 < w_1 < 1$ and $w_2 = 0$ from the definition of L_{12}. Thus, we get

$$v_1 \in (\gamma, 1+\gamma) \text{ and } v_2 \in [-\gamma, \gamma].$$

Since $w_2 = 0$ we have $0 = v_2 + \gamma - 2\lambda\gamma$ and we can determine $\lambda = \frac{v_2}{2\gamma} + \frac{1}{2}$. From the definition of v_2 we see that it holds $\lambda \in [0,1]$.

$$(\partial g^*)_\gamma(p) = \frac{1}{\gamma}\left[p - p + \gamma\left(\begin{pmatrix}1\\-1\end{pmatrix} + \left(\frac{p_2}{2\gamma} + \frac{1}{2}\right)\begin{pmatrix}0\\2\end{pmatrix}\right)\right] = \begin{pmatrix}1\\\frac{p_2}{\gamma}\end{pmatrix}.$$

We can do the same calculation for L_{24}, L_{34} and L_{13}. The sets for the regularization are neighbourhoods of 2γ in one direction and 1 in the other direction of the axes. The Moreau-Yosida regularization is constant in the component that is not 0 and in the other component it is the component divided by γ.

vi) $w \in L_{56}$: In this case $v \in w + \gamma\left(\begin{pmatrix}2\\-2\end{pmatrix} + \lambda\begin{pmatrix}0\\4\end{pmatrix}\right)$ for $\lambda \in [0,1]$. We have again $w_2 = 0$ and $w_1 > 1$ from the definition of L_{56}. Hence, using both conditions

$$v_1 > 1 + 2\gamma \text{ and } v_2 \in [-2\gamma, 2\gamma]\,.$$

We can compute λ as in the case before since $w_2 = 0$ again. Here we get $\lambda = \frac{v_2}{4\gamma} + \frac{1}{2}$ and by the definition of v_2 λ is again in $[0,1]$ and therefore it is well-defined. As the Moreau-Yosida regularization we have

$$(\partial g^*)_\gamma(p) = \frac{1}{\gamma}\left[p - p + \gamma\left(\begin{pmatrix}2\\-2\end{pmatrix} + \left(\frac{p_2}{4\gamma} + \frac{1}{2}\right)\begin{pmatrix}0\\4\end{pmatrix}\right)\right] = \begin{pmatrix}2\\\frac{p_2}{\gamma}\end{pmatrix}.$$

As in the case before we can calculate the sets and regularizations in the same way for L_{68}, L_{78} and L_{57}. The sets are 4γ hoses around the axes and for the regularization we have the same as in case v).

vii) $w \in L_{15}$: In this case we have $\|w\|_{\ell^1} = 1$, $w_1, w_2 > 0$ and $v \in w + \gamma(2-\lambda)\begin{pmatrix}1\\1\end{pmatrix}$ for $\lambda \in [0,1]$. From these restrictions on w we get for v

$$v_1 - 1 < v_2 < v_1 + 1 \text{ and } 1 + 2\gamma \leq v_1 + v_2 \leq 1 + 4\gamma.$$

With the equation $w_1 + w_2 = 1$ we can determine λ as $2 - \frac{v_1+v_2-1}{2\gamma}$. From this we get the following Moreau-Yosida regularization:

$$(\partial g^*)_\gamma(p) = \frac{1}{\gamma}\left[p - p + \gamma\left(2 - 2 + \frac{p_1 + p_2 - 1}{2\gamma}\right)\begin{pmatrix}1\\1\end{pmatrix}\right] = \frac{1}{2\gamma}\left[\begin{pmatrix}p_1 + p_2\\p_1 + p_2\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix}\right].$$

We can proceed analogously for L_{26}, L_{48} and L_{37}.

In figure 7.1 one can see the different sets for the definition of the regularization. To sum up all our definitions we want to state first the sets and then the regularization. The regularized version of Q_i is

$$Q_i^\gamma = \begin{cases} \left\{q \in \mathbb{R}^2 \,|q_1, q_2 > \gamma \wedge q_1 + q_2 < 1 + 2\gamma\right\} & \text{i} = 1, \\ \left\{q \in \mathbb{R}^2 \,|q_1 > \gamma, q_2 < -\gamma \wedge q_1 - q_2 < 1 + 2\gamma\right\} & \text{i} = 2, \\ \left\{q \in \mathbb{R}^2 \,|q_1 < -\gamma, q_2 > \gamma \wedge -q_1 + q_2 < 1 + 2\gamma\right\} & \text{i} = 3, \\ \left\{q \in \mathbb{R}^2 \,|q_1, q_2 < -\gamma \wedge -q_1 - q_2 < 1 + 2\gamma\right\} & \text{i} = 4, \\ \left\{q \in \mathbb{R}^2 \,|q_1, q_2 > 2\gamma \wedge q_1 + q_2 > 1 + 4\gamma\right\} & \text{i} = 5, \\ \left\{q \in \mathbb{R}^2 \,|q_1 > 2\gamma, q_2 < -2\gamma \wedge q_1 - q_2 > 1 + 4\gamma\right\} & \text{i} = 6, \\ \left\{q \in \mathbb{R}^2 \,|q_1 < -2\gamma, q_2 > 2\gamma \wedge -q_1 + q_2 > 1 + 4\gamma\right\} & \text{i} = 7, \\ \left\{q \in \mathbb{R}^2 \,|q_1, q_2 < -2\gamma \wedge -q_1 - q_2 > 1 + 4\gamma\right\} & \text{i} = 8, \end{cases} \tag{7.12}$$

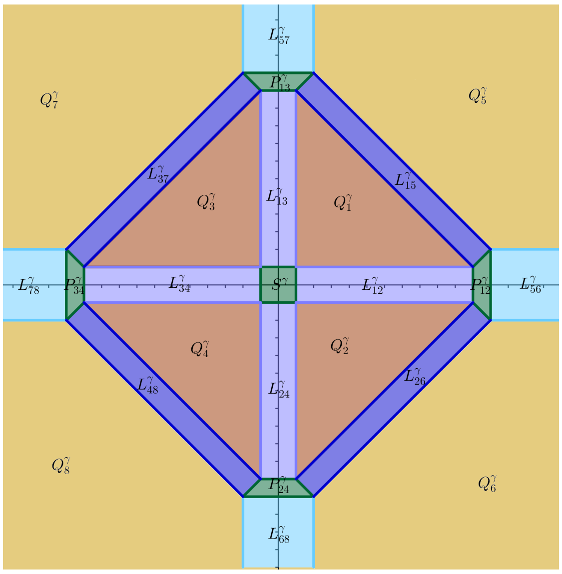


Figure 7.1.: Subdomains for the definition of $(\partial g^*)_\gamma$.

and the one of S

$$S^\gamma = \left\{ q \in \mathbb{R}^2 \,| q \in [-\gamma, \gamma] \times [-\gamma, \gamma] \right\}. \tag{7.13}$$

The regularized sets of the P_{ij}'s and L_{ij}'s are

$$P_{ij}^\gamma = \begin{cases} \left\{ q \in \mathbb{R}^2 \,| q_1 \in [1+\gamma, 1+2\gamma] \wedge q_2 \in [1-q_1, q_1-1] \right\} & \mathrm{i}=1,\, \mathrm{j}=2, \\ \left\{ q \in \mathbb{R}^2 \,| q_1 \in [1+q_2, -1-q_2,] \wedge q_2\, [-1-2\gamma, -1-\gamma] \right\} & \mathrm{i}=2,\, \mathrm{j}=4, \\ \left\{ q \in \mathbb{R}^2 \,| q_1 \in [-1-2\gamma, -1-\gamma] \wedge q_2 \in [1+q_1, -1-q_1] \right\} & \mathrm{i}=3,\, \mathrm{j}=4, \\ \left\{ q \in \mathbb{R}^2 \,| q_1 \in [1-q_2, q_2-1] \wedge q_2 \in [1+\gamma, 1+2\gamma] \right\} & \mathrm{i}=1,\, \mathrm{j}=3, \end{cases} \tag{7.14}$$

$$L_{ij}^{\gamma} = \begin{cases} \left\{q \in \mathbb{R}^2 \,|\gamma < q_1 < 1+\gamma \wedge -\gamma \leq q_2 \leq \gamma\right\} & \text{i = 1, j = 2,} \\ \left\{q \in \mathbb{R}^2 \,| -\gamma \leq q_1 \leq \gamma \wedge -1-\gamma < q_2 < -\gamma\right\} & \text{i = 2, j = 4,} \\ \left\{q \in \mathbb{R}^2 \,| -1-\gamma < q_1 < -\gamma \wedge -\gamma \leq q_2 \leq \gamma\right\} & \text{i = 3, j = 4,} \\ \left\{q \in \mathbb{R}^2 \,| -\gamma \leq q_1 \leq \gamma \wedge \gamma < q_2 < 1+\gamma\right\} & \text{i = 1, j = 3,} \\ \left\{q \in \mathbb{R}^2 \,|q_1 > 1+2\gamma \wedge -2\gamma \leq q_2 \leq 2\gamma\right\} & \text{i = 5, j = 6,} \\ \left\{q \in \mathbb{R}^2 \,| -2\gamma \leq q_1 \leq 2\gamma \wedge q_2 < -1-2\gamma\right\} & \text{i = 6, j = 8,} \\ \left\{q \in \mathbb{R}^2 \,|q_1 < -1-2\gamma \wedge -2\gamma \leq q_2 \leq 2\gamma\right\} & \text{i = 7, j = 8,} \\ \left\{q \in \mathbb{R}^2 \,| -2\gamma \leq q_1 \leq 2\gamma \wedge q_2 > 1+2\gamma\right\} & \text{i = 5, j = 7,} \\ \left\{q \in \mathbb{R}^2 \,|q_1 - 1 < q_2 < q_1 + 1 \wedge 1+2\gamma \leq q_1 + q_2 \leq 1+4\gamma\right\} & \\ & \text{i = 1, j = 5,} \\ \left\{q \in \mathbb{R}^2 \,| -q_1 - 1 < q_2 < -q_1 + 1 \wedge 1+2\gamma \leq q_1 - q_2 \leq 1+4\gamma\right\} & \\ & \text{i = 2, j = 6,} \\ \left\{q \in \mathbb{R}^2 \,|q_1 - 1 < q_2 < q_1 + 1 \wedge 1+2\gamma \leq -q_1 - q_2 \leq 1+4\gamma\right\} & \\ & \text{i = 4, j = 8,} \\ \left\{q \in \mathbb{R}^2 \,| -q_1 - 1 < q_2 < -q_1 + 1 \wedge 1+2\gamma \leq -q_1 + q_2 \leq 1+4\gamma\right\} & \\ & \text{i = 3, j = 7.} \end{cases} \tag{7.15}$$

The Moreau-Yosida regularization reads the following:

$$(\partial g^*)_\gamma(q) = \begin{cases} u_i & \text{if } q \in Q_i^\gamma, \\ \frac{1}{\gamma}p & \text{if } q \in S^\gamma, \\ \frac{1}{\gamma}\left(p - P_{ij}\right) & \text{if } q \in P_{ij}^\gamma, \\ \begin{pmatrix} \operatorname{sgn}(p_1) \\ \frac{p_2}{\gamma} \end{pmatrix} & \text{if } q \in L_{12}^\gamma \vee q \in L_{34}^\gamma, \\ \begin{pmatrix} \frac{p_1}{\gamma} \\ \operatorname{sgn}(p_2) \end{pmatrix} & \text{if } q \in L_{24}^\gamma \vee q \in L_{13}^\gamma, \\ \begin{pmatrix} 2\operatorname{sgn}(p_1) \\ \frac{p_2}{\gamma} \end{pmatrix} & \text{if } q \in L_{56}^\gamma \vee q \in L_{78}^\gamma, \\ \begin{pmatrix} \frac{p_1}{\gamma} \\ 2\operatorname{sgn}(p_2) \end{pmatrix} & \text{if } q \in L_{68}^\gamma \vee q \in L_{57}^\gamma, \\ \frac{p_1+p_2}{2\gamma}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{2\gamma}\begin{pmatrix} \operatorname{sgn}(p_1) \\ \operatorname{sgn}(p_2) \end{pmatrix} & \text{if } q \in L_{15}^\gamma \vee q \in L_{48}^\gamma, \\ \frac{p_1-p_2}{2\gamma}\begin{pmatrix} 1 \\ -1 \end{pmatrix} - \frac{1}{2\gamma}\begin{pmatrix} \operatorname{sgn}(p_1) \\ \operatorname{sgn}(p_2) \end{pmatrix} & \text{if } q \in L_{26}^\gamma \vee q \in L_{37}^\gamma. \end{cases} \tag{7.16}$$

7.2. Semismooth Newton Method

Now we want to compute a solution of our regularized system (7.10) with a semismooth Newton method. We can reduce our regularized system via elimination of u_γ to the state and the dual variable

$$\begin{aligned} \frac{1}{\alpha}A^* p_\gamma &= z - y_\gamma, \\ A y_\gamma &= H_\gamma(p_\gamma). \end{aligned} \tag{7.17}$$

First we want to give the definition of semismoothness. It is recited from [11], Def 2.1.

Definition 7.11. *Let X, Y be Banach spaces and $\mathcal{G}\colon X \to Y$ be a continuous operator. Additionally let $\partial^* \mathcal{G}$ be a generalized differential of $\mathcal{G}$, e.g. the subdifferential if $\mathcal{G}$ is convex or the Clarke differential. Then $\mathcal{G}$ is called semismooth at $x \in X$ if*

$$\sup_{M \in \partial^* \mathcal{G}(x+d)} \|\mathcal{G}(x+d) - \mathcal{G}(x) - Md\|_Y = o(\|d\|_x) \text{ for } \|d\|_x \to 0. \tag{7.18}$$

A Newton method aims to solve the problem:

For a given operator $\mathcal{G}\colon X \to Y$, find $\bar{x} \in X$ such that $\mathcal{G}(\bar{x}) = 0$.

(cf. [10], p. 3)
For a semismooth operator a Newton method is defined via the following algorithm according to [11] chapter 2.4.5.

Algorithm 1 Semismooth Newton method

Choose an initial point $x^0 \in X$ (sufficiently close to the solution $\bar{x}$)
for k = 1,2, ... **do**
  Choose $M_k \in \partial^* \mathcal{G}(x^k)$.
  Solve $M_k \delta x^k = -\mathcal{G}(x^k)$ for δx^k.
  Set $x^{k+1} = x^k + \delta x^k$.
end for

Consequently, we are interested in the solution of the following problem:

$$F(y,p) := \begin{pmatrix} y - z + \frac{1}{\alpha}A^* p \\ Ay - H_\gamma(p) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{7.19}$$

Our next goal is to verify that F is semismooth. Therefore, we want to list two properties of semismoothness. This theorem is also in line with chapter two of [11].

Theorem 7.12. *Let $X, Y, Y_i, i = 1,2$ be Banach spaces.*

i) If the operators $\mathcal{G}_i\colon X \to Y_i, i = 1,2$ are semismooth at x, then the operator $\mathcal{G} := (\mathcal{G}_1, \mathcal{G}_2)\colon X \to Y_1 \times Y_2,\ x \mapsto (\mathcal{G}_1(x), \mathcal{G}_2(x))$ is also semismooth at x. Here $(\partial^* \mathcal{G}_1, \partial^* \mathcal{G}_2)$ is defined as the direct product of the generalized differentials of the components, i.e. for $M \in (\partial^* \mathcal{G}_1, \partial^* \mathcal{G}_2)(x)$ we have $Mv = (M_1 v, M_2 v)$, where $M_1 \in \partial^* \mathcal{G}_1(x)$ and $M_2 \in \partial^* \mathcal{G}_2(x)$.

ii) If the operators $\mathcal{G}_i\colon X \to Y, i = 1,2$ are semismooth at x, then the operator $\mathcal{G}^+ := \mathcal{G}_1 + \mathcal{G}_2\colon x \to Y,\ x \mapsto \mathcal{G}_1(x) + \mathcal{G}_2(x)$ is also semismooth at x. Here $\partial^* \mathcal{G}^+ := \partial^* \mathcal{G}_1 + \partial^* \mathcal{G}_2$.

Proof. i) For all $M \in \partial^* \mathcal{G}(x+d)$ there exist by definition $M_i \in \partial^* \mathcal{G}_i(x+d), i = 1,2$ such that $Mv = (M_1 v, M_2)$. We want to use the following norm on $Y_1 \times Y_2$: $\|\cdot\|_{Y_1 \times Y_2} = \|\cdot\|_{Y_1} + \|\cdot\|_{Y_2}$.

$$\begin{aligned}
&\sup_{M \in \partial^* \mathcal{G}(x+d)} \|\mathcal{G}(x+d) - \mathcal{G}(x) - Md\|_{Y_1 \times Y_2} \\
&= \sup_{(M_1,M_2) \in \partial^* \mathcal{G}(x+d)} \|(\mathcal{G}_1(x+d), \mathcal{G}_2(x+d)) - (\mathcal{G}_1(x), \mathcal{G}_2(x)) - (M_1 d, M_2 d)\|_{Y_1 \times Y_2} \\
&= \sup_{(M_1,M_2) \in \partial^* \mathcal{G}(x+d)} \|\mathcal{G}_1(x+d) - \mathcal{G}_1(x) - M_1 d\|_{Y_1} + \|\mathcal{G}_2(x+d) - \mathcal{G}_2(x) - M_2 d\|_{Y_2} \\
&\leq \sup_{M_1 \in \partial^* \mathcal{G}_1(x+d)} \|\mathcal{G}_1(x+d) - \mathcal{G}_1(x) - M_1 d\|_{Y_1} \\
&\qquad + \sup_{M_2 \in \partial^* \mathcal{G}_2(x+d)} \|\mathcal{G}_2(x+d) - \mathcal{G}_2(x) - M_2 d\|_{Y_2} \\
&= \mathrm{o}(\|d\|_x) \text{ for } \|d\|_x \to 0,
\end{aligned}$$

since both terms are of $\mathrm{o}(\|d\|_x)$ for $\|d\|_x \to 0$.

ii)

$$\begin{aligned}
&\sup_{M \in \partial^* \mathcal{G}^+(x+d)} \left\|\mathcal{G}^+(x+d) - \mathcal{G}^+(x) - Md\right\|_Y \\
&= \sup_{(M_1+M_2) \in \partial^* \mathcal{G}^+(x+d)} \|\mathcal{G}_1(x+d) + \mathcal{G}_2(x+d) - \mathcal{G}_1(x) - \mathcal{G}_2(x) - M_1 d - M_2 d\|_Y \\
&\leq \sup_{M_1 \in \partial^* \mathcal{G}_1(x+d)} \|\mathcal{G}_1(x+d) - \mathcal{G}_1(x) - M_1 d\|_Y \\
&\qquad + \sup_{M_2 \in \partial^* \mathcal{G}_2(x+d)} \|\mathcal{G}_2(x+d) - \mathcal{G}_2(x) - M_2 d\|_Y \\
&= \mathrm{o}(\|d\|_x) \text{ for } \|d\|_x \to 0,
\end{aligned}$$

□

Part i) of this theorem yields that we need to prove that F_1 and F_2 are semismooth. F_1 is continuously differentiable in y and p and because of this it is semismooth. F_2 is also continuously differentiable in y, but not in p. So it remains to verify that H_γ is semismooth. We recall the definition of $H_\gamma(p)$:

$$H_\gamma(p)(x) = h_\gamma(p(x)) := (\partial g^*)_\gamma(p(x)),$$

i.e. H_γ is the superposition operator of h_γ. By theorem 7.9 h_γ is Lipschitz continuous with constant $\frac{1}{\gamma}$. This yields that h_γ is almost everywhere directional differentiable by Rademacher's theorem. Furthermore, h_γ is by its definition (7.16) piecewise differentiable. Thus, it holds for the directional derivative

$$h'_\gamma(q; \delta q) := \lim_{t \to 0} \frac{1}{t} \left(h_\gamma(q + t\delta q) - h_\gamma(q)\right)$$

at q in direction q that

$$\lim_{\|\delta q\|\to 0} \frac{1}{\delta q} \left\| h'_\gamma(q+\delta q;\delta q) - h'_\gamma(q;\delta q)\right\| = 0.$$

Therefore, h_γ is semismooth. [See [14], Prop. 2.7 for a proof that the condition for the directional derivative is equivalent to (7.18) in finite dimensions.] According to Ulbrich [14] Thm. 3.49 the restriction of the superposition H_γ from $H^1_\Gamma \hookrightarrow L^p(\Omega,\mathbb{R}^2)$ to $L^2(\Omega,R^2)$ for any $p>2$ is semismooth and hence Newton differentiable.
At this point we have verified that F defined in (7.19) is semismooth. To apply a semismooth Newton method we need to calculate $\partial^* F$. The challenging part is the derivative of H_γ.
Using the definition that is given by Clason, Ito and Kunisch in [6] the Newton derivative $D_N H_\gamma \colon H^1_\Gamma \to L^2(\Omega,\mathbb{R}^2)$ at p in direction of δp is given pointwise almost everywhere by

$$[D_N H_\gamma(p)\delta p]\,(x) \in \partial_\circ h_\gamma(p(x))\delta p(x).$$

Here $\partial_\circ$ denotes the Clarke derivative as defined in 4.7.
Consequently, for our h_γ stated in (7.16) the Newton derivative at p in direction δp is defined pointwise almost everywhere by

$$[D_N H_\gamma(p)\delta p]\,(x) = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \text{if } p(x) \in Q_i^\gamma, \\ \begin{pmatrix} \frac{1}{\gamma} & 0 \\ 0 & \frac{1}{\gamma} \end{pmatrix} \delta p(x) & \text{if } p(x) \in S^\gamma \vee p(x) \in P_{ij}^\gamma, \\ \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\gamma} \end{pmatrix} \delta p(x) & \text{if } p(x) \in L_{12}^\gamma \vee p(x) \in L_{34}^\gamma, \\ \begin{pmatrix} \frac{1}{\gamma} & 0 \\ 0 & 0 \end{pmatrix} \delta p(x) & \text{if } p(x) \in L_{24}^\gamma \vee p(x) \in L_{13}^\gamma, \\ \begin{pmatrix} 0 & 0 \\ 0 & \frac{2}{\gamma} \end{pmatrix} \delta p(x) & \text{if } p(x) \in L_{56}^\gamma \vee p(x) \in L_{78}^\gamma, \\ \begin{pmatrix} \frac{2}{\gamma} & 0 \\ 0 & 0 \end{pmatrix} \delta p(x) & \text{if } p(x) \in L_{68}^\gamma \vee p(x) \in L_{57}^\gamma, \\ \begin{pmatrix} \frac{1}{2\gamma} & \frac{1}{2\gamma} \\ \frac{1}{2\gamma} & \frac{1}{2\gamma} \end{pmatrix} \delta p(x) & \text{if } p(x) \in L_{15}^\gamma \vee p(x) \in L_{48}^\gamma, \\ \begin{pmatrix} \frac{1}{2\gamma} & -\frac{1}{2\gamma} \\ -\frac{1}{2\gamma} & \frac{1}{2\gamma} \end{pmatrix} \delta p(x) & \text{if } p(x) \in L_{26}^\gamma \vee p(x) \in L_{37}^\gamma. \end{cases} \tag{7.20}$$

For a semismooth Newton step according to 1 we need to obtain $(\delta y, \delta p)$ by solving

$$\begin{pmatrix} \mathrm{Id} & \frac{1}{\alpha}A^* \\ A & -D_N H_\gamma(p^k) \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \end{pmatrix} = \begin{pmatrix} z - y^k - \frac{1}{\alpha}A^* p^k \\ -Ay^k + H_\gamma(p^k) \end{pmatrix} \tag{7.21}$$

for given (y^k, p^k) and set $y^{k+1} = y^k + \delta y$ and $p^{k+1} = p^k + \delta p$.
The last goal of this section is to prove that this iteration converges superlinearly. To reach this goal we need to show uniform solvability of the Newton system (7.21). The proof of this theorem is in line with Clason, Ito and Kunisch [6] and Clason and Kunisch [7].

Theorem 7.13. *For any $p \in H^1_\gamma$ and any $(w_1, w_2) \in (H^1_\gamma)^* \times (H^1_\gamma)^*$ there exists a solution $(\delta y, \delta p) \in H^1_\gamma \times \in H^1_\gamma$ of the system*

$$\begin{cases} \delta y + \frac{1}{\alpha} A^* \delta p & = w_1, \\ A\delta y - D_N H_\gamma(p)\delta p & = w_2, \end{cases} \tag{7.22}$$

which satisfies

$$\|\delta y\|_{H^1} + \|\delta p\|_{H^1} \leq C(\|w_1\|_{H^{-1}} + \|w_2\|_{H^{-1}}).$$

Proof. First observe that $\|[D_N H_\gamma(p)]\,\delta p\|_{L^2} \leq \frac{2}{\gamma}\|\delta p\|_{L^2}$ almost everywhere by its definition (7.20). Inserting $\delta p = A^{-*}(\alpha(w_1 - \delta y))$ into the second equation of (7.22) and applying A^{-1} implies that (7.22) is equivalent to

$$\delta y + \alpha A^{-1} D_N H_\gamma(p) A^{-*} \delta y = A^{-1} w_2 + \alpha A^{-1} D_N H_\gamma(p) A^{-*} w_1. \tag{7.23}$$

The next step is to verify that $D_N H_\gamma$ is maximally monotone.
By theorem 7.9 b) h_γ is maximally monotone. Hence for all $t > 0$, all δq and almost all q it holds

$$0 \leq (h_\gamma(q + t\delta q) - h_\gamma(q)) \cdot (q + t\delta q - q) = \frac{1}{t}\,(h_\gamma(q + t\delta q) - h_\gamma(q)) \cdot (t^2 \delta q).$$

This is equivalent to

$$0 \leq \frac{1}{t}\,(h_\gamma(q + t\delta q) - h_\gamma(q)) \cdot \delta q$$

and as $t \to 0$ we get for the limit

$$h'_\gamma(q; \delta q) \cdot \delta q \geq 0.$$

We need to look at

$$\begin{aligned} \langle D_N H_\gamma(p)\delta p, \delta p\rangle_{L^2} &= \int_\Omega [D_N H_\gamma(p)\delta p]\,(x) \cdot \delta p(x)\,\mathrm{d}x \\ &= \int_\Omega h'_\gamma(p(x); \delta p(x)) \cdot \delta p(x)\,\mathrm{d}x \geq 0, \end{aligned}$$

by the estimation above.
Therefore, $D_N H_\gamma$ is maximally monotone. Since A^{-1} and A^{-*} are isomorphisms the operator $A^{-1} D_N H_\gamma A^{-1}$ is maximally monotone according to theorem 5.9. The arguments are the same as in the proof of 6.2. Minty's theorem 5.4 yields that

$\operatorname{ran}(\operatorname{Id} + A^{-1} D_N H_\gamma A^{-1}) = H^1_\Gamma$ and this implies the existence of a solution δy of (7.23) and thus the existence of a solution δp.
Now we take the inner product of (7.23) with δy and use that A^{-1} and A^{-*} are isomorphisms from $(H^1_\Gamma)^*$ to H^1_Γ and that the embedding $H^1_\Gamma \hookrightarrow L^2(\Omega, \mathbb{R}^2) \hookrightarrow (H^1_\Gamma)^*$ is continuous

$$\begin{aligned}
\|\delta y\|^2_{L^2} &\leq \|\delta y\|^2_{L^2} + \left\langle \alpha D_N H_\gamma(p) A^{-*}\delta y, A^{-*}\delta y \right\rangle_{L^2} \\
&\overset{C.S.}{\leq} \left\|A^{-1} w_2\right\|_{L^2} \|\delta y\|_{L^2} + \left\|\alpha D_N H_\gamma(p) A^{-1} w_1\right\|_{L^2} \left\|A^{-*}\delta y\right\|_{L^2} \\
&\leq \left\|A^{-1} w_2\right\|_{L^2} \|\delta y\|_{L^2} + \frac{2\alpha}{\gamma} \left\|A^{-1} w_1\right\|_{L^2} \left\|A^{-*}\delta y\right\|_{L^2} \\
&\leq C\left(\|w_1\|_{H^{-1}} + \|w_2\|_{H^{-1}}\right) \|\delta y\|_{L^2}.
\end{aligned}$$

A^* is an isomorphism and hence coercive. This implies together with the first equation of (7.22)

$$\|\delta p\|_{H^1} \leq \|A^*\delta p\|_{H^{-1}} \leq \|w_1\|_{H^{-1}} + \|\delta y\|_{H^{-1}} \leq C_1(\|w_1\|_{H^{-1}} + \|\delta y\|_{L^2}).$$

The second equation together with the boundedness of $D_N H_\gamma$ yields

$$\|\delta y\|_{H^1} \leq \|A\delta y\|_{H^{-1}} \leq \|w_2\|_{H^{-1}} + \|D_N H_\gamma(p)\delta p\|_{H^{-1}} \leq C_2(\|w_2\|_{H^{-1}} + \|\delta p\|_{L^2}).$$

Combining the last three estimations we get

$$\begin{aligned}
\|\delta y\|_{H^1} + \|\delta p\|_{H^1} &\leq C_1(\|w_1\|_{H^{-1}} + \|\delta y\|_{L^2}) + C_2(\|w_2\|_{H^{-1}} + \|\delta p\|_{L^2}) \\
&\leq \tilde{C}\left(\|w_1\|_{H^{-1}} + \|w_2\|_{H^{-1}}\right).
\end{aligned}$$

□

The boundedness of the inverse Newton matrix together with the Newton differentiability of H_γ leads according to Hinze, Pinnau, Ulbrich and Ulbrich [11, Thm. 2.12] to the following result:

Theorem 7.14. *The semismooth Newton method* (7.21) *converges locally superlinearly in* $H^1_\Gamma \times H^1_\Gamma$.

We need to define a stopping criterion for our Newton iteration. Therefore, we introduce active sets:

$$\begin{aligned}
\mathcal{A}^\gamma_i(p) &= \{x \in \Omega | p(x) \in Q^\gamma_i\}, \\
\mathcal{A}^\gamma_S(p) &= \{x \in \Omega | p(x) \in S^\gamma\}, \\
\mathcal{A}^\gamma_{Pij}(p) &= \left\{x \in \Omega | p(x) \in P^\gamma_{ij}\right\}, \\
\mathcal{A}^\gamma_{Lij}(p) &= \left\{x \in \Omega | p(x) \in L^\gamma_{ij}\right\}.
\end{aligned}$$

We terminate the iteration if all active sets coincide for p^k and p^{k+1}. At this point the regularized control can be derived by $u^{k+1} = H_\gamma(p^{k+1})$. Thus, our semismooth Newton method reads the following:

Algorithm 2 Semismooth Newton method for the regularized system (7.17)

Start with $\gamma^0 = 1$ and $(y^0, p^0) = (0, 0)$.
while $\gamma^m > 10^{-10}$ and there are nodes in the regularized active sets and # iterations $< max_{it}$ and # iterations > 1 **do**
  while Not all active sets coincide **do**
    Solve the regularized optimality system (7.17) via the semismooth Newton iteration (7.21).
  end while
  Set $\gamma^{m+1} = \frac{1}{10}\gamma^m$.
end while

We observed the same problem that is mentioned by Clason, Ito and Kunisch [6]. In our example the strategy from above failed to provide a sufficient initial guess for the next Newton step. Therefore, we implemented also a backtracking line search in addition to the Newton iteration. Due to this extra routine we need to change our stopping criterion to all active sets coincide and the norm of the residual $< 10^{-6}$.

7.3. Numerical Examples

Now we are going to show the structure of the optimal control for $\Omega = [0,1]^2$ and A the linearised elasticity equation with homogeneous Dirichlet conditions on $\Gamma_0 = \lambda\,(1,0)^T$ for $\lambda \in [0,1]$ and homogeneous Neumann conditions on $\Gamma_1 = \partial\Omega \setminus \Gamma_0$.

We take a combination of the code of Clason, Ito and Kunisch [6] and of Clason and Kunisch [7]. Therefore, we take a uniform triangulation $\mathcal{T}_h$ of the domain Ω with $N_h = 128 \times 128$ nodes. The state y and the adjoint state p are discretized with piecewise linear finite elements. As done in the two papers listed above we approximate the integration over the piecewise defined functions $H_\gamma(p_h)$ and $D_N H_\gamma(p_h)\delta p_h$ in the weak formulation of (7.21) by multiplication of the mass matrix with a vector of nodal points. The space of piecewise linear finite elements based on the interior points $\{x_j\}_{j=1}^{N_h}$ of $\mathcal{T}_h$ is denoted by V_h. Furthermore, let $v \in \mathbb{R}^{N_h}$ be defined by $v_j = v_h(x_j)$ for $1 \le j \le N_h$, $x_j \in \mathcal{T}_h$ and $v_h \in V_h$. $H_\gamma(p) \in \mathbb{R}^{N_h}$ and $DH_\gamma(p) \in \mathbb{R}^{N_h}$ are defined via

$$(H_\gamma(p))_j = h_\gamma(p(x_j)), \qquad (DH_\gamma(p))_j = D_N H_\gamma(p_h)(x_j), \qquad 1 \le j \le N_h.$$

Then the variational formulation of second equation of (7.21) reads

$$\begin{aligned} 2\mu\,\langle \epsilon(\delta y_h), \epsilon(v_h)\rangle_{L^2} + \lambda\,\langle \operatorname{div}\delta y_h, \operatorname{div} v_h\rangle_{L^2} - \langle D_N H_\gamma(p_h^k)\delta p_h, v_h\rangle_{L^2} \\ = -2\mu\langle \epsilon(y_h^k), \epsilon(v_h)\rangle_{L^2} - \lambda\langle \operatorname{div} y_h^k, \operatorname{div} v_h\rangle_{L^2} + \langle H_\gamma(p_h^k), v_h\rangle_{L^2} \quad \forall \;\; v_h \in V_h. \end{aligned}$$

This variational formulation is approximated by

$$2\mu K_h \delta y + \lambda L_h \delta y - M_h(DH_\gamma(p) \circ \delta p) = -2\mu K_h y^k - \lambda L_h y^k + M_h H_\gamma(p^k),$$

where $A_h = 2\mu K_h + \lambda L_h$ is the stiffness matrix of the linearised elasticity equation and M_h is the mass matrix corresponding to V_h and $\circ$ denotes the pointwise product of two matrices. The code that was used for the examples can be found in appendix A.

We want to illustrate two effects of the weight α. Therefore, we take the eight control states given in the introduction of chapter 6

$$\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix},\begin{pmatrix}-1\\-1\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix},\begin{pmatrix}2\\-2\end{pmatrix},\begin{pmatrix}-2\\2\end{pmatrix},\begin{pmatrix}-2\\-2\end{pmatrix}\right\}.$$

Observe that the homogeneous Dirichlet conditions on Γ_0 yield that we can not get a true multi-bang control as a solution since the nodes in Γ_0 have to be in S^γ. Thus, the termination criterion for the continuation is every time the bound of γ or the convergence of the Newton iteration in one step. This boundary condition is visualised in all plots of the control u_γ. Furthermore, we agree that our control is a true multi-bang control if the solution is in the set of control states at all nodes besides the N boundary nodes in Γ_0.
For the first examples the first component of the target is a scaled with $\frac{1}{10}$ version of Matlab's peaks function

$$\begin{aligned} z_1(x_1,x_2) = \frac{3}{10}\left(4-6x_1\right)^2 e^{-(6x_1-3)^2-(6x_2-2)^2} \\ - \left(\frac{1}{5}(6x_1-3)-(6x_1-3)^3-(6x_2-3)^5\right) e^{-(6x_1-3)^2-(6x_2-3)^2} \\ - \frac{1}{30}e^{-(6x_1-2)^2-(6x_2-3)^2} \end{aligned}$$

and for the second component $z_2(x_1,x_2)$ we take the Matrix of the discrete version of z_1 rotated counterclockwise by 90 degrees, i.e. $z_2(x_1,x_2) = rot90(z_1(x_1,x_2))$, see figure 7.2. This function has in the whole domain values that between -1 and 1 in both components.

For $\alpha = 5\cdot 10^3$ we find 149 nodes at which the control is unequal all control states u_i. If we disregard the boundary nodes where u_γ has to be zero, there are 21 nodes that does not fulfil our assumption. Hence we do not have a true multi-bang solution. This result is reached for $\gamma = 10^{-10}$ and the corresponding partition of the domain is shown in figure 7.3(a). The blue area indicates the singular arc and the green area the nodes where u_γ attains on of control states of smaller magnitude, i.e. u_γ of the green nodes is in

$$\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix},\begin{pmatrix}-1\\-1\end{pmatrix}\right\},$$

and the red area the nodes where u_γ attains the control states with larger magnitude. For $\alpha = 50$ we have disregarding the mentioned boundary nodes only 10 nodes at which u_γ is not one of the control states and thus u_γ is again not a true multi-bang control. In figure 7.3(b) one can see the partition of the domain. The continuation is terminated because the bound of $\gamma = 10^{-10}$ is reached. Since α is the weight for the distance between the state and the target this is the result we expected.

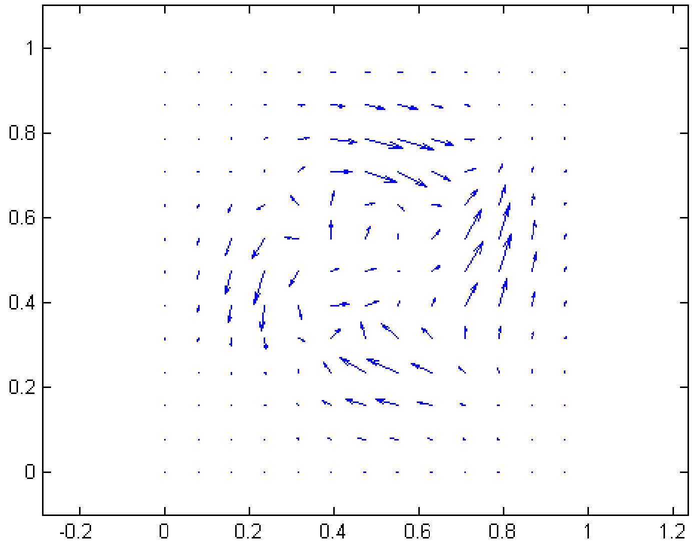
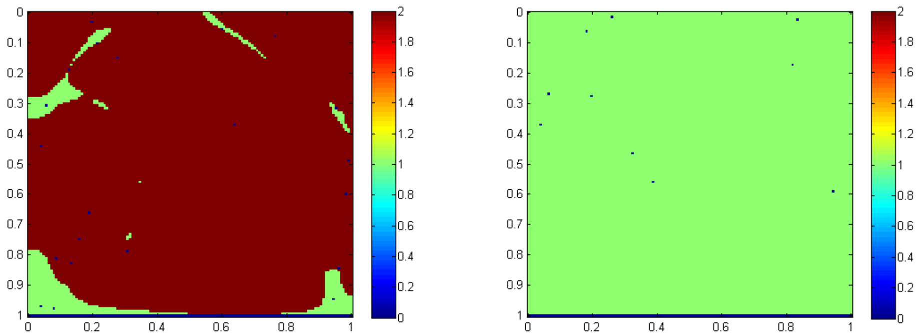


Figure 7.2.: Discrete version of the target z



(a) $\alpha = 5 \cdot 10^3$

(b) $\alpha = 50$

Figure 7.3.: Effect of α on the nodes in regularized sets (blue = singular arc, green = part of multi-bang arc with smaller magnitude, red = multi-bang arc with larger magnitude)

We want to mention that we can not choose $\alpha < 10^{-4}$ in this setting because in this case the Newton iteration in the second continuation step is terminated after the first iteration and therefore the whole continuation is terminated after the second step. Furthermore, we are not able to get a true multi-bang control for this setting.

For the second effect of α we choose again a scaled version of Matlab's peaks function for the first component and the rotated array for the second component. This time our scale factor is 10, thus the first component of the target is

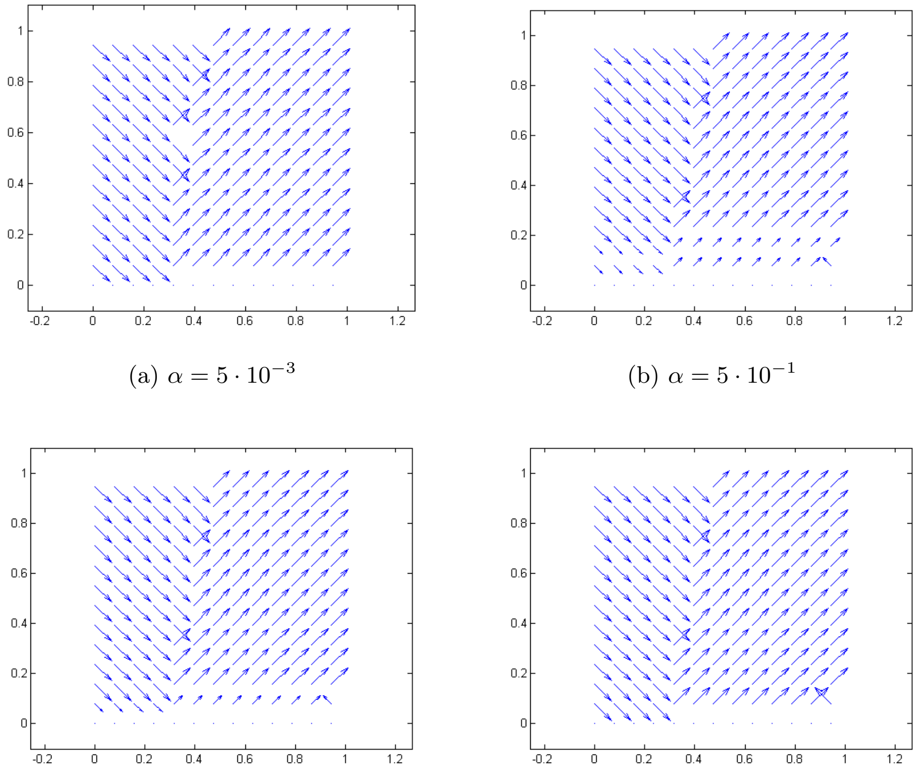
$$\begin{aligned} z_1(x_1,x_2) = 30\,(4-6x_1)^2\, e^{-(6x_1-3)^2-(6x_2-2)^2} \\ - 100\left(\frac{1}{5}(6x_1-3)-(6x_1-3)^3-(6x_2-3)^5\right) e^{-(6x_1-3)^2-(6x_2-3)^2} \\ - \frac{10}{3} e^{-(6x_1-2)^2-(6x_2-3)^2}. \end{aligned}$$

The minimal value of this function in our domain is -65.48 and the maximal value is 81.05. We do not give a plot of this target since one can not see that the vectors are longer than the ones in figure 7.2 due to the scaling of the plot routine.
For $\alpha = 5 \cdot 10^{-3}$ the continuation terminated at $\gamma = 10^{-9}$ and the solution u_γ is a true multi-bang control, see figure 7.4(a). Here we find again that all arrows have the length $\sqrt{2}$. This is shown in figure 7.5(a).
The continuation terminated at $\gamma = 10^{-7}$ for $\alpha = 5 \cdot 10^{-1}$. The solution is also a true multi-bang control, but in this case all control states are attained, see 7.4(b). The control u_γ is visualized by figure 7.4(b).
A true multi-bang solution can also be found for $\alpha = 1$. The continuation is terminated at $\gamma = 10^{-6}$. Again all control states are attained, but as we can see in figure 7.5(c) there are less nodes where the smaller u_i's are attained. The control for this set of parameters is shown in 7.4(c).
We do not get a true multi-bang control for $\alpha = 50$, there is one node beside the boundary nodes in one of the regularized sets. See figure 7.4(d) for the plot of the control. We observe in figure 7.5(d) that this control has only values within the control states with larger magnitude, i.e.

$$\left\{ \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \end{pmatrix} \right\}.$$

This set of example shows that a larger value of α causes that the control states with larger magnitude are attained.

Our next step is to take a look at the convergence of the continuation and of the Newton iteration.
Table 7.1 shows the number of Newton iterations and the number of nodes in the regularized sets for each γ and different values of α. One can see the interplay of γ and α in table 7.1. The smaller α the smaller γ to get the optimal solution. We also want to mention that the continuation fails to converge if we have a solution that is not a true



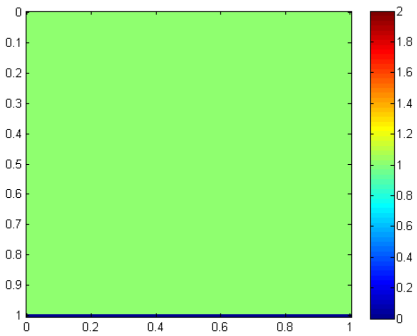
(a) $\alpha = 5 \cdot 10^{-3}$ (b) $\alpha = 5 \cdot 10^{-1}$

(c) $\alpha = 1$ (d) $\alpha = 50$

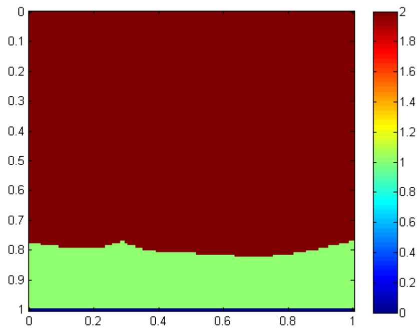
Figure 7.4.: Effect of α on the structure of the control

	$\alpha = 50$		$\alpha = 1$		$\alpha = 0.5$		$\alpha = 0.005$	
γ	It	$n(\gamma)$	It	$n(\gamma)$	It	$n(\gamma)$	It	$n(\gamma)$
1	7	804	6	12508	5	15991	2	16384
1e-1	2	189	3	2256	4	3675	2	16384
1e-2	3	135	3	402	3	604	4	12576
1e-3	3	130	3	149	3	174	3	2358
1e-4	3	129	3	130	2	133	3	436
1e-5	2	129	2	128	2	129	2	169
1e-6	2	129	1	128	2	128	3	132
1e-7	2	129			1	128	2	129
1e-8	2	129					2	128
1e-9	2	129					1	128
1e-10	2	129						

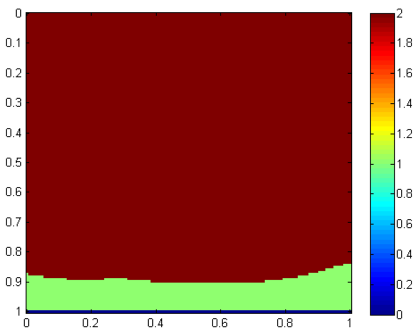
Table 7.1.: Convergence history of continuation for different α (It are the number of Newton iterations and $n(\gamma)$ are in the number of nodes that are in one of the regularized sets at the end of the Newton iteration)



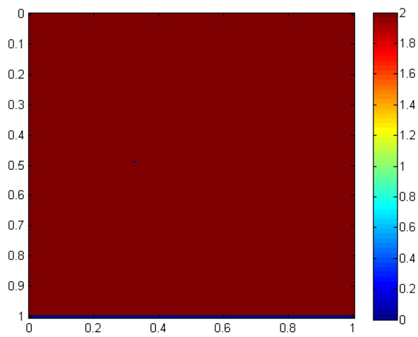
(a) $\alpha = 5 \cdot 10^{-3}$



(b) $\alpha = 5 \cdot 10^{-1}$



(c) $\alpha = 1$



(d) $\alpha = 50$

Figure 7.5.: Effect of α on the control states that are attained (blue = singular arc, green = part of multi-bang arc with smaller magnitude, red = part multi-bang arc with larger magnitude)

Iteration	1	2	3	4	5	6
$\gamma = 10^0$	16384	25920	10236	510	4	0
$\gamma = 10^{-1}$	23846	602	0			
$\gamma = 10^{-2}$	4432	22	0			
$\gamma = 10^{-3}$	558	12	0			
$\gamma = 10^{-4}$	50	2	0			
$\gamma = 10^{-5}$	4	0				
$\gamma = 10^{-6}$	0					

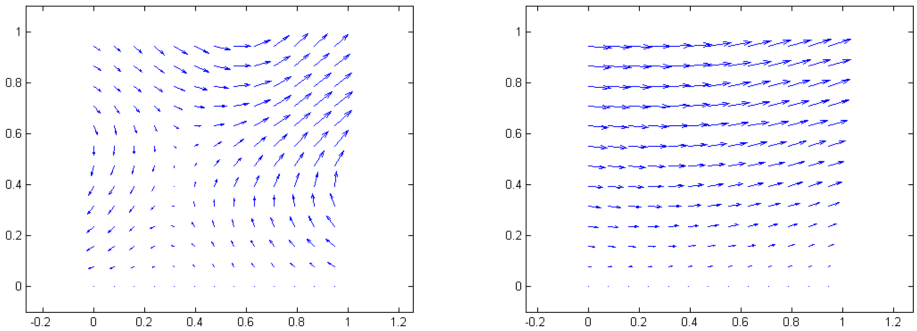
Table 7.2.: Convergence history of the Newton iteration for example in figure 7.4(c) (shown are the number of nodes that change the active set after each iteration)

multi-bang control, see column 2 of 7.1.

Furthermore, we want to take a look at the convergence of the Newton iteration. We picked example 7.4(c). The convergence history looks similar for all cases where a true multi-bang solution is attained. In table 7.2 we can see the typical convergence of a semismooth Newton method. At the beginning there are some steps with very small or even without decrease. Then we reach the superlinear phase in which the algorithm converges in a few steps. Thus, in the next step of the continuation we start in the superlinear phase and therefore the Newton iteration converges within a few steps.

At a last step we want to plot the states for both given targets, each for one value of alpha since they look very similar. See figure 7.6 for the plots with $\alpha = 50$. This figure 7.6 visualises the target's influence on the state.

The last figure shows the distance between the target z and the state y. All plots look similar, they are only scaled with different factors. Figure 7.7 shows the residual for the second version of the target and $\alpha = 1$.



(a) Target with scale factor $\frac{1}{10}$

(b) Target with scale factor 10

Figure 7.6.: State for different target

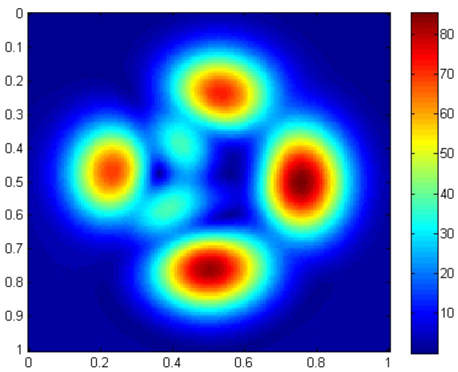


Figure 7.7.: Distance between state and target

8. Conclusion

The goal of this thesis was to provide a vector-valued version of Clason's and Kunisch's multi-bang control of elliptic systems [7]. Before we were able to state our multi-bang control problem, we needed to list some basics.

Our choice for the elliptic PDE was the linearised elasticity equation, since in this PDE the two components are linked. We started with the definition of this equation, motivated by the deformation problem that can be modelled with this equation. Afterwards we derived the weak formulation and proved that the weak formulation has a unique solution for all given right-hand sides.
The next step was to give a short introduction to optimal control theory. We motivated this introduction with the stationary heating of a solid object.

In advance of the main part of this thesis we collected several definitions and theorems from convex analysis and monotone operator theory. We introduced our multi-bang control problem with the linearised elasticity equation as the constraint and fixed set of control states in chapter 6. Then we explained how we derived the primal-dual optimality system for our problem. This included the calculation of a Fenchel conjugate and its subdifferential and the calculation of the Fréchet derivative. In the following we were able to prove the existence of a unique solution of this optimality system. The stability of this solution was also part of this section. In the last section we took a look on the structure of our solution. Clason and Kunisch were able to give a better classification of the structure of the solution. We found an example that proves that this is not possible for the vector-valued problem.

The last part aims to give an explanation for a numerical solution of the primal-dual optimality system. Since one of the equations of the system is set-valued, we needed to introduce the Moreau-Yosida regularization of the optimality system. We also proved that the solution of the regularized system converges weakly to the real solution. Afterwards we defined semismoothness and verified that our regularized system is semismooth. With this knowledge we could define a semismooth Newton method. A next step was to show that the semismooth Newton method converges locally superlinear. As a last point we listed an algorithm that is a combination of the semismooth Newton method and a backtracking. Finally, we gave some numerical examples that illustrate our theory.

A further research topic could be the generalization of our vector-valued multi-bang problem in the sense of arbitrary control states. It should not be too difficult to prove all our theorems for eight arbitrary control states that are the vertices of two different

squares centred in the origin and parallel to the axes. It is more complex to give the generalization of our problem to an arbitrary number of control states.
Clason, Ito and Kunisch were able to find similar results for optimal control with switching structure and a parabolic PDE as the constraint [6]. Thus, it would be interesting to see if one can find similar results to ours for multi-bang problems with non-elliptic PDEs as constraints.

A. Matlab Code

```
function multibang2d
%This function solves the multibang control problem
%  min alpha/2 \|y-yd\|^2 + \int \max(\max(|u1|,|u2|),1)
%      s.t. -\div(2\mu \epsilon(y) + \lambda \tr(\epsilon(y))Id)= u,
%      \max(|u1(x)|,|u2(x)|) <= 2
% using the approach described in my masterthesis.
% The code is a combination of the code from Calson, Ito and Kunisch
% regarding the paper 'A Convex Analysis Approach to Optimal Control with
% Switching Structure for Partial Differential Equations' and the code of
% Clason and Kunisch that belongs to the paper 'Multi-Bang Control of
% Elliptic Systems'.

clear all
close all

%% setup
% problem parameters
N     = 128;                          % number of nodes per dimension
maxit = 300;                          % max number of Newton steps
alpha = 5e3;                          % control cost parameter (L^2)
tmin  = 1e-10;                        % minimal step length for line search
ub = [1 1 -1 -1 2 2 -2 -2;...
      1 -1 1 -1 2 -2 2 -2];           % matrix of control states
d  = length(ub);                      % number of control states
%material parameters
E = 1;                                              % elastic modulus
nu = 0.3;                                           % Poisson's ratio
mu = E / (2*(1 + nu));                              % Lamé constant
lambda = E * nu / ((1 + nu)*(1 - 2*nu));            % Lamé constant


% setup grid, assemble stiffness and mass matrix
[K,L,M,xx,yy] = assemble2dFEM(N);
A = 2 * mu* K + lambda * L;

% setup target
z1 = peaks(N)/10;
z2 = rot90(z1);

% extract every tenth node for the plots
i = 1:10:N;

xxl = xx(i,i);
yyl = yy(i,i);
```

```
z1l = z1(i,i);
z2l = z2(i,i);

% plot every 10th point of target
figure(1)
quiver(xxl,yyl,z1l,z2l);
title('target');
xlim([-0.1 1.1]);
ylim([-0.1 1.1]);
axis equal;

z = [z1(:) ; z2(:)];

% precompute some terms
Mz = M*z(:);    AT = A';    N2 = N*N; a1 = 1/alpha;

%% compute control
% initialize iterates
y  = zeros(N2,2);                     % state variable
p  = zeros(N2,2);                     % dual variable
as = zeros(2*d+3,N2);                 % active sets

% continuation: start with gamma^0 = 1
gamma  = 1;
while gamma > 1e-10
    it = 1;    nold = 1e99;    tau = 1;    tflag = '';

    fprintf('\nCompute solution for gamma = %1.3e:\n',gamma);
    while true
        % update active sets and compute from them Hg
        % and the diagonals of DHg

        as_old = as;

        % Q_i^gamma
        as(1,:) = (p(:,1) > gamma & p(:,2) > gamma &...
                   sum(abs(p),2) < 1 + 2 * gamma);
        as(2,:) = (p(:,1) > gamma & p(:,2) < - gamma & ...
                   sum(abs(p),2) < 1 + 2 * gamma);
        as(3,:) = (p(:,1) < - gamma & p(:,2) > gamma &...
                   sum(abs(p),2) < 1 + 2 * gamma);
        as(4,:) = (p(:,1) < - gamma & p(:,2) < - gamma & ...
                   sum(abs(p),2) < 1 + 2 * gamma);
        as(5,:) = (p(:,1) > 2 * gamma & p(:,2) > 2 * gamma & ...
                   sum(abs(p),2) > 1 + 4 * gamma);
        as(6,:) = (p(:,1) > 2 * gamma & p(:,2) < - 2 * gamma & ...
                   sum(abs(p),2) > 1 + 4 * gamma);
        as(7,:) = (p(:,1) < - 2 * gamma & p(:,2) > 2 * gamma & ...
                   sum(abs(p),2) > 1 + 4 * gamma);
        as(8,:) = (p(:,1) < - 2 * gamma & p(:,2) < - 2 * gamma &...
                   sum(abs(p),2) > 1 + 4 * gamma);

        Hg = as(1:d,:)'*ub';
```

```
% S^gamma
as(d+1,:) = ( -gamma <= p(:,1) & p(:,1) <= gamma & ...
                -gamma <= p(:,2) & p(:,2) <= gamma);
Hg = Hg + (1/gamma * p) .* repmat(as(d+1,:)',1,2);

% P_i,j^gamma
as(d+2,:) = ( 1 + gamma <= p(:,1) & p(:,1) <= 1 + 2*gamma & ...
                1 - p(:,1) <= p(:,2) & p(:,2) <= p(:,1) - 1);
Hg = Hg + ...
       (1/gamma * [(p(:,1) - 1) p(:,2)]).* repmat(as(d+2,:)',1,2);

as(d+3,:) = (1 + p(:,2) <= p(:,1) & p(:,1) <= -1 - p(:,2) &...
                -1 - 2*gamma <= p(:,2) & p(:,2) <= -1 - gamma);
Hg = Hg + ...
        (1/gamma * [p(:,1) (p(:,2) + 1)]).* repmat(as(d+3,:)',1,2);

as(d+4,:) = ( -1 - 2*gamma <= p(:,1) & p(:,1) <= -1 - gamma & ...
                1 + p(:,1) <= p(:,2) & p(:,2) <= -1 - p(:,1));
Hg = Hg + ...
        (1/gamma * [(p(:,1) + 1) p(:,2)]).* repmat(as(d+4,:)',1,2);

as(d+5,:) = ( 1 - p(:,2) <= p(:,1) & p(:,1) <= p(:,2) - 1 &...
                1 + gamma <= p(:,2) & p(:,2) <= 1 + 2*gamma);
Hg = Hg + ...
        (1/gamma * [p(:,1) (p(:,2) - 1)]).* repmat(as(d+5,:)',1,2);

% L_i,j^gamma
as(d+6,:) = ((gamma < p(:,1) & p(:,1) < 1 + gamma) | ...
            (-1 - gamma < p(:,1) & p(:,1) < -gamma)) & ...
            (-gamma <= p(:,2) & p(:,2) <= gamma);
Hg = Hg + [sign(p(:,1)), (p(:,2)/gamma)].* repmat(as(d+6,:)',1,2);

as(d+7,:) = (-gamma <= p(:,1) & p(:,1) <= gamma) & ...
            ((gamma < p(:,2) & p(:,2) < 1 + gamma) | ...
            (-1 - gamma < p(:,2) & p(:,2) < -gamma));
Hg = Hg + [(p(:,1)/gamma), sign(p(:,2))].* repmat(as(d+7,:)',1,2);

as(d+8,:) = (1 + 2*gamma < p(:,1) | p(:,1) < -1 - 2*gamma) &...
            (-2*gamma <= p(:,2) & p(:,2) <= 2*gamma);
Hg = Hg + ...
       [(2*sign(p(:,1))) (p(:,2)/gamma)].* repmat(as(d+8,:)',1,2);
DHg22 = sum(as([d+1:d+6,d+8],:)'/gamma,2);

as(d+9,:) = (-2*gamma <= p(:,1) & p(:,1) <= 2*gamma) & ...
            (1 + 2*gamma < p(:,2) |  p(:,2) < - 1 - 2*gamma);
Hg = Hg + ...
       [(p(:,1)/gamma) (2*sign(p(:,2)))].* repmat(as(d+9,:)',1,2);
DHg11 = sum(as([d+1:d+5,d+7,d+9],:)'/gamma,2);

as(d+10,:) = (p(:,1) - 1 < p(:,2) & p(:,2) < p(:,1) + 1 & ...
              1 + 2*gamma <= sum(abs(p),2) & ...
              sum(abs(p),2) <= 1 + 4*gamma);
Hg = Hg + (repmat((p(:,1) + p(:,2))/(2*gamma),1,2) - ...
```

```
        [sign(p(:,1)) sign(p(:,2))]/(2*gamma))...
        .* repmat(as(d+10,:)',1,2);
    DHg12 = as(d+10,:)' /(2*gamma);
    DHg21 = as(d+10,:)' /(2*gamma);

    as(d+11,:) = (-p(:,1) - 1 < p(:,2) & p(:,2) < -p(:,1) +1 &...
                  1 + 2*gamma <= sum(abs(p),2) & ...
                  sum(abs(p),2) <= 1 + 4*gamma);
    Hg = Hg + (([(p(:,1) - p(:,2)), (-p(:,1) + p(:,2))]  - ...
        [sign(p(:,1)), sign(p(:,2))])/(2*gamma))...
        .* repmat(as(d+11,:)',1,2);
    DHg11 = DHg11 + sum(as([d+10, d+11],:)'/(2 * gamma),2);
    DHg12 = DHg12 - as(d+11,:)' /(2*gamma);
    DHg21 = DHg21 - as(d+11,:)' /(2*gamma);
    DHg22 = DHg22 + sum(as([d+10,d+11],:)'/(2*gamma),2);

    % build up the full Matrix DHg
    DHg = [spdiags(DHg11,0,N2,N2), spdiags(DHg12,0,N2,N2); ...
           spdiags(DHg21,0,N2,N2), spdiags(DHg22,0,N2,N2)];

    % system matrix, right hand side
    C   = [M a1*AT; A -M*DHg];
    rhs = [Mz-M*y(:)-a1*AT*p(:); -A*y(:) + M*Hg(:)];
    nr  = norm(rhs(:));

     % line search
    if nr >= nold               % if no decrease: backtrack
                                %  (never on first iteration)
        tau = tau/2;
        y(:) = y(:) - tau*dx(1:2*N2);
        p(:) = p(:) - tau*dx(1+2*N2:end);
        if tau < tmin     % accept non-monotone step
            tflag = 'n';
        else              % bypass rest of while loop;
            continue;
        end
    end

    % terinate Newton?
    update = nnz((as-as_old));
fprintf('It# %i: update = %i,\t residual = %1.3e,\t tau = %1.3e\n',...
        it,update,nr,tau);
    if update == 0  && nr < 1e-6             % success, solution found
        break;
    elseif it == maxit           % failure, too many iterations
        break;
    end



    % semismooth Newton step
    dx = C\rhs;
    y(:)  = y(:)+dx(1:2*N2);
    p(:)  = p(:)+dx(2*N2+1:end);
```

```
        % otherwise update information, continue
        it = it+1;   nold = nr;   tau = 1;   tflag = '';

    end %newton

    % check convergence
    if it < maxit                         % converged: accept iterate
        u = Hg;

        regnodes = nnz(as(d+1:end,:)); % number of nodes in
                                       % regularized active sets

        fprintf('Solution has %i node(s) in regularized active sets\n',...
                                                         regnodes);

        if regnodes == 0 || it == 1       % solution optimal: terminate
            break;
        else                              % reduce gamma, continue
            gamma = gamma/10;
        end
    else                                  % not converged: reject, terminate
      fprintf('Iterate rejected, returning u_gamma for gamma = %1.3e\n',...
                                                          gamma*10);
        break;
    end

end

% calculate distance of x and z
resyz = sqrt(sum(reshape(y(:)-z,N2,2).^2,2));

u1 = reshape(u(:,1),N,N);
u2 = reshape(u(:,2),N,N);
y1 = reshape(y(:,1),N,N);
y2 = reshape(y(:,2),N,N);

% extrac every 10th points to be able to plot the state and the control
u1l = u1(i,i);
u2l = u2(i,i);

y1l = y1(i,i);
y2l = y2(i,i);

% plot control
figure(2)
quiver(xxl,yyl,u1l,u2l);
title('control');
xlim([-0.1 1.1]);
ylim([-0.1 1.1]);
axis equal;
```

```
% Build color matrix to see the different domains: Singular arc and
% multi-bang arc
fb = zeros(N2,1);
for i = 1:N2
    if sum(as(1:4,i) == 1) == 1
        fb(i) = 1;
    end
    if sum(as(5:8,i) == 1) == 1
        fb(i) = 2;
    end
end
figure(3)

% plot partition of domain
x = linspace(0,1,N);
imagesc(x,x,flipud(reshape(fb,N,N)))
caxis([0,2]);

% plot state
figure(4)
quiver(xxl,yyl,y1l,y2l);
title('state');
xlim([-0.1 1.1]);
ylim([-0.1 1.1]);
axis equal;

% plot residual of y and z
figure(5)
imagesc(x,x,reshape(resyz,N,N));
title('residual');
end

function [K,L,M,xx,yy] = assemble2dFEM(n)
a   = 0;     b = 1;         % computational domain [a,b]^2
nel = 2*(n-1)^2;            % number of nodes
h2  = ((b-a)/(n-1))^2;      % Jacobi determinant of transformation (2*area(T))
n2 = n * n;

% nodes
[xx,yy] = meshgrid(linspace(0,1,n));

% triangulation
tri = zeros(nel,3);
ind = 1;
for i = 1:n-1
    for j = 1:n-1
        node         = (i-1)*n+j+1;               % two triangles at node
        tri(ind,:)   = [node node-1 node+n];      % triangle 1 (lower left)
        tri(ind+1,:) = [node+n-1 node+n node-1]; % triangle 2 (upper right)
        ind = ind+2;
    end
end

% Mass and stiffness matrices
```

```
L11e = 1/2 * [1 0 -1 0 1 0 -1 0 0]';      % elemental block matrices for
L12e = 1/2 * [-1 1 0 0 0 0 1 -1 0]';      % div u * div v
L22e = 1/2 * [1 -1 0 -1 1 0 0 0 0]';
Me = h2/24 * [2 1 1 1 2 1 1 1 2]';        % elemental mass matrix

ent = 9*nel;
row = zeros(ent,1);
col = zeros(ent,1);
valM = zeros(ent,1);
valL11 = zeros(ent,1);
valL12 = zeros(ent,1);
valL22 = zeros(ent,1);
valK11 = zeros(ent,1);
valK12 = zeros(ent,1);
valK22 = zeros(ent,1);

ind = 1;
for el=1:nel
    ll        = ind:(ind+8);            % local node indices
    gl        = tri(el,:);              % global node indices
    row(ll)   = gl([1;1;1],:); rg = gl';
    col(ll)   = rg(:,[1 1 1]);
    valK11(ll) = L11e + 1/2 * L22e;
    valK12(ll) = 1/2 * L12e;
    valK22(ll) = L22e + 1/2 * L11e;
    valL11(ll) = L11e;
    valL12(ll) = L12e;
    valL22(ll) = L22e;
    valM(ll) = Me;
    ind       = ind+9;
end

Ml = sparse(row,col,valM);
M = [ Ml, sparse(n2,n2); sparse(n2,n2), Ml];
L11 = sparse(row,col,valL11);
L12 = sparse(row,col,valL12);
L22 = sparse(row,col,valL22);
L = [ L11 , L12; L12', L22];
K11 = sparse(row,col,valK11);
K12 = sparse(row,col,valK12);
K22 = sparse(row,col,valK22);
K = [K11, K12; K12' , K22];

% modify matrices for homogenenous Dirichlet conditions on Gamma_0
bdnodd = find(abs(yy-a) < eps); %nodes for Dirichlet condition

M([bdnodd; bdnodd + n2],:) = 0;
K([bdnodd; bdnodd + n2],:) = 0;  K(:,[bdnodd; bdnodd + n2]) = 0;
L([bdnodd; bdnodd + n2],:) = 0; L(:,[bdnodd; bdnodd + n2]) = 0;
for j = [bdnodd; bdnodd + n2]'
    K(j,j) = 1;
    L(j,j) = 1;
end
end
```

```matlab
% This function is used to build the stiffness- and the massmatrix to solve
% the problem -\Delta y = u.
% The code is a variation of the code from Clason and Kunisch that belongs
% to the paper 'Multi-Bang Control of Elliptic Systems'.



function [K,M,xx,yy,bdnod0,bdnodD] = assembleFEMLaplace2d(n)
a   = 0;     b = 1;       % computational domain [a,b]^2
nel = 2*(n-1)^2;          % number of nodes
h2  = ((b-a)/(n-1))^2;    % Jacobi determinant of transformation (2*area(T))
n2 = n*n;

% nodes
[xx,yy] = meshgrid(linspace(0,1,n));

% triangulation
tri = zeros(nel,3);
ind = 1;
for i = 1:n-1
    for j = 1:n-1
        node         = (i-1)*n+j+1;              % two triangles at node
        tri(ind,:)   = [node node-1 node+n];      % triangle 1 (lower left)
        tri(ind+1,:) = [node+n-1 node+n node-1]; % triangle 2 (upper right)
        ind = ind+2;
    end
end

% Mass and stiffness matrices
Ke = 1/2 * [2 -1 -1 -1 1 0 -1 0 1]';    % elemental stiffness matrix
Me = h2/24 * [2 1 1 1 2 1 1 1 2]';      % elemental mass matrix

ent = 9*nel;
row = zeros(ent,1);
col = zeros(ent,1);
valk = zeros(ent,1);
valm = zeros(ent,1);

ind = 1;
for el=1:nel
    ll       = ind:(ind+8);              % local node indices
    gl       = tri(el,:);                % global node indices
    row(ll)  = gl([1;1;1],:);
    rg = gl';
    col(ll)  = rg(:,[1 1 1]);
    valk(ll) = Ke;
    valm(ll) = Me;
    ind      = ind+9;
end
M1 = sparse(row,col,valm);
M = [M1 , sparse(n2,n2); sparse(n2,n2), M1];
K1 = sparse(row,col,valk);
K = [K1 , sparse(n2,n2); sparse(n2,n2), K1];
```

```
% modify matrices for homogenenous Dirichlet conditions
bdnod0 = [find(abs(xx-a) < eps); find(abs(xx-b) < eps)]; % find 0 nodes
bdnodD = [find(abs(yy-a) < eps); find(abs(yy-b) < eps)]; % find other
                                                          %       bdnods
M([bdnod0; bdnod0 + n2],:) = 0;
K([bdnod0; bdnod0 + n2],:) = 0;  K(:,[bdnod0; bdnod0 + n2]) = 0;
for j = [bdnod0; bdnod0 + n2]'
    K(j,j) = 1; %#ok<SPRIX>
    M(j,j) = 1;
end
end
```

List of Figures

List of Tables

Bibliography

[1] H. H. Bauschke and P. L. Combettes. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, 2011.

[2] D. Braess. Finite Elemente Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie. Springer, 1992.

[3] H. Brezis, M. G. Crandall, and A. Pazy. Perturbations of nonlinear maximal monotone sets in banach space. COMMUNICATIONS ON PURE AND APPLIED MATHEMATIC, 23:123 – 144, 1970.

[4] F. H. Clarke. Optimization and Nonsmooth Analysis. Society for Industrial and Applied Mathematics, 1990.

[5] C. Clason. Mathematische bildverarbeitung. Universtity Lecture, 2014.

[6] C. Clason, K. Ito, and K. Kunisch. A convex analysis approach to optimal controls with switching structure for partial differential equations. ESAIM: Control, Optimisation and Calculus of Variations, 2015.

[7] C. Clason and K. Kunisch. Multi-bang control of elliptic systems. Annales de l'Institut Henri Poincare (C) Non Linear Analysis, 31:1109 – 1130, 2013.

[8] I. Ekeland and R. Temam. Convex analysis and variational problems, volume 1. North-Holland Publishing company, 1976.

[9] L. C. Evans. Partial Differential Equations. American Mathematical Society, 2010.

[10] M. Hintermüller. Semismooth newton method and applications, 2010.

[11] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. Optimization with PDE constraints, volume 23. Springer, New York, 2009.

[12] D. Kinderlehrer and G. Stampacchia. An Introduction to Variational Inequalities and Their Applications. Society for Industrial and Applied Mathematics, 2000.

[13] W. Schirotzek. Nonsmooth Analysis. Springer, 2007.

[14] M. Ulbrich. Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces. SIAM, 2011.

[15] K. Yosida. Functional Analysis. Springer, 1974.