The Mathematical Analysis of a Micro Scale Model for Lithium-Ion Batteries

Zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften

von der Fakultät für Mathematik des Karlsruher Instituts für Technologie (KIT)

genehmigte Dissertation

von

Markus Maier

Referent: Prof. Dr. Willy Dörfler
Korreferent: Prof. Dr. Wolfgang Reichel
Externer Korreferent: Prof. Dr. Robert Denk

Tag der mündlichen Prüfung: 6. Juli 2016

Acknowledgments

This dissertation emerged during my work at the Institute for Applied and Numerical Mathematics of the Karlsruhe Institute of Technology (KIT). I would like to take this opportunity to thank a number of people who supported me throughout my time as a doctoral student.

First and foremost, I want to thank my advisor Prof. Dr. Willy Dörfler for giving me the opportunity to work on this interesting topic. In countless inspiring mathematical discussions he eagerly shared his deep mathematical knowledge and thereby guided me through my doctoral studies. Furthermore, I thank Prof. Dr. Wolfgang Reichel for being my co-examiner and giving me valuable feedback to preliminary versions of my thesis. Moreover, I would like to give thanks to Prof. Dr. Robert Denk who served as an external referee. I am also very grateful for his invitation to the University of Konstanz where I could present my results. I enjoyed a lot the insightful discussions about mathematical topics related to my work with him and his group. In addition, Prof. Dr. Stefan Volkwein deserves my thanks for kindly inviting me to the University of Konstanz in order to present some of my results. It was a pleasure and very helpful to discuss the numerical aspects of my topic with him and his doctoral students.

I thank Fabian Castelli for the implementation of the numerical methods and for the generation of the numerical results. For the very detailed proof-reading and the explanations on the electro-chemical working principles of lithium-ion batteries I want to express my gratitude to Dr. Marieke Poß. Furthermore, I want to thank Michael Kespe for displaying to me the key aspects of battery modelling from an engineer's point of view and for the good cooperation in general. Moreover, I want to give thanks to my colleagues Dr. Branimir Anić, Carlos Galeano, Dr. Stefan Findeisen, Pascal Kraft and Lars Machinek. Our discussions about all kinds of topics during lunch and coffee breaks were incredibly entertaining and made the every day life at the university a really pleasant experience.

On a more personal note, I thank my family for their ongoing support throughout my studies. Especially I want to mention my grand father Kurt who fascinated me for science by working with me on countless little projects in his workshop. Also I want to give thanks to Eva's lovely family who included me in the most gracious fashion. Finally, I want to express my deepest gratitude to my fiancee Eva for being such a beautiful partner. She supported me through dark times, always believed in my abilities and never got tired of listening to the boring problems I adressed her to. Apart from that I thank her for sacrifycing so much of her precious time for thoroughly proof-reading my thesis.

Contents

1 Introduction

In this thesis we consider a system of partial differential equations arising in the modeling of the electrochemical processes in Li-ion batteries. The model has originally been presented in the article [60] and has been refined in [57, 58].

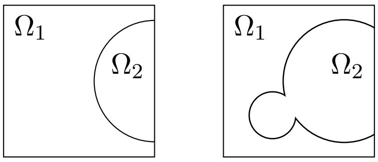

Figure 1.1: Model geometries representing a single particle electrode.

In this model the battery is geometrically represented by the two given disjoint Lipschitz domains Ω_1 and Ω_2 in $\mathbb{R}^d$ for some $d \geq 2$. The domain Ω_1 is the region of the electrolyte and Ω_2 is the region of the solid particles in the lithium metal oxide electrode.

The unknown quantities are the *lithium concentrations* c_i and the so-called *electrochemical potentials* u_i which are both real-valued functions defined on the time-space cylinder $(0,T) \times \Omega_i$ for $i = 1, 2$. The concentration additionally obeys the pointwise bounds $0 < c_1$ and $0 < c_2 < 1$.

For $i = 1, 2$, the following partial differential equations are imposed:

$$\partial_t c_i - \Delta c_i = 0 \qquad \text{in } (0,T) \times \Omega_i, \tag{1.1}$$

$$-\nabla \cdot (\kappa(c_i) \nabla u_i) = 0 \qquad \text{in } (0,T) \times \Omega_i. \tag{1.2}$$

Here, $\kappa(c_i) > 0$ is the electric conductivity. The function κ is a given locally Lipschitz continuous function on $(0, \infty)$.

The interface $I := \partial\Omega_1 \cap \partial\Omega_2$ is supposed to have positive $(d-1)$-dimensional Hausdorff-measure. We denote by ν the unit normal on I pointing from Ω_1 to Ω_2. The following interface conditions are imposed:

$$\partial_\nu c_1 = \beta\, i_{12}, \qquad \partial_\nu c_2 = i_{12} \qquad \text{on } (0,T) \times I, \tag{1.3}$$

$$\kappa(c_1)\partial_\nu u_1 = i_{12}, \qquad \kappa(c_2)\partial_\nu u_2 = i_{12} \qquad \text{on } (0,T) \times I. \tag{1.4}$$

Here, $\beta \in (0,1)$ is a constant and $i_{12} = i_{12}(c_1, c_2, u_2 - u_1)$ is a real-valued $\mathcal{C}^1$-function defined on $(0,\infty) \times (0,1) \times \mathbb{R}$ which is monotone with respect to the third variable. A typical example from the applications is the so-called *Butler–Vollmer nonlinearity*, given by

$$i_{12}(c_1, c_2, z) = \sqrt{c_1}\sqrt{c_2}\sqrt{1 - c_2}\,\sinh(z - t_+ \ln(c_1)) \tag{1.5}$$

where $t_+ = 1 - \beta \in (0, 1)$ is the so-called *transference number*.

Finally, mixed boundary conditions are imposed on $(0, T) \times \partial\overline{\Omega}$ together with an inital condition for the concentration.

For a more detailed description of the model see Chapter 3.

Mathematical Challenges

The model is challenging from a mathematical point of view for several reasons:

First of all, this is a *nonlinear system* of four partial differential equations of two *different types*: The equations (1.1) for the lithium transport are *parabolic* with respect to the concentration whereas the equations for the charge transport (1.2) are *elliptic* with respect to the potential.

Furthermore, the nonlinear functions coupling the equations display a certain singular behavior: All nonlinearities can exhibit singularities with respect to the lithium concentration at 0 and 1. In addition, the interface nonlinearity i_{12} does not satisfy any polynomial growth conditions with respect to the third variable $z = u_2 - u_1$. In fact the Butler–Vollmer nonlinearity (1.5) grows exponentially at $\pm\infty$. Note that a subcritical polynomial growth condition is a frequent standard assumption in the theory of quasilinear elliptic equations, compare for example [42, §8.5.2 and §9.1] and [44, §8.5].

Another difficulty is that the equations are coupled via a set of nonlinear Neumann interface conditions on the lower dimensional submanifold I instead of, say, the right hand side of the differential equations (1.1) and (1.2). This feature relates the model to transmission problems [86, 31, 40, 54].

Finally, the boundary of each subdomain is only assumed to be Lipschitz continuous. For this reason, special attention has to be paid when applying regularity results from the theory of elliptic and parabolic equations, since many of these results require $\mathcal{C}^1$-boundaries of the underlying domain, see for example [42, §6.3.2], [44, Chapter 8] for elliptic and [55, Page 9], [63, §IV.7] for parabolic problems. In particular, the standard literature lacks regularity results for parabolic equations with inhomogeneous Neumann boundary conditions on Lipschitz domains, which are necessary to perform the fixed point argument for the fully coupled system.

Note that the geometries in the applications are in fact not much more regular than Lipschitz continuous. Even in the most basic example where the lithium electrode Ω_2 consists of one particle only, the subdomain Ω_1 corresponding to the electrolyte can have inward corners of arbitrary small angle, see Fig. 1.1.

Both quasilinear elliptic and parabolic equations have been well-understood for a long time now, see for example the textbooks [44, 45] for the elliptic and [55, 63, 36] for the parabolic case. In contrast, the theory of elliptic-parabolic *systems* is far less developed. A central article that provides structural conditions which are sufficient for the well-posedness of a wide class of elliptic-parabolic systems is [87]. However, a central assumption in this article is a subcritical polynomial growth condition which is not satisfied in our case. Moreover it does not include the case of a nonlinear interface condition.

In a similar system which models the electrochemical effects in Li-ion batteries on the macro scale, the Butler–Volmer nonlinearity (1.5) enters the equation as a homogenized source term. For this system, local in time existence of weak and strong solutions has been proved in [88] and [74], respectively. These articles were in fact a guideline for us in establishing the main result of this thesis.

Note that the model considered in *the current* thesis has been investigated numerically both in the mathematical and in the engineering community. Discrete approximation schemes based on the finite elements and the finite volumes method and the respective numerical results were presented in [83, 70, 91, 62, 53] for example, however, discrete stability and convergence proofs are missing.

Due to the above mentioned reasons, well-posedness of the model does not follow *directly* from the standard theories of elliptic and parabolic equations. Furthermore, to the author's best knowledge, no explicit results have been established in the literature, either.

Main Results

Let us summarize the main results of this thesis and briefly sketch the techniques which were used to obtain them:

We will show *local in time existence* of *weak solutions* to the *fully coupled problem*.

In the proof, we exploit the elliptic-parabolic structure of the model: For a fixed concentration $c = (c_1, c_2)$ we consider the equation for the charge-conservation as a problem for the potential $u = (u_1, u_2)$, the so-called *elliptic subproblem*. Note that this subproblem is strongly nonlinear. However, we can exploit the monotonicity of i_{12} with respect to the third variable and apply the theory of monotone operators to an approximate problem. The convergence of the approximate solutions is then derived by showing uniform pointwise bounds with the Stampacchia truncation method.

Then we plug in the elliptic equation (1.2) into the parabolic one (1.1) and perform a Schauder fixed point argument for the resulting equation for the concentration c. A central tool used in this argument is a recently developed maximal parabolic regularity result for the negative Laplacian with inhomogeneous Neumann boundary conditions on non-smooth domains. Note that the Schauder fixed point theorem does not provide uniqueness. However, for $d \leq 3$ we can refine the estimates from the existence proof and use Sobolev embeddings to derive uniqueness of our weak solutions.

As the boundary values used in this thesis model the discharge of the battery at a prescribed constant rate, it is evident that the system might not admit a solution which exists globally in time. This can also be verified by considering one-dimensional geometries where the exact solution is explicitely known, see Section 7.3.3.

As a result, global existence of solutions will not be discussed in this thesis. However, when different boundary values are chosen, the question of global existence is more interesting and might in fact be answered positively. The example for such boundary conditions that comes to our mind is when they model the discharge of the battery via a given electrical resistance.

Auxillary results

Motivated by the elliptic subproblem, we study the finite element discretizations for a class of strongly nonlinear, uniformly monotone elliptic problems.

The reason to study these is that in the theory of finite elements for (quasilinear) elliptic equations, a standard assumption is a subcritical polynomial growth condition which is not satisfied for the elliptic subproblem, see [21, Chapter 5] and [14, §8.7].

For the standard Galerkin approximations we prove optimal convergence. However, the Stampacchia truncation cannot be executed and we can only recover uniform pointwise bounds from the error estimate by invoking inverse inequalities under additional technical assumptions.

To overcome this shortcoming, we introduce a modified version of the Galerkin discretization with linear finite elements which is basically obtained by applying the trapezoidal rule to the nonlinear interface term. For this modified discretization we are still able to prove optimal linear convergence. Additionally, we can use the maximum principle for the discrete Laplacian on non-negative meshes and are able to derive a discrete comparison principle and, more importantly, a uniform pointwise bound for the discrete solutions.

Finally we discuss the concrete numerical solution of the equations considered. For the elliptic subproblem, numerical results are presented which substantiate the theoretical convergence estimates that we have established. For the full system we show up several possible solution strategies which are used for an efficient numerical treatment of the fully coupled system. The numerical results obtained with these strategies are shown in order to illustrate the simulated transport processes in Li-ion batteries. All the numerical results were kindly provided by our master student Fabian Castelli who has worked under my supervision [17].

Outline

In Chapter 2 we introduce some basic notation used throughout the thesis and recall some standard results from calculus which will from then on be used without explicit citation. In Chapter 3 we describe very briefly the modelling of Li-ion batteries and derive the equations under investigation. In Chapter 4 a class of strongly nonlinear elliptic problems is investigated. Weak well-posedness is established and the mapping properties of the solution operator are considered. The problems which are inspected in this chapter are motivated by the elliptic subproblem for the fully coupled system. Then in Chapter 5 the main results of the thesis are proved: Local in time existence and uniqueness of weak solutions to the fully coupled system. The arguments in this chapter rely heavily on the properties of the elliptic subproblem derived in Chapter 4. Chapter 6 is then devoted to discretizations of the elliptic problems considered in Chapter 4. Optimal convergence is shown for both the standard Galerkin approximations and the modified discretization. Additionally, for the modified system, a comparison principle and a uniform pointwise bound is derived. Finally in Chapter 7 we present numerical results both for the elliptic subproblem and the fully coupled system.

2 Prelimaries

In this chapter we will explain the notation which will be adopted throughout the thesis and recall some frequently used concepts and results from calculus.

2.1 Notation

The set of *positive integers* is denoted by $\mathbb{N}$. For the *real numbers* we write $\mathbb{R}$. The *positive part* of $z \in \mathbb{R}$ is denoted by $z_+ := \max\{z, 0\}$. We use the symbol $|\cdot|$ for both *the absolute value* in $\mathbb{R}$ and the 2-norm in $\mathbb{R}^n$. For general $p \geq 1$ the *p-norm in* $\mathbb{R}^n$ is denoted by $|\cdot|_p$. The components of a vector $v \in \mathbb{R}^d$ will mostly be denoted by v_i for $i = 1, \dots, n$, that is, $v = (v_i)_i$. For the dot-product of two vectors $v, w \in \mathbb{R}^n$ we will write $v \cdot w := \sum_i v_i w_i$.

Suppose u is a real-valued function defined on some subset of $\mathbb{R}^n$. If the variable for u is $z = (z_i)_i$, the (possibly distributional) *partial derivatives* of u are denoted by $\partial_{z_i} u$ for $i = 1, \dots, n$.

Throughout this work, the *spatial dimension* will be $d \in \mathbb{N}$ with $d \geq 2$. The spatial variable will be $x = (x_i)_i$. Therefore, the *spatial derivates* of a real-valued function u defined on a subset of $\mathbb{R}^d$ are denoted by $\partial_{x_i} u$ for $i = 1, \dots, d$. The *gradient* of u is $\nabla u = (\partial_{x_i} u)_i$. Suppose $v = (v_i)_i$ is defined on a subset of $\mathbb{R}^d$ and has the d real-valued components v_i for $i = 1, \dots, d$. Then the *divergence* of v is $\nabla \cdot v := \sum_i \partial_{x_i} v_i$. Finally, the *laplacian* of u is $\Delta u := \nabla \cdot (\nabla u)$.

Magnitudes which depend on time and space are represented by functions u defined on subsets of $\mathbb{R} \times \mathbb{R}^d$. The variable for such functions is the couple (t, x) of the *temporal variable* t and the spatial variable x. The derivative of u with respect to time is thus denoted by $\partial_t u$ and its spatial derivatives are again $\partial_{x_i} u$ for $i = 1, \dots, d$. The *spatial* gradient, divergence and laplacian of such functions are denoted by ∇, $\nabla\cdot$ and Δ and they are defined analogously as above. For example the *spatial* gradient is $\nabla u := (\partial_{x_i} u)_i$.

The d-dimensional Lebesgue measure is denoted by μ. The Hausdorff-measure is scaled in a way such that it coincides with the respective Lebesgue surface measures on Lipschitz submanifolds of $\mathbb{R}^d$. For the $(d-1)$-dimensional Hausdorff-measure we will use the symbol σ. See for example [43, Chapter 2] for an introduction into measure theory.

As already indicated, throughout the thesis we will be given two fixed disjoint domains Ω_1 and Ω_2 in $\mathbb{R}^d$. We will define the open but *non-connected* set $\Omega := \Omega_1 \cup \Omega_2$ and identify functions u defined on Ω with couples (u_1, u_2) of functions u_i defined on Ω_i for $i = 1, 2$, by setting $u_i := u|_{\Omega_i}$, or conversely, $u := u_1 \chi_{\Omega_1} + u_2 \chi_{\Omega_2}$. Here, χ_D denotes the characteristic function of a set D, that is, $\chi_D(x) = 1$ if $x \in D$ and $\chi_D(x) = 0$ otherwise. Additionally, the jump of u across $I := \partial\Omega_1 \cap \Omega_2$ will be denoted by $[u] := u_2|_I - u_1|_I$ where $u_i|_I$ is the trace of u_i on I for $i = 1, 2$, see Section 2.1.1.

In general, positive constants will be denoted by C, C_1, C_2, ... Assume some objects $X_1, X_2, \dots, X_m$ and $Y_1, Y_2, \dots, Y_n$ are given. To express that the constant C *only depends* on X_i but *not on* Y_i, we will write $C = C(X_1, X_2, \dots, X_m)$. Furthermore, we will use the symbols $\lesssim_{X_1,\dots,X_m}$ and $\lesssim_{\neg Y_1,\dots,\neg Y_n}$ for the relation on $(0, \infty)$ defined by $a \lesssim_{X_1,\dots,X_n} b$ if there is a positive constant $C = C(X_1, \dots, X_m)$ which only depends on X_i but neither on Y_i nor on a or b such that $a \le Cb$ holds. Note that we will *omit the dependence* on the *fixed data* like the geometry, the coefficient κ and the nonlinearity i_{12} from (1.5).

2.1.1 Spaces

Suppose X is a real Banach space. The topological dual space of X is denoted by X' and equipped with the operator norm. For the evaluation of a bounded linear functional we use angled brackets $\langle\cdot,\cdot\rangle$, that is, $\langle x', x\rangle := x'(x)$ for $x' \in X'$ and $x \in X$. The arrow $\rightharpoonup$ will be used to express weak convergence. Assume, Y is another Banach space. We say that X is continuously embedded into Y if there is an *injective* bounded linear operator from X to Y. In this case we write $X \hookrightarrow Y$. Such an embedding is called *compact* if it is compact in the sense of linear operators. See for example [44, Chapter 5] for a general introduction to functional analysis.

Polynomials

For $p \in \mathbb{N}$ the set of polynomials in d variables of total degree less or equal than p is denoted by $\mathbb{P}_p := \mathbb{P}_p(\mathbb{R}^d)$. If D is a subset of $\mathbb{R}^d$ we write $\mathbb{P}_p(D)$ for the set of restrictions of such polynomials to D.

The set of tensor product polynomials in d variables of degree less or equal than p is $\mathbb{Q}_p := \mathbb{Q}_p(\mathbb{R}^d) := (\mathbb{P}_p(\mathbb{R}))^{\otimes d}$. Again, we use the symbol $\mathbb{Q}_p(D)$ for the set of restrictions of such polynomials to D.

Hölder Spaces

Let D be a bounded open set in $\mathbb{R}^d$. By $\mathcal{C}^0(\overline{D})$ we denote the space of real-valued continuous functions on $\overline{D}$, equipped with the maximum norm $\|\cdot\|_{0,\infty;D}$. For $\alpha \in (0,1)$ we introduce the symbol $\mathcal{C}^\alpha(\overline{D})$ for the space of α-Hölder continuous functions on $\overline{D}$ with corresponding Hölder norm $[\![\cdot]\!]_{\alpha;D}$ [42, §5.1].

Lebesgue Spaces

Let (D, Σ, μ) be a measure space and $p \in [1, \infty]$. Then $L^p(\mu)$ denotes the Lebesgue space of real-valued Σ-measurable functions on D which are p-integrable with respect to μ. For the corresponding p-norm we use the symbol $\|\cdot\|_{0,p;\mu}$.

If D is a Lebesgue measurable subset of $\mathbb{R}^d$ with Hausdorff-dimension $n \in \{0, \dots, d\}$, we choose $\hat{\mu}$ as the n-dimensional Hausdorff measure on D and define $L^p(D) := L^p(\hat{\mu})$ and $\|\cdot\|_{0,p;D} = \|\cdot\|_{0,p;\hat{\mu}}$.

Note that the elements in $L^p(D)$ are not actual functions defined on D but rather equivalence classes of such functions with respect to the relation $v \sim w$ iff $v = w$ $\hat{\mu}$-almost everywhere on D. In this thesis we will switch between the notion of functions and equivalence classes whenever it is convenient.

The construction of these spaces can be found in basically every book on measure theory, see for example [43, Chapter 1].

Sobolev Spaces

Let D be an open subset of $\mathbb{R}^d$. For $s \in [0, \infty)$ and $p \in [1, \infty]$ we denote by $W^{s,p}(D)$ the (possibly fractional) real-valued Sobolev space on D and by $\|\cdot\|_{s,p;D}$ and $|\cdot|_{s,p;D}$ the corresponding Sobolev norm and semi-norm, respectively. More generally, suppose D is an open subset of a submanifold of $\mathbb{R}^d$ with k-times Lipschitz continuously differentiable maps. Then for $s \in [k, k+1)$ and $p \geq 1$ one can define the real-valued Sobolev spaces $W^{s,p}(D)$ of order s and Sobolev exponent p. The corresponding norms and semi-norms are again denoted by $\|\cdot\|_{s,p;D}$ and $|\cdot|_{s,p;D}$, respectively. For $s = 0$ this coincides with the Lebesgue space, that is, $W^{0,p}(D) = L^p(D)$. Other important special cases are the Hilbert spaces $H^s(D) := W^{2,s}(D)$.

Let D be a bounded Lipschitz domain in $\mathbb{R}^d$. Then for $s > 0$ and $p \geq 1$ with either $s - 1/p > 0$ or $s = p = 1$ there is a bounded trace operator $\gamma : W^{s,p}(D) \to L^1(\partial\Omega)$ such that $\gamma(v) = v|_{\partial D}$ for all $v \in W^{s,p}(D) \cap \mathcal{C}^0(\overline{D})$. Then we will write $v|_{\partial D} := \gamma(v)$ for $v \in W^{s,p}(D)$ even if v is not continuous. Note that by this definition, the symbol $\cdot|_{\partial D}$ is overloaded. In fact, for functions in $W^{s,p}(D)$ which are not continuous, $v|_{\partial D}$ in general deviates from the restriction in the sense of mappings.

Suppose S is a measurable subset of $\partial\Omega$. Having introduced the trace operator, we can define $W^{s,p}_S(D)$ as the subspace of all functions $v \in W^{s,p}_S(D)$ which vanish on S in the sense of traces, that is, $v|_S := \gamma(v)|_S = 0$. The functions which vanish on the complete boundary are denoted by $W^{s,p}_0(D) := W^{s,p}_{\partial D}(D)$. For $p = 2$ the spaces $H_S(D)$ and $H_0(D)$ are defined analogously.

The conjugated Sobolev exponent $p' \in [1, \infty]$ is defined by the equation $1/p + 1/p' = 1$. For $s > 0$ we denote by $W^{-s,p}(D) := (W^{s,p'}(D))'$ the negative Sobolev space with corresponding norm $\|\cdot\|_{-s,p;D}$. Additionally, we define $H^{-1}(D) := (H^1(D))'$ and $H^{-1}_S(D) := (H^1_S(D))'$. For sufficiently large $q \in [1, \infty]$ a function $v \in L^q(D)$ will be considered as an element of $W^{-s,p}(D)$ by defining $\langle v, w \rangle := \int_D vw \,\mathrm{d}x$ for all $v \in W^{s,p'}(D)$. The finiteness of the integral is justified by Sobolev embedding and Hölder's inequality, see Section 2.2.

Good references for Sobolev spaces on open subsets of $\mathbb{R}^d$ are [1] and [61], especially [61, Chapter 15] for traces. A nice introduction to Sobolev spaces on manifolds can be found in [45, §1].

Vector-Valued Spaces

Let X be a Banach space and $T > 0$. Similar to the scalar-valued case one can define vector-valued versions of the previously defined spaces. We use the symbols $\mathcal{C}^\alpha([0, T]; X)$,

$L^p((0,T);X)$ and $W^{1,p}((0,T);X)$ for the vector-valued versions of the respective spaces. The corresponding norms are denotes by $\|\cdot\|_{\mathcal{C}^\alpha([0,T];X)}$, $\|\cdot\|_{L^p((0,T);X)}$ and $\|\cdot\|_{W^{1,p}((0,T);X)}$, respectively. The weak derivative of a function $v \in W^{1,p}((0,T);X)$ is an element of $L^p((0,T);X)$ and it will be denoted by v'.

See [42, §5.9] for the precise definitions of the vector-valued Lebesgue and Sobolev-spaces and for example [4, II.1.1] for the respective Hölder-spaces.

Broken Spaces

As indicated in the introduction, throughout the thesis we will be given two disjoint Lipschitz domains $\Omega_1, \Omega_2 \subset \mathbb{R}^d$. We will define $\Omega := \Omega_1 \cup \Omega_2$ and $I := \partial\Omega_1 \cap \Omega_2$ and assume that I has positive Hausdorff-measure. As a consequence, the boundary of Ω is no longer a Lipschitz submanifold of $\mathbb{R}^d$. Therefore, the function spaces on Ω require some explanation, which will be given now:

We define $W^{s,p} := W^{s,p}(\Omega)$ for $s \in \mathbb{R}$ and $p \geq 1$ and, additionally, we set $H^s := W^{s,2}$. Since $\Omega = \Omega_1 \cup \Omega_2$ is not a Lipschitz domain, for *all* $s \geq 0$ functions from $W^{s,p}$ generally admit jumps across I. In fact, $W^{s,p}$ is isomorphic to $W^{s,p}(\Omega_1) \oplus W^{s,p}(\Omega_2)$. The isomorphism is given by the identification described in Section 2.1, that is $v := v_1\chi_{\Omega_1} + v_2\chi_{\Omega_2}$ and, vice versa, $v_i := v|_{\Omega_i}$ for $i = 1, 2$.

For $s > 0$ and $p \geq 1$ with either $s - 1/p > 0$ or $s = p = 1$, the trace of a function $v \in W^{s,p}$ on the *outer boundary* $\partial\overline{\Omega}$ is then defined as $v|_{\partial\overline{\Omega}} = v_i|_{\partial\Omega_i}$ on $\partial\Omega_i$ for $i = 1, 2$. With this convention at hand, we can define $W^{s,p}_S := W^{s,p}_S(\Omega)$ if S is a measurable subset of $\partial\overline{\Omega}$. Note that $W^{s,p}_S$ is isomorphic to $W^{s,p}_{S_1}(\Omega_1) \oplus W^{s,p}_{S_2}(\Omega_2)$, where $S_i := S \cap \partial\Omega_i$ for $i = 1, 2$. Finally, we set $W^{-s,p}_S := (W^{s,p'}_S)'$. The Hilbert spaces H^s_S and H^{-s}_S are defined analogously.

The broken Hölder spaces are defined as $\mathcal{C}^\delta_{\mathrm{b}} := \mathcal{C}^\delta_{\mathrm{b}}(\overline{\Omega}) := \mathcal{C}^\delta(\overline{\Omega}_1) \oplus \mathcal{C}^\delta(\overline{\Omega}_2)$ for $\delta \in [0,1)$. As norms on these space we choose $\|v\|_{\mathcal{C}^\delta_{\mathrm{b}}} := \max_{i=1,2} [\![v_i]\!]_{\delta;\Omega_i}$. Note that in general, functions from $\mathcal{C}^\delta_{\mathrm{b}}(\overline{\Omega})$ are discontinuous across I: $\mathcal{C}^\delta_{\mathrm{b}}(\overline{\Omega}) \supsetneq \mathcal{C}^\delta(\overline{\Omega})$.

2.2 Basic Results

We will now recall some basic results from calculus.

Let $\mathcal{H}$ be an inner product space with inner product $(\cdot,\cdot)$ and induced norm $\|\cdot\|$. Then the *Cauchy–Schwarz inequality* holds. It reads $|(v,w)| \leq \|v\|\|w\|$ for all $v, w \in \mathcal{H}$. Of particular interest for us is the case $\mathcal{H} = L^2(D)$ or $\mathcal{H} = H^1(D)$ with respective inner products defined by $(u,v)_{L^2} = \int_D uv \,\mathrm{d}x$ and $(u,v)_{H^1} = \int_D uv + \nabla u \cdot \nabla v \,\mathrm{d}x$.

For $p > 1$ and $a, b \geq 0$ *Young's inequality* is satisfied: $ab \leq a^p/p + b^{p'}/p'$.

The *Hölder inequality* is a generalization of Cauchy–Schwarz for $p \neq 2$: Let $\hat{\mu}$ be any measure and $p \in [1,\infty]$. Then for all $v \in L^p(\hat{\mu})$ and $w \in L^{p'}(\hat{\mu})$ the pointwise product vw is in $L^1(\hat{\mu})$ and it holds $\|vw\|_{0,1;\hat{\mu}} \leq \|v\|_{0,p;\hat{\mu}}\|w\|_{0,p';\hat{\mu}}$.

Let D be an open Lipschitz domain in $\mathbb{R}^d$ and $S \subset \partial D$ a measurable subset with $\sigma(S) > 0$. Then the *Poincaré inequality* holds. It reads $\|v\|_{1,2;D} \leq C_P\|\nabla v\|_{0,2;D}$ for all $v \in H^1_S(D)$ with a so-called Poincaré constant C_P which does not depend on v.

A good reference for these inequalities is the appendix of [42]. This particular variant of the Poincaré-inequality follows from the abstract result [2, Bemerkung 6.15].

Hölder Embeddings

Let D be a bounded Lipschitz domain in $\mathbb{R}^d$ and $0 \le \alpha < \beta < 1$. Then it holds $\mathcal{C}^\beta(\overline{D}) \subset \mathcal{C}^\alpha(\overline{D})$ and the embedding is compact. More precisely, it holds $[\![v]\!]_{\alpha,D} \le C[\![v]\!]_{\beta,D}$ for all $v \in \mathcal{C}^\beta(\overline{D})$ with the constant $C = \max\{1, \mathrm{diam}(D)^{\beta-\alpha}\}$. Here, $\mathrm{diam}(D)$ denotes the diameter of D. For further details see [2, Ch. 8.6]

In the vector valued case, that is, when X is a Banach space, it still holds $\mathcal{C}^\beta([0,T];X) \subset \mathcal{C}^\alpha([0,T];X)$, but in general the embedding is only *bounded*. However, suppose Y is another Banach space such that X is *compactly embedded* into Y. Then the induced embedding $\mathcal{C}^\beta([0,T];X) \to \mathcal{C}^\alpha([0,T];Y)$ is *compact*. This is a direct consequence of the *Arzéla-Ascoli theorem* for Banach spaces, see for example [3, Lemma 7.2].

Sobolev Embeddings

Let D be a bounded Lipschitz domain in $\mathbb{R}^d$ and $p \in (1,\infty)$. The Sobolev embeddings can be summarized in the following way, see [45, Chapter 1] and [61, Chapter 11].

For $s > t \ge 0$ the embedding $W^{s,p}(D) \hookrightarrow W^{t,p}(D)$ is compact.

For $s \ge t \ge 0$ and $1 \le q < \infty$ satisfying $s - d/p = r - d/q$ it holds $W^{s,p}(D) \hookrightarrow W^{r,q}(D)$.

For $0 < s - d/p < 1$ and $\alpha := s - d/p$ the embedding $W^{s,p}(D) \hookrightarrow \mathcal{C}^\alpha(\overline{D})$ holds.

In the limit case $s = d/p$ it holds $W^{s,p}(D) \hookrightarrow L^q(D)$ for $1 \le q < \infty$.

In the other limit case $s = d/p + 1$ it holds $W^{s,p}(D) \hookrightarrow \mathcal{C}^\alpha(\overline{D})$ for all $\alpha \in [0,1)$. Note that these embeddings are automatically compact by compactness of the Hölder embeddings.

The same results hold on compact d-dimensional Lipschitz submanifold of $\mathbb{R}^d$ with the modification that only $s \in (0,1)$ is allowed.

For $p \in [1,\infty]$ we define the critical Sobolev exponent $p^* := dp/(d-p)$ for $p < d$ and $p^* := \infty$. By the general Sobolev embeddings, p^* is the largest number such that the embeddings $W^{1,q}(D) \hookrightarrow L^q(D)$ are bounded for $1 \le q < p^*$.

We will use the following vector-valued version of the Sobolev embeddings, see [42, §5.9.2]: For $p > 1$ it holds $W^{1,p}((0,T);X) \subset \mathcal{C}^0([0,T];X)$ and the embedding is compact. This embedding is also used to define the pointwise function evaluations $v(t)$ for $t \in [0,T]$ and $v \in W^{1,p}((0,T);X)$.

Trace Operator

Let D be a bounded Lipschitz domain in $\mathbb{R}^d$. Then for $s \in (0,\infty)$ and $p \in (0,1]$ satisfying $s - 1/p > 0$, the trace operator (see Section 2.1.1) is in fact a bounded linear operator $\gamma : W^{s,p}(D) \to W^{s-1/p,p}(\partial D)$. In particular, it is compact as an operator $\gamma : W^{s,p}(D) \to L^p(\partial D)$ [45, §1.5].

3 Problem Statement

3.1 Working Principle of Li-ion Batteries

This introduction to the electrochemical working principles is based on the information provided in [16, 52, 79].

A lithium-ion battery is an electrical power device consisting of two solid electrodes which are separated by an electrical separator soaked with a highly concentrated electrolyte. One of the electrodes consists of graphite particles and the other is composed of lithium metal oxide particles like $LiCoO_2$. The electrolyte is a solution of a lithium salt in an apriotic solvent, for example $LiPF_6$ or $LiBF_4$ in ethylene carbonate or dimethyl carbonate. The schematic construction of such a battery is shown in Fig. 3.1.

For the sake of an easier presentation, we will only consider the case of the battery getting *discharged*. Then the graphite electrode is the *anode* and the lithium metal oxide electrode is the *cathode* of the battery.

At both the anode and the cathode there is mounted a current collector which forms the negative and the positive pole of the battery, respectively.

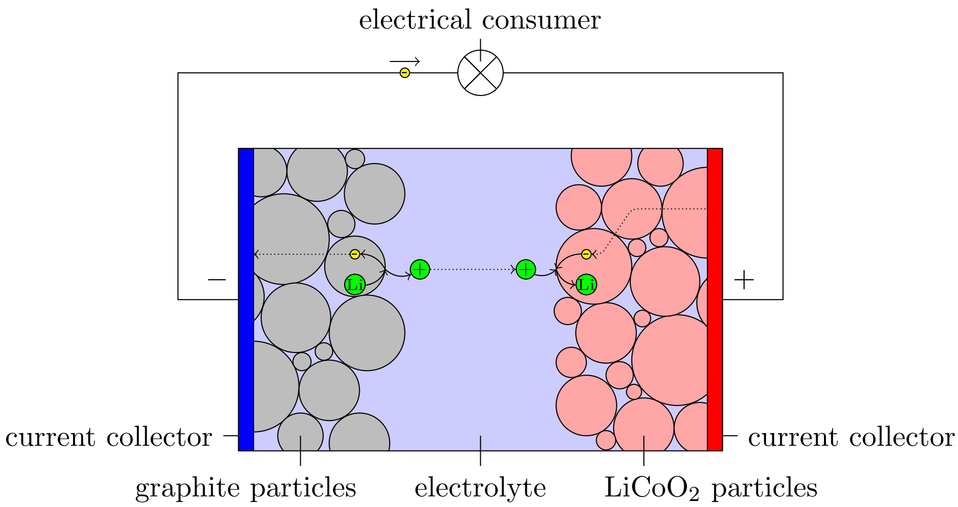

Figure 3.1: ⊖ symbolyze electrons, ⊕ Li^{+}-ions and Li Li-atoms.

When connecting the poles of the battery via an electrical consumer the battery gets discharged: Lithium-atoms migrate from the interior of the graphite particles to the interface between the graphite particle and the electrolyte. At this point, they are oxidized to Li^+-ions and enter the electrolyte. Within the electrolyte, these ions are then transported to the cathode. As soon as they reach the interface between electrolyte and cathode, they are reduced into metallic lithium atoms again which are then intercalated into the lithium oxide particles.

Simultaneous to the transport of lithium, electrons are transported inside the active particles of both anode and cathode and the negative ions from the lithium salt such as $PF_6{}^-$ or $BF_4{}^-$are transported inside the electrolyte. The electrons in the anode are transported away from the interface to the current collector and for every lithium atom being oxidized to a Li^+-ion at the particle-electrolyte interface and leaving the particle, one electron leaves the particles at the anode current collector. In the cathode particles, the transport is opposite: The electrons migrate from the current collector to the interface and for every Li^+-ion from the electrolyte being reduced into a metallic lithium atom at the interface and entering the particle, an electron enters the particle at the cathode current collector.

Within the electrolyte, the charge-transport is purely ionic. As the Li^+-ions move from cathode to anode, the negative ions from the lithium salt migrate into the other direction, that is, from the cathode to the anode. In contrast to the Li^+-ions, however, they are not involved in any electrochemical reaction at the interfaces between the electrolyte and the active particles. As a consequence, these interfaces act as a barrier for the negative ions from the lithium salt.

When applying an electrical voltage between the poles of the battery directed from the positive pole to the negative one, the battery gets charged. In this case, all the above described processes are reversed. Note, however, that in this case the terms anode and cathode are exchanged.

For further reading we refer to [69, 67, 66].

3.2 Review of Battery Models

There is a large number of mathematical models for Li-ion batteries in the literature.

The most simple ones describe the battery as an *equivalent electric current circle* composed of current sources, resistors and capacitors [18]. These models are cheap to solve numerically, however, they lack to give a deeper insight into the electrochemical mechanisms inside the battery.

In *porous-electrode-models*, the electrodes are considered as superimposed continua of the electrolyte and the active particles [29, 30]. They allow to model the electrochemical reactions inside the battery and give good predictions on the macroscopic performance of the battery. Early porous-electrode-models used the assumption of dilute electrolytes and thus were based on the Nernst–Planck equation [6, 7]. As this assumption is in fact not satisfied in the applications, recent works incorporated the theory of concentrated electrolytes into the model [37, 68]. It is possible to analyze the impact of particle size and basic arrangement with this model [25, 65]. However, it fails to provide a deeper understanding of the influence of the detailed electrode micro structure on the battery's behavior.

In contrast to that, the model we will be analyzing resolves the micro structure of the active particles in more detail. In this model, the domain of the battery is composed of three subdomains: the region of the active particles of the anode and the cathode, respectively, and the region of the electrolyte. The equations in the subdomains are coupled across the separating interface via the strongly nonlinear Butler–Volmer condition. In the electrolyte a thermodynamically consistent, refined version of the theory of concentrated electrolytes is used, resulting in a nonlinear coupling between the equation for charge and lithium transport [60, 58, 62].

This model can be considered more general than the porous-electrodes-model, since the latter can be derived from the first by homogenization techniques [22, 56].

Clearly the model is much more expensive to solve numerically than the porous-electrodes-model, as the complex particle geometry of the electrodes has to be resolved by a much finer mesh. However, it gives a more refined prediction of the battery performance and, more importantly, it serves as a starting point for more complex models including other physical effects like heat production [59, 57], intercalation stress [93, 19, 92, 53] or the coexistence of multiple phases in the active particles [90, 32, 33]. Note that it is known that mechanical stress [85, 84, 27] and temperature [82, 15] have a huge impact on the battery's performance and lifetime.

Since these effects are in general highly localized, a model based on volume-averaging can only give a rather rough prediction of them, compared to a model that resolves the electrodes' micro structure in detail.

3.3 The Model Equations

In this section we present the model equations. We will be following the articles [60, 57]. To simplify the presentation, we only treat the case of an electrochemical half-cell where the anode consists of metallic lithium. It will only be modeled by boundary conditions instead of the full transport equations and its complex micro structure [53].

3.3.1 Geometry and Notation

The region Ω of the battery is the union of the two disjoint open domains Ω_1 and Ω_2 in $\mathbb{R}^d$ with $d \geq 2$, see Fig. 3.2. They represent the electrolyte and the active particles in the cathode, respectively. The whole anode together with its current collector is represented by the boundary part $\Gamma_1 \subset \partial\Omega_1 \setminus \partial\Omega_2$. The area where the active particles touch the cathode current collectors is $\Gamma_2 \subset \partial\Omega_2 \setminus \partial\Omega_1$. The remainder of the battery's outer boundary is denoted by $\Gamma_0 := \partial\overline{\Omega} \setminus (\Gamma_1 \cup \Gamma_2)$. The interface between the active particles and the electrolyte is $I := \partial\Omega_1 \cap \partial\Omega_2$. Finally, we choose a unit normal ν on $\partial\Omega = \partial\overline{\Omega} \cup I$ such that it points outwards on $\partial\overline{\Omega}$ and from Ω_1 towards Ω_2 on the interface I.

3.3.2 Transport Equations

In this model, the unknown quantities are the lithium concentration $c : [0,T] \times \Omega \to \mathbb{R}$ satisfying $0 < c < c_{\max}$ and the electrical potential $\Phi : [0,T] \times \Omega \to \mathbb{R}$. Here, $c_{\max}$:

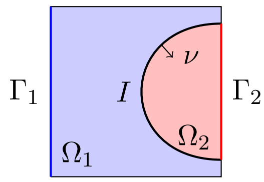

Figure 3.2: Decomposition of the battery.

$\Omega \to (0, \infty]$ is the given maximal concentration of lithium which is constant on Ω_1 and Ω_2 where it takes the value $c_{\max,1} = \infty$ and $c_{\max,2} \in (0, \infty)$, respectively.

Denoting by $\vec{N} : [0,T] \times \Omega \to \mathbb{R}^d$ the unknown lithium flux density and by $\vec{j} : [0,T] \times \Omega \to \mathbb{R}^d$ the unknown electrical current density, the conservation equations for lithium and charge read

$$\partial_t c + \nabla \cdot \vec{N} = 0 \qquad \text{on } (0,T) \times \Omega, \tag{3.1}$$

$$\nabla \cdot \vec{j} = 0 \qquad \text{on } (0,T) \times \Omega. \tag{3.2}$$

Note that in the equation for the charge conservation (3.2), already the important assumptions of local charge neutrality has entered the model. In the active particles this is due to the high mobility of the electrons and in the electrolyte this follows from the theory of concentrated electrolytes [68].

The fluxes $\vec{N}$ and $\vec{j}$ can be eliminated from the equations by expressing them in terms of c and Φ via the *constitutive relations*:

In the electrolyte, we use the version of the theory of concentrated electrolytes which has been derived in [60, 57] and which satisfies the second law of thermodynamics, namely, strictly positive entropy production [39]. The constitutive relations read

$$\vec{N}_1 = -D_1(x, c_1)\nabla c_1 + \frac{t_+(x, c_1)}{F}\vec{j}_1, \tag{3.3}$$

$$\vec{j}_1 = -\kappa_1(x, c_1)\nabla\Phi_1 - \frac{\kappa_1(x, c_1)t_+(x, c_1)}{F}\left(\frac{\partial\mu}{\partial c_1}\right)(x, c_1)\,\nabla c_1. \tag{3.4}$$

for $(t, x) \in (0, T) \times \Omega_1$, where $\vec{N}_1$, $\vec{j}_1$, c_1, Φ_1 and their gradients are evaluated at the point (t, x). Here, R is the universal gas constant, T is the temperature which we assume to be fixed and F is the Faraday constant. Additionally, $D_1(x, c_1) > 0$ is the inter diffusion coefficient and $t_+(x, c) \in (0, 1)$ is the so-called *transference number* of Li^+-ions. $\kappa_1(x, c_1) > 0$ is the ionic electric conductivity. The functions $D_1, \kappa_1 : \Omega \times (0, c_{\max,1}) \to (0, \infty)$ and $t_+ : \Omega \times (0, c_{\max,1}) \to (0, 1)$ are assumed to be given. They can be modeled by fitting a parametrized function to experimental data, for example.

Furthermore, $\mu(x, c_1)$ is the effective chemical potential of the Li-ions. The general form is $\mu(x, c_1) = \mu_0(x) + RT\ln(f_\pm(x, c_1)c_1)$ with the concentration-independent part $\mu_0 : \Omega_1 \to \mathbb{R}$ and the activity coefficient $f_\pm : \Omega_1 \times (0, c_{\max,1}) \to \mathbb{R}$ which are both given functions. In this work, however, we will only consider the case when $f_\pm$ is independent of

c, as it was done in [62, 70, 91] for example. As a consequence, we have $(\partial\mu/\partial c_1)(x, c_1) = RT/c_1$ and thus, (3.4) simplifies to

$$\vec{j}_1 = -\kappa_1(x, c_1)\nabla\Phi_1 - \frac{RT}{F}\frac{\kappa_1(x, c_1)t_+(x, c_1)}{c_1}\nabla c_1 \qquad \text{for } (t, x) \in (0, T) \times \Omega_1. \tag{3.5}$$

In the active particles the transport mechanisms are much simpler. Lithium transport is governed by Fick's law and charge transport is purely electronic and governed by Ohm's law. The constitutive relations thus read

$$\vec{N}_2 = -D_2(x, c_2)\nabla c_2 \qquad \text{for } (t, x) \in (0, T) \times \Omega_2, \tag{3.6}$$

$$\vec{j}_2 = -\kappa_2(x, c_2)\nabla\Phi_2 \qquad \text{for } (t, x) \in (0, T) \times \Omega_2, \tag{3.7}$$

where $D_2(x, c_2) > 0$ is the diffusion coefficient of lithium and $\kappa_2(x, c_2) > 0$ is the electronic conductivity. Again, the functions $D_2, \kappa_2 : \Omega_2 \times (0, c_{\max,2}) \to (0, \infty)$ are given.

Combining the functions D_1, κ_1 defined on Ω_1 and D_2, κ_2 defined on Ω_2 to global ones D, κ defined on $\Omega = \Omega_1 \cup \Omega_2$ and putting $t_+(x, c) = 0$ for $(x, c) \in \Omega_2 \times (0, c_{\max,2})$, we can write (3.3), (3.5), (3.6), (3.7) in a more compact way:

$$\vec{N} = -D(x, c)\nabla c + \frac{t_+(x, c)}{F}\vec{j} \qquad \text{for } (t, x) \in (0, T) \times \Omega, \tag{3.8}$$

$$\vec{j} = -\kappa(x, c)\nabla\Phi - \frac{RT}{F}\frac{\kappa(x, c)t_+(x, c)}{c}\nabla c \qquad \text{for } (t, x) \in (0, T) \times \Omega. \tag{3.9}$$

3.3.3 Interface Condition

The transport equations (3.8), (3.9) in Ω_1 and Ω_2 are coupled across the interface I via a nonlinear set of equations which will be explained in the following.

Since in the electrochemical reactions at the interface lithium and charge is conserved, see Section 3.1, the normal components of the lithium flux and of the electrical current are continuous across I:

$$\vec{N}_1 \cdot \nu = \vec{N}_2 \cdot \nu \qquad \text{on } (0, T) \times I,$$

$$\vec{j}_1 \cdot \nu = \vec{j}_2 \cdot \nu \qquad \text{on } (0, T) \times I.$$

On the other hand, the lithium flux and the electrical current are directly coupled: Consider for example the case when a lithium atom from a cathode particle is reduced to a positive lithium ion which then enters the electrolyte. The remaining electron remains in the particle and the charge of this electron is the negative of the elementary charge. Since this consideration also applies to the reverse reaction we have

$$\vec{N}_1 \cdot \nu = \frac{\vec{j}_1 \cdot \nu}{F} \qquad \text{on } (0, T) \times I.$$

Recall that F denotes the Faraday constant, that is, the negative amount of charge carried by one mole of electrons.

The final building block of the interface condition is an equation which expresses the velocity of the electrochemical reactions occurring at the interfaces in terms of the concentrations of the involved species and the electrical potential. In our case the *Butler-Volmer* equation is used in the version which is presented in [57, 60]. It reads

$$\vec{j}_2 \cdot \nu = -i_{\mathrm{BV}}(c_1, c_2, [\Phi]),$$

where $[\Phi] = \Phi_2 - \Phi_1$ denotes the jump of the electrical potential across I (along ν) and the exchange current density i_{BV} is given by

$$i_{\mathrm{BV}}(c_1, c_2, [\Phi]) = k c_1^{\alpha_1} c_2^{\alpha_1} (c_{\max,2} - c_2)^{\alpha_2} \left(e^{\frac{\alpha_1 F}{RT}([\Phi]-U(c_2))} - e^{\frac{-\alpha_2 F}{RT}([\Phi]-U(c_2))} \right). \tag{3.10}$$

Here, $k > 0$ is a given reaction rate and $\alpha_1, \alpha_2 \in (0,1)$ are the given anodic and cathodic transfer coefficients. Furthermore, $U(c_2)$ is the equilibrium open-circuit potential of the cathode. The function $U : (0, c_{\max,2}) \to \mathbb{R}$ is given. It can be obtained by measuring the battery's voltage when the half-cell is charged very slowly.

Other interface conditions can be found in the articles [68, 58].

3.3.4 Boundary Conditions

We take the most basic approach and model the situation when the battery is discharged (or charged) at a given electrical current density. Other situations are possible and of interest in the application, however, they should not change the basic mathematical properties of the problem.

By making the assumption that the current collectors are ideal conductors and that the electrical contact resistance the active particles and the current collector at the cathode is small, we obtain that the electrical potential Φ_2 needs to be constant in space on Γ_2. Since the electrical potential is only well-defined up to a constant, we will choose

$$\Phi_2 = 0 \qquad \text{on } (0,T) \times \Gamma_2.$$

Furthermore the charging of the battery at a given electrical current is realized by the following Neumann-condition:

$$\vec{j}_1 \cdot \nu = j^{\mathrm{ext}} \qquad \text{on } (0,T) \times \Gamma_1,$$

where $j^{\mathrm{ext}} : (0,T) \times \Gamma_1 \to \mathbb{R}$ is a given electrical current density. Note that $j^{\mathrm{ext}} < 0$ models an *influx of charge* and thus the *discharge* of the battery, see Fig. 3.1. For example we can set $j^{\mathrm{ext}}(t,x) = -I^{\mathrm{ext}}/\sigma(\Gamma_1)$ when we are modeling the charging of the battery at the constant macroscopic current I, say, $I = 1\mathrm{A}$.

Since the transport of charge and lithium is directly coupled across the interface between anode particles and the electrolyte, it is reasonable to impose as well the following:

$$\vec{N}_1 \cdot \nu = N^{\mathrm{ext}} := \frac{j^{\mathrm{ext}}}{F} \qquad \text{on } (0,T) \times \Gamma_1.$$

The current collector Γ_2 is a barrier for Lithium and the outer boundary Γ_0 is a barrier for both lithium and charge and thus we impose homogeneous Neumann boundary conditions for the lithium flux on $\Gamma_0 \cup \Gamma_2$ and for the electrical current on Γ_2:

$$\begin{aligned} \vec{N} \cdot \nu &= 0 \qquad \text{on } (0,T) \times (\Gamma_0 \cup \Gamma_2), \\ \vec{j} \cdot \nu &= 0 \qquad \text{on } (0,T) \times \Gamma_0. \end{aligned}$$

Finally, we require an initial condition for the lithium concentration:

$$c(0,\cdot) = c_0 \qquad \text{in } \Omega.$$

Where $c_0 : \Omega \to (0, c_{\max})$ is a given initial concentration. Note that we do not impose an initial condition for the electrical potential though, which is in accordance with the theory of elliptic-parabolic systems, see [87].

3.3.5 Summary of the Model

Let us very briefly collect all the equations defining our battery model.

Problem 3.3.1. *Find $c, \Phi : [0,T] \times \Omega \to \mathbb{R}$ such that the following conditions hold:*

1. $0 < c < c_{\max}$
2. *Lithium-conservation and local charge neutrality:*

$$\partial_t c + \nabla \cdot \vec{N} = 0, \qquad \text{in } (0,T) \times \Omega, \tag{3.11}$$

$$\nabla \cdot \vec{j} = 0, \qquad \text{in } (0,T) \times \Omega, \tag{3.12}$$

where $\vec{N}$ and $\vec{j}$ are given by:

$$\vec{N} = -D(x,c)\nabla c + \frac{t_+(x,c)}{F}\vec{j}, \tag{3.13}$$

$$\vec{j} = -\kappa(x,c)\nabla\Phi - \frac{RT}{F}\frac{\kappa(x,c)t_+(x,c)}{c}\nabla c \tag{3.14}$$

3. *Butler-Volmer-condition:*

$$\begin{aligned} \vec{N}_1 \cdot \nu = \vec{N}_2 \cdot \nu &= -\frac{i_{\mathrm{BV}}(c_1, c_2, [\Phi])}{F} \qquad \text{on } (0,T) \times I, \\ \vec{j}_1 \cdot \nu = \vec{j}_2 \cdot \nu &= -i_{\mathrm{BV}}(c_1, c_2, [\Phi]) \qquad \text{on } (0,T) \times I, \end{aligned}$$

where i_{BV} is given by:

$$i_{\mathrm{BV}}(c_1, c_2, [\Phi]) = k c_1^{\alpha_1} c_2^{\alpha_1} (c_{\max,2} - c_2)^{\alpha_2} \left(e^{\frac{\alpha_1 F}{RT}([\Phi] - U(c_2))} - e^{\frac{-\alpha_2 F}{RT}([\Phi] - U(c_2))} \right). \tag{3.15}$$

4. Boundary conditions

$$\begin{aligned}
\vec{N} \cdot \nu &= 0, & \vec{j} \cdot \nu &= 0 & &\text{on } (0,T) \times \Gamma_0, \\
\vec{N}_1 \cdot \nu &= N^{ext}, & \vec{j}_1 \cdot \nu &= j^{ext} & &\text{on } (0,T) \times \Gamma_1, \\
\vec{N}_2 \cdot \nu &= 0, & \Phi_2 &= 0 & &\text{on } (0,T) \times \Gamma_2.
\end{aligned} \tag{3.16}$$

5. Initial conditions

$$c(0, \cdot) = c_0 \qquad \text{in } \Omega.$$

3.4 Simplified Model

3.4.1 Assumptions

In order to simplify the analysis, we make several assumptions. Note that our numerical simulations are written for the general case, Problem 3.3.1. First we state the assumptions which are essential for the techniques we apply.

Assumption 3.4.1.

1. *t_+ is constant on $\Omega_i \times (0, c_{\max})$ for $i = 1, 2$.*
2. *$D(x, c)$ is independent of c.*

By the assumption on t_+ and the definition of $\vec{N}$, (3.13), we can plug in the divergence condition for the electrical current (3.12) into the equation for lithium transport (3.11) to obtain the following:

$$\begin{aligned}
\partial_t c &= \nabla \cdot (D(x,c)\nabla c) - \nabla \cdot \left(\frac{t_+(x,c)}{F} \vec{j} \right) \\
&= \nabla \cdot (D(x,c)\nabla c) - \frac{t_+(x,c)}{F} \nabla \cdot \vec{j} \\
&= \nabla \cdot (D(x,c)\nabla c).
\end{aligned} \tag{3.17}$$

As a result, the electrical potential Φ is eliminated from the equation for the lithium transport. Note that Assumption 3.4.1 is quite standard in the applications. For example in [62, 53, 25], constant values are used for $t_{+,1}$. As stated in Section 3.3.2, $t_{+,2} = 0$ is always assumed. In [37] we could find a non-constant transference number $t_{+,1}$. Examples for concentration-independent D are given in [37, 53], however several authors include the concentration-dependence into their considerations, see [62].

For the sake of a cleaner presentation, we make the following additional assumptions:

Assumption 3.4.2.

1. *$R = T = F = 1$*
2. *$k = 1$*

3. $c_{\max,1} = \infty$, $c_{\max,2} = 1$.

4. $\kappa(x, c) = \kappa_i(c)$ *is independent of* $x \in \Omega_i$ *for all* $c \in (0, c_{\max,i})$ *for* $i = 1, 2$.

5. $D(x, c) = 1$ *for* $(x, c) \in \Omega \times (0, c_{\max})$.

6. $j^{ext}(t, x)$ *does not depend on* t.

Note that Assumption 3.4.2 is in general not satisfied in the applications. However, in contrast to Assumption 3.4.1, it does not remove the key difficulties in the theoretical treatment of Problem 3.3.1. In fact, transferring our results to the case when Assumption 3.4.2 is not satisfied is straight forward.

3.4.2 Transformation

By the above assumptions, the transport equation for the lithium has been simplified to a heat equation (3.17) which is coupled in a nonlinear way to the electrical potential via the interface and boundary conditions.

Plugging in the definition (3.14) of $\vec{j}$, however, the equation for the charge-transport still reads

$$-\nabla \cdot \left(\kappa(c)\nabla\Phi + \frac{\kappa(c)t_+}{c}\nabla c \right) = 0 \qquad \text{in } (0, T) \times \Omega, \tag{3.18}$$

with the second order coupling term $\nabla \cdot (\frac{\kappa(c)t_+}{c}\nabla c)$. In order to get rid of this highest order coupling, we make the change of dependent variables $(c, \Phi) \to (c, u)$, where

$$u := \Phi + t_+ \ln(c). \tag{3.19}$$

This is inspired by the change of variables in [88] and [74]. Since $c > 0$, this is well-defined and we have by the chain-rule:

$$\nabla u = \nabla\Phi + t_+\nabla(\ln(c)) = \nabla\Phi + \frac{t_+}{c}\nabla c.$$

As a consequence, (3.18) reads

$$-\nabla \cdot \big(\kappa(c)\nabla u\big) = 0 \qquad \text{in } (0, T) \times \Omega, .$$

in the variables (c, u).

3.4.3 The Simplified Model Equations

With Assumption 3.4.1, Assumption 3.4.2 and the transformation (3.19), our model simplifies to the following:

Problem 3.4.3. *Find* $c, u : [0, T] \times \Omega \to \mathbb{R}$ *such that the following conditions hold:*

1. $0 < c_1$ *and* $0 < c_2 < 1$.

2. Lithium-conservation and local charge neutrality:

$$\partial_t c - \Delta c = 0 \qquad \text{in } (0,T)\times\Omega, \tag{3.20}$$
$$-\nabla\cdot(\kappa(c)\nabla u) = 0 \qquad \text{in } (0,T)\times\Omega. \tag{3.21}$$

3. Nonlinear interface condition:

$$\begin{aligned} \partial_\nu c_1 &= (1-t_+)i_{12}(c_1,c_2,[u]) && \text{on } (0,T)\times I,\\ \partial_\nu c_2 &= i_{12}(c_1,c_2,[u]) && \text{on } (0,T)\times I,\\ \kappa_i(c_i)\partial_\nu u_i &= i_{12}(c_1,c_2,[u]) && \text{on } (0,T)\times I, \end{aligned} \tag{3.22}$$

where i_{12} is (for example) given by $i_{12}(c_1,c_2,[u]) := i_{\mathrm{BV}}(c_1,c_2,[\Phi])$, see (3.15) *for the definition of the Butler–Volmer nonlinearity.*

4. Boundary conditions:

$$\begin{aligned} \partial_\nu c &= 0 & \kappa(c)\partial_\nu u &= 0 && \text{on } (0,T)\times\Gamma_0,\\ \partial_\nu c_1 &= (t_+-1)j^{ext} & \kappa_1(c_1)\partial_\nu u_1 &= -j^{ext} && \text{on } (0,T)\times\Gamma_1,\\ \partial_\nu c_2 &= 0 & u_2 &= 0 && \text{on } (0,T)\times\Gamma_2. \end{aligned}$$

5. Initial condition:

$$c(0,\cdot) = c_0 \qquad \text{in } \Omega \tag{3.23}$$

3.5 Regularity Assumptions

In this section we will state the precise regularity assumptions on the data. Let us begin with the rather weak requirements on the geometry.

Assumption 3.5.1 (Conditions on the geometry)**.**

1. Ω_1 *and* Ω_2 *are* disjoint bounded Lipschitz domains*.*
2. *The interface $I = \partial\Omega_1 \cap \Omega$ satisfies $\sigma(I) > 0$.*
3. *For $i = 1,2$, the boundary part Γ_i is a measurable subset of $\partial\Omega_i \setminus I$.*
4. *It holds $\sigma(\Gamma_2) > 0$.*

For some of our results it is actually necessary to require that the *Dirichlet boundary* Γ_2 and the Neumann-boundary $\partial\Omega_2\setminus\Gamma_2$ additionally satisfy a certain geometrical matching condition, namely, that they are *well-distributed*: The concept of *well-distributed* boundary parts of a Lipschitz-domain D in $\mathbb{R}^d$ has been introduced in [38]. Roughly speaking, a subset $S \subset \partial D$ and its complement $\partial D \setminus S$ are well-distributed in ∂D if the closure of S is a $(d-1)$-dimensional Lipschitz submanifold of ∂D with boundary. The precise definition is based on local coordinates:

Definition 3.5.2. *Let $D \subset \mathbb{R}^d$ be an open set and $S \subset \partial D$ be any subset of the boundary of D. S and $S^c := \partial D \setminus S$ are* well-distributed *in ∂D if for every $x \in \partial D$, there exists an open set U in $\mathbb{R}^d$ containing x and a bi-Lipschitz map*

$$\psi : U \to B := \{x \in \mathbb{R}^d : |x|_2 < 1\},$$

such that

$$\begin{cases} \psi(U \cap D) = B \cap \{x_1 > 0,\ x_2 > 0\} =: B_+ \text{ and} \\ \psi(U \cap \partial D) = \partial B_+ \cap \{x_1 x_2 = 0\} \end{cases}$$

holds, and additionally, either $U \cap \partial D \subset S$, $U \cap \partial D \subset S^c$ or

$$\begin{cases} \psi(U \cap \overline{S}) = \partial B_+ \cap \{x_1 = 0\} \text{ and} \\ \psi(U \cap \overline{S^c}) = \partial B_+ \cap \{x_2 = 0\}. \end{cases} \tag{3.24}$$

From Definition 3.5.2 it follows automatically that D is a Lipschitz-domain. Note that Definition 3.5.2 is not very restrictive and that it is satisfied in all reasonable applications.

Apart from the geometry, the only parameters entering the simplified model Problem 3.4.3, are the functions $\kappa_1, \kappa_2, i_{12}$ and j^{ext}. Let us state the precise assumptions on them.

Assumption 3.5.3 (Conditions on the data)**.**

1. *$j^{ext} : \Gamma_1 \to \mathbb{R}$ is Lebesgue measurable and essentially bounded.*
2. *$\kappa_i : (0, c_{\max,i}) \to (0, \infty)$ is locally Lipschitz continuous for $i = 1, 2$.*
3. *The nonlinearity*

$$\begin{aligned} i_{12} : (0,\infty) \times (0,1) \times \mathbb{R} &\to \mathbb{R} \\ (c_1, c_2, z) &\mapsto i_{12}(c_1, c_2, z) \end{aligned}$$

is continuously differentiable and it satisfies

$$\inf \big\{ \partial_z i_{12}(c_1, c_2, z) \mid c_1 \in K_1,\ c_2 \in K_2 \text{ and } z \in \mathbb{R} \big\} > 0 \tag{3.25}$$

for all compact sets $K_1 \subset (0, \infty)$ and $K_2 \subset (0, 1)$.

Clearly, the assumption 1 on j^{ext} is satisfied in all reasonable applications where j^{ext} is constant on time.

The assumption 2 on κ is satisfied in all the applications we know of. In all articles we read, κ_i is even $\mathcal{C}^\infty$ because it is a polynomial fitted to experimental data or simply a constant, see for example [62, 70, 53].

Let us discuss the assumption 3 on the nonlinearity i_{12}. First we consider the case that i_{12} is given by the Butler–Volmer condition (3.10). Recalling the definition (3.19) of the variable u, we have

$$\begin{aligned} i_{12}(c_1, c_2, z) &= i_{\text{BV}}(c_1, c_2, z + \ln(c_1)) \\ &= c_1^{1/2} c_2^{1/2} (1 - c_2)^{1/2} \Big(e^{(z + t_+ \ln(c_1) - U(c_2))/2} - e^{-(z + t_+ \ln(c_1) - U(c_2))/2} \Big). \end{aligned} \tag{3.26}$$

Clearly, i_{12} is $\mathcal{C}^1$ if the equilibrium potential $U : (0,1) \to \mathbb{R}$ is. However, this is satisfied in basically all the applications since most of the time U is a polynomial [62] or a linear combination of elementary smooth functions defined on $(0,\infty)$ [19, 26]. Additionally it holds

$$\begin{aligned}\partial_z i_{12}(c_1,c_2,z) &= c_1^{1/2} c_2^{1/2}(1-c_2)^{1/2}\left(\frac{1}{2}e^{(z+t_+\ln(c_1)-U(c_2))/2}+\frac{1}{2}e^{-(z+t_+\ln(c_1)-U(c_2))/2}\right)\\ &\geq \frac{1}{2}c_1^{1/2}c_2^{1/2}(1-c_2)^{1/2}\end{aligned}$$

for all $c_1 \in (0,\infty)$, $c_2 \in (0,1)$ and $z \in \mathbb{R}$. As a consequence (3.25) is satisfied for all compact sets $K_1 \subset (0,\infty)$ and $K_2 \subset (0,1)$.

Remark 3.5.4. *In the remainder of the thesis, it is assumed that Assumption 3.5.1 and Assumption 3.5.3 hold. Additionally, all the objects given by these assumptions will be considered constant, that is, we will write*

$$C(X) := C(\Omega_1,\Omega_2,\Gamma_1,\Gamma_2,j^{ext},\kappa_1,\kappa_2,i_{12},X)$$

for constants which depend on the geometry, the data and some object X.

We conclude this section with a remark which makes the statements on κ and i_{12} in Assumption 3.5.3 more concrete.

Remark 3.5.5. *For $M > 0$ define the compact set*

$$K_M := [1/M, M] \times [1/M, 1-1/M] \subset \mathbb{R}^2.$$

Then the following holds:

1. *There exists a positive constant $C_1 = C_1(M)$ such that it holds*

$$C_1^{-1} \leq \kappa_1(c_1),\, \kappa_2(c_2) \leq C_1 \qquad \text{and} \qquad \partial_z i_{12}(c_1,c_2,z) \geq C_1^{-1}$$

   for all $(c_1,c_2) \in K_M$ and all *$z \in \mathbb{R}$.*

2. *For all $R > 0$ there exists a positive constant $C_2 = C_2(M,R)$ such that it holds*

$$|\partial_{c_1} i_{12}(c_1,c_2,z)|,\, |\partial_{c_2} i_{12}(c_1,c_2,z)|,\, |\partial_z i_{12}(c_1,c_2,z)| \leq C_2$$

   for all $(c_1,c_2) \in K_M$ and all $z \in \mathbb{R}$ with $|z| \leq R$.

Proof. This follows directly from Assumption 3.5.3. □

4 A Strongly Nonlinear Elliptic Problem

In this section we consider a class of monotone elliptic problems with a strongly nonlinear interface condition and mixed boundary conditions, see Problem 4.1.1.

The motivation to study these elliptic problems is the fixed point argument for the fully coupled problem: Assuming the concentration $c(t) : \Omega \to (0, c_{\max})$ is given at a certain time $t \in [0, T)$, the problem to determine the potential $u(t) : \Omega \to \mathbb{R}$ satisfying (3.21), (3.22) and the boundary conditions for u in Problem 3.4.3 is called the elliptic subproblem (for the potential). Under certain regularity assumptions on c, which will be made precise in Chapter 5, this problem will fit into the framework of the current chapter.

The structure of this chapter is the following: Firstly, we will present the considered equations and the precise assumptions on the data in Section 4.1. Then we will prove well-posedness and uniform bounds for the solution in Section 4.2. By applying a linear regularity result, we will conclude the piecewise Hölder regularity of the solution in Section 4.3. In Section 4.4 we will show the continuity of the solution operator and in Section 4.5 we will prove a comparision principle.

4.1 Problem Statement

Throughout this chapter we will denote the outer Neumann boundary by $\Gamma_N := \partial\overline{\Omega} \setminus \Gamma_2$. The formal statement of the problem considered is the following:

Problem 4.1.1. *Find $u : \Omega \to \mathbb{R}$ such that the following holds:*

$$-\nabla \cdot (\kappa \nabla u) = G \qquad \text{in } \Omega, \tag{4.1}$$

$$\kappa_i \partial_\nu u_i = f(\cdot, [u]) \qquad \text{on } I, \tag{4.2}$$

$$\kappa \partial_\nu u = 0 \qquad \text{on } \Gamma_N, \tag{4.3}$$

$$u_2 = 0 \qquad \text{on } \Gamma_2. \tag{4.4}$$

Here, the data κ, G and f is assumed to satisfy Assumption 4.1.2. Note that the conditions in Assumption 4.1.2 are motivated by considering the case of the elliptic subproblem of Problem 3.4.3 as already indicated in the introduction to this chapter. The precise relation between Problem 4.1.1 and Problem 3.4.3 is given in Chapter 5, see in particular Remark 5.4.2.

Assumption 4.1.2. *There is a positive constant $M_1 > 0$ and a function $M_2 : (0, \infty) \to (0, \infty)$ such that the following conditions hold:*

1. $\kappa : \Omega \to \mathbb{R}$ *is measurable and*

$$\frac{1}{M_1} \le \kappa \le M_1 \tag{4.5}$$

holds almost everywhere in Ω*.*

2. $G \in W^{-1,\infty}$ *and it satisfies* $\|G\|_{-1,\infty;\Omega} \le M_1$*.*

3. f *is a function* $f : I \times \mathbb{R} \to \mathbb{R},\ (x, z) \mapsto f(x, z)$ *with the following properties:*

   a) $f(\cdot, z)$ *is measurable for all* $z \in \mathbb{R}$*.*

   b) $f(x, \cdot)$ *is continuously differentiable for* σ*-almost all* $x \in I$ *and it holds* $f(x, 0) = 0$,

$$\partial_z f(x, z) \ge 1/M_1 \qquad \text{for all } z \in \mathbb{R} \tag{4.6}$$

   and

$$|\partial_z f(x, z)| \le M_2(R) \qquad \text{for all } R > 0 \text{ and } z \in [-R, R]. \tag{4.7}$$

In the remainder of this chapter we will assume that Assumption 4.1.2 is satisfied for a fixed constant $M_1 > 0$ a fixed function $M_2 : (0, \infty) \to (0, \infty)$.

Note that condition 3 of Assumption 4.1.2 is equivalent to demanding that f is a Carathéodory function which is continuously differentiable with respect to z for almost all $x \in I$ and satisfies (4.6) and (4.7).

Despite the monotonicity property (4.6) of the nonlinearity f, Problem 4.1.1 does not fit completely into the framework of monotone operators due to the lack of a polynomial growth condition for f with respect to z at $\pm\infty$, cf. [89, Chapter 26]. Note that the growth condition (4.7) is only of qualitative nature. In fact, when applying the results of this chapter to the fixed point argument in Chapter 5, f will be a function growing exponentially with respect to z at $\pm\infty$.

Additionally, since G is a distributional right hand side, the Neumann conditions (4.2) and (4.3) in general will not be satisfied in a classical sense. Consider for example the case when G is given by

$$G(v) = \int_{\Gamma_N} g_n v \, \mathrm{d}\sigma \qquad \text{for all } v \in W^{1,1}$$

for a fixed function $g_n \in L^\infty(\Gamma_N)$. Then solutions of Problem 4.1.1 satisfy

$$\kappa_i \partial_\nu u_i = g_n \qquad \text{on } \partial\Omega_i \cap \Gamma_N$$

instead of (4.3).

These considerations point out that it is necessary to establish a mathematically precise formulation of Problem 4.1.1. This will be done in the next section.

4.2 Existence and Uniqueness

Formally multiplying (7.2) with a test function, integrating by parts in each subdomain and then plugging in (4.2) and (4.3), motivates us to state the following weak formulation:

Definition 4.2.1. *Let $V := H^1_{\Gamma_2}(\Omega)$ and $\|\cdot\|_V := \|\cdot\|_{1,2;\Omega}$. Then $u \in V$ is called a* weak solution *of Problem 4.1.1 if $f(\cdot,[u]) \in L^2(I)$ and*

$$\langle A(u), v\rangle := \int_\Omega \kappa \nabla u \cdot \nabla v \,\mathrm{d}x + \int_I f(\cdot,[u])[v]\,\mathrm{d}\sigma = G(v) \tag{4.8}$$

holds for all $v \in V$.

Note that by the trace-theorem and Assumption 4.1.2 all the integrals in (4.8) are finite. Before discussing uniqueness and existence of weak solutions, let us recall the following basic result which is crucial for our analysis, see for example [13]. We will give a brief proof using abstract compactness arguments.

Lemma 4.2.2. *Let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let S be a measurable subset of ∂D with $\sigma(S) > 0$. Then there exists a positive constant $C = C(D,S)$ such that*

$$\|v\|_{0,2;D} \le C\big(\|v\|_{0,2;S} + \|\nabla v\|_{0,2;D}\big)$$

holds for all $v \in H^1(D)$.

Proof. We define $F : H^1(D) \setminus \{0\} \to (0,\infty)$ by

$$F(v) := \frac{\|v\|_{0,2;S} + \|\nabla v\|_{0,2;D}}{\|v\|_{0,2;D}} \qquad \text{for } v \in H^1(D)\setminus\{0\}.$$

Then, F is well-defined and it satisfies

$$F(\alpha v) = F(v) \qquad \text{for all } \alpha \in \mathbb{R}\setminus\{0\}. \tag{4.9}$$

We will now show that

$$F_{\inf} := \inf\big\{F(v) \mid v \in H^1(D)\setminus\{0\}\big\} > 0.$$

To this end assume the contrary, that is, $F_{\inf} = 0$. Then from (4.9) it follows that there exists a sequence $(v_n)_{n\in\mathbb{N}}$ in $H^1(D)$ such that it holds $\|v_n\|_{1,2;D} = 1$ for $n \in \mathbb{N}$ and $F(v_n) \to 0$ for $n \to \infty$.

Since $H^1(D)$ is reflexive we can pick a subsequence of $(v_n)_n$ which converges weakly in $H^1(D)$, say, against $v \in H^1(D)$. Without loss of generality we assume $v_n \rightharpoonup v$ in $H^1(D)$ as $n \to \infty$. By the compactness of the Sobolev embedding $H^1(D) \hookrightarrow L^2(D)$ and the trace operator $H^1(D) \to L^2(S)$ we conclude $v_n \to v$ in $L^2(D)$ and $v_n|_S \to v|_S$ in $L^2(S)$ as $n \to \infty$. For $n \in \mathbb{N}$ it holds $\|v_n\|_{0,2;D} \le \|v_n\|_{1,2;D} = 1$ and thus

$$\|\nabla v_n\|_{0,2;D} \le \frac{\|\nabla v_n\|_{0,2;D}}{\|v_n\|_{0,2;D}} \le \frac{\|v_n\|_{0,2;S} + \|\nabla v_n\|_{0,2;D}}{\|v_n\|_{0,2;D}} = F(v_n).$$

As a consequence, it follows from $F(v_n) \to 0$ that also $\nabla v_n \to 0$ in $L^2(D)$ for $n \to \infty$. Since $v_n \rightharpoonup v$ in $H^1(D)$ and the gradient $\nabla : H^1(D) \to L^2(D)$ is a bounded linear operator, it also holds $\nabla v_n \rightharpoonup \nabla v$ in $L^2(D)$. By the uniqueness of weak limits, this implies $\nabla v = 0$, that is, v is constant, say, $v = c$ almost everywhere on Ω for some $c \in \mathbb{R}$.

From $1 = \|v_n\|^2_{1,2;D} = \|\nabla v_n\|^2_{0,2;D} + \|v_n\|^2_{0,2;D}$, $\nabla v_n \to 0$ in $L^2(D)$ and $v_n \to v = c$ in $L^2(D)$ we conclude $c \neq 0$. It follows:

$$F_{\text{inf}} = \lim_{n\to\infty} F(v_n) \geq \liminf_{n\to\infty} \frac{\|v_n\|_{0,2;S}}{\|v_n\|_{0,2;D}} = \frac{\|v\|_{0,2;S}}{\|v\|_{0,2;D}} = \frac{\|c\|_{0,2;S}}{\|c\|_{0,2;D}} = \frac{\sigma(S)}{\mu(D)} > 0$$

This, however, contradicts our assumption that $F_{\text{inf}} = 0$ and thus the proof is completed. □

An immediate consequence of Lemma 4.2.2 is that we can define an equivalent norm on V which is well-suited for the analysis of Problem 4.1.1:

Lemma 4.2.3. *For $v \in V$ let*

$$|v|_V := \left(\int_\Omega |\nabla v|^2 \, \mathrm{d}x + \int_I [v]^2 \, \mathrm{d}\sigma \right)^{1/2}.$$

Then $|\cdot|_V$ is a norm on V which is equivalent to $\|\cdot\|_V$, that is, there exists a positive constant C only depending on the geometry such that

$$\frac{1}{C}|v|_V \leq \|v\|_V \leq C|v|_V \tag{4.10}$$

holds for all $v \in V$.

Proof. From the trace theorem it follows that it suffices to show the estimate $|v|_V \geq C\|v\|_V$ for all $v \in V$ for some constant $C > 0$ that is independent of v. For this, let $v \in V$. Since V incorporates homogeneous Dirichlet boundary conditions on Γ_2, we have by the standard Poincaré estimate with homogeneous Dirichlet boundary conditions (see for example Section 2.2):

$$\|v_2\|_{1,2;\Omega_2} \leq C_{\mathrm{P}} \|\nabla v_2\|_{0,2;\Omega_2} \leq C_{\mathrm{P}} |v|_V, \tag{4.11}$$

for some positive constant C_P. It remains to bound the L^2-norm of v_1 by $|v|_V$. By Lemma 4.2.2, writing $v_1 = v_2 - [v]$ on I, the trace theorem and (4.11), we have

$$\begin{aligned} \|v_1\|_{0,2,\Omega_1} &\leq C_1 \big(\|\nabla v_1\|_{0,2;\Omega_1} + \|[v]\|_{0,2;I} + \|v_2\|_{0,2;I} \big) \\ &\leq 2C_1 \big(|v|_V + C_{\mathrm{T}} \|v_2\|_{1,2;\Omega_2} \big) \\ &\leq 2C_1 (1 + C_{\mathrm{T}} C_{\mathrm{P}}) |v|_V, \end{aligned}$$

where C_T denotes a positive constant from the trace-inequality. Collecting the estimates for v_1 and v_2 finishes the proof. □

Having established the equivalent norm $|\cdot|_V$, we can derive the uniqueness of the weak solution of Problem 4.1.1 by using the difference between two potential solutions as a test function:

Lemma 4.2.4. *There is at most one weak solution of Problem 4.1.1 in the sense of Definition 4.2.1.*

Proof. Let $u, \widetilde{u} \in V$ be weak solutions of Problem 4.1.1. Using $v = u - \widetilde{u}$ in both the equations (4.8) for u and $\widetilde{u}$ yields, after subtraction,

$$\int_\Omega \kappa |\nabla(u - \widetilde{u})|^2 \, \mathrm{d}x + \int_I \big(f(\cdot, [u]) - f(\cdot, [\widetilde{u}])\big)[u - \widetilde{u}] \, \mathrm{d}\sigma = 0. \tag{4.12}$$

It follows by Assumption 4.1.2 and the mean-value theorem:

$$\begin{aligned} |u - \widetilde{u}|_V^2 &= \int_\Omega |\nabla(u - \widetilde{u})|^2 \, \mathrm{d}x + \int_I [u - \widetilde{u}]^2 \, \mathrm{d}\sigma \\ &\leq M_1 \bigg(\int_\Omega \kappa |\nabla(u - \widetilde{u})|^2 \, \mathrm{d}x + \int_I \big(f(\cdot, [u]) - f(\cdot, [\widetilde{u}])\big)[u - \widetilde{u}] \, \mathrm{d}\sigma \bigg) = 0 \end{aligned}$$

This implies $u = \widetilde{u}$ in V by Lemma 4.2.3 which finishes the proof. □

To prove existence of weak solutions, we approximate the function $f : (x, z) \mapsto f(x, z)$ by functions f_ε for $\varepsilon > 0$ which are globally Lipschitz-continuous with respect to z. For $x \in I$ and $\varepsilon > 0$ we set $f_\varepsilon(x, \cdot)$ to be the unique $\mathcal{C}^1$-function which coincides with $f(x, \cdot)$ inside $[-R_\varepsilon, R_\varepsilon]$, $R_\varepsilon := 1/\varepsilon$, and is affine linear outside this interval, namely

$$f_\varepsilon(x, z) := \begin{cases} f(x, -R_\varepsilon) + \partial_z f(x, -R_\varepsilon)(z + R_\varepsilon), & z < -R_\varepsilon, \\ f(x, z), & |z| \leq R_\varepsilon, \\ f(x, R_\varepsilon) + \partial_z f(x, R_\varepsilon)(z - R_\varepsilon), & z > R_\varepsilon \end{cases} \tag{4.13}$$

for $x \in I$ and $z \in \mathbb{R}$. See Fig. 4.1 for an illustration of this construction.

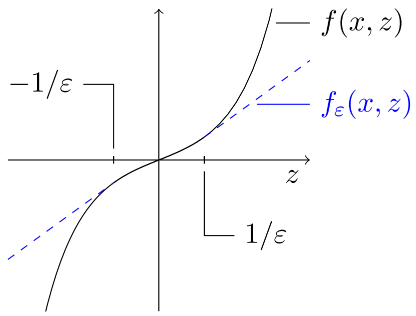

Figure 4.1: The functions $f(x, \cdot)$ and $f_\varepsilon(x, \cdot)$ for fixed $x \in I$

Note that by Assumption 4.1.2 the function $f(x, \cdot) : z \mapsto f(x, z)$ is continuously differentiable for almost all $x \in I$ and thus $f_\varepsilon(x, z)$ is well defined for almost all $x \in I$

and all $z \in \mathbb{R}$. An immediate but very important consequence of this definition is the following:

Remark 4.2.5. *For $\varepsilon > 0$ the function f_ε satisfies condition 3 of Assumption 4.1.2 with the same constant M_1. Moreover,*

$$\partial_z f_\varepsilon(x, z) \leq M_2(1/\varepsilon) \tag{4.14}$$

holds for σ-almost all $x \in I$ and all $z \in \mathbb{R}$.

For fixed $\varepsilon > 0$ we have thus defined an approximation f_ε to the nonlinearity f_ε which satisfies a suitable linear growth condition for $|z| \to \infty$ while maintaining the monotonicity property (4.6).

As a consequence, we can apply the theory of monotone operators to Problem 4.1.1 with f replaced by f_ε. We obtain the existence of unique weak solutions $u_\varepsilon \in V$ to the perturbed equations:

Remark 4.2.6. *For each $\varepsilon > 0$ there exists exactly one $u_\varepsilon \in V$ such that*

$$\int_\Omega \kappa \nabla u_\varepsilon \cdot \nabla v \, \mathrm{d}x + \int_I f_\varepsilon(\cdot, [u_\varepsilon])[v] \, \mathrm{d}\sigma = G(v) \tag{4.15}$$

holds for all $v \in V$. Moreover, there is a positive constant $C = C(M_1)$, independent of ε, such that it holds

$$\|u_\varepsilon\|_V \leq C. \tag{4.16}$$

Proof. Let $\varepsilon > 0$. We will apply the main theorem of monotone operators, see [89, §26.2], to the operator

$$\begin{aligned} &A_\varepsilon : V \to V', \\ &\langle A_\varepsilon(u), v\rangle = \int_\Omega \kappa \nabla u \cdot \nabla v \, \mathrm{d}x + \int_I f_\varepsilon(\cdot, [u])[v] \, \mathrm{d}\sigma \qquad \text{for } u, v \in V. \end{aligned}$$

For $u, \widetilde{u}, v \in V$ it follows by the global Lipschitz-continuity of f_ε with respect to z, see Remark 4.2.5, the Cauchy–Schwarz-inequality and the trace-theorem:

$$\begin{aligned} &|\langle A_\varepsilon(u) - A_\varepsilon(\widetilde{u}), v\rangle| \\ &\leq \int_\Omega |\kappa \nabla (u - \widetilde{u}) \cdot \nabla v| \, \mathrm{d}x + \int_I |f_\varepsilon(\cdot, [u]) - f_\varepsilon(\cdot, \widetilde{u})| \, |[v]| \, \mathrm{d}\sigma \\ &\leq M_1 \int_\Omega |\nabla (u - \widetilde{u}) \cdot \nabla v| \, \mathrm{d}x + M_2(1/\varepsilon) \int_I |[u - \widetilde{u}][v]| \, \mathrm{d}\sigma \\ &\leq (M_1 + M_2(1/\varepsilon) C_T) \|u - \widetilde{u}\|_V \|v\|_V, \end{aligned} \tag{4.17}$$

where C_T denotes a positive constant from the trace-inequality. Note that (4.17) holds especially for $\widetilde{u} = 0$. From the assumption $f(\cdot, 0) = 0$ and the construction of f_ε it follows $f_\varepsilon(\cdot, 0) = 0$ and thus $A_\varepsilon(0) = 0$. This shows that $A_\varepsilon : V \to V'$ is well-defined and continuous.

Furthermore, denoting by C_1 a positive constant from Lemma 4.2.3, we have for $u, v \in V$:

$$
\begin{aligned}
&\langle A_\varepsilon(u) - A_\varepsilon(v), u - v\rangle \\
&= \int_\Omega \kappa \nabla(u-v)\cdot\nabla(u-v)\,\mathrm{d}x + \int_I \big(f_\varepsilon(\cdot,[u]) - f_\varepsilon(\cdot,[v])\big)[u-v]\,\mathrm{d}\sigma \\
&\geq M_1^{-1}\Big(\int_\Omega |\nabla(u-v)|^2\,\mathrm{d}x + \int_I [u-v]^2\,\mathrm{d}\sigma\Big) \\
&\geq C_1^{-1}M_1^{-1}\|u-v\|^2.
\end{aligned}
\tag{4.18}
$$

This shows that A_ε is strictly monotone and taking $v = 0$ in (4.18) shows that A_ε is coercive. Note that the constant $C_2 := (C_1 M_1)^{-1}$ does not depend on ε.

Summing up, we have shown that $A_\varepsilon : V \to V'$ is a monotone, coercive and continuous operator on the real, separable, reflexive Banach space $V = H^1_{\Gamma_2}$. The main theorem of monotone operators, [89, Theorem 26.A], thus implies that A_ε is bijective and that the inverse operator $A_\varepsilon^{-1} : V' \to V$ satisfies

$$
\|A_\varepsilon^{-1}(b) - A_\varepsilon^{-1}(\widetilde{b})\|_V \leq C_2^{-1}\|b - \widetilde{b}\|_{V'} \tag{4.19}
$$

for all $b, \widetilde{b} \in V'$. Since $f_\varepsilon(\cdot, 0) = 0$, it holds $A_\varepsilon^{-1}(0) = 0$. Denoting by C_3 the operator-norm of the embedding $V \hookrightarrow W^{1,1}$ it thus follows for the unique solution $u_\varepsilon = A_\varepsilon^{-1}(G)$ of (4.15):

$$
\|u_\varepsilon\|_V = \|A_\varepsilon^{-1}(G) - 0\|_V \leq C_2^{-1}\|G - 0\|_{V'} \leq C_3 C_2^{-1}\|G\|_{-1,\infty;\Omega}.
$$

Since $C := C_3 C_2^{-1}$ only depends on M_1 (and in particular not on ε), the proof is finished. □

Now we will prove the main result of this chapter: The existence of a *bounded* weak solution to Problem 4.1.1. To this end we will use the Stampacchia truncation method ([76, §4]) to show a uniform in ε a-priori L^∞-bound for u_ε.

Theorem 4.2.7. *There exists a bounded weak solution $u \in V \cap L^\infty(\Omega)$ of Problem 4.1.1. It satisfies*

$$
\|u\|_V,\ \|u\|_{0,\infty;\Omega} \leq C \tag{4.20}
$$

with a positive constant $C = C(M_1)$ only depending on M_1.

Before presenting the proof of Theorem 4.2.7, we will present two elementary results which will be used in the proof. The first result combines basic properties of the positive part $z \mapsto z_+$ with the monotonicity of the nonlinearity f:

Remark 4.2.8. *Let $w_i \in \mathbb{R}$ and define $[w] := w_2 - w_1$ and $[v] := (w_2)_+ - (w_1)_+$. Then it holds*

$$
[w][v] \geq [v]^2. \tag{4.21}
$$

Proof. Let $P : \mathbb{R} \to [0, \infty)$ be the positive part, that is,

$$P(z) := z_+ \qquad \text{for } z \in \mathbb{R}.$$

Clearly, P is a monotone and non-expansive function. The monotonicity of P implies

$$[w][v] = (w_2 - w_1)(P(w_2) - P(w_1)) \geq 0, \tag{4.22}$$

and from the non-expansiveness of P we conclude

$$|[w][v]| = |w_2 - w_1|\,|P(w_2) - P(w_1)| \geq |P(w_2) - P(w_1)|^2 = [v]^2. \tag{4.23}$$

Combining (4.22) and (4.23) thus gives

$$[w][v] \geq [v]^2.$$

□

The second result states the existence of an explicit root of functions satisfying a recursive growth condition like (4.33). The proof can be found in the articles [77] and [75].

Lemma 4.2.9. *[76, Lemme 4.1] Let $k_0 \in \mathbb{R}$ and $\varphi : [k_0, \infty) \to [0, \infty)$ be a non-increasing function with the property that there exist $C, \alpha > 0$ and $\beta > 1$ such that it holds*

$$\varphi(h) \leq \frac{C}{(h-k)^\alpha}\big(\varphi(k)\big)^\beta \qquad \textit{for all } h > k \geq k_0. \tag{4.24}$$

Then $\varphi(k_1) = 0$, where $k_1 = k_1(\varphi(k_0), C, \alpha, \beta)$ is given explicitely by

$$k_1 = k_0 + C^{1/\alpha}\big(\varphi(k_0)\big)^{(\beta-1)/\alpha} 2^{\beta/(\beta-1)}. \tag{4.25}$$

Finally we present the proof of Theorem 4.2.7:

Proof of Theorem 4.2.7. Let $\varepsilon > 0$ and $u_\varepsilon \in V$ be the solution of (4.16). By $\lesssim$ we denote the relation $\lesssim_{M_1}$. Additionally, let $k > 0$ be arbitrary and define

$$v := (u_\varepsilon - k)_+.$$

It follows from the generalized chain-rule, cf. [44, §7.4], that $v \in H^1$ with

$$\nabla v = \chi_{\{u_\varepsilon > k\}} \nabla u_\varepsilon \qquad \text{almost everywhere on } \Omega. \tag{4.26}$$

Additionally, since $u_\varepsilon = 0$ holds σ-almost everywhere on Γ_2, we have

$$v = \max\{u_\varepsilon - k, 0\} = \max\{-k, 0\} = 0$$

σ-almost everywhere on Γ_2 and thus $v \in V = H^1_{\Gamma_2}(\Omega)$. As a consequence, we can use v as a test-function in (4.15) to obtain

$$\int_\Omega \kappa \nabla u_\varepsilon \cdot \nabla v \,\mathrm{d}x + \int_I f_\varepsilon(\cdot, [u_\varepsilon])[v]\,\mathrm{d}\sigma = G(v). \tag{4.27}$$

Since k is a constant, it holds $[u_\varepsilon] = [u_\varepsilon - k]$ on I. Thus it follows from (4.20), Assumption 4.1.2 and Remark 4.2.8 and finally Lemma 4.2.3:

$$\begin{aligned} \int_\Omega \kappa \nabla u_\varepsilon \cdot \nabla v \, \mathrm{d}x + \int_I f_\varepsilon(\cdot, [u_\varepsilon])[v] \, \mathrm{d}\sigma &= \int_\Omega \kappa |\nabla v|^2 \, \mathrm{d}x + \int_I f_\varepsilon(\cdot, [u_\varepsilon - k])[v] \, \mathrm{d}\sigma \\ &\gtrsim \int_\Omega |\nabla v|^2 \, \mathrm{d}x + \int_I [v]^2 \, \mathrm{d}\sigma \\ &\gtrsim \|v\|_V^2 = \|v\|_{1,2;\Omega}^2. \end{aligned} \tag{4.28}$$

Now define

$$A(k) := \{x \in \Omega : u_\varepsilon(x) > k\} \tag{4.29}$$

and $\varphi(k) := \mu(A(k))$. Then we obtain by Assumption 4.1.2 and Hölder's inequality:

$$\begin{aligned} |G(v)| &\lesssim \|v\|_{0,1;\Omega} + \|\nabla v\|_{0,1;\Omega} \\ &\le \mu(A(k))^{1/2} \left(\|v\|_{0,2;\Omega} + \|\nabla v\|_{0,2;\Omega}\right) \\ &\lesssim \varphi(k)^{1/2} \|v\|_{1,2;\Omega}. \end{aligned} \tag{4.30}$$

As a consequence, if $\|v\|_{1,2;\Omega} \neq 0$, we can combine (4.28), (4.27) and (4.30), to obtain

$$\|v\|_{1,2;\Omega} \lesssim \varphi(k)^{1/2}. \tag{4.31}$$

Clearly, (4.31) also holds when $\|v\|_{1,2;\Omega} = 0$.

Note that φ is non-increasing, since $A(\widetilde{k}) \subset A(k)$ holds for $\widetilde{k} > k \geq 1$. We will now show that φ satisfies (4.24). To this end, choose $q \in (2, 2^*)$ and let $\widetilde{k} > k \geq 1$. By Sobolev embedding, the inclusion $A(k) \supset A(\widetilde{k})$ and the definition of v and $A(\widetilde{k})$ it follows:

$$\begin{aligned} \|v\|_{1,2;\Omega}^2 \gtrsim \|v\|_{0,q;\Omega}^2 &= \left(\int_{A(k)} (u-k)^q \, \mathrm{d}x \right)^{2/q} \\ &\geq \left(\int_{A(\widetilde{k})} (u-k)^q \, \mathrm{d}x \right)^{2/q} \\ &\geq \left(\int_{A(\widetilde{k})} (\widetilde{k}-k)^q \, \mathrm{d}x \right)^{2/q} = (\widetilde{k}-k)^2 \big(\mu(A(\widetilde{k}))\big)^{2/q}. \end{aligned} \tag{4.32}$$

Recalling the definition of φ and combining this estimate with (4.31) gives:

$$\varphi(\widetilde{k}) \lesssim \frac{1}{(\widetilde{k}-k)^q} \|v\|_{1,2;\Omega}^q \lesssim \frac{1}{(\widetilde{k}-k)^q} \big(\varphi(k)\big)^{q/2}. \tag{4.33}$$

Thus, the assumptions of Lemma 4.2.9 are satisfied with some positive constant C, which only depends on M_1, $\alpha := q > 0$ and $\beta := q/2 > 1$.

It follows that $\varphi(k_1) = 0$, where $k_1 \geq 1$ is given explicitely by (4.25) and only depending on $\varphi(1)$, C, α and β, whereby the dependence on $\varphi(1)$ is non-decreasing and since $\varphi(1) \leq \mu(\Omega)$ we deduce $k_1 \lesssim 1$. By the definition of $A(k_1)$ we have thus shown that

$$u_\varepsilon \leq k_1 \lesssim 1 \qquad \text{holds almost everywhere in } \Omega. \tag{4.34}$$

A lower bound $u_\varepsilon \geq -k_2$ for some $0 < k_2 \lesssim 1$ follows from the observation that $-u_\varepsilon$ satisfies (4.15) with f_ε replaced by $(x, z) \mapsto (-f_\varepsilon(x, -z))$ and G replaced by $-G$.

We have thus shown that

$$\|u_\varepsilon\|_{0,\infty;\Omega} \lesssim \max\{k_1, k_2\} =: C_3. \tag{4.35}$$

Since $u_\varepsilon \in H^1$ it follows that $u_\varepsilon|_{\partial\Omega}$ satisfies the same pointwise estimate, namely $\|u_\varepsilon\|_{0,\infty;\partial\Omega} \leq C_3$. In particular, for $\varepsilon = 1/C_3$, we have $f_\varepsilon(\cdot, [u_\varepsilon]) = f(\cdot, [u_\varepsilon])$, that is, u_ε satisfies $f(\cdot, [u_\varepsilon]) \in L^\infty(I) \subset L^2(I)$ and it solves (4.8) with the original nonlinearity f.

Thus we have shown the existence of a weak solution of Problem 4.1.1, namely $u = u_\varepsilon$ with $\varepsilon = 1/C_3$. The uniform $\|\cdot\|_V$-estimate in (4.20) follows from Remark 4.2.6. □

4.3 Hölder Regularity

Having established the existence of bounded weak solutions we can now apply the regularity results from [38] to conclude that the weak solution of Problem 4.1.1 is in fact Hölder continuous in each subdomain and that there are uniform bounds for both the Hölder exponent and the Hölder norm.

Note that we do not only point out the Hölder regularity of solutions for its own sake. Instead it is necessary to conclude that the solution operator to Problem 4.1.1 is continuous with respect to the L^∞-norm, Lemma 4.4.2, which is required to apply the Schauder fixed point theorem in the proof of the existence result for the fully coupled Problem 3.4.3 (Theorem 5.6.1).

To apply the results from [38] we need an additional geometrical assumption, which was not necessary for the wellposedness results in Section 4.2.

Lemma 4.3.1. *Assume Γ_2 and $\partial\Omega_2 \setminus \Gamma_2$ are well-distributed in $\partial\Omega_2$ in the sense of Definition 3.5.2 and denote by u the weak solution of Problem 4.1.1. Then there exists a Hölder-exponent $\delta = \delta(M_1) \in (0, 1)$ and a positive constant $C = C(M_1, M_2)$ such that*

$$u_i \in \mathcal{C}^\delta(\overline{\Omega_i}) \qquad \text{and} \qquad [\![u_i]\!]_{\delta;\Omega_i} \leq C \tag{4.36}$$

hold for $i = 1, 2$.

Proof. We apply the Hölder-regularity results from [38] on Ω_1 and Ω_2 separately.

First, let $v_1 \in H^1(\Omega_1)$ be arbitrary. Using $v = (v_1, 0) \in V$ in (4.8) and defining $\widetilde{u}_1 := u_1 - \fint_{\Omega_1} u_1 \,\mathrm{d}x$, we obtain

$$\int_\Omega \kappa_1 \nabla \widetilde{u}_1 \cdot \nabla v_1 \,\mathrm{d}x = \int_I f(\cdot, [u]) v_1 \,\mathrm{d}\sigma + G(v_1, 0) =: \widetilde{G}_1(v_1).$$

Since $\|u\|_{0,\infty;I} \lesssim_{M_1} 1$, by (4.20), we obtain from Assumption 4.1.2:

$$\|f(\cdot,[u])\|_{0,\infty;I} \lesssim_{M_1,M_2} 1.$$

As a consequence, $\widetilde{G}_1$ is an element of $W^{-1,\infty}$ satisfying

$$\|\widetilde{G}_1\|_{-1,\infty;\Omega_1} \lesssim_{M_1,M_2} 1.$$

Now [38, Theorem 2.2] implies that there is some $\delta_1 = \delta_1(M_1) \in (0,1)$ such that $\widetilde{u}_1 \in \mathcal{C}^{\delta_1}(\overline{\Omega}_1)$ and

$$[\![\widetilde{u}_1]\!]_{\delta_1;\Omega_1} \lesssim_{M_1,M_2} 1$$

holds. It follows $u_1 = \widetilde{u}_1 + \int_{\Omega_1} u_1 \,\mathrm{d}x \in \mathcal{C}^{\delta_1}(\overline{\Omega_1})$ and, again from (4.20):

$$[\![u_1]\!]_{\delta_1;\Omega_1} \le [\![\widetilde{u}_1]\!]_{\delta_1;\Omega_1} + \mu(\Omega_1)\|u_1\|_{0,\infty;\Omega} \lesssim_{M_1,M_2} 1.$$

To show the Hölder regularity of u_2, note that for $v_2 \in H^1_{\Gamma_2}(\Omega_2)$ it follows from (4.8):

$$\int_\Omega \kappa_2 \nabla u_2 \cdot \nabla v_2 \,\mathrm{d}x = -\int_I f(\cdot,[u]) v_2 \,\mathrm{d}\sigma + G(0,v_2) =: \widetilde{G}_2(v_2).$$

As for $\widetilde{G}_2$, we can again argue that $\widetilde{G}_2 \in W^{-1,\infty}(\Omega_2)$ with

$$\|\widetilde{G}_1\|_{-1,\infty;\Omega_2} \lesssim_{M_1,M_2} 1.$$

Since Γ_2 and $\partial\Omega_2 \setminus \Gamma_2$ are assumed to be well-separated in the sense of Definition 3.5.2, the assumptions of [38, Theorem 2.1] are satisfied and it follows that there exists some $\delta_2 = \delta_2(M_1) \in (0,1)$ such that $u_2 \in \mathcal{C}^{\delta_2}(\overline{\Omega}_2)$ and

$$[\![u_2]\!]_{\delta_2;\Omega_2} \lesssim_{M_1,M_2} 1.$$

Defining $\delta := \min\{\delta_1, \delta_2\}$ thus finishes the proof. □

4.4 Mapping properties of the solution operator

In this section we write $\boldsymbol{U}(\kappa, f, G) := u$ to express the dependence of the weak solution of Problem 4.1.1 on the data κ, f and G.

Let us denote by $\mathfrak{D}_{M_1,M_2}$ the space of all triples (κ, f, G) satisfying Assumption 4.1.2 for this M_1 and M_2, equipped with the family of metrics which are induced by the norms $\|\cdot\|_R$ defined by

$$|\!|\!|(\kappa, f, G)|\!|\!|_R := \|\kappa\|_{0,\infty;\Omega} + \sup_{|z|\le R} \|f(\cdot,z)\|_{-1/2,2;I} + \|G\|_{-1,2;\Omega} \qquad \text{for } R > 0.$$

We then consider the solution operator $\boldsymbol{U}$ defined on the space $\mathfrak{D}_{M_1,M_2}$.

In Lemma 4.4.1 we show the Lipschitz continuity of $\boldsymbol{U}$ with respect to H^1, and in Lemma 4.4.2 we will derive the continuity of $\boldsymbol{U}$ with respect to the Hölder norms for a suitable Hölder exponent.

The results in this section are immediate consequences of Theorem 4.2.7, Lemma 4.3.1 and the uniform estimates therein, respectively. We will apply Lemma 4.4.1 in the proof of the existence result Theorem 5.6.1 and Lemma 4.4.1 in the proof of the uniqueness result Theorem 5.7.1 for the fully coupled problem Problem 3.4.3.

Lemma 4.4.1. *Let* $(\kappa, f, G), (\widetilde{\kappa}, \widetilde{f}, \widetilde{G}) \in \mathfrak{D}_{M_1,M_2}$ *and define* $u := \boldsymbol{U}(\kappa, f, G)$ *and* $\widetilde{u} := \boldsymbol{U}(\widetilde{\kappa}, \widetilde{f}, \widetilde{G})$*. Then it holds*

$$\begin{aligned} &\|u-\widetilde{u}\|_{1,2;\Omega} \\ &\le C_1\Big(\|\kappa-\widetilde{\kappa}\|_{0,\infty;\Omega} + \sup_{|z|\le C_1} \|(f-\widetilde{f})(\cdot,z)\|_{-1/2,2,I} + \|G-\widetilde{G}\|_{-1,2;\Omega}\Big). \end{aligned} \tag{4.37}$$

for some positive constant $C_1 = C_1(M_1)$ *only depending on* M_1*.*

Proof. Let us denote by $\lesssim$ the relation $\lesssim_{M_1}$. First we consider the case $f = \widetilde{f}$ and $G = \widetilde{G}$. Using $v = u - \widetilde{u}$ in the defining equation (4.8) for both u and $\widetilde{u}$ gives

$$\begin{aligned} 0 = &\int_\Omega \kappa\nabla u\cdot\nabla(u-\widetilde{u})\,\mathrm{d}x - \int_\Omega \widetilde{\kappa}\nabla\widetilde{u}\cdot\nabla(u-\widetilde{u})\,\mathrm{d}x \\ &+ \int_I f(\cdot,[u])[u-\widetilde{u}]\,\mathrm{d}\sigma - \int_I f(\cdot,[\widetilde{u}])[u-\widetilde{u}]\,\mathrm{d}\sigma \\ = &\int_\Omega \kappa|\nabla(u-\widetilde{u})|^2\,\mathrm{d}x + \int_I \big(f(\cdot,[u]) - f(\cdot,[\widetilde{u}])\big)[u-\widetilde{u}]\,\mathrm{d}\sigma \\ &+ \int_\Omega (\kappa-\widetilde{\kappa})\nabla\widetilde{u}\cdot\nabla(u-\widetilde{u})\,\mathrm{d}x. \end{aligned} \tag{4.38}$$

Using Lemma 4.2.3, rearranging (4.38) and using (4.5) and (4.20) we obtain

$$\begin{aligned} \|u-\widetilde{u}\|_V^2 &\lesssim \int_\Omega \kappa|\nabla(u-\widetilde{u})|^2\,\mathrm{d}x + \int_I \big(f(\cdot,[u]) - f(\cdot,[\widetilde{u}])\big)[u-\widetilde{u}]\,\mathrm{d}\sigma \\ &\le \left|\int_\Omega (\kappa-\widetilde{\kappa})\nabla\widetilde{u}\cdot\nabla(u-\widetilde{u})\,\mathrm{d}x\right| \\ &\lesssim \|\kappa-\widetilde{\kappa}\|_{0,\infty;\Omega}\|u-\widetilde{u}\|_V. \end{aligned}$$

Dividing by $\|u-\widetilde{u}\|_V$ shows the claimed estimate in this case.

Now we consider the case $\kappa = \widetilde{\kappa}$ and $G = \widetilde{G}$. Again testing (4.8) for both u and $\widetilde{u}$ with $v = u - \widetilde{u}$ yields:

$$\begin{aligned} 0 &= \int_\Omega \kappa|\nabla(u-\widetilde{u})|^2\,\mathrm{d}x + \int_I f(\cdot,[u])[v]\,\mathrm{d}\sigma - \int_I \widetilde{f}(\cdot,[\widetilde{u}])[v]\,\mathrm{d}\sigma \\ &= \int_\Omega \kappa|\nabla(u-\widetilde{u})|^2\,\mathrm{d}x + \int_I \big(f(\cdot,[u]) - f(\cdot,[\widetilde{u}])\big)[v]\,\mathrm{d}\sigma \\ &\quad+ \int_I \big(f(\cdot,[\widetilde{u}]) - \widetilde{f}(\cdot,[\widetilde{u}])\big)[v]\,\mathrm{d}\sigma \end{aligned} \tag{4.39}$$

By Theorem 4.2.7 there exists a constant $C_1 = C_1(M_1)$, only depending on M_1, such that $\|\widetilde{u}\|_{0,\infty;\Omega} \le C_1$. Thus it follows from Lemma 4.2.3, (4.5), (4.39) and the trace-theorem:

$$\begin{aligned}\|u-\widetilde{u}\|_V^2 &\lesssim \int_\Omega \kappa|\nabla(u-\widetilde{u})|^2\,\mathrm{d}x + \int_I \big(f(\cdot,[u]) - f(\cdot,[\widetilde{u}])\big)[u-\widetilde{u}]\,\mathrm{d}\sigma \\ &\le \left|\int_I \big(f(\cdot,[\widetilde{u}]) - \widetilde{f}(\cdot,[\widetilde{u}])\big)[u-\widetilde{u}]\,\mathrm{d}\sigma\right| \\ &\lesssim \sup_{|z|\le C_1} \|(f-\widetilde{f})(\cdot,z)\|_{-1/2,2,I}\,\|u-\widetilde{u}\|_V.\end{aligned}$$

Again we can divide by $\|u-\widetilde{u}\|_V$ to obtain the claimed estimate.

Now in the case $\kappa = \widetilde{\kappa}$ and $G = \widetilde{G}$, we apply the same technique as above and obtain

$$\begin{aligned}\|u-\widetilde{u}\|_V^2 &\lesssim \int_\Omega \kappa|\nabla(u-\widetilde{u})|^2\,\mathrm{d}x + \int_I \big(f(\cdot,[u]) - f(\cdot,[\widetilde{u}])\big)[u-\widetilde{u}]\,\mathrm{d}\sigma \\ &= (G-\widetilde{G})(u-\widetilde{u}) \le \|G-\widetilde{G}\|_{-1,2;\Omega}\|u-\widetilde{u}\|_V.\end{aligned}$$

The general case can be reduced to the three cases considered by writing

$$\begin{aligned}&u(\kappa,f,G) - u(\widetilde{\kappa},\widetilde{f},\widetilde{G}) \\ &= \big(\boldsymbol{U}(\kappa,f,G) - \boldsymbol{U}(\widetilde{\kappa},f,G)\big) + \big(\boldsymbol{U}(\widetilde{\kappa},f,G) - \boldsymbol{U}(\widetilde{\kappa},\widetilde{f},G)\big) + \big(\boldsymbol{U}(\widetilde{\kappa},\widetilde{f},G) - \boldsymbol{U}(\widetilde{\kappa},\widetilde{f},\widetilde{G})\big)\end{aligned}$$

and using the triangle-inequality. □

Lemma 4.4.2. *Assume Γ_2 and $\partial\Omega_2$ are well-distributed in the sense of Definition 3.5.2. Let $(\kappa,f,G),(\kappa_n,f_n,G_n) \in \mathfrak{D}_{M_1,M_2}$ for $n\in\mathbb{N}$ such that it holds*

$$\|\kappa_n-\kappa\|_{0,\infty;\Omega} + \sup_{|z|\le R}\|f_n(\cdot,z)-f(\cdot,z)\|_{-1/2,2;I} + \|G_n-G\|_{-1,2;\Omega} \to 0 \qquad \text{as } n\to\infty$$

for all $R>0$. *Then it holds*

$$\|\boldsymbol{U}(\kappa_n,f_n,G_n) - \boldsymbol{U}(\kappa,f,G)\|_{\mathcal{C}_\mathrm{b}^{\widetilde{\delta}}} \to 0 \qquad \text{as } n\to\infty$$

for all $\widetilde{\delta}\in(0,\delta)$, where $\delta=\delta(M_1)\in(0,1)$ is a Hölder-exponent from Lemma 4.3.1.

Proof. Let us assume the contrary and define $u := \boldsymbol{U}(\kappa,f,G)$ and $u_n := \boldsymbol{U}(\kappa_n,f_n,G_n)$ for $n\in\mathbb{N}$. Then, by selecting a subsequence, we can assume without loss of generality that there is some $\varepsilon>0$ such that

$$\|u_n-u\|_{\mathcal{C}_\mathrm{b}^{\widetilde{\delta}}} \ge \varepsilon \qquad \text{holds for all } n\in\mathbb{N}. \tag{4.40}$$

By Lemma 4.3.1, $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\mathcal{C}_\mathrm{b}^\delta$ and by the compactness of the embedding $\mathcal{C}_\mathrm{b}^\delta \hookrightarrow \mathcal{C}_\mathrm{b}^{\widetilde{\delta}}$, $(u_n)_{n\in\mathbb{N}}$ contains a subsequence which converges in $\mathcal{C}_\mathrm{b}^{\widetilde{\delta}}$. Again, without loss of generality, we assume that $(u_n)_n$ itself converges in $\mathcal{C}_\mathrm{b}^{\widetilde{\delta}}$, say

$$u_n \to \widetilde{u} \text{ in } \mathcal{C}_\mathrm{b}^{\widetilde{\delta}} \qquad \text{as } n\to\infty. \tag{4.41}$$

We will now show that $u = \widetilde{u}$ holds in Ω, which completes the proof, since it contradicts (4.40) and (4.41). Note that, by Lemma 4.4.1,

$$u_n \to u \text{ in } H^1(\Omega) \qquad \text{as } n \to \infty. \tag{4.42}$$

Since the embeddings $H^1(\Omega) \hookrightarrow L^2(\Omega)$ and $\mathcal{C}_{\mathrm{b}}^{\delta} \hookrightarrow L^2(\Omega)$ are continuous, it follows from (4.41) and (4.42):

$$u_n \to u \qquad \text{and} \qquad u_n \to \widetilde{u}$$

in $L^2(\Omega)$ as $n \to \infty$. As a consequence, $u = \widetilde{u}$ holds almost everywhere in Ω. Since both u and $\widetilde{u}$ are continuous it follows $u = \widetilde{u}$ and thereby, the proof is finished. □

4.5 Comparison Principle

In this section we prove a comparison principle for Problem 4.1.1, that is, the pointwise estimate $\underline{u} \leq u \leq \overline{u}$, where u is the weak solution, $\underline{u}$ a weak sub- and $\overline{u}$ a weak supersolution of Problem 4.1.1. The precise definition of sub- and supersolutions is given in Definition 4.5.1.

First of all, it can be considered as an interesting property on its own as it is known that standard quasilinear elliptic equations of second order satisfy comparision principles as well. An immediate consequence of comparison principles is the uniqueness of the solution because every solution is both a sub- and supersolution. Moreover they can be used to establish maximum principles which then imply a priori bounds for the solution in the L^∞-norm. Finally these a priori bounds can be used in fixed point methods to construct solutions. For the details of these arguments see for example [44, Chapter 10].

Secondly we point out the comparison principle here for *historical* reasons: When we started working on the topic, inspired by [74] and [88], we used the Moser iteration technique ([64]) to prove the uniform L^∞-bound for the approximate solutions u_ε. For technical reasons our proof only worked for $d \leq 3$.

One attempt to generalize the result to arbitrary $d \geq 4$ was to use the comparison principle and construct explicit bounded sub- and supersolutions. In fact we succeeded in proving the comparision principle but we could only construct bounded comparison solutions in very simple cylindrical geometries and under additional cumbersome assumptions on the data.

However the proof of the comparison principle finally motivated us to review the Stampacchia truncation method in [76] and so we were able to generalize the result to arbitrary dimension eventually.

Let us provide our definition of sub- and supersolutions.

Definition 4.5.1. $\widetilde{u} \in W := H^1(\Omega)$ *is called a* weak subsolution (supersolution) *of Problem 4.1.1 if the following conditions are satisfied:*

1. $\widetilde{u}_2 \leq 0$ $(\widetilde{u}_2 \geq 0)$ *holds* σ*-almost everywhere on* Γ_2,
2. $f(\cdot, [\widetilde{u}]) \in L^2(I)$,

3. For all $v \in V$ satisfying $v \geq 0$ almost everywhere in Ω it holds

$$\int_\Omega \kappa \nabla \widetilde{u} \cdot \nabla v \, \mathrm{d}x + \int_I f(\cdot, [\widetilde{u}]), [v] \, \mathrm{d}\sigma \leq (\geq) G(v). \tag{4.43}$$

Theorem 4.5.2. *Let $\underline{u}$ be a weak subsolution, $\overline{u}$ be a weak supersolution and u be the weak solution of Problem 4.1.1. Then*

$$\underline{u} \leq u \leq \overline{u} \tag{4.44}$$

holds almost everywhere in Ω.

Proof. Let $v := (\underline{u} - u)_+$. As in the proof of Theorem 4.2.7, it follows $v \in V$ and

$$\nabla v = \chi_{\{\underline{u} > u\}} \nabla(\underline{u} - u) \qquad \text{almost everywhere on } \Omega. \tag{4.45}$$

Thus we can use v in both (4.8) and (4.44) and obtain after subtracting:

$$\int_\Omega \kappa \nabla(\underline{u} - u) \cdot \nabla v \, \mathrm{d}x + \int_I \big(f(\cdot, [\underline{u}]) - f(\cdot, [u])\big)[v] \, \mathrm{d}\sigma \leq 0. \tag{4.46}$$

On the other hand it follows from (4.45), Assumption 4.1.2 and Remark 4.2.8:

$$\begin{aligned} &\int_\Omega \kappa \nabla(\underline{u} - u) \cdot \nabla v \, \mathrm{d}x + \int_I \big(f(\cdot, [\underline{u}]) - f(\cdot, [u])\big)[v] \, \mathrm{d}\sigma \\ &\geq \int_\Omega \kappa |\nabla v|^2 \, \mathrm{d}x + M_1^{-1} \int_I [\underline{u} - u][v] \, \mathrm{d}\sigma \\ &\geq M_1^{-1} \int_\Omega |\nabla v|^2 \, \mathrm{d}x + M_1^{-1} \int_I [v]^2 \, \mathrm{d}\sigma. \end{aligned} \tag{4.47}$$

Finally, combining (4.46) and (4.47) with Lemma 4.2.3 it follow $\|v\|_V = 0$. By the definition of v, this implies $\underline{u} \leq u$ almost everywhere in Ω, which finishes the proof. □

5 The Fully Coupled Problem

In this chapter we treat the fully coupled system, Problem 3.4.3. We will prove the main result Theorem 5.6.1 which states there exists a $T > 0$ such that Problem 3.4.3 has a weak solution on $(0, T)$ in a certain weak sense, defined in Section 5.3. Furthermore we will show that this solution is unique for $d \leq 3$, see Theorem 5.7.1.

Throughout this chapter we will assume that the data from Section 3.5 is given such that both the geometrical conditions from Assumption 3.5.1 and the regularity conditions from Assumption 3.5.3 are satisfied. We will also omit the dependence on this data, that is, we will consider objects as constant if the *only* depend on this data, compare Remark 3.5.4.

In addition we will impose the following geometrical matching condition between the Dirichlet and Neumann boundary:

Assumption 5.0.1. *Γ_2 and $\partial\Omega_2 \backslash \Gamma_2$ are* well-distributed *in the sense of Definition 3.5.2.*

Let us rephrase Problem 3.4.3. However, in order to emphasize the elliptic-parabolic structure, we rearrange the equations.

Problem 3.4.3. *Find $c, u : [0, T] \times \Omega \to \mathbb{R}$ such that the following conditions hold:*

1. *Lithium-transport:* $0 < c < c_{\max}$ *and*

$$\begin{aligned} \partial_t c - \Delta c &= 0 && \text{in } (0,T) \times \Omega, \\ \partial_\nu c_1 &= (1 - t_+) i_{12}(c, [u]) && \text{on } (0,T) \times I, \\ \partial_\nu c_2 &= i_{12}(c, [u]) && \text{on } (0,T) \times I, \\ \partial_\nu c &= 0 && \text{on } (0,T) \times \Gamma_0, \\ \partial_\nu c &= (t_+ - 1) j^{ext} && \text{on } (0,T) \times \Gamma_1, \\ \partial_\nu c_2 &= 0 && \text{on } (0,T) \times \Gamma_2, \\ c(0, \cdot) &= c_0 && \text{in } \Omega. \end{aligned} \tag{5.1}$$

2. *Charge-transport:*

$$\begin{aligned} -\nabla \cdot (\kappa(c) \nabla u) &= 0 && \text{in } (0,T) \times \Omega, \\ \kappa(c_i) \partial_\nu u_i &= i_{12}(c, [u]) && \text{on } (0,T) \times I, \\ \kappa(c) \partial_\nu u &= 0 && \text{on } (0,T) \times \Gamma_0, \\ \kappa(c_1) \partial_\nu u_1 &= -j^{ext} && \text{on } (0,T) \times \Gamma_1, \\ u_2 &= 0 && \text{on } (0,T) \times \Gamma_2. \end{aligned} \tag{5.2}$$

The idea of the existence proof (Theorem 5.6.1) is the following: Starting with a concentration $\widetilde{c}$, we denote by u the solution of the elliptic subproblem at this concentration, that is, u is the solution of (5.1) with c replaced by $\widetilde{c}$. Then we define c as the solution of the linearized and decoupled parabolic subproblem, that is, c is the the solution of (5.2), where the nonlinear term is replaced by $i_{12}(\widetilde{c}, [u])$.

We will then use the results from Chapter 4 and maximal parabolic regularity of the negative laplacian, see Section 5.5, to show that the operator $\boldsymbol{T}$, which is defined on a suitable subset of the space of the continuous functions and maps $\widetilde{c}$ to c, satisfies the assumptions of the Schauder fixed point theorem. If c is a fixed point of $\boldsymbol{T}$, a solution of Problem 3.4.3 will then be given by the couple (c, u), where u again denotes the solution of the elliptic subproblem.

For the proof of the uniqueness results Theorem 5.7.1 we will use the Lipschitz continuity of the elliptic solution operator and Sobolev embeddings to conclude that $\boldsymbol{T}$ is a contraction for $d \leq 3$.

The structure of this chapter is the following: In Section 5.1 and Section 5.2 we introduce the important function spaces and operators. In Section 5.3 we then give a precise weak formulation of Problem 3.4.3. After rephrasing the important properties of the elliptic subproblem in Section 5.4 and the maximal parabolic regularity results in Section 5.5, we will then prove the existence and uniqueness of our weak solutions in Section 5.6 and Section 5.7, respectively.

5.1 Function Spaces

Let $\beta \in [0, 1)$ and X be a Banach space. Recall that we defined $\mathcal{C}^\beta([0, T]; X)$ as the Banach space of X-valued (Hölder) continuous functions defined on $[0, T]$. For a subset $D \subset X$ we will denote by $\mathcal{C}^\beta([0, T]; D)$ those functions in $u \in \mathcal{C}^\beta([0, T]; X)$ satisfying $u(t) \in D$ for all $t \in [0, T]$. We will consider $\mathcal{C}^\beta([0, T]; D)$ as a topological subspace of $\mathcal{C}^\beta([0, T]; X)$, that is, we equip it with the norm $\| \cdot \|_{\mathcal{C}^\beta([0,T];X)}$. Note that $\mathcal{C}^\beta([0, T]; D)$ is closed (in $\mathcal{C}^\beta([0, T]; D)$) if D is closed in X.

Furthermore for $M > 0$ and $i \in \{1, 2\}$ we define the compact sets

$$K_{M,i} := \Big[\frac{1}{M}, \min\Big\{M, c_{\max,i} - \frac{1}{M}\Big\}\Big], \qquad K_M := M_{M,1} \times K_{M,2} \tag{5.3}$$

and the function spaces

$$\begin{aligned} Z_M &:= \{c \in \mathcal{C}^0_{\mathrm{b}} \mid c_i(x) \in K_{M,i} \text{ for all } x \in \overline{\Omega}_i \text{ and } i = 1, 2\}, \\ Z_{M;T} &:= \mathcal{C}^0([0, T]; Z_M). \end{aligned}$$

Note that both Z_M and $Z_{M;T}$ are nonempty, given that M is sufficiently large. We consider Z_M and $Z_{M;T}$ as subsets of $\mathcal{C}^0_{\mathrm{b}}$ and $\mathcal{C}^0([0, T]; \mathcal{C}^0_{\mathrm{b}})$, respectively. By construction, they are both closed and bounded. Finally, we define

$$Z_\infty := \bigcup_{M>0} Z_M \qquad \text{and} \qquad Z_{\infty;T} := \bigcup_{M>0} Z_{M;T}.$$

We also consider Z_∞ and $Z_{\infty;T}$ as subsets of $\mathcal{C}^0_\mathrm{b}$ and $\mathcal{C}^0([0,T];\mathcal{C}^0_\mathrm{b})$, respectively. However, they are neither closed nor bounded in general.

5.2 Operators

Recall that for $s, p \geq 1$ we defined the negative Sobolev spaces $W^{-s,p}$ as the dual space of $W^{s,p'}$, see Section 2.1.1. This definition coincides with the notation used in our references for the maximal parabolic regularity results, see [80, 35, 8]. We note that in the literature it is quite common to define $W^{-s,p}$ as the dual space of $W^{s,p'}_0$ instead, see for example [42, §5.9.1].

Definition 5.2.1.

1. *Define* $\mathcal{A} : H^1 \to H^{-1}$ *by*

$$\langle \mathcal{A}u, v\rangle := \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x$$

for all $u, v \in H^1$.

Then $-\mathcal{A}$ *is the* Laplace-operator with homogeneous Neumann boundary conditions on $\partial\Omega$, *therefore we will simply write* $\Delta := -\mathcal{A}$ *when there is no danger of confusion.*

2. *For* $q \in (d, \infty)$ *define the operator* $\mathcal{A}_q$ *by*

$$\mathrm{Dom}(\mathcal{A}_q) = \{u \in H^1 \mid \mathcal{A}u \in W^{-1,q}\}$$

and

$$\mathcal{A}_q : \mathrm{Dom}(\mathcal{A}_q) \to W^{-1,q},\ \langle \mathcal{A}_q u, v\rangle = \langle \mathcal{A}u, v\rangle$$

for all $u \in \mathrm{Dom}(\mathcal{A}_q)$ *and all* $v \in W^{1,q'}$.

It holds $\mathrm{Dom}(\mathcal{A}_q) \hookrightarrow \mathcal{C}^\alpha_b$ *for some* $\alpha \in (0,1)$ *only depending on the geometry, see Remark 5.2.2 below. Therefore we can and will consider* $\mathcal{A}_q$ *as an unbounded operator on the Banach-Space* $W^{-1,q}$.

3. *Define* $\mathcal{B} : L^\infty \times H^1_{\Gamma_2} \to H^{-1}_{\Gamma_2}$ *by*

$$\langle \mathcal{B}(\kappa, u), v\rangle := \int_\Omega \kappa \nabla u \cdot \nabla v \,\mathrm{d}x$$

for all $\kappa \in L^\infty$ *and* $u, v \in H^1_{\Gamma_2}$.

$\mathcal{B}(\kappa, \cdot)$ *is the second order differential operator in divergence form* $-\nabla \cdot (\kappa \nabla(\cdot))$ *with homogeneous mixed boundary conditions on* $\partial\Omega$. *When there is no risk of confusion, we will simply write* $-\nabla \cdot (\kappa \nabla u) := \mathcal{B}(\kappa, u)$.

Let us collect some important properties of the operator $\mathcal{A}_q$ from the literature.

Remark 5.2.2. *Let $q \in (d, \infty)$. Then the following hold:*

1. *[80, Theorem 1.1] There exists $\beta_0 = \beta_0(q) \in (0,1)$ such that $\mathrm{Dom}(\mathcal{A}_q) \subset \mathcal{C}_\mathrm{b}^{\beta_0}$ and the embedding is continuous.*
2. *[8, Proposition 4.6] $\mathcal{A}_q$ is a densely defined closed operator on $W^{-1,q}$.*

Now we define the operators which realize the nonlinear Neumann boundary and interface conditions.

Definition 5.2.3.

1. *For $q \in (d, \infty)$ define $\mathcal{N}_q : Z_\infty \times \mathcal{C}_\mathrm{b}^0 \to W^{-1,q}$ by*

$$\begin{aligned}\langle \mathcal{N}_q(c,u), \varphi \rangle := & \int_I (1 - t_+) i_{12}(c, [u]) \varphi_1 \,\mathrm{d}\sigma \\ & - \int_I i_{12}(c, [u]) \varphi_2 \,\mathrm{d}\sigma \\ & + \int_{\Gamma_1} (t_+ - 1) j^{ext} \varphi_1 \,\mathrm{d}\sigma\end{aligned}$$

for all $c \in Z_\infty$, $u \in \mathcal{C}_\mathrm{b}^0$ and $\varphi \in W^{1,q'}$.

When there is no possibility for confusion, we simply write $\mathcal{N} := \mathcal{N}_q$.

2. *Define $\mathcal{J} : Z_\infty \times \mathcal{C}_\mathrm{b}^0 \to H_{\Gamma_2}^{-1}$ by*

$$\langle \mathcal{J}(c,u), \varphi \rangle := - \int_I i_{12}(c, [u]) [\varphi] \,\mathrm{d}\sigma - \int_{\Gamma_1} j^{ext} \varphi_1 \,\mathrm{d}\sigma$$

for all $c \in Z_\infty$, $u \in \mathcal{C}_\mathrm{b}^0$ and $\varphi \in H_{\Gamma_2}^1$.

By the definition of $\mathcal{N}_q$ and the fact that i_{12} is $\mathcal{C}^1$, see Assumption 3.5.3, the operator $\mathcal{N}_q$ satisfies the following Lipschitz condition:

Remark 5.2.4. *For $q \in (d, \infty)$, $M, R \in (0, \infty)$ and $r := q(d-1)/d$ there is a positive constant $C = C(q, M, R)$ such that it holds*

$$\|\mathcal{N}_q(c,u) - \mathcal{N}_q(\widetilde{c}, \widetilde{u})\|_{-1,q;\Omega} \le C\big(\|c - \widetilde{c}\|_{0,r;I} + \|u - \widetilde{u}\|_{0,r;I}\big)$$

for all $c, \widetilde{c} \in Z_M$ and $u, \widetilde{u} \in \mathcal{C}_\mathrm{b}^0$ satisfying $\|u\|_{0,\infty;I}, \|\widetilde{u}\|_{0,\infty;I} \le R$.

Proof. Let $c, \widetilde{c} \in Z_M$ and $u, \widetilde{u} \in \mathcal{C}_\mathrm{b}^0$ satisfy $\|u\|_{0,\infty;I}, \|\widetilde{u}\|_{0,\infty;I} \le R$. Denote by $\lesssim$ the relation $\lesssim_{q,M,R}$. Since i_{12} is continuously differentiable, see Assumption 3.5.3, its partial derivatives $\partial_c i_{12} := (\partial_{c_1} i_{12}, \partial_{c_2} i_{12})$ and $\partial_z i_{12}$ are bounded on the compact set $K_{M,R} := K_M \times [-R, R]$.

For all $x \in I$ it holds $(c(x), [u(x)]) \in K_{M,R}$ and $(\widetilde{c}(x), [\widetilde{u}(x)]) \in K_{M,R}$ and the latter set is convex and only depends on M and R. Thus it follows by the mean value theorem:

$$\begin{aligned}
&\Big| i_{12}\big(c(x), [u(x)]\big) - i_{12}\big(\widetilde{c}(x), [\widetilde{u}(x)]\big) \Big| \\
&\le \max_{(\zeta,z)\in K_{M,R}} |\partial_c i_{12}(\zeta, z)|_1 |c(x) - \widetilde{c}(x)|_\infty + \max_{(\zeta,z)\in K_{M,R}} |\partial_z i_{12}(\zeta, z)| \, |[u(x) - \widetilde{u}(x)]| \\
&\lesssim |c(x) - \widetilde{c}(x)|_\infty + |u(x) - \widetilde{u}(x)|_\infty.
\end{aligned}$$

Since $q' = q/(q-1)$ and $r = q(d-1)/d$ it holds

$$1 - \frac{d}{q'} = \frac{q - qd + d}{q} = -\frac{d-1}{r'}.$$

By the trace theorem it follows that the trace operators $W^{1,q'}(\Omega_i) \to L^{r'}(I)$ are bounded. Thus, it follows from the Hölder inequality for $\varphi \in W^{1,q'} \cong W^{1,q'}(\Omega_1) \oplus W^{1,q'}(\Omega_2)$:

$$\begin{aligned}
\big| \langle \mathcal{N}_q(c,u), \varphi\rangle - \langle \mathcal{N}_q(\widetilde{c}, \widetilde{u}), \varphi\rangle \big| &\lesssim \|i_{12}(c, [u]) - i_{12}(\widetilde{c}, [\widetilde{u}])\|_{0,r;I} \|\varphi\|_{0,r';I} \\
&\lesssim \big(\|c - \widetilde{c}\|_{0,r;I} + \|u - \widetilde{u}\|_{0,r;I}\big) \|\varphi\|_{1,q';\Omega}.
\end{aligned}$$

As $\varphi \in W^{1,q'}$ was arbitrary, the proof is finished. □

5.3 Weak Formulation

Having made these definitions, (5.1) and (5.2) can formally be written in the very compact form

$$\begin{aligned}
c' - \Delta c &= \mathcal{N}(c,u), \\
-\nabla \cdot (\kappa(c)\nabla u) &= \mathcal{J}(c,u), \qquad u|_{\Gamma_2} = 0.
\end{aligned}$$

More precisely, we can define the following weak formulation of Problem 3.4.3:

Definition 5.3.1. *Let $q \in (d, \infty)$, $c_0 \in \mathrm{Dom}(\mathcal{A}_q)$ and $T > 0$. A weak solution of Problem 3.4.3 on the time interval $(0,T)$ is a couple (c,u) with the following properties:*

$$c \in H^1\big((0,T); W^{-1,q}\big) \cap L^2((0,T); \mathrm{Dom}(\mathcal{A}_q)) \cap Z_{\infty;T}, \tag{5.4}$$

$$u \in \mathcal{C}^0([0,T]; H^1_{\Gamma_2} \cap \mathcal{C}^0_{\mathrm{b}}), \tag{5.5}$$

and, additionally, $c(0) = c_0$ and

$$c' + \mathcal{A}_q c = \mathcal{N}_q(c,u), \tag{5.6}$$

$$\mathcal{B}(\kappa(c), u) = \mathcal{J}(c,u) \tag{5.7}$$

almost everywhere on $(0,T)$.

Here, $c' \in L^2((0,T); W^{-1,q})$ denotes the distributional derivative of the function $c \in H^1((0,T); W^{-1,q})$. The inital value $c(0)$ is defined in $W^{-1,q}$ by using the vector-valued Sobolev embedding $H^1((0,T); W^{-1,q}) \hookrightarrow \mathcal{C}^0((0,T); W^{-1,q})$. See Chapter 2 and the references which were given there. Since we additionally require $u \in Z_{\infty;T} \subset \mathcal{C}^0([0,T]; \mathcal{C}^0_{\mathrm{b}})$, the inital value also has the classical interpretation $c(0) \in \mathcal{C}^0_{\mathrm{b}} \subset W^{-1,q}$.

Additionally, note that for $t \in (0,T)$, (5.6) is an identity in $W^{-1,q}$, whereas (5.7) is one in $H^{-1}_{\Gamma_2}$.

5.4 Properties of the Elliptic Subproblem

In this section we consider the situation when $c \in Z_{\infty;T}$ is given and study the properties of the problem to determine the unknown potential u satisfying (5.7). We will relate this problem to the equation studied in Chapter 4 and apply the results from that section to the current context.

Definition 5.4.1. *For $c \in Z_\infty$ define $\boldsymbol{\kappa}(c) := \kappa \circ c : \Omega \to \mathbb{R}$,*

$$\boldsymbol{f}(c) : I \times \mathbb{R} \to \mathbb{R}, \quad (x, z) \mapsto i_{12}(c(x), z) - i_{12}(c, 0)$$

and

$$\boldsymbol{G}(c) : W^{1,1} \to \mathbb{R}, \quad v \mapsto -\int_I i_{12}(c, 0)[v] \,\mathrm{d}\sigma - \int_{\Gamma_1} j^{ext} v_1 \,\mathrm{d}\sigma.$$

Then, for given $c \in Z_{\infty;T}$ and fixed $t \in (0, T)$, (5.7) is equivalent to the weak formulation (4.8) of the elliptic subproblem, Problem 4.1.1, with $\boldsymbol{\kappa}(c(t))$, $\boldsymbol{f}(c(t))$ and $\boldsymbol{G}(c(t))$ as data, compare Remark 5.4.11.

We will now verify the conditions which are necessary to apply the results from Chapter 4 and investigate the properties of the solution operator $\boldsymbol{U}^T$ of the elliptic subproblem, that is, the operator which maps a given concentration $c \in Z_{\infty;T}$ to the solution $u \in \mathcal{C}^0([0, T]; H^1_{\Gamma_2})$ of (5.7).

Remark 5.4.2. *For $M > 0$ there exists a positive constant $M_1 = M_1(M)$ and a function $M_2 = M_2(M) : (0, \infty) \to (0, \infty)$ such that for all $c \in Z_M$ the triple $(\boldsymbol{\kappa}(c), \boldsymbol{f}(c), \boldsymbol{G}(c))$ satisfies the conditions in Assumption 4.1.2.*

Proof. Let $M > 0$ and $c \in Z_M$. Denote by $\lesssim$ the relation $\lesssim_M$. By Assumption 3.5.3, $\kappa_i : (0, c_{\max,i}) \to (0, \infty)$ is Lipschitz continuous on the compact set $K_{M,i}$, see (5.3), and thus it attains its minimum and maximum. It follows

$$0 < \min \kappa_i(K_{M,i}) \le \big(\boldsymbol{\kappa}(c)\big)(x) = \kappa_i(c_i(x)) \le \max \kappa_i(K_{M,i})$$

for all $x \in \Omega_i$. Since $K_{M,1}$ and $K_{M,2}$ only depend on M, it follows $\min \kappa_i(K_{M,i}) \gtrsim 1$ and $\max \kappa_i(K_{M,i}) \lesssim 1$ for $i \in \{1, 2\}$.

To verify the condition on $\boldsymbol{G}(c)$, let $v \in W^{1,1}$. Since i_{12} is continuously differentiable on $(0, c_{\max,1}) \times (0, c_{\max,2}) \times \mathbb{R}$ and $j^{\text{ext}} \in L^\infty(\Gamma_1)$, see Assumption 3.5.3, it follows from the boundedness of the trace operators from $W^{1,1}(\Omega_i) \to L^1(I)$:

$$\begin{aligned}
\big|\langle \boldsymbol{G}(c), v\rangle\big| &\le \int_I |i_{12}(c, 0)[v]| \,\mathrm{d}\sigma + \int_{\Gamma_1} |j^{\text{ext}} v_1| \,\mathrm{d}\sigma \\
&\le \max_{\zeta \in K_M} |i_{12}(\zeta, 0)| \int_I |[v]| \,\mathrm{d}\sigma + \|j^{\text{ext}}\|_{0,\infty;\Gamma_1} \int_{\Gamma_1} |v_1| \,\mathrm{d}\sigma \\
&\lesssim \int_I |[v]| \,\mathrm{d}\sigma + \int_{\Gamma_1} |v_1| \,\mathrm{d}\sigma \lesssim \|v\|_{1,1;\Omega}.
\end{aligned}$$

This shows $\boldsymbol{G}(c) \in W^{-1,\infty}$ with $\|\boldsymbol{G}(c)\|_{-1,\infty;\Omega} \lesssim 1$. Finally, let us check the conditions on $f := \boldsymbol{f}(c)$. By definition it holds

$$f(x,z) = i_{12}\big(c(x),z\big) - i_{12}\big(c(x),0\big) \qquad \text{for all } (x,z) \in I \times \mathbb{R}.$$

Since $c \in \mathcal{C}^0_{\mathrm{b}} = \mathcal{C}^0(\overline{\Omega}_1) \times \mathcal{C}^0(\overline{\Omega}_2)$ and i_{12} is continuously differentiable, it follows that $f(\cdot,z)$ is continuous on I and thus measurable for all $z \in \mathbb{R}$. Additionally it follows that $f(x,\cdot)$ is continuously differentiable for all $x \in I$ with the derivative given by

$$\partial_z f(x,z) = \partial_z i_{12}\big(c(x),z\big) \qquad \text{for all } z \in \mathbb{R}.$$

It remains to show the pointwise estimates for $\partial_z f$. For the upper bound (4.7) let $R > 0$. Since i_{12} is continuously differentiable, $\partial_z i_{12}$ is continuous and since $c(x) = (c_1(x), c_2(x)) \in K_M$ for $x \in I$, it follows

$$|\partial_z f(x,z)| = |\partial_z i_{12}(c(x),z)| \le \max\big\{|\partial_z i_{12}(\zeta,\widetilde{z})| \mid \zeta \in K_M, \widetilde{z} \in [-R,R]\big\} =: M_2(R)$$

for all $x \in I$ and $z \in \mathbb{R}$ satisfying $|z| \le R$. Clearly, the function M_2 which maps $R > 0$ to $M_2(R)$ only depends on M. For the lower bound (4.6) note that for $x \in I$ and $z \in \mathbb{R}$ it holds

$$\partial_z f(x,z) = \partial_z i_{12}(c(x),z) \ge \inf\{\partial_z i_{12}(\zeta,\widetilde{z}) \mid \zeta \in K_M, \widetilde{z} \in \mathbb{R})\big) =: C_1.$$

From (3.25) it follows $C_1 > 0$. Since C_1 only depends on M this implies $C_1 \gtrsim 1$.

Summing up, we have shown that the conditions of Assumption 3.5.3 are satisfied for a constant M_1 and the function $M_2 : (0,\infty) \to (0,\infty)$ which both only depend on M but *not on c*. This finishes the proof. □

Remark 5.4.3. *For every $M > 0$ and $R > 0$ the mappings*

$$\boldsymbol{\kappa} : Z_M \to L^\infty(\Omega), \qquad \boldsymbol{f} : Z_M \to L^\infty(I \times (-R,R)), \qquad \boldsymbol{G} : Z_M \to W^{-1,1}(\Omega)$$

are Lipschitz continuous.

Proof. Let $M > 0$ and $c, \widetilde{c} \in Z_M$ and denote by $\lesssim$ the relation $\lesssim_M$. For $i = 1,2$, by the local Lipschitz continuity of $\kappa_i : (0, c_{\max,i}) \to (0,\infty)$, see Assumption 3.5.3, κ_i is Lipschitz continuous on $K_{M,i}$, say, with a Lipschitz constant $L_i = L_i(M)$. It follows:

$$\begin{aligned} \|\boldsymbol{\kappa}(c) - \boldsymbol{\kappa}(\widetilde{c})\|_{0,\infty;\Omega} &= \sup_{i=1,2} \sup_{x \in \Omega_i} |\kappa_i(c_i(x)) - \kappa_i(\widetilde{c}_i(x))| \\ &\le \sup_{i=1,2} \sup_{x \in \Omega_i} L_i |c_i(x) - \widetilde{c}_i(x)| \\ &\lesssim \|c - \widetilde{c}\|_{0,\infty;\Omega}, \end{aligned}$$

which shows that $\kappa : Z_M \to L^\infty$ is Lipschitz continuous.

Since i_{12} is continuously differentiable and K_M is a convex set, it follows from the mean-value theorem:

$$\begin{aligned} |i_{12}(c(x),0) - i_{12}(\widetilde{c}(x),0)| &\le \max_{\xi \in K_M} |\partial_c i_{12}(\xi,0)|_\infty |c(x) - \widetilde{c}(x)|_1 \\ &\lesssim \|c - \widetilde{c}\|_{0,\infty;\Omega}. \end{aligned}$$

By taking into account the boundedness of the trace-operator from $W^{1,1}(\Omega_i)$ to $L^1(I)$, it follows for $v \in W^{1,1} \cong W^{1,1}(\Omega_1) \oplus W^{1,1}(\Omega_2)$:

$$\begin{aligned} \left|\langle \boldsymbol{G}(c) - \boldsymbol{G}(\widetilde{c}), v\rangle\right| &\leq \int_I |i_{12}(c,0) - i_{12}(\widetilde{c},0)|\,|[v]|\,\mathrm{d}\sigma \\ &\lesssim \|c-\widetilde{c}\|_{0,\infty;\Omega} \int_I |[v]|\,\mathrm{d}\sigma \\ &\lesssim \|c-\widetilde{c}\|_{0,\infty;\Omega}\,\|v\|_{1,1;\Omega}, \end{aligned}$$

which shows the Lipschitz continuity of $\boldsymbol{G} : Z_M \to W^{-1,1}$.

For the Lipschitz continuity of $\boldsymbol{f}$, let $R > 0$. As above it holds for $x \in I$ and $z \in \mathbb{R}$ with $|z| \leq R$:

$$\begin{aligned} &|i_{12}(c(x),z) - i_{12}(\widetilde{c}(x),z)| \\ &\leq \max\left\{|\partial_c i_{12}(\zeta,\widetilde{z})|_\infty \mid \zeta \in K_M, \widetilde{z} \in [-R,R]\right\} |c(x) - \widetilde{c}(x)|_1 \\ &\lesssim \|c-\widetilde{c}\|_{0,\infty;\Omega}. \end{aligned}$$

As a consequence, it follows

$$\begin{aligned} &\|\boldsymbol{f}(c) - \boldsymbol{f}(\widetilde{c})\|_{0,\infty;I\times(-R,R)} \\ &\leq \sup_{|z|\leq R} \|i_{12}(c,z) - i_{12}(\widetilde{c},z)\|_{0,\infty;I} + \|i_{12}(c,0) - i_{12}(\widetilde{c},0)\|_{0,\infty;I} \\ &\lesssim \|c-\widetilde{c}\|_{0,\infty;\Omega}. \end{aligned}$$

This shows that $\boldsymbol{f} : Z_M \to L^\infty(I \times (-R,R))$ is Lipschitz continuous and the proof is finished. □

As a consequence of Remark 5.4.2, it follows from Theorem 4.2.7 (existence) and Lemma 4.2.4 (uniqueness) that there exists a unique weak solution

$$\boldsymbol{U}(c) := u \in H^1_{\Gamma_2}(\Omega) \cap L^\infty(\Omega)$$

of Problem 4.1.1 in the sense of Definition 4.2.1 with the data $\boldsymbol{\kappa}(c)$, $\boldsymbol{f}(c)$ and $\boldsymbol{G}(c)$.

Recall, that throughout this chapter it is assumed that Γ_2 and $\partial\Omega_2 \setminus \Gamma_2$ are *well-distributed*, see Assumption 5.0.1. Therefore we can apply the Hölder-regularity result Lemma 4.3.1. Together with the uniform H^1-bound from Theorem 4.2.7 we obtain the following uniform estimates:

Theorem 5.4.4. *For $M > 0$ and $c \in Z_M$ we have $\boldsymbol{U}(c) \in H^1_{\Gamma_2} \cap \mathcal{C}^\delta_{\mathrm{b}}$ and*

$$\|\boldsymbol{U}(c)\|_{1,2;\Omega},\ \|\boldsymbol{U}(c)\|_{\mathcal{C}^\delta_{\mathrm{b}}} \leq C \tag{5.8}$$

for some $\delta = \delta(M) \in (0,1)$ and a positive constant $C = C(M)$ only depending on the constant M but not on the function $c \in Z_M$.

Combining Remark 5.4.3 and Lemma 4.4.1 we obtain the *Lipschitz continuity* of $\boldsymbol{U} : Z_M \to H^1$, see the following lemma:

Lemma 5.4.5. *For $M > 0$ the solution operator $\boldsymbol{U}$ is Lipschitz continuous as an operator*

$$\boldsymbol{U} : Z_M \to H^1.$$

The Lipschitz constant depends on M in general.

From Remark 5.4.3 and Lemma 4.4.2 we can conclude that the solution operator $\boldsymbol{U}$, considered as a nonlinear operator $\boldsymbol{U} : Z_M \to \mathcal{C}_{\mathrm{b}}^{\delta}$ is continuous for some Hölder-exponent $\delta = \delta(M) \in (0,1)$, see the following lemma. However, we do *not obtain the Lipschitz continuity* of $\boldsymbol{U}$ as an operator between these spaces.

Lemma 5.4.6. *For $M > 0$ there exists $\delta = \delta(M) \in (0,1)$ such that the solution operator $\boldsymbol{U}$ is continuous as an operator*

$$\boldsymbol{U} : Z_M \to \mathcal{C}_{\mathrm{b}}^{\delta}.$$

For $T > 0$ we now consider the constant in time extension $\boldsymbol{U}^T$ of the operator $\boldsymbol{U}$ for $T > 0$, that it, $\boldsymbol{U}^T$ is defined by

$$\big(\boldsymbol{U}^T(c)\big)(t) = \boldsymbol{U}(c(t)) \qquad \text{for } c \in Z_{\infty;T} \text{ and } t \in [0,T].$$

We will frequently simply write $\boldsymbol{U}$ instead of $\boldsymbol{U}^T$ whenever it is convenient.

The continuity properties of $\boldsymbol{U}$ carry over to the time-dependent solution operator $\boldsymbol{U}^T$: From Lemma 5.4.5 we can deduce the Lipschitz continuity of $\boldsymbol{U}^T : Z_{M;T} \to \mathcal{C}^0([0,T];H^1)$, see Lemma 5.4.7. Furthermore, from Lemma 5.4.6 we can derive the continuity of $\boldsymbol{U}^T$ as an operator $\boldsymbol{U}^T : Z_{M;T} \to \mathcal{C}^0([0,T];\mathcal{C}_{\mathrm{b}}^{\delta})$ for the $\delta = \delta(M)$ from Lemma 5.4.6, see Lemma 5.4.9.

Lemma 5.4.7. *For $M > 0$ and $T > 0$ the time-dependent solution operator of the elliptic subproblem $\boldsymbol{U}^T$ is* Lipschitz continuous *as an operator*

$$\boldsymbol{U}^T : Z_{M;T} \to \mathcal{C}^0([0,T];H^1).$$

The Lipschitz constant L depends on M but not on T, that is, $L = L(M)$.

Proof. This follows from the Lipschitz continuity of $\boldsymbol{U} : Z_M \to H^1$ (Lemma 5.4.5) and the abstract Lemma 5.4.8. □

Lemma 5.4.8. *Let X and Y be Banach spaces, $D \subset X$ be a closed subset and $A : D \to Y$ a Lipschitz continuous (nonlinear) operator with Lipschitz constant $L > 0$. Then, for $T > 0$ the operator*

$$\begin{aligned} A^T : \mathcal{C}^0([0,T];D) &\to \mathcal{C}^0([0,T];Y), \\ u &\mapsto A(u(\cdot)) \end{aligned}$$

is also Lipschitz continuous with the same Lipschitz constant L.

Proof. First we show that A^T indeed maps $\mathcal{C}^0([0,T];D)$ into $\mathcal{C}^0([0,T];Y)$. To this end, let $u \in \mathcal{C}^0([0,T];D)$ and $(t_n)_{n\in\mathbb{N}} \subset [0,T]$ be a convergent series, say, $t_n \to t$. Since $u \in \mathcal{C}^0([0,T];D) \subset \mathcal{C}^0([0,T];X)$, we have

$$u(t_n) \to u(t) \text{ in } X \qquad \text{as } n \to \infty.$$

Since $A : D \to Y$ is in particular continuous, this implies

$$\big(A^T(u)\big)(t_n) = A(u(t_n)) \to A(u(t)) = \big(A^T(u)\big)(t) \text{ in } Y \qquad \text{as } n \to \infty,$$

which proves $A^T(u) \in \mathcal{C}^0([0,T];Y)$. Furthermore, we obtain for $u, \widetilde{u} \in \mathcal{C}^0([0,T];D)$:

$$\begin{aligned} \|A^T(u) - A^T(\widetilde{u})\|_{\mathcal{C}^0([0,T];Y)} &= \sup_{t\in[0,T]} \|A(u(t)) - A(\widetilde{u}(t))\|_Y \\ &\le L \sup_{t\in[0,T]} \|u(t) - \widetilde{u}(t)\|_Y \\ &= L\|u - \widetilde{u}\|_{\mathcal{C}^0([0,T];Y)}. \end{aligned}$$

This finishes the proof. □

Lemma 5.4.9. *For $M > 0$ and $T > 0$ the time-dependent solution operator of the elliptic subproblem $\boldsymbol{U}^T$ is* continuous *as a mapping*

$$\boldsymbol{U}^T : Z_{M;T} \to \mathcal{C}^0([0,T];\mathcal{C}^\delta_{\mathrm{b}}),$$

where $\delta = \delta(M) \in (0,1)$ is the Hölder-exponent from Theorem 5.4.4

Proof. This follows from the continuity of $\boldsymbol{U} : Z_M \to \mathcal{C}^\delta_{\mathrm{b}}$ (Lemma 5.4.6) and the abstract Lemma 5.4.10. □

Lemma 5.4.10. *Let X and Y be Banach spaces, $D \subset X$ be a closed subset and A be a continuous (nonlinear) operator $A : D \to Y$. Then for $T > 0$, the operator*

$$\begin{aligned} A^T : \mathcal{C}^0([0,T];D) &\to \mathcal{C}^0([0,T];Y), \\ u &\mapsto A(u(\cdot)) \end{aligned}$$

is also continuous.

Proof. As in the proof of Lemma 5.4.8 it holds that A^T maps $\mathcal{C}^0([0,T];D)$ to $\mathcal{C}^0([0,T];Y)$.

Now let us prove the continuity of A^T between the respective spaces. To this end let $(u_n)_{n\in\mathbb{N}} \subset \mathcal{C}^0([0,T];D)$ be a convergent series, say, against $u \in \mathcal{C}^0([0,T];D)$. We need to show:

$$\sup_{t\in[0,T]} \big\|\big(A^T(u_n)\big)(t) - \big(A^T(u)\big)(t)\big\|_Y \to 0 \qquad \text{as } n \to \infty.$$

Let us assume the contrary. Without loss of generality, we find $\varepsilon > 0$ and for *each* $n \in \mathbb{N}$ some $t_n \in [0,T]$, such that it holds

$$\big\|A\big(u_n(t_n)\big) - A\big(u(t_n)\big)\big\|_Y = \big\|\big(A^T(u_n)\big)(t_n) - \big(A^T(u)\big)(t_n)\big\|_Y \ge \varepsilon. \tag{5.9}$$

By compactness of $[0,T]$ we may assume that $(t_n)_{n\in\mathbb{N}}$ converges, say, against $t \in [0,T]$. Since $u \in \mathcal{C}^0([0,T];D)$ and $D \subset X$, clearly,

$$u(t_n) \to u(t) \text{ in } X \qquad \text{as } n \to \infty.$$

On the other hand, we have:

$$\begin{aligned}\|u_n(t_n) - u(t)\|_X &\le \|u_n(t_n) - u(t_n)\|_X + \|u(t_n) - u(t)\|_X \\ &\le \|u_n - u\|_{\mathcal{C}^0([0,T];X)} + \|u(t_n) - u(t)\|_X.\end{aligned}$$

Since $u_n \to u$ in $\mathcal{C}^0([0,T];X)$ and $t_n \to t$ for $n \to \infty$, the term on the right hand side converges to 0 for $n \to \infty$ and thus it holds

$$u_n(t_n) \to u(t) \text{ in } X \qquad \text{as } n \to \infty.$$

By the continuity of $A : D \to Y$, it follows

$$A(u_n(t_n)) - A(u(t_n)) \to A(u(t)) - A(u(t)) = 0 \text{ in } Y \qquad \text{as } n \to \infty.$$

This contradicts (5.9) and, as a consequence, the proof is finished.

□

Let us conclude this section with a remark that sums up the relation between the weak formulation, Definition 5.3.1, of Problem 3.4.3 and the solution operator $\boldsymbol{U}$ of the elliptic subproblem.

Remark 5.4.11. *Let $c \in Z_{\infty;T}$ and $u \in \mathcal{C}^0([0,T];H^1_{\Gamma_2})$ be given. Then the following holds:*

1. *For every $t \in (0,T)$, (5.7) is equivalent to $u(t) = \boldsymbol{U}(c(t))$.*
2. *If (5.7) holds for* almost all $t \in (0,T)$ *it holds for* all $t \in [0,T]$.
3. *(5.7) is equivalent to $u = \boldsymbol{U}(c)$.*

Proof. 1 is immediate from the definition of the operator $\boldsymbol{U}$. 2 follows from 1 and the continuity of $\boldsymbol{U} : Z_M \to H^1$ for $M > 0$. Finally, 3 is obtained from combining 1 and 2. □

5.5 Maximal Parabolic Regularity

In this section we will present a maximal parabolic regularity result found in [35] which is a central ingredient to the proofs of our existence and uniqueness results Theorem 5.6.1 and Theorem 5.7.1, respectively.

For the sake of a clearer presentation, we will restrict ourselves to *real Banach spaces*, whereas in the articles addressing maximal parabolic regularity, like [35, 80, 8], it is often only considered the case of *complex Banach spaces*. However, since the data in (3.4.3) is purely real-valued, we can still apply the respective results to our problem. Note that the concept of maximal parabolic regularity is not restricted to complex Banach spaces, see for example [4, Chapter III].

Let us start by rephrasing the concept of maximal parabolic regularity.

Definition 5.5.1. *[35, Definition 3.2] Let X be a Banach space, $T > 0$ and $s \in (1, \infty)$. Assume that*

$$B : X \supset \mathrm{Dom}(B) \to X$$

is a densely defined closed operator on X. Then B admits maximal parabolic $L^s((0,T);X)$-regularity *if there is an isomorphism of Banach spaces which maps every $f \in L^s((0,T);X)$ to the unique function*

$$u \in W^{1,s}\big((0,T);X\big) \cap L^s\big((0,T);\mathrm{Dom}(B)\big)$$

such that it holds $u(0) = 0$ and

$$u' + Bu = f \qquad \text{almost everywhere on } (0,T). \tag{5.10}$$

Note that (5.10) is an equality in the space X. Furthermore, $\mathrm{Dom}(B)$ is endowed with the graph norm $\|\cdot\|_{\mathrm{Dom}(B)}$, that is,

$$\|u\|_{\mathrm{Dom}(B)} = \|u\|_X + \|Bu\|_X \qquad \text{for } u \in \mathrm{Dom}(B).$$

Let us again emphasize that $u' \in L^s((0,T);X)$ denotes the weak derivative of $u \in W^{1,s}((0;T);X)$ and that the point evaluation is verified by the vector valued Sobolev embedding $W^{1,s}((0,T);X) \hookrightarrow \mathcal{C}^0([0,T];X)$, see Chapter 2.

If X, Y are Banach spaces forming an interpolation couple (X,Y), we denote by $(X,Y)_{\theta,q}$ the corresponding interpolation space by the real interpolation method with interpolation parameters $\theta \in [0,1]$ and $q \in [1,\infty]$, see [81].

Remark 5.5.2.

1. *The property of B to admit maximal parabolic $L^s\big((0,T);X\big)$-regularity is independent of $s \in (1,\infty)$ and $T > 0$.*

2. *Let $s \in (1,\infty)$ and $T > 0$ and let B admit maximal parabolic $L^s\big((0,T);X\big)$-regularity. Then the following holds:*

   There is a constant $C = C(s,T) > 0$ such that for all $f \in L^s((0,T);X)$ and $u_0 \in (X, \mathrm{Dom}(B))_{1-1/s,s}$ there exists a unique function

   $$u \in W^{1,s}((0,T);X) \cap L^s((0,T);\mathrm{Dom}(B))$$

   such that it holds $u(0) = u_0$ and

   $$u' + Bu = f \qquad \text{almost everywhere on } (0,T).$$

   Addtionally, it holds

   $$\begin{aligned} &\|u\|_{W^{1,s}((0,T);X)} + \|u\|_{L^s((0,T);\mathrm{Dom}(B))} \\ &\le C\big(\|f\|_{L^s((0,T);X)} + \|u_0\|_{(X,\mathrm{Dom}(B))_{1-1/s,s}}\big). \end{aligned} \tag{5.11}$$

3. *The constant $C = C(s,T)$ in (5.11) can be chosen to grow monotonically in $T > 0$ for fixed $s \in (1,\infty)$.*

Proof. 1 is a consequence of [36, Theorem 2.5, Theorem 4.2]. 2 is included in [5, Proposition 2.1]. In order to prove the third part, let $0 < \widetilde{T} < T$ and $\widetilde{f} \in L^s((0,\widetilde{T});X)$. Define $f := \chi_{(0,\widetilde{T})}\widetilde{f} \in L^s((0,T);X)$. By the maximal parabolic $L^s((0,T);X)$-regularity of B, we can choose u as the unique solution of

$$\begin{aligned} u(0) &= u_0, \\ u' + Bu &= f \qquad \text{almost everywhere on } (0,T). \end{aligned}$$

Then $\widetilde{u} := u|_{(0,\widetilde{T})}$ is the unique solution of

$$\begin{aligned} \widetilde{u}(0) &= u_0 \\ \widetilde{u}' + B\widetilde{u} &= \widetilde{f} \qquad \text{almost everywhere on } (0,\widetilde{T}). \end{aligned}$$

Moreover, by (5.11) we have for example

$$\begin{aligned} \|\widetilde{u}\|_{W^{1,s}((0,\widetilde{T});X)} &\le \|u\|_{W^{1,s}((0,T);X)} \\ &\le C(s,T)\big(\|f\|_{L^s((0,T);X)} + \|u_0\|_{(X,\mathrm{Dom}(B))_{1-1/s,s}}\big) \\ &= C(s,T)\big(\|\widetilde{f}\|_{L^s((0,\widetilde{T});X)} + \|u_0\|_{(X,\mathrm{Dom}(B))_{1-1/s,s}}\big). \end{aligned}$$

This shows that $C(s,T)$ is a possible choice for $C(s,\widetilde{T})$. □

Now we are in the position to state the maximal parabolic regularity result for distributional right-hand sides in the space $L^s((0,T);W^{-1,q})$ which we use in the existence and uniqueness proofs in Section 5.6 and Section 5.7, respectively.

It is an immediate consequence of the results in [35] and [8].

Lemma 5.5.3. *For all $q \in (d,\infty)$ there exists $s \in (1,\infty)$ and $\beta \in (0,1)$ such that for all $T > 0$ there is a positive constant $C = C(q,T)$ with the following property:*

For all

$$f \in L^s\big((0,T);W^{-1,q}\big) \qquad \text{and} \qquad u_0 \in \big(W^{-1,q},\mathrm{Dom}(\mathcal{A}_q)\big)_{1-1/s,s}$$

there exists a unique

$$u \in W^{1,s}\big((0,T);W^{-1,q}\big) \cap L^s\big((0,T);\mathrm{Dom}(\mathcal{A}_q)\big)$$

such that it holds $u(0) = u_0$ and

$$u' + \mathcal{A}_q u = f \qquad \text{almost everywhere on } (0,T).$$

Additionally, it holds $u \in \mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})$ and

$$\begin{aligned} &\|u\|_{W^{1,s}((0,T);W^{-1,q})} + \|u\|_{L^s((0,T),\,\mathrm{Dom}(\mathcal{A}_q))} + \|u\|_{\mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})} \\ &\qquad \le C\big(\|f\|_{L^s((0,T);W^{-1,q})} + \|u_0\|_{(W^{-1,q},\mathrm{Dom}(\mathcal{A}_q))_{1-1/s,s}}\big). \end{aligned} \tag{5.12}$$

Proof. Let $q \in (d, \infty)$ and, according to the proof of [35, Theorem 4.5], choose $\beta \in (0,1)$ and $s \in (1, \infty)$ such that the embedding

$$W^{1,s}((0,T); W^{-1,q}) \cap L^s((0,T); \mathrm{Dom}(\mathcal{A}_q)) \hookrightarrow \mathcal{C}^\beta([0,T]; \mathcal{C}^\beta_{\mathrm{b}}) \tag{5.13}$$

is bounded for all $T > 0$. The existence of a unique

$$u \in W^{1,s}\big((0,T); W^{-1,q}\big) \cap L^s\big((0,T); \mathrm{Dom}(\mathcal{A}_q)\big)$$

satisfying $u(0) = u_0$ and

$$u' + \mathcal{A}_q u = f \qquad \text{almost everywhere on } (0,T)$$

and the estimates in $W^{1,s}((0,T); W^{-1,q})$ and $L^s((0,T), \mathrm{Dom}(\mathcal{A}_q))$ follow from the fact that $\mathcal{A}_q$ admits maximal parabolic $L^s((0,T); W^{-1,q})$-regularity, see Remark 5.5.4. Finally, the Hölder-regularity of u follows from the boundedness of the embedding (5.13). □

Remark 5.5.4. *In the situation of Lemma 5.5.3, the following assertions hold:*

1. *$\mathcal{A}_q$ admits maximal parabolic regularity on $W^{-1,q}$.*
2. *The constant $C(q,T)$ can be chosen to grow monotonically in T for fixed $q \in (d, \infty)$.*

Proof.
1. This follows from [8, Theorem 11.5].
2. Let $0 < \widetilde{T} < T$ and $\widetilde{f} \in L^s((0,\widetilde{T}); X)$. Define $u, \widetilde{u}$ and f as in the proof of Remark 5.5.2. It follows with (5.12):

$$\begin{aligned}\|\widetilde{u}\|_{\mathcal{C}^\beta((0,\widetilde{T});\mathcal{C}^\beta_{\mathrm{b}})} &\le \|u\|_{\mathcal{C}^\beta((0,T);\mathcal{C}^\beta_{\mathrm{b}})} \\ &\le C(q,T)\big(\|f\|_{L^s((0,T);W^{-1,q})} + \|u_0\|_{(W^{-1,q},\mathrm{Dom}(B))_{1-1/s,s}}\big) \\ &= C(q,T)\big(\|\widetilde{f}\|_{L^s((0,\widetilde{T});W^{-1,q})} + \|u_0\|_{(W^{-1,q},\mathrm{Dom}(B))_{1-1/s,s}}\big).\end{aligned}$$

   As a consequence, $C(q,T)$ is a possible choice for $C(q,\widetilde{T})$.

□

5.6 Local Existence

In this section we prove the main result of the thesis, the local in time existence of weak solutions to Problem 3.4.3. This is the precise formulation:

Theorem 5.6.1. *For all $q \in (d, \infty)$ and $c_0 \in \mathrm{Dom}(\mathcal{A}_q) \cap Z_\infty$ there exists $T > 0$ such that Problem 3.4.3 has a weak solution (c,u) on the time-interval $(0,T)$ in the sense of Definition 5.3.1. Additionally, it holds*

$$c \in \mathcal{C}^\beta([0,T]; \mathcal{C}^\beta_{\mathrm{b}}) \qquad \text{and} \qquad u \in \mathcal{C}^\beta([0,T]; H^1) \cap \mathcal{C}^0([0,T]; \mathcal{C}^\beta)$$

for some $\beta = \beta(q, c_0) \in (0,1)$.

Before proving this theorem, we give the following characterization of weak solutions of Problem 3.4.3.

Remark 5.6.2. *Let* $q \in (d, \infty)$*,* $c_0 \in \mathrm{Dom}(\mathcal{A}_q) \cap Z_\infty$ *and* $T > 0$*. Then for every* $s \in (1, \infty)$ *a couple* (c, u) *is a weak solution of Problem 3.4.3 on the time interval* $(0, T)$ *in the sense of Definition 5.3.1 if and only if it holds*

$$c \in W^{1,s}\big((0,T); W^{-1,q}\big) \cap L^s((0,T); \mathrm{Dom}(\mathcal{A}_q)) \cap Z_{\infty;T}, \tag{5.14}$$

$$u \in \mathcal{C}^0([0,T]; H^1_{\Gamma_2} \cap \mathcal{C}^0_{\mathrm{b}}), \tag{5.5}$$

and, additionally, $c(0) = c_0$ *and*

$$c' + \mathcal{A}_q c = \mathcal{N}_q(c, u), \tag{5.6}$$

$$\mathcal{B}(\kappa(c), u) = \mathcal{J}(c, u) \tag{5.7}$$

almost everywhere on $(0, T)$*.*

Note that the only difference to Definition 5.3.1 is (5.14).

Proof of Remark 5.6.2. Let $s \in (1, \infty)$ and (c, u) satisfy (5.14), (5.5), $c(0) = c_0$, and (5.6), (5.7) almost everywhere on $(0, T)$. Since $c \in Z_{M;T}$ for some $M > 0$ and $u \in \mathcal{C}^0([0,T]; \mathcal{C}^0_{\mathrm{b}})$, it follows from Remark 3.5.5:

$$\mathcal{N}_q(c, u) \in L^\infty((0,T); W^{-1,q}).$$

By the maximal parabolic regularity of $\mathcal{A}_q$, see Remark 5.5.4 and Remark 5.5.2, there exists a unique

$$\widetilde{c} \in H^1((0,T); W^{-1,q}) \cap L^2((0,T); \mathrm{Dom}(\mathcal{A}_q)) \tag{5.15}$$

satisfying

$$\begin{aligned} \widetilde{c}(0) &= c_0, \\ \widetilde{c}' + \mathcal{A}_q \widetilde{c} &= \mathcal{N}_q(c, u) \qquad \text{almost everywhere on } (0,T). \end{aligned} \tag{5.16}$$

As a consequence, putting $r := \min\{2, s\}$, the difference $c - \widetilde{c}$ satisfies

$$c - \widetilde{c} \in W^{1,r}((0,T); W^{-1,q}) \cap L^r((0,T); \mathrm{Dom}(\mathcal{A}_q))$$

and, additionally,

$$\begin{aligned} (c - \widetilde{c})(0) &= 0 \\ (c - \widetilde{c})' + \mathcal{A}_q(c - \widetilde{c}) &= 0 \qquad \text{almost everywhere on } (0,T). \end{aligned}$$

By the maximal parabolic regularity of $\mathcal{A}_q$, this implies $c(t) = \widetilde{c}(t)$ in $W^{-1,q}$ for almost all $t \in (0, T)$. As a consequence, (5.4) follows from (5.15). We have thus shown that (c, u) is a weak solution of Problem 3.4.3 in the sense of Definition 5.3.1.

The reverse implication can be proved analogously. □

Proof of Theorem 5.6.1. Let $q \in (d, \infty)$ and $c_0 \in \mathrm{Dom}(\mathcal{A}_q) \cap Z_\infty$. Then we find some $N > 0$ such that $c_0 \in Z_N$. According to Remark 5.2.2, we choose a Hölder-exponent $\beta_0 \in (0,1)$ such that the embedding $\mathrm{Dom}(\mathcal{A}_q) \hookrightarrow \mathcal{C}_{\mathrm{b}}^{\beta_0}$ is continuous. Additionally, we define

$$M := 2 \max \left\{ N,\ \|c_0\|_{\mathrm{Dom}(\mathcal{A}_q)} \right\}. \tag{5.17}$$

Suppose $T \in (0,1]$ is arbitrary for now. Its value will be chosen later in the proof. Let us write $\lesssim$ for the relation $\lesssim_{q,M}$. Recall that $\boldsymbol{U}$ denotes the solution operator to the elliptic subproblem, see Section 5.4. By Lemma 5.4.9 and Theorem 5.4.4 it is a continuous nonlinear operator

$$\boldsymbol{U} : Z_{M;T} \to \mathcal{C}^0([0,T]; \mathcal{C}_{\mathrm{b}}^0)$$

and it holds

$$\|\boldsymbol{U}(c)\|_{\mathcal{C}^0([0,T];\mathcal{C}_{\mathrm{b}}^0)} \lesssim 1 \qquad \text{for } c \in Z_{M;T}. \tag{5.18}$$

Note that we are now simply using the symbol $\boldsymbol{U}$ for $\boldsymbol{U}^T$ as we have already suggested in Section 5.4. Recall the operator $\mathcal{N}_q$ introduced in Definition 5.2.3. By Remark 5.2.4 and Lemma 5.4.8,

$$\mathcal{N}_q : Z_{M;T} \times \mathcal{C}^0([0,T]; \mathcal{C}_{\mathrm{b}}^0) \to \mathcal{C}^0([0,T]; W^{-1,q}),$$

is locally Lipschitz continuous and for all $R > 0$ there exists a positive constant $C_1 = C_1(q, M, R)$ such that it holds

$$\|\mathcal{N}_q(c,u)\|_{\mathcal{C}^0([0,T];W^{-1,q})} \le C_1 \left(\|u\|_{\mathcal{C}^0([0,T];\mathcal{C}_{\mathrm{b}}^0)} + 1 \right) \tag{5.19}$$

for all $c \in Z_{M,0,T}$ and $u \in \mathcal{C}^0([0,T]; \mathcal{C}_{\mathrm{b}}^0)$ satisfying $\|u\|_{\mathcal{C}^0([0,T];\mathcal{C}_{\mathrm{b}}^0)} \le R$.

Now denote by $s \in (1, \infty)$ and $\beta_1 \in (0,1)$ constants provided by Lemma 5.5.3. Thus for every $f \in L^\infty((0,T); W^{-1,q})$ there exists a unique

$$\boldsymbol{P} f := c \in W^{1,s}\big((0,T); W^{-1,q}\big) \cap L^s\big((0,T); \mathrm{Dom}(\mathcal{A}_q)\big)$$

satisfying

$$\begin{aligned} c(0) &= c_0, \\ c' + \mathcal{A}_q c &= f \qquad \text{almost everywhere on } (0,T). \end{aligned} \tag{5.20}$$

Moreover, it holds $c \in \mathcal{C}^{\beta_1}\big([0,T]; \mathcal{C}_{\mathrm{b}}^{\beta_1}\big)$ and, taking into account the second statement of Remark 5.5.4,

$$\|c\|_{\mathcal{C}^{\beta_1}([0,T];\mathcal{C}_{\mathrm{b}}^{\beta_1})} \lesssim \big(\|f\|_{L^s((0,T);W^{-1,q})} + \|c_0\|_{(W^{-1,q},\mathrm{Dom}(\mathcal{A}_q))_{1-1/s,s}} \big). \tag{5.21}$$

Since $c_0 \in \mathrm{Dom}(\mathcal{A}_q)$, the function $\bar{c}_0$ defined by

$$\bar{c}_0(t) := c_0 \qquad \text{for } t \in [0,T]$$

is an element of $W^{1,s}\big((0,T);W^{-1,q}\big) \cap L^s\big((0,T);\mathrm{Dom}(\mathcal{A}_q)\big)$ and it holds

$$\begin{aligned}(c-\bar{c}_0)(0) &= 0,\\ (c-\bar{c}_0)' + \mathcal{A}_q(c-\bar{c}_0) &= f - \mathcal{A}_q\bar{c}_0 \qquad \text{almost everywhere on } (0,T).\end{aligned}$$

As a consequence, it follows from (5.12), the Hölder inequality and the definition of M:

$$\begin{aligned}\|c-\bar{c}_0\|_{\mathcal{C}^{\beta_1}([0,T];\mathcal{C}_\mathrm{b}^{\beta_1})} &\lesssim \|f-\mathcal{A}_q\bar{c}_0\|_{L^s((0,T);W^{-1,q})}\\ &\le \|f\|_{L^s((0,T);W^{-1,q})} + \|\mathcal{A}_q\bar{c}_0\|_{L^s((0,T);W^{-1,q})}\\ &\lesssim T^{1/s}(\|f\|_{L^\infty((0,T);W^{-1,q})} + \|c_0\|_{\mathrm{Dom}(\mathcal{A}_q)})\\ &\lesssim T^{1/s}(\|f\|_{L^\infty((0,T);W^{-1,q})} + 1).\end{aligned}$$

We have thus shown that the parabolic solution operator

$$\boldsymbol{P} : L^\infty((0,T);W^{-1,q}) \to \mathcal{C}^{\beta_1}([0,T];\mathcal{C}_\mathrm{b}^{\beta_1})$$

is continuous and satisfies

$$\|\boldsymbol{P}f-\bar{c}_0\|_{\mathcal{C}^{\beta_1}([0,T];\mathcal{C}_\mathrm{b}^{\beta_1})} \lesssim T^{1/s}(\|f\|_{L^\infty((0,T);W^{-1,q})}+1) \tag{5.22}$$

for all $f \in L^\infty((0,T);W^{-1,q})$. Now define $\beta := \min\{\beta_1,\beta_0\}$ and

$$\boldsymbol{T}(c) := \boldsymbol{P}\Big(\mathcal{N}_q\big(c,\boldsymbol{U}(c)\big)\Big) \qquad \text{for } c \in Z_{M;T}. \tag{5.23}$$

By the above considerations, (5.23) defines a continuous nonlinear operator from $Z_{M;T}$ to $\mathcal{C}^\beta([0,T];\mathcal{C}_\mathrm{b}^\beta)$. Therefore, we can and will consider $\boldsymbol{T}$ as a continuous nonlinear operator

$$\boldsymbol{T} : Z_{M;T} \to \mathcal{C}^0([0,T];\mathcal{C}_\mathrm{b}).$$

In order to apply the Schauder fixed point theorem on $\boldsymbol{T}$, we will now construct some $0 < T \le 1$ such that the image of $\boldsymbol{T}$ is again contained in $Z_{M;T}$.

To this end let $c \in Z_{M;T}$. Theorem 5.4.4 implies that $R := \|\boldsymbol{U}(c)\|_{\mathcal{C}^0([0,T];\mathcal{C}_\mathrm{b}^0)} \lesssim 1$ and thus the constant $C_1 = C_1(q,M,R)$ in (5.19) satisfies $C_1(q,M,R) \lesssim 1$. Then it follows from (5.22), (5.19) and (5.18):

$$\begin{aligned}\Big\|\boldsymbol{P}\Big(\mathcal{N}_q\big(c,\boldsymbol{U}(c)\big)\Big)-\bar{c}_0\Big\|_{\mathcal{C}^\beta([0,T];\mathcal{C}_\mathrm{b}^\beta)} &\lesssim T^{1/s}\Big(\big\|\mathcal{N}_q\big(c,\boldsymbol{U}(c)\big)\big\|_{L^\infty((0,T);W^{-1,q})}+1\Big)\\ &\lesssim T^{1/s}\Big(\big\|\boldsymbol{U}(c)\big\|_{\mathcal{C}^0([0,T];\mathcal{C}_\mathrm{b}^0)}+1\Big)\\ &\lesssim T^{1/s}.\end{aligned} \tag{5.24}$$

By the boundedness of the embeddings $\mathrm{Dom}(\mathcal{A}_q) \hookrightarrow \mathcal{C}_\mathrm{b}^{\beta_0} \hookrightarrow \mathcal{C}_\mathrm{b}^\beta$ and the definition (5.17) of M, this implies

$$\|\boldsymbol{T}(c)\|_{\mathcal{C}^\beta([0,T];\mathcal{C}_\mathrm{b}^\beta)} \lesssim T^{1/s} + \|c_0\|_{\mathcal{C}^\beta([0,T];\mathcal{C}_\mathrm{b}^\beta)} \lesssim 1. \tag{5.25}$$

On the other hand, (5.24) shows that there is a positive constant $C_2 = C_2(M, q)$ *which in particular does not depend on* T such that it holds

$$\|\boldsymbol{T}(c) - \bar{c}_0\|_{\mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})} \leq C_2 T^{1/s}. \tag{5.26}$$

Now we choose $0 < T \leq 1$ satisfying

$$T^{1/s} \leq \frac{\min\{\frac{M}{2}, \frac{1}{M}\}}{C_2}. \tag{5.27}$$

It follows by (5.26), the definition (5.17)of M and the choice (5.27) of T :

$$\begin{aligned}
\|\boldsymbol{T}(c)\|_{\mathcal{C}^0([0,T];\mathcal{C}^0_{\mathrm{b}})} &\leq \|\boldsymbol{T}(c) - \bar{c}_0\|_{\mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})} + \|c_0\|_{\mathcal{C}^0_{\mathrm{b}}} \\
&\leq C_2 T^{1/s} + \frac{M}{2} \\
&\leq C_2 \frac{M}{2C_2} + \frac{M}{2} = M.
\end{aligned}$$

Using the same estimates, we obtain on $(0, T) \times \Omega$:

$$\begin{aligned}
\boldsymbol{T}(c) &\geq -\|\boldsymbol{T}(c) - \bar{c}_0\|_{\mathcal{C}^0([0;T];\mathcal{C}^0_{\mathrm{b}})} + \bar{c}_0 \\
&\geq -\|\boldsymbol{T}(c) - \bar{c}_0\|_{\mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})} + \bar{c}_0 \\
&\geq -C_2 T^{1/s} + \frac{2}{M} \\
&\geq -C_2 \frac{1}{MC_2} + \frac{2}{M} = \frac{1}{M}.
\end{aligned}$$

The other pointwise estimate $\boldsymbol{T}(c) \leq c_{\max} - 1/M$ can be shown similarly and thus we obtain $\boldsymbol{T}(c) \in Z_{M;T}$. Since $c \in Z_{M;T}$ was arbitrary, we have therefore shown $\boldsymbol{T}(Z_{M;T}) \subset Z_{M;T}$.

Additionally, from (5.25) it follows that the image of $\boldsymbol{T}$ is bounded in $\mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})$, which itself is compactly embedded into the underlying space $\mathcal{C}^0([0,T];\mathcal{C}^0_{\mathrm{b}})$. This shows that the image $\boldsymbol{T}$ is *precompact* in $\mathcal{C}^0([0,T];\mathcal{C}^0_{\mathrm{b}})$.

Since $Z_{M;T}$ is a *nonempty, closed* and *convex* subset of $\mathcal{C}^0([0,T];\mathcal{C}^0_{\mathrm{b}})$, the Schauder fixed point theorem, [44, Theorem 11.1], can be applied and we obtain the existence of a fixed point $c_* \in Z_{M;T} \cap \mathcal{C}^\beta([0,T];\mathcal{C}^\beta_{\mathrm{b}})$ of the operator $\boldsymbol{T}$. Bearing in mind Remark 5.4.11 and Remark 5.6.2, by construction the couple $(c_*, \boldsymbol{U}(c_*))$ is a weak solution of Problem 3.4.3 in the sense of Definition 5.3.1.

The Hölder regularity of $\boldsymbol{U}(c_*)$ follows from Lemma 5.4.5 and Lemma 5.4.6. This finishes the proof. □

5.7 A Uniqueness Result for $d \leq 3$

Let $d \leq 3$. Then we can use the Lipschitz continuity of the solution operator $\boldsymbol{U}$ of the elliptic subproblem, Lemma 5.4.5, together with Sobolev embeddings to conclude that the operator $\boldsymbol{T}$ defined in the proof of Theorem 5.6.1 is a contraction, given that the final time T is small enough. This implies that there exists at most one weak solution of Problem 3.4.3. The details of this argument are given in Theorem 5.7.1 and its proof, respectively.

Note that these considerations imply that in this case we can alternatively prove existence of a solution using the Banach fixed point theorem instead of the Schauder fixed point theorem.

Theorem 5.7.1. *Let $d \leq 3$ and $c_0 \in \mathrm{Dom}(\mathcal{A}_q) \cap Z_\infty$ for some $q \in (d, \infty)$. Then for every $T > 0$ there is at most one weak solution (c, u) of Problem 3.4.3 on the time-interval $(0, T)$ in the sense of Definition 5.3.1.*

Proof. Let (c, u) and $(\widetilde{c}, \widetilde{u})$ be weak solutions of Problem 3.4.3 on $(0, T)$. From Remark 5.4.11 it follows that $u = \boldsymbol{U}(c)$ and $\widetilde{u} = \boldsymbol{U}(\widetilde{c})$. Thus it remains to show that $c = \widetilde{c}$. We find $M > 0$ such that it holds $c, \widetilde{c} \in Z_{M;T}$. Now define

$$t_0 := \inf\{t \in (0, T] \mid c(t) \neq \widetilde{c}(t)\}. \tag{5.28}$$

Let us first consider the case $t_0 = 0$. Since $W^{-1,q} \hookrightarrow W^{-1,\widetilde{q}}$ holds for $d < \widetilde{q} \leq q$ by Hölder's inequality, without loss of generality, we can assume $q \leq 6$. Now choose $s = s(q) \in (1, \infty)$ as in Lemma 5.5.3. From Remark 5.6.2 it follows, that the difference $\bar{c} := c - \widetilde{c}$ satisfies

$$\bar{c} \in W^{1,s}((0, T); W^{-1,q}) \cap L^s((0, T); \mathrm{Dom}(\mathcal{A}_q)) \cap Z_{M;T}$$

and, additionally,

$$\begin{aligned} &\bar{c}(0) = 0, \\ &\bar{c}' + \mathcal{A}_q \bar{c} = \mathcal{N}_q(c, \boldsymbol{U}(c)) - \mathcal{N}_q(\widetilde{c}, \boldsymbol{U}(\widetilde{c})) \qquad \text{almost everywhere on } (0, T). \end{aligned}$$

Let $S \in (0, T)$ be arbitrary for now. Then it follows from (5.12):

$$\|\bar{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})} \lesssim S^{1/s} \|\mathcal{N}(c, \boldsymbol{U}(c)) - \mathcal{N}(\widetilde{c}, \boldsymbol{U}(\widetilde{c}))\|_{L^\infty((0,S);W^{-1,q})},$$

We will use the symbol $\lesssim$ for the relation $\lesssim_{\neg S}$. Since $c, \widetilde{c} \in Z_{M;T}$, Theorem 5.4.4 implies that $|\boldsymbol{U}(c)|, |\boldsymbol{U}(\widetilde{c})| \lesssim 1$ holds in the pointwise sense in $(0, T) \times \Omega$. Using the Lipschitz continuity of $\mathcal{N}_q$, Remark 5.2.4, then gives:

$$\begin{aligned} \|\bar{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})} &\lesssim S^{1/s} \|\mathcal{N}(c, \boldsymbol{U}(c)) - \mathcal{N}(\widetilde{c}, \boldsymbol{U}(\widetilde{c}))\|_{L^\infty((0,S);W^{-1,q})} \\ &\lesssim S^{1/s} \big(\|c - \widetilde{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})} + \|\boldsymbol{U}(c) - \boldsymbol{U}(\widetilde{c})\|_{L^\infty((0,S);L^r(I))} \big) \end{aligned} \tag{5.29}$$

for $r := q(d-1)/d$.

Clearly, for $d = 2$, the trace operators $H^1(\Omega_i) \to L^r(I)$, $i = 1, 2$, are bounded. For $d = 3$, the trace operator $H^1(\Omega) \to L^4(I)$ is bounded. However, since $q \leq 6$ by definition, it holds $r = 2q/3 \leq 4$. The Hölder inequality therefore implies the continuity of the embedding $L^4(I) \to L^r(I)$ and thus the continuity of the trace operators $H^1(\Omega_i) \to L^r(I)$ for $i = 1, 2$.

As a consequence, it follows from the Lipschitz-continuity of $\boldsymbol{U} : Z_M \to H^1$, see Lemma 5.4.5:

$$\begin{aligned} \|\boldsymbol{U}(c) - \boldsymbol{U}(\widetilde{c})\|_{L^\infty((0,S);L^r(I))} &\lesssim \|\boldsymbol{U}(c) - \boldsymbol{U}(\widetilde{c})\|_{L^\infty((0,S);H^1)} \\ &\lesssim \|c - \widetilde{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})}. \end{aligned} \tag{5.30}$$

Combining (5.29) and (5.30) thus yields

$$\|\bar{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})} \leq C_1 S^{1/s} \|\bar{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})}$$

with a positive constant C_1 which does not depend on S. As a consequence, for $S := (2C)^{-s} > 0$ it follows that

$$\|c - \widetilde{c}\|_{\mathcal{C}^0([0,S];\mathcal{C}^0_{\mathrm{b}})} = 0.$$

This implies, however, $t_0 \geq S > 0$, see (5.28), which contradicts our assumption $t_0 = 0$.

The case $t_0 \in (0, T)$ can be reduced to the case $t_0 = 0$ by the transformation $t \mapsto t - t_0$ in the time-variable. As a consequence it follows $t_0 = T$ and thus $c = \widetilde{c}$ on $[0, T]$, which finishes the proof.

□

6 Discretization of a Strongly Nonlinear Elliptic Problem

Throughout this section we will postulate that κ, f and G are given such that Assumption 4.1.2 is satisfied for some positive constant M_1 and a function $M_2 : (0, \infty) \to (0, \infty)$ which will be fixed throughout the whole chapter.

We investigate the convergence of two possible discretizations of the strongly nonlinear elliptic Problem 4.1.1.

In Section 6.1 we first consider the Galerkin approximations.

As it turns out, well-posedness is immediately obtained by the Brouwer fixed point theorem. However, due to the lack of a suitable polynomial growth condition for the nonlinearity f, the quasi optimality of the discrete solutions cannot be obtained by the standard proof of Céa's lemma [41, Lemma 2.28]. However, by reviewing the proof of the Strang lemma ([41, Lemma 2.25]) we are still able to relate the H^1-error to the best approximation in the discrete subspace. Combining this result with the L^∞-stability of the Clément interpolation error we finally establish convergence at the optimal rate under additional regularity assumptions on the exact solution.

The error analysis for the Galerkin approximation would have been much simpler if it was possible to prove a uniform L^∞-bound for the discrete solutions as well. However, we could not apply the Stampacchia truncation method used for the continuous case to the discrete system, see the proof of Theorem 4.2.7. In order to overcome this issue, in Section 6.2 we present a modified discretization. The basic idea is to restrict to linear finite elements and, in addition, apply the trapeziodal rule to the nonlinear interface term. As it turns out, these modified solutions still converge at the optimal linear rate when the exact solution is in $H^2 \cap W^{1,4}$ on a shape-regular family of triangulations. Additionally, it is possible to use the ideas from [20] to generalize the proof for the continuous comparison principle and, more importantly, the L^∞-bound, to the modified discrete system.

In this whole chapter we will denote by $u \in H^1_{\Gamma_2} \cap L^\infty$ the unique weak solution to Problem 4.1.1, see Theorem 4.2.7. Furthermore, we define $V := H^1_{\Gamma_2}$. Finally, recall that we omit the dependence on the geometry which is defined by the objects from Assumption 3.5.1.

6.1 Standard Galerkin Formulation

Having in mind $\mathcal{C}^0$-conforming finite elements, let $V_h \subset V \cap L^\infty$ be a fixed finite-dimensional subspace. The Galerkin approximation $u_h \in V_h$ is defined by testing the weak equation (4.8) with elements from $V_h \subset V$ only:

Definition 6.1.1. *$u_h \in V_h$ is called a discrete solution of Problem 4.1.1 if*

$$\int_\Omega \kappa \nabla u_h \cdot \nabla v_h \, \mathrm{d}x + \int_I f(\cdot, [u_h])[v_h] \, \mathrm{d}\sigma = G(v_h) \tag{6.1}$$

holds for all $v_h \in V_h$.

By choosing a basis for V_h, the problem to determine u_h is equivalent to finding the root of a nonlinear function $F : \mathbb{R}^{\dim V_h} \to \mathbb{R}^{\dim V_h}$. Since f is continuously differentiable and (6.1) is uniquely solvable by Lemma 6.1.2 below, Newton's method is a canonical candidate for the solution of this problem and, in fact, we use it in our numerical simulations in Chapter 7. See [34] for a review of convergence criteria Newton's method which might be verified for (6.1).

Our focus, however, is to analyze the H^1-error between u_h and u. We start with establishing the well-posedness of the discrete problem (6.1).

Lemma 6.1.2. *There exists exactly one discrete solution of Problem 4.1.1 in the sense of Definition 6.1.1.*

Proof. In this proof, we denote by $\lesssim$ the relation $\lesssim_{M_1}$. Let us define $A : V_h \to V_h'$ by

$$\langle A(v_h), w_h \rangle := \int_\Omega \kappa \nabla v_h \cdot \nabla w_h \, \mathrm{d}x + \int_I f(\cdot, [v_h])[w_h] \, \mathrm{d}\sigma - G(w_h) \tag{6.2}$$

for $v_h, w_h \in V_h$.

From Assumption 4.1.2 and the inclusion $V_h \subset V \cap L^\infty$ it follows that $\langle A(v_h), w_h \rangle$ is well-defined and finite. Clearly, $A(v_h)$ is linear. As V_h is finite-dimensional it follows that A in fact maps V_h into its dual space V_h'.

Since V_h is contained in $V \cap L^\infty$, $\| \cdot \|_{V \cap L^\infty}$ is a norm on V_h. Note that $\| \cdot \|_{V \cap L^\infty}$ is given by

$$\|v\|_{V \cap L^\infty} = \max\{\|v\|_{1,2;\Omega}, \|v\|_{0,\infty;\Omega}\} \qquad \text{for all } v \in V \cap L^\infty.$$

To show continuity of $A : V_h \to V_h'$ let $v_h, \widetilde{v}_h, w_h \in V_h$. By defining

$$C_1 := M_2\big(\max\{\|v_h\|_{0,\infty;\Omega}, \|\widetilde{v}_h\|_{0,\infty;\Omega}\}\big)$$

it follows from Assumption 4.1.2, the mean value theorem and the trace theorem:

$$\begin{aligned}
&|\langle A(v_h) - A(\widetilde{v}_h), w_h \rangle| \\
&\leq \int_\Omega |\kappa \nabla (v_h - \widetilde{v}_h) \cdot \nabla w_h| \, \mathrm{d}x + \int_I |f(\cdot, [v_h]) - f(\cdot, [\widetilde{v}_h])| \, |[w_h]| \, \mathrm{d}\sigma \\
&\leq M_1 \int_\Omega |\nabla (v_h - \widetilde{v}_h) \cdot \nabla w_h| \, \mathrm{d}x + C_1 \int_I |[v_h - \widetilde{v}_h][w_h]| \, \mathrm{d}\sigma \\
&\lesssim (C_1 + 1) \|v_h - \widetilde{v}_h\|_V \|w_h\|_V.
\end{aligned} \tag{6.3}$$

Note that C_1 depends on v_h and $\widetilde{v}_h$. However, (6.3) still implies that A is locally Lipschitz continuous and thus in particular continuous.

Now, let us show that there exists some $R > 0$ such that $\langle A(v_h), v_h \rangle \geq 0$ for all $v_h \in V_h$ satisfying $\|v_h\|_V = R$. Denote by C_2 the positive constant from Lemma 4.2.3, that is, $\|\cdot\|_V \leq C_2 |\cdot|_V$ and recall Assumption 4.1.2. Then we have for $v_h \in V_h$:

$$
\begin{aligned}
\langle A(v_h), v_h \rangle &= \int_\Omega \kappa |\nabla v_h|^2 \,\mathrm{d}x + \int_I f(\cdot, [v_h])[v_h] \,\mathrm{d}\sigma - G(v_h) \\
&\geq M_1^{-1} \int_\Omega |\nabla v_h|^2 \,\mathrm{d}x + M_1^{-1} \int_I [v_h]^2 \,\mathrm{d}\sigma - M_1 \|v_h\|_V \\
&\geq C_2^{-2} M_1^{-1} \|v_h\|_V^2 - M_1 \|v_h\|_V .
\end{aligned}
$$

Thus for $R := (M_1 C_2)^2 > 0$ it follows from the Brouwer fixed point theorem, [71, Theorem 1.58], that there exists a $u_h \in V_h$ satisfying $A(u_h) = 0$ in V_h' and $\|u_h\|_V \leq R$. Note that from the definition of R it follows $R \lesssim 1$ and thus $\|u_h\|_V \leq R \lesssim 1$.

To prove uniqueness of the discrete solution, let $u_h, \widetilde{u}_h \in V_h$ be two discrete solutions, that is, $A(u_h) = A(\widetilde{u}_h) = 0$ in V_h'. Using $v_h = u_h - \widetilde{u}_h$ in both the defining equations (6.1) for u_h and $\widetilde{u}_h$ and subtracting gives by the mean-value theorem, Assumption 4.1.2 and Lemma 4.2.3:

$$
\begin{aligned}
0 &= \int_\Omega \kappa |\nabla (u_h - \widetilde{u}_h)|^2 \,\mathrm{d}x + \int_I \big(f(\cdot, [u_h]) - f(\cdot, [\widetilde{u}_h])\big)[u_h - \widetilde{u}_h] \,\mathrm{d}\sigma \\
&\gtrsim \int_\Omega |\nabla (u_h - \widetilde{u}_h)|^2 \,\mathrm{d}x + \int_I [u_h - \widetilde{u}_h]^2 \,\mathrm{d}\sigma \\
&\gtrsim \|u_h - \widetilde{u}_h\|_V .
\end{aligned}
$$

This implies $u_h = \widetilde{u}_h$. □

Let us for the remainder of Section 6.1 denote by u_h the discrete solution in the subspace V_h in the sense of Definition 6.1.1.

As in the continuous case, we obtain an estimate for the H^1-norm of u_h which does not depend on the subspace V_h but only on the constant M_1 from Assumption 4.1.2.

Remark 6.1.3. *There is a positive constant $C = C(M_1)$, which only depends on M_1 and in particular not on V_h, such that $\|u_h\|_V \leq C$ holds.*

Proof. This follows immediately from the proof of Lemma 6.1.2. □

6.1.1 Abstract Estimates

The first step in establishing convergence of the Galerkin method is relating the error of the Galerkin approximations u_h to the optimal approximation error in the subspace V_h. Due to the lack of a suitable polynomial growth condition we cannot imitate the proof of Céa's Lemma for the linear case. Instead we apply the technique used by Strang to prove the well-known Strang lemma, see [78], and combine it with the already proven L^∞-bound for the exact weak solution u (Theorem 4.2.7) to obtain the following quasi optimality result:

Lemma 6.1.4. *For all $R > 0$ and $v_h \in V_h$ satisfying $\|v_h\|_{0,\infty;\Omega} \leq R$ it holds*

$$\|u_h - v_h\|_V \leq C\|u - v_h\|_V \tag{6.4}$$

with a constant $C = C(M_1, M_2, R)$ depending on M_1, M_2 and R but neither *on V_h nor on v_h.*

By choosing v_h as the Clément interpolant of u in V_h one can derive explicit error estimates from (6.4), see Section 6.1.3 for the details of this argument. Note that, since the Clément interpolation operator is stable with respect to the L^∞-norm and u is bounded on Ω, it is actually not a problem that the constant C in (6.4) depends on the upper bound R for $\|v_h\|_{0,\infty;\Omega}$.

Proof of Lemma 6.1.4. Let $R > 0$ and $v_h \in V_h$ satisfying $\|v_h\|_{0,\infty;\Omega} \leq R$ be arbitrary. We will use the symbol $\lesssim$ for the relation $\lesssim_{M_1,M_2}$ and the symbol $\lesssim_R$ for $\lesssim_{M_1,M_2,R}$. Writing $w_h := u_h - v_h$, we have by Lemma 4.2.3 and Assumption 4.1.2:

$$\begin{aligned}
\|u_h - v_h\|_V^2 &\lesssim \int_\Omega |\nabla(u_h - v_h)|^2 \,\mathrm{d}x + \int_I [u_h - v_h]^2 \,\mathrm{d}\sigma \\
&\lesssim \int_\Omega \kappa\nabla(u_h - v_h)\cdot\nabla w_h \,\mathrm{d}x + \int_I \big(f(\cdot,[u_h]) - f(\cdot,[v_h])\big)[w_h]\,\mathrm{d}\sigma \\
&= \langle A(u_h), w_h\rangle - \langle A(v_h), w_h\rangle.
\end{aligned}$$

Since $w_h \in V_h \subset V$, it follows from (4.8) and (6.1) that

$$\langle A(u), w_h\rangle = \langle A(u_h), w_h\rangle = 0.$$

From the L^∞-estimate for u in Theorem 4.2.7 it follows $|[u]| \lesssim 1$ on I. Then, by the mean value theorem and the trace theorem, it follows from Assumption 4.1.2:

$$\begin{aligned}
\|u_h - v_h\|_V^2 &\lesssim \langle A(u) - A(v_h), w_h\rangle \\
&= \int_\Omega \kappa\nabla(u - v_h)\cdot\nabla w_h \,\mathrm{d}x + \int_I \big(f(\cdot,[u]) - f(\cdot,[v_h])\big)[w_h]\,\mathrm{d}\sigma \\
&\lesssim_R \int_\Omega |\nabla(u - v_h)\cdot\nabla w_h| \,\mathrm{d}x + \int_I \big|[u - v_h][w_h]\big|\,\mathrm{d}\sigma \\
&\lesssim \|u - v_h\|_V \|w_h\|_V.
\end{aligned}$$

Recalling the definition $w_h = u_h - v_h$, dividing by $\|w_h\|_V$ and collecting the estimates thus gives

$$\|u_h - v_h\|_V \lesssim_R \|u - v_h\|_V.$$

This finishes the proof. □

6.1.2 Abstract Convergence Criterion

Now let $(V_h)_{0<h\leq 1}$ be a family of finite dimensional subspaces $V_h \subset V \cap L^\infty$ for $h \in (0,1]$.

Motivated by the quasi optimality result Lemma 6.1.4 we define an abstract criterion which is sufficient for the convergence of the Galerkin approximations at the optimal rate. Roughly speaking, we require the existence of a family $(P_h)_{0<h\leq 1}$ of interpolation operators $P_h : V \to V_h$ which are stable with respect to L^∞ and admit optimal approximation properties with respect to the H^1-norm for functions in $H^{1+s} \cong H^{1+s}(\Omega_1) \oplus H^{1+s}(\Omega_2)$.

Definition 6.1.5. *For $s > 0$ the family $(V_h)_{0<h\leq 1}$ satisfies the* abstract h^s-convergence criterion *if there is a positive constant M_3 and for each $h \in (0,1]$ a map $P_h : V \to V_h$ such that it holds*

$$\|v - P_h(v)\|_V \leq M_3 |v|_{1+s,2;\Omega} h^s \qquad \text{for all } v \in V \cap H^{1+s} \tag{6.5}$$

and

$$\|P_h(v)\|_{0,\infty;\Omega} \leq M_3 \|v\|_{0,\infty;\Omega} \qquad \text{for all } v \in V \cap L^\infty. \tag{6.6}$$

Let us briefly show that this criterion is in fact sufficient for the convergence of u_h towards u at the optimal rate h^s under the additional assumption $u \in H^{1+s}$.

Remark 6.1.6. *Assume that $s > 0$ is given such that it holds $u \in H^{1+s}$ and such that the family $(V_h)_{0<h\leq 1}$ satisfies the abstract h^s-convergence criterion. Then it holds*

$$\|u - u_h\|_V \leq C|u|_{1+s,2;\Omega}\, h^s \qquad \text{for all } h \in (0,1]$$

with a constant $C = C(M_1, M_2, M_3)$ only depending on M_1, M_2 and M_3 but not *on h.*

Proof. Let $h \in (0,1]$ and denote by $\lesssim$ the relation $\lesssim_{M_1,M_2,M_3}$. From Theorem 4.2.7 it follows $u \in L^\infty$ and

$$\|u\|_{0,\infty;\Omega} \lesssim_{M_1} 1.$$

Thus by the assumption on $(V_h)_{0<h\leq 1}$ it holds (6.6) and we obtain

$$\|P_h(u)\|_{0,\infty;\Omega} \lesssim_{M_3} \|u\|_{0,\infty;\Omega} \lesssim_{M_1} 1.$$

Now we can apply the abstract error estimate Lemma 6.1.4 to conclude

$$\|P_h(u) - u_h\|_V \lesssim \|P_h(u) - u\|_V. \tag{6.7}$$

Combining (6.7) with the approximation property (6.5) of the family $(V_h)_{0<h\leq 1}$ thus gives:

$$\begin{aligned} \|u - u_h\|_V &\leq \|u - P_h(u)\|_V + \|P_h(u) - u_h\|_V \\ &\lesssim 2\|u - P_h(u)\|_V \\ &\lesssim |u|_{1+s,2;\Omega}\, h^s. \end{aligned}$$

This completes the proof. □

6.1.3 Finite Elements

It remains to provide an explicit example for a family $(V_h)_{0<h\leq 1}$ of subspaces satisfying the abstract h^s-convergence criterion from Definition 6.1.5.

In this section we will provide the tools for showing that in fact the $\mathcal{C}^0$-conforming finite element method on a simplicial conforming mesh and of arbitrary polynomial degree $p \in \mathbb{N}$ can be used to construct such a family. The important properties of the method are the approximation properties of the Clément interpolation operator and its stability with respect to L^∞.

As it is possible to construct interpolation operators with similar properties for other finite element spaces, it is very likely that the results can be extended to those cases, too, see for example [12, 11]. However, we will not discuss these extensions in the present work.

The details of the construction of V_h itself will be given later in Section 6.1.4. The concepts and notation of this section are following the monograph [41].

Meshes

The first step in the construction of finite element spaces is usually the discretization of the underlying domain $D \subset \mathbb{R}^d$. For the sake of a simpler presentation let us assume that D is a polyhedron, see the following definition.

Definition 6.1.7. *[47], [41, Definition 1.47].*

1. *A* convex polygon *in $\mathbb{R}^d$ is the convex hull of finitely many points in $\mathbb{R}^d$*
2. *A* polyhedron *in $\mathbb{R}^d$ is a Lipschitz domain in $\mathbb{R}^d$ (or more generally, a Lipschitz submanifold of $\mathbb{R}^d$ with boundary) which is the finite union of convex polygons.*

Definition 6.1.8. *A* mesh *for D is a finite collection $\mathcal{T}$ of compact and connected Lipschitz sets in $\mathbb{R}^d$ with non-empty interior such that it holds*

$$\overline{D} = \bigcup_{T\in\mathcal{T}} T \qquad \text{and} \qquad T_1^\circ \cap T_2^\circ = \emptyset \quad \text{für } T_1, T_2 \in \mathcal{T} \text{ with } T_1 \neq T_2.$$

A general mesh can consist of rather arbitrarily shaped elements. As mentioned in the introduction to this section, we only consider simplicial meshes. Let us briefly recall the definition of simplices in $\mathbb{R}^d$.

Definition 6.1.9.

1. *A set $\triangle \subset \mathbb{R}^d$ is called* d-simplex *if there are $d+1$ points $x_0, \dots, x_d$ which are in general position such that $\triangle$ is the convex hull of $x_0, \dots, x_d$.*
2. *The points $x_0, \dots, x_d$ are called the* vertices *of $\triangle$.*
3. *For $k = 0, \dots, d$ a* k-face *of $\triangle$ is the convex hull of $k+1$ of its vertices. The $(d-1)$-faces are simply called faces. Note that the 0-faces are the sets consisting of only one vertex and the only d-face is $\triangle$ itself.*

4. *The* reference d-simplex $\hat{\triangle}$ *is the simplex with the vertices*

$$\hat{x}_i = (\delta_{1i}, \ldots, \delta_{di}) \qquad \textit{for } i = 0, \ldots, d,$$

where δ_{ij} *denotes the Kronecker delta, that is,* $\delta_{ij} = 1$ *if* $i = j$ *and* $\delta_{ij} = 0$ *otherwise.*

Note that the *set of* vertices $\{x_0, \ldots, x_d\}$ of a simplex $\triangle$ is unique. Additionally, a set $\triangle \subset \mathbb{R}^d$ is a d-simplex if and only if there is an invertible affine linear map $F : \mathbb{R}^d \to \mathbb{R}^d$ such that it holds $\triangle = F(\hat{\triangle})$.

Definition 6.1.10.

1. *A* triangulation *for* D *is a mesh* $\mathcal{T}$ *for* D *such that each element* $T \in \mathcal{T}$ *is a* d*-simplex.*

2. *A triangulation* $\mathcal{T}$ *for* D *is called* (geometrically) conforming *if for all* $T_1, T_2 \in \mathcal{T}$ *one of the two following cases holds:*

   a) T_1 *and* T_2 *are disjoint.*

   b) $T_1 \cap T_2$ *is a* k*-face of both* T_1 *and* T_2 *for some* $k \in \{0, \ldots, d\}$.

Function Spaces

Throughout this section let us assume that we are given a conforming triangulation $\mathcal{T}$ for the polyhedron D in $\mathbb{R}^d$.

We will describe how one can construct a finite dimensional subspace $\mathcal{S}$ of $H^1(D)$ and give explicit examples of bases of $\mathcal{S}$ and its dual $\mathcal{S}'$. In order to define the interpolation operators it is necessary to pick the degrees of freedom for $\mathcal{S}'$, that is, a basis of $\mathcal{S}$. For our purpose it is sufficient and convenient to use the nodal function evaluations at the Lagrange nodes. The first step is to define the respective objects on the reference element $\hat{\triangle}$.

Definition 6.1.11 (Reference finite element)**.**

1. *The* reference function space *is* $\hat{\mathcal{S}} := \mathbb{P}_p(\hat{\triangle})$.

2. *Let* $\hat{\mathcal{X}}$ *denote the Lagrange nodes on* $\hat{\triangle}$*, that is,*

$$\hat{\mathcal{X}} = \left\{ \left(\frac{i_1}{p}, \ldots, \frac{i_d}{p} \right) \mid i_j \in \mathbb{N}_0 \textit{ for } j = 1, \ldots, d \textit{ and } i_1 + \ldots + i_d \le p \right\}.$$

   The reference degrees of freedom $\hat{\mathcal{N}}$ *are the function evaluations* $\hat{v} \mapsto \hat{v}(\hat{x})$ *at the Lagrange nodes on* $\hat{\triangle}$*, that is,*

$$\hat{\mathcal{N}} = \{ \hat{v} \mapsto \hat{v}(\hat{x}) \mid \hat{x} \in \hat{\mathcal{X}} \} \subset \hat{\mathcal{S}}'.$$

The triple $(\hat{\triangle}, \hat{\mathcal{S}}, \hat{\mathcal{N}})$ is called the *reference finite element*. It is in fact a finite element in the sense of [41, Definition 1.23].

Let us introduce the following notation: First of all, let us define $\hat{N}_i := i$ for $i \in \hat{\mathcal{N}}$. Furthermore let $\hat{x}_i \in \hat{\triangle}$ for $i \in \hat{\mathcal{N}}$ be the point of evaluation of the functional $\hat{N}_i$, that is, $\hat{N}_i(\hat{v}) = \hat{v}(\hat{x}_i)$ for all $\hat{v} \in \hat{\mathcal{S}}$.

Now for any d-simplex $\triangle$ we define the local finite element $(\triangle, \mathcal{S}, \mathcal{N})$ by transforming the reference finite element using any affine linear bijective mapping $F : \hat{\triangle} \to \triangle$.

Definition 6.1.12 (Local objects)**.** *Let $\triangle$ be any d-simplex and $F : \hat{\triangle} \to \triangle$ an affine linear bijection.*

1. *The* local function space *on $\triangle$ is $\mathcal{S}_\triangle := \mathbb{P}_p(\triangle)$*
2. *The* local degrees of freedom *on $\triangle$ are*

$$\mathcal{N}_\triangle := \{v \mapsto v(x) \mid x \in \mathcal{X}_\triangle\} \subset (\mathcal{S}_\triangle)'$$

where $\mathcal{X}_\triangle := F(\hat{\mathcal{X}})$ are the Lagrange nodes on $\triangle$.

Extending the notation introduced for the reference element, for $i \in \mathcal{N}_\triangle$ we define $N_{\triangle,i} := i$ and $x_{\triangle,i} \in \mathcal{X}_\triangle$ such that $N_{\triangle,i}(v) = v(x_{\triangle,i})$ holds for all $v \in \mathcal{S}_\triangle$. We will omit the subscript $\triangle$ whenever it is clear what the underlying simplex is.

3. *The* local shape functions *are the unique functions $\varphi_j := \varphi_{\triangle,j} \in \mathcal{S}_\triangle$ satisfying*

$$N_i(\varphi_j) = \delta_{ij} \qquad \text{for all } i, j \in \mathcal{N}_\triangle.$$

Then the triple $(\triangle, \mathcal{S}, \mathcal{N})$ is an affine equivalent finite element to the reference element [14, (3.4.1) Definition].

Now the global space $\mathcal{S}$ can be constructed from the local objects. The elements in $\mathcal{S}$ are the continuous functions on $\overline{D}$ whose restrictions to T are in the local spaces $\mathcal{S}_T$ for every element $T \in \mathcal{T}$. By restricting a function $v \in \mathcal{S}$ to an element $T \in \mathcal{T}$, the local degrees of freedom $\mathcal{N}_T$ can be considered as a subset of $\mathcal{S}$. The global degrees of freedom are then defined as the union of all local degrees of freedom $\mathcal{N}_\triangle$ over all elements $T \in \mathcal{T}$.

Definition 6.1.13 (Global objects)**.** *Let $\mathcal{T}$ be a conforming triangulation for D.*

1. *The* global $\mathcal{C}^0$-conforming space *is*

$$\mathcal{S}(\mathcal{T}) := \mathcal{S}^{p,0}(\mathcal{T}) := \{v \in \mathcal{C}^0(\overline{D}) \mid v|_T \in \mathcal{S}_T \text{ for all } T \in \mathcal{T}\}.$$

2. *For each element $T \in \mathcal{T}$ we consider $\mathcal{S}'_\mathcal{T}$ as a subspace of $\mathcal{S}'$ via*

$$N_T(v) = N(v|_T) \qquad \text{for } N_T \in \mathcal{S}'_\mathcal{T} \text{ and } v \in \mathcal{S}.$$

The global degrees of freedom *$\mathcal{N}$ are then defined as $\mathcal{N} := \bigcup_{T\in\mathcal{T}} \mathcal{N}_\mathcal{T}$.*

By defining $N_i := i$ and $x_i := x_{T,i}$ if $i \in \mathcal{N}_T \subset \mathcal{N}$, we canonically extend the notation introduced for the local objects to the global ones

3. *The* global shape functions *are the unique functions* $\varphi_j \in \mathcal{S}$ *satisfying*

$$N_i(\varphi_j) = \delta_{ij} \qquad \text{for all } i, j \in \mathcal{N}.$$

Remark 6.1.14.

1. *In general it does not hold* $(\mathcal{S}_{T_1})' \cap (\mathcal{S}_{T_2})' = \emptyset$ *for* $T_1, T_2 \in \mathcal{T}$ *satisfying* $T_1 \neq T_2$.
2. *It holds* $\varphi_i|_T = \varphi_{T,i}$ *if* $T \in \mathcal{T}$ *satisfies* $i \in \mathcal{N}_T$ *and* $\varphi_i|_T = 0$ *otherwise.*
3. *In particular, the support of* φ_i *is the element patch*

$$\omega_i := \cup\{T \in \mathcal{T} \mid x_i \in T\}.$$

around x_i *for* $i \in \mathcal{N}$.

Until now we have only constructed subspaces of $H^1(D)$. However, since the space $V = H^1_{\Gamma_2}(\Omega)$ incorporates homogeneous Dirichlet boundary values on the part Γ_2 of Ω_2, it is actually necessary to construct subspaces of $H^1_S(D)$ for a given measurable subset S of ∂D.

If $\mathcal{T}$ is conforming with the the partition of ∂D into S and $\partial D \setminus S$, see the following definition, this can be done by setting the nodal values on S to zero.

Definition 6.1.15. *A mesh* $\mathcal{T}$ *for* D *is called* conforming with $S \subset \partial D$ *if for every* $T \in \mathcal{T}$ *the intersection* $T \cap \partial D$ *is either contained in* $\overline{S}$ *or in* $\overline{\partial D \setminus S}$.

Definition 6.1.16.

1. *For* $S \subset D$ *let* $\mathcal{N}_S := \{i \in \mathcal{N} \mid x_i \in S\}$.
2. *Let* $S \subset \partial D$ *be a closed subset such that* $\mathcal{T}$ *is conforming with* S. *Then*

$$\mathcal{S}^{p,0}_S(\mathcal{T}) := \mathcal{S}_S := \mathcal{S} \cap H^1_S(D).$$

Bases for $\mathcal{S}'_S$ *and* $\mathcal{S}_S$ *are given by* $\mathcal{N} \setminus \mathcal{N}_S$ *and* $\{\varphi_i \mid i \in \mathcal{N} \setminus \mathcal{N}_S\}$, *respectively.*

Families of Meshes

Now suppose that for each $h \in (0, 1]$ we are given a mesh $\mathcal{T}_h$ for D.

We will present two important properties concerning the whole family $(\mathcal{T}_h)_{0<h\leq 1}$.

The first property is the so-called shape-regularity. It basically requires that the elements in $\mathcal{T}_h$ are *uniformly* non-degenerate for $h \in [0, 1)$, see Definition 6.1.17. This property is necessary to ensure the required approximation properties of the spaces $\mathcal{S}(\mathcal{T}_h)$ defined previously, see the subsequent Lemma 6.1.22 and Lemma 6.1.24.

Definition 6.1.17. *The family* $(\mathcal{T}_h)_{0<h\leq 1}$ *is called* shape-regular *if there is a positive constant* σ *such that it holds*

$$\sigma_T := \frac{h_T}{\varrho_T} \leq \sigma \qquad \text{for all } T \in \mathcal{T}_h \text{ and } h \in (0, 1]. \tag{6.8}$$

Here $\varrho_T \in [0, T]$ *is the radius of the largest ball contained in the set* $T \subset \mathbb{R}^d$ *and* σ_T *is called the* chunkiness parameter *of* T. *The smallest* σ *such that* (6.8) *is satisfied is the chunkiness parameter of the family* $(\mathcal{T}_h)_{0<h\leq 1}$.

The second property is the quasi-uniformity. It requires that the diameters of all elements in a fixed mesh $\mathcal{T}_h$ are comparable with constants which do not depend on h. We require quasi-uniform meshes in Section 6.1.5 where we use inverse estimates to derive the uniform L^∞-bound from the error estimates for the Galerkin approximations.

Definition 6.1.18. $(\mathcal{T}_h)_{0<h\leq 1}$ *is called* quasi-uniform *if there is a positive constant such that it holds*

$$h_{T_1} \leq C h_{T_2} \qquad \text{for all } T_1, T_2 \in \mathcal{T}_h \text{ and all } h \in (0,1]. \tag{6.9}$$

It is called locally quasi-uniform *if* (6.9) *holds for all* $T_1, T_2 \in \mathcal{T}_h$ *having non-empty intersection.*

Finally, let us note that if we are given a family $(\mathcal{T}_h)_{0<h\leq 1}$ of triangulations for D, we will enhance the symbols for the objects $\mathcal{S}$, $\mathcal{N}$, etc. defined for a single triangulation $\mathcal{T}$ with a subscript h to make clear the dependence on h, for example, we write $\mathcal{S}_h$, $\mathcal{N}_h$, etc. for the respective objects on $\mathcal{T}_h$.

Nodal Interpolation Operator

Let us recall that the nodal interpolant $\mathcal{I}v$ of a function $v \in \mathcal{C}^0(\overline{D})$ is the unique function in $\mathcal{S}$ which coincides with v in every node x_i for $i \in \mathcal{N}$.

We introduce the following symbols for its local and global version:

Definition 6.1.19.

1. *Let* $\triangle$ *be a d-simplex and* $v \in \mathcal{C}^0(\triangle)$*. Then* $\mathcal{I}^\triangle$ *is defined as*

$$\mathcal{I}^\triangle v := \sum_{i \in \mathcal{N}_\triangle} v(x_i)\varphi_i.$$

2. *The nodal interpolant* $\mathcal{I}v$ *of* $v \in \mathcal{C}^0(\overline{D})$ *is defined as*

$$\mathcal{I}v := \sum_{i \in \mathcal{N}} v(x_i)\varphi_i.$$

Note that it holds $\mathcal{I}v|_T = \mathcal{I}^T v$ for all elements $T \in \mathcal{T}$ and all $v \in \mathcal{C}^0(\overline{D})$. Let us repeat the standard stability and approximation properties of the nodal interpolation operator.

Remark 6.1.20. *[14, (4.4.1) Lemma] Let* $\triangle$ *be a d-simplex. Then it holds*

$$\|\mathcal{I}^\triangle v\|_{0,\infty;\triangle} \leq C\|v\|_{0,\infty;\triangle} \qquad \text{for all } v \in \mathcal{C}^0(\triangle)$$

with a positive constant $C = C(p)$ *which only depends on the polynomial degree p and in particular not on* $\triangle$ *or* v*.*

Lemma 6.1.21. *[14, Theorem 4.4.4] Suppose* $\triangle$ *is a* d*-simplex and let* $0 \le k \le s \le p+1$ *and* $1 \le q \le \infty$ *satisfy either* $s - d/q > 0$ *when* $q > 1$ *or* $s - d \ge 0$ *when* $q = 1$*. Then it holds*

$$|v - \mathcal{I}v|_{k,q;\triangle} \le C(\operatorname{diam}\triangle)^{s-k}|v|_{s,q;\triangle} \qquad \text{for all } v \in W^{s,q}(\triangle)$$

with a positive constant C *which only depends on* s*,* d *and the chunkiness parameter* $\sigma_\triangle$ *of* $\triangle$*.*

Lemma 6.1.22. *[14, (4.4.20) Theorem] Let* $(\mathcal{T}_h)_{0<h\le 1}$ *be a shape-regular family of conforming triangulations for* D *such that the maximal diameter of elements in* $\mathcal{T}_h$ *is bounded above by* h *for all* $h \in (0,1]$*. Additionally let* $0 \le q \le s \le p+1$ *and* $1 \le q \le \infty$ *satisfy either* $m - d/q > 0$ *when* $q > 1$ *or* $m - d \ge 0$ *when* $q = 1$*. Then it holds*

$$\Big(\sum_{T\in\mathcal{T}_h} \|v - \mathcal{I}_h v\|_{s,q;T}^q\Big)^{1/q} \le C h^{s-k}|v|_{s,q;D}$$

for all $v \in W^{s,q}(D)$ *with a positive constant* C *which only depends* d*,* s*,* q *and the chunkiness parameter* σ *of the family* $(\mathcal{T}_h)_{0<h\le 1}$*.*

Clément Interpolation Operator

Let us recall the definition of the Clément interpolant $\mathcal{C}v$ of a function $v \in L^2(D)$. Note that in contrast to the nodal interpolation operator it is not required that the function v is continuous. The idea is to approximate v by a polynomial v_i on each element patch ω_i and replace the coefficient $v(x_i)$ in the definition of the nodal interpolation operator by the value $v_i(x_i)$.

Definition 6.1.23. *[23, 12, 11] and [41, Section 1.6]. Let* $v \in L^2(D)$*. For* $i \in \mathcal{N}$ *denote by* v_i *the* $L^2(\omega_i)$*-orthogonal projection of* $v|_{\omega_i}$ *onto* $\mathbb{P}_p(\omega_i)$*. The Clément interpolant* $\mathcal{C}v \in \mathcal{S}_S$ *of* v *is defined by*

$$\mathcal{C}v := \sum_{i\in\mathcal{N}\setminus\mathcal{N}_S} v_i(x_i)\varphi_i.$$

Now assume that we are given a family $(\mathcal{T}_h)_{0<h\le 1}$ of triangulations $\mathcal{T}_h$ of D such that every $\mathcal{T}_h$ is conforming with S and, additionally, that the maximal diameter of elements in $\mathcal{T}_h$ is bounded above by h.

If $(\mathcal{T}_h)_{0<h\le 1}$ is shape-regular, the Clément interpolation operator satisfies similar optimal approximation properties as the nodal interpolation operator:

Lemma 6.1.24. *[23, Theorem 2] Let* $(\mathcal{T}_h)_{0<h\le 1}$ *be shape-regular and* $0 \le k \le s \le p+1$*. Then it holds*

$$|v - \mathcal{C}_h v|_{k,2;D} \le C h^{s-k}|v|_{s,2;D} \qquad \text{for all } v \in H^s(D) \cap H^1_S(D) \text{ and } h \in (0,1]$$

with a positive constant C *which does not depend on* v *or* h*.*

In order to verify the second condition (6.6) of our abstract h^s-convergence criterion it is necessary to show the stability with respect to L^∞. This is done in Lemma 6.1.25. Note that the statement of Lemma 6.1.25 is contained in [12, Theorem 2.1] for example. However we choose to give our own proof here since in [12] a slightly modified version of the Clément interpolation operator was analyzed.

Lemma 6.1.25. *Let $(\mathcal{T}_h)_{0<h\leq 1}$ be shape-regular. Then it holds*

$$\|\mathcal{C}_h v\|_{0,\infty;D} \leq C\|v\|_{0,\infty;D} \qquad \text{for all } v \in L^\infty(D) \text{ and } h \in (0,1]$$

with a positive constant C which does not depend on v or h.

Proof. Let $v \in L^\infty(D)$ and $h \in (0,1]$. We use the symbol $\lesssim$ for the relation $\lesssim_{\neg v,\neg h}$.

Fix $i \in \mathcal{N}_h$, denote by $P_i : L^2(\omega_i) \to \mathbb{P}_p(\omega_i)$ the $L^2(\omega_i)$-orthogonal projection onto $\mathbb{P}^p(\omega_i)$ and let $v_i := P_i(v|_{\omega_i})$. Finally, let $T \in \mathcal{T}_h$ be an element contained in ω_i.

Since $(T_h)_{0<h\leq 1}$ is assumed to be shape-regular, we can apply the local inverse inequality from [41, Lemma 1.138]. From the inclusion $T \subset \omega_i$ it follows

$$\|v_i\|_{0,\infty,T} \lesssim \mu(T)^{-1/2}\|v_i\|_{0,2,T} \leq \mu(T)^{-1/2}\|v_i\|_{0,2,\omega_i}.$$

Since P_i is an $L^2(\omega_i)$-orthogonal projection and $v_i = P_i(v|_{\omega_i})$, we have

$$\|v_i\|_{0,2,\omega_i} = \|P_i(v|_{\omega_i})\|_{0,2,\omega_i} \leq \|v\|_{0,2,\omega_i}.$$

The shape-regularity of the family $(\mathcal{T}_h)_{0<h\leq 1}$ implies the local quasi-uniformity, see [24, Section 2.2], that is,

$$\operatorname{diam}(T_1) \lesssim \operatorname{diam}(T_2)$$

for all $T_1, T_2 \in \mathcal{T}_h$ satisfying $T_1 \cap T_2 \neq 0$, see Definition 6.1.18. This implies, however, $\mu(\omega_i) \lesssim \mu(T)$. Combining these estimates and applying the Hölder-inequality, we obtain

$$\begin{aligned}\|v_i\|_{0,\infty,T} &\lesssim \mu(T)^{-1/2}\|v_i\|_{0,2,\omega_i}\\ &\leq \mu(T)^{-1/2}\|v\|_{0,2,\omega_i}\\ &\leq \mu(T)^{-1/2}\mu(\omega_i)^{1/2}\|v\|_{0,\infty,\omega_i}\\ &\lesssim \|v\|_{0,\infty,D}.\end{aligned}$$

Since $T \subset \omega_i$ was an arbitrary element contained in ω_i, it follows

$$|v_i(x_i)| \leq \|v_i\|_{0,\infty,\omega_i} \lesssim \|v\|_{0,\infty,\Omega}.$$

For arbitrary $x \in D$ we have by Remark 6.1.14:

$$\begin{aligned}\left|\mathcal{C}_h v(x)\right| &\lesssim \|v\|_{0,\infty;\Omega} \sum_{i\in\mathcal{N}} |\varphi_i(x)|\\ &\leq \|v\|_{0,\infty;\Omega} \sum_{i\in\hat{\mathcal{N}}} \|\hat{\varphi}_i\|_{0,\infty;\hat{\triangle}}\\ &\lesssim \|v\|_{0,\infty;\Omega}.\end{aligned}$$

This finishes the proof. □

6.1.4 Optimal Convergence of FEM

Let us apply the results from the previous section 6.1.3 to our actual Problem 4.1.1.

To this end, assume $\overline{\Omega}_i$ is polyhedral and $(\mathcal{T}_{h,i})_{0<h\leq 1}$ is a shape-regular family of conforming triangulations for Ω_i such that the maximal diameter of elements in $\mathcal{T}_{h,i}$ is bounded above by h for $i = 1, 2$. Additionally, assume that $\mathcal{T}_{h,2}$ is conforming with Γ_2 for all $h \in (0, 1]$.

Let us adopt the notation introduced in Section 6.1.3 and denote for $h \in (0, 1]$ and $i = 1, 2$ the respective objects with the same letter followed by a subscript indicating the dependence on h and i. For example the degrees of freedom on $\mathcal{S}_{h,i} := \mathcal{S}^{p,0}(\mathcal{T}_{h,i})$ are denoted by $\mathcal{N}_{h,i}$ for $i = 1, 2$.

We let $\mathcal{S}_h := \mathcal{S}_{h,1} \oplus \mathcal{S}_{h,2}$ and as in Section 6.1.3, we consider $\mathcal{S}'_{h,i}$ as a subspace of $\mathcal{S}'_h$ by applying the projection $(v_1, v_2) \mapsto v_i$ for $i = 1, 2$. Following the naming conventions in Section 6.1.3 we let $\mathcal{N}_h := \mathcal{N}_{h,1} \cup \mathcal{N}_{h,2}$ and canonically extend the mappings

$$N_{h,i} : j \mapsto N_{h,i;j}, \qquad x_{h,i} : j \mapsto x_{h,i;j}, \qquad \text{and} \qquad \varphi_{h,i} : j \mapsto \varphi_{h,i;j}$$

defined on $\mathcal{N}_{h,i}$ for $i = 1, 2$, to $\mathcal{N}_h = \mathcal{N}_{h,1} \cup \mathcal{N}_{h,2}$ in order to obtain mappings $N := N_h$, $x := x_h$ and $\varphi := \varphi_h$ defined on $\mathcal{N}_h$.

Finally let $V_h := \mathcal{S}^{p,0}(\mathcal{T}_{h,1}) \oplus \mathcal{S}^{p,0}_{\Gamma_2}(\mathcal{T}_{h,2})$ and denote by $u_h \in V_h$ the corresponding discrete solution to Problem 4.1.1.

Note that the elements in V_h in general admit jumps across I, just like functions from V or $\mathcal{C}^0_{\mathrm{b}}(\Omega)$. As a consequence, the given finite element discretization can be considered $\mathcal{C}^0_{\mathrm{b}}$-conforming but *not* $\mathcal{C}^0(\overline{\Omega})$-conforming.

Lemma 6.1.26. *For $0 < s \leq p$ the family $(V_h)_{0<h\leq 1}$ satisfies the abstract h^s-convergence criterion from Definition 6.1.5.*

Proof. Denote by $\mathcal{C}_{h,1} : L^2(\Omega_1) \to \mathcal{S}^{p,0}(\mathcal{T}_{h,1})$ and $\mathcal{C}_{h,2} : L^2(\Omega_2) \to \mathcal{S}^{p,0}_{\Gamma_2}(\mathcal{T}_{h,2})$ two Clément interpolation operators from Definition 6.1.23, where $S = \emptyset$ for $\mathcal{C}_{h,1}$ and $S = \Gamma_2$ for $\mathcal{C}_{h,2}$.

Let $\mathcal{C}_h := \mathcal{C}_{h,1} \oplus \mathcal{C}_{h,2}$ be the composed Clément interpolation operator, that is,

$$\mathcal{C}_h : L^2 \to V_h, \; v \mapsto \mathcal{C}_h v := (\mathcal{C}_{h,1} v_1, \mathcal{C}_{h,2} v_2).$$

Now define $P_h := \mathcal{C}_h|_V$. From Lemma 6.1.24 and Lemma 6.1.25 it follows that the conditions (6.5) and (6.6) are satisfied. □

An immediate consequence is of course the convergence of the discrete solutions u_h towards the exact solution u at the optimal rate, see the following corollary.

Corollary 6.1.27. *Assume the weak solution u of Problem 4.1.1 satisfies $u \in H^{1+s}$ for some $0 < s \leq p$. Then it holds*

$$\|u - u_h\|_V \leq C h^s \qquad \textit{for all } h \in (0, 1]$$

with a positive constant C which does not depend on h.

Proof. This is a direct consequence of Lemma 6.1.26 and Remark 6.1.6. □

6.1.5 Uniform L^∞-bound for FEM

Even though the question of convergence has been answered in the previous sections, it is still an interesting question if the L^∞-bound for the exact solution u (Theorem 4.2.7) carries over to an L^∞-bound for the discrete solutions u_h which is uniform in h.

In fact, when we started working on the error estimates we tried to prove such a uniform L^∞-bound first. The reason behind that was that such a bound would have enabled us to generalize the proof for the Céa Lemma in the linear case to the current problem and the error estimates would have followed immediately.

However, it turned out that the proofs for the continuous L^∞-bound could not be generalized to the discrete equation, even for the most simple cases. The basic idea of the Stampacchia truncation method used in the continuous proof was to use $(u-k)_+$ as a test function where k is an arbitrary positive number. However, the function $(u_h-k)_+$ is in general not an element of the finite dimensional subspace V_h. The canonical idea of using a *suitable* approximation $v_h \in V_h$ to $(u_h-k)_+$ could not be used successfully because during this process the important monotonicity properties of the equation got destroyed.

Nevertheless, it is possible to recover a uniform L^∞-bound for the discrete solutions from the error estimate by using inverse inequalities if we make an additional rather artificial regularity assumption on the exact solution u. Furthermore, it is required that the underlying meshes are quasi-uniform, see Definition 6.1.18.

Corollary 6.1.28. *Let $(\mathcal{T}_{h,i})_{0<h\leq 1}$ be shape-regular and quasi-uniform for $i=1,2$ and assume that $u \in H^{d/2+\varepsilon}$ and $p \geq d/2+\varepsilon-1$ for some $\varepsilon > 0$. Then there exists a positive constant C which does not depend on h such that $\|u_h\|_{0,\infty,\Omega} \leq C$ holds.*

Proof. Without loss of generality, we assume $\varepsilon \in (0,1)$. Let $h \in (0,1]$ and denote by $\lesssim$ the relation $\lesssim_{\neg h}$. Defining

$$q := \frac{d}{1-\varepsilon},$$

it holds $q \in (d,\infty)$ and thus it follows from Sobolev embedding:

$$\|u_h - \mathcal{C}_h u\|_{0,\infty;\Omega} \lesssim \|u_h - \mathcal{C}_h u\|_{1,q;\Omega}.$$

By the quasi-uniformity of $(\mathcal{T}_h)_{0<h\leq 1}$ the inverse inequality

$$\|u_h - \mathcal{C}_h u\|_{1,q;\Omega} \lesssim h^{d/q-d/2}\|u_h - \mathcal{C}_h u\|_V$$

holds, see for example [41, Corollary 1.141]. Since $p \geq d/2+\varepsilon-1$, we can apply Lemma 6.1.26 and Lemma 6.1.24 to obtain

$$\|u_h - \mathcal{C}_h u\|_V \lesssim h^{d/2+\varepsilon-1}|u|_{d/2+\varepsilon,2;\Omega}.$$

Combining the above estimates and the definition of q yields:

$$\|u_h - \mathcal{C}_h u\|_{0,\infty;\Omega} \lesssim h^{(d/q-d/2)+(d/2+\varepsilon-1)} = h^{1-\varepsilon+\varepsilon-1} = 1.$$

From the L^∞-stability of the Clément operator (Lemma 6.1.25) and the boundedness of the exact solution u (Theorem 4.2.7) it finally follows:

$$\|u_h\|_{0,\infty;\Omega} \leq \|u_h - \mathcal{C}_h u\|_{0,\infty;\Omega} + \|\mathcal{C}_h u\|_{0,\infty;\Omega} \lesssim 1 + \|u\|_{0,\infty;\Omega} \lesssim 1.$$

This finishes the proof. □

Remark 6.1.29. *The number $d/2 + \varepsilon$ for some $\varepsilon > 0$ is the smallest Sobolev-exponent s such that the Sobolev-embedding $H^s \hookrightarrow \mathcal{C}^0_{\mathrm{b}} \subset L^\infty$ is continuous.*

For $d \in \{1, 2, 3\}$ we have $d/2 - 1 < 1$ and thus the assumption $p > d/2 + \varepsilon - 1$ for some $\varepsilon > 0$ is automatically satisfied in that case.

6.2 Discretization with Quadrature

As it has already been pointed out, the error analysis for the discretization introduced in Section 6.1 would have been a lot easier if it was possible to prove a uniform L^∞-bound for the discrete solutions first, instead of recovering such an estimate from the error estimate using inverse inequalities like in Corollary 6.1.28.

In the article [20] the Stampacchia truncation method is applied to the *linear* finite element method for the Poisson problem on non-negative meshes, see Definition 6.2.15. The basic idea is to test the equation with the nodal interpolant of $(u_h - k)_+$ and to use monotonicity properties of the nodal interpolation operator and the discrete Laplacian which are available for linear finite elements on non-negative meshes [20, 73, 51, 50, 49].

In order to apply the techniques of [20] to our case, however, it is necessary to modify the discrete equations. Instead of exactly evaluating the integral corresponding to the interface nonlinearity, it is approximated by the trapezoidal rule, see (6.10).

As it turns out, it is still possible to establish well-posedness of the modified discrete problem. Under the additional assumption that f is $\mathcal{C}^2$ and the exact solution u is in $H^2 \cap W^{1,4}$, the error introduced by deviating from the continuous equations is of order $\mathcal{O}(h)$. As a result, the modified discrete system still admits convergence at optimal linear rate, see Lemma 6.2.12.

In addition, we can apply the technique of [20] and succeed to prove a comparison principle (Lemma 6.2.19) and a uniform L^∞-bound (Section 6.2.6) for the modified discrete system, at least on *non-negative triangulations*.

6.2.1 Preliminaries

As in Section 6.1, let $\overline{\Omega}_i$ be a polyhedron and $\mathcal{T}_{h,i}$ a conforming triangulation for Ω_i for $i = 1, 2$ such that $\mathcal{T}_{h,2}$ is conforming with Γ_2.

In this section we additionally assume that $\mathcal{T}_{h,i}$ is conforming with I for $i = 1, 2$, and that $\mathcal{T}_h = \mathcal{T}_{h,1} \cup \mathcal{T}_{h,2}$ is a conforming triangulation for $\overline{\Omega}$. This is satisfied if and only if the intersections of triangles $T_1 \in \mathcal{T}_{h,1}$ and $T_2 \in \mathcal{T}_{h,2}$ are either empty or k-faces of both T_1 and T_2 for some $k \in \{0, \dots, d-1\}$. We will use the notation introduced in Section 6.1, however, we only consider the case of linear elements here, that is, $p = 1$.

As mentioned in the introduction to this section, the modified discrete equations are obtained by applying the trapezoidal rule to the interface nonlinearity. Since the trapezoidal rule is defined by exactly integrating the linear interpolant of the function, it is necessary to define the nodal interpolation operator on the interface I:

Definition 6.2.1.

1. $\mathcal{F}_h^I := \{F \subset \overline{I} \mid \exists T \in \mathcal{T}_h$ *such that* F *is a face of* $T\}$.
2. $\mathcal{S}_h^I := \mathcal{S}^{1,0}(\mathcal{F}_h^I)$ *is the trace space of* $\mathcal{S}_h$ *on* I.
3. $\mathcal{I}_h^I : \mathcal{C}^0(\overline{I}) \to \mathcal{S}_h^I$ *is the nodal interpolation operator onto* $\mathcal{S}_h^I$.

Note that by the assumption that $\mathcal{T}_h$ is a conforming triangulation of Ω, $\mathcal{F}_h^I$ is in fact a conforming triangulation for I. Therefore, the trace space $\mathcal{S}_h^I$ and the nodal interpolation operator $\mathcal{I}_h^I$ are well-defined.

Now let us briefly recall two immediate but important monotonicity properties of the nodal interpolation operator for linear elements:

Remark 6.2.2. *Let* $\triangle$ *be a* d*-simplex and* $v \in \mathcal{C}^0(\triangle)$*. Denote by* $\mathcal{I}^{\triangle} v$ *the linear nodal interpolant in* $\mathbb{P}_1(\triangle)$ *of* v *with respect to the Lagrange-nodes* $\mathcal{X}_{\triangle}$*. Then it holds:*

1. $v(x_i) \geq 0$ *for all* $i \in \mathcal{N}_{\triangle}$ *implies* $\mathcal{I}^{\triangle} v \geq 0$ *on* $\triangle$.
2. *If* v *is convex, it holds* $\mathcal{I}^{\triangle} v \geq v$ *on* $\triangle$.

6.2.2 Formulation

In order to apply the trapezoidal rule to a function, it is necessary that the point-values of the function are well-defined, that is, the function needs to be continuous. However, the nonlinearity $f : (x, z) \mapsto f(x, z)$ is only assumed to be measurable with respect to x, see Assumption 4.1.2. Therefore we need to additionally *postulate* the continuity of f with respect to x, see the following assumption.

Assumption 6.2.3. *For every* $z \in \mathbb{R}$*, the mapping* $f(\cdot, z) : x \mapsto f(x, z)$ *is continuous on* $\overline{I}$.

Note that Assumption 6.2.3 is satisfied when $f(x, z) = i_{12}(c(x), z)$, where i_{12} satisfies the properties of Assumption 3.5.3 and $c \in \mathcal{C}_{\mathrm{b}}^0$. This is for example satisfied by the solutions (c, u) of the fully coupled problem provided by Theorem 5.6.1

Now we can finally state the modified discrete formulation:

Definition 6.2.4. *A function* $u_h \in V_h$ *is called a modified discrete solution of Problem 4.1.1 if*

$$\langle A_h(u_h), v_h \rangle := \int_\Omega \kappa \nabla u_h \cdot \nabla v_h \,\mathrm{d}x + \int_I \mathcal{I}_h^I\big(f(\cdot, [u_h])[v_h]\big) \,\mathrm{d}\sigma - G(v_h) = 0 \tag{6.10}$$

holds for all $v_h \in V_h$.

Lemma 6.2.5. *All terms in* (6.10) *are well-defined and finite*

Proof. Since $V_h \subset H^1 \cap \mathcal{C}^0_{\mathrm{b}}$ it suffices to show that $f(\cdot, [u_h])$ is continuous on $\overline{I}$. To this end let $(x_n)_n \subset \overline{I}$ be a convergent sequence, say, $x_n \to x$ for some $x \in \overline{I}$. Then it holds:

$$\begin{aligned}&\big|f(x_n, [u_h(x_n)]) - f(x, [u_h(x)])\big| \\ &\leq \big|f(x_n, [u_h(x_n)]) - f(x_n, [u_h(x)])\big| + \big|f(x_n, [u_h(x)]) - f(x, [u_h(x)])\big| \\ &\leq \sup_{\xi} \|\partial_z f(\cdot, \xi)\|_{0,\infty;I} |x_n - x| + \big|f(x_n, [u_h(x)]) - f(x, [u_h(x)])\big|.\end{aligned}$$

The supremum is taken over all $\xi \in \mathbb{R}$ satisfying $|\xi| \leq \|u_h\|_{0,\infty;\Omega} =: R$. Clearly, this bound does not depend on n and thus it follows by Assumption 4.1.2, Assumption 6.2.3 and the inclusion $V_h \subset \mathcal{C}^0_{\mathrm{b}}$:

$$\begin{aligned}&\big|f(x_n, [u_h(x_n)]) - f(x, [u_h(x)])\big| \\ &\leq M_2(R)|x_n - x| + \big|f(x_n, [u_h(x)] - f(x, u_h(x))\big| \to 0 \qquad \text{as } n \to \infty.\end{aligned}$$

Since $(x_n)_n$ was arbitrary, this shows that $f(\cdot, [u_h])$ is continuous. □

Before proving well-posedness of (6.10), let us note the following basic estimate which is an immediate consequence of the monotonicity of f (Assumption 4.1.2) and the nodal interpolation operator (Remark 6.2.2).

Remark 6.2.6. *For $v_h, w_h \in W_h$ it holds*

$$\mathcal{I}^I_h\Big(\big(f(\cdot, [v_h]) - f(\cdot, [w_h])\big)[v_h - w_h]\Big) \geq M_1^{-1}[v_h - w_h]^2 \qquad \text{on } I.$$

Proof. Let $T \in \mathcal{F}^I_h$ be arbitrary and define $v := v_h|_T$ and $w := w_h|_T$. From Assumption 4.1.2 it follows:

$$\big(f(\cdot, [v]) - f(\cdot, [w])\big)[v - w] \geq M_1^{-1}[v - w]^2.$$

Therefore, Remark 6.2.2 implies

$$\mathcal{I}^I_h\Big(\big(f(\cdot, [v]) - f(\cdot, [w])\big)[v - w]\Big) \geq M_1^{-1}\mathcal{I}^I_h\big([v - w]^2\big). \tag{6.11}$$

Since v and w are linear, $[v - w]^2$ is convex. Thus it follows again from Remark 6.2.2 that

$$\mathcal{I}^I_h\big([v - w]^2\big) \geq [v - w]^2. \tag{6.12}$$

Combining (6.11) and (6.12) completes the proof. □

Lemma 6.2.7. *There exists exactly one modified discrete solution of Problem 4.1.1.*

Proof. Denote by $\lesssim$ the relation $\lesssim_{M_1}$. Let us first prove existence. For $v_h, w_h \in V_h$, by Lemma 6.2.5, $\langle A_h(v_h), w_h\rangle$ is well-defined and finite. Clearly, $A_h(v_h)$ is linear.

Let us show that $A_h : V_h \to V_h'$ is continuous. To this end let $v_h, \widetilde{v}_h, w_h \in V_h$. Note that by the construction of V_h it holds $V_h \subset V \cap L^\infty$. Define $C_1 := \max\{\|v_h\|_{0,\infty;\Omega}, \|\widetilde{v}_h\|_{0,\infty;\Omega}\}$. Then it follows from Hölder's inequality, Remark 6.1.20 and Assumption 4.1.2:

$$\begin{aligned}
&|\langle A(v_h) - A(\widetilde{v}_h), w_h\rangle| \\
&\lesssim \|v_h - \widetilde{v}_h\|_V \|w_h\|_V + \|f(\cdot, [v_h]) - f(\cdot, [\widetilde{v}_h])\|_{0,\infty;I} \|[w_h]\|_{0,\infty;I} \\
&\le \|v_h - \widetilde{v}_h\|_V \|w_h\|_V + M_2(C_1)\|[v_h - \widetilde{v}_h]\|_{0,\infty;I} \|[w_h]\|_{0,\infty;I} \\
&\le M_2(C_1)\|v_h - \widetilde{v}_h\|_{V\cap L^\infty} \|w_h\|_{V\cap L^\infty}.
\end{aligned}$$

This shows that A_h is locally Lipschitz continuous and thus it is in particular continuous.

Now let us show that there exists some $R > 0$ such that $\langle A_h(v_h), v_h\rangle > 0$ holds for all $v_h \in V_h$ satisfying $\|v_h\|_V = R$. Denote by C_2 a positive constant from Lemma 4.2.3, that is, $\|\cdot\|_V \le C_2|\cdot|_V$ and by C_3 the operator norm of the embedding $H^1 \hookrightarrow W^{1,1}$. Then it follows from Assumption 4.1.2 and Remark 6.2.6:

$$\begin{aligned}
\langle A_h(v_h), v_h\rangle &= \int_\Omega \kappa|\nabla v_h|^2 \,\mathrm{d}x + \int_I \mathcal{I}_h^I\big(f(\cdot, [v_h])[v_h]\big) \,\mathrm{d}\sigma - G(v_h) \\
&\ge M_1^{-1} \int_\Omega |\nabla v_h|^2 \,\mathrm{d}x + M_1^{-1} \int_I [v_h]^2 \,\mathrm{d}\sigma - G(v_h) \\
&\ge C_2^{-2} M_1^{-1} \|v_h\|_V^2 - M_1 C_3 \|v_h\|_v.
\end{aligned}$$

Thus for $R := M_1^2 C_2 C_3$ it follows from the Brouwer fixed point theorem that there exists a $v_h \in V_h$ satisfying $A(v_h) = 0$ in V_h' and $\|v_h\|_V \le R$ [71, Theorem 1.58]. Note that it holds $R \lesssim 1$.

Now let us show uniqueness of u_h. To this end, let $\widetilde{u}_h \in V_h$ be another modified discrete solution. Using $v_h := u_h - \widetilde{u}_h \in V_h$ in the defining equations (6.10) for both u_h and $\widetilde{u}_h$ and then taking the difference of the resulting equations gives, by using Assumption 4.1.2 and Remark 6.2.6:

$$\begin{aligned}
0 &= \int_\Omega \kappa|\nabla(u_h - \widetilde{u}_h)|^2 \,\mathrm{d}x + \int_I \mathcal{I}_h^I\Big(\big(f(\cdot, [u_h]) - f(\cdot, [\widetilde{u}_h])\big)[u_h - \widetilde{u}_h]\Big) \,\mathrm{d}\sigma \\
&\gtrsim \int_\Omega |\nabla(u_h - \widetilde{u}_h)|^2 \,\mathrm{d}x + \int_I [u_h - \widetilde{u}_h]^2 \,\mathrm{d}\sigma.
\end{aligned}$$

From Lemma 4.2.3 it follows that $u_h = \widetilde{u}_h$. □

In the remainder of this section let us denote by u_h the modified discrete solution to Problem 4.1.1 in the sense of Definition 6.2.4 with respect to the triangulation $\mathcal{T}_h$.

In the proof of Lemma 6.2.7, R does only depend on M_1 but not on the triangulation $\mathcal{T}_h$. As a consequence, we have actually proven the following lemma.

Lemma 6.2.8. *There is a positive constant $C = C(M_1)$ only depending on M_1 but not on $\mathcal{T}_h$ such that it holds $\|u_h\|_V \le C$.*

6.2.3 Abstract Estimates

Now we can state the following abstract error estimate, which is obtained by reviewing the proof of the Strang lemma [78].

Lemma 6.2.9. *There is a positive constant $C = C(M_1)$ depending on M_1 but not on the triangulation $\mathcal{T}_h$ such that it holds*

$$\begin{aligned} &\|u - u_h\|_V \\ &\leq \inf_{v_h \in V_h} \Bigg\{ \|u - v_h\|_V \\ &\qquad + C \sup_{w_h \in V_h} \frac{1}{\|w_h\|_V} \Big(\int_\Omega \nabla(u - v_h) \cdot \nabla w_h \,\mathrm{d}x \\ &\qquad\qquad + \int_I \big(f(\cdot,[u]) - f(\cdot,[v_h])\big)[w_h] \,\mathrm{d}\sigma \\ &\qquad\qquad + \int_I f(\cdot,[v_h])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[v_h])[w_h]\big) \,\mathrm{d}\sigma \Big) \Bigg\}. \end{aligned} \tag{6.13}$$

Proof. Let $v_h \in V_h$ be arbitrary and define $w_h := u_h - v_h \in V_h$. Denote by $\lesssim$ the relation $\lesssim_{M_1}$. By Lemma 4.2.3, Assumption 4.1.2 and the properties of the nodal interpolation error from Remark 6.2.2, we have:

$$\begin{aligned} \|u_h - v_h\|_V^2 &\lesssim \int_\Omega |\nabla(u_h - v_h)|^2 \,\mathrm{d}x + \int_I [u_h - v_h]^2 \,\mathrm{d}\sigma \\ &\lesssim \int_\Omega \kappa |\nabla(u_h - v_h)|^2 \,\mathrm{d}x + \int_I \mathcal{I}_h^I\Big(\big(f(\cdot,[u_h]) - f(\cdot,[v_h])\big)[u_h - v_h]\Big) \,\mathrm{d}\sigma \\ &= \langle A_h(u_h), w_h \rangle - \langle A_h(v_h), w_h \rangle. \end{aligned}$$

Since u_h is the modified discrete solution and $V_h \subset V$, it holds

$$\langle A_h(u_h), w_h \rangle = \langle A(u), w_h \rangle = 0.$$

It follows:

$$\begin{aligned} \|u_h - v_h\|_V^2 &\lesssim \langle A_h(u_h) - A_h(v_h), w_h \rangle \\ &= \langle A(u) - A_h(v_h), w_h \rangle \\ &= \langle A(u) - A(v_h), w_h \rangle + \langle A(v_h) - A_h(v_h), w_h \rangle. \end{aligned} \tag{6.14}$$

Finally, writing

$$\|u - u_h\|_V \leq \|u - v_h\|_V + \frac{1}{\|w_h\|_V} \|v_h - u_h\|_V^2$$

and first taking the supremum over all $w_h \in V_h$ and then the infimum over all $v_h \in V_h$, the claim follows from (6.14) and the definitions of A_h and A. □

6.2.4 Linear Convergence

For every $h \in (0,1]$ let $\mathcal{T}_h$ be a conforming triangulation for $\overline{\Omega}$ with the properties described in the introduction to this section such that the maximal diameter of elements in $\mathcal{T}_h$ is bounded above by h.

We will use the abstract estimate, (6.13), to prove linear convergence of the modified discrete solutions to the continuous weak solution of Problem 4.1.1. For this, however, we impose an additional regularity assumption on f:

Assumption 6.2.10. *f is $\mathcal{C}^2$ in an open neighborhood of $I \times \mathbb{R} \subset \mathbb{R}^d \times \mathbb{R}$.*

Example 6.2.11. *Suppose $c_i \in \mathcal{C}^2(\overline{\Omega}_i)$ is given satisfying $M_1^{-1} \leq c_i \leq c_{\max,i} - M_1^{-1}$ on Ω_i for $i = 1, 2$ and f is defined as*

$$\begin{aligned} f(\cdot, z) &= i_{\mathrm{BV}}(c_1, c_2, z + \ln(c_1)) \\ &= c_1^{1/2} c_2^{1/2} (1 - c_2)^{1/2} \Big(e^{(z+\ln(c_1) - U(c_2))/2} - e^{-(z+\ln(c_1)-U(c_2))/2} \Big). \end{aligned}$$

with a $\mathcal{C}^2$ function $U : (0,1) \to \mathbb{R}$, see (3.26)*. Then Assumption 6.2.10 is satisfied.*

With the extra assumption on f at hand, we can combine the abstract error estimate Lemma 6.2.9 with the approximation property of the nodal interpolation operator $\mathcal{I}_h^I$ to show convergence of the modified discrete solution at the optimal linear rate if the exact solution u is in H^2.

Lemma 6.2.12. *Let $(T_h)_{0<h\leq 1}$ be shape-regular and assume that $u \in H^2$. Then there is a positive constant C which does not depend on h such that it holds*

$$\|u - u_h\|_V \leq Ch \qquad \text{for } h \in (0,1]. \tag{6.15}$$

Proof. Let $h \in (0,1]$ and use the symbol $\lesssim$ for the relation $\lesssim_{\neg h}$. Denote by $\mathcal{C}_h : L^2 \to V_h$ the composed Clément interpolation operator as in the proof of Lemma 6.1.26. Then, by Lemma 6.2.9, we have:

$$\begin{aligned} \|u - u_h\|_V \lesssim\ & \|u - \mathcal{C}_h u\|_V \\ & + \sup_{w_h \in V_h} \frac{1}{\|w_h\|_V} \int_\Omega \nabla(u - \mathcal{C}_h u) \cdot \nabla w_h \,\mathrm{d}x \\ & + \sup_{w_h \in V_h} \frac{1}{\|w_h\|_V} \int_I \big(f(\cdot,[u]) - f(\cdot,[\mathcal{C}_h u])\big)[w_h] \,\mathrm{d}\sigma \\ & + \sup_{w_h \in V_h} \frac{1}{\|w_h\|_V} \int_I f(\cdot,[\mathcal{C}_h u])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[\mathcal{C}_h u])[w_h]\big) \,\mathrm{d}\sigma \\ =:\ & (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}) + (\mathrm{IV}). \end{aligned}$$

By the boundedness of u (Theorem 4.2.7), the L^∞-stability of the Clément operator and Assumption 4.1.2 we can argue for example as in the proof of Lemma 6.1.4 that it holds

$$(\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}) \lesssim h.$$

Let us now investigate the term (IV). To this end, define $v_h := \mathcal{C}_h u$ and let $q := d/2+1$. Then it follows by the Hölder inequality:

$$
\begin{aligned}
&\int_I f(\cdot,[v_h])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[v_h])[w_h]\big)\,\mathrm{d}\sigma \\
&\le \sum_{F\in\mathcal{F}_h^I} \|f(\cdot,[v_h])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[v_h])[w_h]\big)\|_{0,1;F} \\
&\le \sum_{F\in\mathcal{F}_h^I} \sigma(F)^{1-1/q}\|f(\cdot,[v_h])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[v_h])[w_h]\big)\|_{0,q;F}.
\end{aligned}
$$

By the choice of q, the assumptions of the local error estimate for the nodal interpolation operator Lemma 6.1.21 are satisfied and we obtain

$$
\|f(\cdot,[v_h])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[v_h])[w_h]\big)\|_{0,q;F} \lesssim h_F^2 |f(\cdot,[v_h])[w_h]|_{2,q;F}. \tag{6.16}
$$

Now fix some face $F \in \mathcal{F}_h^I$. For a sufficiently smooth real-valued function v defined on I we denote by $\nabla_F v := \partial_{x_F} v \in \mathbb{R}^d$ the tangential gradient and by $\nabla_F^2 v := \partial_{x_F}^2 v \in \mathbb{R}^{d\times d}$ the tangential Hessian.

For ease of notation we omit the argument $(\cdot,[v_h])$ at every evaluation of f and its derivatives, e.g. $f := f(\cdot,[v_h])$. A straight-forward application of the chain-rule and the Leibniz-rule gives

$$
\begin{aligned}
&\nabla_F^2\big(f(\cdot,[v_h])[w_h]\big) \\
&= (\partial_{x_F}^2 f)[w_h] + (\partial_z\partial_F f)[\nabla_F^\top v_h][w_h] + (\partial_{x_F} f)[\nabla_F^\top w_h] \\
&\quad + [\nabla_F v_h][w_h](\partial_z\partial_{x_F}^\top f) + [\nabla_F v_h][w_h](\partial_z^2 f)[\nabla_F^\top v_h]) + [\nabla_F v_h][\nabla_F^\top w_h](\partial_z f) \\
&\quad + [\nabla_F w_h](\partial_{x_F}^\top f) + [\nabla_F w_h](\partial_z f)[\nabla_F^\top v_h].
\end{aligned}
$$

Since $v_h = \mathcal{C}_h u$ is bounded on Ω, it follows from Assumption 6.2.10:

$$
\begin{aligned}
&|f(\cdot,[v_h])[w_h]|_{2,q;F} \\
&\lesssim \|[w_h]\|_{0,q;F} + \|[\nabla_F^\top v_h][w_h]\|_{0,q;F} + \|[\nabla_F^\top w_h]\|_{0,q;F} \\
&\quad + \|[\nabla_F v_h][w_h]\|_{0,q;F} + \|[\nabla_F v_h][w_h][\nabla_F^\top v_h]\|_{0,q;F} + \|[\nabla_F v_h][\nabla_F^\top w_h]\|_{0,q;F} \\
&\quad + \|[\nabla_F w_h]\|_{0,q;F} + \|[\nabla_F w_h][\nabla_F^\top v_h]\|_{0,q;F}.
\end{aligned} \tag{6.17}
$$

Note that each of the functions on the right hand side is at most linear. Since $\|\cdot\|_{0,q;\hat{F}}$ and $\|\cdot\|_{0,1;\hat{F}}$ are equivalent norms on the finite-dimensional space $\mathbb{P}_1(\hat{F})$, where $\hat{F}$ denotes the reference $(d-1)$-simplex, the following estimate follows from the shape-regularity of $(\mathcal{T}_h)_{0<h\le 1}$:

$$
\|v\|_{0,q;F} \lesssim \sigma(F)^{1/q-1}\|v\|_{0,1;F} \qquad \text{for all } p_h \in \mathbb{P}^1(F). \tag{6.18}
$$

We thus obtain from (6.16), (6.17) and (6.18):

$$
\begin{aligned}
&\sigma(F)^{1-1/q}\|f(\cdot,[v_h])[w_h]-I_h\big(f(\cdot,[v_h])[w_h]\big)\|_{0,q;F}\\
&\lesssim h_F^2\,\big(\|[w_h]\|_{0,1;F} \quad + \|[\nabla_F^\top v_h][w_h]\|_{0,1;F} \quad + \|[\nabla_F^\top w_h]\|_{0,1;F}\\
&\qquad + \|[\nabla_F v_h][w_h]\|_{0,1;F} + \|[\nabla_F v_h][w_h][\nabla_F^\top v_h]\|_{0,1;F} + \|[\nabla_F v_h][\nabla_F^\top w_h]\|_{0,1;F}\\
&\qquad + \|[\nabla_F w_h]\|_{0,1;F} \quad + \|[\nabla_F w_h][\nabla_F^\top v_h]\|_{0,1;F}\big).
\end{aligned}
$$

Now we apply Hölder's inequality to the summands on the right hand side and obtain:

$$
\begin{aligned}
&\sigma(F)^{1-1/p}\|f(\cdot,[v_h])[w_h]-I_h\big(f(\cdot,[v_h])[w_h]\big)\|_{0,p;F}\\
&\lesssim h_F^2\,\big(\|w_h\|_{0,1;F} \quad + \|\nabla_F v_h\|_{0,2;F}\,\|w_h\|_{0,2;F} + \|\nabla_F w_h\|_{0,1;F}\\
&\qquad + \|\nabla_F v_h\|_{0,2;F}\,\|w_h\|_{0,2;F} + \|\nabla_F v_h\|_{0,4;F}^2\,\|w_h\|_{0,2;F} + \|\nabla_F v_h\|_{0,2;F}\,\|\nabla_F w_h\|_{0,2;F}\big)\\
&=: \text{(a)} + \text{(b)} + \ldots + \text{(h)}.
\end{aligned}
\tag{6.19}
$$

Denote by T_F the union of the elements in $\mathcal{T}_h$ which are adjacent to T_F. Then from the shape-regularity of $(\mathcal{T}_h)_{0<h\leq 1}$ it follows

$$
\mu(T_F) \lesssim h_F\sigma(F) \lesssim \mu(T_F). \tag{6.20}
$$

Now we consider the sum over all $F \in \mathcal{F}_h^I$ for each of the summands (a)–(f) in the estimate (6.19) separately. The basic idea is to use that $|\nabla_F v_h|$ is constant and satisfies $|\nabla_F v_h| \leq |\nabla v_h|$ on F.

(a) is straightforward:

$$
\sum_{F\in\mathcal{F}_h^I} h_F^2\|w_h\|_{0,1;F} \lesssim h^2\|w_h\|_{0,2;I} \lesssim h^2\|w_h\|_{1,2;\Omega}.
$$

(b): We use the Cauchy-Schwarz inequality for sums, (6.20) and the trace theorem.

$$
\begin{aligned}
&\sum_{F\in\mathcal{F}_h^I} h_F^2\|\nabla_F v_h\|_{0,2;F}\|w_h\|_{0,2;F}\\
&\leq h^{3/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\|\nabla_F v_h\|_{0,2;F}^2\Big)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I}\|w_h\|_{0,2;F}^2\Big)^{1/2} && \text{(CS)}\\
&= h^{3/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\frac{\sigma(F)}{\mu(T_F)}\|\nabla_F v_h\|_{0,2;T_F}^2\Big)^{1/2}\|w_h\|_{0,2;I}\\
&\lesssim h^{3/2}\Big(\sum_{T\in\mathcal{T}_h}\|\nabla v_h\|_{0,2;T}^2\Big)^{1/2}\|w_h\|_{0,2;I} && \text{(Eq. (6.20))}\\
&\lesssim h^{3/2}\|\nabla v_h\|_{0,2;\Omega}\,\|w_h\|_{1,2;\Omega}. && \text{(Trace-Theorem)}
\end{aligned}
$$

(c) is obtained from Cauchy–Schwarz for integrals and sums and (6.20):

$$\begin{aligned}
\sum_{F\in\mathcal{F}_h^I} h_F^2\|\nabla_F w_h\|_{0,1;F} &\le \sum_{F\in\mathcal{F}_h^I} h_F^2\sigma(F)^{1/2}\|\nabla_F w_h\|_{0;2;F} && \text{(CS)}\\
&\le h^{3/2}\Big(\sum_{F\in\mathcal{F}_h^I}\sigma(F)\Big)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\|\nabla_F w_h\|_{0,2;F}^2\Big)^{1/2} && \text{(CS)}\\
&= h^{3/2}\sigma(I)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\frac{\sigma(F)}{\mu(T_F)}\|\nabla w_h\|_{0,2;T_F}^2\Big)^{1/2}\\
&\lesssim h^{3/2}\Big(\sum_{T\in\mathcal{T}_h}\|\nabla w_h\|_{0,2;T}^2\Big)^{1/2} && \text{(Eq. (6.20))}\\
&\le h^{3/2}\|w_h\|_{1,2;\Omega}
\end{aligned}$$

(d) is estimated with Cauchy–Schwarz and the same technique as for (c):

$$\begin{aligned}
&\sum_{F\in\mathcal{F}_h^I} h_F^2\|\nabla_F v_h\|_{0,2;F}\|w_h\|_{0,2;F}\\
&\le h^{3/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\|\nabla_F v_h\|_{0,2;F}^2\Big)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I}\|w_h\|_{0,2;F}^2\Big)^{1/2} && \text{(CS)}\\
&\lesssim h^{3/2}\|\nabla v_h\|_{0,2;\Omega}\|w_h\|_{1,2;\Omega}. && \text{(as above)}
\end{aligned}$$

(e): Use Cauchy–Schwarz, (6.20) and the trace theorem. However, we are left with the $W^{1,4}$-norm of v_h in the upper bound:

$$\begin{aligned}
&\sum_{F\in\mathcal{F}_h^I} h_F^2\|\nabla_F v_h\|_{0,4;F}^2\|w_h\|_{0,2;F}\\
&\le h^{3/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\|\nabla_F v_h\|_{0,4;F}^4\Big)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I}\|w_h\|_{0,2;F}^2\Big)^{1/2} && \text{(CS)}\\
&= h^{3/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\frac{\sigma(F)}{\mu(T_F)}\|\nabla v_h\|_{0,4;T_F}^4\Big)^{1/2}\|w_h\|_{0,2;I}\\
&\lesssim h^{3/2}\Big(\sum_{T\in\mathcal{T}_h}\|\nabla v_h\|_{0,4;T}^4\Big)^{1/2}\|w_h\|_{0,2;I} && \text{(Eq. (6.20))}\\
&= h^{3/2}\|\nabla v_h\|_{0,4;\Omega}^2\|w_h\|_{1,2;\Omega} && \text{(Trace Theorem)}
\end{aligned}$$

(f) is treated similarly. Note that the upper bound we obtain is of order h which is weaker estimate than for the other terms:

$$
\begin{aligned}
&\sum_{F\in\mathcal{F}_h^I} h_F^2 \|\nabla_F v_h\|_{0,2;F}\|\nabla_F w_h\|_{0,2;F} \\
&\le h\Big(\sum_{F\in\mathcal{F}_h^I} h_F\|\nabla_F v_h\|_{0,2;F}^2\Big)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\|\nabla_F w_h\|_{0,2;F}^2\Big)^{1/2} && \text{(CS)} \\
&\le h\Big(\sum_{F\in\mathcal{F}_h^I} h_F\frac{\sigma(F)}{\mu(T_F)}\|\nabla v_h\|_{0,2;T_F}^2\Big)^{1/2}\Big(\sum_{F\in\mathcal{F}_h^I} h_F\frac{\sigma(F)}{\mu(T_F)}\|\nabla w_h\|_{0,2;T_F}^2\Big)^{1/2} \\
&\lesssim h\Big(\sum_{T\in\mathcal{T}_h}\|\nabla v_h\|_{0,2;T}^2\Big)^{1/2}\Big(\sum_{T\in\mathcal{T}_h}\|\nabla w_h\|_{0,2;T}^2\Big)^{1/2} && \text{(Eq. (6.20))} \\
&\le h\|\nabla v_h\|_{0,2;\Omega}\|w_h\|_{1,2;\Omega}.
\end{aligned}
$$

Condensing the above estimates yields the following:

$$
\begin{aligned}
&\sum_{F\in\mathcal{F}_h^I} \sigma(F)^{1-1/p}\|f(\cdot,[v_h])[w_h] - \mathcal{I}_h^I\big(f(\cdot,[v_h])[w_h]\big)\|_{0,p;F} \\
&\lesssim h\|w_h\|_{1,2;\Omega}\Big(h \quad + h^{1/2}\|\nabla v_h\|_{0,2;\Omega} + h^{1/2} + \\
&\qquad h^{1/2}\|\nabla v_h\|_{0,2;\Omega} + h^{1/2}\|\nabla v_h\|_{0,4;\Omega}^2 + \|\nabla v_h\|_{0,2;\Omega}\Big).
\end{aligned}
$$

By the stability of the Clément interpolation operator with respect to L^∞ (Lemma 6.1.25) and $W^{1,4}$ ([41, Lemma 1.127]) and the boundesness of u (Theorem 4.2.7) we have

$$
\|\nabla v_h\|_{0,2;\Omega} \lesssim \|\nabla u\|_{0,2;\Omega} \lesssim 1 \qquad \text{and} \qquad \|\nabla v_h\|_{0,4;\Omega} \lesssim \|\nabla u\|_{0,4;\Omega}.
$$

Thus we have shown that (IV) $\lesssim h$. This finishes the proof. □

Remark 6.2.13. *For $d \le 4$ it holds $H^2 \hookrightarrow W^{1,4}$ by Sobolev embedding because $2-d/2 \ge 1-d/4$ is equivalent to $d \le 4$. Therefore, the assumption $u \in W^{1,4}$ in Lemma 6.2.12 is redundant with $u \in H^2$ and can be ommit in this case.*

6.2.5 Discrete Comparision Principle

As mentioned in the introduction of the section, it is possible to prove the discrete counterpart of the comparison principle (Theorem 4.5.2) for the modified discrete equation (6.10) when the underlying triangulation is of non-negative type.

Definition 6.2.14. *$\widetilde{u}_h \in W_h$ is called a* discrete subsolution (supersolution) *if the following conditions are satisfied:*

1. *$\widetilde{u}_{h,2} \le 0$ ($\widetilde{u}_{h,2} \ge 0$) holds on Γ_2.*
2. *For all $v_h \in V_h$ satisfying $v_h \ge 0$ on Ω it holds*

$$
\langle A_h(u_h), v_h\rangle \le (\ge) 0. \tag{6.21}
$$

In the remainder of this section we will assume that $\mathcal{T}_h$ is of *non-negative type*. To this end let us first give the precise definition of this property:

Definition 6.2.15. *1. A d-simplex $\triangle$ is* of non-negative type *if the nodal shape functions $\varphi_{\triangle,i} =: \varphi_i$ for $i \in \mathcal{N}_\triangle$ (see Definition 6.1.12) satisfy*

$$\nabla\varphi_i \cdot \nabla\varphi_j \leq 0 \qquad \text{for } i,j \in \mathcal{N}_\triangle \text{ with } i \neq j.$$

2. A triangulation $\mathcal{T}$ is of non-negative type *if all elements $T \in \mathcal{T}$ are of non-negative type.*

Remark 6.2.16. *For $d = 2$ a triangle is of non-negative type if and only if all interior angles are less or equal than $\pi/2$. [20]*

Before stating and proving the discrete comparison principle we will record two remarks which will be used both in the proof of the discrete comparision principle and of the uniform L^∞-bound for the modified discrete system, see Section 6.2.6.

The first remark concerns the divergence term: In Chapter 4 we have used the fact that $\nabla v \cdot \nabla(v_+) = |\nabla(v_+)|^2$ holds for any sufficiently smooth function v. Since in contrast to V, the space V_h is not closed under taking the positive part, we need to estimate $\nabla v_h \cdot \nabla(\mathcal{I}_h(v_{h,+}))$ for $v_h \in V_h$ instead. For linear elements this is possible on non-negative triangulations, see the following remark.

Remark 6.2.17. *Let $\triangle$ be a d-simplex of non-negative type and denote by $\mathcal{I} : \mathcal{C}^0(\triangle) \to \mathbb{P}_1(\triangle)$ the nodal interpolation operator. Then it holds*

$$\nabla v \cdot \nabla\big(\mathcal{I}(v_+)\big) \geq \big|\nabla\big(\mathcal{I}(v_+)\big)\big|^2 \qquad \text{for all } v \in \mathbb{P}_1(\triangle).$$

Proof. Let $\varphi_i := \varphi_{\triangle,i} \in \mathbb{P}_1(\triangle)$ denote the nodal shape function satisfying $\varphi_i(x_j) = \delta_{ij}$ for $i,j \in \mathcal{N}_\triangle$. Then it holds $v = \sum_i v_i\varphi_i$ and $\mathcal{I}(v_+) = \sum_i \widetilde{v}_i\varphi_i$ with $v_i = v(x_i)$ and $\widetilde{v}_i = (v_i)_+$, where we use the abbreviation $\sum_i$ for a sum over $i \in \mathcal{N}_\triangle$. Let us define the sets

$$J := \{i \mid v_i > 0\} \qquad \text{and} \qquad J^{\mathrm{c}} := \{i \mid v_i \leq 0\}$$

and write $a_{ij} := \nabla\varphi_i \cdot \nabla\varphi_j$ for $i,j \in \mathcal{N}_\triangle$. Clearly, it holds $\widetilde{v}_i = 0$ for $i \in J^{\mathrm{c}}$. Therefore, we have

$$\begin{aligned}
\nabla v \cdot \nabla\big(\mathcal{I}(v_+)\big) &= \sum_{i,j} v_i\widetilde{v}_j\nabla\varphi_i\nabla\varphi_j \\
&= \sum_i \sum_{j\in J} v_i v_j a_{ij} \\
&= \sum_{i\in J^{\mathrm{c}}} \sum_{j\in J} v_i v_j a_{ij} + \sum_{i,j\in J} v_i v_j a_{ij}.
\end{aligned}$$

Since for $i \in J^{\mathrm{c}}$ and $j \in J$ it holds $v_i \le 0$, $v_j \ge 0$ and $a_{ij} \le 0$, every summand in the first sum is non-negative. By again using $\widetilde{v}_i = 0$ for $i \in J^{\mathrm{c}}$ it follows:

$$\nabla v \cdot \nabla\big(\mathcal{I}(v_+)\big) \ge \sum_{i,j\in J} v_i v_j a_{ij} = \sum_{i,j} \widetilde{v}_i \widetilde{v}_j a_{ij} = \big|\nabla\big(\mathcal{I}(v_+)\big)\big|^2.$$

This finishes the proof. □

The second remark concerns the interface nonlinearity. In the continuous proof we used the estimate $f(\cdot,[v])[v_+] \ge [v_+]^2$ which holds σ-almost everywhere on I for any $v \in V$. For the discrete proof one again needs to replace v_+ with its nodal interpolant $\mathcal{I}_h^I(v_+)$. The required estimate is provided in Remark 6.2.18.

Remark 6.2.18. *For $v_h, \widetilde{v}_h \in W_h$ let $w_h := \mathcal{I}_h\big((v_h - \widetilde{v}_h)_+\big)$. Then it holds*

$$\mathcal{I}_h^I\Big(\big(f(\cdot,[v_h]) - f(\cdot,[\widetilde{v}_h])\big)[w_h]\Big) \ge M_1^{-1}[w_h]^2 \qquad \text{on } I.$$

Proof. Let $F \in \mathcal{F}_h^I$ be arbitrary and define $v := v_h|_F$, $\widetilde{v} := \widetilde{v}_h|_F$ and $w := w_h|_F = \mathcal{I}^F((v - \widetilde{v})_+)$. From Assumption 4.1.2, Remark 4.2.8 and the definition of w it follows

$$\Big(f\big(x_i,[v(x_i)]\big) - f\big(x_i,[\widetilde{v}(x_i)]\big)\Big)[w(x_i)] \ge M_1^{-1}[w(x_i)]^2$$

for all $i \in \mathcal{N}_F$. Therefore, Remark 6.2.2 implies

$$\mathcal{I}_h^I\Big(\big(f(\cdot,[v]) - f(\cdot,[\widetilde{v}])\big)[w]\Big) \ge M_1^{-1}\mathcal{I}_h^I\big([w]^2\big). \tag{6.22}$$

Since w is linear, $[w]^2$ is convex. Therefore, it follows again by Remark 6.2.2 that

$$\mathcal{I}_h^I\big([w]^2\big) \ge [w]^2. \tag{6.23}$$

Combining (6.22) and (6.23) completes the proof. □

Note that in the situation of Remark 6.2.18 in general it does not hold

$$\big(f(\cdot,[v_h]) - f(\cdot,[\widetilde{v}_h])\big)[w_h] \ge M_1^{-1}[w_h]^2$$

on I, for example, when $\widetilde{v}_h = 0$ and $[v_h]$ has a sign-change. This is the basic reason why the discrete formulation had to be changed in order to prove the comparison principle and the L^∞-bound.

Lemma 6.2.19. *Let $\underline{u}_h$ be a discrete subsolution and $\overline{u}_h$ be a discrete supersolution. Then it holds*

$$\underline{u}_h \le u_h \le \overline{u}_h \qquad \text{on } \Omega. \tag{6.24}$$

Proof. Let $v_h := \mathcal{I}_h\big((\underline{u}_h - u_h)_+\big)$.

Since $u_h \in V_h$ and $\underline{u}_h$ is a discrete subsolution, it holds $(\underline{u}_{h,2} - u_{h,2})_+ = (\underline{u}_{h,2})_+ = 0$ on Γ_2. By the monotonicity of $\mathcal{I}_h$ it follows $v_{h,2} = 0$ on Γ_2 and thus $v_h \in V_h$. As a consequence, we can plug in v_h in the defining equations (6.21) and (6.2.4) for $\underline{u}_h$ and u_h, respectively, to obtain

$$\int_\Omega \kappa\nabla(\underline{u}_h - u_h)\cdot\nabla v_h \,\mathrm{d}x + \int_I \mathcal{I}_h^I\Big(\big(f(\cdot,[\underline{u}_h]) - f(\cdot,[u_h])\big)[v_h]\Big)\,\mathrm{d}\sigma \le 0. \tag{6.25}$$

Then it follows from Assumption 4.1.2, Remark 6.2.17 and Remark 6.2.18 :

$$\begin{aligned} 0 &\ge \int_\Omega \kappa\nabla(\underline{u}_h - u_h)\cdot\nabla v_h \,\mathrm{d}x + \int_I \mathcal{I}_h^I\Big(\big(f(\cdot,[\underline{u}_h]) - f(\cdot,[u_h])\big)[v_h]\Big)\,\mathrm{d}\sigma \\ &\ge M_1^{-1}\Big(\int_\Omega |\nabla v_h|^2\,\mathrm{d}x + \int_I [v_h]^2\,\mathrm{d}\sigma\Big) \end{aligned} \tag{6.26}$$

Thus Lemma 4.2.3 implies $v_h = 0$ which is only possibe if $\underline{u}_h \le u_h$.

The estimate $u_h \le \overline{u}_h$ is shown analogously and thus the proof is finished. □

6.2.6 L^∞ bound

Our initial motivation to introduce the modified discrete formulation (6.10) was to make the Stampacchia truncation method work in the discrete case. In this section we show that in fact we can prove a uniform L^∞-bound for the modified discrete solutions on non-negative triangulations. The proof basically follows the continuous proof with a slight modification which was inspired by the article [20].

Lemma 6.2.20. *There is a positive constant C which does not depend on $\mathcal{T}_h$ such that it holds*

$$\|u_h\|_{0,\infty;\Omega} \le C$$

whenever $\mathcal{T}_h$ is of non-negative type.

In the proof we will use the following basic result from the article [20]:

Lemma 6.2.21. *[20, Lemma 1] Let $\triangle$ be a d-simplex and $q \in [1,\infty)$. Then there exists a positive constant $C = C(q)$ only depending on q and not on $\triangle$ such that it holds*

$$\|v\|_{0,q;\triangle}^q \ge C\mu(\triangle)\sum_{i\in\mathcal{N}_\triangle} |v(x_i)|^q \qquad \text{for all } v \in \mathbb{P}_1(\triangle).$$

Now we present the proof of Lemma 6.2.20:

Proof of Lemma 6.2.20. Let $\lesssim$ denote the relation $\lesssim_{\neg\mathcal{T}_h}$ and for $k \ge 0$ define

$$v_h := \mathcal{I}_h\big((u_h - k)_+\big).$$

Since it holds $[u_h] = [u_h - k]$ on I, it follows from Lemma 4.2.3, Remark 6.2.17, Remark 6.2.18, Assumption 4.1.2 and (6.10):

$$\begin{aligned}\|v_h\|^2_{1,2;\Omega} &\lesssim \int_\Omega |\nabla v_h|^2 \,\mathrm{d}x + \int_I [v_h]^2 \,\mathrm{d}\sigma \\ &\lesssim \int_\Omega \kappa \nabla u_h \cdot \nabla v_h \,\mathrm{d}x + \int_I \mathcal{I}_h\big(f(\cdot,[u_h])[v_h]\big)\,\mathrm{d}\sigma \\ &= G(v_h) \lesssim \|v_h\|_{1,1;\Omega}.\end{aligned} \tag{6.27}$$

Now define $A(k) := \{x \in \Omega \mid v_h > 0\}$ and $\varphi(k) := \mu(A(k))$. Then it holds $v_h = 0$ on $\Omega \setminus A(k)$ and thus it follows from Hölder's inequality:

$$\|v_h\|_{1,1;\Omega} \lesssim \varphi(k)^{1/2} \|v_h\|_{1,2;\Omega}.$$

Thus from (6.27) it follows

$$\|v_h\|_{1,2;\Omega} \lesssim \varphi(k)^{1/2}.$$

Now define

$$\mathcal{N}_h^k := \{i \in \mathcal{N}_h \mid N_i(v_h) > 0\} \qquad \text{and} \qquad \mathcal{T}_h^k := \{T \in \mathcal{T}_h \mid v_h > 0 \text{ on } T^\circ\}.$$

and fix some $q \in (2, 2^*)$, where 2^* denotes the critical Sobolev exponent. Then it follows from Sobolev embedding, the definition of $\mathcal{T}_h^k$ and Lemma 6.2.21:

$$\begin{aligned}\|v_h\|^q_{1,2;\Omega} &\gtrsim \|v_h\|^q_{0,q;\Omega} \\ &= \int_\Omega |v_h|^q \,\mathrm{d}x \\ &= \sum_{T \in \mathcal{T}_h^k} \int_T |v_h|^q \,\mathrm{d}x \\ &\gtrsim \sum_{T \in \mathcal{T}_h^k} \mu(T) \sum_{i \in \mathcal{N}_{h,T}} |v_h(x_i)|^q.\end{aligned}$$

Now let $\widetilde{k} > k$. Using the definition of $\mathcal{T}_h^k$, the sums can be rearranged. Since $\widetilde{k} > k$, the inclusion $A(\widetilde{k}) \subset A(k)$ holds and finally, by the definition of v_h and φ we obtain the following:

$$\begin{aligned}\|v_h\|^q_{1,2;\Omega} &\gtrsim \sum_{i \in \mathcal{N}_h^k} |v_h(x_i)|^q \mu(\omega_i) \\ &\gtrsim (\widetilde{k} - k)^q \sum_{i \in \mathcal{N}_h^{\widetilde{k}}} \mu(\omega_i) \\ &= (\widetilde{k} - k)^q \varphi(\widetilde{k}).\end{aligned}$$

Collecting the estimates yields

$$\varphi(\widetilde{k}) \lesssim \frac{\varphi(k)^{q/2}}{(\widetilde{k}-k)^q}.$$

From Lemma 4.2.9 it follows that there exists $0 < k_0 \lesssim 1$ such that it holds $\varphi(k_0) = 0$. This implies $u_h \leq k_0 \lesssim 1$. As we analogously find some $0 < k_1 \lesssim 1$ such that $-u_h \leq k_1$ holds, the proof is complete. □

7 Numerical Results

In this chapter we present the methods that we used to solve Problem 4.1.1, Problem 3.3.1 and Problem 3.4.3 numerically. The numerical results have been produced by our master student Fabian Castelli as a component of his master thesis [17]. In the code, the open source finite elements framework deal.II was used [10] and the graphics were produced with the open source visualization tool paraview [9].

The chapter is devided into three parts: In Section 7.1 we present the specific meshes, function spaces and degrees of freedom used in the simulations, in Section 7.2 we discuss the numerical results of the finite element discretization of the elliptic problems from Chapter 4 which has been discussed in Section 6.1. Finally, in Section 7.3 we present the numerical solution of the time-dependent, fully coupled systems Problem 3.3.1 and Problem 3.4.3.

7.1 Preliminaries

Since the deal.ii framework works with quadrilateral and hexahedral meshes, our code is *not* a straight forward implementation of the finite element method on simplicial meshes presented in Section 6.1.4. We will therefore explain the mathematical setting which is realized by our code. Even though it is written for arbitrary $d \in \{1, 2, 3\}$, we will restrict our presentation to the two-dimensional case.

7.1.1 Quadrilateral Meshes

In this section we will define precisely the structure and properties of the meshes implemented in deal.ii. Basically, for $d = 2$, the framework works on *generalized* quadrilateral meshes with at most one *hanging node* per edge. Here, the term *generalized quadrilateral* means that the elements in these meshes are transformed copies of the reference quadrilateral $\hat{\square} := [0, 1]^2$. Let us make the following definition:

Definition 7.1.1.

1. *The set of vertices of* $\hat{\square}$ *is denoted by* $\hat{\mathcal{V}}$.
2. *The set of edges of* $\hat{\square}$ *is denoted by* $\hat{\mathcal{E}}$.
3. *The* half-edges *of* $\hat{\square}$ *are*

$$\begin{aligned} &\{0\} \times [0, 1/2], \quad &&\{0\} \times [1/2, 1], \quad &&[0, 1/2] \times \{0\}, \quad &&[1/2, 0] \times \{0\}, \\ &\{1\} \times [0, 1/2], \quad &&\{1\} \times [1/2, 1], \quad &&[0, 1/2] \times \{1\}, \quad &&[1/2, 0] \times \{1\}. \end{aligned}$$

The set of half-edges of $\hat{T}$ *is denoted by* $\hat{\mathcal{E}}_{1/2}$.

In the theoretical setting in Section 6.1.3 the elements in the meshes were copies of the reference Simplex $\hat{\triangle}$ under affine linear injective maps $F : \hat{\triangle} \to \mathbb{R}^2$. This class of transformations is not suitable for quadrilateral meshes though, since the image of the *unit square* $\hat{\square}$ under such a map is always a parallelogram. In fact, for arbitrary quadrilaterals, it is necessary to use injective *bilinear maps* $F \in \mathbb{Q}_1^2(\hat{\square})$.

However, for domains with curved boundaries, in order to preserve the optimal convergence rates of the finite elements method, it is necessary to allow an even larger class of transformations. In our particular examples it will be injective mappings from $\mathbb{Q}_q^2(\hat{\square})$ with sufficiently large polynomial degree $q \in \mathbb{N}$. For the abstract presentation of the method it is convenient to assume that we have chosen a set of transformations $\mathcal{M}$ which satisfies

$$\mathcal{M} \subset \{F : \hat{\square} \to \mathbb{R}^2 \mid F : \hat{\square} \to F(\hat{\square}) \text{ is a } \mathcal{C}^1\text{-diffeomorphism}\}.$$

Then we define a *generalized quadrilateral* as the image of the unit square under such a transformation:

Definition 7.1.2 (Generalized quadrilaterals)**.**

1. *A generalized quadrilateral is a couple* $(\square, F)$ *consisting of a closed subset* $\square \subset \mathbb{R}^2$ *and a mapping* $F_\square := F \in \mathcal{M}$ *such that it holds* $\square = F(\hat{\square})$.

   For a simpler notation we will omit the mapping F *when it is appropriate.*

2. *The set of vertices, edges and half-edges of a generalized quadrilateral* $\square$ *are*

$$\mathcal{V}_\square := F_\square(\hat{\mathcal{V}}), \qquad \mathcal{E}_\square := F_\square(\hat{\mathcal{E}}) \qquad \text{and} \qquad \mathcal{E}_{1/2,\square} := F_\square(\hat{\mathcal{E}}_{1/2}),$$

   respectively.

   Note that this definition depends on the mapping $F_\square$ *in general.*

Now, quadrilateral *meshes* are defined as meshes which consist of generalized quadrilaterals. Note that this notion depends on the choice of the set of admissable transformations $\mathcal{M}$. Let $D \subset \mathbb{R}^2$ be a closed set.

Definition 7.1.3 (Quadrilateral meshes)**.**

1. *A* (generalized) quadrilateral mesh *for* D *is a mesh* $\mathcal{T}$ *for* D *together with a family* $(F_T)_{T\in\mathcal{T}}$ *of mappings* $F_T \in \mathcal{M}$ *for* $T \in \mathcal{T}$, *such that for each* $T \in \mathcal{T}$ *the couple* (T, F_T) *is a generalized quadrilateral.*

2. *The sets of vertices, edges and half-edges in* $\mathcal{T}$ *are*

$$\mathcal{V} := \bigcup_{T\in\mathcal{T}} \mathcal{V}_T, \qquad \mathcal{E} := \bigcup_{T\in\mathcal{T}} \mathcal{E}_T \qquad \text{and} \qquad \mathcal{E}_{1/2} := \bigcup_{T\in\mathcal{T}} \mathcal{E}_{1/2,T}.$$

Now we are able to state the hanging vertex condition which is satisfied by the meshes used within the deal.ii framework.

Definition 7.1.4 (Hanging Vertex Condition). *Let $\mathcal{T}$ be a quadrilateral mesh. Then $\mathcal{T}$ satisfies the* hanging vertex condition *if for all $T, S \in \mathcal{T}$ one of the following conditions is satisfied:*

1. $T \cap S = \emptyset$
2. $T \cap S$ *is a vertex of both T and S.*
3. $T \cap S$ *is an edge of both T and S.*
4. $T \cap S$ *is an edge of T and a half-edge of S or vice versa.*
5. $T = S$

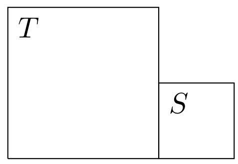

Figure 7.1: Case 4 of Definition 7.1.4.

Suppose $\mathcal{T}$ is a quadrilateral mesh satisfying the hanging vertex condition. In order to identify the global degrees of freedom for the discrete function spaces on $\mathcal{T}$ which will be introduced in Section 7.1.2, it is important that the mappings F_T and F_S of adjacent elements $T, S \in \mathcal{T}$ fit together in a certain sense which is made precise in Definition 7.1.5. See also Definition 7.1.10 and the following remarks.

For bilinear transformations, that is, $\mathcal{M} \subset \mathbb{Q}_1^d$, Definition 7.1.5 is automatically satisfied since the restriction of both F_T^{-1} and F_S^{-1} to the separating edge $T \cap S$ is linear. For arbitrary $\mathcal{M}$ this is no longer automatically fulfilled, see Fig. 7.3 for such an example.

However, the following $\mathcal{C}^0$-compatibility condition is satisfied by all admissible meshes in the deal.ii framework:

Definition 7.1.5. *Let $\mathcal{T}$ be a generalized quadrilateral mesh satisfying the hanging vertex condition. Then $\mathcal{T}$ is called* quasi-conforming *if for all $T, S \in \mathcal{T}$ the change of coordinates*

$$F_T^{-1} \circ F_S : F_S^{-1}(T \cap S) \to F_T^{-1}(T \cap S)$$

is an affine linear map.

Fig. 7.2 shows and example, where Definition 7.1.5 is satisfied but $\mathcal{M} \not\subset \mathbb{Q}_1^d$.

7.1.2 Function Spaces

Now we introduce the function spaces and the corresponding degrees of freedom which we use in our simulations. Roughly speaking, we use $\mathcal{C}^0$-conforming tensor product elements of polynomial degree $p \in \mathbb{N}$. As degrees of freedom we take the function evaluations at

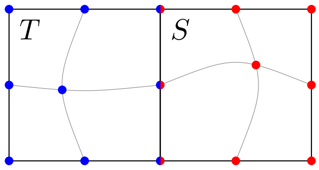

Figure 7.2: Example for an affine linear change of coordinates.

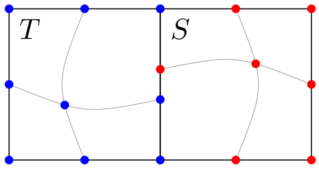

Figure 7.3: Example for non affine linear change of coordinates

the *Lagrange nodes*. The construction is similar to Section 6.1.3: At first, the *reference finite element* is defined, which is then transformed onto each element in the mesh. These *local finite elements* are then pieced together to form the *global finite element*.

Definition 7.1.6 (Reference finite element)**.**

1. *The reference function space is* $\hat{\mathcal{S}} := \mathbb{Q}_p(\hat{\square})$.
2. *Let* $\hat{\mathcal{X}}$ *denote the Lagrange nodes on* $\hat{\square}$, *that is,*

$$\hat{\mathcal{X}} = \{i/p \mid i = 0, \dots, p\}^d.$$

The reference degrees of freedom $\hat{\mathcal{N}}$ *are the function evaluations* $\hat{v} \mapsto \hat{v}(\hat{x})$ *at the Lagrange nodes on* $\hat{\square}$, *that is,*

$$\hat{\mathcal{N}} = \{\hat{v} \mapsto \hat{v}(\hat{x}) \mid \hat{x} \in \hat{\mathcal{X}}\} \subset \hat{\mathcal{S}}'.$$

By these choices, $(\hat{\square}, \hat{\mathcal{S}}, \hat{\mathcal{N}})$ is a finite element in the sense of [14, Chapter 3].
Similar as in Section 6.1.3, for $i \in \hat{\mathcal{N}}$ we set $\hat{N}_i := i$ and, additionally, $\hat{x}_i \in \hat{\square}$ is defined such that $\hat{N}_i(\hat{v}) = \hat{v}(\hat{x}_i)$ holds for all $\hat{v} \in \hat{\mathcal{S}}$.

Definition 7.1.7 (Local objects)**.** *Let* $(\square, F_\square)$ *be a generalized quadrilateral.*

1. *The local function space on* $\square$ *is*

$$\mathcal{S}_\square := \{\hat{v} \circ F_\square^{-1} \mid \hat{v} \in \hat{\mathcal{S}}\}.$$

2. *The local degrees of freedom on* $\square$ *are*

$$\mathcal{N}_\square := \{v \mapsto v(x_\square) \mid x_\square \in \mathcal{X}_\square\} \subset \mathcal{S}_\square',$$

where $\mathcal{X}_\Box := F_\Box(\hat{\mathcal{X}})$ *are the Lagrange nodes on* $\Box$.

Again, for $i \in \mathcal{N}_\Box$ *we define* $N_i := N_{\Box,i} := i$ *and* $x_i := x_{\Box,i} \in \Box$ *by the relation* $N_i(v) = v(x_i)$ *for all* $v \in \mathcal{S}_\Box$.

3. *The local shape functions are the unique functions* $\varphi_i := \varphi_{\Box,i} \in \mathcal{S}_\Box$ *satisfying*

$$N_i(\varphi_j) = \delta_{ij} \qquad \text{for all } i,j \in \mathcal{N}_\Box.$$

Then the triple $(\Box, \mathcal{S}_\Box, \mathcal{N}_\Box)$ is an $\mathcal{M}$-equivalent finite element to the reference element, compare [14, Section 3.4].

Definition 7.1.8 (Global Space)**.** *Let* $\mathcal{T}$ *be a quasi-conforming quadrilateral mesh for* D*. The global* $\mathcal{C}^0$*-conforming space is*

$$\mathcal{S} := \mathcal{S}^{p,0}(\mathcal{T}) := \{v \in \mathcal{C}^0(\overline{D}) \mid v|_T \in \mathcal{S}_T \text{ for all } T \in \mathcal{T}\}.$$

Recall that for the conforming simplicial meshes, the definition of global degrees of freedom was particularly straight forward. We simply collected all local degrees of freedom which yielded a basis of the dual of our global finite element space.

However, when there are elements in the mesh satisfying condition 3 of Definition 7.1.4, there will be so-called hanging nodes, see Definition 7.1.9. As a consequence, the union of all nodal values $\bigcup_{T\in\mathcal{T}} \mathcal{N}_T$ is no longer a linearly independent subset of $\mathcal{S}'$. In order to construct the *global degrees of freedom*, a basis of $\mathcal{S}'$, we remove the nodal values corresponding to hanging nodes from $\bigcup_{T\in\mathcal{T}} \mathcal{N}_T$, see Definition 7.1.10.

Definition 7.1.9 (Hanging nodes)**.** *Let* $\mathcal{T}$ *be a quasi-conforming quadrilateral mesh for* D*.*

1. *The* hanging nodes *on* $T \in \mathcal{T}$ *are*

$$\dot{\mathcal{X}}_T := \bigcup_{S\in\mathcal{T}} (\mathcal{X}_T \cap S) \setminus (\mathcal{X}_S \cap T). \tag{7.1}$$

The situation of nonempty $\mathcal{X}_T$ *is shown in Fig. 7.4 for* $p \in \{1,2\}$.

2. *The (local) degrees of freedom on* T *corresponding to hanging nodes are*

$$\dot{\mathcal{N}}_T := \{N_{T,i} \mid i \in \mathcal{N}_T \text{ and } x_{T,i} \in \dot{\mathcal{X}}_T\}.$$

Definition 7.1.10 (Global Degrees of Freedom)**.** *Let* $\mathcal{T}$ *be a quasi-conforming quadrilateral mesh for* D

1. *As global degrees of freedom on* $\mathcal{S}$ *we choose* $\mathcal{N}$ *defined by*

$$\mathcal{N} := \bigcup_{T\in\mathcal{T}} \mathcal{N}_T \setminus \dot{\mathcal{N}}_T.$$

We canonically extend the notation introduced for the local objects to the global ones: For $i \in \mathcal{N}$ *we define* $N_i := i$ *and we let* $x_i := x_{T,i}$ *if* $T \in \mathcal{T}$ *is such that it holds* $i \in \mathcal{N}_T \subset \mathcal{N}$.

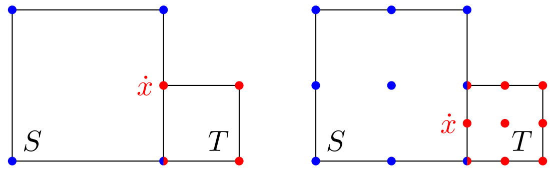

Figure 7.4: Example for hanging nodes $\dot{x}$ for $p = 1$ (left) and $p = 2$ (right).

2. *The global shape functions are the unique functions* $\varphi_j \in \mathcal{S}$ *satisfying*

$$N_i(\varphi_j) = \delta_{ij} \qquad \text{for all } i, j \in \mathcal{N}.$$

From the quasi-conformity of $\mathcal{T}$, see Definition 7.1.5, it follows that $\mathcal{N}$ is in fact a basis of $\mathcal{S}'$. Additionally it holds $\varphi_i|_T = \varphi_{T,i}$ for $i \in \mathcal{N}_T \subset \mathcal{N}$ and $\varphi_i|_T = 0$ for $i \in \mathcal{N} \setminus \mathcal{N}_T$. Such a function is depicted in Fig. 7.5.

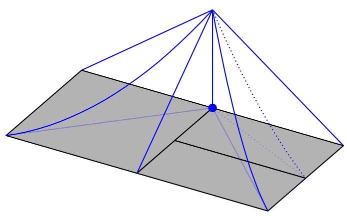

Figure 7.5: Nodal basis function for $p = 1$

Analogue to Definition 6.1.16 we define the subspace of functions vanishing on a part $S \subset \partial D$ of the boundary of D.

Definition 7.1.11. *Let* $S \subset \partial D$ *such that* $\mathcal{T}$ *is conforming with* S *in the sense of Definition 6.1.15. Then, analogously to Definition 6.1.16, we define* $\mathcal{S}_S := \mathcal{S} \cap H^1_S(D)$*. As degrees of freedom for* $\mathcal{S}'_S$ *we choose* $\mathcal{N} \setminus \mathcal{N}_S$*. The corresponding nodal basis is given by the shape functions* φ_i*,* $i \in \mathcal{N} \setminus \mathcal{N}_S$*.*

7.1.3 Broken Discrete Function Spaces

After having presented the general $\mathcal{C}^0$-conforming tensor-product elements on quasi-conforming quadrilateral meshes and having identified the nodal degrees of freedom for these elements, we will now explain how we can use these object in order to construct the concrete discrete spaces which we use in our simulations.

For $i = 1, 2$, let $\mathcal{T}_{h,i}$ be a quasi-conforming quadrilateral mesh for the discrete computational domain $\overline{\Omega}_{h,i} \subset \mathbb{R}^2$ such that $\mathcal{T}_{h,2}$ is conforming with Γ_2. The discrete interface is $I_h := \partial\Omega_{h,1} \cap \partial\Omega_{h,2}$.

Let us adopt the notation introduced in Section 7.1.2 and denote the respective objects with the same letter followed by a subscript h and $i \in \{1, 2\}$. For example, the degrees on freedom on $\mathcal{S}_{h,i} := \mathcal{S}^{p,0}(\mathcal{T}_{h,i})$ are denoted by $\mathcal{N}_{h,i}$.

We let $\mathcal{S}_h := \mathcal{S}_{h,1} \oplus \mathcal{S}_{h,2}$ and, as in Section 6.1.3, we consider $\mathcal{S}'_{h,i}$ as a subspace of $\mathcal{S}'_h$ by applying the projection $(v_1, v_2) \mapsto v_i$ for $i = 1, 2$. Following the naming conventions in Section 6.1.3, we let $\mathcal{N}_h := \mathcal{N}_{h,1} \cup \mathcal{N}_{h,2}$ and canonically extend the mappings

$$N_{h,i} : j \mapsto N_{h,i;j}, \qquad x_{h,i} : j \mapsto x_{h,i;j}, \qquad \text{and} \qquad \varphi_{h,i} : j \mapsto \varphi_{h,i;j}$$

defined on $\mathcal{N}_{h,i}$ for $i = 1, 2$, to $\mathcal{N}_h = \mathcal{N}_{h,1} \cup \mathcal{N}_{h,2}$ in order to obtain mappings $N := N_h$, $x := x_h$ and $\varphi := \varphi_h$ defined on $\mathcal{N}_h$.

Finally we define the spaces $W_h := \mathcal{S}_h$ and $V_h := W_h \cap H^1_{\Gamma_2}$. The degrees of freedom for V_h are denoted by $\mathcal{N}_h^\circ := \mathcal{N}_h \setminus \mathcal{N}_{h,\Gamma_2}$.

7.2 Solving the Elliptic Subproblem

We present the numerical solution of the strongly nonlinear elliptic problems discussed in Chapter 4. Let us briefly recall their formulation:

Problem 4.1.1. *Find $u : \Omega \to \mathbb{R}$ such that the following holds:*

$$\begin{aligned} -\nabla \cdot (\kappa \nabla u) &= G && \text{in } \Omega, \\ \kappa_i \partial_\nu u_i &= f(\cdot, [u]) && \text{on } I, \\ \kappa \partial_\nu u &= 0 && \text{on } \partial\overline{\Omega} \setminus \Gamma_2, \\ u_2 &= 0 && \text{on } \Gamma_2. \end{aligned} \tag{7.2}$$

The data κ, G and f are supposed to satisfy Assumption 4.1.2 and will be explicitly given in the examples in Section 7.2.1 – Section 7.2.2.

In Section 6.1 we analyzed the Galerkin discretization of Problem 4.1.1, see Definition 6.1.1. In a more general form it reads:

Problem 7.2.1. *Find $u_h \in V_h$ such that*

$$\int_{\Omega_h} \kappa \nabla u_h \cdot \nabla v_h \, \mathrm{d}x + \int_{I_h} f(\cdot, [u_h])[v_h] \, \mathrm{d}\sigma = G(v_h) \tag{7.3}$$

holds for all $v_h \in V_h$.

Note that, since the discrete computational domains $\Omega_{h,1}$, $\Omega_{h,2}$ and I_h in general do not coincide with their continuous counterparts Ω_1, Ω_2 and I, the integrals in (7.3) require some explanation. The general technique is to approximate the integrands by suitable functions defined on the *discrete* domains and apply some quadrature rules to evaluate the respective integrals. For our purpose, however, it is sufficient to provide extensions of the functions defined on the continuous domains to the respective discrete domains. These extensions will be explicitly provided in each numerical experiment separately.

In our code, all occuring integrals are in fact approximated by quadrature rules. This issue is called *variational crimes* and has been considered in the literature for a wide range of problems, see for example [14, Chapter 10]. However, for the sake of a simpler presentation we choose to omit the quadrature rules in the formulas and use the same symbol as for the exact integration instead. Since there will be no rigorous proofs in this chapter but only a presentation of the principal numerical method, this should not pose a problem.

In order to solve Problem 7.2.1 numerically, we apply Newton's method [34]:

Problem 7.2.2. *Suppose $u_h^0 \in V_h$ is given. Find $u_h^1 =: N(u_h^0) \in V_h$ such that*

$$\begin{aligned} 0 &= \int_{\Omega_h} \kappa \nabla u_h^1 \cdot \nabla v_h \,\mathrm{d}x + \int_{I_h} \big(f(\cdot,[u_h^0]) + \partial_z f(\cdot,[u_h^0])[u_h^1 - u_h^0]\big)[v_h]\,\mathrm{d}\sigma - G(v_h) \\ &=: \mathrm{Res}(u_h^1, u_h^0; v_h) \end{aligned} \tag{7.4}$$

holds for all $v_h \in V_h$.

For every starting value $u_h^0 \in V_h$, the – possibly finite – sequence of Newton iterations $(u_h^k)_k$ is then given by $u_h^{k+1} := N(u_h^k)$. We terminate the iteration and accept u_h^{k+1} as a valid approximation for u_h when either

$$|u_h^{k+1} - u_h^k|_2 \le \varepsilon_{\mathrm{it}} := 10^{-10} \qquad \text{or} \qquad \big|\big(\mathrm{Res}(u_h^{k+1}, u_h^k; \varphi_i)\big)_i\big|_2 \le \varepsilon_{\mathrm{res}} := 10^{-12}$$

is satisfied. Here, $|\cdot|_2$ is the 2-norm on V_h with respect to the nodal basis $\{\varphi_i \mid i \in \mathcal{N}_h^\circ\}$.

The concrete solution of (7.4) is performed by expanding u_h^1 in terms of the nodal basis functions φ_i, $i \in \mathcal{N}_h^\circ$, that is

$$u_h^1 = \sum_{i \in \mathcal{N}_h^\circ} u_i^1 \varphi_i,$$

where $u_i^1 \in \mathbb{R}$ for $i \in \mathcal{N}_h^\circ$ are the unknown coefficients. Using $v_h = \varphi_j$ for $j \in \mathcal{N}_h^\circ$ in (7.4), we derive the following linear system of equations for the vector of unkowns $u^1 = (u_i^1)_i \in \mathbb{R}^{\mathcal{N}_h^\circ}$:

$$Au^1 = b, \tag{7.5}$$

where $A \in \mathbb{R}^{\mathcal{N}_h^\circ \times \mathcal{N}_h^\circ}$ and $b \in \mathbb{R}^{\mathcal{N}_h^\circ}$ are given by

$$\begin{aligned} A &= \Big(\int_\Omega \kappa \nabla\varphi_j \cdot \nabla\varphi_i \,\mathrm{d}x + \int_{I_h} \partial_z f(\cdot,[u_0])[\varphi_j][\varphi_i]\,\mathrm{d}\sigma\Big)_{ij} \text{ and} \\ b &= \Big(G(\varphi_j) - \int_{I_h} \big(f(\cdot,[u_0]) - \partial_z f(\cdot,[u_0])[u_0]\big)[\varphi_j]\,\mathrm{d}\sigma\Big)_j \end{aligned} \tag{7.6}$$

The matrix A is symmetric and positive definite, compare Lemma 4.2.3, and thus (7.5) can be solved with the method of conjugated gradients [72, Section 9.2]. However, in our code we choose to solve (7.5) by a direct LU-decomposition using the algorithm UMFPACK [28].

7.2.1 Example 1: Radially Symmetric Explicit Solutions

In the case of radially symmetric data we can explicitly calculate the solution of Problem 4.1.1. Let us provide such an example in order to confirm the error estimate in Corollary 6.1.27 and to validate the correctness of our implementation.

The geometry is defined in the following way: The radii $0 < r_1 < r_I < r_2$ are given by $r_1 = 0.1$, $r_I = 0.45$ and $r_2 = 1$. Furthermore, $\Omega_1 := B_{r_I}(0)\setminus\overline{B}_{r_1}(0)$, $\Omega_2 := B_{r_2}(0)\setminus\overline{B}_{r_I}(0)$ and $\Gamma_i := \partial\Omega_i \setminus I$ for $i = 1, 2$, where the interface is $I = \partial\Omega_1 \cap \partial\Omega_2 = \partial B_{r_I}(0)$, see also Fig. 7.6.

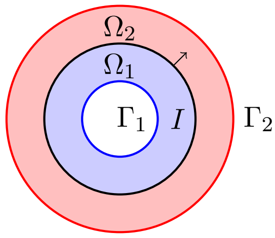

Figure 7.6: Radially symmetric geometry.

We consider the following problem for u:

$$\begin{aligned}
-\Delta u &= 0 && \text{in } \Omega, \\
\partial_\nu u_i &= \sinh([u] - 2) && \text{on } I, \\
\partial_\nu u &= \chi_{\Gamma_1} && \text{on } \partial\overline{\Omega} \setminus \Gamma_2, \\
u_2 &= 0 && \text{on } \Gamma_2.
\end{aligned} \tag{7.7}$$

The exact solution is piecewise smooth and explicitly given by

$$\begin{aligned}
u_1(x) &= -r_1 \ln(|x|_2) + \operatorname{asinh}(r_1/r_2) + 2 && \text{for } x \in \Omega_1, \\
u_2(x) &= -r_1 \ln(|x|_2) && \text{for } x \in \Omega_2,
\end{aligned}$$

see [17]. The problem (7.7) fits into the pattern (7.2) by the definitions

$$\kappa = 1, \qquad f(x, z) = \sinh(z - 2) \qquad \text{and} \qquad G(v) = -\int_{\Gamma_1} v_1 \, \mathrm{d}\sigma.$$

Note that neither Ω_1 nor Ω_2 are polytopes. As a consequence, we must define $\mathcal{M}$ as a proper *superset* of $\{v \in \mathbb{Q}_1^d(\hat{\square}) \mid v \text{ injective}\}$ to maintain the optimal convergence rates of the finite element method for $p \geq 2$. In our simulations we choose $\mathbb{Q}_q^d$-transformations, that is, $\mathcal{M} = \{v \in \mathbb{Q}_q^d(\hat{\square}) \mid v \text{ injective}\}$ where $q = q(p)$ is chosen depending on the polynomial degree of the finite element space according to Table 7.1.

As starting value for the Newton-iterations we simply use $u = 0$.

In Fig. 7.7 the results of the numerical method for (7.7) which has just been described are shown. In these tables, $n_{\text{Dof}} = \dim(V_h)$ denotes the respective number of unknowns

p	q
1	1
2	2
3	2
4	3

Table 7.1: Polynomial degree q of the geometrical transformations.

and n_{it} is the number of Newton-iterations. The experimental orders of convergence of the L^2-error and the H^1-error are calculated with a gliding mean of 2.

The numbers in Fig. 7.7 confirm the error estimate in Corollary 6.1.27: For $p \in \{1, 2, 3, 4\}$ the H^1-error behaves like h^p. Note that this does not follow rigorously from Lemma 6.1.4 since by the approximation of the boundary and the evaluation of the integrals by quadrature rules, the concrete method in our simulations does no longer fit into the framework of Section 6.1. The experimental convergence rate of the L^2-error is h^{p+1} which is not surprising since this is known for *linear* elliptic problems, see [41, Section 2.3.4]. We also observe that the Newton-algorithm sucessfully terminates after the reasonable number of 5-6 iterations.

7.2.2 Example 2: Elliptic Subproblem

Now we want to consider the case when Problem 4.1.1 is equivalent to the elliptic subproblem in Problem 3.4.3 on a somewhat realistic geometry representing a cathode which consists of a single particle.

To this end, let $Q := (0, 1)^2$ and $B := B_{0.4}(0, 0.5)$. Then the electrolyte region is $\Omega_1 := Q \setminus \overline{B}$ and the particle region is $\Omega_2 := Q \cap B$. The respective boundary parts are $\Gamma_i := \partial\Omega_i \cap \partial\Omega$ for $i = 1, 2$. This geometrical situation is outlined in Fig. 7.8.

Motivated by Assumption 3.5.3 and the Butler–Volmer condition (3.15), we use the following parameters:

- $\kappa_1(c_1) = \sqrt{c_1}$,
- $\kappa_2(c_2) = \sqrt{c_2}\sqrt{1 - c_2}$,
- $i_{12}(c_1, c_2, z) = \sqrt{c_1}\sqrt{c_2}\sqrt{1 - c_2}\,\sinh(z + \ln(c_1) - U(c_2))$,
- $U(c_2) = \ln(c_2) + \frac{1}{1-c_2}$.

Additionally, we prescribe the concentrations $c_1 : \Omega_1 \to (0, \infty)$ and $c_2 : \Omega_2 \to (0, 1)$ as the following smooth functions:

$$\begin{aligned} c_1(x) &= 0.5 \sin(2\pi x_1) \sin(3\pi x_2) + 1, \\ c_2(x) &= 0.1 e^{-x_1} \sin(2\pi x_2) + 0.5. \end{aligned} \tag{7.8}$$

These functions are depicted in Fig. 7.9.

	$\lvert\mathcal{T}_h\rvert$	h	n_{DoF}	L^2-error		H^1-error		n_{it}
1	128	3.876e-01	160	3.824e-03	-	6.127e-02	-	6
2	512	2.013e-01	576	9.610e-04	1.99	3.121e-02	0.97	6
3	2048	1.024e-01	2176	2.405e-04	2.00	1.570e-02	0.99	6
4	8192	5.161e-02	8448	6.015e-05	2.00	7.861e-03	1.00	6
5	32768	2.590e-02	33280	1.504e-05	2.00	3.932e-03	1.00	6

	$\lvert\mathcal{T}_h\rvert$	h	n_{DoF}	L^2-error		H^1-error		n_{it}
1	128	3.876e-01	576	6.321e-05	-	4.788e-03	-	6
2	512	2.013e-01	2176	9.315e-06	2.76	1.394e-03	1.78	6
3	2048	1.024e-01	8448	1.232e-06	2.92	3.671e-04	1.92	6
4	8192	5.161e-02	33280	1.566e-07	2.98	9.323e-05	1.98	6
5	32768	2.590e-02	132096	1.966e-08	2.99	2.342e-05	1.99	5

	$\lvert\mathcal{T}_h\rvert$	h	n_{DoF}	L^2-error		H^1-error		n_{it}
1	128	3.876e-01	1248	7.337e-06	-	6.719e-04	-	6
2	512	2.013e-01	4800	5.712e-07	3.68	1.116e-04	2.59	6
3	2048	1.024e-01	18816	3.895e-08	3.87	1.545e-05	2.85	6
4	8192	5.161e-02	74496	2.498e-09	3.96	1.991e-06	2.96	5
5	32768	2.590e-02	296448	1.563e-10	4.00	2.508e-07	2.99	5

	$\lvert\mathcal{T}_h\rvert$	h	n_{DoF}	L^2-error		H^1-error		n_{it}
1	128	3.876e-01	2176	8.701e-07	-	1.192e-04	-	6
2	512	2.013e-01	8448	4.264e-08	4.35	1.104e-05	3.43	6
3	2048	1.024e-01	33280	1.812e-09	4.56	8.029e-07	3.78	6
4	8192	5.161e-02	132096	8.603e-11	4.40	5.257e-08	3.93	5

Figure 7.7: Numerical results for $p = 1, 2, 3, 4$ from top to bottom.

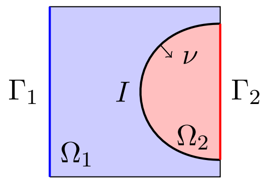

Figure 7.8: Geometry representing a single particle cathode.

We consider the elliptic subproblem (5.2) for the potential u, that is:

$$\begin{aligned} -\nabla \cdot (\kappa(c)\nabla u) &= 0 && \text{in } \Omega, \\ \kappa_i(c_i)\partial_\nu u_i &= i_{12}(c, [u]), && \text{on } \Omega, \\ \kappa(c)\partial_\nu u &= \chi_{\Gamma_1} && \text{on } \partial\overline{\Omega} \setminus \Gamma_2, \\ u_2 &= 0 && \text{on } \Gamma_2. \end{aligned} \tag{7.9}$$

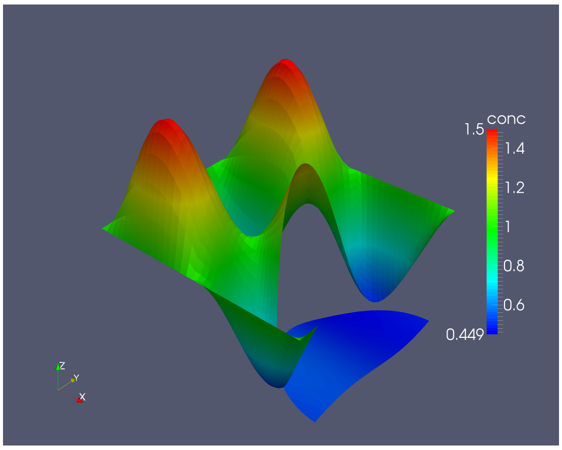

Figure 7.9: Graph of the concentrations given in (7.8).

Here, the exact solution is no longer explicitly known and it is thus not possible to calculate the discretization error $\|u - u_h\|_{H^1}$. As a surrogate, we consider the error $P(0) - T(u_h)$ where $T(v) := \|v\|_{1,2;\Omega}$ denotes the H^1-*energy* of the function $v \in H^1$. The value $P(0)$ is a higher order approximation to the unknown quantity $T(u)$. It is obtained in the following way: We start with a coarse mesh $\mathcal{T}_0$ which we successivly refine to obtain the sequence $\mathcal{T}_0, \ldots, \mathcal{T}_n$ of meshes. Let h_i be the maximal diameter of elements in $\mathcal{T}_i$ for $i = 0, \ldots, n$. Then $P \in \mathbb{P}_n$ is defined as the interpolation polynomial satisfying $P(h_i) = T(u_{h_i})$ for $i = 0, \ldots, n$. From polynomial interpolation theory it follows that, under the assumption that the mapping $h \mapsto T(u_h)$ is sufficiently smooth, the error $P(0) - T(u)$ is negligible compared to $T(u) - T(u_h)$ [46, §36]. As a consequence, the quantity $P(0) - T(u_h)$ is a *reasonable estimate* for the true error $T(u) - T(u_h)$.

The solution of the discrete system is shown in Fig. 7.10. In Table 7.2 the quantity $|P(0) - T(u_h)|$ is presented for a sequence of succesively refined meshes for $p \in \{1, 2\}$. In both cases, the experimental order of convergence (with a gliding mean of 2) is approximately 1. In combination with Corollary 6.1.27 this indicates that the *exact solution satisfies* (at most) $u \in H^2$.

7.3 Solving the Fully Coupled System

In this section we present the numerical solution of the fully coupled problem Problem 3.4.3. As suggested in Chapter 5, we will use the symbols Δ and $\nabla \cdot (\kappa(c)\nabla(\cdot))$ for the respective second order differential operators on Ω *with homogeneous Neumann and mixed boundary values*, respectively, see Definition 5.2.1. Additionally, $\mathcal{N}$ and $\mathcal{J}$ are corresponding to the respective remaining nonlinear Neumann boundary values, see Definition 5.2.3. With these definitions at hand, Problem 3.4.3 can be written formally in the following compact form:

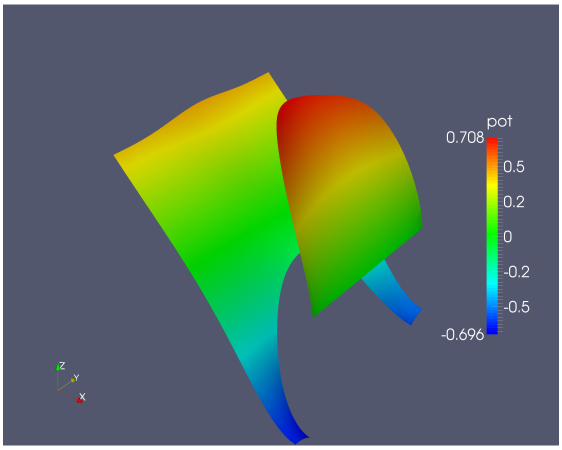

Figure 7.10: Solution of the elliptic subproblem.

$\lvert T_h \rvert$	$p=1$		$p=2$	
24	3.98e-02	-	-	-
70	1.73e-02	0.78	1.00e-02	-
234	5.25e-03	0.99	1.57e-03	1.54
850	1.52e-03	0.96	3.35e-04	1.20
3234	4.28e-04	0.95	8.00e-05	1.07
12610	1.15e-04	0.96	1.99e-05	1.02
49794	2.68e-05	1.06	5.03e-06	1.00

Table 7.2: Approximated energy error $|P(0) - T(u_h)|$.

Problem 3.4.3. *Find* $c, u : [0,T] \times \Omega \to \mathbb{R}$ *such that* $c(0) = c_0$ *and*

$$\partial_t c - \Delta c = \mathcal{N}(c, u), \tag{7.10}$$

$$-\nabla \cdot (\kappa(c)\nabla u) = \mathcal{J}(c, u), \quad u_2|_{\Gamma_2} = 0. \tag{7.11}$$

7.3.1 Semi Discretization in Time

The easiest time discretization is arguably the explicit Euler method, or a bit more general, the *θ-Euler method* for $\theta \in [0,1]$, see [46, §98]. For our problem it constists of determining c^n and u^n for $n \in \mathbb{N}$ such that it holds $c^0 = c_0$ and for $n \in \mathbb{N}_0$:

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^{n+\theta} = \mathcal{N}(c^{n+\theta}, u^{n+\theta}), \tag{7.12}$$

$$-\nabla \cdot (\kappa(c^{n+\theta})\nabla u^{n+\theta}) = \mathcal{J}(c^{n+\theta}, u^{n+\theta}), \quad u_2^{n+\theta}|_{\Gamma_2} = 0, \tag{7.13}$$

where $c^{n+\theta} := (1-\theta)c^n + \theta c^{n+1}$ and $u^{n+\theta} := (1-\theta)u^n + \theta u^{n+1}$.

The functions c^n and u^n are approximations for $c(t_n, \cdot)$ and $u(t_n, \cdot)$ respectively. Here, $t_n = n\tau$ with the *time-step* $\tau > 0$. Note that (7.12), (7.13) can be formally obtained from (7.10), (7.11) by evaluating at $t = t_n$ and replacing the time-derivative with a difference quotient.

Explicit Euler

Let us first discuss the case $\theta = 0$, that is, the *explicit Euler method.* In this case, (7.12) and (7.13) read

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^n = \mathcal{N}(c^n, u^n), \tag{7.14}$$

$$-\nabla \cdot (\kappa(c^n)\nabla u^n) = \mathcal{J}(c^n, u^n), \quad u_2^n|_{\Gamma_2} = 0. \tag{7.15}$$

Note that from these equations we cannot determine u^{n+1}, simply because it does not appear in the equations. However, this is not an issue: Recall that we are given initial values $c^0 = c_0$ for the concentration. Then we can solve the elliptic subproblem (7.15) for $n = 0$ to obtain u^0. Inductively, let us suppose, that we have already calculated c^k and u^k for some $k \in \mathbb{N}_0$. Then c^{k+1} can be obtained by explicitly solving (7.14) for $n = k$. However, to obtain u^{k+1} we need to solve the elliptic subproblem (7.15) for $n = k + 1$.

In order to make clear the order in which the equations are solved, we write down the method in the following equivalent way:

$$c^0 = c_0, \tag{7.16}$$

$$-\nabla \cdot (\kappa(c^0)\nabla u^0) = \mathcal{J}(c^0, u^0), \quad u_2^0|_{\Gamma_2} = 0, \tag{7.17}$$

and for $n \in \mathbb{N}_0$:

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^n = \mathcal{N}(c^n, u^n), \tag{7.18}$$

$$-\nabla \cdot (\kappa(c^{n+1})\nabla u^{n+1}) = \mathcal{J}(c^{n+1}, u^{n+1}), \quad u_2^{n+1}|_{\Gamma_2} = 0. \tag{7.19}$$

Implicit Euler

Now we discuss the case $\theta = 1$, that is, the *implicit Euler method.* It reads

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^{n+1} = \mathcal{N}(c^{n+1}, u^{n+1}), \tag{7.20}$$

$$-\nabla \cdot (\kappa(c^{n+1})\nabla u^{n+1}) = \mathcal{J}(c^{n+1}, u^{n+1}), \quad u_2^{n+1}|_{\Gamma_2} = 0. \tag{7.21}$$

Note that this is a fully implicit system for the unknown functions c^{n+1} and u^{n+1} (given, that c^n is known) and the system cannot be split into the parabolic and the elliptic part as for the case $\theta = 0$. Also it is worth pointing out, that u^0 does not occur in any of the equations (7.20) and (7.21) for $n \in \mathbb{N}_0$ and thus is not computed by this method.

Crank–Nicolson

For the sake of a simpler notation we only consider the case $\theta = 0.5$ instead of a general $\theta \in (0,1)$. Then (7.12), (7.13) read:

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^{n+0.5} = \mathcal{N}(c^{n+0.5}, u^{n+0.5}), \tag{7.22}$$

$$-\nabla \cdot (\kappa(c^{n+0.5}) \nabla u^{n+0.5}) = \mathcal{J}(c^{n+0.5}, u^{n+0.5}), \quad u_2^{n+0.5}|_{\Gamma_2} = 0. \tag{7.23}$$

If c^n and u^n are given for some $n \in \mathbb{N}_0$, (7.22) and (7.23) again yield a fully implicit system for the unknown functions c^{n+1} and u^{n+1}. In contrast to the implicit Euler method, however, the value u^n in fact enters the equation, yet u^0 cannot be determined from just (7.22) and (7.23) and therefore has to be provided. A natural choice for u^0 is the solution of the elliptic subproblem at the given initial concentration $c^0 = c_0$. The system then reads:

$$c^0 = c_0,$$
$$-\nabla \cdot (\kappa(c^0) \nabla u^0) = \mathcal{J}(c^0, u^0), \quad u_2^0|_{\Gamma_2} = 0,$$

and for $n \in \mathbb{N}_0$:

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^{n+0.5} = \mathcal{N}(c^{n+0.5}, u^{n+0.5}),$$
$$-\nabla \cdot (\kappa(c^{n+0.5}) \nabla u^{n+0.5}) = \mathcal{J}(c^{n+0.5}, u^{n+0.5}), \quad u_2^{n+0.5}|_{\Gamma_2} = 0.$$

A Semi-Implicit Method

The purpose of the method presented in this section is to combine the (relatively) low computational cost of the explicit Euler method with the good stability properties of the implicit Euler method by making use of the elliptic-parabolic structure of the system.

It is obtained from the implicit Euler method by replacing the unknown u^{n+1} in the parabolic part (7.20) by u^n. That way, the two equations are decoupled from each other while the discretization of the parabolic part is still an implicit one. The method reads:

$$c^0 = c_0,$$
$$-\nabla \cdot (\kappa(c^0) \nabla u^0) = \mathcal{J}(c^0, u^0), \quad u_2^0|_{\Gamma_2} = 0,$$

and for $n \in \mathbb{N}_0$:

$$\frac{1}{\tau}(c^{n+1} - c^n) - \Delta c^{n+1} = \mathcal{N}(c^{n+1}, u^n), \tag{7.24}$$

$$-\nabla \cdot (\kappa(c^{n+1}) \nabla u^{n+1}) = \mathcal{J}(c^{n+1}, u^{n+1}), \quad u_2^{n+1}|_{\Gamma_2} = 0. \tag{7.25}$$

For given c^n and u^n, the equations (7.24) and (7.25) are resolved in the following way: First, c^{n+1} is determined by solving (7.24) and then u^{n+1} is determined by solving (7.25) which is just the elliptic subproblem at c^{n+1}.

Note that the method presented here is just a basic example for how to combine different methods for the elliptic and parabolic part in order to obtain new methods which combine the advantages of the old ones.

Also the use of higher order methods like Runge–Kutta methods have not been discussed but might be the method of choice if we use higher order elements in space.

7.3.2 Fully Discrete Systems

In order to obtain computable problems with a finite number of unknowns, the semi-discrete systems from Section 7.3.1 still need to be discretized in space.

As for the elliptic subproblem in Section 7.2 we use the finite element method. We will discuss the resulting equations and the specific solution techniques for the explicit and implicit Euler method now.

Explicit Euler

Instead of looking for c^n, u^n for $n \in \mathbb{N}_0$ in some continuous function space satisfying the defining equations (7.17)–(7.19), we are now looking for $c_h^n \in W_h$ and $u_h^n \in V_h$ for $n \in \mathbb{N}_0$ satisfying *discrete versions* of (7.17)-(7.19).

To begin with, c_h^0 is taken to be an appropriate approximation of the continuous inital value c_0 in W_h. The function u_h^0 is then given by the discrete solution of (7.17), with c^0 replaced by c_h^0, as it has been described in Section 7.2

Now suppose, c_h^n and u_h^n are given for some $n \in \mathbb{N}$. Then u_h^{n+1} is the discrete solution of (7.19). The discrete concentration $c_h^{n+1} \in W_h$ is determined by solving the discrete version of (7.18). To be more precise, $c_h^{n+1} \in W_h$ is defined by satisfying

$$\frac{1}{\tau}\int_{\Omega_h}(c_h^{n+1}-c_h^n)w_h\,\mathrm{d}x+\int_{\Omega_h}\nabla c_h^n\cdot\nabla w_h\,\mathrm{d}x=-\int_{I_h} i_{12}(c_h^n,[u_h^n])w_{h,2}\,\mathrm{d}\sigma \tag{7.26}$$

for all $w_h \in W_h$. Expanding $c_h^{n+1}=\sum_i c_i^{n+1}\varphi_i$ in the nodal basis $\{\varphi_i \mid i \in \mathcal{N}_h\}$ of W_h, (7.26) reads for the vector of unknowns $c^{n+1}=(c_i^{n+1})_i \in \mathbb{R}^{\mathcal{N}_h}$:

$$Ac^{n+1}=b, \tag{7.27}$$

where $A \in \mathbb{R}^{\mathcal{N}_h\times\mathcal{N}_h}$ and $b \in R^{\mathcal{N}_h}$ are given by

$$A=\Big(\int_{\Omega_h}\varphi_j\varphi_i\,\mathrm{d}x\Big)_{ij},$$
$$b=\Big(-\tau\int_{I_h} i_{12}(c_h^n,[u_h^n])\varphi_{j,2}\,\mathrm{d}\sigma-\tau\int_{\Omega_h}\nabla c_h^n\cdot\nabla\varphi_j\,\mathrm{d}x+\int_{\Omega_h}c_h^n\varphi_j\,\mathrm{d}x\Big)_j.$$

The solution of the linear system (7.27) is particularly easy since it can be transformed into an equivalent well-conditioned system by Jacobi preconditioning, see for example [48, Section 3] and the references therein.

Implicit Euler

Similar to the fully discrete implicit Euler method, the problem is to determine the discrete approximations $c_h^n \in W_h$ for $n \in \mathbb{N}_0$ and $u_h^n \in V_h$ for $n \in \mathbb{N}$ to their semi-discrete counterparts c^n and u^n defined by (7.20) and (7.21).

For c_h^0 we again take an appropriate approximation of c_0 in W_h. Suppose c_h^n is given for some $n \in \mathbb{N}_0$. Then the finite elements discretization of (7.20) and (7.21) read:

Find $(c_h^{n+1}, u_h^{n+1}) \in W_h \times V_h$ such that

$$\begin{aligned} \frac{1}{\tau}\int_{\Omega_h} (c_h^{n+1} - c_h^n) w_h \,\mathrm{d}x + \int_{\Omega_h} \nabla c_h^{n+1} \cdot \nabla w_h \,\mathrm{d}x &= -\int_{I_h} i_{12}(c_h^{n+1}, [u_h^{n+1}]) w_{h,2} \,\mathrm{d}\sigma, \\ \int_{\Omega_h} \kappa(c_h^{n+1}) \nabla u_h^{n+1} \cdot \nabla v_h \,\mathrm{d}x &= -\int_{I_h} i_{12}(c_h^{n+1}, [u_h^{n+1}])[v_h] \,\mathrm{d}\sigma - \int_{\Gamma_{1,h}} j^{\mathrm{ext}} v_{h,1} \,\mathrm{d}\sigma \end{aligned} \tag{7.28}$$

holds for all $(w_h, v_h) \in W_h \times V_h$.

As it has been done for the elliptic subproblem in Section 7.2, we apply the Newton method to (7.28) and then expand c_h^{n+1} and u_h^{n+1} with respect to the nodal bases $\{\varphi_i \mid i \in \mathcal{N}_h\}$ and $\{\varphi_i \mid i \in \mathcal{N}_h^\circ\}$, respectively, that is

$$c_h^{n+1} = \sum_{i \in \mathcal{N}_h} c_i^{n+1} \varphi_i \quad \text{and} \quad u_h^{n+1} = \sum_{i \in \mathcal{N}_h^\circ} u_i^{n+1} \varphi_i$$

with the unknown coefficients $c_i^{n+1}, u_i^{n+1} \in \mathbb{R}$.

For the vector of unknowns $(c_h^{n+1}, u_h^{n+1}) \in \mathbb{R}^{\mathcal{N}_h} \times \mathbb{R}^{\mathcal{N}_h^\circ}$, the resulting linear system $Ax = b$ that has to be solved in each iteration of the Newton method is symmetric. Again we use the direct solver UMFPACK for the solution of the linear system [28].

For a more detailed description of this method we refer to [17].

7.3.3 Example 1: One-Dimensional Test Case

In order to validate the implementation of the numerical method for the fully-coupled problem which has been presented, we consider the following one-dimensional geometry: The electrolyte region is $\Omega_1 = (-1, 0)$ and the cathode region is $\Omega_2 := (0, 1)$. The boundary part corresponding to the anode is $\Gamma_1 := \{-1\}$ and the cathode current collector is represented by $\Gamma_2 := \{1\}$. Note that throughout the thesis we assumed $d \geq 2$ and therefore this example does not fit completely into the current framework. However, this issue is easily resolved by canonically extending all domains and functions to $\mathbb{R}^d$. For example, we can define $\Omega_1 := (-1, 0) \times Q$ for some $Q \subset \mathbb{R}^{d-1}$ and so on. This construction has also been used to validate the implementation for $d \in \{2, 3\}$.

Our code for the fully-coupled problem is written in terms of the variables c and Φ instead of c and u, see Section 3.4.2. As a consequence, we solve Problem 3.3.1 instead of Problem 3.4.3. We use the following parameters:

- $F = R = T$
- $D_1 = 0.005$, $D_2 = 0.01$

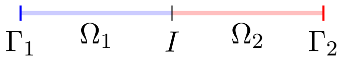

Figure 7.11: One-dimensional geometry.

- $\kappa_1 = 0.1$, $\kappa_2 = 1$
- $t_+ = 0.5$
- $k = 0.01$, $\alpha = 0.5$, $c_{\max,2} = 1$, $U \equiv 1$
- $j^{\text{ext}} = -0.03$
- $c_{0,1} = 0.5$, $c_{0,2} = 0.5$

In this one-dimensional situation, the exact solution can be calculated rather explicitly. The basic idea is that from the Neumann boundary condition on Γ_1, (3.16), it follows $\vec{j} = j^{\text{ext}}$ and $\vec{N} = N^{\text{ext}}$. As a result, the equation for the lithium conservation, (3.11), is then decoupled from the charge conservation equation, (3.12), and it is reduced to a heat equation with inhomogeneous Neumann boundary condition. This can be solved by making a Fourier series ansatz and adapting the coefficients to the initial values. Finally, the potential is obtained by integrating along Ω. The resulting solution is depicted in Fig. 7.12.

In the code we use the *implicit Euler method* described in Section 7.3.2 in combination with bilinear finite elements in space, that is, $p = 1$. Let us denote by $(c_{h,\tau}, \Phi_{h,\tau})$ the discrete solution corresponding to the mesh $\mathcal{T}_h$ and the time-step τ by , where h is the maximal diameter of elements in $\mathcal{T}_h$. We consider the error at the final time in the H^1-norm for both the concentration and the electrical potential Φ, see Fig. 7.13 and Fig. 7.14, respectively. For purely parabolic problems the expected convergence rate is $\mathcal{O}(\tau) + \mathcal{O}(h)$ [94], which is apparently attained in our simulations.

7.3.4 Example 2: Application Case

Let us wrap up this chapter by presenting the solution of the fully coupled system, Problem 3.3.1, for the single particle geometry which has also been used in Section 7.2.2 and which is shown in Fig. 7.9. We use the same parameters as in Section 7.3.3.

Let us point out that by the choice $j^{\text{ext}} < 0$ we simulate the *discharge* of the battery. The implicit Euler method from Section 7.3.2 is applied to this problem. The time-step is $\tau =$ 7.81e-03 and the maximal diameter of elements is $h =$ 5.35e-02 which results in 128 time-steps and $n_{\text{DoF}} = 6468$ degrees of freedom in space.

The results of this computation is shown in Fig. 7.15. In these pictures one can see that, within the electrolyte, the lithium gets transported from the anode to the cathode particles. Simultaneously, lithium accumulates inside the cathode particle. An interesting observation is that there are large differences in the pointwise norm of the

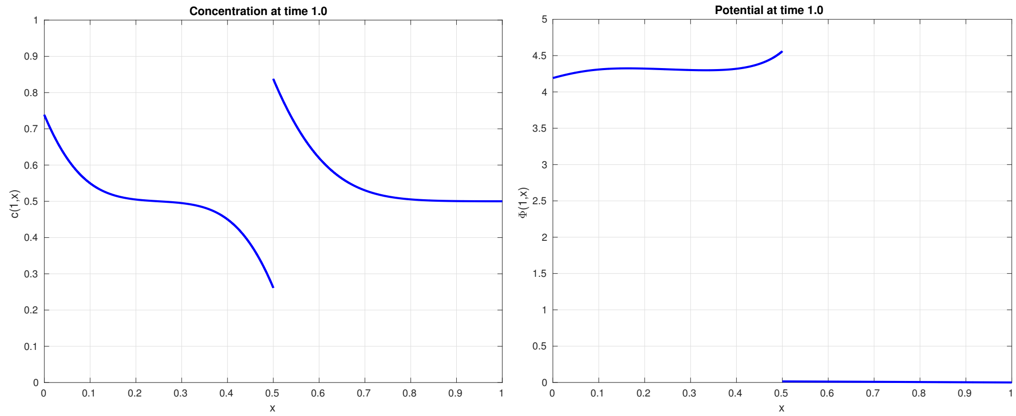

Figure 7.12: Exact solution for the fully coupled system in one dimension.

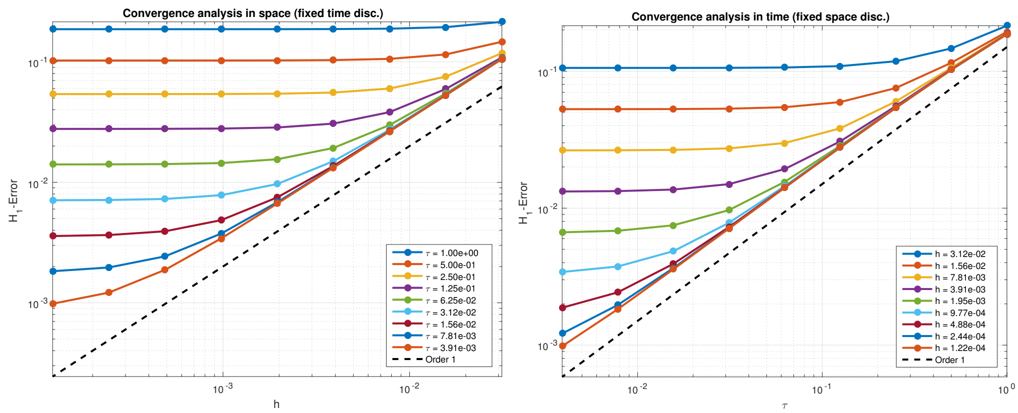

Figure 7.13: Error in the concentration $\|c(T) - c_{h,\tau}(T)\|_{1;2;\Omega}$

gradient: In the interior of both subdomains Ω_1 and Ω_2 it almost vanishes, whereas in the vicinity of the interface I and the boundary part Γ_1 it is very large. This suggest that a local refinement of the mesh in these areas can reduce the computational cost significantly.

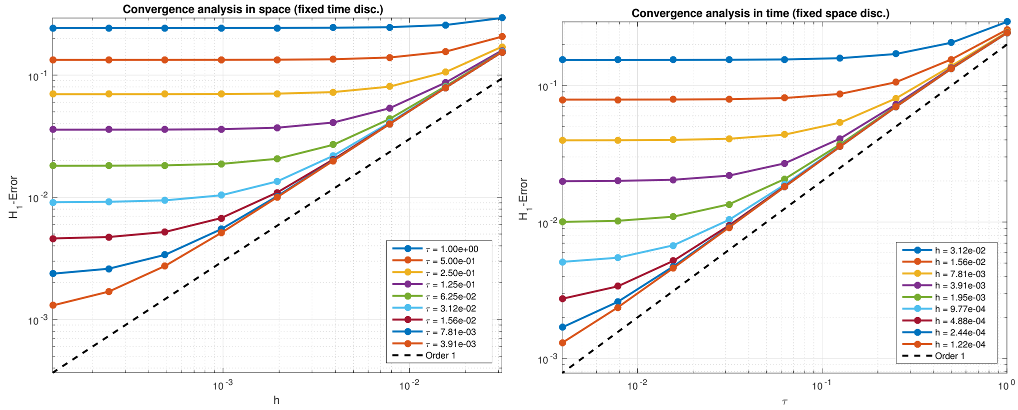

Figure 7.14: Error in the potential $\|\Phi(T) - \Phi^n_{h,\tau}(T)\|_{1;2;\Omega}$

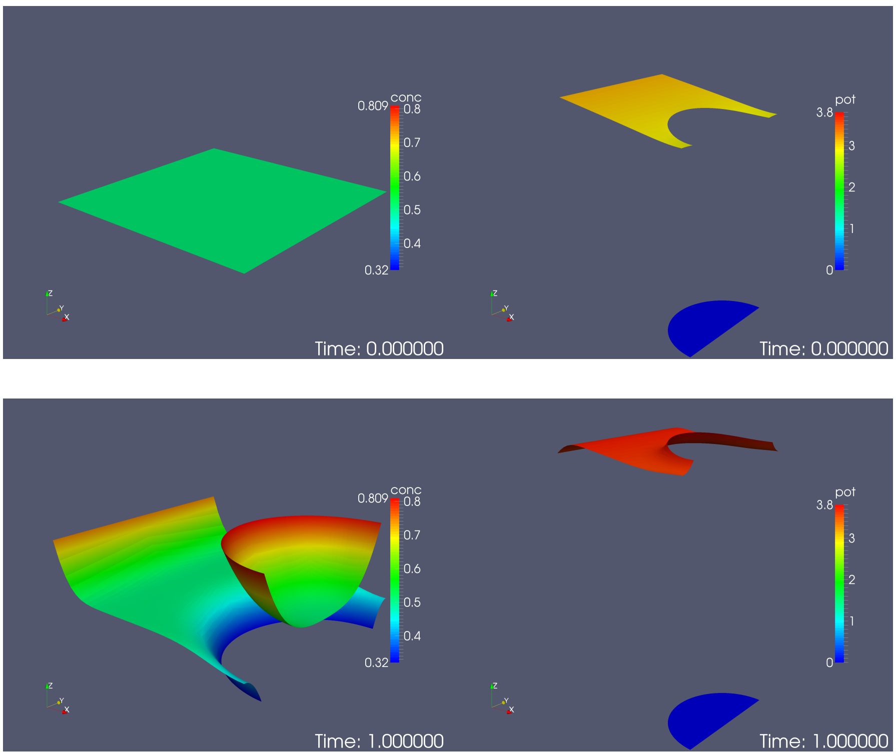

Figure 7.15: Lithium concentration and electrical potential in a single particle domain.

Bibliography

[1] Robert A. Adams and John J. F. Fournier. *Sobolev Spaces*, volume 140. Academic press, 2003.

[2] Hans-Wilhelm Alt. *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*. Springer Verlag, 2012.

[3] Herbert Amann. *Gewöhnliche Differentialgleichungen*. Walter de Gruyter, 1995.

[4] Herbert Amann. *Linear and Quasilinear Parabolic Problems: Volume I: Abstract Linear Theory*, volume 1. Springer Science & Business Media, 1995.

[5] Herbert Amann. Nonautonomous parabolic equations involving measures. *Journal of Mathematical Sciences*, 130(4):4780–4802, 2005.

[6] Sven Atlung, Keld West, and Torben Jacobsen. Dynamic aspects of solid solution cathodes for electrochemical power sources. *Journal of The Electrochemical Society*, 126(8):1311–1321, 1979.

[7] Sven Atlung, Birgit Zachau-Christiansen, Keld West, and Torben Jacobsen. The composite insertion electrode theoretical part. equilibrium in the insertion compound and linear potential dependence. *Journal of The Electrochemical Society*, 131(5):1200–1207, 1984.

[8] Pascal Auscher, Nadine Badr, Robert Haller-Dintelmann, and Joachim Rehberg. The square root problem for second-order, divergence form operators with mixed boundary conditions on L^p. *Journal of Evolution Equations*, 15(1):165–208, 2015.

[9] Utkarsh Ayachit. The paraview guide: A parallel visualization application. 2015.

[10] Wolfgang Bangerth, Ralf Hartmann, and Guido Kanschat. deal.ii - a general-purpose object-oriented finite element library. *ACM Transactions on Mathematical Software (TOMS)*, 33(4):24, 2007.

[11] Christine Bernardi. Optimal finite-element interpolation on curved domains. *SIAM Journal on Numerical Analysis*, 26(5):1212–1240, 1989.

[12] Christine Bernardi and Vivette Girault. A local regularization operator for triangular and quadrilateral finite elements. *SIAM Journal on Numerical Analysis*, 35(5):1893–1916, 1998.

[13] Yermal S. Bhat and Shari Moskow. Homogenization of a nonlinear elliptic boundary value problem modeling galvanic currents. *Multiscale Modeling & Simulation*, 5(1):149–169, 2006.

[14] Susanne C. Brenner and Ridgway Scott. *The Mathematical Theory of Finite Element Methods*, volume 15. Springer Science & Business Media, 2008.

[15] M. Broussely, Ph. Biensan, F. Bonhomme, Ph. Blanchard, S. Herreyre, K. Nechev, and R. J. Staniewicz. Main aging mechanisms in li ion batteries. *Journal of power sources*, 146(1):90–96, 2005.

[16] Peter G. Bruce. Energy storage beyond the horizon: Rechargeable lithium batteries. *Solid State Ionics*, 179(21-26):752–760, sep 2008.

[17] Fabian Castelli. Die numerische Behandlung eines nichtlinearen Systems mit Anwendungen bei Li-ionen Batterien. Master's thesis, Karlsruhe Institute of Technology, 2016.

[18] Min Chen and Gabriel A. Rincon-Mora. Accurate electrical battery model capable of predicting runtime and I-V performance. *Energy conversion, IEEE transactions on Energy C*, 21(2):504–511, 2006.

[19] John Christensen and John Newman. A mathematical model of stress generation and fracture in lithium manganese oxide. *Journal of The Electrochemical Society*, 153(6):A1019–A1030, 2006.

[20] Ph. G. Ciarlet and P.-A. Raviart. Maximum principle and uniform convergence for the finite element method. *Computer Methods in Applied Mechanics and Engineering*, 2(1):17–31, 1973.

[21] Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 40. Siam, 2002.

[22] Francesco Ciucci and Wei Lai. Derivation of micro/macro lithium battery models from homogenization. *Transport in porous media*, 88(2):249–270, 2011.

[23] Ph. Clément. Approximation by finite element functions using local regularization. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(2):77–84, 1975.

[24] Wolfgang Dahmen, Birgit Faermann, Ivan Graham, Wolfgang Hackbusch, and Stefan Sauter. Inverse inequalities on non-quasi-uniform meshes and application to the mortar element method. *Mathematics of computation*, 73(247):1107–1138, 2004.

[25] Robert Darling and John Newman. Modeling a porous intercalation electrode with two characteristic particle sizes. *Journal of The Electrochemical Society*, 144(12):4201–4208, 1997.

[26] Robert Darling and John Newman. Modeling side reactions in composite $Li_yMn_2O_4$ electrodes. *Journal of The Electrochemical Society*, 145(3):990–998, 1998.

[27] W. I. F. David, M. M. Thackeray, L. A. De Picciotto, and J. B. Goodenough. Structure refinement of the spinel-related phases $Li_2Mn_2O_4$ and $Li_{0.2}Mn_2O_4$. *Journal of Solid State Chemistry*, 67(2):316–323, 1987.

[28] Timothy A. Davis. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software (TOMS)*, 30(2):196–199, 2004.

[29] Robert de Levie and Andrew Vogt. On the electrochemical response of rough electrodes: Part I. The current transient following a voltage step in the absence of faradaic processes. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 278(1):25–34, 1990.

[30] Robert de Levie and Andrew Vogt. On the electrochemical response of rough electrodes: Part II. The transient response in the presence of slow faradaic processes. *Journal of electroanalytical chemistry and interfacial electrochemistry*, 281(1):23–28, 1990.

[31] Robert Denk and Tim Seger. L^p-estimates for a transmission problem of mixed elliptic-parabolic type. *Boundary Value Problems*, 2014(1):22, 2014.

[32] Rutooj Deshpande, Yang-Tse Cheng, and Mark W. Verbrugge. Modeling diffusion-induced stress in nanowire electrode structures. *Journal of Power Sources*, 195(15):5081–5088, 2010.

[33] Rutooj Deshpande, Yang-Tse Cheng, Mark W. Verbrugge, and Adam Timmons. Diffusion induced stresses and strain energy in a phase-transforming spherical electrode particle. *Journal of The Electrochemical Society*, 158(6):A718–A724, 2011.

[34] Peter Deuflhard. *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, volume 35. Springer Science & Business Media, 2011.

[35] Karoline Disser, A.F.M. ter Elst, and Joachim Rehberg. Hölder estimates for parabolic operators on domains with rough boundary. *arXiv preprint arXiv:1503.07035*, 2015.

[36] Giovanni Dore. L^p regularity for abstract differential equations. In *Functional Analysis and Related Topics, 1991*, pages 25–38. Springer, 1993.

[37] Marc Doyle, Thomas F. Fuller, and John Newman. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *Journal of the Electrochemical Society*, 140(6):1526–1533, 1993.

[38] Jérôme Droniou. Solving convection–diffusion equations with mixed, Neumann and Fourier boundary conditions and measures as data, by a duality method. *Advances in Differential Equations*, 5(10-12):1341–1396, 2000.

[39] Christof Eck, Harald Garcke, and Peter Knabner. *Mathematische Modellierung*. Springer, 2001.

[40] Johannes Elschner, Joachim Rehberg, and Gunther Schmidt. Optimal regularity for elliptic transmission problems including C^1 interfaces. *Interfaces and Free Boundaries*, 9(2):233–252, 2007.

[41] Alexandre Ern and Jean-Luc Guermond. *Theory and Practice of Finite Elements*, volume 159. Springer Science & Business Media, 2013.

[42] Lawrence C. Evans. *Partial Differential Equations*, volume 19. American Mathematical Society, 1998.

[43] Lawrence Craig Evans and Ronald F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC press, 2015.

[44] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer Verlag, 2015.

[45] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2011.

[46] Martin Hanke-Bourgeois. *Grundlagen der numerischen Mathematik und des wissenschaftlichen Rechnens*. Springer, 2009.

[47] Daniel Hug and Wolfgang Weil. A course on convex geometry. *Vorlesungsskript Universität Karlsruhe*, 2010.

[48] Lennard Kamenski, Weizhang Huang, and Hongguo Xu. Conditioning of finite element equations with arbitrary anisotropic meshes. *Mathematics of Computation*, 83(289):2187–2211, 2014.

[49] János Karátson and Sergey Korotov. Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions. *Numerische Mathematik*, 99(4):669–698, 2005.

[50] János Karátson and Sergey Korotov. Discrete maximum principles for finite element solutions of some mixed nonlinear elliptic problems using quadratures. *Journal of Computational and Applied Mathematics*, 192(1):75–88, 2006.

[51] János Karátson, Sergey Korotov, and Michal Křížek. On discrete maximum principles for nonlinear elliptic problems. *Mathematics and Computers in Simulation*, 76(1):99–108, 2007.

[52] Michael Kespe. Numerische Untersuchungen zur Interkalation bei Lithium-Ionen Batterien. Master's thesis, Karlsruhe Institute of Technology, 2012.

[53] Michael Kespe and Hermann Nirschl. Numerical simulation of lithium-ion battery performance considering electrode microstructure. *International Journal of Energy Research*, 39(15):2062–2074, 2015.

[54] Dorothee Knees. On the regularity of weak solutions of quasi-linear elliptic transmission problems on polyhedral domains. *Zeitschrift für Analysis und ihre Anwendungen*, 23(3):509–546, 2004.

[55] Olga Aleksandrovna Ladyzhenskaia, Vsevolod Alekseevich Solonnikov, and Nina N Uraltseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.

[56] Wei Lai and Francesco Ciucci. Mathematical modeling of porous battery electrodesrevisit of newman's model. *Electrochimica Acta*, 56(11):4369–4377, 2011.

[57] Arnulf Latz and Jochen Zausch. Thermodynamic consistent transport theory of Li-ion batteries. *Journal of Power Sources*, 196(6):3296–3302, 2011.

[58] Arnulf Latz and Jochen Zausch. Thermodynamic derivation of a Butler–Volmer model for intercalation in li-ion batteries. *Electrochimica Acta*, 110:358–362, 2013.

[59] Arnulf Latz and Jochen Zausch. Multiscale modeling of lithium ion batteries: thermal aspects. *Beilstein Journal of Nanotechnology*, 6:987–1007, 2015.

[60] Arnulf Latz, Jochen Zausch, and Oleg Iliev. *Modeling of species and charge transport in li–ion batteries based on non-equilibrium thermodynamics*. Springer, 2011.

[61] Giovanni Leoni. *A First Course in Sobolev Spaces*, volume 105. American Mathematical Society Providence, RI, 2009.

[62] Gregory B. Less, Jung H. Seo, S. Han, Ann Marie Sastry, Jochen Zausch, Arnulf Latz, Sebastian Schmidt, Christian Wieser, Dirk Kehrwald, and Stefan Fell. Microscale modeling of Li-ion batteries: parameterization and validation. *Journal of The Electrochemical Society*, 159(6):A697–A704, 2012.

[63] Gary M. Lieberman. *Second Order Parabolic Differential Equations*, volume 68. World Scientific, 2005.

[64] Jürgen Moser. On Harnack's theorem for elliptic differential equations. *Communications on Pure and Applied Mathematics*, 14(3):577–591, 1961.

[65] Gowri S. Nagarajan, J. W. Van Zee, and R. M. Spotnitz. A mathematical model for intercalation electrode behavior I. effect of particle-size distribution on discharge capacity. *Journal of the Electrochemical Society*, 145(3):771–779, 1998.

[66] Gholam-Abbas Nazri and Gianfranco Pistoia, editors. *Lithium Batteries*. Springer Science & Business Media, 2003.

[67] Gerold Neumann. Lithium-Akkumulatoren: Stand der Technik und zuknftige Entwicklungen. *Chemie Ingenieur Technik*, 83(11):2042–2050, oct 2011.

[68] John Newman and Karen .E Thomas-Alyea. *Electrochemical Systems*. John Wiley & Sons, 2012.

[69] Naoki Nitta, Feixiang Wu, Jung Tae Lee, and Gleb Yushin. Li-ion battery materials: present and future. *Materials Today*, 18(5):252–264, jun 2015.

[70] P. Popov, Yavor Vutov, Svetozar Margenov, and O. Iliev. *Finite volume discretization of equations describing nonlinear diffusion in Li-ion batteries.* Springer, 2011.

[71] Michael Ruzicka. *Nichtlineare Funktionalanalysis: Eine Einführung.* Springer-Verlag, 2006.

[72] Yousef Saad. *Iterative methods for sparse linear systems.* Siam, 2003.

[73] Alfred H. Schatz. A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I. *Mathematics of Computation*, 34(149):77–91, 1980.

[74] Tim Seger. *Elliptic-Parabolic Systems with Applications to Lithium-Ion Battery Models.* PhD thesis, University of Konstanz, 2013.

[75] Guido Stampacchia. Some limit cases of L^p-estimates for solutions of second order elliptic equations. *Communications on Pure and Applied Mathematics*, 16(4):505–510, 1963.

[76] Guido Stampacchia. Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. In *Annales de l'institut Fourier*, volume 15, pages 189–257, 1965.

[77] Guido Stampacchia. Equations elliptiques du second ordre à coefficients discontinus. *Séminaire Jean Leray*, (3):1–77, 1966.

[78] Gilbert Strang. Variational crimes in the finite element method. *The mathematical foundations of the finite element method with applications to partial differential equations*, pages 689–710, 1972.

[79] J.-M. Tarascon and Michel Armand. Issues and challenges facing rechargeable lithium batteries. *Nature*, 414(6861):359–367, 2001.

[80] A. F. M. ter Elst and Joachim Rehberg. Hölder estimates for second-order operators on domains with rough boundary. *Advances in Differential Equations*, 20(3/4):299–360, 2015.

[81] Hans Triebel. *Interpolation Theory, Function Spaces, Differential Operators.* Barth, Heidelberg, 2., rev. and enl. ed. edition, 1995.

[82] J. Vetter, P. Novák, M. R. Wagner, C. Veit, K.-C. Möller, J. O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, and A. Hammouche. Ageing mechanisms in lithium-ion batteries. *Journal of power sources*, 147(1):269–281, 2005.

[83] Stefan Volkwein and Andrea Wesche. The reduced basis method applied to transport equations of a lithium-ion battery. *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 32(5):1760–1772, 2013.

[84] Deyu Wang, Xiaodong Wu, Zhaoxiang Wang, and Liquan Chen. Cracking causing cyclic instability of $LiFePO_4$ cathode material. *Journal of Power Sources*, 140(1):125–128, 2005.

[85] Haifeng Wang, Young-Il Jang, Biying Huang, Donald R. Sadoway, and Yet-Ming Chiang. TEM study of electrochemical cycling-induced damage and disorder in $LiCoO_2$ cathodes for rechargeable lithium batteries. *Journal of the Electrochemical Society*, 146(2):473–480, 1999.

[86] Guo Chun Wen and Heinrich G. W. Begehr. *Boundary value problems for elliptic equations and systems*, volume 46. Chapman & Hall/CRC, 1990.

[87] Hans Wilhelm Alt and Stephan Luckhaus. Quasilinear elliptic-parabolic differential equations. *Mathematische Zeitschrift*, 183(3):311–341, 1983.

[88] Jinbiao Wu, Jinchao Xu, and Henghui Zou. On the well-posedness of a mathematical model for lithium-ion battery systems. *Methods and Applications of Analysis*, 13(3):275–298, 2006.

[89] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications, Vol. II/A Linear Monotone Operators, Vol. II/B Nonlinear Monotone Operators*. Springer Verlag, 1985.

[90] Qi Zhang and Ralph E. White. Moving boundary model for the discharge of a $LiCoO_2$ electrode. *Journal of The Electrochemical Society*, 154(6):A587–A596, 2007.

[91] Shiquan Zhang, Oleg Iliev, Sebastian Schmidt, and Jochen Zausch. Comparison of two approaches for treatment of the interface conditions in FV discretization of pore scale models for Li-ion batteries. In *Finite Volumes for Complex Applications VII-Elliptic, Parabolic and Hyperbolic Problems*, pages 731–738. Springer, 2014.

[92] Xiangchun Zhang. *Multiscale modeling of Li-ion cells: mechanics, heat generation and electrochemical kinetics.* PhD thesis, The University of Michigan, 2009.

[93] Xiangchun Zhang, Wei Shyy, and Ann Marie Sastry. Numerical simulation of intercalation-induced stress in Li-ion battery electrode particles. *Journal of the Electrochemical Society*, 154(10):A910–A916, 2007.

[94] Miloš Zlámal. Finite element methods for parabolic equations. *Mathematics of Computation*, 28(126):393–404, 1974.