INAUGURAL – DISSERTATION

zur

Erlangung der Doktorwürde

der

Gesamtfakultät für Mathematik, Ingenieur- und Naturwissenschaften

der

Ruprecht – Karls – Universität

Heidelberg

vorgelegt von

Seyedehmasoumeh Hashemibarmchi, M.Sc.

aus Rasht, Iran

Tag der mündlichen Prüfung:

Optimal Control of Nonlocal Partial Differential Equations

First Advisor: Prof. Dr. Roland Herzog

Second Advisor: Assoc. Prof. Morteza Fotouhi

To my amazing mom and dad

Bita and Saeed

Abstract

In the past decades, the optimal control of partial differential equations governed by partial differential equations has made significant progress. This work concerns optimal control of nonlocal partial differential equations, raising the natural question as to why this type of partial differential equations are of interest and relevance. Nonlocal partial differential equations abound in modeling of various physical and biological phenomena. In contrast to classical partial differential equations, nonlocal partial differential equations take not only the local spatial or time variables into consideration, but also any possible dependence of the involved quantities on neighboring points as well as preceding times in the evolution of the process under consideration. This type of non-local reliance typically arises from interactions over a distance or from multiple conservation laws.

In this thesis, our primary emphasis was placed on two significant nonlocal partial differential equations: one originating from the field of physics and the other from the biology. In our first study we consider an optimal control problem for the steady-state Kirchhoff equation, a prototype for nonlocal partial differential equations, different from fractional powers of closed operators. Existence and uniqueness of solutions of the state equation, existence of global optimal solutions, differentiability of the control-to-state map and first-order necessary optimality conditions are established. The aforementioned results require the controls to be functions in H^1 and subject to pointwise lower and upper bounds. In order to obtain the Newton differentiability of the optimality conditions, we employ a Moreau-Yosida-type penalty approach to treat the control constraint and study its convergence. The first-order optimality conditions of the regularized problems are shown to be Newton differentiable, and a generalized Newton method is detailed. A discretization of the optimal control problem by piecewise linear finite elements is proposed and numerical results are presented.

In our second study, we delve into an optimal control problem involving a coupled parabolic-elliptic chemotaxis system with a nonlocal logistic growth term. We establish the existence and uniqueness of the state equation. By constructing the corresponding logistic Ordinary Differential Equation (ODE), we determine the maximal existence time to prevent blow-up. Depending on the sign coefficient of the nonlinear term in the ODE, it either blows up in finite time or in infinite time. We demonstrate that the solution y of the Partial Differential Equation (PDE) is bounded above by the solution of the ODE. Subsequently, we provide an a-priori estimate for the solution of the chemotaxis system and establish the existence of an optimal solution. Finally, we demonstrate the Fréchet differentiability of the control-to-state map and derive first-order necessary conditions using the Lagrangian method.

Zusammenfassung

In den letzten Jahrzehnten hat die optimale Steuerung von partiellen Differentialgleichungen, die durch partielle Differentialgleichungen beschrieben werden, erhebliche Fortschritte gemacht. Diese Arbeit befasst sich mit der optimalen Steuerung nichtlokaler partieller Differentialgleichungen und stellt die natürliche Frage, warum diese Art von partiellen Differentialgleichungen von Interesse und Relevanz ist.

Nichtlokale partielle Differentialgleichungen sind reichlich vorhanden bei der Modellierung verschiedener physikalischer und biologischer Phänomene. Im Gegensatz zu klassischen partiellen Differentialgleichungen berücksichtigen nichtlokale partielle Differentialgleichungen nicht nur die lokalen räumlichen oder zeitlichen Variablen, sondern auch jede mögliche Abhängigkeit der beteiligten Größen von benachbarten Punkten sowie vorangegangenen Zeiten im Verlauf des zu betrachtenden Prozesses. Diese Art der nicht-lokalen Abhängigkeit entsteht typischerweise durch Wechselwirkungen über eine Entfernung oder durch mehrere Erhaltungsgesetze.

In dieser Arbeit lag unser Hauptaugenmerk auf zwei bedeutenden nichtlokalen partiellen Differentialgleichungen: eine stammt aus dem Bereich der Physik und die andere aus der Biologie. In unserer ersten Studie betrachten wir ein Optimalsteuerungsproblem für die stationäre Kirchhoff-Gleichung, einen Prototypen für nichtlokale partielle Differentialgleichungen, die von den Bruchteilen geschlossener Operatoren abweicht. Existenz und Eindeutigkeit der Lösungen der Zustandsgleichung, Existenz global optimaler Lösungen, Differenzierbarkeit der Abbildung von der Steuerungs-Zustands-Operator und notwendige Optimalitätsbedingungen erster Ordnung werden etabliert. Die genannten Ergebnisse erfordern, dass die Steuerungen Funktionen in H^1 sind und punktweise untere und obere Schranken unterliegen. Um die Newton-Differenzierbarkeit der Optimalitätsbedingungen zu erhalten, verwenden wir einen Ansatz mit Strafterm im Stil von Moreau-Yosida, um die Steuerungsbeschränkung zu behandeln, und untersuchen ihre Konvergenz. Die Optimalitätsbedingungen erster Ordnung der regularisierten Probleme zeigen sich als Newton-differenzierbar, und eine verallgemeinerte Newton-Methode wird detailliert. Es wird eine Diskretisierung des optimalen Steuerungsproblems durch stückweise lineare Finite-Elemente vorgeschlagen und numerische Ergebnisse werden präsentiert.

In unserer zweiten Studie vertiefen wir uns in ein Optimalsteuerungsproblem, das ein gekoppeltes parabolisch-elliptisches Chemotaxis-System mit einem nichtlokalen logistischen Wachstumsterm betrifft. Wir etablieren die Existenz und Eindeutigkeit der Zustandsgleichung. Durch die Konstruktion der entsprechenden logistischen gewöhnlichen Differentialgleichung (ODE) bestimmen wir die maximale Existenzzeit, um ein Blowup zu verhindern. Abhängig vom Vorzeichen des nichtlinearen Terms in der ODE entsteht ein Blowup entweder in endlicher Zeit oder in unendlich lange Zeit. Wir zeigen, dass die Lösung y der partiellen Differentialgleichung (PDE) von oben durch die Lösung der ODE begrenzt ist. Anschließend geben wir eine a-priori-Abschätzung für die Lösung des Chemotaxis-Systems und etablieren die Existenz einer optimalen Lösung. Schließlich zeigen wir die Fréchet-Differenzierbarkeit der Abbildung von der Steuerung zum Zustand und leiten notwendige Bedingungen erster Ordnung mit Hilfe der Lagrange-Methode her.

Acknowledgements

I would like to express my heartful gratitude to my primary supervisor, Prof. Dr. Roland Herzog, who guided me throughout this journey. A few months after beginning my doctoral research, the world experienced a strange situation. The coronavirus affected the personal and work life of every person. This situation continued for two years, and I personally witnessed that many students found themselves unable to continue their research. I have to admit that Roland Herzog was my biggest encouragement to stay on this path, and without his support, it would have been impossible to persevere.

I would also like to express my deep appreciation to my second supervisor, Assoc. Prof. Morteza Fotouhi, who supported me in this project, especially during my trip to Iran. He provided me with a guest office in Tehran at Sharif University of Technology and assisted me in every scientific discussion that was necessary. I would like to thank Georg Müller, with whom I was in the teaching group, for his professional performance in preparing the exercises. His assistance helped me focus on my research topic, and I appreciate his patience in answering my questions.

Special thanks to Hannes Meinlschmidt for fruitful discussions concerning the optimal control of the chemotaxis equation.

I cannot express my gratitude to my friends and family for how much they trusted me and supported me after I moved to Germany.

Contents

1 Introduction

Optimal control problems governed by partial differential equations progressed quickly in the past decade. Among these, nonlocal partial differential equations are of crucial importance. They provide essential aspects of real-world phenomena by incorporating all available information in the evolution of the observed process. Therefore, one can make more useful and accurate predictions in various application areas through the mathematical study of these equations. An enormous range of such equations has emerged in the literature, finding notable applications initially in physics and later expanding into engineering, astrophysics, and biology. Among the earliest instances of nonlocal equations are those encountered in the filed of phase transitions, connected to theories introduced by Chen, Fife, 2000. Models incorporating nonlocal spatial terms are encountered in various contexts, such as Ohmic heating production Lacey, 1995; Quittner, Souplet, 2007, the theory of gravitational equilibrium of polytropic stars Lacey, 1983, population dynamics Furter, Grinfeld, 1989, and the modeling of cell aggregation through interaction with a chemical substance (chemotaxis) Wolansky, 1997.

1.1 Main Contributions

In this Thesis we study the optimal control of two different nonlocal nonlinear partial differential equations:

- optimal control of stationary Kirchhoff equation
- optimal control a nonlocal chemotaxis system.

1.1.1 Optimal Control of Stationary Kirchhoff Equation

In this work we study an optimal control problem governed by a nonlinear, nonlocal partial differential equation (PDE) of Kirchhoff-type

$$\begin{cases} -M\big(x, \|\nabla y\|^2_{L^2(\Omega)}; u\big)\Delta y = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.1.1}$$

Here, $\Omega \subset \mathbb{R}^N$ is an open and bounded set and the right-hand side f belongs to $L^2(\Omega)$. We focus on the particular case $M(x, s; u) = u(x) + b(x)\, s$, which has been considered previously, e. g., in Figueiredo et al., 2014; Delgado, Figueiredo, et al., 2017. Here u and b are strictly positive functions and u serves as the control. The full set of assumptions is given in section 4.1. We mention that in case u and b are positive *constants*, (1.1.1) has a variational structure; see Figueiredo et al., 2014.

Equation (1.1.1) is the steady-state problem associated with its time-dependent variant

$$\begin{cases} y_{tt} - M\big(x, \|\nabla y\|^2_{L^2(\Omega)}; u\big)\Delta y = f & \text{in } \Omega \times (0,T), \\ y = 0 & \text{on } \partial\Omega \times (0,T), \\ y(x,0) = y_0(x), \quad y_t(x,0) = y_1(x) & \text{in } \Omega. \end{cases} \tag{1.1.2}$$

In one space dimension, problem (1.1.2) models small vertical vibrations of an elastic string with fixed ends, when the density of the material is not constant. Specifically, the control u is proportional to the inverse of the string's cross section; see Ma, 2005; Figueiredo et al., 2014. A physical interpretation of the multi-dimensional problems (1.1.1) and (1.1.2) appears to be missing in the literature.

As mentioned before, PDEs with nonlocal terms play an important role in physics and technology and they can be mathematically challenging. Although in some cases variational reformulations are available, the models (1.1.1), (1.1.2) do not allow this in general. Thus, despite the deceptively simple structure, (1.1.1) requires a set of analytical tools not often employed in PDE-constrained optimization. Existence and uniqueness of solutions for (1.1.1) have been investigated in Figueiredo et al., 2014 and Delgado, Figueiredo, et al., 2017; see also the references therein. For further applications of nonlocal PDEs, we refer the reader to Eringen, 1983; Ahmed, Elgazzar, 2007; Kavallaris, Suzuki, 2018.

The authors in Delgado, Figueiredo, et al., 2017 studied an optimal control problem for (1.1.1) with the following cost functional

$$J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Omega)} \tag{1.1.3}$$

with an admissible set $\mathcal{U}_{\text{ad}} = \{u \in L^2(\Omega) \,|\, u \geq u_a > 0 \text{ a.e. in } \Omega\}$. However we believe that the proof of existence of an optimal solution in this work has a flaw. We give further details in the appendix. Moreover, the proof in Delgado, Figueiredo, et al., 2017 is explicitly tailored to such tracking type functionals. In the present work we see it necessary to modify the control cost term to contain the stronger H^1-norm. We also allow for a more general state dependent term, which leads to the objective

$$J(y,u) = \int_\Omega \varphi(x, y(x))\,\mathrm{d}x + \frac{\lambda}{2}\|u\|^2_{H^1(\Omega)} \tag{1.1.4}$$

and a set of admissible controls in $H^1(\Omega)$. In this setting, we prove the weak-strong continuity of the control-to-state operator into $H^1_0(\Omega) \cap W^{2,q}(\Omega)$ for any $q \in [1,\infty)$ and the existence of a globally optimal solution. Moreover, we work with a pointwise lower bound on admissible controls. This bound has an immediate technological interpretation, representing an upper bound on the string's cross section. On the other hand, we also impose an upper bound on the controls. This is to be able to use the topology of $L^\infty(\Omega)$ in the proof of the Fréchet differentiability of the control-to-state map so that we can derive optimality conditions in a more straightforward way than by the Dubovitskii-Milyoutin formalism utilized in Delgado, Figueiredo, et al., 2017.

The first-order optimality conditions obtained when minimizing (1.1.4) subject to (1.1.1) involve a variational inequality of nonlinear obstacle type in H^1. We choose to relax and penalize the bound constraints via a Moreau-Yosida regularization, which amounts to a quadratic penalty of the bound constraints for the control. In this setting, we can prove the generalized (Newton) differentiability of the optimality system. A similar philosophy, albeit for a different problem, has been pursued by Adam, Hintermüller, Surowiec, 2018. We also mention Ulbrich, 2011, Chapter 9.2 for an approach via a regularized dual obstacle problem. A recent alternative is offered by Christof, Wachsmuth, 2023, where the Newton differentiability of the solution map for unilateral obstacle problems is shown, without the need to penalize the constraint. Indeed, relaxing the lower and upper bounds adds new difficulties, since the existence of a solution of the Kirchhoff equation (1.1.1) can only be guaranteed for positive controls. Therefore, we compose the control-to-state map with a smooth cut-off function. We then study the convergence of global minimizers as the penalty parameter goes to zero, see theorem 3.2.6 for details. We can expect a corresponding result to hold also for locally optimal solutions under an assumption of second-order sufficient optimality conditions, but this is not investigated here.

To summarize our contributions in comparison to Delgado, Figueiredo, et al., 2017, we consider a more general objective, present a simpler proof for the existence of a globally

optimal control, prove the differentiability of the control-to-state map and generalized differentiability of the optimality system for a regularized version of the problem as well as the applicability of a generalized Newton scheme. We also describe a structure preserving finite element discretization of the problem and the discrete counterpart of the generalized Newton method.

1.1.2 Optimal Control a Nonlocal Chemotaxis Model

Taxis refers to the motion of an organism towards or away from an external stimulus. Specifically, when the stimulus is a chemical, it is termed *chemotaxis*. Positive chemotaxis occurs, if the cell movement is toward a higher concentration of the chemical in question, while negative chemotaxis occurs if the movement is in the opposite direction.

To address specific aspects of chemotaxis, numerous mathematical models have been proposed. One of the most crucial and interesting models is the Keller-Segel model, introduced in 1970. This model consists of two equations, forming a parabolic-parabolic system, see Keller, Segel, 1970. This model describes the evolution of the population density $y(x,t)$ of motile cells (or other living organisms) and the concentration $w(x,t)$ of a chemically attracting substance (chemoattractant), which is produced by the cell population itself. In a bounded domain $\Omega \subset \mathbb{R}^N$ and a time interval $[0,T]$

$$\begin{cases} \partial_t y = \Delta y - \chi \operatorname{div}(y \nabla w) + g(y) & \text{in } \Omega \times (0,T), \\ \partial_t w = \Delta w + f(w,y) & \text{in } \Omega \times (0,T), \\ \dfrac{\partial y}{\partial n} = \dfrac{\partial w}{\partial n} = 0 & \text{on } \partial\Omega \times (0,T), \\ y(x,0) = y_0, w(x,0) = w_0 & \text{in } \Omega. \end{cases} \tag{1.1.5}$$

The chemotactic coefficient χ is a positive constant. The term $g(y)$ represents the growth term for the cells and $f(w,y)$ is the kinetics/source term, which may depend on w and y. There is no flux of cells or chmoatractans across the boundary of the domain. The initial concentrations y_0 and w_0 are non-negative functions.

We refer the reader to Hillen, Painter, 2008 for a review of a number of variations of the original Keller-Segel model from a biological perspective. The review article Horstmann, 2004 provides a detailed introduction into the mathematics of the Keller-Segel model for chemotaxis. Arumugam, Tyagi, 2021 discussed some of the most important analytical methods and blow-up criteria for analyzing the solutions of Keller-Segel chemotaxis models. Quite a few of the known results on numerical methods have been discussed in this review. In Egger, Pietschmann, Schlottbom, 2015 a parameter identification of a nonlinear parabolic-elliptic system has been investigated. The existence of global bounded classical solutions has been proved in Tello, Winkler, 2007.

There are some studies for optimal control of Keller-Segel models and chemotaxis equations, see for instance Ryu, Yagi, 2001; Fister, McCarthy, 2003; Rodríguez-Bellido, Rueda-Gómez, Villamizar-Roa, 2018 and the references therein. Recently Liu, Yuan, 2022 addressed a distributed optimal control problem for an attraction-repulsion chemotaxis system, which describes the process of cells interacting with a combination of repulsive and attractive signal chemicals. A numerical investigation of optimal control of self-organisation dynamics in a chemotaxis reaction diffusion system carried out in Lebiedz, Maurer, 2004. Dolgov, Pearson, 2019 considered the efficient numerical solution of an optimal control formulation for the Keller–Segel model, specifically addressing bacterial chemotaxis. A chemotaxis equation can arise in cancer models, such as angiogenesis. Optimal control of such equations has been studied in Delgado, Gayte, Morales Rodrigo, 2021. In Belmiloudi, 2017 a mathematical model describing the dynamics of interaction between tumor and normal cells has been

presented. The paper also delved into an optimal control problem, emphasizing the role of drugs in treating brain tumors.

In this work we study an optimal control problem governed by a nonlinear, nonlocal system of partial differential equations (PDEs) under nonlocal chemotactic effects

$$\begin{cases} \partial_t y - \Delta y = -\chi \operatorname{div}(y \nabla w) + y \left(a_0 - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x \right) & \text{in } \Omega \times (0,T), \\ -\Delta w + \lambda\, w = y & \text{in } \Omega \times (0,T), \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega \times (0,T), \\ y(x,0) = y_0(x) & \text{in } \Omega. \end{cases} \tag{1.1.6}$$

where all the constants χ, a_0, a_1, λ are positive and a_2 belongs to $\mathbb{R}$. This equation has been proposed in Negreanu, Tello, 2013 to study of single-species and two-species system in competition for the case where the chemical is also introduced in the system from outside, i. e., in the elliptic equation is an artificial external chemical force imposed. For further application of nonlocal PDEs, we refer the reader to Eringen, 1983; Ahmed, Elgazzar, 2007; Kavallaris, Suzuki, 2018.

The first equation represents the rate of change of the cell density $\partial_t y$, where $-\Delta y$ and $-\chi \operatorname{div}(y \nabla w)$ describe the diffusion and chemotactic contribution, respectively. The reaction term, namely logistic population growth $y\left(a_0 - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x\right)$, counteracts the blow-up tendency produced by chemotaxis. The nonlocal term $\fint_\Omega y \,\mathrm{d}x \coloneqq \frac{1}{|\Omega|}\int_\Omega y \,\mathrm{d}x$, which describes the total mass of the population, has a balancing effect. In fact, this type of logistic growth describes the competition among cells for environmental resources and their cooperation to survive. The impact of the nonlocal term evidently depends on the sign of a_2. If $a_2 > 0$, individuals of the species compete, hindering the growth of the population. However, when $a_2 < 0$ the effects of both $a_1\, y$ and $a_2 \fint y$ balance the system. In this case, individuals compete locally but cooperate globally.

The second equation, a stationary version of the reaction-diffusion equation for chemoattractant, models that the attractant w diffuses as a chemical and is produced by the cells. It will be assumed that the chemoattractant diffuses much faster than the cell population. This results in an interesting case with $\partial_t w = 0$. The first term in the kinetics/source term $y - \lambda\, w$ represents the spontaneous production of the chemoattractant and is proportional to the number of cells, while $-\lambda\, w$ represents decay of attractant activity; see Murray, 1989, Chapter 11 and Negreanu, Tello, 2013. In this model, we impose a positive flux of the chemoattractant on the boundary of the domain.

Existence and uniqueness of solutions for a chemotaxis system with a local logistic source have been investigated in Tello, Winkler, 2007; see also the references therein. See for instance Ryu, Yagi, 2001; Fister, McCarthy, 2003; Rodríguez-Bellido, Rueda-Gómez, Villamizar-Roa, 2018.

In this work we study an optimal control problem for (1.1.6) with the following standard tracking-type cost functional

$$J(y,u) \coloneqq \frac{1}{2}\int_\Omega |y(x,T) - y_d(x)|^2 \,\mathrm{d}x + \frac{\gamma}{2}\int_0^T \int_{\partial\Omega} |u(x,t)|^2 \,\mathrm{d}s\,\mathrm{d}t.$$

The first term measures the discrepancy between the cell density y and and desired density y_d at final time T, while the second models the control effort.

1.2 Outline of the Thesis

The thesis is structured as follows.

Chapter 2 – Fundamentals of Optimal Control of PDEs

gives a mathematical background into both PDEs and optimal control problems governed by PDEs. This chapter consists of two sections. In section 2.1 we give the main related results about the existence of solutions of linear and nonlinear elliptic and parabolic PDEs. Section 2.2 is devoted to the introducing of the main concepts of optimal control problems governed by PDEs.

Chapter 3 – Optimal Control of the Stationary Kirchhoff Equation

focuses on optimal control of a nonlocal nonlinear elliptic equation, namely the Kirchhoff equation. In section 3.1, we demonstrate the existence and uniqueness of solutions to the Kirchhoff equation, as well as the existence of a solution to the optimal control problem. Fréchet differentiability of the control-to-state map is proved in section 3.2. We also derive a system of necessary optimality conditions for a regularized problem and construct an analytical solution for the optimal control problem. We devote section 3.3 to showing the Newton differentiability of the optimality system and devising a locally superlinearly convergent scheme in appropriate function spaces. We discretize the optimal control problem, its optimality system and the generalized Newton method by a finite element scheme in section 3.4. This chapter ends with describing some numerical experiments in section 3.5.

Chapter 4 – Optimal Control a Nonlocal Chemotaxis Model

focuses on optimal control of a nonlocal nonlinear system of PDEs, namely a parabolic-elliptic chemotaxis system. In section 4.1 we address the existence and uniqueness of solutions to the chemotaxis equation and the existence of an optimal control. Fréchet differentiability of control-to-state map is proved in section 4.2. A first-order optimality system is also derived in this section.

Chapter 5 – Conclusions and Outlook

summarizes the primary results of the thesis and outlines potential avenues for future research.

2 Fundamentals of Optimal Control of PDEs

Contents

All definitions and results in this chapter, without reference, are based on Evans, 1998; Tröltzsch, 2010; Manzoni, Quarteroni, Salsa, 2022. Throughout this chapter, we mention the connection between these and our works in chapter 3 and chapter 4.

This chapter is structured as follows. Section 2.1 provides the fundamentals of partial differential equations and the corresponding spaces employed for the analysis of these equations. In section 2.2 we discuss the general framework for theoretical and numerical analysis of optimal control problems governed by partial differential equations.

2.1 Fundamentals of PDEs

All definitions and theorems in this section are based on Evans, 1998; Tröltzsch, 2010. Our comprehension of the fundamental processes in the natural world relies significantly on partial differential equations. Illustrative examples encompass the vibrations of solids, the flow of fluids, the diffusion of chemicals, the spread of heat, the interactions of photons and electrons, and the radiation of electromagnetic waves.

In the exploration of optimal control problems governed by partial differential equations (PDEs), our initial step involves the introduction of these foundational equations.

Throughout this chapter, $\Omega \subset \mathbb{R}^N$ is a domain, i. e., an open and connected set,whose boundary is generally denoted by $\partial\Omega$.

2.1.1 Partial Differential Equations

Definition 2.1.1. *Let $F\colon \Omega \times \mathbb{R} \times \mathbb{R}^N \times \cdots \times \mathbb{R}^{N^k} \to \mathbb{R}$ be given. An equation of the form*

$$F(x, u(x), Du(x), \dots, D^k u(x)) = 0 \quad x \in \Omega \tag{2.1.1}$$

is called a k^{th}-order partial differential equation, where $D^k u$ is defined in definition B.1.6 and $u\colon \Omega \to \mathbb{R}$ is the unknown. Solving the PDE means to find all u satisfying (2.1.1).

There are different classes of PDEs:

Definition 2.1.2. *(i) Let α be a multiindex of order $|\alpha|$ as defined in definition B.1.6. The partial differential equation (2.1.1) is said to be linear if it is of the form*

$$\sum_{|\alpha| \le k} a_\alpha(x)\, D^\alpha u = f(x)$$

for given functions a_α and f. It is called homogeneous if $f \equiv 0$.

(ii) The partial differential equation (2.1.1) is called semilinear if it is of the form

$$\sum_{|\alpha| = k} a_\alpha(x)\, D^\alpha u + a_0(x, u, Du, \dots, D^{k-1}u) = f(x)$$

(*iii*) *The partial differential equation* (2.1.1) *is quasilinear if it is of the form*

$$\sum_{|\alpha|=k} a_\alpha(x, u, Du, \dots, D^{k-1}u)\, D^\alpha u + a_0(x, u, Du, \dots, D^{k-1}u) = f(x)$$

(*iv*) *The partial differential equation* (2.1.1) *is called nonlinear if it depends nonlinearly upon the highest derivative.*

The theory of partial differential equations requires the spatial domains to have sufficiently smooth boundary.

2.1.2 Sobolev Spaces

The Hölder spaces, unfortunately, do not frequently serve as appropriate settings for a theoretical analysis of partial differential equations as well as for the analysis of some numerical methods for solving such equations. This is due to our typical difficulty to establish adequately accurate analytic estimates, which would demonstrate that our constructed solutions belong to such spaces. Instead, what is required are alternative types of spaces including functions with less smoothness, namely Sobolev spaces.

Weak Derivative

At first, we give the notion of weak derivatives and aim to extend the definition of derivatives.

Let Ω be a bounded Lipschitz domain. The classical integration by parts formula reads

$$\int_\Omega u(x) D^\alpha v(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega D^\alpha u(x) v(x)\,\mathrm{d}x$$

for $u \in C^k(\Omega)$, $v \in C_0^k(\Omega)$ and $|\alpha| \le k$.

The weak derivative is defined in such a way that the integration by parts formula holds, where ∂^α is now a weak differential operator applied to less smooth functions. To this end, we denote by $L^1_{\mathrm{loc}}(\Omega)$ the set of all locally integrable functions in Ω, meaning they are Lebesgue integrable on every compact subset of Ω.

Now we are ready to introduce the definition of a weak derivative.

Definition 2.1.3. *Let $u \in L^1_{loc}(\Omega)$ be given and α be some multi index. $w \in L^1_{loc}(\Omega)$ is termed a weak α^{th}- weak partail derivative of u, provided*

$$\int_\Omega u(x) D^\alpha v(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega w(x) v(x)\,\mathrm{d}x \quad \text{for all } v \in C_0^\infty(\Omega),$$

and we write $D^\alpha u = w$.

The function $v\colon \Omega \to \mathbb{R}$, which is infinitely differentiable with compact support in Ω, is called a test function. We denote the space of all test functions with $C_0^\infty(\Omega)$.

Lemma 2.1.4. *A weak derivative of u, if it exists, is uniquely defined up to a set of measure zero.*

Lemma 2.1.5. *If $u \in C^k(\Omega)$, then the classical partial derivative $D^\alpha u$, for each α with $|\alpha| \le k$, coincides with the α^{th}- weak partial derivative of u.*

Sobolev Spaces

Let $1 \le p \le \infty$ and k be a nonnegative integer. In the following we give the definition of a function space, including functions with weak derivative of various orders in L^p spaces.

A certain degree of regularity of the boundary $\partial\Omega$ of the domain Ω is required for some properties of Sobolev spaces.

Definition 2.1.6. *Suppose that Ω is a bounded domain in $\mathbb{R}^N$ and V denotes a function space on $\mathbb{R}^{N-1}$. $\partial\Omega$ is said to be of class V if for every point $x^0 \in \partial\Omega$, there exists an $r > 0$*

and some function $g \in V$ such that

$$\Omega \cap B(x^0, r) = \{x \in B(x^0, r) \,|\, x_n > g(x_1, \ldots, x_{n-1})\},$$

upon, if necessary, a relabeling and reorienting the coordinate system. Here, $B(x^0, r)$ denotes the n-dimensional open ball centered at x^0 with radius r.

When V particularly comprises Lipschitz continuous functions, namely $C^{0,1}$ functions, and C^k functions, Ω is called Lipschitz domain and C^k domain, respectively.

When V consists of $C^{k,\alpha}$ functions, $0 < \alpha \leq 1$, $\partial\Omega$ is called a Hölder boundary of class $C^{k,\alpha}$.

As $\partial\Omega$ is a compact set in $\mathbb{R}^N$, we we can identify a finite number of points $\{x^i\}_{i=1}^I$ on the boundary such that there exist positive numbers $\{r_i\}_{i=1}^I$ and functions $\{g_i\}_{i=1}^I \subset V$,

$$\Omega \cap B(x^i, r_i) = \{x \in B(x^i, r_i) \,|\, x_n > g_i(x_1, \ldots, x_{n-1})\}$$

upon a transformation of the coordinate system if necessary, and

$$\partial\Omega \subset \bigcup_{i=1}^{I} B(x_i, r_i).$$

Definition 2.1.7. *Let k be a nonnegative integer, $p \in [1, \infty]$. The Sobolev space $W^{k,p}(\Omega)$ consists of all functions $u \in L^p(\Omega)$ such that the weak derivatives $D^\alpha u$, for all multiindices α of length $|\alpha| \leq k$, exist and belong to $L^p(\Omega)$. If $u \in W^{k,p}(\Omega)$, its norm is defined by*

$$\|u\|_{W^{k,p}(\Omega)} := \begin{cases} \left[\displaystyle\sum_{|\alpha| \leq k} \|D^\alpha u\|^p_{L^p(\Omega)} \right]^{1/p} & 1 \leq p < \infty, \\ \displaystyle\max_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)} & p = \infty. \end{cases}$$

When $p = 2$, we write $H^k(\Omega) = W^{k,2}(\Omega)$.

Theorem 2.1.8. *The Sobolev space $W^{k,p}(\Omega)$ is a Banach space.*

A simple consequence of the theorem is the following result.

Corollary 2.1.9. *The Sobolev space $H^k(\Omega)$ is a Hilbert space with the inner product*

$$(u, v)_{H^k(\Omega)} = \int_\Omega \sum_{|\alpha| \leq k} D^\alpha u(x) D^\alpha v(x) \, \mathrm{d}x, \quad u, v \in H^k(\Omega).$$

It is easy to show that the Sobolev space $W^{k,p}(\Omega)$ is reflexive if $1 < p < \infty$.

The space $C_0^\infty(\Omega)$ does not need to be dense in $W^{k,p}(\Omega)$. So we introduce the following definition.

Definition 2.1.10. *$W_0^{k,p}(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ in $W^{k,p}(\Omega)$, that means, $u \in W_0^{k,p}(\Omega)$ if and only if there exist functions $u_n \in C_0^\infty(\Omega)$ such that $u_n \to u$ in $W^{k,p}(\Omega)$. When $p = 2$, we denote the Hilbert space $H_0^k(\Omega) \equiv W_0^{k,2}(\Omega)$.*

The functions in $W_0^{k,p}(\Omega)$ can be interpreted as functions $u \in W^{k,p}(\Omega)$ with the following property

$$D^\alpha u = 0 \text{ on } \partial\Omega \quad \text{for all } \alpha \text{ with } |\alpha| \leq k - 1.$$

The meaning of this statement will be made clear later after introducing the definition of trace operator in theorem 2.1.20.

Sobolev Spaces of Real Order

Definition 2.1.11. *Let Ω be a bounded domain and let $s = k + \lambda$ be a positive non-integer, with $k = [s]$ and $0 < \lambda < 1$. The Slobodetskii space $W^{s,p}(\Omega)$ is defined as the set*

$$\left\{ v \in W^{k,p}(\Omega) \;\middle|\; \frac{|D^\alpha v(x) - D^\alpha v(y)|}{|x-y|^{\lambda+n/p}} \in L^p(\Omega \times \Omega) \text{ for all } \alpha, \text{ with } |\alpha| = k \right\},$$

having the norm

$$\|v\|_{W^{s,p}(\Omega)} = \left[\|v\|^p_{W^{k,p}(\Omega)} + \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{|D^\alpha v(x) - D^\alpha v(y)|^p}{|x-y|^{n+\lambda p}} \,\mathrm{d}x\,\mathrm{d}y \right]^{1/p}.$$

The space $W^{s,p}(\Omega)$ is a Banach space. It is reflexive if and only if $p \in (1,\infty)$. When $p = 2$, we denote $H^s(\Omega)$. This a Hilbert space, equipped with the inner product

$$(u,v)_{H^s(\Omega)} = (u,v)_{H^k(\Omega)} + \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{[D^\alpha u(x) - D^\alpha u(y)]\,[D^\alpha v(x) - D^\alpha v(y)]}{|x-y|^{n+2\lambda}} \,\mathrm{d}x\,\mathrm{d}y.$$

The space $C_0^\infty(\Omega)$ is not, in general, dense in $W^{s,p}(\Omega)$, therefore it is useful to bring the following definition.

Definition 2.1.12. *Let $s \geq 0$. $W_0^{s,p}(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ in $W^{s,p}(\Omega)$. When $p = 2$, we denote the Hilbert space $H_0^s(\Omega) \equiv W_0^{s,2}(\Omega)$.*

With the spaces $W_0^{s,p}(\Omega)$, spaces with negative order can be defined.

Definition 2.1.13. *Let $s \geq 0$, either an integer or a non-integer. Let $p \in [1,\infty)$ and p^* be its conjugate exponent defined by $1/p + 1/p^* = 1$. $W^{-s,p^*}(\Omega)$ is defined as the dual space of $W_0^{s,p}(\Omega)$. In particular, we denote $H^{-s}(\Omega) \equiv W^{-s,2}(\Omega)$.*

Occasionally, we need to use the dual space of $H_0^1(\Omega)$.

Definition 2.1.14. *The dual space of $H_0^1(\Omega)$ is denoted by $H^{-1}(\Omega)$. If $f \in H^{-1}(\Omega)$, a bounded linear functional on $H_0^1(\Omega)$, we define the norm*

$$\|f\|_{H^{-1}(\Omega)} \coloneqq \sup\{\langle f\,, u\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} \,|\, u \in H_0^1(\Omega),\, \|u\|_{H_0^1(\Omega)} \leq 1\}.$$

We have the following Gelfand triplet, which is defined in appendix B.2.3

$$H_0^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega).$$

Then, any function $f \in L^2(\Omega)$ defines a bounded linear functional $f \in H^{-1}(\Omega)$ by the relation

$$\langle f\,, u\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = \int_\Omega f u \,\mathrm{d}x \quad \text{for all } u \in H_0^1(\Omega).$$

Sometimes even when f belong to $H^{-1}(\Omega)$, but not to $L^2(\Omega)$, we write the duality pairing $\langle f\,, u\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}$ as integral $\int_\Omega f u \,\mathrm{d}x$, although integration in this situation does not make sense.

Sobolev Spaces over Boundaries

Definition 2.1.15. *Let $k \geq 0$ be an integer, $\alpha \in (0,1]$, $s \in [0, k+\alpha]$ and $p \in [1,\infty)$. Assume a set of local representations of the boundary given by*

$$\partial\Omega \cap B(x^i, r_i) = \{x \in B(x^i, r_i) \,|\, x_n = g_i(x_1, \ldots, x_{n-1})\},$$

where g_i, $i = 1, \ldots I$, is defined on the open domain $D_i \subset \mathbb{R}^{N-1}$. We assume that every point of $\partial\Omega$ lies in at least one of these local representations. In addition, we assume

$g_i \in C^{k,\alpha}(D_i)$. *The Sobolev space* $W^{s,p}(\partial\Omega)$, $s \leq k+\alpha$ *is defined as follows*

$$W^{s,p}(\partial\Omega) = \{v \in L^2(\partial\Omega) \,|\, v \circ g_i \in W^{s,p}(D_i),\, i = 1, \ldots, I\},$$

endowed with the norm

$$\|v\|_{W^{s,p}(\partial\Omega)} = \max_i \|v \circ g_i\|_{W^{s,p}(D_i)}\,.$$

Sobolev Inequalities

There are embeddings of various Sobolev spaces into other spaces, which are the powerful analytic tools for the regularity of a weak solution of a boundary value problem.

Theorem 2.1.16 (General Sobolev inequalities). *Assume* Ω *is a Lipschitz domain and* $\partial\Omega \in C^1$. *Suppose* $u \in W^{k,p}(\Omega)$. *The we have the following compact embeddings, defined in definition B.2.5*

(*i*) *If* $kp < n$, *then*

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega) \quad \text{for } p \leq q \leq \frac{np}{n-kp}.$$

(*ii*) *If* $kp = n$, *then*

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega) \quad \text{for } p \leq q < \infty.$$

(*iii*) *If* $kp > n$, *then*

$$W^{k,p} \hookrightarrow\hookrightarrow C(\overline{\Omega}).$$

The embedding also result hold for non-integer $m > 0$, see Adams, 1975, theorem 7.57. A direct consequence of this theorem is the following compact embedding result.

Theorem 2.1.17 (Rellich). *Assume that* Ω *is a nonempty bounded Lipschitz domain. Let* k *and* l *be nonnegative integers with* $k > l$, *and let* $p \in [1, \infty]$. *Then*

$$W^{k,p}(\Omega) \hookrightarrow\hookrightarrow W^{l,p}(\Omega).$$

We can infer the following results from Adams, 1975:

Theorem 2.1.18. *Let* $1 < q \leq p \leq \infty$ *and* $s - \frac{n}{q} \geq t - \frac{n}{p}$. *Then the following continuous dense embedding holds:*

$$W^{s,q}(\Omega) \hookrightarrow W^{t,p}(\Omega) \quad s, t \geq 0.$$

Theorem 2.1.19. *Suppose that* $0 \leq s < \sigma$. *The Sobolev- Slobodetskii space* $H^\sigma(\Omega)$ *is continuously, compactly, and densely embedded in* $H^s(\Omega)$.

Traces

It is evident that the values of a function $u \in C(\overline{\Omega})$ on boundary of domain $\partial\Omega$ are understandable in the usual sense. Since a typical function $u \in W^{1,p}(\Omega)$ is only defined almost everywhere in Ω and the boundary $\partial\Omega$ has measure zero, the question arises how we can assign boundary values along $\partial\Omega$ to the function u.

This question is resolved by the notion of *trace operator*.

Theorem 2.1.20 (Trace Theorem). *Suppose* Ω *is a Lipschitz domain and let* $1 \leq p < \infty$. *There exists a continuous linear operator*

$$\tau \colon W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

with the following properties

(*i*) $\tau u = u\big|_{\partial\Omega}$ *if* $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$.

(*ii*) < *For every* $u \in W^{1,p}(\Omega)$

$$\|\tau u\|_{L^p(\partial\Omega)} \leq C \|u\|_{W^{1,p}(\Omega)},$$

where the positive constant C *depends only on* p *and* Ω.

(*iii*) *The mapping* $\tau\colon W^{1,p}(\Omega) \to L^p(\partial\Omega)$ *is compact.*

Definition 2.1.21. *We call* τu *the trace of* u *on* $\partial\Omega$.

The trace operator is neither an injection nor a surjection form $W^{1,p}(\Omega)$ to $L^p(\partial\Omega)$. Its range is smaller than $L^p(\partial\Omega)$, namely $W^{1-1/p,p}(\partial\Omega)$, a positive order Sobolev space over the boundary.

Theorem 2.1.22. *Let* $1 < p < \infty$, $k \geq 1$ *and suppose that* Ω *is a bounded domain of class* $C^{k,1}$. *Then there exist unique bounded linear and surjective mappings* $W^{k,p}(\Omega) \to W^{k-1/p,p}(\partial\Omega)$.

Theorem 2.1.23. *If* $s > \frac{1}{2}$, *then there exists a bounded trace operator which maps each function in* $H^s(\Omega)$ *to its boundary value in* $H^{s-1/2}(\partial\Omega)$.

A function can have zero trace in the following sense:

Theorem 2.1.24. *Suppose* Ω *is a bounded domain with* C^1 *boundary* $\partial\Omega$. *Furthermore, assume that* $u \in W^{1,p}(\Omega)$. *Then*

$$u \in W_0^{1,p}(\Omega) \quad \textit{if and only if } u = 0 \textit{ on } \partial\Omega.$$

Integration by Parts Formula

Here we collect some facts employed for calculus.

Suppose that the boundary of the domain $\partial\Omega$ is of class $C^{0,1}$. As already mentioned, the outward pointing unit normal vector is defined along $\partial\Omega$ denoted by

$$\nu = (\nu_1, \ldots, \nu_n)^T.$$

Theorem 2.1.25 (Gauss-Green Theorem). *Let* $u \in C^1(\overline{\Omega})$. *Then we have*

$$\int_\Omega u_{x_i} \,\mathrm{d}x = \int_{\partial\Omega} u\,\nu_i \,\mathrm{d}s,$$

where ν_i *denotes the* i*th component of the outward unit normal vector* ν *to the boundary.*

We can apply the Gauss-Green theorem to a vector-valued function and obtain the following variant, which is also called Divergence Theorem:

Theorem 2.1.26 (Divergence Theorem). *Let* $u \in C^1(\overline{\Omega})$. *Then we have*

$$\int_\Omega \operatorname{div} u \,\mathrm{d}x = \int_{\partial\Omega} u \cdot \nu \,\mathrm{d}s,$$

where $\operatorname{div} u = \sum_{i=1}^n \frac{\partial u_i}{\partial x_i}$.

Applying the Gauss-Green Theorem to a function product results in the following theorem.

Theorem 2.1.27. *Suppose* $u, v \in C^1(\overline{\Omega})$. *Then we have*

$$\int_\Omega u_{x_i} v \,\mathrm{d}x = -\int_\Omega u v_{x_i} \,\mathrm{d}x + \int_{\partial\Omega} uv\nu_i \,\mathrm{d}s.$$

The classical Gauss's formula can be extended to functions from certain Sobolev spaces so that the smoothness of the function is quite enough for the welldefinedness of the integrals in the Lebesgue sense, which can be proved by so-called density argument.

Proposition 2.1.28. *Let Ω be a Lipschitz domain. Then*

$$\int_\Omega u_{x_i} v \,\mathrm{d}x = -\int_\Omega u v_{x_i} \,\mathrm{d}x + \int_{\partial\Omega} u v \nu_i \,\mathrm{d}s \quad \text{for all } u, v \in H^1(\Omega).$$

For analyzing nonlinear problems, it is beneficial to extend this formula even further. Indeed, we have

$$\int_\Omega u_{x_i} v \,\mathrm{d}x = -\int_\Omega u v_{x_i} \,\mathrm{d}x + \int_{\partial\Omega} u v \nu_i \,\mathrm{d}s \quad \text{for all } u \in W^{1,p}(\Omega), v \in W^{1,p^*}(\Omega),$$

where $p, p^* \in (1,\infty)$ and p^* is the conjugate exponent, defined by $\frac{1}{p^*} + \frac{1}{p} = 1$.

Various other useful formulas can be derived. One of them is

$$\int_\Omega \Delta u v \,\mathrm{d}x = \int_{\partial\Omega} \partial_n u v \,\mathrm{d}s - \int_\Omega \nabla u \cdot \nabla v \quad \text{for all } u \in H^2(\Omega), v \in H^1(\Omega).$$

Here

$$\Delta u \coloneqq \sum_{i=1}^n u_{x_i x_i} = \operatorname{div} \nabla u$$

is the Laplacian operator,

$$\nabla u \coloneqq (u_{x_1}, \dots, u_{x_n})^T$$

is the gradient of u, and

$$\partial_n u \coloneqq \frac{\partial u}{\partial \nu} = \nabla u \cdot \nu$$

is the outward normal derivative.

2.1.3 Spaces Involving Time

This sort of Sobolev spaces, which are essential in linear and nonlinear parabolic PDEs, consist of functions mapping time into Banach spaces.

Definition 2.1.29. *Any mapping from a subset of $\mathbb{R}$ or $\mathbb{R}^N$ into a Banach space is termed a vector-valued function.*

First, we give a generalization of an integrable real-valued functions to vector-valued functions.

Let X and Y be Banach spaces.

Definition 2.1.30. (*i*) *A vector-valued function $s\colon X \to Y$ is said to be simple if there exist finitely many functions $u_i \in Y, 1 \le i \le k$ such that*

$$s(t) = \sum_{i=1}^k \chi_{E_i}(t) u_i, \quad t \in X,$$

where each E_i is a Lebesgue measurable subset of X.

(*ii*) *A function $f\colon X \to Y$ is called measurable if there exists a sequence $\{s_n\}_{n=1}^\infty$ of simple functions $s_n\colon X \to Y$ such that*

$$s_n(t) \to f(t) \quad \text{for a.e. } t \in X.$$

Definition 2.1.31. (*i*) *If $s(t) = \sum_{i=1}^k \chi_{E_i}(t) u_i$ is simple, we define*

$$\int_X s(t) \,\mathrm{d}t \coloneqq \sum_{i=1}^k |E_i| \, u_i.$$

(*ii*) *The measurable function is said to be integrable if there exists a sequence of simple functions* $\{s_n\}_{n=1}^{\infty}$ *such that*

$$\int_X f(t)\,\mathrm{d}t = \lim_{n\to\infty} \int_X s_n(t)\,\mathrm{d}t.$$

Definition 2.1.32. *The linear space*

$$L^p(0,T;X)$$

comprises all measurable vector-valued functions $u\colon [0,T] \to X$ *having the properties*

(*i*) *For* $1 \le p < \infty$

$$\|u\|_{L^p(0,T;X)} \coloneqq \left(\int_0^T \|u(t)\|_X^p\right)^{1/p} < \infty.$$

(*ii*) *For* $p = \infty$

$$\|u\|_{L^\infty(0,T;X)} \coloneqq \operatorname*{ess\,sup}_{0\le t\le T} \|u(t)\|_X < \infty.$$

Definition 2.1.33. *We denote by*

$$C([0,T];X)$$

the linear space of all continuous functions $u\colon [0,T] \to X$ *with*

$$\|u\|_{C([0,T];X)} \coloneqq \max_{0\le t\le T} |u(t)| < \infty.$$

Definition 2.1.34. *Let* $u \in L^1(0,T;X)$. $w \in L^1(0,T;X)$ *is called the weak derivative of* u *denoted by*

$$\partial_t u = w,$$

if

$$\int_0^T \partial_t v(t) u(t)\,\mathrm{d}t = -\int_0^T v(t) w(t)$$

for all real valued test functions $v \in C_0^\infty(0,T)$.

When u and $\partial_t u$ belong to different spaces $u \in L^2(0,T;H_0^1(\Omega))$ and $\partial_t u \in L^2(0,T;H^{-1}(\Omega))$, respectively the following statements hold:

Theorem 2.1.35. *Let* Ω *be a bounded Lipschitz domain,* $u \in L^2(0,T;H_0^1(\Omega))$ *and* $\partial_t u \in L^2(0,T;H^{-1}(\Omega))$.

(*i*) *We have*

$$u \in C([0,T];L^2(\Omega))$$

(*ii*) *The mapping*

$$t \mapsto \|u(t)\|_{L^2(\Omega)}^2$$

is absolutely continuous, with

$$\frac{d}{dt}\|u(t)\|_{L^2(\Omega)}^2 = 2\langle \partial_t u(u)\,, u(t)\rangle$$

for a.e. $0 \le t \le T$.

(*iii*) *There exists some constant* C *such that*

$$\max_{0\le t\le T} \|u(t)\|_{L^2(\Omega)} \le C\left(\|u\|_{L^2(0,T;H_0^1(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))}\right)$$

where C *depends only on* T.

Definition 2.1.36. *We define the parabolic cylinder*

$$\Omega_T := \Omega \times (0,T),$$

and the parabolic boundary of Ω_T

$$\partial\Omega_T := \partial\Omega \times (0,T).$$

2.1.4 Second Order Elliptic Equations

Let Ω be a bounded domain and $f\colon \Omega \to \mathbb{R}$ be given. A linear second order boundary value problem is of the following form

$$\begin{cases} Lu = f & \text{in } \Omega, \\ \partial_{\nu_{\mathcal{A}}} u = g & \text{on } \partial\Omega_1, \\ u = h & \text{on } \partial\Omega_2, \end{cases} \tag{2.1.2}$$

where the boundary $\partial\Omega = \partial\Omega_1 \cup \partial\Omega_2$ is split into two disjoint measurable sets $\partial\Omega_1$ and $\partial\Omega_2$, being relatively closed and open subsets of $\partial\Omega$.

L denotes a second-order partial differential operator having either the divergence form

$$Lu = -\sum_{i,j=1}^{n} \big(a_{ij}(x)u_{x_j}\big)_{x_i} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u \tag{2.1.3}$$

or the nondivergence form

$$Lu = -\sum_{i,j=1}^{n} a_{ij}(x)u_{x_j x_i} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u. \tag{2.1.4}$$

The values of u prescribed at each point on $\partial\Omega_2$ are called Dirichlet condition and the values of normal derivative of u prescribed at each point on $\partial\Omega_1$ are called Neumann conditions.

We note, an operator given in divergence form can be transformed into nondivergence form and vice versa, provided that the highest-order coefficients a_{ij}, $i,j = 1,\dots,n$ are C^1 functions.

We denote the first term in (2.1.3) and (2.1.4) by

$$\mathcal{A}u = -\sum_{i,j=1}^{n} (a_{ij}(x)u_{x_j})_{x_i} \quad \text{and} \quad \mathcal{A}u = -\sum_{i,j=1}^{n} a_{ij}(x)u_{x_i x_j},$$

respectively. The conormal vector is denoted by $\nu_{\mathcal{A}} = A\,\nu$, with the matrix function $A = (a_{ij})$, that means $\partial_{\nu_{\mathcal{A}}} u$ is given by

$$\partial_{\nu_{\mathcal{A}}} u = \sum_{i,j=1}^{n} a_{ij}\, u_{x_i}\, \nu_j.$$

We henceforth assume the coefficient functions a^{ij}, b^i, c, $i,j = 1,\dots,n$ belong to $L^\infty(\Omega)$.

Definition 2.1.37. *The partial differential operator L is said to be (uniformly) elliptic if there exists some constant $\theta > 0$ such that*

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \theta\,|\xi|^2$$

for a.e. $x \in \Omega$ and all $\xi \in \mathbb{R}^N$.

That means the symmetric matrix $A(x) = (a_{ij}(x))$ is positive definite for almost every point $x \in \Omega$, with the smallest eigenvalue greater than or equal to θ.

Weak Solution

For a Banach space V, let us first explore the relation between a linear operator $L\colon V \to V^*$ and a continuous bilinear form $a\colon V \times V \to \mathbb{R}$.

Theorem 2.1.38. *There exists a one-to-one correspondence between linear continuous operators $L\colon V \to V^*$ and continuous bilinear form $a\colon V \times V \to \mathbb{R}$, related by*

$$\langle Lu\,, v\rangle_{V^*,V} = a(u,v).$$

We consider the boundary-value problem

$$\begin{cases} Lu = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.1.5}$$

with the divergence form of the elliptic operator L.

Definition 2.1.39. *(i) The bilinear form $a(\cdot,\cdot)$ associated with L is defined by*

$$a(u,v) \coloneqq \int_\Omega \Big(-\sum_{i,j=1}^n a_{ij}(x) u_{x_i} v_{x_j} + \sum_{i=1}^n b_i(x) u_{x_i} v + c(x)uv \Big)\,\mathrm{d}x \tag{2.1.6}$$

for $u, v \in H_0^1(\Omega)$.

(ii) Let $f \in H^{-1}(\Omega)$ be given. $u \in H_0^1(\Omega)$ is said to be a weak solution of the boundary-value problem (2.1.5) provided

$$a(u,v) = \langle f\,, v\rangle \tag{2.1.7}$$

for all $v \in H_0^1(\Omega)$, where $\langle\cdot\,,\cdot\rangle$ is the duality pairing of $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

The identity (2.1.7) is termed the variational formulation of (2.1.5).

Existence of Weak Solutions

Let H be a real Hilbert space and $\langle\cdot\,,\cdot\rangle$ denote the pairing of H with its dual.

Theorem 2.1.40 (Lax-Milgram Theorem). *Suppose that*

$$a\colon H \times H \to \mathbb{R}$$

is a bilinear mapping, for which there exist constants $\alpha, \beta > 0$ such that

(i) a is H-bounded, that is

$$|a(u,v)| \le \alpha \|u\|_H \|v\|_H \quad \text{for all } u, v \in H,$$

(ii) a is H-coercive, that is

$$a(u,u) \ge \beta \|u\|_H^2 \quad \text{for all } u \in H.$$

Furthermore, assume that

$$f\colon H \to \mathbb{R}$$

is a bounded linear functional on H, i. e. $f \in H^*$.

Then there exists a unique element $u \in H$ such that

$$a(u,v) = \langle f\,, v\rangle_{H^*,H}$$

for all $v \in H$. Moreover, there exists some positive constant c, independent on f, such that

$$\|u\|_H \le c \|f\|_{H^*}.$$

The aforementioned specific bilinear form (2.1.6) satisfy the hypothesis of the Lax-Milgram Theorem.

Theorem 2.1.41 (Energy Estimates). *Let $a(\cdot,\cdot)$ be the bilinear form (2.1.6). There exists constants $\alpha, \beta > 0$ and $\gamma \ge 0$ such that*

(*i*)
$$|a(u,v)| \le \alpha \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)} \quad \text{for all } u, v \in H_0^1(\Omega),$$
(*ii*)
$$\beta \|u\|_{H_0^1(\Omega)} \le a(u,v) + \gamma \|u\|^2_{L^2(\Omega)} \quad \text{for all } u \in H_0^1(\Omega).$$

For $\gamma > 0$, where the hypotheses of Lax-Milgram Theorem can not be precisely satisfied, the following statement verifies the existence of weak solutions.

Theorem 2.1.42. *There exists a number $\gamma \ge 0$ such that for every $\mu \ge \gamma$ and every prescribed function $f \in H^{-1}(\Omega)$, there exists a unique weak solution $u \in H_0^1(\Omega)$ of the following boundary-value problem*
$$\begin{cases} Lu + \mu u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Regularity

The regularity problem for weak solutions addresses the question concerning whether the weak solution u of the PDE
$$Lu = f \quad \text{in } \Omega$$
is in fact smooth.

Let Ω be a bounded domain. Assume that $u \in H_0^1(\Omega)$ is a weak solution of the above PDE, where L has the divergence form
$$Lu = -\sum_{i,j=1}^{n} \left(a^{ij}(x) u_{x_i}\right)_{x_j} + \sum_{i=1}^{n} b^i(x) u_{x_i} + c(x)u$$
Suppose also L is uniformly elliptic.

Theorem 2.1.43 (Boundary H^2-Regularity). *Let $a^{ij} \in C^1(\overline{\Omega})$ and $b^i, c \in L^\infty(\Omega)$, $i, j = 1, \dots, n$. Assume that the boundary-value problem*
$$\begin{cases} Lu = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$
possesses a weak solution $u \in H_0^1(\Omega)$ for a right hand side $f \in L^2(\Omega)$. Furthermore, let $\partial\Omega \in C^{1,1}$. Then $u \in H^2(\Omega)$ and we the following estimate holds
$$\|u\|_{H^2(\Omega)} \le C \left(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \right)$$
where the constant C depends only on Ω and the coefficients of L.

We observe that the above estimate reduces to
$$\|u\|_{H^2(\Omega)} \le C \|f\|_{L^2(\Omega)}$$
provided that $u \in H_0^1(\Omega)$ is the unique weak solution of the PDE.

The Weak Maximum Principle

Let Ω be a bounded domain of class C^1. We consider the bounded bilinear form (2.1.6) on $H^1(\Omega)$, defining a bounded linear operator $L\colon H^1(\Omega) \to H^1(\Omega)^*$
$$\langle Lu\,, v \rangle_{H^1(\Omega), H^1(\Omega)^*} = a(u,v)$$
for $u, v \in H^1(\Omega)$.

We say the *weak maximum principle* holds for $L\colon H^1(\Omega) \to H^1(\Omega)^*$ if any function $u \in H^1(\Omega)$, satisfying
$$Lu \le 0 \quad \text{in } \Omega, \quad u \le 0 \quad \text{on } \partial\Omega,$$

is nonpositive. The condition $Lu \leq 0$ means that

$$a(u, v) \leq 0 \quad \text{for all } 0 \leq v \in V.$$

We have the following result, see Troianiello, 2013, Theorem 2.3.

Theorem 2.1.44. *If the bilinear form (2.1.6) is H^1-coercive, then the weak maximum principle holds for L.*

2.1.5 Nonlinear PDEs

Different methods are available for solving nonlinear PDEs, and one notable approach is the monotone operators method. This method serves as a generalization of Lax-Milgram, specifically applied to nonlinear operators. Another approach is the sub-super solution method, where the idea is to utilize the maximum principle and construct an increasing sequence of subsolutions. The limit of this sequence converges to the solution of the problem under consideration. One of the most important methods applied in this work for analyzing the solvability of such equations is the Banach fixed-point theorem.

Weak Solution

Kirchhoff Equation: The first PDE of interest in our work is a nonlocal nonlinear elliptic equation, namely, the stationary Kirchhoff equation:

$$\begin{cases} -\left(u + b \|\nabla y\|^2_{L^2(\Omega)}\right) \Delta y = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.1.8}$$

where $b \geq b_0$, $f \geq f_0$, for some positive numbers b_0 and f_0 and they belong to $L^\infty(\Omega)$.

If $u \geq u_a > 0$ and belongs to $L^2(\Omega)$, $u + b \|\nabla y\|^2_{L^2(\Omega)}$ is strictly positive, we can write (2.1.8) in the form

$$-\Delta y = \frac{f}{u + b \|\nabla y\|^2_{L^2(\Omega)}}. \tag{2.1.9}$$

We note that

$$\frac{1}{u + b \|\nabla y\|^2_{L^2(\Omega)}} \leq \frac{1}{u_a}$$

and therefore, $1/\left(u + b \|\nabla y\|^2_{L^2(\Omega)}\right)$ belongs to $L^\infty(\Omega)$.

We define the operator E

$$\begin{aligned} E &\colon H^1_0(\Omega) \times L^2(\Omega) \to H^{-1}(\Omega) \\ (y, u) &\mapsto -\Delta y - \frac{f}{u + b \|\nabla y\|^2_{L^2(\Omega)}}, \end{aligned}$$

where

$$\langle E(y, u), v \rangle_{H^{-1}(\Omega), H^1_0(\Omega)} = \int_\Omega \nabla y \cdot \nabla v \, \mathrm{d}x - \int_\Omega \frac{f}{u + b \|\nabla y\|^2_{L^2(\Omega)}} v \, \mathrm{d}x$$

for all $v \in H^1_0(\Omega)$. $y \in H^1_0(\Omega)$ is called a weak solution of this PDE if

$$E(y, u) = 0 \quad \text{in } H^{-1}(\Omega).$$

Chemotaxis System: The second PDE we focus on is a system of PDEs, namely, a system of nonlocal, nonlinear parabolic-elliptic chemotaxis equations.

$$\begin{cases} \partial_t y - \Delta y = -\chi \operatorname{div}(y \nabla w) + y \left(a_0 - a_1 y - a_2 \fint_\Omega y \,\mathrm{d}x \right) & \text{in } \Omega_T, \\ -\Delta w + \lambda w = y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega, \end{cases}$$

where $0 \leq u \in L^\infty(\partial\Omega_T)$.

We have the initial value $y(0) = y_0$, therefore y must be continuous in time. A formulation of the weak solution is the following:

For given functions $y_0 \in L^\infty(\Omega)$ and $u \in L^\infty(\partial\Omega)$, to find (y, w) satisfying

$$y \in L^2(0,T; H^1(\Omega)), \quad \partial_t y \in L^2(0,T; H^1(\Omega)^*), \quad w \in L^\infty(0,T; H^1(\Omega))$$

with

$$\begin{aligned} \int_\Omega \partial_t y \, \varphi \,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla \varphi \,\mathrm{d}x = \chi \int_\Omega y \nabla w \cdot \nabla \varphi \,\mathrm{d}x - \chi \int_{\partial\Omega} u \, y \, \varphi \,\mathrm{d}x \\ + \int_\Omega \left(a_0 - a_1 y - a_2 \fint_\Omega y \,\mathrm{d}x \right) y \, \varphi \,\mathrm{d}x \end{aligned} \tag{2.1.10}$$

and

$$\int_\Omega \nabla w \cdot \nabla \varphi \,\mathrm{d}x - \int_{\partial\Omega} u \, \varphi \,\mathrm{d}s + \lambda \int_\Omega w \, \varphi \,\mathrm{d}x = \int_\Omega y \, \varphi \,\mathrm{d}x \tag{2.1.11}$$

for all $\varphi \in C^\infty(\overline{\Omega})$ and for a.a. $t \in [0,T]$ and $y(x,0) = y_0(x)$.

Obviously, the test function φ can be chosen from $H^1(\Omega)$, due to the density of $C^\infty(\overline{\Omega}) \subset H^1(\Omega)$.

If y merely belongs to $L^2(0,T; H^1(\Omega))$ the condition $y(x,0) = y_0(x)$ need not make sense. The following result will help, see e. g. Temam, 1984, Chapter 3, Section 2.2, Theorem 2.1.

Theorem 2.1.45. *Let X_0, X, X_1 be three Banach spaces such that*

$$X_0 \hookrightarrow X \hookrightarrow X_1,$$

where the embeddings are continuous, X_0 and X_1 are reflexive and the embedding $X_0 \hookrightarrow\hookrightarrow X$ is compact.

Let $T > 0$ be a fixed finite number, and let α_0, α_1 be two finite numbers such that $\alpha_i > 1$, $i = 1, 2$. We consider the space

$$\mathcal{Y} = \mathcal{Y}(0,T; \alpha_0, \alpha_1; X_0, X_1) = \{ v \in L^{\alpha_0}(0,T; X_0) \,|\, \partial_t v \in L^{\alpha_1}(0,T; X_1) \}.$$

The space $\mathcal{Y}$ is a Banach space with the norm

$$\|y\|_{\mathcal{Y}} = \|v\|_{L^{\alpha_0}(0,T;X_0)} + \|\partial_t v\|_{L^{\alpha_1}(0,T;X_1)}.$$

Moreover, $\mathcal{Y}$ is continuously embedded in $C([0,T]; X_1)$ and the embedding $\mathcal{Y}$ into $L^{\alpha_0}(0,T;X)$ is compact.

With $\alpha_0 = \alpha_1 = 2$ and considering the following Gelfand triple

$$V \hookrightarrow H \hookrightarrow V^*$$

the space $\mathcal{Y}$ will be denoted by $W(0,T;V,V^*)$, that is

$$W(0,T;V,V^*) \coloneqq \{ y \in L^2(0,T;V) \,|\, \partial_t y \in L^2(0,T;V^*) \}.$$

This space is composed of functions, which belong to $L^2(0,T;H^1(\Omega))$ and whose partial derivatives with respect to time belong to $L^2(0,T;H^1(\Omega)^*)$.

We note that

$$W(0,T;V,V^*) \hookrightarrow C([0,T];H).$$

A special case of this space, where $V = H^1(\Omega)$ and $H = L^2(\Omega)$ is the well-known space $W(0;T)$, that is

$$W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) \,|\, \partial_t y \in L^2(0,T;H^1(\Omega)^*)\}.$$

Existence of Weak Solution

After we defined a weak solution, we have to analyze existence of such solutions.

Fixed Point Theorems

Let K be a subset of a Banach space X. We consider operator

$$T\colon K \to X$$

The solutions of the equation

$$u = T(u), \quad u \in K$$

are termed fixed points.

The first question is why we are interested in the solving of the fixed-point problem

$$u = T(u), \quad u \in K. \tag{2.1.12}$$

Consider the operator $f\colon K \subset X \to X$. For solving an equation

$$f(u) = 0 \tag{2.1.13}$$

we can set

$$T(v) = v - F(f(v))$$

with an operator $F\colon X \to X$ satisfying

$$F(y) = 0 \quad \text{if and only if } y = 0.$$

Therefore, (2.1.13) can be regarded as a fixed-point problem.

The resolution of a nonlinear partial differential equation can be reduced to solving a nonlinear equation in $\mathbb{R}$.

Before introducing the Banach fixed point theorem in Banach spaces, we present a finite-dimensional version of a particular fixed-point theorem.

Theorem 2.1.46 (Brouwer's Fixed Point). *Suppose that $K \subset \mathbb{R}^N$ is a bounded, closed and convex set. Suppose, further that the operator $T\colon K \to K$ is continuous. Then T has a fixed point in K.*

Kirchhoff Equation: The existence of a solution of the Kirchhoff equation can be shown by applying Brouwer's fixed point theorem.

We define $g(s) \coloneqq s - \int_\Omega |\nabla y_s|^2 \,\mathrm{d}x$, where y_s is the unique solution of the Poisson problem

$$\begin{cases} -\Delta y_s = \dfrac{f}{u + b\,s} & \text{in } \Omega, \\ y_s = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.1.14}$$

Multiplying (2.1.14) with y_s as test function, we obtain

$$\int_\Omega |\nabla y_s|^2 \,\mathrm{d}x = \int_\Omega \frac{f}{u + b\,s}\, y_s \,\mathrm{d}x.$$

Since $g(s) = 0$ if and only if y_s solves (2.1.14), then the problem of solving $g(s) = 0$ can be reduced to finding a fixed point for the following operator

$$T(s) = \int_\Omega \frac{f}{u + bs} y_s \,\mathrm{d}x.$$

However, the defined function $g(s) = s - \|\nabla y_s\|^2$ has as many solutions as the Kirchhoff equation. Therefore, for the existence and uniqueness of the solution of the Kirchhoff equation we will directly use a monotonicity argument to show that g has a unique root in theorem 3.1.6.

Chemotaxis System: In this case Banach's fixed point theorem is applied to nonlinear PDEs with a perturbation. First, we indicate the definition of a contractive operator.

Definition 2.1.47. *Let V be a Banach space. The mapping $A\colon K \subset V \to V$ is said to be a strict contraction if there exists some constant $0 \leq \gamma < 1$ such that*

$$\|Au_1 - Au_2\|_V \leq \gamma \|u_1 - u_2\|_V \quad \text{for all } u_1, u_2 \in K.$$

We note that contractivity implies Lipschitz continuity.

Theorem 2.1.48 (Banach Fixed Point). *Assume that X is a nonempty closed set in a Banach space V and*

$$A\colon X \to X$$

be a nonlinear mapping. Furthermore, suppose that A is a strict contraction. Then A has a unique fixed point.

To analyze the existence of the solution of the chemotaxis system we apply Banach Fixed-Point theorem. We consider the following perturbed problem

$$\begin{cases} \partial_t y - \Delta y = -\chi \nabla y \cdot \nabla w + y \left(-\chi \lambda w + \chi \tilde{y} + a_0 - a_1 \tilde{y} - a_2 \fint_\Omega \tilde{y} \, \mathrm{d}x \right) & \text{in } \Omega_T, \\ -\Delta w + \lambda w = \tilde{y} & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega, \end{cases} \tag{2.1.15}$$

with $\tilde{y} \in \{z \in C([0,T]; L^2(\Omega)) \,|\, 0 \leq z \leq M\}$, and we show that operator $A\colon \tilde{y} \mapsto y$ is a strict contradiction in theorem 4.1.13.

Regularity

At times, the regularity of a function may not be sufficient for our optimal control problem. In such cases, we need to attain a higher regularity.

Kirchhoff Equation: To have a weak solution $y \in H_0^1(\Omega)$ it suffices that f belongs to $L^2(\Omega)$, by virtue of theorem 2.1.42. However, if f belongs to $L^p(\Omega)$, then the right-hand side in (2.1.14) belongs also to $L^p(\Omega)$. In this case we can have an intermediate situation of solutions, namely strong solution, provided that the domain is sufficiently smooth. While a weak solution need only be once weakly differentiable a strong solution is twice weakly differentiable satisfying (2.1.14) almost everywhere in Ω. Indeed, we have the following result, see e. g. Gilbarg, Trudinger, 1977, Theorem 9.15, Theorem 9.17.

Theorem 2.1.49. *Suppose that Ω is a bounded $C^{1,1}$ domain and $\mathcal{A}$ is an elliptic differential operator of the form*

$$\mathcal{A}y(x) = -\sum_{i,j=1}^{n} (a_{ij}(x) y_{x_j}(x))_{x_i} \quad x \in \Omega. \tag{2.1.16}$$

The coefficient functions a_{ij} of $\mathcal{A}$ are assumed to belong to $C^{0,1}(\overline{\Omega})$ and satisfy the symmetric condition $a_{ij}(x) = a_{ji}(x)$ for all $i, j \in \{1, \ldots, n\}$ and $x \in \Omega$.

If the right-hand side function $f \in L^p(\Omega)$, then the weak solution to the following Dirichlet problem

$$\begin{cases} \mathcal{A}y = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases}$$

belongs to $W^{2,p}(\Omega)$.

By virtue of this theorem, Kirchhoff equation has a strong solution $y \in W^{2,p}(\Omega)$.

Chemotaxis System: In the following we assume that $\mathcal{A}$ is a symmetric elliptic operator of the form (2.1.16) and $a_{ij} \in C^{0,1}(\overline{\Omega})$. Moreover, $\lambda \in \mathbb{R}$ is prescribed. We have the following result as a consequence of Grisvard, 1985, Theorem 2.4.2.7:

Theorem 2.1.50. *Let Ω be a bounded $C^{1,1}$ domain, $\lambda > 0$ and $y \in L^p(\Omega)$ be given. The following Neumann problem*

$$\begin{cases} \mathcal{A}w + \lambda w = y & \text{in } \Omega, \\ \partial_{\nu_{\mathcal{A}}} w = u & \text{on } \Omega, \end{cases}$$

has a unique weak solution $w \in W^{2,p}(\Omega)$, provided that $u \in W^{1-1/p,p}(\partial\Omega)$.

If the control u belongs to $H^1(\partial\Omega)$, we can employ theorem 2.1.18 resulting in $u \in W^{1-1/p,p}(\partial\Omega)$.

However, we preferred to work with the controls from $L^\infty(\partial\Omega)$. Fortunately, there is another result from Morrey, 1966, Chapter 5, Section 5.5.

Theorem 2.1.51. *Let Ω be a bounded C^1 domain. For given right-hand side functions $y \in W^{1,6/5}(\Omega)^*$ and $u \in L^6(\partial\Omega)$, the following PDE*

$$\begin{cases} -\Delta w + \lambda\, w = y & \text{in } \Omega, \\ \partial_n w = u & \text{on } \partial\Omega, \end{cases}$$

possesses a unique weak solution $w \in W^{1,6}(\Omega)$.

We note that, we cannot define a normal trace for a first-order Sobolev function without further information, in general. This only works if the Laplacian of the function (defined as a distribution) is an integrable function. This would then be the solution w from the equation. Then you can use it to define a normal trace in a negative Sobolev space, per Gauss theorem. With the weak formulation of the equation for w it then turns out that this normal trace coincides with the u pointwise.

There is also the following result that will help and can be deduced from Grisvard, 1985, Theorem 2.2.2.5.

Theorem 2.1.52. *Let Ω be a $C^{1,1}$ domain, $\lambda > 0$ and $y \in L^2(\Omega)$ be given. The following Neumann problem*

$$\begin{cases} \mathcal{A}w + \lambda w = y & \text{in } \Omega, \\ \partial_{\nu_{\mathcal{A}}} w = 0 & \text{on } \Omega, \end{cases}$$

has a unique weak solution $w \in H^2(\Omega)$.

2.2 Fundamentals of Optimal Control of PDEs

All definitions and results without reference in this section are based on Tröltzsch, 2010; Manzoni, Quarteroni, Salsa, 2022.

Problems arising from applied sciences are frequently represented by PDEs, depending on a set of input data. This data includes physical or material coefficients, boundary and initial conditions, source terms, as well as the geometrical configuration describing the domain where the problem is formulated, which can be regarded as input itself. Frequently,

a problem governed by PDEs, typically referred to as the state system, needs to be controlled or optimized by acting on (one, or more of) these input variables. This is a difficult mathematical task that can involve significant computational challenges.

2.2.1 Optimal Control of PDEs

In the context of addressing a forward problem, the goal is to compute the solution to a specified PDE, referred to as the *state system*. Conversely, solving an optimal control problem governed by a PDE entails the minimizing (or maximizing) of a physical quantity known as the *cost functional*. This functional depends on the PDE solution itself, and is influenced by some suitable control variables; these latter are some of the *data* required by the PDE.

An optimal control problem has the following essential features:

- a *cost functional* to be minimized
- a *control function* exerted on the system under consideration
- a *state problem* (forward problem) explaining the relation between the control and the state variables
- possibly, some constraints on the control (and state).

The solution of the state problem, denoted by y, depends on the control variable u. In this way, the state problem associates to every control a state solution $y = y(u)$. We aim to choose the control u in such a way that the observed variable y approximates a desired value y_d, so-called desired state. An additional (optional) element involves constraints that may act on the control and/or the state, referred to as control constraints or state constraints, respectively. The latter will depend on the problem at hand and will be made precise from case to case.

An illustrative example is the optimal control for heat transfer, represented by a body occupying a region (the domain) $\Omega \subset \mathbb{R}^3$ that requires heating or cooling. In this scenario, the state variable y represents the temperature, and the stationary heat equation models the state system. To regulate the temperature, a heat source u is applied to the volume, serving as the control function. The objective is to select u in a way that the resulting temperature distribution $y = y(u)$ in the domain closely approximates a specified target temperature, while minimizing the effort required for heating or cooling the system. Clearly, the state y, the control u and the target y_d are all functions of the spatial coordinates $x \in \Omega$.

This problem can be modeled by seeking for the solution of the following minimization problem for the cost function

$$\begin{aligned} &\text{Minimize } J(y,u) = \frac{1}{2}\left\|y - y_d\right\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\left\|u\right\|^2_{L^2(\Omega)} \\ &\text{subject to } \begin{cases} -\Delta y = u & \text{in } \Omega, \\ \quad y = 0 & \text{on } \partial\Omega, \end{cases} \end{aligned} \tag{2.2.1}$$

given by a Dirichlet boundary problem for the stationary heat equation.

2.2.2 Optimal Control of Linear Elliptic Equations

Let Y and U be two Hilbert spaces for the state variables and control variables, respectively and $\mathcal{U}_{\text{ad}} \subset U$ be a convex and closed set of admissible controls.

An elliptic boundary value problem can be transformed into an abstract variational problem: find $y = y(u) \in Y$ such that

$$a(y,\varphi) = \langle f\,,\varphi\rangle_{Y^*,Y} \quad \text{for all } \varphi \in Y, \tag{2.2.2}$$

where $a(\cdot,\cdot)\colon Y \times Y \to \mathbb{R}$ is a continuous bilinear form and $f \in Y^*$. Then, there exists a continuous linear mapping $A\colon Y \to Y^*$ satisfying

$$\langle Ay\,, \varphi\rangle_{Y^*,Y} = \langle f\,, \varphi\rangle_{Y^*,Y}.$$

Thus, the variational equality (2.2.2) can be written as

$$Ay = f \quad \text{in } Y^*.$$

Suppose that a is additionally coercive. Thanks to the Lax-Milgram theorem, for each $u \in U$, this equation has a unique solution $y \in Y$, which means A is bijective. Furthermore, there exist some constant $c > 0$, independent of f, such that

$$\|y\|_Y \leq c\,\|f\|_{Y^*} \tag{2.2.3}$$

showing the continuity of A. The inverse operator $A^{-1}\colon Y^* \to Y$ is continuous as well, by virtue of open mapping theorem.

We consider the following optimal control problem

$$\begin{aligned} \text{Minimize} \quad & J(y,u) \\ \text{subject to} \quad & Ay - Bu = 0 \\ \text{and} \quad & u \in \mathcal{U}_{\text{ad}}, \end{aligned}$$

where $J\colon Y \times U \to \mathbb{R}$, and $B\colon U \to Y^*$ is a linear operator, so-called *control operator*. Since A is invertible, the state can be expressed uniquely in terms of the control

$$y = A^{-1}Bu.$$

The mapping $S\colon U \to Y$

$$u \mapsto y = S(u) = A^{-1}Bu$$

is called *control-to-sate operator*.

The well-posedness of the control-to-state operator means that to every given control $u \in \mathcal{U}_{\text{ad}} \subset U$ there exists a unique solution of the state equation, called the associated state.

After showing the well-posedness of control-to-state map we have to prove the existence of an *optimal solution*.

Definition 2.2.1. *The pair $(\bar{y},\bar{u})$ is said to be optimal solution provided that $\bar{u} \in \mathcal{U}_{\text{ad}}$ with the associated state $\bar{y} = \bar{y}(\bar{u})$ satisfy*

$$J(\bar{y},\bar{u}) \leq J(y(u),u).$$

We call $\bar{u}$ optimal control and $\bar{y} = y(\bar{u})$ the associated optimal state.

Optimality System

If the control-to-sate map is well-defined we can rewrite the cost functional as follows

$$j(u) \coloneqq J(y,u) = J(S(u),u) = J(A^{-1}Bu,u),$$

which is called the *reduced cost functional*. We can thus transform the optimal control problem into an optimization problem

$$\min_{u\in\mathcal{U}_{\text{ad}}} j(u). \tag{2.2.4}$$

There is the following fundamental result to derive the optimality condition in the presence of control constraints.

Theorem 2.2.2. *Let C be a nonempty and convex subset of a Banach space U and $\mathcal{U}$ be an open set, with $C \subset \mathcal{U} \subset U$. Moreover, let $j\colon \mathcal{U} \to \mathbb{R}$ be Gâteaux differentiable in $\mathcal{U}$. If $\bar{u} \in C$ is a solution of*

$$\min_{u\in C} j(u),$$

then $\bar{u}$ satisfies the variational inequality

$$\langle j'(\bar{u}), u - \bar{u}\rangle_{U^*\times U} \geq 0 \quad \textit{for all } u \in C. \tag{2.2.5}$$

Conversely, $\bar{u}$ is a solution of the above minimization problem, provided that $\bar{u}$ satisfies the variational inequality (2.2.5) and j is convex.

In the case of convexity the first-order necessary optimality condition is sufficient.

Taking derivative of $j(u) = J(A^{-1}Bu, u)$, we obtain

$$\langle j'(u), v\rangle_{U^*,U} = \langle J_y(y,u), A^{-1}Bv\rangle_{Y^*,Y} + \langle J_u(y,u), v\rangle_{U^*,U}.$$

Let $f_y \in Y^*$ be given. To $A^{-1}\colon Z \to Y$ we correspond the mapping $(A^{-1})^*\colon Y^* \to Z^*$ such that

$$\langle (A^{-1})^* f_y, z\rangle_{Z^*,Z} = \langle f_y, A^{-1}z\rangle_{Y^*,Y}.$$

Then

$$\langle j'(u), v\rangle_{U^*,U} = \langle (A^{-1})^* J_y(y,u), Bv\rangle_{Z^*,Z} + \langle J_u(y,u), v\rangle_{U^*,U}.$$

The operator $(A^{-1})^*$ is referred to as *adjoint operator* and $p := (A^{-1})^* J_y(y,u) \in Z^*$ is called *adjoint state*.

For an optimal solution $(\bar{y}, \bar{u})$, together with the adjoint state p, the first-order optimality system reads

$$\begin{cases} A^*p = J_y(\bar{y}, \bar{u}), \\ A\bar{y} = B\bar{u}, \\ \langle B^*p + J_u(\bar{y}, \bar{u}), u - \bar{u}\rangle_{U,U^*} \geq 0 \quad \text{for all } u \in \mathcal{U}_{\text{ad}}. \end{cases}$$

2.2.3 Optimal Control of Nonlinear Elliptic Equation

So far we have considered linear-quadratic Optimal Control Problems for elliptic Partial Differential Equations (PDEs) in a Hilbert space framework. The well-posedness analysis and derivation of first-order necessary optimality conditions have been achieved. This involved expressing the reduced cost functional as a quadratic functional and exploiting fundamental properties like the continuity and coercivity of certain bilinear forms.

To address a wider range of applications, we need to develop a more general framework for the analysis of optimal control problems where the state equation may comprise a nonlinear PDE, or the cost functional is no longer quadratic. Typical examples of nonlinear PDEs arise, e. g., in fluid dynamics (Navier-Stokes equations) and in structural mechanics (non-elastic materials).

Consider the following general optimal control problem

$$\begin{aligned} \min \quad & J(y,u) \\ \text{subject to} \quad & E(y,u) = 0, \\ \text{and} \quad & u \in \mathcal{U}_{\text{ad}}. \end{aligned}$$

Where $J\colon Y \times U \to \mathbb{R}$, $E\colon Y \times U \to Z$, and Y, U, Z are reflexive Banach spaces and $\mathcal{U}_{\text{ad}}$ is a closed and convex set and represents a *control constraint*. A point (y,u) is called *feasible* if it belongs to $Y \times \mathcal{U}_{\text{ad}}$ and $E(y,u) = 0$.

Exploring the behavior of the control-to-state map $S\colon u \mapsto y$, when it involves a nonlinear term, poses a significant challenge in investigating these types of problems. To deal with the aforementioned challenge we have to be able to answer the following questions:

(*i*) Is the control-to-state map S well-defined?

(*ii*) If the answer to the first question is positive the next question will be whether S is differentiable or not?

In linear problems S is linear and continuous, thereby being differentiable, and its derivative coincides with the operator itself.

Welldefindness of Control-to-State Map

To answer the first question we have to prove that the control-to-sate map

$$\begin{aligned} S\colon \mathcal{U}_{\text{ad}} &\to Y \\ u &\mapsto y(u) = S(u) \end{aligned}$$

assigns to each $u \in \mathcal{U}_{\text{ad}}$ a unique weak solution $y(u)$ to $E(y,u) = 0$.

Optimal Control of KE: We consider the following optimal control problem

$$\begin{aligned} \text{Minimize} \quad & J(y,u) \coloneqq \int_\Omega \varphi(x, y(x))\,\mathrm{d}x + \frac{\lambda}{2}\|u\|^2_{H^1(\Omega)} \\ \text{subject to} \quad & \begin{cases} -\left(u + b\,\|\nabla y\|^2_{L^2(\Omega)}\right)\Delta y = f & \text{in } \Omega, \\ \qquad\qquad\qquad\qquad\quad y = 0 & \text{on } \partial\Omega \end{cases} \\ \text{and} \quad & u \in \mathcal{U}_{\text{ad}} = \{u \in L^2(\Omega) \,|\, u_a(x) \le u(x) \text{ a.e. in } \Omega\}. \end{aligned} \tag{2.2.6}$$

The welldefindness of the control-to-state map means for a given control $u \in \mathcal{U}_{\text{ad}}$, there exists a unique weak solution $y \in H^1_0(\Omega) \cap W^{2,q}(\Omega)$ of the Kirchhoff problem.

As mentioned before the existence of the solution can be shown by means of a monotonicity argument for the controls $u \in \mathcal{U}_{\text{ad}}$. That means the control-to-state map

$$S\colon \mathcal{U}_{\text{ad}} \to H^1_0(\Omega) \cap W^{2,q}(\Omega)$$

is well-defined.

Optimal Control of CS: We consider the following optimal control problem

$$\begin{aligned} \text{Minimize} \quad & J(y,u) \coloneqq \frac{1}{2}\int_\Omega |y(x,T) - y_d(x)|^2\,\mathrm{d}x + \frac{\gamma}{2}\int_0^T \int_{\partial\Omega} |u(x,t)|^2\,\mathrm{d}s\,\mathrm{d}t \\ \text{subject to} \quad & \begin{cases} \partial_t y - \Delta y = -\chi\,\operatorname{div}(y\nabla w) + y\left(a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right) & \text{in } \Omega_T, \\ -\Delta w + \lambda\, w = y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega \end{cases} \\ \text{and} \quad & u \in \{u \in L^\infty(\partial\Omega_T) \,|\, 0 \le u(x,t) \le u_b(x,t) \text{ a.e. on } \partial\Omega_T\}. \end{aligned} \tag{2.2.7}$$

As mentioned earlier, the existence of the solution of the chemotaxis system can be demonstrated through the Banach Fixed-Point theorem for controls $u \in \mathcal{U}_{\text{ad}}$. This implies the well-definedness of the control-to-state map

$$S\colon \mathcal{U}_{\text{ad}} \to W(0,T) \times L^\infty(0,T; W^{1,6}(\Omega)).$$

Existence of Optimal Controls

After finding a unique weak solution, we must address the fundamental question: Does the optimal control problem possess a solution, specifically, an optimal control with an associated optimal state?

We consider again the following constrained optimization problem for proving the existence of an optimal solution:

$$\begin{aligned} \min \quad & J(y,u) \\ \text{subject to} \quad & E(y,u)=0 \\ \text{and} \quad & u \in \mathcal{U}_{\text{ad}}, \end{aligned}$$

The main steps in proving the existence an optimal solution are as follows:

(*i*) To show the existence of a minimizing sequence $\{(y_n,u_n)\}$ of feasible points, that is

$$\lim_{n\to\infty} J(y_n,u_n) = \inf_{u\in\mathcal{U}_{\text{ad}}} J(y,u),$$

where $y_n = S(u_n)$. We can achieve this by showing the boundedness of the cost functional J from below.

(*ii*) To find some candidate $(\bar{y},\bar{u})$ for optimal solution. Here we need to show the boundedness of the minimizing sequence $\{(y_n,u_n)\}$ in $Y\times U$.

(*iii*) To show that $\bar{u}$ is admissible and $\bar{y} = S(\bar{u})$. At this point, it is required to show that the set of feasible points is weakly sequentially closed in $Y\times U$.

(*iv*) To show the weakly lower semi continuity of J, that is

$$J(\bar{y},\bar{u}) \le \liminf_{n\to\infty} J(y_n,u_n).$$

Optimal Control of KE: In optimal control of the Kirchhoff equation we face a challenge when the controls belong to $L^2(\Omega)$. Indeed, $u_n \rightharpoonup u$ in $L^2(\Omega)$ does not imply $S(u_n) \rightharpoonup S(u)$ in $W^{2,q}(\Omega)$.

That is the point why we switch from $L^2(\Omega)$ to $H^1(\Omega)$. Indeed, when $\{u_n\}\subset\mathcal{U}_{\text{ad}}$ with $u_n \rightharpoonup u$ in $H^1(\Omega)$ there exists a subsequence, denoted by the same indices, with $u_n \to u$ in $L^2(\Omega)$. The continuity of the control-to-state map with the $L^2(\Omega)$-topology for the controls can be proved, see theorem 3.1.7.

In what follows, we bring some helpful definitions and results:

Definition 2.2.3. *Let X be a real Banach space and M be a subset of X.*

(*i*) *Let M be a subset of X. M is called sequentially compact if every sequence $\{u_n\}_{n=1}^{\infty}\subset M$ contains a convergent subsequence, with limit in M.*

(*ii*) *The set M is called sequentially relatively compact if it has compact closure in X.*

In a finite-dimensional space, it is well-known that a bounded sequence has a convergent subsequence. In an infinite-dimensional space, we expect only a weaker property; but even the weaker property is still useful in proving many existence results.

Definition 2.2.4. *Let X be a real Banach space and M be a subset of X.*

(*i*) *The set M is called weakly sequentially closed if for every weakly convergent sequence $\{u_n\}_{n=1}^{\infty}\subset M$ to some $u\in X$, we can imply $u\in M$.*
Any weakly sequentially closed set is also strongly closed.

(*ii*) *The set M said to be weakly sequentially relatively compact if every sequence $\{u_n\}_{n=1}^{\infty}\subset M$ contains a weakly convergent subsequence in X.*

(*iii*) *The weakly sequentially relatively compact set M is called weakly sequentially compact, if it is weakly sequentially closed.*

Theorem 2.2.5. *Every bounded subset M of a reflexive Banach space X is weakly sequentially relatively compact. That means, for any bounded sequence $\{u_n\}_{n=1}^{\infty}\subset M$ there exists a subsequence $\{u_{n_k}\}_{k=1}^{\infty}\subset\{u_n\}_{n=1}^{\infty}$ such that*

$$u_{n_k} \rightharpoonup u,$$

where $u\in X$.

In particular, this statement holds for a Hilbert space.

Theorem 2.2.6. *Let X be a Banach space and $M \subset X$ some closed and convex subset of X. Then M is weakly sequentially closed. If X is reflexive and M is additionally bounded in X, then M is weakly sequentially compact.*

Theorem 2.2.7. *Let X be a Banach space and $M \subset X$. Suppose $f\colon M \to \mathbb{R}$ is a continuous and convex function. Then f is weakly sequentially lower semicontinuous, that is, any weakly convergent sequence $\{u_n\}_{n=1}^{\infty} \subset M$, $u_n \rightharpoonup u \in M$ implies*

$$f(u) \leq \liminf_{k\to\infty} f(u_n).$$

It is easily seen that if a function is weakly lower semicontinuous, then it is lower semicontinuous.

Since the cost functional in (2.2.6) is comprised of a Nemytskii operator, we need to study the well-definedness, continuity, and differentiability of such functions.

Nemytskii Operator

All materials about Nemytskii operator are based on Tröltzsch, 2010.

Definition 2.2.8 (Nemytskii Operator). *Let Ω be a bounded and measurable set and assume the function $\varphi = \varphi(x, y)\colon \Omega \times \mathbb{R} \to \mathbb{R}$. The mapping*

$$\Phi\colon y(\cdot) \mapsto \varphi(\cdot, y(\cdot)),$$

which assigns to a function $y(\cdot)\colon \Omega \to \mathbb{R}$ the function $\varphi(\cdot, y(\cdot))\colon \Omega \to \mathbb{R}$ is said to be a Nemytskii operator or superposition operator.

Here we intend to address this question between which spaces this mapping is well-defined or rather continuous and differentiable.

Definition 2.2.9. *We consider the function $\varphi = \varphi(x, y)$*

(*i*) *φ is called Carathéodory if it is for any fixed $y \in \mathbb{R}$ measurable with respect to x and for almost every fixed $x \in \Omega$ continuous with respect to y.*

(*ii*) *φ is said to satisfy the boundedness condition if there exist some constants $K > 0$ such that*

$$|\varphi(x, 0)| \leq K \quad \text{for a.e. } x \in \Omega. \tag{2.2.8}$$

(*iii*) *φ is called locally Lipschitz continuous with respect to y if for any positive constant M there exists some positive constant $L(M)$ such that the estimate*

$$|\varphi(x, y) - \varphi(x, z)| \leq L(M)\,|y - z|\,, \tag{2.2.9}$$

holds, for almost every $x \in \Omega$ and all $y, z \in [-M, M]$.

Example 2.2.10. *Each function in the form $\varphi(x, y) = a(x) + b(x)h(y)$, where $a, b \in L^{\infty}(\Omega)$ and $h \in C^1(\mathbb{R})$, satisfy the all above conditions.*

In the following, we assume that φ meet the requirements concerning definition 2.2.9. These conditions imply the following property for the Nemytskii operator $\Phi(y)$.

Theorem 2.2.11. *Φ is well-defined and continuous in $L^{\infty}(\Omega)$.*

We note this result doesn't hold for a Nemytskii operator $\Phi\colon L^p(\Omega) \to L^p(\Omega)$ for $p \in [1, \infty)$. For instance, to have the function $\Phi(y) = y^3$ in $L^p(\Omega)$, y must belong to $L^{3p}(\Omega)$, which is not necessarily correct for every $y \in L^p(\Omega)$. However, for all $p \in [1, \infty]$, it holds

$$\|\Phi(y) - \Phi(z)\|_{L^p(\Omega)} \leq L(M)\,\|y - z\|_{L^p(\Omega)} \tag{2.2.10}$$

for all $\|y\|_{L^{\infty}(\Omega)} \leq M$ and $\|z\|_{L^{\infty}(\Omega)} \leq M$, which is an evident result of the local Lipschitz continuity of $\varphi(x, y)$.

If φ is in addition globally Lipschitz continuous (not only locally Lipschitz continuous) with respect to y, that is

$$|\varphi(x,y) - \varphi(x,z)| \leq L\,|y-z| \quad \text{for all } y,z \in \mathbb{R},$$

then $\Phi\colon L^p(\Omega) \to L^p(\Omega)$ is well-defined and Φ is Lipschitz continuous in $L^p(\Omega)$, which means the inequality (2.2.10) holds, for all $y, z \in L^p(\Omega)$. For instance $\Phi(y) = \sin(y(\cdot))$ has this property.

Differentiability of the Nemytskii Operator

First, we introduce the concept of differentiability in Banach spaces:

Differentiability in Banach Spaces

In the following, U and V will denote two Banach spaces and $\mathcal{U}$ an open subset of U.

Definition 2.2.12. *We say that the mapping $F\colon \mathcal{U} \to V$ has the directional derivative $\delta F(u,h)$ at $u \in \mathcal{U}$ in the direction $h \in U$, if the limit*

$$\delta F(u,h) \coloneqq \lim_{t\to 0} \frac{F(u+th) - F(u)}{t}$$

exists in V. The mapping $h \mapsto \delta F(u,h)$ is called the first variation of F at u, if this limit exists for all $h \in U$.

We know that the first variation is not necessarily a linear mapping. Therefore, there is a stronger definition as Gâteaux differentiability:

Definition 2.2.13. *Assume that the first variation of F at u exists. F is said to be Gâteaux differentiable at u, if there exists a continuous linear mapping $A\colon U \to V$ such that*

$$\delta F(u,h) = A\,h \quad \text{for all } h \in U,$$

then A is termed the Gâteaux derivative of F at u and we denote $A = F_G(u)$

A function defined in Banach spaces can have even better differentiability property, namely Fréchet differentiability.

Definition 2.2.14. *The mapping $F\colon \mathcal{U} \to V$ is called Fréchet differentiable at a point $u \in \mathcal{U}$ if there exists an operator $A \in \mathcal{L}(U,V)$, such that, for all $h \in U$,*

$$\lim_{h\to 0} \frac{\|F(u+h) - F(u) - A\,h\|_V}{\|h\|_U} = 0 \quad \text{as } \|h\|_U \to 0.$$

In this case A is said to be the Fréchet derivative of F at u and denoted by $A = F'(u)$. If A is Fréchet differentiable at every point $u \in \mathcal{U}$, then A is called Fréchet differentiable in $\mathcal{U}$.

Theorem 2.2.15 (Chain Rule). *Let U, V, and Z be Banach space and $\mathcal{U}$ and $\mathcal{V}$ be the open subsets of U and V respectively. Suppose that $F\colon \mathcal{U} \to \mathcal{V}$ and $G\colon \mathcal{V} \to Z$ are Fréchet differentiable at $u \in \mathcal{U}$ and at $F(u) \in \mathcal{V}$ respectively. The the composition*

$$E \coloneqq G \circ F\colon \mathcal{U} \to Z$$
$$E(u) = G(F(u))$$

is Fréchet differentiable at u, and

$$E'(u) = G'(F(u))F'(u).$$

Definition 2.2.16. *Suppose that $F\colon \mathcal{U} \to V$ be a Fréchet differentiable mapping in an open neighborhood $\mathcal{U}$ of $\bar{u} \in U$. F is called continuously Fréchet differentiable at $\bar{u}$ if the mapping $u \mapsto F'(u)$ from $\mathcal{U}$ into $\mathcal{L}(U,V)$ is continuous at $\bar{u}$,that is*

$$\|u - \bar{u}\|_U \to 0 \quad \text{implies } \left\|F'(u) - F'(\bar{u})\right\|_{\mathcal{L}(U,V)} \to 0.$$

If F is continuously Fréchet differentiable at every point in $\mathcal{U}$, then it is called continuously Fréchet differentiable in $\mathcal{U}$.

For the Gâteaux differentiability of the Nemytskii operator we have to check the existence of the following limit:

$$\Phi'(y)h = \lim_{t\to 0} \frac{\Phi(y+th)-\Phi(y)}{t}.$$

Let this limit exist in $L^\infty(\Omega)$ and call it z, then

$$\frac{1}{t}\lim_{t\to 0} \|\Phi(y+th)-\Phi(y)-tz\|_{L^\infty(\Omega)} \to 0,$$

which means

$$(\Phi'(y)h)(x) = \frac{1}{t}\lim_{t\to 0} |\varphi(x,y(x)+th(x))-\varphi(x,y(x))-tz(x)| \quad \text{for a.e.} x\in\Omega.$$

Therefore, we expect at least the differentiability of $\varphi(x,y)$ with respect to y and for almost every $x\in\Omega$. If φ is differentiable with respect to y, then for every fixed x we have $z(x)=\varphi_y(x,y(x))h(x)$.

This means we found a candidate for the Fréchet derivative, namely

$$D\Phi(\bar{y})\colon L^\infty(\Omega)\to L^\infty(\Omega)$$
$$h\mapsto \varphi_y(\cdot,\bar{y}(\cdot))h(\cdot).$$

For the Fréchet differentiability we must have

$$\|\Phi(\bar{y}+h)-\Phi(\bar{y})-D\Phi(\bar{y})h\|_{L^\infty(\Omega)} = o(\|h\|_{L^\infty(\Omega)}),$$

where $D\Phi(\bar{y})=\varphi_y(\cdot,\bar{y}(\cdot))$ is itself a Nemytskii operator. In order for the function φ_y to be well-defined it must satisfy the aforementioned conditions in definition 2.2.9.

Theorem 2.2.17. *Suppose that the function φ is Carathéodory and for almost every $x\in\Omega$ is differentiable with respect to y. Furthermore, suppose that φ_y satisfy both the boundedness and local Lipschitz condition. Then the Nemytskii $\Phi\colon L^\infty(\Omega)\to L^\infty(\Omega)$ is Fréchet differentiable and we have*

$$(\Phi'(y)h)(x) = \varphi_y(x,y(x))h(x)$$

for almost every $x\in\Omega$ and all $h\in L^\infty(\Omega)$. Moreover, Φ is continuously Fréchet differentiable in $L^\infty(\Omega)$.

We note that every function $\varphi\in C^2(\mathbb{R})$, which depends only on y, satisfies the conditions in above theorem.

Remark 2.2.18. *Let Ω be a bounded and measurable set and suppose that $\varphi\colon\Omega\times\mathbb{R}\to\mathbb{R}$ is Carathéodory and satisfies the growth condition*

$$|\varphi(x,y)|\le \alpha(x)+\beta(x)\,|y|^{p/q},$$

where $\alpha,\beta\in L^\infty(\Omega)$ and $1\le q\le p<\infty$, then the Nemytskii operator

$$\Phi\colon L^p(\Omega)\to L^q(\Omega)$$
$$\Phi(y)=\varphi(\cdot,y(\cdot))$$

is well-defined.

Furthermore, if the operator Φ is well-defined it is automatically continuous.

In addition, if $\varphi_y(x,y)$ exists for almost every $x\in\Omega$ and its corresponding Nemytskii operator is a mapping from $L^p(\Omega)$ into $L^r(\Omega)$, where $1\le q<p<\infty$ and $\frac{1}{r}+\frac{1}{p}=\frac{1}{q}$, then $\Phi\colon L^p(\Omega)\to L^q(\Omega)$ is Fréchet differentiable, and we have

$$(\Phi'(y)h)(x)=\varphi_y(\cdot,y(\cdot))h.$$

Indeed,

$$\int_\Omega |\varphi_y(x,y(x))h(x)|^q \,\mathrm{d}x \le \left(\int_\Omega |\varphi_y(x,y(x))|^{q\times\frac{p}{p-q}}\right)^{1-\frac{q}{p}} \left(\int_\Omega |h(x)|^p \,\mathrm{d}x\right)^{q/p}$$
$$\le \left(\int_\Omega |\varphi_y(x,y(x))|^r \,\mathrm{d}x\right)^{1-\frac{q}{p}} \left(\int_\Omega |h(x)|^p \,\mathrm{d}x\right)^{q/p} < \infty.$$

We observe the following example from Tröltzsch, 2010, Page 205.

Example 2.2.19. *Let Ω be a bounded domain and $k \ge 1$ an integer. The Nemytskii operator generated by $\varphi(y) = y^k$ is for $k \le 5$, Fréchet differentiable form $L^6(\Omega)$ into $L^{6/5}(\Omega)$.*

2.2.4 Fréchet Differentiability of Control-to-State Map

After explaining the general framework it is time to address the second question, namely the differentiability of the control-to-state map.

We assume that $J\colon Y \times U \to \mathbb{R}$, $E\colon Y \times U \to Z$ are continuously Fréchet differentiable. If the partial derivative of E is additionally boundedly invertible with respect to y at $(\bar{y},\bar{u})$, then existence of a locally unique solution $y(u)$ to the state equation $E(y,u) = 0$ in a neighborhood of $(\bar{y},\bar{u})$ is established by the implicit function theorem. Moreover, S is Fréchet differentiable and we have

$$S'(\bar{u}) = -(E_y(\bar{y},\bar{u}))^{-1}E_u(\bar{y},\bar{u}). \tag{2.2.11}$$

Theorem 2.2.20 (Implicit Function Theorem). *Suppose that X, Y and Z are Banach spaces and $U \subset X \times Y$ is an open set and*

$$G\colon U \to Z$$

is a $C^k(U,Z)$ mapping. In addition, we assume for a suitable point $(x_0,y_0) \in U$, $G(x_0,y_0) = 0$ and

$$G_y(x_0,y_0) \in \mathcal{L}(Y,Z),$$

is a bijection. Then there exists a unique function, defined on a neighborhood $B(x_0)$ of y_0

$$g\colon B(x_0) \to Y$$

such that $g \in C^k(B(x_0),Y)$ and

$$\begin{aligned} &g(x_0) = y_0,\\ &G(x,g(x)) = 0 \quad \text{in } B(x_0),\\ &g'(x) = -(G_y(x,g(x)))^{-1}(G_x(x,g(x)). \end{aligned}$$

Optimal Control of KE: We observe that in optimal control problem (2.2.6) the control-to-state map S is well-defined and continuous if the controls belong to $H^1(\Omega)$ with $u_a \le u$. That is an obstacle problem, and we are not able to achieve Fréchet differentiability of S with respect to topology of $H^1(\Omega)$. That is the reason why we impose an upper bound u_b to controls. Indeed, we can utilize the upper bound in admissible set and work with topology $H^1(\Omega) \cap L^\infty(\Omega)$.

We define the corresponding operator to KE

$$E\colon (H_0^1(\Omega) \cap W^{2,q}(\Omega)) \times L^\infty(\Omega) \to L^q(\Omega)$$
$$E(y,u) = -\Delta y - \frac{f}{u + b\,\|\nabla y\|^2},$$

This operator is continuously Fréchet differentiable. To apply the implicit function theorem we need to show invertibility (bijectivity) of E_y at an arbitrary point $(\hat{y}, \hat{u})$

$$E_y(\hat{y}, \hat{u})\, y = -\Delta y + \frac{2\, b\, f\, (\nabla \hat{y}, \nabla y)}{(\hat{u} + b\, \|\nabla \hat{y}\|^2)^2},$$

where $\hat{u} \in \mathcal{U}_{\text{ad}}$ and $\hat{y}$ is the solution of Kirchhoff equation.

This is equivalent to that the following PDE has a unique solution:

$$\begin{cases} -\Delta y - \dfrac{2\, b\, (\nabla \hat{y}, \nabla y) \Delta \hat{y}}{(\hat{u} + b\, \|\nabla \hat{y}\|^2)} = h & \text{in } \Omega, \\ \qquad\qquad\qquad\qquad\qquad y = 0 & \text{on } \partial\Omega. \end{cases}$$

This will be shown in proposition 3.2.1.

Optimal Control of CS: The corresponding operator to the coupled chemotaxis system reads as follows:

$$\begin{aligned} &E \colon Y \times L^\infty(\partial\Omega_T) \to Z \\ &(y, w, u) \mapsto (E_1(y, w, u), E_2(y, w, u), E_3(y, w, u))\,. \end{aligned}$$

where, $Y \coloneqq W(0,T) \times L^\infty(0,T; W^{1,6}(\Omega))$ and $Z \coloneqq L^2(0,T; H^1(\Omega)^*) \times L^\infty(0,T; W^{1,6/5(\Omega)^*}) \times L^2(\Omega))$. Furthermore E_1, E_2 and E_3 are defined as follows:

$$\begin{aligned} \langle E_1(y, w, u)\,, \varphi \rangle \coloneqq &\int_0^T \int_\Omega \partial_t y\, \varphi \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \nabla y \cdot \nabla \varphi \,\mathrm{d}x\,\mathrm{d}t \\ &- \chi \int_0^T \int_\Omega y\, \nabla w \cdot \nabla \varphi \,\mathrm{d}x\,\mathrm{d}t + \chi \int_0^T \int_{\partial\Omega} u\, y\, \varphi \,\mathrm{d}s\,\mathrm{d}t \\ &- \int_0^T \int_\Omega \left(a_0 - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x \right) y\, \varphi \,\mathrm{d}x\,\mathrm{d}t, \quad \varphi \in L^2(0,T; H^1(\Omega)), \end{aligned}$$

$$\begin{aligned} \langle E_2(y, w, u)\,, \psi \rangle \coloneqq &\int_0^T \int_\Omega \nabla w \cdot \nabla \psi \,\mathrm{d}x\,\mathrm{d}t - \int_0^T \int_{\partial\Omega} u\, \psi \,\mathrm{d}s\,\mathrm{d}t + \lambda \int_0^T \int_\Omega w\, \psi \,\mathrm{d}x\,\mathrm{d}t \\ &- \int_0^T \int_\Omega y\, \psi \,\mathrm{d}x\,\mathrm{d}t, \quad \psi \in L^2(0,T; H^1(\Omega)), \end{aligned}$$

and

$$E_3(y, w, u) = y(x, 0) - y_0(x) \quad \text{in } \Omega.$$

These operators are well-defined: $y \in H^1(\Omega) \hookrightarrow L^6(\Omega)$, $\nabla w \in L^3(\Omega)$ and $\nabla \varphi \in L^2(\Omega)$. Hence, $\int_0^T \int_\Omega y\, \nabla w \cdot \nabla \varphi \,\mathrm{d}x\,\mathrm{d}t$ is well-defined. We also need to check the nonlinearity term, namely $\varphi \mapsto \int_0^T \int_\Omega y^2\, \varphi \,\mathrm{d}x\,\mathrm{d}t$. $y \in W(0,T)$ results in $y \in L^4(0,T; L^3(\Omega))$ for $N \leq 3$, by virtue of DiBenedetto, 1993, Proposition 3.4.

Theorem 2.2.21. *For every function $y \in W(0,T)$ we have*

$$\|y\|_{L^p(0,T;L^q(\Omega))} \leq c\, \|y\|_{W(0,T)}$$

for some positive constant c depending on Ω, where the numbers $p, q \geq 1$ are linked by

$$\frac{1}{p} + \frac{N}{2q} = \frac{N}{4}$$

and their admissible range is

$$\begin{cases} q \in (2,\infty],\ p \in [4,\infty); & \text{if } N=1,\\ q \in [2,\infty),\ p \in (\frac{4}{N},\infty]; & \text{if } N=2,\\ q \in [2,\frac{2N}{N-2}],\ p \in [2,\infty]; & \text{if } N \geq 3. \end{cases}$$

That means $y^2 \in L^2(0,T;L^{3/2}(\Omega)) \equiv L^2(0,T;L^3(\Omega)^*)$. Since $\varphi \in L^2(0,T;H^1(\Omega)) \hookrightarrow L^2(0,T;L^6(\Omega))$ for $N \leq 3$ and $L^2(0,T;H^1(\Omega))$ is dense in $L^2(0,T;L^3(\Omega))$ we may extend $\varphi \mapsto \int_0^T \int_\Omega y^2\,\varphi \,\mathrm{d}x\,\mathrm{d}t$ to a continuous functional on $L^2(0,T;H^1(\Omega))$, which concludes the well-definedness of E_1.

The well-definedness of E_2 is obvious. E_3 benefits $W(0,T) \hookrightarrow C([0,T];L^2(\Omega))$ to be well-defined.

Fréchet differentiability of these operators can be confirmed simply. We note that the Nemytskii operator y^2 is Fréchet differentiable from $L^6(\Omega)$ into $L^{6/5}(\Omega)$, by virtue of (2.2.19). Since $H^1(\Omega) \hookrightarrow L^6(\Omega)$ is dense, we have $L^6(\Omega)^* \equiv L^{6/5}(\Omega) \hookrightarrow H^1(\Omega)^*$. Therefor, y^2 is Fréchet differentiability from $H^1(\Omega)$ into $H^1(\Omega)^*$.

2.2.5 Optimality Condition

If $\bar{u} \in \mathcal{U}_{\mathrm{ad}}$ is a local optimal solution of the optimization problem in the reduced form

$$\min_{u \in \mathcal{U}_{\mathrm{ad}}} j(u) = J(y(u),u),$$

then it satisfies the variational inequality

$$\langle j'(\bar{u})\,, u-\bar{u}\rangle_{U^*,U} \geq 0 \quad \text{for all } u \in \mathcal{U}_{\mathrm{ad}},$$

by virtue of theorem 2.2.2. $j'(\bar{u})$ we can be computed by applying the chain rule

$$j'(\bar{u}) = J_y(\bar{y},\bar{u}) \circ S'(u) + J_u(\bar{y},\bar{u}).$$

This means for a direction $h \in U$

$$\langle j'(\bar{u})\,, h\rangle_{U^*,U} = \langle J_y(\bar{y},\bar{u})\,, S'(\bar{u})h\rangle_{Y^*,Y} + \langle J_u(\bar{y},\bar{u})\,, h\rangle_{U^*,U}.$$

By means of the definition of adjoint operator we compute

$$\langle J_y(\bar{y},\bar{u})\,, S'(\bar{u})h\rangle_{Y^*,Y} + \langle J_u(\bar{y},\bar{u})\,, h\rangle_{U^*,U} = \langle (S'(\bar{u}))^* J_y(\bar{y},\bar{u})\,, h\rangle_{U^*,U} + \langle J_u(\bar{y},\bar{u})\,, h\rangle_{U^*,U}$$

and consequently

$$j'(\bar{u}) = (S'(\bar{u}))^* J_y(\bar{y},\bar{u}) + J_u(\bar{y},\bar{u}) \in U^*.$$

As in linear case we need to make the optimality condition more effective, to exploit it numerically. Furthermore, we introduce an appropriate adjoint problem to reexpress the derivative $j'(\bar{u})$ in a more convenient form.

Exploiting (2.2.11) we obtain

$$\begin{aligned} j'(\bar{u})h &= \langle (S'(\bar{u}))^* J_y(\bar{y},\bar{u})\,, h\rangle_{U^*,U} + \langle J_u(\bar{y},\bar{u})\,, h\rangle_{U^*,U} \\ &= \langle E_u(\bar{y},\bar{u})^* (E_y(\bar{y},\bar{u})^{-1})^* J_y(\bar{y},\bar{u})\,, h\rangle + \langle J_u(\bar{y},\bar{u})\,, h\rangle. \end{aligned}$$

Defining a Multiplier, or adjoint state, $p := (E_y(\bar{y},\bar{u})^{-1})^* J_y(\bar{y},\bar{u}) \in Z^*$ we obtain

$$j'(\bar{u})h = \langle E_u(\bar{y},\bar{u})^* p\,, h\rangle + \langle J_u(\bar{y},\bar{u})\,, h\rangle.$$

Consequently, the optimality condition reads

$$j'(\bar{u})(u-\bar{u}) = \langle E_u(\bar{y},\bar{u})^* p + J_u(\bar{y},\bar{u})\,, (u-\bar{u})\rangle \geq 0$$

Definition 2.2.22. *An element $p \in Z^*$ is called the adjoint state associated to $\bar{u}$ if it fulfills the following adjoint equation*

$$E_y(\bar{y}, \bar{u})^* p = -J_y(\bar{y}, \bar{u}), \tag{2.2.12}$$

where $E_y(\bar{y}, \bar{u})^*$ denotes the adjoint operator of $E_y(\bar{y}, \bar{u})$.

Then the optimality system reads as follows

$$\begin{cases} E(\bar{y}, \bar{u}) = 0 & \text{state equation,} \\ E_y(\bar{y}, \bar{u})^* p = -J_y(\bar{y}, \bar{u}) & \text{adjoint equation,} \\ \langle E_u(\bar{y}, \bar{u})^* p + J_u(\bar{y}, \bar{u}) \,, (u - \bar{u}) \rangle_{U^*, U} \geq 0 \quad \text{for all } u \in \mathcal{U}_{\text{ad}} & \text{variational inequality.} \end{cases} \tag{2.2.13}$$

Optimal Control of CS: We have

$$S \colon L^\infty(\partial\Omega_T) \to W(0,T) \times L^\infty(0,T; W^{1,6}(\Omega))$$
$$S(u) = (S_1(u), S_2(u)).$$

Substituting this into J, we obtain the reduced cost functional j,

$$j(u) \coloneqq J(S(u), u).$$

J is Fréchet differentiable and S is also Fréchet differentiable, by virtue of theorem 4.2.3. Therefore, j is Fréchet differentiable in $L^\infty(\partial\Omega)$.

Since $\mathcal{U}_{\text{ad}}$ is convex, every locally optimal control $\bar{u}$ satisfies the variational inequality

$$j'(\bar{u})(u - \bar{u}) \geq 0 \quad \text{for all } u \in \mathcal{U}_{\text{ad}},$$

see Tröltzsch, 2010, lemma 5.10. Using the chain rule

$$j(\bar{u})(u - \bar{u}) = J_y(\bar{y}, \bar{w}, \bar{u}) S_1'(\bar{u})(u - \bar{u}) + J_w(\bar{y}, \bar{w}, \bar{u}) S_2'(\bar{u})(u - \bar{u}) + J_u(\bar{y}, \bar{w}, \bar{u})(u - \bar{u}),$$

we can calculate j', where $(y, w) = (S_1'(\bar{u})(u - \bar{u}), S_2'(\bar{u})(u - \bar{u}))$ is the solution of the linearized problem (4.2.3).

The state variable (y, w) can be eliminated by means of an adjoint variable (p, q), which is the solution to the adjoint problem defined by

$$E_{(y,w)}(\bar{y}, \bar{w}, \bar{u})^*(p, q) = -J_{(y,w)}(\bar{y}, \bar{w}, \bar{u}), \tag{2.2.14}$$

which in our case amounts to (4.2.4).

Lagrangian Method

There is also a formal way to derive the optimality system called Lagrangian method.

We define the Lagrangian function

$$\mathcal{L} \colon Y \times U \times Z^* \to \mathbb{R}$$
$$\mathcal{L}(y, u, p) = J(y, u) + \langle p \,, E(y, u) \rangle_{Z^*, Z}.$$

The variable p is referred to as Lagrange multiplier.

We expect the optimal solution $(\bar{y}, \bar{u})$, together with the Lagrange multiplier p, to satisfy the optimality conditions associated with the problem

$$\min_{u \in \mathcal{U}_{\text{ad}}} \mathcal{L}(y, u, p), \quad y \text{ unconstrained.}$$

Taking derivative of $\mathcal{L}(y, u, p)$ with respect to y in direction h, we obtain

$$\begin{aligned} \mathcal{L}_y(y, u, p)h &= \langle J_y(y, u) \,, h \rangle_{Y^*, Y} + \langle p \,, E_y(y, u)h \rangle_{Z^*, Z} \\ &= \langle J_y(y, u) \,, h \rangle_{Y^*, Y} + \langle E_y(y, u)^* p \,, h \rangle_{Y^*, Y} \end{aligned}$$

and therefore, the adjoint equation in (2.2.13) can be written as

$$\mathcal{L}_y(\bar{y}, \bar{u}, p) = 0.$$

In a similar manner, taking derivative of $\mathcal{L}(y,u,p)$ with respect to u in direction v, we obtain

$$\begin{aligned}\mathcal{L}_u(y,u,p)v &= \langle J_u(y,u)\,,v\rangle_{U^*,U} + \langle p\,, E_u(y,u)h\rangle_{Z^*,Z}\\ &= \langle J_u(y,u)\,,v\rangle_{U^*,U} + \langle E_u(y,u)^*p\,,v\rangle_{U^*,U}.\end{aligned}$$

Applying theorem 2.2.2 the variational inequality in (2.2.13) can be expressed as

$$\mathcal{L}_u(\bar{y},\bar{u},p)(u-\bar{u}) \geq 0 \quad \text{for all } u \in \mathcal{U}_{\text{ad}}.$$

In summary, the first-order optimality system (2.2.13) can be expressed in the following form

$$\begin{cases}\mathcal{L}_y(\bar{y},\bar{u},p) = 0 & \text{(adjoint equation)},\\ \mathcal{L}_p(\bar{y},\bar{u},p) = 0 & \text{(state equation)},\\ \mathcal{L}_u(\bar{y},\bar{u},p)(u-\bar{u}) \geq 0 \quad \text{for all } u \in \mathcal{U}_{\text{ad}} & \text{(variational inequality)}.\end{cases}$$

Optimal Control of KE: We define the Lagrangian function as follows:

$$\begin{aligned}&\mathcal{L}\colon H_0^1(\Omega)\times\mathcal{U}_{\text{ad}}\times H_0^1(\Omega)\to\mathbb{R}\\ &\mathcal{L}(y,u,p) \coloneqq \int_\Omega \varphi(x,y)\,\mathrm{d}x + \frac{\lambda}{2}\|u\|^2_{H^1(\Omega)} + \int_\Omega \nabla y\cdot\nabla p\,\mathrm{d}x - \int_\Omega \frac{f}{u+b\|\nabla y\|^2}\,p\,\mathrm{d}x\end{aligned} \tag{2.2.15}$$

and the optimality system will be derived in subsection 3.2.2. Since we have a box constraint for controls in Kirchhoff problem we obtain a variational inequality in optimality system. We intend to have a system of operator equations, which is Newton differentiable. Therefore, we relax the lower and upper bounds in admissible set and apply the Moreau-Yosida approximation. This will be explained later in subsection 3.2.4.

Finite Element Discretization

After deriving the optimality system, the next step involves discretizing the partial differential equations. This approach, known as optimize-then-discretize, contrasts with discretize-then-optimize, where the equations and the cost functional are initially discretized before being solved using large-scale optimization tools.

One of the most popular method for discretization of partial differential equation is finite element method. Initially, we need a weak formulation of the boundary value problem. The domain is then subdivided by a regular triangulation into finitely many triangles ($N = 2$) or tetrahedra ($N = 3$), each with disjoint interiors, referred to as elements. Additionally, we define a finite element space associated with the triangulation and select basis functions for this finite element space. These basis functions are designed to have small supports, ensuring that the resulting stiffness matrix is sparse.

In unconstrained control case the optimality system (2.2.13) reduces to the following form

$$\begin{cases}E(\bar{y},\bar{u}) = 0,\\ E_y(\bar{y},\bar{u})^*p + J_y(\bar{y},\bar{u}) = 0,\\ E_u(\bar{y},\bar{u})^*p + J_u(\bar{y},\bar{u}) = 0.\end{cases} \tag{2.2.16}$$

Below we detail the construction of a discretized system of (2.2.16). We start with the discretization of the nonlinear adjoint equation and assume $Z = V^*$

$$E(y,u) = 0 \quad \text{in } V^*.$$

We first recall its weak formulation: given $u \in \mathcal{U}$, find $y = y(u) \in V$ such that

$$\langle E(y,u)\,,\varphi\rangle_{V^*,V} = 0 \quad \text{for all } \varphi \in V.$$

The variational form $e(\cdot,u;\cdot)\colon V\times U\times V\to\mathbb{R}$ can be defined as

$$e(y,u;\varphi) = \langle E(y,u)\,,\varphi\rangle_{V^*,V} \quad \text{for all } y,\varphi\in V,$$

and therefore

$$e(y, u; \varphi) = 0 \quad \text{for all } \varphi \in V. \tag{2.2.17}$$

For the numerical approximation of (2.2.17) we introduce two suitable finite-dimensional subspaces $V_h \subset V$ and $U_h \subset U$ of dimension N_V and N_U, respectively. Here the same space V_h is used to approximate both state and adjoint variables. Moreover, V_h is used to approximate the trial space (where we seek the solution) as well as the test space, thus yielding a Galerkin problem for both the state and the adjoint system.

The Galerkin formulation reads as follows: given $u_h \in U_h$, find $y_h = y_h(u_h) \in V_h$ such that

$$e(y_h, u_h; \varphi) = 0 \quad \text{for all } \varphi \in V_h \tag{2.2.18}$$

solving a nonlinear system of N_V equations.

Indeed, we set

$$y_h = \sum_{j=1}^{N_V} y_{h,j}\varphi_j, \quad u_h = \sum_{j=1}^{N_U} u_{h,j}\psi_j,$$

where $\{\varphi_j\}_{j=1}^{N_V}$, $\{\psi_j\}_{j=1}^{N_U}$ are a basis for V_h, U_h, respectively.

Denoting the vectors having as component the unknown coefficients $y_{h,j}$ and $u_{h,j}$ by $\mathbf{y}$ and $\mathbf{u}$, respectively, (2.2.18) is equivalent to: $\mathbf{u} \in \mathbb{R}^{N_U}$, find $\mathbf{y} = \mathbf{y}(\mathbf{u}) \in \mathbb{R}^{N_V}$ such that

$$e(\mathbf{y}, \mathbf{u}) = 0,$$

where the *residual vector* $e(\cdot, \mathbf{u}) \in \mathbb{R}^{N_V}$ is given by

$$(e(\mathbf{y}, \mathbf{u}))_i = e(y_h, u_h; \varphi_i) \quad i = 1, \dots, N_V.$$

To discretize the adjoint equation, let us assume that the same space is used to discretize both state and adjoint variable. We define the *Jacobian matrix* $e_y(\bar{\mathbf{y}}, \mathbf{u}) \in \mathbb{R}^{N_V \times N_V}$ as

$$(e_y(\bar{\mathbf{y}}, \mathbf{u}))_{ij} = e_y(\bar{y}_h)(\varphi_j, \varphi_i) \quad i, j = 1, \dots, N_V,$$

where the partial Fréchet derivative of g with respect to y at $\bar{y} \in V$ and in the direction z is denoted by

$$e_y(\bar{y})(z, u; \varphi) = \langle E_y(\bar{y}, u)z \, , \varphi \rangle_{V^*,V} \quad \text{for all } z \in V, \varphi \in V,$$

with $E_y \colon V \to \mathcal{L}(V, V^*)$. The vector $(J_y(\mathbf{y}, \mathbf{u})) \in \mathbb{R}^{N_V}$ as,

$$(J_y(\mathbf{y}, \mathbf{u}))_i = J_y(y_h, u_h)\varphi_i \quad i = 1, \dots, N_V.$$

Finally, if we don't have any control constraint the optimality condition turns to the following discrete equation

$$J_u(\mathbf{y}, \mathbf{u}) + e_u(\mathbf{y}, \mathbf{u})^T \mathbf{p}.$$

The matrix $e_u(\mathbf{y}, \mathbf{u}) \in \mathbb{R}^{N_V \times N_U}$ is given by

$$(e_u(\mathbf{y}, \mathbf{u}))_{ij} = e_u(\bar{u}_h)(y, \psi_j; \varphi_i) \quad i = 1, \dots, N_V, \; j = 1, \dots, N_U$$

and the partial Fréchet derivative of g, with respect to u at $\bar{u} \in U$ in the direction w, by

$$e_u(\bar{u})(y; w, \varphi) = \langle E_u(y, \bar{u})w \, , \varphi \rangle_{V^*,V} \quad \text{for all } w \in U, \varphi \in V,$$

where $E_u \colon U \to \mathcal{L}(U, V^*)$.

Finally, $J_u(\mathbf{y}, \mathbf{u}) \in \mathbb{R}^{N_U}$ is a vector, whose components are defined as

$$(J_u(\mathbf{y}, \mathbf{u}))_i = J_u(y_h, u_h)\psi_i \quad i = 1, \dots, N_U.$$

In summary, the optimal solution $(\bar{\mathbf{y}}, \bar{\mathbf{u}}, \mathbf{p})$ satisfies the following system of optimality conditions:

$$\begin{cases} e(\mathbf{y}, \mathbf{u}) = 0 & \text{state equation,} \\ (e_y(\mathbf{y}, \mathbf{u}))^T \mathbf{p} = -J_y(\mathbf{y}, \mathbf{u}) & \text{adjoint equation,} \\ J_u(\mathbf{y}, \mathbf{u}) + (e_u(\mathbf{y}, \mathbf{u}))^T \mathbf{p} = 0 & \text{variational equality.} \end{cases} \tag{2.2.19}$$

A system akin to (2.2.19) would also be obtained by employing the discretize-then-optimize approach.

To introduce some discretization matrices we consider the optimal control of the stationary heat source and carry the discretize-then-optimize approach out.

$$\text{Minimize } J(y, u) = \frac{1}{2} \int_\Omega (y - y_d)^2 \, \mathrm{d}x + \frac{\lambda}{2} \int_\Omega u^2 \, \mathrm{d}x \tag{2.2.20a}$$

$$\text{subject to } \begin{cases} -\Delta y = u & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.2.20b}$$

We have already observed the optimize-then-discretize approach in abstract form, recovering first a system of optimality condition at continuous level, in the abstract from. Alternatively to the previous strategy we may first discretize problem (2.2.20) by substituting all functional spaces by finite dimensional ones. That is

$$\text{Minimize } J(y_h, u_h) = \frac{1}{2} \int_\Omega (y_h - y_d)^2 \, \mathrm{d}x + \frac{\lambda}{2} \int_\Omega u_h^2 \, \mathrm{d}x \tag{2.2.21a}$$

$$\text{subject to } (\nabla y_h, \nabla \varphi_h)_{L^2(\Omega)} = (u_h, \varphi_h)_{L^2(\Omega)} \quad \varphi_h \in V_h. \tag{2.2.21b}$$

We introduce *stiffness matrix* $\mathbf{K} \in \mathbb{R}^{N_V \times N_V}$ and the *control matrix* $\mathbf{B} \in \mathbb{R}^{N_V \times N_U}$ as follows:

$$\mathbf{K}_{ij} = (\nabla \varphi_j, \nabla \varphi_i)_{L^2(\Omega)} \quad \text{and} \quad \mathbf{B}_{ij} = (\psi_j, \varphi_i)_{L^2(\Omega)}.$$

This results in the discrete state problem

$$\mathbf{K}\mathbf{y} = \mathbf{B}\mathbf{u}.$$

For the cost functional we have

$$\begin{aligned} J(\mathbf{y}, \mathbf{u}) &= \frac{1}{2} \int_\Omega (\mathbf{y} - \mathbf{y}_d)^2 \, \mathrm{d}x + \frac{\lambda}{2} \int_\Omega \mathbf{u}^2 \, \mathrm{d}x \\ &= \frac{1}{2} (\mathbf{y} - \mathbf{y}_d)^T \mathbf{M} (\mathbf{y} - \mathbf{y}_d) + \frac{\lambda}{2} \mathbf{u}^T \mathbf{D} \mathbf{u}, \end{aligned}$$

where $\mathbf{y}_d \in \mathbb{R}^{N_V}$ with

$$(\mathbf{y}_d)_i = (y_d, \varphi_i)_{L^2(\Omega)}$$

and the mass matrices $\mathbf{M} \in \mathbb{R}^{N_V \times N_V}$ and $\mathbf{D} \in \mathbb{R}^{N_U \times N_U}$ for state and control respectively, are given by

$$\mathbf{M}_{ij} = (\varphi_j, \varphi_i)_{L^2(\Omega)}, \quad \text{and} \quad \mathbf{D}_{ij} = (\psi_j, \psi_i)_{L^2(\Omega)}.$$

In summary, we obtain the following discretized optimal control problem

$$\begin{aligned} &\text{Minimize } \frac{1}{2} (\mathbf{y} - \mathbf{y}_d)^T \mathbf{M} (\mathbf{y} - \mathbf{y}_d) + \frac{\lambda}{2} \mathbf{u}^T \mathbf{D} \mathbf{u} \\ &\text{subject to } \mathbf{K}\mathbf{y} = \mathbf{B}\mathbf{u}. \end{aligned}$$

The Lagrangian of our discretized problem becomes

$$\mathcal{L}(\mathbf{y}, \mathbf{u}, \mathbf{p}) = \frac{1}{2} (\mathbf{y} - \mathbf{y}_d)^T \mathbf{M} (\mathbf{y} - \mathbf{y}_d) + \frac{\lambda}{2} \mathbf{u}^T \mathbf{D} \mathbf{u} + \mathbf{p}^T (\mathbf{K}\mathbf{y} - \mathbf{B}\mathbf{u}).$$

Therefore, the first-order derivatives of $\mathcal{L}$ with respect to y and u are given by

$$\mathcal{L}_y(\mathbf{y}, \mathbf{u}, \mathbf{p}) = \mathbf{M}(\mathbf{y} - \mathbf{y}_d) + \mathbf{K}^T p, \quad \text{and} \quad \mathcal{L}_u(\mathbf{y}, \mathbf{u}, \mathbf{p}) = \lambda \mathbf{D}\mathbf{u} - \mathbf{B}^T \mathbf{p}.$$

Therefore, the associated discretized optimality system reads

$$\begin{cases} \mathbf{K}\mathbf{y} = \mathbf{B}\mathbf{u}, \\ \mathbf{K}^T\mathbf{p} = -\mathbf{M}(\mathbf{y} - \mathbf{y}_d), \\ \lambda\mathbf{D}\mathbf{u} = \mathbf{B}^T\mathbf{p}. \end{cases}$$

Numerical Method

One of the most popular strategies to address nonlinear optimal control problems are the sequential quadratic programming (SQP) method. In unconstrained case $\mathcal{U}_{\text{ad}} \equiv U$, the SQP method can be obtained by applying the Newton method to solve the nonlinear system of optimality conditions (2.2.5); for this reason, the SQP method is also referred to as Lagrange-Newton method.

In this case, the variational inequality in (2.2.5) reduces to an equality and the optimality system can be written as

$$\begin{cases} E(\bar{y}, \bar{u}) = 0, \\ \mathcal{L}_y(\bar{y}, \bar{u}, p) = 0, \\ \mathcal{L}_u(\bar{y}, \bar{u}, p) = 0. \end{cases} \tag{2.2.22}$$

By applying a Newton method for solving this system of equations, we obtain the following linearized system:

$$\begin{bmatrix} \mathcal{L}_{yy}(y_k, u_k, p_k) & \mathcal{L}_{yu}(y_k, u_k, p_k) & E_y(y_k, u_k)^* \\ \mathcal{L}_{uy}(y_k, u_k, p_k) & \mathcal{L}_{uu}(y_k, u_k, p_k) & E_u(y_k, u_k)^* \\ E_y(y_k, u_k, p_k) & E_u(y_k, u_k, p_k) & 0 \end{bmatrix} \begin{bmatrix} \delta y \\ \delta u \\ \delta p \end{bmatrix} = - \begin{bmatrix} E_y(y_k, u_k)^* p_k + J_y(y_k, u_k) \\ E_u(y_k, u_k)^* p_k + J_u(y_k, u_k) \\ E(y_k, u_k) \end{bmatrix} \tag{2.2.23}$$

$$y_{k+1} = y_k + \delta y, \quad u_{k+1} = u_k + \delta u, \quad p_{k+1} = p_k + \delta p.$$

Let us now highlight the structure of the linear-quadratic problem that is indeed generated, and solved, at each step. Introducing the abridged notation

$$\mathcal{L}'' = \begin{bmatrix} \mathcal{L}_{yy} & \mathcal{L}_{yu} \\ \mathcal{L}_{uy} & \mathcal{L}_{uu} \end{bmatrix}, \quad \mathcal{L}' = \begin{bmatrix} \mathcal{L}_y \\ \mathcal{L}_u \end{bmatrix}$$

system (2.2.23) corresponds to the optimality conditions of the following linear-quadratic problem:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\mathcal{L}''(y_k, u_k, p_k)(\delta y, \delta u)^2 + \mathcal{L}'(y_k, u_k)(\delta y, \delta u) \\ \text{subject to} \quad & E_y(y_k, u_k)\delta y + E_u(y_k, u_k)\delta u + E(y_k, u_k) = 0. \end{aligned} \tag{2.2.24}$$

Solving a sequence of quadratic programs under the form (2.2.24) leads to the Sequential Quadratic Programming (SQP) method, which is equivalent to applying a Newton method on the optimality system. Consequently, SQP is also referred to as the Lagrange-Newton method. This method offers a locally quadratic convergent approach for identifying stationary points in constrained optimization problems. This technique will be employed for the solution of Optimal Control Problems involving nonlinear Partial Differential Equations.

Note that, unlike the Newton method, the iterates (y_k, u_k) produced by the SQP method are infeasible for the nonlinear state equation. In other words, the SQP method generates control/state pairs that satisfy the state equation only in the limit.

We also note that system (2.2.23) is well-posed provided the second order sufficient condition (2.2.25) is verified. Indeed, under this assumption, the Newton method generates a unique sequence of iterates converging quadratically to $(\bar{y}, \bar{u}, p)$.

Theorem 2.2.23. *Let $J\colon Y \times U \to \mathbb{R}$ and $E\colon Y \times U \to Z$ be twice continuously Fréchet differentiable and let $(\bar{y}, \bar{u}, p)$ be a solution to the optimality system (2.2.13). If there exists some constant $\delta > 0$ such that*

$$\langle \mathcal{L}''(\bar{y}, \bar{u}, p)(y, u)\,, (y, u)\rangle_{Y^* \times U^*, Y \times U} \geq \delta \|u\|_U^2 \tag{2.2.25}$$

for all $y = y(u) \in Y$ and $u \in U$ that satisfy the linearized equation

$$E_y(\bar{y}, \bar{u})y + E_u(\bar{y}, \bar{u})u = 0 \quad \text{in } Z,$$

then there exists two constants $\varepsilon, \sigma > 0$ such that

$$J(y, u) \geq J(\bar{y}, \bar{u}) + \|u - \bar{u}\|_U^2$$

for all $u \in U$ with $\|u - \bar{u}\|_U \leq \varepsilon$.

In particular, $(\bar{y}, \bar{u})$ is a local minimizer of J.

After we obtained a system of operator equations we need to check a kind of differentiability of these operators.

Semismooth Newton Method

The optimality system (2.2.16) can be considered as an operator equation

$$\begin{aligned} &F\colon X \to Y \\ &F(x) = 0, \end{aligned} \tag{2.2.26}$$

where X and Y are Banach spaces. Assuming F is Fréchet differentiable, we can use a classical Newton method for solving (2.2.26).

In the optimality system of the relaxed Kirchhoff problem, namely (3.2.13), a max function will appear. However, the max function is not Fréchet differentiable, and a standard Newton scheme cannot be applied. This prompts the question: Is it possible to define a weaker differentiability notion for such a function that allows the formulation of a Newton-type iterative scheme.

First, we introduce the definition of a Newton differentiable mapping, see Hintermüller, Ito, Kunisch, 2002, Definition 1, Ito, Kunisch, 2008, Definition 8.10.

Definition 2.2.24. *Let X and Y be two Banach spaces and D be an open subset of X. The mapping $F\colon D \subset X \to Y$ is called Newton differentiable on the open subset $V \subset D$ if there exists a map $G\colon V \to \mathcal{L}(X, Y)$ such that, for every $x \in V$,*

$$\lim_{h \to 0} \frac{1}{\|h\|_X} \|F(x + h) - F(x) - G(x + h)h\|_Y = 0.$$

In this case G is said to be a Newton derivative of F on V.

Theorem 2.2.25. *Let $\bar{x}$ be a solution to (2.2.26). Suppose that F is Newton differentiable in an open neighborhood V containing $\bar{x}$. If*

$$\left\|G(x)^{-1}\right\|_{\mathcal{L}(Y,X)} \leq C$$

for some constant $C > 0$ and all $x \in V$, then the semismooth Newton iteration

$$x_{k+1} = x_k - G(x_k)^{-1} F(x_k)$$

converges superlinearly to $\bar{x}$, provided that, for the initial value x_0, $\|x_0 - \bar{x}\|_X$ is sufficiently small.

Second Order Derivatives

Here we recall the concept of second derivative.

Suppose that $F\colon \mathcal{U} \subset U \to V$ is Fréchet differentiable on the open set $\mathcal{U}$. Then

$$F'\colon \mathcal{U} \to \mathcal{L}(U,V)$$

which means $F'(u) \in \mathcal{L}(U,V)$, for $u \in \mathcal{U}$. If the mapping $u \mapsto F'(u)$ is again Fréchet differentiable at $u \in \mathcal{U}$, then F is twice Fréchet differentiable at u and

$$F''(u) \coloneqq (F')'(u) \in \mathcal{L}(U, \mathcal{L}(U,V)).$$

For given $\bar{u} \in \mathcal{U}$ and $u_1, u_2 \in U$ this means

$$\big\|F'(\bar{u}+u_1) - F'(\bar{u}) - F''(\bar{u})u_1\big\|_{\mathcal{L}(U,V)} = o(\|u_1\|_U)$$

or equivalently

$$\lim_{u_1 \to 0} \sup_{u_2 \in U} \frac{\|F'(\bar{u}+u_1)(u_2) - F'(\bar{u})(u_2) - F''(\bar{u})(u_1)(u_2)\|_V}{\|u_1\|_U \, \|u_2\|_U} = 0,$$

and we have

$$F''(\bar{u})(u_1) \in \mathcal{L}(U,V) \quad \text{and } F''(\bar{u})(u_1)(u_2) \in V.$$

That is the reason why we observe $F''(\bar{u})$ as bilinear form on U and use the notation

$$F''(\bar{u})[u_1,u_2] \coloneqq F''(\bar{u})(u_1)(u_2) \quad \text{and } F''(\bar{u})u^2 \coloneqq F''(\bar{u})[u,u].$$

We can easily calculate

$$F''(\bar{u})[u_1,u_2] = \frac{\mathrm{d}^2}{\mathrm{d}t\,\mathrm{d}s} F(\bar{u} + tu_1 + su_2)|_{(t,s)=0}$$

and

$$F''(\bar{u})[u,u] = \frac{\mathrm{d}^2}{\mathrm{d}t^2} F(\bar{u} + tu_1)|_{t=0}.$$

Second-Order Derivatives of Nemytskii Operators

Suppose that the function $\varphi = \varphi(x,y)\colon \Omega \times \mathbb{R} \to \mathbb{R}$ is Carathéodory. Moreover, assume that φ_{yy} exist and satisfy the boundedness and local Lipschitz condition. Then the associated Nemytskii operator Φ is twice continuously differentiable in $L^\infty(\Omega)$, and we have

$$(\Phi''(y)[h_1,h_2])(x) = \varphi_{yy}(x,y(x))h_1(x)h_2(x).$$

Let $1 \le 2q < p < \infty$. The Nemytskii operator $\Phi\colon L^p(\Omega) \to L^q(\Omega)$ has the second derivative, provided that $y(\cdot) \mapsto \varphi_{yy}(\cdot, y(\cdot))$ maps $L^p(\Omega)$ into $L^r(\Omega)$ with

$$r = \frac{pq}{p-2q}.$$

In this case, we have

$$(\Phi''(y)[h_1,h_2])(x) = \varphi_{yy}(x,y(x))h_1(x)h_2(x),$$

since $h = h_1h_2$ belongs to $L^{\frac{p}{2}}(\Omega)$ for $h_1, h_2 \in L^p(\Omega)$, the formula for r is evident.

3 Optimal Control of the Stationary Kirchhoff Equation

Contents

In this chapter we study an optimal control problem governed by a nonlinear nonlocal elliptic partial differential equation (3.1.1b). As explained in subsection 1.1.1, the equation (3.1.1b) is the steady-state problem corresponding to its time-dependent counterpart, given by (1.1.2). In one space dimension (1.1.2) models small vertical vibrations of an elastic string with fixed ends, when the density of the material is not constant. Specifically, the control u is proportional to the inverse of the string's cross section; see Ma, 2005; Figueiredo et al., 2014.

This chapter is organized as follows. In section 3.1, we review existence and uniqueness results for solutions of the Kirchhoff equation (1.1.1) and prove the existence of a globally optimal control. Subsequently, we prove the Fréchet differentiability of the control-to-state operator and derive a system of necessary optimality conditions for a regularized problem in section 3.2. We also presented an analytical solution in this section. In section 3.3, we prove the Newton differentiability of the optimality system and devise a locally superlinearly convergent scheme in appropriate function spaces. Section 3.4 addresses the discretization of the optimal control problem, its optimality system and the generalized Newton method by a finite element scheme. The chapter concludes with numerical results in section 3.5.

We remark that main part of this chapter has been taken from the following published paper.

- Hashemi Masoumeh, Herzog Roland, Surowiec Thomas M.;
  Optimal Control of the Stationary Kirchhoff Equation;
  Computational Optimization and Applications;
  https://doi.org/10.1007/s10589-023-00463-6.

Here, we provide a list of what is additional to the papar.

(i) Proof of theorem 3.1.6
(ii) Proof of proposition 3.2.1
(iii) An analytical solution of the optimal control problem, example 3.2.4
(iv) Construction of a cut-off function for penalization of the optimal control problem, example 3.2.5
(v) Proof of theorem 3.2.6, statement (i)
(vi) Investigation of the influence of the control cost parameters, as a numerical experiment in subsection 3.5.4

For convenience, we have included the essential results mentioned in the chapter 2 but relevant to the current chapter.

3.1 Optimal Control Problem: Existence of a Solution

In this work we are interested in the study of the following optimal control problem for a stationary nonlinear, nonlocal Kirchhoff equation:

$$\text{Minimize} \quad J(y,u) \coloneqq \int_\Omega \varphi(x,y(x))\,\mathrm{d}x + \frac{\lambda}{2}\|u\|^2_{H^1(\Omega)} \tag{3.1.1a}$$

$$\text{subject to} \quad \begin{cases} -\left(u + b\,\|\nabla y\|^2_{L^2(\Omega)}\right)\Delta y = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega \end{cases} \tag{3.1.1b}$$

$$\text{and} \quad u \in \mathcal{U}_{\mathrm{ad}}. \tag{3.1.1c}$$

The set of admissible controls is given by

$$\mathcal{U}_{\mathrm{ad}} = \{u \in H^1(\Omega) \,|\, u_a(x) \le u(x) \le u_b(x) \text{ a.e. in } \Omega\}. \tag{3.1.2}$$

The following are our standing assumptions.

Assumption 3.1.1. *We assume that $\Omega \subset \mathbb{R}^N$ is a bounded domain of class $C^{1,1}$ with $1 \le N \le 3$; see for instance Tröltzsch, 2010, Chapter 2.2.2. The control cost parameter λ is a positive number. The right-hand side f is a given function in $L^\infty(\Omega)$ satisfying $f \ge f_0$ a.e., where f_0 is a positive real number. The bounds u_a and u_b are functions in $C(\overline{\Omega})$ such that $u_b \ge u_a \ge u_0$ holds for some positive real number u_0. Finally, we assume $b \in L^\infty(\Omega)$ with $b \ge b_0$ a.e. for some positive real number b_0.*

The integrand φ in the objective is assumed to satisfy the following standard assumptions; see for instance Tröltzsch, 2010, Chapter 4.3:

Assumption 3.1.2. *(1) $\varphi\colon \Omega \times \mathbb{R} \to \mathbb{R}$ is Carathéodory and of class C^2, i. e.,*

(i) $\varphi(\cdot, y)\colon \Omega \to \varphi(x,y)$ is measurable for all $y \in \mathbb{R}$,

(ii) $\varphi(x,\cdot)\colon \mathbb{R} \to \varphi(x,y)$ is twice continuously differentiable for a.e. $x \in \Omega$.

(2) φ satisfies the boundedness and local Lipschitz conditions of order 2, i. e., there exists a constant $K > 0$ such that

$$\left|D^\ell_y \varphi(x,0)\right| \le K \quad \text{for all } 0 \le \ell \le 2 \text{ and for a.e. } x \in \Omega,$$

and for every $M > 0$, there exists a Lipschitz constant $L(M) > 0$ such that

$$\left|D^2_y \varphi(x,y_1) - D^2_y \varphi(x,y_2)\right| \le L(M)\,|y_1 - y_2|$$

holds for a.e. $x \in \Omega$ and for all $|y_i| \le M$, $i = 1,2$.

Assumption 3.1.2 implies the following properties for the Nemytskii operator $\Phi(y)(x) \coloneqq \varphi(x, y(x))$.

Lemma 3.1.3 (Tröltzsch, 2010, Lemma 4.11, Lemma 4.12).

(i) Φ is continuous in $L^\infty(\Omega)$. Moreover, for all $r \in [1,\infty]$, we have

$$\|\Phi(y) - \Phi(z)\|_{L^r(\Omega)} \le L(M)\,\|y - z\|_{L^r(\Omega)}$$

for all $y, z \in L^\infty(\Omega)$ such that $\|y\|_{L^\infty(\Omega)} \le M$ and $\|z\|_{L^\infty(\Omega)} \le M$.

(ii) Φ is twice continuously Fréchet differentiable in $L^\infty(\Omega)$, and we have

$$\left(\Phi'(y)\,h\right)(x) = \varphi_y\left(x, y(x)\right) h(x),$$

$$\left(\Phi''(y)\,[h_1,h_2]\right)(x) = \varphi_{yy}\left(x, y(x)\right) h_1(x)\,h_2(x)$$

for a.e. $x \in \Omega$ and $h, h_1, h_2 \in L^\infty(\Omega)$.

We now proceed to define the notion of weak solution of the Kirchhoff problem. Since for any pair $(u, y) \in \mathcal{U}_{\text{ad}} \times H^1(\Omega)$, $u + b\,\|\nabla y\|^2_{L^2(\Omega)}$ is strictly positive, we can write equation (3.1.1b) in the form

$$-\Delta y = \frac{f}{u + b\,\|\nabla y\|^2_{L^2(\Omega)}}. \tag{3.1.3}$$

Here and in the following, we occasionally write $\|\cdot\|$ instead of $\|\cdot\|_{L^2(\Omega)}$. The $L^2(\Omega)$-inner product is denoted by $(\cdot,\cdot)$. Moreover, we denote by $\mathcal{L}(U, V)$ the space of bounded linear operators from U to V.

Multiplication of (3.1.3) with a test function $v \in H^1_0(\Omega)$ and integration by parts yields the following definition.

Definition 3.1.4. *A function $y \in H^1_0(\Omega)$ is called a weak solution of (3.1.3) if it satisfies*

$$\int_\Omega \nabla y \cdot \nabla v \,\mathrm{d}x = \int_\Omega \frac{f\,v}{u + b\,\|\nabla y\|^2} \,\mathrm{d}x \quad \text{for all } v \in H^1_0(\Omega). \tag{3.1.4}$$

The existence of a unique weak solution as well as its $W^{2,q}(\Omega)$-regularity has been shown in Delgado, Figueiredo, et al., 2017, Theorem 2.2. Nevertheless, we briefly sketch the proof since its main idea is utilized again later on.

To achieve $W^{2,q}(\Omega)$-regularity we need the following result, which can be deduced from e. g., Gilbarg, Trudinger, 1977, Theorem 9.15, Theorem 9.17:

Theorem 3.1.5. *Suppose that Ω is a bounded $C^{1,1}$ domain and $\mathcal{A}$ is an elliptic differential operator of the form*

$$\mathcal{A}y(x) = -\sum_{i,j=1}^{n} (a_{ij}(x) y_{x_j}(x))_{x_i}, \quad x \in \Omega.$$

The coefficient functions a_{ij} of $\mathcal{A}$ are assumed to belong to $C^{0,1}(\overline{\Omega})$ and satisfy the condition of symmetry $a_{ij}(x) = a_{ji}(x)$ for all $i, j \in \{1, \dots, n\}$ and $x \in \Omega$.

If the right-hand side function $f \in L^q(\Omega)$, $1 < q < \infty$, then the weak solution to the following Dirichlet problem

$$\begin{cases} \mathcal{A}y = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases}$$

belongs to $W^{2,q}(\Omega)$ and it holds

$$\|y\|_{W^{2,q}(\Omega)} \le C\, \|\mathcal{A}(y)\|_{L^q(\Omega)} \quad \text{for some } C > 0.$$

Theorem 3.1.6. *For any $u \in \mathcal{U}_{\text{ad}}$, there exists a unique weak solution $y \in H^1_0(\Omega)$ of the Kirchhoff problem (3.1.3). Moreover, $y \in W^{2,q}(\Omega)$ holds for all $q \in [1, \infty)$, so it is also a strong solution.*

Proof. Suppose that $u \in \mathcal{U}_{\text{ad}}$ and let $g\colon [0, \infty) \to \mathbb{R}$ be the function defined by

$$g(s) = s - \|\nabla y_s\|^2,$$

where y_s is the unique weak solution of the Poisson problem

$$\begin{cases} -\Delta y_s = \dfrac{f}{u + b\,s} & \text{in } \Omega, \\ y_s = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.1.5}$$

A monotonicity argument can be used to show that g has a unique root.

Step (1): First, we show that g is continuous. Given $s \in [0,\infty)$. The function $u + b\,s$ is strictly greater than zero, consequently $\frac{f}{u+b\,s}$ belongs to $L^\infty(\Omega)$. That means, (3.1.5) possesses a unique weak solution $y_s \in H_0^1(\Omega)$, by virtue of Lax-Milgram Theorem, theorem 2.1.40.
Now, we consider a non-negative number sequence $\{s_k\}$, $k \in \mathbb{N}$, its elements fulfill (3.1.5) and converges to $\bar{s}$, which means $|s_k - \bar{s}| \to 0$, as $k \to 0$. We have

$$\begin{aligned} |g(s_k) - g(\bar{s})| &= \left| s_k - \|\nabla y_{s_k}\|^2 - \bar{s} + \|\nabla y_{\bar{s}}\|^2 \right| \\ &\leq |s_k - \bar{s}| + \left| \|\nabla y_{s_k}\|^2 - \|\nabla y_{\bar{s}}\|^2 \right| \end{aligned} \tag{3.1.6}$$

Set $w_k := y_{s_k} - y_{\bar{s}}$. Inserting this into (3.1.5) results in

$$\begin{aligned} \Delta w_k = \Delta y_{s_k} - \Delta y_{\bar{s}} &= -\frac{f}{u + b\,s_k} + \frac{f}{u + b\,\bar{s}} \\ &= f\left(\frac{1}{u + b\,\bar{s}} - \frac{1}{u + b\,s_k} \right). \end{aligned}$$

By means of Lax-Milgram Theorem and embeddings theorem in L^p spaces we obtain the following estimate

$$\begin{aligned} \|w_k\|_{H_0^1(\Omega)} &\leq c_1 \left\| f\left(\frac{1}{u + b\,\bar{s}} - \frac{1}{u + b\,s_k} \right) \right\|_{L^2(\Omega)} \\ &\leq c_1 \, \|f\|_{L^2(\Omega)} \left\| \frac{1}{u + b\,\bar{s}} - \frac{1}{u + b\,s_k} \right\|_{L^\infty(\Omega)} \\ &\leq c_2 \left\| \frac{b\,(s_k - \bar{s})}{(u + b\,\bar{s})\,(u + b\,s_k)} \right\|_{L^\infty(\Omega)} \\ &\leq c_3 \, |s_k - \bar{s}| \left\| \frac{b}{u_a^2} \right\|_{L^\infty(\Omega)} \leq c \, |s_k - \bar{s}| \,. \end{aligned}$$

Passing to the limit as $k \to 0$, we find $\|w_k\|_{H_0^1(\Omega)} = \|y_{s_k} - y_{\bar{s}}\|_{H_0^1(\Omega)} \to 0$. Since

$$\left| \|y_{s_k}\|_{H_0^1(\Omega)} - \|y_{\bar{s}}\|_{H_0^1(\Omega)} \right| \leq \|y_{s_k} - y_{\bar{s}}\|_{H_0^1(\Omega)}$$

we obtain $\|y_{s_k}\|_{H_0^1(\Omega)} \to \|y_{\bar{s}}\|_{H_0^1(\Omega)}$. This shows the continuity of g, due to (3.1.6).

Step (2): As next step, we prove that g is strictly monotonically increasing. To this end, let s_1 and s_2 be two elements in $[0,\infty)$, with $s_1 < s_2$. Then

$$\frac{f}{u + b\,s_1} \geq \frac{f}{u + b\,s_2} \geq 0,$$

which means $-\Delta y_{s_1} \geq -\Delta y_{s_2} \geq 0$. By the maximum principle, we obtain $y_{s_1} \geq 0$ and $y_{s_2} \geq 0$.
On the other hand $-\Delta y_{s_1} \geq -\Delta y_{s_2}$ results in $-\Delta\,(y_{s_1} - y_{s_2}) \geq 0$ and again by the maximum principle $y_{s_1} - y_{s_2} \geq 0$. Accordingly, we obtain $-\Delta y_{s_1}\, y_{s_1} \geq -\Delta y_{s_2}\, y_{s_2}$, hence $(-\Delta y_{s_1}, y_{s_1}) \geq (-\Delta y_{s_2}, y_{s_2})$. Using Green's formula, we can infer $\|\nabla y_{s_1}\|^2 \geq \|\nabla y_{s_2}\|^2$. From $s_1 < s_2$, follows $s_1 - \|\nabla y_{s_1}\|^2 < s_2 - \|\nabla y_{s_2}\|^2$, which means $g(s_1) < g(s_2)$.

Step (3): As last step, we show that g changes its sign between zero and some $s > 0$, and it can occur merely one time, due to the strict monotonicity of g. $g(0) = -\|\nabla y_0\|^2 <$

0, where y_0 solves (3.1.5). Multiplying this equation by a test function y_s and integrating by parts, we obtain

$$\|\nabla y_s\|^2 = \int_\Omega |\nabla y_s|^2 \,\mathrm{d}x = \int_\Omega \frac{f}{u + b\, s}\, y_s \,\mathrm{d}x.$$

The term $\int_\Omega \frac{f}{u+b\,s}\, y_s \,\mathrm{d}x$ is bounded above by $\int_\Omega \frac{f}{u_a}\, y_0$, which results in the boundedness of $\|\nabla y_s\|^2$, for all $s > 0$. Therefore, $\lim_{s\to\infty} g(s) = \lim_{s\to\infty} \left(s - \|\nabla y_s\|^2\right) = +\infty$ and this concludes the assertion that g has a unique root.

Since y_s solves (3.1.3) if and only if $g(s) = 0$ holds, the uniqueness of the solution of the Kirchhoff equation is guaranteed. Furthermore, due to the boundedness of u from below, the right-hand side $f/(u+b\,s)$ of the Poisson problem above belongs to $L^\infty(\Omega)$. Hence, by virtue of regularity results for the Poisson problem, $y \in W^{2,q}(\Omega)$ holds for any $q \in [1,\infty)$. □

For the proof of existence of a globally optimal control of (3.1.1), we show next that the control-to-state operator $S\colon \mathcal{U}_{\mathrm{ad}} \to H_0^1(\Omega) \cap W^{2,q}(\Omega)$ is continuous.

Theorem 3.1.7. *The control-to-state map S is continuous from $\mathcal{U}_{\mathrm{ad}}$ (with the $L^2(\Omega)$-topology) into $H_0^1(\Omega) \cap W^{2,q}(\Omega)$ for all $q \in [1,\infty)$.*

Proof. The control-to-state map $S\colon \mathcal{U}_{\mathrm{ad}} \to H_0^1(\Omega) \cap W^{2,q}(\Omega)$ is well-defined as a consequence of theorem 3.1.6. To show its continuity, let $\{u_n\} \subset \mathcal{U}_{\mathrm{ad}}$ be a sequence with $u_n \to u$ in $L^2(\Omega)$. Set $y_n := S(u_n)$, that means

$$-\Delta y_n = \frac{f}{u_n + b\,\|\nabla y_n\|^2},$$

then we have the a-priori estimate, by virtue of theorem 3.1.5

$$\begin{aligned}\|y_n\|_{W^{2,q}(\Omega)} &\le c_1 \left\| \frac{f}{u_n + b\,\|\nabla y_n\|^2} \right\|_{L^q(\Omega)} \le c_2 \left\| \frac{f}{u_n + b\,\|\nabla y_n\|^2} \right\|_{L^\infty(\Omega)} \\ &\le c_2 \left\| \frac{f}{u_n} \right\|_{L^\infty(\Omega)} \le c_2 \left\| \frac{f}{u_a} \right\|_{L^\infty(\Omega)} \le C.\end{aligned}$$

From now on, suppose without loss of generality that $q \in [2,\infty)$ holds. Since $W^{2,q}(\Omega)$ is a reflexive Banach space and every bounded subset of a reflexive Banach space is weakly relatively compact, there exists a subsequence y_n, denoted by the same indices, satisfying $y_n \rightharpoonup \hat{y}$ in $W^{2,q}(\Omega)$. The compactness of the embedding $W^{2,q}(\Omega) \hookrightarrow W^{1,q}(\Omega)$ implies the strong convergence $y_n \to \hat{y}$ in $W^{1,q}(\Omega)$ and thus $\nabla y_n \to \nabla \hat{y}$ in $L^q(\Omega)$. From

$$|\|\nabla y_n\| - \|\nabla \hat{y}\|| \le \|\nabla y_n - \nabla \hat{y}\| \le \|\nabla y_n - \nabla \hat{y}\|_{L^q(\Omega)}$$

follows $\|\nabla y_n\| \to \|\nabla \hat{y}\|$.

On the other hand, $u_n \to u$ in $L^2(\Omega)$ implies the existence of a further subsequence u_n, still denoted by the same indices, with $u_n(x) \to u(x)$ for a.e. $x \in \Omega$. Consequently,

$$\frac{f}{u_n + b\,\|\nabla y_n\|^2} \to \frac{f}{u + b\,\|\nabla \hat{y}\|^2} \quad \text{a.e. in } \Omega.$$

Since $\frac{f}{u_n + b\,\|\nabla y_n\|^2}$ is dominated by $\frac{f}{u_a}$, we have

$$\begin{aligned}\left| \frac{f}{u_n + b\,\|\nabla y_n\|^2} - \frac{f}{u + b\,\|\nabla \hat{y}\|^2} \right| &\le \left| \frac{f}{u_n + b\,\|\nabla y_n\|^2} \right| + \left| \frac{f}{u + b\,\|\nabla \hat{y}\|^2} \right| \\ &\le \left| \frac{f}{u_a} \right| + \left| \frac{f}{u_a} \right| = 2 \left| \frac{f}{u_a} \right|,\end{aligned}$$

hence

$$\left| \frac{f}{u_n + b\,\|\nabla y_n\|^2} - \frac{f}{u + b\,\|\nabla \hat{y}\|^2} \right|^q \leq \left| \frac{2f}{u_a} \right|^q.$$

By virtue of the dominated convergence theorem,

$$-\Delta y_n = \frac{f}{u_n + b\,\|\nabla y_n\|^2} \to \frac{f}{u + b\,\|\nabla \hat{y}\|^2} \quad \text{in } L^q(\Omega).$$

On the other hand, from $y_n \rightharpoonup \hat{y}$ in $W^{2,q}(\Omega)$, it follows that $\Delta y_n \rightharpoonup \Delta \hat{y}$ holds in $L^q(\Omega)$. The uniqueness of the weak limit yields

$$-\Delta \hat{y} = \frac{f}{u + b\,\|\nabla \hat{y}\|^2}$$

and from the uniqueness of the solution of (3.1.3) we obtain $\hat{y} = S(u)$. Therefore, $\Delta y_n \to \Delta \hat{y}$ holds in $L^q(\Omega)$ and thereby $y_n \to \hat{y}$ in $W^{2,q}(\Omega)$.

We note that we have proved that for any sequence $\{u_n\} \subset \mathcal{U}_{\text{ad}}$ with $u_n \to u$ in $L^2(\Omega)$ there exists a subsequence $\{u_n\}$, denoted by the same indices, so that $S(u_n) \to S(u)$ in $W^{2,q}(\Omega)$. Thus, we can easily conclude convergence of the entire sequence $S(u_n) \to S(u)$ in $W^{2,q}(\Omega)$. Indeed, if $S(u_n) \not\to S(u)$, then there exist $\delta > 0$ and a subsequence with indices n_k such that

$$\|S(u_{n_k}) - S(u)\|_{W^{2,q}(\Omega)} > \delta \text{ for } k \to \infty.$$

Since $u_{n_k} \to u$ in $L^2(\Omega)$, there exists a further subsequence $\{u_{n_{k_\ell}}\}$ such that $S(u_{n_{k_\ell}}) \to S(u)$, which is a contradiction. Consequently, we obtain $S(u_n) \to S(u)$ as claimed. □

The compact embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$ immediately leads to the following corollary.

Corollary 3.1.8. *The control-to-state map S is weakly-strongly continuous from $\mathcal{U}_{\text{ad}}$ (with the $H^1(\Omega)$-topology) into $H_0^1(\Omega) \cap W^{2,q}(\Omega)$ for all $q \in [1, \infty)$. That is, when $\{u_n\} \subset \mathcal{U}_{\text{ad}}$ with $u_n \rightharpoonup u$ in $H^1(\Omega)$, then $S(u_n) \to S(u)$ in $W^{2,q}(\Omega)$.*

We can now address the existence of a global minimizer of (3.1.1).

Theorem 3.1.9. *Problem (3.1.1) possesses a globally optimal control $\bar{u} \in \mathcal{U}_{\text{ad}}$ with associated optimal state $\bar{y} = S(\bar{u}) \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$ for all $q \in [1, \infty)$.*

Proof. The proof follows the standard route of the direct method, so we can be brief. Step (1): We show that the reduced cost functional

$$j(u) := \int_\Omega \Phi(S(u))\,\mathrm{d}x + \frac{\lambda}{2}\|u\|^2_{H^1(\Omega)}$$

is bounded from below on the set $\mathcal{U}_{\text{ad}}$. To this end, recall

$$\begin{aligned} \|S(u)\|_{W^{2,q(\Omega)}} = \|y\|_{W^{2,q}(\Omega)} &\leq c_1 \left\| \frac{f}{u + b\,\|\nabla y\|^2} \right\|_{L^q(\Omega)} \leq c_2 \left\| \frac{f}{u + b\,\|\nabla y\|^2} \right\|_{L^\infty(\Omega)} \\ &\leq c \left\| \frac{f}{u} \right\|_{L^\infty(\Omega)} \leq c \left\| \frac{f}{u_a} \right\|_{L^\infty(\Omega)} \leq C. \end{aligned}$$

That means, $S(\mathcal{U}_{\text{ad}})$ is bounded in $W^{2,q}(\Omega)$. Due to the embedding $W^{2,q}(\Omega) \hookrightarrow C(\overline{\Omega})$ for $q > N/2$, there exists $M > 0$ such that $\|S(u)\|_{L^\infty(\Omega)} \leq M$ holds for all $u \in \mathcal{U}_{\text{ad}}$. From Assumption 3.1.2 we can obtain the estimate

$$\begin{aligned} |\varphi(x, S(u)(x))| &= |\varphi(x, S(u)(x)) - \varphi(x, 0) + \varphi(x, 0)| \\ &\leq |\varphi(x, 0)| + |\varphi(x, S(u)(x)) - \varphi(x, 0)| \\ &\leq K + L(M)\,|S(u)(x)| \leq K + L(M)\,M. \end{aligned}$$

This implies

$$\int_\Omega \Phi(S(u))\,\mathrm{d}x \geq -\big(K + L(M)\,M\big)\,|\Omega| \tag{3.1.7}$$

for all $u \in \mathcal{U}_{\mathrm{ad}}$. The assertion follows.

Step (2): We construct the tentative minimizer $\bar{u}$. Since j is bounded from below on $\mathcal{U}_{\mathrm{ad}}$, there exists a minimizing sequence $\{u_n\} \subset \mathcal{U}_{\mathrm{ad}}$ so that

$$j(u_n) \searrow \inf_{u\in\mathcal{U}_{\mathrm{ad}}} j(u) =: \beta.$$

$\{u_n\}$ is bounded in $H^1(\Omega)$. Consequently, there exists a subsequence, denoted by the same indices, such that $u_n \rightharpoonup \bar{u}$ in $H^1(\Omega)$. $\mathcal{U}_{\mathrm{ad}}$ is convex and closed in $H^1(\Omega)$ and therefore weakly closed in $H^1(\Omega)$, thus $\bar{u} \in \mathcal{U}_{\mathrm{ad}}$. Now corollary 3.1.8 implies $S(u_n) \to S(\bar{u})$ in $W^{2,q}(\Omega)$.

Step (3): It remains to show the global optimality of $\bar{u}$. Set $F(y) := \int_\Omega \Phi(y)\,\mathrm{d}x$, thus F is composed of a Nemytskii operator and a continuous linear integral operator from $L^1(\Omega)$ into $\mathbb{R}$. By virtue of lemma 3.1.3, Φ is continuous in $L^\infty(\Omega)$. Since $W^{2,q}(\Omega) \hookrightarrow L^\infty(\Omega)$ holds, $\Phi \circ S$ is weakly-strongly continuous on $\mathcal{U}_{\mathrm{ad}}$ w.r.t. the topology of $H^1(\Omega)$. Therefore, $F \circ S = \int_\Omega \Phi \circ S\,\mathrm{d}x$ is weakly-strongly continuous on $\mathcal{U}_{\mathrm{ad}}$.

In summary, exploiting the weak sequential lower semicontinuity of $\|\cdot\|_{H^1}$ we have

$$\begin{aligned}
\beta = \lim_{n\to\infty} j(u_n) &= \lim_{n\to\infty} F(S(u_n)) + \frac{\lambda}{2} \lim_{n\to\infty} \|u_n\|_{H^1}^2 \\
&\geq \lim_{n\to\infty} F(S(u_n)) + \frac{\lambda}{2} \liminf_{n\to\infty} \|u_n\|_{H^1}^2 \\
&\geq F(S(\bar{u})) + \frac{\lambda}{2} \|\bar{u}\|_{H^1(\Omega)}^2 = j(\bar{u}).
\end{aligned}$$

By definition of β and since $\bar{u} \in \mathcal{U}_{\mathrm{ad}} \cap H^1(\Omega)$, we therefore must have $\beta = j(\bar{u})$. □

Remark 3.1.10. *An inspection of the existence theory shows that these results remain valid in the absence of an upper bound u_b on the control. However, the upper bound is of essential importance in the following section, where we prove the Fréchet differentiability of the control-to-state map.*

3.2 Optimality System

In this section we address first-order necessary optimality conditions for local minimizers. We need to overcome several obstacles. First of all, the control-to-state operator

$$S\colon \mathcal{U}_{\mathrm{ad}} \to H_0^1(\Omega) \cap W^{2,q}(\Omega)$$

is well-defined and continuous on $\mathcal{U}_{\mathrm{ad}}$ with respect to the topology of $H^1(\Omega)$, but this operator is not Fréchet differentiable in the $H^1(\Omega)$-topology. The reason is that $\mathcal{U}_{\mathrm{ad}}$ has empty interior w.r.t. this topology except in dimension $N = 1$, which means we cannot define S on any open set with respect to the $H^1(\Omega)$-topology. More precisely, every $H^1(\Omega)$-neighborhood of any control $u \in \mathcal{U}_{\mathrm{ad}}$ contains functions which are arbitrarily negative on sets of small but positive measure. However, the proof of theorem 3.1.6, which establishes the well-definedness of the control-to-state map, is contingent upon the controls to remain positive. In order to overcome this issue, we work with the topology of $H^1(\Omega) \cap L^\infty(\Omega)$. Therefor, it is essential that the controls belong to $L^\infty(\Omega)$, which is evident in the presence of an upper bound.

With regard to an efficient numerical solution method in function spaces, we are aiming to arrive at an optimality system which is Newton differentiable. To this end, we propose

to relax and penalize the control constraint. Notice that this is not straightforward since we need to ensure positivity of the relaxed control in the state equation. We achieve the latter by a smooth cut-off function. The optimality system of the penalized problem then turns out to be Newton differentiable, as we shall show in section 3.3.

The material in this section is structured as follows. In subsection 3.2.1, we prove the Fréchet differentiability of the control-to-state map. We establish the system of first-order necessary optimality conditions for the original problem (3.1.1) in subsection 3.2.2. An analytical solution is constructed in subsection 3.2.3. In subsection 3.2.4 we introduce the penalty approximation and show that for any null sequence of penalty parameters, there exists a subsequence of global solutions to the corresponding penalized problems which converges weakly to a global solution of the original problem; see theorem 3.2.6. subsection 3.2.5 addresses the system of first-order necessary optimality conditions for the penalized problem.

3.2.1 Differentiability of the Control-to-State Map

In this subsection we show the Fréchet differentiability of the control-to-state map S by means of the implicit function theorem. To verify the assumption of the implicit function theorem, we need the following result about the linearization of the Kirchhoff equation (3.1.3).

Proposition 3.2.1. *Suppose that $\hat{u} \in \mathcal{U}_{\mathrm{ad}}$ and $\hat{y} \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$ is the associated unique solution of the Kirchhoff equation (3.1.3) for any $q \in [1, \infty)$. Then, for any $h \in L^q(\Omega)$, the linearized problem*

$$\begin{cases} -\Delta y - \dfrac{2\,b\,(\nabla \hat{y}, \nabla y)\Delta \hat{y}}{(\hat{u} + b\,\|\nabla \hat{y}\|^2)} = h & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.2.1}$$

has a unique solution $y \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$.

Proof. For a given $h \in L^q(\Omega)$, we set $h := h^+ - h^-$, where $h^+ := \max\{0, h\}$ and $h^- := -\min\{0, h\}$ are the positive part and negative part of h, respectively. We use Green's formula for $(\nabla \hat{y}, \nabla y)$ in (3.2.1) and obtain $(\nabla \hat{y}, \nabla y) = (-\Delta \hat{y}, y)$. That means, (3.2.1) is modified in the following form

$$\begin{cases} -\Delta y - \dfrac{2\,b\,(-\Delta \hat{y}, y)\Delta \hat{y}}{(\hat{u} + b\,\|\nabla \hat{y}\|^2)} = h^+ - h^- & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.2.2}$$

We consider the following subproblems

$$\begin{cases} -\Delta y - \dfrac{2\,b\,(-\Delta \hat{y}, y)\Delta \hat{y}}{(\hat{u} + b\,\|\nabla \hat{y}\|^2)} = h^+ & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.2.3}$$

and

$$\begin{cases} -\Delta y - \dfrac{2\,b\,(-\Delta \hat{y}, y)\Delta \hat{y}}{(\hat{u} + b\,\|\nabla \hat{y}\|^2)} = -h^- & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.2.4}$$

Let $g^+ \colon [0, \infty)$ be the function defied by

$$g^+(s) = s - (-\Delta \hat{y}, y_s),$$

where y_s solves

$$\begin{cases} -\Delta y_s - \dfrac{2\,b\,\Delta \hat{y}\,s}{(\hat{u} + b\,\|\nabla \hat{y}\|^2)} = h^+ & \text{in } \Omega, \\ y_s = 0 & \text{on } \partial\Omega, \end{cases}$$

and $g^-\colon (-\infty, 0]$ be the function defied by

$$g^-(s) = s - (-\Delta\hat{y}, y_s),$$

where y_s solves

$$\begin{cases} -\Delta y_s - \dfrac{2\,b\,\Delta\hat{y}\,s}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)} = -h^- & \text{in } \Omega, \\ y_s = 0 & \text{on } \partial\Omega. \end{cases}$$

We use a monotonicity argument again to show that g^+ and g^- have unique roots on their respective domains.

Step (1): We show that g^+ and g^- are continuous. Let $\{s_k\}$ be a number sequence in each case. We have

$$\begin{aligned} \left|g^{\pm}(s_k) - g(s)\right| &= |s_k - (-\Delta\hat{y}, y_{s_k}) - s + (-\Delta\hat{y}, y_s)| \\ &\le |s_k - s| + |(\Delta\hat{y}, y_{s_k} - y_s)| \\ &\le |s_k - s| + \|\Delta\hat{y}\|\ \|y_{s_k} - y_s\| \quad \text{(Cauchy-Schwarz)} \\ &\le |s_k - s| + \|\Delta\hat{y}\|\ \|y_{s_k} - y_s\|_{W^{2,q}(\Omega)}\,. \end{aligned}$$

Since

$$-\Delta\left(y_{s_k} - y_s\right) = \frac{2\,b\,\Delta\hat{y}}{\left(\hat{u} + b\,\|\nabla\hat{y}\|^2\right)}\,(s_k - s)\,,$$

we obtain

$$\|y_{s_k} - y_s\|_{W^{2,q}(\Omega)} \le c\left\|\frac{2\,b\,\Delta\hat{y}}{\left(\hat{u} + b\,\|\nabla\hat{y}\|^2\right)}\,(s_k - s)\right\|_{L^q(\Omega)} \le c'\,|s_k - s|\,,$$

by virtue of theorem 3.1.5, which results in

$$\left|g^{\pm}(s_k) - g(s)\right| \le |y_{s_k} - y_s| + c'\,|y_{s_k} - y_s| \le c\,|y_{s_k} - y_s|\,.$$

Passing to the limit as k converges to zero, we obtain $g^{\pm}(s_k) \to g^{\pm}(s)$.

Step (2): We show that g^+ and g^- are strictly monotonically increasing in each case with $s_1 < s_2$. We have

$$-\Delta\left(y_{s_1} - y_{s_2}\right) = \frac{2\,b\,\Delta\hat{y}}{\left(\hat{u} + b\,\|\nabla\hat{y}\|^2\right)}\,(s_1 - s_2)\,.$$

Since $s_1 - s_2 < 0$ and $\Delta\hat{y} \le 0$, we obtain $-\Delta\left(y_{s_1} - y_{s_2}\right) \ge 0$. By the maximum principle, we can infer $y_{s_1} - y_{s_2} \ge 0$, which results in

$$g^{\pm}(s_1) = s_1 - (-\Delta\hat{y}, y_{s_1}) < s_2 - (-\Delta\hat{y}, y_{s_2}) = g^{\pm}(s_2).$$

Step (3): We show that g^+ and g^- change their sings at least one time on $(0, -\infty)$ and $(-\infty, 0)$, respectively. We have $g^+(0) = -(-\Delta\hat{y}, y_0)$, where y_0 solves

$$\begin{cases} -\Delta y_0 = h^+ & \text{in } \Omega, \\ y_0 = 0 & \text{on } \partial\Omega. \end{cases}$$

This means, $-\Delta y_0 > 0$. By the maximum principle we obtain $y_0 > 0$ and thus $g^+(0) < 0$. Therefor, for all $s > 0$ ($y_s \le y_0$), we obtain $-\Delta\hat{y}\,y_s \le -\Delta\hat{y}\,y_0$ and hence $(-\Delta\hat{y}, y_s) \le (-\Delta\hat{y}, y_0)$. That means

$$\lim_{s\to+\infty} g^+(s) = \lim_{s\to+\infty}\left(s - (-\Delta\hat{y}, y_s)\right) = +\infty.$$

On the other hand, We have $g^-(0) = -(-\Delta\hat{y}, y_0)$, where y_0 solves

$$\begin{cases} -\Delta y_0 = -h^- & \text{in } \Omega, \\ y_0 = 0 & \text{on } \partial\Omega. \end{cases}$$

This means, $-\Delta y_0 < 0$. By the maximum principle we obtain $y_0 < 0$ and thus $g^-(0) > 0$. Therefor, for all $s < 0$ ($y_s \geq y_0$), we obtain $-\Delta\hat{y}\, y_s \geq -\Delta\hat{y}\, y_0$ and hence $(-\Delta\hat{y}, y_s) \geq (-\Delta\hat{y}, y_0)$. That means

$$\lim_{s\to-\infty} g^+(s) = \lim_{s\to-\infty} (s - (-\Delta\hat{y}, y_s)) = -\infty.$$

Now by monotonicity of $g^\pm$, we can conclude that g^+ and g^- possess unique roots s^+ and s^-, respectively.

y_{s^+} solves (3.2.3) and y_{s^-} (3.2.4) if and only if $g^+(s^+) = 0$ and $g^-(s^-) = 0$ hold, respectively. Therefore, the uniqueness of solutions y^+ and y^- for the respective subproblems (3.2.3) and (3.2.4) is guaranteed. At the end, we show that $y \coloneqq y^+ + y^-$ solves (3.2.2).

$$\begin{aligned} -\Delta(y^+ + y^-) - \frac{2\,b\,(-\Delta\hat{y}, y^+ + y^-)\Delta\hat{y}}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)} \\ = -\Delta y^+ - \frac{2\,b\,(-\Delta\hat{y}, y^+)\Delta\hat{y}}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)} - \Delta y^- - \frac{2\,b\,(-\Delta\hat{y}, y^-)\Delta\hat{y}}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)} = h^+ - h^-. \end{aligned}$$

□

Theorem 3.2.2. *Suppose that $\hat{u} \in \mathcal{U}_{\text{ad}}$. Then the control-to-state operator*

$$S\colon \mathcal{U}_{\text{ad}} \to H_0^1(\Omega) \cap W^{2,q}(\Omega)$$

is continuously Fréchet differentiable on an open $L^\infty(\Omega)$-neighborhood of $\hat{u}$ for all $q \in [1, \infty)$.

Proof. Suppose that $\hat{u} \in \mathcal{U}_{\text{ad}}$ is arbitrary and that $\hat{y} \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$ is the associated state. The map $E\colon \big(H_0^1(\Omega) \cap W^{2,q}(\Omega)\big) \times L^\infty(\Omega) \to L^q(\Omega)$ defined by

$$E(y, u) \coloneqq -\Delta y - \frac{f}{u + b\,\|\nabla y\|^2}$$

is continuously Fréchet differentiable with

$$E'(\hat{y}, \hat{u})(y, u) = -\Delta y + \frac{\big(u + 2\,b\,(\nabla\hat{y}, \nabla y)\big) f}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)^2},$$

It remains to show that $E_y(\hat{y}, \hat{u}) \in \mathcal{L}\big(H_0^1(\Omega) \cap W^{2,q}(\Omega), L^q(\Omega)\big)$ has a bounded inverse. To this end, consider

$$E_y(\hat{y}, \hat{u})\, y = -\Delta y + \frac{2\,b\,f\,(\nabla\hat{y}, \nabla y)}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)^2}. \tag{3.2.5}$$

The existence and uniqueness of $y \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$ satisfying (3.2.1), i. e., $E_y(\hat{y}, \hat{u})\, y = h$, is established by virtue of proposition 3.2.1. This implies the bijectivity of $E_y(\hat{y}, \hat{u})$. The open mapping/continuous inverse theorem now yields that the inverse of $E_y(\hat{y}, \hat{u})$ is continuous. Notice that $E(y, u) = 0 \Leftrightarrow E(S(u), u) = 0$ holds for all $u \in \mathcal{U}_{\text{ad}}$. Invoking the implicit function theorem, we obtain that S is continuously differentiable in some $L^\infty(\Omega)$-neighborhood of $\hat{u}$. Since $\hat{u} \in \mathcal{U}_{\text{ad}}$ was arbitrary, S actually extends into an $L^\infty(\Omega)$-neighborhood of $\mathcal{U}_{\text{ad}}$ and it is continuously differentiable there. Moreover, we obtain that $\delta y = S'(\hat{u})\,\delta u$ satisfies $E_y(\hat{y}, \hat{u})\,\delta y = -E_u(\hat{y}, \hat{u})\,\delta u$, i. e.,

$$-\Delta\delta y + \frac{\big(\delta u + 2\,b\,(\nabla\hat{y}, \nabla\delta y)\big)\, f}{(\hat{u} + b\,\|\nabla\hat{y}\|^2)^2} = 0.$$

□

3.2.2 First-Order Optimality Conditions

The optimality system can be derived by using the Lagrangian $\mathcal{L}\colon H_0^1(\Omega) \times \mathcal{U}_{\mathrm{ad}} \times H_0^1(\Omega) \to \mathbb{R}$, defined by

$$\mathcal{L}(y,u,p) \coloneqq \int_\Omega \varphi(x,y)\,\mathrm{d}x + \frac{\lambda}{2}\,\|u\|_{H^1(\Omega)}^2 + \int_\Omega \nabla y \cdot \nabla p\,\mathrm{d}x - \int_\Omega \frac{f}{u + b\,\|\nabla y\|^2}\,p\,\mathrm{d}x \tag{3.2.6}$$

and taking the derivative with respect to the state and the control. In the first case, we obtain

$$\mathcal{L}_y(y,u,p)\,\delta y = \int_\Omega \varphi_y(x,y)\,\delta y\,\mathrm{d}x + \int_\Omega \nabla \delta y \cdot \nabla p\,\mathrm{d}x + \int_\Omega \frac{2\,b\,f\,p\,(\nabla y, \nabla \delta y)}{(u + b\,\|\nabla y\|^2)^2}\,\mathrm{d}x$$

for $\delta y \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$. Integration by parts yields

$$\begin{aligned}\mathcal{L}_y(y,u,p)\,\delta y &= \int_\Omega \varphi_y(x,y)\,\delta y\,\mathrm{d}x + \int_\Omega \nabla \delta y \cdot \nabla p\,\mathrm{d}x + \Big(\nabla y \int_\Omega \frac{2\,b\,f\,p}{(u + b\,\|\nabla y\|^2)^2}\,\mathrm{d}x, \nabla \delta y\Big) \\ &= \int_\Omega \varphi_y(x,y)\,\delta y\,\mathrm{d}x - \int_\Omega \Delta p\,\delta y\,\mathrm{d}x - \Big(\Delta y \int_\Omega \frac{2\,b\,f\,p}{(u + b\,\|\nabla y\|^2)^2}\,\mathrm{d}x, \delta y\Big).\end{aligned}$$

Notice that $\mathcal{L}_y(y,u,p)\,\delta y = 0$ for all $\delta y \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$ represents the strong form of the adjoint equation, which reads

$$\begin{cases} -\Delta p - \Delta y \displaystyle\int_\Omega \frac{2\,b\,f\,p}{(u + b\,\|\nabla y\|^2)^2}\,\mathrm{d}x = -\varphi_y(x,y) & \text{in } \Omega, \\ \qquad\qquad p = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.2.7}$$

We point out that (3.2.7) is again a nonlocal equation. Given $u \in \mathcal{U}_{\mathrm{ad}}$ and $y \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$, (3.2.7) has a unique solution $p \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$. This can be shown either by direct arguments as in proposition 3.2.1, or by exploiting that the bounded invertibility of E_y implies that of its adjoint, see the proof of theorem 3.2.2.

The derivative of the Lagrangian with respect to the control is given by

$$\mathcal{L}_u(y,u,p)\,\delta u = \lambda\,(u, \delta u)_{H^1(\Omega)} + \int_\Omega \frac{f\,p}{(u + b\,\|\nabla y\|^2)^2}\,\delta u\,\mathrm{d}x$$

for $\delta u \in H^1(\Omega)$.

It is now standard to derive the following system of necessary optimality conditions.

Theorem 3.2.3. *Suppose that $(y,u) \in \big(H_0^1(\Omega) \cap W^{2,q}(\Omega)\big) \times \mathcal{U}_{\mathrm{ad}}$ is a locally optimal solution of problem (3.1.1) for any $q \in [1,\infty)$. Then there exists a unique adjoint state $p \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$ for all $q \in [1,\infty)$ such that the following system holds:*

$$\begin{cases} -\Delta p - \Delta y \displaystyle\int_\Omega \frac{2\,b\,f\,p}{(u + b\,\|\nabla y\|^2)^2}\,\mathrm{d}x = -\varphi_y(x,y) & \text{in } \Omega, \\ \qquad\qquad p = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.2.8a}$$

$$\begin{cases} \lambda \displaystyle\int_\Omega \nabla u \cdot \nabla(v-u)\,\mathrm{d}x + \int_\Omega \Big(\frac{f\,p}{(u + b\,\|\nabla y\|^2)^2} + \lambda\,u\Big)\,(v-u)\,\mathrm{d}x \geq 0 \\ \quad \text{for all } v \in \mathcal{U}_{\mathrm{ad}}, \end{cases} \tag{3.2.8b}$$

$$\begin{cases} -\Delta y = \dfrac{f}{u + b\,\|\nabla y\|^2} & \text{in } \Omega, \\ \quad y = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.2.8c}$$

3.2.3 Analytical Solution

At this point we intend to construct some analytical solution for our optimal control problem:

Example 3.2.4. *We consider the following optimal control problem*

$$\text{Minimize } J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{H^1(\Omega)}$$
$$\text{subject to } \begin{cases} -\Delta y = \dfrac{f}{u + b\|\nabla y\|^2} & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases}$$

Let the problem domain be $\Omega = (0,1)^2$. *We choose the desired state as* $y_d = \sin(\pi x_1)\sin(\pi x_2)$ *and set* $y = y_d$. *First, we compute the gradient of* y *and obtain*

$$\nabla y = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \pi\cos(\pi x_1)\sin(\pi x_2) \\ \pi\sin(\pi x_1)\cos(\pi x_2) \end{bmatrix},$$

which results in

$$\begin{aligned}\|\nabla y\|^2_{L^2(\Omega)} &= \int_0^1\int_0^1 \left(\pi\cos(\pi x_1)\sin(\pi x_2)\right)^2 + \left(\pi\sin(\pi x_1)\cos(\pi x_2)\right)^2 \mathrm{d}x_1\,\mathrm{d}x_2 \\ &= \int_0^1\int_0^1 \pi^2\cos^2(\pi x_1)\sin^2(\pi x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 + \int_0^1\int_0^1 \pi^2\sin^2(\pi x_1)\cos^2(\pi x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.\end{aligned}$$

Applying Fubini's Theorem, we obtain

$$\|\nabla y\|^2_{L^2(\Omega)} = 2\pi^2\int_0^1\int_0^1 \cos^2(\pi x_1)\sin^2(\pi x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

Computing the inner integral, we obtain

$$\begin{aligned}\int_0^1 \cos^2(\pi x_1)\sin^2(\pi x_2)\,\mathrm{d}x_1 &= \sin^2(\pi x_2)\int_0^1 \cos^2(\pi x_1)\,\mathrm{d}x_1 = \sin^2(\pi x_2)\int_0^1 \frac{1+\cos(2\pi x_1)}{2}\,\mathrm{d}x_1 \\ &= \sin^2(\pi x_2)\left(\frac{1}{2}x_1 + \frac{1}{4}\sin(2\pi x_1)\right)\Bigg]_0^1 = \frac{1}{2}\sin^2(\pi x_2)\end{aligned}$$

and with that

$$\begin{aligned}\|\nabla y\|^2_{L^2(\Omega)} &= 2\pi^2\int_0^1 \frac{1}{2}\sin^2(\pi x_2)\,\mathrm{d}x_2 = \pi^2\int_0^1 \frac{1-\cos(2\pi x_2)}{2}\,\mathrm{d}x_2 \\ &= \pi^2\left(\frac{1}{2}x_2 - \frac{1}{4}\sin(2\pi x_2)\right)\Bigg]_0^1 = \frac{\pi^2}{2}.\end{aligned}$$

Next, we compute the Laplacian of y. *The Hessian amounts to*

$$D^2 y = \begin{bmatrix} -\pi^2\sin(\pi x_1)\sin(\pi x_2) & \pi^2\cos(\pi x_1)\cos(\pi x_2) \\ \pi^2\cos(\pi x_1)\cos(\pi x_2) & -\pi^2\sin(\pi x_1)\sin(\pi x_2) \end{bmatrix}$$

and the Laplacian

$$\begin{aligned}\Delta y = \operatorname{trace}(D^2 y) &= -\pi^2\sin(\pi x_1)\sin(\pi x_2) - \pi^2\sin(\pi x_1)\sin(\pi x_2) \\ &= -2\pi^2\sin(\pi x_1)\sin(\pi x_2).\end{aligned}$$

Putting $u = u_a \equiv 1$, $u_b \equiv 2$ *and* $b \equiv \frac{2}{\pi^2}$ *and inserting all this in*

$$-\Delta y = \frac{f}{u + b\,\|\nabla y\|^2_{L^2(\Omega)}}$$

we obtain

$$2\pi^2 y = \frac{f}{1 + \frac{2}{\pi^2} \cdot \frac{\pi^2}{2}} = \frac{f}{2},$$

which results in

$$f = 4\pi^2 y = 4\pi^2 \sin(\pi x_1) \sin(\pi x_2).$$

We observe that $p = 0$ *as the unique adjoint sate together with* (y, u) *satisfying the first order necessary optimality conditions* (3.2.8) *for this example. Since* $u = u_a$*, this solution is obviously a local minimizer. Since* $p = 0$*, for all* $u \in \mathcal{U}_{\text{ad}}$ *we have*

$$j'(u)(u - \bar{u}) = J_u(\bar{y}, \bar{u})(u - \bar{u}) = \lambda(\underbrace{\bar{u}}_{\equiv 1}, \underbrace{u - \bar{u}}_{\geq 0})_{L^2(\Omega)} + \lambda(\underbrace{\nabla \bar{u}}_{=0}, \nabla(u - \bar{u}))_{L^2(\Omega)} \geq 0.$$

Notice that (4.2.5b) is a nonlinear obstacle problem for the control variable u originating from the bound constraints in $\mathcal{U}_{\text{ad}}$ and the presence of the H^1-control cost term in the objective. Until recently, the Newton differentiability of the associated solution map was not known. In order to apply a generalized Newton method, we therefore chose to relax and penalize the bound constraints via a quadratic penalty in the following section. This is also known as Moreau-Yosida regularization of the indicator function pertaining to $\mathcal{U}_{\text{ad}}$.

Recently, the authors in Christof, Wachsmuth, 2023 proved a Newton differentiability result for the solution map of unilateral obstacle problems, which also is applicable to the other obstacle-type variational inequalities. This approach offers an alternative route to solving (4.2.5) numerically. It would amount to introducing a fourth unknown satisfying $z = \frac{-f\,p}{(u + b\,\|\nabla y\|^2)^2}$ and replacing (4.2.5b) by $u = G(z)$, where G stands for the solution map of the obstacle problem

$$\lambda \int_\Omega \nabla u \cdot (\nabla v - u) + u\,(v - u) - z\,(v - u)\,\mathrm{d}x \geq 0 \quad \text{for all } v \in \mathcal{U}_{\text{ad}}.$$

We leave the details for future work.

3.2.4 Moreau-Yosida Penalty Approximation

The Moreau-Yosida penalty approximation of problem (3.1.1) consists of the following modifications.

(1) We remove the constraints $u_a \leq u \leq u_b$ from $\mathcal{U}_{\text{ad}}$ and work with controls in $H^1(\Omega)$ which do not necessarily belong to $L^\infty(\Omega)$.

(2) We add the penalty term $\frac{1}{2\varepsilon} \int_\Omega (u_a - u)^2_+ + (u - u_b)^2_+ \,\mathrm{d}x$ to the objective. Here $v_+ = \max\{0, v\}$ is the positive part function and $\varepsilon > 0$ is the penalty parameter.

(3) We replace the control-to-state relation $y = S(u)$ by

$$y = S\big(u_a/2 + \eta_\varepsilon(u - u_a/2)\big),$$

where η_ε is a family of monotone and convex C^3 approximations of the positive part function satisfying $\eta_\varepsilon(t) = t$ for $t > \varepsilon$, $\eta_\varepsilon(t) = 0$ for $t < -\varepsilon$ for some $0 < \varepsilon < u_0/2$ and $\eta'_\varepsilon \in [0, 1]$ everywhere.

Notice that modification (3) is required since the control-to-state map S is guaranteed to be defined only for positive controls; compare theorem 3.1.6. Therefore, we use $u_a/2 + \eta_\varepsilon(u - u_a/2) \geq u_a/2$ as an effective control. In addition, $u_a/2 + \eta_\varepsilon(u - u_a/2) = u$ holds for all $u \in \mathcal{U}_{\text{ad}}$, provided that ε is small enough.

Next, we give an example of such a cut-off function:

Example 3.2.5. *An example of such a function is* $\eta_\varepsilon(t) = \varepsilon\,\eta(\frac{t}{\varepsilon})$*, where*

$$\eta(t) = \begin{cases} 0 & \text{for } t \le -1, \\ 15\left(\frac{t^4}{12} + \frac{t^5}{10} + \frac{t^6}{30}\right) + \frac{1+t}{2} - \frac{1}{4} & \text{for } -1 < t < 0, \\ 15\left(\frac{t^4}{12} - \frac{t^5}{10} + \frac{t^6}{30}\right) + \frac{1+t}{2} - \frac{1}{4} & \text{for } 0 \le t < 1, \\ t & \text{for } t \ge 1. \end{cases}$$

To construct such cut-off function, we begin by considering the following function

$$\eta(t) = \begin{cases} 0 & \text{for } t < -1, \\ t & \text{for } t > 1, \end{cases}$$

satisfying the following properties

$$\eta \in C^3,\ \eta(-1) = 0,\ \eta(1) = 1,\ 0 \le \eta' \le 1,\ \eta'' \ge 0. \tag{3.2.9}$$

Defining the cut-off function η_ε *by* $\eta_\varepsilon(t) = \varepsilon\,\eta(\frac{t}{\varepsilon})$*, we obtain*

$$\eta_\varepsilon(t) = \begin{cases} 0 & \text{for } t < -\varepsilon, \\ t & \text{for } t > \varepsilon, \end{cases}$$

and it inherits the following properties

$$\eta_\varepsilon \in C^3,\ \eta_\varepsilon(-\varepsilon) = 0,\ \eta_\varepsilon(\varepsilon) = \varepsilon,\ 0 \le \eta_\varepsilon' \le 1,\ \eta_\varepsilon'' \ge 0,$$

by means of (3.2.9) *and* $\eta_\varepsilon'(t) = \varepsilon\,\eta'(\frac{t}{\varepsilon})$*. We have to find a function rule for* $\eta(t)$*, for* $-1 < t < 1$*. We set* $\phi(t) \coloneqq \eta'(t)$*. Taking* (3.2.9) *into consideration,* ϕ *has to fulfill* $\phi \in C^2$*,* $0 \le \phi \le 1$*, and* $\phi' \ge 0$*. Since* $\eta(t) = \int_{-1}^{t} \phi(s)\,\mathrm{d}s$*, we intend to have* $\int_{-1}^{1} \phi(s)\,\mathrm{d}s = 1$ *and* $\phi''(0) = 0$*. We can construct a function* ψ*, such that*

$$\psi(t) = \begin{cases} \frac{1}{2} & \text{for } t > 1, \\ 0 & \text{for } t \le 0, \end{cases}$$

such that, $\psi \in C^2$*,* $\psi' \ge 0$*,* $\psi''(0) = 0$ *and* $\psi(t) = \int_0^t \xi(s)\,\mathrm{d}s$*, for* $0 < t \le 1$*. That means,* ξ *have to satisfy*

$$\xi(t) = \begin{cases} 0 & \text{for } t > 1 \\ 0 & \text{for } t < 0, \end{cases}$$

with $\xi \in C^1$ *and* $\xi \ge 0$*. We intend to have* $\int_0^1 \xi(s)\,\mathrm{d}s = \frac{1}{2}$*. In this case, we can define* ϕ *as following*

$$\phi(t) = \begin{cases} \frac{1}{2} + \psi(t) & \text{for } t \ge 0, \\ \frac{1}{2} - \psi(-t) & \text{for } t < 0. \end{cases}$$

We note that

$$\begin{aligned} \int_{-1}^{1} \phi(t)\,\mathrm{d}t &= \int_{-1}^{0} \frac{1}{2} - \psi(-t)\,\mathrm{d}t + \int_{0}^{1} \frac{1}{2} + \psi(t)\,\mathrm{d}t \\ &= \int_{-1}^{1} \frac{1}{2}\,\mathrm{d}t + \int_{-1}^{0} \psi(t)\,\mathrm{d}t + \int_{0}^{1} \psi(t)\,\mathrm{d}t = 1. \end{aligned}$$

Taking $\xi(t) = \alpha t^2\,(1-t)^2$*, we obtain*

$$\psi(t) = \int_0^t \xi(s)\,\mathrm{d}s = \alpha\left(\frac{1}{3}\,t^3 - \frac{1}{2}\,t^4 + \frac{1}{5}\,t^5\right).$$

Consequently,

$$\phi(t) = \begin{cases} \frac{1}{2} + \alpha\left(\frac{1}{3}\,t^3 - \frac{1}{2}\,t^4 + \frac{1}{5}\,t^5\right) & \text{for } t \geq 0, \\ \frac{1}{2} + \alpha\left(\frac{1}{3}\,t^3 + \frac{1}{2}\,t^4 + \frac{1}{5}\,t^5\right) & \text{for } t < 0. \end{cases}$$

Now, $\eta(t) = \int_{-1}^{t} \phi(s)\,\mathrm{d}s$ *results in*

$$\eta(t) = \int_{-1}^{t} \phi(s)\,\mathrm{d}s = \frac{1}{2}\,(t+1) - \frac{\alpha}{60} + \alpha\left(\frac{1}{12}\,t^4 - \frac{1}{10}\,t^5 + \frac{1}{30}\,t^6\right) \quad \text{for } t < 0,$$

and

$$\eta(t) = \int_{-1}^{0} \phi(s)\,\mathrm{d}s + \int_{0}^{t} \phi(s)\,\mathrm{d}s = \frac{1}{2} - \frac{\alpha}{60} + \frac{t}{2} + \alpha\left(\frac{1}{12}\,t^4 - \frac{1}{10}\,t^5 + \frac{1}{30}\,t^6\right) \quad \text{for } t \geq 0,$$

and eventually,

$$\eta(t) = \begin{cases} 0 & \text{for } t \leq -1, \\ \alpha\left(\frac{t^4}{12} + \frac{t^5}{10} + \frac{t^6}{30}\right) + \frac{1+t}{2} - \frac{\alpha}{60} & \text{for } -1 < t < 0, \\ \alpha\left(\frac{t^4}{12} - \frac{t^5}{10} + \frac{t^6}{30}\right) + \frac{1+t}{2} - \frac{\alpha}{60} & \text{for } 0 \leq t < 1, \\ t & \text{for } t \geq 1. \end{cases}$$

From $\int_0^1 \xi(s)\,\mathrm{d}s = \frac{1}{2}$, *follows* $\alpha = 15$ *and this yields* (3.2.5).

We now consider the following relaxed problem:

$$\begin{aligned} \text{Minimize} \quad & J_\varepsilon(y,u) := J(y,u) + \frac{1}{2\varepsilon}\int_\Omega (u_a - u)_+^2 + (u - u_b)_+^2\,\mathrm{d}x \\ \text{where} \quad & y = S\left(u_a/2 + \eta_\varepsilon(u - u_a/2)\right) \\ \text{and} \quad & u \in H^1(\Omega). \end{aligned} \tag{P_ε}$$

The relation between (P_ε) and the original problem (3.1.1) is clarified in the following theorem.

Theorem 3.2.6.

(*i*) *For all* $\varepsilon > 0$, *problem* (P_ε) *possesses a globally optimal solution* $(\bar{y}_\varepsilon, \bar{u}_\varepsilon) \in \left(H_0^1(\Omega) \cap W^{2,q}(\Omega)\right) \times H^1(\Omega)$ *for all* $q \in [1,\infty)$.

(*ii*) *For any sequence* $\varepsilon_n \searrow 0$, *there is a subsequence of* $(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n})$ *which converges weakly to some* (y^*, u^*) *in* $W^{2,q}(\Omega) \times H^1(\Omega)$. *Moreover,* $u^* \in \mathcal{U}_{\mathrm{ad}}$ *holds and* (y^*, u^*) *is a globally optimal solution of* (3.1.1).

Proof. Statement (*i*): The proof of statement (*i*) is divided into several steps.

Step (1): Since

$$y = S\left(u_a/2 + \eta_\varepsilon(u - u_a/2)\right)$$

we have

$$\|y\|_{W^{2,q(\Omega)}} = \|S\left(u_a/2 + \eta_\varepsilon(u - u_a/2)\right)\|_{W^{2,q}(\Omega)} \leq c_1 \left\| \frac{f}{u_a/2 + \eta_\varepsilon(u - u_a/2)} \right\|_{L^q(\Omega)}$$

$$\leq c_2 \left\| \frac{f}{u_a/2 + \eta_\varepsilon(u - u_a/2)} \right\|_{L^\infty(\Omega)} \leq c \left\| \frac{f}{u_a/2} \right\|_{L^\infty(\Omega)} \leq C$$

that means $y = S\left(u_a/2 + \eta_\varepsilon(u - u_a/2)\right)$ is bounded in $W^{2,q}(\Omega)$. Due to the embedding $W^{2,q}(\Omega) \hookrightarrow C(\overline{\Omega})$ for $q > N/2$, there exists $M > 0$ such that $\|y\|_{L^\infty(\Omega)} \leq M$ holds. From Assumption 3.1.2 we can obtain the estimate

$$|\varphi(x, y(x))| = |\varphi(x, y(x)) - \varphi(x,0) + \varphi(x,0)|$$

$$\begin{aligned}&\leq |\varphi(x,0)| + |\varphi(x,y(x)) - \varphi(x,0)|\\ &\leq K + L(M)\,|y(x)| \leq K + L(M)\,M.\end{aligned}$$

This implies

$$\int_\Omega \Phi(S\,(u_a/2 + \eta_\varepsilon(u - u_a/2)))\,\mathrm{d}x \geq -\big(K + L(M)\,M\big)\,|\Omega| \tag{3.2.10}$$

for all $u \in H^1(\Omega)$. Taking also $\|u\|^2_{H^1_0(\Omega)} \geq 0$ and $\frac{1}{2\varepsilon}\int_\Omega (u_a - u)^2_+ + (u - u_b)^2_+\,\mathrm{d}x \geq 0$ into consideration, we can infer that the penalized problem

$$J_\varepsilon(y,u) \coloneqq J(y,u) + \frac{1}{2\varepsilon}\int_\Omega (u_a - u)^2_+ + (u - u_b)^2_+\,\mathrm{d}x$$

is also bounded from below.

Step (2): We construct the tentative minimizer $(\bar y, \bar u)$. Since J_ε is bounded from below, there exists a minimizing sequence $\{(y_n, u_n)\}$, $u_n \in H^1(\Omega)$ so that

$$J_\varepsilon(y_n, u_n) \searrow \inf_{u \in H^1(\Omega)} J_\varepsilon(y,u) =: \beta.$$

The boundedness of $\{u_n\}$ in $H^1(\Omega)$ follows from the radial unboundedness of J_ε. Consequently, there exists a subsequence, denoted by the same indices, such that $u_n \rightharpoonup \bar u$ in $H^1(\Omega)$. As we observed $\{y_n\}$ is bounded in $W^{2,q}(\Omega)$. Since $W^{2,q}(\Omega)$ is a reflexive Banach space and every bounded subset of a reflexive Banach space is weakly relatively compact, there exists a subsequence y_n, denoted by the same indices, satisfying $y_n \rightharpoonup \bar y$ in $W^{2,q}(\Omega)$.
We now aim to show that $\bar y$ is the weak solution associated with $\bar u$. Form $u_n \rightharpoonup \bar u$ in $H^1(\Omega)$ follows $u_n \to \bar u$ in $L^2(\Omega)$ and consequently $u_a/2 + \eta_\varepsilon(u_n - u_a/2) \to u_a/2 + \eta_\varepsilon(\bar u - u_a/2)$. Since $u_a/2 + \eta_\varepsilon(u_n - u_a/2) \geq u_a/2$, the sequence $\{u_a/2 + \eta_\varepsilon(u_n - u_a/2)\}$ belongs to $\mathcal{U}_{\mathrm{ad}}$, now theorem 3.1.7 implies

$$y_n = S(u_a/2 + \eta_\varepsilon(u_n - u_a/2)) \to S(u_a/2 + \eta_\varepsilon(\bar u - u_a/2)).$$

From the uniqueness of the weak limit, we can infer that

$$\bar y = S(u_a/2 + \eta_\varepsilon(\bar u - u_a/2)).$$

Step (3): It remains to show the global optimality of $\bar u$. Set $F(y) \coloneqq \int_\Omega \varphi(x, y(x))\,\mathrm{d}x$, thus F is composed of a Nemytskii operator $\Phi(y) = \varphi(x, y(x))$ and a continuous linear integral operator from $L^1(\Omega)$ into $\mathbb{R}$. By virtue of lemma 3.1.3, Φ is continuous in $L^\infty(\Omega)$. Since $W^{2,q}(\Omega) \hookrightarrow L^\infty(\Omega)$ holds, F is continuous in $H^1(\Omega)$. Therefore, $F \circ S = \int_\Omega \Phi \circ S\,\mathrm{d}x$ is weakly-strongly continuous on $\mathcal{U}_{\mathrm{ad}}$.
In summary, exploiting the weak sequential lower semicontinuity of $\|\cdot\|_{H^1}$ we have

$$\begin{aligned}\beta &= \lim_{n\to\infty} J_\varepsilon(y_n, u_n)\\ &= \lim_{n\to\infty} J(y_n,u_n) + \frac{1}{2\varepsilon}\liminf_{n\to\infty}\Big(\|(u_a - u)_+\|^2_{L^2(\Omega)} + \|(u - u_b)_+\|^2_{L^2(\Omega)}\Big)\\ &= \lim_{n\to\infty} F(y_n) + \frac{\lambda}{2}\liminf_{n\to\infty}\|u_n\|^2_{H^1} + \frac{1}{2\varepsilon}\liminf_{n\to\infty}\Big(\|(u_a - u)_+\|^2_{L^2(\Omega)} + \|(u - u_b)_+\|^2_{L^2(\Omega)}\Big)\\ &\geq F(\bar y) + \frac{\lambda}{2}\|\bar u\|^2_{H^1(\Omega)} + \frac{1}{2\varepsilon}\Big(\|(u_a - \bar u)_+\|^2_{L^2(\Omega)} + \|(\bar u - u_b)_+\|^2_{L^2(\Omega)}\Big)\\ &= J(\bar y, \bar u).\end{aligned}$$

By definition of β and since $\bar u \in \mathcal{U}_{\mathrm{ad}} \cap H^1(\Omega)$, we therefore must have $\beta = j(\bar u)$.

Statement (*ii*): The proof of statement (*ii*) is also divided into several steps. As in the proof of theorem 3.1.9, we define β to be the globally optimal value of the objective in (P_ε). Similarly, we let β_ε denote the globally optimal value of the objective in (P_ε). Suppose that $\varepsilon_n \searrow 0$ is any sequence.

Step (1): We show that $\{(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n})\}$ is bounded in $W^{2,q}(\Omega) \times H^1(\Omega)$.
Suppose that $(\bar{y}, \bar{u})$ is a globally optimal solution of (3.1.1). Owing to the definition of β_ε, we have

$$\beta_\varepsilon \leq J_\varepsilon(\bar{y}, \bar{u}) = J(\bar{y}, \bar{u}) + \frac{1}{2\varepsilon} \int_\Omega (u_a - \bar{u})_+^2 + (\bar{u} - u_b)_+^2 \, \mathrm{d}x = J(\bar{y}, \bar{u}) = \beta. \tag{$*$}$$

The next-to-last equality is true since $\bar{u} \in \mathcal{U}_{\mathrm{ad}}$ holds and therefore, the penalty term vanishes. Moreover, we obtain

$$J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) \leq J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) + \frac{1}{2\varepsilon_n} \int_\Omega (u_a - \bar{u}_{\varepsilon_n})_+^2 + (\bar{u}_{\varepsilon_n} - u_b)_+^2 \, \mathrm{d}x = \beta_{\varepsilon_n} \leq \beta,$$

where the last inequality follows from ($*$). Since

$$\bar{y}_{\varepsilon_n} = S\left(u_a/2 + \eta_{\varepsilon_n}(\bar{u}_{\varepsilon_n} - u_a/2)\right)$$

holds, we obtain $\|\bar{y}_{\varepsilon_n}\|_{W^{2,q}(\Omega)} \leq C$ as in the proof of theorem 3.1.7. Therefore, $\bar{y}_{\varepsilon_n}$ is also bounded in $C(\overline{\Omega})$ and consequently, $\int_\Omega \varphi(x, \bar{y}_{\varepsilon_n}) \, \mathrm{d}x$ is bounded below, see (3.1.7). Finally,

$$J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) = \int_\Omega \varphi(x, \bar{y}_{\varepsilon_n}) \, \mathrm{d}x + \frac{\lambda}{2} \|\bar{u}_{\varepsilon_n}\|_{H^1(\Omega)}^2 \leq \beta$$

implies that $\|\bar{u}_{\varepsilon_n}\|_{H^1(\Omega)}$ is bounded.

Step (2): From Step (1) it follows that there exists a subsequence $\{(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n})\}$, denoted with the same subscript, such that $(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) \rightharpoonup (y^*, u^*)$ in $W^{2,q}(\Omega) \times H^1(\Omega)$. We show that $u^* \in \mathcal{U}_{\mathrm{ad}}$ holds.
We have already shown that $\beta_{\varepsilon_n} \leq \beta$ holds, therefore

$$\int_\Omega (u_a - \bar{u}_{\varepsilon_n})_+^2 + (\bar{u}_{\varepsilon_n} - u_b)_+^2 \, \mathrm{d}x \leq 2\varepsilon_n \left[\beta - J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n})\right].$$

Taking the lim sup in this inequality as $n \to \infty$, we find

$$0 \leq \limsup_{n\to\infty} \int_\Omega (u_a - \bar{u}_{\varepsilon_n})_+^2 + (\bar{u}_{\varepsilon_n} - u_b)_+^2 \, \mathrm{d}x \leq 0 - 2 \liminf_{n\to\infty} \varepsilon_n \, J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}). \tag{$**$}$$

From $(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) \rightharpoonup (y^*, u^*)$ in $W^{2,q}(\Omega) \times H^1(\Omega)$ we conclude $\bar{u}_{\varepsilon_n} \to u^*$ in $L^2(\Omega)$ and

$$J(y^*, u^*) \leq \liminf_{n\to\infty} J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) \tag{$***$}$$

as in the proof of theorem 3.1.9. Passing with $n \to \infty$ in ($**$) yields

$$\int_\Omega (u_a - u^*)_+^2 + (u^* - u_b)_+^2 \, \mathrm{d}x = 0$$

and consequently, $u^* \in \mathcal{U}_{\mathrm{ad}}$ follows.

Step (3): To obtain the convergence $\eta_{\varepsilon_n}(\bar{u}_{\varepsilon_n} - u_a/2) \to u^* - u_a/2$ in $L^2(\Omega)$, it suffices to note that the assumptions on η_ε imply that, for all $t \in \mathbb{R}$, $\eta_{\varepsilon_n}(t) \to \max\{0, t\}$ holds as $n \to \infty$ and that η_{ε_n} has a Lipschitz constant of 1 for all n. In combination

with $u^* \geq u_a$, the triangle inequality, and the dominated convergence theorem, this gives

$$\begin{aligned}
&\|\eta_{\varepsilon_n}(\bar{u}_{\varepsilon_n} - u_a/2) - (u^* - u_a/2)\|_{L^2(\Omega)} \\
&\quad \leq \|\eta_{\varepsilon_n}(\bar{u}_{\varepsilon_n} - u_a/2) - \eta_{\varepsilon_n}(u^* - u_a/2)\|_{L^2(\Omega)} \\
&\qquad + \|\eta_{\varepsilon_n}(u^* - u_a/2) - (u^* - u_a/2)\|_{L^2(\Omega)} \\
&\quad \leq \|\bar{u}_{\varepsilon_n} - u^*\|_{L^2(\Omega)} \\
&\qquad + \|\eta_{\varepsilon_n}(u^* - u_a/2) - \max\{0, u^* - u_a/2\}\|_{L^2(\Omega)} \to 0
\end{aligned}$$

as desired. The continuity of S on $\mathcal{U}_{\mathrm{ad}}$ w.r.t. the $L^2(\Omega)$-topology now implies

$$\bar{y}_{\varepsilon_n} = S\left(u_a/2 + \eta_{\varepsilon_n}\left(\bar{u}_{\varepsilon_n} - u_a/2\right)\right) \to S\left(u^*\right).$$

From Step (2) we have the weak convergence of $\bar{y}_{\varepsilon_n}$ to y^*. The uniqueness of the weak limit shows $y^* = S(u^*)$.

Step (4): Since $J(\bar{y}_{\varepsilon_n}, \bar{u}_{\varepsilon_n}) \leq \beta$ holds, we obtain $J(y^*, u^*) \leq \beta$ by invoking (***). Moreover, since (y^*, u^*) is admissible for (3.1.1), the definition of β implies $J(y^*, u^*) = \beta$, which completes the proof. □

3.2.5 First-Order Optimality Conditions for the Penalized Problem

The derivation of optimality conditions for (P_ε) proceeds along the same lines as in subsection 3.2.2 and the details are omitted. Notice that the use of the cut-off function in the control-to-state map resolves the difficulty with differentiability of this map with respect to $H^1(\Omega)$-topology in appropriate function spaces. For simplicity, we drop the index $\cdot_\varepsilon$ from now on and denote states, controls, and associated adjoint states by (y, u, p).

The optimality system can be derived by using the Lagrangian $\mathcal{L}\colon H_0^1(\Omega) \times H^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$, defined by

$$\begin{aligned}
\mathcal{L}(y, u, p) \coloneqq &\int_\Omega \varphi(x, y)\,\mathrm{d}x + \frac{\lambda}{2}\|u\|_{H^1(\Omega)}^2 + \frac{1}{2\varepsilon}\int_\Omega (u_a - u)_+^2 + (u - u_b)_+^2\,\mathrm{d}x \\
&+ \int_\Omega \nabla y \cdot \nabla p\,\mathrm{d}x - \int_\Omega \frac{f}{u_a/2 + \eta_\varepsilon(u - u_a/2) + b\|\nabla y\|^2}\,p\,\mathrm{d}x
\end{aligned} \tag{3.2.11}$$

and taking the derivative with respect to the state and the control. In the first case, we obtain

$$\mathcal{L}_y(y, u, p)\,\delta y = \int_\Omega \varphi_y(x, y)\,\delta y\,\mathrm{d}x + \int_\Omega \nabla \delta y \cdot \nabla p\,\mathrm{d}x + \int_\Omega \frac{2\,b\,f\,p\,(\nabla y, \nabla \delta y)}{\left(u_a/2 + \eta_\varepsilon(u - u_a/2)\right)^2}\,\mathrm{d}x$$

for $\delta y \in H_0^1(\Omega) \cap W^{2,q}(\Omega)$. Integration by parts yields

$$\begin{aligned}
\mathcal{L}_y(y, u, p)\,\delta y &= \int_\Omega \varphi_y(x, y)\,\delta y\,\mathrm{d}x + \int_\Omega \nabla \delta y \cdot \nabla p\,\mathrm{d}x \\
&\quad + \left(\nabla y \int_\Omega \frac{2\,b\,f\,p}{\left(u_a/2 + \eta_\varepsilon(u - u_a/2) + b\|\nabla y\|^2\right)^2}\,\mathrm{d}x, \nabla \delta y\right) \\
&= \int_\Omega \varphi_y(x, y)\,\delta y\,\mathrm{d}x - \int_\Omega \Delta p\,\delta y\,\mathrm{d}x \\
&\quad - \left(\Delta y \int_\Omega \frac{2\,b\,f\,p}{\left(u_a/2 + \eta_\varepsilon(u - u_a/2) + b\|\nabla y\|^2\right)^2}\,\mathrm{d}x, \delta y\right).
\end{aligned}$$

Notice that $\mathcal{L}_y(y,u,p)\,\delta y = 0$ for all $\delta y \in H_0^1(\Omega)\cap W^{2,q}(\Omega)$ represents the strong form of the adjoint equation, which reads

$$\begin{cases} -\Delta p - \Delta y \displaystyle\int_\Omega \frac{2\,b\,f\,p}{\left(u_a/2+\eta_\varepsilon(u-u_a/2)+b\,\|\nabla y\|^2\right)^2}\,\mathrm{d}x = -\varphi_y(x,y) & \text{in } \Omega,\\ p = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.2.12}$$

We point out that (3.2.12) is again a nonlocal equation. Given $u \in H^1(\Omega)$ and $y \in H_0^1(\Omega)\cap W^{2,q}(\Omega)$, (3.2.12) has a unique solution $p \in H_0^1(\Omega)\cap W^{2,q}(\Omega)$. This can be shown either by direct arguments as in proposition 3.2.1, or by exploiting that the bounded invertibility of E_y implies that of its adjoint, see the proof of theorem 3.2.2.

The derivative of the Lagrangian with respect to the control is given by

$$\begin{aligned}\mathcal{L}_u(y,u,p)\,\delta u = \lambda\,(u,\delta u)_{H^1(\Omega)} - \frac{1}{\varepsilon}\int_\Omega \left((u_a-u)_+ - (u-u_b)_+\right)\delta u\,\mathrm{d}x\\ + \int_\Omega \frac{f\,p\,\eta_\varepsilon'(u-u_a/2)}{\left(u_a/2+\eta_\varepsilon(u-u_a/2)+b\,\|\nabla y\|^2\right)^2}\,\delta u\,\mathrm{d}x\end{aligned}$$

for $\delta u \in H^1(\Omega)$.

We obtain the following regularized system of necessary optimality conditions.

Theorem 3.2.7. *Suppose that $(y,u) \in \left(H_0^1(\Omega)\cap W^{2,q}(\Omega)\right)\times H^1(\Omega)$ is a locally optimal solution of problem (P_ε) for any $q \in [1,\infty)$. Then there exists a unique adjoint state $p \in H_0^1(\Omega)\cap W^{2,q}(\Omega)$ for all $q \in [1,\infty)$ such that the following system holds:*

$$\begin{cases} -\Delta p - \Delta y \displaystyle\int_\Omega \frac{2\,b\,f\,p}{\left(u_a/2+\eta_\varepsilon\,(u-u_a/2)+b\,\|\nabla y\|^2\right)^2}\,\mathrm{d}x = -\varphi_y(x,y) & \text{in } \Omega,\\ p = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.2.13a}$$

$$\begin{cases} \lambda\displaystyle\int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}x + \int_\Omega \left(\frac{f\,p\,\eta_\varepsilon'(u-u_a/2)}{\left(u_a/2+\eta_\varepsilon\,(u-u_a/2)+b\,\|\nabla y\|^2\right)^2} + \lambda\,u\right)v\,\mathrm{d}x\\ \quad -\dfrac{1}{\varepsilon}\displaystyle\int_\Omega \left((u_a-u)_+ - (u-u_b)_+\right)v\,\mathrm{d}x = 0 \quad \text{for all } v \in H^1(\Omega), \end{cases} \tag{3.2.13b}$$

$$\begin{cases} -\Delta y = \dfrac{f}{u_a/2+\eta_\varepsilon\,(u-u_a/2)+b\,\|\nabla y\|^2} & \text{in } \Omega,\\ y = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.2.13c}$$

Remark 3.2.8. *We note that under a second-order sufficient condition, which is not investigated in this work, every solution of (3.2.13) is a strict local minimizer of (P_ε). According to theorem 3.2.6, applied to a modified problem with a suitable localization term, the local minimizer of the penalized problem under consideration converges to a local minimizer of the original optimal control problem as $\varepsilon \to 0$. This technique is well known; see for instance Casas, Mateos, Raymond, 2007, Section 4. Therefore, under second-order sufficient optimality conditions, the solutions of the optimality system of (P_ε) converge to the solutions of the optimality system of (3.1.1).*

Corollary 3.2.9. *The terms*

$$\left(\frac{f\,p\,\eta_\varepsilon'(u-u_a/2)}{\left(u_a/2+\eta_\varepsilon\,(u-u_a/2)+b\,\|\nabla y\|^2\right)^2} + \lambda\,u\right) - \frac{1}{\varepsilon}\left((u_a-u)_+ - (u-u_b)_+\right)$$

in (3.2.13b) belong to $L^\infty(\Omega)$ and therefore, any locally optimal control of (P_ε) belongs to $W^{2,q}(\Omega)$ for any $q \in [1,\infty)$.

Proof. We only elaborate on the case $N = 3$ since the cases $N \in \{1, 2\}$ are similar. We first consider the numerator of the first term. Here $f \in L^\infty(\Omega)$ holds by Assumption 4.1.1 and $p \in L^\infty(\Omega)$ by virtue of the embedding $W^{2,q}(\Omega) \hookrightarrow L^\infty(\Omega)$ for $q > 3/2$. Moreover, η_ε' maps into $[0, 1]$ and therefore $\eta_\varepsilon'(u - u_a)$ belongs to $L^\infty(\Omega)$ as well. The denominator is bounded below by $u_a/2$, and therefore, the first term belongs to $L^\infty(\Omega)$.
The second term, $\frac{1}{\varepsilon}\big((u_a - u)_+ - (u - u_b)_+\big)$, belongs to $L^6(\Omega)$ due to the embedding $H^1(\Omega) \hookrightarrow L^6(\Omega)$. Inserting this into (3.2.13b) with the differential operator $\lambda(-\Delta + \mathrm{id})$ and the remaining terms on the right-hand side shows $u \in W^{2,6}(\Omega)$, which in turn embeds into $L^\infty(\Omega)$. Repeating this procedure one more time implies $u \in W^{2,q}(\Omega)$. □

3.3 Generalized Newton Method

In this section we show that the optimality system (3.2.13) of the penalized problem is differentiable in a generalized sense, referred to as Newton differentiability. This allows us to formulate a generalized Newton method. Due to its similarity with the concept of semismoothness, see Ulbrich, 2011, such methods are sometimes referred to as a semismooth Newton method.

Definition 3.3.1 (Hintermüller, Ito, Kunisch, 2002, Definition 1, Ito, Kunisch, 2008, Definition 8.10). *Let X and Y be two Banach spaces and D be an open subset of X. The mapping $F\colon D \subset X \to Y$ is called Newton differentiable on the open subset $V \subset D$ if there exists a map $G\colon V \to \mathcal{L}(X, Y)$ such that, for every $x \in V$,*

$$\lim_{h \to 0} \frac{1}{\|h\|_X} \|F(x + h) - F(x) - G(x + h)h\|_Y = 0.$$

In this case G is said to be a Newton derivative of F on V.

We formulate the optimality system (3.2.13) in terms of an operator equation $F = 0$ where

$$F\colon X := \big(W^{2,q}(\Omega) \cap H_0^1(\Omega)\big) \times W^{2,q}_\diamond(\Omega) \times \big(W^{2,q}(\Omega) \cap H_0^1(\Omega)\big) \to L^q(\Omega)^3 =: Y \tag{3.3.1}$$

and $q \in [\max\{1, N/2\}, \infty)$ is arbitrary but fixed.

Here $W^{2,q}_\diamond(\Omega)$ is defined as

$$W^{2,q}_\diamond(\Omega) := \left\{ u \in W^{2,q}(\Omega) \,\middle|\, \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \right\},$$

which is directly motivated by the integration by parts of equation (3.2.13b). The component F_1 represents the adjoint equation (3.2.13a) in strong form, i. e.,

$$F_1(y, u, p) = -\Delta p - \Delta y \int_\Omega \frac{2\,b\,f\,p}{\big(u_a/2 + \eta_\varepsilon\,(u - u_a/2) + b\,\|\nabla y\|^2\big)^2}\,\mathrm{d}x + \varphi_y(x, y).$$

The continuous Fréchet differentiability of F_1 is a standard result, which uses lemma 3.1.3 and the embedding $W^{2,q}(\Omega) \hookrightarrow L^\infty(\Omega)$. The directional derivative is given by

$$\begin{aligned}
&F_1'(y, u, p)\,(\delta y, \delta u, \delta p) \\
&\quad = -\Delta\delta p - \Delta\delta y \int_\Omega \frac{2\,b\,f\,p}{\big(u_a/2 + \eta_\varepsilon\,(u - u_a/2) + b\,\|\nabla y\|^2\big)^2}\,\mathrm{d}x \\
&\qquad - \Delta y \int_\Omega \frac{2\,b\,f\,\delta p}{\big(u_a/2 + \eta_\varepsilon\,(u - u_a/2) + b\,\|\nabla y\|^2\big)^2}\,\mathrm{d}x \\
&\qquad + \Delta y \int_\Omega \frac{4\,b\,f\,p\,(\eta_\varepsilon'(u - u_a/2)\,\delta u + 2\,b\,(\nabla y, \nabla\delta y))}{\big(u_a/2 + \eta_\varepsilon\,(u - u_a/2) + b\,\|\nabla y\|^2\big)^3}\,\mathrm{d}x + \varphi_{yy}(x, y)\,\delta y.
\end{aligned}$$

Similarly, F_3 represents the state equation (3.2.13c), i. e.,

$$F_3(y,u,p) = -\Delta y - \frac{f}{u_a/2 + \eta_\varepsilon\,(u - u_a/2) + b\,\|\nabla y\|^2}$$

and its continuous Fréchet derivative is given by

$$F_3'(y,u,p)(\delta y,\delta u,\delta p) = -\Delta\delta y + \frac{f\,[\eta_\varepsilon'(u-u_a/2)\,\delta u + 2\,b\,(\nabla y,\nabla\delta y)]}{(u_a/2 + \eta_\varepsilon\,(u-u_a/2) + b\,\|\nabla y\|^2)^2}.$$

Finally, in order to define F_2 we integrate (3.2.13b) by parts, which is feasible due to corollary 3.2.9. This results in the equivalent formulation $F_2 = 0$, where

$$F_2(y,u,p) = -\lambda\Delta u + \frac{f\,p\,\eta_\varepsilon'(u-u_a/2)}{(u_a/2 + \eta_\varepsilon\,(u-u_a/2) + b\,\|\nabla y\|^2)^2} + \lambda\,u$$
$$- \frac{1}{\varepsilon}\left(\max\{u_a - u, 0\} - \max\{u - u_b, 0\}\right),$$

and the boundary conditions $\frac{\partial u}{\partial n} = 0$, which are included in the definition of $W^{2,q}_\diamond(\Omega)$.

In order to establish the Newton differentiability of F_2, we invoke the following classical result.

Theorem 3.3.2 (Hintermüller, Ito, Kunisch, 2002, Proposition 4.1, Ito, Kunisch, 2008, Example 8.14). *The mapping*

$$\max\{0,\cdot\}\colon L^p(\Omega) \to L^q(\Omega), \quad 1 \le q < p \le \infty$$

is Newton differentiable on $L^p(\Omega)$ with generalized derivative

$$G_{\max}\colon L^p(\Omega) \to \mathcal{L}(L^p(\Omega), L^q(\Omega))$$

given by

$$G_{\max}(u)\,\delta u = \begin{cases} \delta u(x), & \text{where } u(x) \ge 0, \\ 0, & \text{where } u(x) < 0. \end{cases}$$

Using theorem 3.3.2 and the embedding $W^{2,q}(\Omega) \hookrightarrow L^\infty(\Omega)$, it follows that F_2 is Newton differentiable on the entire space X with generalized derivative

$$G_2(y,u,p)(\delta y,\delta u,\delta p)$$
$$= -\lambda\Delta\delta u + \frac{f\,\delta p\,\eta_\varepsilon'(u-u_a/2) + f\,p\,\eta_\varepsilon''(u-u_a/2)\,\delta u}{(u_a/2 + \eta_\varepsilon\,(u-u_a/2) + b\,\|\nabla y\|^2)^2} + \lambda\,\delta u$$
$$- \frac{2\,f\,p\,\eta_\varepsilon'(u-u_a/2)\,[\eta_\varepsilon'(u-u_a/2)\,\delta u + 2\,b\,(\nabla y,\nabla\delta y)]}{(u_a/2 + \eta_\varepsilon\,(u-u_a/2) + b\,\|\nabla y\|^2)^3} + \frac{1}{\varepsilon}\chi_{A(u)}\delta u.$$

Here χ_A stands for the indicator function of the set

$$A(u) = \{x \in \Omega \,|\, u_a - u \ge 0 \text{ or } u - u_b \ge 0\}.$$

We are now in a position to state a basic generalized Newton method; see Algorithm 3.3.3. Following well-known arguments, we can show its local well-posedness and superlinear convergence to local minimizers satisfying second-order sufficient conditions. We refrain from repeating the details and refer the interested reader to, e. g., Ito, Kunisch, 2008, Chapter 7, Hinze, Pinnau, et al., 2009, Chapter 2.4–2.5 and Ulbrich, 2011, Chapter 10. It is also possible to globalize the method using a line search approach; see, e. g., Hinze, Vierling, 2012.

Algorithm 3.3.3 (Basic semismooth Newton method for the solution of problem (P_ε)).

Input: *initial guess* $(y_0,u_0,p_0) \in X$

Output: *approximate stationary point of* (P_ε)

1: *Set* $k \coloneqq 0$
2: ***while*** *not converged* ***do***
3: *Determine the active set* $A(u_k)$
4: *Solve the Newton system*

$$\begin{aligned} G_1(y_k, u_k, p_k)(\delta y, \delta u, \delta p) &= -F_1(y_k, u_k, p_k) \\ G_2(y_k, u_k, p_k)(\delta y, \delta u, \delta p) &= -F_2(y_k, u_k, p_k) \\ G_3(y_k, u_k, p_k)(\delta y, \delta u, \delta p) &= -F_3(y_k, u_k, p_k) \end{aligned} \tag{3.3.2}$$

5: *Update the iterates by setting*

$$y_{k+1} \coloneqq y_k + \delta y, \quad u_{k+1} \coloneqq u_k + \delta u, \quad p_{k+1} \coloneqq p_k + \delta p$$

6: *Set* $k \coloneqq k + 1$
7: ***end while***

An appropriate criterion for the convergence of Algorithm 3.3.3 is the smallness of $\|F_1(y_k, u_k, p_k)\|_{L^q(\Omega)}$, $\|F_2(y_k, u_k, p_k)\|_{L^q(\Omega)}$ and $\|F_3(y_k, u_k, p_k)\|_{L^q(\Omega)}$, either in absolute terms or relative to the initial values.

Remark 3.3.4. *We remark that all previous results can be generalized to convex domains* $\Omega \subset \mathbb{R}^N$ *where* $1 \leq N \leq 3$. *In this case, we can invoke the* H^2*-regularity result for the Poisson problem on convex domains in the proof of theorem 3.1.6, which can be deduced from Grisvard, 1985, Theorem 3.2.1.2:*

Theorem 3.3.5. *Let* Ω *be a bounded and convex domain. If the right-hand side function* $f \in L^2(\Omega)$, *then the solution of the Poisson problem*

$$\begin{cases} -\Delta y = f & \text{in } \Omega, \\ \quad y = 0 & \text{on } \partial\Omega, \end{cases}$$

belongs to $H^2(\Omega)$.

Consequently, we have to replace $q \in [1, \infty)$ *by* $q = 2$ *in theorem 3.1.6 and all subsequent results. The requirement* $N \leq 3$ *ensures the validity of the embedding* $H^2(\Omega) \hookrightarrow C(\overline{\Omega})$.

3.4 Discretization and Implementation

In this section we address the discretization of the relaxed optimal control problem (P_ε). We then follow a discretize–then–optimize approach and derive the associated discrete optimality system, as well as a discrete version of the generalized Newton method. In order to simplify the implementation, we employ the original control-to-state map $y = S(u)$. In other words, we choose $\eta_\varepsilon = \mathrm{id}$ in (P_ε), which no longer approximates the positive part function. Consequently, the controls appearing in the control-to-state map are no longer guaranteed to be bounded below by u_a. This simplification is justified a posteriori, provided that the control iterates happen to remain positive and bounded away from zero and thus still permit the state equation to be uniquely solvable, or rather its linearized counterpart appearing in the generalized Newton method. We numerically observed this to be the case for all examples. In addition, we allow the addition of an upper bound on the constraint in our implementation, which is treated via the same penalty approach as the lower bound.

Our discretization method of choice is the finite element method. We employ piecewise linear, globally continuous finite elements on geometrically conforming triangulations of the domain Ω. More precisely, we use the space

$$V_h \coloneqq \{v \in H^1(\Omega) \cap C(\overline{\Omega}) \,|\, v \text{ is linear on all triangles}\} \subset H^1(\Omega)$$

to discretize the control, the state and adjoint state variables. We use the usual Lagrangian basis and refer to the basis functions as $\{\varphi_j\}$, where $j = 1, \ldots, N_V$ and N_V denotes the

number of vertices in the mesh. The coefficient vector, e. g., for the discrete control variable $u \in V_h$, will be denoted by $\mathbf{u}$, so we have

$$u = \sum_{j=1}^{N_V} \mathbf{u}_j \varphi_j.$$

In order to formulate the discrete optimal control problem, we introduce the mass and stiffness matrices $\mathbf{M}$ and $\mathbf{K}$ as follows:

$$\mathbf{M}_{ij} = \int_\Omega \varphi_i \, \varphi_j \, \mathrm{d}x \quad \text{and} \quad \mathbf{K}_{ij} = \int_\Omega \nabla\varphi_i \cdot \nabla\varphi_j \, \mathrm{d}x.$$

We also make use of the diagonally lumped mass matrix $\mathbf{M}_{\text{lumped}}$ with entries $\mathbf{M}_{\text{lumped}_{ii}} = \sum_{j=1}^{N_V} \mathbf{M}_{ij}$. Suppose that the right-hand side f and coefficient b have been discretized and represented by their coefficient vectors $\mathbf{f}$ and $\mathbf{b}$ in V_h. Using the lumped mass matrix, the weak formulation (3.1.4) of the state equation can be written in preliminary discrete form as

$$\mathbf{K}\,\mathbf{y} = \mathbf{M}_{\text{lumped}} \left[\frac{\mathbf{f}_i}{\mathbf{u}_i + \mathbf{b}_i(\mathbf{y}^{\mathrm{T}}\mathbf{K}\,\mathbf{y})} \right]_{i=1}^{N_V}.$$

In order to incorporate the Dirichlet boundary conditions, we introduce the boundary projector $\mathbf{P}_\Gamma$. This is a diagonal $N_V \times N_V$-matrix which has ones along the diagonal in entries pertaining to boundary vertices, and zeros otherwise. We also introduce the interior projector $\mathbf{P}_\Omega \coloneqq \mathrm{id} - \mathbf{P}_\Gamma$. We can thus state the discrete form of the state equation (3.1.4) as

$$\mathbf{P}_\Omega\mathbf{K}\,\mathbf{y} - \mathbf{P}_\Omega\mathbf{M}_{\text{lumped}} \left[\frac{\mathbf{f}_i}{\mathbf{u}_i + \mathbf{b}_i(\mathbf{y}^{\mathrm{T}}\mathbf{K}\,\mathbf{y})} \right]_{i=1}^{N_V} + \mathbf{P}_\Gamma\mathbf{y} = \mathbf{0}. \tag{3.4.1}$$

In order to simplify the notation, we introduce further diagonal matrices

$$\mathbf{F} \coloneqq \mathrm{diag}(\mathbf{f}), \quad \mathbf{B} \coloneqq \mathrm{diag}(\mathbf{b}) \quad \text{and} \quad \mathbf{D}(\mathbf{y},\mathbf{u}) \coloneqq \mathrm{diag}(\mathbf{u}) + (\mathbf{y}^{\mathrm{T}}\mathbf{K}\,\mathbf{y})\,\mathbf{B}.$$

Using these matrices, we can write (3.4.1) more compactly as

$$e(\mathbf{y},\mathbf{u}) \coloneqq \mathbf{P}_\Omega\mathbf{K}\,\mathbf{y} - \mathbf{P}_\Omega\mathbf{M}_{\text{lumped}}\,\mathbf{F}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}\mathbf{1} + \mathbf{P}_\Gamma\mathbf{y} = \mathbf{0}, \tag{3.4.2}$$

where $\mathbf{1}$ and $\mathbf{0}$ denote column vectors of all ones and all zeros, respectively.

To be specific, we focus on a tracking-type objective and choose $\varphi(x,y) = \frac{1}{2}(y - y_d)^2$ in (1.1.4) and thus also in (P_ε). In addition, we distinguish two positive control cost parameters λ_1 and λ_2, which leads to discrete problems of the form

$$\begin{aligned} J(\mathbf{y},\mathbf{u}) = {} & \frac{1}{2}(\mathbf{y}-\mathbf{y}_d)^{\mathrm{T}}\mathbf{M}(\mathbf{y}-\mathbf{y}_d) + \frac{\lambda_1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{K}\mathbf{u} + \frac{\lambda_2}{2}\mathbf{u}^{\mathrm{T}}\mathbf{M}\mathbf{u} \\ & + \frac{1}{2\varepsilon}(\mathbf{u}_a-\mathbf{u})_+^{\mathrm{T}}\mathbf{M}_{\text{lumped}}(\mathbf{u}_a-\mathbf{u})_+ + \frac{1}{2\varepsilon}(\mathbf{u}-\mathbf{u}_b)_+^{\mathrm{T}}\mathbf{M}_{\text{lumped}}(\mathbf{u}-\mathbf{u}_b)_+ \end{aligned} \tag{3.4.3}$$

and the Lagrangian of our discretized problem becomes

$$\begin{aligned} \mathcal{L}(\mathbf{y},\mathbf{u},\mathbf{p}) = {} & \frac{1}{2}(\mathbf{y}-\mathbf{y}_d)^{\mathrm{T}}\mathbf{M}(\mathbf{y}-\mathbf{y}_d) + \frac{\lambda_1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{K}\mathbf{u} + \frac{\lambda_2}{2}\mathbf{u}^{\mathrm{T}}\mathbf{M}\mathbf{u} \\ & + \frac{1}{2\varepsilon}(\mathbf{u}_a-\mathbf{u})_+^{\mathrm{T}}\mathbf{M}_{\text{lumped}}(\mathbf{u}_a-\mathbf{u})_+ + \frac{1}{2\varepsilon}(\mathbf{u}-\mathbf{u}_b)_+^{\mathrm{T}}\mathbf{M}_{\text{lumped}}(\mathbf{u}-\mathbf{u}_b)_+ \\ & + \mathbf{p}^{\mathrm{T}}\mathbf{P}_\Omega\mathbf{K}\,\mathbf{y} - \mathbf{p}^{\mathrm{T}}\mathbf{P}_\Omega\mathbf{M}_{\text{lumped}}\,\mathbf{F}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}\mathbf{1} + \mathbf{p}^{\mathrm{T}}\mathbf{P}_\Gamma\mathbf{y}. \end{aligned} \tag{3.4.4}$$

Before we state the first- and second-order derivatives of the Lagrangian, we address the nonlinear term $\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}$ first. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{y}}\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}\delta\mathbf{y} = -2\,(\mathbf{y}^{\mathrm{T}}\mathbf{K}\,\delta\mathbf{y})\,\mathbf{B}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-2} \text{ and thus } \frac{\mathrm{d}}{\mathrm{d}\mathbf{y}}\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}\mathbf{1} = -2\,\mathbf{B}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\mathbf{1}\,\mathbf{y}^{\mathrm{T}}\mathbf{K},$$
$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{u}}\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}\delta\mathbf{u} = -\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\,\mathrm{diag}(\delta\mathbf{u}) \qquad \text{and thus } \frac{\mathrm{d}}{\mathrm{d}\mathbf{u}}\mathbf{D}(\mathbf{y},\mathbf{u})^{-1}\mathbf{1} = -\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}.$$

Therefore, the first-order derivatives of $\mathcal{L}$ (written as column vectors) are given by

$$\mathcal{L}_{\mathbf{y}}(\mathbf{y},\mathbf{u},\mathbf{p}) = \mathbf{M}(\mathbf{y}-\mathbf{y}_d) + \mathbf{K}\,\mathbf{P}_\Omega\mathbf{p} + 2\,\mathbf{K}\,\mathbf{y}\,\mathbf{1}^{\mathrm{T}}\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\mathbf{B}\,\mathbf{F}\,\mathbf{M}_{\mathrm{lumped}}\,\mathbf{P}_\Omega\mathbf{p} + \mathbf{P}_\Gamma\mathbf{p}, \tag{3.4.5a}$$

and

$$\begin{aligned}\mathcal{L}_{\mathbf{u}}(\mathbf{y},\mathbf{u},\mathbf{p}) = \lambda_1\mathbf{K}\mathbf{u} + \lambda_2\mathbf{M}\mathbf{u} - \frac{1}{\varepsilon}\mathbf{D}_{A_-}(\mathbf{u})\,\mathbf{M}_{\mathrm{lumped}}\,\mathbf{D}_{A_-}(\mathbf{u})\,(\mathbf{u}_a-\mathbf{u})\\ + \frac{1}{\varepsilon}\mathbf{D}_{A_+}(\mathbf{u})\,\mathbf{M}_{\mathrm{lumped}}\,\mathbf{D}_{A_+}(\mathbf{u})\,(\mathbf{u}-\mathbf{u}_b) + \mathbf{F}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\mathbf{M}_{\mathrm{lumped}}\,\mathbf{P}_\Omega\mathbf{p}.\end{aligned} \tag{3.4.5b}$$

Here $\mathbf{D}_{A_+}(\mathbf{u})$ and $\mathbf{D}_{A_-}(\mathbf{u})$ are diagonal (active-set) matrices with entries

$$[\mathbf{D}_{A_+}(\mathbf{u})]_{ii} = \begin{cases}1 & \text{where } [\mathbf{u}_a-\mathbf{u}]_i \geq 0,\\ 0 & \text{otherwise,}\end{cases} \qquad [\mathbf{D}_{A_-}(\mathbf{u})]_{ii} = \begin{cases}1 & \text{where } [\mathbf{u}-\mathbf{u}_b]_i \geq 0,\\ 0 & \text{otherwise,}\end{cases}$$

and we set $\mathbf{D}_A(\mathbf{u}) = \mathbf{D}_{A_+}(\mathbf{u}) + \mathbf{D}_{A_-}(\mathbf{u})$.

In order to solve the discrete optimality system consisting of (3.4.1) and (3.4.5), we employ a finite-dimensional semismooth Newton method (Algorithm 3.4.1). This requires the evaluation of first-order derivatives of the state equation (3.4.1) as well as second-order derivatives of the Lagrangian (3.4.4). The following expressions are obtained.

$$e_{\mathbf{y}}(\mathbf{y},\mathbf{u}) = \mathbf{P}_\Omega\mathbf{K} + 2\,\mathbf{P}_\Omega\,\mathbf{M}_{\mathrm{lumped}}\,\mathbf{F}\,\mathbf{B}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\mathbf{1}\,\mathbf{y}^{\mathrm{T}}\,\mathbf{K} + \mathbf{P}_\Gamma, \tag{3.4.6a}$$

$$e_{\mathbf{u}}(\mathbf{y},\mathbf{u}) = \mathbf{P}_\Omega\mathbf{M}_{\mathrm{lumped}}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\mathbf{F}, \tag{3.4.6b}$$

$$\begin{aligned}\mathcal{L}_{\mathbf{yy}}(\mathbf{y},\mathbf{u},\mathbf{p}) = \mathbf{M} - 8\mathbf{p}^{\mathrm{T}}\mathbf{P}_\Omega\mathbf{M}_{\mathrm{lumped}}\,\mathbf{F}\,\mathbf{B}^2\mathbf{D}(\mathbf{y},\mathbf{u})^{-3}\mathbf{1}\,\mathbf{K}\,\mathbf{y}\,\mathbf{y}^{\mathrm{T}}\mathbf{K}\\ + 2\mathbf{p}^{\mathrm{T}}\mathbf{P}_\Omega\mathbf{M}_{\mathrm{lumped}}\,\mathbf{F}\,\mathbf{B}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-2}\mathbf{1}\,\mathbf{K},\end{aligned} \tag{3.4.6c}$$

$$\mathcal{L}_{\mathbf{yu}}(\mathbf{y},\mathbf{u},\mathbf{p}) = -4\,\mathbf{K}\,\mathbf{y}\,\mathbf{p}^{\mathrm{T}}\mathbf{P}_\Omega\mathbf{M}_{\mathrm{lumped}}\,\mathbf{F}\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-3}\mathbf{B}, \tag{3.4.6d}$$

$$\begin{aligned}\mathcal{L}_{\mathbf{uu}}(\mathbf{y},\mathbf{u},\mathbf{p}) = \lambda_1\mathbf{K} + \lambda_2\mathbf{M} + \frac{1}{\varepsilon}\mathbf{D}_A(\mathbf{u})\,\mathbf{M}_{\mathrm{lumped}}\,\mathbf{D}_A(\mathbf{u})\\ - 2\,\mathrm{diag}(\mathbf{M}_{\mathrm{lumped}}\,\mathbf{p})\,\mathbf{D}(\mathbf{y},\mathbf{u})^{-3}\mathbf{F}.\end{aligned} \tag{3.4.6e}$$

Notice that the expression for $\mathcal{L}_{\mathbf{uu}}$ is the generalized derivative of $\mathcal{L}_{\mathbf{u}}$ in the sense of definition 3.3.1.

The discrete generalized Newton system has the following form:

$$\begin{bmatrix}\mathcal{L}_{\mathbf{yy}}(\mathbf{y},\mathbf{u},\mathbf{p}) & \mathcal{L}_{\mathbf{yu}}(\mathbf{y},\mathbf{u},\mathbf{p}) & e_{\mathbf{y}}(\mathbf{y},\mathbf{u})^{\mathrm{T}}\\ \mathcal{L}_{\mathbf{uy}}(\mathbf{y},\mathbf{u},\mathbf{p}) & \mathcal{L}_{\mathbf{uu}}(\mathbf{y},\mathbf{u},\mathbf{p}) & e_{\mathbf{u}}(\mathbf{y},\mathbf{u})^{\mathrm{T}}\\ e_{\mathbf{y}}(\mathbf{y},\mathbf{u}) & e_{\mathbf{u}}(\mathbf{y},\mathbf{u}) & \mathbf{0}\end{bmatrix}\begin{pmatrix}\delta\mathbf{y}\\ \delta\mathbf{u}\\ \delta\mathbf{p}\end{pmatrix} = -\begin{pmatrix}\mathcal{L}_{\mathbf{y}}(\mathbf{y},\mathbf{u},\mathbf{p})\\ \mathcal{L}_{\mathbf{u}}(\mathbf{y},\mathbf{u},\mathbf{p})\\ e(\mathbf{y},\mathbf{u})\end{pmatrix}. \tag{3.4.7}$$

The well-posedness of the system (3.4.7) can be shown in a neighborhood of a locally optimal solution satisfying second-order sufficient optimality conditions, under the additional assumption that $\mathbf{u}$ remains positive. This is a well established technique and it applies both to the continuous as well as to the discrete setting; see for instance Alt, 1990; Tröltzsch, 1999; Rösch, Wachsmuth, 2008. In contrast to standard optimal control problems which do not feature a nonlocal PDE, some of the blocks in (3.4.7) are no longer sparse. This comment applies to $e_{\mathbf{y}}$ due to the second summand in (3.4.6a), to $\mathcal{L}_{\mathbf{yy}}$ due to the second

summand in (3.4.6c) as well as to $\mathcal{L}_{\mathbf{yu}}$ given by (3.4.6d). For a high performance implementation, it is therefore important to not assemble the blocks in (3.4.7) as matrices, but rather to provide matrix-vector products and use a preconditioned iterative solver such as MINRES (Paige, Saunders, 1975) to solve (3.4.7). We defer the design and analysis of a suitable preconditioner to future work. For the time being we resort to the direct solution of (3.4.7) using MATLAB's direct solver, which is still feasible on moderately fine discretizations of two-dimensional domains.

Our implementation of the semismooth Newton method is described in Algorithm 3.4.1. In contrast to Algorithm 3.3.3, we added an additional step in which we solve the discrete nonlinear state equation (3.4.2) for $\mathbf{y}_{k+1}$ once per iteration for increased robustness; see line 6 in Algorithm 3.4.1. Notice that the preliminary linear update to $\mathbf{y}_{k+1}$ in line 5 is still useful since it provides an initial guess for the subsequent solution of $e(\mathbf{y}_{k+1}, \mathbf{u}_{k+1}) = 0$. We mention that nonlinear state updates have been analyzed in the closely related context of SQP methods, e. g., in Ulbrich, 2007; Clever et al., 2011. We also added a rudimentary damping strategy which improves the convergence behavior. In our examples, it suffices to choose $\gamma = 1/2$ when $\|\mathcal{L}_{\mathbf{u}}(\mathbf{y}_k, \mathbf{u}_k, \mathbf{p}_k)\|_{(\mathbf{K}+\mathbf{M})^{-1}} > 1/10$ and $\gamma = 1$ otherwise.

The stopping criterion we employ in line 2 measures the three components of the residual, i. e., the right-hand side in (3.4.7). Since each component represents an element of the dual space of $H^1(\Omega)$, we evaluate the (squared) $H^1(\Omega)^*$-norm of all residual components, which amounts to

$$R^2(\mathbf{y}, \mathbf{u}, \mathbf{p}) \coloneqq \|\mathcal{L}_{\mathbf{y}}(\mathbf{y}, \mathbf{u}, \mathbf{p})\|^2_{(\mathbf{K}+\mathbf{M})^{-1}} + \|\mathcal{L}_{\mathbf{u}}(\mathbf{y}, \mathbf{u}, \mathbf{p})\|^2_{(\mathbf{K}+\mathbf{M})^{-1}} + \|e(\mathbf{y}, \mathbf{u})\|^2_{(\mathbf{K}+\mathbf{M})^{-1}}. \tag{3.4.8}$$

Algorithm 3.4.1 is stopped when

$$R(\mathbf{y}, \mathbf{u}, \mathbf{p}) \leq 10^{-6} \tag{3.4.9}$$

is reached. Moreover, we impose a tolerance of $\|e(\mathbf{y}, \mathbf{u})\|_{(\mathbf{K}+\mathbf{M})^{-1}} \leq 10^{-10}$ for the solution of the forward problem in line 6.

Algorithm 3.4.1 (Discrete semismooth Newton method with nonlinear state update for the solution of a discretized instance of problem (P_ε)).

Input: *initial guess* $(\mathbf{y}_0, \mathbf{u}_0, \mathbf{p}_0) \in V_h \times V_h \times V_h$
Output: *approximate stationary point of the discretized instance of* (P_ε)

1: *Set* $k \coloneqq 0$
2: ***while*** *not converged* ***do***
3: *Determine the active sets* $A_+(\mathbf{u}_k)$ *and* $A_-(\mathbf{u}_k)$
4: *Solve the Newton system* (3.4.7) *for* $(\delta\mathbf{y}, \delta\mathbf{u}, \delta\mathbf{p})$, *given* $(\mathbf{y}_k, \mathbf{u}_k, \mathbf{p}_k)$
5: *Update the iterates by setting*

$$\mathbf{y}_{k+1} \coloneqq \mathbf{y}_k + \gamma\,\delta\mathbf{y}, \quad \mathbf{u}_{k+1} \coloneqq \mathbf{u}_k + \gamma\,\delta\mathbf{u}, \quad \mathbf{p}_{k+1} \coloneqq \mathbf{p}_k + \gamma\,\delta\mathbf{p}$$

where $\gamma \in (0, 1]$ *is a suitable damping parameter.*
6: *Solve the nonlinear state equation* (3.4.2) *for the state* $\mathbf{y}_{k+1}$, *given the control* $\mathbf{u}_{k+1}$
7: *Set* $k \coloneqq k + 1$
8: ***end while***

3.5 Numerical Experiments

In the following, we describe a number of numerical experiments. The first experiment introduce in subsection 3.5.1, serves the purpose of demonstrating the influence of the nonlocality parameter b. In the second experiment subsection 3.5.2, we numerically confirm the mesh independence of our algorithm. The third experiment presented in subsection 3.5.3, is dedicated to studying the impact of the penalty parameter ε. The last example presented

in subsection 3.5.4 is devoted to studying the influence of the control cost parameters λ_1 and λ_2.

As mentioned in section 3.4, our implementation of Algorithm 3.4.1 employs a direct solver for the linear systems arising in line 4 and is therefore only suitable for relatively coarse discretization of two-dimensional domains. Unless otherwise mentioned, the following experiments are obtained on a mesh discretizing a square domain with $N_V = 665$ vertices and $N_T = 1248$ triangles. Notice that convex domains are covered by our theory due to remark 3.3.4. The typical run-time for Algorithm 3.4.1 is around 3 s.

3.5.1 Influence of the Non-Locality Parameter

Our initial example builds on the two-dimensional problem presented in Delgado, Figueiredo, et al., 2017. The problem domain is $\Omega = (-0.5, 0.5)^2$; notice that this is slightly incorrectly stated in Delgado, Figueiredo, et al., 2017. Moreover, we have right-hand side $f(x, y) \equiv 100$ and desired state $y_d(x, y) \equiv 0$. The lower bound for the control is given as $u_a(x, y) = -3x - 3y + 10$ and the upper bound is $u_b \equiv \infty$. Moreover, the control cost parameters are $\lambda_1 = 0$ and $\lambda_2 = 4 \cdot 10^{-5}$. We choose $\varepsilon = 10^{-2}$ as our penalty parameter. The coefficient function determining the degree of non-locality is set to $b(x, y) = \alpha\,(x^2 + y^2)$, where α varies in $\{0, 10^0, 10^1, 10^2, 10^3\}$. We point out that these settings violate Assumption 4.1.1 due to $\lambda_1 = 0$, i. e., the cost term is only of L^2-type, and since b is not uniformly positive inside Ω. The lack of an upper bound in this example is of no concern because we could assign a posteriori a sufficiently large upper bound which does not become active. Nonetheless, we present this experiment in order to reproduce the results in Delgado, Figueiredo, et al., 2017, which correspond to the case $\alpha = 1$.

For each value of α, we start from an initial guess constructed as follows. We initialize $\mathbf{u}_0$ to the lower bound $\mathbf{u}_a$ and set $\mathbf{y}_0$ to the numerical solution of the forward problem with control $\mathbf{u}_0$. The adjoint state is initialized to $\mathbf{p}_0 = \mathbf{0}$.

Figure 3.5.1 shows some of the optimal state and control functions obtained. We notice that the solution in case of a local problem ($\alpha = 0$) is visually indistinguishable from the setting $\alpha = 1$ considered in Delgado, Figueiredo, et al., 2017. We therefore compare it to the case $\alpha = 10^3$ of significantly more pronounced non-local effects. Clearly, an increase in the non-local parameter aids the control in this example, so the control effort can decrease, as reflected in figure 3.5.1. Also, we observe that the number of iterations of the discrete semismooth Newton method (Algorithm 3.4.1) decreases slightly as α increases; see table 3.5.1.

α	iterations
0.00e+00	10
1.00e+00	9
1.00e+01	7
1.00e+02	7
1.00e+03	6

Table 3.5.1. Number of iterations of the discrete semismooth Newton method (Algorithm 3.4.1) for various values of the non-locality parameter α in the example from subsection 3.5.1.

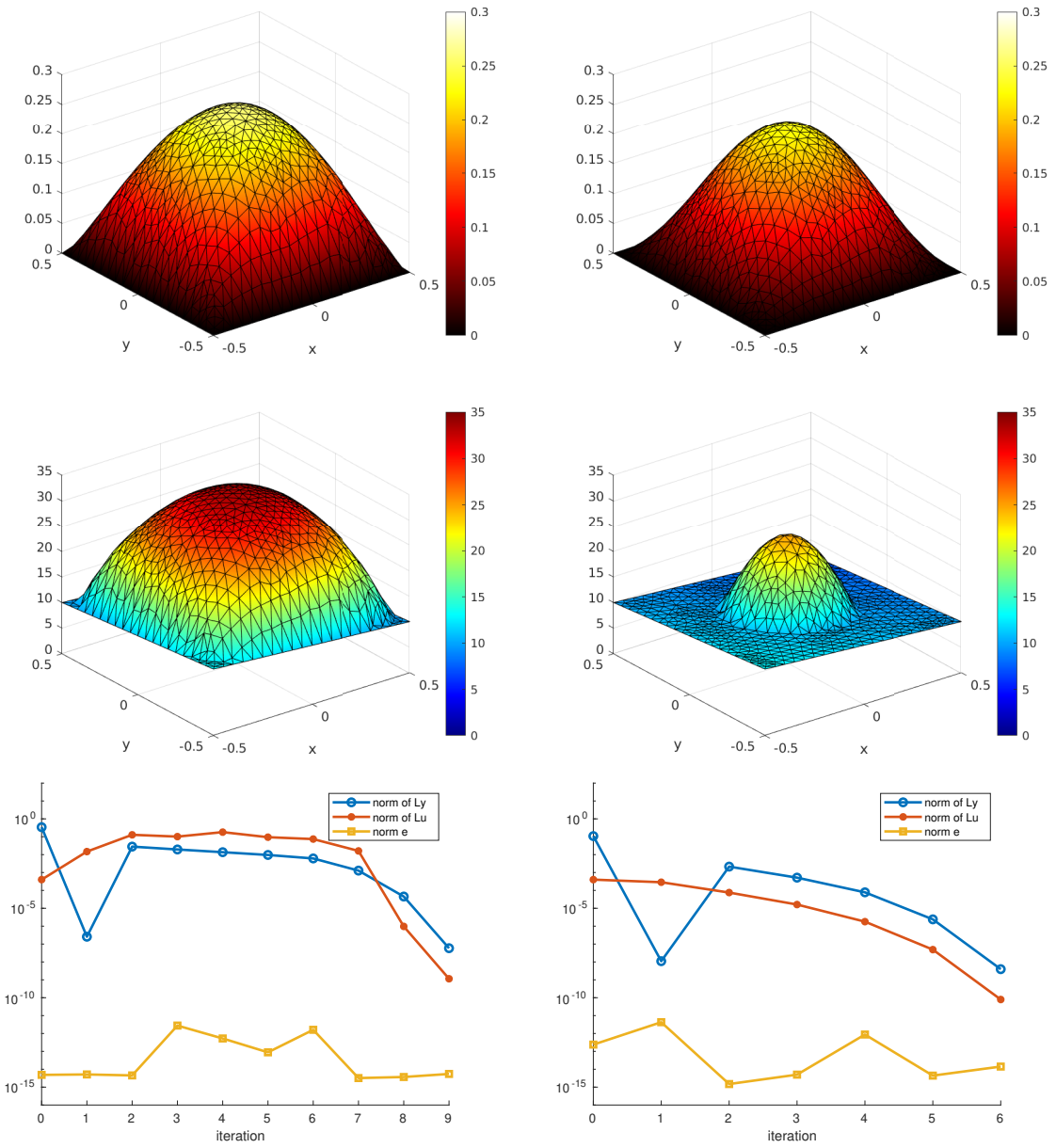

Figure 3.5.1. Optimal states y (top row), optimal controls u (middle row) and convergence history (bottom row) obtained for the example from subsection 3.5.1 for $\alpha = 1$ (left column) and $\alpha = 10^3$ (right column). The three norms shown in the convergence plots correspond to the three terms in (3.4.8), i. e., $\|\mathcal{L}_{\mathbf{y}}(\mathbf{y},\mathbf{u},\mathbf{p})\|_{(\mathbf{K}+\mathbf{M})^{-1}}$, $\|\mathcal{L}_{\mathbf{u}}(\mathbf{y},\mathbf{u},\mathbf{p})\|_{(\mathbf{K}+\mathbf{M})^{-1}}$ and $\|e(\mathbf{y},\mathbf{u})\|_{(\mathbf{K}+\mathbf{M})^{-1}}$.

3.5.2 Dependence on the Discretization

In this experiment we study the dependence of the number of semismooth Newton steps in Algorithm 3.4.1 on the refinement level of the underlying discretization. To this end, we consider a coarse mesh and two uniform refinements; see table 3.5.2.

The problem is similar as in subsection 3.5.1. The domain is $\Omega = (-0.5, 0.5)^2$. We use $f(x, y) \equiv 100$ as right-hand side and the desired state is $y_d(x, y) \equiv 0$. The lower bound for the control is now given as $u_a(x, y) = -10x - 10y + 20$ and the upper bound is $u_b = u_a + 5$. Moreover, the control cost parameters are $\lambda_1 = 10^{-7}$ and $\lambda_2 = 4 \cdot 10^{-5}$. We choose $\varepsilon = 10^{-2}$ as our penalty parameter. The coefficient function determining the degree of non-locality is set to $b(x, y) \equiv 10$. Notice that Assumption 4.1.1 is satisfied for this experiment.

For each mesh, we start from an initial guess constructed as follows. We initialize $\mathbf{u}_0$ to the lower bound $\mathbf{u}_a$ and set $\mathbf{y}_0$ to the numerical solution of the forward problem with control $\mathbf{u}_0$. The adjoint state is initialized to $\mathbf{p}_0 = \mathbf{0}$. In this example, both the lower and upper bounds are relevant on all mesh levels. Nonetheless, we observe a mesh-independent convergence behavior; see figure 3.5.2.

level	N_V	N_T	iterations
1	177	312	11
2	665	1248	8
3	2577	4992	8

Table 3.5.2. Number of iterations of the discrete semismooth Newton method (Algorithm 3.4.1) for various mesh levels in the example from subsection 3.5.2.

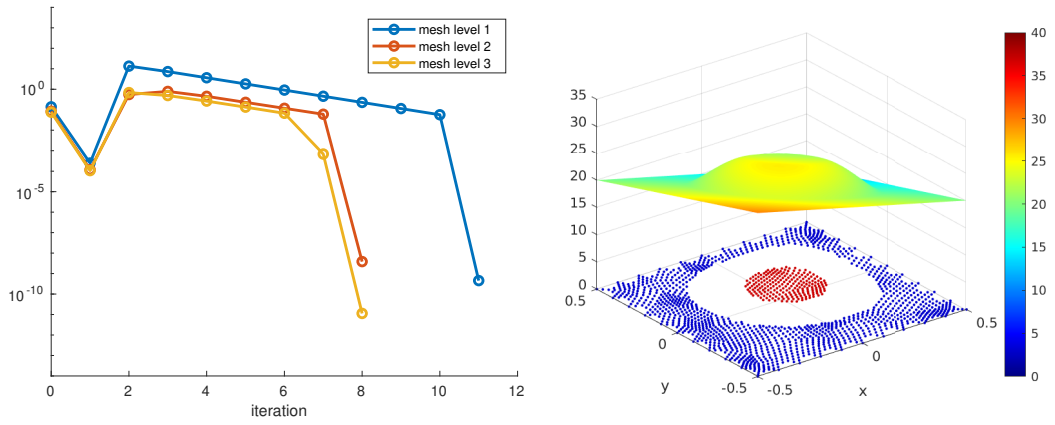

Figure 3.5.2. The convergence plot (left column) shows the total residual norm $R(\mathbf{y}, \mathbf{u}, \mathbf{p})$ as in (3.4.8) on all mesh levels for the example from subsection 3.5.2. The control on the finest level is shown in the right column. Nodes where $u = u_b$ and $u = u_a$ holds are shown in red and blue, respectively.

3.5.3 Influence of the Penalty Parameters

In this experiment, we study the behavior of Algorithm 3.4.1 and the solutions to the penalized problem (P_ε) in dependence of the penalty parameter ε. We solve similar problems as before, with domain $\Omega = (-0.5, 0.5)^2$, right-hand side $f(x, y) \equiv 100$ and desired state $y_d(x, y) \equiv 0$. The lower bound for the control is $u_a(x, y) = -10x - 10y + 20$ and the upper bound is $u_b = u_a + 8$. Moreover, the control cost parameters are $\lambda_1 = 10^{-7}$ and $\lambda_2 = 4 \cdot 10^{-5}$. The penalty parameter varies in $\{10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. The coefficient function determining the degree of non-locality is set to $b(x, y) \equiv 10$.

The construction of an initial guess is the same as in subsection 3.5.2. The experiment is split into two parts. First, we consider Algorithm 3.4.1 without warmstarts. The corresponding results are shown in table 3.5.3. As expected, the number of Newton steps increases as $\varepsilon \searrow 0$ while the norm of the bound violation decreases. Second, we repeat the same experiment with warmstarts. That is, we use the initialization as described above only for the initial value of ε. Subsequent runs of Algorithm 3.4.1 are initialized with the final iterates obtained for the previous value of ε. This strategy is very effective, as shown in figure 3.5.3 (right column).

ε	iterations	$\|(u_a - u)_+\|_{L^\infty(\Omega)}$	$\|(u - u_b)_+\|_{L^\infty(\Omega)}$
1.00e+00	4	1.32e-03	6.36e-05
1.00e-01	4	1.32e-04	6.37e-06
1.00e-02	6	1.32e-05	6.39e-07
1.00e-03	10	1.32e-06	6.40e-08
1.00e-04	13	1.32e-07	6.40e-09

Table 3.5.3. Number of iterations of the discrete semismooth Newton method (Algorithm 3.4.1, without warmstart) for various values of the penalty parameter ε in the example from subsection 3.5.3. The terms $\|(u_a - u)_+\|_{L^\infty(\Omega)}$ and $\|(u - u_b)_+\|_{L^\infty(\Omega)}$ refer to the maximal positive nodal values of $u_a - u$ and $u - u_b$, respectively.

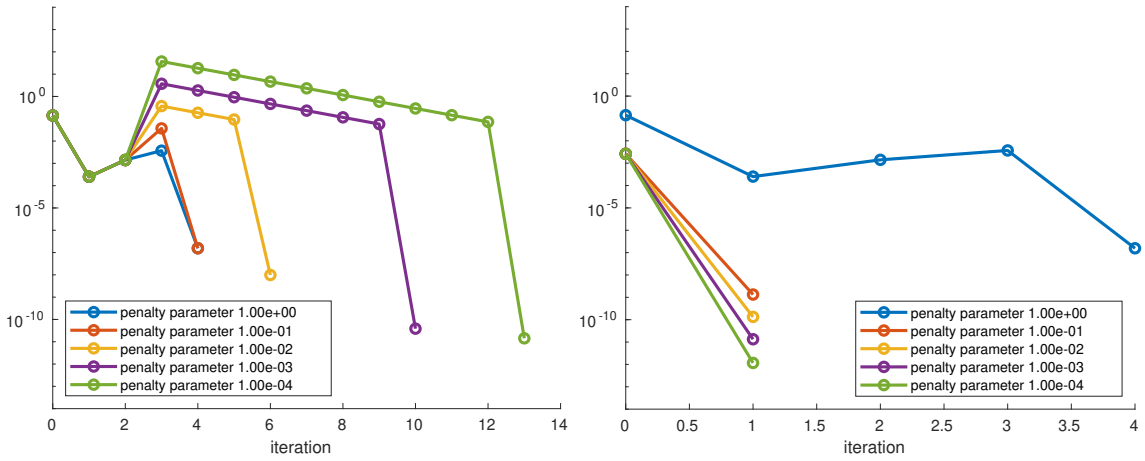

Figure 3.5.3. The convergence plot shows the total residual norm $R(\mathbf{y}, \mathbf{u}, \mathbf{p})$ as in (3.4.8) for all values of the penalty parameter ε. In the left plot, the same initial guess was used for all penalty parameters. With warmstarting, convergence can be achieved in one semismooth Newton step.

3.5.4 Influence of the Control Cost Parameters

In this final experiment, we study variations of the control cost parameters. Specifically, we consider $\lambda_1 \in \{0, 10^{-9}\}$ and $\lambda_2 \in \{10^{-8}, 10^{-7}\}$. The other problem is similar as in subsection 3.5.1. Specifically, we use $\Omega = (-0.5, 0.5)^2$, $f(x, y) = 100$ and desired state $y_d(x, y) = 0$. The lower bound for the control is given again as $u_a(x, y) = -3x - 3y = 10$ and we choose $\varepsilon = 10^{-2}$ as our penalty parameter. The coefficient function determining the degree of non-locality is set to $b(x, y) = 100(x^2 + y^2)$.

figure 3.5.4 shows the controls obtained for each choice of λ_1, λ_2 mentioned above.

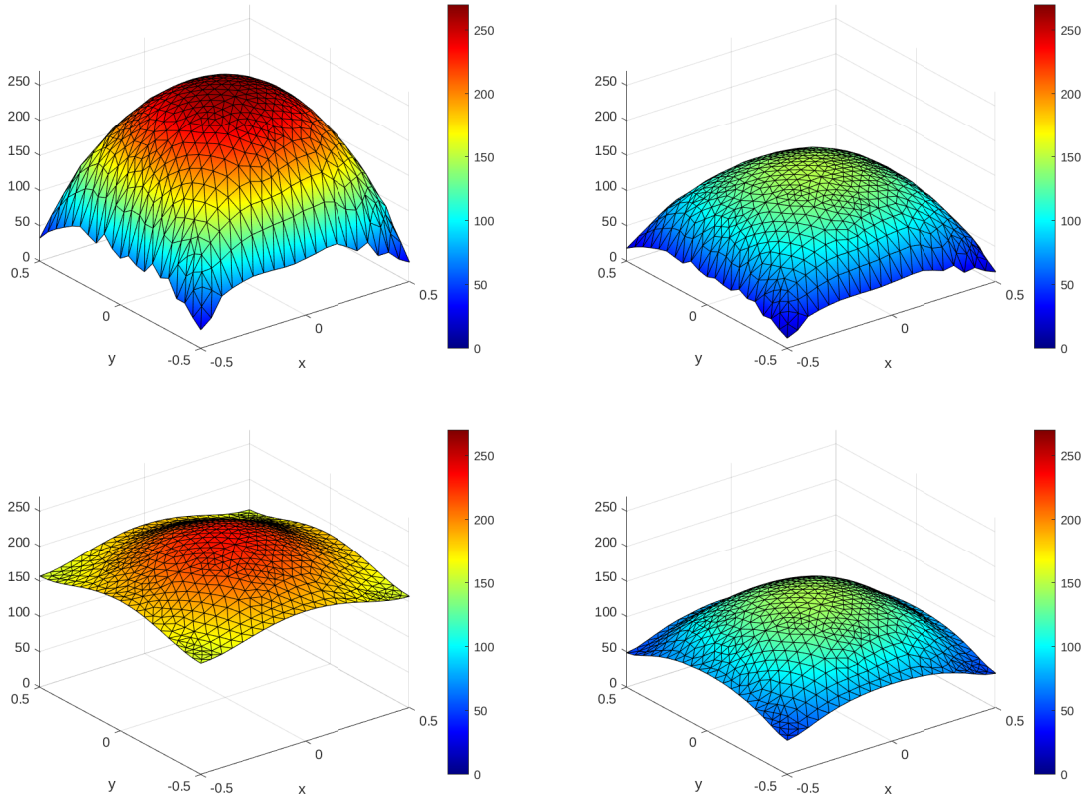

Figure 3.5.4. Optimal control u obtained for the example from subsection 3.5.4 for $(\lambda_1, \lambda_2) = (0, 10^{-8})$ (top left), $(\lambda_1, \lambda_2) = (0, 10^{-7})$ (top right), $(\lambda_1, \lambda_2) = (10^{-9}, 10^{-8})$ (bottom left) and $(\lambda_1, \lambda_2) = (10^{-9}, 10^{-7})$ (bottom right).

4 Optimal Control a Nonlocal Chemotaxis Model

Contents

This chapter is dedicated to an optimal control problem governed by a nonlinear nonlocal parabolic-elliptic chemotaxis system of partial differential equations (4.1.1b). As explained in subsection 1.1.2, chemotaxis is the directed movement of a motile cell or living organisim in response to a chemical concentration gradient. The equation (4.1.1b) describes the evolution of a cell population under chemotactic effects with a logistic growth reaction term defined in terms of the total mass of the population. The variables y and w represent the cell density and the chemoattractant concentration, which is produced by the cell population itself, respectively. The nonlocal term $a_2 \fint_\Omega y \,\mathrm{d}x$, which describes the total mass of the population, has a balancing effect and its impact depends on the sign of a_2. The variable u represents the control imposed on the boundary of the domain as a positive flux of the chemoattractant. The tracking-type cost functional (4.1.1a) measure the discrepancy between the cell density y and desired density y_d at final time T and the control cost.

This chapter is divided into two sections. In the first section, section 4.1, we establish the existence and uniqueness of solutions to the chemotaxis system (4.1.1b), as well as the existence of a solution to the optimal control problem (4.1.1). The first-order necessary optimality conditions are derived in the second section, section 4.2.

For convenience, we have included the main results discussed in chapter 2 that are relevant to the current chapter.

4.1 Optimal Control Problem: Existence Theory

The existence theory is divided into two subsections. In subsection 4.1.1, we introduce a definition of a weak solution to the state equation and prove its existence. The existence of an optimal solution is demonstrated in subsection 4.1.2.

In this work, we are interested in studying the following optimal control problem for a nonlocal, nonlinear parabolic-elliptic system

$$\text{Minimize} \quad J(y,u) \coloneqq \frac{1}{2}\int_\Omega |y(x,T) - y_d(x)|^2 \,\mathrm{d}x + \frac{\gamma}{2}\int_0^T \int_{\partial\Omega} |u(x,t)|^2 \,\mathrm{d}s\,\mathrm{d}t \tag{4.1.1a}$$

$$\text{subject to} \quad \begin{cases} \partial_t y - \Delta y = -\chi \operatorname{div}(y\nabla w) + y\left(a_0 - a_1 y - a_2 \fint_\Omega y \,\mathrm{d}x\right) & \text{in } \Omega_T, \\ -\Delta w + \lambda w = y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega \end{cases} \tag{4.1.1b}$$

$$\text{and} \quad u \in \mathcal{U}_{\text{ad}}. \tag{4.1.1c}$$

The set of admissible controls is given by

$$\mathcal{U}_{\text{ad}} = \{u \in L^\infty(\partial\Omega_T) \,|\, 0 \le u(x,t) \le u_b(x,t) \text{ a.e. on } \partial\Omega_T\}. \tag{4.1.2}$$

According to definition 2.1.36, $\Omega_T := \Omega \times (0,T)$ and $\partial\Omega_T := \partial\Omega \times (0,T)$.

Throughout this chapter, we maintain the following standing assumptions.

Assumption 4.1.1. *We assume that $\Omega \subset \mathbb{R}^N$, $N \ge 1$, is a bounded domain of class $C^{1,1}$ with $N \in \{2,3\}$. Furthermore, the initial cell density y_0 is a non-negative function in $L^\infty(\Omega) \cap H^1(\Omega)$ and the upper bound $u_b(x,t)$ in the set of admissible controls $\mathcal{U}_{\text{ad}}$ belongs to $L^\infty(\partial\Omega_T)$ for almost every $(x,t) \in \partial\Omega_T$. Moreover, the chemotactic coefficient χ, the control cost parameter γ and the reproduction rate λ of the chemoattractant are positive constants. Finally, the constants a_0, a_1 in the logistic growth term are positive, and $a_2 \in \mathbb{R}$ holds. Furthermore, we require*

$$-a_1 + [a_2]_- < 0, \tag{4.1.3}$$

where $[a_2]_- = -\min\{a_2, 0\} \ge 0$ denotes the negative part of a_2.

4.1.1 Existence of a Weak Solution

We begin our analysis with the definition of the weak solutions to problem (4.1.1b). We work with the well-known space

$$W(0,T) = \left\{ y \in L^2(0,T; H^1(\Omega)) \,\middle|\, \partial_t y \in L^2(0,T; H^1(\Omega)^*) \right\},$$

equipped with the norm

$$\|y\|_{W(0,T)} := \left(\int_0^T \|y\|^2_{H^1(\Omega)} + \|\partial_t y\|^2_{H^1(\Omega)^*} \,\mathrm{d}t \right)^{1/2}.$$

Definition 4.1.2. *Let $u \in L^\infty(\partial\Omega_T)$. The pair $(y,w) \in W(0,T) \times L^\infty(0,T; H^1(\Omega))$ is said to be a weak solution of (4.1.1b) if it satisfies*

$$\begin{aligned} \int_\Omega \partial_t y\, \varphi \,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla \varphi \,\mathrm{d}x &= \chi \int_\Omega y \,\nabla w \cdot \nabla \varphi \,\mathrm{d}x - \chi \int_{\partial\Omega} u\, y\, \varphi \,\mathrm{d}x \\ &\quad + \int_\Omega \left(a_0 - a_1\, y - a_2 \fint_\Omega y dx \right) y\, \varphi \,\mathrm{d}x, \end{aligned} \tag{4.1.4}$$

$$\int_\Omega \nabla w \cdot \nabla \varphi \,\mathrm{d}x - \int_{\partial\Omega} u\, \varphi \,\mathrm{d}s + \lambda \int_\Omega w\, \varphi \,\mathrm{d}x = \int_\Omega y\, \varphi \,\mathrm{d}x, \tag{4.1.5}$$

$$y(x,0) = y_0(x) \tag{4.1.6}$$

for all $\varphi \in C^\infty(\overline{\Omega})$ and for a.a. $t \in [0,T]$.

Expanding $\operatorname{div}(y \nabla w) = \nabla y \cdot \nabla w + y\, \Delta w$ in the first equation of (4.1.1b) and inserting the second equation of (4.1.1b) into the first one, we can rewrite (4.1.1b) as follows:

$$\begin{cases} \partial_t y - \Delta y = -\chi\, \nabla y \cdot \nabla w + y \left(-\chi\, \lambda w + \chi\, y + a_0 - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x \right) & \text{in } \Omega_T, \\ -\Delta w + \lambda\, w = y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega. \end{cases} \tag{4.1.7}$$

In the following, we define the weak solution to problem (4.1.7) and observe that definition 4.1.2 and definition 4.1.3 are equivalent.

Definition 4.1.3. *Let $u \in L^\infty(\partial\Omega_T)$. The pair $(y, w) \in W(0,T) \times L^\infty(0,T;H^1(\Omega))$ is said to be a weak solution of (4.1.7) if it satisfies*

$$\int_\Omega \partial_t y\,\varphi\,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla\varphi\,\mathrm{d}x = -\chi \int_\Omega \nabla y \cdot \nabla w\,\varphi\,\mathrm{d}x + \int_\Omega \left(-\chi\,\lambda w + \chi\,y + a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right) y\,\varphi\,\mathrm{d}x, \tag{4.1.8}$$

$$\int_\Omega \nabla w \cdot \nabla\varphi\,\mathrm{d}x - \int_{\partial\Omega} u\,\varphi\,\mathrm{d}s + \lambda \int_\Omega w\,\varphi\,\mathrm{d}x = \int_\Omega y\,\varphi\,\mathrm{d}x, \tag{4.1.9}$$

$$y(x,0) = y_0(x) \tag{4.1.10}$$

for all $\varphi \in C^\infty(\overline{\Omega})$ and for a.a. $t \in [0,T]$.

Remark 4.1.4. *It can be easily proven that the weak solutions of (4.1.1b) and (4.1.7) are equivalent. The equivalence is demonstrated by establishing the similarity of the first conditions in both definition 4.1.2 and definition 4.1.3. Indeed, multiplying the first equation in (4.1.1b) with the test function $\varphi \in C^\infty(\overline{\Omega})$ and the second one with $y\,\varphi \in H^1(\Omega)$ for a.a. $t \in [0,T]$, we obtain, using integration by parts,*

$$\int_\Omega \partial_t y\,\varphi\,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla\varphi\,\mathrm{d}x = \chi \int_\Omega y\,\nabla w \cdot \nabla\varphi\,\mathrm{d}x - \chi \int_{\partial\Omega} u\,y\,\varphi\,\mathrm{d}s + \int_\Omega \left(a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right) y\,\varphi\,\mathrm{d}x \tag{4.1.11a}$$

and

$$\int_\Omega \nabla w \cdot \nabla(y\,\varphi)\,\mathrm{d}x - \int_{\partial\Omega} u\,y\,\varphi\,\mathrm{d}s + \lambda \int_\Omega w\,y\,\varphi\,\mathrm{d}x = \int_\Omega y^2\,\varphi\,\mathrm{d}x. \tag{4.1.11b}$$

Inserting (4.1.11a) into (4.1.11b) yields

$$\begin{aligned}\int_\Omega \partial_t y\,\varphi\,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla\varphi\,\mathrm{d}x &= \chi \int_\Omega y\,\nabla w \cdot \nabla\varphi\,\mathrm{d}x - \chi \int_\Omega \nabla w \cdot \nabla\,(y\,\varphi)\,\mathrm{d}x \\ &\quad - \chi\,\lambda \int_\Omega w\,y\,\varphi\,\mathrm{d}x + \chi \int_\Omega y^2\,\varphi\,\mathrm{d}x + \int_\Omega y \left(a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right)\varphi\,\mathrm{d}x \\ &= -\chi \int_\Omega \nabla w \cdot \nabla y\,\varphi\,\mathrm{d}x + \int_\Omega \left(-\chi\,\lambda w + \chi\,y + a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right) y\,\varphi\,\mathrm{d}x\end{aligned}$$

and this is exactly definition 4.1.3.

For the reverse direction, we multiply the first equation in (4.1.7) with the test function $\varphi \in C^\infty(\overline{\Omega})$ and the second one with $y\,\varphi \in H^1(\Omega)$ and obtain, using integration by parts,

$$\int_\Omega \partial_t y\,\varphi\,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla\varphi\,\mathrm{d}x = -\chi \int_\Omega \nabla y \cdot \nabla w\,\varphi\,\mathrm{d}x + \int_\Omega \left(-\chi\,\lambda\,w + \chi\,y + a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right) y\,\varphi\,\mathrm{d}x \tag{4.1.12a}$$

and

$$\int_\Omega \nabla w \cdot \nabla(y\,\varphi)\,\mathrm{d}x - \int_{\partial\Omega} u\,y\,\varphi\,\mathrm{d}s + \lambda \int_\Omega w\,y\,\varphi\,\mathrm{d}x = \int_\Omega y^2\,\varphi\,\mathrm{d}x. \tag{4.1.12b}$$

Inserting (4.1.12a) into (4.1.12b) yields

$$\int_\Omega \partial_t y\,\varphi\,\mathrm{d}x + \int_\Omega \nabla y \cdot \nabla\varphi\,\mathrm{d}x = -\chi \int_\Omega \nabla y \cdot \nabla w\,\varphi\,\mathrm{d}x + \chi \int_\Omega \nabla w \cdot \nabla\,(y\,\varphi)\,\mathrm{d}x$$

$$
\begin{aligned}
&-\chi \int_{\partial\Omega} u\, y\, \varphi \,\mathrm{d}s + \int_\Omega \left(a_0 - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x \right) y\, \varphi \,\mathrm{d}x \\
&= \chi \int_\Omega y\, \nabla w \cdot \nabla \varphi \,\mathrm{d}x + \int_\Omega \left(a_0 \; - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x \right) y\, \varphi \,\mathrm{d}x
\end{aligned}
$$

and this is exactly definition 4.1.2.

We can attain a higher regularity for w, which is also essential for our proofs throughout this chapter.

The following theorem is a deduced result from Grisvard, 1985, Theorem 2.2.2.5:

Theorem 4.1.5. *Suppose that Ω is a bounded $C^{1,1}$ domain and $\mathcal{A}$ is an elliptic differential operator of the form*

$$
\mathcal{A}y(x) = -\sum_{i,j=1}^{n} (a_{ij}(x) y_{x_j}(x))_{x_i}, \quad x \in \Omega.
$$

The coefficient functions a_{ij} of $\mathcal{A}$ are assumed to belong to $C^{0,1}(\overline{\Omega})$ and satisfy the symmetric condition $a_{ij}(x) = a_{ji}(x)$ for all $i, j \in \{1, \dots, n\}$ and $x \in \Omega$.

If $\lambda > 0$ and the right-hand side function $y \in L^2(\Omega)$, then the weak solution to the following Neumann problem

$$
\begin{cases} \mathcal{A}w + \lambda w = y & \text{in } \Omega, \\ \dfrac{\partial w}{\partial n} = 0 & \text{on } \partial\Omega, \end{cases}
$$

belongs to $H^2(\Omega)$.

In addition, the following theorem is an inferred result from Morrey, 1966, Chapter 5, Section 5.5.

Theorem 4.1.6. *Let Ω be a bounded C^1 domain. For given right-hand side functions $y \in W^{1,6/5}(\Omega)^*$ and $u \in L^6(\partial\Omega)$, the following PDE*

$$
\begin{cases} -\Delta w + \lambda\, w = y & \text{in } \Omega, \\ \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega, \end{cases}
$$

possesses a unique weak solution $w \in W^{1,6}(\Omega)$.

In our work $y(t)$ belongs to $H^1(\Omega) \hookrightarrow L^2(\Omega)$. By virtue of theorem 2.1.16, $W^{1,6/5}(\Omega)$ is densely embedded in $L^2(\Omega)$, which implies $L^2(\Omega) \hookrightarrow W^{1,6/5}(\Omega)^*$. Therefore, the above result is applicable to our work.

For our purpose, it suffices that w belongs to $W^{1,4}(\Omega)$. Nevertheless, having $w \in W^{1,6}(\Omega)$, could simplify some estimates in our proofs.

We now proceed to discuss the well-posedness of (4.1.7). To this end, we rely on some classical results. First, we introduce the definition of a closed form:

Definition 4.1.7. *Let $T > 0$ and V, H to be Hilbert spaces over $\mathbb{R}$ such that V is continuously and densely embedded in H. We consider*

$$
\mathfrak{a} \colon [0, T] \times V \times V \to \mathbb{R}.
$$

We say that $\mathfrak{a}$ is a closed form if the following properties hold:

- *(i) $\mathfrak{a}$ is a non-autonomous form, i. e., $\mathfrak{a}(t; \cdot, \cdot)$ is bilinear for all $t \in [0, T]$, and $\mathfrak{a}(\cdot, y, v)$ is measurable for all y, $v \in V$.*
- *(ii) $\mathfrak{a}$ is V-bounded, i. e., there exists some time-independent constant $\alpha = \alpha(T) > 0$.*

$$
|\mathfrak{a}(t; y, v)| \leq \alpha \, \|y\|_V \, \|v\|_V \quad \text{for all } t \in [0, T] \text{ and } y,\, v \in V.
$$

(*iii*) $\mathfrak{a}$ *is quasi-coercive, i. e., there exist time-independent constants* $\beta > 0$ *and* $\omega \in \mathbb{R}$ *such that*

$$\mathfrak{a}(t; y, y) + \omega \|y\|_H^2 \geq \beta \|y\|_V^2 \quad \text{for all } t \in [0, T] \text{ and } y \in V.$$

We note that for each $t \in [0, T]$ the bounded non-autonomous form $\mathfrak{a}(t; y, v)$ defines an operator $\mathcal{A}(t) \in \mathcal{L}(V, V^*)$

$$\langle \mathcal{A}(t) y \,, v \rangle = (\mathfrak{a} + \omega)(t; y, v) \quad \text{for all } y,\ v \in V.$$

Here $\langle \cdot \,, \cdot \rangle$ denotes the duality between V and its dual space V^* and the form $(\mathfrak{a} + \omega)(t; y, v)$ is defined by

$$(\mathfrak{a} + \omega)(t; y, v) \coloneqq \mathfrak{a}(t; y, v) + \omega (y, v)_H.$$

Now a classical result states the following, see Dautray, Lions, 2000, Chapter XVIII, Section 3.

In the following theorem, we suppose that V is dense in H, and we identify H with its dual H^*, meaning we have

$$V \hookrightarrow H \hookrightarrow V^*$$

and $y \in W(0, T; V, V^*)$ is defined by

$$W(0, T; V, V^*) \coloneqq \{ y \in L^2(0, T; V) \,|\, \partial_t y \in L^2(0, T; V^*) \}.$$

We note that for $V = H^1(\Omega)$, this space coincide with $W(0, T)$.

Theorem 4.1.8. *Assume that* $\mathfrak{a}(t; y, v)$ *to be a closed form with its associated operator* $\mathcal{A}(t)$. *For every* $f \in L^2(0, T; V^*)$ *and* $y_0 \in H$ *there exists a unique solution* $y \in W(0, T; V, V^*)$ *such that*

$$\begin{cases} \partial_t y + \mathcal{A}(t)\, y = f(t), \\ \qquad\quad y(0) = y_0. \end{cases} \tag{4.1.13}$$

We consider the following linear parabolic equation

$$\begin{cases} \partial_t y - \Delta y = -\chi \, \nabla w \cdot \nabla y + g \, y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \partial\Omega_T, \\ y(x, 0) = y_0(x) & \text{in } \Omega, \end{cases} \tag{4.1.14}$$

where $\nabla w \in L^\infty(0, T; L^4(\Omega))$, $g \in L^\infty(0, T; L^2(\Omega))$, and $y_0 \in L^2(\Omega)$ are given functions and χ is some given constant.

For $T > 0$, we set

$$\mathcal{A}(t) \coloneqq -\Delta y + \chi \nabla w \cdot \nabla y - g \, y \quad t \in [0, T]$$

and define

$$\begin{aligned} &\mathfrak{a} \colon [0, T] \times H^1(\Omega) \times H^1(\Omega) \to \mathbb{R} \\ &\mathfrak{a}(t; y, v) = \int_\Omega \nabla y \cdot \nabla v + \chi \int_\Omega \nabla w \cdot \nabla y \, v \, \mathrm{d}x - \int_\Omega g \, y \, v \, \mathrm{d}x. \end{aligned} \tag{4.1.15}$$

Now, we can show the existence of a unique solution to (4.1.14).

Theorem 4.1.9. *Let* $y_0 \in L^2(\Omega)$, $\nabla w \in L^\infty(0, T; L^4(\Omega))$ *and* $g \in L^\infty(0, T; L^2(\Omega))$ *be given functions. Then,* (4.1.14) *has a unique weak solution* $y \in W(0, T)$.

Proof. Obviously, it holds

$$H^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^1(\Omega)^*.$$

In the following, we will verify the three properties in definition 4.1.7, asserting that $\mathfrak{a}$ is a closed form. As explained in definition B.2.5, we use the notation $\lesssim$ to avoid excessive estimation constants in embeddings.

(*i*) Obviously, $\mathfrak{a}(t;\cdot,\cdot)$ is bilinear for each $t \in [0,T]$, and $\mathfrak{a}(\cdot,y,v)$ is measurable for every y, $v \in H^1(\Omega)$.

(*ii*) For each $t \in [0,T]$, $\mathfrak{a}$ is H^1-bounded with t-independent bound: We apply Hölder's inequality and Sobolev embeddings, resulting in:

$$\begin{aligned}
|\mathfrak{a}(t;y,v)| &= \left| \int_\Omega \nabla y \cdot \nabla v \,\mathrm{d}x + \chi \int_\Omega \nabla w \cdot \nabla y \, v \,\mathrm{d}x - \int_\Omega g \, y \, v \,\mathrm{d}x \right| \\
&\le \|\nabla y\|_{L^2(\Omega)} \, \|\nabla v\|_{L^2(\Omega)} + \chi \, \|\nabla w\|_{L^3(\Omega)} \, \|\nabla y\|_{L^2(\Omega)} \, \|v\|_{L^6(\Omega)} \\
&\quad + \|g\|_{L^2(\Omega)} \, \|y\|_{L^3(\Omega)} \, \|v\|_{L^6(\Omega)} \\
&\lesssim \|y\|_{H^1(\Omega)} \, \|v\|_{H^1(\Omega)} + \|\nabla w\|_{L^\infty(0,T;L^3(\Omega))} \, \|y\|_{H^1(\Omega)} \, \|v\|_{H^1(\Omega)} \\
&\quad \|g\|_{L^\infty(0,T;L^2(\Omega))} \, \|y\|_{H^1(\Omega)} \, \|v\|_{H^1(\Omega)} \\
&\lesssim c \, \|y\|_{H^1(\Omega)} \, \|v\|_{H^1(\Omega)} .
\end{aligned}$$

(*iii*) For each $t \in [0,T]$, $\mathfrak{a}$ is quasi-coercive: We apply Hölder's and Young's inequalities, as well as Sobolev embeddings, and obtain:

$$\begin{aligned}
\mathfrak{a}(t;y,y) &= \int_\Omega |\nabla y|^2 \,\mathrm{d}x + \chi \int_\Omega \nabla w \cdot \nabla y \, y \,\mathrm{d}x - \int_\Omega g \, y^2 \,\mathrm{d}x \\
&\ge \|\nabla y\|^2_{L^2(\Omega)} - \chi \, \|\nabla w\|_{L^4(\Omega)} \, \|\nabla y\|_{L^2(\Omega)} \, \|y\|_{L^4(\Omega)} - \|g\|_{L^2(\Omega)} \, \|y\|^2_{L^4(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi \, \|\nabla w\|_{L^4(\Omega)} \, \|\nabla y\|_{L^2(\Omega)} \, \|y\|^{1/4}_{L^2(\Omega)} \, \|y\|^{3/4}_{L^6(\Omega)} \\
&\quad - \|g\|_{L^2(\Omega)} \left(\|y\|^{1/4}_{L^2(\Omega)} \, \|y\|^{3/4}_{L^6(\Omega)} \right)^2 \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi \, \|\nabla w\|_{L^4(\Omega)} \, \|\nabla y\|_{L^2(\Omega)} \, \|y\|^{1/4}_{L^2(\Omega)} \left(\|y\|_{L^2(\Omega)} + \|\nabla y\|_{L^2(\Omega)} \right)^{3/4} \\
&\quad - \|g\|_{L^2(\Omega)} \, \|y\|^{1/2}_{L^2(\Omega)} \, \|y\|^{3/2}_{H^1(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi \, \|\nabla w\|_{L^4(\Omega)} \, \|\nabla y\|_{L^2(\Omega)} \, \|y\|^{1/4}_{L^2(\Omega)} \left(\|y\|^{3/4}_{L^2(\Omega)} + \|\nabla y\|^{3/4}_{L^2(\Omega)} \right) \\
&\quad - \|g\|_{L^2(\Omega)} \, \|y\|^{1/2}_{L^2(\Omega)} \, \|y\|^{3/2}_{H^1(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi \, \|\nabla w\|_{L^4(\Omega)} \, \|\nabla y\|_{L^2(\Omega)} \, \|y\|_{L^2(\Omega)} \\
&\quad - \chi \, \|\nabla w\|_{L^4(\Omega)} \, \|\nabla y\|^{7/4}_{L^2(\Omega)} \, \|y\|^{1/4}_{L^2(\Omega)} - \|g\|_{L^2(\Omega)} \, \|y\|^{1/2}_{L^2(\Omega)} \, \|y\|^{3/2}_{H^1(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi \, \|\nabla w\|_{L^4(\Omega)} \left[\varepsilon \, \|\nabla y\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|y\|^2_{L^2(\Omega)} \right] \\
&\quad - \chi \, \|\nabla w\|_{L^4(\Omega)} \left[\varepsilon \left(\|\nabla y\|^{7/4}_{L^2(\Omega)} \right)^{8/7} + C(\varepsilon) \left(\|y\|^{1/4}_{L^2(\Omega)} \right)^8 \right] \\
&\quad - \|g\|_{L^2(\Omega)} \left[\varepsilon \left(\|y\|^{3/2}_{H^1(\Omega)} \right)^{4/3} + C(\varepsilon) \left(\|y\|^{1/2}_{L^2(\Omega)} \right)^4 \right] \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - 2\,\chi \, \|\nabla w\|_{L^\infty(0,T;L^4(\Omega))} \left[\varepsilon \, \|\nabla y\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|y\|^2_{L^2(\Omega)} \right] \\
&\quad - \|g\|_{L^\infty(0,T;L^2(\Omega))} \left[\varepsilon \, \|\nabla y\|^2_{L^2(\Omega)} + (\varepsilon + C(\varepsilon)) \, \|y\|^2_{L^2(\Omega)} \right] \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \left[2\,\varepsilon\chi \, \|\nabla w\|_{L^\infty(0,T;L^4(\Omega))} + \varepsilon \, \|g\|_{L^\infty(0,T;L^2(\Omega))} \right] \|\nabla y\|^2_{L^2(\Omega)} \\
&\quad - \left[2\,C(\varepsilon)\chi \, \|\nabla w\|_{L^\infty(0,T;L^4(\Omega))} + (\varepsilon + C(\varepsilon)) \, \|g\|_{L^\infty(0,T;L^2(\Omega))} \right] \|y\|^2_{L^2(\Omega)} .
\end{aligned}$$

Choosing $\varepsilon = 1/\left(2\left[2\chi \left\|\nabla w\right\|_{L^\infty(0,T;L^4(\Omega))} + \left\|g\right\|_{L^\infty(0,T;L^2(\Omega))}\right]\right)$ we obtain

$$\mathfrak{a}(t;y,y) \geq \frac{1}{2}\left\|\nabla y\right\|^2_{L^2(\Omega)} - \left[C(\chi, \left\|\nabla w\right\|_{L^\infty(0,T;L^3(\Omega))}, \left\|g\right\|_{L^\infty(0,T;L^2(\Omega))})\right]\left\|y\right\|^2_{L^2(\Omega)}$$

and consequently

$$\mathfrak{a}(t;y,y) + \left[\frac{1}{2} + C(\chi, \left\|\nabla w\right\|_{L^\infty(0,T;L^3(\Omega))}, \left\|g\right\|_{L^\infty(0,T;L^2(\Omega))})\right]\left\|y\right\|^2_{L^2(\Omega)} \geq \frac{1}{2}\left\|y\right\|^2_{H^1(\Omega)}.$$

By setting $\omega := \frac{1}{2} + C(\chi, \left\|\nabla w\right\|_{L^\infty(0,T;L^3(\Omega))}, \left\|g\right\|_{L^\infty(0,T;L^2(\Omega))})$, $\mathfrak{a}$ is quasi-coercive.

Due to the closed form nature of $\mathfrak{a}$, the existence of a unique solution $y \in W(0,T)$ is readily established by applying theorem 4.1.8 to the problem defined in (4.1.17). □

In what follows, we aim to establish the unique solution to (4.1.14) is both positive and bounded from above.

First, we recall some standard notations. Given $v \in L^2(\Omega)$, we set

$$v_+ := \max\{v(x), 0\},\ v_- := -\min\{v(x), 0\}, v \wedge 1 := \min\{v(x), 1\} \quad \text{for } a.e.\ x \in \Omega.$$

We write $v \geq 0$ as shorthand for $v(x) \geq 0$ for a.e. $x \in \Omega$ and $L^2(\Omega)_+ := \{v \in L^2(\Omega) \,|\, v \geq 0\}$.

To ensure the positivity of the solution, we refer to the result presented in Arendt, Dier, Ouhabaz, 2014, Proposition 3.1.

Proposition 4.1.10. *Let $\mathfrak{a}$ be a closed form and V be a sublattice of $L^2(\Omega)$, i. e., $v \in V$ implies $v_+ \in V$. Assume $\mathfrak{a}(t; v_+, v_-) \leq 0$ for a.e. $t \in [0,T]$ and all $v \in V$.*

We additionally suppose that $y_0 \in V_+ := L^2(\Omega)_+ \cap V$ and $f \geq 0$. Then, the solution of (4.1.13) satisfies $y(t) \geq 0$ for a.e. $t \in [0,T]$.

To establish the boundedness of the solution from above, we refer to the result presented in Arendt, Dier, Ouhabaz, 2014, Proposition 3.2.

Proposition 4.1.11. *Let $\mathfrak{a}$ be a closed form and $v \wedge 1 \in V$. Assume $\mathfrak{a}(t; v \wedge 1, (v-1)_+) \geq 0$ for all $t \in [0,T]$, $v \in V$.*

We additionally suppose that $y_0 \in L^2(\Omega)$, such that $y_0(x) \leq 1$ for a.e. $x \in \Omega$ and $f \leq 0$. Then, the solution y of (4.1.13) satisfies $y(x,t) \leq 1$.

It is noteworthy that if $y_0(x) \leq M$, where M is a positive number, then $y(x,t) \leq M$, provided that $v \wedge M \in V$ and $\mathfrak{a}(t; v \wedge M, (v-M)_+) \geq 0$. This is an immediate consequence of proposition 4.1.11 according to the proof of Arendt, Dier, Ouhabaz, 2014, Proposition 3.2. and is explicitly stated therein.

Theorem 4.1.12. *Suppose that the assumptions in theorem 4.1.9 hold, and $y_0 \in H^1(\Omega)_+ := L^2(\Omega)_+ \cap H^1(\Omega)$. Then, the solution of (4.1.14) satisfies $y(t) \geq 0$ for a.e. $t \in [0,T]$.*

Suppose additionally, that $g_+ \in L^\infty(0,T;L^\infty(\Omega))$ and $y_0(x) \leq M$ for a.e. $x \in \Omega$. Then, the solution of (4.1.14) satisfies $y(x,t) \leq M$ for a.e. $x \in \Omega$ and for all $t \in [0,T]$.

Proof. To show the positivity of the solution we verify the assumption of proposition 4.1.10. Since $v \in H^1(\Omega)$ implies $v_+ \in H^1(\Omega)$, $H^1(\Omega)$ is a sublattice of $L^2(\Omega)$.

$$\begin{aligned}
\mathfrak{a}(t; v_+, v_-) &= \int_\Omega \nabla v_+ \cdot \nabla v_- \,\mathrm{d}x + \chi \int_\Omega \nabla w \cdot \nabla v_+ v_- \,\mathrm{d}x - \int_\Omega g\, v_+\, v_- \,\mathrm{d}x \\
&= \int_\Omega \sum_{k=1}^n \partial_k v\, \chi_{\{v>0\}} \left(-\partial_k v\, \chi_{\{v<0\}}\right) \mathrm{d}x + \int_\Omega \sum_{k=1}^n \partial_k w\, \partial_k v\, \chi_{\{v>0\}}\, v_- \,\mathrm{d}x \\
&\quad - \int_\Omega g\, v\, \chi_{\{v>0\}}\, v\, \chi_{\{v<0\}} \,\mathrm{d}x = 0,
\end{aligned}$$

where χ stands for the indicator function.

This implies that when $y_0 \in H^1(\Omega)_+ := L^2(\Omega)_+ \cap H^1(\Omega)$ and $f \geq 0$, the solution of (4.1.14) satisfies $y(t) \geq 0$ for a.e. $t \in [0,T]$, by virtue of proposition 4.1.10.

To show the boundedness of the solution we verify the assumption of proposition 4.1.11. Obviously $v \wedge M \in H^1(\Omega)$. We show $(\mathfrak{a}+\omega)(t; v\wedge M, (v-M)_+) \geq 0$ for all $t \in [0,T]$, $v \in H^1(\Omega)$.

$$
\begin{aligned}
(\mathfrak{a}+\omega)(t; v\wedge M, (v-M)_+) &= \mathfrak{a}(t; v\wedge M, (v-M)_+) + \omega(v\wedge M, (v-M)_+)_{L^2(\Omega)} \\
&= \int_\Omega \sum_k^n \partial_k (v\wedge M)\, \partial_k (v-M)_+ \,\mathrm{d}x + \chi \int_\Omega \sum_{k=1}^n \partial_k w\, \partial_k (v\wedge M)(v-M)_+ \,\mathrm{d}x \\
&\quad - \int_\Omega g\,(v\wedge M)(v-M)_+ \,\mathrm{d}x + \omega \int_\Omega (v\wedge M)(v-M)_+ \,\mathrm{d}x \\
&= -M \int_\Omega g\,(v-M)_+ \,\mathrm{d}x + \omega\, M \int_\Omega (v-M)_+ \,\mathrm{d}x \\
&\geq \int_\Omega (\omega - g_+)(v-M)_+ \,\mathrm{d}x \geq 0
\end{aligned}
$$

We note that $g_+ \in L^\infty(0,T;L^\infty(\Omega))$, therefore ω can be calculated in terms of $\|g_+\|_{L^\infty(0,T;L^\infty(\Omega))}$ as follows.

$$
\begin{aligned}
\mathfrak{a}(t;y,y) &= \int_\Omega |\nabla y|^2 \,\mathrm{d}x + \chi \int_\Omega \nabla w \cdot \nabla y\, y \,\mathrm{d}x - \int_\Omega g\, y^2 \,\mathrm{d}x \\
&\geq \|\nabla y\|^2_{L^2(\Omega)} - \chi\, \|\nabla w\|_{L^4(\Omega)}\, \|\nabla y\|_{L^2(\Omega)}\, \|y\|_{L^4(\Omega)} - \|g_+\|_{L^\infty(\Omega)}\, \|y\|^2_{L^2(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi\, \|\nabla w\|_{L^4(\Omega)}\, \|\nabla y\|_{L^2(\Omega)}\, \|y\|^{1/4}_{L^2(\Omega)}\, \|y\|^{3/4}_{L^6(\Omega)} \\
&\quad - \|g_+\|_{L^\infty(\Omega)}\, \|y\|^2_{L^2(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi\, \|\nabla w\|_{L^4(\Omega)}\, \|\nabla y\|_{L^2(\Omega)}\, \|y\|^{1/4}_{L^2(\Omega)} \left(\|y\|_{L^2(\Omega)} + \|\nabla y\|_{L^2(\Omega)}\right)^{3/4} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi\, \|\nabla w\|_{L^4(\Omega)}\, \|\nabla y\|_{L^2(\Omega)}\, \|y\|^{1/4}_{L^2(\Omega)} \left(\|y\|^{3/4}_{L^2(\Omega)} + \|\nabla y\|^{3/4}_{L^2(\Omega)}\right) \\
&\quad - \|g_+\|_{L^\infty(\Omega)}\, \|y\|^2_{L^2(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi\, \|\nabla w\|_{L^4(\Omega)}\, \|\nabla y\|_{L^2(\Omega)}\, \|y\|_{L^2(\Omega)} \\
&\quad - \chi\, \|\nabla w\|_{L^4(\Omega)}\, \|\nabla y\|^{7/4}_{L^2(\Omega)}\, \|y\|^{1/4}_{L^2(\Omega)} - \|g_+\|_{L^\infty(\Omega)}\, \|y\|^2_{L^2(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \chi\, \|\nabla w\|_{L^4(\Omega)} \left[\varepsilon\, \|\nabla y\|^2_{L^2(\Omega)} + C(\varepsilon)\, \|y\|^2_{L^2(\Omega)}\right] \\
&\quad - \chi\, \|\nabla w\|_{L^4(\Omega)} \left[\varepsilon \left(\|\nabla y\|^{7/4}_{L^2(\Omega)}\right)^{8/7} + C(\varepsilon) \left(\|y\|^{1/4}_{L^2(\Omega)}\right)^8\right] \\
&\quad - \|g_+\|_{L^\infty(\Omega)}\, \|y\|^2_{L^2(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - 2\,\chi\, \|\nabla w\|_{L^\infty(0,T;L^4(\Omega))} \left[\varepsilon\, \|\nabla y\|^2_{L^2(\Omega)} + C(\varepsilon)\, \|y\|^2_{L^2(\Omega)}\right] \\
&\quad - \|g_+\|_{L^\infty(0,T;L^\infty(\Omega))}\, \|y\|^2_{L^2(\Omega)} \\
&\gtrsim \|\nabla y\|^2_{L^2(\Omega)} - \left[2\,\varepsilon\,\chi\, \|\nabla w\|_{L^\infty(0,T;L^4(\Omega))}\right] \|\nabla y\|^2_{L^2(\Omega)} \\
&\quad - \left[2\, C(\varepsilon)\chi\, \|\nabla w\|_{L^\infty(0,T;L^4(\Omega))} + \|g_+\|_{L^\infty(0,T;L^\infty(\Omega))}\right] \|y\|^2_{L^2(\Omega)}\,.
\end{aligned}
$$

Choosing an appropriate ε we obtain

$$\mathfrak{a}(t;y,y)+\left[\frac{1}{2}+C(\chi,\|\nabla w\|_{L^\infty(0,T;L^3(\Omega))})+\|g_+\|_{L^\infty(0,T;L^\infty(\Omega))}\right]\|y\|^2_{L^2(\Omega)}\geq\frac{1}{2}\,\|y\|^2_{H^1(\Omega)}\,.$$

Choosing $\varepsilon=1/\left(4\,\chi\,\|\nabla w\|_{L^\infty(0,T;L^4(\Omega))}\right)$ we obtain

$$\mathfrak{a}(t;y,y)\geq\frac{1}{2}\,\|\nabla y\|^2_{L^2(\Omega)}-\left[C(\chi,\|\nabla w\|_{L^\infty(0,T;L^3(\Omega))})+\|g\|_{L^\infty(0,T;L^2(\Omega))}\right]\|y\|^2_{L^2(\Omega)}$$

and consequently

$$\mathfrak{a}(t;y,y)+\left[\frac{1}{2}+C(\chi,\|\nabla w\|_{L^\infty(0,T;L^4(\Omega))})+\|g_+\|_{L^\infty(0,T;L^\infty(\Omega))}\right]\|y\|^2_{L^2(\Omega)}\geq\frac{1}{2}\,\|y\|^2_{H^1(\Omega)}\,.$$

This implies that $\mathfrak{a}$ is quasi-coercive with $\omega:=\frac{1}{2}+C(\chi,\|\nabla w\|_{L^\infty(0,T;L^4(\Omega))})+\|g_+\|_{L^\infty(0,T;L^\infty(\Omega))}$.

Now, we can conclude that $y(x,t)\leq M$, by virtue of proposition 4.1.11 and the remark after that. □

In proving the existence of a weak solution for (4.1.7), we will employ Banach's fixed-point theorem in the space $C([0,T];L^2(\Omega))$ with the norm

$$\|z\|_{C(0,T;L^2(\Omega))}:=\max_{0\leq t\leq T}\|z(t)\|_{L^2(\Omega)}\,.$$

Theorem 4.1.13. *Let $u\in\mathcal{U}_{\mathrm{ad}}$. For any initial data $0\leq y_0\leq M$, there exists some $T>0$ such that (4.1.7) has a unique weak solution $(y,w)\in W(0,T)\times C([0,T];W^{1,6}(\Omega))$.*

Proof. For $T>0$ which will be defined specifically later, we consider

$$X:=\{z\in C([0,T];L^2(\Omega))\,|\,0\leq z\leq M\},$$

with the norm $\|z\|_X:=\|z\|_{C([0,T];L^2(\Omega))}$. To a given function $\tilde{y}\in X$, we define $A\colon X\to X$ by $A\tilde{y}=y$, where y solves the following auxiliary problem in the weak sense:

$$\begin{cases}\partial_t y-\Delta y=-\chi\,\nabla y\cdot\nabla w+y\left(-\chi\,\lambda\,w+\chi\,\tilde{y}+a_0-a_1\,\tilde{y}-a_2\fint_\Omega\tilde{y}\,\mathrm{d}x\right) & \text{in } \Omega_T,\\ -\Delta w+\lambda\,w=\tilde{y} & \text{in } \Omega_T,\\ \dfrac{\partial y}{\partial n}=0\quad\text{and}\quad\dfrac{\partial w}{\partial n}=u & \text{on } \partial\Omega_T,\\ y(x,0)=y_0(x) & \text{in } \Omega.\end{cases}\tag{4.1.16}$$

For future reference, let us define $g:=-\chi\,\lambda\,w+\chi\,\tilde{y}+a_0-a_1\,\tilde{y}-a_2\fint_\Omega\tilde{y}\,\mathrm{d}x$.

Step (1): We prove that A is well-defined. Since $\tilde{y}$ is bounded, it belongs to $L^p(\Omega)$, we can infer from

$$\begin{cases}-\Delta w+\lambda\,w=\tilde{y} & \text{in } \Omega,\\ \dfrac{\partial w}{\partial n}=u & \text{on } \partial\Omega,\end{cases}$$

that $w(t)$ belongs to $W^{1,6}(\Omega)$, by virtue of theorem 4.1.6. First, we observe the regularity of g:

$$\begin{aligned}\|g\|_{L^\infty(0,T;L^2(\Omega))}&=\left\|-\chi\,\lambda\,w+\chi\,\tilde{y}+a_0\,-a_1\,\tilde{y}-a_2\fint_\Omega\tilde{y}\,\mathrm{d}x\right\|_{L^\infty(0,T;L^2(\Omega))}\\&\lesssim\chi\,\lambda\,\|w\|_{L^\infty(0,T;L^3(\Omega))}+|\chi-a_1|\,\|\tilde{y}\|_{L^\infty(0,T;L^\infty(\Omega))}\\&\quad+a_0+\frac{|a_2|}{|\Omega|}\,|\Omega|\,\|\tilde{y}\|_{L^\infty(0,T;L^\infty(\Omega))}\end{aligned}$$

$$\lesssim \chi\,\lambda + a_0 + M\,(|\chi - a_1| + |a_2|)\,.$$

Now, we can apply theorem 4.1.9, resulting in the existence of a unique weak solution $y \in W(0,T)$ to the following problem:

$$\begin{cases} \partial_t y - \Delta y = -\chi\,\nabla w \cdot \nabla y + g\,y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega. \end{cases} \tag{4.1.17}$$

On the other hand, we observe that $w \geq 0$. Indeed, $-\Delta w + \lambda w = \tilde{y}$ and $\tilde{y}$ and u are positive, by virtue of the maximum principle we can deduce the positivity of w. This results in the following estimate

$$\begin{aligned} \|g_+\|_{L^\infty(\Omega_T)} &= \left\| \left(-\chi\,\lambda\,w + \chi\,\tilde{y} + a_0 - a_1\,\tilde{y} - a_2 \fint_\Omega \tilde{y}\,\mathrm{d}x \right)_+ \right\|_{L^\infty(\Omega_T)} \\ &\leq \left\| \chi\tilde{y} + a_0 - a_2 \fint_\Omega \tilde{y}\,\mathrm{d}x \right\|_{L^\infty(\Omega_T)} \\ &\leq \chi\,\|\tilde{y}\|_{L^\infty(\Omega_T)} + a_0 + \frac{|a_2|}{|\Omega|}\,|\Omega|\,\|\tilde{y}\|_{L^\infty(\Omega_T)} \\ &\leq a_0 + M\,(\chi + |a_2|)\,. \end{aligned}$$

Therefore, the solution y of (4.1.17) is positive and bounded above by M, by virtue of theorem 4.1.12. Therefore, A is well-defined.

Step (2): We prove that A is a strict contraction when T is small enough. To this end, let $\tilde{y}_1$, $\tilde{y}_2$ be given in X, where $y_1 = A\,\tilde{y}_1$, $y_2 = A\,\tilde{y}_2$. We have to show $\|y_1 - y_2\|_X \leq \beta\,\|\tilde{y}_1 - \tilde{y}_2\|_X$, for some positive constant $\beta \in [0,1)$ independent of $\tilde{y}_1$ and $\tilde{y}_2$. Let

$$\begin{cases} -\Delta w_i + \lambda w_i = \tilde{y}_i & \text{in } \Omega, \\ \dfrac{\partial w_i}{\partial n} = u & \text{on } \partial\Omega, \end{cases}$$

for $i = 1, 2$. By virtue of theorem 4.1.6, $w_i(t) \in W^{1,6}(\Omega)$ for a.e. $t \in [0,T]$, which means $\|\nabla w_i(t)\|_{L^6(\Omega)} \leq C$ uniformly for all $\tilde{y}_i \in X$, $i = 1, 2$. We set

$$\begin{aligned} &\bar{y} \coloneqq y_1 - y_2,\ \bar{w} \coloneqq w_1 - w_2, \\ &g_i = -\chi\,\lambda\,w_i + \chi\,\tilde{y}_i + a_0 - a_1\,\tilde{y}_i - a_2 \fint_\Omega \tilde{y}_i\,\mathrm{d}x. \end{aligned}$$

This setting results in the following equation:

$$\begin{aligned} \partial_t \bar{y} - \Delta \bar{y} &= -\chi\,(\nabla y_1 \cdot \nabla w_1 - \nabla y_2 \cdot \nabla w_2) + y_1\,g_1 - y_2\,g_2 \\ &= -\chi\,(\nabla \bar{y} \cdot \nabla w_1 + \nabla y_2 \cdot \nabla \bar{w}) + \bar{y}\,g_1 + y_2\,(g_1 - g_2)\,. \end{aligned} \tag{4.1.18}$$

with the boundary condition $\frac{\partial \bar{y}}{\partial n} = 0$ and

$$\begin{cases} -\Delta \bar{w} + \lambda \bar{w} = \tilde{y}_1 - \tilde{y}_2 & \text{in } \Omega, \\ \dfrac{\partial \bar{w}}{\partial n} = 0 & \text{on } \partial\Omega. \end{cases}$$

By virtue of Grisvard, 1985, Theorem 2.2.2.5, $\bar{w}$ belongs to $H^2(\Omega)$.

Similar to g, the regularity of g_i is as follows

$$\begin{aligned}\|g_i\|_{L^2(\Omega)} &= \left\| -\chi\,\lambda\, w_i + \chi\,\tilde{y_i} + a_0 \; - a_1\,\tilde{y_i} - a_2 \int_\Omega \tilde{y_i}\,\mathrm{d}x \right\|_{L^2(\Omega)} \\ &\lesssim \chi\,\lambda\,\|w_i\|_{L^3(\Omega)} + |\chi - a_1|\;\|\tilde{y_i}\|_{L^\infty(\Omega)} + a_0 + |a_2|\;\|\tilde{y_i}\|_{L^\infty(\Omega)} \\ &\lesssim \chi\,\lambda + a_0 + M\,(|\chi - a_1| + |a_2|)\,.\end{aligned}$$

Multiplying (4.1.18) by $\bar{y}$, integrating by part with $\frac{\partial \bar{w}}{\partial n} = 0$, we obtain

$$\begin{aligned}\frac{1}{2}\int_\Omega \partial_t\,|\bar{y}|^2\,\mathrm{d}x + \int_\Omega |\nabla \bar{y}|^2\,\mathrm{d}x &= -\chi \int_\Omega \nabla \bar{y}\cdot\nabla w_1\,\bar{y}\,\mathrm{d}x - \chi\int_\Omega \nabla y_2\cdot\nabla\bar{w}\,\bar{y}\,\mathrm{d}x \\ &+ \int_\Omega g_1\;|\bar{y}|^2\,\mathrm{d}x + \int_\Omega y_2\,(g_1 - g_2)\,\bar{y}\,\mathrm{d}x \\ &= -\chi\int_\Omega \nabla\bar{y}\cdot\nabla w_1\,\bar{y}\,\mathrm{d}x + \chi\int_\Omega y_2\;\mathrm{div}\,(\bar{y}\,\nabla\bar{w})\,\mathrm{d}x \\ &+ \int_\Omega g_1\;|\bar{y}|^2\,\mathrm{d}x + \int_\Omega y_2\,(g_1 - g_2)\,\bar{y}\,\mathrm{d}x \\ &= -\chi\int_\Omega \nabla\bar{y}\cdot\nabla w_1\,\bar{y}\,\mathrm{d}x + \chi\int_\Omega y_2\,(\nabla\bar{y}\cdot\nabla\bar{w} + \bar{y}\,\Delta\bar{w})\,\mathrm{d}x \\ &+ \int_\Omega g_1\;|\bar{y}|^2\,\mathrm{d}x + \int_\Omega y_2\,(g_1 - g_2)\,\bar{y}\,\mathrm{d}x \\ &= -\chi\int_\Omega \nabla\bar{y}\cdot\nabla w_1\,\bar{y}\,\mathrm{d}x + \chi\int_\Omega y_2\Big(\nabla\bar{y}\cdot\nabla\bar{w} + \bar{y}\,(\tilde{y_1} - \tilde{y_2}) + \lambda\,\bar{y}\,\bar{w}\Big)\,\mathrm{d}x \\ &+ \int_\Omega g_1\;|\bar{y}|^2\,\mathrm{d}x + \int_\Omega y_2\,(g_1 - g_2)\,\bar{y}\,\mathrm{d}x.\end{aligned}$$

We observe the following estimate, exploiting interpolation's inequality equation (B.1.2) for $r = 4$, $p = 2$, $q = 6$ and choosing $\theta = \frac{3}{4}$ and Young's inequalities equation (B.1.1)

$$\begin{aligned}-\chi\int_\Omega \nabla\bar{y}\cdot\nabla w_1\,\bar{y}\,\mathrm{d}x &\le \chi\;\|\nabla w_1\|_{L^4(\Omega)}\;\|\nabla\bar{y}\|_{L^2(\Omega)}\;\|\bar{y}\|_{L^4(\Omega)} \\ &\le \chi\;\|\nabla w_1\|_{L^4(\Omega)}\;\|\nabla\bar{y}\|_{L^2(\Omega)}\,\|\bar{y}\|^{1/4}_{L^2(\Omega)}\;\|\bar{y}\|^{3/4}_{L^6(\Omega)} \\ &\lesssim \chi\;\|\nabla w_1\|_{L^4(\Omega)}\;\|\nabla\bar{y}\|_{L^2(\Omega)}\,\|\bar{y}\|^{1/4}_{L^2(\Omega)}\left[\|\bar{y}\|^{3/4}_{L^2(\Omega)} + \|\nabla\bar{y}\|^{3/4}_{L^2(\Omega)}\right] \\ &\lesssim \chi\;\|\nabla w_1\|_{L^4(\Omega)}\left[\varepsilon\,\|\nabla\bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon)\,\|\bar{y}\|^2_{L^2(\Omega)}\right] \\ &+ \chi\;\|\nabla w_1\|_{L^4(\Omega)}\left[\varepsilon\left(\|\nabla\bar{y}\|^{7/4}_{L^2(\Omega)}\right)^{8/7} + C(\varepsilon)\left(\|\bar{y}\|^{1/4}_{L^2(\Omega)}\right)^8\right] \\ &\lesssim 2\,\chi\;\|\nabla w_1\|_{L^4(\Omega)}\left[\varepsilon\,\|\nabla\bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon)\,\|\bar{y}\|^2_{L^2(\Omega)}\right]\end{aligned}$$

and

$$\begin{aligned}\int_\Omega g_1\;|\bar{y}|^2\,\mathrm{d}x &\le \|g_1\|_{L^2(\Omega)}\;\|\bar{y}\|^2_{L^4(\Omega)} \le \|g_1\|_{L^2(\Omega)}\;\|\bar{y}\|^{1/2}_{L^2(\Omega)}\;\|\bar{y}\|^{3/2}_{L^6(\Omega)} \\ &\lesssim \|g_1\|_{L^2(\Omega)}\left[\varepsilon\left(\|\bar{y}\|^{3/2}_{H^1(\Omega)}\right)^{4/3} + C(\varepsilon)\left(\|\bar{y}\|^{1/2}_{L^2(\Omega)}\right)^4\right] \\ &\lesssim \|g_1\|_{L^2(\Omega)}\left[\varepsilon\;\|\nabla\bar{y}\|^2_{L^2(\Omega)} + (\varepsilon + C(\varepsilon))\,\|\bar{y}\|^2_{L^2(\Omega)}\right].\end{aligned}$$

Employing Young's inequality for all other norm products results in

$$\begin{aligned}
\frac{1}{2}\int_\Omega \partial_t |\bar{y}|^2 \,\mathrm{d}x + \int_\Omega |\nabla \bar{y}|^2 \,\mathrm{d}x &\le \chi \, \|\bar{y}\|_{L^4(\Omega)} \, \|\nabla \bar{y}\|_{L^2(\Omega)} \, \|\nabla w_1\|_{L^4(\Omega)} \\
&+ \chi \, \|y_2\|_{L^\infty(\Omega)} \Big(\|\nabla \bar{w}\|_{L^2(\Omega)} \, \|\nabla \bar{y}\|_{L^2(\Omega)} + \|\bar{y}\|_{L^2(\Omega)} \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} + \lambda \, \|\bar{y}\|_{L^2(\Omega)} \, \|\bar{w}\|_{L^2(\Omega)} \Big) \\
&+ \|g_1\|_{L^2(\Omega)} \, \|\bar{y}\|^2_{L^4(\Omega)} + \|y_2\|_{L^\infty(\Omega)} \, \|g_1 - g_2\|_{L^2(\Omega)} \, \|\bar{y}\|_{L^2(\Omega)} \\
&\lesssim 2\,\chi \, \|\nabla w_1\|_{L^4(\Omega)} \Big[\varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|\bar{y}\|^2_{L^2(\Omega)} \Big] \\
&+ \chi \, M \Big(\frac{1}{4\,\varepsilon} \, \|\nabla \bar{w}\|^2_{L^2(\Omega)} + \varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + \|\bar{y}\|^2_{L^2(\Omega)} \\
&+ \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} + \lambda \big(\|\bar{y}\|^2_{L^2(\Omega)} + \|\bar{w}\|^2_{L^2(\Omega)} \big) \Big) \\
&+ \|g_1\|_{L^2(\Omega)} \, \Big[\varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + (\varepsilon + C(\varepsilon)) \, \|\bar{y}\|^2_{L^2(\Omega)} \Big] + M \, \Big(\|g_1 - g_2\|^2_{L^2(\Omega)} + \|\bar{y}\|^2_{L^2(\Omega)} \Big)
\end{aligned}$$

for a.a. $t \in [0,T]$. Choosing some appropriate $\varepsilon > 0$, we find

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega |\bar{y}|^2 \,\mathrm{d}x \le c \, \Big(\|\bar{y}\|^2_{L^2(\Omega)} + \|\nabla \bar{w}\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} + \|\bar{w}\|^2_{L^2(\Omega)} + \|g_1 - g_2\|^2_{L^2(\Omega)} \Big), \tag{4.1.19}$$

where c depends on M, Ω, a_0, a_1, a_2, χ and λ.
Since $-\Delta \bar{w} + \lambda \, \bar{w} = \tilde{y}_1 - \tilde{y}_2$ holds with $\frac{\partial \bar{w}}{\partial n} = 0$, we also have

$$\|\bar{w}\|^2_{H^1(\Omega)} = \|\bar{w}\|^2_{L^2(\Omega)} + \|\nabla \bar{w}\|^2_{L^2(\Omega)} \le c_1 \, \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} .$$

On the other hand and

$$g_1 - g_2 = -\chi \, \lambda \, (w_1 - w_2) + \chi \, (\tilde{y}_1 - \tilde{y}_2) - a_1 \, (\tilde{y}_1 - \tilde{y}_2) - a_2 \fint_\Omega (\tilde{y}_1 - \tilde{y}_2) \,\mathrm{d}x.$$

Hence, it follows

$$\begin{aligned}
\|g_1 - g_2\|_{L^2(\Omega)} &\le \chi \, \lambda \, \|\bar{w}\|_{L^2(\Omega)} + |\chi - a_1| \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} + \frac{|a_2|}{|\Omega|} \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^1(\Omega)} \\
&\le c_2 \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} .
\end{aligned}$$

Now, we can deduce from (4.1.19) that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega |\bar{y}|^2 \,\mathrm{d}x \le C \, \Big(\|\bar{y}\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} \Big) .$$

This implies

$$\frac{\mathrm{d}}{\mathrm{d}t} \Big(\exp(-C\,t) \, \|\bar{y}(t)\|^2_{L^2(\Omega)} \Big) \le C \, \exp(-C\,t) \, \|(\tilde{y}_1 - \tilde{y}_2)\,(t)\|^2_{L^2(\Omega)} .$$

Hence, we find

$$\begin{aligned}
\exp(-C\,t) \, \|\bar{y}(t)\|^2_{L^2(\Omega)} &\le C \int_0^t \exp(-C\,s) \, \|(\tilde{y}_1 - \tilde{y}_2)\,(s)\|^2_{L^2(\Omega)} \,\mathrm{d}s \\
&\le (1 - \exp(-C\,t)) \, \|\tilde{y}_1 - \tilde{y}_2\|^2_X ,
\end{aligned}$$

for each $0 \le t \le T$, which results in

$$\max_{0 \le t \le T} \|\bar{y}(t)\|^2_{L^2(\Omega)} = \|\bar{y}\|^2_X \le (\exp(C\,T) - 1) \, \|\tilde{y}_1 - \tilde{y}_2\|^2_X .$$

Consequently, A is a strict contraction, provided $T > 0$ is sufficiently small, so that $\exp(C\,T) < 2$.

Now, we can apply Banach's fixed point theorem to find a weak solution (y, w) of the problem (4.1.7) existing on the time interval $[0, T]$.

To show the uniqueness of the solution, consider two solutions y_1 and y_2 of problem (4.1.7). From existence proof we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\|A\,\tilde{y}_1 - A\,\tilde{y}_2\|^2_{L^2(\Omega)} = \frac{\mathrm{d}}{\mathrm{d}t}\,\|y_1 - y_2\|^2_{L^2(\Omega)} \le C\left(\|y_1 - y_2\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)}\right).$$

Since y_1 and y_2 are fixed points, we have $y_1 = \tilde{y}_1$, $y_2 = \tilde{y}_2$, respectively. Consequently,

$$\|y_1 - y_2\|^2_{L^2(\Omega)} \le 2\,C \int_0^t \|y_1(s) - y_2(s)\|^2\,\mathrm{d}s,$$

for $0 \le s \le T$. According to Gronwall's inequality, $y_1 = y_2$. □

4.1.2 Existence of an Optimal Solution

In the following, we will prove that the function $y \in W(0, T)$, as a solution of (4.1.7), is bounded above by the function $\Upsilon(t)$, which satisfies the following logistic ODE

$$\begin{cases} \dot{\Upsilon} = \Upsilon\big(a_0 + (\chi - a_1 + [a_2]_-)\,\Upsilon\big), \\ \Upsilon(0) = \|y_0\|_{L^\infty(\Omega)}\,. \end{cases} \tag{4.1.20}$$

The solution of this ODE will be computed as follows.

The ODE (4.1.20) is of the form

$$\frac{\mathrm{d}\Upsilon}{\mathrm{d}t} = \Upsilon\,(a_0 + C_0\,\Upsilon)\,,$$

where $C_0 = \chi - a_1 + [a_2]_-$. This implies

$$\left(\frac{1/a_0}{\Upsilon} - \frac{C_0/a_0}{a_0 + C_0\,\Upsilon}\right)\mathrm{d}\Upsilon = \mathrm{d}t$$

Multiplying this equation by a_0 and integrating from 0 to t, results in

$$\ln\frac{\Upsilon}{a_0 + C_0\,\Upsilon} = a_0 t + K_0,$$

and consequently

$$\Upsilon(t) = \frac{a_0}{K\,\mathrm{e}^{-a_0\,t} - C_0} \quad \text{for some constant } K.$$

Exploiting the initial condition $\Upsilon(0) = \|y_0\|_{L^\infty(\Omega)}$, K amounts to

$$K = C_0 + \frac{a_0}{\|y_0\|_{L^\infty(\Omega)}}.$$

Remark 4.1.14. *The solution of the ODE (4.1.20) is given by*

$$\Upsilon(t) = \frac{a_0}{K\,\mathrm{e}^{-a_0\,t} - C_0}, \quad \textit{where } K = C_0 + \frac{a_0}{\|y_0\|_{L^\infty(\Omega)}}.$$

When $C_0 < 0$ holds, then the solution Υ is bounded, so its maximal time of existence is $T_* = \infty$. On the other hand, when $C_0 \ge 0$ holds (in particular, when $-a_1 + [a_2]_- \ge 0$), there exists a maximal existence time $T_* \in (0, \infty)$ such that $\Upsilon(t)$ satisfies

$$\lim_{t \nearrow T_*} \Upsilon(t) = \infty.$$

This shows $T_* = -\frac{1}{a_0}\ln\frac{C_0}{K} > 0$ is finite in this case.

In the following, we will refer to $T_* \in (0, \infty]$ as defined above.

Theorem 4.1.15. *Assume that $y \in W(0,T)$ is a solution of (4.1.7) and $\frac{\chi}{2} - a_1 + [a_2]_- < 0$. Then y is uniformly bounded by $\Upsilon(t)$,*

$$\|y(t)\|_{L^\infty(\Omega)} \le \Upsilon(t) \quad \text{for } t \in [0, T_*),$$

where $\Upsilon(t)$ satisfies the logistic ODE (4.1.20).

Proof. First, we consider the ODE (4.1.20). Setting

$$z(x,t) \coloneqq y(x,t) - \Upsilon(t),$$

we obtain

$$\begin{aligned}
\partial_t z - \Delta z &= -\chi \nabla y \cdot \nabla w + y \left(-\chi \lambda w + \chi y + a_0 - a_1 y - a_2 \fint_\Omega y \,\mathrm{d}x \right) \\
&\quad - \Upsilon \Big(a_0 + (\chi - a_1 + [a_2]_-) \Upsilon \Big) \\
&= -\chi \nabla z \cdot \nabla w - \chi \lambda w y + a_0 (y - \Upsilon) + (\chi - a_1)(y + \Upsilon) z - a_2 y \fint_\Omega y \,\mathrm{d}x - [a_2]_- \Upsilon^2 \\
&= -\chi \nabla z \cdot \nabla w - \chi \lambda w (z + \Upsilon) + a_0 z + (\chi - a_1)(z + 2\Upsilon) z \\
&\quad - a_2 (z + \Upsilon) \fint_\Omega (z + \Upsilon) \,\mathrm{d}x - [a_2]_- \Upsilon^2.
\end{aligned}$$

Multiplying this equation with z_+, integrating by parts yields

$$\begin{aligned}
\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega z_+^2 \,\mathrm{d}x + \int_\Omega |\nabla z_+|^2 \,\mathrm{d}x &\le -\frac{\chi}{2} \int_\Omega \nabla w \cdot \nabla (z_+^2) - \chi \lambda \int_\Omega w (z + \Upsilon) z_+ \,\mathrm{d}x \\
&\quad + a_0 \int_\Omega z_+^2 \,\mathrm{d}x + (\chi - a_1) \int_\Omega (z + 2\Upsilon) z_+^2 \,\mathrm{d}x \\
&\quad + [a_2]_- \Big(\fint_\Omega (z + \Upsilon) \,\mathrm{d}x \Big) \int_\Omega (z + \Upsilon) z_+ \,\mathrm{d}x - [a_2]_- \Upsilon^2 \int_\Omega z_+ \,\mathrm{d}x \\
&\le \frac{\chi}{2} \int_\Omega \big(\lambda w - z - \Upsilon \big) z_+^2 \,\mathrm{d}x - \int_{\partial\Omega} u z_+^2 \,\mathrm{d}x - \chi \lambda \int_\Omega w z_+^2 \,\mathrm{d}x \\
&\quad + a_0 \int_\Omega z_+^2 \,\mathrm{d}x + (\chi - a_1) \int_\Omega z_+^3 \,\mathrm{d}x + \Upsilon \Big(2(\chi - a_1) + [a_2]_- \Big) \int_\Omega z_+^2 \,\mathrm{d}x \\
&\quad + [a_2]_- \fint_\Omega z_+ \,\mathrm{d}x \int_\Omega z_+^2 \,\mathrm{d}x + [a_2]_- \Upsilon \fint_\Omega z_+ \,\mathrm{d}x \int_\Omega z_+ \,\mathrm{d}x \\
&\le -\lambda \frac{\chi}{2} \int_\Omega w z_+^2 \,\mathrm{d}x + a_0 \int_\Omega z_+^2 \,\mathrm{d}x \\
&\quad + \Upsilon \Big(2(\chi - a_1) + [a_2]_- - \frac{\chi}{2} \Big) \int_\Omega z_+^2 \,\mathrm{d}x - \Big(a_1 - \frac{\chi}{2} \Big) \int_\Omega z_+^3 \,\mathrm{d}x \\
&\quad + [a_2]_- \fint_\Omega z_+ \,\mathrm{d}x \int_\Omega z_+^2 \,\mathrm{d}x + [a_2]_- \Upsilon \fint_\Omega z_+ \,\mathrm{d}x \int_\Omega z_+ \,\mathrm{d}x.
\end{aligned}$$

By means of Hölder's inequality for the terms $\int_\Omega z_+ \,\mathrm{d}x$ and $\int_\Omega z_+^2 \,\mathrm{d}x$, namely

$$\begin{aligned}
\int_\Omega z_+ \,\mathrm{d}x &\le |\Omega|^{\frac{2}{3}} \left(\int_\Omega |z_+|^3 \,\mathrm{d}x \right)^{\frac{1}{3}}, \\
\int_\Omega z_+^2 \,\mathrm{d}x &\le |\Omega|^{\frac{1}{3}} \left(\int_\Omega |z_+|^3 \,\mathrm{d}x \right)^{\frac{2}{3}}, \\
\int_\Omega z_+ \,\mathrm{d}x &\le |\Omega|^{\frac{1}{2}} \left(\int_\Omega |z_+|^2 \,\mathrm{d}x \right)^{\frac{1}{2}},
\end{aligned}$$

and the negativity of $\frac{\chi}{2} - a_1 + [a_2]_-$ and exploiting the positivity of w, we obtain

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega z_+^2\,\mathrm{d}x + \int_\Omega |\nabla z_+|^2\,\mathrm{d}x &\le a_0\int_\Omega z_+^2\,\mathrm{d}x - \left(a_1 - \frac{\chi}{2}\right)\int_\Omega z_+^3\,\mathrm{d}x \\
&+ \Upsilon\left(2\,(\chi - a_1) + [a_2]_- - \frac{\chi}{2}\right)\int_\Omega z_+^2\,\mathrm{d}x + \frac{[a_2]_-}{|\Omega|}\,|\Omega|\int_\Omega z_+^3\,\mathrm{d}x + \frac{[a_2]_-}{|\Omega|}\,|\Omega|\,\Upsilon\int_\Omega z_+^2\,\mathrm{d}x \\
&= a_0\int_\Omega z_+^2\,\mathrm{d}x + \left(\frac{\chi}{2} - a_1 + [a_2]_-\right)\int_\Omega z_+^3\,\mathrm{d}x + 2\,\Upsilon\left(\frac{3\,\chi}{4} - a_1 + [a_2]_-\right)\int_\Omega z_+^2\,\mathrm{d}x \\
&\le a_0\int_\Omega z_+^2\,\mathrm{d}x + 2\,\chi\,\Upsilon\int_\Omega z_+^2\,\mathrm{d}x.
\end{aligned}$$

Finally, the estimate above attains the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega z_+^2\,\mathrm{d}x \le 2\,a_0\int_\Omega z_+^2\,\mathrm{d}x + 4\,\chi\,\Upsilon\int_\Omega z_+^2\,\mathrm{d}x,$$

and consequently,

$$\int_\Omega z_+^2(t)\,\mathrm{d}x \le \exp\left[2\,a_0\,t + 4\,\chi\int_0^t \Upsilon(s)\,\mathrm{d}s\right]\int_\Omega z_+^2(0)\,\mathrm{d}x.$$

Since $z_+(0) = 0$ by virtue of Gronwall's inequality we obtain $z_+(t) = 0$, which results in $y(x,t) \le \Upsilon(t)$. □

Remark 4.1.16. *We note that in ODE (4.1.20) the boundedness of Υ is just guaranteed by the negativity of $C_0 = \chi - a_1 + [a_2]_-$. In other words if $C_0 \ge 0$, in particular when (4.1.3) fails, the L^∞-Norm of the solution of PDE (4.1.7) may blow up.*

Remark 4.1.17. *We observe that by virtue of theorem 4.1.13 and theorem 4.1.15, there exists a maximal existence time $T_{\max} \in (0,\infty]$ and a unique weak solution $(y,w) \in W(0,T_{\max}) \times L^\infty(0,T_{\max};W^{1,6}(\Omega))$ of problem (4.1.7). Notice that $T_{\max} \ge T_*$, where T_* is the maximal existence time of ODE (4.1.20).*

In the following we prove the existence of the solution to the optimal control problem (4.1.1). For this we require the following a-priori estimate.

Theorem 4.1.18. *Assume (y,w) be a weak solution of problem (4.1.7) and $T < T_*$. If $u \in \mathcal{U}_{\mathrm{ad}}$, then the following estimate holds:*

$$\operatorname*{ess\,sup}_{0\le t\le T} \|\nabla w\|_{L^6(\Omega)} + \|y\|_{L^\infty(\Omega_T)} + \|y\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t y\|_{L^2(0,T;H^1(\Omega)^*)} \le C_{T_*},$$

where C_{T_*} is a positive constant which depends on u, χ, T and $\|y_0\|_{L^\infty(\Omega)}$.

Proof. The boundedness of y in $L^\infty(0,T;L^\infty(\Omega))$ is a straightforward result from theorem 4.1.15, namely,

$$\|y\|_{L^\infty(0,T;L^\infty(\Omega))} \le \|\Upsilon\|_{L^\infty(0,T)}.$$

According to theorem 4.1.15, y is bounded, it belongs then to $L^p(\Omega)$. The regularity of w follows evidently from $-\Delta w + \lambda\,w = y$ with $\frac{\partial w}{\partial n} = u$, by virtue of theorem 4.1.6. We have

$$\|w\|_{L^\infty(0,T;W^{1,6}(\Omega))} \le C\left(\|y\|_{L^\infty(0,T;L^6(\Omega))} + \|u\|_{L^\infty(0,T;L^6(\partial\Omega))}\right).$$

In the second step we recall the first equation of (4.1.7) with setting $g := -\chi\,\lambda w + \chi\,y + a_0 - a_1\,y - a_2\fint_\Omega y\,\mathrm{d}x$, that is

$$\partial_t y - \Delta y = -\chi\nabla y\cdot\nabla w + y\,g. \tag{4.1.21}$$

We observe the following estimate for g

$$\begin{aligned}\|g\|_{L^\infty(0,T;L^2(\Omega))} &= \left\| -\chi\,\lambda\,w + \chi\,y + a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x \right\|_{L^\infty(0,T;L^2(\Omega))} \\ &\lesssim \chi\,\lambda\,\|w\|_{L^\infty(0,T;L^2(\Omega))} + |\chi - a_1|\,\|y\|_{L^\infty(0,T;L^2(\Omega))} + a_0 + |a_2|\,\|y\|_{L^\infty(0,T;L^\infty(\Omega)} \\ &\lesssim \chi\,\lambda\,\|w\|_{L^\infty(0,T;L^2(\Omega))} + a_0 + (|\chi - a_1| + |a_2|)\,\|\Upsilon\|_{L^\infty(0,T)} \le C_1.\end{aligned}$$

Multiplying equation (4.1.21) by y and integrating by parts and using Hölder's and interpolation inequalities, we obtain

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |y|^2\,\mathrm{d}x + \int_\Omega |\nabla y|^2\,\mathrm{d}x &= -\chi\int_\Omega \nabla y\cdot\nabla w\,y\,\mathrm{d}x + \int_\Omega g\,|y|^2\,\mathrm{d}x \\ &\le \chi\,\|\nabla w\|_{L^4(\Omega)}\,\|\nabla y\|_{L^2(\Omega)}\,\|y\|_{L^4(\Omega)} + \|g\|_{L^2(\Omega)}\,\|y\|^2_{L^4(\Omega)} \\ &\le \chi\,\|\nabla w\|_{L^4(\Omega)}\,\|\nabla y\|_{L^2(\Omega)}\,\|y\|^{1/4}_{L^2(\Omega)}\,\|y\|^{3/4}_{L^6(\Omega)} + \|g\|_{L^2(\Omega)}\left(\|y\|^{1/4}_{L^2(\Omega)}\,\|y\|^{3/4}_{L^6(\Omega)}\right)^2 \\ &\lesssim \chi\,\|\nabla w\|_{L^4(\Omega)}\,\|\nabla y\|_{L^2(\Omega)}\,\|y\|^{1/4}_{L^2(\Omega)}\left[\|y\|^{3/4}_{L^2(\Omega)} + \|\nabla y\|^{3/4}_{L^2(\Omega)}\right] \\ &\quad + \|g\|_{L^2(\Omega)}\,\|y\|^{1/2}_{L^2(\Omega)}\,\|y\|^{3/2}_{L^6(\Omega)} \\ &\lesssim \chi\,\|\nabla w\|_{L^4(\Omega)}\left[\varepsilon\,\|\nabla y\|^2_{L^2(\Omega)} + C(\varepsilon)\,\|y\|^2_{L^2(\Omega)}\right] \\ &\quad + \chi\,\|\nabla w\|_{L^4(\Omega)}\left[\varepsilon\left(\|\nabla y\|^{7/4}_{L^2(\Omega)}\right)^{8/7} + C(\varepsilon)\left(\|y\|^{1/4}_{L^2(\Omega)}\right)^8\right] \\ &\quad + \|g\|_{L^2(\Omega)}\left[\varepsilon\left(\|y\|^{3/2}_{H^1(\Omega)}\right)^{4/3} + C(\varepsilon)\left(\|y\|^{1/2}_{L^2(\Omega)}\right)^4\right] \\ &\lesssim C_2\left(\varepsilon\,\|\nabla y\|^2_{L^2(\Omega)} + C(\varepsilon)\,\|y\|^2_{L^2(\Omega)}\right) + C_3\left[\varepsilon\,\|\nabla y\|^2_{L^2(\Omega)} + (\varepsilon + C(\varepsilon))\,\|y\|^2_{L^2(\Omega)}\right].\end{aligned}$$

Choosing some appropriate ε and small enough, we find

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y\|^2_{L^2(\Omega)} + \|\nabla y\|^2_{L^2(\Omega)} \le \frac{1}{2}\|\nabla y\|^2_{L^2(\Omega)} + C_4\,\|y\|^2_{L^2(\Omega)},$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\|y\|^2_{L^2(\Omega)} + \|\nabla y\|^2_{L^2(\Omega)} \le 2\,C_4\,\|y\|^2_{L^2(\Omega)}.$$

We integrate this inequality from 0 to T to find

$$\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\,\|y(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \int_0^T \|\nabla y(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t \le 2\,C_4\int_0^T \|y(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t,$$

and consequently

$$\|y(T)\|^2_{L^2(\Omega)} - \|y_0\|^2_{L^2(\Omega)} + \int_0^T \|\nabla y(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t \le 2\,C_4\int_0^T \|y(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t.$$

In summary, we can infer

$$\begin{aligned}\|\nabla y\|^2_{L^2(0,T;L^2(\Omega))} &\le 2\,C_4\,\|y\|^2_{L^2(0,T;L^2(\Omega))} + \|y_0\|^2_{L^2(\Omega)} \\ &\le 2\,C_4\,T\,\|y\|^2_{L^\infty(0,T;L^2(\Omega))} + \|y_0\|^2_{L^2(\Omega)} \\ &\le 2\,C_4\,T\,\|\Upsilon\|^2_{L^\infty(0,T)} + \|y_0\|^2_{L^2(\Omega)} \le C_5.\end{aligned}$$

This estimate and taking the boundedness of y in Ω_T into account result in its boundedness in $L^2(0,T;H^1(\Omega))$.

Finally, we multiply (4.1.21) by $\varphi \in L^2(0,T;H^1(\Omega))$ and integrate over $\Omega \times (0,T)$

$$\int_0^T \int_\Omega \partial_t y\, \varphi \,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \int_\Omega \nabla y \cdot \nabla \varphi \,\mathrm{d}x\,\mathrm{d}t - \chi \int_0^T \int_\Omega \nabla y \cdot \nabla w\, \varphi \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega g\, y\, \varphi \,\mathrm{d}x\,\mathrm{d}t.$$

That implies

$$\begin{aligned}\left| \int_0^T \int_\Omega \partial_t y\, \varphi \,\mathrm{d}x\,\mathrm{d}t \right| &\le \|\nabla y\|_{L^2(0,T;L^2(\Omega))}\, \|\nabla \varphi\|_{L^2(0,T;L^2(\Omega))} \\ &\quad + \|\nabla w\|_{L^\infty(0,T;L^3(\Omega))}\, \|\nabla y\|_{L^2(0,T;L^2(\Omega))}\, \|\varphi\|_{L^2(0,T;L^6(\Omega))} \\ &\quad + \|g\|_{L^\infty(0,T;L^2(\Omega))}\, \|y\|_{L^2(0,T;L^3(\Omega))}\, \|\varphi\|_{L^2(0,T;L^6(\Omega))} \\ &\lesssim c_1\, \|\varphi\|_{L^2(0,T;H^1(\Omega))} + c_2\, \|\varphi\|_{L^2(0,T;H^1(\Omega))} + c_3\, \|\varphi\|_{L^2(0,T;H^1(\Omega))} \\ &\lesssim c\, \|\varphi\|_{L^2(0,T;H^1(\Omega))}\,.\end{aligned}$$

That means

$$\|\partial_t y\|_{L^2(0,T;H^1(\Omega)^*)} \le C.$$

This complete the proof. □

Theorem 4.1.19. *Suppose $T < T_*$. Then the optimal control problem (4.1.1) has at least one optimal solution $\bar{u} \in \mathcal{U}_{\mathrm{ad}}$ with associated optimal state $(\bar{y}, \bar{w}) \in W(0,T) \times L^\infty(0,T;W^{1,6}(\Omega))$.*

Proof. We construct the tentative minimizer $(\bar{y}, \bar{u})$. Since J is bounded from below on $W(0,T) \times \mathcal{U}_{\mathrm{ad}}$, therefore the infimum

$$\beta := \inf_{u \in \mathcal{U}_{\mathrm{ad}}} J(y,u)$$

is finite. Let $\{(y_n, u_n)\} \subset W(0,T) \times \mathcal{U}_{\mathrm{ad}}$ be a minimizing sequence, that means

$$\lim_{n\to\infty} J(y_n, u_n) = \beta.$$

$\mathcal{U}_{\mathrm{ad}}$ is a bounded subset of $L^\infty(\partial\Omega_T)$ and thus bounded in $L^2(\partial\Omega_T)$. $L^2(\partial\Omega_T)$ is a reflexive Banach space, consequently there exists a subsequence, without loss of generality $\{u_n\}$ itself, such that

$$u_n \rightharpoonup \bar{u} \quad \text{in } L^2(\partial\Omega_T).$$

$\mathcal{U}_{\mathrm{ad}}$ is closed and convex in $L^2(\partial\Omega_T)$, then it is weakly closed in $L^2(\partial\Omega_T)$, consequently $\bar{u} \in \mathcal{U}_{\mathrm{ad}}$.

From theorem 4.1.18 we have the following estimate

$$\|y_n\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t y_n\|_{L^2(0,T;H^1(\Omega)^*)} \le C.$$

From the boundedness of $\{y_n\}$ and $\{\partial_t y_n\}$ in the corresponding reflexive Banach spaces follows the existence of subsequences, denoted by the same indices, such that

$$\begin{aligned} y_n &\rightharpoonup \bar{y} \quad \text{in } L^2(0,T;H^1(\Omega)), \\ \partial_t y_n &\rightharpoonup \partial_t \bar{y} \quad \text{in } L^2(0,T;H^1(\Omega)^*). \end{aligned}$$

Since $W(0,T) \hookrightarrow L^2(\Omega_T)$ is compact, see e. g. Temam, 1984, Chapter 3, Section 2.2, Theorem 2.1, then

$$y_n \to \bar{y} \quad \text{in } L^2(\Omega_T).$$

On the other hand, $W(0,T)$ is compactly embedded in $L^2(0,T;H^s(\Omega))$. Indeed, $H^1(\Omega) \hookrightarrow H^s(\Omega) \hookrightarrow H^1(\Omega)^*$, with $s < 1$ and $H^1(\Omega) \hookrightarrow H^s(\Omega)$ is a compact embedding by virtue of theorem 2.1.19. Hence, we can apply the result in Simon, 1986, Theorem 5 to imply that $W(0,T)$ is compactly embedded in $L^2(0,T;H^s(\Omega))$. Consequently, for $s > \frac{1}{2}$, the trace

operator $W(0,T) \to L^2(0,T;L^2(\partial\Omega))$, which exists by virtue of theorem 2.1.23, is compact. This results in

$$y_n \to \bar{y} \quad \text{in } L^2(0,T;L^2(\partial\Omega)).$$

Moreover, $y_n \rightharpoonup \bar{y}$ in $L^2(0,T;H^1(\Omega))$ results in

$$\nabla y_n \rightharpoonup \nabla \bar{y} \quad \text{in } L^2(0,T;L^2(\Omega)).$$

In the same way, we have by virtue of theorem 4.1.18 $\|w_n\|_{L^2(0,T;W^{1,6}(\Omega))} \leq C$. Since $L^2(0,T;W^{1,6})$ is reflexive, there exist a subsequence, denoted by the same indices, satisfying $w_n \rightharpoonup \bar{w}$ in $L^2(0,T;W^{1,6}(\Omega))$. This means

$$\nabla w_n \rightharpoonup \nabla \bar{w} \quad \text{in } L^2(0,T;L^6(\Omega)).$$

Now, we recall definition of a weak solution for problem (4.1.1) but in the original form (4.1.2):

$$\begin{aligned}\int_\Omega \partial_t y_n\, \varphi \,\mathrm{d}x + \int_\Omega \nabla y_n \cdot \nabla \varphi \,\mathrm{d}x &= \chi \int_\Omega y_n \nabla w_n \cdot \nabla \varphi \,\mathrm{d}x - \chi \int_{\partial\Omega} u_n\, y_n\, \varphi \,\mathrm{d}s \\ &\quad + \int_\Omega \left(a_0 - a_1\, y_n - a_2 \fint_\Omega y_n \,\mathrm{d}x \right) y_n\, \varphi \,\mathrm{d}x \quad \text{for all } \varphi \in C^\infty(\Omega),\end{aligned}$$

and

$$\int_\Omega \nabla w_n \cdot \nabla \varphi \,\mathrm{d}x - \int_{\partial\Omega} u_n\, \varphi \,\mathrm{d}s + \lambda \int_\Omega w_n\, \varphi \,\mathrm{d}x = \int_\Omega y_n\, \varphi \,\mathrm{d}x \quad \text{for all } \varphi \in C^\infty(\Omega).$$

Passing to the limit in each term and employing above convergence results we can conclude that $(\bar{y}, \bar{w})$ is a weak solution of problem (4.1.1) associated with the control function $\bar{u}$.

In summary, by exploiting the weak sequential lower semicontinuity of the cost functional we have

$$J(\bar{y}, \bar{u}) \leq \lim_{n\to\infty} \inf J(y_n, u_n) = \lim_{n\to\infty} J(y_n, u_n) = \beta.$$

By definition of β, we infer that $\beta = J(\bar{y}, \bar{u})$. □

4.2 Optimality System

In this section, we derive the optimality system of the chemotaxis system. We first discuss Fréchet differentiabilty of the linearized problem (4.2.1) to the chemotaxis system in subsection 4.2.1 and then prove the Fréchet differentiabilty of the control-to-state map. Subsequently, we apply the Lagrangian method to derive a system of necessary optimality system in subsection 4.2.2.

4.2.1 Differentiability of Control-to-State Map

In this section we show the Fréchet differentiability of the control-to-state map S by means of the implicit function theorem. The control-to-state operator

$$S\colon \mathcal{U}_{\mathrm{ad}} \to W(0,T) \times L^\infty(0,T;W^{1,6}(\Omega))$$

is well-defined according to theorem 4.1.13. To verify the assumption of the implicit function theorem, we need the following result about the linearization of the chemotaxis equation (4.1.1b).

Form now on, we assume that $T < T_*$ to grantee the boundedness of Υ and employ the estimate in theorem 4.1.18.

Proposition 4.2.1. *Suppose $\hat{u} \in \mathcal{U}_{\text{ad}}$ and $(\hat{y}, \hat{w}) \in W(0,T) \times L^\infty(0,T; W^{1,6}(\Omega))$ to be the associated unique solution of the chemotaxis equation (4.1.1b). Then, for any $L^2(\partial\Omega_T)$ and $h = (h_1, h_2, y_0) \in Z := L^2(0,T; H^1(\Omega)^*) \times C([0,T]; W^{1,6/5}(\Omega)^*) \times L^2(\Omega)$, the linearized problem*

$$\begin{cases} \partial_t y - \Delta y = -\chi \operatorname{div}(y \nabla \hat{w} + \hat{y} \nabla w) + \left(a_0 - a_1 \hat{y} - a_2 \fint_\Omega \hat{y} \,\mathrm{d}x\right) y \\ \qquad\qquad - \left(a_1 y + a_2 \fint_\Omega y \,\mathrm{d}x\right) \hat{y} + h_1, & \text{in } \Omega_T, \\ -\Delta w + \lambda w = y + h_2 & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0 & \text{in } \Omega, \end{cases} \tag{4.2.1}$$

has a unique solution $(y, w) \in Y := W(0,T) \times L^\infty(0,T; W^{1,6}(\Omega))$.

Proof. Inserting $-\Delta\hat{w} + \lambda\hat{w} = \hat{y}$ and the second equation of (4.2.1) in the first one, we obtain

$$\begin{aligned} \partial_t y - \Delta y = &-\chi\Big(\nabla y \cdot \nabla\hat{w} + \nabla\hat{y} \cdot \nabla w\Big) + \Big(-\chi\lambda\hat{w} + a_0 - a_1\hat{y} - a_2 \fint_\Omega \hat{y}\,\mathrm{d}x\Big) y \\ &+ \Big(-\chi\lambda w + 2\chi y - a_1 y - a_2 \fint_\Omega y\,\mathrm{d}x\Big)\hat{y} + \chi h_2 \hat{y} + h_1. \end{aligned}$$

We will apply Banach's fixed point theorem on the space $X = C([0,T]; L^2(\Omega))$. Let define $A\colon X \to X$ by setting $A\tilde{y} = y$, where y solves

$$\begin{aligned} \partial_t y - \Delta y = &-\chi(\nabla y \cdot \nabla\hat{w} + \nabla\hat{y} \cdot \nabla w) + \left(-\chi\lambda\hat{w} + a_0 - a_1\hat{y} - a_2 \fint_\Omega \hat{y}\,\mathrm{d}x\right) y \\ &+ \left(-\chi\lambda w + 2\chi\tilde{y} - a_1\tilde{y} - a_2 \fint_\Omega \tilde{y}\,\mathrm{d}x\right)\hat{y} + \chi h_2 \hat{y} + h_1, \end{aligned}$$

with $\frac{\partial y}{\partial n} = 0$, and

$$\begin{cases} -\Delta w + \lambda w = \tilde{y} + h_2 & \text{in } \Omega, \\ \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega. \end{cases}$$

Since $L^2(\Omega) \hookrightarrow W^{1,6/5}(\Omega)^*$, the right-hand side $\tilde{y} + h_2$ belongs to $W^{1,6/5}(\Omega)^*$, that means $w \in W^{1,6}(\Omega)$, by virtue of theorem 4.1.6. On the other hand, setting

$$\mathcal{A}(t) := -\Delta y - \chi\nabla\hat{w} \cdot \nabla y + \left(-\chi\lambda\hat{w} + a_0 - a_1\hat{y} - a_2 \fint_\Omega \tilde{y}\,\mathrm{d}x\right) y,$$

and

$$f(t) := -\chi\nabla\hat{y} \cdot \nabla w + \left(-\chi\lambda w + 2\chi\tilde{y} - a_1\tilde{y} - a_2 \fint_\Omega \tilde{y}\,\mathrm{d}x\right)\hat{y} + \chi h_2 \hat{y} + h_1,$$

we obtain that $y \in W(0,T)$, by virtue of theorem 4.1.8.

We claim if $T > 0$ is small enough, then A is a strict contraction. Let $\tilde{y}_1$ and $\tilde{y}_2$ to be two elements in X, with $y_1 = A\tilde{y}_1$, $y_2 = A\tilde{y}_2$. We define $\bar{y} := y_1 - y_2$ and $\bar{w} := w_1 - w_2$, with $-\Delta w_i + \lambda w_i = \tilde{y}_i + h_2$, $i = 1, 2$. This implies

$$\begin{cases} -\Delta\bar{w} + \lambda\bar{w} = \tilde{y}_1 - \tilde{y}_2 & \text{in } \Omega, \\ \dfrac{\partial\bar{w}}{\partial n} = 0 & \text{on } \Omega, \end{cases}$$

which results in $\|\bar{w}\|_{H^2(\Omega)} \leq C \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)}$, see Grisvard, 1985, Theorem 2.2.2.5

This setting reads

$$\begin{aligned}\partial_t \bar{y} - \Delta \bar{y} = &- \chi \, (\nabla \bar{y} \cdot \nabla \hat{w} + \nabla \hat{y} \cdot \nabla \bar{w}) + \left(-\chi \, \lambda \, \hat{w} + a_0 - a_1 \, \hat{y} - a_2 \fint_\Omega \hat{y} \, \mathrm{d}x \right) \bar{y} \\ &+ \left(-\chi \, \lambda \, \bar{w} + 2 \, \chi \, (\tilde{y}_1 - \tilde{y}_2) - a_1 \, (\tilde{y}_1 - \tilde{y}_2) - a_2 \fint_\Omega (\tilde{y}_1 - \tilde{y}_2) \, \mathrm{d}x \right) \hat{y}.\end{aligned}$$

Multiplying this equation with $\bar{y}$ and integrating by parts yields, we obtain

$$\begin{aligned}\frac{1}{2} \int_\Omega \partial_t \, |\bar{y}|^2 \, \mathrm{d}x + \int_\Omega |\nabla \bar{y}|^2 \, \mathrm{d}x &= -\chi \int_\Omega (\nabla \bar{y} \cdot \nabla \hat{w}) \; \bar{y} \, \mathrm{d}x - \chi \int_\Omega (\nabla \hat{y} \cdot \nabla \bar{w}) \, \bar{y} \, \mathrm{d}x \\ &+ \int_\Omega \left(-\chi \, \lambda \, \hat{w} + a_0 - a_1 \, \hat{y} - a_2 \fint_\Omega \hat{y} \, \mathrm{d}x \right) \bar{y}^2 \, \mathrm{d}x \\ &+ \int_\Omega \left(-\chi \, \lambda \, \bar{w} + 2 \, \chi \, (\tilde{y}_1 - \tilde{y}_2) - a_1 \, (\tilde{y}_1 - \tilde{y}_2) - a_2 \fint_\Omega (\tilde{y}_1 - \tilde{y}_2) \, \mathrm{d}x \right) \hat{y} \, \bar{y} \, \mathrm{d}x \\ &= -\chi \int_\Omega (\nabla \bar{y} \cdot \nabla \hat{w}) \; \bar{y} \, \mathrm{d}x + \chi \int_\Omega (\nabla \bar{y} \cdot \nabla \bar{w}) \, \hat{y} \, \mathrm{d}x + \chi \int_\Omega \bar{y} \, \Delta \bar{w} \, \hat{y} \, \mathrm{d}x \\ &- \chi \, \lambda \int_\Omega \hat{w} \, \bar{y}^2 + \int_\Omega \left(a_0 - a_1 \, \hat{y} - a_2 \fint_\Omega \hat{y} \, \mathrm{d}x \right) \bar{y}^2 \, \mathrm{d}x \\ &+ \int_\Omega \left(-\chi \, \lambda \, \bar{w} + 2 \, \chi \, (\tilde{y}_1 - \tilde{y}_2) - a_1 \, (\tilde{y}_1 - \tilde{y}_2) - a_2 \fint_\Omega (\tilde{y}_1 - \tilde{y}_2) \, \mathrm{d}x \right) \hat{y} \, \bar{y} \, \mathrm{d}x.\end{aligned}$$

Using the Hölder's and interpolation inequalities and exploiting $\|\bar{w}\|_{H^2(\Omega)} \leq C \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)}$, we obtain

$$\begin{aligned}\frac{1}{2} \int_\Omega \partial_t \, |\bar{y}|^2 \, \mathrm{d}x + \int_\Omega |\nabla \bar{y}|^2 \, \mathrm{d}x &\leq \chi \, \|\nabla \hat{w}\|_{L^4(\Omega)} \, \|\bar{y}\|_{L^4(\Omega)} \, \|\nabla \bar{y}\|_{L^2(\Omega)} \\ &+ \chi \, \|\hat{y}\|_{L^\infty(\Omega)} \, \|\nabla \bar{y}\|_{L^2(\Omega)} \, \|\nabla \bar{w}\|_{L^2(\Omega)} + \chi \, \|\hat{y}\|_{L^\infty(\Omega)} \, \|\bar{y}\|_{L^2(\Omega)} \, \|\Delta \bar{w}\|_{L^2(\Omega)} \\ &+ \chi \, \lambda \, \|\hat{w}\|_{L^2(\Omega)} \, \|\bar{y}\|^2_{L^4(\Omega)} + \left\| a_0 - a_1 \, \hat{y} - a_2 \fint_\Omega \hat{y} \, \mathrm{d}x \right\|_{L^\infty(\Omega)} \|\bar{y}\|^2_{L^2(\Omega)} \\ &+ \|\hat{y}\|_{L^\infty(\Omega)} \left\| -\chi \, \lambda \, \bar{w} + 2 \, \chi \, (\tilde{y}_1 - \tilde{y}_2) - a_1 \, (\tilde{y}_1 - \tilde{y}_2) - a_2 \fint_\Omega (\tilde{y}_1 - \tilde{y}_2) \, \mathrm{d}x \right\|_{L^2(\Omega)} \|\bar{y}\|_{L^2(\Omega)} \\ &\lesssim \chi \, \|\nabla \hat{w}\|_{L^4(\Omega)} \, \|\nabla \bar{y}\|_{L^2(\Omega)} \, \|\bar{y}\|^{1/4}_{L^2(\Omega)} \, \|\bar{y}\|^{3/4}_{L^6(\Omega)} + \chi \, \Upsilon \, \|\nabla \bar{y}\|_{L^2(\Omega)} \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} \\ &+ \chi \, \Upsilon \, \|\bar{y}\|_{L^2(\Omega)} \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} + \chi \, \lambda \, \|\hat{w}\|_{L^2(\Omega)} \, \|\bar{y}\|^{1/2}_{L^2(\Omega)} \, \|\bar{y}\|^{3/2}_{L^6(\Omega)} \\ &+ (a_0 + a_1 \, \Upsilon + |a_2| \; \Upsilon) \, \|\bar{y}\|^2_{L^2(\Omega)} \\ &+ \Upsilon \left(\chi \, \lambda \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} + 2 \, \chi \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} \right. \\ &\left. + a_1 \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} + \frac{|a_2|}{|\Omega|} \, \|\tilde{y}_1 - \tilde{y}_2\|_{L^1(\Omega)} \right) \|\bar{y}\|_{L^2(\Omega)}.\end{aligned}$$

In summary, using the Young's inequality, we obtain

$$\begin{aligned}\frac{1}{2} \int_\Omega \partial_t \, |\bar{y}|^2 \, \mathrm{d}x + \int_\Omega |\nabla \bar{y}|^2 \, \mathrm{d}x &\lesssim \varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|\bar{y}\|^2_{L^2(\Omega)} \\ &+ \varepsilon \left(\|\nabla \bar{y}\|^{7/4}_{L^2(\Omega)} \right)^{8/7} + C(\varepsilon) \left(\|\bar{y}\|^{1/4}_{L^2(\Omega)} \right)^8 \\ &+ \varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} + \|\bar{y}\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)}\end{aligned}$$

$$
\begin{aligned}
&+ \varepsilon \left(\|\bar{y}\|^{3/2}_{H^1(\Omega)} \right)^{4/3} + C(\varepsilon) \left(\|\bar{y}\|^{1/2}_{L^2(\Omega)} \right)^4 + \|\bar{y}\|^2_{L^2(\Omega)} \\
&+ \|\tilde{y}_1 - \tilde{y}_2\|_{L^2(\Omega)} \, \|\bar{y}\|_{L^2(\Omega)} \\
&\leq C_1 \left(\varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|\bar{y}\|^2_{L^2(\Omega)} \right) + C_2 \left(\varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + C(\varepsilon) \, \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} \right) \\
&+ C_3 \left(\|\bar{y}\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} \right) + C_4 \left[\varepsilon \, \|\nabla \bar{y}\|^2_{L^2(\Omega)} + (\varepsilon + C(\varepsilon)) \, \|\bar{y}\|^2_{L^2(\Omega)} \right].
\end{aligned}
$$

Choosing the appropriate ε and small enough, the estimate reads as follows

$$
\frac{\mathrm{d}}{\mathrm{d}t} \|\bar{y}\|^2_{L^2(\Omega)} \leq C \left(\|\bar{y}\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} \right), \tag{4.2.2}
$$

where constant C merely depends on C_{T_*} in theorem 4.1.18. Consequently,

$$
\frac{\mathrm{d}}{\mathrm{d}t} \left(\exp(-C\,t) \, \|\bar{y}\|^2_{L^2(\Omega)} \right) \leq C \, \exp(-C\,t) \, \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} .
$$

That implies

$$
\begin{aligned}
\exp(-C\,t) \, \|\bar{y}\|^2_{L^2(\Omega)} &\leq \int_0^t C \, \exp(-C\,s) \, \|(\tilde{y}_1 - \tilde{y}_2)(s)\|^2_{L^2(\Omega)} \, \mathrm{d}s \\
&= (1 - \exp(-C\,t)) \, \|(\tilde{y}_1 - \tilde{y}_2)(s)\|^2_{L^2(\Omega)} .
\end{aligned}
$$

Hence,

$$
\|\bar{y}\|^2_X \leq (\exp(C\,T) - 1) \, \|\tilde{y}_1 - \tilde{y}_2\|^2_X .
$$

We just need to have $\exp(C\,T) - 1 < 1$ a.e., therefore $\exp(C\,T) < 2$.

Let any $T > 0$ to be given. We choose $T_1 > 0$ such that $\exp(C\,T_1) < 2$. Applying Banach Fixed Point Theorem we may find a weak solution $(\hat{y}, \hat{w})$ of the problem (4.2.1), which exists on the time interval $[0, T_1]$. Repeating this argument the solution will be extended to the time interval $[T_1, 2\,T_1]$. We can continue this finitely many times to construct a weak solution existing on the full interval $[0, T]$.

For the uniqueness, let y_1 and y_2 be two solutions of the problem (4.2.1), from the existence proof regarding equation (4.2.2) we have

$$
\frac{\mathrm{d}}{\mathrm{d}t} \|A\,\tilde{y}_1 - A\,\tilde{y}_2\|^2_{L^2(\Omega)} = \frac{\mathrm{d}}{\mathrm{d}t} \|y_1 - y_2\|^2_{L^2(\Omega)} \leq C \left(\|y_1 - y_2\|^2_{L^2(\Omega)} + \|\tilde{y}_1 - \tilde{y}_2\|^2_{L^2(\Omega)} \right).
$$

Since y_1 and y_2 are fixed points, then $y_1 = \tilde{y}_1$, $y_2 = \tilde{y}_2$, respectively. Therefore,

$$
\|y_1 - y_2\|^2_{L^2(\Omega)} \leq 2\,C \int_0^t \|y_1(s) - y_2(s)\|^2 \, \mathrm{d}s
$$

for $0 \leq s \leq T$. According to Gronwall's inequality, $y_1 = y_2$. □

For our next proof we need the following embedding result, which can be deduced from DiBenedetto, 1993, Proposition 3.4:

Theorem 4.2.2. *For every function $y \in W(0,T)$ we have*

$$
\|y\|_{L^p(0,T;L^q(\Omega))} \leq c \, \|y\|_{W(0,T)}
$$

for some positive constant c depending on Ω, where the numbers $p, q \geq 1$ are linked by

$$
\frac{1}{p} + \frac{N}{2q} = \frac{N}{4}
$$

and their admissible range is

$$\begin{cases} q \in (2,\infty],\ p \in [4,\infty); & \text{if } N = 1, \\ q \in [2,\infty),\ p \in (\frac{4}{N},\infty]; & \text{if } N = 2, \\ q \in [2,\frac{2N}{N-2}],\ p \in [2,\infty]; & \text{if } N \geq 3. \end{cases}$$

Theorem 4.2.3. *The control-to-state map $S\colon \mathcal{U}_{\mathrm{ad}} \to W(0,T)\times L^\infty(0,T;W^{1,6}(\Omega))$ is of class C^1.*

Proof. The proof of the theorem is based on the application of the implicit function theorem. Suppose that $u \in \mathcal{U}_{\mathrm{ad}}$ is arbitrary and that $(y,w) \in W(0,T) \times L^\infty(0,T;W^{1,6}(\Omega))$ is the associated state. We consider the operator E

$$\begin{aligned} &E\colon Y \times L^\infty(\partial\Omega_T) \to Z \\ &(y,w,u) \mapsto (E_1(y,w,u), E_2(y,w,u), E_3(y,w,u)), \end{aligned}$$

where $Y \coloneqq W(0,T)\times C([0,T];W^{1,6}(\Omega))$ and $Z \coloneqq L^2(0,T;H^1(\Omega)^*)\times C([0,T];W^{1,6/5}(\Omega)^*)\times L^2(\Omega)$. Furthermore, E_1, E_2 and E_3 are defined as follows:

$$\begin{aligned} \langle E_1(y,w,u), \varphi\rangle \coloneqq &\int_0^T\int_\Omega \partial_t y\,\varphi \,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega \nabla y\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t \\ &- \chi\int_0^T\int_\Omega y\,\nabla w\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t + \chi\int_0^T\int_{\partial\Omega} u\,y\,\varphi\,\mathrm{d}s\,\mathrm{d}t \\ &- \int_0^T\int_\Omega \left(a_0 - a_1\,y - a_2 \fint_\Omega y\,\mathrm{d}x\right) y\,\varphi\,\mathrm{d}x\,\mathrm{d}t, \quad \varphi \in L^2(0,T;H^1(\Omega)), \end{aligned}$$

$$\begin{aligned} \langle E_2(y,w,u), \psi\rangle \coloneqq &\int_0^T\int_\Omega \nabla w\cdot\nabla\psi\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_{\partial\Omega} u\,\psi\,\mathrm{d}s\,\mathrm{d}t + \lambda\int_0^T\int_\Omega w\,\psi\,\mathrm{d}x\,\mathrm{d}t \\ &- \int_0^T\int_\Omega y\,\psi\,\mathrm{d}x\,\mathrm{d}t, \quad \psi \in L^2(0,T;H^1(\Omega)), \end{aligned}$$

and

$$E_3(y,w,u) = y(x,0) - y_0(x) \quad \text{in } \Omega.$$

The operator E_1 is well-defined. Indeed, $y \in W(0,T)$ results in $y \in L^4(0,T;L^3(\Omega))$ for $N \leq 3$, by virtue of DiBenedetto, 1993, Proposition 3.4. That means $y^2 \in L^2(0,T;L^{3/2}(\Omega)) \equiv L^2(0,T;L^3(\Omega)^*)$. Since $\varphi \in L^2(0,T;H^1(\Omega)) \hookrightarrow L^2(0,T;L^6(\Omega))$ for $N \leq 3$ and $L^2(0,T;H^1(\Omega))$ is dense in $L^2(0,T;L^3(\Omega))$ we may extend $\varphi \mapsto \int_0^T\int_\Omega y^2\,\varphi\,\mathrm{d}x\,\mathrm{d}t$ to a continuous functional on $L^2(0,T;H^1(\Omega))$, which means $E_1(y,w,u) \in L^2(0,T;H^1(\Omega)^*)$.

E is continuously Fréchet differentiable with

$$\begin{aligned} E_1'(y,w,u)(\delta y,\delta w,\delta u) = &\int_0^T\int_\Omega \partial_t\delta y\,\varphi\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega \nabla\delta y\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t \\ &- \chi\int_0^T\int_\Omega (\delta y\,\nabla w + y\,\nabla\delta w)\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t + \chi\int_0^T\int_{\partial\Omega}(\delta u\,y + u\,\delta y)\,\varphi\,\mathrm{d}s\,\mathrm{d}t \\ &- \int_0^T\int_\Omega \left(a_0 - a_1\,y - a_2\fint_\Omega y\,\mathrm{d}x\right)\delta y\,\varphi\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega \left(a_1\,\delta y + a_2\fint_\Omega \delta y\,\mathrm{d}x\right) y\,\varphi\,\mathrm{d}x\,\mathrm{d}t, \end{aligned}$$

$$E_2'(y,w,u)(\delta y,\delta w,\delta u) = \int_0^T\int_\Omega \nabla\delta w\cdot\nabla\psi\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_{\partial\Omega}\delta u\,\psi\,\mathrm{d}s\,\mathrm{d}t$$

$$+ \lambda \int_0^T \int_\Omega \delta w \, \psi \, \mathrm{d}x \, \mathrm{d}t - \int_0^T \int_\Omega \delta y \, \psi \, \mathrm{d}x \, \mathrm{d}t,$$

and

$$E_3'(y, w, u)(\delta y, \delta w, \delta u) = \delta y(x, 0).$$

We note that well definedness and differentiability of the operator E_3 follows from $W(0,T) \hookrightarrow C([0,T]; L^2(\Omega))$.

Since for any $u \in \mathcal{U}_{\mathrm{ad}}$, $E(y, w, u) = 0$ is equivalent to $S(u) = (y, w)$, the control-to-state map is C^1 provided that $E_{(y,w)}(y, w, u) \in \mathcal{L}(Y, Z)$ is boundedly invertible. The existence and uniqueness of $(\delta y, \delta w) \in Y$ satisfying (4.2.1), i. e., $E_{(y,w)}(y, w, u)(\delta y, \delta w) = h = (h_1, h_2, u, y_0)$, is established by virtue of proposition 4.2.1. This yields that $E_{(y,w)}(y, w, u)$ is bijective.

It follows from the continuous inverse theorem that the inverse of $E_{(y,w)}(y, w, u)$ is continuous. By the implicit function theorem, there exists a unique continuous function

$$\begin{aligned} &S \colon L^\infty(\partial\Omega_T) \to W(0,T) \times L^\infty(0,T; W^{1,6}(\Omega)) \\ &S(u) = (S_1(u), S_2(u)), \end{aligned}$$

such that $y = S_1(u)$ and $w = S_2(u)$.

In addition, S is differentiable and $\delta y = S_1'(u)\delta u$ and $\delta w = S_2'(u)\delta u$ satisfies

$$E_{(y,w)}(y, w, u)(\delta y, \delta w) = -E_u(y, w, u)\delta u.$$

For any $\delta u \in L^\infty(\partial\Omega_T)$ this equation has a unique solution $(\delta y, \delta w)$ and this solution satisfies the following problem

$$\begin{cases} \partial_t \delta y - \Delta \delta y = -\chi \operatorname{div}(\delta y \nabla w + y \nabla \delta w) + \left(a_0 - a_1 y - a_2 \fint_\Omega y \, \mathrm{d}x\right) \delta y \\ \qquad\qquad - \left(a_1 \delta y + a_2 \fint_\Omega \delta y \, \mathrm{d}x\right) y, & \text{in } \Omega_T, \\ -\Delta \delta w + \lambda \delta w = \delta y & \text{in } \Omega_T, \\ \dfrac{\partial \delta y}{\partial n} = 0, \quad \dfrac{\partial \delta w}{\partial n} = \delta u & \text{on } \partial\Omega_T, \\ \delta y(x, 0) = 0 & \text{in } \Omega. \end{cases} \tag{4.2.3}$$

□

4.2.2 First Order Optimality Condition

In this subsection we aim to derive the first order optimality conditions for a locally optimal solution $(\bar{y}, \bar{w}, \bar{u})$ of problem (4.1.1).

We can derive the optimality system by using the reduced cost functional as explained in subsection 2.2.5. However, we apply the formal Lagrangian method to derive the first order optimality condition. The Lagrangian function $\mathcal{L}\colon Y \times L^\infty(\partial\Omega_T) \times Y \to \mathbb{R}$ is defined by

$$\begin{aligned}\mathcal{L}(y,w,u,p,q) = {} & \frac{1}{2}\int_\Omega |y(x,T) - y_d(x)|^2 \,\mathrm{d}x + \frac{\gamma}{2}\int_0^T\int_{\partial\Omega} |u(x,t)|^2 \,\mathrm{d}s\,\mathrm{d}t \\ & + \int_0^T\int_\Omega \partial_t y\, p \,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega \nabla y \cdot \nabla p \,\mathrm{d}x\,\mathrm{d}t - \chi\int_0^T\int_\Omega y\,\nabla w \cdot \nabla p \,\mathrm{d}x\,\mathrm{d}t \\ & + \chi\int_0^T\int_{\partial\Omega} u\, y\, p \,\mathrm{d}s\,\mathrm{d}t - \int_0^T\int_\Omega \left(a_0 - a_1\, y - a_2 \fint_\Omega y\right) y\, p \,\mathrm{d}x\,\mathrm{d}t \\ & + \int_0^T\int_\Omega \nabla w \cdot \nabla q \,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_{\partial\Omega} u\, q \,\mathrm{d}s\,\mathrm{d}t + \lambda\int_0^T\int_\Omega w\, q \,\mathrm{d}x\,\mathrm{d}t \\ & - \int_0^T\int_\Omega y\, q \,\mathrm{d}x\,\mathrm{d}t + \int_\Omega (y(x,0) - y_0(x))\, p(x,0) \,\mathrm{d}x.\end{aligned}$$

Taking the derivative with respect to the state y, we obtain

$$\begin{aligned}\mathcal{L}_y(y,w,u,p,q)\delta y = {} & \int_\Omega \big(y(x,T) - y_d(x)\big)\, \delta y(x,T) \,\mathrm{d}x + \int_0^T\int_\Omega \partial_t \delta y\, p \,\mathrm{d}x\,\mathrm{d}t \\ & + \int_0^T\int_\Omega \nabla\delta y \cdot \nabla p \,\mathrm{d}x\,\mathrm{d}t - \chi\int_0^T\int_\Omega \nabla w \cdot \nabla p\, \delta y \,\mathrm{d}x\,\mathrm{d}t + \chi\int_0^T\int_{\partial\Omega} u\, p\, \delta y \,\mathrm{d}s\,\mathrm{d}t \\ & + a_1\int_0^T\int_\Omega p\, y\, \delta y \,\mathrm{d}x\,\mathrm{d}t + a_2\int_0^T\int_\Omega p\, y \left(\fint_\Omega \delta y \,\mathrm{d}x\right)\mathrm{d}x\,\mathrm{d}t \\ & - \int_0^T\int_\Omega p\left(a_0 - a_1\, y - a_2\fint_\Omega y\,\mathrm{d}x\right)\delta y \,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_\Omega q\,\delta y \,\mathrm{d}x\,\mathrm{d}t \\ & + \int_\Omega p(x,0)\,\delta y(x,0)\,\mathrm{d}x.\end{aligned}$$

Integrating by parts and using

$$\int_0^T\int_\Omega \partial_t \delta y\, p \,\mathrm{d}x\,\mathrm{d}t = \int_\Omega p\,\delta y \,\mathrm{d}x\,\big|_0^T - \int_0^T\int_\Omega \partial_t p\,\delta y \,\mathrm{d}x\,\mathrm{d}t,$$

we can infer that

$$\begin{aligned}
\mathcal{L}_y(y,w,u,p,q)\delta y &= \int_\Omega (y(x,T)-y_d(x))\,\delta y(x,T)\,\mathrm{d}x + \int_\Omega p\,\delta y\,\mathrm{d}x\,|_0^T - \int_0^T\int_\Omega \partial_t p\,\delta y\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega \Delta p\,\delta y\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\partial\Omega} \partial_n p\,\delta y\,\mathrm{d}s\,\mathrm{d}t \\
&\quad - \chi\int_0^T\int_\Omega \nabla w\cdot\nabla p\,\delta y\,\mathrm{d}x\,\mathrm{d}t + \chi\int_0^T\int_{\partial\Omega} u\,p\,\delta y\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + a_1\int_0^T\int_\Omega p\,y\,\delta y\,\mathrm{d}x\,\mathrm{d}t + a_2\int_0^T \fint_\Omega \left(\int_\Omega p\,y\,\mathrm{d}x\right)\delta y\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega p\left(a_0 - a_1\,y - a_2\fint_\Omega y\,\mathrm{d}x\right)\delta y\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_\Omega q\,\delta y\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_\Omega p(x,0)\,\delta y(x,0)\,\mathrm{d}x \\
&= \int_\Omega \Big(y(x,T)-y_d(x)+p(x,T)\Big)\delta y(x,T)\,\mathrm{d}x \\
&\quad - \int_\Omega p(x,0)\,\delta y(x,0)\,\mathrm{d}x + \int_\Omega p(x,0)\,\delta y(x,0)\,\mathrm{d}x - \int_0^T\int_\Omega \partial_t p\,\delta y\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega \Delta p\,\delta y\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\partial\Omega} \partial_n p\,\delta y\,\mathrm{d}s\,\mathrm{d}t \\
&\quad - \chi\int_0^T\int_\Omega \nabla w\cdot\nabla p\,\delta y\,\mathrm{d}x\,\mathrm{d}t + \chi\int_0^T\int_{\partial\Omega} u\,p\,\delta y\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + a_1\int_0^T\int_\Omega p\,y\,\delta y\,\mathrm{d}x\,\mathrm{d}t + a_2\int_0^T\int_\Omega \left(\fint_\Omega p\,y\,\mathrm{d}x\right)\delta y\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega p\left(a_0 - a_1\,y - a_2\fint_\Omega y\,\mathrm{d}x\right)\delta y\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_\Omega q\,\delta y\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

We notice that

$$\begin{aligned}
a_2\int_0^T\int_\Omega p\,y\left(\fint_\Omega \delta y\,\mathrm{d}x\right)\mathrm{d}x\,\mathrm{d}t &= a_2\fint_\Omega\left(\int_0^T\int_\Omega p\,y\,\mathrm{d}x\,\mathrm{d}t\right)\delta y\,\mathrm{d}x \\
&= a_2\int_0^T\int_\Omega\left(\fint_\Omega p\,y\,\mathrm{d}x\right)\delta y\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Now, we take the derivative of Lagrangian with respect to w and obtain

$$\begin{aligned}
&\mathcal{L}_w(y,w,u,p,q)\delta w \\
&\quad = \chi\int_0^T\int_\Omega y\,\nabla p\cdot\nabla\delta w\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_\Omega \nabla q\cdot\nabla\delta w\,\mathrm{d}x\,\mathrm{d}t - \lambda\int_0^T\int_\Omega q\,\delta w\,\mathrm{d}x\,\mathrm{d}t \\
&\quad = -\chi\int_0^T\int_\Omega \operatorname{div}(y\,\nabla p)\,\delta w\,\mathrm{d}x\,\mathrm{d}t + \chi\int_0^T\int_{\partial\Omega} y\,\partial_n p\delta w\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + \int_0^T\int_\Omega \Delta q\,\delta w\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_{\partial\Omega}\partial_n q\,\delta w\,\mathrm{d}s\,\mathrm{d}t - \lambda\int_0^T\int_\Omega q\,\delta w\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

In summary, if (y, w, u) is a solution of problem (4.1.1), then there exists a Lagrange multiplier $(p, q) \in Y$ such that:

$$\begin{cases} \partial_t p + \Delta p + \chi \nabla w \cdot \nabla p + q = -\left(a_0 - 2\, a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x\right) p + a_2 \fint_\Omega p\, y \,\mathrm{d}x, & \text{in } \Omega_T, \\ \Delta q - \lambda\, q = \chi \operatorname{div}(y \nabla p) & \text{in } \Omega_T, \\ \dfrac{\partial p}{\partial n} = -\chi\, u\, p, \quad \dfrac{\partial q}{\partial n} = -\chi^2\, u\, p\, y & \text{on } \partial\Omega_T, \\ p(x, T) = y(x, T) - y_d(x) & \text{in } \Omega. \end{cases} \tag{4.2.4}$$

We note that the existence of a unique solution (p, q) of (4.2.4) follows from Tröltzsch, 2010, lemma 3.17.

At the end, taking the derivative of Lagrangian with respect to control, we obtain

$$\begin{aligned} \mathcal{L}_u(y, w, u, p, q)\delta u = &\gamma \int_0^T \int_{\partial\Omega} u\, \delta u \,\mathrm{d}s \,\mathrm{d}t + \chi \int_0^T \int_{\partial\Omega} y\, p\, \delta u \,\mathrm{d}s \,\mathrm{d}t \\ &- \int_0^T \int_{\partial\Omega} q\, \delta u \,\mathrm{d}s \,\mathrm{d}t. \end{aligned}$$

Given an optimal control $\bar{u}$ and corresponding state $(\bar{y}, \bar{w})$, then there exists a solution $(p, q) \in Y$ of (4.2.4) such that:

$$\int_0^T \int_{\partial\Omega} (\gamma\, \bar{u} + \chi\, \bar{y}\, p - q)\,(u - \bar{u}) \,\mathrm{d}s \,\mathrm{d}t \geq 0 \quad \text{for all } u \in \mathcal{U}_{\mathrm{ad}}.$$

This implies

$$\bar{u} = \operatorname{proj}_{\mathcal{U}_{\mathrm{ad}}} \left(-\frac{1}{\gamma}\, \chi\, \bar{y}\, p + \frac{1}{\gamma}\, q\right),$$

where $\operatorname{proj}\colon L^1(\partial\Omega_T) \to \mathcal{U}_{\mathrm{ad}}$ is the metric Projection operator. We note that $\frac{1}{\gamma}\, \chi\, \bar{y}\, p - \frac{1}{\gamma}\, q \in L^1(\partial\Omega_T)$.

Now we can derive the following system of necessary optimality conditions.

Theorem 4.2.4. *Suppose that $(y, w, u) \in W(0, T) \times L^\infty(0, T; W^{1,6}(\Omega)) \times \mathcal{U}_{\mathrm{ad}}$ is a locally optimal solution of problem (4.1.1). Then there exists a unique adjoint state $(p, q) \in W(0, T) \times L^\infty(0, T; W^{1,6}(\Omega))$ such that the following system holds:*

$$\begin{cases} \partial_t p + \Delta p + \chi \nabla w \cdot \nabla p + q = -\left(a_0 - 2\, a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x\right) p + a_2 \fint_\Omega p\, y \,\mathrm{d}x, & \textit{in } \Omega_T, \\ \Delta q - \lambda\, q = \chi \operatorname{div}(y \nabla p) & \textit{in } \Omega_T, \\ \dfrac{\partial p}{\partial n} = -\chi\, u\, p, \quad \dfrac{\partial q}{\partial n} = -\chi^2\, u\, p\, y & \textit{on } \partial\Omega_T, \\ p(x, T) = y(x, T) - y_d(x) & \textit{in } \Omega. \end{cases} \tag{4.2.5a}$$

$$\left\{ u = \operatorname{proj}_{\mathcal{U}_{\mathrm{ad}}} \left(\frac{1}{\gamma}\, \chi\, y\, p - \frac{1}{\gamma}\, q\right), \quad \textit{for all } u \in \mathcal{U}_{\mathrm{ad}}, \right. \tag{4.2.5b}$$

$$\begin{cases} \partial_t y - \Delta y = -\chi \operatorname{div}(y \nabla w) + y \left(a_0 - a_1\, y - a_2 \fint_\Omega y \,\mathrm{d}x \right), & \text{in } \Omega_T, \\ -\Delta w + \lambda\, w = y & \text{in } \Omega_T, \\ \dfrac{\partial y}{\partial n} = 0 \quad \text{and} \quad \dfrac{\partial w}{\partial n} = u & \text{on } \partial\Omega_T, \\ y(x,0) = y_0(x) & \text{in } \Omega. \end{cases} \tag{4.2.5c}$$

5 Conclusions and Outlook

In this thesis, we analyzed optimal control problems governed by partial differential equations with nonlocal terms and focused on two different types of partial differential equations arising in physics and biology.

We provided a framework for essential concepts of partial differential equations and optimal control problems governed by such equations in chapter 2. In section 2.1 we collected the essential definitions and results relevant to our work. Section 2.2 is dedicated to pertinent results for analyzing optimal control problems governed by partial differential equations.

Chapter 3 and chapter 4 are regarded as the main body of our work. We devoted chapter 3 to the optimal control of a nonlocal nonlinear elliptic equation, namely a stationary Kirchhoff equation. In section 3.1 we inspected existence and uniqueness of the solutions of the Kirchhoff equation in theorem 3.1.6. We proved that the Kirchhoff equation has strong solution y in $W^{2,q}(\Omega)$. The existence of global optimal solutions of the optimal control problem was also proved in theorem 3.1.9. It turned out that having controls in $H^1(\Omega)$ was of essential importance. We can however omit the upper bound for the existence theory.

We derived the necessary optimality system of first-order for local minimizers in section 3.2. We noticed that the control-to-state map is not Fréchet differentiable with respect the topology of $H^1(\Omega)$. In the presence of the upper bound for the controls they evidently belong to $L^\infty(\Omega)$, resolving the issue of differentiability of control-to-state map by working with the topology of $H^1(\Omega) \cap L^\infty(\Omega)$. First, in subsection 3.2.1 we proved Fréchet differentiability of the linearized problem in proposition 3.2.1 which was employed to show the Fréchet differentiability of the control-to-state map using implicit function theorem in theorem 3.2.2. Afterwards, we applied the formal Lagrangian method to derive the optimality system in subsection 3.2.2. In subsection 3.2.3 we introduced some analytical solution for optimal control problem, example 3.2.4. We observed that the optimality system contains a nonlinear obstacle problem for the control variable, which can be explained by the presence of the bound constraints in $H^1(\Omega)$ and the H^1 control cost term in the cost functional. To employ an efficient numerical method we required the optimality system to be differentiable in some sense. We chose to relax the bound constraints, penalize the optimal control problem employing the Moreau-Yosida penalty approximation in subsection 3.2.4. This followed some modifications in the optimal control problem. To begin with, we eliminated the bound constraints from the admissible set. However, due to this omission, specifically in the absence of the lower bound, the well-definedness of the control-to-state map can no longer be guaranteed. Additionally, without the upper bound, the controls may not necessarily belong to $L^\infty(\Omega)$, preventing us from working with the topology of $H^1(\Omega) \cap L^\infty(\Omega)$ for the differentiability of the control-to-state map. To address this, we modify the control-to-state map by employing a cut-off function–a family of approximations of the positive part function that fulfills specific properties. Overcoming these difficulties is achieved through the use of the cut-off function. An example of such a cut-off function was given in example 3.2.5. Finally, we added a quadratic penalty term to the objective, resulting in (P_ε). To explore the relationship between the original optimal control problem and the relaxed problem (P_ε), we demonstrated in theorem 3.2.6 that for any null sequence of penalty parameters, there exists a subsequence of global solutions to the corresponding penalized problems that converges weakly to a global solution of the original problem. At the end of this section, in

subsection 3.2.5, we derived first order optimality conditions for the penalized problem. It also turned out in corollary 3.2.9 that the optimal control belongs to $L^\infty(\Omega)$.

In section 3.3, we demonstrated the system of first order optimality system for the penalized problem is differentiable in generalized sense, referred to as Newton differentiability. Furthermore, we introduced a basic semismooth Newton algorithm for solving the penalized problem in Algorithm 3.3.3.

In section 3.4 we discretized the relaxed optimal control problem by means of finite element method. We followed a discretized-then-optimized approach, that is, we first formulated the discrete optimal control problem and then derived the associated discrete optimality system. A discrete semismooth Newton algorithm with nonlinear state update for the solution of a discretized instance of the penalized problem was also introduced in Algorithm 3.4.1.

We concluded chapter 3 with section 3.5, which was devoted to numerical experiments. We investigated the influence of the nonlocality parameter in subsection 3.5.1. It turned out that as α increases the number of iterations of the discrete semismooth Newton method decreases. Furthermore, we studied the dependence of the number of the semismooth Newton steps on the discretization in subsection 3.5.2. We considered three refinement levels and observed a mesh-independent convergence behavior. In subsection 3.5.3 we investigated the behaviour of Algorithm 3.4.1 and observed how the solution to the penalized is affected vy variation of the penalty parameter. As final experiment, we studied influence of the control cost parameters in subsection 3.5.4.

Chapter 4 was devoted to the optimal control of a nonlocal nonlinear parabolic-elliptic system, namely chemotaxis system. The existence theory in section 4.1 was divided into two subsections. In subsection 4.1.1 we demonstrated the existence and uniqueness of a weak solution by virtue of Banach fixed point argument in theorem 4.1.13. Indeed, we constructed a perturbed linear system associated to the nonlinear parabolic-elliptic one and proved the corresponding operator is a strict contradiction for a specific time. Since the solution y of the chemotaxis system can blow up we cannot extend this argument to any given time. We proved the existence of an optimal solution in subsection 4.1.2. First, we constructed the corresponding ordinary differential equation (ODE), namely (4.1.20), to the parabolic equation in the chemotaxis system and determined its maximal existence time (blow-up time of existence) with respect to the initial condition of the PDE. We observed that the occurrence of finite or infinite time blow-up depends on the coefficient of the nonlinear term in the ODE, see remark 4.1.14. Subsequently, we demonstrated in theorem 4.1.15 that solution y of the chemotaxis system is uniformly bounded by solution Υ of the ODE. Finally, we demonstrated some a-priori estimate for the solution of the chemotaxis system and proved the existence of an optimal solution.

We proceeded by deriving first-order necessary optimality conditions in section 4.2. We discussed Fréchet differentiability of the control-to-state operator in subsection 4.2.1. To achieve this, we initially established the Fréchet differentiability of the linearized chemotaxis system in proposition 4.2.1. Afterwards, we demonstrated in theorem 4.2.3 the Fréchet differentiability of the control-to-state map. Finally, a first-order optimality system was derived in subsection 4.2.2.

At the end, we point out some possible future research lines.

(i) Optimal Control of Kirchhoff Equation
Design and analysis of a suitable preconditioner: As mentioned earlier in section 3.4, unlike standard optimal control problems without a nonlocal PDE, certain blocks in the discrete generalized Newton system lose sparsity. For an efficient implementation, it is crucial not to assemble these blocks as matrices. Instead,

providing matrix-vector products and utilizing a preconditioned iterative solver, such as MINRES (Paige, Saunders, 1975), becomes essential for solving the discrete generalized Newton system. The design and analysis of a suitable preconditioner are deferred to future work.

(*ii*) Optimal Control of Chemotaxis System
Discretization and Implementation: In the future we intend to apply a numerical method to solve the optimality system of the chemotaxis equation. Discretization and numerical experiments will be deferred to future work.

A Appendix: Comment on the Proof of Existence of an Optimal Solution in Delgado, Figueiredo, et al., 2017

We believe that the proof concerning the existence of an optimal solution in Theorem 2.5 of Delgado, Figueiredo, et al., 2017 contains a flaw. That proof uses the direct method of the calculus of variations and begins by constructing two sequences $\{u_n\}$ and $\{y_n\}$ satisfying the state equation (3.1.3) and converging weakly in $L^2(\Omega)$. The proof then proceeds to show that the weak limit satisfies the state equati on as well. That claim, however, is incorrect. Indeed, we construct below a counterexample showing that the control-to-state map is not continuous in any meaningful sense w.r.t. the weak L^2-convergence of the controls. We acknowledge that this argument was suggested by one of the reviewers.

It suffices to consider (3.1.3) in the setting $\Omega = (0,1) \subset \mathbb{R}$ with data $b \equiv 1$ and $f \equiv 1$. We consider the sequence of controls $\{u_n\} \subset L^2(\Omega)$ defined by $u_n(x) := 1 + 2\chi(nx)$, where $\chi(x)$ is 1-periodic function on $\mathbb{R}$ defined by

$$\chi(x) := \begin{cases} 0, & 0 \le x \le 1/2, \\ 1, & 1/2 < x \le 1. \end{cases}$$

This sequence clearly satisfies $u_n \rightharpoonup \bar{u} := 2$ in $L^2(\Omega)$; see, for instance, Cioranescu, Donato, 1999, Theorem 2.6.

We now show that $y_n := S(u_n)$ does not converge to $S(\bar{u}) =: \bar{y}$. To this end, we note that $\{y_n\}$ is bounded in $H^2(\Omega)$ and thus a subsequence (which we denote the same) converges weakly in $H^2(\Omega)$ and strongly in $H_0^1(\Omega)$ to some $y^* \in H^2(\Omega) \cap H_0^1(\Omega)$. This implies that

$$\begin{aligned} &\left\| \frac{1}{u_n + \|\nabla y_n\|^2_{L^2(\Omega)}} - \frac{1}{u_n + \|\nabla y^*\|^2_{L^2(\Omega)}} \right\|^2_{L^2(\Omega)} \\ &\quad = \left\| \frac{\|\nabla y^*\|^2_{L^2(\Omega)} - \|\nabla y_n\|^2_{L^2(\Omega)}}{\big(u_n + \|\nabla y_n\|^2_{L^2(\Omega)}\big)\big(u_n + \|\nabla y^*\|^2_{L^2(\Omega)}\big)} \right\|^2_{L^2(\Omega)} \\ &\quad \le C \left(\|\nabla y^*\|^2_{L^2(\Omega)} - \|\nabla y_n\|^2_{L^2(\Omega)} \right)^2 \to 0 \end{aligned}$$

for $n \to \infty$. The estimate employs that the terms in the denominator are bounded below by 1. Consequently,

$$-\Delta y_n = \frac{1}{u_n + \|\nabla y_n\|^2_{L^2(\Omega)}} = \frac{1}{u_n + \|\nabla y^*\|^2_{L^2(\Omega)}} + r_n \tag{A.0.1}$$

holds with some $\|r_n\|_{L^2(\Omega)} \to 0$. Since $\big(u_n + \|\nabla y^*\|^2_{L^2(\Omega)}\big)^{-1}$ oscillates between the values $\big(1 + \|\nabla y^*\|^2_{L^2(\Omega)}\big)^{-1}$ and $\big(3 + \|\nabla y^*\|^2_{L^2(\Omega)}\big)^{-1}$, the right-hand side of (A.0.1) converges weakly in $L^2(\Omega)$ to the function $\frac{1}{2}\big(1 + \|\nabla y^*\|^2_{L^2(\Omega)}\big)^{-1} + \frac{1}{2}\big(3 + \|\nabla y^*\|^2_{L^2(\Omega)}\big)^{-1}$. The passage to the

limit implies

$$-\Delta y^* = \frac{1}{2}\Big(\frac{1}{1+\|\nabla y^*\|^2_{L^2(\Omega)}} + \frac{1}{3+\|\nabla y^*\|^2_{L^2(\Omega)}}\Big).$$

Now if $S(\bar{u}) = \bar{y} = y^*$ held, then

$$-\Delta y^* = \frac{1}{2}\Big(\frac{1}{1+\|\nabla y^*\|^2_{L^2(\Omega)}} + \frac{1}{3+\|\nabla y^*\|^2_{L^2(\Omega)}}\Big) = \frac{1}{2+\|\nabla y^*\|^2_{L^2(\Omega)}}$$

would follow. This, however, is impossible due to the strict convexity of the function $(0,\infty) \ni t \mapsto 1/\big(t+\|\nabla y^*\|^2_{L^2(\Omega)}\big)$.

Consequently, $\bar{y} \neq y^*$ and we obtain that $u_n \rightharpoonup u$ in $L^2(\Omega)$ does not imply $S(u_n) \to S(u)$ in any meaningful sense. Therefore, the proof of Theorem 2.5 of Delgado, Figueiredo, et al., 2017 cannot be correct, since it implies the weak L^2-continuity of the control-to-state map. The issues appears to be in step four of the proof on page 779, where the authors conclude that

$$\sum_{i\in I_n} \lambda_i a_i(x_m) - \hat{a}(x_m) \geq \delta$$

holds *for all* $n \in \mathbb{N}$. This, however, is not the case, and therefore, the desired contradiction is not obtained.

Given the lack of weak L^2-continuity of the control-to-state operator, the direct method of the calculus of variations cannot be applied in the setting of Delgado, Figueiredo, et al., 2017, where only an L^2-cost term is present. We overcome this issue by choosing a stronger norm for the control cost term, so that we can use the strong L^2-continuity of the control-to-state map proved in theorem 3.1.7.

B Appendix: Toolbox of Functional Analysis

This chapter provides a quick outline of some fundamentals of functional analysis and is based on Atkinson, Han, 2010; Tröltzsch, 2010; Evans, 1998.

This chapter is divided into two sections. Appendix B.1 provides all related definitions and theorems in linear spaces and appendix B.2 collects important results for linear operators.

B.1 Linear Spaces

In this section, we collect essential concepts and results related to various aspects of linear spaces, with a particular emphasis on significant linear spaces like Banach spaces, Hilbert spaces, and specific function spaces used in this work. In appendix B.1.1 we introduce Banach and Hilbert spaces. The definition of continuously differentiable function spaces and Hölder spaces are presented in appendix B.1.2. Appendix B.1.3 is devoted to introduce L^p-spaces and some important inequalities applicable in these spaces.

B.1.1 Banach and Hilbert Spaces

Definition B.1.1 (Normed Space). *Let X be a linear space over $\mathbb{R}$. A norm on X is a mapping $\|\cdot\| : X \to [0,\infty)$ such that the following properties hold:*

(i) $\|u\| = 0$ *if and only if* $u = 0$,
(ii) $\|\lambda u\| = |\lambda|\,\|u\|$ *for all* $u \in X, \lambda \in \mathbb{R}$ *(homogeneity),*
(iii) $\|u+v\| \le \|u\| + \|v\|$ *for all* $u, v \in X$ *(triangular inequality).*

A normed space is a linear space X endowed with a norm $\|\cdot\|$, denoted by $\{X, \|\cdot\|\}$.

Hereafter we assume X is a normed space.

Definition B.1.2. *A sequence $\{u_n\}_{n=1}^{\infty} \subset X$ is said to be convergent to $u \in X$, denoted by $u_n \to u$, provided*

$$\lim_{n\to\infty} \|u_n - u\| = 0.$$

Definition B.1.3 (Banach Space). *(i) A sequence $\{u_n\}_{n=1}^{\infty}$ is said to be a Cauchy sequence if*

$$\lim_{n,m\to\infty} \|u_n - u_m\| = 0.$$

(ii) X is said to be complete if every Cauchy sequence in X converges, that is, there exists $u \in X$ such that $\{u_n\}_{n=1}^{\infty}$ converges to u.
(iii) A complete normed space is called a Banach space.

Definition B.1.4 (Inner Product Space). *Let H be a real linear space. A mapping $(\cdot,\cdot)\colon H \times H \to \mathbb{R}$ is called an inner (scalar) product on H if the following conditions hold:*

(i) $(u,u) \ge 0$ *for all* $u \in H$,
(ii) $(u,u) = 0$ *if and only if* $u = 0$,
(iii) $(u,v) = (v,u)$ *for all* $u, v \in H$,
(iv) the mapping $u \mapsto (u,v)$ *is linear for all* $v \in H$.

The space H together with the inner product $(\cdot,\cdot)$ is called an inner product space.

If $(\cdot,\cdot)_H$ is an inner product, the induced norm is defined by

$$\|u\| \coloneqq (u,u)_H^{1/2} \quad u \in H.$$

We easily verify this defines a norm on H, using Cauchy-Schwarz inequality

$$|(u,v)| \leq \|u\|_H \|v\|_H \quad \text{for all } u,v \in H.$$

Definition B.1.5 (Hilbert Space). *An inner product space is called a Hilbert space if it is complete wit respect to the induced norm.*

B.1.2 Spaces of Continuously Differentiable Functions

Let Ω denote a bounded domain of $\mathbb{R}^n$ and $u\colon \Omega \to \mathbb{R}$ a function of several variables

$$u = u(x_1, \cdots, x_n).$$

For multi-variable functions, it is convenient to use the multi-index notation for partial derivatives, they are vectors $(\alpha_1, \cdots, \alpha_n)$ having non-negative integer components.

Definition B.1.6. (*i*) *An ordered collection $\alpha = (\alpha_1, \cdots, \alpha_n)$ of non-negative integers α_i is said to be a multiindex of order*

$$|\alpha| = \alpha_1 + \cdots \alpha_n.$$

(*ii*) *Given a multiindex α with $|\alpha| \leq k$, $k \in \mathbb{N} \cup \{0\}$, then for a k-times differentiable function u we define*

$$D^\alpha u(x) \coloneqq \frac{\partial^{|\alpha|} u(x)}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}} = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} u.$$

(*iii*) *If k is a nonnegative integer,*

$$D^k u(x) \coloneqq \{D^\alpha u(x) \mid |\alpha| = k\},$$

denotes the set of partial derivative of order k.

A function from $C(\Omega)$ consisting of real-valued and continuous functions on Ω may exhibit non-smooth behavior as the variable approaches the boundary of Ω.

Let $C(\overline{\Omega})$ be space of continuous functions up to the boundary. This is a Banach space with norm

$$\|u\|_{C(\overline{\Omega})} = \sup_{x\in\Omega} |u(x)| \equiv \max_{x\in\overline{\Omega}} u(x).$$

$C^k(\Omega)$ denotes the space of functions which together with their derivatives of order less than or equal to k, are continuous on Ω; that is,

$$C^k(\Omega) = \{u \in C(\Omega) \mid D^\alpha u \in C(\Omega) \text{ for } |\alpha| \leq k\}.$$

$C^k(\overline{\Omega})$ denotes the space of functions which are continuous up to the boundary, together with their derivatives of order less than or equal to k; that is,

$$C^k(\overline{\Omega}) = \{u \in C(\overline{\Omega}) \mid D^\alpha u \in C(\overline{\Omega}) \text{ for } |\alpha| \leq k\}.$$

This is a Banach space with the norm

$$\|u\|_{C^k(\overline{\Omega})} = \max_{|\alpha|\leq k} \|D^\alpha u\|_{C(\overline{\Omega})}.$$

Definition B.1.7. (*i*) *The closure of a set $E \subset X$ is defined by*

$$\overline{E} = \{x \in X \mid \text{there exists some sequence } \{x_n\}_{n=1}^\infty \subset E, \text{with } x_n \to x\}.$$

The set $E \subset X$ is said to be dense in X if $\overline{E} = X$.

(*ii*) *Support of a function v on Ω is defined to be*

$$\operatorname{supp}(u) = \overline{\{x \in \Omega \mid u(x) \neq 0\}}.$$

(*iii*) *We say that u has a compact support if* $\operatorname{supp}(u)$ *is a proper subset of* Ω.

We denote $C_0^\infty(\Omega)$ as the space of infinitely differentiable functions having compact supports.

Hölder Spaces

We continue with the definition of the less complex Hölder spaces:

Definition B.1.8. *A function $u\colon \Omega \to \mathbb{R}$ is called Lipschitz continuous if*

$$|u(x) - u(y)| \leq c\,|x - y| \quad \text{for all } x, y \in \Omega.$$

For some constant c. The smallest constant in the above inequality is called the Lipschitz constant and is defined by

$$Lip(u) = \sup_{x,y\in\Omega,\, x\neq y} \frac{|u(x) - u(y)|}{|x - y|}.$$

We can generalize this definition to the Hölder continuous functions:

Definition B.1.9. *A function $u\colon \Omega \to \mathbb{R}$ is said to be Hölder continuous with exponent $0 < \gamma \leq 1$, if there exists some constant $C > 0$ such that*

$$|u(x) - u(y)| \leq C\ |x - y|^\gamma\,, \quad x, y \in \Omega.$$

The Hölder space $C^{0,\gamma}(\overline{\Omega})$ is referred to as the subspace of $C(\overline{\Omega})$ functions that are Hölder continuous with the exponent γ. This is a Banach space with the norm

$$\|u\|_{C^{0,\gamma}(\overline{\Omega})} \coloneqq \|u\|_{C(\overline{\Omega})} + [u]_{C^{0,\gamma}(\overline{\Omega})}\,,$$

where

$$[u]_{C^{0,\gamma}(\overline{\Omega})} \coloneqq \sup_{x,y\in\Omega,\, x\neq y} \frac{|u(x) - u(y)|^\gamma}{|x - y|}$$

is the γ^{th}-Hölder seminorm. Similarly, we can define the Hölder space $C^{k,\gamma}(\overline{\Omega})$ consisting of all k-times continuously differentiable functions whose k^{th}-partial derivatives are bounded and Hölder continuous with exponent γ, that is

$$C^{k,\gamma}(\overline{\Omega}) = \{u \in C^k(\overline{\Omega}) \mid D^\alpha u \in C^{0,\gamma}(\overline{\Omega}) \text{ for all } \alpha \text{ with } |\alpha| = k\}.$$

This is a Banach space with the norm

$$\|u\|_{C^{k,\gamma}(\overline{\Omega})} = \|u\|_{C^k(\overline{\Omega})} + \sum_{|\alpha|=k} [D^\alpha u]_{C^{0,\gamma}(\overline{\Omega})}.$$

B.1.3 L^p Spaces

We do not introduce formally the concepts of measurable set and measurable function. Intuitively, the measure of a set $D \subset \mathbb{R}^n$ is its length, area, volume, or suitable generalization. To define the measurable function, we begin by introducing a step function:

Definition B.1.10. *A real-valued function s defined on a measurable set E is called a step function if E can be decomposed into a finite number of pairwise disjoint measurable subsets $E_1, \cdots, E_k$ such that s is constant $s(x) = \alpha_i$ over each E_i, $1 \leq i \leq k$. That means, s can be written as*

$$s(x) = \sum_{i=1}^{k} \alpha_i \chi_{E_i}(x), \quad x \in E,$$

where $\alpha_1, \cdots, \alpha_k$ are scalars and the characteristic function χ_{E_i} is defined by

$$\chi_{E_i}(x) = \begin{cases} 1 & x \in E_i, \\ 0 & x \notin E_i. \end{cases}$$

The function χ_i is measurable if and only if E_i is a measurable set.

Now, we can introduce the definition of a measurable function:

Definition B.1.11. *A function u defined on E is called a measurable function if it is the pointwise limit of a sequence of step functions s_n over E, that means*

$$u(x) = \lim_{n\to\infty} s_n(x), \quad x \in E.$$

Two measurable functions are said to be equal almost everywhere if the set of points on which their function values differ is a set of measure zero. We denote that by

$$u = v \quad a.e.$$

Given a measurable function u on E. An equivalent class of equivalent functions is defined by

$$[u] = \{v \,|\, v \text{ is measurable on } E \text{ and } u = v \, a.e.\}$$

This chapter primarily focuses on the L^p spaces and Sobolev spaces, which frequently emerges as the suitable context for employing concepts from functional analysis to extract insights related to partial differential equations.

Definition B.1.12. *The Lebesgue integral of a step function s over E, is defined by*

$$\int_E s(x)\,\mathrm{d}x = \sum_{j=1}^{k} \alpha_i \left|E_i\right|.$$

The Lebesgue integral of a Lebesgue measurable function u over E, is defined by

$$\int_E u(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_E s_n(x)\,\mathrm{d}x.$$

In the following, we bring an important property of Lebesgue integration:

Theorem B.1.13. *Suppose $\{f_n\}$ is a sequence of Lebesgue integrable functions converging a.e. to f on a measurable set E. If there exists a Lebesgue integrable function g such that*

$$|f_n(x)| \le g(x) \quad a.e. \text{ in } E,\ n \ge 1,$$

then the limit f is Lebesgue integrable and

$$\lim_{n\to\infty} \int_E f_n(x)\,\mathrm{d}x = \int_E f(x)\,\mathrm{d}x.$$

Now, it is time to introduce L^p functions:

Definition B.1.14. *We define the* essential supremum *of a real-valued measurable function $u\colon \Omega \to \mathbb{R}$ to be*

$$\begin{aligned}\operatorname*{ess\,sup}_{x\in\Omega} u(x) &\coloneqq \inf\{\mu \in \mathbb{R} \,|\, |\{u(x) > \mu\}| = 0\} \\ &= \inf\{\sup_{x\in\Omega\setminus\Omega'} u(x) \,|\, \left|\Omega'\right| = 0\}.\end{aligned}$$

Definition B.1.15. *For $1 \le p \le \infty$, we define $L^p(\Omega)$ to be the linear space of all measurable functions $u\colon \Omega \to \mathbb{R}$ for which the following norm is finite:*

$$\|u\|_{L^p(\Omega)} \coloneqq \begin{cases} \left[\int_\Omega |u(x)|^p \,\mathrm{d}x\right]^{1/p} & \text{if } 1 \le p < \infty \\ \operatorname*{ess\,sup}_{x\in\Omega} |u(x)| & \text{if } p = \infty. \end{cases}$$

If $u \in L^\infty(\Omega)$, it called an essentially bounded measurable function.

This is Banach space for $1 \le p \le \infty$ and reflexive for $1 < p < \infty$.

In the following we summarized some properties of L^p spaces.

Theorem B.1.16. *(i) Every Cauchy sequence in $L^p(\Omega)$, $1 \le p \le \infty$ has a subsequence converging pointwise almost everywhere on Ω.*
(ii) If $1 \le p \le \infty$, then $L^q(\Omega) \subset L^q(\Omega)$ and we have

$$\|u\|_{L^p(\Omega)} \le |\Omega|^{1/p-1/q} \|u\|_{L^q(\Omega)} \quad \text{for all } u \in L^q(\Omega).$$

Some Elementary Inequalities

There are some fundamental inequalities employed in our work.

Young's Inequality

Let $a, b \ge 0$ and $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}.$$

Young's Inequality with ϵ

Let $a, b \ge 0$, $\epsilon > 0$ and $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \le \epsilon a^p + C(\epsilon) b^q, \tag{B.1.1}$$

where $C(\epsilon) = (\epsilon p)^{-p/q} q^{-1}$.
For $p = q = 2$ this inequality reads

$$ab \le \epsilon a^2 + \frac{1}{4\epsilon} b^2.$$

Hölder's Inequality

Suppose $1 \le p \le \infty$ with $\frac{1}{q} + \frac{1}{q} = 1$. Then for any $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$ we have

$$\int_\Omega |uv| \, \mathrm{d}x \le \|u\|_{L^p(\Omega)} \|v\|_{L^p(\Omega)}$$

for $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$.

Here, we adopt the convention $1/\infty = 0$. We observe $1 < q < \infty$ if $1 < p < \infty$, $q = 1$ if $p = \infty$, and $q = \infty$ if $p = 1$.

Minkowski's Inequality

Suppose $1 \le p, q \le \infty$ and $u, v \in L^p(\Omega)$. Then we have

$$\|u + v\|_{L^p(\Omega)} \le \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}.$$

Interpolation Inequality

Suppose that $1 \le p \le r \le q \le \infty$, and we choose $0 \le \theta \le 1$ such that

$$\frac{1}{r} = \frac{\theta}{p} + \frac{(1-\theta)}{q}.$$

Then for all $u \in L^q(\Omega)$ we have

$$\|u\|_{L^r(\Omega)} \le \|u\|^{\theta}_{L^p(\Omega)} \|u\|^{1-\theta}_{L^q(\Omega)}. \tag{B.1.2}$$

Gronwall's Inequality

Assume that $\xi(t)$ is a nonnegative and integrable function on $[0, T]$ satisfying

$$\xi(t) \le C \int_0^T \xi(s) \, \mathrm{d}s$$

for a.e. $0 \le t \le T$ and some positive constant. Then

$$\xi(t) = 0 \quad \text{for } a.e.\, t \in [0, T].$$

B.2 Linear Operators on Banach Spaces

Numerous fundamental problems in applied mathematics exhibit linearity, and the utilization of linear spaces and operators offers a comprehensive and valuable framework for the analysis of such problems. While more intricate applications may introduce nonlinear operators, and a study of linear operators also offers some useful tools for the analysis of nonlinear operators.

In this chapter we review some basic results on linear operators. In appendix B.2.1 we present definition of continuous linear operators and the important result of boundedness of inverse operators. Linear functionals and Riesz representation theorem and adjoint operators are introduced in appendix B.2.2. Appendix B.2.4 is devoted to the definition of weak convergence and reflexive space and some characterization of this space.

B.2.1 Continuous Linear Operators

Let X and Y be two real Banach spaces.

Definition B.2.1 (Linear Operator). *A mapping $L\colon X \to Y$ is said to be a linear operator if*

$$L(\lambda u + \mu v) = \lambda L(u) + \mu L(v)$$

for all $u, v \in X$ and $\lambda, \mu \in \mathbb{R}$.

Definition B.2.2. *A linear mapping $L\colon X \to Y$ is called continuous on X if any convergent sequence $u_n \to u$ in X, implies $Lu_n \to Lu$ in Y*

Definition B.2.3 (Bounded Operator). *A linear operator $L\colon X \to Y$ is called bounded if there exist some constant $c > 0$ such that*

$$\|Lu\|_Y \leq c\,\|u\|_X \quad \text{for all } u \in X.$$

The operator norm of a bounded linear operator is defined by

$$\|L\|_{\mathcal{L}(X,Y)} \coloneqq \sup\{\|Lu\|_Y \mid \|u\|_X \leq 1\},$$

which is a finite number.

Theorem B.2.4. *A linear operator is bounded if and only if it is continuous.*

Definition B.2.5. (*i*) *Let V and W be two Banach spaces with $V \subset W$. The space V is said to be continuously embedded in W, denoted by $V \hookrightarrow W$, if*

$$\|v\|_W \leq c\,\|v\|_V \quad \text{for all } v \in V$$

for some $c > 0$. To simplify the computation of estimates and avoid excessive constants, we use the notation $\lesssim$.

(*ii*) *The embedding $V \hookrightarrow W$ is called compact, denoted by $V \hookrightarrow\hookrightarrow W$, if every bounded sequence in V has a subsequence converging in W.*

The normed space of all linear and bounded mapping from X into Y is denoted by $\mathcal{L}(X, Y)$. We write $\mathcal{L}(X)$ if $X = Y$.

The space $\mathcal{L}(X, Y)$ is complete and consequently a Banach space if Y is complete.

The following theorem is widely used in obtaining boundedness of inverse operators.

Theorem B.2.6 (Open Mapping Theorem). *Let X and Y be two Banach spaces. If $L \in \mathcal{L}(X, Y)$ is bijective, then $L^{-1} \in \mathcal{L}(Y, X)$.*

B.2.2 Linear Functionals

Definition B.2.7 (Dual Space). *Let X be a normed space.*

(*i*) *A bounded linear operator $u^*\colon X \to \mathbb{R}$ is said to be a bounded linear functional on X.*

(*ii*) *The space of all bounded linear functionals on X is called dual space of X and is denoted by X^*.*

(*iii*) *If $u \in X$ and $u^* \in X^*$ we denote the pairing of X^* and X by $\langle u^*, u\rangle_{X^*,X}$, which is the real number $u^*(u)$. For an element $u^* \in X^*$ we define*

$$\|u^*\|_{X^*} := \sup\{\langle u^*, u\rangle_{X^*,X} \mid \|u\|_X \leq 1\}.$$

Since $\mathbb{R}$ is a complete space, the dual space X^* is always a Banach space.

Now let H be a real Hilbert space, with inner product $(\cdot,\cdot)$. Every $u \in H$ defines

$$f_u(v) := (u, v)_H,$$

which is a linear functional $f_u \in H^*$ with $\|f_u\|_{H^*} = \|u\|_H$. The converse also holds as the following theorem states.

Theorem B.2.8 (Riesz Representation Theorem). *H^* can be canonically identified with H. More precisely, for any bounded linear functional $u^* \in H^*$ there exists a unique element $u \in H$ such that*

$$\langle u^*, v\rangle_{H^*,H} = (u, v)_H \quad \text{for all } v \in H$$

and $\|u^*\|_{H^*} = \|u\|_H$. The mapping $u^* \mapsto u$ is a linear isomorphism from H^* onto H.

Definition B.2.9 (Adjoint Operator). (*i*) *Let X and Y be two Banach spaces and Let $L\colon X \to Y$ be a bounded linear operator. The mapping $L^*\colon Y^* \to X^*$ is called dual operator of L if*

$$\langle v^*, Lu\rangle_{Y^*,Y} = \langle L^*v^*, u\rangle_{X^*,X} \quad \text{for all } v^* \in Y^*, u \in X.$$

(*ii*) *Let H_1 and H_2 be two linear spaces and $L\colon H_1 \to H_2$ be a bounded linear operator. The operator $L^*\colon H_2 \to H_1$ is called the adjoint operator if it satisfies*

$$(Lu, v)_{H_2} = (u, L^*v)_{H_1} \quad \text{for all } u, v \in H.$$

L is called symmetric if $L^* = L$.

B.2.3 Gelfand triplets

In the analysis of boundary value problems, we frequently encounter a pair of Hilbert spaces, denoted as V and H, having the following properties.

(*i*) V is *continuously embedded* in H, $V \hookrightarrow H$.

(*ii*) V is dense in H.

Using Riesz representation, we may identify H with H^*, writing $H \equiv H^*$. Therefore, H can be continuously embedded into V^*, such that any element in H can be considered as an element of V^* through

$$\langle u, v\rangle_{V^*,V} = (u, v)_H \quad \text{for all } v \in v.$$

Finally, V and therefore also H is dense in V^*. In summary, we have

$$V \hookrightarrow H \hookrightarrow V^* \tag{B.2.1}$$

with dense embeddings. (B.2.1) is called a *Gelfand triplet*.

B.2.4 Weak Convergence

Let X be a real Banach space.

Definition B.2.10 (Weak Convergence). *Let X be a normed space and X^* its dual space. We say a sequence $\{u_n\}_{k=1}^{\infty} \subset X$ converges weakly to $u \in X$, denoted by $u_n \rightharpoonup u$, if*

$$\langle u^*, u_n\rangle_{X^*,X} \to \langle u^*, u\rangle_{X^*,X}$$

for any $u^* \in X^*$.

If a sequence $\{u_n\}_{n=1}^{\infty} \subset X$ converges strongly to $u \in X$, then it also converges weakly to u.

For a weakly convergent sequence $\{u_n\}_{n=1}^{\infty}$ to u in X we have

$$\sup_n \|u_n\|_X < \infty.$$

Thus, every weakly convergent sequence is bounded.

Strong convergence implies weak convergence, but not vice versa. A notable exception is when the space X is finite-dimensional.

Since X^* is a normed space, actually a Banach space, we can consider its dual $X^{**} := (X^*)^*$, called *bidual* of X. This Banach space is normed by

$$\|u^{**}\|_{X^{**}} = \sup\{\langle u^{**}, u^*\rangle_{X^{**},X^*} \mid \|u^*\|_{X^*} \leq 1\}.$$

Between X and X^{**} there exists a *canonical embedding* $J\colon X \to X^{**}$ defined by

$$\langle u^{**}, u^*\rangle_{X^{**},X^*} = \langle u^*, u\rangle_{X^*,X} \quad \text{for all } u^* \in X^*.$$

J is indeed an *isometry*, which refers to $\|Ju\|_{X^{**}} = \|u\|_X$.

Definition B.2.11 (Reflexive Space). *A normed space is called reflexive if $J(X) = X^{**}$.*

Thus, if X is reflexive we can identify X with X^{**} through the *canonical isometry*. An immediate consequence of this definition is that a reflexive normed space must be complete i. e. a Banach space.

Theorem B.2.12. *A Banach space X is reflexive if and only if any bounded sequence in X has a subsequence weakly converging to an element in X.*

Definition B.2.13. *Let X and Y be two real Banach spaces. A mapping $T\colon X \to Y$ is called weakly sequentially continuous if weakly convergence of a sequence $\{u_n\}_{n=1}^{\infty}$ to some $u \in X$ implies that $\{T(u_n)\}_{n=1}^{\infty} \subset Y$ converges weakly to $T(u) \in Y$, this means*

$$u_n \rightharpoonup u \quad \textit{implies } T(u_n) \rightharpoonup T(u)$$

as $n \to \infty$.

We can easily verify that every continuous linear operator $A\colon X \to Y$ is weakly sequentially continuous.

Bibliography

Adam, L.; M. Hintermüller; T. M. Surowiec (2018). "A semismooth Newton method with analytical path-following for the H^1-projection onto the Gibbs simplex". *IMA Journal of Numerical Analysis* 39.3, pp. 1276–1295. DOI: 10.1093/imanum/dry034.

Adams, R. (1975). *Sobolev Spaces*. New York: Academic Press.

Ahmed, E.; A. S. Elgazzar (2007). "On fractional order differential equations model for nonlocal epidemics". *Physica A: Statistical Mechanics and its Applications* 379.2, pp. 607–614. DOI: 10.1016/j.physa.2007.01.010.

Alt, W. (1990). "The Lagrange-Newton method for infinite-dimensional optimization problems". *Numerical Functional Analysis and Optimization* 11, pp. 201–224. DOI: 10.1080/01630569008816371.

Arendt, W.; D. Dier; E. M. Ouhabaz (2014). "Invariance of convex sets for non-autonomous evolution equations governed by forms". *Journal of the London Mathematical Society* 89.3, pp. 903–916. DOI: 10.1112/jlms/jdt082.

Arumugam, G.; J. Tyagi (2021). "Keller-Segel chemotaxis models: A review". *Acta Applicandae Mathematicae* 171.1, pp. 1–82. DOI: 10.1007/s10440-020-00374-2.

Atkinson, K.; W. Han (2010). *Theoretical Numerical Analysis: A Functional Analysis Framework*. Texts in Applied Mathematics. Springer New York. DOI: 10.1007/978-1-4419-0458-4.

Belmiloudi, A. (2017). "Mathematical modeling and optimal control problems in brain tumor targeted drug delivery strategies". *International Journal of Biomathematics* 10.04, p. 1750056. DOI: 10.1142/S1793524517500565.

Casas, E.; M. Mateos; J.-P. Raymond (2007). "Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier–Stokes equations". *SIAM Journal on Control and Optimization* 46.3, pp. 952–982. DOI: 10.1137/060649999.

Chen, C.; P. C. Fife (2000). "Nonlocal models of phase transitions in solids". *Advances in Mathematical Sciences and Applications* 10.2, pp. 821–849. URL: api.semanticscholar.org/CorpusID:989567.

Christof, C.; G. Wachsmuth (2023). "Semismoothness for Solution Operators of Obstacle-Type Variational Inequalities with Applications in Optimal Control". *SIAM Journal on Control and Optimization* 61.3, pp. 1162–1186. DOI: 10.1137/21M1467365.

Cioranescu, D.; P. Donato (1999). *An Introduction to Homogenization*. Vol. 17. Oxford Lecture Series in Mathematics and its Applications. The Clarendon Press, Oxford University Press, New York.

Clever, D.; J. Lang; S. Ulbrich; C. Ziems (2011). "Generalized multilevel SQP-methods for PDAE-constrained optimization based on space-time adaptive PDAE solvers". *International Series of Numerical Mathematics*. Springer Basel, pp. 51–74. DOI: 10.1007/978-3-0348-0133-1_4.

Dautray, R.; J.-L. Lions (2000). *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I*. Berlin: Springer. DOI: 10.1007/978-3-642-58090-1.

Delgado, M.; I. Gayte; C. Morales Rodrigo (2021). "Optimal control of a chemotaxis equation arising in angiogenesis". *Mathematics in Engineering* 4 (6), pp. 1–25. DOI: 10.3934/mine.2022047.

Delgado, M.; G. M. Figueiredo; I. Gayte; C. Morales-Rodrigo (2017). "An optimal control problem for a Kirchhoff-type equation". *ESAIM. Control, Optimisation and Calculus of Variations* 23.3, pp. 773–790. DOI: 10.1051/cocv/2016013.

DiBenedetto, E. (1993). *Degenerate Parabolic Equations*. Berlin: Springer. DOI: 10.1007/978-1-4612-0895-2.

Dolgov, S.; J. W. Pearson (2019). "Preconditioners and tensor product solvers for optimal control problems from chemotaxis". *SIAM Journal on Scientific Computing* 41.6, B1228–B1253. DOI: 10.1137/18M1198041.

Egger, H.; J.-F. Pietschmann; M. Schlottbom (2015). "Identification of chemotaxis models with volume-filling". *SIAM Journal on Applied Mathematics* 75.2, pp. 275–288. DOI: 10.1137/140967222.

Eringen, A. C. (1983). "On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves". *Journal of Applied Physics* 54.9, pp. 4703–4710. DOI: 10.1063/1.332803.

Evans, L. C. (1998). *Partial Differential Equations*. Vol. 19. Graduate Studies in Mathematics. Providence, Rhode Island: American Mathematical Society. DOI: 10.1090/gsm/019.

Figueiredo, G. M.; C. Morales-Rodrigo; J. R. Santos Júnior; A. Suárez (2014). "Study of a nonlinear Kirchhoff equation with non-homogeneous material". *Journal of Mathematical Analysis and Applications* 416.2, pp. 597–608. DOI: 10.1016/j.jmaa.2014.02.067.

Fister, K. R.; C. M. McCarthy (2003). "Optimal control of a chemotaxis system". *Quarterly of Applied Mathematics* 61.2, pp. 193–211. DOI: 10.1090/qam/1976365.

Furter, J.; M. Grinfeld (1989). "Local vs. non-local interactions in population dynamics". *Journal of Mathematical Biology* 27, pp. 65–80. DOI: 10.1007/BF00276081.

Gilbarg, D.; N. S. Trudinger (1977). *Elliptic Differential Equations of Second Order*. New York: Springer.

Grisvard, P. (1985). *Elliptic Problems in Nonsmooth Domains*. Boston: Pitman.

Hillen, T.; K. J. Painter (2008). "A user's guide to PDE models for chemotaxis". *Journal of Mathematical Biology* 58.1-2, pp. 183–217. DOI: 10.1007/s00285-008-0201-3.

Hintermüller, M.; K. Ito; K. Kunisch (2002). "The primal-dual active set strategy as a semismooth Newton method". *SIAM Journal on Optimization* 13.3, pp. 865–888. DOI: 10.1137/s1052623401383558.

Hinze, M.; R. Pinnau; M. Ulbrich; S. Ulbrich (2009). *Optimization with PDE Constraints*. Berlin: Springer. DOI: 10.1007/978-1-4020-8839-1.

Hinze, M.; M. Vierling (2012). "The semi-smooth Newton method for variationally discretized control constrained elliptic optimal control problems; implementation, convergence and globalization". *Optimization Methods & Software* 27.6, pp. 933–950. DOI: 10.1080/10556788.2012.676046.

Horstmann, D. (2004). "From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. II". *Jahresbericht der Deutschen Mathematiker-Vereinigung* 106.2, pp. 51–69.

Ito, K.; K. Kunisch (2008). *Lagrange Multiplier Approach to Variational Problems and Applications*. Vol. 15. Advances in Design and Control. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM). DOI: 10.1137/1.9780898718614.

Kavallaris, N. I.; T. Suzuki (2018). *Non-Local Partial Differential Equations for Engineering and Biology*. Vol. 31. Mathematics for Industry (Tokyo). Mathematical Modeling and Analysis. Springer, Cham, pp. xix–300. DOI: 10.1007/978-3-319-67944-0.

Keller, E. F.; L. A. Segel (1970). "Initiation of slime mold aggregation viewed as an instability". *Journal of Theoretical Biology* 26.3, pp. 399–415. DOI: 10.1016/0022-5193(70)90092-5.

Lacey, A. A. (1983). "Mathematical Analysis of Thermal Runaway for Spatially Inhomogeneous Reactions". *SIAM Journal on Applied Mathematics* 43.6, pp. 1350–1366. DOI: 10.1137/0143090.

Lacey, A. A. (1995). "Thermal runay in a non-local problem modelling Ohmic heating. Part II: General proof of blow-up and asymptotics of runaway". *European Journal of Applied Mathematics* 6.3, pp. 201–224. DOI: 10.1017/S0956792500001807.

Lebiedz, D.; H. Maurer (2004). "External optimal control of self-organisation dynamics in a chemotaxis reaction diffusion system". *Systems Biology* 1.2, pp. 222–229. DOI: 10.1049/sb:20045022.

Liu, C.; Y. Yuan (2022). "Optimal control of a fully parabolic attraction-repulsion chemotaxis model with logistic source in 2D". *Applied Mathematics & Optimization* 85.1, p. 7. DOI: 10.1007/s00245-022-09845-4.

Ma, T. F. (2005). "Remarks on an elliptic equation of Kirchhoff type". *Nonlinear Analysis: Theory, Methods & Applications* 63.5-7, e1967–e1977. DOI: 10.1016/j.na.2005.03.021.

Manzoni, A.; A. Quarteroni; S. Salsa (2022). *Optimal Control of Partial Differential Equations: Analysis, Approximation, and Applications*. Applied Mathematical Sciences. Springer International Publishing. DOI: 10.1007/978-3-030-77226-0.

Morrey, C. B. (1966). *Multiple Integrals in the Calculus of Variations*. 1. Springer Berlin, Heidelberg, pp. XI–506. DOI: 10.1007/978-3-540-69952-1.

Murray, J. D. (1989). *Mathematical Biology*. 2nd ed. New York: Springer. DOI: 10.1007/978-3-662-08539-4.

Negreanu, M.; J. I. Tello (2013). "On a competitive system under chemotactic effects with non-local terms". *Nonlinearity* 26.4, pp. 1083–1103. DOI: 10.1088/0951-7715/26/4/1083.

Paige, C.; M. Saunders (1975). "Solution of sparse indefinite systems of linear equations". *SIAM Journal on Numerical Analysis* 12.4, pp. 617–629. DOI: 10.1137/0712047.

Quittner, P.; P. Souplet (2007). *Superlinear Parabolic Problems: Blow-up, Global Existence and Steady States*. Birkhäuser Cham, pp. XXII–719. DOI: 10.1007/978-3-7643-8442-5.

Rodríguez-Bellido, M. Á.; D. A. Rueda-Gómez; É. J. Villamizar-Roa (2018). "On a distributed control problem for a coupled chemotaxis-fluid model". *Discrete and Continuous Dynamical Systems. Series B* 23.2, pp. 557–571. DOI: 10.3934/dcdsb.2017208.

Rösch, A.; D. Wachsmuth (2008). "Numerical verification of optimality conditions". *SIAM Journal on Control and Optimization* 47.5, pp. 2557–2581. DOI: 10.1137/060663714.

Ryu, S.-U.; A. Yagi (2001). "Optimal control of Keller–Segel equations". *Journal of Mathematical Analysis and Applications* 256.1, pp. 45–66. DOI: 10.1006/jmaa.2000.7254.

Simon, J. (1986). "Compact sets in the space $L^p(0,T;B)$". *Annali di Matematica Pura ed Applicata* 146 (1), pp. 65–96. DOI: 10.1007/BF01762360.

Tello, J. I.; M. Winkler (2007). "A chemotaxis system with logistic source". *Communications in Partial Differential Equations* 32.4-6, pp. 849–877. DOI: 10.1080/03605300701319003.

Temam, R. (1984). *Navier-Stokes Equations, Theory and Numerical Analysis*. Amsterdam: North-Holland. DOI: 10.1090/chel/343.

Troianiello, G. M. (2013). *Elliptic Differential Equations and Obstacle Problems*. Springer New York, NY, pp. XVI–354. DOI: 10.1007/978-1-4899-3614-1.

Tröltzsch, F. (1999). "On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations". *SIAM Journal on Control and Optimization* 38.1, pp. 294–312. DOI: 10.1137/s0363012998341423.

Tröltzsch, F. (2010). *Optimal Control of Partial Differential Equations*. Vol. 112. Graduate Studies in Mathematics. Providence: American Mathematical Society. DOI: 10.1090/gsm/112.

Ulbrich, M. (2011). *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. Vol. 11. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA. DOI: 10.1137/1.9781611970692.

Ulbrich, S. (2007). "Generalized SQP methods with "parareal" time-domain decomposition for time-dependent PDE-constrained optimization". *Real-Time PDE-Constrained Optimization*. Vol. 3. Computational Science and Engineering. SIAM, Philadelphia, PA, pp. 145–168. DOI: 10.1137/1.9780898718935.ch7.

Wolansky, G. (1997). "A critical parabolic estimate and application to nonlocal equations arising in chemotaxis". *Applicable Analysis* 66.3-4, pp. 291–321. DOI: 10.1080/00036819708840588.