



 **Universität Trier**

Optimal Control of Partial Integro-Differential Equations
and Analysis of the Gaussian Kernel

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

Dem Fachbereich IV der Universität Trier
vorgelegt von

Lukas Aaron Zimmer

Köln, Juni 2018

Eingereicht am 12.06.2018

Berichterstatter: Prof. Dr. Ekkehard Sachs
Prof. Dr. Leonhard Frerick

Tag der mündlichen Prüfung: 27.08.2018

Contents

German Summary (Zusammenfassung in deutscher Sprache)	VII
Preface	IX
Introduction	1
I. Necessary Optimality Conditions for Semilinear Partial Integro-Differential Equations	9
1. Preliminaries	11
1.1. Theorems from Analysis	11
1.2. Partial Differential Equations	12
1.3. Function Spaces	14
1.4. Evolution Equations and Abstract Functions	16
2. Semigroup Theory	19
2.1. Strongly Continuous Semigroups	19
2.2. Sectorial Operators and Analytic Semigroups	21
2.3. Fractional Powers of Sectorial Operators	24
2.4. Existence and Uniqueness of Solutions for Evolution Equations	26
2.4.1. Linear Abstract Cauchy Problems	26
2.4.2. Semilinear Evolution Equations	30
2.4.3. Semilinear Initial Boundary Value Problems	34
3. Necessary Optimality Conditions	37
3.1. Necessary Optimality Conditions in Banach Spaces	37
3.2. Intermezzo on Semilinear Boundary Control	45
3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations	52
3.3.1. Abstract Setting	55

3.3.2. Global Existence	62
3.3.3. Adjoint Operators and the Adjoint Equation	64
3.3.4. Necessary Optimality Conditions for the Partial Integro-Differential Equation	68
II. On the Gaussian Kernel: Diffusive Effect, Spectrum and Discretization	71
4. Aspects of the Gaussian Kernel	73
4.1. The Diffusive Effect of the Gaussian Kernel	74
5. The Gaussian Kernel in a Partial Integro-Differential Equation	79
5.1. Discretization and Spectrum of the Coefficient Matrix	80
5.2. Preconditioning	86
5.2.1. Circulant Preconditioners	87
5.2.2. Preconditioning of the Gaussian Kernel	88
5.3. Alternative Discretization Scheme	89
5.4. Numerical Results	96
Conclusion	103
Bibliography	105

German Summary

(Zusammenfassung in deutscher Sprache)

Ein wichtiges Themengebiet der angewandten Mathematik ist die Simulation komplexer finanzmathematischer, mechanischer, chemischer, physikalischer oder medizinischer Prozesse mit mathematischen Modellen. Neben der reinen modellhaften Abbildung der Prozesse ist die gleichzeitige Optimierung einer Zielfunktion durch Änderung der Modellparameter das eigentliche Ziel. Modelle in Bereichen der Finanzmathematik, Biologie und Medizin profitieren von dieser Parameteroptimierung.

Während viele Prozesse mit gewöhnlichen Differentialgleichungen abgebildet werden können, bedarf es partieller Differentialgleichungen zur Optimierung von Wärmeleitungs- und Strömungseigenschaften, der Ausbreitung von Tumorzellen oder auch von Optionspreisberechnungen. Eine partielle Integro-Differentialgleichung ist eine Differentialgleichung mit nicht-lokalen Termen in den Ortsvariablen, z.B. einer Faltung mit einem Integralkern. Zur Modellierung von adhäsiven Kräften zwischen Zellen oder zur Simulation sprunghafter Optionen werden beispielsweise partielle Integro-Differentialgleichungen benötigt.

In beiden Teilen dieser Dissertation untersuchen wir eine partielle Integro-Differentialgleichung. Im ersten Teil stellen wir notwendige Optimalitätsbedingungen für Steuerungsprobleme mit semilinearen partiellen Integro-Differentialgleichungen in der Nebenbedingung auf. Im zweiten analysieren wir eine lineare partielle Integro-Differentialgleichung, die eine Faltung mit dem Gaußkern enthält.

Necessary Optimality Conditions for Semilinear Partial Integro-Differential Equations

Im ersten Teil dieser Arbeit beschäftigen wir uns mit Steuerungsproblemen einer bestimmten Klasse von Integro-Differentialgleichungen, den semilinearen parabolischen Integro-Differentialgleichungen, und greifen für die Anwendung auf ein Modell aus Armstrong et al. (2006) zurück, das die Aggregation von zwei Zelltypen unter Berücksichtigung adhäsiver Kräfte beschreibt.

Die Frage nach der Optimalität einer Steuerungsfunktion ist wesentlich. Es gilt also notwendige Bedingungen zu finden, die von optimalen Steuerungen erfüllt werden. Wir werden in dieser Arbeit Bedingungen an die Modellparameter stellen, unter denen zwei Zellarten eine vorgegebene Aggregation bilden. Dabei werden Armstrongs Modellparameter, die konstant und zeitinvariant gewählt sind, durch zeitabhängige Steuerungsfunktionen ersetzt.

Ziel des ersten Teils der vorliegenden Arbeit ist es also, notwendige Optimalitätsbedingungen an die Steuerungsfunktion eines partiellen Integro-Differentialgleichungssystems zu stellen, so dass dessen Lösung zum Zeitpunkt T einen beobachteten Zustand möglichst genau abbildet. Dazu benötigen wir zunächst Existenz- und Eindeutigkeitsaussagen bezüglich einer Lösung des betrachteten partiellen Integro-Differentialgleichungssystems. Wir bedienen uns hierzu der Halbgruppen Theorie und verwenden das Konzept der milden Lösung einer Evolutionsgleichung. Wir stellen mit Hilfe der Ajungiertengleichung die notwendigen Optimalitätsbedingungen in Banach Räumen auf und wenden dieses Resultat anschließend auf das Zell-Adhäsionsmodell an.

On the Gaussian Kernel: Diffusive Effect, Spectrum and Discretization

Im zweiten Teil der Arbeit untersuchen wir verschiedenen Aspekte des Gaußkerns. Dieser findet in unterschiedlichen Themengebieten der Mathematik Anwendung: als Normalverteilung in der Wahrscheinlichkeitstheorie, als Gausscher Unschärfefilter in der Bildbearbeitung oder als Faltungskern zur Preisberechnung von sprunghaften Optionen in der Finanzmathematik.

Ein Beispiel des letztgenannten Anwendungsfeldes ist das Modell von Merton, das sich als partielle Integro-Differentialgleichung darstellen lässt. Das durch die Diskretisierung des Merton Modells entstehende Gleichungssystem ist auf Grund des Faltung mit dem Gaußkern dicht besetzt. Zur effizienten Berechnung wird ein geeigneter Prädiktionierer benötigt.

Die Betrachtung des Gaußkerns in dieser Arbeit wurde durch zwei Resultate motiviert: Das erste ist der Vergleich zweier verschiedener Prädiktionierer in Ye (2013). Der Autor zeigt, dass ein tridiagonaler Prädiktionierer, der nur auf den diffusiven Term wirkt, den Prädiktionierer von Strang (1986) hinsichtlich der Effizienz übertrifft. Dieses Ergebnis unterstützt die intuitive Einschätzung, dass der Integralterm auf Grund seiner glättenden Eigenschaft keine Prädiktionierung benötigt. Das zweite Ergebnis ist Proposition 6.1 aus Briani et al. (2004), welches besagt, dass der verwendete Integralterm die zweite Ableitung im Fall $\delta \ll 1$ approximiert. Dieses Ergebnis steht in Kontrast zu dem von Ye und führt zu der Vermutung, dass der Integralterm für kleine δ analog zum Laplace Operator prädiktioniert werden sollte.

Wir zeigen in dieser Arbeit zwei Erweiterungen des Ergebnisses von Briani et al. und analysieren anschließend das Eigenwertspektrum des diskretisierten Integralterms in Abhängigkeit von δ . Darüber hinaus regen wir ein alternatives Diskretisierungsschema für die numerische Berechnung der Faltung mit dem Gaußkern an.

Preface

First of all, I would like to express my special gratitude to my supervisor Prof. Dr. Ekkehard W. Sachs for his support during my time as a PhD student and for giving me the opportunity of writing this thesis. I am also thankful to Prof. Dr. Leonhard Frerick not only for serving as examiner but also for many fruitful discussions.

Moreover, I would like to thank everyone who supported me during the compilation of this thesis: my colleagues of the Department of Mathematics at Trier University for the moral support and the great atmosphere they provided; Thomas Bastuck, Ulf Friedrich and Dennis Schwarz for proof reading and practical advise; my uncle Manfred for letting me stay at his place during my time in Trier.

Further, I am very grateful to my wife, Simone, for her enduring support over the last years and my parents, Claudia and Hans-Michael, for facilitating my studies with their encouragement and financial support.

Lukas Aaron Zimmer
Köln, 2018

Introduction

An important field of applied mathematics is the simulation of complex financial, mechanical, chemical, physical or medical processes with mathematical models. In addition to the pure modeling of the processes, the simultaneous optimization of an objective function by changing the model parameters is often the actual goal. Models in fields such as finance, biology or medicine benefit from this optimization step. Such an optimization problem is of the form

$$\begin{aligned} & \min_{\lambda \in \Lambda} J(u, \lambda) \\ & \text{s.t. } G(u, \lambda) = 0, \quad (u, \lambda) \in Z \times \Lambda, \end{aligned}$$

with objective functional J , state function u and control function λ .

While many processes can be modeled using an ordinary differential equation (ODE), partial differential equations (PDEs) are needed to optimize heat conduction and flow characteristics, spreading of tumor cells in tissue as well as option prices. A partial integro-differential equation (PIDE) is a partial differential equation involving an integral operator, e.g., the convolution of the unknown function with a given kernel function. PIDEs occur for example in models that simulate adhesive forces between cells or option prices with jumps.

In order to check whether a function satisfies a PIDE at a particular point $x \in \Omega$, not only the values of the function and its derivatives at x are required. On the contrary, the function needs to be evaluated for all points y that lie within a sensing radius R around x . The evaluation of the neighborhood $U_R(x)$ has an impact on the numerical analysis as well as the numerical calculation of PIDE models.

PIDEs are a generalization of partial differential equations. A calculus for pure nonlocal models, i.e., models that involve no (spatial) derivatives at all, has been developed, see Gunzburger and Lehoucq (2010); Burch and Lehoucq (2011); Du et al. (2013b). Pure nonlocal models have been considered in applications in Zhou and Du (2010); Du and Zhou (2011); Du et al. (2013a); Emmrich et al. (2013); Lehoucq et al. (2014).

In each of the two parts of this thesis, a certain PIDE is the main object of interest. In the first part, we study a semilinear PIDE-constrained optimal control problem with the aim to derive necessary optimality conditions. In the second, we analyze a linear PIDE that includes the convolution of the unknown function with the Gaussian kernel.

Necessary Optimality Conditions for Semilinear Partial Integro-Differential Equations

Checking whether a control function λ is indeed optimal is the core task that has to be solved in control problems. It is therefore inevitable to find necessary conditions that are met by optimal control functions. The research literature on PDE-constrained optimal control is extensive. A comprehensive introduction can be found in Tröltzsch (2009).

In this part, we consider optimal control problems of a particular class of PIDEs: the semilinear parabolic integro-differential equations. The continuous cell-cell adhesion model by Armstrong et al. (2006) serves in the following as an example. The model describes the aggregation of two cell types taking into account adhesive forces. The adhesion parameters, which are constant and time-invariant, are considered as time-dependent control functions in this work. We choose a least squares function and obtain the following control problem for the control function $\lambda = (S_u, S_v, C)$ on the domain $\Omega = (a, b)$ with $a, b \in \mathbb{R}$, $a < b$:

$$\min_{\lambda \in \Lambda} \frac{1}{2} \int_{\Omega} (u(T, x; \lambda) - u_{obs}(x))^2 + (v(T, x; \lambda) - v_{obs}(x))^2 dx. \quad (0.1)$$

The functions $u(\cdot, \cdot)$ and $v(\cdot, \cdot)$ solve the initial value problem

$$\begin{aligned} u_t &= u_{xx} - (uK_u(u, v))_x, & u(0, x) &= u_0(x), & u(t, a) &= u(t, b), \\ v_t &= v_{xx} - (vK_v(u, v))_x, & v(0, x) &= v_0(x), & v(t, a) &= v(t, b), \end{aligned} \quad (0.2)$$

with periodic boundary conditions and integral operators

$$\begin{aligned} K_u(u, v)(t, x) &= \int_{-1}^1 S_u(t)u(t, x+y)\omega(y) + C(t)v(t, x+y)\omega(y) dy, \\ K_v(u, v)(t, x) &= \int_{-1}^1 S_v(t)v(t, x+y)\omega(y) + C(t)u(t, x+y)\omega(y) dy, \end{aligned}$$

where $\omega \in L^1([-1, 1])$ is a given odd function and $u_0, v_0 \in H^1(\Omega)$ are initial values. The functions u_{obs} and v_{obs} are cell aggregations that have been observed at time T . The goal is to choose the parameter functions S_u, S_v and C in such a way that the solutions u and v are closest to the observed cell aggregations at time T .

First, we prove that a unique solution of the model exists. We do so by utilizing the theory of analytic semigroups. We reformulate the model at hand as an evolution equation

$$\nu'(t) + A\nu(t) = F(\nu(t)\lambda(t)), \quad t \in (0, T), \quad \nu(0) = u_0,$$

and consider its state solution $\nu = (u, v) \in X$ and control $\lambda = (S_u, S_v, C) \in \Lambda$ as abstract functions in Banach spaces. In order to allow discontinuous parameter functions for the calibration, we consider the mild solution $\nu \in C([0, T], X)$ of the evolution equation, which is given by the integral equation

$$\nu(t) = e^{-tA}\nu_0 + \int_0^t e^{-(t-s)A}F(\nu(s), \lambda(s)) ds, \quad t \in [0, T],$$

where $\mathcal{T} = \{e^{-tA}\}_{t \geq 0} \subset \mathcal{L}(X)$ is the analytic semigroup generated by $-A$.

The abstract calibration problem is then given by

$$\begin{aligned} \min_{\lambda \in \Lambda} J(\nu, \lambda) &= \frac{1}{2} \|\nu(T; \lambda) - \nu_{obs}\|_{(L^2(\Omega))^2}^2 \\ \text{s.t. } G(\nu, \lambda) &= \nu - e^{-\cdot A} \nu_0 - \int_0^\cdot e^{-(\cdot-s)A} F(\nu(s), \lambda(s)) ds = 0 \\ (\nu, \lambda) &\in C([0, T], (H^1(\Omega))^2) \times \Lambda, \end{aligned}$$

with Λ being a closed and convex set of essentially bounded and integrable functions.

We apply well-known results from Zowe and Kurcyusz (1979) and use the adjoint approach to formulate the necessary optimality conditions in Banach spaces,

$$\int_0^T (F_\lambda(t)^* [p(t)] + g_\lambda(t), \lambda(t) - \bar{\lambda}(t)) dt \geq 0, \quad \lambda \in \Lambda,$$

where $p \in Z^*$ solves

$$p(s) = e^{-(T-s)A^*} \nabla h + \int_s^T e^{-(t-s)A^*} (F_u(t)^* [p(t)] + g_u(t)) dt.$$

Finally, we apply the previous results to provide necessary optimality conditions for the continuous cell-cell adhesion model. Preliminary results on necessary optimality conditions have been published Frerick et al. (2015).

On the Gaussian Kernel: Diffusive Effect, Spectrum and Discretization

In Part II of the thesis, we study different aspects of the one-dimensional Gaussian kernel $\Gamma_\delta: \mathbb{R} \rightarrow \mathbb{R}$,

$$\Gamma_\delta(x) = \frac{1}{\sqrt{2\pi\delta}} e^{-\frac{|x|^2}{2\delta^2}}.$$

The Gaussian kernel is used in various fields of mathematics, e.g., the Gaussian (or normal) distribution in probability theory, the Gaussian blur filter in image processing or as a convolution kernel for simulating option prices with jump diffusion models in finance.

An example in finance is Merton's jump-diffusion model, which can be formulated as a PIDE. The system resulting from the discretization of that model is dense due to the integral term. A suitable preconditioner is required for an efficient computation.

The motivation to analyze the Gaussian kernel was triggered by two results: The first is the comparison of two different types of preconditioners for the jump-diffusion model conducted in Ye (2013). The author shows that a tridiagonal preconditioner that only acts on the diffusive part of the model outperforms Strang's preconditioner. This result supports the intuitive assessment, that the integral term

$$\Gamma_\delta * u - u \tag{0.3}$$

does not need preconditioning due to its smoothing nature. The second is Proposition 6.1 in Briani et al. (2004) which states for $\delta \ll 1$, that the integral term (0.3) approximates the diffusion term u_{xx} . This result is in contrast to the result of Ye and leads to the conjecture, that the discretization of $\Gamma_\delta * u - u$ needs preconditioning for small values of δ as the discretized Laplace operator does.

We show two extensions to the result of Briani et al. and then analyze the eigenvalue spectrum of the discretized integral term with respect to the width δ of the Gaussian kernel. Further, we propose an alternative discretization scheme for the numerical computation of the integral term (0.3).

Structure of the Thesis

In Chapter 1, we first compile elementary results. We cite the theorem on implicit functions and Banach's fixed-point theorem since they will serve as important tools in later proofs. The function spaces on which the partial integro-differential equations will act are introduced in Section 1.3, especially the Sobolov spaces $W_k^{m,p}$ and H_k^m . In Section 1.4, we introduce evolution equations using the concept of abstract functions. $C([a, b], X)$ is the space of continuous abstract functions and $L^p([a, b]; X)$ is the space of p -integrable abstract functions with values in the Banach space X . To make the concept of abstract functions tangible, we conclude this chapter by applying this concept to the one-dimensional heat equation.

Chapter 2 treats the semigroup theory. We use the concept of mild solutions to present existence and uniqueness results for semilinear evolution equations,

$$u'(t) + Au(t) = F(t, u), \quad t \in (0, T), \quad u(0) = u_0,$$

with sectorial operator A and semilinear operator F . For the sake of self-containedness, we first introduce C_0 -semigroups and then analytic semigroups. We also differentiate between C_0 -semigroups and analytic semigroups in Section 2.4.2 to present several regularity conditions for the nonlinear function F . Theorem 2.27 shows that the domain of fractional powers of sectorial operators $X_\alpha := D(A^\alpha)$ is a Banach-space and Theorem 2.29 provides an analogous statement of Sobolev's embedding theorem for these spaces. We show for $2\alpha \in \mathbb{N}_0$, that the spaces $D(-\Delta^\alpha)$ and the Sobolov spaces coincide. In Theorem 2.51, we use the characteristics of X_α and the Banach fixed-point theorem to show, under certain conditions on $F: X_\alpha \rightarrow X$, the local existence and uniqueness of a mild solution

$$u(t) = e^{-tA}u_0 + \int_0^t e^{-(t-s)A}F(s, u(s)) ds,$$

of the semilinear evolution equation on $[0, T)$ with initial value $u_0 \in X_\alpha$. We conclude this chapter with a result from Amann (1988) that provides a mild solution to evolution equations with semilinear boundary conditions.

Chapter 3 is dedicated to the derivation of necessary optimality conditions. In Section 3.1,

we formulate a control problem in an infinite Banach space,

$$\begin{aligned} \min_{\lambda \in \Lambda} J(u, \lambda) &= \int_0^T g(t, u(t), \lambda(t)) dt + h(u(T)) \\ \text{s.t. } G(u, \lambda) &= 0, \quad (u, \lambda) \in Z \times \Lambda, \end{aligned}$$

subject to a semilinear evolution equation

$$u'(t) + Au(t) = F(u(t), \lambda(t)), \quad t \in (0, T), \quad u(0) = u_0.$$

We derive a general approach to formulate the necessary optimality conditions for the control function λ using the adjoint equation in Banach spaces. The results from Theorem 3.3 and Theorem 3.5 meet the prerequisites of Theorem 4.1 from Zowe and Kurcyusz (1979) which is key in the proof of Theorem 3.6. Our main results in Part I is then Theorem 3.9, which provides a representation of the linear functional $l \in Z^*$. Using that representation, we conclude this section with the necessary optimality conditions,

$$\int_0^T (F_\lambda(t)^*[p(t)] + g_\lambda(t), \lambda(t) - \bar{\lambda}(t)) dt \geq 0, \quad \lambda \in \Lambda,$$

where $p \in Z^*$ satisfies

$$p(s) = e^{-(T-s)A^*} \nabla h + \int_s^T e^{-(t-s)A^*} (F_u(t)^*[p(t)] + g_u(t)) dt.$$

Section 3.2 serves as an intermezzo and extends the result from the previous section to a control problem with boundary control. We do that along the lines of Tröltzsch (1989).

In Section 3.3, we first take a closer look at our motivating example, the cell-cell adhesion model of Armstrong et al. (2006) and discuss structural similarities to the nonlocal swarm model of Mogilner and Edelstein-Keshet (1999). Then, we apply the theory developed in the previous sections to the cell-cell adhesion model. This section provides our main result, the computation of the necessary optimality conditions for the two population adhesion model in one dimension. We formulate the cell-cell adhesion model in the abstract setting of Section 3.1 and show that Assumption 3.1 is met. Lemma 3.21 then yields a representation of the linear functional,

$$\begin{aligned} l(\nu) &= - \int_0^T \int_\Omega \left(\pi(t) + \mathcal{K}(\bar{\nu}(t), \bar{\lambda}(t)) D_x \pi(t, x) - \mathcal{K}(D_x \pi(t) \odot \bar{\nu}(t), \bar{\lambda}(t)) \mathbf{id} \right)^\top \nu(t, x) dx dt \\ &\quad - \int_\Omega (\bar{\nu}(T, x) - \nu_{obs}(x))^\top \nu(T, x) dx, \end{aligned}$$

where the adjoint state $\pi = (p, q) \in L^2(0, T; (H^1(\Omega))^2)$ is the mild solution of the terminal value problem

$$\begin{aligned} -\pi'(t) &= D_{xx} \pi(t) + \mathcal{K}(\bar{\nu}(t), \bar{\lambda}(t)) D_x \pi(t) - \mathcal{K}(D_x \pi(t) \odot \bar{\nu}(t), \bar{\lambda}(t)) \mathbf{id} \\ \pi(T) &= \bar{\nu}(T) - \nu_{obs} \end{aligned}$$

for all $t \in (0, T)$. Then, Corollary 3.22 provides the necessary optimality conditions for Equation (0.1). If \bar{u} and \bar{v} are solutions to Equation (0.2) and \bar{S}_u, \bar{S}_v and \bar{C} are optimal controls of Equation (0.1), then the following inequality holds for all admissible controls S_u, S_v and C

$$\int_0^T \int_{\Omega} p(t, x) (\bar{u}(t, x) \hat{K}_u(\bar{u}, \bar{v})(t, x))_x + q(t, x) (\bar{v}(t, x) \hat{K}_v(\bar{u}, \bar{v})(t, x))_x \, dx \, dt \geq 0,$$

with

$$\hat{K}_u(u, v)(t, x) = \int_{-1}^1 (S_u(t) - \bar{S}_u(t)) u(t, x + y) \omega(y) + (C(t) - \bar{C}(t)) v(t, x + y) \omega(y) \, dy$$

and

$$\hat{K}_v(u, v)(t, x) = \int_{-1}^1 (S_v(t) - \bar{S}_v(t)) v(t, x + y) \omega(y) + (C(t) - \bar{C}(t)) u(t, x + y) \omega(y) \, dy,$$

where $p(\cdot, \cdot)$ and $q(\cdot, \cdot)$ solve the terminal value problem

$$\begin{aligned} -p_t &= p_{xx} + p_x K_u(\bar{u}, \bar{v}) - K_u(p_x \bar{u}, q_x \bar{v}), & p(T, x) &= \bar{u}(T, x) - u_{obs}(x), & p(t, a) &= p(t, b), \\ -q_t &= q_{xx} + q_x K_v(\bar{u}, \bar{v}) - K_v(p_x \bar{u}, q_x \bar{v}), & q(T, x) &= \bar{v}(T, x) - v_{obs}(x), & q(t, a) &= q(t, b). \end{aligned}$$

Chapter 4 introduces Part II and we start this chapter by establishing a connection to the first part of this thesis: The convolution with the Gaussian kernel Γ_{δ} itself is a semigroup that solves the heat equation $u_{\delta} = u_{xx}$, with δ being the temporal variable. Section 4.1 is motivated by Proposition 6.1 in Briani et al. (2004), which shows a diffusive effect of the Gaussian kernel. If the integral term $\Gamma_{\delta} * u - u$ in a partial differential equation is replaced locally by the diffusion term $\frac{\delta^2}{2} u_{xx}$, the resulting error in the solution of the PDE can be expressed in terms of δ^2 . In the course of this section, we obtain similar results with respect to the operators $\Lambda_{\delta} - \text{id} - \frac{\delta^2}{2} \Delta$, with $\Lambda_{\delta} u = \Gamma_{\delta} * u$. Theorem 4.2 provides an estimate in the space of three-times differentiable functions and Theorem 4.5 gives an error estimation in the Fourier space.

In Chapter 5, we briefly present the Merton model as an application which employs the Gaussian kernel in its PIDE representation, but focus on the integral term and subsequently analyze

$$u_t(t, x) + \lambda \left(u(t, x) - \int_{-\infty}^{\infty} u(t, x) \Gamma_{\delta}(z - x) \, dz \right) = 0 \quad \text{on } (0, T] \times \mathbb{R}.$$

The discretization scheme proposed in Sachs and Strauss (2008) results in a Toeplitz system

$$T_n u^p = b^p$$

for every time step p . After a brief introduction to the special structure and characteristics of Toeplitz matrices, we derive an estimate for the spectrum of the coefficient matrix T_n in Section 5.1 by analyzing the range of the generating function $g^{(n)}$. The combination of results from Brown and Hewitt (1984) and Gawronski and Stadtmüller (1982) sharpens the

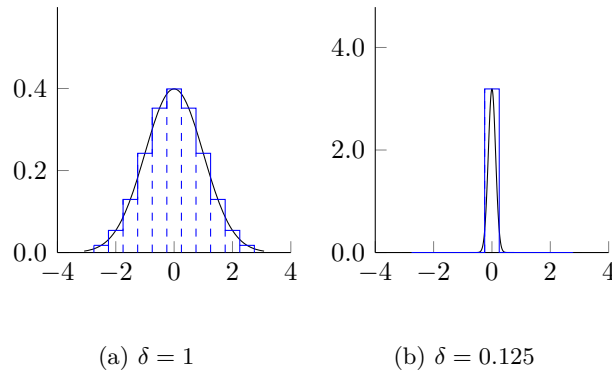


Figure: Increasing discretization error of the Gaussian kernel.

estimate of Ye (2013) for small values of δ . In Section 5.3, we illustrate the increasing error of the discretization scheme for decreasing values of δ . Then, we propose an alternative discretization scheme, that also results in second order accuracy. We use the implementation of the error function for the numerical integration instead of quadrature rules to overcome increasing discretization errors at $\Gamma_\delta(0)$ for $\delta \rightarrow 0$. We conclude the chapter with some numerical experiments in Section 5.4.

The final chapter offers a reflection of the major results. Furthermore, some possibilities for subsequent work within the scope of this thesis are presented.

Part I.

Necessary Optimality Conditions for Semilinear Partial Integro-Differential Equations

Chapter 1

Preliminaries

In this chapter, we provide elementary results and notations that are used in the course of this part. We cite relevant theorems from analysis, introduce the required function spaces and the concept of abstract functions.

1.1 Theorems from Analysis

To start, we recite well-known theorems from analysis that we will use later in this thesis.

Definition 1.1. Let X and E be linear normed spaces and let $U \subset X$ be an open subset. Furthermore, let $x_0 \in U$, a map $f: U \rightarrow E$ as well as a map $df(x_0): X \rightarrow E$ be given. We call f *Fréchet-differentiable* in x_0 with Fréchet-derivative $df(x_0)$, if $df(x_0)$ is continuous and linear and if

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - df(x_0)[h]\|_E}{\|h\|} = 0.$$

We define analogously to the real-valued case the partial Fréchet-differential:

Definition 1.2. Let X , Y and E be linear normed spaces, $U \subset X \times Y$ be open and $f: U \rightarrow E$ be Fréchet-differentiable. We define

$$d_x f(x, y)[r] := df(x, y)[r, 0], \quad r \in X$$

as the *partial Fréchet-differential* with respect to x and

$$d_y f(x, y)[s] := df(x, y)[0, s], \quad s \in Y$$

as the partial Fréchet-differential with respect to y . From the linearity of the Fréchet-derivative $df(x, y)$ follows

$$df(x, y)[r, s] = d_x f(x, y)[r] + d_y f(x, y)[s], \quad (r, s) \in X \times Y.$$

Theorem 1.3 (Implicit function theorem). *Let X, Y and E be Banach spaces, $U \subset X \times Y$ be open and $(x_0, y_0) \in U$. Moreover, let $g: U \rightarrow E$ be continuously Fréchet-differentiable with $g(x_0, y_0) = 0$ such that $d_y g(x_0, y_0): Y \rightarrow E$ is isomorphic. Then there are constants $\epsilon, \delta > 0$ and a continuously differentiable operator $f: U_\delta(x_0) \rightarrow U_\epsilon(y_0)$ such that*

- i) $y = f(x)$ is the unique solution of $g(x, y) = 0$ for all $x \in U_\delta(x_0)$,
- ii) $df(x_0) = -d_y g(x_0, y_0)^{-1} \circ d_x g(x_0, y_0)$.

In other words: If $f(x) = y$ holds for $x \in U_\delta(x_0)$ and $y \in U_\epsilon(y_0)$, then $g(x, y) = 0$ holds. Otherwise if $(x, y) \in U_\delta(x_0) \times U_\epsilon(y_0)$ solves $g(x, y) = 0$, $y = f(x)$ follows.

The implicit function theorem is used in optimal control theory to obtain a operator, that yields a solution to the constraint equation for a given control.

We conclude this section with the Banach fixed-point theorem, which is key for the proof of the existence and uniqueness of a solution for an evolution equation.

Theorem 1.4 (Banach fixed-point theorem). *Let X be a nonempty Banach space and $\phi: X \rightarrow X$ a contraction, i.e., there is a $\gamma < 1$ such that $\|\phi(x) - \phi(y)\| \leq \gamma\|x - y\|$ for all $x, y \in X$. Then ϕ admits a unique fixed-point $x^* \in X$, therefore $\phi(x^*) = x^*$.*

1.2 Partial Differential Equations

Let $\Omega \subset \mathbb{R}^n$ be open, with $n \geq 2$. Let $F: \Omega \times \mathbb{C} \times \mathbb{C}^n \times (\mathbb{C}^{n \times n}) \times \dots \times (\mathbb{C}^{n \times \dots \times n}) \rightarrow \mathbb{C}$ be continuous. We call

$$F(x, u(x), (\partial_1 u(x), \dots, \partial_n u(x)), (\partial^\alpha u(x))_{|\alpha|=2}, \dots, (\partial^\alpha u(x))_{|\alpha|=m}) = 0 \quad (1.1)$$

a *partial differential equation* of order m and $u: \Omega \rightarrow \mathbb{C}$ a (*classical*) *solution* to the partial differential equation, if u satisfies (1.1) for all $x \in \Omega$. We provide fundamentals of partial differential equations in this section.

We begin with the definition of a linear differential operator and declare the symbol of such an operator afterwards.

The tuple $\alpha = (\alpha_1, \dots, \alpha_n)$, with $\alpha \in \mathbb{N}_0^n$, is a multi-index of order $|\alpha| := \sum_{j=1}^n \alpha_j$ with factorial $\alpha! := \prod_{j=1}^n \alpha_j!$. For an m times continuously differentiable function $f \in C^m(\Omega)$, a multi-index $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq m$ and $x \in \Omega \subset \mathbb{R}^n$ we define

$$f^{(\alpha)}(x) := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f(x).$$

For the sake of brevity we set $\partial_j := \frac{\partial}{\partial x_j}$, $\partial_{jk}^2 := \partial_j \partial_k$ as well as $\partial = (\partial_1, \dots, \partial_n)$ and $\partial^\alpha := \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$ resulting in the notation $f^{(\alpha)} = \partial^\alpha f$. It follows that $\partial^\alpha: C^m(\Omega) \rightarrow C(\Omega)$, $u \mapsto \partial^\alpha u$ is well defined and linear.

Definition 1.5. Let $a_\alpha \in C(\Omega)$ be a continuous function for $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq m$. Then,

$$L: C^m(\Omega) \rightarrow C(\Omega), \quad u \mapsto \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u$$

is a *linear differential operator (of order m) with continuous coefficients*. If all a_α are constant, we call L a *linear differential operator with constant coefficients*.

The next definition justifies the characterization of a polynomial as a linear differential operator with constant coefficients.

Definition 1.6. Let L be a linear differential operator. Then we denote by

$$P: \Omega \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad (x, \xi) \mapsto \sum_{|\alpha| \leq m} a_\alpha(x) i^{|\alpha|} \xi^\alpha$$

the *symbol* of L and $P_m(x, \xi) := \sum_{|\alpha|=m} a_\alpha(x) i^m \xi^\alpha$ the *leading symbol*. If L is a linear differential operator with constant coefficients, then P maps from \mathbb{C}^n to \mathbb{C} .

Remark 1.7. Let L be a linear differential operator and P its symbol. Then the Fourier transformation \mathcal{F} yields $L = \mathcal{F}^{-1} P \mathcal{F}$. Indeed, from the differentiation property of the Fourier transform follows

$$\partial^\alpha u(x) = \mathcal{F}^{-1}(\mathcal{F}(\partial^\alpha u))(x) = \mathcal{F}^{-1}(i^{|\alpha|}(\cdot)^\alpha \mathcal{F}(u))(x).$$

Alternatively, with $\tilde{D}_j := \frac{1}{i} \partial_j$ and $\tilde{D}^\alpha := \tilde{D}_1^{\alpha_1} \cdots \tilde{D}_n^{\alpha_n} = \frac{1}{i^{|\alpha|}} \partial^\alpha$ as well as applying the symbol P to \tilde{D} instead of ξ , we receive $L(u) = P(\tilde{D})u$ for linear differential operators with constant coefficients.

We summarize the remark above in the following theorem.

Theorem 1.8. Let $L = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha$ be a linear differential operator with constant coefficients and symbol P . We denote by $V(P) := \{\zeta \in \mathbb{C}^n : P(\zeta) = 0\}$ the set of its roots. The function $e_\zeta: \Omega \rightarrow \mathbb{C}$, $x \mapsto \exp(i\langle x, \zeta \rangle)$ is a solution to the homogeneous differential equation $0 = L(u) = P(\tilde{D})(u)$ for $\zeta \in V(P)$.

Proof. Let $\zeta \in V(P)$. Then, with e_ζ as above we have

$$\begin{aligned} L(e_\zeta)(x) &= \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha \exp(i\langle x, \zeta \rangle) = \sum_{|\alpha| \leq m} a_\alpha i^{|\alpha|} \zeta^\alpha \exp(i\langle x, \zeta \rangle) \\ &= \sum_{|\alpha| \leq m} a_\alpha i^{|\alpha|} \zeta^\alpha e_\zeta(x) = P(\zeta) e_\zeta(x) = 0, \end{aligned}$$

since ζ is a root of P . □

From now on, we assume L to be a linear differential operator of second order with constant coefficients. We use the leading symbol for a classification of linear differential operators.

With Schwarz's theorem (see Rudin, 1976, 9.41 Theorem), a linear differential operator L can be written as

$$L(u) = \sum_{|\alpha| \leq 2} a_\alpha \partial^\alpha u = \sum_{j,k=1}^n a_{j,k} \partial_{j,k}^2 u + \sum_{j=1}^n b_j \partial_j u + cu,$$

whereas $a_{j,k} = a_{k,j}$, i.e., the matrix $A := (a_{j,k})_{j,k=1}^n$ is symmetric.

Definition 1.9. A linear differential operator L given in the notation above is called

- i) *elliptic*, if A is definite,
- ii) *parabolic*, if A is semidefinite, but not definite,
- iii) *hyperbolic*, if A is indefinite and has $n - 1$ positive or $n - 1$ negative eigenvalues,
- iv) *ultrahyperbolic* for the remaining cases.

1.3 Function Spaces

In this section, we provide the function spaces which we will use in the following chapters. Throughout this section, let $\Omega \subset \mathbb{R}^n$ ($n \in \mathbb{N}$) be open and bounded and $k \in \mathbb{N}$. First, we define Lebesgue spaces and their dual spaces. Then, we introduce Sobolev spaces using weak derivatives.

Definition 1.10. Let $k \in \mathbb{N}$. The *Lebesgue space* $L^p(\Omega)^k$ is defined as

$$L^p(\Omega)^k := L^p(\Omega, \mathbb{R}^k) := \{f: \Omega \rightarrow \mathbb{R}^k \text{ measurable} : |f|^p \text{ Lebesgue integrable}\}$$

for $1 \leq p < \infty$ and

$$L^\infty(\Omega)^k := L^\infty(\Omega, \mathbb{R}^k) := \{f: \Omega \rightarrow \mathbb{R}^k \text{ measurable} : f \text{ essentially bounded}\},$$

where we call a measurable function $f: \Omega \rightarrow \mathbb{R}^k$ essentially bounded, if

$$\operatorname{ess\,sup}_\Omega |f| := \sup_{x \in \Omega} |f(x)| < \infty$$

holds almost everywhere, i.e., except on a set N of measure zero. If Ω is clear from context, we briefly write $(L^p)^k$.

The Lebesgue space $L^p(\Omega)^k$ equipped with the norm

$$\|f\|_{L^p(\Omega)^k} := \left(\int_\Omega |f(x)|^p \, dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \quad \|f\|_{L^\infty(\Omega)^k} := \operatorname{ess\,sup}_\Omega |f(x)|$$

is a Banach space (see Meise and Vogt, 1997, Riesz-Fischer theorem 13.5).

Let $p, q \in \mathbb{R}$ with $p, q \geq 1$. Then p and q are conjugate exponents if $\frac{1}{p} + \frac{1}{q} = 1$ holds; furthermore, 1 and ∞ are conjugate exponents. Let p and q be conjugate exponents, then $L^q(\Omega)^k$ is the dual space of $L^p(\Omega)^k$, whereas the dual space of a real normed space X is defined as the space $\mathcal{L}(X, \mathbb{R})$ of continuous linear functionals from X to \mathbb{R} . This relationship follows from the Riesz representation theorem.

Theorem 1.11 (Riesz representation theorem for $L^p(\Omega)$). *Let $1 \leq p < \infty$ and q such that p and q are conjugate exponents. Let $T \in (L^p(\Omega)^k)^* = (L^p(\Omega)^*)^k$. Then a unique $f \in L^q(\Omega)^k$ exists with*

$$T(g) = \int_{\Omega} f(x)g(x) \, dx \quad \text{for all } g \in L^p(\Omega)^k.$$

The norms of both spaces satisfy $\|f\|_{L^q(\Omega)^k} = \|T\|_{(L^p(\Omega)^*)^k}$.

Definition 1.12. Let X and Y be normed spaces and $T \in \mathcal{L}(X, Y)$. The *adjoint operator* $T^*: Y^* \rightarrow X^*$ is then defined as $(T^*y^*)(x) = y^*(Tx)$.

The special case $p = 2$ yields that $L^2(\Omega)^k$ is a Hilbert space with inner product

$$(f, g) = \int_{\Omega} f(x)g(x) \, dx.$$

Let $\phi \in \mathcal{D}(\Omega)$ be a test function, whereas

$$\mathcal{D}(\Omega) := \{\phi \in C^\infty(\Omega) : \text{supp}(\phi) := \overline{\{x : \phi(x) \neq 0\}} \subset \Omega \text{ compact}\}$$

and $f \in C^1(\overline{\Omega})$. With partial integration we then obtain

$$\int_{\Omega} f(x)\phi'(x) \, dx = (-1) \int_{\Omega} f'(x)\phi(x) \, dx,$$

since ϕ vanishes on the border of Ω due to its compact support. These considerations motivate the definition of a weak derivative that generalizes the classical derivative.

Definition 1.13. Let $f, f_\alpha \in L^p(\Omega)^k$ and $\alpha \in \mathbb{N}_0^n$. The function f_α is called α -th weak derivative of f in Ω , if

$$\int_{\Omega} f(x)\partial^\alpha \phi(x) \, dx = (-1)^{|\alpha|} \int_{\Omega} f_\alpha(x)\phi(x) \, dx$$

holds for all test functions $\phi \in \mathcal{D}(\Omega)$. The weak derivative is unique and equal to the classical derivative if the latter exists. Thus, we denote the weak derivative also by ∂^α as well.

We can now define Sobolev spaces:

Definition 1.14. Let $m \in \mathbb{N}_0$, $1 \leq p \leq \infty$. The space

$$W^{m,p}(\Omega)^k := \{f \in L^p(\Omega)^k : \partial^\alpha f \in L^p(\Omega)^k \text{ for all } |\alpha| \leq m\}$$

is called *Sobolev space*. If Ω is given by context, we briefly write $(W^{m,p})^k$. For $m = 0$ we obviously obtain $W^{0,p}(\Omega)^k = L^p(\Omega)^k$.

Equipped with the norm $\|\cdot\|_{(W^{m,p})^k}$, given by

$$\|f\|_{(W^{m,p})^k} := \left(\sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{(L^p)^k}^p \right)^{1/p}, \quad 1 \leq p < \infty,$$

$$\|f\|_{(W^{m,\infty})^k} := \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{(L^\infty)^k}, \quad p = \infty,$$

the Sobolev space $(W^{m,p})^k$ is a Banach space as well. In the case $p = 2$ the Sobolev space $H^m(\Omega)^k := W^{m,2}(\Omega)^k$ is again a Hilbert space with inner product

$$(f, g)_{H^m(\Omega)^k} = \sum_{|\alpha| \leq m} \int_{\Omega} \partial^\alpha f \partial^\alpha g \, dx$$

(see Dobrowolski, 2010, Satz 5.10 and Korollar 5.11).

Substantial for Part I is the following theorem (Kaballo, 2011, Satz 5.12).

Theorem 1.15 (Sobolev embedding theorem). *Let $\Omega := (a, b) \subset \mathbb{R}$ be a real interval. Then it follows:*

(i) *The inclusion $W^{m,1}(\Omega)^k \rightarrow C^{m-1}(\overline{\Omega})^k$ is continuous for $m \in \mathbb{N}$.*

(ii) *For conjugate exponents $1 < p, q < \infty$ the inclusion*

$$W^{m,p}(\Omega)^k \rightarrow C^{m-1, \frac{1}{q}}(\overline{\Omega})^k \subset C^{m-1}(\overline{\Omega})^k$$

is continuous.

Remark 1.16. The space $C^{m-1, \frac{1}{q}}(\overline{\Omega})^k$, defined as

$$C^{m,\theta}(\overline{\Omega})^k := \{f \in C^m(\overline{\Omega})^k : \sup\left\{ \frac{|\partial^\alpha f(x) - \partial^\alpha f(y)|}{|x - y|^\theta} : x \neq y \in \overline{\Omega} \right\} < \infty \text{ for all } |\alpha| = m\},$$

with $\theta \in (0, 1]$ is called *Hölder space*. With $\nu = m + \theta$ we briefly write $(C^\nu)^k$ for $(C^{m,\theta})^k$. If $m = 0$ we write $C^\theta(\overline{\Omega})^k$ instead of $C^{m,\theta}(\overline{\Omega})^k$.

1.4 Evolution Equations and Abstract Functions

Evolution equations are temporally unsteady equations that assign the temporal change of an abstract function at time $t \in (a, b)$ to that time t . Frequently, evolution equations represent partial differential equations describing the temporal change of a spatial quantity, e.g., the heat in an object or the concentration of a substance in a body. For this purpose, the domain $Q \subset \mathbb{R}^{n+1}$ of such a partial differential equation is considered as the product of the time interval $[a, b]$ and the spatial domain $\Omega \subset \mathbb{R}^n$. The function, which is the solution to the partial differential equation, is conceived as an abstract function.

In the subsequent, let $[a, b]$ be a real interval. A map from $[a, b]$ to a Banach space $X = X(\Omega)$ is called an *abstract* or *vector-valued function*. If we choose the real axis \mathbb{R} as the Banach space X , we obtain the real functions of one variable. If, however, we choose the Banach space $H^m(\Omega)$, then the value of an abstract function f is a function itself for every $t \in [a, b]$: $x \mapsto f(t, x) \in H^m(\Omega)$. Choosing $X = (H^m(\Omega))^k$, yields $x \mapsto f(t, x) \in (H^m(\Omega))^k$ and $f(t)$ is a function that maps to any $x \in \Omega$ to a vector in \mathbb{R}^k .

We now introduce function spaces of abstract functions that are relevant for the subsequent chapter.

Definition 1.17. Let X be a real Banach space with norm $\|\cdot\|_X$. $C([a, b], X)$ is then the *space of continuous abstract functions* $f: [a, b] \rightarrow X$ which, equipped with the norm

$$\|f\|_{C([a,b],X)} = \max_{t \in [a,b]} \|f(t)\|_X,$$

is a Banach space itself.

Defining measurable functions analogously to the case $X = \mathbb{R}$, we can formulate Lebesgue spaces for abstract functions.

Definition 1.18. Let X be a Banach with norm $\|\cdot\|_X$. We define the *Lebesgue spaces* $L^p(a, b; X)$ on X as

$$\begin{aligned} L^p(a, b; X) &:= \{f: [a, b] \rightarrow X \text{ measurable} : \int_a^b \|f(t)\|_X^p dt < \infty\}, \quad 1 \leq p < \infty, \\ L^\infty(a, b; X) &:= \{f: [a, b] \rightarrow X \text{ measurable} : \operatorname{ess\,sup}_{[a,b]} \|f(t)\|_X < \infty\}. \end{aligned}$$

Equipped with the corresponding norm $\|\cdot\|_{L^p([a,b],X)}$,

$$\begin{aligned} \|f\|_{L^p(a,b;X)} &:= \left(\int_a^b \|f(t)\|_X^p dt \right)^{1/p}, \quad 1 \leq p < \infty, \\ \|f\|_{L^\infty(a,b;X)} &:= \operatorname{ess\,sup}_{[a,b]} \|f(t)\|_X, \end{aligned}$$

the L^p spaces of abstract functions are also Banach spaces (see Tröltzsch, 2009, p. 114 f.).

Example 1.19 (cf. Tröltzsch, 2009, p. 115). A map $f \in L^2([0, T], (H^1(\Omega))^k)$ can be considered as an abstract square-integrable function with values $f(t)$ in $(H^1(\Omega))^k$ or as a function $f: [0, T] \times \Omega \rightarrow \mathbb{R}^k$, $(t, x) \mapsto f(t, x)$. It follows

$$\begin{aligned} \|f\|_{L^2([0,T],(H^1(\Omega))^k)} &= \left(\int_0^T \|f(t)\|_{(H^1(\Omega))^k}^2 dt \right)^{\frac{1}{2}} \\ &= \left(\int_0^T \int_\Omega (|f(t, x)|^2 + |\nabla_x f(t, x)|^2) dx dt \right)^{\frac{1}{2}}. \end{aligned}$$

The (one-dimensional) heat equation

$$u_t(t, x) = au_{xx}(t, x)$$

1. Preliminaries

can be interpreted as an evolution equation with operator $A \in \mathcal{L}(L^2(\Omega))$, $Au(t) = au_{xx}(t)$ applied to an abstract function $u \in C([0, T], H^2(\Omega))$. On the other hand, it can be treated as a partial differential equation involving a function $u: [0, T] \times \Omega \rightarrow \mathbb{R}$.

Semigroup Theory

The semigroup theory can be interpreted as a generalization of the theory of ordinary differential equations. One of their aims is the investigation of linear and nonlinear partial differential equations and the existence, uniqueness, and regularity of their solutions. For this purpose, partial differential equations are interpreted as abstract evolution equations. We are particularly interested in the existence and uniqueness of the solution of semilinear evolution equations of the form

$$u'(t) + Au(t) = F(t, u), \quad t \in (0, T), \quad u(0) = u_0.$$

We start with the definition of strongly continuous semigroups and then consider the most important class of semigroups, the analytic semigroups. We introduce fractional powers of the generators of analytic semigroups and utilize them in the results on the existence and uniqueness of solutions of semilinear evolutionary equations. The results presented in this chapter are mostly taken from Pazy (1983), Engel and Nagel (2001), Henry (1981) and Haase (2006).

Throughout this chapter, let $X := X(\Omega)$ with $\Omega \subset \mathbb{R}^n$ open, $n \in \mathbb{N}$, be a Banach space with norm $\|\cdot\|$.

2.1 Strongly Continuous Semigroups

Definition 2.1. A family $\mathcal{T} = \{T(t)\}_{t \geq 0} \subset \mathcal{L}(X)$ is called a *strongly continuous semigroup* or *C_0 -semigroup* if

- (i) $T(0) = I$, (I is the identity operator on X),
- (ii) $T(t + s) = T(t)T(s)$ for all $t, s \geq 0$ (the semigroup property),
- (iii) $\lim_{t \rightarrow 0^+} T(t)x = x$ for every $x \in X$ (the strong continuity property).

The semigroup property means that in evolution processes there is no difference between directly evolving from the initial state to the state in $t + s$ or evolving t time units first and then s time units in a second step.

Another way of ensuring the strong continuity property is given in the following lemma.

Lemma 2.2 (Engel and Nagel, 2001, Proposition I.5.3). *The strong continuity property (Property (iii)) in Definition 2.1 is equivalent to the existence of $\delta > 0$, $M \geq 1$, and a dense subset $D \subset X$ such that*

$$(i) \|T(t)\|_{\mathcal{L}(X)} \leq M \text{ for all } t \in [0, \delta],$$

$$(ii) \lim_{t \rightarrow 0^+} T(t)x = x \text{ for all } x \in D.$$

Remark 2.3. The uniform boundedness of the operators $T(t)$ for $t \in [0, \delta]$ is obvious in most cases. Therefore, it is sufficient to only check the (right) continuity of the orbit map ξ_x at $t = 0$ for a dense set $D \subset X$ in order to obtain the strong continuity of the semigroup \mathcal{T} . See Remark 2.8 for the definition of the orbit map.

Definition 2.4. The *infinitesimal generator* $A: D(A) \subset X \rightarrow X$ of a C_0 -semigroup \mathcal{T} , $\mathcal{T} = \{T(t)\}_{t \geq 0}$, is defined by

$$Ax = \lim_{t \rightarrow 0^+} \frac{1}{t}(T(t)x - x)$$

with domain $D(A) = \{x \in X : \lim_{t \rightarrow 0^+} \frac{1}{t}(T(t)x - x) < \infty\}$.

Theorem 2.5 (Pazy, 1983, Corollary 1.2.5, Theorem 1.2.6). *The generator $(A, D(A))$ of a strongly continuous semigroup $\mathcal{T} = \{T(t)\}_{t \geq 0}$ is a closed and densely defined linear operator that determines the semigroup uniquely.*

Theorem 2.6 (Pazy, 1983, Theorem 1.2.2; Engel and Nagel, 2001, Definition I.3.11, Definition I.5.6). *Let $\mathcal{T} = \{T(t)\}_{t \geq 0}$ be a C_0 -semigroup. Then $\{T(t)\}_{t \geq 0}$ is exponentially bounded, i.e., there exist constants $\omega \in \mathbb{R}$ and $M \geq 1$ such that $\|T(t)\|_{\mathcal{L}(X)} \leq Me^{\omega t}$ for $t \in \mathbb{R}^+$. We say that the C_0 -semigroup is*

(i) *bounded, if $\omega = 0$,*

(ii) *contractive, if $\omega = 0$ and $M = 1$ and*

(iii) *exponentially stable, if $\omega < 0$.*

Theorem 2.7 (Pazy, 1983, Theorem 1.2.4). *Let $\mathcal{T} = \{T(t)\}_{t \geq 0}$ be a C_0 -semigroup and $(A, D(A))$ its generator. For $x \in D(A)$ and $t \geq 0$ we have that $T(t)x \in D(A)$, $t \mapsto T(t)x$ is continuously differentiable and*

$$T'(t)x = AT(t)x = T(t)Ax.$$

Remark 2.8. Let $\mathcal{T} = \{T(t)\}_{t \geq 0}$ be a C_0 -semigroup and $x \in X$. We define the orbit map as $\xi_x: [0, \infty) \rightarrow X$, $x \mapsto T(t)x$. The orbit map ξ_x is a continuous function. We derive the generator A of the C_0 -semigroup as the derivative of the orbit map at $t = 0$, i.e., $Ax = \xi'_x(0)$. Its domain includes every $x \in X$ such that ξ_x is differentiable.

2.2 Sectorial Operators and Analytic Semigroups

If we look for semigroups that accept not only real nonnegative parameters but can also be extended into certain domains in the complex plane, we will find analytic semigroups. Remark 2.18 will elucidate the difference between C_0 -semigroups and analytic semigroups. The generator of an analytic semigroup is called a sectorial operator, which is not uniformly defined in the literature. Throughout this thesis we will use the definition provided in Henry (1981) which is also applied in Haase (2006). On Banach spaces, sectorial operators concur with the concept of nonnegative operators (see Haase, 2006, Section 2.8). We start this section with said definition. For that, we recall the definition of a sector of angle θ .

Definition 2.9. For $\theta \in (0, \pi]$ and $\omega \in \mathbb{R}$ we call

$$\Sigma_{\omega, \theta} := \{\lambda \in \mathbb{C} \setminus \{\omega\}; |\arg(\lambda - \omega)| < \theta\}$$

a sector of angle θ at ω . We denote by $\Sigma_{\omega, \theta}^-$ the reflection of that sector along a line through ω parallel to the y -axis, i.e.,

$$\Sigma_{\omega, \theta}^- = \{\lambda \in \mathbb{C} \setminus \{\omega\}; \theta < |\arg(\lambda - \omega)| \leq \pi\}.$$

If $\omega = 0$, we denote the sector $\Sigma_{0, \theta}$ by Σ_θ and its reflection $\Sigma_{0, \theta}^-$ by Σ_θ^- for short.

Definition 2.10. Let $A: D(A) \subset X \rightarrow X$ be a closed linear operator with dense domain $D(A)$ in a Banach space X . A is called a *negative quasi-sectorial operator* (of angle δ), if there exist $\omega \in \mathbb{R}$, $\delta \in (0, \frac{\pi}{2}]$ and $M > 0$ such that

- (i) the sector $\Sigma_{\omega, \frac{\pi}{2} + \delta}$ is contained in the resolvent set of $\rho(A)$,
- (ii) the resolvent $R(\lambda, A) = (\lambda - A)^{-1}$ satisfies

$$\|R(\lambda, A)\|_{\mathcal{L}(X)} \leq \frac{M}{|\lambda - \omega|}, \quad \text{for } \lambda \in \Sigma_{\omega, \frac{\pi}{2} + \delta}.$$

An operator A is called *negative sectorial* (of angle δ) if $\omega = 0$ is possible. A is called a *(quasi-)sectorial operator*, if $-A$ is a negative (quasi-)sectorial operator.

Remark 2.11. (i) The definition above yields that the spectrum of a sectorial operator A is contained in the sector $\overline{\Sigma_{\frac{\pi}{2} - \delta}^-}$, i.e., $|\arg(\sigma(A))| \leq \frac{\pi}{2} - \delta$. Hence, we have $\operatorname{Re} \sigma(A) \geq 0$. Moreover, $\Sigma_{\frac{\pi}{2} - \delta}^- \subset \rho(A)$ and a similar estimate for the resolvent is satisfied: If $\lambda \in \Sigma_{\frac{\pi}{2} - \delta}^-$, then $-\lambda \in \Sigma_{\frac{\pi}{2} + \delta}$. Since $-A$ is negative sectorial, we have

$$\|R(-\lambda, -A)\|_{\mathcal{L}(X)} = \|(-\lambda + A)^{-1}\|_{\mathcal{L}(X)} \leq \frac{M}{|-\lambda|},$$

which is equivalent to $\|(\lambda - A)^{-1}\|_{\mathcal{L}(X)} \leq \frac{M}{|\lambda|}$. Thus, $\|R(\lambda, A)\|_{\mathcal{L}(X)} \leq \frac{M}{|\lambda|}$ for $\lambda \in \Sigma_{\frac{\pi}{2} - \delta}^-$.

- (ii) In Pazy (1983) and Engel and Nagel (2001), a sectorial operator is defined in the way we define a negative sectorial operator.

Definition 2.12. A family $\mathcal{T} = \{T(t)\}_{t \in \Sigma_\delta \cup \{0\}} \subset \mathcal{L}(X)$ of continuous linear operators is called an *analytic semigroup* (of angle $\delta \in (0, \frac{\pi}{2})$), if

- (i) $T(0) = I$ and $T(t)T(s) = T(t+s)$ for $t, s \in \Sigma_\delta$,
- (ii) $t \mapsto T(t) \in \mathcal{L}(X)$ is analytic in Σ_δ ,
- (iii) $\lim_{\Sigma_\delta \ni t \rightarrow 0} T(t)x = x$ for all $x \in X$.

The (infinitesimal) generator $(A, D(A))$ of \mathcal{T} is defined the same way as in Definition 2.4.

Remark 2.13. The restriction of an analytic semigroup to $(0, \infty)$ is a C_0 -semigroup.

Theorem 2.14 (Engel and Nagel, 2001, Section II.4.a). *The linear operator A generates an analytic semigroup of angle δ , if and only if A is negative sectorial of angle $\delta > 0$.*

Remark 2.15. If A would be negative sectorial of angle $\delta = 0$, i.e., only the right half plane is contained in the resolvent set $\rho(A)$, then A is the generator of a C_0 -semigroup.

We will define fractional powers of sectorial operators in the next section. Therefore, we consider from now on sectorial operators $(A, D(A))$ (of angle δ) and analytic semigroups $T(t)$ generated by $-A$ and write $T(t) = e^{-tA}$. We give that expression a meaning by the functional calculus of sectorial operators—that is, roughly speaking, $f(A) = \frac{1}{2\pi i} \int_\Gamma f(\lambda)R(\lambda, A) d\lambda$, where the function f is holomorphic in $\Sigma_{\frac{\pi}{2}-\delta'}$, $0 < \delta' < \delta$, and has rapid decay at ∞ and Γ is a suitable positive oriented path around $\Sigma_{\frac{\pi}{2}-\delta}$ passing through ∞ , cf. Haase (2006).

Theorem 2.16 (Henry, 1981, Theorem 1.3.4). *Let $(A, D(A))$, $A: D(A) \subset X \rightarrow X$ be a sectorial operator of angle δ . For $t \in \Sigma_\delta$, the analytic semigroup $T(t)$ generated by $-A$ is given as*

$$T(t) = e^{-tA} := \frac{1}{2\pi i} \int_\Gamma (\lambda + A)^{-1} e^{\lambda t} d\lambda = \frac{1}{2\pi i} \int_\Gamma R(\lambda, -A) e^{\lambda t} d\lambda,$$

where Γ is any piecewise smooth curve in $\Sigma_{\frac{\pi}{2}+\delta} \subset \rho(-A)$ with $\arg \lambda \rightarrow \pm \frac{\pi}{2} + \delta'$ as $|\lambda| \rightarrow \infty$ for some $\delta' \in (0, \delta)$.

We summarize basic properties of analytic semigroups in the following theorem

Theorem 2.17 (Henry, 1981, Theorem 1.3.4; Engel and Nagel, 2001, Theorem II.4.3). *Let the operator $(A, D(A))$, $A: D(A) \subset X \rightarrow X$ be sectorial of angle δ .*

- (i) $\|e^{-zA}\|_{\mathcal{L}(X)}$ is uniformly bounded for $z \in \Sigma_{\delta'}$ if $0 < \delta' < \delta$.
- (ii) If $\operatorname{Re} \sigma(A) > \omega$, a constant $C > 0$ exists such that

$$\begin{aligned} \|e^{-tA}\|_{\mathcal{L}(X)} &\leq C e^{-\omega t}, \quad t \geq 0, \\ \|Ae^{-tA}\|_{\mathcal{L}(X)} &\leq \frac{C}{t} e^{-\omega t}, \quad t > 0. \end{aligned}$$

Hence, if $\operatorname{Re} \sigma(A) > 0$ the estimates read $\|e^{-tA}\|_{\mathcal{L}(X)} \leq C$ and $\|Ae^{-tA}\|_{\mathcal{L}(X)} \leq \frac{C}{t}$ for $t > 0$.

(iii) $T'(t) = \frac{d}{dt}e^{-tA} = -Ae^{-tA}$ for $t > 0$.

(iv) $e^{-tA}x \in D(A)$ for all $t > 0$ and $x \in X$.

Remark 2.18. We note that for a C_0 -semigroup $\mathcal{T} = \{T(t)\}_{t \geq 0}$ we have $T(t)x \in D(A)$ only for $x \in D(A)$. If $T(t)x \in D(A)$ for $t > 0$ and all $x \in X$, \mathcal{T} is bounded, i.e., $\|T(t)\|_{\mathcal{L}(X)} \leq C$, and $\|AT(t)\|_{\mathcal{L}(X)} \leq \frac{C}{t}$, then \mathcal{T} is an analytic semigroup.

Example 2.19. As a final example we show that the negative Laplace-operator $-\Delta$ on $L^2(\mathbb{R}^n)$ is a sectorial operator and derive a representation of the generated analytic semigroup $e^{t\Delta}$ (cf. Engel and Nagel, 2001, Example II.4.9). The symbol

$$a(\xi) = \sum_{|\alpha|=2} a_\alpha i^{|\alpha|} \xi^\alpha = \sum_{j=1}^n \xi_j^2 = |\xi|^2$$

of $-\Delta$ implies

$$\begin{aligned} \xi \neq 0 &\Rightarrow a(\xi) \neq 0, \\ -1 &\notin a(\mathbb{R}^n) \end{aligned}$$

and from (Haase, 2006, Theorem 8.2.1) follows, that $-\Delta$ is a sectorial operator with angle $\delta = \frac{\pi}{2}$, spectrum $\sigma(A) = a(\mathbb{R}^n) = [0, \infty)$ and domain $D(-\Delta) = H^2(\mathbb{R}^n)$. Theorem 2.14 yields that Δ generates the analytic semigroup $\{e^{t\Delta}\}_{t \geq 0}$ that can be continued analytically into the sector $\Sigma_{\frac{\pi}{2}}$. We either obtain a representation of the semigroup by convolution with the Gauss-Weierstrass kernel $G_t \in L^1(\mathbb{R}^n)$,

$$G_t(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}},$$

which yields the Gauss-Weierstrass semigroup

$$e^{t\Delta}x = G_t * x, \quad \operatorname{Re} t > 0, \quad x \in L^2(\mathbb{R}^n)$$

or by using the Fourier transformation $\mathcal{F}: L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$:

$$e^{t\Delta}x = \mathcal{F}^{-1}(e^{-t|\xi|^2} \mathcal{F}(x)), \quad \operatorname{Re} t > 0, \quad x \in L^2(\mathbb{R}^n),$$

cf. Remark 1.7 and (Haase, 2006, Proposition 8.3.1). The solution of the heat equation

$$u_t(t, x) - u_{xx}(t, x) = 0, \quad u(0, x) = u_0(x),$$

with initial data u_0 is then given by

$$u(t, x) = e^{t\Delta}u_0(x) = G_t * u_0(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} u_0(y) dy.$$

2.3 Fractional Powers of Sectorial Operators

For a sectorial operator A with $\operatorname{Re} \sigma(A) > 0$ and $\alpha > 0$, the functional calculus for sectorial operators (see Haase, 2006) yields the following representation for fractional powers $A^{-\alpha}$ of sectorial operators,

$$A^{-\alpha} = \frac{1}{2\pi i} \int_{\Gamma} \lambda^{-\alpha} R(\lambda, A) d\lambda.$$

Since $-A$ is the generator of an analytic semigroup, another representation of $A^{-\alpha}$ can be obtained. We will use that representation in the subsequent.

Definition 2.20. For a sectorial operator $A: D(A) \subset X \rightarrow X$ with $\operatorname{Re} \sigma(A) > 0$ and $\alpha > 0$ we define

$$A^{-\alpha}: D(A^{-\alpha}) \rightarrow X, \quad A^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} t^{\alpha-1} e^{-tA} dt.$$

Theorem 2.21 (Henry, 1981, Theorem 1.4.2). *Let $A: D(A) \subset X \rightarrow X$ be a sectorial operator with $\operatorname{Re} \sigma(A) > 0$ and $\alpha \in (0, 1)$. Then we can represent $A^{-\alpha}$ using the resolvent $R(\lambda, -A)$ as*

$$A^{-\alpha} = \frac{\sin(\pi\alpha)}{\pi} \int_0^{\infty} \lambda^{-\alpha} (\lambda + A)^{-1} d\lambda = \frac{\sin(\pi\alpha)}{\pi} \int_0^{\infty} \lambda^{-\alpha} R(\lambda, -A) d\lambda.$$

We state basic properties of fractional powers of sectorial operators.

Theorem 2.22 (Pazy, 1983, Lemma 2.6.2 to Lemma 2.6.4). *Let $A: D(A) \subset X \rightarrow X$ be a sectorial operator with $\operatorname{Re} \sigma(A) > 0$.*

- (i) *For any $\alpha > 0$, $A^{-\alpha}$ is a bounded linear operator which is one-to-one.*
- (ii) *$A^{-\alpha} A^{-\beta} = A^{-(\alpha+\beta)}$ for all $\alpha, \beta > 0$.*
- (iii) *There exists a constant C such that $\|A^{-\alpha}\|_{\mathcal{L}(X)} \leq C$ for $\alpha \in (0, 1)$.*
- (iv) *$\lim_{\alpha \rightarrow 0} A^{-\alpha} x = x$ for every $x \in X$.*

Corollary 2.23 (Engel and Nagel, 2001, Theorem II.5.29). *We conclude from Theorem 2.22 that the family $\{A^{-t}\}_{t \geq 0}$ is a C_0 -semigroup of bounded linear operators on $D(A)$.*

Definition 2.24. Let $A: D(A) \subset X \rightarrow X$ be a sectorial operator with $\operatorname{Re} \sigma(A) > 0$.

- (i) For $\alpha > 0$ we define

$$A^{\alpha}: D(A^{\alpha}) = R(A^{-\alpha}) \rightarrow X, \quad A^{\alpha} = (A^{-\alpha})^{-1}$$

and $A^{\alpha} = I$ for $\alpha = 0$.

- (ii) For $\alpha \geq 0$ and $x \in D(A^{\alpha})$ the *graph norm* of $D(A^{\alpha})$ is defined as

$$\|x\|_{\alpha} := \|A^{\alpha} x\| := \|x\| + \|A^{\alpha} x\| = \|x\| + \|x\|_{\alpha}.$$

Remark 2.25. For $x \in D(A) \subset D(A^\alpha)$, $\alpha \in (0, 1)$, we have an explicit representation for $A^\alpha x$ using Theorem 2.21,

$$A^\alpha = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \lambda^{\alpha-1} A(\lambda + A)^{-1} d\lambda = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \lambda^{\alpha-1} AR(\lambda, -A) d\lambda.$$

The domain $D(A^\alpha)$ equipped with its graph norm is a Banach space. The graph norm $\|\cdot\|_\alpha$ of $D(A^\alpha)$ is equivalent to the norm $\|\cdot\|_\alpha$ since A^α is invertible. Thus, $D(A^\alpha)$ endowed with $\|\cdot\|_\alpha$ is a Banach space which we denote by X_α . The spaces X_α will provide the basic topology for the solution of semilinear evolution equations. Therefore, we assume without loss of generality $\operatorname{Re} \sigma(A) > 0$ for the rest of this section.

Remark 2.26. If A does not meet the condition $\operatorname{Re} \sigma(A) > 0$, let $X_\alpha = D(A_\omega^\alpha)$, with $A_\omega := A + \omega$ and $\omega > 0$ being the smallest value satisfying $\operatorname{Re} \sigma(A_\omega) > 0$. For different values of ω , the norms $\|\cdot\|_\alpha$ are equivalent (see Henry, 1981, Definition 1.4.7, Theorem 1.4.6).

The above-defined spaces X_α are indeed Banach spaces.

Theorem 2.27 (Henry, 1981, Theorem 1.4.8). *Let $A: D(A) \subset X \rightarrow X$ be a sectorial operator. The space X_α , equipped with the norm $\|\cdot\|_\alpha$, is a Banach space for $\alpha \geq 0$. For $\alpha > \beta \geq 0$, X_α is dense in X_β with continuous inclusion. If A has compact resolvent, the inclusion $X_\alpha \subset X_\beta$ is compact when $\alpha > \beta \geq 0$.*

The following theorem shows a regularization effect of sectorial operators, which we utilize in the proof of the local existence and uniqueness of a solution to semilinear evolution equations see Henry, 1981, Theorem 1.4.3 and 1.4.4:

Theorem 2.28. *Let $A: D(A) \subset X \rightarrow X$ be a sectorial operator with $\operatorname{Re} \sigma(A) > \rho > 0$.*

(i) *For $\alpha \geq 0$, $t > 0$ a constant $C_\alpha < \infty$ exists such that*

$$\|e^{-tA}\|_{\mathcal{L}(X_\alpha)} = \|A^\alpha e^{-tA}\|_{\mathcal{L}(X)} \leq C_\alpha t^{-\alpha} e^{-\rho t} \leq C_\alpha t^{-\alpha},$$

i.e., the operator $A^\alpha e^{-tA}$ is bounded. If $\alpha \in (0, 1]$ and $x \in D(A^\alpha)$ we have

$$\|(e^{-tA} - 1)x\| \leq \frac{1}{\alpha} C_{1-\alpha} t^\alpha \|A^\alpha x\| = \frac{1}{\alpha} C_{1-\alpha} t^\alpha \|x\|_\alpha.$$

(ii) *For $\alpha \in [0, 1]$, $x \in D(A)$ the interpolation inequality*

$$\|x\|_\alpha = \|A^\alpha x\| \leq C \|Ax\|^\alpha \|x\|^{1-\alpha} = C \|x\|_1^\alpha \|x\|^{1-\alpha}$$

holds with a positive C independent of α .

Similar to Sobolev spaces and Theorem 1.15, various embedding properties can be shown for the Banach spaces X_α (see Henry, 1981, Theorem 1.6.1):

Theorem 2.29 (Embedding theorem for X_α). *Assume that $\Omega \subset \mathbb{R}^n$ is an sufficiently smooth open set, $1 \leq p < \infty$, and $A: D(A) \subset X \rightarrow X = L^p(\Omega)$ is a sectorial operator. If $m \geq 1$ exists such that $D(A) = X_1 \subset W^{m,p}(\Omega)$, then we have for $\alpha \in [0, 1]$*

$$\begin{aligned} X_\alpha &\subset W^{k,q}(\Omega) \quad \text{when} \quad k - \frac{n}{q} < m\alpha - \frac{n}{p}, \quad q \geq p, \\ X_\alpha &\subset C^\nu(\Omega) \quad \text{when} \quad 0 \leq \nu < m\alpha - \frac{n}{p}. \end{aligned}$$

Example 2.30. Let $X := L^2(\mathbb{R}^n)$. We calculate the spaces X_α for the negative Laplace operator. Since $0 \in \sigma(-\Delta)$, fractional powers of $-\Delta$ are not defined. Thus, we consider $-\Delta + 1$ instead of $-\Delta$. From Remark 1.7 follows $(1 - \Delta)^\alpha = \mathcal{F}^{-1}(1 + |\xi|^2)^\alpha \mathcal{F}$ (see Haase, 2006, p. 225). Since \mathcal{F} is an isometric isomorphism on $L^2(\mathbb{R}^n)$ (see Werner, 2007, Korollar V.4.14), we have $\|\mathcal{F}^{-1}(1 + |\xi|^2)^\alpha \mathcal{F}u\|^2 = \|(1 - \Delta)^\alpha u\|^2 < \infty$. We define the Bessel potential space

$$\hat{H}^s(\mathbb{R}^n) = \{x \in L^2(\mathbb{R}^n) : \mathcal{F}^{-1}(1 + |\xi|^2)^{\frac{s}{2}} \mathcal{F}x \in L^2(\mathbb{R}^n)\}$$

and see, that if $x \in X_\alpha$, then $x \in \hat{H}^{2\alpha}(\mathbb{R}^n)$. Moreover, we receive that $D(-\Delta) = H^2(\mathbb{R}^n)$ and $D(-\Delta^{\frac{1}{2}}) = H^1(\mathbb{R}^n)$ due to the result of A. P. Calderón, that the Bessel potential spaces for $s \in \mathbb{N}$ coincide with the Sobolev spaces (see Adams and Hedberg, 1996, Theorem 1.2.3).

2.4 Existence and Uniqueness of Solutions for Evolution Equations

We start this section with results on the existence and uniqueness of solutions for linear abstract Cauchy problems using C_0 -semigroups. These results can be extended to semilinear evolution equations if the nonlinear function meets certain prerequisites. We then show that these prerequisites can be weakened for equations that involve sectorial operators if we use the theory of analytic semigroups.

2.4.1 Linear Abstract Cauchy Problems

The homogeneous abstract Cauchy problem

$$u'(t) = Au(t), \quad t \in (0, T), \quad u(0) = u_0 \tag{2.1}$$

is given by a closed and densely defined linear operator $A: D(A) \subset X \rightarrow X$, an initial value $u_0 \in X$ and a time period $(0, T)$, whereby $T = \infty$ is permitted.

Definition 2.31. A function $u \in C([0, T], X) \cap C^1((0, T), X)$ with $u(t) \in D(A)$ for $t \in (0, T)$ satisfying (2.1) on $[0, T)$ is called a (*classical*) *solution* of Equation (2.1).

For $u_0 \in D(A)$ follows from Theorem 2.7 that $u(t) = T(t)u_0$ is a solution of Equation (2.1) with $T(t)$ being the C_0 -semigroup generated by A . It remains to show the uniqueness of the solution: Let \tilde{u} be another solution of Equation (2.1) on $(0, T)$ with the same $u_0 \in D(A)$.

For $0 \leq s \leq t < T$ we define $y(t, s) := T(t - s)\tilde{u}(s)$. Then $s \mapsto y(t, s)$ is continuous on $[0, t]$ and differentiable on $(0, t)$. Since \tilde{u} solves Equation (2.1), we have

$$y_s(t, s) = -AT(t - s)\tilde{u}(s) + T(t - s)\tilde{u}'(s) = -AT(t - s)\tilde{u}(s) + AT(t - s)\tilde{u}(s) = 0$$

for $s \in (0, t]$. Hence, $s \mapsto y(t, s)$ is constant and $y(t, 0) = y(t, t)$ in particular. Thus, $u(t) = T(t)u_0 = T(t)\tilde{u}(0) = y(t, 0) = y(t, t) = \tilde{u}(t)$ (cf. Henry, 1981, p. 50).

Remark 2.32. For $u_0 \notin D(A)$ the existence of a (classical) solution is not guaranteed if A generates a C_0 -semigroup that cannot be extended analytically. In this case, we call $t \rightarrow T(t)u_0$ a *mild solution* (for $u \in D(A)$ the mild solutions is the (classical) solution). If A can be extended analytically, i.e A is a negative sectorial operator, then $u(t) = T(t)u_0$ is the unique (classical) solution of Equation (2.1) for every $x \in X$. We will define the concept of mild solutions in the following for inhomogeneous abstract Cauchy problems and semilinear evolution equations and leave this concept for the homogeneous abstract Cauchy problem without a definition.

We consider in the following the inhomogeneous abstract Cauchy problem

$$u'(t) = Au(t) + f(t), \quad t \in (0, T), \quad u(0) = u_0, \quad (2.2)$$

which involves an inhomogeneous term $f: [0, T] \rightarrow X$.

Let u be a classical solution of Equation (2.2) and $T(t)$ the C_0 -semigroup generated by A . We define $y(s) := T(t - s)u(s)$, then $y'(s) = T(t - s)f(s)$ and if $f \in L^1(0, T; X)$ integration from 0 to t provides

$$u(t) = T(t)u_0 + \int_0^t T(t - s)f(s) ds.$$

As in the homogeneous case, depending on the initial data but also depending on the characteristics of f , a (classical) solution of Equation (2.2) does not necessarily exist. Therefore, we introduce a weaker concept, the mild solution.

Definition 2.33. Let A be the generator of a C_0 -semigroup $T(t)$ on X . Let $u_0 \in X$ and $f \in L^1(0, T; X)$. The function $u \in C([0, T], X)$ given by

$$u(t) = T(t)u_0 + \int_0^t T(t - s)f(s) ds, \quad t \in [0, T],$$

is the *mild solution* of the initial value problem (2.2) on $[0, T]$.

For $f \in L^1(0, T; X)$, Equation (2.2) has a unique mild solution by definition. We will now collect conditions on f such that the mild solution of Equation (2.2) becomes the (classical) solution for every $x \in D(A)$.

Theorem 2.34 (Pazy, 1983, Theorem 4.2.4). *Let A be the generator of a C_0 -semigroup $T(t)$ on X , let $f \in L^1(0, T; X)$ be continuous on $(0, T)$ and define*

$$F(t) = \int_0^t T(t - s)f(s) ds, \quad t \in [0, T].$$

Equation (2.2) has a solution u on $(0, T)$ for every $u_0 \in D(A)$ if one of the following conditions is satisfied:

2. Semigroup Theory

(i) $F \in C^1((0, T), X)$,

(ii) $F(t) \in D(A)$ for $t \in (0, T)$ and $AF(t)$ is continuous on $(0, T)$.

If Equation (2.2) has a solution u on $[0, T)$ for some $u_0 \in D(A)$ then F satisfies both conditions (i) and (ii).

Corollary 2.35 (Pazy, 1983, Corollary 4.2.5). *Let A be the generator of a C_0 -semigroup $T(t)$ on X . If $f \in C^1((0, T), X)$ then Equation (2.2) has a solution u on $[0, T)$ for every $u_0 \in D(A)$.*

Corollary 2.36 (Pazy, 1983, Corollary 4.2.6). *Let A be the generator of a C_0 -semigroup $T(t)$ on X and $f \in L^1(0, T; X) \cap C((0, T), X)$. If $f(t) \in D(A)$ for $t \in (0, T)$ and $Af \in L^1(0, T; X)$ then Equation (2.2) has a solution u on $[0, T)$ for every $u_0 \in D(A)$.*

We can pose some more conditions on f if we consider another concept of solution, the strong solution.

Definition 2.37. Let the function u be differentiable almost everywhere on $[0, T]$ such that $u' \in L^1(0, T; X)$. Then u is called a *strong solution* of Equation (2.2) if $u(0) = u_0$ and $u'(t) = Au(t) + f(t)$ almost everywhere on $[0, T]$.

Theorem 2.38 (Pazy, 1983, Theorem 4.2.9). *Let A be the generator of a C_0 -semigroup $T(t)$ on X , let $f \in L^1(0, T; X)$ and define*

$$F(t) = \int_0^t T(t-s)f(s) \, ds, \quad t \in [0, T].$$

Equation (2.2) has a strong solution u on $[0, T]$ for every $u_0 \in D(A)$ if one of the following conditions is satisfied:

(i) F is differentiable almost everywhere on $[0, T]$ and $F' \in L^1(0, T; X)$,

(ii) $F(t) \in D(A)$ almost everywhere on $[0, T]$ and $AF \in L^1(0, T; X)$.

If Equation (2.2) has a strong solution u on $[0, T]$ for some $u_0 \in D(A)$ then F satisfies both conditions (i) and (ii).

Corollary 2.39 (Pazy, 1983, Corollary 4.2.10). *Let A be the generator of a C_0 -semigroup $T(t)$ on X . If f is differentiable almost everywhere on $[0, T]$ and $f \in L^1(0, T; X)$ then Equation (2.2) has a strong solution u on $[0, T]$ for every $u_0 \in D(A)$.*

Corollary 2.40 (Pazy, 1983, Corollary 4.2.11). *Let X be a reflexive Banach space and let A generate a C_0 -semigroup $T(t)$ on X . If f is Lipschitz continuous on $[0, T]$ then Equation (2.2) has a strong solution u on $[0, T]$ for every $u_0 \in D(A)$.*

As in the homogeneous case, $u_0 \in D(A)$ is necessary for the existence of a classical or strong solution, if A generates a C_0 -semigroup that cannot be extended analytically. However, if A is negative sectorial, i.e., it generates an analytic semigroup, we collect conditions on f under that Equation (2.2) has a unique solution u on $[0, T)$ for all $u_0 \in X$. One condition is, that f is Hölder continuous.

We start with a condition on the modulus of continuity of f .

Theorem 2.41 (Pazy, 1983, Theorem 4.3.1). *Let A be the generator of an analytic semigroup $T(t)$ on X . Let $f \in L^1(0, T; X)$ and suppose that for every $t \in (0, T)$ there is a $\delta_t > 0$ and a continuous function $w_t: [0, \infty) \rightarrow [0, \infty)$ such that*

$$\|f(t) - f(s)\| \leq w_t(|t - s|)$$

and

$$\int_0^{\delta_t} \frac{w_t(\tau)}{\tau} d\tau < \infty.$$

Then the mild solution of Equation (2.2) is a (classical) solution for every $x \in X$.

As a direct consequence, we receive the already indicated Hölder continuity condition.

Corollary 2.42 (Pazy, 1983, Corollary 4.3.3). *Let A be the generator of an analytic semigroup $T(t)$ on X . If $f \in L^1(0, T; X)$ is locally Hölder continuous on $(0, T]$ then Equation (2.2) has a unique solution for every $x \in X$.*

Remark 2.43. The prerequisite on f in the Corollary above can be weakened such that f needs to be integrable on an interval $[0, t_1]$ for some $t_1 > 0$ instead of $f \in L^1(0, T; X)$ (cf. Henry, 1981, Theorem 3.2.2).

Theorem 2.44 (Pazy, 1983, Theorem 4.3.6). *Let A be the generator of an analytic semigroup $T(t)$ on X and let $0 \in \rho(A)$. If $f(t)$ is continuous, $f(t) \in D((-A)^\alpha)$, $\alpha \in (0, 1]$ and $\|f(t)\|_\alpha$ is bounded, then the mild solution of Equation (2.2) is a (classical) solution for every $x \in X$.*

We conclude this subsection with a regularity result for the case where A generates an analytic semigroup and f is Hölder continuous.

Theorem 2.45 (Pazy, 1983, Theorem 4.3.5). *Let A be the generator of an analytic semigroup $T(t)$ on X and let $f \in C^\theta([0, T], X)$. If u is the solution of Equation (2.2) on $[0, T]$ then,*

- (i) for every $\delta > 0$, $Au \in C^\theta([\delta, T], X)$ and $u' \in C^\theta([\delta, T], X)$,
- (ii) if $u_0 \in D(A)$ then Au and u' are continuous on $[0, T]$,
- (iii) if $u_0 = 0$ and $f(0) = 0$ then $Au, u' \in C^\theta([0, T], X)$.

2.4.2 Semilinear Evolution Equations

In the following, let $X := X(\Omega)$ with $\Omega \subset \mathbb{R}^n$ open and sufficiently smooth, $n \in \mathbb{N}$, be a Banach space with norm $\|\cdot\|$. We consider the following semilinear evolution equation:

$$u'(t) + Au(t) = F(t, u(t)), \quad t \in (0, T), \quad u(0) = u_0, \quad (2.3)$$

where $-A: D(A) \subset X \rightarrow X$ is the generator of a C_0 -semigroup and $F: [0, T] \times X \rightarrow X$ is continuous in t and Lipschitz continuous in u .

Analogously to Definition 2.33 we define the *mild solution* of Equation (2.3) as the function $u \in C([0, T], X)$ given by

$$u(t) = T(t)u_0 + \int_0^t T(t-s)F(s, u(s)) \, ds, \quad t \in [0, T].$$

We now state results assuring the existence of mild solutions of Equation (2.3). Essential in those results is the Lipschitz continuity of f .

Theorem 2.46 (Pazy, 1983, Theorem 6.1.2, Corollary 6.1.3). *Let $-A$ be the generator of a C_0 -semigroup $T(t)$ on X and let F be continuous in t on $[0, T]$ and uniformly Lipschitz continuous on X . Then, for every $u_0 \in X$, there exist a unique (global) mild solution $u \in C([0, T], X)$ of Equation (2.3). Moreover, the mapping $u_0 \mapsto u$ is Lipschitz continuous from X into $C([0, T], X)$. In general, the integral equation*

$$w(t) = g(t) + \int_0^t T(t-s)F(s, w(s)) \, ds$$

has a unique solution $w \in C([0, T], X)$ for every $g \in C([0, T], X)$.

The local Lipschitz continuity in u is sufficient for the existence of a unique (local) mild solution.

Theorem 2.47 (Pazy, 1983, Theorem 6.1.4). *Let $-A$ be the generator of a C_0 -semigroup $T(t)$ on X and let $F: [0, \infty) \times X \rightarrow X$ be continuous in t for $t \geq 0$ and locally Lipschitz continuous in u , uniformly in t on bounded intervals. Then for every $u_0 \in X$ there is a $t_{max} \leq \infty$ such that Equation (2.3) has a unique mild solution u on $[0, t_{max}]$. Moreover, if $t_{max} < \infty$ then*

$$\lim_{t \rightarrow t_{max}} \|u(t)\| = \infty.$$

We are now interested in checking whether such a mild solution is also a classical or strong solution. In general, that is not the case.

Theorem 2.48 (Pazy, 1983, Theorem 6.1.5, Theorem 6.1.6). *(i) Let $-A$ be the generator of a C_0 -semigroup $T(t)$ on X . If $F \in C^1([0, T] \times X, X)$, then the mild solution of Equation (2.3) with $u_0 \in D(A)$ is a (classical) solution.*

(ii) If F is just Lipschitz continuous in both t and u and X is reflexive, then the mild solution of Equation (2.3) with $u_0 \in D(A)$ is a strong solution.

We denote by X_1 the domain $D(A)$ endowed with the graph norm $\|\cdot\|_1$ which is a Banach space. Then we receive the following result:

Theorem 2.49 (Pazy, 1983, Theorem 6.1.7). *Let $-A$ be the generator of a C_0 -semigroup $T(t)$ on X .*

- (i) *Let $F: [0, T] \times X_1 \rightarrow X_1$ be uniformly Lipschitz in u and for every $x \in X_1$ let $F(t, x)$ be continuous from $[0, T]$ into X_1 . Then the mild solution of Equation (2.3) with $u_0 \in D(A)$ is a (classical) solution on $[0, T]$.*
- (ii) *If F is just locally Lipschitz continuous in u uniformly in t on $[0, T]$, then the mild solution of Equation (2.3) with $u_0 \in D(A)$ on a maximal interval $[0, t_{max})$ is a (classical) solution. If $t_{max} < \infty$ then*

$$\lim_{t \rightarrow t_{max}} \|u(t)\| + \|Au(t)\| = \infty.$$

Similar to the linear case, we can further reduce the requirements on the regularity of F if we consider analytic semigroups. Therefore, let us now assume that A is a sectorial operator. Thus, $-A$ generates an analytic semigroup.

According to Section 2.3, A^α is well defined for $\alpha \in [0, 1)$ and $X_\alpha = D(A^\alpha)$ is a Banach space dense in X . Note, that if A does not meet the condition $\operatorname{Re} \sigma(A) > 0$, let X_α be the domain $D(A_\omega^\alpha)$ of the operator $A_\omega := A + \omega I$, see Remark 2.26. We now adjust the definition of the classical and mild solution to incorporate the space X_α : A solution $u \in C([0, T], X_\alpha) \cap C^1((0, T), X)$ of Equation (2.3) on $(0, T)$ with $u(t) \in D(A)$ for $t \in (0, T)$ is called a (classical) solution. A function $u \in C((0, T), X_\alpha)$ is called a mild solution of Equation (2.3) on $(0, T)$ if it solves the integral equation

$$u(t) = e^{-tA}u_0 + \int_0^t e^{-(t-s)A}F(s, u(s)) \, ds \quad (2.4)$$

for all $t \in (0, T)$.

We show the existence and uniqueness of a mild solution of Equation (2.3) and specify conditions for the mild solution to be also the classical solution. Therefore, we impose the following prerequisites on F .

We are mostly interested in assumptions on F that are sufficient for the existence of a mild solution of Equation (2.3).

Assumption 2.50. We suppose $F \in L^1((0, T), C(X_\alpha, X))$, i.e., F is a map from $(0, T) \times X_\alpha$ to X which is

- (i) locally integrable with respect to t , i.e., there is $t_0 > 0$ such that

$$\int_0^{t_0} \|F(s, u(s))\| \, ds < \infty,$$

2. Semigroup Theory

- (ii) locally Lipschitz-continuous with respect to u with Lipschitz constant C independent from t , i.e., for all $u_0 \in X_\alpha$ there is a $\delta > 0$ and a constant C such that for all $u, v \in U_\delta(u_0)$ and $t \in (0, T)$, the inequality $\|F(t, u) - F(t, v)\| \leq C\|u - v\|_\alpha$ holds.

Since $D(A) = X_1 \subset X_\alpha = D(A^\alpha)$, F is of lower order than A and thus Equation (2.3) remains semilinear.

Given an initial value $u_0 \in X_\alpha$, the following result provides the local existence of a mild solution:

Theorem 2.51. *Let $-A$ be the generator of an analytic semigroup and $u_0 \in X_\alpha$. Suppose that F satisfies Assumption 2.50. Then there exist a $T' \in (0, T)$ and a unique mild solution $u \in C([0, T'], X_\alpha)$ of Equation (2.3) in $(0, T')$.*

Proof. Most of the literature on semigroups postulate assumptions for the existence of a (classical) solution in the prerequisites of the existence theorem. Therefore, we present the proof for the case that F satisfies only Assumption 2.50.

Let $\delta > \delta' > 0, T'' > 0$ and $M := \sup_{t \in [0, T'']} \|F(t, u_0)\|$. From the local Lipschitz continuity with respect to u follows

$$\|F(t, u_1) - F(t, u_2)\| \leq L\|u_1 - u_2\|_\alpha$$

for $t \in [0, T'']$ and $u_1, u_2 \in X_\alpha$ with $\|u_i - u_0\|_\alpha < \delta$ ($i = 1, 2$). We choose $T' \in (0, T'')$ sufficiently small such that

$$\begin{aligned} \|(e^{-tA} - I)A^\alpha u_0\| &\leq \frac{\delta'}{2}, \quad t \in [0, T'], \\ C_\alpha(L\delta' + M) \int_0^t (t-s)^{-\alpha} ds &\leq \frac{\delta'}{2}, \quad t \in [0, T'] \end{aligned} \tag{2.5}$$

with C_α from Theorem 2.28.

Consider $B := B_{\delta'}(u_0) \subset C([0, T'], X_\alpha) =: Z$. As B is a closed subset of the Banach space Z , it is a Banach space itself (cf. Rudin, 1976, Theorem 3.11). We define the mapping $G: B \rightarrow B$ by

$$G(w)(t) = e^{-tA}u_0 + \int_0^t e^{-(t-s)A}F(s, w(s)) ds.$$

If we can show that G is a contraction, the Banach fixed-point theorem is applicable which then provides a unique mild solution for Equation (2.3).

We first show that G is a self mapping of B and then prove that G is a contraction. For $w \in B$ we have

$$\begin{aligned} \|F(s, w(s))\| &\leq \|F(s, w(s)) - F(s, u_0)\| + \|F(s, u_0)\| \\ &\leq L\|w(s) - u_0\|_\alpha + M \leq L\delta' + M \end{aligned}$$

and it follows that

$$\begin{aligned}
 \|G(w)(t) - u_0\|_\alpha &\leq \|(e^{-tA} - I)u_0\|_\alpha + \int_0^t \|e^{-(t-s)A}F(s, w(s))\|_\alpha \, ds \\
 &\leq \|(e^{-tA} - I)A^\alpha u_0\| + \int_0^t \|A^\alpha e^{-(t-s)A}\|_{\mathcal{L}(X)} \|F(s, w(s))\| \, ds \\
 &\leq \|(e^{-tA} - I)A^\alpha u_0\| + \int_0^t \|A^\alpha e^{-(t-s)A}\|_{\mathcal{L}(X)} (L\delta' + M) \, ds.
 \end{aligned}$$

From Theorem 2.28 and $\operatorname{Re} \sigma(A) > \rho > 0$ we obtain

$$\begin{aligned}
 (L\delta' + M) \int_0^t \|A^\alpha e^{-(t-s)A}\|_{\mathcal{L}(X)} \, ds &\leq (L\delta' + M) \int_0^t C_\alpha (t-s)^{-\alpha} e^{-\rho(t-s)} \, ds \\
 &\leq C_\alpha (L\delta' + M) \int_0^t (t-s)^{-\alpha} \, ds.
 \end{aligned}$$

Finally, Equation (2.5) provides $\|G(w)(t) - u_0\|_\alpha \leq \delta'$ for $t \in [0, T']$. Since $(t-s)^\alpha$ is integrable and $F(\cdot, w(\cdot)): [0, T'] \rightarrow X$ is so by assumption the continuity of $G(w)$ is obtained. Thus, $G(w) \in B$.

It remains to show the contraction property: Let $v, w \in B$ and $t \in [0, T']$. Then

$$\begin{aligned}
 \|G(w)(t) - G(v)(t)\|_\alpha &\leq \int_0^t \|A^\alpha e^{-(t-s)A}\|_{\mathcal{L}(X)} \|F(s, w(s)) - F(s, v(s))\| \, ds \\
 &\leq C_\alpha \int_0^t (t-s)^\alpha \, ds \sup_{s \in [0, T']} \|F(s, w(s)) - F(s, v(s))\| \\
 &\leq C_\alpha L \int_0^t (t-s)^\alpha \, ds \sup_{s \in [0, T']} \|w(s) - v(s)\|_\alpha \\
 &\leq C_\alpha L \int_0^t (t-s)^\alpha \, ds \|w - v\|_Z \\
 &\leq \frac{\delta'}{2} \|w - v\|_Z < \|w - v\|_Z.
 \end{aligned}$$

Hence, G is a contraction and admits a unique fixed-point $G(w) = w$ which is the mild solution of Equation (2.3) by definition of G . \square

The local Lipschitz-continuity yields only the existence of a solution on a rather short time interval $(0, T)$. In general, one can extend the solution to a maximal existence interval $(0, \hat{T})$, i.e., no solution exists on the interval (\hat{T}, \tilde{T}) with $\tilde{T} > \hat{T}$. We state a condition on F such that the solution can either be extended on $(0, \infty)$ or is unbounded:

Theorem 2.52 (Henry, 1981, Theorem 3.3.4). *If the prerequisites of Theorem 2.51 are given and if F maps every closed and bounded set $B \subset U$ onto a bounded set $F(B) \subset X$, then the solution u either has a maximal existence interval $(0, \hat{T})$ with $\hat{T} = \infty$ or there is a sequence T_k with $T_k \nearrow \hat{T}$ and $(T_k, u(T_k)) \rightarrow \partial U$, i.e., $\|u(T_k)\|_\alpha \rightarrow \infty$ for $T_k \nearrow \hat{T}$.*

We reformulate the condition on F so that the solution exists for all times if the condition holds.

Corollary 2.53 (Henry, 1981, Corollary 3.3.5). *Let A and F be given as in Theorem 2.51. If there is a continuous $\kappa: [0, \infty) \rightarrow \mathbb{R}$ with*

$$\|F(t, u)\| \leq \kappa(t)(1 + \|u\|_\alpha),$$

then the solution u exists on $(0, \infty)$.

If we add to the prerequisites of the previous theorems that $F: (0, T) \times X_\alpha \rightarrow X$ is locally Hölder-continuous with respect to t , then the mild solution u of Equation (2.3) is also a (classical) solution (see Henry, 1981, Lemma 3.3.2). Let F be defined on an open set U with $(0, u_0) \in U \subset \mathbb{R} \times X_\alpha$.

Theorem 2.54 (Henry, 1981, Theorem 3.3.3). *Let A be the generator of an analytic semigroup, $\alpha \in [0, 1)$ and let $F: U \rightarrow X$ satisfy Assumption 2.50 and be locally Hölder-continuous with respect to t . Then for each $(0, u_0) \in U$ there exists a $T > 0$ such that Equation (2.3) has a unique (classical) solution $u \in C([0, T], X_\alpha)$ on $(0, T)$ with initial value $u(0) = u_0 \in X_\alpha$.*

With the additional prerequisite that F is locally Hölder-continuous with respect to t , the results from Theorem 2.52 and Corollary 2.53 yield the unique classical solution.

2.4.3 Semilinear Initial Boundary Value Problems

In many applications, evolution equations represent partial differential equations in an abstract setting. The initial value of the partial differential equation serves as the initial value of the evolution equation. However, the evolution equation itself is not equipped with boundary conditions. For a (linear) boundary condition, e.g., Neumann or Dirichlet boundary condition, the domain of the linear operator is restricted to functions satisfying the boundary condition. That approach is not practicable for semilinear boundary conditions. We present a generalized variation of constants formula, derived heuristically in Amann (1986) and proved in Amann (1988), that represents the mild solution satisfying semilinear boundary conditions.

We consider equations of the form

$$u'(t) + Au(t) = F(t, u(t)), \quad Bu(t) = G(t, u(t)), \quad t \in (0, T), \quad u(0) = u_0 \quad (2.6)$$

where $-A$ is the generator of an analytic semigroup, $B: D(B) \rightarrow Y$ is a linear operator with $D(B) = D(A)$, F meets Assumption 2.50 and $G: (0, T) \times X_\alpha(\Gamma) \rightarrow Y$. The Banach space Y is defined over $\Gamma = \partial\Omega$. Its specific definition depends on the definition of X and $D(A)$, e.g., $X = L^2(\Omega)$ and $D(A) = H^2(\Omega)$ yield $Y = H^{1/2}(\Gamma)$ (Amann, 1986, Section 3).

In the case of linear boundary conditions, i.e., $G = 0$, we restrict the domain of A to those functions, that satisfy $Bu = 0$ and consider the operator $A_B := A|_{\ker(B)}: \ker(B) \rightarrow X$ with domain $D(A_B) = \{u \in D(A); Bu = 0\}$. Then, the mild solution of

$$u'(t) + A_B u(t) = F(t, u(t)), \quad t \in (0, T), \quad u(0) = u_0$$

is given by Equation (2.4) with A replaced by A_B . For semilinear boundary conditions, we assume that both operators share the same domain

$$W := D(A) = D(B) \hookrightarrow X,$$

which is by Theorem 2.27 a Banach space that is continuously embedded in X . We write W_A and W_B instead of $\ker(A)$ and $\ker(B)$, respectively. Moreover, we assume that B is surjective and A_B is an isomorphism. Amann shows that $R_B := (B|_{W_A})^{-1} : Y \rightarrow W_A$ is an isomorphism and that the mild solution of Equation (2.6) is given by

$$u(t) = e^{-tA_B} u_0 + \int_0^t e^{-(t-s)A_B} \left(F(s, u(s)) - A_B R_B G(s, u(s)) \right) ds.$$

A detailed proof can be found in Amann (1988).

Necessary Optimality Conditions

We consider optimal control problems of the form

$$\begin{aligned} \min_{\lambda \in \Lambda} J(u, \lambda) &= \int_0^T g(t, u(t), \lambda(t)) dt + h(u(T)) \\ \text{s.t. } G(u, \lambda) &= 0, \quad (u, \lambda) \in Z \times \Lambda, \end{aligned}$$

where $G = 0$ satisfies a semilinear evolution equation

$$u'(t) + Au(t) = F(u(t), \lambda(t)), \quad t \in (0, T), \quad u(0) = u_0, \quad (3.1)$$

with an abstract function u with $u(t) \in X$ and control function $\lambda \in \Lambda$. We present assumptions under which necessary first order optimality conditions can be formulated in Banach spaces. The formulation of those conditions itself follows the adjoint approach.

To ensure the well-posedness of the control problem, we apply the theory presented in Chapter 2. Therefore, let $A: D(A) \subset X \rightarrow X$ be a sectorial operator and hence, the generator of an analytic semigroup. According to Remark 2.26, we can define A such that the real part of its spectrum, $\text{Re } \sigma(A)$, is positive. We can define $X_\alpha = D(A^\alpha)$ for $\alpha \in [0, 1)$ and obtain the Banach space $(X_\alpha, \|\cdot\|_\alpha)$ with $\|u\|_\alpha := \|A^\alpha u\|$, see Section 2.3 for details. Let

$$F: X_\alpha \times \mathbb{R}^d \rightarrow X \quad (3.2)$$

be a semilinear mapping depending on the values of the state $u(t)$ and the control $\lambda(t)$. We know from Section 2.4.2 that the local solution u for Equation (3.1) with initial value $u_0 \in X_\alpha$ is contained in $C^0([0, T], X_\alpha)$ and $u(t) \in D(A)$ for $t \in [0, T]$. If we can show that the local solution is also a global solution that can be extended to the nonnegative real axis, the above statements hold for the compact interval $[0, T]$ (see Corollary 2.53).

3.1 Necessary Optimality Conditions in Banach Spaces

In the following, let X be a real Banach space over a bounded domain $\Omega \in \mathbb{R}^n$ with norm $\|\cdot\|$ and L be a real Banach spaces over a given time interval $[0, T]$ with norm $\|\cdot\|_L$. X is

3. Necessary Optimality Conditions

the range space of the abstract function u , L is the space of the control function λ and $\Lambda \subset L = L^1([0, T], \mathbb{R}^d)$ is a nonempty, closed and convex set of admissible control functions. Moreover, let an initial value $u_0 \in X_\alpha$ be given.

Let $Z := C([0, T], X_\alpha)$ be the Banach space of continuous functions $u: [0, T] \rightarrow X_\alpha$ and $Z' := C([0, T], X)$, both equipped with their corresponding uniform norm,

$$\|u\|_Z = \sup_{t \in [0, T]} \|u(t)\|_\alpha \text{ and } \|u\|_{Z'} = \sup_{t \in [0, T]} \|u(t)\|,$$

respectively.

We obtain the control problem

$$\begin{aligned} \min_{\lambda \in \Lambda} J(u, \lambda) &= \int_0^T g(t, u(t), \lambda(t)) dt + h(u(T)) \\ \text{s.t. } G(u, \lambda) &= 0, \quad (u, \lambda) \in Z \times \Lambda, \end{aligned} \quad (3.3)$$

with a continuous and convex objective functional $J: Z \times L \rightarrow \mathbb{R}$ with terminal costs function $h: X_\alpha \rightarrow \mathbb{R}$ and running costs function $g: [0, T] \times X_\alpha \times \mathbb{R}^d \rightarrow \mathbb{R}$. The constraint $G: Z \times L \rightarrow Z$,

$$G(u, \lambda)(t) = u(t) - e^{-tA}u_0 - \int_0^t e^{-(t-s)A}F(u(s), \lambda(s)) ds = 0,$$

represents the mild solution of Equation (3.1).

To formulate the necessary optimality conditions, we assume the following:

Assumption 3.1. (i) J is Fréchet-differentiable on $Z \times L$.

(ii) F is continuously Fréchet-differentiable on $X_\alpha \times \mathbb{R}^d$.

(iii) F satisfies Assumption 2.50, i.e., F is locally integrable with respect to t and locally Lipschitz-continuous with respect to u .

(iv) There exists a $\kappa \in C([0, \infty), \mathbb{R})$ such that for every $t > 0$ and $u(t) \in X_\alpha$

$$\|\tilde{F}(t, u(t))\| = \|F(u(t), \lambda(t))\| \leq \kappa(t)(1 + \|u(t)\|_\alpha),$$

with $\tilde{F}: [0, T] \times X_\alpha \rightarrow X$ (see the prerequisites of Corollary 2.53).

Remark 3.2. Assumption 3.1 (iii) yields the local existence and uniqueness of a mild solution of (3.1). Indeed, we define $\tilde{F}_\lambda: (0, T) \times X_\alpha \rightarrow X$, $\tilde{F}_\lambda(t, u(t)) = F(u(t), \lambda(t))$ for fixed $\lambda \in L$ and consider the evolution equation $u'(t) + Au(t) = \tilde{F}_\lambda(t, u(t))$ which coincides with Equation (2.3). The global existence follows directly from Assumption 3.1 (iv) by the same argument as above.

Theorem 3.3. Assumption 3.1 (ii) yields the continuous Fréchet-differentiability of G on $Z \times L$ with derivative $dG(u, \lambda): Z \times L \rightarrow Z$,

$$dG(u, \lambda)[\Delta u, \Delta \lambda](t) = \Delta u(t) - \int_0^t e^{-(t-s)A} dF(u(s), \lambda(s))[\Delta u(s), \Delta \lambda(s)] ds. \quad (3.4)$$

Proof. G is Fréchet-differentiable in (u, λ) if and only if

$$\frac{1}{\|(\Delta u, \Delta \lambda)\|_{Z \times L}} \|G(u + \Delta u, \lambda + \Delta \lambda) - G(u, \lambda) - dG(u, \lambda)[\Delta u, \Delta \lambda]\|_Z \rightarrow 0$$

follows from $\|(\Delta u, \Delta \lambda)\|_{Z \times L} \rightarrow 0$.

For $(\Delta u, \Delta \lambda) \in Z \times \Lambda$ we obtain

$$\begin{aligned} & G(u + \Delta u, \lambda + \Delta \lambda)(t) - G(u, \lambda)(t) \\ &= u(t) + \Delta u(t) + e^{-tA}u_0 - \int_0^t e^{-(t-s)A}F(u(s) + \Delta u(s), \lambda(s) + \Delta \lambda(s)) \, ds \\ &\quad - u(t) - e^{-tA}u_0 + \int_0^t e^{-(t-s)A}F(u(s), \lambda(s)) \, ds \\ &= \Delta u(t) - \int_0^t e^{-(t-s)A} \left(F(u(s) + \Delta u(s), \lambda(s) + \Delta \lambda(s)) - F(u(s), \lambda(s)) \right) \, ds. \end{aligned}$$

Together with (3.4), this yields

$$\tilde{G}(t) = - \int_0^t e^{-(t-s)A} \tilde{F}(s) \, ds,$$

with

$$\tilde{G}(t) := G(u + \Delta u, \lambda + \Delta \lambda)(t) - G(u, \lambda)(t) - dG(u, \lambda)[\Delta u, \Delta \lambda](t)$$

and

$$\tilde{F}(s) := F(u(s) + \Delta u(s), \lambda(s) + \Delta \lambda(s)) - F(u(s), \lambda(s)) - dF(u(s), \lambda(s))[\Delta u(s), \Delta \lambda(s)].$$

We consider on both sides of the equation the corresponding norm and apply Theorem 2.28. Thus, we can estimate

$$\begin{aligned} \|\tilde{G}\|_Z &= \left\| \int_0^{\cdot} e^{-(\cdot-s)A} \tilde{F}(s) \, ds \right\|_Z = \sup_{t \in [0, T]} \left\| \int_0^t e^{-(t-s)A} \tilde{F}(s) \, ds \right\|_\alpha \\ &\leq \sup_{t \in [0, T]} \left| \int_0^t \|e^{-(t-s)A} \tilde{F}(s)\|_\alpha \, ds \right| \leq \sup_{t \in [0, T]} \left| \int_0^t \|A^\alpha e^{-(t-s)A} \tilde{F}(s)\| \, ds \right| \\ &\leq \sup_{t \in [0, T]} \left| \int_0^t \|A^\alpha e^{-(t-s)A}\|_{\mathcal{L}(X)} \|\tilde{F}(s)\| \, ds \right| \stackrel{2.28(i)}{\leq} \sup_{t \in [0, T]} \left| \int_0^t C(t-s)^{-\alpha} \|\tilde{F}(s)\| \, ds \right| \\ &\leq \sup_{t \in [0, T]} \sup_{s \in [0, t]} \|\tilde{F}(s)\| \int_0^t C(t-s)^{-\alpha} \, ds \leq \sup_{s \in [0, T]} \|\tilde{F}(s)\| \underbrace{\sup_{t \in [0, T]} \int_0^t C(t-s)^{-\alpha} \, ds}_{=: \tilde{C} < \infty} \\ &\leq \tilde{C} \|\tilde{F}\|_{Z'}. \end{aligned}$$

The estimate for \tilde{C} holds due to the integrability of $C(t-s)^{-\alpha}$ for $\alpha \in [0, 1)$. Assumption 3.1 (ii) and (Henry, 1981, Lemma 3.4.3) yield the Fréchet-differentiability of

3. Necessary Optimality Conditions

$\mathcal{F}: Z \times L \rightarrow Z'$, $\mathcal{F}(u, \lambda)(t) = F(u(t), \lambda(t))$ on $Z \times L$. Finally, we have

$$\begin{aligned} & \frac{1}{\|(\Delta u, \Delta \lambda)\|_{Z \times L}} \|G(u + \Delta u, \lambda + \Delta \lambda) - G(u, \lambda) - dG(u, \lambda)[\Delta u, \Delta \lambda]\|_Z \\ & \leq \frac{\tilde{C}}{\|(\Delta u, \Delta \lambda)\|_{Z \times L}} \|\mathcal{F}(u + \Delta u, \lambda + \Delta \lambda) - \mathcal{F}(u, \lambda) - d\mathcal{F}(u, \lambda)[\Delta u, \Delta \lambda]\|_{Z'} \\ & \longrightarrow 0, \quad \|(\Delta u, \Delta \lambda)\|_{Z \times L} \rightarrow 0 \end{aligned}$$

□

Remark 3.4. To be precise, we observe that (Henry, 1981, Lemma 3.4.3) only yields the differentiability of \mathcal{F} on Z . The extension of that result to the differentiability on $Z \times L$ can easily be shown. For a given $(u, \lambda) \in Z \times L$ we have

$$\begin{aligned} \|(u(t), \lambda(t))\|_{X_\alpha \times \mathbb{R}^d}^2 &= \|u(t)\|_\alpha^2 + |\lambda(t)|^2 \leq \sup_{t \in [0, T]} \|u(t)\|_\alpha^2 + \sup_{t \in [0, T]} |\lambda(t)|^2 \\ &= \|u\|_Z^2 + \|\lambda\|_L^2 = \|(u, \lambda)\|_{Z \times L}^2 \end{aligned}$$

for all $t \in [0, T]$. And further follows

$$\begin{aligned} \frac{1}{\|(\Delta u, \Delta \lambda)\|_{Z \times L}} \|\tilde{\mathcal{F}}(t)\| &\leq \frac{1}{\|(\Delta u(t), \Delta \lambda(t))\|_{X_\alpha \times \mathbb{R}^d}} \|\tilde{\mathcal{F}}(t)\| \longrightarrow 0, \\ &\|(\Delta u, \Delta \lambda)\|_{X_\alpha \times \mathbb{R}^d} \rightarrow 0, \quad (3.5) \end{aligned}$$

with $\tilde{\mathcal{F}}(t)$ being defined analogously to $\tilde{G}(t)$ and $\tilde{F}(s)$.

Equation (3.5) still holds if we consider the supremum of all $t \in [0, T]$ due to the compactness of $[0, T]$. The same argument yields $\|(\Delta u, \Delta \lambda)\|_{Z \times L} \rightarrow 0$ from $\|(\Delta u, \Delta \lambda)\|_{X_\alpha \times \mathbb{R}^d} \rightarrow 0$. Both result in

$$\frac{1}{\|(\Delta u, \Delta \lambda)\|_{Z \times L}} \|\tilde{\mathcal{F}}(t)\|_{Z'} \longrightarrow 0, \quad \|(\Delta u, \Delta \lambda)\|_{Z \times L} \rightarrow 0.$$

We obtain for the partial derivative with respect to u the following result.

Theorem 3.5. *Let Assumption 3.1 (ii) be true. Then the partial derivative*

$$d_u G(u, \lambda): Z \rightarrow Z, \quad d_u G(u, \lambda)[v] = v - \int_0^\cdot e^{-(\cdot-s)A} d_u F(u(s), \lambda(s))[v(s)] ds,$$

is an isomorphism.

Proof. Consider for given right-hand side $w \in Z$ the inhomogeneous equation

$$v = \int_0^\cdot e^{-(\cdot-s)A} d_u F(u(s), \lambda(s))[v(s)] ds + w. \quad (3.6)$$

If $w = 0$, then the uniqueness of solutions yields that $v = 0$ is the only solution. For $0 \neq w \in Z$ we look at

$$y = \int_0^{\cdot} e^{-(\cdot-s)A} d_u F(u(s), \lambda(s)) [y(s) + w(s)] ds. \quad (3.7)$$

Define the mapping $\tilde{F}: (0, T) \times X_\alpha \rightarrow X$, $\tilde{F}(t, y) = d_u F(u, \lambda)[y + w(t)]$, which is continuous and linear, hence, Lipschitz-continuous in y . Thus, a unique mild solution $y \in Z$ exists and replacing $v := y + w$ in (3.7) shows that v satisfies Equation (3.6). Therefore, $d_u G(u, \lambda)$ is an isomorphism. \square

We derive the necessary optimality conditions based on a result in Zowe and Kurcyusz (1979).

Theorem 3.6. *Let Assumption 3.1 be true for the optimal state and control $(\bar{u}, \bar{\lambda}) \in Z \times \Lambda$, then*

$$\langle d_\lambda J(\bar{u}, \bar{\lambda}) + d_\lambda G(\bar{u}, \bar{\lambda})^* [l], \lambda - \bar{\lambda} \rangle_{L^*, L} \geq 0, \quad \lambda \in \Lambda, \quad (3.8)$$

where $l \in Z^*$ is the unique solution of the adjoint equation

$$d_u G(\bar{u}, \bar{\lambda})^* [l] = -d_u J(\bar{u}, \bar{\lambda}). \quad (3.9)$$

Alternatively, $l \in Z^*$ is uniquely defined by

$$l(v) = -\langle d_u J(\bar{u}, \bar{\lambda}), d_u G(\bar{u}, \bar{\lambda})^{-1} [v] \rangle_{Z^*, Z} \quad \text{for all } v \in Z. \quad (3.10)$$

Proof. The authors consider in Zowe and Kurcyusz (1979) the following optimization problem

$$\min f(x), \quad x \in C, \quad g(x) \in K$$

where $f: X \rightarrow \mathbb{R}$, $g: X \rightarrow Y$ with a closed convex set $C \subset X$ and a cone $K \subset Y$. In our case this reads as

$$\min J(u, \lambda), \quad \lambda \in \Lambda, \quad G(u, \lambda) = 0,$$

with $J: Z \times L \rightarrow \mathbb{R}$, $G: Z \times L \rightarrow Z$, $\Lambda \subset L$ is a closed convex set and

$$X = Z \times L, \quad Y = Z, \quad C = Z \times \Lambda, \quad K = \{0_Z\}.$$

A point \bar{x} is a regular point in the sense of Zowe and Kurcyusz, if

$$g'(\bar{x})C(\bar{x}) - K(g(\bar{x})) = Y$$

where

$$C(\bar{x}) = \{\lambda(c - \bar{x}) : c \in C, \lambda \geq 0\}, \quad K(y) = \{k - \lambda y : k \in K, \lambda \geq 0\}.$$

In our case we have

$$C(\bar{u}, \bar{\lambda}) = Z \times \Lambda(\bar{\lambda}), \quad K(G(\bar{u}, \bar{\lambda})) = 0_Z.$$

3. Necessary Optimality Conditions

This means that we have a regular point $(\bar{u}, \bar{\lambda})$, if

$$d_u G(\bar{u}, \bar{\lambda})[Z] + d_\lambda G(\bar{u}, \bar{\lambda})[\Lambda(\bar{\lambda})] = Z.$$

In other words, for each $v \in Z$ we have to find a $u \in Z$ and $\mu \in \Lambda(\bar{\lambda})$ such that

$$d_u G(\bar{u}, \bar{\lambda})[u](t) = v(t) - d_\lambda G(\bar{u}, \bar{\lambda})[\mu](t). \quad (3.11)$$

Choosing, for example, $\mu = 0$, this holds due to Theorem 3.5.

By (Zowe and Kurcyusz, 1979, Theorem 4.1) there exists a Lagrange multiplier $y^* \in Y^*$, i.e., by definition (Zowe and Kurcyusz, 1979, (1.1)), we have

- (i) $y^* \in K^+ = \{y^* \in Y^* : \langle y^*, k \rangle \geq 0 \text{ for all } k \in K\}$,
- (ii) $\langle y^*, g(\bar{x}) \rangle = 0$,
- (iii) $f'(\bar{x}) - g'(\bar{x})^* y^* \in C(\bar{x})^+$.

In our case this leads to an $\tilde{l} \in K^+ = Z^*$ where (ii) holds trivially since $g(\bar{x}) = 0$. For the third condition note that

$$C(\bar{u}, \bar{\lambda})^+ = Z^+ \times \Lambda(\bar{\lambda})^+ = \{0_Z\} \times \{\lambda^* \in L^* : \langle \lambda^*, \lambda - \bar{\lambda} \rangle \geq 0 \text{ for all } \lambda \in \Lambda\}.$$

Hence,

$$d_u J(\bar{u}, \bar{\lambda}) - d_u G(\bar{u}, \bar{\lambda})^*[\tilde{l}] = 0$$

and

$$\langle d_\lambda J(\bar{u}, \bar{\lambda}) - d_\lambda G(\bar{u}, \bar{\lambda})^*[\tilde{l}], \lambda - \bar{\lambda} \rangle_{L^*, L} \geq 0 \text{ for all } \lambda \in \Lambda.$$

Setting $l = -\tilde{l} \in K^- = Z^*$ concludes the proof of Equations (3.8) and (3.9).

With u and v from Equation (3.11) (with $\mu = 0$) and the invertibility of $d_u G(\bar{u}, \bar{\lambda})$ we obtain

$$\begin{aligned} -\langle d_u J(\bar{u}, \bar{\lambda}), d_u G(\bar{u}, \bar{\lambda})^{-1}[v] \rangle_{Z^*, Z} &= -\langle d_u J(\bar{u}, \bar{\lambda}), u \rangle_{Z^*, Z} \\ &= \langle d_u G(\bar{u}, \bar{\lambda})^*[l], u \rangle_{Z^*, Z} \\ &= \langle l, d_u G(\bar{u}, \bar{\lambda})[u] \rangle_{Z^*, Z} \\ &= \langle l, v \rangle_{Z^*, Z} \\ &= l(v). \end{aligned}$$

□

Remark 3.7. Zowe and Kurcyusz provide in (Zowe and Kurcyusz, 1979, Theorem 4.1) the existence of a nonempty bounded set of Lagrange multipliers for a regular point \bar{x} . To apply their result to our control problem (3.3), the surjectivity of $d_u G(\bar{u}, \bar{\lambda})$ would be sufficient. However, the injectivity implies the uniqueness of the Lagrange multiplier $l \in Z^*$.

Remark 3.8. A different approach to retrieve the Lagrange multiplier l is using the implicit function theorem: Given Assumption 3.1, we can apply the implicit function theorem on a pair $(\bar{u}, \bar{\lambda}) \in Z \times \Lambda$ with $G(\bar{u}, \bar{\lambda}) = 0$. Thus, there exist $\epsilon, \delta > 0$, neighborhoods $U_\delta(\bar{\lambda}) \subset \Lambda$ and $U_\epsilon(\bar{u}) \subset Z$ and a continuously Fréchet-differentiable function $s: U_\delta(\bar{\lambda}) \rightarrow U_\epsilon(\bar{u})$ with derivative $ds(\lambda) \in \mathcal{L}(\Lambda, Z)$ such that $u = s(\lambda) \in U_\epsilon(\bar{u})$ is the unique solution of the constraint $G(u, \lambda) = 0$ and for $\lambda \in U_\delta(\bar{\lambda})$ the equation

$$d_u G(s(\lambda), \lambda) \circ ds(\lambda) = -d_\lambda G(s(\lambda), \lambda) \quad (3.12)$$

holds. Hence, we can represent (3.3) as an unconstrained control problem

$$\min_{\lambda \in \Lambda} f(\lambda) = J(s(\lambda), \lambda).$$

The first derivative of f includes the derivative of s , which is only given by an implicit definition. However, with the use of (3.12), a representation of the linear functional $df(\lambda)$ that does not rely on the implicit function s can be derived. That representation concurs with the result of Theorem 3.6, cf. Tröltzsch (1984) and Tröltzsch (2009).

For the optimal pair $(\bar{u}, \bar{\lambda})$ we write in the following $G_u(t), G_\lambda(t), F_u(t), F_\lambda(t), J_u(t), J_\lambda(t), g_u(t), g_\lambda(t), \nabla h$ instead of $d_u G(\bar{u}, \bar{\lambda})(t), d_\lambda G(\bar{u}, \bar{\lambda})(t), d_u F(t, \bar{u}(t), \bar{\lambda}(t)), d_\lambda F(t, \bar{u}(t), \bar{\lambda}(t)), d_u J(\bar{u}, \bar{\lambda})(t), d_\lambda J(\bar{u}, \bar{\lambda})(t), d_u g(t, \bar{u}(t), \bar{\lambda}(t)), d_\lambda g(t, \bar{u}(t), \bar{\lambda}(t)), dh(\bar{u}(T))$, respectively. Furthermore, $\langle \cdot, \cdot \rangle$ is the duality product between X_α^* and X_α , $\langle \cdot, \cdot \rangle_{X_\alpha^*, X_\alpha}$, i.e., $\langle F, u \rangle$ refers to the value of $F \in X_\alpha^*$ applied to $u \in X_\alpha$, and (\cdot, \cdot) is the scalar product in \mathbb{R}^d .

In the following theorem, we give a representation of the linear functional $l \in Z^*$.

Theorem 3.9. *The linear functional $l \in Z^*$ in the setting outlined above is given by*

$$l(v) = - \int_0^T \langle F_u(t)^*[p(t)], v(t) \rangle dt - \langle \nabla h, v(T) \rangle - \int_0^T \langle g_u(t), v(t) \rangle dt \quad (3.13)$$

where $p \in Z^*$,

$$p(s) = e^{-(T-s)A^*} \nabla h + \int_s^T e^{-(t-s)A^*} (F_u(t)^*[p(t)] + g_u(t)) dt, \quad (3.14)$$

is the (mild) solution of the terminal value evolution equation

$$-p'(t) + A^*p(t) = F_u(t)^*[p(t)] + g_u(t), \quad t \in (0, T), \quad p(T) = \nabla h. \quad (3.15)$$

Proof. Let $v \in Z$ be given and let $u \in Z$ be the solution of

$$u(t) = \int_0^t e^{-(t-s)A} F_u(s)[u(s)] ds + v(t). \quad (3.16)$$

From Equation (3.10) follows

$$l(v) = -\langle d_u J(\bar{u}, \bar{\lambda}), u \rangle_{Z^*, Z} = - \int_0^T \langle g_u(t), u(t) \rangle dt - \langle \nabla h, u(T) \rangle. \quad (3.17)$$

3. Necessary Optimality Conditions

We use Equation (3.16) to show that Equation (3.13) is equal to Equation (3.17).

$$\begin{aligned}
l(v) &= - \int_0^T \langle F_u(t)^* [p(t)], v(t) \rangle dt - \langle \nabla h, v(T) \rangle - \int_0^T \langle g_u(t), v(t) \rangle dt \\
&= \int_0^T \langle -F_u(t)^* [p(t)], u(t) - \int_0^t e^{-(t-s)A} F_u(s) [u(s)] ds \rangle dt \\
&\quad + \langle \nabla h, \int_0^T e^{-(T-s)A} F_u(s) [u(s)] ds \rangle + \int_0^T \langle g_u(t), \int_0^t e^{-(t-s)A} F_u(s) [u(s)] ds \rangle dt \\
&\quad - \langle \nabla h, u(T) \rangle - \int_0^T \langle g_u(t), u(t) \rangle dt.
\end{aligned} \tag{3.18}$$

We isolate u in every term of Equation (3.18) and use Equation (3.10). We have

$$\langle \nabla h, \int_0^T e^{-(T-t)A} F_u(t) [u(t)] dt \rangle = \int_0^T \langle F_u(t)^* [e^{-(T-t)A^*} \nabla h], u(t) \rangle dt.$$

Fubini's theorem yields

$$\begin{aligned}
&\int_0^T \langle F_u(t)^* [p(t)], \int_0^t e^{-(t-s)A} F_u(s) [u(s)] ds \rangle dt \\
&= \int_0^T \langle \int_t^T e^{-(s-t)A^*} F_u(s)^* [p(s)] ds, F_u(t) [u(t)] \rangle dt \\
&= \int_0^T \langle F_u(t)^* [\int_t^T e^{-(s-t)A^*} F_u(s)^* [p(s)] ds], u(t) \rangle dt
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^T \langle g_u(t), \int_0^t e^{-(t-s)A} F_u(s) [u(s)] ds \rangle dt \\
&= \int_0^T \langle \int_t^T e^{-(s-t)A^*} g_u(s) ds, F_u(t) [u(t)] \rangle dt \\
&= \int_0^T \langle F_u(t)^* [\int_t^T e^{-(s-t)A^*} g_u(s) ds], u(t) \rangle dt.
\end{aligned}$$

Hence,

$$\begin{aligned}
l(v) &= - \langle \nabla h, u(T) \rangle - \int_0^T \langle g_u(t), u(t) \rangle dt \\
&\quad + \int_0^T \langle F_u(t)^* [\{ -p(t) + e^{-(T-t)A^*} \nabla h \\
&\quad\quad\quad + \int_t^T e^{-(s-t)A^*} (F_u(s)^* [p(s)] + g_u(s)) ds \}], u(t) \rangle dt.
\end{aligned} \tag{3.19}$$

Comparing (3.19) to (3.17), we receive that the term in curly brackets must be equal to zero, which is true if and only if Equation (3.14) holds.

We have not yet defined a solution concept for an adjoint system. In Tröltzsch (1989), the author shows in Theorem 3 that p is a mild solution of (3.15) in the sense that $q(t) = p(T-t)$ is a mild solution to the corresponding initial value evolution equation

$$q'(t) + A^*q(t) = F_u(T-t)^*[q(t)] + g_u(T-t), \quad t \in (0, T), \quad q(0) = \nabla h.$$

□

We can now insert the result of Theorem 3.9 into Equation (3.8) and state the necessary optimality conditions for an optimal pair $(\bar{u}, \bar{\lambda}) \in Z \times \Lambda$.

Theorem 3.10. *Let Assumption 3.1 be true for the optimal state and control $(\bar{u}, \bar{\lambda}) \in Z \times \Lambda$, then*

$$\int_0^T (F_\lambda(t)^*[p(t)] + g_\lambda(t), \lambda(t) - \bar{\lambda}(t)) dt \geq 0, \quad \lambda \in \Lambda, \quad (3.20)$$

where $p \in Z^*$ satisfies Equation (3.14).

Proof. We set $\mu = \lambda - \bar{\lambda}$ and calculate $l(G_\lambda[\mu])$:

$$\begin{aligned} l(G_\lambda[\mu]) &= \int_0^T \langle F_u(t)^*[p(t)], \int_0^t e^{-(t-s)A} F_\lambda(s)[\mu(s)] ds \rangle dt \\ &\quad + \langle \nabla h, \int_0^T e^{-(T-s)A} F_\lambda(s)[\mu(s)] ds \rangle + \int_0^T \langle g_u(t), \int_0^t e^{-(t-s)A} F_\lambda(s)[\mu(s)] ds \rangle dt \\ &= \int_0^T (F_\lambda(t)^*[\int_t^T e^{-(s-t)A^*} F_u(s)^*[p(s)] ds], \mu(t)) dt \\ &\quad + \int_0^T (F_\lambda(t)^*[e^{-(T-t)A^*} \nabla h + \int_t^T e^{-(s-t)A^*} g_u(s) ds], \mu(t)) dt \\ &= \int_0^T (F_\lambda(t)^*[e^{-(T-t)A^*} \nabla h + \int_t^T e^{-(s-t)A^*} (F_u(s)^*[p(s)] + g_u(s)) ds], \mu(t)) dt \\ &= \int_0^T (F_\lambda(t)^*[p(t)], \mu(t)) dt. \end{aligned}$$

It is easy to see that $\langle J_\lambda, \mu \rangle_{L^*, L} = \int_0^T (g_\lambda(t), \mu(t)) dt$, which concludes the proof. □

3.2 Intermezzo on Semilinear Boundary Control

In this section we extend the control problem (3.3) with boundary control and derive necessary optimality conditions for the boundary control problem. Such a boundary control problem is the following model problem (cf. Tröltzsch, 1989):

$$\begin{aligned} \min J(u, \lambda) &= \int_0^T \int_\Omega \tilde{g}(t, u(t, x), \lambda(t)) dx dt \\ &\quad + \int_0^T \int_\Gamma \tilde{g}^b(t, u(t, x), \lambda^b(t)) dS_x dt + h(u(T, \cdot)) \end{aligned} \quad (3.21)$$

3. Necessary Optimality Conditions

subject to

$$\begin{aligned} u_t(t, x) &= (\Delta u)(t, x) + f(u(t, x), \lambda(t)) \quad \text{in } (0, T] \times \Omega, \\ u(0, x) &= u_0(x) \quad \text{on } \Omega, \\ \frac{\partial u(t, x)}{\partial n} &= b(u(t, x), \lambda^b(t)) \quad \text{on } (0, T] \times \Gamma, \end{aligned} \tag{3.22}$$

with control functions $\lambda, \lambda^b \in \mathcal{R}([0, T], \mathbb{R})$ bounded by upper and lower bounds,

$$\lambda \leq \lambda(t) \leq \bar{\lambda}, \quad \lambda^b \leq \lambda^b(t) \leq \bar{\lambda}^b. \tag{3.23}$$

Here, $\mathcal{R}([0, T], \mathbb{R})$ denotes the space of regulated functions. Let $\Omega \in \mathbb{R}^n$ be a bounded domain with sufficient smooth boundary Γ . We denote by Δ the Laplace operator and $\partial u / \partial n$ is the conormal derivative. The terminal costs $h \in C^1(L^p(\Omega), \mathbb{R})$ and running costs $\tilde{g}: [0, T] \times \mathbb{R} \times [\lambda, \bar{\lambda}] \rightarrow \mathbb{R}$ are real functions analogously to the previous section. The real function $\tilde{g}^b: [0, T] \times \mathbb{R} \times [\lambda^b, \bar{\lambda}^b] \rightarrow \mathbb{R}$ represents running costs on the boundary. The semilinear real functions f and b are defined on $\mathbb{R} \times [\lambda, \bar{\lambda}]$ and $\mathbb{R} \times [\lambda^b, \bar{\lambda}^b]$, respectively.

The solution to the problem (3.22) is the mild solution $u \in C([0, T], W^{\sigma,p}(\Omega))$, where p and σ are chosen such that $p > n - 1$ and $n/p < \sigma < 1 + 1/p$. The existence of such a solution is guaranteed by the following assumption.

Assumption 3.11. The functions $\tilde{g}, \tilde{g}^b, f, b$ satisfy the Carathéodory type condition: They are continuously partially differentiable with respect to u, λ, λ^b for fixed (t, x) and t and they and their derivatives are measurable with respect to (t, x) and t for fixed (u, λ) and (u, λ^b) . Moreover, these functions and their derivatives are supposed to be bounded if (u, λ) and (u, λ^b) run through a bounded subset of \mathbb{R}^2 (cf. Tröltzsch, 1989).

As in the previous section, $(\cdot, \cdot)_D$ is the pairing between $L^p(D)$ and $L^q(D)$, ($1/p + 1/q = 1$), and $\langle f, x \rangle_{X^*, X}$ is the value of $f \in X^*$ applied to $x \in X$.

The derivation of the abstract representation is along the lines of Tröltzsch (1989), adjusted to fit into the general setting of Section 3.1. To formulate the problem (3.21) – (3.23) in the abstract setting, we define the linear operator A in $X = L^p(\Omega)$ by

$$D(A) = \left\{ u \in W^{2,p}(\Omega); \frac{\partial u}{\partial n} = 0 \text{ on } \Gamma \right\}, \quad Au = -\Delta u + au \text{ on } D(A),$$

with $0 < a \in \mathbb{R}$ such that $\operatorname{Re} \sigma(A) > 0$. Thus, the results of Chapter 2 and especially Section 2.3 can be applied. Note, that in case of Dirichlet boundary control the domain $D(A)$ would be defined accordingly.

We define the mappings $F: W^{\sigma,p}(\Omega) \times \mathbb{R} \rightarrow L^p(\Omega)$ and $B: [0, T] \times W^{\sigma-1/p,p}(\Gamma) \times \mathbb{R} \rightarrow L^p(\Gamma)$ by

$$\begin{aligned} F(u(t, \cdot), \lambda(t))(x) &= f(u(t, x), \lambda(t)) + au(t, x), \\ B(u(t, \cdot), \lambda^b(t))(x) &= b(u(t, x), \lambda^b(t)) \end{aligned}$$

Then, we have that $u \in C([0, T], W^{\sigma,p}(\Omega))$,

$$u(t) = e^{-tA}u_0 + \int_0^t e^{-(t-s)A}F(u(s), \lambda(s)) ds + \int_0^t Ae^{-(t-s)A}NB(\tau u(s), \lambda^b(s)) ds, \quad (3.24)$$

is a mild solution of (3.22). Here, $\tau: W^{\sigma,p}(\Omega) \rightarrow W^{\sigma-1,p}(\Gamma)$ is the trace operator and $N: L^p(\Gamma) \rightarrow W^{\sigma,p}(\Omega)$ assigns to $b \in L^p(\Gamma)$ the solution u of

$$\Delta u - au = 0 \text{ on } \Omega, \quad \frac{\partial u}{\partial n} = b \text{ on } \Gamma.$$

To express the objective functional in the abstract setting, we define

$$\begin{aligned} g(t, u(t), \lambda(t)) &= \int_{\Omega} \tilde{g}(t, u(t, x), \lambda(t, x)) dx \quad (u(t), \lambda(t)) \in W^{\sigma,p}(\Omega) \times \mathbb{R} \\ g^b(t, u(t), \lambda^b(t)) &= \int_{\Gamma} \tilde{g}^b(t, u(t, x), \lambda^b(t, x)) dS_x \quad (u(t), \lambda(t)) \in W^{\sigma-1/p,p}(\Gamma) \times \mathbb{R} \end{aligned}$$

and obtain the optimal control problem

$$\begin{aligned} \min J(u, \lambda) &= \int_0^T g(t, u(t), \lambda(t)) dt + \int_0^T g^b(t, u(t), \lambda^b(t)) dt + h(u(T)) \\ \text{s.t. } G(u, \lambda) &= u(t) - e^{-tA}u_0 - \int_0^t e^{-(t-s)A}F(u(s), \lambda(s)) ds \\ &\quad - \int_0^t Ae^{-(t-s)A}NB(\tau u(s), \lambda^b(s)) ds = 0. \end{aligned} \quad (3.25)$$

We derive the surjectivity of G_u with the following general linear result by Tröltzsch (1989), general in the sense that p is replaced by $1 < r < \infty$.

Lemma 3.12. *Given $F \in L^\infty(0, T; \mathcal{L}(L^r(\Omega)))$, $B \in L^\infty(0, T; \mathcal{L}(L^r(\Gamma)))$ and an abstract function $v: [0, T] \rightarrow W_r^\sigma(\Omega)$, $1/r < \sigma < 1 + 1/r$. Further we have either $v \in L^r(0, T; W_r^\sigma(\Omega))$ or $v \in C([0, T], W_r^\sigma(\Omega))$. Then the integral equation*

$$u(t) = v(t) + \int_0^t e^{-(t-s)A_r}F(s)u(s) ds + \int_0^t A_r e^{-(t-s)A_r}N_r B(s)\tau u(s) ds$$

has a unique solution in $L^r(0, T; W_r^\sigma(\Omega))$, which is continuous on $[0, T]$ if v is continuous on $[0, T]$.

Lemma 3.13 (Adjoint operators). *The adjoints of the operators that occur in the boundary control problem are as follows*

- (i) $A^* = A_q$ and $e^{-tA^*} = e^{-tA_q}$, with $1/p + 1/q = 1$
- (ii) $(\tau Ae^{-tA}N)^* = \tau A_q e^{-tA_q} N_q \in \mathcal{L}(L^q(\Gamma))$
- (iii) $(Ae^{-tA}N)^* = \tau e^{-tA_q} \in \mathcal{L}(L^q(\Omega), L^q(\Gamma))$

3. Necessary Optimality Conditions

$$(iv) (\tau e^{-tA})^* = A_q e^{-tA_q} N_q \in \mathcal{L}(L^q(\Gamma), L^q(\Omega))$$

Proof. See (Tröltzsch, 1989, Lemma 2 – Lemma 4) □

Let in the following the triple $(\bar{u}, \bar{\lambda}, \bar{\lambda}^b)$ be locally optimal. We use the shorter notation for the Fréchet-derivatives, e.g., we write $g_u(t)$ for $d_u g(t, \bar{u}(t), \bar{\lambda}(t))$, cf. Section 3.1. For the boundary control problem (3.21) – (3.23), the linear functional $l \in C([0, T], W^{\omega, p}(\Omega))^*$ from Theorem 3.9 is given by

$$\begin{aligned} l(v) = & - \int_0^T (F_u(t)^*[p(t)] + g_u(t), v(t))_{\Omega} dt - (\nabla h, v(T))_{\Omega} \\ & - \int_0^T (B_u(t)^*[\tau p(t)] + g_u^b(t), \tau v(t))_{\Gamma} dt, \end{aligned}$$

where $p \in C([0, T], W^{\omega, p}(\Omega))^*$ solves

$$\begin{aligned} p(s) = & e^{-(T-s)A_q} \nabla h + \int_s^T e^{-(t-s)A_q} (F_u(t)^*[p(t)] + g_u(t)) dt \\ & + \int_s^T A_q e^{-(t-s)A_q} N_q (B_u(t)^*[\tau p(t)] + g_u^b(t)) dt. \end{aligned} \tag{3.26}$$

Proof. We recall Equation (3.10): For $v \in Z$ we have the representation

$$l(v) = -(\nabla h, u(T))_{\Omega} - \int_0^T (g_u(t), u(t))_{\Omega} + (g_u^b(t), \tau u(t))_{\Gamma} dt, \tag{3.27}$$

where $u \in C([0, T], W^{\omega, p}(\Omega))$ solves the equation

$$u(t) = v(t) + \int_0^t e^{-(t-s)A} F_u(s)[u(s)] ds + \int_0^t A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds.$$

Analogously to the proof of Theorem 3.9, we calculate

$$\begin{aligned}
 l(v) &= - \int_0^T (F_u(t)^*[p(t)] + g_u(t), v(t))_{\Omega} dt - (\nabla h, v(T))_{\Omega} \\
 &\quad - \int_0^T (B_u(t)^*[\tau p(t)] + g_u^b(t), \tau v(t))_{\Gamma} dt \\
 &= - \int_0^T (F_u(t)^*[p(t)] + g_u(t), u(t) - \int_0^t e^{-(t-s)A} F_u(s)[u(s)] ds)_{\Omega} dt \\
 &\quad + \int_0^T (F_u(t)^*[p(t)] + g_u(t), \int_0^t A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_{\Omega} dt \\
 &\quad - (\nabla h, u(T) - \int_0^T e^{-(T-t)A} F_u(t)[u(t)] dt - \int_0^T A e^{-(T-t)A} N B_u(t)[\tau u(t)] dt)_{\Omega} \\
 &\quad - \int_0^T (B_u(t)^*[\tau p(t)] + g_u^b(t), \tau u(t) - \int_0^t \tau e^{-(t-s)A} F_u(s)[u(s)] ds)_{\Gamma} dt \\
 &\quad + \int_0^T (B_u(t)^*[\tau p(t)] + g_u^b(t), \int_0^t \tau A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_{\Gamma} dt \\
 &= - (\nabla h, u(T))_{\Omega} - \int_0^T (g_u(t), u(t))_{\Omega} + (g_u^b(t), \tau u(t))_{\Gamma} dt \\
 &\quad - \int_0^T (F_u(t)^*[p(t)], u(t))_{\Omega} dt \\
 &\quad + \int_0^T (F_u(t)^*[p(t)], \int_0^t e^{-(t-s)A} F_u(s)[u(s)] ds)_{\Omega} dt \\
 &\quad + \int_0^T (g_u(t), \int_0^t e^{-(t-s)A} F_u(s)[u(s)] ds)_{\Omega} dt \\
 &\quad + \int_0^T (B_u(t)^*[\tau p(t)], \int_0^t \tau e^{-(t-s)A} F_u(s)[u(s)] ds)_{\Gamma} dt \\
 &\quad + \int_0^T (g_u^b(t), \int_0^t \tau e^{-(t-s)A} F_u(s)[u(s)] ds)_{\Gamma} dt \\
 &\quad + (\nabla h, \int_0^T e^{-(T-t)A} F_u(t)[u(t)] dt)_{\Omega} \\
 &\quad - \int_0^T (B_u(t)^*[\tau p(t)], \tau u(t))_{\Gamma} dt \\
 &\quad + \int_0^T (F_u(t)^*[p(t)], \int_0^t A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_{\Omega} dt \\
 &\quad + \int_0^T (g_u(t), \int_0^t A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_{\Omega} dt \\
 &\quad + \int_0^T (B_u(t)^*[\tau p(t)], \int_0^t \tau A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_{\Gamma} dt \\
 &\quad + \int_0^T (g_u^b(t), \int_0^t \tau A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_{\Gamma} dt \\
 &\quad + (\nabla h, \int_0^T A e^{-(T-t)A} N B_u(t)[\tau u(t)] dt)_{\Omega}.
 \end{aligned}$$

(3.28)

3. Necessary Optimality Conditions

Now we show that Equation (3.27) and Equation (3.28) are the same. To do so, we isolate u and τu in every term of Equation (3.28) and use Equation (3.26). We only present those terms that have not been covered in the proof of Theorem 3.9. From Lemma 3.13 (iv) and Fubini's theorem follows

$$\begin{aligned} \int_0^T (B_u(t)^*[\tau p(t)], \int_0^t \tau e^{-(t-s)A} F_u(s)[u(s)] ds)_\Gamma dt \\ = \int_0^T (F_u(t)^*[\int_t^T A_q e^{-(s-t)A_q} N_q B_u(s)^*[\tau p(s)] ds], u(t))_\Omega dt \end{aligned}$$

and

$$\begin{aligned} \int_0^T (g_u^b(t), \int_0^t \tau e^{-(t-s)A} F_u(s)[u(s)] ds)_\Gamma dt \\ = \int_0^T (F_u(t)^*[\int_t^T A_q e^{-(s-t)A_q} N_q G_u^2(s) ds], u(t))_\Omega dt. \end{aligned}$$

From Lemma 3.13 (iii) follows

$$(\nabla h, \int_0^T A e^{-(T-t)A} N B_u(t)[\tau u(t)] dt)_\Omega = \int_0^T (B_u(t)^*[\tau e^{-(T-t)A_q} \nabla h], \tau u(t))_\Gamma dt$$

and together with Fubini's theorem we derive

$$\begin{aligned} \int_0^T (F_u(t)^*[p(t)], \int_0^t A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_\Omega dt \\ = \int_0^T (B_u(t)^*[\tau \int_t^T e^{-(s-t)A_q} F_u(s)^*[p(s)] ds], \tau u(t))_\Gamma dt \end{aligned}$$

and

$$\begin{aligned} \int_0^T (g_u(t), \int_0^t A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_\Omega dt \\ = \int_0^T (B_u(t)^*[\tau \int_t^T e^{-(s-t)A_q} G_u^1(s) ds], \tau u(t))_\Gamma dt. \end{aligned}$$

Finally, Lemma 3.13 (ii) yields

$$\begin{aligned} \int_0^T (B_u(t)^*[\tau p(t)], \int_0^t \tau A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_\Gamma dt \\ = \int_0^T (B_u(t)^*[\tau \int_t^T A_q e^{-(s-t)A_q} N_q B_u(s)^*[\tau p(s)] ds], \tau u(t))_\Gamma dt \end{aligned}$$

and

$$\begin{aligned} \int_0^T (g_u^b(t), \int_0^t \tau A e^{-(t-s)A} N B_u(s)[\tau u(s)] ds)_\Gamma dt \\ = \int_0^T (B_u(t)^*[\tau \int_t^T A_q e^{-(s-t)A_q} N_q G_u^2(s) ds], \tau u(t))_\Gamma dt. \end{aligned}$$

Hence,

$$\begin{aligned}
 l(v) = & -(\nabla h, u(T))_{\Omega} - \int_0^T (g_u(t), u(t))_{\Omega} + (g_u^b(t), \tau u(t))_{\Gamma} dt \\
 & + \int_0^T \left(F_u(t)^* \left[\left\{ e^{-(T-t)A_q} \nabla h + \int_t^T e^{-(s-t)A_q} (F_u(s)^* [p(s)] + g_u(s)) ds \right. \right. \right. \\
 & \quad \left. \left. \left. + \int_t^T A_q e^{-(s-t)A_q} N_q (B_u(s)^* [\tau p(s)] + g_u^b(s)) ds - p(t) \right\} \right], u(t) \right) dt \quad (3.29) \\
 & + \int_0^T \left(B_u(t)^* \left[\tau \left\{ e^{-(T-t)A_q} \nabla h + \int_t^T e^{-(s-t)A_q} (F_u(s)^* [p(s)] + g_u(s)) ds \right. \right. \right. \\
 & \quad \left. \left. \left. + \int_t^T A_q e^{-(s-t)A_q} N_q (B_u(s)^* [\tau p(s)] + g_u^b(s)) ds - p(t) \right\} \right], \tau u(t) \right)_{\Gamma} dt.
 \end{aligned}$$

Comparing the equation above with Equation (3.27) we conclude that the term in the curly brackets must be zero, which is equivalent to Equation (3.26). \square

Theorem 3.14. *For an optimal triple $(\bar{u}, \bar{\lambda}, \bar{\lambda}^b)$ we have*

$$\begin{aligned}
 & \int_0^T (F_{\lambda}(t)^* [p(t)] + g_{\lambda}(t)) (\lambda(t) - \bar{\lambda}(t)) \\
 & \quad + (B_{\lambda}(t)^* [\tau p(t)] + g_{\lambda}^b(t)) (\lambda^b(t) - \bar{\lambda}^b(t)) dt \geq 0.
 \end{aligned}$$

Proof. We need to calculate $\mathcal{I} \equiv l(-\int_0^{\cdot} e^{-(\cdot-s)A} F_{\lambda}(s) [\mu(s)] + A e^{-(\cdot-s)A} N B_{\lambda}(s) [\mu^b(s)] ds)_{\Omega}$ with $(\mu, \mu^b) = (\lambda - \bar{\lambda}, \lambda^b - \bar{\lambda}^b)$ and $(\lambda, \lambda^b) \in U^{ad} \subset \mathcal{R}([0, T], \mathbb{R}) \times \mathcal{R}([0, T], \mathbb{R})$. We apply the same techniques as in the proof above and receive

$$\begin{aligned}
 \mathcal{I} = & \int_0^T (F_u(t)^* [p(t)] + g_u(t), \int_0^t e^{-(t-s)A} F_{\lambda}(s) [\mu(s)] ds)_{\Omega} \\
 & + (F_u(t)^* [p(t)] + g_u(t), \int_0^t A e^{-(t-s)A} N B_{\lambda}(s) [\mu^b(s)] ds)_{\Omega} \\
 & + (\nabla h, e^{-(T-t)A} F_{\lambda}(t) [\mu(t)])_{\Omega} + (\nabla h, A e^{-(T-t)A} N B_{\lambda}(t) [\mu^b(t)])_{\Omega} \\
 & + (B_u(t)^* [\tau p(t)] + g_u^b(t), \int_0^t \tau e^{-(t-s)A} F_{\lambda}(s) [\mu(s)] ds)_{\Gamma} \\
 & + (B_u(t)^* [\tau p(t)] + g_u^b(t), \int_0^t \tau A e^{-(t-s)A} N B_{\lambda}(s) [\mu^b(s)] ds)_{\Gamma} dt \\
 = & \int_0^T F_{\lambda}(t)^* \left[\int_t^T e^{-(s-t)A_q} (F_u(s)^* [p(s)] + g_u(s)) ds \right] \mu(t) \\
 & + B_{\lambda}(t)^* \left[\tau \int_t^T e^{-(s-t)A_q} (F_u(s)^* [p(s)] + g_u(s)) ds \right] \mu^b(t) \\
 & + F_{\lambda}(t)^* [e^{-(T-t)A_q} \nabla h] \mu(t) + B_{\lambda}(t)^* [\tau e^{-(T-t)A_q} \nabla h] \mu^b(t) \\
 & + F_{\lambda}(t)^* \left[\int_t^T A_q e^{-(s-t)A_q} N_q (B_u(s)^* [\tau p(s)] + g_u^b(s)) ds \right] \mu(t)
 \end{aligned}$$

$$\begin{aligned}
& + B_\lambda(t)^* [\tau \int_t^T A_q e^{-(s-t)A_q} N_q(B_u(s)^* [\tau p(s)] + g_u^b(s)) ds] \mu^b(t) dt \\
= & \int_0^T F_\lambda(t)^* [\{ e^{-(T-t)A_q} \nabla h + \int_t^T e^{-(s-t)A_q} (F_u(s)^* [p(s)] + g_u(s)) ds \\
& + \int_t^T A_q e^{-(s-t)A_q} N_q(B_u(s)^* [\tau p(s)] + g_u^b(s)) ds \}] \mu(t) \\
& + B_\lambda(t)^* [\tau \{ e^{-(T-t)A_q} \nabla h + \int_t^T e^{-(s-t)A_q} (F_u(s)^* [p(s)] + g_u(s)) ds \\
& + \int_t^T A_q e^{-(s-t)A_q} N_q(B_u(s)^* [\tau p(s)] + g_u^b(s)) ds \}] \mu^b(t) dt \\
= & \int_0^T F_\lambda(t)^* [p(t)] \mu(t) + B_\lambda(t)^* [\tau p(t)] \mu^b(t) dt.
\end{aligned}$$

Adding the above to $(J_\lambda(u, \lambda), (\mu, \mu^b)) = \int_0^T g_\lambda(t) \mu(t) + g_\lambda^b(t) \mu^b(t) dt$ concludes the proof. \square

3.3 Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

Nonlocal models in the form of a partial integro-differential equation (PIDE) arise in various fields and become more and more important. In mechanics, the peridynamics theory was introduced in order to model surfaces with cracks. In finance, in particular for option pricing, existing models were extended with Lévy processes in order to model jumps, like those that occurred during the financial crisis in 2007. Recent work on numerical treatment of such jump-diffusion PIDEs or the corresponding calibration problems can be found, for example, in Andersen and Andreasen (2000), Matache et al. (2004), Cont and Voltchkova (2005), Briani et al. (2007), Sachs and Strauss (2008), or Schu (2012). Many biological models benefit from nonlocal terms. Biological applications of PIDEs are discussed, for example, in Armstrong et al. (2006), Anderson et al. (2000), Gerisch (2010), or Mogilner and Edelstein-Keshet (1999).

First, we take a closer look at a motivating example: In biology, cell adhesion describes the binding between two cells or between a cell and the extracellular matrix through certain proteins, called cell-adhesion molecules. Cell adhesion is responsible for tissue formation, tissue stability and—in case of loss of the adhesion—cell breakdown. In 1962, Steinberg showed that two different cell populations can aggregate in four different ways: mixing, engulfment, partial engulfment and complete sorting, Armstrong et al. (2006).

Armstrong et al. simulate that process with a continuous model, while they stress that all previous models were discrete ones. They model the adhesion driven cell-movement with a nonlocal term, which results in an integro-differential equation.

Without considering cell birth and cell death, mass conservation implies

$$u_t(t, x) = -J_x(t, x)$$

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

for the variation of cell concentration u in x over time. Armstrong et al. split up the flux J of the cells in

$$\text{random diffusion } J^{(d)} = -D(u_x) \text{ and adhesive forces } J^{(a)} = \frac{\phi}{R}uF,$$

where D is the diffusion coefficient, ϕ is a viscosity related constant, R is the sensing radius of the cells and F is the force that is acting on the cells within that radius. The force acting on the cell at x , that is created by a cell at position $x + y$, is given by

$$f(x) = \alpha g(u(x + y))\omega(y), \quad (3.30)$$

where g describes the nature of the forces and their dependence on the local cell density at $x + y$. The authors provide two possible examples for g : a simple linear one ($g(u) = u$) and one of logistic form ($g(u) = u(1 - u/M)$ for $u < M$ and 0 otherwise). The function $\omega(y)$ describes how the direction and magnitude of the force alters according to y (thus, ω is an odd function), a simple example would be $\omega(y) = \text{sign}(y)$. In that case, ω only provides the direction, not the magnitude of the force. α is a positive parameter reflecting the strength of adhesive force between the cells. The total force F is derived as the sum of the local forces

$$F(x) = \int_{-R}^R \alpha g(u(x + y))\omega(y) dy.$$

Together with random diffusion, we obtain the model of Armstrong et al. in one dimension and for one population:

$$u_t = Du_{xx} - (uK(u))_x = Du_{xx} - u_xK(u) - uK_x(u) \quad (3.31)$$

with

$$K(u)(x) = \frac{\phi}{R} \int_{-R}^R \alpha g(u(x + y))\omega(y) dy.$$

With two transformations, $\tau := \frac{D}{R^2}t$ and $\xi := \frac{x}{R}$, a nondimensional version can be formulated (see Armstrong et al., 2006, Section 2): If u solves (3.31), then

$$v(\tau, \xi) := \frac{R\phi}{D}u\left(\frac{R^2}{D}\tau, R\xi\right)$$

is the solution of

$$v_\tau = v_{\xi\xi} - (v\kappa(v))_\xi, \quad (3.32)$$

with

$$\kappa(v)(\xi) = \alpha \int_{-1}^1 g(v(\xi + \zeta))\omega(\zeta) d\zeta.$$

The remaining nondimensional parameter α is a measure for the adhesion strength. Armstrong et al. showed that, if α is below a certain threshold, no cell aggregation will occur.

When observing two cell populations, a distinction is made between homogeneous and heterogeneous cell adhesion. The parameters S_u , S_v and C represent the self-adhesive

strength of u , the self-adhesive strength of v and the cross-adhesive strength between u and v , respectively. Armstrong et al. find suitable parameter combinations for a system of two cell populations to model all four different cell aggregations that Steinberg proposed, i.e., mixing, engulfment, partial engulfment and complete sorting (cf. Armstrong et al., 2006, Fig. 7).

Remark 3.15. In their nonlocal model for a swarm, the authors model the nonlocal term very similar in Mogilner and Edelstein-Keshet (1999). Instead of a smaller sensing radius, the whole domain Ω serves as the sensing radius and instead of an odd function an even kernel function ω is applied. However, the authors also discuss the effects of the nonlocal force in case of an odd kernel function (Equations (3.35) and (3.36) yield how the nonlocal term in the adhesion model (3.32) can be expressed as a convolution). This shows that the nonlocal term involved in both models is of universal nature.

Given the results of the previous section, our aim in this section is to calculate the necessary optimality conditions for the nonlocal adhesion model for two populations in one space dimension. The objective of the control problem is to determine the optimal adhesion parameters to model an observed cell aggregation. Instead of constant adhesion parameters, we consider them to be time-dependent.

Armstrong et al. simulate the model on an interval with periodic boundary conditions. Hence, we choose $\Omega = (a, b)$ with $a, b \in \mathbb{R}$, $-\infty < a < b < \infty$. We choose a least squares type function for the objective function, which results in the following control problem:

$$\frac{1}{2} \int_{\Omega} (u(T, x) - u_{obs}(x))^2 + (v(T, x) - v_{obs}(x))^2 dx. \quad (3.33a)$$

The functions $u(\cdot, \cdot)$ and $v(\cdot, \cdot)$ solve the initial value problem

$$\begin{aligned} u_t &= u_{xx} - (uK_u(u, v))_x, & u(0, x) &= u_0(x), & u(t, a) &= u(t, b), \\ v_t &= v_{xx} - (vK_v(u, v))_x, & v(0, x) &= v_0(x), & v(t, a) &= v(t, b), \end{aligned} \quad (3.33b)$$

with periodic boundary conditions and integral operators

$$\begin{aligned} K_u(u, v)(t, x) &= \int_{-1}^1 S_u(t)u(t, x + y)\omega(y) + C(t)v(t, x + y)\omega(y) dy, \\ K_v(u, v)(t, x) &= \int_{-1}^1 S_v(t)v(t, x + y)\omega(y) + C(t)u(t, x + y)\omega(y) dy, \end{aligned}$$

where $\omega \in L^1([-1, 1])$ is an odd function and $u_0, v_0 \in H^1(\Omega)$ are initial values.

The functions u_{obs} and v_{obs} are cell aggregations that have been observed at time T . The objective is to choose the parameter functions S_u, S_v and C in such a way that the solutions u and v of the integro-differential equation system (3.33b) are closest to the observed cell aggregations at time T .

3.3.1 Abstract Setting

In order to be able to compute the optimal parameters, we need to formulate the necessary optimality condition. We set

$$X := (L^2(\Omega))^2, \quad L := (L^1(0, T))^3$$

and choose $\Lambda \subset (L^1(0, T))^3$ as a closed and convex set of essentially bounded functions $\lambda \in (L^1(0, T))^3$. Further, we set $\alpha = \frac{1}{2}$.

Let the operator $A: D(A) \subset (L^2(\Omega))^2 \rightarrow (L^2(\Omega))^2$ be defined as the self-adjoint extension of $1 - D_{xx}$ with domain of definition $D(A) = (H^2(\Omega))^2 \subset (L^2(\Omega))^2$. Then, $-A$ is the generator of an analytic semigroup. Since $\operatorname{Re} \sigma(-D_{xx}) \geq 0$, fractional powers of $-D_{xx}$ are not defined and thus, we consider $1 - D_{xx}$ instead of $-D_{xx}$. We are now able to consider $X_{1/2} = D(A^{1/2}) = (H^1(\Omega))^2$, see Example 2.30. Therefore, we set

$$Z = C([0, T], (H^1(\Omega))^2).$$

Let $\nu := (u, v) \in Z$ be a vector of the functions u and v and let $\lambda := (S_u, S_v, C) \in L$ be a vector of time-dependent adhesion parameters. The operator $\mathcal{K}: (L^2(\Omega))^2 \times \mathbb{R}^3 \rightarrow (L^2(\Omega))^{2 \times 2}$ is defined as

$$\mathcal{K}(\nu, \lambda) = \begin{pmatrix} K(u, v, S_u, C) & 0 \\ 0 & K(u, v, C, S_v) \end{pmatrix},$$

with $K: (L^2(\Omega))^2 \times \mathbb{R}^2 \rightarrow L^2(\Omega)$,

$$K(u, v, S, C)(x) = \int_{-1}^1 Su(x+y)\omega(y) + Cv(x+y)\omega(y) dy.$$

Finally, we define the semilinear mapping $F: (H^1(\Omega))^2 \times \mathbb{R}^3 \rightarrow (L^2(\Omega))^2$ as set in (3.2) as

$$F(\nu(t), \lambda(t)) := \nu(t) - D_x(\mathcal{K}(\nu(t), \lambda(t))\nu(t)). \quad (3.34)$$

Remark 3.16. Operators of the form $\tilde{K}(u)(x) = \int_{-1}^1 u(x+y)\omega(y) dy$ are not well defined for $x \in [a, a+1] \cup [b-1, b]$. Due to the boundary condition we can extend u periodically to \mathbb{R} such that $u(t, x) = u(t, b-a+x) = u(t, a-b+x)$ for $x \in \mathbb{R}$. For $x \in [a, a+1]$ we then have

$$\tilde{K}(u)(x) = \int_{-1}^{a-x} u(b-a+x+y)\omega(y) dy + \int_{a-x}^1 u(x+y)\omega(y) dy$$

and accordingly for $x \in [b-1, b]$

$$\tilde{K}(u)(x) = \int_{-1}^{b-x} u(x+y)\omega(y) dy + \int_{b-x}^1 u(a-b+x+y)\omega(y) dy.$$

This representation has no effect on the derivative $\tilde{K}(u)_x$, since from Leibniz' integral rule follows for $x \in [a+1, b-1]$ that

$$\tilde{K}(u)_x(x) = \int_{-1}^1 u_x(x+y)\omega(y) dy$$

3. Necessary Optimality Conditions

and for $x \in [a, a + 1]$ we have

$$\begin{aligned}
\tilde{K}(u)_x(x) &= u(b - a + x + a - x)\omega(a - x)(-1) + \int_{-1}^{a-x} u_x(b - a + x + y)\omega(y) \, dy \\
&\quad - u(x + a - x)\omega(a - x)(-1) + \int_{a-x}^1 u_x(x + y)\omega(y) \, dy \\
&= \int_{-1}^{a-x} u_x(b - a + x + y)\omega(y) \, dy + \int_{a-x}^1 u_x(x + y)\omega(y) \, dy \\
&\quad - u(b)\omega(a - x) + u(a)\omega(a - x) \\
&= \int_{-1}^{a-x} u_x(b - a + x + y)\omega(y) \, dy + \int_{a-x}^1 u_x(x + y)\omega(y) \, dy.
\end{aligned}$$

Analogously, we obtain the corresponding result for $x \in [b - 1, b]$. These considerations justify to write

$$\tilde{K}(u)(x) = \int_{-1}^1 u(x + y)\omega(y) \, dy$$

for all $x \in [a, b]$.

Lemma 3.17. *The mappings*

$$K : (L^2(\Omega))^2 \times \mathbb{R}^2 \rightarrow L^2(\Omega) \text{ and } K : (H^1(\Omega))^2 \times \mathbb{R}^2 \rightarrow H^1(\Omega)$$

are well defined and linear.

Proof. In this proof we omit the argument t in the functions u, v, S_u, S_v, C for the sake of readability. We extend ω from $[-1, 1]$ to $[a, b]$ with 0, hence

$$\int_{-1}^1 u(x + y)\omega(y) \, dy = \int_{\Omega} u(x + y)\omega(y) \, dy.$$

Since ω is an odd function, we obtain

$$\int_{\Omega} u(x + y)\omega(y) \, dy = \int_{\Omega} u(x - y)\omega(-y) \, dy = - \int_{\Omega} u(x - y)\omega(y) \, dy = -u \otimes \omega(x). \quad (3.35)$$

Thus, we can represent $K(u, v, S_u, S_v, C)$ as a sum of two convolutions

$$K(u, v, S_u, S_v, C) = -S_u(u \otimes \omega) - C(v \otimes \omega). \quad (3.36)$$

Where \otimes is the convolution on a torus, which is well-defined due to the periodic boundary condition on u , whereby u_x is periodic as well, cf. Remark 3.19 (ii). With Young's convolution inequality (see Meise and Vogt, 1997, p. 117f), $u \in L^2(\Omega)$ and $\omega \in L^1(\Omega)$, we have

$$\|u \otimes \omega\|_{L^2(\mathbb{R})} \leq \|u\|_{L^2(\mathbb{R})} \|\omega\|_{L^1(\mathbb{R})}. \quad (3.37)$$

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

For $u, v \in H^1(\Omega)$ follows

$$\begin{aligned} D_x K(u, v, S_u, S_v, C)(x) &= -S_u D_x(\omega \otimes u)(x) - C D_x(\omega \otimes v)(x) \\ &= -S_u(\omega \otimes D_x u)(x) - C(\omega \otimes D_x v)(x) \quad \text{for all } x \in \Omega \end{aligned}$$

and thus

$$\|D_x K(u, v, S_u, S_v, C)\|_{L^2(\Omega)} \leq |S_u| \|\omega\|_{L^1(\Omega)} \|D_x u\|_{L^2(\Omega)} + |C| \|\omega\|_{L^1(\Omega)} \|D_x v\|_{L^2(\Omega)} < \infty.$$

Therefore, $D_x K(u, v, S_u, S_v, C) \in L^2(\Omega)$ and we receive that $K: (H^1(\Omega))^2 \times \mathbb{R}^2 \rightarrow H^1(\Omega)$ is well-defined. The linearity of K follows directly from the linearity of the convolution operator \otimes . \square

Lemma 3.18. *The mappings*

$$\mathcal{K}: (L^2(\Omega))^2 \times \mathbb{R}^3 \rightarrow (L^2(\Omega))^{2 \times 2}, \quad \mathcal{K}: (H^1(\Omega))^2 \times \mathbb{R}^3 \rightarrow (H^1(\Omega))^{2 \times 2}$$

and

$$F: (H^1(\Omega))^2 \times \mathbb{R}^3 \rightarrow (L^2(\Omega))^2$$

are well defined. F is locally Lipschitz-continuous with respect to ν and continuously Fréchet-differentiable on $(H^1(\Omega))^2 \times \mathbb{R}^3$ with

$$dF(\bar{\nu}, \bar{\lambda})(\nu, \lambda) = \nu - D_x(\mathcal{K}(\bar{\nu}, \bar{\lambda})\nu) - D_x(\mathcal{K}(\nu, \bar{\lambda})\bar{\nu}) - D_x(\mathcal{K}(\bar{\nu}, \lambda)\bar{\nu}).$$

Proof. As in the previous proof, we omit the argument t in the functions u, v, S_u, S_v, C for the sake of readability. Before we start with the proof, we carry out some calculations. We know from Theorem 2.29 that $u \in H^1(\Omega)$ yields $u \in C(\bar{\Omega})$. We consider

$$\begin{aligned} &\|u - D_x((S_u(t) \int_{-1}^1 u(\cdot + y)\omega(y) dy + C(t) \int_{-1}^1 v(\cdot + y)\omega(y) dy) u(\cdot))\|_{L^2(\Omega)} \\ &\leq \|u\|_{L^2(\Omega)} + \|S_u(t) \int_{-1}^1 u_x(\cdot + y)\omega(y) dy u\|_{L^2(\Omega)} + \|C(t) \int_{-1}^1 v_x(\cdot + y)\omega(y) dy u\|_{L^2(\Omega)} \\ &\quad + \|S_u(t) \int_{-1}^1 u(\cdot + y)\omega(y) dy u_x\|_{L^2(\Omega)} + \|C(t) \int_{-1}^1 v(\cdot + y)\omega(y) dy u_x\|_{L^2(\Omega)} \\ &= \|u\|_{L^2(\Omega)} + I_{1,u} + I_{1,v} + I_{2,u} + I_{2,v}. \end{aligned} \tag{3.38}$$

We start with the first term and receive

$$\begin{aligned} I_{1,u} &\leq \|(S_u(t) \int_{-1}^1 u_x(\cdot + y)\omega(y) dy \max_{x \in \bar{\Omega}} u)\|_{L^2(\Omega)} \\ &\leq \max_{x \in \bar{\Omega}} |u| |S_u(t)| \|\int_{-1}^1 u_x(\cdot + y)\omega(y) dy\|_{L^2(\Omega)} \\ &\leq \|u\|_{C(\bar{\Omega})} |S_u(t)| \|\int_{\Omega} u_x(\cdot + y)\omega(y) dy\|_{L^2(\Omega)} \\ &\leq \|u\|_{C(\bar{\Omega})} |S_u(t)| \|u_x \otimes \omega\|_{L^2(\Omega)} \\ &\leq \|u\|_{C(\bar{\Omega})} |S_u(t)| \|\omega\|_{L^1(\Omega)} \|u_x\|_{L^2(\Omega)} \\ &\leq \|u\|_{C(\bar{\Omega})} |S_u(t)| \|\omega\|_{L^1(-1,1)} \|u_x\|_{L^2(\Omega)} \\ &\leq \|u\|_{C(\bar{\Omega})} |S_u(t)| \|\omega\|_{L^1(-1,1)} \|u\|_{H^1(\Omega)} \end{aligned} \tag{3.39}$$

3. Necessary Optimality Conditions

using Equations (3.35) to (3.37) from the proof of Lemma 3.17. Analogously, it follows that

$$I_{1,v} \leq \|u\|_{C(\bar{\Omega})} |C(t)| \|\omega\|_{L^1(-1,1)} \|v\|_{H^1(\Omega)}. \quad (3.40)$$

For the second term we obtain

$$\begin{aligned} I_{2,u} &\leq \|(S_u(t) \int_{-1}^1 \max_{z \in [-1,1]} u(\cdot + z) \omega(y) dy u_x)\|_{L^2(\Omega)} \\ &\leq \|(S_u(t) \int_{-1}^1 \omega(y) dy \max_{z \in [-1,1], \tilde{x} \in \bar{\Omega}} u(\tilde{x} + z) u_x)\|_{L^2(\Omega)} \\ &\leq |S_u(t)| \|\omega\|_{L^1(-1,1)} \max_{x \in \bar{\Omega}} |u(x)| \|u_x\|_{L^2(\Omega)} \\ &\leq |S_u(t)| \|\omega\|_{L^1(-1,1)} \|u\|_{C(\bar{\Omega})} \|u_x\|_{L^2(\Omega)} \\ &\leq |S_u(t)| \|\omega\|_{L^1(-1,1)} \|u\|_{C(\bar{\Omega})} \|u\|_{H^1(\Omega)}. \end{aligned} \quad (3.41)$$

And again it follows analogously that

$$I_{2,v} \leq |C(t)| \|\omega\|_{L^1(-1,1)} \|v\|_{C(\bar{\Omega})} \|u\|_{H^1(\Omega)}. \quad (3.42)$$

Hence,

$$\begin{aligned} \left\| \left(\boldsymbol{\nu} - D_x(\mathcal{K}(\boldsymbol{\nu}, \boldsymbol{\lambda}(t))\boldsymbol{\nu}) \right)_1 \right\|_{L^2(\Omega)} &\leq \|u\|_{L^2(\Omega)} + \|u\|_{C(\bar{\Omega})} |S_u(t)| \|\omega\|_{L^1(-1,1)} \|u\|_{H^1(\Omega)} \\ &\quad + \|u\|_{C(\bar{\Omega})} |C(t)| \|\omega\|_{L^1(-1,1)} \|v\|_{H^1(\Omega)} \\ &\quad + |S_u(t)| \|\omega\|_{L^1(-1,1)} \|u\|_{C(\bar{\Omega})} \|u\|_{H^1(\Omega)} \\ &\quad + |C(t)| \|\omega\|_{L^1(-1,1)} \|v\|_{C(\bar{\Omega})} \|u\|_{H^1(\Omega)} \\ &\leq C_1 \|u\|_{H^1(\Omega)} + C_{1,u} |S_u(t)| \|u\|_{H^1(\Omega)} \\ &\quad + C_{1,v} |C(t)| \|v\|_{H^1(\Omega)} + C_{1,uv} |C(t)| \|u\|_{H^1(\Omega)} \\ &\leq (C_1 + C_{1,u} |S_u(t)| + C_{1,uv} |C(t)|) \|u\|_{H^1(\Omega)} \\ &\quad + C_{1,v} |C(t)| \|v\|_{H^1(\Omega)} \\ &\leq C_{1,\nu} \kappa_1(t) (\|u\|_{H^1(\Omega)} + \|v\|_{H^1(\Omega)}), \end{aligned} \quad (3.43)$$

where $\kappa_1 \in C(\mathbb{R})$ with $C_{1,\nu} \kappa_1(t) \geq C_1 + C_{1,u} |S_u(t)| + (C_{1,uv} + C_{1,v}) |C(t)|$ for $t \in [0, T]$ and $C_{1,\nu}$ depends on $\boldsymbol{\nu}$. With the same considerations and calculations applied to the second component of F follows

$$\begin{aligned} \left\| \left(\boldsymbol{\nu}(t) - D_x(\mathcal{K}(\boldsymbol{\nu}(t), \boldsymbol{\lambda}(t))\boldsymbol{\nu}) \right)_2 \right\|_{L^2(\Omega)} &\leq C_2 \|v\|_{H^1(\Omega)} + C_{2,u} |C(t)| \|u\|_{H^1(\Omega)} \\ &\quad + C_{2,v} |S_v(t)| \|v\|_{H^1(\Omega)} + C_{2,uv} |C(t)| \|v\|_{H^1(\Omega)} \\ &\leq (C_2 + C_{2,v} |S_v(t)| + C_{1,uv} |C(t)|) \|v\|_{H^1(\Omega)} \\ &\quad + C_{2,u} |C(t)| \|u\|_{H^1(\Omega)} \\ &\leq C_{2,\nu} \kappa_2(t) (\|u\|_{H^1(\Omega)} + \|v\|_{H^1(\Omega)}), \end{aligned}$$

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

where $\kappa_2 \in C(\mathbb{R})$ with $C_{2,\nu}\kappa_2(t) \geq C_2 + C_{2,v}|S_v(t)| + (C_{2,uv} + C_{2,u})|C(t)|$ for $t \in [0, T]$ and $C_{2,\nu}$ depends on ν . Before we combine the previous results, we recall the elementary inequality $2ab \leq a^2 + b^2$ for real $a, b > 0$ and obtain

$$\begin{aligned} (\|u\|_{H^1(\Omega)} + \|v\|_{H^1(\Omega)})^2 &= \|u\|_{H^1(\Omega)}^2 + \|v\|_{H^1(\Omega)}^2 + 2\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \\ &\leq 2(\|u\|_{H^1(\Omega)}^2 + \|v\|_{H^1(\Omega)}^2). \end{aligned} \quad (3.44)$$

Finally, we choose $\kappa \in C(\mathbb{R})$ with $C_\nu\kappa(t) \geq \sqrt{2(C_{1,\nu}^2\kappa_1(t)^2 + C_{2,\nu}^2\kappa_2(t)^2)}$ and obtain

$$\begin{aligned} \|\nu - D_x(\mathcal{K}(\nu, \lambda(t))\nu)\|_{L^2(\Omega)}^2 &\leq (C_{1,\nu}^2\kappa_1(t)^2 + C_{2,\nu}^2\kappa_2(t)^2)(\|u\|_{H^1(\Omega)} + \|v\|_{H^1(\Omega)})^2 \\ &\leq C_\nu^2\kappa(t)^2(\|u\|_{H^1(\Omega)}^2 + \|v\|_{H^1(\Omega)}^2) \\ &= C_\nu^2\kappa(t)^2\|\nu\|_{(H^1(\Omega))^2}^2. \end{aligned} \quad (3.45)$$

Well-defined: With Lemma 3.17 we receive

$$\begin{aligned} \|\mathcal{K}(\nu, \lambda)\|_{(L^2(\Omega))^{2 \times 2}}^2 &= \|(K(u, v, S_u, S_v, C), 0)^\top\|_{(L^2(\Omega))^2}^2 + \|(0, K(v, u, S_v, S_u, C))^\top\|_{(L^2(\Omega))^2}^2 \\ &= \|K(u, v, S_u, S_v, C)\|_{L^2(\Omega)}^2 + \|K(v, u, S_v, S_u, C)\|_{L^2(\Omega)}^2 < \infty, \end{aligned}$$

and analogously

$$\|D_x\mathcal{K}(\nu, \lambda)\|_{(L^2(\Omega))^{2 \times 2}}^2 = \|\mathcal{K}(D_x\nu, \lambda)\|_{(L^2(\Omega))^{2 \times 2}}^2 < \infty.$$

Theorem 2.29 yields $u \in C(\bar{\Omega})$ for $u \in H^1(\Omega)$ and we obtain with the calculation above

$$\begin{aligned} \|F(\nu, \lambda)\|_{(L^2(\Omega))^2} &= \|\nu - \mathcal{K}(\nu, \lambda)D_x\nu - \mathcal{K}(D_x\nu, \lambda)\nu\|_{(L^2(\Omega))^2} \\ &\leq \|\nu\|_{(L^2(\Omega))^2} + \|\mathcal{K}(\nu, \lambda)D_x\nu\|_{(L^2(\Omega))^2} + \|\mathcal{K}(D_x\nu, \lambda)\nu\|_{(L^2(\Omega))^2} \\ &\leq C_\nu\kappa(t)\|\nu\|_{(H^1(\Omega))^2} < \infty. \end{aligned}$$

Lipschitz-continuity with respect to ν : Let $r > 0$ and let any $\nu_1, \nu_2 \in U_r(\nu_0) \subset (H^1(\Omega))^2$ be given. We have to show that there exists a constant $c > 0$ with

$$\|F(\nu_1, \lambda) - F(\nu_2, \lambda)\|_{(L^2(\Omega))^2} \leq c\|\nu_1 - \nu_2\|_{(H^1(\Omega))^2}.$$

We consider the first component of F , the same result for the second component is obtained analogously. Let $\nu_1 = (u_1, v_1), \nu_2 = (u_2, v_2) \in U_r(\nu_0)$, $\tilde{u} = u_1 - u_2, \tilde{v} = v_1 - v_2$ and $\lambda \in \Lambda$. Since $\nu_1, \nu_2 \in U_r(\nu_0)$, we have

$$\|D_x\tilde{u}\|_{L^2(\Omega)} \leq \|\tilde{u}\|_{H^1(\Omega)} < 2r, \quad \|D_x\tilde{v}\|_{L^2(\Omega)} \leq \|\tilde{v}\|_{H^1(\Omega)} < 2r. \quad (3.46)$$

We begin the proof with the following basic estimate

$$\begin{aligned} \|(F(\nu_1, \lambda))_1 - (F(\nu_2, \lambda))_1\|_{L^2(\Omega)} &\leq \|\tilde{u}\|_{L^2(\Omega)} + \|K(\tilde{u}, \tilde{v}, S_u, C)D_x\tilde{u}\|_{L^2(\Omega)} \\ &\quad + \|K(D_x\tilde{u}, D_x\tilde{v}, S_v, C)\tilde{u}\|_{L^2(\Omega)} \\ &\leq \|\tilde{u}\|_{L^2(\Omega)} + I_{2,\tilde{u}} + I_{2,\tilde{v}} + I_{1,\tilde{u}} + I_{1,\tilde{v}} \\ &\leq C_{1,\tilde{\nu}}\kappa_1(t)(\|\tilde{u}\|_{H^1(\Omega)} + \|\tilde{v}\|_{H^1(\Omega)}). \end{aligned}$$

3. Necessary Optimality Conditions

The terms $I_{i,j}$ ($i = 1, 2, j = \tilde{u}, \tilde{v}$) are defined accordingly to the terms $I_{i,j}$ in Equation (3.38). With Equation (3.44) follows

$$\|(F(\boldsymbol{\nu}_1, \boldsymbol{\lambda}))_1 - (F(\boldsymbol{\nu}_2, \boldsymbol{\lambda}))_1\|_{L^2(\Omega)} \leq C_{1,\tilde{\nu}\kappa_1}(t) \|\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2\|_{H^1(\Omega)^2}.$$

Analogously follows for the second component

$$\|(F(\boldsymbol{\nu}_1, \boldsymbol{\lambda}))_2 - (F(\boldsymbol{\nu}_2, \boldsymbol{\lambda}))_2\|_{L^2(\Omega)} \leq C_{2,\tilde{\nu}\kappa_1}(t) \|\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2\|_{H^1(\Omega)^2}.$$

Finally, we obtain

$$\|F(\boldsymbol{\nu}_1, \boldsymbol{\lambda}) - F(\boldsymbol{\nu}_2, \boldsymbol{\lambda})\|_{L^2_2(\Omega)} \leq C_{\tilde{\nu}\kappa_1}(t) \|\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2\|_{H^1(\Omega)^2}.$$

Fréchet-differentiability: For $u \in H^1$, $v \in L^2$ and $\omega \in L^1$ follows

$$\begin{aligned} \|(\omega * u)v\|_{L^2(\Omega)} &= \left(\int_{\Omega} \left(\int_{\mathbb{R}} u(x-y)\omega(y) dy v(x) \right)^2 dx \right)^{\frac{1}{2}} \\ &= \left(\int_{\Omega} \left(\int_{\mathbb{R}} u(x-y)\omega(y) dy \right)^2 (v(x))^2 dx \right)^{\frac{1}{2}} \\ &\leq \left(\sup_{x \in \Omega} \int_{\mathbb{R}} |u(x-y)\omega(y)| dy \right) \left(\int_{\Omega} v(x)^2 dx \right)^{\frac{1}{2}} \\ &\leq \sup_{x \in \Omega} |u(x)| \int_{\mathbb{R}} |\omega(y)| dy \|v\|_{L^2(\Omega)} \\ &\leq \|u\|_{C(\Omega)} \|\omega\|_{L^1(\Omega)} \|v\|_{L^2(\Omega)}. \end{aligned} \quad (3.47)$$

We set $\boldsymbol{z} := (\boldsymbol{\nu}, \boldsymbol{\lambda}) \in (H^1(\Omega))^2 \times \mathbb{R}^3$ with $\|\boldsymbol{z}\| := (\|\boldsymbol{\nu}\|_{(H^1(\Omega))^2}^2 + |\boldsymbol{\lambda}|^2)^{\frac{1}{2}}$. For $S_u \in \mathbb{R}$ and $u \in H^1(\Omega)$ we have then

$$|S_u| \leq |\boldsymbol{\lambda}| \leq \|\boldsymbol{z}\|, \quad \|u\|_{H^1(\Omega)} \leq \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2} \leq \|\boldsymbol{z}\| \quad (3.48a)$$

and

$$|\boldsymbol{\lambda}| \|u\|_{H^1(\Omega)} \leq |\boldsymbol{\lambda}| \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2} \leq |\boldsymbol{\lambda}|^2 + \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2}^2 \leq \|\boldsymbol{z}\|^2, \quad (3.48b)$$

since for $a, b \in \mathbb{R}$ either $|a||b| \leq |a|^2$ or $|a||b| \leq |b|^2$ holds. We need to show, that with

$$\begin{aligned} dF(\boldsymbol{\nu}, \boldsymbol{\lambda})[\boldsymbol{\Delta}\boldsymbol{\nu}, \boldsymbol{\Delta}\boldsymbol{\lambda}] &= \boldsymbol{\Delta}\boldsymbol{\nu} - \mathcal{K}(\boldsymbol{\nu}, \boldsymbol{\lambda})D_x\boldsymbol{\Delta}\boldsymbol{\nu} - \mathcal{K}(\boldsymbol{\Delta}\boldsymbol{\nu}, \boldsymbol{\lambda})D_x\boldsymbol{\nu} - \mathcal{K}(D_x\boldsymbol{\nu}, \boldsymbol{\lambda})\boldsymbol{\Delta}\boldsymbol{\nu} \\ &\quad - \mathcal{K}(D_x\boldsymbol{\Delta}\boldsymbol{\nu}, \boldsymbol{\lambda})\boldsymbol{\nu} - \mathcal{K}(\boldsymbol{\nu}, \boldsymbol{\Delta}\boldsymbol{\lambda})D_x\boldsymbol{\nu} - \mathcal{K}(D_x\boldsymbol{\nu}, \boldsymbol{\Delta}\boldsymbol{\lambda})\boldsymbol{\nu} \end{aligned} \quad (3.49)$$

the estimate

$$\|F(\boldsymbol{\nu} + \boldsymbol{\Delta}\boldsymbol{\nu}, \boldsymbol{\lambda} + \boldsymbol{\Delta}\boldsymbol{\lambda}) - F(\boldsymbol{\nu}, \boldsymbol{\lambda}) - dF(\boldsymbol{\nu}, \boldsymbol{\lambda})[\boldsymbol{\Delta}\boldsymbol{\nu}, \boldsymbol{\Delta}\boldsymbol{\lambda}]\|_{(L^2(\Omega))^2} \leq o(\|\boldsymbol{\Delta}\boldsymbol{z}\|)$$

holds. We consider the left side of the inequality, which we denote by $r(\boldsymbol{\Delta}\boldsymbol{z})$, component-wise and obtain

$$\begin{aligned} \|(r(\boldsymbol{\Delta}\boldsymbol{z}))_1\|_{L^2(\Omega)} &= \|K(\Delta u, \Delta v, \Delta S_u, \Delta C)D_x u + K(\Delta u, \Delta v, \Delta S_u, \Delta C)D_x \Delta u \\ &\quad + K(\Delta u, \Delta v, S_u, C)D_x \Delta u + K(u, v, \Delta S_u, \Delta C)D_x \Delta u \\ &\quad + K(D_x \Delta u, D_x \Delta v, \Delta S_u, \Delta C)u + K(D_x \Delta u, D_x \Delta v, \Delta S_u, \Delta C)\Delta u \\ &\quad + K(D_x \Delta u, D_x \Delta v, S_u, C)\Delta u + K(D_x u, D_x v, \Delta S_u, \Delta C)\Delta u\|_{(L^2(\Omega))^2}. \end{aligned}$$

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

We estimate each term separately; to do that, we use Equations (3.36), (3.37), (3.39) to (3.43), (3.47) and (3.48). For the sake of readability, we omit the domain Ω in the norm. With constants c , that depend only on ν , λ and ω respectively, follows

$$\begin{aligned} \|K(\Delta u, \Delta v, \Delta S_u, \Delta C)D_x u\|_{L^2} &\leq |\Delta S_u| \|(\omega * \Delta u)D_x u\|_{L^2} + |\Delta C| \|(\omega * \Delta v)D_x u\|_{L^2} \\ &\leq |\Delta \lambda| \|D_x u\|_{L^2} \|\omega\|_{L^1} (\|\Delta u\|_C + \|\Delta v\|_C) \\ &\leq 2c \|\Delta z\| \|\Delta \nu\|_{(C)^2} \leq o(\|\Delta z\|), \end{aligned}$$

$$\begin{aligned} \|K(u, v, \Delta S_u, \Delta C)D_x \Delta u\|_{L^2} &\leq |\Delta \lambda| \|D_x \Delta u\|_{L^2} \|\omega\|_{L^1} (\|u\|_C + \|v\|_C) \\ &\leq c |\Delta \lambda| \|\Delta \nu\|_{H_2^1} \leq c \|\Delta z\|^2 \leq o(\|\Delta z\|), \end{aligned}$$

$$\begin{aligned} \|K(\Delta u, \Delta v, S_u, C)D_x \Delta u\|_{L^2} &\leq |\lambda| \|D_x \Delta u\|_{L^2} \|\omega\|_{L^1} (\|\Delta u\|_C + \|\Delta v\|_C) \\ &\leq c \|\Delta u\|_{H_2^1} (\|\Delta u\|_C + \|\Delta v\|_C) \\ &\leq 2c \|\Delta z\| \|\Delta \nu\|_{(C)^2} \leq o(\|\Delta z\|), \end{aligned}$$

$$\begin{aligned} \|K(\Delta u, \Delta v, \Delta S_u, \Delta C)D_x \Delta u\|_{L^2} &\leq |\Delta \lambda| \|D_x \Delta u\|_{L^2} \|\omega\|_{L^1} (\|\Delta u\|_C + \|\Delta v\|_C) \\ &\leq 2c \|\Delta z\|^2 \|\Delta \nu\|_{(C)^2} \leq o(\|\Delta z\|), \end{aligned}$$

$$\begin{aligned} \|K(D_x \Delta u, D_x \Delta v, \Delta S_u, \Delta C)u\|_{L^2} &\leq \|u\|_C |\Delta \lambda| \|\omega\|_{L^1} (\|D_x \Delta u\|_{L^2} + \|D_x \Delta v\|_{L^2}) \\ &\leq c |\Delta \lambda| (\|\Delta \nu\|_{H_2^1} + \|\Delta \nu\|_{H_2^1}) \leq 2c |\Delta \lambda| \|\Delta \nu\|_{H_2^1} \\ &\leq 2c \|\Delta z\|^2 \leq o(\|\Delta z\|), \end{aligned}$$

$$\begin{aligned} \|K(D_x u, D_x v, \Delta S_u, \Delta C)\Delta u\|_{L^2} &\leq \|\Delta u\|_C |\Delta \lambda| \|\omega\|_{L^1} (\|D_x u\|_{L^2} + \|D_x v\|_{L^2}) \\ &\leq c |\Delta \lambda| \|\Delta \nu\|_{C^0} \leq \|\Delta z\| \|\Delta \nu\|_{(C)^2} \leq o(\|\Delta z\|), \end{aligned}$$

$$\begin{aligned} \|K(D_x \Delta u, D_x \Delta v, S_u, C)\Delta u\|_{L^2} &\leq \|\Delta u\|_C |\lambda| \|\omega\|_{L^1} (\|D_x \Delta u\|_{L^2} + \|D_x \Delta v\|_{L^2}) \\ &\leq 2c \|\Delta \nu\|_{C^0} \|\Delta \nu\|_{H_2^1} \leq 2c \|\Delta z\| \|\Delta \nu\|_{(C)^2} \leq o(\|\Delta z\|) \end{aligned}$$

and

$$\begin{aligned} \|K(D_x \Delta u, D_x \Delta v, \Delta S_u, \Delta C)\Delta u\|_{L^2} &\leq \|\Delta u\|_C |\Delta \lambda| \|\omega\|_{L^1} (\|D_x \Delta u\|_{L^2} + \|D_x \Delta v\|_{L^2}) \\ &\leq 2c \|\Delta z\|^2 \|\Delta \nu\|_{(C)^2} \leq o(\|\Delta z\|). \end{aligned}$$

We show $\|(r(\Delta z))_2\|_{L^2(\Omega)} \leq o(\|\Delta z\|)$ with the same approach and conclude the proof. \square

Remark 3.19. (i) The boundedness of the control functions comes here into play, because it ensures $\int_0^{t_0} |F(\boldsymbol{\nu}(s), \boldsymbol{\lambda}(s))| ds < \infty$, and with the results of Section 2.4.2 we derive the local existence of a unique solution of (3.33b).

(ii) The periodic boundary conditions of (3.33) have also to be included in the definition of the spaces. Consider for $k \in \mathbb{N}$ $(H_{\Omega}^k)^2$ the space of $|\Omega|$ -periodic functions whose restrictions to Ω are in $(H^k(\Omega))^2$. Without loss of generality, let $\Omega = [-\pi, \pi]$. For $s \geq 0$ the Sobolev-space of 2π -periodic functions is defined as

$$(\hat{H}_{2\pi}^s)^2 := \{\boldsymbol{\nu} \in (L_{2\pi}^2)^2 : \|\boldsymbol{\nu}\|_{(\hat{H}_{2\pi}^s)^2} := \sum_{k=-\infty}^{\infty} (1 + |k|^2)^s |\hat{\boldsymbol{\nu}}(k)|^2 < \infty\},$$

where $\hat{\boldsymbol{\nu}}$ is the Fourier transform of $\boldsymbol{\nu}$. Since the Fourier transform is an isometry, $(\hat{H}_{2\pi}^s)^2$ is a Hilbert-space for $s \geq 0$ and $(\hat{H}_{2\pi}^k)^2 = (H_{2\pi}^k)^2$ for $k \in \mathbb{N}$ (cf. Kaballo, 2011, p. 114 and Satz 6.7). We obtain that $(H^1(\Omega))^2$ is together with the boundary conditions equal to $(\hat{H}_{2\pi}^1)^2$ and $(H_{\Omega}^1)^2$.

We can now formulate (3.33) in the framework of (3.3):

$$\begin{aligned} \min_{\boldsymbol{\lambda} \in \Lambda} J(\boldsymbol{\nu}, \boldsymbol{\lambda}) &= \frac{1}{2} \|\boldsymbol{\nu}(T; \boldsymbol{\lambda}) - \boldsymbol{\nu}_{obs}\|_{(L^2(\Omega))^2}^2 \\ \text{s.t. } G(\boldsymbol{\nu}, \boldsymbol{\lambda}) &= \boldsymbol{\nu} - e^{-\cdot A} \boldsymbol{\nu}_0 - \int_0^{\cdot} e^{-(\cdot-s)A} F(\boldsymbol{\nu}(s), \boldsymbol{\lambda}(s)) ds = 0 \\ (\boldsymbol{\nu}, \boldsymbol{\lambda}) &\in C([0, T], (H^1(\Omega))^2) \times \Lambda. \end{aligned} \quad (3.50)$$

It is easy to see that Assumption 3.1 (i) is satisfied for a tracking type objective function with

$$dJ(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})[\boldsymbol{\nu}, \boldsymbol{\lambda}] = \int_{\Omega} (\bar{\boldsymbol{\nu}}(T, x; \boldsymbol{\lambda}) - \boldsymbol{\nu}_{obs}(x))^{\top} \boldsymbol{\nu}(T, x) dx. \quad (3.51)$$

Assumption 3.1 (ii) and (iii) are met by Lemma 3.18 and we obtain the partial derivatives

$$G_{\boldsymbol{\nu}}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})[\boldsymbol{\nu}] = \boldsymbol{\nu} + \int_0^{\cdot} e^{-(\cdot-s)A} [D_x(\mathcal{K}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})\boldsymbol{\nu}) + D_x(\mathcal{K}(\boldsymbol{\nu}, \bar{\boldsymbol{\lambda}})\bar{\boldsymbol{\nu}})] ds$$

and

$$G_{\boldsymbol{\lambda}}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})[\boldsymbol{\lambda}] = \int_0^{\cdot} e^{-(\cdot-s)A} D_x(\mathcal{K}(\bar{\boldsymbol{\nu}}, \boldsymbol{\lambda})\bar{\boldsymbol{\nu}}) ds.$$

3.3.2 Global Existence

Assumption 3.1 (iii) and Theorem 2.51 guarantee only local existence of a solution of Equation (3.33b) on an interval $(0, T')$ with possibly $T' < T$. In order to receive global existence (and thus existence on $[0, T]$), we need to verify Assumption 3.1 (iv). Hence, we need to show the existence of a continuous $\kappa: \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\|\boldsymbol{\nu}(t) - D_x(\mathcal{K}(\boldsymbol{\nu}(t), \boldsymbol{\lambda}(t))\boldsymbol{\nu}(t))\|_{(L^2(\Omega))^2} \leq \kappa(t)(1 + \|\boldsymbol{\nu}(t)\|_{(H^1(\Omega))^2})$$

for all $t > 0$ and $\boldsymbol{\nu}(t) \in (H^1(\Omega))^2$.

A first attempt would be to use the calculations we carried out at the beginning of the proof of Lemma 3.18 (Equations (3.38) to (3.45)). However, this would result in

$$\|F(\boldsymbol{\nu}(t), \boldsymbol{\lambda}(t))\|_{(L^2(\Omega))^2} \leq C_{\boldsymbol{\nu}} \kappa(t) \|\boldsymbol{\nu}(t)\|_{(H^1(\Omega))^2},$$

with $C_{\boldsymbol{\nu}}$ depending on $\boldsymbol{\nu}$ —this is due to the fact that there exist no upper bound for $\|\boldsymbol{\nu}(t)\|_{C(\Omega)^2}$ for all $\boldsymbol{\nu}$ and t . Therefore, Assumption 3.1 (iv) would not hold.

A second attempt would be to modify the force function g in Equation (3.30), e.g., we could use the logistic type function given in the model introduction or even just

$$g(u(x)) = \begin{cases} 0 & \text{for } u(x) < 0 \\ u(x) & \text{for } 0 \leq u(x) < M \\ 0 & \text{for } u(x) \geq M, \end{cases}$$

with sufficiently big $M \in \mathbb{R}$. Then, for $u(x) \in (0, M)$, the derivative of g with respect to x is given by $g_x(u(x)) = u_x(x)$. Since u_x can be unbounded for bounded u we would not obtain an upper bound of g_x for all u . Thus, the estimates of $I_{1,u}$ and $I_{1,v}$ from Equation (3.38) would still contain constants that depend on $\boldsymbol{\nu}$ —furthermore, we would need to smooth out the nondifferentiability at $u(x) = 0$ and $u(x) = M$, e.g., with suitable polynomials, so that Assumption 3.1 (ii) and (iii) remains satisfied.

Therefore, we rewrite Equation (3.34) as

$$F(\boldsymbol{\nu}(t), \boldsymbol{\lambda}(t)) = \boldsymbol{\nu}(t) - \mathcal{K}(D_x \boldsymbol{\nu}(t), \boldsymbol{\lambda}(t)) \boldsymbol{\nu}(t) - \mathcal{K}(\boldsymbol{\nu}(t), \boldsymbol{\lambda}(t)) D_x \boldsymbol{\nu}(t)$$

with

$$\mathcal{K}(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \begin{pmatrix} \tilde{K}(u, v, S_u, C) & 0 \\ 0 & \tilde{K}(u, v, C, S_v) \end{pmatrix}$$

and $\tilde{K}: (L^2(\Omega))^2 \times \mathbb{R}^2 \rightarrow L^2(\Omega)$,

$$\tilde{K}(u, v, S, C)(x) = \begin{cases} 0 & \text{if } \min_{z \in [x-1, x+1]} \{Su(z) + Cv(z)\} \leq 0, \\ 0 & \text{if } \max_{z \in [x-1, x+1]} \{Su(z) + Cv(z)\} \geq (S+C)M, \\ K(u, v, S, C)(x) & \text{otherwise,} \end{cases}$$

with $M \in \mathbb{R}$ big enough. For the sake of readability, we write u and v instead of $u(t)$ and $v(t)$ in the following. The definition of \tilde{K} implies that $\tilde{K}(u, v, S_u(t), C(t))(x)$ is nonzero if $0 < S(t)u(z) + C(t)v(z) < (S(t) + C(t))M$ for all $z \in [x-1, x+1]$. This yields for all $x \in \Omega$ the estimate

$$\begin{aligned} \tilde{K}(u, v, S(t), C(t))(x) &= \int_{-1}^1 S(t)u(x+y)\omega(y) + C(t)v(x+y)\omega(y) dy \\ &\leq \int_{-1}^1 |\omega(y)| dy (S(t) + C(t))M \\ &\leq c_\omega M(S(t) + C(t)). \end{aligned} \tag{3.52}$$

We have

$$\begin{aligned}
 \|\mathcal{K}(D_x \boldsymbol{\nu}, \boldsymbol{\lambda}(t)) \boldsymbol{\nu}\|_{(L^2(\Omega))^2}^2 &= \|\tilde{K}(u_x, v_x, S_u(t), C(t))u\|_{L^2(\Omega)}^2 + \|\tilde{K}(u_x, v_x, C(t), S_v(t))v\|_{L^2(\Omega)}^2 \\
 &\leq c_\omega^2 M^2 (|S_u(t) + C(t)|^2 \|u\|_{L^2(\Omega)}^2 + |C(t) + S_v(t)|^2 \|v\|_{L^2(\Omega)}^2) \\
 &\leq \frac{1}{3} \kappa^2(t) (\|u\|_{H^1(\Omega)}^2 + \|v\|_{H^1(\Omega)}^2) \\
 &\leq \frac{1}{3} \kappa^2(t) \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2}^2,
 \end{aligned}$$

where $\kappa(t) \geq 3 \max\{c_\Omega, 2c_\omega \max\{1, c_\Omega\} M |\boldsymbol{\lambda}(t)|\}$ and c_Ω is the Poincaré constant. With the same arguments, we obtain

$$\begin{aligned}
 \|\mathcal{K}(\boldsymbol{\nu}, \boldsymbol{\lambda}(t)) D_x \boldsymbol{\nu}\|_{(L^2(\Omega))^2}^2 &= \|\tilde{K}(u, v, S_u(t), C(t))u_x\|_{L^2(\Omega)}^2 + \|\tilde{K}(u, v, C(t), S_v(t))v_x\|_{L^2(\Omega)}^2 \\
 &\leq \frac{1}{3} \kappa^2(t) \|D_x \boldsymbol{\nu}\|_{(L^2(\Omega))^2}^2 \\
 &\leq \frac{1}{3} \kappa^2(t) \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2}^2,
 \end{aligned}$$

Hence,

$$\begin{aligned}
 \|F(\boldsymbol{\nu}, \boldsymbol{\lambda}(t))\|_{(L^2(\Omega))^2} &\leq \|\boldsymbol{\nu}\|_{(L^2(\Omega))^2} + \frac{2}{3} \kappa(t) \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2} \\
 &\leq \kappa(t) \|\boldsymbol{\nu}\|_{(H^1(\Omega))^2}.
 \end{aligned}$$

In order for Assumption 3.1 (ii) and (iii) remaining satisfied, suitable smoothing functions need to be applied to \tilde{K} around the nondifferentiability and noncontinuity points and the proofs of Lemma 3.17 and Lemma 3.18 need to be adapted accordingly. However, this is not in the scope of this thesis and we leave that for potential future research.

In Rankin (1993) assumptions are presented under which semilinear evolution equations of the form $u'(t) + Au(t) = F(u(t))$ with $F = A^\alpha G$ are solved both locally and globally. Extending that result to permit controls $\lambda(t)$ as an additional argument of F and G is a direction of future research.

3.3.3 Adjoint Operators and the Adjoint Equation

For the rest of this chapter, let $\bar{\boldsymbol{\nu}} \in C([0, T], (H^1(\Omega))^2)$ and $\bar{\boldsymbol{\lambda}} \in (L^1(0, T))^3$ be an optimal state and optimal control, respectively. In order to apply Theorems 3.9 and 3.10 to Equation (3.50) we first calculate the adjoints of the partial derivatives of $F(\bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t))$ and use the relations $((L^2(\Omega))^2)^* = (L^2(\Omega))^2$ and $((H^1(\Omega))^2)^* = (H^1(\Omega))^2$. With respect to the notation in Section 3.1, we write $F_\nu(t)$ instead of $F_\nu(\bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t))$ and $F_\lambda(t)$ instead of $F_\lambda(\bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t))$ and use the notations

$$\langle \cdot, \cdot \rangle = (\cdot, \cdot)_{(H^1(\Omega))^2} \quad \text{and} \quad (\cdot, \cdot) = (\cdot, \cdot)_{\mathbb{R}^3}.$$

Theorem 3.20. *We have the following adjoints*

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

(i) $F_\nu(t)^*: (L^2(\Omega))^2 \rightarrow (H^1(\Omega))^2$ is given by

$$F_\nu(t)^* = \text{id}_{(H^1(\Omega))^2} + \mathcal{K}(\bar{\nu}(t), \bar{\lambda}(t))D_x - \mathcal{K}(D_x(\cdot) \odot \bar{\nu}(t), \bar{\lambda}(t))\mathbf{id}.$$

(ii) $F_\lambda(t)^*: (L^2(\Omega))^2 \rightarrow \mathbb{R}^3$ is given by

$$F_\lambda(t)^* = \langle \hat{\mathcal{K}}(\bar{\nu}(t)), \cdot \rangle,$$

whereas \mathbf{id} is the identity in $(H^1(\Omega))^2$ and $a \odot b$ is the Hadamard product or elementwise product of two vector valued functions a and b , $\hat{\mathcal{K}}(\nu) \in (L^2(\Omega))^{3 \times 2}$ is defined as

$$\hat{\mathcal{K}}(\nu) = \begin{pmatrix} D_x(uK(u, \nu, 1, 0)) & 0 \\ 0 & D_x(vK(\nu, u, 1, 0)) \\ D_x(uK(u, \nu, 0, 1)) & D_x(vK(\nu, u, 0, 1)) \end{pmatrix}$$

and

$$\langle \hat{\mathcal{K}}(\nu), \pi \rangle = \begin{pmatrix} \langle D_x(uK(u, \nu, 1, 0)), p \rangle \\ \langle D_x(vK(\nu, u, 1, 0)), q \rangle \\ \langle D_x(uK(u, \nu, 0, 1)), p \rangle + \langle D_x(vK(\nu, u, 0, 1)), q \rangle \end{pmatrix}.$$

Proof. In this proof, we omit the dependence on t for the sake of readability. Let some $\pi = (p, q)^\top \in (H^1(\Omega))^2$ be given.

(i) Consider

$$\langle \pi, F_\nu(\bar{\nu}, \bar{\lambda})[\nu] \rangle = \langle \pi, \nu - D_x(\mathcal{K}(\bar{\nu}, \bar{\lambda})\nu + \mathcal{K}(\nu, \bar{\lambda})\bar{\nu}) \rangle.$$

It is well known, that D_x is a skew-adjoint operator. Moreover, $\mathcal{K}(\bar{\nu}, \bar{\lambda})$ is symmetric. Hence,

$$\begin{aligned} \langle \pi, \nu - D_x(\mathcal{K}(\bar{\nu}, \bar{\lambda})\nu + \mathcal{K}(\nu, \bar{\lambda})\bar{\nu}) \rangle &= \langle \pi, \nu \rangle + \langle D_x\pi, \mathcal{K}(\bar{\nu}, \bar{\lambda})\nu + \mathcal{K}(\nu, \bar{\lambda})\bar{\nu} \rangle \\ &= \langle \pi, \nu \rangle + \langle \mathcal{K}(\bar{\nu}, \bar{\lambda})D_x\pi, \nu \rangle + \langle D_x\pi, \mathcal{K}(\nu, \bar{\lambda})\bar{\nu} \rangle. \end{aligned}$$

It remains to isolate ν in $\langle D_x\pi, \mathcal{K}(\nu, \bar{\lambda})\bar{\nu} \rangle$. With

$$\int_{\Omega} f(x)g(x+y) dx = \int_{y+\Omega} f(z-y)g(z) dz = \int_{\Omega} f(x-y)g(x),$$

for $f, g \in L^2(\Omega)$ (the last equality is due to the periodic boundary conditions, see Remark 3.19 (ii)) and with the antisymmetry of ω , we derive

$$\begin{aligned} & \int_{-1}^1 \int_{\Omega} D_x p(x) S_u \bar{u}(x) u(x+y) \omega(y) dx dy \\ &= \int_{-1}^1 \int_{\Omega} D_x p(x-y) S_u \bar{u}(x-y) u(x) \omega(y) dx dy \\ &= - \int_{\Omega} u(x) \int_{-1}^1 S_u D_x p(x+y) \bar{u}(x+y) \omega(y) dy dx. \end{aligned}$$

3. Necessary Optimality Conditions

Therefore,

$$\begin{aligned}
& \int_{\Omega} D_x p(x) \bar{u}(x) K(u, v, S_u, C)(x) dx \\
&= - \int_{\Omega} u(x) \int_{-1}^1 S_u D_x p(x+y) \bar{u}(x+y) \omega(y) dy \\
& \quad + v(x) \int_{-1}^1 C D_x p(x+y) \bar{u}(x+y) \omega(y) dy dx.
\end{aligned} \tag{3.53}$$

Overall, we have

$$\begin{aligned}
\langle D_x p, (\mathcal{K}(\boldsymbol{\nu}, \bar{\boldsymbol{\lambda}}) \bar{\boldsymbol{\nu}})_1 \rangle &= - \int_{\Omega} u(x) \int_{-1}^1 S_u D_x p(x+y) \bar{u}(x+y) \omega(y) dy \\
& \quad + v(x) \int_{-1}^1 C D_x p(x+y) \bar{u}(x+y) \omega(y) dy dx, \\
\langle D_x q, (\mathcal{K}(\boldsymbol{\nu}, \bar{\boldsymbol{\lambda}}) \bar{\boldsymbol{\nu}})_2 \rangle &= - \int_{\Omega} v(x) \int_{-1}^1 S_v D_x q(x+y) \bar{v}(x+y) \omega(y) dy \\
& \quad + u(x) \int_{-1}^1 C D_x q(x+y) \bar{v}(x+y) \omega(y) dy dx.
\end{aligned}$$

After some rearrangements, we receive

$$\langle D_x \boldsymbol{\pi}, \mathcal{K}(\boldsymbol{\nu}, \bar{\boldsymbol{\lambda}}) \bar{\boldsymbol{\nu}} \rangle = - \langle \mathcal{K}(D_x \boldsymbol{\pi} \odot \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}}) \mathbf{id}, \boldsymbol{\nu} \rangle.$$

And finally, $F_{\boldsymbol{\nu}}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})^*[\boldsymbol{\pi}] = \boldsymbol{\pi} + \mathcal{K}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}}) D_x \boldsymbol{\pi} - \mathcal{K}(D_x \boldsymbol{\pi} \odot \bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}}) \mathbf{id}$.

(ii) We start with

$$\langle \boldsymbol{\pi}, F_{\boldsymbol{\lambda}}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})[\boldsymbol{\lambda}] \rangle = \langle \boldsymbol{\pi}, -D_x(\mathcal{K}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}}) \bar{\boldsymbol{\nu}}) \rangle.$$

The first component yields

$$\begin{aligned}
& \int_{\Omega} p(x) \left(D_x(\bar{u}(x) K(\bar{u}, \bar{v}, S_u, C)(x)) \right) dx \\
&= \left(\int_{\Omega} \int_{-1}^1 p(x) D_x(\bar{u}(x) \bar{u}(x+y)) \omega(y) dy dx S_u \right. \\
& \quad \left. + \int_{\Omega} \int_{-1}^1 p(x) D_x(\bar{u}(x) \bar{v}(x+y)) \omega(y) dy dx C \right) \\
&= \left(\int_{\Omega} p(x) \hat{\mathcal{K}}_u(\bar{\boldsymbol{\nu}})(x) dx, \boldsymbol{\lambda} \right) \\
&= \left(\langle \hat{\mathcal{K}}_u(\bar{\boldsymbol{\nu}}), p \rangle, \boldsymbol{\lambda} \right).
\end{aligned}$$

Analogously, for the second component holds

$$\int_{\Omega} q(x) \left(D_x(\bar{v}(x) K(\bar{v}, \bar{u}, S_v, C)(x)) \right) dx = \left(\langle \hat{\mathcal{K}}_v(\bar{\boldsymbol{\nu}}), q \rangle, \boldsymbol{\lambda} \right),$$

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

with $\hat{\mathcal{K}}_u(\boldsymbol{\nu})$ and $\hat{\mathcal{K}}_v(\boldsymbol{\nu})$ being the first and second column of $\hat{\mathcal{K}}(\boldsymbol{\nu})$, respectively. Hence,

$$\langle \boldsymbol{\pi}, F_\lambda(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\lambda}})[\boldsymbol{\lambda}] \rangle = \langle \hat{\mathcal{K}}(\bar{\boldsymbol{\nu}}), \boldsymbol{\pi}, \boldsymbol{\lambda} \rangle.$$

□

Theorem 3.10 provides a Lagrange functional l and we can formulate the necessary optimality conditions for the control problem (3.50). First, we present a representation of the functional $l \in Z^* = (C([0, T], (H^1(\Omega))^2))^*$. The embedding of $C([0, T], (H^1(\Omega))^2)$ into $L^2(0, T; (H^1(\Omega))^2)$ is continuous with dense image. Thus, we obtain the smaller dual space $(L^2(0, T; (H^1(\Omega))^2))^*$. We further assume, that l can be represented with the adjoint function $\boldsymbol{\pi} \in L^2(0, T; (H^1(\Omega))^2)$.

Lemma 3.21. *The linear functional $l \in (C([0, T], (H^1(\Omega))^2))^*$ in the setting outlined above is given by*

$$\begin{aligned} l(\boldsymbol{\nu}) = & - \int_0^T \int_\Omega \left(\boldsymbol{\pi}(t) + \mathcal{K}(\bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t)) D_x \boldsymbol{\pi}(t, x) - \mathcal{K}(D_x \boldsymbol{\pi}(t) \odot \bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t)) \mathbf{id} \right)^\top \boldsymbol{\nu}(t, x) \, dx \, dt \\ & - \int_\Omega (\bar{\boldsymbol{\nu}}(T, x) - \boldsymbol{\nu}_{obs}(x))^\top \boldsymbol{\nu}(T, x) \, dx, \end{aligned} \quad (3.54)$$

where the adjoint state $\boldsymbol{\pi} = (p, q) \in L^2(0, T; (H^1(\Omega))^2)$ is the mild solution

$$\begin{aligned} \boldsymbol{\pi}(t) = & e^{-(T-t)A} (\bar{\boldsymbol{\nu}}(T) - \boldsymbol{\nu}_{obs}) \\ & + \int_t^T e^{-(s-t)A} \left(\boldsymbol{\pi}(s) + \mathcal{K}(\bar{\boldsymbol{\nu}}(s), \bar{\boldsymbol{\lambda}}(s)) D_x \boldsymbol{\pi}(s) - \mathcal{K}(D_x \boldsymbol{\pi}(s) \odot \bar{\boldsymbol{\nu}}(s), \bar{\boldsymbol{\lambda}}(s)) \mathbf{id} \right) \, ds \end{aligned} \quad (3.55)$$

of the terminal value problem

$$\begin{aligned} -\boldsymbol{\pi}'(t) &= D_{xx} \boldsymbol{\pi}(t) + \mathcal{K}(\bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t)) D_x \boldsymbol{\pi}(t) - \mathcal{K}(D_x \boldsymbol{\pi}(t) \odot \bar{\boldsymbol{\nu}}(t), \bar{\boldsymbol{\lambda}}(t)) \mathbf{id} \\ \boldsymbol{\pi}(T) &= \bar{\boldsymbol{\nu}}(T) - \boldsymbol{\nu}_{obs} \end{aligned}$$

for all $t \in (0, T)$, with \mathbf{id} being the identity in $(H^1(\Omega))^2$ and \odot being the Hadamard product.

Proof. From Theorem 3.9 we have

$$\boldsymbol{\pi}(t) = e^{-(T-t)A} (\bar{\boldsymbol{\nu}}(T) - \boldsymbol{\nu}_{obs}) + \int_t^T e^{-(s-t)A} F_\nu(\bar{\boldsymbol{\nu}}(s), \bar{\boldsymbol{\lambda}}(s))^* [\boldsymbol{\pi}(s)] \, ds, \quad t \in (0, T),$$

since $A = 1 - D_{xx} = A^*$ is a self-adjoint operator in $(L^2(\Omega))^2$. Theorem 3.20 (i) concludes the proof. □

With Theorem 3.10, we obtain the necessary optimality conditions for the control problem (3.50).

Corollary 3.22. *Given optimal control $\bar{\lambda} \in \Lambda$ and optimal state $\bar{\nu} \in L^2(0, T; (H^1(\Omega))^2)$, then the equation*

$$\int_0^T \langle \hat{K}(\bar{\nu}(t)), \pi(t) \rangle, \lambda(t) - \bar{\lambda}(t) \geq 0 \quad (3.56)$$

holds for all $\lambda \in \Lambda$, where $\pi \in L^2(0, T; (H^1(\Omega))^2)$ is the mild solution given by Equation (3.55).

Proof. We set $\mu = \lambda - \bar{\lambda}$. The corollary follows then directly from $J_\lambda(\bar{\nu}, \bar{\lambda})(\lambda) = 0$, Theorem 3.20 (ii) and Theorem 3.10 \square

3.3.4 Necessary Optimality Conditions for the Partial Integro-Differential Equation

With the results on the abstract version of Equation (3.33) we are now able to postulate necessary optimality conditions on the PIDE version that we restate here for the sake of completeness. Let the time dependent control problem with the least squares type objective functional

$$\frac{1}{2} \int_\Omega (u(T, x) - u_{obs}(x))^2 + (v(T, x) - v_{obs}(x))^2 dx,$$

be given. The functions $u(\cdot, \cdot)$ and $v(\cdot, \cdot)$ solve the initial value problem

$$\begin{aligned} u_t &= u_{xx} - (uK_u(u, v))_x, & u(0, x) &= u_0(x), & u(t, a) &= u(t, b) \\ v_t &= v_{xx} - (vK_v(u, v))_x, & v(0, x) &= v_0(x), & v(t, a) &= v(t, b) \end{aligned}$$

with

$$K_u(u, v)(t, x) = \int_{-1}^1 S_u(t)u(t, x + y)\omega(y) + C(t)v(t, x + y)\omega(y) dy$$

and

$$K_v(u, v)(t, x) = \int_{-1}^1 S_v(t)v(t, x + y)\omega(y) + C(t)u(t, x + y)\omega(y) dy.$$

Moreover, let optimal control functions $\bar{S}_u(t)$, $\bar{S}_v(t)$ and $\bar{C}(t)$ and optimal state functions $\bar{u}(t)$ and $\bar{v}(t)$ be given. It then follows from Lemma 3.21 and Corollary 3.22, that all admissible control functions $S_u(t)$, $S_v(t)$ and $C(t)$ satisfy

$$\int_0^T \int_\Omega p(t, x) (\bar{u}(t, x) \hat{K}_u(\bar{u}, \bar{v})(t, x))_x + q(t, x) (\bar{v}(t, x) \hat{K}_v(\bar{u}, \bar{v})(t, x))_x dx dt \geq 0$$

with

$$\hat{K}_u(u, v)(t, x) = \int_{-1}^1 (S_u(t) - \bar{S}_u(t))u(t, x + y)\omega(y) + (C(t) - \bar{C}(t))v(t, x + y)\omega(y) dy$$

and

$$\hat{K}_v(u, v)(t, x) = \int_{-1}^1 (S_v(t) - \bar{S}_v(t))v(t, x + y)\omega(y) + (C(t) - \bar{C}(t))u(t, x + y)\omega(y) dy,$$

3.3. Necessary Optimality Conditions for a Class of Partial Integro-Differential Equations

where $p(\cdot, \cdot)$ and $q(\cdot, \cdot)$ solve the terminal value problem

$$\begin{aligned} -p_t &= p_{xx} + p_x K_u(\bar{u}, \bar{v}) - K_u(p_x \bar{u}, q_x \bar{v}), & p(T, x) &= \bar{u}(T, x) - u_{obs}(x), & p(t, a) &= p(t, b), \\ -q_t &= q_{xx} + q_x K_v(\bar{u}, \bar{v}) - K_v(p_x \bar{u}, q_x \bar{v}), & q(T, x) &= \bar{v}(T, x) - v_{obs}(x), & q(t, a) &= q(t, b). \end{aligned}$$

Part II.

On the Gaussian Kernel: Diffusive Effect, Spectrum and Discretization

Aspects of the Gaussian Kernel

In this chapter, we study characteristics of the Gaussian kernel for diminishing widths. The one dimensional Gaussian kernel is defined as

$$\Gamma_\delta(x) = \frac{1}{\sqrt{2\pi\delta}} e^{-\frac{x^2}{2\delta^2}}, \quad (4.1)$$

where δ specifies its width. Throughout this chapter, we provide results showing that the equation

$$u_t(t, x) = \Gamma_\delta * u(t, x) - u(t, x) \quad (4.2)$$

is well approximated by the heat equation with initial data $u_0(x)$ for small values of δ .

Before we do so, we establish a connection to the first part of this thesis by recalling the Gauss-Weierstrass semigroup $\{(e^{t\Delta})_{\text{Re } t > 0}\}$ that we introduced in Example 2.19. The Gauss-Weierstrass semigroup solves the heat equation, $u_t(t, x) = u_{xx}(t, x)$, with initial data $u_0(x)$ and diffusion constant $d = 1$. The solution $u \in C^{1,2}((0, T] \times \mathbb{R}) \cap C^0([0, T] \times \mathbb{R})$ can be represented by the convolution of the initial data with the Gauss-Weierstrass kernel G_t ,

$$u(t, x) = G_t * u_0(x),$$

where G_t is defined as

$$G_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}.$$

In case of an arbitrary diffusion constant d , we use an adapted version of the Gauss-Weierstrass kernel $\tilde{G}_{t,d}(x) = G_{td}(x)$. Hence, for the diffusion constant $\frac{\delta^2}{2}$ the solution to $u_t(t, x) = \frac{\delta^2}{2} u_{xx}(t, x)$ with initial data u_0 is then

$$u(t, x) = \tilde{G}_{t, \frac{\delta^2}{2}} * u_0(x) = G_{t\frac{\delta^2}{2}} * u_0(x) = \frac{1}{\sqrt{2\pi t\delta}} \int_{\mathbb{R}} u_0(y) e^{-\frac{(x-y)^2}{2t\delta^2}} dy.$$

If we set $t = \delta^2/2$, the Gauss-Weierstrass kernel $G_{\delta^2/2}$ becomes the Gaussian kernel. We consider the partial differential equation $v_\delta(\delta, x) = \delta v_{xx}(\delta, x)$ with the same initial data

$u_0(x)$. Then, $v(\delta, x) = \Gamma_\delta * u_0(x)$ is the solution to this equation. Indeed, it is obvious that $v(\delta, x) = G_{\delta^2/2} * u_0(x) = u(\delta^2/2, x)$. Therefore,

$$v_\delta(\delta, x) = \delta u_t(\delta^2/2, x), \quad v_x(\delta, x) = u_x(\delta^2/2, x), \quad v_{xx}(\delta, x) = u_{xx}(\delta^2/2, x),$$

and since u is the solution to the heat equation, we receive $v_\delta(\delta, x) = \delta v_{xx}(\delta, x)$.

If the convolution with Γ_δ was a semigroup generated by $\delta\Delta$, then the definition of the generator would yield

$$\lim_{\delta \rightarrow 0} \frac{\Gamma_\delta * u_0 - u_0}{\delta} = \delta\Delta u_0$$

and

$$\lim_{\delta \rightarrow 0} \Gamma_\delta * u_0 - u_0 - \delta^2\Delta u_0 = 0;$$

but since the operator A of the corresponding evolution equation depends on δ , $A(\delta) = \delta\Delta$, the theory of quasi-linear evolution equations would need to be applied here in order to get a sophisticated solution. Nevertheless, this lax calculation and the result on the diffusive effect of the Gaussian kernel by Briani et al. (2004) motivate for the following considerations of this chapter.

4.1 The Diffusive Effect of the Gaussian Kernel

To apply Proposition 6.1 from Briani et al. (2004), which shows a diffusive effect of the Gaussian kernel, we perform a simple transformation of Equation (4.2). A similar result is then established for both right-hand side operators in the Fourier space.

We denote by Λ_δ the convolution with the Gaussian kernel, $\Lambda_\delta u(x) = \Gamma_\delta * u(x)$. With $z = -y$, the symmetry of Γ_δ and the unity of $\int_{\mathbb{R}} \Gamma_\delta(z) dz$ we obtain

$$\begin{aligned} \Lambda_\delta u(t, x) - u(t, x) &= \int_{-\infty}^{\infty} u(t, x - y) \Gamma_\delta(y) dy - u(t, x) \\ &= \int_{\infty}^{-\infty} -u(t, x + z) \Gamma_\delta(-z) dz - u(t, x) \\ &= \int_{-\infty}^{\infty} u(t, x + z) \Gamma_\delta(z) dz - u(t, x) \\ &= \int_{-\infty}^{\infty} [u(t, x + z) - u(t, x)] \Gamma_\delta(z) dz. \end{aligned} \tag{4.3}$$

We can now apply (Briani et al., 2004, Proposition 6.1) which shows a diffusive effect of the Gaussian kernel.

Theorem 4.1 (Briani et al., 2004, Proposition 6.1). *Let $u(t, x)$ be the solution to*

$$u_t(t, x) + au_x(t, x) - bu_{xx}(t, x) + cu(t, x) = \int_{-\infty}^{\infty} [u(t, x + z) - u(t, x)] \Gamma_\delta(z) dz \tag{4.4}$$

and $v(t, x)$ be the solution to

$$v_t(t, x) + av_x(t, x) - bv_{xx}(t, x) + cv(t, x) = \frac{\delta^2}{2}v_{xx}(t, x)$$

on $(0, T) \times \mathbb{R}$ with the same initial condition $u_0(x) = u(0, x) = v(0, x)$, $u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Then, if $\delta \ll 1$, there holds

$$\|u - v\|_{L^\infty(0, T; L^1(\mathbb{R}))} \leq \mathcal{O}(T\delta^3).$$

We obtain an analogous result for the operator

$$\Lambda_\delta - \text{id} - \frac{\delta^2}{2}\Delta \tag{4.5}$$

in the space of three-times differentiable functions.

Theorem 4.2. *The operator norm of $\Lambda_\delta - \text{id} - \frac{\delta^2}{2}\Delta$ on $C^3(\mathbb{R})$ satisfies the estimate*

$$\|\Lambda_\delta - \text{id} - \frac{\delta^2}{2}\Delta\|_{C^3(\mathbb{R}) \rightarrow C^1(\mathbb{R})} \leq \frac{1}{\sqrt{2\pi}} \frac{2}{3} \delta^3 = \mathcal{O}(\delta^3).$$

Proof. Let $\phi \in C^3(\mathbb{R})$ be a three-times differentiable function and $x, y \in \mathbb{R}$ be given. Taking the Taylor expansion of $\phi(x + y)$ around x leads to

$$\phi(x + y) = \phi(x) + y\phi_x(x) + \frac{y^2}{2}\phi_{xx}(x) + \frac{y^3}{6}\phi_{xxx}(\xi)$$

with a $\xi \in [x, x + y]$. Equation (4.3) then yields

$$\begin{aligned} \Lambda_\delta\phi(x) - \phi(x) - \frac{\delta^2}{2}\Delta\phi(x) &= \int_{-\infty}^{\infty} [\phi(x + y) - \phi(x) - \frac{\delta^2}{2}\phi_{xx}(x)]\Gamma_\delta(y) \, dy \\ &= \int_{-\infty}^{\infty} [y\phi_x(x) + \frac{y^2 - \delta^2}{2}\phi_{xx}(x) + \frac{y^3}{6}\phi_{xxx}(\xi)]\Gamma_\delta(y) \, dy. \end{aligned}$$

With simple calculations, we obtain

$$\int_{-\infty}^{\infty} [y\phi_x(x) + \frac{y^2 - \delta^2}{2}\phi_{xx}(x)]\Gamma_\delta(y) \, dy = 0$$

and

$$|\int_{-\infty}^{\infty} \frac{y^3}{6}\phi_{xxx}(\xi)\Gamma_\delta(y) \, dy| \leq \frac{1}{\sqrt{2\pi}} \|\phi\|_{C^3(\mathbb{R})} \int_{-\infty}^{\infty} \frac{|y|^3}{6\delta} e^{-\frac{y^2}{2\delta^2}} \, dy = \frac{\delta^3}{\sqrt{2\pi}} \frac{2}{3} \|\phi\|_{C^3(\mathbb{R})}.$$

Indeed, L'Hôpital's rule yields

$$\int_{-\infty}^{\infty} [y\phi_x(x) + \frac{y^2 - \delta^2}{2}\phi_{xx}(x)]e^{-\frac{y^2}{2\delta^2}} \, dy = \lim_{R \rightarrow \infty} \left(-\delta^2 e^{-\frac{y^2}{2\delta^2}} (\phi_x(x) + \frac{y}{2}\phi_{xx}(x)) \Big|_{-R}^R \right) = 0$$

and with standard calculus we receive

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{|y|^3}{\delta} e^{-\frac{y^2}{2\delta^2}} dy &= \int_0^{\infty} \frac{y^3}{\delta} e^{-\frac{y^2}{2\delta^2}} dy - \int_{-\infty}^0 \frac{y^3}{\delta} e^{-\frac{y^2}{2\delta^2}} dy \\ &= \lim_{R \rightarrow \infty} \left(-\delta(2\delta^2 + y^2)e^{-\frac{y^2}{2\delta^2}} \Big|_0^R + \delta(2\delta^2 + y^2)e^{-\frac{y^2}{2\delta^2}} \Big|_{-R}^0 \right) = 4\delta^3. \end{aligned}$$

Hence,

$$\|\Lambda_\delta \phi - \phi - \frac{\delta^2}{2} \Delta \phi\|_{C^1(\mathbb{R})} \leq \frac{\delta^3}{\sqrt{2\pi}} \frac{2}{3} \|\phi\|_{C^3(\mathbb{R})},$$

which concludes the proof. \square

For the next result, we apply the Fourier transform to the operator $\Lambda_\delta - \text{id} - \frac{\delta^2}{2} \Delta$ and study its behavior with respect to $\delta \rightarrow 0$. For the calculation of the Fourier transform of $I - \Lambda_\delta$, we need the following lemma.

Lemma 4.3. *The Fourier transform of $\Gamma_\delta(x) = \frac{1}{\sqrt{2\pi\delta}} e^{-\frac{x^2}{2\delta^2}}$ is given by*

$$\mathcal{F}(\Gamma_\delta)(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2} \xi^2}.$$

Whether we first apply the Fourier transform to Γ_δ and then consider the limit as δ approaches zero or first let δ approach zero and then apply the Fourier transform, we obtain the same result, which allows us to draw the commutative diagram, where δ_0 is the Dirac delta function:

$$\begin{array}{ccc} \Gamma_\delta & \xrightarrow{\mathcal{F}} & \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2}(\cdot)^2} \\ \delta \rightarrow 0 \downarrow & & \downarrow \delta \rightarrow 0 \\ \delta_0 & \xrightarrow{\mathcal{F}} & \frac{1}{\sqrt{2\pi}} \mathbf{1}_{\mathbb{R}} \end{array}$$

Proof. We present two different approaches to derive the Fourier transform of the Gaussian kernel. The first approach makes use of differential polynomials. We define the polynomial P_Γ as

$$P_\Gamma(x) = \frac{x}{\delta^2}.$$

Γ_δ solves the ordinary differential equation

$$\Gamma'_\delta + \frac{x}{\delta^2} \Gamma_\delta = 0, \quad \Gamma_\delta(0) = \frac{1}{\sqrt{2\pi\delta}}.$$

Applying the Fourier transform to both sides of the equation and using (Folland, 1992, Theorem 7.5) we obtain

$$\begin{aligned} 0 &= \mathcal{F}(0)(\xi) = \mathcal{F}(\Gamma'_\delta + \frac{x}{\delta^2} \Gamma_\delta)(\xi) = \mathcal{F}(\Gamma'_\delta)(\xi) + \mathcal{F}(\frac{x}{\delta^2} \Gamma_\delta)(\xi) \\ &= \delta^2 \mathcal{F}(P_\Gamma(\delta) \Gamma_\delta)(\xi) + \mathcal{F}(P_\Gamma(x) \Gamma_\delta)(\xi) = \delta^2 i \frac{\xi}{\delta^2} \widehat{\Gamma}_\delta(\xi) + \frac{i}{\delta^2} \widehat{\Gamma}'_\delta(\xi) = i \xi \widehat{\Gamma}_\delta(\xi) + \frac{i}{\delta^2} \widehat{\Gamma}'_\delta(\xi). \end{aligned}$$

Consequently, $\widehat{\Gamma}_\delta$ solves the ordinary differential equation

$$\widehat{\Gamma}'_\delta + \delta^2 \xi \widehat{\Gamma}_\delta = 0, \quad \widehat{\Gamma}_\delta(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Gamma_\delta(x) dx = \frac{1}{\sqrt{2\pi}}$$

which yields

$$\widehat{\Gamma}_\delta(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2} \xi^2}.$$

The second approach is a direct calculation of the Fourier transform. By definition, we have

$$\begin{aligned} \mathcal{F}(\Gamma_\delta)(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\delta} e^{-\frac{x^2}{2\delta^2}} e^{-i\xi x} dx \\ &= \frac{1}{2\pi\delta} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\delta^2} - i\xi x} dx. \end{aligned}$$

We note, that $-\left(\frac{x^2}{2\delta^2} + i\xi x\right) = -\left(\frac{x}{\sqrt{2}\delta} + i\sqrt{2}\delta\frac{\xi}{2}\right)^2 - \frac{\delta^2\xi^2}{2}$ by completing the square. Thus,

$$\widehat{\Gamma}_\delta(\xi) = \frac{1}{2\pi\delta} \int_{-\infty}^{\infty} e^{-\left(\frac{x}{\sqrt{2}\delta} + i\sqrt{2}\delta\frac{\xi}{2}\right)^2 - \frac{\delta^2\xi^2}{2}} dx.$$

Substituting $\beta = \frac{x}{\sqrt{2}\delta} + i\sqrt{2}\delta\frac{\xi}{2}$ we obtain

$$\begin{aligned} \widehat{\Gamma}_\delta(\xi) &= \frac{\sqrt{2}\delta}{2\pi\delta} e^{-\frac{\delta^2\xi^2}{2}} \int_{-\infty}^{\infty} e^{-\beta^2} d\beta \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2\xi^2}{2}} \sqrt{\pi} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2} \xi^2}. \end{aligned}$$

In order to proof that the commutative diagram is well defined, we need to proof the convergence of Γ_δ for $\delta \rightarrow 0$, since $\mathcal{F}\delta_0 = \sqrt{2\pi}\mathbb{1}_{\mathbb{R}}$ is a well-known result. We consider Γ_δ as a distribution and apply it to a test function $\phi \in \mathcal{D}$,

$$\Gamma_\delta(\phi) = \int \frac{1}{\sqrt{2\pi}\delta} e^{-\frac{x^2}{2\delta^2}} \phi(x) dx.$$

With the substitution $x = \delta y$ we have

$$\lim_{\delta \rightarrow 0} \Gamma_\delta(\phi) = \lim_{\delta \rightarrow 0} \int \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \phi(\delta y) dy = \frac{1}{\sqrt{2\pi}} \int_{\delta \rightarrow 0} e^{-\frac{y^2}{2}} dy \quad \phi(0) = \phi(0) = \delta_0(\phi).$$

□

Theorem 4.4. *The Fourier transform of the operators at hand, $\frac{\delta^2}{2}\Delta$ and $\Lambda_\delta - \text{id}$, are given by*

$$\mathcal{F}\left(\frac{\delta^2}{2}\Delta u\right)(\xi) = -\frac{\delta^2}{2}\xi^2 \widehat{u}(\xi), \tag{4.6}$$

$$\mathcal{F}(\Lambda_\delta u - u)(\xi) = e^{-\frac{\delta^2}{2}\xi^2} \widehat{u}(\xi) - \widehat{u}(\xi). \tag{4.7}$$

Proof. In order to calculate the Fourier transform of $\frac{\delta^2}{2}\Delta$ we define the polynomial P_Δ as

$$P_\Delta(x) = \frac{\delta^2}{2}x^2, \quad P_\Delta(\partial)u = \frac{\delta^2}{2}u''$$

and with (Folland, 1992, Theorem 7.5) we conclude

$$\mathcal{F}\left(\frac{\delta^2}{2}\Delta u\right)(\xi) = \mathcal{F}(P_\Delta(\partial)u)(\xi) = P_\Delta(i\xi)\widehat{u}(\xi) = -\frac{\delta^2}{2}\xi^2\widehat{u}(\xi).$$

With Lemma 4.3 we receive

$$\begin{aligned} \mathcal{F}(\Lambda_\delta u - u)(\xi) &= \mathcal{F}(\Lambda_\delta u)(\xi) - \mathcal{F}(u)(\xi) = \mathcal{F}(\Gamma_\delta * u)(\xi) - \mathcal{F}(u)(\xi) \\ &= \sqrt{2\pi}\mathcal{F}(\Gamma_\delta)(\xi)\mathcal{F}(u)(\xi) - \mathcal{F}(u)(\xi) = e^{-\frac{\delta^2}{2}\xi^2}\widehat{u}(\xi) - \widehat{u}(\xi) \end{aligned}$$

□

We next derive a result on the L^2 operator norm of the Fourier transform of Equation (4.5) on a specific interval.

Theorem 4.5. *Let $\delta \ll 1$ and $I := [-\frac{1}{\delta}, \frac{1}{\delta}]$. Then for the operator given by Equation (4.5), the following estimate holds*

$$\|\mathcal{F}(\Lambda_\delta - \text{id} - \frac{\delta^2}{2}\Delta)\|_{L^2(I) \rightarrow L^2(I)} = \mathcal{O}(\delta^4).$$

Proof. Let us first examine the Fourier transform of $(\Lambda_\delta - \text{id})u - \frac{\delta^2}{2}\Delta u$ in $L^2(\mathbb{R})$ and use the series expansion of the exponential function. Let $\xi \in I$ be given, then

$$\mathcal{F}\left((\Lambda_\delta - \text{id} - \frac{\delta^2}{2}\Delta)u\right)(\xi) = (e^{-\frac{\delta^2}{2}\xi^2} - 1 + \frac{\delta^2}{2}\xi^2)\widehat{u}(\xi) = \sum_{\nu=2}^{\infty} \frac{(-\frac{\delta^2}{2}\xi^2)^\nu}{\nu!}\widehat{u}(\xi),$$

and

$$\left\| \sum_{\nu=2}^{\infty} \frac{(-\frac{\delta^2}{2}(\cdot)^2)^\nu}{\nu!}\widehat{u} \right\|_{L^2(I)} \leq \sup_{\xi \in I} \left| \sum_{\nu=2}^{\infty} \frac{(-\frac{\delta^2}{2}\xi^2)^\nu}{\nu!} \right| \|\widehat{u}\|_{L^2(I)},$$

since

$$\sum_{\nu=2}^{\infty} \frac{(-\frac{\delta^2}{2}\xi^2)^\nu}{\nu!} < e^{-\frac{\delta^2}{2}\xi^2}$$

for all $\xi \in I$ and the series converges on that interval. So,

$$\begin{aligned} \|\mathcal{F}(\Lambda_\delta - \text{id} - \frac{\delta^2}{2}\Delta)\|_{L^2(I) \rightarrow L^2(I)} &\leq \sup_{\xi \in I} \left| \sum_{\nu=2}^{\infty} \frac{(-\frac{\delta^2}{2}\xi^2)^\nu}{\nu!} \right| \\ &\leq \sup_{\xi \in I} \frac{\delta^4 \xi^4}{8} + \left| \sum_{\nu=3}^{\infty} \frac{(-\frac{\delta^2}{2}\xi^2)^\nu}{\nu!} \right| \\ &= \mathcal{O}(\delta^4). \end{aligned}$$

□

The Gaussian Kernel in a Partial Integro-Differential Equation

The Black-Scholes model, that simulates the price for European call options, has been extended in many ways since its release by Black and Scholes (1973). One extension among others is the jump-diffusion model in Merton (1976). Under certain assumptions this model leads to a partial integro-differential equation that involves the Gaussian kernel as the nonlocal integral term.

With a variable transformation, the authors overcome in Sachs and Strauss (2008) the numerical instability caused by the model's convection term. This transformation results in an equivalent problem with the aim to find a solution $w \in C^{1,2}((0, T] \times \mathbb{R}) \cap C^0([0, T] \times \mathbb{R})$ to the partial integro-differential equation

$$w_t - \frac{1}{2}\sigma^2 w_{xx} + (r + \lambda)w - \lambda \int_{-\infty}^{\infty} w(t, z) \Gamma_{\delta}(z - x) dz = 0 \quad \text{on } (0, T] \times \mathbb{R}, \quad (5.1)$$

$$w(0, x) = H(e^x), \quad \text{for all } x \in \mathbb{R},$$

where $H(x) = \max\{0, x - K\}$ is the payoff function with strike price K and $\sigma, \lambda, T > 0$ as well as $r \geq 0$ are constants, cf. Sachs and Strauss (2008) for more details.

The numerical solution of Equation (5.1) requires the computation of linear systems with dense coefficient matrices. Due to the structure of those coefficient matrices, the conjugate gradient (CG) method can be used. The convergence behavior of the CG method relies on the condition of the coefficient matrix and therefore circulant preconditioners are used in Sachs and Strauss (2008). In Ye (2013), the author suggests a tridiagonal preconditioner that only acts on the PDE part of the coefficient matrix.

In this chapter, we focus on the integral part of Equation (5.1) and study the equation

$$u_t(t, x) + \lambda \left(u(t, x) - \int_{-\infty}^{\infty} u(t, x) \Gamma_{\delta}(z - x) dz \right) = 0 \quad \text{on } (0, T] \times \mathbb{R}, \quad (5.2)$$

$$u(0, x) = h(x), \quad \text{for all } x \in \mathbb{R},$$

with some initial function $h \in C(\mathbb{R})$. We note that Equation (5.2) is equal to Equation (4.2) for $\lambda = 1$. In the first section, we discuss the numerical discretization and analyze the spectrum of the resulting coefficient matrix. In the second section, we discuss the need for preconditioning and compare the eigenvalue spectra of the linear system and the preconditioned system for different values of δ . In the third section, we focus on the discretization of the term in brackets of Equation (5.2) and propose an alternative discretization scheme that is more suitable for small values of δ . We conclude this chapter by providing some numerical results.

5.1 Discretization and Spectrum of the Coefficient Matrix

To solve Equation (5.2) numerically, we need to use a proper truncation and discretization of \mathbb{R} . We restrict the domain to $\Omega := (-R, R)$ with R sufficiently large. For a given $\epsilon > 0$ the sensing radius R_δ of Γ_δ around $x = 0$ is given by $R_\delta = \sqrt{-2\delta^2 \ln(\epsilon\delta\sqrt{2\pi})}$, which is derived by simple calculation, i.e., $\Gamma_\delta < \epsilon$ for $|x| > R_\delta$. For the subsequent discussion, we set $R \gg R_\delta$ and impose Dirichlet boundary conditions, i.e., $u(t, x) \rightarrow 0$ for $x \rightarrow \pm\infty$. That results in

$$u_t(t, x) + \lambda u(t, x) - \lambda \int_{-R}^R u(t, z) \Gamma_\delta(z - x) dz = \lambda \hat{R}(t, x, R), \quad \text{on } (0, T] \times \Omega, \quad (5.3)$$

with $\hat{R}(t, x, R) = \int_R^\infty u(t, z) \Gamma_\delta(z - x) + u(t, -z) \Gamma_\delta(-z - x) dz$. Due to the boundary conditions and the sensing radius R_δ of Γ_δ the truncation error \hat{R} is negligible for our subsequent discussion.

We follow the discretization scheme that is used in Sachs and Strauss (2008): We set

$$\begin{aligned} x_i &:= -R + (i - 1)h & \text{with } i &= 1, \dots, n + 2, & h &= 2R/(n + 1), \\ t_p &:= p\tau & \text{with } p &= 1, \dots, m, & \tau &= T/m, \end{aligned}$$

with $n \in \mathbb{N}$ odd and denote by u_i^p the approximation of the true solution $u(t_p, x_i)$. To obtain second order accuracy and stability in time, we use a backward difference formula of second order (BDF2) for $p \geq 2$, which is

$$u_t(t_p, x_i) \approx \begin{cases} (\frac{3}{2}u_i^p - 2u_i^{p-1} + \frac{1}{2}u_i^{p-2})/\tau, & \text{for } p \geq 2, \\ (u_i^p - u_i^{p-1})/\tau, & \text{for } p = 1, \end{cases}$$

and use the composite trapezoidal rule to approximate the integral on Ω ,

$$\begin{aligned} \int_{-R}^R u(t_p, z) \Gamma_\delta(z - x_i) dz &\approx \frac{h}{2} (u(t_p, x_1) \Gamma_\delta(x_1 - x_i) + 2 \sum_{j=2}^{n+1} u(t_p, x_j) \Gamma_\delta(x_j - x_i) \\ &\quad + u(t_p, x_{n+2}) \Gamma_\delta(x_{n+2} - x_i)). \end{aligned} \quad (5.4)$$

Using the boundary conditions to eliminate the first and last equation of the resulting linear system and re-indexing the subscripts, i.e., $x_i = -R + ih$ with $i = 1, \dots, n$, we receive a $n \times n$ Toeplitz system

$$T_n u^p = b^p, \quad (5.5)$$

with $b^p = 2u^{p-1} - \frac{1}{2}u^{p-2}$ for $p \geq 2$ and $b^1 = u^{p-1}$. Without loss of generality, we assume $p \geq 2$ in the following. The $2n - 1$ coefficients of the Toeplitz matrix T_n are given by

$$\begin{aligned} t_0 &= \frac{3}{2} + \lambda\tau - \frac{\lambda\tau h}{\sqrt{2\pi\delta}}, \\ t_i &= t_{-i} = -\frac{\lambda\tau h}{\sqrt{2\pi\delta}} e^{-i^2 h^2 / (2\delta^2)}, \end{aligned} \quad (5.6)$$

with $i = 1, \dots, n - 1$. Depending on the context, we will formulate the coefficients in terms of n and m ,

$$\begin{aligned} t_0 &= \frac{3}{2} + \frac{\lambda T}{m} - \frac{\lambda T \alpha}{(n+1)m\delta}, \\ t_i &= t_{-i} = -\frac{\lambda T \alpha}{(n+1)m\delta} e^{-\beta i^2 / ((n+1)\delta)^2}, \end{aligned} \quad (5.7)$$

with $\alpha = 2R/\sqrt{2\pi}$ and $\beta = (2R)^2/2$. A third way we will consider the Toeplitz matrix T_n is as the sum

$$T_n = \frac{3}{2}I + \lambda\tau(I - D_n), \quad (5.8)$$

where I is the identity matrix and $D = (d_{ij})_{i,j=1}^n$ with

$$d_{ij} = \frac{h}{\sqrt{2\pi\delta}} e^{-(i-j)^2 h^2 / (2\delta^2)}.$$

Toeplitz Matrices

We provide a brief introduction to Toeplitz matrices and their generating functions in this subsection.

Definition 5.1. A matrix $T_n \in \mathbb{C}^{n \times n}$ is called *Toeplitz*, if T_n is determined by the $2n - 1$ scalars $t_{-(n-1)}, \dots, t_{-1}, t_0, t_1, \dots, t_{n-1}$ with $T_{ij} = t_{i-j}$, i.e., the Toeplitz matrix is of the following form:

$$T_n = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{2-n} & t_{1-n} \\ t_1 & t_0 & \ddots & & t_{2-n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ t_{n-2} & & \ddots & \ddots & t_{-1} \\ t_{n-1} & t_{n-2} & \cdots & t_1 & t_0 \end{pmatrix}. \quad (5.9)$$

The function $g: [-\pi, \pi] \rightarrow \mathbb{R}$, defined by the Fourier series

$$g(x) = \sum_{j=-\infty}^{\infty} t_j e^{-ijx}, \quad (5.10)$$

is called a *generating function* of the indefinite Toeplitz matrix T_∞ . In other words, the entries of T_∞ are given by

$$t_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{ijx} dx, \quad j = 0, \pm 1, \pm 2, \dots \quad (5.11)$$

If g is real-valued, we have

$$t_{-j} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{-ijx} dx = \bar{t}_j, \quad j = 0, \pm 1, \pm 2, \dots, \quad (5.12)$$

and hence, T_∞ must be Hermitian. If g is also an even function, i.e., $g(-x) = g(x)$, then T_∞ is also real and symmetric.

For finite matrices T_n we set $t_k = 0$ for $k \geq n$, which results in a finite series for the generating function

$$g^{(n)}(x) = \sum_{j=-(n-1)}^{n-1} t_j e^{-ijx}.$$

Theorem 5.2 (Grenander and Szegö, 1984). *Let g be a real-valued function in $L^1[-\pi, \pi]$. Then the spectrum $\sigma(T_n)$ of T_n satisfies*

$$\sigma(T_n) \subset [g_{\min}, g_{\max}], \quad \text{for all } n \geq 1,$$

where g_{\min} and g_{\max} are the essential infimum and the essential supremum of g , respectively. Moreover, if $g_{\max} > g_{\min}$, then

$$g_{\min} < \lambda_{\min}(T_n) \leq \lambda_{\max}(T_n) < g_{\max}.$$

In particular, if $g_{\min} > 0$, then T is positive definite for all $n \in \mathbb{N}$.

Definition 5.3. Let g be a real-valued function in $L^1[-\pi, \pi]$. A sequence $(\alpha_k^{(n)})$ is said to be *equally distributed* as g if

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n F(\alpha_k^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(g(x)) dx$$

for any continuous function F with bounded support.

Theorem 5.4 (Grenander and Szegö, 1984). *Let $g \in L^2[-\pi, \pi]$. Then the singular values of the matrices T_n generated by g are equally distributed as $|g(x)|$. In particular, for real-valued g , the eigenvalues of T_n are equally distributed as $g(x)$. Thus, for g Riemann integrable, the sets of values*

$$\{\lambda_i(T_n)\}_{i=0}^n \quad \text{and} \quad \left\{ g \left(-\pi + \frac{2i\pi}{n+1} \right) \right\}_{i=0}^n$$

are equally distributed, i.e., for every eigenvalue $\lambda_i(T_n)$ exists a $\xi_i \in [0, \pi]$ with $g(\xi_i) = \lambda_i(T_n)$.

Analysis of the Spectrum

The generating function $g^{(n)}$ of the matrix T_n is given by

$$g^{(n)}(x) = \frac{3}{2} + \lambda\tau \left(1 - \frac{h}{\sqrt{2\pi\delta}} \left(1 + 2 \sum_{j=1}^{n-1} e^{-j^2 h^2 / (2\delta^2)} \cos(jx) \right) \right). \quad (5.13)$$

In order to get an estimate of the spectrum $\sigma(T_n)$, we will calculate upper and lower bounds of the corresponding generating function $g^{(n)}$ on $[-\pi, \pi]$ and apply Theorem 5.2.

Adjusting the estimate of Ye (2013) to the generating function (5.13) we receive

$$\begin{aligned} g^{(n)}(x) &\geq \frac{3}{2} + \lambda\tau \left(1 - \frac{h}{\sqrt{2\pi\delta}} - \operatorname{erf} \left(\frac{(n-1)h}{\sqrt{2\delta}} \right) \right) \geq \frac{3}{2} - \lambda\tau \frac{h}{\sqrt{2\pi\delta}}, \\ g^{(n)}(x) &\leq \frac{3}{2} + \lambda\tau \left(1 - \frac{h}{\sqrt{2\pi\delta}} + \operatorname{erf} \left(\frac{(n-1)h}{\sqrt{2\delta}} \right) \right) \leq \frac{3}{2} + \lambda\tau \left(2 - \frac{h}{\sqrt{2\pi\delta}} \right), \end{aligned} \quad (5.14)$$

where

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$$

is the Gauss error function.

Remark 5.5. The above argument of the error function is independent of n due to the definition of h and only depends on δ . For $\delta \ll 1$, we have

$$\operatorname{erf} \left(\frac{(n-1)h}{\sqrt{2\delta}} \right) = \operatorname{erf} \left(\frac{(n-1)}{(n+1)\delta} \sqrt{\beta} \right) \approx 1$$

for any h , which yields the estimate

$$\sigma(T_n) \subset \frac{3}{2} + \lambda\tau \left[-\frac{h}{\sqrt{2\pi\delta}}, 2 - \frac{h}{\sqrt{2\pi\delta}} \right];$$

this is equal to

$$\sigma(T_n) \subset \frac{3}{2} + \frac{\lambda T}{m} \left[-\frac{\alpha}{(n+1)\delta}, 2 - \frac{\alpha}{(n+1)\delta} \right]$$

in terms of n and m . Thus, the matrix T_n is positive definite if $(n+1)\delta$ or m is sufficiently large.

In the subsequent of this subsection, we want to improve the estimate of $\sigma(T_n)$ for $\delta \ll 1$. To this end, we first present two results that we will use to estimate the range of

$$g_D^{(n)}(x) = \frac{h}{\sqrt{2\pi\delta}} \left(1 + 2 \sum_{j=1}^{n-1} e^{-j^2 h^2 / (2\delta^2)} \cos(jx) \right),$$

which is the generating function of the Toeplitz matrix D_n from Equation (5.8).

The first result is from Brown and Hewitt and provides an estimate for certain cosine sums.

Theorem 5.6 (Brown and Hewitt, 1984). *Suppose that $(a_k)_{k=0}^{\infty}$ is a nonincreasing sequence of nonnegative real numbers such that $a_0 > 0$ and*

$$a_{2k} \leq \frac{2k}{2k+1} a_{2k-1}, \quad \text{for } k = 1, 2, 3, \dots \quad (5.15)$$

Then, for all positive integers N , we have

$$a_0 + a_1 \cos \theta + a_2 \cos 2\theta + \dots + a_n \cos N\theta > 0 \quad \text{for } 0 \leq \theta < \pi. \quad (5.16)$$

The second result is by Gawronski and Stadtmüller. The authors use the Fourier series expansion of a certain theta series and reformulate an identity that can be found for example in Butzer and Nessel (1971) or Becker et al. (1976).

Lemma 5.7 (Gawronski and Stadtmüller, 1982). *The functions*

$$\bar{g}_{j,n}(x) := \frac{1}{\sqrt{2\pi n \delta_n}} \exp\left(-\frac{(j-nx)^2}{2n^2 \delta_n^2}\right)$$

with $j \in \mathbb{Z}$, $n \in \mathbb{N}$, $x \in \mathbb{R}$ and $\delta_n > 0$ satisfy the identity

$$\begin{aligned} \bar{g}^{(n)}(x) &= \sum_{j=-\infty}^{\infty} \bar{g}_{j,n}(x) \\ &= \sum_{\nu=-\infty}^{\infty} e^{-2\pi^2 \nu^2 n^2 \delta_n^2} e^{2\pi i \nu x n} \\ &= 1 + 2 \sum_{\nu=1}^{\infty} e^{-2\pi^2 \nu^2 n^2 \delta_n^2} \cos(2\pi \nu x n). \end{aligned} \quad (5.17)$$

We are now able to formulate the following Lemma, which is our main result in this subsection.

Lemma 5.8. *Let D_n be the Toeplitz matrix defined in Equation (5.8) and*

$$\delta \leq \frac{h}{\sqrt{2 \ln 2}}, \quad (5.18)$$

then

$$\sigma(D_n) \subset (0, 1 + c], \quad (5.19)$$

with $c = 2 \frac{e^{-2\pi^2(\delta/h)^2}}{1 - e^{-2\pi^2(\delta/h)^2}}$. If we choose $\delta = \frac{\gamma h}{\sqrt{2 \ln 2}}$ with $\gamma \in (0, 1]$, i.e., for decreasing values of δ we proportionally increase n , then c is independent of n .

Proof. We first proof the lower bound. Let $a_0 = 1$, $a_k = 2e^{-k^2 h^2 / (2\delta^2)}$ for $1 \leq k \leq n-1$ and $a_k = 0$ for $k \geq n$. We show that the coefficients a_k satisfy the prerequisites of Theorem 5.6 and receive $g_D^{(n)}(x) > 0$.

It can be easily seen that $a_i \geq a_j$ for $1 \leq i < j$ and $a_0 \geq a_1$ follows directly from Equation (5.18),

$$\delta \leq \frac{h}{\sqrt{2 \ln 2}} \Rightarrow \delta^2 \leq \frac{h^2}{2 \ln 2} \Leftrightarrow -\frac{h^2}{2\delta^2} \leq \ln\left(\frac{1}{2}\right) \Leftrightarrow e^{-\frac{h^2}{2\delta^2}} \leq \frac{1}{2}.$$

Hence, the sequence $(a_k)_{k=0}^\infty$ is nonincreasing and it is left to show that Equation (5.15) holds. For $x > \frac{1}{4}$ we define $f(x) := \frac{4x-1}{2 \ln(1+1/(2x))}$ and notice that Equation (5.18) can be written as $\delta \leq h\sqrt{f(\frac{1}{2})}$. It is obvious, that (5.15) holds for $2k \geq n$. For $1 \leq 2k \leq n-1$ we have

$$\begin{aligned} 2e^{-\frac{(2k)^2 h^2}{2\delta^2}} &\leq \frac{2k}{2k+1} 2e^{-\frac{(2k-1)^2 h^2}{2\delta^2}} \\ \Leftrightarrow -\frac{(2k)^2 h^2}{2\delta^2} &\leq \ln\left(\frac{2k}{2k+1}\right) - \frac{(2k-1)^2 h^2}{2\delta^2} \\ \Leftrightarrow \ln\left(\frac{2k+1}{2k}\right) &\leq \frac{(2k)^2 h^2}{2\delta^2} - \frac{(2k-1)^2 h^2}{2\delta^2} \\ \Leftrightarrow \ln\left(1 + \frac{1}{2k}\right) &\leq \frac{h^2}{2\delta^2}(4k-1) \\ \Leftrightarrow \delta^2 &\leq h^2 \frac{4k-1}{2 \ln(1+1/(2k))} \\ \Leftrightarrow \delta &\leq h\sqrt{\frac{4k-1}{2 \ln(1+1/(2k))}} \\ \Leftrightarrow \delta &\leq h\sqrt{f(k)}. \end{aligned}$$

With simple calculations, it can be shown that $f(x)$ is positive and increasing for $x > \frac{1}{4}$. Thus, (5.15) holds for $2k \geq n$.

To proof the upper bound of the estimate, we notice that the function $g_D^{(n)}$ has its maximum at $x = 0$. By setting $\delta_n = \bar{\delta} = \delta/(2R)$ for all $n \in \mathbb{N}$ and using $h = \frac{2R}{n+1}$, we receive that $g_D^{(n)}(0) = \sum_{j=-n+1}^{n-1} \bar{g}_{j,n+1}(0)$ with $\bar{g}_{j,n+1}$ as in Lemma 5.7 and $g_D^{(n)}(0) \leq \bar{g}^{(n+1)}(0)$ follows. For all $n \in \mathbb{N}$ and $\bar{\delta} > 0$, we have $e^{-2\pi^2 n^2 \bar{\delta}^2} < 1$ resulting in $e^{-2\pi^2 n^2 \bar{\delta}^2 \nu^2} \leq e^{-2\pi^2 n^2 \bar{\delta}^2 \nu}$ for $\nu \geq 1$. Consequently, we can estimate

$$\begin{aligned} g_D^{(n)}(0) &\leq 1 + 2 \sum_{\nu=1}^{\infty} e^{-2\pi^2 \nu^2 (n+1)^2 \bar{\delta}^2} \\ &\leq 1 + 2 \sum_{\nu=1}^{\infty} e^{-2\pi^2 (n+1)^2 \bar{\delta}^2 \nu} \\ &= 1 + 2 \frac{e^{-2\pi^2 (n+1)^2 \bar{\delta}^2}}{1 - e^{-2\pi^2 (n+1)^2 \bar{\delta}^2}} \\ &= 1 + 2 \frac{e^{-2\pi^2 (\delta/h)^2}}{1 - e^{-2\pi^2 (\delta/h)^2}} \\ &= 1 + c. \end{aligned} \tag{5.20}$$

On the other hand, estimate (5.20) for the upper bound is sharp as $g_D^{(n)}(x) > 1$ for some $x \in [-\pi, \pi]$. \square

As a direct consequence from this lemma, we obtain the following corollary.

Corollary 5.9. *Let T_n be the Toeplitz matrix defined in Equation (5.6) and*

$$\delta \leq \frac{h}{\sqrt{2 \ln 2}},$$

then

$$\sigma(T_n) \subset \frac{3}{2} + \lambda\tau[-c, 1), \quad (5.21)$$

with $c = 2 \frac{e^{-2\pi^2(\delta/h)^2}}{1 - e^{-2\pi^2(\delta/h)^2}}$.

Remark 5.10. From (Ye, 2013, Lemma 3.1.6) we know that $\|D_n\|_2 < \alpha(\frac{1}{2\delta} + \sqrt{\frac{\pi}{\beta}})$, which is independent from n but is also unbounded for $\delta \rightarrow 0$. If we adjust the estimates in the proof of the Lemma, we obtain

$$\|D_n\|_2 < \frac{h}{\delta} + 1.$$

Thus, $\|D_n\|_2$ is bounded for any value of h . In order to bound the norm of the coefficient matrix for $\delta \rightarrow 0$, we need to decrease h accordingly, which leads to an increased computational effort due to a higher dimensional system. Therefore, instead of solving a higher dimensional system, preconditioning might be a suitable alternative.

5.2 Preconditioning

Motivated by the numerical experiments in Sachs and Strauss (2008), the author theoretically shows in Ye (2013) the need for preconditioning the linear system that is obtained after the discretization of Equation (5.1). He provides estimates for the condition number of the unpreconditioned coefficient matrix T_n as well as the condition number of two differently preconditioned systems, $\hat{T}_n^{-1}T_n$ and $D_n^{-1}T_n$. One is preconditioned with Strang's preconditioner, the other with a tridiagonal matrix, where the latter acts only on the elliptic part of the coefficient matrix while the dense coefficients obtained by the integral part remains unpreconditioned. Numerical experiments show that using the tridiagonal preconditioner, the CG solver requires less iterations and less CPU time, due to less numerical effort.

From Chapter 4 we know that the Gaussian kernel approximates the Laplace operator for $\delta \rightarrow 0$. Therefore, we want to analyze the eigenvalue spectra with respect to δ of the unpreconditioned coefficient matrix, the Strang preconditioned coefficient matrix and the preconditioned coefficient matrix using two different tridiagonal preconditioner: one consists of the main and secondary diagonals of T_n , the other one is the discretized Laplace operator with coefficient $\frac{\delta^2}{2}$.

5.2.1 Circulant Preconditioners

By \widehat{T}_n we denote the preconditioner constructed in Strang (1986), whose entries on the diagonals are given by

$$s_k = \begin{cases} t_k, & 0 \leq k \leq \text{floor}(n/2), \\ t_{k-n}, & \text{floor}(n/2) < k < n, \\ s_{n+k}, & 0 < -k < n, \end{cases} \quad (5.22)$$

where $\text{floor}(x) = \max\{i \in \mathbb{Z}: i \leq x\}$. For $n = 2l + 1 (l \in \mathbb{N})$, we have

$$\widehat{T}_n = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{-l} & t_l & \cdots & t_1 \\ t_1 & t_0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & t_l \\ t_l & & \ddots & \ddots & \ddots & & t_{-l} \\ t_{-l} & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & t_0 & t_{-1} \\ t_{-1} & \cdots & t_{-l} & t_l & \cdots & t_1 & t_0 \end{pmatrix}.$$

This preconditioner is optimal in the following sense:

Theorem 5.11 (Chan, 1989). *Let T_n be a Hermitian Toeplitz matrix. The circulant matrix \widehat{T}_n whose entries are given by (5.22) minimizes $\|C_n - T_n\|_1 = \|C_n - T_n\|_\infty$ over all possible Hermitian circulant matrices C_n .*

Theorem 5.12 (Chan, 1989). *Let g be a positive function in the Wiener class, that means its Fourier coefficients are absolutely summable. Then for large n the circulant matrices \widehat{T}_n and \widehat{T}_n^{-1} are bounded in the l_2 -norm. In fact, for large n , the spectrum $\sigma(\widehat{T}_n)$ of \widehat{T}_n satisfies*

$$\sigma(\widehat{T}_n) \subset [g_{\min}, g_{\max}].$$

By construction, \widehat{T}_n is symmetric if T_n is. Furthermore, if T_n has a positive generating function, the theorem above yields that \widehat{T}_n is positive definite, which makes it suitable as a preconditioner.

Theorem 5.13 (Chan, 1989). *Let g be a positive function in the Wiener class and $\{T_n\}$ be the sequence of Toeplitz matrices generated by g . Then for all $\epsilon > 0$, there exists $N > 0$ such that for all $n > N$, at most $2N$ eigenvalues of $\widehat{T}_n - T_n$ have absolute values exceeding ϵ .*

Combining the last two theorems and using the identity

$$\widehat{T}_n^{-1}T_n - I_n = \widehat{T}_n^{-1}(T_n - \widehat{T}_n),$$

we receive the following result on the eigenvalue distribution of the preconditioned matrix $\widehat{T}_n^{-1}T_n - I_n$.

Corollary 5.14 (Chan, 1989). *Let g be a positive function in the Wiener class, then for all $\epsilon > 0$, there exists $N > 0$ such that for all $n > N$, at most $2N$ eigenvalues of $\widehat{T}_n^{-1}T_n - I_n$ have absolute values exceeding ϵ .*

In Sachs and Strauss (2008) various other circulant preconditioners have been considered, all of which are optimal in a different sense and cluster most eigenvalues of the preconditioned coefficient matrix around 1. Those preconditioners and the optimality proofs can be found in Chan (1988); Chan et al. (1991a,b); Chan and Jin (2007); Tismenetsky (1991); Tyrtshnikov (1992); Huckle (1992, 1993). We confine our analysis to Strang preconditioned systems, since it is almost free to construct and outperforms in numerical experiments with Merton's model, see Sachs and Strauss (2008).

5.2.2 Preconditioning of the Gaussian Kernel

We show corresponding results for the Strang preconditioned system with T_n given by Equation (5.6) adjusting the results in Ye (2013) and expressing the result with respect to δ .

Lemma 5.15. *If the Toeplitz matrix T_n is defined by (5.6) and \widehat{T}_n is the corresponding Strang preconditioner, then $\|T_n^{-1}\|_\infty < 1$ and $\|\widehat{T}_n^{-1}\|_\infty < 1$*

Proof. (cf. Ye, 2013, Lemma 3.1.21) It can easily be seen that

$$\xi := \min_k \{ |t_{kk}| - \sum_{j \neq k} |t_{kj}| \} > \frac{3}{2} + \lambda\tau(1 - \|D_n\|_\infty) = \frac{3}{2} - \lambda\tau \frac{h}{\delta}.$$

Thus, $\xi > 1$ if $\frac{1}{2} > \lambda\tau \frac{h}{\delta}$ which can be obtained for sufficiently small τ . With results from Varah (1975), we receive $\|T_n^{-1}\|_\infty < 1$, with similar calculations we obtain $\|\widehat{T}_n^{-1}\|_\infty < 1$. \square

Theorem 5.16 (Ye, 2013, Theorem 3.1.22). *Let τ be sufficiently small. For all $\epsilon > 0$, there exists $N(\epsilon) > 0$ such that for all $n \geq N(\epsilon)$, at most $2N(\epsilon)$ eigenvalues of $\widehat{T}_n^{-1}T_n$ be outside of the interval $(1 - \epsilon, 1 + \epsilon)$. Here $N(\epsilon)$ is given by*

$$N(\epsilon) = \begin{cases} 1, & \text{if } \frac{\lambda\tau}{2} < \epsilon, \\ \text{ceil} \left(\frac{\sqrt{2}\delta}{h} \text{erf}^{-1} \left(1 - \frac{2\epsilon}{\lambda\tau} \right) \right), & \text{otherwise,} \end{cases} \quad (5.23)$$

where $\text{ceil}(x) = \min\{i \in \mathbb{Z} : i \geq x\}$.

Remark 5.17. (i) The essential prerequisite for Theorem 5.16 is $\lambda\tau \frac{h}{\delta} < \frac{1}{2}$, which is satisfied if $\delta > 2\lambda\tau h$. Moreover, the prerequisite for Lemma 5.8 is $\delta \leq \frac{h}{\sqrt{2 \ln 2}}$. Hence,

$$2\lambda\tau h < \delta \leq \frac{h}{\sqrt{2 \ln 2}}.$$

This can only hold if $2\lambda\sqrt{2 \ln 2}\tau < 1$ or $m > 2\sqrt{2 \ln 2}\lambda T$, respectively. However, this is true for almost all numerical settings. On the other hand, if $\delta > h$, Lemma 5.8 is

not applicable and we have no estimate for $\sigma(T_n)$. We don't know if the matrix D_n is positive definite on its own. Nevertheless, with a sufficiently large m , the matrix T_n remains positive definite and the Strang preconditioned system still has most of its eigenvalues clustered around 1 (if $\lambda\tau < 2\epsilon$ all except two).

- (ii) A tridiagonal preconditioner is proposed in Ye (2013) and it outperforms Strang's preconditioner for a given parameter set that is used in the numerical experiments. By \bar{T}_n we denote the matrix consisting of the first and secondary diagonals of T_n :

$$\bar{T}_n = \begin{pmatrix} t_0 & t_{-1} & 0 & \cdots & 0 \\ t_1 & t_0 & t_{-1} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & t_1 & t_0 & t_{-1} \\ 0 & \cdots & 0 & t_1 & t_0 \end{pmatrix}.$$

In Section 5.4 we study the eigenvalue spectrum of the Strang preconditioned system and the system preconditioned by the tridiagonal matrix \bar{T} . In addition to that, we consider the discretized Laplace operator as a preconditioner due to the results in Section 4.1.

5.3 Alternative Discretization Scheme

In this section we consider the right-hand side of Equation (4.2) and discuss its numerical discretization. Since $\Gamma_\delta(z-x) = \Gamma_\delta(x-z)$, we are able to rewrite $u(x) - \Gamma_\delta * u(x) = 0$ as

$$u(x) - \int_{-\infty}^{\infty} u(z)\Gamma_\delta(z-x)dz = 0, \quad x \in \mathbb{R}, \quad (5.24)$$

in order to align with the notation in Equation (5.2).

For the numerical calculation of the integral

$$\int_{\Omega} u(z)\Gamma_\delta(z-x_i)dz,$$

we started with the composite trapezoidal rule and received

$$\frac{h}{2}(u(x_1)\Gamma_\delta(x_1-x_i) + 2\sum_{j=2}^{n+1} u(x_j)\Gamma_\delta(x_j-x_i) + u(x_{n+2})\Gamma_\delta(x_{n+2}-x_i)),$$

with $x_j = -R + (j-1)h$ for $j = 1, \dots, n+2$. By using the boundary condition, we eliminated the first and last equation of the resulting linear system and after re-indexing the subscripts, i.e., $x_j = -R + jh$ for $j = 1, \dots, n$, we obtained the composite midpoint rule,

$$h \sum_{j=1}^n u(x_j)\Gamma_\delta(x_j-x_i),$$

for the interval $I := (-R + h/2, R - h/2)$; see Figure 5.1 for an illustration.

However, for $\delta \rightarrow 0$ (and fixed n), the weight $h\Gamma_\delta(0)$ gets more and more inaccurate, since the value $\Gamma_\delta(0)$ rapidly increases and so does the descent at $x = 0$. Hence,

$$h\Gamma_\delta(0) > \frac{1}{\sqrt{2\pi}\delta} \int_{-h/2}^{h/2} e^{-z^2/(2\delta^2)} dz$$

and this error contributes most to the discretization error, see Figure 5.2.

An adaptive step size control could handle that problem, but we would lose the Toeplitz structure of the matrix T_n , while a dense matrix without a special structure would remain. Since the critical interval is around $x = 0$, we could adjust the step size to a smaller $h' \ll h$ only in the interval $(-2h, 2h)$. For $h' = \frac{1}{2}h$, this would yield

$$h \sum_{\substack{j < \lceil \frac{n}{2} \rceil - 2 \\ j > \lceil \frac{n}{2} \rceil + 2}} \Gamma_\delta(x_j) + h' \sum_{k=1}^7 \Gamma_\delta(x'_k) + \frac{h+h'}{2} (\Gamma_\delta(x_{\lceil \frac{n}{2} \rceil - 2}) + \Gamma_\delta(x_{\lceil \frac{n}{2} \rceil + 2}))$$

as an approximation for $\int_I \Gamma_\delta(z) dz$, with $x'_k = -2h + kh'$, $k = 1, \dots, 7$ (cf. Figure 5.3). Yet, the approximation of $\int_I u(z)\Gamma_\delta(z - x_i)$ would be

$$\begin{aligned} & h \sum_{\substack{j < \lceil \frac{n}{2} \rceil - 2 \\ j > \lceil \frac{n}{2} \rceil + 2}} u(x_j)\Gamma_\delta(x_j - x_i) + h' \sum_{k=1}^7 u(x'_k)\Gamma_\delta(x'_k - x_i) \\ & + \frac{h+h'}{2} (u(x_{\lceil \frac{n}{2} \rceil - 2})\Gamma_\delta(x_{\lceil \frac{n}{2} \rceil - 2} - x_i) + u(x_{\lceil \frac{n}{2} \rceil + 2})\Gamma_\delta(x_{\lceil \frac{n}{2} \rceil + 2} - x_i)). \end{aligned}$$

Hence, the structure of the integrand would lead to a finer discretization of u around $x = 0$, the function Γ_δ would be evaluated with the finer step size h' around x_i and this approach would only work for $x_i = 0$. And as mentioned above, the resulting matrix $T_{n,h'}$ would lose its Toeplitz structure.

Instead of using the composite midpoint rule and receiving the weights $h\Gamma_\delta(x_j)$, we use the error function to approximate the integral of Γ_δ on the interval $I_j = (x_j - h/2, x_j + h/2)$. For the critical point $x_j = 0$, that approach yields the weight

$$\frac{1}{\sqrt{2\pi}\delta} \int_{-h/2}^{h/2} e^{-\frac{x^2}{2\delta^2}} dx = \frac{2}{\sqrt{2\pi}\delta} \int_0^{h/2} e^{-\frac{x^2}{2\delta^2}} dx = \frac{2}{\sqrt{\pi}} \int_0^{\frac{h}{2\delta\sqrt{2}}} e^{-y^2} dy = \operatorname{erf}\left(\frac{h}{2\delta\sqrt{2}}\right).$$

For any other discretization point $x_j > 0$ follows

$$\begin{aligned} \frac{1}{\sqrt{2\pi}\delta} \int_{x_j-h/2}^{x_j+h/2} e^{-\frac{x^2}{2\delta^2}} dx &= \frac{1}{\sqrt{\pi}} \int_0^{\frac{x_j+h/2}{\delta\sqrt{2}}} e^{-y^2} dy - \frac{1}{\sqrt{\pi}} \int_0^{\frac{x_j-h/2}{\delta\sqrt{2}}} e^{-y^2} dy \\ &= \frac{1}{2} \left(\operatorname{erf}\left(\frac{x_j+h/2}{\delta\sqrt{2}}\right) - \operatorname{erf}\left(\frac{x_j-h/2}{\delta\sqrt{2}}\right) \right) \\ &= \frac{1}{2} \left(\operatorname{erf}\left(\frac{-2R+(2j+1)h}{2\delta\sqrt{2}}\right) - \operatorname{erf}\left(\frac{-2R+(2j-1)h}{2\delta\sqrt{2}}\right) \right). \end{aligned}$$

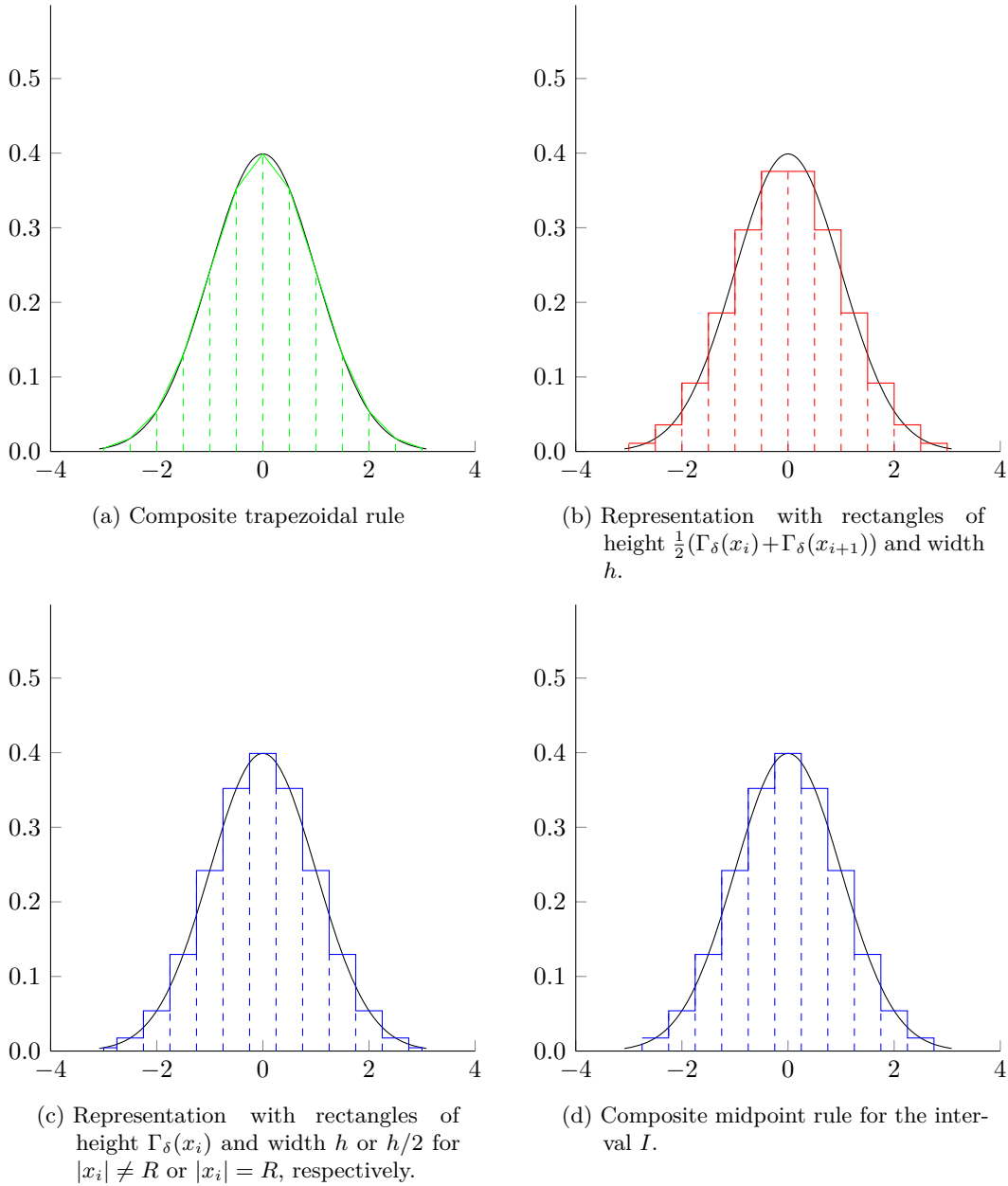


Figure 5.1.: Discretization of the Gaussian kernel using the composite trapezoidal rule on Ω , (a)–(c), and the composite midpoint rule on I , (d), with $R = 3$, $n = 11$, $h = 0.5$ and $\delta = 1$. The approximations shown in (a)–(c) are equivalent, but (c) illustrates the transition from the composite trapezoidal rule to the composite midpoint rule best. The first and last row of the resulting linear system corresponds to the first and last rectangle in (c). Their elimination then leads to the composite midpoint rule visualized in (d).

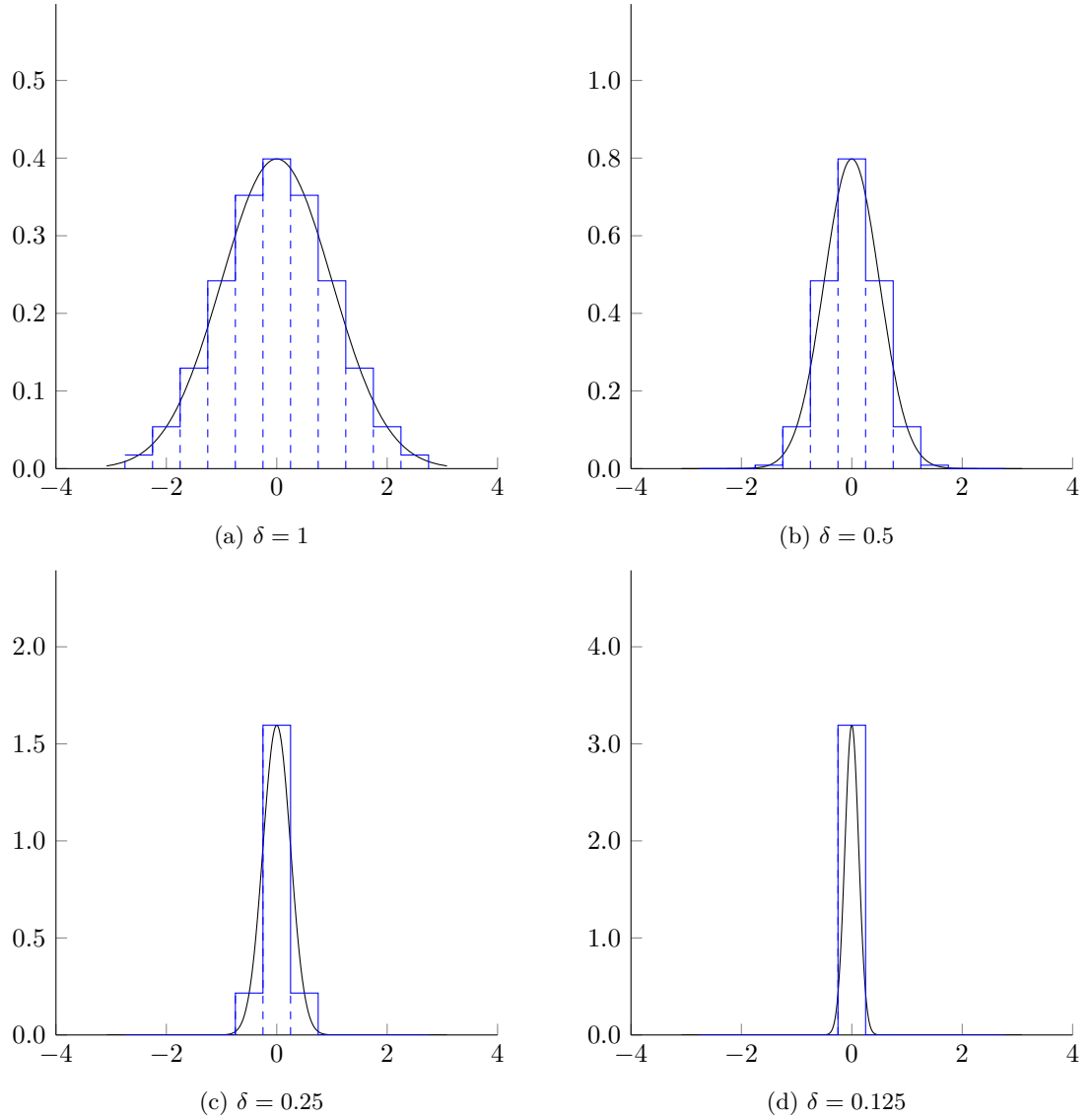


Figure 5.2.: Discretization of the Gaussian kernel using the composite midpoint rule on I with $R = 3$, $n = 11$, $h = 0.5$ and decreasing values of δ .

We receive the same weights for $x_j < 0$ and finally, we obtain the discretization scheme

$$\begin{aligned}
 \int_I u(z)\Gamma_\delta(z - x_i) dz &\approx u(x_i) \operatorname{erf}\left(\frac{h}{2\delta\sqrt{2}}\right) \\
 &+ \sum_{\substack{j=1 \\ j \neq i}}^n u(x_j) \frac{1}{2} \left(\operatorname{erf}\left(\frac{(2|j-i|+1)h}{2\delta\sqrt{2}}\right) - \operatorname{erf}\left(\frac{(2|j-i|-1)h}{2\delta\sqrt{2}}\right) \right).
 \end{aligned} \tag{5.25}$$

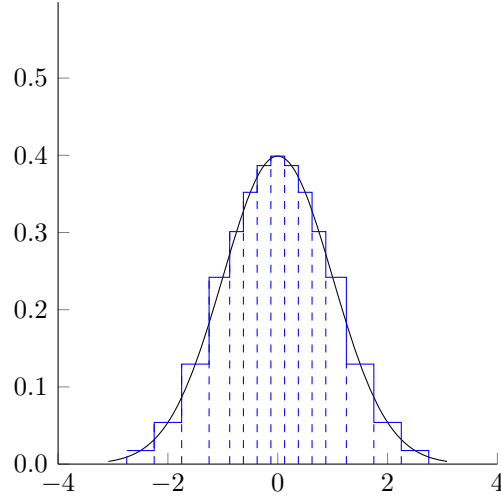


Figure 5.3.: Discretization of the Gaussian kernel using the composite midpoint rule on I using a finer step size $h' = h/2$ in $(-2h, 2h)$ with $R = 3$, $n = 11$, $h = 0.5$ and $\delta = 1$.

The change of the discretization scheme does not change the order of accuracy.

Lemma 5.18. *Both discretization schemes, the midpoint rule presented in Section 5.1 as well as the alternative discretization scheme given by Equation (5.25), result in second order accuracy. The errors E_M and E_{erf} differ only by constants*

$$C_M = \max_{x \in I} |(u(x)\Gamma_\delta(x - x_i))''|$$

and

$$C_{erf} = \max_{x \in I} \left| \frac{u''(x)}{\sqrt{2\pi\delta}} \right|.$$

Proof. We calculate the error E_M produced by the midpoint rule

$$E_M = \int_{-R+h/2}^{R-h/2} u(x)\Gamma_\delta(x - x_i) dx - h \sum_{j=1}^n u(x_j)\Gamma_\delta(x_j - x_i).$$

First, we calculate the discretization error for each subinterval I_j by using the Taylor expansion of $u(x)\Gamma_\delta(x - x_i)$ around x_j ,

$$\begin{aligned} & \int_{I_j} u(x)\Gamma_\delta(x - x_i) dx \\ &= \int_{I_j} u(x_j)\Gamma_\delta(x_j - x_i) + (x - x_j)(u(x_j)\Gamma_\delta(x_j - x_i))' + \frac{(x - x_j)^2}{2}(u(\xi)\Gamma_\delta(\xi - x_i))'' dx \\ &= hu(x_j)\Gamma_\delta(x_j - x_i) + \int_{I_j} \frac{(x - x_j)^2}{2}(u(\xi)\Gamma_\delta(\xi - x_i))'' dx, \end{aligned}$$

with $\xi \in I_j$ depending on x . Note that $\int_{I_j} (x - x_j) dx = 0$, due to the antisymmetry of $(x - x_j)$ in I_j . Thus, we have for the interval I_j the estimate

$$\begin{aligned} & \left\| \int_{I_j} u(x) \Gamma_\delta(x - x_i) dx - hu(x_j) \Gamma_\delta(x_j - x_i) \right\| \\ & \leq \frac{1}{2} \left\| \int_{I_j} (x - x_j)^2 (u(\xi) \Gamma_\delta(\xi - x_i))'' dx \right\| \\ & \leq \frac{C_M}{6} \left((x_j + \frac{h}{2} - x_j)^3 - (x_j - \frac{h}{2} - x_j)^3 \right) \\ & = \frac{C_M}{24} h^3, \end{aligned}$$

with $C_M = \max_{x \in I} |(u(x) \Gamma_\delta(x - x_i))''|$. The discretization error for the interval I is then

$$\begin{aligned} & \left\| \int_{-R+h/2}^{R-h/2} u(x) \Gamma_\delta(x - x_i) dx - \sum_{j=1}^n hu(x_j) \Gamma_\delta(x_j - x_i) \right\| \\ & \leq \left\| \sum_{j=1}^n \int_{x_j-h/2}^{x_j+h/2} u(x) \Gamma_\delta(x - x_i) dx - hu(x_j) \Gamma_\delta(x_j - x_i) \right\| \\ & = \sum_{j=1}^n \frac{C_M}{24} h^3 = \frac{C_M}{24} nh^3 \\ & = \frac{n}{n+1} \frac{2RC_M}{24} h^2 = \mathcal{O}(h^2). \end{aligned}$$

Now, we estimate the error E_{erf} of the alternative discretization scheme

$$E_{erf} = \int_{-R+h/2}^{R-h/2} u(x) \Gamma_\delta(x - x_i) dx - \sum_{j=1}^n u(x_j) \int_{x_j-h/2}^{x_j+h/2} \Gamma_\delta(x - x_i) dx.$$

Again, we start with the calculation of the discretization error for each subinterval I_j . This time we perform the Taylor expansion around x_j only on u ,

$$\begin{aligned} & \int_{I_j} u(x) \Gamma_\delta(x - x_i) dx \\ & = \int_{I_j} u(x_j) \Gamma_\delta(x - x_i) + (x - x_j) u'(x_j) \Gamma_\delta(x - x_i) + \frac{(x - x_j)^2}{2} u''(\xi) \Gamma_\delta(x - x_i) dx \\ & = u(x_j) \int_{I_j} \Gamma_\delta(x - x_i) dx + u'(x_j) \int_{I_j} (x - x_j) \Gamma_\delta(x - x_i) dx + \\ & \quad + \int_{I_j} \frac{(x - x_j)^2}{2} u''(\xi) \Gamma_\delta(x - x_i) dx, \end{aligned}$$

with $\xi \in I_j$ depending on x . For the interval I_j , we have the estimate

$$\begin{aligned}
 & \left\| \int_{I_j} u(x) \Gamma_\delta(x - x_i) dx - u(x_j) \int_{I_j} \Gamma_\delta(x - x_i) dx \right\| \\
 & \leq |u'(x_j) \Gamma_\delta(0) \int_{I_j} (x - x_j) dx| + \frac{1}{2} \left\| \int_{I_j} (x - x_j)^2 u''(\xi) \Gamma_\delta(x - x_i) dx \right\| \\
 & \leq \frac{1}{2} \left\| \int_{I_j} (x - x_j)^2 u''(\xi) \Gamma_\delta(0) dx \right\| \\
 & \leq \frac{C_{erf}}{6} \left((x_j + \frac{h}{2} - x_j)^3 - (x_j - \frac{h}{2} - x_j)^3 \right) \\
 & = \frac{C_{erf}}{24} h^3,
 \end{aligned}$$

with $C_{erf} = \max_{x \in I} |u''(x)| \Gamma_\delta(0) = \max_{x \in I} \frac{|u''(x)|}{\sqrt{2\pi\delta}}$. Thus, the discretization error for the interval I is

$$\begin{aligned}
 & \left\| \int_{-R+h/2}^{R-h/2} u(x) \Gamma_\delta(x - x_i) dx - \sum_{j=1}^n u(x_j) \int_{x_j-h/2}^{x_j+h/2} \Gamma_\delta(x - x_i) dx \right\| \\
 & = \left\| \sum_{j=1}^n \int_{x_j-h/2}^{x_j+h/2} u(x) \Gamma_\delta(x - x_i) dx - \sum_{j=1}^n u(x_j) \int_{x_j-h/2}^{x_j+h/2} \Gamma_\delta(x - x_i) dx \right\| \\
 & \leq \sum_{j=1}^n \frac{C_{erf}}{24} h^3 = \frac{C_{erf}}{24} n h^3 \\
 & = \frac{n}{n+1} \frac{2RC_{erf}}{24} h^2 = \mathcal{O}(h^2).
 \end{aligned}$$

We conclude that both discretization schemes result in second order accuracy. \square

The alternative discretization scheme yields the Toeplitz matrix T_n^{erf} with coefficients

$$\begin{aligned}
 t_0^{erf} &= 1 - \operatorname{erf}\left(\frac{h}{2\delta\sqrt{2}}\right), \\
 t_j^{erf} &= -\frac{1}{2} \left(\operatorname{erf}\left(\frac{(2j+1)h}{2\delta\sqrt{2}}\right) - \operatorname{erf}\left(\frac{(2j-1)h}{2\delta\sqrt{2}}\right) \right), \\
 t_{-j}^{erf} &= t_j^{erf},
 \end{aligned} \tag{5.26}$$

with $j = 1, \dots, n-1$. The generating function $g_{erf}^{(n)}$ of the matrix T_n^{erf} is given by

$$g_{erf}^{(n)}(x) = 1 - \left(\operatorname{erf}\left(\frac{h}{2\delta\sqrt{2}}\right) + \sum_{j=1}^{n-1} \left(\operatorname{erf}\left(\frac{(2j+1)h}{2\delta\sqrt{2}}\right) - \operatorname{erf}\left(\frac{(2j-1)h}{2\delta\sqrt{2}}\right) \right) \cos(jx) \right).$$

Tian and Du propose similar discretization schemes in Tian and Du (2013) for nonlocal diffusion equations. Those equations involve symmetric nonlocal kernels $\gamma_\delta(x, y): \Omega \times \Omega \rightarrow \mathbb{R}$ with compact support, i.e., $\gamma_\delta(x, y) = \gamma_\delta(y, x)$ and $\gamma_\delta(x, y) = 0$ if $y \notin B_\delta(x)$. They show the convergence of the discrete schemes to the nonlocal problem as $h \rightarrow 0$ with δ fixed and the convergence of the discrete schemes to the corresponding local problem as h and δ vanish.

5.4 Numerical Results

The results of our numerical experiments support the analytic results from Part II. We show the eigenvalue distribution of D_n from Equation (5.8) for decreasing values of δ in comparison to the eigenvalue distribution of the discretized Laplace operator. Then we plot the smallest and largest eigenvalue with the minimum and maximum of the generating function as a numerical illustration of Lemma 5.8. We then present the spectral properties of the preconditioned systems. We conclude this section with numerical experiments on the alternative discretization scheme (5.25). To be able to compare this scheme with the midpoint rule scheme, we apply it to Merton's model, since a semi-closed form solution exists.

Analysis of the Spectrum

For the discretization of the Laplace operator we use the second order central difference scheme,

$$v_{xx}(x_i) \approx \frac{1}{h^2}(v_{i+1} - 2v_i + v_{i-1}),$$

and obtain the system $\Delta_n v = 0$ as the discretization of $-\frac{\delta^2}{2}v_{xx} = 0$. The matrix Δ_n is a symmetric tridiagonal matrix with constant diagonal $\frac{\delta^2}{h^2}$ and constant subdiagonals $-\frac{\delta^2}{2h^2}$. It is well known, that the eigenvalues λ_k of the Laplace operator with Dirichlet boundary conditions are given by $\lambda_k = k^2\pi^2$ and the spectrum $\sigma(-\Delta) \subset [\pi^2, \infty)$ is unbounded. For Δ_n we have $\lambda_k = \frac{2\delta^2}{h^2} \sin^2(\frac{k\pi h}{2})$ and $\sigma(-\Delta_n) \subset [0, \frac{\delta^2}{R^2}(n+1)^2]$.

The distribution of the eigenvalues of $I - D_n$ along with its generating function $1 - g_D^{(n)}$ is shown in Figure 5.4 for decreasing values of δ next to the eigenvalue distribution and the generating function of $-\frac{\delta^2}{2}\Delta_n$. For larger values of δ the eigenvalues of $I - D_n$ cluster around 1, e.g., for $\delta = 1$ there are only three exceptions, whereas the eigenvalue distribution approaches the eigenvalue distribution of $-\frac{\delta^2}{2}\Delta_n$ for decreasing values of δ . Moreover, the plots illustrate that Equation (5.19) is a good estimate for the given parameter set even for $\delta > \frac{h}{\sqrt{2\ln 2}}$.

Figure 5.5 shows the improvement of estimate (5.19) over (5.14) for small values of δ , but also shows the advancement of (5.14) for larger values of δ . As a takeaway from this experiment, we consider the estimate

$$\sigma(T_n) \subset \frac{3}{2} + \lambda\tau \left[\max \left\{ -c, 1 - \frac{h}{\sqrt{2\pi}\delta} - \operatorname{erf} \left(\frac{(n-1)h}{\sqrt{2}\delta} \right) \right\}, 1 \right], \quad (5.27)$$

which is independent of δ . Although Equation (5.19) is only valid if $\delta \leq \frac{h}{\sqrt{2\ln 2}}$, Table 5.1 indicates a negligible deviation if $\delta > \frac{h}{\sqrt{2\ln 2}}$. The table shows the corresponding values for D_n since the values for T_n are scaled by $\tau = T/m$ and therefore less meaningful.

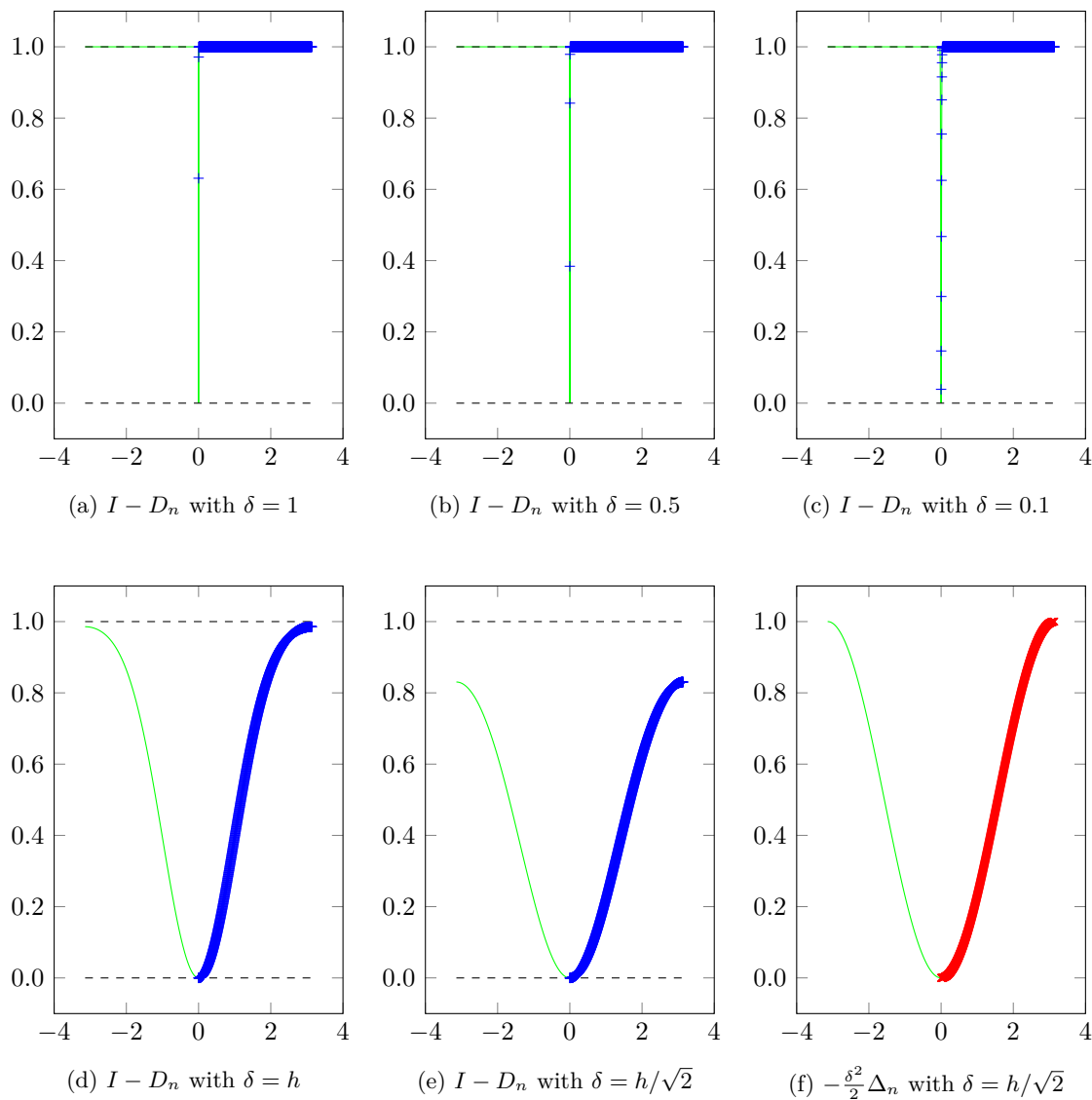


Figure 5.4.: Generating function (green) and eigenvalues of $I - D_n$ (blue), (a)–(e), with different values of δ and generating function (green) and eigenvalues of $-\frac{\delta^2}{2}\Delta_n$ (red), (f), with $\delta = h/\sqrt{2}$; $n = 1000$, $m = 10$ and $T = 1$ in all six plots. The black dashed lines mark the interval $[-c, 1]$ with c from Lemma 5.8.

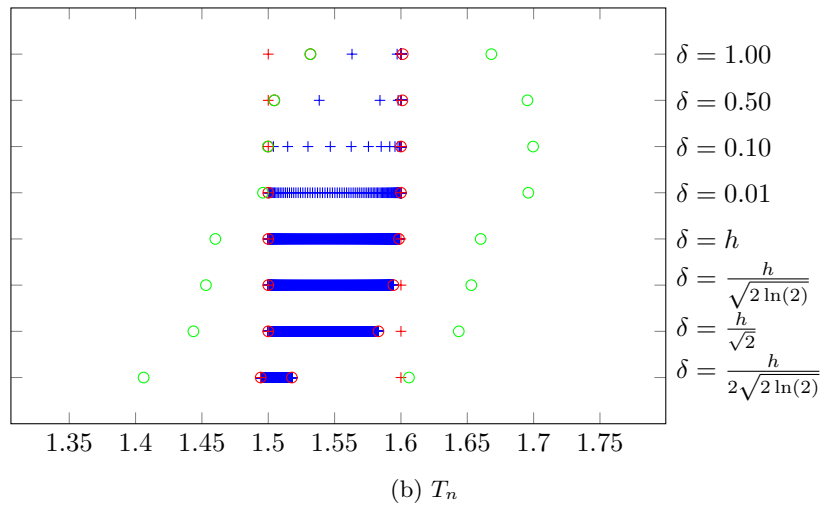
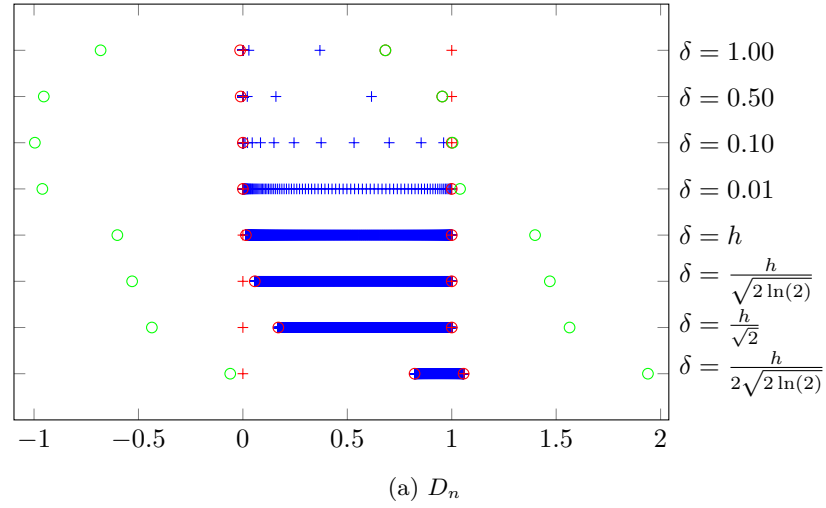


Figure 5.5.: Eigenvalue spectrum of D_n , (a), and T_n , (b), (blue cross), as well as the minimum and maximum of the corresponding generating functions $g_D^{(n)}$ and $g^{(n)}$ (red circle), the estimates from Lemma 5.8 and Corollary 5.9, (red cross), and the estimates in Equation (5.14) (green circle); with $n = 1000$, $m = 10$ and $T = 1$ in both plots.

δ	$g_{D,min}^{(n)}$	0	$\min g_D^{(n)}(x)$	$\min g_D(x)$	$\text{eig}(D_n)_1$	$\text{eig}(D_n)_n$	$\max g_D(x)$	$\max g_D^{(n)}(x)$	$1 + c$	$g_{D,max}^{(n)}$
1.000	-0.68132	0	-1.317e-02	-9.889e-15	-1.152e-16	0.36886	1.00000	0.68245	1.00000	0.68212
0.500	-0.95327	0	-1.053e-02	-4.542e-15	-2.006e-16	0.61586	1.00000	0.95439	1.00000	0.95486
0.100	-0.99601	0	-1.820e-15	-1.823e-15	-2.280e-16	0.96113	1.00000	1.00000	1.00000	1.00399
0.010	-0.96015	0	-3.192e-16	-3.192e-16	-3.338e-16	0.99952	1.00000	1.00000	1.00000	1.03985
h	-0.60106	0	0.01438	0.01438	0.01438	1.00000	1.00000	1.00000	1.00000	1.39894
$\frac{h}{\sqrt{2}}$	-0.53028	0	0.05689	0.05689	0.05690	1.00000	1.00000	1.00000	1.00000	1.46972
$\frac{h}{\sqrt{2 \ln 2}}$	-0.43581	0	0.16961	0.16961	0.16961	1.00010	1.00010	1.00010	1.00010	1.56419
$\frac{h}{2\sqrt{2 \ln 2}}$	-0.06056	0	0.82204	0.82204	0.82204	1.05690	1.05690	1.05690	1.05856	1.93944

Table 5.1.: Smallest and largest eigenvalue of D_n ($\text{eig}(D_n)_1$ and $\text{eig}(D_n)_n$), minimum and maximum of the corresponding generating function ($\min g_D^{(n)}(x)$ and $\max g_D^{(n)}(x)$) and of a modified generating function with infinite sum ($\min g_D(x)$ and $\max g_D(x)$) as well as upper and lower bounds of the eigenvalue spectrum estimates given by Equation (5.19) (0 and $1 + c$ with c from Lemma 5.8) and by Equation (5.14) adjusted to D_n ($g_{D,min}^{(n)}$ and $g_{D,max}^{(n)}$), each for different values of δ and with $n = 1000$, $m = 10$ and $T = 1$. Equation (5.19) seems to hold for $\delta \in (\frac{h}{\sqrt{2 \ln 2}}, h]$, although we have no theoretical proof. For $\delta > h$ the smallest eigenvalue of D_n is only slightly below 0. For $\delta > \frac{1}{2}$ the table shows that $g_{D,max}^{(n)}$ is more accurate than $1 + c$ as an estimate for the upper bound of $\sigma(D_n)$.

Preconditioning

In numerical experiments Ye shows that a tridiagonal preconditioner that only acts on the differential terms of the Merton model outperforms Strang's preconditioner with respect to the iteration number of the preconditioned CG method. The previous results show, that the spectrum of T_n is evenly distributed for decreasing values of δ . Consequently we assume that preconditioning is crucial for small values of δ .

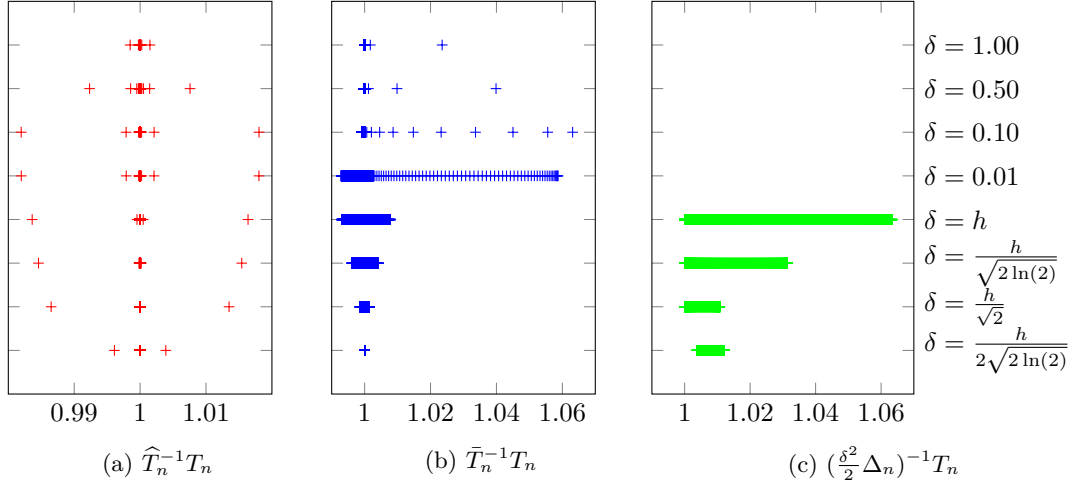


Figure 5.6.: Eigenvalue spectrum of preconditioned system with three different preconditioner for different values of δ ; $n = 1000$, $m = 10$ and $T = 1$.

The eigenvalue spectra of the preconditioned systems $\hat{T}_n^{-1}T_n$, $\bar{T}_n^{-1}T_n$ and $(\frac{\delta^2}{2}\Delta_n)^{-1}T_n$ are plotted in Figure 5.6. We see the expected clustering around 1 except few outliers in the spectrum of the Strang preconditioned systems. For $\delta < h$ the tridiagonal preconditioner also clusters the eigenvalues around 1, however, the eigenvalues remain distributed. If we use $\frac{\delta^2}{2}\Delta_n$ as a preconditioner, the eigenvalue spectrum gets worse compared to that of T_n for $\delta > \frac{h}{\sqrt{2\ln 2}}$ and we omit the corresponding plots, but for $\delta < \frac{h}{\sqrt{2\ln 2}}$ the eigenvalues approaches 1 from the right. Nonetheless, the eigenvalue spectrum has the largest distance between the smallest and largest eigenvalue compared to Strang's preconditioner and the tridiagonal preconditioner.

Alternative Discretization Scheme

In our last numeric experiment we compare the midpoint rule for the numerical integration from Section 5.1 with the alternative scheme given by Equation (5.25). We apply both schemes to the Merton model (5.1) using the same discretization for the time and spatial derivatives. We use the Merton model, because a semi-closed form solution is available:

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda\eta\tau} (\lambda\eta\tau)^k}{k} V_{BS}(S, \sigma_k, r_k, \tau, K), \quad (5.28)$$

with the call option's Black-Scholes price V_{BS} (cf. Black and Scholes, 1973), price S , strike price K , time to maturity $\tau = T - t$, $\sigma_k = \sqrt{\sigma^2 + k\delta^2/\tau}$, $r_k = r - \lambda(\eta - 1) + k \ln \eta/\tau$ and $\eta = e^{\frac{\delta^2}{2}}$. The derivation of pricing formula (5.28) can be found in (Joshi, 2003, Chapter 15.3). The differences between the numerical solutions and the semi-closed form solution are shown in Figure 5.7.

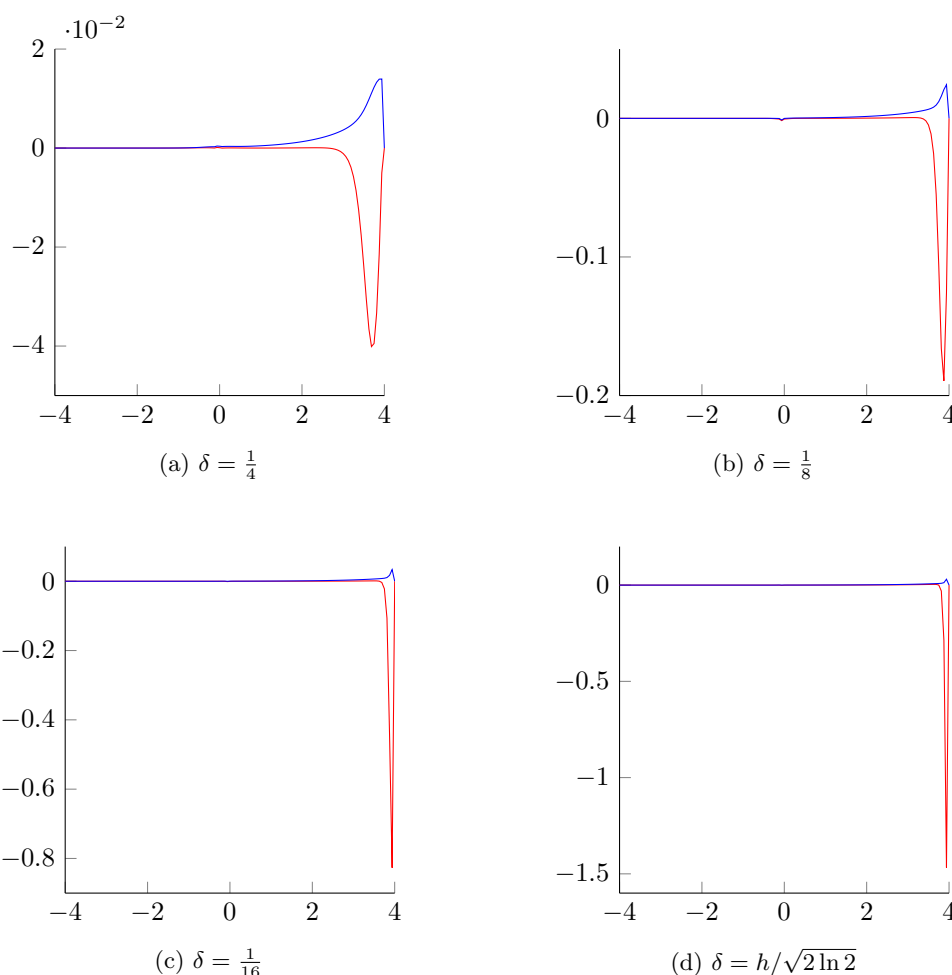


Figure 5.7.: Difference between discretization scheme from Section 5.3 and pricing formula (5.28) (blue) and difference between discretization scheme from Section 5.1 and pricing formula (5.28) (red) for different values of δ with parameters $n = 128$, $m = 10$, $T = 1$, $K = 1$, $r = 0.07$, $\sigma = 0.01$, $\lambda = 1$; the infinite sum in Equation (5.28) has been cut off at $k = 100$. The alternative discretization scheme seems to be more accurate for decreasing values of δ . The solutions obtained with the standard discretization scheme tend to be smaller than the semi-closed form solution whereas the solutions received by the alternative discretization scheme tend to be larger.

Conclusion

The aim of Part I was to derive necessary optimality conditions for the control of a class of semilinear partial integro-differential equations. Therefore, we introduced a PIDE system as it appears in an application in biology. We considered a PIDE-constrained optimization problem, where the objective function is a tracking type functional and the controls are certain time-dependent adhesion control functions. In the sequel, we first derived necessary optimality conditions for semilinear evolution equations in Banach spaces and extended our work that has been published in Frerick et al. (2015). The key result of Part I is Theorem 3.9. From that we deduced the optimality conditions for an optimal point of the PIDE system which include adjoint differential equations. These are also of a partial integro-differential equation type which run backwards in time with final conditions coming from the tracking type of the objective function. While we proved the existence and uniqueness of a solution of the PIDE system, the existence of optimal control functions is left for future research. Possible approaches are provided in Tröltzsch (2009) and Hinze et al. (2009).

The major result of Part II is Lemma 5.8, which provides a sharp estimate of the eigenvalue spectrum of the convolution with the Gaussian kernel. Numerical results on the eigenvalue spectra of different preconditioned systems correspond to the iteration number of the preconditioned conjugate gradient method shown in Table 5.2.

δ	iter T_n	iter $\hat{T}_n^{-1}T_n$	iter $(\frac{\delta^2}{2}\Delta_n)^{-1}T_n$
0.5	15	11	277
0.25	26	13	214
0.05	102	13	66

Table 5.2.: Iterations for solving Equation (5.5) using the cg-Method and the preconditioned cg-Method with Strang’s preconditioner and the discretized Laplace operator as preconditioner for different values of δ ; $n = 500$, $m = 10$ and $T = 1$.

Further, we proposed an alternative discretization scheme for the convolution with the Gaussian kernel to overcome numerical errors in case of small values of δ . The numerical results are encouraging. A thoroughful analysis of that discretization scheme is of great interest and in the direction of future research.

Bibliography

- D. R. Adams and L. I. Hedberg (1996). *Function Spaces and Potential Theory*. Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg.
- H. Amann (1986). Parabolic evolution equations with nonlinear boundary conditions. In F. E. Browder (editor) *Nonlinear functional analysis and its applications - Part 1*, volume 45, 17–27, American Mathematical Soc.
- H. Amann (1988). Parabolic evolution equations and nonlinear boundary conditions. *Journal of Differential Equations*, 72(2):201–269.
- L. Andersen and J. Andreasen (2000). Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Option Pricing. *Review of Derivatives Research*, 4(3):231–262.
- A. R. A. Anderson, M. A. J. Chaplain, E. L. Newman, R. J. C. Steele, and A. M. Thompson (2000). Mathematical Modelling of Tumour Invasion and Metastasis. *Journal of Theoretical Medicine*, 2(2):129–154.
- N. Armstrong, K. Painter, and J. Sherratt (2006). A continuum approach to modelling cell–cell adhesion. *Journal of theoretical biology*, 243(1):98–113.
- M. Becker, P. L. Butzer, and R. J. Nessel (1976). Saturation for Favard operators in weighted function spaces. *Studia Math.*, 59:139–153.
- F. Black and M. Scholes (1973). Black and Scholes, The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654.
- M. Briani, C. La Chioma, and R. Natalini (2004). Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. *Numerische Mathematik*, 98(4):607–646.
- M. Briani, R. Natalini, and G. Russo (2007). Implicit-explicit numerical schemes for jump-diffusion processes. *Calcolo*, 44(1):33–57.
- G. Brown and E. Hewitt (1984). A Class of Positive Trigonometric Sums. *Mathematische Annalen*, 268:91–122.

- N. Burch and R. Lehoucq (2011). Classical, Nonlocal, and Fractional Diffusion Equations on Bounded Domains. *International Journal for Multiscale Computational Engineering*, 9(6):661–674.
- P. L. Butzer and R. J. Nessel (1971). *Fourier analysis and approximation*. Birkhäuser, Basel.
- R. Chan (1989). Circulant Preconditioners for {Hermitian} {Toeplitz} Systems. *SIAM J. Matrix Anal. Appl.*, 10:542–550.
- R. H. Chan and X.-Q. Jin (2007). *An Introduction to Iterative Toeplitz Systems*. SIAM, Philadelphia.
- R. H. Chan, X.-Q. Jin, and M.-C. Yeung (1991a). The circulant operator in the banach algebra of matrices. *Linear Algebra and its Applications*, 149:41–53.
- R. H. Chan, X.-Q. Jin, and M.-C. Yeung (1991b). The Spectra of Super-Optimal Circulant Preconditioned Toeplitz Systems.
- T. F. Chan (1988). An Optimal Circulant Preconditioner for Toeplitz Systems.
- R. Cont and E. Voltchkova (2005). A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM Journal on Numerical Analysis*, 43(4):1596–1626.
- M. Dobrowolski (2010). *Angewandte Funktionalanalysis*. Springer-Verlag Berlin Heidelberg, 2nd edition.
- Q. Du, M. Gunzburger, R. Lehoucq, and K. Zhou (2013a). *A Nonlocal Vector Calculus, Nonlocal Volume-Constrained Problems, and Nonlocal Balance Laws*, volume 23.
- Q. Du, M. Gunzburger, R. Lehoucq, and K. Zhou (2013b). Analysis of the volume-constrained peridynamic navier equation of linear elasticity. *Journal of Elasticity*, 113(2):193–217.
- Q. Du and K. Zhou (2011). Mathematical analysis for the peridynamic nonlocal continuum theory. *Esaim-Mathematical Modelling and Numerical Analysis-Modelisation Mathématique Et Analyse Numérique*, 45(2):217–234.
- E. Emmrich, R. Lehoucq, and D. Puhst (2013). Peridynamics: A nonlocal continuum theory. *Lecture Notes in Computational Science and Engineering*, 89 LNCSE(1):45–65.
- K.-J. Engel and R. Nagel (2001). One-parameter semigroups for linear evolution equations. *Semigroup Forum*, 63(2):278–280.
- G. B. Folland (1992). *Fourier Analysis and Its Applications*. Pure and applied undergraduate texts, American Mathematical Society.

-
- L. Frerick, E. W. Sachs, and L. A. Zimmer (2015). Necessary optimality conditions for the control of partial integro-differential equations. In *Springer Proceedings in Mathematics and Statistics*, volume 134, 77–94.
- W. Gawronski and U. Stadtmüller (1982). Approximation of continuous functions by generalized Favard operators. *Journal of Approximation Theory*, 34(4):384–396.
- A. Gerisch (2010). On the approximation and efficient evaluation of integral terms in PDE models of cell adhesion. *IMA Journal of Numerical Analysis*, 30(1):173–194.
- U. Grenander and G. Szegő (1984). *Toeplitz Forms and Their Applications*. Chelsea Publishing, New York, 2nd edition.
- M. Gunzburger and R. Lehoucq (2010). A nonlocal vector calculus with application to nonlocal boundary value problems. *Multiscale Modeling & Simulation*, 8(5):1581–1598.
- M. Haase (2006). *The Functional Calculus for Sectorial Operators*. Birkhäuser, Basel.
- D. Henry (1981). *Geometric theory of semilinear parabolic equations, Lecture Notes in Mathematics, Vol. 840*. Springer, Heidelberg, 61st edition.
- M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich (2009). *Optimization with PDE Constraints*. Mathematical modelling–theory and applications, Springer.
- T. Huckle (1992). Circulant and Skewcirculant Matrices for Solving Toeplitz Matrix Problems. *SIAM Journal on Matrix Analysis and Applications*, 13(3):767–777.
- T. Huckle (1993). Some Aspects of Circulant Preconditioners. *SIAM J. Sci. Comput.*, 14(3):531–541.
- M. S. Joshi (2003). *The Concepts and Practice of Mathematical Finance*. Mathematics, Finance and Risk, Cambridge University Press.
- W. Kaballo (2011). *Grundkurs Funktionalanalysis*. Spektrum Akademischer Verlag.
- R. Lehoucq, Z. Huang, and Q. Du (2014). Nonlocal convection-diffusion volume-constrained problems and jump processes. *Discrete and Continuous Dynamical Systems - Series B*, 19(4):961–977.
- A.-M. Matache, T. von Petersdorff, and C. Schwab (2004). Fast deterministic pricing of options on Lévy driven assets. *ESAIM: Mathematical Modelling and Numerical Analysis*, 38:37–71.
- R. Meise and D. Vogt (1997). *Introduction to functional analysis*. Oxford graduate texts in mathematics, Clarendon Press.
- R. C. Merton (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144.

- A. Mogilner and L. Edelstein-Keshet (1999). A non-local model for a swarm. *Journal of Mathematical Biology*, 38(6):534–570.
- A. Pazy (1983). Semigroups of linear operators and applications to partial differential equations. *Springer-Verlag, New York. Applied Mathematical Sciences*.
- S. M. Rankin (1993). Semilinear Evolution Equations in Banach Spaces with Application to Parabolic Partial Differential Equations. *Transactions of the American Mathematical Society*, 336(2):523–535.
- W. Rudin (1976). *Principles of Mathematical Analysis*. International series in pure and applied mathematics, McGraw-Hill.
- E. Sachs and A. K. Strauss (2008). Efficient solution of a partial integro-differential equation in finance. *Applied Numerical Mathematics*, 58(11):1687–1703.
- M. Schu (2012). *Adaptive Trust-Region POD Methods and their Application in Finance*. Ph.D. thesis, Universität Trier.
- G. Strang (1986). A Proposal for Toeplitz Matrix Calculations. *Stud. Appl. Math.*, 74(2):171–176.
- X. Tian and Q. Du (2013). Analysis and Comparison of Different Approximations to Nonlocal Diffusion and Linear Peridynamic Equations. *SIAM Journal on Numerical Analysis*, 51(6):3458–3482.
- M. Tismenetsky (1991). A decomposition of Toeplitz matrices and optimal circulant preconditioning.
- F. Tröltzsch (1984). *Optimality conditions for parabolic control problems and applications*. Teubner-Texte zur Mathematik, B.G. Teubner.
- F. Tröltzsch (1989). On the Semigroup Approach for the Optimal Control of Semilinear Parabolic Equations Including Distributed and Boundary Control. *Zeitschrift für Analysis und ihre Anwendungen*, 8(5):431–443.
- F. Tröltzsch (2009). *Optimale Steuerung partieller Differentialgleichungen: Theorie, Verfahren und Anwendungen*. Vieweg + Teubner.
- E. E. Tyrtyshnikov (1992). Optimal and Superoptimal Circulant Preconditioners. *SIAM J. Matrix Anal. Appl.*, 13(2):459–473.
- J. Varah (1975). A lower bound for the smallest singular value of a matrix. *Linear Algebra and its Applications*, 11(1):3–5.
- D. Werner (2007). *Funktionalanalysis*. Springer-Lehrbuch, Springer.
- X. Ye (2013). *Some Preconditioners in Optimization with Partial Differential Equations*. Ph.D. thesis, Universität Trier.

- K. Zhou and Q. Du (2010). Mathematical and Numerical Analysis of Linear Peridynamic Models with Nonlocal Boundary Conditions. *SIAM Journal on Numerical Analysis*, 48(5):1759–1780.
- J. Zowe and S. Kurcyusz (1979). Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics & Optimization*, 5(1):49–62.