ASC Report No. 30/2008

hp-Finite Element Method for Variational Inequalities

Philipp Dörsek

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-00-1

Most recent ASC Reports

29/2008 *Bertram Düring, Daniel Matthes, Guiseppe Toscani*
A Boltzmann-type Approach to the Formation of Wealth Distribution Curves

28/2008 *Heinrich Freistühler, Peter Szmolyan*
Spectral Stability of Small-Amplitude Viscous Shock Waves in Several Space Dimensions

27/2008 *Stephen Schecter, Peter Szmolyan*
Rarefactions in the Dafermos Regularization of a System of Conservation Laws

26/2008 *Ilona Gucwa, Peter Szmolyan*
Geometric Singular Perturbation Analysis of an Autocatalator Model

25/2008 *Christoph Ortner, Dirk Praetorius*
On the Convergence of Adaptive Non-Conforming Finite Element Methods

24/2008 *Christoph Ortner, Dirk Praetorius*
A Non-Conforming Finite Element Method for Convex Variational Problems

23/2008 *Samuel Ferraz-Leite, Dirk Praetorius*
Adaptive Boundary Element Methods Based on Accurate a Posteriori Error Estimation

22/2008 *Jeff Cash, Georg Kitzhofer, Othmar Koch, Gerald Moore, Ewa Weinmüller*
Numerical Solutions of singular Two Point BVPs

21/2008 *Samuel Ferraz-Leite, Dirk Praetorius*
A Posteriori Fehlerschätzer für die Symmsche Integralgleichung in 3D

20/2008 *Bertram Düring*
Asset Pricing under Information with Stochastic Volatility

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-00-1

ASC
TU WIEN

Unterschrift des Betreuers

DIPLOMARBEIT

hp-Finite Element Methods for Variational Inequalities

Ausgeführt am Institut für

Analysis und Scientific Computing
der Technischen Universität Wien

unter der Anleitung von Univ.Prof. Jens Markus Melenk, PhD

durch

Philipp Dörsek

Döblinger Hauptstraße 7A/64
A-1190 Wien

Datum

Unterschrift

Abstract

This diploma thesis presents the mathematical theory and numerical analysis of the contact problem with Tresca friction in plane elasticity. We give an overview of the mathematical formulation of this problem as a variational inequality of the second kind, and prove the existence and uniqueness of the corresponding displacement field using methods of convex analysis. Furthermore, we introduce a primal-dual formulation, where the nonlinear friction functional is replaced by using a Lagrange multiplier function on the contact boundary.

Next, we analyse how the given problem can be appropriately discretised. It is well known that p-finite element methods can yield exponential convergence, but only if the exact solution is smooth on all elements of the employed mesh. As the displacement field is expected to be nonsmooth near those parts of the contact boundary where the boundary conditions change from sticking to sliding, in addition to corners and transitions between Dirichlet and Neumann boundaries, however, this assumption is not justified for the presented problem. Therefore, in the numerical analysis, we focus on hp-methods. These methods combine fine grids at points where the solution is irregular with high polynomial degrees on elements where it is smooth. We prove a general convergence result for hp-finite element approximations on meshes with arbitrary element size and polynomial degree distributions. Furthermore, given sufficient regularity, we obtain convergence rates using a novel hp-mortar projection operator, which uses a discontinuous Lagrange multiplier space on the boundary.

As the information on the regularity of the exact solution, which is necessary to construct an appropriate mesh, is not available a priori, we apply an error indicator of residual type, generalised to our context, to determine those elements where the local error appears high and which thus should be refined in an adaptive computation. For these elements, we then estimate the local regularity of the solution using the rate of decay of the Legendre series coefficients of the given numerical approximation. Based on this, we decide whether to subdivide the element or increase the polynomial degree.

We finally show numerical results which confirm our analysis. The adaptive methods are able to resolve the irregularities of the solution properly, and give rates of convergence that are significantly higher than those of uniform mesh refinements. In particular, the hp-method empirically yields exponential convergence.

Acknowledgements

First and foremost, I want to thank my advisor, Professor Markus Melenk, for his support while tackling the problem analysed in this thesis. Without his help, this work would not have been possible.

Second, I want to thank Professor Matthias Maischak, for providing me with his finite element/boundary element code `maiprogs`, and also for implementing several additional features during the course of this diploma thesis which were very useful to me.

Last, I want to express my sincerest gratitude to my parents, for aiding me and letting me enough freedom at the same time. Without their continued moral and financial support, I would not be standing where I am now.

Contents

Introduction

Today, mathematical models are an indispensable tool in science and engineering. They give the practitioner the opportunity to predict the behaviour of complex systems, and as such save significant amounts of money, as only those prototypes will be built that have proved usable in the model.

Typically, these models are very complex. They are often given as systems of ordinary or partial differential equations, and can in general not be solved explicitly. Thus, numerical methods are central to the simulation of technical processes, and this makes the development of efficient numerical schemes for a large range of mathematical problems necessary.

The present work deals with a certain kind of problem arising in technical applications. We want to simulate an elastic body which has frictional contact with a fixed object. Due to the presence of the frictional contact, we do not have a partial differential equation, but a partial differential inequality, more specifically, a variational inequality of the second kind; this formulation was pioneered in [DL76]. Moving from an equation to an inequality leads to several difficulties in the numerical simulation. Because it is straightforward to solve a linear system of equations, the Newton algorithm is a standard approach for solving nonlinear system of equations, and there are several high-performance methods for solving linear variational inequalities of the first kind. The presence of a nontrivial, convex, nondifferentiable functional in variational inequalities of the second kind, however, makes it necessary to use different algorithms. One approach is the primal method, described in detail in [Kor97], where also some ways to accelerate the convergence rate through the use of adequate preconditioners are given.

We focus on a different idea: Due to the special structure of the nondifferentiable functional, it is possible to construct a primal-dual formulation as a saddle point problem, which is described in detail in [HHNL88]. This leads to a coupled system of variational inequalities of the first kind, and under certain assumptions can be reduced to a single variational inequality of the first kind on the contact boundary by first solving the problem on the domain and using the Schur complement of the system matrix, as done in [Sin06].

Furthermore, we investigate the use of high order *hp*-methods in this context. For these, it is essential to have a well-constructed mesh, and this can only be done by either knowing the problem relatively well, or using an adaptive process, as we expect the solution to have singularities, not only at corners and transitions between Dirichlet and Neumann boundaries, but also at the unknown points of the contact boundary where the boundary conditions change from sticking to sliding. Adaptive algorithms are based on local error indicators and those in *hp*-FEM typically use an estimation of the local regularity of the solution.

The error indicator employed in this work is a standard residual error indicator, generalised appropriately to the primal-dual formulation as suggested in [Han05]. These indicators are relatively easy to implement, and deliver acceptable results to steer the refinement process. As a stopping criterion, however, their reliability and efficiency properties are not good enough. For methods enlarging the polynomial degree, in particular, the fact that the error indicator might overestimate the error by up to a factor p is a significant problem.

To decide whether to do an h-refinement, more appropriate for a singular, or a p-

refinement, more appropriate for a regular solution, on an element where the estimated error is large, the estimation of the local smoothness of the solution is done by expanding the approximate solution into a Legendre series. Depending on the decay rate of the coefficients in this series, calculated by a least squares method, we do an h- or a p-refinement. This method, based on theoretical results on Legendre series for analytic functions in [Dav63, HS05, EM07], yields very good empiric convergence rates in our numerical experiments, and can thus be recommended for practical computations: In particular, we empirically obtain exponential convergence in one model problem.

This diploma thesis is arranged as follows. In Chapter 1, we give an overview of the mathematical basics which are necessary for an understanding of the later topics. This includes, in particular, several results on Sobolev spaces.

Chapter 2 contains the mathematical formulation of the frictional contact problem. In particular, we give a primal-dual or saddle point formulation, where we can reduce the variational inequality of the second kind on the domain to a variational inequality of the first kind on the boundary.

Based on the primal-dual formulation, in Chapter 3, we construct an hp-finite element approximation of the frictional contact problem. We use a theorem by Glowinski to prove the strong convergence of the method, and a theorem by Haslinger on mixed methods for variational inequalities, together with a new kind of hp-mortar protection operators, to obtain an a priori estimate on the convergence rate.

Chapter 4 contains the formulation of the residual error indicator as given in [Han05] for the hp-approximation of the frictional contact problem. We obtain reliability and efficiency up to a factor p and certain terms which can be expected to be of higher order if the mesh is well chosen.

Finally, in Chapter 5, we show some numerical experiments which support our theoretical results.

Chapter 1

Mathematical Preliminaries

The aim of this chapter is to collect the mathematical tools needed in the following parts of this work.

1.1 Vectors and Tensors

For the convenience of the reader, we quickly repeat the notation that we already used in [Dör07].

We use the *Einstein summation convention*, that is, if there is a repeated index in a single term, we sum over it. Letting $d \in \mathbb{N}$ be the dimension, we say that a *change of coordinates* is an affine mapping

$$y_i = a_{ij}x_j + c_j, \quad i = 1, \ldots, d, \tag{1.1}$$

where the linear part is given by an orthogonal matrix $\mathbf{A} = (a_{ij})_{i,j=1,\ldots,d} \in \mathbb{R}^{d \times d}$, that is, $a_{ij}a_{kj} = \delta_{ik}$, and $\mathbf{c} = (c_i)_{i=1,\ldots,d} \in \mathbb{R}^d$. The inverse change of coordinates is then given as

$$x_j = a_{ji}y_i - a_{ji}c_i, \quad j = 1, \ldots, d. \tag{1.2}$$

A *tensor of order* (or *rank*) N is a mapping $\mathbf{T} = (T_{i_1 \ldots i_N})$, $i_k = 1, \ldots, d$, $k = 1, \ldots, N$ from the set of Cartesian coordiate systems to $(\mathbb{R}^d)^N$ which transforms by the rule

$$T'_{i_1 \ldots i_N} = a_{i_1 j_1} \ldots a_{i_N j_N} T_{j_1 \ldots j_N}, \quad i_k = 1, \ldots, d, \; k = 1, \ldots, N, \tag{1.3}$$

whenever we apply the change of coordinates $y_i = a_{ij}x_j + c_i$, $i = 1, \ldots, d$.

The *trace* (or *contraction*) of a tensor of order $N \geqslant 2$ is obtained by setting two different indices equal. Thus, the trace of a tensor of order N is a tensor of order $N - 2$. If $\mathbf{T} = (T_{ij})_{i,j=1,\ldots,d}$ is a matrix, we recover the usual trace $\operatorname{tr} \mathbf{T} = T_{ii}$, the sum over the diagonal elements.

For vectors in $\mathbb{R}^d$, we define the *Euclidean inner product*

$$\mathbf{x} \cdot \mathbf{y} := x_i y_i \tag{1.4}$$

with its induced *Euclidean norm* $|\mathbf{x}| := (\mathbf{x} \cdot \mathbf{x})^{1/2}$, and for matrices in $\mathbb{R}^{d \times d}$, we define the *Frobenius inner product*

$$\mathbf{A} : \mathbf{B} := a_{ij} b_{ij} \tag{1.5}$$

with its induced norm $|\mathbf{A}| := (\mathbf{A} : \mathbf{A})^{1/2}$, the *Frobenius norm*.

Partial derivatives are denoted by

$$v_{,j} := \frac{\partial v}{\partial x_j}, \quad j = 1, \ldots, d. \tag{1.6}$$

A *multi-index* $\boldsymbol{\alpha} = (\alpha_k)_{k=1,\ldots,d}$ is an element of $\mathbf{N}_0^d$. Its *order* is denoted by $|\boldsymbol{\alpha}| := \sum_{k=1}^{d} \alpha_k$, and its maximum by $\max \boldsymbol{\alpha} := \max_{k=1,\ldots,d} \alpha_k$. For a (sufficiently regular) function v, we set $D^{\boldsymbol{\alpha}} v := v_{,1^{\alpha_1}\ldots d^{\alpha_d}}$, that is,

$$D^{\boldsymbol{\alpha}} v := \frac{\partial^{|\boldsymbol{\alpha}|} v}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}}. \tag{1.7}$$

1.2 Sobolev Spaces

We give a thorough introduction into the part of the theory of Sobolev spaces which we need in the following. Standard references are [Ada75, Eva98, RR04, Bre83, Gri85].

Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain, and $\mathrm{L}^2(\Omega)$ the Hilbert space of (equivalence classes of) real valued, square-integrable functions on $\Omega \subseteq \mathbb{R}^d$ endowed with the norm

$$\|v\|_{\mathrm{L}^2(\Omega)} := \left(\int_\Omega v^2 \mathrm{d}\mathbf{x} \right)^{1/2}. \tag{1.8}$$

The *Sobolev space* $\mathrm{H}^1(\Omega)$ is the Hilbert space of all elements of $\mathrm{L}^2(\Omega)$ such that the weak derivatives are again in $\mathrm{L}^2(\Omega)$, and carries the seminorm and norm

$$|v|_{\mathrm{H}^1(\Omega)} := \left(\sum_{j=1}^{d} \|v_{,j}\|^2_{\mathrm{L}^2(\Omega)} \right)^{1/2}, \tag{1.9}$$

$$\|v\|_{\mathrm{H}^1(\Omega)} := \left(\|v\|^2_{\mathrm{L}^2(\Omega)} + |v|^2_{\mathrm{H}^1(\Omega)} \right)^{1/2}. \tag{1.10}$$

Starting from $\mathrm{H}^1(\Omega)$, higher order Sobolev spaces can be defined recursively; then, for $k \geqslant 1$, $\mathrm{H}^{k+1}(\Omega)$ is the Hilbert space of all elements of $\mathrm{H}^1(\Omega)$ such that its weak derivatives are in $\mathrm{H}^k(\Omega)$.

Analogous definitions are also possible for unbounded domains, but then, for several theorems formulated below, in particular the trace theorem, additional assumptions are required.

Theorem 1.1 (Meyers-Serrin). *Assume that $\Omega \subseteq \mathbb{R}^d$ is open, bounded, and has a* Lipschitz boundary*, that is, $\partial\Omega$ can be locally parametrised by Lipschitz functions, and Ω is locally on one side of its boundary $\partial\Omega$.*

Then, for $k \in \mathbb{N}_0$, the set $C^\infty(\overline{\Omega})$ of functions which are infinitely often differentiable on a neighbourhood of $\overline{\Omega}$ is dense in $\mathrm{H}^k(\Omega)$.

Furthermore, we shall define *fractional order Sobolev spaces* in the following way. Let $s = k + \theta$, where $k \in \mathbb{N}$ and $\theta \in (0, 1)$ (for integer s, we use the above definitions). Denoting the *Slobodeckij seminorm* by

$$|v|_{\mathrm{H}^s(\Omega)} := \left(\sum_{|\boldsymbol{\alpha}|=k} \int_\Omega \int_\Omega \frac{|D^{\boldsymbol{\alpha}} v(\mathbf{x}) - D^\alpha v(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|^{d+2\theta}} \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \right)^{1/2}, \tag{1.11}$$

we set

$$\mathrm{H}^s(\Omega) := \left\{ v \in \mathrm{H}^k(\Omega) \colon\ |v|_{\mathrm{H}^s(\Omega)} < \infty \right\}, \tag{1.12}$$

and endow $\mathrm{H}^s(\Omega)$ with the norm

$$\|v\|_{\mathrm{H}^s(\Omega)} := \left(\|v\|^2_{\mathrm{H}^k(\Omega)} + |v|^2_{\mathrm{H}^s(\Omega)} \right)^{1/2}. \tag{1.13}$$

For ease of notation, we also define $\mathrm{H}^0(\Omega) := \mathrm{L}^2(\Omega)$ and

$$|v|_{\mathrm{H}^0(\Omega)} := \|v\|_{\mathrm{L}^2(\Omega)} . \tag{1.14}$$

Using local parametrisations, it is possible to define Sobolev spaces $\mathrm{H}^s(\Gamma)$ for $\Gamma \subseteq \mathbb{R}^d$ a $(d-1)$-dimensional manifold. We shall make use of these spaces in the following; the more technical details of their definition can be found in [Sch98].

In order to be able to deal with boundary conditions, we need to check in which sense it is possible to evaluate functions in Sobolev spaces on the boundary. This question is answered by the following results.

Theorem 1.2 (Trace theorem)**.** *Let $\Omega \subseteq \mathbb{R}^d$ be open, bounded with Lipschitz boundary, and $\Gamma_1 \subseteq \Gamma := \partial\Omega$ be relatively open with positive surface measure, and $s \in (1/2, 3/2)$.*

Then, there exists a continuous linear operator $\gamma_{0,\Gamma_1} : \mathrm{H}^s(\Omega) \to \mathrm{L}^2(\Gamma_1)$, the trace operator*, satisfying $\gamma_{0,\Gamma_1} u = u|_{\Gamma_1}$ whenever $u \in \mathrm{C}^0(\overline{\Omega}) \cap \mathrm{H}^s(\Omega)$.*

Note that the above result does not yet characterise the range of γ_{0,Γ_1}. Setting $\gamma_0 := \gamma_{0,\Gamma}$, the next result shows when it is possible to lift a boundary condition to the domain.

Theorem 1.3 (Inverse trace theorem)**.** *Under the assumptions of the trace theorem, we have that $\gamma_0 \mathrm{H}^s(\Omega) = \mathrm{H}^{s-1/2}(\Gamma)$.*

In particular, there is a linear, continuous lifting operator *$Z\colon \mathrm{H}^{s-1/2}(\Gamma) \to \mathrm{H}^s(\Omega)$ such that $\gamma_0 Z v = v$ for all $v \in \mathrm{H}^{s-1/2}(\Gamma)$.*

A similar problem is to extend a function given on a bounded domain $\Omega \subseteq \mathbb{R}^d$ to $\mathbb{R}^d$. This is possible in a very general way; we shall only need the following result, which is given in [Ada75, Theorem 4.26].

Theorem 1.4 (Extension operator in one dimension)**.** *There exists an* extension operator *$\hat{E}\colon \mathrm{L}^2(0,1) \to \mathrm{L}^2(\mathbb{R})$ such that*

$$\left\|\hat{E}v\right\|_{\mathrm{L}^2(\mathbb{R})} \leqslant C \|v\|_{\mathrm{L}^2(0,1)} \quad \text{for } v \in \mathrm{L}^2(0,1), \tag{1.15}$$

$$\left\|\hat{E}v\right\|_{\mathrm{H}^1(\mathbb{R})} \leqslant C \|v\|_{\mathrm{H}^1(0,1)} \quad \text{for } v \in \mathrm{H}^1(0,1), \tag{1.16}$$

and

$$(\hat{E}v)|_{(0,1)} = v \quad \text{for all } v \in \mathrm{L}^2(0,1). \tag{1.17}$$

In general, the H^s-seminorm is clearly not a norm, as it vanishes on constant functions. For $s \leqslant 1$, however, these are the only functions for which this happens, as the following result shows.

Theorem 1.5 (Generalised Deny-Lions lemma)**.** *For $0 \leqslant s \leqslant 1$, there exists a constant $C > 0$ such that*

$$\inf_{z \in \mathbb{R}} \|v - z\|_{\mathrm{H}^s(\Omega)} \leqslant C |v|_{\mathrm{H}^s(\Omega)} . \tag{1.18}$$

For $s = 0$, this is trivial; for $0 < s < 1$, this follows from [DS80, Theorem 6.1]; and for $s = 1$, this is the well-known Poincaré inequality (see [Eva98, Section 5.8.1, Theorem 1]). A similar result, also known as Poincaré inequality and given in [Eva98, Section 5.6.1, Theorem

3], uses the fact that the function vanishes on the boundary of the domain to deduce the fact that the H^1-seminorm is actually a norm. We define $\mathrm{H}_0^1(\Omega) := \{v \in \mathrm{H}^1(\Omega) \colon \gamma_0 v = 0\}$. Equivalently, $\mathrm{H}_0^1(\Omega)$ is given as the closure of the space $\mathrm{D}(\Omega)$ of *test functions*, that is, of infinitely often differentiable functions with support strictly contained in Ω, with respect to the norm of $\mathrm{H}^1(\Omega)$.

Theorem 1.6 (Poincaré inequality)**.** *There exists a constant $C > 0$ such that for all bounded domains $\Omega \subseteq \mathbb{R}^d$,*

$$\|v\|_{\mathrm{H}^1(\Omega)} \leqslant C(1 + \operatorname{diam}\Omega)\, |v|_{\mathrm{H}^1(\Omega)} \quad \textit{for all } v \in \mathrm{H}_0^1(\Omega). \tag{1.19}$$

Scaling arguments make use how Sobolev norms behave if we map the domain Ω to another domain. First of all, we note that all the Sobolev norms are, due to the translation invariance of the Lebesgue measure, equally translation invariant. If we scale the domain Ω, that is, we consider the mapping $F \colon \Omega \to r\Omega$, $\mathbf{x} \mapsto F(\mathbf{x}) := r\mathbf{x}$, then we have, by the transformation theorem for multidimensional integrals, that for all $v \in \mathrm{L}^2(r\Omega)$,

$$\|v\|_{\mathrm{L}^2(r\Omega)}^2 = r^d \,\|v \circ F\|_{\mathrm{L}^2(\Omega)}^2 , \tag{1.20}$$

for all $v \in \mathrm{H}^1(r\Omega)$,

$$|v|_{\mathrm{H}^1(r\Omega)}^2 = \sum_{j=1}^{d} \|v_{,j}\|_{\mathrm{L}^2(r\Omega)}^2 = \sum_{j=1}^{d} r^d \left\|r^{-1}(v \circ F)_{,j}\right\|_{\mathrm{L}^2(\Omega)}^2 = r^{d-2}\,|v \circ F|_{\mathrm{H}^1(\Omega)}^2 , \tag{1.21}$$

and for fractional order Sobolev spaces, $s \in (0,1)$, $v \in \mathrm{H}^s(r\Omega)$,

$$\begin{aligned} |v|_{\mathrm{H}^s(r\Omega)}^2 &= \int_{r\Omega}\int_{r\Omega} \frac{|v(\mathbf{x}) - v(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^{d+2s}} \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} = r^{d-2s} \int_{\Omega}\int_{\Omega} \frac{|v(F(\mathbf{s})) - v(F(\mathbf{t}))|}{|\mathbf{s} - \mathbf{t}|^{d+2s}} \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{t} \\ &= r^{d-2s}\,|v \circ F|_{\mathrm{H}^s(\Omega)}^2 . \end{aligned} \tag{1.22}$$

Remark 1.7. Similar results also hold true if F is a more complicated, one-to-one and onto function $F \colon \hat{\Omega} \to \Omega$, where $\hat{\Omega}$ is usually called the *reference element*; then, we only obtain inequalities. The constants appearing only depend on the product of the Frobenius norms of the Jacobian $DF = (F_{i,j})_{i,j=1,\dots,d}$ and $D(F^{-1}) = ((F^{-1})_{i,j})_{i,j=1,\dots,d}$; in particular, if we have *regular* meshes (see Section 3.2), the powers of the diameter of the domain are the same. Together with approximation results on the reference element such as Theorem 1.5, we see that this yields convergence of h-versions. For hp-versions, we have to use finer results on the reference element which make the dependence of the estimate on the polynomial degree explicit.

The question of the smoothness of functions in Sobolev spaces is answered by the following results.

Theorem 1.8 (Sobolev embedding theorem)**.** *Assume that $\Omega \subseteq \mathbb{R}^d$ is open, bounded with Lipschitz boundary, and that $2(s - m) > d$.*

Then, we have $\mathrm{H}^s(\Omega) \subseteq \mathrm{C}^m(\overline{\Omega})$ with continuous embedding.

Theorem 1.9 (Gagliardo-Nirenberg-Sobolev inequality)**.** *Let $(a,b) \subseteq \mathbb{R}^1$ be a bounded interval.*

Then, there exists a constant $C > 0$ such that for all $u \in \mathrm{H}^1(a,b)$,

$$\|u\|_{\mathrm{L}^\infty(a,b)} \leqslant C\,\|u\|_{\mathrm{L}^2(a,b)}^{1/2}\,\|u\|_{\mathrm{H}^1(a,b)}^{1/2} . \tag{1.23}$$

Consider an open interval $(a, b) \subseteq \mathbb{R}^1$. Then, by Theorem 1.8, we see that $\mathrm{H}^s(a, b)$ consists of continuous functions if $s > 1/2$. One can show that $\mathrm{H}^{1/2}(a, b)$ contains functions which are not continuous. Another feature of $\mathrm{H}^{1/2}(a, b)$ functions in $\mathrm{H}^{1/2}(a, b)$ cannot be extended by zero to functions in $\mathrm{H}^{1/2}(\mathbb{R})$. We define therefore the spaces of functions $\mathrm{H}^{1/2}_{00}(a, b)$ which consists of those functions in $\mathrm{H}^{1/2}(a, b)$ which can be extended by zero to functions in $\mathrm{H}^{1/2}(\mathbb{R})$, $\mathrm{H}^{1/2}_{(0}(a, b)$ which consists of functions which can be extended by zero to functions in $\mathrm{H}^{1/2}(-\infty, b)$, and $\mathrm{H}^{1/2}_{0)}(a, b)$ similarly with $\mathrm{H}^{1/2}(a, \infty)$. On these spaces, we define seminorms by

$$|v|_{\mathrm{H}^{1/2}_{00}(a,b)} := \left(|v|^2_{\mathrm{H}^{1/2}(a,b)} + \int_a^b \frac{|v(x)|^2}{\operatorname{dist}(x, \{a, b\})} \mathrm{d}x \right)^{1/2}, \tag{1.24}$$

$$|v|_{\mathrm{H}^{1/2}_{(0}(a,b)} := \left(|v|^2_{\mathrm{H}^{1/2}(a,b)} + \int_a^b \frac{|v(x)|^2}{x - a} \mathrm{d}x \right)^{1/2}, \quad \text{and} \tag{1.25}$$

$$|v|_{\mathrm{H}^{1/2}_{0)}(a,b)} := \left(|v|^2_{\mathrm{H}^{1/2}(a,b)} + \int_a^b \frac{|v(x)|^2}{b - x} \mathrm{d}x \right)^{1/2}, \tag{1.26}$$

and norms by

$$\|v\|_{\mathrm{H}^{1/2}_{00}(a,b)} := \left(\|v\|^2_{\mathrm{L}^2(a,b)} + |v|^2_{\mathrm{H}^{1/2}_{00}(a,b)} \right)^{1/2}, \tag{1.27}$$

$$\|v\|_{\mathrm{H}^{1/2}_{(0}(a,b)} := \left(\|v\|^2_{\mathrm{L}^2(a,b)} + |v|^2_{\mathrm{H}^{1/2}_{(0}(a,b)} \right)^{1/2}, \quad \text{and} \tag{1.28}$$

$$\|v\|_{\mathrm{H}^{1/2}_{0)}(a,b)} := \left(\|v\|^2_{\mathrm{L}^2(a,b)} + |v|^2_{\mathrm{H}^{1/2}_{0)}(a,b)} \right)^{1/2}. \tag{1.29}$$

Note that, actually, the seminorms on these spaces are already equivalent to the full norms, as the weighted L^2-norms are upper bounds for the standard L^2-norms.

We define *negative order Sobolev spaces* as dual spaces of Sobolev spaces with positive order. We set $\tilde{\mathrm{H}}^{-s}(\Omega) := \mathrm{H}^s(\Omega)^*$ and $\mathrm{H}^{-1}(\Omega) := \mathrm{H}^1_0(\Omega)^*$, and correspondingly for manifolds. Further spaces will be defined by interpolation.

Finally, we note that vector-valued spaces can always be defined using product spaces, as for a space V of functions $\Omega \to M$, we see that the product space V^m can be interpreted as a space of functions $\Omega \to M^m$.

1.2.1 Interpolation Spaces

Above, we defined integer and fractional order Sobolev spaces. In practice, it is typically easier to show certain estimates for integer order spaces. It is thus interesting to check whether it is possible to generalise such results to the fractional order case. A very general approach is given by the theory of interpolation spaces. Here, we define, based on two Banach spaces A_0 and A_1, some kind of "intermediate" spaces $(A_0, A_1)_{\theta,q}$ with the property that all operators T which are defined on both A_0 and A_1 and coincide on $A_0 \cap A_1$ can be extended to operators on the intermediate spaces, and admit bounds on the norms based on the norms on A_0 and A_1 and the parameters q and θ. Further details on interpolation theory in general and its applications to Sobolev spaces in particular can be found in [BS08, BL76, Tri95, Sch98].

Definition 1.10. Let $A_1 \subseteq A_0$ be two Banach spaces where the embedding is continuous, $0 < \theta < 1$ and $1 \leqslant q \leqslant \infty$, and define the *K-functional* by

$$K(t, v) := \inf_{w \in A_1} \left[\|v - w\|_{A_0} + t \|w\|_{A_1} \right]. \tag{1.30}$$

Then, we define the *interpolation space* $(A_0, A_1)_{\theta,q}$ by

$$A_{\theta,q} := (A_0, A_1)_{\theta,q} := \left\{ v \in A_0 \colon \|v\|_{A_{\theta,q}} < \infty \right\}, \tag{1.31}$$

where

$$\|v\|_{A_{\theta,q}} := \left(\int_0^\infty \left[t^{-\theta} K(t, v) \right]^q \frac{\mathrm{d}t}{t} \right)^{1/q} \quad \text{for } q \in [1, \infty), \tag{1.32}$$

$$\|v\|_{A_{\theta,\infty}} := \sup_{0<t<\infty} \left[t^{-\theta} K(t, v) \right]. \tag{1.33}$$

The most important properties are collected in the following results.

Theorem 1.11. *1. For $0 < \theta_2 \leqslant \theta_1 < 1$, $1 \leqslant q_1 \leqslant q_2 \leqslant \infty$,*

$$A_1 \subseteq A_{\theta_1,q_1} \subseteq A_{\theta_2,q_2} \subseteq A_0. \tag{1.34}$$

2. If $A_0 = A_1$, then for all $0 < \theta < 1$ and $1 \leqslant q \leqslant \infty$,

$$A_0 = A_{\theta,q} = A_1. \tag{1.35}$$

3. For $v \in A_1$, $0 < \theta < 1$ and $1 \leqslant q \leqslant \infty$,

$$\|v\|_{A_{\theta,q}} \leqslant C(\theta, q) \|v\|_{A_0}^{1-\theta} \|v\|_{A_1}^{\theta}. \tag{1.36}$$

Theorem 1.12 (Interpolation of operators)**.** *Let A_i, B_i be two pairs of Banach spaces as above, and assume that $T_i \colon A_i \to B_i$ are continuous and linear, $i = 0, 1$, with $T_0|_{A_1} = T_1$.*

Then, the operator $T_{\theta,q} \colon A_{\theta,q} \to B_{\theta,q}$ is well-defined and continuous for every θ and q, coincides with T_1 on A_1, and satisfies

$$\|Tv\|_{B_{\theta,q}} \leqslant \|Tv\|_{B_0}^{1-\theta} \|Tv\|_{B_1}^{\theta}. \tag{1.37}$$

Theorem 1.13 (Reiteration theorem)**.** *For $0 < \theta_0 < \theta_1 < 1$, $1 \leqslant q_0, q_1, q \leqslant \infty$, $0 < \theta < 1$, we have that*

$$\left((A_0, A_1)_{\theta_0,q_0}, (A_0, A_1)_{\theta_1,q_1} \right)_{\theta,q} = (A_0, A_1)_{(1-\theta)\theta_0+\theta\theta_1,q}. \tag{1.38}$$

Theorem 1.14 (Dual spaces)**.** *For A_1 dense in A_0, $0 < \theta < 1$ and $1/p + 1/q = 1$,*

$$(A_0, A_1)_{\theta,q}^* = (A_1^*, A_0^*)_{1-\theta,p}. \tag{1.39}$$

The fundamental theorem which allows us to apply the above theory to Sobolev spaces is:

Theorem 1.15. *Let $s = k + \theta$, where $k \in \mathbb{N}_0$ and $0 < \theta < 1$, and assume that $\Omega \subseteq \mathbb{R}^n$ is a bounded domain with Lipschitz boundary.*

Then,

$$\mathrm{H}^s(\Omega) = \left(\mathrm{H}^k(\Omega), \mathrm{H}^{k+1}(\Omega) \right)_{\theta,2}, \tag{1.40}$$

and the interpolation norm is equivalent to the Slobodečkij norm given above.

We want to see how the spaces satisfying weak boundary conditions fit into this interpolation framework. This is answered by the following result.

Theorem 1.16. *Let $(a, b) \subseteq \mathbb{R}$ be a bounded interval.*
Then,

$$\mathrm{H}^{1/2}_{00}(a, b) = \big(\mathrm{L}^2(a, b), \mathrm{H}^1_0(a, b)\big)_{1/2,2}\,, \tag{1.41}$$

and the interpolation norm is equivalent to the $\mathrm{H}^{1/2}_{00}$-norm.

Similarly, if we let $\mathrm{H}^1_{(0}(a, b)$ and $\mathrm{H}^1_{0)}(a, b)$ be the spaces of functions in $\mathrm{H}^1(a, b)$ vanishing at a or b, respectively, then

$$\mathrm{H}^{1/2}_{(0}(a, b) = \Big(\mathrm{L}^2(a, b), \mathrm{H}^1_{(0}(a, b)\Big)_{1/2,2}\,, \tag{1.42}$$

$$\mathrm{H}^{1/2}_{0)}(a, b) = \Big(\mathrm{L}^2(a, b), \mathrm{H}^1_{0)}(a, b)\Big)_{1/2,2}\,, \tag{1.43}$$

where again the interpolation norms are equivalent to the natural norms of the respective space.

The proof for $\mathrm{H}^{1/2}_{00}(a, b)$ is given in [LM72, Chapter 1, Theorem 1.7]. The result for $\mathrm{H}^{1/2}_{(0}(a, b)$ follows by using the reflection operator R mapping functions on the interval (a, b) to functions on the interval $(a, 2b-a)$, that is, for $f\colon (a, b) \to \mathbb{R}$, we define $Rf\colon (a, 2b-a) \to \mathbb{R}$ by

$$Rf(x) := \begin{cases} f(x), & x \leqslant b, \\ f(2b - x), & x > b, \end{cases} \tag{1.44}$$

and its left inverse, the restriction operator S, which is defined by $Sf(x) := f(x)$ for $x \in (a, b)$, where $f\colon (a, 2b - a) \to \mathbb{R}$. The result for $\mathrm{H}^{1/2}_{0)}(a, b)$ is proved analogously.

Define $\mathrm{H}^s_0(\Omega) := \big(\mathrm{L}^2(\Omega), \mathrm{H}^1_0(\Omega)\big)_{s,2}$ for $s \in (0, 1)\backslash\{1/2\}$. It can be shown that the interpolation norm of $\mathrm{H}^s_0(\Omega)$ is equivalent to the norm of $\mathrm{H}^s(\Omega)$, and the space equals the closure of $\mathrm{D}(\Omega)$ in this norm. For $s = 1/2$, we see by the last theorem that this result obviously cannot hold. Furthermore, for $s < 1/2$, $\mathrm{H}^s_0(\Omega) = \mathrm{H}^s(\Omega)$.

Using the duality theorem, we set $\mathrm{H}^{-s}(\Omega) := \big(\mathrm{H}^{-1}(\Omega), \mathrm{L}^2(\Omega)\big)_{s,2}$. We note that $\mathrm{H}^{-1/2}(\Gamma_{\mathrm{C}})$ is thus the dual space of $\mathrm{H}^{1/2}_{00}(\Gamma_{\mathrm{C}})$, where Γ_{C} is a piece of $\Gamma = \partial\Omega$.

1.2.2 Inverse Inequalities

In general, in the finite element method, we have two kinds of inequalities: First *direct inequalities*, which give approximation rates for sufficiently regular functions, and second *inverse inequalities*, which yield an estimate of a stronger norm by a weaker norm. Clearly, such a statement is only possible on finite-dimensional spaces. In particular, we shall focus here on spaces of polynomials. Let, thus, $\mathcal{P}^q$ be the vector space of polynomials of degree q.

The following results can be found in [Sch98, Sections 3.6, 4.6].

Theorem 1.17. *There exists a constant $C > 0$ such that for all $p \in \mathbb{N}$ and all $v \in \mathcal{P}^p$, we have that*

$$\|v\|_{\mathrm{L}^\infty(-1,1)} \leqslant C(p + 1)\,\|v\|_{\mathrm{L}^2(-1,1)}\,. \tag{1.45}$$

Theorem 1.18. *For all $p \in \mathbb{N}$ and all $v \in \mathcal{P}^p$ with $v(-1) = v(1) = 0$, we have that*

$$\left\|(1 - \cdot^2)^{1/2} v'\right\|_{\mathrm{L}^2(-1,1)} \leqslant (p+1) \|v\|_{\mathrm{L}^2(-1,1)} . \tag{1.46}$$

Theorem 1.19. *There exists a constant $C > 0$ such that for all $p \in \mathbb{N}$ and all $v \in \mathcal{P}^p$ with $v(-1) = v(1) = 0$, we have that*

$$\|v\|_{\mathrm{H}^{1/2}_{00}(-1,1)} \leqslant C \ln(p+1) \|v\|_{\mathrm{H}^{1/2}(-1,1)} . \tag{1.47}$$

The next result follows from [BDM07, Proposition 4.1].

Theorem 1.20. *There exists a constant $C > 0$ such that for all $p \in \mathbb{N}$ and all $v \in \mathcal{P}^p$ with $v(-1) = v(1) = 0$, we have that*

$$\|v\|_{\mathrm{H}^{1/2}(-1,1)} \leqslant C(p+1) \|v\|_{\mathrm{L}^2(-1,1)} . \tag{1.48}$$

For the reference interval $I = [-1, 1]$, we define the *edge bubble function* as $\psi_I(x) := \mathrm{dist}(x, \partial I)$, and similarly, for the reference square $S := I^2$, we define the *element bubble function* as $\psi_S(x) := \mathrm{dist}(x, \partial S)$. Using the edge bubble function, we can formulate the following result, which is similar to Theorem 1.18.

Theorem 1.21. *Let $-1 < \alpha < \beta$ and $\delta \in [0, 1]$.*
Then, there exists a constant $C > 0$ such that for all $p \in \mathbb{N}$ and all polynomials $v \in \mathcal{P}^p$,

$$\left\|\psi_I^{1/2} v'\right\|_{\mathrm{L}^2(-1,1)} \leqslant Cp \|v\|_{\mathrm{L}^2(-1,1)} , \tag{1.49}$$

$$\left\|\psi_I^{\alpha/2} v\right\|_{\mathrm{L}^2(-1,1)} \leqslant Cp^{\beta-\alpha} \left\|\psi_I^{\beta/2} v\right\|_{\mathrm{L}^2(-1,1)} , \tag{1.50}$$

$$\left\|\psi_I^{\delta} v'\right\|_{\mathrm{L}^2(-1,1)} \leqslant Cp^{2-\delta} \left\|\psi_I^{\delta/2} v\right\|_{\mathrm{L}^2(-1,1)} . \tag{1.51}$$

If, furthermore, $v(-1) = v(1) = 0$, then

$$\|v'\|_{\mathrm{L}^2(-1,1)} \leqslant Cp \left\|\psi_I^{-1/2} v\right\|_{\mathrm{L}^2(-1,1)} . \tag{1.52}$$

The following inverse inequalities are the two-dimensional analogues.

Theorem 1.22. *For $-1 < \alpha < \beta$ and $\delta \in [0, 1]$, there exists a constant $C > 0$ such that for all $p \in \mathbb{N}$ and all polynomials $v \in \mathcal{P}^p$,*

$$\left\|\psi_S^{1/2} \nabla v\right\|_{\mathrm{L}^2((-1,1)^2)} \leqslant Cp \|v\|_{\mathrm{L}^2((-1,1)^2)} , \tag{1.53}$$

$$\left\|\psi_S^{\alpha/2} v\right\|_{\mathrm{L}^2((-1,1)^2)} \leqslant Cp^{\beta-\alpha} \left\|\psi_S^{\beta/2} v\right\|_{\mathrm{L}^2((-1,1)^2)} , \tag{1.54}$$

$$\left\|\psi_S^{\delta} \nabla v\right\|_{\mathrm{L}^2((-1,1)^2)} \leqslant Cp^{2-\delta} \left\|\psi_S^{\delta/2} v\right\|_{\mathrm{L}^2((-1,1)^2)} . \tag{1.55}$$

If, furthermore, $v = 0$ on ∂S, then

$$\|\nabla v\|_{\mathrm{L}^2((-1,1)^2)} \leqslant Cp \left\|\psi_S^{-1/2} v\right\|_{\mathrm{L}^2((-1,1)^2)} . \tag{1.56}$$

The last two theorems can be found in [MW01, Lemma 2.4, Theorem 2.5].

1.3 Functional Analysis, Variational Inequalities and Duality Theory

We collect some fundamental results from functional analysis, variational inequalities and duality theory. Standard references are [KS80, Kor97, Zei85].

Let V be a Hilbert space, $\mathcal{K} \subseteq V$ convex and closed, $a\colon V \times V \to \mathbb{R}$ a symmetric, continuous bilinear form with

$$a(v,v) \geqslant 0 \quad \text{for all } v \in V, \tag{1.57}$$

$L\colon V \to \mathbb{R}$ a continuous linear functional, and $j\colon \mathcal{K} \to \mathbb{R}$ a continuous, convex, but possibly nonlinear functional. We define the *energy functional* $J\colon \mathcal{K} \to \mathbb{R}$ by

$$J(v) := \frac{1}{2}a(v,v) - L(v) + j(v) \quad \text{for all } v \in V. \tag{1.58}$$

We then have:

Theorem 1.23. *u is a minimiser of J over $\mathcal{K}$ if and only if*

$$a(u, v-u) + j(v) - j(u) \geqslant L(v-u) \quad \textit{for all } v \in \mathcal{K}. \tag{1.59}$$

Proof. First, assume that $\mathbf{u}$ minimises J. For $\mathbf{v} \in \mathcal{K}$ and $t \in (0,1)$, we see that $\mathbf{u} + t(\mathbf{v}-\mathbf{u}) = (1-t)\mathbf{u} + t\mathbf{v} \in \mathcal{K}$ by the convexity of $\mathcal{K}$, and thus

$$J(\mathbf{u} + t(\mathbf{v}-\mathbf{u})) - J(\mathbf{u}) \geqslant 0, \tag{1.60}$$

that is,

$$\begin{aligned} 0 &\leqslant \frac{1}{2}a(\mathbf{u}+t(\mathbf{v}-\mathbf{u}), \mathbf{u}+t(\mathbf{v}-\mathbf{u})) - L(\mathbf{u}+t(\mathbf{v}-\mathbf{u})) + j(\mathbf{u}+t(\mathbf{v}-\mathbf{u})) \\ &\quad - \frac{1}{2}a(\mathbf{u},\mathbf{u}) + L(\mathbf{u}) - j(\mathbf{u}) \\ &= ta(\mathbf{u}, \mathbf{v}-\mathbf{u}) + t^2 a(\mathbf{v}-\mathbf{u}, \mathbf{v}-\mathbf{u}) - tL(\mathbf{v}-\mathbf{u}) + j(\mathbf{u}+t(\mathbf{v}-\mathbf{u})) - j(\mathbf{u}). \end{aligned} \tag{1.61}$$

Applying the convexity of j, we obtain that

$$j((1-t)\mathbf{u} + t\mathbf{v}) \leqslant (1-t)j(\mathbf{u}) + tj(\mathbf{v}), \tag{1.62}$$

which yields

$$0 \leqslant ta(\mathbf{u}, \mathbf{v}-\mathbf{u}) + t^2 a(\mathbf{v}-\mathbf{u}, \mathbf{v}-\mathbf{u}) - tL(\mathbf{v}-\mathbf{u}) + tj(\mathbf{v}) - tj(\mathbf{u}). \tag{1.63}$$

Dividing by t and letting $t \to 0$, we see that $\mathbf{u}$ satisfies the variational inequality.

Conversely, for a solution $\mathbf{u}$ of the variational inequality, we see that for all $\mathbf{v} \in \mathcal{K}$,

$$\begin{aligned} J(\mathbf{v}) - J(\mathbf{u}) &= \frac{1}{2}a(\mathbf{v},\mathbf{v}) - L(\mathbf{v}) + j(\mathbf{v}) - \frac{1}{2}a(\mathbf{u},\mathbf{u}) + L(\mathbf{u}) - j(\mathbf{u}) \\ &= \frac{1}{2}a(\mathbf{v}, \mathbf{v}-\mathbf{u}) + \frac{1}{2}a(\mathbf{u}, \mathbf{v}-\mathbf{u}) - L(\mathbf{v}-\mathbf{u}) + j(\mathbf{v}) - j(\mathbf{u}) \\ &= \frac{1}{2}a(\mathbf{v}-\mathbf{u}, \mathbf{v}-\mathbf{u}) + a(\mathbf{u}, \mathbf{v}-\mathbf{u}) - L(\mathbf{v}-\mathbf{u}) + j(\mathbf{v}) - j(\mathbf{u}) \\ &\geqslant 0, \end{aligned} \tag{1.64}$$

that is, $\mathbf{u}$ minimises J. $\square$

To analyse the solvability of the above problems, let $J\colon \mathcal{K} \to \mathbb{R}$ be a general convex functional which is continuous and *coercive*, that is,

$$\lim_{\substack{\|v\|_V \to \infty \\ v \in \mathcal{K}}} J(v) = \infty. \tag{1.65}$$

The next result follows from [Zei85, Proposition 41.8].

Lemma 1.24. *Let $J\colon \mathcal{K} \to \mathbb{R}$ be continuous and convex.*

Then, J is weakly sequentially lower semicontinuous*, that is, $J(v) \leqslant \liminf_{n\to\infty} J(v_n)$ whenever $v_n \rightharpoonup v$.*

Here, as usual, $v_n \rightharpoonup v$ denotes convergence in the weak topology, that is, we say that $v_n \rightharpoonup v$ if and only if $\lim_{n\to\infty} \langle v_n, w\rangle_V = \langle v, w\rangle_V$ for all $w \in V$; similarly, we write $v_n \to v$ for strong convergence, that is, $v_n \to v$ if and only if $\lim_{n\to\infty} \|v - v_n\|_V = 0$. With this, we obtain:

Theorem 1.25. *Let $\mathcal{K} \subseteq V$ be a closed, convex set and $J\colon \mathcal{K} \to \mathbb{R}$ be continuous, convex and coercive.*

Then, there exists $u \in \mathcal{K}$ such that

$$J(u) \leqslant J(v) \quad \textit{for all } v \in \mathcal{K}. \tag{1.66}$$

If, furthermore, J is strictly convex*, that is,*

$$J(tu + (1-t)v) < tJ(u) + (1-t)J(v) \quad \textit{for } u, v \in \mathcal{K} \textit{ with } u \neq v \textit{ and } t \in (0,1), \tag{1.67}$$

then the minimiser u is unique.

We shall need in the proof the following compactness result, which is proved in [Yos80, p. 126, Theorem 1].

Lemma 1.26. *Let V be a Hilbert space, and let (v_n) be a sequence in V which is bounded in norm.*

Then, there exists a subsequence $(v_{n'})$ of (v_N) converging weakly to some $v \in V$.

Proof of Theorem 1.25. Set $\alpha := \inf_{v\in\mathcal{K}} J(v) \in \bar{\mathbb{R}}$, and choose a sequence $(u_n)_{n\in\mathbb{N}} \subseteq \mathcal{K}$ such that $\alpha = \lim_{n\to\infty} J(u_n)$. By the coercivity of J, we see that (u_n) is necessarily bounded, and thus, by Lemma 1.26, admits a weakly convergent subsequence $(u_{n'})$. Denote this limit by u.

By Lemma 1.24, we see that J is weakly sequentially lower semicontinuous, and thus

$$\alpha \leqslant J(u) \leqslant \liminf_{n'\to\infty} J(u_{n'}) = \lim_{n'\to\infty} J(u_{n'}) = \alpha, \tag{1.68}$$

that is, $\alpha \in \mathbb{R}$, $J(u) = \alpha$, and u is a minimiser of J.

For the second part, assume that J is strictly convex, and let $u_1 \neq u_2$ be two minimisers. Then, letting again α denote the minimum and noting that $tu_1 + (1-t)u_2 \in \mathcal{K}$ as $\mathcal{K}$ is convex, for $t \in (0,1)$,

$$\alpha \leqslant J(tu_1 + (1-t)u_2) < tJ(u_1) + (1-t)J(u_2) = \alpha, \tag{1.69}$$

a contradiction. Thus, $u_1 = u_2$, and the minimiser is unique. □

To derive an error indicator, we shall apply duality theory. We collect here the definition of the conjugate function, and the basic theorem on solvability of the primal and the dual problem and their connection.

Definition 1.27. Let V be a Hilbert space, and $f\colon V \to \bar{\mathbb{R}}$.

Then, the *conjugate function* $f^*\colon V^* \to \bar{\mathbb{R}}$ of f is defined by

$$f^*(v^*) := \sup_{v\in V} [\langle v^*, v\rangle_{V^*} - f(v)] \quad \text{for all } v^* \in V^*. \tag{1.70}$$

The next result is a consequence of [Han05, Theorem 2.39].

Theorem 1.28. *Let V, Z be Hilbert spaces, $\mathcal{L}\colon V \to Z$ linear and bounded with $\mathcal{L}\colon Z^* \to V^*$ its adjoint operator, $F\colon V \to \bar{\mathbb{R}}$, $G\colon Z \to \bar{\mathbb{R}}$ lower semicontinuous, convex functions such that there exists $v_0 \in V$ with $F(v_0) < \infty$, $G(\mathcal{L}v_0) < \infty$, and $q \mapsto G(q)$ is continuous at $\mathcal{L}v_0$, and $v \mapsto F(v) + G(\mathcal{L}v)$ is coercive on V.*

Denoting $J(v, q) := F(v) + G(q)$, the conjugate function of J is given by $J^*(v^*, q^*) = F^*(v^*) + G^*(q^*)$. Furthermore, there exist $u \in V$ and $p^* \in Z^*$ with

$$J(u, \mathcal{L}u) = \inf_{v\in V} J(v, \mathcal{L}v), \tag{1.71}$$

$$-J^*(\mathcal{L}^*p^*, -p^*) = \sup_{q^*\in Z^*} [-J^*(\mathcal{L}^*q^*, -q^*)], \tag{1.72}$$

and

$$J(u, \mathcal{L}u) = -J^*(\mathcal{L}^*p^*, -p^*). \tag{1.73}$$

Moreover, if $v \mapsto J(v, \mathcal{L}v)$ is strictly convex, then the minimiser u is unique.

1.4 Measure Theory

We shall make use of the following version of the Riesz representation theorem, which is proved in [Yos80, p. 115, Example 3].

Theorem 1.29 (Riesz representation theorem)**.** *Let (X, μ) be a σ-finite measure space.*

Then, for every continuous linear functional $\ell\colon \mathrm{L}^1(X) \to \mathbb{R}$, there exists a function $f \in \mathrm{L}^\infty(X)$ such that

$$\ell(g) = \int_X fg\mathrm{d}\mu \quad \textit{for all } g \in \mathrm{L}^1(X) \tag{1.74}$$

and

$$\sup_{\substack{g\in \mathrm{L}^1(X) \\ \|g\|_{\mathrm{L}^1(X)}=1}} |\ell(g)| = \|f\|_{\mathrm{L}^\infty(X)}. \tag{1.75}$$

The following result is given in [Yos80, p. 53, Corollary to Proposition 2].

Theorem 1.30. *Let (X, μ) be a σ-finite measure space.*

Then, for every Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathrm{L}^p(X)$, there exists a subsequence $(f_{n'})$ of (f_n) which converges almost everywhere on X.

Theorem 1.31 (Dominated convergence theorem)**.** *Let (X, μ) be a measure space, $(f_n)_{n\in\mathbb{N}}$ a sequence of measurable functions converging almost everywhere on X to f, and assume that there exists a function $g \in \mathrm{L}^1(X)$ such that $|f_n(x)| \leqslant g(x)$ for almost every $x \in X$.*

Then, $f \in \mathrm{L}^1(X)$, $\lim_{n\to\infty} \int_X |f_n - f| \,\mathrm{d}\mu = 0$, and $\lim_{n\to\infty} \int_X f_n\mathrm{d}\mu = \int_X f\mathrm{d}\mu$.

The proof can be found in [Rud87, Theorem 1.34].

Chapter 2

An Introduction to Elastic Contact with Friction

In the present chapter, we want to give a short introduction to the mathematical formulation of elastic contact problems with friction. We shall briefly describe the elements of small-strain elasticity, and then focus on the effects of friction. For a more detailed account on elasticity, we refer the reader to the author's work [Dör07] and the references cited therein. A standard reference for contact problems, with and without friction, is [KO88], and the formulation as a variational inequality is detailed in [DL76]. The model given here corresponds to the friction model used in [Han05, Example 1.27]. A good, short introduction to Coulomb friction is given in [Sin06].

2.1 Small-Strain Elasticity

In this section, we introduce the objects necessary to work with problems of elasticity.

2.1.1 The Basic Equations

Consider a *body*, that is, a domain $\Omega \subseteq \mathbb{R}^3$. A *body force* $\mathbf{F} = (F_1, F_2, F_3)$ is an $\mathbb{R}^3$-valued function defined on Ω, a *force density*. The *stress vector* $\mathbf{T}(\mathbf{x}, \mathbf{z})$ is defined as the density of internal forces at $\mathbf{x}$ in the direction $\mathbf{z}$. Using $\mathbf{T}$, we can define the *stress tensor* $\boldsymbol{\sigma} = (\sigma_{ij})$ as

$$\sigma_{ij}(\mathbf{x}) := T_j(\mathbf{x}, \mathbf{e}_i) \quad \text{for } i, j = 1, 2, 3 \text{ and } \mathbf{x} \in \Omega. \tag{2.1}$$

For a body in equilibrium, the internal forces $\boldsymbol{\sigma}$ and the external forces $\mathbf{F}$ have to balance, and from this, one can show that the *equations of equilibrium*

$$\sigma_{ji,j}(\mathbf{x}) + F_i(\mathbf{x}) = 0 \quad \text{for } i = 1, 2, 3 \text{ and } \mathbf{x} \in \Omega \tag{2.2}$$

hold true. In the dynamic case, one has to add inertial terms, that is, $\rho\ddot{\mathbf{u}}$ with ρ the density of the material, $\mathbf{u}$ the displacement and $\ddot{\mathbf{u}} = \partial^2\mathbf{u}/\partial t^2$ the acceleration, in the above equation.

Additionally, from the equilibrium of moments, we obtain the symmetry of the stress tensor,

$$\sigma_{ij} = \sigma_{ji} \quad \text{for } i, j = 1, 2, 3. \tag{2.3}$$

Let Ω be deformed into another body Ω', and assume that this deformation is realised by a diffeomorphism $\mathbf{y}\colon \Omega \to \Omega'$, that is, $\mathbf{y}$ is one-to-one, onto, and $\mathbf{y}$ and its inverse are differentiable. Then, we define the *displacement vector* by $\mathbf{u}(\mathbf{x}) := \mathbf{y}(\mathbf{x}) - \mathbf{x}$. Comparing the lengths of the line segment from $\mathbf{x}$ to $\mathbf{x} + t\Delta\mathbf{x}$ with the line segment after the deformation, $\mathbf{y}(\mathbf{x})$ to $\mathbf{y}(\mathbf{x} + t\Delta\mathbf{x})$, we see that

$$\varphi(t) := |\mathbf{y}(\mathbf{x} + t\Delta\mathbf{x}) - \mathbf{y}(\mathbf{x})|^2 - |t\Delta\mathbf{x}|^2$$

$$
\begin{aligned}
&= |\mathbf{u}(\mathbf{x}+t\Delta\mathbf{x}) - \mathbf{u}(\mathbf{x}) + t\Delta x|^2 - t^2 |\Delta\mathbf{x}|^2 \\
&= \sum_{i=1}^{3} \Big[(u_i(\mathbf{x}+t\Delta\mathbf{x}) - u_i(\mathbf{x}))^2 + 2t\,(u_i(\mathbf{x}) + t\Delta\mathbf{x}) - u_i(\mathbf{x}))\,\Delta x_i \Big] \qquad (2.4) \\
&= t^2 \left[\sum_{i=1}^{3} \left(\int_0^1 u_{i,j}(\mathbf{x}+t\tau\Delta\mathbf{x})\Delta x_j \mathrm{d}\tau \right)^2 + 2\int_0^1 u_{i,j}(\mathbf{x}+t\tau\Delta\mathbf{x})\Delta x_i \Delta x_j \mathrm{d}\tau \right].
\end{aligned}
$$

Then, $\frac{1}{2}\varphi''(0) = 2\varepsilon_{ij}^{\text{Finite}} \Delta x_i \Delta x_j$, where $(\varepsilon_{ij}^{\text{Finite}})_{i,j=1,2,3}$ is the *finite strain tensor* defined by

$$\varepsilon_{ij}^{\text{Finite}} := \frac{1}{2}\left(u_{k,i}u_{k,j} + u_{i,j} + u_{j,i}\right) \quad \text{for } i,j = 1,2,3. \tag{2.5}$$

Assuming that $u_{i,j}$ is small, we see that the term $u_{k,i}u_{k,j}$ is of higher order and can be neglected, which gives the *small strain tensor*

$$\varepsilon_{ij} := \frac{1}{2}\left(u_{i,j} + u_{j,i}\right) \quad \text{for } i,j = 1,2,3. \tag{2.6}$$

We want to relate $\boldsymbol{\varepsilon} := (\varepsilon_{ij})$ and $\boldsymbol{\sigma}$; this is done by a *material law*. As we restrict ourselves to small deformations, we can assume that the relation between $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ is linear, from which it follows that there exists a 4-tensor $\mathbf{C} := (c_{ijkl})$ with $\sigma_{ij} = c_{ijkl}\varepsilon_{kl}$ for i, j, k, $l = 1, 2, 3$; in short $\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon}$. From the symmetry of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ and an assumption of *hyperelasticity*, it follows that

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{klij} \quad \text{for } i,j,k,l = 1,2,3, \tag{2.7}$$

and furthermore, we suppose that $\mathbf{C}$ is positive definite, that is, $\boldsymbol{\varepsilon} : \mathbf{C}\boldsymbol{\varepsilon} > 0$ for all strains $\boldsymbol{\varepsilon}$. For a general material, $\mathbf{C}$ depends on the point $\mathbf{x}$ and on the choice of the coordinate system. For simplicity, we shall assume a *homogeneous* and *isotropic material*, that is, $\mathbf{C}$ neither depends on the point or on the choice of (Cartesian) coordinate system. These assumptions yield the existence of *Lamé coefficients* λ, $\mu \in \mathbb{R}$ such that the *generalised Hooke's law*

$$\sigma_{ij}(\mathbf{x}) = \lambda \operatorname{tr} \boldsymbol{\varepsilon}(\mathbf{x})\delta_{ij} + 2\mu\varepsilon_{ij}(\mathbf{x}) \quad \text{for } i,j = 1,2,3 \tag{2.8}$$

holds true. Typically, for physical materials, the Lamé coefficients are not directly given, but instead the *Young modulus E* and the *Poisson ratio ν*. From these, the Lamé coefficients can be calculated as

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{1+\nu}. \tag{2.9}$$

2.1.2 Boundary Conditions

We describe the different kinds of boundary conditions considered in the following. Decompose the boundary Γ of Ω into three disjoint, relatively closed subsets Γ_{D}, Γ_{N} and Γ_{C} which are the closures of their interiors such that the respective interiors have empty intersection, and $\Gamma = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{C}}$. Then, we prescribe *kinematic boundary conditions* on Γ_{D}, that is, the displacement is given, $u_j = u_{0j}$, $j = 1,2,3$, which corresponds to Dirichlet boundary conditions. To guarantee the unique solvability of the problems which will be formulated below, we shall always assume that $|\Gamma_{\mathrm{D}}| > 0$, as then, the Korn inequality as given in Theorem 2.8 holds, and the fundamental bilinear form describing the inner energy of an elastic body is coercive.

On Γ_N, we prescribe *static boundary conditions*, that is, the stress vector is given, $\sigma_{ij}\nu_j = G_i$, $i = 1, 2, 3$, where $\boldsymbol{\nu} = (\nu_j)$ is the outer unit normal vector on Γ_N and $\mathbf{G} = (G_i)$ is a given function. Lastly, on Γ_C, we assume *contact conditions with friction*. These will be the topic of Section 2.2.

2.1.3 Plane Problems in Elasticity

There are two typical ways of reducing problems in elasticity to two dimensions. The first is the *plane strain assumption*, where the body is considered to be infinitely large in one dimension, and the strain $\boldsymbol{\varepsilon}$ is assumed to be planar, that is, $\varepsilon_{i3} = 0$ for $i = 1, 2, 3$. The second is the *plane stress assumption* which we will describe now.

Let the 3-dimensional body be given as $\Omega \times [-h, h]$ with a domain $\Omega \subseteq \mathbb{R}^2$, and assume that the boundary conditions and volume forces do not depend on x_3. Furthermore, assume that h is small, u_{03} is an odd function of x_3, that is, $u_{03}(-x_3) = -u_{03}(x_3)$, and that $\sigma_{3i} = 0$ at $x_3 = \pm h$ for $i = 1, 2, 3$. Then, the assumption that $\sigma_{3i} = 0$, $i = 1, 2, 3$, on $\Omega \times [-h, h]$ is justified, see [LL70, pp. 53], and the stress tensor can be described by a 2×2 matrix. Thus, as $\sigma_{33} = 0$,

$$0 = \lambda\left(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}\right) + 2\mu\varepsilon_{33}, \tag{2.10}$$

which yields

$$\varepsilon_{33} = \frac{-\lambda}{\lambda + 2\mu}\left(\varepsilon_{11} + \varepsilon_{22}\right), \tag{2.11}$$

and the generalised Hooke's law reads

$$\sigma_{ij} = \lambda^*\left(\varepsilon_{11} + \varepsilon_{22}\right)\delta_{ij} + 2\mu\varepsilon_{ij}, \quad i, j = 1, 2, \tag{2.12}$$

where $\lambda^* := \lambda\frac{2\mu}{\lambda+2\mu}$, that is, we have to solve a problem analogous to the three-dimensional system, but λ is replaced by λ^*.

2.2 Contact with Friction

In this section, we first describe the full three-dimensional setup for contact with friction, and then explain how the system can be reduced under the plane stress assumption.

2.2.1 The 3-Dimensional Situation

We shall now define the Signorini contact conditions with friction on Γ_C. Set for $\mathbf{x} \in \Gamma_\mathrm{C}$ and $\mathbf{v}$ sufficiently regular on Ω

$$v_n(\mathbf{x}) := v_i(\mathbf{x})\nu_i(\mathbf{x}), \tag{2.13}$$

$$v_{tj}(\mathbf{x}) := v_j(\mathbf{x}) - v_n(\mathbf{x})\nu_j(\mathbf{x}), \tag{2.14}$$

$$T_n(\mathbf{v})(\mathbf{x}) := \sigma_{ij}(\mathbf{v})\nu_i(\mathbf{x})\nu_j(\mathbf{x}), \tag{2.15}$$

$$T_{tj}(\mathbf{v})(\mathbf{x}) := \sigma_{jk}(\mathbf{v})\nu_k(\mathbf{x}) - T_n(\mathbf{v})(\mathbf{x})\nu_j(\mathbf{x}), \quad j = 1, 2, 3. \tag{2.16}$$

Assume that the body with which contact is possible and the *coefficient of friction* are given by functions $\mathbf{u}_0 \in \mathrm{H}^{1/2}(\Gamma_\mathrm{C})$ and $f \in \mathrm{L}^\infty(\Gamma_\mathrm{C})$ on Γ_C, $\mathbf{u}_0 = u_0\boldsymbol{\nu}$ and $f \geqslant 0$, respectively. We see that necessarily, $u_n \leqslant u_0$ on Γ_C. Furthermore, if, at $\mathbf{x} \in \Gamma_\mathrm{C}$, $u_n(\mathbf{x}) < u_0(\mathbf{x})$, we do not have contact, and thus, the stresses have to vanish, $T_n(\mathbf{u})(\mathbf{x}) = 0$ and $T_{tj}(\mathbf{u})(\mathbf{x}) = 0$, $j = 1, 2, 3$, as this corresponds to zero static boundary conditions. If, however, at $\mathbf{x} \in \Gamma_\mathrm{C}$,

$u_n(\mathbf{x}) = u_0(\mathbf{x})$, then the normal stresses have to point inwards, that is, $T_n(\mathbf{u})(\mathbf{x}) \leqslant 0$. If, additionally, for $\mathbf{T}_t(\mathbf{u})(\mathbf{x})$, $|\mathbf{T}_t(\mathbf{u})(\mathbf{x})| < f(\mathbf{x})\,|T_n(\mathbf{u})(\mathbf{x})|$, then $\mathbf{u}_t = 0$. If $|\mathbf{T}_t(\mathbf{u})(\mathbf{x})| = f(\mathbf{x})\,|T_n(\mathbf{u})(\mathbf{x})|$, then there exists $h \geqslant 0$ such that $\mathbf{u}_t(\mathbf{x}) = -h\mathbf{T}_t(\mathbf{u})(\mathbf{x})$. This model is called *Coulomb friction*.

Thus, we obtain: For every $\mathbf{x} \in \Gamma_C$:

$u_n(\mathbf{x}) < u_0(\mathbf{x})$. Then, $\mathbf{T}(\mathbf{u})(\mathbf{x}) = 0$.

$u_n(\mathbf{x}) = u_0(\mathbf{x})$. Then, $T_n(\mathbf{u})(\mathbf{x}) \leqslant 0$, and only the following two cases can occur:

$|\mathbf{T}_t(\mathbf{u})(\mathbf{x})| < f(\mathbf{x})\,|T_n(\mathbf{u})(\mathbf{x})|$; then, $\mathbf{u}_t(\mathbf{x}) = 0$, and

$|\mathbf{T}_t(\mathbf{u})(\mathbf{x})| = f(\mathbf{x})\,|T_n(\mathbf{u})(\mathbf{x})|$; then, there exists $h \geqslant 0$ with $\mathbf{u}_t(\mathbf{x}) = -h\mathbf{T}_t(\mathbf{u})(\mathbf{x})$.

All the different cases can also be summed up in an equivalent formulation given as

$$u_n \leqslant u_0, \tag{2.17a}$$

$$T_n(\mathbf{u}) \leqslant 0, \tag{2.17b}$$

$$(u_n - u_0)T_n(\mathbf{u}) = 0, \tag{2.17c}$$

$$|\mathbf{T}_t(\mathbf{u})| \leqslant f\,|T_n(\mathbf{u})|\,, \tag{2.17d}$$

$$(f\,|T_n(\mathbf{u})| - |\mathbf{T}_t(\mathbf{u})|)\mathbf{u}_t = 0, \tag{2.17e}$$

$$\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{u}_t + f\,|T_n(\mathbf{u})|\,|\mathbf{u}_t| = 0. \tag{2.17f}$$

To facilitate the discretisation of the given problem, we shall give an equivalent formulation as a variational inequality. To that end, define the space of displacements

$$V := \left\{\mathbf{v} \in \mathrm{H}^1(\Omega)^3 \colon \gamma_{0,\Gamma_D}(\mathbf{v}) = \mathbf{u}_0\right\}, \tag{2.18}$$

the closed, convex set of admissible displacements

$$\mathcal{K} := \{\mathbf{v} \in V \colon \gamma_{0,\Gamma_C}(\mathbf{v}) \cdot \boldsymbol{\nu} \leqslant u_0 \text{ almost everywhere on } \Gamma_C\}\,, \tag{2.19}$$

the bilinear form $a \colon V \times V \to \mathbb{R}$ by

$$a(\mathbf{v}, \mathbf{w}) := \int_\Omega \boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{w}) \mathrm{d}\mathbf{x} \quad \text{for } \mathbf{v},\, \mathbf{w} \in V, \tag{2.20}$$

and the linear functional $L \colon V \to \mathbb{R}$ by

$$L(\mathbf{v}) := \int_\Omega \mathbf{F} \cdot \mathbf{v} \mathrm{d}\mathbf{x} + \int_{\Gamma_N} \mathbf{G} \cdot \gamma_{0,\Gamma_N} \mathbf{v} \mathrm{d}s_{\mathbf{x}} \quad \text{for } \mathbf{v} \in V. \tag{2.21}$$

Note that in the case $\Gamma_C \cap \Gamma_D \neq \varnothing$ it can happen that the conditions $\gamma_{0,\Gamma_D}(v) = \mathbf{u}_0$ and $\gamma_{0,\Gamma_C}(v) \cdot \boldsymbol{\nu} \leqslant u_0$ are incompatible. We shall therefore assume in the following that $\mathbf{u}_0$ and u_0 are chosen in such a way that $\mathcal{K} \neq \varnothing$.

To incorporate the friction terms, we introduce the friction functional $j \colon V \times V \to \mathbb{R}$ by

$$j(\mathbf{v}, \mathbf{w}) := \int_{\Gamma_C} f\,|T_n(\mathbf{v})|\,|\mathbf{w}_t|\,\mathrm{d}s_{\mathbf{x}} \quad \text{for all } \mathbf{v},\, \mathbf{w} \in V. \tag{2.22}$$

Note that not only j, but even the mapping $\mathbf{v} \mapsto j(\mathbf{v}, \mathbf{v})$ is nondifferentiable and nonconvex.

Consider the following two problems.

Problem 2.1 (Classical formulation, Coulomb friction)**.** Find $\mathbf{u} \in (\mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega}))^3$ such that

$$\sigma_{ij,j}(\mathbf{u}) + F_i = 0, \qquad \text{in } \Omega, \quad i = 1,2,3, \tag{2.23a}$$
$$u_i = u_{0i} \qquad \text{on } \Gamma_\mathrm{D}, \quad i = 1,2,3, \tag{2.23b}$$
$$\sigma_i j(\mathbf{u})\nu_j = G_i \qquad \text{on } \Gamma_\mathrm{N}, \quad i = 1,2,3, \tag{2.23c}$$

and on Γ_C,

$$u_n \leqslant u_0, \tag{2.23d}$$
$$T_n(\mathbf{u}) \leqslant 0, \tag{2.23e}$$
$$(u_n - u_0)T_n(\mathbf{u}) = 0, \tag{2.23f}$$
$$|\mathbf{T}_t(\mathbf{u})| \leqslant f\,|T_n(\mathbf{u})|\,, \tag{2.23g}$$
$$(f\,|T_n(\mathbf{u})| - |\mathbf{T}_t(\mathbf{u})|)\mathbf{u}_t = 0, \tag{2.23h}$$
$$\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{u}_t + f\,|T_n(\mathbf{u})|\,|\mathbf{u}_t| = 0. \tag{2.23i}$$

Problem 2.2 (Variational formulation, Coulomb friction)**.** Find $\mathbf{u} \in \mathcal{K}$ such that for all $\mathbf{v} \in \mathcal{K}$,

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) + j(\mathbf{u}, \mathbf{v}) - j(\mathbf{u}, \mathbf{u}) \geqslant L(\mathbf{v} - \mathbf{u}). \tag{2.24}$$

Theorem 2.3. *The Problems 2.1 and 2.2 are equivalent in the following sense: If Γ and Γ_C are smooth enough, and $\mathbf{u} \in (\mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega}))^3$ solves one of the two above problems, it solves the other one, as well.*

This is [KO88, Theorem 10.1].

Proof. Integration by parts and the symmetry of $\boldsymbol{\sigma}(\mathbf{u})$ yield for an arbitrary $\mathbf{v} \in \mathcal{K}$

$$\begin{aligned} a(\mathbf{u}, \mathbf{v} - \mathbf{u}) &= \int_\Omega \sigma_{ij}(\mathbf{u})(v_{i,j} - u_{i,j})\mathrm{d}\mathbf{x} \\ &= \int_\Gamma \sigma_{ij}(\mathbf{u})\nu_j\gamma_{0,\Gamma}(v_i - u_i)\mathrm{d}s_\mathbf{x} - \int_\Omega \sigma_{ij,j}(\mathbf{u})(v_i - u_i)\mathrm{d}\mathbf{x}. \end{aligned} \tag{2.25}$$

Let $\mathbf{u}$ be a solution of the classical formulation. Then, $\sigma_{ij,j}(\mathbf{u}) = -F_i$, $v_i = u_i = 0$ on Γ_D and $\sigma_{ij}(\mathbf{u})\nu_j = G_i$ on Γ_N, from which we obtain

$$\begin{aligned} a(\mathbf{u}, \mathbf{v} - \mathbf{u}) &= \int_{\Gamma_\mathrm{N}} G_i\gamma_{0,\Gamma_\mathrm{N}}(v_i - u_i)\mathrm{d}s_\mathbf{x} + \int_{\Gamma_\mathrm{C}} \sigma_{ij}(\mathbf{u})\nu_j\gamma_{0,\Gamma_\mathrm{C}}(v_i - u_i)\mathrm{d}s_\mathbf{x} \\ &\quad + \int_\Omega F_i(v_i - u_i)\mathrm{d}\mathbf{x} \\ &= L(\mathbf{v} - \mathbf{u}) + \int_{\Gamma_\mathrm{C}} \sigma_{ij}(\mathbf{u})\nu_j\gamma_{0,\Gamma_\mathrm{C}}(v_i - u_i)\mathrm{d}s_\mathbf{x}. \end{aligned} \tag{2.26}$$

As $\sigma_{ij}(\mathbf{u})\nu_j = T_{ti}(\mathbf{u}) + T_n(\mathbf{u})\nu_i$, we see that

$$\begin{aligned} \int_{\Gamma_\mathrm{C}} \sigma_{ij}(\mathbf{u})\nu_j\gamma_{0,\Gamma_\mathrm{C}}(v_i - u_i)\mathrm{d}s_\mathbf{x} &= \int_{\Gamma_\mathrm{C}} (T_{ti}(\mathbf{u}) + T_n(\mathbf{u})\nu_i)\gamma_{0,\Gamma_\mathrm{C}}(v_i - u_i)\mathrm{d}s_\mathbf{x} \\ &= \int_{\Gamma_\mathrm{C}} \left[\mathbf{T}_t(\mathbf{u}) \cdot (\mathbf{v}_t - \mathbf{u}_t) + T_n(\mathbf{u})(v_n - u_n)\right] \mathrm{d}s_\mathbf{x} \end{aligned}$$

$$
\begin{aligned}
&= \int_{\Gamma_C} [\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{v}_t + T_n(\mathbf{u})(v_n - u_n)] \, ds_{\mathbf{x}} \\
&\quad - \int_{\Gamma_C} \mathbf{T}_t(\mathbf{u}) \cdot \mathbf{u}_t ds_{\mathbf{x}} \\
&= \int_{\Gamma_C} [\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{v}_t + T_n(\mathbf{u})(v_n - u_n)] \, ds_{\mathbf{x}} \\
&\quad + \int_{\Gamma_C} f \, |T_n(\mathbf{u})| \, |\mathbf{u}_t| \, ds_{\mathbf{x}} \\
&= \int_{\Gamma_C} [\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{v}_t + T_n(\mathbf{u})(v_n - u_n)] \, ds_{\mathbf{x}} + j(\mathbf{u}, \mathbf{u}).
\end{aligned} \tag{2.27}
$$

Note that for $\mathbf{v} \in \mathcal{K}$,

$$T_n(\mathbf{u})(v_n - u_n) = T_n(\mathbf{u})(v_n - u_0) + T_n(\mathbf{u})(u_0 - u_n) \geqslant T_n(\mathbf{u})(u_0 - u_n) = 0 \tag{2.28}$$

and

$$\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{v}_t \geqslant - |\mathbf{T}_t(\mathbf{u})| \, |\mathbf{v}_t| \geqslant -f \, |T_n(\mathbf{u})| \, |\mathbf{v}_t| \, , \tag{2.29}$$

which in turn yields

$$\int_{\Gamma_C} [\mathbf{T}_t(\mathbf{u}) \cdot \mathbf{v}_t + T_n(\mathbf{u})(v_n - u_n)] \geqslant -j(\mathbf{u}, \mathbf{v}). \tag{2.30}$$

Thus, $\mathbf{u}$ satisfies the variational formulation.

For the converse, pick $\mathbf{u}$ satisfying the variational formulation. Clearly, the Dirichlet conditions and the condition $u_n \leqslant u_0$ on Γ_C are defined by definition. For $\boldsymbol{\phi} \in \mathrm{D}(\Omega)^3$, we see easily that $\mathbf{v} := \mathbf{u} + \boldsymbol{\phi} \in \mathcal{K}$, and as the two functions $\mathbf{u}$ and $\mathbf{v}$ coincide on the boundary, we obtain by integration by parts that $\sigma_{ij,j}(\mathbf{u}) + F_i = 0$ on Ω.

Next, choose $\boldsymbol{\phi} = (\varphi_i) \in C^\infty(\overline{\Omega})$ such that there exists $U \subseteq \mathbb{R}^3$ open with $\operatorname{supp} \boldsymbol{\phi} \subseteq U$, $U \cap \Gamma \subseteq \Gamma_N$, and $\phi_j(\mathbf{x}) \neq 0$ for some $\mathbf{x} \in \Gamma_N$. Such a function can be found for every $\mathbf{x} \in \Gamma_N$ and $j = 1, 2, 3$ as Γ_N is relatively open in Γ. As $\mathbf{u}$ and $\mathbf{u} + \boldsymbol{\phi}$ coincide on Γ_C, this in turn yields $T_j(\mathbf{u}) = G_j$ on Γ_N. Thus, we obtain that for all $\mathbf{v} \in \mathcal{K}$,

$$\int_{\Gamma_C} [f \, |T_n(\mathbf{u})| \, (|\mathbf{v}_t| - |\mathbf{u}_t|) + \mathbf{T}_t(\mathbf{u}) \cdot (\mathbf{v}_t - \mathbf{u}_t) + T_n(\mathbf{u})(v_n - u_n)] \, ds_{\mathbf{x}} \geqslant 0. \tag{2.31}$$

Choose $\boldsymbol{\phi} \in C^\infty(\overline{\Omega})$ such that $\boldsymbol{\phi}_t = 0$, that is, $\boldsymbol{\phi} = \phi_n \boldsymbol{\nu}$ on Γ_C, and $\mathbf{u} + \boldsymbol{\phi} \in \mathcal{K}$. Then, we obtain

$$\int_{\Gamma_C} T_n(\mathbf{u}) \phi_n ds_{\mathbf{x}} \geqslant 0. \tag{2.32}$$

We can always choose $\phi_n \leqslant 0$, which entails $T_n(\mathbf{u}) \leqslant 0$. Furthermore, if at some point $\mathbf{x} \in \Gamma_C$, $u_n(\mathbf{x}) < u_0(\mathbf{x})$, then we can also choose $\boldsymbol{\phi}$ with $\phi_n(x) = u_0(\mathbf{x}) - u_n(\mathbf{x}) > 0$, and this gives $T_n(\mathbf{u})(\mathbf{x}) = 0$.

Choose $\boldsymbol{\phi} \in C^\infty(\overline{\Omega})$ such that $\phi_n = 0$. Clearly, $\mathbf{u} + \alpha \boldsymbol{\phi} \in \mathcal{K}$ for all $\alpha \in \mathbb{R}$, and thus

$$\int_{\Gamma_C} [f \, |T_n(\mathbf{u})| \, (|\mathbf{u}_t + \alpha \boldsymbol{\phi}_t| - |\mathbf{u}_t|) + \alpha \mathbf{T}_t(\mathbf{u}) \cdot \boldsymbol{\phi}_t] \, ds_{\mathbf{x}} \geqslant 0. \tag{2.33}$$

Choose $\boldsymbol{\phi}$ in such a way that $\boldsymbol{\phi}_t = 0$ whenever $\mathbf{u}_t = 0$. Then, $\frac{\partial}{\partial \alpha}|_{\alpha=0} \, |\mathbf{u}_t + \alpha \boldsymbol{\phi}_t| = \frac{\mathbf{u}_t \cdot \boldsymbol{\phi}_t}{|\mathbf{u}_t|}$ whereever $\mathbf{u}_t \neq 0$, and 0 otherwise. By Theorem 1.31, we obtain

$$\int_{\Gamma_C} f \, |T_n(\mathbf{u})| \, \frac{\mathbf{u}_t \cdot \phi_t}{|\mathbf{u}_t|} ds_{\mathbf{x}} \geqslant - \int_{\Gamma_C} \mathbf{T}_t(\mathbf{u}) \cdot \boldsymbol{\phi}_t ds_{\mathbf{x}}, \tag{2.34}$$

and thus, by replacing $\boldsymbol{\phi}$ by $-\boldsymbol{\phi}$,

$$f\,|T_n(\mathbf{u})|\,\mathbf{u}_t = -\mathbf{T}_t(\mathbf{u})\,|\mathbf{u}_t| \quad \text{whenever } \mathbf{u}_t \neq 0. \tag{2.35}$$

Taking absolute values and dividing by $|\mathbf{u}_t|$, we obtain

$$(f\,|T_n(\mathbf{u})| - |\mathbf{T}_t(\mathbf{u})|)\,\mathbf{u}_t = 0 \quad \text{on } \Gamma_\mathrm{C}, \tag{2.36}$$

and taking the scalar product with $\mathbf{u}_t$ and dividing by $|\mathbf{u}_t|$, we obtain

$$f\,|T_n(\mathbf{u})|\,|\mathbf{u}_t| = -\mathbf{T}_t(\mathbf{u})\cdot\mathbf{u}_t, \tag{2.37}$$

as in both cases, the assertion is trivial for $\mathbf{u}_t = 0$.

Thus, again for $\boldsymbol{\phi} \in C^\infty(\overline{\Omega})$ with $\phi_n = 0$, setting $\mathbf{v} := \mathbf{u} + \boldsymbol{\phi}$,

$$\int_{\Gamma_\mathrm{C}} \left[f\,|T_n(\mathbf{u})|\,(|\mathbf{v}_t| - |\mathbf{u}_t|) + \mathbf{T}_t(\mathbf{u})\cdot(\mathbf{v}_t - \mathbf{u}_t)\right] \mathrm{d}s_\mathbf{x} \geqslant 0. \tag{2.38}$$

Choose $\boldsymbol{\phi}$ in such a way that $\boldsymbol{\phi} = 0$ whenever $\mathbf{u}_t \neq 0$. Then, $|\mathbf{v}_t| - |\mathbf{u}_t| = |\boldsymbol{\phi}_t|$, and plugging in both $\boldsymbol{\phi}$ and $-\boldsymbol{\phi}$, we obtain

$$\left|\int_{\Gamma_\mathrm{C}} \mathbf{T}_t(\mathbf{u})\cdot\boldsymbol{\phi}_t \mathrm{d}s_\mathbf{x}\right| \leqslant \int_{\Gamma_\mathrm{C}} f\,|T_n(\mathbf{u})|\,|\boldsymbol{\phi}_t|\,\mathrm{d}s_\mathbf{x}. \tag{2.39}$$

By approximating the function $\operatorname{sign}(\mathbf{T}_t(\mathbf{u})\cdot\boldsymbol{\phi}_t)$ with smooth functions bounded by 1, we see that

$$\int_{\Gamma_\mathrm{C}} |\mathbf{T}_t(\mathbf{u})\cdot\boldsymbol{\phi}_t|\,\mathrm{d}s_\mathbf{x} \leqslant \int_{\Gamma_\mathrm{C}} f\,|T_n(\mathbf{u})|\,|\boldsymbol{\phi}_t|\,\mathrm{d}s_\mathbf{x}. \tag{2.40}$$

First, we note that obviously, from the above, $\mathbf{T}_t(\mathbf{u}) = 0$ whenever $f\,|\mathbf{T}_n(\mathbf{u})| = 0$. Second, using the weighted measure $\mathrm{d}\mu := f\,|T_n(\mathbf{u})|\,\mathrm{d}s_\mathbf{x}$, we see that, as $\mathbf{T}_t(\mathbf{u})\cdot\boldsymbol{\psi}_n = 0$ for any $\boldsymbol{\psi}$,

$$\boldsymbol{\psi} \mapsto \int_{\Gamma_\mathrm{C}} \mathbf{T}_t(\mathbf{u})\cdot\boldsymbol{\psi}\mathrm{d}s_\mathbf{x}, \tag{2.41}$$

defined on the smooth functions, can be extended to a linear functional of norm $\leqslant 1$ on the space $\mathrm{L}^1(\Gamma_\mathrm{C};\mu)^3$ of measurable functions integrable with respect to μ. The Riesz representation theorem, Theorem 1.29, yields the existence of a function $\mathbf{H} \in \mathrm{L}^\infty(\Gamma_\mathrm{C})$ with $\|\mathbf{H}\|_{\mathrm{L}^\infty(\Gamma_\mathrm{C})} \leqslant 1$ and

$$\mathbf{T}_t(\mathbf{u}) = f\,|T_n(\mathbf{u})|\,\mathbf{H}; \tag{2.42}$$

in particular, $|\mathbf{T}_t\mathbf{u}| \leqslant f\,|T_n(\mathbf{u})|$ whenever $f\,|T_n(\mathbf{u})| > 0$.

Combining the results, we see that $\mathbf{u}$ is a solution of the classical formulation. □

Due to the fact that $\mathbf{v} \mapsto j(\mathbf{v},\mathbf{v})$ is nonconvex, we shall analyse only a simplified problem. The simplified friction law is known as *Tresca friction* and corresponds to a friction functional $j\colon V \to \mathbb{R}$ defined by

$$j(\mathbf{v}) := \int_{\Gamma_\mathrm{C}} g\,|\mathbf{v}_t|\,\mathrm{d}s_\mathbf{x}. \tag{2.43}$$

This corresponds to the assumption that $g = f\,|T_n(\mathbf{u})|$ is constant, that is, the normal component of the normal stresses is replaced by a given slip stress. This functional is still nondifferentiable, but convex.

Thus, we define:

Problem 2.4 (Variational formulation, Tresca friction)**.** Find $\mathbf{u} \in \mathcal{K}$ such that for all $\mathbf{v} \in \mathcal{K}$,

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) + j(\mathbf{v}) - j(\mathbf{u}) \geqslant L(\mathbf{v} - \mathbf{u}). \tag{2.44}$$

The equivalent conditions on the contact boundary with friction are here

$$u_n \leqslant u_0, \tag{2.45a}$$
$$T_n(\mathbf{u}) \leqslant 0, \tag{2.45b}$$
$$(u_n - u_0)T_n(\mathbf{u}) = 0, \tag{2.45c}$$
$$|\mathbf{T}_t| \leqslant g, \tag{2.45d}$$
$$(g - |\mathbf{T}_t|)\mathbf{u}_t = 0, \tag{2.45e}$$
$$\mathbf{u}_t \cdot \mathbf{T}_t \leqslant 0. \tag{2.45f}$$

For the proof of existence and uniqueness of the solution of 2.4, it is helpful to rewrite the problem as a minimisation of a convex functional. Define therefore $J\colon V \to \mathbb{R}$ by

$$J(\mathbf{v}) := \frac{1}{2}a(\mathbf{v}, \mathbf{v}) - L(\mathbf{v}) + j(\mathbf{v}). \tag{2.46}$$

Problem 2.5 (Minimisation formulation, Tresca friction)**.** Find $\mathbf{u} \in \mathcal{K}$ such that

$$J(\mathbf{u}) = \min_{\mathbf{v} \in \mathcal{K}} J(\mathbf{v}). \tag{2.47}$$

The following result clearly is a consequence of Theorem 1.23.

Theorem 2.6. *Problems 2.4 and 2.5 are equivalent.*

We shall prove:

Theorem 2.7. *There exists a unique solution of Problem 2.5.*

Thus, by Theorem 2.6, there exists a unique solution of Problem 2.4. The proof uses the following variant of a Korn inequality, which is proved in [KO88, Lemma 6.2].

Theorem 2.8 (Korn's inequality)**.** *For* $|\Gamma_{\mathrm{D}}| > 0$*, there exists a constant* $C > 0$ *such that*

$$\|\mathbf{v}\|^2_{\mathrm{H}^1(\Omega)} \leqslant Ca(\mathbf{v}, \mathbf{v}) \quad \text{for all } \mathbf{v} \in V. \tag{2.48}$$

Proof of Theorem 2.7. Clearly, J defines a convex functional which is continuous with respect to the strong topology of $\mathcal{K} \subseteq V$. By Korn's inequality and the boundedness of L, we see that

$$J(\mathbf{v}) = \frac{1}{2}a(\mathbf{v}, \mathbf{v}) - L(\mathbf{v}) + j(\mathbf{v}) \geqslant C^{-1} \|\mathbf{v}\|^2_{\mathrm{H}^1(\Omega)} - C \|\mathbf{v}\|_{\mathrm{H}^1(\Omega)}, \tag{2.49}$$

and this expression tends to ∞ as $\|\mathbf{v}\|_{\mathrm{H}^1(\Omega)} \to \infty$. This yields the coercivity of J. Moreover, for $\mathbf{v} \neq \mathbf{w}$ and $t \in (0, 1)$, by the convexity of j,

$$\begin{aligned} J(t\mathbf{v} + (1-t)\mathbf{w}) &= \frac{1}{2}\left[ta(\mathbf{v}, \mathbf{v}) + (1-t)a(\mathbf{w}, \mathbf{w}) - t(1-t)a(\mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w})\right] \\ &\quad - tL(\mathbf{v}) - (1-t)L(\mathbf{w}) + j(t\mathbf{v} + (1-t)\mathbf{w}) \\ &\leqslant \frac{1}{2}\left[ta(\mathbf{v}, \mathbf{v}) + (1-t)a(\mathbf{w}, \mathbf{w}) - t(1-t)a(\mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w})\right] \\ &\quad - tL(\mathbf{v}) - (1-t)L(\mathbf{w}) + tj(\mathbf{v}) + (1-t)j(\mathbf{w}). \end{aligned} \tag{2.50}$$

Applying Korn's inequality, we see that $t(1-t)a(\mathbf{v}-\mathbf{w},\mathbf{v}-\mathbf{w}) > 0$, and we obtain

$$J(t\mathbf{v} + (1-t)\mathbf{w}) < tJ(\mathbf{v}) + (1-t)J(\mathbf{w}), \tag{2.51}$$

that is, J is strictly convex.

Thus, Theorem 1.25 yields the result. □

2.2.2 Friction under the Plane Stress Assumption

Consider the plane stress assumption as described in Subsection 2.1.3. As the boundary conditions are independent of x_3, we see that the contact boundary is given as $\Gamma_\mathrm{C} \times [-h, h]$, where $\Gamma_\mathrm{C} \subseteq \Gamma := \partial\Omega$. For h small, we can assume that all functions are constant with respect to x_3. Thus, we see that the friction term satisfies

$$j(v) = \int_{-h}^{h} \int_{\Gamma_\mathrm{C}} g\,|\mathbf{v}_t|\,\mathrm{d}s_{(x_1,x_2)}\mathrm{d}x_3 = 2h \int_{\Gamma_\mathrm{C}} g\,|\mathbf{v}_t|\,\mathrm{d}s_{(x_1,x_2)}. \tag{2.52}$$

Similarly, as $\sigma_{3i}(\mathbf{v}) = 0$,

$$\begin{aligned} a(\mathbf{v},\mathbf{w}) &= \int_{-h}^{h} \int_{\Omega} [\sigma_{11}(\mathbf{v})\varepsilon_{11}(\mathbf{w}) + 2\sigma_{12}(\mathbf{v})\varepsilon_{12}(\mathbf{w}) + \sigma_{22}(\mathbf{v})\varepsilon_{22}(\mathbf{w})]\,\mathrm{d}(x_1,x_2)\mathrm{d}x_3 \\ &= 2h \int_{\Omega} [\sigma_{11}(\mathbf{v})\varepsilon_{11}(\mathbf{w}) + 2\sigma_{12}(\mathbf{v})\varepsilon_{12}(\mathbf{w}) + \sigma_{22}(\mathbf{v})\varepsilon_{22}(\mathbf{w})]\,\mathrm{d}(x_1,x_2) \end{aligned} \tag{2.53}$$

and, as $v_3(0) = 0$,

$$\begin{aligned} L(\mathbf{v}) &= \int_{-h}^{h} \left[\int_{\Omega} \mathbf{F}\cdot\mathbf{v}\mathrm{d}(x_1,x_2) + \int_{\Gamma_\mathrm{N}} \mathbf{G}\cdot\gamma_{0,\Gamma_\mathrm{N}}\mathbf{v}\mathrm{d}s_{(x_1,x_2)}\right]\mathrm{d}x_3 \\ &= 2h\left[\int_{\Omega} (F_1v_1 + F_2v_2)\mathrm{d}(x_1,x_2) + \int_{\Gamma_\mathrm{N}} (G_1\gamma_{0,\Gamma_\mathrm{N}}v_1 + G_2\gamma_{0,\Gamma_\mathrm{N}}v_2)\mathrm{d}s_{(x_1,x_2)}\right]. \end{aligned} \tag{2.54}$$

Thus, setting

$$\tilde{V} := \left\{\mathbf{v} \in \mathrm{H}^1(\Omega)^2 \colon \gamma_{0,\Gamma_\mathrm{D}}\mathbf{v} = 0\right\} \tag{2.55}$$

and

$$\tilde{\mathcal{K}} := \left\{\mathbf{v} \in \tilde{V} \colon \gamma_{0,\Gamma_\mathrm{C}}(\mathbf{v})\cdot\boldsymbol{\nu} \leqslant u_0 \text{ almost everywhere on } \Gamma_\mathrm{C}\right\} \tag{2.56}$$

and defining the bilinear form $\tilde{a}\colon \tilde{V}\times\tilde{V} \to \mathbb{R}$, the linear form $\tilde{L}\colon \tilde{V} \to \mathbb{R}$ and the nonlinear functional $\tilde{j}\colon \tilde{V} \to \mathbb{R}$ with $\tilde{\mathbf{F}} := (F_1, F_2)$ and $\tilde{\mathbf{G}} := (G_1, G_2)$ by

$$\tilde{a}(\mathbf{v},\mathbf{w}) := \int_{\Omega} (\boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{w}))\,\mathrm{d}(x_1,x_2), \tag{2.57}$$

$$\tilde{L}(\mathbf{v}) := \int_{\Omega} \tilde{\mathbf{F}}\cdot\mathbf{v}\mathrm{d}(x_1,x_2) + \int_{\Gamma_\mathrm{N}} \tilde{\mathbf{G}}\cdot\gamma_{0,\Gamma_\mathrm{N}}\mathbf{v}\mathrm{d}s_{(x_1,x_2)}, \tag{2.58}$$

and

$$\tilde{j}(\mathbf{v}) := \int_{\Gamma_\mathrm{C}} g\,|\mathbf{v}_t|\,\mathrm{d}s_{(x_1,x_2)}, \tag{2.59}$$

where $\mathbf{v}_t$ is defined for $\mathbf{v} = (v_1, v_2)$ as in the three-dimensional case using the unit normal vector $\boldsymbol{\nu} = (\nu_1, \nu_2)$ on the two-dimensional boundary $\Gamma = \partial\Omega$, together with the energy functional $\tilde{J} \colon \tilde{V} \to \mathbb{R}$

$$\tilde{J}(\mathbf{v}) := \frac{1}{2}\tilde{a}(\mathbf{v}, \mathbf{v}) - \tilde{L}(\mathbf{v}) + \tilde{j}(\mathbf{v}), \tag{2.60}$$

we have that

$$J(\mathbf{v}) = 2h\tilde{J}(\mathbf{v}). \tag{2.61}$$

This gives:

Problem 2.9 (Tresca friction, minimisation formulation, plane stress)**.** Find $\mathbf{u} \in \tilde{\mathcal{K}}$ such that

$$\tilde{J}(\mathbf{u}) = \min_{\mathbf{v} \in \tilde{\mathcal{K}}} \tilde{J}(\mathbf{v}). \tag{2.62}$$

Thus, the problems in the plane stress situation are of the same type as in the full three-dimensional situation, we just replace the sets and the operators accordingly. It is now also possible to formulate a two-dimensional analogue of Problem 2.4 based on the above minimisation problem.

Note that the solution $\mathbf{u}$ of the plane stress problem does not necessarily satisfy a corresponding three-dimensional problem, as we applied an approximation while going from J to $\tilde{J}$.

2.3 The Primal-Dual Formulation

It is possible to discretise Problem 2.4 directly, but the solution algorithms for problems of this kind are unsatisfying in the *hp*-context that will be the focus of the later chapters of the present work. Instead, it is possible to give an equivalent formulation as a saddle point problem. For this, note first that

$$j(\mathbf{v}) = \sup_{\boldsymbol{\mu} \in \Lambda} b(\mathbf{v}, \boldsymbol{\mu}), \tag{2.63}$$

where

$$\Lambda := \left\{ \boldsymbol{\mu} \in \mathrm{L}^\infty(\Gamma_\mathrm{C})^3 \colon |\boldsymbol{\mu}| \leqslant 1 \text{ and } \mu_n = 0 \text{ almost everywhere on } \Gamma_\mathrm{C} \right\}, \tag{2.64}$$

and the bilinear form $b \colon V \times \Lambda \to \mathbb{R}$ is given by

$$b(\mathbf{v}, \boldsymbol{\mu}) := \int_{\Gamma_\mathrm{C}} g\mathbf{v}_t \cdot \boldsymbol{\mu} \mathrm{d}s_{\mathbf{x}}. \tag{2.65}$$

From a functional analytic point of view, it is sensible to define a surrounding space W of Λ which has better analytic properties than $\mathrm{L}^\infty(\Gamma_\mathrm{C})$. The largest reasonable space is

$$\left(\gamma_{0,\Gamma_\mathrm{C}} V\right)^*, \tag{2.66}$$

the dual space of the traces on Γ_C of functions in V, as for these functions, the mapping b can still be defined in the sense of a duality product. For the moment, though, we shall choose

$$W := \mathrm{L}^2(\Gamma_\mathrm{C})^3. \tag{2.67}$$

Here, the inclusion $\Lambda \subseteq W$ is trivial, and we can also define $b \colon V \times W \to \mathbb{R}$ by continuous extension.

Furthermore, Λ is closed (as all sequences converging in $L^2(\Gamma_C)$ admit a subsequence converging almost everywhere; see Theorem 1.30) and convex, but clearly not a linear space. We define the Lagrange functional $\mathcal{L}\colon V \times W \to \mathbb{R}$ by

$$\mathcal{L}(\mathbf{v}, \boldsymbol{\mu}) := \frac{1}{2} a(\mathbf{v}, \mathbf{v}) - L(\mathbf{v}) + b(\mathbf{v}, \boldsymbol{\mu}) \tag{2.68}$$

and note that, by (2.63),

$$J(\mathbf{v}) = \sup_{\boldsymbol{\mu} \in \Lambda} \mathcal{L}(\mathbf{v}, \boldsymbol{\mu}), \tag{2.69}$$

and thus, for the unique minimiser $\mathbf{u}$ of J,

$$J(\mathbf{u}) = \inf_{\mathbf{v} \in \mathcal{K}} J(\mathbf{v}) = \inf_{\mathbf{v} \in \mathcal{K}} \sup_{\boldsymbol{\mu} \in \Lambda} \mathcal{L}(\mathbf{v}, \boldsymbol{\mu}). \tag{2.70}$$

We consider the following formulation of our problem.

Problem 2.10 (Primal-dual formulation, Tresca friction)**.** Find $(\mathbf{u}, \boldsymbol{\lambda}) \in \mathcal{K} \times \Lambda$ such that for all $(\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{K} \times \Lambda$,

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) + b(\mathbf{v} - \mathbf{u}, \boldsymbol{\lambda}) \geqslant L(\mathbf{v} - \mathbf{u}), \tag{2.71a}$$

$$b(\mathbf{u}, \boldsymbol{\mu} - \boldsymbol{\lambda}) \leqslant 0. \tag{2.71b}$$

Theorem 2.11. *For every solution* $(\mathbf{u}, \boldsymbol{\lambda})$ *of Problem 2.10,* $\mathbf{u}$ *solves Problem 2.5.*

Proof. First of all, note that the primal-dual formulation implies

$$\mathcal{L}(\mathbf{u}, \boldsymbol{\mu}) \leqslant \mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) \leqslant \mathcal{L}(\mathbf{v}, \boldsymbol{\lambda}) \quad \text{for all } (\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{K} \times \Lambda, \tag{2.72}$$

as the first inequality follows from (2.71b) and the second one from (2.71a). Furthermore, for any such pair $(\mathbf{u}, \boldsymbol{\lambda})$, it is clear that $\mathbf{u}$ minimises J, as obviously $\mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) = J(\mathbf{u})$, and thus

$$J(\mathbf{v}) \geqslant \mathcal{L}(\mathbf{v}, \boldsymbol{\lambda}) \geqslant \mathcal{L}(\mathbf{u}, \boldsymbol{\lambda}) = J(\mathbf{u}). \tag{2.73}$$

We see thus that for any solution $(\mathbf{u}, \boldsymbol{\lambda})$ of the primal-dual formulation, $\mathbf{u}$ is a minimiser of J. ☐

It follows that there exists at most one solution to Problem 2.10. Next, we want to prove existence of a solution, which also establishes the equivalence of the three formulations of the equations of elasticity with Tresca friction given above.

Theorem 2.12. *There exists a solution to Problem 2.10.*

Proof. Consider the space $\mathfrak{V} := V \times W$ endowed with the norm

$$\|(\mathbf{v}, \boldsymbol{\mu})\|_{\mathfrak{V}} := \left(\|\mathbf{v}\|^2_{H^1(\Omega)} + \|\boldsymbol{\mu}\|^2_{L^2(\Gamma_C)} \right)^{1/2}. \tag{2.74}$$

Defining the bilinear form $\mathfrak{a}\colon \mathfrak{V} \times \mathfrak{V} \to \mathbb{R}$,

$$\mathfrak{a}((\mathbf{v}, \boldsymbol{\mu}), (\mathbf{w}, \boldsymbol{\eta})) := a(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, \boldsymbol{\mu}) - b(\mathbf{v}, \boldsymbol{\eta}) \tag{2.75}$$

and the linear form $\mathfrak{L}\colon \mathfrak{V} \to \mathbb{R}$,

$$\mathfrak{L}((\mathbf{v}, \boldsymbol{\mu})) := L(\mathbf{v}), \tag{2.76}$$

it is readily seen that $\mathfrak{a}$ and $\mathfrak{L}$ are bounded. The set $\mathfrak{K} := \mathcal{K} \times \Lambda \subseteq \mathfrak{V}$ being convex and closed, the given problem is equivalent to finding $\mathfrak{u} = (\mathbf{u}, \boldsymbol{\lambda}) \in \mathfrak{K}$ such that for all $\mathfrak{v} = (\mathbf{v}, \boldsymbol{\mu}) \in \mathfrak{K}$,

$$\mathfrak{a}(\mathfrak{u}, \mathfrak{v} - \mathfrak{u}) \geqslant \mathfrak{L}(\mathfrak{v} - \mathfrak{u}). \tag{2.77}$$

For all $\mathfrak{v} = (\mathbf{v}, \boldsymbol{\mu}) \in \mathfrak{V}$,

$$\mathfrak{a}(\mathfrak{v}, \mathfrak{v}) = a(\mathbf{v}, \mathbf{v}), \tag{2.78}$$

and as Λ is bounded and a is coercive, it follows that $\mathfrak{v} \mapsto \frac{1}{2}\mathfrak{a}(\mathfrak{v}, \mathfrak{v}) - \mathfrak{L}(\mathfrak{v})$ is coercive, convex and continuous. Thus, by Theorem 1.25, there exists a minimiser $\mathfrak{u} = (\mathbf{u}, \boldsymbol{\lambda})$ of

$$\frac{1}{2}\mathfrak{a}(\mathfrak{v}, \mathfrak{v}) - \mathfrak{L}(\mathfrak{v}) \tag{2.79}$$

in $\mathfrak{K}$, and this pair solves the primal-dual formulation. □

Note that we did not have to prove the strict coercivity of the energy functional: The uniqueness of the solution follows from Theorem 2.7.

Theorem 2.13. *The Lagrange multiplier* $\boldsymbol{\lambda}$ *of the solution of Problem 2.10 satisfies* $g\boldsymbol{\lambda} = \mathbf{T}_t(\mathbf{u})$.

If, additionally, $g > 0$ *on* Γ_C*,* $\boldsymbol{\lambda}$ *is unique.*

Proof. First, note that we have shown already that $\mathbf{u}$ is unique. Thus, the first assertion follows from an integration by parts.

The second statement follows trivially by dividing by g. □

Remark 2.14. In a similar way, it is possible to derive a primal-dual formulation of the minimisation formulation in Problem 2.9. Here, we have

$$\tilde{W} := \mathrm{L}^2(\Gamma_C)^2, \tag{2.80}$$

$$\tilde{\Lambda} := \left\{ \boldsymbol{\mu} \in \tilde{W} \colon \|\boldsymbol{\mu}\|_{\mathrm{L}^\infty(\Gamma_C)} \leqslant 1 \text{ and } \mu_n = 0 \text{ almost everywhere on } \Gamma_C \right\}, \tag{2.81}$$

and the bilinear form $\tilde{b} \colon \tilde{V} \times \tilde{W} \to \mathbb{R}$ is given by

$$\tilde{b}(\mathbf{v}, \boldsymbol{\mu}) = \int_{\Gamma_C} g\mathbf{v}_t \cdot \boldsymbol{\mu} \mathrm{d}s_{(x_1,x_2)}. \tag{2.82}$$

This will be the basis of the numerical methods developed in the subsequent parts of the present work.

Chapter 3

Finite Element Methods for Mixed Variational Inequalities

In this chapter, we describe how finite element methods have to be formulated for variational inequalities, with a focus on mixed variational problems. The classical references for this topic are [GLT81] and [Glo84], where first order approximations are discussed. These approaches, however, cannot be generalised to an *hp*-context. For high order polynomials, important results were given in [Mai01a] in the context of boundary element methods, generalising earlier results in [GS93] to a true *p*-method. The idea of using Gauss-Lobatto nodes in spectral collocation methods was then already well known, see, for example, [BM92], and here, this was taken to the logical conclusion by also discretising inequality constraints by restricting them to Gauss-Lobatto points, and using the positivity-preserving Bernstein polynomials of nonnegative test functions in conjunction with the exactness of Gauss-Lobatto quadrature to prove the convergence in the sense of Glowinski of the discretised convex sets. Based on these, in [MS05, MS07, Kre04, KS07], some further applications to boundary and finite element methods were analysed, all of which restrict themselves to variational inequalities of the first kind, but also for nonlinear operators.

Some advances with respect to *hp*-finite element methods for variational inequalities of the second kind were done by Chernov in [Che06], where he considered penalty approaches, and in [CMS08], where the focus lies on an a priori estimate. The latter article, however, has a significant deficit: The estimates are done for a primal formulation which, in itself, is clearly numerically infeasible due to the necessity of determining certain integrals of absolute values of polynomials exactly. Thus, the actual calculation is done using a primal-dual formulation, which is *not* equivalent to the formulation for which the convergence rate is proved, and for this method, the estimates are not directly applicable.

The aim of this chapter is therefore also to give an a priori convergence rate result directly for a discrete formulation which can be solved numerically. Note that nevertheless, there is the implicit assumption that the solver for the discrete problem yields the exact solution, which typically is not the case: In practice, one uses an iterative solution algorithm — which may or may not have a finite termination property — and stops the algorithm as soon as the solution is “good enough” in some appropriate sense.

3.1 Abstract Finite Elements for Mixed Variational Inequalities

Let V, W be Hilbert spaces with $\mathcal{K} \subseteq V$, $\Lambda \subseteq W$ nonempty, closed and convex, $a\colon V \times V \to \mathbb{R}$ bilinear, symmetric, bounded and coercive, $b\colon V \times W \to \mathbb{R}$ bilinear and bounded, $F\colon V \to \mathbb{R}$, $G\colon W \to \mathbb{R}$ linear and bounded. We consider the following mixed variational inequality problem:

Problem 3.1 (Continuous abstract variational inequality)**.** Find $(u, \lambda) \in \mathcal{K} \times \Lambda$ with

$$\begin{aligned} a(u, v-u) + b(v-u, \lambda) &\geqslant F(v-u), \quad v \in \mathcal{K}, && (3.1a)\\ b(u, \mu-\lambda) &\leqslant G(\mu-\lambda), \quad \mu \in \Lambda. && (3.1b) \end{aligned}$$

We shall assume that Problem 3.1 has a unique solution (u, λ). This is guaranteed, for example, if we consider the situation as in Section 2.3.

Consider sequences $(V_N)_N$, $(W_N)_N$ of finite-dimensional subspaces of V and W, $V_N \subseteq V$, $W_N \subseteq W$, where the index N runs over an infinite subset of $\mathbb{N}$. Typically, one sets $N := \dim V_N$. From these, define closed convex subsets $(\mathcal{K}_N)_N$, $(\Lambda_N)_N$ such that $\mathcal{K}_N \subseteq V_N$ and $\Lambda_N \subseteq W_N$.

If we were able to consider $\mathcal{K}_N := \mathcal{K} \cap V_N \subseteq \mathcal{K}$ and $\Lambda_N := \Lambda \cap W_N \subseteq \Lambda$, the situation would be relatively simple, but this is for practical problems, especially for p-versions, not possible in an actual implementation of the algorithm: Consider, for example, that $V = \mathrm{H}^1(\Omega)$ for $\Omega = (0,1) \subseteq \mathbb{R}^1$ and

$$\mathcal{K} := \{v \in V : v \geqslant 0 \text{ almost everywhere}\}, \tag{3.2}$$

and the approximation $V_N := \mathcal{P}^N$. Then, the constraint $v_N \geqslant 0$ almost everywhere cannot easily be checked numerically. Thus, we cannot assume that the approximations $\mathcal{K}_N$ and Λ_N satisfy $\mathcal{K}_N \subseteq \mathcal{K}$ and $\Lambda_N \subseteq \Lambda$, respectively, that is, we have to deal with a *non-conforming approximation*. The fundamental notion in this context is:

Definition 3.2 ([AG00, Definition 2.1], [Glo84, p. 9])**.** Let V be a Hilbert space, $\mathcal{K} \subseteq V$ nonempty, convex and closed, and $(\mathcal{K}_N)_N$ a sequence of convex, closed subsets of V.

Then, $\mathcal{K}_N$ *converges to* $\mathcal{K}$ *in the sense of Glowinski*, $\mathcal{K}_N \xrightarrow{\mathrm{Gl}} \mathcal{K}$, if and only if

- for all sequences (v_N) with $v_N \in \mathcal{K}_N$ for all N, and such that v_N converges weakly to some $v \in V$, the limit satisfies $v \in \mathcal{K}$, and
- there exists a dense subset $D \subseteq \mathcal{K}$ such that for all $v \in D$, there exists a sequence (v_N) with $v_N \in \mathcal{K}_N$ for all N such that v_N converges strongly to v.

In some sense, this means that $\mathcal{K}_N$ is neither "too large" nor "too small" compared to $\mathcal{K}$. Note that, as $\mathcal{K}$ is nonempty, $\mathcal{K}_N$ has to be nonempty, as well.

The discretisation of Problem 3.1 is given as:

Problem 3.3 (Discretised abstract variational inequality)**.** Find $(u_N, \lambda_N) \in \mathcal{K}_N \times \Lambda_N$ such that

$$\begin{aligned} a(u_N, v_N-u_N) + b(v_N-u_N, \lambda_N) &\geqslant F(v_N-u_N), \quad v_N \in \mathcal{K}_N, && (3.3a)\\ b(u_N, \mu_N-\lambda_N) &\leqslant G(\mu_N-\lambda_N), \quad \mu_N \in \Lambda_N. && (3.3b) \end{aligned}$$

We shall also assume that this discretisation has a unique solution for all N. This can be shown similarly as in the continuous situation.

Using these definitions, we are able to show a relatively general convergence result.

Theorem 3.4 ([HHNL88, Section 1.1.52, Theorem 5.3], [Glo84, Chapter I, Theorem 5.2])**.** *Assume that* Λ, Λ_N *are uniformly bounded, that is, there exists a constant* $C > 0$ *such that*

$$\begin{aligned} \|\mu\|_W &\leqslant C \quad \textit{for all } \mu \in \Lambda \textit{ and} && (3.4)\\ \|\mu_N\|_W &\leqslant C \quad \textit{for all } \mu_N \in \Lambda_N \textit{ and all } N. && (3.5) \end{aligned}$$

Furthermore, suppose that $\mathcal{K}_N \xrightarrow{\text{Gl}} \mathcal{K}$ *and* $\Lambda_N \xrightarrow{\text{Gl}} \Lambda$.

Then, the solutions (u_N, λ_N) *of Problems 3.3 and the solution* (u, λ) *of Problem 3.1 satisfy* $u_N \to u$ *in* V *and* $\lambda_N \rightharpoonup \lambda$ *in* W.

Proof. Assume that the dense subsets given by the convergence in the sense of Glowinski are denoted by $D \subseteq \mathcal{K}$ and $M \subseteq \Lambda$, respectively.

By definition, (λ_N) is bounded in W. Choose an arbitrary $v_0 \in D$ and $v_N \in \mathcal{K}_N$ with $v_N \to v_0$ in V, which is possible due to $\mathcal{K}_N \xrightarrow{\text{Gl}} \mathcal{K}$. Then, there exists $C > 0$ such that $\|v_N\|_V \leqslant C$ for all N. Plugging v_N into (3.3a), we see that

$$a(u_N, u_N) \leqslant a(u_N, v_N) + b(v_N - u_N, \lambda_N) - F(v_N - u_N), \tag{3.6}$$

and the boundedness of λ_N yields the existence of a constant $C > 0$ such that for all N,

$$a(u_N, u_N) \leqslant C(1 + \|u_N\|_V). \tag{3.7}$$

Applying the coercivity of a, we obtain boundedness of $\|u_N\|_V$.

Thus, we see that we can choose a subsequence $(u_{N'}, \lambda_{N'})$ of (u_N, λ_N) converging weakly to some (u^*, λ^*) in $V \times W$. As $\mathcal{K}_N \xrightarrow{\text{Gl}} \mathcal{K}$ and $\Lambda_N \xrightarrow{\text{Gl}} \Lambda$, we see that $(u^*, \lambda^*) \in \mathcal{K} \times \Lambda$. We now want to show that (u^*, λ^*) solves the continuous problem.

Let $(v, \mu) \in D \times M$ and choose $(v_N, \mu_N) \in \mathcal{K}_N \times \Lambda_N$ with $(v_N, \mu_N) \to (v, \mu)$ strongly in $V \times W$. Passing to the limit inferior in (3.3a) and (3.3b), we obtain

$$\begin{aligned} \liminf_{N'\to\infty} a(u_{N'}, u_{N'}) + \liminf_{N'\to\infty} b(u_{N'}, \lambda_{N'}) &\leqslant \liminf_{N'\to\infty} [a(u_{N'}, u_{N'}) + b(u_{N'}, \lambda_{N'})] \\ &\leqslant \liminf_{N'\to\infty} [F(u_{N'} - v_{N'}) \\ &\qquad + a(u_{N'}, v_{N'}) + b(v_{N'}, \lambda_{N'})], \end{aligned} \tag{3.8}$$

$$\begin{aligned} \liminf_{N'\to\infty} b(u_{N'}, \mu_{N'}) + \liminf_{N'\to\infty} G(\lambda_{N'} - \mu_{N'}) &\leqslant \liminf_{N'\to\infty} [b(u_{N'}, \mu_{N'}) + G(\lambda_{N'} - \mu_{N'})] \\ &\leqslant \liminf_{N'\to\infty} b(u_{N'}, \lambda_{N'}). \end{aligned} \tag{3.9}$$

From the lower semicontinuity of a, which is a consequence of Lemma 1.24, the continuity of F and the strong convergence of the sequence (v_N) (and thus also $(v_{N'})$), we obtain that $\lim_{N'\to\infty} F(u_{N'} - v_{N'}) = F(u^* - v)$, $\lim_{N'\to\infty} b(v_{N'}, \lambda_{N'}) = b(v, \lambda^*)$ and $\lim_{N'\to\infty} a(u_{N'}, v_{N'}) = a(u^*, v)$ together with $a(u^*, u^*) \leqslant \liminf_{N'\to\infty} a(u_{N'}, u_{N'})$, which yields

$$a(u^*, u^*) + \liminf_{N'\to\infty} b(u_{N'}, \lambda_{N'}) \leqslant F(u^* - v) + a(u^*, v) + b(v, \lambda^*), \tag{3.10}$$

and due to the continuity of G and the strong convergence of (μ_N),

$$b(u^*, \mu) + G(\lambda^* - \mu) \leqslant \liminf_{N'\to\infty} b(u_{N'}, \lambda_{N'}). \tag{3.11}$$

Letting $\mu \to \lambda^*$ in the last inequality, we have that

$$b(u^*, \lambda^*) \leqslant \liminf_{N'\to\infty} b(u_{N'}, \lambda_{N'}), \tag{3.12}$$

and thus with (3.10),

$$a(u^*, u^*) + b(u^*, \lambda^*) \leqslant F(u^* - v) + a(u^*, v) + b(v, \lambda^*). \tag{3.13}$$

Letting $v \to u^*$ in (3.10), we obtain

$$\liminf_{N' \to \infty} b(u_{N'}, \lambda_{N'}) \leqslant b(u^*, \lambda^*), \tag{3.14}$$

which gives in (3.11) that

$$b(u^*, \mu) + G(\lambda^* - \mu) \leqslant b(u^*, \lambda^*). \tag{3.15}$$

Finally, choose an arbitrary pair $(v, \mu) \in \mathcal{K} \times \Lambda$ and a sequence $(v_k, \mu_k) \in D \times M$ converging strongly to (v, μ). Thus, we see that (u^*, λ^*) solves Problem 3.1, and due to the uniqueness, $u^* = u$ and $\lambda^* = \lambda$. As this argument works for all subsequences $(u_{N'}, \lambda_{N'})$, the entire sequence (u_N, λ_N) converges weakly, as well.

To prove the strong convergence of u_N to u, we note that, by the coercivity of a, there exists a constant $C > 0$ such that

$$\|u - u_N\|_V^2 \leqslant C^2 a(u - u_N, u - u_N) = C^2 \left[a(u, u) - 2a(u, u_N) + a(u_N, u_N)\right]. \tag{3.16}$$

With a sequence (v_N) such that $v_N \in \mathcal{K}_N$ converging strongly to $v \in D$,

$$a(u_N, u_N) \leqslant a(u_N, v_N) + b(v_N - u_N, \lambda_N) - F(v_N - u_N), \tag{3.17}$$

and thus,

$$\limsup_{N \to \infty} a(u_N, u_N) \leqslant a(u, v) + b(v - u, \lambda) - F(v - u), \tag{3.18}$$

where we applied (3.12). Letting $v \to u$,

$$\limsup_{N \to \infty} a(u_N, u_N) \leqslant a(u, u). \tag{3.19}$$

Therefore, by the weak convergence $u_N \rightharpoonup u$,

$$\begin{aligned} 0 &\leqslant \liminf_{N \to \infty} \|u - u_N\|_V^2 \leqslant \limsup_{N \to \infty} \|u - u_N\|_V^2 \\ &\leqslant C^2 \limsup_{N \to \infty} \left[a(u, u) - 2a(u, u_N) + a(u_N, u_N)\right] \\ &\leqslant 0, \end{aligned} \tag{3.20}$$

and the assertion follows. □

Remark 3.5. The above proof shows that actually, the uniqueness of u and λ is not necessary for a corresponding result to hold. In particular, if λ and λ_N are not unique, we see that still $u_N \to u$ in V, but for λ_N, we only have convergence of subsequences to some solution of the continuous problem. If, however, λ_N is not necessarily unique, but λ is, then we still obtain weak convergence of the entire sequence.

The last result has the disadvantage that only convergence is ensured, but no rate is given. As such, it is not very useful for the practitioner. To give an a priori result on the rate of the convergence, we need to introduce the concept of *inf-sup conditions*, which is also fundamental for finite elements for mixed variational equations; see [BF91] for applications.

For simplicity, we shall restrict ourselves to the situation that the convex subset $\mathcal{K} \subseteq V$ is actually the entire space. In this situation, we have the advantage that we can formulate the discrete problem solved by u_N on V_N. Furthermore, by substituting v by $u \pm v$ and v_N by $u_N \pm v_N$, respectively, the equalities (3.1a) and (3.3a) simplify to

$$a(u, v) + b(v, \lambda) = F(v) \quad \text{for all } v \in V \tag{3.1a'}$$

and

$$a(u_N, v_N) + b(v_N, \lambda_N) = F(v_N) \quad \text{for all } v_N \in V_N, \tag{3.3a'}$$

respectively, where we stress that (3.1a') can also be used with $v := v_N \in V_N$ as $V_N \subseteq V$. This assumption is not necessary, however, in the special case that K is a convex cone given as

$$\mathcal{K} = \{v \in V : c(v, \varphi) \leqslant H(\varphi) \text{ for all } \varphi \in \Phi\}, \tag{3.21}$$

where Z is another Hilbert space, $c: V \times Z \to \mathbb{R}$ is a bounded bilinear form, $H: Z \to \mathbb{R}$ is a bounded linear functional, and $\Phi \subseteq Z$ is a closed, convex set. Here, the inequality constraint in $\mathcal{K}$ can be again formulated by duality, and we obtain a *dual-dual formulation* (*two-fold saddle point problem*; see [GM01] for other applications of dual-dual formulations. This is, for example, true for the contact problem with friction we are considering. Details can be found in [HHNL88, pp. 204].

Definition 3.6 ([BF91])**.** The bounded bilinear form b is said to satisfy an *inf-sup condition on $V \times W$* if and only if there exists a constant $\beta > 0$ such that

$$\inf_{\mu \in W} \sup_{v \in V} \frac{b(v, \mu)}{\|v\|_V \|\mu\|_W} \geqslant \beta. \tag{3.22}$$

b is said to satisfy a *(non-uniform) discrete inf-sup condition on $(V_N \times W_N)_N$* (*Babuška-Brezzi condition*) if and only if there exists a sequence $(\beta_N)_N$, $\beta_N > 0$ for all N, such that

$$\inf_{\mu_N \in W_N} \sup_{v_N \in V_N} \frac{b(v_N, \mu_N)}{\|v_N\|_V \|\mu_N\|_W} \geqslant \beta_N \quad \text{for all } N. \tag{3.23}$$

Note that as b is bounded, β_N stays bounded.

The next result is a generalisation of [Has81, Theorem 6] to the case of non-uniform discrete inf-sup conditions.

Theorem 3.7 (A priori error estimate for the abstract primal-dual formulation)**.** *Assume that b satisfies a non-uniform discrete inf-sup condition on $(V_N, W_N)_N$ with constants $(\beta_N)_N$, that $\mathcal{K}_N = V_N \subseteq \mathcal{K} = V$, and that Λ and $(\Lambda_N)_N$ are uniformly bounded.*

Then, there exists $C > 0$ such that for all $\mu \in \Lambda$, $\mu_N \in \Lambda_N$ and $v_N \in V_N$,

$$\begin{aligned}\|u - u_N\|_V^2 \leqslant C\Big[& b(u, \lambda_N - \mu) - G(\lambda_N - \mu) + b(u, \lambda - \mu_N) - G(\lambda - \mu_N) \\ &+ \beta_N^{-2}\left(\|u - v_N\|_V^2 + \|\lambda - \mu_N\|_W^2\right)\Big]\end{aligned} \tag{3.24}$$

and

$$\|\lambda - \lambda_N\|_W \leqslant C\beta_N^{-1}\left(\|u - u_N\|_V + \|\lambda - \mu_N\|_W\right). \tag{3.25}$$

Proof. Define $\mathfrak{V} := V \times W$, endowed with the norm

$$\|(v,\mu)\|_{\mathfrak{V}} := \left(\|v\|_V^2 + \|\mu\|_W^2\right)^{1/2}, \tag{3.26}$$

$\mathfrak{K} := V \times \Lambda$ and $\mathfrak{K}_N := V_N \times \Lambda_N$, the bilinear form $\mathfrak{a}\colon \mathfrak{V} \times \mathfrak{V} \to \mathbb{R}$ by

$$\mathfrak{a}((v,\mu),(w,\eta)) := a(v,w) + b(w,\mu) - b(v,\eta), \tag{3.27}$$

and the linear form $\mathfrak{L}\colon \mathfrak{V} \to \mathbb{R}$

$$\mathfrak{L}((v,\mu)) := F(v) - G(\mu). \tag{3.28}$$

It is readily seen that

$$\mathfrak{a}((v,\mu),(v,\mu)) = a(v,v) \quad \text{for all } v \in V \text{ and } \mu \in W \tag{3.29}$$

and that $\mathfrak{a}$ and $\mathfrak{L}$ are bounded on $\mathfrak{V}$. Furthermore, due to (3.1) and (3.3), $\mathfrak{u} := (u,\lambda)$ and $\mathfrak{u}_N := (u_N, \lambda_N)$ satisfy

$$\mathfrak{a}(\mathfrak{u}, \mathfrak{v} - \mathfrak{u}) \geqslant \mathfrak{L}(\mathfrak{v} - \mathfrak{u}) \quad \text{for all } \mathfrak{v} \in \mathfrak{K} \tag{3.30}$$

and

$$\mathfrak{a}(\mathfrak{u}_N, \mathfrak{v}_N - \mathfrak{u}_N) \geqslant \mathfrak{L}(\mathfrak{v}_N - \mathfrak{u}_N) \quad \text{for all } \mathfrak{v}_N \in \mathfrak{K}_N, \tag{3.31}$$

respectively. Thus, by the coercivity of a, there exists a constant $C > 0$ such that for all $\mathfrak{v} = (v,\mu) \in \mathfrak{K}$ and $\mathfrak{v}_N = (v_N, \mu_N) \in \mathfrak{K}_N$,

$$\begin{aligned}
\|u - u_N\|_V^2 &\leqslant C\mathfrak{a}(\mathfrak{u} - \mathfrak{u}_N, \mathfrak{u} - \mathfrak{u}_N) \\
&\leqslant C\left(\mathfrak{a}(\mathfrak{u},\mathfrak{u}) - \mathfrak{a}(\mathfrak{u},\mathfrak{u}_N) - \mathfrak{a}(\mathfrak{u}_N,\mathfrak{u}) + \mathfrak{a}(\mathfrak{u}_N,\mathfrak{u}_N)\right) \\
&\leqslant C\big(\mathfrak{L}(\mathfrak{u} - \mathfrak{v}) + \mathfrak{a}(\mathfrak{u},\mathfrak{v}) + \mathfrak{L}(\mathfrak{u}_N - \mathfrak{v}_N) + \mathfrak{a}(\mathfrak{u}_N,\mathfrak{v}_N) \\
&\qquad - \mathfrak{a}(\mathfrak{u},\mathfrak{u}_N) - \mathfrak{a}(\mathfrak{u}_N,\mathfrak{u})\big) \\
&= C\big(\mathfrak{L}(\mathfrak{u} - \mathfrak{v}_N) + \mathfrak{L}(\mathfrak{u}_N - \mathfrak{v}) + \mathfrak{a}(\mathfrak{u},\mathfrak{v} - \mathfrak{u}_N) \\
&\qquad + \mathfrak{a}(\mathfrak{u}_N - \mathfrak{u}, \mathfrak{v}_N - \mathfrak{u}) + \mathfrak{a}(\mathfrak{u}, \mathfrak{v}_N - \mathfrak{u})\big).
\end{aligned} \tag{3.32}$$

By the boundedness of $\mathfrak{a}$, we obtain the existence of a constant $C > 0$ such that for all $\varepsilon > 0$,

$$\begin{aligned}
\|u - u_N\|_V^2 &\leqslant C\big(\mathfrak{L}(\mathfrak{u} - \mathfrak{v}_N) + \mathfrak{L}(\mathfrak{u}_N - \mathfrak{v}) + \mathfrak{a}(\mathfrak{u},\mathfrak{v} - \mathfrak{u}_N) \\
&\qquad + \frac{C}{2}\varepsilon \|\mathfrak{u}_N - \mathfrak{u}\|_{\mathfrak{V}}^2 + \frac{C}{2}\varepsilon^{-1} \|\mathfrak{v}_N - \mathfrak{u}\|_{\mathfrak{V}}^2 + \mathfrak{a}(\mathfrak{u}, \mathfrak{v}_N - \mathfrak{u})\big) \\
&= C\big(\mathfrak{L}(\mathfrak{u} - \mathfrak{v}_N) + \mathfrak{L}(\mathfrak{u}_N - \mathfrak{v}) + \mathfrak{a}(\mathfrak{u},\mathfrak{v} - \mathfrak{u}_N) \\
&\qquad + \frac{C}{2}\varepsilon \|u_N - u\|_V^2 + \frac{C}{2}\varepsilon \|\lambda_N - \lambda\|_W^2 \\
&\qquad + \frac{C}{2}\varepsilon^{-1} \|\mathfrak{v}_N - \mathfrak{u}\|_{\mathfrak{V}}^2 + \mathfrak{a}(\mathfrak{u}, \mathfrak{v}_N - \mathfrak{u})\big).
\end{aligned} \tag{3.33}$$

Due to the discrete inf-sup condition, we obtain that for arbitrary $\mu_N \in \Lambda_N$, there exists $v_N \in V_N$ with

$$\|\mu_N - \lambda_N\|_W \leqslant \frac{1}{2\beta_N} \frac{b(v_N, \mu_N - \lambda_N)}{\|v_N\|_V}, \tag{3.34}$$

and from (3.3a') and (3.1a') and the boundedness of a and b, we see that there exists a constant $C > 0$ such that for all $v \in V$,

$$\begin{aligned} b(v_N, \mu_N - \lambda_N) &= b(v_N, \mu_N) - b(v_N, \lambda_N) \\ &= b(v_N, \mu_N) + a(u_N, v_N) - F(v_N) \\ &= b(v_N, \mu_N - \lambda) + a(u_N - u, v_N) \\ &\leqslant C \|v_N\|_V (\|\mu_N - \lambda\|_W + \|u_N - u\|_V). \end{aligned} \tag{3.35}$$

Thus, by the triangle inequality, there exists $C > 0$ such that

$$\|\lambda - \lambda_N\|_W \leqslant \|\lambda - \mu_N\|_W + \|\mu_N - \lambda_N\|_W \leqslant C\beta_N^{-1}(\|\lambda - \mu_N\|_W + \|u - u_N\|_V). \tag{3.36}$$

Plugging this into (3.33), we see that there exists a constant $C > 0$ such that

$$\begin{aligned} \|u - u_N\|_V^2 \leqslant C\big(&\mathfrak{a}(\mathfrak{u}, \mathfrak{v}_N - \mathfrak{u}) - \mathfrak{L}(\mathfrak{v}_N - \mathfrak{u}) + \mathfrak{a}(\mathfrak{u}, \mathfrak{v} - \mathfrak{u}_N) - \mathfrak{L}(\mathfrak{v} - \mathfrak{u}_N) \\ &+ \varepsilon\beta_N^{-2} \|u - u_N\|_V + \varepsilon\beta_N^{-2} \|\lambda - \mu_N\|_W + \varepsilon^{-1} \|\mathfrak{u} - \mathfrak{v}_N\|_{\mathfrak{V}}\big). \end{aligned} \tag{3.37}$$

Set $\varepsilon := (2C)^{-1}\beta_N^2$. Choosing $v = u_N$ and applying again (3.1a'), the result follows. □

3.2 Mixed hp-Finite Elements for Elasticity with Tresca Friction

We now want to apply the theory developed in the last section to the problem with frictional contact introduced in Chapter 2. We shall restrict ourselves to the two-dimensional situation.

Recall that $\Omega \subseteq \mathbb{R}^2$ is a domain, Γ_D, Γ_N and Γ_C partition the boundary $\Gamma := \partial\Omega$, and that the continuous spaces are given by

$$V = \left\{\mathbf{v} \in \mathrm{H}^1(\Omega)^2 \colon \gamma_{0,\Gamma_\mathrm{D}}(\mathbf{v}) = 0\right\} \tag{3.38}$$

where we assume for simplicity that the Dirichlet data vanishes, $\mathbf{u}_0 = 0$ on Γ_D, and

$$W = \mathrm{L}^2(\Gamma_\mathrm{C})^2, \tag{3.39}$$

and that the closed, convex sets of admissible functions are

$$\mathcal{K} = \{\mathbf{v} \in V \colon \gamma_{0,\Gamma_\mathrm{C}}(\mathbf{v}) \cdot \boldsymbol{\nu} \leqslant u_0 \text{ almost everywhere on } \Gamma_\mathrm{C}\} \tag{3.40}$$

and

$$\Lambda = \left\{\boldsymbol{\mu} \in W \colon \|\boldsymbol{\mu}\|_{\mathrm{L}^\infty(\Gamma_\mathrm{C})} \leqslant 1 \text{ and } \mu_n = 0 \text{ almost everywhere on } \Gamma_\mathrm{C}\right\}. \tag{3.41}$$

We define the mappings $a \colon V \times V \to \mathbb{R}$, $L \colon V \to \mathbb{R}$ and $b \colon V \times W \to \mathbb{R}$ directly for the two-dimensional situation as

$$a(\mathbf{v}, \mathbf{w}) := \int_\Omega \boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{w}) \mathrm{d}\mathbf{x}, \tag{3.42}$$

$$L(\mathbf{v}) := \int_\Omega \mathbf{F} \cdot \mathbf{v} \mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \mathbf{G} \cdot \gamma_{0,\Gamma_\mathrm{N}} \mathbf{v} \mathrm{d}s_\mathbf{x}, \tag{3.43}$$

and

$$b(\mathbf{v},\boldsymbol{\mu}) := \int_{\Gamma_C} g\mathbf{v}_t \cdot \boldsymbol{\mu} \mathrm{d}s_{\mathbf{x}}. \tag{3.44}$$

For simplicity, we shall assume that g is constant on Γ_C.

The following definitions are based on the formulation given in [Sch98, pp. 169].

Definition 3.8. 1. A finite set $\mathcal{T}$ is called a *partition of* Ω *into quadrilaterals* if and only if

a) $\overline{\Omega} = \bigcup_{K\in\mathcal{T}} K$,

b) every $K \in \mathcal{T}$ is a closed, convex quadrilateral,

c) for all K, $K' \in \mathcal{T}$, $\operatorname{int} K \cap \operatorname{int} K' = \emptyset$.

2. Let $\mathcal{T}$ be a partition of Ω into quadrilaterals, $K \in \mathcal{T}$, and $\mathbf{x}$ a vertex of K.

   Then, $\mathbf{x}$ is called an *unconstrained* or *regular node* if and only if it is a vertex of all elements $K' \in \mathcal{T}$ with $\mathbf{x} \in K'$. Otherwise, $\mathbf{x}$ is called a *constrained* or *irregular node*.

3. Let $\mathcal{T}$ be a partition of Ω into quadrilaterals.

   $\mathcal{T}$ is said to satisfy the *one hanging node rule* if and only if for every element K and every edge E of K, there is at most one constrained node $\mathbf{x}$ with $\mathbf{x} \in \operatorname{int} E$.

4. Let $\mathcal{T}$ be a partition of Ω into quadrilaterals satisfying the one hanging node rule.

   A constrained node $\mathbf{x}$ is called *singly-constrained* if and only if there exists an element K and an edge E of K such that $\mathbf{x} \in \operatorname{int} E$, and the two vertices of E are unconstrained. Otherwise, $\mathbf{x}$ is called *multiply constrained*.

5. Let $\mathcal{T}$ be a partition of Ω into quadrilaterals satisfying the one hanging node rule.

   Then, $\mathcal{T}$ is said to be a 1-*irregular partition* if and only if all nodes $\mathbf{x}$ of $\mathcal{T}$ are either unconstrained or singly-constrained.

6. Let $(\mathcal{T}_N)_N$ be a sequence of partitions of Ω into quadrilaterals.

   $(\mathcal{T}_N)_N$ is said to be *shape-regular with constant* κ if and only if

$$\sup_N \sup_{K\in\mathcal{T}_N} \frac{h_K}{\rho_K} \leqslant \kappa < \infty, \tag{3.45}$$

   where h_K is the length of the longest edge of K and ρ_K is the diameter of the largest circle lying entirely in K.

7. Let $(\mathcal{T}_N)_N$ be a sequence of partitions of Ω into quadrilaterals.

   We say that $(\mathcal{T}_N)_N$ is *regular in the sense of Ciarlet* (see [Cia78, Exercise 4.3.9] if and only if there exist constants $\sigma' > 0$ and $\gamma < 1$ such that

$$\sup_N \sup_{K\in\mathcal{T}_N} \frac{h_K}{h'_K} \leqslant \sigma' \quad \text{and} \quad \gamma_K \leqslant \gamma, \tag{3.46}$$

   where h'_K is the length of the shortest side of K and $\gamma_K := \max_{i=1,2,3,4} |\cos(\beta_i)|$, where β_i are the inner angles of K.

Note that the regularity in the sense of Ciarlet prevents the quadrilaterals from degenerating to triangles and implies that the family of partitions is shape-regular (see [Cia78, Exercise 4.3.9(iii)]).

Clearly, for a partition $\mathcal{T}$ of Ω into quadrilaterals to be possible, Ω has to be a polygon. We shall assume this, and furthermore, we assume that the partition ensures that for each element $K \in \mathcal{T}$ and each edge E of K, either $\operatorname{int} E \subseteq \Omega$ or $\operatorname{int} E \subseteq \Gamma_{\mathrm{D}}$, $\operatorname{int} E \subseteq \Gamma_{\mathrm{N}}$ or $\operatorname{int} E \subseteq \Gamma_{\mathrm{C}}$; that is, the boundary conditions are completely resolved by $\mathcal{T}$. In particular, this ensures that Γ_{C} is piecewise affine, and we can define Gauss quadrature on edges on Γ_{C}.

From now on, let $(\mathcal{T}_N)_N$ be a sequence of 1-irregular partitions of Ω, regular in the sense of Ciarlet, and for every N, let $(p_{N,K})_{K\in\mathcal{T}_N}$ be a vector of polynomial degrees; that is, $p_{N,K} \in \mathbb{N}$ with $p_{N,K} \geqslant 2$ for all $K \in \mathcal{T}_N$.

It is proved in [Mel05, Lemma 2.3] that for such meshes, there exists a constant $C > 0$ with $h_K/h_{K'} \leqslant C$ for all K, $K' \in \mathcal{T}_N$ with $K \cap K' \neq \emptyset$, that is, the elements of $\mathcal{T}_N$ are *locally of comparable size*. Similarly, we shall require that the elements of $\mathcal{T}_N$ are *locally of comparable polynomial degree*, that is, there exists $C > 0$ with $p_{N,K}/p_{N,K'} \leqslant C$ for all K, $K' \in \mathcal{T}_N$ with $K \cap K' \neq \emptyset$.

We shall also assume that $\mathcal{T}_{N'}$ is a refinement of $\mathcal{T}_N$ for $N' > N$, that is, for all $K' \in \mathcal{T}_{N'}$ there is some $K \in \mathcal{T}_N$ such that $K' \subseteq K$, and for $K \in \mathcal{T}_N$, $K' \in \mathcal{T}_{N'}$ with $K' \subseteq K$, $p_{N,K} \leqslant p_{N',K'}$. Furthermore, set

$$\mathcal{E}_{\mathrm{I},N} := \left\{\partial K \cap \partial K' \colon K,\, K' \in \mathcal{T}_N,\, K \neq K'\right\}, \tag{3.47}$$

$$\mathcal{E}_{\mathrm{D},N} := \{\Gamma_{\mathrm{D}} \cap \partial K \colon K \in \mathcal{T}_N \text{ with } \Gamma_{\mathrm{C}} \cap \partial K \neq \emptyset\}, \tag{3.48}$$

$$\mathcal{E}_{\mathrm{N},N} := \{\Gamma_{\mathrm{N}} \cap \partial K \colon K \in \mathcal{T}_N \text{ with } \Gamma_{\mathrm{C}} \cap \partial K \neq \emptyset\}, \tag{3.49}$$

$$\mathcal{E}_{\mathrm{C},N} := \{\Gamma_{\mathrm{C}} \cap \partial K \colon K \in \mathcal{T}_N \text{ with } \Gamma_{\mathrm{C}} \cap \partial K \neq \emptyset\}, \tag{3.50}$$

and for every $E \in \mathcal{E}_{\mathrm{C},N}$, let K_E be the (unique) element with $E = \Gamma_{\mathrm{C}} \cap \partial K_E$. The set of all edges is denoted by $\mathcal{E}_N := \mathcal{E}_{\mathrm{I},N} \cup \mathcal{E}_{\mathrm{D},N} \cup \mathcal{E}_{\mathrm{N},N} \cup \mathcal{E}_{\mathrm{C},N}$. To be able to apply certain kinds of *hp*-Clément operators, we require the assumption [Mel05, (M4)], which basically says that all elements adjacent to a node can be together mapped to a reference patch, requiring only affine maps in between.

The *element maps* are defined in the following way: Let $K \in \mathcal{T}_N$ be the convex hull of its vertices $(\mathbf{N}_i)_{i=1}^4$, $S := [-1,1]^2$ the reference square, and $\psi_i \colon S \to \mathbb{R}$ the hat functions on S, that is,

$$\psi_1(t_1,t_2) := \frac{1}{4}(1-t_1)(1-t_2), \tag{3.51}$$

$$\psi_2(t_1,t_2) := \frac{1}{4}(1+t_1)(1-t_2), \tag{3.52}$$

$$\psi_3(t_1,t_2) := \frac{1}{4}(1+t_1)(1+t_2), \tag{3.53}$$

$$\psi_4(t_1,t_2) := \frac{1}{4}(1-t_1)(1+t_2). \tag{3.54}$$

Then, the *element map* $F_K \colon S \to K$ is given by $F_K(\mathbf{t}) := \sum_{i=1}^4 \mathbf{N}_i\psi_i(\mathbf{t})$. The regularity in the sense of Ciarlet yields that F_K is one-to-one and onto, and that there exists a constant $C > 0$ such that for all N and all $K \in \mathcal{T}_N$, for the Frobenius norms of the Jacobians DF_K and DF_K^{-1} of the element map and its inverse,

$$C^{-1}h_K \leqslant |DF_K| \leqslant Ch_k \tag{3.55}$$

and

$$C^{-1}h_K^{-1} \leqslant \left|DF_K^{-1}\right| \leqslant Ch_k^{-1}. \tag{3.56}$$

This ensures that all scaling arguments work as expected.

We consider the following *hp*-approximations: For $\mathcal{Q}^q := \left\{\sum_{i,j=0}^{q} \alpha_{ij} x_1^i x_2^j \colon \alpha_{ij} \in \mathbb{R}\right\}$ the space of tensor product polynomials, we let

$$V_N := \left\{\mathbf{v}_N \in V \colon v_N|_K \circ F_K \in (\mathcal{Q}^{p_{N,K}})^2 \text{ for all } K \in \mathcal{T}_N\right\} \tag{3.57}$$

and

$$W_N := \left\{\boldsymbol{\mu}_N \in W \colon \mu_N|_E \in \left(\mathcal{P}^{p_{N,K_E}-2}\right)^2 \text{ for all } E \in \mathcal{E}_{\mathrm{C},N}\right\}. \tag{3.58}$$

To be able to define discretisations of $\mathcal{K}$ and Λ appropriate for an *hp*-context, we recapitulate the main properties of Gauss and Gauss-Lobatto quadrature. The basic properties can be found in books on numerical analysis such as [SB93, QSS07]. The interpolation error estimates in fractional order Sobolev spaces are taken from [BM92, Theorem 3.4, Theorem 4.2].

Theorem 3.9 (Gauss quadrature)**.** *For every $q \in \mathbb{N}$, there exists a unique set $\mathrm{G}_q^I \subseteq (-1,1)$ of cardinality $q+1$ and, for every $x \in \mathrm{G}_q^I$, a corresponding weight $w_{I,x}^{\mathrm{G},q}$, such that the induced quadrature formula is of exactness $2q+1$, that is,*

$$\int_{-1}^{1} v(x)\mathrm{d}x = \sum_{x \in \mathrm{G}_q^I} v(x) w_{I,x}^{\mathrm{G},q} \quad \textit{for all } v \in \mathcal{P}^{2q+1}. \tag{3.59}$$

Furthermore,

1. *all weights are positive, $w_{I,x}^{\mathrm{G},q} > 0$ for all $x \in \mathrm{G}_q^I$, and*
2. *letting $j_q \colon \mathrm{C}^0([-1,1]) \to \mathcal{P}^q$ denote the interpolation operator at G_q^I, we have that for all $\varepsilon > 0$, there exists a constant $C > 0$ such that*

$$\|v - j_q v\|_{\mathrm{L}^2(-1,1)} \leqslant Cq^{-(1/2+\varepsilon)} \|v\|_{\mathrm{H}^{1/2+\varepsilon}(-1,1)} \quad \textit{for all } v \in \mathrm{H}^{1/2+\varepsilon}(-1,1). \tag{3.60}$$

Theorem 3.10 (Gauss-Lobatto quadrature)**.** *For every $q \in \mathbb{N}$, there exists a unique set $\mathrm{GL}_q^I \subseteq [-1,1]$ with -1, $+1 \in \mathrm{GL}_q^I$ of cardinality $q+1$ and, for every $x \in \mathrm{GL}_q^I$, a corresponding weight $w_{I,x}^{\mathrm{GL},q}$, such that the induced quadrature formula is of exactness $2q-1$, that is,*

$$\int_{-1}^{1} v(x)\mathrm{d}x = \sum_{x \in \mathrm{GL}_q^I} v(x) w_{I,x}^{\mathrm{GL},q} \quad \textit{for all } v \in \mathcal{P}^{2q-1}. \tag{3.61}$$

Furthermore,

1. *all weights are positive, $w_{I,x}^{\mathrm{GL},q} > 0$ for all $x \in \mathrm{GL}_q^I$, and*
2. *letting $i_q \colon \mathrm{C}^0([-1,1]) \to \mathcal{P}^q$ denote the interpolation operator at GL_q^I, we have that for all $\varepsilon > 0$, there exists a constant $C > 0$ such that*

$$\|v - i_q v\|_{\mathrm{L}^2(-1,1)} \leqslant Cq^{-(1/2+\varepsilon)} \|v\|_{\mathrm{H}^{1/2+\varepsilon}(-1,1)} \quad \textit{for all } v \in \mathrm{H}^{1/2+\varepsilon}(-1,1). \tag{3.62}$$

Let $E \subseteq \mathbb{R}^2$ be an arbitrary segment. Then, we can easily define Gauss and Gauss-Lobatto points and weights on E by G_q^E with $w_{E,\mathbf{x}}^{\mathrm{G},q}$ for $\mathbf{x} \in \mathrm{G}_q^E$ and GL_q^E with $w_{E,\mathbf{x}}^{\mathrm{GL},q}$ for $\mathbf{x} \in \mathrm{GL}_q^E$, respectively, by using an affine, one-to-one and onto mapping $F_E \colon [-1,1] \to E$ and setting

$$\mathrm{G}_q^E := \left\{F_E(x) \colon x \in \mathrm{G}_q^I\right\}, \tag{3.63}$$

$$w_{E,\mathbf{x}}^{\mathrm{G},q} := \frac{|E|}{2} w_{I,x}^{\mathrm{G},q} \quad \text{for } \mathbf{x} = F_E(x) \text{ with } x \in \mathrm{G}_q^E, \tag{3.64}$$

$$\mathrm{GL}_q^E := \left\{F_E(x) \colon x \in \mathrm{GL}_q^I\right\}, \tag{3.65}$$

$$w_{E,\mathbf{x}}^{\mathrm{GL},q} := \frac{|E|}{2} w_{I,x}^{\mathrm{GL},q} \quad \text{for } \mathbf{x} = F_E(x) \text{ with } x \in \mathrm{GL}_q^E. \tag{3.66}$$

We note that these formulas yield the same exactness when integrating polynomials as the corresponding formulas on I. In the same way, we define the Gauss and Gauss-Lobatto interpolation operators $j_q^E \colon \mathrm{C}^0(E) \to \mathcal{P}^q$ and $i_q^E \colon \mathrm{C}^0(E) \to \mathcal{P}^q$.

Note that, if $v \in \mathrm{C}^0(\Gamma_\mathrm{C})$ and we do a piecewise interpolation with respect to the Gauss-Lobatto points on the edges $E \in \mathcal{E}_{\mathrm{C},N}$, the resulting piecewise polynomial is also a continuous function, as the boundary points of the edges $E \in \mathcal{E}_{\mathrm{C},N}$ are always contained in the points where we do the interpolation.

The discretisation of $\mathcal{K}$ and Λ is done by setting

$$\mathcal{K}_N := \left\{\mathbf{v}_N \in V_N \colon \mathbf{v}_N(\mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x}) \leqslant u_0(\mathbf{x}) \text{ for all } \mathbf{x} \in \mathrm{GL}_{p_{N,K_E}} \text{ and all } E \in \mathcal{E}_{\mathrm{C},N}\right\}, \tag{3.67}$$

and

$$\Lambda_N := \Big\{\boldsymbol{\mu}_N \in W_N \colon |\boldsymbol{\mu}_N(\mathbf{x})| \leqslant 1 \text{ for all } \mathbf{x} \in \mathrm{G}_{p_{N,K_E}-2} \text{ and all } E \in \mathcal{E}_{\mathrm{C},N}, \\ \text{and } (\mu_N)_n = 0\Big\}. \tag{3.68}$$

Furthermore, we introduce the local mesh width $h_N \colon \Omega \to (0,\infty)$ and local polynomial degree $p_N \colon \Omega \to \mathbb{N}$ and $q_N \colon \Gamma_\mathrm{C} \to \mathbb{N}$ as

$$h_N(\mathbf{x}) := \sup_{\substack{K \in \mathcal{T}_N: \\ \mathbf{x} \in K}} h_K, \tag{3.69}$$

$$p_N(\mathbf{x}) := \sup_{\substack{K \in \mathcal{T}_N: \\ \mathbf{x} \in K}} p_{N,K}, \tag{3.70}$$

and

$$q_N(\mathbf{x}) := \sup_{\substack{E \in \mathcal{E}_{\mathrm{C},N}: \\ \mathbf{x} \in E}} (p_{N,K_E} - 2). \tag{3.71}$$

The discrete problem is:

Problem 3.11 (Discrete primal-dual formulation, Tresca friction)**.** Find $(\mathbf{u}_N, \boldsymbol{\lambda}_N) \in \mathcal{K}_N \times \Lambda_N$ such that for all $(\mathbf{v}_N, \boldsymbol{\mu}_N) \in \mathcal{K}_N \times \Lambda_N$,

$$a(\mathbf{u}_N, \mathbf{v}_N - \mathbf{u}_N) + b(\mathbf{v}_N - \mathbf{u}_N, \boldsymbol{\lambda}_N) \geqslant L(\mathbf{v}_N - \mathbf{u}_N), \tag{3.72a}$$

$$b(\mathbf{u}_N, \boldsymbol{\mu}_N - \boldsymbol{\lambda}_N) \leqslant 0. \tag{3.72b}$$

The following theorem, which is given in [Mel05, Theorem 3.3], ensures the existence of *hp*-interpolation operators of Clément type. Here, $K_{\text{patch}} = \bigcup K'$, where the union is taken over all elements K' which are near to K; for details, see the article cited above. In particular, K_{patch} is chosen large enough such that $E_{\text{patch}} \subseteq K_{\text{patch}}$ for all edges $E \subseteq \partial K$, where $E_{\text{patch}} := \bigcup_{\partial K' \cap E \neq \varnothing} K'$.

Theorem 3.12 (*hp*-Clément interpolant)**.** *There exists a sequence $(i_N)_N$ of linear operators $i_N\colon V \to V_N$ such that for all $K \in \mathcal{T}_N$ and edges $E \subseteq \partial K$,*

$$\begin{aligned} \|v - i_N v\|_{\mathrm{L}^2(K)} + \frac{h_K}{p_K} \|\nabla i_N v\|_{\mathrm{L}^2(K)} + \sqrt{\frac{h_K}{p_K}} \|v - i_N v\|_{\mathrm{L}^2(E)} \\ \leqslant C \frac{h_K}{p_K} \|\nabla v\|_{\mathrm{L}^2(K_{patch})} \quad \text{for all } v \in V. \end{aligned} \tag{3.73}$$

We shall also need the following variant, which is given in [Mel05, Theorem 3.4].

Theorem 3.13 (*hp*-Clément interpolant preserving polynomials on Γ_{C})**.** *Set*

$$\hat{V}_N := \{\mathbf{v} \in V \colon \gamma_{0,\Gamma_{\mathrm{C}}} \mathbf{v}|_E \in \mathcal{P}^{p_{N,K_E}} \text{ for all } E \in \mathcal{E}_{\mathrm{C},N}, \text{ and } \gamma_{0,\Gamma_{\mathrm{C}}} \mathbf{v} \in \mathrm{C}^0(\Gamma_{\mathrm{C}})\}. \tag{3.74}$$

Then, there exists a sequence $(\hat{i}_N)_N$ of linear operators $\hat{i}_N\colon \hat{V}_N \to V_N$ such that for all $K \in \mathcal{T}_N$ and edges $E \subseteq \partial K$,

$$\begin{aligned} \left\|v - \hat{i}_N v\right\|_{\mathrm{L}^2(K)} + \frac{h_K}{p_K} \left\|\nabla \hat{i}_N v\right\|_{\mathrm{L}^2(K)} + \sqrt{\frac{h_K}{p_K}} \left\|v - \hat{i}_N v\right\|_{\mathrm{L}^2(E)} \\ \leqslant C \frac{h_K}{p_K} \|\nabla v\|_{\mathrm{L}^2(K_{patch})} \quad \text{for all } v \in \hat{V}_N, \end{aligned} \tag{3.75}$$

and $\gamma_{0,\Gamma_{\mathrm{C}}} \mathbf{v} = \gamma_{0,\Gamma_{\mathrm{C}}} \hat{i}_N \mathbf{v}$ for all $\mathbf{v} \in \hat{V}_N$.

3.2.1 Strong Convergence

In the following, we suppose that the function u_0 given on Γ_{C} satisfies $u_0 \in \mathrm{H}^{1/2+\varepsilon}(\Gamma_{\mathrm{C}})$. Actually, we shall see below that it would be sufficient that the piecewise interpolation polynomials of u_0 at Gauss-Lobatto nodes converges in $\mathrm{L}^1(\Gamma_{\mathrm{C}})$ to u_0.

Theorem 3.14. *Suppose that $\sup_\Omega h_N/p_N \to 0$.*

Then for the solutions $(\mathbf{u}_N, \boldsymbol{\lambda}_N)$ of the discrete primal-dual formulation given in Problem 3.11 and the solution $(\mathbf{u}, \boldsymbol{\lambda})$ of the corresponding continuous problem, we have that $\mathbf{u}_N \to \mathbf{u}$ in V and $\boldsymbol{\lambda}_N \rightharpoonup \boldsymbol{\lambda}$ in W.

We shall apply the variant of Theorem 3.4 suggested in Remark 3.5. Thus, we need to prove the convergence in the sense of Glowinski of $\mathcal{K}_N$ and Λ_N, and also that the solution $\mathbf{u}_N$ is unique.

Lemma 3.15. *We have that $\mathcal{K}_N \xrightarrow{\mathrm{Gl}} \mathcal{K}$.*

This follows similarly as in [MS05, Theorem 1], where they prove a corresponding result for a Sobolev space on the boundary, which is used in a boundary element formulation of the Signorini problem. We recall the following basic convergence result on Bernstein polynomials; see [DL93, Chapter 10] for further properties.

Theorem 3.16 (Bernstein operators)**.** *For $q \in \mathbb{N}$, define the q-th* Bernstein operator *as*

$$B_q\colon \mathrm{C}^0([0,1]) \to \mathcal{P}^q, \quad B_q f := \sum_{j=0}^{n} f(j/n) p_{n,j}, \tag{3.76}$$

where

$$p_{n,j}(x) := \binom{n}{j} x^j (1-x)^{n-j}. \tag{3.77}$$

Then, $B_q f$ converges uniformly to f on $[0,1]$ for all $f \in \mathrm{C}^0([0,1])$, that is, B_q converges to the identity operator in the strong operator topology.

Furthermore, for $q \geqslant 1$, $B_q f(0) = f(0)$, $B_q f(1) = f(1)$, and if $f \geqslant 0$, then $B_q f \geqslant 0$, that is, B_q is a positive operator*.*

Similarly as for the interpolation operators at Gauss and Gauss-Lobatto points, it is possible to transform the Bernstein operators into operators on the space $\mathrm{C}^0(E)$ for $E \subseteq \mathbb{R}^2$ an arbitrary segment. We shall denote these operators by $B_q^E\colon \mathrm{C}^0(E) \to \mathcal{P}^q$. Using these operators, we can construct approximations of continuous functions:

Theorem 3.17. *Given a function $\varphi \in \mathrm{C}^2(\Gamma_\mathrm{C})$, define $\varphi_N \in \mathrm{C}^0(\Gamma_\mathrm{C})$ piecewise by $\varphi_N|_E = B^E_{p_{N,K_E}-1}\varphi|_E$.*

Then, for $N \to \infty$, φ_N converges uniformly to φ.

Proof. The continuity of φ_N follows from the fact that $B_q f(0) = f(0)$ and $B_q f(1) = f(1)$ for all $q \geqslant 1$.

Recall that $\lim_{N\to\infty} \sup_{\mathbf{x}\in\Gamma_\mathrm{C}} h_N(\mathbf{x}) p_N(\mathbf{x})^{-1} = 0$. By [DL93, p. 308, (3.4)], we have that

$$\|f - B_q f\|_{\mathrm{C}^0([0,1])} \leqslant C(q+1)^{-1} \left\|f''\right\|_{\mathrm{C}^0([0,1])}, \quad f \in \mathrm{C}^2([0,1]). \tag{3.78}$$

Thus, for every N and every $E \in \mathcal{E}_{\mathrm{C},N}$, with an affine, one-to-one and onto mapping $F_E\colon [0,1] \to E$,

$$\begin{aligned} \|\varphi|_E - \varphi_N|_E\|_{\mathrm{C}^0(E)} &= \|\varphi|_E \circ F_E - \varphi_N|_E \circ F_E\|_{\mathrm{C}^0([0,1])} \\ &= \left\|\varphi|_E \circ F_E - B_{p_{N,K_E}-1}(\varphi|_E \circ F_E)\right\|_{\mathrm{C}^0([0,1])} \\ &\leqslant C p_{N,K_E}^{-1} \left\|(\varphi|_E \circ F_E)''\right\|_{\mathrm{C}^0([0,1])} \\ &= C h_E^2 p_{N,K_E}^{-1} \left\|(\varphi|_E)''\right\|_{\mathrm{C}^0(E)} \\ &\leqslant C h_E^2 p_{N,K_E}^{-1} \|\varphi\|_{\mathrm{C}^2(\Gamma_\mathrm{C})}. \end{aligned} \tag{3.79}$$

The regularity of the mesh yields the convergence. □

Proof of Lemma 3.15. First, let $(\mathbf{v}_N)$ be a sequence of functions with $\mathbf{v}_N \in \mathcal{K}_N$ for all N, and assume that $\mathbf{v}_N \rightharpoonup \mathbf{v}$. We need to prove that $\mathbf{v} \in \mathcal{K}$, that is, $\gamma_{0,\Gamma_\mathrm{C}}\mathbf{v} \cdot \boldsymbol{\nu} \leqslant u_0$ almost everywhere on Γ_C. By Theorem 1.1, this follows if we can show

$$\int_{\Gamma_\mathrm{C}} (\gamma_{0,\Gamma_\mathrm{C}}\mathbf{v} \cdot \boldsymbol{\nu} - u_0)\, \varphi \mathrm{d}s_\mathbf{x} \leqslant 0 \quad \text{for all } \varphi \in \mathrm{C}^2(\Gamma_\mathrm{C}) \text{ with } \varphi \geqslant 0. \tag{3.80}$$

As the mapping $\mathbf{v} \mapsto \gamma_{0,\Gamma_\mathrm{C}}\mathbf{v} \cdot \boldsymbol{\nu}$ is continuous as a mapping $V \to \mathrm{L}^2(\Gamma_\mathrm{C})$, we see that $\gamma_{0,\Gamma_\mathrm{C}}\mathbf{v}_N \cdot \boldsymbol{\nu} \rightharpoonup \gamma_{0,\Gamma_\mathrm{C}}\mathbf{v} \cdot \boldsymbol{\nu}$ in $\mathrm{L}^2(\Gamma_\mathrm{C})$.

Define, for a given $\varphi \in C^0(\Gamma_C)$, $\varphi \geqslant 0$, the function φ_N by $\varphi_N|_E = B^E_{p_{N,K_E}-1}\varphi|_E$, that is, piecewise by Bernstein polynomials, and u_{0N} by $u_{0N}|_E = i^E_{p_{N,K_E}} u_0$, that is, piecewise by interpolation at the Gauss-Lobatto points. By Theorem 3.16, we see that $\varphi_N \geqslant 0$. By the exactness of the Gauss-Lobatto quadrature given in Theorem 3.10 and the definition of $\mathcal{K}_N$, we see that for every $E \in \mathcal{E}_{C,N}$, as $u_{0N}\varphi_N \in \mathcal{P}^{2p_{N,K_E}-1}$,

$$\begin{aligned}\int_E \gamma_{0,\Gamma_C}\mathbf{v}_N \cdot \boldsymbol{\nu}\varphi_N \mathrm{d}s_{\mathbf{x}} &= \sum_{\mathbf{x}\in \mathrm{GL}^E_{p_{N,K_E}}} \mathbf{v}_N(\mathbf{x})\cdot\boldsymbol{\nu}(\mathbf{x})\varphi_N(\mathbf{x}) w_{E,\mathbf{x}}^{\mathrm{GL},p_{N,K_E}} \\ &\leqslant \sum_{\mathbf{x}\in \mathrm{GL}^E_{p_{N,K_E}}} u_0(\mathbf{x})\varphi_N(\mathbf{x}) w_{E,\mathbf{x}}^{\mathrm{GL},p_{N,K_E}} \\ &= \sum_{\mathbf{x}\in \mathrm{GL}^E_{p_{N,K_E}}} u_{0N}(\mathbf{x})\varphi_N(\mathbf{x}) w_{E,\mathbf{x}}^{\mathrm{GL},p_{N,K_E}} \\ &= \int_E u_{0N}(\mathbf{x})\varphi_N(\mathbf{x})\mathrm{d}s_{\mathbf{x}}.\end{aligned} \tag{3.81}$$

As $u_{0N} \in H^{1/2+\varepsilon}(\Gamma_C)$, we can apply the estimate in Theorem 3.10 together with a scaling argument to see that u_{0N} converges to u_0 strongly in $L^2(\Gamma_C)$. Furthermore, by Theorem 3.17, φ_N converges uniformly to φ, and thus also strongly in $L^2(\Gamma_C)$. Taking the limit $N \to \infty$, we therefore obtain that $\mathbf{v} \in \mathcal{K}$.

To show the second property, let $\mathbf{v} \in C^\infty(\overline{\Omega})^2 \cap \mathcal{K}$ be given arbitrarily, where we note that $C^\infty(\overline{\Omega})^2 \cap \mathcal{K}$ is dense in $\mathcal{K}$ by [Glo84, Chapter II, Lemma 4.2]. We see that the construction as in [DGS+98, p. 150] yielding, on Γ_C, interpolation polynomials at the Gauss-Lobatto points, produces a sequence $(\mathbf{v}_N)$ satisfying $\mathbf{v}_N \in \mathcal{K}_N$, and, as stated in the beforementioned article, $\mathbf{v}_N \to \mathbf{v}$. This proves the claimed result. □

The following measure-theoretic result is essential in showing the convergence in the sense of Glowinski of Λ_N.

Lemma 3.18. *Let (X, μ) be a finite measure space. For $1 \leqslant q \leqslant \infty$, let q' be the conjugate of q, that is, $\frac{1}{q} + \frac{1}{q'} = 1$ for $q \in (1, \infty)$, $q' = \infty$ for $q = 1$ and $q' = 1$ for $q = \infty$ Let $D \subseteq L^{q'}(X)$ be a dense subspace, and define $p\colon L^q(X) \to [0, \infty]$ by*

$$p(f) := \sup\left\{\int_X fg\mathrm{d}\mu\colon g \in D \text{ and } \|g\|_{L^1(X)} \leqslant 1\right\}. \tag{3.82}$$

Then, for $f \in L^q(X)$, if $p(f) < \infty$, it follows that $f \in L^\infty(X)$ and $p(f) = \|f\|_{L^\infty(X)}$.

Proof. We see that, if $p(f) < \infty$, the linear functional

$$\tilde{\ell}\colon D \subseteq L^1(X) \to \mathbb{R}, \quad g \mapsto \int_X fg\mathrm{d}\mu \tag{3.83}$$

is continuous. As the injection $L^{q'}(X) \hookrightarrow L^1(X)$ is continuous, we see that the space of characteristic functions is contained in the $L^1(X)$-closure of D as it is contained in $L^{q'}(X)$, and therefore, D is dense in $L^1(X)$. Thus, there exists a continuous extension $\ell\colon L^1(X) \to \mathbb{R}$, and by the definition of p, $\|\ell\|_{L^1(X)^*} = p(f)$. By the Riesz representation theorem, Theorem 1.29, we see that there exists $\tilde{f} \in L^\infty(X)$ such that

$$\ell(g) = \int_X \tilde{f}g\mathrm{d}\mu \quad \text{for all } g \in L^1(X) \tag{3.84}$$

and $\left\|\tilde{f}\right\|_{L^\infty(X)} = \|\ell\|_{L^1(X)^*} = p(f)$. It is clear that

$$\int_X \tilde{f} g \mathrm{d}\mu = \int_X f g \mathrm{d}\mu \quad \text{for all } g \in D. \tag{3.85}$$

By continuity and density, this equality extends to $L^{q'}(X)$, and we obtain that $f = \tilde{f}$ almost everywhere in X, from which it follows that $f \in L^\infty(X)$ and $\|f\|_{L^\infty(X)} = p(f)$. □

Remark 3.19. A version of Lemma 3.18 which holds for vector-valued functions f follows analogously.

Lemma 3.20. *We have that* $\Lambda_N \xrightarrow{\text{Gl}} \Lambda$.

Proof. The set $M := C^\infty(\Gamma_C)^2 \cap \Lambda$ is dense in Λ with respect to the L^2-topology, which follows from [Dör07, Lemma 4.1]. For $\boldsymbol{\mu} \in M$, we can choose $\boldsymbol{\mu}_N \in \Lambda_N$ as the interpolant of $\boldsymbol{\mu}$ at the Gauss points of every edge $E \in \mathcal{E}_{C,N}$. Then, by applying Theorem 3.9 together with a scaling argument, we see that $\boldsymbol{\mu}_N$ converges strongly to $\boldsymbol{\mu}$.

For the second property, consider a sequence $(\boldsymbol{\mu}_N)$, $\boldsymbol{\mu}_N \in \Lambda_N$, and assume that $\boldsymbol{\mu}_N \rightharpoonup \boldsymbol{\mu}$ in W. We shall prove that $\boldsymbol{\mu} \in \Lambda$. Following Remark 3.19, we only need to prove that $\int_{\Gamma_C} \boldsymbol{\mu} \cdot \boldsymbol{\eta} \mathrm{d}s_{\mathbf{x}} \leqslant 1$ for all $\boldsymbol{\eta} \in C^\infty(\Gamma_C)^2$ with $\|\boldsymbol{\eta}\|_{L^1(\Gamma_C)} = 1$. Note, furthermore, that $\boldsymbol{\mu} \cdot \boldsymbol{\nu} = 0$, and thus, we can assume that $\boldsymbol{\eta} \cdot \boldsymbol{\nu} = 0$. Choose therefore such a $\boldsymbol{\eta}$. Choose a sequence $(\boldsymbol{\eta}_N)$, $\boldsymbol{\eta}_N \in W_N$, piecewise on each $E \in \mathcal{E}_{N,N}$, componentwise as the Bernstein polynomial of order $p_{N,K_E} - 1$. By Theorem 3.17, it follows that $\boldsymbol{\eta}_N$ converges to $\boldsymbol{\eta}$ uniformly, and thus also in $L^2(\Gamma_C)$. In particular, this yields that the norms converge, $\|\boldsymbol{\eta}_N\|_{L^1(\Gamma_C)} \to \|\boldsymbol{\eta}\|_{L^1(\Gamma_C)} = 1$.

By the exactness of Gauss quadrature as given in Theorem 3.9, as $\boldsymbol{\mu}_N \cdot \boldsymbol{\eta}_N \in \mathcal{P}^{2p_{N,K_E}-3}$,

$$\begin{aligned}
\int_{\Gamma_C} \boldsymbol{\mu}_N \cdot \boldsymbol{\eta}_N \mathrm{d}s_{\mathbf{x}} &= \sum_{E \in \mathcal{E}_{C,N}} \sum_{\mathbf{x} \in G^E_{p_{N,K_E}-2}} \boldsymbol{\mu}_N(\mathbf{x}) \cdot \boldsymbol{\eta}_N(\mathbf{x}) w_{E,\mathbf{x}}^{G,p_{N,K_E}-2} \\
&\leqslant \sum_{E \in \mathcal{E}_{C,N}} \sum_{\mathbf{x} \in G^E_{p_{N,K_E}-2}} |\boldsymbol{\eta}_N(\mathbf{x})| \, w_{E,\mathbf{x}}^{G,p_{N,K_E}-2} \\
&\leqslant \sum_{E \in \mathcal{E}_{C,N}} \sum_{\mathbf{x} \in G^E_{p_{N,K_E}-2}} |\boldsymbol{\eta}_N(\mathbf{x}) - \boldsymbol{\eta}(\mathbf{x})| \, w_{E,\mathbf{x}}^{G,p_{N,K_E}-2} \\
&\quad + \sum_{E \in \mathcal{E}_{C,N}} \sum_{\mathbf{x} \in G^E_{p_{N,K_E}-2}} |\boldsymbol{\eta}(\mathbf{x})| \, w_{E,\mathbf{x}}^{G,p_{N,K_E}-2} \\
&\leqslant |\Gamma_C| \, \|\boldsymbol{\eta}_N - \boldsymbol{\eta}\|_{L^\infty(\Gamma_C)} + \sum_{E \in \mathcal{E}_{C,N}} \sum_{\mathbf{x} \in G^E_{p_{N,K_E}-2}} |\boldsymbol{\eta}(\mathbf{x})| \, w_{E,\mathbf{x}}^{G,p_{N,K_E}-2}.
\end{aligned} \tag{3.86}$$

As $\mathbf{x} \mapsto |\boldsymbol{\eta}(\mathbf{x})|$ is in $H^1(\Gamma_C)$ and $\boldsymbol{\eta}_N \to \boldsymbol{\eta}$ uniformly, we see that the above expression converges to $\|\boldsymbol{\eta}\|_{L^1(\Gamma_C)} = 1$.

On the other hand, $\boldsymbol{\mu}_N \rightharpoonup \boldsymbol{\mu}$ and $\boldsymbol{\eta}_N \to \boldsymbol{\eta}$ in $L^2(\Gamma_C)$, and finally, we obtain

$$\lim_{N \to \infty} \int_{\Gamma_C} \boldsymbol{\mu}_N \cdot \boldsymbol{\eta}_N \mathrm{d}s_{\mathbf{x}} = \int_{\Gamma_C} \boldsymbol{\mu} \cdot \boldsymbol{\eta} \mathrm{d}s_{\mathbf{x}} \leqslant 1, \tag{3.87}$$

that is, $\boldsymbol{\mu} \in \Lambda$. □

Proof of Theorem 3.14. First, we show that for $(\mathbf{u}_N, \boldsymbol{\lambda}_N)$, $\mathbf{u}_N$ is unique. Let $(\mathbf{u}'_N, \boldsymbol{\lambda}'_N)$ be another solution. Then,

$$\begin{aligned} a(\mathbf{u}_N, \mathbf{u}'_N - \mathbf{u}_N) + b(\mathbf{u}'_N - \mathbf{u}_N, \boldsymbol{\lambda}_N) &\geqslant L(\mathbf{u}'_N - \mathbf{u}_N), &(3.88)\\ a(\mathbf{u}'_N, \mathbf{u}_N - \mathbf{u}'_N) + b(\mathbf{u}_N - \mathbf{u}'_N, \boldsymbol{\lambda}'_N) &\geqslant L(\mathbf{u}_N - \mathbf{u}'_N), &(3.89)\\ b(\mathbf{u}_N, \boldsymbol{\lambda}'_N - \boldsymbol{\lambda}_N) &\leqslant 0, &(3.90) \end{aligned}$$

and

$$b(\mathbf{u}'_N, \boldsymbol{\lambda}_N - \boldsymbol{\lambda}'_N) \leqslant 0. \tag{3.91}$$

Thus, adding the respective inequalities,

$$a(\mathbf{u}_N - \mathbf{u}'_N, \mathbf{u}_N - \mathbf{u}'_N) + b(\mathbf{u}_N - \mathbf{u}'_N, \boldsymbol{\lambda}_N - \boldsymbol{\lambda}'_N) \leqslant 0 \tag{3.92}$$

and

$$b(\mathbf{u}_N - \mathbf{u}'_N, \boldsymbol{\lambda}_N - \boldsymbol{\lambda}'_N) \geqslant 0, \tag{3.93}$$

which gives that $\mathbf{u}_N = \mathbf{u}'_N$ by the coercivity of a.

Next, clearly,

$$\|\boldsymbol{\lambda}\|_{\mathrm{L}^2(\Gamma_\mathrm{C})} \leqslant |\Gamma_\mathrm{C}|^{1/2} \|\boldsymbol{\lambda}\|_{\mathrm{L}^\infty(\Gamma_\mathrm{C})} \leqslant |\Gamma_\mathrm{C}|^{1/2}, \quad \text{for all } \boldsymbol{\lambda} \in \Lambda, \tag{3.94}$$

and also

$$\begin{aligned} \|\boldsymbol{\lambda}_N\|_{\mathrm{L}^2(\Gamma_\mathrm{C})} &= \left(\sum_{E \in \mathcal{E}_{\mathrm{C},N}} \sum_{\mathbf{x} \in \mathrm{G}^E_{p_{N,K_E}-2}} |\boldsymbol{\lambda}_N(\mathbf{x})|^2 \, w_{E,\mathbf{x}}^{p_{N,K_E}-2} \right)^{1/2} \\ &\leqslant \left(\sum_{E \in \mathcal{E}_{\mathrm{C},N}} \sum_{\mathbf{x} \in \mathrm{G}^E_{p_{N,K_E}-2}} w_{E,\mathbf{x}}^{p_{N,K_E}-2} \right)^{1/2} = \left(\int_{\Gamma_\mathrm{C}} \mathrm{d}s_\mathbf{x} \right)^{1/2} \\ &= |\Gamma_\mathrm{C}|^{1/2} \quad \text{for all } \boldsymbol{\lambda}_N \in \Lambda_N, \end{aligned} \tag{3.95}$$

which gives the uniform boundedness of both Λ and Λ_N.

Applying Lemma 3.15 and Lemma 3.20 together with Remark 3.5 and Theorem 3.4, we obtain the claimed result. □

3.2.2 An *hp*-Mortar Projection Operator with Slowly Growing Bound

Our aim is to apply Theorem 3.7 to the friction problem under consideration. For this, we need to show an inf-sup condition. A usual approach to do this is outlined in [BS08, Lemma 12.5.22] and is based on the construction of sequences of operators conserving scalar products with elements of the Lagrange multiplier space, and having a calculable operator norm. This is done in this section.

Note that the operators constructed here correspond to the kind of operators used in mortar finite element methods. They can be compared to the mortar projection operators constructed in [SS00]. There, a space of continuous functions on the boundary is used as the mortar space, which leads to an operator norm growing as $\max_i p_i^{3/4}$, which was numerically shown to be sharp by Seshaiyer and Suri in the above reference. We use a discontinuous mortar space, and obtain that the operator norms are of the order $\max_i(p_i^{1/2}\log^{3/2} p_i)$, that is, they grow more slowly.

Theorem 3.21. *There exists a sequence of polynomials $(L_n)_{n\geqslant 0}$, the* Legendre polynomials*, with $L_n \in \mathcal{P}^n$ for all $n \geqslant 0$ and such that*

$$\langle L_n, L_m\rangle_{\mathrm{L}^2(-1,1)} = w_n\delta_{nm} \quad \textit{for all } n,\, m \geqslant 0, \tag{3.96}$$

where

$$w_n := \frac{2}{2n+1}. \tag{3.97}$$

Furthermore, with the above normalisation,

$$\|L_q\|_{\mathrm{H}^{1/2}(-1,1)} \leqslant C \log^{1/2} q \tag{3.98}$$

for some constant $C > 0$ independent of $q \geqslant 2$.

The first part is well known, and can be found in [Sch98, Appendix C]. The estimate for the $\mathrm{H}^{1/2}$-norm of the Legendre polynomials is contained in the proof of [AMW99, Lemma 10].

We consider an operator $P_q\colon \mathrm{H}^{1/2}(-1,1) \to \mathrm{H}^{1/2}(-1,1)$ which satisfies $P_q v \in \mathcal{P}^q$ for all $v \in \mathrm{H}^{1/2}(-1,1)$. Furthermore, we require that

$$\langle P_q v, w\rangle_{\mathrm{L}^2(-1,1)} = \langle v, w\rangle_{\mathrm{L}^2(-1,1)} \quad \text{for all } w \in \mathcal{P}^{q-2}. \tag{3.99}$$

This implies that

$$P_q v = \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{q-2}} v + \varphi_{q,q-1}(v)L_{q-1} + \varphi_{q,q}(v)L_q, \tag{3.100}$$

where $\Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^j}$ is the orthogonal projector onto $\mathcal{P}^j$ with respect to the scalar product of $\mathrm{L}^2(-1,1)$, and $\varphi_{j,k}$ are continuous linear functionals on $\mathrm{H}^{1/2}(-1,1)$. The aim is therefore to determine $\varphi_{j,k}$ in such a way that the norm of P_q does not grow too quickly, but at the same time, we obtain a globally continuous approximation if, for a given mesh, we apply P_q element by element.

Consider therefore a mesh $(x_i)_{i=0,\dots,m}$, $a = x_0 < x_1 < \cdots < x_m = b$, together with a polynomial degree distribution $(p_i)_{i=1,\dots,m}$, that is, $p_i \in \mathbb{N}$, $p_i \geqslant 2$, satisfying

$$C^{-1} \leqslant h_i/h_{i+1} \leqslant C \quad \text{for } i = 1, \dots, m-1 \tag{3.101}$$

and

$$C^{-1} \leqslant p_i/p_{i+1} \leqslant C \quad \text{for } i = 1, \dots, m-1. \tag{3.102}$$

and let $F_i\colon (-1,1) \to (x_{i-1}, x_i)$ be defined by $F_i(t) := x_{i-1} + \frac{t+1}{2}(x_i - x_{i-1})$. Define the operator $P_N\colon \mathrm{H}^{1/2}(a,b) \to \mathrm{H}^{1/2}(a,b)$ by

$$\begin{aligned}(P_N v)|_{(x_{i-1},x_i)} = \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v\circ F_i)\circ F_i^{-1}\\ + \varphi^i_{p_i,p_i-1}(v)L_{p_i-1}\circ F_i^{-1} + \varphi^i_{p_i,p_i}(v)L_{p_i}\circ F_i^{-1}.\end{aligned} \tag{3.103}$$

Let $J_N \colon \mathrm{H}^{1/2}(a,b) \to V_N$, where

$$V_N := \left\{ v_N \in \mathrm{H}^{1/2}(a,b) \colon v_N|_{(x_{i-1},x_i)} \in \mathcal{P}^{p_i} \text{ for } i = 1, \dots, m \right\}, \tag{3.104}$$

be an arbitrary, continuous operator. We define $\varphi^i_{p_i,p_i-1}$ and $\varphi^i_{p_i,p_i}$ in such a way that $P_N v(x_i) = J_N v(x_i)$ for $i = 0, \dots, m$. This is achieved by solving the linear system of equations

$$J_N v(x_{i-1}) - \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v \circ F_i)(x_{i-1}) = (-1)^{p_i-1}\varphi^i_{p_i,p_i-1}(v) + (-1)^{p_i}\varphi^i_{p_i,p_i}(v), \tag{3.105}$$

$$J_N v(x_i) - \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v \circ F_i)(x_i) = \varphi^i_{p_i,p_i-1}(v) + \varphi^i_{p_i,p_i}(v). \tag{3.106}$$

In particular, as $J_N v$ is necessarily continuous, $P_N v$ is so, as well, and $P_N v \in V_N$.

We have the following variant of a von Petersdorff inequality, which is also given in [AMT99, Theorem 4.1].

Lemma 3.22. *There exists a constant $C > 0$ independent of the mesh such that for all $v \in \mathrm{H}^{1/2}(a,b)$ such that $v|_{(x_{i-1},x_i)} \in \mathrm{H}^{1/2}_{00}(x_{i-1},x_i)$ for all $i = 1, \dots, m$,*

$$|v|^2_{\mathrm{H}^{1/2}(a,b)} \leqslant C \sum_{i=1}^m \left| v|_{(x_{i-1},x_i)} \right|^2_{\mathrm{H}^{1/2}_{00}(x_{i-1},x_i)}. \tag{3.107}$$

Conversely, for $v \in \mathrm{H}^{1/2}(a,b)$,

$$\sum_{i=1}^m \left| v|_{(x_{i-1},x_i)} \right|^2_{\mathrm{H}^{1/2}(x_{i-1},x_i)} \leqslant |v|^2_{\mathrm{H}^{1/2}(a,b)}. \tag{3.108}$$

Proof. By definition of the $\mathrm{H}^{1/2}$-seminorm (see Section 1.2),

$$\begin{aligned} |v|^2_{\mathrm{H}^{1/2}(a,b)} &= \int_{(a,b)} \int_{(a,b)} \frac{|v(x)-v(y)|^2}{|x-y|^2} \mathrm{d}x\mathrm{d}y \\ &= \sum_{i,j=1}^m \int_{(x_{i-1},x_i)} \int_{(x_{j-1},x_j)} \frac{|v(x)-v(y)|^2}{|x-y|^2} \mathrm{d}x\mathrm{d}y \\ &= \sum_{i=1}^m \int_{(x_{i-1},x_i)} \int_{(x_{i-1},x_i)} \frac{|v(x)-v(y)|^2}{|x-y|^2} \mathrm{d}x\mathrm{d}y \\ &\quad + 2 \sum_{i=1}^m \sum_{j \neq i} \int_{(x_{i-1},x_i)} \int_{(x_{j-1},x_j)} \frac{|v(x)-v(y)|^2}{|x-y|^2} \mathrm{d}x\mathrm{d}y. \end{aligned} \tag{3.109}$$

By the triangle inequality and the Fubini theorem,

$$\begin{aligned} &\sum_{i=1}^m \sum_{j \neq i} \int_{(x_{i-1},x_i)} \int_{(x_{j-1},x_j)} \frac{|v(x)-v(y)|^2}{|x-y|^2} \mathrm{d}x\mathrm{d}y \\ &\quad \leqslant 2 \sum_{i=1}^m \sum_{j \neq i} \int_{(x_{i-1},x_i)} \int_{(x_{j-1},x_j)} \frac{|v(x)|^2 + |v(y)|^2}{|x-y|^2} \mathrm{d}x\mathrm{d}y \\ &\quad \leqslant 4 \sum_{i=1}^m \sum_{j \neq i} \int_{(x_{i-1},x_i)} \int_{(x_{j-1},x_j)} \frac{|v(x)|^2}{|x-y|^2} \mathrm{d}y\mathrm{d}x \end{aligned} \tag{3.110}$$

$$= 4\sum_{i=1}^{m}\int_{(x_{i-1},x_i)}|v(x)|^2\int_{(a,b)\setminus(x_{i-1},x_i)}\frac{1}{|x-y|^2}\mathrm{d}y\mathrm{d}x.$$

As $x \in (x_{i-1}, x_i)$,

$$\begin{aligned}\int_{(a,b)\setminus(x_{i-1},x_i)}\frac{1}{|x-y|^2}\mathrm{d}y &= \int_{(a,x_{i-1})}\frac{1}{(x-y)^2}\mathrm{d}y + \int_{(x_i,b)}\frac{1}{(x-y)^2}\mathrm{d}y\\ &= \frac{1}{x-x_{i-1}} - \frac{1}{x-a} - \frac{1}{b-x} + \frac{1}{x_i-x}\\ &\leqslant \frac{1}{x-x_{i-1}} + \frac{1}{x_i-x}\\ &\leqslant \frac{2}{\mathrm{dist}(x,\{x_{i-1},x_i\})}.\end{aligned} \tag{3.111}$$

Thus,

$$\begin{aligned}|v|^2_{\mathrm{H}^{1/2}(a,b)} \leqslant{}& \sum_{i=1}^{m}\int_{(x_{i-1},x_i)}\int_{(x_{i-1},x_i)}\frac{|v(x)-v(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y\\ &+ 2\sum_{i=1}^{m}\sum_{j\neq i}\int_{(x_{i-1},x_i)}\int_{(x_{j-1},x_j)}\frac{|v(x)-v(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y\\ \leqslant{}& \sum_{i=1}^{m}\left[\left|v|_{(x_{i-1},x_i)}\right|^2_{\mathrm{H}^{1/2}(x_{i-1},x_i)} + 16\int_{(x_{i-1},x_i)}\frac{|v(x)|}{\mathrm{dist}(x,\{x_{i-1},x_i\})}\mathrm{d}x\right]\\ \leqslant{}& 16\sum_{i=1}^{m}\left[\left|v|_{(x_{i-1},x_i)}\right|^2_{\mathrm{H}^{1/2}_{00}(x_{i-1},x_i)}\right].\end{aligned} \tag{3.112}$$

Therefore, the result holds true with $C = 16$.

For the second estimate, note that

$$\begin{aligned}|v|^2_{\mathrm{H}^{1/2}(a,b)} &= \int_{(a,b)}\int_{(a,b)}\frac{|v(x)-v(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y\\ &= \sum_{i,j=1}^{m}\int_{(x_{i-1},x_i)}\int_{(x_{j-1},x_j)}\frac{|v(x)-v(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y\\ &\geqslant \sum_{i=1}^{m}\int_{(x_{i-1},x_i)}\int_{(x_{i-1},x_i)}\frac{|v(x)-v(y)|^2}{|x-y|^2}\mathrm{d}x\mathrm{d}y\\ &= \sum_{i=1}^{m}\left|v|_{(x_{i-1},x_i)}\right|^2_{\mathrm{H}^{1/2}(x_{i-1},x_i)},\end{aligned} \tag{3.113}$$

from which the result follows. □

Lemma 3.23. *We have the estimate*

$$\begin{aligned}\|P_N v\|_{\mathrm{H}^{1/2}(a,b)} \leqslant{}& \|J_N v\|_{\mathrm{H}^{1/2}(a,b)}\\ &+ C\max_{i=1,\ldots,m}\log^{3/2}p_i\left(\sum_{i=1}^{m}p_i^2\left\|\Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v\circ F_i) - J_N v\circ F_i\right\|^2_{\mathrm{L}^2(-1,1)}\right)^{1/2}.\end{aligned} \tag{3.114}$$

Proof. First, note that

$$\|P_N v\|_{\mathrm{H}^{1/2}(a,b)} \leqslant \|J_N v\|_{\mathrm{H}^{1/2}(a,b)} + \|P_N v - J_N v\|_{\mathrm{H}^{1/2}(a,b)} . \tag{3.115}$$

As $P_N v(x_i) = J_N v(x_i)$, we can apply Lemma 3.22 together with a scaling argument to obtain

$$\begin{aligned}
\|P_N v - J_N v\|^2_{\mathrm{H}^{1/2}(a,b)} &= \|P_N v - J_N v\|^2_{\mathrm{L}^2(a,b)} + |P_N v - J_N v|^2_{\mathrm{H}^{1/2}(a,b)} \\
&\leqslant C \sum_{i=1}^{m} \Big[\|P_N v - J_N v\|^2_{\mathrm{L}^2(x_{i-1},x_i)} \\
&\qquad + |P_N v - J_N v|^2_{\mathrm{H}^{1/2}_{00}(x_{i-1},x_i)} \Big] \\
&= C \sum_{i=1}^{m} \Big[(x_i - x_{i-1}) \|P_N v \circ F_i - J_N v \circ F_i\|^2_{\mathrm{L}^2(-1,1)} \\
&\qquad + |P_N v \circ F_i - J_N v \circ F_i|^2_{\mathrm{H}^{1/2}_{00}(-1,1)} \Big] \\
&\leqslant C(b-a) \sum_{i=1}^{m} \|P_N v \circ F_i - J_N v \circ F_i\|^2_{\mathrm{H}^{1/2}_{00}(-1,1)} .
\end{aligned} \tag{3.116}$$

Theorem 1.19 yields the existence of a constant $C > 0$ such that

$$\begin{aligned}
&\|P_N v \circ F_i - J_N v \circ F_i\|_{\mathrm{H}^{1/2}_{00}(-1,1)} \\
&\quad \leqslant C \log p_i \|P_N v \circ F_i - J_N v \circ F_i\|_{\mathrm{H}^{1/2}(-1,1)} .
\end{aligned} \tag{3.117}$$

By the triangle inequality,

$$\begin{aligned}
\|P_N v \circ F_i - J_N v \circ F_i\|_{\mathrm{H}^{1/2}(-1,1)} \leqslant & \left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}} (v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{H}^{1/2}(-1,1)} \\
&+ \left|\varphi^i_{p_i,p_i-1}(v)\right| \|L_{p_i-1}\|_{\mathrm{H}^{1/2}(-1,1)} \\
&+ \left|\varphi^i_{p_i,p_i}(v)\right| \|L_{p_i}\|_{\mathrm{H}^{1/2}(-1,1)} .
\end{aligned} \tag{3.118}$$

Furthermore, Theorem 1.20 gives, with some constant $C > 0$,

$$\begin{aligned}
&\left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}} (v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{H}^{1/2}(-1,1)} \\
&\quad \leqslant C p_i \left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}} (v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} .
\end{aligned} \tag{3.119}$$

Next, it is easy to see that, by Theorem 1.17, there exists a constant $C > 0$ such that

$$\begin{aligned}
\left|\varphi^i_{p_i,q}(v)\right| &\leqslant \left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}} (v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^\infty(-1,1)} \\
&\leqslant C p_i \left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}} (v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} .
\end{aligned} \tag{3.120}$$

Together with the estimate of the $\mathrm{H}^{1/2}$-norms of the Legendre polynomials in 3.21, this yields the claimed estimate. □

Thus, the fundamental point is to construct the operator J_N in such a way that

$$\left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} \quad \text{and} \quad \| J_N v \|_{\mathrm{H}^{1/2}(a,b)} \tag{3.121}$$

can be bounded. This is done in the following result.

Theorem 3.24. *There exists a sequence (J_N) of operators $J_N \colon \mathrm{H}^{1/2}(a,b) \to V_N$ such that*

$$\begin{aligned} \left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} &\leqslant C p_i^{-1/2} \, |v|_{\mathrm{H}^{1/2}(x_{i-2},x_{i+1})} \\ &\qquad \textit{for } i = 2, \dots, m-1, \end{aligned} \tag{3.122}$$

$$\left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_1-2}}(v \circ F_1) - J_N v \circ F_1 \right\|_{\mathrm{L}^2(-1,1)} \leqslant C p_1^{-1/2} \, |v|_{\mathrm{H}^{1/2}(x_0,x_2)} \,, \tag{3.123}$$

$$\left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_m-2}}(v \circ F_m) - J_N v \circ F_m \right\|_{\mathrm{L}^2(-1,1)} \leqslant C p_m^{-1/2} \, |v|_{\mathrm{H}^{1/2}(x_{m-2},x_m)} \,, \tag{3.124}$$

and

$$\| J_N v \|_{\mathrm{H}^{1/2}(a,b)} \leqslant C \, \| v \|_{\mathrm{H}^{1/2}(a,b)} \,. \tag{3.125}$$

The construction uses certain operators giving simultaneous approximation in Sobolev spaces with different exponents.

Lemma 3.25. *There exists a sequence $(\pi_q)_{q=0}^{\infty}$ of operators $\pi_q \colon \mathrm{L}^2(-1,1) \to \mathcal{P}^q$ and a constant $C > 0$ such that for all $q \geqslant 0$,*

$$\begin{aligned} \| v - \pi_q v \|_{\mathrm{H}^r(-1,1)} &\leqslant C (q+1)^{-(s-r)} \, |v|_{\mathrm{H}^s(-1,1)} \\ &\qquad \textit{for } 0 \leqslant r \leqslant s \leqslant 1 \textit{ and } v \in \mathrm{H}^s(-1,1), \end{aligned} \tag{3.126}$$

and

$$\pi_q v = v \quad \textit{for } v \in \mathcal{P}^0. \tag{3.127}$$

In particular, for a function $v \in \mathrm{H}^s(-1,1)$, this yields a simultaneous approximation in all spaces $\mathrm{H}^r(-1,1)$ with $r < s$, and stability in $\mathrm{H}^s(-1,1)$.

Proof. Define the linear operator $\pi_q \colon \mathrm{L}^2(-1,1) \to \mathcal{P}_q$ as the operator constructed in [Mel05, Proposition A.2] for $R = 1$ and $N = q$. Then, the results follow by combining Theorems 1.15 and 1.12, where we apply Theorem 1.5 to obtain the seminorms. □

Proof of Theorem 3.24. We shall do this based on partition of unity methods; for general ideas of this method, see the articles [MB96, BM97].

Let $\psi_k \colon (a,b) \to \mathbb{R}$ be functions piecewise affine on the given mesh such that $\psi_k(x_i) = \delta_{ki}$ for i, $k = 0, \dots, m$. We remark that $\operatorname{supp} \psi_k = [x_{k-1}, x_{k+1}]$ for $k = 1, \dots, m-1$, $\operatorname{supp} \psi_0 = [x_0, x_1]$ and $\operatorname{supp} \psi_m = [x_{m-1}, x_m]$, and that on (x_{i-1}, x_i), $\psi_{i-1} + \psi_i = 1$.

Define $\tilde{p}_i := \min(p_i, p_{i+1})$ for $i = 1, \dots, m-1$, $\tilde{p}_0 := p_1$ and $\tilde{p}_m := p_m$, the mappings $\tilde{F}_k \colon (-1,1) \to (x_{k-1}, x_{k+1})$, $\tilde{F}_k(t) := x_{k-1} + \frac{t+1}{2}(x_{k+1} - x_{k-1})$ for $k = 1, \dots, m-1$, $\tilde{F}_0 := F_1$ and $\tilde{F}_m := F_m$,

$$\pi_N^i v := \begin{cases} \pi_{\tilde{p}_i - 1}(v \circ \tilde{F}_i) \circ \tilde{F}_i^{-1} & \text{on } \tilde{F}_i((-1,1)), \\ 0, & \text{elsewhere}, \end{cases} \tag{3.128}$$

and J_N by

$$J_N v := \sum_{i=0}^{m} (\pi_N^i v)\psi_i. \tag{3.129}$$

Note that by the choice of the $\tilde{p}_j$, $(J_N v)|_{(x_{i-1},x_i)}$ is a polynomial of degree not more than p_i, and by construction of the ψ_k, $J_N v$ is continuous. Thus, $J_N \colon \mathrm{H}^{1/2}(a,b) \to V_N$ is well-defined. By the triangle inequality,

$$\begin{aligned} &\left\| \Pi_{\mathcal{P}^{p_i-2}}^{\mathrm{L}^2(-1,1)}(v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant \left\| \Pi_{\mathcal{P}^{p_i-2}}^{\mathrm{L}^2(-1,1)}(v \circ F_i) - v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\quad + \left\| v \circ F_i - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)}, \end{aligned} \tag{3.130}$$

and by the properties of the orthogonal projection,

$$\begin{aligned} \left\| \Pi_{\mathcal{P}^{p_i-2}}^{\mathrm{L}^2(-1,1)}(v \circ F_i) - v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} &\leqslant \left\| \pi_{p_i-2}(v \circ F_i) - v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant C(p_i - 1)^{-1/2} \left| v \circ F_i \right|_{\mathrm{H}^{1/2}(-1,1)} \\ &\leqslant C(p_i - 1)^{-1/2} \left| v \right|_{\mathrm{H}^{1/2}(x_{i-1},x_i)}. \end{aligned} \tag{3.131}$$

As $0 \leqslant \psi_k \leqslant 1$, by scaling, there exists, due to (3.101), a constant $C > 0$ such that for $i = 2, \ldots, m-1$,

$$\begin{aligned} &\left\| v \circ F_i - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant C h_i^{-1/2} \left\| v - J_N v \right\|_{\mathrm{L}^2(x_{i-1},x_i)} \\ &= C h_i^{-1/2} \left\| v(\psi_{i-1} + \psi_i) - J_N v \right\|_{\mathrm{L}^2(x_{i-1},x_i)} \\ &\leqslant C h_i^{-1/2} \left\| (v - \pi_N^{i-1} v)\psi_{i-1} + (v - \pi_N^i v)\psi_i \right\|_{\mathrm{L}^2(x_{i-2},x_{i+1})} \\ &\leqslant C h_i^{-1/2} \Big(\left\| (v - \pi_N^{i-1} v)\psi_{i-1} \right\|_{\mathrm{L}^2(x_{i-2},x_i)} \\ &\qquad + \left\| (v - \pi_N^i v)\psi_i \right\|_{\mathrm{L}^2(x_{i-1},x_{i+1})} \Big) \\ &\leqslant C h_i^{-1/2} \Big(\left\| v - \pi_N^{i-1} v \right\|_{\mathrm{L}^2(x_{i-2},x_i)} \\ &\qquad + \left\| v - \pi_N^i v \right\|_{\mathrm{L}^2(x_{i-1},x_{i+1})} \Big) \\ &\leqslant C \Big(\left\| v \circ \tilde{F}_{i-1} - \pi_{\tilde{p}_{i-1}-1}(v \circ \tilde{F}_{i-1}) \right\|_{\mathrm{L}^2(-1,1)} \\ &\quad + \left\| v \circ \tilde{F}_i - \pi_{\tilde{p}_i-1}(v \circ \tilde{F}_i) \right\|_{\mathrm{L}^2(-1,1)} \Big). \end{aligned} \tag{3.132}$$

By the properties of π_q and the invariance of the $\mathrm{H}^{1/2}$-norm with respect to scaling,

$$\begin{aligned} \left\| v \circ \tilde{F}_i - \pi_{\tilde{p}_i-1}(v \circ \tilde{F}_i) \right\|_{\mathrm{L}^2(-1,1)} &\leqslant C \tilde{p}_i^{-1/2} \left| v \circ \tilde{F}_i \right|_{\mathrm{H}^{1/2}(-1,1)} \\ &\leqslant C \tilde{p}_i^{-1/2} \left| v \right|_{\mathrm{H}^{1/2}(x_{i-1},x_{i+1})}. \end{aligned} \tag{3.133}$$

Clearly, there exists a constant $C > 0$ such that

$$(p_i - 1)^{-1/2} \geqslant C p_i^{-1/2} \tag{3.134}$$

and

$$\min(\tilde{p}_{i-1}, \tilde{p}_i)^{-1/2} \leqslant C p_i^{-1/2}, \tag{3.135}$$

and thus

$$\left\| \Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i - 2}}(v \circ F_i) - J_N v \circ F_i \right\|_{\mathrm{L}^2(-1,1)} \leqslant C p_i^{-1/2} \left| v \right|_{\mathrm{H}^{1/2}(x_{i-2}, x_{i+1})} \tag{3.136}$$

with some $C > 0$; corresponding results hold for $i = 1$ and $i = m$.

For the stability of the operators J_N, note first that

$$\begin{aligned} \|J_N v\|_{\mathrm{H}^{1/2}(a,b)} &\leqslant \|v\|_{\mathrm{H}^{1/2}(a,b)} + \|J_N v - v\|_{\mathrm{H}^{1/2}(a,b)} \\ &\leqslant \|v\|_{\mathrm{H}^{1/2}(a,b)} + \sum_{i=0}^{m} \left\| (\pi_N^i v - v)\psi_i \right\|_{\mathrm{H}^{1/2}(a,b)}. \end{aligned} \tag{3.137}$$

As, by (3.126) with $r = 0$ and $s = 0$ and a scaling argument, for $i = 1, \dots, m-1$,

$$\begin{aligned} \left\| (\pi_N^i v - v)\psi_i \right\|_{\mathrm{L}^2(a,b)} &= \left\| (\pi_N^i v - v)\psi_i \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\leqslant \left\| \pi_N^i v - v \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\leqslant C(h_i + h_{i+1})^{1/2} \left\| \pi_{\tilde{p}_i - 1}(v \circ \tilde{F}_i) - v \circ \tilde{F}_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant C(h_i + h_{i+1})^{1/2} \left\| v \circ \tilde{F}_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant C \left\| v \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})}, \end{aligned} \tag{3.138}$$

and again correspondingly for $i = 0$ and $i = m$, which yields, as locally, there are at most two i such that $\psi_i \neq 0$,

$$\|J_N v - v\|_{\mathrm{L}^2(a,b)} \leqslant C \|v\|_{\mathrm{L}^2(a,b)}. \tag{3.139}$$

In a similar fashion, by a scaling argument together with (3.126) with $r = 0$ and $s = 1$, for $i = 1, \dots, m-1$,

$$\begin{aligned} \left\| [(\pi_N^i v - v)\psi_i]' \right\|_{\mathrm{L}^2(a,b)} &= \left\| [(\pi_N^i v - v)\psi_i]' \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &= \left\| (\pi_N^i v - v)'\psi_i + (\pi_N^i v - v)\psi_i' \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\leqslant \left\| (\pi_N^i v - v)'\psi_i \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\quad + \left\| (\pi_N^i v - v)\psi_i' \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\leqslant \left\| (\pi_N^i v - v)' \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\quad + C(h_i + h_{i+1})^{-1} \left\| \pi_N^i v - v \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\leqslant C(h_i + h_{i+1})^{-1/2} \left\| (\pi_{\tilde{p}_i - 1}(v \circ \tilde{F}_i) - v \circ \tilde{F}_i)' \right\|_{\mathrm{L}^2(-1,1)} \\ &\quad + C(h_i + h_{i+1})^{-1/2} \left\| \pi_{\tilde{p}_i - 1}(v \circ \tilde{F}_i) - v \circ \tilde{F}_i \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant C(h_i + h_{i+1})^{-1/2} \left\| (v \circ \tilde{F}_i)' \right\|_{\mathrm{L}^2(-1,1)} \\ &\leqslant C \left\| v' \right\|_{\mathrm{L}^2(x_{i-1}, x_{i+1})} \\ &\leqslant C \left\| v \right\|_{\mathrm{H}^1(x_{i-1}, x_{i+1})}, \end{aligned} \tag{3.140}$$

and correspondingly for $i = 0$ and $i = m$, which yields, by the finite overlap of the (ψ_i),

$$\|J_N v - v\|_{\mathrm{H}^1(a,b)} \leqslant C \|v\|_{\mathrm{H}^1(a,b)} . \tag{3.141}$$

Using Theorems 1.15 and 1.12 to interpolate the above estimates in $\mathrm{L}^2(a,b)$ and $\mathrm{H}^1(a,b)$, we obtain

$$\|J_N v\|_{\mathrm{H}^{1/2}(a,b)} \leqslant C \|v\|_{\mathrm{H}^{1/2}(a,b)} , \tag{3.142}$$

that is, the stability of the operators J_N. □

Combining the above results, we obtain:

Theorem 3.26. *There exists a sequence (P_N) of operators $P_N \colon \mathrm{H}^{1/2}(a,b) \to V_N$ such that for all $v \in \mathrm{H}^{1/2}(a,b)$,*

$$\|P_N v\|_{\mathrm{H}^{1/2}(a,b)} \leqslant C \left(1 + \max_{i=1,\dots,m} (p_i^{1/2} \log^{3/2} p_i)\right) \|v\|_{\mathrm{H}^{1/2}(a,b)} , \tag{3.143}$$

and for all $w \in \mathcal{P}^{p_i-2}$ and $i = 1,\dots,m$,

$$\langle P_N v, w\rangle_{\mathrm{L}^2(x_{i-1},x_i)} = \langle v, w\rangle_{\mathrm{L}^2(x_{i-1},x_i)} . \tag{3.144}$$

Proof. Plug in the operators J_N constructed in Theorem 3.24 into the estimate given in Lemma 3.23. This, together with Lemma 3.22, yields

$$\begin{aligned} \|P_N v\|_{\mathrm{H}^{1/2}(a,b)} &\leqslant \|J_N v\|_{\mathrm{H}^{1/2}(a,b)} + C \max_{i=1,\dots,m} \log^{3/2} p_i \times \\ &\quad \times \left(\sum_{i=1}^m p_i^2 \left\|\Pi^{\mathrm{L}^2(-1,1)}_{\mathcal{P}^{p_i-2}}(v \circ F_i) - J_N v \circ F_i\right\|^2_{\mathrm{L}^2(-1,1)}\right)^{1/2} \\ &\leqslant C \|v\|_{\mathrm{H}^{1/2}(a,b)} + C \max_{i=1,\dots,m} \log^{3/2} p_i \Bigg(p_1 |v|^2_{\mathrm{H}^{1/2}(x_0,x_2)} \\ &\qquad + \sum_{i=2}^{m-1} p_i |v|^2_{\mathrm{H}^{1/2}(x_{i-2},x_{i+1})} + p_m |v|^2_{\mathrm{H}^{1/2}(x_{m-2},x_m)}\Bigg)^{1/2} \\ &\leqslant C \|v\|_{\mathrm{H}^{1/2}(a,b)} \\ &\quad + C \max_{i=1,\dots,m} (p_i^{1/2} \log^{3/2} p_i) \left(\sum_{i=1}^m |v|^2_{\mathrm{H}^{1/2}(x_{i-1},x_i)}\right)^{1/2} \\ &\leqslant C(1 + \max_{i=1,\dots,m} (p_i^{1/2} \log^{3/2} p_i)) \|v\|_{\mathrm{H}^{1/2}(a,b)} . \end{aligned} \tag{3.145}$$

The second property follows from the definition of P_N in (3.103). □

Remark 3.27. Analysing the above proofs, we see that actually, given an operator J_N mapping a space of functions vanishing weakly at one or both end points into polynomials vanishing there as well and satisfying analogous approximation properties will ensure that P_N also satisfies these boundary conditions.

Consider therefore the space $\mathrm{H}^{1/2}_{(0}(a,b)$ as defined in Section 1.2. We then only have to define a modification $\tilde{\pi}^1_N$ of the local approximation operator π^1_N on the element (a, x_1) which satisfies $(\tilde{\pi}^1_N v)(a) = 0$ for all $v \in \mathrm{H}^{1/2}_{(0}(a,b)$.

The first step is the following argument, which is done on the reference element.

Lemma 3.28. *There exists a sequence* $(\tilde{\pi}_q)_{q=0}^{\infty}$ *of operators* $\tilde{\pi}_q\colon \mathrm{L}^2(0,1) \to \mathcal{P}^q$ *and a constant* $C > 0$ *which satisfy*

$$\begin{aligned} \|v - \tilde{\pi}_q v\|_{\mathrm{L}^2(0,1)} &\leqslant C \|v\|_{\mathrm{L}^2(0,1)} \\ &\quad \textit{for } v \in \mathrm{L}^2(0,1), \end{aligned} \tag{3.146}$$

$$\begin{aligned} \|v - \tilde{\pi}_q v\|_{\mathrm{H}^1(0,1)} &\leqslant C \|v\|_{\mathrm{H}^1(0,1)} \\ &\quad \textit{for } v \in \mathrm{H}^1(0,1), \end{aligned} \tag{3.147}$$

$$\begin{aligned} \|v - \tilde{\pi}_q v\|_{\mathrm{L}^2(0,1)} &\leqslant C(q+1)^{-1/2} |v|_{\mathrm{H}^{1/2}_{(0}(0,1)} \\ &\quad \textit{for } v \in \mathrm{H}^{1/2}_{(0}(0,1), \end{aligned} \tag{3.148}$$

$$\begin{aligned} \|v - \tilde{\pi}_q v\|_{\mathrm{L}^2(0,1)} &\leqslant C(q+1)^{-1} |v|_{\mathrm{H}^{1}_{(0}(0,1)} \\ &\quad \textit{for } v \in \mathrm{H}^{1}_{(0}(0,1), \end{aligned} \tag{3.149}$$

and

$$(\tilde{\pi}_q v)(a) = 0 \quad \textit{for } v \in \mathrm{H}^{1/2}_{(0}(0,1). \tag{3.150}$$

Proof. By Theorem 1.4, there exists an extension operator $\hat{E}\colon \mathrm{L}^2(0,1) \to \mathrm{L}^2(\mathbb{R})$ such that

$$\left\|\hat{E}v\right\|_{\mathrm{L}^2(\mathbb{R})} \leqslant C \|v\|_{\mathrm{L}^2(0,1)} \quad \text{for } v \in \mathrm{L}^2(0,1) \tag{3.151}$$

and

$$\left\|\hat{E}v\right\|_{\mathrm{H}^1(\mathbb{R})} \leqslant C \|v\|_{\mathrm{H}^1(0,1)} \quad \text{for } v \in \mathrm{H}^1(0,1). \tag{3.152}$$

Define the operator $E\colon \mathrm{L}^2(0,1) \to \mathrm{L}^2(-1,1)$ by

$$Ev := (\hat{E}v)|_{(-1,1)}. \tag{3.153}$$

Then,

$$\left\|\hat{E}v\right\|_{\mathrm{L}^2(-1,1)} \leqslant C \|v\|_{\mathrm{L}^2(0,1)} \quad \text{for } v \in \mathrm{L}^2(0,1) \tag{3.154}$$

and

$$\left\|\hat{E}v\right\|_{\mathrm{H}^1(-1,1)} \leqslant C \|v\|_{\mathrm{H}^1(0,1)} \quad \text{for } v \in \mathrm{H}^1(0,1). \tag{3.155}$$

Interpolating the above estimates using Theorems 1.15 and 1.12 yields

$$\|Ev\|_{\mathrm{H}^s(-1,1)} \leqslant C \|v\|_{\mathrm{H}^s(0,1)} \quad \text{for all } v \in \mathrm{H}^s(0,1) \text{ and all } s \in [0,1]. \tag{3.156}$$

Define $\tilde{\pi}_q\colon \mathrm{L}^2(0,1) \to \mathrm{L}^2(0,1)$ by

$$\tilde{\pi}_q u(x) := \pi_q E u(x) - (\pi_q E u)(0)(1-x)^q \quad \text{for } x \in (0,1). \tag{3.157}$$

We see that, as by definition $(Eu)|_{(0,1)} = u$,

$$\|\tilde{\pi}_q u - u\|_{L^2(0,1)} \leqslant \|\pi_q Eu - Eu\|_{L^2(-1,1)} + |\pi_q Eu(0)| \, \|(1-\cdot)^q\|_{L^2(0,1)} \tag{3.158}$$

and

$$\|\tilde{\pi}_q u - u\|_{H^1(0,1)} \leqslant \|\pi_q Eu - Eu\|_{H^1(-1,1)} + |\pi_q Eu(0)| \, \|(1-\cdot)^q\|_{H^1(0,1)} . \tag{3.159}$$

A simple calculation shows that

$$\|(1-\cdot)^q\|_{L^2(0,1)} \sim (q+1)^{-1/2} \tag{3.160}$$

and

$$\|(1-\cdot)^q\|_{H^1(0,1)} \sim (q+1)^{1/2}. \tag{3.161}$$

By Theorem 1.9, we see that, as $(1-x^2)|_{x=0} = 1$ and $0 \leqslant (1-x^2)^\gamma \leqslant 1$ for $x \in [-1,1]$ and $\gamma \geqslant 0$,

$$\begin{aligned} |\pi_q Eu(0)| &\leqslant \left\|(1-\cdot^2)\pi_q Eu\right\|_{L^\infty(-1,1)} \\ &\leqslant C \left\|(1-\cdot^2)\pi_q Eu\right\|_{L^2(-1,1)}^{1/2} \times \\ &\quad \times \left[\left\|(1-\cdot^2)\pi_q Eu\right\|_{L^2(-1,1)}^{1/2} + \left\|\left((1-\cdot^2)\pi_q Eu\right)'\right\|_{L^2(-1,1)}^{1/2} \right] \\ &\leqslant C \|\pi_q Eu\|_{L^2(-1,1)}^{1/2} \left[\|\pi_q Eu\|_{L^2(-1,1)}^{1/2} + \left\|(1-\cdot^2)^{1/2}(\pi_q Eu)'\right\|_{L^2(-1,1)}^{1/2} \right]. \end{aligned} \tag{3.162}$$

Theorem 1.18 yields

$$\left\|(1-\cdot^2)^{1/2}(\pi_q Eu)'\right\|_{L^2(-1,1)} \leqslant C(q+1) \|\pi_q Eu\|_{L^2(-1,1)} , \tag{3.163}$$

and thus

$$|\pi_q Eu(0)| \leqslant C(q+1)^{1/2} \|\pi_q Eu\|_{L^2(-1,1)} . \tag{3.164}$$

Choose first $u \in L^2(0,1)$; then, it follows by the L^2-stability of π_q that

$$\|\tilde{\pi}_q u - u\|_{L^2(0,1)} \leqslant C \|u\|_{L^2(0,1)} . \tag{3.165}$$

Next, choose $u \in H^1(0,1)$ with $u(0) = 0$, which also yields $Eu(0) = 0$. Then, again by Theorem 1.9,

$$\begin{aligned} |\pi_q Eu(0)| = |\pi_q Eu(0) - Eu(0)| &\leqslant \|\pi_q Eu - Eu\|_{L^\infty(-1,1)} \\ &\leqslant C \|\pi_q Eu - Eu\|_{L^2(-1,1)}^{1/2} \|\pi_q Eu - Eu\|_{H^1(-1,1)}^{1/2} . \end{aligned} \tag{3.166}$$

By the approximation properties and stability of π_q and the continuity of E, we have that

$$\|\pi_q Eu - Eu\|_{L^2(-1,1)} \leqslant C(q+1)^{-1} \|Eu\|_{H^1(-1,1)} \leqslant C(q+1)^{-1} \|u\|_{H^1(0,1)} \tag{3.167}$$

and

$$\|\pi_q Eu - Eu\|_{H^1(-1,1)} \leqslant C \|Eu\|_{H^1(-1,1)} \leqslant C \|u\|_{H^1(0,1)} , \tag{3.168}$$

Thus,

$$|\pi_q Eu(0)| \leqslant C(q+1)^{1/2} \|u\|_{H^1(0,1)}, \tag{3.169}$$

which yields both

$$\|\tilde{\pi}_q u - u\|_{L^2(0,1)} \leqslant C(q+1)^{-1} \|u\|_{H^1(0,1)}, \tag{3.170}$$

and

$$\|\tilde{\pi}_q u - u\|_{H^1(0,1)} \leqslant C \|u\|_{H^1(0,1)}, \tag{3.171}$$

Interpolating these estimates using Theorems 1.16 and 1.12, the results follow. □

Thus, defining $\tilde{\pi}_N^0$ by

$$\tilde{\pi}_N^0 v := \begin{cases} \tilde{\pi}_{\tilde{p}_0 - 1}(v \circ \tilde{F}_0) \circ \tilde{F}_0^{-1} & \text{on } (a, x_1), \\ 0, & \text{elsewhere,} \end{cases} \tag{3.172}$$

and $\tilde{J}_N$ by

$$\tilde{J}_N v := (\tilde{\pi}_N^0 v)\psi_0 + \sum_{i=1}^{m} (\pi_N^i v)\psi_i, \tag{3.173}$$

we obtain, with

$$\tilde{V}_N := \{v \in V_N : v(a) = 0\}, \tag{3.174}$$

the following result corresponding to Theorem 3.26:

Theorem 3.29. *There exists a sequence $(\tilde{P}_N)$ of operators $\tilde{P}_N : H_{(0}^{1/2}(a,b) \to \tilde{V}_N$ such that for all $v \in H_{(0}^{1/2}(a,b)$,*

$$\|P_N v\|_{H_{(0}^{1/2}(a,b)} \leqslant C(1 + \max_{i=1,\dots,m}(p_i^{1/2} \log^{3/2} p_i)) \|v\|_{H_{(0}^{1/2}(a,b)}, \tag{3.175}$$

and for all $w \in \mathcal{P}^{p_i - 2}$ and $i = 1, \dots, m$,

$$\langle P_N v, w\rangle_{L^2(x_{i-1},x_i)} = \langle v, w\rangle_{L^2(x_{i-1},x_i)}. \tag{3.176}$$

The proof is analogous to the steps done before; we only have to note that

$$\|P_N v - J_N v\|_{H_{(0}^{1/2}(a,b)} \leqslant \|P_N v - J_N v\|_{H_{00}^{1/2}(a,b)}. \tag{3.177}$$

Remark 3.30. Corresponding results also hold true for $H_{0)}^{1/2}(a,b)$ and $H_{00}^{1/2}(a,b)$.

3.2.3 A priori Error Estimates for the Frictional Contact Problem

We are now ready to give an a priori convergence rate result for the finite element method formulated in Problem 3.11. For simplicity, however, we shall assume that contact holds on the complete contact boundary and that $u_0 = 0$, that is, we solve the continuous problem on the closed convex set

$$\mathcal{K}' := \{\mathbf{v} \in V : \gamma_{0,\Gamma_C}(\mathbf{v}) \cdot \boldsymbol{\nu} = 0 \text{ almost everywhere on } \Gamma_C\}, \tag{3.178}$$

and for the discrete problem,

$$\mathcal{K}'_N := \{\mathbf{v}_N \in V_N \colon \gamma_{0,\Gamma_{\mathrm{C}}}(\mathbf{v}_N) \cdot \boldsymbol{\nu} = 0 \text{ almost everywhere on } \Gamma_{\mathrm{C}}\}\,. \tag{3.179}$$

In particular, we note that $\mathcal{K}'$ and $\mathcal{K}'_N$ are linear spaces and $\mathcal{K}'_N \subseteq \mathcal{K}'$. Thus, the continuous and discrete primal-dual problems become:

Problem 3.31 (Primal-dual formulation, Tresca friction, forced contact)**.** Find $(\mathbf{u}, \boldsymbol{\lambda}) \in \mathcal{K}' \times \Lambda$ such that for all $(\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{K}' \times \Lambda$,

$$a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, \boldsymbol{\lambda}) = L(\mathbf{v}), \tag{3.180}$$

$$b(\mathbf{u}, \boldsymbol{\mu} - \boldsymbol{\lambda}) \leqslant 0. \tag{3.181}$$

Problem 3.32 (Discrete primal-dual formulation, Tresca friction, forced contact)**.** Find $(\mathbf{u}_N, \boldsymbol{\lambda}_N) \in \mathcal{K}'_N \times \Lambda_N$ such that for all $(\mathbf{v}_N, \boldsymbol{\mu}_N) \in \mathcal{K}'_N \times \Lambda_N$,

$$a(\mathbf{u}_N, \mathbf{v}_N) + b(\mathbf{v}_N, \boldsymbol{\lambda}_N) = L(\mathbf{v}_N), \tag{3.182}$$

$$b(\mathbf{u}_N, \boldsymbol{\mu}_N - \boldsymbol{\lambda}_N) \leqslant 0. \tag{3.183}$$

The existence, uniqueness and basic convergence results can be proved similarly as in the situation considered above.

We now prove an inf-sup condition using the results of the last section. For simplicity, we restrict ourselves to the situation that Γ_{C} consists of a single affine line.

Theorem 3.33. *Assume that $\overline{\Gamma_{\mathrm{C}}} \cap \overline{\Gamma_{\mathrm{D}}} = \varnothing$.*

Then, we have the discrete inf-sup condition

$$\inf_{\boldsymbol{\mu}_N \in W_N} \sup_{\mathbf{v}_N \in V_N} \frac{b(\mathbf{v}_N, \boldsymbol{\mu}_N)}{\|\mathbf{v}_N\|_{\mathrm{H}^1(\Omega)} \|\boldsymbol{\mu}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_{\mathrm{C}})}} \geqslant \frac{C}{\max_{E \in \mathcal{E}_{\mathrm{C},N}} (p_{N,K_E}^{1/2} \log^{3/2} p_{N,K_E})}, \tag{3.184}$$

where the constant $C > 0$ is independent of N.

In the proof of this theorem, we need to extend a function v given on a part Γ_{C} of the boundary Γ of Ω to a function $\tilde{v}$ on Ω which satisfies the (homogeneous) Dirichlet boundary conditions. If $v \in \mathrm{H}^{1/2}_{00}(\Gamma_{\mathrm{C}})$, this can be done by simply setting $\tilde{v} = 0$ on $\Gamma \backslash \Gamma_{\mathrm{C}}$ and applying Theorem 1.3. If, however, $v \in \mathrm{H}^{1/2}(\Gamma_{\mathrm{C}})$, we need to ensure that $\tilde{v} = 0$ on Γ_{D}. In principle, the theory developed in [Gri85, Section 1.5.2] shows that this is possible. We prove this result directly for polygons.

Lemma 3.34. *Let $\Omega \subseteq \mathbb{R}^2$ be a polygon, and let Γ_{C}, Γ_{D} be disjoint, relatively open parts of $\partial\Omega$ consisting each of a finite number of edges of Ω.*

Then, there exists a continuous operator $E \colon \mathrm{H}^{1/2}(\Gamma_{\mathrm{C}}) \to \mathrm{H}^1(\Omega)$ such that $\gamma_{0,\Gamma_{\mathrm{C}}} E v = v$ and $\gamma_{0,\Gamma_{\mathrm{D}}} E v = 0$ for all $v \in \mathrm{H}^{1/2}(\Gamma_{\mathrm{C}})$.

Proof. We only prove the result under the simplifying assumption that $\Omega := (a, b) \times (c, d)$, $\Gamma_{\mathrm{C}} := (a, b) \times \{c\}$ and $\Gamma_{\mathrm{D}} := (a, b) \times \{d\}$. In the case of a general polygon, a similar, but much more technical, argument is possible.

Similarly as in the proof of Lemma 3.28, we see that there exists a continuous extension operator $\tilde{E} \colon \mathrm{H}^{1/2}(\Gamma_{\mathrm{C}}) \to \mathrm{H}^{1/2}(\mathbb{R} \times \{c\})$. By Theorem 1.3, we see that there exists a continuous lifting operator $Z \colon \mathrm{H}^{1/2}(\mathbb{R} \times \{c\}) \to \mathrm{H}^1(\mathbb{R}^2)$. Choose a function $\varphi \in \mathrm{C}^\infty_0(\mathbb{R}^2)$ with $0 \leqslant \varphi \leqslant 1$

such that $\varphi(x,c) = 1$ and $\varphi(x,d) = 0$ for $x \in (a,b)$. Then, the operator $E\colon \mathrm{H}^{1/2}(\Gamma_\mathrm{C}) \to \mathrm{H}^1(\Omega)$ defined by

$$Ev := (\varphi Z \tilde{E} v)|_\Omega \tag{3.185}$$

satisfies the claimed properties. □

A similar result holds true for a space of functions vanishing at one boundary point of Γ_C.

Proof of Theorem 3.33. As Γ_C consists of a single affine line, we can assume without loss of generality that $\Gamma_\mathrm{C} = \{(t,0)\colon t \in (a,b)\}$. Furthermore, we see that the partition $\mathcal{E}_{\mathrm{C},N}$ of Γ_C is given by points $(x_i)_{i=0,\dots,m}$ with $a = x_0 < x_1 < \dots < x_m = b$ such that for every $E \in \mathcal{E}_{\mathrm{C},N}$ there exists $i \in \{1,\dots,m\}$ with $E = (x_{i-1}, x_i)$ and polynomial degrees $p_{N,i} := p_{N,K_E}$. The fact that g is constant yields that $b = g\langle\cdot,\cdot\rangle_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}$, that is, b is just a scalar multiple of the duality product of $\mathrm{H}^{1/2}(\Gamma_\mathrm{C}) = \gamma_{0,\Gamma_\mathrm{C}} V$ and $\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C}) = (\mathrm{H}^{1/2}(\Gamma_\mathrm{C}))^*$, and thus, for a given $\boldsymbol{\mu}_N \in W_N$ there exists $\mathbf{v} \in V$ with

$$\frac{b(\mathbf{v}, \boldsymbol{\mu}_N)}{\|\mathbf{v}\|_{\mathrm{H}^1(\Omega)} \|\boldsymbol{\mu}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}} \geqslant \frac{g}{2}. \tag{3.186}$$

Applying the operators P_N constructed in Theorem 3.26 componentwise and using a bounded extension operator $E\colon \mathrm{H}^{1/2}(\Gamma_\mathrm{C}) \to V$ as given in Lemma 3.34, we see that due to Theorem 1.2, together with the operators $\hat{i}_N$ given in Theorem 3.13,

$$\left\|\hat{i}_N E P_N \gamma_{0,\Gamma_\mathrm{C}} \mathbf{v}\right\|_{\mathrm{H}^1(\Omega)} \leqslant C \max_{E \in \mathcal{E}_{\mathrm{C},N}} \left(p_{N,K_E}^{1/2} \log^{3/2} p_{N,K_E}\right) \|\mathbf{v}\|_{\mathrm{H}^1(\Omega)}\,. \tag{3.187}$$

Thus, there exists $\mathbf{v}_N := \hat{i}_N E P_N \gamma_{0,\Gamma_\mathrm{C}} \mathbf{v} \in V_N$ with $\gamma_{0,\Gamma_\mathrm{C}} \mathbf{v}_N = P_N \gamma_{0,\Gamma_\mathrm{C}} \mathbf{v}$ such that

$$\begin{aligned}
&\frac{b(\mathbf{v}_N, \boldsymbol{\mu}_N)}{\|\mathbf{v}_N\|_{\mathrm{H}^1(\Omega)} \|\boldsymbol{\mu}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}} \\
&\quad = \frac{b(\mathbf{v}, \boldsymbol{\mu}_N)}{\|\mathbf{v}_N\|_{\mathrm{H}^1(\Omega)} \|\boldsymbol{\mu}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}} \\
&\quad \geqslant \frac{b(\mathbf{v}, \boldsymbol{\mu}_N)}{C \max_{E \in \mathcal{E}_{\mathrm{C},N}} \left(p_{N,K_E}^{1/2} \log^{3/2} p_{N,K_E}\right) \|\mathbf{v}_N\|_{\mathrm{H}^1(\Omega)} \|\boldsymbol{\mu}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}},
\end{aligned} \tag{3.188}$$

and the assertion follows. □

Theorem 3.35. *Assume that* $\overline{\Gamma_\mathrm{C}} \cap \overline{\Gamma_\mathrm{D}} = \varnothing$, *and set*

$$\beta_N := \max_{E \in \mathcal{E}_{\mathrm{C},N}} \left(p_{N,K_E}^{1/2} \log^{3/2} p_{N,K_E}\right)^{-1}. \tag{3.189}$$

There exists a constant $C > 0$ *independent of* N *such that for* $(\mathbf{u}, \boldsymbol{\lambda})$ *the solution of Problem 3.31 and* $(\mathbf{u}_N, \boldsymbol{\lambda}_N)$ *the solution of Problem 3.32,*

$$\begin{aligned}
\|\mathbf{u} - \mathbf{u}_N\|^2_{\mathrm{H}^1(\Omega)} \leqslant C\Big[&b(\mathbf{u}, \boldsymbol{\lambda}_N - \boldsymbol{\mu}) + b(\mathbf{u}, \boldsymbol{\lambda} - \boldsymbol{\mu}_N) \\
&+ \beta_N^{-2}\left(\|\mathbf{u} - \mathbf{v}_N\|^2_{\mathrm{H}^1(\Omega)} + \|\boldsymbol{\lambda} - \boldsymbol{\mu}_N\|^2_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}\right)\Big]
\end{aligned} \tag{3.190}$$

and

$$\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})} \leqslant C\beta_N^{-1}\left(\|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(\Omega)} + \|\boldsymbol{\lambda} - \boldsymbol{\mu}_N\|_{\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})}\right). \tag{3.191}$$

Proof. Applying Theorem 3.7 together with Theorem 3.33 yields the result. □

Remark 3.36. Applying the results of Remark 3.27 instead of Theorem 3.26, we see that the assumption $\overline{\Gamma_\mathrm{C}} \cap \overline{\Gamma_\mathrm{D}} = \varnothing$ is not necessary. We then only have to replace the space $\tilde{\mathrm{H}}^{-1/2}(\Gamma_\mathrm{C})$ by $\mathrm{H}^{-1/2}(\Gamma_\mathrm{C})$ or, potentially, a space satisfying zero boundary conditions at one of the points in $\overline{\Gamma_\mathrm{C}} \cap \overline{\Gamma_\mathrm{D}}$.

Chapter 4

The Residual Error Indicator for the Frictional Contact Problem

Error indicators are a fundamental part of modern finite element implementations. The two goals which one aims for with error indicators are

1. to be able to determine whether the numerical approximation which was obtained by the finite element method is "good enough" for the application under consideration, and

2. to find the elements where the error is high, and which should therefore be refined in an adaptive computation.

Today, there are several very versatile error indicators available. In this work, we will focus on the classical residual error indicator, which was generalised to the *hp*-context by Melenk and Wohlmuth in [MW01]. To be able to apply this error indicator for the variational inequality under consideration, we shall make use of the dual approach to a posteriori error indication given in [Han05, Chapter 6]. Using this error indicator, we propose an *hp*-adaptive mesh refinement strategy based on local estimation of solution regularity through the decay of Legendre coefficients, which was developed in [HS05] and [EM07].

As in Subsection 3.2.3, we assume in the following that $\gamma_{0,\Gamma_C}(\mathbf{v}) \cdot \boldsymbol{\nu} = 0$ almost everywhere on Γ_C.

4.1 Duality-Based Error Estimation

Let $d = 2$ or $d = 3$, $Z_1 := \left(\mathrm{L}^2(\Omega)\right)^{d \times d}$ with the scalar product

$$\langle \mathbf{q}, \mathbf{r} \rangle_{Z_1} := \int_\Omega \mathbf{C}\mathbf{q} : \mathbf{r} \mathrm{d}\mathbf{x}, \tag{4.1}$$

$Z_2 := \left(\mathrm{L}^2(\Gamma_C)\right)^d$, and $Z := Z_1 \times Z_2$. Defining a function $\mathcal{J} \colon V \times Z \to \mathbb{R}$ by

$$\mathcal{J}(\mathbf{v}, \mathbf{q}) := \frac{1}{2} \langle \mathbf{q}_1, \mathbf{q}_1 \rangle_{Z_1} \mathrm{d}\mathbf{x} - L(\mathbf{v}) + \int_{\Gamma_C} g \, |\mathbf{q}_2| \, \mathrm{d}s_{\mathbf{x}}, \tag{4.2}$$

where $\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2)$, we see easily that, by introducing $\mathcal{L} \colon V \to Z$ through

$$\mathcal{L}\mathbf{v} := (\boldsymbol{\varepsilon}(\mathbf{v}), \gamma_{0,\Gamma_C}\mathbf{v}), \tag{4.3}$$

we have that, as $a(\mathbf{v}, \mathbf{w}) = \langle \boldsymbol{\varepsilon}(\mathbf{v}), \boldsymbol{\varepsilon}(\mathbf{w}) \rangle_{Z_1}$,

$$J(\mathbf{v}) = \mathcal{J}(\mathbf{v}, \mathcal{L}\mathbf{v}) \quad \text{for all } \mathbf{v} \in V. \tag{4.4}$$

Thus, in particular, the minimisation formulation in Problem 2.5 (or, for $d = 2$, Problem 2.9) can be written as: Find $\mathbf{u} \in V$ such that

$$\mathcal{J}(\mathbf{u}, \mathcal{L}\mathbf{u}) = \inf_{\mathbf{v} \in V} \mathcal{J}(\mathbf{v}, \mathcal{L}\mathbf{v}). \tag{4.5}$$

With $\mathcal{L}^* \colon Z^* \to V^*$ the adjoint of $\mathcal{L}$, we want to calculate the conjugate function $\mathcal{J}^* \colon V^* \times Z^* \to \overline{\mathbb{R}}$, which is given on $\{(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*) \colon \mathbf{q}^* \in Z^*\}$ by

$$\mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*) = \sup_{(\mathbf{v},\mathbf{q}) \in V \times Z} \left[\mathcal{L}^*\mathbf{q}^*(\mathbf{v}) - \mathbf{q}^*(\mathbf{q}) - \mathcal{J}(\mathbf{v}, \mathbf{q})\right]. \tag{4.6}$$

Plugging in the definition of $\mathcal{J}$,

$$\begin{aligned}
&\sup_{(\mathbf{v},\mathbf{q}) \in V \times Z} \left[\mathcal{L}^*\mathbf{q}^*(\mathbf{v}) - \mathbf{q}^*(\mathbf{q}) - \mathcal{J}(\mathbf{v}, \mathbf{q})\right] \\
&\qquad = \sup_{(\mathbf{v},\mathbf{q}) \in V \times Z} \left[\mathcal{L}^*\mathbf{q}^*(\mathbf{v}) - \mathbf{q}^*(\mathbf{q}) - \frac{1}{2}\langle \mathbf{q}_1, \mathbf{q}_1 \rangle_{Z_1} + L(\mathbf{v}) - \int_{\Gamma_C} g\,|\mathbf{q}_2|\,\mathrm{d}s_{\mathbf{x}}\right] \\
&\qquad = \sup_{\mathbf{v} \in V} \left[\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + L(\mathbf{v})\right] \\
&\qquad\quad + \sup_{\mathbf{q}_1 \in Z_1} \left[-\left\langle \mathbf{q}_1^* + \frac{1}{2}\mathbf{q}_1, \mathbf{q}_1 \right\rangle_{Z_1}\right] \\
&\qquad\quad + \sup_{\mathbf{q}_2 \in Z_2} \left[-\langle \mathbf{q}_2^*, \mathbf{q}_2 \rangle_{L^2(\Gamma_C)} - \int_{\Gamma_C} g\,|\mathbf{q}_2|\,\mathrm{d}s_{\mathbf{x}}\right].
\end{aligned} \tag{4.7}$$

Clearly,

$$\sup_{\mathbf{q}_1 \in Z_1} \left[-\left\langle \mathbf{q}_1^* + \frac{1}{2}\mathbf{q}_1, \mathbf{q}_1 \right\rangle_{Z_1}\right] = \frac{1}{2}\langle \mathbf{q}_1^*, \mathbf{q}_1^* \rangle_{Z_1}. \tag{4.8}$$

Next,

$$\begin{aligned}
&\sup_{\mathbf{q}_2 \in Z_2} \left[-\langle \mathbf{q}_2^*, \mathbf{q}_2 \rangle_{L^2(\Gamma_C)} - \int_{\Gamma_C} g\,|\mathbf{q}_2|\,\mathrm{d}s_{\mathbf{x}}\right] \\
&\qquad = \begin{cases} 0 & \text{if } |\mathbf{q}_2^*| \leqslant g \text{ almost everywhere on } \Gamma_C, \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned} \tag{4.9}$$

Finally, it is easy to see that

$$\sup_{\mathbf{v} \in V} \left[\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + L(\mathbf{v})\right] = +\infty \tag{4.10}$$

unless $\mathbf{q}^* \in Z^*$ is chosen in such a way that

$$\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + L(\mathbf{v}) = 0 \quad \text{for all } \mathbf{v} \in V, \tag{4.11}$$

and in that case, the supremum is trivially 0. We define the set of admissible dual functions by

$$\begin{aligned}
Z_0^* := \Big\{\mathbf{q}^* \in Z^* \colon &\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + L(\mathbf{v}) = 0 \text{ for all } \mathbf{v} \in V, \\
&\text{and } |\mathbf{q}_2^*| \leqslant g \text{ almost everywhere on } \Gamma_C\Big\}.
\end{aligned} \tag{4.12}$$

Thus, the conjugate function is given by

$$\mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*) = \begin{cases} \frac{1}{2}\langle \mathbf{q}_1^*, \mathbf{q}_1^* \rangle_{Z_1} & \text{for } \mathbf{q}^* \in Z_0^*, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.13}$$

Consider the dual problem of finding $\mathbf{p}^* \in Z_0^*$ such that

$$-\mathcal{J}^*(\mathcal{L}^*\mathbf{p}^*, -\mathbf{p}^*) = \sup_{\mathbf{q}^* \in Z_0^*} \left[-\mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*)\right]. \tag{4.14}$$

To prove the strict convexity of the dual functional, let $\mathbf{q}^*$, $\tilde{\mathbf{q}}^* \in Z_0^*$ with $\mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*)$, $\mathcal{J}^*(\mathcal{L}^*\tilde{\mathbf{q}}^*, -\tilde{\mathbf{q}}^*) < \infty$ and $s \in (0, 1)$ such that

$$\begin{aligned} \mathcal{J}^*(\mathcal{L}^*(s\mathbf{q}^* + (1-s)\tilde{\mathbf{q}}^*), -(s\mathbf{q}^* + (1-s)\tilde{\mathbf{q}}^*)) \\ = s\mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*) + (1-s)\mathcal{J}^*(\mathcal{L}^*\tilde{\mathbf{q}}^*, -\tilde{\mathbf{q}}^*). \end{aligned} \tag{4.15}$$

This implies

$$\langle s\mathbf{q}_1^* + (1-s)\tilde{\mathbf{q}}_1^*, s\mathbf{q}_1^* + (1-s)\tilde{\mathbf{q}}_1^* \rangle_{Z_1} = s\langle \mathbf{q}_1^*, \mathbf{q}_1^* \rangle_{Z_1} + (1-s)\langle \tilde{\mathbf{q}}_1^*, \tilde{\mathbf{q}}_1^* \rangle_{Z_1}, \tag{4.16}$$

and thus $\mathbf{q}_1^* = \tilde{\mathbf{q}}_1^*$. Therefore,

$$\langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} = \langle \gamma_{0,\Gamma_C}\mathbf{v}, \tilde{\mathbf{q}}_2^* \rangle_{L^2(\Gamma_C)} \quad \text{for all } \mathbf{v} \in V, \tag{4.17}$$

and we obtain $\mathbf{q}^* = \tilde{\mathbf{q}}^*$.

It follows similarly that the mapping $\mathbf{q}^* \mapsto \mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, -\mathbf{q}^*)$ is coercive, and we can apply Theorem 1.28 to obtain the existence of a unique solution $\mathbf{p}^* \in Z_0^*$ of the above problem which satisfies

$$\mathcal{J}(\mathbf{u}, \mathcal{L}\mathbf{u}) = -\mathcal{J}^*(\mathcal{L}^*\mathbf{p}^*, -\mathbf{p}^*), \tag{4.18}$$

where $\mathbf{u}$ is the minimiser of J.

For an arbitrary $\mathbf{w} \in V$, by the variational inequality formulation in Problem 2.4,

$$\begin{aligned} \frac{1}{2}a(\mathbf{u}-\mathbf{w}, \mathbf{u}-\mathbf{w}) &= \frac{1}{2}a(\mathbf{w}, \mathbf{w}) - a(\mathbf{u}, \mathbf{w}) + \frac{1}{2}a(\mathbf{u}, \mathbf{u}) \\ &= \frac{1}{2}a(\mathbf{w}, \mathbf{w}) - a(\mathbf{u}, \mathbf{w}-\mathbf{u}) - \frac{1}{2}a(\mathbf{u}, \mathbf{u}) \\ &\leqslant \frac{1}{2}a(\mathbf{w}, \mathbf{w}) + j(\mathbf{w}) - j(\mathbf{u}) - L(\mathbf{w}-\mathbf{u}) - \frac{1}{2}a(\mathbf{u}, \mathbf{u}) \\ &= J(\mathbf{w}) - J(\mathbf{u}). \end{aligned} \tag{4.19}$$

As $\mathbf{p}^*$ solves the dual problem, we obtain that

$$J(\mathbf{u}) = \mathcal{J}(\mathbf{u}, \mathcal{L}\mathbf{u}) = -\mathcal{J}^*(\mathcal{L}^*\mathbf{p}^*, -\mathbf{p}^*) \geqslant -\mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, \mathbf{q}^*) \quad \text{for all } \mathbf{q}^* \in Z_0^*, \tag{4.20}$$

and thus

$$\frac{1}{2}a(\mathbf{u}-\mathbf{w}, \mathbf{u}-\mathbf{w}) \leqslant J(\mathbf{w}) + \mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, \mathbf{q}^*) \quad \text{for all } \mathbf{q}^* \in Z_0^*. \tag{4.21}$$

Let $\mathbf{r}_1^* \in Z_1^*$ be arbitrary, then

$$\begin{aligned} J(\mathbf{w}) + \mathcal{J}^*(\mathcal{L}^*\mathbf{q}^*, \mathbf{q}^*) = \frac{1}{2}\langle \varepsilon(\mathbf{w}) + \mathbf{r}_1^*, \varepsilon(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} - \langle \varepsilon(\mathbf{w}), \mathbf{r}_1^* \rangle_{Z_1} - L(\mathbf{w}) \\ + j(\mathbf{w}) + \frac{1}{2}\left[\langle \mathbf{q}_1^*, \mathbf{q}_1^* \rangle_{Z_1} - \langle \mathbf{r}_1^*, \mathbf{r}_1^* \rangle_{Z_1}\right]. \end{aligned} \tag{4.22}$$

As $\mathbf{q}^* \in Z_0^*$, we have that

$$\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + L(\mathbf{w}) = 0, \tag{4.23}$$

which yields

$$\begin{aligned} J(\mathbf{w}) + \mathcal{J}^*(\mathcal{L}^* \mathbf{q}^*, \mathbf{q}^*) &= \frac{1}{2} \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} - \langle \boldsymbol{\varepsilon}(\mathbf{w}), \mathbf{r}_1^* \rangle_{Z_1} \\ &\quad + \langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} \\ &\quad + j(\mathbf{w}) + \frac{1}{2} \langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} + \langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{r}_1^* \rangle_{Z_1} \\ &= \frac{1}{2} \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} + \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} \\ &\quad + \frac{1}{2} \langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + j(\mathbf{w}). \end{aligned} \tag{4.24}$$

Furthermore,

$$\begin{aligned} &\frac{1}{2} \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} + \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} + \frac{1}{2} \langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} \\ &\quad \leqslant \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} + \langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1}. \end{aligned} \tag{4.25}$$

Thus, we have proved the following result.

Theorem 4.1. *Let $\mathbf{u} \in V$ be the solution of the continuous minimisation formulation in Problem 2.5, and $w \in V$ arbitrary.*

Then, for all $\mathbf{r}_1^* \in Z_1^*$,

$$\begin{aligned} \frac{1}{2} a(\mathbf{u} - \mathbf{w}, \mathbf{u} - \mathbf{w}) &\leqslant \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} \\ &\quad + \inf_{\mathbf{q}^* \in Z_0^*} \Big[\langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + j(\mathbf{w}) \Big]. \end{aligned} \tag{4.26}$$

Define the residual $\mathcal{R} \colon Z_2^* \times Z_1^* \to \mathbb{R}$ by

$$\mathcal{R}(\mathbf{q}_2^*, \mathbf{r}_1^*) := \sup_{\mathbf{v} \in V} a(\mathbf{v}, \mathbf{v})^{-1/2} \Big[\langle \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{v}) \rangle_{Z_1} + L(\mathbf{v}) + \langle \mathbf{q}_2^*, \gamma_{0,\Gamma_C} \mathbf{v} \rangle_{L^2(\Gamma_C)} \Big]. \tag{4.27}$$

As $\mathbf{q}^* \in Z_0^*$ if and only if $|\mathbf{q}_2^*| \leqslant g$ almost everywhere on Γ_C and

$$\sup_{\mathbf{v} \in V} \Big[\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v}) \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{v}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + L(\mathbf{v}) \Big] = 0 \tag{4.28}$$

and the above expression is infinite otherwise, we see that, setting

$$Z_\#^* := \{ \mathbf{q}_2^* \in Z_2^* : |\mathbf{q}_2^*| \leqslant g \text{ almost everywhere on } \Gamma_C \} \tag{4.29}$$

and replacing $\mathbf{q}_1^* - \mathbf{r}_1^*$ by $\mathbf{q}_1^*$ in the infimum over $Z_\#^*$, we get

$$\begin{aligned} &\inf_{\mathbf{q}^* \in Z_0^*} \Big[\langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + j(\mathbf{w}) \Big] \\ &\quad = \inf_{(\mathbf{q}_1^*, \mathbf{q}_2^*) \in Z_1^* \times Z_\#^*} \sup_{\mathbf{v} \in V} \Big[\langle \mathbf{q}_1^* - \mathbf{r}_1^*, \mathbf{q}_1^* - \mathbf{r}_1^* \rangle_{Z_1} + \langle \gamma_{0,\Gamma_C} \mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + j(\mathbf{w}) \end{aligned}$$

$$+ \langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + L(\mathbf{v})\Big] \tag{4.30}$$

$$= \inf_{(\mathbf{q}_1^*,\mathbf{q}_2^*)\in Z_1^*\times Z_\#^*} \sup_{\mathbf{v}\in V} \Big[\langle \mathbf{q}_1^*, \mathbf{q}_1^*\rangle_{Z_1} + \langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})$$
$$+ \langle \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + L(\mathbf{v})\Big].$$

The above expression can be simplified by applying the *principle of complementary energy* (see [BF91, p. 20], [Bra07, p. 293]); we do the calculation here using Theorem 1.28. Define, for $\mathbf{q}_2^* \in Z_\#^*$ fixed, the linear functional $\ell_{\mathbf{q}_2^*}\colon V \to \mathbb{R}$ by

$$\ell_{\mathbf{q}_2^*}(\mathbf{v}) := \langle \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + L(\mathbf{v}). \tag{4.31}$$

We see that the above simplifies to

$$\inf_{\mathbf{q}_2^*\in Z_\#^*} \inf_{\mathbf{q}_1^*\in Z_1^*} \sup_{\mathbf{v}\in V} \Big[\langle \mathbf{q}_1^*, \mathbf{q}_1^*\rangle_{Z_1} + \langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w}) + \ell_{\mathbf{q}_2^*}(\mathbf{v})\Big]. \tag{4.32}$$

Clearly, the inner supremum is infinite unless

$$\langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \ell_{\mathbf{q}_2^*}(\mathbf{v}) = 0 \quad \text{for all } \mathbf{v} \in V, \tag{4.33}$$

and thus, defining $F\colon Z_1^* \to \mathbb{R}$ by $F(\mathbf{q}_1^*) := \langle \mathbf{q}_1^*, \mathbf{q}_1^*\rangle_{Z_1}$ and $G\colon V^* \to \bar{\mathbb{R}}$ by $G(-\ell_{\mathbf{q}_2^*}) := 0$ and $G(\ell) := +\infty$ for $\ell \neq \ell_{\mathbf{q}_2^*}$, we obtain, with the linear operator $\mathcal{M}\colon Z_1^* \to V^*$ given by $(\mathcal{M}\mathbf{q}_1^*)(\mathbf{v}) := \langle \mathbf{q}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1}$,

$$\inf_{\mathbf{q}_2^*\in Z_\#^*} \inf_{\mathbf{q}_1^*\in Z_1^*} \Big(F(\mathbf{q}_1^*) + G(\mathcal{M}\mathbf{q}_1^*) + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})\Big). \tag{4.34}$$

As $F^*\colon Z_1 \to \mathbb{R}$ and $G^*\colon V \to \mathbb{R}$ are given by

$$F^*(\mathbf{q}_1) = \sup_{\mathbf{q}_1^*\in Z_1^*} \big[\langle \mathbf{q}_1^*, \mathbf{q}_1\rangle_{Z_1} - \langle \mathbf{q}_1^*, \mathbf{q}_1^*\rangle_{Z_1}\big] = \frac{1}{4}\langle \mathbf{q}_1, \mathbf{q}_1\rangle_{Z_1} \tag{4.35}$$

and

$$G^*(\mathbf{v}) = -\ell_{\mathbf{q}_2^*}(\mathbf{v}), \tag{4.36}$$

respectively, F is coercive, and $\mathcal{M}^*\colon V \to Z_1$ is given by $\mathcal{M}^*\mathbf{v} = \boldsymbol{\varepsilon}(\mathbf{v})$, we can apply Theorem 1.28 to obtain

$$\inf_{\mathbf{q}_2^*\in Z_\#^*} \inf_{\mathbf{q}_1^*\in Z_1^*} \Big(F(\mathbf{q}_1^*) + G(\mathcal{M}\mathbf{q}_1^*) + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})\Big)$$
$$= \inf_{\mathbf{q}_2^*\in Z_\#^*} \sup_{\mathbf{v}\in V} \Big(-\frac{1}{4}\langle \boldsymbol{\varepsilon}(\mathbf{v}), \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \ell_{\mathbf{q}_2^*}(\mathbf{v}) + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})\Big)$$
$$= \inf_{\mathbf{q}_2^*\in Z_\#^*} \sup_{\mathbf{v}\in V} \Big[-\frac{1}{4}\langle \boldsymbol{\varepsilon}(\mathbf{v}), \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})$$
$$+ \langle \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{v})\rangle_{Z_1} + \langle \gamma_{0,\Gamma_C}\mathbf{v}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + L(\mathbf{v})\Big] \tag{4.37}$$
$$\leqslant \inf_{\mathbf{q}_2^*\in Z_\#^*} \sup_{\mathbf{v}\in V} \Big[-\frac{1}{4}a(\mathbf{v},\mathbf{v}) + \mathcal{R}(\mathbf{q}_2^*, \mathbf{r}_1^*)a(\mathbf{v},\mathbf{v})^{1/2} + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})\Big]$$
$$= \inf_{\mathbf{q}_2^*\in Z_\#^*} \Big[\mathcal{R}(\mathbf{q}_2^*, \mathbf{r}_1^*)^2 + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^*\rangle_{L^2(\Gamma_C)} + j(\mathbf{w})\Big].$$

This yields:

Theorem 4.2. *Let* $\mathbf{u} \in V$ *be the solution of the continuous minimisation formulation in Problem 2.5, and* $w \in V$ *arbitrary.*

Then, for all $\mathbf{r}_1^* \in Z_1^*$,

$$\begin{aligned}\frac{1}{2}a(\mathbf{u}-\mathbf{w},\mathbf{u}-\mathbf{w}) \leqslant{}& \langle \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^*, \boldsymbol{\varepsilon}(\mathbf{w}) + \mathbf{r}_1^* \rangle_{Z_1} \\ &+ \inf_{\mathbf{q}_2^* \in Z_\#^*} \Big[\mathcal{R}(\mathbf{q}_2^*, \mathbf{r}_1^*)^2 + \langle \gamma_{0,\Gamma_C}\mathbf{w}, \mathbf{q}_2^* \rangle_{L^2(\Gamma_C)} + j(\mathbf{w}) \Big].\end{aligned} \tag{4.38}$$

4.2 Reliability and Efficiency of the Residual Error Indicator

Selecting $\mathbf{w} := \mathbf{u}_N$, $\mathbf{r}_1^* := -\boldsymbol{\varepsilon}(\mathbf{u}_N)$ and $\mathbf{q}_2^* := -g\tilde{\boldsymbol{\lambda}}_N$ in Theorem 4.2 and applying Theorem 2.8, we have the error estimate

$$\|\mathbf{u}-\mathbf{u}_N\|_{H^1(\Omega)} \leqslant C \sup_{\mathbf{v}\in V} \|\mathbf{v}\|_{H^1(\Omega)}^{-1} \Big(-a(\mathbf{u}_N, \mathbf{v}) + L(\mathbf{v}) - b(\mathbf{v}, \tilde{\boldsymbol{\lambda}}_N) \Big), \tag{4.39}$$

where $\tilde{\boldsymbol{\lambda}}_N$ can be chosen arbitrarily in Λ satisfying

$$j(\mathbf{u}_N) = b(\mathbf{u}_N, \tilde{\boldsymbol{\lambda}}_N). \tag{4.40}$$

It is easy to see that such a $\tilde{\boldsymbol{\lambda}}_N$ exists, for example by choosing $\tilde{\boldsymbol{\lambda}}_N := \mathbf{u}_{N,t}/|\mathbf{u}_{N,t}|$ whenever $\mathbf{u}_{N,t} \neq 0$, and 0 otherwise. Inserting the $\boldsymbol{\lambda}_N \in \Lambda_N$ obtained by solving Problem 3.11, we obtain by the definition of the $H^{-1/2}$-norm that

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_N\|_{H^1(\Omega)} \leqslant C\Big[& \sup_{\mathbf{v}\in V} \|\mathbf{v}\|_{H^1(\Omega)}^{-1} \left[-a(\mathbf{u}_N, \mathbf{v}) + L(\mathbf{v}) - b(\mathbf{v}, \boldsymbol{\lambda}_N) \right] \\ &+ \Big\| \boldsymbol{\lambda}_N - \tilde{\boldsymbol{\lambda}}_N \Big\|_{H^{-1/2}(\Gamma_C)} \Big].\end{aligned} \tag{4.41}$$

Applying the definition of the discrete problem, we can insert $\mathbf{v}_N \in V_N$ and substitute $\mathbf{v}$ by $-\mathbf{v}$, which yields

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_N\|_{H^1(\Omega)} \leqslant C\Big[& \sup_{\mathbf{v}\in V} \|\mathbf{v}\|_{H^1(\Omega)}^{-1} \left[a(\mathbf{u}_N, \mathbf{v}-\mathbf{v}_N) - L(\mathbf{v}-\mathbf{v}_N) + b(\mathbf{v}-\mathbf{v}_N, \boldsymbol{\lambda}_N) \right] \\ &+ \Big\| \boldsymbol{\lambda}_N - \tilde{\boldsymbol{\lambda}}_N \Big\|_{H^{-1/2}(\Gamma_C)} \Big].\end{aligned} \tag{4.42}$$

Decomposing the integrals and integrating by parts on each element, we obtain, defining the vector divergence operator by $\mathbf{div}(\boldsymbol{\sigma}(\mathbf{u}_N)) := (\boldsymbol{\sigma}_{ji,j}(\mathbf{u}_N))_{i=1,\ldots,d}$,

$$\begin{aligned}&\|\mathbf{u}-\mathbf{u}_N\|_{H^1(\Omega)} \\ &\leqslant C\Bigg[\sup_{\mathbf{v}\in V} \frac{1}{\|\mathbf{v}\|_{H^1(\Omega)}} \sum_{K\in\mathcal{T}_N} \Bigg[\int_K [-\,\mathbf{div}\,\boldsymbol{\sigma}(\mathbf{u}_N) - \mathbf{F}](\mathbf{v}-\mathbf{v}_N)\mathrm{d}\mathbf{x} \\ &\qquad + \int_{\Omega\cap\partial K} \boldsymbol{\sigma}(\mathbf{u}_N)\cdot\boldsymbol{\nu}(\mathbf{v}-\mathbf{v}_N)\mathrm{d}s_{\mathbf{x}} \\ &\qquad + \int_{\Gamma_C\cap\partial K} [\boldsymbol{\sigma}(\mathbf{u}_N)_\tau + g(\boldsymbol{\lambda}_N)_\tau](\mathbf{v}-\mathbf{v}_N)\mathrm{d}s_{\mathbf{x}}\end{aligned} \tag{4.43}$$

$$+ \int_{\Gamma_\mathrm{N} \cap \partial K} [\boldsymbol{\sigma}(\mathbf{u}_N) \cdot \boldsymbol{\nu} - \mathbf{G}](\mathbf{v} - \mathbf{v}_N) \mathrm{d}s_\mathbf{x} \Bigg]$$

$$+ \left\| \boldsymbol{\lambda}_N - \tilde{\boldsymbol{\lambda}}_N \right\|_{\mathrm{H}^{-1/2}(\Gamma_\mathrm{C})} \Bigg].$$

Define for $K \in \mathcal{T}_N$ the interior residuals by

$$\mathbf{r}_K := -\,\mathbf{div}\, \boldsymbol{\sigma}(\mathbf{u}_N) - \mathbf{F} \tag{4.44}$$

and for $E \in \mathcal{E}_N$ the boundary residuals by

$$\mathbf{R}_E := \begin{cases} \frac{1}{2} \left[\boldsymbol{\sigma}(\mathbf{u}_N) \cdot \boldsymbol{\nu}\right]_E & \text{if } E \in \mathcal{E}_{\mathrm{I},N}, \\ \boldsymbol{\sigma}(\mathbf{u}_N)_\tau + g \boldsymbol{\lambda}_{N,\tau} & \text{if } E \in \mathcal{E}_{\mathrm{C},N} \\ \boldsymbol{\sigma}(\mathbf{u}_N) \cdot \boldsymbol{\nu} - \mathbf{G} & \text{if } E \in \mathcal{E}_{\mathrm{N},N}, \\ 0, & \text{if } E \in \mathcal{E}_{\mathrm{D},N}, \end{cases} \tag{4.45}$$

where

$$[\boldsymbol{\sigma}(\mathbf{u}_N) \cdot \boldsymbol{\nu}]_E := \boldsymbol{\sigma}(\mathbf{u}_N)|_{K_{E,1}} \cdot \boldsymbol{\nu}_{K_{E,1}} + \boldsymbol{\sigma}(\mathbf{u}_N)|_{K_{E,2}} \cdot \boldsymbol{\nu}_{K_{E,2}} \tag{4.46}$$

is the boundary jump with $E = K_{E,1} \cap K_{E,2}$ and $\boldsymbol{\nu}_{K_{E,1}}$ pointing from $K_{E,1}$ to $K_{E,2}$, and $\boldsymbol{\nu}_{K_{E,2}} = -\boldsymbol{\nu}_{K_{E,1}}$. Applying the Cauchy-Schwarz inequality and regrouping the interior boundary terms, we thus obtain

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(\Omega)} &\leqslant C \sup_{v \in V} \frac{1}{\|\mathbf{v}\|_{\mathrm{H}^1(\Omega)}} \sum_{K \in \mathcal{T}_N} \Bigg[\int_K \mathbf{r}_K \cdot (\mathbf{v} - \mathbf{v}_N) \mathrm{d}\mathbf{x} \\ &\qquad + \sum_{E \subseteq \partial K} \int_E \mathbf{R}_E \cdot (\mathbf{v} - \mathbf{v}_N) \mathrm{d}s_\mathbf{x} \Bigg] \\ &\leqslant C \sup_{v \in V} \frac{1}{\|\mathbf{v}\|_{\mathrm{H}^1(\Omega)}} \sum_{K \in \mathcal{T}_N} \Bigg[\|\mathbf{r}_K\|_{\mathrm{L}^2(K)} \|\mathbf{v} - \mathbf{v}_N\|_{\mathrm{L}^2(K)} \\ &\qquad + \sum_{E \subseteq \partial K} \|\mathbf{R}_E\|_{\mathrm{L}^2(E)} \|\mathbf{v} - \mathbf{v}_N\|_{\mathrm{L}^2(E)} \Bigg]. \end{aligned} \tag{4.47}$$

Plugging in the hp-Clément operator of Theorem 3.12 for $\mathbf{v}_N$,

$$\|\mathbf{v} - \mathbf{v}_N\|_{\mathrm{L}^2(K)} \leqslant C h_{K_\mathrm{patch}} p_{K_\mathrm{patch}}^{-1} \|\mathbf{v}\|_{\mathrm{H}^1(K_\mathrm{patch})}, \tag{4.48}$$

$$\|\mathbf{v} - \mathbf{v}_N\|_{\mathrm{L}^2(E)} \leqslant C h_{K_\mathrm{patch}}^{1/2} p_{K_\mathrm{patch}}^{-1/2} \|\mathbf{v}\|_{\mathrm{H}^1(K_\mathrm{patch})}. \tag{4.49}$$

Defining the *local error indicators* by

$$\eta_{N,K} := \left[h_K^2 p_K^{-2} \|\mathbf{r}_K\|_{\mathrm{L}^2(K)}^2 + h_K p_K^{-1} \sum_{E \subseteq \partial K} \|\mathbf{R}_E\|_{\mathrm{L}^2(E)}^2 \right]^{1/2} \tag{4.50}$$

and the *global error indicator* by

$$\eta_N := \left[\sum_{K \in \mathcal{T}_N} \eta_{N,K} \right]^{1/2}, \tag{4.51}$$

we obtain due to the finite overlap and local comparability of h and p (see the assumptions on the mesh in Section 3.2):

Theorem 4.3 (Reliability)**.** *There exists a constant $C > 0$ such that for all $\tilde{\boldsymbol{\lambda}}_N \in \Lambda$ with $j(\mathbf{u}_N) = b(\mathbf{u}_N, \tilde{\boldsymbol{\lambda}}_N)$, the residual error indicator satisfies*

$$\|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(\Omega)} \leqslant C \left[\eta_N + \left\| \boldsymbol{\lambda}_N - \tilde{\boldsymbol{\lambda}}_N \right\|_{\mathrm{H}^{-1/2}(\Gamma_{\mathrm{C}})} \right] \quad \textit{for all } N. \tag{4.52}$$

Therefore, if $\left\| \boldsymbol{\lambda}_N - \tilde{\boldsymbol{\lambda}}_N \right\|_{\mathrm{H}^{-1/2}(\Gamma_{\mathrm{C}})}$ is of higher order, η_N is reliable up to higher order terms. This can be expected if the mesh is chosen adaptively: Then, we actually presume that $|\boldsymbol{\lambda}_N| \leqslant 1$ on "most" of Γ_{C}, and there, we can choose $\tilde{\boldsymbol{\lambda}}_N = \boldsymbol{\lambda}_N$. Thus, as the elements where $|\boldsymbol{\lambda}_N| > 1$ should altogether have a size of order h, we should obtain a power of h in the above estimate, additionally to the rate obtained by estimating the error $\left\| \boldsymbol{\lambda}_N - \tilde{\boldsymbol{\lambda}}_N \right\|_{\mathrm{H}^{-1/2}(\Gamma_{\mathrm{C}})}$.

We now want to prove an efficiency result, which shows that the above error indicator will not overestimate the true error too much. Let $F_K \colon S \to K$ be the *element map* for K, that is, F_K is one-to-one and onto and bilinear, and assume that F_K maps I, interpreted as an edge of the reference element, to the edge E of K. Then, using the bubble functions on the reference interval and element given in Subsection 1.2.2, we define the *element bubble function* on K and the *edge bubble function* on E by

$$\psi_K := c_K \psi_S \circ F_K^{-1}, \quad \psi_E := c_E \psi_I \circ F_K^{-1}, \tag{4.53}$$

where the scaling factors c_K, $c_E > 0$ are chosen in such a way that

$$\int_K \psi_K \mathrm{d}\mathbf{x} = |K|, \quad \int_E \psi_E \mathrm{d}s_{\mathbf{x}} = |E|. \tag{4.54}$$

For the proof of the efficiency, we shall need the following lifting theorem, which is proved in [MW01, Lemma 2.6]. Recall that $S = [-1, 1]^2$ is the reference square.

Theorem 4.4. *Set $E := [-1, 1] \times \{-1\}$.*

For every $\alpha \in (1/2, 1]$, there exists a constant $C > 0$ such that for every p, every $\varepsilon \in (0, 1]$ and every polynomial $v \in \mathcal{P}^p$, there exists a function $\tilde{v} \in \mathrm{H}^1((-1, 1)^2)$ such that

$$\gamma_{0,E} \tilde{v} = v \cdot \psi_E^{\alpha}, \tag{4.55}$$

$$\gamma_{0,\partial S \backslash E} \tilde{v} = 0, \tag{4.56}$$

$$\|\tilde{v}\|^2_{\mathrm{L}^2((-1,1)^2)} \leqslant C\varepsilon \left\| \psi_I^{\alpha/2} v \right\|^2_{\mathrm{L}^2(-1,1)}, \tag{4.57}$$

$$\|\nabla \tilde{v}\|^2_{\mathrm{L}^2((-1,1)^2)} \leqslant C \left(\varepsilon p^{2(2-\alpha)} + \varepsilon^{-1} \right) \left\| \psi_I^{\alpha/2} v \right\|^2_{\mathrm{L}^2(-1,1)}. \tag{4.58}$$

Consider (2.71a) and (3.72a) and integrate by parts on each element to obtain

$$\begin{aligned} a(\mathbf{u} - \mathbf{u}_N, \mathbf{v}) &= a(\mathbf{u}, \mathbf{v}) - a(\mathbf{u}_N, \mathbf{v}) = L(\mathbf{v}) - b(\mathbf{v}, \boldsymbol{\lambda}) - a(\mathbf{u}_N, \mathbf{v}) \\ &= - \sum_{K \in \mathcal{T}_N} \left[\int_K \mathbf{r}_K \mathbf{v} \mathrm{d}\mathbf{x} + \sum_{E \subseteq \partial K} \int_E \mathbf{R}_E \mathbf{v} \mathrm{d}s_{\mathbf{x}} \right] + g \int_{\Gamma_{\mathrm{C}}} (\boldsymbol{\lambda}_N - \boldsymbol{\lambda}) \mathbf{v} \mathrm{d}s_{\mathbf{x}}. \end{aligned} \tag{4.59}$$

Choose $\beta \in (1/2, 1]$ arbitrary, but fixed. Let $\mathbf{v} := \psi_K^{\beta} \bar{\mathbf{r}}_K$, where $\bar{\mathbf{r}}_K$ is a polynomial approximation of $\mathbf{r}_K$ of degree $p_{N,K}$. Plugging this into (4.59) yields

$$a(\mathbf{u} - \mathbf{u}_N, \psi_K^{\beta} \bar{\mathbf{r}}_K) = - \int_K \mathbf{r}_K \psi_K^{\beta} \bar{\mathbf{r}}_K \mathrm{d}\mathbf{x}. \tag{4.60}$$

Thus,

$$\int_K \psi_K^\beta \bar{\mathbf{r}}_K^2 \mathrm{d}\mathbf{x} = \int_K \psi_K^\beta \bar{\mathbf{r}}_K (\bar{\mathbf{r}}_K - \mathbf{r}_K) \mathrm{d}\mathbf{x} - a(\mathbf{u} - \mathbf{u}_N, \psi_K^\beta \bar{\mathbf{r}}_K), \tag{4.61}$$

and the Cauchy-Schwarz inequality and the boundedness of a give, together with Theorem 1.6, which is applicable due to the fact that $\psi_K^\beta \bar{\mathbf{r}}_K = 0$ on Γ,

$$\begin{aligned}\int_K \psi_K^\beta \bar{\mathbf{r}}_K^2 \mathrm{d}\mathbf{x} &\leqslant \left\| \psi_K^{\beta/2} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)} \left\| (\bar{\mathbf{r}}_K - \mathbf{r}_K) \psi_K^{\beta/2} \right\|_{\mathrm{L}^2(K)} \\ &\quad + C \left\| \mathbf{u} - \mathbf{u}_N \right\|_{\mathrm{H}^1(K)} \left\| \psi_K \bar{\mathbf{r}}_K \right\|_{\mathrm{H}^1(K)} \\ &\leqslant \left\| \psi_K^{\beta/2} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)} \left\| (\bar{\mathbf{r}}_K - \mathbf{r}_K) \psi_K^{\beta/2} \right\|_{\mathrm{L}^2(K)} \\ &\quad + C \left\| \mathbf{u} - \mathbf{u}_N \right\|_{\mathrm{H}^1(K)} \left| \psi_K \bar{\mathbf{r}}_K \right|_{\mathrm{H}^1(K)} .\end{aligned} \tag{4.62}$$

Applying Theorem 1.22 together with a scaling argument, we see that

$$\begin{aligned}\left| \psi_K^\beta \bar{\mathbf{r}}_K \right|_{\mathrm{H}^1(K)}^2 = \left\| \nabla(\psi_K^\beta \bar{\mathbf{r}}_K) \right\|_{\mathrm{L}^2(K)}^2 &\leqslant 2 \left[\left\| (\nabla \psi_K^\beta) \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)}^2 + \left\| \psi_K^\beta \nabla \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)}^2 \right] \\ &\leqslant C \left[h_K^{-2} \left\| \psi_K^{\beta-1} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)}^2 + h_K^{-2} p_{N,K}^{2(2-\beta)} \left\| \psi_K^{\beta/2} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)}^2 \right] \\ &\leqslant C h_K^{-2} p_{N,K}^{2(2-\beta)} \left\| \psi_K^{\beta/2} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)}^2 .\end{aligned} \tag{4.63}$$

Inserting this in (4.62), we get

$$\begin{aligned}\left\| \psi_K^{\beta/2} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)} &\leqslant C \left[\left\| (\bar{\mathbf{r}}_K - \mathbf{r}_K) \psi_K^{\beta/2} \right\|_{\mathrm{L}^2(K)} + h_K^{-1} p_{N,K}^{2-\beta} \left\| \mathbf{u} - \mathbf{u}_N \right\|_{\mathrm{H}^1(K)} \right] \\ &\leqslant C \left[\left\| \bar{\mathbf{r}}_K - \mathbf{r}_K \right\|_{\mathrm{L}^2(K)} + h_K^{-1} p_{N,K}^{2-\beta} \left\| \mathbf{u} - \mathbf{u}_N \right\|_{\mathrm{H}^1(K)} \right].\end{aligned} \tag{4.64}$$

Finally, by the triangle inequality and Theorem 1.22,

$$\begin{aligned}\left\| \mathbf{r}_K \right\|_{\mathrm{L}^2(K)} &\leqslant \left\| \bar{\mathbf{r}}_K - \mathbf{r}_K \right\|_{\mathrm{L}^2(K)} + \left\| \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)} \\ &\leqslant \left\| \bar{\mathbf{r}}_K - \mathbf{r}_K \right\|_{\mathrm{L}^2(K)} + C p_K^\beta \left\| \psi_K^{\beta/2} \bar{\mathbf{r}}_K \right\|_{\mathrm{L}^2(K)} \\ &\leqslant C \left[(1 + p_K^\beta) \left\| \bar{\mathbf{r}}_K - \mathbf{r}_K \right\|_{\mathrm{L}^2(K)} + h_K^{-1} p_{N,K}^2 \left\| \mathbf{u} - \mathbf{u}_N \right\|_{\mathrm{H}^1(K)} \right] \\ &\leqslant C \left[p_K^\beta \left\| \bar{\mathbf{r}}_K - \mathbf{r}_K \right\|_{\mathrm{L}^2(K)} + h_K^{-1} p_{N,K}^2 \left\| \mathbf{u} - \mathbf{u}_N \right\|_{\mathrm{H}^1(K)} \right].\end{aligned} \tag{4.65}$$

As the next step, we shall estimate $\|\mathbf{R}_E\|_{\mathrm{L}^2(E)}$ for $E \in \mathcal{E}_N$. Let $\mathbf{v}$ be an extension to E_{patch} of $\psi_E^\beta \bar{\mathbf{R}}_E$ with $\bar{\mathbf{R}}_E$ a polynomial approximation of $\bar{\mathbf{R}}_E$ of degree $p_{N,E}$, constructed by applying Theorem 4.4 together with a scaling argument, and patching the results for the two neighbouring elements of E together. Plugging this into (4.59), we obtain

$$\begin{aligned}a(\mathbf{u} - \mathbf{u}_N, \psi_E^\beta \bar{\mathbf{R}}_E) = &- \sum_{K \subseteq E_{\mathrm{patch}}} \left[\int_K \mathbf{r}_K \psi_E^\beta \bar{\mathbf{R}}_E \mathrm{d}\mathbf{x} + \int_{E \cap \partial K} \mathbf{R}_E \psi_E^\beta \bar{\mathbf{R}}_E \mathrm{d}s_{\mathbf{x}} \right] \\ &+ g \int_{E \cap \Gamma_{\mathrm{C}}} (\boldsymbol{\lambda}_N - \boldsymbol{\lambda}) \psi_K^\beta \bar{\mathbf{R}}_E \mathrm{d}s_{\mathbf{x}}.\end{aligned} \tag{4.66}$$

Assume first that $E \in \mathcal{E}_{\mathrm{I},N}$. Then, $|E \cap \Gamma_{\mathrm{C}}| = 0$, the integrals over each E appear twice, and thus

$$\begin{aligned}\int_E \psi_E^\beta \bar{\mathbf{R}}_E^2 \mathrm{d}\mathbf{x} &= \int_E \psi_E^\beta \bar{\mathbf{R}}_E (\bar{\mathbf{R}}_E - \mathbf{R}_E) \mathrm{d}\mathbf{x} + \int_E \mathbf{R}_E \psi_E^\beta \bar{\mathbf{R}}_E \mathrm{d}s_{\mathbf{x}} \\ &= \int_E \psi_E^\beta \bar{\mathbf{R}}_E (\bar{\mathbf{R}}_E - \mathbf{R}_E) \mathrm{d}\mathbf{x} - \frac{1}{2} a(\mathbf{u} - \mathbf{u}_N, \psi_E^\beta \bar{\mathbf{R}}_E) \\ &\quad - \frac{1}{2} \int_{E_{\mathrm{patch}}} \mathbf{r}_{E_{\mathrm{patch}}} \psi_E^\beta \bar{\mathbf{R}}_E \mathrm{d}\mathbf{x}.\end{aligned} \tag{4.67}$$

The Cauchy-Schwarz inequality and the boundedness of a yield, together with Theorem 1.6,

$$\begin{aligned}\int_E \psi_E^\beta \bar{\mathbf{R}}_E^2 \mathrm{d}\mathbf{x} \leqslant{} & \left\| \psi_E^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)} \left\| (\bar{\mathbf{R}}_E - \mathbf{R}_E) \psi_E^{\beta/2} \right\|_{\mathrm{L}^2(E)} \\ & + C \Big[\|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(E_{\mathrm{patch}})} \left\| \psi_E^\beta \bar{\mathbf{R}}_E \right\|_{\mathrm{H}^1(E_{\mathrm{patch}})} \\ & \qquad + \left\| \mathbf{r}_{E_{\mathrm{patch}}} \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \left\| \psi_E^\beta \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \Big] \\ \leqslant{} & \left\| \psi_E^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)} \left\| (\bar{\mathbf{R}}_E - \mathbf{R}_E) \psi_E^{\beta/2} \right\|_{\mathrm{L}^2(E)} \\ & + C \Big[\|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(E_{\mathrm{patch}})} \left| \psi_E^\beta \bar{\mathbf{R}}_E \right|_{\mathrm{H}^1(E_{\mathrm{patch}})} \\ & \qquad + \left\| \mathbf{r}_{E_{\mathrm{patch}}} \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \left\| \psi_E^\beta \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \Big].\end{aligned} \tag{4.68}$$

Applying Theorem 4.4 together with a scaling argument, we obtain, as $\beta > 1/2$, for $\varepsilon > 0$, and as we assumed in Section 3.2 that the mesh is regular,

$$\left| \psi_E^\beta \bar{\mathbf{R}}_E \right|_{\mathrm{H}^1(E_{\mathrm{patch}})}^2 \leqslant C h_E^{-1} (\varepsilon p_{N,E}^{2(2-\beta)} + \varepsilon^{-1}) \left\| \psi_E^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)}^2, \tag{4.69}$$

$$\left\| \psi_E^\beta \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})}^2 \leqslant C h_E \varepsilon \left\| \psi_E^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)}^2, \tag{4.70}$$

and thus

$$\begin{aligned}\left\| \psi_E^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)} \leqslant C \Big[& \left\| \bar{\mathbf{R}}_E - \mathbf{R}_E \right\|_{\mathrm{L}^2(E)} + h_E^{-1/2} (\varepsilon p_{N,E}^{2(2-\beta)} + \varepsilon^{-1})^{1/2} \|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(E_{\mathrm{patch}})} \\ & + h_E^{1/2} \varepsilon^{1/2} \left\| \mathbf{r}_{E_{\mathrm{patch}}} \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \Big].\end{aligned} \tag{4.71}$$

Choosing $\varepsilon = p_E^{-2}$ yields

$$\begin{aligned}\left\| \psi^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)} \leqslant C \Big[& \left\| \bar{\mathbf{R}}_E - \mathbf{R}_E \right\|_{\mathrm{L}^2(E)} + h_E^{-1/2} p_{N,E} \|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(E_{\mathrm{patch}})} \\ & + h_E^{1/2} p_{N,E}^{-1} \left\| \mathbf{r}_{E_{\mathrm{patch}}} \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \Big].\end{aligned} \tag{4.72}$$

Using the triangle inequality, we obtain with Theorem 1.21 that

$$\begin{aligned}\|\mathbf{R}_E\|_{\mathrm{L}^2(E)} &\leqslant \left\| \bar{\mathbf{R}}_E - \mathbf{R}_E \right\|_{\mathrm{L}^2(E)} + \left\| \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)} \\ &\leqslant C \left[\left\| \bar{\mathbf{R}}_E - \mathbf{R}_E \right\|_{\mathrm{L}^2(E)} + p_{N,E}^\beta \left\| \psi_E^{\beta/2} \bar{\mathbf{R}}_E \right\|_{\mathrm{L}^2(E)} \right] \\ &\leqslant C \Big[p_{N,E}^\beta \left\| \bar{\mathbf{R}}_E - \mathbf{R}_E \right\|_{\mathrm{L}^2(E)} + h_E^{-1/2} p_{N,E}^{1+\beta} \|\mathbf{u} - \mathbf{u}_N\|_{\mathrm{H}^1(E_{\mathrm{patch}})} \\ &\qquad + h_E^{1/2} p_{N,E}^{-1+\beta} \left\| \mathbf{r}_{E_{\mathrm{patch}}} \right\|_{\mathrm{L}^2(E_{\mathrm{patch}})} \Big],\end{aligned} \tag{4.73}$$

and $\left\|\mathbf{r}_{E_{\text{patch}}}\right\|_{\mathrm{L}^2(E_{\text{patch}})}$ is estimated using (4.65), giving, due to the local comparability of h and p,

$$
\begin{aligned}
\|\mathbf{R}_E\|_{\mathrm{L}^2(E)} &\leqslant C\Big[p_{N,E}^{\beta}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}+h_E^{-1/2}p_{N,E}^{1+\beta}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(E_{\text{patch}})}\\
&\qquad+h_E^{1/2}p_{N,E}^{-1+\beta}\left\|\mathbf{r}_{E_{\text{patch}}}\right\|_{\mathrm{L}^2(E_{\text{patch}})}\Big]\\
&\leqslant C\Big[p_{N,E}^{\beta}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}+h_E^{-1/2}p_{N,E}^{1+\beta}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(E_{\text{patch}})}\\
&\qquad+h_E^{1/2}p_{N,E}^{-1+\beta}\big(p_{N,E}^{\beta}\left\|\bar{\mathbf{r}}_{E_{\text{patch}}}-\mathbf{r}_{E_{\text{patch}}}\right\|_{\mathrm{L}^2(E_{\text{patch}})}\\
&\qquad\qquad+h_E^{-1}p_{N,K}^{2}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(E_{\text{patch}})}\big)\Big]\\
&\leqslant C\Big[p_{N,E}^{\beta}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}+h_E^{-1/2}p_{N,E}^{1+\beta}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(E_{\text{patch}})}\\
&\qquad+h_E^{1/2}p_{N,E}^{-1+2\beta}\left\|\bar{\mathbf{r}}_{E_{\text{patch}}}-\mathbf{r}_{E_{\text{patch}}}\right\|_{\mathrm{L}^2(E_{\text{patch}})}\Big].
\end{aligned}
\tag{4.74}
$$

The proof for $E\in\mathcal{E}_{\mathrm{N},N}$ is done analogously. For $E\in\mathcal{E}_{\mathrm{C},N}$, we have to add the term corresponding to the contact boundary, $gp_{N,K}^{\beta}\|\boldsymbol{\lambda}_N-\boldsymbol{\lambda}\|_{\mathrm{L}^2(E)}$.

Plugging the results together, we obtain for $|\partial K\cap\Gamma_{\mathrm{C}}|=0$ due to the local comparability of h and p with an adequate element patch $K_{\text{patch}}\supseteq E_{\text{patch}}$ for all $E\subseteq\partial K$, as $\beta>1/2$,

$$
\begin{aligned}
\eta_{N,K}^2 &= h_K^2p_{N,K}^{-2}\|\mathbf{r}_K\|_{\mathrm{L}^2(K)}^2+h_Kp_{N,K}^{-1}\sum_{E\subseteq\partial K}\|\mathbf{R}_E\|_{\mathrm{L}^2(E)}^2\\
&\leqslant C\Big[h_K^2p_{N,K}^{-2}\left(p_{N,K}^{2\beta}\|\bar{\mathbf{r}}_K-\mathbf{r}_K\|_{\mathrm{L}^2(K)}^2+h_K^{-2}p_{N,K}^{4}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(K)}^2\right)\\
&\qquad+h_Kp_{N,K}^{-1}\sum_{E\subseteq\partial K}\Big(p_{N,K}^{2\beta}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}^2\\
&\qquad\qquad+h_K^{-1}p_{N,K}^{2(1+\beta)}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(E_{\text{patch}})}^2\\
&\qquad\qquad+h_Kp_{N,K}^{-2(1-2\beta)}\left\|\bar{\mathbf{r}}_{E_{\text{patch}}}-\mathbf{r}_{E_{\text{patch}}}\right\|_{\mathrm{L}^2(E_{\text{patch}})}^2\Big)\Big]\\
&\leqslant C\Big[(p_{N,K}^2+p_{N,K}^{1+2\beta})\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(K_{\text{patch}})}^2\\
&\qquad+h_K^2p_{N,K}^{-2(1-\beta)}(1+p_{N,K}^{2\beta-1})\left\|\bar{\mathbf{r}}_{K_{\text{patch}}}-\mathbf{r}_{K_{\text{patch}}}\right\|_{\mathrm{L}^2(K_{\text{patch}})}^2\\
&\qquad+h_Kp_{N,K}^{2\beta-1}\sum_{E\subseteq\partial K}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}\Big]\\
&\leqslant C\Big[p_{N,K}^{1+2\beta}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(K_{\text{patch}})}^2\\
&\qquad+h_K^2p_{N,K}^{-3+4\beta}\left\|\bar{\mathbf{r}}_{K_{\text{patch}}}-\mathbf{r}_{K_{\text{patch}}}\right\|_{\mathrm{L}^2(K_{\text{patch}})}^2\\
&\qquad+h_Kp_{N,K}^{2\beta-1}\sum_{E\subseteq\partial K}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}^2\Big]\\
&\leqslant Cp_{N,K}^{2\beta}\bigg(p_{N,K}\|\mathbf{u}-\mathbf{u}_N\|_{\mathrm{H}^1(K_{\text{patch}})}^2+h_K^2p_{N,K}^{-3+2\beta}\left\|\bar{\mathbf{r}}_{K_{\text{patch}}}-\mathbf{r}_{K_{\text{patch}}}\right\|_{\mathrm{L}^2(K_{\text{patch}})}\\
&\qquad+h_Kp_K^{-1}\sum_{E\subseteq\partial K}\left\|\bar{\mathbf{R}}_E-\mathbf{R}_E\right\|_{\mathrm{L}^2(E)}^2\bigg).
\end{aligned}
\tag{4.75}
$$

For $E \subseteq \partial K \cap \Gamma_{\mathrm{C}}$, we obtain in a similar fashion

$$\begin{aligned}\eta_{N,K}^2 \leqslant Cp_{N,K}^{2\beta}\Big(&p_{N,K}\left\|\mathbf{u}-\mathbf{u}_N\right\|_{\mathrm{H}^1(K_{\mathrm{patch}})}^2 + h_K^2 p_{N,K}^{-3+2\beta}\left\|\bar{\mathbf{r}}_{K_{\mathrm{patch}}} - \mathbf{r}_{K_{\mathrm{patch}}}\right\|_{\mathrm{L}^2(K_{\mathrm{patch}})} \\ &+ h_K p_K^{-1}\sum_{E\subseteq\partial K}\left\|\bar{\mathbf{R}}_E - \mathbf{R}_E\right\|_{\mathrm{L}^2(E)}^2 \\ &+ g^2 h_K p_K^{-1}\left\|\boldsymbol{\lambda}_N - \boldsymbol{\lambda}\right\|_{\mathrm{L}^2(\partial K\cap\Gamma_{\mathrm{C}})}^2\Big).\end{aligned} \tag{4.76}$$

Summing up, and noting that the term involving $\boldsymbol{\lambda}_N$ vanishes whenever $\partial K \cap \Gamma_{\mathrm{C}} = \varnothing$, we have:

Theorem 4.5 (Efficiency)**.** *There exists a constant $C > 0$ such that residual error indicator satisfies*

$$\begin{aligned}\eta_{N,K}^2 \leqslant Cp_{N,K}^{2\beta}\Big(&p_{N,K}\left\|\mathbf{u}-\mathbf{u}_N\right\|_{\mathrm{H}^1(K_{patch})}^2 + h_K^2 p_{N,K}^{-3+2\beta}\left\|\bar{\mathbf{r}}_{K_{patch}} - \mathbf{r}_{K_{patch}}\right\|_{\mathrm{L}^2(K_{patch})} \\ &+ h_K p_K^{-1}\sum_{E\subseteq\partial K}\left\|\bar{\mathbf{R}}_E - \mathbf{R}_E\right\|_{\mathrm{L}^2(E)}^2 + g^2 h_K p_K^{-1}\left\|\boldsymbol{\lambda}_N - \boldsymbol{\lambda}\right\|_{\mathrm{L}^2(\partial K\cap\Gamma_{\mathrm{C}})}^2\Big)\end{aligned} \tag{4.77}$$

for all N and $K \in \mathcal{T}_N$.

Thus, up to the term containing $\left\|\boldsymbol{\lambda}_N - \boldsymbol{\lambda}\right\|_{\mathrm{L}^2(E)}$, we obtain the same efficiency result as in [MW01]. Note, however, that the presence of this additional term is not surprising: It corresponds to the error done in the approximation of $\boldsymbol{\lambda}$, and is as such, at least from a heuristic point of view, acceptable in the efficiency estimate.

4.3 An *hp*-Adaptive Mesh Refinement Algorithm

We shall describe a mesh refinement algorithm which is aimed at producing good meshes for the problem under consideration. The basic approach in the refinement is always the one discussed, for example, in [BC04, p. 98], that is,

`SOLVE` ⇒ `ESTIMATE` ⇒ `MARK` ⇒ `REFINE`,

that is, we *solve* the discrete problem to find an approximate solution $\mathbf{u}_N \in V_N$, *estimate* the error based on the results of the preceding sections, *mark* those elements where the error indicator is high, and *refine* those elements. For the refinement itself, we use two distinct refinement strategies:

1. In the *h-adaptive refinement*, we divide all elements which have been marked into four new elements. For simplicity, we restrict ourselves to halving the quadrilaterals which are the basis of our partition of the domain in both directions simultaneously.

2. In the *hp-adaptive refinement*, we decide first whether to do a bisection of a marked element, or whether to increase the polynomial degree on the given element by one. Again, for simplicity, we increase the polynomial degree in both directions simultaneously.

A fundamental question in the theory of fully automatic *hp*-adaptive algorithms is how to decide whether an *h*- or a *p*-refinement should be applied on a given element. Several different strategies were compared in [EM07] for the case of triangular meshes; we shall use the approach of expanding the numerically computed solution into Legendre polynomials and estimating the decay of the coefficients, as was proposed therein and also in [HS05]. At the heart of this heuristic, we have the following result, which can be found in [Mel02, Lemma 3.2.7] and is based on the corresponding one-dimensional result given in [Dav63, Theorem 12.4.7].

Theorem 4.6. *Let $I := [-1, 1]$, $S := I^2$ and $u\colon S \to \mathbb{R}$ be real-analytic, and satisfy for some C_u, $\gamma > 0$, $h_x, h_y \in (0, 1]$*

$$\|D^\alpha u\|_{\mathrm{L}^\infty((-1,1)^2)} \leqslant C_u h_x^{\alpha_1} h_y^{\alpha_2} \gamma^{|\alpha|} \alpha! \quad \text{for all } \alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0^2 \backslash \{(0, 0)\}. \tag{4.78}$$

Then, u can be expanded in a Legendre series on S, and there are C, $\sigma > 0$ depending only on γ such that

$$u(x, y) = \sum_{i,j=0}^{\infty} u_{ij} L_i(x) L_j(y) \quad \text{uniformly on } S, \tag{4.79}$$

$$|u_{ij}| \leqslant C_u C (1 + \sigma/h_x)^{-i} (1 + \sigma/h_y)^{-j} \quad \text{for } (i, j) \neq (0, 0). \tag{4.80}$$

In particular, setting $b := \ln(\min(1 + \sigma/h_x, 1 + \sigma/h_y))$, we see that for an analytic function u, the Legendre coefficients u_{ij} satisfy

$$|u_{ij}| \leqslant C_u C \exp(-b(i + j)). \tag{4.81}$$

Calculating b from the given Legendre coefficients of the local approximation on a single element with the above formula therefore gives a heuristic estimate of how regular the function is locally: We expect that the behaviour of the true, unknown Legendre coefficients is reflected by the known Legendre coefficients of the given approximation.

We prove a result showing that based on b, we obtain fast convergence of the local polynomial approximations, which shows heuristically why the above approach is reasonable. We follow the proof for [Mel02, Proposition 3.2.8], but in contrast to that result, we are actually working on rectangular elements and can therefore directly use the L^2-projection operator, which is for u as above given by

$$\Pi_p^{\mathrm{L}^2((-1,1)^2)} u(x, y) := \sum_{i,j=0}^{p} u_{ij} L_i(x) L_j(y). \tag{4.82}$$

Theorem 4.7. *Under the assumptions of the last theorem, there exist constants $C > 0$, $\sigma > 0$ depending only on $\gamma > 0$ such that*

$$\begin{aligned} \left\| u - \Pi_p^{\mathrm{L}^2((-1,1)^2)} u \right\|_{\mathrm{L}^\infty((-1,1)^2)} &+ \left\| \partial_x \left(u - \Pi_p^{\mathrm{L}^2((-1,1)^2)} u \right) \right\|_{\mathrm{L}^\infty((-1,1)^2)} \\ &+ \left\| \partial_y \left(u - \Pi_p^{\mathrm{L}^2((-1,1)^2)} u \right) \right\|_{\mathrm{L}^\infty((-1,1)^2)} \\ &\leqslant C C_u \left[\left(\frac{h_x}{h_x + \sigma} \right)^{p+1} + \left(\frac{h_y}{h_y + \sigma} \right)^{p+1} \right]. \end{aligned} \tag{4.83}$$

The proof uses the following version of Markov's inequality, which can be found in [DL93, Theorem 4.1.4].

Theorem 4.8 (Markov's inequality)**.** *For all polynomials* $v \in \mathcal{P}^p$*,*

$$\left\|v'\right\|_{\mathrm{L}^\infty(-1,1)} \leqslant p^2 \left\|v\right\|_{\mathrm{L}^\infty(-1,1)} . \tag{4.84}$$

Proof of Theorem 4.7. We only prove the estimate for $\partial_x \left(u - \Pi_p^{\mathrm{L}^2((-1,1)^2)} u\right)$; the other inequalities follow similarly. Clearly,

$$\begin{aligned}\left\|\partial_x \left(u - \Pi_p^{\mathrm{L}^2((-1,1)^2)} u\right)\right\|_{\mathrm{L}^\infty((-1,1)^2)} \leqslant &\sum_{\substack{i\geqslant p+1\\ j\geqslant 0}} |u_{ij}| \left\|L_i'\right\|_{\mathrm{L}^\infty(-1,1)} \left\|L_j\right\|_{\mathrm{L}^\infty(-1,1)} \\ &+ \sum_{\substack{i\geqslant 0\\ j\geqslant p+1}} |u_{ij}| \left\|L_i'\right\|_{\mathrm{L}^\infty(-1,1)} \left\|L_j\right\|_{\mathrm{L}^\infty(-1,1)} .\end{aligned} \tag{4.85}$$

By (4.84) and (4.80), after setting $\alpha_x := (1+\sigma/h_x)^{-1}$ and $\alpha_y := (1+\sigma/h_y)^{-1}$ and decreasing σ as needed to absorb the additional factor i^2,

$$\begin{aligned}\left\|\partial_x \left(u - \Pi_p^{\mathrm{L}^2((-1,1)^2)} u\right)\right\|_{\mathrm{L}^\infty((-1,1)^2)} &\leqslant C_u C \left[\sum_{\substack{i\geqslant p+1\\ j\geqslant 0}} \alpha_x^i \alpha_y^j + \sum_{\substack{i\geqslant 0\\ j\geqslant p+1}} \alpha_x^i \alpha_y^j\right] \\ &= C_u C \left[\frac{\alpha_x^{p+1}}{1-\alpha_x} \sum_{j=0}^{\infty} \alpha_y^j + \frac{\alpha_y^{p+1}}{1-\alpha_y} \sum_{i=0}^{\infty} \alpha_x^i\right] \\ &= \frac{C_u C}{(1-\alpha_x)(1-\alpha_y)} \left[\alpha_x^{p+1} + \alpha_y^{p+1}\right],\end{aligned} \tag{4.86}$$

and the result follows. □

Chapter 5

Numerical Experiments

We now present some numerical experiments. The implementation was done using Fortran 90 and is based on the software package `maiprogs` by Matthias Maischak (see [Mai01b]). The problem descriptions were given in the `maiprogs` control language `BCL`.

The discrete problems which are solved numerically are those given in Problem 3.32. We need to chooses bases of the spaces V_N and W_N. For V_N, we choose tensor products of antiderivatives of Legendre polynomials piecewise with respect to the mesh, corrected on the inter-element boundaries by the minimum rule to deal with differing polynomial degrees and hanging nodes. For W_N, we are able to use discontinuous basis functions, and thus, we select Lagrange interpolation polynomials at the shifted and scaled Gauss points on each boundary piece on Γ_C. Denoting these bases as $(\mathbf{v}_i)_{i=1,\dots,N}$ and $(\mathbf{w}_j)_{j=1,\dots,N'}$, respectively, we define the matrices $A^{(N)} \in \mathbb{R}^{N\times N}$ and $B^{(N)} \in \mathbb{R}^{N'\times N}$ by

$$A^{(N)}_{kl} := a(\mathbf{v}_l, \mathbf{v}_k) \quad \text{and} \quad B^{(N)}_{mn} := b(\mathbf{v}_n, \mathbf{w}_m) \tag{5.1}$$

and the vector $f^{(N)} \in \mathbb{R}^N$ by

$$f^{(N)}_k := L(\mathbf{v}_k), \tag{5.2}$$

then we see that in matrix notation, we obtain:

Problem 5.1 (Discrete primal-dual formulation, Tresca friction, forced contact, matrix formulation)**.** Find $(x^{(N)}, z^{(N)}) \in \mathbb{R}^N \times [-1,+1]^{N'}$ such that for all $w^{(N)} \in [-1,+1]^{N'}$,

$$A^{(N)}x^{(N)} + {B^{(N)}}^T z^{(N)} = f^{(N)}, \tag{5.3}$$

$$(w^{(N)} - z^{(N)})B^{(N)}x^{(N)} \leqslant 0. \tag{5.4}$$

It is then easy to see that Problem 5.1 has a unique solution, which corresponds to the unique solution of Problem 3.32 through

$$\mathbf{u}_N = \sum_{i=1}^{N} x_i^{(N)}\mathbf{v}_i \quad \text{and} \quad \boldsymbol{\lambda}_N = \sum_{j=1}^{N'} z_j^{(N)}\mathbf{w}_j. \tag{5.5}$$

To be able to solve this problem efficiently, we use a similar approach as proposed in [Sin06]: Defining the *Schur complement matrix* by

$$S^{(N)} := B^{(N)}{A^{(N)}}^{-1}{B^{(N)}}^T, \tag{5.6}$$

we have that the above system corresponds to

$$(w^{(N)} - z^{(N)})S^{(N)}z^{(N)} \geqslant (w^{(N)} - z^{(N)})B^{(N)}{A^{(N)}}^{-1}f^{(N)} \quad \text{for all } w^{(N)} \in [-1,+1]^{N'}. \tag{5.7}$$

Due to the inf-sup condition proved in Theorem 3.33 and the symmetry and positive definiteness of $A^{(N)}$, we see that $S^{(N)}$ is symmetric and positive definite. Thus, the reduced variational inequality (5.7) admits a unique solution for $z^{(N)}$, and given this solution, we can find $x^{(N)}$ by solving (5.3). The numerical computation of $S^{(N)}$ is done by first calculating a Cholesky factorisation of $A^{(N)}$ using the PARDISO library and METIS (see [SG02, SG06, KK98]).

We are therefore led to the question of how to solve the variational inequality (5.7) efficiently. As detailed in [Sin06, Section 4.2], the MPRGP algorithm given in [DS05] can be generalised to such a problem with two-sided constraints; see Section 5.1 for a short review of the algorithm.

We give some numbers below showing that for the *hp*-version, the matrices $S^{(N)}$ are very ill-conditioned, but diagonal scaling (which can easily be applied, even with variational inequalities) was sufficient for the two-dimensional problems under consideration (where the boundary is one-dimensional), as here, N', and thus the dimensionality of the variational inequality, stays relatively small.

For the numerical problems under consideration, we consider *h*-uniform and *h*-adaptive versions with polynomial degrees 2, 3 and 4 (`hup2`, `hup3` and `hup4` and `hap2`, `hap3` and `hap4`, respectively), the uniform *p*-version (`pu`), and two different *hp*-adaptive versions. All methods use the same initial mesh.

In the *h*-adaptive methods, we refine all elements where the local error indicator is larger than 1/2 times the mean local error indicator.

The first *hp*-adaptive version (`hpa1`) is similar as in [MS05], and works by sorting elements by the local error indicator and *h*-refining the first and *p*-refining the second ten percent. The second *hp*-adaptive version (`hpa2`) works, as described in more detail in Section 4.3, by estimating the decay of the coefficients in the local expansion of the numerical solution into tensor products of Legendre polynomials. We refine the 20 percent of the elements with the highest local error indicator. Calculating b, which is given as in (4.81), by a linear regression, we do a *p*-refinement if $b > 1$, and an *h*-refinement otherwise. Here, we start with a uniform polynomial degree of 3 to be able to obtain a useful estimate of the rate of decay.

We do not consider methods using a polynomial degree of 1, as the convergence rate results in Subsection 3.2.3 only apply for splines of local polynomial degree $\geqslant 2$.

Note that there are no exact solutions for all problems we consider below. Therefore, we cannot give any estimates including the actual error, but only the estimated error as given by the residual error indicator. This means, in particular, that for all *p*-refining methods, the given numbers might overestimate the actual error by as much as a factor p in addition to all effects resulting from the terms in the reliability and efficiency estimates which cannot be directly controlled.

Due to the fact that we do adaptive calculations, we plot all error estimates against the number of degrees of freedom on the domain, and also use these numbers when giving convergence rates. This means that we do not include the number of degrees of freedom for the Lagrange multiplier space in the analysis below, but this should not be significant, as the mesh on Γ_C is constructed from the mesh on Ω, and the bulk of the total problem size stems from the domain discretisation; in particular, the total number of degrees of freedom $\dim V_N + \dim W_N$ can be bounded by $2N$, where $N = \dim V_N$.

5.1 The MPRGP Algorithm with Two-Sided Constraints

The MPRGP algorithm as used in our numerical computations is given in Algorithm 1.The functions $\beta := (\beta_j)_{j=1}^N$, $\varphi := (\varphi_j)_{j=1}^N$ and $\nu := (\nu_j)_{j=1}^N$ depend on a vector $y \in \mathbb{R}^N$ (for which we always plug in the current approximation $x \in \mathbb{R}^N$) and on the residual $r = Ax - b$ and are defined by

$$\beta_j(y) := \begin{cases} \min(r_j, 0), & y_j = \ell_j, \\ \max(r_j, 0), & y_j = u_j, \\ 0, & \text{otherwise}, \end{cases} \tag{5.8}$$

$$\varphi_j(y) := \begin{cases} r_j, & \ell_j < y_j < u_j, \\ 0, & \text{otherwise}, \end{cases} \tag{5.9}$$

and

$$\nu_j(y) := \beta_j(y) + \varphi_j(y). \tag{5.10}$$

For $z \in \mathbb{R}^N$, we define $\tilde{z} \in \mathbb{R}^N$, which depends on the current approximation $x \in \mathbb{R}^N$, by

$$\tilde{z} := \bar{\alpha}^{-1} \left(x - \max(\ell, \min(u, x - \bar{\alpha} * z)) \right). \tag{5.11}$$

5.2 Problems and Results

We analyse the following two problems, which are both taken from [Han05, Section 6.6].

Example 5.2 ([Han05, Example 6.12]). Consider the domain $\Omega := (0, 4) \times (0, 4)$, where space is measured in millimeters, under the plane stress assumption, with the boundary decomposed into the Dirichlet boundary $\Gamma_\mathrm{D} := \{4\} \times (0, 4)$, the Neumann boundary $\Gamma_\mathrm{N} := (\{0\} \times (0, 4)) \cup ((0, 4) \times \{4\})$, and the contact boundary $\Gamma_\mathrm{C} := (0, 4) \times \{4\}$. The elastic constants are $E = 15\mathrm{kN/mm}^2$ and $\nu = 0.4$, the frictional constant is $g = 4.5\mathrm{kN/mm}^2$. The volume forces vanish, $\mathbf{F} := (0, 0)\mathrm{kN/mm}^2$, and on the Neumann boundary, we have $\mathbf{G}(x_1, x_2) := (1.5(5 - x_2), -.75)\mathrm{kN/mm}^2$ on $\{0\} \times (0, 4)$, and no surface forces on $(0, 4) \times \{4\}$, $\mathbf{G} := (0, 0)\mathrm{kN/mm}^2$.

The convergence plot is given in Figure 5.1, and Figures 5.2 and 5.3 show the deformed mesh and Lagrange multiplier for the second *hp*-adaptive method, which can be expected to deliver the best results. The advantage of the adaptive methods compared to the uniform methods is very clear: Neither the uniform *h*-versions nor the uniform *p*-version can deliver a satisfying convergence rate, but the adaptive *h*-methods give good results.

It seems that the second *hp*-adaptive method `hpa2` yields an exponential convergence rate. Numerically, assuming that the error approximately follows a behaviour of the type $e_N \approx C \exp(-\beta N^{1/3})$ and doing a linear least squares estimate for $\ln e_N = \ln C - \beta N^{1/3}$, we obtain $\beta \approx 0.3476$. Such a behaviour is expected from *hp*-FEM in two space dimensions for appropriate meshes; see [GB86]. The condition numbers of the Schur complement become as high as $5.5 \cdot 10^8$ for 22690 degrees of freedom for $\mathbf{u}_N$; by a diagonal scaling, however, it is reduced to $2.7 \cdot 10^3$. Table 5.1 gives the number of elements with the different polynomial degrees. Figure 5.4 shows a zoom of the mesh in the final refinement step, together with

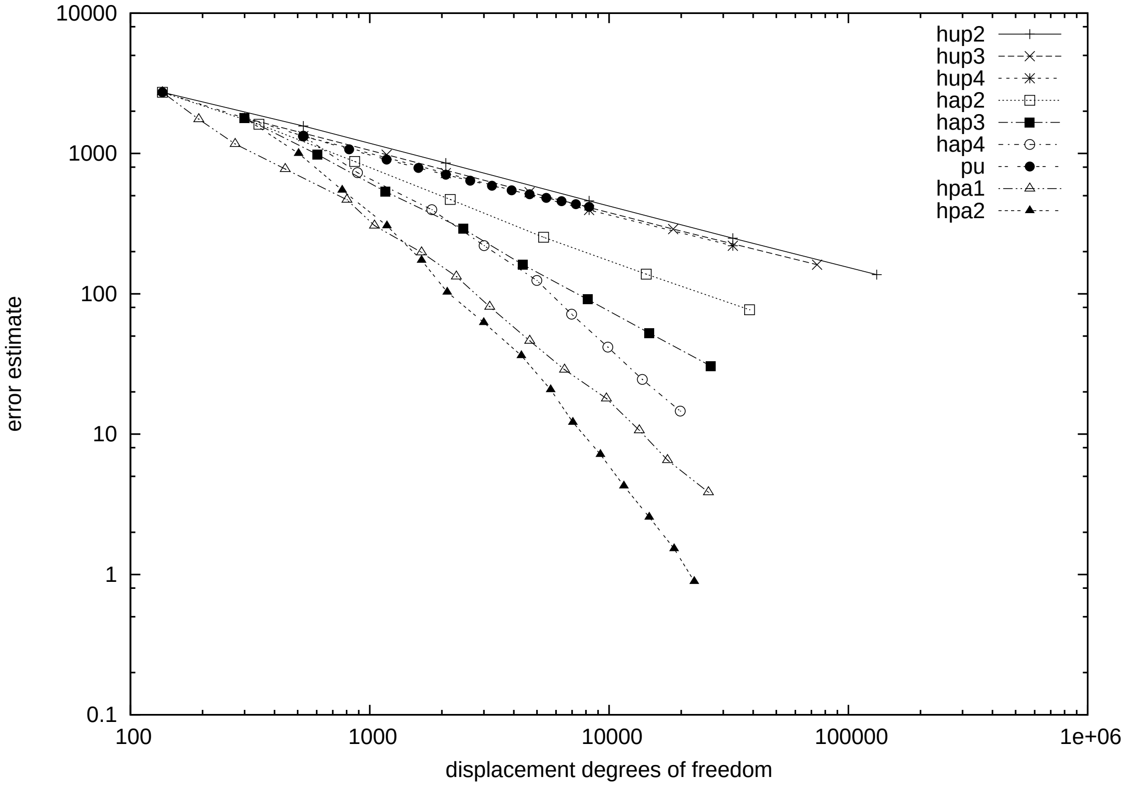

Figure 5.1: Error plot, Example 5.2

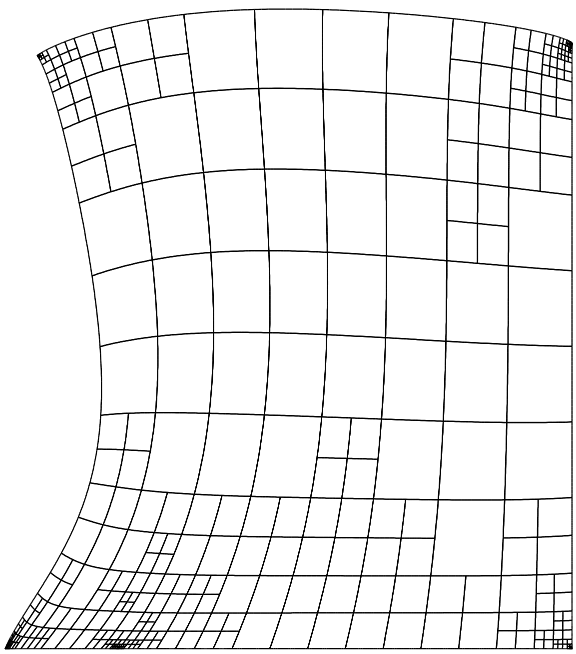

Figure 5.2: Deformed mesh, Example 5.2, method `hpa2`

Algorithm 1: MPRGP algorithm for minimising $\frac{1}{2}x^T Ax - x^T b$ subject to $u \geqslant x \geqslant \ell$

Data: $A \in \mathbb{R}^{N\times N}$ positive definite, $b, u, \ell \in \mathbb{R}^N$ with $\ell < u$, parameter $\bar{\alpha} > 0$
Result: minimiser $x \in \mathbb{R}^N$ of $\frac{1}{2}x^T Ax - x^T b$

$x := 0$;
$r := b$;
$p := \varphi(x)$;
while $|\nu(x)| > \varepsilon$ **do**
- **if** $\beta(x) \cdot \tilde{\beta}(x) < \Gamma^2 \varphi(x) \cdot \tilde{\varphi}(x)$ **then**
  - `// trial CG step`
  - $\alpha_{\mathrm{CG}} := \frac{p\cdot r}{p\cdot Ap}$;
  - $\alpha_{\min} := \min\{\alpha\colon \ell \leqslant x - \alpha p \leqslant u\}$;
  - $\alpha_{\max} := \max\{\alpha\colon \ell \leqslant x - \alpha p \leqslant u\}$;
  - **if** $\alpha_{\min} \leqslant \alpha_{\mathrm{CG}} \leqslant \alpha_{\max}$ **then**
    - `// CG step`
    - $x := x - \alpha_{\mathrm{CG}} * p$;
    - $r := r - \alpha_{\mathrm{CG}} * Ap$;
    - $p := \varphi(x) - \frac{\varphi(x)\cdot Ap}{p\cdot Ap}$;
  - **else**
    - `// expansion step`
    - **if** $\alpha_{\mathrm{CG}} < \alpha_{\min}$ **then**
      - $x := x - \alpha_{\min} * p$;
      - $r := r - \alpha_{\min} * Ap$;
    - **else**
      - `//` $\alpha_{\mathrm{CG}} > \alpha_{\max}$
      - $x := x - \alpha_{\max} * p$;
      - $r := r - \alpha_{\max} * Ap$;
    - **end**
    - $x := \max(\ell, \min(u, x - \bar{\alpha} * \varphi(x))$;
    - $r := Ax - b$;
    - $p := \varphi(x)$;
  - **end**
- **else**
  - `// proportioning step`
  - $\alpha_{\mathrm{CG}} := \frac{r\cdot\beta(x)}{\beta(x)\cdot A\beta(x)}$;
  - $x := \max(\ell, \min(u, x - \alpha_{\mathrm{CG}}\beta(x))$;
  - $r := Ax - b$;
  - $p := \varphi(x)$;
- **end**

end

the polynomial degrees, near the point of the contact boundary where the transition from sticking to sliding happens.

The uniform h-versions all deliver convergence rates of about 0.45, which corresponds to the fact that the solution is not very regular. The adaptive h-versions with polynomial degree 2, 3 and 4 deliver convergence rates of 0.6, 0.9 and 1.5, respectively, which is still not optimal, but significantly better. The uniform p-version only yields a convergence rate of

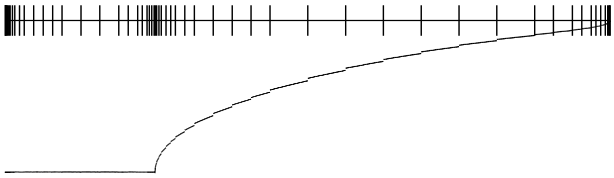

Figure 5.3: Lagrange multiplier, Example 5.2, method hpa2

refinement step	$\|\mathcal{T}\|$	3	4	5	6	7	8
1	16	16					
2	28	28					
3	43	42	1				
4	64	58	6				
5	85	70	15				
6	97	66	28	3			
7	127	78	41	6	2		
8	157	79	57	12	9		
9	196	83	83	16	14		
10	226	75	107	25	19		
11	271	67	131	49	22	2	
12	319	59	165	61	26	8	
13	358	54	158	87	36	19	4
14	418	40	180	113	44	37	4
15	472	34	176	145	64	44	9

Table 5.1: Number of elements of different polynomial degrees, Example 5.2, method hpa2

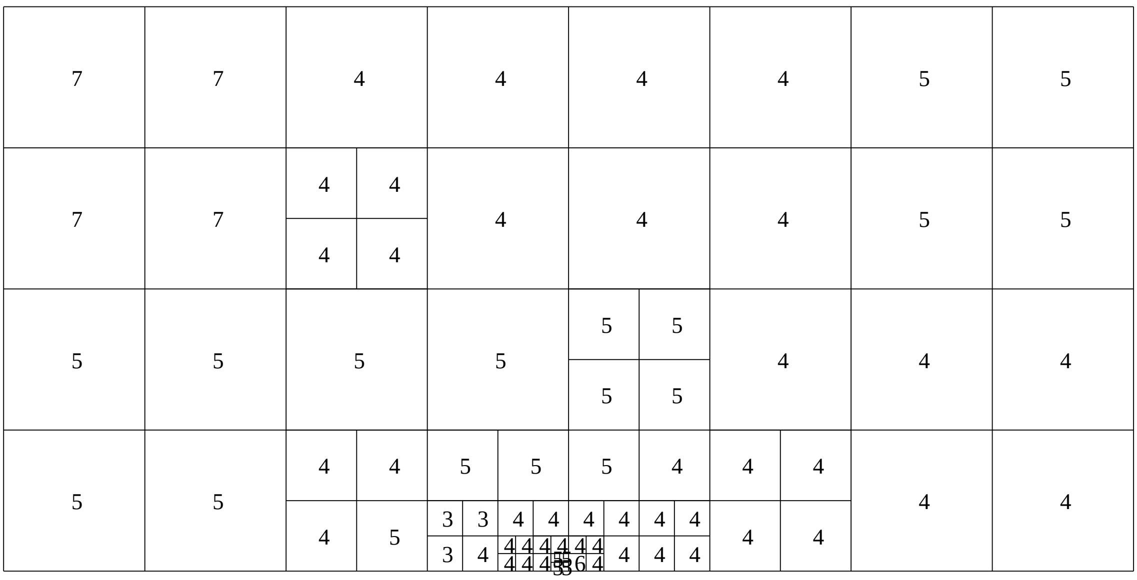

Figure 5.4: Polynomial degree distribution, zoom onto $[0.5, 1.5] \times [0.0, 0.5]$, Example 5.2, method hpa2, final refinement step

about 0.35, which can be conjectured to stem from the fact that a uniform p-mesh will not be appropriate do decompose the contact boundary properly into the sliding and the sticking region.

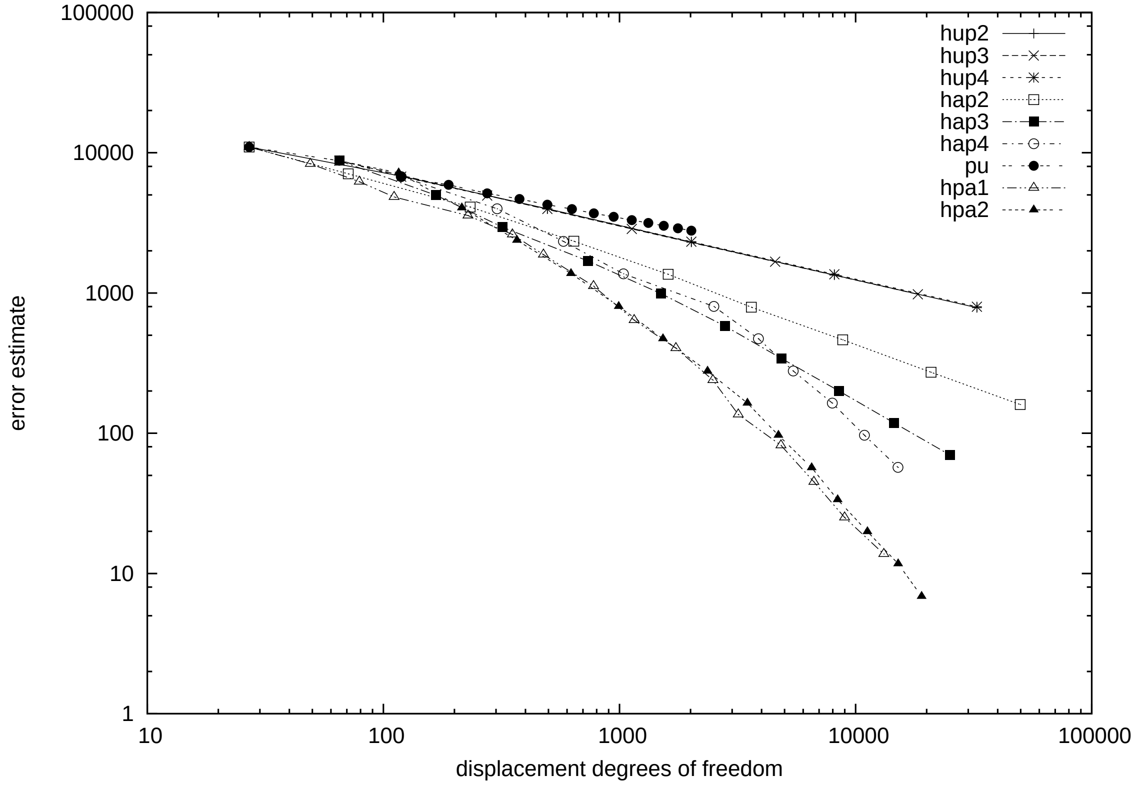

Figure 5.5: Error plot, Example 5.3

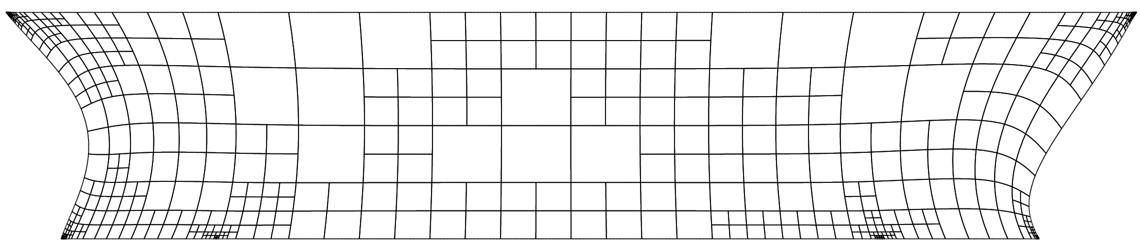

Figure 5.6: Deformed mesh, Example 5.3, method `hpa2`

Example 5.3 ([Han05, Example 6.13]). In the second example, we again use millimeters as spacial unit, and consider $\Omega := (0, 10) \times (0, 2)$ with plane stress, $\Gamma_D := (0, 10) \times \{2\}$, $\Gamma_N := (\{0\} \times (0, 2)) \cup (\{10\} \times (0, 2))$ and $\Gamma_C := (0, 10) \times \{0\}$. The elastic constants are $E = 10\text{kN/mm}^2$ and $\nu = 0.3$, the frictional constant is $g = 1.75\text{kN/mm}^2$. The volume forces vanish, $\mathbf{F} := (0, 0)\text{kN/mm}^2$, the surface forces on $\{0\} \times (0, 2)$ are $\mathbf{G}(x_1, x_2) := (5, 0)\text{kN/mm}^2$, and on $\{10\} \times (0, 2)$, $\mathbf{G}(x_1, x_2) := (2.5x_2 - 7.5, -1)\text{kN/mm}^2$.

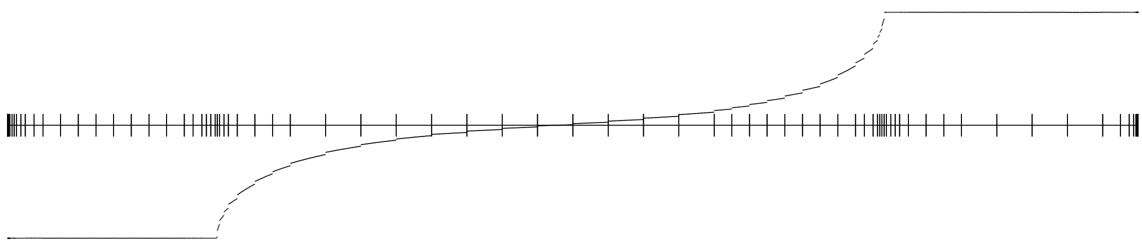

Figure 5.7: Lagrange multiplier, Example 5.3, method `hpa2`

refinement step	$\|\mathcal{T}\|$	3	4	5	6	7
1	4	4				
2	7	7				
3	13	13				
4	22	22				
5	37	37				
6	58	57	1			
7	85	78	7			
8	124	104	20			
9	175	135	40			
10	226	161	63	2		
11	280	171	94	11	4	
12	331	170	132	22	7	
13	400	167	173	45	14	1
14	502	189	211	74	24	4
15	580	170	254	116	33	7

Table 5.2: Number of elements of different polynomial degrees, Example 5.3, method `hpa2`

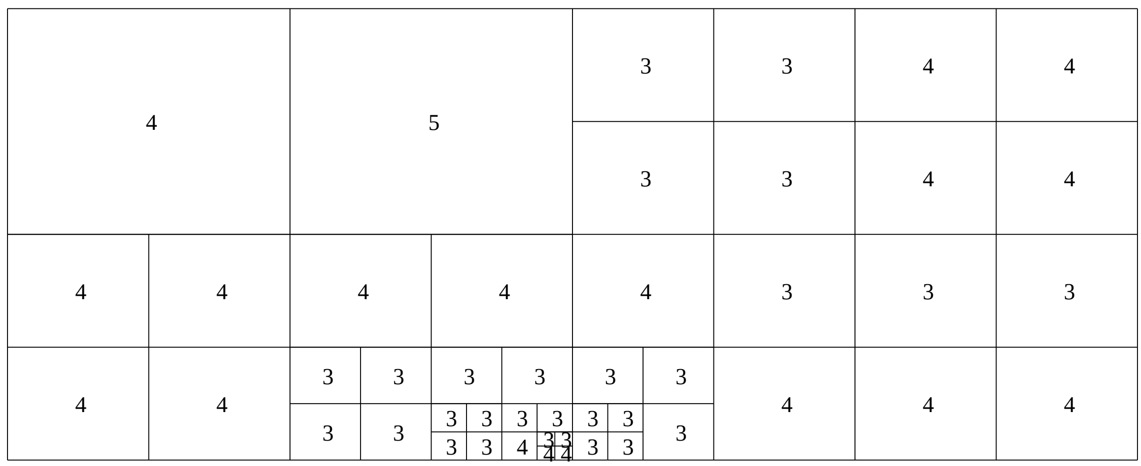

Figure 5.8: Polynomial degree distribution, zoom onto $[1.25, 2.5] \times [0.0, 0.5]$, Example 5.3, method `hpa2`, final refinement step

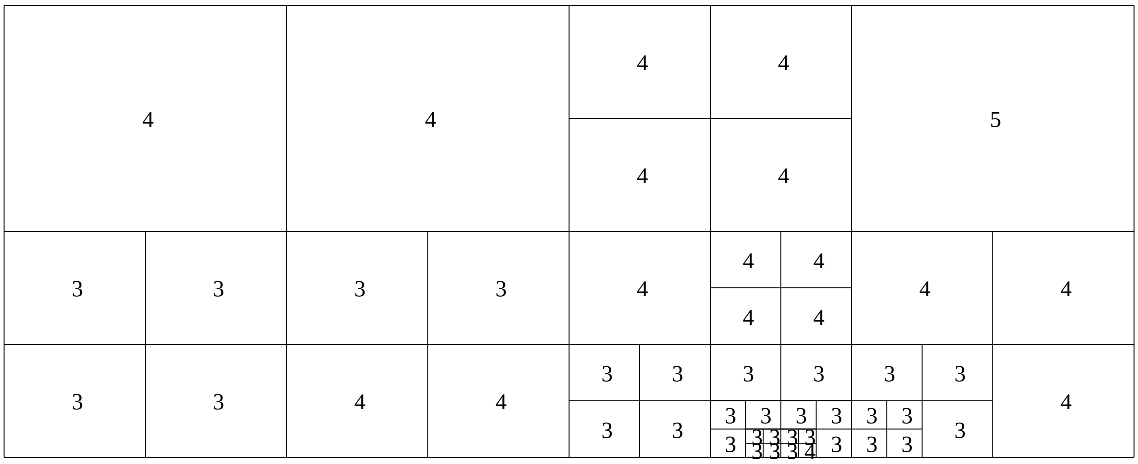

Figure 5.9: Polynomial degree distribution, zoom onto $[6.875, 8.125] \times [0.0, 0.5]$, Example 5.3, method `hpa2`, final refinement step

The convergence plot is shown in Figure 5.5, and Figures 5.6 and 5.7 give the deformed mesh and Lagrange multiplier for the second *hp*-adaptive version. Again, only the adaptive methods yield acceptable convergence rates, the uniform methods converge too slowly.

Here, the second *hp*-adaptive version appears to yield an exponential convergence rate, and for an error behaviour of the form $e_N \approx C \exp(-\beta N^{1/3})$, linear least squares for $\ln e_N$ yields $\beta \approx 0.3167$. The condition number of the unmodified Schur complement for 19045 degrees of freedom on the domain is $5.3 \cdot 10^7$; after diagonal scaling, we obtain a condition number of $1.2 \cdot 10^3$. Table 5.2 gives the number of elements with the different polynomial degrees. Figures 5.8 and 5.9 show zooms of the mesh in the final refinement step, together with the polynomial degrees, near the two points of the contact boundary where the transition from sticking to sliding happens.

The uniform *h*-versions all deliver a convergence rate of about 0.38. The adaptive *h*-versions with polynomial degrees 2, 3 and 4 yield convergence rates of 0.6, 0.95 and 1.6, respectively, again a significant enhancement. The uniform *p*-version delivers a convergence rate of about 0.3, again significantly worse than all the other methods.

5.3 Comments

When comparing different choices for the mesh of a finite element method, it is not only relevant to compare the errors of the different methods, but also the computational work. In our numerical experiments, it showed that the different adaptive methods take a comparable time, as the main task is calculating a Cholesky decomposition of the stiffness matrix $A^{(N)}$ while constructing the Schur complement, and the iterations of the MPRGP algorithm are rather negligible due to the fact that the Lagrange multiplier only has about 300 degrees of freedom. Therefore, we can recommend the fully automatic second *hp*-adaptive strategy proposed above for practical computations.

Compared to the results given in [Han05], we see that the higher-order elements we employ yield significantly better convergence rates in adaptive computations. This is what we expected from the a priori error estimate in Theorem 3.35, together with the reliability and efficiency of the error indicator given in Theorems 4.3 and 4.5. The non-optimal terms in these results might also be the reason why we do not obtain the full approximation rates expected from the h-adaptive methods used. Note here, however, that due to the fact that we use a smaller polynomial degree for the approximation of $\boldsymbol{\lambda}$, we can only expect a rate of up to $N^{1/2p-1/4}$ for a pure h-version; but even this, lower, rate is not fully attained. For hp-versions, the reduced rate for the approximation of the Lagrange multiplier is irrelevant, as we anyways expect exponential convergence.

Conclusion

In this diploma thesis, we gave an account of the mathematical theory and numerical analysis of an elastic contact problem with friction, focussing on the latter. In Chapter 2, we saw that the existence and uniqueness theory for the problem with Tresca friction is simplified by the fact that we only have to deal with a problem of convex minimisation.

The numerical analysis, however, is more complicated. Due to the presence of the non-differentiable functional, which contains an integration which cannot be done exactly when using polynomials of higher degrees, we use a primal-dual formulation in the discretisation. This allows us to use a very elegant approach for the actual solution of the discrete problem, based on the Schur complement of the system matrix and using the very powerful MPRGP algorithm developed in [DS05], as suggested in [Sin06]. The primal-dual formulation leads to certain difficulties in the proof of an a priori convergence rate result. Using a new kind of *hp*-mortar projection operator, constructed in Subsection 3.2.2, we can remedy these problems, and obtain an error estimate given in Theorem 3.33 which shows (Theorem 3.35) that a well-chosen *hp*-mesh can be expected to yield exponential convergence.

In Chapter 4, we use the duality approach suggested in [Han05] to construct a variant of the residual error indicator for the displacement $\mathbf{u}$ of our primal-dual formulation. We obtain reliability and efficiency (Theorems 4.3 and 4.5) by the methods developed in [MW01] up to terms which we expect to be of higher order for adapted meshes, and up to a factor p, which is also present in the error indicator for linear systems given in [MW01]. Furthermore, we recapitulate the ideas of the *hp*-adaptive strategies given in [HS05, EM07] using estimation of the decay of the coefficients of the Legendre series of the discrete solution for the decision of whether to h- or to p-refine a certain element.

This combination turns out to be very effective, as can be seen in the numerical experiments given in Chapter 5. We see that higher order h-adaptive methods yield very good convergence rates compared to uniform methods. The uniform p-version, in particular, is not able to deliver an acceptable convergence rate, which is likely due to the fact that the (a priori unknown) points of nondifferentiability of $\boldsymbol{\lambda}$ are not resolved properly by the mesh. With the *hp*-adaptivity using the decay of the Legendre coefficients, in contrast, we can even obtain exponential convergence. This proves that even for relatively complicated nonlinear problems, fully automatic *hp*-adaptivity is very effective, and if the implementational complexity appears to be too high, one should at least try to use h-adaptive methods, by which one can obtain the maximal order of convergence of an h-version even for non-smooth solutions.

Bibliography

[Ada75] Robert A. Adams, *Sobolev spaces*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975, Pure and Applied Mathematics, Vol. 65. MR MR0450957 (56 #9247)

[AG00] Samir Adly and Daniel Goeleven, *A discretization theory for a class of semi-coercive unilateral problems*, Numer. Math. **87** (2000), no. 1, 1–34. MR MR1800152 (2001h:49013)

[AMT99] Mark Ainsworth, William McLean, and Thanh Tran, *The conditioning of boundary element equations on locally refined meshes and preconditioning by diagonal scaling*, SIAM J. Numer. Anal. **36** (1999), no. 6, 1901–1932 (electronic). MR MR1712149 (2001d:65063)

[AMW99] Yves Achdou, Yvon Maday, and Olof B. Widlund, *Iterative substructuring preconditioners for mortar element methods in two dimensions*, SIAM J. Numer. Anal. **36** (1999), no. 2, 551–580 (electronic). MR MR1675257 (99m:65233)

[BC04] S. Brenner and C. Carstensen, *Finite element methods*, Encyclopedia of Computational Mechanics (E. Stein, R. de Borst, and T.J.R. Hughes, eds.), vol. 1, John Wiley & Sons, 2004, pp. 74–118.

[BDM07] Christine Bernardi, Monique Dauge, and Yvon Maday, *Polynomials in the Sobolev World*, Preprints of the Laboratories J.-L. Lions (2007).

[BF91] Franco Brezzi and Michel Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991. MR MR1115205 (92d:65187)

[BL76] Jöran Bergh and Jörgen Löfström, *Interpolation spaces. An introduction*, Springer-Verlag, Berlin, 1976, Grundlehren der Mathematischen Wissenschaften, No. 223. MR MR0482275 (58 #2349)

[BM92] Christine Bernardi and Yvon Maday, *Polynomial interpolation results in Sobolev spaces*, J. Comput. Appl. Math. **43** (1992), no. 1-2, 53–80, Orthogonal polynomials and numerical methods. MR MR1193294 (93k:65010)

[BM97] I. Babuška and J. M. Melenk, *The partition of unity method*, Internat. J. Numer. Methods Engrg. **40** (1997), no. 4, 727–758. MR MR1429534 (97j:73071)

[Bra07] Dietrich Braess, *Finite elements. Theory, fast solvers and applications in elasticity theory. (Finite Elemente. Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie.) 4th revised and extended ed.*, Berlin: Springer. xviii, 357 p. EUR 44.95; SFR 73.50 , 2007 (German).

[Bre83] Haïm Brezis, *Analyse fonctionnelle*, Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree], Masson, Paris, 1983, Théorie et applications. [Theory and applications]. MR MR697382 (85a:46001)

[BS08] Susanne C. Brenner and L. Ridgway Scott, *The mathematical theory of finite element methods*, third ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR MR2373954

[Che06] A. Chernov, *Nonconforming boundary elements and finite elements for interface and contact problems with friction — hp-version for mortar, penalty and Nitsches methods*, Ph.D. thesis, Universität Hannover, 2006.

[Cia78] Philippe G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978, Studies in Mathematics and its Applications, Vol. 4. MR MR0520174 (58 #25001)

[CMS08] A. Chernov, M. Maischak, and E.P. Stephan, *hp-Mortar boundary element method for two-body contact problems with friction*, Math. Meth. Appl. Sci. (2008).

[Dav63] Philip J. Davis, *Interpolation and approximation*, Blaisdell Publishing Co. Ginn and Co. New York-Toronto-London, 1963. MR MR0157156 (28 #393)

[DGS+98] L. Demkowicz, K. Gerdes, C. Schwab, A. Bajer, and T. Walsh, *HP90: A general and flexible Fortran 90 hp-FE code.*, Comput. Vis. Sci. **1** (1998), no. 3, 145–163 (English).

[DL76] G. Duvaut and J.-L. Lions, *Inequalities in mechanics and physics*, Springer-Verlag, Berlin, 1976, Translated from the French by C. W. John, Grundlehren der Mathematischen Wissenschaften, 219. MR MR0521262 (58 #25191)

[DL93] Ronald A. DeVore and George G. Lorentz, *Constructive approximation*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 303, Springer-Verlag, Berlin, 1993. MR MR1261635 (95f:41001)

[Dör07] Philipp Dörsek, *h- and p-version Finite Element Methods for Elastoplasticity*, Master's thesis, Brunel University West London, Sep 2007.

[DS80] Todd Dupont and Ridgway Scott, *Polynomial approximation of functions in Sobolev spaces*, Math. Comp. **34** (1980), no. 150, 441–463. MR MR559195 (81h:65014)

[DS05] Zdeněk Dostál and Joachim Schöberl, *Minimizing quadratic functions subject to bound constraints with the rate of convergence and finite termination*, Comput. Optim. Appl. **30** (2005), no. 1, 23–43. MR MR2122181 (2006e:90082)

[EM07] T. Eibner and J. M. Melenk, *An adaptive strategy for hp-FEM based on testing for analyticity*, Comput. Mech. **39** (2007), no. 5, 575–595. MR MR2288643 (2008d:65132)

[Eva98] Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR MR1625845 (99e:35001)

[GB86] B. Guo and I. Babuška, *The h-p version of the finite element method. I- The basic approximation results*, Comput. Mech. **1** (1986), no. 1, 21–41.

[Glo84] Roland Glowinski, *Numerical methods for nonlinear variational problems*, Springer Series in Computational Physics, Springer-Verlag, New York, 1984. MR MR737005 (86c:65004)

[GLT81] Roland Glowinski, Jacques-Louis Lions, and Raymond Trémolières, *Numerical analysis of variational inequalities*, Studies in Mathematics and its Applications, vol. 8, North-Holland Publishing Co., Amsterdam, 1981, Translated from the French. MR MR635927 (83k:49014)

[GM01] Gabriel N. Gatica and Salim Meddahi, *A fully discrete Galerkin scheme for a two-fold saddle point formulation of an exterior nonlinear problem*, Numer. Funct. Anal. Optim. **22** (2001), no. 7-8, 885–912. MR MR1871866 (2002k:65187)

[Gri85] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985. MR MR775683 (86m:35044)

[GS93] J. Gwinner and E. P. Stephan, *A boundary element procedure for contact problems in plane linear elastostatics*, RAIRO Modél. Math. Anal. Numér. **27** (1993), no. 4, 457–480. MR MR1230830 (94h:73080)

[Han05] Weimin Han, *A posteriori error analysis via duality theory*, Advances in Mechanics and Mathematics, vol. 8, Springer-Verlag, New York, 2005, With applications in modeling and numerical approximations. MR MR2101057 (2005k:65003)

[Has81] Jaroslav Haslinger, *Mixed formulation of elliptic variational inequalities and its approximation*, Apl. Mat. **26** (1981), no. 6, 462–475, With a loose Russian summary. MR MR634283 (83g:49009)

[HHNL88] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek, *Solution of variational inequalities in mechanics*, Applied Mathematical Sciences, vol. 66, Springer-Verlag, New York, 1988, Translated from the Slovak by J. Jarník. MR MR952855 (89i:73003)

[HS05] Paul Houston and Endre Süli, *A note on the design of hp-adaptive finite element methods for elliptic partial differential equations*, Comput. Methods Appl. Mech. Engrg. **194** (2005), no. 2-5, 229–243. MR MR2105162 (2005h:65211)

[KK98] George Karypis and Vipin Kumar, *A fast and high quality multilevel scheme for partitioning irregular graphs*, SIAM J. Sci. Comput. **20** (1998), no. 1, 359–392 (electronic). MR MR1639073 (99f:68158)

[KO88] N. Kikuchi and J. T. Oden, *Contact problems in elasticity: a study of variational inequalities and finite element methods*, SIAM Studies in Applied Mathematics, vol. 8, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1988. MR MR961258 (89j:73097)

[Kor97] Ralf Kornhuber, *Adaptive monotone multigrid methods for nonlinear variational problems*, Advances in Numerical Mathematics, B. G. Teubner, Stuttgart, 1997. MR MR1469497 (98e:65054)

[Kre04] A. Krebs, *On Solving Nonlinear Variational Inequalities by p-Version Finite Elements*, Ph.D. thesis, Fachbereich Mathematik der Universität Hannover, 2004.

[KS80] David Kinderlehrer and Guido Stampacchia, *An introduction to variational inequalities and their applications*, Pure and Applied Mathematics, vol. 88, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. MR MR567696 (81g:49013)

[KS07] Andreas Krebs and Ernst P. Stephan, *A p-version finite element method for nonlinear elliptic variational inequalities in 2D*, Numer. Math. **105** (2007), no. 3, 457–480. MR MR2266833 (2007k:65089)

[LL70] L. D. Landau and E. M. Lifshitz, *Theory of elasticity*, second ed., Course of Theoretical Physics, Vol. 7. Translated by J. B. Sykes and W. H. Reid, Pergamon Press, London, 1970.

[LM72] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. I*, Springer-Verlag, New York, 1972, Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. MR MR0350177 (50 #2670)

[Mai01a] M. Maischak, *FEM/BEM methods for Signorini-type problems — error analysis, adaptivity, preconditioners*, Habilitation thesis, Institut für Angewandte Mathematik, Universität Hannover, 2001.

[Mai01b] Matthias Maischak, *Manual of the software package maiprogs*, Tech. Report 48, Institut für Angewandte Mathematik, Universität Hannover, 2001.

[MB96] J. M. Melenk and I. Babuška, *The partition of unity finite element method: basic theory and applications*, Comput. Methods Appl. Mech. Engrg. **139** (1996), no. 1-4, 289–314. MR MR1426012 (97k:65258)

[Mel02] Jens M. Melenk, *hp-finite element methods for singular perturbations*, Lecture Notes in Mathematics, vol. 1796, Springer-Verlag, Berlin, 2002. MR MR1939620 (2003i:65108)

[Mel05] J. M. Melenk, *hp-interpolation of nonsmooth functions and an application to hp-a posteriori error estimation*, SIAM J. Numer. Anal. **43** (2005), no. 1, 127–155 (electronic). MR MR2177138 (2006g:65178)

[MS05] Matthias Maischak and Ernst P. Stephan, *Adaptive hp-versions of BEM for Signorini problems*, Appl. Numer. Math. **54** (2005), no. 3-4, 425–449. MR MR2149362 (2006a:65174)

[MS07] ———, *Adaptive hp-versions of boundary element methods for elastic contact problems*, Comput. Mech. **39** (2007), no. 5, 597–607. MR MR2288644 (2008b:74082)

[MW01] J. M. Melenk and B. I. Wohlmuth, *On residual-based a posteriori error estimation in hp-FEM*, Adv. Comput. Math. **15** (2001), no. 1-4, 311–331 (2002), A posteriori error estimation and adaptive computational methods. MR MR1887738 (2002m:65126)

[QSS07] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri, *Numerical mathematics*, second ed., Texts in Applied Mathematics, vol. 37, Springer-Verlag, Berlin, 2007. MR MR2265914 (2007f:65001)

[RR04] Michael Renardy and Robert C. Rogers, *An introduction to partial differential equations*, second ed., Texts in Applied Mathematics, vol. 13, Springer-Verlag, New York, 2004. MR MR2028503 (2004j:35001)

[Rud87] Walter Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987. MR MR924157 (88k:00002)

[SB93] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, second ed., Texts in Applied Mathematics, vol. 12, Springer-Verlag, New York, 1993, Translated from the German by R. Bartels, W. Gautschi and C. Witzgall. MR MR1295246 (95i:65006)

[Sch98] Ch. Schwab, *p- and hp-finite element methods*, Numerical Mathematics and Scientific Computation, The Clarendon Press Oxford University Press, New York, 1998, Theory and applications in solid and fluid mechanics. MR MR1695813 (2000d:65003)

[SG02] Olaf Schenk and Klaus Gärtner, *Solving unsymmetric sparse systems of linear equations with PARDISO*, Computational science—ICCS 2002, Part II (Amsterdam), Lecture Notes in Comput. Sci., vol. 2330, Springer, Berlin, 2002, pp. 355–363. MR MR2045251

[SG06] ______, *On fast factorization pivoting methods for sparse symmetric indefinite systems*, Electron. Trans. Numer. Anal. **23** (2006), 158–179 (electronic). MR MR2268558 (2008c:65079)

[Sin06] Astrid Sinwel, *Numerical Simulation of Contact Problems with Coulomb Friction*, Master's thesis, Johannes Kepler Universität, Jan 2006.

[SS00] Padmanabhan Seshaiyer and Manil Suri, *Uniform hp convergence results for the mortar finite element method*, Math. Comp. **69** (2000), no. 230, 521–546. MR MR1642762 (2000i:65192)

[Tri95] Hans Triebel, *Interpolation theory, function spaces, differential operators*, second ed., Johann Ambrosius Barth, Heidelberg, 1995. MR MR1328645 (96f:46001)

[Yos80] Kôsaku Yosida, *Functional analysis*, sixth ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 123, Springer-Verlag, Berlin, 1980. MR MR617913 (82i:46002)

[Zei85] Eberhard Zeidler, *Nonlinear functional analysis and its applications. III*, Springer-Verlag, New York, 1985, Variational methods and optimization, Translated from the German by Leo F. Boron. MR MR768749 (90b:49005)

Index