



UNIVERSITÄT
BAYREUTH

**OPTIMAL CONTROL OF TIME-DISCRETIZED
CONTACT PROBLEMS**

Von der Universität Bayreuth
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)
genehmigte Abhandlung

von

Georg Arne Müller

geboren in Berlin

1. Gutachter: Prof. Dr. Anton Schiela
2. Gutachter: Prof. Dr. Roland Herzog
3. Gutachter: Prof. Dr. Christian Meyer

Tag der Einreichung: 13.11.2018

Tag des Kolloquiums: 08.05.2019

Mai 2019

Zusammenfassung

Die dynamische Wechselwirkung zwischen einem viskoelastischen Körper und einem starren Hindernis ist von Natur aus nichtlinear, und Optimierungsprobleme mit diesen Nebenbedingungen sind komplexe, nichtglatte Probleme mit Komplementaritätsbeschränkungen. Optimierungsalgorithmen für solche Probleme zu entwickeln und deren Potential korrekt einzuschätzen erfordert ein detailliertes Verständnis der Problemstruktur, insbesondere bezüglich der Beschaffenheit von den Kontaktbereichen als Ursache der Nichtglattheit. In dieser Arbeit werden umfangreiche Sensitivitätsuntersuchungen für zeitdiskretisierte Einkörperkontaktprobleme in reibungsfreier, linearer Viskoelastizität vorgestellt und Existenz sowie starke Stationarität von Minimierern der dazugehörigen Optimalsteuerungsprobleme gezeigt. Die Ergebnisse werden für die Entwicklung, Implementierung und Auswertung von zwei Optimierungsalgorithmen, die auf einem adjungierten Problem basieren, genutzt.

Die Analyse des Problems hängt stark von der schwachen Formulierung der Kontaktbedingungen auf dem Gebietsrand ab. Verschiedene Sobolev-Kapazitäten werden hinsichtlich ihres Verhaltens am Rand und ihrer Eignung für die Verwendung in der Formulierung der Nebenbedingung untersucht. Unter schwachen Voraussetzungen an die Daten wird bewiesen, dass alle geeigneten Kapazitätsbegriffe äquivalent sind. Weiterhin wird gezeigt, dass die Formulierungen der Kontaktbedingungen, die auf dem quasi-überall-Sinn dieser Kapazitäten beruhen, mit der klassischen, maßtheoretischen Formulierung übereinstimmen. In der darauffolgenden Analysis ermöglicht das die Verwendung einer Vielzahl verschiedener Resultate aus den jeweiligen, bisher voneinander unabhängigen Ansätzen.

Aus dem zeitkontinuierlichen Kontaktproblem wird über nichtkonforme Finite-Elemente ein kontaktimplizites zeitdiskretisiertes Problem hergeleitet, von dessen Lösungsoperator die Hadamard-Differenzierbarkeit gezeigt wird. Eine lokalisierte Darstellung der Kontaktkräfte ermöglicht eine punktweise Charakterisierung der linearisierten Randbedingungen, durch welche die Punkte der Gâteaux-(Nicht-)Differenzierbarkeit des Operators identifiziert werden können. Mit Hilfe der Differenzierbarkeitsinformationen werden die Existenz von Minimierern des Optimierungsproblems sowie deren Stationaritätsbedingungen bewiesen.

Aufbauend auf dem adjungierten Problem wird eine subgradientenartige Suchrichtung für die Verwendung in einer Liniensuche und einem Impulsverfahren berechnet. Das Verhalten von deren Implementierungen wird anhand dreier Testprobleme numerisch ausgewertet. Insbesondere wird die Geometrie- und Zielfunktionalabhängigkeit der problemspezifischen Nichtglattheit und ihrer Einflüsse auf das Verhalten der Algorithmen untersucht.

Abstract

The dynamic interaction of a viscoelastic body and a rigid obstacle is inherently nonlinear, and optimization problems with these constraints are complex, nonsmooth complementarity constrained problems. Developing optimization algorithms for these problems and correctly assessing their limitations requires detailed knowledge of the problem structure, specifically with respect to the characteristics of contact patches as the source of non-smoothness. In this thesis, a comprehensive sensitivity analysis of frictionless, linearly viscoelastic, time-discretized one-body contact problems is presented, and existence as well as strong stationarity of minimizers of the associated optimal control problems are shown. The results are used in the design, implementation and evaluation of two adjoint-based optimization schemes.

The analysis of the problem is strongly dependent on the formulation of the contact constraints on the boundary of the domain in the framework of weak solutions. Several Sobolev capacities are examined with respect to their boundary behavior and their suitability for the constraint's formulation. Under mild regularity assumptions on the data, all reasonable notions of capacity are proven to be equivalent. Additionally, the formulation of the contact constraints with respect to the corresponding quasi everywhere sense is shown to coincide with the formulation using the classic, measure theoretical sense. This enables the use of a wide range of results from these previously unrelated approaches in the subsequent analysis.

The time-continuous contact problem is discretized using nonconforming finite elements, and the solution operator to the contact implicit time-discretized problem is shown to be Hadamard differentiable. A localized representation of the contact forces is derived to obtain a pointwise characterization of the linearized boundary conditions, which allows for the identification of the operator's points of (non-)differentiability in the sense of Gâteaux. Existence of minimizers of the optimization problem and their stationarity conditions are derived using the differentiability information on the operator.

Based on the adjoint problem in the stationarity condition, a subgradient-type search direction is computed as part of a line search method and a corresponding momentum method. The behavior of their implementations is evaluated numerically using three test configurations. Particularly, the dependence of the problem-specific nonsmoothness and its effects on the algorithms on the contact boundaries' geometries and the objective functional are examined.

Acknowledgments

I am sincerely grateful to my supervisor Anton Schiela for introducing me to this interesting topic and for giving me the opportunity to pursue this line of research. His feedback and advice were a substantial help during our years of joint research. I remain impressed by his ability to seamlessly switch from unrelated topics to involved, technical discussion with me in an instant.

For committing their time to refereeing this thesis, I would also like to thank Roland Herzog and Christian Meyer.

Thanks go to Constantin Christof for taking the time to co-author our publication, which the third chapter in this thesis is based on, and to Gerd Wachsmuth for pointing out some of the key literature that is referenced in the analysis in Chapter 4. I also greatly appreciate the feedback of my colleagues Manuel Schaller and Matthias Stöcklein.

Finally, I am thankful for the support of my family and especially of my wife Veronika.

Contents

Zusammenfassung	iii
Abstract	v
Acknowledgments	vii
1 Introduction	1
1.1 Contributions and Outline	3
1.2 Notation	4
2 Modeling of Contact Problems	11
2.1 Linear Viscoelasticity	12
2.2 Contact Constraints	19
2.2.1 Non-Penetration and Contact Stresses	19
2.2.2 Discussion	21
2.3 The Variational One-Body Contact Problem	23
3 Capacities on the Boundary	27
3.1 Abstract Setting	28
3.2 Sobolev Capacities	35
3.3 Equivalence of Sobolev Capacities	41
3.3.1 Equivalences on $\mathcal{P}(\bar{\Omega})$	41
3.3.2 Equivalences on $\mathcal{P}(\partial\Omega)$	48
3.4 Conclusions and Implications	56
3.5 The Weak Non-Penetration Condition	62
4 Time Discretization and Analysis of the Contact Problem	65
4.1 Time Discretization of the Continuous Problem	67
4.2 The Time-Discretized Contact Problem	72
4.3 Solutions to the Time-Discretized Problem	75
4.4 Differentiability of the Solution Operator	78
4.4.1 Hadamard Differentiability	78
4.4.2 Linearized Boundary Conditions	86
4.4.3 Gâteaux and Fréchet Differentiability	92

5	Optimal Control of the Time-Discretized Contact Problem	95
5.1	Time-Discretized Controls	96
5.2	Existence of Minimizers	97
5.3	Stationarity Conditions	98
5.4	Structure of the Adjoint Problem	101
5.4.1	Adjoint Time Stepping	101
5.4.2	Boundary Conditions and Existence of Adjoint States	105
5.5	Differentiability of the Reduced Objective Functional	107
6	Numerical Optimization	111
6.1	Adjoint-Based Optimization Algorithms	112
6.1.1	Search Directions	115
6.1.2	Step Lengths	119
6.2	Numerical Results	122
6.2.1	Implementation Details	124
6.2.2	Example 1: A Well-Behaved Inverse Problem	125
6.2.3	Example 2: A Nonsmooth Positioning Problem	134
6.2.4	Example 3: A Parameter Balancing Problem	145
6.3	Evaluation	151
7	Conclusions and Outlook	153
A	Function Analytical Results	155
A.1	Integration Theory	155
A.2	Sobolev Traces	156
A.3	Regularity of the Contact Normal	161
B	Capacity Related Results	167
B.1	Nested Exceptional Sets	167
B.2	Quasi Open and Quasi Closed Sets	168
B.3	Quasi Continuous Functions	171
C	Miscellaneous	175
C.1	Singer's Theorem	175
C.2	Cones in Product Spaces	175
C.3	Hadamard Differentiability	178
C.4	Gâteaux and Fréchet Differentiability	179
C.5	Auxiliary Numerical Results	180
C.5.1	Influence of the Forward Solver's Tolerance	180
C.5.2	Momentum Restarts of the Accelerated Scheme in Example 1	182
C.6	Boundedness of the Boundary-To-Domain Extension Operator	183
	Lists of Symbols, Figures, Tables and Algorithms	187
	Bibliography	197
	Publications	211

Chapter 1

Introduction

Contact problems involving viscoelastic and rigid bodies are an essential part of solid mechanics with a great deal of engineering and (bio)mechanical applications. The mathematical research on the topic dates back to the early works of Poisson, Saint-Venant, Voight and Hertz in the mid 1800s, see [110, Sec. 1.3], and has since produced a large number of publications that cover various different models and aspects of mechanical contact, including static and dynamic effects, punches, viscosity, friction, adhesion, damage, plasticity and thermal effects in one- and multi-body configurations. Refer to the monographs [62, 84, 110, 125] for an overview of the topic. In this thesis, the optimal control of dynamic contact involving a linearly viscoelastic body and a rigid obstacle in the absence of friction is considered, a setting that is well suited for situations where the elastic deformations are expected to be small and where frictional effects at the contact patches can be disregarded, e.g., due to lubrication. The linearized non-penetration condition employed in the following is a variation of the condition proposed by Signorini in 1933 ([175]), which remains one of the prevalently used constraint models, see, e.g., [4, 27, 29, 57, 98, 109, 149, 179] and many others. Resulting from Signorini's condition are variational inequalities that feature the inherent nonlinearity and nonsmoothness of the contact condition, owing to one of the fundamental issues of contact mechanics — the a priori unknown contact patches.

Existence of a Lipschitz continuous solution operator for static contact problems is covered by the general result for elliptic variational inequalities proven by Lions and Stampacchia in [130]. Further, the sensitivity analysis in Mignot's work [138] shows that directional differentiability of the operator is obtained if the set of admissible displacements is polyhedral — a property that has since become increasingly important in inequality constrained optimal control. An extensive survey of the concept can be found in [195]. Numerically, contact problems can be solved with optimal complexity by the multigrid techniques developed in [116, 117] or, alternatively, by combining regularization approaches and semismooth Newton methods, see [98, 183, 191].

Analytic results for time dependent, hyperbolic contact problems have been developed for related models, such as viscoelasticity with singular memory [102] and frictional contact [47, 103, 123]. The frictionless, viscoelastic setting of this thesis, on the other hand, is rarely analyzed. The only apparent publication [4] investigates the existence of a possibly

non-unique solution to the hyperbolic variational inequality. However, several publications address the time discretization of the dynamic contact problem based on modifications of the Newmark time stepping scheme that was introduced in [150], such as the well-known energy dissipative, contact implicit scheme that was introduced by Kane et al. in [106]. Various stabilizing techniques that are based on this method and that deal with spurious oscillations introduced by spatial discretization of the problem are considered in [52, 81, 112, 113, 114, 118]. Overviews of time discretization schemes can be found in [57, 118]. An alternative spatial discretization of the contact constraint is proposed in [41], and an additional class of methods that stress conservation principles and also cover nonlinear contact problems is studied in [125, 126]. Spatial and temporal adaptivity for contact problems based on Newmark schemes are investigated in [29] and [112], respectively.

Optimal control problems with general elliptic inequality constraints are considered in, e.g., [138, 140], and optimization of the closely related — but scalar — obstacle problem are investigated in [31, 138, 194, 196]. Sensitivity analysis of elliptic variational inequalities in shape optimization, including the obstacle and the static Signorini problem, is addressed in [182]. Betz's recently published work [28] contains stationarity results for optimal control of frictionless, viscoelastic, static contact problems in a similar setting as the one considered in this thesis but involves a more restrictive contact condition that requires increased regularity assumptions on the data. Strong stationarity results and constraint qualifications for general classes of complementarity constrained problems in Banach spaces have been examined in [193]. Numerical results for the optimization of contact problems are scarce, except for the algorithmic considerations for the optimization of a static contact problem in a medical design application presented in [198]. Given the lack of information on existence of unique solutions to the dynamic contact problem, to the best knowledge of the author, there are no publications dealing with optimal control of either the corresponding hyperbolic variational inequality or a time-discretized counterpart. Optimal control of parabolic variational inequalities, on the other hand, has been addressed in, e.g., [18, 42, 101, 139].

The technical analysis of the contact problem in this thesis relies heavily on a proper understanding of the active contact sets and is therefore strongly dependent on the precise formulation of the contact constraint in the framework of weak solutions. Compared to the obstacle problem, which extends the scalar Poisson problem by inequality constraints on the interior of the domain, contact problems extend vector valued problems in linear elasticity by constraints on the boundary of the domain, which generally introduces mixed boundary conditions. While the measure theoretical boundary trace sense for Sobolev functions is commonly used in the analysis of the weak contact problem formulation, the considerations in the literature regarding differentiability of solution operators to variational inequalities and the optimal control results for the obstacle problem suggest that a capacity-based formulation is advantageous in the sensitivity and optimal control analysis, see, e.g., [31, 138, 194]. Several different approaches to the notion of a capacity with a reasonable boundary behavior that deals with the mixed boundary constraints can be found in the literature, such as Dirichlet-space-based approaches [138], approaches based on dropping the Dirichlet boundary conditions on the boundary [85, 138] as well as the technique of choosing a sufficiently large superdomain to avoid effects introduced by the mixed boundary conditions, e.g., [28], cf. [88]. These approaches are initially unconnected.

1.1 Contributions and Outline

The main objective of this thesis is the development of optimal control theory for problems governed by frictionless, time-discretized, dynamic one-body contact problems in linear viscoelasticity and the implementation and evaluation of adjoint-based optimization schemes based on an extensive differential sensitivity analysis of the constraints. A solid foundation for the analysis is established by a thorough examination of the boundary behavior and the equivalence of several Sobolev capacities regarding their use in the weak formulation of the contact constraint. The structure is as follows.

Chapter 1 – Introduction. The remainder of this chapter provides an overview of the notation and of the tools from topology, function theory and measure theory that are required in the following analysis. References for further reading are included.

Chapter 2 – Modeling of Contact Problems. This chapter contains a brief introduction into the modeling of solid mechanics, Kelvin-Voight-type materials in linear viscoelasticity and a variant of Signorini’s linearized non-penetration condition that uses a contact normal direction based on an independent contact mapping instead of the commonly employed geometric normal — similarly to the way the contact condition is employed in the setting of two-body contact problems. After a discussion regarding the benefits of the decoupling of the body’s geometry and the contact direction, the strong and the weak form of the time dependent, hyperbolic contact problem are presented. All physical assumptions on the model are fixed within this chapter.

Chapter 3 – Capacities on the Boundary. As a foundation for this chapter, we review basic facts from capacity theory in an abstract setting that is tailored to the study of Sobolev capacities. We examine several approaches to these capacities with respect to their behavior at and close to the boundary of the domain to find that not all of them — especially not the notion commonly employed in the analysis of the obstacle problem — are suitable for the treatment of contact conditions on the boundary. Under mild regularity assumptions on the data, those capacities that show reasonably stable behavior near the boundary are shown to be equivalent on the closure of the domain, and we establish that the corresponding quasi everywhere sense on the boundary coincides with the boundary measure theoretical sense for Sobolev traces. As a consequence, either sense can be employed in the formulation of the contact constraint, which allows for the use of pre-existing results from all approaches in the analysis of the subsequent chapters.

Chapter 4 – Time Discretization and Analysis of the Contact Problem. We derive a temporal finite element discretization that essentially corresponds to the contact implicit Newmark scheme by Kane [106] and allows for the consistent derivation of an adjoint time stepping scheme. The resulting time discrete problem is comprised of a sequence of elliptic variational inequalities and is shown to allow for a Lipschitz continuous solution operator. We prove directional differentiability of this solution operator in the sense of

Hadamard by establishing polyhedricity of the set of admissible displacements. Further, we derive a localized representation of the contact forces and a pointwise characterization of the linearized boundary conditions that allow for the identification of the points of Gâteaux differentiability of the operator based on strong and weak contact patches.

Chapter 5 – Optimal Control of the Time-Discretized Contact Problem. After establishing the optimal control problem and the temporal finite element discretization of the control, we show existence of minimizers for the time-discretized problem under mild conditions. Assuming dense controls, we obtain first order optimality conditions of strong stationarity type for the control problem. The chapter is concluded by a discussion of the structure of the adjoint problem in the stationarity condition and the differentiability properties of the reduced objective functional.

Chapter 6 – Numerical Optimization. Based on the pointwise representation of the linearized boundary conditions, we modify the boundary conditions in the adjoint problem to be able to compute a subgradient-type search direction, which we use in a line search method and an accelerated momentum method with different step length computation schemes. The behavior of the algorithms' implementations for three tracking-type test configurations are compared to show that the influence of nonsmooth effects on the behavior of the solvers and on the solutions are strongly dependent on the geometries of the contact boundaries and the objective functional.

Chapter 7 – Conclusions and Outlook. The results of this thesis are summarized and put in perspective with respect to open questions and current research on related topics.

Appendix. The appendix contains auxiliary results on measure theory and transformation results for Lebesgue functions and Sobolev traces that mostly support the capacity theoretical analysis of Chapter 3. Some results on the connection between Gâteaux and Fréchet differentiability in Banach spaces are shown. We briefly discuss the relation between Lipschitz continuous functions and $W^{1,\infty}(\Omega)$ -functions and the implications for the regularity of the contact normal. Some additional numerical results are presented.

1.2 Notation

The purpose of this section is to introduce the notation and provide references for the elements of geometry, topology, the analysis of partial differential equations and of functional analysis that are employed throughout this thesis. All concepts that either require a more detailed introduction or are only relevant within a limited scope will be introduced when used for the first time. A list of symbols is included after the appendix.

The natural, rational and real numbers are denoted by the symbols \mathbb{N} , \mathbb{Q} and \mathbb{R} , respectively, and $\overline{\mathbb{R}}$ means the extended real numbers $\mathbb{R} \cup \{\pm\infty\}$. We will consider contact

problems in d -dimensional Euclidean space \mathbb{R}^d with d in $\{2, 3\}$ ¹ and with a fixed orthonormal basis $\{e_i : i = 1, \dots, d\}$. There will be no further distinction between a point \mathbf{x} in space and its coordinates with respect to the basis. Time parameters are denoted by t , and in the time-discretized setting, the time step size is denoted by τ .

In the following, domains will be understood to be open and connected subsets of \mathbb{R}^d . A domain $\Omega \subset \mathbb{R}^d$ is called a strong Lipschitz domain if its boundary is strongly Lipschitz, i.e., if its boundary can be locally represented as the graph of a Lipschitz continuous function. The precise definition of this regularity condition is of some importance for the results in Chapter 3. It can be found in Definition A.3.10 in the appendix as stated in [65, Def. 4.4], see also [121, Def. 6.2.2]. For a distinction between strong and weak Lipschitz boundaries, see the “two brick” example, e.g., [158, Ex. 2.2].

Vectors, Tensors and Functions. Vectors and tensors in d -dimensional space as well as quantities that are vector or tensor valued are written in bold characters, e.g., \mathbf{y} and $\boldsymbol{\sigma}$. Their components are represented by y_i and $\sigma_{i,j}$, $i, j = 1, \dots, d$. The elements of sequences or N -tuples of such quantities are denoted \mathbf{y}_i and $\boldsymbol{\sigma}_i$. When subsequences are extracted from sequences, they will generally be denoted by the same symbol as the initial sequence.

Further, the symbols $\partial_i y(\mathbf{x}) := \frac{\partial y}{\partial x_i}(\mathbf{x})$ denote the (weak) spatial derivatives of a function y at \mathbf{x} with respect to the i^{th} component of \mathbf{x} , and $\nabla y(\mathbf{x})$ is understood to be its (weak) first derivative. The order of a multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$ is defined as $|\alpha| := \sum_{k=0}^d \alpha_k$, and the mixed partial α -derivative of the function y is denoted $D^\alpha y$. For quantities y that are time dependent, $\dot{y} := \frac{\partial y}{\partial t}$ and $\ddot{y} := \frac{\partial^2 y}{\partial t^2}$ are the first and second (weak) time derivative, respectively. When a time dependent quantity is vector valued, its (weak) time derivative $\dot{\mathbf{y}}$ will be written in bold characters, even though it may take values in dual spaces of Sobolev spaces instead of \mathbb{R}^d , in order to avoid notational confusion. When functions depend on both time and space, i.e., $y = y(\mathbf{x}, t)$, the partial evaluations $y(\mathbf{x})$ and $y(t)$ mean the respective functions that result from fixing the space or time component to \mathbf{x} or t , respectively. The derivative of an operator S of a single argument is denoted by S' .

Tensors are identified with their representations in higher order matrix form with respect to the fixed basis. As commonly done in the engineering literature, tensor contractions are denoted by $\boldsymbol{\sigma} \cdot \mathbf{y}$ for second and first order tensors $\boldsymbol{\sigma}$ and \mathbf{y} and by $\boldsymbol{\sigma} : \boldsymbol{\varepsilon} := \sum_{i,j=1}^d \sigma_{i,j} \varepsilon_{i,j}$ for two second order tensors $\boldsymbol{\sigma}$, $\boldsymbol{\varepsilon}$. Tensor contractions are assumed to act on the last indices of the representations, i.e., $(C\boldsymbol{\sigma})_{i,j} := \sum_{k,l=1}^d C_{i,j,k,l} \sigma_{k,l}$ for a fourth and second order tensor C , $\boldsymbol{\sigma}$. Details on tensors as multilinear mappings can be found in [32, 35].

Topological and Normed Spaces. In a topological space $(X, \mathcal{O}(X))$, the interior, closure and boundary of a set $A \subset X$ are denoted by \bar{A} , $\text{int}(A)$ and ∂A , respectively. Subsets of X are assumed to be endowed with the subset topology unless otherwise stated. The topology on the extended real numbers is understood as the order topology $\mathcal{O}(\bar{\mathbb{R}})$. Note that the subset topology induced by $\mathcal{O}(\bar{\mathbb{R}})$ on \mathbb{R} coincides with the one induced by the standard absolute value metric. For information on the basics in topology, see [10, 144, 202, 203].

¹A two-dimensional problem description is commonly used as an approximation for either very thick or very thin objects in three dimensions, see Section 2.1.

All normed spaces $(X, \|\cdot\|_X)$ are assumed to be defined over the real numbers, and their topology is the one induced by $\|\cdot\|_X$. On Hilbert spaces $(X, (\cdot, \cdot)_X)$, the norm is always induced by the scalar product $(\cdot, \cdot)_X : X \times X \rightarrow \mathbb{R}$, unless otherwise stated. Note that for a set $A \subset X$, the topology induced by the restriction of the metric induced by $\|\cdot\|_X$ to A coincides with the subset topology. On \mathbb{R}^d , the equivalent p -norms for $1 \leq p \leq \infty$ are denoted $\|\cdot\|_p$, and $\|\cdot\| := \|\cdot\|_2$.

For x in $(X, \|\cdot\|_X)$ and $r > 0$, the ball of radius r with center x is written as $B_X(r, x)$, and $B_X(r) := B_X(r, 0)$. When X is clear from context, the index is omitted. When both $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed spaces, continuous linear mappings from X to Y are denoted $\mathcal{L}(X, Y)$.

The topological dual and bidual spaces of $(X, \|\cdot\|_X)$ are denoted X^* and X^{**} , and the dual pairing is given by $\langle \cdot, \cdot \rangle_X : X^* \times X \rightarrow \mathbb{R}$. When $(H_k)_{k=1}^N$ are Hilbert spaces, the product space $H = \prod_{k=1}^N H_k$ is endowed with the scalar product that is given by the sum of the componentwise scalar products in H_k , which induces the two-norm on the vector of the componentwise norms in H_k on the product space. As usual, the dual space of the product of Banach spaces is identified with the product of the corresponding dual spaces, see Section C.2. The distance function for subsets A, B of a normed space X is defined as

$$\text{dist}(A, B) := \inf_{a \in A, b \in B} \|a - b\|_X,$$

and the same notation is employed for the distance function between points and sets. When $(X, (\cdot, \cdot)_X)$ and $(Y, (\cdot, \cdot)_Y)$ are two Banach spaces and a bounded operator L in $\mathcal{L}(X, Y)$ is given, then the adjoint operator to L is written as $L^* : Y^* \rightarrow X^*$.

For reflexive Banach spaces $(X, \|\cdot\|_X)$, the polar cone and annihilator for subsets $A_1 \subset X$ and $A_2 \subset X^*$ are defined as

$$\begin{aligned} A_1^\circ &:= \{ f \in X^* : \langle f, y \rangle_X \leq 0 \ \forall y \in A_1 \}, & A_1^\perp &:= \{ f \in X^* : \langle f, y \rangle_X = 0 \ \forall y \in A_1 \}, \\ A_2^\circ &:= \{ y \in X : \langle f, y \rangle_X \leq 0 \ \forall f \in A_2 \}, & A_2^\perp &:= \{ y \in X : \langle f, y \rangle_X = 0 \ \forall f \in A_2 \}. \end{aligned}$$

For a convex subset A of X , the radial and tangent cone to A at $y \in A$ are

$$\mathcal{R}_A(y) := \bigcup_{\alpha > 0} \alpha(A - y), \quad \mathcal{T}_A(y) := \overline{\mathcal{R}_A(y)},$$

and for y in A , and f in $\mathcal{T}_A(y)^\circ$, the critical cone to A with respect to (y, f) is

$$\mathcal{C}_A(y, f) := \mathcal{T}_A(y) \cap \{ f \}^\perp.$$

See [119, 136] for introductions to functional analysis.

Measures and Integration. The power set of a set X is denoted by $\mathcal{P}(X)$, and the Borel σ -algebra of a topological space $(X, \mathcal{O}(X))$ is written as $\mathcal{B}(X)$. We write χ_A for the indicator function of a set A .

For $d \in \mathbb{N}$, $d \geq 1$ and $0 \leq s < \infty$, the d -dimensional Lebesgue measure and the s -dimensional Hausdorff measure on $\mathcal{B}(\mathbb{R}^d)$ are denoted by

$$\mathcal{L}^d, \mathcal{H}^s: \mathcal{B}(\mathbb{R}^d) \rightarrow [0, \infty],$$

respectively. The same symbols are used for the corresponding restrictions of these measures to subsets of $\mathcal{B}(\mathbb{R}^d)$. The Hausdorff measure is understood to be defined with an appropriate scaling factor as in, e.g., [65, Def. 2.1]. This ensures that due to the area formula [65, Thm 3.8], its restriction to $\mathcal{B}(\partial\Omega)$ coincides with the surface measure based on the $(d-1)$ -dimensional Lebesgue measure for bounded strong Lipschitz domains. For details, see, e.g., [65, Sec. 3.3], [121, Sec. 6.3], [148, Sec. 2.4] and [110, Sec. 5.3].

When (X, Σ, μ) is an appropriate measure space, the set A is in Σ and $f: X \rightarrow Y$ is a (Lebesgue/Bochner/Bartle)- μ -integrable function for a Banach space $(Y, \|\cdot\|_Y)$, then

$$\int_A f \, d\mu = \int_A f(x) \, d\mu(x)$$

means the μ -integral of f over A . The variable x will be indicated explicitly if the reduced notation admits ambiguity. In the special case where $\Omega \subset \mathbb{R}^d$ is a bounded strong Lipschitz domain and the set $\Gamma \in \mathcal{B}(\partial\Omega)$ is a part of its boundary, integration with respect to the Lebesgue and the Hausdorff measures is abbreviated

$$\int_{\Omega} f(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) \, d\mathcal{L}^d(\mathbf{x}) \quad \int_{\Gamma} g(\mathbf{x}) \, d\mathbf{x} = \int_{\Gamma} g(\mathbf{x}) \, d\mathcal{H}^{d-1}(\mathbf{x}),$$

for \mathcal{L}^d -integrable f and \mathcal{H}^{d-1} -integrable g . In Chapter 3, the meaning of the integrals is crucial, therefore the lengthy notation will be employed for the sake of a clearer presentation. As usual, a property is said to hold μ -almost everywhere (μ -a.e.) if it is violated only on sets of μ -measure zero.

See [7, 30, 65, 66, 200] for real valued measure theory and [54, 55, 58] for vector valued measures. Be advised that the term ‘‘measure’’ is ambiguous in the literature and occasionally refers to either outer measures (countably subadditive set functions) or, in the vector valued case, to finitely additive set functions.

Function Spaces. The set of continuous functions between two given topological spaces $(X, \mathcal{O}(X))$ and $(Y, \mathcal{O}(Y))$ is denoted $C(X, Y)$. When Y is a normed space, then the functions in $C(X, Y)$ are assumed to be bounded, and the usual supremum norm

$$\|u\|_{\infty} = \sup_{x \in X} \|u(x)\|_Y$$

is employed for u in $C(X, Y)$. If $Y = \mathbb{R}$, then $C(X) := C(X, \mathbb{R})$.

Now let $(X, \|\cdot\|_X)$ be a normed space and $A \subset X$. The spaces of all functions from A to \mathbb{R}^d that are (locally) L -Lipschitz for some $L > 0$ are denoted by $C^{0,1}(A, \mathbb{R}^d)$ and $C_{\text{loc}}^{0,1}(A, \mathbb{R}^d)$, respectively, and when $d = 1$, then they are abbreviated $C^{0,1}(A) := C^{0,1}(A, \mathbb{R})$ and $C_{\text{loc}}^{0,1}(A) := C_{\text{loc}}^{0,1}(A, \mathbb{R})$. When O is an open subset of X and $0 \leq k \leq \infty$, then the

symbol $C^k(\Omega)$ means all k -times continuously differentiable functions from O to \mathbb{R} . For bounded and open $O \subset X$, functions that possess k^{th} derivatives on O that are continuous on \overline{O} form the space $C^m(\overline{O})$. Finally, we set $C_0^k(O) := \{f \in C^k(O) : \text{supp}(f) \text{ compact}\}$ where $\text{supp}(f) := \{x \in O : f(x) \neq 0\}$.

When (X, Σ, μ) is a measure space and $1 \leq p \leq \infty$, then $L^p(X, \Sigma; \mu)$ stands for the Lebesgue spaces of equivalence classes of (extended) real valued functions defined on X that are finitely (μ, p) -integrable, or essentially bounded when $p = \infty$, with respect to μ . The standard norms are

$$\|v\|_{L^p(X, \Sigma; \mu)} := \left(\int_X |v|^p \, d\mu \right)^{\frac{1}{p}}, \quad \|v\|_{L^\infty(X, \Sigma; \mu)} := \text{ess sup}_{\mathbf{x} \in X} |v(\mathbf{x})|,$$

for $p < \infty$ and $p = \infty$, respectively. When elements of these spaces are referred to as functions, this is understood to mean the entire class of functions. In the case where $X \in \mathcal{B}(\mathbb{R}^d)$, we abbreviate $L^p(X; \mu) := L^p(X, \mathcal{B}(X); \mu)$.

Now, letting $\Omega \subset \mathbb{R}^d$ be an open set, the Sobolev space $W^{k,p}(\Omega; \mu)$ contains all functions in $L^p(\Omega; \mu)$ with finitely (μ, p) -integrable weak derivatives up to order k in \mathbb{N} , i.e.,

$$W^{k,p}(\Omega; \mu) := \{v \in L^p(\Omega; \mu) : D^\alpha v \in L^p(\Omega; \mu) \, \forall |\alpha| \leq k\}$$

endowed with the usual norms

$$\|v\|_{W^{k,p}(\Omega; \mu)} := \left(\sum_{0 \leq |\alpha| \leq k} \|D^\alpha v\|_{L^p(\Omega; \mu)}^p \right)^{\frac{1}{p}}, \quad \|v\|_{W^{k,\infty}(\Omega; \mu)} := \max_{|\alpha| \leq k} \|D^\alpha v\|_{L^\infty(\Omega; \mu)},$$

and the fractional Sobolev spaces $W^{s,p}(\Omega; \mu)$ contain all functions v in $L^p(\Omega; \mu)$ for which the Sobolev–Slobodeckij norm

$$|v|_{W^{s,p}(\Omega; \mu)} := \left(\|v\|_{L^p(\Omega; \mu)}^p + \int_\Omega \int_\Omega \frac{|v(\mathbf{x}) - v(\mathbf{y})|^p}{|\mathbf{x} - \mathbf{y}|^{d-1+sp}} \, d\mu(\mathbf{x}) \, d\mu(\mathbf{y}) \right)^{\frac{1}{p}} \quad (1.1)$$

is finite. When $\mu = \mathcal{L}^d$, we abbreviate

$$L^p(\Omega) := L^p(\Omega; \mathcal{L}^d), \quad W^{k,p}(\Omega) := W^{k,p}(\Omega; \mathcal{L}^d), \quad W^{s,p}(\Omega) := W^{s,p}(\Omega; \mathcal{L}^d),$$

and we define $W_0^{k,p}(\Omega)$ as the closure of $C_0^\infty(\Omega)$ in $W^{k,p}(\Omega)$. Additionally, when $\Omega \subset \mathbb{R}^d$ is a bounded strong Lipschitz domain with boundary $\partial\Omega$, then $L^p(\partial\Omega) := L^p(\partial\Omega, \mathcal{B}(\partial\Omega); \mathcal{H}^{d-1})$ and the fractional Sobolev space $W^{s,p}(\partial\Omega) := W^{s,p}(\partial\Omega; \mathcal{H}^{d-1})$ is defined in the same manner as $W^{s,p}(\Omega; \mu)$ above when Ω is replaced by $\partial\Omega$ and \mathcal{H}^{d-1} is set as the measure.

For bounded strong Lipschitz domains $\Omega \subset \mathbb{R}^d$, the (surjective) trace operator is

$$\text{tr}: W^{1,p}(\Omega) \rightarrow W^{1-1/p,p}(\partial\Omega),$$

and $W^{1-1/p,p}(\partial\Omega)$ is considered to be endowed with the quotient norm

$$\|v\|_{W^{1-1/p,p}(\partial\Omega)} := \inf_{\substack{w \in W^{1,p}(\Omega) \\ \text{tr}(w) = v}} \|w\|_{W^{1,p}(\Omega)}, \quad (1.2)$$

which is equivalent to the Sobolev–Slobodeckij norm (1.1), see Lemma A.2.3. Refer to Section A.2 for a short summary on the trace operator and its properties. Also, note that

$$W_0^{1,p}(\Omega) = \left\{ y \in W^{1,p}(\Omega) : \operatorname{tr}(y) = 0 \text{ } \mathcal{H}^{d-1}\text{- a.e. on } \partial\Omega \right\},$$

see [148, Thm. 2.4.10]. In the same spirit, when $\Gamma_D \subset \partial\Omega$ is a Dirichlet boundary part, Dirichlet values are incorporated into the space via

$$W_D^{1,p}(\Omega) := \left\{ y \in W^{1,p}(\Omega) : \operatorname{tr}(y) = 0 \text{ } \mathcal{H}^{d-1}\text{- a.e. on } \Gamma_D \right\}.$$

For the index $p = 2$, k in \mathbb{N} and s in $(0, 1)$, the following standard abbreviations are applied

$$H^k(\Omega) := W^{k,2}(\Omega), \quad H^s(\partial\Omega) := W^{s,2}(\partial\Omega).$$

For \mathbb{R}^d valued functions whose component functions belong to any of the spaces X above, we define the product spaces $\mathbf{X} := \prod_{i=1}^d X$, i.e.,

$$\mathbf{L}^p(\Omega) := \prod_{i=1}^d L^p(\Omega), \quad \mathbf{W}^{k,p}(\Omega) := \prod_{i=1}^d W^{k,p}(\Omega), \quad \mathbf{W}^{s,p}(\partial\Omega) := \prod_{i=1}^d W^{s,p}(\partial\Omega),$$

etc. The product spaces are endowed with norms that result from application of any of the equivalent norms on \mathbb{R}^d to the vector of norms on the component functions. With a slight abuse of notation, the trace operator is evaluated componentwise when applicable and the symbol tr is used in the vector valued case as well.

For Lebesgue spaces and fractional Sobolev spaces on open subsets Γ of the boundary, see, e.g., [110, Sec. 5.3] and [148, Sec. 2.4]. Cf. [3, 11, 77, 121, 148] for general information on Sobolev spaces and [110] for a detailed analysis of various trace operators in the context of linear elasticity.

Finally, for a Banach space X , the symbol $L^p(0, T; X)$ means the Bochner-Sobolev space of (\mathcal{L}, p) -Bochner-integrable functions on $[0, T]$ with values in X equipped with the norms

$$\|y\|_{L^p(0,T;X)} := \left(\int_0^T \|y(t)\|_X^p \, d\mathcal{L}(t) \right)^{\frac{1}{p}}, \quad \|y\|_{L^\infty(0,T;X)} := \operatorname{ess\,sup}_{t \in [0,T]} \|y(t)\|_X.$$

Based on the standard Gelfand triple $H^1(\Omega) \subset L^2(\Omega) \subset H^1(\Omega)^*$, we consider weak time derivatives using the space

$$\mathbf{W}(0, T) := \left\{ y \in L^2(0, T; \mathbf{H}^1(\Omega)) : \dot{y} \in L^2(0, T; \mathbf{H}^1(\Omega)^*) \right\},$$

$$\|y\|_{\mathbf{W}(0,T)} := \|y\|_{L^2(0,T;\mathbf{H}^1(\Omega))} + \|\dot{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega)^*)}.$$

See [207, Sec. 23.6 ff.] for the treatment of weak time derivatives and [71, 204] for Bochner-Lebesgue and Bochner-Sobolev spaces.

Miscellaneous. The set of real, orthogonal $d \times d$ -matrices with determinant equal to one, i.e., rotations in d -dimensional space, is denoted $\operatorname{SO}(d)$, and \mathbb{S}^d means the set of symmetric second order tensors. We write $\mathcal{S}_1 := \{ \mathbf{y} \in \mathbb{R}^d : \|\mathbf{y}\| = 1 \}$ for the unit sphere in \mathbb{R}^d with respect to the euclidean two-norm. The Bachmann-Landau “O”-symbol is used when describing the asymptotic behavior of terms. Finally, $\operatorname{Tr}(\cdot)$ means the trace of a $d \times d$ -matrix.

Chapter 2

Modeling of Contact Problems

The literature on contact problems spans a wide range of models and settings, including, e.g., punch problems [34, 105, 115] and various combinations of (quasi-)static [48, 72, 107], dynamic [112, 126, 134, 141, 206], adhesive [40], thermal [9, 48], plastic or damaging [37, 39, 178] and frictionless [4, 112] or frictional effects with and without wear [103, 104, 134] in viscoelasticity. In this thesis, we consider the contact between a single linearly viscoelastic body with Kelvin-Voigt-type response and a rigid obstacle in a frictionless and time dependent setting. This chapter provides an introductory overview of the basic concepts of solid mechanics, viscoelasticity and contact dynamics required for the mathematical modeling of these problems. All quantities are assumed to be sufficiently regular in order for the expressions to be meaningful, for now. References to the literature that treat the contents of this chapter rigorously and in greater detail will be given along the way. A detailed account of the modeling of contact problems can be found in [110].

The assumptions made in this chapter are limited to the material's behavior and some minimal regularity requirements for the weak formulation of the problem. Note that while the results are formulated for one-body contact problems, the techniques of this thesis can be applied in settings involving the contact between two viscoelastic bodies as well, and the main results are expected to carry over. See, e.g., [61, 163, 185, 198] for the modeling and optimization of two-body contact problems.

Structure. Section 2.1 provides a brief overview of the basic concepts of solid mechanics for one-body contact problems and linear viscoelasticity. The equilibrium equations of the viscoelastic body's unconstrained movement under load are stated. In Section 2.2, the linear non-penetration condition is introduced with a short discussion of its properties that are relevant in the following analysis. Finally, Section 2.3 contains the one-body contact problem in both strong and weak form.

2.1 Linear Viscoelasticity

In solid mechanics, objects are modeled as continuous media with particles of infinitesimal volume and a “continuous” distribution of their mass in space, disregarding the material’s properties on the level of particles. See, e.g., the monographs [49, 97, 173] for introductions into the topic. Out of the wide range of physical properties that are considered in solid mechanics, the theory of viscoelasticity focuses on the body’s deformation under load and the connection between external loads and internal stresses, so-called constitutive laws. Viscoelasticity is a long established field in mathematics and engineering, and there is extensive literature on the matter in both finite and infinitesimal strain theory, see the monographs [17, 45, 74, 79, 80, 177, 180, 190, 197]. This section addresses the basic concepts for the description of the movement and deformation of a single solid body on a given time interval $I = [0, T]$ with $T > 0$ and provides a brief introduction into viscoelasticity with infinitesimal strain and linear strain-stress-relationships.

Reference Configuration and Mass. The body is assumed to be in its undeformed shape at initial time $t = 0$ and the space occupied by the body is described by the set of points in $\bar{\Omega}$ for an open, nonempty set $\Omega \subset \mathbb{R}^d$ in d -dimensional space with $d \in \{2, 3\}$. The set $\bar{\Omega}$ is called the *reference configuration* and on its boundary, we identify three disjoint subsets $\Gamma_D, \Gamma_N, \Gamma_C \subset \partial\Omega$ for Dirichlet, Neumann and contact conditions — where the body is clamped, exposed to boundary forces or may experience contact, respectively — such that $\bar{\Gamma}_D \cup \bar{\Gamma}_N \cup \bar{\Gamma}_C = \partial\Omega$.

The *mass distribution* of the body is given by the mass density function

$$\rho: \bar{\Omega} \rightarrow (0, \infty),$$

which gives mass per unit volume in the reference configuration. Accordingly, when $A \subset \Omega$ is a Lebesgue measurable subset of the reference configuration $\bar{\Omega}$, its mass is obtained by evaluating

$$\int_A \rho(\mathbf{x}) \, d\mathbf{x}.$$

The immovable obstacle is identified with another open, nonempty set $\mathcal{O} \subset \mathbb{R}^d$, which may be unbounded and not connected. No further modeling of the obstacle is required for now because it will be incorporated in the boundary conditions in Section 2.2. In order to maintain consistency of the model, the two domains are assumed to have empty intersection.

Deformation and Displacement. When the viscoelastic body moves, the corresponding transformation of its mass to its *deformed configuration* $\bar{\Omega}^{\varphi, t}$ at time t is described by the *deformation vector field*

$$\varphi: \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}^d \quad \text{with} \quad \bar{\Omega}^{\varphi, t} := \varphi(\bar{\Omega}, t).$$

The mapping is assumed to be orientation preserving and injective up to the boundary $\partial\Omega$ in order to be physically meaningful. Its first derivative reads as

$$\nabla\varphi(\mathbf{x}, t) = \begin{pmatrix} \frac{\partial\varphi_1}{\partial x_1}(\mathbf{x}, t) & \cdots & \frac{\partial\varphi_1}{\partial x_d}(\mathbf{x}, t) \\ \vdots & \ddots & \vdots \\ \frac{\partial\varphi_d}{\partial x_1}(\mathbf{x}, t) & \cdots & \frac{\partial\varphi_d}{\partial x_d}(\mathbf{x}, t) \end{pmatrix}.$$

Remark 2.1.1 *Regardless of the nomenclature, translations and rotations (rigid body movements), which leave the shape of the body unchanged, are generally described by non-zero deformation fields.*

Instead of using the new absolute position of the body to describe its properties, it will often be more convenient to work with its new relative position and introduce the so-called *displacement field*

$$\mathbf{y}: \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}^d, \quad \mathbf{y}(\mathbf{x}, t) := \varphi(\mathbf{x}, t) - \mathbf{x},$$

with the spatial derivative

$$\nabla\mathbf{y}(\mathbf{x}, t) = \nabla\varphi(\mathbf{x}, t) - \mathbf{I}.$$

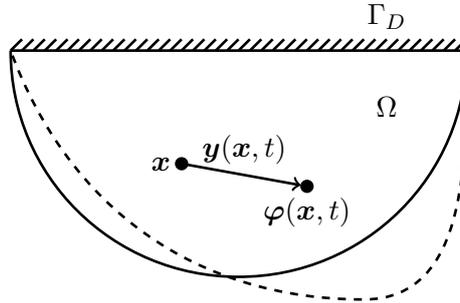


Figure 2.1: Deformation and displacement. The reference configuration is semicircular and clamped on the planar part of its boundary.

Forces and Strain. The external loads on the body are composed of *volume* and *boundary forces*, which are assumed to be independent of the body's deformation (dead loads) and can therefore be represented by their corresponding force densities in the reference configuration

$$\mathbf{f}_\Omega: \Omega \times [0, T] \rightarrow \mathbb{R}^d, \quad \mathbf{f}_{\Gamma_N}: \Gamma_N \times [0, T] \rightarrow \mathbb{R}^d,$$

which give force per unit volume and per unit area in the reference configuration, respectively.

When external forces or initial velocities deform the body, we can identify two types of resulting movements, see [17, Sec. 2.1.1].

1. **Rigid body movements:** Any movement that leaves the distance between all points in the domain unchanged, i.e.,

$$|\varphi(\mathbf{x}, t) - \varphi(\tilde{\mathbf{x}}, t)| = |\mathbf{x} - \tilde{\mathbf{x}}| \quad \forall \mathbf{x}, \tilde{\mathbf{x}} \in \bar{\Omega}, t \in [0, T].$$

These coincide with the movements that are the sum of a translation and a rotation [79, P. 49], i.e., continuous deformations of the form

$$\varphi(\mathbf{x}, t) = \mathbf{a}(t) + \mathbf{Q}(t)\mathbf{x}, \quad \mathbf{a}(t) \in \mathbb{R}^d, \mathbf{Q}(t) \in \text{SO}(d, \mathbb{R}).$$

2. **Distorting movements:** Any movements that are not rigid body movements. The change in distance between points in the body means that sections of the body can be compressed while others might be stretched.

The latter are crucial in (visco-)elasticity theory, because the *strain* introduced by the distortions causes the viscoelastic resistances in the body. To motivate the introduction of the strain as a measure of the distortion, i.e., the change in angles and lengths, we follow the exposition in [192, P. 25] and consider two points $\mathbf{x}, \tilde{\mathbf{x}} \in \Omega$ that are moved to $\varphi(\mathbf{x}, t)$ and $\varphi(\tilde{\mathbf{x}}, t)$, respectively. Their distances in the reference and the deformed configuration are $\delta\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}$ and $\delta\varphi(t) = \varphi(\tilde{\mathbf{x}}, t) - \varphi(\mathbf{x}, t)$. From Taylor's theorem, we obtain that

$$\delta\varphi(t) = \nabla\varphi(\mathbf{x}, t)\delta\mathbf{x} + o(\|\delta\mathbf{x}\|).$$

Temporarily suppressing the dependencies on space and time, the change in the squared distances in the reference and the deformed configuration is therefore

$$\begin{aligned} \|\delta\varphi\|^2 - \|\delta\mathbf{x}\|^2 &= \delta\varphi^T\delta\varphi - \delta\mathbf{x}^T\delta\mathbf{x} \\ &= \delta\mathbf{x}^T\nabla\varphi^T\nabla\varphi\delta\mathbf{x} - \delta\mathbf{x}^T\delta\mathbf{x} + o(\|\delta\mathbf{x}\|^2) \\ &= \delta\mathbf{x}^T(\nabla\varphi^T\nabla\varphi - \mathbf{I})\delta\mathbf{x} + o(\|\delta\mathbf{x}\|^2) \\ &= \delta\mathbf{x}^T(\nabla\mathbf{y} + \nabla\mathbf{y}^T + \nabla\mathbf{y}^T\nabla\mathbf{y})\delta\mathbf{x} + o(\|\delta\mathbf{x}\|^2), \end{aligned}$$

which motivates the introduction of the *finite strain tensor*

$$\mathbf{E} := \frac{1}{2}(\nabla\varphi^T\nabla\varphi - \mathbf{I}) = \frac{1}{2}(\nabla\mathbf{y} + \nabla\mathbf{y}^T + \nabla\mathbf{y}^T\nabla\mathbf{y})$$

as a local measure of the body's distortion compared to reference configuration. Clearly, the strain tensor is nonlinear in the displacement \mathbf{y} and generally non-constant in time and space. Additionally, the tensor \mathbf{E} vanishes if and only if the deformation is a rigid body motion [79, P. 45].

When the displacements \mathbf{y} and their derivatives $\nabla\mathbf{y}$ are (infinitesimally) small, the first order terms in \mathbf{E} dominate those of second order, and the strain in the body may be approximated by its linearization at the vanishing displacement, i.e., by the *infinitesimal strain tensor*

$$\boldsymbol{\epsilon} := \frac{1}{2}(\nabla\mathbf{y} + \nabla\mathbf{y}^T), \quad (2.1)$$

which is clearly linear, symmetric and approximates \mathbf{E} up to terms in $o(\|\nabla\mathbf{y}\|^2)$. Note that $\boldsymbol{\epsilon}$ is generally non-zero when φ is a rigid body movement with non-vanishing rotational part.

Stress and Equilibrium of Forces. The internal resistance that counteracts the distortion of a deformed viscoelastic body and acts towards returning the body to its resting position in the original shape is known as the *stress*. While the stresses of purely elastic materials are the result of the current deformation and thus depend only on the deformation gradient at the time, the viscous stresses are a response to the change in the deformation during loading and unloading and depend on the gradient of the velocity.

The existence of stresses that balance the forces and moments in the deformed configuration is postulated in the fundamental Stress Principle of Cauchy and Euler [45, Axiom 2.2-1] for static, elastic bodies in the form of a stress vector

$$\mathbf{S}^\varphi: \overline{\Omega}^\varphi \times \mathcal{S}_1 \rightarrow \mathbb{R}^d,$$

where $\mathbf{S}^\varphi(\mathbf{x}^\varphi, \mathbf{n})$ gives the force per unit area, i.e., measured in units of pressure, on infinitesimal surfaces normal to a vector \mathbf{n} in \mathcal{S}_1 and at a point \mathbf{x}^φ in the deformed configuration $\overline{\Omega}^\varphi$ as a reaction to the load on the body. As it turns out, the relation between the stresses and the external forces is of divergence form.

Theorem 2.1.2 (Cauchy's Theorem [45, Thm. 2.3-1]) *Consider an elastic body in its deformed configuration $\overline{\Omega}^\varphi$ under force densities $\mathbf{f}_\Omega^\varphi: \Omega^\varphi \rightarrow \mathbb{R}^d$ and $\mathbf{f}_{\Gamma_N}^\varphi: \Gamma_N^\varphi \rightarrow \mathbb{R}^d$. Assuming sufficient regularity of the data, there exists a tensor field*

$$\mathbf{T}^\varphi: \overline{\Omega}^\varphi \rightarrow \mathbb{S}^d$$

such that

$$\mathbf{S}^\varphi(\mathbf{x}^\varphi, \mathbf{n}) = \mathbf{T}^\varphi(\mathbf{x}^\varphi)\mathbf{n} \quad \text{for all } (\mathbf{x}^\varphi, \mathbf{n}) \text{ in } \overline{\Omega}^\varphi \times \mathcal{S}_1$$

and

$$\begin{aligned} -\nabla^\varphi \mathbf{T}^\varphi(\mathbf{x}^\varphi) &= \mathbf{f}_\Omega^\varphi(\mathbf{x}^\varphi) && \text{for all } \mathbf{x}^\varphi \text{ in } \overline{\Omega}^\varphi, \\ \mathbf{T}^\varphi(\mathbf{x}^\varphi)\boldsymbol{\nu}^\varphi &= \mathbf{f}_{\Gamma_N}^\varphi(\mathbf{x}^\varphi) && \text{for all } \mathbf{x}^\varphi \text{ in } \Gamma_N^\varphi, \end{aligned}$$

where $\boldsymbol{\nu}^\varphi$ is the unit outer normal vector on the deformed Neumann boundary Γ_N^φ .

The differential equations obtained for the *Cauchy stress tensor* \mathbf{T}^φ are stated on the deformed configuration and in terms of the unknown $\mathbf{x}^\varphi = \boldsymbol{\varphi}(\mathbf{x})$ and is therefore commonly transformed to obtain a problem on the reference configuration using the *second Piola-Kirchhoff stress tensor* $\boldsymbol{\sigma}: \overline{\Omega} \rightarrow \mathbb{S}^d$ (see [45, P. 72]) which is given by

$$\boldsymbol{\sigma}(\mathbf{x}) = \det(\nabla \boldsymbol{\varphi}(\mathbf{x})) \boldsymbol{\varphi}(\mathbf{x})^{-1} \mathbf{T}^\varphi(\boldsymbol{\varphi}(\mathbf{x})) \nabla \boldsymbol{\varphi}(\mathbf{x})^{-1}.$$

In infinitesimal strain theory, the dependencies are linearized by assuming that the second order error introduced by taking the quantities on the reference configuration instead of their transformed counterparts and in $\boldsymbol{\varphi} \approx \mathbf{I}$ can be disregarded. Consequently, the linearized system in the time dependent setting that includes the acceleration terms to account for the dynamics requires the identity

$$\mathbf{S}(\mathbf{x}, t, \mathbf{n}) = \boldsymbol{\sigma}(\mathbf{x}, t) \cdot \mathbf{n}$$

for the time dependent stress vectors and tensors and the well-known equilibrium of force densities condition in linear elasticity:

$$\begin{aligned} \rho \ddot{\mathbf{y}} - \operatorname{div} \boldsymbol{\sigma} &= \mathbf{f}_\Omega && \text{in } \Omega \times [0, T], \\ \boldsymbol{\sigma} \cdot \boldsymbol{\nu} &= \mathbf{f}_{\Gamma_N} && \text{on } \Gamma_N \times [0, T], \end{aligned}$$

where $\boldsymbol{\nu}$ denotes the geometric normal on the boundary and the divergence operator is again understood as acting row-wise on the $d \times d$ matrix representing $\boldsymbol{\sigma}$.

See [45, Sec. 2.4-2.6] and [62, 177] for more details.

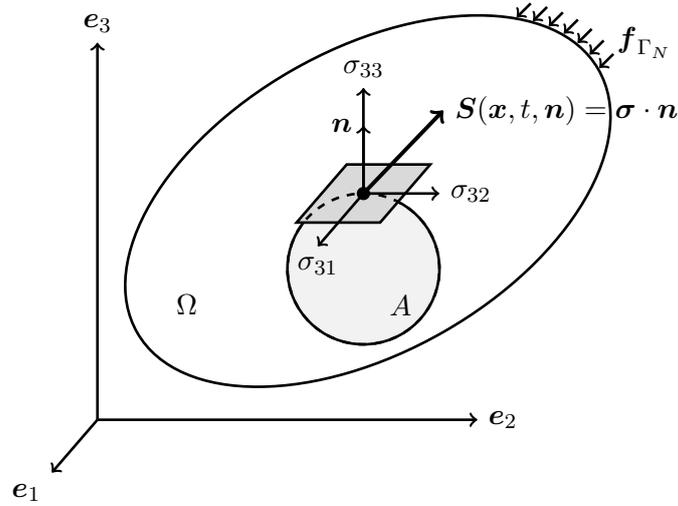


Figure 2.2: Stress on the surface of a subdomain A of the domain Ω .

Stress-Strain-Relation. What remains to be established is a connection between the measure of distortion (strains and change in strain) and the corresponding response of the material (stresses), which clearly needs to depend on the material properties. Different concepts for the combination of elastic and viscous properties are generally needed to model different materials, and we restrict our considerations to *Kelvin-Voigt*-type materials, where the viscoelastic response is modeled in a parallel manner. The total stress consequently decomposes into the sum of the elastic stresses $\boldsymbol{\sigma}_E$ and viscous stresses $\boldsymbol{\sigma}_V$, i.e.,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_E + \boldsymbol{\sigma}_V.$$

In one dimension, this approach corresponds to the simple model of a spring and a dashpot being connected in parallel.

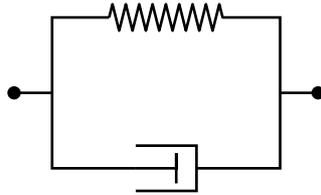


Figure 2.3: One dimensional model for Kelvin-Voigt-type materials.

The equations that link strain and strain rate to the elastic and viscous stresses in the body are referred to as the constitutive equations of the model. While there are plenty of constitutive models, cf. [45, Sec. 4.10], the considerations in this thesis are limited to materials whose stress is linearly dependent on the strain, specifically, where the stress-strain-relationship can be modeled with two fourth order tensor fields, the *elasticity* and the *viscosity tensor fields* $\mathbf{C}, \mathbf{V}: \bar{\Omega} \rightarrow \mathcal{L}(\mathbb{S}^d, \mathbb{S}^d)$, with

$$\boldsymbol{\sigma}_E(\mathbf{x}, t) := \mathbf{C}(\mathbf{x})\boldsymbol{\epsilon}(\mathbf{x}, t), \quad (\sigma_E(\mathbf{x}, t))_{i,j} = \sum_{k,l=1}^d C_{i,j,k,l}(\mathbf{x})\epsilon_{k,l}(\mathbf{x}, t), \quad (2.3a)$$

$$\boldsymbol{\sigma}_V(\mathbf{x}, t) := \mathbf{V}(\mathbf{x})\dot{\boldsymbol{\epsilon}}(\mathbf{x}, t), \quad (\sigma_V(\mathbf{x}, t))_{i,j} = \sum_{k,l=1}^d V_{i,j,k,l}(\mathbf{x})\dot{\epsilon}_{k,l}(\mathbf{x}, t), \quad (2.3b)$$

and consequently

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \mathbf{C}(\mathbf{x})\boldsymbol{\epsilon}(\mathbf{x}, t) + \mathbf{V}(\mathbf{x})\dot{\boldsymbol{\epsilon}}(\mathbf{x}, t) \quad \forall \mathbf{x} \in \bar{\Omega}, t \in [0, T]. \quad (2.4)$$

In elasticity, equation (2.3a) is known as *Hooke's law*, while (2.3b) is known as *Newton's law* of viscosity. The tensor fields are assumed to be symmetric, uniformly bounded and uniformly positive definite in the sense that there exist constants $c_0, C_0, v_0, V_0 > 0$ with

$$C_{i,j,k,l} = C_{k,l,i,j} = C_{i,j,l,k}, \quad V_{i,j,k,l} = V_{k,l,i,j} = V_{i,j,l,k} \quad \forall i, j, k, l \in \{1, \dots, d\},$$

$$\mathbf{C}\mathbf{v} : \mathbf{w} \leq c_0 \|\mathbf{v}\|_{\mathbb{S}^d} \|\mathbf{w}\|_{\mathbb{S}^d}, \quad \mathbf{V}\mathbf{v} : \mathbf{w} \leq v_0 \|\mathbf{v}\|_{\mathbb{S}^d} \|\mathbf{w}\|_{\mathbb{S}^d} \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{S}^d, \quad (2.5a)$$

$$\mathbf{C}\mathbf{v} : \mathbf{v} \geq C_0 \|\mathbf{v}\|_{\mathbb{S}^d}^2, \quad \mathbf{V}\mathbf{v} : \mathbf{v} \geq V_0 \|\mathbf{v}\|_{\mathbb{S}^d}^2 \quad \forall \mathbf{v} \in \mathbb{S}^d \quad (2.5b)$$

for one of the equivalent norms on the matrix representations of the symmetric second order tensors \mathbf{v}, \mathbf{w} and everywhere in $\bar{\Omega} \times [0, T]$. The representations of the tensors \mathbf{C} and \mathbf{V} with respect to the fixed basis generally include d^4 components. However, combining the symmetry requirements on \mathbf{C} and \mathbf{V} shows that for either tensor, only 6 or 21 of their components, in two and three dimensions, respectively, are independent. In the important case of isotropic and homogeneous materials, where the stress is additionally independent of the orientation and of the location in the solid, the independent parameters even reduce to two positive parameters for each of the tensors and the spatial dependence of \mathbf{C} and \mathbf{V} can be dropped, see [45, Sec. 3.7-3.8], simplifying (2.3a)–(2.3b) to read as

$$\boldsymbol{\sigma}_E(\mathbf{x}, t) = \frac{E\nu_{\text{poi}}}{(1 + \nu_{\text{poi}})(1 - 2\nu_{\text{poi}})} \text{Tr}(\boldsymbol{\epsilon}(\mathbf{x}, t)) \mathbf{I} + \frac{E}{1 + \nu_{\text{poi}}} \boldsymbol{\epsilon}(\mathbf{x}, t),$$

$$\boldsymbol{\sigma}_V(\mathbf{x}, t) = \left(\mu_{\text{bulk}} - \frac{2}{3}\mu_{\text{shear}} \right) \text{Tr}(\dot{\boldsymbol{\epsilon}}(\mathbf{x}, t)) \mathbf{I} + 2\mu_{\text{shear}} \dot{\boldsymbol{\epsilon}}(\mathbf{x}, t).$$

The chosen notation is general consensus in the literature on viscoelasticity. The parameter $E > 0$ is called *Young's modulus* and $\nu_{\text{poi}} > 0$ is called *Poisson's ratio*. They give a linear correspondence for the change in stress and the change in width in a body, respectively, when the body is stretched or compressed along an axis. The parameters $\mu_{\text{bulk}}, \mu_{\text{shear}} > 0$ are the *bulk* and *shear viscosity* of the material.

Plane Strain. Problems in three-dimensional elasticity that possess additional structure in the geometry of the configuration can oftentimes be reduced to two-dimensional planar models, so-called plane strain and plane stress situations. In the numerics of Chapter 6, we will consider plane strains, which occur when the displacements \mathbf{y} are constant in one spatial direction and when there is no displacement in the same direction, i.e., the displacements can be written as

$$\mathbf{y}(\mathbf{x}, t) = \begin{pmatrix} y_1(x_1, x_2, t) \\ y_2(x_1, x_2, t) \\ 0 \end{pmatrix},$$

assuming the basis has been chosen sensibly. The strain tensor $\boldsymbol{\epsilon}(\mathbf{x}, t)$ corresponding to these displacements therefore reduces to a tensor acting in the x_1 - x_2 -plane, i.e.,

$$\boldsymbol{\epsilon}(\mathbf{x}, t) = \begin{pmatrix} \epsilon_{11}(x_1, x_2, t) & \epsilon_{12}(x_1, x_2, t) & 0 \\ \epsilon_{21}(x_1, x_2, t) & \epsilon_{22}(x_1, x_2, t) & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

For isotropic, homogeneous materials, the corresponding stress tensor accordingly takes the form

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \begin{pmatrix} \sigma_{11}(x_1, x_2, t) & \sigma_{12}(x_1, x_2, t) & 0 \\ \sigma_{21}(x_1, x_2, t) & \sigma_{22}(x_1, x_2, t) & 0 \\ 0 & 0 & \sigma_{33}(x_1, x_2, t) \end{pmatrix}.$$

Note that the component $\sigma_{33}(x_1, x_2, t)$ is generally non-zero and ensures that there is no deformation in the direction \mathbf{e}_3 .

Plane strain situations typically correspond to configurations where the geometry of the body corresponds to that of one or multiple coaxial cylinders on which the forces act perpendicularly to the axis \mathbf{e}_3 and the x_1 - x_2 -cutting-planes of the obstacle are independent of x_3 . If the cylinder is either very long or it is clamped at its bases, a plane strain situation arises.

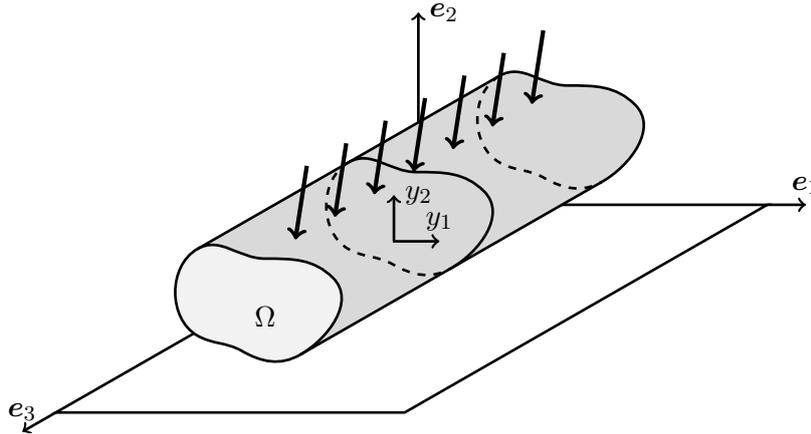


Figure 2.4: Typical plane strain configuration.

Similarly, bodies whose geometry resembles that of a thin plate, with one dimension being significantly smaller than the remaining two, can often be found to be in a state of plane stress, where the stress in the direction of the small expansion vanishes.

More information on the well documented plane strain and plane stress models can be found in most introductory books on the matter of solid mechanics and linear elasticity, e.g., [180, Par. 5.66], [197, Sec. 9.1] or [190, Sec. C.VII].

Unconstrained Viscoelastic Problem. By adding Dirichlet boundary conditions on $\Gamma_D \subset \partial\Omega$ and initial conditions to the equilibrium conditions (2.2), we arrive at the strong form of the unconstrained, dynamic problem of viscoelasticity:

$$\begin{aligned} \rho \ddot{\mathbf{y}} - \operatorname{div} \boldsymbol{\sigma} &= \mathbf{f}_\Omega && \text{on } \Omega \times [0, T] \\ \boldsymbol{\sigma} \cdot \boldsymbol{\nu} &= \mathbf{f}_{\Gamma_N} && \text{on } \Gamma_N \times [0, T] \\ \mathbf{y} &= 0 && \text{on } \Gamma_D \times [0, T] \\ \mathbf{y}(0) &= \mathbf{y}_{\text{ini}}, \dot{\mathbf{y}}(0) = \mathbf{v}_{\text{ini}} && \text{on } \Omega. \end{aligned}$$

The obstacle will be incorporated into the model by an additional set of boundary conditions on the boundary section Γ_C .

2.2 Contact Constraints

Contact constraints implement the physical restriction that no two objects can occupy the same space at the same time. Aside from enforcing the non-penetration of masses, the constraints also introduce the contact stresses on the active contact boundary sections that are required to maintain an equilibrium of forces into the unconstrained model. The contact constraint that we will consider is a version of Signorini's linearized condition that was first presented in 1933 ([175]) in the static one-body context. The particular model was addressed in [61], cf. the overviews in [52, 112]. We present its concept and briefly discuss some of its properties. For additional information on contact problems and their modeling, refer to [62] and notably [110, Sec. 2.2].

2.2.1 Non-Penetration and Contact Stresses

A priori, there is no information on which of the parts of the boundaries of the viscoelastic body and the obstacle come into contact at what time. The foundation of the contact condition is the assumption that the boundary section of the body that may come into contact with the obstacle is confined to subsets of the contact boundary Γ_C .

Non-Penetration of Masses. Assume the existence of a smooth, injective *contact mapping* $\Phi: \Gamma_C \rightarrow \partial\mathcal{O}$ that maps every point on the contact boundary to an associated point on the boundary of the obstacle. Using the contact mapping, the *contact normal* is defined as

$$\boldsymbol{\nu}_\Phi: \Gamma_C \rightarrow \mathbb{R}^d, \quad \boldsymbol{\nu}_\Phi(\mathbf{x}) := \begin{cases} \frac{\Phi(\mathbf{x}) - \mathbf{x}}{\|\Phi(\mathbf{x}) - \mathbf{x}\|}, & \mathbf{x} \neq \Phi(\mathbf{x}) \\ \boldsymbol{\nu}(\mathbf{x}), & \mathbf{x} = \Phi(\mathbf{x}) \end{cases},$$

and the *initial gap* is defined as

$$\Psi: \Gamma_C \rightarrow \overline{\mathbb{R}}, \quad \Psi(\mathbf{x}) := \|\mathbf{x} - \Phi(\mathbf{x})\|.$$

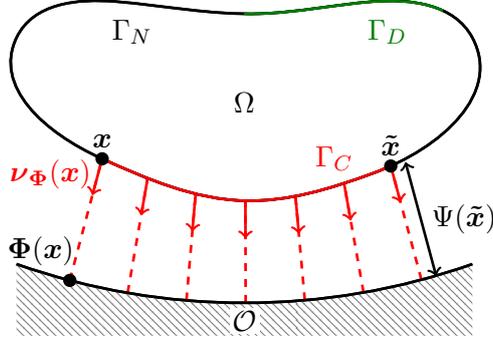


Figure 2.5: One-body reference configuration with contact mapping and contact normal.

The linearized non-penetration condition for a displacement $\mathbf{y}: \overline{\Omega} \rightarrow \mathbb{R}^d$ then requires that

$$\mathbf{y}(\mathbf{x}, t) \cdot \boldsymbol{\nu}_{\Phi}(\mathbf{x}) \leq \Psi(\mathbf{x}) \quad \forall \mathbf{x} \in \Gamma_C, t \in [0, T], \quad (2.8)$$

which means that the displacement on the contact boundary in direction of the contact normal is bounded by the initial gap.

Note that the contact mapping Φ did not enter the problem description explicitly and can be interpreted as “user provided input” that supplies the contact normal $\boldsymbol{\nu}_{\Phi}$ and the gap function Ψ for the problem description. Providing a contact normal and gap function is equally sufficient for the formulation of the one-body constraint. Nevertheless, the contact mapping was included in the description of the constraint here because of its role in the two-body case, where the contact constraint is based on the relative displacement of the two bodies and therefore mentions the contact mapping explicitly. The behavior of the constrained problem is determined by the choice of these quantities, hence, they should incorporate as much information on the geometry of the problem as possible, which will be discussed in the following section (Section 2.2.2).

Contact Stresses. As well as the contact patches, the required contact forces are unknown. They are described by the stresses they induce in the viscoelastic body. We refer to

$$\boldsymbol{\sigma}_{\nu}: \partial\Omega \times [0, T] \rightarrow \mathbb{R}^d, \quad \boldsymbol{\sigma}_{\nu}(\mathbf{x}, t) := \boldsymbol{\sigma}(\mathbf{x}, t) \cdot \boldsymbol{\nu}(\mathbf{x})$$

as the *boundary stresses*. Recall that $\boldsymbol{\sigma}_{\nu}(\mathbf{x}, t)$ gives the boundary force density (a pressure) at $\mathbf{x} \in \partial\Omega$ and $t \in [0, T]$. On the contact boundary, their contact normal parts

$$\boldsymbol{\sigma}_{\nu_{\Phi}}: \Gamma_C \times [0, T] \rightarrow \mathbb{R}^d, \quad \boldsymbol{\sigma}_{\nu_{\Phi}}(\mathbf{x}, t) := (\boldsymbol{\sigma}_{\nu}(\mathbf{x}, t) \cdot \boldsymbol{\nu}_{\Phi}(\mathbf{x})) \boldsymbol{\nu}_{\Phi}(\mathbf{x})$$

are called the *contact stresses* and the corresponding *tangent stresses* are given by

$$\boldsymbol{\sigma}_t: \Gamma_C \times [0, T] \rightarrow \mathbb{R}^d, \quad \boldsymbol{\sigma}_t(\mathbf{x}, t) := \boldsymbol{\sigma}_{\nu}(\mathbf{x}, t) - \boldsymbol{\sigma}_{\nu_{\Phi}}(\mathbf{x}, t).$$

In the absence of adhesive forces, only compressive forces can be transmitted between the objects at the times of contact. Accordingly, we obtain the inequality constraint

$$\boldsymbol{\sigma}_{\boldsymbol{\nu}_{\Phi}}(\mathbf{x}, t) \cdot \boldsymbol{\nu}_{\Phi}(\mathbf{x}) \leq 0 \quad \forall \mathbf{x} \in \Gamma_C, t \in [0, T]$$

for the contact stresses on the contact boundary. Additionally, contact stresses may only be transmitted when contact is established, i.e., we require the *complementarity condition*

$$(\boldsymbol{\sigma}_{\boldsymbol{\nu}_{\Phi}}(\mathbf{x}, t) \cdot \boldsymbol{\nu}_{\Phi}(\mathbf{x})) (\mathbf{y}(\mathbf{x}, t) \cdot \boldsymbol{\nu}_{\Phi}(\mathbf{x}) - \Psi(\mathbf{x})) = 0 \quad \forall \mathbf{x} \in \Gamma_C, t \in [0, T].$$

In our frictionless setting, tangential stresses vanish, hence the final requirement on the stresses reads as

$$\boldsymbol{\sigma}_t(\mathbf{x}, t) = 0 \quad \forall \mathbf{x} \in \Gamma_C, t \in [0, T].$$

2.2.2 Discussion

In comparison to a physically accurate non-penetration requirement, such as, e.g., the complex set-based condition

$$\mathbf{x} + \mathbf{y}(\mathbf{x}, t) \notin \mathcal{O} \quad \forall \mathbf{x} \in \bar{\Omega}, t \in [0, T], \quad (2.9)$$

the linear constraint (2.8) is significantly easier to handle mathematically. An obvious but natural drawback of the linear condition is the lack of higher order information, which generally makes the linear constraint (2.8) a viable replacement of (2.9) for small deformations only because otherwise, non-physical movements can appear feasible to the linear constraint.

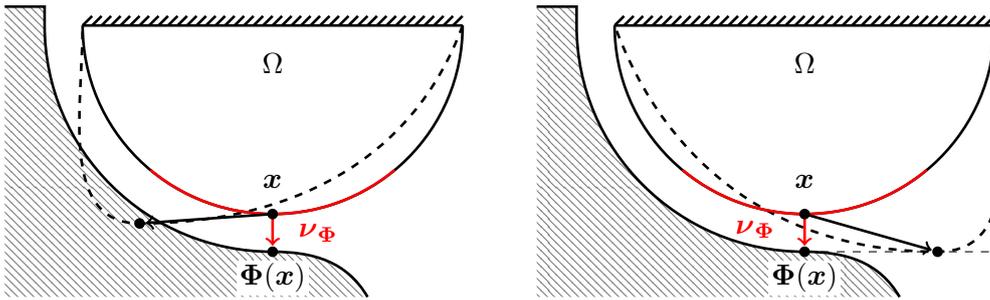


Figure 2.6: Modeling error in the linearized non-penetration constraint. Physical penetration is possible and contact may be incorrectly recognized because the linear constraint does not account for curved obstacle boundaries.

While the inability to incorporate curvature information in the constraint is clearly inherent to the linear description and can be expected for nonlinear obstacle boundaries, the aforementioned inconsistencies may similarly occur for uncurved obstacles when the contact normal is chosen improperly, see Figure 2.7.

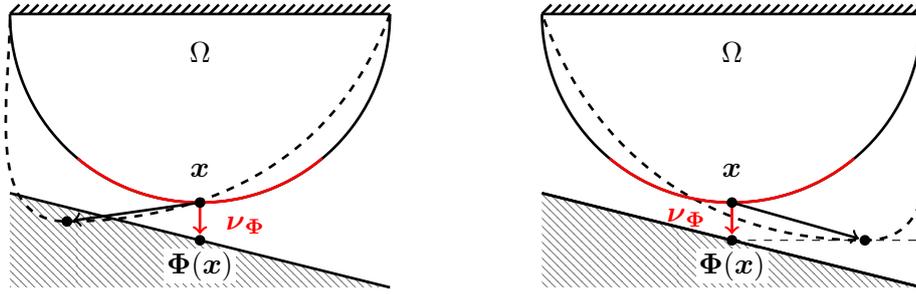


Figure 2.7: Modeling error in the linearized non-penetration constraint. Incorrect choice of the contact normal introduces errors even for planar boundaries.

Clearly, the specific choice of the contact mapping Φ , and consequently the contact normal ν_{Φ} , strongly influence the quality of the linear constraint (2.8) as a replacement of (2.9), and the input needs to be chosen with care. The main aspects to consider are twofold.

1. **Recognition of Contact:** Since contact in the linear constraint is assumed to occur at a point x in Γ_C if $\mathbf{y}(x) \cdot \nu_{\Phi}(x) = \Psi(x)$, the contact mapping that defines the contact normal should map points on the boundary of the body to the corresponding point on the rigid obstacle that it will in fact collide with during the deformation.
2. **Orientation of Contact Forces:** Because the displacements in contact normal direction on the contact boundary are restricted while tangential movements remain unrestricted, the contact forces that are introduced in the mathematical representation act in the direction of the contact normal, see Lemma 4.4.17 for a more precise description. In order to obtain physically meaningful contact forces, the contact normal direction $\nu_{\Phi}(x)$ at x in Γ_C , should coincide with the geometric normal of the obstacle \mathcal{O} at the point of contact during the deformation.

Aside from the fact that the corresponding point of contact on the obstacle for a point on the contact boundary is of course unknown in advance, the requirements for the recognition of contact and the orientation of the contact forces generally contradict each other. Accordingly, for complex geometries, a reasonable compromise is essential and needs to be made depending on the specific structure at hand.

In the classic formulation of the Signorini condition, there is no mention of a contact mapping or contact normal, and the geometric normal is used in its place, see [175] and [110]. This approach works well in so-called *detachment problems*, where the initial gap vanishes, and for small displacements and positive, small gap where the respective boundaries are necessarily very close and essentially “locally parallel”. Since small deformations are a requirement in order to obtain reasonable results from the linearized viscoelasticity model, the use of the geometric normal as the contact normal is generally accepted practice in the literature on (static) one-body problems, see, e.g., [4, 28, 62, 182], while the contact normal approach is more commonly taken in dynamic and two-body problems, see [112, 163]. There are, however, some decisive benefits to decoupling the geometry of the domain and the orientation of the contact normal. For one, in order to be able to resort to using a

Lipschitz continuous geometric normal, the domain of the body, or at the very least its contact boundary, is required to be of class $C^{1,1}$. Using a generally unrelated contact normal allows us to work with only $C^{0,1}$ -regular domains, i.e., Lipschitz domains. Furthermore, a closer look at the definition of ϵ in (2.1) reveals that while the linearized strain tensor may be unsuited to treat rotations and large distortions, it accurately treats large translations, i.e., $\epsilon = 0$ for all translations. This property is quite practical when considering dynamic problems where the body may remain undeformed while traveling towards the obstacle for longer periods of times. When considering large translations, the geometric normal is insufficient for the use as a contact normal, a problem that can be remedied to some degree if an appropriate contact mapping for the problem configuration can be found. In the special case where the obstacle's (relevant) boundary is *planar*, and can thus be fully described by first order information, this is always possible by choosing the contact normal parallel to the geometric normal of the planar obstacle boundary. Therefore, as one would expect, modeling a linear boundary with the linearized constraint with the contact normal approach over the geometric normal approach allows for full accuracy, both with respect to the orientation of the forces as well as the recognition of contact.

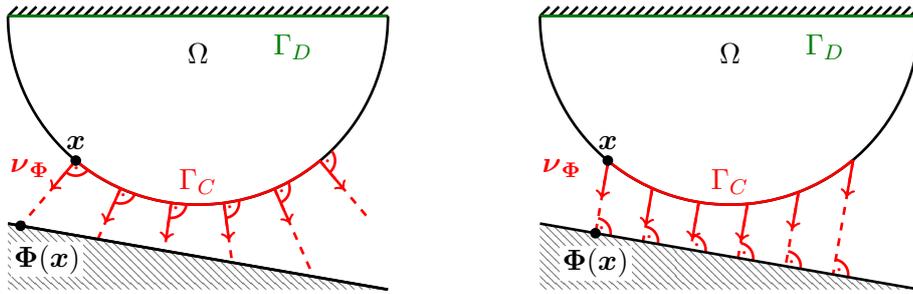


Figure 2.8: Geometric normal vs. decoupled contact normal. Using the geometric normal leads to errors in the model while using the normal of the obstacle is consistent in the case of planar obstacles.

In more complex configurations, a reasonable contact normal can be obtained by simulating the movement of the body without constraints first and taking the contact mapping that generates the geometric normal approach of the configuration at the first time at which the initial gap vanishes.

2.3 The Variational One-Body Contact Problem

Combining the unconstrained viscoelastic problem (2.7) with the requirements on the stresses in Section 2.2.1 amounts to the strong form of the one-body contact problem in linear, frictionless viscoelasticity of finding a displacement $\mathbf{y}: \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}^d$ that solves

the system

$$\rho \ddot{\mathbf{y}} - \operatorname{div} \boldsymbol{\sigma} = \mathbf{f}_\Omega \quad \text{on } \Omega \times [0, T] \quad (2.10a)$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{\nu} = \mathbf{f}_{\Gamma_N} \quad \text{on } \Gamma_N \times [0, T] \quad (2.10b)$$

$$\mathbf{y} = 0 \quad \text{on } \Gamma_D \times [0, T] \quad (2.10c)$$

$$\boldsymbol{\sigma}_t = 0 \quad \text{on } \Gamma_C \times [0, T] \quad (2.10d)$$

$$\boldsymbol{\sigma}_{\boldsymbol{\nu}_\Phi} \cdot \boldsymbol{\nu}_\Phi \leq 0 \quad \text{on } \Gamma_C \times [0, T] \quad (2.10e)$$

$$\mathbf{y} \cdot \boldsymbol{\nu}_\Phi - \Psi \leq 0 \quad \text{on } \Gamma_C \times [0, T] \quad (2.10f)$$

$$(\boldsymbol{\sigma}_{\boldsymbol{\nu}_\Phi} \cdot \boldsymbol{\nu}_\Phi)(\mathbf{y} \cdot \boldsymbol{\nu}_\Phi - \Psi) = 0 \quad \text{on } \Gamma_C \times [0, T] \quad (2.10g)$$

$$\mathbf{y}(0) = \mathbf{y}_{\text{ini}}, \quad \dot{\mathbf{y}}(0) = \mathbf{v}_{\text{ini}} \quad \text{on } \Omega. \quad (2.10h)$$

The problem consists of the equilibrium of force densities conditions (2.10a)–(2.10b), the Dirichlet boundary conditions (2.10c) and frictionless boundary conditions (2.10d), the contact conditions (2.10e)–(2.10g) and the initial conditions (2.10h).

The formulation of a corresponding weak variational form of (2.10) requires additional notation for the generalized forces and boundary conditions, which we will address now. Elastic stresses and viscous stresses are represented by the bilinear forms

$$a, b: \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}, \quad a_I, b_I: L^2(0, T; \mathbf{H}^1(\Omega)) \times L^2(0, T; \mathbf{H}^1(\Omega)) \rightarrow \mathbb{R}$$

of the form

$$\begin{aligned} a(\mathbf{v}, \mathbf{w}) &:= \int_\Omega \sum_{i,j,k,l=1}^d C_{ijkl} \partial_j v_i \partial_l w_k \, d\mathcal{L}^d, & a_I(\mathbf{y}, \mathbf{v}) &:= \int_0^T a(\mathbf{y}(t), \mathbf{v}(t)) \, dt, \\ b(\mathbf{v}, \mathbf{w}) &:= \int_\Omega \sum_{i,j,k,l=1}^d V_{ijkl} \partial_j v_i \partial_l w_k \, d\mathcal{L}^d, & b_I(\mathbf{y}, \mathbf{v}) &:= \int_0^T b(\mathbf{y}(t), \mathbf{v}(t)) \, dt. \end{aligned} \quad (2.11)$$

The bilinear forms above are bounded due to (2.5a) and when $\mathcal{H}^{d-1}(\Gamma_D) > 0$, coercivity is ensured by the uniform coercivity of the tensors (2.5b) and Korn's second inequality, see [110, Thm. 5.13, Lem. 6.2] or [46, 151] for a proof. As usual, Dirichlet boundary conditions are incorporated into the choice of the function space, and we employ a weak form of the non-penetration condition to fix the set of admissible displacements as the closed and convex set

$$\mathbf{K}_\Phi := \left\{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \operatorname{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_\Phi \leq \Psi \text{ } \mathcal{H}^{d-1}\text{- a.e. on } \Gamma_C \right\}. \quad (2.12)$$

Recall that we have assumed the obstacle and the viscoelastic body to occupy disjoint domains in the reference configuration, therefore \mathbf{K}_Φ contains the vanishing displacement and is therefore nonempty. Though the boundary measure theoretical trace sense for Sobolev functions is a common approach, see [28, 110, 112, 163], we will come back to this formulation after the considerations in Chapter 3, which allow us to find an alternative characterization of the set that simplifies the sensitivity analysis of the time-discretized problem. See Definition 3.5.4 for a formal redefinition of the set. Finally, combining the function space counterparts

$$f_\Omega \in L^2(0, T; \mathbf{H}^1(\Omega)^*), \quad f_{\Gamma_N} \in L^2\left(0, T; \mathbf{H}^{\frac{1}{2}}(\Gamma_N)^*\right)$$

of the boundary and volume force densities yields the external loads $f \in L^2(0, T; \mathbf{H}^1(\Omega))^*$ that are given by

$$\langle f, \mathbf{y} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} = \int_0^T \langle f_\Omega(t), \mathbf{y}(t) \rangle_{\mathbf{H}^1(\Omega)} + \langle f_{\Gamma_N}(t), \text{tr}(\mathbf{y}(t)) \rangle_{\mathbf{H}^{\frac{1}{2}}(\Gamma_N)} dt.$$

Assuming the expressions involved are sufficiently regular, the weak formulation of (2.10) can then be found by testing with suitable test functions and integrating using Green's formula with the appropriate boundary conditions, see [62, 110], which amounts to finding a function

$$\mathbf{y} \in \{ \mathbf{y} \in L^2(0, T; \mathbf{H}^1(\Omega)) : \dot{\mathbf{y}} \in \mathbf{W}(0, T) \}$$

with $\mathbf{y}(0) = \mathbf{y}_{\text{ini}}$, $\dot{\mathbf{y}}(0) = \mathbf{v}_{\text{ini}}$ and $\mathbf{y}(t) \in \mathbf{K}_\Phi$ for almost all t in $[0, T]$ that satisfies

$$\langle \rho \ddot{\mathbf{y}} - f, \mathbf{v} - \mathbf{y} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} + a_I(\mathbf{y}, \mathbf{v} - \mathbf{y}) + b_I(\dot{\mathbf{y}}, \mathbf{v} - \mathbf{y}) \geq 0 \quad (2.13)$$

for all $\mathbf{v} \in L^2(0, T; \mathbf{H}^1(\Omega))$ such that $\mathbf{v}(t) \in \mathbf{K}_\Phi$ for almost all t in $[0, T]$. Note that the quantities in the action inequality (2.13) are measured in Js, and the condition is directly related to the principle of stationary action. The non-equilibrium is of course due to the contact forces $f_{\text{con}} \in L^2(0, T; \mathbf{H}^1(\Omega))^*$ that are implicitly given by the residual in (2.13) as

$$f_{\text{con}}(t) := \langle \rho \ddot{\mathbf{y}}(t) - f(t), \cdot \rangle_{\mathbf{H}^1(\Omega)} + a(\mathbf{y}(t), \cdot) + b(\dot{\mathbf{y}}(t), \cdot) \quad (2.14)$$

for almost all t in $[0, T]$.

The main assumptions on the physical framework presented in this chapter can be summarized as follows:

Assumption 2.3.1 (Assumptions on the Model)

- (a) The spatial dimension d is in $\{2, 3\}$.
- (b) The mass density ρ is positive and constant in space and will not be stated explicitly again until Chapter 6.
- (c) The strain is measured by the infinitesimal strain tensor (2.1).
- (d) The stress is given as the sum of elastic and viscous stresses, and the constitutive relation is linear (2.4).
- (e) The elasticity and viscosity tensors are symmetric, uniformly bounded and positive definite (2.5).

Refer to the beginning of Chapter 4 for the technical assumptions on the data required to perform the forward analysis, and see Chapter 5 for the remaining assumptions in the optimal control framework.

Chapter 3

Capacities on the Boundary

Pointwise properties of Lebesgue functions can only be studied in the *almost everywhere* (*a.e.*) sense, i.e., up to sets of measure zero, which makes these sets negligible to some degree. Sobolev functions initially inherit this measure theoretical property. Their increased regularity, however, suggests that their properties can be studied in a refined sense. This is also indicated by the classic trace theorem (Theorem A.2.2), which defines boundary values on sets of d -dimensional Lebesgue measure zero up to surface measure zero. The implication is that the underlying measure theoretical sense is too coarse to determine which sets are negligible in the context of Sobolev functions. The appropriate tool to identify these sets are *capacities*, which are outer measures on $\mathcal{P}(\overline{\Omega})$ whose definition is directly based on Sobolev functions. Capacities account for the additional regularity and precisely identify the negligible sets as those of capacity zero — so-called *polar sets*. Properties holding up to polar sets are said to hold *quasi everywhere* (*q.e.*).

The analysis of the contact problem in Chapters 4 to 6 heavily relies on finding an adequate, Sobolev-capacity-based reformulation of the set of admissible displacements in (2.12) that is advantageous in the sensitivity of the (time-discretized) dynamic contact problem, e.g., in order to apply the results in [85, 138] on directional differentiability of projections in Dirichlet spaces and in order to obtain (*q.e.*) pointwise characterizations of the active contact set — similar to those in [194, Lem. A.5.] — for algorithmic purposes. Since the linearized non-penetration condition (2.8) that defines the set of admissible displacements is formulated on the contact boundary $\Gamma_C \subset \partial\Omega$, it is imperative that the quasi everywhere sense in the reformulation remain meaningful on the domain's boundary.

For the analysis of partial differential equations, obstacle problems and fine properties of Sobolev functions on the domain in general, the capacity based on the functions in $W_0^{1,p}(\Omega)$ is most commonly used [11, 65, 88, 194]. However, the zero boundary values prescribed on the entire boundary render it incapable of characterizing negligible sets on the boundary correctly, cf. the discussion after Lemma 3.2.5. Therefore, alternative notions of Sobolev capacity are required in order to correctly characterize negligible sets on the boundary. Several different approaches have been taken to this end. In [138], Mignot employed the theory of Dirichlet spaces — which is only applicable in the case $p = 2$ — and the capacity of the space $W^{1,2}(\Omega)$ to obtain a meaningful capacity on $\mathcal{P}(\overline{\Omega})$, see [138, P. 150, Ex. 2]

and cf. [85]. In [28], on the other hand, Betz worked with the $W_0^{1,p}(\Omega')$ -capacity of a domain Ω' satisfying $\bar{\Omega} \subset \Omega'$ to define a reasonable quasi everywhere sense on $\bar{\Omega}$. Lastly, another natural definition of a suitable capacity on the boundary is based on the space $W^{1-1/p,p}(\partial\Omega)$, see [138].

In this chapter, we will establish that the capacities associated with the Sobolev spaces $W^{1,p}(\Omega)$, $W_0^{1,p}(\Omega')$, $W^{1,p}(\mathbb{R}^d)$ and $W^{1-1/p,p}(\partial\Omega)$ are equivalent on the power set of the boundary $\mathcal{P}(\partial\Omega)$ of any Lipschitz domain Ω and that the corresponding quasi everywhere senses therefore coincide. When a positive distance to the Dirichlet boundary is maintained, we obtain equivalence to the $W_D^{1,p}(\Omega)$ -capacity as well, see the summary in Theorem 3.4.1. For the contact setting, this means that all previous approaches to Sobolev capacities that are capable of characterizing polar sets on the boundary actually coincide and are equally suited for the reformulation of the non-penetration condition. Given mild regularity assumptions on the initial gap function Ψ , the preliminary set of admissible displacements even coincides with its reformulation in the quasi everywhere sense, as we will see.

For a general overview of Sobolev capacities, refer to [11, 31, 88, 208]. See [70, 154] for approaches to capacities from the point of view of Dirichlet spaces and [1, 2, 208] for their treatment in (nonlinear) potential theory. Equivalences of Sobolev capacities have rarely been addressed previously, but basic equivalencies can be found in [11, 88]. The results, however, concern few very specific situations that are unrelated to the boundary effects that are of interest in the context of contact problems and their mixed boundary conditions.

Structure. In Section 3.1, the notion of a capacity is introduced in an abstract function space setting that covers the Sobolev settings that will be compared subsequently. We review several known results from capacity theory that are required for the corresponding analysis. The Sobolev settings that we will compare with respect to equivalence is specified in Section 3.2, and we examine the relation between zero boundary conditions of the underlying Sobolev functions and degenerate boundary behavior of the capacities. We show equivalence of the respective capacities on $\mathcal{P}(\bar{\Omega})$ and $\mathcal{P}(\partial\Omega)$ in Section 3.3. The results are collected for an overview in Section 3.4, and we derive some interesting conclusions for polar sets and boundary traces. Finally, in Section 3.5, we compare the weak forms of the non-penetration condition defined by the a.e. sense of the surface measure and the quasi everywhere sense introduced by any of the equivalent capacities examined in the previous sections.

The results of Sections 3.1–3.4 have previously been published in [44] and were modified to fit the framework of this thesis where appropriate. Additional results are obtained regarding the implications for the modeling of contact problems.

3.1 Abstract Setting

In this section, we will review basic results from capacity theory in an abstract setting that is tailored to the study of the spaces $W^{1,p}(\Omega)$, $W^{1,p}(\mathbb{R}^d)$, $W^{1-1/p,p}(\partial\Omega)$ and $W_0^{1,p}(\Omega')$ considered in Sections 3.2–3.4. The main results of this chapter concern the existence, the

uniqueness and the behavior of quasi continuous representatives of the Sobolev functions, see Corollary 3.1.14, Lemma 3.1.15 and Theorem 3.1.17 below. The approach employed in this section builds on the framework of topological spaces and is heavily inspired by the analysis in [69]. Consider the following setting:

Assumption 3.1.1 (Standing Assumptions for the Abstract Setting)

- (a) $(X, \mathcal{O}(X))$ is a topological space.
- (b) μ is a positive measure on $\mathcal{B}(X)$ with $\mu(A) > 0$ for all $A \in \mathcal{O}(X) \setminus \{\emptyset\}$.
- (c) $1 \leq p < \infty$ is arbitrary but fixed.
- (d) $V \subset L^p(X; \mu)$ is a Banach space such that
 - (i) $V \cap C(X)$ is $\|\cdot\|_V$ -dense in V ,
 - (ii) $\max(0, v) \in V$ and $\|\max(0, v)\|_V \leq \|v\|_V$ for all $v \in V$,
 - (iii) there exists a constant $\kappa > 0$ with $\|v\|_{L^p(X; \mu)} \leq \kappa \|v\|_V$ for all $v \in V$.

For details on the topological concepts in Assumption 3.1.1, refer to [203] or any of the references in Section 1.2. With slight abuse of notation, the Nemytskii operator to $\max: \mathbb{R}^2 \rightarrow \mathbb{R}$ is denoted by the same symbol, i.e., we assume the operator to act pointwise μ -almost everywhere on functions in V .

To avoid misunderstandings in the following, we fix the notion of neighborhoods of sets.

Definition 3.1.2 (Neighborhood) *Given $A \subset X$, a set O in $\mathcal{O}(X)$ satisfying $A \subset O$ is called a neighborhood of A .*

According to the definition, neighborhoods are necessarily open sets. In general topology, neighborhoods are occasionally defined as (possibly non-open) supersets of the sets defined in Definition 3.1.2, cf. [202]. However, the definition chosen above is slightly notationally advantageous in the following sections. Both notions of a neighborhood yield exactly the same capacity, which is evident from the following definition.

Definition 3.1.3 (Capacity) *The set function*

$$\begin{aligned} \text{cap}(\cdot; X, V, \mu): \mathcal{P}(X) &\rightarrow [0, \infty] \\ A &\mapsto \inf \{ \|v\|_V : v \in V, v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \} \end{aligned} \quad (3.1)$$

is called the capacity generated by the triple (X, V, μ) .

For improved readability, the dependency on the triple (X, V, μ) will be omitted in the remainder of this section and we simply write $\text{cap}(\cdot)$ instead of $\text{cap}(\cdot; X, V, \mu)$.

Remark 3.1.4 *In capacity theory of Sobolev and Dirichlet spaces, the term $\|v\|_V$ in the definition (3.1) is commonly raised to a suitable power. If, e.g., $p = 2$ and V is a Dirichlet space, the capacity is typically defined as*

$$\inf \left\{ \|v\|_V^2 : v \in V, v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \right\},$$

cf. [69, Sec. 3.1]. Given the setting in Assumption 3.1.1, however, where no further information about the space V and its norm $\|\cdot\|_V$ is available, such an approach is unnatural. Replacing the term $\|v\|_V$ in (3.1) with, e.g., $\|v\|_V^p$ would even cause the resulting capacity to

be non-subadditive in general, cf. the proof of Lemma 3.1.6 (d) below. However, the equivalency estimates obtained in Sections 3.2 to 3.4 using Definition 3.1.3 can easily be transformed to conform to the definitions of Sobolev capacity employed in [11, 31, 65, 88, 138] and others, see Lemma 3.4.8.

Since $\max(0, v) \in V$ and $\|\max(0, v)\|_V \leq \|v\|_V$ for all $v \in V$, we immediately obtain an alternative representation of the capacity.

Corollary 3.1.5 *The capacity in Definition 3.1.3 can equivalently be computed by*

$$\text{cap}(A) = \inf \{ \|v\|_V : v \in V, v \geq 0 \text{ } \mu\text{-a.e. in } X, v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \}$$

for all $A \subset X$.

By adapting the proofs in [88, Sec. 2] and [11, Sec. 5.8.2], we obtain the following properties of the capacity.

Lemma 3.1.6 (Elementary Properties of the Capacity)

- (a) *If $\text{cap}(A) = 0$ for $A \in \mathcal{B}(X)$, then $\mu(A) = 0$ as well.*
- (b) *If $A_1 \subset A_2 \subset X$, then $\text{cap}(A_1) \leq \text{cap}(A_2)$.*
- (c) *If $A_i, i = 1, \dots, n$, is a finite collection of subsets of X , then*

$$\frac{1}{n} \sum_{i=1}^n \text{cap}(A_i) \leq \text{cap} \left(\bigcup_{i=1}^n A_i \right).$$

- (d) *If $A_i, i \in \mathbb{N}$, is a countable collection of subsets of X , then*

$$\text{cap} \left(\bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} \text{cap}(A_i). \quad (3.2)$$

Proof. Due to Assumption 3.1.1 (d), there exists a $\kappa > 0$ such that

$$\begin{aligned} 0 \leq \frac{1}{\kappa} \mu(A)^{\frac{1}{p}} &\leq \frac{1}{\kappa} \inf \left\{ \|v\|_{L^p(X; \mu)} : v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \right\} \\ &\leq \inf \{ \|v\|_V : v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \} = \text{cap}(A) \quad \forall A \in \mathcal{B}(X), \end{aligned}$$

immediately yielding part (a). The monotonicity property in (b) is true since the set of functions over which the infimum in the definition of $\text{cap}(A_2)$ is taken is a subset of the set in the definition of $\text{cap}(A_1)$. Part (c) is obtained by adding up the inequalities $\text{cap}(A_i) \leq \text{cap}(\bigcup_{j=1}^n A_j)$, for $i = 1, \dots, n$ and dividing by n .

It remains to prove part (d). To this end, let $A_i \subset X, i \in \mathbb{N}$, be a countable collection of sets. We may assume w.l.o.g. that the series on the right hand side of (3.2) is finite, otherwise the inequality holds trivially. Consider an arbitrary but fixed $\varepsilon > 0$. From the alternative representation of $\text{cap}(\cdot)$ in Corollary 3.1.5, we obtain that for every $i \in \mathbb{N}$, we can find a v_i in V such that $v_i \geq 1$ μ -a.e. in a neighborhood of A_i , $v_i \geq 0$ μ -a.e. in X and

$$\text{cap}(A_i) \leq \|v_i\|_V \leq \text{cap}(A_i) + \frac{\varepsilon}{2^i}.$$

Since V is Banach, we can define $v := \sum_{i=1}^{\infty} v_i \in V$ and obtain

$$v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } \bigcup_{i=1}^{\infty} A_i, \quad \text{cap} \left(\bigcup_{i=1}^{\infty} A_i \right) \leq \|v\|_V \leq \varepsilon + \sum_{i=1}^{\infty} \text{cap}(A_i).$$

Letting $\varepsilon \rightarrow 0$ above yields (3.2) and completes the proof. \square

Due to its construction (see Definition 3.1.3), the capacity $\text{cap}(\cdot)$ has just the right level of sensitivity that is needed to properly identify those subsets of X that are negligible in the pointwise characterization of V -functions. Analogously to the classical almost everywhere sense, we can define a quasi everywhere sense as in [69, Chap. 3].

Definition 3.1.7 (Polar Sets and Quasi Everywhere (q.e.)) *If a subset N of X satisfies $\text{cap}(N) = 0$, then N is called a polar set. A condition depending on $x \in A \subset X$ is said to hold quasi everywhere (q.e.) in A if there exists a polar set $N \subset X$ such that the condition is satisfied for all x in $A \setminus N$.*

Corollary 3.1.8 *Let $A \subset X$ with $\text{cap}(A) = 0$, then there exists A' in $\mathcal{B}(X)$ such that $A \subset A'$ and $\text{cap}(A') = 0$.*

Proof. Due to Definition 3.1.3, for all $n \in \mathbb{N} \setminus \{0\}$ there exist $v_n \in V$ and $G_n \in \mathcal{O}(X)$ such that $\|v_n\|_V \leq \frac{1}{n}$, $A \subset G_n$ and $v_n \geq 1$ μ -a.e. on G_n , hence $\text{cap}(G_n) \leq \frac{1}{n}$. Set $A' := \bigcap_{n=1}^{\infty} G_n$, then $A' \in \mathcal{B}(X)$ and $A \subset A' \subset G_n$ for all n in $\mathbb{N} \setminus \{0\}$ and the monotonicity in Lemma 3.1.6 (b) yields that $\text{cap}(A') = 0$. \square

As a consequence, we may always assume the exceptional set in the Definition 3.1.7 of the quasi everywhere sense to be Borel measurable. Lemma 3.1.6 (a) and Corollary 3.1.8 further imply that the quasi everywhere sense is always at least as strict as the almost everywhere sense of the measure space $(X, \mathcal{B}(X), \mu)$.

Remark 3.1.9 *The notion of quasi everywhere becomes less restrictive with decreasing regularity of the functions in V . Analogously to Definition 3.1.3, a capacity for the base space $L^p(X; \mu)$ could be defined by setting*

$$\text{cap}_{L^p}(\cdot; X, \mu): \mathcal{P}(X) \rightarrow [0, \infty]$$

$$A \mapsto \inf \left\{ \|v\|_{L^p(X; \mu)} : v \in L^p(X; \mu), v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \right\}$$

When μ is an outer regular measure on $\mathcal{B}(X)$, we obtain $\mu(A)^{\frac{1}{p}} = \text{cap}_{L^p}(A)$ for all A in $\mathcal{B}(X)$, hence the a.e. and the “ L^p -q.e.” senses coincide in this case.

The (semi-)continuity of functions up to polar sets is of particular importance for the following analysis.

Definition 3.1.10 (Quasi Continuity) *A function $v: A \rightarrow \overline{\mathbb{R}}$ is called quasi (lower/upper semi-)continuous on a set $A \subset X$ if for every $\varepsilon > 0$ there exists a set G_ε in $\mathcal{O}(X)$ such that*

$$\text{cap}(G_\varepsilon) < \varepsilon \text{ and } v: A \setminus G_\varepsilon \rightarrow \overline{\mathbb{R}} \text{ is (lower/upper semi-)continuous.}$$

The terms “(lower/upper semi-) continuous” are understood in the topological sense, see [203, Sec. 7K] for details on the topological concepts. Recall that subsets of a topological space are endowed with the subspace topology. Due to their topological definition, quasi continuous functions are quite obviously connected to quasi open sets.

Definition 3.1.11 (Quasi Open and Quasi Closed Sets) *A set $A \subset X$ is quasi open if for every $\varepsilon > 0$ there exists a set G_ε in $\mathcal{O}(X)$ such that*

$$\text{cap}(G_\varepsilon) \leq \varepsilon \quad \text{and} \quad A \cup G_\varepsilon \in \mathcal{O}(X).$$

A set $A \subset X$ is quasi closed if $X \setminus A$ is quasi open.

Additional properties of quasi open sets and quasi continuous functions are discussed in Chapter B of the appendix.

Uniform convergence of sequences of functions is extended to quasi uniform convergence in a straight forward manner.

Definition 3.1.12 (Quasi Uniform Convergence) *A sequence of functions $v_n: X \rightarrow \mathbb{R}$ converges quasi uniformly in X to a function $v: X \rightarrow \mathbb{R}$ if for every $\varepsilon > 0$ there exists a set G_ε in $\mathcal{O}(X)$ such that*

$$\text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad \lim_{n \rightarrow \infty} \left(\sup_{x \in X \setminus G_\varepsilon} |v_n(x) - v(x)| \right) = 0.$$

An immediate consequence of Definition 3.1.3 and the properties of V is the following lemma.

Lemma 3.1.13 *Let $(v_n) \subset V \cap C(X)$ be a $\|\cdot\|_V$ -Cauchy sequence. Then there exists a subsequence (v_{n_k}) whose continuous representatives converge quasi uniformly in X to a quasi continuous and Borel measurable function $u: X \rightarrow \mathbb{R}$.*

Note that the continuous representative of an element of $V \cap C(X)$ is indeed unique, see Assumption 3.1.1 (b). The proof of the lemma is completely analogous to the classic Egorov theorem, cf. [7, Lem. 2.19] and also [88, Thm. 4.3], and even yields a non-increasing nested sequence of sets $(G_k) \subset \mathcal{O}(X)$ with $\text{cap}(G_k) < \frac{1}{k}$. See also Lemma B.1.1.

Proof. Since $(v_n) \subset V \cap C(X)$ is a Cauchy sequence in V , there exists a subsequence (denoted by the same symbol) such that

$$\sum_{n=1}^{\infty} 2^n \|v_n - v_{n+1}\|_V < \infty.$$

The above implies that for every k in \mathbb{N} there exists an N_k in \mathbb{N} with

$$\sum_{n=N_k}^{\infty} 2^n \|v_n - v_{n+1}\|_V < \frac{1}{2k}.$$

W.l.o.g., we assume that $N_k \leq N_{k+1}$ for all k in \mathbb{N} and define

$$E_n := \{x \in X : |v_n(x) - v_{n+1}(x)| > 2^{-n}\}.$$

Due to the continuity of the functions $|v_n - v_{n+1}|$, the sets E_n are in $\mathcal{O}(X)$, and

$$\begin{aligned} \text{cap}(E_n) &\leq \|2^n |v_n - v_{n+1}|\|_V \\ &\leq 2^n (\|\max(0, v_n - v_{n+1})\|_V + \|\min(0, v_n - v_{n+1})\|_V) \\ &\leq 2^{n+1} \|v_n - v_{n+1}\|_V. \end{aligned}$$

Setting

$$G_k := \bigcup_{n=N_k}^{\infty} E_n \in \mathcal{O}(X),$$

Lemma 3.1.6 (d) yields that

$$\text{cap}(G_k) \leq \sum_{n=N_k}^{\infty} \text{cap}(E_n) \leq 2 \sum_{n=N_k}^{\infty} 2^n \|v_n - v_{n+1}\|_V < \frac{1}{k} < \varepsilon$$

for arbitrary ε and sufficiently large k . Further, for all $N_k \leq m_1 \leq m_2$ we have that

$$\sup_{x \in X \setminus G_k} |v_{m_1}(x) - v_{m_2}(x)| \leq \sum_{n=m_1}^{m_2} \sup_{x \in X \setminus G_k} |v_n(x) - v_{n+1}(x)| \leq \sum_{n=m_1}^{m_2} 2^{-n} \xrightarrow{m_1 \rightarrow \infty} 0.$$

Accordingly, for all $k \geq 1$, the sequence $(v_n|_{X \setminus G_k}) \subset C(X \setminus G_k)$ is uniformly Cauchy for all $k \geq 1$, and from the uniform limit theorem ([202, Thm. 4.2.10]), we can deduce that $v_n \rightarrow u_k$ uniformly in $X \setminus G_k$ for some $u_k \in C(X \setminus G_k)$. Note that for $k_1 \geq k_2$, we have $N_{k_1} \geq N_{k_2}$ and, consequently, $G_{k_1} \subset G_{k_2}$. Therefore,

$$u_{k_1}(x) = u_{k_2}(x) \quad \forall x \in X \setminus G_{k_2} \quad \forall k_1 \geq k_2.$$

By setting

$$N := \bigcap_{k=1}^{\infty} G_k, \quad u(x) := \begin{cases} u_k(x), & \text{if } x \in X \setminus G_k \text{ for some } k \\ 0, & \text{if } x \in N \end{cases}, \quad (3.3)$$

we get a well-defined function $u: X \rightarrow \mathbb{R}$ that is obviously quasi continuous, and v_n converges to u quasi uniformly in X , according to its construction. Borel measurability of u is an immediate consequence of the representation in (3.3). \square

Using Lemma 3.1.13, it is straight forward to prove the following result on quasi continuity of V -limits.

Corollary 3.1.14 *If $(v_n) \subset V \cap C(X)$ is a sequence with $v_n \rightarrow v$ in V , then there exists a subsequence (v_{n_k}) and a Borel measurable, quasi continuous function $\tilde{v}: X \rightarrow \mathbb{R}$ such that the continuous representatives of v_{n_k} converge quasi uniformly in X to \tilde{v} and such that $v = \tilde{v}$ μ -a.e. in X .*

Proof. Lemma 3.1.13 implies that there exists a subsequence (v_{n_k}) such that the continuous representatives of v_{n_k} converge quasi uniformly in X to a quasi continuous, Borel measurable function $\tilde{v}: X \rightarrow \mathbb{R}$. In particular, this implies that $v_{n_k} \rightarrow \tilde{v}$ pointwise q.e. in

X and therefore $v_{n_k} \rightarrow \tilde{v}$ pointwise μ -a.e. in X as well, see Lemma 3.1.6 (a) and Corollary 3.1.8. From Assumption 3.1.1 (d) and by possibly passing over to another subsequence, we readily obtain that $v_{n_k} \rightarrow v$ μ -a.e. in X . Consequently, the functions v and \tilde{v} coincide μ -a.e., proving the claim. \square

Assumption 3.1.1 (d) ensures that $V \cap C(X)$ is dense in V , therefore Corollary 3.1.14 particularly implies that every v in V possesses a quasi continuous representative $\tilde{v}: X \rightarrow \mathbb{R}$, cf. Lusin's theorem in Lebesgue theory. The following lemma proves that this representative is even unique up to polar sets.

Lemma 3.1.15 *Let $A \subset X$ be quasi open and $u: A \rightarrow \overline{\mathbb{R}}$ be a quasi lower semi-continuous function satisfying $u \leq 0$ μ -a.e. in A . Then $u \leq 0$ q.e. in A as well.*

Proof. We proceed analogously to the proof of [194, Lem. 2.3]. Let $\varepsilon > 0$ and $G_\varepsilon, H_\varepsilon \in \mathcal{O}(X)$ such that

$$\begin{aligned} \text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad A \cup G_\varepsilon \in \mathcal{O}(X), \\ \text{cap}(H_\varepsilon) < \varepsilon \quad \text{and} \quad u: A \setminus H_\varepsilon \rightarrow \overline{\mathbb{R}} \text{ is lower semi-continuous,} \end{aligned}$$

and define $U_\varepsilon := G_\varepsilon \cup H_\varepsilon$, which is in $\mathcal{O}(X)$.

The lower semi-continuity of u yields that $\{x \in A \setminus H_\varepsilon : u(x) > 0\} \in \mathcal{O}(A \setminus H_\varepsilon)$, i.e., there exists a set O in $\mathcal{O}(X)$ such that $\{x \in A \setminus H_\varepsilon : u(x) > 0\} = O \cap A \setminus H_\varepsilon$ and accordingly $\{x \in A \setminus U_\varepsilon : u(x) > 0\} = O \cap A \setminus U_\varepsilon$. Using the quasi openness of A , we obtain that

$$\begin{aligned} \{x \in A \setminus U_\varepsilon : u(x) > 0\} \cup U_\varepsilon &= (O \cap A \setminus U_\varepsilon) \cup U_\varepsilon \\ &= (O \cup U_\varepsilon) \cap (A \cup G_\varepsilon \cup H_\varepsilon) \in \mathcal{O}(X). \end{aligned}$$

Since $\{x \in A : u(x) > 0\}$ is a set of μ -measure zero, every v in V with $v \geq 1$ μ -a.e. in a neighborhood of U_ε also satisfies $v \geq 1$ μ -a.e. in the set $\{x \in A \setminus U_\varepsilon : u(x) > 0\} \cup U_\varepsilon \in \mathcal{O}(X)$, which is a neighborhood of itself. Therefore, Definition 3.1.3 and the monotonicity and subadditivity of the capacity imply that

$$\text{cap}(\{x \in A : u(x) > 0\}) \leq \text{cap}(\{x \in A \setminus U_\varepsilon : u(x) > 0\} \cup U_\varepsilon) \leq \text{cap}(U_\varepsilon) < 2\varepsilon.$$

Letting $\varepsilon \rightarrow 0$ yields the claim. \square

Since X itself is open in X and therefore quasi open, we immediately obtain the following corollary.

Corollary 3.1.16 *Let $u: X \rightarrow \overline{\mathbb{R}}$ be a quasi lower semi-continuous function satisfying $u \leq 0$ μ -a.e. in X . Then $u \leq 0$ q.e. in X as well.*

By combining the results obtained in this section, we arrive at the following theorem.

Theorem 3.1.17 *Every function $v \in V$ admits a quasi continuous representative $\tilde{v}: X \rightarrow \mathbb{R}$ that is unique up to polar sets.*

Proof. The existence of a quasi continuous representative immediately follows from Corollary 3.1.14 and the density of $V \cap C(X)$ in V , as noted above. In order to prove uniqueness,

let $\tilde{v}_1, \tilde{v}_2: X \rightarrow \mathbb{R}$ be two quasi continuous representatives of v . Then $\tilde{v}_1 - \tilde{v}_2$ is quasi continuous with $\tilde{v}_1 - \tilde{v}_2 = 0$ μ -a.e. in X and we may employ Corollary 3.1.16 to deduce that $\tilde{v}_1 - \tilde{v}_2 = 0$ is satisfied q.e. in X . \square

The uniqueness of quasi continuous representatives up to polar sets implies a well defined sense of quasi everywhere behavior of a function $v \in V$.

Definition 3.1.18 (Quasi Everywhere Behavior) *A function $v \in V$ satisfies a pointwise condition quasi everywhere in $A \subset X$ if the respective condition is satisfied quasi everywhere in $A \subset X$ by all quasi continuous representatives of v .*

The quasi everywhere sense defined above provides the most natural setting for the study of pointwise properties of functions in V because it takes the regularity of the underlying function space into account.

3.2 Sobolev Capacities

In the next section, the results of the abstract setting are applied in the analysis of Sobolev capacities and the suitability of their corresponding q.e. senses for an adequate study of the active sets in the contact problem — and other obstacle type variational inequalities in Sobolev spaces, cf. [85, 138, 194, 196]. Based on the results of the previous section, we can now introduce and compare the Sobolev capacities that are relevant in the formulation of the contact constraint. We derive some additional results for Sobolev capacities and shortly discuss the degenerative behavior on and close to the boundary $\partial\Omega$ caused by zero boundary conditions. Considered the following setting:

Assumption 3.2.1 (Standing Assumptions for the Sobolev Setting)

- (a) $d \geq 2$ and $1 < p \leq d$.
- (b) $\Omega \subset \mathbb{R}^d$ is a bounded strong Lipschitz domain (Definition A.3.10).
- (c) $\Gamma_D \subset \partial\Omega$ is relatively open.
- (d) $\Omega' \subset \mathbb{R}^d$ is an open set satisfying $\overline{\Omega} \subset \Omega'$.

Remark 3.2.2 *The majority of the results proven in the following sections also hold for $p = 1$. That case is excluded here to avoid discussing the problems and peculiarities that arise in the context of Hardy's inequality and the inverse trace theorem when $W^{1,1}(\Omega)$ -capacities are considered, cf. [63] and [156].*

Note that Assumption 3.2.1 (a) excludes the cases where $W^{1,p}(\Omega)$ embeds into the function space $C(\overline{\Omega})$. When $W^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega})$, the only $W^{1,p}(\Omega)$ -polar set is the empty set and the study of Sobolev capacities becomes somewhat academic. The capacities that will be examined in the remainder of this chapter are the following:

$$\begin{aligned} \text{cap}_\Omega(\cdot) &:= \text{cap} \left(\cdot ; \overline{\Omega}, W^{1,p}(\Omega), \mathcal{L}^d \right), & \text{cap}_{\Omega,D}(\cdot) &:= \text{cap} \left(\cdot ; \overline{\Omega}, W_D^{1,p}(\Omega), \mathcal{L}^d \right), \\ \text{cap}_{\mathbb{R}^d}(\cdot) &:= \text{cap} \left(\cdot ; \mathbb{R}^d, W^{1,p}(\mathbb{R}^d), \mathcal{L}^d \right), & \text{cap}_{\Omega',0}(\cdot) &:= \text{cap} \left(\cdot ; \overline{\Omega'}, W_0^{1,p}(\Omega'), \mathcal{L}^d \right), \\ \text{cap}_{\partial\Omega}(\cdot) &:= \text{cap} \left(\cdot ; \partial\Omega, W^{1-\frac{1}{p},p}(\partial\Omega), \mathcal{H}^{d-1} \right). \end{aligned} \quad (3.4)$$

Remark 3.2.3

- (a) We always employ the Euclidean topology on \mathbb{R}^d and the associated subset topologies on $\partial\Omega$, $\bar{\Omega}$ and $\bar{\Omega}'$. Recall that for a bounded strong Lipschitz domain Ω , the subset topology on $\partial\Omega$ and the topology induced by the atlas of $\partial\Omega$ coincide, see [127, Lem. 1.30, Lem. 1.35]. Additionally, the boundary of Ω has d -dimensional Lebesgue measure zero, cf. [65, Lem. 2.2, Thm. 2.5].
- (b) For all $w \in W^{1,p}(\Omega)$ and $v \in W^{1-1/p,p}(\partial\Omega)$, we have that $\max(0, w) \in W^{1,p}(\Omega)$, $\max(0, v) \in W^{1-1/p,p}(\partial\Omega)$ and

$$\begin{aligned} \|\max(0, w)\|_{W^{1,p}(\Omega)} &\leq \|w\|_{W^{1,p}(\Omega)}, \\ \|\max(0, v)\|_{W^{1-\frac{1}{p},p}(\partial\Omega)} &\leq \|v\|_{W^{1-\frac{1}{p},p}(\partial\Omega)}, \end{aligned}$$

cf. [11, Sec. 5.8.1] and Section A.2.

- (c) Recall that $\|\cdot\|_{W^{1-\frac{1}{p},p}(\partial\Omega)}$ means the quotient norm, see Section 1.2.

With the observations in Remark 3.2.3 and other standard results from the theory of Sobolev spaces, e.g., [3, 11, 64, 121], it is easy to confirm that the triples (X, V, μ) in (3.4) all satisfy the conditions in Assumption 3.1.1. Accordingly, the theory of Section 3.1 is applicable and we may indeed discuss polar sets and quasi continuous representatives with respect to cap_Ω , $\text{cap}_{\Omega, D}$, $\text{cap}_{\mathbb{R}^d}$, $\text{cap}_{\Omega', 0}$ and $\text{cap}_{\partial\Omega}$. Note that all of the capacities in 3.4 can be encountered in the literature — most commonly raised to the power p , cf. Remark 3.1.4. The first capacity, cap_Ω , appears, e.g., in [138]. The second capacity ($\text{cap}_{\Omega, D}$) is commonly used in the study of partial differential equations, cf. [11, 65, 85, 88, 138], and $\text{cap}_{\mathbb{R}^d}$ is addressed in [69, 85, 88, 138]. The fourth one ($\text{cap}_{\Omega', 0}$) can be found in [28]. Lastly, the capacity of the trace space $W^{1-1/p,p}(\partial\Omega)$ has been considered in [85, Ex. 6]. We begin our study of the relationship between the capacities in (3.4) with the following elementary result.

Proposition 3.2.4 *The capacities satisfy*

$$\text{cap}_\Omega(\bar{\Omega}) < \infty, \text{cap}_{\Omega', 0}(\bar{\Omega}) < \infty, \text{cap}_{\mathbb{R}^d}(\bar{\Omega}) < \infty \text{ and } \text{cap}_{\partial\Omega}(\partial\Omega) < \infty.$$

Proof. Let $\varphi \in C_0^\infty(\mathbb{R}^d)$ be a cut-off function satisfying $\varphi \equiv 1$ in a neighborhood of $\bar{\Omega}$ and $\text{supp } \varphi \subset \Omega'$. Then Definition 3.1.3 implies

$$\begin{aligned} \text{cap}_\Omega(\bar{\Omega}) &\leq \|\varphi|_\Omega\|_{W^{1,p}(\Omega)}, & \text{cap}_{\partial\Omega}(\partial\Omega) &\leq \|\varphi|_{\partial\Omega}\|_{W^{1-\frac{1}{p},p}(\partial\Omega)}, \\ \text{cap}_{\mathbb{R}^d}(\bar{\Omega}) &\leq \|\varphi\|_{W^{1,p}(\mathbb{R}^d)}, & \text{cap}_{\Omega', 0}(\bar{\Omega}) &\leq \|\varphi|_{\Omega'}\|_{W^{1,p}(\Omega')}. \end{aligned}$$

Since all norms are finite, so are the capacities in the claim. \square

Because of Proposition 3.2.4 and Lemma 3.1.6, we know that the capacities cap_Ω , $\text{cap}_{\Omega', 0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\partial\Omega}$ define finite outer measures on the closure $\bar{\Omega}$ and the boundary $\partial\Omega$, respectively. Note that this is certainly untrue for $\text{cap}_{\Omega, D}$ and nonempty Γ_D since $\text{cap}_{\Omega, D}(A) = \infty$ for every $A \subset \bar{\Omega}$ with $A \cap \bar{\Gamma}_D \neq \emptyset$, which is an immediate consequence of Definition 3.1.3. In fact, if Γ_D is a $(d-1)$ -set ([63, Def. 4.1]), i.e., there exist constants $c, C > 0$ such that

$$cr^{d-1} \leq \mathcal{H}^{d-1}(\bar{\Gamma}_D \cap B(r, \mathbf{x})) \leq Cr^{d-1}$$

for all \mathbf{x} in $\overline{\Gamma_D}$ and r in $(0, 1]$, then $\text{cap}_{\Omega, D}(A)$ may even be infinite for sets A that do not intersect $\overline{\Gamma_D}$, as the following lemma shows.

Lemma 3.2.5 *If $\overline{\Gamma_D}$ is a $(d - 1)$ -set in the sense of [63, Def. 4.1], then there exists a constant $C > 0$ such that*

$$\left(\int_A \frac{1}{\text{dist}(\mathbf{x}, \Gamma_D)^p} \, d\mathbf{x} \right)^{\frac{1}{p}} \leq C \text{cap}_{\Omega, D}(A) \quad \forall A \subset \overline{\Omega}. \quad (3.5)$$

Proof. Let $v \in W_D^{1,p}(\Omega)$ be an arbitrary but fixed function satisfying $v \geq 1$ \mathcal{L}^d -a.e. in a (relative) neighborhood of A . Then Hardy's inequality [63, Thm. 3.2] implies

$$\int_A \frac{1}{\text{dist}(\mathbf{x}, \Gamma_D)^p} \, d\mathbf{x} \leq \int_{\Omega} \frac{|v(\mathbf{x})|^p}{\text{dist}(\mathbf{x}, \Gamma_D)^p} \, d\mathbf{x} \leq C \|v\|_{W^{1,p}(\Omega)}^p. \quad (3.6)$$

Taking the infimum over all v in (3.6) yields (3.5) as claimed. \square

Lemma 3.2.5 and [63, Thm. 3.4] immediately yield that $\text{cap}_{\Omega, D}(\Omega) = \infty$ whenever $\overline{\Gamma_D}$ is a $(d - 1)$ -set. Moreover, using Fatou's lemma and (3.5), we obtain that $\text{cap}_{\Omega, D}(A_k) \rightarrow \infty$ as $k \rightarrow \infty$ for every compact exhaustion (A_k) of Ω and every $1 < p \leq d$. This implies that $\text{cap}_{\Omega, D}$ cannot be equivalent to any of the other capacities in (3.4), i.e., there cannot exist constants $c, C > 0$ with, e.g.,

$$c \text{cap}_{\Omega}(A) \leq \text{cap}_{\Omega, D}(A) \leq C \text{cap}_{\Omega}(A) \quad \forall A \subset \Omega. \quad (3.7)$$

Instead, the following type of pseudo-equivalency estimate can be obtained.

Proposition 3.2.6 *The estimate*

$$\text{cap}_{\Omega}(A) \leq \text{cap}_{\Omega, D}(A) \leq \left(1 + \frac{d^{\frac{1}{p}}}{\text{dist}(A, \Gamma_D)} \right) \text{cap}_{\Omega}(A) \quad (3.8)$$

is satisfied for all $A \subset \overline{\Omega}$ that satisfy $\text{dist}(A, \Gamma_D) > 0$ or $\text{cap}_{\Omega}(A) > 0$.

Proof. The first inequality in (3.8) is trivial since $W_D^{1,p}(\Omega)$ is a subset of $W^{1,p}(\Omega)$ and since the capacities cap_{Ω} and $\text{cap}_{\Omega, D}$ are both defined w.r.t. the subset topology of $\overline{\Omega}$, see Definition 3.1.3 and (3.4). To prove the second estimate, let $A \subset \overline{\Omega}$ be arbitrary but fixed. Note that (3.8) holds trivially if $\text{dist}(A, \Gamma_D) = 0$ and $\text{cap}_{\Omega}(A) > 0$ and if $\text{dist}(A, \Gamma_D) > 0$ but $\Gamma_D = \emptyset$. Therefore, we may assume w.l.o.g. that $\text{dist}(A, \Gamma_D) > 0$ and that Γ_D is nonempty. For sufficiently small $\varepsilon > 0$, consider

$$\delta_{\varepsilon}: \overline{\Omega} \rightarrow [0, 1], \quad \delta_{\varepsilon}(\mathbf{x}) := \min \left(\max \left(0, (1 + 2\varepsilon) \frac{\text{dist}(\mathbf{x}, \Gamma_D)}{\text{dist}(A, \Gamma_D)} - \varepsilon \right), 1 \right),$$

and let $v \in W^{1,p}(\Omega)$ be an arbitrary function satisfying $v \geq 1$ \mathcal{L}^d -a.e. in a (relative) neighborhood of A . Then $\delta_{\varepsilon} \in C^{0,1}(\overline{\Omega})$ and $\delta_{\varepsilon} \equiv 0$ in a neighborhood of $\overline{\Gamma_D}$ as well as $\delta_{\varepsilon} \equiv 1$ in a neighborhood of \overline{A} . Therefore, the product $v\delta_{\varepsilon}$ is an element of $W_D^{1,p}(\Omega)$, cf. [28, Cor. A18], and $v\delta_{\varepsilon} \geq 1$ \mathcal{L}^d -a.e. in a neighborhood of A . Further, if we denote

$\|\nabla\delta_\varepsilon\|_{\mathbf{L}^\infty(\Omega)} = \max_{k=1,\dots,d} \|\partial_k\delta_\varepsilon\|_{L^\infty(\Omega)}$, then we obtain the estimate

$$\begin{aligned}
\|v\delta_\varepsilon\|_{W^{1,p}(\Omega)} &= \left(\|v\delta_\varepsilon\|_{L^p(\Omega)}^p + \sum_{k=1}^d \|\partial_k v \delta_\varepsilon + v \partial_k \delta_\varepsilon\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} \\
&\leq \left(\|v\delta_\varepsilon\|_{L^p(\Omega)}^p + \sum_{k=1}^d \left(\|\partial_k v \delta_\varepsilon\|_{L^p(\Omega)} + \|v \partial_k \delta_\varepsilon\|_{L^p(\Omega)} \right)^p \right)^{\frac{1}{p}} \\
&\leq \left(\|v\|_{L^p(\Omega)}^p + \sum_{k=1}^d \left(\|\partial_k v\|_{L^p(\Omega)} + \|\partial_k \delta_\varepsilon\|_{L^\infty(\Omega)} \|v\|_{L^p(\Omega)} \right)^p \right)^{\frac{1}{p}} \\
&\leq \left(\|v\|_{L^p(\Omega)}^p + \sum_{k=1}^d \|\partial_k v\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^d \|\partial_k \delta_\varepsilon\|_{L^\infty(\Omega)} \|v\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} \\
&\leq \left(1 + d^{\frac{1}{p}} \|\nabla\delta_\varepsilon\|_{\mathbf{L}^\infty(\Omega)} \right) \|v\|_{W^{1,p}(\Omega)} \\
&\leq \left(1 + d^{\frac{1}{p}} \frac{1 + 2\varepsilon}{\text{dist}(A, \Gamma_D)} \right) \|v\|_{W^{1,p}(\Omega)}, \tag{3.9}
\end{aligned}$$

where Minkowski's inequality in \mathbb{R}^{d+1} was applied to the third line. Taking the infimum over all v in (3.9), we obtain

$$\text{cap}_{\Omega, D}(A) \leq \left(1 + d^{\frac{1}{p}} \frac{1 + 2\varepsilon}{\text{dist}(A, \Gamma_D)} \right) \text{cap}_\Omega(A) \quad \forall \varepsilon > 0.$$

Passing to the limit $\varepsilon \rightarrow 0$ in the above completes the proof. \square

Note that we have explicitly excluded the degenerative case of cap_Ω -polar sets A that intersect $\overline{\Gamma_D}$, for which clearly $\text{cap}_{\Omega, D}(A) = \infty$, in order to avoid notational confusion regarding the conventions for multiplication of zero and infinity. Nonetheless, we immediately obtain the following inequalities.

Corollary 3.2.7 *Let Γ_D be a $(d-1)$ -set. Then there exist constants $c, C > 0$ such that*

$$c \left(\int_A \frac{1}{\text{dist}(\mathbf{x}, \Gamma_D)^p} \, d\mathbf{x} \right)^{\frac{1}{p}} \leq \text{cap}_{\Omega, D}(A) \leq C \left(1 + \frac{d^{\frac{1}{p}}}{\text{dist}(A, \Gamma_D)} \right) \tag{3.10}$$

for all $A \subset \overline{\Omega}$.

Proof. The left hand inequality is from Lemma 3.2.5. For $\text{dist}(A, \Gamma_D) > 0$ the right hand inequality is an immediate consequence of Propositions 3.2.6 and 3.2.4 and for $\text{dist}(A, \Gamma_D) = 0$, the right hand inequality is trivial. \square

This shows that the qualitative behavior of the $W_D^{1,p}(\Omega)$ -capacity is directly related to that of the distance function $A \ni \mathbf{x} \mapsto \text{dist}(\mathbf{x}, \Gamma_D) \in [0, \infty]$. Observe that the second estimate in (3.10) is not optimal since there exist sets $A \subset \overline{\Omega}$ with $\text{dist}(A, \Gamma_D) = 0$ and $\text{cap}_{\Omega, D}(A) < \infty$. While it is interesting to study the geometry of sets with the latter two properties more generally, the following exemplary result will suffice to give an idea of what sets with these properties look like.

Theorem 3.2.8 *Let $\Omega := (0, 1) \times (0, 1)$ and $\Gamma_D := (0, 1) \times \{0\}$. Then the sets*

$$A_\alpha := \{(x_1, x_2) \in \Omega : x_1^\alpha < x_2\} \subset \Omega \text{ with } \alpha > 0$$

satisfy the following:

- (a) A_α is open and $\text{dist}(A_\alpha, \Gamma_D) = 0$ for all $\alpha > 0$.
- (b) For $1 < p < 2$:

$$\text{cap}_{\Omega, D}(A_\alpha) = \text{cap}\left(A_\alpha; \overline{\Omega}, W_D^{1,p}(\Omega), \mathcal{L}^2\right) \begin{cases} < \infty, & \alpha < \frac{1}{p-1} \\ = \infty, & \alpha \geq \frac{1}{p-1} \end{cases}. \quad (3.11)$$

- (c) For $p = 2$:

$$\text{cap}_{\Omega, D}(A_\alpha) = \text{cap}\left(A_\alpha; \overline{\Omega}, W_D^{1,2}(\Omega), \mathcal{L}^2\right) = \infty \text{ for all } \alpha > 0.$$

Proof. Part (a) is obvious from the definition of A_α . To obtain (b), note that $\overline{\Gamma_D} = [0, 1] \times \{0\}$ is a 1-set and that according to Lemma 3.2.5, for all $1 < p \leq 2$, there exists a constant $C > 0$ with

$$C \text{cap}_{\Omega, D}(A_\alpha)^p \geq \int_{A_\alpha} \frac{1}{\text{dist}(\mathbf{x}, \Gamma_D)^p} d\mathbf{x} = \int_0^1 \int_0^{x_2^{1/\alpha}} \frac{1}{x_2^p} dx dx_2 = \int_0^1 x_2^{-p+1/\alpha} dx_2.$$

This implies $\text{cap}_{\Omega, D}(A_\alpha) = \infty$ for all $\alpha \geq 1/(p-1)$ and proves the second case in (3.11). It remains to show that $\text{cap}_{\Omega, D}(A_\alpha) < \infty$ for $0 < \alpha < 1/(p-1)$ and $1 < p < 2$. To this end, assume that $p-1 < \alpha < 1/(p-1)$ for now and define

$$v_\alpha : \Omega \rightarrow \mathbb{R}, \quad v_\alpha(\mathbf{x}) := \min\left(1, \frac{x_2}{x_1^\alpha}\right).$$

Then v_α is in $L^\infty(\Omega) \cap W_{\text{loc}}^{1,\infty}(\Omega)$, and in the distributional sense,

$$\nabla v_\alpha(x_1, x_2) = \begin{cases} \left(-\alpha \frac{x_2}{x_1^{\alpha+1}}, \frac{1}{x_1^\alpha}\right) & \mathcal{L}^2\text{-a.e. in } \{(x_1, x_2) \in \Omega : x_2 < x_1^\alpha\} \\ (0, 0) & \mathcal{L}^2\text{-a.e. in } \{(x_1, x_2) \in \Omega : x_2 \geq x_1^\alpha\} \end{cases}.$$

A straight forward computation yields that

$$\int_\Omega |\partial_{x_1} v_\alpha|^p + |\partial_{x_2} v_\alpha|^p d\mathcal{L}^2 = \int_0^1 \frac{\alpha^p}{p+1} x_1^{\alpha-p} + x_1^{\alpha(1-p)} dx_1 < \infty,$$

and therefore $v_\alpha \in W^{1,p}(\Omega)$. Further, our construction yields $v_\alpha \geq 1$ \mathcal{L}^2 -a.e. in A_α and $\text{tr}(v_\alpha) = 0$ \mathcal{H}^1 -a.e. on Γ_D , where the latter follows from the continuity of v_α on $\overline{\Omega} \setminus \{0\}$, the properties of the trace operator and a localization argument using cut-off functions around $(0, 0)$. Since A_α is open, we obtain

$$v_\alpha \in \left\{ v \in W_D^{1,p}(\Omega) : v \geq 1 \text{ } \mathcal{L}^2\text{-a.e. in a nbhd. of } A_\alpha \right\} \neq \emptyset.$$

Accordingly, we know that $\text{cap}_{\Omega,D}(A_\alpha) < \infty$, and the second case in (3.11) is proven for all $1 < p < 2$ and $(p-1) < \alpha < 1/(p-1)$. For the remaining α , equation (3.11) follows from the monotonicity of $\text{cap}_{\Omega,D}$, see Lemma 3.1.6 (b), and the fact that $A_{\alpha_1} \subset A_{\alpha_2}$ for all $0 < \alpha_1 \leq \alpha_2$. This completes the proof of (b).

It remains to show that $\text{cap}_{\Omega,D}(A_\alpha) = \infty$ for all $\alpha > 0$ in the case $p = 2$, which we prove by contradiction.

Assuming the existence of an $\alpha > 0$ with $\text{cap}_{\Omega,D}(A_\alpha) < \infty$, we can find at least one function $v \in W_D^{1,2}(\Omega)$ satisfying $v \geq 1$ \mathcal{L}^2 -a.e. in a neighborhood of A_α . Linearity of the trace operator and [28, Lem. 3.27] imply that $\text{tr}(\min(v, 1)) = \min(\text{tr}(v), 1)$ \mathcal{H}^2 -a.e. on $\partial\Omega$ and consequently $\text{tr}(\min(v, 1)) = 0$ \mathcal{H}^1 -a.e. on $(0, 1) \times \{0\}$.

Additionally, $\text{tr}(\min(v, 1)) = 1$ \mathcal{H}^1 -a.e. on $\{0\} \times (0, 1)$ can be confirmed because $\min(v, 1) = 1$ \mathcal{L}^2 -a.e. in a relative neighborhood of every point $x \in \{0\} \times (0, 1)$. Therefore, we can construct a sequence $v_n \in C^\infty(\bar{\Omega})$ converging to $\min(v, 1)$ in $W^{1,2}(\Omega)$ and $v_n \rightarrow 1$ \mathcal{H}^2 -a.e. on $\{0\} \times (0, 1)$, see [121, Thm. 5.5.9].

Using the Sobolev–Slobodeckij norm (1.1), one can immediately compute that a function that is locally a step function can not be an element of $W^{\frac{1}{2},2}(\partial\Omega)$. This contradiction yields that the set $\{v \in W_D^{1,2}(\Omega) : v \geq 1 \text{ } \mathcal{L}^2\text{-a.e. in a nbhd. of } A_\alpha\}$ is empty for all $\alpha > 0$, which yields (c). \square

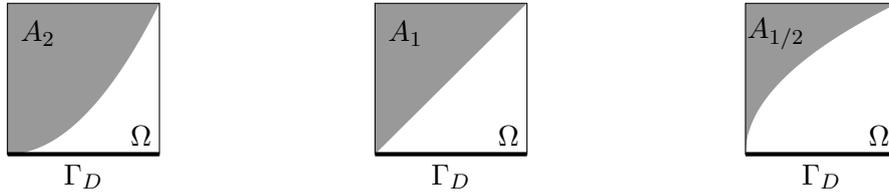


Figure 3.1: Sets with degenerative capacity behavior at the boundary. The sets $A_\alpha \subset \Omega$ for $\alpha = 2, 1, 1/2$, left to right.

Remark 3.2.9 *Theorem 3.2.8 not only demonstrates that there exist configurations where $\text{dist}(A, \Gamma_D) = 0$ and $\text{cap}_{\Omega,D}(A) < \infty$, but it also shows that the capacities $\text{cap}_{\Omega,D}$ are typically non-equivalent (in the sense of (3.7)) for different values of p .*

The singular behavior exhibited by the set function $\text{cap}_{\Omega,D}$ on and near the Dirichlet boundary part Γ_D is the main reason why $W_D^{1,p}(\Omega)$ -capacities and, consequently, the $W_0^{1,p}(\Omega)$ -capacity as a special case, are unfit for applications that require an adequate study of subsets of the boundary $\partial\Omega$ that are not sufficiently separated from the Dirichlet boundary. Note that in contrast to $\text{cap}_{\Omega,D}$, the capacities cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\partial\Omega}$ are all able to meaningfully measure subsets of $\partial\Omega$, and as mentioned above, all of these capacities have been used in the literature as a substitute for $\text{cap}_{\Omega,0}$ at one point or another. In the following sections, equivalence of the latter four capacities on $\mathcal{P}(\partial\Omega)$ is established. This implies that they all give rise to the same quasi everywhere sense on the boundary $\partial\Omega$.

3.3 Equivalence of Sobolev Capacities

The equivalence results for the capacities in (3.4) are split in two parts. We start out showing the equivalence of the capacities cap_Ω , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$ on $\mathcal{P}(\overline{\Omega})$ in Section 3.3.1, and in Section 3.3.2, we show that $\text{cap}_{\partial\Omega}$ and $\text{cap}_{\Omega',0}$ are additionally equivalent on $\mathcal{P}(\partial\Omega)$. Consequently, equivalence of all four of these capacities on $\mathcal{P}(\partial\Omega)$ will be established. We employ the notions of Section 3.1 and Assumption 3.2.1 is assumed to hold for the entire section.

3.3.1 Equivalences on $\mathcal{P}(\overline{\Omega})$

We will establish equivalence of cap_Ω , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$ on $\mathcal{P}(\overline{\Omega})$ by showing that there exists a constant $C = C(\Omega, \Omega') > 0$ such that

$$\text{cap}_\Omega(A) \leq \text{cap}_{\mathbb{R}^d}(A) \leq \text{cap}_{\Omega',0}(A) \leq C \text{cap}_\Omega(A) \quad \forall A \subset \overline{\Omega}. \quad (3.12)$$

The first two parts of this estimate can be obtained in a straight forward manner.

Lemma 3.3.1 *The estimate*

$$\text{cap}_\Omega(A) \leq \text{cap}_{\mathbb{R}^d}(A) \leq \text{cap}_{\Omega',0}(A) \quad (3.13)$$

is satisfied for all $A \subset \overline{\Omega}$.

Proof. Using restriction and the definitions of the subset topologies on $\overline{\Omega'}$ and $\overline{\Omega}$, we obtain that for all $A \subset \overline{\Omega}$:

$$\begin{aligned} \text{cap}_\Omega(A) &= \inf \{ \|v\|_{W^{1,p}(\Omega)} : v \in W^{1,p}(\Omega), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \cap \overline{\Omega} \} \\ &\leq \inf \{ \|v|_\Omega\|_{W^{1,p}(\Omega)} : v \in W^{1,p}(\mathbb{R}^d), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \} \\ &\leq \inf \{ \|v\|_{W^{1,p}(\mathbb{R}^d)} : v \in W^{1,p}(\mathbb{R}^d), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \} \\ &= \text{cap}_{\mathbb{R}^d}(A), \end{aligned}$$

and in the same manner, using extension by zero ([3, Lem. 3.27]) and the definition of the subset topology, we conclude

$$\begin{aligned} \text{cap}_{\mathbb{R}^d}(A) &= \inf \{ \|v\|_{W^{1,p}(\mathbb{R}^d)} : v \in W^{1,p}(\mathbb{R}^d), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \} \\ &\leq \inf \{ \|v\|_{W^{1,p}(\Omega')} : v \in W_0^{1,p}(\Omega'), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \} \\ &= \inf \{ \|v\|_{W^{1,p}(\Omega')} : v \in W_0^{1,p}(\Omega'), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \cap \Omega' \} \\ &\leq \text{cap}_{\Omega',0}(A). \end{aligned} \quad \square$$

In order to show the equivalence (3.12), it remains to show that there exists a constant $C(\Omega, \Omega') > 0$ such that

$$\text{cap}_{\Omega', 0}(A) \leq C \text{cap}_{\Omega}(A) \quad \forall A \subset \overline{\Omega}. \quad (3.14)$$

Unfortunately, the derivation of (3.14) is not as straight forward as that of (3.13). The proof of (3.13) is comparatively simple because the restriction $v|_{\Omega}$ of a function v satisfying $v \geq 1$ a.e. on an \mathbb{R}^d - or $\overline{\Omega'}$ -relatively open set always satisfies $v|_{\Omega} \geq 1$ a.e. on an $\overline{\Omega}$ -open set — an immediate consequence of the definition of the subset topology.

To prove (3.14), one has to recover the condition “ $v \geq 1$ a.e. in a $\overline{\Omega'}$ -relative neighborhood” from the condition “ $v \geq 1$ a.e. in a $\overline{\Omega}$ -relative neighborhood” for an extension of $v \in W^{1,p}(\Omega)$ to $W_0^{1,p}(\Omega')$, i.e., the transition from the subset topology of $\overline{\Omega}$ to the topology of the ambient space $\overline{\Omega'}$ needs to be handled. The mere existence of a linear and continuous extension operator $E: W^{1,p}(\Omega) \rightarrow W_0^{1,p}(\Omega')$ is insufficient if this topological requirement is not met.

In the remainder of this section, we first prove the inequality (3.14) in a prototypical situation where the boundary corresponds to a subset of a hyperplane. With a stability result for capacities under domain transformations, the general case is then obtained by localization and rectification arguments.

We start out with the prototypical situation, where we can construct a bounded extension that is compatible with the topological constraints based on a reflection.

Lemma 3.3.2 *Let $B(s)$ denote the open ball in \mathbb{R}^{d-1} with radius $s > 0$ centered at the origin and let*

$$U(s, t) := B(s) \times (-t, t), \quad V(s, t) := B(s) \times (0, t), \quad W(s, t) := B(s) \times [0, t).$$

Then for all $r, \varepsilon > 0$, there exists a constant $C = C(r, \varepsilon)$ such that

$$\text{cap}_{U(3r, 3\varepsilon), 0}(A) \leq C \text{cap}_{V(3r, 3\varepsilon)}(A) \quad \forall A \subset W(r, \varepsilon).$$

Proof. Let $r, \varepsilon > 0$ and $A \subset W(r, \varepsilon)$ be arbitrary but fixed. Assume that a function $v \in W^{1,p}(V(3r, 3\varepsilon))$ and an open set $G \in \mathcal{O}(\mathbb{R}^d)$ are given such that

$$A \subset G \quad \text{and} \quad v \geq 1 \quad \mathcal{L}^d\text{-a.e. in } G \cap \overline{V(3r, 3\varepsilon)}. \quad (3.15)$$

We will use this v to construct a function w in $W_0^{1,p}(U(3r, 3\varepsilon))$ that satisfies $w \geq 1$ \mathcal{L}^d -a.e. in a $U(3r, 3\varepsilon)$ -neighborhood of A and $\|w\|_{W^{1,p}(U(3r, 3\varepsilon))} \leq C(r, \varepsilon) \|v\|_{W^{1,p}(V(3r, 3\varepsilon))}$ for a suitable, positive constant $C(r, \varepsilon)$.

Let $\psi \in C_0^\infty(\mathbb{R}^d)$ be an arbitrary but fixed cut-off function that satisfies

$$\psi \equiv 1 \text{ in } U(r, \varepsilon), \quad 0 \leq \psi \leq 1 \text{ in } U(2r, 2\varepsilon) \setminus U(r, \varepsilon) \quad \text{and} \quad \psi \equiv 0 \text{ in } \mathbb{R}^d \setminus U(2r, 2\varepsilon).$$

Then the product ψv satisfies

$$\psi v \in W^{1,p}(V(3r, 3\varepsilon)), \quad \|\psi v\|_{V(3r, 3\varepsilon)} \leq C(r, \varepsilon) \|v\|_{W^{1,p}(V(3r, 3\varepsilon))}$$

for a positive constant $C = C(\psi)$, cf. [77, Lem. 1.4.1.1], and

$$\psi v = 0 \text{ in } V(3r, 3\varepsilon) \setminus V(2r, 2\varepsilon), \quad \psi v = v \geq 1 \text{ in } G \cap \overline{V(r, \varepsilon)}$$

both hold in the \mathcal{L}^d -a.e. sense.

We define the extension w of v for almost all $\mathbf{x} = (\mathbf{x}', x_d)$ in $U(3r, 3\varepsilon)$ by reflection across the hyperplane $\mathbb{R}^{d-1} \times \{0\}$, i.e.,

$$w(\mathbf{x}) := \begin{cases} (\psi v)(\mathbf{x}), & \text{if } \mathbf{x} \in V(3r, 3\varepsilon) \\ (\psi v)(\mathbf{x}', -x_d), & \text{if } (\mathbf{x}', -x_d) \in V(3r, 3\varepsilon) \end{cases},$$

so that

$$w \in W^{1,p}(U(3r, 3\varepsilon)), \quad w = 0 \text{ } \mathcal{L}^d\text{-a.e. in } U(3r, 3\varepsilon) \setminus U(2r, 2\varepsilon), \\ \|w\|_{W^{1,p}(U(3r, 3\varepsilon))} \leq 2 \|\psi v\|_{W^{1,p}(V(3r, 3\varepsilon))}.$$

It remains to check that $w \geq 1$ \mathcal{L}^d -a.e. in a $U(3r, 3\varepsilon)$ -relative neighborhood. We define the set

$$H := (G \cap W(r, \varepsilon)) \cup \left\{ \mathbf{x} \in \mathbb{R}^d : (\mathbf{x}', -x_d) \in G \cap W(r, \varepsilon) \right\} \subset U(r, \varepsilon),$$

see Figure 3.2. Note that because $A \subset G \cap W(r, \varepsilon)$, we know that $A \subset H$. We will confirm that H is open in the following proof by contradiction. Assume that there exists an \mathbf{x} in H such that there is no ball of radius $\delta > 0$ around \mathbf{x} that is contained in H . Then we can find a sequence $(\mathbf{x}_n) \subset U(r, \varepsilon) \setminus H$ with $\mathbf{x}_n \rightarrow \mathbf{x}$. Since H is symmetric w.r.t. the last component by definition, we can either find a subsequence of (\mathbf{x}_n) , denoted by the same symbol, which is contained in $W(r, \varepsilon) \setminus H$ or we consider the reflection of both \mathbf{x} and \mathbf{x}_n across the hyperplane, which then satisfy $\mathbf{x} \in H$, $\mathbf{x}_n \in W(r, \varepsilon) \setminus H$ and $\mathbf{x}_n \rightarrow \mathbf{x}$.

From convergence of the d^{th} component of the sequence, we can deduce that $\mathbf{x} \in G \cap W(r, \varepsilon) \subset G \cap U(r, \varepsilon)$, which is an open set, and

$$\mathbf{x}_n \in W(r, \varepsilon) \setminus H = W(r, \varepsilon) \setminus (G \cap W(r, \varepsilon)) = W(r, \varepsilon) \setminus (G \cap U(r, \varepsilon)),$$

i.e., $\mathbf{x}_n \notin (G \cap U(r, \varepsilon))$, contradicts the convergence $\mathbf{x}_n \rightarrow \mathbf{x}$.

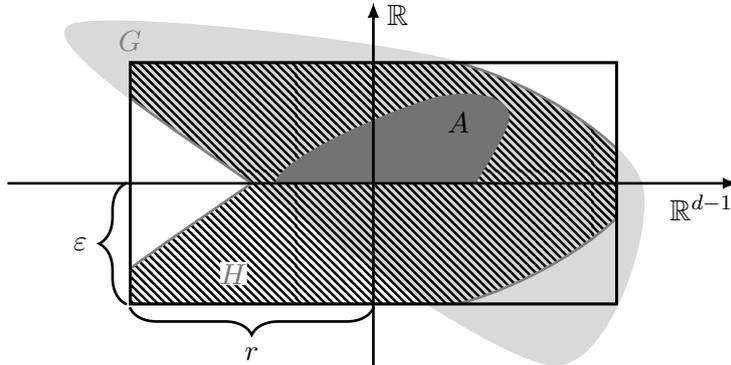


Figure 3.2: The extension-set H and its construction.

Consequently, the function w satisfies

$$\begin{aligned} w &\in W_0^{1,p}(U(3r, 3\varepsilon)), \\ \|w\|_{W^{1,p}(U(3r, 3\varepsilon))} &\leq 2 \|\psi v\|_{W^{1,p}(V(3r, 3\varepsilon))}, \\ w &\geq 1 \mathcal{L}^d\text{-a.e. in } H, \quad H \subset U(3r, 3\varepsilon) \text{ open}, \quad A \subset H. \end{aligned} \quad (3.16)$$

By (3.16) and taking the infimum over all v in $W^{1,p}(V(3r, 3\varepsilon))$ that satisfy (3.15) for some open set G , we obtain

$$\begin{aligned} \text{cap}_{U(3r, 3\varepsilon), 0}(A) &= \inf\{ \|w\|_{W^{1,p}(U(3r, 3\varepsilon))} : w \in W_0^{1,p}(U(3r, 3\varepsilon)) \text{ and} \\ &\quad w \geq 1 \mathcal{L}^d\text{-a.e. in a } \overline{U(3r, 3\varepsilon)}\text{-nbhd. of } A \} \\ &\leq \inf\{ 2 \|\psi v\|_{W^{1,p}(V(3r, 3\varepsilon))} : v \in W^{1,p}(V(3r, 3\varepsilon)) \text{ and} \\ &\quad v \geq 1 \mathcal{L}^d\text{-a.e. in a } \overline{V(3r, 3\varepsilon)}\text{-nbhd. of } A \} \\ &\leq 2C(\psi) \text{cap}_{V(3r, 3\varepsilon)}(A) \end{aligned}$$

with a constant $C = C(\psi) = C(r, \varepsilon)$. □

With the case of a “flat” boundary handled by Lemma 3.3.2, we use the following result on the stability of $W^{1,p}$ - and $W_0^{1,p}$ -capacities under bi-Lipschitz coordinate transformations in order to locally reduce the general case to the former.

Lemma 3.3.3 *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^d$ be bounded strong Lipschitz domains and let $\Theta: \overline{\Omega_1} \rightarrow \overline{\Omega_2}$ be a bi-Lipschitz mapping with $\Theta(\overline{\Omega_1}) = \overline{\Omega_2}$ and $\Theta(\Omega_1) = \Omega_2$. Then there exist constants $c, C > 0$ depending only on Θ such that*

$$c \text{cap}_{\Omega_1}(A) \leq \text{cap}_{\Omega_2}(\Theta(A)) \leq C \text{cap}_{\Omega_1}(A), \quad (3.17a)$$

$$c \text{cap}_{\Omega_1, 0}(A) \leq \text{cap}_{\Omega_2, 0}(\Theta(A)) \leq C \text{cap}_{\Omega_1, 0}(A) \quad (3.17b)$$

for all $A \subset \overline{\Omega_1}$.

Proof. Clearly, it suffices to show one of the inequalities in (3.17a) and one of the inequalities in (3.17b). The remaining estimates follow analogously by reversing the roles of Θ and Θ^{-1} . We show the left one of either of the inequalities.

To that end, let $A \subset \overline{\Omega_1}$ such that $\Theta(A) \subset \overline{\Omega_2}$, and assume that G in $\mathcal{O}(\overline{\Omega_2})$ and v in $W^{1,p}(\Omega_2)$ are given such that

$$\Theta(A) \subset G \quad \text{and} \quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G.$$

Then $w := v \circ \Theta \in W^{1,p}(\Omega_1)$, and $\|w\|_{W^{1,p}(\Omega_1)} \leq c \|v\|_{W^{1,p}(\Omega_2)}$ for a constant $c = c(\Theta)$, see [148, Lem. 2.3.2] and [208, Thm. 2.2.2], as well as

$$A \subset \Theta^{-1}(G) \in \mathcal{O}(\overline{\Omega_1}) \quad \text{and} \quad w \geq 1 \mathcal{L}^d\text{-a.e. in } \Theta^{-1}(G).$$

Accordingly,

$$\begin{aligned}
\text{cap}_{\Omega_1}(A) &= \inf \{ \|w\|_{W^{1,p}(\Omega_1)} : w \in W^{1,p}(\Omega_1), \exists H \in \mathcal{O}(\overline{\Omega_1}) \text{ s.t. } A \subset H \text{ and} \\
&\quad w \geq 1 \mathcal{L}^d\text{-a.e. in } H \} \\
&\leq \inf \{ \|v \circ \Theta\|_{W^{1,p}(\Omega_1)} : v \in W^{1,p}(\Omega_2), \exists G \in \mathcal{O}(\overline{\Omega_2}) \text{ s.t. } \Theta(A) \subset G \text{ and} \\
&\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \} \\
&\leq c \inf \{ \|v\|_{W^{1,p}(\Omega_2)} : v \in W^{1,p}(\Omega_2), \exists G \in \mathcal{O}(\overline{\Omega_2}) \text{ s.t. } \Theta(A) \subset G \text{ and} \\
&\quad v \geq 1 \mathcal{L}^d\text{-a.e. in } G \} \\
&\leq c \text{cap}_{\Omega_2}(\Theta(A)),
\end{aligned}$$

dividing by c and relabeling the constant proves (3.17a).

The estimate in (3.17b) follows by the same argument, but we need to check that $v \circ \Theta$ is in $W_0^{1,p}(\Omega_1)$ for all v in $W_0^{1,p}(\Omega_2)$. Note that due to Lemma A.2.8, we have that

$$\text{tr}(v \circ \Theta) = \text{tr}(v) \circ \Theta = 0 \quad \forall v \in W_0^{1,p}(\Omega_2)$$

and since Ω_1 is Lipschitz, [148, Lem. 2.4.10] yields that $v \circ \Theta \in W_0^{1,p}(\Omega_1)$. \square

Remark 3.3.4 *The capacities in Lemma 3.3.3 can also be defined if Ω_1 and Ω_2 do not have Lipschitz boundaries. Since $v \circ \Theta \in W_0^{1,p}(\Omega_1)$ for all $v \in W_0^{1,p}(\Omega_2)$ can be confirmed for arbitrary bounded domains by a lengthy approximation argument without the use of the trace operator, the previous Lemma can be stated without the Lipschitz regularity of the boundaries.*

Remark 3.3.5 *The requirement $\Theta(\Omega_1) = \Omega_2$ ensures that the transformation Θ also maps $\partial\Omega_1$ to $\partial\Omega_2$. When there exist open domains Ω'_1 and Ω'_2 such that $\overline{\Omega_i} \subset \Omega'_i$, $i = 1, 2$ and $\Theta: \Omega'_1 \rightarrow \Omega'_2$ is bi-Lipschitz with $\Theta(\overline{\Omega_1}) = \overline{\Omega_2}$, the former holds.*

We can now combine the equivalence in the special case and the transformation result that were proven in the previous lemmas for a proof of (3.14) in the general case.

Proposition 3.3.6 *There exists a constant $C = C(\Omega, \Omega') > 0$ such that*

$$\text{cap}_{\Omega',0}(A) \leq C \text{cap}_{\Omega}(A) \quad \forall A \subset \overline{\Omega}.$$

Proof. The strategy of the following proof is to employ a decomposition of the domain into a finite cover of the boundary and a remaining interior subdomain with positive distance to the boundary. The inequality will be treated locally on each of the subsets. Monotonicity of the capacity then allows for a recombination of the local estimates to obtain the claim.

Part 1 (Decomposition): Recall that according to the definition of a strong Lipschitz domain, e.g., [65, Def. 4.4], for every $\mathbf{q} \in \partial\Omega$, there exist an orthogonal transformation $R_{\mathbf{q}}$ in $\text{SO}(d)$, an open ball $B_{\mathbf{q}} \subset \mathbb{R}^{d-1}$ with midpoint $\mathbf{x}'_{\mathbf{q}}$ in \mathbb{R}^{d-1} , an open interval $J_{\mathbf{q}} = (a_{\mathbf{q}}, b_{\mathbf{q}})$ and a Lipschitz map $h_{\mathbf{q}}: B_{\mathbf{q}} \rightarrow J_{\mathbf{q}}$ such that

$$\mathbf{q} \in R_{\mathbf{q}}(B_{\mathbf{q}} \times J_{\mathbf{q}}) \text{ and } \Omega \cap R_{\mathbf{q}}(B_{\mathbf{q}} \times J_{\mathbf{q}}) = R_{\mathbf{q}}(\{(\mathbf{x}', x_d) : \mathbf{x}' \in B_{\mathbf{q}}, h_{\mathbf{q}}(\mathbf{x}') < x_d < b_{\mathbf{q}}\}).$$

Further, note that since $\bar{\Omega} \subset \Omega'$, we can always reduce the size of the sets $B_{\mathbf{q}}$ and $J_{\mathbf{q}}$ sufficiently to ensure that $\overline{R_{\mathbf{q}}(B_{\mathbf{q}} \times J_{\mathbf{q}})}$ is a subset of Ω' . Due to continuity of $h_{\mathbf{q}}$, we can further find $\varepsilon_{\mathbf{q}}, r_{\mathbf{q}} > 0$ such that

$$\{(\mathbf{x}', x_d) : \|\mathbf{x}' - \mathbf{x}'_{\mathbf{q}}\| < 4r_{\mathbf{q}}, |x_d - h_{\mathbf{q}}(\mathbf{x}')| < 4\varepsilon_{\mathbf{q}}\} \subset B_{\mathbf{q}} \times J_{\mathbf{q}}.$$

We fix a choice of $R_{\mathbf{q}}, B_{\mathbf{q}}, J_{\mathbf{q}}, h_{\mathbf{q}}, r_{\mathbf{q}}$ and $\varepsilon_{\mathbf{q}}$ for every $\mathbf{q} \in \partial\Omega$, and for all $0 < s \leq 4r_{\mathbf{q}}$ and all $0 < t \leq 4\varepsilon_{\mathbf{q}}$, we define

$$\begin{aligned} \tilde{U}_{\mathbf{q}}(s, t) &:= R_{\mathbf{q}}(\{(\mathbf{x}', x_d) : \|\mathbf{x}' - \mathbf{x}'_{\mathbf{q}}\| < s, |x_d - h_{\mathbf{q}}(\mathbf{x}')| < t\}), \\ \tilde{V}_{\mathbf{q}}(s, t) &:= R_{\mathbf{q}}(\{(\mathbf{x}', x_d) : \|\mathbf{x}' - \mathbf{x}'_{\mathbf{q}}\| < s, h_{\mathbf{q}}(\mathbf{x}') < x_d < h_{\mathbf{q}}(\mathbf{x}') + t\}) = \tilde{U}_{\mathbf{q}}(s, t) \cap \Omega, \\ \tilde{W}_{\mathbf{q}}(s, t) &:= R_{\mathbf{q}}(\{(\mathbf{x}', x_d) : \|\mathbf{x}' - \mathbf{x}'_{\mathbf{q}}\| < s, h_{\mathbf{q}}(\mathbf{x}') \leq x_d < h_{\mathbf{q}}(\mathbf{x}') + t\}) = \tilde{U}_{\mathbf{q}}(s, t) \cap \bar{\Omega}. \end{aligned}$$

Then $\{\tilde{U}_{\mathbf{q}}(r_{\mathbf{q}}, \varepsilon_{\mathbf{q}}) : \mathbf{q} \in \partial\Omega\}$ is an open cover of the compact set $\partial\Omega$, and we can find a finite number of points $\mathbf{q}_i, i = 1, \dots, n$, such that

$$\partial\Omega \subset \bigcup_{i=1}^n \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}). \quad (3.18)$$

Now consider an arbitrary but fixed set $A \subset \bar{\Omega}$ and define

$$A_i := A \cap \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}) \quad \forall i = 1, \dots, n, \quad \text{and} \quad A_0 := A \setminus \bigcup_{i=1}^n \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}),$$

then

$$A_i \subset \tilde{W}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}) \quad \forall i = 1, \dots, n, \quad A_0 \subset \Omega \setminus \bigcup_{i=1}^n \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}) \quad \text{and} \quad A = \bigcup_{i=0}^n A_i.$$

Part 2 (Local Estimates): Note that since

$$\text{dist}(A_0, \partial\Omega) \geq \text{dist}\left(\Omega \setminus \bigcup_{i=1}^n \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}), \partial\Omega\right) > 0,$$

we obtain a strictly positive lower bound on the distance of A_0 and $\partial\Omega$ that is independent of the set A . Proposition 3.2.6 therefore yields the existence of a constant $C_0 = C_0(\Omega, \Omega') > 0$ such that

$$\text{cap}_{\Omega, 0}(A_0) \leq C_0 \text{cap}_{\Omega}(A_0),$$

and since the extension of $W_0^{1,p}(\Omega)$ -functions to $W_0^{1,p}(\Omega')$ -functions by zero is an isometry, we deduce

$$\text{cap}_{\Omega', 0}(A_0) \leq \text{cap}_{\Omega, 0}(A_0) \leq C_0 \text{cap}_{\Omega}(A_0). \quad (3.19)$$

Accordingly, the remaining task is to find similar estimates for the contributions of the sets A_i , $i = 1, \dots, n$. To this end, fix an i in $\{1, \dots, n\}$, assume w.l.o.g. that $R_{q_i} = I$ and define

$$U(s, t) := B(s) \times (-t, t), \quad V(s, t) := B(s) \times (0, t), \quad W(s, t) := B(s) \times [0, t),$$

as in Lemma 3.3.2. The bi-Lipschitz transformations Θ_{q_i} with

$$\begin{aligned} \Theta_{q_i} : U(4r_{q_i}, 4\varepsilon_{q_i}) &\rightarrow \tilde{U}_{q_i}(4r_{q_i}, 4\varepsilon_{q_i}), & (\mathbf{x}', y) &\mapsto (\mathbf{x}'_{q_i} + \mathbf{x}', y + h_{q_i}(\mathbf{x}'_{q_i} + \mathbf{x}')), \\ \Theta_{q_i}^{-1} : \tilde{U}_{q_i}(4r_{q_i}, 4\varepsilon_{q_i}) &\rightarrow U(4r_{q_i}, 4\varepsilon_{q_i}), & (\mathbf{x}', y) &\mapsto (-\mathbf{x}'_{q_i} + \mathbf{x}', y - h_{q_i}(\mathbf{x}')), \end{aligned}$$

locally flatten the boundary of the distorted covering boxes and satisfy

$$\begin{aligned} \Theta_{q_i}(\overline{U(3r_{q_i}, 3\varepsilon_{q_i})}) &= \overline{\tilde{U}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i})}, & \Theta_{q_i}(U(3r_{q_i}, 3\varepsilon_{q_i})) &= \tilde{U}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i}), \\ \Theta_{q_i}(\overline{V(3r_{q_i}, 3\varepsilon_{q_i})}) &= \overline{\tilde{V}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i})}, & \Theta_{q_i}(V(3r_{q_i}, 3\varepsilon_{q_i})) &= \tilde{V}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i}), \\ \Theta_{q_i}^{-1}(W(r_{q_i}, \varepsilon_{q_i})) &= \tilde{W}_{q_i}(r_{q_i}, \varepsilon_{q_i}) \supseteq A_i. \end{aligned}$$

The stability results in Lemma 3.3.3 yield a constant $C = C(r_{q_i}, \varepsilon_{q_i}, h_{q_i}) > 0$ such that

$$\begin{aligned} \text{cap}_{\tilde{U}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i}), 0}(A_i) &\leq C \text{cap}_{U(3r_{q_i}, 3\varepsilon_{q_i}), 0}(\Theta_{q_i}^{-1}(A_i)), \\ \text{cap}_{V(3r_{q_i}, 3\varepsilon_{q_i})}(\Theta_{q_i}^{-1}(A_i)) &\leq C \text{cap}_{\tilde{V}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i})}(A_i). \end{aligned}$$

For the transformed domains with $\Theta_{q_i}^{-1}(A_i) \subset W(r_{q_i}, \varepsilon_{q_i})$, the prototypical result in Lemma 3.3.2 implies the existence of a constant $\tilde{C} = \tilde{C}(r_{q_i}, \varepsilon_{q_i}) > 0$ with

$$\text{cap}_{U(3r_{q_i}, 3\varepsilon_{q_i}), 0}(\Theta_{q_i}^{-1}(A_i)) \leq \tilde{C} \text{cap}_{V(3r_{q_i}, 3\varepsilon_{q_i})}(\Theta_{q_i}^{-1}(A_i)).$$

By combining the last three inequalities, we obtain the existence of a constant $C_i = C_i(r_{q_i}, \varepsilon_{q_i}, h_{q_i}) > 0$ such that

$$\text{cap}_{\tilde{U}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i}), 0}(A_i) \leq C_i \text{cap}_{\tilde{V}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i})}(A_i).$$

The same extension-by-zero argument as in (3.19) yields

$$\text{cap}_{\Omega', 0}(A_i) \leq \text{cap}_{\tilde{U}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i}), 0}(A_i),$$

and by restriction of $W^{1,p}(\Omega)$ functions to $W^{1,p}(\tilde{V}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i}))$ functions, we immediately obtain that

$$\text{cap}_{\tilde{V}_{q_i}(3r_{q_i}, 3\varepsilon_{q_i})}(A_i) \leq \text{cap}_{\Omega}(A_i).$$

Accordingly, the last three inequalities show that

$$\text{cap}_{\Omega', 0}(A_i) \leq C_i \text{cap}_{\Omega}(A_i).$$

Part 3 (Recombination): Part 2 of the proof establishes the existence of constants $C_i = C_i(\Omega, \Omega') > 0$ for $i = 0, \dots, n$ such that

$$\text{cap}_{\Omega', 0}(A_i) \leq C_i \text{cap}_{\Omega}(A_i).$$

The monotonicity and the subadditivity properties of the capacities, see Lemma 3.1.6 (b) and (d), therefore yield that

$$\text{cap}_{\Omega',0}(A) \leq \sum_{i=0}^n \text{cap}_{\Omega',0}(A_i) \leq \sum_{i=0}^n C_i \text{cap}_{\Omega}(A_i) \leq (n+1) \max_{i=0,\dots,n} (C_i) \text{cap}_{\Omega}(A). \quad \square$$

Accordingly, the capacities cap_{Ω} , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$ are indeed equivalent on $\bar{\Omega}$, which constitutes the main result of this section.

Theorem 3.3.7 *Let Assumption 3.2.1 hold, and let cap_{Ω} , $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\Omega',0}$ be defined as in (3.4). Then there exists a constant $C(\Omega, \Omega') > 0$ such that*

$$\text{cap}_{\Omega}(A) \leq \text{cap}_{\mathbb{R}^d}(A) \leq \text{cap}_{\Omega',0}(A) \leq C \text{cap}_{\Omega}(A) \quad \forall A \subset \bar{\Omega}.$$

Proof. Lemma 3.3.1 and Proposition 3.3.6 immediately imply the theorem. \square

3.3.2 Equivalences on $\mathcal{P}(\partial\Omega)$

A trivial implication of the previous section is the equivalence of the capacities cap_{Ω} , $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\Omega',0}$ in Theorem 3.3.7 on $\mathcal{P}(\partial\Omega)$, since it is a subset of $\mathcal{P}(\Omega)$. We will now see that these capacities are additionally equivalent to the capacity $\text{cap}_{\partial\Omega}$ on $\mathcal{P}(\partial\Omega)$. Specifically, we will prove that there exists a constant $C = C(\Omega, \Omega') > 0$ such that

$$\frac{1}{C} \text{cap}_{\Omega',0}(A) \leq \text{cap}_{\partial\Omega}(A) \leq \text{cap}_{\Omega}(A) \quad \forall A \subset \partial\Omega \quad (3.20)$$

and then use Theorem 3.3.7, see the previous subsection. In this subsection, we will encounter the Sobolev-Slobodeckij norm on several occasions. Recall the equivalence to the quotient norm (Lemma A.2.3).

As in the previous subsection, one of the estimates in (3.20) is easily obtained.

Lemma 3.3.8 *The estimate*

$$\text{cap}_{\partial\Omega}(A) \leq \text{cap}_{\Omega}(A)$$

is satisfied for all sets $A \subset \partial\Omega$.

Proof. First, as in Theorem 3.2.8, for every v in $W^{1,p}(\Omega)$ and G in $\mathcal{O}(\mathbb{R}^d)$ such that $v = 1$ \mathcal{L}^d -a.e. in $G \cap \Omega$, we can construct a sequence $(v_n) \subset C^\infty(\bar{\Omega})$ such that $v_n \rightarrow v$ in $W^{1,p}(\Omega)$ and $v_n \rightarrow 1$ \mathcal{H}^{d-1} -a.e. on $G \cap \partial\Omega$, and hence $\text{tr}(v) = 1$ \mathcal{H}^{d-1} -a.e. in $G \cap \partial\Omega$.

Further, the definition of the quotient norm on $W^{1-1/p,p}(\partial\Omega)$ implies that

$$\|\text{tr}(v)\|_{W^{1-1/p,p}(\partial\Omega)} := \inf \left\{ \|w\|_{W^{1,p}(\Omega)} : v \in W^{1,p}(\Omega), \text{tr}(w) = v \right\} \leq \|v\|_{W^{1,p}(\Omega)}$$

for all v in $W^{1,p}(\Omega)$, see (1.2).

Lastly, because Ω is bounded, any constant function is in $W^{1,p}(\Omega)$, and [11, Cor. 5.8.2.] ensures that $\|\min(v, 1)\|_{W^{1,p}(\Omega)} \leq \|v\|_{W^{1,p}(\Omega)}$, cf. Remark 3.2.3 (b).

Applying these arguments in the same order, we obtain that

$$\begin{aligned}
\text{cap}_{\partial\Omega}(A) &= \inf\{ \|w\|_{W^{1-\frac{1}{p},p}(\partial\Omega)} : w \in W^{1-\frac{1}{p},p}(\partial\Omega), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\
&\quad w = 1 \mathcal{H}^{d-1}\text{-a.e. in } G \cap \partial\Omega \} \\
&\leq \inf\{ \|\text{tr}(v)\|_{W^{1-\frac{1}{p},p}(\partial\Omega)} : v \in W^{1,p}(\Omega), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\
&\quad v = 1 \mathcal{L}^d\text{-a.e. in } G \cap \bar{\Omega} \} \\
&\leq \inf\{ \|v\|_{W^{1,p}(\Omega)} : v \in W^{1,p}(\Omega), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\
&\quad v = 1 \mathcal{L}^d\text{-a.e. in } G \cap \bar{\Omega} \} \\
&= \inf\{ \|v\|_{W^{1,p}(\Omega)} : v \in W^{1,p}(\Omega), \exists G \in \mathcal{O}(\mathbb{R}^d) \text{ s.t. } A \subset G \text{ and} \\
&\quad v \leq 1 \mathcal{L}^d\text{-a.e. in } G \cap \bar{\Omega} \} \\
&= \text{cap}_{\Omega}(A)
\end{aligned}$$

for all $A \subset \partial\Omega$. □

The remaining estimate on the left of (3.20) is shown by an argument similar to the one in Section 3.3.1, i.e., the proof is split up into the proof for the case of a flat boundary and a stability result with respect to Lipschitz transformations of local boundary sections, which are then combined with localization and rectification arguments for the general case.

In contrast to the proof of Lemma 3.3.2, the reflection-based (local) extension operator is replaced by a boundary-to-domain extension operator based on the regularization operator presented in [148, Sec. 2.5.5].

Lemma 3.3.9 ([148, Lem. 2.5.6]) *Let $B(s)$ denote the open ball in \mathbb{R}^{d-1} with radius $s > 0$ centered at the origin and let ρ be a mollifying kernel on \mathbb{R}^{d-1} , i.e.,*

$$0 \leq \rho \in C^\infty(\mathbb{R}^{d-1}), \quad \text{supp}(\rho) \subset B(1), \quad \int_{\mathbb{R}^{d-1}} \rho \, d\mathcal{L}^{d-1} = 1.$$

Let $v \in W^{1-1/p,p}(\mathbb{R}^{d-1} \times \{0\}) \cong W^{1-1/p,p}(\mathbb{R}^{d-1})$ be a function such that $v = 0$ \mathcal{L}^{d-1} -a.e. in $\mathbb{R}^{d-1} \setminus B(r)$ for an $r > 0$, and let $R > 0$ be arbitrary but fixed. Then there exists a constant $C > 0$ that is independent of v such that the function

$$w(\mathbf{x}', x_d) := \frac{1}{x_d^{d-1}} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \rho\left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d}\right) d\mathcal{L}^{d-1}(\mathbf{y}'), \quad (\mathbf{x}', x_d) \in \mathbb{R}^{d-1} \times (0, R)$$

satisfies

$$\begin{aligned}
&w \in C^\infty(\mathbb{R}^{d-1} \times (0, R)), \quad \text{tr}(w) = v \text{ in } \mathbb{R}^{d-1} \times \{0\}, \\
&w = 0 \mathcal{L}^d\text{-a.e. in } \left\{ (\mathbf{x}', x_d) \in \mathbb{R}^{d-1} \times (0, R) : \|\mathbf{x}'\| \geq r + x_d \right\}, \\
&\|w\|_{W^{1,p}(\mathbb{R}^{d-1} \times (0, R))} \leq C |v|_{W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1})}.
\end{aligned}$$

Recall that $|\cdot|_{W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1})}$ denotes the Sobolev–Slobodeckij norm.

Proof. The properties of the extension operator for a pyramidal extension domain are discussed in [148, Sec. 2.5.5] and the arguments can be transferred in a straight forward manner, therefore they will only be outlined for convenience. The regularity of w is a standard results of the regularization operator, see, e.g., [148, Sec. 2.3.1.] or [3, Thm. 2.29]. Further, on every hyperplane with x_d fixed, the extension is defined by the convolution with a mollifying kernel whose support is included in a ball of radius x_d , hence the support properties of w are clear as well.

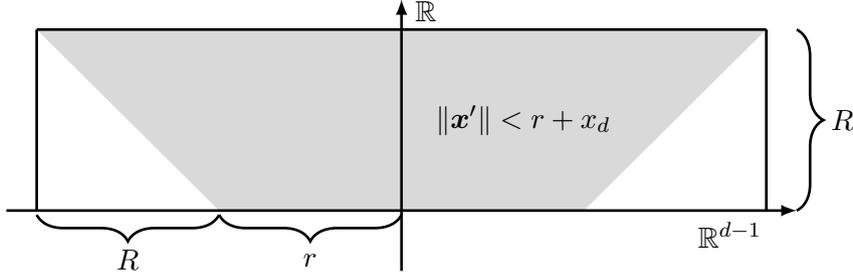


Figure 3.3: Maximum support of the extension w .

We write $Ev := w$ and note that $E: W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1}) \rightarrow W^{1,p}(\mathbb{R}^{d-1} \times (0, R))$ is a linear operator. Boundedness of E follows from a somewhat lengthy but straight forward computation using the properties of the mollifying kernel and the Sobolev-Slobodeckijj norm that closely follows the original arguments in [148, Lem. 2.5.6]. For the sake of completeness, the computations are included in Section C.6, see Lemma C.6.1. Ultimately, one obtains a constant $C(d, r, R, p, \rho) > 0$ such that $\|Ev\|_{W^{1,p}(\mathbb{R}^{d-1} \times (0, R))} \leq C|v|_{W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1})}$.

The remaining trace property is proven using a density argument. To that end, let $v \in C(\mathbb{R}^{d-1}) \cap W^{1-1/p,p}(\mathbb{R}^{d-1})$ satisfy the properties required in the claim and define the extension

$$\tilde{v}(\mathbf{x}', x_d) := \begin{cases} w(\mathbf{x}', x_d) & (\mathbf{x}', x_d) \in \mathbb{R}^{d-1} \times (0, R] \\ v(\mathbf{x}') & (\mathbf{x}', x_d) \in \mathbb{R}^{d-1} \times \{0\} \end{cases},$$

which is in $C^\infty(\mathbb{R}^{d-1} \times (0, R)) \cap C(\mathbb{R}^{d-1} \times [0, R])$. To confirm continuity on $\mathbb{R}^{d-1} \times \{0\}$, let $\mathbb{R}^{d-1} \times (0, R) \ni (\mathbf{x}'_n, y_n) \rightarrow (\mathbf{x}', 0) \in \mathbb{R}^{d-1} \times \{0\}$. If $y_n = 0$ f.a.a. n in \mathbb{N} , then continuity follows from the continuity of v . If there are infinitely many $y_n > 0$, consider the subsequence corresponding to the indices where $y_n > 0$ (denoted by the same symbol),

and we obtain that

$$\begin{aligned}
|\tilde{v}(\mathbf{x}', 0) - w(\mathbf{x}'_n, y_n)| &= |v(\mathbf{x}') - w(\mathbf{x}'_n, y_n)| \\
&= \left| v(\mathbf{x}') - \frac{1}{y_n^{d-1}} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \rho\left(\frac{\mathbf{y}' - \mathbf{x}'_n}{y_n}\right) d\mathbf{y}' \right| \\
&\leq \frac{1}{y_n^{d-1}} \int_{\mathbb{R}^{d-1}} |v(\mathbf{y}') - v(\mathbf{x}')| \rho\left(\frac{\mathbf{y}' - \mathbf{x}'_n}{y_n}\right) d\mathbf{y}' \\
&\leq \sup_{\mathbf{y}' \in B(y_n, \mathbf{x}'_n)} |v(\mathbf{y}') - v(\mathbf{x}')| \frac{1}{y_n^{d-1}} \int_{B(y_n, \mathbf{x}'_n)} \rho\left(\frac{\mathbf{y}' - \mathbf{x}'_n}{y_n}\right) d\mathbf{y}' \\
&= \sup_{\mathbf{y}' \in B(y_n, \mathbf{x}'_n)} |v(\mathbf{y}') - v(\mathbf{x}')| \rightarrow 0.
\end{aligned}$$

The trace property $\text{tr}(w) = v$ is to be understood in the sense of the restriction of w to $\mathcal{D} := B(r+R) \times (0, R)$, which is obviously a bounded Lipschitz domain. According to the above, the function \tilde{v} is in $C(\overline{\mathcal{D}}) \cap W^{1,p}(\mathcal{D})$ and therefore $\text{tr}(\tilde{v}) = \tilde{v}|_{\partial\mathcal{D}}$ with $\text{tr}(Ev) = \text{tr}(\tilde{v}) = v$ \mathcal{L}^{d-1} -a.e. in $B(r+R) \times \{0\}$.

Finally, for any $v \in W^{1-1/p,p}(\mathbb{R}^{d-1})$, the density of the space $C^\infty(\mathbb{R}^{d-1}) \cap W^{1-1/p,p}(\mathbb{R}^{d-1})$ in $W^{1-1/p,p}(\mathbb{R}^{d-1})$ yields an approximating sequence $(v_n) \subset C(\mathbb{R}^{d-1}) \cap W^{1-1/p,p}(\mathbb{R}^{d-1})$ (see [3, Sec. 7.32], [77, Thm. 1.4.2.1]), and we obtain

$$\begin{aligned}
|v - \text{tr}(Ev)|_{W^{1-\frac{1}{p},p}(B(r+R))} &\leq |v - v_n|_{W^{1-\frac{1}{p},p}(B(r+R))} + |v_n - \text{tr}(Ev)|_{W^{1-\frac{1}{p},p}(B(r+R))} \\
&= |v - v_n|_{W^{1-\frac{1}{p},p}(B(r+R))} + |\text{tr}(Ev_n) - \text{tr}(Ev)|_{W^{1-\frac{1}{p},p}(B(r+R))} \\
&\leq |v - v_n|_{W^{1-\frac{1}{p},p}(B(r+R))} + |\text{tr}(Ev_n) - \text{tr}(Ev)|_{W^{1-\frac{1}{p},p}(\partial\mathcal{D})} \\
&\leq (1+C) |v - v_n|_{W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1})} \rightarrow 0.
\end{aligned}$$

Extending the restriction of $\text{tr}(Ev)$ to \mathcal{D} onto $\mathbb{R}^{d-1} \times \{0\}$ by zero, we can also write $v = \text{tr}(Ev) = \text{tr}(w)$ \mathcal{L}^{d-1} -a.e. in \mathbb{R}^{d-1} . \square

A slight modification of the operator in the previous theorem is sufficient to prove the desired inequality in the prototypical case.

Lemma 3.3.10 *Let $B(s)$ denote the open ball in \mathbb{R}^{d-1} with radius $s > 0$ centered at the origin and let*

$$U(s, t) := B(s) \times (-t, t), \quad V(s, t) := B(s) \times (0, t), \quad R(s) := B(s) \times \{0\}.$$

Then for all $r, \varepsilon > 0$, there exists a constant $C > 0$ independent of v such that

$$\text{cap}_{U(3r, 3\varepsilon), 0}(A) \leq C \text{cap}_{\partial V(3r, 3\varepsilon)}(A) \quad \text{for all } A \subset R(r).$$

Proof. Let $r, \varepsilon > 0$ and $A \subset R(r)$ be arbitrary but fixed, and suppose that a function v in $W^{1-1/p,p}(\partial V(3r, 3\varepsilon))$ and a set $G \in \mathcal{O}(\mathbb{R}^d)$ are given such that

$$A \subset G \quad \text{and} \quad v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } G \cap \partial V(3r, 3\varepsilon). \quad (3.21)$$

As we have done in the proof of Lemma 3.3.2, we will use v to construct a function w in $W_0^{1,p}(U(3r, 3\varepsilon))$ that satisfies $w \geq 1$ \mathcal{L}^d -a.e. in a $U(3r, 3\varepsilon)$ -neighborhood of A and $\|w\|_{W^{1,p}(U(3r, 3\varepsilon))} \leq C(r, \varepsilon) \|v\|_{W^{1-1/p,p}(\partial V(3r, 3\varepsilon))}$ for a suitable constant $C(r, \varepsilon)$.

Let $\psi \in C_0^\infty(\mathbb{R}^d)$ be an arbitrary but fixed cut-off function that satisfies

$$\psi \equiv 1 \text{ in } U(r, \varepsilon), \quad 0 \leq \psi \leq 1 \text{ in } U(2r, 2\varepsilon) \setminus U(r, \varepsilon) \quad \text{and} \quad \psi \equiv 0 \text{ in } \mathbb{R}^d \setminus U(2r, 2\varepsilon).$$

Then the function ψv , or rather its extension by zero onto $\mathbb{R}^{d-1} \times \{0\}$, satisfies

$$\psi v \in W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1} \times \{0\}), \quad |\psi v|_{W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1} \times \{0\})} \leq C \|v\|_{W^{1-\frac{1}{p},p}(\partial V(3r, 3\varepsilon))}$$

for a positive constant $C = C(r, \varepsilon)$, cf. [148, Lem. 2.5.5], and

$$\psi v = 0 \text{ in } (\mathbb{R}^{d-1} \times \{0\}) \setminus R(2r), \quad \psi v = v \geq 1 \text{ in } G \cap R(r)$$

both hold in the \mathcal{H}^{d-1} -a.e. sense.

For $(\mathbf{x}', x_d) \in \mathbb{R}^{d-1} \times ((-3\varepsilon, 3\varepsilon) \setminus \{0\})$, we define the extension w of v by

$$w(\mathbf{x}', x_d) := \frac{\psi(\mathbf{x}', x_d)}{|x_d|^{d-1}} \int_{\mathbb{R}^{d-1}} (\psi v)(\mathbf{y}') \rho\left(\frac{\mathbf{y}' - \mathbf{x}'}{|x_d|}\right) d\mathcal{L}^{d-1}(\mathbf{y}'),$$

where ρ denotes a mollifying kernel on \mathbb{R}^{d-1} as in Lemma 3.3.9, so that

$$\begin{aligned} w &\in W^{1,p}(U(3r, 3\varepsilon)), \quad w = 0 \text{ } \mathcal{L}^d\text{-a.e. in } U(3r, 3\varepsilon) \setminus U(2r, 2\varepsilon), \\ \|w\|_{W^{1,p}(U(3r, 3\varepsilon))} &\leq \tilde{C} \|v\|_{W^{1-\frac{1}{p},p}(\partial V(3r, 3\varepsilon))} \end{aligned}$$

for a constant $\tilde{C} = \tilde{C}(r, \varepsilon) > 0$. It remains to check that $w \geq 1$ is satisfied \mathcal{L}^d -a.e. in a neighborhood of A . To this end, for any \mathbf{z}' in $B(r)$, we define

$$g(\mathbf{z}') := \sup \{ s > 0 : (\mathbf{z}' + B(s)) \times \{0\} \subset G \cap R(r) \} \in \{-\infty\} \cup (0, r)$$

and

$$H(\mathbf{z}') := \{ (\mathbf{x}', x_d) : \|\mathbf{x}' - \mathbf{z}'\| < g(\mathbf{z}') - |x_d| \}.$$

Then $H(\mathbf{z}')$ is open for all \mathbf{z}' in $B(r)$ and therefore the set

$$H := U(r, \varepsilon) \cap \bigcup_{\mathbf{z}' \in B(r)} H(\mathbf{z}')$$

is open. The set H contains A because $g(\mathbf{z}') > 0$ for all \mathbf{z}' in $A \subset G \cap R(r)$. For all $(\mathbf{x}', x_d) \in H(\mathbf{z}')$, we further know that

$$\begin{aligned} \rho\left(\frac{\mathbf{y}' - \mathbf{x}'}{|x_d|}\right) \neq 0 &\Rightarrow \left\| \frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right\| \leq 1 \\ &\Rightarrow \|\mathbf{y}' - \mathbf{z}'\| \leq \|\mathbf{y}' - \mathbf{x}'\| + \|\mathbf{x}' - \mathbf{z}'\| < g(\mathbf{z}') \\ &\Rightarrow (\mathbf{y}', 0) \in G \cap R(r), \end{aligned}$$

and accordingly,

$$w(\mathbf{x}', x_d) = \frac{\psi(\mathbf{x}', x_d)}{|x_d|^{d-1}} \int_{\mathbb{R}^{d-1}} (\psi v)(\mathbf{y}') \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{|x_d|} \right) d\mathcal{L}^{d-1}(\mathbf{y}') \geq 1$$

for all (\mathbf{x}', x_d) in $H(\mathbf{z}') \cap U(r, \varepsilon) \setminus (\mathbb{R}^{d-1} \times \{0\})$, and thus $w \geq 1$ \mathcal{L}^d -a.e. in H .

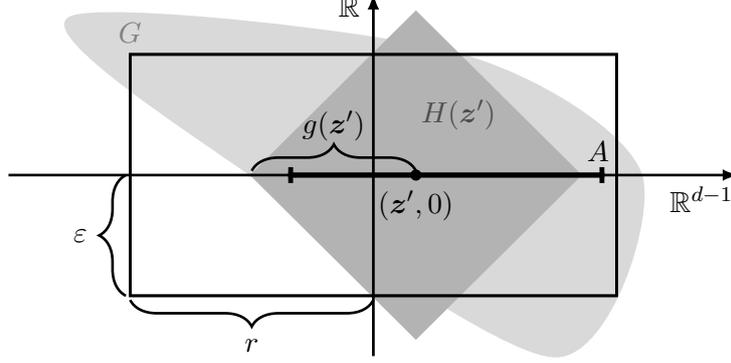


Figure 3.4: The extension-sets $H(\mathbf{z}')$ and their construction.

Consequently, the function w satisfies

$$\begin{aligned} w &\in W_0^{1,p}(U(3r, 3\varepsilon)), \\ \|w\|_{W^{1,p}(U(3r, 3\varepsilon))} &\leq \tilde{C} \|v\|_{W^{1-\frac{1}{p},p}(\partial V(3r, 3\varepsilon))}, \\ w &\geq 1 \text{ } \mathcal{L}^d\text{-a.e. in } H, \quad H \subset U(3r, 3\varepsilon) \text{ open, } A \subset H. \end{aligned} \quad (3.22)$$

By (3.22) and taking the infimum over all v in $W^{1-1/p,p}(\partial V(3r, 3\varepsilon))$ that satisfy (3.21) for some open set G , we obtain

$$\begin{aligned} \text{cap}_{U(3r, 3\varepsilon), 0}(A) &= \inf \{ \|w\|_{W^{1,p}(U(3r, 3\varepsilon))} : w \in W_0^{1,p}(U(3r, 3\varepsilon)) \text{ and} \\ &\quad w \geq 1 \text{ } \mathcal{L}^d\text{-a.e. in a } \overline{U(3r, 3\varepsilon)}\text{-nbhd. of } A \} \\ &\leq \tilde{C} \inf \{ \|v\|_{W^{1-\frac{1}{p},p}(\partial V(3r, 3\varepsilon))} : v \in W^{1-\frac{1}{p},p}(\partial V(3r, 3\varepsilon)) \text{ and} \\ &\quad v \geq 1 \text{ } \mathcal{H}^{d-1}\text{-a.e. in a} \\ &\quad \partial V(3r, 3\varepsilon)\text{-nbhd. of } A \} \\ &= \tilde{C} \text{cap}_{V(3r, 3\varepsilon)}(A) \end{aligned}$$

with a constant $\tilde{C} = \tilde{C}(r, \varepsilon)$. □

To obtain a result for Lipschitz boundaries from the prototypical Lemma 3.3.10, we again use a stability property for the boundary capacity similar to that in Lemma 3.3.3.

Lemma 3.3.11 *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^d$ be bounded strong Lipschitz domains and let $\Theta: \overline{\Omega_1} \rightarrow \overline{\Omega_2}$ be a bi-Lipschitz mapping with $\Theta(\overline{\Omega_1}) = \overline{\Omega_2}$ and $\Theta(\Omega_1) = \Omega_2$. Then there exist constants $c, C > 0$ depending only on Θ such that*

$$c \text{cap}_{\partial\Omega_1}(A) \leq \text{cap}_{\partial\Omega_2}(\Theta(A)) \leq C \text{cap}_{\partial\Omega_1}(A) \quad (3.23)$$

for all $A \subset \partial\Omega_1$.

Proof. Again, it suffices to show one of the two inequalities, the other one follows from interchanging the roles of Θ and Θ^{-1} . Note that since $\Theta(\Omega_1) = \Omega_2$, we know that $\Theta(\partial\Omega_1) = \partial\Omega_2$ as well and $\Theta|_{\partial\Omega_1}, \Theta|_{\partial\Omega_2}^{-1}$ are homeomorphisms as well. We show the left one of the inequalities.

Let $A \subset \partial\Omega_1$, then $\Theta(A) \subset \partial\Omega_2$ and assume that G in $\mathcal{O}(\partial\Omega_2)$ and v in $W^{1-\frac{1}{p},p}(\partial\Omega_2)$ are given such that

$$\Theta(A) \subset G \quad \text{and} \quad v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } G.$$

Then $w := v \circ \Theta \in W^{1-\frac{1}{p},p}(\partial\Omega_1)$ and $\|w\|_{W^{1,p}(\Omega_1)} \leq C \|v\|_{W^{1,p}(\Omega_2)}$ for a constant $C = C(\Theta)$, see Lemma A.2.9, as well as

$$A \subset \Theta^{-1}(G) \in \mathcal{O}(\partial\Omega_1) \quad \text{and} \quad w \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } \Theta^{-1}(G).$$

Accordingly,

$$\begin{aligned} \text{cap}_{\partial\Omega_1}(A) &= \inf \{ \|w\|_{W^{1-\frac{1}{p},p}(\partial\Omega_1)} : w \in W^{1-\frac{1}{p},p}(\partial\Omega_1), \exists G \in \mathcal{O}(\partial\Omega_1) \text{ s.t. } A \subset G, \\ &\quad w \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } G \} \\ &\leq \inf \{ \|v \circ \Theta\|_{W^{1-\frac{1}{p},p}(\partial\Omega_1)} : v \in W^{1-\frac{1}{p},p}(\partial\Omega_2), \exists G \in \mathcal{O}(\partial\Omega_2) \text{ s.t.} \\ &\quad \Theta(A) \subset G, v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } G \} \\ &\leq C \inf \{ \|v\|_{W^{1-\frac{1}{p},p}(\partial\Omega_2)} : v \in W^{1-\frac{1}{p},p}(\partial\Omega_2), \exists G \in \mathcal{O}(\partial\Omega_2) \text{ s.t. } \Theta(A) \subset G, \\ &\quad v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } G \} \\ &\leq C \text{cap}_{\partial\Omega_2}(\Theta(A)). \end{aligned}$$

Dividing by C proves (3.23) with $c = 1/C$. \square

Combining the previous lemmas with the same localization and transformation arguments as in the last subsection yields the remaining inequality in (3.20) for the general case.

Proposition 3.3.12 *There exists a constant $C = C(\Omega, \Omega') > 0$ such that*

$$\text{cap}_{\Omega',0}(A) \leq C \text{cap}_{\partial\Omega}(A) \quad \forall A \subset \partial\Omega.$$

Proof. We follow the same steps as in the proof of Proposition 3.3.6.

Part 1 (Decomposition): Assume that the quantities $\mathbf{q}_i, R_{\mathbf{q}_i}, B_{\mathbf{q}_i}, \Theta_{\mathbf{q}_i}$ and $\tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}), \tilde{V}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i})$ are chosen the same way as in the proof of Proposition 3.3.6 for $i = 1, \dots, n$, see (3.18) especially.

Let $A \subset \partial\Omega$ be arbitrary but fixed and define

$$A_i := A \cap \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}), \quad i = 1, \dots, n,$$

then $A_i \subset \partial\Omega \cap \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i})$ for all i in $\{1, \dots, n\}$. Note that since $A \subset \partial\Omega$, no remaining inner part needs to be considered.

Part 2 (Local Estimates): Like in Proposition 3.3.6, we want to show that there exist constants $C_i = C_i(\Omega, \Omega') > 0$ such that

$$\text{cap}_{\Omega', 0}(A_i) \leq C_i \text{cap}_{\partial\Omega}(A_i) \quad \forall i \in \{1, \dots, n\}.$$

To this end, fix an $i \in \{1, \dots, n\}$ and assume $R_{\mathbf{q}_i} = \mathbf{I}$, w.l.o.g. The stability results in Lemmas 3.3.3 and 3.3.11 yield a constant $C = C(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}, h_{\mathbf{q}_i}) > 0$ such that

$$\begin{aligned} \text{cap}_{\tilde{U}_{\mathbf{q}_i}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i}), 0}(A_i) &\leq C \text{cap}_{U(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i}), 0}(\Theta_{\mathbf{q}_i}^{-1}(A_i)), \\ \text{cap}_{\partial V(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})}(\Theta_{\mathbf{q}_i}^{-1}(A_i)) &\leq C \text{cap}_{\partial\tilde{V}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})}(A_i). \end{aligned}$$

For the transformed domains with $\Theta_{\mathbf{q}_i}^{-1}(A_i) \subset R(r_{\mathbf{q}_i})$, the prototypical result in Lemma 3.3.10 implies the existence of a constant $\tilde{C} = \tilde{C}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}) > 0$ such that

$$\text{cap}_{U(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i}), 0}(\Theta_{\mathbf{q}_i}^{-1}(A_i)) \leq \tilde{C} \text{cap}_{\partial V(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})}(\Theta_{\mathbf{q}_i}^{-1}(A_i)).$$

By combining the last three inequalities and the same extension-by-zero argument as in (3.19), we obtain the existence of a constant $\tilde{C}_i = \tilde{C}_i(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}, h_{\mathbf{q}_i}) > 0$ with

$$\text{cap}_{\Omega', 0}(A_i) \leq \text{cap}_{\tilde{U}_{\mathbf{q}_i}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i}), 0}(A_i) \leq \tilde{C}_i \text{cap}_{\partial\tilde{V}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})}(A_i). \quad (3.24)$$

Finally, let $v \in W^{1-\frac{1}{p}, p}(\partial\Omega)$ and an open set $G \in \mathcal{O}(\mathbb{R}^d)$ be given such that

$$A_i \subset G \quad \text{and} \quad v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in } G \cap \partial\Omega,$$

and choose a cut-off function $\psi_i \in C_0^\infty(\mathbb{R}^d)$ that satisfies

$$\begin{aligned} \psi_i &\equiv 1 \text{ in } \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}), \quad 0 \leq \psi_i \leq 1 \text{ in } \tilde{U}_{\mathbf{q}_i}(2r_{\mathbf{q}_i}, 2\varepsilon_{\mathbf{q}_i}) \setminus \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}) \quad \text{and} \\ \psi_i &\equiv 0 \text{ in } \mathbb{R}^d \setminus \tilde{U}_{\mathbf{q}_i}(2r_{\mathbf{q}_i}, 2\varepsilon_{\mathbf{q}_i}). \end{aligned}$$

Then the restriction of $\psi_i v \in W^{1-\frac{1}{p}, p}(\partial\Omega)$ to $\partial\Omega \cap \tilde{U}_{\mathbf{q}_i}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})$ satisfies

$$\psi_i v = 0 \text{ in } \partial\Omega \setminus \tilde{U}_{\mathbf{q}_i}(2r_{\mathbf{q}_i}, 2\varepsilon_{\mathbf{q}_i}), \quad \psi_i v = v \geq 1 \text{ in } G \cap \tilde{U}_{\mathbf{q}_i}(r_{\mathbf{q}_i}, \varepsilon_{\mathbf{q}_i}) \cap \partial\Omega$$

in the \mathcal{H}^{d-1} -a.e. sense and the extension to $\partial\tilde{V}_{\mathbf{q}_i}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})$ by zero satisfies

$$\psi_i v \in W^{1-\frac{1}{p}, p}(\partial V_{\mathbf{q}_i}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})), \quad \|\psi_i v\|_{W^{1-\frac{1}{p}, p}(\partial\tilde{V}_{\mathbf{q}_i}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i}))} \leq \tilde{C} \|v\|_{W^{1-\frac{1}{p}, p}(\partial\Omega)}$$

for a positive constant $\tilde{C} = \tilde{C}(\Omega)$, cf. the proof in [148, Lem. 2.5.5]. Accordingly, we obtain

$$\begin{aligned} \text{cap}_{\partial\tilde{V}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})}(A_i) &= \inf \left\{ \|v\|_{W^{1-\frac{1}{p}, p}(\partial\tilde{V}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i}))} : v \in W^{1-\frac{1}{p}, p}(\partial\tilde{V}(3r_{\mathbf{q}_i}, 3\varepsilon_{\mathbf{q}_i})), \right. \\ &\quad \left. v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in a nbhd. of } A_i \right\} \\ &\leq \inf \left\{ \|\psi_i v\|_{W^{1-\frac{1}{p}, p}(\partial\Omega)} : v \in W^{1-\frac{1}{p}, p}(\partial\Omega), \right. \\ &\quad \left. v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in a nbhd. of } A_i \right\} \\ &\leq \tilde{C}(\Omega) \inf \left\{ \|v\|_{W^{1-\frac{1}{p}, p}(\partial\Omega)} : v \in W^{1-\frac{1}{p}, p}(\partial\Omega), \right. \\ &\quad \left. v \geq 1 \quad \mathcal{H}^{d-1}\text{-a.e. in a nbhd. of } A_i \right\} \\ &= \tilde{C}(\Omega) \text{cap}_{\partial\Omega}(A_i), \end{aligned}$$

where the neighborhoods mean the relative neighborhoods on the corresponding domains and subdomains. The combination with (3.24) concludes Part 2 of the proof.

Part 3 (Recombination): In the previous part, we have established the existence of constants $C_i = C_i(\Omega, \Omega') > 0$ for $i = 0, \dots, n$ such that

$$\text{cap}_{\Omega',0}(A_i) \leq C_i \text{cap}_{\partial\Omega}(A_i).$$

The monotonicity and the subadditivity properties of the capacities, see Lemma 3.1.6 (b) and (d), therefore yield that

$$\text{cap}_{\Omega',0}(A) \leq \sum_{i=1}^n \text{cap}_{\Omega',0}(A_i) \leq \sum_{i=1}^n C_i \text{cap}_{\partial\Omega}(A_i) \leq n \max_{i=1,\dots,n} (C_i) \text{cap}_{\partial\Omega}(A). \quad \square$$

As (3.20) is proven, we can state the main result of this subsection.

Theorem 3.3.13 *Let Assumption 3.2.1 hold, and let cap_Ω , $\text{cap}_{\mathbb{R}^d}$, $\text{cap}_{\Omega',0}$ and $\text{cap}_{\partial\Omega}$ be defined as in (3.4). Then there exists a constant $C(\Omega, \Omega') > 0$ such that*

$$\text{cap}_{\partial\Omega}(A) \leq \text{cap}_\Omega(A) \leq \text{cap}_{\mathbb{R}^d}(A) \leq \text{cap}_{\Omega',0}(A) \leq C \text{cap}_{\partial\Omega}(A) \quad \forall A \subset \partial\Omega.$$

Proof. Lemma 3.3.8 and Proposition 3.3.12 imply (3.20). Lemma 3.3.1 yields the claim. \square

3.4 Conclusions and Implications

We will now collect the results from the previous technical sections and develop some interesting implications for the boundary traces of Sobolev functions. First of all, the equivalences established in the previous sections amount to the following Theorem.

Theorem 3.4.1 (Equivalence of Sobolev Capacities) *Let Assumption 3.2.1 hold, and let cap_Ω , $\text{cap}_{\Omega,D}$, $\text{cap}_{\mathbb{R}^d}$, $\text{cap}_{\Omega',0}$ and $\text{cap}_{\partial\Omega}$ be defined as in (3.4). Then there exists a constant $C(\Omega, \Omega') > 0$ such that*

$$\begin{aligned} \text{cap}_\Omega(A) \leq \text{cap}_{\mathbb{R}^d}(A) &\leq \text{cap}_{\Omega',0}(A) \leq C \text{cap}_\Omega(A) && \forall A \subset \bar{\Omega}, \\ \text{cap}_{\partial\Omega}(A) \leq \text{cap}_\Omega(A) &\leq \text{cap}_{\mathbb{R}^d}(A) \leq \text{cap}_{\Omega',0}(A) \leq C \text{cap}_{\partial\Omega} && \forall A \subset \partial\Omega, \end{aligned}$$

and, additionally,

$$\text{cap}_\Omega(A) \leq \text{cap}_{\Omega,D}(A) \leq \left(1 + \frac{d^{\frac{1}{p}}}{\text{dist}(A, \Gamma_D)} \right) \text{cap}_\Omega(A)$$

for all $A \subset \bar{\Omega}$ that satisfy $\text{dist}(A, \Gamma_D) > 0$ or $\text{cap}_\Omega(A) > 0$.

Proof. The claim for $A \subset \bar{\Omega}$ follows from Theorem 3.3.7 and Proposition 3.2.6. See Theorem 3.3.13 for the case $A \subset \partial\Omega$. \square

Several corollaries can be drawn from these equivalences. First, note that $\text{cap}_{\Omega,D}$ and the other capacities in Theorem 3.4.1 are generally not equivalent on $\mathcal{P}(\bar{\Omega})$, as discussed after Lemma 3.2.5. However, the previous Theorem ensures equivalence for sets that maintain a uniform distance to Γ_D .

Corollary 3.4.2 *Let Assumption 3.2.1 hold and let $U \subset \overline{\Omega}$ with $\text{dist}(U, \Gamma_D) > 0$. Then there exist two constants $C(\Omega, \Omega'), C_D(\Omega, \Omega', U) > 0$ such that*

$$\text{cap}_\Omega(A) \leq \text{cap}_{\mathbb{R}^d}(A) \leq \text{cap}_{\Omega', 0}(A) \leq C \text{cap}_{\Omega, D}(A) \leq C_D \text{cap}_\Omega(A) \quad \forall A \subset U.$$

Proof. Since $\text{dist}(U, \Gamma_D) \leq \text{dist}(A, \Gamma_D)$ for all $A \subset U$, the equivalence is an immediate consequence of Theorem 3.4.1 with $C_D = C \left(1 + \frac{d^{\frac{1}{p}}}{\text{dist}(U, \Gamma_D)}\right)$. \square

In the application to contact problems, this is especially relevant whenever the distance between the contact boundary and the Dirichlet boundary is positive, since all of the capacities mentioned above are then equivalent for subsets of the contact boundary.

Trivially, the equivalence result in Corollary 3.4.2 above implies that when the distance requirement is satisfied, a set $A \subset \overline{\Omega}$ is a polar set for any one of the capacities if and only if it is a polar set with respect to all of the equivalent capacities. While Theorem 3.2.8 showed that equivalence can not hold for arbitrary sets when the distance requirement for the sets and the Dirichlet boundary is dropped, the conditions can easily be refined for polar sets.

Corollary 3.4.3 *Let Assumption 3.2.1 hold. A set $A \subset \overline{\Omega} \setminus \overline{\Gamma_D}$ is a polar set w.r.t. to one of the capacities cap_Ω , $\text{cap}_{\Omega, D}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\Omega', 0}$ if and only if it is a polar set w.r.t. all four.*

A set $A \subset \partial\Omega \setminus \overline{\Gamma_D}$ is a polar set w.r.t. to one of the capacities cap_Ω , $\text{cap}_{\Omega, D}$, $\text{cap}_{\mathbb{R}^d}$, $\text{cap}_{\Omega', 0}$ and $\text{cap}_{\partial\Omega}$ if and only if it is a polar set w.r.t. all five.

Proof. Due to Theorem 3.4.1, it suffices to check that every set $A \subset \overline{\Omega} \setminus \overline{\Gamma_D}$ with $\text{cap}_\Omega(A) = 0$ satisfies $\text{cap}_{\Omega, D}(A) = 0$. Let such an A be given. Since $\overline{\Gamma_D}$ is a closed set, every element of the complement $\overline{\Omega} \setminus \overline{\Gamma_D}$ has positive distance to Γ_D and we may write

$$A = \bigcup_{n=1}^{\infty} A_n \quad \text{where} \quad A_n := \left\{ \mathbf{x} \in A : \text{dist}(\mathbf{x}, \Gamma_D) \geq \frac{1}{n} \right\}.$$

The subadditivity of $\text{cap}_{\Omega, D}$ and the monotonicity of cap_Ω , see Lemma 3.1.6 (d) and (b), imply that

$$\text{cap}_{\Omega, D}(A) = \text{cap}_{\Omega, D} \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \text{cap}_{\Omega, D}(A_n) \leq \sum_{n=1}^{\infty} \left(1 + d^{\frac{1}{p}} n\right) \text{cap}_\Omega(A_n) = 0,$$

which proves the claim. \square

Like in Proposition 3.2.6, we can generally not expect the previous result for sets that satisfy $A \subset \overline{\Omega} \setminus \Gamma_D$ instead of $A \subset \overline{\Omega} \setminus \overline{\Gamma_D}$. Conflicts arise for cap_Ω -polar sets A that intersect $\overline{\Gamma_D}$ and are clearly not $\text{cap}_{\Omega, D}$ -polar since $\text{cap}_{\Omega, D}(A) = \infty$.

A direct consequence of the results presented above is the following generalization of the results concerning equivalence of polar sets in [194, Lem. A.2] and [88, Cor. 2.39].

Corollary 3.4.4 (Equivalent Notions of Quasi Everywhere) *Let Assumption 3.2.1 hold, then the associated notions of quasi everywhere satisfy the following:*

- (a) *A condition holds q.e. in $\overline{\Omega}$ w.r.t. one of the capacities cap_{Ω} , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$ if and only if it holds q.e. in $\overline{\Omega}$ w.r.t. all three.*
- (b) *A condition holds q.e. in $\overline{\Omega} \setminus \overline{\Gamma_D}$ w.r.t. one of the capacities cap_{Ω} , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\Omega,D}$ if and only if it holds q.e. in $\overline{\Omega} \setminus \overline{\Gamma_D}$ w.r.t. all four.*
- (c) *A condition holds q.e. in $\partial\Omega$ w.r.t. one of the capacities cap_{Ω} , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\partial\Omega}$ if and only if it holds q.e. in $\partial\Omega$ w.r.t. all four.*
- (d) *A condition holds q.e. in $\partial\Omega \setminus \overline{\Gamma_D}$ w.r.t. one of the capacities cap_{Ω} , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$, $\text{cap}_{\Omega,D}$ and $\text{cap}_{\partial\Omega}$ if and only if it holds q.e. in $\partial\Omega \setminus \overline{\Gamma_D}$ w.r.t. all five.*

Proof. Theorem 3.4.1 and Corollary 3.4.3 immediately yield the claim. \square

Following the conclusions for the q.e. senses with respect to the equivalent capacities, the remainder of this section is dedicated to establishing corresponding results for quasi continuous representatives. The first result is an interesting connection between Sobolev functions and their traces on the boundary.

Corollary 3.4.5 (Quasi Continuous Traces) *Let Assumption 3.2.1 (a), (b) and (d) be satisfied and $E: W^{1,p}(\Omega) \rightarrow W^{1,p}(\mathbb{R}^d)$ be a continuous extension operator such that $E(W^{1,p}(\Omega) \cap C(\overline{\Omega})) \subseteq C(\mathbb{R}^d)$. Then*

$$v = Ev = \text{tr}(v) \quad \text{q.e. on } \partial\Omega \text{ for all } v \in W^{1,p}(\Omega)$$

in the sense that for every cap_{Ω} -quasi continuous representative $\tilde{v}: \overline{\Omega} \rightarrow \mathbb{R}$ of $v \in W^{1,p}(\Omega)$, every $\text{cap}_{\partial\Omega}$ -quasi continuous representative $\widetilde{\text{tr}(v)}: \partial\Omega \rightarrow \mathbb{R}$ of $\text{tr}(v) \in W^{1-1/p,p}(\partial\Omega)$ and every $\text{cap}_{\mathbb{R}^d}$ -quasi continuous representative $\widetilde{Ev}: \mathbb{R}^d \rightarrow \mathbb{R}$ of $Ev \in W^{1,p}(\mathbb{R}^d)$, there exists a set $N \subseteq \partial\Omega$ such that

$$\text{cap}_{\Omega}(N) = \text{cap}_{\Omega',0}(N) = \text{cap}_{\mathbb{R}^d}(N) = \text{cap}_{\partial\Omega}(N) = 0$$

and

$$\tilde{v}(\mathbf{x}) = \widetilde{\text{tr}(v)}(\mathbf{x}) = (\widetilde{Ev})(\mathbf{x}) \quad \forall \mathbf{x} \in \partial\Omega \setminus N.$$

Proof. Let $(v_n) \subset W^{1,p}(\Omega) \cap C(\overline{\Omega})$ be a sequence such that $v_n \rightarrow v$ in $W^{1,p}(\Omega)$. Then we know that

$$\begin{aligned} C(\partial\Omega) \cap W^{1-\frac{1}{p},p}(\partial\Omega) \ni \text{tr } v_n &\rightarrow \text{tr } v \in W^{1-\frac{1}{p},p}(\partial\Omega), \\ C(\mathbb{R}^d) \cap W^{1,p}(\mathbb{R}^d) \ni Ev_n &\rightarrow Ev \in W^{1,p}(\mathbb{R}^d). \end{aligned}$$

Corollary 3.1.14 yields the existence of a subsequence, denoted by the same symbol, such that

$$\begin{aligned} v_n(\mathbf{x}) &\rightarrow \tilde{v}(\mathbf{x}) \quad \forall \mathbf{x} \in \overline{\Omega} \setminus N_1, \\ \text{tr}(v_n)(\mathbf{x}) &\rightarrow \widetilde{\text{tr}(v)}(\mathbf{x}) \quad \forall \mathbf{x} \in \partial\Omega \setminus N_2, \\ (Ev_n)(\mathbf{x}) &\rightarrow (\widetilde{Ev})(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^d \setminus N_3, \end{aligned}$$

where \tilde{v} , $\widetilde{\text{tr}(v)}$ and (\widetilde{Ev}) are quasi continuous representatives of v , $\text{tr } v$ and Ev , respectively, and where $N_1 \subseteq \overline{\Omega}$, $N_2 \subseteq \partial\Omega$ and $N_3 \subseteq \mathbb{R}^d$ satisfy

$$\text{cap}_\Omega(N_1) = \text{cap}_{\partial\Omega}(N_2) = \text{cap}_{\mathbb{R}^d}(N_3) = 0,$$

and the subadditivity of the capacities and Corollary 3.1.8 imply that the set $N = (N_1 \cup N_2 \cup N_3) \cap \partial\Omega$ is a polar set with respect to all of the involved capacities. The claim follows since $v_n = \text{tr}(v_n) = Ev_n$ on $\partial\Omega$ for all n and because quasi continuous representatives are unique up to polar sets, see Theorem 3.1.17. \square

Corollary 3.4.5 shows that the restriction of a cap_Ω -quasi continuous representative of a $W^{1,p}(\Omega)$ -function to the boundary $\partial\Omega$ is always a $\text{cap}_{\partial\Omega}$ -quasi continuous representative of the trace, a result that accords very well with the intuition. See [208, Rem. 4.4.5, Ex. 4.2] for an approach to the trace operator by extensions of Sobolev functions that are unique up to sets of \mathcal{H}^{d-1} -measure zero.

Remark 3.4.6 *An example of extension operators that Corollary 3.4.5 may be applied to are so-called total extension operators, as defined in [3, Def. 5.17], which map $C^m(\overline{\Omega})$ to $C^m(\mathbb{R}^d)$. The existence of total extension operators for strong bounded Lipschitz domains is ensured by Stein's extension theorem [184, Chap. 6], see also [3, Thm. 5.24].*

Of course, there is an analogous result to Corollary 3.4.4 for the quasi continuity of functions.

Corollary 3.4.7 (Equivalences of Quasi Continuity) *Let Assumption 3.2.1 be satisfied, $U \subset \overline{\Omega}$, $V \subset \partial\Omega$ such that $\text{dist}(U, \Gamma_D), \text{dist}(V, \Gamma_D) > 0$ as well as $u, \tilde{u}: \overline{\Omega} \rightarrow \overline{\mathbb{R}}$ and $v: \partial\Omega \rightarrow \overline{\mathbb{R}}$. Then:*

- (a) *u is quasi continuous on $\overline{\Omega}$ w.r.t. one of the capacities cap_Ω , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$ if and only if it is quasi continuous on $\overline{\Omega}$ w.r.t. all three.*
- (b) *u is quasi continuous on U w.r.t. one of the capacities $\text{cap}_{\Omega,D}$, cap_Ω , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$ if and only if it is quasi continuous on U w.r.t. all four.*
- (c) *v is quasi continuous on $\partial\Omega$ w.r.t. one of the capacities cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\partial\Omega}$ if and only if it is quasi continuous on $\partial\Omega$ w.r.t. all four.*
- (d) *v is quasi continuous on V w.r.t. one of the capacities $\text{cap}_{\Omega,D}$, cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\partial\Omega}$ if and only if it is quasi continuous on V w.r.t. all five.*

Additionally:

- (e) *If u is quasi continuous on $\overline{\Omega}$ w.r.t. $\text{cap}_{\Omega,D}$, then it is quasi continuous on $\overline{\Omega}$ w.r.t. cap_Ω , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$.*
- (f) *If u is quasi continuous on $\overline{\Omega}$ w.r.t. $\text{cap}_{\Omega,D}$ and \tilde{u} is quasi continuous on $\overline{\Omega}$ w.r.t. cap_Ω and $u = \tilde{u}$ \mathcal{L}^d -a.e. on $\overline{\Omega}$, then $u = \tilde{u}$ q.e. on $\overline{\Omega} \setminus \overline{\Gamma_D}$ w.r.t. $\text{cap}_{\Omega,D}$, cap_Ω , $\text{cap}_{\Omega',0}$ and $\text{cap}_{\mathbb{R}^d}$.*

Proof. Parts (a), (c) and (e) are direct consequences of the definition and Theorem 3.4.1.

To confirm part (b), it suffices to check that if u is quasi continuous on U w.r.t. cap_Ω , then it is quasi continuous on U w.r.t. $\text{cap}_{\Omega,D}$, due to Theorem 3.4.1. Let such a u be given and assume that $\varepsilon > 0$. By definition, there exists $G_\varepsilon \in \mathcal{O}(\overline{\Omega})$ such that

$$\text{cap}_\Omega(G_\varepsilon) < \varepsilon \quad \text{and} \quad u: U \setminus G_\varepsilon \rightarrow \overline{\mathbb{R}} \text{ is continuous.}$$

Since the claim holds trivially for empty Γ_D , assume w.l.o.g. that $\Gamma_D \neq \emptyset$, let $\delta := \text{dist}(U, \Gamma_D)/2 > 0$ and consider the relatively open superset of U given by

$$U_\delta := \bigcup_{\mathbf{x} \in U} B(\delta, \mathbf{x}) \cap \bar{\Omega} \in \mathcal{O}(\bar{\Omega}) \quad \text{with} \quad \text{dist}(U_\delta, \Gamma_D) > 0.$$

The set $\tilde{G}_\varepsilon := G_\varepsilon \cap U_\delta \in \mathcal{O}(\bar{\Omega})$ and $\tilde{G}_\varepsilon \subset U_\delta$. Corollary 3.4.2 implies that there exists a constant $C_D(\Omega, U, \Gamma_D) > 0$ independent of G_ε such that

$$\text{cap}_{\Omega, D}(\tilde{G}_\varepsilon) < C_D \varepsilon \quad \text{and} \quad u: U \setminus \tilde{G}_\varepsilon \rightarrow \bar{\mathbb{R}} \text{ is continuous.}$$

Accordingly, the function u is quasi continuous on U w.r.t. $\text{cap}_{\Omega, D}$ as well, which proves part (b). An analogous argument shows (d).

To confirm part (f), consider u and \tilde{u} as stated in the claim. Part (e) implies that both u and \tilde{u} are quasi continuous w.r.t. cap_Ω . Due to Lemma 3.1.15, they coincide q.e. on $\bar{\Omega}$ w.r.t. cap_Ω and therefore coincide q.e. on $\bar{\Omega} \setminus \bar{\Gamma}_D$ w.r.t. $\text{cap}_{\Omega, D}$, cap_Ω , $\text{cap}_{\Omega', 0}$ and $\text{cap}_{\mathbb{R}^d}$, see Corollary 3.4.4. \square

The previous Corollary ensures that for functions in $W_D^{1,p}(\Omega)$, representatives that are $\text{cap}_{\Omega, D}$ -quasi continuous are also cap_Ω -quasi continuous representatives. Further, every cap_Ω -quasi continuous representative is $\text{cap}_{\Omega, D}$ quasi continuous when restricted to subsets with positive distance to Γ_D and coincides with any $\text{cap}_{\Omega, D}$ -quasi continuous representative q.e. in $\bar{\Omega} \setminus \bar{\Gamma}_D$. Note that all $\text{cap}_{\Omega, D}$ -quasi continuous representatives are necessarily continuous (and therefore constant on $\bar{\Gamma}_D$), which is satisfied only up to cap_Ω -polar sets for cap_Ω -quasi continuous representatives.

The absence of the exponent p in Definition 3.1.3 of the abstract capacity was shortly addressed in Remark 3.1.4. Including the exponent p in the definition generally leads to loss of subadditivity — a result that is well known to hold in the Sobolev case regardless, e.g. [88, Thm. 2.2, Thm. 2.37]. The relation between the definitions including and excluding the exponent p , respectively, is quite obvious.

Lemma 3.4.8 (Sobolev Capacities with Power p) *Let Assumption 3.1.1 be satisfied, then*

$$\inf \{ \|v\|_V^p : v \in V, v \geq 1 \text{ } \mu\text{-a.e. in a nbhd. of } A \} = \text{cap}(A; X, V, \mu)^p \quad \forall A \subset X.$$

Proof. This is an immediate consequence of the monotonicity and continuity of the exponential function. \square

Accordingly, both definitions yield the same notion of quasi continuity and therefore the same quasi continuous representatives, the same polar sets and the same sense of quasi everywhere, cf. Definition 3.1.10. The trace results in Corollary 3.4.5 remain valid when the alternative definition of the capacity (including the exponent p) is considered and Corollaries 3.4.3, 3.4.4 and 3.4.7 can even be extended to include the alternative definitions as well. The equivalency results in Theorem 3.4.1 and Corollary 3.4.2 carry over to the alternative definition with different constants, again due to the monotonicity of the exponential function by raising both sides of the inequalities to the exponent p . This implies that the

approaches to the differential sensitivity analysis of Signorini-type problems employed in [138, Ex. 2, P. 150], [85, Ex. 6] and [28, Sec. 3.2] are equivalent.

The considerations in this section allow for the following extension of Definition 3.1.18, which allows us to simply require that Sobolev functions satisfy a pointwise condition q.e. on a set without having to specify the sense further.

Definition 3.4.9 (Generalized Quasi Everywhere Behavior) *Let Assumption 3.2.1 hold.*

A function $v \in W^{1,p}(\Omega)$ satisfies a pointwise condition quasi everywhere on $A \subset \overline{\Omega}$ if the respective condition is satisfied quasi everywhere on A by all quasi continuous representatives of v with respect to cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and their counterparts involving the exponent p .

A function $v \in W_D^{1,p}(\Omega)$ satisfies a pointwise condition quasi everywhere on $A \subset \overline{\Omega}$ with $\text{dist}(A, \Gamma_D) > 0$ if the respective condition is satisfied quasi everywhere on A by all quasi continuous representatives of v with respect to $\text{cap}_{\Omega,D}$, cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and their counterparts involving the exponent p .

A function $v \in W^{1-1/p,p}(\Omega)$ satisfies a pointwise condition quasi everywhere on $A \subset \partial\Omega$ if the respective condition is satisfied quasi everywhere on A by all quasi continuous representatives of v with respect to cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and their counterparts involving the exponent p .

The last result of this section is the following alternative characterization of zero boundary conditions in the q.e. sense, which will be helpful later on.

Corollary 3.4.10 *Let Assumption 3.2.1 be satisfied. Then there exists a set $\widetilde{\Gamma}_D \subset \partial\Omega$ such that $\Gamma_D \subset \widetilde{\Gamma}_D \subset \overline{\Gamma}_D$ and*

$$\begin{aligned} W_D^{1,p}(\Omega) &:= \{ v \in W^{1,p}(\Omega) : \text{tr}(v) = 0 \text{ } \mathcal{H}^{d-1}\text{-a.e. on } \Gamma_D \} \\ &= \{ v \in W^{1,p}(\Omega) : \text{tr}(v) = 0 \text{ q.e. on } \Gamma_D \} \\ &= \{ v \in W^{1,p}(\Omega) : v = 0 \text{ q.e. on } \Gamma_D \} \\ &= \{ v \in W^{1,p}(\Omega) : v = 0 \text{ q.e. on } \widetilde{\Gamma}_D \}. \end{aligned}$$

The set $\widetilde{\Gamma}_D$ is quasi closed in $\partial\Omega$ and unique up to polar sets w.r.t. cap_Ω , $\text{cap}_{\Omega',0}$, $\text{cap}_{\mathbb{R}^d}$ and $\text{cap}_{\partial\Omega}$.

Proof. The first equality is a direct consequence of Lemma 3.1.15 applied to the representatives of $\text{tr}(v) \in W^{1-1/p,p}(\partial\Omega)$.

The second equality follows since the restrictions of the quasi continuous representatives of $v \in W^{1,p}(\Omega)$ to $\partial\Omega$ coincide with quasi continuous representatives of $\text{tr}(v)$ up to polar sets, see Corollary 3.4.5.

To confirm the last equality, note that one set inclusion is trivial since $\Gamma_C \subset \widetilde{\Gamma}_C$ is claimed. For the remaining inclusion, let $v \in W_D^{1,p}(\Omega)$, which then satisfies $v = 0$ q.e. on Γ_D . Since $W^{1,p}(\Omega)$ is a separable normed space, there exists a countable family of functions $(v_n) \subset W_D^{1,p}(\Omega)$ that is dense in $W_D^{1,p}(\Omega)$. For every $n \in \mathbb{N}$, the set $v_n^{-1}(\{0\})$ is a quasi closed

set that is determined up to polar sets and contains Γ_D up to polar sets w.r.t. to either of the capacities, cf. Lemma B.3.1. Accordingly, the intersection $\widetilde{\Gamma}_D := \bigcap_{n=0}^{\infty} v_n^{-1}(\{0\}) \cap \partial\Omega$ is quasi closed as well as a superset of Γ_D and determined up to polar sets, see Lemma B.2.3, and $v_n = 0$ q.e. on $\widetilde{\Gamma}_D$ for all n in \mathbb{N} . Now fix a sequence $v_{n_k} \xrightarrow{k \rightarrow \infty} v$ in $W_D^{1,p}(\Omega)$. Then Proposition B.3.5 yields a subsequence of (v_{n_k}) that converges pointwise quasi everywhere to v , which implies $v = 0$ q.e. on $\widetilde{\Gamma}_D$.

It remains to prove uniqueness of $\widetilde{\Gamma}_D$ and the inclusion in $\overline{\Gamma}_D$ up to polar sets. To that end, let $A \subset \partial\Omega$ with $\Gamma_D \subset A$ be quasi closed in $\partial\Omega$. Due to Lemma B.2.5, the set A is $\text{cap}_{\mathbb{R}^d}$ -quasi closed in \mathbb{R}^d as well and [111, Thm. 1.5] yields a non-negative function $v \in W^{1,p}(\mathbb{R}^d)$ such that $v^{-1}(\{0\})$ coincides with A up to $\text{cap}_{\mathbb{R}^d}$ -polar sets. The restriction $v|_{\overline{\Omega}}$ satisfies $v = 0$ q.e. on Γ_D and therefore on $\widetilde{\Gamma}_D$, as shown above, which proves that $\widetilde{\Gamma}_D \subset A$ up to polar sets and especially that $\widetilde{\Gamma}_D \subset \overline{\Gamma}_D$ up to polar sets.

When $A \subset \partial\Omega$ and satisfies $\Gamma_D \subset A$ and

$$W_D^{1,p}(\Omega) = \{ v \in W^{1,p}(\Omega) : v = 0 \text{ q.e. on } A \},$$

the same argument with reversed roles for A and the quasi closed set $\widetilde{\Gamma}_D$ yields that $A \subset \widetilde{\Gamma}_D$ and therefore $A = \widetilde{\Gamma}_D$ up to polar sets, which proves uniqueness of $\widetilde{\Gamma}_D$. \square

The first two equalities of the previous corollary are quite natural. The last equality can be viewed as an extension of the fact that continuous functions that map into Hausdorff spaces and are constant on a set are necessarily constant on the closure of the same set.

Remark 3.4.11 *Up to polar sets, the set $\widetilde{\Gamma}_D$ is the smallest quasi closed set that contains Γ_D . It coincides with the notion of the quasi closure of Γ_D as defined in [68, Sec. 2.8], which is omitted in the scope of this thesis. Note that the arbitrary intersection of all quasi closed sets that contain a set generally is not quasi closed and even yields the interior of a set, whenever point sets have capacity zero, see Example B.2.4.*

3.5 The Weak Non-Penetration Condition

Due to the trace results of the previous section, we can now give an alternative weak characterization of the admissible displacements in the contact problem that is based on the quasi everywhere sense implied by the capacities defined in (3.4), i.e., of the type employed in [138, P. 150,151], which the analysis of the contact problem and the corresponding optimal control problem in Chapters 4 to 6 relies on quite extensively. In light of their application in the contact problem, we fix $p = 2$ and consider the results in this section w.r.t. $H_D^1(\Omega)$. For notational simplicity, we write

$$\begin{aligned} \text{cap}_{\Omega,2}(\cdot) &:= \text{cap}(\cdot; \overline{\Omega}, H^1(\Omega), \mathcal{L}^d), \\ \text{cap}_{\partial\Omega,2}(\cdot) &:= \text{cap}(\cdot; \partial\Omega, H^{1/2}(\partial\Omega), \mathcal{H}^{d-1}), \end{aligned}$$

for the specialization of the capacities defined in (3.4).

The reformulation of the admissible displacements via the q.e. sense that we will consider is given by

$$\mathbf{K}_{\Phi}^{q.e.} := \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi \text{ q.e. in } \Gamma_C \}.$$

With the tools from the previous sections, the relation to the classical formulation and the properties of the alternative set can be discussed rather easily.

Proposition 3.5.1 *Let Assumption 3.2.1 (a)–(c) hold. Further, let $\Gamma_C \subset \partial\Omega$ be relatively open with $\text{dist}(\Gamma_C, \Gamma_D) > 0$, $\boldsymbol{\nu}_{\Phi}: \Gamma_C \rightarrow \mathbb{R}^d$ be continuous, and let $\Psi: \Gamma_C \rightarrow \overline{\mathbb{R}}$. Then*

$$\begin{aligned} \mathbf{K}_{\Phi} &= \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \text{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi \text{ } \mathcal{H}^{d-1}\text{-a.e. on } \Gamma_C \} \\ &\supset \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \text{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi \text{ q.e. on } \Gamma_C \} \\ &= \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi \text{ q.e. on } \Gamma_C \} = \mathbf{K}_{\Phi}^{q.e.}. \end{aligned}$$

When Ψ is additionally $\text{cap}_{\partial\Omega,2}$ -quasi continuous, all sets coincide.

Proof. The first and the last equality are simply a restatement of the definition of the set of admissible displacements \mathbf{K}_{Φ} in (2.12) and the reformulated set above. The second to last equality is a direct consequence of Corollary 3.4.5 since the restriction to Γ_C of all quasi continuous representatives of \mathbf{y} and $\text{tr}(\mathbf{y})$ are equal up to polar sets.

To confirm the inclusion, let $\mathbf{y} \in \mathbf{H}_D^1(\Omega)$ such that $\text{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_{\Phi} \leq \psi$ $\text{cap}_{\partial\Omega,2}$ -q.e. on Γ_C . By definition and Corollary 3.1.8, we know that for every quasi continuous representative of $\text{tr}(\mathbf{y})$ there exists a polar set $A \in \mathcal{B}(\partial\Omega)$ and $A \subset \Gamma_C$ such that the inequality is satisfied for all \mathbf{x} in $\Gamma_C \setminus A$. Due to Lemma 3.1.6 (a), we know that $\mathcal{H}^{d-1}(A) = 0$ is satisfied for every polar set $A \in \mathcal{B}(\partial\Omega)$, and therefore $\text{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_{\Phi} \leq \psi$ \mathcal{H}^{d-1} -a.e. on Γ_C , which proves the inclusion in the claim.

Now let Ψ additionally be $\text{cap}_{\partial\Omega,2}$ -quasi continuous. To prove the inverse inclusion, let $\mathbf{y} \in \mathbf{H}_D^1(\Omega)$ such that $\text{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi$ \mathcal{H}^{d-1} -a.e. on Γ_C , and consider the function

$$g: \Gamma_C \rightarrow \overline{\mathbb{R}}, \quad \text{with } g := \widetilde{\text{tr}(\mathbf{y})} \cdot \boldsymbol{\nu}_{\Phi} - \Psi,$$

where $\widetilde{\text{tr}(\mathbf{y})}$ is composed of $\text{cap}_{\partial\Omega,2}$ -quasi continuous representatives of the component functions y_i with i in $\{1, \dots, d\}$. Due to the regularity of Ψ and since $\boldsymbol{\nu}_{\Phi}$ is continuous, we know that g is $\text{cap}_{\partial\Omega,2}$ -quasi continuous on Γ_C as the product and difference of quasi continuous functions on Γ_C , cf. Lemma B.3.2. Clearly, we have that $g \leq 0$ \mathcal{H}^{d-1} -a.e. on Γ_C , and because Γ_C is relatively open in $\partial\Omega$, Lemma 3.1.15 implies that $g \leq 0$ $\text{cap}_{\partial\Omega,2}$ -q.e. on Γ_C as well. Therefore $\text{tr}(\mathbf{y}) \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi$ $\text{cap}_{\partial\Omega,2}$ -q.e. on Γ_C . \square

Remark 3.5.2 *In the proof of Proposition 3.5.1, quasi upper semi-continuity of Ψ , quasi continuity of the components of $\boldsymbol{\nu}_{\Phi}$ and quasi openness of Γ_C in $\partial\Omega$ are in fact sufficient to obtain equality of the sets in the claim.*

Note that Corollaries 3.4.2 and 3.4.7 imply that the quasi everywhere sense on $\partial\Omega \setminus \overline{\Gamma_D}$ can be chosen somewhat arbitrarily, but in the proof of the previous proposition, it is crucial that the capacity $\text{cap}_{\partial\Omega,2}$ generated by trace functions is used in order to be able to establish the connection to the Hausdorff measure and in order to apply Lemma 3.1.15.

Accordingly, it was imperative to include the trace capacity $\text{cap}_{\partial\Omega}$ in the considerations of this chapter.

The set of admissible displacements \mathbf{K}_{Φ} defined in (2.12) is always convex and closed in $\mathbf{H}_D^1(\Omega)$ and $\mathbf{H}^1(\Omega)$, which is implied by the existence of pointwise \mathcal{H}^{d-1} -a.e. convergent subsequences of convergent sequences in the trace spaces. When the regularity requirements on Ψ in Proposition 3.5.1 are violated, then $\mathbf{K}_{\Phi}^{q.e.}$ may be a strict subset of \mathbf{K}_{Φ} . However, the set of admissible displacements is always convex and closed, which is why the subsequent analysis of the contact problem with the constraint formulated in the quasi everywhere sense can be carried out even if the regularity assumption is violated.

Proposition 3.5.3 (Convexity and Closedness) *Let Assumption 3.2.1 (a) - (c) hold, and let $\Gamma_C \subset \partial\Omega$ as well as $\nu_{\Phi}: \Gamma_C \rightarrow \mathbb{R}^d$ and $\Psi: \Gamma_C \rightarrow \overline{\mathbb{R}}$. Then $\mathbf{K}_{\Phi}^{q.e.}$ is convex and closed in $\mathbf{H}_D^1(\Omega)$ and $\mathbf{H}^1(\Omega)$.*

Proof. Convexity clearly holds because of the linearity of the euclidean scalar product and since the inequality is enforced pointwise quasi everywhere. For closedness, consider a sequence $(\mathbf{y}_n) \subset \{\mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \nu_{\Phi} \leq \Psi \text{ q.e. in } \Gamma_C\}$ such that $\mathbf{y}_n \rightarrow \mathbf{y}$ in $\mathbf{H}_D^1(\Omega)$. Due to Proposition B.3.5, there exists a subsequence of (\mathbf{y}_n) converging pointwise quasi everywhere, therefore $\mathbf{y} \cdot \nu_{\Phi} \leq \Psi$ is satisfied q.e. in Γ_C . Continuity and linearity of the trace operator imply that $\mathbf{H}_D^1(\Omega)$ is convex and closed in $\mathbf{H}^1(\Omega)$. \square

For the remainder of the thesis, we will assume the initial gap Ψ to be quasi continuous. Accordingly, the measure theoretical definition coincides with the capacity theoretical reformulation, and we can fix the definition of the set of admissible displacements.

Definition 3.5.4 (Set of Admissible Displacements) *Let Assumption 3.2.1 (a)-(c) hold. Further, let $\Gamma_C \subset \partial\Omega$ be relatively open with $\text{dist}(\Gamma_C, \Gamma_D) > 0$, $\nu_{\Phi}: \Gamma_C \rightarrow \mathbb{R}^d$ be continuous, and let $\Psi: \Gamma_C \rightarrow \overline{\mathbb{R}}$ be quasi continuous. The set*

$$\begin{aligned} \mathbf{K}_{\Phi} &:= \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \nu_{\Phi} \leq \Psi \text{ q.e. in } \Gamma_C \} \\ &= \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \text{tr}(\mathbf{y}) \cdot \nu_{\Phi} \leq \Psi \text{ } \mathcal{H}^{d-1}\text{-a.e. in } \Gamma_C \} \end{aligned}$$

is called the set of admissible displacements.

From this point on, when no specific capacity is referenced, the capacity $\text{cap}_{\mathbb{R}^d}$ for $p = 2$ is employed in the examination of the fine properties of Sobolev functions. We will use the equivalent capacity $\text{cap}_{\Omega,2}$ on occasion.

Chapter 4

Time Discretization and Analysis of the Contact Problem

With the capacity-based reformulation of the set of admissible displacements in Definition 3.5.4, the dynamic obstacle problem amounts to the problem of finding a function \mathbf{y} in $\mathbf{K}_{\Phi}^T \cap \mathbf{Y}$ that solves the variational problem

$$\begin{aligned} \mathbf{y}(0) = \mathbf{y}_{\text{ini}}, \quad \dot{\mathbf{y}}(0) = \mathbf{v}_{\text{ini}} \\ \langle \rho \ddot{\mathbf{y}} - f, \mathbf{v} - \mathbf{y} \rangle_{L^2(0,T; \mathbf{H}^1(\Omega))} + a_I(\mathbf{y}, \mathbf{v} - \mathbf{y}) + b_I(\dot{\mathbf{y}}, \mathbf{v} - \mathbf{y}) \geq 0 \end{aligned} \quad (4.1)$$

for all $\mathbf{v} \in \mathbf{K}_{\Phi}^T$, with the sets

$$\begin{aligned} \mathbf{Y} &:= \{ \mathbf{y} \in L^2(0, T; \mathbf{H}^1(\Omega)) : \dot{\mathbf{y}} \in \mathbf{W}(0, T) \}, \\ \mathbf{K}_{\Phi}^T &:= \{ \mathbf{y} \in L^2(0, T; \mathbf{H}^1(\Omega)) : \mathbf{y}(t) \in \mathbf{K}_{\Phi} \text{ f.a.a. } t \in [0, T] \}, \\ \mathbf{K}_{\Phi} &:= \{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \boldsymbol{\nu}_{\Phi} \leq \Psi \text{ q.e. on } \Gamma_C \}. \end{aligned} \quad (4.2)$$

Results concerning the existence, uniqueness and regularity of solutions for this fully dynamic and frictionless problem in the current literature are incomplete at best, cf. the overview in Chapter 1. While the optimal control theory for the former may therefore lack a solid foundation to build on, several problem specific temporal discretization schemes for the problem at hand are well established and provide an adequate basis for considerations in optimization.

A popular basis for discretization schemes suited for the treatment of contact constraints is the well-known family of *Newmark schemes*. For appropriate parameter choices, Newmark schemes are known to be second-order consistent, unconditionally stable and also momentum and energy preserving when applied to the unconstrained elastic problem. See, e.g., [99, 124] and [112, Sec. 2.4] for conservation properties in viscoelasticity. When applied in the constrained setting, the original scheme may lead to energy blow-ups as well as oscillations of the solution at the contact boundary, see [52, 118, 124]. An energy dissipative modification that deals with the contact condition implicitly was introduced by Kane et al. in [106], and improvements that deal with the oscillations while preserving dissipativity

are discussed in [52, 112, 113, 114, 118]. See [112, Sec. 2.1-2.4] and the references within for a more detailed overview of these methods.

In this chapter, we will analyze a time-discretized version of the continuous problem (4.1) with respect to its solutions and sensitivities. We will develop a nonconforming, discontinuous temporal finite element discretization of the contact problem that leads to a slightly modified version of the *contact implicit* Newmark scheme in [106] and allows for the consistent derivation of an adjoint time stepping scheme. For the discretized problem, we then establish existence of a Lipschitz continuous solution operator and establish its differentiability properties based on the active sets. In addition to the assumptions on the physical setting in Assumption 2.3.1, we can fix standing regularity requirements for the remainder of the thesis, which are assumed to hold unless otherwise stated.

Assumption 4.0.1 (Standing Assumptions)

- (a) *Assumption 2.3.1 on the physical setting holds.*
- (b) *The reference configuration $\Omega \subset \mathbb{R}^d$ is a bounded strong Lipschitz domain.*
- (c) *The boundary sections $\Gamma_D, \Gamma_N, \Gamma_C \subset \partial\Omega$ are pairwise disjoint, relatively open in $\partial\Omega$, and $\overline{\Gamma_D} \cup \overline{\Gamma_N} \cup \overline{\Gamma_C} = \partial\Omega$.*
- (d) *The boundary sections Γ_D and Γ_C have positive distance.*
- (e) *The contact normal $\boldsymbol{\nu}_\Phi$ is in $\mathbf{C}^{0,1}(\overline{\Gamma_C})$, and $\|\boldsymbol{\nu}_\Phi(\mathbf{x})\| = 1$ for all \mathbf{x} in Γ_C .*
- (f) *The gap function Ψ is quasi continuous w.r.t. $\text{cap}(\cdot; \partial\Omega, H^{1/2}(\partial\Omega), \mathcal{H}^{d-1})$ in the sense of Definition 3.1.10 and non-negative.*
- (g) *The bilinear forms a, b, a_I and b_I are given as in (2.11), symmetric and bounded.*
- (h) *The external forces are given by $f \in L^2(0, T; \mathbf{H}^1(\Omega))^*$.*

Recall that the mass density ρ was assumed to be constant in space. For the sake of a clearer presentation, the mass density ρ is assumed to equal 1 in the analysis until the mass regains relevance in the numerics chapter, where we will explicitly include its symbol again. Further, note that the assumptions for the capacity-based reformulation in Proposition 3.5.1 and Definition 3.5.4 are included in Assumption 4.0.1.

Structure. In Section 4.1, we reformulate the hyperbolic variational inequality (4.1) as a first order system. We apply the temporal finite element discretization to obtain a time-discretized contact problem that corresponds to a slightly modified contact implicit Newmark scheme that maintains energy dissipativity. The time-discretized setting is collected and reduced to a formulation in the displacements in Section 4.2. Using the time stepping structure of the discretization, we can then establish the existence of the Lipschitz continuous solution operator to the time-discretized problem in Section 4.3. Section 4.4 deals with the differentiability properties of the solution operator. As usual, the operator can not be expected to be Fréchet differentiable everywhere, but we establish directional differentiability using the polyhedricity-based approaches in [138]. A localized characterization of the contact forces allows us to interpret the contact conditions of the linearized problem as boundary conditions, and the pointwise characterization is used to investigate the points of Gâteaux and Fréchet differentiability of the solution operator.

For the most part, the results in this chapter have been published in [143], but we will see refined results concerning polyhedricity of the set of admissible displacements in Sec-

tion 4.4.1 that make due with simplified assumptions, and the analysis of the sensitivities and the linearized boundary conditions in Sections 4.4.2 and 4.4.3 is significantly extended.

4.1 Time Discretization of the Continuous Problem

The fundamental difficulty in the (finite-element-based) discretization of the contact problem (4.1) is the treatment of the constraint. Due to the constraint, the external and internal forces in the second line of the problem are generally imbalanced, which forces us to deal with the inequality structure. We will employ an extension of the temporal part of the Petrov-Galerkin discretization presented in [120], where the authors present a finite element discretization for optimal control of the wave equation that results in a Crank-Nicolson scheme. We start out by introducing the multiplier $\lambda \in \mathcal{T}_{\mathbf{K}_{\Phi}^T}(\mathbf{y})^\circ \subset L^2(0, T; \mathbf{H}^1(\Omega))^*$ into the problem, which yields an equilibrium condition, i.e., the problem of finding $\mathbf{y} \in \mathbf{Y}$ and $\lambda \in L^2(0, T; \mathbf{H}^1(\Omega))^*$ that solve

$$\begin{aligned} \mathbf{y} \in \mathbf{K}_{\Phi}^T, \quad \lambda \in \mathcal{T}_{\mathbf{K}_{\Phi}^T}(\mathbf{y})^\circ, \quad \mathbf{y}(0) = \mathbf{y}_{\text{ini}}, \quad \dot{\mathbf{y}}(0) = \mathbf{v}_{\text{ini}} \\ \ddot{\mathbf{y}} + a_I(\mathbf{y}, \cdot) + b_I(\dot{\mathbf{y}}, \cdot) - f + \lambda = 0. \end{aligned}$$

As in [120], we capitalize on the well-known fact that Newmark-type schemes for the second order problems are closely linked to the Crank-Nicolson scheme for their respective reformulations as first order systems and explicitly introduce the velocities $\mathbf{v} \in \mathbf{W}(0, T)$ to obtain the following first order reformulation of the contact problem:

Problem 4.1.1 (Continuous Contact Problem) *Given forces $f \in L^2(0, T; \mathbf{H}^1(\Omega))^*$ and initial values $\mathbf{y}_{\text{ini}} \in \mathbf{K}_{\Phi} \subset \mathbf{H}_D^1(\Omega)$ and $\mathbf{v}_{\text{ini}} \in \mathbf{H}^1(\Omega)$, find a state $\mathbf{y} \in \mathbf{Y}$, its velocity $\mathbf{v} \in \mathbf{W}(0, T)$ and a multiplier $\lambda \in L^2(0, T; \mathbf{H}^1(\Omega))^*$ that solve*

$$\mathbf{y} \in \mathbf{K}_{\Phi}^T \quad (4.3a)$$

$$\lambda \in \mathcal{T}_{\mathbf{K}_{\Phi}^T}(\mathbf{y})^\circ \quad (4.3b)$$

$$\langle \dot{\mathbf{v}}, \mathbf{p} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} + a_I(\mathbf{y}, \mathbf{p}) + b_I(\mathbf{v}, \mathbf{p}) + \langle \lambda - f, \mathbf{p} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} = 0 \quad (4.3c)$$

$$\langle \dot{\mathbf{y}} - \mathbf{v}, \mathbf{q} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} = 0 \quad (4.3d)$$

$$\langle \mathbf{y}(0) - \mathbf{y}_{\text{ini}}, \mathbf{q}_0 \rangle_{L^2(\Omega)} = 0 \quad (4.3e)$$

$$\langle \mathbf{v}(0) - \mathbf{v}_{\text{ini}}, \mathbf{p}_0 \rangle_{L^2(\Omega)} = 0 \quad (4.3f)$$

for all $\mathbf{p}, \mathbf{q} \in L^2(0, T; \mathbf{H}^1(\Omega))$ and $\mathbf{p}_0, \mathbf{q}_0 \in L^2(\Omega)$.

The initial value conditions (4.3e)–(4.3f) are to be understood in the sense of the embeddings $\mathbf{W}(0, T) \hookrightarrow C([0, T], L^2(\Omega))$ and $\mathbf{H}^1(\Omega) \hookrightarrow L^2(\Omega)$. As a foundation for the finite element discretization, assume a subdivision of the time interval $I = [0, T]$ into $N \geq 1$ subintervals I_k with

$$0 = t_0 < t_1 < \dots < t_N = T, \quad I_k = (t_{k-1}, t_k], \quad k = 1, \dots, N$$

to be fixed. For ease of presentation, we will only consider the equidistant case with $\tau := |I_k| > 0$. Non-equidistant discretizations can be handled in a straight forward manner.

Remark 4.1.2 *The multiplier of course satisfies $\lambda = -f_{con}$ for the contact forces defined in (2.14) and is introduced separately mainly for the sign convention.*

Discretization of Ansatz and Test Functions. Following [120], we discretize displacements and velocities using piecewise linear and continuous ansatz functions while the test functions are chosen piecewise constant and discontinuous, i.e.,

$$\begin{aligned} \mathbf{y}, \mathbf{v} &\in \{ \mathbf{y} \in C([0, T], \mathbf{H}^1(\Omega)) : \mathbf{y}|_{I_k} \in \mathcal{P}_1(I_k, \mathbf{H}^1(\Omega)) \}, \\ \mathbf{p}, \mathbf{q} &\in \{ \mathbf{p} \in L^2(0, T; \mathbf{H}^1(\Omega)) : \mathbf{p}|_{I_k} \in \mathcal{P}_0(I_k, \mathbf{H}^1(\Omega)), \mathbf{p}(0) = \mathbf{p}(t_1) \}, \end{aligned} \quad (4.4)$$

where $\mathcal{P}_0(I_k, X)$ and $\mathcal{P}_1(I_k, X)$ denote constant and affine linear functions on I_k with values in X , respectively. The discretization is nonconforming in the sense that the displacements are discretized piecewise affine linearly, which implies piecewise constant discontinuous weak derivatives, but the velocities themselves are discretized piecewise affine linearly as well. This leads to a symmetric averaging of implicit and explicit information when the states are updated from the velocities in the time stepping scheme.

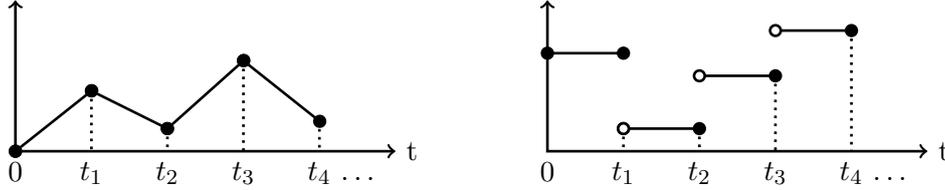


Figure 4.1: Discretization of the primal variables. Time-continuous ansatz functions (left) and time-discontinuous test functions (right).

Discretization of Dual Quantities. Having discretized the primal variables, the forces in the discretized setting are usually obtained by applying the continuous dual quantities to the discretized primal variables. This way, for all testfunctions $\mathbf{p} = \sum_{k=1}^N \mathbf{p}_k \chi_{I_k}$, we obtain that

$$\langle f, \mathbf{p} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} = \sum_{k=1}^N \tau \langle f_k, \mathbf{p}_k \rangle_{\mathbf{H}^1(\Omega)},$$

where the external forces in the discretized setting are computed from the impulses by evaluating

$$\langle f_k, \mathbf{p}_k \rangle_{\mathbf{H}^1(\Omega)} := \frac{1}{\tau} \int_{t_{k-1}}^{t_k} \langle f(t), \mathbf{p}_k \rangle_{\mathbf{H}^1(\Omega)} dt.$$

Note that for the piecewise constant test functions, this is equivalent to discretizing the forces $f \in L^2(0, T; \mathbf{H}^1(\Omega)^*)$ explicitly and piecewisely constant by $f = \sum_{k=1}^N f_k \chi_{I_k}$, which corresponds to distributing the impulse $\int_{t_{k-1}}^{t_k} \langle f(t), \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)} dt$ evenly over the interval I_k . This will become relevant in the discretization of the controls in Chapter 5 again, and we will adhere to this convention whenever referring to forces that correspond to discrete impulses, particularly after discretizing the multiplier λ .

The discrete representation of λ is the key component in the discretization. Wanting to end up with a dissipative time stepping scheme, we require a discrete counterpart of λ that appears implicitly in the time steps of the action equality but satisfies a reasonable discretized version of the constraints. For motivation, observe that $\{ \mathbf{v} \in C([0, T], \mathbf{H}^1(\Omega)) : \mathbf{v}(t) \in \mathbf{K}_\Phi \text{ for all } t \in [0, T] \}$ is dense in $\{ \mathbf{v} \in L^2(0, T; \mathbf{H}^1(\Omega)) : \mathbf{v}(t) \in \mathbf{K}_\Phi \text{ f.a.a. } t \in [0, T] \}$, as one can check using the fact that $C([0, T], \mathbf{H}^1(\Omega))$ is dense in $L^2(0, T; \mathbf{H}^1(\Omega))$ [207, Prop. 23.2 c)] and the norm projection in $\mathbf{H}^1(\Omega)$ onto the set of admissible displacements. In fact, due to the embedding

$$\{ \mathbf{y} \in L^2(0, T; \mathbf{H}^1(\Omega)) : \dot{\mathbf{y}} \in \mathbf{W}(0, T) \} \hookrightarrow C([0, T], \mathbf{H}^1(\Omega)),$$

e.g., [192, Thm. 3.1.41], it suffices to consider continuous solutions as well, hence the test functions and the solution in Problem 4.1.1 can be replaced by continuous functions. For bounded, linear functionals on continuous functions, Singer's Theorem (see Theorem B.3) establishes that

$$C([0, T], \mathbf{H}^1(\Omega))^* \cong \mathcal{M}([0, T], \mathbf{H}^1(\Omega)^*),$$

where the latter denotes the $\mathbf{H}^1(\Omega)^*$ -valued regular Borel measures of bounded variation on $[0, T]$, and evaluation of a linear functional at a continuous function corresponds to the integration of the function with respect to the measure. Accordingly, restricting its domain to continuous functions, the multiplier $\lambda \in C([0, T], \mathbf{H}^1(\Omega))^*$ admits such a measure representation, which motivates the discretization

$$\lambda \in \left\{ \lambda \in \mathcal{M}([0, T], \mathbf{H}^1(\Omega)^*) : \lambda \in \text{lin}_{\mathbf{H}^1(\Omega)^*}(\delta_{t_k}, k = 1, \dots, N) \right\},$$

where $\text{lin}_{\mathbf{H}^1(\Omega)^*}(\delta_{t_k}, k = 1, \dots, N)$ is the span of all $\mathbf{H}^1(\Omega)^*$ -valued Dirac measures that are concentrated on the temporal discretization points t_k , $k = 1, \dots, N$. The discrete degrees of freedom of λ are therefore

$$\lambda_k := \lambda(\{t_k\}), \quad k = 1, \dots, N.$$

As one would expect from Dirac measures, the discrete quantities λ_k are impulses, and according to the convention above, the time dependent contact forces that correspond to these impulses are given by $-\sum_{k=1}^N \frac{1}{\tau} \lambda_k \chi_{I_k}$.

Of course, the discretization of the contact multiplier is nonconforming — a fact that can only be remedied by allowing $C([0, T], \mathbf{H}^1(\Omega))^*$ multipliers in the problem formulation of Problem 4.1.1 to begin with, which changes the sense of weak solutions. Note that the test functions are discontinuous but piecewisely defined, therefore using the identification with the measures, we adhere to the convention that with a slight abuse of notation, “ $\langle \lambda, p \rangle_{C([0, T], \mathbf{H}^1(\Omega))}$ ” means the integration of p with respect to λ , which is well defined. We will shortly discuss the necessity of this part of the discretization after addressing the components of the action equality and the constraint.

Discretized Operators. Evaluating the respective (bi-)linear forms in (4.3c) leads to:

$$\int_0^T a(\mathbf{y}, \mathbf{p}) \, d\mathcal{L} = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} a(\mathbf{y}, \mathbf{p}) \, d\mathcal{L} = \frac{\tau}{2} \sum_{k=0}^{N-1} a(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) + a(\mathbf{y}_k, \mathbf{p}_{k+1}) \quad (4.5a)$$

$$\int_0^T b(\mathbf{v}, \mathbf{p}) \, d\mathcal{L} = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} b(\mathbf{v}, \mathbf{p}) \, d\mathcal{L} = \frac{\tau}{2} \sum_{k=0}^{N-1} b(\mathbf{v}_{k+1}, \mathbf{p}_{k+1}) + b(\mathbf{v}_k, \mathbf{p}_{k+1}) \quad (4.5b)$$

$$\langle f, \mathbf{p} \rangle_{L^2(0,T; \mathbf{H}^1(\Omega))} = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \langle f, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)} \, d\mathcal{L} = \sum_{k=0}^{N-1} \tau \langle f_{k+1}, \mathbf{p}_{k+1} \rangle_{\mathbf{H}^1(\Omega)} \quad (4.5c)$$

$$\langle \dot{\mathbf{v}}, \mathbf{p} \rangle_{L^2(0,T; \mathbf{H}^1(\Omega))} = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \langle \dot{\mathbf{v}}, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)} \, d\mathcal{L} = \sum_{k=0}^{N-1} (\mathbf{v}_{k+1}, \mathbf{p}_{k+1})_{L^2(\Omega)} - (\mathbf{v}_k, \mathbf{p}_{k+1})_{L^2(\Omega)} \quad (4.5d)$$

$$\langle \lambda, \mathbf{p} \rangle_{C([0,T], \mathbf{H}^1(\Omega))} \hat{=} \int_0^T \mathbf{p} \, d\lambda = \sum_{k=0}^{N-1} \langle \lambda_{k+1}, \mathbf{p}_{k+1} \rangle_{\mathbf{H}^1(\Omega)} \quad (4.5e)$$

The same argument yields the discretized form of the velocity update (4.3d):

$$\begin{aligned} \langle \dot{\mathbf{y}} - \mathbf{v}, \mathbf{q} \rangle_{L^2(0,T; \mathbf{H}^1(\Omega))} &= \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \langle \dot{\mathbf{y}} - \mathbf{v}, \mathbf{q} \rangle_{\mathbf{H}^1(\Omega)} \, d\mathcal{L} \\ &= \sum_{k=0}^{N-1} (\mathbf{y}_{k+1} - \mathbf{y}_k, \mathbf{q}_{k+1})_{L^2(\Omega)} - \frac{\tau}{2} (\mathbf{v}_{k+1} + \mathbf{v}_k, \mathbf{q}_{k+1})_{L^2(\Omega)}. \end{aligned} \quad (4.6)$$

Equations (4.5a)–(4.5b) follow because the integrand is affine linear on each subinterval, while equations (4.5c)–(4.5d) follow because of the constant integrands. Note that in accordance with the Gelfand triple $\mathbf{H}^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow \mathbf{H}^1(\Omega)^*$, all $\mathbf{H}^1(\Omega)$ -functions are identified with $\mathbf{H}^1(\Omega)^*$ -functionals by use of the $L^2(\Omega)$ -Riesz isomorphism, instead of the $\mathbf{H}^1(\Omega)$ -isomorphism, which leads to the $L^2(\Omega)$ -scalar products in (4.5d) and (4.6). Since \mathbf{p} is discretized discontinuous in time, the evaluation of the linear functional λ in (4.5e) is to be understood in the sense of Singer's theorem, as explained above.

Discretized Constraints. The set \mathbf{K}_{Φ}^T is convex, therefore any \mathbf{y} in $C([0, T], \mathbf{H}^1(\Omega))$ that is piecewise affine linear satisfies $\mathbf{y} \in \mathbf{K}_{\Phi}^T$ if and only if $\mathbf{y}_k \in \mathbf{K}_{\Phi}$ for all $k = 0, \dots, N$. This leaves the constraint on the multiplier λ to be discussed, which enforces the condition $\lambda \in \mathcal{T}_{\mathbf{K}_{\Phi}^T}(\mathbf{y})^\circ$ for the piecewise affine linear solution \mathbf{y} , see (4.3b). The choice of discretization yields that

$$\langle \lambda, \varphi - \mathbf{y} \rangle_{C([0,T], \mathbf{H}^1(\Omega))} = \sum_{k=1}^N \langle \lambda_k, \varphi_k - \mathbf{y}_k \rangle_{\mathbf{H}^1(\Omega)} \leq 0 \quad (4.7)$$

for all pointwisely defined $\varphi \in \mathbf{K}_{\Phi}^T$. The variation in φ includes the choice of continuous, piecewise linear φ with $\varphi_i = \mathbf{y}_i \in \mathbf{K}_{\Phi}$, $i = 1, \dots, N$, $i \neq k$, therefore (4.7) decouples and yields the componentwise condition that

$$\langle \lambda_k, \varphi - \mathbf{y}_k \rangle_{\mathbf{H}^1(\Omega)} \leq 0 \quad \forall \varphi \in \mathbf{K}_{\Phi}, k = 1, \dots, N,$$

which can equivalently be rewritten as:

$$\lambda_k \in \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y}_k)^\circ, \quad k = 1, \dots, N.$$

The tangent cone $\mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})$ coincides with the product of the tangent cones. More specifically, we obtain the following intuitive correspondence.

Lemma 4.1.3 *Let $N \geq 1$ and $K \subset Y$ be a closed and convex subset of a Banach space Y , and let $y \in K^N$, $\lambda \in \mathcal{T}_{K^N}(y)$. Then:*

- (a) $\mathcal{T}_{K^N}(y) = \prod_{k=1}^N \mathcal{T}_K(y_k)$
- (b) $\mathcal{T}_{K^N}(y)^\circ = \prod_{k=1}^N \mathcal{T}_K(y_k)^\circ$
- (c) $\mathcal{C}_{K^N}(y, \lambda) = \prod_{k=1}^N \mathcal{C}_K(y_k, \lambda_k)$
- (d) $\mathcal{C}_{K^N}(y, \lambda)^\circ = \prod_{k=1}^N \mathcal{C}_K(y_k, \lambda_k)^\circ$

Proof. See Lemma C.2.2. □

Accordingly, we can write the discretized problem in its compact variational form, which amounts to the following:

Problem 4.1.4 (Discretized Contact Problem) *Given $f \in (\mathbf{H}^1(\Omega)^N)^*$, $\mathbf{y}_{ini} \in \mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega)$ and $\mathbf{v}_{ini} \in \mathbf{H}^1(\Omega)$, find $\mathbf{y}_0, \mathbf{v}_0 \in \mathbf{H}^1(\Omega)$, $\mathbf{y}, \mathbf{v} \in \mathbf{H}^1(\Omega)^N$ and $\lambda \in (\mathbf{H}^1(\Omega)^*)^N$ that solve*

$$\begin{aligned} & \mathbf{y} \in \mathbf{K}_\Phi^N \\ & \lambda \in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ \\ & (\mathbf{y}_0 - \mathbf{y}_{ini}, \mathbf{q}_0)_{L^2(\Omega)} + (\mathbf{v}_0 - \mathbf{v}_{ini}, \mathbf{p}_0)_{L^2(\Omega)} \\ & + \sum_{k=0}^{N-1} \left[(\mathbf{v}_{k+1}, \mathbf{p}_{k+1})_{L^2(\Omega)} - (\mathbf{v}_k, \mathbf{p}_{k+1})_{L^2(\Omega)} \right. \\ & + \frac{\tau}{2} (a(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) + a(\mathbf{y}_k, \mathbf{p}_{k+1})) \\ & + \frac{\tau}{2} (b(\mathbf{v}_{k+1}, \mathbf{p}_{k+1}) + b(\mathbf{v}_k, \mathbf{p}_{k+1})) \\ & - \tau \langle f_{k+1}, \mathbf{p}_{k+1} \rangle_{\mathbf{H}^1(\Omega)} + \langle \lambda_{k+1}, \mathbf{p}_{k+1} \rangle_{\mathbf{H}^1(\Omega)} \\ & \left. + (\mathbf{y}_{k+1} - \mathbf{y}_k, \mathbf{q}_{k+1})_{L^2(\Omega)} - \frac{\tau}{2} (\mathbf{v}_{k+1} + \mathbf{v}_k, \mathbf{q}_{k+1})_{L^2(\Omega)} \right] = 0 \end{aligned}$$

for all $\mathbf{p}, \mathbf{q} \in \mathbf{H}^1(\Omega)_{L^2(\Omega)}$.

Note that the discontinuity of the test functions in the discretization of (4.3) leads to a set of equations that is decoupled with respect to the test functions' degrees of freedom and sequentially coupled in the ansatz functions. It yields the following modified Crank-Nicolson time stepping scheme in the values $(\mathbf{y}_k, \mathbf{v}_k) \in \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega)$, $k = 0, \dots, N$.

$$\begin{aligned}
(\mathbf{y}_0, \mathbf{q}_0)_{\mathbf{L}^2(\Omega)} &= (\mathbf{y}_{\text{ini}}, \mathbf{q}_0)_{\mathbf{L}^2(\Omega)} & \forall \mathbf{q}_0 \in \mathbf{H}^1(\Omega) \\
(\mathbf{v}_0, \mathbf{p}_0)_{\mathbf{L}^2(\Omega)} &= (\mathbf{v}_{\text{ini}}, \mathbf{p}_0)_{\mathbf{L}^2(\Omega)} & \forall \mathbf{p}_0 \in \mathbf{H}^1(\Omega) \\
(\mathbf{y}_{k+1}, \mathbf{p})_{\mathbf{L}^2(\Omega)} &= (\mathbf{y}_k, \mathbf{p})_{\mathbf{L}^2(\Omega)} + \frac{\tau}{2} (\mathbf{v}_{k+1} + \mathbf{v}_k, \mathbf{p})_{\mathbf{L}^2(\Omega)} & \forall \mathbf{p} \in \mathbf{H}^1(\Omega) \\
(\mathbf{v}_{k+1}, \mathbf{q})_{\mathbf{L}^2(\Omega)} &= (\mathbf{v}_k, \mathbf{q})_{\mathbf{L}^2(\Omega)} - \frac{\tau}{2} (a(\mathbf{y}_{k+1}, \mathbf{q}) + a(\mathbf{y}_k, \mathbf{q})) \\
&\quad - \frac{\tau}{2} (b(\mathbf{v}_{k+1}, \mathbf{q}) + b(\mathbf{v}_k, \mathbf{q})) \\
&\quad + \tau \langle f_{k+1}, \mathbf{q} \rangle_{\mathbf{H}^1(\Omega)} - \langle \lambda_{k+1}, \mathbf{q} \rangle_{\mathbf{H}^1(\Omega)} & \forall \mathbf{q} \in \mathbf{H}^1(\Omega) \\
\mathbf{y}_{k+1} &\in \mathbf{K}_{\Phi}, \quad \lambda_{k+1} \in \mathcal{T}_{\mathbf{K}_{\Phi}}(\mathbf{y}_{k+1})^\circ.
\end{aligned}$$

This scheme is a slightly modified version of the contact implicit scheme first presented in [106]. It treats the external forces implicitly as well and is energy dissipative in the absence of external forces. The proof of energy dissipativity can be obtained by minor modifications of [52, Thm. 2.1] and its extension in [112, Thm. 2.4.2].

Discussion of the Modified Discretization. Due to the poor stability properties of the classic symmetric Newmark scheme in the contact constrained case, our aim was to obtain an energy dissipative, contact implicit time stepping scheme. In order to obtain a set of time independent variational inequalities that can be solved sequentially, the discretization scheme was required to decouple the inequality constraint on the multiplier. By nature of the variational inequality, the multiplier condition in the continuous formulation is tested with a difference of two ansatz functions from the admissible set, meaning

$$\lambda \in \mathcal{T}_{\mathbf{K}_{\Phi}^T}(\mathbf{y})^\circ \Leftrightarrow \langle \lambda, \varphi - \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)} \leq 0 \quad \forall \varphi \in \mathbf{K}_{\Phi}^T.$$

As an ansatz function, the displacement \mathbf{y} in $\mathbf{K}_{\Phi}^T \cap \mathbf{Y}$ is discretized piecewise linearly and continuously, which requires the Dirac-type discretization introduced above, since otherwise the coupling of \mathbf{y}_k and \mathbf{y}_{k+1} introduces a dependence of the constraint on the multiplier λ at the time t_{k+1} on both y_k and y_{k+1} . Allowing the contact forces to only act locally at the times of discretization to respect the contact constraints at those specific times is justified due to the convexity of the set of piecewise linear admissible displacements.

4.2 The Time-Discretized Contact Problem

The results of the discretization that was introduced above will be collected in this section. We introduce the required notation and state the time-discretized problem. Problem 4.1.4, which is formulated in tuples of states and velocities $(\mathbf{y}_{\text{ini}}, \mathbf{v}_{\text{ini}}), (\mathbf{y}_k, \mathbf{v}_k), k = 0, \dots, N$, will be reduced to a formulation in displacements only. Subsequently, we include initial values explicitly in the problem, replacing the additional variational equations that enforce the initial conditions. In the distributed control setting of the following chapter, these structural modifications yield density of the control space's image under the appropriate operator in the dual of the test space, which allows us to apply the stationarity concepts in [193] to obtain optimality conditions of first order. Additionally, it being the key aspect of the formulation, we will elaborate on the time stepping structure of Problem 4.1.4.

Eliminating the velocities from the unknowns can be done easily via a direct computation based on the coupling (4.6).

Lemma 4.2.1 *Let $\mathbf{y}_0, \mathbf{v}_0 \in \mathbf{H}^1(\Omega)$ and $\mathbf{v}, \mathbf{y} \in \mathbf{H}^1(\Omega)^N$. Then*

$$\sum_{k=0}^{N-1} (\mathbf{y}_{k+1} - \mathbf{y}_k, \mathbf{q}_{k+1})_{L^2(\Omega)} - \frac{\tau}{2} (\mathbf{v}_{k+1} + \mathbf{v}_k, \mathbf{q}_{k+1})_{L^2(\Omega)} = 0$$

for all $\mathbf{q} \in \mathbf{H}^1(\Omega)^N$ if and only if

$$\mathbf{v}_k = (-1)^k \mathbf{v}_0 + \frac{2}{\tau} \sum_{j=1}^k (-1)^{k+j} (\mathbf{y}_j - \mathbf{y}_{j-1}) \text{ for } k = 0, \dots, N.$$

Proof. The claim can be confirmed by a straight forward induction argument using the form of the coupling. \square

As expected, the velocities \mathbf{v}_k are dependent only on the initial values (which equal $\mathbf{y}_0, \mathbf{v}_0$) and the displacements $\mathbf{y}_1, \dots, \mathbf{y}_k$ at the time steps up to the current step. Note that when $\mathbf{y}_0, \mathbf{v}_0$ are replaced with the initial values $\mathbf{y}_{\text{ini}}, \mathbf{v}_{\text{ini}}$, the velocities are only affine linear in $\mathbf{y}_1, \dots, \mathbf{y}_k$. In order to separate the given data from the unknowns in the reduced formulation and obtain a linear operator representing the time stepping scheme, we will employ the following decomposition into linear and constant parts.

Definition 4.2.2 (Notation) *Let $f \in (\mathbf{H}^1(\Omega)^N)^*$, and let $\mathbf{y}_{\text{ini}} \in \mathbf{K}_{\Phi} \subset \mathbf{H}_D^1(\Omega)$ and $\mathbf{v}_{\text{ini}} \in \mathbf{H}^1(\Omega)$.*

- (a) *The parts of the velocities that are independent of the initial data are represented by the bounded linear operators $\bar{v}_k: \mathbf{H}^1(\Omega)^k \rightarrow \mathbf{H}^1(\Omega)$ with*

$$\bar{v}_k(\mathbf{y}) := \frac{2}{\tau} ((-1)^{k+1} \mathbf{y}_1 + \sum_{j=2}^k (-1)^{k+j} (\mathbf{y}_j - \mathbf{y}_{j-1})),$$

for $k = 1, \dots, N$.

- (b) *The internal forces and the time stepping structure are represented by the bounded linear operator*

$$\begin{aligned} A_{\tau}: \mathbf{H}^1(\Omega)^N &\rightarrow (\mathbf{H}^1(\Omega)^N)^* \\ \langle A_{\tau} \mathbf{y}, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} &:= (\mathbf{y}_1, \mathbf{p}_1)_{L^2(\Omega)} + \frac{\tau^2}{4} a(\mathbf{y}_1, \mathbf{p}_1) + \frac{\tau}{2} b(\mathbf{y}_1, \mathbf{p}_1) \\ &+ \sum_{k=1}^{N-1} (\mathbf{y}_{k+1} - \mathbf{y}_k - \tau \bar{v}_k(\mathbf{y}_1, \dots, \mathbf{y}_k), \mathbf{p}_{k+1})_{L^2(\Omega)} \\ &+ \frac{\tau^2}{4} (a(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) + a(\mathbf{y}_k, \mathbf{p}_{k+1})) \\ &+ \frac{\tau}{2} (b(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) - b(\mathbf{y}_k, \mathbf{p}_{k+1})). \end{aligned}$$

(c) The initial data is represented by

$$\begin{aligned} f_{ini} &\in (\mathbf{H}^1(\Omega)^N)^* \\ \langle f_{ini}, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} &:= (\mathbf{y}_{ini}, \mathbf{p}_1)_{\mathbf{L}^2(\Omega)} - \frac{\tau^2}{4} a(\mathbf{y}_{ini}, \mathbf{p}_1) + \frac{\tau}{2} b(\mathbf{y}_{ini}, \mathbf{p}_1) \\ &\quad + \tau \sum_{k=0}^{N-1} (-1)^k (\mathbf{v}_{ini}, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} + 2 \sum_{k=1}^{N-1} (-1)^k (\mathbf{y}_{ini}, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)}. \end{aligned}$$

(d) The combined external data is represented by

$$f_\tau := f_{ini} + \frac{\tau^2}{2} f.$$

Note that all initial data, i.e., the constant parts of the operators involved in the problem formulation, are included in $f_{ini} \in (\mathbf{H}^1(\Omega)^N)^*$. With the notation above, Problem 4.1.4 can be reduced to the following compact formulation in the displacements only, which is examined in the remainder of this chapter.

Problem 4.2.3 (Reduced Discretized Contact Problem) Given $\mathbf{y}_{ini} \in \mathbf{K}_\Phi$, $\mathbf{v}_{ini} \in \mathbf{H}^1(\Omega)$ and $f \in (\mathbf{H}^1(\Omega)^N)^*$, find $\mathbf{y} \in \mathbf{H}^1(\Omega)^N$ and $\lambda \in (\mathbf{H}^1(\Omega)^N)^*$ that solve

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi^N \\ \lambda &\in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ \end{aligned} \quad \text{or equivalently} \quad \begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi^N \\ f_\tau - A_\tau \mathbf{y} &\in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ. \end{aligned}$$

$$A_\tau \mathbf{y} - f_\tau + \lambda = 0$$

The time stepping structure is clearly preserved in this reduced and discretized form. The problem is equivalent to the contact implicit symmetric Newmark scheme, which means it corresponds to a sequence of elliptic variational inequalities where the right hand sides depend on all previous displacements. More specifically, consider the bilinear form

$$\begin{aligned} d: \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) &\rightarrow \mathbb{R} \\ d(\mathbf{y}_{k+1}, \cdot) &:= (\mathbf{y}_{k+1}, \cdot)_{\mathbf{L}^2(\Omega)} + \frac{\tau^2}{4} a(\mathbf{y}_{k+1}, \cdot) + \frac{\tau}{2} b(\mathbf{y}_{k+1}, \cdot) \end{aligned}$$

and the linear, bounded maps $l_{k+1}: \mathbf{H}^1(\Omega)^k \times (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^*$ with

$$\begin{aligned} l_1(w) &:= \langle w_1, \cdot \rangle_{\mathbf{H}^1(\Omega)}, \\ l_{k+1}(\mathbf{y}, w) &:= (\mathbf{y}_k, \cdot)_{\mathbf{L}^2(\Omega)} - \frac{\tau^2}{4} a(\mathbf{y}_k, \cdot) + \frac{\tau}{2} b(\mathbf{y}_k, \cdot) + \tau (\bar{v}_k(\mathbf{y}), \cdot)_{\mathbf{L}^2(\Omega)} \\ &\quad \langle w_{k+1}, \cdot \rangle_{\mathbf{H}^1(\Omega)} \end{aligned}$$

for $k = 1, \dots, N-1$, which map a right hand side of the time stepping problem to the

right hand side of a timestep. Then we have the representation

$$\begin{aligned}
\langle A_\tau \mathbf{y} - w, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} &= (\mathbf{y}_1, \mathbf{p}_1)_{\mathbf{L}^2(\Omega)} + \frac{\tau^2}{4} a(\mathbf{y}_1, \mathbf{p}_1) + \frac{\tau}{2} b(\mathbf{y}_1, \mathbf{p}_1) \\
&\quad + \sum_{k=1}^{N-1} [(\mathbf{y}_{k+1} - \mathbf{y}_k - \tau \bar{\mathbf{v}}_k(\mathbf{y}_1, \dots, \mathbf{y}_k), \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} \\
&\quad \quad + \frac{\tau^2}{4} (a(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) + a(\mathbf{y}_k, \mathbf{p}_{k+1})) \\
&\quad \quad + \frac{\tau}{2} (b(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) - b(\mathbf{y}_k, \mathbf{p}_{k+1}))] \\
&\quad - \langle w, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} \\
&= d(\mathbf{y}_1, \mathbf{p}_1) - \langle l_1(w), \mathbf{p}_1 \rangle_{\mathbf{H}^1(\Omega)} \\
&\quad + \sum_{k=1}^{N-1} d(\mathbf{y}_{k+1}, \mathbf{p}_{k+1}) - \langle l_{k+1}(\mathbf{y}_1, \dots, \mathbf{y}_k, w), \mathbf{p}_{k+1} \rangle_{\mathbf{H}^1(\Omega)}
\end{aligned} \tag{4.9}$$

for all w in $(\mathbf{H}^1(\Omega)^N)^*$. Accordingly, having solved for $\mathbf{y}_1, \dots, \mathbf{y}_k$, the displacement \mathbf{y}_{k+1} of the following timestep can be computed by solving the variational inequality

$$\begin{aligned}
&\mathbf{y}_{k+1} \in \mathbf{K}_\Phi \\
&l_{k+1}(\mathbf{y}_1, \dots, \mathbf{y}_k, w) - d(\mathbf{y}_{k+1}, \cdot) \in \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y}_{k+1})^\circ.
\end{aligned} \tag{4.10}$$

Setting $w := f_\tau = f_{\text{ini}} + \frac{\tau^2}{2} f$ results in the time-discretized contact problem (Problem 4.2.3).

4.3 Solutions to the Time-Discretized Problem

This section focuses on the existence of unique solutions to the time-discretized contact problem. Based on the time stepping interpretation of the variational inequality, we will establish the existence of a Lipschitz continuous solution operator to Problem 4.2.3. We start out proving existence of a solution operator to the variational inequalities in each time step first and use these in the representation of the solution operator to the entire variational inequality.

Lemma 4.3.1 (Preliminaries)

(a) *The bilinear form*

$$\begin{aligned}
d: \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) &\rightarrow \mathbb{R} \\
d(\cdot, \cdot) &:= (\cdot, \cdot)_{\mathbf{L}^2(\Omega)} + \frac{\tau^2}{4} a(\cdot, \cdot) + \frac{\tau}{2} b(\cdot, \cdot)
\end{aligned}$$

is bounded and coercive. The associated linear operator

$$\begin{aligned}
D: \mathbf{H}^1(\Omega) &\rightarrow \mathbf{H}^1(\Omega)^* \\
D\mathbf{y} &:= d(\mathbf{y}, \cdot)
\end{aligned}$$

is surjective and has a bounded linear inverse.

(b) There exists a Lipschitz continuous solution operator

$$s: \mathbf{H}^1(\Omega)^* \rightarrow \mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega)$$

that maps $l \in \mathbf{H}^1(\Omega)^*$ to the solution $\mathbf{y} \in \mathbf{K}_\Phi$ of the variational inequality

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi \\ l - D\mathbf{y} &\in \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})^\circ. \end{aligned}$$

(c) For $k \in \{0, \dots, N-1\}$, the operators $l_{k+1}: \mathbf{H}^1(\Omega)^k \times (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^*$ defined by

$$\begin{aligned} l_1(w) &:= \langle w_1, \cdot \rangle_{\mathbf{H}^1(\Omega)}, \\ l_{k+1}(\mathbf{y}, w) &:= (\mathbf{y}_k, \cdot)_{\mathbf{L}^2(\Omega)} - \frac{\tau^2}{4} a(\mathbf{y}_k, \cdot) + \frac{\tau}{2} b(\mathbf{y}_k, \cdot) + \tau (\bar{v}_k(\mathbf{y}), \cdot)_{\mathbf{L}^2(\Omega)} \\ &\quad + \langle w_{k+1}, \cdot \rangle_{\mathbf{H}^1(\Omega)} \end{aligned}$$

for $k \in \{1, \dots, N-1\}$ are linear and bounded.

Proof. In part (a), boundedness of d is an immediate consequence of the boundedness of the bilinear forms $(\cdot, \cdot)_{\mathbf{L}^2(\Omega)}$, a and b , since

$$|d(\mathbf{y}, \mathbf{v})| \leq \left| (\mathbf{y}, \mathbf{v})_{\mathbf{L}^2(\Omega)} \right| + \frac{\tau^2}{4} |a(\mathbf{y}, \mathbf{v})| + \frac{\tau}{2} |b(\mathbf{y}, \mathbf{v})| \leq C \|\mathbf{y}\|_{\mathbf{H}^1(\Omega)} \|\mathbf{v}\|_{\mathbf{H}^1(\Omega)}$$

for a constant $C = C(\tau)$ and all $\mathbf{y}, \mathbf{v} \in \mathbf{H}^1(\Omega)$. Due to Assumption 4.0.1 (g), Korn's second inequality [110, Thm. 5.13] yields

$$d(\mathbf{y}, \mathbf{y}) = (\mathbf{y}, \mathbf{y})_{\mathbf{L}^2(\Omega)} + \frac{\tau^2}{4} a(\mathbf{y}, \mathbf{y}) + \frac{\tau}{2} b(\mathbf{y}, \mathbf{y}) \geq c \|\mathbf{y}\|_{\mathbf{H}^1(\Omega)}^2$$

for a constant $c = c(\tau)$ and all $\mathbf{y} \in \mathbf{H}^1(\Omega)$, proving coercivity. The Lax-Milgram Lemma yields the first proposition.

Part (b) follows since \mathbf{K}_Φ is closed and convex in $\mathbf{H}^1(\Omega)$, see Proposition 3.5.3. Therefore the fundamental theorem of variational analysis by Lions and Stampacchia [130, Thm. 2.1] yields the existence of the Lipschitz continuous solution operator $s: \mathbf{H}^1(\Omega)^* \rightarrow \mathbf{K}_\Phi$.

For l_1 , the claim in part (c) is clearly satisfied. For l_{k+1} with $k \in \{1, \dots, N-1\}$, linearity is due to the linearity of \bar{v}_k . The boundedness of the bilinear forms a and b (see Assumption 4.0.1 (g)) additionally yields the existence of $C_a, C_b > 0$ such that

$$\begin{aligned} \left| \langle l_{k+1}(\mathbf{y}, w), \varphi \rangle_{\mathbf{H}^1(\Omega)} \right| &\leq \|\mathbf{y}_k\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} + \frac{\tau^2}{4} C_a \|\mathbf{y}_k\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} \\ &\quad + \frac{\tau}{2} C_b \|\mathbf{y}_k\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} + \tau \|\bar{v}_k(\mathbf{y})\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} \\ &\quad + \|w_{k+1}\|_{\mathbf{H}^1(\Omega)^*} \|\varphi\|_{\mathbf{H}^1(\Omega)} \end{aligned}$$

for all $\varphi \in \mathbf{H}^1(\Omega)$, which yields well-definedness and boundedness of the operator due to the boundedness of $\bar{v}_k: \mathbf{H}^1(\Omega)^k \rightarrow \mathbf{H}^1(\Omega)$, see Definition 4.2.2 (a). \square

Existence of a solution to each of the time steps in the time stepping scheme is direct consequence of the previous lemma.

Lemma 4.3.2 (Time Step Solution) *Let $w \in (\mathbf{H}^1(\Omega)^N)^*$, and let additionally $k \in \{0, \dots, N-1\}$ and $\mathbf{y}_i \in \mathbf{K}_\Phi$ for $1 \leq i \leq k$ be the solutions of the timesteps i of the time stepping scheme in the problem*

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi^N \\ w - A_\tau \mathbf{y} &\in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ. \end{aligned}$$

Then there exists a unique solution $\mathbf{y}_{k+1} \in \mathbf{K}_\Phi$ to time step $k+1$, which is given by

$$\mathbf{y}_{k+1} = \begin{cases} s(l_1(w)), & \text{if } k = 0 \\ s(l_{k+1}(\mathbf{y}_1, \dots, \mathbf{y}_k, w)), & \text{if } 1 \leq k \leq N-1. \end{cases}$$

Proof. The time stepping structure implied by (4.9) and (4.10) and Lemma 4.3.1 (b) yields the existence of unique solutions and their specific form. \square

The solution to the complete discretized problem naturally follows from the time step solutions.

Theorem 4.3.3 (Solution Operator) *Let Assumption 4.0.1 be satisfied. Then there exists a Lipschitz continuous solution operator*

$$S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{K}_\Phi^N \subset \mathbf{H}_D^1(\Omega)^N$$

that maps $w \in (\mathbf{H}^1(\Omega)^N)^$ to the solution $\mathbf{y} \in \mathbf{K}_\Phi^N$ of the variational inequality*

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi^N \\ w - A_\tau \mathbf{y} &\in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ. \end{aligned}$$

Proof. We can recursively define the solution operator to the state problem as

$$\begin{aligned} S: (\mathbf{H}^1(\Omega)^N)^* &\rightarrow \mathbf{K}_\Phi^N \subset \mathbf{H}_D^1(\Omega)^N \\ w &\mapsto \mathbf{y} = S(w), \end{aligned}$$

where $\mathbf{y}_k = S_k(w)$ with

$$\begin{aligned} S_k: (\mathbf{H}^1(\Omega)^N)^* &\rightarrow \mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega) \\ S_k(w) &= s\left(\tilde{l}_k(w)\right) \end{aligned}$$

for $k = 1, \dots, N$ is the solution to each of the elliptic variational inequalities in the time steps of the scheme. The right hand sides in each step are given by

$$\begin{aligned} \tilde{l}: (\mathbf{H}^1(\Omega)^N)^* &\rightarrow \mathbf{H}^1(\Omega)^* \\ \tilde{l}_1(w) &= l_1(w), \\ \tilde{l}_k(w) &= l_k(S_1(w), \dots, S_{k-1}(w), w) \quad \text{for } k \in \{2, \dots, N\}, \end{aligned} \tag{4.11}$$

see Lemma 4.3.1 (c) and (4.9). Lemmas 4.3.1 and 4.3.2 imply that $s: \mathbf{H}^1(\Omega)^* \rightarrow \mathbf{K}_\Phi$ and

the right hand sides $l_k: \mathbf{H}^1(\Omega)^k \times (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^*$ are Lipschitz continuous. Since the component mappings S_k , $k = 1, \dots, N$, are the composition of Lipschitz continuous mappings, Lipschitz continuity immediately follows inductively. The Lipschitz continuity of $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{K}_{\Phi}^N$ is due to the Lipschitz continuity of each component mapping. \square

We have established the existence of a Lipschitz continuous solution operator using the time stepping structure (Lemma 4.3.2 and Theorem 4.3.3). The computed Lipschitz constant of the operator $S: \mathbf{H}^1(\Omega)^{N^*} \rightarrow \mathbf{H}^1(\Omega)^N$ is of therefore dependent on the Lipschitz constant of the time stepping solution operators. They depend on the constants of the solution operators $s: \mathbf{H}^1(\Omega)^* \rightarrow \mathbf{H}^1(\Omega)$ and the operators $\tilde{l}_k: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^*$, i.e., generally on the coercivity constants of a and b as well as the time step size τ and the number of time steps. A straight forward estimation of the Lipschitz constants yields no bound w.r.t. τ and N , and this is expected to be an aspect that requires a significant amount of work in approaches of passing to the time-continuous problem in the limit.

4.4 Differentiability of the Solution Operator

Having established the solution operator, we can now address the question of differentiability, which is of importance in the next chapter, where first order optimality conditions are derived. In the first subsection, we will establish Hadamard differentiability, and therefore directional differentiability, of the solution operator $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$. Similarly to the previous section, the proof is based on the time stepping structure, i.e., Hadamard differentiability of the time stepping solution operator $s: \mathbf{H}^1(\Omega)^* \rightarrow \mathbf{K}_{\Phi}$ is established and extended to S . A closer examination of the cones that arise in the linearized problems yields an interpretation as a contact problem with sliding boundary conditions, which shows that Gâteaux differentiability of the solution operator can generally not hold everywhere. However, we can guarantee Fréchet differentiability on a dense subset of $(\mathbf{H}^1(\Omega)^N)^*$.

4.4.1 Hadamard Differentiability

Due to the recursive definition of the component mappings S_k in (4.11), obtaining differentiability of S from the differentiability of the time stepping solution operator is reliant on the validity of a chain rule, which is generally invalid for directionally differentiable functions, see the example [174, P. 484] and [14, Ex. 1.22]. The key differentiability concept to obtain a chain rule for functions that are not Fréchet differentiable is the concept of *Hadamard differentiability*, which has been addressed in, e.g., [51, 59, 174]

Definition 4.4.1 (Hadamard Differentiability) *Let X, Y be Banach spaces, and let $x, dx \in X$. An operator $F: X \rightarrow Y$ is called directionally differentiable in the sense of Hadamard at x in direction dx if*

$$\lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{F(x + \delta x(t)) - F(x)}{t} = F'(x, dx) \in X$$

for all functions $\delta x: [0, \infty) \rightarrow X$ with $\delta x(0) = 0$ and $\frac{\delta x(t) - t dx}{t} \xrightarrow{t \rightarrow 0, t > 0} 0$. The element $F'(x, dx) \in X$ is called the directional derivative. F is called Hadamard differentiable at x if it is Hadamard differentiable at x in all directions $\delta x \in X$ and F is called Hadamard differentiable if it is Hadamard differentiable at all $x \in X$.

The functions δx parameterize curves with end point 0 that are tangential to dx . Of course, choosing $\delta x(t) = t dx$, we can see that Hadamard differentiability immediately implies directional differentiability and that the Hadamard directional derivative always coincides with the directional derivative if they exist. The essential properties of Hadamard differentiable functionals for our purpose are recalled in the following lemma for convenience of the reader.

Lemma 4.4.2 *Let X, Y, Z be Banach spaces, let the mappings $F: Y \rightarrow Z$ and $G: X \rightarrow Y$ be given and let $x, dx \in X$.*

- (a) *If G is Lipschitz continuous and directionally differentiable at $x \in X$, then G is Hadamard differentiable at x .*
- (b) *If F is Hadamard differentiable at $G(x)$ and G is directionally differentiable at x , then $F \circ G: X \rightarrow Z$ is directionally differentiable at x and*

$$(F \circ G)'(x, dx) = F'(G(x), G'(x, dx)).$$

- (c) *If F is Hadamard differentiable at $G(x)$ and G is Hadamard differentiable at x , then $F \circ G: X \rightarrow Z$ is Hadamard differentiable at x and*

$$(F \circ G)'(x, dx) = F'(G(x), G'(x, dx)).$$

Proof. See Lemma C.3.1 of the appendix. Cf. also [174]. □

Hadamard Differentiability and Polyhedricity

As the operators $s: \mathbf{H}_D^1(\Omega)^* \rightarrow \mathbf{K}_\Phi$ in the time stepping scheme are Lipschitz continuous, Hadamard differentiability follows from directional differentiability, which we will establish using Mignot's central result in [138, Sec. 2]. A key property in that respect is that of *polyhedricity* of the set of admissible displacements, which is defined as follows:

Definition 4.4.3 (Polyhedricity) *Let Y be a Banach space. A closed and convex set $K \subset Y$ is called polyhedric at $(y, \lambda) \in K \times Y^*$ if*

$$\overline{\mathcal{R}_K(y) \cap \{\lambda\}^\perp} = \overline{\mathcal{R}_K(y)} \cap \{\lambda\}^\perp. \quad (4.12)$$

The set K is called polyhedric at $y \in K$ if it is polyhedric at (y, λ) for all $\lambda \in Y^$ and K is called polyhedric if it is polyhedric at all $y \in K$.*

Polyhedricity ensures that elements in the tangent cone that are contained in a “hyperplane” (the annihilator) can be approximated by elements in the intersection of the radial cone and the same hyperplane. Intuitively, a set is locally polyhedric if it locally contains the parts of the hyperplane that it is tangent to. Polyhedricity should not be confused with *polyhedricity* — the property of a set being the finite intersection of half spaces — as even in finite dimensions, there are polyhedric sets that are not polyhedral. Confer Wachsmuth’s recent work [195] for an overview of the state-of-the-art of polyhedricity, various new results on the polyhedricity and related concepts as well as several interesting (counter-)examples.

Remark 4.4.4 *Note that polyhedricity is commonly defined and examined for elements $\lambda \in \mathcal{T}_K(y)^\circ$ only, but by [195, Lem. 4.1], the condition (4.12) holds for all $\lambda \in \mathcal{T}_K(y)^\circ$ if and only if it holds for all $\lambda \in Y^*$.*

Mignot’s result shows directional differentiability of projections onto polyhedric sets in Dirichlet spaces.

Theorem 4.4.5 ([138, Thm. 2.1]) *Let V be a Hilbert space, $d: V \times V \rightarrow \mathbb{R}$ be bilinear, bounded and coercive, the set $K \subset V$ be closed and convex and $v \in V$. Assume $y = P_K^d(v)$ to be the $d(\cdot, \cdot)$ -orthogonal projection of v onto K . If K is polyhedric w.r.t. $(y, d(y - v, \cdot))$, then the projection operator P_K^d is directionally differentiable at v , and the derivative at v in direction δv can be computed as the $d(\cdot, \cdot)$ -orthogonal projection of δv onto $\mathcal{C}_K(y, d(y - v, \cdot))$ — the critical cone to K w.r.t. $(y, d(y - v, \cdot))$.*

As stated in Lemma 4.3.1 (a), the operator $D: \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}^1(\Omega)^*$ has a bounded linear inverse, therefore \mathbf{y} solves

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi \\ l - D\mathbf{y} &\in \mathcal{T}_{\mathbf{K}_\Phi}(y)^\circ \end{aligned}$$

for $l \in \mathbf{H}^1(\Omega)^*$ if and only if

$$d(D^{-1}l - \mathbf{y}, v - \mathbf{y}) \geq 0 \quad \forall v \in \mathbf{K}_\Phi,$$

which means that $s = P_{\mathbf{K}_\Phi}^d \circ D^{-1}$. Accordingly, the previous theorem yields the directional differentiability of $s: \mathbf{H}_D^1(\Omega)^* \rightarrow \mathbf{K}_\Phi$ together with an explicit expression for its derivative, provided that the set of admissible displacements \mathbf{K}_Φ in Definition 3.5.4 is polyhedric.

Polyhedricity of sets in Lebesgue and Sobolev spaces has been considered on various occasions. See [181] for results in $H^{-1/2}(\partial\Omega)$, [160] for $H^2(\Omega)$ and [85, 138] for bounded sets in Dirichlet spaces. Polyhedricity in vector valued Sobolev spaces is slightly more complicated but was examined previously as well. An approach using measure representations of dual space elements can be found in [28, Sec. 3.2.2]. Other works share the common approach of investigating the transfer of the polyhedricity property between sets under linear operators, i.e., bounded linear operators $L: Y \rightarrow Z$ between Banach spaces and sets $K \subset Y$, $C \subset Z$ with $K = L^{-1}C$ such that polyhedricity of C implies polyhedricity of K , see [182, Sec. 4.6.1], [31, Prop. 3.54] and [143, Sec. 4.3]. All of the above are limited by a surjectivity requirement at some point in the argument. In the setting of contact problems, the set C is usually chosen as a subset of scalar valued Sobolev functions, and the surjectivity

requirement directly translates to restrictions on the geometry, i.e., higher regularity of the domain's boundary [28] or the ability to extend the contact normal to either Ω or at least a small extension subdomain of Ω *while preserving the pointwise normalization and the Lipschitz continuity* [143, 182]. However, Wachsmuth [195] managed to weaken the requirement of surjectivity.

Lemma 4.4.6 ([195, Lem. 3.3]) *Let Y, Z be Banach spaces and $L: Y \rightarrow Z$ be linear and bounded. Let $C \subset Z$ be closed and convex and $K = L^{-1}C$, $y \in K$ and $\lambda \in \mathcal{T}_K(y)^\circ$. If*

$$LY - \mathcal{R}_C(Ly) = Z,$$

then there exists $\mu \in \mathcal{T}_C(Ly)^\circ$ such that $\lambda = L^\mu$. If additionally*

$$LY - (\mathcal{R}_C(Ly) \cap \{\mu\}^\perp) = Z$$

and C is polyhedral at (Ly, μ) , then K is polyhedral at (y, λ) .

The corresponding setting for the contact problem, which does not require further assumptions on the geometry, is developed in the next paragraph.

Polyhedricity of the Set of Admissible Displacements

Since $\nu_\Phi \in C^{0,1}(\Gamma_C)$, the McShance-Whitney-Kirszbraun Theorem (Theorem A.3.9) guarantees the existence of an extension $\tilde{\nu}_\Phi \in C^{0,1}(\bar{\Omega})$ with $\tilde{\nu}_\Phi = \nu_\Phi$ on Γ_C , which is assumed to be fixed from now on. Note that while $\|\tilde{\nu}_\Phi(\mathbf{x})\| = 1$ on Γ_C (Assumption 4.0.1 (e)), $\|\tilde{\nu}_\Phi(\mathbf{x})\| > 0$ is not guaranteed on $\bar{\Omega}$, so $\tilde{\nu}_\Phi$ need not be normalizable, which generally impedes surjectivity of the operator

$$\begin{aligned} L: \mathbf{H}_D^1(\Omega) &\rightarrow H_D^1(\Omega) \\ Ly &:= \mathbf{y} \cdot \tilde{\nu}_\Phi. \end{aligned} \tag{4.13}$$

Extending the initial gap $\Psi: \Gamma_C \rightarrow \bar{\mathbb{R}}$ as

$$\begin{aligned} \tilde{\Psi}: \bar{\Omega} &\rightarrow \bar{\mathbb{R}} \\ \tilde{\Psi}(\mathbf{x}) &:= \begin{cases} \Psi(\mathbf{x}), & \mathbf{x} \in \Gamma_C \\ \infty, & \mathbf{x} \in \bar{\Omega} \setminus \Gamma_C, \end{cases} \end{aligned}$$

we can rewrite the set of admissible displacements in Definition 3.5.4 as

$$\mathbf{K}_\Phi = \left\{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \tilde{\nu}_\Phi \leq \tilde{\Psi} \text{ q.e. on } \bar{\Omega} \right\} \subset \mathbf{H}_D^1(\Omega). \tag{4.14}$$

Note that the reformulation maintains consistency with respect to the notions of quasi everywhere in Definition 3.4.9.

The setting for the transfer of polyhedricity amounts to the following:

Lemma 4.4.7 *Let $\mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega)$, $L: \mathbf{H}_D^1(\Omega) \rightarrow H_D^1(\Omega)$ as in (4.13)–(4.14) and*

$$C := \{v \in H_D^1(\Omega) : v \leq \Psi \text{ q.e. on } \Gamma_C\} = \left\{v \in H_D^1(\Omega) : v \leq \tilde{\Psi} \text{ q.e. on } \bar{\Omega}\right\}. \quad (4.15)$$

Then:

- (a) \mathbf{K}_Φ and C are closed and convex.
- (b) L is linear and bounded.
- (c) $\mathbf{K}_\Phi = L^{-1}C$.

Proof. For closedness and convexity of \mathbf{K}_Φ and C , see the proof in Proposition 3.5.3.

Linearity of L is clear, and due to the Lipschitz continuity of $\tilde{\nu}_\Phi$, Corollary B.3.4 yields well definedness, i.e., preservation of the Dirichlet boundary values and boundedness of L . Therefore, the inclusion $\mathbf{K}_\Phi \subset L^{-1}C$ in part (c) follows, and the choice of the domain of L ensures that $\mathbf{K}_\Phi \supset L^{-1}C$ as well. \square

Note that choosing $\mathbf{H}_D^1(\Omega)$ as the domain of L instead of $\mathbf{H}^1(\Omega)$ is (technically) essential to obtain part (c) of the previous lemma, as $\tilde{\nu}_\Phi$ might take the value zero on Γ_D . This turns out to be an unproblematic restriction, however, as we will see shortly.

Verification of the assumptions on K , C and L in Lemma 4.4.6 is straight forward in the setting at hand. In fact, we have the following Lemma.

Lemma 4.4.8 *Let $\mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega)$, $C \subset H_D^1(\Omega)$, $L: \mathbf{H}_D^1(\Omega) \rightarrow H_D^1(\Omega)$ be given as in (4.13)–(4.15). Then*

$$L\mathbf{H}_D^1(\Omega) - (\mathcal{R}_C(L\mathbf{y}) \cap \{\mu\}^\perp) = H_D^1(\Omega) \quad \forall \mathbf{y} \in \mathbf{K}_\Phi, \mu \in \mathcal{T}_C(L\mathbf{y})^\circ.$$

Proof. Let $\mathbf{y} \in \mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega)$, $\mu \in \mathcal{T}_C(L\mathbf{y})^\circ \subset H_D^1(\Omega)^*$ and $v \in H_D^1(\Omega)$. Again, Lipschitz continuity of $\tilde{\nu}_\Phi$ and Corollary B.3.4 yield that $\tilde{\mathbf{v}} := v\tilde{\nu}_\Phi$ is in $\mathbf{H}_D^1(\Omega)$, and of course

$$v = L\tilde{\mathbf{v}} - (L\tilde{\mathbf{v}} - v). \quad (4.16)$$

As $\tilde{\nu}_\Phi$ coincides with the contact *normal* ν_Φ on Γ_C , we know that $L\tilde{\mathbf{v}} - v = v\tilde{\nu}_\Phi \cdot \tilde{\nu}_\Phi - v = 0$ q.e. on Γ_C , cf. Lemma B.3.3. By definition of C and because $L\mathbf{y} \in C$, we deduce that

$$L\mathbf{y} + \alpha(L\tilde{\mathbf{v}} - v) \in C \quad \forall \alpha \in \mathbb{R},$$

which means that

$$L\tilde{\mathbf{v}} - v \in \mathcal{R}_C(L\mathbf{y}) \quad \text{and} \quad v - L\tilde{\mathbf{v}} \in \mathcal{R}_C(L\mathbf{y}).$$

Because $\mu \in \mathcal{T}_C(L\mathbf{y})^\circ$, we immediately obtain that $\langle \mu, L\tilde{\mathbf{v}} - v \rangle_{H_D^1(\Omega)} = 0$ and therefore $L\tilde{\mathbf{v}} - v \in \mathcal{R}_C(L\mathbf{y}) \cap \{\mu\}^\perp$, and (4.16) yields the claim. \square

Lemma 4.4.8 of course implies that $L\mathbf{H}_D^1(\Omega) - \mathcal{R}_C(L\mathbf{y}) = H_D^1(\Omega)$ for all $\mathbf{y} \in \mathbf{K}_\Phi$ as well. Hence, we can transfer polyhedricity from the set C of scalar valued functions — which is easily established — to the set \mathbf{K}_Φ of vector valued functions.

Lemma 4.4.9 *The set of admissible displacements $\mathbf{K}_\Phi \subset \mathbf{H}_D^1(\Omega)$ is polyhedral as a subset of $\mathbf{H}_D^1(\Omega)$, i.e.,*

$$\overline{\mathcal{R}_K(\mathbf{y}) \cap \{\lambda\}^\perp} = \overline{\mathcal{R}_K(\mathbf{y})} \cap \{\lambda\}^\perp \quad \forall \mathbf{y} \in \mathbf{K}_\Phi, \lambda \in \mathbf{H}_D^1(\Omega)^*,$$

where $\{\lambda\}^\perp = \left\{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \langle \lambda, \mathbf{y} \rangle_{\mathbf{H}_D^1(\Omega)} = 0 \right\}$ is the annihilator in $\mathbf{H}_D^1(\Omega)$.

Proof. $\mathbf{H}_D^1(\Omega)$ is a Hilbert space in which the natural pointwise ordering, equivalently in the \mathcal{L}^d -a.e. and the $\text{cap}_{\Omega, D, 2}$ -q.e. sense, see Lemma 3.1.15, induces a lattice structure. The corresponding max-operator is bounded [11, Sec. 5.8.1], and the set C is a set with upper bound in the sense of [195, Def. 4.9]. Therefore, [195, Thm. 4.18, Lem. 4.1] yield the polyhedricity of C in $\mathbf{H}_D^1(\Omega)$.

Due to Lemma 4.4.8 and the reformulation of the admissible set in (4.14), we can employ the transfer of polyhedricity in Lemma 4.4.6 to obtain polyhedricity of \mathbf{K}_Φ in $\mathbf{H}_D^1(\Omega)$. \square

Note that we initially obtain polyhedricity in $\mathbf{H}_D^1(\Omega)$ due to the technical restrictions needed for Lemma 4.4.7. Of course, with $\mathbf{H}_D^1(\Omega)$ being a closed subspace of $\mathbf{H}^1(\Omega)$, we immediately obtain polyhedricity of \mathbf{K}_Φ in $\mathbf{H}^1(\Omega)$ as well.

Corollary 4.4.10 (Polyhedricity of \mathbf{K}_Φ) *The set of admissible displacements*

$$\mathbf{K}_\Phi = \left\{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq \Psi \text{ q.e. on } \Gamma_C \right\}$$

is polyhedral in $\mathbf{H}^1(\Omega)$.

Proof. Since $\mathbf{H}_D^1(\Omega)$ is a closed subspace of $\mathbf{H}^1(\Omega)$, we have $\mathcal{R}_{\mathbf{K}_\Phi}(\mathbf{y}) \subset \overline{\mathcal{R}_{\mathbf{K}_\Phi}(\mathbf{y})} \subset \mathbf{H}_D^1(\Omega)$. Due to Lemma 4.4.9,

$$\begin{aligned} \overline{\mathcal{R}_K(\mathbf{y}) \cap \{\lambda\}^\perp} &= \overline{\mathcal{R}_K(\mathbf{y}) \cap (\{\lambda\}^\perp \cap \mathbf{H}_D^1(\Omega))} \\ &= \overline{\mathcal{R}_K(\mathbf{y})} \cap (\{\lambda\}^\perp \cap \mathbf{H}_D^1(\Omega)) = \overline{\mathcal{R}_K(\mathbf{y})} \cap \{\lambda\}^\perp \end{aligned}$$

for all $\mathbf{y} \in \mathbf{K}_\Phi, \lambda \in \mathbf{H}^1(\Omega)^* \subset \mathbf{H}_D^1(\Omega)^*$, where

$$\{\lambda\}^\perp = \left\{ \mathbf{y} \in \mathbf{H}^1(\Omega) : \langle \lambda, \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)} = 0 \right\}$$

is the annihilator in $\mathbf{H}^1(\Omega)$. \square

Having established the polyhedricity of the set of admissible displacements, we can now address the Hadamard differentiability of the solution operator for the time stepping scheme.

Hadamard Differentiability of S

Hadamard differentiability of the solution operators is obtained as sketched in the introduction of this subsection, i.e., via Hadamard differentiability of the time stepping solution.

Lemma 4.4.11 *Let Assumption 4.0.1 be satisfied. Then the operator*

$$s: \mathbf{H}^1(\Omega)^* \rightarrow \mathbf{K}_\Phi$$

that maps $l \in \mathbf{H}^1(\Omega)^$ to the unique solution $\mathbf{y} := s(l) \in \mathbf{K}_\Phi$ of*

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_\Phi \\ l - D\mathbf{y} &\in \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})^\circ \end{aligned}$$

is Lipschitz continuous and Hadamard differentiable. Further, the Hadamard derivative $\delta\mathbf{y} := s'(l, \delta l)$ is the unique solution to the problem

$$\begin{aligned} \delta\mathbf{y} &\in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, (l - D\mathbf{y})) \\ \delta l - D\delta\mathbf{y} &\in \mathcal{T}_{\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, (l - D\mathbf{y}))}(\delta\mathbf{y})^\circ, \end{aligned}$$

and $\delta l \mapsto s'(l, \delta l)$ is Lipschitz continuous.

Proof. See Lemma 4.3.1 (b) for the existence of the Lipschitz continuous solution operator. Since D has a bounded linear inverse, we can write $s = P_{\mathbf{K}_\Phi}^d \circ D^{-1}$, where $P_{\mathbf{K}_\Phi}^d$ is the d -orthogonal projection onto \mathbf{K}_Φ . Due to polyhedricity of \mathbf{K}_Φ , see Corollary 4.4.10, Mignot's theorem (Theorem 4.4.5) yields directional differentiability of $P_{\mathbf{K}_\Phi}^d: \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}^1(\Omega)$ with

$$P_{\mathbf{K}_\Phi}^d{}'(l, \delta l) = P_{\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, (l - D\mathbf{y}))}^d(\delta l).$$

Lipschitz continuity of s and D imply Lipschitz continuity of $P_{\mathbf{K}_\Phi}^d = s \circ D$, and Hadamard differentiability of $P_{\mathbf{K}_\Phi}^d$ follows from Lemma 4.4.2 (a). Since the inverse mapping D^{-1} is linear and bounded, it is trivially Fréchet and therefore Hadamard differentiable, and Lemma 4.4.2 (c) implies Hadamard differentiability of s with

$$s'(l, \delta l) = P_{\mathbf{K}_\Phi}^d{}'(l, D^{-1}\delta l) = P_{\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, (l - D\mathbf{y}))}^d(D^{-1}\delta l).$$

This is precisely the solution operator to the variational equation in the claim, for which [130, Thm. 2.1] provides Lipschitz continuity because the critical cone $\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, (l - D\mathbf{y}))$ is closed and convex. \square

Clearly, the critical cone of the time stepping problem and the critical cones in each of the time steps correspond, see Lemma C.2.2 and the definitions of A_τ and \tilde{l} in (4.9) and (4.11), which yield

$$\begin{aligned} \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, w - A_\tau\mathbf{y}) &= \prod_{k=1}^N \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, (w - A_\tau\mathbf{y})_k) = \prod_{k=1}^N \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \tilde{l}_k(w) - D\mathbf{y}_k), \\ \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, w - A_\tau\mathbf{y})^\circ &= \prod_{k=1}^N \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, (w - A_\tau\mathbf{y})_k)^\circ = \prod_{k=1}^N \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \tilde{l}_k(w) - D\mathbf{y}_k)^\circ \end{aligned}$$

for $w \in (\mathbf{H}^1(\Omega)^N)^*$ and $\mathbf{y} = S(w) \in \mathbf{H}^1(\Omega)^N$. The results for the time step solution operator s can therefore be applied in order to obtain the corresponding information for the solution operator to the time stepping scheme.

Theorem 4.4.12 *Let Assumption 4.0.1 be satisfied. Then the operator*

$$S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{K}_{\Phi}^N \subset \mathbf{H}_D^1(\Omega)^N$$

that maps $w \in (\mathbf{H}^1(\Omega)^N)^$ to the unique solution $\mathbf{y} := S(w) \in \mathbf{K}_{\Phi}^N$ of*

$$\begin{aligned} \mathbf{y} &\in \mathbf{K}_{\Phi}^N \\ w - A_{\tau}\mathbf{y} &\in \mathcal{T}_{\mathbf{K}_{\Phi}^N}(\mathbf{y})^{\circ} \end{aligned}$$

is Lipschitz continuous and Hadamard differentiable. Further, the Hadamard derivative $\delta\mathbf{y} := S'(w, \delta w)$ is the unique solution to the problem

$$\begin{aligned} \delta\mathbf{y} &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y}) \\ \delta w - A_{\tau}\delta\mathbf{y} &\in \mathcal{T}_{\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y})}(\delta\mathbf{y})^{\circ}, \end{aligned} \tag{4.17}$$

and $\delta w \mapsto S'(w, \delta w)$ is Lipschitz continuous.

Proof. Let $w \in (\mathbf{H}^1(\Omega)^N)^*$. Theorem 4.3.3 established the Lipschitz continuity of S with

$$S(w) = \begin{pmatrix} S_1(w) \\ \vdots \\ S_N(w) \end{pmatrix} = \begin{pmatrix} s(\tilde{l}_1(w)) \\ \vdots \\ s(\tilde{l}_N(w)) \end{pmatrix} = \begin{pmatrix} s(l_1(w)) \\ \vdots \\ s(l_N(S_1(w), \dots, S_{N-1}(w), w)) \end{pmatrix},$$

where all l_k are linear and bounded, see Lemma 4.3.1 (c), and therefore Fréchet differentiable with $l'_k = l_k$, which we use in the following induction argument.

For $k = 1$, we have $S_1(w) = s(\tilde{l}_1(w)) = s(l_1(w))$ and Hadamard differentiability of s and l_1 implies Hadamard differentiability of $S_1: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)$ with

$$S'_1(w, \delta w) = s'(\tilde{l}_1(w), \tilde{l}'_1(w, \delta w)) = s'(\tilde{l}_1(w), l_1(\delta w)) \quad \forall \delta w \in (\mathbf{H}^1(\Omega)^N)^*,$$

which is Lipschitz continuous in δw , see Lemma 4.4.11.

For $1 < k \leq N$, we have Hadamard differentiability of $S_i: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)$ with

$$S'_i(w, \delta w) = s'(\tilde{l}_i(w), \tilde{l}'_i(w, \delta w)) \quad \forall \delta w \in (\mathbf{H}^1(\Omega)^N)^*, \quad 1 \leq i < k,$$

which all are Lipschitz continuous in δw . From the representation

$$S_k(w) = s(l_k(S_1(w), \dots, S_{k-1}(w), w)),$$

we again obtain Hadamard differentiability of S_k from Hadamard differentiability of s and l_k with

$$S'_k(w, \delta w) = s'(\tilde{l}_k(w), \tilde{l}'_k(w, \delta w)) = s'(\tilde{l}_k(w), l_k(S'_1(w, \delta w), \dots, S'_{k-1}(w, \delta w), \delta w)),$$

and $\delta w \mapsto S'_k(w, \delta w)$ is Lipschitz continuous in δw .

Hadamard differentiability and Lipschitz continuity of the derivative of S immediately follows the same properties of the component mappings with

$$S'(w, \delta w) = \begin{pmatrix} S'_1(w, \delta w) \\ \vdots \\ S'_N(w, \delta w) \end{pmatrix} = \begin{pmatrix} s'(\tilde{l}_1(w), \tilde{l}'_1(w, \delta w)) \\ \vdots \\ s'(\tilde{l}_N(w), \tilde{l}'_N(w, \delta w)) \end{pmatrix}.$$

Lemma 4.4.11 and the form of the \tilde{l}_k and A_τ yield the representation of the directional derivative. \square

Recall that solving the time-discretized dynamic contact problem for given initial values simply corresponds to setting a specific right hand side for the problems examined above.

Corollary 4.4.13 *Let Assumption 4.0.1 be satisfied, $\mathbf{y}_{ini} \in \mathbf{K}_\Phi$ and $\mathbf{v}_{ini} \in \mathbf{H}^1(\Omega)$. Then there exists a Lipschitz continuous and Hadamard differentiable solution operator that maps the external forces $f \in (\mathbf{H}^1(\Omega)^N)^*$ to the solution to Problem 4.2.3.*

Proof. With $f_\tau, f_{ini} \in (\mathbf{H}^1(\Omega)^N)^*$ as in Definition 4.2.2, the solution operator is

$$S(f_\tau(f)) = S(f_{ini} + \frac{\tau^2}{2}f),$$

and since $f \mapsto f_\tau = f_{ini} + \frac{\tau^2}{2}f$ is clearly Fréchet differentiable, Theorem 4.4.12 yields the claim. \square

4.4.2 Linearized Boundary Conditions

We will now derive a quasi everywhere pointwise characterization of the tangent and critical cones to the set of admissible displacements. The pointwise conditions can be interpreted as boundary conditions for the problem of computing the directional derivatives (4.17) and the adjoint problem that arises as part of the optimality conditions in Section 5.3. Along the way, we develop a measure-based characterization of the normal cones that allows for a more localized interpretation of the residual multiplier λ in its representation of the contact forces on the active contact set in the forward problem. The local characterization allows us to decompose the active contact patch into a subsection where contact forces actively prevent penetration and another where contact is established without participation of contact forces.

We will repeatedly have to deal with sets depending on classes of functions in $H^1(\Omega)$ that are unique for every quasi continuous representative. For a clearer presentation, we adhere to the following convention.

Definition 4.4.14 (Classes of Sets) *Let A denote a class of sets that is uniquely determined up to polar sets. Then a condition involving A is said to hold if and only if it holds up to polar sets for every representative of A .*

In this sense, any set in the class can equivalently be used to represent the class, so we will not distinguish between the set and its representatives explicitly.

Definition 4.4.15 (Active Contact Set) *Let $\mathbf{y} \in \mathbf{K}_\Phi$. Then*

$$\mathcal{A}_\Psi(\mathbf{y}) := \{ \mathbf{x} \in \Gamma_C : \mathbf{y}(\mathbf{x}) \cdot \boldsymbol{\nu}_\Phi(\mathbf{x}) = \Psi(\mathbf{x}) \}$$

is called the (active) contact set of \mathbf{y} .

Recall that Theorem 3.1.17 guarantees that the active contact set is unique up to polar sets.

The pointwise characterization of the cones are obtained from the results in the scalar valued case in [138] using a transfer argument similar to the one applied to obtain polyhedricity in Section 4.4.1. We can reuse the setting based on the operator L in (4.13) that was used in the transfer of polyhedricity. In the case where L is surjective, the correspondence between the scalar and vector valued cones can be computed in a straight forward manner, see [143]. When surjectivity is not guaranteed, the literature provides the following result, cf. Lemma 4.4.6.

Lemma 4.4.16 *Let Y, Z be Banach spaces and $L: Y \rightarrow Z$ be linear and bounded. Let $C \subset Z$ be closed and convex and $K = L^{-1}C$, $y \in K$ and $\lambda \in \mathcal{T}_K(y)^\circ$. If*

$$LY - \mathcal{R}_C(Ly) = Z, \quad (4.18)$$

then

$$\mathcal{T}_K(y) = L^{-1}\mathcal{T}_C(Ly) \quad \text{and} \quad \mathcal{T}_K(y)^\circ = L^*\mathcal{T}_C(Ly)^\circ.$$

Proof. See [31, Cor. 2.91, Prop. 2.95] for the first and [122, Thm. 2.1] for the second part. \square

Hence, we obtain the following characterization of the tangent and the normal cone.

Lemma 4.4.17 *Let $\mathbf{y} \in \mathbf{K}_\Phi$. Then*

$$\mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y}) = \{ \boldsymbol{\delta}\mathbf{y} \in \mathbf{H}_D^1(\Omega) : \boldsymbol{\delta}\mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}) \}.$$

Additionally, a $\mathbf{H}^1(\Omega)^$ -functional λ is in $\mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})^\circ$ if and only if there exists a positive radon measure $\xi \in \mathcal{M}_+(\overline{\Omega}, \mathbb{R})$ that satisfies the following:*

- (a) $\xi(A) = 0$ for all $A \subset \overline{\Omega}$ with $\text{cap}_{\Omega,2}(A) = 0$.
- (b) ξ is concentrated on $\mathcal{A}_\Psi(\mathbf{y})$.
- (c) $\mathbf{H}_D^1(\Omega)$ is continuously embedded into $L^1(\overline{\Omega}; \xi)$.
- (d) We have that

$$\langle \lambda, \boldsymbol{\delta}\mathbf{y} \rangle_{\mathbf{H}^1(\Omega)} = \int_{\mathcal{A}_\Psi(\mathbf{y})} \boldsymbol{\delta}\mathbf{y} \cdot \boldsymbol{\nu}_\Phi \, d\xi \quad \forall \boldsymbol{\delta}\mathbf{y} \in \mathbf{H}_D^1(\Omega),$$

where the integration means any quasi continuous representative of $\boldsymbol{\delta}\mathbf{y}$.

Proof. Recall the scalar-to-vector transfer setting from (4.13)–(4.15), where

$$\begin{aligned} L: \mathbf{H}_D^1(\Omega) &\rightarrow H_D^1(\Omega), \quad L\mathbf{y} := \mathbf{y} \cdot \tilde{\boldsymbol{\nu}}_\Phi, \\ \mathbf{K}_\Phi &= \left\{ \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \mathbf{y} \cdot \tilde{\boldsymbol{\nu}}_\Phi \leq \tilde{\Psi} \text{ q.e. on } \overline{\Omega} \right\}, \\ C &= \left\{ v \in H_D^1(\Omega) : v \leq \tilde{\Psi} \text{ q.e. on } \overline{\Omega} \right\}. \end{aligned}$$

Let $\mathbf{y} \in \mathbf{K}_\Phi$, then we deduce

$$\mathcal{T}_C(L\mathbf{y}) = \{ v \in H_D^1(\Omega) : v \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}) \}$$

from [138, Lem. 3.2] by use of the standard Riesz isomorphism and because $H_D^1(\Omega)$ with the standard $H^1(\Omega)$ -scalar product constitutes a Dirichlet space. See [138, P. 146,147] for a brief overview and [70] for a thorough account of Dirichlet spaces. Due to Lemma 4.4.8, we know that (4.18) is satisfied, therefore Lemma 4.4.16 yields the form of the tangent cone $\mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})$.

In the same $H_D^1(\Omega)$ -Dirichlet space setting, [138, Lem. 3.1] shows that a $\mathbf{H}_D^1(\Omega)^*$ -functional μ is in $\mathcal{T}_C(L\mathbf{y})^\circ$ if and only if there exists a positive radon measure ξ that is concentrated on $\mathcal{A}_\Psi(L\mathbf{y})$ and, in the nomenclature used in the analysis of Dirichlet spaces, is of finite $H_D^1(\Omega)$ -energy integral ([70, Sec. 2.2]) and that satisfies

$$\langle \mu, v \rangle_{\mathbf{H}_D^1(\Omega)} = \int_{\bar{\Omega}} v \, d\xi \quad \forall v \in \mathbf{H}_D^1(\Omega).$$

The fact that $H_D^1(\Omega)$ is continuously embedded in $L^1(\bar{\Omega}; \xi)$ for any ξ of finite $H_D^1(\Omega)$ -energy integral and that $\text{cap}_{\Omega, D, 2}$ -polar sets are null sets of these measures can be found in [70, P. 85,86]. Since ξ is concentrated on $\mathcal{A}_\Psi(L\mathbf{y}) \subset \Gamma_C$ and $\text{cap}_{\Omega, D, 2}$ is equivalent to $\text{cap}_{\Omega, 2}$ on Γ_C , $\text{cap}_{\Omega, 2}$ -polar sets are ξ -null sets as well, cf. Corollary 3.4.7. Accordingly, Lemma 4.4.16 yields the claim. \square

The representation of $\lambda \in \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})^\circ$ as a localized measure provides an interpretation as the force or the impulse being exerted on the body during contact to enforce the non-penetration condition. It shows that the contact forces can only act in the contact normal direction. Note that the representation is only valid for the multiplier evaluated at $\mathbf{H}_D^1(\Omega)$ -functions. Evaluating the residual λ at $\mathbf{H}^1(\Omega)$ -functions, we can expect another localized part on the Dirichlet boundary that can be interpreted as the associated “force” required to ensure the Dirichlet condition. In the scalar problem, we can obtain a signed measure that is concentrated on Γ_D by splitting the Dirichlet conditions into two constraints of lower and upper bound, respectively.

As with all mathematical problems with complementarity constraints, degenerative behavior is observed when both complementary constraints are active. In the contact problem, this corresponds to the body and the obstacle being in contact without contact forces separating the objects. We can obtain a pointwise quasi everywhere characterization of this *biactive* set and the corresponding *weakly* and *strongly* active sets. The argument is based on Stollmann’s characterization of closed ideals in Dirichlet spaces in [186].

Definition 4.4.18 (Ideal) *Let \mathcal{I} be a subspace of a vector lattice (Y, \leq) . If $f \in \mathcal{I}$ and $|g| \leq |f|$ imply that $g \in \mathcal{I}$, then \mathcal{I} is called a closed (order) ideal.*

See [132] for more information. For the measure characterization of the multiplier λ , we now obtain the following.

Lemma 4.4.19 *Let $\xi \in \mathcal{M}_+(\bar{\Omega}, \mathbb{R})$ satisfy Lemma 4.4.17 (a). Then the set*

$$\mathcal{I}_\xi := \{ v \in H_D^1(\Omega) : v = 0 \text{ } \xi\text{-a.e. on } \bar{\Omega} \}$$

is a closed ideal in $H^1(\Omega)$ with respect to the pointwise partial ordering.

Proof. Since $\text{cap}_{\Omega,2}$ -polar sets are ξ -null sets, the set \mathcal{I}_ξ is well defined in the sense that the equality is satisfied by all $\text{cap}_{\Omega,2}$ -quasi continuous representatives. Clearly, \mathcal{I}_ξ is a linear subspace of $H^1(\Omega)$.

Let $f \in \mathcal{I}_\xi$ and $g \in H^1(\Omega)$ with $|g| \leq |f|$ \mathcal{L}^d -a.e. on $\bar{\Omega}$. Due to Lemma 3.1.15, we know that $|g| \leq |f|$ $\text{cap}_{\Omega,2}$ -q.e. on $\bar{\Omega}$ as well, and $|g| \leq |f|$ ξ -a.e. on $\bar{\Omega}$ due to Lemma 4.4.17 (a). Hence, we have $g = 0$ ξ -a.e. on $\bar{\Omega}$, and $g \in H_D^1(\Omega)$ follows using the quasi everywhere characterization of boundary conditions, see Corollary 3.4.10 \square

We can now use Stollmann's characterization of closed ideals in Dirichlet spaces to obtain the following characterization of the strongly active contact set.

Lemma 4.4.20 *Let $\mathbf{y} \in \mathbf{K}_\Phi$ and $\xi \in \mathcal{M}_+(\bar{\Omega}, \mathbb{R})$ satisfy Lemma 4.4.17 (a)–(c). Then there exists a set $\mathcal{S}_\Psi(\mathbf{y}) \subset \bar{\Omega}$, determined and maximal up to $\text{cap}_{\Omega,2}$ -polar sets, such that:*

- (a) $\mathcal{S}_\Psi(\mathbf{y}) \subset \mathcal{A}_\Psi(\mathbf{y})$
- (b) ξ is concentrated on $\mathcal{S}_\Psi(\mathbf{y})$
- (c) $\mathcal{I}_\xi = \{v \in H_D^1(\Omega) : v = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y})\}$

When $\mathcal{A}_\Psi(\mathbf{y})$ is quasi closed, then $\mathcal{S}_\Psi(\mathbf{y})$ is quasi closed.

Proof. Since \mathcal{I}_ξ is a closed ideal in the Dirichlet space $H^1(\Omega)$ with the standard scalar product, [186, Thm. 1, Lem. 2] ensure the existence of a non-negative generating element $F \in \mathcal{I}_\xi \subset H_D^1(\Omega)$ of the ideal such that

$$S := \{ \mathbf{x} \in \bar{\Omega} : F(\mathbf{x}) = 0 \},$$

$$\mathcal{I}_\xi = \{ v \in H^1(\Omega) : v = 0 \text{ q.e. on } S \}.$$

The class of sets S is determined up to $\text{cap}_{\Omega,2}$ -polar sets and quasi closed (Lemma B.3.1). We use S to define

$$\mathcal{S}_\Psi(\mathbf{y}) := S \cap \mathcal{A}_\Psi(\mathbf{y}),$$

which is clearly defined up to $\text{cap}_{\Omega,2}$ -polar sets and satisfies part (a).

Since $F \in \mathcal{I}_\xi$, we know that $F = 0$ ξ -a.e. on $\bar{\Omega}$, and therefore

$$\xi(\bar{\Omega} \setminus S) = \xi(\{ \mathbf{x} \in \bar{\Omega} : F(\mathbf{x}) > 0 \}) = 0.$$

Hence, the measure ξ is concentrated on both S and $\mathcal{A}_\Psi(\mathbf{y})$. It is therefore concentrated on $\mathcal{S}_\Psi(\mathbf{y})$, which proves part (b).

Clearly, the set $\mathcal{S}_\Psi(\mathbf{y})$ is contained in S up to polar sets, therefore

$$\mathcal{I}_\xi = \{ v \in H_D^1(\Omega) : v = 0 \text{ q.e. on } S \} \subset \{ v \in H_D^1(\Omega) : v = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}) \}.$$

Because $\text{cap}_{\Omega,2}$ -polar sets are ξ -null sets and ξ is concentrated on $\mathcal{S}_\Psi(\mathbf{y})$, we additionally obtain that

$$\{ v \in H_D^1(\Omega) : v = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}) \} \subset \{ v \in H_D^1(\Omega) : v = 0 \text{ } \xi\text{-a.e. on } \mathcal{S}_\Psi(\mathbf{y}) \} = \mathcal{I}_\xi,$$

which shows part (c).

When $\mathcal{A}_\Psi(\mathbf{y})$ is quasi closed, then $\mathcal{S}_\Psi(\mathbf{y}) = S \cap \mathcal{A}_\Psi(\mathbf{y})$ is of course quasi closed because of quasi closedness of both S and $\mathcal{A}_\Psi(\mathbf{y})$ and Lemma B.2.3, which yields quasi closedness of $\mathcal{A}_\Psi(\mathbf{y})$.

Since both S and $\mathcal{A}_\Psi(\mathbf{y})$ are determined up to polar sets, the set $\mathcal{S}_\Psi(\mathbf{y})$ is as well. Additionally, because $F \in \mathcal{I}_\xi$ and $F > 0$ q.e. on $\mathcal{A}_\Psi(\mathbf{y}) \setminus S$ by definition, we know that any set $\tilde{\mathcal{S}} \subset \mathcal{A}_\Psi(\mathbf{y})$ such that

$$\mathcal{I}_\xi = \left\{ v \in H_D^1(\Omega) : v = 0 \text{ q.e. on } \tilde{\mathcal{S}} \right\}$$

is contained in $\mathcal{S}_\Psi(\mathbf{y})$ up to polar sets, which means that $\mathcal{S}_\Psi(\mathbf{y})$ is a largest set satisfying the claim, i.e., maximal in that sense. \square

Combining the previous lemmas, we obtain the pointwise characterization of the critical cone.

Proposition 4.4.21 *Let Assumption 4.0.1 be satisfied, $\mathbf{y} \in \mathbf{K}_\Phi$ and $\lambda \in \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y})^\circ$. Then there exists a set $\mathcal{S}_\Psi(\mathbf{y}) \subset \mathcal{A}_\Psi(\mathbf{y})$, determined and maximal up to $\text{cap}_{\Omega,2}$ -polar sets, such that*

$$\begin{aligned} \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, \lambda) = \{ \delta \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}), \\ \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}) \}. \end{aligned}$$

When $\mathcal{A}_\Psi(\mathbf{y})$ is quasi closed, then $\mathcal{S}_\Psi(\mathbf{y})$ is quasi closed.

Proof. Let ξ denote the positive radon measure representing λ on $\mathbf{H}_D^1(\Omega)$ as in Lemma 4.4.17 (d) and $\mathcal{S}_\Psi(\mathbf{y})$ the corresponding set from Lemma 4.4.20. Then we obtain

$$\begin{aligned} \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, \lambda) &= \mathcal{T}_{\mathbf{K}_\Phi}(\mathbf{y}) \cap \{ \lambda \}^\perp \\ &= \left\{ \delta \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}), \int_{\mathcal{A}_\Psi(\mathbf{y})} \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \, d\xi = 0 \right\} \\ &= \left\{ \delta \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}), \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi = 0 \text{ } \xi\text{-a.e. on } \bar{\Omega} \right\} \\ &= \left\{ \delta \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}), \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}) \right\}, \end{aligned}$$

where the third equality is valid because ξ is a positive measure concentrated on $\mathcal{A}_\Psi(\mathbf{y})$ and $\delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0$ q.e. on $\mathcal{A}_\Psi(\mathbf{y})$ implies $\delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0$ ξ -a.e. on $\mathcal{A}_\Psi(\mathbf{y})$, cf. Lemma 4.4.17. \square

Quasi closedness of $\mathcal{A}_\Psi(\mathbf{y})$ can be verified by imposing the following condition.

Corollary 4.4.22 *Let the assumptions of Lemma 4.4.20 hold, and let Γ_C be quasi closed. Then $\mathcal{A}_\Psi(\mathbf{y})$ and $\mathcal{S}_\Psi(\mathbf{y})$ are quasi closed.*

Proof. Due to quasi continuity of \mathbf{y} and the initial gap Ψ , the active contact set $\mathcal{A}_\Psi(\mathbf{y})$ is quasi closed in Γ_C , see Definition 4.4.15 and Lemma B.2.2. The set Γ_C is quasi closed in $\partial\Omega$ if and only if it is quasi closed in $\bar{\Omega}$, and in that case, the set $\mathcal{A}_\Psi(\mathbf{y})$ is quasi closed in Γ_C if and only if it is quasi closed in $\bar{\Omega}$, see Lemma B.2.5 (b). \square

Accordingly, we can define the decomposition of the active contact set into the sets of weak and strong contact.

Definition 4.4.23 (Weakly and Strongly Active Contact Sets) For a displacement $\mathbf{y} \in \mathbf{K}_\Phi$, the set $\mathcal{S}_\Psi(\mathbf{y})$ in Proposition 4.4.21 is called the set of strongly active contact, and the set

$$\mathcal{W}_\Psi(\mathbf{y}) := \mathcal{A}_\Psi(\mathbf{y}) \setminus \mathcal{S}_\Psi(\mathbf{y})$$

is called the set of weakly active contact.

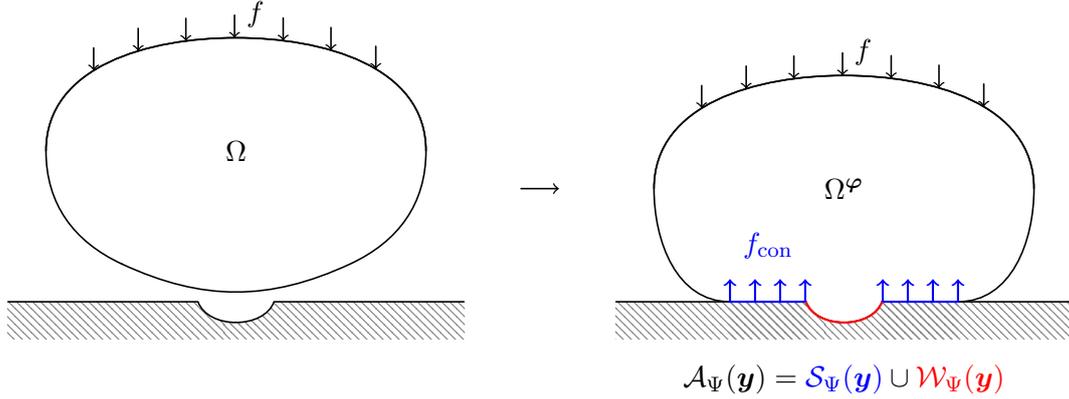


Figure 4.2: Illustration of active contact patches. Reference configuration (left) and deformed configuration for displacement \mathbf{y} with strongly active set on the planar part of the obstacle boundary (right). Weak contact is established if the center section of the obstacle is shaped like the deformed body.

The obvious implication for the boundary conditions of the linearized problem is the following pointwise characterization.

Corollary 4.4.24 Let Assumption 4.0.1 be satisfied, $w \in (\mathbf{H}^1(\Omega)^N)^*$, $\mathbf{y} := S(w) \in \mathbf{K}_\Phi^N$ and $\lambda := w - A_\tau \mathbf{y}$. Then $\delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)$ if and only if

$$\begin{aligned} \delta \mathbf{y}_k &= 0 \text{ q.e. on } \Gamma_D, \\ \delta \mathbf{y}_k \cdot \boldsymbol{\nu}_\Phi &= 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}_k), \\ \delta \mathbf{y}_k \cdot \boldsymbol{\nu}_\Phi &\leq 0 \text{ q.e. on } \mathcal{W}_\Psi(\mathbf{y}_k) \end{aligned}$$

for all $k = 1, \dots, N$.

Proof. This is an immediate consequence of Proposition 4.4.21 and the product structure of $\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)$, see Lemma C.2.2. \square

Hence, the linearized problem

$$\begin{aligned} \delta \mathbf{y} &\in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, w - A_\tau \mathbf{y}) \\ \delta w - A_\tau \delta \mathbf{y} &\in \mathcal{T}_{\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, w - A_\tau \mathbf{y})}(\delta \mathbf{y})^\circ \end{aligned}$$

that arises in the computation of the directional derivative of the solution operator includes standard Dirichlet boundary conditions on the Dirichlet boundary of the forward problem as well as sliding boundary conditions (with respect to the contact normal) on the strongly

active contact set of the corresponding time step's displacement and, finally, a linearized non-penetration condition with vanishing initial gap on the weak contact set of the same displacement. The structure is similar to a detachment problem of the body Ω under the forces δw , but the non-penetration condition is imposed on parts of the contact boundary that generally *varies* over all time steps.

Note that [194, Lem. A.5] contains similar results to those above for the scalar valued obstacle problem with Dirichlet boundary conditions prescribed on the entire boundary using the notion of the fine-support of a $H_0^1(\Omega)^*$ -functional, which is unavailable in our setting with mixed boundary conditions. Similarly, [28, Lem. 3.31, 3.32] contains pointwise interpretations of the tangent and normal cones that correspond to those in Lemma 4.4.17, based on a specific extension operator in a comparatively restricted setting for contact problems. In comparison to these two rather specific approaches, the technique above that is based on the ideal representation is relatively flexible and can be applied to obtain a pointwise representation of a constraint anytime a measure with the required properties used above is at hand.

4.4.3 Gâteaux and Fréchet Differentiability

Due to the previous subsection, we can now characterize points where the Hadamard differentiable solution operator S is Gâteaux differentiable in terms of the biactive set.

Proposition 4.4.25 *Let Assumption 4.0.1 be satisfied, and let $w \in (\mathbf{H}^1(\Omega)^N)^*$, $\mathbf{y} = S(w) \in \mathbf{K}_{\Phi}^N$ and $\mathcal{W}_{\Psi}(\mathbf{y}_k) = \emptyset$ for all $k = 1, \dots, N$. Then $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$ is Gâteaux differentiable at w .*

Proof. Theorem 4.4.12 yields Hadamard differentiability of $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$ at w and Lipschitz continuity of $S'(w, \cdot)$. Additionally, the directional derivative $\delta \mathbf{y} := S'(w, \delta w)$ for δw in $(\mathbf{H}^1(\Omega)^N)^*$ is the unique solution to the problem

$$\begin{aligned} \delta \mathbf{y} &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y}) \\ \delta w - A_{\tau} \delta \mathbf{y} &\in \mathcal{T}_{\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y})}(\delta \mathbf{y})^{\circ}. \end{aligned} \quad (4.19)$$

When the biactive sets vanish, the critical cone $\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y})$ is a closed subspace of $\mathbf{H}^1(\Omega)^N$, see Corollary 4.4.24 and Proposition B.3.5, and therefore

$$\mathcal{T}_{\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y})}(\delta \mathbf{y})^{\circ} = \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y})^{\perp},$$

which is a linear subspace of $(\mathbf{H}^1(\Omega)^N)^*$. In order to confirm linearity of $S'(w, \cdot)$, note that

$$\begin{aligned} S(w, \delta w_1) + \alpha S(w, \delta w_2) &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y}), \\ \delta w_1 + \alpha \delta w_2 - A_{\tau}(S(w, \delta w_1) + \alpha S(w, \delta w_2)) \\ &= \delta w_1 - A_{\tau} S(w, \delta w_1) + \alpha(\delta w_2 - A_{\tau} S(w, \delta w_2)) \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y})^{\perp} \end{aligned}$$

for all $\delta w_1, \delta w_2$ in $(\mathbf{H}^1(\Omega)^N)^*$ and for all α in \mathbb{R} , due to linearity of the set $\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau} \mathbf{y})$. The linearized problem (4.19) is uniquely solvable, therefore, the directional derivatives

$S'(w, \cdot)$ are linear and (Lipschitz) continuous, which implies boundedness and therefore Gâteaux differentiability. \square

In Section 5.5, we will see that Gâteaux differentiability of the solution operator can be used to infer Fréchet differentiability of the control-to-state operator under additional assumptions. As a converse result to the Gâteaux differentiability in the absence of a biactive set, we obtain the following:

Proposition 4.4.26 *Let Assumption 4.0.1 be satisfied, and let $w \in (\mathbf{H}^1(\Omega)^N)^*$ with associated state $\mathbf{y} = S(w) \in \mathbf{K}_{\Phi}^N$. If $\mathcal{S}_{\Psi}(\mathbf{y}_k)$ is quasi closed and $\mathcal{W}_{\Psi}(\mathbf{y}_k) \neq \emptyset$ for a k in $\{1, \dots, N\}$, then $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$ is not Gâteaux differentiable at w .*

Proof. Let $\widetilde{\Gamma}_D \subset \overline{\Gamma}_D \subset \partial\Omega \setminus \overline{\Gamma}_C$ denote the quasi closed superset of Γ_D in Corollary 3.4.10 (the “quasi closure” of Γ_D). Then $\widetilde{\Gamma}_D \cup \mathcal{S}_{\Psi}(\mathbf{y}_k)$ is quasi closed in $\partial\Omega$, and due to Lemma B.2.5, it is quasi closed in \mathbb{R}^d . [111, Thm. 1.5] therefore yields a non-positive function $v \in H^1(\mathbb{R}^d)$ such that $v^{-1}(\{0\}) = \widetilde{\Gamma}_D \cup \mathcal{S}_{\Psi}(\mathbf{y}_k)$ up to $\text{cap}_{\Omega,2}$ -polar sets.

Using the extended contact normal in (4.13), we define $\mathbf{v} := v|_{\overline{\Omega}} \tilde{\nu}_{\Phi}$, which satisfies

$$\begin{aligned} \mathbf{v} &= 0 \text{ q.e. on } \Gamma_D, \\ \mathbf{v} \cdot \nu_{\Phi} &= v = 0 \text{ q.e. on } \mathcal{S}_{\Psi}(\mathbf{y}_k), \\ \mathbf{v} \cdot \nu_{\Phi} &= v < 0 \text{ q.e. on } \mathcal{W}_{\Psi}(\mathbf{y}_k). \end{aligned}$$

We can now define $\tilde{\mathbf{v}} := (0, \dots, 0, \mathbf{v}, 0, \dots, 0) \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y}) \subset \mathbf{H}^1(\Omega)^N$, where the non-zero component is at index k . Then $S'(w, A_{\tau}\tilde{\mathbf{v}}) = \tilde{\mathbf{v}}$, but $S'(w, -A_{\tau}\tilde{\mathbf{v}}) \neq -\tilde{\mathbf{v}}$ since $\tilde{\mathbf{v}} \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y})$, but $-\tilde{\mathbf{v}} \notin \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y})$ because $\mathcal{W}_{\Psi}(\mathbf{y}_k) \neq \emptyset$. Accordingly, the mapping $S'(w, \cdot)$ is clearly nonlinear. \square

Remark 4.4.27 *Recall that Corollary 4.4.24 ensures quasi closedness of $\mathcal{S}_{\Psi}(\mathbf{y}_k)$ when Γ_C is quasi closed.*

Since Gâteaux differentiability of the solution operator fails at all right hand sides $w \in H_D^1(\Omega)^*$ that lead to a state with non-vanishing weak contact sets, Fréchet differentiability obviously fails as well. However, Fréchet differentiability can be obtained on a dense subset of the right hand sides.

Proposition 4.4.28 *Let Assumption 4.0.1 be satisfied. Then $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$ is Fréchet differentiable on a dense subset of $(\mathbf{H}^1(\Omega)^N)^*$.*

Proof. Note that the space $\mathbf{H}^1(\Omega)^N$ with its associated scalar product is a Hilbert space and therefore both $\mathbf{H}^1(\Omega)^N$ itself and $(\mathbf{H}^1(\Omega)^N)^{**}$ are reflexive and possess the Radon-Nikodym property [53, P. 5]. Hence, [205, Thm. 6] implies that $(\mathbf{H}^1(\Omega)^N)^*$ is an Asplund space, and since $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$ is Lipschitz continuous, [157, Thm. 2.5] yields the Fréchet differentiability on a dense subset. \square

Chapter 5

Optimal Control of the Time-Discretized Contact Problem

Optimal control of contact problems falls into the complex class of complementarity constrained optimization in function space. Complementarity constrained problems are generally nonsmooth, and their analysis requires special concepts and techniques, especially concerning optimality conditions. For overview of stationarity concepts in the finite-dimensional setting, see [129, 165], and see [193] for work in Banach spaces. This chapter is dedicated to the development of optimal control theory for the time-discretized contact problem. Based on the forward analysis of Chapter 4, we will address existence of solutions in a Banach space \mathbf{U} and first order optimality conditions for the time-discretized version of the optimization problem

$$\begin{aligned} \min J(\mathbf{y}, \mathbf{u}) \\ \text{s.t. } (\mathbf{y}, \mathbf{u}) \in \mathbf{Y} \times \mathbf{U} \\ \mathbf{y} \in \mathbf{K}_{\Phi}^T, \quad \mathbf{y}(0) = \mathbf{y}_{\text{ini}}, \quad \dot{\mathbf{y}}(0) = \mathbf{v}_{\text{ini}} \\ \langle \ddot{\mathbf{y}} - f - B\mathbf{u}, \mathbf{v} - \mathbf{y} \rangle_{L^2(0,T;\mathbf{H}^1(\Omega))} + a_I(\mathbf{y}, \mathbf{v} - \mathbf{y}) + b_I(\dot{\mathbf{y}}, \mathbf{v} - \mathbf{y}) \geq 0 \quad \forall \mathbf{v} \in \mathbf{K}_{\Phi}^T \end{aligned} \tag{5.1}$$

for a bounded, linear operator $B: \mathbf{U} \rightarrow L^2(0,T;\mathbf{H}^1(\Omega))^*$ and an objective functional $J: \mathbf{Y} \times \mathbf{U} \rightarrow \mathbb{R}$, cf. the time-continuous contact problem (4.1). Using the differentiability properties of the solution operator to the time-discretized forward problem, we will elaborate on the differentiability properties of the corresponding state reduced objective functional. Additionally, we will introduce adjoint-based formulations for the functional's derivatives that are used to compute the search directions in the optimization algorithms presented in Chapter 6.

Due to the insufficient analytical results for the hyperbolic variational inequality (uniqueness of solutions has not been established yet), results concerning the optimal control of these problems are consequently non-existent. Fundamentals for general complementarity constrained problems in Banach spaces and the constraint qualifications and optimality conditions tailored to their specific structure, however, are developed in [194, 196] and notably [193], whose techniques are key in the following analysis. Additionally, optimal control problems governed by elliptic variational inequalities have previously been addressed in,

e.g., [19, 96, 138, 140], and optimization results on parabolic variational inequalities and complementarity constrained control problems are examined in [18, 101, 139, 188, 192]. Mignot’s work [138, Sec. 6] treats an optimization problem with a simplified, static, scalar “contact problem”, i.e., an obstacle problem with an obstacle on the boundary of the domain. In [28], Betz obtains first order optimality conditions for the static contact problem and $C^{1,1}$ -boundaries, where the geometric normal coincides with the contact normal.

The results of this chapter have essentially been published in [143]. Sections 5.4 and 5.5 include refined results on the existence of the adjoint states and on the differentiability of the reduced objective functional.

Structure. We discretize the control and the objective functional to obtain the time-discretized optimization problem that corresponds to (5.1) in Section 5.1, and existence of minimizers to the discretized problem is addressed in Section 5.2. In Section 5.3, we establish first order optimality conditions of strong-stationary-type for dense controls. We analyze the backwards-in-time time stepping structure and the boundary conditions of the corresponding adjoint problem in Section 5.4, and in Section 5.5, we prove directional differentiability of the objective functional as well as Fréchet differentiability whenever the biactive set vanishes, in which case the adjoint problem admits a unique solution that can be used as a representation of the derivative.

5.1 Time-Discretized Controls

Applying the time discretization that was described in Section 4.1 to the time-continuous optimal control problem in (5.1) results in a natural correspondence between the controls in the continuous and the discretized setting. Recall that the discretization of the test- and ansatz functions in (4.4) yields the mappings

$$\begin{aligned} i_0: \mathbf{H}^1(\Omega)^N &\rightarrow L^2(0, T; \mathbf{H}^1(\Omega)), & i_1: \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega)^N &\rightarrow L^2(0, T; \mathbf{H}^1(\Omega)), \\ i_0(\mathbf{p})(t) &:= \sum_{k=1}^N \mathbf{p}_k \chi_{I_k}(t), & i_1(\mathbf{y}_0, \mathbf{y})(t) &:= \sum_{k=0}^{N-1} \left(\mathbf{y}_k + \frac{t-t_k}{\tau} \mathbf{y}_{k+1} \right) \chi_{I_{k+1}}(t), \end{aligned}$$

which map the degrees of freedom in the time-discretized setting to their time-continuous counterparts. Both i_0 and i_1 are clearly linear and bounded — with constants depending on τ and N — therefore

$$\begin{aligned} B_\tau: \mathbf{U} &\rightarrow (\mathbf{H}^1(\Omega)^N)^* & J_\tau: \mathbf{H}^1(\Omega)^N \times \mathbf{U} &\rightarrow \mathbb{R} \\ B_\tau &:= i_0^* B & J_\tau(\mathbf{y}, \mathbf{u}) &:= J(i_1(\mathbf{y}_{\text{ini}}, \mathbf{y}), i_0(\mathbf{u})), \end{aligned}$$

immediately yield a bounded, linear *control-to-force* operator and an objective functional for the discretized system. In practice, however, the control will generally be time dependent as well and may be discretized for algorithmic purposes. As this is of course dependent on the application, this chapter is based on the following abstract assumption that extends the standing assumptions made in the previous chapter and are assumed to hold throughout this one.

Assumption 5.1.1

- (a) Assumption 4.0.1 holds.
- (b) \mathbf{U}_τ is a reflexive Banach space.
- (c) $B_\tau: \mathbf{U}_\tau \rightarrow (\mathbf{H}^1(\Omega)^N)^*$ is linear and bounded.
- (d) $J_\tau: \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau \rightarrow \mathbb{R}$ is Fréchet differentiable.

Employing the notation established in the previous chapter, the time-discretized optimal control problem can then be stated as follows:

Problem 5.1.2 (Time-Discretized Optimal Control Problem) *Given $\mathbf{y}_{ini} \in \mathbf{K}_\Phi$, $\mathbf{v}_{ini} \in \mathbf{H}^1(\Omega)$ and $f \in (\mathbf{H}^1(\Omega)^N)^*$ find $\mathbf{u} \in \mathbf{U}_\tau$ and $\mathbf{y} \in \mathbf{H}^1(\Omega)^N$ that solve*

$$\begin{aligned} \min J_\tau(\mathbf{y}, \mathbf{u}) \\ \text{s.t. } (\mathbf{y}, \mathbf{u}) \in \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau \\ \mathbf{y} \in \mathbf{K}_\Phi^N \\ B_\tau \mathbf{u} + f_\tau - A_\tau \mathbf{y} \in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ. \end{aligned}$$

For future reference, some notation for the reduced quantities is introduced.

Definition 5.1.3 (Reduced Objective Functional) *The mapping*

$$\begin{aligned} \tilde{S}: \mathbf{U}_\tau &\rightarrow \mathbf{K}_\Phi^N \subset \mathbf{H}_D^1(\Omega)^N \\ \tilde{S}(\mathbf{u}) &:= S(B_\tau \mathbf{u} + f_\tau), \end{aligned}$$

where S denotes the solution operator from Theorem 4.3.3, is called the control-to-state operator for the contact problem in Problem 5.1.2, and

$$\begin{aligned} \tilde{J}_\tau: \mathbf{U}_\tau &\rightarrow \mathbb{R} \\ \tilde{J}_\tau(\mathbf{u}) &:= J(\tilde{S}(\mathbf{u}), \mathbf{u}) \end{aligned}$$

is called the reduced objective functional.

5.2 Existence of Minimizers

Under mild assumptions, we can establish existence of minimizers by the standard technique used in optimization problems in reflexive Banach spaces.

Theorem 5.2.1 (Existence of Minimizers) *Let Assumption 5.1.1 be satisfied and let $B_\tau \in \mathcal{L}(\mathbf{U}_\tau, (\mathbf{H}^1(\Omega)^N)^*)$ be compact, the objective functional $J_\tau: \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau \rightarrow \mathbb{R}$ be lower semi-continuous for strongly/weakly convergent sequences of tuples, and let*

$$\lim_{\|(\mathbf{y}, \mathbf{u})\|_{\mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau} \rightarrow \infty} J_\tau(\mathbf{y}, \mathbf{u}) = \infty. \quad (5.2)$$

Then Problem 5.1.2, i.e.,

$$\begin{aligned} \min J_\tau(\mathbf{y}, \mathbf{u}) \\ \text{s.t. } (\mathbf{y}, \mathbf{u}) \in \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau \\ \mathbf{y} \in \mathbf{K}_\Phi^N \\ B_\tau \mathbf{u} + f_\tau - A_\tau \mathbf{y} \in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ, \end{aligned}$$

has at least one solution $(\bar{\mathbf{y}}, \bar{\mathbf{u}})$.

Proof. Let $(\mathbf{u}_k) \subset \mathbf{U}_\tau$ be a minimizing sequence such that

$$J_\tau(S(B_\tau \mathbf{u}_k + f_\tau), \mathbf{u}_k) = \tilde{J}_\tau(\mathbf{u}_k) \rightarrow \inf_{\mathbf{u} \in \mathbf{U}_\tau} \tilde{J}_\tau(\mathbf{u}) = \inf_{\mathbf{u} \in \mathbf{U}_\tau} J_\tau(S(B_\tau(\mathbf{u}) + f_\tau), \mathbf{u}).$$

The coercivity of J_τ in (5.2) implies that the sequence (\mathbf{u}_k) is necessarily bounded in \mathbf{U}_τ . Due to reflexivity of \mathbf{U}_τ , we obtain a weakly convergent subsequence (denoted by the same symbol) and $\mathbf{u} \in \mathbf{U}_\tau$ such that $\mathbf{u}_k \rightharpoonup \bar{\mathbf{u}}$.

Compactness of B_τ implies that we can pass to another subsequence such that

$$B_\tau \mathbf{u}_k \rightarrow B_\tau \bar{\mathbf{u}} \quad \text{in } (\mathbf{H}^1(\Omega)^N)^*,$$

therefore continuity of $S: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^N$ yields that

$$\mathbf{y}_k := S(B_\tau \mathbf{u}_k + f_\tau) \rightarrow S(B_\tau \bar{\mathbf{u}} + f_\tau) = \bar{\mathbf{y}} \quad \text{in } \mathbf{H}^1(\Omega)^N.$$

Finally, due to the lower semi-continuity of J_τ , we obtain

$$\inf_{\mathbf{u} \in \mathbf{U}_\tau} \tilde{J}_\tau(\mathbf{u}) = \liminf_{k \rightarrow \infty} J_\tau(\mathbf{y}_k, \mathbf{u}_k) \geq \tilde{J}_\tau(\bar{\mathbf{u}}) \geq \inf_{\mathbf{u} \in \mathbf{U}_\tau} \tilde{J}_\tau(\mathbf{u}). \quad \square$$

Remark 5.2.2 Strongly/weakly convergent tuples are tuples $(\mathbf{y}_k, \mathbf{u}_k)$ where \mathbf{y}_k is strongly convergent in $\mathbf{H}^1(\Omega)^N$ and \mathbf{u}_k is weakly convergent in \mathbf{U}_τ .

Remark 5.2.3 The existence of a nonempty, bounded level set of the problem is a slightly more general assumption than the coercivity condition above but equally sufficient for the existence of minimizers.

5.3 Stationarity Conditions

In this section, we will develop first order optimality conditions of strong stationarity type for the time-discretized optimal control problem. The proof of the stationarity result is a straight forward extension of the argument in [193, Sec. 5.1] — which is based on a linearization of the optimal control problem and a density argument — to the discretized contact problem and its time stepping structure. The essential properties for the extension follow from the results of Chapter 4, where we derived the specific form of the directional derivative of the solution operator to the state problem.

Note that we can naturally identify (adjoint) states $p \in (\mathbf{H}^1(\Omega)^N)^{**}$ with an element of the primal space $\tilde{p} \in \mathbf{H}^1(\Omega)^N$ due to reflexivity of $\mathbf{H}^1(\Omega)^N$. There will be no further distinctions made between the two.

Theorem 5.3.1 (First Order Optimality Conditions) *Let Assumption 5.1.1 be satisfied, $\overline{B_\tau \mathbf{U}_\tau} = (\mathbf{H}^1(\Omega)^N)^*$ and $\bar{\mathbf{x}} := (\bar{\mathbf{y}}, \bar{\mathbf{u}})$ be a local minimizer to Problem 5.1.2, i.e.,*

$$\begin{aligned} & \min J_\tau(\mathbf{y}, \mathbf{u}) \\ & \text{s.t. } (\mathbf{y}, \mathbf{u}) \in \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau \\ & \quad \mathbf{y} \in \mathbf{K}_\Phi^N \\ & \quad B_\tau \mathbf{u} + f_\tau - A_\tau \mathbf{y} \in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ. \end{aligned} \quad (5.3)$$

Then there exist multipliers $\mathbf{p} \in \mathbf{H}^1(\Omega)^N$, $\mu \in (\mathbf{H}^1(\Omega)^N)^*$ with

$$\partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}}) + \mu - A_\tau^* \mathbf{p} = 0, \quad \mathbf{p} \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, B_\tau \bar{\mathbf{u}} + f_\tau - A_\tau \bar{\mathbf{y}}), \quad (5.4a)$$

$$\partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}) + B_\tau^* \mathbf{p} = 0, \quad \mu \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, B_\tau \bar{\mathbf{u}} + f_\tau - A_\tau \bar{\mathbf{y}})^\circ. \quad (5.4b)$$

Proof. Theorem 4.4.12 supplies the Hadamard differentiable solution operator with Lipschitz continuous directional derivative S' to the contact problem. Linearity of B_τ and local optimality of $(\bar{\mathbf{y}}, \bar{\mathbf{u}}) = (S(B_\tau \bar{\mathbf{u}} + f_\tau), \bar{\mathbf{u}})$ therefore imply that

$$\langle \partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}}), S'(B_\tau \bar{\mathbf{u}} + f_\tau, B_\tau \delta \mathbf{u}) \rangle_{\mathbf{H}^1(\Omega)^N} + \langle \partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}), \delta \mathbf{u} \rangle_{\mathbf{U}_\tau} \geq 0 \quad \forall \delta \mathbf{u} \in \mathbf{U}_\tau. \quad (5.5)$$

Testing the previous line with $\pm \delta \mathbf{u}$ as proposed in [138, 193], we obtain the existence of a constant $C > 0$ such that

$$\begin{aligned} -\langle \partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}), \delta \mathbf{u} \rangle_{\mathbf{U}_\tau} & \leq \langle \partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}}), S'(B_\tau \bar{\mathbf{u}} + f_\tau, B_\tau \delta \mathbf{u}) \rangle_{\mathbf{H}^1(\Omega)^N} \\ & \leq \|\partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}})\|_{(\mathbf{H}^1(\Omega)^N)^*} \|S'(B_\tau \bar{\mathbf{u}} + f_\tau, B_\tau \delta \mathbf{u})\|_{\mathbf{H}^1(\Omega)^N} \\ & \leq \|\partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}})\|_{(\mathbf{H}^1(\Omega)^N)^*} L_{S'} \|B_\tau \delta \mathbf{u}\|_{(\mathbf{H}^1(\Omega)^N)^*} = C \|B_\tau \delta \mathbf{u}\|_{(\mathbf{H}^1(\Omega)^N)^*}, \end{aligned}$$

and analogously

$$\begin{aligned} \langle \partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}), \delta \mathbf{u} \rangle_{\mathbf{U}_\tau} & \leq \langle \partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}}), S'(B_\tau \bar{\mathbf{u}} + f_\tau, -B_\tau \delta \mathbf{u}) \rangle_{\mathbf{H}^1(\Omega)^N} \\ & \leq \|\partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}})\|_{(\mathbf{H}^1(\Omega)^N)^*} \|S'(B_\tau \bar{\mathbf{u}} + f_\tau, -B_\tau \delta \mathbf{u})\|_{\mathbf{H}^1(\Omega)^N} \\ & \leq \|\partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}})\|_{(\mathbf{H}^1(\Omega)^N)^*} L_{S'} \|-B_\tau \delta \mathbf{u}\|_{(\mathbf{H}^1(\Omega)^N)^*} = C \|B_\tau \delta \mathbf{u}\|_{(\mathbf{H}^1(\Omega)^N)^*}, \end{aligned}$$

where $L_{S'}$ denotes the Lipschitz constant of $S'(B_\tau \bar{\mathbf{u}} + f_\tau, \cdot)$. Accordingly,

$$|\langle \partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}), \delta \mathbf{u} \rangle_{\mathbf{U}_\tau}| \leq C \|B_\tau \delta \mathbf{u}\|_{(\mathbf{H}^1(\Omega)^N)^*} \quad \text{for all } \delta \mathbf{u} \in \mathbf{U}_\tau.$$

Therefore, the kernel of B_τ is a subset of the kernel of $\partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}})$ and the mapping $B_\tau \delta \mathbf{u} \mapsto \partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}) \delta \mathbf{u}$ is well defined, linear and bounded on the image of B_τ . Due to the density of $\text{im}(B_\tau)$, we can extend the mapping to a functional $\mathbf{p} \in (\mathbf{H}^1(\Omega)^N)^{**} \cong \mathbf{H}^1(\Omega)^N$, see [119, Thm. 2.7-11]. Then

$$\partial_{\mathbf{u}} J_\tau(\bar{\mathbf{x}}) = B_\tau^* \mathbf{p}, \quad (5.6)$$

and the density of $\text{im}(B_\tau)$ in $(\mathbf{H}^1(\Omega)^N)^*$ in combination with (5.5) yields that

$$\langle \partial_{\mathbf{y}} J_\tau(\bar{\mathbf{x}}), S'(B_\tau \bar{\mathbf{u}} + f_\tau, \delta w) \rangle_{\mathbf{H}^1(\Omega)^N} + \langle \mathbf{p}, \delta w \rangle_{(\mathbf{H}^1(\Omega)^N)^*} \geq 0 \quad \text{for all } \delta w \in (\mathbf{H}^1(\Omega)^N)^*,$$

implying that $(\delta \mathbf{y}, \delta w) = (0, 0)$ is a global minimizer to the problem

$$\begin{aligned} & \min \langle \partial_{\mathbf{y}} J_{\tau}(\bar{\mathbf{x}}), \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} + \langle \mathbf{p}, \delta w \rangle_{(\mathbf{H}^1(\Omega)^N)^*} \\ & \text{s.t. } (\delta \mathbf{y}, \delta w) \in \mathbf{H}^1(\Omega)^N \times (\mathbf{H}^1(\Omega)^N)^* \\ & \quad \delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, B_{\tau} \mathbf{u} + f_{\tau} - A_{\tau} \mathbf{y}) \\ & \quad \delta w - A_{\tau} \delta \mathbf{y} \in \mathcal{T}_{\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, B_{\tau} \mathbf{u} + f_{\tau} - A_{\tau} \mathbf{y})}(\delta \mathbf{y})^{\circ}, \end{aligned}$$

because its functional value is zero and because of the form of S' established in Theorem 4.4.12. We can split the last condition of the problem up since $\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, B_{\tau} \mathbf{u} + f_{\tau} - A_{\tau} \mathbf{y})$ is a cone, see [31, Ex. 2.62], and the linearized problem is therefore equivalent to the complementarity constrained problem

$$\begin{aligned} & \min \langle \partial_{\mathbf{y}} J_{\tau}(\bar{\mathbf{x}}), \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} + \langle \mathbf{p}, \delta w \rangle_{(\mathbf{H}^1(\Omega)^N)^*} \\ & \text{s.t. } (\delta \mathbf{y}, \delta w) \in \mathbf{H}^1(\Omega)^N \times (\mathbf{H}^1(\Omega)^N)^* \\ & \quad \delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\bar{\mathbf{y}}, B_{\tau} \bar{\mathbf{u}} + f_{\tau} - A_{\tau} \bar{\mathbf{y}}) \\ & \quad \delta w - A_{\tau} \delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\bar{\mathbf{y}}, B_{\tau} \bar{\mathbf{u}} + f_{\tau} - A_{\tau} \bar{\mathbf{y}})^{\circ} \\ & \quad \langle \delta w - A_{\tau} \delta \mathbf{y}, \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} = 0. \end{aligned}$$

The mapping

$$\mathbf{H}^1(\Omega)^N \times (\mathbf{H}^1(\Omega)^N)^* \ni (\delta \mathbf{y}, \delta w) \mapsto (\delta \mathbf{y}, \delta w - A_{\tau} \delta \mathbf{y}) \in \mathbf{H}^1(\Omega)^N \times (\mathbf{H}^1(\Omega)^N)^*$$

is linear and surjective, hence the constraint qualification for [193, Prop. 4.8] is satisfied and the strong stationarity conditions for the auxiliary problem yield multipliers μ, ζ such that

$$\begin{aligned} \partial_{\mathbf{y}} J_{\tau}(\bar{\mathbf{x}}) + \mu - A_{\tau}^* \zeta &= 0, & \mu &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\bar{\mathbf{y}}, B_{\tau} \bar{\mathbf{u}} + f_{\tau} - A_{\tau} \bar{\mathbf{y}})^{\circ}, \\ \mathbf{p} + \zeta &= 0, & \zeta &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\bar{\mathbf{y}}, B_{\tau} \bar{\mathbf{u}} + f_{\tau} - A_{\tau} \bar{\mathbf{y}}), \end{aligned}$$

cf. [193, Def. 4.1]. The claim follows from the identity (5.6), eliminating ζ and reversing the sign of \mathbf{p} . \square

Remark 5.3.2 *The sign of the adjoint state \mathbf{p} is reversed for a more compact notation. We initially obtain the negative of what we refer to as the adjoint state because of the multiplier's role in the system of strong stationarity.*

The density of the image of B_{τ} enters the linearization in the previous proof twice. When density is not guaranteed, the auxiliary problem can in turn only be formulated in $\mathbf{H}^1(\Omega)^N \times \text{im}(B_{\tau})$, and the constraint qualification may fail in general. Therefore we can not evaluate strong stationarity conditions for the minimizer $(0, 0)$ to the auxiliary problem in that case.

Combined with the forward problem (5.3), the adjoint problem (5.4a) and the stationarity condition (5.4b) form the first order optimality system. When (5.4a) is referred to the adjoint problem, this is meant to include the constraint on the multiplier μ from now on.

5.4 Structure of the Adjoint Problem

Similarly to the forward problem, the adjoint problem in the optimality conditions to the optimal control problem in Theorem 5.3.1 exhibits a time stepping structure. The time stepping is reversed in time, as we will see in this section.

Problem 5.4.1 *Given $\mathbf{u} \in \mathbf{U}_\tau$, $\mathbf{y} = S(B_\tau \mathbf{u} + f_\tau)$ and the residual $\lambda = B_\tau \mathbf{u} + f_\tau - A_\tau \mathbf{y}$, find $\mathbf{p} \in \mathbf{H}^1(\Omega)^N$ and $\mu \in (\mathbf{H}^1(\Omega)^N)^*$ that solve*

$$\begin{aligned} \mathbf{p} &\in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda) \\ \mu &\in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^\circ \quad \text{or equivalently} \quad \mathbf{p} \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda) \\ \partial_\varphi J_\tau(\mathbf{y}, \mathbf{u}) + \mu - A_\tau^* \mathbf{p} &= 0 \quad A_\tau^* \mathbf{p} - \partial_\varphi J_\tau(\mathbf{y}, \mathbf{u}) \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^\circ. \end{aligned}$$

Based on the pointwise characterization of the critical cone in Section 4.4.2, we obtain boundary conditions for the adjoint problem and address the existence of solutions to the adjoint problem. Reflexivity of $\mathbf{H}^1(\Omega)^N$ and the identification of $\mathbf{p} \in (\mathbf{H}^1(\Omega)^N)^{**}$ with its primal representative are repeatedly used in this section.

5.4.1 Adjoint Time Stepping

Taking a closer look at the form of the adjoint operator A_τ^* reveals the backwards-in-time time stepping structure. In order to obtain a variational form of $A_\tau^* \mathbf{p} - J$ for every $J \in (\mathbf{H}^1(\Omega)^N)^*$, we test with φ in $\mathbf{H}^1(\Omega)^N$, which yields

$$\begin{aligned} \langle A_\tau \varphi, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} - \langle J, \varphi \rangle_{\mathbf{H}^1(\Omega)^N} & \tag{5.7} \\ &= (\varphi_1, \mathbf{p}_1)_{L^2(\Omega)} + \frac{\tau^2}{4} a(\varphi_1, \mathbf{p}_1) + \frac{\tau}{2} b(\varphi_1, \mathbf{p}_1) \\ & \quad \sum_{k=1}^{N-1} \left[(\varphi_{k+1} - \varphi_k - \tau \bar{v}_k(\varphi_1, \dots, \varphi_k), \mathbf{p}_{k+1})_{L^2(\Omega)} \right. \\ & \quad \quad \left. + \frac{\tau^2}{4} (a(\varphi_{k+1}, \mathbf{p}_{k+1}) + a(\varphi_k, \mathbf{p}_{k+1})) \right. \\ & \quad \quad \left. + \frac{\tau}{2} (b(\varphi_{k+1}, \mathbf{p}_{k+1}) - b(\varphi_k, \mathbf{p}_{k+1})) \right] - \langle J, \varphi \rangle_{\mathbf{H}^1(\Omega)^N}. \end{aligned}$$

Confer the forward problem (4.9), where variation was with respect to \mathbf{p} . This expression decouples with respect to the components of the test functions, and we will introduce some additional notation for the right hand sides. To that end, note that the form of the “velocities” \bar{v}_k , see Definition 4.2.2 (a), induces terms that can be interpreted as *adjoint velocities*.

Lemma 5.4.2 *The mappings*

$$\begin{aligned} \bar{q}_k &: \mathbf{H}^1(\Omega)^{N-k+1} \rightarrow \mathbf{H}^1(\Omega) \\ \bar{q}_k(\mathbf{p}_k, \dots, \mathbf{p}_N) &:= \frac{2}{\tau} \left((-1)^{N+k+1} \mathbf{p}_N + \sum_{j=k}^{N-1} (-1)^{j+k} (\mathbf{p}_{j+1} - \mathbf{p}_j) \right) \end{aligned} \quad (5.8)$$

are linear and bounded for all k in $\{1, \dots, N\}$. The form in (5.8) holds if and only if

$$\begin{aligned} \bar{q}_N(\mathbf{p}_N) &= -\frac{2}{\tau} \mathbf{p}_N, \\ \mathbf{p}_{k+1} - \mathbf{p}_k &= \frac{\tau}{2} (\bar{q}_{k+1}(\mathbf{p}_{k+1}, \dots, \mathbf{p}_N) + \bar{q}_k(\mathbf{p}_k, \dots, \mathbf{p}_N)) \end{aligned} \quad (5.9)$$

for all k in $\{1, \dots, N-1\}$ and $\mathbf{p} \in \mathbf{H}^1(\Omega)^N$. Additionally,

$$\sum_{k=1}^{N-1} (\bar{v}_k(\varphi_1, \dots, \varphi_k), \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} = - \sum_{k=1}^{N-1} (\bar{q}_{k+1}(\mathbf{p}_{k+1}, \dots, \mathbf{p}_N), \varphi_k)_{\mathbf{L}^2(\Omega)}. \quad (5.10)$$

Proof. Linearity and boundedness are clear and the equivalence of (5.8) and (5.9) is a straight forward induction argument, cf. the corresponding result for the primal velocities in Lemma 4.2.1. The identity (5.10) can be confirmed using the symmetry of the scalar product and by reversing the roles of the test and ansatz functions. To that end, note that φ_N has no part in the equation

$$\begin{aligned} &\sum_{k=1}^{N-1} (\bar{v}_k(\varphi_1, \dots, \varphi_k), \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} \\ &= \frac{2}{\tau} \sum_{k=1}^{N-1} \left[(-1)^{k+1} (\varphi_1, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} + \sum_{j=2}^k (-1)^{k+j} (\varphi_j - \varphi_{j-1}, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} \right]. \end{aligned}$$

Now fix an m in $\{1, \dots, N-1\}$ and note that φ_m appears in the summands of the outer sum as φ_j or φ_1 , respectively, whenever $k \geq m$ and as φ_{j-1} whenever $k \geq m+1$. Therefore, we have that

$$\begin{aligned} &\sum_{k=1}^{N-1} (\bar{v}_k(\varphi_1, \dots, \varphi_k), \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} \\ &= \frac{2}{\tau} \sum_{m=1}^{N-1} \left[\sum_{k=m}^{N-1} (-1)^{k+m} (\varphi_m, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} + \sum_{k=m+1}^{N-1} (-1)^{k+m} (\varphi_m, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} \right] \\ &= \frac{2}{\tau} \sum_{m=1}^{N-1} \left[\sum_{k=m+1}^N (-1)^{k+m+1} (\varphi_m, \mathbf{p}_k)_{\mathbf{L}^2(\Omega)} + \sum_{k=m+1}^{N-1} (-1)^{k+m} (\varphi_m, \mathbf{p}_{k+1})_{\mathbf{L}^2(\Omega)} \right] \\ &= \frac{2}{\tau} \sum_{m=1}^{N-1} \left[(-1)^{N+m+1} (\mathbf{p}_N, \varphi_m)_{\mathbf{L}^2(\Omega)} + \sum_{k=m+1}^{N-1} (-1)^{k+m+1} (\mathbf{p}_k - \mathbf{p}_{k+1}, \varphi_m)_{\mathbf{L}^2(\Omega)} \right] \\ &= - \sum_{m=1}^{N-1} (\bar{q}_{m+1}(\mathbf{p}_{m+1}, \dots, \mathbf{p}_N), \varphi_m)_{\mathbf{L}^2(\Omega)}. \quad \square \end{aligned}$$

Note that equation (5.9) corresponds to a Crank-Nicolson discretization of $\dot{\mathbf{p}} = \mathbf{q}$, which motivates the interpretation as adjoint velocities.

The adjoint velocities and the remaining right hand terms are now collected for a compact representation of the right hand sides in the time stepping scheme of the adjoint equation. We obtain combined right hand sides r_k in the adjoint problem that correspond to the right hand sides l_k in the forward problem.

Lemma 5.4.3 *The operators*

$$r_k : \mathbf{H}^1(\Omega)^{N-k} \times (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^*$$

defined by

$$\begin{aligned} r_N(J) &:= \langle J_N, \cdot \rangle_{\mathbf{H}^1(\Omega)}, \\ r_k(\mathbf{p}, J) &:= (\mathbf{p}_1, \cdot)_{\mathbf{L}^2(\Omega)} - \frac{\tau^2}{4} a(\mathbf{p}_1, \cdot) + \frac{\tau}{2} b(\mathbf{p}_1, \cdot) \\ &\quad - \tau (\bar{q}_{k+1}(\mathbf{p}), \cdot)_{\mathbf{L}^2(\Omega)} + \langle J_k, \cdot \rangle_{\mathbf{H}^1(\Omega)} \end{aligned}$$

for k in $\{1, \dots, N-1\}$, respectively, are linear and bounded.

Proof. For r_N , the claim is certainly true. For r_k with k in $\{1, \dots, N-1\}$, linearity is due to the linearity of \bar{q}_{k+1} , see Lemma 5.4.2. The boundedness of the bilinear forms a and b (Assumption 4.0.1 (g)) additionally yields the existence of positive constants $C_a, C_b > 0$ such that the estimate

$$\begin{aligned} \left| \langle r_k(\mathbf{p}, J), \varphi \rangle_{\mathbf{H}^1(\Omega)} \right| &\leq \|\mathbf{p}_1\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} + \frac{\tau^2}{4} C_a \|\mathbf{p}_1\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} \\ &\quad + \frac{\tau}{2} C_b \|\mathbf{p}_1\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} + \tau \|\bar{q}_{k+1}(\mathbf{p})\|_{\mathbf{H}^1(\Omega)} \|\varphi\|_{\mathbf{H}^1(\Omega)} \\ &\quad + \|J_k\|_{\mathbf{H}^1(\Omega)^*} \|\varphi\|_{\mathbf{H}^1(\Omega)} \end{aligned}$$

is satisfied for all $\varphi \in \mathbf{H}^1(\Omega)$. This yields well-definedness and boundedness of the operator due to the boundedness of $\bar{q}_{k+1} : \mathbf{H}^1(\Omega)^k \rightarrow \mathbf{H}^1(\Omega)$ — again, see Lemma 5.4.2. \square

Using the symmetry of a , b and the $\mathbf{L}^2(\Omega)$ -scalar product and the representation of the adjoint right hand sides in Lemma 5.4.3, we can rewrite the representation of the adjoint

problem's variational form in (5.7) and obtain the equality

$$\begin{aligned}
& \langle A_\tau \boldsymbol{\varphi}, \boldsymbol{p} \rangle_{\mathbf{H}^1(\Omega)^N} - \langle J, \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)^N} \\
&= (\boldsymbol{p}_1, \boldsymbol{\varphi}_1)_{\mathbf{L}^2(\Omega)} + \frac{\tau^2}{4} a(\boldsymbol{p}_1, \boldsymbol{\varphi}_1) + \frac{\tau}{2} b(\boldsymbol{p}_1, \boldsymbol{\varphi}_1) \\
&\quad \sum_{k=1}^{N-1} \left[(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_{k+1} - \boldsymbol{\varphi}_k)_{\mathbf{L}^2(\Omega)} + \tau (\bar{q}_{k+1}(\boldsymbol{p}_{k+1}, \dots, \boldsymbol{p}_N), \boldsymbol{\varphi}_k)_{\mathbf{L}^2(\Omega)} \right. \\
&\quad \quad + \frac{\tau^2}{4} (a(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_{k+1}) + a(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_k)) \\
&\quad \quad \left. + \frac{\tau}{2} (b(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_{k+1}) - b(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_k)) \right] - \langle J, \boldsymbol{\varphi} \rangle_{(\mathbf{H}^1(\Omega)^N)^*} \\
&= (\boldsymbol{p}_N, \boldsymbol{\varphi}_N)_{\mathbf{L}^2(\Omega)} + \frac{\tau^2}{4} a(\boldsymbol{p}_N, \boldsymbol{\varphi}_N) + \frac{\tau}{2} b(\boldsymbol{p}_N, \boldsymbol{\varphi}_N) \tag{5.11} \\
&\quad \sum_{k=1}^{N-1} \left[(\boldsymbol{p}_k - \boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_k)_{\mathbf{L}^2(\Omega)} + \tau (\bar{q}_{k+1}(\boldsymbol{p}_{k+1}, \dots, \boldsymbol{p}_N), \boldsymbol{\varphi}_k)_{\mathbf{L}^2(\Omega)} \right. \\
&\quad \quad + \frac{\tau^2}{4} (a(\boldsymbol{p}_k, \boldsymbol{\varphi}_k) + a(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_k)) \\
&\quad \quad \left. + \frac{\tau}{2} (b(\boldsymbol{p}_k, \boldsymbol{\varphi}_k) - b(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}_k)) \right] - \langle J, \boldsymbol{\varphi} \rangle_{(\mathbf{H}^1(\Omega)^N)^*} \\
&= \sum_{k=1}^N d(\boldsymbol{p}_k, \boldsymbol{\varphi}_k) - \langle r_k(\boldsymbol{p}_{k+1}, \dots, \boldsymbol{p}_N, J), \boldsymbol{\varphi}_k \rangle_{\mathbf{H}^1(\Omega)}
\end{aligned}$$

for all $J \in (\mathbf{H}^1(\Omega)^N)^*$ and all $\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega)^N$, where $d: \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}$ means the bounded, coercive bilinear form that appeared in the time steps in the forward problem (4.9). This clearly shows the decoupling in the $\boldsymbol{\varphi}_k$ and therefore the time stepping structure of the adjoint problem, cf. 4.9.

Including the representation (5.9) of the adjoint velocities, the adjoint problem (Problem 5.4.1) can be reformulated as the following equivalent backwards-in-time time stepping scheme in the quantities $\boldsymbol{p}, \boldsymbol{q} \in \mathbf{H}^1(\Omega)^N$ and $\mu \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\boldsymbol{y}, \lambda)^\circ$:

$$\begin{aligned}
d(\boldsymbol{p}_N, \boldsymbol{\varphi}) - \langle \mu_N, \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)} &= \langle \partial_{\boldsymbol{y}_N} J_\tau(\boldsymbol{y}, \boldsymbol{u}), \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)} \\
\boldsymbol{q}_N &= -\frac{2}{\tau} \boldsymbol{p}_N \\
d(\boldsymbol{p}_k, \boldsymbol{\varphi}) - \langle \mu_k, \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)} &= (\boldsymbol{p}_{k+1}, \boldsymbol{\varphi})_{\mathbf{L}^2(\Omega)} - \frac{\tau^2}{4} a(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}) + \frac{\tau}{2} b(\boldsymbol{p}_{k+1}, \boldsymbol{\varphi}) \\
&\quad - \tau (\boldsymbol{q}_{k+1}, \boldsymbol{\varphi})_{\mathbf{L}^2(\Omega)} + \langle \partial_{\boldsymbol{y}_k} J_\tau(\boldsymbol{y}, \boldsymbol{u}), \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)} \\
\boldsymbol{q}_k &= -\boldsymbol{q}_{k+1} + \frac{2}{\tau} (\boldsymbol{p}_{k+1} - \boldsymbol{p}_k).
\end{aligned}$$

Introducing the adjoint velocities in the viscous forces of the update, we can rewrite the

update step for $k = 1, \dots, N-1$ as the following backwards-in-time Newmark-type scheme:

$$\begin{aligned} (\mathbf{q}_k, \boldsymbol{\varphi})_{\mathbf{L}^2(\Omega)} &= (\mathbf{q}_{k+1}, \boldsymbol{\varphi})_{\mathbf{L}^2(\Omega)} + \frac{\tau}{2} (a(\mathbf{p}_{k+1}, \boldsymbol{\varphi}) + a(\mathbf{p}_k, \boldsymbol{\varphi}) - b(\mathbf{q}_k, \boldsymbol{\varphi}) - b(\mathbf{q}_{k+1}, \boldsymbol{\varphi})) \\ &\quad - \frac{2}{\tau} \langle \mu_k, \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)} - \frac{2}{\tau} \langle \partial_{\mathbf{y}_k} J_\tau(\mathbf{y}, \mathbf{u}), \boldsymbol{\varphi} \rangle_{\mathbf{H}^1(\Omega)} \\ \mathbf{p}_k &= \mathbf{p}_{k+1} - \frac{\tau}{2} (\mathbf{q}_{k+1} + \mathbf{q}_k) \\ \mathbf{p}_k &\in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k), \quad \mathbf{q}_k \in \mathbf{H}^1(\Omega), \quad \mu_k \in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k)^\circ \end{aligned}$$

Note that as in the forward problem, the backward time steps are *implicit* in the multiplier μ and that the reversal of time introduces an opposing sign in the viscosity part and in the update of the adjoint states compared to the forward scheme.

5.4.2 Boundary Conditions and Existence of Adjoint States

The boundary conditions of the adjoint problem coincide with those of the linearized problem (4.17), since both involve the critical cone $\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, w - A_\tau \mathbf{y})$, which we have characterized in Corollary 4.4.24.

Corollary 5.4.4 *Let $\mathbf{u} \in \mathbf{U}_\tau$, $\mathbf{y} := S(B_\tau \mathbf{u} + f_\tau) \in \mathbf{K}_\Phi$ and $\lambda := B_\tau \mathbf{u} + f_\tau - A_\tau \mathbf{y}$. Then $\mathbf{p} \in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}, \lambda)$ if and only if*

$$\begin{aligned} \mathbf{p}_k &= 0 \text{ q.e. on } \Gamma_D, \\ \mathbf{p}_k \cdot \boldsymbol{\nu}_\Phi &= 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}_k), \\ \mathbf{p}_k \cdot \boldsymbol{\nu}_\Phi &\leq 0 \text{ q.e. on } \mathcal{W}_\Psi(\mathbf{y}_k) \end{aligned}$$

for all $k \in \{1, \dots, N\}$.

Hence, the adjoint problem includes standard Dirichlet boundary conditions on the Dirichlet boundary of the forward problem as well as sliding boundary conditions, with respect to the contact normal, on the strongly active contact set of the corresponding time step's displacement. Non-penetration is enforced on the weak contact set of the same displacement.

Theorem 5.3.1 ensures the existence of a solution to the adjoint problem only at a local minimizer to the optimal control problem but not for every admissible pairs of control and state. Specifically, consider $\mathbf{u} \in \mathbf{U}_\tau$ and $\mathbf{y} = S(B_\tau \mathbf{u} + f_\tau) \in \mathbf{H}^1(\Omega)^N$ that are not necessarily optimal for the optimal control problem (Problem 5.1.2). Using the time stepping structure (5.11), we know that solving the adjoint problem (Problem 5.4.1) consists of solving a sequence of time stepping problems of the type

$$\begin{aligned} \mathbf{p}_k &\in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k) && \text{or equivalently} && \mathbf{p}_k &\in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k) \\ \mu_k &\in \mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k)^\circ && && r_k - D\mathbf{p}_k &\in -\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k)^\circ, \\ D\mathbf{p}_k - \mu_k - r_k &= 0, && && & \end{aligned}$$

where D is the linear, bounded, coercive operator introduced in the forward problem, cf. Lemma 4.3.1, and r_k denotes the right hand side in the time step, cf. Lemma 5.4.3. Though the linearized non-penetration-type boundary condition for \mathbf{p}_k on $\mathcal{W}_\Psi(\mathbf{y}_k)$ in Corollary 5.4.4

suggest a contact problem structure, this is clearly not the case in general, as there is not necessarily any complementarity involving \mathbf{p}_k and the multiplier μ_k . Additionally, the sign on the multiplier is reversed. The multiplier μ_k can of course act as a contact force at the optimal pair $(\bar{u}, \bar{\mathbf{y}})$, establishing an equilibrium of adjoint forces, even though this is generally not the case for arbitrary pairs of admissible controls and displacements.

The adjoint problem is linear and uniquely solvable whenever the biactive set vanishes up to polar sets.

Lemma 5.4.5 *Let \mathcal{C} be a closed linear subspace of $\mathbf{H}^1(\Omega)^N$. Then there exists a linear and bounded solution operator that maps $J \in (\mathbf{H}^1(\Omega)^N)^*$ to the unique solution to the problem*

$$\begin{aligned} \mathbf{p} &\in \mathcal{C} \\ J - A_\tau^* \mathbf{p} &\in -\mathcal{C}^\circ. \end{aligned}$$

Proof. Since \mathcal{C} is a closed linear subspace of $\mathbf{H}^1(\Omega)^N$, it is a Hilbert space with the restricted scalar product of the product space, and we know that

$$-\mathcal{C}^\circ = \mathcal{C}^\perp = \prod_{k=1}^N \mathcal{C}_k^\perp,$$

where the component spaces \mathcal{C}_k are closed subspaces of $\mathbf{H}^1(\Omega)$ and therefore Hilbert with the standard $\mathbf{H}^1(\Omega)$ scalar product.

Due to (5.11) in the time stepping analysis of the adjoint problem, we know that

$$\langle J - A_\tau^* \mathbf{p}, \varphi \rangle_{\mathbf{H}^1(\Omega)^N} = \sum_{k=1}^N \langle r_k(\mathbf{p}_{k+1}, \dots, \mathbf{p}_N, J) - D\mathbf{p}_k, \varphi_k \rangle_{\mathbf{H}^1(\Omega)}$$

for all $\varphi \in \mathbf{H}^1(\Omega)^N$, where $D: \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}^1(\Omega)^*$ is defined by the bounded, coercive, bilinear form d in Lemma 4.3.1 (a), and the right hand sides $r_k: \mathbf{H}^1(\Omega)^{N-k} \times (\mathbf{H}^1(\Omega)^N)^* \rightarrow \mathbf{H}^1(\Omega)^*$ are as in Lemma 5.4.3 and therefore bounded and linear.

For $k = N$, we obtain a bounded linear solution operator that maps $J \in (\mathbf{H}^1(\Omega)^N)^*$ to the unique solution to the problem

$$\begin{aligned} \mathbf{p}_N &\in \mathcal{C}_N \\ r_N(J) - D\mathbf{p}_N &\in \mathcal{C}_N^\perp \end{aligned}$$

from the Lax-Milgram Lemma because the bilinear form d is bounded and coercive on the subspace $\mathcal{C}_N \subset \mathbf{H}^1(\Omega)$, and $r_N(J) = J_N$.

For the induction argument in $k < N$, assume $\mathbf{p}_i = \mathbf{p}_i(J)$ to be the image of the bounded linear solution operators that maps $J \in (\mathbf{H}^1(\Omega)^N)^*$ to the solution \mathbf{p}_i of the problem

$$\begin{aligned} \mathbf{p}_i &\in \mathcal{C}_i \\ r_i(\mathbf{p}_{i+1}(J), \dots, \mathbf{p}_N(J), J) - D\mathbf{p}_i &\in \mathcal{C}_i^\perp \end{aligned}$$

for i in $\{k+1, \dots, N\}$. Again, Lax-Milgram's Lemma yields a solution operator for the problem

$$\begin{aligned} \mathbf{p}_k &\in \mathcal{C}_k \\ r_k(\mathbf{p}_{k+1}(J), \dots, \mathbf{p}_N(J), J) - D\mathbf{p}_k &\in \mathcal{C}_k^\perp \end{aligned}$$

that is bounded and linear in the right hand side r_k , which are bounded and linear themselves. Linearity and boundedness of the solution operator in $J \in (\mathbf{H}^1(\Omega)^N)^*$ follows from the induction hypothesis. \square

Accordingly, given an admissible pair of control and state where the active set of the state contains no weakly active part, we can compute a corresponding adjoint state.

Corollary 5.4.6 *Let $\mathbf{u} \in U_\tau$, $\mathbf{y} = S(B_\tau \mathbf{u} + f_\tau) \in \mathbf{K}_\Phi^N$, $\lambda = B_\tau \mathbf{u} + f_\tau - A_\tau \mathbf{y} \in \mathcal{T}_{\mathbf{K}_\Phi^N}(\mathbf{y})^\circ$ and $\mathcal{W}_\Psi(\mathbf{y}_k) = \emptyset \forall k = 1, \dots, N$. Then there exists a unique solution to the adjoint problem*

$$\begin{aligned} \mathbf{p} &\in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda) \\ \partial_{\mathbf{y}} J_\tau(\mathbf{y}, \mathbf{u}) - A_\tau^* \mathbf{p} &\in -\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^\circ. \end{aligned}$$

Proof. The assumption $\mathcal{W}_\Psi(\mathbf{y}) = \emptyset$ implies that

$$\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda) = \{ \delta \mathbf{y} \in H_D^1 : \delta \mathbf{y} \cdot \nu_\Phi = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}) \},$$

which is a closed linear space, see also Corollary 5.4.4 and Proposition B.3.5, therefore Lemma 5.4.5 yields the claim. \square

5.5 Differentiability of the Reduced Objective Functional

Section 4.4 concerned the differentiability properties of the solution operator S of the variational inequality. We will now examine the implications for the differentiability properties of the control-to-state operator and subsequently the reduced objective functional, and we will see how to compute the Fréchet derivatives of the objective functional using the adjoint states. The results will be used in the following numerics chapter to compute search directions in iterative, adjoint-based optimization schemes.

Since the control-to-force operator $B_\tau: U_\tau \rightarrow (\mathbf{H}^1(\Omega)^N)^*$ is linear and bounded, the control-to-state operator $S \circ B_\tau$ of course inherits the differentiability properties of S . When B_τ is additionally assumed to be compact, we can derive Fréchet differentiability of the control-to-state map from the Gâteaux differentiability of S in Section 4.4 by a similar argument as in the standard proof in finite dimensions.

Lemma 5.5.1 *Let X, Y, Z be Banach spaces, $G: X \rightarrow Y$, $F: Y \rightarrow Z$, and let $x \in X$. If*

- (a) G is affine linear, continuous and compact,
- (b) F is Lipschitz continuous and
- (c) F is Gâteaux differentiable at $G(x)$,

then $F \circ G$ is Fréchet differentiable at x .

Note that G is not required to be a *linear* bounded operator.

Proof. See Lemma C.4.1 of the appendix for the technical proof. \square

Compactness of the control-to-force operator therefore implies Fréchet differentiability of the control-to-state operator and the reduced objective functional in the optimal control problem.

Theorem 5.5.2 *Let Assumption 5.1.1 be satisfied, $\mathbf{u} \in \mathbf{U}_\tau$ and let $\mathbf{y} = \tilde{S}(\mathbf{u}) \in \mathbf{K}_\Phi^N \subset \mathbf{H}^1(\Omega)^N$ be the corresponding displacement. Then \tilde{J}_τ is Hadamard differentiable at \mathbf{u} , and*

$$\tilde{J}_\tau'(\mathbf{u}, \delta\mathbf{u}) = \tilde{J}_\tau'(\mathbf{u})\tilde{S}'(\mathbf{u}, \delta\mathbf{u}) = \tilde{J}_\tau'(\mathbf{u})S'(B_\tau\mathbf{u} + f_\tau, B_\tau\delta\mathbf{u}).$$

If B_τ is compact, and $\mathcal{W}_\Psi(\mathbf{y}_k) = \emptyset$ for all k in $\{1, \dots, N\}$, then $\tilde{S}: \mathbf{U}_\tau \rightarrow \mathbf{K}_\Phi^N$ and $\tilde{J}_\tau: \mathbf{U}_\tau \rightarrow \mathbb{R}$ are Fréchet differentiable at \mathbf{u} , and

$$\tilde{J}_\tau'(\mathbf{u}) = B_\tau^* \mathbf{p} + \partial_u J_\tau(\mathbf{y}, \mathbf{u}),$$

where \mathbf{p} is the unique solution to the adjoint problem (Problem 5.4.1).

Proof. Since J_τ is assumed to be Fréchet differentiable, the Hadamard differentiability of \tilde{J}_τ and the form of its derivative is an implication of Theorem 4.4.12 and Lemma 4.4.2 (c).

When B_τ is compact, then the operator $\mathbf{u} \mapsto B_\tau\mathbf{u} + f_\tau$ is affine linear, continuous and compact. The solution operator S to the discretized problem was established to be Lipschitz continuous in Theorem 4.3.3, and the vanishing biactive sets of the corresponding displacement \mathbf{y} yields that S is Gâteaux differentiable at $w := B_\tau\mathbf{u} + f_\tau$, see Proposition 4.4.25. Lemma 5.5.1 therefore ensures Fréchet differentiability of $\tilde{S} = S(B_\tau(\cdot) + f_\tau)$ at \mathbf{u} .

Since J_τ is assumed to be Fréchet differentiable, we immediately obtain Fréchet differentiability of the reduced objective functional \tilde{J}_τ at \mathbf{u} with

$$\tilde{J}_\tau'(\mathbf{u}) = \partial_y J_\tau(\mathbf{y}, \mathbf{u})S'(w)B_\tau + \partial_u J_\tau(\mathbf{y}, \mathbf{u}) = B_\tau^* S'(w)^* \partial_y J_\tau(\mathbf{y}, \mathbf{u}) + \partial_u J_\tau(\mathbf{y}, \mathbf{u}).$$

The adjoint state \mathbf{p} is the unique solution to Problem 5.4.1, which reads as

$$\begin{aligned} \mathbf{p} &\in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda) \\ \partial_y J_\tau(\mathbf{y}, \mathbf{u}) - A_\tau^* \mathbf{p} &\in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^\perp \end{aligned} \tag{5.12}$$

for $\lambda := w - A_\tau\mathbf{y}$, see Corollary 5.4.6, because the absence of the biactive sets for \mathbf{y} implies that the critical cone is a closed linear space, see Proposition 4.4.21. We can therefore confirm the representation of the Fréchet derivative of the reduced objective functional in the claim by showing that the primal representative of $S'(w)^* \partial_y J_\tau(\mathbf{y}, \mathbf{u}) \in (\mathbf{H}^1(\Omega)^N)^{**}$ solves the adjoint problem and therefore coincides with the adjoint state. Because the critical cone is a closed linear space, the primal representative of $S'(w)^* \partial_y J_\tau(\mathbf{y}, \mathbf{u})$ is in $\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)$ if and only if

$$S'(w)^* \partial_y J_\tau(\mathbf{y}, \mathbf{u}) \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^{\perp\perp}, \tag{5.13}$$

as a consequence of the Hahn-Banach Theorem. We will prove (5.13) by showing that $\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^\perp \subset \ker(S'(w))$. To that end, let $\delta w \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}, \lambda)^\perp$. As we have seen in the

proof of Proposition 4.4.25, the Gâteaux derivative $S'(w)$ maps δw to the unique solution $\delta \mathbf{y} := S'(w)\delta w$ of

$$\begin{aligned} \delta \mathbf{y} &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y}) \\ \delta w - A_{\tau}\delta \mathbf{y} &\in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y})^{\perp}. \end{aligned} \quad (5.14)$$

Testing the second line with $\delta \mathbf{y}$, we obtain $\langle A_{\tau}\delta \mathbf{y}, \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} = 0$, and the specific form of A_{τ} yields

$$\begin{aligned} 0 &= \langle A_{\tau}\delta \mathbf{y}, \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} = d(\delta \mathbf{y}_1, \delta \mathbf{y}_1) - \langle l_1(0), \delta \mathbf{y}_1 \rangle_{\mathbf{H}^1(\Omega)} \\ &\quad + \sum_{k=1}^{N-1} d(\delta \mathbf{y}_{k+1}, \delta \mathbf{y}_{k+1}) - \langle l_{k+1}(\delta \mathbf{y}_1, \dots, \delta \mathbf{y}_k, 0), \delta \mathbf{y}_{k+1} \rangle_{\mathbf{H}^1(\Omega)}, \end{aligned}$$

see (4.9). Linearity of the right hand sides l_k and the coercivity of the bilinear form d inductively yields that $\delta \mathbf{y} = 0$.

In order to confirm the second line of the adjoint problem (5.12) for $S'(w)^*\partial_y J_{\tau}(\mathbf{y}, \mathbf{u})$, we again use that $S'(w)$ is the solution operator to the linearized problem (5.14), which implies

$$S'(w)A_{\tau}\delta \mathbf{y} = \delta \mathbf{y} \quad \forall \delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y}).$$

Therefore,

$$\begin{aligned} &\langle \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} - \langle A_{\tau}^* S'(w)^* \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} \\ &= \langle \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), \delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} - \langle \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), S'(w)A_{\tau}\delta \mathbf{y} \rangle_{\mathbf{H}^1(\Omega)^N} = 0 \end{aligned}$$

is satisfied for all $\delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}, w - A_{\tau}\mathbf{y})$. Accordingly, the element $S'(w)^*\partial_y J_{\tau}(\mathbf{y}, \mathbf{u})$ solves the uniquely solvable adjoint problem, i.e., it coincides with the adjoint state. \square

Chapter 6

Numerical Optimization

In the previous chapter, we have seen that minimizers of the contact constrained optimal control problem exist in a reasonable setting, and we have linked the representations of the reduced objective functional's derivatives to the adjoint problem of the first order conditions. We will now focus on the implementation and evaluation of two adjoint-based optimization schemes and their application to the contact constrained optimal control problem to compute its minimizers. Several combinations of (accelerated) (sub-)gradient-type methods and different line search approaches will be applied to solve three test problems with distributed controls and tracking-type functionals. The corresponding results are discussed with respect to the performance of the algorithms and the implications on the problem structure of the respective test configurations.

Complementarity constrained optimization problems in *finite* dimensions have been treated numerically by a number of authors and algorithms. E.g., Fletcher and Leyffer [67] present the numerical results for several NLP solvers, a filter SQP method and a large number of (finite-dimensional) mathematical problems with complementarity constraints. Several different regularization techniques and interior-point methods are employed in [128, 131, 159], and [86, 171, 198] contain trust-region bundle methods for nonsmooth problems and some applications to discretized optimization problems with variational inequalities in function space. Algorithms and numerical results for complementarity constrained problems that are formulated in function space are less commonly found in the literature. Refer to [192] for the solution of problems in quasistatic plasticity by a nonlinear CG-type method, and see [187] for a problem in thermoviscoplasticity, including extensive numerical results. Christof et al. characterize the Bouligand subdifferential(s) of the solution operator to an elliptic variational inequality with max-type nonsmoothness and $L^2(\Omega)$ -controls and present numerical results for tracking-type test configurations based on regularization and a semismooth Newton method in [43]. Similarly, Rauls and Wachsmuth characterize the Bouligand subdifferential(s) of the solution operator to the obstacle problem as an operator from $H_0^1(\Omega)^*$ to $H^1(\Omega)$ in [162]. In an optimal control setting including more general controls, such as, e.g., $L^2(\Omega)$ -controls, a way to compute a single subgradient for the reduced functional via adjoint techniques is developed by Rauls and S. Ulbrich in [161]. Hertlein and M. Ulbrich present corresponding numerical results using an inexact bundle method in Hilbert space [90].

Structure. The first section of this chapter is dedicated to the description of the (sub-) gradient-type optimization scheme and its accelerated modification with a focus on the adjoint-based computation of the search direction. We employ several (well-known) line search algorithms, which makes for a total of five different implementations that are compared for the test problem setting of tracking-type functionals with distributed controls and plane strain states. The setting is briefly reviewed in Section 6.2. The numerical results of the algorithms' implementations are presented and analyzed in Sections 6.2.2–6.2.4, and an evaluation of the algorithms and their performance in Section 6.3 closes off the chapter.

The optimization algorithm based on the search directions introduced in Section 6.1.1 and the example of Section 6.2.2 were previously published in [143]. This chapter contains unpublished results for two momentum approaches for accelerating the line search algorithm and an extensive amount of additional analysis on the behavior of the algorithms and the implications on the problem structure.

6.1 Adjoint-Based Optimization Algorithms

The solution operator to the contact problem is generally only Hadamard differentiable and Fréchet differentiable on a dense subset, see Section 4.4, which renders the problem unsuited for treatment by fast second order methods. Due to the benefit of low computational cost per iteration, first order methods have experienced an increase in interest lately. Popular methods that have recently been considered include proximal and augmented Lagrangian methods [12, 13, 95, 108, 133], (nonlinear) conjugate (sub-)gradient schemes [82, 93, 94], accelerated schemes [26, 146, 147, 153, 166] as well as (bundle) trust region methods [43, 90, 95, 171]. See [152] and the recently published [25] for an overview of first order methods.

In the remainder of this chapter, we will study a basic (sub-)gradient-type line search method in function space as well as an accelerated momentum method with and without restart. We are interested in the behavior of their implementations in application to multiple test configurations of the optimal control problem stated in Problem 5.1.2, i.e., the unconstrained problem

$$\min_{\mathbf{u} \in \mathcal{U}_\tau} \tilde{J}_\tau(\mathbf{u}), \quad (6.1)$$

cf. the reduced form in Definition 5.1.3.

Standard Line Search Method. Line search methods form a large class of optimization schemes for unconstrained optimization problems. Generally, they consist of the computation of a search direction and a step length for the update in each iteration. Popular examples are, e.g., (quasi-)Newton methods and preconditioned (sub-)gradient schemes. For more information on line search methods, including higher order search directions, see [152, Chap. 3].

Algorithm 6.1 Line Search**Require:** $\mathbf{u}_0 \in \mathcal{U}_\tau$, $k = 0$

- 1: **while** stopping criterion not satisfied **do**
- 2: Compute search direction $\delta \mathbf{u}_k$ (Algorithm 6.4)
- 3: Compute step length s_k (Section 6.1.2)
- 4: Update $\mathbf{u}_{k+1} = \mathbf{u}_k + s_k \delta \mathbf{u}_k$
- 5: $k \leftarrow k + 1$
- 6: **return** $\bar{\mathbf{u}} = \mathbf{u}_k$

We will apply Algorithm 6.1 to (6.1) using the adjoint-based search direction presented in Section 6.1.1 below and several algorithms for the choice of the step length, see Section 6.1.2.

Accelerated Line Search Methods. Accelerated momentum-type methods as a modification of first order methods go back to Nesterov’s work [145] in convex optimization. They rely on an auxiliary step that adds an additional correction to the update in line 4 of Algorithm 6.1. Usually, the influence of this step is chosen to increase within certain bounds as the algorithm advances. The accelerated line search method described below is essentially taken from [145].

Algorithm 6.2 Accelerated Line Search without Restart (AccNoRe)**Require:** $\mathbf{u}_0, \mathbf{v}_0 \in \mathcal{U}_\tau$, $\mathbf{v}_0 = 0$, $k = 0$, $t_0 = 1$

- 1: **while** stopping criterion not satisfied **do**
- 2: Compute search direction $\delta \mathbf{u}_k$ (Algorithm 6.4)
- 3: Compute step length s_k (Algorithm 6.6)
- 4: $\mathbf{v}_{k+1} = \mathbf{u}_k + s_k \delta \mathbf{u}_k$
- 5: $t_{k+1} = \frac{1}{2}(1 + \sqrt{1 + 4t_k^2})$
- 6: $\mathbf{u}_{k+1} = \mathbf{v}_{k+1} + \frac{t_k - 1}{t_{k+1}}(\mathbf{v}_{k+1} - \mathbf{v}_k)$
- 7: $k \leftarrow k + 1$
- 8: **return** $\bar{\mathbf{u}} = \mathbf{u}_k$

Steps 2–4 of Algorithm 6.1 remain unchanged in principal, but the line search update is used for the update of an auxiliary variable in the accelerated scheme. The update of the actual iterate then consists of the line search correction and an additional correction that moves the iterate in the direction of the difference of the auxiliary results, cf. step 6.

Remark 6.1.1 *Nesterov calls the update step 6 a “ravine” step [145, P. 372], presumably for its potential to reduce zig-zagging effects in narrow valleys.*

Clearly, the parameters t_k that are determined in step 5 of Algorithm 6.2 are strictly monotonically increasing and unbounded, hence the momentum parameter

$$m_{k+1} := \frac{t_k - 1}{t_{k+1}}$$

for the additional correction in the update step is non-negative and converges to 1 as $k \rightarrow \infty$, which increases the influence of the momentum term as the algorithm progresses.

Note that the first iteration of the algorithm is always a standard line search step while the second is a simple “overshooting” of the intermediate step.

Accelerated schemes like Algorithm 6.2 generally are not descent schemes even if the computed corrections $\delta \mathbf{u}_k$ are descent directions. A characteristic behavior of these algorithms are “ripples” in the values of the objective functional over the course of the run, i.e., blocks of iterations that decrease and increase the objective value alternate while the objective values decrease on a larger scale, see Figure C.3. A simple but effective way to counteract this behavior and increase the performance of these methods is by adding the option of *restarting* the algorithm as proposed in [78, 147, 153]. When an appropriate condition is satisfied in an iteration, a standard line search step is carried out instead of the accelerated step, and the momentum is reset by setting $t_{k+1} = 1$ (and therefore $m_{k+1} = 0$). The next update in the algorithm is another standard line search step, and the momentum starts building up from zero. Several restarting conditions are investigated in [153], including restarts after a fixed number of iterations as well as depending on conditions involving the objective values and the derivatives. We will consider the modification of the initial scheme that restarts whenever ascent is detected, thus ensuring a descent algorithm.

Algorithm 6.3 Accelerated Line Search with Restart at Ascent (AccRe)

Require: $\mathbf{u}_0, \mathbf{v}_0 \in \mathcal{U}_\tau$, $\mathbf{v}_0 = \mathbf{0}$, $k = 0$, $t_0 = 1$

- 1: **while** stopping criterion not satisfied **do**
 - 2: Compute search direction $\delta \mathbf{u}_k$ (Algorithm 6.4)
 - 3: Compute step length s_k (Algorithm 6.6)
 - 4: $\mathbf{v}_{k+1} = \mathbf{u}_k + s_k \delta \mathbf{u}_k$
 - 5: $t_{k+1} = \frac{1}{2}(1 + \sqrt{1 + 4t_k^2})$
 - 6: $m_{k+1} = \frac{t_k - 1}{t_{k+1}}$
 - 7: **if** $\tilde{J}_\tau(\mathbf{v}_{k+1} + m_{k+1}(\mathbf{v}_{k+1} - \mathbf{v}_k)) < \tilde{J}_\tau(\mathbf{u}_k)$ **then**
 - 8: $\mathbf{u}_{k+1} = \mathbf{v}_{k+1} + m_{k+1}(\mathbf{v}_{k+1} - \mathbf{v}_k)$
 - 9: **else**
 - 10: $\mathbf{u}_{k+1} = \mathbf{v}_{k+1}$
 - 11: $t_{k+1} = 1$
 - 12: $k \leftarrow k + 1$
 - 13: **return** $\bar{\mathbf{u}} = \mathbf{u}_k$
-

In the restarted algorithm, steps 2–5 remain unchanged from the accelerated method in Algorithm 6.2, but the accelerated update is only accepted if it yields descent in the objective values. Otherwise, a standard line search iteration is carried out, and the parameter t_{k+1} is reset to 1, as described above.

Remark 6.1.2 *To a degree, taking prior steps into account for the current correction adds a memory effect to the line search method. While this approach and the possible restart management may be reminiscent of nonlinear conjugated gradient methods, e.g., [152, Sec. 5.2] and [50], the behavior of nonlinear CG updates and the momentum methods above can already be observed to differ significantly in smooth problems.*

6.1.1 Search Directions

A standard approach for obtaining search directions in line search methods is to construct local models $m_k(\delta \mathbf{u})$ of the functional $\tilde{J}_\tau(\mathbf{u}_k + \delta \mathbf{u})$ that allow for a relatively uncomplicated computation of unique minimizers, so that

$$\delta \mathbf{u}_k := \arg \min_{\delta \mathbf{u} \in \mathcal{U}_\tau} m_k(\delta \mathbf{u})$$

can be used in the update step. Of course, Taylor-type expansions of the functional are popular models. Choosing the order of the expansion yields a straight forward method to balance the amount of information included in the model against the work required to compute the minimizer of the model.

In the case of the optimal control problem (6.1), a local model that incorporates the first order information of the problem completely can be constructed using the Hadamard derivative of the reduced objective functional from Section 5.5. A quadratic term in the model, which acts as a preconditioner, can be chosen depending on the control and the specific problem structure.

Model 6.1.3 *Let Assumptions 5.1.1 hold, let $\mathbf{u}_k \in \mathcal{U}_\tau$ and $q: \mathcal{U}_\tau \times \mathcal{U}_\tau \rightarrow \mathbb{R}$ be a bounded, coercive, symmetric bilinear form, and define*

$$m_k(\delta \mathbf{u}) := \tilde{J}_\tau(\mathbf{u}_k) + \tilde{J}_\tau'(\mathbf{u}_k, \delta \mathbf{u}) + \frac{1}{2}q(\delta \mathbf{u}, \delta \mathbf{u}).$$

Using the representation of the Hadamard derivative of the reduced objective functional and the solution operator, we can rewrite the minimization of m_k to find the search direction $\delta \mathbf{u}_k$ as the optimal control problem

$$\begin{aligned} \min & J_\tau(\mathbf{y}_k, \mathbf{u}_k) + J_\tau'(\mathbf{y}_k, \mathbf{u}_k)(\delta \mathbf{y}, \delta \mathbf{u}) + \frac{1}{2}q(\delta \mathbf{u}, \delta \mathbf{u}) \\ \text{s.t. } & (\delta \mathbf{y}, \delta \mathbf{u}) \in \mathbf{H}^1(\Omega)^N \times \mathcal{U}_\tau \\ & \delta \mathbf{y} \in \mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}_k, \lambda_k) \\ & B_\tau \delta \mathbf{u} - A_\tau \delta \mathbf{y} \in \mathcal{T}_{\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}_k, \lambda_k)}(\delta \mathbf{y})^\circ, \end{aligned} \tag{6.2}$$

where $\mathbf{y}_k = S(B_\tau \mathbf{u}_k + f_\tau)$, $\lambda_k = B_\tau \mathbf{u}_k + f_\tau - A_\tau \mathbf{y}$ and $\mathcal{C}_{\mathbf{K}_\Phi^N}(\mathbf{y}_k, \lambda_k)$ is the closed and convex critical cone to \mathbf{K}_Φ^N w.r.t. (\mathbf{y}, λ_k) . Clearly, including the full first order information of the nonsmooth problem comes at the cost of a highly nonlinear model, seeing as the directional derivatives are generally only positively homogeneous but not linear. Using the techniques employed in Chapter 5, we could again establish the existence of minimizers to (6.2) and obtain optimality conditions of first order, but the model's minimizers are generally non-unique, and it is unclear how to compute one from the first order conditions. Consequently, we rely on a linearization of the constraint in the model.

Linearized Local Models. Recall that Proposition 4.4.21 provides an explicit, point-wise representation of the critical cone as

$$\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k) = \prod_{l=1}^N \{ \delta \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi \leq 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}_{k,l}), \\ \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi = 0 \text{ q.e. on } \mathcal{S}_\Psi(\mathbf{y}_{k,l}) \},$$

where $\mathbf{y}_{k,l}$ denotes the l^{th} component of the iterate \mathbf{y}_k . The nonlinearity in the model is due to the fact that $\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k)$ is a linear set if and only if the weak contact sets $\mathcal{W}_\Psi(\mathbf{y}_{k,l}) = \mathcal{A}_\Psi(\mathbf{y}_{k,l}) \setminus \mathcal{S}_\Psi(\mathbf{y}_{k,l})$ have vanishing capacity for all time steps $l = 1, \dots, N$, cf. Section 4.4.2. We therefore modify the constraints of the model (6.2) by enforcing the sliding boundary conditions on the entire contact sets in each of the time steps, i.e., by setting

$$\hat{\mathcal{C}}(\mathbf{y}_k) := \prod_{l=1}^N \{ \delta \mathbf{y} \in \mathbf{H}_D^1(\Omega) : \delta \mathbf{y} \cdot \boldsymbol{\nu}_\Phi = 0 \text{ q.e. on } \mathcal{A}_\Psi(\mathbf{y}_{k,l}) \}, \quad (6.3)$$

which is a closed linear subspace of $\mathcal{C}_{\mathbf{K}_\Phi}(\mathbf{y}_k, \lambda_k)$, cf. Proposition B.3.5. The corresponding modified local model is given by

$$\begin{aligned} \min J_\tau(\mathbf{y}_k, \mathbf{u}_k) + J'_\tau(\mathbf{y}_k, \mathbf{u}_k)(\delta \mathbf{y}, \delta \mathbf{u}) + \frac{1}{2}q(\delta \mathbf{u}, \delta \mathbf{u}) \\ \text{s.t. } (\delta \mathbf{y}, \delta \mathbf{u}) \in \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau \\ \delta \mathbf{y} \in \hat{\mathcal{C}}(\mathbf{y}_k) \\ B_\tau \delta \mathbf{u} - A_\tau \delta \mathbf{y} \in \hat{\mathcal{C}}(\mathbf{y}_k)^\circ. \end{aligned} \quad (6.4)$$

Since (6.4) is a linear-quadratic problem, existence of a unique minimizer and the corresponding first order optimality conditions are easily obtained. As we have done before, we identify the bidual elements of reflexive Banach spaces with their primal representatives.

Proposition 6.1.4 *Let Assumption 5.1.1 be satisfied, $k \in \mathbb{N}$, $\mathbf{u}_k \in \mathbf{U}_\tau$ and let $\mathbf{y}_k = \tilde{S}(\mathbf{u}_k) \in \mathbf{K}_\Phi^N \subset \mathbf{H}^1(\Omega)^N$ be the corresponding displacement. Then (6.4) has a unique solution $(\delta \mathbf{y}_k, \delta \mathbf{u}_k) \in \mathbf{H}^1(\Omega)^N \times \mathbf{U}_\tau$, and there exists a unique adjoint state $\delta \mathbf{p}_k \in \mathbf{H}^1(\Omega)^N$ that solves*

$$\begin{aligned} \delta \mathbf{p}_k \in \hat{\mathcal{C}}(\mathbf{y}_k) \\ \partial_{\mathbf{y}} J(\mathbf{y}_k, \mathbf{u}_k) - A_\tau^* \delta \mathbf{p}_k \in \hat{\mathcal{C}}(\mathbf{y}_k)^\circ. \end{aligned} \quad (6.5)$$

The adjoint state and the solution satisfy the stationarity condition

$$B_\tau^* \delta \mathbf{p}_k + \partial_{\mathbf{u}} J(\mathbf{y}_k, \mathbf{u}_k) + q(\delta \mathbf{u}_k, \cdot) = 0. \quad (6.6)$$

Proof. The claim is a standard result in convex analysis, and the proof will only be reproduced quickly for the readers convenience. First of all, recall that we have established the Lipschitz continuous solution operator S to the contact problem in Theorem 4.3.3 without

the use of the specific form of \mathbf{K}_{Φ}^N , but we only required the set in the constraint to be closed and convex. We can apply the arguments verbatim for the system

$$\begin{aligned}\delta \mathbf{y} &\in \hat{\mathcal{C}}(\mathbf{y}_k) \\ \omega - A_{\tau} \delta \mathbf{y} &\in \hat{\mathcal{C}}(\mathbf{y}_k)^{\circ}\end{aligned}$$

in the model (6.4) with $\omega \in (\mathbf{H}^1(\Omega)^N)^*$ to obtain a Lipschitz continuous solution operator $S_{\hat{\mathcal{C}}}: (\mathbf{H}^1(\Omega)^N)^* \rightarrow \hat{\mathcal{C}}(\mathbf{y}_k)$. Its linearity is an immediate consequence of the linearity of A_{τ} and $\hat{\mathcal{C}}(\mathbf{y}_k)$. The reduced objective functional of the model

$$\tilde{m}_k(\delta \mathbf{u}) := J_{\tau}(\mathbf{y}_k, \mathbf{u}_k) + J_{\tau}'(\mathbf{y}_k, \mathbf{u}_k)(S_{\hat{\mathcal{C}}} B_{\tau} \delta \mathbf{u}, \delta \mathbf{u}) + \frac{1}{2} q(\delta \mathbf{u}, \delta \mathbf{u}) \quad (6.7)$$

is therefore continuous, proper and strictly convex, which makes it weakly lower semi-continuous in $\delta \mathbf{u}$ [20, Prop. 2.10]. The coercivity of the quadratic part q and the strict convexity therefore yields the existence of a unique minimizer $(\delta \mathbf{u}_k, \delta \mathbf{y}_k)$ [20, Thm. 2.11, Rem. 2.12/2.13].

Since the reduced objective functional of the model in (6.7) is Fréchet differentiable, we obtain the optimality condition

$$B_{\tau}^* S_{\hat{\mathcal{C}}}^* \partial_y \tilde{J}_{\tau}(\mathbf{y}_k, \mathbf{u}_k) + \partial_u \tilde{J}_{\tau}(\mathbf{y}_k, \mathbf{u}_k) + q(\delta \mathbf{u}_k, \cdot) = 0$$

in \mathbf{U}_{τ}^* at $\delta \mathbf{u}_k$. It remains to show that $S_{\hat{\mathcal{C}}}^* \partial_y \tilde{J}_{\tau}(\mathbf{y}_k, \mathbf{u}_k)$ solves the adjoint system in the claim, i.e., the system

$$\begin{aligned}\delta \mathbf{p}_k &\in \hat{\mathcal{C}}(\mathbf{y}_k) \\ \partial_y J(\mathbf{y}_k, \mathbf{u}_k) - A_{\tau}^* \delta \mathbf{p}_k &\in \hat{\mathcal{C}}(\mathbf{y}_k)^{\circ},\end{aligned}$$

which we know to be uniquely solvable due to Lemma 5.4.5.

Analogously to the proof of Theorem 5.5.2, we can now argue that since $\hat{\mathcal{C}}(\mathbf{y}_k)$ is closed, $S_{\hat{\mathcal{C}}}^* \partial_y \tilde{J}_{\tau}(\mathbf{y}_k, \mathbf{u}_k) \in \hat{\mathcal{C}}(\mathbf{y}_k)^{\perp}$ if and only if its primal representative is in $\hat{\mathcal{C}}(\mathbf{y}_k)$, which is satisfied because clearly $\hat{\mathcal{C}}(\mathbf{y}_k)^{\perp} \subset \ker(S_{\hat{\mathcal{C}}})$. Of course

$$\begin{aligned}\langle \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), \varphi \rangle_{\mathbf{H}^1(\Omega)^N} - \langle A_{\tau}^* S_{\hat{\mathcal{C}}}^* \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), \varphi \rangle_{\mathbf{H}^1(\Omega)^N} \\ = \langle \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), \varphi \rangle_{\mathbf{H}^1(\Omega)^N} - \langle \partial_y J_{\tau}(\mathbf{y}, \mathbf{u}), S_{\hat{\mathcal{C}}} A_{\tau} \varphi \rangle_{\mathbf{H}^1(\Omega)^N} = 0\end{aligned}$$

for all $\varphi \in \hat{\mathcal{C}}(\mathbf{y}_k)$, which finalizes the proof. \square

Discussion of the Search Direction. According to the modified model, we compute the search direction $\delta \mathbf{u}_k$ at iteration k from the stationarity condition (6.6) using the q -Riesz isomorphism on \mathbf{U}_{τ} .

Algorithm 6.4 Computation of Search Directions $\delta \mathbf{u}_k$ **Require:** $k \in \mathbb{N}$, $\mathbf{u}_k \in U_\tau$

- 1: Compute state $\mathbf{y}_k = S(B_\tau(\mathbf{u}_k) + f_\tau)$
- 2: Compute modified adjoint state $\delta \mathbf{p}_k$ by solving (6.5)
- 3: Compute search direction $\delta \mathbf{u}_k$ as the q -representative of $-B_\tau^* \delta \mathbf{p}_k - \partial_u J(\mathbf{y}_k, \mathbf{u}_k)$
- 4: **return** $\delta \mathbf{u}_k$

The price of the linearization in step 2 is of course that the search directions may not be first order consistent in a sufficiently large neighborhood of the respective iterates if nonsmoothness in the vicinity can not be neglected.

Note that the modification of the critical cone that we employed to obtain the linear set $\hat{\mathcal{C}}(\mathbf{y}_k)$ is somewhat arbitrary. In fact, defining $\hat{\mathcal{C}}(\mathbf{y}_k)$ by enforcing the sliding boundary conditions on any set $\mathcal{B}_{k,l} \subset \Gamma_C$ that satisfies $\mathcal{S}_\Psi(\mathbf{y}_{k,l}) \subset \mathcal{B}_{k,l} \subset \mathcal{A}_\Psi(\mathbf{y}_{k,l})$ up to polar sets yields a closed linear subspace of the critical cone. Our choice in (6.3) yields the smallest of these subspaces, which is the easiest choice to implement numerically since weakly active sets are inherently difficult to detect.

Every specific choice of $\mathcal{B}_{k,l}$ as described above adds a certain bias to the search direction of the next update step and leaves some room for tracking the change of the strongly active sets over the course of the algorithm and improving the quality of the search direction by choosing an advantageous $\mathcal{B}_{k,l}$. In fact, in finite dimensions, the works of Outrata et al. [155, Prop. 7.14] and Christof et al. [43, Sec. 4.1] suggest that choosing any subset \mathcal{B}_k as above results in the computation of a subgradient. Similarly, Rauls and Wachsmuth [162] show that the Bouligand subdifferential of the solution operator to the infinite-dimensional contact problem with right hand sides in $H_0^1(\Omega)^*$ is comprised of all solution operators to linearized problems that include zero boundary conditions on corresponding sets $\mathcal{B}_{k,l}$ as above. It is unclear, however, whether this assessment holds for more restricted right hand sides, e.g., controls in $L^2(\Omega)$. Nonetheless, Rauls and Ulbrich [161] show that a subgradient of the reduced objective functional in the optimal control of the contact problem can be computed by choosing $\mathcal{B}_{k,l}$ equal to the entire active sets, as we have done above. While this issue has not been considered for contact problems so far, this is a strong indication that computing a search direction as described in Algorithm 6.4 corresponds to the computation of a q -subgradient of the reduced objective functional at the iterate, which essentially makes Algorithm 6.1 a subgradient scheme.

Finally, notice that when the weakly active set has vanishing capacity and the reduced objective functional is Fréchet differentiable at the current iterate, the method performs a standard q -gradient step.

Proposition 6.1.5 *Let Assumption 5.1.1 be satisfied, B_τ be compact and $\mathbf{u} \in U_\tau$. Let $\mathbf{y} = \tilde{S}(\mathbf{u}) \in \mathbf{K}_{\Phi}^N \subset \mathbf{H}^1(\Omega)^N$ be the corresponding displacement and $\mathcal{W}_\Psi(\mathbf{y}_k) = \emptyset$ for all $k = 1, \dots, N$. Then Algorithm 6.4 computes the q -gradient of the reduced objective functional at \mathbf{u}_k .*

Proof. Since the weak contact sets vanish up to capacity zero, Theorem 5.5.2 ensures that $\tilde{J}_\tau: U_\tau \rightarrow \mathbb{R}$ is Fréchet differentiable at \mathbf{u}_k and that the derivative can be computed as

$$\tilde{J}_\tau'(\mathbf{u}_k) = B_\tau^* \mathbf{p}_k + \partial_u J_\tau(\mathbf{y}_k, \mathbf{u}_k),$$

where \mathbf{p}_k solves the adjoint problem (Problem 5.4.1). The vanishing weak contact sets also yield that the critical cone $\mathcal{C}_{\mathbf{K}_{\Phi}^N}(\mathbf{y}_k, \lambda_k)$ with $\lambda_k = B_{\tau}\mathbf{u}_k + f_{\tau} - A_{\tau}\mathbf{y}$ is a linear space and coincides with the set of modified boundary conditions $\hat{\mathcal{C}}(\mathbf{y}_k)$ that was introduced in (6.3), i.e., $\mathbf{p}_k = \delta\mathbf{p}_k$ in step 2 of Algorithm 6.4, which yields the claim. \square

Again, note that Proposition 6.1.5 holds for any choice of the set $\mathcal{B}_{k,l}$ as described above.

Remark 6.1.6 *The bilinear form q on \mathbf{U}_{τ} that is required in this subsection defines a scalar product on \mathbf{U}_{τ} that reproduces the topology on \mathbf{U}_{τ} , i.e., \mathbf{U}_{τ} can equivalent be endowed with the Hilbert space structure. When \mathbf{U}_{τ} is a Hilbert space to begin with, the bilinear form q can simple be chosen as the standard scalar product on \mathbf{U}_{τ} .*

6.1.2 Step Lengths

Unlike Newton-type search directions, the search directions computed by Algorithm 6.4 do not supply a “natural” scaling. In fact, arbitrarily scaling the bilinear form $q: \mathbf{U}_{\tau} \times \mathbf{U}_{\tau} \rightarrow \mathbb{R}$ results in the same search direction but with scaled length. The procedure for choosing the step lengths in step 3 are therefore essential for the algorithms presented in Section 6.1. We will employ three procedures for the line search Algorithm 6.1 and fix one of them for the accelerated Algorithms 6.2 and 6.3 in order to keep the presentation concise. Since the approaches used to find the step lengths are well known and well documented, the following is restricted to a short overview of the algorithms and the appropriate references.

Armijo Backtracking. Backtracking is a relatively robust approach that starts out with a user supplied, positive step length that is iteratively decreased by multiplication with a (fixed) parameter until the quality of the step length is acceptable. This usually includes a *sufficient decrease* condition, such as the Armijo-Goldstein condition, cf. [152, Sec. 3.1], which can easily be modified to fit the nonsmooth setting in a straight forward manner. As most sufficient decrease conditions, the Armijo-Goldstein condition relies on a second model of the functional values that is used to predict the descent in the following step and relatively to which the actual descent should be sufficient. The condition uses the affine linear parts of the model \tilde{m}_k , which was introduced to compute the search direction, expecting that $\tilde{J}_{\tau}(\mathbf{u}_k + \delta\mathbf{u})$ can be computed by evaluating

$$\tilde{J}_{\tau}(\mathbf{u}_k + \delta\mathbf{u}) \approx \tilde{m}_k^{\Delta}(\delta\mathbf{u}) := J_{\tau}(\mathbf{y}_k, \mathbf{u}_k) + J_{\tau}'(\mathbf{y}_k, \mathbf{u}_k)(S_{\hat{c}}B_{\tau}\delta\mathbf{u}, \delta\mathbf{u}).$$

The sufficient (relative) decrease condition for a descent direction $\delta\mathbf{u}_k$ accordingly requires the step length s to satisfy

$$\eta(s) := \frac{\tilde{J}_{\tau}(\mathbf{u}_k + s\delta\mathbf{u}_k) - \tilde{J}_{\tau}(\mathbf{u}_k)}{\tilde{m}_k^{\Delta}(s\delta\mathbf{u}_k) - \tilde{J}_{\tau}(\mathbf{u}_k)} = \frac{\tilde{J}_{\tau}(\mathbf{u}_k + s\delta\mathbf{u}_k) - \tilde{J}_{\tau}(\mathbf{u}_k)}{sJ_{\tau}'(\mathbf{y}_k, \mathbf{u}_k)(S_{\hat{c}}B_{\tau}\delta\mathbf{u}_k, \delta\mathbf{u}_k)} \geq \bar{\eta}, \quad (6.8)$$

where $\bar{\eta}$ is a parameter in $(0, 1)$. I.e., the actual decrease in objective values relative to the decrease predicted by the linear model is expected to exceed a given threshold.

Algorithm 6.5 Armijo Backtracking

Require: $k \in \mathbb{N}$, $i = 0$, $\mathbf{u}_k, \delta \mathbf{u}_k \in \mathcal{U}_\tau$, $\eta(s_k^0), \bar{\eta}, \beta \in (0, 1)$, $s_k^0 > 0$

- 1: **while** $\eta(s_k^i) < \bar{\eta}$ **do**
- 2: $s_k^{i+1} = \beta s_k^i$
- 3: $i \leftarrow i + 1$
- 4: **return** $s_k = s_k^i$

Quadratic Regularization. For second order consistent search directions, an effective alternative to backtracking and computations in trust-region methods is that of cubic regularization, see [76] and the descriptions in [38, 199] and [167]. In cubic regularization, the model used to predict the objective values in the search direction consists of the second order model used to compute the search direction and an additional third order term, i.e., a cubic error model that is scaled by a procedural parameter to be determined over the course of the algorithm. The quadratic regularization approach can be seen as a lower order version of the same technique that takes the lower order consistency of the search direction into account. Assuming that the first order information included in the local model (6.4) is accurate in a sufficiently large neighborhood of the iterate, we expect the error of the model \tilde{m}_k and the reduced objective functional to be of second order due to the lack of second order information in the model. The model that is used for the computations of the predicted objective functional value (along the search direction) should therefore include a scaled second order part, i.e.,

$$\tilde{m}_k^{\Delta, \omega}(\delta \mathbf{u}) = J_\tau(\mathbf{y}_k, \mathbf{u}_k) + J_\tau'(\mathbf{y}_k, \mathbf{u}_k)(S_{\hat{c}} B_\tau \delta \mathbf{u}, \delta \mathbf{u}) + \frac{\omega}{2} q(\delta \mathbf{u}, \delta \mathbf{u}) \quad (6.9)$$

for a parameter $\omega > 0$. Proposition 6.1.4 yields that the unique minimizer to this model is $\frac{1}{\omega} \delta \mathbf{u}_k$, i.e., the previously computed search direction scaled by the step length $1/\omega$.

For each iteration $k \in \mathbb{N}$ of the outer optimization algorithm, we want to compute a parameter ω_k (and therefore the step length) that satisfies a sufficient decrease condition relative to the rescaled quadratic model, i.e., such that

$$\begin{aligned} \eta(\omega) &:= \frac{\tilde{J}_\tau(\mathbf{u}_k + \frac{1}{\omega} \delta \mathbf{u}_k) - \tilde{J}_\tau(\mathbf{u}_k)}{\tilde{m}_k^{\Delta, \omega}(\frac{1}{\omega} \delta \mathbf{u}_k) - \tilde{J}_\tau(\mathbf{u}_k)} \\ &= \frac{\tilde{J}_\tau(\mathbf{u}_k + \frac{1}{\omega} \delta \mathbf{u}_k) - \tilde{J}_\tau(\mathbf{u}_k)}{\frac{1}{\omega} J_\tau'(\mathbf{y}_k, \mathbf{u}_k)(S_{\hat{c}} B_\tau \delta \mathbf{u}_k, \delta \mathbf{u}_k) + \frac{\omega}{2} q(\frac{1}{\omega} \delta \mathbf{u}, \frac{1}{\omega} \delta \mathbf{u})} \geq \bar{\eta} \end{aligned} \quad (6.10)$$

is satisfied for an $\bar{\eta} \in (0, 1)$. Provided a user supplied initial omega $\omega_{\text{init}} > 0$ for the first update, we start out with the omega of the previous iteration and perform a series of updates of ω where the update is essentially computed from the approximation of the objective functional in (6.9). Hence, we compute

$$\omega_{\text{temp}} := 2 \frac{\tilde{J}_\tau(\mathbf{u}_k + \frac{1}{\omega_k^i} \delta \mathbf{u}_k) - \tilde{J}_\tau(\mathbf{u}_k) - \frac{1}{\omega_k^i} J_\tau'(\mathbf{y}_k, \mathbf{u}_k)(S_{\hat{c}} B_\tau \delta \mathbf{u}_k, \delta \mathbf{u}_k)}{q(\frac{1}{\omega_k^i} \delta \mathbf{u}_k, \frac{1}{\omega_k^i} \delta \mathbf{u}_k)}. \quad (6.11)$$

This basic update of ω_k^i is complemented by a simple safe-guard constraint that ensures sufficient growth of the parameter if no descent can be established by the current choice

of ω_k^i and ensures that the initial parameter for the following outer iteration is not chosen overly enthusiastically small (which would result in an unreasonably large step length). See [167, Sec. 3] for more possible modifications of the basic update.

Algorithm 6.6 Quadratic Regularization

Require: $k \in \mathbb{N}$, $i = 0$, $\mathbf{u}_k, \delta\mathbf{u}_k \in \mathcal{U}_\tau$, $\omega_0^0 > 0$, $\bar{\eta} \in (0, 1)$, $0 < \sigma_2 < 1 < \sigma_1$

- 1: **while** $\eta < \bar{\eta}$ **do**
 - 2: Compute ω_{temp}
 - 3: Update $\omega_k^{i+1} = \max(\sigma_1 \omega_k^i, \omega_{\text{temp}})$
 - 4: $i \leftarrow i + 1$
 - 5: Compute ω_{temp}
 - 6: Compute and save $\omega_{k+1}^0 = \max(\sigma_2 \omega_k^i, \omega_{\text{temp}})$
 - 7: **return** $s = \frac{1}{\omega_k^i}$
-

Note that this procedure coincides with simply rescaling the preconditioner used in the model \tilde{m}_k to compute a scaled search direction and adaptively updating the scaling parameter until the quadratic model is a sufficiently accurate approximation of the reduced objective functional. The method can therefore equivalently be thought of as a procedure for finding a reasonably scaled preconditioner to compute the search direction.

Brent’s Method. In order to find the minimum of a function on the real numbers, combining the reliability of golden section search and the efficiency of parabolic interpolation was proposed in Brent’s method, see the description in [33, Sec. 8, “localmin”]. We use Brent’s method as an approximation of an “exact” step length computation, i.e., for finding a step length

$$s_k \in \arg \min_{s>0} \tilde{J}_\tau(\mathbf{u}_k + s\delta\mathbf{u}_k)$$

for the reduced objective functional along the search direction. Since our problem is nonlinear, we can generally only expect to compute a local minimum of the objective functional in the search direction. Compared to the step length computations presented above, Brent’s method requires significantly more function evaluations, which results in a computationally expensive line search method. The specific implementation of the algorithm used was Kaskade 7.2’s *Fmin*, which is translated from the public domain Fortran routine with the same name.

Behavior of Step Lengths. Monitoring the behavior of the step lengths and the parameters ω_k over the course of the outer optimization loop is an effective way to obtain information on the quality of the search direction and the underlying local model. Both the Armijo backtracking rule and the quadratic regularization rate the quality of the model used to compute the search direction in sufficient decrease conditions, see (6.8) and (6.10). They assess the accuracy of the descent predicted by the local model compared to the actual functional value decrease in the direction of the search direction. Step lengths that tend

to be constant over parts of the run indicate that the search directions remain first order consistent during these iterations, i.e., that there are sufficiently large neighborhoods of the iterates that are free of nonsmooth effects. A decrease in step lengths, however, suggests that the step length computation model (6.7) is only first order consistent in increasingly small neighborhoods of the iterates. In the quadratic regularization approach, this effect is amplified because of the step length update in (6.11), which is directly linked to the order of the model's error. The update suggests that when the first order information included in the model at the current iterate is accurate, then the error in the numerator should actually be of order two, hence the parameter ω should be more or less stable or stay within a few orders of magnitude. When nonsmooth effects influence the vicinity of the iterate, however, then the error in the numerator of the update will be of an order lower than two, and the parameters should noticeably tend to infinity or grow by several orders of magnitude over the course of several iterations without a stabilizing behavior. Brent's algorithm, as sort of a "black-box approach", leaves less room for interpretation. Provided that the algorithm computes a reasonable local minimizer along the search direction, however, an inconsistent search direction will lead to decrease in the search directions as well.

6.2 Numerical Results

The test problems considered in the numerical examples in Sections 6.2.2 – 6.2.4 consist of different configurations of tracking-type objective functionals with final time observation, L^2 -regularization and distributed controls, i.e., problems of the following form.

Problem 6.2.1 *Given initial values $\mathbf{y}_{ini} \in \mathbf{K}_{\Phi}$, $\mathbf{v}_{ini} \in \mathbf{H}^1(\Omega)$, a desired displacement $\mathbf{y}_d: [0, T] \rightarrow \mathbf{L}^2(\Omega)$, the control space $\mathbf{U}_{\tau} = L^2(0, T; \mathbf{L}^2(\Omega))$, a mass density $\rho > 0$, the constant of approximated gravitational acceleration $c_g = 9.81 \text{ m/s}^2$ and parameters $\alpha, \beta \geq 0$, find $\mathbf{u} \in \mathbf{U}_{\tau}$ and $\mathbf{y} \in \mathbf{H}^1(\Omega)^N$ that solve*

$$\begin{aligned}
& \min \frac{1}{2} \int_0^T \|\mathbf{y}(t) - \mathbf{y}_d(t)\|_{\mathbf{L}^2(\Omega)}^2 dt + \frac{\alpha}{2} \|\mathbf{y}(T) - \mathbf{y}_d(T)\|_{\mathbf{L}^2(\Omega)}^2 + \frac{\beta}{2} \int_0^T \|\mathbf{u}(t)\|_{\mathbf{L}^2(\Omega)}^2 dt \\
& \text{s.t. } (\mathbf{y}, \mathbf{u}) \in \mathbf{Y} \times L^2(0, T; \mathbf{L}^2(\Omega)) \\
& \quad \mathbf{y} \in \mathbf{K}_{\Phi}^T, \quad \mathbf{y}(0) = \mathbf{y}_{ini}, \quad \dot{\mathbf{y}}(0) = \mathbf{v}_{ini} \\
& \quad \rho \langle \ddot{\mathbf{y}} - c_g \mathbf{u}, \mathbf{v} - \mathbf{y} \rangle_{L^2(0, T; \mathbf{H}^1(\Omega))} + a_I(\mathbf{y}, \mathbf{v} - \mathbf{y}) + b_I(\dot{\mathbf{y}}, \mathbf{v} - \mathbf{y}) \geq 0 \\
& \quad \text{for all } \mathbf{v} \in \mathbf{K}_{\Phi}^T.
\end{aligned} \tag{6.12}$$

See (4.2) for the definition of the time dependent state space \mathbf{Y} . Note that we have reintroduced the constant mass density ρ into the system, cf. Assumption 2.3.1 (b), as it clearly plays a role in the numerics. The scaling $c_g \rho$ on the right hand side of the variational inequality is the gravitational force density for the body with density ρ measured in N/m^3 and serves to "normalize" the magnitude of the controls, as we will consider problems where the involved forces are of first order relatively to the gravitational forces.

When we apply the time discretization from Section 4.1 and discretize \mathbf{u} using functions that are piecewise constant and discontinuous in time on the subintervals $I_k = (t_{k-1}, t_k]$, i.e.,

$$\mathbf{u} \in \left\{ \mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega)) : \mathbf{u}|_{I_k} \in \mathcal{P}_0(I_k, \mathbf{L}^2(\Omega)), \mathbf{u}(0) = \mathbf{u}(t_1) \right\},$$

we obtain the time-discretized control space $\mathbf{U}_\tau = \mathbf{L}^2(\Omega)^N$. Exact integration yields the form of the corresponding control-to-force operator

$$B_\tau: \mathbf{L}^2(\Omega)^N \rightarrow (\mathbf{H}^1(\Omega)^N)^*, \quad \langle B_\tau \mathbf{u}, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} = \tau c_g \rho \sum_{k=1}^N (\mathbf{u}_k, \mathbf{p}_k)_{\mathbf{L}^2(\Omega)}. \quad (6.13)$$

Depending on the application, the discretized objective functional can now either be computed exactly or determined using an approximation. We will employ the trapezoidal rule in the following, which yields

$$J_\tau: \mathbf{H}^1(\Omega)^N \times \mathbf{L}^2(\Omega)^N \quad (6.14)$$

$$J_\tau(\mathbf{y}, \mathbf{u}) = \frac{\tau}{2} \sum_{k=1}^{N-1} \|\mathbf{y}_k - \mathbf{y}_{d,k}\|_{\mathbf{L}^2(\Omega)}^2 + \frac{\tau + 2\alpha}{4} \|\mathbf{y}_N - \mathbf{y}_{d,N}\|_{\mathbf{L}^2(\Omega)}^2 + \frac{\tau\beta}{2} \sum_{k=1}^N \|\mathbf{u}_k\|_{\mathbf{L}^2(\Omega)}^2.$$

Note that the part corresponding to \mathbf{y}_0 can be disregarded in the minimization problem due to the fixed initial value of the displacements \mathbf{y} . For this setting, we can easily verify the assumptions made in Section 5.2 for the existence of minimizers of the discretized problem.

Corollary 6.2.2 *Let Assumption 4.0.1 be satisfied and $\mathbf{U}_\tau, B_\tau, J_\tau$ be given as in (6.13)-(6.14). Then Problem 5.1.2 has at least one solution.*

Proof. Due to the quadratic tracking and regularization parts, coercivity as in (5.2) is guaranteed, and the functional is bounded from below by zero, convex and continuous and therefore weakly lower semi-continuous [20, Prop. 2.10]. The operator

$$\langle B_\tau \mathbf{u}, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} = \tau c_g \rho \sum_{k=1}^N (\mathbf{u}_k, \mathbf{p}_k)_{\mathbf{L}^2(\Omega)}$$

corresponds to the (scaled) composition of the Riesz isomorphism in the Hilbert space $\mathbf{L}^2(\Omega)^N$ and the embedding of $(\mathbf{L}^2(\Omega)^N)^*$ into $(\mathbf{H}^1(\Omega)^N)^*$. The isomorphism maps weakly convergent subsequences in $\mathbf{L}^2(\Omega)^N$ to weakly convergent subsequences in $(\mathbf{L}^2(\Omega)^N)^*$. Rellich-Kondrachov's Theorem (see [3, Thm. 6.4]) ensures that the standard embedding $\mathbf{H}^1(\Omega)^N \hookrightarrow \mathbf{L}^2(\Omega)^N$ is compact, and therefore application of Schauder's Theorem (see [119, Thm. 8.2-5]) yields compactness of the embedding $(\mathbf{L}^2(\Omega)^N)^* \hookrightarrow (\mathbf{H}^1(\Omega)^N)^*$ and ultimately compactness of B_τ . Hence, Theorem 5.2.1 can be applied. \square

Additionally, the density of the distributed controls in the dual of the state space ensures that the first order optimality conditions hold at the respective minimizers.

Corollary 6.2.3 *Let Assumption 4.0.1 be satisfied, $\mathbf{U}_\tau, B_\tau, J_\tau$ be given as in (6.13)-(6.14), and let $(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ be a local minimizer of Problem 5.1.2. Then there exist multipliers $\mathbf{p} \in \mathbf{H}^1(\Omega)^N$ and $\mu \in (\mathbf{H}^1(\Omega)^N)^*$ such that the first order optimality condition (5.4a)-(5.4b) is satisfied.*

Proof. Clearly, the tracking-type objective functional J_τ in (6.14) is Fréchet differentiable. Therefore, the only assumption for Theorem 5.3.1 that remains to be checked is density of the image of $B_\tau: \mathbf{L}^2(\Omega)^N \rightarrow (\mathbf{H}^1(\Omega)^N)^*$, where

$$\langle B_\tau \mathbf{u}, \mathbf{p} \rangle_{\mathbf{H}^1(\Omega)^N} = \tau c_g \rho \sum_{k=1}^N (\mathbf{u}_k, \mathbf{p}_k)_{\mathbf{L}^2(\Omega)}.$$

Recall that $B(\mathbf{U}_\tau) = (\mathbf{L}^2(\Omega)^N)^* \hat{=} (\mathbf{L}^2(\Omega)^*)^N$ and that the operator $E: \mathbf{H}^1(\Omega)^N \rightarrow \mathbf{L}^2(\Omega)^N$ has trivial kernel. By identification of $\mathbf{H}^1(\Omega)^N$ with its bidual, we obtain that $E^{**}: (\mathbf{H}^1(\Omega)^N)^{**} \rightarrow (\mathbf{L}^2(\Omega)^N)^{**}$ has trivial kernel as well. Due to the closed range theorem [200, Thm. III.4.5] in application to the adjoint operator $E^*: (\mathbf{L}^2(\Omega)^N)^* \rightarrow (\mathbf{H}^1(\Omega)^N)^*$, we obtain that

$$\overline{\text{im } E^*} = (\ker E^{**})^\perp,$$

with

$$\begin{aligned} \overline{\text{im } E^*} &= \overline{(\mathbf{L}^2(\Omega)^N)^*}, \\ \ker E^{**} &= \{0\}_{(\mathbf{H}^1(\Omega)^N)^{**}}, \end{aligned}$$

which yields the claim. \square

Accordingly, the optimal control problem (6.12) fits into the framework we have considered in the previous chapters, making it an appropriate setting to be considered in the numerical computations. The properties of the problems in the setting described above can still vary quite significantly. The specific choice of the parameters, initial values, desired state, etc. influences the characteristics of the test problem considerably, as we will see in the application of the algorithms described above for the examples in the following sections.

6.2.1 Implementation Details

For all computations, a *plane strain* state is assumed, see Section 2.1. Accordingly, the computations are carried out for a model in two spatial dimensions. We identify the three-dimensional domains with their two-dimensional cutting plane representatives in the description for simplicity.

The optimization algorithms described in Section 6.1 were implemented in C++, and the contact problems were discretized using the time stepping scheme described in Section 4.1 and piecewise linear continuous finite element functions on triangular meshes for the spatial discretization of the control and the states. The resulting mass matrices were lumped. (See [209, Sec. 17.2, App. I] for an introduction into mass lumping.) Seeing as the controls are in $\mathbf{L}^2(\Omega)^N$, the bilinear form q in the implementation of the algorithm was chosen as the $\mathbf{L}^2(\Omega)^N$ -scalar product.

For the numerical realization, the finite element toolbox Kaskade 7.2 [73] and the underlying Distributed Unified Numerics Environment DUNE 2.3.1 [21, 22, 23] were employed. The discretized complementarity problems arising in each of the time steps were solved up to

a relative tolerance of 10^{-9} using the monotone multigrid solver provided by the *dune-solvers* module and the projected block Gauss Seidel (PBGS) solver as the smoother and base solver. For a description of the monotone multigrid solver, see [75]. The influence of the forward solver tolerance is shortly addressed in Section C.5.1 of the appendix. The computations were performed on an Intel Xeon E3-1271 at 3.60 GHz with 32 GB of memory under Ubuntu 14.04.5.

The parameters of the material model for the viscoelastic bodies in the examples were chosen to model a rubber-like body with moderate viscous damping. Table 6.1 lists the common parameters of the examples.

ρ (kg/m ³)	E (Pa)	ν_{poi}	μ_{bulk} (Pa s)	μ_{shear} (Pa s)	Multigrid Tolerance	PBGS Tolerance
10^3	10^7	0.45	50	50	10^{-9}	10^{-7}

Table 6.1: Parameters for the numerical examples.

Remark 6.2.4 *The material was chosen as rubber-like with $\nu_{\text{poi}} = 0.45$ instead of assuming less compressability in order to avoid the degenerative effects associated with the employed finite element approximation for ν_{poi} close to 0.5. As $\nu_{\text{poi}} \rightarrow 0.5$, non-robust finite element approximations tend to show an overly stiff response and a corresponding deterioration of the accuracy of the approximation. These effects appear in a number of problems and discretization scheme constellations and are known as locking. See [209, Sec. 11.3.2], [15, 16, 56] and [24, P. 279 ff.] for more information on the topic.*

6.2.2 Example 1: A Well-Behaved Inverse Problem

As we have seen in the previous chapters, nonsmooth effects arise in the contact problem when the body is in contact with the obstacle and the biactive set has positive capacity, i.e., when the body and the obstacle are in contact and sufficiently large sections of the contact patch are not subjected to contact forces that separate the body and the obstacle. Intuitively, constellations where the contact boundaries of the body and the obstacle are dissimilar are less prone to introducing these nonsmooth effects into the optimization problem. In this first example, we examine the regularized inverse problem of tracking the trajectory of a body with a curved boundary falling onto a planar obstacle due to gravitational forces. We will see that the proposed algorithms' performances on these well behaved contact problems is comparable to their performances on smooth problems.

Reference Configuration. Consider a viscoelastic, rubber-like body with the material parameters as listed in Table 6.1 and with a reference configuration in the shape of a semicircle with radius 1 m,

$$\begin{aligned}\Omega &:= \{ \mathbf{x} \in \mathbb{R}^2 : x_1^2 + (x_2 - 1.05)^2 < 1 \wedge x_2 < 1.05 \}, \\ \mathcal{O} &:= \{ \mathbf{x} \in \mathbb{R}^2 : x_2 < 0 \},\end{aligned}$$

which corresponds to the base of a long semicylinder positioned 0.05 m above a planar, rigid obstacle. The contact boundary is taken as the middle third of the curved boundary

section, and free boundary conditions are assumed on the remaining boundary sections. As discussed in Section 2.2.2, the contact normal is chosen according to the normal of the obstacle. We set vanishing initial values and fix a time span of 0.2 seconds, which allows the body to fully release after rebounding off the obstacle when subjected to gravitational pull.

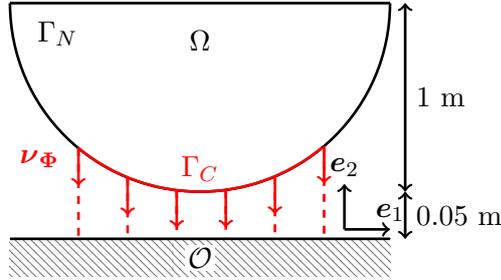


Figure 6.1: Reference configuration in Example 1.

Objective Functional. The desired state \mathbf{y}_d is taken as the result of running the forward simulation for the discretized contact problem with vanishing initial values and a force density of

$$\mathbf{f}_g := \begin{pmatrix} 0 \\ -9.81 \cdot 10^3 \end{pmatrix} \text{N/m}^3,$$

i.e., the state \mathbf{y}_d resembles the displacement generated by an approximation of the gravitational pull on the earth's surface. Recall that the control that corresponds to this right hand side is $\mathbf{u} = (0, -1)$ due to the scaling in the control-to-force operator, see (6.13).

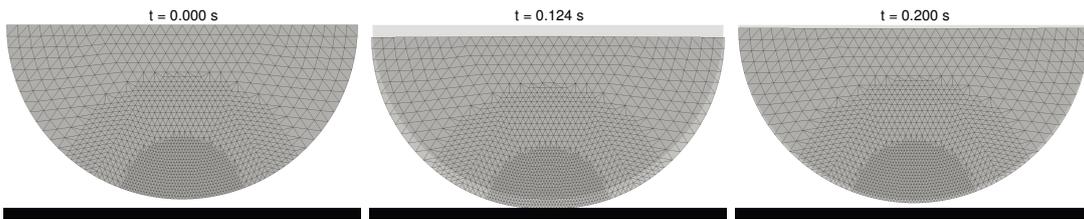


Figure 6.2: Deformation of the desired state. The reference configuration is shaded in light gray.

The position of the semi-circular viscoelastic body that corresponds to the desired state is depicted in Figure 6.2 for the initial time, the time where the contact patch is largest and for the final time. Figure 6.3 shows the behavior of the point of the semi-circle that is in contact with the obstacle the longest (the “south pole”) and the two points of the contact patch that are in contact for the shortest amount of time (the far ends of the active contact patch). The left hand axis corresponds to the vertical component of the positions and the right hand axis is for the vertical component of the contact force density.

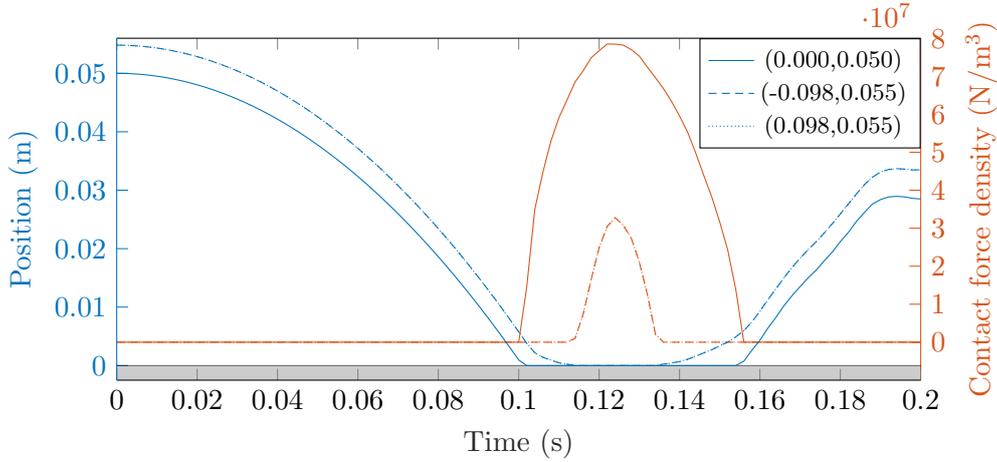


Figure 6.3: Vertical components of position and contact force density of three points on the contact patch over time. The legend shows the coordinates of the points in the reference configuration and gray indicates the obstacle.

Note that until contact is established at $t_{\text{con}} := \sqrt{1/10c_g} \text{s} \approx 0.10096 \text{s}$, the trajectory of the desired state can simply be computed by the standard parabola of an object under constant acceleration. Since t_{con} is in the interior of the temporal intervals of discretization, the time of contact is not accurately resolved by the time grid, i.e., contact is recognized at $t = 1.002 \text{s}$.

T (s)	τ (s)	\mathbf{y}_{ini} (m)	\mathbf{v}_{ini} (m/s)	$\Psi(\mathbf{x})$ (m)	ν_{Φ}	\mathbf{y}_d (m)
0.2	$2 \cdot 10^{-3}$	(0, 0)	(0, 0)	x_2	(0, -1)	$S(\mathbf{f}_g)$

Table 6.2: Specific parameters of Ex. 1.

Solver Behavior. For the examination of the solver behavior, we fix the final time observation parameter $\alpha = 0$, the regularization parameter $\beta = 10^{-4}$ and constant zero as the initial control. The behavior is consistent for different choices of these parameters, and the aforementioned setting is merely chosen as a representative for the possible combinations. We will now take a closer look at the algorithmic parameters of the respective optimization algorithms. All algorithms were terminated when an iterate was reached where no step length yielded a strict decrease in the functional value, i.e., until either the tolerance of the forward solver would have needed to be decreased or the computed search direction is not a descent direction.

The left hand side of Figure 6.4 shows the difference of the objective functional values at each iteration and the final functional value of the respective optimization algorithm. Recall that the difference between the functional value at each iteration and the minimal functional value coincides with the energy norm for well-posed quadratic problems.

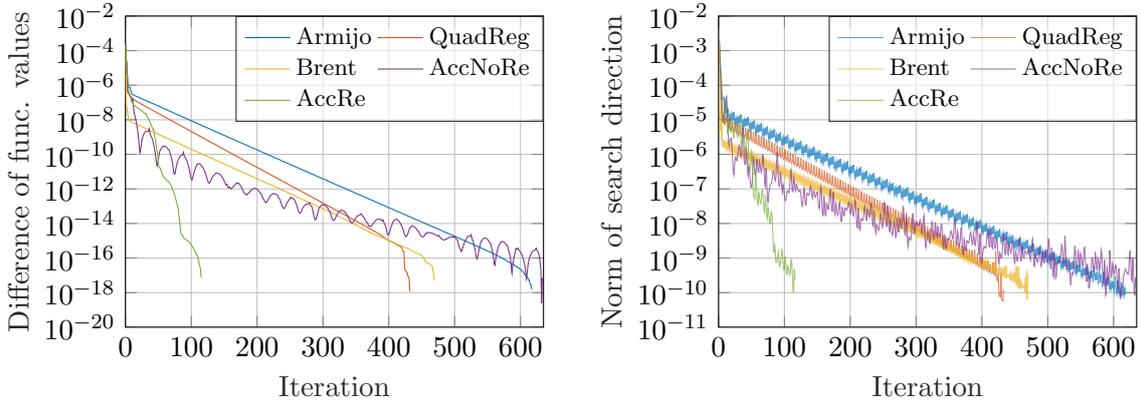


Figure 6.4: The difference of the current functional values and the final functional values of each algorithm (left) and the $L^2(\Omega)^N$ -norm of the search direction (right) over the iterations.

The three line search methods with the standard (non-accelerated) scheme all achieve rapid decrease in the functional values over the first iterations while exhibiting linear “convergence” in the log-plot of the functional values in the long run. We observe that Brent’s method — as an approximation of the exact step length — allows for the fastest descent in the first couple of iterations but shows relatively slow convergence in the functional values as the iterations progress. The Armijo rule and the quadratic regularization method are both strongly parameter dependent step length computation algorithms, hence the first iterations yield noticeably smaller descent in the objective values than Brent’s method. However, later on the objective functional values of the former decrease faster than those of the latter. The behavior of the accelerated momentum methods is quite different from the standard scheme, where the non-restarted version shows the characteristic ripples and the overall tendency that the decrease in functional values slows down when getting closer to the final control. The non-restarted algorithm performs the first ascent step in iteration eleven, which is where the momentum of the restarted scheme is reset. Restarting the momentum slows down the decrease in functional values between iterations eleven and 50, while convergence closer to the final control is significantly faster compared to the non-restarted algorithm or the standard scheme with any of the step length computations. The restarted algorithm uses the initial acceleration of the momentum steps while the functional values descent providing the fastest functional value decrease.

On the right hand side of Figure 6.4 the norm of the (unscaled) search direction over the iterations shows an overall linear decrease in the log-plot and the well-known zig-zagging effects that the standard gradient schemes exhibits in narrow valleys, cf., e.g., the Rosenbrock functional. Recall that the search direction coincides with the gradient of the reduced objective functional at every iterate whose state does not feature biactive sets at any point in time. In that case, the adjoint problem is uniquely solvable, see Corollary 5.4.6, hence the search direction also coincides with the residual of the stationarity condition, see (5.4a)–(5.4b). Assuming no weak contact occurs for the final iterate, the first order optimality condition is therefore satisfied up to a tolerance of approximately 10^{-10} in the norm as the algorithms terminate.

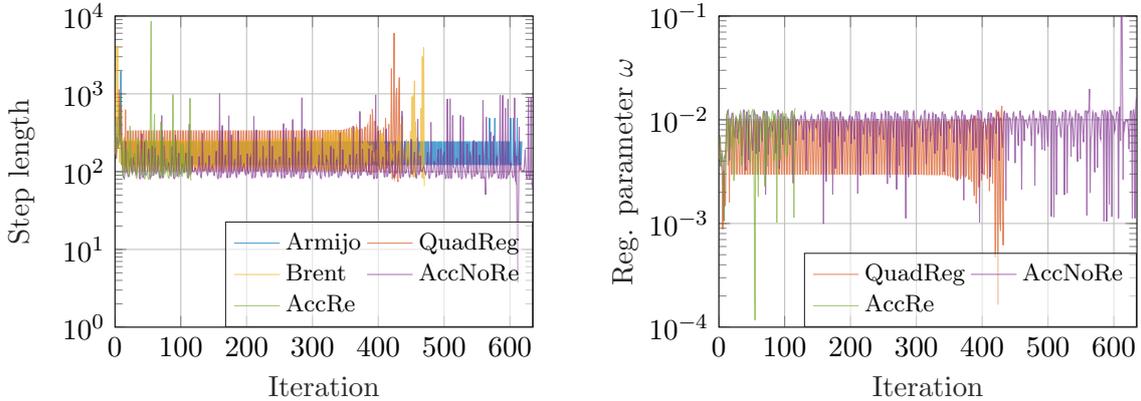


Figure 6.5: The computed step lengths for all algorithms (left) and the regularization parameters ω for the three optimization algorithms that use quadratic regularization (right).

The computed step lengths are shown on the left hand side of Figure 6.5, and the right hand side shows the corresponding parameters ω for the algorithms that use the quadratic regularization scheme. Recall that the step length s is given by $s = 1/\omega$, and note that while the values fluctuate rapidly — which is another effect that can be attributed to gradient-type zig-zagging — the overall tendency of the step lengths and the parameter ω is constant, and all five algorithms use step lengths that stay within a range of up to half an order of magnitude, depending on the step length method. This indicates that the chosen search direction provides reliable first order information for the functional in this problem. Furthermore, the step lengths show that the improved initial decrease of the functional values in Brent’s algorithm over the Armijo scheme and the quadratic regularization is due to the fact that the step length in the initial iterations is chosen larger than in the respective alternatives.

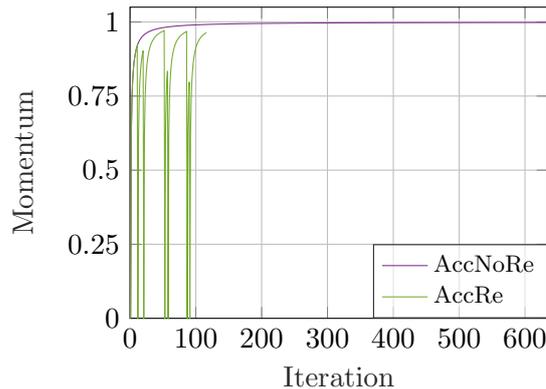


Figure 6.6: The momentum of the accelerated schemes over the iterations.

Figure 6.6 shows the momentum of the accelerated schemes over the iterations. The momentum terms of the non-restarted version converge to one in a predetermined manner, see Algorithm 6.3 steps 5 and 6. Note that the momentum grows quite quickly such that the restarts of the momentum in the restarted scheme, which are executed every three to 30 iterations in this example, yield momentum terms of 0.79 to 0.97 in the restarted scheme.

Discussion of the Solution. All five algorithms compute the same solution up to a reasonable tolerance, as Table 6.3 below shows. The table shows maximum, minimum and mean of the $L^2(0, T; \mathbf{L}^2(\Omega))$ -norms of the differences in the controls and states and of the absolute value of the difference in functional values — both absolutely and relatively — taken pairwise for all algorithms. As we can see, the error is largest in the controls, while the functional values coincide up to numerical accuracy.

	Absolute norm of differences			Relative norm of differences		
	Max	Min	Mean	Max	Min	Mean
Controls	$6.7 \cdot 10^{-7}$	$1.4 \cdot 10^{-7}$	$4.5 \cdot 10^{-7}$	$2.0 \cdot 10^{-6}$	$4.4 \cdot 10^{-7}$	$1.4 \cdot 10^{-6}$
States	$2.0 \cdot 10^{-9}$	$8.1 \cdot 10^{-10}$	$1.4 \cdot 10^{-9}$	$9.5 \cdot 10^{-8}$	$3.9 \cdot 10^{-8}$	$6.8 \cdot 10^{-8}$
Values	$1.4 \cdot 10^{-17}$	$1.5 \cdot 10^{-18}$	$6.6 \cdot 10^{-18}$	$2.2 \cdot 10^{-12}$	$2.4 \cdot 10^{-13}$	$1.1 \cdot 10^{-12}$

Table 6.3: The $L^2(0, T; \mathbf{L}^2(\Omega))$ -norms of the differences of the computed solution controls and solution states and the differences of the corresponding objective values taken pairwise for the five compared algorithms. The relative values are taken relatively to the smaller of the two respective values.

Figure 6.7 shows the vertical position of the semicircle’s “south pole” and the vertical component of the contact force density, both for the computed solution state and desired state for the parameters set above.

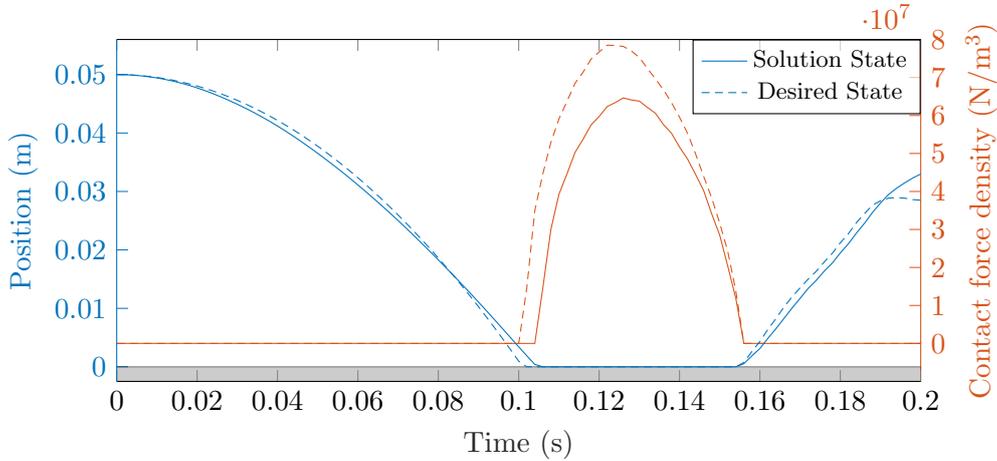


Figure 6.7: Vertical component of the position and the contact force density corresponding to the south pole for desired and solution state over time. Gray indicates the obstacle.

The computed solution initially accelerates towards the obstacle faster than the desired state, but the acceleration is reduced before the contact phase, such that contact is established at a later time. During the contact phase, the contact forces are smaller than those corresponding to the desired state. While the contact phase ends at the same time that it does for the desired state, the computed solution is not accelerated towards the obstacle at the end of the contact phase and the time interval as strongly as the desired state.

Of course, the behavior of the solution is influenced by the choice of parameters, and we will outline some of the characteristics of the solutions for different parameter sets in this paragraph. As we can see on the left hand side of Figure 6.8, increasing the final time observation parameter to values $\alpha > 0$ causes the differences of the computed solution state and the desired state to decrease at the end of the time interval, as expected.

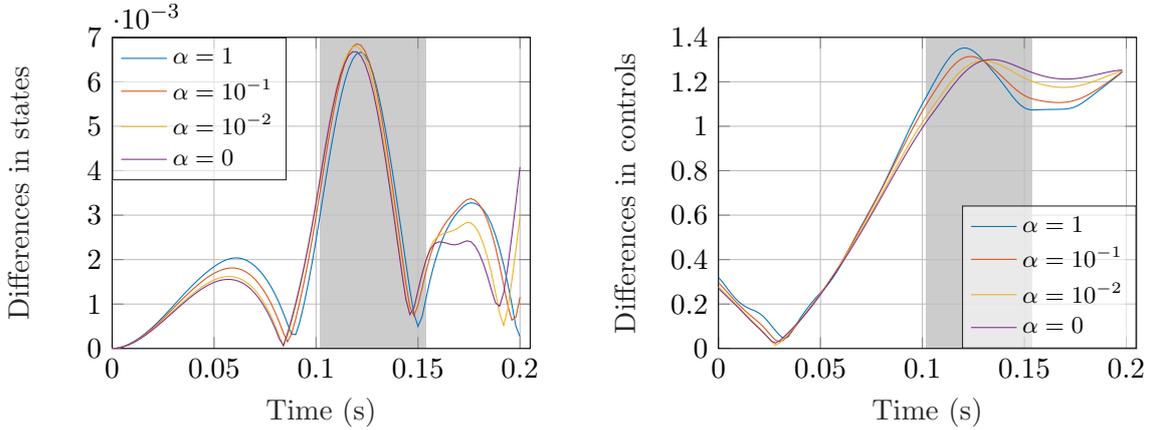


Figure 6.8: The $L^2(\Omega)$ -norm of the differences of the computed solution state and the desired state (left) and the computed solution control and the control associated with the gravitational pull used to compute the desired state (right) for several values of the final time observation scaling α and regularization $\beta = 10^{-4}$. Time of contact of the desired state is shaded.

The difference of the states exhibits three characteristic local maximums — qualitatively independently of the final time scaling — and the largest difference in the states can be observed during the time of contact. The local minima coincide to the points in time when the desired state catches up to the solution (and vice versa) after the acceleration phases and the contact phase, cf. Figure 6.7. The $L^2(\Omega)$ -norm of the difference of the solution control and the control associated with the gravitational forces \mathbf{f}_g are shown on the right hand side of the same figure, where the difference in the controls has its characteristic maximum at the time of contact as well. This effect can be observed because the amount of external forces enters the objective functional in the regularization term and the external forces acting towards the obstacle during the time of contact are met by corresponding reactive contact forces introduced by the obstacle that restrict the deformation of the body, making it uneconomic to invest forces during the time of contact. We will encounter the same effect again later on. Note that the final time observation parameter influences the amount of the control input after the detachment of the body from the obstacle.

The plots in Figure 6.9 show the behavior of controls and states in the $L^2(\Omega)$ -norm over time depending on the regularization parameter β for two fixed values of the final time scaling α . As observed above, the difference in the solution state and the desired state over time in the top line are quite similar for both values of the final time scaling. As expected, with β being the parameter that determines the amount of cost that an active control introduces into the objective functional, the difference in the states decreases with β , while the three characteristic phases of maximum difference remain. The center and the

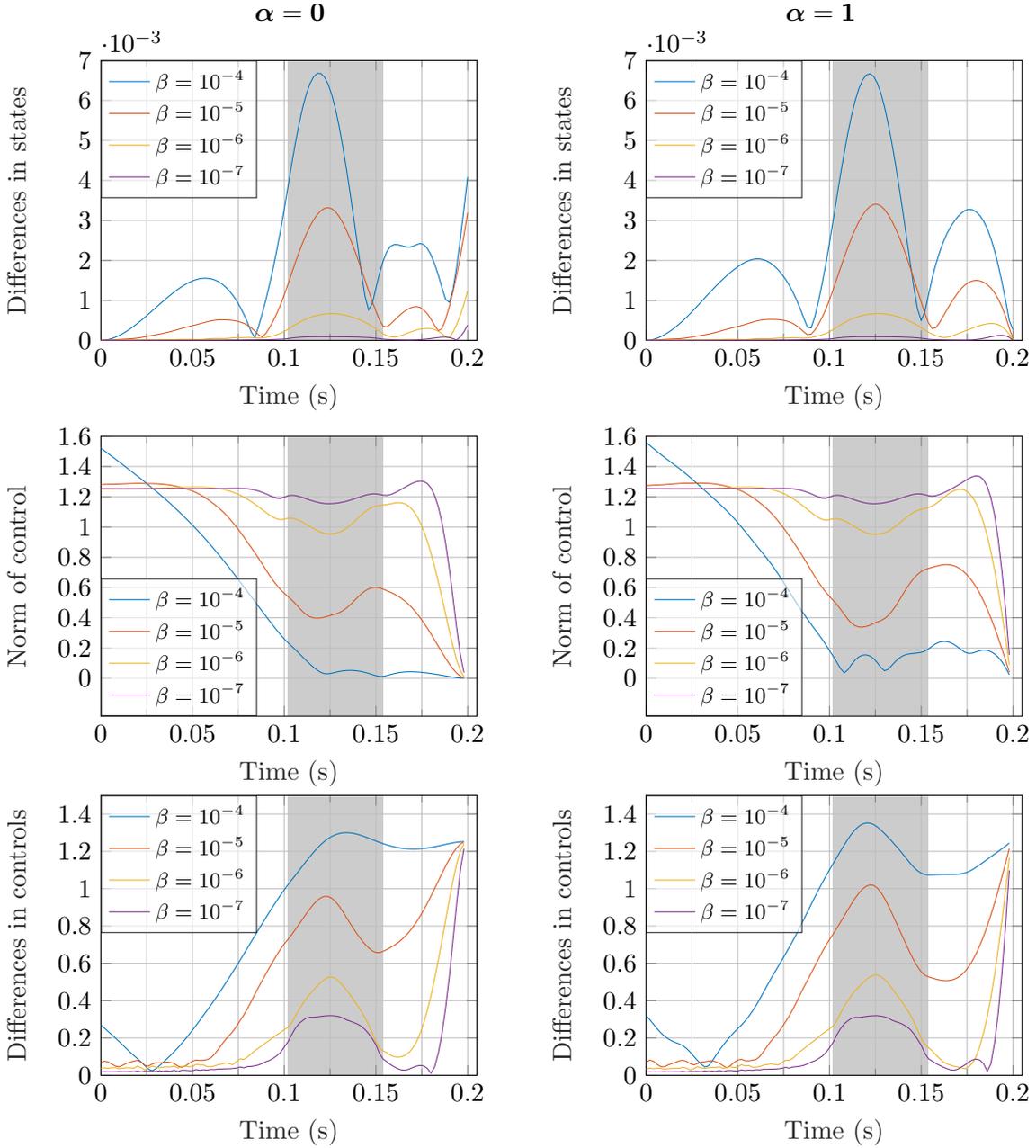


Figure 6.9: The $L^2(\Omega)$ -norm of the differences of the computed solution state and the desired state (top), the norm of the computed solution control (center) and the norm of the difference of the computed solution control and the gravitational pull used to compute the desired state (bottom) over time for final time observation values $\alpha = 0$ (left) and $\alpha = 1$ (right). Time of contact of the desired state is shaded.

bottom line show the behavior of the controls, and we observe that $\beta = 10^{-4}$ is the smallest value for which the resulting solution forces initially surpasses the gravitational forces, while for all other β , the norm of the forces remains smaller than that of the gravitational forces. For all β , the forces are noticeably reduced at the time of contact, which is when the contact forces introduced by the obstacle additionally influence the displacements. With decreasing β , i.e., reduced cost for controls, the restriction of the forces during the time of contact becomes less distinct and the forces increase more strongly after detachment. At the final time $T = 0.2$ s, all forces are significantly reduced. Note that the characteristics of the control's behavior are independent of whether final time observation is present, but after the detachment phase, the time of maximum norm of the forces shifts to later times when final time observation is introduced.

As these previous plots show, controls (external forces) are decreased during the contact phase, since application of forces in the direction of the obstacle causes corresponding contact forces to counteract. With decreasing cost for the controls, we have seen that the decrease of external forces during the contact phase becomes less noticeable in the norm. Figure 6.10 shows the deformation of the body associated with the solution at $t = 0.134$ s for no final time observation and the various values of the regularization β as above with the body colored by the magnitude of the controls. We can see that the decrease in cost for the controls not only leads to less decrease in the external forces at the time of contact, but that the decrease of the external forces locally concentrates around the contact patch, which is where the contact forces take effect.

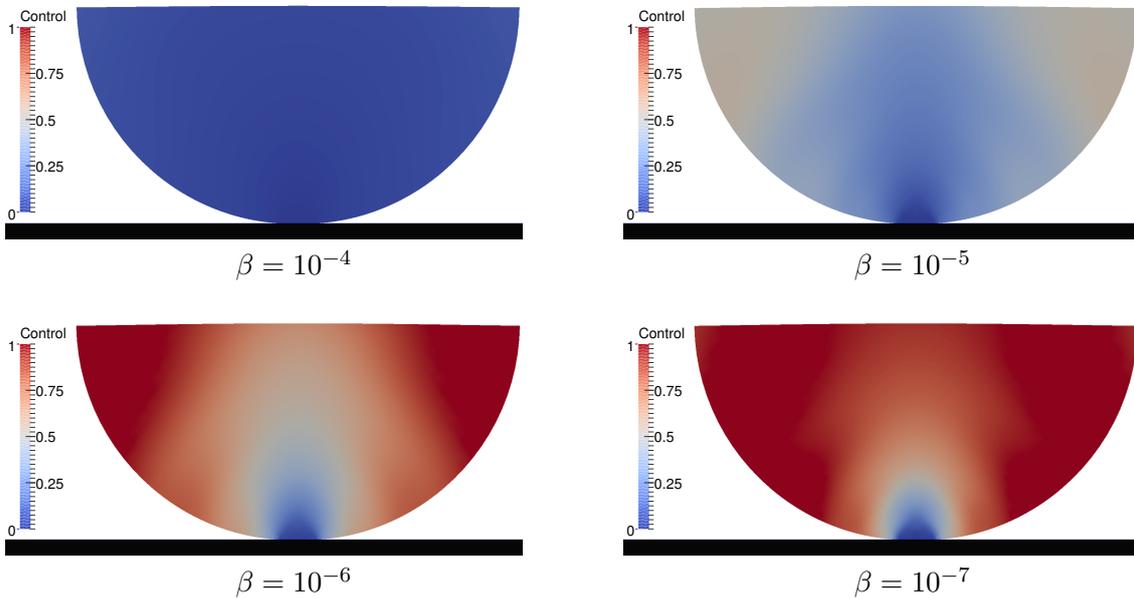


Figure 6.10: The body corresponding to the computed solution at $t = 0.134$ s colored by the magnitude of the solution controls for different regularization parameters and final time observation parameter $\alpha = 0$.

Finally, we can observe that the reduction of the external forces during the contact phase can be observed on a longer time interval of $T = 0.35$ s as well, but the controls reengage after the body detaches for the first time until the second contact phase is established.

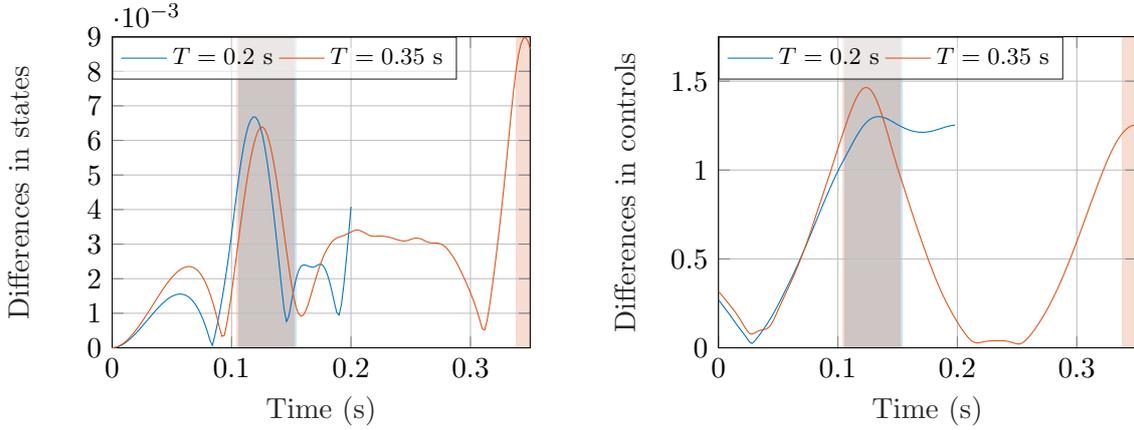


Figure 6.11: The $L^2(\Omega)$ -norm of the differences of the computed solution state and the desired state (left) and the computed solution control and the gravitational pull used to compute the desired state (right) for two time horizons and $\alpha = 0$, $\beta = 10^{-4}$. Time of contact of the computed solutions is shaded.

6.2.3 Example 2: A Nonsmooth Positioning Problem

Both the behavior of the optimization algorithms and the characteristics of the computed solution in the previous example are in accordance with the behavior observed in smooth problems, hence the problem can reasonably be deemed unaffected by nondifferentiability effects. In this second example, we construct a setting where the nonsmoothness introduced by biactive sets clearly causes the progress of the optimizers to deteriorate. Recall that in order for nondifferentiability to affect the optimization schemes, it is not necessary for any iterate to in fact be a point of non-Fréchet differentiability, but if the iterates are sufficiently “close” to such points, in the sense that the model used to compute the search direction by Algorithm 6.4 is no longer first order consistent in a sufficiently large neighborhood of the iterate. We consider a reference configuration and an obstacle with parallel, planar contact faces, and we use a static desired state that is in weak contact with the obstacle in order to incentivize the optimizers to position the body in such a biactive state.

Reference Configuration. Again, assume the body to consist of the viscoelastic material described in Table 6.1, but consider a square reference configuration with sides of length 1 m positioned 0.05 m above a planar, rigid obstacle, i.e.,

$$\begin{aligned}\Omega &:= \{ \mathbf{x} \in \mathbb{R}^2 : \max(|x_1|, |x_2 - 0.55|) < 0.5 \}, \\ \mathcal{O} &:= \{ \mathbf{x} \in \mathbb{R}^2 : x_2 < 0 \},\end{aligned}$$

which corresponds to the three-dimensional setting of a cuboid above a plane. The contact boundary is taken as the bottom face of the square and free boundary conditions are assumed on the remaining boundary segments. With both contact surfaces being planar and parallel, the contact normal is chosen vertically and is parallel to the geometric normals of both the surface of the body and the obstacle, see Figure 6.12. Initial values for the state are set to vanish, and we consider a time span of 0.125 s.

Objective Functional. The desired state in the objective functional corresponds to a shift of the body downwards by the 0.05 m of the initial gap such that contact between the body and the obstacle is established, and the desired state is assumed constant over time.

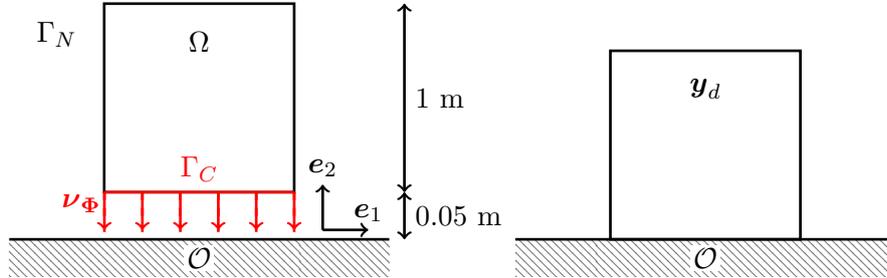


Figure 6.12: Reference configuration (left) and desired state (right) for the positioning problem in Example 2.

T (s)	τ (s)	\mathbf{y}_{ini} (m)	\mathbf{v}_{ini} (m/s)	$\Psi(\mathbf{x})$ (m)	ν_{Φ}	\mathbf{y}_d (m)
0.125	$1.25 \cdot 10^{-3}$	(0, 0)	(0, 0)	x_2	(0, -1)	(0, -0.05)

Table 6.4: Specific parameters of Example 2.

As mentioned in Section 6.2, the objective functional consists of a part that penalizes the difference of the states and a second part that penalizes the magnitude of the external forces. Hence, for this specific configuration, the objective functional values decrease when the body is brought into contact with the obstacle with minimal use of external forces. Accordingly, a setting where the body is stably in contact with the obstacle with no remaining momentum and no additional external forces is pursued. In the absence of external forces or momentum in the body, the contact forces vanish as well, leading to a state of weak contact for at least part of the time interval. While it is unlikely that such a biactive state is reached exactly at an iterate due to the momentum of the body, we can at least expect that the iterates' proximity to such states suffices to corrupt the consistency of the search direction. We can therefore expect the behavior of the optimizers to differ from the behavior in the previous example in a more erratic manner. The results will again be strongly parameter dependent, and we will address the final time observation parameters $\alpha = 0, 1$ and 10 . Recall that an increase of α promotes seeking weak contact between the body and the obstacle towards the end of the time interval, increasing the advantages of states in weak contact and therefore a more noticeable effect on the solvers.

Solver Behavior. As in Example 1, the algorithms were terminated at iterates where no step length yielded a strict decrease in the functional value. Seeing as the configuration was explicitly chosen to challenge the optimizers, an additional maximum of 15000 iterations was enforced, a number that allows to obtain a good read on the behavior on the algorithmic parameters while limiting the computation times in the event that the iterates do not converge. Computation times for the computationally most expensive algorithm — Brent's line search for $\alpha = 10$ — were well in excess of a week using the hardware described above.

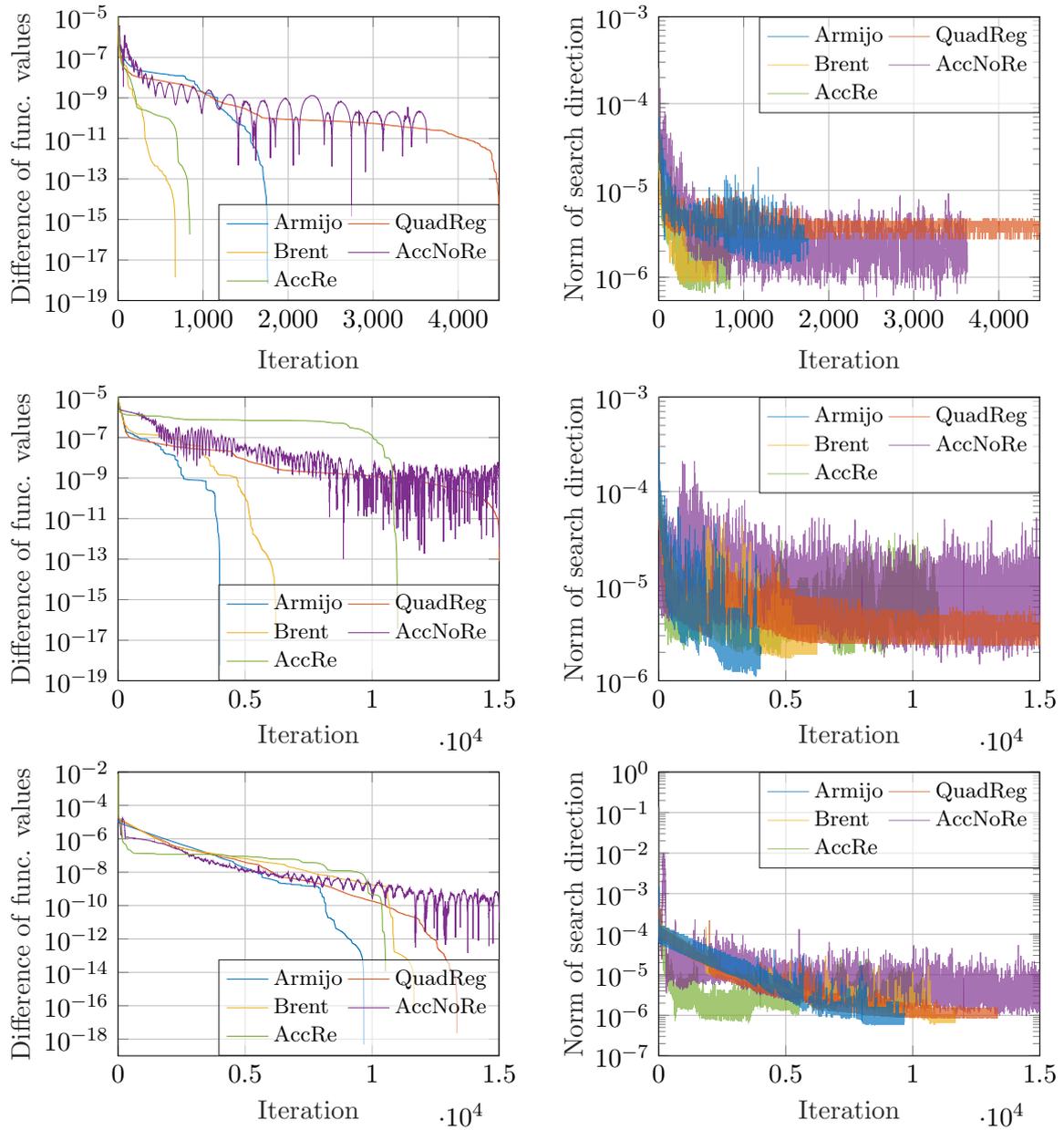


Figure 6.13: The difference of the current functional values and the final functional values of each algorithm (left) and the $L^2(\Omega)^N$ -norm of the search direction (right) over the iterations for $\alpha = 0, 1, 10$ (top to bottom).

The left hand side of Figure 6.13 shows the difference of the objective functional values at each iteration and the final functional value of the respective optimization algorithm for the final time observation parameters α equal to 0, 1 and 10 (top to bottom). Due to the decrease in regularity of the problem that comes with an increase in the final time observation parameter α , it is unsurprising that the algorithms generally require more iterations to terminate when α is increased. The decrease in the objective functional can be observed to behave quite differently from the decrease in Example 1, cf. the corresponding figure (Figure 6.4). Particularly, the behavior of most algorithms is less consistent. We can observe that the decrease in the functional values' differences for the standard line search method exhibits plateaus, see the plots for α equal to 1 and 10, especially for Brent's method and the Armijo rule, instead of showing a consistent, linear decrease. At the beginning of each run, the behavior of the function values for the algorithms is noticeably different with respect to steepness of the descent, i.e., the initial progress compared to the progress over the remainder of the iterations is of different quality for different algorithms. Both accelerated schemes show the rippling structure that could be observed in Example 1. The frequency of the ripples can be observed to increase when α is increased, owing to the reduced regularity of the problems. Note that the non-restarted accelerated algorithm reaches the user-specified 15000 iteration threshold for $\alpha = 1, 10$, and the line search method with the quadratic regularization does so for $\alpha = 1$ but not for $\alpha = 10$.

The right hand side of the figure shows the norm of the (unscaled) search direction over the iterations for the same final time observation parameters. Again, we do not observe the clear tendency for linear decrease that was encountered in Example 1, see the corresponding figure (Figure 6.4). The norms for the search directions in Figure 6.13 tend to stagnate at or near the value 10^{-6} , all of them after an initial phase of decrease in norms in the beginning of the run. The phase of initial decrease extends, relatively to the total number of iterations needed, as the final time observation increases. This indicates that the general region of the problem where the norm of the search directions is not reduced further is located close to the final iterate, and it is of course reached later in the run for the less regular problem configurations. The search directions' norms for the Armijo rule and the quadratic regularization algorithm for $\alpha = 10$ are especially noteworthy. The norms in the log-plots appear to decrease in an approximately linear manner — as they have done in Example 1 — before the decreasing tendency disappears when the respective threshold is reached, which indicates “regular” behavior until the second phase of the run is entered. Comparing with the functional values on the left hand side of the figure, we can observe that the iterations which introduce a change in the behavior of the objective values and the norm of the search direction generally coincide — see, e.g., the plots for the Armijo rule step length computation at 8250 iterations and the quadratic regularization approach at 11866 iterations for $\alpha = 10$ in the bottom right hand side of the figure. While the optimization algorithm is terminated relatively quickly after reaching the respective thresholds when the Armijo rule or Brent's algorithm is used, the algorithm that use the quadratic regularization tend to run on for several additional iterations.

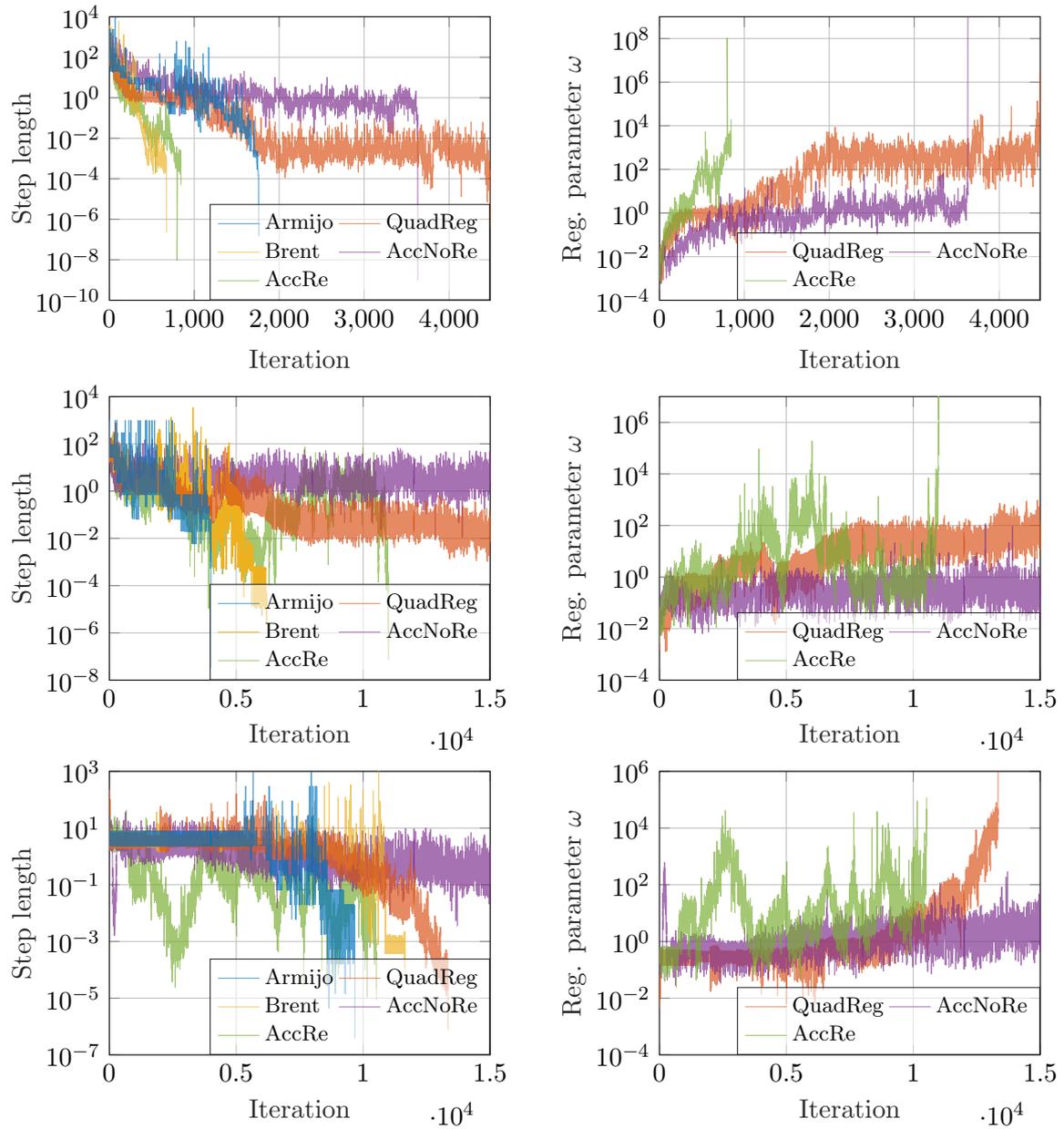


Figure 6.14: The computed step lengths for all algorithms (left) and the regularization parameters ω for the three optimization algorithms that use quadratic regularization (right) for $\alpha = 0, 1, 10$ (top to bottom).

Figure 6.14 shows the computed step lengths over the iterations on the left hand side and the corresponding regularization parameters ω on the right hand side, both for the final time observation parameters α equal to 0, 1 and 10 (top to bottom). As in the previous figure, we observe that the behavior is *significantly* different from the behavior in Example 1, see the corresponding figure (Figure 6.5). Instead of an overall tendency to remain within a fixed range of a few order of magnitudes for all algorithms, the computed step lengths exhibit behavior that is dependent on the specific algorithm and the problem configuration.

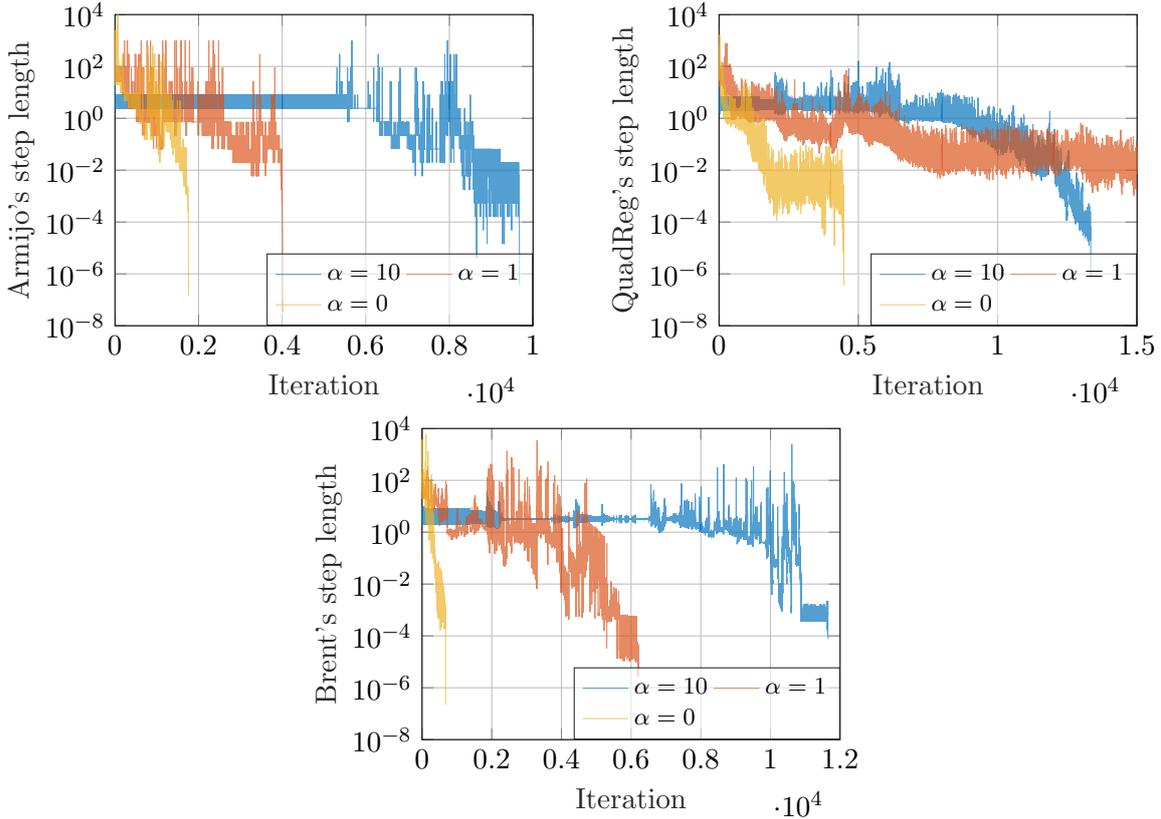


Figure 6.15: The computed step lengths for the Armijo step length algorithm (top left), the QuadReg algorithm (top right) and Brent's algorithm (bottom).

The step lengths that are computed as part of the standard line search algorithm by Armijo's rule, QuadReg and Brent's algorithm behave similarly. They are depicted as plots by algorithm that include the runs for all values of α in Figure 6.15. As we can see, Armijo's rule produces step lengths that show a clear decrease over the course of the algorithm, regardless of the parameter α . For the value $\alpha = 10$, however, we can see that the step lengths tend to be constant for the first 5263 iterations, where the decrease sets in. The step lengths computed by the quadratic regularization approach settle into a relatively constant behavior after an initial phase of adaptation for $\alpha = 0$ and 1 but show a clear change in the constant behavior towards an overall decrease after the first 6000 iterations for $\alpha = 10$. The behavior of the step lengths computed by Brent's algorithm is very similar to that of the Armijo rule, since the magnitude of the step length is significantly decreased

over the course of the iterations. For the value $\alpha = 10$, we can again observe an initially constant tendency and a subsequent decrease in the step lengths in the long term behavior from 8000 iterations on. Note that these points of change in the behaviors correlate with the points of the changed behavior of the norm of the search directions, which remains constant in its general tendency after the change in behavior, see Figure 6.13. These observations suggest that the optimization algorithms are influenced by nonsmooth effects around the final iterates, which leads to a deterioration in the quality of the search direction, causing the norm of the search directions to stay within a few orders of magnitude instead of decreasing, while the search directions are required to correct the overly confident length of the search directions by decreasing. Cf. the overview in Section 6.1.2. The fact that the decrease in the step lengths for $\alpha = 10$ tends to set in after an initial phase of fairly constant behavior can be attributed to the low regularity of the problem, which causes the nonsmooth regions near the desired state to be reached much slower than they are in the more regular problems.

Since both the restarted and the non-restarted accelerated scheme repeatedly use the computed search direction as an *intermediate* update only, the behavior of the computed step lengths is not directly linked to the quality of the search direction, which it is in the standard line search. However, we can observe that the step lengths of the non-restarted scheme tend to be fairly constant over the course of the algorithm, independently of the final time observation parameter α — a behavior that we have seen in Example 1 as well and that is usually observed in smooth problems. In the restarted scheme, on the other hand, the step lengths exhibit an overall tendency to remain fairly constant but with increased volatility. We can observe a connection between the restarting patterns of the momentum and the structure of the decrease of the functional value and the computed step lengths, see Figures 6.16 and 6.17.

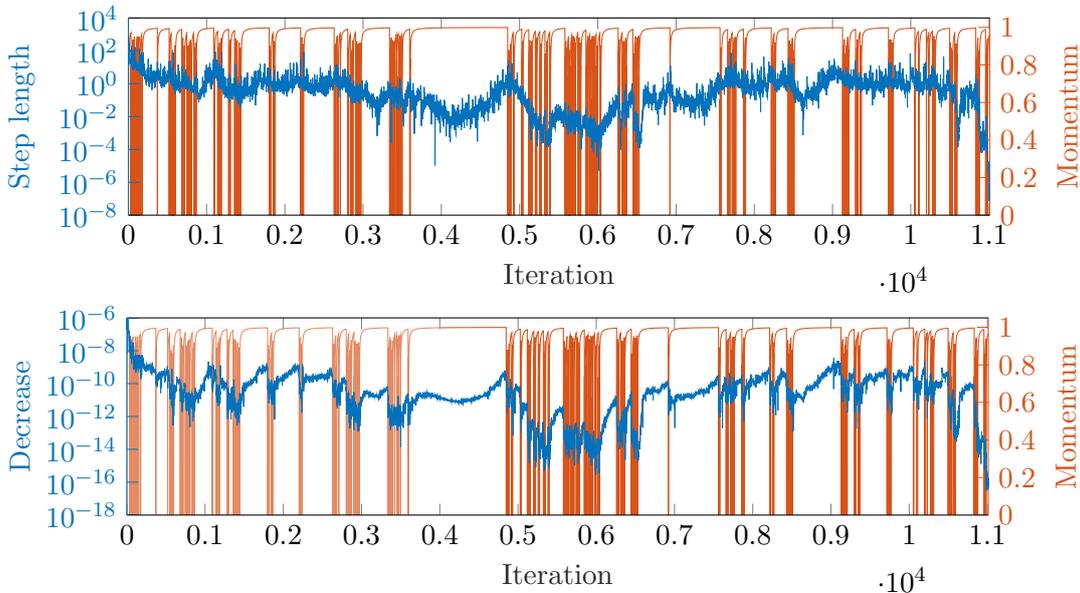


Figure 6.16: The momentum of the restarted accelerated scheme vs. the computed step lengths (top) and the objective functional decrease (bottom) for $\alpha = 1$.

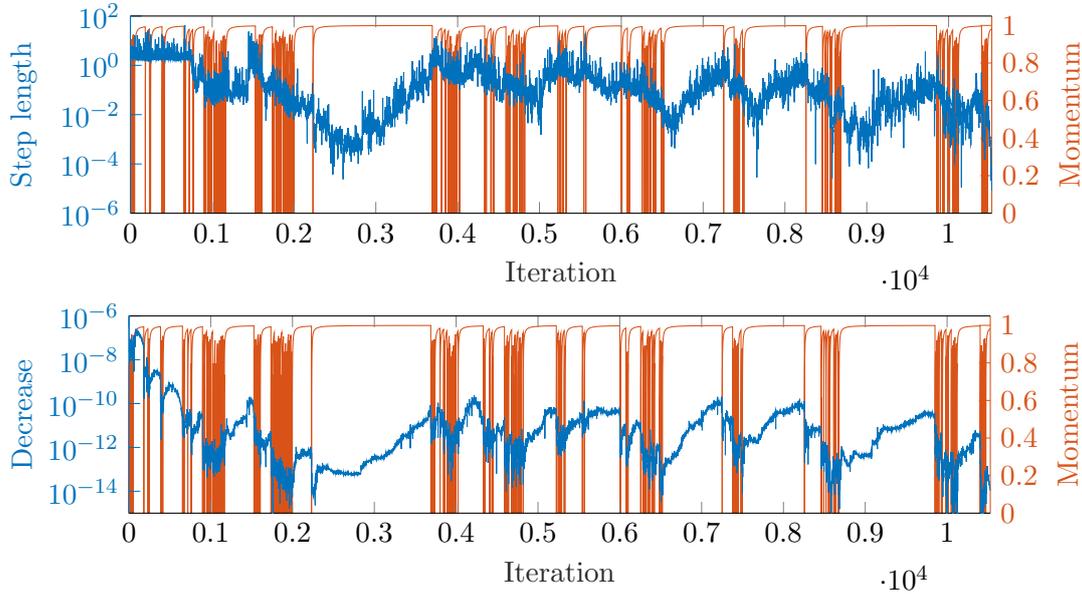


Figure 6.17: The momentum of the restarted accelerated scheme vs. the computed step lengths (top) and the objective functional decrease (bottom) for $\alpha = 10$.

We can observe that the build up in momentum correlates with an increase in the decrease of the objective functional value and a V-shaped development of the step lengths. It is unsurprising that the build up of momentum corresponds to the increase in the decrease in the functional value because the restarts are triggered whenever ascent is detected in the non-restarted scheme, which commonly happens right after very good progress was made — see the ripples in the functional values in Figure 6.13. In a sense, this effect can also be observed in Example 1 as well. Since the overall tendency in well behaved, smooth problems is for the decrease to decrease over the course of the algorithm, this means that one can observe that the decrease stays within the same order of magnitude during the momentum build up phase, and that it abruptly plummets when the momentum is restarted, see Figure C.3. This same behavior is observed in the first 800 iterations in Figure 6.17. The correspondence between the build up of momentum and the computed step lengths suggests that the search directions are not first order consistent around the iterates at the beginning of one “momentum cycle”. In the beginning of one of these cycles, the accelerated updates are close to the corrections that are computed in the standard line search method with the quadratic regularization approach for computing the step lengths. After the cycle has progressed and several steps that include the momentum update have been taken, the iterates appear to reach a region in which the model for the computation of the search directions is accurate in a greater neighborhood of the respective iterates, which allows for an increase in the step lengths. The high momentum influence in these parts of the cycle ultimately leads to ascent and a restart of the momentum. These effects are more distinct after the initial phase of the algorithm. The behavior implicates that the step lengths are decreased to compensate for an inconsistent model when the momentum’s influence is low. With increasing momentum, the algorithm adheres to the search directions less and potentially progresses to regions where the quality of the model improves.

Discussion of the Solution(s). In Example 1, all of the five algorithms computed the same solutions up to a relative error in the controls of $2 \cdot 10^{-6}$, cf. Table 6.3. That is not the case for this example, where the algorithms appear to compute quite different solutions with a maximum relative error of order 1 in the case of $\alpha = 1$, see Table 6.5. Accordingly, discussing only one solution is insufficient.

	Absolute norm of differences			Relative norm of differences		
	Max	Min	Mean	Max	Min	Mean
$\alpha = 0$						
Controls	$1.6 \cdot 10^{-1}$	$3.7 \cdot 10^{-2}$	$9.1 \cdot 10^{-2}$	$7.7 \cdot 10^{-1}$	$1.8 \cdot 10^{-1}$	$4.3 \cdot 10^{-1}$
States	$8.7 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$	$4.3 \cdot 10^{-4}$	$6.2 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$	$3.0 \cdot 10^{-2}$
Values	$3.8 \cdot 10^{-7}$	$5.6 \cdot 10^{-8}$	$1.9 \cdot 10^{-7}$	$1.5 \cdot 10^{-2}$	$2.3 \cdot 10^{-3}$	$7.7 \cdot 10^{-3}$
$\alpha = 1$						
Controls	$3.5 \cdot 10^{-1}$	$5.3 \cdot 10^{-3}$	$1.9 \cdot 10^{-1}$	$1.1 \cdot 10^0$	$1.6 \cdot 10^{-2}$	$6.0 \cdot 10^{-1}$
States	$1.3 \cdot 10^{-3}$	$4.8 \cdot 10^{-6}$	$6.6 \cdot 10^{-4}$	$9.1 \cdot 10^{-2}$	$3.4 \cdot 10^{-4}$	$4.6 \cdot 10^{-2}$
Values	$2.4 \cdot 10^{-6}$	$6.4 \cdot 10^{-9}$	$1.2 \cdot 10^{-6}$	$9.1 \cdot 10^{-2}$	$2.2 \cdot 10^{-4}$	$4.5 \cdot 10^{-2}$
$\alpha = 10$						
Controls	$1.9 \cdot 10^{-1}$	$3.0 \cdot 10^{-3}$	$9.1 \cdot 10^{-2}$	$5.9 \cdot 10^{-1}$	$8.9 \cdot 10^{-3}$	$2.8 \cdot 10^{-1}$
States	$4.7 \cdot 10^{-4}$	$3.0 \cdot 10^{-6}$	$2.0 \cdot 10^{-4}$	$3.3 \cdot 10^{-2}$	$2.1 \cdot 10^{-4}$	$1.4 \cdot 10^{-2}$
Values	$1.1 \cdot 10^{-6}$	$8.7 \cdot 10^{-11}$	$4.7 \cdot 10^{-7}$	$4.0 \cdot 10^{-2}$	$3.0 \cdot 10^{-6}$	$1.7 \cdot 10^{-2}$

Table 6.5: The $L^2(0, T; \mathbf{L}^2(\Omega))$ -norms of the differences of the computed solution controls and solution states and the differences of the corresponding objective values taken pairwise for the five compared algorithms for final time observation $\alpha = 0, 1, 10$. Relative values are taken relatively to the smaller of the two respective values.

Figure 6.18 shows the $\mathbf{L}^2(\Omega)$ -norms of the difference of the desired and the result state of the respective algorithms on the left hand side and the $\mathbf{L}^2(\Omega)$ -norm of the result controls over time on the right hand side. Qualitatively, the states and controls behave similarly for all α , but there is a noticeable difference in the timing of the solutions. In all solutions, the body is brought into contact with the obstacle rapidly by downward forces that are reduced during contact and re-engage in the rebound phase in order to control the rebound and keep the body close to the obstacle. For $\alpha = 1$ and 10, some qualitative differences are noticeable between the solutions computed by the accelerated and the standard schemes. The accelerated schemes generally produce controls that have an oscillating component and states that tend to be in contact early and show stronger rebound.

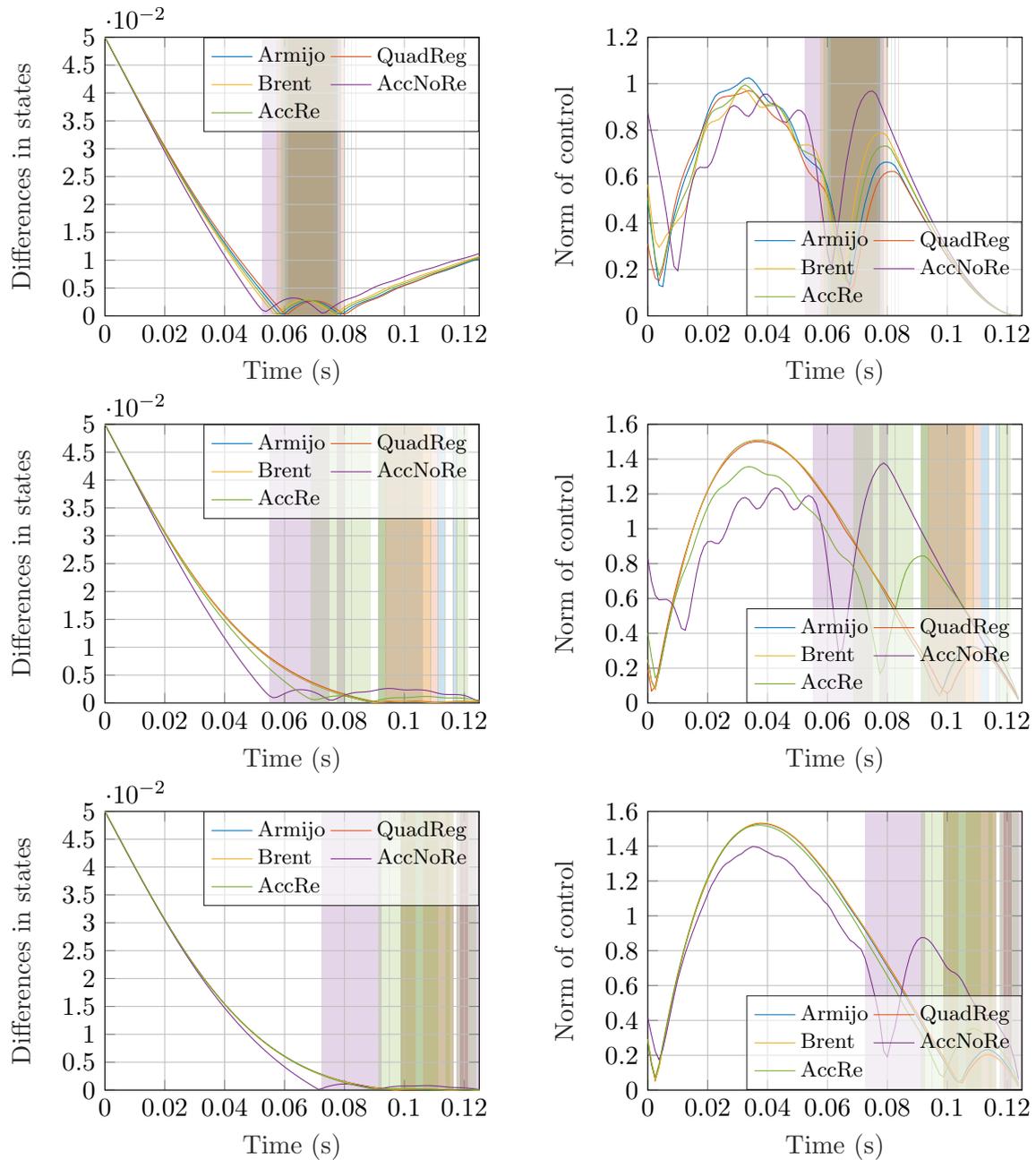


Figure 6.18: The $L^2(\Omega)$ -norm of the differences of the computed solution state and the desired state (left) and the $L^2(\Omega)$ -norm of the computed solution control (right) over time for $\alpha = 0, 1, 10$ (top to bottom). Time of contact of the computed solutions is shaded.

The typical behavior of the solution states and the corresponding contact forces can be found in Figure 6.19, which shows the vertical position of the solution state at the center of the lower side of the square domain and the corresponding component of the contact force density of the solutions computed by Armijo's rule for the range of considered parameters α . When final time observation is absent ($\alpha = 0$), the body moves towards the obstacle, establishes contact and initial contact forces and then rebounds off the obstacle. After the detachment phase, a slight vibration in the material leads to contact without any active contact forces at this node. For active final time observation, the trajectory of the body towards the obstacle is less aggressive. This results in a delay in the engagement of contact compared to the case of $\alpha = 0$. The magnitude of the contact forces decrease as α increases with a difference of a full order of magnitude between $\alpha = 0$ and $\alpha = 1$. Additionally, the body does not simply bounce off the obstacle once but rebounds several times. Each time contact is established, the contact forces acting on the body are decreased and the body's movement tends to stabilize near the obstacle with low or no contact forces at all.

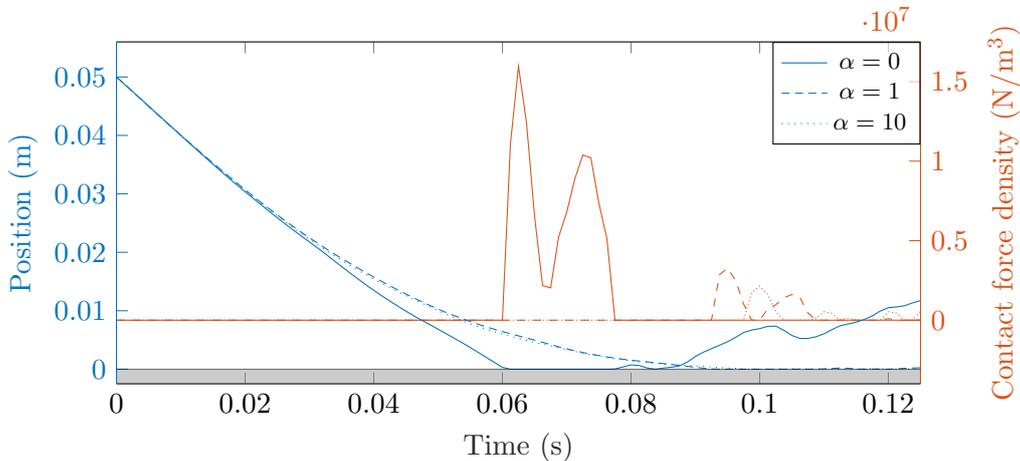


Figure 6.19: Vertical component of the position and of the contact force density corresponding to the center of the bottom side of the domain for the solutions computed by Algorithm 6.1 with the Armijo rule. Gray indicates the obstacle.

Finally, Table 6.6 shows the lowest objective values that were computed by the five algorithms for the respective parameters α . It is worth to note that the non-restarted, accelerated scheme manages to achieve the overall lowest functional value of the algorithms, independently of the final time observation. This is especially interesting because the algorithm modifies the initial search direction that is computed from the linearized model (6.4) the most compared to the rest of the algorithms. Recall that the search directions of the algorithm were the only ones to show a general tendency to remain constant.

	Armijo	QuadReg	Brent	AccNoRe	AccRe
$\alpha = 0$	$2.493 \cdot 10^{-5}$	$2.507 \cdot 10^{-5}$	$2.474 \cdot 10^{-5}$	$2.469 \cdot 10^{-5}$	$2.482 \cdot 10^{-5}$
$\alpha = 1$	$2.859 \cdot 10^{-5}$	$2.867 \cdot 10^{-5}$	$2.860 \cdot 10^{-5}$	$2.629 \cdot 10^{-5}$	$2.730 \cdot 10^{-5}$
$\alpha = 10$	$2.882 \cdot 10^{-5}$	$2.882 \cdot 10^{-5}$	$2.881 \cdot 10^{-5}$	$2.771 \cdot 10^{-5}$	$2.868 \cdot 10^{-5}$

Table 6.6: The lowest objective functional value achieved by the respective algorithms in Example 2.

6.2.4 Example 3: A Parameter Balancing Problem

In the previous two examples, we have focused on the effects of nonsmoothness on the performance of the optimization algorithms. In this third and last example, we shortly examine the influence of the regularization parameter on the application of external forces in a problem where the movement of the body can be manipulated by either external forces, which come at a cost, or by contact forces, which are free of charge. We revisit the geometry introduced in the first example, in which nonsmoothness seemed not to impede the performance of the optimization algorithms, which all obtained the same solution.

Reference Configuration and Objective Functional. Reproducing the semicylindric setting that we have examined in the first example, we set

$$\begin{aligned}\Omega &:= \{ \mathbf{x} \in \mathbb{R}^2 : x_1^2 + (x_2 - 1.05)^2 < 1 \wedge x_2 < 1.05 \}, \\ \mathcal{O} &:= \{ \mathbf{x} \in \mathbb{R}^2 : x_2 < 0 \},\end{aligned}$$

which yields the reference configuration in Figure 6.20, cf. Figure 6.1 for the identical reference configuration in Example 1.

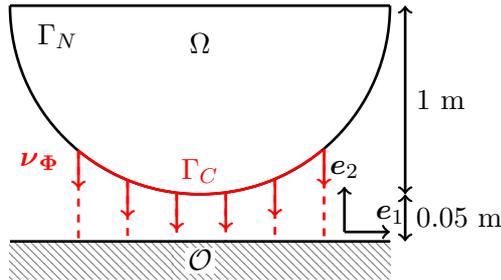


Figure 6.20: Reference configuration in Example 3. (Same as in Example 1.)

The initial displacement for the state is assumed to vanish, but we consider an initial velocity of $\mathbf{v}_{\text{ini}} = (0, -1)$ m/s. The desired state is set as vanishing displacement on the time interval of 0 to 0.12s, i.e., the desired position corresponds to the reference configuration.

T (s)	τ (s)	\mathbf{y}_{ini} (m)	\mathbf{v}_{ini} (m/s)	$\Psi(\mathbf{x})$ (m)	$\boldsymbol{\nu}_{\Phi}$	\mathbf{y}_d (m)
0.12	$1.2 \cdot 10^{-3}$	(0, 0)	(0, -1)	x_2	(0, -1)	(0, 0)

Table 6.7: Parameters in Example 3.

We are interested in the behavior of the solution when the regularization parameter β is varied and the final time observation parameter is fixed. From the plethora of parameter settings to choose from, we consider $\beta \in \{10^{-2}, 10^{-5}\}$ for $\alpha = 1$ here.

The setting described above corresponds to the body traveling towards the obstacle and away from its desired position. The optimizers strive to find external forces that position the body at its reference configuration, with an increased interest in the body's position at the fixed end of the time interval, while including the cost of the external forces depending on the regularization parameter β . For varying β , the question is whether to use external forces

to decelerate the body (catch) or to apply minor external forces and allow the resulting contact forces to return the body to its reference configuration (bounce). Catching the body allows for a smaller difference between the displacements but requires costly external forces while allowing the bounce increases the difference in the states but reduces the cost for external forces because the momentum of the body is reversed by the contact forces, which are introduced by the initial velocity and come at no cost. For a visualization of the setup, think of a basketball player trying to control an incoming ball either at the expense of his own forces by catching the ball directly or by bouncing it off the floor.

Solver Behavior. Both problems were solved with the restarted accelerated line search method (AccRe) from a vanishing initial control since the method performed best in Example 1. The results suggest sufficient reliability of the algorithm *for the configuration at hand*. The algorithms were stopped when the step length computation was unable to produce a step length that yielded strict decrease and after a maximum of 1000 iterations. This ensures that the $L^2(0, T; \mathbf{L}^2(\Omega))$ -norm of the search direction — which coincides with the gradient at all iterates where the reduced objective functional is Fréchet differentiable — is reduced below a threshold of 10^{-8} .

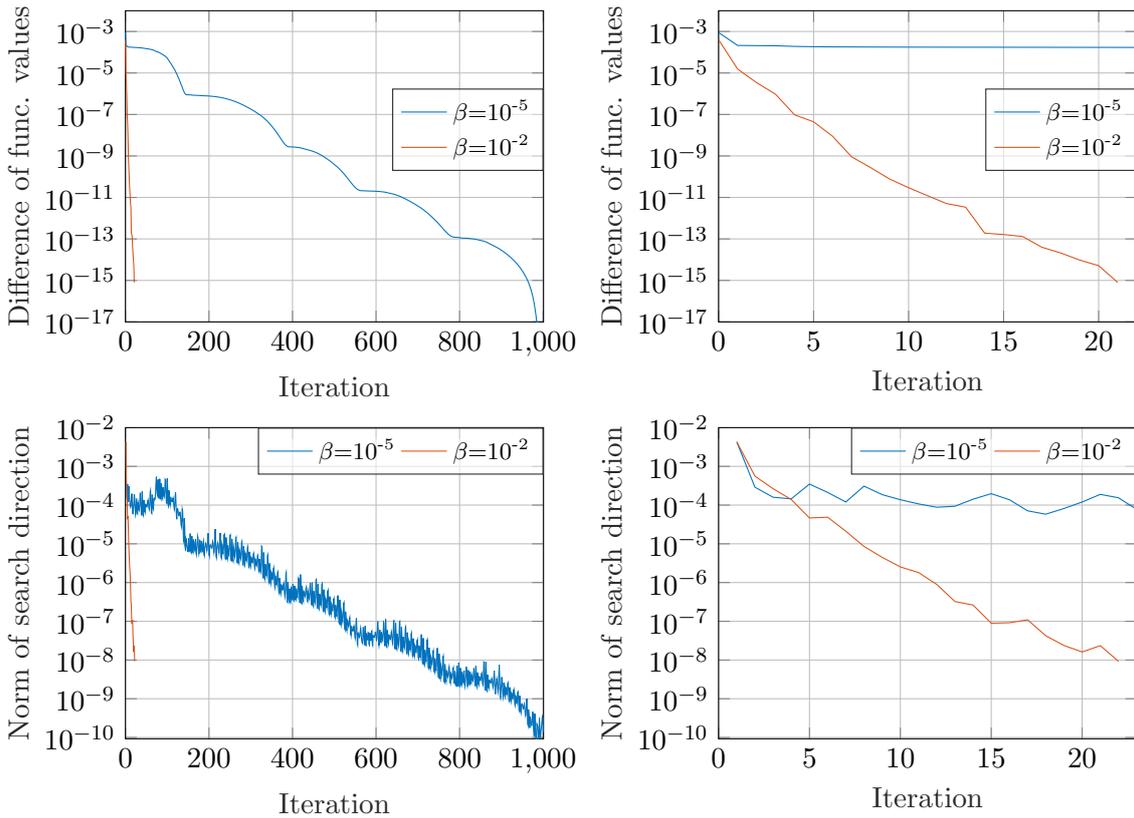


Figure 6.21: The difference of the current functional values and the final functional values over the iterations (top) and the norm of the search direction over the iterations (bottom). The right hand side shows a modified range of iterations.

The performance of the algorithm in the two settings surely differs quantitatively, since larger values of β yield a more regular problem, which is easier to solve. However, the results clearly coincide qualitatively. The behavior of the objective functional values' difference and the norm of the search directions is in line with the results of Example 1, showing the characteristic cyclic decrease in both functional values and search direction norm, see Figure 6.21. Similarly, the progression of the restarted scheme's momentum is comparable to the one in Example 1. As we can see in Figure 6.22, the number of restart intervals of the momentum increases with lower regularity of the problem, which can be expected since lower regularity "flattens" the graph of the reduced objective functional.

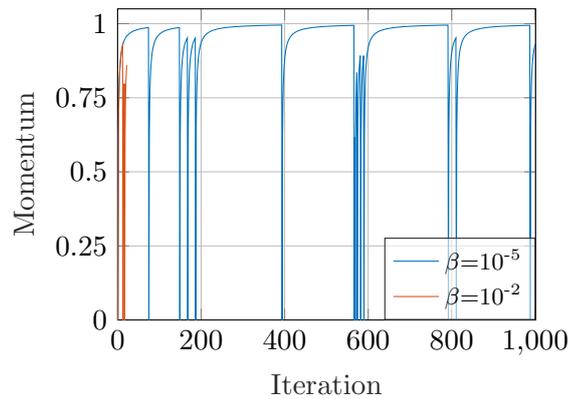


Figure 6.22: The momentum of the accelerated scheme over the iterations.

Figure 6.23 shows that the overall tendency of the computed step lengths and the quadratic regularization parameter ω is stable and that the values remain within the same order of magnitude for both settings. In combination with the decrease of the norm of the search directions, this suggests that the search direction is first order consistent and that the objective functional model yields a first order approximation of the objective functional in a reasonably large neighborhood of the iterates.

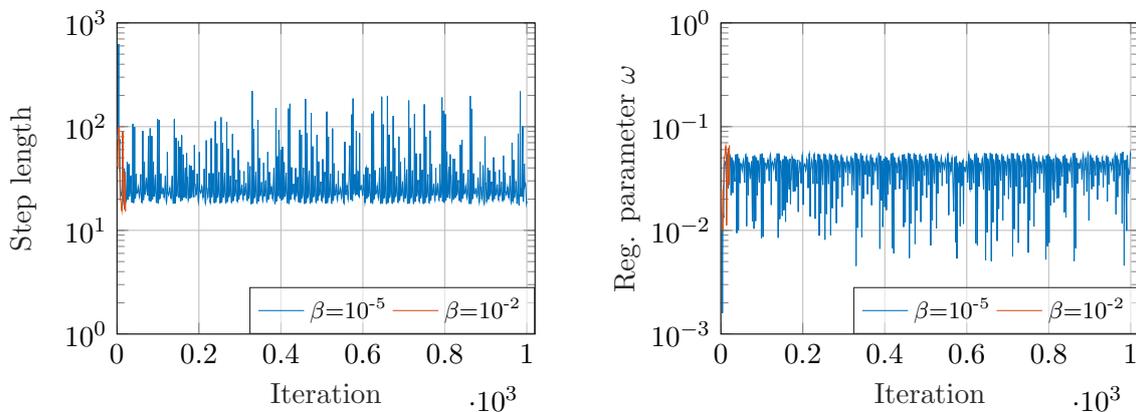


Figure 6.23: The computed step lengths (left) and the regularization parameters ω (right) over the iterations.

Discussion of the Solution(s). As expected, the two respective solutions behave quite differently, as Figure 6.24 shows, where the vertical position of the south pole of the semi-circle and the vertical component of the contact force density are displayed. The solution for $\beta = 10^{-2}$ shows a similar behavior as the gravity-induced trajectory in Example 1, first moving towards the obstacle, establishing contact with the obstacle — which introduces contact forces — and finally disengaging and moving away from obstacle again. When the magnitude of the forces is penalized *less* in the objective functional by $\beta = 10^{-5}$, the body never engages the obstacle but is moved back up, away from the obstacle, with the downward movement turning into an upwards motion at $t = 0.0216$ s, moving above its intended position of 0.05 m above the obstacle at $t = 0.0804$ s and the upward motion once again becoming a downward motion at $t = 0.108$ s. Cf. the plot of the $L^2(\Omega)$ -norm of the displacement on the left hand side of Figure 6.8 for a better understanding of the external forces involved. Of course, there are no contact forces present, since contact is never established.

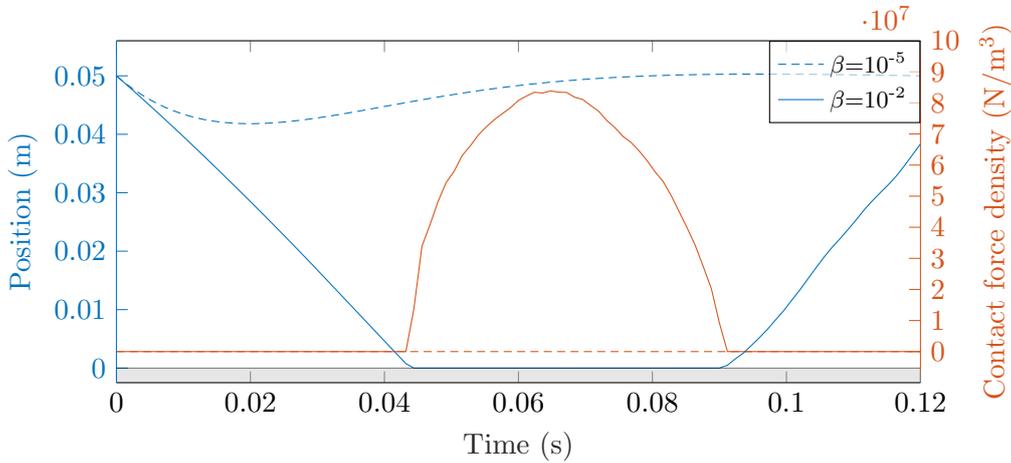


Figure 6.24: Vertical component of the position and of the contact force density corresponding to the south pole. Gray indicates the obstacle.

We can see that the $L^2(\Omega)$ -norm of the controls for both constellations are qualitatively comparable, see the right hand side of Figure 6.25. The controls are largest in the beginning of the time interval, and both results have a strictly positive local minimum before the controls are shut off at the end of the time interval. For $\beta = 10^{-2}$, the magnitude of the forces is moderate and within the same order of magnitude over a large part of the time interval with the local minimum at the time of contact. For $\beta = 10^{-5}$, significantly larger forces are applied prior to the local minimum compared to the forces after the minimum.

The different behavior of the controls can be analyzed further with the information from Figure 6.26, which shows the orientation and magnitude of the forces on the body. Since the magnitude of the controls differ significantly over time, the arrows are *not* scaled by the magnitude of the forces to indicate orientation of the force vectors, but the magnitude is represented by the coloring according to the legend. For $\beta = 10^{-2}$ on the left hand side, the initial external forces are applied downwards, further accelerating the body towards the

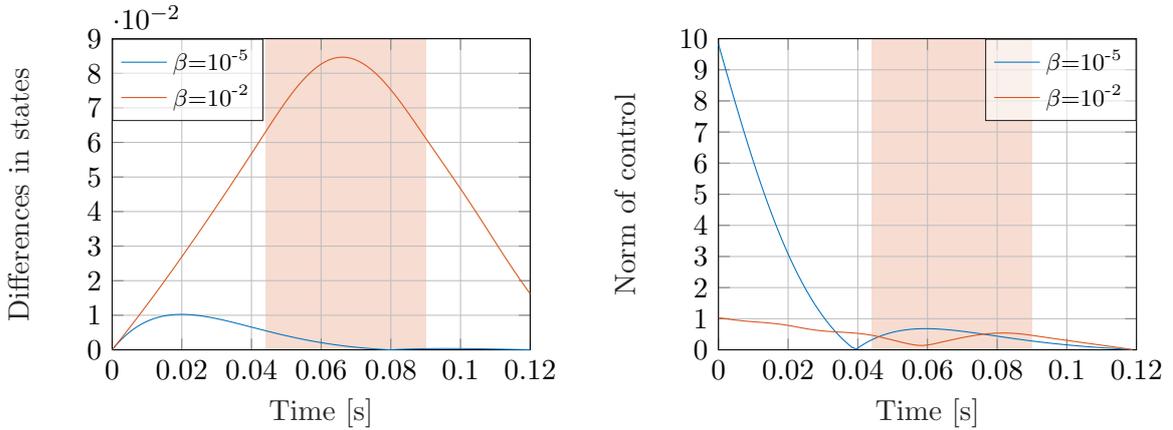


Figure 6.25: The $L^2(\Omega)$ -norm of the differences of the computed solution state and the desired state (left) and the computed solution control (right). Time of contact is shaded.

obstacle. During the time of contact, the external forces are significantly reduced, and the momentum of the body is turned upward by the contact forces while weak external forces counteract the distortion of the bodies shape introduced by the elasticity of the material. Once the body is in an upward motion, external forces are applied in upward direction as well, assisting in the return of the body to its desired reference configuration until the forces are eventually shut down. The right hand side shows the solution for $\beta = 10^{-5}$, where external forces are less expensive. In the beginning of the time interval, large forces are applied upward to change the momentum of the body, which moves upward from time $t = 0.0204$ s, see Figure 6.24. Acceleration is maintained upward until $t = 0.04$ s, over-accelerating the body towards its desired position, see the images in the second and third row of the figure. Until the end of the time interval, the body is accelerated downwards again to counter the over acceleration and correct the overshooting of the desired state, see Figure 6.24.

As expected, depending on the cost of the external forces, it can be beneficial to use the contact forces — that are essentially introduced by the initial velocity — to return the body to its reference configuration. Since the solution state corresponding to $\beta = 10^{-5}$ never even establishes contact and the solution corresponding to $\beta = 10^{-2}$ is both dynamic with contact forces acting over the entire contact patch with the exception of two points at a time, neither solution involves weak contact/biactive sets, hence the reduced objective functional is Fréchet differentiable at the solution, and Figure 6.21 shows that the norm of the gradients at the solution is less than 10^{-8} when the algorithm terminates. The behavior of the optimization algorithm further indicates that nondifferentiability of the objective functional seems to not be an issue over the course of the runs. Since contact plays a role in both settings, the problems are still highly nonlinear. The problem corresponding to $\beta = 10^{-5}$ still requires the solver to deal with iterates where the body establishes contact with the obstacle because the initial iterate's state collides with the obstacle.

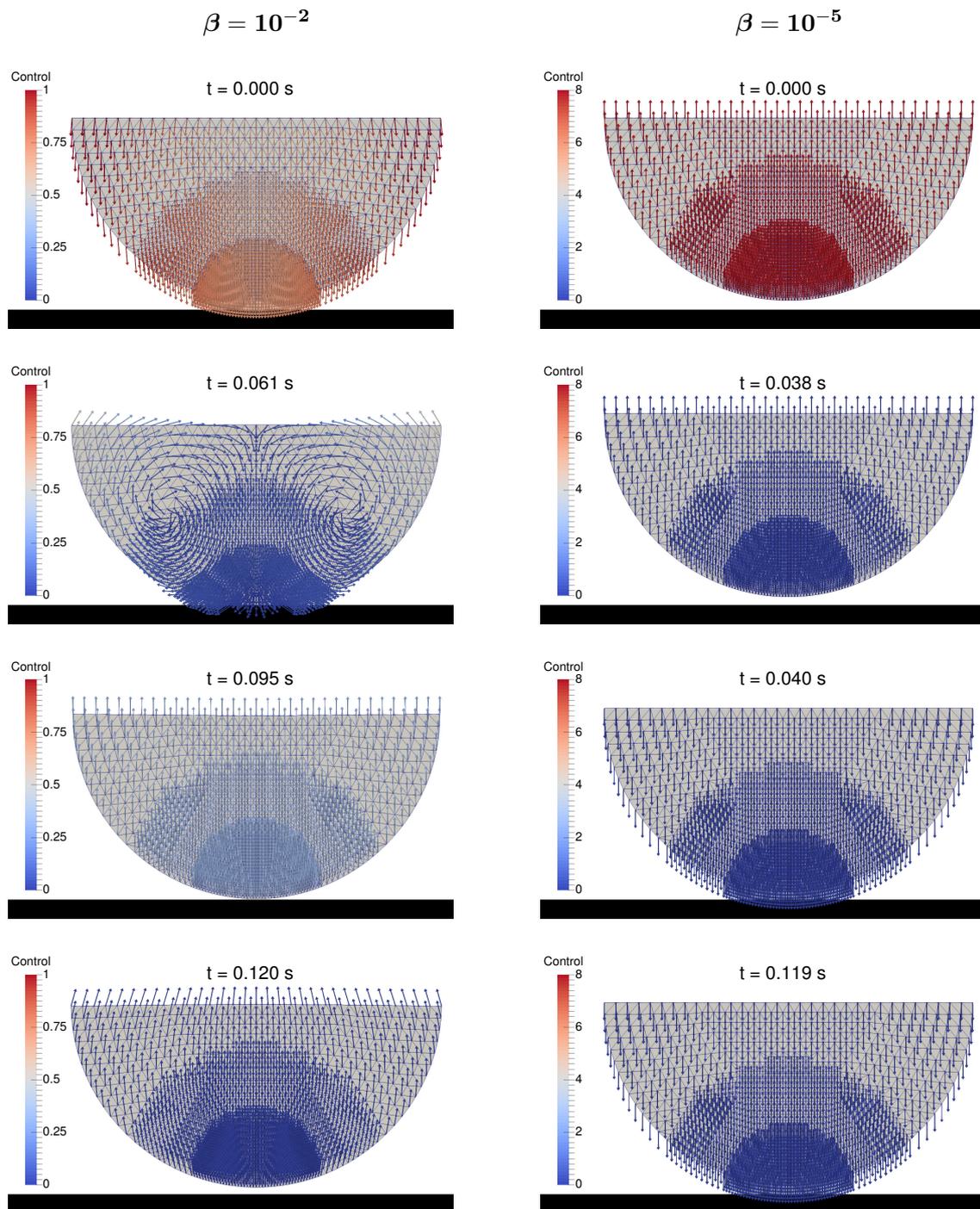


Figure 6.26: The computed solutions for $\beta = 10^{-2}$ (left) and $\beta = 10^{-5}$ (right) at selected times. Direction of the arrows shows orientation, and color shows magnitude.

6.3 Evaluation

In Chapter 5, we have seen that the reduced objective functional is Fréchet differentiable when the biactive sets vanish. In the chapter above, we have investigated the conjecture that the existence of biactive sets is strongly dependent on the geometric similarity of the contact boundaries of the body and the obstacle as well as the form of the objective functional.

The first example involved a problem with dissimilar contact boundary shapes and an objective functional that is unlikely to produce deformations that allow for weak contact patches. As expected, the behavior of the optimization algorithms, when applied in this setting, resembled the standard gradient scheme's behavior in smooth, nonlinear problems with a consistent decrease in the norm of the search directions and a tendency for step lengths that remain constant by and large. Up to sufficient accuracy, the solutions computed by the algorithms coincide and exhibit behavior that is to be expected in this type of inverse problems, taking into account the rigidity of the obstacle. As the name suggests, the accelerated momentum scheme improves the decrease provided by the gradient scheme, especially within the first couple of iterations. With increasing number of iterations, the initial advantage of the accelerated method decreases. Since it comes at the very small cost of one additional comparison of functional values but offers significantly better performance, the restarted accelerated method is clearly preferable to the standard and the non-restarted accelerated algorithm.

Example 2 consisted of a body and an obstacle with parallel contact boundaries and an objective functional that incentivizes optimal solutions to establish weak contact. The nonlinearity in this problem is quite obvious because the results computed by the different algorithms differ significantly. We have seen that the goal of establishing relatively steady contact with low contact forces could be achieved for higher final time observation parameters. When the accelerated scheme is not restarted, the functional values' behavior resembles that of the well-behaved Example 1 and the algorithm consistently reaches the lowest functional value of all algorithms. All other algorithms behave more erratic, and the norm of the search directions does not decrease steadily but stagnates well before the algorithms terminate. The step lengths of the algorithms that include an update or a sufficiency condition that evaluate the model with respect to its order of consistency around the iterate indicate that the computed search directions are not first order consistent in this example — an effect that is easier observed for higher final time observation parameters as well. Restarting the accelerated scheme, we have observed that over a single momentum cycle, the quality of the search direction seems to not be first order consistent until several momentum assisted update steps have been taken and the region of relevance of the model used to compute the search direction increases. These observations suggest that the problem with higher final time observation parameters is affected by nonsmoothness around the minimizer and that using the computed search direction in a straight forward line search algorithm leads the iterates closer to these points of nonsmoothness, compromising the quality of the search direction. This results in deteriorated performance of the algorithms and generally gives the accelerated schemes, which adhere to the computed search direction the least, an advantage.

In Example 3, we have shortly addressed the influence of the regularization parameter in the regularized tracking-type functional with respect to the balance of external forces and contact forces. We have seen that the contact forces will be used to augment the external forces when those become sufficiently expensive.

As far as the line search options are concerned, the large number of required functional evaluations for the optimal step length approximation by Brent's algorithm and its parameter dependency makes it impractical compared to the other two options. The Armijo rule suffers from a similar drawback when the step length is computed from the same starting step length, since reasonable step lengths in the first couple of iterations differ from the step lengths in the iterations later in the run, making lots of functional evaluations necessary. This effect could of course be remedied somewhat by introducing a memory-type effect into the method. The method remains highly parameter dependent, however, requiring some information on the problem structure in advance, especially since the scaling of the search directions is arbitrary. The only user-provided parameters needed in the quadratic regularization algorithm are those of the safe-guard implementations that ensure that the regularization parameter is not automatically increased by too many orders of magnitude in a single outer iteration. In practice, the algorithm requires only few, usually less than five, evaluations of the objective functional to compute a reasonable step length for the given search direction, making the quadratic regularization preferable to the former options.

Chapter 7

Conclusions and Outlook

In this thesis, we have seen analytical and numerical results for optimal control problems governed by time-discretized dynamic contact problems.

We have constructed a solid foundation for modeling these problems by establishing that the contact constraint in the weak framework can equivalently be formulated in the quasi everywhere sense induced by several previously unconnected Sobolev capacities as well as the boundary measure theoretical sense. This enabled the use of function theoretical results from all approaches in the analysis, which was based on a temporal finite element method for the contact problem that yields a variant of the contact implicit Newmark time stepping scheme and induces a consistent adjoint time stepping scheme. We have, however, restricted ourselves to a setting in *linear* viscoelasticity. Therefore, the deformation of the bodies is required to be small and rotation free in order for the model to be suitable, and larger deformations will require a more involved approach. Contact problems in large deformation elasticity are a challenging topic that has been considered in, e.g., [164, 206], and optimal control of static contact problems in finite strain elasticity is currently a topic of the research in [168]. Additionally, before optimization can be addressed in the time-continuous setting, a reasonable framework w.r.t. weak solutions remains to be established to carry out the required analysis, which is expected to be a demanding task.

Exploiting a transfer in the polyhedricity properties of sets under linear mappings, we have proven the existence of a Lipschitz continuous and Hadamard differentiable solution operator in the time-discretized setting, though the Lipschitz constants are expected to be strongly reliant on the specific time discretization. We were able to precisely characterize the points of nondifferentiability in the sense of Gâteaux in terms of the biactive sets, owing to a localized representation of the contact forces and the resulting pointwise characterization of the linearized boundary conditions. Based on the analysis of the solution operator and the stationarity concepts in [193], we have obtained optimality conditions of strong-stationarity-type for the time-discretized optimal control problem — similarly to the conditions derived by Mignot [138] in a scalar valued setting — but required the image of the control-to-force operator to be dense in the right hand sides of the problem. This restriction to sufficiently “rich” controls — like distributed controls — usually appears in the literature when strong stationarity conditions are derived, cf. [92, 137, 140].

When those assumptions are dropped, one is commonly reduced to obtain C-stationarity, see, e.g., [91, 169], which remains to be proven for the contact problem at hand. Furthermore, Rauls and Wachsmuth just recently developed a strengthened C-stationarity condition based on the optimal control of the obstacle problem [162]. See also [196] for results on M-stationarity concepts for the control constrained obstacle problem, and see [185] for numerical results for the optimal control of a two-body contact problem in linear viscoelasticity with Neumann boundary controls.

For the two proposed line-search-type optimization algorithms in the numerics section, we have used the differentiability results for the reduced objective functional to compute an adjoint-based search direction that coincides with the $L^2(\Omega)$ -gradient of the objective functional whenever the functional is Fréchet differentiable. It remains to prove that this search direction is a subgradient whenever the reduced objective functional is not differentiable, as is suggested by the recent publications [43, 161, 162] and the works of Outrata et al. [155]. We have seen that the modification of the boundary conditions used for the computation of the search direction in Section 6.1.1 is only one of multiple options for computing a search direction that coincides with the gradient in points of differentiability. The connection between the corresponding search directions and subgradients for the reduced problem remains to be investigated for now, but the flexibility in the choice of boundary condition can potentially be exploited in the computation of search directions, e.g., for improved approximation properties of the subdifferential in bundle-type methods or, as Rauls and Ulbrich suggest [161], in inexact methods, see [90].

The connection between biactive sets and nondifferentiability of the solution operator (Proposition 4.4.25) and the numerical results in Sections 6.2.2 – 6.2.4 support the conjecture that the influence of nonsmoothness on the performance of the algorithms and their respective solutions is strongly dependent on the combination of the involved objects' geometries and the objective functional. They suggest that a rather simple subgradient-type approach, such as the one presented in this thesis, can be viable for solving problems of the type that arise in total knee replacements, where rounded surfaces on the one hand face planar surfaces on the other and where external forces steadily keep the objects in contact, see, e.g., [100] or think of ligaments exerting forces to keep the components engaged. On the other hand, efficiently solving problems where objects with similar boundary geometries tend to change from contact to non-contact phases frequently and gaps between the objects are desired to be minimal, such as the design of break discs and pads, is likely to require more sophisticated optimization approaches, such as the bundle methods developed in [43, 90, 171, 172].

Since nonlinear CG methods are known to generally perform on par with or superiorly to the restarted accelerated scheme for convex optimal control problems, a CG-type approach based on the search directions' computation above is another promising option when bundle information is unessential. Cf. [187, 192] for CG applications in nonsmooth, infinite dimensional optimization. One has to keep in mind, however, that reliability and increased performance of nonlinear CG-based methods are satisfying mostly when the functional to be minimized is locally close to a quadratic problem, whereas contact problems are highly nonlinear.

Appendix A

Function Analytical Results

This chapter contains some auxiliary results that mostly support the Sobolev capacity theory in Chapter 3. The results cover integration and trace theory and address the relationship between functions in the Sobolev space $W^{1,\infty}$ and Lipschitz continuous functions with respect to the contact normal and its extension.

A.1 Integration Theory

In this first section, measure and integration theoretic results are proven to support the Lemmas in the next section. For introductory literature into measures and integration theory, see, e.g., [30, 65, 66, 201].

The first result is the following intuitive result on integration.

Lemma A.1.1 (Equivalent Measures) *Let (X, Σ) be a measurable space, $C > 0$ and μ_1, μ_2 be measures on Σ that satisfy*

$$\mu_1(A) \leq C \mu_2(A) \quad \forall A \in \Sigma.$$

Then for any Σ - $\mathcal{B}(\overline{\mathbb{R}})$ measurable function $f: X \rightarrow [0, \infty]$ and $A \in \Sigma$,

$$\int_A f \, d\mu_1 \leq C \int_A f \, d\mu_2.$$

Proof. Let $g: X \rightarrow [0, \infty)$ be a simple, Σ - $\mathcal{B}(\overline{\mathbb{R}})$ measurable function of the form

$$g = \sum_{k=0}^n y_k \chi_{g^{-1}(y_k)}$$

and $A \in \Sigma$. Then

$$\int_A g \, d\mu_1 = \sum_{k=0}^n y_k \mu_1(g^{-1}(y_k) \cap A) \leq C \sum_{k=0}^n y_k \mu_2(g^{-1}(y_k) \cap A) = C \int_A g \, d\mu_2.$$

Now let $f: X \rightarrow [0, \infty]$ be Σ - $\mathcal{B}(\overline{\mathbb{R}})$ measurable, then by definition

$$\begin{aligned} \int_A f \, d\mu_1 &= \sup \left\{ \int_A g \, d\mu_1 : 0 \leq g \leq f, g < \infty \text{ simple and measurable} \right\} \\ &\leq C \sup \left\{ \int_A g \, d\mu_2 : 0 \leq g \leq f, g < \infty \text{ simple and measurable} \right\} \\ &= \int_A f \, d\mu_2. \end{aligned} \quad \square$$

Particularly, the Lemma implies that if $v \in L^p(X, \Sigma; d\mu_2)$, then $v \in L^p(X, \Sigma; d\mu_1)$. We apply the Lemma in order to obtain the following transformation result for Lebesgue integrable functions on the boundary of a strong Lipschitz domain. Recall that \mathcal{H}^s is merely an outer measure on $\mathcal{P}(\mathbb{R}^d)$ for any $d \geq 1$ but yields a measure when restricted to any σ -algebra that is contained in the \mathcal{H}^s -measurable sets, particularly, restricted to $\mathcal{B}(A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$.

Lemma A.1.2 (Transformations and $L^p(\partial\Omega)$) *Let $d \geq 1$, Ω_1 and $\Omega_2 \subset \mathbb{R}^d$ be bounded strong Lipschitz domains and let $\Theta: \partial\Omega_1 \rightarrow \partial\Omega_2$ be a bi-Lipschitz mapping with $\Theta(\partial\Omega_1) = \partial\Omega_2$. Then $v \circ \Theta \in L^p(\partial\Omega_1, \mathcal{B}(\partial\Omega_1); \mathcal{H}^{d-1})$ for all $v \in L^p(\partial\Omega_2, \mathcal{B}(\partial\Omega_2); \mathcal{H}^{d-1})$, and there exists a constant $C = C(\Theta)$ such that*

$$\|v \circ \Theta\|_{L^p(\partial\Omega_1, \mathcal{B}(\partial\Omega_1); \mathcal{H}^{d-1})} \leq C \|v\|_{L^p(\partial\Omega_2, \mathcal{B}(\partial\Omega_2); \mathcal{H}^{d-1})}.$$

Proof. Since $\partial\Omega_i \in \mathcal{B}(\mathbb{R}^d)$, we know that $\mathcal{B}(\partial\Omega_i) \subset \mathcal{B}(\mathbb{R}^d)$ and the restriction of the outer measure \mathcal{H}^{d-1} to $\mathcal{B}(\partial\Omega_i)$, $i = 1, 2$, is in fact a measure, cf. [64, Thm. 2.1]. Since Θ is bi-Lipschitz, the functions Θ and Θ^{-1} are also $(\mathcal{B}(\partial\Omega_1), \mathcal{B}(\partial\Omega_2))$ - and $(\mathcal{B}(\partial\Omega_2), \mathcal{B}(\partial\Omega_1))$ -measurable, respectively, and map measurable sets into measurable sets as well.

Further, we can extend Θ^{-1} to a Lipschitz continuous function on \mathbb{R}^d , see Theorem A.3.9, and [65, Thm. 2.8] implies that

$$\mathcal{H}^{d-1}(\Theta^{-1}(A)) \leq L_{\Theta^{-1}}^{d-1} \mathcal{H}^{d-1}(A) \quad \forall A \in \mathcal{B}(\partial\Omega_2)$$

and for every Lipschitz constant $L_{\Theta^{-1}}$ of Θ^{-1} .

The set function $\mathcal{H}^{d-1} \circ \Theta^{-1}$ is a measure on $\mathcal{B}(\partial\Omega_2)$ and known as the *image measure* of \mathcal{H}^{d-1} under Θ , e.g., [30, Sec. 3.6], and we know that $v \in L^p(\partial\Omega_1, \mathcal{B}(\partial\Omega_1), \mathcal{H}^{d-1})$ if and only if $v \circ \Theta \in L^p(\partial\Omega_2, \mathcal{B}(\partial\Omega_2), \mathcal{H}^{d-1} \circ \Theta)$. Further, we can apply the transformation formula [30, Thm. 3.6.1], [201, Satz IV.9.14] and Lemma A.1.1 to obtain

$$\int_{\partial\Omega_1} |v \circ \Theta|^p \, d\mathcal{H}^{d-1} = \int_{\partial\Omega_2} |v|^p \, d\mathcal{H}^{d-1} \circ \Theta^{-1} \leq L_{\Theta^{-1}}^{d-1} \int_{\partial\Omega_2} |v|^p \, d\mathcal{H}^{d-1}. \quad \square$$

A similar transformation result for open domains can be found in [148, Lem. 2.3.1].

A.2 Sobolev Traces

This section is dedicated to some trace and trace space related facts about Sobolev functions that are used repeatedly.

Definition A.2.1 (Positive and Negative Parts) Let $d \geq 1$, $\Omega \subset \mathbb{R}^d$ and $f: \Omega \rightarrow \overline{\mathbb{R}}$. Then $f^+, f^-: \Omega \rightarrow [0, \infty]$ with

$$f^+(\mathbf{x}) := \max(0, f(\mathbf{x})), \quad f^-(\mathbf{x}) := \max(0, -f(\mathbf{x}))$$

are the positive and negative part of f . We write $|f| := f^+ + f^-$.

Theorem A.2.2 (Trace Operator [148, Sec. 2]) Let $d \geq 1$, $1 \leq p \leq d$ and $\Omega \subset \mathbb{R}^d$ be a bounded strong Lipschitz domain. There exists a unique, linear and bounded operator

$$\text{tr}: W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega) \quad \text{with} \quad \text{tr}(y) = y|_{\partial\Omega} \quad \forall y \in W^{1,p}(\Omega) \cap C(\overline{\Omega}).$$

When $p > 1$, then $\text{tr}: (W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}(\Omega)}) \rightarrow (W^{1-1/p,p}(\partial\Omega), |\cdot|_{W^{1-1/p,p}(\partial\Omega)})$ is linear and bounded, and there exists a linear, bounded right inverse

$$\text{tr}^{-1}: (W^{1-1/p,p}(\partial\Omega), |\cdot|_{W^{1-1/p,p}(\partial\Omega)}) \rightarrow (W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}(\Omega)})$$

$$\text{with } \text{tr}(\text{tr}^{-1}(v)) = v \quad \forall v \in W^{1-1/p,p}(\partial\Omega).$$

Proof. The existence of the trace operator is proven in [148, Thm. 2.4.2, Thm. 2.4.6] and the equality to the restriction

$$\text{tr}(y) = y|_{\partial\Omega} \quad \forall y \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$$

can be found in [148, Ex. 2.4.1] and is also implied by [6, Thm. A 6.6].

For $p > 1$, [148, Thm. 2.5.5] implies that $\text{tr}(W^{1,p}(\Omega)) \subset W^{1-1/p,p}(\partial\Omega)$. Surjectivity on $W^{1-1/p,p}(\partial\Omega)$ follows from the existence of the right inverse [148, Thm. 2.5.7]. Note that [148] introduces $W^{s,p}(\partial\Omega)$ for non-integer $s \in (0, 1)$ using the norm introduced by the atlas of Ω , which is equivalent to the Sobolev-Slobodeckij norm (1.1). \square

For the case of $d < p < \infty$, the Sobolev embedding theorem immediately yields the existence of traces, see [3, Thm. 4.12].

Using the trace operator, we can easily show equivalence of the Sobolev-Slobodeckij norm and the quotient norm on $W^{1-1/p,p}(\partial\Omega)$.

Lemma A.2.3 (Equivalent Norms on the Trace Space) Let $d \geq 1$, $1 < p \leq d$ and $\Omega \subset \mathbb{R}^d$ be a bounded strong Lipschitz domain. Then there exist constants $c, C > 0$ for the Sobolev-Slobodeckij norm and the quotient norm on $W^{1-1/p,p}(\partial\Omega)$, i.e.,

$$|v|_{W^{1-1/p,p}(\partial\Omega)} = \left(\|v\|_{L^p(\Omega)}^p + \int_{\partial\Omega} \int_{\partial\Omega} \frac{|v(\mathbf{x}) - v(\mathbf{y})|^p}{|\mathbf{x} - \mathbf{y}|^{d+p-2}} d\mathcal{H}^{d-1}(\mathbf{x}) d\mathcal{H}^{d-1}(\mathbf{y}) \right)^{\frac{1}{p}},$$

$$\|v\|_{W^{1-1/p,p}(\partial\Omega)} = \inf_{\substack{w \in W^{1,p}(\Omega) \\ \text{tr}(w) = v}} \|w\|_{W^{1,p}(\Omega)},$$

such that

$$c \|v\|_{W^{1-1/p,p}(\partial\Omega)} \leq |v|_{W^{1-1/p,p}(\partial\Omega)} \leq C \|v\|_{W^{1-1/p,p}(\partial\Omega)} \quad \forall v \in W^{1-1/p,p}(\partial\Omega).$$

Proof. Let $v \in W^{1-\frac{1}{p},p}(\partial\Omega)$ be fixed. For all $w \in W^{1,p}(\Omega)$ with $\text{tr}(w) = v$, we obtain

$$\begin{aligned} \inf_{\substack{w \in W^{1,p}(\Omega) \\ \text{tr}(w)=v}} \|w\|_{W^{1,p}(\Omega)} &\leq \|\text{tr}^{-1} v\|_{W^{1,p}(\Omega)} \leq \|\text{tr}^{-1}\|_{\mathcal{L}(W^{1-\frac{1}{p},p}(\partial\Omega), W^{1,p}(\Omega))} |v|_{W^{1-\frac{1}{p},p}(\partial\Omega)} \\ &= \|\text{tr}^{-1}\|_{\mathcal{L}(W^{1-\frac{1}{p},p}(\partial\Omega), W^{1,p}(\Omega))} |\text{tr}(w)|_{W^{1-\frac{1}{p},p}(\partial\Omega)} \\ &\leq \|\text{tr}^{-1}\|_{\mathcal{L}(W^{1-\frac{1}{p},p}(\partial\Omega), W^{1,p}(\Omega))} \|\text{tr}\|_{\mathcal{L}(W^{1,p}(\Omega), W^{1-\frac{1}{p},p}(\partial\Omega))} \|w\|_{W^{1,p}(\Omega)}, \end{aligned}$$

where the fractional Sobolev space in the linear operator norms is understood to be equipped with the Sobolev-Slobodeckij norm, see Theorem A.2.2. Accordingly, taking the infimum over all these w on the right hand side completes the proof. \square

Lemma A.2.4 (Truncation on $W^{1,p}(\Omega)$) *Let $d \geq 1$, $1 \leq p < \infty$, $\Omega \subset \mathbb{R}^d$ be an open set and $w \in W^{1,p}(\Omega)$, then:*

1. $w^+, w^-, |w| \in W^{1,p}(\Omega)$
2. $w_n^+ \rightarrow w^+, w_n^- \rightarrow w^-, |w_n| \rightarrow |w|$, for any sequence $w_n \rightarrow w$

Proof. Part 1. can be found in [11, Sec. 5.8.1], where $\nabla(\max(0, v)) = \mathbf{1}_{\{v>0\}} \nabla v$ is additionally proven. For Part 2., it suffices to show the first convergence. We closely follow the proof of [28, Lem. 3.15] for the case of $p = 2$ but account for the variable exponent. Let $w_n \rightarrow w$ in $W^{1,p}(\Omega)$, then we decompose $\Omega = A_n \cup B_n \cup C_n \cup D_n$ with

$$\begin{aligned} A_n &:= \{ \mathbf{x} \in \Omega : w_n(\mathbf{x}) > 0, w(\mathbf{x}) \geq 0 \}, & B_n &:= \{ \mathbf{x} \in \Omega : w_n(\mathbf{x}) \leq 0, w(\mathbf{x}) > 0 \}, \\ C_n &:= \{ \mathbf{x} \in \Omega : w_n(\mathbf{x}) > 0, w(\mathbf{x}) < 0 \}, & D_n &:= \{ \mathbf{x} \in \Omega : w_n(\mathbf{x}) \leq 0, w(\mathbf{x}) \leq 0 \}. \end{aligned}$$

Then we obtain

$$\begin{aligned} \|w_n^+ - w^+\|_{W^{1,p}(\Omega)}^p &= \int_{\Omega} |w_n^+ - w^+|^p \, d\mathcal{L}^d + \int_{\Omega} \|\nabla w_n^+ - \nabla w^+\|_p^p \, d\mathcal{L}^d \\ &= \sum_{I_n \in \{A_n, B_n, C_n, D_n\}} \int_{I_n} |w_n^+ - w^+|^p + \|\nabla w_n^+ - \nabla w^+\|_p^p \, d\mathcal{L}^d \\ &\leq \|w_n - w\|_{W^{1,p}(\Omega)}^p + \int_{B_n} |w|^p + \|\nabla w\|_p^p \, d\mathcal{L}^d \\ &\quad + \int_{C_n} |w_n|^p + \|\nabla w_n\|_p^p \, d\mathcal{L}^d \\ &\leq \|w_n - w\|_{W^{1,p}(\Omega)}^p + \int_{B_n} |w|^p + \|\nabla w\|_p^p \, d\mathcal{L}^d \\ &\quad + \left(\int_{C_n} |w_n - w|^p \, d\mathcal{L}^{d\frac{1}{p}} + \int_{C_n} |w|^p \, d\mathcal{L}^{d\frac{1}{p}} \right)^p \\ &\quad + \left(\int_{C_n} \|\nabla w_n - \nabla w\|_p^p \, d\mathcal{L}^{d\frac{1}{p}} + \int_{C_n} \|\nabla w\|_p^p \, d\mathcal{L}^{d\frac{1}{p}} \right)^p \end{aligned}$$

$$\begin{aligned}
&\leq \|w_n - w\|_{W^{1,p}(\Omega)}^p + \int_{B_n} |w|^p + \|\nabla w\|_p^p \, d\mathcal{L}^d \\
&\quad + \left(\|w_n - w\|_{W^{1,p}(\Omega)} + \int_{C_n} |w|^p \, d\mathcal{L}^{d\frac{1}{p}} \right)^p \\
&\quad + \left(\|w_n - w\|_{W^{1,p}(\Omega)} + \int_{C_n} \|\nabla w\|_p^p \, d\mathcal{L}^{d\frac{1}{p}} \right)^p.
\end{aligned}$$

[28, Lem 3.13] implies $\mathcal{L}^d(B_n) \rightarrow 0$ and $\mathcal{L}^d(C_n) \rightarrow 0$, and [6, Lem. A 1.17] concludes the proof for the positive part. We apply the result to $w_n^- = (-w_n)^+$ and $|w_n| = w_n^+ - w_n^-$ to finalize the proof. \square

Lemma A.2.5 (Trace Operator and Truncation) *Let $d \geq 1$, $1 < p \leq d$ and $\Omega \subset \mathbb{R}^d$ be a bounded strong Lipschitz domain. Then $\text{tr}(w^+) = \text{tr}(w)^+ \mathcal{H}^{d-1}$ -a.e. in $\partial\Omega$ for all w in $W^{1,p}(\Omega)$.*

Proof. Let $w \in W^{1,p}(\Omega)$, then we can find a sequence $(w_n) \subset W^{1,p}(\Omega) \cap C(\bar{\Omega})$ with $w_n \rightarrow w$ [3, Thm. 3.22]. By Theorem A.2.2, we know that $\text{tr}(w_n^+) = \text{tr}(w_n)^+$. Since $|v^+| \leq |v|$ for all $\forall v \in W^{1-1/p,p}(\partial\Omega)$, we know $\text{tr}(w_n)^+, \text{tr}(w)^+ \in L^p(\partial\Omega)$ and

$$\begin{aligned}
\|\text{tr}(w_n^+) - \text{tr}(w)^+\|_{L^p(\partial\Omega)} &= \|\text{tr}(w_n)^+ - \text{tr}(w)^+\|_{L^p(\partial\Omega)} \\
&= \frac{1}{2} \|\text{tr}(w_n) + |\text{tr}(w_n)| - \text{tr}(w) - |\text{tr}(w)|\|_{L^p(\partial\Omega)} \\
&\leq \|\text{tr}(w_n) - \text{tr}(w)\|_{L^p(\partial\Omega)} \\
&\leq \|\text{tr}\|_{\mathcal{L}(W^{1,p}(\Omega), L^p(\partial\Omega))} \|w_n - w\|_{W^{1,p}(\Omega)} \rightarrow 0,
\end{aligned}$$

which provides a subsequence of $\text{tr}(w_n^+)$ (denoted by the same symbol) that converges to $\text{tr}(w)^+ \mathcal{H}^{d-1}$ -a.e. on $\partial\Omega$. Lemma A.2.4 and the continuity of the trace operator on the other hand imply that $\text{tr}(w_n^+) \rightarrow \text{tr}(w^+)$ in $W^{1-1/p,p}(\partial\Omega)$ and therefore also in $L^p(\partial\Omega)$ for the same subsequence. Extracting another subsequence of $\text{tr}(w_n^+)$ that converges to $\text{tr}(w^+)$ \mathcal{H}^{d-1} -a.e. in $\partial\Omega$ yields that $\text{tr}(w^+) = \text{tr}(w)^+$ holds \mathcal{H}^{d-1} -a.e. in $\partial\Omega$. \square

Confer [28, Lem. 3.27] for a similar result in the case $p = 2$.

Corollary A.2.6 (Truncation on $W^{1-1/p,p}(\partial\Omega)$) *Let $d \geq 1$, $1 < p \leq d$, $\Omega \subset \mathbb{R}^d$ be a bounded strong Lipschitz domain and $v \in W^{1-1/p,p}(\partial\Omega)$. Then*

$$v^+ \in W^{1-1/p,p}(\partial\Omega) \text{ and } \|v^+\|_{W^{1-1/p,p}(\partial\Omega)} \leq \|v\|_{W^{1-1/p,p}(\partial\Omega)}.$$

Proof. Theorem A.2.2 and Lemma A.2.5 imply $v^+ \in W^{1-1/p,p}(\partial\Omega)$, because $\text{tr}^{-1}(v)^+$ is a preimage in $W^{1,p}(\Omega)$ whose image under the trace operator is in $W^{1-1/p,p}(\partial\Omega)$. Additionally,

$$\|v^+\|_{W^{1-1/p,p}(\partial\Omega)} = \inf_{\substack{w \in W^{1,p}(\Omega) \\ \text{tr } w = v^+}} \|w\|_{W^{1,p}(\Omega)} \leq \inf_{\substack{w \in W^{1,p}(\Omega) \\ \text{tr } w = v}} \|w^+\|_{W^{1,p}(\Omega)} \leq \|v\|_{W^{1-1/p,p}(\partial\Omega)}. \quad \square$$

Remark A.2.7 *The result of the previous Corollary can also be obtained directly for the equivalent Sobolev-Slobodeckij norm, since*

$$|v^+| \leq |v| \text{ and } |v^+(\mathbf{x}) - v^+(\mathbf{y})| \leq |v(\mathbf{x}) - v(\mathbf{y})| \quad \forall \mathbf{x}, \mathbf{y} \in \partial\Omega.$$

Accordingly, the result holds in $W^{s,p}(\Omega)$ equipped with the Sobolev Slobodeckij norm for arbitrary $0 < s < 1$, $p \geq 1$ and open $\Omega \subset \mathbb{R}^d$ as well.

In addition to the truncation results, we obtain stability results for bi-Lipschitz coordinate transformations on the boundary of Lipschitz domains.

Lemma A.2.8 (Trace Operator and Transformations) *Let $d \geq 1$, Ω_1 and $\Omega_2 \subset \mathbb{R}^d$ be bounded strong Lipschitz domains and let $\Theta: \overline{\Omega_1} \rightarrow \overline{\Omega_2}$ be a bi-Lipschitz mapping with $\Theta(\overline{\Omega_1}) = \overline{\Omega_2}$ and $\Theta(\Omega_1) = \Omega_2$. Then $\text{tr}(w \circ \Theta) = \text{tr}(w) \circ \Theta$ \mathcal{H}^{d-1} -a.e. in $\partial\Omega_1$ for all w in $W^{1,p}(\Omega_2)$.*

Proof. Note that the trace operators in the claim denote the two trace operators defined for each of the domains Ω_1 and Ω_2 .

Let $w \in W^{1,p}(\Omega_2)$. We obtain that $w \circ \Theta \in W^{1,p}(\Omega_1)$ as well as the existence of a positive constant $C = C(\Theta)$ such that

$$\|w \circ \Theta\|_{W^{1,p}(\Omega_1)} \leq C \|w\|_{W^{1,p}(\Omega_2)}$$

from [148, Lem. 2.3.2], cf. also [208, Thm. 2.2.2] and [135, P. 46].

We find a sequence $(w_n) \subset W^{1,p}(\Omega_2) \cap C(\overline{\Omega_2})$ with $w_n \rightarrow w$ in $W^{1,p}(\Omega_2)$ and therefore

$$\begin{aligned} w_n \circ \Theta &\rightarrow w \circ \Theta && \text{in } W^{1,p}(\Omega_1), \\ \text{tr}(w_n \circ \Theta) &\rightarrow \text{tr}(w \circ \Theta) && \text{in } W^{1-\frac{1}{p},p}(\partial\Omega_1). \end{aligned}$$

Therefore, we can find a subsequence of (w_n) (still denoted by the same symbol) such that $\text{tr}(w_n \circ \Theta) \rightarrow \text{tr}(w \circ \Theta)$ \mathcal{H}^{d-1} -a.e. in $\partial\Omega_1$.

Since $w_n \circ \Theta \in W^{1,p}(\Omega_1) \cap C(\overline{\Omega_1})$, we have that

$$\text{tr}(w_n \circ \Theta) = w_n|_{\partial\Omega_2} \circ \Theta|_{\partial\Omega_1} = \text{tr}(w_n) \circ \Theta.$$

Lemma A.1.2 implies that $\text{tr}(w_n) \circ \Theta \in L^p(\partial\Omega_1)$ and $\text{tr}(w_n) \circ \Theta \rightarrow \text{tr}(w) \circ \Theta$ in $L^p(\partial\Omega_1)$. Passing to another subsequence, we obtain that

$$\text{tr}(w_n) \circ \Theta \rightarrow \text{tr}(w) \circ \Theta \quad \mathcal{H}^{d-1}\text{-a.e. in } \partial\Omega_1,$$

and therefore $\text{tr}(w \circ \Theta) = \text{tr}(w) \circ \Theta$ \mathcal{H}^{d-1} -a.e. in $\partial\Omega_1$. □

Accordingly, we obtain the following norm estimate for transformations.

Lemma A.2.9 (Transformations for $W^{1-1/p,p}(\partial\Omega)$) *Let $d \geq 1$, Ω_1 and $\Omega_2 \subset \mathbb{R}^d$ be bounded strong Lipschitz domains and let $\Theta: \overline{\Omega_1} \rightarrow \overline{\Omega_2}$ be a bi-Lipschitz mapping with $\Theta(\overline{\Omega_1}) = \overline{\Omega_2}$ and $\Theta(\Omega_1) = \Omega_2$. Then there exists a positive constant $C = C(\Theta)$ such that*

$$v \circ \Theta \in W^{1-\frac{1}{p},p}(\partial\Omega_1) \quad \text{and} \quad \|v \circ \Theta\|_{W^{1-\frac{1}{p},p}(\partial\Omega_1)} \leq C \|v\|_{W^{1-\frac{1}{p},p}(\partial\Omega_2)}$$

for all $v \in W^{1-\frac{1}{p},p}(\partial\Omega_2)$.

Proof. For every $w \in W^{1,p}(\Omega_2)$, [148, Lem. 2.3.2] ensures that $w \circ \Theta \in W^{1,p}(\Omega_1)$ and the existence of a constant $C = C(\Theta)$ such that $\|w \circ \Theta\|_{W^{1,p}(\Omega_1)} \leq C \|w\|_{W^{1,p}(\Omega_2)}$.

Let $v \in W^{1-\frac{1}{p},p}(\partial\Omega_2)$. Then there exists an inverse image $w = \text{tr}^{-1}(v) \in W^{1,p}(\Omega_2)$, for which $\text{tr}(w \circ \Theta) = \text{tr}(w) \circ \Theta = v \circ \Theta$ due to Lemma A.2.8. The range of the trace operator implies that $v \circ \Theta \in W^{1-\frac{1}{p},p}(\partial\Omega)$, and we compute

$$\begin{aligned} \|v \circ \Theta\|_{W^{1-\frac{1}{p},p}(\partial\Omega_1)} &= \inf_{w \in W^{1,p}(\Omega_1) : \text{tr}(w) = v \circ \Theta} \|w\|_{W^{1,p}(\Omega_1)} \\ &\leq \inf_{w \in W^{1,p}(\Omega_2) : \text{tr}(w) = v} \|w \circ \Theta\|_{W^{1,p}(\Omega_1)} \leq C \|v\|_{W^{1-\frac{1}{p},p}(\partial\Omega_2)} \quad \square \end{aligned}$$

A more detailed discussion of trace theorems in the context of the analysis of contact problems can be found in [110, Sec. 5.3].

A.3 Regularity of the Contact Normal

As mentioned in the modeling chapter (Chapter 2), a commonly employed contact model assumes $C^{1,1}$ -regularity of the domain, i.e., a piecewisely continuously differentiable boundary, and uses the geometric normal as the contact normal in the Signorini condition. The geometric normal is Lipschitz continuous and allows for a normal trace operator. In the model used in this thesis, the geometry of the domain and the regularity of the contact normal decouple, and the regularity of the boundary can be reduced to $C^{0,1}$. We assume Lipschitz continuity of the contact normal for two reasons, the first being that we depend on the existence of a well-defined trace operator as well. The second reason is that we can rely on the extension results of standard Lipschitz analysis to find an extension that allows for the technical transformations in the transfer of polyhedricity between sets of scalar and vector valued Sobolev functions. In the literature, the required regularity of these extensions can be found to be $W^{1,\infty}(\Omega)$ instead of $C^{0,1}(\Omega)$, cf. [181, P. 212] and [143, P. 269]. This section will be used for an overview of the relationship between Lipschitz functions and $W^{1,\infty}$ -functions, which is well understood. We present the required extension results of the literature and follow [87] for the majority of this section. We start out with the basic definitions required for the analysis.

Definition A.3.1 (*C*-Quasi Convexity) *A set $\Omega \subset \mathbb{R}^d$ is called C -quasi convex for a constant $C \geq 1$ if each two points $\mathbf{x}, \mathbf{y} \in \Omega$ can be connected by a rectifiable path $\gamma: [0, 1] \rightarrow \Omega$ whose length satisfies*

$$\text{len}(\gamma) \leq C \|\mathbf{x} - \mathbf{y}\|.$$

The set Ω is called quasi convex if it is C -quasi convex for some $C \geq 1$.

Recall that the length of a rectifiable path is defined as

$$\sup \sum_{k=0}^{N-1} \|\gamma(t_k) - \gamma(t_{k+1})\| < \infty$$

with the supremum taken over all partitions $0 = t_0 < \dots < t_N = 1$ of $[0, 1]$.

Remark A.3.2 *The nomenclature in Definition A.3.1 may be misleading within the scope of this thesis. The notion of C -quasi convexity is in no way based on or connected to the “quasi” properties of capacity theory in Chapter 3. In the literature, the property of C -quasi convexity is oftentimes not given a name at all ([51, 121]). Other names and related concepts include bounded turning and uniform domains ([5, 36]) as well as chord arc, Lavrentiev and quasi smooth ([60]) sets or domains. The name “ C -quasi convex” fits this section best and is most consistently used ([83, 87]).*

We consider three distinct notions of Lipschitz continuity of different strength.

Definition A.3.3 (Lipschitz Continuity) *Let $\mathbf{f}: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ be a function, and let $L \geq 0$ be a constant.*

- (a) *The function \mathbf{f} is called L -Lipschitz (continuous) if $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|$ for all \mathbf{x}, \mathbf{y} in Ω .*
- (b) *The function \mathbf{f} is called locally L -Lipschitz (continuous) if every point \mathbf{x} in Ω has a relative neighborhood on which \mathbf{f} is L -Lipschitz continuous.*
- (c) *The function \mathbf{f} is called weakly locally Lipschitz (continuous) if every point \mathbf{x} in Ω has a relative neighborhood on which \mathbf{f} is $L(\mathbf{x})$ -Lipschitz continuous for a constant $L = L(\mathbf{x}) \geq 0$ that may depend on \mathbf{x} .*

It is essential that the Lipschitz constant in the definition of local L -Lipschitz continuity remains constant in all neighborhoods. Before we address the connection between Lipschitz functions and $W^{1,\infty}$ -functions, we quickly prove the following observation.

Lemma A.3.4 *Let $\{I_i : i = 1, \dots, M\}$ be a finite covering of $[0, 1]$ by open intervals I_i . Then for each pair of points x, y in $[0, 1]$ satisfying*

$$0 < |x - y| < \min_{i,j: I_i \cap I_j \neq \emptyset} \text{diam}(I_i \cap I_j), \quad (\text{A.1})$$

there exists an index $i \in \{1, \dots, M\}$ such that both s and t are contained in I_i .

Proof. Assuming $x < y$ and that no interval contains both x and y , set

$$I_+ := \max_{i: x \in I_i} (\sup(I_i)) > x, \quad I_- := \min_{i: y \in I_i} (\inf(I_i)) < y, \quad (\text{A.2})$$

and i_+, i_- as the indices of any two intervals containing x, y whose right or left interval boundary is I_+ or I_- , respectively. If $I_- < I_+$, we obtain a contradiction to (A.1) because $\text{diam}(I_{i_+} \cap I_{i_-})$ is strictly less than the minimum of all diameters of nonempty intersections. When $I_- \geq I_+$, however, $I_+ \in [0, 1]$ is contained in a third open interval I , which by (A.2) contains neither x nor y , therefore $\text{diam}(I \cap I_{i_+}) < |x - y|$ contradicts (A.1). \square

The first result from Lipschitz analysis is an easy-to-see connection between Lipschitz and locally Lipschitz functions on quasi convex sets. Since the proof is left as an exercise in both [87, Lem. 2.2] and [189, P. 154] and could not be found elsewhere, it is included here for the sake of completeness.

Lemma A.3.5 (E.g. [87, Lem. 2.2]) *Let $d, n \geq 1$ and $C, L > 0$. If $\Omega \subset \mathbb{R}^d$ is C -quasi convex, then every locally L -Lipschitz function $\mathbf{f}: \Omega \rightarrow \mathbb{R}^n$ is CL -Lipschitz on Ω .*

Proof. Let $\mathbf{x}, \mathbf{y} \in \Omega$. By Definition A.3.1, we have a path $\gamma: [0, 1] \rightarrow \Omega$ that connects \mathbf{x} and \mathbf{y} with $\text{len}(\gamma) \leq C \|\mathbf{x} - \mathbf{y}\|$. We show the existence of a partition $0 = t_0 < t_1 < \dots < t_N = 1$ of $[0, 1]$, where for each k in $\{0, \dots, N-1\}$ the points $\gamma(t_k)$ and $\gamma(t_{k+1})$ are contained in a neighborhood of a point on the path on which \mathbf{f} is L -Lipschitz.

For each $t \in [0, 1]$, there exists a radius $\varepsilon_{\gamma(t)} > 0$ such that \mathbf{f} is L -Lipschitz on the set $B(\varepsilon_{\gamma(t)}, \gamma(t)) \cap \Omega$. From continuity of γ , we obtain corresponding radii $\delta_t > 0$ and open intervals $I_t = (t - \delta_t, t + \delta_t)$ with $\gamma(I_t) \subset B(\varepsilon_{\gamma(t)}, \gamma(t)) \cap \Omega$. Compactness of the interval $[0, 1]$ yields a finite, open covering $\{I_i : i = 1, \dots, M\}$ of $[0, 1]$, and we fix any equidistant partition $0 = t_0 < t_1 < \dots < t_N = 1$ of $[0, 1]$ with

$$0 < |t_{k+1} - t_k| < \min_{i,j: I_i \cap I_j \neq \emptyset} \text{diam}(I_i \cap I_j).$$

Lemma A.3.4 ensures that for all $k = 0, \dots, N-1$, there exists an interval from the finite covering that contains both t_k and t_{k+1} , meaning that $\gamma(t_k)$ and $\gamma(t_{k+1})$ are contained in a subset of Ω that \mathbf{f} is L -Lipschitz on. Consequently,

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \sum_{k=0}^{N-1} \|\mathbf{f}(\gamma(t_{k+1})) - \mathbf{f}(\gamma(t_k))\| \leq L \sum_{k=0}^{N-1} \|\gamma(t_{k+1}) - \gamma(t_k)\| \leq LC \|\mathbf{x} - \mathbf{y}\|.$$

□

The well known relation between real valued (locally) Lipschitz functions and functions in $W_{(\text{loc})}^{1,\infty}(\Omega)$ can now be summarized as follows:

Theorem A.3.6 ($C_{(\text{loc})}^{0,1}$ and $W_{(\text{loc})}^{1,\infty}$) *Let $d \geq 2$ and $\Omega \subset \mathbb{R}^d$ be a domain.*

- (a) $W_{\text{loc}}^{1,\infty}(\Omega) = \{\mathbf{f} : \Omega \rightarrow \mathbb{R} : \mathbf{f} \text{ is weakly locally Lipschitz}\}$
- (b) $W_{\text{loc}}^{1,\infty}(\Omega) = \{\mathbf{f} \in C_{\text{loc}}^{0,1}(\Omega) : \mathbf{f} \text{ is bounded}\}$

When Ω is additionally quasi convex, then additionally:

- (c) $W^{1,\infty}(\Omega) = \{\mathbf{f} \in C_{\text{loc}}^{0,1}(\Omega) : \mathbf{f} \text{ is bounded}\} = \{\mathbf{f} \in C^{0,1}(\Omega) : \mathbf{f} \text{ is bounded}\}$

Proof. Equality (a) can be found in [65, Thm. 4.5]. The equalities (b) and (c) can be found in [87, Thm. 4.1], [189, P. 154], and part (c) follows from Lemma A.3.5. □

Accordingly, the requirements of $W^{1,\infty}$ -regularity and Lipschitz continuity of the extensions coincides for quasi convex domains. Note that the boundedness restrictions in Theorem A.3.6 are not only relevant for unbounded Ω . While any Lipschitz function on a bounded set is obviously bounded, locally Lipschitz functions do not share this property. The following example is a minor modification of an example communicated by Tero Kilpeläinen at the University of Jyväskylä, Finland, in connection with the capacity theory presented in Chapter 3.

Example A.3.7 (Loc. Lipschitz, Unbounded Function) Consider the bounded domain $\Omega := (0, 1)^2 \setminus (I \cup J) \subset \mathbb{R}^2$, where the set difference removes all line segments

$$I = \bigcup_{k=0}^{\infty} \left\{ 2^{-2k-1} \right\} \times \left[0, \frac{2}{3} \right], \quad J = \bigcup_{k=0}^{\infty} \left\{ 2^{-2k} \right\} \times \left[\frac{1}{3}, 1 \right]$$

from the unit square. Define the function $f: \Omega \rightarrow \mathbb{R}$ as

$$f(x_1, x_2) := \begin{cases} 2k, & (x_1, x_2) \in (2^{-2k-1}, 2^{-2k+1}) \times (0, \frac{1}{3}) \\ 2k + (-1)^k 3(x_2 - \frac{1}{3}), & (x_1, x_2) \in (2^{-2k-1}, 2^{-2k}) \times (\frac{1}{3}, \frac{2}{3}) \\ 2k + (-1)^k 3(x_2 - \frac{1}{3}), & (x_1, x_2) \in (2^{-2k}, 2^{-2k+1}) \times (\frac{1}{3}, \frac{2}{3}) \\ 2k + 1, & (x_1, x_2) \in (2^{-2k-2}, 2^{-2k}) \times (\frac{2}{3}, 1) \end{cases}$$

for $k \in \mathbb{N}$, which is piecewise constant for all (x_1, x_2) with $x_2 \notin (\frac{1}{3}, \frac{2}{3})$ and increases linearly whenever $x_2 \in (\frac{1}{3}, \frac{2}{3})$, i.e., its graph can be thought of as a "parking-garage" ramp with decreasing width as one progresses through the domain from right to left.

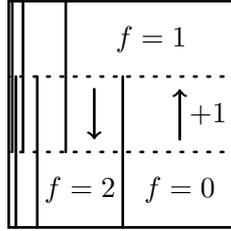


Figure A.1: Locally Lipschitz and unbounded f with "ramp-graph"

The function clearly is locally $\frac{1}{3}$ -Lipschitz continuous but unbounded.

Remark A.3.8 Due to Theorem A.3.6 (c), it is clear that an example like A.3.7 can only be obtained on a non-quasi-convex domain, which Ω in the example clearly is. Other examples of non-quasi-convex domains are the slit domain in two dimensions and the domain between two infinite spirals, cf. [121, Ex. 1.2.13]. On both, bounded locally Lipschitz functions that are not Lipschitz on the entire domain can be constructed.

The central McShane-Whitney-Kirszbraun Extension Theorem ensures the existence of the required extensions of the Lipschitz continuous contact normal on the contact boundary.

Theorem A.3.9 (Extension Theorem) Let $d, \tilde{d} \geq 1$, $\Omega \subset \mathbb{R}^d$ and $f: \Omega \rightarrow \mathbb{R}^{\tilde{d}}$ be an L -Lipschitz function. Then there exists an L -Lipschitz function $F: \mathbb{R}^d \rightarrow \mathbb{R}^{\tilde{d}}$ with $F|_{\Omega} = f$.

Accordingly, we can extend any Lipschitz function between euclidean spaces of arbitrary finite-dimensional euclidean space to be defined on the entire space without altering the Lipschitz constant. See [87, Thm. 2.3] for an easily readable proof of the case where $\tilde{d} = 1$ and, e.g., [66, Thm. 2.10.43] for the general case. Since the extension is necessarily unique on $\overline{\Omega}$, the extension theorem especially implies that $C^{0,1}(\Omega)$ is equivalent to $C^{0,1}(\overline{\Omega})$ for any set Ω , i.e., Lipschitz continuity of ν_{Φ} on Γ_C implies that $\nu_{\Phi} \in C^{0,1}(\overline{\Gamma_C})$ as well.

We have seen that Lipschitz continuity generally is a stronger assumption than requiring $W^{1,\infty}$ -regularity and that the notions coincide for quasi convex domains. On bounded

strong Lipschitz domains, Theorem A.3.6 (c) in fact transfers to the boundary sense as well. Recall the definition of a Lipschitz domain, e.g., [51, 65, 110].

Definition A.3.10 (Lipschitz Domain) *Let $d \geq 2$, $\Omega \subset \mathbb{R}^d$ be an open set. The boundary $\partial\Omega$ is said to be strongly Lipschitz if for each $\mathbf{x} \in \partial\Omega$ there exist $r > 0$, $L > 0$ and an L -Lipschitz continuous function $h \in C^{0,1}(\mathbb{R}^{d-1})$ such that in a local coordinate system, i.e., after rotation, we can write*

$$\Omega \cap Q(\mathbf{x}, r) = \{ \mathbf{y} = (\mathbf{y}', y_d) : h(\mathbf{y}') < y_d \} \cap Q(\mathbf{x}, r),$$

where

$$S(\mathbf{x}, r) := \left\{ \mathbf{y} \in \mathbb{R}^{d-1} : |x_i - y_i| < r, i = 1, \dots, d-1 \right\},$$

$$Q(\mathbf{x}, r) := S(\mathbf{x}, r) \times (x_d - r, x_d + r).$$

If Ω is a (bounded) domain whose boundary is strongly Lipschitz, then Ω is referred to as a (bounded) strong Lipschitz domain.

When the domain Ω is additionally bounded, its boundary is compact, so finitely many (rotated) open sets Q_i in Definition A.3.10 are sufficient to cover $\partial\Omega$. We assume (S_i, h_i) , $i = 1, \dots, M$ from the corresponding atlas and appropriate orthogonal transformations R_i in $\text{SO}(d)$ to be fixed. Then Theorem A.3.6 (c) is easily adapted to accommodate the boundary structure.

Corollary A.3.11 (Quasi Convex Contact Boundary) *Let $d \geq 2$, and let $\Omega \subset \mathbb{R}^d$ be a bounded strong Lipschitz domain with relatively open boundary section $\Gamma \subset \partial\Omega$. If Γ is quasi convex, then $C^{0,1}(\Gamma) = W^{1,\infty}(\Gamma)$.*

Proof. We show local Lipschitz continuity and boundedness of all functions in $W^{1,\infty}(\Gamma)$ on Γ and apply Lemma A.3.5. Note that $W^{1,\infty}(\Gamma)$ consists of all functions that are essentially bounded in the sense introduced by the atlas. Now let $f \in W^{1,\infty}(\Gamma)$. For all $i \in \{1, \dots, M\}$ with associated (S_i, h_i) and rotation R_i , define $\mathbf{g}_i: S_i \rightarrow \mathbb{R}^d$ as $\mathbf{g}_i(\mathbf{x}') := (\mathbf{x}', h_i(\mathbf{x}'))$, and note that its inverse is continuous. By definition (e.g. [110, Sec. 5.3]), we know that

$$f \circ R_i \circ \mathbf{g}_i \in W^{1,\infty}(S_i \cap \mathbf{g}_i^{-1}(R_i^{-1}\Gamma)) = \{ v \in C_{\text{loc}}^{0,1}(S_i \cap \mathbf{g}_i^{-1}(R_i^{-1}\Gamma)) : v \text{ bounded} \}$$

for all i in $\{1, \dots, M\}$. Take $L > 0$ as the maximum of the constants $L_i > 0$ such that $f \circ R_i \circ \mathbf{g}_i$ is locally L_i -Lipschitz continuous on its domain of definition in \mathbb{R}^{d-1} , and note that f is bounded.

Now let $\mathbf{x} \in \Gamma$ and $i \in \{1, \dots, M\}$ such that $\mathbf{x} = R_i(\mathbf{x}', h_i(\mathbf{x}'))$ for an $\mathbf{x}' \in S_i$. There exists $\delta > 0$ such that $f \circ R_i \circ \mathbf{g}_i$ is L -Lipschitz continuous on $B(\delta, \mathbf{x}') \subset S_i \cap \mathbf{g}_i^{-1}(R_i^{-1}\Gamma)$. As \mathbf{g}_i is the identity map in the first $d-1$ components,

$$B(\delta, \mathbf{g}_i(\mathbf{x}')) \cap R_i^{-1}\Gamma \subset \mathbf{g}_i(B(\delta, \mathbf{x}') \cap S_i).$$

Since R_i is an isometric rotation,

$$B(\delta, \mathbf{x}) \cap \Gamma = R_i(B(\delta, \mathbf{g}_i(\mathbf{x}')) \cap R_i^{-1}\Gamma) \subset R_i \mathbf{g}_i(B(\delta, \mathbf{x}') \cap S_i),$$

hence f is L -Lipschitz on the relative neighborhood $B(\delta, \mathbf{x}) \cap \Gamma$ of \mathbf{x} . Therefore, the function f is locally L -Lipschitz continuous on the quasi convex set Γ , and by Lemma A.3.5, it is Lipschitz continuous on Γ . \square

A crucial assumption in the previous corollary is that the boundary segment Γ itself is quasi convex. Generally, it may not even be chosen as path connected at all, even though the entire boundary may be quasi convex. On the other hand, Lipschitz domains with a boundary that is not path connected are covered if Γ is a quasi convex boundary segment.

Note that bounded, path connected strong Lipschitz domains as well as path connected (entire) boundaries of bounded strong Lipschitz domains are quasi convex [51, Cor. 1., Thm. 5.8].

Appendix B

Capacity Related Results

This chapter provides some additional facts about quasi open sets and quasi continuous functions, the exceptional sets in quasi continuity, quasi openness and quasi uniform convergence. These results, however interesting, are inessential to the analysis that was required in Chapter 3 and were therefore omitted in the latter. This chapter is stated within the abstract setting of Section 3.1 defined in Assumption 3.1.1.

B.1 Nested Exceptional Sets

The first observation concerns the Definitions 3.1.10, 3.1.11 and 3.1.12 of quasi continuous functions, quasi open sets and quasi uniform convergence. All three definitions rely on an exceptional set of arbitrarily small capacity. Naturally, these sets can be chosen in a non-increasing, nested manner.

Lemma B.1.1 *Let $0 < \delta < \varepsilon$, $v: X \rightarrow \overline{\mathbb{R}}$ be quasi continuous and $G_\delta, G_\varepsilon \in \mathcal{O}(X)$ such that*

$$\begin{aligned} \text{cap}(G_\delta) < \delta \quad \text{and} \quad v: X \setminus G_\delta \rightarrow \overline{\mathbb{R}} \text{ is continuous,} \\ \text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad v: X \setminus G_\varepsilon \rightarrow \overline{\mathbb{R}} \text{ is continuous.} \end{aligned}$$

Then $G_\delta \subset G_\varepsilon$ w.l.o.g.

Proof. From the monotonicity of the capacity, we obtain that

$$G_\delta \cap G_\varepsilon \in \mathcal{O}(X), \quad G_\delta \cap G_\varepsilon \subset G_\varepsilon \quad \text{and} \quad \text{cap}(G_\delta \cap G_\varepsilon) \leq \text{cap}(G_\delta) < \delta.$$

Since both $X \setminus G_\delta$ and $X \setminus G_\varepsilon$ are closed sets, basic topology yields that the restriction of v to the set $X \setminus (G_\delta \cap G_\varepsilon) = X \setminus G_\delta \cup X \setminus G_\varepsilon$ is continuous as well. \square

Lemma B.1.2 *Let $0 < \delta < \varepsilon$ and $A \subset X$ be quasi open $G_\delta, G_\varepsilon \in \mathcal{O}(X)$ such that*

$$\begin{aligned} \text{cap}(G_\delta) < \delta \quad \text{and} \quad A \cup G_\delta \in \mathcal{O}(X), \\ \text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad A \cup G_\varepsilon \in \mathcal{O}(X). \end{aligned}$$

Then $G_\delta \subset G_\varepsilon$ w.l.o.g.

Proof. From the monotonicity of the capacity, we obtain that

$$G_\delta \cap G_\varepsilon \in \mathcal{O}(X), \quad G_\delta \cap G_\varepsilon \subset G_\varepsilon \quad \text{and} \quad \text{cap}(G_\delta \cap G_\varepsilon) \leq \text{cap}(G_\varepsilon) < \varepsilon.$$

Additionally, $A \cup (G_\delta \cap G_\varepsilon) = A \cup G_\delta \cap A \cup G_\varepsilon \in \mathcal{O}(X)$. \square

Lemma B.1.3 *Let $0 < \delta < \varepsilon$, $v_n, v: X \rightarrow \mathbb{R}$ be functions such that $v_n \rightarrow v$ quasi uniformly in X and $G_\delta, G_\varepsilon \in \mathcal{O}(X)$ such that*

$$\begin{aligned} \text{cap}(G_\delta) < \delta \quad \text{and} \quad \lim_{n \rightarrow \infty} \sup_{X \setminus G_\delta} |v_n - v| &= 0, \\ \text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad \lim_{n \rightarrow \infty} \sup_{X \setminus G_\varepsilon} |v_n - v| &= 0. \end{aligned}$$

Then $G_\delta \subset G_\varepsilon$ w.l.o.g.

Proof. From the monotonicity of the capacity, we obtain that

$$G_\delta \cap G_\varepsilon \in \mathcal{O}(X), \quad G_\delta \cap G_\varepsilon \subset G_\varepsilon \quad \text{and} \quad \text{cap}(G_\delta \cap G_\varepsilon) \leq \text{cap}(G_\varepsilon) < \varepsilon,$$

and additionally,

$$\sup_{x \in X \setminus (G_\delta \cap G_\varepsilon)} |v_n(x) - v(x)| \leq \sup_{x \in X \setminus G_\delta} |v_n(x) - v(x)| + \sup_{x \in X \setminus G_\varepsilon} |v_n(x) - v(x)| \rightarrow 0. \quad \square$$

Therefore, it is unsurprising that the proof of Lemma 3.1.13 naturally yields a sequence of non-increasing nested exceptional sets $(G_k) \subset \mathcal{O}(X)$ with $\text{cap}(G_k) < \frac{1}{k}$.

B.2 Quasi Open and Quasi Closed Sets

In the literature on Sobolev capacities, different equivalent definitions of quasi openness and quasi closedness of sets can be found. The equivalence of the definitions can easily be proven with the following lemmas.

Lemma B.2.1 *Let $A \subset X$. For every $\varepsilon > 0$, there exists a set U_ε in $\mathcal{O}(X)$ such that $A \subset U_\varepsilon$ and $\text{cap}(U_\varepsilon) \leq \text{cap}(A) + \varepsilon$.*

Proof. By Definition 3.1.3, for any $\varepsilon > 0$, there exists a $v \in V$ such that $\|v\|_V \leq \text{cap}(A) + \varepsilon$ and $v \geq 1$ μ -a.e. on a set $U_\varepsilon \in \mathcal{O}(X)$ with $A \subset U_\varepsilon$. Since v qualifies as one of the functions that the infimum of norms in the definition of $\text{cap}(U_\varepsilon)$ is formed over, we have that $\text{cap}(U_\varepsilon) \leq \|v\|_V \leq \text{cap}(A) + \varepsilon$. \square

Lemma B.2.2 (Quasi Openness of Sets) *A set $A \subset X$ is quasi open in X if one of the following equivalent conditions holds for every $\varepsilon > 0$.*

- (a) *There exists G in $\mathcal{O}(X)$ such that $\text{cap}(G) \leq \varepsilon$ and $A \setminus G$ is rel. open in $X \setminus G$.*
- (b) *There exists G in $\mathcal{O}(X)$ such that $\text{cap}(G) \leq \varepsilon$ and $A \cup G \in \mathcal{O}(X)$.*
- (c) *There exists U in $\mathcal{O}(X)$ such that $A \subset U$ and $\text{cap}(U \setminus A) \leq \varepsilon$.*
- (d) *There exists U in $\mathcal{O}(X)$ such that $\text{cap}(U \setminus A \cup A \setminus U) \leq \varepsilon$.*

Proof. Let $\varepsilon > 0$ be arbitrary but fixed for the remainder of the proof.

(a) \Rightarrow (b): Let $G \in \mathcal{O}(X)$ be given as in (a). Since $A \setminus G$ is relatively open in $X \setminus G$, there exists a set $U \in \mathcal{O}(X)$ with $A \setminus G = U \setminus G$ and consequently $A \cup G = U \cup G \in \mathcal{O}(X)$.

(b) \Rightarrow (c): Let $G \in \mathcal{O}(X)$ be given as in (b) and define $U := A \cup G \in \mathcal{O}(X)$, then we have that $\text{cap}(U \setminus A) \leq \text{cap}(G) \leq \varepsilon$.

(c) \Rightarrow (d): Holds trivially with the same U .

(d) \Rightarrow (a): Choose $U \in \mathcal{O}(X)$ as in (d) for $\frac{\varepsilon}{2} > 0$ and find $G \in \mathcal{O}(X)$ as in Lemma B.2.1 with $U \setminus A \cup A \setminus U \subset G$ and $\text{cap}(G) \leq \varepsilon$. Since $A \setminus (U \setminus A \cup A \setminus U) = A \cap U = U \setminus (U \setminus A \cup A \setminus U)$, we also know that $A \setminus G = U \setminus G$, which means A is relatively open in $X \setminus G$. \square

Parts (a) and (c) of the previous lemma are addressed in [2, Sec. 6.4], part (b) appears in both [51, Chap. 6] and [111], and part (d) is found in [11, Sec. 5.8.3]. Since quasi closed sets are defined as the complements of quasi open sets, the equivalences in B.2.2 have obvious corresponding formulations for quasi closed sets.

A standard observation is that countable unions or intersections of quasi open or closed sets are themselves quasi open or -closed, respectively.

Lemma B.2.3 *Let $A_n \subset X$ be a countable family of sets.*

(a) *If A_n is quasi open for all $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n$ is quasi open.*

(b) *If A_n is quasi closed for all $n \in \mathbb{N}$, then $\bigcap_{n=1}^{\infty} A_n$ is quasi closed.*

Proof. We prove B.2.3 (a) while (b) follows from considering the complements. Let $\varepsilon > 0$ and $G_n \in \mathcal{O}(X)$, $n \geq 1$ be given such that

$$\text{cap}(G_n) \leq \frac{\varepsilon}{2^n} \quad \text{and} \quad A_n \cup G_n \in \mathcal{O}(X).$$

Then $\bigcup_{n=1}^{\infty} G_n \in \mathcal{O}(X)$ as well as

$$\text{cap}\left(\bigcup_{n=1}^{\infty} G_n\right) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon \quad \text{and} \quad \bigcup_{n=1}^{\infty} A_n \cup \bigcup_{n=1}^{\infty} G_n = \bigcup_{n=1}^{\infty} (A_n \cup G_n) \in \mathcal{O}(X). \quad \square$$

Note that quasi open and quasi closed sets are generally unstable under arbitrary unions and intersections, as the following example shows.

Example B.2.4 *Let $d \geq 2$, $1 < p < d$ and consider the (standard Sobolev) capacity cap defined by the triple $(\mathbb{R}^d, W^{1,p}(\mathbb{R}^d), \mathcal{L}^d)$. Let $A \subset \mathbb{R}^d$ be a set that is not quasi open, e.g., the closed unit ball, then A can be written as the uncountable union of quasi open sets.*

Proof. In the setting of the example, sets that consist of single points are polar w.r.t. cap , cf. [88, Thm. 2.27, Thm. 2.38]. Lemma B.2.1 implies that every polar set is quasi open and clearly $A = \bigcup_{x \in A} x$. \square

Accordingly, quasi open subsets of X do not necessarily define a topology on X , despite the terminology, and a counterexample to the arbitrary intersection of quasi closed sets can be found in the same manner as before. When considering $W_0^{1,p}$ -capacities, however, for each quasi open set, we can find sets in the fine topology such that the capacity of

the symmetric difference has arbitrarily small capacity, i.e., the sets are equivalent in that sense [2, Sec. 6.4], [68, Sec. 4], [111]. Additionally, the following property of topologically open and closed sets applies to quasi open and quasi closed subsets as well.

Lemma B.2.5 *Let $U \subset X$ be quasi open in X and $V \subset X$ be quasi closed in X . Then:*

- (a) *A set $A \subset U$ is quasi open in U if and only if it is quasi open in X .*
- (b) *A set $A \subset V$ is quasi closed in V if and only if it is quasi closed in X .*

Proof. Let $\varepsilon > 0$ be fixed for the proof. For part (a), let A be quasi open in X . By definition there exists a set $G_\varepsilon \in \mathcal{O}(X)$ such that

$$\text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad A \cup G_\varepsilon \in \mathcal{O}(X).$$

Accordingly, $G_\varepsilon \cap U \in \mathcal{O}(U)$, and due to monotonicity of the capacity,

$$\text{cap}(G_\varepsilon \cap U) < \varepsilon \quad \text{and} \quad A \cup G_\varepsilon \cap U = (A \cup G_\varepsilon) \cap U \in \mathcal{O}(U).$$

For the inverse implication, let an A be quasi open in U , i.e., by the definition of the subset topology, there exist sets $G_\varepsilon \in \mathcal{O}(U)$ and $F, H_\varepsilon \in \mathcal{O}(X)$ such that

$$\begin{aligned} G_\varepsilon = F \cap U, \quad \text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad A \cup G_\varepsilon = G \cap U \in \mathcal{O}(U), \\ \text{cap}(H_\varepsilon) < \varepsilon \quad \text{and} \quad U \cup H_\varepsilon \in \mathcal{O}(X). \end{aligned}$$

Then $G_\varepsilon \cup H_\varepsilon = (F \cap U) \cup H_\varepsilon = (F \cap (U \cup H_\varepsilon)) \cup H_\varepsilon \in \mathcal{O}(X)$, and due to monotonicity of the capacity, we obtain $\text{cap}(G_\varepsilon \cup H_\varepsilon) < 2\varepsilon$ with

$$A \cup (G_\varepsilon \cup H_\varepsilon) = (G \cap U) \cup H_\varepsilon = (G \cap (U \cup H_\varepsilon)) \cup H_\varepsilon \in \mathcal{O}(X).$$

To confirm part (b), let A be quasi closed in X , i.e., there exists a set $G_\varepsilon \in \mathcal{O}(X)$ such that

$$\text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad X \setminus A \cup G_\varepsilon \in \mathcal{O}(X).$$

Accordingly, $G_\varepsilon \cap V \in \mathcal{O}(V)$, and due to monotonicity of the capacity,

$$\text{cap}(G_\varepsilon \cap V) < \varepsilon \quad \text{and} \quad V \setminus A \cup G_\varepsilon \cap V = (X \setminus A \cup G_\varepsilon) \cap V \in \mathcal{O}(V).$$

Therefore A is quasi closed in V . For the converse, let A be quasi closed in V , i.e., there exist sets $G_\varepsilon \in \mathcal{O}(V)$, $G, H_\varepsilon \in \mathcal{O}(X)$ such that

$$\begin{aligned} \text{cap}(G_\varepsilon) < \varepsilon \quad \text{and} \quad V \setminus A \cup G_\varepsilon = G \cap V \in \mathcal{O}(V), \\ \text{cap}(H_\varepsilon) < \varepsilon \quad \text{and} \quad X \setminus V \cup H_\varepsilon \in \mathcal{O}(X) \end{aligned}$$

Due to Lemma B.2.1 we can find $G_{2\varepsilon} \in \mathcal{O}(X)$ such that

$$G_\varepsilon \subset G_{2\varepsilon} \quad \text{and} \quad \text{cap}(G_{2\varepsilon}) < 2\varepsilon,$$

and therefore $G_{2\varepsilon} \cup H_\varepsilon \in \mathcal{O}(X)$ and $\text{cap}(G_{2\varepsilon} \cup H_\varepsilon) \leq 3\varepsilon$. Additionally,

$$\begin{aligned} (X \setminus A) \cup (G_{2\varepsilon} \cup H_\varepsilon) &= (X \setminus V \cup V \setminus A) \cup (G_{2\varepsilon} \cup H_\varepsilon) \\ &= X \setminus V \cup H_\varepsilon \cup V \setminus A \cup G_{2\varepsilon} \\ &= X \setminus V \cup H_\varepsilon \cup (G \cap V) \cup G_{2\varepsilon} \\ &= X \setminus V \cup H_\varepsilon \cup G \cup G_{2\varepsilon} \in \mathcal{O}(X) \end{aligned}$$

implies that $X \setminus A$ is quasi open in X and therefore the claim. \square

B.3 Quasi Continuous Functions

The following proposition is an extension of the result [2, Prop. 6.4.10] for finite valued functions in the Sobolev case, whose proof carries over.

Lemma B.3.1 (Quasi Openness and Quasi Continuity) *A function $u: X \rightarrow \overline{\mathbb{R}}$ is quasi continuous if and only if the inverse image of each open sets in $\mathcal{O}(\overline{\mathbb{R}})$ under u is quasi open in X .*

Proof. Let $u: X \rightarrow \overline{\mathbb{R}}$ be quasi continuous and $O \in \mathcal{O}(\overline{\mathbb{R}})$ w.r.t. the order topology. For all $\varepsilon > 0$, there exists G such that $\text{cap}(G) \leq \varepsilon$ and $u: X \setminus G \rightarrow \overline{\mathbb{R}}$ is continuous, i.e., for each $O \in \mathcal{O}(\overline{\mathbb{R}})$, we know that $u^{-1}(O) \setminus G$ is relatively open in $X \setminus G$ which means that $u^{-1}(O) \setminus G$ is quasi open, see Lemma B.2.2.

For the converse, let $\varepsilon > 0$ and O in $\mathcal{O}(\overline{\mathbb{R}})$ be given. We will show that there exists a $G \in \mathcal{O}(X)$ with $\text{cap}(G) \leq \varepsilon$ such that $u^{-1}(O) \setminus G$ is relatively open in $X \setminus G$.

First note that the collection \mathcal{I} of all open intervals

$$(a, b), \{x > a\}, \{x < b\} \text{ for } a, b \text{ in } \mathbb{R} \text{ with } a < b$$

form a base of the (order) topology on $\overline{\mathbb{R}}$, hence there exists a subcollection \mathcal{J} of \mathcal{I} with $O = \bigcup_{J \in \mathcal{J}} J$. Since \mathbb{Q} is dense in \mathbb{R} , every J in \mathcal{J} can be represented as the numerable union $J = \bigcup_{l=0}^{\infty} J_l$, where $J_l \in \mathcal{I}$ with rational a, b .

Let $\{I_k : k \in \mathbb{N}\}$ denote an enumeration of all intervals I in \mathcal{I} with a, b in \mathbb{Q} . The inverse image $u^{-1}(I_k)$ is quasi open for all k , so there exist $G_k \in \mathcal{O}(X)$ with $\text{cap}(G_k) \leq \frac{\varepsilon}{2^{k+1}}$ such that $u^{-1}(I_k) \setminus G_k$ is relatively open in $X \setminus G_k$. Define the set $G := \bigcup_{k=0}^{\infty} G_k$, then

$$\text{cap}(G) \leq \sum_{k=0}^{\infty} \frac{\varepsilon}{2^{k+1}} = \varepsilon \text{ and } G \in \mathcal{O}(X).$$

The sets $u^{-1}(I_k) \setminus G$ are relatively open in $X \setminus G$ because $G_k \subset G$ for all k . Therefore, $u^{-1}(J) \setminus G = \bigcup_{l=0}^{\infty} u^{-1}(J_l) \setminus G$ is relatively open in $X \setminus G$ for any $J \in \mathcal{J}$ as the union of relatively open sets. By the same argument, the set $u^{-1}(O) \setminus G = \bigcup_{j \in \mathcal{J}} u^{-1}(J) \setminus G$ is relatively open. \square

The same argument obviously holds for quasi lower-/upper semi-continuous functions and the preimages of sets $\{x \in \mathbb{R} : x > a\}$ and $\{x \in \mathbb{R} : x < b\}$, respectively.

It is also quite clear that the standard calculus for continuous functions from topological spaces into topological vector spaces immediately transfers to quasi continuous functions as well.

Lemma B.3.2 (Calculus for Quasi Continuous Functions) *Let $u, v: A \subset X \rightarrow \overline{\mathbb{R}}$ be quasi continuous, then whenever $u + v$ is defined, it is quasi continuous. When u, v are finite valued, then uv is quasi continuous as well.*

Proof. Let $\varepsilon > 0$ and $G_\varepsilon, H_\varepsilon > 0$ such that $\text{cap}(G_\varepsilon), \text{cap}(H_\varepsilon) < \varepsilon$ and

$$u: A \setminus G_\varepsilon \rightarrow \overline{\mathbb{R}} \quad \text{and} \quad v: A \setminus H_\varepsilon \rightarrow \overline{\mathbb{R}} \quad \text{are continuous.}$$

Then their restrictions to $A \setminus (G_\varepsilon \cup H_\varepsilon)$ are continuous as well and continuity of the sum and the product in the finite valued case on the restricted domain are standard results from topology. Due to the monotonicity of the capacity, $\text{cap}(G_\varepsilon \cup H_\varepsilon) < 2\varepsilon$, which concludes the proof. \square

In the case of Sobolev capacities, this leads to the following:

Lemma B.3.3 *Let $d \geq 1$, $\Omega \subset \mathbb{R}^d$ be a bounded strong Lipschitz domain, $\nu \in C^{0,1}(\Omega)$ and $y \in W^{1,p}(\Omega)$. Then $y\nu \in W^{1,p}(\Omega)$, and there exists $C(\nu) > 0$, independent of y , such that $\|y\nu\|_{W^{1,p}(\Omega)} \leq C \|y\|_{W^{1,p}(\Omega)}$. Additionally,*

$$\tilde{y}\nu = \widetilde{y\nu} \text{ q.e. on } \Omega$$

for all cap_Ω -quasi continuous representatives \tilde{y} and $\widetilde{y\nu}$ of y and $y\nu$.

Proof. The fact that $y\nu \in W^{1,p}(\Omega)$ as well the boundedness is due to [148, Lem. 2.5.5]. Theorem 3.1.17 yields that y has a finite valued quasi continuous representative \tilde{y} and since ν is continuous, Lemma B.3.2 yields that $\tilde{y}\nu$ is a finite valued quasi continuous representative of $y\nu$. Uniqueness up to polar sets of all quasi continuous representatives of $y\nu$ in Theorem 3.1.17 yields the claim. \square

Particularly, we know the following:

Corollary B.3.4 *Let $d \geq 1$, $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $\Gamma_D \subset \partial\Omega$ be relatively open, $\nu \in C^{0,1}(\Omega)$ and $y \in W_D^{1,p}(\Omega)$. Then $y\nu \in W_D^{1,p}(\Omega)$.*

Proof. Due to Corollary 3.4.10, we know that $y = 0$ q.e. on Γ_D and Lemma B.3.3 yields that

$$y\nu = 0 \text{ q.e. on } \Gamma_D,$$

and therefore $y\nu \in W_D^{1,p}(\Omega)$. \square

Note that [28, Lem. A.18] implies the previous corollary without the use of capacity theory.

Another intuitive result in the abstract setting is the existence of pointwise quasi everywhere convergent subsequences of sequences that converge in V , cf. [31, Lem. 6.52]. The following proof is attributable to a communication with Constantin Christof during our work on [44].

Proposition B.3.5 (Q.E. Pointwise Convergent Subsequences) *Let $(v_n) \subset V$ be a sequence in V and $v \in V$ such that $v_n \rightarrow v$ in V . Then there exists a subsequence (v_{n_k}) of (v_n) that converges pointwise quasi everywhere in X to v .*

Proof. Density of $V \cap C(X)$ in V yields sequences $(w_{n,k})$ such that $w_{n,k} \xrightarrow{k \rightarrow \infty} v_n$ in V for all $n \in \mathbb{N}$. Corollary 3.1.14 implies that for every $n \in \mathbb{N}$, we can consider a subsequence (still denoted $w_{n,k}$) such that $w_{n,k} \xrightarrow{k \rightarrow \infty} \tilde{v}_n$ quasi uniformly in X for a quasi continuous representative \tilde{v}_n of v_n . Due to the quasi uniform convergence, for every $n \in \mathbb{N}$, there is a set $G_n \in \mathcal{O}(X)$ such that

$$\text{cap}(G_n) < \frac{1}{2^{n+1}} \quad \text{and} \quad \lim_{k \rightarrow \infty} \left(\sup_{x \in X \setminus G_n} |w_{n,k}(x) - \tilde{v}_n(x)| \right) = 0$$

For every $n \geq 1$, choose a k_n , such that

$$\|w_{n,k_n} - v_n\|_V \leq \frac{1}{n} \quad \text{and} \quad \sup_{x \in X \setminus G_n} |w_{n,k_n}(x) - \tilde{v}_n(x)| \leq \frac{1}{n}.$$

Clearly, $(w_{n,k_n})_n \subset V \cap C(X)$ and $w_{n,k_n} \xrightarrow{n \rightarrow \infty} v$ in V . Again, Corollary 3.1.14 yields a subsequence of (w_{n,k_n}) and the corresponding subsequence of (v_n) (both still denoted by the same symbols) such that $w_{n,k_n} \xrightarrow{n \rightarrow \infty} \tilde{v}$ quasi uniformly in X for a quasi continuous representative \tilde{v} of v , which implies $w_{n,k_n} \xrightarrow{n \rightarrow \infty} v$ pointwise in $X \setminus N$ for a polar set N .

By defining

$$H_l := N \cup \bigcup_{n \geq l} G_n \quad \text{and} \quad H := \bigcap_{l \in \mathbb{N}} H_l$$

and by monotonicity of the capacity, we obtain that

$$H_{l+1} \subset H_l, \quad \text{cap}(H_l) \leq \sum_{n=l}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2^l} \quad \text{for all } l \text{ in } \mathbb{N}, \text{ and } \text{cap}(H) = 0.$$

Accordingly, for any $x \in X \setminus H$, we know that $x \in X \setminus N$ and that there exists $l \in \mathbb{N}$ such that $x \in X \setminus H_l$, and therefore, for all $n \geq l$, we have that

$$\begin{aligned} |\tilde{v}_n(x) - \tilde{v}(x)| &\leq |\tilde{v}_n(x) - w_{n,k_n}(x)| + |w_{n,k_n}(x) - \tilde{v}(x)| \\ &\leq \frac{1}{n} + |w_{n,k_n}(x) - \tilde{v}(x)| \xrightarrow{n \rightarrow \infty} 0. \end{aligned} \quad \square$$

We close this section with a condition that allows for the extension of quasi continuous functions.

Lemma B.3.6 (Extensions of Quasi Continuous Functions) *Let $A \subset X$ be quasi open and quasi closed and $u: A \rightarrow \overline{\mathbb{R}}$ be quasi continuous. Then there exists a function $v: X \rightarrow \overline{\mathbb{R}}$ that is quasi continuous and $v|_A = u$. If u is finite valued, then so is v .*

Proof. Define the extension as

$$v(x) := \begin{cases} u(x), & x \in A \\ 0, & x \in X \setminus A, \end{cases}$$

and let $\varepsilon > 0$ and $I_\varepsilon, G_\varepsilon, H_\varepsilon \in \mathcal{O}(X)$ such that

$$\begin{aligned} \text{cap}(G_\varepsilon) &< \varepsilon \quad \text{and} \quad X \setminus A \cup G_\varepsilon \in \mathcal{O}(X), \\ \text{cap}(H_\varepsilon) &< \varepsilon \quad \text{and} \quad A \cup H_\varepsilon \in \mathcal{O}(X), \\ \text{cap}(I_\varepsilon) &< \varepsilon \quad \text{and} \quad g: A \setminus I_\varepsilon \rightarrow \overline{\mathbb{R}} \text{ is continuous.} \end{aligned}$$

Set $U_\varepsilon := I_\varepsilon \cup G_\varepsilon \cup H_\varepsilon$ and note that for every $O \in \mathcal{O}(\overline{\mathbb{R}})$, if $0 \notin O$, we know from the continuity of u on $A \setminus U_\varepsilon$ and the definition of the subset topology that there exists a set $G \in \mathcal{O}(X)$ with

$$v^{-1}(O) \setminus U_\varepsilon = u^{-1}(O) \setminus U_\varepsilon = G \cap A \setminus U_\varepsilon = G \cap (A \cup H_\varepsilon) \setminus U_\varepsilon \in \mathcal{O}(X \setminus U_\varepsilon).$$

If $0 \in O$, then the quasi closedness of $X \setminus A$ additionally yields that

$$\begin{aligned} v^{-1}(O) \setminus U_\varepsilon &= (u^{-1}(O) \cup X \setminus A) \setminus U_\varepsilon \\ &= (G \cap (A \cup H_\varepsilon) \cup (X \setminus A \cup G_\varepsilon)) \setminus U_\varepsilon \in \mathcal{O}(X \setminus U_\varepsilon). \end{aligned}$$

Hence the restriction of v to $X \setminus U_\varepsilon$ is continuous and monotonicity of the capacity ensures $\text{cap}(U_\varepsilon) < 3\varepsilon$ which concludes the proof. \square

Of course, when the boundary of a set has capacity zero, then the set is both quasi open and quasi closed and the constant extension above preserves quasi continuity.

Corollary B.3.7 *Let $A \subset X$ with $\text{cap}(\partial A) = 0$. Then A is quasi open and quasi closed.*

Proof. Let $\varepsilon > 0$ and $G_\varepsilon \in \mathcal{O}(X)$ such that $\text{cap}(G_\varepsilon) < \varepsilon$ with $\partial A \subset G$, see Lemma B.2.1. Then $A \cup G_\varepsilon = \text{int}(A) \cup G_\varepsilon \in \mathcal{O}(X)$, so A is quasi open. Additionally, $X \setminus A \cup G_\varepsilon = X \setminus \bar{A} \cup G_\varepsilon \in \mathcal{O}(X)$, so A is quasi closed. \square

Appendix C

Miscellaneous

We collect the remaining auxiliary results and theorems that needed referencing in the main content of this thesis.

C.1 Singer's Theorem

The Riesz-Markov-Kakutani representation theorem establishes the well-known connection between bounded linear functionals on the scalar valued continuous functions of a Hausdorff space X and the Radon measures on X . Singer's representation theorem extends the result to Banach space valued continuous functions. The result as stated in [142, Thm. 1.2] reads as follows.

Theorem C.1.1 (Singer Representation Theorem) *Let X be a compact Hausdorff space and V be a Banach space. Then there is a one-to-one correspondence between the linear and bounded functionals $f \in C([0, T], \mathbf{H}^1(\Omega))^*$ and the $\mathbf{H}^1(\Omega)^*$ -valued radon measures on $[0, T]$ with bounded variation, denoted $\mu \in \mathcal{M}([0, T], \mathbf{H}^1(\Omega)^*)$, that is given by*

$$\langle f, \mathbf{p} \rangle_{C([0, T], \mathbf{H}^1(\Omega))} = \int_X \mathbf{p} \, d\mu \quad \forall \mathbf{p} \in C([0, T], \mathbf{H}^1(\Omega)),$$

and the variation of μ coincides with $\|f\|_{C([0, T], \mathbf{H}^1(\Omega))^}$.*

See also [89], [170, Thm. III.3.4] or the original source [176] in French.

Remark C.1.2 *In the literature, the space of regular (countably additive) Borel measures with bounded variation $\mathcal{M}([0, T], \mathbf{H}^1(\Omega)^*)$ is also denoted $\text{racbv}([0, T], \mathbf{H}^1(\Omega)^*)$.*

C.2 Cones in Product Spaces

For the most part, the contact problem and its control are analyzed in the time-discretized setting in Chapters 4 and 5. Naturally, subsets of product spaces are encountered fre-

quently. First of all, recall that we identify the dual of the product of Banach spaces with the product of the dual spaces in the following sense.

Lemma C.2.1 *Let $N \geq 1$, $(Y_k, \|\cdot\|_{Y_k})$, $k = 1, \dots, N$ be Banach spaces, and endow $Y = \prod_{k=1}^N Y_k$ with $\|\cdot\|_\infty$ and $\prod_{k=1}^N Y_k^*$ with $\|\cdot\|_1$, where*

$$\|y\|_\infty := \max_{k \in \{1, \dots, N\}} \|y_k\|_{Y_k} \quad \text{and} \quad \|f\|_1 := \sum_{k=1}^N \|f_k\|_{Y_k^*}.$$

Then the mapping

$$\zeta: \left(\prod_{k=1}^N (Y_k, \|\cdot\|_{Y_k})^*, \|\cdot\|_1 \right) \rightarrow ((Y, \|\cdot\|_\infty)^*, \|\cdot\|_{Y^*})$$

$$\langle \zeta(f_1, \dots, f_N), y \rangle_Y := \sum_{k=1}^N \langle f_k, y_k \rangle_{Y_k}$$

is a linear isometry. Particularly, there is an isomorphism between Y^* and $\prod_{k=1}^N Y_k^*$ when either of the Y_k , Y or $\prod_{k=1}^N Y_k^*$ are endowed with equivalent norms.

Proof. Well definedness of ζ is quickly verified, since

$$|\langle \zeta(f_1, \dots, f_N), y \rangle_Y| = \left| \sum_{k=1}^N \langle f_k, y_k \rangle_{Y_k} \right| \leq \sum_{k=1}^N \|f_k\|_{Y_k^*} \|y_k\|_{Y_k} \leq \|(f_1, \dots, f_N)\|_1 \|y\|_\infty,$$

and linearity of the image as well as linearity of ζ itself clearly holds. Surjectivity of ζ holds since the preimage of $f \in Y^*$ can explicitly be constructed by $\zeta^{-1}(f) := (f_1, \dots, f_N)$ with $\langle f_k, y_k \rangle_{Y_k} := \langle f, (0, \dots, 0, y_k, 0, \dots, 0) \rangle_Y$, so we are left with the task of checking isometry. To that end, let $\varepsilon > 0$ be arbitrary and $\tilde{y} \in Y$ such that $\tilde{y}_k \in Y_k$, $k = 1, \dots, N$ are normalized elements such that

$$\|f_k\|_{Y_k^*} = \sup_{\substack{y_k \in Y_k \\ \|y_k\|_{Y_k} = 1}} \langle f_k, y_k \rangle_{Y_k} \leq \langle f_k, \tilde{y}_k \rangle_{Y_k} + \frac{\varepsilon}{N}.$$

Setting $f = \zeta(f_1, \dots, f_N)$, we have that

$$\begin{aligned} \|f\|_{Y^*} &= \sup_{\substack{y \in Y \\ \|y\|_\infty = 1}} \langle f, y \rangle_Y = \sup_{\substack{y \in Y \\ \|y\|_\infty = 1}} \sum_{k=1}^N \langle f_k, y_k \rangle_{Y_k} \leq \sup_{\substack{y \in Y \\ \|y\|_\infty = 1}} \sum_{k=1}^N \|f_k\|_{Y_k^*} \|y_k\|_{Y_k} \\ &\leq \sum_{k=1}^N \|f_k\|_{Y_k^*} = \|(f_1, \dots, f_N)\|_1 = \sum_{k=1}^N \sup_{\substack{y_k \in Y_k \\ \|y_k\|_{Y_k} = 1}} \langle f_k, y_k \rangle_{Y_k} \\ &\leq \sum_{k=1}^N \langle f_k, \tilde{y}_k \rangle_{Y_k} + \frac{\varepsilon}{N} = \langle f, \tilde{y} \rangle_Y + \varepsilon \leq \|f\|_{Y^*} + \varepsilon. \end{aligned}$$

Letting $\varepsilon \rightarrow 0$ above shows the isometry. □

We obtain the following intuitive correspondences of the properties of the component sets and their product counterparts.

Lemma C.2.2 *Let $N \geq 1$ and K_k be subsets of Banach spaces Y_k for $k = 1, \dots, N$, $K = \prod_{k=1}^N K_k$, $y \in K$ and $\lambda \in \mathcal{T}_K(y)^\circ$. Then:*

- (a) K is closed and convex in Y if and only if the sets K_k are closed and convex in Y_k for all $k = 1, \dots, N$.
- (b) $\mathcal{T}_K(y) = \prod_{k=1}^N \mathcal{T}_{K_k}(y_k)$
- (c) $\mathcal{T}_K(y)^\circ = \prod_{k=1}^N \mathcal{T}_{K_k}(y_k)^\circ$
- (d) $\mathcal{C}_K(y, \lambda) = \prod_{k=1}^N \mathcal{C}_{K_k}(y_k, \lambda_k)$
- (e) $\mathcal{C}_K(y, \lambda)^\circ = \prod_{k=1}^N \mathcal{C}_{K_k}(y_k, \lambda_k)^\circ$
- (f) K is polyhedral if and only if the sets K_k are polyhedral for all $k = 1, \dots, N$.

Proof. Convexity in part (a) is a trivial assertion, and the closedness is due to the definition of the product topology.

Part (b) is a consequence of forming the componentwise closure in the product topology because $\mathcal{R}_K(y) = \prod_{k=1}^N \mathcal{R}_{K_k}(y_k)$.

With the identification of Y^* and $\prod_{k=1}^N Y_k^*$, the inclusion $\mathcal{T}_K(y)^\circ \supset \prod_{k=1}^N \mathcal{T}_{K_k}(y_k)^\circ$ is immediately clear, while the converse follows because $0 \in \mathcal{T}_{K_k}(y_k)$ for all $k = 1, \dots, N$, which shows part (c).

For part (d), let $v \in \mathcal{T}_K(y)$ and $\lambda \in \mathcal{T}_K(y)^\circ$, and note that the tangent and polar cone characterizations of the previous parts imply that $v \in \{\lambda\}^\perp$, i.e.,

$$\langle \lambda, v \rangle_Y = \sum_{k=1}^N \underbrace{\langle \lambda_k, v_k \rangle_{Y_k}}_{\leq 0} = 0,$$

if and only if $v_k \in \{\lambda_k\}^\perp$, hence

$$\mathcal{C}_K(y, \lambda) = \mathcal{T}_K(y) \cap \{\lambda\}^\perp = \prod_{k=1}^N \mathcal{T}_{K_k}(y_k) \cap \{\lambda_k\}^\perp = \prod_{k=1}^N \mathcal{C}_{K_k}(y_k, \lambda_k).$$

The characterization of part (e) can be proven analogously to the proof of part (c) above, because $0 \in \mathcal{C}_{K_k}(y_k, \lambda_k)$ for all k in $\{1, \dots, N\}$.

For the polyhedricity claim in (f), note that the previous parts established that for $y \in K$ and $\lambda \in \mathcal{T}_K(y)^\circ$ we have that

$$\mathcal{T}_K(y) \cap \{\lambda\}^\perp = \prod_{k=1}^N \mathcal{T}_{K_k}(y_k) \cap \{\lambda_k\}^\perp,$$

and by the exact same argument, we obtain that

$$\mathcal{R}_K(y) \cap \{\lambda\}^\perp = \prod_{k=1}^N \mathcal{R}_{K_k}(y_k) \cap \{\lambda_k\}^\perp.$$

Since the closure of the product sets is taken componentwise, we have equivalence of polyhedricity at arbitrary $(y, \lambda) \in K \times \mathcal{T}_K(y)^\circ$. Due to [195, Lem. 4.1], equivalence holds for all $\lambda \in Y^*$ as well. \square

C.3 Hadamard Differentiability

The following properties of Hadamard differentiable functions are used in Section 4.4.1.

Lemma C.3.1 *Let X, Y, Z be Banach spaces, let the mappings $F: Y \rightarrow Z$ and $G: X \rightarrow Y$ be given and let $x, dx \in X$.*

- (a) *If G is Lipschitz continuous and directionally differentiable at $x \in X$, then G is Hadamard differentiable at x .*
- (b) *If F is Hadamard differentiable at $G(x)$ and G is directionally differentiable at x , then $F \circ G: X \rightarrow Z$ is directionally differentiable at x , and*

$$(F \circ G)'(x, dx) = F'(G(x), G'(x, dx)).$$

- (c) *If F is Hadamard differentiable at $G(x)$ and G is Hadamard differentiable at x , then $F \circ G: X \rightarrow Z$ is Hadamard differentiable at x , and*

$$(F \circ G)'(x, dx) = F'(G(x), G'(x, dx)).$$

Proof. Let $x, dx \in X$ and $\delta x: [0, \infty) \rightarrow X$ as in Definition 4.4.1 be arbitrary but fixed for the proof. For part (a), let $L_G \geq 0$ denote the Lipschitz constant of G . Then

$$\begin{aligned} \left\| \frac{G(x + \delta x(t)) - G(x)}{t} - G'(x, dx) \right\|_Y &\leq \left\| \frac{G(x + \delta x(t)) - G(x + tdx)}{t} \right\|_Y \\ &\quad + \left\| \frac{G(x + tdx) - G(x)}{t} - G'(x, dx) \right\|_Y \\ &\leq L_G \left\| \frac{\delta x(t) - tdx}{t} \right\|_X \\ &\quad + \left\| \frac{G(x + tdx) - G(x)}{t} - G'(x, dx) \right\|_Y \xrightarrow{t \rightarrow 0} 0, \end{aligned}$$

which proves (a). The proofs of (b) and (c) are virtually the same, therefore we only consider (c), where

$$\begin{aligned} &\left\| \frac{F(G(x + \delta x(t))) - F(G(x))}{t} - F'(G(x), G'(x, dx)) \right\|_Z \\ &= \left\| \frac{F(G(x) - G(x) + G(x + \delta x(t))) - F(G(x))}{t} - F'(G(x), G'(x, dx)) \right\|_Z \xrightarrow{t \rightarrow 0} 0 \end{aligned}$$

because the Hadamard differentiability of G at x implies

$$\left\| \frac{G(x + \delta x(t)) - G(x) - tG'(x, dx)}{t} \right\|_Y \xrightarrow{t \rightarrow 0} 0,$$

i.e., $\delta G(t) := G(x + \delta x(t)) - G(x)$ parameterizes a curve tangential to $G'(x, dx)$ such that $\delta G(0) = G(x) - G(x) = 0$. \square

C.4 Gâteaux and Fréchet Differentiability

In arbitrary Banach spaces, Fréchet differentiability can be obtained from Gâteaux differentiability with the help of a compactness argument similar to the standard finite-dimensional case, treated, e.g., in [8, Prop. A.4].

Lemma C.4.1 *Let X, Y, Z be Banach spaces, $G: X \rightarrow Y$, $F: Y \rightarrow Z$ and let $x \in X$. If*

- (a) G is affine linear, continuous and compact,
- (b) F is Lipschitz continuous and
- (c) F is Gâteaux differentiable at $G(x)$,

then $F \circ G$ is Fréchet differentiable at x .

Proof. First off, note that continuity and affine linearity of G hold if and only if its linear part is bounded, which implies that G is both Lipschitz continuous and Fréchet differentiable everywhere. As F is Lipschitz continuous and Gâteaux differentiable at $G(x)$, it is Hadamard differentiable at $G(x)$ as well, cf. Lemma C.3.1, which also implies that $F \circ G$ is Gâteaux differentiable at x .

Let $\varepsilon > 0$ be arbitrary but fixed for the proof. Since $\overline{G(B_X(1, 0))}$ is relatively compact, there exist a finite number of points $\delta x_k \in X$, $k = 1, \dots, m$, such that

$$\overline{G(B_X(1, 0))} \subset \bigcup_{k=1}^m B_Y(\varepsilon, G(\delta x_k)). \quad (\text{C.1})$$

Gâteaux differentiability of $F \circ G$ at x implies that we can set $h(\varepsilon) := \min_{k=1, \dots, m} h_k(\varepsilon)$, where $h_k(\varepsilon) > 0$ are chosen such that

$$\|F \circ G(x + h\delta x_k) - F \circ G(x) - h(F \circ G)'(x)\delta x_k\|_Z \leq h\varepsilon.$$

for all $h < h_k(\varepsilon)$. Now let $\delta x \in X$, then there exists one of the δx_k such that

$$\begin{aligned} \|G'\delta x - \|\delta x\|_X G'\delta x_k\|_Y &= \|G(\delta x) - G(\|\delta x\|_X \delta x_k)\|_Y \\ &= \|\delta x\|_X \left\| G\left(\frac{\delta x}{\|\delta x\|_X}\right) - G(\delta x_k) \right\|_Y \leq \|\delta x\|_X \varepsilon \end{aligned}$$

due to affine linearity of G and (C.1). Denoting a Lipschitz constant of F by L_F and using the affine linearity of G once again, we obtain

$$\begin{aligned}
& \|F \circ G(x + \delta x) - F \circ G(x) - (F \circ G)'(x)\delta x\|_Z \\
& \leq \|F \circ G(x + \delta x) - F \circ G(x + \|\delta x\|_X \delta x_k)\|_Z \\
& \quad + \|F \circ G(x + \|\delta x\|_X \delta x_k) - F \circ G(x) - \|\delta x\|_X (F \circ G)'(x)\delta x_k\|_Z \\
& \quad + \|\|\delta x\|_X (F \circ G)'(x)\delta x_k - (F \circ G)'(x)\delta x\|_Z \\
& \leq L_F \|G(\delta x) - G(\|\delta x\|_X \delta x_k)\|_Y \\
& \quad + \|\delta x\|_X \varepsilon \\
& \quad + \|F'(G(x), \cdot)\|_{\mathcal{L}(Y,Z)} \|\|\delta x\|_X G'(x)\delta x_k - G'(x)\delta x\|_Y \\
& \leq (L_F + 1 + \|F'(G(x), \cdot)\|_{\mathcal{L}(Y,Z)}) \varepsilon \|\delta x\|_X
\end{aligned}$$

for all $\|\delta x\|_X < h(\varepsilon)$. Since the constant in front of ε in the last line is independent of the direction δx , we have Fréchet differentiability of $F \circ G$ at x . \square

C.5 Auxiliary Numerical Results

Finally, we collect some additional results pertaining to the numerics section.

C.5.1 Influence of the Forward Solver's Tolerance

The tolerance of the multigrid solver that is used to solve the contact problem should of course be set to a user desired value in practice. As far as the examinations in this thesis are concerned, the tolerance is not expected to influence the qualitative behavior of the results, as suggested by the following results. We will take a closer look at the behavior of the functional values depending on the tolerance.

Figures C.1 and C.2 show the difference of the functional values of the line search Algorithm 6.1 with the Armijo rule for the setting of Example 1 for different values of the forward solver tolerance and the lowest functional value achieved by the same algorithm with 10^{-14} as a reference forward solver tolerance. We can observe that the number of iterations needed until no further descent can be detected are similar and that the decrease in the first iterations coincides, but the trajectories in the plot split up after a maximum of 30 iterations and the minimum functional values reached by the optimizer decreases when the tolerance decreases. This suggests that as one would expect, an increase in accuracy of the forward solver becomes beneficial in the optimization considerations only at the later stages of the algorithms when the norm of the corrections reduces below a given threshold that is dependent on the tolerance.

Figure C.2 indicates that this threshold reduces approximately linearly with the tolerance and that for the chosen tolerance of 10^{-9} , no degenerative effects are to be expected.

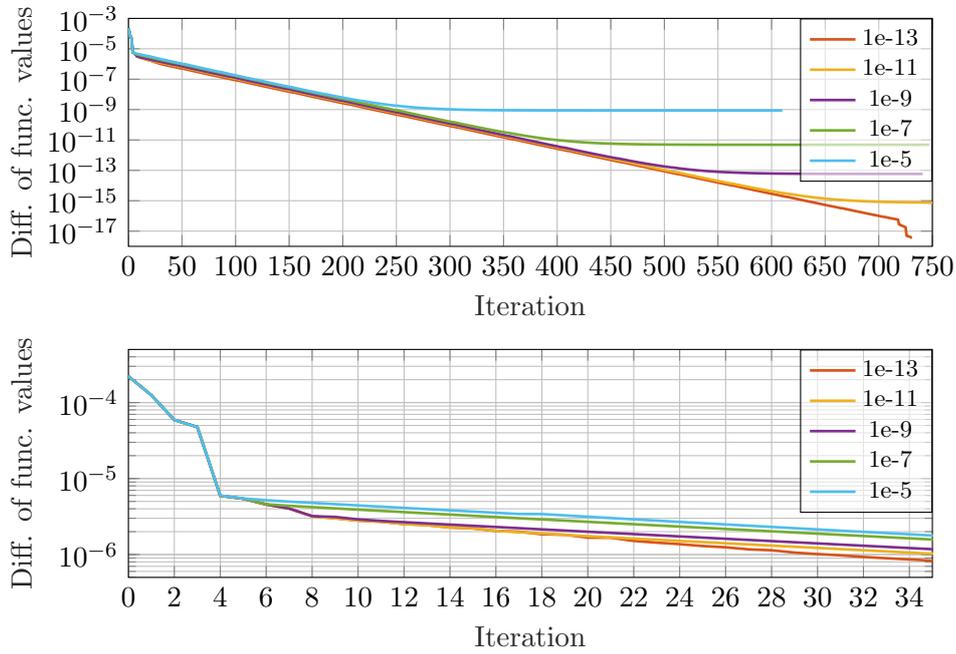


Figure C.1: Difference of the functional values computed by the line search Algorithm 6.1 and the Armijo rule with several forward solver tolerances between 10^{-5} and 10^{-13} and the lowest functional value achieved by the same setting with multigrid solver tolerance set to 10^{-14} as a reference.

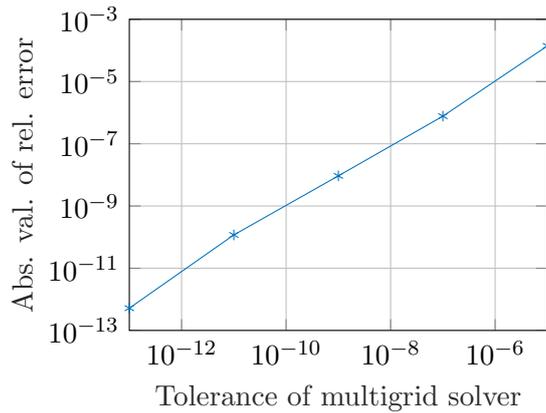


Figure C.2: Difference of the lowest functional values computed by the line search Algorithm 6.1 and the Armijo rule with several forward solver tolerances between 10^{-5} and 10^{-13} and the lowest functional value achieved by the same setting with multigrid solver tolerance set to 10^{-14} as a reference.

C.5.2 Momentum Restarts of the Accelerated Scheme in Example 1

In the second numerical example, we have seen that the structure of the restarted accelerated scheme's cycles correlates with the structure of the functional value decrease and the computed step lengths. The behavior suggests that the quality of the model used to compute the search direction changes over the course of the cycle. In Example 1, a similar structure can be observed for most parameter configurations. The configuration with final time observation $\alpha = 0$ and regularization $\beta = 10^{-4}$, however, only requires very few iterations so that these effects are barely noticeably. The correspondence is clearer when more iterations are required in less regular problem configurations, e.g., for $\alpha = 1$, which we include here in the appendix to keep the presentation in Example 1 concise.

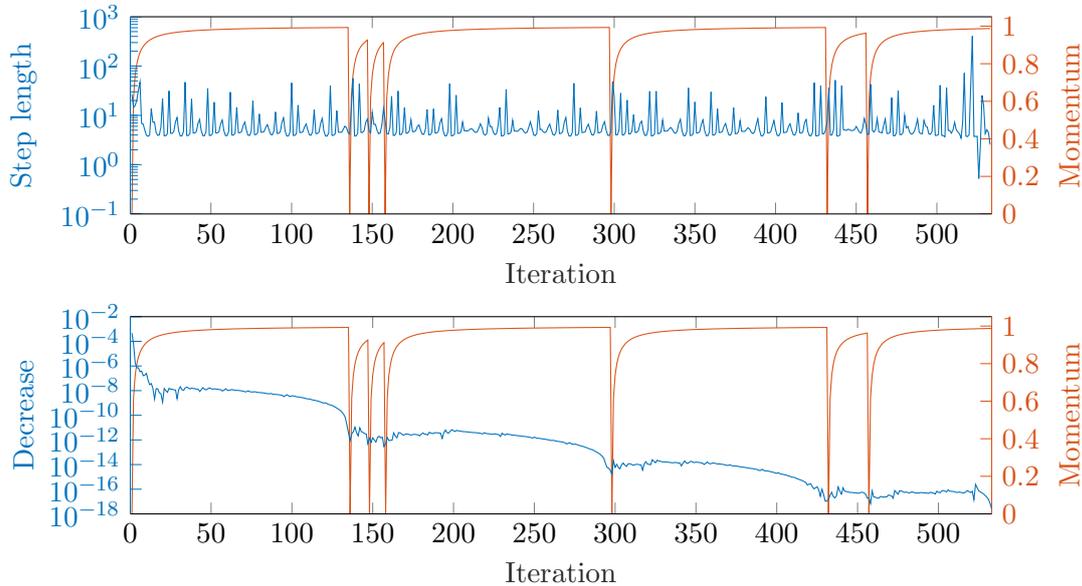


Figure C.3: The momentum of the restarted accelerated scheme vs. the computed step lengths (top) and the objective functional decrease (bottom) for $\alpha = 10$ and $\beta = 10^{-4}$ in Example 1.

We can observe a similar correlation between the momentum and the decrease as in the nonsmooth Example 2 in this case. However, the decrease tends to stay constant during most of the momentum cycle and abruptly drop at the restarts instead of an increase in the decrease as we have seen it in Example 2. This is due to the overall decrease of the decrease in the functional values over the iterations. Since the step lengths remain constant in Example 1, no significant connection between the momentum cycles and the step lengths is noticeable.

C.6 Boundedness of the Boundary-To-Domain Extension Operator

The proof of Lemma 3.3.9 requires a lengthy computation to prove that the boundary-to-domain extension operator by Nečas is indeed bounded. It was omitted there for the sake of presentation and is included here for completeness. The arguments closely follow those of the original work and are attributable to the joint work with Constantin Christof. The proof was omitted from the resulting publication [44] and is in large parts due to him.

Lemma C.6.1 *Let the assumptions of Lemma 3.3.9 hold. Then the operator*

$$E: \left(W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1}), |\cdot|_{W^{1-\frac{1}{p},p}(\mathbb{R}^{d-1})} \right) \mapsto \left(W^{1,p}(\mathbb{R}^{d-1} \times (0, R)), \|\cdot\|_{W^{1,p}(\mathbb{R}^{d-1} \times (0, R))} \right)$$

$$Ev(\mathbf{x}', x_d) := \frac{1}{x_d^{d-1}} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}', \quad (\mathbf{x}', x_d) \in \mathbb{R}^{d-1} \times (0, R)$$

is bounded with a constant depending on d, r, R, p, ρ .

Proof. Denote $w = Ev$ as in the Lemma, and observe that

$$\begin{aligned} \|w\|_{L^p(\mathbb{R}^{d-1} \times (0, R))}^p &= \int_0^R \int_{\mathbb{R}^{d-1}} |w(\mathbf{x}', x_d)|^p d\mathbf{x}' dx_d \\ &= \int_0^R \int_{\mathbb{R}^{d-1}} \left| \frac{1}{x_d^{d-1}} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}' \right|^p d\mathbf{x}' dx_d \\ &= \int_0^R \int_{\mathbb{R}^{d-1}} \left| \int_{\mathbb{R}^{d-1}} v(\mathbf{x}' + x_d \mathbf{z}') \rho(\mathbf{z}') d\mathbf{z}' \right|^p d\mathbf{x}' dx_d \\ &\leq \int_0^R \int_{\|\mathbf{x}'\| \leq r+R} \left(\int_{\|\mathbf{z}'\| \leq 1} |v(\mathbf{x}' + x_d \mathbf{z}') \rho(\mathbf{z}')| d\mathbf{z}' \right)^p d\mathbf{x}' dx_d \\ &\leq \|\rho\|_\infty^p \int_0^R \int_{\|\mathbf{x}'\| \leq r+R} \left(\int_{\|\mathbf{z}'\| \leq 1} |v(\mathbf{x}' + x_d \mathbf{z}')| d\mathbf{z}' \right)^p d\mathbf{x}' dx_d \\ &\leq \|\rho\|_\infty^p C(p) \int_0^R \int_{\|\mathbf{x}'\| \leq r+R} \int_{\|\mathbf{z}'\| \leq 1} |v(\mathbf{x}' + x_d \mathbf{z}')|^p d\mathbf{z}' d\mathbf{x}' dx_d \\ &\leq \|\rho\|_\infty^p C(p) \int_0^R \int_{\|\mathbf{x}'\| \leq r+R} \int_{\mathbb{R}^{d-1}} |v(\mathbf{y}')|^p d\mathbf{y}' d\mathbf{x}' dx_d \\ &= \|\rho\|_\infty^p C(p) \mathcal{L}^{d-1}(B(r+R)) R \|v\|_{L^p(\mathbb{R}^{d-1})}^p \\ &= C(d, p, r, R, \rho) \|v\|_{L^p(\mathbb{R}^{d-1})}^p. \end{aligned}$$

For all derivatives $\frac{\partial}{\partial x_i} w$, $i = 1, \dots, d-1$, we have that

$$\frac{\partial}{\partial x_i} w(\mathbf{x}', x_d) = -\frac{1}{x_d^d} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \frac{\partial \rho}{\partial z_i} \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}'.$$

By adding the artificial zero

$$\begin{aligned} 0 &= v(\mathbf{x}') \frac{\partial}{\partial x_i} (-1) = v(\mathbf{x}') \frac{\partial}{\partial x_i} \left(\frac{-1}{x_d^{d-1}} \int_{\mathbb{R}^{d-1}} \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}' \right) \\ &= \frac{1}{x_d^d} \int_{\mathbb{R}^{d-1}} v(\mathbf{x}') \frac{\partial \rho}{\partial z_i} \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}', \end{aligned}$$

we obtain that

$$\begin{aligned} \frac{\partial}{\partial x_i} w &= \int_{\{\|\mathbf{x}' - \mathbf{y}'\| < x_d\}} \frac{\partial \rho}{\partial z_i} \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) \frac{v(\mathbf{x}') - v(\mathbf{y}')}{x_d^d} d\mathbf{y}' \\ &= \int_{\{\|\mathbf{z}'\| < 1\}} \frac{\partial \rho}{\partial z_i}(\mathbf{z}') \frac{v(\mathbf{x}') - v(\mathbf{x}' + x_d \mathbf{z}')}{x_d} d\mathbf{z}'. \end{aligned}$$

Thus, using that

$$\begin{aligned} &\{(\mathbf{y}', x_d) : \|\mathbf{y}' - \mathbf{x}'\| < x_d, 0 < x_d < R\} = \\ &\{(\mathbf{y}', x_d) : \|\mathbf{y}' - \mathbf{x}'\| < R, \|\mathbf{y}' - \mathbf{x}'\| < x_d < R\} \end{aligned}$$

for all \mathbf{x}' in \mathbb{R}^{d-1} , we gather that

$$\begin{aligned} \left\| \frac{\partial w}{\partial x_i} \right\|_{L^p(\mathbb{R}^{d-1} \times (0, R))}^p &= \int_0^R \int_{\mathbb{R}^{d-1}} \left| \left(\frac{\partial w}{\partial x_i} \right) \right|^p d\mathbf{x}' dx_d \\ &\leq C(p, \rho) \int_0^R \int_{\mathbb{R}^{d-1}} \int_{\{\|\mathbf{z}'\| < 1\}} \left| \frac{v(\mathbf{x}') - v(\mathbf{x}' + x_d \mathbf{z}')}{x_d} \right|^p d\mathbf{z}' d\mathbf{x}' dx_d \\ &= C(p, \rho) \int_{\mathbb{R}^{d-1}} \int_0^R \int_{\{\|\mathbf{y}' - \mathbf{x}'\| < x_d\}} \frac{|v(\mathbf{x}') - v(\mathbf{y}')|^p}{x_d^{d+p-1}} d\mathbf{y}' dx_d d\mathbf{x}' \\ &= C(p, \rho) \int_{\mathbb{R}^{d-1}} \int_{\{\|\mathbf{y}' - \mathbf{x}'\| < R\}} \int_{\|\mathbf{x}' - \mathbf{y}'\|}^R \frac{|v(\mathbf{x}') - v(\mathbf{y}')|^p}{x_d^{d+p-1}} dx_d d\mathbf{y}' d\mathbf{x}' \\ &\leq C(p, \rho) \int_{\mathbb{R}^{d-1}} \int_{\{\|\mathbf{y}' - \mathbf{x}'\| < R\}} \frac{|v(\mathbf{x}') - v(\mathbf{y}')|^p}{(d+p-2) \|\mathbf{x}' - \mathbf{y}'\|^{d+p-2}} d\mathbf{y}' d\mathbf{x}' \\ &\leq C(d, p, \rho) |v|_{W^{1-\frac{1}{p}, p}(\mathbb{R}^{d-1})}^p. \end{aligned}$$

For $i = d$, we can follow the same argument for

$$\begin{aligned} \frac{\partial}{\partial x_d} w(\mathbf{x}', x_d) &= -\frac{d-1}{x_d^d} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}' \\ &\quad - \sum_{i=1}^{d-1} \frac{1}{x_d^d} \int_{\mathbb{R}^{d-1}} v(\mathbf{y}') \frac{y_i - x_i}{x_d} \frac{\partial \rho}{\partial z_i} \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}' \end{aligned}$$

by adding the artificial zero

$$\begin{aligned}
0 = v(\mathbf{x}') \frac{\partial}{\partial x_d}(-1) &= -v(\mathbf{x}') \frac{\partial}{\partial x_d} \left(\frac{1}{x_d^{d-1}} \int_{\mathbb{R}^{d-1}} \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}' \right) \\
&= \frac{d-1}{x_d^d} \int_{\mathbb{R}^{d-1}} v(\mathbf{x}') \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}' \\
&\quad + \sum_{i=1}^{d-1} \frac{1}{x_d^d} \int_{\mathbb{R}^{d-1}} \frac{y_i - x_i}{x_d} v(\mathbf{x}') \frac{\partial \rho}{\partial z_i} \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) d\mathbf{y}',
\end{aligned}$$

to obtain that

$$\begin{aligned}
\left\| \frac{\partial w}{\partial x_d} \right\|_{L_p(\mathbb{R}^{d-1} \times (0, R))}^p &= \int_0^R \int_{\mathbb{R}^{d-1}} \left| \frac{\partial w}{\partial x_d} \right|^p d\mathbf{x}' dx_d \\
&\leq \int_0^R \int_{\mathbb{R}^{d-1}} \left(\int_{\|\mathbf{y}' - \mathbf{x}'\| < x_d} (d-1) \frac{|v(\mathbf{y}') - v(\mathbf{x}')|}{x_d^d} \left| \rho \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) \right| \right. \\
&\quad \left. + \sum_{i=1}^{d-1} \frac{|v(\mathbf{y}') - v(\mathbf{x}')|}{x_d^d} \underbrace{\left| \frac{y_i - x_i}{x_d} \right|}_{\leq \left\| \frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right\| \leq 1} \left| \frac{\partial \rho}{\partial z_i} \left(\frac{\mathbf{y}' - \mathbf{x}'}{x_d} \right) \right| d\mathbf{y}' \right)^p d\mathbf{x}' dx_d \\
&\leq C(d, p, \rho) \int_0^R \int_{\mathbb{R}^{d-1}} \int_{\|\mathbf{y}' - \mathbf{x}'\| < x_d} \frac{|v(\mathbf{y}') - v(\mathbf{x}')|^p}{x_d^{d+p-1}} d\mathbf{y}' d\mathbf{x}' dx_d \\
&\leq C(d, p, \rho) \int_{\mathbb{R}^{d-1}} \int_{B(R, x_d)} \frac{|v(\mathbf{y}') - v(\mathbf{x}')|^p}{(d+p-2) \|\mathbf{y}' - \mathbf{x}'\|^{d+p-2}} dx_d d\mathbf{y}' d\mathbf{x}' \\
&\leq C(d, p, \rho) |v|_{W^{1-\frac{1}{p}, p}(\mathbb{R}^{d-1})}^p.
\end{aligned}$$

Consequently, combining these components in the sum for the norm, cf. Section 1.2, we obtain that operator is well defined and bounded with a constant as claimed. \square

List of Symbols

Mechanics, Elasticity and Contact

a_I	Bilinear form associated with dynamic elastic forces	24
\mathcal{A}_Ψ	Active contact set	87
a	Bilinear form associated with elastic forces	24
b_I	Bilinear form associated with dynamic viscous forces	24
b	Bilinear form associated with viscous forces	24
\mathbf{C}	Elasticity tensor	17
d	Spatial dimension	5
\mathbf{E}	(Finite) Strain tensor	14
E	Young's modulus	17
e_i	Basis vector of d -dimensional euclidean space	5
f_{con}	Contact forces	25
f_{ini}	Initial data in the time stepping scheme	74
\mathbf{f}_{Γ_N}	Boundary force density	13
\mathbf{f}_Ω	Volume force density	13
I	Time interval $[0, T]$	12
I_k	Section $(t_{k-1}, t_k]$ of the time interval in the discretization	67
\mathbf{K}_Φ^T	Set of time dependent admissible displacements	65
\mathbf{K}_Φ	Set of admissible displacements	24
\mathcal{O}	Obstacle	12
\mathcal{S}_Ψ	Strongly active contact set	91
T	Final time	12
t	Time parameter	5
\mathbf{V}	Viscosity tensor	17
\mathbf{v}	Velocity	67
\mathbf{v}_{ini}	Initial velocity	19
\mathcal{W}_Ψ	Weakly active contact set	91
\mathbf{x}	Points in \mathbb{R}^d	5
\mathbf{y}	Displacement field	13
\mathbf{y}_d	Desired displacement	122
\mathbf{y}_{ini}	Initial displacement	19
Γ_C	Contact boundary	12
Γ_D	Dirichlet boundary	12
Γ_N	Neumann boundary	12

ϵ	(Infinitesimal) Strain tensor	14
μ_{bulk}	Bulk viscosity	17
μ_{shear}	Shear viscosity	17
ν	Geometric outer normal	16
ν_{poi}	Poisson's ratio	17
ν_{Φ}	Contact normal	19
ρ	Mass density in reference configuration	12
σ	Stress tensor	15
σ_E	Elastic stress tensor	16
σ_V	Viscous stress tensor	16
τ	Time step size	5
φ	Deformation field	12
Φ	Contact mapping	19
χ_A	Indicator function of a set A	6
Ψ	Initial gap function	20
Ω	Reference configuration	12
$\overline{\Omega^{\varphi,t}}$	Deformed configuration	12

Measures and Function Spaces

C	Set of continuous functions	7
C^k	Set of k times continuously differentiable functions	8
C_0^k	Compactly supported functions in C^k	8
H^s	(Bochner-)Sobolev space $W^{s,2}$ with exponents $s, 2$	9
\mathcal{H}^s	Hausdorff measure for $0 \leq s < \infty$	7
$\mathcal{L}(X, Y)$	Space of linear maps from X to Y	6
lin	Span of vectors	69
\mathcal{L}^d	Lebesgue measure in d dimensions	7
L^p	Space of p -(Bochner-)Lebesgue-integrable functions	8
\mathcal{M}	Space of vector valued Radon measures	69
$\mathcal{P}_k(I, X)$	Space of order $k = 0, 1$ polynomials from I to X	68
$W^{k,p}$	(Bochner-)Sobolev space with exponents k, p	8
$W_0^{k,p}$	Closure of C_0^∞ in $W^{k,p}$	8
$\mathbf{W}(0, T)$	Space of functions weakly differentiable in time and space	9
$W_D^{1,p}$	Space of $W^{1,p}$ -functions with D -boundary-conditions	9

Sets

B	Ball in normed space	6
$\mathcal{B}(\cdot)$	Borel σ -algebra	6
$\mathcal{C}(\cdot)$	Critical cone	6
c_g	Approximation of constant gravitational acceleration	122
\mathbb{N}	Natural numbers	4
$\mathcal{O}(\cdot)$	Topology on set X	5
$\mathcal{P}(\cdot)$	Power set	6
\mathbb{Q}	Rational numbers	4
$\mathcal{R}(\cdot)$	Radial cone	6
\mathbb{R}	Real numbers	4

$\overline{\mathbb{R}}$	Extended real numbers	4
\mathcal{S}_1	Unit sphere in \mathbb{R}^d	9
\mathbb{S}^d	Set of second order tensors in \mathbb{R}^d	9
$\text{SO}(d)$	Special orthogonal group of rotations in \mathbb{R}^d	9
$\mathcal{T}(\cdot)$	Tangent cone	6
Miscellaneous		
α	Final time observation parameter in the numerics	122
β	$L^2(\Omega)$ -regularization parameter in the numerics	122
$\partial_i(\cdot)$	Derivative with respect to component i	5
$\nabla(\cdot)$	First spatial derivative	5
$\dot{(\cdot)}$	First time derivative	5
$\ddot{(\cdot)}$	Second time derivative	5
$\langle \cdot, \cdot \rangle_X$	Dual pairing in space X	6
$(\cdot, \cdot)_X$	Scalar product in (pre-) Hilbert space X	6
$\ \cdot\ _X$	Norm on space X	6
$\ \cdot\ _p$	p -Norm on \mathbb{R}^d	6
$\ \cdot\ $	2-Norm on \mathbb{R}^d	6
dist	Distance function for subsets of normed space	6
P_K^d	Projection onto K with respect to norm induced by d	80
$\text{Tr}(\cdot)$	Trace of a $d \times d$ -matrix	9
tr	Trace operator for Sobolev functions	8
supp	Support of a function	8
$(\cdot)^\perp$	Annihilator of a set	6
$\partial(\cdot)$	Boundary of a set	5
int (\cdot)	Interior of a set	5
$\bar{\cdot}$	Closure of a set	5
$(\cdot)^\circ$	Polar cone of a set	6
$(\cdot)^*$	Adjoint of an operator	6
$(\cdot)^*$	Dual of a normed space	6
$(\cdot)^{**}$	Bidual of a normed space	6
\cdot	Scalar product in \mathbb{R}^d / Tensor contraction	5
$:$	Tensor contraction	5

List of Figures

2.1	Deformation and displacement in solid mechanics	13
2.2	Stress in linear viscoelasticity	16
2.3	One-dimensional Kelvin-Voigt model	16
2.4	Typical plane strain configuration	18
2.5	One-body reference configuration	20
2.6	Modeling error in the linearized contact model — curved obstacle	21
2.7	Modeling error in the linearized contact model — planar obstacle	22
2.8	Correct and incorrect choice of the contact normal in the planar case	23
3.1	Sets with degenerative capacitory behavior at the boundary	40
3.2	Construction of the extension-set in the proof of Lemma 3.3.2	43
3.3	Maximum support of the extension in Lemma 3.3.9	50
3.4	Construction of the extension-sets in the proof of Lemma 3.3.10	53
4.1	Discretization of the primal variables.	68
4.2	Illustration of active contact patches	91
6.1	Reference configuration in Ex. 1	126
6.2	Desired state in Ex. 1	126
6.3	Displacement and contact force density of desired state in Ex. 1	127
6.4	Functional values and norm of search direction in Ex. 1	128
6.5	Step lengths and regularization parameters in Ex. 1	129
6.6	Momentum in Ex. 1	129
6.7	Position and contact forces of desired state and solution in Ex. 1	130
6.8	State and control differences in Ex. 1 — varying α	131
6.9	State and control differences in Ex. 1 — varying β	132
6.10	Local behavior of controls in Ex. 1	133
6.11	State and control differences in Ex. 1	134
6.12	Reference configuration and desired state in Ex. 2	135
6.13	Functional values and norm of search direction in Ex. 2	136
6.14	Step lengths and regularization parameters in Ex. 2	138
6.15	Damping of Armijo, QuadReg and Brent in Ex. 2	139
6.16	Step lengths, decrease and momentum for $\alpha = 1$ in Ex. 2	140
6.17	Step lengths, decrease and momentum for $\alpha = 10$ in Ex. 2	141
6.18	State and control differences in Ex. 2	143

6.19	Position of solution and contact force density in Ex. 2	144
6.20	Reference configuration in Ex. 3	145
6.21	Functional values and norm of search directions in Ex. 3	146
6.22	Momentum in Ex. 3	147
6.23	Step lengths and regularization parameters in Ex. 3	147
6.24	Position and contact forces of south pole in Ex. 3	148
6.25	State and control differences in Ex. 3	149
6.26	Local behavior of controls in Ex. 3	150
A.1	Locally Lipschitz but unbounded function	164
C.1	Influence of the forward solver tolerance on the obtained values	181
C.2	Influence of the forward solver tolerance on the obtained values	181
C.3	Step lengths, decrease and momentum in Ex. 1	182

List of Tables

6.1	Parameters for the numerical examples	125
6.2	Parameters in Ex. 1	127
6.3	Norm comparison of the solutions in Example 1	130
6.4	Parameters in Ex. 2	135
6.5	Norm comparison of the solutions in Ex. 2	142
6.6	Lowest objective functional values in Ex. 2	144
6.7	Parameters in Ex. 3	145

List of Algorithms

6.1	Line Search	113
6.2	Accelerated Line Search without Restart (AccNoRe)	113
6.3	Accelerated Line Search with Restart at Ascent (AccRe)	114
6.4	Computation of Search Directions $\delta \mathbf{u}_k$	118
6.5	Armijo Backtracking	120
6.6	Quadratic Regularization	121

Bibliography

- [1] D. Adams. Choquet integrals in potential theory. *Publ. Mat.*, 42(1):3–66, 1998.
- [2] D. Adams and L. Hedberg. *Function Spaces and Potential Theory*. Springer Verlag, Berlin, 1999.
- [3] R. Adams and J. Fournier. *Sobolev Spaces*. Pure and Applied Mathematics. Elsevier Science, 2003.
- [4] J. Ahn and D. E. Stewart. Dynamic frictionless contact in linear viscoelasticity. *IMA J. Numer. Anal.*, 29(1):43–71, 2009.
- [5] H. Aikawa. Potential analysis on nonsmooth domains—Martin boundary and boundary Harnack principle. In *Complex analysis and potential theory*, volume 55 of *CRM Proc. Lecture Notes*, pages 235–253. Amer. Math. Soc., Providence, RI, 2012.
- [6] H. Alt. *Lineare Funktionalanalysis*. Springer-Lehrbuch Masterclass. Springer Berlin Heidelberg, sixth edition, 2012.
- [7] L. Ambrosio, G. Da Prato, and A. Mennucci. *Introduction to Measure Theory and Integration*. Edizioni Della Normale, Pisa, 2011.
- [8] B. Andrews and C. Hopper. *The Ricci flow in Riemannian geometry*, volume 2011 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011.
- [9] K. T. Andrews, M. Shillor, S. Wright, and A. Klarbring. A dynamic thermoviscoelastic contact problem with friction and wear. *Internat. J. Engrg. Sci.*, 35(14):1291–1309, 1997.
- [10] M. A. Armstrong. *Basic topology*. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1983.
- [11] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2014.
- [12] H. Attouch and M. Soueycatt. Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. Applications to games, PDE’s and control. *Pac. J. Optim.*, 5(1):17–37, 2009.

- [13] A. Auslender and M. Teboulle. Interior gradient and proximal methods for convex and conic optimization. *SIAM J. Optim.*, 16(3):697–725, 2006.
- [14] V. I. Averbukh and O. G. Smolyanov. The theory of differentiation in linear topological spaces. *Russian Math. Surveys*, 22(6):201, 1967.
- [15] I. Babuška and M. Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62(4):439–463, 1992.
- [16] I. Babuška and M. Suri. On locking and robustness in the finite element method. *SIAM J. Numer. Anal.*, 29(5):1261–1293, 1992.
- [17] H. T. Banks, S. Hu, and Z. R. Kenz. A brief review of elasticity and viscoelasticity for solids. *Adv. Appl. Math. Mech.*, 3(1):1–51, 2011.
- [18] V. Barbu. Necessary conditions for distributed control problems governed by parabolic variational inequalities. *SIAM J. Control Optim.*, 19(1):64–86, 1981.
- [19] V. Barbu. Necessary conditions for nonconvex distributed control problems governed by elliptic variational inequalities. *J. Math. Anal. Appl.*, 80(2):566–597, 1981.
- [20] V. Barbu and T. Precupanu. *Convexity and optimization in Banach spaces*. Springer Monographs in Mathematics. Springer, Dordrecht, fourth edition, 2012.
- [21] P. Bastian, M. Blatt, A. Dedner, C. Engwer, J. Fahlke, C. Gräser, R. Klöforn, M. Nolte, M. Ohlberger, and O. Sander. DUNE Web page, 2011. <http://www.dune-project.org>.
- [22] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöforn, M. Ohlberger, and O. Sander. A Generic Grid Interface for Parallel and Adaptive Scientific Computing. Part I: Abstract Framework. *Computing*, 82(2–3):103–119, 2008.
- [23] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöforn, M. Ohlberger, and O. Sander. A generic interface for adaptive and parallel scientific computing. Part II: Implementation and tests in dune. *Computing*, 82(2–3):121–138, 2008.
- [24] K. Bathe. *Finite Element Procedures*. Prentice Hall, second edition, 2014.
- [25] A. Beck. *First-order methods in optimization*, volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2017.
- [26] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.
- [27] F. Ben Belgacem and Y. Renard. Hybrid finite element methods for the Signorini problem. *Math. Comp.*, 72(243):1117–1145, 2003.
- [28] T. Betz. *Optimal control of two variational inequalities arising in solid mechanics*. PhD thesis, Technische Universität Dortmund, 2015.

- [29] H. Blum, A. Rademacher, and A. Schröder. Space adaptive finite element methods for dynamic Signorini problems. *Comput. Mech.*, 44(4):481–491, 2009.
- [30] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.
- [31] J. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2000.
- [32] R. M. Bowen and C.-C. Wang. *Introduction to vectors and tensors*. Dover Publications, Inc., Mineola, NY, second edition, 2008. Vol. 1. Linear and multilinear algebra, Vol. 2. Vector and tensor analysis.
- [33] R. P. Brent. *Algorithms for minimization without derivatives*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1973.
- [34] M. Brokate and A. H. Siddiqi. Sensitivity in the rigid punch problem. *Adv. Math. Sci. Appl.*, 2(2):445–456, 1993.
- [35] A. Browder. *Mathematical analysis*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1996. An introduction.
- [36] A. Brudnyi and Y. Brudnyi. *Methods of geometric analysis in extension and trace problems. Volume 2*, volume 103 of *Monographs in Mathematics*. Birkhäuser/Springer Basel AG, Basel, 2012.
- [37] M. Campo, J. R. Fernández, K. L. Kuttler, M. Shillor, and J. M. Viaño. Numerical analysis and simulations of a dynamic frictionless contact problem with damage. *Comput. Methods Appl. Mech. Engrg.*, 196(1-3):476–488, 2006.
- [38] C. Cartis, N. I. M. Gould, and P. L. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part I: motivation, convergence and numerical results. *Math. Program.*, 127(2, Ser. A):245–295, 2011.
- [39] O. Chau, J. R. Fernández-García, W. Han, and M. Sofonea. A frictionless contact problem for elastic-viscoplastic materials with normal compliance and damage. *Comput. Methods Appl. Mech. Engrg.*, 191(44):5007–5026, 2002.
- [40] O. Chau, M. Shillor, and M. Sofonea. Dynamic frictionless contact with adhesion. *Z. Angew. Math. Phys.*, 55(1):32–47, 2004.
- [41] F. Chouly, P. Hild, and Y. Renard. A Nitsche finite element method for dynamic contact: 1. Space semi-discretization and time-marching schemes. *ESAIM Math. Model. Numer. Anal.*, 49(2):481–502, 2015.
- [42] C. Christof. Sensitivity analysis and optimal control of obstacle-type evolution variational inequalities. Preprint SPP1962-055, 4 2018.
- [43] C. Christof, J. C. de los Reyes, and C. Meyer. A non-smooth trust-region method for locally lipschitz functions with application to optimization problems constrained by variational inequalities. Preprint SPP1962-051, 02 2018.

- [44] C. Christof and G. Müller. A note on the equivalence and the boundary behavior of a class of sobolev capacities. *GAMM-Mitt.*, 40(3):238–266, Apr 2018.
- [45] P. G. Ciarlet. *Mathematical elasticity. Vol. I*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988. Three-dimensional elasticity.
- [46] P. G. Ciarlet. On Korn’s inequality. *Chin. Ann. Math. Ser. B*, 31(5):607–618, 2010.
- [47] M. Cocu and J.-M. Ricaud. Analysis of a class of implicit evolution inequalities associated to viscoelastic dynamic contact problems with friction. *Internat. J. Engrg. Sci.*, 38(14):1535–1552, 2000.
- [48] M. I. M. Copetti and M. Aouadi. A quasi-static contact problem in thermoviscoelastic diffusion theory. *Appl. Numer. Math.*, 109:157–183, 2016.
- [49] S. Crandall. *An Introduction to Mechanics of Solids*. Mcgraw Hill Higher Education, 2012.
- [50] J. W. Daniel. The conjugate gradient method for linear and nonlinear operator equations. *SIAM J. Numer. Anal.*, 4:10–26, 1967.
- [51] M. Delfour and J. Zolésio. *Shapes and Geometries*. Society for Industrial and Applied Mathematics, second edition, 2011.
- [52] P. Deuffhard, R. Krause, and S. Ertel. A contact-stabilized newmark method for dynamical contact problems. *Internat. J. Numer. Methods Engrg.*, 73(9):1274 – 1290, 2007.
- [53] J. Diestel and J. J. Uhl, Jr. The Radon-Nikodym theorem for Banach space valued measures. *Rocky Mountain J. Math.*, 6(1):1–46, 1976.
- [54] J. Diestel and J. J. Uhl, Jr. *Vector measures*. American Mathematical Society, Providence, R.I., 1977.
- [55] N. Dinculeanu. *Vector measures*. International Series of Monographs in Pure and Applied Mathematics, Vol. 95. Pergamon Press, Oxford-New York-Toronto, Ont.; VEB Deutscher Verlag der Wissenschaften, Berlin, 1967.
- [56] S. Doll, R. Hauptmann, K. Schweizerhof, and C. Freischläger. On volumetric locking of low-order solid and solid-shell elements for finite elastoviscoplastic deformations and selective reduced integration. 2007.
- [57] D. Doyen, A. Ern, and S. Piperno. Time-Integration Schemes for the Finite Element Dynamic Signorini Problem. *SIAM J. Sci. Comput.*, 33(1):223–249, Jan. 2011.
- [58] N. Dunford and J. T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988.
- [59] J. Durdil. On Hadamard differentiability. *Comment. Math. Univ. Carolinae*, 14:457–470, 1973.

- [60] E. B. Dynkin. *Superdiffusions and positive solutions of nonlinear partial differential equations*, volume 34 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2004.
- [61] C. Eck. *Existenz und Regularität der Lösungen für Kontaktprobleme mit Reibung*. Mathematisches Institut A der Universität Stuttgart, Stuttgart, 1996. PhD-Thesis.
- [62] C. Eck, J. Jarušek, and M. Krbec. *Unilateral contact problems*, volume 270 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2005. Variational methods and existence theorems.
- [63] M. Egert, R. Haller-Dintelmann, and J. Rehberg. Hardy's inequality for functions vanishing on a part of the boundary. *Potential Anal.*, 43(1):49–78, 2015.
- [64] L. C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998.
- [65] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.
- [66] H. Federer. *Geometric measure theory*. Grundlehren der mathematischen Wissenschaften. Springer, 1969.
- [67] R. Fletcher and S. Leyffer. Solving mathematical programs with complementarity constraints as nonlinear programs. *Optim. Methods Softw.*, 19(1):15–40, 2004.
- [68] B. Fuglede. The quasi topology associated with a countably subadditive set function. *Ann. Inst. Fourier (Grenoble)*, 21(fasc. 1):123–169, 1971.
- [69] M. Fukushima. *Dirichlet forms and Markov Processes*. North-Holland, Amsterdam, 1980.
- [70] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.
- [71] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Mathematische Lehrbücher und Monographien. Akademie-Verlag Berlin, 1974.
- [72] R. P. Gilbert, P. Shi, and M. Shillor. A quasi-static contact problem in linear thermoelasticity. *Rend. Mat. Appl. (7)*, 10(4):785–808 (1991), 1990.
- [73] S. Götschel, M. Weiser, and A. Schiela. Solving optimal control problems with the kaskade 7 finite element toolbox. In A. Dedner, B. Flemisch, and R. Klöforn, editors, *Advances in DUNE*, pages 101 – 112. 2012.
- [74] P. Gould. *Introduction to Linear Elasticity*. Introduction to Linear Elasticity. Springer, 1993.
- [75] C. Graeser and R. Kornhuber. Multigrid methods for obstacle problems. *J. Comput. Math.*, 27(1):1–44, 2009.

- [76] A. Griewank. The modification of newton's method for unconstrained optimization by bounding cubic terms. Technical Report NA/12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 1981.
- [77] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.
- [78] M. Gu, L.-H. Lim, and C. J. Wu. ParNes: a rapidly convergent algorithm for accurate recovery of sparse and approximately sparse signals. *Numer. Algorithms*, 64(2):321–347, 2013.
- [79] M. E. Gurtin. *An introduction to continuum mechanics*, volume 158 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.
- [80] Y. Haddad. *Viscoelasticity of Engineering Materials*. Chapman & Hall, 1995.
- [81] C. Hager, S. Hüeber, and B. I. Wohlmuth. A stable energy-conserving approach for frictional contact problems based on quadrature formulas. *Int. J. Numer. Meth. Eng.*, 73(2):205–225, 2008.
- [82] W. W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods. *Pac. J. Optim.*, 2(1):35–58, 2006.
- [83] H. Hakobyan and D. A. Herron. Euclidean quasiconvexity. *Ann. Acad. Sci. Fenn. Math.*, 33(1):205–230, 2008.
- [84] W. Han and M. Sofonea. *Quasistatic contact problems in viscoelasticity and viscoplasticity*, volume 30 of *AMS/IP Studies in Advanced Mathematics*. American Mathematical Society, Providence, RI; International Press, Somerville, MA, 2002.
- [85] A. Haraux. How to differentiate the projection on a convex set in hilbert space. Some applications to variational inequalities. *J. Math. Soc. Japan*, 29(4):615–631, 10 1977.
- [86] J. Haslinger and T. Roubíček. Optimal control of variational inequalities. Approximation theory and numerical realization. *Appl. Math. Optim.*, 14(3):187–201, 1986.
- [87] J. Heinonen. *Lectures on Lipschitz analysis*, volume 100 of *Report. University of Jyväskylä Department of Mathematics and Statistics*. University of Jyväskylä, Jyväskylä, 2005.
- [88] J. Heinonen, T. Kilpeläinen, and O. Martio. *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Dover Books on Mathematics Series. Dover Publications, 2012.
- [89] W. Hensgen. A simple proof of Singer's representation theorem. *Proc. Amer. Math. Soc.*, 124(10):3211–3212, 1996.
- [90] L. Hertlein and M. Ulbrich. An inexact bundle algorithm for nonconvex nondifferentiable functions in hilbert space. Preprint SPP1962-084, 10 2018.

- [91] R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *SIAM J. Control Optim.*, 50(5):3052–3082, 2012.
- [92] R. Herzog, C. Meyer, and G. Wachsmuth. B- and strong stationarity for optimal control of static plasticity with hardening. *SIAM J. Optim.*, 23(1):321–352, 2013.
- [93] R. Herzog and E. Sachs. Preconditioned conjugate gradient method for optimal control problems with control and state constraints. *SIAM J. Matrix Anal. Appl.*, 31(5):2291–2317, 2010.
- [94] R. Herzog and W. Wollner. A conjugate direction method for linear systems in Banach spaces. *J. Inverse Ill-Posed Probl.*, 25(5):553–572, 2017.
- [95] M. Hintermüller. A proximal bundle method based on approximate subgradients. *Comput. Optim. Appl.*, 20(3):245–266, 2001.
- [96] M. Hintermüller and T. Surowiec. First-order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis. *SIAM J. Optim.*, 21(4):1561–1593, 2011.
- [97] G. A. Holzapfel. *Nonlinear solid mechanics*. John Wiley & Sons, Ltd., Chichester, 2000. A continuum approach for engineering.
- [98] S. Hübner, G. Stadler, and B. I. Wohlmuth. A primal-dual active set algorithm for three-dimensional contact problems with Coulomb friction. *SIAM J. Sci. Comput.*, 30(2):572–596, 2008.
- [99] T. J. R. Hughes. *The finite element method*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1987.
- [100] International Organization for Standardization. Implants for surgery - wear of total knee-joint prostheses. ISO-Standard 14243:2009, ISO, 2009.
- [101] K. Ito and K. Kunisch. Optimal control of parabolic variational inequalities. *J. Math. Pures Appl. (9)*, 93(4):329–360, 2010.
- [102] J. Jarušek and C. Eck. Remark to dynamic contact problems for bodies with a singular memory. *Comment.Math.Univ.Carolin. 3*, 39(3):545–550, 1998.
- [103] J. Jarušek and C. Eck. Dynamic contact problems with small Coulomb friction for viscoelastic bodies. Existence of solutions. *Math. Models Methods Appl. Sci.*, 9(1):11–34, 1999.
- [104] J. Jarušek. Dynamic contact problems with given friction for viscoelastic bodies. *Czechoslovak Math. J.*, 46(3):475–487, 1996.
- [105] L. Johansson and A. Klarbring. The rigid punch problem with friction using variational inequalities and linear complementarity. *Mech. Structures Mach.*, 20(3):293–319, 1992.

- [106] C. Kane, E. Repetto, M. Ortiz, and J. Marsden. Finite element analysis of nonsmooth contact. *Comput. Methods Appl. Mech. Engrg.*, 180(1-2):1–26, 1999.
- [107] Y. Kanno, J. A. C. Martins, and A. Pinto da Costa. Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem. *Internat. J. Numer. Methods Engrg.*, 65(1):62–83, 2006.
- [108] C. Kanzow and D. Steck. A generalized proximal-point method for convex optimization problems in Hilbert spaces. *Optimization*, 66(10):1667–1676, 2017.
- [109] A. M. Khludnev and J. Sokołowski. *Modelling and control in solid mechanics*, volume 122 of *International Series of Numerical Mathematics*. Birkhäuser Verlag, Basel, 1997.
- [110] N. Kikuchi and J. T. Oden. *Contact Problems in Elasticity*. Society for Industrial and Applied Mathematics, Philadelphia, 1988.
- [111] T. Kilpeläinen and J. Malý. Supersolutions to degenerate elliptic equation on quasi open sets. *Comm. Partial Differential Equations*, 17(3-4):371–405, 1992.
- [112] C. Klapproth. *Adaptive Numerical Integration of Dynamical Contact Problems*. PhD thesis, Freie Universität Berlin, 2010.
- [113] C. Klapproth, A. Schiela, and P. Deuffhard. Consistency results on newmark methods for dynamical contact problems. *Numer. Math.*, 116(1):65–94, 2010.
- [114] C. Klapproth, A. Schiela, and P. Deuffhard. Adaptive timestep control for the contact-stabilized Newmark method. *J. Numer. Math.*, 119(1):49–81, 2011.
- [115] A. Klarbring, A. Mikelić, and M. Shillor. The rigid punch problem with friction. *Internat. J. Engrg. Sci.*, 29(6):751 – 768, 1991.
- [116] R. Kornhuber and R. Krause. Adaptive Multigrid Methods for Signorini’s Problem in Linear Elasticity. *Comput. Vis. Sci.*, 4:9–20, 2001.
- [117] R. Krause. *Monotone Multigrid Methods for Signorini’s Problem with Friction*. phdthesis, Freie Universität Berlin, Fachbereich Mathematik und Informatik, 2001.
- [118] R. Krause and M. Walloth. Presentation and comparison of selected algorithms for dynamic contact based on the newmark scheme. *Appl. Numer. Math.*, 62(10):1393–1410, Oct. 2012.
- [119] E. Kreyszig. *Introductory functional analysis with applications*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1989.
- [120] A. Kröner, K. Kunisch, and B. Vexler. Semismooth newton methods for optimal control of the wave equation with control constraints. *SIAM J. Control Optim.*, 49(2):830–858, 2011.
- [121] A. Kufner, O. John, and S. Fučík. *Function spaces*. Noordhoff International Publishing, Leyden; Academia, Prague, 1977. Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis.

- [122] S. Kurcyusz. On the existence and non-existence Lagrange multipliers in Banach spaces. *J. Optimization Theory Appl.*, 20(1):81–110, 1976.
- [123] K. Kuttler and M. Shillor. Dynamic contact with Signorini’s condition and slip rate dependent friction. *Electron. J. Differential Equations*, pages No. 83, 21, 2004.
- [124] T. A. Laursen. *Computational contact and impact mechanics*. Springer-Verlag, Berlin, 2002. Fundamentals of modeling interfacial phenomena in nonlinear finite element analysis.
- [125] T. A. Laursen. *Computational Contact and Impact Mechanics: Fundamentals of Modeling Interfacial Phenomena in Nonlinear Finite Element Analysis*. Springer Berlin Heidelberg, 2013.
- [126] T. A. Laursen and V. Chawla. Design of energy conserving algorithms for frictionless dynamic contact problems. *Internat. J. Numer. Methods Engrg.*, 40(5):863–886, 1997.
- [127] J. M. Lee. *Manifolds and differential geometry*, volume 107 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2009.
- [128] S. Leyffer, G. López-Calva, and J. Nocedal. Interior methods for mathematical programs with complementarity constraints. *SIAM J. Optim.*, 17(1):52–77, 2006.
- [129] S. Leyffer and T. Munson. A globally convergent filter method for mpecs. Preprint ANL/MCS-P1457-0907, Argonne National Laboratory, 10 2007.
- [130] J. L. Lions and G. Stampacchia. Variational inequalities. *Communications on Pure and Applied Mathematics*, 20(3):493–519, 1967.
- [131] X. Liu and J. Sun. Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints. *Math. Program.*, 101(1, Ser. B):231–261, 2004.
- [132] W. A. J. Luxemburg and A. C. Zaanen. *Riesz spaces. Vol. I*. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., New York, 1971.
- [133] Y. Malitsky. Proximal extrapolated gradient methods for variational inequalities. *Optim. Methods Softw.*, 33(1):140–164, 2018.
- [134] J. A. C. Martins and J. T. Oden. Existence and uniqueness results for dynamic contact problems with nonlinear normal and friction interface laws. *Nonlinear Anal.*, 11(3):407–428, 1987.
- [135] V. G. Maz’ya and S. V. Poborchi. *Differentiable functions on bad domains*. World Scientific Publishing Co., Inc., River Edge, NJ, 1997.
- [136] R. Meise and D. Vogt. *Introduction to functional analysis*, volume 2 of *Oxford Graduate Texts in Mathematics*. The Clarendon Press, Oxford University Press, New York, 1997. Translated from the German by M. S. Ramanujan and revised by the authors.
- [137] C. Meyer and L. M. Susu. Optimal control of nonsmooth, semilinear parabolic equations. *SIAM J. Control Optim.*, 55(4):2206–2234, 2017.

- [138] F. Mignot. Contrôle dans les inéquations variationelles elliptiques. *J. Functional Analysis*, 22(2):130–185, 1976.
- [139] F. Mignot and J.-P. Puel. Contrôle optimal d’un système gouverné par une inéquation variationnelle parabolique. *C. R. Acad. Sci. Paris Sér. I Math.*, 298(12):277–280, 1984.
- [140] F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM J. Control Optim.*, 22(3):466–476, 1984.
- [141] S. Migórski and A. Ochal. A unified approach to dynamic contact problems in viscoelasticity. *J. Elasticity*, 83(3):247–275, 2006.
- [142] J. Mujica. Representation of analytic functionals by vector measures. In R. M. Aron and S. Dineen, editors, *Vector Space Measures and Applications II*, pages 147–161, Berlin, Heidelberg, 1978. Springer Berlin Heidelberg.
- [143] G. Müller and A. Schiela. On the control of time discretized dynamic contact problems. *Comput. Optim. Appl.*, 68(2):243–287, Nov 2017.
- [144] J. Munkres. *Topology*. Featured Titles for Topology Series. Prentice Hall, Incorporated, 2000.
- [145] Y. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Dokl. Akad. Nauk SSSR*, 269(3):543–547, 1983.
- [146] Y. Nesterov. *Introductory lectures on convex optimization*, volume 87 of *Applied Optimization*. Kluwer Academic Publishers, Boston, MA, 2004.
- [147] Y. Nesterov. Gradient methods for minimizing composite functions. *Math. Program.*, 140(1, Ser. B):125–161, 2013.
- [148] J. Nečas. *Direct methods in the theory of elliptic equations*. Springer Monographs in Mathematics. Springer, Heidelberg, 2012.
- [149] J. Nečas, J. Jarušek, and J. Haslinger. On the solution of the variational inequality to the Signorini problem with small friction. *Boll. Un. Mat. Ital. B (5)*, 17(2):796–811, 1980.
- [150] N. M. Newmark. A method of computation for structural dynamics. *J. Eng. Mech. Div-ASCE.*, 85(3), 1959.
- [151] J. A. Nitsche. On Korn’s second inequality. *RAIRO Anal. Numér.*, 15(3):237–248, 1981.
- [152] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, 2nd edition, 2006.
- [153] B. O’Donoghue and E. Candès. Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.*, 15(3):715–732, 2015.

- [154] Y. Oshima. *Semi-Dirichlet forms and Markov processes*, volume 48 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2013.
- [155] J. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*. Nonconvex Optimization and Its Applications. Springer US, 2013.
- [156] J. Peetre. A counterexample connected with Gagliardo’s trace theorem. *Comment. Math. Special Issue*, 2:277–282, 1979.
- [157] D. Preiss. Differentiability of Lipschitz functions on Banach spaces. *J. Funct. Anal.*, 91(2):312–345, 1990.
- [158] T. Qian, T. Hempfling, A. McIntosh, and F. Sommen. *Advances in Analysis and Geometry: New Developments Using Clifford Algebras*. Trends in Mathematics. Birkhäuser Basel, 2012.
- [159] D. Ralph and S. J. Wright. Some properties of regularization and penalization schemes for MPECs. *Optim. Methods Softw.*, 19(5):527–556, 2004.
- [160] M. Rao and J. Sokołowski. Polyhedricity of convex sets in Sobolev space $H_0^2(\Omega)$. *Nagoya Math. J.*, 130:101–110, 1993.
- [161] A.-T. Rauls and S. Ulbrich. Subgradient computation for the solution operator of the obstacle problem. Preprint SPP1962-056, 5 2018.
- [162] A.-T. Rauls and G. Wachsmuth. Generalized derivatives for the solution operator of the obstacle problem. Preprint SPP1962-057, 6 2018.
- [163] O. Sander. *Multidimensional coupling in a human knee model*. PhD thesis, Freie Universität Berlin, 2008.
- [164] O. Sander, C. Klapproth, J. Youett, R. Kornhuber, and P. Deuffhard. Towards an efficient numerical simulation of complex 3D knee joint motion. *Comput. Vis. Sci.*, 16(3):119–138, 2013.
- [165] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. *Math. Oper. Res.*, 25(1):1–22, 2000.
- [166] K. Scheinberg, D. Goldfarb, and X. Bai. Fast first-order methods for composite convex optimization with backtracking. *Found. Comput. Math.*, 14(3):389–417, 2014.
- [167] A. Schiela. A flexible framework for cubic regularization algorithms for non-convex optimization in function space. 2014.
- [168] A. Schiela and M. Stöcklein. Optimal control of static contact in finite strain elasticity - SPP 1962 Project 18, 2018.
- [169] A. Schiela and D. Wachsmuth. Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints. *ESAIM Math. Model. Numer. Anal.*, 47(3):771–787, 2013.

- [170] J. Schmets. *Spaces of vector-valued continuous functions*, volume 1003 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.
- [171] H. Schramm. *Eine Kombination von Bundle- und Trust-Region-Verfahren zur Lösung nichtdifferenzierbarer Optimierungsprobleme*. Bayreuther mathematische Schriften. Univ. Bayreuth, 1989.
- [172] H. Schramm and J. Zowe. A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. *SIAM J. Optim.*, 2(1):121–152, 1992.
- [173] I. Shames. *Introduction To Solid Mechanics*. Prentice Hall College Division, 1989.
- [174] A. Shapiro. On concepts of directional differentiability. *J. Optim. Theory Appl.*, 66(3):477–487, 1990.
- [175] A. Signorini. Sopra alcune questioni di elastostatic. *Soc. Ital. per il Prog. delle Sci. Atti.*, 1933.
- [176] I. Singer. Sur les applications linéaires intégrales des espaces de fonctions continues. I. *Rev. Math. Pures Appl.*, 4:391–401, 1959.
- [177] W. S. Slaughter. *The linearized theory of elasticity*. Birkhäuser Boston, Inc., Boston, MA, 2002.
- [178] M. Sofonea, W. Han, and M. Shillor. *Analysis and approximation of contact problems with adhesion or damage*, volume 276 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2006.
- [179] M. Sofonea and Y. Souleiman. Analysis of a sliding frictional contact problem with unilateral constraint. *Math. Mech. Solids*, 22(3):324–342, 2017.
- [180] I. S. Sokolnikoff. *Mathematical theory of elasticity*. Robert E. Krieger Publishing Co., Inc., Melbourne, Fla., second edition, 1983.
- [181] J. Sokołowski. Sensitivity analysis of contact problems with prescribed friction. *Appl. Math. Optim.*, 18(2):99–117, 1988.
- [182] J. Sokołowski and J.-P. Zolesio. *Introduction to shape optimization : shape sensitivity analysis*. Springer series in computational mathematics. Springer, New York, Heidelberg, Paris, 1992.
- [183] G. Stadler. Path-following and augmented Lagrangian methods for contact problems in linear elasticity. *J. Comput. Appl. Math.*, 203(2):533–547, 2007.
- [184] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
- [185] M. Stöcklein. Optimal control of static contact problems in linear elasticity - M.Sc. thesis, July 2016.
- [186] P. Stollmann. Closed ideals in dirichlet spaces. *Potential Anal.*, 2(3):263–268, 1993.

- [187] A. Stötzner. *Optimal Control of Thermoviscoplasticity*. PhD thesis, Technische Universität Chemnitz, 2018. To appear.
- [188] L. M. Susu. Optimal control of a viscous two-field gradient damage model. *GAMM-Mitt.*, 40(4):287–311, 2018.
- [189] N. S. Trudinger and X.-J. Wang. Quasilinear elliptic equations with signed measure data. *Discrete Contin. Dyn. Syst.*, 23(1-2):477–494, 2009.
- [190] C. Truesdell, S. Antman, D. Carlson, G. Fichera, M. Gurtin, and P. Naghdi. *Mechanics of Solids: Volume II: Linear Theories of Elasticity and Thermoelasticity, Linear and Nonlinear Theories of Rods, Plates, and Shells*. Springer Berlin Heidelberg, 1984.
- [191] M. Ulbrich, S. Ulbrich, and D. Koller. A multigrid semismooth newton method for contact problems in linear elasticity. Technical report, Technical Report, Department of Mathematics, TU Darmstadt, 2013. submitted.
- [192] G. Wachsmuth. *Optimal control of quasistatic plasticity - An MPCC in function space*. PhD thesis, Technische Universität Chemnitz, 2011.
- [193] G. Wachsmuth. Mathematical programs with complementarity constraints in banach spaces. *J. Optim. Theory Appl.*, pages 1–28, 2014.
- [194] G. Wachsmuth. Strong stationarity for optimal control of the obstacle problem with control constraints. *SIAM J. Optim.*, pages 1914–1932, 2014.
- [195] G. Wachsmuth. A guided tour of polyhedral sets. *Preprint*, 2016.
- [196] G. Wachsmuth. Towards M-stationarity for optimal control of the obstacle problem with control constraints. *SIAM J. Control Optim.*, 54(2):964–986, 2016.
- [197] J. Ward. *Solid Mechanics: An Introduction*. Solid Mechanics and Its Applications. Springer Netherlands, 1992.
- [198] J. Wehrstedt. *Formoptimierung mit Variationsungleichungen als Nebenbedingung und eine Anwendung in der Kieferchirurgie*. PhD thesis, 2007.
- [199] M. Weiser, P. Deuffhard, and B. Erdmann. Affine conjugate adaptive Newton methods for nonlinear elastomechanics. *Optim. Methods Softw.*, 22(3):413–431, 2007.
- [200] D. Werner. *Funktionalanalysis*. Springer Verlag, Heidelberg, third edition, 2000.
- [201] D. Werner. *Einführung in die höhere Analysis: Topologische Räume, Funktionentheorie, Gewöhnliche Differentialgleichungen, Maß- und Integrationstheorie, Funktionalanalysis Index.- Literaturverzeichnis*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2009.
- [202] A. Wilanski. *Topology for Analysis*. Ginn, Waltham, MA, 1970.
- [203] S. Willard. *General Topology*. Addison-Wesley, London, 1970.

-
- [204] K. Yosida. *Functional analysis*. Classics in Mathematics. Springer-Verlag, Berlin, 1995.
- [205] D. Yost. Asplund spaces for beginners. *Acta Univ. Carolin. Math. Phys.*, 34(2):159–177, 1993.
- [206] J. Youett. *Dynamic large deformation contact problems and applications in virtual medicine*. PhD thesis, Freie Universität Berlin, 2015.
- [207] E. Zeidler. *Nonlinear functional analysis and its applications. II/A*. Springer-Verlag, New York, 1990. Linear monotone operators.
- [208] W. P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.
- [209] O. C. Zienkiewicz and R. L. Taylor. *The finite element method. Vol. 3*. Butterworth-Heinemann, Oxford, fifth edition, 2000. Fluid dynamics.

Publications

- [44] C. Christof and G. Müller. A note on the equivalence and the boundary behavior of a class of sobolev capacities. *GAMM-Mitt.*, 40(3):238–266, Apr 2018, doi:10.1002/-gamm.201730005
- [143] G. Müller and A. Schiela. On the control of time discretized dynamic contact problems. *Comput. Optim. Appl.*, 68(2):243–287, Nov 2017, doi: 10.1007/s10589-017-9918-5

Several parts of this thesis are contained in the publications above. The introduction in Chapter 1 is based on the overview in [143]. Chapter 3 is based on the results that have been published in [44] and includes additional results that concern the implications for the modeling of contact problems. The analysis in [143] serves as the foundation of Chapters 4 and 5, which include extended and refined results in the sensitivity analysis of the forward problem and the adjoint problem. Additionally, the initial algorithmic approach of the numerical results that are presented in Chapter 6 and the example in Section 6.2.2 are contained in [143] as well. This thesis contains a considerable extension of the numerics of this example as well as additional optimization algorithms and numerical results for other test configurations.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Weiterhin erkläre ich, dass ich die Hilfe von gewerblichen Promotionsberatern bzw. Promotionsvermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe, noch künftig in Anspruch nehmen werde.

Ich erkläre zudem, dass ich keine früheren Promotionsversuche unternommen habe.

Konstanz, den _____

(Georg Müller)