Numerieke technieken voor partiële differentiaalvergelijkingen in supergeleiding en thermo-elasticiteit

Numerical Techniques for Partial Differential Equations in Superconductivity and Thermoelasticity

Karel Van Bockstal

Promotor: prof. dr. M. Slodička
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Wiskundige Ingenieurstechnieken

Vakgroep Wiskundige Analyse
Voorzitter: prof. dr. M. Slodička
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2015 - 2016

UNIVERSITEIT GENT

ISBN 978-90-8578-867-6
NUR 919
Wettelijk depot: D/2015/10.500/111

Ghent University
Faculty of Engineering and Architecture
Department of Mathematical Analysis

Supervisor: prof. dr. Marián Slodička

Ghent University
Faculty of Engineering and Architecture

Department of Mathematical Analysis (NaM^2)
Galglaan 2 - S22, B-9000 Ghent, Belgium

http://www.nam2.ugent.be

Members of the examination board:

prof. dr. ir. Rik Van de Walle (chairman)	Ghent University, Belgium
prof. dr. ir. Joris Degrieck (secretary)	Ghent University, Belgium
prof. dr. Marián Slodička (supervisor)	Ghent University, Belgium
prof. dr. Denis Constales	Ghent University, Belgium
prof. dr. ir. Guillaume Crevecoeur	Ghent University, Belgium
prof. dr. Christophe Geuzaine	University of Liège, Belgium
prof. dr. Tomas Johansson	Linköping University, Sweden
prof. RNDr. Jozef Kačur, DrSc.	Comenius University, Slovakia

Dissertation to obtain the degree of
Doctor of Mathematical Engineering
Academic year 2015-2016

Voorwoord

In jouw handen bevindt zich mijn doctoraatsproefschrift dat ik heb geschreven in een boeiende en verrijkende periode van mijn leven. Het is het resultaat van ongeveer vier jaar onderzoek dat ik nu met ‘de wereld’ kan delen. Misschien spreek ik beter over ‘eilanden in de onderzoekswereld’ dan over ‘de wereld’, want sommige onder jullie zullen vermoedelijk na dit voorwoord niet verder lezen omdat wiskunde toch wel een vak apart is.

Sinds ik onder de deskundige leiding van Marián Slodička in september 2010 als onderzoeker begon aan de vakgroep Wiskundige Analyse van de Faculteit Ingenieurswetenschappen en Architectuur van de Universiteit Gent ben ik veel gegroeid (figuurlijk, niet door het vele resto-eten). Ik heb veel bijgeleerd, zowel over mezelf als over numerieke wiskunde. Na twee jaar onderzoek als doctoraatsbursaal werd ik in september 2012 assistent. Ik geloof dat ieder beginnend onderzoeker baat kan hebben van een afwisseling tussen lesgeven en onderzoek. Enkel tijdens de schrijffase kwam dit niet altijd even goed uit. Belangrijk is dat ik mijn onderzoek en mijn studies als een hobby heb ervaren. Voor mij is dat de essentie en ik ben er dan ook voluit voor gegaan. Ik ben dan ook trots op het resultaat.

De zin van het schrijven van een doctoraatsproefschrift kan gerust in vraag gesteld worden. Het is een tijdrovende bezigheid en het aantal lezers van een doctoraatsproefschrift is heel beperkt. Ook over de invulling ervan valt te discussiëren. De ene persoon wil maximum 150 bladzijden lezen, iemand anders wil enkel achter elkaar gekleefde artikels en nog een andere wil een stevig boek met plaats voor reflectie en verbanden. Ik heb geprobeerd om hierin een tussenweg te vinden. Tijdens mijn doctoraat heb ik veel gehad aan de raadgevingen van mijn promotor, maar het aantal bladzijden heb ik niet kunnen en willen beperken.

Mijn proefschrift start met een lange inleiding op de Rothemethode, een methode die aan onze onderzoeksgroep gebruikt wordt om partiële differentiaalvergelijkingen op te lossen. Dergelijke vergelijkingen worden gebruikt om fysische processen te modelleren. Bijvoorbeeld, de warmtevergelijking is een elementaire partiële differentiaalvergelijking die de variatie van de temperatuur in een gegeven gebied in de tijd beschrijft. Wie de ideeën uit het inleidende hoofdstuk begrijpt, moet in staat zijn om ook het vervolg van mijn proefschrift te begrijpen en te doorgronden.

Daarin bestudeer ik, in twee gescheiden delen, wiskundige problemen die voorkomen in respectievelijk thermo-elasticiteit en supergeleiding. Toch is de gemaakte analyse in beide delen verbonden met elkaar. Als u het eerste deel niet snapt, hoeft u dan ook niet verder te lezen. Het resultaat is een document van ongeveer 300 bladzijden met heel mooie wiskundige technieken om de verschillende problemen op te lossen. Ik heb het hier bewust over 'wiskundige technieken', omdat de inhoud van mijn proefschrift (nog) niet fysisch gevalideerd of getest is. Er is nog veel onderzoek mogelijk in de richting van de fysische implementatie van de verschillende modellen. Dit was echter niet het doel van mijn onderzoek, wat ik achteraf gezien wel jammer vind.

In dit voorwoord wil ik het ook even hebben over de universiteitsomgeving: een uitdagende omgeving, maar ook een bijwijlen harde wereld met competitie tussen verschillende deelgebieden die de wetenschap niet altijd vooruit helpt. Het was voor mij aanvankelijk niet gemakkelijk om in deze wereld mijn weg te vinden. Net zoals tijdens een voetbalmatch kan je veel speelplezier hebben, maar kan je ook wel eens een tackle langs achter of een elleboogstoot verwachten.

Tenslotte is dit voorwoord ook de plaats om enkele mensen te bedanken. Mijn grote inspanningen werden beloond, mede door de steun van en de ontspanning met de juiste personen op de juiste momenten. Bij deze:

> Aan allen die iets hebben bijgedragen: bedankt! Aan allen die niets bijdroegen: ... ook bedankt!

Karel Van Bockstal

Inspirational quotes:

> An expert is a person who has found out by his own painful experience all the mistakes that one can make in a very narrow field.
>
> Niels Bohr

> The life is nonlinear but it can be sometimes simplified.
>
> Marián Slodička

> Crashing is part of cycling as crying is part of love.
>
> Johan Museeuw

Table of Contents

List of Figures

List of Tables

Nederlandse samenvatting –Summary in Dutch–

Dit proefschrift onderzoekt numerieke technieken om wiskundige problemen, die partiële differentiaalvergelijkingen bevatten, op te lossen. Deze problemen hebben toepassingen in supergeleiding en in thermo-elasticiteit. Meer specifiek, voorwaartse problemen voor niet-lokale supergeleiding en inverse bronproblemen voor thermo-elasticiteit worden bestudeerd. Het onderzoek combineert bestaande technieken zoals de Rothemethode en regularisatiemethoden met nieuwe inzichten. Dit proefschrift bevat de resultaten van vijf publicaties die zijn opgenomen in Web of Science.

In het eerste hoofdstuk wordt de achtergrond van het onderzoek geschetst en een overzicht gegeven van de rest van de dissertatie. Deze studie heeft twee algemene doelstellingen. De eerste doelstelling is om wiskundige modellen op te stellen voor niet-lokale supergeleiding en om deze modellen te analyseren met behulp van de Rothemethode. De tweede doelstelling is om op basis van een bijkomende meting numerieke schema's te ontwikkelen om onbekende bronnen in thermo-elastische systemen te reconstrueren. Deze doelstellingen zijn in twee aparte delen in dit proefschrift behandeld.

Hoofdstuk 2 introduceert de Rothemethode. Dit is een hulpmiddel voor het oplossen van problemen geassocieerd met tijdsafhankelijke partiële differentiaalvergelijkingen. Een grondige introductie op partiële differentiaalvergelijkingen en de gerelateerde aspecten wordt gegeven, waarna de belangrijkste ideeën van deze methode op een voorbeeld geïllustreerd worden in Sectie 2.12. Een uitgebreid overzicht van notaties, definities en stellingen wordt gegeven. Deze zijn onmisbaar voor een theoretische en numerieke analyse van tijdsafhankelijke partiële differentiaalvergelijkingen met behulp van de Rothemethode. Dit inleidende hoofdstuk bevat ook een uitbreiding van een algemeen Aubin en Lions lemma. Dit lemma is cruciaal voor het bewijzen van de convergentie van de Rothemethode.

Het eerste deel, genaamd 'niet-lokale problemen voor supergeleiding', bestaat uit vier hoofdstukken (Hoofdstukken 3-6). In Hoofdstuk 3 wordt een theoretische inleiding op supergeleiding gegeven, inclusief de onderverdeling in zijn twee hoofdtypen: type-I en type-II supergeleiders. Het modelleringsgedeelte van dit proefschrift komt aan bod in Sectie 3.3. Hierin worden drie nieuwe macroscopische

modellen voor niet-lokale supergeleiding ontwikkeld: een niet-lokaal lineair parabolisch model voor type-I supergeleiding, een niet-lokaal lineair hyperbolisch model voor type-I supergeleiding en een niet-lokaal niet-lineair parabolisch model voor een tussentoestand tussen type-I en type-II supergeleiding. Alle modellen bevatten een ruimtelijke convolutie met singuliere kern en worden geanalyseerd met behulp van de Rothemethode. Deze methode helpt om vast te stellen of een probleem goed gesteld is, i.e. dat een oplossing bestaat, dat deze uniek is en dat de oplossing continu afhangt van de data. De behandeling van de convolutiekern is de belangrijkste nieuwigheid van de analyse.

Het niet-lokale lineaire parabolisch probleem voor type-I supergeleiders wordt bestudeerd in Hoofdstuk 4. Twee tijdsdiscrete numerieke schema's worden ontwikkeld om het probleem op te lossen. Het verschil tussen beide schema's ligt in een expliciete en een impliciete behandeling van de convolutieterm. De foutenschattingen corresponderend met de tijdsdiscretisatie worden afgeleid voor beide schema's. Daarna wordt een nieuwe convolutiekern berekend onder de aanname dat de normaalcomponent van het onbekende vectorveld gelijk is aan nul op de rand van de supergeleider. Het positief-definiet zijn van deze kern wordt bewezen. Met behulp van de extra veronderstelling wordt aangetoond dat onder hogere regulariteit de oplossing van het oorspronkelijk model voldoet aan een eenvoudiger probleem, dat gemakkelijker te implementeren is. Voor dit probleem blijven de tijdsdiscrete schema's geldig. Betere foutenschattingen worden verkregen voor het impliciete schema waarin de convolutieterm impliciet wordt behandeld. Een numeriek experiment voor het semi-impliciete schema ondersteunt de verkregen theoretische resultaten. De tijdsdiscrete problemen worden opgelost met behulp van de eindige elementenmethode. Een discrete convolutie benadert de convolutieintegraal (over een bol) zodanig dat de singulariteit in de kern vermeden wordt. Bovendien wordt de convergentie van een volledig discreet eindige elementen schema (discretisatie in tijd en ruimte) naar de oplossing van het probleem aangetoond. Op soortgelijke wijze als voor het tijdsdiscrete schema wordt uitgelegd hoe de foutenschattingen verbeteren onder hogere regulariteit van de data.

In Hoofdstuk 5 wordt een analoge analyse gemaakt van het lineaire hyperbolisch probleem. Twee tijdsdiscrete schema's (op basis van een expliciete en impliciete behandeling van de convolutieterm) om het magnetisch veld te bepalen worden opgesteld. Zoals in het parabolisch geval voldoet de oplossing van het oorspronkelijke model aan een eenvoudiger probleem in de veronderstelling dat de normale component van het onbekende vectorveld gelijk is aan nul op de rand van de supergeleider. Er worden geen betere foutenschattingen voor de tijdsdiscretisatie verkregen, ondanks het positief-definiet zijn van de kern.

Een niet-lokaal niet-lineair parabolisch probleem voor een tussentoestand tussen type-I en type-II supergeleiding wordt geanalyseerd in Hoofdstuk 6. Het idee achter dit model is de recente ontdekking van een nieuw soort supergeleider: type-1.5 supergeleiders. Dergelijke supergeleiders ontstaan bijvoorbeeld in multiband supergeleiders, welke bestaan uit meerdere supergeleidende condensaten binnen eenzelfde materiaal. Het bestaan van een unieke oplossing voor het probleem wordt bekomen onder lage regulariteitseisen op de data.

Het tweede doel van dit proefschrift wordt onderzocht in het tweede deel 'Inverse bronproblemen voor thermo-elasticiteit'. Dit deel bestaat uit drie hoofdstukken (Hoofdstukken 7-9). Een introductie tot thermo-elasticiteit en regularisatiemethoden voor inverse problemen wordt gegeven in Hoofdstuk 7. Een klassiek thermoelastisch systeem bestaat uit twee vergelijkingen die gekoppeld zijn: een parabolische (warmte) vergelijking en een vectoriële hyperbolische vergelijking voor de uitwijking. Inverse problemen zijn vaak slecht gesteld. Er zijn zogenaamde regularisatietechnieken ontwikkeld die toch tot bruikbare resultaten leiden. De bekendste regularisatiemethode is afkomstig van Tichonoff.

Hoofdstuk 8 focust op de reconstructie van een vectorbron, die enkel plaatsafhankelijk is, door middel van een meting van de verplaatsing op het eindtijdstip. Deze meting zorgt ervoor dat het inverse probleem een unieke oplossing bezit als een dempingsterm toegevoegd wordt aan de hyperbolische vergelijking voor de verplaatsing in het klassieke thermo-elastische systeem. Het probleem is slecht gesteld omdat de oplossing instabiel is. Een algoritme gebaseerd op een stabiele iteratieve regularisatiemethode (een rij van goed gestelde directe problemen) wordt voorgesteld om de onbekende bron terug te vinden wanneer de dempingsterm lineair is. De directe problemen worden op elke iteratiestap opgelost met behulp van de eindige elementenmethode. De instabiliteit wordt overwonnen door het stoppen van de iteraties op de eerste iteratie waarvoor het discrepantiebeginsel voldaan is. Numerieke resultaten worden gepresenteerd voor een aantal testvoorbeelden.

Hoofdstuk 9 besteedt uitgebreid aandacht aan de reconstructie van een louter tijdsafhankelijke bron in een eendimensionaal thermo-elastisch systeem op basis van een meting in de tijd van de gemiddelde temperatuur in het lichaam. De nieuwigheid van de analyse betreft het herformuleren van het inverse bronprobleem in een geschikte directe formulering. Dit gebeurt door eliminatie van de onbekende bron aan de hand van de bijkomende meting. Het goed gesteld zijn van het probleem wordt aangetoond. Het voorgestelde numerieke schema maakt gebruik van de semi-discretisatie in de tijd volgens de Rothemethode.

De dissertatie wordt afgesloten in Hoofdstuk 10 met een samenvatting van de belangrijkste resultaten en enkele perspectieven voor verder onderzoek. Deze studie levert drie nieuwe modellen voor niet-lokale supergeleiding. Numerieke schema's zijn ontwikkeld om een benaderende oplossing van de betreffende problemen te vinden. In dit proefschrift zijn ook twee specifieke inverse bronproblemen voor thermo-elasticiteit bestudeerd. Ook voor deze problemen is een manier gevonden om een benadering van de oplossing te bekomen.

English summary

This dissertation investigates numerical techniques to solve mathematical problems containing partial differential equations. These problems arise in superconductivity and in thermoelasticity. More precisely, forward problems for nonlocal superconductivity and inverse source problems for thermoelasticity are studied. The research combines existing techniques such as Rothe's method and regularization methods with new insights. This thesis contains results from five publications, which are included in Web of Science.

The first chapter sketches the background of the study and provides an outline of the rest of the dissertation. Two research objectives are proposed, dividing this PhD-thesis into two parts. The first objective is to establish mathematical models for nonlocal superconductivity and to analyse these models using Rothe's method. The second objective is to recover unknown sources in thermoelastic systems from additional data.

Chapter 2 introduces Rothe's method, a tool for solving problems associated with time-dependent partial differential equations. Before illustrating the main ideas of this method on a example in Section 2.12, a thorough introduction on partial differential equations and its aspects is given. A comprehensive overview of notations, definitions and theorems is given. These preliminaries are indispensable for a theoretical and numerical analysis of time-dependent partial differential equations when using Rothe's method as solving tool. This introductory chapter also contains an extension of a generalized Aubin and Lions lemma, which is crucial to prove the convergence of Rothe's method.

The first part, named 'Nonlocal problems for superconductivity', consists of four chapters (Chapters 3-6). In Chapter 3, an introduction on superconductivity is given including the subdivision into its two main types: type-I and type-II superconductors. The modelling part is contained in Section 3.3. In that section, three new macroscopic models for nonlocal superconductivity are developed: a nonlocal linear parabolic problem for type-I superconductivity, a nonlocal linear hyperbolic problem for type-I superconductivity and a nonlocal nonlinear parabolic model for an intermediate state between type-I and type-II superconductivity. All models contain a space convolution with singular kernel and are analysed using Rothe's method in separated chapters. This method helps to establish the well-posedness of the different problems, i.e. to prove the existence, uniqueness, and continuous

dependence of the solution on the given data. The handling of the convolution kernel is the main novelty of this analysis.

The nonlocal linear parabolic problem for type-I superconductors is studied in Chapter 4. Two time-discrete schemes are established to solve the problem, based on an explicit and implicit handling of the convolution term respectively. The error estimates corresponding to the time discretization are derived for both schemes. Afterwards, a new convolution kernel is deduced under the assumption that the normal component of the unknown vector field equals zero on the boundary of the superconductor. The positive definiteness of this kernel is shown. With the aid of the additional assumption, it is demonstrated that under higher regularity the solution of the original model satisfies a simpler problem, which is easier to implement. For this problem, both time-discrete schemes remain valid. Better error estimates are obtained for the implicit scheme in which the convolution term is handled implicitly. A numerical experiment for the semi-implicit scheme supports the obtained theoretical results. The time-discrete problems are solved using the finite element method. By a space-discrete convolution, the convolution integral (over a ball) is approximated in such a way that the singularity in the kernel is avoided. Moreover, the convergence of a fully discrete finite element scheme (discretization in time and space) to the solution of the problem is shown. In a similarly way to the time-discrete scheme, it is demonstrated how to improve the error estimates under higher regularity of the data.

In Chapter 5, an analogue analysis is made for the linear hyperbolic problem. Two time-discrete schemes (based on an explicit and an implicit handling of the convolution term) to approximate the magnetic field are established. As in the parabolic case, the solution of the original model satisfies a simpler problem under the assumption that the normal component of the unknown vector field equals zero on the boundary of the superconductor. No better error estimates for the time discretization are obtained despite the positive definiteness of the kernel.

A nonlocal nonlinear parabolic problem for an intermediate state between type-I and type-II superconductivity is analysed in Chapter 6. The idea behind this model is the recent discovery of a new type of superconductivity: type-1.5 superconductivity. This type of superconductivity arises for instance in multiband superconductors, which have at least two superconducting components. The existence of a unique solution to the problem is shown under low regularity assumptions on the data.

The second goal of this dissertation is studied in the second part ‘Inverse source problems in thermoelasticity’. This part consists of three chapters (Chapters 7-9). An introduction to thermoelasticity and regularization methods for inverse problems is given in Chapter 7. A classical thermoelastic system consists of two equations that are coupled: a parabolic (heat) equation and a vectorial hyperbolic equation for the displacement. Regularization methods can deal with the natural ill-posedness of linear and nonlinear inverse problems.

In Chapter 8, a solely space-dependent vector source is reconstructed using information from a final in time measurement of the displacement. This measurement ensures that the inverse problem has a unique solution when a damping term is added in the hyperbolic equation for the displacement in the classical thermoelastic system. The problem is ill-posed since the solution is unstable. An algorithm based on a stable iterative regularization method (thus on a sequence of well-posed direct problems) is proposed to recover the unknown source in the case that the damping term is linear. The direct problems are solved at each iteration step using the finite element method. The instability is overcome by stopping the iterations at the first iteration for which the discrepancy principle is satisfied. Numerical results are presented for some test examples.

Chapter 9 focuses on the reconstruction of a purely time-dependent source in a one-dimensional thermoelastic system from a measurement in time of the average temperature inside the body. The novelty of the analysis is the reformulation of the inverse source problem into an appropriate direct formulation. This is done by eliminating the unknown source by taking the additional measurement into account. The well-posedness of the problem is shown. The proposed numerical scheme involves the semidiscretization in time by Rothe's method.

Chapter 10 gives an overview of the findings of this study and concludes with some suggestions for future research. This study delivers three new models for nonlocal superconductivity. Numerical schemes are developed to approximate the solution to the related problems. Moreover, two specific inverse source problems for thermoelasticity are considered. Also for these problems, a way of retrieving the solution is established.

1 Introduction

This PhD-thesis consists of two parts. In the first part, forward problems for nonlocal superconductivity are studied. The second part concerns inverse source problems for thermoelasticity. Although these two parts are separated, they do have something in common, i.e. all problems in these parts, except one, are solved using Rothe's method, which is a tool for solving evolution problems. The introduction to this method is the main goal of Chapter 2: the mathematical background. More specifically, in Section 2.12, Rothe's method is applied on a simple example. This is done to illustrate the main ideas of this method and to give some useful insights. Also an extension of a generalized Aubin and Lions lemma is given. This lemma is a useful tool when applying Rothe's method. The following section provides an introduction and outline of the main parts of this thesis.

1.1 Goals and outline

Partial differential equations (PDEs) are equations involving unknown functions of two or more variables and certain of their partial derivatives. They form an indispensable part of mathematical modeling of a wide variety of phenomena such as the propagation of sound or heat, electrostatics, electrodynamics, fluid flow and elasticity.

There is no general theory known concerning the solvability of all partial differential equations. Such a theory is extremely unlikely to exist, given the rich variety of physical, geometric, and probabilistic phenomena which can be modeled by PDEs [1]. Research focuses on various particular (systems of) PDEs that are important for applications, for instance the Laplace equation, the heat (diffusion) equation, the wave equation, the porous medium equation and the Maxwell equations.

There are different notions of what is a solution of a PDE. A classical solution of a PDE is a function having continuous partial derivatives of any order involved in the equation. Many PDEs cannot be solved in the classical sense. To deal with this issue, the smoothness property can be abandoned and the solution can be searched for in a wider class of candidates. Even for those PDEs that turn out to be classically solvable, it is often recommended to search initially for an appropriate kind of a so-called weak solution. It might then be easier to establish the well-posedness of the problem associated with the PDE, i.e. to prove the existence, uniqueness and continuous dependence of the solution on the given data.

Also the availability of powerful computers is shifting the emphasis in partial differential equations away from the analytical computation of solutions towards existence questions and their numerical computation.

This leads to the research topic of this dissertation: the development of numerical techniques to solve problems for partial differential equations appearing in superconductivity and in thermoelasticity. Superconductivity is a phenomenon of exactly zero electrical resistance and expulsion of magnetic fields occurring in certain materials when cooled below a characteristic critical temperature. Thermoelasticity is the change in the size and shape of a solid object as the temperature of that object fluctuates.

In the following, the goals of this study and the organization of the dissertation are provided.

Goal I: To establish mathematical models for nonlocal superconductivity and to analyse these models using Rothe's method.

The study of this goal builds on Chapter 11 'Nonlocal electromagnetism and superconductivity' of the book of Fabrizio and Morro [2], in which two nonlocal models for superconductivity are given: Pippard's and Eringen's nonlocal law. The results can be found in the first part 'Nonlocal problems for superconductivity' of this thesis, which contains four chapters (Chapters 3-6).

The phenomena of superconductivity and the subdivision in type-I and type-II superconductivity are explained in more details in Chapter 3 (Section 3.1 and 3.2). Section 3.3 contains the modelling part of the thesis. In this section, two new macroscopic models (a parabolic and hyperbolic one) for type-I superconductivity in terms of the magnetic field are derived, using Eringen's law and Maxwell's equations. Moreover, a macroscopic model for an intermediate state between type-I and type-II superconductivity is proposed. The idea behind the last model is the recent discovery of a new type of superconductivity arising for instance in multi-band superconductors (superconductors with at least two superconducting components): type-1.5 superconductivity. The three models contain a space convolution with singular kernel. The parabolic model for type-I superconductivity is studied in detail in Chapter 4, the hyperbolic model in Chapter 5. The intermediate state model is analysed in Chapter 6. All these models are solved using Rothe's method. The advantage of Rothe's method is that it contains a numerical algorithm to retrieve the unknown magnetic field.

Goal II: To recover unknown sources in thermoelastic systems from additional data.

The second goal of this dissertation is studied in the second part 'Inverse source problems in thermoelasticity' of this dissertation. This part consists of three chapters (Chapters 7-9).

An introduction to thermoelasticity and regularization methods for inverse problems is given in Chapter 7. Thermoelastic systems consist of two equations that are coupled: a parabolic (heat) equation and a vectorial hyperbolic equation for the displacement. Inverse problems are, roughly speaking, those where from measured data of a system one aims to recover the unknown model parameters of the system. Regularization methods can deal with the natural ill-posedness of inverse problems. Ill-posedness means that there is either no solution, or if there is any, then it might not be unique or might not depend continuously on the data. In this dissertation, two inverse source problems for thermoelasticity are considered. In Chapter 8, the goal is to determine a solely space-dependent vector source using an iterative regularization method. Chapter 9 focuses on the reconstruction of a time-dependent source using Rothe's method.

Chapter 10 summarizes the findings of the thesis. The two general goals of this dissertation are reviewed. Moreover, the significance of the present study and directions for future research are also addressed. This introduction ends with an overview of the publications related to the work presented in this dissertation.

1.2 Publications

The results contained in the first part of this thesis are published in:

- M. Slodička and K. Van Bockstal. *A nonlocal parabolic model for type-I superconductors*. Numerical Methods for Partial Differential Equations, 30(6):1821–1853, 2014;
- K. Van Bockstal and M. Slodička. *Error estimates for the full discretization of a nonlocal parabolic model for type-I superconductors*. Journal of Computational and Applied Mathematics, 275(0):516–526, 2015;

- K. Van Bockstal and M. Slodička. *The well-posedness of a nonlocal hyperbolic model for type-I superconductors*. Journal of Mathematical Analysis and Applications, 421(1):697 – 717, 2015;
- K. Van Bockstal and M. Slodička. *A macroscopic model for an intermediate state between type-I and type-II superconductivity*. Numerical Methods for Partial Differential Equations, 31(5):1551–1567, 2015.

The second part is based on two articles:

- K. Van Bockstal and M. Slodička. *Recovery of a space-dependent vector source in thermoelastic systems*. Inverse Problems in Science and Engineering, 23(6):956–968, 2015;
- K. Van Bockstal and M. Slodička. *Recovery of a time-dependent heat source in one-dimensional thermoelasticity of type-III*. Inverse Problems in Science and Engineering (submitted), 2015.

The results of the subsequent papers are not contained in this thesis:

- K. Van Bockstal and M. Slodička. *Determination of an unknown diffusion coefficient in a semilinear parabolic problem*. Journal of Computational and Applied Mathematics, 246(0):104–112, 2013. Fifth International Conference on Advanced COmputational Methods in ENgineering (ACOMEN 2011);
- K. Van Bockstal and M. Slodička. *Determination of a time-dependent diffusivity in a nonlinear parabolic problem*. Inverse Problems in Science and Engineering, 23(2):307–330, 2015;
- R.H. De Staelen, K. Van Bockstal, and M. Slodička. *Error analysis in the reconstruction of a convolution kernel in a semilinear parabolic problem with integral overdetermination*. Journal of Computational and Applied Mathematics, 275(0):382 – 391, 2015;
- K. Van Bockstal, R.H. De Staelen, and M. Slodička. *Identification of a memory kernel in a semilinear integrodifferential parabolic problem with applications in heat conduction with memory*. Journal of Computational and Applied Mathematics, 289(0):196–207, 2015. Sixth International Conference on Advanced Computational Methods in Engineering (ACOMEN 2014).

2 Mathematical background

The objective of this chapter is to introduce notations, definitions and theorems that are indispensable for a theoretical and numerical analysis of partial differential equations when using Rothe's method. Some theoretical results are illustrated with examples. The reader who is familiar with the topics in this chapter should be able to understand the main chapters of this dissertation. The reader is expected to be comfortable with linear algebra (e.g. vector spaces), real mathematical analysis (e.g. partial derivatives) and Lebesgue theory. The underlying scalar field of a (linear) vector space is denoted by $\mathbb{F}$ and stands for the real numbers ($\mathbb{R}$) or the complex numbers ($\mathbb{C}$). For more details the reader is referred for instance to [1, 3–14].

The outline of this chapter is as follows. In Section 2.1 and 2.2, differential operators and the basic properties of functions are recapitulated. Section 2.3 contains a list of useful inequalities that are used in this dissertation. Afterwards, Section 2.4 includes the main aspects of functional analysis. First, normed vector spaces and inner product spaces over $\mathbb{F}$ are studied. Then, the concept completeness is introduced. This leads to complete normed spaces that are called Banach spaces. An important example is a Hilbert space, where the norm arises from an inner product. The section is continued with a overview of the properties of continuous and compact linear operators in normed vector spaces (including embeddings) and it ends with the difference between strong and weak convergence. In Section 2.5, the basic function spaces $\mathrm{C}^m(\Omega)$ and $\mathrm{C}^m(\overline{\Omega})$ are introduced. Domains in the d-dimensional real space that have a Lipschitz continuous boundary are defined in Section 2.6. This type of domains play an important role in the analysis of partial differential equations. For instance, the famous Green theorems are valid on a Lipschitz domain. The Lebesgue spaces and its characteristics are introduced in Section 2.7.

The important notion of weak partial derivative is introduced in Section 2.8. This notion plays a crucial role in the definition of the Sobolev spaces, see Section 2.9. Sobolev spaces and their properties are discussed, including embedding theorems, traces of functions, Sobolev spaces for vector fields and Green's formulas for functions in Sobolev spaces. These abstract function spaces can be used for solving partial differential equations (PDEs). Section 2.10 concerns the classification of partial differential equations and their associated conditions. It is explained how to solve linear and nonlinear elliptic PDEs in Section 2.11, which leads to the famous Lax-Milgram lemma for linear elliptic PDEs and the monotone operator theory for nonlinear elliptic PDEs.

Finally, in Section 2.12, Rothe's method is presented as a tool for solving evolutionary (time-dependent) PDEs. This is done by explaining this method using an example. Rothe's method is based on a time discretization and is the basis tool used in this thesis for solving the PDE under consideration. In fact, evolution problems are approximated by a sequence of corresponding elliptic problems that can be solved by the finite element method for elliptic equations, see Section 2.13.

2.1 Differential operators in Cartesian coordinates

Definition 2.1.1 (The space $\mathbb{R}^d$). *Let $d \in \mathbb{N} := \{1, 2, 3, \ldots\}$. The standard basis in the d-dimensional real space $\mathbb{R}^d$ is denoted as*

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \quad \ldots \quad , \mathbf{e}_d = \underbrace{(0, \ldots, 1)}_{d-tuple}.$$

The standard origin in $\mathbb{R}^d$ is the d-tuple $(0, 0, \ldots, 0)$. Each element $\mathbf{v} \in \mathbb{R}^d$ can be written in a unique way as a linear combination,

$$\mathbf{v} = \sum_{i=1}^{d} v_i \mathbf{e}_i.$$

The components $(v_1, v_2, \ldots, v_d)$ of a vector $\mathbf{v}$ in $\mathbb{R}^d$ are called the Cartesian coordinates with respect to the basis $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d\}$. The Euclidean inner product of two vectors $\mathbf{u} = \sum_{i=1}^{d} u_i \mathbf{e}_i$ and $\mathbf{v} = \sum_{i=1}^{d} v_i \mathbf{e}_i$ is defined by

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{d} u_i v_i.$$

The Euclidean norm of a vector $\mathbf{v}$ *in* $\mathbb{R}^d$ *is expressed by*

$$|\mathbf{v}|_{\mathrm{e}} = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\sum_{i=1}^{d} v_i^2}.$$

For $d = 1$*, this is the absolute value of a real number, i.e.* $|x|_{\mathrm{e}} = |x|$ *for all* $x \in \mathbb{R}$*. The standard basis* $\mathcal{B}$ *is orthonormal, meaning that*

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

The symbol δ_{ij} *is known as Kronecker's delta. The cross or vector product of two vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ *is defined as*

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \sum_{i,j,k=1}^{3} \mathcal{E}_{ijk} u_j v_k \mathbf{e}_i,$$

where $\mathcal{E}_{ijk}$ *denotes the three-dimensional Levi-Civita symbol defined by*

$$\mathcal{E}_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of } 123, \\ -1 & \text{if } ijk \text{ is an odd permutation of } 123, \\ 0 & \text{otherwise}. \end{cases}$$

In $\mathbb{R}^3$*, the following vector identities are frequently used*

$$\begin{aligned} \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}), \\ \mathbf{a}\times(\mathbf{b}\times\mathbf{c}) &= (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}, \\ (\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}). \end{aligned}$$

Definition 2.1.2 (The space $\mathbb{C}^d$)**.** *The space of complex numbers is denoted by* $\mathbb{C}$*. This space consists of complex numbers* z *that can be written in the form* $z = x+iy$ *with* x *and* y *real numbers and where* i *is the imaginary unit satisfying* $i^2 = -1$*. The complex conjugate of a complex number, denoted* $\overline{z}$*, is defined as* $\overline{z} = x - iy$*. The modulus of a complex number, denoted* $|z|_{\mathrm{c}}$*, is defined as*

$$|z|_{\mathrm{c}} = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}.$$

The standard basis $\mathcal{B}$ *for* $\mathbb{R}^d$ *is also the standard basis for* $\mathbb{C}^d$*. Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be vectors in* $\mathbb{C}^d$*. The Hermitian inner product of* $\mathbf{u}$ *and* $\mathbf{v}$ *is given by*

$$\mathbf{u}\cdot\mathbf{v} = \sum_{i=1}^{d} u_i \overline{v}_i.$$

The Hermitian norm of $\mathbf{v}$ *in* $\mathbb{C}^d$ *is defined as*

$$|\mathbf{v}|_{\mathrm{e}} = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\sum_{i=1}^{d} |v_i|_{\mathrm{c}}^2},$$

where the same notation is used as for the Euclidean norm.

Definition 2.1.3 (Cartesian product)**.** *The Cartesian product of two sets* X *and* Y *is the set of all ordered pairs, written* (x, y)*, where* x *is an element of* X *and* y *is an element of* Y*. It is specified by the following notation*

$$X \times Y = \{\, (x, y) \mid x \in X \;\; \text{and} \;\; y \in Y \,\}.$$

Definition 2.1.4 (Scalar Field, Vector Field, Gradient, Divergence, Rotor, Laplacian)**.** *Let* Ω *be an open subset of* $\mathbb{R}^d$*. In what follows,* $\mathbf{x}$ *is a short notation for the* d*-tuple* $(x_1, \ldots, x_d) \in \Omega$*. A scalar field* ϕ *is defined on* $\Omega \subset \mathbb{R}^d$ *as*

$$\phi : \Omega \to \mathbb{F} : \mathbf{x} \to \phi(\mathbf{x}),$$

where $\mathbb{F} = \mathbb{R}$ *or* $\mathbb{C}$*. A vector field* $\mathbf{f}$ *is defined on* Ω *as*

$$\mathbf{f} : \Omega \to \mathbb{F}^d : \mathbf{x} \to \mathbf{f}(\mathbf{x}) := f_1(\mathbf{x})\mathbf{e}_1 + \ldots + f_d(\mathbf{x})\mathbf{e}_d.$$

The operator ∇ *is the vector differential operator*

$$\nabla = \mathbf{e}_1 \partial_{x_1} + \ldots + \mathbf{e}_d \partial_{x_d} = (\partial_{x_1}, \ldots, \partial_{x_d}).$$

This operator can act on a differentiable scalar field ϕ *and on a differentiable vector field* $\mathbf{f}$ *as follows*

$$\nabla \phi = (\partial_{x_1}\phi, \partial_{x_2}\phi, \ldots, \partial_{x_d}\phi),$$

$$\nabla \cdot \mathbf{f} = \sum_{i=1}^{d} \partial_{x_i} f_i$$

and on a 3-dimensional vector field $\mathbf{f}$ *as follows*

$$\nabla \times \mathbf{f} = \sum_{i,j,k=1}^{3} \mathcal{E}_{ijk} \partial_{x_j} f_k \mathbf{e}_i = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_x & \partial_y & \partial_z \\ f_1 & f_2 & f_3 \end{vmatrix}.$$

These operations are called respectively gradient, divergence and rotor (curl). Moreover, the gradient of a differentiable vector field $\mathbf{f}$ *is defined by*

$$\nabla \mathbf{f} = \sum_{i,j=1}^{d} \partial_{x_j} f_i E_{ij},$$

with E_{ij} matrices in $\mathbb{R}^{d\times d}$ with 1 at position (i,j) and 0 everywhere else. The gradient of a vector field is the Jacobian matrix

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_d}{\partial x_1} & \cdots & \dfrac{\partial f_d}{\partial x_d} \end{pmatrix}.$$

For a differentiable vector field $\mathbf{v}$ and a differentiable vector field $\mathbf{w}$, the convective term is given by

$$(\nabla \mathbf{w})\,\mathbf{v} = \begin{pmatrix} \frac{\partial w_1}{\partial x_1} & \cdots & \frac{\partial w_1}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial w_d}{\partial x_1} & \cdots & \frac{\partial w_d}{\partial x_d} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} = \sum_{i=1}^{d} (\nabla w_i \cdot \mathbf{v})\mathbf{e}_i.$$

The Laplace operator or Laplacian is a differential operator given by

$$\Delta\phi = \nabla^2\phi = \nabla\cdot\nabla\phi = \sum_{i=1}^{d} \frac{\partial^2\phi}{\partial x_i^2},$$

in the case of a twice-differentiable scalar field ϕ. The vector Laplace operator of a twice-differentiable vector field $\mathbf{f}$ in $\mathbb{R}^d$ is defined by

$$\Delta\mathbf{f} = (\Delta f_1, \Delta f_2, \ldots, \Delta f_d).$$

In $\mathbb{R}^3$, the following equality holds

$$\Delta\mathbf{f} = \nabla(\nabla\cdot\mathbf{f}) - \nabla\times(\nabla\times\mathbf{f}).$$

2.2 Properties of functions

In this section, basic properties of functions are reviewed.

Definition 2.2.1 (Function, domain, range, codomain)**.** *A function (mapping) from a set X to a set Y is an object f such that every x in X is uniquely associated with an object $f(x)$ in Y. A function is therefore a many-to-one (or sometimes one-to-one) relation. The set X of values at which a function is defined is called its domain, while the set $f(X) \subset Y$ of values that the function can produce is called its range. The set Y is called the codomain of f.*

Definition 2.2.2 (Metric space)**.** *Let X be a set. A nonnegative function d_X defined on $X\times X$ is called a metric if it satisfies the following properties for any $x, y, z \in X$:*

- $d_X(x,y) = 0$ *if and only if (iff)* $x = y$;

- $d_X(x, y) = d_X(y, x)$;
- $d_X(x, z) \leqslant d_X(x, y) + d_X(y, z)$.

A set X with a metric d_X is called a metric space.

Example 2.2.1. *The space of the real numbers and the space of complex numbers with the distance function $d(x, y) = |y - x|_{\mathrm{c}}$ are metric spaces. More generally, the d-dimensional Euclidean space with the Euclidean distance $d(\mathbf{x}, \mathbf{y}) = |\mathbf{y} - \mathbf{x}|_{\mathrm{e}}$ is a metric space.*

Definition 2.2.3 (Bounded subset, bounded metric space)**.** *A subset S of a metric space (X, d_X) is bounded if it is contained in a ball of finite radius, i.e. if there exists x in X and $r > 0$ such that for all s in S, we have that $d_X(x, s) < r$. X is a bounded metric space (or d_X is a bounded metric) if X is bounded as a subset of itself.*

Definition 2.2.4 (Properties of functions)**.** *A mapping f between two metric spaces (X, d_X) and (Y, d_Y) is*

- *bounded if $f(X)$ is a bounded subset of Y;*
- *continuous if*

$$(\forall y \in X)(\forall \varepsilon > 0)(\exists \delta(\varepsilon, y) > 0) \\ (\forall x \in X : d_X(x, y) < \delta \Rightarrow d_Y\left(f(x), f(y)\right) < \varepsilon);$$

- *uniform continuous if*

$$(\forall \varepsilon > 0)(\exists \delta(\varepsilon) > 0)(\forall x, y \in X : d_X(x, y) < \delta \Rightarrow d_Y\left(f(x), f(y)\right) < \varepsilon);$$

- *Lipschitz continuous if there exists a positive real number $L > 0$ such that*

$$d_Y\left(f(x), f(y)\right) \leqslant L\, d_X(x, y),\ \forall x, y \in X;$$

- *Hölder continuous if there exists positive real numbers $L > 0$ and $\alpha \in (0, 1]$ such that*

$$d_Y\left(f(x), f(y)\right) \leqslant L\, d_X(x, y)^{\alpha},\ \forall x, y \in X;$$

- *a contraction if the map is Lipschitz continuous with Lipschitz constant $L < 1$;*
- *open if it maps open sets to open sets (image is also open);*
- *closed if it maps closed sets to closed sets.*

Remark 2.2.1. *Pay attention on the possible values for α in the definition of Hölder continuity. If $\alpha = 0$ in the inequality, then the mapping f is bounded. If $\alpha > 1$, then only constant mappings can satisfy this inequality.*

A function that is differentiable everywhere (in one dimension: a function whose first derivative exists at each point in its domain) is continuous. The following lemma gives the condition for an everywhere differentiable function to be a Lipschitz continuous function.

Lemma 2.2.1. *An everywhere differentiable function $g : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous (with $L = \sup |g'(x)|$) iff it has bounded first derivative.*

Absolute continuity is a property of functions that is stronger than continuity and uniform continuity.

Definition 2.2.5 (Absolutely continuous function)**.** *Let I be an interval in $\mathbb{R}$ and (X, d_X) a metric space. A function $f : I \to X$ is absolutely continuous on I if for every positive number ε, there is a positive number δ such that whenever a finite sequence of pairwise disjoint subintervals (x_k, y_k) of I satisfies*

$$\sum_k |y_k - x_k| < \delta \Rightarrow \sum_k d_X\left(f(y_k) - f(x_k)\right) < \varepsilon.$$

The following conditions on a real-valued function f on a compact interval $[a, b]$ are equivalent [15, Theorem 20.8]:

- *f is absolutely continuous;*
- *f has a derivative f' almost everywhere (a.e., differentiable at every point outside a set of Lebesgue measure zero), the derivative is Lebesgue integrable, and*

$$f(x) = f(a) + \int_a^x f'(t)\,dt$$

  for all x in $[a, b]$;
- *there exists a Lebesgue integrable function g on $[a, b]$ such that*

$$f(x) = f(a) + \int_a^x g(t)\,dt$$

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

If these equivalent conditions are satisfied then necessarily $g = f'$ a.e..

Lemma 2.2.2. *If $f : I \subset \mathbb{R} \to \mathbb{R}$ is absolutely continuous and satisfies $|f'(x)| \leqslant K$ for almost all (a.a.) $x \in I$, then f is Lipschitz continuous with the Lipschitz constant at most K.*

The next property is valid for Hölder continuous functions.

Lemma 2.2.3. *If $0 < \alpha \leqslant \beta \leqslant 1$ then all β-Hölder continuous functions on a bounded set $I \subset \mathbb{R}$ are also α-Hölder continuous.*

Remark 2.2.2. *This lemma includes $\beta = 1$ and therefore all Lipschitz continuous functions on a bounded set are also α-Hölder continuous with $\alpha \in (0, 1]$.*

The following diagram is valid between the different types of continuity for functions over a compact (closed and bounded) subset of the real line, with $0 < \alpha \leqslant 1$:

$$
\begin{tikzcd}
& \text{Continuously differentiable} \arrow[d, Rightarrow] & \\
& \text{Lipschitz continuous} \arrow[d, Rightarrow] \arrow[dr, Rightarrow] & \\
\text{Differentiable a.e.} & \text{Absolutely continuous} \arrow[l, Rightarrow] \arrow[d, Rightarrow] & \alpha\text{-Hölder continuous} \arrow[dl, Rightarrow] \\
& \text{Uniformly continuous} \arrow[d, Rightarrow] & \\
& \text{Continuous.} &
\end{tikzcd}
$$

Note that a function is continuously differentiable if its derivative exists and is itself a continuous function. The following examples clarify the diagram.

Example 2.2.2.

(i) *Let $a, b \in \mathbb{R}$. The linear function $f : \mathbb{R} \to \mathbb{R} : x \mapsto ax + b$ is Lipschitz continuous with the Lipschitz constant $|a|$.*

(ii) *The function $f(x) = \sqrt{x^2 + \alpha}$ with $\alpha > 0$ defined for all real numbers is Lipschitz continuous with the Lipschitz constant 1, because it is differentiable everywhere and the absolute value of the derivative is bounded above by 1.*

(iii) *The function $f(x) = |x|$ defined for all real numbers is Lipschitz continuous with the Lipschitz constant equal to 1, by the reverse triangle inequality. This is an example of a Lipschitz continuous function that is not differentiable everywhere.*

(iv) *The function $f(x) = \sqrt{x}$ defined on $[0, 1]$ is not Lipschitz continuous because there exists no positive constant C such that $\left|\sqrt{x} - \sqrt{0}\right| \leqslant C\,|x - 0|$ for all $x \in [0, 1]$. However, this function is Hölder continuous with exponent $\alpha \in \left(0, \frac{1}{2}\right]$, see example (vii).*

(v) *The exponential function becomes arbitrarily steep as $x \to \infty$, and therefore is not globally (uniformly, in each point $x \in \mathbb{R}$) Lipschitz continuous, despite being an absolutely continuous function.*

(vi) *The function $f(x) = x^2$ with domain $\mathbb{R}$ is not Lipschitz continuous, not absolutely continuous and also not uniform continuous. This function becomes arbitrarily steep as x approaches infinity. It is however locally Lipschitz continuous (for every x in $\mathbb{R}$ there exists a neighborhood U of x such that f restricted to U is Lipschitz continuous). Any continuously differentiable function is locally Lipschitz continuous.*

(vii) *The function $f(x) = x^\beta$ (with $0 < \beta \leqslant 1$) defined on $[0, \infty)$ is β-Hölder continuous. The function f is β-Hölder continuous if there exists a constant $L > 0$ such that for all $x, y \in [0, \infty)$, it holds that*

$$\left|x^\beta - y^\beta\right| \leqslant L\,|x - y|^\beta .$$

If $x = 0$, $y = 0$ or $x = y$, then the inequality is obvious. Without loss of *generality, let $x > y$ and $x \neq 0$. Then*

$$\frac{\left|x^\beta - y^\beta\right|}{|x-y|^\beta} = \frac{\left|1 - \left(\frac{y}{x}\right)^\beta\right|}{\left|1 - \frac{y}{x}\right|^\beta}.$$

For $0 \leqslant \hat{x} \leqslant 1$ and $0 \leqslant \alpha \leqslant 1$, it holds that $0 \leqslant \hat{x} \leqslant \hat{x}^\alpha \leqslant 1$. Therefore, it is clear that $\left|1 - \frac{y}{x}\right|^\beta \geqslant \left|1 - \frac{y}{x}\right|$ and $\left|1 - \left(\frac{y}{x}\right)^\beta\right| \leqslant \left|1 - \frac{y}{x}\right|$. Hence,

$$\frac{\left|x^\beta - y^\beta\right|}{|x-y|^\beta} \leqslant \frac{\left|1 - \frac{y}{x}\right|}{\left|1 - \frac{y}{x}\right|} = 1 =: L.$$

If the function f is defined on $[0,1]$, then f is α-Hölder continuous for $0 < \alpha \leqslant \beta$ thanks to Lemma 2.2.3. Moreover, f is absolutely continuous on $[0,1]$ since $f(x) = \int_0^x \beta x^{\beta-1}\,\mathrm{d}x$. Furthermore, note that if $\beta > 1$, then the function $f(x) = x^\beta$ defined on $[0,1]$ is Lipschitz continuous by Lemma 2.2.2.

(viii) The function $f(x) = |x|^\beta$ with domain all real numbers and $\beta \in \mathbb{R}$ is
- *not continuous if $\beta < 0$ (not defined in $x = 0$);*
- *Lipschitz continuous if $\beta = 0$ (providing $0^0 = 1$);*
- *absolutely continuous if $\beta > 0$;*
- *not Lipschitz continuous but Hölder continuous if $\beta \in (0,1)$;*
- *Lipschitz continuous if $\beta = 1$;*
- *locally Lipschitz continuous if $\beta > 1$.*

The following applications of continuous functions in one dimension are very important.

Theorem 2.2.3 (Extreme value theorem)**.** *A continuous function $f : \mathbb{R} \to \mathbb{R}$ attains its extreme values (maximum and minimum) on any closed and bounded interval $[a,b]$.*

Theorem 2.2.4 (Rolle's theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b), and $f(a) = f(b)$, then there exists a point $\eta \in (a,b)$ such that $f'(\eta) = 0$.*

Theorem 2.2.5 (Mean value theorem)**.** *Let $f : [a,b] \to \mathbb{R}$ be a continuous function on the closed interval $[a,b]$, and differentiable on the open interval (a,b). Then there exists some $\eta \in (a,b)$ such that*

$$\frac{f(b) - f(a)}{b - a} = f'(\eta).$$

The following assertion is usually called the first mean value theorem for the Lebesgue integral [16, Theorem 2.12.16].

Theorem 2.2.6 (First mean value theorem for Lebesgue integration)**.** *If a real-valued function $f \geqslant 0$ is integrable on $[a, b]$ and a real-valued function g is continuous on $[a, b]$, then there exists a constant $\xi \in [a, b]$ such that*

$$\int_a^b f(t)g(t)\,\mathrm{d}t = g(\xi)\int_a^b f(t)\,\mathrm{d}t.$$

The first part of the fundamental theorem of calculus and Leibniz's integral rule can be combined into a more general result.

Theorem 2.2.7 (Differentiation under the integral sign: Leibniz's rule)**.** *Let $f(x, t)$ be a real-valued function such that both $f(x, t)$ and its partial derivative $f_x(x, t)$ are continuous in t and x in some region of the (x, t)-plane, including $a(x) \leqslant t \leqslant b(x)$, $x_0 \leqslant x \leqslant x_1$. Also suppose that the functions $a(x)$ and $b(x)$ are both continuously differentiable for $x_0 \leqslant x \leqslant x_1$. Then for $x_0 \leqslant x \leqslant x_1$:*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}&\left(\int_{a(x)}^{b(x)} f(x, t)\,\mathrm{d}t\right)\\ &= f(x, b(x))b'(x) - f(x, a(x))a'(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x, t)\,\mathrm{d}t.\end{aligned}$$

This section finishes with the definition of convex and concave function.

Definition 2.2.6 (Convex, concave)**.** *A real-valued function f on a convex set X in a real vector space (i.e. for all x and y in X and all t in the interval $[0, 1]$, the point $(1 - t)x + ty$ also belongs to X) is*

- *convex if*

$$f(\lambda x + (1 - \lambda)y) \leqslant \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in X; \forall\lambda \in [0, 1];$$

- *concave if*

$$f(\lambda x + (1 - \lambda)y) \geqslant \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in X; \forall\lambda \in [0, 1];$$

  or when $-f$ is convex.

Example 2.2.8. *The function $f : \mathbb{R} \to \mathbb{R} : x \mapsto |x|^p$ is convex if $p \in [1, \infty)$ and concave if $p \in [0, 1)$. A concave or convex function does not need to be differentiable everywhere.*

2.3 Important (in)equalities

This section contains a list of useful inequalities that are used in the text [14]. For instance, Minkowski's inequality plays a central role in the theory of $\mathrm{L}^p(\Omega)$-spaces, see Section 2.7. This section starts with the following crucial remark.

Remark 2.3.1. *Throughout the dissertation, the values C, ε and C_ε are generic and positive constants independent of any other parameter. The value ε is small and $C_\varepsilon = C + C\varepsilon + C\varepsilon^{-1}$. These constants can be different from place to place. To reduce the number of arbitrary constants, the notation $a \lesssim b$ is used if there exists a positive constant C such that $a \leqslant Cb$.*

Lemma 2.3.1 (Useful inequalities)**.** *For all $a, b \in \mathbb{R}$, the following inequality holds true*

$$(a+b)^2 \leqslant 2(a^2+b^2).$$

For all $a, b \in \mathbb{R}$ and $p \in [1,\infty)$, it holds that

$$|a+b|^p \leqslant 2^{p-1}(|a|^p + |b|^p).$$

Also, for all $a, b \in [0,+\infty)$, it holds that

$$\sqrt{a+b} \leqslant \sqrt{a} + \sqrt{b}.$$

Moreover, for any $a, b, y, z \in [0,+\infty)$, it is true that

$$4ab\left(y^{\frac{a+b}{2}} - z^{\frac{a+b}{2}}\right)^2 \leqslant (a+b)^2(y^a - z^a)(y^b - z^b).$$

Example 2.3.1. *Using the third inequality, it is clear that the mapping $x \mapsto \sqrt{x}$ is Hölder continuous of order $\frac{1}{2}$ on the interval $[0,+\infty)$. Indeed, for $x, y \in [0,+\infty)$ with $x \neq y$, it holds that*

$$\left|\sqrt{x} - \sqrt{y}\right| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \leqslant \frac{\sqrt{|x|+|y|}}{\sqrt{x}+\sqrt{y}}\sqrt{|x-y|} \leqslant \sqrt{|x-y|}.$$

Lemma 2.3.2 (Cauchy-Schwarz inequality for integrals)**.** *Let Ω be an open subset of $\mathbb{R}^d$. Let $f : \Omega \to \mathbb{F}$ and $g : \Omega \to \mathbb{F}$ be two measurable functions that satisfy $\int_\Omega |f(\mathbf{x})|_{\mathrm{c}}^2 \, \mathrm{d}\mathbf{x} < \infty$ and $\int_\Omega |g(\mathbf{x})|_{\mathrm{c}}^2 \, \mathrm{d}\mathbf{x} < \infty$. Then*

$$\left|\int_\Omega f(\mathbf{x})\overline{g(\mathbf{x})} \, \mathrm{d}\mathbf{x}\right|_{\mathrm{c}}^2 \leqslant \left(\int_\Omega |f(\mathbf{x})|_{\mathrm{c}}^2 \, \mathrm{d}\mathbf{x}\right)\left(\int_\Omega |g(\mathbf{x})|_{\mathrm{c}}^2 \, \mathrm{d}\mathbf{x}\right).$$

Lemma 2.3.3 (Cauchy-Schwarz inequality: discrete)**.** *Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ for $n \in \mathbb{N} \cup \{\infty\}$ be any sequences of elements in $\mathbb{F}$ satisfying $\sum_{k=1}^n |a_k|_{\mathrm{c}}^2 < \infty$ and $\sum_{k=1}^n |b_k|_{\mathrm{c}}^2 < \infty$. Then*

$$\left|\sum_{k=1}^n a_k \bar{b}_k\right|_{\mathrm{c}}^2 \leqslant \left(\sum_{k=1}^n |a_k|_{\mathrm{c}}^2\right)\left(\sum_{k=1}^n |b_k|_{\mathrm{c}}^2\right).$$

Lemma 2.3.4 (Hölder inequality for integrals)**.** *Let Ω be an open subset of $\mathbb{R}^d$. Assume that $p \in (1,\infty)$ and $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Let $f : \Omega \to \mathbb{F}$ and $g : \Omega \to \mathbb{F}$ be two measurable functions such that $|f|_{\mathrm{c}}^p$ and $|g|_{\mathrm{c}}^q$ are integrable in $\Omega \subset \mathbb{R}^d$. Then*

$$\int_\Omega |f(\mathbf{x})g(\mathbf{x})|_{\mathrm{c}} \, \mathrm{d}\mathbf{x} \leqslant \left(\int_\Omega |f(\mathbf{x})|_{\mathrm{c}}^p \, \mathrm{d}\mathbf{x}\right)^{\frac{1}{p}}\left(\int_\Omega |g(\mathbf{x})|_{\mathrm{c}}^q \, \mathrm{d}\mathbf{x}\right)^{\frac{1}{q}}.$$

Lemma 2.3.5 (Hölder inequality: discrete)**.** *Assume that* $n \in \mathbb{N} \cup \{\infty\}$*,* $p \in (1, \infty)$ *and* $\dfrac{1}{p} + \dfrac{1}{q} = 1$*. If* $a_1, a_2, \ldots, a_n$ *and* $b_1, b_2, \ldots, b_n$ *are any sequences of elements in* $\mathbb{F}$ *that satisfy* $\sum_{k=1}^{n} |a_k|_{\mathrm{c}}^p < \infty$ *and* $\sum_{k=1}^{n} |b_k|_{\mathrm{c}}^q < \infty$*, then*

$$\sum_{k=1}^{n} |a_k b_k|_{\mathrm{c}} \leqslant \left(\sum_{k=1}^{n} |a_k|_{\mathrm{c}}^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |b_k|_{\mathrm{c}}^q \right)^{\frac{1}{q}}.$$

Lemma 2.3.6 (Minkowski inequality for integrals)**.** *Let* Ω *be an open subset of* $\mathbb{R}^d$*. Assume that* $p \in [1, \infty)$*. Let* $f : \Omega \to \mathbb{F}$ *and* $g : \Omega \to \mathbb{F}$ *be two measurable functions such that* $|f|_{\mathrm{c}}^p$ *and* $|g|_{\mathrm{c}}^p$ *are integrable in* $\Omega \subset \mathbb{R}^d$*. Then*

$$\left(\int_\Omega |f(\mathbf{x}) + g(\mathbf{x})|_{\mathrm{c}}^p \;\mathrm{d}\mathbf{x} \right)^{\frac{1}{p}} \leqslant \left(\int_\Omega |f(\mathbf{x})|_{\mathrm{c}}^p \;\mathrm{d}\mathbf{x} \right)^{\frac{1}{p}} + \left(\int_\Omega |g(\mathbf{x})|_{\mathrm{c}}^p \;\mathrm{d}\mathbf{x} \right)^{\frac{1}{p}}.$$

Lemma 2.3.7 (Minkowski inequality: discrete)**.** *Let* $n \in \mathbb{N} \cup \{\infty\}$ *and* $p \in [1, \infty)$*. If* $a_1, a_2, \ldots, a_n$ *and* $b_1, b_2, \ldots, b_n$ *are two sequences of scalars in* $\mathbb{F}$ *that satisfy* $\sum_{k=1}^{n} |a_k|_{\mathrm{c}}^p < \infty$ *and* $\sum_{k=1}^{n} |b_k|_{\mathrm{c}}^q < \infty$*. Then*

$$\left(\sum_{k=1}^{n} |a_k + b_k|_{\mathrm{c}}^p \right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{n} |a_k|_{\mathrm{c}}^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |b_k|_{\mathrm{c}}^p \right)^{\frac{1}{p}}.$$

Lemma 2.3.8 (Young inequality)**.** *Suppose that* $a, b \in [0, \infty)$*,* $p \in (1, \infty)$ *and* $\dfrac{1}{p} + \dfrac{1}{q} = 1$*. Then*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}.$$

Lemma 2.3.9 (ε-Young inequality)**.** *Assume that* $a, b, \varepsilon \in [0, \infty)$*,* $p \in (1, \infty)$ *and* $\dfrac{1}{p} + \dfrac{1}{q} = 1$*. Then*

$$ab \leqslant \varepsilon a^p + (\varepsilon p)^{-\frac{q}{p}} q^{-1} b^q = \varepsilon a^p + C_\varepsilon b^q.$$

The following two lemmas relate the value of a convex function of an integral to the integral of the convex function. It was proved by Jensen in 1906 [17, 18].

Lemma 2.3.10 (Jensen's inequality: continuous)**.** *Let* Ω *be an open connected set in* $\mathbb{R}^d$ *with finite measure* $|\Omega| = \int_\Omega \;\mathrm{d}\mathbf{x}$*. If* f *is a real-valued integrable function on* Ω *and if* φ *is a convex function on the real line, then*

$$\varphi\left(\frac{1}{|\Omega|} \int_\Omega f(\mathbf{x}) \;\mathrm{d}\mathbf{x} \right) \leqslant \frac{1}{|\Omega|} \int_\Omega \varphi(f(\mathbf{x})) \;\mathrm{d}\mathbf{x}.$$

If $\Omega = (a, b)$ *and* $f : (a, b) \to \mathbb{R}$ *is a non-negative Lebesgue-integrable function, then*

$$\varphi\left(\int_a^b f(x) \,\mathrm{d}x \right) \leqslant \frac{1}{b - a} \int_a^b \varphi((b - a) f(x)) \,\mathrm{d}x.$$

Lemma 2.3.11 (Jensen's inequality: discrete)**.** *For a real-valued convex function φ, real numbers $x_1, x_2 \ldots, x_n$ in its domain ($n \in \mathbb{N}$), and positive weights $a_i \in \mathbb{R}$, Jensen's inequality can be stated as*

$$\varphi\left(\frac{\sum_{i=1}^{n} a_i x_i}{\sum_{j=1}^{n} a_j}\right) \leqslant \frac{\sum_{i=1}^{n} a_i \varphi(x_i)}{\sum_{j=1}^{n} a_j}$$

and the inequality is reversed if φ is concave, which is

$$\varphi\left(\frac{\sum_{i=1}^{n} a_i x_i}{\sum_{j=1}^{n} a_j}\right) \geqslant \frac{\sum_{i=1}^{n} a_i \varphi(x_i)}{\sum_{j=1}^{n} a_j}.$$

The equalities hold iff $x_1 = x_2 = \cdots = x_n$ or if φ is linear.

Example 2.3.2. *For $n \in \mathbb{N}$ and real numbers $x_1, x_2 \ldots, x_n$, it holds that*

$$\left(\sum_{i=1}^{n} x_i\right)^2 \leqslant n \sum_{i=1}^{n} x_i^2.$$

The following integral inequalities are a crucial tool in the study of various classes of equations. The first prototype was proved by Grönwall in 1919 [19].

Lemma 2.3.12. *Let $u : [\alpha, \alpha + h] \to \mathbb{R}$ be a continuous function satisfying the inequality*

$$0 \leqslant u(t) \leqslant \int_{\alpha}^{t} [a + bu(s)] \, \mathrm{d}s, \qquad \text{for } t \in [\alpha, \alpha + h],$$

where a and b are nonnegative constants. Then $u(t) \leqslant ah \exp(bh)$ for $t \in [\alpha, \alpha + h]$.

For more information on Grönwall type integral inequalities, see [20]. The following two lemmas contain the most important inequalities.

Lemma 2.3.13 (Grönwall: continuous [21])**.** *Let $r(t), h(t), y(t)$ be continuous real functions defined on $[a, b]$ satisfying $r(t), h(t) > 0$.*

(i) If

$$y(t) \leqslant h(t) + \int_{a}^{t} r(s) y(s) \, \mathrm{d}s \quad \text{for} \quad a \leqslant t \leqslant b,$$

then

$$y(t) \leqslant h(t) + \int_{a}^{t} h(s) r(s) \exp\left(\int_{s}^{t} r(\tau) \, \mathrm{d}\tau\right) \mathrm{d}s$$

is true for all $t \in [a, b]$.

(ii) If $r(s) = C$ and the function h is non-decreasing then

$$y(t) \leqslant h(t)\exp(C(t-a)) \quad \textit{for} \quad a \leqslant t \leqslant b.$$

Lemma 2.3.14 (Grönwall: discrete [22])**.** *Let $\{A_i\}, \{a_i\}$ be sequences of nonnegative real numbers and $q \geqslant 0$. Assume that*

$$a_i \leqslant A_i + \sum_{j=1}^{i-1} a_j q,$$

for $i \in \mathbb{N}$, then

$$a_i \leqslant A_i + \exp(qi)\sum_{j=1}^{i-1} A_j q, \qquad i \in \mathbb{N}.$$

Remark 2.3.2. *The condition in the discrete Grönwall lemma is often*

$$a_i \leqslant A_i + \sum_{j=1}^{i} a_j q,$$

for $i \in \mathbb{N}$. From this, one can easily derive that

$$a_i \leqslant (1-q)^{-1}\left(A_i + \sum_{j=1}^{i-1} a_j q\right)$$

such that an application of Grönwall's lemma gives that

$$a_i \leqslant \frac{A_i}{1-q} + \frac{\exp\left(\frac{q}{1-q}i\right)}{(1-q)^2}\sum_{j=1}^{i-1} A_j q \qquad \textit{if} \quad q < 1.$$

The following summation rules are analogue to the integration by parts formula [23].

Lemma 2.3.15 (Abel's summation rule and summation by parts formula)**.** *Let $\{a_i\}_{i=1}^n$ be a sequence of real numbers with $n \geqslant 1$. Then*

$$2\sum_{i=1}^{n}(a_i - a_{i-1})a_i = a_n^2 - a_0^2 + \sum_{i=1}^{n}(a_i - a_{i-1})^2.$$

Moreover, for any real sequences $\{z_i\}_{i=0}^n$ and $\{w_i\}_{i=0}^n$, it holds that

$$\sum_{i=1}^{n} z_i\left(w_i - w_{i-1}\right) = z_n w_n - z_0 w_0 - \sum_{i=1}^{n}\left(z_i - z_{i-1}\right) w_{i-1}. \tag{2.1}$$

In the following lemma, the monotonicity and Lipschitz continuity of $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^d : \mathbf{x} \mapsto |\mathbf{x}|_{\mathrm{e}}^{\beta-1}\, \mathbf{x}$ are studied. This mapping is continuous for $\beta \geqslant 0$. The proof of this lemma can be found in Lemma A.1.1 in Appendix A.

Lemma 2.3.16. *For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, it holds that*

(i)

$$|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{y}|_{\mathrm{e}}^2 + \mathbf{x} \cdot \mathbf{y} \geqslant C_* \,|\mathbf{x} - \mathbf{y}|_{\mathrm{e}}^2\,, \qquad C_* \in \left[-\frac{1}{2}, \frac{1}{4}\right],$$

(ii)

$$\left(|\mathbf{x}|_{\mathrm{e}}^{\beta-1}\, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1}\, \mathbf{y}\right) \cdot (\mathbf{x} - \mathbf{y}) \geqslant \frac{1}{4 \cdot 12^{\frac{\beta+1}{2}}}\, |\mathbf{x} - \mathbf{y}|_{\mathrm{e}}^{\beta+1}\,, \quad \beta \in [1, +\infty),$$

(iii)

$$\begin{aligned}\left(|\mathbf{x}|_{\mathrm{e}}^{\beta-1}\, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1}\, \mathbf{y}\right) \cdot \left(|\mathbf{x}|_{\mathrm{e}}^{\alpha-1}\, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\alpha-1}\, \mathbf{y}\right) \\ \geqslant \frac{4\alpha\beta}{(\alpha+\beta)^2} \left(|\mathbf{x}|_{\mathrm{e}}^{\frac{\alpha+\beta}{2}} - |\mathbf{y}|_{\mathrm{e}}^{\frac{\alpha+\beta}{2}}\right)^2 \geqslant 0,\end{aligned}$$

for $\alpha, \beta \in [0, +\infty)$,

(iv) there exists a positive constant C such that

$$\left| |\mathbf{x}|_{\mathrm{e}}^{\beta-1}\, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1}\, \mathbf{y} \right|_{\mathrm{e}} \leqslant C\, |\mathbf{x} - \mathbf{y}|_{\mathrm{e}}^{\beta}\,, \qquad \beta \in (0, 1].$$

2.4 Functional analysis

This section deals with some important aspects in functional analysis, which are necessary for a thorough understanding of the results in this dissertation. No proofs are given. For more details, the reader is referred to e.g. [6, 7, 14, 24]. The usage of the word ‘functional’ goes back to the calculus of variations, implying a function whose argument is again a function.

2.4.1 Normed linear spaces

Definition 2.4.1 (Norm, normed linear space)**.** *A vector space X on which a norm is defined is called a normed vector space or normed linear space. A norm on X is a function*

$$\|\cdot\|_X : X \to \mathbb{R},$$

that satisfies the following three conditions:

- $\|x\|_X > 0$ *for all $x \in X$ if $x \neq 0$, with equality only for $x = 0$;*
- $\|\alpha x\|_X = |\alpha|_{\mathrm{c}} \|x\|_X, \qquad \forall x \in X, \forall \alpha \in \mathbb{F}$*;*
- *the triangle inequality holds: $\|x + y\|_X \leqslant \|x\|_X + \|y\|_X$ for any vectors x and $y \in X$.*

A seminorm, on the other hand, is allowed to assign zero 'length' to some nonzero vectors (in addition to the zero vector). The triangle inequality states that a norm is convex. The reverse triangle inequality is an elementary consequence of the triangle inequality: for all vectors x and $y \in X$ it holds that

$$|\|x\|_X - \|y\|_X| \leqslant \|x-y\|_X.$$

Every normed space is a metric space with the metric $d_X : X \times X \to \mathbb{R}$ given by $d_X(x,y) = \|x-y\|_X$.

Example 2.4.1. *The vector space $\mathbb{F}^d$ becomes a normed vector space when it is equipped with the norm $|\cdot|_\mathrm{e}$. In fact, the mapping $\|\cdot\|_p$ defined for all $x = \{x_i\}_{i=1}^d \in \mathbb{F}^d$ by $\|x\|_p = \left(\sum_{i=1}^d |x_i|_\mathrm{c}^p\right)^{\frac{1}{p}}$ if $p \in [1,\infty)$ and $\|x\|_p = \sup_{1\leqslant i \leqslant d} |x_i|_\mathrm{c}$ if $p = \infty$ are norms on $\mathbb{F}^d$. This follows basically from Minkowski's inequality* (2.3.7). *Note that the space of rational numbers $\mathbb{Q}$ is not a vector space over $\mathbb{R}$ with the usual addition and multiplication.*

Example 2.4.2. *For $p \in [1,\infty]$, let l^p denote the set of all infinite sequences $x = \{x_i\}_{i=1}^\infty$ of scalars $x_i \in \mathbb{F}$ that satisfy*

$$\sum_{i=1}^\infty |x_i|_\mathrm{c}^p < \infty \quad \text{if } 1 \leqslant p < \infty, \quad \text{or} \quad \sup_{i\geqslant 1} |x_i|_\mathrm{c} < \infty \quad \text{if } p = \infty.$$

For each $p \in [1,\infty]$, the set l^p is a vector space, and the mapping $\|\cdot\|_{l^p}$ defined by

$$\begin{aligned} x = \{x_i\}_{i=1}^\infty \in l^p &\mapsto \|x\|_{l^p} = \left(\sum_{i=1}^\infty |x_i|_\mathrm{c}^p\right)^{\frac{1}{p}} && \text{if } 1 \leqslant p < \infty, \\ x = \{x_i\}_{i=1}^\infty \in l^p &\mapsto \|x\|_{l^p} = \sup_{i\geqslant 1} |x_i|_\mathrm{c} && \text{if } p = \infty, \end{aligned}$$

is a norm on l^p.

A normed vector space inherits all the definitions and properties of metric spaces since it is a metric space. The most important ones are given next.

Definition 2.4.2 (Balls, open sets, closed sets, bounded sets)**.** *Let X be a normed linear space, $x_0 \in X$, $r > 0$. The set*

$$B(x_0, r) = \{x \in X \ : \ \|x - x_0\|_X < r\}$$

is called an open ball, with center x_0 and radius r. A subset $M \subset X$ is called an open set in X if for every $x_0 \in M$ there exists an $r = r(x_0) > 0$ such that $B(x_0, r) \subset M$. A subset $M \subset X$ is called a closed set in X if its complement $X \setminus M$ is an open set in X. The closed ball with center x_0 and radius r is denoted by

$$\overline{B}(x_0, r) = \{x \in X \ : \ \|x - x_0\|_X \leqslant r\}.$$

A subset $M \subset X$ is bounded if there exists a constant $R > 0$ such that $\|x\|_X \leqslant R$ for all $x \in M$.

Definition 2.4.3 (Convergences and Cauchy sequences)**.** *Let X be a normed vector space with norm $\|\cdot\|_X$. A sequence $x_1, x_2, x_3, \ldots$ in X converges to $x \in X$ iff $\|x_n - x\|_X \to 0$ as $n \to \infty$. A sequence $x_1, x_2, x_3, \ldots$ is called a fundamental or a Cauchy sequence, if for every positive real number $\varepsilon > 0$ there is a positive integer N such that for all positive integers $m, n > N$, it holds that $\|x_m - x_n\|_X < \varepsilon$.*

Every convergent sequence is a Cauchy sequence. The converse is not generally true.

Lemma 2.4.1. *Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in a normed vector space X and let $\{x_{n_k}\}_{k=1}^{\infty}$ be a subsequence of it. If $x_{n_k} \to x$ in X, then $x_n \to x$ in X.*

Definition 2.4.4 (Closure)**.** *Let M be a subset of a normed linear space X. The closure of M in X is denoted by $\overline{M}$ and is defined as the set of all elements $x \in X$ such that there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in M with the property that*

$$x_n \to x \quad \textit{in} \quad X.$$

Clearly, $M \subset \overline{M} \subset X$. The set M is closed in X if $M = \overline{M}$.

Definition 2.4.5 (Density)**.** *A subset $M \subset X$ is said to be dense in X if $\overline{M} = X$. This means that for every $x \in X$ and every $\varepsilon > 0$, there exists a $y \in M$ such that $\|x - y\|_X < \varepsilon$.*

Definition 2.4.6 (Separability)**.** *A normed linear space X is called separable if there exists a countable set $M \subset X$ that is dense in X, i.e. there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of elements of the space such that every nonempty open subset of the space contains at least one element of the sequence.*

Example 2.4.3. *Any normed linear space that is itself finite or countably infinite is separable because the whole space is a countable dense subset of itself. An important example of an uncountable separable space is the real line, in which the rational numbers form a countable dense subset. Similarly, the set of all vectors $(r_1, \ldots, r_d) \in \mathbb{R}^d$ in which r_i is rational for all i is a countable dense subset of $\mathbb{R}^d$; therefore, the d-dimensional Euclidean space is separable.*

Example 2.4.4. *The space l^p defined in Definition 2.4.2 is separable for $p \in [1, \infty)$. Any dense subset of l^∞ is uncountably infinite [14, Theorem 2.4-2].*

The following definition is about the difference between a subset and a subspace.

Definition 2.4.7 (Subspace)**.** *A subset of a normed linear space X is called a subspace of X if it is a linear set which is closed in X.*

Remark 2.4.1. *A linear subset does not need to be closed in X.*

Lemma 2.4.2.

(i) A subspace M of a normed linear space $(X, \|\cdot\|_X)$ is again a normed linear space with the norm $\|\cdot\|_M$ defined by

$$\|u\|_M = \|u\|_X \qquad \textit{for} \quad u \in M;$$

(ii) A subspace of a separable normed linear space X is also a separable normed linear space.

The following lemma is a normed-space substitute for one aspect of orthogonality in Hilbert spaces, see [25, p. 218]. This aspect can be found in Corollary 2.4.1.

Lemma 2.4.3 (Riesz)**.** *Let L be a subspace of a normed space (X, d_X), where $L \neq X$ (L is not dense in X). For any $\varepsilon \in (0,1)$, there exists a $z_\varepsilon \in X \setminus L$ with $\|z_\varepsilon\| = 1$ such that*

$$d_X(z_\varepsilon, L) \geqslant 1 - \varepsilon.$$

Definition 2.4.8 (Cartesian product)**.** *Let $n \in \mathbb{N}$ and let $X_1, X_2, \ldots, X_n$ be normed linear spaces (over the same field $\mathbb{F}$). The Cartesian product $X = X_1 \times X_2 \times \ldots \times X_n$ of $X_1, X_2, \ldots, X_n$ is the set of all ordered n-tuples $x = (x_1, \ldots, x_n)$ such that $x_i \in X_i$, $i = 1, \ldots, n$. Then X is also a normed linear space when X is equipped with any one of the following norms*

$$\|x\|_X = \left(\sum_{i=1}^{n} \|x_i\|_{X_i}^p \right)^{\frac{1}{p}}, \quad 1 \leqslant p < \infty, \quad \textit{or} \quad \|x\|_X = \max_{i=1,\ldots,n} \|x_i\|_{X_i}.$$

Lemma 2.4.4. *Let $X_1, X_2, \ldots, X_n$ be separable normed linear spaces. Then the product space $X = X_1 \times X_2 \times \ldots \times X_n$ is also a separable normed linear space.*

Definition 2.4.9 (Compact, relatively compact)**.** *A subset M in a normed linear space X is said to be compact if every sequence in M contains a subsequence converging to an element of M. A set M in a normed linear space X is said to be relatively compact if its closure $\overline{M}$ is compact.*

Remark 2.4.2. *If M is closed and relatively compact then M is compact. Compact sets are closed and bounded, but not vice versa, in general.*

Lemma 2.4.5. *A closed subset of a compact normed space is a compact normed space.*

2.4.2 Inner product spaces

Definition 2.4.10 (Inner product, inner product space)**.** *An inner product space is a vector space X together with an inner product, i.e., with a map $\langle \cdot, \cdot \rangle_X : X \times X \to \mathbb{F}$ that satisfies the following three properties for all vectors $x, y, z \in X$ and all scalars $a \in \mathbb{F}$*

- *Conjugate symmetry:* $\langle x, y\rangle_X = \overline{\langle y, x\rangle}_X$. *Note that if* $\mathbb{F} = \mathbb{R}$, *this means symmetry;*

- *Linearity in the first argument:* $\langle ax, y\rangle_X = a\langle x, y\rangle_X$, $\langle x + y, z\rangle_X = \langle x, z\rangle_X + \langle y, z\rangle_X$;

- *Positive-definiteness:* $\langle x, x\rangle_X \geqslant 0$ *with equality only for* $x = 0$.

The inner product $\langle\cdot,\cdot\rangle_X$ induces a norm which is given by $\|\cdot\|_X = \sqrt{\langle\cdot,\cdot\rangle}_X$. Therefore, every inner product space is a normed space.

Example 2.4.5. *The space* $\mathbb{F}^d$ *equipped with the Euclidean inner product is the simplest example of a real or complex inner product space.*

Example 2.4.6. *The vector space consisting of all real, resp. complex,* $m \times n$ *matrices, equipped with the matrix inner product defined by*

$$\begin{cases} \mathbf{A} : \mathbf{B} := \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ij} & \text{if } \mathbb{F} = \mathbb{R}, \text{ resp.} \\ \mathbf{A} : \mathbf{B} := \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} \overline{b_{ij}} & \text{if } \mathbb{F} = \mathbb{C}, \end{cases}$$

for all $m \times n$ *matrices* $\mathbf{A} = (a_{ij})$ *and* $\mathbf{B} = (b_{ij})$ *is an inner product space. The norm* $\|\cdot\|_F$ *induced by this inner product, thus defined by*

$$\|\mathbf{A}\|_F := (\mathbf{A} : \mathbf{A})^{\frac{1}{2}} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|_{\mathrm{c}}^2\right)^{\frac{1}{2}}$$

for any $m \times n$ *matrix* $\mathbf{A} = (a_{ij})$ *is called the Frobenius norm.*

Example 2.4.7. *The real or complex space* l^2 *defined in Definition 2.4.2 is an inner product space when equipped with the inner product defined by*

$$\begin{cases} (x, y)_{l^2} := \displaystyle\sum_{i=1}^{\infty} x_i y_i & \text{for all } x = \{x_i\}_{i=1}^{\infty}, y = \{y_i\}_{i=1}^{\infty} \in l^2 \text{ if } \mathbb{F} = \mathbb{R}, \\ (x, y)_{l^2} := \displaystyle\sum_{i=1}^{\infty} x_i \overline{y}_i & \text{for all } x = \{x_i\}_{i=1}^{\infty}, y = \{y_i\}_{i=1}^{\infty} \in l^2 \text{ if } \mathbb{F} = \mathbb{C}. \end{cases}$$

The associated norm is defined by

$$\|x\|_{l^2} := \sqrt{(x, x)_{l^2}} = \left(\sum_{i=1}^{\infty} |x_i|_{\mathrm{c}}^2\right)^{\frac{1}{2}} \quad \text{for all } x = \{x_i\}_{i=1}^{\infty} \in l^2.$$

2.4.3 Banach and Hilbert spaces

Banach and Hilbert spaces play a central role in linear and nonlinear functional analysis. In this subsection, some definitions and basic properties are listed.

Definition 2.4.11 (Completeness)**.** *A normed linear space X is said to be complete if every Cauchy sequence in X is convergent, i.e. it has a limit in X.*

Definition 2.4.12 (Banach space)**.** *A complete normed linear space is called a Banach space.*

Example 2.4.8. *The space $\mathbb{F}^d$ is a Banach space when associated with the Euclidean or Hermitian norm respectively. Also the vector space consisting of all real, resp. complex, $m \times n$ matrices is a Banach space. In fact, every finite dimensional normed vector space is a Banach space.*

Example 2.4.9. *The spaces $(l^p, \|\cdot\|_{l^p})^{\frac{1}{p}}$, $p \in [1, \infty]$, defined in Definition 2.4.2 are Banach spaces.*

Example 2.4.10. *The set of continuous functions on the unit interval $[0, 1]$ with norm $\displaystyle\int_0^1 |f(x)|\,\mathrm{d}x$ is not a Banach space because it is not complete. For instance, the Cauchy sequence of functions*

$$f_n = \begin{cases} 1 & \textit{for } x \leqslant \frac{1}{2} \\ \left(\frac{1}{2} - x\right) n + 1 & \textit{for } \frac{1}{2} < x \leqslant \frac{1}{2} + \frac{1}{n} \\ 0 & \textit{for } x > \frac{1}{2} + \frac{1}{n} \end{cases}$$

does not converge to a continuous function.

Lemma 2.4.6. *Each Banach space is closed. A subspace of a Banach space is also a Banach space.*

Lemma 2.4.7.

(i) Let $X_1, X_2, \ldots, X_n$ be Banach spaces. Then the product space $X = X_1 \times X_2 \times \ldots \times X_n$ is also a Banach space.

(ii) Let M_i be a relatively compact subset of a Banach space X_i, $i = 1, \ldots, n$. Then $M_1 \times \ldots \times M_n$ is a relatively compact subset of $X_1 \times \ldots \times X_n$.

Definition 2.4.13 (Hilbert space)**.** *A Hilbert space is an inner product space which is complete with respect to the induced norm.*

Example 2.4.11. *The space $\mathbb{F}^d$ is a Hilbert space, with the classical associated dot product. Also the vector space consisting of all real, resp. complex, $m \times n$ matrices is a Hilbert space.*

Example 2.4.12. *The space l^p defined in Definition 2.4.2 is a Hilbert space only for $p = 2$ (see Example 2.4.7).*

Each Hilbert space is a Banach space. The latest example shows that the converse is not generally true. The orthogonal complement of a subset of a Hilbert space can be defined. A property of a closed subset (i.e. subspace) in a Hilbert space is given in the direct sum theorem.

Definition 2.4.14 (Orthogonal, orthogonal complement)**.** *Let $(X, (\cdot,\cdot)_X)$ be a real or complex inner-product space. Two vectors x and $y \in X$ are said to be orthogonal if $(x, y)_X = 0$, and the orthogonal complement of any nonempty subset U of X is defined as*

$$U^\perp := \{x \in X : (x, u)_X = 0, \forall u \in U\}.$$

Lemma 2.4.8. *Let M be a nonempty subset of an inner-product space X. Then the set $M^\perp$ is a subspace of X. Besides, $\left(\overline{M}\right)^\perp = M^\perp$, and $M \cap M^\perp = \{0\}$ if $0 \in M$ and $M \cap M^\perp = \varnothing$ if $0 \notin M$.*

Theorem 2.4.13 (Direct sum theorem)**.** *Let U be a subspace of a Hilbert space X. Then the space X is the direct sum $X = U \oplus U^\perp$, i.e. for each $x \in X$, there exists unique elements $u \in U$ and $v \in U^\perp$ such that $x = u + v$.*

This implies the following analogue of Riesz's lemma 2.4.3.

Corollary 2.4.1. *In a Hilbert space Y, given a non-dense subspace X, there is a $y \in Y$ with $\inf_{x\in X} \|x - y\|_Y = 1$, by taking y from the orthogonal complement of X.*

The definition of separability for a Hilbert space can be expressed via an orthogonal sequence.

Definition 2.4.15 (Orthogonal sequence)**.** *An orthogonal sequence (or orthogonal system) $\{x_n\}_{n\in\mathbb{N}}$ in an inner product space $(X, (\cdot,\cdot))$ (finite or infinite dimensional) is one in which $(x_n, x_m) = 0$ whenever $n \neq m$. An orthonormal sequence (or orthonormal system) $\{x_n\}_{n\in\mathbb{N}}$ in an inner product space $(X, (\cdot,\cdot))$ is an orthogonal sequence with $\|x_n\|_X = 1$ for all $n \in \mathbb{N}$. An orthonormal sequence $\{x_n\}_{n\in\mathbb{N}}$ in an inner product space $(X, (\cdot,\cdot))$ is called complete if for each $x \in X$ it holds that*

$$x = \sum_{n=1}^{\infty} (x, x_n) x_n.$$

Lemma 2.4.9. *A Hilbert space is separable iff it has a complete orthonormal sequence.*

Example 2.4.14. *The Hilbert space $\mathbb{F}^d$ and the Hilbert space consisting of all real, resp. complex, $m \times n$ matrices are separable (since they are finite-dimensional). The space l^2 provides an example of an infinite-dimensional separable Hilbert space.*

2.4.4 Operators

This subsection reviews linear and nonlinear operators in normed vector spaces. In what follows X and Y are two vector spaces over the same field $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

Definition 2.4.16 (Operator, domain, range, kernel)**.** *Let X and Y be linear spaces, M a set in X. An operator A on M is a mapping from M to Y. For every $u \in M$, the corresponding element in Y is denoted by Au or $A(u)$. The set M is called the domain of the operator A and is denoted by $\mathcal{D}(A)$, the set*

$$\mathcal{R}(A) = \{y \in Y : y = Ax \quad \textit{with} \quad x \in M\}$$

is called the range or direct image of the operator A. The kernel of A is the subset of X defined by $\ker(A) = \{x \in \mathcal{D}(A) : Ax = 0\}$.

Definition 2.4.17 (Injective, surjective, bijective)**.** *Let X and Y be linear spaces. An operator A from X into Y is called surjective (or onto) if for every $y \in \mathcal{R}(A)$, there exists at least one element $x \in \mathcal{D}(A)$ such that $y = A(x)$. An operator A from X into Y is called injective if for all x_1 and x_2 in $\mathcal{D}(A)$ such that $A(x_1) = A(x_2)$, it holds that $x_1 = x_2$. An operator A from X into Y is called bijective if it is both surjective and injective.*

Remark 2.4.3. *The operator A from Definition 2.4.16 is an operator from X to Y that maps M into Y. If $\mathcal{R}(A) = Y$, the operator A maps M onto Y.*

Definition 2.4.18 (Restriction, extension)**.** *Let X and Y be linear spaces, $A : X = \mathcal{D}(A) \to Y$ and M a subset of X. The operator $A|_M : M \to Y$ defined by $A|_M(x) = A(x)$ for all $x \in M$ is the restriction of A to M. Let $B : M = \mathcal{D}(B) \to Y$ be an operator, where M is a subset of X. An operator $A : X \to Y$ is an extension of B if $A|_M = B$.*

Definition 2.4.19 (Linear operator, semilinear operator, nonlinear operator)**.** *Let X and Y be linear spaces. An operator A from X into Y is called a linear operator if $\mathcal{D}(A)$ is a linear set and if for any $x, y \in \mathcal{D}(A)$ and any scalar $\alpha \in \mathbb{F}$ holds that*

$$A(x + \alpha y) = A(x) + \alpha A(y).$$

If $\mathbb{F} = \mathbb{C}$, a mapping $A : X \to Y$ is semilinear (also called antilinear) if

$$A(x + \alpha y) = A(x) + \overline{\alpha} A(y) \quad \textit{for all } x, y \in \mathcal{D}(A) \textit{ and } \alpha \in \mathbb{F}.$$

An operator that is neither linear nor semilinear is called nonlinear.

Remark 2.4.4. *For a linear and semilinear operator $A : X \to Y$, it holds that $A(0) = 0$. Moreover,* $\ker(A)$ *is a subspace of X and $\mathcal{R}(A)$ is a subspace of Y.*

Definition 2.4.20 (Eigenvalue, eigenvector, eigenspace)**.** *Let X be a linear space. Let $A : X \to X$ be a linear operator. Then a scalar $\lambda \in \mathbb{F}$ is an eigenvalue of A if there exists a vector $x \in X$ such that*

$$Ax = \lambda x \quad \textit{and} \quad x \neq 0.$$

Such a nonzero vector is then called an eigenvector of A corresponding to the eigenvalue λ, and the nontrivial subspace

$$\{x \in X \,:\, Ax = \lambda x\}$$

of X is called the eigenspace corresponding to the eigenvalue λ.

Definition 2.4.21 (Continuous operator)**.** *Let X and Y be normed linear spaces. An operator A from X into Y is continuous in $x \in \mathcal{D}(A)$ if $x_n \to x$ in X implies that $A(x_n) \to A(x)$ in Y for any sequence $\{x_n\}_{n=1}^{\infty} \in \mathcal{D}(A)$. An operator A is continuous in $\mathcal{D}(A)$ if A is continuous in any point of $\mathcal{D}(A)$.*

Definition 2.4.22 (Bounded operator)**.** *Let X and Y be normed linear spaces. An operator $A : X \to Y$ is called bounded if it maps any bounded set in X into a bounded set in Y. Equivalently, an operator $A : X \to Y$ is bounded if there exists a positive constant C such that $\|Ax\|_Y \leqslant C\|x\|_X$ for any $x \in \mathcal{D}(A)$.*

Lemma 2.4.10. *Let X and Y be normed linear spaces. A linear operator $A : X \to Y$ is bounded iff it is continuous.*

Lemma 2.4.11. *Let X and Y be normed linear spaces. If X is finite-dimensional, then any linear operator from X into Y is continuous.*

Remark 2.4.5. *A continuous linear operator $A : \mathcal{D}(A) \subset X \to Y$ is uniformly continuous. This follows from the relation*

$$\|Ax - A\tilde{x}\| \leqslant C\,\|x - \tilde{x}\| \quad \text{for all } x, \tilde{x} \in \mathcal{D}(A).$$

Definition 2.4.23 (The space $\mathcal{L}(X,Y)$)**.** *Let X and Y be normed linear spaces. The space of all linear bounded operators from X into Y is a linear space (over the same field as X), which is denoted by $\mathcal{L}(X,Y)$, when equipped with the following addition and scalar multiplication*

$$\begin{aligned}(A_1 + A_2)(x) &= A_1(x) + A_2(x),\\ (aA_1)(x) &= aA_1(x),\end{aligned}$$

for all $A_1, A_2 : X \to Y$, $x \in X$ and $a \in \mathbb{F}$. The norm of the operator $A \in \mathcal{L}(X,Y)$ is defined as follows

$$\|A\|_{\mathcal{L}(X,Y)} = \sup_{\substack{x\in\mathcal{D}(A)\\ \|x\|_X\leqslant 1}} \|Ax\|_Y = \sup_{\substack{x\in\mathcal{D}(A)\\ \|x\|_X = 1}} \|Ax\|_Y = \sup_{\substack{x\in\mathcal{D}(A)\\ x\neq 0}} \frac{\|Ax\|_Y}{\|x\|_X}.$$

Theorem 2.4.15. *Let X be a normed space and Y be a Banach space. Then $\mathcal{L}(X,Y)$ is a Banach space.*

Remark 2.4.6. *In the proof of this theorem (see for instance [26, Theorem 6.6]), the following consequence is hidden: let $\{A_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence (a convergent sequence is a Cauchy sequence) in $\mathcal{L}(X,Y)$, then $\{\|A_n\|\}_{n\in\mathbb{N}}$ is uniformly bounded, i.e.*

$$\|A_n\|_{\mathcal{L}(X,Y)} \leqslant C, \quad \forall n \in \mathbb{N}.$$

This is an analogue of the next result in $\mathbb{R}^d$: every convergent sequence in $\mathbb{R}^d$ is bounded.

In the following theorem, the properties of X and Y are switched.

Theorem 2.4.16 (Uniform boundedness principle)**.** *Let X be a Banach space and Y be a normed space. Consider a sequence of operators $A_n \in \mathcal{L}(X,Y)$ with the property that for any $x \in X$ there exists a $C_x > 0$ such that $\sup_{n\in\mathbb{N}} \|A_n x\|_Y \leqslant C_x$ (i.e. $\{A_n x\}_{n\in\mathbb{N}}$ is bounded). Then the sequence $\{\|A_n\|\}_{n\in\mathbb{N}}$ is bounded.*

The Banach-Steinhaus theorem is about strong pointwise convergence in the space $\mathcal{L}(X,Y)$ when X is a Banach space and Y is a normed space. If Y is a Banach space and X is a normed space, then any Cauchy sequence converges pointwise strongly, see Theorem 2.4.15.

Theorem 2.4.17 (Banach-Steinhaus)**.** *Let X be a Banach space and Y be a normed space. Consider a sequence of operators $A_n \in \mathcal{L}(X,Y)$ with $\mathcal{D}(A_n) = X$, $n \in \mathbb{N}$. Then $A_n x \to Ax$ for every $x \in X$ with $A \in \mathcal{L}(X,Y)$ iff*
(i) the sequence $\{\|A_n\|\}_{n\in\mathbb{N}}$ is bounded;
(ii) $A_n x \to Ax$ for every $x \in V$, with V a dense subset of X.

Definition 2.4.24 (Bijective operator, inverse operator)**.** *Let X and Y be normed linear spaces. An operator $A : X \to Y$ is called bijective (or one-to-one) if for any $y \in \mathcal{R}(A)$ there exists a unique $x \in \mathcal{D}(A)$ such that $A(x) = y$. An inverse operator $A^{-1} : Y \to X$ is then defined as follows: $A^{-1} : Y \to X$ and*

$$\mathcal{D}(A^{-1}) = \mathcal{R}(A), \qquad \mathcal{R}(A^{-1}) = \mathcal{D}(A).$$

Lemma 2.4.12. *The operator $A : \mathcal{D}(A) \to \mathcal{R}(A)$ is one-to-one iff $\ker(A) = \{0\}$.*

Theorem 2.4.18. *There exists a continuous inverse operator $A^{-1} : \mathcal{R}(A) \to \mathcal{D}(A)$ iff there exists a positive constant C_0 such that*

$$\|Ax\|_Y \geqslant C_0 \|x\|_X, \qquad \forall x \in \mathcal{D}(A).$$

Theorem 2.4.19 (On continuity of the inverse operator (Banach Theorem))**.** *Let X and Y be Banach spaces and assume that $A \in \mathcal{L}(X,Y)$ with $\mathcal{D}(A) = X$. Then $A^{-1} \in \mathcal{L}(Y,X)$.*

Definition 2.4.25 (Identity operator)**.** *Let X and Y be two normed linear spaces and let*

$$X \subset Y.$$

The identity operator $I : X \to Y$ with $\mathcal{D}(I) = \mathcal{R}(I) = X$ is defined as the operator which maps every element $x \in X$ onto itself: $Ix = x$, regarded as an element of Y.

Remark 2.4.7. *The identity operator is linear and is an injection.*

Definition 2.4.26 (Compact operators)**.** *Let X and Y be normed linear spaces and assume that $A : X \to Y$ is a linear operator with $\mathcal{D}(A) = X$. The operator A is said to be compact*

- *if the range of a closed unit ball in X is a compact set in Y;*
- *or equivalently if it maps every bounded set in X onto a relatively compact set in Y;*
- *or equivalently if for any bounded sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$, the sequence $\{Ax_n\}_{n\in\mathbb{N}} \subset Y$ contains a convergent subsequence in Y.*

The set of all compact operators from X to Y is denoted by $\sigma(X, Y)$.

Remark 2.4.8. *The identity operator $I : X \to X$ is a compact operator if X is finite dimensional. In that case, X is isomorphic (see Definition 2.4.28) with some $\mathbb{R}^n$, $n \in \mathbb{N}$. By the Heine-Borel theorem, a set is compact in $\mathbb{R}^n$ if it is bounded and closed. A direct consequence of Riesz's lemma 2.4.3 is that this is not true in infinite dimensional spaces. Indeed, Riesz's lemma implies that the closed unit ball is not compact in an infinite dimensional space.*

Theorem 2.4.20 (Riesz)**.** *A linear normed space X is locally compact (every element of X has a compact neighbourhood) iff it is finite dimensional.*

Lemma 2.4.13. *$\sigma(X, Y)$ is a subspace of $\mathcal{L}(X, Y)$, i.e. every compact operator is continuous.*

Lemma 2.4.14. *Let X and Y be normed linear spaces and assume that $A : X \to Y$ is a linear operator. If X is finite-dimensional, then A is compact.*

Lemma 2.4.15. *Let X and Y be normed linear spaces. If X or Y is a finite dimensional space, then $\sigma(X, Y) = \mathcal{L}(X, Y)$.*

The definition of compact operator (see Definition (2.4.26)) is also valid for non-linear operators. If, in addition, the operator is continuous, then the operator is said to be completely continuous.

Definition 2.4.27 (Completely continuous operator)**.** *Let X and Y be normed linear spaces. The operator $A : X \to Y$ is called completely continuous if it is compact and continuous.*

Remark 2.4.9. *A linear compact operator is automatically completely continuous.*

2.4.5 Isomorphisms and embeddings

Definition 2.4.28 (Isomorphism)**.** *Two normed linear spaces X and Y (both real or both complex) are said to be isomorphic (notation: $X \cong Y$) if there exists a continuous linear operator A such that $\mathcal{D}(A) = X$, $\mathcal{R}(A) = Y$, and A^{-1} exists and is continuous. This operator A is called an isomorphism mapping or briefly an isomorphism between X and Y.*

Definition 2.4.29 (Isometry). *Two normed linear spaces X and Y are said to be isometrically isomorphic if there exists a linear operator A such that $\mathcal{D}(A) = X$, $\mathcal{R}(A) = Y$ and*

$$\|u - v\|_X = \|Au - Av\|_Y$$

for every pair $u, v \in X$.

Remark 2.4.10. *Two isometrically isomorphic spaces are isomorphic.*

Lemma 2.4.16. *Let X and Y be two isomorphic normed linear spaces.*
(i) If X is separable then Y is also separable.
(ii) If X is complete then Y is also complete.

Definition 2.4.30 (Embedding operator). *If the identity operator from Definition 2.4.25 is continuous, then it is called the embedding operator (or shortly embedding) from X to Y. An embedding from X into Y is denoted by*

$$X \hookrightarrow Y.$$

Continuity of an embedding implies that there exists a constant $C > 0$ such that

$$\|x\|_Y \leqslant C\|x\|_X \quad \textit{for every} \quad x \in X.$$

If simultaneously

$$X \hookrightarrow Y \qquad \textit{and} \qquad Y \hookrightarrow X,$$

then this is written as

$$X \rightleftarrows Y.$$

Definition 2.4.31 (Compact embedding). *If the embedding operator in Definition 2.4.30 is compact, i.e. each bounded sequence in X has a convergent subsequence in Y, then X is compactly embedded in Y. This is denoted by*

$$X \hookrightarrow\hookrightarrow Y.$$

Definition 2.4.32 (Equivalent norms). *Let X be a vector space and suppose that $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on X. These norms are equivalent if there exist constants $C_1 > 0$ and $C_2 > 0$ such that*

$$C_1\|x\|_1 \leqslant \|x\|_2 \leqslant C_2\|x\|_1 \qquad \textit{for all} \quad x \in X.$$

Remark 2.4.11. *The norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent iff $(X, \|\cdot\|_1) \rightleftarrows (X, \|\cdot\|_2)$. In particular, the embedding from $(X, \|\cdot\|_1)$ into $(X, \|\cdot\|_2)$ is an isomorphism.*

2.4.6 Continuous linear functionals and dual spaces

Definition 2.4.33 (Linear functionals)**.** *Let X be a real (or complex) normed linear space, Y the real line $\mathbb{R}$ (or the complex plane $\mathbb{C}$). A linear operator from X into Y is called a real (complex) linear functional.*

Remark 2.4.12. *From now on, the adjective 'real' or 'complex' is omitted. The value of a functional $f : X \to \mathbb{F}$ at $x \in X$ is usually denoted by $f(x)$ or $\langle f, x\rangle$. Since a functional is an operator, some concepts introduced in connection with operators can be transfered to functionals. For instance, a norm of a linear functional f is defined by*

$$\|f\|_* = \sup_{\substack{x\in\mathcal{D}(f)\\ \|x\|_X\leqslant 1}} |f(x)|_c. \tag{2.2}$$

From now on, only continuous linear functionals are considered. These are elements of the space $\mathcal{L}(X, \mathbb{F})$.

Continuous linear functionals defined on linear subsets of X can be extended to the entire space X.

Theorem 2.4.21. *Let f be a continuous linear functional defined on a linear set M which is dense in X. Then there exists a uniquely determined continuous linear functional $\tilde{f}$ defined on X such that*

$$\tilde{f}(u) = f(u) \qquad \textit{for all} \quad u \in M.$$

Theorem 2.4.22 (Hahn-Banach theorem)**.** *Let f be a continuous linear functional defined on a linear subset M of a normed linear space X. Then there exists a continuous linear functional $\tilde{f}$ defined on X such that*

$$\tilde{f}(u) = f(u) \qquad \textit{for all} \quad u \in M$$

and

$$\|\tilde{f}\|_{X^*} = \sup_{\substack{x\in X\\ \|x\|_X\leqslant 1}} \langle \tilde{f}, x\rangle = \sup_{\substack{x\in M\\ \|x\|_X\leqslant 1}} \langle f, x\rangle = \|f\|_{M^*}.$$

The set of all linear bounded functionals defined on a normed space X is a Banach space $\mathcal{L}(X, \mathbb{F})$, see Theorem 2.4.15. This leads to the definition of a dual space.

Definition 2.4.34 (Dual space)**.** *Let X be a linear normed space. The Banach space $\mathcal{L}(X, \mathbb{F})$ is called the dual space of X (or briefly the dual of X) and is denoted by X^*. The norm in this space is denoted by $\|\cdot\|_{X^*}$ or $\|\cdot\|_*$.*

Remark 2.4.13. *The expressions 'adjoint space', 'conjugate space' and the notation X' are also used in the literature instead of 'dual space' and X^*. The value of a functional $f : X \to \mathbb{F}$ at $x \in X$ can also be denoted as $\langle f, x\rangle_{X^*\times X}$. A sequence $\{f_n\}_{n=1}^{\infty}$ in X^* converges strongly to $f \in X^*$ iff $\|f_n - f\|_* \to 0$ as $n \to \infty$.*

Dual spaces play a central role in linear functional analysis. Now, a basic example of a dual space and an important theorem are presented.

Example 2.4.23 (Dual of l^p)**.** *Given a real number $p \in [1, \infty)$, let $q \in (1, \infty]$ denote its conjugate exponent, i.e. $q = \frac{p}{p-1}$. Then, given any element $a = \{a_i\}_{i=1}^{\infty} \in l^q$, the relation*

$$x^*(x) = \sum_{i=1}^{\infty} a_i x_i \quad \textit{for all } x = \{x_i\}_{i=1}^{\infty} \in l^p$$

defines a continuous linear functional x^* on l^p. Also $\|x^*\|_{(l^p)^*} = \|a\|_{l^q}$. The linear isometry $a \in l^q \to x^* \in (l^p)^*$ defined in this fashion is bijective, i.e. given any continuous linear functional x^* on l^p, there exists exactly one element $a = \{a_i\}_{i=1}^{\infty} \in l^q$ such that $x^*(x) = \sum_{i=1}^{\infty} a_i x_i$ for all $x = \{x_i\}_{i=1}^{\infty} \in l^p$. Consequently, for any $p \in [1, \infty)$, the dual space of l^p can be identified with the space l^q.

Theorem 2.4.24. *Consider two normed linear spaces X and Y. It holds that*

- *$X \subset Y$ implies that $Y^* \subset X^*$;*
- *If the dual space X^* is separable, then X itself is separable [27, p. 245].*

One of the basic advantages of a Hilbert space is that it can be identified with its dual space by means of a specific linear isometry. This is the content of the fundamental Riesz' representation theorem.

Theorem 2.4.25 (Riesz' representation theorem)**.** *Let H be a Hilbert space with the inner product $(\cdot, \cdot)$. Each linear bounded functional f on H can be written as*

$$\langle f, v \rangle = (v, u), \qquad \textit{for all } v \in H.$$

The element $u \in H$ is uniquely determined. Moreover, $\|f\|_{H^*} = \|u\|_H$. The isometry

$$\sigma : l \in H^* \to \sigma(l) = u \in H$$

is a bijection, which is linear if $\mathbb{F} = \mathbb{R}$ and semilinear if $\mathbb{F} = \mathbb{C}$. Consequently, any Hilbert space can be identified with its dual space H^* by means of this isometry. Moreover, the dual space H^* becomes a Hilbert space when it is equipped with the inner product $(\cdot, \cdot)_{H^*} : H^* \times H^* \to \mathbb{F}$ defined by

$$(x^*, y^*)_{H^*} := \overline{(\sigma(x^*), \sigma(y^*))} \quad \textit{for each } x^*, y^* \in H^*.$$

Let us also consider the definition of dual and adjoint operator [14, Theorem 5.11-1].

Definition 2.4.35 (Dual operator)**.** *Let X and Y be normed linear spaces and assume that $A \in \mathcal{L}(X, Y)$. The dual operator $A^* \in \mathcal{L}(Y^*, X^*)$ is defined as follows*

$$\langle f, Ax \rangle_{Y^* \times Y} = \langle A^* f, x \rangle_{X^* \times X}, \qquad \textit{for all } x \in X, f \in Y^*.$$

Theorem 2.4.26. *Let X and Y be normed linear spaces. Given any operator $A \in \mathcal{L}(X,Y)$, there exists exactly one dual operator $A^* \in \mathcal{L}(Y^*, X^*)$. Moreover, it holds that $\|A\|_{\mathcal{L}(X,Y)} = \|A^*\|_{\mathcal{L}(Y^*,X^*)}$.*

Definition 2.4.36 (Adjoint operator)**.** *Let X and Y be two Hilbert spaces, then $A^* \in \mathcal{L}(Y, X)$ is called the adjoint operator and is uniquely defined as follows*

$$(f, Ax)_Y = (A^* f, x)_X\,, \qquad \text{for all } x \in X, f \in Y.$$

Also, $\|A\|_{\mathcal{L}(X,Y)} = \|A^*\|_{\mathcal{L}(Y,X)}$.

Example 2.4.27. *The $n \times m$ transpose matrix $\mathbf{A}^\mathsf{T}$ of any real matrix $\mathbf{A} = (a_{ij})$, which is defined by $(\mathbf{A}^\mathsf{T})_{ij} = a_{ji}$ satisfies*

$$\mathbf{y} \cdot \mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{y} \cdot \mathbf{x} \qquad \text{for all } \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m.$$

For a complex $m \times n$ matrix $\mathbf{A} = (a_{ij})$, the $n \times m$ adjoint matrix $\mathbf{A}^* = \overline{\mathbf{A}}^\mathsf{T}$ can also be defined as the unique $n \times m$ complex matrix that satisfies

$$\mathbf{y} \cdot \mathbf{A}\mathbf{x} = \mathbf{A}^*\mathbf{y} \cdot \mathbf{x} \qquad \text{for all } \mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m.$$

Definition 2.4.37. *Let X be a complex Hilbert space. The operator $A : X \to X$ is said to be self-adjoint if $A = A^*$.*

Remark 2.4.14. *If X is a Hilbert space, then any self-adjoint operator from X into X is continuous [14, Theorem 5.7-2].*

Example 2.4.28. *Let $\mathbf{A} = (a_{ij})$ for $i, j = 1, \ldots, n$ be a real matrix. Then $\mathbf{A}^* = \mathbf{A}^\mathsf{T} = \mathbf{A}$ iff $\mathbf{A}$ is symmetric.*

2.4.7 Strong and weak convergence

The strong convergence in a normed linear space X was defined in Definition 2.4.3. Remark 2.4.8 implies that in infinite-dimensional Banach spaces there exist bounded sequences that have no convergent subsequence, which is in contrast to bounded sequences in a finite-dimensional normed space. In order to overcome this difficulty, the following notion of convergence is very important. It was introduced by Hilbert around 1906 for a Hilbert space, but can be generalized to normed linear spaces.

Definition 2.4.38 (Weak convergence in X)**.** *Let X be a normed linear space, $\{x_n\}_{n=1}^\infty$ a sequence in X. The sequence x_n converges weakly to $x \in X$ as $n \to \infty$ (notation: $x_n \rightharpoonup x$) if*

$$f(x_n) \to f(x) \quad \text{for all } f \in X^*.$$

Linear forms and functionals can also be considered on the Banach space X^*.

Definition 2.4.39. *Let X be a Banach space. The dual of X^* is the Banach space $(X^*)^*$ denoted by X^{**}. Let us denote the elements of X^{**} by x^{**}. The operator J from X into X^{**} with $\mathcal{D}(J) = X$ defined by*

$$(Jx)(f) = f(x) \qquad \textit{for} \quad f \in X^*, x \in X$$

is called the canonical mapping from X into X^{**}. Note that Jx is the element $x^{**} \in X^{**}$ which satisfies $x^{**}(f) = f(x)$. Denote by $J(X)$ the image of the canonical mapping J.

Theorem 2.4.29. *Let X be a Banach space. Then X is isometrically embedded into X^{**}, i.e. J is an isometric isomorphism between X and $J(X)$.*

Definition 2.4.40 (Reflexive)**.** *A normed space X is called reflexive if $J(X) = X^{**}$ or shortly $X \cong X^{**}$.*

A reflexive space X is a Banach space, since X is then isometric to the Banach space X^{**}. This is the reason why in Definition 2.4.39 the space X is a Banach space. The following lemma follows from Riesz' representation theorem 2.4.25.

Lemma 2.4.17. *Any Hilbert space is a reflexive Banach space.*

Lemma 2.4.18.

(i) Let X be a reflexive Banach space. Then every subspace of X is also reflexive.

(ii) The product of a finite number of reflexive Banach spaces is a reflexive Banach space.

(iii) Let X and Y be two isomorphic normed linear spaces. Moreover, let X be a reflexive Banach space. Then Y is also a reflexive Banach space.

(iv) A Banach space X is reflexive iff X^* is reflexive.

(v) Let X be a separable reflexive Banach space. Then its dual X^* is separable.

The following result of Eberlein (1947) and Šmuljan (1940) is crucial to overcome the difficulty that the unit ball in an infinite dimensional Banach space is not compact [7, Theorem 21.D].

Theorem 2.4.30 (Weak compactness of reflexive spaces)**.** *Let X be a reflexive Banach space. Then each bounded set in X is weakly compact. This means that every bounded sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ contains a weak convergent subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ with limit $x \in X$, i.e.*

$$f(x_{n_k}) \to f(x), \qquad \forall f \in X^*.$$

Example 2.4.31. *A countably infinite orthonormal family $\{e_i\}$ in a Hilbert space is bounded since $\|e_i\| = 1$ for all i. This sequence cannot contain any convergent subsequence since $\|e_i - e_j\| = \sqrt{2}$ if $i \neq j$. On the other hand, this sequence has a weak convergent subsequence thanks to the previous theorem.*

The following properties are also important [7, Proposition 21.23].

Lemma 2.4.19 (Properties of weak convergence)**.** *Let X be a Banach space and Y be a normed space.*

(i) Every weakly convergent sequence in X is bounded;

(ii) Strong convergence in X implies weak convergence;

(iii) If X is finite dimensional, then weak convergence implies strong convergence;

(iv) Weakly lower semicontinuity of the norm: the norm $\|\cdot\|_X$ in a Banach space X is weakly sequentially lower semicontinuous on X, i.e

$$\|x\|_X \leqslant \liminf_{n\to\infty} \|x_n\|_X := \lim_{n\to\infty}\left(\inf_{m\geq n}\|x_n\|_X\right) \quad \text{if } x_n \rightharpoonup x,$$

where 'lim inf' is the limit inferior and 'inf' the infimum;

(v) If $A \in \mathcal{L}(X,Y)$ and $x_n \rightharpoonup x$ then $Ax_n \rightharpoonup Ax$;

(vi) If $A \in \sigma(X,Y)$ and $x_n \rightharpoonup x$ then $Ax_n \to Ax$;

(vii) Let X be a Hilbert space with norm $\|\cdot\|_X$. Assume that $x_n \rightharpoonup x$ in X and $\|x_n\|_X \to \|x\|_X$. Then $x_n \to x$ in X.

(viii) It follows from

$$\begin{aligned} &u_n \rightharpoonup u \quad \text{in } X && \text{as } n\to\infty,\\ &f_n \to f \quad \text{in } X^* && \text{as } n\to\infty,\end{aligned}$$

that $\langle f_n, u_n\rangle \to \langle f, u\rangle$ as $n\to\infty$.

(ix) If X is reflexive, then it follows from

$$\begin{aligned} &u_n \to u \quad \text{in } X && \text{as } n\to\infty,\\ &f_n \rightharpoonup f \quad \text{in } X^* && \text{as } n\to\infty,\end{aligned}$$

that $\langle f_n, u_n\rangle \to \langle f, u\rangle$ as $n\to\infty$.

This crucial convergence theorem of Eberlein and Šmuljan (Theorem 2.4.30) is only valid in reflexive Banach spaces. In nonreflexive Banach spaces, one can replace weak convergence by weak* convergence.

Definition 2.4.41 (Weak* convergence in X^*)**.** *Let X be a normed linear space with norm $\|\cdot\|_X$. A sequence $\{f_n\}_{n=1}^{\infty}$ in X^* converges weakly to $f \in X^*$ as $n \to \infty$ (notation: $f_n \overset{*}{\rightharpoonup} f$) if*

$$f_n(x) \to f(x) \quad \text{for all } x \in X.$$

The following theorems can be found in [7, Theorem 21.E and Proposition 21.26].

Theorem 2.4.32. *Let X be a separable Banach space. Then, each bounded sequence $\{f_n\}_{n\in\mathbb{N}}$ in X^* has a weakly* convergent subsequence.*

Theorem 2.4.33 (Properties of weak* and strong convergence)**.** *Let X be a Banach space.*

- *Every weakly* convergent sequence in X^* is bounded and*
$$\|f\|_{X^*} \leqslant \liminf_{n\to\infty} \|f_n\|_{X^*} \qquad \textit{if } f_n \overset{*}{\rightharpoonup} f.$$
- *The strong convergence in X^* implies the weak* convergence;*
- *It follows from*
$$\begin{aligned} u_n &\to u \quad \textit{in } X \qquad \textit{as } n \to \infty, \\ f_n &\overset{*}{\rightharpoonup} f \quad \textit{in } X^* \qquad \textit{as } n \to \infty, \end{aligned}$$
that $\langle f_n, u_n\rangle \to \langle f, u\rangle$ as $n \to \infty$.
- *If X is reflexive, then $f_n \overset{*}{\rightharpoonup} f$ in X^* as $n \to \infty$ is equivalent to $f_n \rightharpoonup f$ in X^* as $n \to \infty$.*

The following results are applied very frequently [6, Proposition 10.13]. They generalize well-known convergence properties of sequences of real numbers.

Lemma 2.4.20 (Convergence principles in Banach spaces)**.** *A sequence $\{x_n\}_{n=1}^\infty$ in a Banach space X has the following convergence properties:*

(i) Let x be a fixed element of a Banach space X. If every subsequence of $\{x_n\}_{n=1}^\infty$ has, in turn, a subsequence which converges strongly to x, then the original sequence converges strongly to x, i.e., $x_n \to x$ as $n \to \infty$.

(ii) Let x be a fixed element in a Banach space X. If every subsequence of $\{x_n\}_{n=1}^\infty$ has, in turn, a subsequence which converges weakly to x, then the original sequence converges weakly to x, i.e., $x_n \rightharpoonup x$ as $n \to \infty$.

(iii) Let $\{x_n\}_{n=1}^\infty$ be a bounded sequence in a reflexive Banach space X. If all the weakly convergent subsequences of $\{x_n\}_{n=1}^\infty$ have the same limit x, then $x_n \rightharpoonup x$ as $n \to \infty$.

2.5 Spaces of continuous functions

In this section, the spaces of continuous and Hölder-continuous functions are presented [24].

Definition 2.5.1 (Multi-indices, derivatives)**.** *Let $N \in \mathbb{N}$. A vector*
$$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)$$
with components $\alpha_i \in \mathbb{N} \cup \{0\}$ is said to be a multi-index of dimension N. The number
$$|\boldsymbol{\alpha}|_1 = \sum_{i=1}^N \alpha_i$$
is called the length of the multi-index $\boldsymbol{\alpha}$. The concept of the classical partial derivative of a function of N real variables u is well-known. The following notation is used for a multi-index $\boldsymbol{\alpha}$:
$$D^{\boldsymbol{\alpha}} u = \frac{\partial^{|\boldsymbol{\alpha}|_1} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_N^{\alpha_N}}.$$

Remark 2.5.1. *For simplicity, in this section, only real-valued functions are considered.*

Definition 2.5.2 (Domain)**.** *A subset $\Omega \subset \mathbb{R}^d$ is said to be a domain if it is nonempty, open and connected (i.e. every two points in Ω can be connected by a continuous curve that lies in Ω).*

Definition 2.5.3. *Let Ω be a domain in $\mathbb{R}^d$.*

- $\mathrm{C}(\Omega)$ *or* $\mathrm{C}^0(\Omega)$ *denotes the set of all functions defined and continuous on* Ω*;*
- $\mathrm{C}^k(\Omega)$ *with $k \in \mathbb{N}$ denotes the set of all functions defined on Ω that have continuous derivatives up to the order k on Ω;*
- $\mathrm{C}^\infty(\Omega)$ *denotes the set of all functions defined on Ω that have continuous derivatives of any order on Ω;*
- *Let u be a function on Ω. The set*
$$\mathit{supp}(f) := \overline{\{\mathbf{x} \in \Omega \ : \ u(\mathbf{x}) \neq 0\}},$$
where the bar denotes the closure in the space $\mathbb{R}^d$, is called the support of the function u;
- *Let $k \in \mathbb{N} \cup \{0, \infty\}$. Then $\mathrm{C}_0^k(\Omega)$ denotes the set of all functions $u \in \mathrm{C}^k(\Omega)$ whose supports are compact subsets of Ω, i.e., $u(\mathbf{x}) = 0$ in a small vicinity of $\partial\Omega$;*
- $\mathrm{C}(\overline{\Omega})$ *or* $\mathrm{C}^0(\overline{\Omega})$ *denotes the set of all functions $u \in \mathrm{C}(\Omega)$ that are bounded and uniformly continuous on Ω;*
- $\mathrm{C}^k(\overline{\Omega})$ *with $k \in \mathbb{N}$ denotes the set of all functions $u \in \mathrm{C}^k(\Omega)$ such that $D^{\boldsymbol{\alpha}} u \in \mathrm{C}(\overline{\Omega})$ for all muti-indices $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}|_1 \leqslant k$;*
- $\mathrm{C}^\infty(\overline{\Omega})$ *denotes the set of all functions $u \in \mathrm{C}^k(\Omega)$ such that $D^{\boldsymbol{\alpha}} u \in \mathrm{C}(\overline{\Omega})$ for any order k.*

The space $\mathrm{C}_0^\infty(\Omega)$ is a very important space and is called the *Schwartz space*. Sometimes, this space is also denoted by $\mathcal{D}(\Omega)$. A function in this space is often called a *test function*. Note that a uniform continuous function is continuous, but the opposite is not true. Therefore, there is a clear difference between $\mathrm{C}(\Omega)$ and $\mathrm{C}(\overline{\Omega})$ even if $\Omega = \overline{\Omega}$. Moreover, a bounded and uniform continuous function on a domain $\Omega \subset \mathbb{R}^d$ has a uniquely determined extension on $\overline{\Omega}$, which is again uniformly continuous [26, Lemma 3.11]. This justifies the previous notations.

Remark 2.5.2. *If the domain Ω is bounded, then $\overline{\Omega}$ is compact and every continuous function on $\overline{\Omega}$ is uniform continuous and bounded. Then, for instance, the space $\mathrm{C}(\overline{\Omega})$ can be defined as the set of functions that are continuous in $\overline{\Omega}$.*

Lemma 2.5.1. *The norm of elements in the spaces* $\mathrm{C}(\overline{\Omega})$ *and* $\mathrm{C}^k(\overline{\Omega})$ *can be defined as follows*

$$\|u\|_{\mathrm{C}(\overline{\Omega})} = \sup_{\mathbf{x}\in\Omega} |u(\mathbf{x})|,$$

$$\|u\|_{\mathrm{C}^k(\overline{\Omega})} = \sum_{n=0}^{k} \sum_{|\boldsymbol{\alpha}|_1=n} \sup_{\mathbf{x}\in\Omega} |D^{\boldsymbol{\alpha}} u(\mathbf{x})| = \sum_{|\boldsymbol{\alpha}|_1\leqslant k} \|D^{\boldsymbol{\alpha}} u\|_{\mathrm{C}(\overline{\Omega})}.$$

Definition 2.5.4 (Uniform convergence)**.** *A sequence* $\{f_n\}_{n=1}^{\infty}$ *of functions* $f_n \in \mathrm{C}(\overline{\Omega})$ *is said to converge uniformly as* $n \to \infty$ *to a function* $f \in \mathrm{C}(\overline{\Omega})$ *if* $\lim_{n\to\infty} \|f_n - f\|_{\mathrm{C}(\overline{\Omega})} = 0$.

This uniform convergence of the sequence $\{f_n\}_{n=1}^{\infty}$ to f as $n \to \infty$ implies the pointwise convergence, i.e. that

$$f_n(x) \to f(x) \text{ as } n \to \infty \text{ for every } x \in X.$$

Remark 2.5.3. *The space* $\mathrm{C}(\overline{\Omega})$ *equipped with the inner product*

$$(f, g)_\Omega = \int_\Omega f(\mathbf{x})g(\mathbf{x})\ \mathrm{d}\mathbf{x}$$

is not complete.

Theorem 2.5.1. *Let* $k \in \mathbb{N} \cup \{0\}$. *The spaces* $\mathrm{C}^k(\overline{\Omega})$ *are Banach spaces but not Hilbert spaces.*

Definition 2.5.5 (Spaces $\mathrm{C}^{k,\lambda}(\overline{\Omega})$)**.** *Take* $k \in \mathbb{N} \cup \{0\}$. *Let* $\lambda \in (0, 1]$ *be a given real parameter. Let* Ω *be a domain in* $\mathbb{R}^d$. *We denote by* $\mathrm{C}^{k,\lambda}(\overline{\Omega})$ *the set of all functions* u *from* $\mathrm{C}^k(\overline{\Omega})$ *satisfying*

$$H_{\boldsymbol{\alpha},\lambda}(u) = \sup_{\substack{\mathbf{x},\mathbf{y}\in\Omega \\ \mathbf{x}\neq\mathbf{y}}} \frac{|D^{\boldsymbol{\alpha}} u(\mathbf{x}) - D^{\boldsymbol{\alpha}} u(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^{\lambda}} < \infty$$

for all multi-indices $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ *with the length* $|\boldsymbol{\alpha}|_1 = k$.

Remark 2.5.4. *Note that the space* $\mathrm{C}^{0,\lambda}(\overline{\Omega})$ *with* $\lambda \in (0, 1]$ *is the space of Hölder-continuous functions with exponent* λ. *The space* $\mathrm{C}^{0,1}(\overline{\Omega})$ *is the space of Lipschitz continuous functions.*

Lemma 2.5.2. *A norm in the vector space* $\mathrm{C}^{k,\lambda}(\overline{\Omega})$ *is*

$$\|u\|_{\mathrm{C}^{k,\lambda}(\overline{\Omega})} = \|u\|_{\mathrm{C}^k(\overline{\Omega})} + \sum_{|\boldsymbol{\alpha}|_1=k} H_{\boldsymbol{\alpha},\lambda}(u), \qquad u \in \mathrm{C}^{k,\lambda}(\overline{\Omega}).$$

Theorem 2.5.2. *Let* $k \in \mathbb{N} \cup \{0\}$ *and* $\lambda \in (0, 1]$. *The space* $\mathrm{C}^{k,\lambda}(\overline{\Omega})$ *is a Banach space.*

The fundamental Arzelà-Ascoli theorem [24, Theorem 1.5.3] gives a characterization of compact subsets in $\mathrm{C}(\overline{\Omega})$ using the equiboundedness and uniform equicontinuity properties.

Definition 2.5.6. *Let Ω be a bounded domain in $\mathbb{R}^d$, $d \in \mathbb{N}$. Then the subset K of $\mathrm{C}(\overline{\Omega})$ is equicontinuous in $\mathbf{x}_0 \in \overline{\Omega}$ if*

$$(\forall \varepsilon > 0)(\exists \delta(\varepsilon, \mathbf{x}_0) > 0)$$
$$(\forall f \in K; \forall \mathbf{y} \in \overline{\Omega} : |\mathbf{y} - \mathbf{x}_0|_{\mathrm{e}} < \delta \Rightarrow |f(\mathbf{y}) - f(\mathbf{x}_0)| < \varepsilon).$$

The set K is equicontinuous if K is equicontinuous in each point $\mathbf{x}_0 \in \overline{\Omega}$. The set K is uniform equicontinuous if

$$(\forall \varepsilon > 0)(\exists \delta(\varepsilon) > 0)(\forall f \in K; \forall \mathbf{x}, \mathbf{y} \in \overline{\Omega} : |\mathbf{x} - \mathbf{y}|_{\mathrm{e}} < \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon).$$

The set K is equibounded if

$$|f(\mathbf{x})| \leqslant C, \qquad \forall \mathbf{x} \in \overline{\Omega}, \forall f \in K.$$

Theorem 2.5.3 (Arzelà-Ascoli theorem)**.** *Let Ω be a bounded domain in $\mathbb{R}^d$. A subset K of $\mathrm{C}(\overline{\Omega})$ is relatively compact iff it is equibounded and uniform equicontinuous.*

Theorem 2.5.4. *Let Ω be a bounded domain in $\mathbb{R}^d$ and $k \in \mathbb{N} \cup \{0\}$. A subset K of $\mathrm{C}^k(\overline{\Omega})$ is relatively compact iff the following conditions are satisfied:*
(i) K is equibounded in $\mathrm{C}^k(\overline{\Omega})$;
(ii) the sets $K_{\boldsymbol{\alpha}} := \{D^{\boldsymbol{\alpha}} u \ : \ u \in K\}$ are uniform equicontinuous for all $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}|_1 \leqslant k$.

Theorem 2.5.5. *Let Ω be a bounded domain in $\mathbb{R}^d$. The space $\mathrm{C}^k(\overline{\Omega})$, $k \in \mathbb{N} \cup \{0\}$, is separable. The space $\mathrm{C}^{k,\lambda}(\overline{\Omega})$ with $k \in \mathbb{N} \cup \{0\}$ and $\lambda \in (0,1]$ is not separable. The spaces $\mathrm{C}^k(\overline{\Omega})$ and $\mathrm{C}^{k,\lambda}(\overline{\Omega})$ with $k \in \mathbb{N} \cup \{0\}$ and $\lambda \in (0,1]$ are not reflexive.*

2.5.1 The space $\mathrm{C}^m([0,T], X)$

First, the notion of abstract function is considered.

Definition 2.5.7. *Let Ω be a domain in $\mathbb{R}^d$ and X a normed linear space. A mapping u from Ω (or from $\overline{\Omega}$) into X is said to be an abstract function (a function with values in X) defined on Ω (or on $\overline{\Omega}$).*

The focus is on $d = 1$. The following function space is frequently used.

Definition 2.5.8. *Let X be a Banach space with norm $\|\cdot\|_X$ over the field $\mathbb{F}$ and $0 < T < \infty$. The space $\mathrm{C}^m([0,T], X)$ with $m \in \mathbb{N} \cup \{0\}$ consists of all continuous functions $u : [0,T] \to X$ that have continuous derivatives up to order m on $[0,T]$. Together with the norm*

$$\|u\|_{\mathrm{C}^m([0,T],X)} = \sum_{i=0}^{m} \max_{0 \leqslant t \leqslant T} \left\| u^{(i)}(t) \right\|_X, \tag{2.3}$$

the space $\mathrm{C}^m([0,T],X)$ *forms a Banach space. Only the right-hand and the left-hand derivatives need to exist at the boundary points* $t=0$ *and* $t=T$*, respectively. In* (2.3)*,* $u^{(0)}$ *means* u*. The space* $\mathrm{C}^0([0,T],X)$ *is denoted by* $\mathrm{C}([0,T],X)$*.*

Remark 2.5.5. *Note that the completeness of* X *is used in the construction of these spaces.*

The following generalization of the Arzelà-Ascoli theorem 2.5.3 holds [24, Theorem 1.6.9]. The definition of uniform equicontinuity (see Definition 2.5.6) can be extended to abstract functions by replacing $|f(x)-f(y)|$ by $\|f(x)-f(y)\|_X$.

Theorem 2.5.6. *Let* X *be a Banach space. Then a set* K *in* $\mathrm{C}([0,T],X)$ *is relatively compact iff*
(i) K *is uniform equicontinuous,*
(ii) the set $K(t) := \{u(t) \,:\, u \in K\}$ *is relatively compact in* X *for any* $t \in [0,T]$*.*

Theorem 2.5.7. *Let* X *be a separable Banach space. Then, the space* $\mathrm{C}^m([0,T],X)$ *is separable for* $m \in \mathbb{N}\cup\{0\}$*.*

Remark 2.5.6. *Analogously to Theorem 2.5.5, the space* $\mathrm{C}^m([0,T],X)$ *with* $m \in \mathbb{N}\cup\{0\}$ *is not reflexive.*

2.6 Domains

Some properties of function spaces require a certain degree of regularity of the boundary of a bounded domain $\Omega \subset \mathbb{R}^d$. A boundary of a domain is often denoted by Γ or $\partial\Omega$ and can be interpreted as a $(d-1)$-dimensional object (manifold, i.e. locally Euclidean). A boundary Γ can be viewed locally as the graph of a function φ. Then, the properties of Γ are specified through the properties of φ. In this section, a distinction is made between Lipschitz continuous and (piecewise) smooth boundaries. The exact definition of Lipschitz continuity for boundaries of domains in $\mathbb{R}^d$ is rather technical [10].

Definition 2.6.1. *Let* Ω *be an open subset of* $\mathbb{R}^d$*,* $d \geqslant 2$*. Its boundary* Γ *is Lipschitz continuous if for every* $\mathbf{x} \in \Gamma$ *there exist a neighbourhood* V *of* $\mathbf{x}$ *in* $\mathbb{R}^d$ *and new orthogonal coordinates* $\{y_1,\ldots,y_d\}$ *such that*
(a) V *is a hypercube in the new coordinates:*

$$V = \{(y_1,\ldots,y_d) \,:\, -a_i < y_i < a_i, 1 \leqslant i \leqslant d\};$$

(b) there exists a Lipschitz continuous function φ*, defined in*

$$V' = \{(y_1,\ldots,y_{d-1}) \,:\, -a_i < y_i < a_i, 1 \leqslant i \leqslant d-1\}$$

such that

$$|\varphi(y')| \leqslant \frac{a_d}{2} \text{ for every } y' = (y_1,\ldots,y_{d-1}) \in V',$$

$$\Omega \cap V = \{y = (y', y_d) \in V \,:\, y_d < \varphi(y')\},$$
$$\Gamma \cap V = \{y = (y', y_d) \in V \,:\, y_d = \varphi(y')\}.$$

In other words, in a neighbourhood of $\mathbf{x}, \Omega$ *is below the graph of* φ *and consequently the boundary* Γ *is the graph of* φ*, see Figure 2.1 [10, 28].*

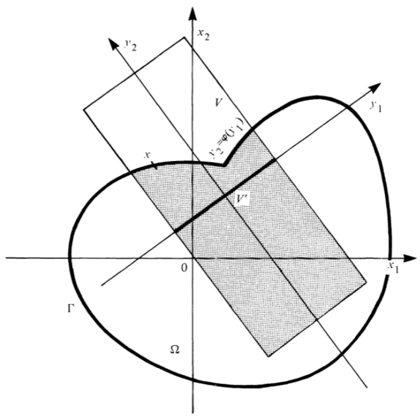


Figure 2.1: Illustration of the definition of Lipschitz continuous domain.

Definition 2.6.2 ((Piecewise) smooth boundary)**.** *If the function* φ *from Definition 2.6.1 belongs to* $\mathrm{C}^\infty(V')$ *for every* $\mathbf{x} \in \Gamma$*, then the boundary is called smooth. A bounded domain* $\Omega \subset \mathbb{R}^d$ *has a piecewise smooth boundary* $\partial\Omega$ *if* $\partial\Omega$ *consists of a finite number of smooth parts.*

Definition 2.6.3. *The bounded domain* $\Omega \subset \mathbb{R}^d$ *is said to be of class* $\mathrm{C}^{m,1}$*, for an integer* $m \geqslant 1$*, if the mappings* φ *from Definition 2.6.1 can be chosen* m*-times differentiable with Lipschitz-continuous partial derivatives of order* m*.*

This implies that a Lipschitz domain Ω is sometimes denoted as an element of the class $\mathrm{C}^{0,1}(\Omega)$. This means that Definition 2.6.1 holds with maps $\varphi \in \mathrm{C}^{0,1}(V')$ for each V' in the definition. Therefore, a smooth domain is automatically a Lipschitz domain (this is not the case for a piecewise smooth domain). A key property of a Lipschitz domain is contained in the following lemma, see [3] and [29, Lemma 4.2].

Lemma 2.6.1. *A Lipschitz domain* $\Omega \subset \mathbb{R}^d$ *has a well-defined unit outward normal vector denoted by* $\boldsymbol{n}$ *(or* $\boldsymbol{\nu}$*) at almost every point of* $\partial\Omega = \Gamma$ *.*

The reason for using Lipschitz polyhedral domains is that they can be covered by a mesh of tetrahedra, which is more interesting for practical implementations. For simply-connected domains (i.e., domains $\Omega \subset \mathbb{R}^d$ such that $\mathbb{R}^d \setminus \Omega$ is connected), the Lipschitz continuity is equivalent to the cone condition [10, p. 414].

Definition 2.6.4 (Cone condition)**.** *The boundary Γ of a d-dimensional bounded domain satisfies the cone condition if there exist constants $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that for every point $\mathbf{x} \in \Gamma$ there are two open d-dimensional cones $C_{int}(\mathbf{x}, \gamma_1, h_1)$ and $C_{ext}(\mathbf{x}, \gamma_2, h_2)$ sharing the vertex $\mathbf{x}$, with vertex angles $0 < \varepsilon_1 \leqslant \gamma_1, \gamma_2$ and heights $0 < \varepsilon_2 \leqslant h_1, h_2$, such that $C_{int}(\mathbf{x}, \gamma_1, h_1) \subset \Omega$ and $C_{ext}(\mathbf{x}, \gamma_2, h_2) \subset \mathbb{R}^d \setminus \Omega$.*

Example 2.6.1. *Figure 2.2 gives a example of a domain whose boundary is Lipschitz continuous. Every bounded convex open set satisfies the uniform cone property and hence is a Lipschitz domain.*

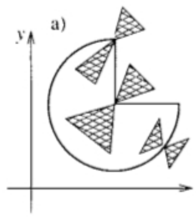


Figure 2.2: 2D domain with a Lipschitz continuous boundary.

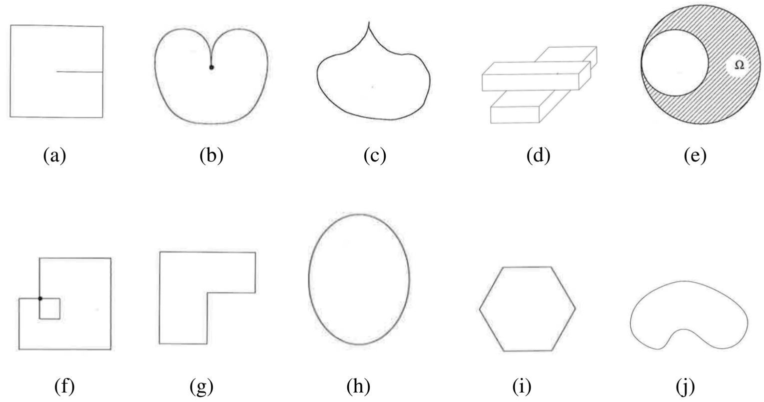


Figure 2.3: Pictures of domains: (a) non-smooth, non-Lipschitz; (b,c,d,e,f) piecewise smooth, non-Lipschitz; (g,i) piecewise smooth, Lipschitz; (h,j) smooth.

In engineering applications most of the domains are Lipschitz.

2.7 Lebesgue spaces

In this section, Lebesgue spaces and its properties are reviewed [6, 14, 24].

Definition 2.7.1. *Let $p \in [1, \infty)$. The vector space $\mathcal{L}^p(\Omega)$ is the set of all measurable functions f from the bounded domain Ω to $\mathbb{F}$, for which*

$$\|f\|_{p,\Omega} = \left(\int_\Omega |f(\mathbf{x})|_c^p \ \mathrm{d}\mathbf{x} \right)^{\frac{1}{p}} < \infty. \tag{2.4}$$

By itself, this space is not a Banach space because there are non-zero functions of which the norm is zero.

Definition 2.7.2. *An equivalence relation is defined as follows: the functions f and g are equivalent if the norm of $f - g$ vanishes ($f(\mathbf{x}) = g(\mathbf{x})$ for a.a. $\mathbf{x} \in \Omega$). The set of equivalence classes forms the Lebesgue space $\mathrm{L}^p(\Omega)$.*

Example 2.7.1. *Let $\Omega = [-1, 1]$. The following functions f and g defined by*

$$f(x) = \begin{cases} 1 & x \geqslant 0 \\ 0 & x < 0 \end{cases} \quad \textit{and} \quad g(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leqslant 0 \end{cases}$$

only differ on a set of measure zero, i.e. $\|f - g\|_{p,\Omega} = 0$ for $p \in [1, \infty)$.

Theorem 2.7.2. *The space $\mathrm{L}^p(\Omega)$ is a Banach space with the norm $\|f\|_{\mathrm{L}^p(\Omega)} = \|f\|_{p,\Omega}$ for $p \in [1, \infty)$.*

Note that, Minkowski's inequality (see Lemma 2.3.6) is in fact the triangle inequality for $\mathrm{L}^p(\Omega)$-spaces. Only for $p = 2$, the space $\mathrm{L}^p(\Omega)$ is a Hilbert space (see [24, Chapter II]) with a scalar product

$$(f, g)_\Omega = \int_\Omega f(\mathbf{x})\overline{g(\mathbf{x})} \ \mathrm{d}\mathbf{x}.$$

This illustrates also that each Hilbert space is a Banach space and that the converse is not generally true. This space plays an essential role in applications. The following lemma results from Hölder's inequality.

Lemma 2.7.1. *The inclusion $\mathrm{L}^{p_1}(\Omega) \subset \mathrm{L}^{p_2}(\Omega)$ holds for $1 \leqslant p_2 < p_1 < \infty$ with*

$$\|f\|_{p_2,\Omega} \leqslant |\Omega|^{\frac{p_1 - p_2}{p_1 p_2}} \|f\|_{p_1,\Omega},$$

where $|\Omega|$ denotes the measure of Ω, i.e. $\mathrm{L}^{p_1}(\Omega) \hookrightarrow \mathrm{L}^{p_2}(\Omega)$.

This embedding is not compact, see [24, p. 90]. The following mean continuity property holds for functions in $\mathrm{L}^p(\Omega)$.

Theorem 2.7.3. *Let $f \in \mathrm{L}^p(\Omega)$ and $f(\mathbf{x}) = 0$ for $\mathbf{x} \notin \Omega$. Then for all $\varepsilon > 0$, there exists a $\delta > 0$ such that*

$$\left(\int_\Omega |f(\mathbf{x}+\mathbf{z}) - f(\mathbf{x})|_c^p \ \mathrm{d}\mathbf{x} \right)^{\frac{1}{p}} < \varepsilon$$

provided $\mathbf{z} \in \mathbb{R}^d$ with $|\mathbf{z}|_e < \delta$, i.e. each function $f \in \mathrm{L}^p(\Omega)$ is p-mean continuous.

Remark 2.7.1. *From this point on, only real-valued functions are considered. In $\mathrm{L}^2(\Omega)$, the norm $\|f\|_{2,\Omega}$ is denoted by $\|f\|$ and the inner product $(f,g)_\Omega$ by (f,g).*

The following theorem sometimes allows us to prove results for smooth functions and to extend them by limiting arguments to more general functions.

Theorem 2.7.4. *The space $\mathrm{C}_0^\infty(\overline{\Omega})$ is dense in $\mathrm{L}^p(\Omega)$ for all $1 \leqslant p < \infty$.*

Corollary 2.7.1. *The space $\mathrm{C}_0^\infty(\Omega)$ is dense in $\mathrm{L}^p(\Omega)$ for all $1 \leqslant p < \infty$.*

Theorem 2.7.5. *Let $p \in [1, \infty)$. Then the Banach space $\mathrm{L}^p(\Omega)$ is separable.*

Unfortunately, $\mathrm{L}^p(\Omega)$ is not a (relatively) compact space as it is an infinite dimensional linear normed space. This makes the analysis in these spaces more difficult. The space $\mathrm{L}^p(\Omega)$ for $p = 2$ is a reflexive Banach space thanks to Corollary 2.4.17. In this case, there is no problem to apply Theorem 2.4.30. When $p \neq 2$, the Riesz' representation theorem for $\mathrm{L}^p(\Omega)$-spaces implies the reflexivity of $\mathrm{L}^p(\Omega)$ for $1 < p < \infty$. Note that the classical Riesz' representation theorem (Theorem 2.4.25) is not applicable if $p \neq 2$.

Theorem 2.7.6 (Riesz' representation theorem in $\mathrm{L}^p(\Omega)$)**.** *Let Ω be a nonempty bounded open subset of $\mathbb{R}^d$. Let φ be a bounded linear functional on $\mathrm{L}^p(\Omega)$ with $1 < p < \infty$. Then there exists exactly one $g \in \mathrm{L}^q(\Omega)$ with $\frac{1}{p} + \frac{1}{q} = 1$ such that*

$$\varphi(f) = \int_\Omega g(\mathbf{x}) f(\mathbf{x}) \ \mathrm{d}\mathbf{x}$$

for every $f \in \mathrm{L}^p(\Omega)$. Moreover

$$\|\varphi\|_{\mathrm{L}^p(\Omega)^*} = \|g\|_{\mathrm{L}^q(\Omega)}.$$

Corollary 2.7.2. *The dual space of $\mathrm{L}^p(\Omega)$ for $1 < p < \infty$ is $\mathrm{L}^q(\Omega)$, with $\frac{1}{p} + \frac{1}{q} = 1$.*

Corollary 2.7.3. *The space $\mathrm{L}^p(\Omega)$ is reflexive for $1 < p < \infty$.*

From this corollary it follows that Theorem 2.4.30 is valid for $\mathrm{L}^p(\Omega)$, with $p \in (1, \infty)$. Hence, each bounded sequence $\{f_n\}$ in $\mathrm{L}^p(\Omega)$ has as subsequence $\{f_{n_k}\}$ with

$$f_{n_k} \rightharpoonup f \quad \text{in } \mathrm{L}^p(\Omega) \qquad \text{as } k \to \infty$$

or

$$\int_\Omega f_{n_k}(\mathbf{x})g(\mathbf{x})\ \mathrm{d}\mathbf{x} \to \int_\Omega f(\mathbf{x})g(\mathbf{x})\ \mathrm{d}\mathbf{x} \quad \text{as } k \to \infty \quad \text{for all} \quad g \in \mathrm{L}^q(\Omega),$$

with $\frac{1}{p} + \frac{1}{q} = 1$.

The space $\mathrm{L}^1(\Omega)$ is a non-reflexive Banach space. The following important theorem about this space can be found in [24, p. 88].

Theorem 2.7.7 (Convergence in $\mathrm{L}^1(\Omega)$)**.** *Any convergent sequence in $\mathrm{L}^1(\Omega)$ converges in measure, i.e. it has a convergent subsequence a.e. in Ω (i.e. pointwise convergence outside a set of measure zero).*

Example 2.7.8. *Let Ω be a nonempty bounded set in $\mathbb{R}^d$, $d \in \mathbb{N}$ and $1 \leqslant p < \infty$. Suppose that*

$$u_n \to u \quad \textit{in } \mathrm{L}^p(\Omega) \qquad \textit{as } n \to \infty,$$

i.e.

$$\int_\Omega |u_n(\mathbf{x}) - u(\mathbf{x})|^p\ \mathrm{d}\mathbf{x} \to 0 \qquad \textit{as } n \to \infty.$$

By Lemma 2.7.1 there is also convergence in $\mathrm{L}^1(\Omega)$, which implies

$$\int_\Omega u_n(\mathbf{x})\ \mathrm{d}\mathbf{x} \to \int_\Omega u(\mathbf{x})\ \mathrm{d}\mathbf{x} \qquad \textit{as } n \to \infty.$$

From this convergence, it follows (Theorem 2.7.7) that there exists a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$u_{n_k}(\mathbf{x}) \to u(\mathbf{x}) \qquad \textit{as } k \to \infty \qquad \textit{for a.a.} \quad \mathbf{x} \in \Omega.$$

Moreover, there exists a function $v \in \mathrm{L}^p(\Omega)$ such that

$$|u_{n_k}(\mathbf{x})| \leqslant v(\mathbf{x}) \qquad \textit{for all } n_k \quad \textit{and a.a.} \quad \mathbf{x} \in \Omega.$$

The following theorem is an application of Lemma 2.4.19.

Theorem 2.7.9 (A convergence theorem)**.** *Let $1 < p, q < +\infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. From*

$$\begin{aligned} u_n &\to u \quad \textit{in } \mathrm{L}^p(\Omega) \qquad \textit{as } n \to \infty, \\ v_n &\rightharpoonup v \quad \textit{in } \mathrm{L}^q(\Omega) \qquad \textit{as } n \to \infty, \end{aligned}$$

it follows that

$$\int_\Omega u_n(\mathbf{x})v_n(\mathbf{x})\ \mathrm{d}\mathbf{x} \to \int_\Omega u(\mathbf{x})v(\mathbf{x})\ \mathrm{d}\mathbf{x} \quad \textit{as } n \to \infty.$$

The following theorem about relatively compact subsets of $\mathrm{L}^p(\Omega)$ is proved by Riesz (one-dimensional) and Kolmogorov (multidimensional), see [24, Theorem 2.13.1].

Theorem 2.7.10 (Riesz-Kolmogorov)**.** *Let $1 \leqslant p < \infty$. The set $K \subset \mathrm{L}^p(\Omega)$ is relatively compact iff the following conditions are satisfied:*

(i) The set K is bounded,

(ii) The set K is p-mean equicontinuous, i.e., for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that

$$\int_\Omega |f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} < \varepsilon^{p-1},$$

for each $f \in K$ and $\mathbf{h} \in \mathbb{R}^d$ with $|\mathbf{h}|_{\mathrm{e}} < \delta$.

This section finishes with the definition of the space $\mathrm{L}^\infty(\Omega)$.

Definition 2.7.3. *$\mathcal{L}_\infty(\Omega)$ denotes the set of all Lebesgue measurable functions f defined a.e. in Ω for which there exist a constant $K > 0$ and a set $E \subset \Omega$ of measure zero (K and E both depending on f), with the property*

$$|f(\mathbf{x})| < K$$

for all $\mathbf{x} \in \Omega \setminus E$.

$\mathcal{L}_\infty(\Omega)$ is a real (or complex) vector space. Equivalence classes are defined on $\mathcal{L}_\infty(\Omega)$ by saying that $f_1 = f_2$ if $f_1(\mathbf{x}) = f_2(\mathbf{x})$ for a.a. $\mathbf{x} \in \Omega$. The vector space of these classes is $\mathrm{L}^\infty(\Omega)$. Then

$$\begin{aligned}\|f\|_{\infty,\Omega} &= \text{ess sup}_{x\in\Omega}|f(\mathbf{x})| = \inf\{a \in \mathbb{R} : \mu(\{\mathbf{x} \in \Omega : f(\mathbf{x}) > a\}) = 0\} \\ &= \lim_{p\to\infty} \|f\|_{p,\Omega}\end{aligned}$$

is a norm in $\mathrm{L}^\infty(\Omega)$. This norm is called the essential supremum of f. The space $\mathrm{C}(\overline{\Omega})$ is a closed linear subset of $\mathrm{L}^\infty(\Omega)$.

Theorem 2.7.11. *The space $\mathrm{L}^\infty(\Omega)$ is a nonreflexive nonseparable Banach space. Moreover, $\mathrm{L}^\infty(\Omega) \ncong \mathrm{L}^1(\Omega)^*$.*

2.8 Weak derivatives

This section introduces the notion of weak derivative. It is an extension of the classical derivative. This extension is necessary because physically relevant problems often have solutions that are not regular enough to have classical partial derivatives (Burger's equation, Porous media equation,...). First, the space $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ is defined.

Definition 2.8.1 (Locally integrable). *Let Ω be an open subset of $\mathbb{R}^d$. The space $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ consists of all functions that are integrable (of which the integral is finite) on any compact subset of their domain of definition, i.e. the restriction $f|_K$ of f to any compact subset K of Ω belongs to the space $\mathrm{L}^1(K)$. Equivalently, a function $f : \Omega \to \mathbb{R}$ such that*

$$\int_\Omega |f(\mathbf{x})\varphi(\mathbf{x})| \ \mathrm{d}\mathbf{x} < +\infty,$$

for each test function $\varphi \in \mathrm{C}_0^\infty(\Omega)$ is called locally integrable.

Remark 2.8.1. *Every function f belonging to $\mathrm{L}^p(\Omega), 1 \leqslant p \leqslant \infty$, where Ω is an open subset of $\mathbb{R}^d$, is locally integrable, since for any compact subset K of Ω, it holds that*

$$\int_K |f(\mathbf{x})| \ \mathrm{d}\mathbf{x} \leqslant \|f\|_{\mathrm{L}^1(\Omega)} < \infty.$$

Clearly, any function in the space $\mathrm{C}(\Omega)$ is locally integrable on Ω, since for any compact subset K of Ω, it holds that

$$\int_K |f(\mathbf{x})| \ \mathrm{d}\mathbf{x} \leqslant \left(\int_K 1 \ \mathrm{d}\mathbf{x}\right) \sup_{\mathbf{x}\in K} |f(\mathbf{x})| < \infty.$$

Theorem 2.8.1. *Let $v \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$, with Ω a nonempty open set in $\mathbb{R}^d$. If*

$$\int_\Omega v(\mathbf{x})\varphi(\mathbf{x}) \ \mathrm{d}\mathbf{x} = 0, \qquad \forall \varphi \in \mathrm{C}_0^\infty(\Omega),$$

then $v = 0$ a.e. in Ω.

The definition of generalized derivative is a generalization of the classical integration by parts formula. For more information, the reader is referred to [7, Section 21.1].

Definition 2.8.2. *Let Ω be a nonempty open set in $\mathbb{R}^d$ and assume that $u, v \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. The function $v = u^{(\boldsymbol{\alpha})} = D^{\boldsymbol{\alpha}} u$ is called the $\boldsymbol{\alpha}$-th weak or generalized derivative of the function u if*

$$\int_\Omega v(\mathbf{x})\varphi(\mathbf{x}) \ \mathrm{d}\mathbf{x} = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega u(\mathbf{x}) \left(D^{\boldsymbol{\alpha}}\varphi\right)(\mathbf{x}) \ \mathrm{d}\mathbf{x}$$

holds for all $\varphi \in \mathrm{C}_0^\infty(\Omega)$.

Example 2.8.2. *Let $u : [a, b] \to \mathbb{R} : x \mapsto |x|$, with $0 \in (a, b)$. Then the function*

$$w(x) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0, \\ \alpha & \text{if } x = 0, \end{cases}$$

where α is a fixed arbitrary real number, is the generalized derivative of u on $]a, b[$. At the point $x = 0$, where the classical derivative of u does not exist, the value of α

can be chosen arbitrarily. The function w has no generalized derivative. However, a distributional derivative exists. In what follows, only generalized derivatives which can be represented by functions are used.

Theorem 2.8.3 (Uniqueness of generalized derivatives)**.** *A weak derivative, if it exists, is uniquely defined up to a set of measure zero.*

Theorem 2.8.4. *If a function has a strong derivative, then its weak derivative is identical to it.*

The classical rules for derivatives of sums and products of functions also hold for weak derivatives. A main difference in comparison to the classical definition of the derivative of a function u is that to define the weak derivative $D^{\boldsymbol{\alpha}}u$, the existence of derivatives of order $\boldsymbol{\beta}$ with $|\boldsymbol{\beta}|_1 < |\boldsymbol{\alpha}|_1$ is not needed.

The following theorem shows that generalized derivatives are compatible with weak limits. This fact is important for proving the convergence of Galerkin/Rothe's method for evolution equations, see [7, p. 300].

Theorem 2.8.5 (Generalized derivatives and weak convergence: A)**.** *Let $\boldsymbol{\alpha}$ be a fixed multi-index and Ω a nonempty open set in $\mathbb{R}^d$. It follows from*

$$\begin{aligned} u_n &\rightharpoonup u \quad \text{in } \mathrm{L}^1(\Omega) \quad \text{as } n \to \infty, \\ D^{(\boldsymbol{\alpha})}u_n &\rightharpoonup v \quad \text{in } \mathrm{L}^1(\Omega) \quad \text{as } n \to \infty, \end{aligned}$$

that

$$D^{(\boldsymbol{\alpha})}u = v \quad \text{on} \quad \Omega.$$

The previous result is stronger then the following relation between uniform convergence and differentiation [30, Theorem 7.17].

Theorem 2.8.6. *Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of functions, differentiable on $[a,b]$ such that $\{f_n(x_0)\}_{n\in\mathbb{N}}$ converges for some point x_0 on $[a,b]$. If $\{f_n'\}_{n\in\mathbb{N}}$ converges uniformly on $[a,b]$, then $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to a function f, and $f'(x) = \lim_{n\to\infty} f_n'(x)$ for $x \in [a,b]$.*

2.9 Sobolev spaces

The sets of functions of which the generalized derivatives up to a fixed order k are elements of L^p-spaces have been investigated by many authors, see for instance Morrey [31], Gagliardo [32] and Sobolev [33]. There exist various definitions of so-called Sobolev spaces. Some of them are presented in the following subsections. These definitions are generally not equivalent. However, in many reasonable cases, depending on the domain where the functions are defined, the spaces coincide [24].

2.9.1 Sobolev Spaces $\mathrm{W}^{k,p}(\Omega)$

Definition 2.9.1. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. Take $i \in \mathbb{N} \cup \{0\}$ and $p \in [1, +\infty)$. The seminorm of the i-th derivative of $u \in \mathrm{C}^\infty(\overline{\Omega})$ is*

$$|u|_{i,p,\Omega} = \left(\sum_{|\boldsymbol{\alpha}|_1 = i} \int_\Omega |D^{\boldsymbol{\alpha}} u(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right)^{\frac{1}{p}}. \tag{2.5}$$

The summation goes through all multi-indices $\boldsymbol{\alpha}$ with the length i. The norm of $u \in \mathrm{C}^k(\overline{\Omega})$ with $k \in \mathbb{N} \cup \{0\}$ is defined in terms of the seminorms as follows

$$\|u\|_{k,p,\Omega} = \left(\sum_{i=0}^{k} |u|^p_{i,p,\Omega} \right)^{\frac{1}{p}}. \tag{2.6}$$

The space $\mathrm{W}^{k,p}(\Omega)$ is the closure of $\mathrm{C}^\infty(\overline{\Omega})$ with respect to the norm (2.6). *Hence, for any sequence $\{f_n\}_{n \in \mathbb{N}} \subset \mathrm{C}^\infty(\overline{\Omega})$ such that*

$$\lim_{n \to \infty} \|f_n - f\|_{k,p,\Omega} = 0,$$

it holds that $f \in \mathrm{W}^{k,p}(\Omega)$. This is a so-called Sobolev space. The Sobolev space $\mathrm{W}^{k,p}_0(\Omega)$ can be introduced in an analogous way as a closure of $\mathrm{C}^\infty_0(\Omega)$ with respect to the norm (2.6). *Analogously, the spaces $\mathrm{W}^{k,\infty}(\Omega)$ and $\mathrm{W}^{k,\infty}_0(\Omega)$ can be defined by using*

$$\|u\|_{k,\infty,\Omega} = \max_{|\boldsymbol{\alpha}|_1 \leqslant k} \|D^{\boldsymbol{\alpha}} u\|_{\mathrm{L}^\infty(\Omega)}.$$

The following theorem is taken from [24, Theorem 5.2.2].

Theorem 2.9.1. *Let $k \in \mathbb{N} \cup \{0\}$ and $p \in [1, +\infty]$. Then*

(i) $\mathrm{W}^{k,p}(\Omega)$ *and* $\mathrm{W}^{k,p}_0(\Omega)$ *are Banach spaces;*

(ii) if $p \in (1, \infty)$ then $\mathrm{W}^{k,p}(\Omega)$ *and* $\mathrm{W}^{k,p}_0(\Omega)$ *are reflexive;*

(iii) if $p \in [1, \infty]$ then $\mathrm{W}^{k,p}(\Omega)$ *and* $\mathrm{W}^{k,p}_0(\Omega)$ *are separable;*

(iv) $\mathrm{W}^{k,p}_0(\Omega) \subset \mathrm{W}^{k,p}(\Omega)$.

For the space $\mathrm{W}^{k,p}(\Omega)$ with $p = \infty$, the following result is known [24, Lemma 5.2.3].

Lemma 2.9.1. *Let $k \in \mathbb{N} \cup \{0\}$. Then,*

(i) $\mathrm{W}^{k,\infty}(\Omega) \rightleftarrows \mathrm{C}^k(\overline{\Omega})$*;*

(ii) $\mathrm{W}^{k,\infty}_0(\Omega) \rightleftarrows \mathrm{C}^k_0(\Omega)$.

The following inclusion relations provide an ordening among the different Sobolev spaces.

Theorem 2.9.2. *Let $k, k_1, k_2 \in \mathbb{N}$ such that $1 \leqslant k_1 \leqslant k_2$ and $p, p_1, p_2 \in [1, +\infty]$ such that $1 \leqslant p_2 \leqslant p_1 \leqslant \infty$, then*

$$\begin{aligned} \mathrm{W}^{k_2,p}(\Omega) \hookrightarrow \mathrm{W}^{k_1,p}(\Omega) \hookrightarrow \mathrm{W}^{1,p}(\Omega), \\ \mathrm{W}_0^{k_2,p}(\Omega) \hookrightarrow \mathrm{W}_0^{k_1,p}(\Omega) \hookrightarrow \mathrm{W}_0^{1,p}(\Omega), \\ \mathrm{W}^{k,\infty}(\Omega) \hookrightarrow \mathrm{W}^{k,p_1}(\Omega) \hookrightarrow \mathrm{W}^{k,p_2}(\Omega) \hookrightarrow \mathrm{W}^{1,1}(\Omega), \\ \mathrm{W}_0^{k,\infty}(\Omega) \hookrightarrow \mathrm{W}_0^{k,p_1}(\Omega) \hookrightarrow \mathrm{W}_0^{k,p_2}(\Omega) \hookrightarrow \mathrm{W}_0^{k,1}(\Omega). \end{aligned}$$

2.9.2 Sobolev Spaces $\mathrm{H}^k(\Omega)$ and $\mathrm{H}_0^k(\Omega)$

Take $p = 2$. The norm (2.6) induces the following scalar product in $\mathrm{W}^{k,2}(\Omega)$

$$(u, v)_{k,\Omega} = \sum_{|\boldsymbol{\alpha}|_1 \leqslant k} \int_\Omega (D^{\boldsymbol{\alpha}} u)(\mathbf{x})\,(D^{\boldsymbol{\alpha}} v)(\mathbf{x})\ \mathrm{d}\mathbf{x}. \tag{2.7}$$

It holds that

$$(u, u)_{k,\Omega} = \|u\|_{k,2,\Omega}^2 .$$

The space $\mathrm{W}^{k,2}(\Omega)$ is a Hilbert space and it is frequently denoted by the symbol $\mathrm{H}^k(\Omega)$. The Hilbert space $\mathrm{W}_0^{k,2}(\Omega)$ is also denoted by $\mathrm{H}_0^k(\Omega)$.

2.9.3 Sobolev Spaces $\mathrm{H}^{k,p}(\Omega)$

Now, a linear space $\mathrm{H}^{k,p}(\Omega)$ for $k \in \mathbb{N} \cup \{0\}$ and $p \in [1, \infty]$ is introduced.

Definition 2.9.2. *Suppose that Ω is a nonempty open subset of $\mathbb{R}^d$. The function $u \in \mathrm{L}^p(\Omega)$ is an element of $\mathrm{H}^{k,p}(\Omega)$ if all generalized derivatives of u up to the order k exist and belong to $\mathrm{L}^p(\Omega)$. The norm in this space is*

$$||u||_{\mathrm{H}^{k,p}(\Omega)}^p = \sum_{|\boldsymbol{\alpha}|_1 \leqslant k} \left\| u^{(\boldsymbol{\alpha})} \right\|_{p,\Omega}^p .$$

The closure of $\mathrm{C}_0^\infty(\Omega)$ in the norm of $\mathrm{H}^{k,p}(\Omega)$ is denoted by $\mathrm{H}_0^{k,p}(\Omega)$. The norm in $\mathrm{H}_0^{k,p}(\Omega)$ is the same as the one in $\mathrm{H}^{k,p}(\Omega)$. Moreover, $\mathrm{H}^{0,p}(\Omega) = \mathrm{L}^p(\Omega)$.

The following result is proved in a paper by Meyers and Serrin [34].

Theorem 2.9.3. *Let Ω be any open bounded set in $\mathbb{R}^d$. Then $\mathrm{C}^\infty(\overline{\Omega}) \cap \mathrm{H}^{k,p}(\Omega)$ is dense in $\mathrm{H}^{k,p}(\Omega)$ for $p \in [1, \infty)$ and $k \in \mathbb{N}$.*

2.9.4 Relationship between the different Sobolev spaces

There exists a very close relationship between the spaces $\mathrm{H}^{k,p}(\Omega)$ and $\mathrm{W}^{k,p}(\Omega)$, and between the spaces $\mathrm{H}_0^{k,p}(\Omega)$ and $\mathrm{W}_0^{k,p}(\Omega)$, cf. [24, Theorem 5.4.4 and 5.5.9].

Theorem 2.9.4. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. It holds that*

(i) $\mathrm{H}_0^{k,p}(\Omega) \rightleftarrows \mathrm{W}_0^{k,p}(\Omega)$ *for* $k \in \mathbb{N} \cup \{0\}$ *and* $p \in [1, \infty]$,

(ii) $\mathrm{H}^{k,p}(\Omega) \rightleftarrows \mathrm{W}^{k,p}(\Omega)$ *for* $k \in \mathbb{N} \cup \{0\}$ *and* $p \in [1, \infty)$ *if the domain* Ω *has a Lipschitz continuous boundary.*

Thus the properties of both spaces can be interchanged if the domain Ω has a sufficiently well-behaved boundary.

Corollary 2.9.1. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ and assume that $p \in [1, \infty)$ and $k, l \in \mathbb{N}$. The space $\mathrm{C}^\infty(\overline{\Omega})$ is dense in $\mathrm{H}^{k,p}(\Omega)$. The Schwartz space $\mathrm{C}_0^\infty(\Omega)$ is dense in $\mathrm{H}_0^{k,p}(\Omega)$ but not in $\mathrm{H}^{k,p}(\Omega)$. Moreover, let $1 \leqslant k < l$, then $\mathrm{H}^{l,p}(\Omega) \subset \mathrm{H}^{k,p}(\Omega)$.*

Example 2.9.5. *Assume that $\Omega = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}|_{\mathrm{e}} < 1\}$ is the open unit ball ($d \geqslant 2$) and let $v(\mathbf{x}) = |\mathbf{x}|_{\mathrm{e}}^\lambda$, where λ is real. Moreover, let $p \in [1, \infty)$. It is possible to show that $v \in \mathrm{W}^{k,p}(\Omega)$ if $\lambda > k - \frac{d}{p}$. Consider*

$$u(\mathbf{x}) = \sum_{k=1}^{\infty} \frac{1}{2^k} \left|\mathbf{x} - \boldsymbol{r}_k\right|_{\mathrm{e}}^{-\alpha}, \qquad \mathbf{x} \in \Omega, \quad \alpha > 0,$$

where $\{\boldsymbol{r}_k\}_{k=1}^\infty$ is a countable, dense subset of Ω. This series converges uniformly. If $0 < \alpha < \frac{d-kp}{p}$, then $u \in \mathrm{W}^{k,p}(\Omega)$. However, this function is unbounded on each open subset of Ω. This illustrates that although a function u belonging to a Sobolev space possesses certain smoothness properties, it can still behave badly in other ways.

Example 2.9.6. *Let $\Omega \subset \mathbb{R}^2$ be the open ball $B(0, \beta)$ in $\mathbb{R}^2$ with $\beta \in (0, 1)$ fixed. The unbounded function $v(\mathbf{x}) = \ln\left(\ln\left(\frac{1}{|\mathbf{x}|_{\mathrm{e}}}\right)\right)$ belongs to $\mathrm{H}^1(\Omega)$.*

Example 2.9.7. *The following interesting example can be found in [24, Example 5.2.7 and 5.4.8]. Let $\Omega = (0, 1) \times (-1, 1)$, $\Omega_1 = (0, 1) \times (0, 1)$, $\Omega_2 = (0, 1) \times (-1, 0)$ and put*

$$u(x, y) = \begin{cases} 0 & \text{if} \quad y \leqslant 0, \\ 1 & \text{if} \quad y > 0. \end{cases}$$

Then, for all $p \in [1, \infty]$, it holds that $u \in \mathrm{W}^{1,p}(\Omega_1)$, $u \in \mathrm{W}^{1,p}(\Omega_2)$, $u \in \mathrm{H}^{1,p}(\Omega_1 \cup \Omega_2)$, but $u \notin \mathrm{W}^{1,p}(\Omega)$, $u \notin \mathrm{W}^{1,p}(\Omega_1 \cup \Omega_2)$ and $u \notin \mathrm{H}^{1,p}(\Omega)$.

Remark 2.9.1. *The Schwartz space $\mathcal{D}(\Omega)$ is dense in $\mathrm{L}^p(\Omega)$ for all $p \in [1, +\infty)$. Let $k \in \mathbb{N} \cup \{0\}$. The spaces $\mathrm{H}_0^{k,p}(\Omega)$ and $\mathrm{H}^{k,p}(\Omega)$ are dense in $\mathrm{L}^p(\Omega)$ by the relation*

$$\mathcal{D}(\Omega) \subset \mathrm{H}_0^{k,p}(\Omega) \subset \mathrm{H}^{k,p}(\Omega) \subset \mathrm{L}^p(\Omega).$$

Remark that $\mathrm{H}_0^{k,p}(\Omega)$ is not dense in $\mathrm{H}^{k,p}(\Omega)$.

2.9.5 Fractional spaces $\mathrm{W}^{k,p}(\Omega)$

In this subsection, the definition of a Sobolev space $\mathrm{W}^{k,p}(\Omega)$ for a non-integer k is given. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary. The floor function of a real number k, denoted by $\lfloor k \rfloor$, is a function that returns the largest integer smaller than or equal to k. Recall the inequality $\lfloor k \rfloor \leqslant k < \lfloor k \rfloor + 1$.

Definition 2.9.3. *Take $k > 0$, $k \notin \mathbb{Z}$, $p \in [1, +\infty)$. The symbol $\mathrm{W}^{k,p}(\Omega)$ denotes the set of all $u \in \mathrm{W}^{\lfloor k \rfloor,p}(\Omega)$, such that*

$$\|u\|_{k,p,\Omega}^p = \|u\|_{\lfloor k \rfloor,p,\Omega}^p + \sum_{|\boldsymbol{\alpha}|=\lfloor k \rfloor} \int_\Omega \int_\Omega \frac{|D^{\boldsymbol{\alpha}} u(\mathbf{x}) - D^{\boldsymbol{\alpha}} u(\mathbf{y})|^p}{|\mathbf{x}-\mathbf{y}|^{d+p(k-\lfloor k \rfloor)}} \,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} < \infty.$$

This formula defines the norm in $\mathrm{W}^{k,p}(\Omega)$.

From this definition, it is clear that $\mathrm{W}^{k,p}(\Omega) \hookrightarrow \mathrm{W}^{\lfloor k \rfloor,p}(\Omega)$. Fractional Sobolev spaces are linear normed spaces with similar properties as those of $\mathrm{W}^{\lfloor k \rfloor,p}(\Omega)$.

Theorem 2.9.8. *If $k > 0, k \notin \mathbb{Z}$ and $p \in [1, +\infty)$, then*

(i) $\mathrm{W}^{k,p}(\Omega)$ *is a separable Banach space.*

(ii) $\mathrm{W}^{k,p}(\Omega)$ *is reflexive if $p > 1$.*

2.9.6 Dual Sobolev spaces

Definition 2.9.4. *Assume that $1 \leqslant p < \infty$. Let $k \in \mathbb{Z}$ be a negative number. Take $q \in \mathbb{R}$ such that $\frac{1}{p} + \frac{1}{q} = 1$. The Sobolev space $\mathrm{W}^{k,p}(\Omega)$ is defined as the dual space to $\mathrm{W}^{-k,q}(\Omega)$, i.e.,*

$$\mathrm{W}^{k,p}(\Omega) = \left(\mathrm{W}^{-k,q}(\Omega)\right)^*.$$

The norm is

$$\|u\|_{\mathrm{W}^{k,p}(\Omega)} = \sup_{\substack{v \in \mathrm{W}^{-k,q}(\Omega) \\ v \neq 0}} \frac{\langle u, v \rangle}{\|v\|_{\mathrm{W}^{-k,q}(\Omega)}}.$$

2.9.7 The space $\mathrm{L}^p\left((0,T), X\right)$

As for the space of continuous functions, it is also possible to define Lebesgue spaces for abstract functions whose values are elements of a general Banach space X. The generalization of the Lebesgue integral for abstract functions is the so called Bochner integral. For more information, the reader is referred to [35, p. 124], [24, Section 2.19] or [36, p. 20]. The following spaces are crucial in the analysis of time-dependent partial differential equations.

Definition 2.9.5 (Bochner spaces)**.** *Let X be a Banach space with norm $\|\cdot\|_X$ over the field $\mathbb{F}$ and $0 < T < \infty$.*

- *The space* $\mathrm{L}^p((0,T),X)$ *with* $1 \leqslant p < \infty$ *consists of all measurable functions* $u:(0,T] \to X$ *for which*

$$\|u\|_{\mathrm{L}^p((0,T),X)} = \left(\int_0^T \|u(t)\|_X^p \,\mathrm{d}t\right)^{\frac{1}{p}} < \infty. \tag{2.8}$$

  holds. Two functions $f,g \in \mathrm{L}^p((0,T),X)$ *are equal iff* $f(t) = g(t)$ *in* X *for a.a.* $t \in (0,T)$.
- *The space* $\mathrm{L}^\infty((0,T),X)$ *consists of all measurable functions* $u:(0,T) \to X$ *that are essentially bounded, i.e. there exists a number* B *such that*

$$\|u(t)\|_X \leqslant B \qquad \text{for a.a.} \quad t \in (0,T).$$

  Precisely all the numbers B *with this property are called essential bounds of* u. *Moreover, we set*

$$\|u\|_{\mathrm{L}^\infty((0,T),X)} = \inf\{B\}, \tag{2.9}$$

  where the infimum is taken over all the essential bounds of u.

The subsequent theorem and definition are necessary for understanding evolution equations, see [7, Section 23.5].

Theorem 2.9.9. *Let* X *be a Banach space. Then it follows from* $v \in \mathrm{L}^1((0,T),X)$ *and*

$$\int_0^T v(t)\varphi(t)\,\mathrm{d}t = 0, \qquad \forall \varphi \in \mathrm{C}_0^\infty((0,T)),$$

that $v = 0$ *in* $\mathrm{L}^1((0,T),X)$, *i.e.*

$$v(t) = 0 \qquad \text{for a.a.} \quad t \in (0,T).$$

Definition 2.9.6. *Let* X *be a Banach space and assume that* $u,v \in \mathrm{L}^1((0,T),X)$. *The function* $v = u^{(n)}$ *is called the* n*-th weak or generalized derivative of the function* u *on* $(0,T)$ *if*

$$\int_0^T v(t)\varphi(t)\,\mathrm{d}t = (-1)^n \int_0^T u(t)\varphi^{(n)}(t)\,\mathrm{d}t$$

holds for all $\varphi \in \mathrm{C}_0^\infty((0,T))$.

The Bochner spaces play an essential role in the theory of evolution equations because of their interesting properties, see below.

Theorem 2.9.10. *Let* $1 \leqslant p < \infty$ *and let* X *and* Y *be Banach spaces over* $\mathbb{R}$. *Then, it holds that* *(i)* *the space* $\mathrm{L}^p\left((0,T),X\right)$ *is a Banach space. Moreover, the set of all step functions* $u:[0,T]\to X$ *is dense in* $\mathrm{L}^p\left((0,T),X\right)$. *The set of all polynomials* $w:[0,T]\to X$, *i.e.*

$$w(t)=a_0+a_1t+\cdots+a_nt^n$$

with $a_i\in X$ *for all* i *and* $n=0,1,\ldots$ *is dense in* $\mathrm{C}\left([0,T],X\right)$ *and* $\mathrm{L}^p\left((0,T),X\right)$*;*

(ii) *the space* $\mathrm{C}\left([0,T],X\right)$ *is dense in* $\mathrm{L}^p\left((0,T),X\right)$ *and the embedding*

$$\mathrm{C}\left([0,T],X\right)\subset \mathrm{L}^p\left((0,T),X\right),\qquad 1\leqslant p<\infty,$$

is continuous;

(iii) *the space* $\mathrm{L}^p\left((0,T),X\right)$ *is separable if* X *is separable and* $p\in[1,+\infty)$*;*

(iv) *the space* $\mathrm{L}^p\left((0,T),X\right)$ *with* $p\in(1,+\infty)$ *is reflexive if* X *is reflexive and separable. Moreover,*

$$\mathrm{L}^p\left((0,T),X\right)^*\cong \mathrm{L}^q\left((0,T),X^*\right)\qquad \textit{with } \frac{1}{p}+\frac{1}{q}=1;$$

(v) *if the embedding* $X\subset Y$ *is continuous, then the embedding*

$$\mathrm{L}^r\left((0,T),X\right)\subset \mathrm{L}^q\left((0,T),Y\right),\qquad 1\leqslant q\leqslant r<\infty,$$

is continuous;

(vi) *the embedding*

$$\mathrm{L}^\infty\left((0,T),X\right)\subset \mathrm{L}^p\left((0,T),X\right),\qquad 1\leqslant p\leqslant\infty,$$

is continuous;

(vii) *if the function* $u:(0,T)\to X$ *is continuous a.e. and bounded, i.e.*

$$\sup_{t\in(0,T)}\|u(t)\|_X<\infty,$$

then $u\in\mathrm{L}^p\left((0,T),X\right)$ *for all* $1\leqslant p\leqslant\infty$*;*

(viii) *a function* $u:[0,T]\to X$ *with* $\partial_t u\in\mathrm{L}^p\left((0,T),X\right)$ *for fixed* $p\in[1,\infty)$ *belongs to* $\mathrm{C}\left([0,T],X\right)$*;*

(ix) *if* X *is a Hilbert space with scalar product* $\langle\cdot,\cdot\rangle_X$, *then* $\mathrm{L}^2\left((0,T),X\right)$ *is also a Hilbert space with the scalar product*

$$\langle u,v\rangle_{\mathrm{L}^2((0,T),X)}=\int_0^T\langle u(t),v(t)\rangle_X\,\mathrm{d}t.$$

Remark 2.9.2. *The proof of (viii) follows from the frequently used estimate*

$$\begin{aligned}\|u(t)-u(s)\|_X&\leqslant\int_s^t\|\partial_t u(\xi)\|_X\,\mathrm{d}\xi\\&\leqslant\left(\int_s^t 1^{\frac{p}{p-1}}\,\mathrm{d}\xi\right)^{\frac{p-1}{p}}\left(\int_s^t\|\partial_t u(\xi)\|_X^p\,\mathrm{d}\xi\right)^{\frac{1}{p}}\\&\leqslant C\,|t-s|^{\frac{p-1}{p}},\end{aligned}$$

which implies the Hölder continuity of u.

The following notations can be introduced

$$\langle v, u\rangle_{\mathrm{L}^p((0,T),X)} = \int_0^T \langle v(t), u(t)\rangle_X \, \mathrm{d}t,$$
$$\|v\|_{\mathrm{L}^q((0,T),X^*)} = \left(\int_0^T \|v(t)\|_{X^*}^q \, \mathrm{d}t\right)^{\frac{1}{q}}$$

for all $u \in \mathrm{L}^p\left((0,T),X\right)$ and $v \in \mathrm{L}^q\left((0,T),X^*\right)$ with $\frac{1}{p}+\frac{1}{q}=1$. Remember that $\langle v(t), u(t)\rangle_X$ denotes the value of the functional $v(t)$ at the point $u(t)$. The following results are worth mentioning [7, Proposition 23.9 and 23.19].

Lemma 2.9.2 (Limit relations for integrals: A)**.** *Let X be a Banach space. Furthermore, let $1 < p, q < \infty$, $\frac{1}{p}+\frac{1}{q}=1$ and $0 \leqslant t \leqslant T < +\infty$. Then the following statements are valid:*

(i) If $u \in \mathrm{L}^p\left((0,T),X\right)$, then

$$\left\langle x^*, \int_0^t u(s)\ \mathrm{d}s \right\rangle_X = \int_0^t \langle x^*, u(s)\rangle_X\ \mathrm{d}s \qquad \textit{for all} \quad x^* \in X^*.$$

(ii) If $u \in \mathrm{L}^p\left((0,T),X^*\right)$, then

$$\left\langle \int_0^t u(s)\ \mathrm{d}s, x \right\rangle_X = \int_0^t \langle u(s), x\rangle_X\ \mathrm{d}s \qquad \textit{for all} \quad x \in X.$$

(iii) From $u_n \to u$ in $\mathrm{L}^p\left((0,T),X\right)$ as $n \to \infty$ if follows that

$$\int_0^t u_n(s)\ \mathrm{d}s \to \int_0^t u(s)\ \mathrm{d}s \qquad \textit{in } X \quad \textit{as } n \to \infty.$$

Lemma 2.9.3 (Limit relations for integrals: B)**.** *Let X be a reflexive and separable Banach space. Furthermore, let $1 < p, q < \infty$, $\frac{1}{p}+\frac{1}{q}=1$ and $0 \leqslant t \leqslant T < +\infty$. Then the following statements are valid:*

(i) From

$$\begin{aligned} u_n &\to u \quad \textit{in } \mathrm{L}^p\left((0,T),X\right) && \textit{as } n \to \infty,\\ f_n &\rightharpoonup f \quad \textit{in } \mathrm{L}^q\left((0,T),X^*\right) && \textit{as } n \to \infty, \end{aligned}$$

it follows that

$$\int_0^t \langle f_n(s), u_n(s)\rangle_X\ \mathrm{d}s \to \int_0^t \langle f(s), u(s)\rangle_X \quad \textit{as } n \to \infty.$$

(ii) From

$$\begin{aligned} u_n &\rightharpoonup u \quad \textit{in } \mathrm{L}^p\left((0,T),X\right) && \textit{as } n \to \infty,\\ f_n &\to f \quad \textit{in } \mathrm{L}^q\left((0,T),X^*\right) && \textit{as } n \to \infty, \end{aligned}$$

it follows that

$$\int_0^t \langle f_n(s), u_n(s)\rangle_X \ \mathrm{d}s \to \int_0^t \langle f(s), u(s)\rangle_X \quad \text{as } n \to \infty.$$

Theorem 2.9.11 (Generalized derivatives and weak convergence: B)**.** *Let Y and Z be Banach spaces such that $Y \hookrightarrow Z$. Then it follows from*

$$u_k^{(n)} = v_k \quad \text{on } (0,T) \qquad \text{for all } k \text{ and fixed } n \geqslant 1$$

and

$$u_k \rightharpoonup u \quad \text{in } \mathrm{L}^p\left((0,T), Y\right) \quad \text{as } k \to \infty,$$
$$v_k \rightharpoonup v \quad \text{in } \mathrm{L}^q\left((0,T), Z\right) \quad \text{as } k \to \infty, \quad 1 \leqslant p, q < +\infty, \frac{1}{p} + \frac{1}{q} = 1,$$

that

$$u^{(n)} = v \quad \text{on } (0,T).$$

The next theorem is about an abstract Lipschitz continuous function [7, Corollary 23.22] and is analogous to Definition 2.2.5.

Lemma 2.9.4. *Let H be a real Hilbert space and let $u : [0,T] \to H$ be Lipschitz continuous. Then the following holds:*

(i) For a.a. $t \in [0,T]$, the function u has a derivative,

$$u'(t) = \lim_{h\to 0} \frac{u(t+h) - u(t)}{h} \qquad \text{in } H$$

and

$$u(t) = u(0) + \int_0^t u'(s) \ \mathrm{d}s \qquad \text{in } H \text{ for all } t \in [0,T].$$

(ii) For a.a. $t \in [0,T]$,

$$\|u'(t)\|_H \leqslant L,$$

and u' is the generalized derivative of u on $(0,T)$.

Definition 2.9.7. *Let H be a real Hilbert space. The space of Lipschitz continuous functions $u : [0,T] \to H$ is denoted by $\mathrm{Lip}([0,T], H)$.*

The following theorem concerns the interchange of the integral and limit signs.

Theorem 2.9.12 (Lebesgue dominated convergence theorem)**.** *Let Y be a Banach space with norm $\|\cdot\|_Y$ and $\{f_n : M \to Y\}_{n\in\mathbb{N}}$ a sequence of measurable functions on $M \subset \mathbb{R}^d$. Then*

$$\lim_{n\to\infty} \int_M f_n(x)\,\mathrm{d}x = \int_M \lim_{n\to\infty} f_n(x)\,\mathrm{d}x,$$

where all the integrals and limits exist, provided the following conditions are satisfied:

(i) $\|f_n(x)\|_Y \leqslant g(x)$ *for a.a.* $x \in M$ *and all* $n \in \mathbb{N}$, *and* $\int_M g(x)\,\mathrm{d}x$ *exists or*
$\|f_n(x)\|_Y \leqslant g_n(x)$ *for a.a.* $x \in M$ *and all* $n \in \mathbb{N}$. *All the functions* $g_n : M \to \mathbb{R}$ *are integrable and* g_n *converges to* $g : M \to \mathbb{R}$ *a.e. in* M *as* $n \to \infty$ *along with*

$$\lim_{n\to\infty} \int_M g_n(x)\,\mathrm{d}x = \int_M g(x)\,\mathrm{d}x;$$

(ii) $\lim_{n\to\infty} f_n(x)$ *exists for a.a.* $x \in M$.

2.9.8 Sobolev-Bochner spaces $\mathrm{W}^{1,2,2}([0,T];V_1,V_2)$

Now, the Sobolev-Bochner space $\mathrm{W}^{1,2,2}([0,T];V_1,V_2)$ is defined, see also [7, Chapter 23.6]. First, the definition of evolution triple is given [7, Chapter 23.4]. The use of evolution triples (sometimes called Gelfand triples) has a central role in the formulation of elliptic, parabolic and hyperbolic PDEs in abstract spaces.

The following theorems are crucial in determining the structure of a Gelfand triple. The first theorem gives as sufficient condition for a dual operator to be compact. The definition of dual operator can be found in Definition 2.4.35. The Arzelà-Ascoli theorem plays an essential role in the proof, see [14, Theorem 5.11-2]. The second theorem gives a crucial fact from operator theory, cf. [14, Theorem 5.11-3] or [37, Theorem 2.9.1].

Theorem 2.9.13. *Let* X *and* Y *two real normed vector spaces and let* $A : X \to Y$ *be a compact linear operator. Then the dual operator* $A^* : Y^* \to X^*$ *is also compact.*

Theorem 2.9.14. *Suppose that* X *and* Y *are normed linear spaces, and that* $A \in \mathcal{L}(X,Y)$. *Then* $R(A)$ *is dense in* Y *iff the adjoint operator* $A^* : Y^* \to X^*$ *is injective.*

Using the canonical injection, the following corollary follows now immediately form the definition of dual operator and Theorem 2.9.13.

Corollary 2.9.2. *Consider two normed linear spaces* X *and* Y. *It holds that*

- $X \hookrightarrow Y$ *implies that* $Y^* \hookrightarrow X^*$;
- $X \hookrightarrow\hookrightarrow Y$ *implies that* $Y^* \hookrightarrow\hookrightarrow X^*$.

Now, the definition of evolution triple can be introduced.

Definition 2.9.8. *An evolution triple*

$$V \subseteq H \cong H^* \subseteq V^*$$

is understood to satisfy the following:

(i) *V is a real, separable and reflexive Banach space,*
(ii) *H is a real, separable Hilbert space with inner product $(\cdot,\cdot)_H$,*
(iii) *The embedding $V \subseteq H$ is continuous, i.e. $V \hookrightarrow H$, and V is dense in H.*

Remark 2.9.3. *An evolution triple can also be denoted as*

$$V \hookrightarrow H \cong H^* \hookrightarrow V^*,$$

due to the fact that the adjoint mapping $i^* : H^* \cong H \to V^*$ of the embedding $i : V \to H$ is continuous and injective, i.e.

$$u_1 \neq u_2 \quad \Rightarrow \quad i^* u_1 \neq i^* u_2 \Leftrightarrow \exists v \in V : \langle u_1, v\rangle \neq \langle u_2, v\rangle.$$

A list of additional remarks can be made:

- *The Hilbert space H in an evolution triple is called the pivot. The identification $H^* \cong H$ follows from the Riesz' representation theorem 2.4.25, i.e. every $h^* \in H^*$ can be represented by*

$$h^*(h) = \left(\hat{h}^*, h\right)_H, \qquad h \in H,$$

  for some unique $\hat{h}^* \in H$;
- *It is agreed to identify $i^* h$ with h if $h \in H$. Then, for $h \in H$, one can define $i^*(h) = h \in V^*$ by*

$$i^*(h)(v) = h(v) = (h, v)_H$$

  for $v \in V$;
- *Theorem 2.9.14 implies that in Definition 2.9.8 the space V is dense in H;*
- *The embedding $H \subset V^*$ is dense by Theorem 2.9.14 and the reflexivity of V;*
- *For any $h \in H$ and $v \in V$, it holds that*

$$(h, v)_H = \langle h, v\rangle_{H^*\times H} = \langle h, iv\rangle_{H^*\times H} = \langle i^* h, v\rangle_{V^*\times V} = \langle h, v\rangle_{V^*\times V}.$$

  These equalities follow subsequently form the identification of H with H^*, the embedding $V \subset H$, the definition of the adjoint operator i^* and the identification of $i^* h$ with h.

The last two remarks say that the duality pairing on $V^* \times V$ can be viewed as the continuous extension of the inner product $(\cdot,\cdot)_H$ acting on $H \times V$ [37, Theorem 2.9.2].

Theorem 2.9.15. *Let $V \subseteq H \cong H^* \subseteq V^*$ be an evolution triple.*
(i) *For any $v^* \in V^*$, there exists a sequence $\{h_n\}_{n\in\mathbb{N}} \subset H$ such that $h_n \to v^*$ in V^* and*

$$\langle v^*, v\rangle_{V^*\times V} = \lim_{n\to\infty} (h_n, v)_H, \qquad \forall v \in V,$$

(ii) For any $h \in H$ and $v \in V$, it holds that

$$\langle h, v \rangle_{V^* \times V} = (h, v)_H \,.$$

Definition 2.9.9. *Let V_1 and V_2 be Banach spaces with $V_1 \subset V_2$. Define the space*

$$\mathrm{W}^{1,2,2}([0,T]; V_1, V_2) := \left\{ u \in \mathrm{L}^2\left((0,T), V_1\right) : \frac{\mathrm{d}u}{\mathrm{d}t} \in \mathrm{L}^2\left((0,T), V_2\right) \right\},$$

with $\frac{\mathrm{d}u}{\mathrm{d}t}$ denoting the generalized derivative of u with respect to t. The space $\mathrm{W}^{1,2,2}([0,T]; V_1, V_2)$ is a Banach space if equipped with the norm

$$\|u\|_{\mathrm{W}^{1,2,2}([0,T];V_1,V_2)} := \|u\|_{\mathrm{L}^2((0,T),V_1)} + \left\| \frac{\mathrm{d}u}{\mathrm{d}t} \right\|_{\mathrm{L}^2((0,T),V_2)} .$$

This space is often denoted by $\mathrm{H}^1((0,T), X)$ when $V_1 = V_2 = X$.

This space has interesting properties that are contained in the following lemma [38, Lemma 7.1, 7.2 and 7.3].

Lemma 2.9.5.

(i) Let $V_1 \hookrightarrow V_2$. Then $\mathrm{W}^{1,2,2}([0,T]; V_1, V_2) \hookrightarrow \mathrm{C}\left([0,T], V_2\right)$. This implies that if $u \in \mathrm{W}^{1,2,2}([0,T]; V_1, V_2)$, there exists a uniquely determined continuous function $u_1 : [0,T] \to V_2$ which coincides a.e. in $[0,T]$ with the original function u.

(ii) Let $V_1 \hookrightarrow V_2$. Then the space $\mathrm{C}^1\left([0,T], V_1\right)$ is dense $\mathrm{W}^{1,2,2}([0,T]; V_1, V_2)$.

(iii) Let $V \hookrightarrow H \cong H^* \hookrightarrow V^*$ be an evolution triple. Then

$$\mathrm{W}^{1,2,2}([0,T]; V, V^*) \hookrightarrow \mathrm{C}\left([0,T], H\right).$$

Moreover, for all $u, v \in \mathrm{W}^{1,2,2}([0,T]; V, V^*)$ and any $0 \leqslant t_1 \leqslant t_2 \leqslant T$, the following generalized integration by parts formula holds true:

$$\begin{aligned}(u(t_2), v(t_2))_H &- (u(t_1), v(t_1))_H \\ &= \int_{t_1}^{t_2} \left\langle \frac{\mathrm{d}u(t)}{\mathrm{d}t}, v(t) \right\rangle_{V^* \times V} \mathrm{d}t + \int_{t_1}^{t_2} \left\langle u(t), \frac{\mathrm{d}v(t)}{\mathrm{d}t} \right\rangle_{V^* \times V} \mathrm{d}t,\end{aligned}$$

where $u(t)$ and $v(t)$ are values of the continuous functions $u, v : [0,T] \to H$ at $t \in [0,T]$. For $u = v$, this formula gives

$$\tfrac{1}{2} \|u(t_2)\|_H^2 - \tfrac{1}{2} \|u(t_1)\|_H^2 = \int_{t_1}^{t_2} \left\langle \frac{\mathrm{d}u(t)}{\mathrm{d}t}, u(t) \right\rangle_{V^* \times V} \mathrm{d}t.$$

The function $t \mapsto \frac{1}{2} \|u(t)\|_H^2$ is absolutely continuous. Hence, its derivative exists a.e. in $[0,T]$, i.e.

$$\tfrac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|u(t)\|_H^2 = \left\langle \frac{\mathrm{d}u(t)}{\mathrm{d}t}, u(t) \right\rangle_{V^* \times V} \qquad \text{for a.a. } t \in [0,T].$$

2.9.9 Sobolev embedding theorems

The relationship between the Sobolev spaces, the Lebesgue spaces and the spaces of continuous functions are formulated in the so-called ‘embeddings theorems’, see [24, Section 5.6, 5.7 and 5.8]. A distinction is made between continuous and compact embeddings, see Section 2.4.5.

Theorem 2.9.16 (1D-case)**.** *Take* $\Omega = (a, b)$ *and* $p \in [1, \infty)$*. Then* $\mathrm{H}^{1,p}(\Omega)$ *is the set of absolutely continuous functions, which derivative exists a.e. in* Ω *and this derivative belongs to* $\mathrm{L}^p(\Omega)$*, i.e.* $\mathrm{H}^{1,p}(a, b) \subset \mathrm{C}([a, b])$*.*

Example 2.9.17. *Let* $\Omega = (a, b) \subset \mathbb{R}$*. Then*

$$\mathrm{H}^1(\Omega) \subset \mathrm{C}(\overline{\Omega}) \subset \mathrm{C}(\Omega) \text{ and } \mathrm{H}^2(\Omega) \subset \mathrm{C}^1(\overline{\Omega}) \subset \mathrm{C}^1(\Omega).$$

This is not valid in higher dimensions, see Example 2.9.5. The following example illustrates that the inclusion above cannot be reversed: for the function $f : [0, 1] \to \mathbb{R} : x \mapsto x^\alpha$*,* $\alpha \in [0, \infty)$*, it holds that*

$$f \in \mathrm{H}^1(0, 1) \Leftrightarrow \alpha > \tfrac{1}{2},$$

although $f \in \mathrm{C}([0, 1])$ *for every* $\alpha \in [0, \infty)$*.*

A function in $\mathrm{H}^{k,p}(\Omega)$ belongs to $\mathrm{L}^p(\Omega)$. However, the following theorem (embedding to $\mathrm{L}^q(\Omega)$) sometimes gives higher regularity.

Theorem 2.9.18 (Embedding to $\mathrm{L}^q(\Omega)$)**.** *Let* $\Omega \subset \mathbb{R}^d$ *be a bounded Lipschitz domain. Assume that* $d \geqslant 2$ *and* $1 \leqslant p < \infty$*. Then*

$$\mathrm{H}^{k,p}(\Omega) \hookrightarrow \mathrm{L}^q(\Omega),$$

if at least one of the following conditions is fulfilled

(i) $kp < d$*,* $1 \leqslant q \leqslant \frac{dp}{d-kp}$*,*

(ii) $kp = d$*,* $q \in [1, \infty)$*.*

The following theorem covers the case $kp > d$.

Theorem 2.9.19 (Embedding into continuous functions)**.** *Let* $\Omega \subset \mathbb{R}^d$ *be a bounded Lipschitz domain. Assume that* $d \geqslant 2$ *and* $1 \leqslant p < \infty$*,* $k \in \mathbb{N}$*,* $kp > d$*. Then*

$$\mathrm{H}^{k,p}(\Omega) \hookrightarrow \mathrm{C}^{0,\lambda}(\overline{\Omega}),$$

where

(i) $\lambda = k - \frac{d}{p}$ *if* $k - \frac{d}{p} < 1$*,*

(ii) $\lambda \in (0, 1)$ *if* $k - \frac{d}{p} = 1$ *and* $p > 1$*,*

(iii) $\lambda = 1$ *if* $k - \frac{d}{p} > 1$*.*

An element of a Sobolev space is an equivalence class of functions that are equal a.e. in Ω. Thus, be careful with the interpretation of these continuous embeddings (and the compact embeddings in the following theorems). For instance, the embedding $\mathrm{H}^{k,p}(\Omega) \hookrightarrow \mathrm{C}^{0,\lambda}(\overline{\Omega})$ means that there exists a constant $C > 0$ such that, in each equivalence class of the space $\mathrm{H}^{k,p}(\Omega)$, there is a (unique) representative v that belongs to the space $\mathrm{C}^{0,\lambda}(\overline{\Omega})$ and satisfies $\|v\|_{\mathrm{C}^{0,\lambda}(\overline{\Omega})} \leqslant C\,\|v\|_{\mathrm{H}^{k,p}(\Omega)}$.

Theorem 2.9.20 (Compact embedding to $\mathrm{L}^q(\Omega)$)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary. Assume that $d \geqslant 2$ and $1 \leqslant p < \infty$. Then*

$$\mathrm{H}^{k,p}(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^q(\Omega),$$

if at least one of the following conditions is fulfilled

(i) $kp < d$, $1 \leqslant q < \frac{dp}{d-kp}$,

(ii) $kp = d$, $q \in [1, \infty)$.

The next theorem covers the case $kp > d$.

Theorem 2.9.21 (Compact embedding to continuous functions)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary. Assume that $d \geqslant 2$ and $1 \leqslant p < \infty$, $k \in \mathbb{N}$, $kp > d$. Then*

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

Moreover,

$$\mathrm{H}^{k,p}(\Omega) \hookrightarrow\hookrightarrow \mathrm{C}^{0,\lambda}(\overline{\Omega}),$$

where

(i) $\lambda < k - \frac{d}{p}$ *if* $k - \frac{d}{p} < 1$,

(ii) $\lambda \in (0,1)$ *if* $k - \frac{d}{p} \geqslant 1$.

The following theorem is a consequence of the previous results and can be found in [1, p. 272] or [29, Theorem 6.3]. This theorem is independent of the dimension.

Theorem 2.9.22 (Rellich-Kondrachov Compactness Theorem)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. The space $\mathrm{H}^{1,p}(\Omega)$ is compactly embedded in $\mathrm{L}^p(\Omega)$ for $p \in [1, +\infty]$. Also $\mathrm{H}_0^{1,p}(\Omega)$ is compactly embedded in $\mathrm{L}^p(\Omega)$. In fact,*

$$\mathrm{H}^{1,p}(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^q(\Omega) \qquad \textit{with} \quad 1 \leqslant q \leqslant p.$$

The embedding theorems for fractional-order spaces are not as powerful as for integer-order spaces. The following result can be found in [39, Theorem 3.7].

Theorem 2.9.23. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. If $t, s \in \mathbb{R}$ such that $0 \leqslant t < s < \infty$, then the embedding*

$$\mathrm{H}^s(\Omega) \hookrightarrow\hookrightarrow \mathrm{H}^t(\Omega)$$

holds.

Compact embeddings are very important when approximating the solutions of PDEs by solutions of the corresponding discretized systems, see Rothe's method in Section 2.12.

2.9.10 Traces of functions on the boundary

In this subsection, the behaviour of a function at the boundary of a domain is studied. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. If a function u belongs to $\mathrm{C}(\overline{\Omega})$, then the trace of u on the boundary can be defined by the values of u on the boundary. In 1D, it holds that $\mathrm{H}^k(\Omega) \subset \mathrm{C}(\overline{\Omega})$ for all $k \in \mathbb{N}$. Therefore, for every function of $\mathrm{H}^k(\Omega)$, with $\Omega \subset \mathbb{R}$, the values of u on the boundary can be defined. However, in 2D and 3D, $\mathrm{H}^k(\Omega) \subset \mathrm{C}(\overline{\Omega})$ only holds for all $k \in \mathbb{N}$ with $k \geqslant 2$. To define the trace of any function $u \in \mathrm{H}^k(\Omega)$ with $\Omega \subset \mathbb{R}^d$ and $k \in \mathbb{N}$, a new function space $\mathrm{L}^p(\Gamma)$ is defined.

Definition 2.9.10. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. The Banach spaces $\mathrm{L}^p(\Gamma)$ on the boundary are defined analogously as the spaces $\mathrm{L}^p(\Omega)$ for $p \geqslant 1$. The space $\mathrm{L}^2(\Gamma)$ is a Hilbert space with inner product*

$$(u, v)_\Gamma = \int_\Gamma u(s)v(s)\ \mathrm{d}s.$$

Denote by ϕ the function of which the graph describes the boundary Γ (ϕ exists because the domain Ω is a Lipschitz domain). Then the integral

$$\int_\Gamma f(s)\ \mathrm{d}s := \int_{\mathbb{R}^{d-1}} f(\mathbf{x}, \phi(\mathbf{x}))\sqrt{1 + |\nabla\phi(\mathbf{x})|_{\mathrm{e}}^2}\ \mathrm{d}\mathbf{x},$$

is well-defined.

First, the trace theorem for functions in the space $\mathrm{H}^1(\Omega) = \{u \in \mathrm{L}^2(\Omega) : \nabla u \in (\mathrm{L}^2(\Omega))^d\}$ is given. It is not possible to define the trace of a function that belongs to $\mathrm{L}^2(\Omega)$, because a function of $\mathrm{L}^2(\Omega)$ does not change as an element of $\mathrm{L}^2(\Omega)$ if it is changed in points of a set with measure zero on the boundary.

Theorem 2.9.24. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. There exists a unique linear map, defined by*

$$\tilde{\gamma} : \mathrm{H}^1(\Omega) \to \mathrm{L}^2(\Gamma),$$

that has the following properties:
(i) $\tilde{\gamma}(u) = u|_\Gamma, \quad \forall u \in \mathrm{H}^1(\Omega) \cap \mathrm{C}(\overline{\Omega})$*;*
(ii) there exists a $C(\Omega) > 0$ such that

$$\|\tilde{\gamma}(u)\|_{\mathrm{L}^2(\Gamma)} \leqslant C\,\|u\|_{\mathrm{H}^1(\Omega)}, \quad \forall u \in \mathrm{H}^1(\Omega).$$

The above theorem assures that with each $u \in \mathrm{H}^1(\Omega)$ a function $\tilde{\gamma}(u) \in \mathrm{L}^2(\Gamma)$ corresponds. This function is called the trace of u and is also denoted by $u(S)$ with $S \in \Gamma$. The inequality in the theorem expresses that the map $\tilde{\gamma}$ is a continuous or bounded operator, i.e.

$$u_k \to u \text{ in } \mathrm{H}^1(\Omega) \Rightarrow \tilde{\gamma}(u_k) \to \tilde{\gamma}(u) \text{ in } \mathrm{L}^2(\Gamma).$$

Insight into the previous theorem can be gained by the fact that $\overline{\mathrm{C}^\infty(\overline{\Omega})} = \mathrm{H}^1(\Omega)$. This implies that $u \in \mathrm{H}^1(\Omega)$ is the limit in the $\mathrm{H}^1(\Omega)$-norm of a sequence $\{u_k\} \subset \mathrm{C}^\infty(\overline{\Omega})$. The trace theorem assures that the sequence of traces $\{u_k(S)\} \subset C^\infty(\Gamma)$ converges in $\mathrm{L}^2(\Gamma)$ to a function $u(S) \in \mathrm{L}^2(\Gamma)$ regardless of the chosen sequence $\{u_k\}$. The map $\tilde{\gamma}$ is not surjective as the following example shows [24, Example 6.6.2].

Example 2.9.25. *Let Ω be the unit ball in $\mathbb{R}^2$. Define*

$$\begin{aligned} u(x,y) &= \sum_{n=1}^{\infty} 2^{-n} \rho^{2^{2n}} \cos\left(2^{2n}\theta\right) \quad \textit{on } B(0,1) \setminus \{(0,0)\}, \\ u(0,0) &= 0, \end{aligned}$$

where

$$x = \rho\cos\theta, \quad y = \rho\sin\theta \quad \textit{for } \rho \in (0,1], \theta \in [0, 2\pi).$$

The series converges uniformly on $\overline{B(0,1)}$. Thus it defines a continuous function on this set, i.e. $u \in \mathrm{L}^q(\Gamma)$ for every $q \geqslant 1$. On the other hand, $u \notin \mathrm{H}^1(\Omega)$. Another example can be found in Example 2.9.7.

Because

$$\mathrm{H}^1(\Gamma) \subset \tilde{\gamma}\left(\mathrm{H}^1(\Omega)\right) \subsetneqq \mathrm{L}^2(\Gamma),$$

the space $\mathrm{H}^{\frac{1}{2}}(\Gamma) := \tilde{\gamma}\left(\mathrm{H}^1(\Omega)\right)$ is well-defined. This is the Hilbert space spanned by all traces of functions from $\mathrm{H}^1(\Omega)$. Therefore, the map

$$\gamma : \mathrm{H}^1(\Omega) \to \mathrm{H}^{\frac{1}{2}}(\Gamma)$$

is surjective. A norm in $\mathrm{H}^{\frac{1}{2}}(\Gamma)$ is defined as

$$\|u\|_{\mathrm{H}^{\frac{1}{2}}(\Gamma)} := \inf_{\varphi \in \mathrm{H}^1(\Omega), \varphi|_\Gamma = u} \|\varphi\|_{\mathrm{H}^1(\Omega)}$$

It can been shown that

$$\|u\|_{\mathrm{H}^{\frac{1}{2}}(\Gamma)} = \|w\|_{\mathrm{H}^1(\Omega)}, \tag{2.10}$$

where w is the unique solution of the Dirichlet problem

$$\begin{aligned} w - \Delta w &= 0 \quad \text{in } \Omega, \\ w &= u \quad \text{on } \Gamma. \end{aligned}$$

The norm on the dual space $\mathrm{H}^{-\frac{1}{2}}(\Gamma)$ can be written as [40, p. 98]

$$\|f\|_{\mathrm{H}^{-\frac{1}{2}}(\Gamma)} = \sup_{g \in \mathrm{H}^{\frac{1}{2}}(\Gamma)} \frac{|\langle f, g\rangle_\Gamma|}{\|g\|_{\mathrm{H}^{\frac{1}{2}}(\Gamma)}} = \sup_{g \in \mathrm{H}^{\frac{1}{2}}(\Gamma)} \frac{|(f, g)_\Gamma|}{\|g\|_{\mathrm{H}^{\frac{1}{2}}(\Gamma)}}.$$

From (2.10) and the continuity of $\tilde{\gamma}$, it follows that the following inclusion is valid:

$$\mathrm{H}^{\frac{1}{2}}(\Gamma) \hookrightarrow \mathrm{L}^2(\Gamma) \cong \left(\mathrm{L}^2(\Gamma)\right)^* \hookrightarrow \mathrm{H}^{-\frac{1}{2}}(\Gamma).$$

Moreover, it holds by the Rellich-Kondrachov Compactness Theorem 2.9.22 that

$$\mathrm{H}^{\frac{1}{2}}(\Gamma) \hookrightarrow \mathrm{H}^1(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Omega) \cong \left(\mathrm{L}^2(\Omega)\right)^* \hookrightarrow \mathrm{H}^{-\frac{1}{2}}(\Gamma),$$

therefore

$$\mathrm{H}^{\frac{1}{2}}(\Gamma) \hookrightarrow\hookrightarrow \mathrm{H}^{-\frac{1}{2}}(\Gamma).$$

Also traces of other functions, for instance, of functions in $\mathrm{H}^2(\Omega)$ can be defined.

Now, the definition of the following function spaces should be more familiar [29, Theorem 4.12].

Definition 2.9.11. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. Then*

$$\mathrm{H}_0^1(\Omega) := \{u \in \mathrm{H}^1(\Omega) : \gamma(u) = 0\} = \{u \in \mathrm{H}^1(\Omega) : u = 0 \textit{ on } \Gamma\}$$

and

$$\mathrm{H}_0^2(\Omega) := \{u \in \mathrm{H}^2(\Omega) : u = 0 \textit{ on } \Gamma \textit{ and } \nabla u \cdot \boldsymbol{\nu} = \frac{\partial u}{\partial \boldsymbol{\nu}} = 0 \textit{ on } \Gamma\}.$$

The following two theorems are 'trace theorems' for general Sobolev spaces [24, Theorem 6.4.2 and 6.4.3].

Theorem 2.9.26. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. Take $p \geqslant 1, k \in \mathbb{N}$. Let at least one of the following assumptions be fulfilled*

(i) $kp < d$ and $q = \frac{dp-p}{d-kp}$,

(ii) $kp \geqslant d$ and $q \geqslant 1$.

Then there exists a unique continuous mapping $\mathcal{T} : \mathrm{W}^{k,p}(\Omega) \to \mathrm{L}^q(\Gamma)$ such that $\mathcal{T}u = u|_\Gamma$ for all $u \in \mathrm{C}^\infty(\overline{\Omega})$.

Theorem 2.9.27 (Inverse trace theorem)**.** *Take $p > 1$. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. Then there exists a continuous linear mapping*

$$T : \mathrm{W}^{1-\frac{1}{p},p}(\Gamma) \to \mathrm{W}^{1,p}(\Omega)$$

such that

$$(v = Tu) \Longrightarrow (v = u \textit{ on the boundary } \Gamma).$$

2.9.11 Poincarè or Friedrichs inequality, Nečas inequality

The following theorem is a generalization of the famous Friedrichs theorem.

Theorem 2.9.28 (Friedrichs inequality). *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary and let $\Gamma \subset \partial\Omega$ be a part of the boundary with a positive measure $|\Gamma| > 0$. Then there exists a positive constant $C = C(\Omega, \Gamma)$ such that*

$$\|u\|^2_{\mathrm{H}^1(\Omega)} \leqslant C \left(\|\nabla u\|^2 + \int_\Gamma u^2 \,\mathrm{d}\Gamma \right)$$

is valid for all $u \in \mathrm{H}^1(\Omega)$.

Thanks to the embedding $\mathrm{H}^1(\Omega) \subset \mathrm{C}(\overline{\Omega})$ for $\Omega = (a, b)$, the Friedrichs inequality in one dimension can be written as

$$\|u\|^2_{\mathrm{H}^1(a,b)} \leqslant C \left[\int_a^b (u'(x))^2 \,\mathrm{d}x + u^2(a) + u^2(b) \right], \qquad \forall u \in \mathrm{H}^1(a, b).$$

Corollary 2.9.3 (Friedrichs, Poincarè inequality). *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary. For all $\varphi \in \mathrm{H}^1_0(\Omega)$, there exists a positive constant C such that*

$$\|\varphi\|^2 \leqslant \|\varphi\|^2_{\mathrm{H}^1_0(\Omega)} \leqslant C \|\nabla\varphi\| . \tag{2.11}$$

The following inequality was proved by Nečas [29, see proof Theorem 1.2, p. 5].

Theorem 2.9.29. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary and let $\Gamma \subset \partial\Omega$ be a part of the boundary with a positive measure $|\Gamma| > 0$. Then*

$$\|z\|^2_\Gamma \leqslant \varepsilon \|\nabla z\|^2 + C_\varepsilon \|z\|^2 , \qquad \forall z \in \mathrm{H}^1(\Omega), \quad 0 < \varepsilon < \varepsilon_0. \tag{2.12}$$

In one dimension, the Nečas inequality can be written as

$$u^2(a) + u^2(b) \leqslant \varepsilon \|z'\|^2 + C_\varepsilon \|z\|^2 , \qquad \forall u \in \mathrm{H}^1(a, b), \quad 0 < \varepsilon < \varepsilon_0.$$

It is possible to interpret the Nečas inequality as a weighted trace inequality. Using Nečas inequality, the compactness of the trace map in Theorem 2.9.24 is shown in Corollary 2.9.30. For a more general result and an alternative proof, the reader is referred to [29, Theorem 6.2].

Theorem 2.9.30. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. Then*

$$\mathrm{H}^1(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Gamma).$$

Proof. Let u_n be a bounded sequence in $\mathrm{H}^1(\Omega)$. By the reflexivity of this space and Theorem 2.4.30, there exists an element $u \in \mathrm{H}^1(\Omega)$ such that $u_n \rightharpoonup u$ in $\mathrm{H}^1(\Omega)$ for $n \to +\infty$ (up to a subsequence). Moreover, the Rellich-Kondrachov

Compactness Theorem 2.9.22 implies that $u_n \to u$ in $\mathrm{L}^2(\Omega)$. This, together with the Nečas inequality (2.12) gives that

$$\|u_n - u\|_\Gamma \leqslant \varepsilon \, \|\nabla(u_n - u)\| + C_\varepsilon \, \|u_n - u\| \, .$$

Passing to the limit $n \to \infty$ implies that

$$\lim_{n\to\infty} \|u_n - u\|_\Gamma \leqslant \varepsilon,$$

which is valid for any small $\varepsilon > 0$. Hence,

$$\lim_{n\to\infty} \|u_n - u\|_\Gamma = 0 \text{ and } u_n \to u \text{ a.e. on } \Gamma.$$

□

The following corollary follows from the embedding $\mathrm{H}^{\frac{1}{2}}(\Gamma) \hookrightarrow \mathrm{H}^1(\Omega)$ and the previous theorem.

Corollary 2.9.4. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary Γ. Then*

$$\mathrm{H}^{\frac{1}{2}}(\Gamma) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Gamma).$$

Nečas inequality follows also from the following interesting theorem [41, Theorem 7.6] when $E = \mathrm{L}^2(\Omega)$, $F = \mathrm{H}^1(\Omega)$, $G = \mathrm{L}^2(\Gamma)$ and $\phi = \tilde{\gamma}$ (trace map).

Theorem 2.9.31. *Let E, F and G be three Banach spaces. Suppose that $F \hookrightarrow E$. If $\phi \in \sigma(F, G)$, then*

$$\forall \varepsilon > 0, \quad \exists C_\varepsilon > 0, \quad \forall u \in F: \qquad \|\phi(u)\|_G \leqslant \varepsilon \, \|u\|_F + C_\varepsilon \, \|u\|_E \, .$$

2.9.12 Equivalent norms

For applications, it is convenient to find different norms that are equivalent to the standard norm in $\mathrm{H}^k(\Omega)$. By Lemma 2.3.1, the following lemma is satisfied.

Lemma 2.9.6. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. The following norms are equivalent in $\mathrm{H}^k(\Omega)$ with $k \in \mathbb{N} \cup \{0\}$:*

$$\|u\|_{\mathrm{H}^k(\Omega)} = \left(\sum_{|\boldsymbol{\alpha}|_1 \leqslant k} \left\| D^{(\boldsymbol{\alpha})}(u) \right\|^2 \right)^{\frac{1}{2}} \text{ and } \|u\|_{\mathrm{H}^k(\Omega)} = \sum_{|\boldsymbol{\alpha}|_1 \leqslant k} \left\| D^{(\boldsymbol{\alpha})}(u) \right\| .$$

Thanks to the Friedrichs inequality 2.9.28, the following lemma is valid.

Lemma 2.9.7. *In $\mathrm{H}^1_0(\Omega)$ the norms $\|u\|_{\mathrm{H}^1(\Omega)}$ and $\|\nabla u\|$ are equivalent.*

It is possible to define other equivalent norms in the space $\mathrm{H}^k(\Omega)$ thanks to [29, Theorem 1.8 and 1.10].

Lemma 2.9.8. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with boundary Γ and $k \in \mathbb{N} \cup \{0\}$. Then, it holds for $u \in \mathrm{H}^k(\Omega)$ that there exists a positive constant C such that*

$$\|u\|_{\mathrm{H}^k(\Omega)} \leqslant C \left(\|u\|^2 + \sum_{|\boldsymbol{\alpha}|_1 = k} \left\| D^{(\boldsymbol{\alpha})} u \right\|^2 \right)^{\frac{1}{2}}.$$

If $u \in \mathrm{H}^2(\Omega)$, then there exists a positive constant C such that

$$\|u\|_{\mathrm{H}^2(\Omega)} \leqslant C \left(\|u\|^2_{\mathrm{L}^2(\Gamma)} + \sum_{|\boldsymbol{\alpha}|_1 = 2} \left\| D^{(\boldsymbol{\alpha})} u \right\|^2 \right)^{\frac{1}{2}}.$$

Corollary 2.9.5. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. The following norms are equivalent in $\mathrm{H}^2(\Omega)$:*

$$\|u\|_{\mathrm{H}^2(\Omega)} \text{ and } \|u\|_{\mathrm{H}^2(\Omega)} = \left(\|u\|^2 + \sum_{|\boldsymbol{\alpha}|_1 = 2} \left\| D^{(\boldsymbol{\alpha})} u \right\|^2 \right)^{\frac{1}{2}}.$$

In the subspace $\mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, an equivalent norm is given by

$$\|u\|_{\mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)} = \left(\sum_{|\boldsymbol{\alpha}|_1 = 2} \left\| D^{(\boldsymbol{\alpha})} u \right\|^2 \right)^{\frac{1}{2}}.$$

For the space $\mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, a stronger result can be obtained [42, Theorem 1].

Theorem 2.9.32. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. In the space $\mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, the norms $\|\Delta u\|$ and $\|u\|_{\mathrm{H}^2(\Omega)}$ are equivalent.*

In the case of homogeneous Neumann boundary conditions, the following theorem is valid [43, Theorem 2.50].

Theorem 2.9.33. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. The norms $\|u\| + \|\Delta u\|$ and $\|u\|_{\mathrm{H}^2(\Omega)}$ are equivalent for $u \in \mathrm{H}^2(\Omega)$ satisfying $\nabla u \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$.*

2.9.13 Sobolev spaces for vector fields

Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with Lipschitz continuous boundary Γ. The inner product in $\mathrm{L}^2(\Omega)$ can be extended trivially to vector functions. Suppose that $\mathbf{u} = (u_1, u_2, u_3) \in \mathbf{L}^2(\Omega) := (\mathrm{L}^2(\Omega))^3$ and $\mathbf{v} = (v_1, v_2, v_3) \in \mathbf{L}^2(\Omega)$. Then the inner product in $\mathbf{L}^2(\Omega)$ is defined as

$$(\mathbf{u}, \mathbf{v}) = \int_\Omega \sum_{j=1}^3 u_j \overline{v}_j = \sum_{j=1}^3 (u_j, v_j)_{\mathrm{L}^2(\Omega)}.$$

From the definition of weak derivative, it is easy to derive the following expressions:

- $\mathbf{v} \in \mathbf{L}^p(\Omega)$ is the (weak) rotor of $\mathbf{u} \in \mathbf{L}^p(\Omega)$ (notation $\mathbf{v} = \nabla \times \mathbf{u}$) if

$$\int_\Omega \mathbf{v} \cdot \boldsymbol{\varphi} = \int_\Omega \mathbf{u} \cdot \nabla \times \boldsymbol{\varphi}, \qquad \forall \boldsymbol{\varphi} \in (\mathrm{C}_0^\infty(\Omega))^3.$$

- $\mathbf{v} \in \mathbf{L}^p(\Omega)$ is the (weak) gradient of $u \in \mathrm{L}^p(\Omega)$ (notation $\mathbf{v} = \nabla u$) if

$$\int_\Omega \mathbf{v} \cdot \boldsymbol{\varphi} = -\int_\Omega u \nabla \cdot \boldsymbol{\varphi}, \qquad \forall \boldsymbol{\varphi} \in (\mathrm{C}_0^\infty(\Omega))^3.$$

- $v \in \mathrm{L}^p(\Omega)$ is the (weak) divergence of $\mathbf{u} \in \mathbf{L}^p(\Omega)$ (notation $v = \nabla \cdot \mathbf{u}$) if

$$\int_\Omega v\phi = -\int_\Omega \mathbf{u} \cdot \nabla \phi, \qquad \forall \phi \in \mathrm{C}_0^\infty(\Omega).$$

Using Theorem 2.8.1 and the previous definitions, the following standard identities stay also valid for weak operators:

$$\begin{aligned} \nabla \times \nabla u = \mathbf{0} \quad & \forall u \in \mathrm{C}_0^\infty(\Omega)^*, \\ \nabla \cdot (\nabla \times \mathbf{u}) = 0 \quad & \forall \mathbf{u} \in (\mathrm{C}_0^\infty(\Omega)^*)^3. \end{aligned}$$

Definition 2.9.12. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^3$. The standard Sobolev spaces for vector fields $\mathbf{H}^1(\Omega), \mathbf{H}(\mathbf{curl};\Omega)$ and $\mathbf{H}(\mathbf{div};\Omega)$ are defined as*

$$\begin{aligned} \mathbf{H}^1(\Omega) &:= \left\{ \mathbf{u} \in \mathbf{L}^2(\Omega) : \nabla \mathbf{u} \in \left(\mathrm{L}^2(\Omega)\right)^{3\times 3} \right\}, \\ \mathbf{H}(\mathbf{curl};\Omega) &:= \left\{ \mathbf{u} \in \mathbf{L}^2(\Omega) : \nabla \times \mathbf{u} \in \mathbf{L}^2(\Omega) \right\}, \\ \mathbf{H}(\mathbf{div};\Omega) &:= \left\{ \mathbf{u} \in \mathbf{L}^2(\Omega) : \nabla \cdot \mathbf{u} \in \mathrm{L}^2(\Omega) \right\}, \end{aligned}$$

and are respectively equipped with the norms

$$\begin{aligned} \|\mathbf{u}\|_{\mathbf{H}^1(\Omega)} &= \left(\|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)} + \|\nabla \mathbf{u}\|^2_{(\mathrm{L}^2(\Omega))^{3\times 3}} \right)^{\frac{1}{2}}, \\ \|\mathbf{u}\|_{\mathbf{H}(\mathbf{curl};\Omega)} &= \left(\|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)} + \|\nabla \times \mathbf{u}\|^2_{\mathbf{L}^2(\Omega)} \right)^{\frac{1}{2}}, \\ \|\mathbf{u}\|_{\mathbf{H}(\mathbf{div};\Omega)} &= \left(\|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)} + \|\nabla \cdot \mathbf{u}\|^2_{\mathrm{L}^2(\Omega)} \right)^{\frac{1}{2}}. \end{aligned}$$

Remark 2.9.4. *Remark that*

$$\begin{aligned} \|\mathbf{u}\|^2_{\mathbf{L}^2(\Omega)} &= \sum_{j=1}^3 \|u_j\|^2_{\mathrm{L}^2(\Omega)}; \\ \|\nabla \mathbf{u}\|^2_{(\mathrm{L}^2(\Omega))^{3\times 3}} &= \sum_{j=1}^3 \|\nabla u_j\|^2_{\mathbf{L}^2(\Omega)} = \sum_{j=1}^3 \sum_{i=1}^3 \|\partial_{x_i} u_j\|^2_{\mathrm{L}^2(\Omega)}. \end{aligned}$$

The following lemma can be found in [39, Theorem 3.22 and Theorem 3.26].

Lemma 2.9.9. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^3$. The space $\left(\mathrm{C}^\infty(\overline{\Omega})\right)^3$ is dense in $\mathbf{H}(\mathbf{div};\Omega)$ and $\mathbf{H}(\mathbf{curl};\Omega)$. The spaces $\mathbf{H}(\mathbf{div};\Omega)$ and $\mathbf{H}(\mathbf{curl};\Omega)$ are Hilbert spaces with associated inner products*

$$\begin{aligned}(\mathbf{u},\mathbf{v})_{\mathbf{H}(\mathbf{div};\Omega)} &= (\mathbf{u},\mathbf{v})_{\mathbf{L}^2(\Omega)} + (\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})_{\mathbf{L}^2(\Omega)},\\ (\mathbf{u},\mathbf{v})_{\mathbf{H}(\mathbf{curl};\Omega)} &= (\mathbf{u},\mathbf{v})_{\mathbf{L}^2(\Omega)} + (\nabla\times\mathbf{u},\nabla\times\mathbf{v})_{\mathbf{L}^2(\Omega)}.\end{aligned}$$

Therefore, both spaces are reflexive.

Also the following Sobolev spaces play an important role.

Definition 2.9.13. *The space $\mathbf{H}_0(\mathbf{div};\Omega)$ is defined as the closure of $(\mathrm{C}_0^\infty(\Omega))^3$ in the $\mathbf{H}(\mathbf{div};\Omega)$ norm. The space $\mathbf{H}_0(\mathbf{curl};\Omega)$ is defined as the closure of $(\mathrm{C}_0^\infty(\Omega))^3$ in the $\mathbf{H}(\mathbf{curl};\Omega)$ norm.*

Remark 2.9.5. *Let $k\in\mathbb{N}$. The Hilbert spaces $\mathbf{H}_0^k(\Omega)$, $\mathbf{H}^k(\Omega)$, $\mathbf{H}_0(\mathbf{div};\Omega)$, $\mathbf{H}_0(\mathbf{curl};\Omega)$, $\mathbf{H}(\mathbf{div};\Omega)$ and $\mathbf{H}(\mathbf{curl};\Omega)$ are dense in $\mathbf{L}^2(\Omega)$ by the density of $(\mathrm{C}_0^\infty(\Omega))^3$ in $\mathbf{L}^2(\Omega)$.*

Remark 2.9.6. *As a consequence of Lemma 2.4.2, it is clear that the spaces $\mathbf{H}_0(\mathbf{div};\Omega)$, $\mathbf{H}_0(\mathbf{curl};\Omega)$, $\mathbf{H}(\mathbf{div};\Omega)$ and $\mathbf{H}(\mathbf{curl};\Omega)$ are separable.*

The following two theorems give alternative characterizations of $\mathbf{H}_0(\mathbf{div};\Omega)$ and $\mathbf{H}_0(\mathbf{curl};\Omega)$ [39, Theorem 3.25, Lemma 3.27 and Theorem 3.33].

Theorem 2.9.34. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^3$ with boundary Γ and unit outward normal vector $\boldsymbol{\nu}$. Then*

$$\begin{aligned}\mathbf{H}_0(\mathbf{curl};\Omega) &= \{\mathbf{u}\in\mathbf{H}(\mathbf{curl};\Omega) : \mathbf{u}\times\boldsymbol{\nu} = \mathbf{0} \textit{ on } \Gamma\}\\ &= \{\mathbf{u}\in\mathbf{H}(\mathbf{curl};\Omega) :\\ &\qquad (\mathbf{u},\nabla\times\boldsymbol{\varphi}) = (\nabla\times\mathbf{u},\boldsymbol{\varphi}) \textit{ for all } \boldsymbol{\varphi}\in\left(\mathrm{C}^\infty(\overline{\Omega})\right)^3\},\\ \mathbf{H}_0(\mathbf{div};\Omega) &= \{\mathbf{u}\in\mathbf{H}(\mathbf{div};\Omega) : \mathbf{u}\cdot\boldsymbol{\nu} = 0 \textit{ on } \Gamma\}\,.\end{aligned}$$

The spaces $\mathbf{H}(\mathbf{curl};\Omega)$ and $\mathbf{H}(\mathbf{div};\Omega)$ are not compactly embedded in $\mathbf{L}^2(\Omega)$ [44]. The space $\mathbf{H}(\mathbf{curl};\Omega)\cap\mathbf{H}(\mathbf{div};\Omega)$ is a Banach space with the graph norm

$$\|\mathbf{u}\|_{\mathbf{H}(\mathbf{curl};\Omega)\cap\mathbf{H}(\mathbf{div};\Omega)} := (\|\mathbf{u}\| + \|\nabla\times\mathbf{u}\| + \|\nabla\cdot\mathbf{u}\|)^{\frac{1}{2}}\,.$$

The space $\left(\mathrm{C}^\infty(\overline{\Omega})\right)^3$ is dense in $\mathbf{H}(\mathbf{curl};\Omega)\cap\mathbf{H}(\mathbf{div};\Omega)$. The embedding of $\mathbf{H}(\mathbf{curl};\Omega)\cap\mathbf{H}(\mathbf{div};\Omega)$ in $\mathbf{L}^2(\Omega)$ is not compact [44, Proposition 2.7]. The subspace $\mathbf{H}_0(\mathbf{div};\Omega)\cap\mathbf{H}_0(\mathbf{curl};\Omega)$ inherits the norm of $\mathbf{H}(\mathbf{curl};\Omega)\cap\mathbf{H}(\mathbf{div};\Omega)$. By the identity

$$\mathbf{v} = \boldsymbol{\nu}\times(\mathbf{v}\times\boldsymbol{\nu}) + (\mathbf{v}\cdot\boldsymbol{\nu})\boldsymbol{\nu},$$

where $\boldsymbol{\nu}$ is a unit normal vector, the following equality holds [44, Theorem 2.5]:

$$\mathbf{H}_0^1(\Omega) = \mathbf{H}_0(\mathbf{div};\Omega)\cap\mathbf{H}_0(\mathbf{curl};\Omega),$$

However, in general, it is only possible to confirm that

$$\mathbf{H}^1(\Omega) \subset \mathbf{H}(\mathbf{div};\Omega) \cap \mathbf{H}(\mathbf{curl};\Omega). \tag{2.13}$$

The reverse inclusion is only valid for convex or smooth bounded domains if $\mathbf{u} \times \boldsymbol{\nu} = \mathbf{0}$ on the boundary. The most important results are summarized in the following theorem [44, Theorem 2.9, 2.12 and 2.17], [45], [46, Theorem 6.1], [47, Lemma 3.3 and Theorem 3.7].

Theorem 2.9.35. *Assume that the bounded domain $\Omega \subset \mathbb{R}^3$ is of class $\mathrm{C}^{1,1}$ or is convex. Then, it holds that*

$$\mathbf{H}(\mathbf{curl};\Omega) \cap \mathbf{H}_0(\mathbf{div};\Omega) \hookrightarrow \mathbf{H}^1(\Omega)$$

and

$$\mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}(\mathbf{div};\Omega) \hookrightarrow \mathbf{H}^1(\Omega).$$

For a Lipschitz domain the reverse inclusion of (2.13) does not hold and the compactness of $\mathbf{H}^1(\Omega)$ in $\mathbf{L}^2(\Omega)$ cannot be used. A more useful result is the following regularity estimate due to Costabel [48, Theorem 2], which is valid for a general Lipschitz domain.

Theorem 2.9.36. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^3$ with boundary Γ. Suppose that $\mathbf{u} \in \mathbf{H}(\mathbf{div};\Omega) \cap \mathbf{H}(\mathbf{curl};\Omega)$ and $\mathbf{u} \times \boldsymbol{\nu} \in \mathbf{L}^2(\Gamma)$. Then $\mathbf{u} \in \mathbf{H}^{\frac{1}{2}}(\Omega) := \left(\mathrm{H}^{\frac{1}{2}}(\Omega)\right)^3$ and the following norm estimate holds:*

$$\|\mathbf{u}\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} \lesssim \|\mathbf{u}\|_{\mathbf{L}^2(\Omega)} + \|\nabla \times \mathbf{u}\|_{\mathbf{L}^2(\Omega)} + \|\nabla \cdot \mathbf{u}\|_{\mathbf{L}^2(\Omega)} + \|\mathbf{u} \times \boldsymbol{\nu}\|_{\mathbf{L}^2(\Gamma)}.$$

Similarly, suppose that $\mathbf{u} \in \mathbf{H}(\mathbf{div};\Omega) \cap \mathbf{H}(\mathbf{curl};\Omega)$ and $\mathbf{u} \cdot \boldsymbol{\nu} \in \mathbf{L}^2(\Gamma)$. Then $\mathbf{u} \in \mathbf{H}^{\frac{1}{2}}(\Omega)$ and the following norm estimate holds:

$$\|\mathbf{u}\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} \lesssim \|\mathbf{u}\|_{\mathbf{L}^2(\Omega)} + \|\nabla \times \mathbf{u}\|_{\mathbf{L}^2(\Omega)} + \|\nabla \cdot \mathbf{u}\|_{\mathbf{L}^2(\Omega)} + \|\mathbf{u} \cdot \boldsymbol{\nu}\|_{\mathbf{L}^2(\Gamma)}.$$

The space $\mathbf{H}^{\frac{1}{2}}(\Omega)$ is separable and reflexive by Theorem 2.9.8. For this fractional Sobolev space $\mathbf{H}^{\frac{1}{2}}(\Omega)$, the following compactness argument is available, see [49, Lemma 10] or [39, Theorem 3.7].

Theorem 2.9.37. *Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain. Then*

$$\mathbf{H}^{\frac{1}{2}}(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega).$$

Corollary 2.9.6. *Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain. Then, it holds that*

$$\mathbf{H}(\mathbf{curl};\Omega) \cap \mathbf{H}_0(\mathbf{div};\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega)$$

and

$$\mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}(\mathbf{div};\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega).$$

2.9.14 Integral identities

In this subsection, basic integral identities for the spaces under consideration are listed. First, the fundamental Green's formula and one of its consequences, the divergence theorem of Gauss, are given. The fundamental Green's formula is the multidimensional extension of the well-known integration by parts formula in 1D:

$$\int_a^b f'(t)g(t)\,\mathrm{d}t = -\int_a^b f(t)g'(t)\,\mathrm{d}t + f(b)g(b) - f(a)g(a),$$

with f and g two continuously differentiable functions.

Theorem 2.9.38 (Fundamental Green's theorem)**.** *Let Ω be a domain in $\mathbb{R}^d$ and let $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_d)$ denote the unit outward normal vector field along the Lipschitz-continuous boundary Γ of Ω. Then, given any functions $u, v \in \mathrm{C}^1(\Omega) \cap \mathrm{C}(\overline{\Omega})$, it holds that*

$$\int_\Omega u\partial_{x_i} v \;\mathrm{d}\mathbf{x} = -\int_\Omega v\partial_{x_i} u \;\mathrm{d}\mathbf{x} + \int_\Gamma uv\nu_i \,\mathrm{d}\Gamma,$$

for all $1 \leqslant i \leqslant d$.

Theorem 2.9.39 (Green's formula)**.** *Let Ω be a domain in $\mathbb{R}^d$ and let $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_d)$ denote the unit outward normal vector field along the Lipschitz continuous boundary Γ of Ω. Then, for all $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$ and $v \in \mathrm{C}^1(\Omega) \cap \mathrm{C}(\overline{\Omega})$, it holds that*

$$\int_\Omega v\Delta u \;\mathrm{d}\mathbf{x} = -\int_\Omega \nabla u \cdot \nabla v \;\mathrm{d}\mathbf{x} + \int_\Gamma v\nabla u \cdot \boldsymbol{\nu} \,\mathrm{d}\Gamma.$$

Theorem 2.9.40 (Divergence theorem of Gauss)**.** *Let $\Omega \subset \mathbb{R}^d$, with boundary Γ and unit outward normal vector $\boldsymbol{\nu}$, be a bounded Lipschitz domain. Let $\boldsymbol{\varphi} \in \left(\mathrm{C}^1(\Omega)\right)^d \cap \left(\mathrm{C}(\overline{\Omega})\right)^d$. Then*

$$\int_\Omega \nabla \cdot \boldsymbol{\varphi} \;\mathrm{d}\mathbf{x} = \int_\Gamma \boldsymbol{\varphi} \cdot \boldsymbol{\nu} \,\mathrm{d}\Gamma.$$

The following theorem contains an extension of the integration by parts formula to suitable Sobolev spaces. For a more general version, the reader is referred to [29, Theorem 1.1].

Theorem 2.9.41 (Green's theorem)**.** *For $u, v \in \mathrm{H}^1(\Omega)$, the following Green's theorem is valid:*

$$\int_\Omega u\partial_{x_i} v \;\mathrm{d}\mathbf{x} = -\int_\Omega v\partial_{x_i} u \;\mathrm{d}\mathbf{x} + \int_\Gamma uv\nu_i \,\mathrm{d}\Gamma,$$

where Γ is the Lipschitz continuous boundary of the bounded domain $\Omega \subset \mathbb{R}^d$ and $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_d)$ is the unit outward normal vector.

With this identity, the following Green's theorems are valid.

Theorem 2.9.42. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz continuous boundary Γ and unit outward normal vector $\boldsymbol{\nu}$.*

(i) For all $u \in \mathrm{H}^2(\Omega)$ and $v \in \mathrm{H}^1(\Omega)$, it holds that

$$\int_\Omega v\Delta u \, \mathrm{d}\mathbf{x} = -\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}\mathbf{x} + \int_\Gamma v \nabla u \cdot \boldsymbol{\nu} \, \mathrm{d}\Gamma.$$

(ii) For all $u, v \in \mathrm{H}^2(\Omega)$, it holds that

$$\int_\Omega (u\Delta v - v\Delta u) \, \mathrm{d}\mathbf{x} = \int_\Gamma (u\nabla v \cdot \boldsymbol{\nu} - v\nabla u \cdot \boldsymbol{\nu}) \, \mathrm{d}\Gamma.$$

(iii) For all $\mathbf{u} \in \mathbf{H}(\mathbf{div}; \Omega)$ and $\phi \in \mathrm{H}^1(\Omega)$, it holds that

$$\int_\Omega \phi \nabla \cdot \mathbf{u} \, \mathrm{d}\mathbf{x} = -\int_\Omega \mathbf{u} \cdot \nabla \phi \, \mathrm{d}\mathbf{x} + \int_\Gamma \phi \mathbf{u} \cdot \boldsymbol{\nu} \, \mathrm{d}\Gamma.$$

(iv) For all $\mathbf{u} \in \mathbf{H}^2(\Omega)$ and $\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega)$, the following Green's theorem is valid:

$$\int_\Omega \Delta \mathbf{u} \cdot \boldsymbol{\varphi} \, \mathrm{d}\mathbf{x} = -\int_\Omega \nabla \mathbf{u} : \nabla \boldsymbol{\varphi} \, \mathrm{d}\mathbf{x} + \int_\Gamma \boldsymbol{\varphi} \cdot [(\nabla \mathbf{u})\, \boldsymbol{\nu}] \, \mathrm{d}\Gamma.$$

Also the following integral equalities are frequently used.

Theorem 2.9.43. *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with Lipschitz continuous boundary Γ and unit outward normal $\boldsymbol{\nu}$.*

(i) For all $\mathbf{u} \in \mathbf{H}(\mathbf{curl}; \Omega)$ and $\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega)$, the following Green's theorem is valid:

$$\int_\Omega \nabla \times \mathbf{u} \cdot \boldsymbol{\varphi} \, \mathrm{d}\mathbf{x} = \int_\Omega \mathbf{u} \cdot \nabla \times \boldsymbol{\varphi} \, \mathrm{d}\mathbf{x} + \int_\Gamma \boldsymbol{\nu} \times \mathbf{u} \cdot \boldsymbol{\varphi} \, \mathrm{d}\Gamma.$$

(ii) For all $\mathbf{u}, \boldsymbol{\varphi} \in \mathbf{H}(\mathbf{curl}; \Omega)$, the following Green's theorem is valid:

$$\int_\Omega \nabla \times \mathbf{u} \cdot \boldsymbol{\varphi} \, \mathrm{d}\mathbf{x} = \int_\Omega \mathbf{u} \cdot \nabla \times \boldsymbol{\varphi} \, \mathrm{d}\mathbf{x} + \int_\Gamma [\boldsymbol{\nu} \times \mathbf{u}] \cdot [(\boldsymbol{\nu} \times \boldsymbol{\varphi}) \times \boldsymbol{\nu}] \, \mathrm{d}\Gamma.$$

2.10 Partial differential equations

A partial differential equation is an equation involving an unknown function of several independent variables and its partial derivatives with respect to those variables [10, 50]. If the unknown function depends on just a single variable, then the relation is called an ordinary differential equation. In partial differential equations, two or more independent variables appear.

Definition 2.10.1. *Let Ω be a domain in $\mathbb{R}^d$. An expression of the form*

$$F\left(\mathbf{x}, D^{(k,0,\dots,0)}u, \dots, D^{\boldsymbol{\alpha}}u, \dots, D^{(0,0,\dots,0)}u\right) = 0, \qquad \mathbf{x} \in \Omega,$$

with $|\boldsymbol{\alpha}|_1 \leqslant k$ is called a PDE of the order $k \in \mathbb{N}$.

PDEs are used to formulate and solve problems that involve unknown functions of several variables, such as the propagation of sound or heat, electrostatics, electrodynamics, fluid flow, elasticity or, more generally, any process that is distributed in space or distributed in space and time. Completely different physical problems may have identical or similar mathematical formulations.

PDEs can be classified by:
- Order of the PDE;
- Linear, semilinear, quasilinear, and fully nonlinear.

The order of a PDE is the order of the highest order derivative that appears in the PDE. If the relation F from Definition 2.10.1 is linear, then the PDE is called linear. Otherwise, the PDE is called nonlinear.

Every linear PDE can be written in the form

$$L[u] = f, \tag{2.14}$$

where $u \mapsto L[u]$ is a linear mapping and f is a function of independent variables (space and time variable). Thus, in a linear PDE all the coefficients are independent of the unknown function u, but they can still be time and space-dependent. Linear PDEs can be classified into two subgroups: homogeneous and nonhomogeneous PDEs. The linear PDE (2.14) is said to be *homogeneous* if $f \equiv 0$. Otherwise, it is called a *nonhomogeneous* linear PDE.

Nonlinear PDEs can be subdivided as follows:
- *quasilinear*: a PDE of order m is called quasilinear if it is linear in the derivatives of order m, but with coefficients that depend on the independent variables and on the derivatives of the unknown function of order strictly smaller than m;
- *semilinear*: a quasilinear PDE in which the coefficients of derivatives of order m are functions of only the independent variables is called a semilinear PDE;
- *fully nonlinear*: a PDE which is not quasilinear is called a fully nonlinear PDE (for instance if the highest order derivative is nonlinear).

The above division can be expressed in the diagram below:

$$\text{linear PDE} \subsetneqq \text{semilinear PDE} \subsetneqq \text{quasilinear PDE} \subsetneqq \text{PDE.}$$

Example 2.10.1. *The second order PDEs*

$$\frac{\partial u}{\partial x}\left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 + 2xu\frac{\partial^2 u}{\partial y^2} - 3xy\frac{\partial u}{\partial y} - u = 0,$$
$$\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial x^2} - 3x^2 u\frac{\partial^2 u}{\partial x \partial y} + 2\frac{\partial u}{\partial x} - f(x,y)u = 0$$

and

$$x\frac{\partial^2 u}{\partial x^2} - x^2\left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial u}{\partial x}\frac{\partial u}{\partial y} = f$$

are fully nonlinear, quasilinear and semilinear respectively.

2.10.1 Classification of linear PDEs of second order

A general form of a linear PDE of second order in $\Omega \subset \mathbb{R}^d$ reads as

$$-\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\frac{\partial^2 u(\mathbf{x})}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(\mathbf{x})\frac{\partial u(\mathbf{x})}{\partial x_i} + c(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x}) \tag{2.15}$$

along with the given coefficient functions $a_{ij}, b_i, c, f : \Omega \to \mathbb{R}$. The main part of this equation is the part containing the highest derivatives of the unknown function u, namely

$$\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\frac{\partial^2 u(\mathbf{x})}{\partial x_i \partial x_j}.$$

In this multidimensional case the character of the PDE is determined by the eigenvalues (EVs) of the matrix $A = (a_{ij})_{i,j=1,\ldots,d}$.

Definition 2.10.2. *The PDE* (2.15) *is called*

- *parabolic if at least one eigenvalue is* 0*,*
- *elliptic if all EVs are* $\neq 0$ *and all have the same sign,*
- *hyperbolic if all EVs are* $\neq 0$ *and all but one have the same sign.*

Note that there are cases, where the classification is not straightforward, e.g., if all EVs are nonzero but several have a different sign. Now, some examples according to this classification are given.

Example 2.10.2.

- *The Poisson equation* $-\Delta u = f$ *is an elliptic PDE;*
- *The heat equation* $\partial_t u - \Delta u = f$ *is a parabolic PDE;*
- *The wave equation* $\partial_{tt} u - \Delta u = f$ *is a hyperbolic PDE.*

2.10.2 Associated conditions

PDEs have in general infinitely many solutions. In order to obtain a unique solution one must supplement the equation with additional conditions. The kind of conditions that should be added depend on the type of PDE under consideration. In general, PDEs are accompanied by initial conditions and boundary conditions.

2.10.2.1 Initial conditions

Consider the heat equation

$$\partial_t u(\mathbf{x},t) - \Delta u(\mathbf{x},t) = f(\mathbf{x},t), \qquad \mathbf{x} \in \Omega \subset \mathbb{R}^d, t \in (-\infty,\infty).$$

Often the temperature distribution at some initial time (say $t = 0$) is given, i.e.

$$u(\mathbf{x},0) = u_0(\mathbf{x}).$$

Then the temperature distribution at later times is derived. A problem together with an initial condition (IC) is called an *initial value problem*. The heat equation contains only the first derivative of the unknown function u with respect to t. Equations involving second derivatives of u with respect to t require two initial conditions.

2.10.2.2 Boundary conditions

Another type of constraints for PDEs that appear in many applications are the so-called boundary conditions (BCs). As the name indicates, these are conditions on the behaviour of the solution (or its derivative) at the boundary of the domain under consideration. For instance, consider again the heat equation but now on a bounded domain $\Omega \subset \mathbb{R}^d$ with boundary $\partial\Omega$. The most frequently used BCs are

- *Dirichlet boundary condition*: the values of the solution (temperature) on the boundary are given (e.g. trough measurements), i.e. for instance

$$u(\mathbf{x},t) = f(\mathbf{x},t), \quad \mathbf{x} \in \partial\Omega, \quad t > 0.$$

- *Neumann boundary condition*: the normal derivative of the solution (flux through the boundary) on the boundary is given, i.e. for instance

$$\frac{\partial u}{\partial \boldsymbol{\nu}}(\mathbf{x},t) = f(\mathbf{x},t), \quad \mathbf{x} \in \partial\Omega, \quad t > 0.$$

- *Condition of the third kind or Robin boundary condition*: a linear combination between the boundary values of u and its normal derivative is given, i.e. for instance

$$au(\mathbf{x},t) + b\frac{\partial u}{\partial \boldsymbol{\nu}}(\mathbf{x},t) = f(\mathbf{x},t), \quad \mathbf{x} \in \partial\Omega, \quad t > 0, \quad a,b \in \mathbb{R} \setminus \{0\}.$$

- *Mixed boundary condition*: for example, when the values of u at some parts of the boundary are given and the values of its normal derivative at the rest of the boundary is given.
- *Nonlocal boundary condition*: for instance, one can provide a boundary condition relating the solution at each point on the boundary to the integral of the solution over the whole boundary.

A partial differential equation together with boundary conditions is named a *boundary value problem* (BVP). A problem together with initial and boundary conditions is called an *initial and boundary value problem* (IBVP).

2.10.3 Well-posedness

The analysis of PDEs has many facets [50]. The classical approach that dominated the nineteenth century was to develop methods for finding explicit (classical) solutions (the method of characteristics invented by Hamilton, Fourier method,...). However, it is not always possible to find a classical solution, even for simple realistic technological applications. This shifted the focus to questions about finding a generalized solution, establishing the uniqueness of a solution and the way to find an approximate solution.

Since practical computers became available, numerical methods are introduced that allow the use of computers to solve PDEs. The technical advances are followed by theoretical progress aimed at understanding the solution's structure. The goal is to discover some of the solution's properties before actually computing it. However, it should be stressed that there exist very complex equations that cannot be solved even with the aid of supercomputers. In addition, the formulation of the equation and its associated side conditions are studied. In general, the equation originates from a model of a physical or engineering problem. It is not automatically obvious that the model is consistent in the sense that it leads to a solvable PDE. Furthermore, it is in most cases desired that the solution is unique, and that the solution is stable under small perturbations of the data. In fact, the continuous dependence of the solution on auxiliary data is frequently related to the uniqueness of the solution. In numerical models this can pose a problem because a non-continuous dependence of the solution on the data (boundary, initial or boundary conditions,...) implies that small errors in the data cause large changes in the solution. Since a numerical method is not arbitrary exact, it is not clear which of the possible approximate solutions is a good one. A theoretical understanding of the equation enables us to check whether these conditions are satisfied.

Hence, one of the fundamental theoretical questions is when the problem consisting of the equation and its associated side conditions is well-posed. The French mathematician Jacques Hadamard (1865–1963) coined the notion of well-posed problem [51].

Definition 2.10.3. *A mathematical problem is said to be well-posed in the sense of Hadamard if*

- *the solution exists,*
- *the solution is unique,*
- *the solution depends continuously on the data (boundary, conditions, coefficients, right-hand side,...): a small change in the equation or in the side conditions gives rise to a small change in the solution.*

Examples of well-posed problems include the Dirichlet problem for Laplace's equation ($\Delta u = 0$), and the heat equation with specified initial conditions. Problems that are not well-posed in the sense of Hadamard are called ill-posed. In practice, such problems are unsolvable.

Remark 2.10.1. *The notion of continuous dependence on the data is important because, in applications, the data is usually obtained through measurements and therefore, it might be noisy. Consider a problem in the general form: find a solution u such that*

$$F(u, f) = 0,$$

with data f. Denote by δf a small perturbation on the data and by δu the modification in the solution that occurred because of this perturbation such that

$$F(u + \delta u, f + \delta f) = 0.$$

The solution depends continuously on the data means that

$$\forall \eta > 0, \exists C(\eta, f) : \quad \|\delta f\| < \eta \Rightarrow \|\delta u\| \leqslant C(\eta, f) \|\delta f\|,$$

with $\|\cdot\|$ a suitable norm.

Remark 2.10.2. *In this thesis, by well-posedness of a problem, it is meant that a solution exists and is unique. The continuous dependence of the solution on the data can be studied in the same way as the uniqueness of the solution.*

Remark 2.10.3. *The existence and uniqueness of solutions of ordinary differential equations (ODEs) has a very satisfactory answer in the Picard-Lindelöf theorem, that is far from the case for PDEs.*

2.10.3.1 Inverse problems

In an inverse problem one usually has to determine an unknown coefficient (information of interest) in the PDE from additional measurements inside the observed domain or on its boundary. Inverse problems (IPs) are often ill-posed. This means that there is either no solution in a classical sense or if there is any, then it might not be unique or might not depend continuously on the data.

In literature, three main types of inverse problems are distinguished:

(a) parameter identification, where the material parameters appearing in the equation are not known and should be reconstructed, e.g., diffusion coefficients, source terms, etc.,
(b) boundary value inverse problems, where direct measurements on the boundary (or a part of it) are unfeasible and have to be determined,
(c) evolutionary inverse problems in which the initial conditions are not known and have to be reconstructed.

Inverse problems are inherently driven by applications and they arise in a vast variety of practical situations such as biomedical engineering, image processing and non-destructive material evaluation.

There are two big goals in IPs: (global or local) existence of a solution and its uniqueness. The usual methodology in IPs relies on a suitable parametrization of the problem and involving the continuous dependence of a parameterized solution on the parameter itself. In that case, a cost functional capturing the error between the parameterized and the exact solutions at a given place is constructed and minimized. The lack of convexity of this functional disturbs the uniqueness of a solution. Therefore a suitable (Tikhonov) regularization of the functional is applied to guarantee its convexity, cf. [52, 53]. The minimization of the functional is based on the theory of monotone operators and is numerically peformed by adequate approximation techniques, such as the steepest descend, Ritz or Newton or Levenberg-Marquardt method.

2.11 Elliptic problems

In this section, the goal is to investigate the existence of a (unique) solution to a PDE of the form

$$Au = f, \tag{2.16}$$

where A is a given linear or nonlinear differential operator of the second order and f is a given continuous function. To find a classical solution means to find a function u that has continuous derivatives up to the second order, i.e. $u \in \mathrm{C}^2(\Omega)$. In such case, this PDE is fulfilled pointwise in all points of the domain under consideration. This means that equation (2.16) can be seen as an equation in the set of real numbers. In many applications, the classical solution does not exist. Therefore, a solution is searched with worse properties. For this, the notion of strong and weak solution is introduced, cf. for instance [10].

2.11.1 Strong and weak solution

For instance, let us solve equation (2.16) with $A = -\Delta + I$ on a bounded domain $\Omega \subset \mathbb{R}^d$ accompanied by a homogeneous Dirichlet condition $u = 0$ on $\partial\Omega$ or a Neumann condition $\nabla u \cdot \boldsymbol{\nu} = g$ on $\partial\Omega$. Then a classical solution belongs respectively to $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$ or $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$. In order to reduce this regularity assumptions, the concepts of weak solution and weak variational formulation are introduced. The derivation of the weak formulation consists of three main steps:

- Multiply the given PDE with a test function and integrate the result over the domain,
- Apply a theorem of Green (integration by parts),
- Choose a suitable test space.

As a first example, the derivation of the weak formulation is done in the case of a homogeneous Dirichlet BC. It is assumed that $u \in \mathrm{C}_0^\infty(\Omega)$ because u is zero on the boundary. The derivation of the weak formulation starts with multiplying equation (2.16) with a test function $\varphi \in \mathrm{C}_0^\infty(\Omega)$ and integrating the result over Ω, i.e.

$$-\int_\Omega \Delta u\varphi + \int_\Omega u\varphi = \int_\Omega f\varphi, \quad \text{for all } \varphi \in \mathrm{C}_0^\infty(\Omega). \tag{2.17}$$

The strong regularity assumptions $u, \varphi \in \mathrm{C}_0^\infty(\Omega)$ can be weakened to $u \in \mathrm{H}^2(\Omega)$ and $\varphi \in \mathrm{L}^2(\Omega)$. Then, if also $f \in \mathrm{L}^2(\Omega)$, all integrals in equation (2.17) remain finite. A *strong solution* is a function $u \in \mathrm{H}^2(\Omega)$ satisfying the previous relation (2.17). Thus a strong solution of problem (2.16) can be seen as the solution of the equation when the equation is considered in a dual space as an integral identity, which has to be satisfied for all test functions $\varphi \in \mathrm{L}^2(\Omega)$. Notice that it is no longer necessary that the second derivatives of u exist pointwise. The following step is to lower the regularity of the solution by transferring 'one derivative' from the solution to the test function φ. To do this, it is common to use a certain integration by parts formula on the problem under considerations. Performing the Green formula from Theorem 2.9.39 on relation (2.17), it is clear that

$$\int_\Omega \nabla u \cdot \nabla\varphi + \int_\Omega u\varphi = \int_\Omega f\varphi, \quad \text{for all } \varphi \in \mathrm{C}_0^\infty(\Omega).$$

This equation remains valid when $u, \varphi \in \mathrm{H}_0^1(\Omega)$ and $f \in \mathrm{L}^2(\Omega)$. The weak form of the homogeneous Dirichlet problem is stated as follows: given $f \in \mathrm{L}^2(\Omega)$, find a function $u \in \mathrm{H}_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla\varphi + \int_\Omega u\varphi = \int_\Omega f\varphi, \quad \text{for all } \varphi \in \mathrm{H}_0^1(\Omega). \tag{2.18}$$

A *weak or variational solution* is a solution $u \in \mathrm{H}_0^1(\Omega)$ of the latter integral equation. The existence and uniqueness of a solution is discussed in the following subsections.

Remark 2.11.1. *The assumption $f \in \mathrm{L}^2(\Omega)$ can be weakened further to $f \in \mathrm{H}_0^1(\Omega)^*$, where the dual space $\mathrm{H}_0^1(\Omega)^*$ is a larger space than $\mathrm{L}^2(\Omega)$. Then the integral $\int_\Omega f\varphi$ in* (2.18) *is interpreted as the duality pairing $\langle f, \varphi\rangle$ between $\mathrm{H}_0^1(\Omega)^*$ and $\mathrm{H}_0^1(\Omega)$.*

Remark 2.11.2 (Equivalence of the classical and weak solutions)**.** *The classical solution (if it exists) to problem* (2.16) *solves the weak formulation* (2.18)*. Conversely, if the weak solution* u *of* (2.18) *is sufficiently regular, i.e. in the case that* $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}(\overline{\Omega})$*, it also satisfies the classical formulation* (2.16)*.*

Remark 2.11.3 (Linear forms)**.** *In the study of PDEs, it is common to rewrite the variational formulation* (2.18) *in the following form: find a function* $u \in V$ *such that*

$$a(u, v) = l(v), \qquad \forall v \in V,$$

where the form $a : V \times V \to \mathbb{R}$ *is defined by*

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v + \int_\Omega uv$$

and the linear form $l \in V^*$ *by*

$$l(v) = \int_\Omega fv.$$

As a second example, the case of a nonhomogeneous Neumann BC is considered. Assume that $u \in \mathrm{C}^\infty(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$. Equation (2.16) is multiplied with a test function $\varphi \in \mathrm{C}^\infty(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$ and the result is integrated over Ω. Afterwards, the Green formula is applied. It is clear that

$$\int_\Omega \nabla u \cdot \nabla \varphi - \int_{\partial\Omega} \varphi \nabla u \cdot \boldsymbol{\nu} + \int_\Omega u\varphi = \int_\Omega f\varphi, \quad \text{for all } \varphi \in \mathrm{C}^\infty(\Omega) \cap \mathrm{C}^1(\overline{\Omega}).$$

Substituting the BC into the integral over $\partial\Omega$ and weakening the regularity assumptions, the following weak formulation is obtained: given $f \in \mathrm{L}^2(\Omega)$ and $g \in \mathrm{L}^2(\Gamma)$, find $u \in \mathrm{H}^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla \varphi + \int_\Omega u\varphi = \int_\Omega f\varphi + \int_{\partial\Omega} \varphi g, \quad \text{for all } \varphi \in \mathrm{H}^1(\Omega).$$

Remark 2.11.4. *The Dirichlet BCs are often also called essential BCs when considering second order equations because they explicitly appear in the variational formulation, namely in the definition of the test space* V*. The Neumann BCs are called natural BCs because they are implicitly incorporated in the variational formulation. They have no influence on the choice of the test space* V*.*

2.11.2 Solving a linear elliptic equation

Definition 2.11.1. *Let* V *be a normed space. The mapping* $a : V \times V \to \mathbb{R}$ *is a bilinear form on* V*, if*

$$\begin{aligned} &a(\alpha_1 u_1 + \beta_1 v_1, \alpha_2 u_2 + \beta_2 v_2) \\ &\qquad = \alpha_1\alpha_2 a(u_1, u_2) + \beta_1\beta_2 a(v_1, v_2) + \sum_{\substack{i,j=1 \\ i \neq j}}^{2} \alpha_i \beta_j a(u_i, v_j) \end{aligned}$$

holds for all $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ *and* $u_1, u_2, v_1, v_2 \in V$. *The bilinear form* $a(\cdot,\cdot)$ *is*

- *symmetric if*
$$a(u,v) = a(v,u), \quad \forall u, v \in V,$$
- *bounded (or continuous) if there exists a constant* $C_M > 0$ *such that*
$$|a(u,v)| \leqslant C_M \|u\|_V \|v\|_V, \quad \forall u, v \in V,$$
- V*-elliptic if there exists a constant* $C_m > 0$ *such that*
$$C_m \|u\|_V^2 \leqslant a(u,u), \quad \forall u \in V.$$

The bilinear form a (see Remark 2.11.3) on the test space V corresponding with the variational formulation of Problem (2.16) (induced by the linear differential operator A) is associated (one-to-one correspondence) with a unique linear operator $A : V \to V^*$ such that

$$\langle Au, \varphi\rangle := a(u,\varphi) = \int_\Omega f\varphi =: \langle l, \varphi\rangle, \quad \forall \varphi \in V.$$

The existence of a solution to this problem is guaranteed by the following well known Lax-Milgram lemma, see for instance [54, Theorem 1.1.3] or [7, Theorem 18.E].

Theorem 2.11.1 (Lax-Milgram lemma)**.** *Let* $a(\cdot,\cdot)$ *be a* V*-elliptic and continuous bilinear form in the Hilbert space* V. *Assume that* $l \in V^*$. *Then there exists a unique weak solution to the variational problem defined by*

$$a(u,\varphi) = \langle l, \varphi\rangle, \quad \textit{for all } \varphi \in V.$$

Moreover, $\|u\|_V \leqslant \frac{1}{C_m} \|l\|_{V^*}$.

In strong form, the Lax-Milgram lemma is as follows.

Theorem 2.11.2. *Let* X *be a Hilbert space. Let the operator* $A : X \to X^*$ *be linear, bounded and* X*-elliptic, i.e.*

$$\|Av\|_{X^*} \leqslant C_1 \|v\|_X, \qquad \forall v \in X$$

and

$$\langle Av, v\rangle \geqslant C_2 \|v\|_X^2, \qquad \forall v \in X.$$

Then, for each $f \in X^*$ *there exists a unique solution* $u \in X$ *of the operator equation* $Au = f$.

Example 2.11.3. *Let* $\Omega \subset \mathbb{R}^d$ *be a bounded Lipschitz domain. Consider*

$$a(u,v) = \int_\Omega \left(\sum_{i=1}^d \frac{\partial u(\mathbf{x})}{\partial x_i} \frac{\partial v(\mathbf{x})}{\partial x_i} + u(\mathbf{x})v(\mathbf{x}) \right) \mathrm{d}\mathbf{x}, \qquad u, v \in \mathrm{H}^1(\Omega).$$

Clearly, $a(\cdot,\cdot)$ *is* $\mathrm{H}^1(\Omega)$*-elliptic. The corresponding operator is* $-\Delta + I$.

Example 2.11.4. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with boundary $\partial\Omega$. Consider*

$$a(u,v) = \int_\Omega \left(\sum_{i=1}^d \frac{\partial u(\mathbf{x})}{\partial x_i} \frac{\partial v(\mathbf{x})}{\partial x_i} \right) \, \mathrm{d}\mathbf{x}.$$

This bilinear form $a(\cdot,\cdot)$ is not $\mathrm{H}^1(\Omega)$-elliptic. Indeed, if u is a constant function, it holds that $a(u,u) = 0$. However, this form is $\left\{\varphi \in \mathrm{H}^1(\Omega) : \varphi = 0 \text{ on } \Gamma_1\right\}$-elliptic with $\Gamma_1 \subset \partial\Omega$ by the Friedrichs inequality (2.11).

Example 2.11.5. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Define the operator*

$$A = -\sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(a_{ij}(\mathbf{x}) \frac{\partial}{\partial x_j} \right) + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial}{\partial x_i} + c(\mathbf{x})$$

where $b_i \in \mathrm{C}^1(\overline{\Omega})$ and where a_{ij}, b_i and c are real functions. Assume that A is strongly elliptic, i.e.

$$\sum_{i,j=1}^d a_{ij}(\mathbf{x}) \xi_i \xi_j \geqslant \alpha \left|\boldsymbol{\xi}\right|_{\mathrm{e}}^2, \quad \alpha > 0, \quad \forall \mathbf{x} \in \Omega, \quad \boldsymbol{\xi} \in \mathbb{R}^d,$$

and assume that

$$c(\mathbf{x}) - \tfrac{1}{2} \sum_{i=1}^d \frac{\partial b_i(\mathbf{x})}{\partial x_i} \geqslant 0.$$

If $V = \mathrm{H}_0^1(\Omega)$, then the bilinear form

$$\int_\Omega \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial u(\mathbf{x})}{\partial x_i} \frac{\partial v(\mathbf{x})}{\partial x_j} \, \mathrm{d}\mathbf{x} + \int_\Omega \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial u(\mathbf{x})}{\partial x_i} v(\mathbf{x}) \, \mathrm{d}\mathbf{x} + \int_\Omega c(\mathbf{x}) u(\mathbf{x}) v(\mathbf{x}) \, \mathrm{d}\mathbf{x}$$

is V-elliptic due to

$$\int_\Omega \sum_{i=1}^d b_i \frac{\partial v(\mathbf{x})}{\partial x_i} v(\mathbf{x}) \, \mathrm{d}\mathbf{x} = -\tfrac{1}{2} \int_\Omega \sum_{i=1}^d \frac{\partial b_i(\mathbf{x})}{\partial x_i} v^2(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

Example 2.11.6. *Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain. The bilinear form*

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega \mathbf{u}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x}) \, \mathrm{d}\mathbf{x} + \int_\Omega (\nabla \times \mathbf{u})(\mathbf{x}) \cdot (\nabla \times \mathbf{v})(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

is $\mathbf{H}(\mathbf{curl};\Omega)$-elliptic.

When the differential operator A is symmetric, then the following theorem is also valid.

Theorem 2.11.7. *Let $a(\cdot,\cdot)$ be a symmetric, V-elliptic and continuous bilinear form in the Hilbert space V. Assume that $f \in V^*$. Then $u \in V$ is the unique weak solution to the variational problem*

$$a(u,\varphi) = \langle f, \varphi \rangle, \quad \text{for all } \varphi \in V,$$

iff u minimizes the energy functional

$$\inf_{v \in V} \left(\tfrac{1}{2} a(v,v) - \langle f, v \rangle \right).$$

2.11.3 Solving a nonlinear elliptic equation

The theory of monotone operators can be seen as a natural nonlinear extension of the ideas behind the (linear) Lax-Milgram theorem. A fundamental result on monotone operators was proved independently by Minty and Browder in 1963 [55, 56]. For a nice introduction on the theory of monotone operators, the reader is referred to [4].

Theorem 2.11.8 (Main theorem on monotone operators)**.** *Let X be a reflexive Banach space. Consider a nonlinear operator $A : X \to X^*$. If this operator A is*

- *monotone, i.e. $\langle Ax - Ay, x - y \rangle \geqslant 0, \qquad \forall x, y \in X$,*
- *coercive, i.e.*
$$\langle Ax, x \rangle \geqslant C \|x\|_X^2, \quad \forall x \in X$$
or equivalently
$$\lim_{\|x\|_X \to \infty} \frac{\langle Ax, x \rangle}{\|x\|_X} \to \infty,$$
- *hemicontinuous, i.e. that the real function*
$$t \mapsto \langle A(u + tv), w \rangle$$
is continuous on the interval $[0, 1]$ for all $u, v, w \in X$,

then the operator equation $Ax = b$ has a solution $x \in X$ for every $b \in X^*$. This solution is unique if the operator A is strictly monotone. i.e.

$$\langle Ax - Ay, x - y \rangle = 0 \Rightarrow x = y.$$

If an operator $A : X \to X^*$ is monotone, then there is an alternative characterization for hemicontinuity.

Definition 2.11.2. *Let X and Y be normed spaces. A mapping $A : X \to Y$ is called demicontinuous iff for any $x_n, x \in X$ satisfying $x_n \to x$, it holds that $Ax_n \rightharpoonup Ax$.*

Theorem 2.11.9. *If the operator $A : X \to X^*$ is monotone, then hemicontinuity is equivalent with demicontinuity*

Example 2.11.10. *Let $\Omega \subset \mathbb{R}^d$ a bounded Lipschitz domain. Consider the operator $A : \mathrm{H}^1(\Omega) \to \mathrm{H}^1(\Omega)^*$ defined by*

$$\langle Au, v \rangle := (\beta(u), v) + (\nabla u, \nabla v)\,,$$

with $\beta : \mathbb{R} \to \mathbb{R}$ a nonlinear, continuous and a.e. differentiable function with $\beta(0) = 0$. By the mean value theorem 2.2.5, the operator A is monotone if $\beta'(s) \geqslant 0$ a.e. in $\mathbb{R}$. The operator A is strictly monotone and coercive if $\beta'(s) \geqslant \beta_0 > 0$ a.e. in $\mathbb{R}$. If $A : \mathrm{H}^1_0(\Omega) \to \mathrm{H}^1_0(\Omega)^*$, then A is strictly monotone and coercive if $\beta'(s) \geqslant 0$ a.e. in $\mathbb{R}$ thanks to Friedrichs inequality (2.11).

Example 2.11.11. *Let $\Omega \subset \mathbb{R}^3$ a bounded Lipschitz domain. Consider the operator $A : \mathbf{H}(\mathbf{curl}; \Omega) \to \mathbf{H}(\mathbf{curl}; \Omega)^*$ defined by*

$$\langle A\mathbf{u}, \mathbf{v} \rangle := (\mathbf{J}(\mathbf{u}), \mathbf{v}) + (\nabla \times \mathbf{u}, \nabla \times \mathbf{v})\,,$$

with $\mathbf{J} : \mathbb{R}^3 \to \mathbb{R}^3$ nonlinear. Then A is strictly monotone and coercive when $\mathbf{J}$ is coercive, i.e.

$$\mathbf{J}(\mathbf{x}) \cdot \mathbf{x} \geqslant |\mathbf{x}|_{\mathrm{e}}^2\,, \qquad \forall \mathbf{x} \in \mathbb{R}^3.$$

2.11.3.1 Fréchet and Gâteaux derivatives

The Fréchet derivative generalizes the derivative of a real-valued function of a single real variable to functions on Banach spaces. The Fréchet derivative should be contrasted to the more general Gâteaux derivative, which is a generalization of the classical directional derivative. The definitions of these derivatives are used to define the derivative of a functional.

Assume that X and Y are Banach spaces. Let A be a nonlinear operator on X such that $A : \mathcal{D}(A) = X \to Y$ (or densely defined). Take any $x \in X$. If

$$A(x + h) - A(x) = dA(x, h) + \omega(x, h),$$

where $dA(x, h)$ is a linear operator in $h \in X$ and

$$\lim_{h \to 0} \frac{\omega(x, h)}{\|h\|_X} = 0,$$

then $dA(x, h)$ is called the Fréchet differential of A at x and $\omega(x, h)$ is the so-called remainder of the differential. It is assumed that the operator $dA(x, h)$ is bounded in h, i.e.

$$dA(x, h) = A'(x)h.$$

The operator $A'(x) \in \mathcal{L}(X, Y)$ is called the Fréchet derivative.

A directional derivative of a scalar function or a vector field can be generalized to nonlinear operators. The Gâteaux differential (or weak variation) $VA(x, h)$ of A in the point $x \in X$ and direction $h \in X$ is defined as

$$VA(x, h) = \frac{\mathrm{d}}{\mathrm{d}t} A(x + th)\bigg|_{t=0} = \lim_{t \to 0} \frac{A(x + th) - A(x)}{t}.$$

The Gâteaux differential is homogeneous ($VA(x, \alpha h) = \alpha VA(x, h)$ with $\alpha \in \mathbb{R}$) but not always additive ($VA(x, h_1 + h_2) = VA(x, h_1) + VA(x, h_2)$). If the weak variation is linear in h then it can be written as

$$VA(x, h) = A'(x)h,$$

where $A'(x)$ is the so-called Gâteaux derivative of A at x.

Remark 2.11.5. *There is not a single Gâteaux differential at each point. Rather, at each point x there is a Gâteaux differential for each direction h.*

The following generalized mean value theorem holds true.

Theorem 2.11.12 (Generalized mean value theorem)**.** *Let $A : X \to Y$ be a nonlinear operator between two Banach spaces X and Y having a linear Gâteaux differential. Then the following generalized Lagrange formula takes place*

$$\langle A(x + h) - A(x), e\rangle = \langle VA(x + \theta h, h), e\rangle, \qquad 0 < \theta = \theta(e) < 1,$$

where $e \in Y^*$.

If A is Fréchet differentiable at x, it is also Gâteaux differentiable at x. The converse generally is not true.

Lemma 2.11.1. *Let X and Y be Banach spaces and assume that the operator $F : X \to Y$ has a linear Gâteaux differential. Then, a continuous Gâteaux derivative of F is a Fréchet derivative of F.*

The definition of Fréchet and Gâteaux derivative can be repeated for a functional. Also the gradient of a functional can be defined. Let f be a nonlinear functional on a Banach space X such that $f : \mathcal{D}(f) = X \to \mathbb{R}$. The Gâteaux differential $Vf(x, h)$ of f in the point $x \in X$ and direction $h \in X$ is defined as

$$Vf(x, h) = \lim_{t \to 0} \frac{f(x + th) - f(x)}{t}.$$

At each point $x \in X$, the Gâteaux differential defines a function $Vf(x, \cdot) : X \to \mathbb{R}$. If the Gâteaux differential $Vf(x, h)$ is linear in h, it can be written as

$$Vf(x, h) = f'(x)h,$$

where f' satisfies $f' : X \to \{z \,|\, z : X \to \mathbb{R} \text{ is linear}\}$ and where $f'(x)$ is called the Gâteaux derivative of f at x. Note that $f'(x)$ is linear for fixed x and the domain of definition is $\mathcal{D}(f)$. Moreover, if $Vf(x, h)$ is bounded in h (thus $f'(x)$ is a linear bounded functional for fixed x, i.e. $f'(x) \in X^*$), then the gradient of the functional f at x can be defined as

$$Vf(x, h) = \langle \text{grad } f(x), h\rangle.$$

Remark that $\mathrm{grad}\, f : X \to X^*$. The gradient of f in the point x is the Gâteaux derivative of f at x. This is the generalization of the gradient of a scalar function or a vector field. These definitions can be extended to functions from X to $\mathbb{R}^3$. The generalized mean value theorem 2.11.12 takes the following form.

Theorem 2.11.13 (Generalized mean value theorem for functionals)**.** *Let f be a nonlinear functional defined on the normed space X. There exists a θ, $0 < \theta < 1$, such that*

$$f(x+h) - f(x) = Vf(x+\theta h, h), \qquad \forall x, h \in X;$$

if the Gâteaux differential of f is linear. Moreover, if the Gâteaux differential of f is also bounded, then there exists a θ, $0 < \theta < 1$, such that

$$f(x+h) - f(x) = \langle \mathrm{grad}\ f(x+\theta h), h \rangle, \qquad \forall x, h \in X.$$

Example 2.11.14. *Let X be a Hilbert space with a norm $\|\cdot\|^2 = (\cdot, \cdot)$. Then for $f(x) = \|x\|^2$, the Gâteaux differential is given by*

$$Vf(x,h) = \lim_{t \to 0} \frac{f(x+th) - f(x)}{t} = \lim_{t \to 0} \frac{\|x+th\|^2 - \|x\|^2}{t} = 2\,(x,h)\,.$$

The Gâteaux differential of f is linear and bounded in h for fixed x, i.e.

$$\mathrm{grad}\, f(x) = 2x.$$

Furthermore, $\mathrm{grad}\,\|x\| = \frac{x}{\|x\|}$. *Thus* $\|\mathrm{grad}\,\|x\|\| = 1$, $(\mathrm{grad}\,\|x\|\,, x) = \|x\|$ *and* $\mathrm{grad}\,\|\alpha x\| = \mathrm{sign}(\alpha)\,\|x\|$. *The latter results are also valid when X is not Hilbert, but then the derivation is more complicated due to the lack of an inner product.*

Example 2.11.15. *Consider the function $\mathbf{a}(\mathbf{x}) = m(|\mathbf{x}|_{\mathrm{e}})\mathbf{x}$ from $\mathbb{R}^3$ to $\mathbb{R}^3$, where $m(s)$ is a known real differentiable function. Then*

$$V\mathbf{a}(\mathbf{x},\mathbf{h}) = m\left(|\mathbf{x}|_{\mathrm{e}}\right)\mathbf{h} + m'\left(|\mathbf{x}|_{\mathrm{e}}\right)\frac{\mathbf{x}\cdot\mathbf{h}}{|\mathbf{x}|_{\mathrm{e}}}\mathbf{x}.$$

This is also valid when, instead of $\mathbb{R}^3$, a Hilbert space X is considered. Note that $V\mathbf{a}(\mathbf{x},\cdot)$ is linear and bounded for fixed $\mathbf{x} \in \mathbb{R}^3$. Therefore,

$$V\mathbf{a}(\mathbf{x},\mathbf{h}) = \langle \mathrm{grad}\ \mathbf{a}(\mathbf{x}), \mathbf{h} \rangle = \mathrm{grad}\ \mathbf{a}(\mathbf{x}) \cdot \mathbf{h},$$

where

$$\mathrm{grad}\ \mathbf{a}(\mathbf{x}) = m\left(|\mathbf{x}|_{\mathrm{e}}\right) I_3 + \frac{m'\left(|\mathbf{x}|_{\mathrm{e}}\right)}{|\mathbf{x}|_{\mathrm{e}}}\mathbf{x} \otimes \mathbf{x},$$

with I_3 the three-dimensional unit matrix and $\mathbf{x} \otimes \mathbf{x}$ defined by $(\mathbf{x} \otimes \mathbf{x})_{ij} = x_i x_j$. Moreover, by the generalized mean value theorem, it is possible to deduce that there exists some $\theta \in (0,1)$ such that

$$\begin{aligned}
[\mathbf{a}(\mathbf{x}+\mathbf{h}) - \mathbf{a}(\mathbf{x})] \cdot \mathbf{h} &= \langle \mathrm{grad}\ \mathbf{a}(\mathbf{x}+\theta\mathbf{h}), \mathbf{h} \rangle \cdot \mathbf{h} \\
&= m\left(|\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}}\right)|\mathbf{h}|_{\mathrm{e}}^2 + m'\left(|\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}}\right)\frac{\left((\mathbf{x}+\theta\mathbf{h})\cdot\mathbf{h}\right)^2}{|\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}}}.
\end{aligned}$$

Now, a specific choice for $m(s)$ is made. If $m(s) = s^{\alpha-1}$, then it is holds trivially that $[\mathbf{a}(\mathbf{x}+\mathbf{h}) - \mathbf{a}(\mathbf{x})] \cdot \mathbf{h} \geqslant 0$ for all $\mathbf{x}, \mathbf{h} \in \mathbb{R}^3$ when $\alpha \geqslant 1$. For $\alpha \in [0, 1)$, it holds that

$$\begin{aligned}[\mathbf{a}(\mathbf{x}+\mathbf{h}) - \mathbf{a}(\mathbf{x})] \cdot \mathbf{h} &\geqslant [m(|\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}}) - |m'(|\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}})| \, |\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}}] \, |\mathbf{h}|_{\mathrm{e}}^2 \\ &= \alpha \, |\mathbf{x}+\theta\mathbf{h}|_{\mathrm{e}}^{\alpha-1} \, |\mathbf{h}|_{\mathrm{e}}^2 .\end{aligned}$$

Thus $\mathbf{a}$ is monotone if $m(s) = s^{\alpha-1}$ for $\alpha \geqslant 0$. From Lemma 2.3.16 follows that $\mathbf{a}$ is strictly monotone for $\alpha \in [1, \infty)$. To obtain strict monotonicity when $\alpha \in (0, 1)$, the function m must be bounded below. Let us redefine the function m when $\alpha \in (0, 1)$ as $m(s) = \max\left\{s^{\alpha-1}, R^{\alpha-1}\right\}$ for $s > 0$ and given large $R > 0$. Then, it holds that

$$m(s) - |m'(s)| \, s \geqslant \frac{\min\{\alpha, 1\}}{R^{\alpha-1}} = \frac{\alpha}{R^{\alpha-1}}, \qquad \forall s > 0.$$

Thus

$$[\mathbf{a}(\mathbf{x}+\mathbf{h}) - \mathbf{a}(\mathbf{x})] \cdot \mathbf{h} \geqslant \frac{\alpha}{R^{\alpha-1}} \, |\mathbf{h}|_{\mathrm{e}}^2 ,$$

i.e. $\mathbf{a}$ is strictly monotone.

2.11.3.2 Handling a nonlinear function: auxiliary tools

Definition 2.11.3. *The potential of a continuous function $\beta : \mathbb{R} \to \mathbb{R}$ is defined by*

$$\Phi_\beta(z) = \int_0^z \beta(s) \, \mathrm{d}s, \qquad z \in \mathbb{R}.$$

Lemma 2.11.2. *Let $\Phi_\beta(z)$ be as in Definition 2.11.3 and β a.e. differentiable.*

(i) If β is monotone, i.e. $\beta'(s) \geqslant 0$ for a.a. $s \in \mathbb{R}$, then for all $z_1, z_2 \in \mathbb{R}$, it holds that

$$\beta(z_1)(z_2 - z_1) \leqslant \Phi_\beta(z_2) - \Phi_\beta(z_1) \leqslant \beta(z_2)(z_2 - z_1). \tag{2.19}$$

Moreover, $\Phi_\beta(z)$ is convex.

(ii) If $\beta(0) = 0$, β is monotone and Lipschitz continuous with Lipschitz constant L_β, i.e. $0 \leqslant \beta'(s) \leqslant L_\beta$ for a.a. $s \in \mathbb{R}$, then

$$0 \leqslant \frac{\beta^2(z)}{2L_\beta} \leqslant \Phi_\beta(z) \leqslant \frac{L_\beta z^2}{2}, \quad \forall z \in \mathbb{R}. \tag{2.20}$$

(iii) If $\beta(0) = 0$, β is strictly monotone, i.e. $\beta'(s) \geqslant \beta_0 > 0$ for a.a. $s \in \mathbb{R}$ and, β satisfies the growth condition $|\beta(s)| \leqslant \beta_1(1 + |s|)$ for all $s \in \mathbb{R}$, then

$$0 \leqslant \frac{\beta_0 z^2}{2} \leqslant \Phi_\beta(z) \leqslant \beta_1 \left(|z| + \frac{z^2}{2} \right), \qquad \forall z \in \mathbb{R}. \tag{2.21}$$

(iv) If $\beta(0) = 0$, β is strictly monotone and Lipschitz continuous with Lipschitz constant L_β, i.e. $0 < \beta_0 \leqslant \beta'(s) \leqslant L_\beta$ for a.a. $s \in \mathbb{R}$, then

$$0 \leqslant \frac{\beta_0 z^2}{2} \leqslant \Phi_\beta(z) \leqslant \frac{L_\beta z^2}{2}, \quad \forall z \in \mathbb{R}. \tag{2.22}$$

Proof. See for instance [57, Lemma 3.1]. □

The proof of the following lemma can be found in Lemma A.1.2 in Appendix A.

Lemma 2.11.3. *Let Ω be a nonempty bounded set in $\mathbb{R}^d$, $d \in \mathbb{N}$, and let $1 \leqslant p < \infty$. Suppose that*

$$u_n \to u \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad n \to \infty.$$

(i) If $h : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, then

$$h(u_n) \to h(u) \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad n \to \infty.$$

(ii) If $h : \mathbb{R} \to \mathbb{R}$ is continuous and satisfies the growth condition $|h(s)| \leqslant C_0(1 + s)$ for all $s \in \mathbb{R}$ with $C_0 > 0$, then there exists a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$h(u_{n_k}) \to h(u) \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad k \to \infty.$$

(iii) If $h : \mathbb{R} \to \mathbb{R}$ is continuous and linear (thus bounded), then there exists a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$h(u_{n_k}) \to h(u) \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad k \to \infty.$$

2.12 Rothe's method for evolution equations

The aim of this section is to present Rothe's method as a tool for solving evolution problems. Rothe's method is based on a time discretization and was introduced by Rothe [58] in 1930. Evolution problems are solved by a relatively simple technique using the results of corresponding elliptic problems. From these problems, an approximate solution for the original evolution problem is constructed. Afterwards, the convergence of the approximate solutions towards the exact solution of the evolution problem is proved. The advantage of Rothe's method is twofold: next to the existence and possible uniqueness of a solution to the original problem, also a numerical algorithm is contained in this method. The lecture notes of Kačur [36] (Rothe's method for monotone operators) and the book of Rektorys [59] (Rothe's method for linear operators) are good introductions to this topic.

Let us introduce Rothe's method for evolution problems with a simple example

$$\left\{ \begin{array}{rll} \partial_t u(\mathbf{x}, t) - \Delta u(\mathbf{x}, t) & = f(\mathbf{x}) & (\mathbf{x}, t) \in \Omega \times (0, T] =: Q_T, \\ u(\mathbf{x}, t) & = 0 & (\mathbf{x}, t) \in \partial\Omega \times (0, T] =: \Sigma_T, \\ u(\mathbf{x}, 0) & = u_0(\mathbf{x}) & \mathbf{x} \in \Omega, \end{array} \right. \tag{2.23}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and T denotes the final time. The time interval $(0,T]$ is divided into $n \in \mathbb{N}$ equidistant subintervals $(t_{i-1}, t_i]$ with time step $\tau = \frac{T}{n} < 1$, thus $t_i = i\tau, i = 0, \dots, n$. The backward Euler method is used to compute the approximate solution of the problem. The following standard notations for the discretized fields are introduced

$$u_i = u_i(\mathbf{x}) \approx u(\mathbf{x}, t_i) \quad \text{and} \quad \partial_t u(t_i) \approx \delta u_i = \frac{u_i - u_{i-1}}{\tau}.$$

For $i = 1, \dots, n$, the following elliptic equations are solved

$$\left\{ \begin{array}{rll} \delta u_i - \Delta u_i & = f & \text{in } \Omega, \\ u_i & = 0 & \text{on } \partial\Omega, \end{array} \right. \tag{2.24}$$

provided that the solution u_{i-1} on the previous time step is known and $u_0 = u_0(\mathbf{x})$.

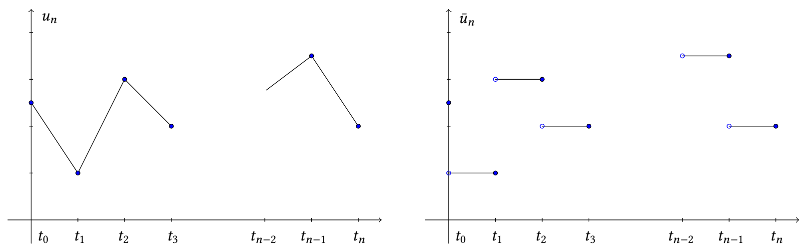
Next, the following so-called Rothe's functions are constructed (see Figure 2.4): the piecewise linear in time functions u_n

$$u_n(t) = \begin{cases} u_0 & \text{for } t = 0 \\ u_{i-1} + (t - t_{i-1})\delta u_i & \text{for } t \in (t_{i-1}, t_i] \end{cases}$$

and the step functions $\overline{u}_n$

$$\overline{u}_n(t) = \begin{cases} u_0 & \text{for } t = 0 \\ u_i & \text{for } t \in (t_{i-1}, t_i]. \end{cases}$$

Due to the construction of these functions, this method is also called the semidis-



(a) (b)

Figure 2.4: (a) Rothe's piecewise linear in time function u_n; (b) Rothe's piecewise constant function $\overline{u}_n$.

cretization in time.

The goal of Rothe's method is to obtain the existence of a weak solution to problem (2.23). Hence, the variational formulation of this problem needs to be well-defined (each term needs to be finite). When establishing a variational formulation, it is common to obtain some information about the solution by studying its natural stability. In doing this, it is assumed for the moment that the solution u to problem (2.23) exists.

The given PDE in (2.23) is multiplied with u and the result is integrated over the space and time domain, i.e.

$$\int_0^\eta \int_\Omega \partial_t u(\mathbf{x},t) u(\mathbf{x},t) \;\mathrm{d}\mathbf{x}\,\mathrm{d}t - \int_0^\eta \int_\Omega \Delta u(\mathbf{x},t) u(\mathbf{x},t) \;\mathrm{d}\mathbf{x}\,\mathrm{d}t \\ = \int_0^\eta \int_\Omega f(\mathbf{x}) u(\mathbf{x},t) \;\mathrm{d}\mathbf{x}\,\mathrm{d}t, \qquad \eta \in (0,T]. \quad (2.25)$$

From now on, to simplify the notations, the standard inner products in $\mathrm{L}^2(\Omega)$ and $\mathbf{L}^2(\Omega)$ are denoted by $(\cdot,\cdot)$ and their induced norms are denoted by $\|\cdot\|$. Using Cauchy's and Young's inequalities, the right-hand side (RHS) of (2.25) can be estimated as

$$\left|\int_0^\eta (f, u(t))\,\mathrm{d}t\right| \leqslant \frac{T}{2}\|f\|^2 + \frac{1}{2}\int_0^\eta \|u(t)\|^2\,\mathrm{d}t.$$

Note that $u\partial_t u = \frac{1}{2}\partial_t u^2$. Assuming $f, u_0 \in \mathrm{L}^2(\Omega)$, using Green's theorem and Grönwall's inequality, it holds that there exists a positive constant C such that

$$\max_{t\in[0,T]} \|u(t)\|^2 + \int_0^T \|\nabla u(t)\|^2\,\mathrm{d}t \leqslant C\left(\|f\|, \|u_0\|\right). \quad (2.26)$$

Alternatively, using the ε-Young's and the Friedrichs inequality, the RHS can be estimated as

$$\left|\int_0^\eta (f, u(t))\,\mathrm{d}t\right| \leqslant C_\varepsilon \|f\|^2 + \varepsilon \int_0^\eta \|\nabla u(t)\|^2\,\mathrm{d}t.$$

Using this estimate, Grönwall's argument is not needed to obtain estimate (2.26) (fix $\varepsilon < 1$).

Remark 2.12.1. *The information* (2.26) *about the solution can also be obtained without the use of both the Grönwall inequality and the Friedrichs inequality. Using the identity*

$$(\partial_t u(t), u(t)) = \tfrac{1}{2}\partial_t \|u(t)\|^2 = \|u(t)\|\,\partial_t\|u(t)\|,$$

it holds that (multiply (2.23) *with u, integrate the result over the space domain and apply Green's theorem)*

$$\|u(t)\|\,\partial_t\|u(t)\| + \|\nabla u(t)\|^2 \leqslant \|f\|\,\|u(t)\|, \qquad t \in (0,T].$$

It follows that

$$\partial_t \|u(t)\| \leqslant \partial_t \|u(t)\| + \frac{\|\nabla u(t)\|^2}{\|u(t)\|} \leqslant \|f\|, \qquad t \in (0,T],$$

assuming that $\|u(t)\| \neq 0$ *for all* $t \in [0,T]$. *Therefore,*

$$\|u(t)\| \leqslant t\|f\| + \|u_0\| \leqslant C(\|f\|, \|u_0\|), \qquad \forall t \in [0,T].$$

Due to (2.25)*, this estimate implies that*

$$\int_0^T \|\nabla u(t)\|^2 \,\mathrm{d}t \leqslant C(\|f\|, \|u_0\|).$$

The estimate (2.26) gives no information about $\partial_t u$. It holds that

$$(\partial_t u(t), \varphi) = (f, \varphi) - (\nabla u(t), \nabla \varphi)$$

for all $\varphi \in \mathrm{H}_0^1(\Omega)$ and for a.a. $t \in (0,T)$. The integral in the LHS can be interpreted in the sense of duality, i.e. seeing $\partial_t u(t)$ as an operator from $\mathrm{H}_0^1(\Omega)$ to $\mathbb{R}$. The dual norm in $\mathrm{H}_0^1(\Omega)^*$ is defined as

$$\|\partial_t u(t)\|_{\mathrm{H}_0^1(\Omega)^*} = \sup_{\varphi \in \mathrm{H}_0^1(\Omega)} \frac{(\partial_t u(t), \varphi)}{\|\varphi\|_{\mathrm{H}_0^1(\Omega)}}.$$

Therefore, using (2.26), it holds that

$$\int_0^T \|\partial_t u(s)\|^2_{\mathrm{H}_0^1(\Omega)^*} \,\mathrm{d}s \leqslant C. \tag{2.27}$$

This means that $\partial_t u \in \mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)^*\right)$ if $u_0 \in \mathrm{L}^2(\Omega)$. Another approach is to multiply the given PDE in (2.23) with $\partial_t u$ and to integrate the result over the space and time domain. Assuming $f \in \mathrm{L}^2(\Omega)$ and $u_0 \in \mathrm{H}^1(\Omega)$, using Green's theorem and the Cauchy's and ε-Young's inequality, there exists a positive constant C such that

$$\int_0^T \|\partial_t u(s)\|^2 \,\mathrm{d}s + \max_{t \in [0,T]} \|\nabla u(t)\|^2 \leqslant C\left(\|f\|, \|u_0\|_{\mathrm{H}^1(\Omega)}\right). \tag{2.28}$$

The estimates (2.26), (2.27) and (2.28) show that the definition of the variational formulation depends on the assumption on the initial condition.

Definition 2.12.1. *The variational formulation of Problem* (2.23) *is the following:*

(i) Given $f, u_0 \in \mathrm{L}^2(\Omega)$*, find* $u(t) \in \mathrm{H}_0^1(\Omega)$ *with* $\partial_t u(t) \in \mathrm{H}_0^1(\Omega)^*$ *such that*

$$(\partial_t u(t), \varphi) + (\nabla u(t), \nabla \varphi) = (f, \varphi), \qquad \forall \varphi \in \mathrm{H}_0^1(\Omega), \tag{2.29}$$

for a.a. $t \in (0,T)$*;*

(ii) Given $f \in \mathrm{L}^2(\Omega)$ and $u_0 \in \mathrm{H}^1(\Omega)$, find $u(t) \in \mathrm{H}^1_0(\Omega)$ with $\partial_t u(t) \in \mathrm{L}^2(\Omega)$ such that

$$(\partial_t u(t), \varphi) + (\nabla u(t), \nabla \varphi) = (f, \varphi), \qquad \forall \varphi \in \mathrm{H}^1_0(\Omega), \tag{2.30}$$

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

In the following, a solution to the variational formulation in Definition 2.12.1(i) is searched. The time derivative in equations (2.29) and (2.30) is interpreted as the generalized derivative on $(0, T)$ and so both equations need to be satisfied only for a.a. $t \in (0, T)$. The solution method is divided into the following steps:

- uniqueness of a solution;
- solving elliptic problems;
- a priori estimates (Grönwall argument);
- convergence of the Rothe's functions (compactness argument);
- error estimates.

2.12.1 Uniqueness of a solution

In this section, the uniqueness of a solution is proved. If a solution to an evolution problem is not unique, then the convergence of the Rothe's functions to the exact solution cannot be guaranteed. Assume that there are two solutions u_1 and $u_2 \in \mathrm{H}^1_0(\Omega)$. Then the difference $u = u_1 - u_2$ satisfies (2.29) with $f = u_0 = 0$. Putting $\varphi = u(t)$ in (2.29) and integrating the result in time over $(0, \eta) \subset (0, T)$ gives

$$\frac{1}{2} \int_0^\eta \frac{\mathrm{d}}{\mathrm{d}s} \|u(s)\|^2 \, \mathrm{d}s + \int_0^\eta \|\nabla u(s)\|^2 \, \mathrm{d}s = 0. \tag{2.31}$$

This implies that $\|u(\eta)\|^2 = 0$ for a.a. $\eta \in (0, T)$. Therefore, the solution is unique, i.e $u_1 = u_2$ a.e. in Q_T. Note that the first term in (2.31) makes sense due to Lemma 2.9.5(iii).

2.12.2 Solving elliptic problems

The variational formulation of the discrete problems (2.24) (put $t = t_i$ in equation (2.29)) is: given $f, u_0 \in \mathrm{L}^2(\Omega)$, find $u_i \in \mathrm{H}^1_0(\Omega)$ with $\delta u_i \in \mathrm{H}^1_0(\Omega)^*$ such that

$$(\delta u_i, \varphi) + (\nabla u_i, \nabla \varphi) = (f, \varphi), \quad \forall \varphi \in \mathrm{H}^1_0(\Omega), \quad i = 1, \ldots, n. \tag{2.32}$$

This problem is equivalent to solving the equation

$$a(u_i, \varphi) = F_{i-1}(\varphi), \qquad \forall \varphi \in \mathrm{H}^1_0(\Omega),$$

for any $i = 1, \ldots, n$, where

$$\begin{aligned} a &: \mathrm{H}^1_0(\Omega) \times \mathrm{H}^1_0(\Omega) \to \mathbb{R} : (u, v) \mapsto \left(\frac{u}{\tau}, v\right) + (\nabla u, \nabla v), \\ F_{i-1} &: \mathrm{H}^1_0(\Omega) \to \mathbb{R} : \varphi \mapsto (f, \varphi) + \left(\frac{u_{i-1}}{\tau}, \varphi\right). \end{aligned}$$

The bilinear form a is $\mathrm{H}_0^1(\Omega)$-elliptic and continuous. The function F_{i-1} is a bounded linear functional on $\mathrm{H}_0^1(\Omega)$ if $u_{i-1} \in \mathrm{L}^2(\Omega)$ and $f \in \mathrm{L}^2(\Omega)$. The existence and uniqueness of $u_i \in \mathrm{H}_0^1(\Omega)$, $i = 1, \ldots, n$, follows from Lemma 2.11.1 if $u_0 \in \mathrm{L}^2(\Omega)$. The next step is to study the stability (regularity) of u_i.

2.12.3 A priori estimates

The a priori estimates proved in this subsection serve as uniform bounds to prove convergence. They depend on the regularity of the initial condition u_0. First, suppose that $u_0 \in \mathrm{L}^2(\Omega)$. Putting $\varphi = u_i\tau$ in (2.32) and summing up the result for $i = 1, \ldots j$ with $1 \leqslant j \leqslant n$ gives

$$\sum_{i=1}^{j} (\delta u_i, u_i)\, \tau + \sum_{i=1}^{j} \|\nabla u_i\|^2 \tau = \sum_{i=1}^{j} (f, u_i)\, \tau.$$

An application of Abel's summation rule (see Lemma 2.3.15) gives

$$\sum_{i=1}^{j} (\delta u_i, u_i)\, \tau = \frac{1}{2} \|u_j\|^2 - \frac{1}{2} \|u_0\|^2 + \frac{1}{2} \sum_{i=1}^{j} \|u_i - u_{i-1}\|^2 .$$

The RHS can be estimated by using the Cauchy's, ε-Young's and Friedrichs' inequality as follows

$$\sum_{i=1}^{j} (f, u_i)\, \tau \leqslant C_\varepsilon \sum_{i=1}^{j} \|f\|^2 \tau + \varepsilon \sum_{i=1}^{j} \|\nabla u_i\|^2 \tau \leqslant C_\varepsilon + \varepsilon \sum_{i=1}^{j} \|\nabla u_i\|^2 \tau.$$

Fixing ε sufficiently small implies the existence of a positive constant C such that

$$\max_{1 \leqslant j \leqslant n} \|u_j\|^2 + \sum_{i=1}^{n} \|\nabla u_i\|^2 \tau + \sum_{i=1}^{n} \|u_i - u_{i-1}\|^2 \leqslant C. \tag{2.33}$$

Due to the application of the Friedrichs inequality, the use of Grönwall's argument could be avoided for this problem. This is rarely the case and depends on the problem under consideration. Moreover, the equation (2.32) implicitly defines

$$(\delta u_i, \varphi) = (f, \varphi) - (\nabla u_i, \nabla \varphi), \qquad \varphi \in \mathrm{H}_0^1(\Omega).$$

The integral in the LHS can be interpreted in the sense of duality, i.e. seeing δu_i as an element of $\mathrm{H}_0^1(\Omega)^*$. Using

$$\|\delta u_i\|_{\mathrm{H}_0^1(\Omega)^*} = \sup_{\varphi \in \mathrm{H}_0^1(\Omega)} \frac{(\delta u_i, \varphi)}{\|\varphi\|_{\mathrm{H}_0^1(\Omega)}}$$

and (2.33), it is easy to check that

$$\sum_{i=1}^{n} \|\delta u_i\|^2_{\mathrm{H}_0^1(\Omega)^*} \tau \leqslant C. \tag{2.34}$$

Secondly, let $u_0 \in \mathrm{H}^1(\Omega)$ and put $\varphi = \delta u_i \tau$ in (2.32). It is not difficult to prove that there exists a positive constant C such that

$$\sum_{i=1}^{n} \|\delta u_i\|^2 \tau + \max_{1 \leqslant j \leqslant n} \|\nabla u_j\|^2 + \sum_{i=1}^{n} \|\nabla u_i - \nabla u_{i-1}\|^2 \leqslant C. \tag{2.35}$$

Finally, when $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega)$, a uniform bound on δu_i can be obtained. Subtracting (2.32) for $i = i-1$ from (2.32), afterwards setting $\varphi = \delta u_i$ and take the sum for $i = 1, \ldots j$ the following equality is obtained

$$\sum_{i=1}^{j} (\delta u_i - \delta u_{i-1}, \delta u_i) + \sum_{i=1}^{j} (\nabla u_i - \nabla u_{i-1}, \nabla \delta u_i) = 0. \tag{2.36}$$

Note that δu_0 is not defined in this equality. To overcome this difficulty, a so-called compatibility condition is needed. In fact, it is necessary to assume that the variational formulation (2.29) is satisfied at $t = 0$, i.e.

$$(\partial_t u(0), \varphi) + (\nabla u_0, \nabla \varphi) = (f, \varphi), \quad \forall \varphi \in \mathrm{H}_0^1(\Omega).$$

For this, it is required that $u_0 \in \mathrm{H}_0^1(\Omega)$. Set $\delta u_0 := \partial_t u(0)$. Applying Green's theorem backwards gives for all $\varphi \in \mathrm{H}_0^1(\Omega)$ that

$$(\delta u_0, \varphi) = (f, \varphi) + (\Delta u_0, \varphi). \tag{2.37}$$

The term δu_0 can be seen as a functional on $\mathrm{H}_0^1(\Omega)$. The RHS is a linear and bounded functional on $\mathrm{H}_0^1(\Omega)$ if $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega)$. This implies that the RHS can be extended to a functional $\widetilde{\delta u_0}$ on $\mathrm{L}^2(\Omega)$ by the Hahn-Banach theorem. Moreover, it holds that

$$\left\|\widetilde{\delta u_0}\right\| = \sup_{\substack{\varphi \in \mathrm{L}^2(\Omega) \\ \|\varphi\| \leqslant 1}} \left(\widetilde{\delta u_0}, \varphi\right) = \sup_{\substack{\varphi \in \mathrm{H}_0^1(\Omega) \\ \|\varphi\| \leqslant 1}} (\delta u_0, \varphi) \lesssim 1,$$

i.e. $\widetilde{\delta u_0} \in \mathrm{L}^2(\Omega)$. Finally, employing the density of $\mathrm{C}_0^\infty(\Omega)$ in $\mathrm{H}_0^1(\Omega)$ and Theorem 2.8.1 on equation (2.37) result in

$$\widetilde{\delta u_0} = f + \Delta u_0, \quad \text{a.e. in } \Omega \quad \text{or} \quad \widetilde{\delta u_0} = f + \Delta u_0 \in \mathrm{L}^2(\Omega).$$

From now on, $\widetilde{\delta u_0}$ and δu_0 are identified. Then, after an application of Abel's summation rule on the first term in the LHS of (2.36), the following a priori estimated is obtained

$$\max_{1 \leqslant j \leqslant n} \|\delta u_j\|^2 + \sum_{i=1}^{n} \|\delta u_i - \delta u_{i-1}\|^2 + \sum_{i=1}^{n} \|\nabla \delta u_i\|^2 \tau \leqslant C. \tag{2.38}$$

The results of the a priori estimates are summarized in the following lemma.

Lemma 2.12.1. *Suppose that $f \in \mathrm{L}^2(\Omega)$.*

(i) Let $u_0 \in \mathrm{L}^2(\Omega)$. Then, there exists a positive constant C such that

$$\max_{1\leqslant j\leqslant n} \|u_j\|^2 + \sum_{i=1}^n \|\nabla u_i\|^2 \tau + \sum_{i=1}^n \|u_i - u_{i-1}\|^2 \leqslant C.$$

Moreover, there exists a positive constant C such that

$$\sum_{i=1}^n \|\delta u_i\|^2_{\mathrm{H}^1_0(\Omega)^*} \tau \leqslant C.$$

(ii) Let $u_0 \in \mathrm{H}^1(\Omega)$. Then, there exists a positive constant C such that

$$\sum_{i=1}^n \|\delta u_i\|^2 \tau + \max_{1\leqslant j\leqslant n} \|\nabla u_j\|^2 + \sum_{i=1}^n \|\nabla u_i - \nabla u_{i-1}\|^2 \leqslant C.$$

(iii) Let $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$. Then, there exists a positive constant C such that

$$\max_{1\leqslant j\leqslant n} \|\delta u_j\|^2 + \sum_{i=1}^n \|\delta u_i - \delta u_{i-1}\|^2 + \sum_{i=1}^n \|\nabla \delta u_i\|^2 \tau \leqslant C.$$

The next step is the reformulation of the a priori estimates to the Rothe functions $u_n : [0,T] \to \mathrm{L}^2(\Omega)$ and $\overline{u}_n : [0,T] \to \mathrm{L}^2(\Omega)$. The discrete variational formulation (2.32) can be rewritten for a.a. $t \in [0,T]$ in terms of u_n and $\overline{u}_n$ as follows

$$(\partial_t u_n(t), \varphi) + (\nabla \overline{u}_n(t), \nabla\varphi) = (f, \varphi), \quad \varphi \in \mathrm{H}^1_0(\Omega). \tag{2.39}$$

If $u_0 \in \mathrm{L}^2(\Omega)$, then a priori estimates (2.33) and (2.34) can be rewritten for any $t \in [0,T]$ and $n \in \mathbb{N}$ as

$$\begin{aligned} \|u_n(t)\|^2 + \|\overline{u}_n(t)\|^2 &+ \int_0^T \|\overline{u}_n(s)\|^2_{\mathrm{H}^1_0(\Omega)} \,\mathrm{d}s \\ &+ \int_0^T \|\partial_t u_n(s)\|^2_{\mathrm{H}^1_0(\Omega)^*} \,\mathrm{d}s + \sum_{i=1}^n \left\| \int_{t_{i-1}}^{t_i} \partial_t u_n(s) \,\mathrm{d}s \right\|^2 \leqslant C. \end{aligned} \tag{2.40}$$

If $u_0 \in \mathrm{H}^1(\Omega)$, then a priori estimates (2.33) and (2.35) can be rewritten for any $t \in [0,T]$ and $n \in \mathbb{N}$ as

$$\begin{aligned} \|u_n(t)\|^2_{\mathrm{H}^1_0(\Omega)} + \|\overline{u}_n(t)\|^2_{\mathrm{H}^1_0(\Omega)} &+ \int_0^T \|\partial_t u_n(s)\|^2 \,\mathrm{d}s \\ &+ \sum_{i=1}^n \left\| \int_{t_{i-1}}^{t_i} \partial_t u_n(s) \,\mathrm{d}s \right\|^2_{\mathrm{H}^1_0(\Omega)} \leqslant C. \end{aligned} \tag{2.41}$$

If $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, then for any $t \in [0,T]$ and $n \in \mathbb{N}$ it holds that

$$\|\partial_t u_n(t)\|^2 + \int_0^T \|\nabla \partial_t u_n(s)\|^2 \, \mathrm{d}s \leqslant C. \tag{2.42}$$

The a priori estimates also imply certain relations between the different Rothe functions.

2.12.3.1 Relations between the different Rothe functions

Note that $u_n(0) - \overline{u}_n(0) = 0$. For all $t \in (t_{i-1}, t_i]$ with $1 \leqslant i \leqslant n$, it holds that

$$\begin{aligned} |u_n(t) - \overline{u}_n(t)| &= |u_{i-1} + (t - t_{i-1})\delta u_i - u_i| \\ &= |(t - t_{i-1} - \tau)\delta u_i| \\ &= |(t - t_i)\delta u_i| \\ &\leqslant \tau\, |\delta u_i| \qquad (2.43) \\ &= |u_i - u_{i-1}|. \qquad (2.44) \end{aligned}$$

Using a priori estimate (2.33) gives (if $u_0 \in \mathrm{L}^2(\Omega)$)

$$\lim_{n\to\infty} \|u_n - \overline{u}_n\|^2_{\mathrm{L}^2((0,T),\mathrm{L}^2(\Omega))} \leqslant \lim_{n\to\infty} \tau \sum_{i=1}^n \|u_i - u_{i-1}\|^2 \leqslant \lim_{n\to\infty} \frac{C}{n} = 0,$$

such that $\overline{u}_n$ and u_n have the same limit (if it exists) in $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$. This result can also be obtained as follows. A priori estimate (2.33) also implies that (if the series $\sum a_n$ converges, then $\lim_{n\to\infty} a_n = 0$)

$$\lim_{n\to\infty} \|u_n(t) - \overline{u}_n(t)\| = 0, \qquad \forall t \in [0,T].$$

Now, the order of the difference between the piecewise linear interpolant and the piecewise constant interpolant is summarized depending on the regularity of the initial condition. Employing (2.44) and a priori estimate (2.33), it holds that

$$\int_0^T \|u_n(t) - \overline{u}_n(t)\|^2 \, \mathrm{d}t \leqslant \tau \sum_{i=1}^n \|u_i - u_{i-1}\|^2 \lesssim \tau \tag{2.45}$$

if $u_0 \in \mathrm{L}^2(\Omega)$. Moreover, using (2.43), (2.44) and (2.35), it is clear that

$$\int_0^T \|u_n(t) - \overline{u}_n(t)\|^2 \, \mathrm{d}t \leqslant \tau^2 \sum_{i=1}^n \|\delta u_i\|^2 \tau \lesssim \tau^2 \tag{2.46}$$

and

$$\int_0^T \|\nabla u_n(t) - \nabla \overline{u}_n(t)\|^2 \, \mathrm{d}t \leqslant \tau \sum_{i=1}^n \|\nabla (u_i - u_{i-1})\|^2 \lesssim \tau \tag{2.47}$$

if $u_0 \in \mathrm{H}^1(\Omega)$. Finally, if $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, then it follows from (2.38) that

$$\max_{t\in[0,T]} \|u_n(t) - \overline{u}_n(t)\| \leqslant \tau \max_{1\leqslant i\leqslant n} \max_{t\in(t_{i-1},t_i]} \|\delta u_i\| \lesssim \tau, \tag{2.48}$$

and

$$\int_0^T \|\nabla u_n(t) - \nabla \overline{u}_n(t)\|^2 \,\mathrm{d}t \leqslant \tau^2 \sum_{i=1}^n \|\nabla \delta u_i\|^2 \,\tau \lesssim \tau^2. \tag{2.49}$$

2.12.4 Convergence

The next step is to pass to the limit $n \to \infty$ in (2.39). For this, a compactness argument and certain convergence principles are needed.

The compactness argument (which leads to an evolution triple) in this example is

$$\mathrm{H}^1_0(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Omega) \cong \mathrm{L}^2(\Omega)^* \hookrightarrow\hookrightarrow \mathrm{H}^1_0(\Omega)^*,$$

see the Rellich-Kondrachov Compactness Theorem 2.9.22. Now, a distinction is made based on the properties of the initial condition.

If $u_0 \in \mathrm{L}^2(\Omega)$, as a consequence of (2.40) and Remark 2.9.2, it holds that $\{u_n\}$ is uniformly equicontinuous and $u_n(t)$ is uniformly bounded for all $t \in [0, T]$, i.e.

$$\|u_n(t) - u_n(s)\|_{\mathrm{H}^1_0(\Omega)^*} \leqslant C\,|t-s|^{\frac{1}{2}} \quad \text{and} \quad \max_{t\in[0,T]} \|u_n(t)\| \leqslant C,$$

uniformly for $n \in \mathbb{N}$ and $t, s \in [0, T]$. The uniform boundedness of $u_n(t)$ in $\mathrm{L}^2(\Omega)$ and the compactness argument $\mathrm{L}^2(\Omega) \hookrightarrow\hookrightarrow \mathrm{H}^1_0(\Omega)^*$ yield that there exists a subsequence $\{u_{n_k}(t)\}$ of $\{u_n(t)\}$ such that

$$u_{n_k}(t) \to u(t) \quad \text{in} \quad \mathrm{H}^1_0(\Omega)^* \quad \text{for all } t \in [0, T],$$

i.e. the set $\{u_n(t)\}$ is relatively compact in $\mathrm{H}^1_0(\Omega)^*$ for any $t \in [0, T]$. Therefore, the generalized Arzelà-Ascoli theorem 2.5.6 implies the existence of a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that

$$u_{n_k} \to u \quad \text{in } \mathrm{C}\left([0,T], \mathrm{H}^1_0(\Omega)^*\right).$$

Furthermore, the reflexivity of $\mathrm{L}^2(\Omega)$ implies by Theorem 2.4.30 that also

$$u_{n_k}(t) \rightharpoonup u(t) \quad \text{in} \quad \mathrm{L}^2(\Omega) \quad \text{for all } t \in [0, T].$$

By

$$\lim_{n\to\infty} \|u_n(t) - \overline{u}_n(t)\| = 0, \quad \text{for all } t \in [0, T],$$

it is also clear that

$$\overline{u}_{n_k}(t) \rightharpoonup u(t) \quad \text{in} \quad \mathrm{L}^2(\Omega) \quad \text{for all } t \in [0, T].$$

Moreover,

$$\overline{u}_{n_k}(t) \rightharpoonup u(t) \quad \text{in } \mathrm{H}_0^1(\Omega) \quad \text{for a.a. } t \in [0,T], \tag{2.50}$$

thanks to $\int_0^T \|\overline{u}_n(s)\|^2_{\mathrm{H}_0^1(\Omega)} \,\mathrm{d}s \leqslant C$. Therefore, $u \in \mathrm{L}^\infty\left((0,T), \mathrm{H}_0^1(\Omega)\right)$. Finally, a convergence result is needed for $\partial_t u_n$. The space $\mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)^*\right)$ is reflexive and separable because $\mathrm{H}_0^1(\Omega)$ is separable and reflexive (apply successively Theorem 2.9.1, Lemma 2.4.18 and Theorem 2.9.10). Therefore, the estimate $\int_0^T \|\partial_t u_n(s)\|^2_{\mathrm{H}_0^1(\Omega)^*} \,\mathrm{d}s$ from (2.40) implies that

$$\partial_t u_{n_k} \rightharpoonup \partial_t u \quad \text{in} \quad \mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)^*\right), \tag{2.51}$$

due to Theorem 2.9.11.

Remark 2.12.2. *Note that also*

$$u_{n_k} \to u \quad \textit{in } \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right) \textit{ as } k \to +\infty,$$

when $u_0 \in \mathrm{L}^2(\Omega)$. As a consequence (by Example 2.7.8), there exists a subsequence $\{u_{n_{k_l}}\}$ of $\{u_{n_k}\}$ such that

$$u_{n_{k_l}}(\mathbf{x},t) \to u(\mathbf{x},t) \quad \textit{as } l \to \infty, \qquad \textit{a.e. in } \Omega \times (0,T).$$

This follows from $u_{n_k} \rightharpoonup u$ in $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ and $\|u_{n_k}\|_{\mathrm{L}^2((0,T),\mathrm{L}^2(\Omega))} \to \|u\|_{\mathrm{L}^2((0,T),\mathrm{L}^2(\Omega))}$ due to Lemma 2.4.19(vii). Let us check the second convergence result. The duality pairing $\langle \cdot,\cdot\rangle_{\mathrm{H}_0^1(\Omega)^*\times \mathrm{H}_0^1(\Omega)}$ can be seen as a continuous extension of the inner product on $\mathrm{L}^2(\Omega)$ (see Theorem 2.9.15). This implies that

$$\int_0^T \|u_{n_k}(t)\|^2 \,\mathrm{d}t = \int_0^T (u_{n_k}(t), u_{n_k}(t) - \overline{u}_{n_k}(t)) \,\mathrm{d}t + \int_0^T \langle u_{n_k}(t), \overline{u}_{n_k}(t)\rangle_{\mathrm{H}_0^1(\Omega)^*\times \mathrm{H}_0^1(\Omega)}. \tag{2.52}$$

From equation (2.33)*, it follows that*

$$\left|\int_0^T (u_{n_k}(t), u_{n_k}(t) - \overline{u}_{n_k}(t)) \,\mathrm{d}t\right| \leqslant \sqrt{\int_0^T \|u_{n_k}(t)\|^2 \,\mathrm{d}t}\sqrt{\int_0^T \|u_{n_k}(t) - \overline{u}_{n_k}(t)\|^2 \,\mathrm{d}t} \lesssim \sqrt{\tau}.$$

Hence, taking the limit $\tau \to 0$ in (2.52) *implies that*

$$\int_0^T \|u_{n_k}(t)\|^2 \,\mathrm{d}t \to \int_0^T \|u(t)\|^2 \,\mathrm{d}t$$

because $\overline{u}_{n_k} \rightharpoonup u$ in $\mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)\right)$ and $u_{n_k} \to u$ in $\mathrm{C}\left([0,T], \mathrm{H}_0^1(\Omega)^*\right)$.

If $u_0 \in \mathrm{H}^1(\Omega)$, then the reasoning from the case where u_0 belongs to $\mathrm{L}^2(\Omega)$ can be repeated. Now, due to (2.41), it holds that

$$\|u_n(t) - u_n(s)\| \leqslant C\,|t-s|^{\frac{1}{2}} \quad \text{and} \quad \max_{t\in[0,T]} \|u_n(t)\|_{\mathrm{H}_0^1(\Omega)} \leqslant C.$$

The compactness argument is $\mathrm{H}_0^1(\Omega) \hookrightarrow\hookrightarrow \mathrm{L}^2(\Omega)$. Using the Arzelà-Ascoli theorem 2.5.6, it can be shown that there exists a

$$u \in \mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right) \cap \mathrm{L}^\infty\left((0,T), \mathrm{H}_0^1(\Omega)\right)$$

with $\partial_t u \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ and a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that

$$\begin{cases} u_{n_k} \to u, & \text{in} \quad \mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right), \\ u_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{H}_0^1(\Omega) \text{ for all } t \in [0,T], \\ \overline{u}_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{H}_0^1(\Omega) \text{ for all } t \in [0,T], \qquad (2.53a) \\ \partial_t u_{n_k} \rightharpoonup \partial_t u, & \text{in} \quad \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right). \qquad (2.53b) \end{cases}$$

By Example 2.7.8, there exists a subsequence $\{u_{n_{k_l}}\}$ of $\{u_{n_k}\}$ such that

$$u_{n_{k_l}}(\mathbf{x},t) \to u(\mathbf{x},t) \quad \text{as } l \to \infty, \qquad \forall t \in [0,T], \text{ for a.a. } \mathbf{x} \in \Omega.$$

Moreover, if $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega)$, then by the a priori estimate (2.41) and (2.42), it holds that

$$\|u_n(t) - u_n(s)\| \leqslant C\,|t-s| \quad \text{and} \quad \max_{t\in[0,T]} \|u_n(t)\|_{\mathrm{H}_0^1(\Omega)} \leqslant C.$$

Therefore, it holds that $\partial_t u \in \mathrm{L}^\infty\left((0,T), \mathrm{L}^2(\Omega)\right)$ and $u : [0,T] \to \mathrm{L}^2(\Omega)$ is Lipschitz continuous.

Now, the goal is to prove that the function u is a weak solution to problem (2.29). This can already be done when $u_0 \in \mathrm{L}^2(\Omega)$. The discrete variational formulation (2.32) can be rewritten for a.a. $t \in (0,T)$ in terms of u_{n_k} and $\overline{u}_{n_k}$ as follows

$$(\partial_t u_{n_k}(t), \varphi) + (\nabla \overline{u}_{n_k}(t), \nabla\varphi) = (f, \varphi)\,, \quad \forall \varphi \in \mathrm{H}_0^1(\Omega). \qquad (2.54)$$

Before passing to the limit $k \to \infty$ in (2.54), this equality is integrated in time over $(0,\eta) \subset [0,T]$ to obtain for all $\varphi \in \mathrm{H}_0^1(\Omega)$ that

$$\int_0^\eta (\partial_t u_{n_k}(t), \varphi)\,\mathrm{d}t + \int_0^\eta (\nabla \overline{u}_{n_k}(t), \nabla\varphi)\,\mathrm{d}t = \eta\,(f,\varphi)\,. \qquad (2.55)$$

Using (2.51), (2.50) and applying Theorem 2.7.9, it is clear that for $k \to \infty$, it holds that

$$\int_0^\eta (\partial_t u_{n_k}(t), \varphi)\,\mathrm{d}t \to \int_0^\eta (\partial_t u(t), \varphi)\,\mathrm{d}t,$$
$$\int_0^\eta (\nabla \overline{u}_{n_k}(t), \nabla\varphi)\,\mathrm{d}t \to \int_0^\eta (\nabla u(t), \nabla\varphi)\,\mathrm{d}t.$$

Therefore, it holds for all $\varphi \in \mathrm{H}_0^1(\Omega)$ that

$$\int_0^\eta (\partial_t u(t), \varphi)\,\mathrm{d}t + \int_0^\eta (\nabla u(t), \nabla\varphi)\,\mathrm{d}t = \eta\,(f, \varphi)\,.$$

which is valid for all $\eta \in [0, T]$. Differentiating this with respect to the time variable gives (2.29) for a.a. $t \in [0, T]$. The existence of a weak solution is proved. The convergence of Rothe's functions towards the weak solution (2.29) has been shown for a subsequence $\{u_{n_k}\}$. Nevertheless, taking into account the uniqueness of a solution and Lemma 2.4.20, it is clear that the whole Rothe's sequence $\{u_n\}$ converges in $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ towards the solution. If $u_0 \in \mathrm{H}^1(\Omega)$, then Rothe's sequence $\{u_n\}$ converges in $\mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right)$ towards the solution.

Theorem 2.12.1. *Let $f \in \mathrm{L}^2(\Omega)$.*

- *Assume that $u_0 \in \mathrm{L}^2(\Omega)$. Then there exists a unique weak solution*

  $$u \in \mathrm{L}^\infty\left((0,T), \mathrm{H}_0^1(\Omega)\right)$$

  with

  $$\partial_t u \in \mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)^*\right)$$

  to Problem (2.23).
- *Assume that $u_0 \in \mathrm{H}^1(\Omega)$. Then there exists a unique weak solution*

  $$u \in \mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right) \cap \mathrm{L}^\infty\left((0,T), \mathrm{H}_0^1(\Omega)\right)$$

  with

  $$\partial_t u \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$$

  to Problem (2.23).
- *Assume that $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega)$. Then there exists a unique weak solution*

  $$u \in \mathrm{Lip}((0,T), \mathrm{L}^2(\Omega)) \cap \mathrm{L}^\infty\left((0,T), \mathrm{H}_0^1(\Omega)\right)$$

  with

  $$\partial_t u \in \mathrm{L}^\infty\left((0,T), \mathrm{L}^2(\Omega)\right)$$

  to Problem (2.23).

Remark 2.12.3. *The proof that the whole Rothe's sequence $\{u_n\}$ converges to the solution is by contradiction. If it is not the case that $u_n \to u$ in $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ as $n \to \infty$, then there exists an $\varepsilon_0 > 0$ and a subsequence $\{u_{n'}\}$ such that*

$$\|u - u_{n'}\|_{\mathrm{L}^2((0,T),\mathrm{L}^2(\Omega))} > \varepsilon_0 \qquad \text{for all } u_{n'}. \tag{2.56}$$

The analysis in this subsection can be repeated on the sequence $\{u_{n'}\}$. By the uniqueness of the solution, the sequence $\{u_{n'}\}$ has a subsequence $\{u_{n''}\}$ convergent to u. That is a contradiction with (2.56).

2.12.5 Error estimates

In this subsection, error estimates between the semidiscrete solutions obtained by Rothe's method and the solution of the original problem are established. To obtain an error estimate in the example, it is sufficient that $u_0 \in \mathrm{L}^2(\Omega)$.

There are different possibilities to derive an error estimate. The classical approach is to subtract the variational formulation (2.39) from (2.29), i.e.

$$(\partial_t(u-u_n)(t),\varphi)+(\nabla(u-\overline{u}_n)(t),\nabla\varphi)=0, \qquad \forall\varphi\in \mathrm{H}^1_0(\Omega), \tag{2.57}$$

to make a choice for the test function and to integrate in time. Let us follow this approach. Choose $\varphi = u(t)-u_n(t)$ in (2.57) and integrate in time over $(0,\eta)\subset[0,T]$ to obtain

$$\int_0^\eta (\partial_t(u-u_n),u-u_n)+\int_0^\eta (\nabla(u-\overline{u}_n),\nabla(u-u_n))=0.$$

Now, the trick is to rewrite the second term by adding $\pm u_n$. This implies

$$\tfrac12\left\|(u-u_n)(\eta)\right\|^2+\int_0^\eta \left\|\nabla(u-u_n)\right\|^2=\int_0^\eta\left(\nabla(\overline{u}_n-u_n),\nabla(u-u_n)\right).$$

The RHS can be estimated as

$$\left|\int_0^\eta (\nabla(\overline{u}_n-u_n),\nabla(u-u_n))\right| \leqslant C_\varepsilon\int_0^\eta\left\|\nabla(\overline{u}_n-u_n)\right\|^2+\varepsilon\int_0^\eta\left\|\nabla(u-u_n)\right\|^2.$$

Therefore, fixing ε sufficiently small and employing bounds (2.47) and (2.49), it is clear that

$$\max_{t\in[0,T]}\left\|u(t)-u_n(t)\right\|^2+\int_0^T\left\|\nabla(u-u_n)\right\|^2\lesssim\tau$$

if $u_0\in \mathrm{H}^1(\Omega)$ and

$$\max_{t\in[0,T]}\left\|u(t)-u_n(t)\right\|^2+\int_0^T\left\|\nabla(u-u_n)\right\|^2\lesssim\tau^2$$

if $u_0\in \mathrm{H}^2(\Omega)\cap \mathrm{H}^1_0(\Omega)$. A convergence rate of order $\mathcal{O}\left(\sqrt{\tau}\right)$, respectively $\mathcal{O}\left(\tau\right)$, is called suboptimal, respectively optimal. If $u_0\in \mathrm{H}^1(\Omega)$, then the error between the semidiscrete solution obtained by Rothe's method and the solution of the original problem is of order $\mathcal{O}\left(\sqrt{\tau}\right)$ in the norm $\mathrm{C}\left([0,T],\mathrm{L}^2(\Omega)\right)$ for the piecewise linear interpolant and in the norm $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$ for the piecewise constant interpolant (see (2.46)). The convergence rate is for both interpolants of order $\mathcal{O}\left(\sqrt{\tau}\right)$ in the space $\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$. If $u_0\in \mathrm{H}^2(\Omega)\cap \mathrm{H}^1_0(\Omega)$, then the convergence rate is optimal in the different spaces. Using this approach, no error estimates can be obtained when $u_0\in \mathrm{L}^2(\Omega)$.

The same convergence rates can be obtained when choosing as test function $\varphi = u(t) - \overline{u}_n(t)$. Then, it holds that

$$\tfrac{1}{2}\left\|(u-u_n)(\eta)\right\|^2 + \int_0^\eta \left\|\nabla(u-\overline{u}_n)\right\|^2 = \int_0^\eta \left(\partial_t(u-u_n), \overline{u}_n - u_n\right).$$

Because of $\partial_t u, \partial_t u_n \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ and (2.46), the RHS can be estimated as

$$\left|\int_0^\eta \left(\partial_t(u-u_n), \overline{u}_n - u_n\right)\right| \leqslant \sqrt{\int_0^\eta \left\|\partial_t(u-u_n)\right\|^2}\sqrt{\int_0^\eta \left\|\overline{u}_n - u_n\right\|^2} \lesssim \tau,$$

if $u_0 \in \mathrm{H}^1(\Omega)$. The convergence rate is suboptimal. Now, the error estimate is derived when $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$. From equation (2.57), it follows for a.a. $t \in (0,T)$ that

$$\begin{aligned}
\left\|\partial_t(u-u_n)(t)\right\|_{\mathrm{H}^1_0(\Omega)^*} &= \sup_{\substack{\varphi \in \mathrm{H}^1_0(\Omega) \\ \|\varphi\|_{\mathrm{H}^1_0(\Omega)} \leqslant 1}} \left(\partial_t(u-u_n)(t), \varphi\right) \\
&= \sup_{\substack{\varphi \in \mathrm{H}^1_0(\Omega) \\ \|\varphi\|_{\mathrm{H}^1_0(\Omega)} \leqslant 1}} \left(\nabla(\overline{u}_n - u)(t), \nabla\varphi\right) \\
&\leqslant \left\|\nabla(\overline{u}_n - u)(t)\right\|.
\end{aligned}$$

Therefore, the RHS can be estimated as follows

$$\begin{aligned}
\left|\int_0^\eta \left(\partial_t(u-u_n), \overline{u}_n - u_n\right)\right| &\leqslant \int_0^\eta \left\|\partial_t(u-u_n)\right\|_{\mathrm{H}^1_0(\Omega)^*} \left\|\overline{u}_n - u_n\right\|_{\mathrm{H}^1_0(\Omega)} \\
&\leqslant \int_0^\eta \left\|\nabla(\overline{u}_n - u)\right\| \left\|\overline{u}_n - u_n\right\|_{\mathrm{H}^1_0(\Omega)} \\
&\leqslant \varepsilon \int_0^\eta \left\|\nabla(u-\overline{u}_n)\right\|^2 + C_\varepsilon \int_0^\eta \left\|\overline{u}_n - u_n\right\|^2_{\mathrm{H}^1_0(\Omega)}.
\end{aligned}$$

Fixing ε sufficiently small and using bounds (2.48) and (2.49), it is clear that

$$\max_{t\in[0,T]} \left\|u(t) - u_n(t)\right\|^2 + \int_0^T \left\|\nabla(u-\overline{u}_n)\right\|^2 \lesssim \tau^2$$

if $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$.

An alternative approach is to integrate (2.57) first over the time variable $t \in (0,\xi) \subset (0,T)$. Then put $\varphi = u(\xi) - \overline{u}_n(\xi)$ and integrate again in time over $(0,\eta) \subset (0,T)$. It holds after some rearrangements in the terms that

$$\begin{aligned}
\int_0^\eta \left\|u-u_n\right\|^2 + \int_0^\eta \left(\int_0^\xi (\nabla u(t) - \nabla\overline{u}_n(t))\,\mathrm{d}t, \nabla u(\xi) - \nabla\overline{u}_n(\xi)\right)\,\mathrm{d}\xi& \\
= \int_0^\eta \left(u(\xi) - u_n(\xi), \overline{u}_n(\xi) - u_n(\xi)\right)\,\mathrm{d}\xi&.
\end{aligned}$$

and thus

$$\int_0^\eta \|u-u_n\|^2 + \tfrac{1}{2}\left\|\int_0^\eta (\nabla u(t) - \nabla \overline{u}_n(t))\,\mathrm{d}t\right\|^2 \leqslant \varepsilon \int_0^\eta \|u-u_n\|^2 + C_\varepsilon \int_0^\eta \|\overline{u}_n - u_n\|^2 .$$

Using (2.45), this results in the error estimate

$$\int_0^T \|u-u_n\|^2 \lesssim \tau$$

if $u_0 \in \mathrm{L}^2(\Omega)$ and using (2.46) this yields

$$\int_0^T \|u-u_n\|^2 \lesssim \tau^2$$

if $u_0 \in \mathrm{H}^1(\Omega)$. This implies that the convergence to the solution of the original problem in the space $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ is of order $\mathcal{O}(\tau)$ if $u_0 \in \mathrm{H}^1(\Omega)$ for the piecewise linear interpolant and the piecewise constant interpolant. This is better than the previously obtained result ($\mathcal{O}(\sqrt{\tau})$) if $u_0 \in \mathrm{H}^1(\Omega)$. Moreover, the convergence in the space $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ is of order $\mathcal{O}(\sqrt{\tau})$ if $u_0 \in \mathrm{L}^2(\Omega)$ for both interpolants.

Theorem 2.12.2. *Let $f \in \mathrm{L}^2(\Omega)$.*

- *Assume that $u_0 \in \mathrm{L}^2(\Omega)$. Then there exists a positive constant C such that*

$$\int_0^T \|u-u_n\|^2 \leqslant C\tau.$$

- *Assume that $u_0 \in \mathrm{H}^1(\Omega)$. Then there exists a positive constant C such that*

$$\max_{t\in[0,T]} \|u(t)-u_n(t)\|^2 + \sqrt{\int_0^T \|u-u_n\|^2} + \int_0^T \|\nabla(u-u_n)\|^2 \leqslant C\tau.$$

- *Assume that $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$. Then there exists a positive constant C such that*

$$\max_{t\in[0,T]} \|u(t)-u_n(t)\|^2 + \int_0^T \|\nabla(u-u_n)\|^2 \leqslant C\tau^2.$$

Remark 2.12.4. *Instead of using the backward Euler method, the problem can be approximated by using for instance the trapezoidal rule (Crank-Nicolson). This implicit method is a second-order method in time. Then, the sequence of time discrete problems is given by*

$$\left\{\begin{array}{rll} \delta u_i - \frac{1}{2}\left(\Delta u_i + \Delta u_{i-1}\right) & = f & \text{in } \Omega, \\ u_i & = 0 & \text{on } \partial\Omega. \end{array}\right.$$

In this example, the same estimates can be obtained as in the case of using the backward Euler method if $u_0 \in \mathrm{H}^1(\Omega)$ (for more details, the reader is referred to Appendix A.2). This implies that the convergence results stay valid (but $u_0 \in \mathrm{H}^1(\Omega)$ is needed instead of $u_0 \in \mathrm{L}^2(\Omega)$) and that there exists a unique weak solution to the problem under consideration by using the Crank-Nicolson scheme. Note that from (2.35) *and* (2.38)*, it follows that*

$$\int_0^T \|\overline{u}_n(t) - \overline{u}_n(t-\tau)\|^2_{\mathrm{H}^1(\Omega)} \,\mathrm{d}t \leqslant \tau \sum_{i=1}^{n} \|u_i - u_{i-1}\|^2_{\mathrm{H}^1(\Omega)} \leqslant C\tau$$

if $u_0 \in \mathrm{H}^1(\Omega)$. The error estimate is independent of the choice of the discretization method in time. To obtain better estimates it is crucial to increase the regularity on the data in the spatial domain (in the case of linear problems) [60].

2.12.6 Rothe's method for a more general setting

For the limit passage in Rothe's method a compactness argument and certain convergence principles are needed. This section presents such arguments to make the analysis easier.

In general, solving an evolution equation requires often the use of two spaces V and H. Here, the space H is obtained in connection with the time derivative and V results form the elliptic term $-\Delta u$ in (2.23) and the boundary condition $u = 0$ on Σ_T. This leads in fact to the concept of evolution triple that was defined in Definition 2.9.8.

The following lemmas are crucial for proving the convergence of Rothe's method. The first two lemmas in this section can be found in Kačur [36, Lemma 1.3.10 and 1.3.13]. Lemma 2.12.2 is a modification of the Arzelà-Ascoli theorem and can also be found in [61]. Lemma 2.12.3 is a generalized version of lemmas from Aubin [62] and Lions [63].

Definition 2.12.2. *The space $\mathrm{C}_w((0,T),X)$ consists of all $u : (0,T) \to X$ such that $\langle f, u(\cdot)\rangle$ is continuous as a real function on $(0,T)$ for any $f \in X^*$.*

Lemma 2.12.2.

(i) Let X be a reflexive Banach space and let $u_n : [0,T] \to X (n \in \mathbb{N})$ be equibounded and uniform equicontinuous. Then there exists a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ such that $u_{n_k}(t) \rightharpoonup u(t)$ in X for all $t \in [0,T]$, with $u \in \mathrm{C}_w((0,T),X) \cap \mathrm{L}^\infty((0,T),X)$.

(ii) Let X be a reflexive Banach space, Y a Banach space and let the embedding $X \subset Y$ be compact, i.e $X \hookrightarrow\hookrightarrow Y$. If $u_n : [0,T] \to X(n \in \mathbb{N})$ is equibounded and $u_n : [0,T] \to Y(n \in \mathbb{N})$ is uniform equicontinuous, then there exists a function $u \in \mathrm{C}([0,T],Y) \cap \mathrm{L}^\infty((0,T),X)$ and a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ such that $u_{n_k} \to u$ in $\mathrm{C}([0,T],Y)$ and $u_{n_k}(t) \rightharpoonup u(t)$ in X for a.a. $t \in [0,T]$.

Lemma 2.12.3. *Let V and Y be reflexive Banach spaces and let the embedding $V \subset Y$ be compact, i.e. $V \hookrightarrow\hookrightarrow Y$. If the estimates ($u_n$ and $\overline{u}_n$ are Rothe functions)*

$$\int_0^T \|\partial_t u_n(s)\|_Y^2 \ \mathrm{d}s \leqslant C, \qquad \|\overline{u}_n(t)\|_V \leqslant C \quad \text{for all } t \in [0,T]$$

hold for all $n \geqslant n_0 > 0$, then there exists a function $u \in \mathrm{C}([0,T],Y) \cap \mathrm{L}^\infty((0,T),V)$ with $\partial_t u \in \mathrm{L}^2((0,T),Y)$ (i.e. u is differentiable a.e. in $[0,T]$) and a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$\begin{cases} u_{n_k} \to u, & \text{in} \quad \mathrm{C}([0,T],Y), \\ u_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad V \text{ for all } t \in [0,T], \\ \overline{u}_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad V \text{ for all } t \in [0,T], \\ \partial_t u_{n_k} \rightharpoonup \partial_t u, & \text{in} \quad \mathrm{L}^2((0,T),Y). \end{cases}$$

Moreover, if

$$\|\partial_t u_n(t)\|_Y \leqslant C \qquad \text{for all } n \geqslant n_0 \text{ and a.a. } t \in [0,T],$$

then $\partial_t u \in \mathrm{L}^\infty((0,T),Y)$ and $u \in \mathrm{Lip}([0,T],Y)$.

Also the following lemma is a generalized Aubin and Lions lemma and can be found in [38, Lemma 7.7]. In fact, the lemma given here is an extension of [38, Lemma 7.7]. In the proof of [38, Lemma 7.7], the weak convergence of $\partial_t u_{n_k}$ is not studied in detail. Hence, the lemma is accompanied by a proof, see Lemma A.1.3 in Appendix A. The space $\mathrm{W}^{1,2,2}$ was defined in Definition 2.9.9.

Lemma 2.12.4. *Let V, Y and W be Banach spaces, with V separable and reflexive,*

$$V \hookrightarrow\hookrightarrow Y \qquad \text{and} \qquad Y \hookrightarrow W.$$

Then $\mathrm{W}^{1,2,2}([0,T];V,W) \hookrightarrow\hookrightarrow \mathrm{L}^2((0,T),Y)$. For every bounded sequence $\{u_n\}_{n\in\mathbb{N}}$ in $\mathrm{W}^{1,2,2}([0,T];V,W)$ there exist a function $u \in \mathrm{L}^2((0,T),Y)$ and a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$\begin{cases} u_{n_k} \to u & \text{in} \quad \mathrm{L}^2((0,T),Y), \\ u_{n_k} \rightharpoonup u & \text{in} \quad \mathrm{L}^2((0,T),V). \end{cases}$$

Moreover,

$$u \in \mathrm{C}([0,T],Y) \quad \text{and} \quad \partial_t u_{n_k} \rightharpoonup \partial_t u \quad \text{in} \quad \mathrm{L}^2((0,T),W)$$

if also the following evolution triple is satisfied

$$V \subset V_1 \hookrightarrow Y \cong Y^* \hookrightarrow W = V_1^* \subset V^*,$$

i.e. W has a predual, with

- V_1 *a reflexive and separable Banach space;*
- Y *Hilbert;*
- V_1 *dense in* Y.

Lemma 2.12.4 is more general than Lemma 2.12.3 because it can also be used outside Rothe's method.

2.12.6.1 Example

Consider again problem (2.23). If $u_0 \in \mathrm{H}^1(\Omega)$, then by a priori estimate (2.41) and Lemma 2.12.3 with $V = \mathrm{H}^1(\Omega)$ and $Y = \mathrm{L}^2(\Omega)$, there exists a

$$u \in \mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right) \cap \mathrm{L}^\infty\left((0,T), \mathrm{H}^1(\Omega)\right)$$

with $\partial_t u \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ and a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that

$$\begin{cases} u_{n_k} \to u, & \text{in} \quad \mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right), \\ u_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{H}^1_0(\Omega) \text{ for all } t \in [0,T], \\ \overline{u}_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{H}^1_0(\Omega) \text{ for all } t \in [0,T], \\ \partial_t u_{n_k} \rightharpoonup \partial_t u, & \text{in} \quad \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right). \end{cases}$$

Moreover, if $u_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, then by the a priori estimate (2.42) it holds that $\partial_t u \in \mathrm{L}^\infty\left((0,T), \mathrm{L}^2(\Omega)\right)$ and $u : [0,T] \to \mathrm{L}^2(\Omega)$ is Lipschitz continuous.

If $u_0 \in \mathrm{L}^2(\Omega)$, then by a priori estimate (2.40) and Lemma 2.12.3 with $V = \mathrm{L}^2(\Omega)$ and $Y = \mathrm{H}^1_0(\Omega)^*$, there exist a $u \in \mathrm{C}\left([0,T], \mathrm{H}^1_0(\Omega)^*\right) \cap \mathrm{L}^\infty\left((0,T), \mathrm{L}^2(\Omega)\right)$ with $\partial_t u \in \mathrm{L}^2\left((0,T), \mathrm{H}^1_0(\Omega)^*\right)$ and a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that

$$\begin{cases} u_{n_k} \to u, & \text{in} \quad \mathrm{C}\left([0,T], \mathrm{H}^1_0(\Omega)^*\right), \\ u_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{L}^2(\Omega) \text{ for all } t \in [0,T], \\ \overline{u}_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{L}^2(\Omega) \text{ for all } t \in [0,T], \\ \partial_t u_{n_k} \rightharpoonup \partial_t u, & \text{in} \quad \mathrm{L}^2\left((0,T), \mathrm{H}^1_0(\Omega)^*\right). \end{cases}$$

Note that in this case the result of Lemma 2.12.3 is not sufficient to prove the existence of a solution. For this, also

$$\overline{u}_{n_k}(t) \rightharpoonup u(t) \quad \text{in } \mathrm{H}^1_0(\Omega) \text{ for all } t \in [0,T],$$

is needed, which is valid due to $\displaystyle\int_0^T \|\overline{u}_n(s)\|^2_{\mathrm{H}^1(\Omega)}\, \mathrm{d}s \leqslant C$.

2.13 Finite element method

After the theoretical analysis of PDEs, the focus is now on the finite element method, which is a numerical technique for finding approximate solutions to elliptic boundary value problems. Finite element methods (FEM) are based on a variational formulation of the partial differential equation to be solved. First, Galerkin's method is introduced. The finite element method is an important special case of this method and is discussed afterwards. The reader can find more information about the FEM for instance in [10–12, 39, 54].

2.13.1 Galerkin method

Let V be a Hilbert space, $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ a bilinear form (coming, e.g., from the weak formulation of a PDE) and $f \in V^*$ (representing, e.g., the right-hand side of a PDE). The goal is to find $u \in V$ such that

$$a(u,v) = f(v), \qquad \forall v \in V. \tag{2.62}$$

Suppose that the properties of the Lax-Milgram lemma 2.11.1 are satisfied such that the weak problem (2.62) has a unique solution. Problem (2.62) is stated in an infinite dimensional space V. Therefore, in general, its exact solution is impossible to find (as a function of infinitely many unknown parameters). The finite dimensional (numerical) approximation of such problems was first studied by Galerkin in 1915.

The Galerkin method is based on the construction of a sequence of finite dimensional subspaces $\{V_h\}_{h=1}^{\infty} \subset V$ that fill the space V in the limit. The dimension of V_h is denoted by $\dim(V_h) = N_h$, $h \in \mathbb{N} \cup \{\infty\}$. It is possible to see h as $\frac{1}{N_h}$ such that $h \in (0,1]$ without loss of generality. The parameter h describes the quality of the approximation. In each finite-dimensional space V_h, problem (2.62) is solved exactly. Every finite dimensional subspace of a Hilbert space is closed and therefore a Hilbert space. The Galerkin approximate problem is usually called the discrete problem.

Definition 2.13.1 (Discrete problem)**.** *Find $u_h \in V_h$ such that*

$$a(u_h,v) = f(v), \qquad \forall v \in V_h. \tag{2.63}$$

All properties of the bilinear form a that are valid in V are automatically valid in V_h. Therefore, the existence and uniqueness of a solution to the discrete problem directly follow from the Lax-Milgram lemma. The solution $u_h \in V_h$, to the discrete problem (2.63) can be found explicitly thanks to the fact that the space V_h has a finite basis $\{\varphi_i\}_{i=1}^{N_h}$. The solution u_h can be written as a linear combination of these basis functions with unknown coefficients

$$u_h = \sum_{i=1}^{N_h} c_i \varphi_i.$$

Let us evaluate this combination in the discrete problem (2.63), i.e.

$$a\left(\sum_{i=1}^{N_h} c_i \varphi_i, v\right) = \sum_{i=1}^{N_h} c_i a(\varphi_i, v) = f\left(v\right), \qquad \forall v \in V_h. \tag{2.64}$$

Substituting the basis functions $\varphi_j, j = 1, \ldots, N_h$ in (2.64) yields

$$\sum_{i=1}^{N_h} c_i a(\varphi_i, \varphi_j) = f(\varphi_j), \qquad j = 1, \ldots, N_h. \tag{2.65}$$

The relations (2.64) and (2.65) are equivalent. Now, from equation (2.65), it follows that $\boldsymbol{c} = (c_1, \ldots, c_{N_h})^T$ is the solution of the algebraic system

$$M\boldsymbol{c} = \boldsymbol{f}, \tag{2.66}$$

with $\boldsymbol{f} = (f(\varphi_1), \ldots, f(\varphi_{N_h}))^T$ and

$$M = \begin{pmatrix} a(\varphi_1, \varphi_1) & a(\varphi_2, \varphi_1) & \ldots & a(\varphi_{N_h}, \varphi_1) \\ a(\varphi_1, \varphi_2) & a(\varphi_2, \varphi_2) & \ldots & a(\varphi_{N_h}, \varphi_2) \\ \vdots & \vdots & \vdots & \vdots \\ a(\varphi_1, \varphi_{N_h}) & a(\varphi_2, \varphi_{N_h}) & \ldots & a(\varphi_{N_h}, \varphi_{N_h}) \end{pmatrix}.$$

This matrix M is positive definite thanks to the V-ellipticity of a. Therefore, the matrix M is regular and there exists a unique solution of (2.66). The discrete problem (2.63) can thus be reduced to an algebraic system (2.66) that can be solved by appropriate algebraic solver. In this way, a uniquely determined u_h as the solution of (2.63) is determined. The error $e_h = u - u_h$ of the solution to the discrete problem (2.63) has the following orthogonality property.

Lemma 2.13.1 (Orthogonality of the error for elliptic problems)**.** *Let $u \in V$ be the exact solution of the continuous problem* (2.62) *and u_h the exact solution of the discrete problem* (2.63). *Then the error $e_h = u - u_h$ satisfies*

$$a(u - u_h, v) = 0, \qquad \textit{for all } v \in V_h.$$

The interpretation is as follows. When the bilinear form a is symmetric, then it induces an inner product and a norm $\|v\|_e = \sqrt{a(v,v)}$ for all $v \in V$. Therefore, the previous lemma implies that the error of the Galerkin approximation $e_h := u - u_h$ is orthogonal to the Galerkin subspace V_h. Hence, the approximate solution $u_h \in V_h$ is an orthogonal projection of the exact solution $u \in V$ onto the Galerkin subspace V_h, i.e.

$$\|u - u_h\|_e = \inf_{v \in V_h} \|u - v\|_e .$$

Cea's lemma establishes the relation between the error of the approximation e_h and the subspace V_h [64].

Lemma 2.13.2 (Cea's lemma)**.** *Let V be a Hilbert space, $a : V \times V \to \mathbb{R}$ a bilinear bounded V-elliptic form and $f \in V^*$. Let $u \in V$ be the solution of problem* (2.62). *Furthermore, let V_h be a subspace of V and $u_h \in V_h$ the solution of the Galerkin approximation* (2.63). *Let C_M and C_m be the continuity and V-ellipticity constants of the form a respectively. Then,*

$$\|u - u_h\|_V \leqslant \frac{C_M}{C_m} \inf_{v \in V_h} \|u - v\|_V .$$

Cea's lemma states that the approximation error e_h, depends on the choice of the Galerkin subspace V_h, but it does not depend on the choice of its basis. Therefore, it is reasonable to consider such a V_h that approximates V with sufficient accuracy.

Remark 2.13.1. *In the symmetric case, it is true that*

$$\|u - u_h\|_e = \inf_{v \in V_h} \|u - v\|_e . \tag{2.67}$$

Using

$$\sqrt{C_m}\, \|v\|_V \leqslant \sqrt{a(v,v)} = \|v\|_e \leqslant \sqrt{C_M}\, \|v\|_V , \qquad v \in V,$$

it holds that

$$\begin{aligned} \|u - u_h\|_V &\leqslant \frac{1}{\sqrt{C_m}} \|u - u_h\|_e \\ &= \frac{1}{\sqrt{C_m}} \inf_{v \in V_h} \|u - v\|_e \\ &\leqslant \sqrt{\frac{C_M}{C_m}} \inf_{v \in V_h} \|u - v\|_V \\ &\leqslant \frac{C_M}{C_m} \inf_{v \in V_h} \|u - v\|_V , \end{aligned}$$

because $C_M \geqslant C_m$. Thus, in the symmetric case, Cea's relation can be derived from (2.67).

The convergence of the sequence of approximate solutions $\{u_h\}$, $u_h \in V_h$, to the exact solution of problem (2.62) follows from Cea's lemma.

Theorem 2.13.1. *Let V be a Hilbert space and V_h a sequence of finite dimensional subspaces $V_h \subset V$ for which*

$$\inf_{v \in V_h} \|u - v\|_V \to 0 \qquad \textit{as } h \to 0, \forall u \in V,$$

where $u \in V$ is the solution of problem (2.62). *Let $u_h \in V_h$ be the solution of the Galerkin approximation* (2.63). *Let $a : V \times V \to \mathbb{R}$ be a bilinear bounded V-elliptic form and $f \in V^*$. Then*

$$\lim_{h \to 0} \|u - u_h\|_V = 0,$$

i.e. the Galerkin method for problem (2.62) *converges.*

2.13.2 Finite element method

Let $\Omega \subset \mathbb{R}^d$ be an open bounded set. If the Hilbert space V consists of functions defined in Ω and the Galerkin subspaces $V_h \subset V$ comprise piecewise-polynomial functions, the Galerkin method is called the finite element method. In what follows the main ideas of the finite element method are explained in 1D, 2D and 3D. For a nice introduction on the finite element method in 3D when Ω is a polyhedral Lipschitz continuous domain, the reader is referred to [39, Chapter 5].

2.13.3 Finite element method in 1D

The main idea of the FEM is to split the whole domain $\Omega = (a, b)$ into a finite system of N disjoint open parts $\Omega_i = (x_{i-1}, x_i)$ of length $h_i = x_i - x_{i-1}$, $i = 1, \ldots, N$, i.e.

$$a = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = b.$$

The largest length (mesh diameter) of all subintervals is denoted by

$$h = \max_{1 \leqslant i \leqslant N} h_i.$$

Let $P_k(\Omega_i)$ be the set of all polynomials of a degree $\leqslant k \in \mathbb{N}$ that are defined on Ω_i. Introduce the so-called Lagrange FEM space as

$$V_h^k := \left\{ \varphi \in \mathrm{C}(\overline{\Omega}) : \varphi_{|\Omega_i} \in P_k(\Omega_i), i = 1, \ldots, N \right\}.$$

For each segment there are $k + 1$ nodes. In general case, a polynomial $p_k(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_k x^k$ of degree k can be determined from its values in $(k + 1)$ nodes $z_j, j = 0, \ldots k$ in $\overline{\Omega}_i$. The basis functions for the determination of $p_k(x)$ are $\varphi_j \in P_k(\overline{\Omega}_i)$ that are associated with a node z_j in such a way that $\varphi_j(z_l) = \delta_{lj}$, $0 \leqslant j, l \leqslant k$. These are the Lagrange interpolation operators given by

$$\varphi_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^{k} \frac{x - z_i}{z_j - z_i}, \quad \forall x \in \overline{\Omega}_i.$$

Example 2.13.2 (Piecewise linear polynomials)**.** *The hat functions $\varphi_i \in V_h^1$, $i = 0, \ldots, N$, given by*

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_i} & x \in \overline{\Omega}_i, \\ 1 - \frac{x - x_i}{h_{i+1}} & x \in \overline{\Omega}_{i+1}, \\ 0 & x \notin \overline{\Omega}_i \cup \overline{\Omega}_{i+1}, \end{cases}$$

create a basis for $V_h^1 \subset \mathrm{H}^1(a, b) \subset \mathrm{C}([a, b])$. For each $i = 0, \ldots, N$, the requirement $\varphi_i(x_j) = \delta_{ij}$ is satisfied, $1 \leqslant i, j \leqslant N$. Define the operator of linear interpolation

$$\Pi_h : \mathrm{C}([a, b]) \to V_h^1$$

as follows

$$\Pi_h v := \sum_{j=0}^{N} v(x_j) \varphi_j,$$

i.e. $(\Pi_h v)(x_i) = v(x_i)$, $\forall i = 1, \ldots, N$. The function $\Pi_h v$ is continuous and piecewise linear. The error of linear interpolation is given by [39, Theorem 5.48]

$$\|u - \Pi_h u\|_{\mathrm{H}^1(a,b)} \leqslant Ch \|u''\|,$$

if $u \in \mathrm{H}^2(a,b)$.

Now, let u *be a variational solution of* (2.62) *with test space* $V \subset \mathrm{H}^1(a,b)$*. It is assumed that the bilinear form* a *in* (2.62) *is* H^1*-elliptic and bounded. The right-hand side is supposed to be a linear bounded functional in* $\mathrm{H}^1(a,b)$*. Choose* $V_h^1 \subset V$ *as the finite dimensional space (i.e. the finite element method is conforming) in which we seek an approximate solution* u_h *of* u*. If* $u \in \mathrm{H}^2(a,b)$*, then by Cea's lemma 2.13.2, there exists a positive constant* C *such that*

$$\|u - u_h\|_{\mathrm{H}^1(a,b)} \leqslant C \|u - \Pi_h u\|_{\mathrm{H}^1(a,b)} \leqslant Ch \|u''\| \,.$$

Therefore, under the regularity assumption $u \in \mathrm{H}^2(a,b)$*, a convergence rate* $\mathcal{O}(h)$ *in the space* $\mathrm{H}^1(a,b)$ *is obtained. However, when the solution is less regular, namely* $u \in \mathrm{H}^1(a,b)$*, only the convergence of Galerkin's approximations to the solution can be proved by the density of* $\mathrm{H}^2(a,b)$ *in* $\mathrm{H}^1(a,b)$*, i.e.*

$$\|u - u_h\|_{\mathrm{H}^1(a,b)} \to 0 \qquad \text{as } h \to 0.$$

Note that the space $V_h^k \subset \mathrm{H}^1(a,b) \subset \mathrm{C}([a,b])$ for all $k \in \mathbb{N}$.

2.13.4 Finite element method in 2D

Let Ω be a bounded domain in $\mathbb{R}^2$ with polygonal boundary Γ.

Definition 2.13.2. *A finite system of open triangles* $\mathcal{T}$ *in* Ω *is said to be a triangulation* $\mathcal{T}_h$ *if*

(i) $\mathcal{T} \subset \Omega$ *for all* $\mathcal{T} \in \mathcal{T}_h$*,*

(ii) $\bigcup_{\mathcal{T} \in \mathcal{T}_h} \overline{\mathcal{T}} = \overline{\Omega}$*,*

(iii) only one of the following possibilities holds for any couple of different triangles

$$\overline{\mathcal{T}^i} \cap \overline{\mathcal{T}^j} = \begin{cases} \varnothing \\ \textit{common node} \\ \textit{common edge of two triangles.} \end{cases}$$

The parameter h *is given by*

$$h = \max_{\mathcal{T} \in \mathcal{T}_h} \textit{diam}(\overline{\mathcal{T}}),$$

where diam is defined as $\sup_{\mathbf{x},\mathbf{y} \in \mathcal{T}} |\mathbf{x} - \mathbf{y}|_{\mathrm{e}}$*, i.e.* h *is the longest edge of the set of all edges.*

In fact, a system of triangulations $\mathcal{T}_h$, $h \in (0,1]$, is chosen to be regular.

Definition 2.13.3. *A system of triangulations $\{\mathcal{T}_h\}_{h\in(0,1]}$ is said to be regular if there exists a positive constant C such that*

$$\frac{h_{\mathcal{T}}}{\rho_{\mathcal{T}}} \leqslant C, \quad \forall \mathcal{T} \in \mathcal{T}_h \text{ and } \forall h \in (0,1],$$

where $h_{\mathcal{T}} = diam(\overline{\mathcal{T}})$ and $\rho_{\mathcal{T}}$ is the diameter of the incircle of $\overline{\mathcal{T}}$.

The regularity of a system of triangulations can be also expressed using the minimal angle condition, i.e. there exists a constant θ_0 such that

$$\theta_{\mathcal{T}} \geqslant \theta_0 > 0 \qquad \forall \mathcal{T} \in \mathcal{T}_h, \forall h \in (0,1],$$

where $\theta_{\mathcal{T}}$ is the minimal angle of the triangle $\mathcal{T}$.

Example 2.13.3. *Define the following Lagrange FEM space*

$$X_h^k := \left\{\varphi \in \mathrm{C}(\overline{\Omega}) : \varphi_{|\mathcal{T}} \in P_k(\mathcal{T}), \forall \mathcal{T} \in \mathcal{T}_h\right\},$$

with $P_k(\mathcal{T})$ the set of all polynomials of a degree $\leqslant k$, which are defined on $\mathcal{T}$. It holds that $X_h^k \subset \mathrm{H}^1(\Omega)$ for all $k \in \mathbb{N}$ [65, p. 1.125].

2.13.5 Finite element method in 3D

The first step is to generate a finite element mesh that covers the domain Ω, which is a bounded polyhedral Lipschitz continuous domain in $\mathbb{R}^3$. The domain Ω can be subdivided into a finite set of distinct tetrahedra $\mathcal{T}_h = \{\mathcal{T}\}$ such that $\overline{\Omega} = \bigcup\limits_{\mathcal{T}\in\mathcal{T}_h} \overline{\mathcal{T}}$, see [66]. Remark that $h = \max\limits_{\mathcal{T}\in\mathcal{T}_h} h_{\mathcal{T}}$, where $h_{\mathcal{T}}$ is the diameter of the smallest sphere containing $\mathcal{T}$. If $\mathcal{T}^1 \in \mathcal{T}_h$ and $\mathcal{T}^2 \in \mathcal{T}_h$ with $\overline{\mathcal{T}^1} \cap \overline{\mathcal{T}^2} \neq \varnothing$, then the elements meet in one of the following ways:

- the elements meet at a single point that is a vertex for both elements;
- the elements meet along a common edge and the endpoints of that edge are vertices of the two elements;
- the elements meet at a common face and the vertices of that face are vertices of both elements.

Usually, it is assumed that there is a regular family of meshes or triangulations $\{\mathcal{T}_h : h > 0\}$, where h denotes the mesh parameter.

Definition 2.13.4. *A family of triangulations $\{\mathcal{T}_h\}_{h\in(0,1]}$ of the domain $\Omega \subset \mathbb{R}^3$ is called regular if there exists a constant $C > 0$ such that*

$$\frac{h_{\mathcal{T}}}{\rho_{\mathcal{T}}} \leqslant C, \quad \forall \mathcal{T} \in \mathcal{T}_h \text{ and } \forall h \in (0,1],$$

where $\rho_{\mathcal{T}}$ is the supremum of the diameters of the spheres inscribed into $\mathcal{T}$.

Example 2.13.4 (First-order Lagrange finite elements). *In this example, the first-order Lagrange finite elements for the space discretization are considered. The finite element space $V_h \subset \mathrm{H}^1(\Omega)$ is given by $V_h = \{v_h \in \mathrm{H}^1(\Omega) \,:\, v_h|_{\mathcal{T}} \in P_1(\mathcal{T}), \quad \forall \mathcal{T} \in \mathcal{T}_h\}$, with $P_1(\mathcal{T})$ the space of componentwise first-order polynomials. The coefficients of these polynomials are determined by $v_h(\mathbf{a}_i)$ with $\mathbf{a}_i, i = 1, \ldots, 4$, the vertices of $\mathcal{T}$. The total number of vertices of $\mathcal{T}_h$ is set equal to M and the ith vertex of $\mathcal{T}_h$ is put equal to $\mathbf{x}_i$. The linear basis functions $\{\varphi_j\}_{j=1}^M$, such that*

$$\varphi_j(\mathbf{x}_i) = \delta_{ij} \qquad i, j = 1, \ldots M;$$

span the finite element space $V_h \subset \mathrm{H}^1(\Omega)$.

Example 2.13.5 (Lowest order curl-conforming Nédélec edge elements [67]). *The finite element space $\mathbf{V}^h$ is given by*

$$\mathbf{V}^h = \{\mathbf{v}^h \in \mathbf{H}(\mathbf{curl}; \Omega) \,:\, \mathbf{v}^h\big|_{\mathcal{T}}(\mathbf{x}) = \mathbf{a}_{\mathcal{T}} + \mathbf{b}_{\mathcal{T}} \times \mathbf{x}, \quad \forall \mathcal{T} \in \mathcal{T}_h\},$$

where $\mathbf{a}_{\mathcal{T}}$ and $\mathbf{b}_{\mathcal{T}}$ are constants in $\mathbb{R}^3$. The components of $\mathbf{a}_{\mathcal{T}}$ and $\mathbf{b}_{\mathcal{T}}$ are determined by the degrees of freedom $\int_e \mathbf{v}^h \cdot \hat{\boldsymbol{\tau}}$ on the six edges of a tetrahedron $\mathcal{T}$ with $\hat{\boldsymbol{\tau}}$ a unit vector along the edge e of $\mathcal{T}$. The space $\mathbf{V}^h$ is spanned by the basis functions

$$\boldsymbol{\varphi}_e = \varphi_i \nabla \varphi_j - \varphi_j \nabla \varphi_i,$$

where φ_i and φ_j denote the first-order Lagrange basis functions corresponding to nodes i resp. j of the mesh and e is the edge connecting i and j. For every edge e and e' it holds that $\int_{e'} \boldsymbol{\varphi}_e \cdot \hat{\boldsymbol{\tau}} = \delta_{ee'}$. The vector fields $\boldsymbol{\varphi}_e$ are called the basis functions dual to the degrees of freedom.

2.13.6 The finite element: general definition

The classical definition of a finite element is from Ciarlet, cf. [54].

Definition 2.13.5. *A finite element in $\mathbb{R}^d$ is a triple $(K, \mathcal{P}, \mathcal{N})$ where*

(i) K is a closed bounded set in $\mathbb{R}^d$ with nonempty interior and piecewise smooth boundary (the element domain),

(ii) $\mathcal{P}$ is a finite dimensional space of functions over the set K (the space of shape (basis) functions),

(iii) $\mathcal{N} = \{N_1, \ldots, N_k\}$ is a set of linear functionals on $\mathcal{P}$ (these linear functionals are called the degrees of freedom of the finite element). By definition, the set $\mathcal{N}$ is $\mathcal{P}$-unisolvent, i.e. $\mathcal{N}$ can be taken as a basis for the dual space $\mathcal{P}^*$.

Remark 2.13.2. *The set K is usually an interval in 1D, a triangle in 2D and a tetrahedron in 3D.*

Definition 2.13.6. *Let $(K, \mathcal{P}, \mathcal{N})$ be a finite element. The basis $\{\varphi_1, \varphi_2, \ldots, \varphi_k\}$ of $\mathcal{P}$ dual to $\mathcal{N}$ (i.e. $N_i(\varphi_j) = \delta_{ij}$) is called the nodal basis of $\mathcal{P}$.*

Example 2.13.6 (The 1-dimensional Lagrange element)**.** *Set $K = [0, 1]$ (reference element), $\mathcal{P}$ the set of linear polynomials and $\mathcal{N} = \{N_1, N_2\}$, where $N_1(v) = v(0)$ and $N_2(v) = v(1)$ for all $v \in \mathcal{P}$. Then $(K, \mathcal{P}, \mathcal{N})$ is a finite element and the nodal basis consists of $\varphi_1(x) = 1 - x$ and $\varphi_2(x) = x$. In general, set $K = [a, b]$ and $\mathcal{P}_k$ the set of all polynomials of degree $\leqslant k$. Let $\mathcal{N}_k = \{N_1, \ldots, N_k\}$ with $N_i(v) = v\,(a + (b - a)i/k)$ for all $v \in \mathcal{P}_k$ and $i = 0, 1, \ldots, k$. Then $(K, \mathcal{P}_k, \mathcal{N}_k)$ is a finite element. The verification of this uses Lemma 2.13.3. Other examples (in more dimensions) can be found in [12, Chapter 3].*

This section is finished with the following equivalence relation and the definitions of local and global interpolant [12, Lemma 3.1.4, Definition 3.3.1 and Definition 3.3.9].

Lemma 2.13.3. *Let $\mathcal{P}$ be a d-dimensional vector space and let $\{N_1, \ldots, N_d\}$ be a subset of the dual space $\mathcal{P}^*$. Then the following two statements are equivalent*

(a) $\{N_1, \ldots, N_d\}$ is a basis for $\mathcal{P}^*$;

(b) Given $v \in \mathcal{P}$ with $N_i v = 0$ for $i = 1, 2, \ldots, d$, then $v \equiv 0$.

Example 2.13.7. *Condition (iii) in Definition 2.13.5 is equivalent to (a) in Lemma 2.13.3, which can be verified by checking (b) in Lemma 2.13.3. For instance, in Example 2.13.6, $v \in \mathcal{P}_1$ means $v = a + bx$; $N_1(v) = N_2(v) = 0$ implies that $a = 0$ and $a + b = 0$. Hence, $v \equiv 0$. More generally, if $v \in \mathcal{P}_k$ and $N_i(v) = v\,(a + (b - a)i/k) = 0$ for all $i = 0, 1, \ldots, k$, then $v \equiv 0$ by the fundamental theorem of algebra. Thus $(K, \mathcal{P}_k, \mathcal{N}_k)$ is a finite element.*

Definition 2.13.7 (Local interpolant)**.** *Given a finite element $(K, \mathcal{P}, \mathcal{N})$, let the set $\{\varphi_i : 1 \leqslant i \leqslant k\} \subset \mathcal{P}$ be the basis dual to $\mathcal{N}$. Then the local interpolant is given by*

$$\mathcal{I}_K v := \sum_{i=1}^{k} N_i(v)\varphi_i$$

if v is a function for which all $N_i \in \mathcal{N}$, $i = 1, \ldots, k$, are defined.

Definition 2.13.8 (Global interpolant)**.** *Let Ω be a domain with a subdivision $\mathcal{T}$, i.e. there exists a finite collection of element domains $\{K_i\}$ such that*

- *int$(K_i) \cap$ int$(K_j) = \varnothing$ if $i \neq j$ and*
- *$\bigcup K_i = \overline{\Omega}$.*

Assume that each element domain K_i in the subdivision is equipped with some type of shape functions $\mathcal{P}$ and nodal variables $\mathcal{N}$ such that $(K_i, \mathcal{P}, \mathcal{N})$ forms a finite element. Let m be the order of the highest partial derivatives involved in the nodal variables. For $f \in \mathrm{C}^m(\overline{\Omega})$, the global interpolant $\mathcal{I}_\mathcal{T}$ is defined by

$$\mathcal{I}_\mathcal{T} f|_{K_i} = \mathcal{I}_{K_i} f$$

for all $K_i \in \mathcal{T}$.

Remark 2.13.3. *Without further assumptions on a subdivision, no continuity properties can be asserted for the global interpolant. For instance, in 2D, a triangulation of the domain is needed, cf. Definition 2.13.2. This is a subdivision consisting of triangles having the property that no vertex of any triangle lies in the interior of an edge of another triangle. Analogous definitions and results can be formulated in three dimensions. A nice example of a global interpolant in 2D can be found in [12, p. 80].*

2.13.7 The finite element method for nonlinear problems

The proposed Galerkin method is only valid for linear problems. In the case of a nonlinear elliptic problem, a possible approach is to linearize the problem using Picard iteration or Newton method [68].

2.13.8 Finite element libraries

There are several good software packages available for solving partial differential equations using the finite element method: Agros, ALBERTA, COMSOL, DUNE, FEniCS Project, freeFEM, GetDP, Hermes, kaskade, PLTMG, UG,....

The finite element library DOLFIN [69, 70] from the FEniCS Project [71] is used for the implementation of the results in this thesis. The FEniCS Project is a collection of free software with an extensive list of features for automated, efficient solution of differential equations. PDEs can be specified in near-mathematical notation (as finite element variational problems) and solved automatically. FEniCS can be programmed both in C++ and Python. More information can be found on

`http://fenicsproject.org/`.

Part I

Nonlocal problems for superconductivity

3 Superconductivity: overview and new models

In 1911, the Dutch physicist Kamerlingh Onnes discovered that the resistance of mercury not only decreases with temperature, but also has a sudden drop when cooled below 4.1K (-269°C). Afterwards, it was detected that also other metals (tin, lead,...) and intermetallic compounds (made of two or more metallic elements) lose all electric resistance below a certain critical temperature T_c. Kamerlingh Onnes called this a superconducting state in contrast to a normal state. Materials that exhibit such behaviour are called superconductors [72–74]. The fact that the resistance is zero in a superconducting state has been demonstrated by sustaining currents in superconducting lead rings for many years with no measurable reduction. This phenomenon is not the only characteristic of a superconductor. A brief introduction to the basic phenomena of superconductivity is given in Section 3.1.

There exist two main types of superconductors. The classification into type-I and type-II superconductors is made in Section 3.2. Finally, Section 3.3 concerns the modelling part of this chapter. Two new macroscopic models in terms of the magnetic field for nonlocal superconductors of type-I are derived in Subsection 3.3.1. These models are obtained from the Maxwell equations, the two-fluid model of London and London, and the nonlocal representation of the superconductive current by Eringen. Moreover, after giving an overview of the available macroscopic models for type-II superconductors in Subsection 3.3.2, a macroscopic model for an intermediate state between type-I and type-II superconductivity is proposed in Subsection 3.3.3.

The well-posedness of the different models is discussed into more details in the following chapters. The important question is whether an engineer or a physicist can use these models and how these models behave in comparison with other models (what are the advantages?). In Chapter 4, it becomes clear that a numerical implementation of these models with aid of the finite element method is very challenging.

This study builds on the book of Fabrizio and Morro [2, Chapter 11]. Further readings on superconductors can be found in [75–79] from which some of the passages are extracted.

3.1 Basic phenomena

Kamerlingh Onnes discovered in 1911 that for various cooled down materials the electrical resistance does not only decrease with temperature, but also suddenly drops at some critical absolute temperature T_c [72–74]. Materials that exhibit such

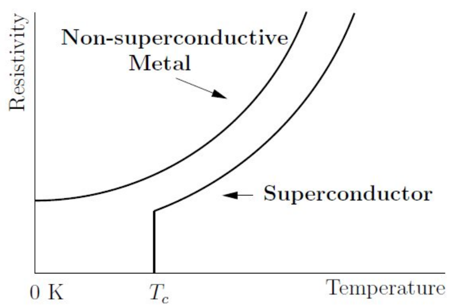


Figure 3.1: Resistivity of a typical material as a function of the temperature. For a non-superconducting metal (such as copper or gold) the resistivity approaches a finite value at zero temperature, while for a superconductor (such as lead or mercury) all signs of resistance suddenly disappear below a certain temperature, T_c.

behaviour are called superconducting materials or in short superconductors. Well-known examples of superconducting materials are mercury, lead and tin. Electric currents started in these materials persist for a long time. The resistivity of a material as a function of the temperature is depicted in Fig. 3.1 [80]. Thus, *perfect conductivity* (i.e. zero resistance) is an important feature of superconductivity.

Another fundamental property of superconductivity is *perfect diamagnetism* (i.e. expulsion of the magnetic flux from the superconducting material), which was discovered in 1933 by Meissner and Ochsenfeld [81]. This so-called Meissner-Ochsenfeld effect of superconductivity is often demonstrated by cooling a disk made of a superconducting material with liquid nitrogen below the critical temperature T_c. A magnet placed above the disk is repelled, i.e., it is levitated above the superconductor. This phenomenon is caused by the mentioned Meissner-Ochsenfeld effect that is related to the fact that a superconductor excludes magnetic fields from the interior of the superconducting material. All these features are, however, ceased to exist when the temperature surpasses T_c and when the magnitude of the magnetic field exceeds a temperature-dependent critical magnetic field magnitude $H_c(T)$. For practical applications, it is also important that the magnitude of the current density $\mathbf{J}$ stays below a critical value J_c. The following subsections illustrate the difference between perfect conductivity and perfect diamagnetism in more detail.

3.1.1 Zero resistivity

In the presence of conducting materials, electric fields induce an electric current. The conductivity σ is defined by Ohm's law

$$\mathbf{J} = \sigma \mathbf{E}. \tag{3.1}$$

Here, $\mathbf{J}$ is the electric current density associated with the external electric field $\mathbf{E}$. The resistivity ϱ is simply the reciprocal of the conductivity, i.e. $\varrho = \frac{1}{\sigma}$.

A superconductor is a perfect conductor. This means that the resistivity is zero and that the conductivity σ appears to become infinite below T_c. From Ohm's law (3.1), it follows that $\mathbf{E} \to \mathbf{0}$ if the current density $\mathbf{J}$ is bounded. This suggests that in a perfect conductor the electric field vanishes. There is current flow without electric field. Then, as a consequences of Faraday's law, the magnetic flux density doesn't depend on time at all points inside a superconductor, i.e.

$$\partial_t \mathbf{B} = \mathbf{0}.$$

Thus perfect conductivity implies that a change in the magnetic flux enclosed in the material is not possible.

This consequence of zero resistivity in a material is illustrated by the following experiment. A scheme of the experiment is presented in Fig. 3.2 [76]. Suppose that a sample of a material is initially held at temperature $T > T_c$ and is placed in a zero external magnetic field $\mathbf{H}_{\text{ext}} = \mathbf{0}$. Then, the temperature is cooled down below T_c. Thereafter, an external field is turned on with a magnitude less than the magnitude of the critical magnetic field $\mathbf{H}_c$. It is for instance an externally applied current in coils that produces the external field. The magnetic flux density inside the sample remains zero due to the fact that $\partial_t \mathbf{B} = \mathbf{0}$ inside a perfect conductor. Thus, by applying the external magnetic field to the sample after it is already in the superconducting state, the state with zero magnetic field everywhere in the sample is conserved (this by induced screening currents).

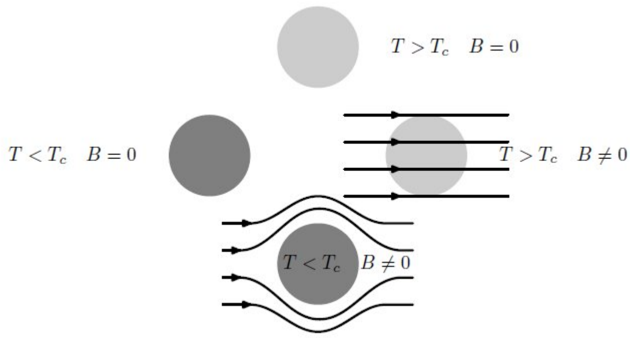


Figure 3.2: The Meissner-Ochsenfeld effect in superconductors. If a sample which is initially at high temperature and in zero magnetic field (top) is first cooled (left) and then placed in a magnetic field (bottom), then the magnetic field cannot enter the material (bottom). This is a consequence of zero resistivity. On the other hand, a normal sample (top) can be first placed in a magnetic field (right) and then cooled (bottom). In this case the magnetic field is expelled from the system.

3.1.2 The Meissner-Ochsenfeld Effect

Nowadays, the fundamental evidence that superconductivity occurs in a given material is the fact that a superconductor expels a weak external magnetic field. This fact was detected by Meissner and Ochsenfeld in 1933 [81]. The experiment of the previous subsection is reconsidered (see Fig. 3.2), but the different steps are switched in order. Suppose that the initial temperature of the sample is above T_c and first turn on the external field $\mathbf{H}_{\text{ext}}$. In this case, the magnetic field easily penetrates into the sample, $\mathbf{H} = \mathbf{H}_{\text{ext}}$. Then, the sample is cooled down below T_c. The observation is that the magnetic field is expelled from the interior of the material. This fact cannot be deduced from zero resistivity. In a material with zero resistance, the magnetic flux would remain unchanged ($\partial_t \mathbf{B} = \mathbf{0}$). The observation is a new and separate physical phenomenon associated with superconductors. Accordingly, infinite conductivity is a necessary but not a sufficient condition for the existence of the Meissner effect. The Meissner effect thus consists of expulsion of any magnetic field from the interior of a superconductor, whether it was there before the material became superconducting or not. When a sample is in a static external magnetic field, the condition $\mathbf{B} = \mathbf{0}$ inside the sample is maintained by superficial superconducting currents $\mathbf{J}_{\text{int}}$. These currents produce a magnetic field which is equal and opposite to the applied external field, leaving zero field in total. The screening currents $\mathbf{J}_{\text{int}}$ also produce a magnetization $\mathbf{M}$ in the sample defined by

$$\nabla \times \mathbf{M} = \mathbf{J}_{\text{int}}.$$

The three vectors $\mathbf{M}, \mathbf{H}_{\text{ext}}$ and $\mathbf{B}$ are related by

$$\mathbf{B} = \mu_0(\mathbf{H}_{\text{ext}} + \mathbf{M}). \tag{3.2}$$

Imposing the Meissner condition $\mathbf{B} = \mathbf{0}$ in (3.2) immediately leads to $\mathbf{M} = -\mathbf{H}_{\text{ext}}$. In what follows, the external magnetic field is denoted by $\mathbf{H}$ instead of $\mathbf{H}_{\text{ext}}$.

3.1.3 Threshold field

In 1914, Kamerlingh Onnes discovered that the resistance of a superconductor could be restored to its value in the normal state by the application of a large magnetic field and by increasing the temperature. The value $H_c = |\mathbf{H}_c|_{\text{e}}$ denotes the magnitude of the magnetic field $\mathbf{H}$ at which the jump in resistance occurs. It is termed the threshold value. This value depends on the temperature by the following parabolic law

$$H_c(T) = H_c(0)\left[1 - \left(\frac{T}{T_c}\right)^2\right],$$

where $H_c(0)$ is the threshold field at the absolute zero.

It was also found that the magnitude of the magnetization, denoted by $M = |\mathbf{M}|_{\text{e}}$, abruptly increases from the value $-H$ to zero when H reaches the threshold value. Gorter and Casimir [82] proved that the threshold field curve provides a phase diagram in the T-H plane. Any point below the curve $H_c(T)$, namely, the set of pairs $\{(T,H) : H < H_c(T)\}$, specifies states in which the material is locally in the superconductive phase; the points above define states of the normal phase, see Fig. 3.3(a).

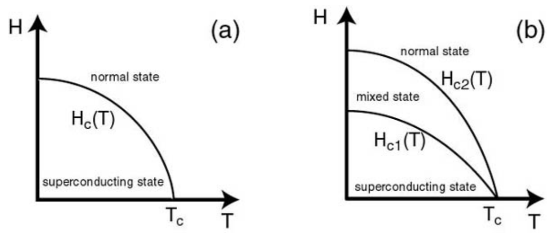


Figure 3.3: Critical magnetic field as a function of the temperature for (a) type I superconductors and (b) type II superconductors.

3.2 Type-I and Type-II superconductivity

There are two main types of superconductors: type-I and type-II superconductors. They behave similarly for a very weak external magnetic field when the temperature $T < T_c$ is fixed, but as the field becomes stronger it turns out that different outcomes can show up.

In the first case, the $\mathbf{B}$ field remains zero inside the superconductor until suddenly, as the critical field H_c is reached, the superconductivity is destroyed. Such materials are called type-I superconductors. The way the magnetization $\mathbf{M}$ changes with $\mathbf{H}$ in a type-I superconductor is shown in Fig. 3.4 [76]. The magnetization obeys the relation $\mathbf{M} = -\mathbf{H}$ in the superconductive state and becomes $\mathbf{0}$ in the normal state.

In the second case, for type-II superconductors, a mixed state occurs in addition to the superconductive and the normal state. In this mixed state, there is a partial magnetic penetration (on the macroscopic level). A phase diagram is depicted in Fig. 3.3(b). There are two different critical fields, the lower critical field H_{c1} and the upper critical field H_{c2}. For small values of the applied field $\mathbf{H}$, the Meissner-Ochsenfeld effect leads to $\mathbf{B} = \mathbf{0}$ and $\mathbf{M} = -\mathbf{H}$ inside the sample. Once the magnitude of the applied magnetic field exceeds H_{c1}, magnetic flux starts to enter the superconductor and hence $\mathbf{B} \neq \mathbf{0}$. Therefore, the magnitude of $\mathbf{M}$ is closer to zero than the magnitude of $-\mathbf{H}$. There is a large number of small normal regions (tubes) being produced in the superconducting material. Upon increasing the magnitude of the field $\mathbf{H}$ further, the magnetic flux density gradually increases, until finally at H_{c2} the superconductivity is destroyed and $\mathbf{M} = \mathbf{0}$.

A type-I superconductor is usually made of a pure metal (e.g. lead, mercury, niobium, tin), while type-II superconductors are usually alloys (a material composed of two or more metals or a metal and a nonmetal, e.g. niobium tin, titanium niobium). High-temperature superconductors are an important subclass of type-II superconductors. These materials have a superconducting transition temperature above 30K (-240°C) up to 130K (-140°C). For this sort of materials is the cooling more efficient and less expensive. More information about high-temperature superconductors can be found in [83]. Our study is mainly focused on type-I superconductors.

3.3 Nonlocal macroscopic models for superconductivity

Although a large number of studies have been devoted to the microscopic theory of superconductivity (the first microscopic theory was the BCS theory by Bardeen, Cooper and Schrieffer [84]), the macroscopic theory seems to have less attention in the literature. Nonetheless, since the discovery of high-temperature superconductors in 1986 [85], industrial applications require macroscopic models and their mathematical analysis for superconductivity.

In Subsection 3.3.1, the available macroscopic models for type-I superconductivity are overviewed. Afterwards, two new models for nonlocal superconductivity are derived and discussed. Next, in Subsection 3.3.2, the available macroscopic models for type-II superconductivity are given. Finally, in Subsection 3.3.3, based

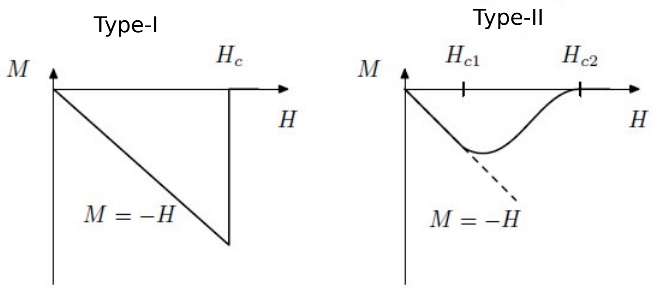


Figure 3.4: The magnetization M as a function of H in type-I and type-II superconductors in 1D. For type-I, perfect Meissner diamagnetism is continued until H_c, beyond which superconductivity is destroyed. For type-II materials, perfect diamagnetism occurs only below H_{c1}. Between H_{c1} and H_{c2}, the material is still superconducting.

on these macroscopic models, a model for an intermediate state between type-I and type-II superconductivity is proposed.

3.3.1 Macroscopic models for type-I superconductivity

In their phenomenological theory of superconductivity in 1935, London and London explained that a macroscopic description of type-I superconductors involves a two-fluid model [2, 86]. One fluid consists of normal electrons and the other one of superconducting electrons. Superconducting electrons cross the metal without experiencing any resistance, in contrast to electrons in a normal material, which scatter resistance along their motion. Below the critical temperature T_c, when the superconductive material loses all resistivity, the current consists of superconducting electrons and normal electrons. Above the critical temperature only normal electrons occur. Accordingly, the current density $\mathbf{J}$ is supposed to be the sum of a normal part $\mathbf{J}_n$ and a superconducting part $\mathbf{J}_s$, that is

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_s.$$

The superfluid electrons 'short circuit' the normal ones and make the overall resistivity equal to zero [76].

From now on, it is assumed that a superconductive material occupies a bounded domain $\Omega \subset \mathbb{R}^3$ with a Lipschitz continuous boundary $\partial\Omega$. The symbol $\boldsymbol{\nu}$ denotes the outward unit normal vector on $\partial\Omega$. The full Maxwell's equations ($\tilde{\delta} = 1$) and quasi-static Maxwell's equations ($\tilde{\delta} = 0$) for linear materials are considered. Thus, a linear dependence of the magnetic induction $\mathbf{B}$ and the electric displacement field $\mathbf{D}$ on respectively the magnetic field $\mathbf{H}$ and the electric field $\mathbf{E}$ is assumed, namely

$$\mathbf{B} = \mu\mathbf{H} \qquad \text{and} \qquad \mathbf{D} = \epsilon\mathbf{E}, \tag{3.3}$$

where the constant $\mu > 0$ stands for the magnetic permeability and the constant $\epsilon > 0$ for the electric permittivity of the material. In agreement with our previous notations, the quasi-static (also called the eddy current approximation of the Maxwell equations) and the full Maxwell's equations can be combined as

$$\nabla \times \mathbf{H} = \mathbf{J} + \tilde{\delta}\partial_t \mathbf{D} = \mathbf{J}_n + \mathbf{J}_s + \tilde{\delta}\epsilon\partial_t \mathbf{E}, \qquad \text{Ampère's law} \qquad (3.4)$$

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B} = -\mu\partial_t \mathbf{H}. \qquad \text{Faraday's law} \qquad (3.5)$$

Applying the divergence operator to Faraday's law (3.5) and integrating in time gives

$$\nabla \cdot \mathbf{H}(t) = \nabla \cdot \mathbf{H}(t = 0).$$

Therefore, assuming $\nabla \cdot \mathbf{H}(t = 0) = 0$, it is ensured that the magnetic field remains divergence free for any time. The normal density current $\mathbf{J}_n$ is required to satisfy Ohm's law

$$\mathbf{J}_n = \sigma \mathbf{E},$$

$\sigma > 0$ being the conductivity of the normal electrons.

London and London postulated two equations, in addition to Maxwell's equations, governing the electromagnetic field in a superconductor [86]:

$$\partial_t \mathbf{J}_s = \Lambda^{-1}\mathbf{E} \qquad \text{and} \qquad \nabla \times \mathbf{J}_s = -\Lambda^{-1}\mathbf{B}, \qquad (3.6)$$

where $\Lambda = \frac{m_e}{n_s e^2}$, with n_s the number of superelectrons per unit volume, m_e and $-e$ the mass and the electric charge of an electron respectively. These equations provide an accurate description of the two fundamental features of superconductors: perfect conductivity and perfect diamagnetism (Meissner-Ochsenfeld effect) [2]. The first equation in (3.6) explains the perfect conductivity aspect of the superconductor (the superconducting electrons suffer no resistance). It follows that

$$\mathbf{J}_s \cdot \mathbf{E} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}\Lambda \left|\mathbf{J}_s\right|_{\mathrm{e}}^2\right),$$

such that, since $\mathbf{E} = \mathbf{0}$ implies that the magnitude of $\mathbf{J}_s$ is constant, i.e. $\mathbf{J}_s$ is conservative and does not dissipate.

The second London equation in (3.6) and the quasi-static Maxwell equation $\nabla \times \mathbf{H} = \mathbf{J}_s$ give

$$\nabla \times \nabla \times \mathbf{B} = -\frac{1}{\beta}\mathbf{B},$$

where $\beta = \frac{\Lambda}{\mu}$. But $\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \Delta\mathbf{B} = -\Delta\mathbf{B}$, since $\nabla \cdot \mathbf{B} = 0$. Thus

$$\Delta\mathbf{B} = \frac{1}{\beta}\mathbf{B}. \qquad (3.7)$$

Now, the perfect diamagnetism property is demonstrated. Suppose that the surface of a superconductor lies in the y-z plane. A magnetic field is applied in the z direction parallel to the surface, $\mathbf{B} = (0, 0, B_a)$. Given that inside the superconductor the magnetic field is a function of x only, $\mathbf{B} = (0, 0, B_z(x))$, equation (3.7) is equivalent with

$$\frac{\mathrm{d}^2 B_z(x)}{\mathrm{d}x^2} = \frac{1}{\beta} B_z(x).$$

The solution as $x \geqslant 0$ is

$$B_z(x) = B_a \exp\left(-\frac{x}{\sqrt{\beta}}\right).$$

This result shows that a magnetic field is exponentially decayed at the surface of a superconductor. That is the Meissner effect. Roughly, $\mathbf{B}$ penetrates in the half-space $x \geqslant 0$ of a distance $\sqrt{\beta}$. When $x = \sqrt{\beta}$, then 63% of the flux density has declined. This is why the quantity $\lambda_L = \sqrt{\beta}$ is called the London penetration depth. For more complicated shapes and directions of the applied magnetic field, the drop in intensity of the magnetic flux density can be expressed by the second London equation.

Since $\mathbf{B}$ is divergence free, there exists a magnetic vector potential $\mathbf{A}$ such that $\mathbf{B} = \nabla \times \mathbf{A}$ and $\nabla \cdot \mathbf{A} = 0$, cf. [39]. If the domain Ω is simply connected, then $\mathbf{A}$ is uniquely determined when $\mathbf{A} \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$. This is the so-called London gauge. Therefore, the second London equation can be rewritten in the local form

$$\mathbf{J}_s = -\Lambda^{-1}\mathbf{A}. \tag{3.8}$$

Up to now, everything said so far about the electrodynamics of superconductors falls into the category of the so-called local electrodynamics. An improved form of the London equations was proposed by Pippard.

In 1953, Pippard discovered that the penetration depth in impure tin (by addition of indium) vary much more than in pure tin on application of a magnetic field and that it is possible by the addition of impurity to alter considerably the penetration depth in zero magnetic field (without producing a corresponding change in the thermodynamical properties of the material). From the London theory follows that β depends only on constants of the metal and not on any parameter that may be modified by the addition of small amounts of impurity. Therefore, Pippard stated that the London theory as it stands is unable to account for the observed variation of the penetration depth [87]. Pippard proposed the following modification of the local expression (3.8):

$$\mathbf{J}_{s,p}(\mathbf{x}, t) = \int_\Omega Q(\mathbf{x} - \mathbf{x}')\mathbf{A}(\mathbf{x}', t)\ \mathrm{d}\mathbf{x}', \qquad (\mathbf{x}, t) \in Q_T := \Omega \times (0, T),$$

with

$$Q(\mathbf{x} - \mathbf{x}')\mathbf{A}(\mathbf{x}', t) = -\widetilde{C} \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|_{\mathrm{e}}^4} \left[\mathbf{A}(\mathbf{x}', t) \cdot (\mathbf{x} - \mathbf{x}')\right] \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|_{\mathrm{e}}}{r_0}\right),$$

where $\widetilde{C} := \frac{3}{4\pi\xi_0\Lambda} > 0$. The integral is taken over the whole volume of the metal. The length ξ_0 is called the coherence length of the material. The parameter r_0 is defined by

$$\frac{1}{r_0} = \frac{1}{\xi_0} + \frac{1}{l} \Rightarrow r_0 = \frac{\xi_0 l}{\xi_0 + l},$$

with l the mean free path of the electrons in the material. This nonlocal expression is based on Chambers nonlocal Ohm's law [88].

Pippard's nonlocal law satisfactorily fits the experimental data. However, it fails to explain the vanishing of electrical resistance [2]. For this reason, the nonlocal representation of the superconductive current by Eringen [89] is considered (1984). This representation identifies the state of the superconductor, at time t, with the field $\mathbf{H}(\cdot, t)$ and is given by the linear functional

$$\mathbf{J}_{s,e}(\mathbf{x}, t) = \int_\Omega \sigma_0 \left(|\mathbf{x} - \mathbf{x}'|_{\mathrm{e}}\right) (\mathbf{x}-\mathbf{x}') \times \mathbf{H}(\mathbf{x}', t)\ \mathrm{d}\mathbf{x}' =: -(\mathcal{K}_0 \star \mathbf{H})(\mathbf{x}, t), \quad (3.9)$$

where $\sigma_0 : (0, \infty) \to \mathbb{R}$ is defined by

$$\sigma_0(s) = \begin{cases} \frac{\widetilde{C}}{2s^2} \exp\left(-\frac{s}{r_0}\right) & s < r_0, \\ 0 & s \geqslant r_0. \end{cases} \quad (3.10)$$

The dependence of $\mathbf{J}_s$ on time t is solely through $\mathbf{H}$. The points which contribute to the integral are separated by distances of order r_0 or less. The function σ_0 becomes unbounded for $\mathbf{x}' = \mathbf{x}$. Moreover, σ_0 is chosen such that it is possible to recover the London equations and the form given by Pippard from (3.9), see [2,89]. Consequently, the form by Eringen is a more direct generalization of the London theory in comparison with Pippard's nonlocal law. For this reason, the nonlocal law of Eringen is considered and is denoted by $\mathbf{J}_s$ instead of $\mathbf{J}_{s,e}$.

Taking the curl of (3.4) and the time derivative of (3.5) result into the following parabolic ($\tilde{\delta} = 0$) and hyperbolic ($\tilde{\delta} = 1$) integro-differential equation

$$\tilde{\delta}\epsilon\mu\partial_{tt}\mathbf{H} + \sigma\mu\partial_t\mathbf{H} + \nabla \times \nabla \times \mathbf{H} + \nabla \times (\mathcal{K}_0 \star \mathbf{H}) = \mathbf{0}. \quad (3.11)$$

The well-posedness of the nonlocal parabolic model ($\tilde{\delta} = 0$ in (3.11)) and the nonlocal hyperbolic model ($\tilde{\delta} = 1$ in (3.11)) is studied in detail in Chapter 4 and Chapter 5 respectively, using Rothe's method (cf. Section 2.12). Recent engineering applications about nonlocal effects in superconductors can be found in [90–93].

3.3.2 Available macroscopic models for type-II superconductivity

One of the first macroscopic models for type-II superconductors was Bean's critical-state model [94]. This model imposes that a current either flows at the critical level

J_c (in the mixed state) or doesn't flow at all (perfect diamagnetism), i.e. the magnitude of $\mathbf{J}$ is equal to its critical value J_c at all points $\mathbf{x}$ where the electric field $\mathbf{E}(\mathbf{x})$ is not equal to zero. The critical state J_c is a constant determined by the properties of the superconductive material. This model is used in numerical simulations as follows: $\mathbf{J} = J_c \frac{\mathbf{E}}{|\mathbf{E}|_\mathrm{e}}$ if $|\mathbf{E}|_\mathrm{e} \neq 0$ and $\frac{\partial \mathbf{J}}{\partial t} = \mathbf{0}$ if $|\mathbf{E}|_\mathrm{e} = 0$. Many authors have studied this model [95–100]. Bean's critical-state model is fairly accurate for simple geometries as plane slabs and circular cross-section cylinders. Unfortunately, it is not fully applicable to superconductors with smooth current-voltage characteristics. In [101], the author states that type-II superconductors can be treated as electrically nonlinear conductors due to processes of magnetic hysteresis and nonzero resistivity. From this sight, another model frequently used in the modelling of type-II superconductors is the power law constitutive relation by Rhyner [101, 102]:

$$\mathbf{E} = \sigma_c^{-n} \left|\mathbf{J}\right|_\mathrm{e}^{n-1} \mathbf{J}, \qquad n \in (7, 1000), \tag{3.12}$$

where σ_c is some parameter that coordinates the dimensions of both sides in the expression. The value of n depends on the superconducting material and is a measure of the sharpness of the resistive transition. If $n = 1$, the relation (3.12) leads to the linear Ohm's law. If $n \to \infty$, the solution to the power law formulation converges to the solution to Beans critical-state formulation [97, 98]. This relation in combination with the eddy current approximation of the Maxwell's equations is investigated in [103–107]. Employing (3.4) and taking the curl of (3.12) lead to the following equation for the magnetic field:

$$\mu \partial_t \mathbf{H} + \sigma_c^{-n} \nabla \times \Big(\left|\nabla \times \mathbf{H}\right|_\mathrm{e}^{n-1} \nabla \times \mathbf{H} \Big) = \mathbf{0}. \tag{3.13}$$

Note that (3.12) was firstly derived with the intention to model the soft transition of the current density, before it had been justified [103]. When $\tilde{\delta} = 1$, no equation in terms of only the magnetic field can be obtained.

3.3.3 Macroscopic model for an intermediate state between type-I and type-II superconductivity

Recently, there has been an increased interest in superconductors with several superconducting components. They arise for instance in multiband superconductors. The classification into types-I and II is insufficient for such multicomponent superconductors [108]. For instance, physicists have found that the material 'magnesium diboride' combines the characteristics of both types [109–111]. This leads to a complete new kind of superconductors, the so-called type-1.5 superconductors [110], which allow coexistence of various properties of type-I and type-II superconductors. Type-1.5 materials can be made by placing a thin layer of type-I material onto a thin layer of type-II material [112–115]. For more articles about type-1.5 superconductors, the reader is referred to [108, 116, 117]. From this viewpoint, by introducing a real parameter $\beta \geqslant 1$ and a real function $f(\beta)$, it is proposed to combine equations (3.11) for $\tilde{\delta} = 0$ and (3.13) to

$$\begin{aligned}\mu\partial_t\mathbf{H} + \sigma^{-1}f(\beta)\nabla\times\nabla\times\mathbf{H} + \sigma_c^{-\beta}g(\beta)\nabla\times\left(|\nabla\times\mathbf{H}|_{\mathrm{e}}^{\beta-1}\nabla\times\mathbf{H}\right)\\ + \sigma^{-1}f(\beta)\nabla\times(\mathcal{K}_0\star\mathbf{H}) = \mathbf{0},\end{aligned} \quad (3.14)$$

with

$$g(\beta) := 1 - f(\beta)$$

and where $f \in \mathrm{C}([1,\infty))$ is monotone decreasing and satisfies $f(1) = 1$ and $0 \leqslant f(\beta) \leqslant 1$ for $\beta > 1$. Moreover, suppose that f is zero or sufficiently small for $\beta > 7$. For instance, f can take the form

$$f(\beta) = \begin{cases} \frac{(-1)^\alpha}{6^\alpha}(\beta-7)^\alpha & 1 \leqslant \beta \leqslant 7,\\ 0 & \beta > 7; \end{cases}$$

$$f(\beta) = \exp(-k\beta),$$

with $\alpha \in \mathbb{N}$ and where $k > 0$ represents the speed of convergence to zero. This implies that $g \in \mathrm{C}([1,\infty))$ is monotone increasing with $g(1) = 0$ and $0 \leqslant g(\beta) \leqslant 1$. Equation (3.14) simplifies to equation (3.11) with $\tilde{\delta} = 0$ for type-I superconductors in the case that $\beta = 1$. If $7 < \beta < 1000$, then equation (3.14) equals or approximates equation (3.13) for type-II superconductivity depending on the choice of f. Note that in practical applications β is less than 1000, but in the analysis presented in Chapter 6 β is allowed to be larger. The intermediate phase ($1 < \beta \leqslant 7$) is attributed to an intermediate state between type-I and type-II superconductivity. The focus of Chapter 6 is on the mathematical analysis of equation (3.14) and not on its implementation.

4

Nonlocal parabolic problem for type-I superconductivity

This chapter is based on the articles [118] and [119], which are published in
Numerical Methods for Partial Differential Equations
and
Journal of Computational and Applied Mathematics.

The aim of this chapter is to address the well-posedness of the following parabolic problem in terms of the magnetic field $\mathbf{H}$:

$$\begin{cases} \partial_t \mathbf{H} + \nabla \times \nabla \times \mathbf{H} + \nabla \times (\mathcal{K}_0 \star \mathbf{H}) & = \mathbf{F} \quad \text{in } Q_T, \\ \mathbf{H} \times \boldsymbol{\nu} & = \mathbf{0} \quad \text{on } \Sigma_T, \\ \mathbf{H}(\mathbf{x}, 0) & = \mathbf{H}_0 \quad \text{in } \Omega. \end{cases} \tag{4.1}$$

This is done by using Rothe's method, which is based on a semidiscretization in time, cf. Section 2.12. This method includes the development of a numerical scheme for computations. Error estimates for both the time and the space discretization are derived.

The domain $\Omega \subset \mathbb{R}^3$ occupying a type-I superconductor is a bounded Lipschitz domain. Note that $Q_T = \Omega \times (0, T]$ and $\Sigma_T = \partial\Omega \times (0, T]$, with T the final time. The problem (4.1) is obtained by setting $\mu = \sigma = 1$ in (3.11) in the case that $\tilde{\delta} = 0$, where

$$(\mathcal{K}_0 \star \mathbf{H})(\mathbf{x}, t) = -\int_\Omega \sigma_0 \left(|\mathbf{x} - \mathbf{x}'|_{\mathrm{e}}\right) (\mathbf{x} - \mathbf{x}') \times \mathbf{H}(\mathbf{x}', t) \; \mathrm{d}\mathbf{x}', \tag{4.2}$$

with $\sigma_0 : (0, \infty) \to \mathbb{R}$ defined by

$$\sigma_0(s) = \begin{cases} \frac{\widetilde{C}}{2s^2} \exp\left(-\frac{s}{r_0}\right) & s < r_0, \\ 0 & s \geqslant r_0. \end{cases}$$

The parameters $\widetilde{C}$ and r_0 depend on the material under consideration. Also a possible source term $\mathbf{F}$ is considered in the right-hand side. To obtain the magnetic boundary condition in (4.1), it is assumed that the magnetic field equals zero outside the domain Ω [39, p. 8]. Nevertheless, from mathematical point of view, it is also possible to consider a boundary condition of the form $\mathbf{H} \times \boldsymbol{\nu} = \mathbf{g}$.

Problem (4.1) is based on the eddy current approximation of the Maxwell equations. This approximation is valid in highly conductive media [120, 121]. A mathematical analysis of integro-differential equations arising from the nonlocal theory of superconductivity has been carried out for smooth electromagnetic fields in [122]. The model considered in that article was written in terms of the vector potential of $\mathbf{H}$. The analysis was based on the spectral analysis and expansion in terms of eigenfunctions. In this chapter, a variational approach is proposed, which can be applied to general geometries without knowledge of the spectrum. The mathematics of the eddy-current approximation has recently been developed in some other settings, see for instance [123–125]. The main difference in the analysis of problem (4.1), in comparison with the available results, is caused by the nonlocal term in (4.1).

This chapter is organized as follows. Section 4.1 contains crucial estimates containing the kernel $\mathcal{K}_0$. The uniqueness of a solution to problem (4.1) is addressed in detail in Section 4.2 and the well-posedness of the problem is shown in Section 4.3. A time-discrete numerical scheme is developed. The existence of a weak solution for each time step is shown. Also the convergence of the method is discussed and error estimates are derived. A modified scheme is considered in Section 4.4. In Section 4.5, another convolution kernel is derived under an additional assumption. The positive definiteness of this kernel is shown. Using the obtained expression, it is demonstrated that the solution of the original model satisfies a simpler equation, which is described and analysed in Subsection 4.5.2. A comparison with the London equations is given into Subsection 4.5.4. Moreover, some numerical experiments are developed in Section 4.6. Finally, a fully discrete approximation scheme is proposed in Section 4.7.

4.1 Useful estimates

The analysis starts with the derivation of some useful estimates on the kernels. The integral in (4.2) is taken over the subdomain $\Omega_{\mathbf{x}} := \Omega \cap B(\mathbf{x}, r_0)$. Using spherical coordinates, one can see that the singular vectorial field $\sigma_0(|\mathbf{x}|)\mathbf{x}$ belongs to $\mathbf{L}^p(\Omega)$ for $1 \leqslant p < 3$:

$$\begin{aligned}
\int_{\Omega} |\sigma_0(|\mathbf{x}|)\,\mathbf{x}|^p \, \mathrm{d}\mathbf{x} &= \int_{\Omega_{\mathbf{0}}} |\sigma_0(|\mathbf{x}|)\,\mathbf{x}|^p \, \mathrm{d}\mathbf{x} \\
&\overset{(3.10)}{\leqslant} \int_{B(\mathbf{0},r_0)} \frac{C}{|\mathbf{x}|^{2p}} \left| \exp\left(-\frac{|\mathbf{x}|}{r_0} \right) \right|^p |\mathbf{x}|^p \, \mathrm{d}\mathbf{x} \\
&\leqslant C \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin(\theta)\,\mathrm{d}\theta \int_0^{r_0} r^{2-p}\,\mathrm{d}r \\
&\leqslant C \left[\frac{r^{3-p}}{3-p} \right]_0^{r_0} < \infty.
\end{aligned}$$

Consequently, using Hölder's inequality, it is easily checked that

$$\begin{aligned}
|(\mathcal{K}_0 \star \mathbf{H})(\mathbf{x},t)| &\leqslant \int_{\Omega} |\sigma_0(|\mathbf{x}-\mathbf{x}'|)(\mathbf{x}-\mathbf{x}')|\,|\mathbf{H}(\mathbf{x}',t)| \, \mathrm{d}\mathbf{x}' \\
&\leqslant \sqrt[p]{\int_{\Omega} |\sigma_0(|\mathbf{x}-\mathbf{x}'|)(\mathbf{x}-\mathbf{x}')|^p \, \mathrm{d}\mathbf{x}'} \sqrt[q]{\int_{\Omega} |\mathbf{H}(\mathbf{x}',t)|^q \, \mathrm{d}\mathbf{x}'} \\
&\leqslant C(q) \|\mathbf{H}(t)\|_q, \qquad (4.3)
\end{aligned}$$

for all $q > \frac{3}{2}$ and $(\mathbf{x},t) \in Q_T$. Hence, the Cauchy and Young inequalities together with (4.3) for $q = 2$ yield that

$$(\mathcal{K}_0 \star \mathbf{H}_1, \nabla \times \mathbf{H}_2) \leqslant C_\varepsilon \|\mathbf{H}_1\|^2 + \varepsilon \|\nabla \times \mathbf{H}_2\|^2 \qquad (4.4)$$

for all $\mathbf{H}_1 \in \mathbf{L}^2(\Omega)$ and $\mathbf{H}_2 \in \mathbf{H}(\mathbf{curl};\Omega)$. The position of the positive constants ε and C_ε can be interchanged.

4.2 Uniqueness of a solution

The goal of this section is to prove the uniqueness of the solution. The variational formulation of (4.1) reads as:

Given $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ and $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, find $\mathbf{H}(t) \in \mathbf{H}_0(\mathbf{curl};\Omega)$ with $\partial_t \mathbf{H}(t) \in \mathbf{H}_0^{-1}(\mathbf{curl};\Omega) := \mathbf{H}_0(\mathbf{curl};\Omega)^*$ such that

$$(\partial_t \mathbf{H}(t), \boldsymbol{\varphi}) + (\nabla \times \mathbf{H}(t), \nabla \times \boldsymbol{\varphi}) + ((\mathcal{K}_0 \star \mathbf{H})(t), \nabla \times \boldsymbol{\varphi}) = (\mathbf{F}(t), \boldsymbol{\varphi}), \qquad (4.5)$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$ and a.a. $t \in (0,T)$.

That this variational formulation is well-defined follows from the natural stability of the solution $\mathbf{H}$ of (4.1).

Theorem 4.2.1 (Stability). *Suppose that* $\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$ *and that* $\mathbf{H}$ *solves* (4.1).

(i) If $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}(t)\|^2 + \int_0^T \|\nabla\times\mathbf{H}\|^2 \leqslant C.$$

(ii) If $\nabla\cdot\mathbf{F}(t) = 0 = \nabla\cdot\mathbf{H}_0$ *for any* $t\in[0,T]$, *then* $\nabla\cdot\mathbf{H}(t) = 0$ *for any* $t\in[0,T]$. *Moreover, we have that*

$$\int_0^T \|\partial_t\mathbf{H}\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C$$

and $\mathbf{H}\in\mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$.

(iii) If $\mathbf{H}_0\in\mathbf{H}(\mathbf{curl};\Omega)$ *then*

$$\max_{t\in[0,T]} \|\nabla\times\mathbf{H}(t)\|^2 + \int_0^T \|\partial_t\mathbf{H}\|^2 \leqslant C.$$

(iv) If $\mathbf{F}(0)\in\mathbf{L}^2(\Omega), \partial_t\mathbf{F}\in\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right),\ \nabla\times(\mathcal{K}_0\star\mathbf{H}_0)\in\mathbf{L}^2(\Omega),\ \mathbf{H}_0\in\mathbf{H}_0(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0\in\mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\partial_t\mathbf{H}(t)\|^2 + \int_0^T \|\nabla\times\partial_t\mathbf{H}\|^2 \leqslant C.$$

Proof. (i) Setting $\boldsymbol{\varphi} = \mathbf{H}(t)$ in (4.5) and integrating in time over $t\in(0,\xi)\subset(0,T)$, we get that

$$\tfrac{1}{2}\|\mathbf{H}(\xi)\|^2 + \int_0^\xi \|\nabla\times\mathbf{H}\|^2 = \tfrac{1}{2}\|\mathbf{H}_0\|^2 + \int_0^\xi (\mathbf{F},\mathbf{H}) - \int_0^\xi (\mathcal{K}_0\star\mathbf{H},\nabla\times\mathbf{H}).$$

Using estimate (4.4), we obtain

$$\|\mathbf{H}(\xi)\|^2 + \int_0^\xi \|\nabla\times\mathbf{H}\|^2 \leqslant C + C_\varepsilon\int_0^\xi \|\mathbf{H}\|^2 + \varepsilon\int_0^\xi \|\nabla\times\mathbf{H}\|^2.$$

We obtain the desired result choosing a sufficiently small positive ε and involving the Grönwall argument.

(ii) Take the divergence of (4.1) and integrate in time to arrive at $\nabla\cdot\mathbf{H}(t) = \nabla\cdot\mathbf{H}_0 = 0$ for all $t\in[0,T]$. We rewrite (4.5) for $\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)$ as follows

$$(\partial_t\mathbf{H}(t),\boldsymbol{\varphi}) = (\mathbf{F}(t),\boldsymbol{\varphi}) - (\nabla\times\mathbf{H}(t),\nabla\times\boldsymbol{\varphi}) - ((\mathcal{K}_0\star\mathbf{H})(t),\nabla\times\boldsymbol{\varphi}).$$

The integral in the LHS can be interpreted in the sense of duality, i.e. seeing $\partial_t \mathbf{H}(t)$ as an element of $\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)$. A simple calculation implies

$$|\left(\mathbf{F}(t),\boldsymbol{\varphi}\right)| \leqslant \|\mathbf{F}(t)\|\,\|\boldsymbol{\varphi}\|\,, \quad |\left(\nabla\times\mathbf{H}(t),\nabla\times\boldsymbol{\varphi}\right)| \leqslant \|\nabla\times\mathbf{H}(t)\|\,\|\nabla\times\boldsymbol{\varphi}\|$$

and

$$|\left((\mathcal{K}_0\star\mathbf{H})(t),\nabla\times\boldsymbol{\varphi}\right)| \overset{(4.3)}{\lesssim} \|\mathbf{H}(t)\|\,\|\nabla\times\boldsymbol{\varphi}\|\,.$$

The dual norm in $\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)$ is

$$\|\partial_t\mathbf{H}(t)\|_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} = \sup_{\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)} \frac{(\partial_t\mathbf{H}(t),\boldsymbol{\varphi})}{\|\boldsymbol{\varphi}\|_{\mathbf{H}_0(\mathbf{curl};\Omega)}}.$$

Therefore, using (i), we deduce that

$$\int_0^T \|\partial_t\mathbf{H}(s)\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)}\,\mathrm{d}s \leqslant C.$$

Consider the following evolution triple (or sometimes called Gelfand's triple) of spaces

$$\mathbf{H}_0(\mathbf{curl};\Omega) \hookrightarrow \mathbf{L}^2(\Omega) \cong \mathbf{L}^2(\Omega)^* \hookrightarrow \mathbf{H}_0^{-1}(\mathbf{curl};\Omega).$$

We know that

$$\mathbf{H}\in \mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right) \quad \text{and} \quad \partial_t\mathbf{H}\in\mathrm{L}^2\left((0,T),\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)\right).$$

Applying Lemma 2.9.5(iii), we get that $\mathbf{H}\in\mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$.

(iii) Now, we set $\boldsymbol{\varphi}=\partial_t\mathbf{H}(t)$ in (4.5) and integrate in time over $t\in(0,\xi)\subset(0,T)$ to obtain that

$$\begin{aligned}\int_0^\xi \|\partial_t\mathbf{H}\|^2 + \tfrac{1}{2}\|\nabla\times\mathbf{H}(\xi)\|^2 & \\ = \tfrac{1}{2}\|\nabla\times\mathbf{H}_0\|^2 + \int_0^\xi(\mathbf{F},\partial_t\mathbf{H}) - \int_0^\xi(\mathcal{K}_0\star\mathbf{H},\nabla\times\partial_t\mathbf{H})\,.\end{aligned}$$

The second term on the RHS can be estimated using the Cauchy and Young inequalities as follows

$$\int_0^\xi(\mathbf{F},\partial_t\mathbf{H}) \leqslant \varepsilon\int_0^\xi\|\partial_t\mathbf{H}\|^2 + C_\varepsilon\int_0^\xi\|\mathbf{F}\|^2 \leqslant \varepsilon\int_0^\xi\|\partial_t\mathbf{H}\|^2 + C_\varepsilon.$$

Using the integration by parts formula, we may write

$$\begin{aligned}\int_0^\xi(\mathcal{K}_0\star\mathbf{H},\nabla\times\partial_t\mathbf{H}) &= (\mathcal{K}_0\star\mathbf{H},\nabla\times\mathbf{H})|_0^\xi - \int_0^\xi(\mathcal{K}_0\star\partial_t\mathbf{H},\nabla\times\mathbf{H}) \\ &\overset{(4.4)}{\leqslant} C+\varepsilon\|\nabla\times\mathbf{H}(\xi)\|^2 + C_\varepsilon\|\mathbf{H}(\xi)\|^2 \\ &\quad +\varepsilon\int_0^\xi\|\partial_t\mathbf{H}\|^2 + C_\varepsilon\int_0^\xi\|\nabla\times\mathbf{H}\|^2 \\ &\overset{(i)}{\leqslant} \varepsilon\|\nabla\times\mathbf{H}(\xi)\|^2 + \varepsilon\int_0^\xi\|\partial_t\mathbf{H}\|^2 + C_\varepsilon.\end{aligned}$$

Collecting all considerations above and fixing a sufficiently small positive ε, we conclude the proof.

(iv) First, we differentiate (4.5) with respect to the time variable. Then, we set $\boldsymbol{\varphi} = \partial_t \mathbf{H}(t)$ and integrate in time over $t \in (0,\xi) \subset (0,T)$ to get

$$\begin{aligned}\tfrac{1}{2}\left\|\partial_t \mathbf{H}(\xi)\right\|^2 &+ \int_0^\xi \left\|\nabla\times\partial_t\mathbf{H}\right\|^2 \\ &= \tfrac{1}{2}\left\|\partial_t\mathbf{H}(0)\right\|^2 + \int_0^\xi \left(\partial_t\mathbf{F}, \partial_t\mathbf{H}\right) - \int_0^\xi \left(\mathcal{K}_0 \star \partial_t\mathbf{H}, \nabla\times\partial_t\mathbf{H}\right).\end{aligned}$$

Employing the Cauchy and Young inequalities, (4.4) and (iii) to the RHS, we get

$$\left\|\partial_t\mathbf{H}(\xi)\right\|^2 + \int_0^\xi \left\|\nabla\times\partial_t\mathbf{H}\right\|^2 \leqslant C_\varepsilon + \left\|\partial_t\mathbf{H}(0)\right\|^2 + \varepsilon\int_0^\xi \left\|\nabla\times\partial_t\mathbf{H}\right\|^2.$$

Fixing a small ε and applying Grönwall's argument, we arrive at

$$\max_{t\in[0,T]}\left\|\partial_t\mathbf{H}(t)\right\|^2 + \int_0^T \left\|\nabla\times\partial_t\mathbf{H}\right\|^2 \lesssim 1 + \left\|\partial_t\mathbf{H}(0)\right\|^2.$$

To find a bound for $\left\|\partial_t\mathbf{H}(0)\right\|$, it is assumed that the variational formulation (4.5) is satisfied at $t=0$, i.e.

$$\left(\partial_t\mathbf{H}(0),\boldsymbol{\varphi}\right) + \left(\nabla\times\mathbf{H}_0,\nabla\times\boldsymbol{\varphi}\right) + \left(\mathcal{K}_0\star\mathbf{H}_0,\nabla\times\boldsymbol{\varphi}\right) = \left(\mathbf{F}(0),\boldsymbol{\varphi}\right)$$

for all $\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)$. For this, it is required that $\mathbf{H}_0\in\mathbf{H}_0(\mathbf{curl};\Omega)$. Applying Green's theorem in a backward way gives for all $\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)$ that

$$\left(\partial_t\mathbf{H}(0),\boldsymbol{\varphi}\right) = \left(\mathbf{F}(0),\boldsymbol{\varphi}\right) - \left(\nabla\times\nabla\times\mathbf{H}_0,\boldsymbol{\varphi}\right) - \left(\nabla\times\left(\mathcal{K}_0\star\mathbf{H}_0\right),\boldsymbol{\varphi}\right).$$

The term $\partial_t\mathbf{H}(0)$ can be seen as a functional on $\mathbf{H}_0(\mathbf{curl};\Omega)$. The RHS is a linear and bounded functional on $\mathbf{H}_0(\mathbf{curl};\Omega)$. This implies that the RHS can be extended to a functional $\widetilde{\partial_t\mathbf{H}(0)}$ on $\mathbf{L}^2(\Omega)$ by the Hahn-Banach theorem. Moreover,

$$\left\|\widetilde{\partial_t\mathbf{H}(0)}\right\| = \sup_{\substack{\boldsymbol{\varphi}\in\mathbf{L}^2(\Omega)\\ \|\boldsymbol{\varphi}\|\leqslant 1}} \left(\widetilde{\partial_t\mathbf{H}(0)},\boldsymbol{\varphi}\right) = \sup_{\substack{\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)\\ \|\boldsymbol{\varphi}\|\leqslant 1}} \left(\partial_t\mathbf{H}(0),\boldsymbol{\varphi}\right) \lesssim 1,$$

i.e. $\widetilde{\partial_t\mathbf{H}(0)}\in\mathbf{L}^2(\Omega)$. Therefore, thanks to the density of $(\mathrm{C}_0^\infty(\Omega))^3$ in $\mathbf{H}_0(\mathbf{curl};\Omega)$ and Theorem 2.8.1, it holds that

$$\widetilde{\partial_t\mathbf{H}(0)} = \mathbf{F}(0) - \nabla\times\nabla\times\mathbf{H}_0 - \nabla\times\left(\mathcal{K}_0\star\mathbf{H}_0\right), \quad \text{a.e. in } \Omega.$$

Finally, we identify $\widetilde{\partial_t\mathbf{H}(0)}$ and $\partial_t\mathbf{H}(0)$. $\square$

Theorem 4.2.2 (Uniqueness)**.** *The problem* (4.1) *admits at most one solution* $\mathbf{H} \in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathbf{H}_0(\mathbf{curl};\Omega)\right)$.

Proof. Assume that we have two solutions $\mathbf{H}_1$ and $\mathbf{H}_2$. Then $\mathbf{H} = \mathbf{H}_1 - \mathbf{H}_2$ fulfils (4.1) with $\mathbf{H}_0 = \mathbf{0} = \mathbf{F}$. Setting $\boldsymbol{\varphi} = \mathbf{H}(t)$ in (4.5) and integrating in time over $t \in (0,\xi) \subset (0,T)$, we find that

$$\tfrac{1}{2}\left\|\mathbf{H}(\xi)\right\|^2 + \int_0^\xi \left\|\nabla \times \mathbf{H}\right\|^2 + \int_0^\xi (\mathcal{K}_0 \star \mathbf{H}, \nabla \times \mathbf{H}) = 0.$$

Using inequality (4.4) for the last term, we arrive at

$$\left\|\mathbf{H}(\xi)\right\|^2 + \int_0^\xi \left\|\nabla \times \mathbf{H}\right\|^2 \leqslant \varepsilon \int_0^\xi \left\|\nabla \times \mathbf{H}\right\|^2 + C_\varepsilon \int_0^\xi \left\|\mathbf{H}\right\|^2.$$

Fixing a sufficiently small positive ε and applying Grönwall's argument, we get that $\mathbf{H} = \mathbf{0}$ a.e. in Q_T. ☐

4.3 Existence of a solution

To address the existence of a solution to (4.1), a semidiscretization in time is employed. This discretization is based on Rothe's method, cf. Section 2.12. The interval $[0,T]$ is divided into n equidistant subintervals $[t_{i-1}, t_i]$ with time step $\tau = \frac{T}{n}$, thus $t_i = i\tau, i = 0, \ldots, n$. With the standard notation for the discretized fields

$$\mathbf{h}_i \approx \mathbf{H}(t_i), \qquad \delta\mathbf{h}_i = \frac{\mathbf{h}_i - \mathbf{h}_{i-1}}{\tau},$$

the following linear recurrent implicit scheme is proposed to approximate the original problem ($\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$)

$$\left\{ \begin{aligned} (\delta\mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) + (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) &= (\mathbf{F}_i, \boldsymbol{\varphi}), \\ \mathbf{h}_0 &= \mathbf{H}_0, \end{aligned} \right. \tag{4.6}$$

which is equivalent to

$$\begin{aligned} a(\mathbf{h}_i, \boldsymbol{\varphi}) := \left(\frac{\mathbf{h}_i}{\tau}, \boldsymbol{\varphi}\right) + (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) + (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) \\ = (\mathbf{F}_i, \boldsymbol{\varphi}) + \left(\frac{\mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

Theorem 4.3.1. *Suppose that* $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$. *Then the variational problem* (4.6) *admits a unique solution* $\mathbf{h}_i \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *for any* $i = 1, \ldots, n$ *if* $\tau < \tau_0$.

Proof. The bilinear form a is elliptic for $\tau < \tau_0$:

$$\begin{aligned} a(\mathbf{h},\mathbf{h}) &\geqslant \frac{1}{\tau}\|\mathbf{h}\|^2 + \|\nabla\times\mathbf{h}\|^2 - |(\mathcal{K}_0 \star \mathbf{h}, \nabla\times\mathbf{h})| \\ &\overset{(4.4)}{\geqslant} \left(\frac{1}{\tau} - C_\varepsilon\right)\|\mathbf{h}\|^2 + (1-\varepsilon)\|\nabla\times\mathbf{h}\|^2 \\ &\geqslant C(\tau)\|\mathbf{h}\|^2_{\mathbf{H}_0(\mathbf{curl};\Omega)}, \end{aligned}$$

with $\varepsilon < 1$ fixed. Moreover, a is continuous in $\mathbf{H}_0(\mathbf{curl};\Omega)$. The functional f_i is linear and bounded in $\mathbf{H}_0(\mathbf{curl};\Omega)$ if $\mathbf{h}_{i-1} \in \mathbf{L}^2(\Omega)$. Therefore, applying Lax-Milgram's lemma 2.11.1 gives the existence of a unique solution to (4.6) for any $i = 1,\dots,n$ if $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$. $\square$

4.3.1 A priori estimates

First, basic stability results for $\mathbf{h}_i$ are derived. The first two a priori estimates in the following theorem can serve as uniform bounds to prove convergence (see Remark 4.3.2).

Lemma 4.3.1 (A priori estimates)**.** *Suppose that* $\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$.

(i) Let $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$*. Then, there exists a positive constant* C *such that*

$$\max_{1\leqslant i\leqslant n}\|\mathbf{h}_i\|^2 + \sum_{i=1}^n \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\nabla\times\mathbf{h}_i\|^2\tau \leqslant C$$

for all $\tau < \tau_0$.

(ii) If $\nabla\cdot\mathbf{H}_0 = 0 = \nabla\cdot\mathbf{F}_i$ *for* $i = 1,\dots,n$*, then* $\nabla\cdot\mathbf{h}_i = 0$ *for all* $i = 1,\dots,n$*. Moreover, we have that*

$$\tau\sum_{i=1}^n \|\delta\mathbf{h}_i\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C$$

for all $\tau < \tau_0$.

(iii) If $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n}\|\nabla\times\mathbf{h}_i\|^2 + \sum_{i=1}^n \|\nabla\times\mathbf{h}_i - \nabla\times\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\delta\mathbf{h}_i\|^2\tau \leqslant C$$

for all $\tau < \tau_0$.

(iv) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$*,* $\partial_t\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$*,* $\nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega)$*,* $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n}\|\delta\mathbf{h}_i\|^2 + \sum_{i=1}^n \|\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\nabla\times\delta\mathbf{h}_i\|^2\tau \leqslant C$$

for all $\tau < \tau_0$.

Proof. (i) Setting $\boldsymbol{\varphi} = \mathbf{h}_i$ in (4.6), multiplying by τ and summing up for $i = 1, \ldots, j$ with $1 \leqslant j \leqslant n$, we have that

$$\sum_{i=1}^{j} (\delta \mathbf{h}_i, \mathbf{h}_i)\, \tau + \sum_{i=1}^{j} \|\nabla \times \mathbf{h}_i\|^2 \tau + \sum_{i=1}^{j} (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \mathbf{h}_i)\, \tau = \sum_{i=1}^{j} (\mathbf{F}_i, \mathbf{h}_i)\, \tau.$$

For the first term on the left-hand side (LHS), we use Abel's summation rule

$$2 \sum_{i=1}^{j} (\delta \mathbf{h}_i, \mathbf{h}_i)\, \tau = \|\mathbf{h}_j\|^2 - \|\mathbf{H}_0\|^2 + \sum_{i=1}^{j} \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 .$$

For the third term on the LHS, the relation (4.4) implies that

$$\left| \sum_{i=1}^{j} (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \mathbf{h}_i)\, \tau \right| \leqslant \varepsilon \sum_{i=1}^{j} \|\nabla \times \mathbf{h}_i\|^2 \tau + C_\varepsilon \sum_{i=1}^{j} \|\mathbf{h}_i\|^2 \tau.$$

For the RHS, we apply the Cauchy and Young inequalities to get for a fixed small ε that

$$\|\mathbf{h}_j\|^2 + \sum_{i=1}^{j} \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^{j} \|\nabla \times \mathbf{h}_i\|^2 \tau \lesssim 1 + \sum_{i=1}^{j} \|\mathbf{F}_i\|^2 \tau + \sum_{i=1}^{j} \|\mathbf{h}_i\|^2 \tau.$$

Applying Grönwall's argument, we conclude the proof.

(ii) The result can be readily obtained applying the divergence operator to

$$\delta \mathbf{h}_i + \nabla \times \nabla \times \mathbf{h}_i + \nabla \times (\mathcal{K}_0 \star \mathbf{h}_i) = \mathbf{F}_i.$$

It holds for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ that

$$(\delta \mathbf{h}_i, \boldsymbol{\varphi}) = (\mathbf{F}_i, \boldsymbol{\varphi}) - (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) - (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) .$$

The integral in the LHS has to be interpreted in the sense of duality, i.e. seeing $\delta \mathbf{h}_i$ as an operator from $\mathbf{H}_0(\mathbf{curl}; \Omega)$ to $\mathbb{R}$. We may write

$$|\,(\mathbf{F}_i, \boldsymbol{\varphi})\,| \leqslant \|\mathbf{F}_i\| \, \|\boldsymbol{\varphi}\| , \qquad |\,(\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})\,| \leqslant \|\nabla \times \mathbf{h}_i\| \, \|\nabla \times \boldsymbol{\varphi}\|$$

and

$$|\,(\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})\,| \overset{(4.3)}{\lesssim} \|\mathbf{h}_i\| \, \|\nabla \times \boldsymbol{\varphi}\| .$$

Thus using the dual norm

$$\|\delta \mathbf{h}_i\|_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} = \sup_{\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)} \frac{(\delta \mathbf{h}_i, \boldsymbol{\varphi})}{\|\boldsymbol{\varphi}\|_{\mathbf{H}_0(\mathbf{curl};\Omega)}}$$

and (i), we deduce that

$$\tau \sum_{i=1}^{n} \|\delta \mathbf{h}_i\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C.$$

(iii) Setting $\boldsymbol{\varphi} = \delta \mathbf{h}_i$ in (4.6), multiplying by τ and summing up for $i = 1, \ldots, j$, we have

$$\sum_{i=1}^{j} \|\delta \mathbf{h}_i\|^2 \tau + \sum_{i=1}^{j} (\nabla \times \mathbf{h}_i, \nabla \times \mathbf{h}_i - \nabla \times \mathbf{h}_{i-1}) + \sum_{i=1}^{j} (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \delta \mathbf{h}_i) \tau = \sum_{i=1}^{j} (\mathbf{F}_i, \delta \mathbf{h}_i) \tau.$$

Abel's summation rule helps us to get

$$2 \sum_{i=1}^{j} (\nabla \times \mathbf{h}_i, \nabla \times \mathbf{h}_i - \nabla \times \mathbf{h}_{i-1}) = \|\nabla \times \mathbf{h}_j\|^2 - \|\nabla \times \mathbf{H}_0\|^2 + \sum_{i=1}^{j} \|\nabla \times \mathbf{h}_i - \nabla \times \mathbf{h}_{i-1}\|^2$$

and

$$\sum_{i=1}^{j} (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \delta \mathbf{h}_i) \tau = (\mathcal{K}_0 \star \mathbf{h}_j, \nabla \times \mathbf{h}_j) - (\mathcal{K}_0 \star \mathbf{h}_0, \nabla \times \mathbf{h}_0) - \sum_{i=1}^{j} (\mathcal{K}_0 \star \delta \mathbf{h}_i, \nabla \times \mathbf{h}_{i-1}) \tau.$$

Hence, using (i), we obtain

$$\left| \sum_{i=1}^{j} (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \delta \mathbf{h}_i) \tau \right| \overset{(4.4)}{\leqslant} C_\varepsilon + \varepsilon \|\nabla \times \mathbf{h}_j\|^2 + \varepsilon \sum_{i=1}^{j} \|\delta \mathbf{h}_i\|^2 \tau.$$

The RHS can be estimated using the Cauchy and Young inequalities as follows

$$\sum_{i=1}^{j} (\mathbf{F}_i, \delta \mathbf{h}_i) \tau \leqslant \varepsilon \sum_{i=1}^{j} \|\delta \mathbf{h}_i\|^2 \tau + C_\varepsilon \sum_{i=1}^{j} \|\mathbf{F}_i\|^2 \tau \leqslant C_\varepsilon + \varepsilon \sum_{i=1}^{j} \|\delta \mathbf{h}_i\|^2 \tau.$$

Putting things together and fixing a sufficiently small positive ε, we conclude the proof.

(iv) We set

$$\delta \mathbf{h}_0 := \partial_t \mathbf{H}(0) = \mathbf{F}(0) - \nabla \times \nabla \times \mathbf{H}_0 - \nabla \times (\mathcal{K}_0 \star \mathbf{H}_0).$$

We subtract (4.6) for $i = i - 1$ from (4.6), then we set $\boldsymbol{\varphi} = \delta \mathbf{h}_i$ and we sum the result up for $i = 1, \ldots, j$ with $1 \leqslant j \leqslant n$ to get

$$\sum_{i=1}^{j} \left(\delta^2 \mathbf{h}_i, \delta \mathbf{h}_i\right) \tau + \sum_{i=1}^{j} \left\|\nabla \times \delta \mathbf{h}_i\right\|^2 \tau + \sum_{i=1}^{j} \left(\mathcal{K}_0 \star \delta \mathbf{h}_i, \nabla \times \delta \mathbf{h}_i\right) \tau = \sum_{i=1}^{j} \left(\delta \mathbf{F}_i, \delta \mathbf{h}_i\right) \tau.$$

Next, we follow the same way as in (i) when considering $\delta \mathbf{h}_i$ instead of $\mathbf{h}_i$. □

4.3.2 Convergence

The existence of a weak solution is proved using Rothe's method. The following piecewise linear in time vector fields $\mathbf{H}_n : [0, T] \to \mathbf{L}^2(\Omega)$

$$\begin{aligned} \mathbf{H}_n(0) &= \mathbf{H}_0 \\ \mathbf{H}_n(t) &= \mathbf{h}_{i-1} + (t - t_{i-1})\delta \mathbf{h}_i \qquad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \ldots, n \end{aligned}$$

and the piecewise constant in time fields $\overline{\mathbf{H}}_n : [0, T] \to \mathbf{L}^2(\Omega)$ are introduced

$$\overline{\mathbf{H}}_n(0) = \mathbf{H}_0, \qquad \overline{\mathbf{H}}_n(t) = \mathbf{h}_i, \qquad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \ldots, n.$$

Similarly, the vector field $\overline{\mathbf{F}}_n$ is defined.

The variational formulation (4.6) can be rewritten for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ and a.a. $t \in (0, T)$ as

$$\left(\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}\right) + \left(\nabla \times \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) + \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right). \tag{4.7}$$

Now, the convergence of the sequences $\mathbf{H}_n$ and $\overline{\mathbf{H}}_n$ to the unique weak solution of (4.1) is proved if $\tau \to 0$ or $n \to \infty$.

Theorem 4.3.2 (Existence)**.** *Let* $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl}; \Omega)$ *and* $\mathbf{F} \in \mathrm{L}^2\left((0, T), \mathbf{L}^2(\Omega)\right)$*. Assume that* $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *for any* $t \in [0, T]$*. Then there exists a vector field* $\mathbf{H} \in \mathrm{C}\left([0, T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0, T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right)$ *with* $\partial_t \mathbf{H} \in \mathrm{L}^2\left((0, T), \mathbf{L}^2(\Omega)\right)$*, which solves* (4.5)*.*

Proof. First, let us integrate (4.7) in time to get (for any $\eta \in (0, T)$)

$$\int_0^\eta \left(\partial_t \mathbf{H}_n, \boldsymbol{\varphi}\right) + \int_0^\eta \left(\nabla \times \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right) + \int_0^\eta \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right) = \int_0^\eta \left(\overline{\mathbf{F}}_n, \boldsymbol{\varphi}\right). \tag{4.8}$$

We want to find the limit of (4.8) as $n \to \infty$. By the reflexivity of the space $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$, we have the existence of a subsequence of $\overline{\mathbf{F}}_n$, which we denote with the same symbol again, such that $\overline{\mathbf{F}}_n \rightharpoonup \mathbf{F}$ in $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$. Lemma 4.3.1(i) implies that the corresponding sequence $\{\overline{\mathbf{H}}_n\}$ is uniformly bounded in $\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right)$. Due to the reflexivity of this space, the sequence $\{\overline{\mathbf{H}}_n\}$ has a weak convergent subsequence, which we denote with the same symbol again, i.e.

$$\overline{\mathbf{H}}_n \rightharpoonup \mathbf{H} \qquad \text{in } \mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right).$$

Note that both terms $\int_0^\eta \left(\nabla\times\overline{\mathbf{H}}_n, \nabla\times\boldsymbol{\varphi}\right)$ and $\int_0^\eta \left(\mathcal{K}_0\star\overline{\mathbf{H}}_n, \nabla\times\boldsymbol{\varphi}\right)$ are linear bounded functionals in the space $\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right)$. An easy calculation gives

$$\left\|\mathbf{H}_n(t)-\overline{\mathbf{H}}_n(t)\right\| \leqslant \|\mathbf{h}_i-\mathbf{h}_{i-1}\| \qquad \text{for } t\in[t_{i-1},t_i].$$

By Lemma 4.3.1(i), we get that

$$\lim_{n\to\infty}\left\|\mathbf{H}_n(t)-\overline{\mathbf{H}}_n(t)\right\| = 0, \qquad \text{for all } t\in[0,T]. \tag{4.9}$$

It follows that $\mathbf{H}_n$ and $\overline{\mathbf{H}}_n$ have the same limit in $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$ (if it exists).

The convergence of Rothe's method is based on a compactness argument. Using Theorem 2.9.37, we see that

$$\mathbf{H}^{\frac{1}{2}}(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega).$$

Lemma 4.3.1(i-iii) and the assumption $\mathbf{H}_0\in\mathbf{H}(\mathbf{curl};\Omega)$ give for all $n\in\mathbb{N}$ that

$$\int_0^T \|\mathbf{H}_n(t)\|^2_{\mathbf{H}_0(\mathbf{curl};\Omega)}\,\mathrm{d}t \leqslant C, \quad \nabla\cdot\mathbf{H}_n(t)=0, \quad \forall t\in[0,T].$$

Reviewing Theorem 2.9.36, we see that

$$\int_0^T \|\mathbf{H}_n(t)\|^2_{\mathbf{H}^{\frac{1}{2}}(\Omega)}\,\mathrm{d}t \leqslant C.$$

Taking into account the fact that

$$\int_0^T \|\partial_t\mathbf{H}_n(t)\|^2\,\mathrm{d}t \leqslant C$$

and using the generalized Aubin-Lions lemma 2.12.4 with $V=\mathbf{H}^{\frac{1}{2}}(\Omega), Y=W=\mathbf{L}^2(\Omega)$, we get that $\{\mathbf{H}_n\}$ is compact in the space $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$. Therefore, there exists a subsequence of $\{\mathbf{H}_n\}$ (denoted by the same symbol again) for which we have by (4.9) that

$$\mathbf{H}_n(\mathbf{x},t)\to\mathbf{H}(\mathbf{x},t) \qquad \text{for a.a. } (\mathbf{x},t)\in Q_T.$$

Moreover, $\mathbf{H} \in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$, $\mathbf{H}_n \rightharpoonup \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right)$ and $\partial_t \mathbf{H}_n \rightharpoonup \partial_t \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

Now, we can pass to the limit for $n \to \infty$ in (4.8) to arrive at

$$\int_0^\eta (\partial_t \mathbf{H}, \boldsymbol{\varphi}) + \int_0^\eta (\nabla \times \mathbf{H}, \nabla \times \boldsymbol{\varphi}) + \int_0^\eta (\mathcal{K}_0 \star \mathbf{H}, \nabla \times \boldsymbol{\varphi}) = \int_0^\eta (\mathbf{F}, \boldsymbol{\varphi}). \quad (4.10)$$

This is valid for all $\eta \in (0,T)$. Differentiating the result of (4.10) with respect to the time variable, we get the existence of a solution to (4.5). Taking into account the uniqueness of a solution and Lemma 2.4.20, it is clear that the whole Rothe's sequence $\{\mathbf{H}_n\}$ converges in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ towards the solution. $\square$

Remark 4.3.1. *Instead of using the generalized Aubin-Lions lemma, one can also use Kačur's lemma 2.12.3 because it holds that*

$$\max_{t\in[0,T]} \left\|\overline{\mathbf{H}}_n(t)\right\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} \leqslant C \quad \textit{and} \quad \partial_t \mathbf{H}_n \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$$

if $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$. *An analogous result as in the previous theorem can be obtained. Then, the convergence result* $\mathbf{H}_n \to \mathbf{H}$ *is in the space* $\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$ *instead of* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

Remark 4.3.2. *The existence result can also be obtained by using the test space*

$$\mathcal{W} := \{\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}(\mathbf{div};\Omega) \,:\, \nabla \cdot \boldsymbol{\varphi} = 0 \textit{ in } \Omega\}$$

in the variational formulation (4.5)*. The use of this space is allowed because the solution of problem* (4.5) *is divergence free. The space* $\mathcal{W}$ *is compact in* $\mathbf{L}^2(\Omega)$ *[39, Corollary 3.49], i.e. we have that*

$$\mathcal{W} \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega) \cong (\mathbf{L}^2(\Omega))^* \hookrightarrow\hookrightarrow \mathcal{W}^*.$$

Using these spaces, one can show the existence of a solution under the lower regularity assumption that $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$*. An analogous result as in Lemma 4.3.1(i) and (ii) can be obtained replacing* $\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)$ *by* $\mathcal{W}^*$*. Then, it is clear that*

$$\max_{t\in[0,T]} \left\|\overline{\mathbf{H}}_n(t)\right\|^2 + \sum_{i=1}^n \left\| \int_{t_{i-1}}^{t_i} \partial_t \mathbf{H}_n(s)\, \mathrm{d}s \right\|^2 + \int_0^T \left\|\nabla \times \overline{\mathbf{H}}_n(t)\right\|^2 \mathrm{d}t \leqslant C \quad (4.11)$$

and

$$\int_0^T \|\partial_t \mathbf{H}_n(s)\|_{W^*}^2 \, \mathrm{d}s \leqslant C, \quad \forall n \in \mathbb{N}.$$

According to Kačur's lemma 2.12.3 with $V = \mathbf{L}^2(\Omega)$ *and* $Y = \mathcal{W}^*$*, there exist a field* $\mathbf{H} \in \mathrm{C}([0,T], \mathcal{W}^*) \cap \mathrm{L}^\infty\left((0,T), \mathbf{L}^2(\Omega)\right)$ *solving* (4.5) *with* $\partial_t \mathbf{H} \in \mathrm{L}^2((0,T), \mathcal{W}^*)$ *and a subsequence* $\{\mathbf{H}_n\}$ *of* $\{\mathbf{H}_n\}$ *such that*

$$\begin{cases} \mathbf{H}_n \to \mathbf{H}, & \textit{in} \quad \mathrm{C}([0,T], \mathcal{W}^*), \\ \mathbf{H}_n(t) \rightharpoonup \mathbf{H}(t), & \textit{in} \quad \mathbf{L}^2(\Omega) \textit{ for all } t \in [0,T], \\ \overline{\mathbf{H}}_n \rightharpoonup \mathbf{H}, & \textit{in} \quad \mathrm{L}^2((0,T), \mathcal{W}), \\ \partial_t \mathbf{H}_n \rightharpoonup \partial_t \mathbf{H}, & \textit{in} \quad \mathrm{L}^2((0,T), \mathcal{W}^*). \end{cases}$$

This result is sufficient to prove the existence of a weak solution to problem (4.5)*. However, no error estimates can be obtained when* $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$*. Note that also* $\mathbf{H}_n \to \mathbf{H}$ *in* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *as* $n \to +\infty$*. This follows from* $\mathbf{H}_n \rightharpoonup \mathbf{H}$ *in* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and* $\|\mathbf{H}_n\|_{\mathrm{L}^2((0,T),\mathbf{L}^2(\Omega))} \to \|\mathbf{H}\|_{\mathrm{L}^2((0,T),\mathbf{L}^2(\Omega))}$ *due to Lemma 2.4.19.*

Let us check that $\|\mathbf{H}_n\|_{\mathrm{L}^2((0,T),\mathbf{L}^2(\Omega))} \to \|\mathbf{H}\|_{\mathrm{L}^2((0,T),\mathbf{L}^2(\Omega))}$ *as* $n \to \infty$*. The duality pairing* $\langle \cdot, \cdot \rangle_{\mathcal{W}^* \times \mathcal{W}}$ *can be considered as a continuous extension of the inner product on* $\mathbf{L}^2(\Omega)$*, see Theorem 2.9.15. This implies that*

$$\int_0^T \|\mathbf{H}_n(t)\|^2 \, \mathrm{d}t = \int_0^T \left(\mathbf{H}_n(t), \mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t)\right) \mathrm{d}t + \int_0^T \langle \mathbf{H}_n(t), \overline{\mathbf{H}}_n(t) \rangle_{\mathcal{W}^* \times \mathcal{W}}. \tag{4.13}$$

From (4.11)*, it follows that*

$$\left| \int_0^T \left(\mathbf{H}_n(t), \mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t)\right) \mathrm{d}t \right| \lesssim \sqrt{\tau}.$$

Hence, passing to the limit $\tau \to 0$ *in* (4.13) *implies that*

$$\int_0^T \|\mathbf{H}_n(t)\|^2 \, \mathrm{d}t \to \int_0^T \|\mathbf{H}(t)\|^2 \, \mathrm{d}t$$

because $\overline{\mathbf{H}}_n \rightharpoonup \mathbf{H}$ *in* $\mathrm{L}^2((0,T), \mathcal{W})$ *and* $\mathbf{H}_n \to \mathbf{H}$ *in* $\mathrm{C}([0,T], \mathcal{W}^*)$.

4.3.3 Error estimates

The following theorem addresses the error estimates for the time discretization, i.e. the error estimates between the unique solution of the original problem and the solution of Rothe's problem.

Theorem 4.3.3 (Error)**.** *Suppose that* $\mathbf{F} \in \mathrm{Lip}([0,T], \mathbf{L}^2(\Omega))$.

(i) If $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl}; \Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \int_0^T \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 \leqslant C\tau.$$

(ii) If $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega),\ \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ *and* $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \int_0^T \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 \leqslant C\tau^2.$$

Please note that the positive constant C *in these estimates is of the form* Ce^{CT}.

Proof. We subtract (4.5) from (4.7), set $\boldsymbol{\varphi} = \mathbf{H}_n(t) - \mathbf{H}(t)$ and integrate in time over $t \in (0, \eta) \subset (0, T)$ to get

$$\begin{aligned}
\tfrac{1}{2} \|\mathbf{H}_n(\eta) - \mathbf{H}(\eta)\|^2 &+ \int_0^\eta \|\nabla \times \mathbf{H}_n - \nabla \times \mathbf{H}\|^2 \\
&\qquad + \int_0^\eta (\mathcal{K}_0 \star [\mathbf{H}_n - \mathbf{H}], \nabla \times [\mathbf{H}_n - \mathbf{H}]) \\
= \int_0^\eta \left(\overline{\mathbf{F}}_n - \mathbf{F}, \mathbf{H}_n - \mathbf{H}\right) &+ \int_0^\eta \left(\nabla \times [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n - \mathbf{H}]\right) \\
&\qquad + \int_0^\eta \left(\mathcal{K}_0 \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n - \mathbf{H}]\right). \qquad (4.14)
\end{aligned}$$

Due to the Lipschitz continuity of $\mathbf{F}$, we may write that

$$\begin{aligned}
\left|\int_0^\eta \left(\overline{\mathbf{F}}_n - \mathbf{F}, \mathbf{H}_n - \mathbf{H}\right)\right| &\lesssim \int_0^\eta \left\|\overline{\mathbf{F}}_n - \mathbf{F}\right\|^2 + \int_0^\eta \|\mathbf{H}_n - \mathbf{H}\|^2 \\
&\leqslant C\tau^2 + \int_0^\eta \|\mathbf{H}_n - \mathbf{H}\|^2.
\end{aligned}$$

It holds that

$$\left\|\mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t)\right\| \leqslant \tau \|\partial_t \mathbf{H}_n(t)\| \qquad \text{for } t \in [0, T].$$

The last term of (4.14) can be estimated using (4.4) and Lemma 4.3.1(iii) as follows

$$\begin{aligned}
&\left|\int_0^\eta \left(\mathcal{K}_0 \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n - \mathbf{H}]\right)\right| \\
&\overset{(4.4)}{\leqslant} \quad \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \int_0^\eta \left\|\mathbf{H}_n - \overline{\mathbf{H}}_n\right\|^2 \\
&\leqslant \quad \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \tau^2.
\end{aligned}$$

Analogously, we have that

$$\left| \int_0^\eta (\mathcal{K}_0 \star [\mathbf{H}_n - \mathbf{H}], \nabla \times [\mathbf{H}_n - \mathbf{H}]) \right| \leqslant \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \int_0^\eta \|\mathbf{H}_n - \mathbf{H}\|^2 . \quad (4.15)$$

It remains to estimate the second term on the RHS in (4.14). Here, we have to distinguish between two cases depending on the a priori estimates we have (see Lemma 4.3.1(iii) and (iv)):

(i)

$$\begin{aligned} \left| \int_0^\eta \left(\nabla \times [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n - \mathbf{H}]\right) \right| &\leqslant \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla \times [\mathbf{H}_n - \overline{\mathbf{H}}_n]\right\|^2 \\ &\leqslant \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \tau, \end{aligned}$$

(ii)

$$\begin{aligned} \left| \int_0^\eta \left(\nabla \times [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n - \mathbf{H}]\right) \right| &\leqslant \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla \times [\mathbf{H}_n - \overline{\mathbf{H}}_n]\right\|^2 \\ &\leqslant \varepsilon \int_0^\eta \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 + C_\varepsilon \tau^2. \end{aligned}$$

Putting all things together, choosing a sufficiently small positive ε and applying Grönwall's argument, we conclude the proof. □

4.4 Modified scheme

In this section, the following time-discrete scheme that represents a slight modification of (4.6) is considered for $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl}; \Omega)$:

$$\left\{ \begin{aligned} (\delta \mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) &= (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi}), \\ \mathbf{h}_0 &= \mathbf{H}_0. \end{aligned} \right. \quad (4.16)$$

This scheme is equivalent to

$$\begin{aligned} a(\mathbf{h}_i, \boldsymbol{\varphi}) := \left(\frac{\mathbf{h}_i}{\tau}, \boldsymbol{\varphi}\right) + (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) \\ = (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi}) + \left(\frac{\mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

Here, the convolution term is taken explicitly (from the last time step), while (4.6) considers an implicit form (from the actual time step). In other words, the scheme (4.16) is semi-implicit. The bilinear form $a(\cdot,\cdot)$ is elliptic and continuous in $\mathbf{H}_0(\mathbf{curl};\Omega)$. According to (4.3), the functional $f_i(\cdot)$ is linear and bounded in $\mathbf{H}_0(\mathbf{curl};\Omega)$ if $\mathbf{H}_{i-1} \in \mathbf{L}^2(\Omega)$. Thus, if $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$, an application of the Lax-Milgram lemma 2.11.1 gives the existence of a unique solution to (4.16) for any $i = 1, \dots, n$ and any $\tau > 0$.

Handling this scheme is very similar to the way used for (4.6). For brevity, only the differences between both algorithms are pointed out.

Lemma 4.4.1 (A priori estimates)**.** *Suppose that* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

(i) Let $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$. *Then, there exists a positive constant* C *such that*

$$\max_{1\leqslant i\leqslant n} \|\mathbf{h}_i\|^2 + \sum_{i=1}^{n} \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^{n} \|\nabla\times\mathbf{h}_i\|^2 \tau \leqslant C.$$

(ii) If $\nabla\cdot\mathbf{H}_0 = 0 = \nabla\cdot\mathbf{F}_i$ *for* $i = 1, \dots, n$, *then* $\nabla\cdot\mathbf{h}_i = 0$ *for all* $i = 1, \dots, n$. *Furthermore, we have that*

$$\tau \sum_{i=1}^{n} \|\delta\mathbf{h}_i\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C.$$

(iii) If $\nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n} \|\nabla\times\mathbf{h}_i\|^2 + \sum_{i=1}^{n} \|\nabla\times\mathbf{h}_i - \nabla\times\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^{n} \|\delta\mathbf{h}_i\|^2 \tau \leqslant C.$$

(iv) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$, $\partial_t\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, $\nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n} \|\delta\mathbf{h}_i\|^2 + \sum_{i=1}^{n} \|\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^{n} \|\nabla\times\delta\mathbf{h}_i\|^2 \tau \leqslant C.$$

Proof. (i) We follow Lemma 4.3.1(i). Using (4.4), we have that

$$\left|\sum_{i=1}^{j} (\mathcal{K}_0\star\mathbf{h}_{i-1}, \nabla\times\mathbf{h}_i)\,\tau\right| \leqslant \varepsilon \sum_{i=1}^{j} \|\nabla\times\mathbf{h}_i\|^2 \tau + C_\varepsilon \sum_{i=0}^{j} \|\mathbf{h}_i\|^2 \tau.$$

After fixing a sufficiently small positive ε, an application of Grönwall's lemma completes the proof.

(ii) The proof is the same as in Lemma 4.3.1(ii) replacing $(\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})$ by $(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi})$.

(iii) The verification is the same as in Lemma 4.3.1(iii) replacing $(\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})$ by $(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi})$. Remark that

$$\sum_{i=1}^{j} (\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \delta \mathbf{h}_i)\,\tau = (\mathcal{K}_0 \star \mathbf{h}_j, \nabla \times \mathbf{h}_j) - (\mathcal{K}_0 \star \mathbf{h}_0, \nabla \times \mathbf{h}_0) - \sum_{i=1}^{j} (\mathcal{K}_0 \star \delta \mathbf{h}_i, \nabla \times \mathbf{h}_i)\,\tau.$$

(iv) We set

$$\delta \mathbf{h}_0 := \partial_t \mathbf{H}(0) = \mathbf{F}(0) - \nabla \times \nabla \times \mathbf{H}_0 - \nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \quad \text{a.e. in } \Omega,$$

and

$$\mathbf{h}_{-1} := \mathbf{h}_0 - \delta \mathbf{h}_0 \tau \quad \text{a.e. in } \Omega.$$

Please note that $\delta \mathbf{h}_0$ and $\mathbf{h}_{-1} \in \mathbf{L}^2(\Omega)$. The proof follows very closely the proof of Lemma 4.3.1(iv), except for the appearance of the term $(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi})$ instead of $(\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})$. $\square$

The variational formulation (4.16) can be rewritten for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ and a.a. $t \in (0, T)$ as

$$(\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}) + \left(\nabla \times \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right) - \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n(t-\tau), \nabla \times \boldsymbol{\varphi}\right).$$

The existence theorem 4.3.2 stays valid. The next theorem derives the error estimates for the scheme (4.16).

Theorem 4.4.1 (Error)**.** *Suppose that* $\mathbf{F} \in \operatorname{Lip}([0,T], \mathbf{L}^2(\Omega))$.

(i) If $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl}; \Omega)$ *then*

$$\max_{t \in [0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \int_0^T \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 \leqslant C\tau.$$

(ii) If $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega),\ \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ *and* $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t \in [0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \int_0^T \|\nabla \times [\mathbf{H}_n - \mathbf{H}]\|^2 \leqslant C\tau^2.$$

Please note that the positive constant C *in these estimates is of the form* Ce^{CT}.

Proof. We follow Theorem 4.3.3. We get (4.14) in which the term

$$\int_0^t \left(\mathcal{K}_0 \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n - \mathbf{H}]\right)$$

is replaced by

$$\int_0^t \left(\mathcal{K}_0 \star [\mathbf{H}_n(s) - \overline{\mathbf{H}}_n(s-\tau)], \nabla \times [\mathbf{H}_n(s) - \mathbf{H}(s)]\right) \, \mathrm{d}s.$$

This can be handled using (4.4) and Lemma 4.4.1(iii) as follows

$$\begin{aligned}
&\left|\int_0^t \left(\mathcal{K}_0 \star [\mathbf{H}_n(s) - \overline{\mathbf{H}}_n(s-\tau)], \nabla \times [\mathbf{H}_n(s) - \mathbf{H}(s)]\right) \, \mathrm{d}s\right| \\
&\quad \leqslant \quad \varepsilon \int_0^t \left\|\nabla \times [\mathbf{H}_n(s) - \mathbf{H}(s)]\right\|^2 \, \mathrm{d}s + C_\varepsilon \int_0^t \left\|\mathbf{H}_n(s) - \overline{\mathbf{H}}_n(s-\tau)\right\|^2 \, \mathrm{d}s \\
&\quad \leqslant \quad \varepsilon \int_0^t \left\|\nabla \times [\mathbf{H}_n(s) - \mathbf{H}(s)]\right\|^2 \, \mathrm{d}s + C_\varepsilon \tau^2.
\end{aligned}$$

The rest of the proof is the same as in Theorem 4.3.3. □

This second scheme is considered, because it is easier to implement than the first scheme and it gives the same order of convergence. Moreover, the finite element matrix corresponding with the LHS of (4.26) is sparser and hence less memory is needed.

4.5 Higher regularity

The problem (4.1) is nonsymmetric due to the convolution term. This means that the term $(\mathcal{K}_0 \star \mathbf{H}, \nabla \times \mathbf{H})$ in the variational formulation consists of two terms with derivatives of a different order. The unique solution of problem (4.1) can be approached by the scheme (4.6) or (4.16). Theorems 4.3.3 and 4.4.1 claim to have optimal convergence rates $\mathcal{O}(\tau)$ in the space

$$\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathbf{H}(\mathbf{curl}; \Omega)\right)$$

under appropriate conditions. These error estimates have been obtained using a priori estimates, which were based on Grönwall's argument. In fact, it is the nonsymmetric term who leads to the use of Grönwall's lemma. Therefore, $\mathcal{O}(\tau) = e^{CT}\tau$, which means that the constant e^{CT} might be large.

The exponential (in time) growth character of this constant can be overcome by making a symmetrification of the problem when using an analogue of the implicit scheme (4.6). This is done by incorporation of the curl operator $\nabla \times \mathbf{J}_s$ into a new convolution kernel, cf. [89] and [2, §11.7]. Then, the term $(\mathcal{K}_0 \star \mathbf{H}, \nabla \times \mathbf{H})$ can be replaced by $(\mathcal{K} \star \mathbf{H}, \mathbf{H})$ in the variational formulation, where $\mathcal{K}$ is defined in the following lemma. The time dependency of $\mathbf{H}$ is ignored in the following lemma and subsection.

Lemma 4.5.1. *Let $\mathbf{J}_s$ be defined as in* (3.9)*. Suppose that $\nabla \cdot \mathbf{H} = 0$ in Ω and $\mathbf{H} \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$. Then,*

$$\nabla \times \mathbf{J}_s(\mathbf{x}) = -\int_\Omega \mathcal{K}(\mathbf{x}, \mathbf{x}')\mathbf{H}(\mathbf{x}')\ \mathrm{d}\mathbf{x}' =: -\left(\mathcal{K} \star \mathbf{H}\right)(\mathbf{x}), \qquad \mathbf{x} \in \Omega,$$

where the kernel $\mathcal{K}$ is defined by

$$\mathcal{K} : \Omega \times \Omega \to \mathbb{R} : (\mathbf{x}, \mathbf{x}') \mapsto \kappa(|\mathbf{x} - \mathbf{x}'|), \tag{4.17}$$

with

$$\kappa : (0, \infty) \to \mathbb{R} : s \mapsto \begin{cases} \frac{\widetilde{C}}{2s^2}\left(1 - \frac{s}{r_0}\right)\exp\left(-\frac{s}{r_0}\right) & s < r_0; \\ 0 & s \geqslant r_0. \end{cases}$$

Proof. The proof is adapted from [89]. Let $\mathbf{r} = \mathbf{x} - \mathbf{x}'$ and $r = |\mathbf{r}|_{\mathrm{e}}$. The first component is given by

$$\begin{aligned}\left(\nabla \times \mathbf{J}_s\right)_1(\mathbf{x}) = \int_\Omega &[H_2(\mathbf{x}')\partial_{x_2}\left(\sigma_0(r)r_1\right) + H_3(\mathbf{x}')\partial_{x_3}\left(\sigma_0(r)r_1\right) \\ &-H_1(\mathbf{x}')\partial_{x_2}\left(\sigma_0(r)r_2\right) - H_1(\mathbf{x}')\partial_{x_3}\left(\sigma_0(r)r_3\right)]\ \mathrm{d}\mathbf{x}'.\end{aligned}$$

Notice that $\partial_{x_i}\sigma_0(r) = \sigma_0'(r)\frac{r_i}{r}$ and $\partial_{x_i'}\sigma_0(r) = -\sigma_0'(r)\frac{r_i}{r}$. A simple calculation gives that $\partial_{x_j} = -\partial_{x_j'}$ on the term $\sigma_0(r)r_i$. We use this on the first two terms of the integrand together with the product rule

$$\partial_{x_j'}\left(\sigma_0(r)r_l H_m(\mathbf{x}')\right) = \sigma_0(r)r_l\partial_{x_j'}H_m(\mathbf{x}') + H_m(\mathbf{x}')\partial_{x_j'}\left(\sigma_0(r)r_l\right).$$

Employing $\nabla \cdot \mathbf{H} = 0$ on the terms corresponding with the first term on the RHS of the product rule gives

$$\begin{aligned}\left(\nabla \times \mathbf{J}_s\right)_1(\mathbf{x}) = \int_\Omega &\left[-\partial_{x_2'}\left(\sigma_0(r)r_1H_2(\mathbf{x}')\right) - \partial_{x_3'}\left(\sigma_0(r)r_1H_3(\mathbf{x}')\right)\right. \\ &\left.-\sigma_0(r)r_1\partial_{x_1'}H_1(\mathbf{x}') - H_1(\mathbf{x}')\partial_{x_2}\left(\sigma_0(r)r_2\right) - H_1(\mathbf{x}')\partial_{x_3}\left(\sigma_0(r)r_3\right)\right]\ \mathrm{d}\mathbf{x}'.\end{aligned}$$

An application of the divergence theorem on the first two terms in the integrand and an integration by parts on the third term in the integrand give that $\left(\nabla \times \mathbf{J}_s\right)_1(\mathbf{x})$ equals

$$\begin{aligned}\int_{\partial\Omega} &[-\sigma_0(r)r_1H_2(\mathbf{x}')\nu_2(\mathbf{x}') - \sigma_0(r)r_1H_3(\mathbf{x}')\nu_3(\mathbf{x}') \\ &-\sigma_0(r)r_1H_1(\mathbf{x}')\nu_1(\mathbf{x}')]\ \mathrm{d}\mathbf{x}' + \int_\Omega \left[H_1(\mathbf{x}')\partial_{x_1'}\left(\sigma_0(r)r_1\right)\right. \\ &\left.-H_1(\mathbf{x}')\partial_{x_2}\left(\sigma_0(r)r_2\right) - H_1(\mathbf{x}')\partial_{x_3}\left(\sigma_0(r)r_3\right)\right]\ \mathrm{d}\mathbf{x}'.\end{aligned}$$

Firstly, we use $\partial_{x'_j} = -\partial_{x_j}$ on the first term in the second integral. Secondly, the surface and volume integrals are combined. Equivalently, we can prove an analogous result for the second and third component of the curl of $\mathbf{J}_s$. Consequently,

$$(\nabla \times \mathbf{J}_s)_i(\mathbf{x}) = -\int_\Omega H_i(\mathbf{x}')\nabla \cdot (\sigma_0(r)\mathbf{r})\ \mathrm{d}\mathbf{x}' - \int_{\partial\Omega} \sigma_0(r) r_i \mathbf{H}(\mathbf{x}') \cdot \boldsymbol{\nu}(\mathbf{x}')\ \mathrm{d}\mathbf{x}',$$

for $i = 1, 2, 3$. From the assumption $\mathbf{H} \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$, we conclude the proof. Moreover,

$$\kappa(r) = \kappa(|\mathbf{r}|) = \nabla \cdot (\sigma_0(r)\mathbf{r}) = \partial_{\mathbf{r}} \cdot (\sigma_0(r)\mathbf{r}) = r\sigma_0'(r) + 3\sigma_0(r).$$

A simple calculation using (3.10) gives the exact form of the kernel $\mathcal{K}$. □

4.5.1 Properties of the kernel $\mathcal{K}$

The goal of this section is to show that the kernel $\mathcal{K}$, defined in (4.17), is positive definite. This new characteristic is useful for simplifying proofs and for avoiding the Grönwall argument. The starting point is a definition of a positive definite kernel in the sense of Mercer [126] and a radial function [127].

Definition 4.5.1. *Let $X \subset \mathbb{R}^d,\ d \geqslant 1$. A symmetric kernel $\hat{\mathcal{K}} : X \times X \to \mathbb{R}$ is called positive definite if*

$$\int_X \left(\int_X \hat{\mathcal{K}}(\mathbf{x}, \mathbf{x}')\mathbf{F}(\mathbf{x}')\ \mathrm{d}\mathbf{x}' \right) \mathbf{F}(\mathbf{x})\ \mathrm{d}\mathbf{x} \geqslant \mathbf{0}, \qquad \forall \mathbf{F} \in \mathbf{L}^1(X). \tag{4.18}$$

Definition 4.5.2. *A function $\Psi : \mathbb{R}^d \to \mathbb{R},\ d \geqslant 1$, is called radial provided there exists a function $\varphi : [0, \infty) \to \mathbb{R}$ such that*

$$\Psi(\mathbf{x}) = \varphi(r) \text{ where } r = |\mathbf{x}|\,.$$

Remark 1. $\mathcal{K}$ is a radial function with $\varphi = \kappa$ and $r = |\mathbf{x} - \mathbf{x}'|$, see Theorem 4.5.1.

The following important lemma is easy to show using spherical coordinates, so its proof is omitted.

Lemma 4.5.2. $\mathcal{K}(\mathbf{x}, \cdot) \in \mathrm{L}^p(\Omega)$ *if* $1 \leqslant p < \frac{3}{2}$, $\forall \mathbf{x} \in \Omega$.

The theory of completely monotone functions is involved to prove that $\mathcal{K}$ is positive definite.

Definition 4.5.3. *A function $\varphi : (0, a) \to \mathbb{R}$ that is an element of $\mathrm{C}(0, a)$ and that satisfies*

$$(-1)^l \varphi^{(l)}(x) \geq 0, \quad x > 0, \quad l = 0, 1, 2, \ldots,$$

is called completely monotone on $(0, a)$. The limit $\varphi^{(l)}(0) = \lim_{x \searrow 0} \varphi^{(l)}(x)$, finite or infinite, exists.

Schoenberg has pointed out the following connection between positive definite radial and completely monotone functions [128, Thm. 3]. A more recent reference is [129].

Theorem 4.5.1 (Schoenberg interpolation theorem)**.** *A function φ is completely monotone on $(0,\infty)$ if and only if $\Psi = \varphi(|\cdot|^2)$ is positive definite and radial on $\mathbb{R}^d$ for all $d \geqslant 1$.*

Lemma 4.5.3. *The kernel $\mathcal{K}$, defined in Lemma 4.5.1, is positive definite on $\Omega \times \Omega \subset \mathbb{R}^d \times \mathbb{R}^d$.*

Proof. One can proof by induction that the following functions are completely monotone on $(0,\infty)$:

- $\hat{\kappa}_1(r) = \frac{C}{2}, \quad C \geqslant 0$;
- $\hat{\kappa}_2(r) = \frac{1}{r}$, since for $l = 0,1,2,\ldots$
$$(-1)^l \hat{\kappa}_2^{(l)}(r) = (-1)^{2l} l! r^{-(l+1)} \geqslant 0, \quad r > 0;$$
- $\hat{\kappa}_3(r) = \exp\left(-\frac{\sqrt{r}}{r_0}\right) \geq 0$, since for $l = 1,2,\ldots$
$$(-1)^l \hat{\kappa}_3^{(l)}(r) = (-1)^{2l} 2^{-l} \exp\left(-\frac{\sqrt{r}}{r_0}\right) \sum_{i=0}^{l-1} C_i^{(l)} r_0^{i-l} r^{-\frac{l+i}{2}} \geqslant 0, \quad r > 0,$$
with
$$C_i^{(l)} = \begin{cases} C_i^{(l-1)} & i = 0; \\ C_i^{(l-1)} + C_{i-1}^{(l-1)}(l+i-2) & 1 \leqslant i \leqslant l-1; \\ C_{i-1}^{(l-1)}(l+i-2) & i = l-1; \end{cases}$$
and
$$C_0^{(1)} = 1;$$
- $\hat{\kappa}_4(r) = \begin{cases} 1 - \frac{\sqrt{r}}{r_0} & r < r_0^2; \\ 0 & r \geqslant r_0^2; \end{cases}$ since for $l = 1,2,\ldots$
$$(-1)^l \hat{\kappa}_4^{(l)}(r) = \begin{cases} (-1)^{2l} \frac{\Gamma\left(l-\frac{1}{2}\right)}{2\sqrt{\pi} r_0 r^{l-\frac{1}{2}}} \geq 0 & 0 < r < r_0^2; \\ 0 & r \geqslant r_0^2; \end{cases}$$
where Γ denotes the gamma function.

From the previous calculations, it follows that the function
$$\hat{\kappa}(r) := \hat{\kappa}_1(r)\hat{\kappa}_2(r)\hat{\kappa}_3(r)\hat{\kappa}_4(r)$$
is completely monotone on $(0,\infty)$ (the product of two completely monotone functions is completely monotone, cf. [130]). The Schoenberg interpolation Theorem 4.5.1 tells us that $\mathcal{K}(\mathbf{x},\mathbf{x}') = \hat{\kappa}(|\mathbf{x}-\mathbf{x}'|^2)$ is positive definite and radial on $\mathbb{R}^d \times \mathbb{R}^d$ for all d. $\square$

4.5.2 Symmetric problem

The solution of problem (4.1) is divergence free for any $t \in [0,T]$ if $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ for any $t \in [0,T]$, see Theorem 4.2.1. From now on, it is assumed that $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$ for any $t \in [0,T]$. This is a natural condition due to (3.3) and the assumption that the magnetic field outside the domain equals zero [39, p. 8]. Note that $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$, $\nabla \cdot \mathbf{H}_0 = 0$ and $\mathbf{H}_0 \cdot \boldsymbol{\nu} = \mathbf{0}$ on $\partial\Omega$ imply that

$$\mathbf{H}_0 \in \{\boldsymbol{\varphi} \in \mathbf{H}(\mathbf{curl};\Omega) \cap \mathbf{H}_0(\mathbf{div};\Omega) \,:\, \nabla \cdot \boldsymbol{\varphi} = 0\}\,.$$

Using the identity

$$-\Delta \mathbf{H} = \nabla \times (\nabla \times \mathbf{H}) - \nabla(\nabla \cdot \mathbf{H})$$

and Lemma 4.5.1, the solution of problem (4.1) also satisfies

$$\begin{cases} \partial_t \mathbf{H} - \Delta \mathbf{H} + \mathcal{K} \star \mathbf{H} = \mathbf{F} & \text{in } Q_T, \\ \mathbf{H} = \mathbf{0} & \text{on } \partial\Omega \times (0,T), \\ \mathbf{H}(\mathbf{x},0) = \mathbf{H}_0 & \text{in } \Omega, \\ \nabla \cdot \mathbf{H}_0 = \mathbf{0} & \text{in } \Omega. \end{cases} \tag{4.19}$$

This problem is a vector Laplace equation. Therefore, it is natural that the further analysis takes place in the Hilbert spaces $\mathbf{H}^1(\Omega)$ and $\mathbf{H}^2(\Omega)$. First, a useful inequality is stated, which follows from Lemma 4.5.2, namely

$$|\left(\mathcal{K} \star \mathbf{H}\right)(\mathbf{x})| = \left| \int_\Omega \mathcal{K}(\mathbf{x},\mathbf{x}')\mathbf{H}(\mathbf{x}')\ \mathrm{d}\mathbf{x}' \right| \leqslant C(q) \left\|\mathbf{H}\right\|_q, \quad \forall \mathbf{x} \in \Omega, \tag{4.20}$$

for all $q > 3$. Due to the Sobolev embedding theorem 2.9.18 in $\mathbb{R}^3$, it holds that $\mathbf{H}_0^1(\Omega) \hookrightarrow \mathbf{L}^6(\Omega)$. This, together with the Friedrichs inequality, see Theorem 2.9.28, gives for all $\mathbf{h}_1 \in \mathbf{H}_0^1(\Omega)$ and $\mathbf{h}_2 \in \mathbf{L}^2(\Omega)$ that

$$(\mathcal{K} \star \mathbf{h}_1, \mathbf{h}_2) \leqslant C_\varepsilon \left\|\mathbf{h}_1\right\|^2_{\mathbf{H}_0^1(\Omega)} + \varepsilon \left\|\mathbf{h}_2\right\|^2 \leqslant C_\varepsilon \left\|\nabla \mathbf{h}_1\right\|^2 + \varepsilon \left\|\mathbf{h}_2\right\|^2. \tag{4.21}$$

The position of ε and C_ε can be switched. The positive definiteness of $\mathcal{K}$ implies that

$$(\mathcal{K} \star \mathbf{h}, \mathbf{h}) \geqslant 0, \qquad \forall \mathbf{h} \in \mathbf{H}_0^1(\Omega).$$

Under the additional assumption that $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$ for any $t \in [0,T]$, the solution to problem (4.1) obeys for all $\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$ that

$$(\partial_t \mathbf{H}(t), \boldsymbol{\varphi}) + (\nabla \mathbf{H}(t), \nabla \boldsymbol{\varphi}) + ((\mathcal{K} \star \mathbf{H})(t), \boldsymbol{\varphi}) = (\mathbf{F}(t), \boldsymbol{\varphi})\,. \tag{4.22}$$

The following theorem is analogous to Theorem 4.2.1.

Theorem 4.5.2 (Enhanced stability)**.** *Assume that $\nabla \cdot \mathbf{F}(t) = 0 = \nabla \cdot \mathbf{H}_0$ for any $t \in [0,T]$, $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ and $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$ for any $t \in [0,T]$. The solution to problem* (4.1) *obeys* *(i) If* $\mathbf{H}_0 \in \mathbf{H}^1(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}(t)\|^2_{\mathbf{H}^1(\Omega)} + \int_0^T \|\partial_t \mathbf{H}\|^2 \leqslant C.$$

(ii) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$, $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and* $\mathbf{H}_0 \in \mathbf{H}^1_0(\Omega) \cap \mathbf{H}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\partial_t \mathbf{H}(t)\|^2 + \int_0^T \|\nabla \partial_t \mathbf{H}\|^2 \leqslant C.$$

Proof. The proof is straightforward if we follow the line from Theorem 4.2.1. Please note that we employ the positive definiteness of the convolution kernel $\mathcal{K}$ in order to avoid the use of Grönwall's lemma. We only point out how to handle the differences.

(i) We follow Theorem 4.2.1(i) and (iii). Note that for each $\xi \in (0,T)$, it holds that

$$\int_0^\xi (\mathcal{K} \star \mathbf{H}, \mathbf{H}) \geqslant 0.$$

It is not possible to estimate this term by using (4.21) without needing Grönwall's lemma afterwards. Thanks to the Friedrich's inequality, we get that

$$\int_0^\xi (\mathbf{F}, \mathbf{H}) \leqslant C_\varepsilon \int_0^\xi \|\mathbf{F}\|^2 + \varepsilon \int_0^\xi \|\nabla \mathbf{H}\|^2.$$

From this, it is easy to see that

$$\max_{t\in[0,T]} \|\mathbf{H}(t)\|^2 + \int_0^T \|\nabla \mathbf{H}\|^2 \leqslant C.$$

According to (4.21) and this estimate, we successively deduce that

$$\int_0^\xi (\mathcal{K} \star \mathbf{H}, \partial_t \mathbf{H}) \leqslant \varepsilon \int_0^\xi \|\partial_t \mathbf{H}\|^2 + C_\varepsilon \int_0^\xi \|\nabla \mathbf{H}\|^2 \leqslant \varepsilon \int_0^\xi \|\partial_t \mathbf{H}\|^2 + C_\varepsilon.$$

(ii) Note that $\displaystyle\int_0^\xi (\mathcal{K} \star \partial_t \mathbf{H}, \partial_t \mathbf{H}) \geqslant 0$ for each $\xi \in (0,T)$. Analogously as in Theorem 4.2.1(iv), we arrive at

$$\max_{t\in[0,T]} \|\partial_t \mathbf{H}(t)\|^2 + \int_0^T \|\nabla \partial_t \mathbf{H}\|^2 \lesssim 1 + \|\partial_t \mathbf{H}(0)\|^2.$$

Taking into account (4.20) and $\mathbf{H}^1_0(\Omega) \hookrightarrow \mathbf{L}^6(\Omega)$, we get $\|\mathcal{K} \star \mathbf{H}_0\| \leqslant C$. Therefore, it holds that

$$\partial_t \mathbf{H}(0) = \mathbf{F}(0) - \Delta \mathbf{H}_0 - \mathcal{K} \star \mathbf{H}_0 \quad \text{a.e. in } \Omega,$$

if $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$. Hence,

$$\|\partial_t \mathbf{H}(0)\| \lesssim 1.$$

□

The implicit scheme (4.6) takes the form

$$\left\{ \begin{array}{rl} (\delta \mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) + (\mathcal{K} \star \mathbf{h}_i, \boldsymbol{\varphi}) & = (\mathbf{F}_i, \boldsymbol{\varphi}), \quad \boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega), \\ \mathbf{h}_0 & = \mathbf{H}_0, \end{array} \right. \tag{4.23}$$

which is equivalent to

$$\begin{aligned} a(\mathbf{h}_i, \boldsymbol{\varphi}) &:= \left(\frac{\mathbf{h}_i}{\tau}, \boldsymbol{\varphi} \right) + (\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) + (\mathcal{K} \star \mathbf{h}_i, \boldsymbol{\varphi}) \\ &= (\mathbf{F}_i, \boldsymbol{\varphi}) + \left(\frac{\mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi} \right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

The following lemma is analogous to Lemma 4.3.1. As in Theorem 4.5.2, the application of Grönwall's argument can be avoided by the positive definiteness of the kernel $\mathcal{K}$.

Lemma 4.5.4 (Enhanced a priori estimates)**.** *Assume that* $\nabla \cdot \mathbf{F}_i = 0 = \nabla \cdot \mathbf{H}_0$ *for* $i = 1, \ldots, n$, $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and* $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any* $t \in [0,T]$.

(i) Let $\mathbf{H}_0 \in \mathbf{H}^1(\Omega)$. *Then, there exists a positive constant* C *such that*

$$\max_{1 \leqslant i \leqslant n} \|\mathbf{h}_i\|_{\mathbf{H}^1(\Omega)}^2 + \sum_{i=1}^n \|\mathbf{h}_i - \mathbf{h}_{i-1}\|_{\mathbf{H}^1(\Omega)}^2 + \sum_{i=1}^n \|\delta \mathbf{h}_i\|^2 \tau \leqslant C.$$

(ii) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$, $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and* $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ *then*

$$\max_{1 \leqslant i \leqslant n} \|\delta \mathbf{h}_i\|^2 + \sum_{i=1}^n \|\delta \mathbf{h}_i - \delta \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\nabla \delta \mathbf{h}_i\|^2 \tau \leqslant C.$$

Therefore, the same stability results are obtained as in Lemma 4.3.1, where the curl-spaces are replaced by analogous $\mathbf{H}^s(\Omega)$-spaces. The variational formulation (4.23) can be rewritten in terms of the Rothe functions as

$$(\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}) + \left(\nabla \overline{\mathbf{H}}_n(t), \nabla \boldsymbol{\varphi}\right) + \left(\mathcal{K} \star \overline{\mathbf{H}}_n(t), \boldsymbol{\varphi}\right) = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right), \tag{4.24}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$ and a.a. $t \in (0,T)$.

Theorem 4.5.3 (Enhanced existence)**.** *Let* $\mathbf{H}_0 \in \mathbf{H}^1(\Omega)$ *and* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. *Assume that* $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *for any* $t \in [0,T]$. *If* $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any* $t \in [0,T]$, *then the solution* $\mathbf{H}$ *to problem* (4.1) *belongs to* $\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right)$ *with* $\partial_t \mathbf{H} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

Proof. The proof follows the same line as in Theorem 4.3.2. The main point of this theorem is the embedding

$$\mathbf{H}_0^1(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega).$$

Lemma 4.5.4(i) gives

$$\mathbf{H}_n \in \mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right) \quad \text{and} \quad \partial_t \mathbf{H}_n \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right).$$

The generalized Aubin-Lions lemma 2.12.4 implies that $\{\mathbf{H}_n\}$ is compact in the space $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. Therefore, there exists a subsequence of $\mathbf{H}_n$ (denoted by the same symbol again) for which we have

$$\mathbf{H}_n(\mathbf{x},t) \to \mathbf{H}(\mathbf{x},t) \qquad \text{for a.a. } (\mathbf{x},t) \in Q_T.$$

Moreover, $\mathbf{H} \in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$, $\mathbf{H}_n \rightharpoonup \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right)$ and
$\partial_t \mathbf{H}_n \rightharpoonup \partial_t \mathbf{H} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. □

The following error estimates have smaller constant C in comparison with the constants appearing in Theorem 4.3.3 because Grönwall's argument is avoided.

Theorem 4.5.4 (Error)**.** *Suppose that the assumptions of Theorem 4.5.3 are satisfied. Moreover, assume that* $\mathbf{F} \in \mathrm{Lip}([0,T], \mathbf{L}^2(\Omega))$.

(i) If $\mathbf{H}_0 \in \mathbf{H}^1(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \int_0^T \|\nabla[\mathbf{H}_n - \mathbf{H}]\|^2 \leqslant C\tau.$$

(ii) If $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \int_0^T \|\nabla[\mathbf{H}_n - \mathbf{H}]\|^2 \leqslant C\tau^2.$$

Proof. We subtract (4.22) from (4.24), set $\boldsymbol{\varphi} = \mathbf{H}_n(t) - \mathbf{H}(t)$ and integrate in time over $t \in (0,\eta) \subset (0,T)$ to get

$$\begin{aligned}\tfrac{1}{2}\|\mathbf{H}_n(\eta) - \mathbf{H}(\eta)\|^2 &+ \int_0^\eta \|\nabla[\mathbf{H}_n - \mathbf{H}]\|^2 + \int_0^\eta (\mathcal{K} \star [\mathbf{H}_n - \mathbf{H}], \mathbf{H}_n - \mathbf{H}) \\ &= \int_0^\eta \left(\overline{\mathbf{F}}_n - \mathbf{F}, \mathbf{H}_n - \mathbf{H}\right) + \int_0^\eta \left(\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla[\mathbf{H}_n - \mathbf{H}]\right) \\ &\quad + \int_0^\eta \left(\mathcal{K} \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \mathbf{H}_n - \mathbf{H}\right). \quad (4.25)\end{aligned}$$

The last term on the LHS is non-negative due to the positive definiteness of $\mathcal{K}$. Furthermore,

$$\begin{aligned}\left|\int_0^\eta \left(\overline{\mathbf{F}}_n - \mathbf{F}, \mathbf{H}_n - \mathbf{H}\right)\right| &\leqslant C_\varepsilon \int_0^\eta \left\|\overline{\mathbf{F}}_n - \mathbf{F}\right\|^2 + \varepsilon \int_0^\eta \|\mathbf{H}_n - \mathbf{H}\|^2 \\ &\leqslant C_\varepsilon \tau^2 + \varepsilon \int_0^\eta \|\nabla[\mathbf{H}_n - \mathbf{H}]\|^2.\end{aligned}$$

It holds that

$$\left\|\nabla[\mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t)]\right\| \leqslant \tau \left\|\partial_t \nabla \mathbf{H}_n(t)\right\| \qquad \text{for } t \in [0,T].$$

For the last term of (4.25), we may write using (4.21) and Friedrichs inequality that

$$\begin{aligned}
&\left|\int_0^\eta \left(\mathcal{K} \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \mathbf{H}_n - \mathbf{H}\right)\right| \\
&\quad\leqslant \quad \varepsilon \int_0^\eta \left\|\mathbf{H}_n - \mathbf{H}\right\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n]\right\|^2 \\
&\quad\leqslant \quad \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n]\right\|^2.
\end{aligned}$$

Thus, employing Lemma 4.5.4(i) and (ii), we see that

$$\text{(i)} \quad \left|\int_0^\eta \left(\mathcal{K} \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \mathbf{H}_n - \mathbf{H}\right)\right| \leqslant \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \tau$$

$$\text{(ii)} \quad \left|\int_0^\eta \left(\mathcal{K} \star [\mathbf{H}_n - \overline{\mathbf{H}}_n], \mathbf{H}_n - \mathbf{H}\right)\right| \leqslant \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \tau^2.$$

It remains to estimate the second term on the RHS in (4.25). Depending on the stability results, we have that

(i)

$$\begin{aligned}
&\left|\int_0^\eta \left(\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla[\mathbf{H}_n - \mathbf{H}]\right)\right| \\
&\quad\leqslant \quad \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n]\right\|^2 \\
&\quad\leqslant \quad \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \tau
\end{aligned}$$

(ii)

$$\begin{aligned}
&\left|\int_0^\eta \left(\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla[\mathbf{H}_n - \mathbf{H}]\right)\right| \\
&\quad\leqslant \quad \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \overline{\mathbf{H}}_n]\right\|^2 \\
&\quad\leqslant \quad \varepsilon \int_0^\eta \left\|\nabla[\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \tau^2.
\end{aligned}$$

Putting things together and choosing a sufficiently small positive ε, we conclude the proof. $\square$

4.5.3 Modified scheme in $\mathbf{H}^1(\Omega)$

Now, the following semi-implicit time-discrete scheme is considered, which represents a slight modification of (4.23):

$$\left\{\begin{aligned} (\delta \mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) &= (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K} \star \mathbf{h}_{i-1}, \boldsymbol{\varphi}), \\ \mathbf{h}_0 &= \mathbf{H}_0, \end{aligned}\right. \tag{4.26}$$

which is equivalent to ($\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$)

$$\begin{aligned} a(\mathbf{h}_i, \boldsymbol{\varphi}) := \left(\frac{\mathbf{h}_i}{\tau}, \boldsymbol{\varphi}\right) + (\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) \\ = (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K} \star \mathbf{h}_{i-1}, \boldsymbol{\varphi}) + \left(\frac{\mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

The existence of a unique solution to (4.26) is obtained for any $i = 1, \ldots, n$ and any $\tau > 0$ if

$$\mathbf{H}_0 \in \mathbf{L}^q(\Omega) \cap \{\boldsymbol{\varphi} \in \mathbf{H}(\mathbf{curl}; \Omega) \cap \mathbf{H}_0(\mathbf{div}; \Omega) \, : \, \nabla \cdot \boldsymbol{\varphi} = 0\}, \quad q > 3.$$

The scheme (4.26) can be analysed in the same way as (4.23). Therefore, further details are omitted. It is important to note that the use of Grönwall's lemma with exponential in time character of the constant cannot be avoided despite the positive definiteness of $\mathcal{K}$. Nevertheless, the error estimates from Theorem 4.5.4 remain valid for (4.26). The main difference is that the constants C are larger. In the following section, the convergence of this scheme is illustrated in some numerical experiments. Future research can concern the implementation of the other schemes. This section is concluded with a comparison of the obtained results with the London equations.

4.5.4 Comparison with the London equations

Consider a thin superconducting slab, of thickness $2L$, as shown in Fig. 4.1. The

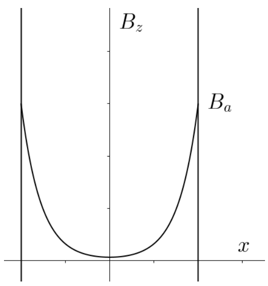


Figure 4.1: The magnetic field inside a superconducting slab of thickness $2L$.

second London equation is given by $\nabla \times \mathbf{J}_s = -\Lambda^{-1}\mathbf{B}$. This equation, combined with the quasi-static Maxwell equation $\nabla \times \mathbf{H} = \mathbf{J}_s$ and equation (3.3), gives

$$\nabla \times \nabla \times \mathbf{B} = -\frac{1}{\beta}\mathbf{B},$$

where $\beta := \frac{\Lambda}{\mu} = \frac{m_e}{\mu n_s e^2}$. But $\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \Delta\mathbf{B} = -\Delta\mathbf{B}$ since $\nabla \cdot \mathbf{B} = 0$. Hence,

$$\Delta\mathbf{B} = \frac{1}{\beta}\mathbf{B}. \tag{4.27}$$

Suppose that an external parallel magnetic field is applied in the z direction parallel to the slab surfaces, i.e. $\mathbf{B} = (0, 0, B_a)$. Given that the magnetic field is a function of only x inside the superconductor, $\mathbf{B} = (0, 0, B_z(x))$, equation (4.27) is equivalent with

$$\frac{d^2 B_z(x)}{dx^2} = \frac{1}{\beta} B_z(x). \tag{4.28}$$

Solving equation (4.28) with the boundary conditions that $B_z = B_a$ at the two surfaces at $x = \pm L$, the solution inside the slab becomes

$$B_z(x) = B_a \frac{\cosh\left(\frac{x}{\sqrt{\beta}}\right)}{\cosh\left(\frac{L}{\sqrt{\beta}}\right)}. \tag{4.29}$$

This result shows that a magnetic field is exponentially decayed at the surface of a superconductor (Meissner effect). The magnetic field $\mathbf{B}$ penetrates at the surfaces corresponding with $x = \pm L$ over a distance of approximately $\sqrt{\beta}$ (London penetration depth). Analogue, choosing $\mathbf{J}_n = \mathbf{0}$ in the calculation of model (4.19), the magnetic field $\mathbf{H}$ satisfies the following elliptic integro-differential equation for type-I superconductivity

$$\Delta\mathbf{H} = \mathcal{K} \star \mathbf{H} \qquad \text{or} \qquad \Delta\mathbf{B} = \mathcal{K} \star \mathbf{B}. \tag{4.30}$$

If $r_0 \to 0$, which means that the support of $\mathcal{K}$ becomes smaller and smaller, $\mathcal{K}$ becomes a Dirac delta function. Consequently, $\mathcal{K} \to \frac{\widetilde{C}}{2}\delta(\mathbf{x} - \mathbf{x}')$ and (4.30) converts to

$$\Delta\mathbf{B} = \frac{\widetilde{C}}{2}\mathbf{B}.$$

The identification $\frac{\widetilde{C}}{2} = \frac{\mu}{\Lambda} = \frac{1}{\beta}$ yields the London equation (4.27). An interesting area for future research is to compare the numerical solution of problem (4.30) for small r_0 with the exact solution (4.29) in the case of an infinite slab.

4.6 Numerical experiments

The numerical experiments follow closely the theoretical analysis. The solution of problem (4.1) is approximated by solving problem (4.19) using the semi-implicit scheme (4.26). Backward Euler's method (Rothe's method) for an equidistant time-discretization with time step $\tau = 2^{-j}$, $2 \leqslant j \leqslant 7$, is used. The resulting elliptic BVPs

$$\left\{ \begin{aligned} (\mathbf{h}_i, \boldsymbol{\varphi}) + \tau\,(\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) \ &= \tau\,(\mathbf{F}_i, \boldsymbol{\varphi}) \\ &\quad -\tau\,(\mathcal{K} \star \mathbf{h}_{i-1}, \boldsymbol{\varphi}) + (\mathbf{h}_{i-1}, \boldsymbol{\varphi})\,, \\ \mathbf{h}_0 \ &= \mathbf{H}_0, \end{aligned} \right. \tag{4.31}$$

$\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$, $i = 1, \ldots, 2^j T$, are solved using the FEM with first-order Lagrange finite elements.

The problems (4.31) are standard except of the appearance of the convolution term. It is this term that complicates the space discretization. It is assumed that the domain $\Omega \subset \mathbb{R}^3$ is a Lipschitz polyhedron. Then, the domain Ω can be triangulated into a finite set of tetrahedra $\mathcal{T}_h$ such that $\overline{\Omega} = \bigcup_{\mathcal{T} \in \mathcal{T}_h} \overline{\mathcal{T}}$, see [66]. Note that $h = \max_{\mathcal{T} \in \mathcal{T}_h} h_{\mathcal{T}}$, where $h_{\mathcal{T}}$ is the diameter of the smallest sphere containing $\mathcal{T}$. The total number of vertices is set equal to M and the ith vertex of $\mathcal{T}_h$ is put equal to $\mathbf{x}_i$. Denote by $\mathbf{x}_{m,\mathcal{T}}$ and Vol$(\mathcal{T})$ the midpoint and the volume of a tetrahedron $\mathcal{T} \in \mathcal{T}_h$ respectively. Define the set

$$\mathcal{T}_{\mathbf{x}} := \{\mathcal{T} \in \mathcal{T}_h : |\mathbf{x}_{m,\mathcal{T}} - \mathbf{x}| < r_0\} \subset \mathcal{T}_h.$$

The convolution integral arising in the numerical experiments is solved numerically as follows

$$\boldsymbol{\mathcal{C}}(\mathbf{x}, \boldsymbol{\varphi}) := \mathcal{K}(\mathbf{x}, \cdot) \star \boldsymbol{\varphi} \approx \sum_{\mathcal{T} \in \mathcal{T}_{\mathbf{x}}} \mathrm{Vol}(\mathcal{T}) \mathcal{K}(\mathbf{x} - \mathbf{x}_{m,\mathcal{T}}) \boldsymbol{\varphi}(\mathbf{x}_{m,\mathcal{T}}). \tag{4.32}$$

This is a way to avoid the singularity in the kernel. The term $(\mathcal{K} \star \mathbf{h}_{i-1}, \boldsymbol{\varphi})$ in the variational formulation can be considered as

$$\int_\Omega \boldsymbol{\mathcal{C}}(\mathbf{x}, \mathbf{h}_{i-1}) \cdot \boldsymbol{\varphi}(\mathbf{x}) \ \mathrm{d}\mathbf{x}.$$

As mentioned before, at each time step, the resulting elliptic BVP (4.31) is solved numerically by the FEM. For each component of the unknown vector field first-order Lagrange elements on tetrahedra are used, see Example 2.13.4. An algebraic system of the form

$$AH^i = b^i$$

is solved, where H^i contains the degrees of freedom of the different components of $\mathbf{H}$.

The algorithm for the determination of the form of the matrix A and the vector b^i is written down. Remark that the matrix A is time independent and can be computed before the time stepping.

Algorithm

- Choose T, τ, $\widetilde{C}$ and r_0. Prescribe $\mathbf{F}$ and $\mathbf{H}_0$.
- Set $t = \tau$;
- Define separate forms $a_M(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \mathbf{v})$ and $a_S(\mathbf{u}, \mathbf{v}) = (\nabla \mathbf{u}, \nabla \mathbf{v})$;
- Assemble a_M to the mass matrix M and a_S to the stiffness matrix S;
- Compute $A = M + \tau S$;
- While $t \leqslant T$:
  - Interpolate the solution on the previous time step, $\mathbf{h}_{i-1}$, componentwise to a finite element function H^{i-1} in $\boldsymbol{\mathcal{V}}_h := (\mathcal{V}_h)^3$;
  - Interpolate the formula for $\mathbf{F}$ on time step t_i, i.e. $\mathbf{F}_i$, componentwise to a finite element function F^i in $\boldsymbol{\mathcal{V}}_h$.
  - The convolution $(\mathcal{K} \star \mathbf{h}_{i-1})(\mathbf{x})$ can be approximated by $\boldsymbol{\mathcal{C}}(\mathbf{x}, \mathbf{h}_{i-1})$. Interpolate $\boldsymbol{\mathcal{C}}(\mathbf{x}, \mathbf{h}_{i-1})$ componentwise to a finite element function $\mathcal{C}_p$ in $\boldsymbol{\mathcal{V}}_h$;
  - Set $b^i = \tau M F^i - \tau M \mathcal{C}_p + M H^{i-1}$;
  - Solve $A H^i = b^i$ for H^i and store in $\mathbf{h}_i$;
  - $t \leftarrow t + \tau$;
  - $\mathbf{h}_{i-1} \leftarrow \mathbf{h}_i$.

The following assumptions are made in the experiments: $T = 1$, $\Omega = (0,1) \times (0,1) \times (0,1)$ and $r_0 = 0.1$. Moreover, two values for the parameter $\widetilde{C}$ are used, namely $\widetilde{C} = 2$ and $\widetilde{C} = 150$. For the space discretization, a fixed uniform mesh that gives a good approximation for the convolution integral is needed. The number of cells in each direction is chosen to be equal, namely (n_x, n_x, n_x). The volume of the sphere S with center $\mathbf{x}$ and radius r_0, denoted by Vol(S), is compared with the volume of the set $\mathcal{T}_\mathbf{x}$ that is defined by

$$\mathrm{Vol}(\mathcal{T}_\mathbf{x}) = \sum_{\mathcal{T} \in \mathcal{T}_\mathbf{x}} \mathrm{Vol}(\mathcal{T}),$$

where $\mathbf{x}$ is a node in $(r_0, 1-r_0) \times (r_0, 1-r_0) \times (r_0, 1-r_0)$. The results are given in Table 4.1. The best approximation for $n_x \leqslant 30$ is 25. For this reason, the total number of tetrahedra in the experiments is 93750. For every time step τ, the error

$$E = \max_{t \in [0,T]} \|\mathbf{h}_n(t) - \mathbf{H}(t)\|^2 \tag{4.33}$$

is computed. For all experiments, the exact solution belongs to $\boldsymbol{\mathcal{V}}_h$. Therefore, the error is only due to the time discretization and the approximation of the convolution. Note that in the numerical experiments, the assumptions $\mathbf{H} \times \boldsymbol{\nu} = \mathbf{0}$ and $\mathbf{H} \cdot \boldsymbol{\nu} = \mathbf{0}$ do not need to be satisfied. The proposed schemes are also valid for a more general boundary condition.

n_x	$\lvert\mathrm{Vol}(S) - \mathrm{Vol}(\mathcal{T}_{\mathbf{x}})\rvert$	n_x	$\lvert\mathrm{Vol}(S) - \mathrm{Vol}(\mathcal{T}_{\mathbf{x}})\rvert$
11	0.000319	21	0.000346
12	0.000441	22	0.000244
13	0.000547	23	0.000408
14	0.000184	24	0.000138
15	0.000633	**25**	**0.000093**
16	0.000206	26	0.000312
17	0.000289	27	0.000124
18	0.000416	28	0.000271
19	0.000185	29	0.000403
20	0.000561	30	0.000108

Table 4.1: $\lvert Vol(S) - Vol(\mathcal{T}_{\mathbf{x}})\rvert$ for $10 < n_x \leqslant 30$.

4.6.1 Experiment 1

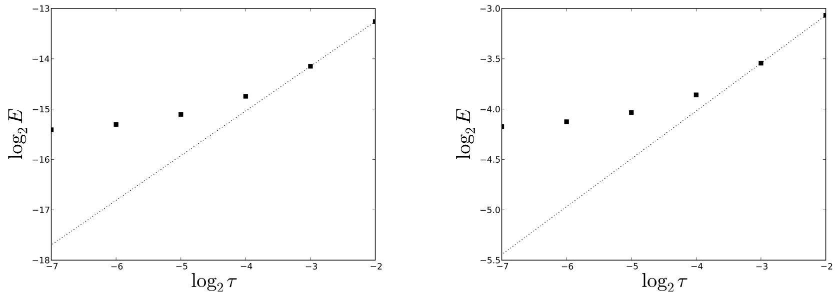
In the first experiment,

$$\mathbf{H}^{\mathrm{ex}} = (1+t^2)\begin{pmatrix}1\\1\\1\end{pmatrix}$$

is used as exact solution. The RHS $\mathbf{F}$ can be calculated exactly in $(r_0, 1-r_0) \times (r_0, 1-r_0) \times (r_0, 1-r_0)$, namely

$$\mathbf{F}^{\mathrm{ex}} = 2t + (2+2t^2)\pi\widetilde{C}r_0\exp(-1.0)\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

The error (4.33) is computed for $\Omega = (r_0, 1-r_0) \times (r_0, 1-r_0) \times (r_0, 1-r_0)$ and $\tau = 2^{-j}$, $2 \leqslant j \leqslant 7$, and is depicted in Figure 4.2 for $\widetilde{C} = 2$ and $\widetilde{C} = 150$. The error $\log_2 E$ is plotted as a function of $\log_2 \tau$ because then the order of convergence corresponds with the slope of the regression line. The expected convergence rate for smooth functions is predicted in Theorem 4.5.4: $E \sim \mathcal{O}\left(\tau^2\right)$. The linear regression line for the first two data points are given by $\log_2 E = 0.8881 \log_2 \tau - 11.4836$ and $\log_2 E = 0.4752 \log_2 \tau - 2.1577$ for $\widetilde{C} = 2$, respectively $\widetilde{C} = 150$. Therefore, the expected linear behaviour is not obtained. With decreasing time step, the error in the approximation of the convolution starts dominating over the time discretization error. However, the error in the approximation of the convolution is sufficiently small. In the following two experiments, the error due to the numerical convolution is cancelled out because the numerical convolution is also used in the determination of $\mathbf{F}$.



(a) $\widetilde{C} = 2$ (b) $\widetilde{C} = 150$

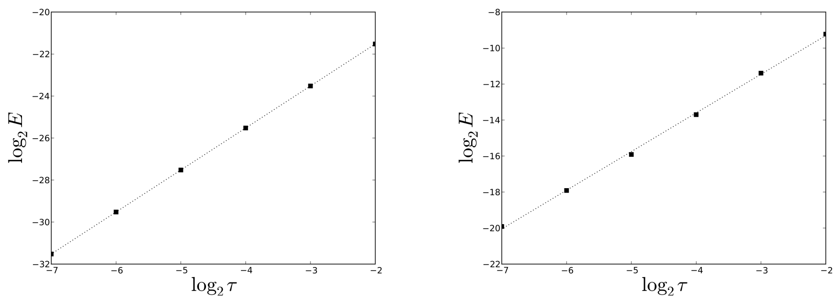
Figure 4.2: Results of numerical experiment 1: (a) convergence rate for experiment 1 with $\widetilde{C} = 2$; (b) convergence rate for experiment 1 with $\widetilde{C} = 150$.

4.6.2 Experiment 2

In this experiment, the following exact solution is defined

$$\mathbf{H}^{\mathrm{ex}} = (1+t)\begin{pmatrix} y - z \\ z - x \\ x - y \end{pmatrix}.$$

The error (4.33) is again computed for $\tau = 2^{-j}$, $2 \leqslant j \leqslant 7$, and is shown in Figure 4.3 for $\widetilde{C} = 2$ and $\widetilde{C} = 150$. Now, a linear regression line is calculated through all the obtained data points: $\log_2 E = 2\log_2 \tau - 17.521$ and $\log_2 E = 2.148\log_2 \tau - 5.0102$ for $\widetilde{C} = 2$ and $\widetilde{C} = 150$ respectively. This is in accordance with the predicted convergence rate of $\mathcal{O}(\tau^2)$ in Theorem 4.5.4.



(a) $\widetilde{C} = 2$ (b) $\widetilde{C} = 150$

Figure 4.3: Results of numerical experiment 2: (a) convergence rate for experiment 2 with $\widetilde{C} = 2$; (b) convergence rate for experiment 2 with $\widetilde{C} = 150$.

4.6.3 Experiment 3

In the last experiment, the following exact solution is taken

$$\mathbf{H}^{\text{ex}} = (1+t^2)\begin{pmatrix} y-z \\ z-x \\ x-y \end{pmatrix}.$$

The linear regression lines are $\log_2 E = 1.9753 \log_2 \tau - 12.858$ and $\log_2 E = 1.9678 \log_2 \tau - 4.0842$ for $\widetilde{C} = 2$ and $\widetilde{C} = 150$ respectively, see Figure 4.4. The predicted convergence rate of $\mathcal{O}\left(\tau^2\right)$ is not exactly obtained since the exact solution is quadratic in time.

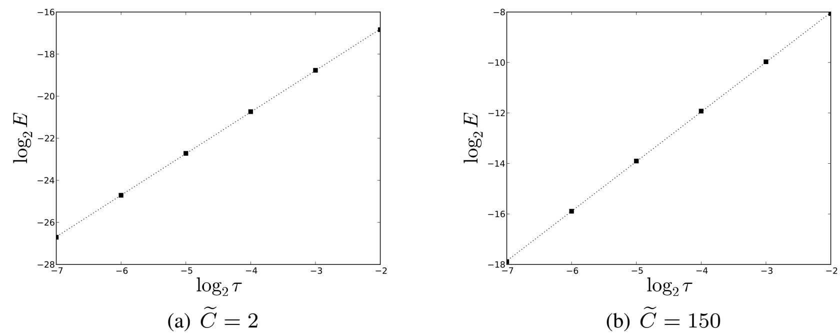


Figure 4.4: Results of numerical experiment 3: (a) convergence rate for experiment 3 with $\widetilde{C} = 2$; (b) convergence rate for experiment 3 with $\widetilde{C} = 150$.

Acknowledgement The computational resources (Stevin Supercomputer Infrastructure) and services used in this section were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, the Hercules Foundation and the Flemish Government - department EWI.

4.7 Full discretization

Numerical experiments are performed in the previous section. It is investigated how the error behaves with decreasing time step. From theoretical point of view, it is also interesting to analyse the space discretization error. In doing this, a linear numerical scheme discretized in time and space for finding an approximation of the solution to problem (4.1) is proposed.

The first step is to generate a finite element mesh that covers the domain Ω, which is a bounded polyhedral Lipschitz continuous domain in $\mathbb{R}^3$. It is assumed that there exists a regular family of meshes or triangulations $\{\mathcal{T}_h : h > 0\}$, where h denotes the mesh parameter. The purpose of this section is to analyse the error as h decreases.

The second step is the consideration of a finite element subspace $\mathbf{V}^h$ of $\mathbf{H}(\mathbf{curl};\Omega)$. In order to take the boundary condition $\mathbf{H} \times \boldsymbol{\nu} = \mathbf{0}$ into account, the finite dimensional subspace $\mathbf{V}_0^h = \{\mathbf{v}^h \in \mathbf{V}^h : \mathbf{v}^h \times \boldsymbol{\nu} = \mathbf{0} \text{ on } \partial\Omega\}$ of $\mathbf{H}_0(\mathbf{curl};\Omega)$ is considered. The following projection operators into $\mathbf{V}_0^h$ are used. Let $\mathbf{P}_h : \mathbf{L}^2(\Omega) \to \mathbf{V}_0^h$ be the orthogonal projection operator such that if $\mathbf{u} \in \mathbf{L}^2(\Omega)$ then $\mathbf{P}_h\mathbf{u} \in \mathbf{V}_0^h$ satisfies

$$(\mathbf{u}, \mathbf{v}_h) = (\mathbf{P}_h\mathbf{u}, \mathbf{v}_h), \quad \forall \mathbf{v}_h \in \mathbf{V}_0^h. \tag{4.34}$$

Analogously, let $\widetilde{\mathbf{P}}_h : \mathbf{H}_0(\mathbf{curl};\Omega) \to \mathbf{V}_0^h$ be the orthogonal projection operator such that if $\mathbf{u} \in \mathbf{H}_0(\mathbf{curl};\Omega)$ then $\widetilde{\mathbf{P}}_h\mathbf{u} \in \mathbf{V}_0^h$ satisfies

$$\begin{aligned}(\mathbf{u}, \mathbf{v}_h) &+ (\nabla \times \mathbf{u}, \nabla \times \mathbf{v}_h) \\ &= \left(\widetilde{\mathbf{P}}_h\mathbf{u}, \mathbf{v}_h\right) + \left(\nabla \times \widetilde{\mathbf{P}}_h\mathbf{u}, \nabla \times \mathbf{v}_h\right), \quad \forall \mathbf{v}_h \in \mathbf{V}_0^h.\end{aligned} \tag{4.35}$$

Choosing $\mathbf{v}_h = \mathbf{P}_h\mathbf{u}$ in (4.34) and $\mathbf{v}_h = \widetilde{\mathbf{P}}_h\mathbf{u}$ in (4.35), it is easy to proof that $\mathbf{P}_h$ and $\widetilde{\mathbf{P}}_h$ are linear bounded operators.

At this point, a fully discrete scheme can be defined. After time and space discretization, the following approximation of problem (4.1) can be obtained: find $\mathbf{h}_i^h \in \mathbf{V}_0^h$, $1 \leqslant i \leqslant n$, such that

$$\left\{\begin{aligned} \left(\delta\mathbf{h}_i^h, \boldsymbol{\varphi}^h\right) + \left(\nabla \times \mathbf{h}_i^h, \nabla \times \boldsymbol{\varphi}^h\right)& \\ + \left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla \times \boldsymbol{\varphi}^h\right) &= \left(\mathbf{P}_h\mathbf{F}_i, \boldsymbol{\varphi}^h\right) = \left(\mathbf{F}_i, \boldsymbol{\varphi}^h\right), \\ \mathbf{h}_0^h &= \widetilde{\mathbf{P}}_h\mathbf{H}_0, \end{aligned}\right. \tag{4.36}$$

is satisfied for all $\boldsymbol{\varphi}^h \in \mathbf{V}_0^h$. This problem is equivalent with solving $a^h(\mathbf{h}_i^h, \boldsymbol{\varphi}^h) = f^h(\boldsymbol{\varphi}^h)$ for all $\boldsymbol{\varphi}^h \in \mathbf{V}_0^h$, where $a^h : \mathbf{V}_0^h \times \mathbf{V}_0^h \to \mathbb{R}$ and $f^h : \mathbf{V}_0^h \to \mathbb{R}$ are defined by

$$a^h(\mathbf{h}_i^h, \boldsymbol{\varphi}^h) = \left(\frac{\mathbf{h}_i^h}{\tau}, \boldsymbol{\varphi}^h\right) + \left(\nabla \times \mathbf{h}_i^h, \nabla \times \boldsymbol{\varphi}^{\mathbf{h}}\right) + \left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla \times \boldsymbol{\varphi}^h\right)$$

and

$$f^h(\boldsymbol{\varphi}^h) = \left(\mathbf{F}_i, \boldsymbol{\varphi}^h\right) + \left(\frac{\mathbf{h}_{i-1}^h}{\tau}, \boldsymbol{\varphi}^h\right).$$

Note that $\mathbf{h}_i^h$ denotes the finite element solution at time $t = t_i$.

Theorem 4.7.1. *Suppose that* $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$. *Then the variational problem* (4.36) *admits a unique solution* $\mathbf{h}_i^h \in \mathbf{V}_0^h$ *for any* $i = 1, \ldots, n$ *if* $\tau < \tau_0$.

Proof. This is an easy application of the Lax-Milgram lemma 2.11.1 for any $i = 1, \ldots, n$. It holds that

$$a^h(\mathbf{v}^h, \mathbf{v}^h) \geqslant \frac{1}{\tau}\left\|\mathbf{v}^h\right\|^2 + \left\|\nabla\times\mathbf{v}^h\right\|^2 - \left|\left(\mathcal{K}_0 \star \mathbf{v}^h, \nabla\times\mathbf{v}^h\right)\right|$$
$$\overset{(4.4)}{\geqslant} \left(\frac{1}{\tau} - C_\varepsilon\right)\left\|\mathbf{v}^h\right\|^2 + (1-\varepsilon)\left\|\nabla\times\mathbf{v}^h\right\|^2.$$

Fixing $\varepsilon < 1$ proofs that the bilinear form $a^h(\cdot,\cdot)$ is elliptic in the Hilbert space $\mathbf{V}_0^h$ for $\tau < \tau_0$. Moreover, a^h is continuous in $\mathbf{V}_0^h$. If $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$, then the functional $f^h(\cdot)$ is linear and bounded in $\mathbf{V}_0^h$. □

A stability analysis is needed to derive the error estimates for the full discretization.

Lemma 4.7.1 (Stability analysis)**.** *Suppose that* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

(i) Let $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$. *Then, there exists a positive constant* C *such that for all* $\tau < \tau_0$ *it holds that*

$$\max_{1\leqslant i\leqslant n}\left\|\mathbf{h}_i^h\right\|^2 + \sum_{i=1}^n \left\|\mathbf{h}_i^h - \mathbf{h}_{i-1}^h\right\|^2 + \sum_{i=1}^n \left\|\nabla\times\mathbf{h}_i^h\right\|^2\tau \leqslant C.$$

(ii) If $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *then for all* $\tau < \tau_0$ *it holds that*

$$\max_{1\leqslant i\leqslant n}\left\|\nabla\times\mathbf{h}_i^h\right\|^2 + \sum_{i=1}^n \left\|\nabla\times\mathbf{h}_i^h - \nabla\times\mathbf{h}_{i-1}^h\right\|^2 + \sum_{i=1}^n \left\|\delta\mathbf{h}_i^h\right\|^2\tau \leqslant C.$$

(iii) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$, $\partial_t\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, $\nabla\times(\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then for all* $\tau < \tau_0$

$$\max_{1\leqslant i\leqslant n}\left\|\delta\mathbf{h}_i^h\right\|^2 + \sum_{i=1}^n \left\|\delta\mathbf{h}_i^h - \delta\mathbf{h}_{i-1}^h\right\|^2 + \sum_{i=1}^n \left\|\nabla\times\delta\mathbf{h}_i^h\right\|^2\tau \leqslant C.$$

Proof. (i) First, we set $\boldsymbol{\varphi}^h = \mathbf{h}_i^h$ in (4.36). Then, we multiply the result by τ and sum it up for $i = 1, \ldots, j$ $(1 \leqslant j \leqslant n)$ to arrive at

$$\sum_{i=1}^j \left(\delta\mathbf{h}_i^h, \mathbf{h}_i^h\right)\tau + \sum_{i=1}^j \left\|\nabla\times\mathbf{h}_i^h\right\|^2\tau + \sum_{i=1}^j \left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla\times\mathbf{h}_i^h\right)\tau = \sum_{i=1}^j \left(\mathbf{F}_i, \mathbf{h}_i^h\right)\tau.$$

For the first term on the left-hand side (LHS), we use Abel's summation rule

$$2\sum_{i=1}^j \left(\delta\mathbf{h}_i^h, \mathbf{h}_i^h\right)\tau = \left\|\mathbf{h}_j^h\right\|^2 - \left\|\widetilde{\mathbf{P}}_h\mathbf{H}_0\right\|^2 + \sum_{i=1}^j \left\|\mathbf{h}_i^h - \mathbf{h}_{i-1}^h\right\|^2.$$

For the third term on the LHS, using (4.4), we have that

$$\left|\sum_{i=1}^{j}\left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla \times \mathbf{h}_i^h\right)\tau\right| \leqslant \varepsilon \sum_{i=1}^{j}\left\|\nabla \times \mathbf{h}_i^h\right\|^2 \tau + C_\varepsilon \sum_{i=1}^{j}\left\|\mathbf{h}_i^h\right\|^2 \tau.$$

For the RHS, we apply the Cauchy and Young inequalities. Fixing ε sufficiently small and applying Grönwall's argument gives the proof.

(ii) Now, we put $\boldsymbol{\varphi}^h = \delta \mathbf{h}_i^h$ in (4.36). Again, we multiply by τ and sum up for $i = 1, \dots, j$ $(1 \leqslant j \leqslant n)$

$$\sum_{i=1}^{j}\left\|\delta \mathbf{h}_i^h\right\|^2 \tau + \sum_{i=1}^{j}\left(\nabla \times \mathbf{h}_i^h, \nabla \times \mathbf{h}_i^h - \nabla \times \mathbf{h}_{i-1}^h\right) \\ + \sum_{i=1}^{j}\left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla \times \delta \mathbf{h}_i^h\right)\tau = \sum_{i=1}^{j}\left(\mathbf{F}_i, \delta \mathbf{h}_i^h\right)\tau.$$

Abel's summation rule gives

$$2\sum_{i=1}^{j}\left(\nabla \times \mathbf{h}_i^h, \nabla \times \mathbf{h}_i^h - \nabla \times \mathbf{h}_{i-1}^h\right) \\ = \left\|\nabla \times \mathbf{h}_j^h\right\|^2 - \left\|\nabla \times \widetilde{\mathbf{P}}_h \mathbf{H}_0\right\|^2 + \sum_{i=1}^{j}\left\|\nabla \times \mathbf{h}_i^h - \nabla \times \mathbf{h}_{i-1}^h\right\|^2$$

and

$$\sum_{i=1}^{j}\left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla \times \delta \mathbf{h}_i^h\right)\tau = \left(\mathcal{K}_0 \star \mathbf{h}_j^h, \nabla \times \mathbf{h}_j^h\right) \\ - \left(\mathcal{K}_0 \star \widetilde{\mathbf{P}}_h \mathbf{H}_0, \nabla \times \widetilde{\mathbf{P}}_h \mathbf{H}_0\right) - \sum_{i=1}^{j}\left(\mathcal{K}_0 \star \delta \mathbf{h}_i^h, \nabla \times \mathbf{h}_{i-1}^h\right)\tau.$$

Hence, using (i) and (4.4), we obtain that

$$\left|\sum_{i=1}^{j}\left(\mathcal{K}_0 \star \mathbf{h}_i^h, \nabla \times \delta \mathbf{h}_i^h\right)\tau\right| \leqslant C_\varepsilon + \varepsilon\left\|\nabla \times \mathbf{h}_j^h\right\|^2 + \varepsilon \sum_{i=1}^{j}\left\|\delta \mathbf{h}_i^h\right\|^2 \tau.$$

Fixing a sufficiently small positive ε concludes the proof.

(iii) We define the following compatibility condition

$$\delta \mathbf{h}_0^h := \mathbf{P}_h \partial_t \mathbf{H}(0) = \mathbf{P}_h \mathbf{F}(0) - \mathbf{P}_h\left(\nabla \times \nabla \times \mathbf{H}_0\right) - \mathbf{P}_h\left(\nabla \times \left(\mathcal{K}_0 \star \mathbf{H}_0\right)\right).$$

We subtract (4.36) for $i = i - 1$ from (4.36). Then, we set $\boldsymbol{\varphi}^h = \delta \mathbf{h}_i^h$ and we sum the result up for $i = 1, \dots, j$ with $1 \leqslant j \leqslant n$ to get

$$\sum_{i=1}^{j} \left(\delta^2 \mathbf{h}_i^h, \delta \mathbf{h}_i^h\right) \tau + \sum_{i=1}^{j} \left\|\nabla \times \delta \mathbf{h}_i^h\right\|^2 \tau + \sum_{i=1}^{j} \left(\mathcal{K}_0 \star \delta \mathbf{h}_i^h, \nabla \times \delta \mathbf{h}_i^h\right) \tau = \sum_{i=1}^{j} \left(\delta \mathbf{F}_i, \delta \mathbf{h}_i^h\right) \tau.$$

Furthermore, we follow the same lines as in (i) when considering $\delta \mathbf{h}_i^h$ instead of $\mathbf{h}_i^h$. $\square$

Now, the following piecewise linear in time vector fields $\mathbf{H}_n^h$ and the piecewise constant in time fields $\overline{\mathbf{H}}_n^h$ are defined

$$\begin{aligned}
\mathbf{H}_n^h(0) &= \widetilde{\mathbf{P}}_h \mathbf{H}_0, \quad \mathbf{H}_n^h(t) = \mathbf{h}_{i-1}^h + (t - t_{i-1}) \delta \mathbf{h}_i^h \qquad && \text{for } t \in (t_{i-1}, t_i], \\
\overline{\mathbf{H}}_n^h(0) &= \widetilde{\mathbf{P}}_h \mathbf{H}_0, \quad \overline{\mathbf{H}}_n^h(t) = \mathbf{h}_i^h, && \text{for } t \in (t_{i-1}, t_i],
\end{aligned}$$

$i = 1, \dots, n$. The full discretized system (4.36) can be rewritten by Rothe's notation for all $\boldsymbol{\varphi}^h \in \mathbf{V}_0^h$ and a.a. $t \in (0, T)$ as follows

$$\left\{ \begin{aligned}
\left(\partial_t \mathbf{H}_n^h(t), \boldsymbol{\varphi}^h\right) + \left(\nabla \times \overline{\mathbf{H}}_n^h(t), \nabla \times \boldsymbol{\varphi}^h\right) & \\
+ \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n^h(t), \nabla \times \boldsymbol{\varphi}^h\right) &= \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}^h\right); \\
\mathbf{H}_n^h(0) &= \widetilde{\mathbf{P}}_h \mathbf{H}_0.
\end{aligned} \right. \tag{4.37}$$

The next theorem summarizes the error estimate for the full discretization.

Theorem 4.7.2. *Suppose that* $\mathbf{F} \in \operatorname{Lip}([0, T], \mathbf{L}^2(\Omega))$.

(i) Let the weak solution $\mathbf{H}$ *of* (4.1) *at time t and the initial condition* $\mathbf{H}_0$ *satisfy* $\mathbf{H}(t), \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl}; \Omega)$. *Then for any* $\tau < \tau_0$, *there exists a constant* C *such that*

$$\begin{aligned}
&\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \\
&\qquad \leqslant C \left(\tau + \left\|\mathbf{H}_0 - \widetilde{\mathbf{P}}_h \mathbf{H}_0\right\|^2 + \sqrt{\int_0^\eta \left\|\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right\|^2} \right. \\
&\qquad\qquad \left. + \int_0^\eta \left\|\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right)\right\|^2 \right)
\end{aligned}$$

is valid for any $\eta \in [0, T]$;

(ii) Let the weak solution $\mathbf{H}$ *of* (4.1) *at time t and the initial condition* $\mathbf{H}_0$ *satisfy* $\mathbf{H}(t), \partial_t \mathbf{H}(t), \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl}; \Omega)$. *Then, for any* $\tau < \tau_0$, *there exists a*

constant C such that

$$\begin{aligned}&\left\|\mathbf{H}(\eta)-\mathbf{H}_n^h(\eta)\right\|^2+\int_0^\eta\left\|\nabla\times\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\right\|^2\\&\leqslant C\left(\tau+\left\|\mathbf{H}_0-\widetilde{\mathbf{P}}_h\mathbf{H}_0\right\|^2+\left\|\mathbf{H}(\eta)-\widetilde{\mathbf{P}}_h\mathbf{H}(\eta)\right\|^2+\int_0^\eta\left\|\mathbf{H}-\widetilde{\mathbf{P}}_h\mathbf{H}\right\|^2\right.\\&\qquad\left.+\int_0^\eta\left\|\partial_t\left(\mathbf{H}-\widetilde{\mathbf{P}}_h\mathbf{H}\right)\right\|^2+\int_0^\eta\left\|\nabla\times\left(\mathbf{H}-\widetilde{\mathbf{P}}_h\mathbf{H}\right)\right\|^2\right)\end{aligned}$$

is valid for any $\eta\in[0,T]$;

(iii) If the initial condition also satisfies $\nabla\times\mathbf{H}_0$ and $\mathcal{K}_0\star\mathbf{H}_0\in\mathbf{H}_0(\mathbf{curl};\Omega)$, then the estimates in (i) and (ii) are satisfied with τ^2 instead of τ.

Please note that the positive constant C in these estimates is of the form Ce^{CT}.

Proof. (i) We subtract (4.37) from (4.5) for $\boldsymbol{\varphi}=\boldsymbol{\varphi}^h$. We set $\boldsymbol{\varphi}^h=\widetilde{\mathbf{P}}_h\mathbf{H}(t)-\mathbf{H}_n^h(t)$ and integrate in time over $(0,\eta)$ for $\eta\in[0,T]$ to get

$$\begin{aligned}&\int_0^\eta\left(\partial_t\mathbf{H}-\partial_t\mathbf{H}_n^h,\widetilde{\mathbf{P}}_h\mathbf{H}-\mathbf{H}_n^h\right)\\&\qquad+\int_0^\eta\left(\nabla\times\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\widetilde{\mathbf{P}}_h\mathbf{H}-\mathbf{H}_n^h\right)\right)\\&+\int_0^\eta\left(\mathcal{K}_0\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\widetilde{\mathbf{P}}_h\mathbf{H}-\mathbf{H}_n^h\right)\right)=\int_0^\eta\left(\mathbf{F}-\overline{\mathbf{F}}_n,\widetilde{\mathbf{P}}_h\mathbf{H}-\mathbf{H}_n^h\right).\end{aligned}$$

We rearrange the terms by adding $\pm\mathbf{H}$ and $\pm\overline{\mathbf{H}}_n^h$ to obtain

$$\begin{aligned}&\tfrac{1}{2}\left\|\mathbf{H}(\eta)-\mathbf{H}_n^h(\eta)\right\|^2-\tfrac{1}{2}\left\|\mathbf{H}_0-\widetilde{\mathbf{P}}_h\mathbf{H}_0\right\|^2+\int_0^\eta\left\|\nabla\times\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\right\|^2\\&=\int_0^\eta\left(\partial_t\mathbf{H}-\partial_t\mathbf{H}_n^h,\mathbf{H}-\widetilde{\mathbf{P}}_h\mathbf{H}\right)\\&+\int_0^\eta\left(\nabla\times\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\mathbf{H}-\widetilde{\mathbf{P}}_h\mathbf{H}\right)\right)\\&+\int_0^\eta\left(\nabla\times\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\mathbf{H}_n^h-\overline{\mathbf{H}}_n^h\right)\right)\\&+\int_0^\eta\left(\mathcal{K}_0\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\mathbf{H}-\widetilde{\mathbf{P}}_h\mathbf{H}\right)\right)\\&+\int_0^\eta\left(\mathcal{K}_0\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\overline{\mathbf{H}}_n^h-\mathbf{H}\right)\right)\\&+\int_0^\eta\left(\mathcal{K}_0\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\times\left(\mathbf{H}_n^h-\overline{\mathbf{H}}_n^h\right)\right)\\&+\int_0^\eta\left(\mathbf{F}-\overline{\mathbf{F}}_n,\widetilde{\mathbf{P}}_h\mathbf{H}-\mathbf{H}\right)+\int_0^\eta\left(\mathbf{F}-\overline{\mathbf{F}}_n,\mathbf{H}-\mathbf{H}_n^h\right)=:\sum_{i=1}^8 S_i.\end{aligned}$$

The following inequality is useful during the term by term estimation of the previous equality

$$\left\|\mathbf{H}_n^h(t) - \overline{\mathbf{H}}_n^h(t)\right\| \leqslant \tau \left\|\partial_t \mathbf{H}_n^h(t)\right\| \qquad \text{for } t \in [0,T].$$

Using Hölder's inequality, Lemma 4.2.1(iii) and Lemma 4.7.1(ii) gives that

$$S_1 \leqslant \sqrt{\int_0^\eta \|\partial_t \mathbf{H} - \partial_t \mathbf{H}_n^h\|^2} \sqrt{\int_0^\eta \left\|\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right\|^2} \lesssim \sqrt{\int_0^\eta \left\|\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right\|^2}.$$

Using Young's inequality gives

$$S_2 \leqslant \varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right)\right\|^2.$$

Moreover, applying Lemma 4.7.1(ii) yields that

$$\begin{aligned} S_3 &\leqslant \varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 + C_\varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H}_n^h - \overline{\mathbf{H}}_n^h\right)\right\|^2 \\ &\leqslant \varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 + C_\varepsilon \tau. \end{aligned}$$

For the term S_4, we get that

$$S_4 \overset{(4.4)}{\lesssim} \int_0^\eta \left\|\mathbf{H} - \overline{\mathbf{H}}_n^h\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right)\right\|^2.$$

Adding $\pm \mathbf{H}_n^h$ in the first term of the RHS of the previous inequality and employing Lemma 4.7.1(ii) gives

$$S_4 \lesssim \tau^2 + \int_0^\eta \left\|\mathbf{H} - \mathbf{H}_n^h\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \widetilde{\mathbf{P}}_h \mathbf{H}\right)\right\|^2.$$

In the same way as for the term S_4, we get thanks to Lemma 4.7.1(ii) that

$$\begin{aligned} S_5 &\overset{(4.4)}{\leqslant} C_\varepsilon \int_0^\eta \left\|\mathbf{H} - \overline{\mathbf{H}}_n^h\right\|^2 + \varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \\ &\leqslant C_\varepsilon \tau^2 + C_\varepsilon \int_0^\eta \left\|\mathbf{H} - \mathbf{H}_n^h\right\|^2 + \varepsilon \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2. \end{aligned}$$

and

$$\begin{aligned} S_6 &\overset{(4.4)}{\lesssim} \int_0^\eta \left\|\mathbf{H} - \overline{\mathbf{H}}_n^h\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H}_n^h - \overline{\mathbf{H}}_n^h\right)\right\|^2 \\ &\lesssim \tau^2 + \int_0^\eta \left\|\mathbf{H} - \mathbf{H}_n^h\right\|^2 + \tau \\ &\lesssim \tau + \int_0^\eta \left\|\mathbf{H} - \mathbf{H}_n^h\right\|^2. \end{aligned}$$

The terms S_7 and S_8 can be estimated due to the Lipschitz continuity of $\mathbf{F}$ by

$$S_7 \lesssim \int_0^\eta \left\|\mathbf{F} - \overline{\mathbf{F}}_n\right\|^2 + \int_0^\eta \left\|\widetilde{\mathbf{P}}_h\mathbf{H} - \mathbf{H}\right\|^2 \lesssim \tau^2 + \int_0^\eta \left\|\mathbf{H} - \widetilde{\mathbf{P}}_h\mathbf{H}\right\|^2$$

and

$$S_8 \lesssim \tau^2 + \int_0^\eta \left\|\mathbf{H} - \mathbf{H}_n^h\right\|^2.$$

Fixing a sufficiently small $\varepsilon > 0$, an application of the Grönwall argument concludes the proof.

(ii) The only difference with part (i) of the proof is the handling of the term S_1. Integration by parts gives

$$\begin{aligned} S_1 &= \left(\mathbf{H}(t) - \mathbf{H}_n^h(t), \mathbf{H}(t) - \widetilde{\mathbf{P}}_h\mathbf{H}(t)\right)\Big|_0^\eta \\ &\quad - \int_0^\eta \left(\mathbf{H}(t) - \mathbf{H}_n^h(t), \partial_t\left(\mathbf{H}(t) - \widetilde{\mathbf{P}}_h\mathbf{H}(t)\right)\right) \\ &\leqslant \varepsilon\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + C_\varepsilon\left\|\mathbf{H}(\eta) - \widetilde{\mathbf{P}}_h\mathbf{H}(\eta)\right\|^2 + \left\|\mathbf{H}_0 - \widetilde{\mathbf{P}}_h\mathbf{H}_0\right\|^2 \\ &\quad + C\int_0^\eta \left\|\mathbf{H} - \mathbf{H}_n^h\right\|^2 + C\int_0^\eta \left\|\partial_t\left(\mathbf{H} - \widetilde{\mathbf{P}}_h\mathbf{H}\right)\right\|^2. \end{aligned}$$

The rest of the proof follows closely the lines of (i).

(iii) The term τ^2 can be obtained by an application of Lemma 4.7.1(iii) instead of Lemma 4.7.1(ii) on the terms S_3 and S_6. □

4.7.1 Example: Nédélec's first family of curl-conforming finite elements of first order

Due to their practical importance, in the first example the lowest order Nédélec edge elements are considered, see Example 2.13.5. The finite element space $\mathbf{V}^h$ is then given by

$$\mathbf{V}^h = \{\mathbf{v}^h \in \mathbf{H}(\mathbf{curl};\Omega) \,:\, \mathbf{v}^h\big|_K(\mathbf{x}) = \mathbf{a}_K + \mathbf{b}_K \times \mathbf{x}, \quad \forall K \in \mathcal{T}_h\},$$

where $\mathbf{a}_K$ and $\mathbf{b}_K$ are constants in $\mathbb{R}^3$. Let us denote by $\mathbf{r}_h$ the interpolation operator valued in $\mathbf{V}_0^h$, defined element by element using $\mathbf{r}_h\mathbf{u}|_K = \mathbf{r}_K\mathbf{u}$ for all $K \in \mathcal{T}_h$, with $\mathbf{r}_K$ the element-wise interpolant given by

$$\int_e (\mathbf{u} - \mathbf{r}_K\mathbf{u}) \cdot \hat{\boldsymbol{\tau}} = 0, \qquad \text{for all edges } e \text{ of } K.$$

Unfortunately, the integrals appearing in this definition are not well defined for functions from $\mathbf{H}(\mathbf{curl};\Omega)$. The interpolation operator $\mathbf{r}_h$ is defined in

$$\mathbf{H}^s(\mathbf{curl}, \Omega) := \{\mathbf{u} \in \mathbf{H}^s(\Omega) : \nabla \times \mathbf{u} \in \mathbf{H}^s(\Omega)\}$$

for any $s > \frac{1}{2}$ [120, Lemma 5.1]. Moreover, there exists a constant $C > 0$, independent of h such that [120, Proposition 5.6]

$$\|\mathbf{H} - \mathbf{r}_h\mathbf{H}\| + \|\nabla \times (\mathbf{H} - \mathbf{r}_h\mathbf{H})\| \leqslant Ch^s \left(\|\mathbf{H}\|_{\mathbf{H}^s(\Omega)} + \|\nabla \times \mathbf{H}\|_{\mathbf{H}^s(\Omega)} \right),$$

for each $\mathbf{H} \in \mathbf{H}^s(\mathbf{curl}, \Omega)$ with $s \in \left(\frac{1}{2}, 1\right]$. Cea's lemma [54] implies that the projection operator $\widetilde{\mathbf{P}}_h$ defined in (4.35) for any $s \in \left(\frac{1}{2}, 1\right]$ has the property

$$\left\|\mathbf{u} - \widetilde{\mathbf{P}}_h\mathbf{u}\right\|_{\mathbf{H}(\mathbf{curl};\Omega)} \leqslant \|\mathbf{u} - \mathbf{r}_h\mathbf{u}\|_{\mathbf{H}(\mathbf{curl};\Omega)} \lesssim h^s \|\mathbf{u}\|_{\mathbf{H}^s(\mathbf{curl},\Omega)},$$

for all $\mathbf{u} \in \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)$. Now, the following corollary of Theorem 4.7.2 can be stated without proof.

Corollary 4.7.1. *Take* $s \in \left(\frac{1}{2}, 1\right]$. *Let* $\mathbf{F} \in \mathrm{Lip}([0,T], \mathbf{L}^2(\Omega))$.

(i) Suppose that $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)$ *and that the weak solution* $\mathbf{H}$ *of* (4.1) *satisfies*

$$\mathbf{H} \in \mathbf{L}^2\left((0,T), \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)\right).$$

Then there exists a constant C independent of both the time step τ and the mesh size h such that

$$\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \leqslant C\left(\tau + h^s\right)$$

is valid for any $\eta \in [0,T]$.

(ii) Suppose that $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)$ *and that the weak solution* $\mathbf{H}$ *of* (4.1) *satisfies*

$$\mathbf{H} \in \mathrm{H}^1\left((0,T), \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)\right).$$

Then there exists a constant C independent of both the time step τ and the mesh size h such that

$$\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + \int_0^\eta \left\|\nabla \times \left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \leqslant C\left(\tau + h^{2s}\right)$$

is valid for any $\eta \in [0,T]$.

(iii) If the initial condition also satisfies $\nabla \times \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)$ *and* $\mathcal{K}_0 \star \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega) \cap \mathbf{H}^s(\mathbf{curl},\Omega)$, *then the estimates in (i) and (ii) are satisfied with τ^2 instead of τ.*

Please note that the positive constant C in these estimates is of the form Ce^{CT}.

Thus, if $\tau \to 0$ and $h \to 0$, the convergence of the Rothe sequence $\mathbf{H}_n^h$ to the unique weak solution $\mathbf{H}$ of problem (4.1) in $\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$ is proved.

4.7.2 Higher regularity

The error estimates in the previous section have been obtained using a priori estimates that were based on Grönwall's argument. The convergence rates are of order $\mathcal{O}(\tau, h) = e^{CT}(\tau + h)$ in the space $\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$ under appropriate conditions. To get rid of the exponential character of this constant, the use of Grönwall's lemma should be avoided. This is again done by the incorporation of the curl operator $\nabla \times \mathbf{J}_s$ into the convolution kernel $\mathcal{K}$, see Lemma 4.5.1, under the assumption that $\mathbf{H} \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$. The solution of problem (4.1) also satisfies problem (4.19). Now, $\mathbf{V}_0^h$ is a finite dimensional subspace of $\mathbf{H}_0^1(\Omega)$. The linear bounded operator $\overline{\mathbf{P}}_h : \mathbf{H}_0^1(\Omega) \to \mathbf{V}_0^h$ is defined such that if $\mathbf{u} \in \mathbf{H}_0^1(\Omega)$, then $\overline{\mathbf{P}}_h \mathbf{u} \in \mathbf{V}_0^h$ satisfies

$$(\mathbf{u}, \mathbf{v}_h) + (\nabla \mathbf{u}, \nabla \mathbf{v}_h) = \left(\overline{\mathbf{P}}_h \mathbf{u}, \mathbf{v}_h\right) + \left(\nabla \overline{\mathbf{P}}_h \mathbf{u}, \nabla \mathbf{v}_h\right)$$

for all $\mathbf{v}_h \in \mathbf{V}_0^h$. The following fully discrete linear recurrent scheme is proposed to approximate (4.19): find $\mathbf{h}_i^h \in \mathbf{V}_0^h$ such that

$$\left\{ \begin{array}{rl} \left(\delta \mathbf{h}_i^h, \boldsymbol{\varphi}^h\right) + \left(\nabla \mathbf{h}_i^h, \nabla \boldsymbol{\varphi}^h\right) & \\ + \left(\mathcal{K} \star \mathbf{h}_i^h, \boldsymbol{\varphi}^h\right) & = \left(\mathbf{P}_h \mathbf{F}_i, \boldsymbol{\varphi}^h\right) = \left(\mathbf{F}_i, \boldsymbol{\varphi}^h\right), \\ \mathbf{h}_0^h & = \overline{\mathbf{P}}_h \mathbf{H}_0 \end{array} \right. \tag{4.38}$$

is satisfied for all $\boldsymbol{\varphi}^h \in \mathbf{V}_0^h$. Due to the positive definiteness of $\mathcal{K}$, an application of the Lax-Milgram lemma 2.11.1 gives the existence of a unique solution in $\mathbf{V}_0^h$ of (4.38) for any $i = 1, \ldots, n$ and any $\tau > 0$ if $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega)$.

The same stability results are obtained as in Lemma 4.7.1, where the curl-spaces are replaced by analogous $\mathbf{H}^s(\Omega)$-spaces. Now, the use of Grönwall's argument is avoided.

Lemma 4.7.2 (Enhanced stability)**.** *Assume that* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$*,* $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *and* $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any time* $t \in [0,T]$*.*

(i) Let $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega)$*. Then, there exists a positive constant* C *such that for all* $\tau > 0$ *it holds that*

$$\max_{1 \leqslant i \leqslant n} \left\|\mathbf{h}_i^h\right\|^2 + \sum_{i=1}^n \left\|\mathbf{h}_i^h - \mathbf{h}_{i-1}^h\right\|^2 + \sum_{i=1}^n \left\|\nabla \mathbf{h}_i^h\right\|^2 \tau \leqslant C.$$

(ii) If $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega)$*, then for all* $\tau > 0$ *it holds that*

$$\max_{1 \leqslant i \leqslant n} \left\|\nabla \mathbf{h}_i^h\right\|^2 + \sum_{i=1}^n \left\|\nabla \mathbf{h}_i^h - \nabla \mathbf{h}_{i-1}^h\right\|^2 + \sum_{i=1}^n \left\|\delta \mathbf{h}_i^h\right\|^2 \tau \leqslant C.$$

(iii) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$*,* $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and* $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$*, then for all* $\tau > 0$ *it holds that*

$$\max_{1 \leqslant i \leqslant n} \left\|\delta \mathbf{h}_i^h\right\|^2 + \sum_{i=1}^n \left\|\delta \mathbf{h}_i^h - \delta \mathbf{h}_{i-1}^h\right\|^2 + \sum_{i=1}^n \left\|\nabla \delta \mathbf{h}_i^h\right\|^2 \tau \leqslant C.$$

Proof. (i) Set $\boldsymbol{\varphi}^h = \mathbf{h}_i^h$ in (4.38). Multiply the result by τ and sum it up for $i = 1, \ldots, j$ ($1 \leqslant j \leqslant n$) to get

$$\sum_{i=1}^{j} \left(\delta \mathbf{h}_i^h, \mathbf{h}_i^h\right) \tau + \sum_{i=1}^{j} \left\|\nabla \mathbf{h}_i^h\right\|^2 \tau + \sum_{i=1}^{j} \left(\mathcal{K} \star \mathbf{h}_i^h, \mathbf{h}_i^h\right) \tau = \sum_{i=1}^{j} \left(\mathbf{F}_i, \mathbf{h}_i^h\right) \tau.$$

The use of Grönwall's argument can be avoided by employing the positive definiteness of $\mathcal{K}$ and Friedrichs inequality. Indeed, it holds that

$$\sum_{i=1}^{j} \left(\mathcal{K} \star \mathbf{h}_i^h, \mathbf{h}_i^h\right) \tau \geqslant 0$$

and

$$\begin{aligned}\left|\sum_{i=1}^{j} \left(\mathbf{F}_i, \mathbf{h}_i^h\right) \tau\right| &\leqslant C_\varepsilon \sum_{i=1}^{j} \left\|\mathbf{F}_i\right\|^2 \tau + \varepsilon \sum_{i=1}^{j} \left\|\mathbf{h}_i^h\right\|^2 \tau \\ &\leqslant C_\varepsilon \sum_{i=1}^{j} \left\|\mathbf{F}_i\right\|^2 \tau + \varepsilon \sum_{i=1}^{j} \left\|\nabla \mathbf{h}_i^h\right\|^2 \tau.\end{aligned}$$

Fixing ε sufficiently small gives the proof.

(ii) We put $\boldsymbol{\varphi}^h = \delta \mathbf{h}_i^h$ in (4.38). Again, we multiply by τ and sum up for $i = 1, \ldots, j$ ($1 \leqslant j \leqslant n$)

$$\begin{aligned}\sum_{i=1}^{j} \left\|\delta \mathbf{h}_i^h\right\|^2 \tau + \sum_{i=1}^{j} \left(\nabla \mathbf{h}_i^h, \nabla \mathbf{h}_i^h - \nabla \mathbf{h}_{i-1}^h\right) + \sum_{i=1}^{j} \left(\mathcal{K} \star \mathbf{h}_i^h, \delta \mathbf{h}_i^h\right) \tau& \\ = \sum_{i=1}^{j} \left(\mathbf{F}_i, \delta \mathbf{h}_i^h\right) \tau.&\end{aligned}$$

Using (i), we obtain

$$\begin{aligned}\left|\sum_{i=1}^{j} \left(\mathcal{K} \star \mathbf{h}_i^h, \delta \mathbf{h}_i^h\right) \tau\right| &\overset{(4.21)}{\leqslant} C_\varepsilon \sum_{i=1}^{j} \left\|\nabla \mathbf{h}_i^h\right\|^2 \tau + \varepsilon \sum_{i=1}^{j} \left\|\delta \mathbf{h}_i^h\right\|^2 \tau \\ &\leqslant C_\varepsilon + \varepsilon \sum_{i=1}^{j} \left\|\delta \mathbf{h}_i^h\right\|^2 \tau.\end{aligned}$$

(iii) The proof is the same as in Lemma 4.7.1(iii). Now, the following compatibility condition is needed

$$\delta \mathbf{h}_0^h := \mathbf{P}_h \mathbf{F}(0) - \mathbf{P}_h \left(\Delta \mathbf{H}_0\right) - \mathbf{P}_h \left(\mathcal{K} \star \mathbf{H}_0\right).$$

$\square$

Using the Rothe's functions, the variational formulation (4.38) can be rewritten for $\boldsymbol{\varphi}^h \in \mathbf{V}_0^h$ and a.a. $t \in (0,T)$ as

$$\left(\partial_t \mathbf{H}_n^h(t), \boldsymbol{\varphi}^h\right) + \left(\nabla \overline{\mathbf{H}}_n^h(t), \nabla \boldsymbol{\varphi}^h\right) + \left(\mathcal{K} \star \overline{\mathbf{H}}_n^h(t), \boldsymbol{\varphi}^h\right) = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}^h\right). \quad (4.39)$$

The following error estimates have smaller constant C in comparison with the constants appearing in Theorem 4.7.2 because Grönwall's argument is avoided thanks to the positive definiteness of $\mathcal{K}$.

Theorem 4.7.3. *Suppose that* $\mathbf{F} \in \operatorname{Lip}([0,T], \mathbf{L}^2(\Omega))$, $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *and* $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any time* $t \in [0,T]$.

(i) Let the weak solution $\mathbf{H}$ *of* (4.1) *at time* t *and the initial condition* $\mathbf{H}_0$ *satisfy* $\mathbf{H}(t), \mathbf{H}_0 \in \mathbf{H}_0^1(\Omega)$. *Then for any* $\tau < \tau_0$, *there exists a constant* C *independent of both the time step* τ *and the mesh size* h, *such that*

$$\begin{aligned}\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 &+ \int_0^\eta \left\|\nabla\left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \\ &\leqslant C\left(\tau + \left\|\mathbf{H}_0 - \overline{\mathbf{P}}_h \mathbf{H}_0\right\|^2 + \sqrt{\int_0^\eta \left\|\mathbf{H} - \overline{\mathbf{P}}_h \mathbf{H}\right\|^2}\right. \\ &\qquad \left. + \int_0^\eta \left\|\mathbf{H} - \overline{\mathbf{P}}_h \mathbf{H}\right\|^2 + \int_0^\eta \left\|\nabla\left(\mathbf{H} - \overline{\mathbf{P}}_h \mathbf{H}\right)\right\|^2\right)\end{aligned}$$

is valid for any $\eta \in [0,T]$;

(ii) Let the weak solution $\mathbf{H}$ *of* (4.1) *at time* t *and the initial condition* $\mathbf{H}_0$ *satisfy* $\mathbf{H}(t), \partial_t \mathbf{H}(t), \mathbf{H}_0 \in \mathbf{H}_0^1(\Omega)$. *Then for any* $\tau < \tau_0$, *there exists a constant* C *independent of both the time step* τ *and the mesh size* h, *such that*

$$\begin{aligned}&\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + \int_0^\eta \left\|\nabla\left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \\ &\leqslant C\left(\tau + \left\|\mathbf{H}_0 - \overline{\mathbf{P}}_h \mathbf{H}_0\right\|^2 + \left\|\mathbf{H}(\eta) - \overline{\mathbf{P}}_h \mathbf{H}(\eta)\right\|^2 + \int_0^\eta \left\|\mathbf{H} - \overline{\mathbf{P}}_h \mathbf{H}\right\|^2\right. \\ &\qquad \left. + \int_0^\eta \left\|\partial_t\left(\mathbf{H} - \overline{\mathbf{P}}_h \mathbf{H}\right)\right\|^2 + \int_0^\eta \left\|\nabla\left(\mathbf{H} - \overline{\mathbf{P}}_h \mathbf{H}\right)\right\|^2\right)\end{aligned}$$

is valid for any $\eta \in [0,T]$;

(iii) If the initial condition satisfies $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$, *then the estimates in (i) and (ii) are satisfied with* τ^2 *instead of* τ.

Proof. (i) We subtract (4.39) from (4.22) for $\boldsymbol{\varphi} = \boldsymbol{\varphi}^h$. We set $\boldsymbol{\varphi}^h = \overline{\mathbf{P}}_h \mathbf{H}(t) - \mathbf{H}_n^h(t)$ and integrate in time over $(0,\eta)$ for $\eta \in [0,T]$ and rearrange the terms to obtain

$$\begin{aligned}
&\tfrac{1}{2}\left\|\mathbf{H}(\eta)-\mathbf{H}_n^h(\eta)\right\|^2-\tfrac{1}{2}\left\|\mathbf{H}_0-\overline{\mathbf{P}}_h\mathbf{H}_0\right\|^2\\
&\quad+\int_0^\eta\left\|\nabla\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\right\|^2+\int_0^\eta\left(\mathcal{K}\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\\
&\qquad=\int_0^\eta\left(\partial_t\mathbf{H}-\partial_t\mathbf{H}_n^h,\mathbf{H}-\overline{\mathbf{P}}_h\mathbf{H}\right)\\
&+\int_0^\eta\left(\nabla\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\left(\mathbf{H}-\overline{\mathbf{P}}_h\mathbf{H}\right)\right)+\int_0^\eta\left(\nabla\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\nabla\left(\mathbf{H}_n^h-\overline{\mathbf{H}}_n^h\right)\right)\\
&\quad+\int_0^\eta\left(\mathcal{K}\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\mathbf{H}-\overline{\mathbf{P}}_h\mathbf{H}\right)+\int_0^\eta\left(\mathcal{K}\star\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right),\mathbf{H}_n^h-\overline{\mathbf{H}}_n^h\right)\\
&\qquad+\int_0^\eta\left(\mathbf{F}-\overline{\mathbf{F}}_n,\overline{\mathbf{P}}_h\mathbf{H}-\mathbf{H}\right)+\int_0^\eta\left(\mathbf{F}-\overline{\mathbf{F}}_n,\mathbf{H}-\mathbf{H}_n^h\right)=:\sum_{i=1}^7 S_i.
\end{aligned}$$

The terms S_1, S_2, S_3, S_6 and S_7 can be handled in the same way as in Theorem 4.7.2. For the others terms, we get that

$$S_4\overset{(4.21)}{\leqslant}\varepsilon\int_0^\eta\left\|\nabla\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\right\|^2+C_\varepsilon\int_0^\eta\left\|\mathbf{H}-\overline{\mathbf{P}}_h\mathbf{H}\right\|^2.$$

and

$$\begin{aligned}
S_5&\overset{(4.21)}{\leqslant}\varepsilon\int_0^\eta\left\|\nabla\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\right\|^2+C_\varepsilon\int_0^\eta\left\|\mathbf{H}_n^h-\overline{\mathbf{H}}_n^h\right\|^2\\
&\leqslant\varepsilon\int_0^\eta\left\|\nabla\left(\mathbf{H}-\overline{\mathbf{H}}_n^h\right)\right\|^2+C_\varepsilon\tau.
\end{aligned}$$

Fixing a sufficiently small $\varepsilon>0$ concludes the proof.

(ii) and (iii) The proof follows the same lines as in Theorem 4.7.2(ii) and (iii). □

4.7.2.1 Example: Lagrangian finite elements

In this example, the first-order Lagrange finite elements for the space discretization are considered. The finite element space $\mathbf{V}^h$ is now given by $\mathbf{V}^h=\{\mathbf{v}^h\in\mathbf{H}^1(\Omega)\,:\,\mathbf{v}^h\big|_K\in\mathbf{P}_1(K),\quad\forall K\in\mathcal{T}_h\}$, with $\mathbf{P}_1(K)$ the space of componentwise first-order polynomials. The coefficients of these polynomials are determined by the degrees of freedom $\mathbf{v}^h(\mathbf{a}_i)$ with $\mathbf{a}_i, i=1,\ldots,4$, the vertices of K. Note that $\mathbf{V}_0^h=\{\mathbf{v}^h\in\mathbf{V}^h:\mathbf{v}^h=\mathbf{0}\text{ on }\partial\Omega\}$. The corresponding interpolation operator is denoted by π_h. The Sobolev Embedding theorem in $\mathbb{R}^3$ [131, Theorem 7.57] implies that $\mathbf{H}^s(\Omega)\subset\mathbf{C}(\overline{\Omega})$ if $s>\frac{3}{2}$. Thus $\pi_h:\mathbf{H}^s(\Omega)\to\mathbf{V}_0^h, s>\frac{3}{2}$, to ensure that the vertex values are well defined. Then, [39, Theorem 5.48] gives that there exists a constant $C>0$ independent of h such that

$$\left\|\mathbf{u}-\pi_h\mathbf{u}\right\|_{\mathbf{H}^1(\Omega)}\leqslant Ch^{s-1}\left\|\mathbf{u}\right\|_{\mathbf{H}^s(\Omega)},\tag{4.40}$$

for each $\mathbf{u} \in \mathbf{H}^s(\Omega)$ with $\frac{3}{2} < s \leqslant 2$. This section finishes with the following corollary.

Corollary 4.7.2. *Take* $s \in \left(\frac{3}{2}, 2\right]$. *Let* $\mathbf{F} \in \mathrm{Lip}\left([0,T], \mathbf{L}^2(\Omega)\right)$, $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *and* $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any time* $t \in [0,T]$.

(i) Suppose that $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^s(\Omega)$ *and that the weak solution* $\mathbf{H}$ *of* (4.1) *satisfies*

$$\mathbf{H} \in \mathbf{L}^2\left((0,T), \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^s(\Omega)\right).$$

Then the error estimate

$$\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + \int_0^\eta \left\|\nabla\left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \lesssim \tau + h^{s-1}$$

is valid for any $\eta \in [0,T]$.

(ii) Suppose that $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^s(\Omega)$ *and that the weak solution* $\mathbf{H}$ *of* (4.1) *satisfies*

$$\mathbf{H} \in \mathrm{H}^1\left((0,T), \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^s(\Omega)\right).$$

Then the error estimate

$$\left\|\mathbf{H}(\eta) - \mathbf{H}_n^h(\eta)\right\|^2 + \int_0^\eta \left\|\nabla\left(\mathbf{H} - \overline{\mathbf{H}}_n^h\right)\right\|^2 \lesssim \tau + h^{2(s-1)}$$

is valid for any $\eta \in [0,T]$.

(iii) If the initial condition satisfies $\nabla \times \mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^s(\Omega)$ *and* $\mathcal{K}_0 \star \mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^s(\Omega)$, *then the estimates in (i) and (ii) are satisfied with* τ^2 *instead of* τ.

4.8 Conclusion

A vectorial nonlocal linear parabolic problem (4.1) in terms of the magnetic field with applications in superconductors of type-I has been studied. This model has been obtained from the eddy current version of the Maxwell equations, the two-fluid model of London and London, and the nonlocal representation of the superconductive current by Eringen. The nonlocal term has been given by a space convolution with a singular kernel.

Two time-discrete numerical schemes based on backward Euler's method have been developed to approximate the solution of problem (4.1). The existence of a weak solution for each time step has been shown. Also the convergence of the method has been discussed and error estimates have been derived. In the first scheme, the convolution has been taken implicitly (from the actual time step). In the second one, the convolution has been taken explicitly (from the previous time step). This second scheme has been considered, because it is easier to implement than the first scheme and it gives the same order of convergence. For both schemes, the error estimates for the time discretization have been obtained using a priori estimates that were based on Grönwall's argument. The convergence rates are of order $\mathcal{O}(\tau) = e^{CT}\tau$ in the space $\mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T),\mathbf{H}(\mathbf{curl};\Omega)\right)$ under appropriate conditions, where τ is the discretization parameter. To get rid of the exponential (in time) character of this constant, the use of Grönwall's lemma should be avoided.

For this reason, a new convolution kernel has been derived under the assumption that the normal component of the unknown vector field equals zero on the boundary of the superconductor. With the help of the additional assumption, it has been demonstrated that under higher regularity the solution of the original model satisfies a simpler problem which is easier to implement. Both time-discrete schemes stay valid. One major advantage is the positive definiteness of the kernel. Using this property, better error estimates have been obtained for the implicit scheme (the convergence rate is of order $\mathcal{O}(\tau) = C\tau$).

A numerical experiment for the semi-implicit scheme supports the obtained theoretical results. Also the convergence of a fully discrete finite element scheme (4.36) to the solution of problem (4.1) has been shown. In a similarly way to the time-discrete scheme, it has been demonstrated how to improve the error estimates under higher regularity.

5

Nonlocal hyperbolic problem for type-I superconductivity

This chapter is based on the article [132], which is published in *Journal of Mathematical Analysis and Applications*.

The aims of this chapter are to address the well-posedness of the following hyperbolic problem in terms of the magnetic field $\mathbf{H}$

$$\left\{\begin{array}{rll} \partial_{tt}\mathbf{H} + \partial_t\mathbf{H} + \nabla\times\nabla\times\mathbf{H} + \nabla\times(\mathcal{K}_0\star\mathbf{H}) & = \mathbf{F} & \text{in } Q_T, \\ \mathbf{H}\times\boldsymbol{\nu} & = \mathbf{0} & \text{on } \Sigma_T, \\ \mathbf{H}(\mathbf{x},0) & = \mathbf{H}_0 & \text{in } \Omega, \\ \partial_t\mathbf{H}(\mathbf{x},0) & = \mathbf{H}_0' & \text{in } \Omega, \end{array}\right. \tag{5.1}$$

to design a scheme for its numerical approximation and to derive error estimates for the time discretization.

The domain $\Omega\subset\mathbb{R}^3$ occupying a type-I superconductor is a bounded Lipschitz domain. Note that $Q_T=\Omega\times(0,T]$ and $\Sigma_T=\partial\Omega\times(0,T]$, with T the final time. This problem is obtained by setting $\tilde{\delta}=\epsilon=\mu=\sigma=1$ (without loss of generality) in (3.11), where

$$(\mathcal{K}_0\star\mathbf{H})(\mathbf{x},t) = -\int_\Omega \sigma_0\left(|\mathbf{x}-\mathbf{x}'|_{\mathrm{e}}\right)(\mathbf{x}-\mathbf{x}')\times\mathbf{H}(\mathbf{x}',t)\ \mathrm{d}\mathbf{x}',$$

with $\sigma_0 : (0, \infty) \to \mathbb{R}$ defined by

$$\sigma_0(s) = \begin{cases} \frac{\widetilde{C}}{2s^2} \exp\left(-\frac{s}{r_0}\right) & s < r_0, \\ 0 & s \geqslant r_0. \end{cases}$$

The parameters $\widetilde{C}$ and r_0 depend on the material under consideration. A source term $\mathbf{F}$ is added in the right-hand side. To obtain the magnetic boundary condition in (4.1), it is assumed that the magnetic field outside the domain Ω equals zero [39, p. 8]. The linear time-dependent full Maxwell's equations and related models are studied in several papers [133–136]. The main difference in the analysis of problem (5.1), in comparison with the available results, is caused by the nonlocal term in (5.1).

The techniques applied in this chapter are similar to the techniques presented in Chapter 4. However, each step in the solution process is more challenging and complicated due to the hyperbolicity of the problem. The uniqueness of a solution to problem (5.1) is studied in Section 5.1. The well-posedness of the problem is shown in Section 5.2. A time-discrete numerical scheme is developed. The existence of a weak solution for each time step is shown. Also the convergence of the method is discussed and error estimates are derived. A modified scheme is considered in Section 5.3. Finally, under additional assumptions, it is demonstrated that the solution of problem (5.1) satisfies a simpler equation, which is shortly described and analysed in Subsection 5.4.

5.1 Uniqueness of a solution

The variational formulation of (5.1) is:

Given $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ and $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, find $\mathbf{H}(t) \in \mathbf{H}_0(\mathbf{curl};\Omega)$ with $\partial_t \mathbf{H}(t) \in \mathbf{L}^2(\Omega)$ and $\partial_{tt}\mathbf{H}(t) \in \mathbf{H}_0^{-1}(\mathbf{curl};\Omega) := \mathbf{H}_0(\mathbf{curl};\Omega)^*$ such that

$$\begin{aligned}(\partial_{tt}\mathbf{H}(t), \boldsymbol{\varphi}) &+ (\partial_t \mathbf{H}(t), \boldsymbol{\varphi}) \\ &+ (\nabla \times \mathbf{H}(t), \nabla \times \boldsymbol{\varphi}) + ((\mathcal{K}_0 \star \mathbf{H})(t), \nabla \times \boldsymbol{\varphi}) = (\mathbf{F}(t), \boldsymbol{\varphi}), \end{aligned} \tag{5.2}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$ and a.a. $t \in [0,T]$.

The estimates on the kernel $\mathcal{K}_0$ from Section 4.1 stay valid, i.e.

$$|(\mathcal{K}_0 \star \mathbf{H})(\mathbf{x},t)| \leqslant C(q) \|\mathbf{H}(t)\|_q, \qquad \forall q > \frac{3}{2} \text{ and } (\mathbf{x},t) \in Q_T \tag{5.3}$$

and

$$(\mathcal{K}_0 \star \mathbf{H}_1, \nabla \times \mathbf{H}_2) \leqslant C_\varepsilon \|\mathbf{H}_1\|^2 + \varepsilon \|\nabla \times \mathbf{H}_2\|^2 \tag{5.4}$$

for all $\mathbf{H}_1 \in \mathbf{L}^2(\Omega)$ and $\mathbf{H}_2 \in \mathbf{H}(\mathbf{curl};\Omega)$. The position of the positive constants ε and C_ε can be interchanged.

The natural stability of the solution $\mathbf{H}$ of (5.1) is addressed in the following theorem.

Theorem 5.1.1 (Stability)**.** *Suppose that* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and that* $\mathbf{H}$ *is the solution to* (5.1)*.*

(i) If $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\| \nabla \times \int_0^t \mathbf{H} \right\|^2 \leqslant C;$$

(ii) If $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\partial_t \mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \|\nabla \times \mathbf{H}(t)\|^2 \leqslant C;$$

(iii) If $\nabla \cdot \mathbf{F}(t) = \nabla \cdot \mathbf{H}_0 = \nabla \cdot \mathbf{H}_0' = 0$ *for any* $t \in [0,T]$*, then* $\nabla \cdot \mathbf{H}(t) = 0$ *for any* $t \in [0,T]$*. Moreover, we have that*

$$\int_0^T \|\partial_{tt}\mathbf{H}\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C;$$

(iv) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$*,* $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$*,* $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega)$*,* $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$*,* $\mathbf{H}_0' \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\partial_{tt}\mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \|\nabla \times \partial_t \mathbf{H}(t)\|^2 \leqslant C.$$

Proof. (i) We first integrate (5.2) in time over $t \in (0,\xi) \subset (0,T)$ and then we set $\boldsymbol{\varphi} = \mathbf{H}(\xi)$. We integrate the result over the time variable $\xi \in (0,\eta) \subset (0,T)$ to get

$$\begin{aligned} &\frac{\|\mathbf{H}(\eta)\|^2}{2} - \frac{\|\mathbf{H}_0\|^2}{2} + \int_0^\eta \|\mathbf{H}(\xi)\|^2 \\ &+ \int_0^\eta \left(\nabla \times \int_0^\xi \mathbf{H}(t), \nabla \times \mathbf{H}(\xi)\right) + \int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi \mathbf{H}(t), \nabla \times \mathbf{H}(\xi)\right) \\ &\quad = \int_0^\eta \left(\int_0^\xi \mathbf{F}(t), \mathbf{H}(\xi)\right) + \int_0^\eta (\mathbf{H}_0', \mathbf{H}(\xi)) + \int_0^\eta (\mathbf{H}_0, \mathbf{H}(\xi)). \end{aligned} \tag{5.5}$$

Integration by parts yields

$$\begin{aligned} &\int_0^\eta \left(\nabla \times \int_0^\xi \mathbf{H}, \nabla \times \mathbf{H}(\xi)\right) \\ &\qquad = \left(\nabla \times \int_0^\eta \mathbf{H}, \int_0^\eta \nabla \times \mathbf{H}\right) - \int_0^\eta \left(\nabla \times \mathbf{H}(\xi), \int_0^\xi \nabla \times \mathbf{H}\right). \end{aligned}$$

Therefore, the third term in the left-hand side of (5.5) can be rewritten as

$$\int_0^\eta \left(\nabla\times\int_0^\xi \mathbf{H}(t), \nabla\times\mathbf{H}(\xi)\right) = \frac{1}{2}\left\|\nabla\times\int_0^\eta \mathbf{H}(t)\right\|^2.$$

Using the integration by parts formula again, we may write

$$\begin{aligned}
&\int_0^\eta \left(\mathcal{K}_0\star\int_0^\xi \mathbf{H}, \nabla\times\mathbf{H}(\xi)\right)\\
&\quad= \left(\mathcal{K}_0\star\int_0^\eta \mathbf{H}, \nabla\times\int_0^\eta \mathbf{H}\right) - \int_0^\eta\left(\mathcal{K}_0\star\mathbf{H}(\xi), \nabla\times\int_0^\xi \mathbf{H}\right)\\
&\quad\overset{(5.4)}{\leqslant} \varepsilon\left\|\nabla\times\int_0^\eta \mathbf{H}\right\|^2 + C_\varepsilon\left\|\int_0^\eta \mathbf{H}(t)\right\|^2\\
&\qquad\qquad +C\int_0^\eta \|\mathbf{H}\|^2 + C\int_0^\eta\left\|\nabla\times\int_0^\xi \mathbf{H}\right\|^2\\
&\quad\leqslant \varepsilon\left\|\nabla\times\int_0^\eta \mathbf{H}\right\|^2 + C_\varepsilon\int_0^\eta \|\mathbf{H}\|^2 + C\int_0^\eta\left\|\nabla\times\int_0^\xi \mathbf{H}\right\|^2.
\end{aligned}$$

Moreover, using the Cauchy and Young inequalities, we obtain

$$\begin{aligned}
\int_0^\eta\left(\int_0^\xi \mathbf{F}, \mathbf{H}(\xi)\right) &\lesssim \int_0^\eta\left\|\int_0^\xi \mathbf{F}\right\|^2 + \int_0^\eta \|\mathbf{H}\|^2\\
&\lesssim \int_0^\eta\int_0^T \|\mathbf{F}\|^2 + \int_0^\eta \|\mathbf{H}\|^2\\
&\lesssim 1 + \int_0^\eta \|\mathbf{H}\|^2
\end{aligned}$$

and

$$\int_0^\eta (\mathbf{H}_0, \mathbf{H}(\xi)) \lesssim 1 + \int_0^\eta \|\mathbf{H}\|^2; \quad \int_0^\eta (\mathbf{H}_0', \mathbf{H}(\xi)) \lesssim 1 + \int_0^\eta \|\mathbf{H}\|^2.$$

Collecting all the estimates gives

$$\begin{aligned}
&\frac{\|\mathbf{H}(\eta)\|^2}{2} + \int_0^\eta \|\mathbf{H}(\xi)\|^2 + \left(\frac{1}{2}-\varepsilon\right)\left\|\nabla\times\int_0^\eta \mathbf{H}(t)\right\|^2\\
&\qquad\qquad \leqslant C + C_\varepsilon\int_0^\eta \|\mathbf{H}\|^2 + C\int_0^\eta\left\|\nabla\times\int_0^\xi \mathbf{H}\right\|^2.
\end{aligned}$$

Choosing a sufficiently small positive ε and involving the Grönwall argument, we obtain the desired result.

(ii) Setting $\boldsymbol{\varphi} = \partial_t \mathbf{H}(t)$ and integrating in time over $t \in (0, \eta) \subset (0, T)$, we get

$$\begin{aligned}\tfrac{1}{2}\left\|\partial_t \mathbf{H}(\eta)\right\|^2 + \int_0^\eta \left\|\partial_t \mathbf{H}\right\|^2 + \tfrac{1}{2}\left\|\nabla\times\mathbf{H}(\eta)\right\|^2 = \tfrac{1}{2}\left\|\mathbf{H}_0'\right\|^2 + \\ \tfrac{1}{2}\left\|\nabla\times\mathbf{H}_0\right\|^2 + \int_0^\eta (\mathbf{F}, \partial_t\mathbf{H}) - \int_0^\eta (\mathcal{K}_0 \star \mathbf{H}, \nabla\times\partial_t\mathbf{H}).\end{aligned}$$

The third term in the RHS can be estimated by

$$\int_0^\eta (\mathbf{F}, \partial_t\mathbf{H}) \leqslant \varepsilon \int_0^\eta \left\|\partial_t\mathbf{H}\right\|^2 + C_\varepsilon \int_0^\eta \left\|\mathbf{F}\right\|^2 \leqslant \varepsilon \int_0^\eta \left\|\partial_t\mathbf{H}\right\|^2 + C_\varepsilon.$$

For the fourth term in the RHS, we obtain, using the integration by parts formula that

$$\begin{aligned}&\int_0^\eta (\mathcal{K}_0 \star \mathbf{H}, \nabla\times\partial_t\mathbf{H}) \\ &= \quad (\mathcal{K}_0\star\mathbf{H}, \nabla\times\mathbf{H})|_0^\eta - \int_0^\eta (\mathcal{K}_0\star\partial_t\mathbf{H}, \nabla\times\mathbf{H}) \\ &\overset{(5.4)}{\leqslant} \quad C + \varepsilon\left\|\nabla\times\mathbf{H}(\eta)\right\|^2 + C_\varepsilon\left\|\mathbf{H}(\eta)\right\|^2 \\ &\qquad\qquad + \varepsilon\int_0^\eta \left\|\partial_t\mathbf{H}\right\|^2 + C_\varepsilon\int_0^\eta\left\|\nabla\times\mathbf{H}\right\|^2 \\ &\overset{(i)}{\leqslant} \quad C_\varepsilon + \varepsilon\left\|\nabla\times\mathbf{H}(\eta)\right\|^2 + \varepsilon\int_0^\eta\left\|\partial_t\mathbf{H}\right\|^2 + C_\varepsilon\int_0^\eta\left\|\nabla\times\mathbf{H}\right\|^2.\end{aligned}$$

Collecting all considerations above and fixing a sufficiently small positive ε, an application of the Grönwall argument concludes the proof.

(iii) Take the divergence of (5.1) and integrate in time over $t \in (0, \xi) \subset (0, T)$ and over $\xi \in (0, \eta) \subset (0, T)$ to arrive at

$$\nabla\cdot\mathbf{H}(\eta) + \int_0^\eta \nabla\cdot\mathbf{H}(\xi) = 0.$$

Taking the absolute value, the second power and integrating over the domain Ω, we get

$$\left\|\nabla\cdot\mathbf{H}(\eta)\right\|^2 \leqslant \int_0^\eta \left\|\nabla\cdot\mathbf{H}\right\|^2.$$

An application of Grönwall's lemma gives $\left\|\nabla\cdot\mathbf{H}(\eta)\right\| = 0$. Hence, $\nabla\cdot\mathbf{H}(\eta) = 0$ a.e. in Ω for all $\eta \in [0, T]$. Now, we rewrite (5.2) for $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$ as follows

$$\begin{aligned}(\partial_{tt}\mathbf{H}(t), \boldsymbol{\varphi}) = (\mathbf{F}(t), \boldsymbol{\varphi}) - (\partial_t\mathbf{H}(t), \boldsymbol{\varphi}) \\ - (\nabla\times\mathbf{H}(t), \nabla\times\boldsymbol{\varphi}) - ((\mathcal{K}_0\star\mathbf{H})(t), \nabla\times\boldsymbol{\varphi}).\end{aligned}$$

The integral in the LHS has to be interpreted in the sense of duality, i.e. seeing $\partial_{tt}\mathbf{H}(t)$ as an operator from $\mathbf{H}_0(\mathbf{curl};\Omega)$ to $\mathbb{R}$. A simple calculation implies

$$|\,(\mathbf{F}(t),\boldsymbol{\varphi})\,| \leqslant \|\mathbf{F}(t)\|\,\|\boldsymbol{\varphi}\|\,, \quad |\,(\nabla\times\mathbf{H}(t),\nabla\times\boldsymbol{\varphi})\,| \leqslant \|\nabla\times\mathbf{H}(t)\|\,\|\nabla\times\boldsymbol{\varphi}\|\,,$$

$$|\,(\partial_t\mathbf{H}(t),\boldsymbol{\varphi})\,| \leqslant \|\partial_t\mathbf{H}(t)\|\,\|\boldsymbol{\varphi}\|\,,\ |\,((\mathcal{K}_0\star\mathbf{H})(t),\nabla\times\boldsymbol{\varphi})\,| \overset{(5.3)}{\lesssim} \|\mathbf{H}(t)\|\,\|\nabla\times\boldsymbol{\varphi}\|\,.$$

Remember that

$$\|\partial_{tt}\mathbf{H}(t)\|_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} = \sup_{\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)} \frac{(\partial_{tt}\mathbf{H}(t),\boldsymbol{\varphi})}{\|\boldsymbol{\varphi}\|_{\mathbf{H}_0(\mathbf{curl};\Omega)}}.$$

Therefore, using (i) and (ii), we deduce that

$$\int_0^T \|\partial_{tt}\mathbf{H}(s)\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)}\ \mathrm{d}s \leqslant C.$$

(iv) First, we differentiate (5.2) with respect to the time variable. Then we set $\boldsymbol{\varphi} = \partial_{tt}\mathbf{H}(t)$ and integrate in time over $t\in(0,\eta)\subset(0,T)$ to get

$$\begin{aligned}\tfrac{1}{2}\|\partial_{tt}\mathbf{H}(\eta)\|^2 + \int_0^\eta \|\partial_{tt}\mathbf{H}\|^2 + \tfrac{1}{2}\|\nabla\times\partial_t\mathbf{H}(\eta)\|^2 &= \tfrac{1}{2}\|\partial_{tt}\mathbf{H}(0)\|^2\\ + \tfrac{1}{2}\|\nabla\times\mathbf{H}_0'\|^2 + \int_0^\eta (\partial_t\mathbf{F},\partial_{tt}\mathbf{H}) &- \int_0^\eta (\mathcal{K}_0\star\partial_t\mathbf{H},\nabla\times\partial_{tt}\mathbf{H})\,.\end{aligned}$$

For the last two terms in the RHS, we follow the same lines as in (ii) when considering $\partial_t\mathbf{H}$ instead of $\mathbf{H}$. We have, using the upper bound (ii), that

$$\begin{aligned}\int_0^\eta (\mathcal{K}_0\star\partial_t\mathbf{H},\nabla\times\partial_{tt}\mathbf{H}) &\leqslant C_\varepsilon + \varepsilon\|\nabla\times\partial_t\mathbf{H}(\eta)\|^2\\ &\quad + \varepsilon\int_0^\eta\|\partial_{tt}\mathbf{H}\|^2 + C_\varepsilon\int_0^\eta\|\nabla\times\partial_t\mathbf{H}\|^2\,.\end{aligned}$$

and

$$\int_0^\eta (\partial_t\mathbf{F},\partial_{tt}\mathbf{H}) \leqslant \varepsilon\int_0^\eta\|\partial_{tt}\mathbf{H}\|^2 + C_\varepsilon.$$

Therefore, we get

$$\begin{aligned}\|\partial_{tt}\mathbf{H}(\eta)\|^2 + \int_0^\eta\|\partial_{tt}\mathbf{H}\|^2 + \|\nabla\times\partial_t\mathbf{H}(\eta)\|^2\\ \leqslant C_\varepsilon + \|\partial_{tt}\mathbf{H}(0)\|^2 + \varepsilon\|\nabla\times\partial_t\mathbf{H}(\eta)\|^2\\ + \varepsilon\int_0^\eta\|\partial_{tt}\mathbf{H}\|^2 + C_\varepsilon\int_0^\eta\|\nabla\times\partial_t\mathbf{H}\|^2\,.\end{aligned}$$

Fixing a small ε and applying Grönwall's argument, we arrive at

$$\max_{t\in[0,T]} \|\partial_{tt}\mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \|\nabla\times\partial_t\mathbf{H}(t)\|^2 + \int_0^T \|\partial_{tt}\mathbf{H}\|^2 \lesssim 1 + \|\partial_{tt}\mathbf{H}(0)\|^2 .$$

To find a bound for $\|\partial_{tt}\mathbf{H}(0)\|$, the variational formulation (5.2) needs to be satisfied at $t=0$, i.e.

$$(\partial_{tt}\mathbf{H}(0),\boldsymbol{\varphi}) + (\mathbf{H}_0',\boldsymbol{\varphi}) + (\nabla\times\mathbf{H}_0,\nabla\times\boldsymbol{\varphi}) \\ + (\mathcal{K}_0\star\mathbf{H}_0,\nabla\times\boldsymbol{\varphi}) = (\mathbf{F}(0),\boldsymbol{\varphi})$$

for all $\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)$ when $\mathbf{H}_0\in\mathbf{H}_0(\mathbf{curl};\Omega)$ and $\mathbf{H}_0'\in\mathbf{L}^2(\Omega)$. Applying Green's theorem in a backward way gives for all $\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)$ that

$$(\partial_{tt}\mathbf{H}(0),\boldsymbol{\varphi}) = (\mathbf{F}(0),\boldsymbol{\varphi}) - (\mathbf{H}_0',\boldsymbol{\varphi}) \\ - (\nabla\times\nabla\times\mathbf{H}_0,\boldsymbol{\varphi}) - (\nabla\times(\mathcal{K}_0\star\mathbf{H}_0),\boldsymbol{\varphi}) .$$

The term $\partial_{tt}\mathbf{H}(0)$ can be seen as a functional on $\mathbf{H}_0(\mathbf{curl};\Omega)$. The RHS is a linear and bounded functional on $\mathbf{H}_0(\mathbf{curl};\Omega)$. This implies that the RHS can be extended to a functional $\widetilde{\partial_{tt}\mathbf{H}(0)}$ on $\mathbf{L}^2(\Omega)$ by the Hahn-Banach theorem. Moreover,

$$\left\|\widetilde{\partial_{tt}\mathbf{H}(0)}\right\| = \sup_{\substack{\boldsymbol{\varphi}\in\mathbf{L}^2(\Omega)\\ \|\boldsymbol{\varphi}\|\leqslant 1}} \left(\widetilde{\partial_{tt}\mathbf{H}(0)},\boldsymbol{\varphi}\right) = \sup_{\substack{\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)\\ \|\boldsymbol{\varphi}\|\leqslant 1}} (\partial_{tt}\mathbf{H}(0),\boldsymbol{\varphi}) \lesssim 1,$$

i.e. $\widetilde{\partial_{tt}\mathbf{H}(0)}\in\mathbf{L}^2(\Omega)$. The density of $(\mathrm{C}_0^\infty(\Omega))^3$ in $\mathbf{H}_0(\mathbf{curl};\Omega)$ and Theorem 2.8.1 imply that

$$\widetilde{\partial_{tt}\mathbf{H}(0)} = \mathbf{F}(0) - \mathbf{H}_0' - \nabla\times\nabla\times\mathbf{H}_0 - \nabla\times(\mathcal{K}_0\star\mathbf{H}_0), \quad \text{a.e. in } \Omega.$$

The proof concludes by identifying $\widetilde{\partial_{tt}\mathbf{H}(0)}$ and $\partial_{tt}\mathbf{H}(0)$. ☐

Theorem 5.1.2 (Uniqueness)**.** *The problem* (5.1) *admits at most one solution* $\mathbf{H}\in\mathrm{C}([0,T],\mathbf{H}_0(\mathbf{curl};\Omega))$ *such that* $\partial_t\mathbf{H}\in\mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$.

Proof. Assume that we have two different solutions $\mathbf{H}_1$ and $\mathbf{H}_2$. Then $\mathbf{H}=\mathbf{H}_1-\mathbf{H}_2$ fulfils (5.2) with $\mathbf{H}_0'=\mathbf{H}_0=\mathbf{F}=\mathbf{0}$. We set $\boldsymbol{\varphi}=\partial_t\mathbf{H}(t)$ and integrate over the time variable $t\in(0,\eta)\subset(0,T)$ to get

$$\tfrac{1}{2}\|\partial_t\mathbf{H}(\eta)\|^2 + \int_0^\eta \|\partial_t\mathbf{H}\|^2 + \tfrac{1}{2}\|\nabla\times\mathbf{H}(\eta)\|^2 = -\int_0^\eta (\mathcal{K}_0\star\mathbf{H},\nabla\times\partial_t\mathbf{H}) .$$

For the term in the RHS, using the integration by parts formula, we obtain that

$$\begin{aligned}
&\int_0^\eta (\mathcal{K}_0 \star \mathbf{H}, \nabla \times \partial_t \mathbf{H}) \\
&= \quad (\mathcal{K}_0 \star \mathbf{H}, \nabla \times \mathbf{H})|_0^\eta - \int_0^\eta (\mathcal{K}_0 \star \partial_t \mathbf{H}, \nabla \times \mathbf{H}) \\
&\overset{(5.4)}{\leqslant} \quad \varepsilon \left\|\nabla \times \mathbf{H}(\eta)\right\|^2 + C_\varepsilon \left\|\mathbf{H}(\eta)\right\|^2 \\
&\qquad\qquad + C \int_0^\eta \left\|\partial_t \mathbf{H}\right\|^2 + C \int_0^\eta \left\|\nabla \times \mathbf{H}\right\|^2 \\
&\leqslant \quad \varepsilon \left\|\nabla \times \mathbf{H}(\eta)\right\|^2 + C_\varepsilon \int_0^\eta \left\|\partial_t \mathbf{H}\right\|^2 + C \int_0^\eta \left\|\nabla \times \mathbf{H}\right\|^2 .
\end{aligned}$$

In the last step, we have used that $\mathbf{H}(\eta) = \displaystyle\int_0^\eta \partial_t \mathbf{H}$ because $\mathbf{H}_0 = \mathbf{0}$. Using the previous estimate, we arrive at

$$\begin{aligned}
\tfrac{1}{2} \left\|\partial_t \mathbf{H}(\eta)\right\|^2 + \int_0^\eta \left\|\partial_t \mathbf{H}\right\|^2 + \left(\tfrac{1}{2} - \varepsilon\right) \left\|\nabla \times \mathbf{H}(\eta)\right\|^2 & \\
\leqslant C_\varepsilon \int_0^\eta \left\|\partial_t \mathbf{H}\right\|^2 + C \int_0^\eta \left\|\nabla \times \mathbf{H}\right\|^2 .&
\end{aligned}$$

Fixing a sufficiently small positive ε and applying Grönwall's argument, we get that $\partial_t \mathbf{H} = \mathbf{0}$ a.e. in Q_T. Therefore, due to $\mathbf{H}_0 = \mathbf{0}$, we have that $\mathbf{H} = \mathbf{0}$ a.e. in Q_T. Thus, $\mathbf{H}_1$ and $\mathbf{H}_2$ are identical. $\square$

5.2 Existence of a solution

Rothe's method is employed to address the existence of a solution to (5.1). The interval $[0, T]$ is divided into n equidistant subintervals $[t_{i-1}, t_i]$ with time step $\tau = \frac{T}{n} < 1$, thus $t_i = i\tau, i = 1, \ldots, n$. With the standard notation for the discretized fields

$$\mathbf{h}_i \approx \mathbf{H}(t_i), \quad \delta \mathbf{h}_i = \frac{\mathbf{h}_i - \mathbf{h}_{i-1}}{\tau}, \quad \delta^2 \mathbf{h}_i = \frac{\delta \mathbf{h}_i - \delta \mathbf{h}_{i-1}}{\tau} = \frac{\mathbf{h}_i}{\tau^2} - \frac{\mathbf{h}_{i-1}}{\tau^2} - \frac{\delta \mathbf{h}_{i-1}}{\tau},$$

the following linear recurrent scheme is proposed to approximate the original problem

$$\left\{ \begin{aligned} \left(\delta^2 \mathbf{h}_i, \boldsymbol{\varphi}\right) + (\delta \mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) & \\ + (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) & = (\mathbf{F}_i, \boldsymbol{\varphi}), \\ \mathbf{h}_0 & = \mathbf{H}_0 \end{aligned} \right. \tag{5.6}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl}; \Omega)$, which is equivalent to

$$\begin{aligned}
a(\mathbf{h}_i, \boldsymbol{\varphi}) := \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right) (\mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) + (\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}) & \\
= (\mathbf{F}_i, \boldsymbol{\varphi}) + \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right) (\mathbf{h}_{i-1}, \boldsymbol{\varphi}) + \left(\frac{\delta \mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). &
\end{aligned}$$

Theorem 5.2.1. *Suppose that* $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$. *Then the variational problem* (5.6) *admits a unique solution* $\mathbf{h}_i \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *for* $i = 1,\dots,n$ *if* $\tau < \tau_0$.

Proof. The bilinear form a is elliptic for $\tau < \tau_0$:

$$\begin{aligned} a(\mathbf{h},\mathbf{h}) &\geqslant \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right)\|\mathbf{h}\|^2 + \|\nabla\times\mathbf{h}\|^2 - |(\mathcal{K}_0 \star \mathbf{h}, \nabla\times\mathbf{h})| \\ &\overset{(5.4)}{\geqslant} \left(\frac{1}{\tau} - C_\varepsilon\right)\|\mathbf{h}\|^2 + \frac{1}{\tau^2}\|\mathbf{h}\|^2 + (1-\varepsilon)\|\nabla\times\mathbf{h}\|^2 \\ &\geqslant C(\tau)\|\mathbf{h}\|^2_{\mathbf{H}_0(\mathbf{curl};\Omega)}, \end{aligned}$$

with $\varepsilon < 1$ fixed. Moreover, a is continuous in $\mathbf{H}_0(\mathbf{curl};\Omega)$. The functional $f_i(\boldsymbol{\varphi})$ is linear and bounded in $\mathbf{H}_0(\mathbf{curl};\Omega)$ if $\mathbf{h}_{i-1} \in \mathbf{L}^2(\Omega)$ and $\delta\mathbf{h}_{i-1} \in \mathbf{L}^2(\Omega)$. Therefore, if $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$, applying Lax-Milgram's lemma 2.11.1 gives the existence of a unique solution to (5.6) for any $i = 1,\dots,n$. $\square$

5.2.1 A priori estimates

First, basic stability results for $\mathbf{h}_i$ are derived. The a priori estimates in parts (i), (ii) and (iii) of the following theorem serve as uniform bounds to prove convergence.

Lemma 5.2.1 (A priori estimates)**.** *Suppose that* $\mathbf{F} : [0,T] \to \mathbf{L}^2(\Omega)$ *obeys* $\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$.

(i) Let $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$. *Then, there exists a positive constant* C *such that*

$$\max_{1\leqslant j\leqslant n}\|\mathbf{h}_j\|^2 + \sum_{i=1}^n \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\tau\nabla\times\mathbf{h}_i\|^2 \leqslant C$$

for all $\tau < \tau_0$;

(ii) If $\mathbf{H} \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$, *then*

$$\max_{1\leqslant i\leqslant n}\|\delta\mathbf{h}_i\|^2 + \max_{1\leqslant i\leqslant n}\|\nabla\times\mathbf{h}_i\|^2 + \sum_{i=1}^n\|\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n\|\nabla\times(\mathbf{h}_i - \mathbf{h}_{i-1})\|^2 \leqslant C$$

for all $\tau < \tau_0$;

(iii) If $\nabla\cdot\mathbf{F}_i = \nabla\cdot\mathbf{H}_0 = \nabla\cdot\mathbf{H}_0'$ *for* $i = 1,\dots,n$, *then* $\nabla\cdot\mathbf{h}_i = 0$ *for all* $i = 1,\dots,n$. *Moreover, we have that*

$$\tau\sum_{i=1}^n\left\|\delta^2\mathbf{h}_i\right\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C$$

for all $\tau < \tau_0$;

(iv) *If* $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$, $\mathbf{H}_0' \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n} \left\|\delta^2 \mathbf{h}_i\right\|^2 + \max_{1\leqslant i\leqslant n} \left\|\nabla \times \delta \mathbf{h}_i\right\|^2 + \sum_{i=1}^n \left\|\delta^2 \mathbf{h}_i - \delta^2 \mathbf{h}_{i-1}\right\|^2 + \sum_{i=1}^n \left\|\nabla \times (\delta \mathbf{h}_i - \delta \mathbf{h}_{i-1})\right\|^2 \leqslant C$$

for all $\tau < \tau_0$.

Proof. (i) First, we multiply (5.6) by τ and sum it up for $i = 1, \ldots, k$, $1 \leqslant k \leqslant n$. We define the sequence $\mathbf{s}_k : \Omega \to \mathbb{R}$ by

$$\mathbf{s}_k = \sum_{i=1}^k \tau \nabla \times \mathbf{h}_i, \quad k \geqslant 1; \qquad \mathbf{s}_0 = 0.$$

Using this notation, we can write for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$ that

$$(\delta \mathbf{h}_k, \boldsymbol{\varphi}) + (\mathbf{h}_k, \boldsymbol{\varphi}) + (\mathbf{s}_k, \nabla \times \boldsymbol{\varphi}) + \left(\sum_{i=1}^k \tau \mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}\right) = \left(\sum_{i=1}^k \tau \mathbf{F}_i, \boldsymbol{\varphi}\right) + (\mathbf{H}_0', \boldsymbol{\varphi}) + (\mathbf{H}_0, \boldsymbol{\varphi}).$$

Then, we put $\boldsymbol{\varphi} = \mathbf{h}_k$, multiply this by τ, sum it up for $k = 1, \ldots j$, $1 \leqslant j \leqslant n$, and obtain

$$\sum_{k=1}^j (\delta \mathbf{h}_k, \mathbf{h}_k)\,\tau + \sum_{k=1}^j \|\mathbf{h}_k\|^2 \tau + \sum_{k=1}^j (\mathbf{s}_k, \delta \mathbf{s}_k)\,\tau + \sum_{k=1}^j \left(\sum_{i=1}^k \tau \mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \mathbf{h}_k\right)\tau = \sum_{k=1}^j \left(\sum_{i=1}^k \tau \mathbf{F}_i, \mathbf{h}_k\right)\tau + \sum_{k=1}^j (\mathbf{H}_0', \mathbf{h}_k)\,\tau + \sum_{k=1}^j (\mathbf{H}_0, \mathbf{h}_k)\,\tau.$$

For the first and third terms on the LHS, we use Abel's summation rule

$$2\sum_{k=1}^j (\delta \mathbf{h}_k, \mathbf{h}_k)\,\tau = \|\mathbf{h}_j\|^2 - \|\mathbf{H}_0\|^2 + \sum_{k=1}^j \|\mathbf{h}_k - \mathbf{h}_{k-1}\|^2,$$

$$2\sum_{k=1}^j (\delta \mathbf{s}_k, \mathbf{s}_k)\,\tau = \|\mathbf{s}_j\|^2 + \sum_{k=1}^j \|\mathbf{s}_k - \mathbf{s}_{k-1}\|^2 = \left\|\sum_{k=1}^j \tau \nabla \times \mathbf{h}_k\right\|^2 + \sum_{k=1}^j \|\tau \nabla \times \mathbf{h}_k\|^2.$$

For the last term on the LHS, we apply Cauchy's and Young's inequalities

$$\begin{aligned}
&\left|\sum_{k=1}^{j}\left(\sum_{i=1}^{k}\tau\mathcal{K}_0\star\mathbf{h}_i,\tau\nabla\times\mathbf{h}_k\right)\right| \\
&\quad\leqslant\quad C_\varepsilon\sum_{k=1}^{j}\left\|\sum_{i=1}^{k}\tau\mathcal{K}_0\star\mathbf{h}_i\right\|^2+\varepsilon\sum_{k=1}^{j}\|\tau\nabla\times\mathbf{h}_k\|^2 \\
&\quad\overset{(5.3)}{\leqslant}\quad C_\varepsilon\sum_{k=1}^{j}\left(\sum_{i=1}^{k}\|\mathbf{h}_i\|^2\tau\right)\tau+\varepsilon\sum_{k=1}^{j}\|\tau\nabla\times\mathbf{h}_k\|^2 \\
&\quad\leqslant\quad C_\varepsilon\sum_{i=1}^{j}\|\mathbf{h}_i\|^2\tau+\varepsilon\sum_{k=1}^{j}\|\tau\nabla\times\mathbf{h}_k\|^2.
\end{aligned}$$

Analogically, for the first term on the RHS, we apply Cauchy's and Young's inequalities together with the assumption on the source function

$$\left|\sum_{k=1}^{j}\left(\sum_{i=1}^{k}\tau\mathbf{F}_i,\mathbf{h}_k\right)\tau\right|\leqslant\sum_{k=1}^{j}\tau\sum_{i=1}^{k}\left(\frac{\|\mathbf{F}_i\|^2+\|\mathbf{h}_k\|^2}{2}\right)\tau\ \lesssim 1+\sum_{k=1}^{j}\|\mathbf{h}_k\|^2\tau.$$

Using the assumptions on the initial conditions, we can easily deduce that

$$\left|\sum_{k=1}^{j}\left(\mathbf{H}_0',\mathbf{h}_k\right)\tau\right|\lesssim 1+\sum_{k=1}^{j}\|\mathbf{h}_k\|^2\tau\ \text{ and }\ \left|\sum_{k=1}^{j}\left(\mathbf{H}_0,\mathbf{h}_k\right)\tau\right|\lesssim 1+\sum_{k=1}^{j}\|\mathbf{h}_k\|^2\tau.$$

Eventually, we arrive at the following inequality (after changing summation indices)

$$\begin{aligned}
\|\mathbf{h}_j\|^2+\sum_{i=1}^{j}\|\mathbf{h}_i-\mathbf{h}_{i-1}\|^2+\sum_{i=1}^{j}\|\mathbf{h}_i\|^2\tau+\left\|\sum_{i=1}^{j}\tau\nabla\times\mathbf{h}_i\right\|^2+\sum_{i=1}^{j}\|\tau\nabla\times\mathbf{h}_i\|^2 \\
\leqslant C+C\|\mathbf{H}_0\|^2+\varepsilon\sum_{i=1}^{j}\|\tau\nabla\times\mathbf{h}_i\|^2+C_\varepsilon\sum_{i=1}^{j}\|\mathbf{h}_i\|^2\tau.
\end{aligned}$$

Fixing ε sufficiently small and applying Grönwall's argument, we conclude the proof.

(ii) Setting $\boldsymbol{\varphi}=\delta\mathbf{h}_i$ in (5.6), multiplying by τ and summing it up for $i=1,\dots,j$, $1\leqslant j\leqslant n$, we have that

$$\begin{aligned}
\sum_{i=1}^{j}\left(\delta^2\mathbf{h}_i,\delta\mathbf{h}_i\right)\tau+\sum_{i=1}^{j}\|\delta\mathbf{h}_i\|^2\tau+\sum_{i=1}^{j}\left(\nabla\times\mathbf{h}_i,\nabla\times\delta\mathbf{h}_i\right)\tau \\
+\sum_{i=1}^{j}\left(\mathcal{K}_0\star\mathbf{h}_i,\nabla\times\delta\mathbf{h}_i\right)\tau=\sum_{i=1}^{j}\left(\mathbf{F}_i,\delta\mathbf{h}_i\right)\tau.
\end{aligned}$$

For the first and third terms on the LHS, we use Abel's summation rule, which gives us

$$2\sum_{i=1}^{j}\left(\delta^2\mathbf{h}_i,\delta\mathbf{h}_i\right)\tau = \|\delta\mathbf{h}_j\|^2 - \|\mathbf{H}_0'\|^2 + \sum_{i=1}^{j}\|\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1}\|^2$$

and

$$2\sum_{i=1}^{j}\left(\nabla\times\mathbf{h}_i,\nabla\times\delta\mathbf{h}_i\right)\tau$$

$$= \|\nabla\times\mathbf{h}_j\|^2 - \|\nabla\times\mathbf{H}_0\|^2 + \sum_{i=1}^{j}\|\nabla\times(\mathbf{h}_i - \mathbf{h}_{i-1})\|^2.$$

Also the following partial summation formula is satisfied

$$\sum_{i=1}^{j}\left(\mathcal{K}_0\star\mathbf{h}_i,\nabla\times\delta\mathbf{h}_i\right)\tau$$

$$= \left(\mathcal{K}_0\star\mathbf{h}_j,\nabla\times\mathbf{h}_j\right) - \left(\mathcal{K}_0\star\mathbf{H}_0,\nabla\times\mathbf{H}_0\right) - \sum_{i=1}^{j}\left(\mathcal{K}_0\star\delta\mathbf{h}_i,\nabla\times\mathbf{h}_{i-1}\right)\tau.$$

Hence, using (i) and (5.4), we obtain

$$\left|\sum_{i=1}^{j}\left(\mathcal{K}_0\star\mathbf{h}_i,\nabla\times\delta\mathbf{h}_i\right)\tau\right|$$

$$\leqslant C_\varepsilon + \varepsilon\|\nabla\times\mathbf{h}_j\|^2 + C\sum_{i=1}^{j}\|\delta\mathbf{h}_i\|^2\tau + C\sum_{i=1}^{j}\|\nabla\times\mathbf{h}_i\|^2\tau.$$

The RHS can be estimated using Cauchy's and Young's inequalities as follows

$$\left|\sum_{i=1}^{j}\left(\mathbf{F}_i,\delta\mathbf{h}_i\right)\tau\right| \leqslant C\sum_{i=1}^{j}\|\mathbf{F}_i\|^2\tau + C\sum_{i=1}^{j}\|\delta\mathbf{h}_i\|^2\tau \lesssim 1 + \sum_{i=1}^{j}\|\delta\mathbf{h}_i\|^2\tau.$$

Combining the previous results gives the following inequality

$$\|\delta\mathbf{h}_j\|^2 + \sum_{i=1}^{j}\|\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^{j}\|\delta\mathbf{h}_i\|^2\tau + \|\nabla\times\mathbf{h}_j\|^2$$

$$+ \sum_{i=1}^{j}\|\nabla\times(\mathbf{h}_i - \mathbf{h}_{i-1})\|^2 \leqslant C_\varepsilon + C\|\mathbf{H}_0'\|^2 + C\|\nabla\times\mathbf{H}_0\|^2$$

$$+ \varepsilon\|\nabla\times\mathbf{h}_j\|^2 + C\sum_{i=1}^{j}\|\delta\mathbf{h}_i\|^2\tau + C\sum_{i=1}^{j}\|\nabla\times\mathbf{h}_i\|^2\tau.$$

Fixing a sufficiently small positive ε, an application of Grönwall's lemma concludes the proof.

(iii) Take the divergence of the strong formulation

$$\delta^2\mathbf{h}_i + \delta\mathbf{h}_i + \nabla\times\nabla\times\mathbf{h}_i + \nabla\times(\mathcal{K}_0\star\mathbf{h}_i) = \mathbf{F}_i,$$

Then, multiply the result by τ and sum it up for $i = 1,\ldots,j$ to arrive at

$$\nabla\cdot\delta\mathbf{h}_j + \nabla\cdot\mathbf{h}_j = 0 \qquad \text{or} \qquad (1+\tau)\nabla\cdot\mathbf{h}_j = \nabla\cdot\mathbf{h}_{j-1},$$

where $1 \leqslant j \leqslant n$. Therefore, $\nabla\cdot\mathbf{h}_i = 0$ for $i = 1,\ldots,n$ if $\nabla\cdot\mathbf{H}_0 = 0$. It holds that

$$\left(\delta^2\mathbf{h}_i,\boldsymbol{\varphi}\right) = (\mathbf{F}_i,\boldsymbol{\varphi}) - (\delta\mathbf{h}_i,\boldsymbol{\varphi}) - (\nabla\times\mathbf{h}_i,\nabla\times\boldsymbol{\varphi}) - (\mathcal{K}_0\star\mathbf{h}_i,\nabla\times\boldsymbol{\varphi}),$$

with $\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)$. The integral in the LHS has to be interpreted in the sense of duality, i.e. seeing $\delta^2\mathbf{h}_i$ as an element of $\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)$. Furthermore, we may write

$$|\,(\mathbf{F}_i,\boldsymbol{\varphi})\,| \leqslant \|\mathbf{F}_i\|\,\|\boldsymbol{\varphi}\|, \quad |\,(\nabla\times\mathbf{h}_i,\nabla\times\boldsymbol{\varphi})\,| \leqslant \|\nabla\times\mathbf{h}_i\|\,\|\nabla\times\boldsymbol{\varphi}\|$$

and

$$|\,(\delta\mathbf{h}_i,\boldsymbol{\varphi})\,| \leqslant \|\delta\mathbf{h}_i\|\,\|\boldsymbol{\varphi}\|, \quad |\,(\mathcal{K}_0\star\mathbf{h}_i,\nabla\times\boldsymbol{\varphi})\,| \overset{(5.3)}{\lesssim} \|\mathbf{h}_i\|\,\|\nabla\times\boldsymbol{\varphi}\|.$$

Thus using the dual norm in $\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)$

$$\left\|\delta^2\mathbf{h}_i\right\|_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} = \sup_{\boldsymbol{\varphi}\in\mathbf{H}_0(\mathbf{curl};\Omega)} \frac{\left(\delta^2\mathbf{h}_i,\boldsymbol{\varphi}\right)}{\|\boldsymbol{\varphi}\|_{\mathbf{H}_0(\mathbf{curl};\Omega)}},$$

(i) and (ii), we deduce that

$$\tau\sum_{i=1}^n \left\|\delta^2\mathbf{h}_i\right\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C.$$

(iv) First, we set

$$\delta^2\mathbf{h}_0 := \partial_{tt}\mathbf{H}(0) = \mathbf{F}(0) - \mathbf{H}_0' - \nabla\times\nabla\times\mathbf{H}_0 - \nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega).$$

We subtract (5.6) for $i = i-1$ from (5.6), then we set $\boldsymbol{\varphi} = \delta^2\mathbf{h}_i$ and we sum the result up for $i = 1,\ldots,j$ with $1 \leqslant j \leqslant n$ to get

$$\sum_{i=1}^j \left(\delta^3\mathbf{h}_i,\delta^2\mathbf{h}_i\right)\tau + \sum_{i=1}^j \left\|\delta^2\mathbf{h}_i\right\|^2\tau + \sum_{i=1}^j \left(\nabla\times\delta\mathbf{h}_i,\nabla\times\delta^2\mathbf{h}_i\right)\tau$$
$$+ \sum_{i=1}^j \left(\mathcal{K}_0\star\delta\mathbf{h}_i,\nabla\times\delta^2\mathbf{h}_i\right)\tau = \sum_{i=1}^j \left(\delta\mathbf{F}_i,\delta^2\mathbf{h}_i\right)\tau.$$

Moreover, we follow the same way as in (ii) considering $\delta^2\mathbf{h}_i$ instead of $\delta\mathbf{h}_i$. $\square$

5.2.2 Convergence

The existence of a weak solution is proved using Rothe's method. The following piecewise linear in time vector fields $\mathbf{H}_n$ and $\mathbf{V}_n$

$$\begin{aligned} \mathbf{H}_n(0) &= \mathbf{H}_0 \\ \mathbf{H}_n(t) &= \mathbf{h}_{i-1} + (t - t_{i-1})\delta\mathbf{h}_i \qquad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \dots, n; \end{aligned}$$

$$\begin{aligned} \mathbf{V}_n(0) &= \mathbf{H}'_0 \\ \mathbf{V}_n(t) &= \delta\mathbf{h}_{i-1} + (t - t_{i-1})\delta^2\mathbf{h}_i \qquad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \dots, n \end{aligned}$$

and the piecewise constant in time fields $\overline{\mathbf{H}}_n$ and $\overline{\mathbf{V}}_n$ are introduced

$$\overline{\mathbf{H}}_n(0) = \mathbf{H}_0, \qquad \overline{\mathbf{H}}_n(t) = \mathbf{h}_i, \quad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \dots, n;$$

$$\overline{\mathbf{V}}_n(0) = \mathbf{H}'_0, \qquad \overline{\mathbf{V}}_n(t) = \delta\mathbf{h}_i, \quad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \dots, n.$$

Similarly, the vector field $\overline{\mathbf{F}}_n$ is defined. Note that $\overline{\mathbf{V}}_n = \partial_t \mathbf{H}_n$.

The variational formulation (5.6) can be rewritten for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ and a.a. $t \in (0, T)$ as

$$(\partial_t \mathbf{V}_n(t), \boldsymbol{\varphi}) + (\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}) + \left(\nabla \times \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) \\ + \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right). \quad (5.7)$$

Now, the convergence of the sequences $\mathbf{H}_n$ and $\overline{\mathbf{H}}_n$ to the unique weak solution of (5.1) is proved as $\tau \to 0$ or $n \to \infty$.

Theorem 5.2.2 (Existence)**.** *Let* $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl}; \Omega), \mathbf{H}'_0 \in \mathbf{L}^2(\Omega)$, $\mathbf{F} : [0, T] \to \mathbf{L}^2(\Omega)$ *and* $\mathbf{F} \in \mathrm{L}^2\left((0, T), \mathbf{L}^2(\Omega)\right)$. *Assume that* $\nabla \cdot \mathbf{H}_0 = \nabla \cdot \mathbf{H}'_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *for any time* $t \in [0, T]$. *Then there exists a vector field* $\mathbf{H}$ *such that*

(i) $\overline{\mathbf{H}}_n \rightharpoonup \mathbf{H}$ *in* $\mathrm{L}^2\left((0, T), \mathbf{H}_0(\mathbf{curl}; \Omega)\right)$,
$\mathbf{H}_n \rightharpoonup \mathbf{H}$ *in* $\mathrm{L}^2\left((0, T), \mathbf{H}_0(\mathbf{curl}; \Omega)\right)$*;*

(ii) $\mathbf{H}_n(t) \rightharpoonup \mathbf{H}(t)$ *in* $\mathbf{L}^2(\Omega)$ *for any* $t \in [0, T]$*;*

(iii) $\partial_t \mathbf{H}_n = \overline{\mathbf{V}}_n \rightharpoonup \partial_t \mathbf{H}$ *in* $\mathrm{L}^2\left((0, T), \mathbf{L}^2(\Omega)\right)$,
$\mathbf{V}_n \to \partial_t \mathbf{H}$ *in* $\mathrm{L}^2\left((0, T), \mathbf{L}^2(\Omega)\right)$, $\mathbf{V}_n \rightharpoonup \partial_t \mathbf{H}$ *in* $\mathrm{L}^2\left((0, T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right)$
and
$\partial_t \mathbf{V}_n \rightharpoonup \partial_{tt} \mathbf{H}$ *in* $\mathrm{L}^2\left((0, T), \mathbf{H}_0^{-1}(\mathbf{curl}; \Omega)\right)$*;*

(iv) $\mathbf{H}$ *is a weak solution of* (5.2)*;*

(v) $\mathbf{H}_n \to \mathbf{H}$ *in* $\mathrm{C}\left([0, T], \mathbf{L}^2(\Omega)\right)$,
$\mathbf{H} \in \mathrm{C}\left([0, T], \mathbf{H}_0(\mathbf{curl}; \Omega)\right) \cap \mathrm{L}^\infty\left((0, T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right)$ *and*
$\partial_t \mathbf{H} \in \mathrm{L}^2\left((0, T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right) \cap \mathrm{C}\left([0, T], \mathbf{L}^2(\Omega)\right)$.

Proof. (i) Thanks to Lemma 5.2.1(i) and (ii), the sequences $\{\overline{\mathbf{H}}_n\}$ and $\{\mathbf{H}_n\}$ are bounded in $\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right)$. Therefore, due to the reflexivity of this space, the sequence $\{\overline{\mathbf{H}}_n\}$ contains a weakly convergence subsequence (denoted by the same symbol again) such that $\overline{\mathbf{H}}_n \rightharpoonup \mathbf{H}$ in $\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right)$. The sequences $\{\overline{\mathbf{H}}_n\}$ and $\{\mathbf{H}_n\}$ have the same limit in the space $\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right)$. Employing Lemma 5.2.1(i) and (ii) gives

$$\begin{aligned}\lim_{n\to\infty}\|\mathbf{H}_n-\overline{\mathbf{H}}_n\|^2_{\mathrm{L}^2((0,T),\mathbf{H}_0(\mathbf{curl};\Omega))} &\leqslant \lim_{n\to\infty}\tau\sum_{i=1}^n\|\mathbf{h}_i-\mathbf{h}_{i-1}\|^2_{\mathbf{H}_0(\mathbf{curl};\Omega)}\\ &\leqslant \lim_{n\to\infty}\frac{C}{n}=0.\end{aligned}$$

Hence, $\mathbf{H}_n \rightharpoonup \mathbf{H}$ in $\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right)$.

(ii) The sequence $\{\mathbf{H}_n\}$ with $\mathbf{H}_n:[0,T]\to\mathbf{L}^2(\Omega), n\in\mathbb{N}$, is equibounded and uniform equicontinuous. For every $n\in\mathbb{N}$ and $\forall t,t_1,t_2\in[0,T]$, we have, using Lemma 5.2.1(i) and (ii), that

$$\|\mathbf{H}_n(t)\|\leqslant C$$

and

$$\|\mathbf{H}_n(t_2)-\mathbf{H}_n(t_1)\|\leqslant\sqrt{|t_2-t_1|}\sqrt{\int_{t_1}^{t_2}\|\partial_t\mathbf{H}_n(t)\|^2\,\mathrm{d}t}\lesssim\sqrt{|t_2-t_1|}.$$

An application of Lemma 2.12.2 gives that $\mathbf{H}_n(t)\rightharpoonup\mathbf{H}(t)$ in $\mathbf{L}^2(\Omega)$ for any $t\in[0,T]$.

(iii) The sequence $\partial_t\mathbf{H}_n$ is bounded in the reflexive space $\mathrm{L}^2\big((0,T),\mathbf{L}^2(\Omega)\big)$ by Lemma 5.2.1(ii). Hence, $\partial_t\mathbf{H}_n=\overline{\mathbf{V}}_n\rightharpoonup\mathbf{z}$ in $\mathrm{L}^2\big((0,T),\mathbf{L}^2(\Omega)\big)$. Thanks to (ii), we get

$$\begin{array}{ccc}(\mathbf{H}_n(t)-\mathbf{H}_0,\boldsymbol{\varphi}) & = & \displaystyle\int_0^t(\partial_t\mathbf{H}_n,\boldsymbol{\varphi})\\ \downarrow & & \downarrow\\ (\mathbf{H}(t)-\mathbf{H}_0,\boldsymbol{\varphi}) & = & \displaystyle\int_0^t(\mathbf{z},\boldsymbol{\varphi}),\end{array}$$

which is valid for any $t\in[0,T]$. Thus $\mathbf{z}=\partial_t\mathbf{H}$ in $\mathrm{L}^2\big((0,T),\mathbf{L}^2(\Omega)\big)$. Lemma 5.2.1(ii) implies

$$\lim_{n\to\infty}\|\mathbf{V}_n-\overline{\mathbf{V}}_n\|^2_{\mathrm{L}^2((0,T),\mathbf{L}^2(\Omega))}\leqslant\lim_{n\to\infty}\tau\sum_{i=1}^n\|\delta\mathbf{h}_i-\delta\mathbf{h}_{i-1}\|^2\leqslant\lim_{n\to\infty}\frac{C}{n}=0.$$

Thus $\mathbf{V}_n\rightharpoonup\partial_t\mathbf{H}$ in $\mathrm{L}^2\big((0,T),\mathbf{L}^2(\Omega)\big)$. Lemma 5.2.1(i), (ii) and (iii) give

$$\mathbf{V}_n\in\mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right),\quad\nabla\cdot\mathbf{V}_n(t)=0\quad\forall t\in[0,T].$$

Consequently, reviewing Theorem 2.9.36, we see that $\mathbf{V}_n \in \mathrm{L}^2\left((0,T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right)$. Using Theorem 2.9.37, we obtain that

$$\mathbf{H}^{\frac{1}{2}}(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega) \cong \mathbf{L}^2(\Omega)^* \hookrightarrow \mathbf{H}_0^{-1}(\mathbf{curl};\Omega).$$

Taking into account the fact that $\partial_t \mathbf{V}_n \in \mathrm{L}^2\left((0,T), \mathbf{H}_0^{-1}(\mathbf{curl};\Omega)\right)$, see Lemma 5.2.1(iii), and using the generalized Aubin-Lions lemma 2.12.4, we get that $\mathbf{V}_n \to \partial_t \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, $\mathbf{V}_n \rightharpoonup \partial_t \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{H}^{\frac{1}{2}}(\Omega)\right)$ and $\partial_t \mathbf{V}_n \rightharpoonup \partial_{tt}\mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{H}_0^{-1}(\mathbf{curl};\Omega)\right)$. Therefore, there exists a subsequence of $\mathbf{V}_n$ (denoted by the same symbol again) for which we have

$$\mathbf{V}_n(\mathbf{x},t) \to \partial_t \mathbf{H}(\mathbf{x},t) \qquad \text{a.e. in } \Omega \times (0,T).$$

(iv) Let us integrate (5.7) in time to get (for any $\xi \in (0,T)$)

$$\int_0^\xi (\partial_t \mathbf{V}_n, \boldsymbol{\varphi}) + \int_0^\xi (\partial_t \mathbf{H}_n, \boldsymbol{\varphi}) + \int_0^\xi \left(\nabla \times \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right) + \int_0^\xi \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right) = \int_0^\xi \left(\overline{\mathbf{F}}_n, \boldsymbol{\varphi}\right). \tag{5.8}$$

Clearly $\overline{\mathbf{F}}_n \rightharpoonup \mathbf{F}$ in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. Both terms $\int_0^\xi \left(\nabla \times \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right)$ and $\int_0^\xi \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right)$ are linear bounded functionals in the space $\mathrm{L}^2\left((0,T), \mathbf{H}_0(\mathbf{curl};\Omega)\right)$. Now, we can pass to the limit for $n \to \infty$ in (5.8). Thanks to (i) and (iii), we get

$$\int_0^\xi (\partial_{tt} \mathbf{H}, \boldsymbol{\varphi}) + \int_0^\xi (\partial_t \mathbf{H}, \boldsymbol{\varphi}) + \int_0^\xi (\nabla \times \mathbf{H}, \nabla \times \boldsymbol{\varphi}) + \int_0^\xi (\mathcal{K}_0 \star \mathbf{H}, \nabla \times \boldsymbol{\varphi}) = \int_0^\xi (\mathbf{F}, \boldsymbol{\varphi}).$$

Differentiating this equality with respect to the time variable gives the existence of a weak solution to (5.2).

(v) We recall that $\mathbf{H}^{\frac{1}{2}}(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega)$. The sequence $\{\mathbf{H}_n\}$ with $\mathbf{H}_n : [0,T] \to \mathbf{H}^{\frac{1}{2}}(\Omega)$, $n \in \mathbb{N}$, is equibounded. Using Theorem 2.9.36, Theorem 5.2.1 and Lemma 5.2.1(i-iii) gives for all $t \in (t_{i-1}, t_i]$ that

$$\begin{aligned}
\|\mathbf{H}_n(t)\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} &\leqslant \|\mathbf{h}_{i-1}\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} + \|\mathbf{h}_i - \mathbf{h}_{i-1}\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} \\
&\lesssim \|\mathbf{h}_{i-1}\| + \|\nabla \times \mathbf{h}_{i-1}\| + \|\mathbf{h}_i - \mathbf{h}_{i-1}\| + \|\nabla \times (\mathbf{h}_i - \mathbf{h}_{i-1})\| \\
&\leqslant C.
\end{aligned}$$

Thus

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t)\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} \leqslant C.$$

In part (ii) of the proof, we have shown that the sequence $\{\mathbf{H}_n\}$ with $\mathbf{H}_n : [0,T] \to \mathbf{L}^2(\Omega), n \in \mathbb{N}$, is uniform equicontinuous. Now, Lemma 2.12.2(ii), implies that $\mathbf{H}_n \to \mathbf{H}$ in $\mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$ and $\mathbf{H} \in \mathrm{L}^\infty\left((0,T),\mathbf{H}^{\frac{1}{2}}(\Omega)\right)$. Consider the following evolution triple (or sometimes called Gelfand's triple) of spaces

$$\mathbf{H}_0(\mathbf{curl};\Omega) \hookrightarrow \mathbf{L}^2(\Omega) \cong \mathbf{L}^2(\Omega)^* \hookrightarrow \mathbf{H}_0^{-1}(\mathbf{curl};\Omega).$$

We know that

$$\mathbf{H}, \partial_t\mathbf{H} \in \mathrm{L}^2\left((0,T),\mathbf{H}_0(\mathbf{curl};\Omega)\right) \text{ and } \partial_{tt}\mathbf{H} \in \mathrm{L}^2\left((0,T),\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)\right).$$

Applying Lemma 2.9.5(i) and (iii) gives that $\mathbf{H} \in \mathrm{C}\left([0,T],\mathbf{H}_0(\mathbf{curl};\Omega)\right)$ and $\partial_t\mathbf{H} \in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$, which concludes the proof. □

5.2.3 Error estimates

The following theorem addresses the error estimates for the time discretization.

Theorem 5.2.3 (Error)**.** *Suppose that* $\mathbf{F} \in \mathrm{Lip}([0,T],\mathbf{L}^2(\Omega))$.

(i) If $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t)-\mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\|\nabla\times\int_0^t [\mathbf{H}_n-\mathbf{H}]\right\|^2 \leqslant C\tau.$$

(ii) If $\nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$, $\mathbf{H}_0' \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t)-\mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\|\nabla\times\int_0^t [\mathbf{H}_n-\mathbf{H}]\right\|^2 \leqslant C\tau^2.$$

Please note that the positive constant C *in these estimates is of the form* Ce^{CT}.

Proof. We subtract (5.2) from (5.7) and integrate the result in time over $t \in (0,\xi) \subset (0,T)$ to get

$$\begin{aligned}&(\mathbf{V}_n(\xi)-\partial_t\mathbf{H}(\xi),\boldsymbol{\varphi}) + (\mathbf{H}_n(\xi)-\mathbf{H}(\xi),\boldsymbol{\varphi})\\ &+\left(\nabla\times\int_0^\xi [\overline{\mathbf{H}}_n(t)-\mathbf{H}(t)],\nabla\times\boldsymbol{\varphi}\right) + \left(\mathcal{K}_0\star\int_0^\xi [\overline{\mathbf{H}}_n(t)-\mathbf{H}(t)],\nabla\times\boldsymbol{\varphi}\right)\\ &\qquad= \left(\int_0^\xi [\overline{\mathbf{F}}_n(t)-\mathbf{F}(t)],\boldsymbol{\varphi}\right).\end{aligned}$$

Now, putting $\boldsymbol{\varphi} = \mathbf{H}_n(\xi) - \mathbf{H}(\xi)$, using $\overline{\mathbf{V}}_n = \partial_t \mathbf{H}_n$ and integrating in time over the variable $\xi \in (0, \eta) \subset (0, T)$, we arrive at

$$\begin{aligned}
\frac{1}{2} \left\| \mathbf{H}_n(\eta) - \mathbf{H}(\eta) \right\|^2 + \int_0^\eta \left\| \mathbf{H}_n - \mathbf{H} \right\|^2 + \frac{1}{2} \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 \\
+ \int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n(t) - \mathbf{H}(t)], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right) \\
= \int_0^\eta \left(\int_0^\xi [\overline{\mathbf{F}}_n(t) - \mathbf{F}(t)], \mathbf{H}_n(\xi) - \mathbf{H}(\xi) \right) + \int_0^\eta \left(\overline{\mathbf{V}}_n - \mathbf{V}_n, \mathbf{H}_n - \mathbf{H} \right) \\
+ \int_0^\eta \left(\nabla \times \int_0^\xi [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right) \\
+ \int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right). \quad (5.9)
\end{aligned}$$

Due to the Lipschitz continuity of $\mathbf{F}$, we may write that

$$\left| \int_0^\eta \left(\int_0^\xi [\overline{\mathbf{F}}_n(t) - \mathbf{F}(t)], \mathbf{H}_n(\xi) - \mathbf{H}(\xi) \right) \right| \lesssim \tau^2 + \int_0^\eta \left\| \mathbf{H}_n - \mathbf{H} \right\|^2.$$

The integration by parts formula gives the following estimate

$$\begin{aligned}
& \int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n(t) - \mathbf{H}(t)], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right) \\
= & \left(\mathcal{K}_0 \star \int_0^\eta [\mathbf{H}_n - \mathbf{H}], \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right) \\
& - \int_0^\eta \left(\mathcal{K}_0 \star [\mathbf{H}_n - \mathbf{H}], \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right) \\
\overset{(5.4)}{\leqslant} & \varepsilon \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 + C_\varepsilon \int_0^\eta \left\| \mathbf{H}_n - \mathbf{H} \right\|^2 \\
& + C \int_0^\eta \left\| \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right\|^2.
\end{aligned}$$

It holds that

$$\left\| \mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t) \right\| \lesssim \tau \left\| \partial_t \mathbf{H}_n(t) \right\| \qquad \text{for } t \in [0, T]$$

and

$$\left\| \mathbf{V}_n(t) - \overline{\mathbf{V}}_n(t) \right\| \lesssim \tau \left\| \partial_t \mathbf{V}_n(t) \right\| \qquad \text{for } t \in [0, T].$$

Analogue as in the previous estimate, using Lemma 5.2.1(ii), we get that

$$
\begin{aligned}
&\int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right) \\
&\quad \leqslant \ \varepsilon \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 \\
&\qquad + C_\varepsilon \int_0^\eta \left\| \mathbf{H}_n - \overline{\mathbf{H}}_n \right\|^2 + C \int_0^\eta \left\| \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right\|^2 \\
&\quad \leqslant \ \varepsilon \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 + C_\varepsilon \tau^2 + C \int_0^\eta \left\| \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right\|^2 .
\end{aligned}
$$

It remains to estimate the second and third terms on the RHS in (5.9). We have to distinguish between two cases depending on the assumptions on the initial conditions. If $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl}; \Omega)$ and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$, we get

$$
\begin{aligned}
\left| \int_0^\eta \left(\overline{\mathbf{V}}_n - \mathbf{V}_n, \mathbf{H}_n - \mathbf{H} \right) \right| &\leqslant \varepsilon \int_0^\eta \| \mathbf{H}_n - \mathbf{H} \|^2 + C_\varepsilon \int_0^\eta \left\| \overline{\mathbf{V}}_n - \mathbf{V}_n \right\|^2 \\
&\leqslant \varepsilon \int_0^\eta \| \mathbf{H}_n - \mathbf{H} \|^2 + C_\varepsilon \tau
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_0^\eta \left(\nabla \times \int_0^\xi [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right) \\
&\quad \leqslant \ \varepsilon \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 \\
&\qquad + C_\varepsilon \int_0^\eta \left\| \nabla \times [\mathbf{H}_n - \overline{\mathbf{H}}_n] \right\|^2 + C \int_0^\eta \left\| \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right\|^2 \\
&\quad \leqslant \ \varepsilon \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 + C_\varepsilon \tau + C \int_0^\eta \left\| \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right\|^2 .
\end{aligned}
$$

If $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega), \ \mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl}; \Omega), \ \mathbf{H}_0' \in \mathbf{H}(\mathbf{curl}; \Omega)$ and $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ then

$$
\left| \int_0^\eta \left(\overline{\mathbf{V}}_n - \mathbf{V}_n, \mathbf{H}_n - \mathbf{H} \right) \right| \leqslant \varepsilon \int_0^\eta \| \mathbf{H}_n - \mathbf{H} \|^2 + C_\varepsilon \tau^2
$$

and

$$
\begin{aligned}
&\int_0^\eta \left(\nabla \times \int_0^\xi [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)] \right) \\
&\quad \leqslant \ \varepsilon \left\| \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}] \right\|^2 + C_\varepsilon \tau^2 + C \int_0^\eta \left\| \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}] \right\|^2 .
\end{aligned}
$$

Combining the previous results, choosing a sufficiently small positive ε and applying Grönwall's argument, we conclude the proof. $\square$

From this estimate, the uniqueness of the solution can also be proved. If $\mathbf{H}_1$ and $\mathbf{H}_2$ satisfy (5.2), then (if $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$ and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$)

$$\max_{\eta\in[0,T]} \|\mathbf{H}_1(\eta) - \mathbf{H}_2(\eta)\| \\ \leqslant \max_{\eta\in[0,T]} \|\mathbf{H}_n(\eta) - \mathbf{H}_1(\eta)\| + \max_{\eta\in[0,T]} \|\mathbf{H}_n(\eta) - \mathbf{H}_2(\eta)\| \lesssim \sqrt{\tau},$$

which is arbitrarily small.

5.3 Modified scheme

In this section, the following semi-implicit time-discrete scheme is considered, which represents a slight modification of (5.6)

$$\left\{\begin{array}{rl} \left(\delta^2\mathbf{h}_i, \boldsymbol{\varphi}\right) + (\delta\mathbf{h}_i, \boldsymbol{\varphi}) & \\ + (\nabla\times\mathbf{h}_i, \nabla\times\boldsymbol{\varphi}) & = (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla\times\boldsymbol{\varphi}), \\ \mathbf{h}_0 & = \mathbf{H}_0, \end{array}\right. \tag{5.10}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$, which is equivalent to

$$\begin{aligned} a(\mathbf{h}_i, \boldsymbol{\varphi}) &:= \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right)(\mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla\times\mathbf{h}_i, \nabla\times\boldsymbol{\varphi}) \\ &= (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla\times\boldsymbol{\varphi}) \\ &\quad + \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right)(\mathbf{h}_{i-1}, \boldsymbol{\varphi}) + \left(\frac{\delta\mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

The convolution term in this scheme is taken explicitly (from the last time step), while in the scheme (5.6) an implicit form (from the actual time step) is considered. The easier implementation is the advantage of this scheme in comparison to (5.6).

An application of the Lax-Milgram lemma 2.11.1 yields the existence of a unique solution to (5.10) in $\mathbf{H}_0(\mathbf{curl};\Omega)$ for any $i = 1,\ldots,n$ and any $\tau > 0$. Indeed, the bilinear form a is elliptic and continuous in $\mathbf{H}_0(\mathbf{curl};\Omega)$. Moreover, according to (5.3), the functional f_i ($i = 1,\ldots,n$) is linear and bounded in $\mathbf{H}_0(\mathbf{curl};\Omega)$ if $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$.

Handling this scheme is very similar to the way used for (5.6). To be short, only the differences between both algorithms are pointed out.

Lemma 5.3.1 (A priori estimates)**.** *Suppose that* $\mathbf{F} : [0,T] \to \mathbf{L}^2(\Omega)$ *obeys* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

(i) Let $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$*. Then, there exists a positive constant* C *such that*

$$\max_{1\leqslant j\leqslant n} \|\mathbf{h}_j\|^2 + \sum_{i=1}^n \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\tau\nabla\times\mathbf{h}_i\|^2 \leqslant C$$

for all $\tau < \tau_0$*;*

(ii) If $\mathbf{H}_0 \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$*, then*

$$\max_{1\leqslant i\leqslant n} \|\delta\mathbf{h}_i\|^2 + \max_{1\leqslant i\leqslant n} \|\nabla\times\mathbf{h}_i\|^2 + \sum_{i=1}^n \|\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\nabla\times(\mathbf{h}_i - \mathbf{h}_{i-1})\|^2 \leqslant C$$

for all $\tau < \tau_0$*;*

(iii) If $\nabla\cdot\mathbf{F}_i = \nabla\cdot\mathbf{H}_0 = \nabla\cdot\mathbf{H}_0'$ *for* $i = 1,\ldots,n$*, then* $\nabla\cdot\mathbf{h}_i = 0$ *for all* $i = 1,\ldots,n$*. Moreover, we have that*

$$\tau\sum_{i=1}^n \left\|\delta^2\mathbf{h}_i\right\|^2_{\mathbf{H}_0^{-1}(\mathbf{curl};\Omega)} \leqslant C$$

for all $\tau < \tau_0$*;*

(iv) If $\partial_t\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$*,* $\nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega)$*,* $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$*,* $\mathbf{H}_0' \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n} \left\|\delta^2\mathbf{h}_i\right\|^2 + \max_{1\leqslant i\leqslant n} \|\nabla\times\delta\mathbf{h}_i\|^2 + \sum_{i=1}^n \left\|\delta^2\mathbf{h}_i - \delta^2\mathbf{h}_{i-1}\right\|^2 + \sum_{i=1}^n \|\nabla\times(\delta\mathbf{h}_i - \delta\mathbf{h}_{i-1})\|^2 \leqslant C$$

for all $\tau < \tau_0$.

Proof. (i) We follow Lemma 5.2.1(i). Using (5.4), we have that

$$\left|\sum_{k=1}^j \left(\sum_{i=1}^k \tau\mathcal{K}_0\star\mathbf{h}_{i-1}, \tau\nabla\times\mathbf{h}_k\right)\right| \leqslant C_\varepsilon \sum_{i=0}^j \|\mathbf{h}_i\|^2\tau + \varepsilon\sum_{k=1}^j \|\tau\nabla\times\mathbf{h}_k\|^2.$$

After fixing a sufficiently small positive ε, an application of Grönwall's lemma completes the proof.

(ii) Note that

$$\sum_{i=1}^j (\mathcal{K}_0\star\mathbf{h}_{i-1}, \nabla\times\delta\mathbf{h}_i)\,\tau = (\mathcal{K}_0\star\mathbf{h}_j, \nabla\times\mathbf{h}_j) - (\mathcal{K}_0\star\mathbf{H}_0, \nabla\times\mathbf{H}_0) - \sum_{i=1}^j (\mathcal{K}_0\star\delta\mathbf{h}_i, \nabla\times\mathbf{h}_i)\,\tau.$$

The rest of the proof runs as before.

(iii) The proof is the same as in Lemma 5.2.1(iii) replacing $(\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})$ by $(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi})$.

(iv) We set

$$\delta^2\mathbf{h}_0 := \mathbf{F}(0) - \mathbf{H}_0' - \nabla \times \nabla \times \mathbf{H}_0 - \nabla \times (\mathcal{K}_0 \star \mathbf{H}_0), \quad \mathbf{h}_{-1} := \mathbf{h}_0 - \delta\mathbf{h}_0\tau.$$

Note that $\delta\mathbf{h}_0, \mathbf{h}_{-1} \in \mathbf{L}^2(\Omega)$. The proof follows very closely the proof of Lemma 5.2.1(iv), except for the appearance of the term $(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi})$ instead of $(\mathcal{K}_0 \star \mathbf{h}_i, \nabla \times \boldsymbol{\varphi})$. □

The variational formulation (5.10) can be rewritten for all $\boldsymbol{\varphi} \in \mathbf{H}_0(\mathbf{curl};\Omega)$ and a.a. $t \in (0,T)$ as

$$\begin{aligned}(\partial_t \mathbf{V}_n(t), \boldsymbol{\varphi}) + (\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}) + \left(\nabla \times \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right)& \\ = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right) - \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n(t-\tau), \nabla \times \boldsymbol{\varphi}\right).&\end{aligned}$$

Next theorem derives the error estimates for the scheme (5.10). The same convergence rate is obtained as in the error estimates in Theorem 5.2.3.

Theorem 5.3.1 (Error)**.** *Suppose that* $\mathbf{F} \in \mathrm{Lip}([0,T], \mathbf{L}^2(\Omega))$.

(i) If $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\|\nabla \times \int_0^t [\mathbf{H}_n - \mathbf{H}]\right\|^2 \leqslant C\tau.$$

(ii) If $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{H}_0(\mathbf{curl};\Omega)$, $\mathbf{H}_0' \in \mathbf{H}(\mathbf{curl};\Omega)$ *and* $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *then*

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\|\nabla \times \int_0^t [\mathbf{H}_n - \mathbf{H}]\right\|^2 \leqslant C\tau^2.$$

Please note that the positive constant C *in these estimates is of the form* Ce^{CT}.

Proof. The proof follows the same lines as Theorem 5.2.3. The term $\displaystyle\int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n - \overline{\mathbf{H}}_n], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)]\right)$ in (5.9) is now replaced by $\displaystyle\int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t-\tau)], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)]\right)$. This can be handled using integration by parts, (5.4) and Lemma 5.3.1(ii) as follows

$$
\begin{aligned}
&\left|\int_0^\eta \left(\mathcal{K}_0 \star \int_0^\xi [\mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t-\tau)], \nabla \times [\mathbf{H}_n(\xi) - \mathbf{H}(\xi)]\right) \mathrm{d}\xi\right| \\
&= \left|\left(\mathcal{K}_0 \star \int_0^\eta [\mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t-\tau)], \nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}]\right)\right. \\
&\qquad \left. - \int_0^\eta \left(\mathcal{K}_0 \star [\mathbf{H}_n(\xi) - \overline{\mathbf{H}}_n(\xi-\tau)], \nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}]\right) \mathrm{d}\xi\right| \\
&\leqslant \varepsilon \left\|\nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}]\right\|^2 \\
&\qquad + C_\varepsilon \int_0^\eta \left\|\mathbf{H}_n(t) - \overline{\mathbf{H}}_n(t-\tau)\right\|^2 + C \int_0^\eta \left\|\nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}]\right\|^2 \\
&\leqslant \varepsilon \left\|\nabla \times \int_0^\eta [\mathbf{H}_n - \mathbf{H}]\right\|^2 + C_\varepsilon \tau^2 + C \int_0^\eta \left\|\nabla \times \int_0^\xi [\mathbf{H}_n - \mathbf{H}]\right\|^2 .
\end{aligned}
$$

The rest of the proof is the same as the proof of Theorem 5.2.3. □

5.4 Higher regularity

The solution of problem (5.1) is divergence free for any $t \in [0,T]$ if $\nabla \cdot \mathbf{H}_0 = \nabla \cdot \mathbf{H}_0' = 0 = \nabla \cdot \mathbf{F}(t)$ for any time $t \in [0,T]$, see Theorem 5.1.1. From now on, it is assumed that $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ on $\partial\Omega$ for any $t \in [0,T]$. Then, thanks to Lemma 4.5.1, it holds that

$$
\nabla \times (\mathcal{K}_0 \star \mathbf{H})(\mathbf{x}) = \int_\Omega \mathcal{K}(\mathbf{x}, \mathbf{x}')\mathbf{H}(\mathbf{x}')\ \mathrm{d}\mathbf{x}' =: (\mathcal{K} \star \mathbf{H})(\mathbf{x}), \qquad \mathbf{x} \in \Omega,
$$

where the kernel $\mathcal{K}$ is defined by

$$
\mathcal{K} : \Omega \times \Omega \to \mathbb{R} : (\mathbf{x}, \mathbf{x}') \mapsto \kappa(|\mathbf{x} - \mathbf{x}'|),
$$

with

$$
\kappa : (0,\infty) \to \mathbb{R} : s \mapsto \begin{cases} \frac{\widetilde{C}}{2s^2}\left(1 - \frac{s}{r_0}\right)\exp\left(-\frac{s}{r_0}\right) & s < r_0, \\ 0 & s \geqslant r_0. \end{cases}
$$

Therefore, the solution of problem (5.1) satisfies

$$
\begin{cases}
\partial_{tt}\mathbf{H} + \partial_t\mathbf{H} - \Delta\mathbf{H} + \mathcal{K} \star \mathbf{H} = \mathbf{F} & \text{in } Q_T, \\
\mathbf{H} = \mathbf{0} & \text{on } \partial\Omega \times (0,T), \\
\mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0 & \text{in } \Omega, \\
\partial_t\mathbf{H}(\mathbf{x}, 0) = \mathbf{H}_0' & \text{in } \Omega, \\
\nabla \cdot \mathbf{H}_0 = \nabla \cdot \mathbf{H}_0' = \mathbf{0} & \text{in } \Omega.
\end{cases} \tag{5.11}
$$

In the hyperbolic problem (5.1), the use of Grönwall's lemma implies that the constants in the error estimates of Theorem 5.2.3 and Theorem 5.3.1 depend exponentially on the final time. For the parabolic problem (4.1), the positive definiteness of $\mathcal{K}$, which was proven in Theorem 4.5.3, was employed to establish better error estimates for the implicit scheme.

However, for the hyperbolic problem 5.11, the use of Grönwall's lemma cannot be avoided despite the positive definiteness of $\mathcal{K}$. Therefore, the analysis follows closely the lines of the analysis of problem (5.1). The same results are obtained as in the previous section, where the curl-spaces are replaced by analogous $\mathbf{H}^s(\Omega)$-spaces.

Under the additional assumption that $\mathbf{H}(t)\cdot\boldsymbol{\nu} = 0$ on $\partial\Omega$ for any $t \in [0,T]$, the solution to problem (5.1) obeys for all $\boldsymbol{\varphi} \in \mathbf{H}^1_0(\Omega)$ and a.a. $t \in (0,T)$ that

$$\begin{aligned}(\partial_{tt}\mathbf{H}(t),\boldsymbol{\varphi}) &+ (\partial_t\mathbf{H}(t),\boldsymbol{\varphi}) \\ &+ (\nabla\mathbf{H}(t),\nabla\boldsymbol{\varphi}) + ((\mathcal{K}\star\mathbf{H})(t),\boldsymbol{\varphi}) = (\mathbf{F}(t),\boldsymbol{\varphi})\,. \end{aligned} \tag{5.12}$$

As before, the following linear recurrent scheme (convolution implicitly) is proposed for $\boldsymbol{\varphi} \in \mathbf{H}^1_0(\Omega)$:

$$\left\{ \begin{array}{rl} \left(\delta^2\mathbf{h}_i,\boldsymbol{\varphi}\right) + (\delta\mathbf{h}_i,\boldsymbol{\varphi}) + (\nabla\mathbf{h}_i,\nabla\boldsymbol{\varphi}) + (\mathcal{K}\star\mathbf{h}_i,\boldsymbol{\varphi}) &= (\mathbf{F}_i,\boldsymbol{\varphi})\,,\\ \mathbf{h}_0 &= \mathbf{H}_0, \end{array}\right. \tag{5.13}$$

which is equivalent to

$$\begin{aligned} a(\mathbf{h}_i,\boldsymbol{\varphi}) &:= \left(\frac{1}{\tau^2}+\frac{1}{\tau}\right)(\mathbf{h}_i,\boldsymbol{\varphi}) + (\nabla\mathbf{h}_i,\nabla\boldsymbol{\varphi}) + (\mathcal{K}\star\mathbf{h}_i,\boldsymbol{\varphi}) \\ &= (\mathbf{F}_i,\boldsymbol{\varphi}) + \left(\frac{1}{\tau^2}+\frac{1}{\tau}\right)(\mathbf{h}_{i-1},\boldsymbol{\varphi}) + \left(\frac{\delta\mathbf{h}_{i-1}}{\tau},\boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

The bilinear form $a(\mathbf{h},\boldsymbol{\varphi})$ is elliptic and continuous in $\mathbf{H}^1_0(\Omega)$ due to the positive definiteness of $\mathcal{K}$, inequality (4.20) and $\mathbf{H}^1_0(\Omega) \hookrightarrow \mathbf{L}^6(\Omega)$. If $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$, then the functional f_i is linear and bounded in $\mathbf{H}^1_0(\Omega)$, $i = 1,\dots,n$. An application of the Lax-Milgram lemma 2.11.1 gives the well-posedness of (5.13) for any $i = 1,\dots,n$ and any $\tau > 0$. The following lemma is analogous to Lemma 5.2.1.

Lemma 5.4.1 (Enhanced a priori estimates)**.** *Assume that* $\nabla\cdot\mathbf{F}_i = 0 = \nabla\cdot\mathbf{H}_0 = \nabla\cdot\mathbf{H}_0'$ *for* $i = 1,\dots,n$ *and* $\mathbf{H}(t)\cdot\boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any* $t \in [0,T]$. *Suppose that* $\mathbf{F} : [0,T] \to \mathbf{L}^2(\Omega)$ *obeys* $\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$.

(i) Let $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$. *Then, there exists a positive constant* C *such that*

$$\max_{1\leqslant j\leqslant n}\|\mathbf{h}_j\|^2 + \sum_{i=1}^n\|\mathbf{h}_i-\mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n\|\tau\nabla\mathbf{h}_i\|^2 \leqslant C$$

for all $\tau < \tau_0$*;*

(ii) If $\mathbf{H}_0 \in \mathbf{H}^1(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$*, then*

$$\max_{1\leqslant i\leqslant n} \|\delta \mathbf{h}_i\|^2 + \max_{1\leqslant i\leqslant n} \|\nabla \mathbf{h}_i\|^2 + \sum_{i=1}^n \|\delta \mathbf{h}_i - \delta \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^n \|\nabla \mathbf{h}_i - \nabla \mathbf{h}_{i-1}\|^2 \leqslant C$$

for all $\tau < \tau_0$*;*

(iii) Moreover, we have that

$$\tau \sum_{i=1}^n \left\|\delta^2 \mathbf{h}_i\right\|_{\mathbf{H}^{-1}(\Omega)}^2 \leqslant C;$$

(iv) If $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$*,* $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ *and* $\mathbf{H}_0' \in \mathbf{H}^1(\Omega)$ *then*

$$\max_{1\leqslant i\leqslant n} \left\|\delta^2 \mathbf{h}_i\right\|^2 + \max_{1\leqslant i\leqslant n} \|\nabla \delta \mathbf{h}_i\|^2 + \sum_{i=1}^n \left\|\delta^2 \mathbf{h}_i - \delta^2 \mathbf{h}_{i-1}\right\|^2 + \sum_{i=1}^n \|\nabla \delta \mathbf{h}_i - \nabla \delta \mathbf{h}_{i-1}\|^2 \leqslant C$$

for all $\tau < \tau_0$*.*

The variational formulation (5.13) can be rewritten as

$$(\partial_t \mathbf{V}_n(t), \boldsymbol{\varphi}) + (\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}) + \left(\nabla \overline{\mathbf{H}}_n(t), \nabla \boldsymbol{\varphi}\right) + \left(\mathcal{K} \star \overline{\mathbf{H}}_n(t), \boldsymbol{\varphi}\right) = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right), \ \boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega). \quad (5.14)$$

The main point of the existence theorem is the embedding

$$\mathbf{H}_0^1(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega) \cong \mathbf{L}^2(\Omega)^* \hookrightarrow \mathbf{H}^{-1}(\Omega).$$

Theorem 5.4.1 (Enhanced existence)**.** *Let* $\mathbf{H}_0 \in \mathbf{H}^1(\Omega), \mathbf{H}_0' \in \mathbf{L}^2(\Omega)$*,* $\mathbf{F} : [0,T] \to \mathbf{L}^2(\Omega)$ *and* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$*. Assume that* $\nabla \cdot \mathbf{H}_0 = \nabla \cdot \mathbf{H}_0' = 0 = \nabla \cdot \mathbf{F}(t)$ *for any time* $t \in [0,T]$ *and* $\mathbf{H}(t) \cdot \boldsymbol{\nu} = 0$ *on* $\partial\Omega$ *for any* $t \in [0,T]$*. Then the solution* $\mathbf{H}$ *of problem* (5.1) *belongs to* $\mathrm{C}\left([0,T], \mathbf{H}_0^1(\Omega)\right)$ *with* $\partial_t \mathbf{H} \in \mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right) \cap \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$ *and* $\partial_{tt} \mathbf{H} \in \mathrm{L}^2\left((0,T), \mathbf{H}^{-1}(\Omega)\right)$*.*

Now, the following error estimates can be derived. There might be no smaller constants C in comparison with the constants appearing in Theorem 5.2.3 because Grönwall's argument with exponential in time character of the constant cannot be avoided.

Theorem 5.4.2 (Error)**.** *Assume that* $\mathbf{F} \in \mathrm{Lip}([0,T], \mathbf{L}^2(\Omega))$*.*

(i) If $\mathbf{H}_0 \in \mathbf{H}^1(\Omega)$ and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$ then

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\|\nabla \int_0^t [\mathbf{H}_n - \mathbf{H}]\right\|^2 \leqslant C\tau.$$

(ii) If $\mathbf{H}_0 \in \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ and $\mathbf{H}_0' \in \mathbf{H}^1(\Omega)$ then

$$\max_{t\in[0,T]} \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 + \max_{t\in[0,T]} \left\|\nabla \int_0^t [\mathbf{H}_n - \mathbf{H}]\right\|^2 \leqslant C\tau^2.$$

Please note that the positive constant C in these estimates is of the form Ce^{CT}.

5.4.1 Modified scheme in $\mathrm{H}^1(\Omega)$

Last, the following semi-implicit time-discrete scheme for $\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$ is stated, where the convolution term is taken explicitly (from the last time step)

$$\left\{ \begin{aligned} \left(\delta^2 \mathbf{h}_i, \boldsymbol{\varphi}\right) + (\delta \mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) &= (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K} \star \mathbf{h}_{i-1}, \boldsymbol{\varphi}), \\ \mathbf{h}_0 &= \mathbf{H}_0, \end{aligned} \right. \tag{5.15}$$

which is equivalent to

$$\begin{aligned} a(\mathbf{h}_i, \boldsymbol{\varphi}) &:= \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right)(\mathbf{h}_i, \boldsymbol{\varphi}) + (\nabla \mathbf{h}_i, \nabla \boldsymbol{\varphi}) \\ &= (\mathbf{F}_i, \boldsymbol{\varphi}) - (\mathcal{K} \star \mathbf{h}_{i-1}, \boldsymbol{\varphi}) \\ &\qquad + \left(\frac{1}{\tau^2} + \frac{1}{\tau}\right)(\mathbf{h}_{i-1}, \boldsymbol{\varphi}) + \left(\frac{\delta \mathbf{h}_{i-1}}{\tau}, \boldsymbol{\varphi}\right) =: f_i(\boldsymbol{\varphi}). \end{aligned}$$

Via the Lax-Milgram lemma 2.11.1, the existence of a unique solution in $\mathbf{H}_0^1(\Omega)$ to (5.15) is obtained for any $i = 1, \ldots, n$ and any $\tau > 0$ if $\mathbf{H}_0 \in \mathbf{L}^q(\Omega)$, $q > 3$, and $\mathbf{H}_0' \in \mathbf{L}^2(\Omega)$. The scheme (5.15) can be analysed in the same way as (5.13). Remark that the error estimates from Theorem 5.4.2 are also valid for (5.15).

5.5 Conclusion

The well-posedness of a vectorial nonlocal linear hyperbolic problem (5.1) with applications in superconductors of type-I has been addressed. This model has been derived from the full Maxwell's equations, the two-fluid model of London and London, and the nonlocal representation (by a space convolution with a singular kernel) of the superconductive current by Eringen.

Two time-discrete schemes (based on an explicit and implicit handling of the convolution term) have been presented. The error estimates have been derived for both schemes. The solution of the original model satisfies a simpler problem (5.11) under the assumption that the normal component of the unknown vector field equals zero on the boundary of the superconductor. The convolution kernel in that problem is positive definite, but this does not lead to better error estimates for the time discretization. Up to now, no numerical experiments have been performed.

6

Macroscopic model for an intermediate state between type-I and type-II superconductivity

This chapter is based on the article [137], which is published in
Numerical Methods for Partial Differential Equations.

The aims of this chapter are to address the well-posedness of the following problem in terms of the magnetic field $\mathbf{H}$ for $\beta \geqslant 1$

$$\left\{ \begin{array}{rll} \partial_t \mathbf{H} + g(\beta)\nabla \times \left(|\nabla \times \mathbf{H}|^{\beta-1}\nabla \times \mathbf{H}\right) & & \\ +f(\beta)\nabla \times \nabla \times \mathbf{H} + f(\beta)\nabla \times (\mathcal{K}_0 \star \mathbf{H}) & = \mathbf{F} & \text{in } Q_T, \\ \mathbf{H} \times \boldsymbol{\nu} & = \mathbf{0} & \text{on } \Sigma_T, \\ \mathbf{H}(\mathbf{x}, 0) & = \mathbf{H}_0 & \text{in } \Omega, \end{array} \right. \tag{6.1}$$

to design a scheme for its numerical approximation and to derive error estimates for the time discretization.

The superconducting domain $\Omega \subset \mathbb{R}^3$ is a bounded Lipschitz domain. Note that $Q_T = \Omega \times (0, T]$ and $\Sigma_T = \partial\Omega \times (0, T]$, with T the final time. The value of β depends on the superconducting material. This problem is obtained by setting $\mu = \sigma = \sigma_c = 1$ (without loss of generality) in (3.14), where

$$(\mathcal{K}_0 \star \mathbf{H})(\mathbf{x}, t) = -\int_\Omega \sigma_0 \left(|\mathbf{x} - \mathbf{x}'|_{\mathrm{e}}\right) (\mathbf{x} - \mathbf{x}') \times \mathbf{H}(\mathbf{x}', t)\ \mathrm{d}\mathbf{x}',$$

with $\sigma_0 : (0, \infty) \to \mathbb{R}$ defined by

$$\sigma_0(s) = \begin{cases} \frac{\widetilde{C}}{2s^2} \exp\left(-\frac{s}{r_0}\right) & s < r_0, \\ 0 & s \geqslant r_0. \end{cases}$$

The parameters $\widetilde{C}$ and r_0 depend on the material under consideration. A source term $\mathbf{F}$ is added in the right-hand side of (3.14). The real function $f \in \mathrm{C}([1, \infty))$ is monotone decreasing and satisfies $f(1) = 1$ and $0 \leqslant f(\beta) \leqslant 1$ for $\beta > 1$. Moreover, it is supposed that f is zero or sufficiently small for $\beta > 7$. This implies that $g \in \mathrm{C}([1, \infty))$ defined by $g(\beta) = 1 - f(\beta)$ is monotone increasing with $g(1) = 0$ and $0 \leqslant g(\beta) \leqslant 1$. To obtain the magnetic boundary condition in (6.1), it is assumed that the magnetic field outside the domain Ω equals zero [39, p. 8]. For more information, the reader is referred to Subsection 3.3.3.

The techniques applied in this chapter are similar to the techniques presented in Chapter 4 and Chapter 5. To handle the nonlinearity in the model, the monotonicity methods and the Minty-Browder argument are employed. The outline of this chapter is as follows. Firstly, Section 6.1 studies the uniqueness of a solution to problem (6.1). Next, the well-posedness of the problem is shown in Section 6.2. More specifically, a semi-implicit time-discrete numerical scheme is developed. Also the existence of a weak solution for each time step is shown. Finally, the convergence of the method is discussed and error estimates for the time-discretization are derived.

6.1 Uniqueness of a solution

First, a variational formulation of (6.1) has to be established. The suitable choice for the space of test functions is ($\beta \geqslant 1$)

$$\mathbf{V}_0 = \left\{ \boldsymbol{\varphi} \in \mathbf{L}^2(\Omega) \, : \, \nabla \times \boldsymbol{\varphi} \in \mathbf{L}^{\beta+1}(\Omega) \text{ and } \boldsymbol{\varphi} \times \boldsymbol{\nu} = \mathbf{0} \text{ on } \Gamma \right\}, \tag{6.2}$$

which is a subset of $\mathbf{H}_0(\mathbf{curl}; \Omega)$. This is a closed subspace of the space

$$\mathbf{V} = \left\{ \boldsymbol{\varphi} \in \mathbf{L}^2(\Omega) \, : \, \nabla \times \boldsymbol{\varphi} \in \mathbf{L}^{\beta+1}(\Omega) \right\} \subset \mathbf{H}(\mathbf{curl}; \Omega), \tag{6.3}$$

and is endowed with the same graph norm

$$\|\boldsymbol{\varphi}\|_{\mathbf{V}} = \|\boldsymbol{\varphi}\|_{\mathbf{V}_0} = \|\boldsymbol{\varphi}\|_{\mathbf{L}^2(\Omega)} + \|\nabla \times \boldsymbol{\varphi}\|_{\mathbf{L}^{\beta+1}(\Omega)} . \tag{6.4}$$

Multiplying (6.1) by any $\boldsymbol{\varphi} \in \mathbf{V}_0$, integrating over the domain Ω and involving the Green theorem, it holds for a.a. $t \in (0, T]$ that

$$\begin{aligned} (\partial_t \mathbf{H}(t), \boldsymbol{\varphi}) + f(\beta) \left(\nabla \times \mathbf{H}(t), \nabla \times \boldsymbol{\varphi} \right) & \\ + g(\beta) \left(|\nabla \times \mathbf{H}(t)|^{\beta-1} \nabla \times \mathbf{H}(t), \nabla \times \boldsymbol{\varphi} \right) & \\ + f(\beta) \left(\mathcal{K}_0 \star \mathbf{H}(t), \nabla \times \boldsymbol{\varphi} \right) &= (\mathbf{F}(t), \boldsymbol{\varphi}) . \end{aligned} \tag{6.5}$$

For each $t \in (0,T]$, there is looked for a solution $\mathbf{H}(t) \in \mathbf{V}_0$. The estimates on the kernel $\mathcal{K}_0$ from Section 4.1 stay valid, for instance

$$(\mathcal{K}_0 \star \mathbf{H}_1, \nabla \times \mathbf{H}_2) \leqslant C_\varepsilon \|\mathbf{H}_1\|^2 + \varepsilon \|\nabla \times \mathbf{H}_2\|^2 \tag{6.6}$$

for all $\mathbf{H}_1 \in \mathbf{L}^2(\Omega)$ and $\mathbf{H}_2 \in \mathbf{H}(\mathbf{curl};\Omega)$. The position of the positive constants ε and C_ε can be interchanged.

Note that each term of (6.5) has to be well-defined for any $\mathbf{H}(t)$ and $\boldsymbol{\varphi} \in \mathbf{V}_0$. This can be easily checked by using the Cauchy and Hölder's inequality as follows ($\beta \geqslant 1$)

$$\begin{aligned}
|(\partial_t \mathbf{H}(t), \boldsymbol{\varphi})| &\leqslant \|\partial_t \mathbf{H}(t)\| \, \|\boldsymbol{\varphi}\|, \\
|(\nabla \times \mathbf{H}(t), \nabla \times \boldsymbol{\varphi})| &\leqslant \|\nabla \times \mathbf{H}(t)\| \, \|\nabla \times \boldsymbol{\varphi}\| \\
&\lesssim \|\nabla \times \mathbf{H}(t)\|_{\mathbf{L}^{\beta+1}(\Omega)} \|\nabla \times \boldsymbol{\varphi}\|_{\mathbf{L}^{\beta+1}(\Omega)}, \\
\left|\left(|\nabla \times \mathbf{H}(t)|^{\beta-1} \nabla \times \mathbf{H}(t), \nabla \times \boldsymbol{\varphi}\right)\right| &\leqslant \int_\Omega |\nabla \times \mathbf{H}(t)|^\beta |\nabla \times \boldsymbol{\varphi}| \\
&\leqslant \|\nabla \times \mathbf{H}(t)\|^\beta_{\mathbf{L}^{\beta+1}(\Omega)} \|\nabla \times \boldsymbol{\varphi}\|_{\mathbf{L}^{\beta+1}(\Omega)}, \\
|(\mathcal{K}_0 \star \mathbf{H}(t), \nabla \times \boldsymbol{\varphi})| &\overset{(6.6)}{\lesssim} \|\mathbf{H}(t)\|^2 + \|\nabla \times \boldsymbol{\varphi}\|^2 \\
&\lesssim \|\mathbf{H}(t)\|^2 + \|\nabla \times \boldsymbol{\varphi}\|^2_{\mathbf{L}^{\beta+1}(\Omega)}, \\
|(\mathbf{F}(t), \boldsymbol{\varphi})| &\leqslant \|\mathbf{F}(t)\| \, \|\boldsymbol{\varphi}\|.
\end{aligned}$$

The following lemma states the reflexivity of the spaces $\mathbf{V}$ and $\mathbf{V}_0$.

Lemma 6.1.1. *The vector spaces $\mathbf{V}$ and $\mathbf{V}_0$ are reflexive Banach spaces.*

Proof. The proof follows closely the proof of [138, Lemma 1]. The space $\mathbf{L}^p(\Omega)$ is a Banach space for $p \geqslant 1$ and is reflexive for $p > 1$. Employing this together with the definition of the vector space $\mathbf{V}$ and its norm implies that $\mathbf{V}$ is a Banach space. Then $\mathbf{V}_0$ is also a Banach space as closed subspace of $\mathbf{V}$, see Lemma 2.4.6. The proof of the reflexivity of $\mathbf{V}$ and $\mathbf{V}_0$ is given in more details below.

Let us define a vector space $\mathbf{X}$ as follows,

$$\mathbf{X} = \mathbf{L}^2(\Omega) \times \mathbf{L}^{\beta+1}(\Omega).$$

This space is a reflexive Banach space as the product of a finite number of reflexive Banach spaces, see Lemma 2.4.18(ii). Consider the following subset of $\mathbf{X}$:

$$\widetilde{\mathbf{V}} = \{(\mathbf{v}, \nabla \times \mathbf{v}) \subset \mathbf{X}\}.$$

Let $\{(\mathbf{v}_n, \nabla \times \mathbf{v}_n)\}$ be an arbitrary Cauchy sequence in $\widetilde{\mathbf{V}}$. Then $\{\mathbf{v}_n\}$ is a Cauchy sequence in $\mathbf{L}^2(\Omega)$. Therefore, there exists a $\mathbf{v} \in \mathbf{L}^2(\Omega)$ such that $\mathbf{v}_n \to \mathbf{v}$ in $\mathbf{L}^2(\Omega)$. Similarly, there exists a $\mathbf{f} \in \mathbf{L}^{\beta+1}(\Omega)$ such that $\nabla \times \mathbf{v}_n \to \mathbf{f}$ in $\mathbf{L}^{\beta+1}(\Omega)$.

From the definition of the curl-operator in the distributional sense, we directly obtain that $\mathbf{f} = \nabla \times \mathbf{v}$ in the sense of functionals on $\mathbf{C}_0^\infty(\Omega)$. Using the density of $\mathbf{C}_0^\infty(\Omega)$ in $\mathbf{L}^{\beta+1}(\Omega)$ and the Hahn-Banach theorem, $\nabla \times \mathbf{v}$ can be extended (in a unique way) to the whole space $\mathbf{L}^{\beta+1}(\Omega)$. As $\mathbf{f} \in \mathbf{L}^{\beta+1}(\Omega)$, we get that $\mathbf{f} = \nabla \times \mathbf{v}$ in $\mathbf{L}^{\beta+1}(\Omega)$. Thus $(\mathbf{v}, \mathbf{f}) \in \widetilde{\mathbf{V}}$. Consequently the set $\widetilde{\mathbf{V}}$ is a closed subset of $\mathbf{X}$ and following Lemma 2.4.18(i) it is a reflexive space. As $\mathbf{V}$ is isomorphic to $\widetilde{\mathbf{V}}$, the space $\mathbf{V}$ is also a reflexive Banach space, see Lemma 2.4.18(iii). Again, the space $\mathbf{V}_0$ is also reflexive as closed subspace of $\mathbf{V}$. $\square$

The following technical lemma is crucial in the proofs. The interested reader is referred to Lemma 2.3.16 for the proof.

Lemma 6.1.2. *Assume that $\alpha \geqslant 1$. There exists a positive constant $C_0(\alpha) = \frac{1}{4 \cdot 12^{\frac{\alpha+1}{2}}}$ such that for any $\mathbf{H}_1, \mathbf{H}_2 \in \mathbf{V}$ it holds that*

$$\begin{aligned}\left(|\nabla \times \mathbf{H}_1|^{\alpha-1} \nabla \times \mathbf{H}_1 - |\nabla \times \mathbf{H}_2|^{\alpha-1} \nabla \times \mathbf{H}_2, \nabla \times (\mathbf{H}_1 - \mathbf{H}_2)\right) \\ \geqslant C_0(\alpha) \left\| \nabla \times (\mathbf{H}_1 - \mathbf{H}_2) \right\|_{\mathbf{L}^{\alpha+1}(\Omega)}^{\alpha+1}.\end{aligned}$$

The following theorem describes the natural stability of the solution $\mathbf{H}$ of (6.1).

Theorem 6.1.1 (Stability)**.** *Suppose that $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ and that $\mathbf{H}$ solves* (6.1).

(i) If $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$, then

$$\max_{t \in [0,T]} \|\mathbf{H}(t)\|^2 + \int_0^T \|\nabla \times \mathbf{H}\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} \leqslant C.$$

(ii) If $\nabla \cdot \mathbf{F}(t) = 0 = \nabla \cdot \mathbf{H}_0$ for any $t \in [0,T]$, then $\nabla \cdot \mathbf{H}(t) = 0$ for any $t \in [0,T]$.

(iii) If $\mathbf{H}_0 \in \mathbf{V}$, then

$$\max_{t \in [0,T]} \|\nabla \times \mathbf{H}(t)\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} + \int_0^T \|\partial_t \mathbf{H}\|^2 \leqslant C.$$

(iv) If $\mathbf{F}(0) \in \mathbf{L}^2(\Omega)$, $\partial_t \mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, $\nabla \times (\mathcal{K}_0 \star \mathbf{H}_0) \in \mathbf{L}^2(\Omega)$, $\mathbf{H}_0 \in \mathbf{V}_0$, $\nabla \times \nabla \times \mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ and $\nabla \times \left[|\nabla \times \mathbf{H}_0|^{\beta-1} \nabla \times \mathbf{H}_0\right] \in \mathbf{L}^2(\Omega)$, then

$$\max_{t \in [0,T]} \|\partial_t \mathbf{H}(t)\|^2 \leqslant C.$$

Proof. (i) Setting $\boldsymbol{\varphi} = \mathbf{H}(t)$ in (6.5) and integrating in time over $t \in (0, \xi) \subset (0,T)$, we get, due to Lemma 6.1.2, that

$$\begin{aligned}\frac{\|\mathbf{H}(\xi)\|^2}{2} + f(\beta) \int_0^\xi \|\nabla \times \mathbf{H}\|^2 + g(\beta) C_0(\beta) \int_0^\xi \|\nabla \times \mathbf{H}\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} \\ \leqslant \frac{\|\mathbf{H}_0\|^2}{2} + \int_0^\xi (\mathbf{F}, \mathbf{H}) - f(\beta) \int_0^t (\mathcal{K}_0 \star \mathbf{H}, \nabla \times \mathbf{H}).\end{aligned}$$

Using the Cauchy and Young inequalities and inequality (6.6) for the last term on the RHS, we obtain that

$$\frac{\|\mathbf{H}(\xi)\|^2}{2} + f(\beta)\int_0^\xi \|\nabla\times\mathbf{H}\|^2 + g(\beta)C_0(\beta)\int_0^\xi \|\nabla\times\mathbf{H}\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} \leqslant C + C_\varepsilon\int_0^\xi \|\mathbf{H}\|^2 + \varepsilon\int_0^\xi \|\nabla\times\mathbf{H}\|^2 .$$

Now, we consider four cases:

- $\beta = 1$: then $f(\beta) = 1$ and $g(\beta) = 0$. Fixing a sufficiently small positive ε and applying Grönwall's argument, we get the asked estimate;
- $1 < \beta < 7$: then f and g are strict positive. Again, fixing a sufficiently small positive ε and applying the Grönwall argument gives the result;
- $\beta \geqslant 7$ and $f(\beta) = 0$ for $\beta \geqslant 7$: thus $g(\beta) = 1$ and the convolution term disappears from the problem. We trivially obtain the estimate;
- $\beta \geqslant 7$ and $f(\beta) > 0$ for $\beta \geqslant 7$ but sufficiently small: analogously as the case $1 < \beta < 7$.

(ii) Take the divergence of (6.1) or set $\boldsymbol{\varphi} = \nabla\phi$ with $\phi \in \mathrm{C}_0^\infty(\Omega)$ in (6.5). Then, integrate in time to arrive at $\nabla\cdot\mathbf{H}(t) = \nabla\cdot\mathbf{H}_0 = 0$ for all $t \in [0,T]$.

(iii) Note that $|\mathbf{u}|^{\beta-1}\mathbf{u}\cdot\partial_t\mathbf{u} = \partial_t\frac{|\mathbf{u}|^{\beta+1}}{\beta+1}$ due to $\partial_t|\mathbf{u}| = \frac{\mathbf{u}}{|\mathbf{u}|}\cdot\partial_t\mathbf{u}$. Now, we set $\boldsymbol{\varphi} = \partial_t\mathbf{H}(t)$ in (6.5) and integrate in time over $t \in (0,\xi) \subset (0,T)$ to obtain

$$\begin{aligned}\int_0^\xi \|\partial_t\mathbf{H}\|^2 + \frac{f(\beta)}{2}\|\nabla\times\mathbf{H}(\xi)\|^2 + \frac{g(\beta)}{\beta+1}\|\nabla\times\mathbf{H}(\xi)\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} \\ = \frac{f(\beta)}{2}\|\nabla\times\mathbf{H}_0\|^2 + \frac{g(\beta)}{\beta+1}\|\nabla\times\mathbf{H}_0\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} \\ + \int_0^\xi (\mathbf{F},\partial_t\mathbf{H}) - f(\beta)\int_0^\xi (\mathcal{K}_0\star\mathbf{H},\nabla\times\partial_t\mathbf{H}) .\end{aligned}$$

The last term in the RHS can be estimated like in Theorem 4.2.1(iii) by using integration by parts. Afterwards, the result follows the lines from (i).

(iv) We differentiate (6.5) with respect to the time variable. Therefore, we need that (6.5) is fulfilled for $t = 0$. Knowing that

$$\begin{aligned}&\nabla\times\nabla\times\mathbf{H}_0 \in \mathbf{L}^2(\Omega), && \nabla\times\left[|\nabla\times\mathbf{H}_0|^{\beta-1}\nabla\times\mathbf{H}_0\right] \in \mathbf{L}^2(\Omega),\\ &\nabla\times(\mathcal{K}_0\star\mathbf{H}_0) \in \mathbf{L}^2(\Omega), && \mathbf{H}_0 \in \mathbf{V}_0,\end{aligned}$$

we may define

$$\begin{aligned}\partial_t\mathbf{H}(0) = \mathbf{F}(0) - f(\beta)\nabla\times\nabla\times\mathbf{H}_0 \\ - g(\beta)\nabla\times\left(|\nabla\times\mathbf{H}_0|^{\beta-1}\nabla\times\mathbf{H}_0\right) - f(\beta)\nabla\times(\mathcal{K}_0\star\mathbf{H}_0),\end{aligned}$$

i.e.

$$\|\partial_t \mathbf{H}(0)\| \lesssim 1.$$

Now, we set $\boldsymbol{\varphi} = \partial_t \mathbf{H}(t)$ and integrate in time over $t \in (0,\xi) \subset (0,T)$ to get that

$$\begin{aligned}
&\tfrac{1}{2}\left\|\partial_t \mathbf{H}(\xi)\right\|^2 + f(\beta)\int_0^\xi \left\|\nabla\times\partial_t\mathbf{H}\right\|^2 \\
&\quad + g(\beta)\int_0^\xi\int_\Omega |\nabla\times\mathbf{H}|^{\beta-1}\left[|\nabla\times\partial_t\mathbf{H}|^2 + (\beta-1)\left(\partial_t|\nabla\times\mathbf{H}|\right)^2\right] \\
&\qquad = \tfrac{1}{2}\left\|\partial_t\mathbf{H}(0)\right\|^2 + \int_0^\xi \left(\partial_t\mathbf{F},\partial_t\mathbf{H}\right) - f(\beta)\int_0^\xi \left(\mathcal{K}_0 \star \partial_t\mathbf{H}, \nabla\times\partial_t\mathbf{H}\right).
\end{aligned}$$

In the last step, we have used that

$$\begin{aligned}
\partial_t\left(|\mathbf{u}|^{\alpha-1}\mathbf{u}\right)\cdot\partial_t\mathbf{u} &= |\partial_t\mathbf{u}|^2|\mathbf{u}|^{\alpha-1} + (\alpha-1)|\mathbf{u}|^{\alpha-3}|\mathbf{u}\cdot\partial_t\mathbf{u}|^2 \\
&= \left[|\partial_t\mathbf{u}|^2 + (\alpha-1)\left(\partial_t|\mathbf{u}|\right)^2\right]|\mathbf{u}|^{\alpha-1}.
\end{aligned}$$

Employing the Cauchy and Young inequalities, (6.6) and (iii) to the RHS, and applying Grönwall's argument (depending on the value of β), we arrive at the result. Note that the second and third term in the LHS cannot be zero together. $\square$

Remark 2. To obtain higher regularity of the solution, higher regularity of the known data is required, see Lemma 6.1.1(iv).

Remark 3. It holds that $\mathbf{H}(\mathbf{curl};\Omega) \hookrightarrow \mathbf{L}^2(\Omega)$. Moreover, it is true that

$$\mathbf{H} \in \mathrm{L}^2\left((0,T),\mathbf{H}(\mathbf{curl};\Omega)\right) \quad \text{and} \quad \partial_t\mathbf{H} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$$

if $\mathbf{H}_0 \in \mathbf{V}$ and $\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$. Applying Lemma 2.9.5(i), we get that $\mathbf{H} \in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$.

Now, it is possible to define the following weak formulation.

Definition 6.1.1. *Let $\beta \geqslant 1$, $\mathbf{H}_0 \in \mathbf{V}$ and $\mathbf{F} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$ be given. The variational formulation of* (6.1) *reads as: find $\mathbf{H} \in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$ with $\nabla\times\mathbf{H} \in \mathrm{L}^{\beta+1}\left((0,T),\mathbf{L}^{\beta+1}(\Omega)\right)$ and $\partial_t\mathbf{H} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$ such that*

$$\begin{aligned}
&\left(\partial_t\mathbf{H}(t),\boldsymbol{\varphi}\right) + f(\beta)\left(\nabla\times\mathbf{H}(t),\nabla\times\boldsymbol{\varphi}\right) \\
&\qquad + g(\beta)\left(|\nabla\times\mathbf{H}(t)|^{\beta-1}\nabla\times\mathbf{H}(t),\nabla\times\boldsymbol{\varphi}\right) \\
&\qquad + f(\beta)\left(\mathcal{K}_0\star\mathbf{H}(t),\nabla\times\boldsymbol{\varphi}\right) = \left(\mathbf{F}(t),\boldsymbol{\varphi}\right), \quad \forall\boldsymbol{\varphi}\in\mathbf{V}_0, \qquad (6.7)
\end{aligned}$$

for a.a. $t \in [0,T]$.

The following theorem guarantees the uniqueness of the solution to problem (6.1).

Theorem 6.1.2 (Uniqueness)**.** *The problem* (6.1) *admits at most one solution $\mathbf{H} \in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right)$ with $\nabla\times\mathbf{H} \in \mathrm{L}^{\beta+1}\left((0,T),\mathbf{L}^{\beta+1}(\Omega)\right)$.*

Proof. Assume that we have two solutions $\mathbf{H}_1$ and $\mathbf{H}_2$. Set $\mathbf{H} = \mathbf{H}_1 - \mathbf{H}_2$. Then $\mathbf{H}_0 = \mathbf{0}$. Subtract equation (6.7) for $\mathbf{H} = \mathbf{H}_1$ from (6.7) for $\mathbf{H} = \mathbf{H}_2$. Setting $\boldsymbol{\varphi} = \mathbf{H}(t)$ into the resulting equation and integrating in time over $t \in (0, \xi) \subset (0, T)$, we find thanks to Lemma 6.1.2 that

$$\begin{aligned}\tfrac{1}{2}\left\|\mathbf{H}(\xi)\right\|^2 + f(\beta)\int_0^\xi \left\|\nabla\times\mathbf{H}\right\|^2 \\ + g(\beta)C_0(\beta)\int_0^\xi \left\|\nabla\times\mathbf{H}\right\|^{\beta+1}_{\mathbf{L}^{\beta+1}(\Omega)} \leqslant -f(\beta)\int_0^\xi \left(\mathcal{K}_0 \star \mathbf{H}, \nabla\times\mathbf{H}\right).\end{aligned}$$

Using inequality (6.6) for the term on the RHS, we arrive at

$$\begin{aligned}\tfrac{1}{2}\left\|\mathbf{H}(\xi)\right\|^2 + f(\beta)\int_0^\xi \left\|\nabla\times\mathbf{H}\right\|^2 \\ + g(\beta)C_0(\beta)\int_0^\xi \left\|\nabla\times\mathbf{H}\right\|^{\beta+1}_{\mathbf{L}^{\beta+1}(\Omega)} \leqslant C_\varepsilon\int_0^\xi \left\|\mathbf{H}\right\|^2 + \varepsilon\int_0^\xi \left\|\nabla\times\mathbf{H}\right\|^2.\end{aligned}$$

We consider again four cases:

- $\beta = 1$: then $f(\beta) = 1$ and $g(\beta) = 0$. Fixing a sufficiently small positive ε and applying Grönwall's argument, we get that $\mathbf{H} = \mathbf{0}$ a.e. in Q_T;
- $1 < \beta < 7$: then f and g are strict positive. Again fixing a sufficiently small positive ε and applying Grönwall's argument gives that $\mathbf{H} = \mathbf{0}$ a.e. in Q_T;
- $\beta \geqslant 7$ and $f(\beta) = 0$ for $\beta \geqslant 7$: thus $g(\beta) = 1$ and the convolution term disappears from the problem. We immediately obtain that $\mathbf{H} = \mathbf{0}$ a.e. in Q_T;
- $\beta \geqslant 7$ and $f(\beta) > 0$ for $\beta \geqslant 7$ but sufficiently small: analogously as the case $1 < \beta < 7$. □

Remark 4. In the previous theorems, four cases are considered depending on the value of the parameter β. These situations are not repeated in the remainder of the chapter but should be reconsidered by the reader in the a priori estimates and in the convergence result.

6.2 Existence of a solution

To address the existence of a solution to (6.1), a semidiscretization in time is employed, which is based on Rothe's method. The interval $[0, T]$ is divided into n equidistant subintervals $[t_{i-1}, t_i]$, $i = 1, \ldots, n$, with time step $\tau = \frac{T}{n} < 1$, thus $t_i = i\tau, i = 0, \ldots, n$. With the standard notation for the discretized fields

$$\mathbf{h}_i \approx \mathbf{H}(t_i) \quad \text{and} \quad \delta\mathbf{h}_i = \frac{\mathbf{h}_i - \mathbf{h}_{i-1}}{\tau},$$

the following linear recurrent semi-implicit scheme is proposed to approximate the original problem

$$\left\{\begin{array}{rl} (\delta \mathbf{h}_i, \boldsymbol{\varphi}) & + g(\beta)\left(|\nabla \times \mathbf{h}_i|^{\beta-1}\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}\right) \\ +f(\beta)\left(\nabla \times \mathbf{h}_i, \nabla \times \boldsymbol{\varphi}\right) & = (\mathbf{F}_i, \boldsymbol{\varphi}) - f(\beta)\left(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \boldsymbol{\varphi}\right), \\ \mathbf{h}_0 & = \mathbf{H}_0, \end{array}\right. \tag{6.8}$$

which is equivalent to solving, on each time step, the operator equation $A(\mathbf{u}) = \mathbf{F}_i^*$ in which $A : \mathbf{V}_0 \to \mathbf{V}_0^*$ is defined by

$$\langle A(\mathbf{u}), \mathbf{v}\rangle = \left(\frac{\mathbf{u}}{\tau}, \mathbf{v}\right) + f(\beta)\left(\nabla \times \mathbf{u}, \nabla \times \mathbf{v}\right) + g(\beta)\left(|\nabla \times \mathbf{u}|^{\beta-1}\nabla \times \mathbf{u}, \nabla \times \mathbf{v}\right)$$

and $\mathbf{F}_i^* : \mathbf{V}_0 \to \mathbb{R}$ by

$$\langle \mathbf{F}_i^*, \mathbf{v}\rangle = (\mathbf{F}_i, \mathbf{v}) - f(\beta)\left(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \mathbf{v}\right) + \left(\frac{\mathbf{h}_{i-1}}{\tau}, \mathbf{v}\right). \tag{6.9}$$

The solution at the previous time step is substituted in the convolution term instead of the solution at the actual time step because this should be easier to implement. The focus in this chapter is not on the implementation of the numerical scheme, but on its analysis.

The existence and uniqueness of a weak solution on each time step is guaranteed by the following theorem.

Theorem 6.2.1 (Uniqueness on a single time step)**.** *Assume that* $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$ *and* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. *Then there exists a* $\tau_0 > 0$ *such that the variational problem* (6.8) *has a unique solution for any* $i = 1, \ldots, n$ *and any* $\tau < \tau_0$.

Proof. The space $\mathbf{V}_0^*$ is a reflexive Banach space, see Lemma (6.1.1). Therefore, applying Theorem 2.11.8, the operator equation $A(\mathbf{u}) = \mathbf{F}_i^*$ has a unique solution on each time step because A is a strictly monotone, coercive, demicontinuous operator and $\mathbf{F}_i^* \in \mathbf{V}_0^*$. In particular, the strict monotonicity of A follows from Lemma 6.1.2. □

First, basic stability result for $\mathbf{h}_i$ are derived. The a priori estimates in parts (i) and (iii) of Lemma 6.2.2 serve as uniform bounds to prove convergence. In the proofs, the following lemma is needed, see [139, Lemma 2.3].

Lemma 6.2.1. *Let* $g : \mathbb{R} \to \mathbb{R}$ *be a nonnegative continuous function such that* $G(s) := g(s)s$ *is monotone increasing. Let* Φ_G *be the primitive function of G, i.e.* $\Phi_G(s) = \int_0^s G(s)\ \mathrm{d}s$. *Then for any* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, *it holds that*

$$\Phi_G\left(|\mathbf{y}|\right) - \Phi_G\left(|\mathbf{x}|\right) \leqslant g(|\mathbf{y}|)\mathbf{y} \cdot (\mathbf{y} - \mathbf{x}).$$

Lemma 6.2.2 (A priori estimates)**.** *Suppose that* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

(i) Assume that $\mathbf{H}_0 \in \mathbf{L}^2(\Omega)$*. Then, there exists a positive constant* C *such that*

$$\max_{1\leqslant i\leqslant n} \|\mathbf{h}_i\|^2 + \sum_{i=1}^{n} \|\mathbf{h}_i - \mathbf{h}_{i-1}\|^2 + \sum_{i=1}^{n} \|\nabla\times\mathbf{h}_i\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1}\,\tau \leqslant C$$

for all $\tau < \tau_0$*.*

(ii) If $\nabla\cdot\mathbf{H}_0 = 0 = \nabla\cdot\mathbf{F}_i$ *for all* $i = 1,\ldots,n$*, then* $\nabla\cdot\mathbf{h}_i = 0$ *for all* $i = 1,\ldots,n$*.*

(iii) If $\mathbf{H}_0 \in \mathbf{V}$ *then*

$$\max_{1\leqslant i\leqslant n} \|\nabla\times\mathbf{h}_i\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} + \sum_{i=1}^{n} \|\delta\mathbf{h}_i\|^2\,\tau \leqslant C$$

for all $\tau < \tau_0$*.*

Proof. (i) Setting $\boldsymbol{\varphi} = \mathbf{h}_i$ in (6.8), multiplying by τ and summing the result up for $i = 1,\ldots,j$ $(1 \leqslant j \leqslant n)$, we have that

$$\begin{aligned}\sum_{i=1}^{j} (\delta\mathbf{h}_i,\mathbf{h}_i)\,\tau + f(\beta)\sum_{i=1}^{j} \|\nabla\times\mathbf{h}_i\|^2\,\tau \\ + g(\beta)\sum_{i=1}^{j}\left(|\nabla\times\mathbf{h}_i|^{\beta-1}\nabla\times\mathbf{h}_i,\nabla\times\mathbf{h}_i\right)\tau \\ = \sum_{i=1}^{j}(\mathbf{F}_i,\mathbf{h}_i)\,\tau - f(\beta)\sum_{i=1}^{j}(\mathcal{K}_0\star\mathbf{h}_{i-1},\nabla\times\mathbf{h}_i)\,\tau.\end{aligned}$$

The first term on the left-hand side (LHS) can be rewritten using Abel's summation rule. The third term can be estimated below thanks to Lemma 6.1.2. The second term on the RHS can be estimated as in Lemma 4.4.1(i). An application of Grönwall's lemma completes the proof.

(ii) The result can be readily obtained by applying the divergence operator to

$$\begin{aligned}\delta\mathbf{h}_i + f(\beta)\nabla\times\nabla\times\mathbf{h}_i + g(\beta)\nabla\times\left(|\nabla\times\mathbf{h}_i|^{\beta-1}\nabla\times\mathbf{h}_i\right) \\ + f(\beta)\nabla\times(\mathcal{K}_0\star\mathbf{h}_{i-1}) = \mathbf{F}_i,\end{aligned}$$

or setting $\boldsymbol{\varphi} = \nabla\phi$ with $\phi \in \mathrm{C}_0^\infty(\Omega)$ in (6.8).

(iii) Setting $\boldsymbol{\varphi} = \delta\mathbf{h}_i$ in (6.8), multiplying by τ and summing it up for $i = 1,\ldots,j$

($1 \leqslant j \leqslant n$), we have that

$$\sum_{i=1}^{j} \|\delta \mathbf{h}_i\|^2 \tau + f(\beta) \sum_{i=1}^{j} \left(\nabla \times \mathbf{h}_i, \nabla \times \mathbf{h}_i - \nabla \times \mathbf{h}_{i-1}\right)$$
$$+ g(\beta) \sum_{i=1}^{j} \left(|\nabla \times \mathbf{h}_i|^{\beta-1} \nabla \times \mathbf{h}_i, \nabla \times \delta \mathbf{h}_i\right) \tau$$
$$= \sum_{i=1}^{j} \left(\mathbf{F}_i, \delta \mathbf{h}_i\right) \tau - f(\beta) \sum_{i=1}^{j} \left(\mathcal{K}_0 \star \mathbf{h}_{i-1}, \nabla \times \delta \mathbf{h}_i\right) \tau.$$

The second term in the LHS can be estimated by Abel's summation rule. The third term on the LHS can be estimated below by using Lemma 6.2.1 with $g(s) = s^{\beta-1}$. We obtain

$$\begin{aligned}
&\sum_{i=1}^{j} \left(|\nabla \times \mathbf{h}_i|^{\beta-1} \nabla \times \mathbf{h}_i, \nabla \times \delta \mathbf{h}_i\right) \tau \\
&\quad \geqslant \quad \frac{1}{\beta+1} \sum_{i=1}^{j} \int_\Omega \left[|\nabla \times \mathbf{h}_i|^{\beta+1} - |\nabla \times \mathbf{h}_{i-1}|^{\beta+1}\right] \\
&\quad = \quad \frac{1}{\beta+1} \left(\|\nabla \times \mathbf{h}_j\|^{\beta+1}_{\mathbf{L}^{\beta+1}(\Omega)} - \|\nabla \times \mathbf{H}_0\|^{\beta+1}_{\mathbf{L}^{\beta+1}(\Omega)}\right).
\end{aligned}$$

The last term on the RHS can be estimated as in Lemma 4.4.1(iii). Using (i), we conclude the proof. □

The existence of a weak solution is proved using Rothe's method. The following piecewise linear in time vector fields $\mathbf{H}_n$ and the piecewise constant in time fields $\overline{\mathbf{H}}_n$ are introduced as

$$\begin{aligned}
\mathbf{H}_n(0) &= \mathbf{H}_0 \\
\mathbf{H}_n(t) &= \mathbf{h}_{i-1} + (t - t_{i-1})\delta \mathbf{h}_i \qquad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \ldots, n
\end{aligned}$$

and

$$\overline{\mathbf{H}}_n(0) = \mathbf{H}_0, \qquad \overline{\mathbf{H}}_n(t) = \mathbf{h}_i, \quad \text{for } t \in (t_{i-1}, t_i], \qquad i = 1, \ldots, n$$

respectively. Similarly, the vector field $\overline{\mathbf{F}}_n$ is defined. The variational formulation (6.8) can be rewritten for all $\boldsymbol{\varphi} \in \mathbf{V}_0$ and a.a. $t \in (0, T)$ as

$$\begin{aligned}
&(\partial_t \mathbf{H}_n(t), \boldsymbol{\varphi}) + f(\beta) \left(\nabla \times \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) \\
&\quad + g(\beta) \left(|\nabla \times \overline{\mathbf{H}}_n(t)|^{\beta-1} \nabla \times \overline{\mathbf{H}}_n(t), \nabla \times \boldsymbol{\varphi}\right) \\
&\quad = \left(\overline{\mathbf{F}}_n(t), \boldsymbol{\varphi}\right) - f(\beta) \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n(t-\tau), \nabla \times \boldsymbol{\varphi}\right). \qquad (6.10)
\end{aligned}$$

Now, the convergence of the sequences $\{\mathbf{H}_n\}$ and $\{\overline{\mathbf{H}}_n\}$ to the unique weak solution of (6.1) is proved as $\tau \to 0$ or $n \to \infty$.

Theorem 6.2.2 (Existence)**.** *Let* $\mathbf{H}_0 \in \mathbf{V}$ *and* $\mathbf{F} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. *Assume that* $\nabla \cdot \mathbf{H}_0 = 0 = \nabla \cdot \mathbf{F}(t)$ *for any time* $t \in [0,T]$. *Then there exists a weak solution* $\mathbf{H}$ *such that*

(i) $\overline{\mathbf{H}}_n \rightharpoonup \mathbf{H}$ *in* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, $\nabla \times \overline{\mathbf{H}}_n \rightharpoonup \nabla \times \mathbf{H}$ *in* $\mathrm{L}^{\beta+1}\left((0,T), \mathbf{L}^{\beta+1}(\Omega)\right)$ *and* $\mathbf{H}_n \rightharpoonup \mathbf{H}$ *in* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$;

(ii) $\mathbf{H}_n \to \mathbf{H}$ *in* $\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$, $\partial_t \mathbf{H}_n \rightharpoonup \partial_t \mathbf{H}$ *in* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ *and* $\overline{\mathbf{H}}_n \to \mathbf{H}$ *in* $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$;

(iii) $|\nabla \times \overline{\mathbf{H}}_n|^{\beta-1} \nabla \times \overline{\mathbf{H}}_n \rightharpoonup |\nabla \times \mathbf{H}|^{\beta-1} \nabla \times \mathbf{H}$ *in* $\mathrm{L}^{\frac{\beta+1}{\beta}}\left((0,T), \mathbf{L}^{\frac{\beta+1}{\beta}}(\Omega)\right)$

(iv) $\mathbf{H} \in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$ *is a weak solution of* (6.7).

Proof. (i) The spaces $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ and $\mathrm{L}^{\beta+1}\left((0,T), \mathbf{L}^{\beta+1}(\Omega)\right)$ are reflexive Banach spaces. Thanks to Lemma 6.2.2(i) and (iii), the sequence $\{\overline{\mathbf{H}}_n\}$ is bounded in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$, the sequence $\{\nabla \times \overline{\mathbf{H}}_n\}$ is bounded in $\mathrm{L}^{\beta+1}\left((0,T), \mathbf{L}^{\beta+1}(\Omega)\right)$ and the sequence $\{\mathbf{H}_n\}$ is bounded in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. Therefore, the sequence $\{\overline{\mathbf{H}}_n\}$ contains a weakly convergence subsequence (denoted by the same symbol again) such that $\overline{\mathbf{H}}_n \rightharpoonup \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$ and $\nabla \times \overline{\mathbf{H}}_n \rightharpoonup \mathbf{z}$ in $\mathrm{L}^{\beta+1}\left((0,T), \mathbf{L}^{\beta+1}(\Omega)\right)$. According to the Hahn-Banach theorem, it is easy to show that $\mathbf{z} = \nabla \times \mathbf{H} \in \mathrm{L}^{\beta+1}\left((0,T), \mathbf{L}^{\beta+1}(\Omega)\right)$. Employing Lemma 6.2.2(i) gives

$$\lim_{n\to\infty} \|\mathbf{H}_n - \overline{\mathbf{H}}_n\|^2_{\mathrm{L}^2((0,T),\mathbf{L}^2(\Omega))} = 0.$$

Thus $\{\overline{\mathbf{H}}_n\}$ and $\{\mathbf{H}_n\}$ have the same limit in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$. Therefore, $\mathbf{H}_n \rightharpoonup \mathbf{H}$ in $\mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$.

(ii) Lemma 6.2.2 implies for $i = 1, \ldots, n$ that

$$\mathbf{h}_i \in \mathbf{L}^2(\Omega), \quad \nabla \times \mathbf{h}_i \in \mathbf{L}^2(\Omega), \quad \nabla \cdot \mathbf{h}_i = 0 \text{ in } \Omega, \quad \mathbf{h}_i \times \boldsymbol{\nu} = \mathbf{0} \text{ on } \Gamma.$$

Employing Theorem 2.9.36, we see that $\mathbf{h}_i \in \mathbf{H}^{\frac{1}{2}}(\Omega)$, $i = 1, \ldots, n$. Also $\mathbf{H}_0$ belongs to $\mathbf{H}^{\frac{1}{2}}(\Omega)$. Thus $\max\limits_{t\in[0,T]} \left\|\overline{\mathbf{H}}_n(t)\right\|_{\mathbf{H}^{\frac{1}{2}}(\Omega)} \leqslant C$. Thanks to Lemma 6.2.2(iii), we have that $\displaystyle\int_0^T \|\partial_t \mathbf{H}_n\|^2 \leqslant C$. Using Theorem 2.9.37, we see that

$$\mathbf{H}^{\frac{1}{2}}(\Omega) \hookrightarrow\hookrightarrow \mathbf{L}^2(\Omega).$$

Then, applying Lemma 2.12.3, there exists a $\mathbf{H} \in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$ and a subsequence of $\{\mathbf{H}_n\}$ (denoted by the same symbol again) for which we have that

$$\begin{cases} \mathbf{H}_n \to \mathbf{H} & \text{in } \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right), \\ \partial_t \mathbf{H}_n \rightharpoonup \partial_t \mathbf{H} & \text{in } \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right). \end{cases}$$

From $\displaystyle\int_0^T \|\partial_t \mathbf{H}_n\|^2 \leqslant C$, it also follows that $\overline{\mathbf{H}}_n \to \mathbf{H}$ in $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$.

(iii) In this part of the proof, we apply Minty-Browder's trick [1, Chapter 9]. Due to the monotonicity, see Lemma 6.1.2, we can write that

$$\int_0^T \left(|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n - |\nabla\times\mathbf{u}|^{\beta-1}\nabla\times\mathbf{u}, \nabla\times\overline{\mathbf{H}}_n - \nabla\times\mathbf{u}\right) \geqslant 0, \quad (6.11)$$

for all $\mathbf{u}$ with $\nabla\times\mathbf{u}\in\mathrm{L}^{\beta+1}\left((0,T),\mathbf{L}^{\beta+1}(\Omega)\right)$. We want to take the limit $n\to\infty$ in (6.11). Because $\nabla\times\overline{\mathbf{H}}_n \rightharpoonup \nabla\times\mathbf{H}$ in $\mathrm{L}^{\beta+1}\left((0,T),\mathbf{L}^{\beta+1}(\Omega)\right)$, we have that

$$\lim_{n\to\infty}\int_0^T \left(|\nabla\times\mathbf{u}|^{\beta-1}\nabla\times\mathbf{u}, \nabla\times\overline{\mathbf{H}}_n - \nabla\times\mathbf{u}\right) = \int_0^T \left(|\nabla\times\mathbf{u}|^{\beta-1}\nabla\times\mathbf{u}, \nabla\times\mathbf{H} - \nabla\times\mathbf{u}\right).$$

From Lemma 6.2.2(iii), we get that

$$|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n \in \mathrm{L}^{\frac{\beta+1}{\beta}}\left((0,T),\mathbf{L}^{\frac{\beta+1}{\beta}}(\Omega)\right),$$

which is a reflexive Banach space. Therefore, $|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n \rightharpoonup \mathbf{z}$ in $\mathrm{L}^{\frac{\beta+1}{\beta}}\left((0,T),\mathbf{L}^{\frac{\beta+1}{\beta}}(\Omega)\right)$. Hence,

$$\lim_{n\to\infty}\int_0^T \left(|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n, \nabla\times\mathbf{u}\right) = \int_0^T (\mathbf{z},\nabla\times\mathbf{u}).$$

Note that $\overline{\mathbf{F}}_n \rightharpoonup \mathbf{F}$ in $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$. Furthermore, due to (i) and (ii), we obtain that

$$\begin{aligned}
&\lim_{n\to\infty} g(\beta)\int_0^T \left(|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n, \nabla\times\overline{\mathbf{H}}_n\right)\\
&\overset{(6.10)}{=} \lim_{n\to\infty}\int_0^T \Big[\left(\overline{\mathbf{F}}_n,\overline{\mathbf{H}}_n\right) - \left(\partial_t\mathbf{H}_n,\overline{\mathbf{H}}_n\right)\\
&\qquad -f(\beta)\left(\nabla\times\overline{\mathbf{H}}_n,\nabla\times\overline{\mathbf{H}}_n\right) - f(\beta)\left(\mathcal{K}_0\star\overline{\mathbf{H}}_n(t-\tau),\nabla\times\overline{\mathbf{H}}_n\right)\Big]\\
&\overset{(\star)}{\leqslant} \int_0^T \left[(\mathbf{F},\mathbf{H}) - (\partial_t\mathbf{H},\mathbf{H}) - f(\beta)\left[(\nabla\times\mathbf{H},\nabla\times\mathbf{H}) - (\mathcal{K}_0\star\mathbf{H},\nabla\times\mathbf{H})\right]\right]\\
&= \lim_{n\to\infty}\int_0^T \Big[(\mathbf{F}_n,\mathbf{H}) - (\partial_t\mathbf{H}_n,\mathbf{H})\\
&\qquad -f(\beta)\left(\nabla\times\overline{\mathbf{H}}_n,\nabla\times\mathbf{H}\right) - f(\beta)\left(\mathcal{K}_0\star\overline{\mathbf{H}}_n(t-\tau),\nabla\times\mathbf{H}\right)\Big]\\
&\overset{(6.10)}{=} \lim_{n\to\infty} g(\beta)\int_0^T \left(|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n, \nabla\times\mathbf{H}\right)\\
&= g(\beta)\int_0^T (\mathbf{z},\nabla\times\mathbf{H}).
\end{aligned}$$

The inequality $(\star)$ is valid by the weak lower semicontinuity of the norm, more specifically, $\nabla \times \overline{\mathbf{H}}_n \rightharpoonup \nabla \times \mathbf{H}$ in $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$ implies that

$$\|\nabla \times \mathbf{H}\|^2 \leqslant \liminf_{n\to\infty} \left\|\nabla \times \overline{\mathbf{H}}_n\right\|^2.$$

Therefore, passing to the limit for $n \to \infty$ in (6.11), we get

$$\int_0^T \left(\mathbf{z} - |\nabla \times \mathbf{u}|^{\beta-1}\nabla \times \mathbf{u}, \nabla \times \mathbf{H} - \nabla \times \mathbf{u}\right) \geqslant 0. \tag{6.12}$$

Now, we continue with the last step of the Minty-Browder trick. We show that $\mathbf{z} = |\nabla \times \mathbf{H}|^{\beta-1}\nabla \times \mathbf{H}$. Firstly, we put $\mathbf{u} = \mathbf{H} + \varepsilon\mathbf{v}$ for any $\mathbf{v}$ with $\nabla \times \mathbf{v} \in \mathrm{L}^{\beta+1}\left((0,T),\mathbf{L}^{\beta+1}(\Omega)\right)$ and $\varepsilon > 0$. Then, we get for (6.12) after dividing by $-\varepsilon$ that

$$\int_0^T \left(\mathbf{z} - |\nabla \times (\mathbf{H}+\varepsilon\mathbf{v})|^{\beta-1}\nabla \times (\mathbf{H}+\varepsilon\mathbf{v}), \nabla \times \mathbf{v}\right) \leqslant 0.$$

Next, taking the limit $\varepsilon \to 0$, we get

$$\int_0^T \left(\mathbf{z} - |\nabla \times \mathbf{H}|^{\beta-1}\nabla \times \mathbf{H}, \nabla \times \mathbf{v}\right) \leqslant 0.$$

The reverse inequality also holds true ($\mathbf{v} \leftrightarrow -\mathbf{v}$) and therefore

$$\int_0^T \left(\mathbf{z} - |\nabla \times \mathbf{H}|^{\beta-1}\nabla \times \mathbf{H}, \nabla \times \mathbf{v}\right) = 0,$$

for all $\mathbf{v}$ with $\nabla \times \mathbf{v} \in \mathrm{L}^{\beta+1}\left((0,T),\mathbf{L}^{\beta+1}(\Omega)\right)$. From this, we conclude that $\mathbf{z} = |\nabla \times \mathbf{H}|^{\beta-1}\nabla \times \mathbf{H}$ a.e. in Q_T.

(iv) Let us integrate (6.10) in time to get for any $\xi \in (0,T)$ and $\boldsymbol{\varphi} \in \mathbf{V}_0$ that

$$\begin{aligned}\int_0^\xi (\partial_t \mathbf{H}_n, \boldsymbol{\varphi}) + f(\beta)\int_0^\xi \left(\nabla \times \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right) \\ + g(\beta)\int_0^\xi \left(|\nabla \times \overline{\mathbf{H}}_n|^{\beta-1}\nabla \times \overline{\mathbf{H}}_n, \nabla \times \boldsymbol{\varphi}\right) \\ = \int_0^\xi \left(\overline{\mathbf{F}}_n, \boldsymbol{\varphi}\right) - f(\beta)\int_0^\xi \left(\mathcal{K}_0 \star \overline{\mathbf{H}}_n(t-\tau), \nabla \times \boldsymbol{\varphi}\right).\end{aligned}$$

We pass to the limit for $n \to \infty$. On the LHS, we use for the first term (ii), for the second term (i) and finally for the third term we apply (iii). For the RHS, we apply that $\overline{\mathbf{F}}_n \rightharpoonup \mathbf{F}$ in $\mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$ and (ii). We arrive at

$$\begin{aligned}\int_0^\xi (\partial_t \mathbf{H},\boldsymbol{\varphi}) + f(\beta)\int_0^\xi (\nabla\times\mathbf{H},\nabla\times\boldsymbol{\varphi})&\\ + g(\beta)\int_0^\xi \left(|\nabla\times\mathbf{H}|^{\beta-1}\nabla\times\mathbf{H},\nabla\times\boldsymbol{\varphi}\right)&\\ = \int_0^\xi (\mathbf{F},\boldsymbol{\varphi}) - f(\beta)\int_0^\xi (\mathcal{K}_0\star\mathbf{H},\nabla\times\boldsymbol{\varphi})&.\end{aligned}$$

Finally, differentiating the resulting identity with respect to the time variable ξ shows that $\mathbf{H}$ is a weak solution of (6.7). Up to now, we only have proven the convergence of the approximate solution for a subsequence of $\{\overline{\mathbf{H}}_n\}$. But, if we take into account Theorem 6.1.2, we obtain the convergence of the whole sequence to the unique weak solution of (6.7) in corresponding spaces. □

The following theorem addresses the error estimates for the time discretization.

Theorem 6.2.3 (Error)**.** *Suppose that* $\mathbf{F}\in\mathrm{Lip}([0,T],\mathbf{L}^2(\Omega))$. *If* $\mathbf{H}_0\in\mathbf{V}$ *then*

$$\max_{t\in[0,T]}\|\mathbf{H}_n(t)-\mathbf{H}(t)\|^2 + \int_0^T \left\|\nabla\times[\overline{\mathbf{H}}_n-\mathbf{H}]\right\|_{\mathbf{L}^{\beta+1}(\Omega)}^{\beta+1} \leqslant C\tau.$$

Please note that the positive constant C in this estimate is of the form Ce^{CT}.

Proof. We subtract (6.7) from (6.10), set $\boldsymbol{\varphi}=\overline{\mathbf{H}}_n(t)-\mathbf{H}(t)$ and integrate in time over $t\in(0,\xi)\subset(0,T)$ to get

$$\begin{aligned}&\tfrac{1}{2}\|\mathbf{H}_n(\xi)-\mathbf{H}(\xi)\|^2 + f(\beta)\int_0^\xi \left\|\nabla\times\overline{\mathbf{H}}_n-\nabla\times\mathbf{H}\right\|^2\\ &+ g(\beta)\int_0^\xi \left(|\nabla\times\overline{\mathbf{H}}_n|^{\beta-1}\nabla\times\overline{\mathbf{H}}_n - |\nabla\times\mathbf{H}|^{\beta-1}\nabla\times\mathbf{H},\nabla\times(\overline{\mathbf{H}}_n-\mathbf{H})\right)\\ &\qquad = \int_0^\xi \left(\overline{\mathbf{F}}_n-\mathbf{F},\overline{\mathbf{H}}_n-\mathbf{H}\right) + \int_0^\xi \left(\partial_t\mathbf{H}_n-\partial_t\mathbf{H},\mathbf{H}_n-\overline{\mathbf{H}}_n\right)\\ &\qquad\quad + \int_0^\xi \left(\mathcal{K}_0\star[\overline{\mathbf{H}}_n(t-\tau)-\mathbf{H}(t)],\nabla\times[\overline{\mathbf{H}}_n(t)-\mathbf{H}(t)]\right)\mathrm{d}t. \qquad (6.13)\end{aligned}$$

In the following estimates, we frequently use that

$$\left\|\mathbf{H}_n(t)-\overline{\mathbf{H}}_n(t)\right\| \leqslant \tau\left\|\partial_t\mathbf{H}_n(t)\right\| \qquad \text{for } t\in[0,T].$$

The third term in the LHS can be bounded below by Lemma 6.1.2. The first term in the RHS can be estimated by employing the Lipschitz continuity of $\mathbf{F}$, see

Theorem 4.3.3. For the last term of (6.13), we calculate that

$$\left|\int_0^\xi \left(\mathcal{K}_0 \star [\overline{\mathbf{H}}_n(t-\tau) - \mathbf{H}(t)], \nabla \times [\overline{\mathbf{H}}_n(t) - \mathbf{H}(t)]\right) \mathrm{d}t\right|$$
$$\overset{(6.6)}{\leqslant} \varepsilon \int_0^\xi \left\|\nabla \times [\overline{\mathbf{H}}_n(t) - \mathbf{H}(t)]\right\|^2 \mathrm{d}t + C_\varepsilon \int_0^\xi \left\|\overline{\mathbf{H}}_n(t-\tau) - \mathbf{H}(t)\right\|^2 \mathrm{d}t$$
$$\leqslant \varepsilon \int_0^\xi \left\|\nabla \times [\overline{\mathbf{H}}_n(t) - \mathbf{H}(t)]\right\|^2 \mathrm{d}t + C_\varepsilon \int_0^\xi \|\mathbf{H}_n(t) - \mathbf{H}(t)\|^2 \ \mathrm{d}s + C_\varepsilon \tau^2.$$

It remains to estimate the second term on the RHS in (6.13). Employing Lemma 6.2.2 (iii), we obtain

$$\left|\int_0^\xi \left(\partial_t \mathbf{H}_n - \partial_t \mathbf{H}, \mathbf{H}_n - \overline{\mathbf{H}}_n\right)\right|$$
$$\leqslant \sqrt{\int_0^\xi \|\partial_t \mathbf{H}_n - \partial_t \mathbf{H}\|^2} \sqrt{\int_0^\xi \left\|\mathbf{H}_n - \overline{\mathbf{H}}_n\right\|^2} \lesssim \tau.$$

Putting things together, choosing a sufficiently small positive ε and applying Grönwall's argument, we conclude the proof. $\square$

6.3 Conclusion

In this chapter, a vectorial nonlocal nonlinear parabolic problem (6.1) in terms of the magnetic field for an intermediate state between type-I and type-II superconductivity has been analysed. This model was obtained from the eddy current version of the Maxwell equations, the two-fluid model of London and London, the nonlocal representation (by a space convolution with a singular kernel) of the superconductive current by Eringen and the power law by Rhyner. A semi-implicit time-discrete scheme based on the backward Euler method in which the convolution is taken explicitly has been developed. The well-posedness of the problem has been shown under low regularity assumptions and suboptimal error estimates have been derived for the time-discretization.

Part II

Inverse source problems in thermoelasticity

7 Introduction on thermoelasticity

Thermoelasticity is the change in the size and shape of a solid object (thermal stresses) as the temperature of that object fluctuates. A material that is elastic expands when heated and contracts when cooled. These interactions between the changes in the shape of an object and the fluctuations in the temperature are modeled by mathematical systems. These so-called thermoelastic systems consist of two equations that are coupled: a parabolic (heat) equation and a vectorial hyperbolic equation for the displacement.

The problem of transmission of heat flow in rigid or elastic materials has attracted considerable attention in the past decades. Coupled thermomechanical problems arise from many important fields of application including casting, metal forming, manufacturing processes, structural models, etc.

Biot [140–142] started laying variational principles for coupled problems of thermoelasticity. A large number of papers follows Biot's principle, cf. [46, 143] and the references therein. By use of Gurtin's method of convolution [144] other variational principles have been formulated [145]. Development of new models in thermoelasticity has generated a response in applied mathematics. Many papers have been devoted to theoretical and numerical analyses of such problems, e.g [146–152].

Green and Naghdi [153] used a general entropy balance to describe the heat flow in materials. The characterization of material response for such thermal phenomena is labeled as type-I, type-II and type-III thermoelasticity. Type-I, after linearisation of the theory, is the same as the classical heat conduction theory (based on Fourier's law). This theory has the shortcoming that a thermal disturbance at one point of the body is instantly felt everywhere (infinite speed of propagation phenomena). This is physically not acceptable for materials with memory and is overcome by taking memory effects into account in the models for type-II and type-III thermoelasticity. Therefore, the type-II and -III thermoelasticity allow propagation of thermoelastic disturbances with a finite speed. The main difference between type-II and type-III is that in type-II thermoelasticity the heat conduction is independent of the present values of the temperature gradient. For forward problems related with the theoretical and computational aspects of thermoelasticity, it is worth to refer to [154–158].

For the mathematical analysis, it is assumed that an isotropic and homogeneous thermoelastic body occupies an open and bounded domain $\Omega \subset \mathbb{R}^d, d \geqslant 1$, with Lipschitz continuous boundary Γ. Let $Q_T = \Omega \times (0,T)$ and $\Sigma_T = \Gamma \times (0,T)$ for a given final time $T > 0$. The convolution product in time of a kernel k and a function θ is denoted with the sign '$*$'

$$(k * \theta)(\mathbf{x},t) := \int_0^t k(t-s)\theta(\mathbf{x},s)\mathrm{d}s, \qquad (\mathbf{x},t) \in Q_T.$$

The coupled thermoelastic system of type-III describing both the elastic and the thermal behaviours in Ω is given by [159]

$$\begin{cases} \partial_{tt}\mathbf{u} - \alpha\Delta\mathbf{u} - \beta\nabla\left(\nabla\cdot\mathbf{u}\right) + \gamma\nabla\theta &= \mathbf{f} \qquad \text{in } Q_T, \\ \partial_t\theta - \rho\Delta\theta - k*\Delta\theta + \gamma\nabla\cdot\partial_t\mathbf{u} &= h \qquad \text{in } Q_T. \end{cases} \tag{7.1}$$

Here, $\mathbf{u} = (u_1,\ldots,u_d)^\mathsf{T}$ and θ denote respectively the displacement and the temperature difference from the reference value (in Kelvin) of the solid elastic material at the location $\mathbf{x}$ and time t. The vector source $\mathbf{f}$ is a load (body force) vector and the source h is a heat source. The Lamé parameters α and β, the coupling (absorbing) coefficient γ and the thermal coefficient ρ are assumed to be positive constants because the medium is supposed to be isotropic homogeneous. Note that the coefficient γ in the first and second equation of (7.1) is in general different. The kernel function (also called relaxation function) $k \in \mathrm{C}([0,T])$ is supposed to decay to zero as the time goes to infinity. Usually k takes the form [157]

$$k(t) = a\exp(-bt), \qquad t > 0,$$

with a and b two positive constants. In type-I thermoelasticity $k \equiv 0$ and $\rho \neq 0$. For type-II thermoelasticity, it holds that $\rho = 0$ and $k \not\equiv 0$.

In the following two chapters, two inverse source problems (ISPs) for thermoelasticity are studied. In Chapter 8, the goal is to determine the vector source $\mathbf{f}(\mathbf{x})$ from a final in time measurement of the displacement. In Chapter 9, the reconstruction of a solely time-dependent heat source $h(t)$ is studied in a one-dimensional thermoelastic system of type-III.

Inverse problems are often ill-posed in the sense of Hadamard [52, 53, 160]. They might not have a solution in the strict sense, solutions might not be unique and/or might not depend continuously on the data. This creates (mostly because of the discontinuous dependence of solutions on the data) numerical problems. A typical example of an ill-posed problem is the operator equation

$$F(x) = y, \tag{7.2}$$

where $F : D(F) \subset X \to Y$ is a completely continuous operator between the Banach spaces X and Y. In the case of an inverse problem the operator F is associated with the forward (direct) problem, which relates the model parameters to the measured data. The direct problem is well-posed in the sense of Hadamard. If the problem (7.2) is well-posed, then the operator $F : D(F) \to R(F)$ is injective and the inverse operator $F^{-1} : R(F) \to D(F)$ is continuous. It follows that the identity operator $I = F^{-1}F$ is compact (the composition of a continuous and compact operator is compact). This is in contradiction with the fundamental Riesz' theorem if the domain $D(F)$ is not finite dimensional. Therefore, the problem (7.2) is ill-posed when the domain $D(F)$ is not finite dimensional.

Regularization methods can deal with the ill-posedness of linear and nonlinear inverse problems. In general terms, the idea is to approximate the ill-posed problem (7.2) by a family of neighbouring well-posed problems [52]. The main goal is to find the best approximate solution for the problem (7.2), where one assumes that only the noisy data y^δ of the exact data y are available, i.e.

$$\left\| y - y^\delta \right\| \leqslant \delta$$

with δ being the noise level in some norm.

The Tikhonov regularization method is the most commonly used method of regularization of ill-posed problems [161, 162]. The best approximate solution for (7.2) is searched by minimizing a certain Tikhonov functional

$$\mathcal{T}_\alpha(x) = \left\| F(x) - y^\delta \right\|_Y^2 + \alpha \left\| x \right\|_X^2, \qquad \alpha > 0.$$

The regularization (stabilizing) term $\|x\|^2$ is added to rectify the possible absence of convexity. This term introduces the a priori knowledge about the solution. For $\alpha > 0$, the functional $\mathcal{T}_\alpha$ has a unique minimizer ($\mathcal{T}_\alpha$ is convex) when F is linear [52, Theorem 5.1]. If F is nonlinear, then the solution is in general not unique. The regularization parameter α can be chosen by using Morozov's discrepancy principle [163], which basically compares the residual (or discrepancy) error $\left\| F(x_\alpha^\delta) - y^\delta \right\|_Y$ for the solution x_α^δ of the minimization problem (7.2) with the noise level δ, i.e. determining $\alpha(\delta, x_\alpha^\delta)$ from the condition

$$\left\| F(x_\alpha^\delta) - y^\delta \right\|_Y \approx \delta. \tag{7.3}$$

For nonlinear problems, (7.3) only has a solution under very restrictive assumptions on the regularized solutions [164]. Note that the problem of minimizing the functional $\mathcal{T}_\alpha$ is stable in the sense of continuous dependence of the solutions on the data y^δ [52, Theorem 10.2].

To numerically find the minimizer x_α^δ of $\mathcal{T}_\alpha$, gradient-based (steepest descent) methods are often used. The approximative sequence $\{x_k\}$ for x_α^δ is constructed as follows

$$x_k = x_{k-1} - \omega \mathcal{T}_\alpha'(x_{k-1}), \qquad k \in \mathbb{N}, \quad x_0 \in X (\text{initial guess}),$$

where $\mathcal{T}_\alpha'$ is the Fréchet derivative of $\mathcal{T}_\alpha$ and ω is a suitable step length.

Recently, inverse source problems related to the classic thermoelastic system (7.1) have been studied in [165, 166]. Without taking memory effects into account, i.e. $k \equiv 0$ and $\rho \neq 0$, Bellassoued and Yamamoto [165] investigated an inverse heat source problem for type-I thermoelasticity. The main subject of the paper is the inverse problem of determining the space-dependent heat source $h(\mathbf{x})$. This is done by measuring $\mathbf{u}_{|\omega\times(0,T)}$ and $\theta(\cdot, t_0)$, where ω is a subdomain of Ω such that $\Gamma \subset \partial\omega$ and $t_0 \in (0,T)$. No data for $\mathbf{u}(\cdot, t_0)$ is needed over the whole domain Ω. Using a Carleman estimate, a Hölder stability for the inverse source problem is proved, which implies the uniqueness of the inverse source problem. Wu and Liu [166] studied an inverse source problem of determining $\mathbf{f}(\mathbf{x})$ for type-II thermoelasticity, i.e. $k \not\equiv 0$ and $\rho = 0$. Based on a Carleman estimate, again a Hölder stability for the inverse source problem has been established from a displacement measurement $\mathbf{u}_{|\omega\times(0,T)}$ provided that $\mathbf{f}$ is known in a neighbourhood ω_0 of Γ. Note that no temperature measurement is needed. Note also [167], in which Wu et al. studied the uniqueness and stability of a spatially varying thermal kernel function in a thermoelastic system of type-III by means of a Carleman estimate.

In both contributions [165,166], no numerical scheme has been provided to recover the unknown source. This is in contrast to the main goal of the following chapters, namely, the development of a scheme to approximate the solution to the inverse problems.

8

Recovery of a space-dependent vector source in thermoelastic systems

This chapter is based on the article [168], which is published in *Inverse Problems in Science and Engineering*.

In this chapter, an inverse problem of determining a space-dependent source in a thermoelastic system of type-III using information from a supplementary measurement at a given single instant of time is studied. The mathematical setting is the following.

An isotropic and homogeneous thermoelastic body occupying an open and bounded domain $\Omega \subset \mathbb{R}^d, d \geqslant 1$, with Lipschitz continuous boundary Γ is considered. Let $Q_T = \Omega \times (0,T)$ and $\Sigma_T = \Gamma \times (0,T)$ for a given final time $T > 0$. The convolution product in time of a kernel k and a function θ is denoted by

$$(k * \theta)(\mathbf{x}, t) := \int_0^t k(t-s)\theta(\mathbf{x}, s) \, \mathrm{d}s, \qquad (\mathbf{x}, t) \in Q_T.$$

The following thermoelastic system of type-III describing the elastic and thermal behaviour in Ω is considered:

$$\left\{ \begin{array}{rll} \partial_{tt}\mathbf{u} + \mathbf{g}\left(\partial_t \mathbf{u}\right) - \alpha\Delta\mathbf{u} - \beta\nabla\left(\nabla\cdot\mathbf{u}\right) + \gamma\nabla\theta & = \mathbf{f} & \text{in } Q_T, \\ \partial_t\theta - \rho\Delta\theta - k * \Delta\theta + \gamma\nabla\cdot\partial_t\mathbf{u} & = h & \text{in } Q_T, \\ \mathbf{u}(\mathbf{x}, t) & = \mathbf{0} & \text{on } \Sigma_T, \\ \theta(\mathbf{x}, t) & = 0 & \text{on } \Sigma_T, \end{array} \right. \tag{8.1}$$

together with the initial conditions:

$$\mathbf{u}(\mathbf{x},0)=\overline{\mathbf{u}}_0(\mathbf{x}),\quad \partial_t\mathbf{u}(\mathbf{x},0)=\overline{\mathbf{u}}_1(\mathbf{x}),\quad \theta(\mathbf{x},0)=\overline{\theta}_0(\mathbf{x}),\quad \mathbf{x}\in\Omega, \tag{8.2}$$

where it is assumed that the unknown vector source $\mathbf{f}$ is of the form

$$\mathbf{f}(\mathbf{x},t)=\mathbf{p}(\mathbf{x})+\mathbf{r}(\mathbf{x},t),\qquad (\mathbf{x},t)\in Q_T, \tag{8.3}$$

with the vector field $\mathbf{r}$ known and $\mathbf{p}$ unknown. The heat source h is known. Note that $\mathbf{u}=(u_1,\ldots,u_d)^{\mathsf{T}}$ and θ denote respectively the displacement and the temperature difference from the reference value (in Kelvin) of the solid elastic material at the location $\mathbf{x}$ and time t. The Lamé parameters α and β, the coupling (absorbing) coefficient γ and the thermal coefficient ρ are assumed to be positive constants because the medium is supposed to be isotropic homogeneous.

The goal is to determine the spatial vector function $\mathbf{p}(\mathbf{x})$ with aid of an additional measurement (the condition of final overdetermination), i.e.

$$\mathbf{u}_T(\mathbf{x}):=\mathbf{u}(\mathbf{x},T)=\boldsymbol{\xi}_T(\mathbf{x}),\qquad \mathbf{x}\in\Omega. \tag{8.4}$$

This means that the displacement is measured at the final time.

The kernel function (also called relaxation function) $k\in \mathrm{C}^2([0,T])$ is decaying to zero as the time goes to infinity. Moreover, it is assumed that

$$k'(t)\not\equiv 0\quad \text{ and }\quad (-1)^j k^{(j)}(t)\geqslant 0,$$

with $j=0,1,2$ denoting the order of the derivative. These assumptions imply that k is strongly positive definite [169, Corollary 7.2.1], which is equivalent with the existence of a positive constant C_0 independent of T such that [169, Lemma 7.2.2]- [170]

$$\int_0^T \phi(t)(k*\phi)(t)\,\mathrm{d}t\geqslant C_0\int_0^T (k*\phi)^2(t)\,\mathrm{d}t,\qquad \forall T>0,\forall\phi\in \mathrm{L}^1_{\mathrm{loc}}(\Omega).$$

These assumptions on the kernel k are natural, because usually k takes the form

$$k(t)=a\exp(-bt),\qquad t>0,$$

with a and b two positive constants [157]. Note that a damping term

$$\mathbf{g}\left(\partial_t\mathbf{u}\right)=\left(g_1(\partial_t\mathbf{u}),\ldots,g_d(\partial_t\mathbf{u})\right),$$

$g_i:\mathbb{R}^d\to\mathbb{R}$, $i=1,\ldots,d$, is added in the hyperbolic equation of the classic thermoelasticity system (7.1). This term is also considered in [169, Chapter 9]- [171, 172] and is essential to establish the uniqueness of a solution to the inverse problem under consideration, see Theorem 8.1.1.

The remainder of this chapter is organized as follows. The uniqueness of a solution to the inverse problem under consideration is established in Section 8.1 under the assumption that the possibly nonlinear function $\mathbf{g}$ is componentwise strictly monotone increasing. This is done using a variational approach instead of using a Carleman estimate like in [165–167]. The inverse problem (8.1)-(8.4) is ill-posed since small errors present in any practical measurements give rise to large errors into the solutions. In Section 8.2, an iterative regularization method in the form of a convergent and stable algorithm for the recovery of the unknown vector source is proposed in the case that $\mathbf{g}$ is linear. This method is based on a sequence of well-posed direct problems, which are numerically solved at each iteration step by using the finite element method. The instability of this inverse source problem is overcome by stopping the iterations using the discrepancy principle [163]. The scheme is of the Landweber-Fridman type [173, 174] and is similar to that of Johansson and Lesnic for the heat conduction equation [175]. This procedure is also used for the heat conduction equation with time-dependent coefficients in [176]. Note that the recovery of the unknown source is not achieved by minimizing a cost functional, which is typical for IPs. Finally, some numerical experiments are developed in Section 8.3.

8.1 Uniqueness

Using Green's formulas, the following coupled variational formulation for (8.1) is obtained:

Given $\overline{\mathbf{u}}_0, \overline{\mathbf{u}}_1, \boldsymbol{\xi}_T \in \mathbf{L}^2(\Omega)$, $\overline{\theta}_0 \in \mathrm{L}^2(\Omega)$, $\mathbf{r} \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$ and $h \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$,

find $\langle \mathbf{u}(t), \theta(t), \mathbf{p}\rangle \in \mathbf{H}_0^1(\Omega) \times \mathrm{H}_0^1(\Omega) \times \mathbf{L}^2(\Omega)$

such that

$$
\begin{aligned}
&(\partial_{tt}\mathbf{u}(t), \boldsymbol{\varphi}) + (\mathbf{g}(\partial_t \mathbf{u}(t)), \boldsymbol{\varphi}) + \alpha(\nabla \mathbf{u}(t), \nabla \boldsymbol{\varphi}) \\
&\qquad + \beta(\nabla \cdot \mathbf{u}(t), \nabla \cdot \boldsymbol{\varphi}) + \gamma(\nabla \theta(t), \boldsymbol{\varphi}) = (\mathbf{p} + \mathbf{r}(t), \boldsymbol{\varphi})
\end{aligned} \tag{8.5}
$$

and

$$
\begin{aligned}
&(\partial_t \theta(t), \psi) + \rho(\nabla \theta(t), \nabla \psi) + ((k * \nabla \theta)(t), \nabla \psi) \\
&\qquad - \gamma(\partial_t \mathbf{u}(t), \nabla \psi) = (h(t), \psi),
\end{aligned} \tag{8.6}
$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$, $\psi \in \mathrm{H}_0^1(\Omega)$ and a.a. $t \in (0,T]$.

The space-dependent measurement (8.4) ensures that the inverse problem has a unique solution. This is stated in the following theorem.

Theorem 8.1.1 (Uniqueness). *Let $\overline{\mathbf{u}}_0, \overline{\mathbf{u}}_1, \boldsymbol{\xi}_T \in \mathbf{L}^2(\Omega)$, $\overline{\theta}_0 \in \mathrm{L}^2(\Omega)$ and $\mathbf{g}' > \mathbf{0}$ componentwise. Then there exists at most one triplet $\langle \mathbf{u}(t), \theta(t), \mathbf{p}\rangle \in \mathbf{H}_0^1(\Omega) \times \mathrm{H}_0^1(\Omega) \times \mathbf{L}^2(\Omega)$ such that problem* (8.1) *together with condition* (8.4) *holds.*

Proof. We use a classical variational approach to establish the uniqueness of a solution. The proof is by contradiction. Suppose that there are two solutions $\langle \mathbf{u}_1, \theta_1, \mathbf{p}_1 \rangle$ and $\langle \mathbf{u}_2, \theta_2, \mathbf{p}_2 \rangle$ to (8.1)-(8.4). Subtract, equation by equation, the variational formulation (8.5)-(8.6) corresponding with the solution $\langle \mathbf{u}_2, \theta_2, \mathbf{p}_2 \rangle$ from the variational formulation for $\langle \mathbf{u}_1, \theta_1, \mathbf{p}_1 \rangle$. Set $\mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2$, $\mathbf{p} = \mathbf{p}_1 - \mathbf{p}_2$ and $\theta = \theta_1 - \theta_2$. Then $\mathbf{u}(\mathbf{x}, 0) = \mathbf{0}$, $\mathbf{u}(\mathbf{x}, T) = \mathbf{0}$, $\partial_t \mathbf{u}(\mathbf{x}, 0) = \mathbf{0}$ and $\theta(\mathbf{x}, 0) = 0$. We obtain that

$$(\partial_{tt}\mathbf{u}(t), \boldsymbol{\varphi}) + (\mathbf{g}(\partial_t \mathbf{u_1}(t)) - \mathbf{g}(\partial_t \mathbf{u_2}(t)), \boldsymbol{\varphi}) + \alpha (\nabla \mathbf{u}(t), \nabla \boldsymbol{\varphi}) \\ + \beta (\nabla \cdot \mathbf{u}(t), \nabla \cdot \boldsymbol{\varphi}) + \gamma (\nabla \theta(t), \boldsymbol{\varphi}) = (\mathbf{p}, \boldsymbol{\varphi}) \quad (8.7)$$

and

$$(\partial_t \theta(t), \psi) + \rho (\nabla \theta(t), \nabla \psi) \\ + ((k * \nabla \theta)(t), \nabla \psi) - \gamma (\partial_t \mathbf{u}(t), \nabla \psi) = 0, \quad (8.8)$$

for all $\boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega)$ and $\psi \in \mathrm{H}_0^1(\Omega)$. First, we prove that $\mathbf{u} = \mathbf{0}$ and $\theta = 0$. Afterwards, we show that $\mathbf{p} = \mathbf{0}$. For this reason, in the first part of the proof, we want to get rid of $\mathbf{p}$. This can be done in a simple way. The main idea is

$$\int_0^T \mathbf{p}(\mathbf{x}) \cdot \partial_t \mathbf{u}(\mathbf{x}, t) \, \mathrm{d}t = \mathbf{p}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}, T) - \mathbf{p}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}, 0) = 0.$$

Indeed, putting $\boldsymbol{\varphi} = \partial_t \mathbf{u}(t)$ in (8.7) and integrating in time over $(0, T)$ gives

$$\frac{1}{2} \|\partial_t \mathbf{u}(T)\|^2 + \int_0^T (\mathbf{g}(\partial_t \mathbf{u_1}) - \mathbf{g}(\partial_t \mathbf{u_2}), \partial_t \mathbf{u}_1 - \partial_t \mathbf{u}_2) \\ + \gamma \int_0^T (\nabla \theta, \partial_t \mathbf{u}) = 0, \quad (8.9)$$

because $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}(\mathbf{x}, T) = \partial_t \mathbf{u}(\mathbf{x}, 0) = \mathbf{0}$. Taking $\psi = \theta(t)$ in (8.8) yields

$$\frac{\|\theta(T)\|^2}{2} + \rho \int_0^T \|\nabla \theta\|^2 + \int_0^T (k * \nabla \theta, \nabla \theta) - \gamma \int_0^T (\partial_t \mathbf{u}, \nabla \theta) = 0, \quad (8.10)$$

due to $\theta(\mathbf{x}, 0) = 0$. Now, adding (8.9) and (8.10) implies

$$\frac{1}{2} \|\partial_t \mathbf{u}(T)\|^2 + \int_0^T (\mathbf{g}(\partial_t \mathbf{u_1}) - \mathbf{g}(\partial_t \mathbf{u_2}), \partial_t \mathbf{u}_1 - \partial_t \mathbf{u}_2) \\ + \frac{\|\theta(T)\|^2}{2} + \rho \int_0^T \|\nabla \theta\|^2 + \int_0^T (k * \nabla \theta, \nabla \theta) = 0. \quad (8.11)$$

Moreover, the strongly positive definiteness of k implies that

$$\int_0^T (k * \nabla \theta, \nabla \theta) \geqslant C_0 \int_0^T \|k * \nabla \theta\|^2 .$$

Thus, from (8.11) follows that

$$\|\partial_t \mathbf{u}(T)\|^2 + \int_0^T \left(\mathbf{g}\left(\partial_t \mathbf{u_1}\right) - \mathbf{g}\left(\partial_t \mathbf{u_2}\right), \partial_t \mathbf{u}_1 - \partial_t \mathbf{u}_2\right) = 0$$

and

$$\|\theta(T)\|^2 + \int_0^T \|\nabla\theta\|^2 + \int_0^T \|k * \nabla\theta\|^2 = 0. \tag{8.12}$$

Due to the fact that $\theta = 0$ on $\partial\Omega$, we deduce that

$$\theta = 0 \text{ a.e. in } Q_T.$$

Here, we can also see why the damping term is necessary. Without this term, we would only have that $\|\partial_t \mathbf{u}(T)\| = 0$, which gives no guarantee that $\mathbf{u} = \mathbf{0}$. Employing the fact that the vector field $\mathbf{g}$ is componentwise strictly monotone increasing, we get that $\mathbf{u}_t = \mathbf{0}$, i.e. $\mathbf{u}$ is constant in time. Therefore,

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{0} \Rightarrow \mathbf{u}(\mathbf{x}, t) = \mathbf{0} \text{ a.e. in } Q_T.$$

Substituting the obtained information in (8.7) gives

$$(\mathbf{p}, \boldsymbol{\varphi}) = 0, \qquad \forall \boldsymbol{\varphi} \in \mathbf{H}_0^1(\Omega).$$

From this, we conclude that $\mathbf{p} = \mathbf{0}$ in $\mathbf{L}^2(\Omega)$. □

Remark 8.1.1. *From the previous theorem also the uniqueness of a solution to the inverse problem corresponding with type-I thermoelasticity ($k = 0, \rho \neq 0$) follows. For type-II thermoelasticity ($\rho = 0, k \neq 0$) the proof of uniqueness of a solution is less straightforward. Then* (8.12) *becomes*

$$\|\theta(T)\|^2 + \int_0^T \|k * \nabla\theta\|^2 = 0.$$

Therefore, $\displaystyle\int_0^t k(t-s)\nabla\theta(\mathbf{x}, s)\ \mathrm{d}s = 0$ *for all* $t \in [0, T]$ *and* $\mathbf{x} \in \Omega$*. Hence, since the Laplace transform is one-to-one, one can derive that* $\nabla\theta = 0$ *in* Q_T*. The uniqueness of a solution follows from* $\theta = 0$ *on* $\partial\Omega$*.*

Remark 8.1.2. *In fact, to prove the uniqueness of a solution in the case of type-III thermoelasticity, it is sufficient that the kernel k is positive definite instead of strongly positive definite. But, the strongly positive definiteness of k is immediately considered because under this assumption the uniqueness of a solution to the inverse problem is valid for all types of thermoelasticity, as mentioned in Remark* 8.1.1.

Remark 8.1.3. *The main trick of the proof cannot be applied if the heat source $h(\mathbf{x})$ would be unknown, i.e.* $\int_0^T h(\mathbf{x})\theta(\mathbf{x}, t)\,\mathrm{d}t \neq 0$.

Remark 8.1.4. *The uniqueness of the solution can also be obtained for more general coefficients. For instance, for*

$$\begin{cases} \partial_{tt}\mathbf{u} + \mathbf{g}\left(\partial_t \mathbf{u}\right) - \nabla \cdot (\alpha(\mathbf{x})\nabla \mathbf{u}) - \nabla\left(\beta(\mathbf{x})\nabla \cdot \mathbf{u}\right) + \gamma\nabla\theta & = \mathbf{f} \qquad \text{in } Q_T, \\ \partial_t\theta - \nabla \cdot (\rho(\mathbf{x})\nabla\theta) - k * \Delta\theta + \gamma\nabla \cdot \partial_t\mathbf{u} & = h \qquad \text{in } Q_T, \end{cases}$$

the following assumptions have to be satisfied for a.a. $x \in \Omega$*:*

$$0 \leqslant \alpha(\mathbf{x}) \leqslant \alpha_1, \quad 0 \leqslant \beta(\mathbf{x}) \leqslant \beta_1 \quad \text{and} \quad 0 < \rho_0 \leqslant \rho(\mathbf{x}) \leqslant \rho_1.$$

8.2 Reconstruction of the source term in a linear case

First, the well-posedness of problem (8.1) is discussed (thus for given $\mathbf{p}$, i.e. take $\mathbf{p} = \mathbf{0}$ for ease of exposition). Rivera and Qin [159] proved the global existence and uniqueness of solutions to problem (8.15) in one dimension when $\mathbf{r} \equiv \mathbf{0} \equiv \mathbf{g}$ and $h = 0$. In the same situation, a more-dimensional case for type-III thermoelasticity is studied in [177]. The following lemma summarizes the available results. For a more general setting, see also [151, 152, 178]. The derivation of the corresponding a priori estimates can be found in Theorem A.1.4 in Appendix A.1 (such that Rothe's method can be applied to show the existence of a solution).

Lemma 8.2.1.

(i) Assume that $\mathbf{r} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$*,* $h \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$*,* $\overline{\mathbf{u}}_0 \in \mathbf{H}^1(\Omega)$*,* $\overline{\mathbf{u}}_1 \in \mathbf{L}^2(\Omega)$*,* $\overline{\theta}_0 \in \mathrm{L}^2(\Omega)$*,* $\mathbf{g}(\mathbf{0}) = \mathbf{0}$*,* $\mathbf{g}' > \mathbf{0}$ *and* $|\mathbf{g}(s)| \leqslant C(1 + |s|)$ *a.e. in* $\mathbb{R}$*. Then* (8.15)*, has a unique solution* $\langle \mathbf{u}, \theta\rangle$ *such that*

$$\begin{aligned} \mathbf{u} &\in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right), \\ \partial_t \mathbf{u} &\in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right), \partial_{tt}\mathbf{u} \in \mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)^*\right), \\ \theta &\in \mathrm{C}\left([0,T], \mathrm{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)\right), \\ \partial_t\theta &\in \mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)^*\right). \end{aligned}$$

(ii) Assume that $\mathbf{r}(0) \in \mathbf{L}^2(\Omega)$*,* $h(0) \in \mathrm{L}^2(\Omega)$*,* $\partial_t\mathbf{r} \in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right)$*,* $\partial_t h \in \mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$*,* $\overline{\mathbf{u}}_0 \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$*,* $\overline{\mathbf{u}}_1 \in \mathbf{H}^1(\Omega)$*,* $\overline{\theta}_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega)$*,* $\mathbf{g}(\mathbf{0}) = \mathbf{0}$ *and* $\mathbf{0} < \mathbf{g}'(s) \leqslant \mathbf{C}$ *a.e. in* $\mathbb{R}$*. Then* (8.15)*, has a unique solution* $\langle \mathbf{u}, \theta\rangle$ *such that*

$$\begin{aligned} \mathbf{u} &\in \mathrm{C}\left([0,T], \mathbf{H}_0^1(\Omega)\right), \\ \partial_t\mathbf{u} &\in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right), \\ \partial_{tt}\mathbf{u} &\in \mathrm{L}^2\left((0,T), \mathbf{L}^2(\Omega)\right), \\ \theta &\in \mathrm{C}\left([0,T], \mathrm{H}_0^1(\Omega)\right), \partial_t\theta \in \mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)\right). \end{aligned}$$

In the special situation that $\overline{\mathbf{u}}_0 = \mathbf{0}$, $\overline{\mathbf{u}}_1 = \mathbf{0}$, $\overline{\theta}_0 = 0$, $h = 0$ *and* $\mathbf{r} = \mathbf{r}(\mathbf{x})$, *the following estimate is valid*

$$\max_{t\in[0,T]} \left\{ \|\partial_{tt}\mathbf{u}(t)\|^2 + \|\nabla\partial_t\mathbf{u}(t)\|^2 + \|\nabla\cdot\partial_t\mathbf{u}(t)\|^2 \right.$$
$$\left. + \|\partial_t\theta(t)\|^2 + \|\nabla\theta(t)\|^2 \right\} + \int_0^T \|\nabla\partial_t\theta(s)\|^2 \, \mathrm{d}s \leqslant C\|\mathbf{r}\|^2 . \quad (8.13)$$

The unknown source can be reconstructed if the assumptions on the function $\mathbf{g}$ are strengthened. From now on, it is assumed that $\mathbf{g}$ is linear (without loss of generality suppose that $\mathbf{g} = \mathbf{I}$) such that the principle of linear superposition is applicable on problem (8.1)-(8.2)-(8.3)-(8.4). The situation is much more difficult when $\mathbf{g}$ is nonlinear. The solution $\langle \mathbf{u}, \theta, \mathbf{p}\rangle$ is given by $\langle \mathbf{u}_1+\mathbf{u}_2, \theta_1+\theta_2, \mathbf{p}\rangle$, where $\langle \mathbf{u}_1, \theta_1, \mathbf{p}\rangle$ is a solution to

$$\begin{cases} \partial_{tt}\mathbf{u} + \partial_t\mathbf{u} - \alpha\Delta\mathbf{u} - \beta\nabla(\nabla\cdot\mathbf{u}) + \gamma\nabla\theta &= \mathbf{p} \quad \text{in } Q_T, \\ \partial_t\theta - \rho\Delta\theta - k*\Delta\theta + \gamma\nabla\cdot\partial_t\mathbf{u} &= 0 \quad \text{in } Q_T, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \Sigma_T, \\ \theta &= 0 \quad \text{on } \Sigma_T, \\ \mathbf{u}(\mathbf{x},0) = \partial_t\mathbf{u}(\mathbf{x},0) = \mathbf{0}, \quad \theta(\mathbf{x},0) &= 0 \quad \mathbf{x}\in\Omega \end{cases} \quad (8.14)$$

and $\langle \mathbf{u}_2, \theta_2\rangle$ is solving

$$\begin{cases} \partial_{tt}\mathbf{u} + \partial_t\mathbf{u} - \alpha\Delta\mathbf{u} - \beta\nabla(\nabla\cdot\mathbf{u}) + \gamma\nabla\theta &= \mathbf{r} \quad \text{in } Q_T, \\ \partial_t\theta - \rho\Delta\theta - k*\Delta\theta + \gamma\nabla\cdot\partial_t\mathbf{u} &= h \quad \text{in } Q_T, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \Sigma_T, \\ \theta &= 0 \quad \text{on } \Sigma_T, \\ \mathbf{u}(\mathbf{x},0) = \overline{\mathbf{u}}_0(\mathbf{x}), \quad \partial_t\mathbf{u}(\mathbf{x},0) = \overline{\mathbf{u}}_1(\mathbf{x}), \quad \theta(\mathbf{x},0) &= \overline{\theta}_0(\mathbf{x}) \quad \mathbf{x}\in\Omega. \end{cases} \quad (8.15)$$

In the remainder of the chapter, next to the linearity of $\mathbf{g}$, it is assumed that the assumptions of Lemma 8.2.1(ii) are valid. This is important because in this situation the boundary conditions are satisfied since $\mathbf{u}(\cdot,t) \in \mathbf{H}_0^1(\Omega)$ and $\theta(\cdot,t) \in \mathrm{H}_0^1(\Omega)$ for $t \in [0,T]$. Moreover, also the restriction $\mathbf{u}(\mathbf{x},t_0)$ is well-defined for $t_0 \in [0,T]$. This means in particular that the final displacement measurement $\mathbf{u}(\mathbf{x},T) \in \mathbf{H}_0^1(\Omega)$ is well-defined. Following Theorem 8.1.1 and Lemma 8.2.1(ii), the solution $\langle \mathbf{u}_1, \theta_1, \mathbf{p}\rangle$ to problem (8.14) is unique if the additional final measurement is satisfied, i.e.

$$\mathbf{u}_1(\mathbf{x},T) = \boldsymbol{\xi}_T(\mathbf{x}) - \mathbf{u}_2(\mathbf{x},T) =: \widetilde{\boldsymbol{\xi}}_T(\mathbf{x}), \qquad \mathbf{x}\in\Omega, \quad (8.16)$$

where $\mathbf{u}_2$ is the solution to problem (8.15). Note that for given $\mathbf{p}$, problem (8.14) is a special case of problem (8.15). In the following subsection, an algorithm for the recovery of the unknown source term is proposed.

Remark 8.2.1. *The well-posedness of problem* (8.15) *can be obtained for more general coefficients. In particular, when the coefficients are space-dependent, the following assumptions have to be satisfied*

$$0 < \alpha_0 \leqslant \alpha(\mathbf{x}) \leqslant \alpha_1, \qquad 0 < \beta_0 \leqslant \beta(\mathbf{x}) \leqslant \beta_1,$$
$$0 < \rho_0 \leqslant \rho(\mathbf{x}) \leqslant \rho_1, \qquad 0 < \gamma_0 \leqslant \gamma(\mathbf{x}) \leqslant \gamma_1.$$

8.2.1 Algorithm for finding the source term

In this section, an algorithm for finding the source term is described. This algorithm is based on an iterative regularization method (Landweber-Fridman iteration) instead of using the Tikhonov regularization.

Let $\langle \mathbf{v}, \zeta \rangle$ be the unique solution to (8.14) for given $\mathbf{p}$, see Lemma 8.2.1. Define the corresponding operator $M(t) \in \mathcal{L}\left(\mathbf{L}^2(\Omega), \mathbf{L}^2(\Omega)\right)$ by

$$M(t)\mathbf{p} = \mathbf{v}(\cdot, t).$$

Finding a solution to the inverse problem is then equivalent to solving the following operator equation

$$M(T)\mathbf{p} = \widetilde{\boldsymbol{\xi}}_T \tag{8.17}$$

or equivalent to solving the fixed point equation

$$\mathbf{p} = \mathbf{p} + \kappa M(T)\left(\widetilde{\boldsymbol{\xi}}_T - M(T)\mathbf{p}\right), \qquad \kappa > 0,$$

due to the linearity of the operator $M(T)$. The parameter κ is called a relaxation parameter. The method of successive approximations can be applied to this latter equation as follows

$$\mathbf{p}_k := \mathbf{p}_{k-1} - \kappa M(T)\left(M(T)\mathbf{p}_{k-1} - \widetilde{\boldsymbol{\xi}}_T\right), \qquad k \in \mathbb{N},$$

with an initial guess $\mathbf{p}_0$, which plays the same role as in the Tikhonov regularization.

This gives rise to the following procedure for the stable reconstruction of the solution $\langle \mathbf{u}, \theta \rangle$ and the source term $\mathbf{p}$ of problem (8.1)-(8.2)-(8.3)-(8.4), which is similar to the one presented in [175, 176, 179]. It runs as follows:

(i) Solve problem (8.15) and determine the transformed final overdetermination $\widetilde{\boldsymbol{\xi}}_T(\mathbf{x})$, see equation (8.16). Denote the solution by $\langle \mathbf{u}_*, \theta_* \rangle$;

(ii) Choose an initial guess $\mathbf{p}_0 \in \mathbf{L}^2(\Omega)$. Let $\langle \mathbf{v}_0, \zeta_0 \rangle$ be the solution to (8.14) with $\mathbf{p} = \mathbf{p}_0$;

(iii) Assume that $\mathbf{p}_k$ and $\langle \mathbf{v}_k, \zeta_k \rangle$ have been constructed. Let $\langle \mathbf{w}_k, \eta_k \rangle$ solve (8.14) with $\mathbf{p}(\mathbf{x}) = \mathbf{v}_k(\mathbf{x}, T) - \widetilde{\boldsymbol{\xi}}_T(\mathbf{x})$;

(iv) Define
$$\mathbf{p}_{k+1}(\mathbf{x}) = \mathbf{p}_k(\mathbf{x}) - \kappa \mathbf{w}_k(\mathbf{x}, T), \quad \mathbf{x} \in \Omega,$$
where $\kappa > 0$, and let $\langle \mathbf{v}_{k+1}, \zeta_{k+1} \rangle$ solve (8.14) with $\mathbf{p} = \mathbf{p}_{k+1}$;

(v) Repeate steps (ii) and (iii) until a desired level of accuracy is achieved, see Subsection 8.2.2. Suppose that the algorithm stopped after $\tilde{k}$ iterations. Denote the corresponding solution by $\langle \mathbf{v}_{\tilde{k}}, \zeta_{\tilde{k}}, \mathbf{p}_{\tilde{k}} \rangle$. Then, the approximating solution to the original problem (8.1)-(8.2)-(8.3)-(8.4) is given by $\langle \mathbf{u}_* + \mathbf{v}_{\tilde{k}}, \theta_* + \zeta_{\tilde{k}}, \mathbf{p}_{\tilde{k}} \rangle$.

The problems used in this iterative procedure are well-posed, see Lemma 8.2.1. Moreover, the restrictions of the solutions are well-defined. The following theorem shows the convergence of the proposed algorithm.

Theorem 8.2.1 (Existence)**.** *Assume that the assumptions of Lemma 8.2.1(ii) are satisfied and suppose that the relaxation parameter κ satisfies $0 < \kappa < \|M(T)\|^{-2}$. Denote by $\langle \mathbf{u}, \theta, \mathbf{p}\rangle = \langle \mathbf{u}_*+\mathbf{v}, \theta_*+\zeta, \mathbf{p}\rangle$ the unique solution to the original inverse problem* (8.1)-(8.2)-(8.3)-(8.4)*, where $\langle \mathbf{u}_*, \theta_*\rangle$ is the solution to problem* (8.15) *and $\langle \mathbf{v}, \zeta, \mathbf{p}\rangle$ is solving* (8.14)-(8.16)*. Let $\langle \mathbf{v}_k, \zeta_k, \mathbf{p}_k\rangle$ be the k-th approximation in the iterative algorithm of Subsection 8.2.1. Then*

$$\lim_{k\to\infty} \left\{ \|\mathbf{v} - \mathbf{v}_k\|_{\mathrm{C}\left([0,T],\mathbf{H}_0^1(\Omega)\right)} + \|\zeta - \zeta_k\|_{\mathrm{C}\left([0,T],\mathrm{H}_0^1(\Omega)\right)} \right\} = 0$$

and

$$\lim_{k\to\infty} \left\{ \|\partial_t\mathbf{v} - \partial_t\mathbf{v}_k\|_{\mathrm{L}^2\left((0,T),\mathbf{H}_0^1(\Omega)\right)} + \|\partial_t\zeta - \partial_t\zeta_k\|_{\mathrm{L}^2\left((0,T),\mathrm{H}_0^1(\Omega)\right)} \right\} = 0$$

for every function $\mathbf{p}_0 \in \mathbf{L}^2(\Omega)$.

Proof. From the iterative algorithm and the linearity of the operator $M(t)$, it is possible to deduce that

$$\begin{aligned}
\mathbf{p}_{k+1} &= \mathbf{p}_k - \kappa \mathbf{w}_k(\cdot, T) \\
&= \mathbf{p}_k - \kappa M(T)\left(\mathbf{v}_k(\cdot, T) - \widetilde{\boldsymbol{\xi}}_T\right) \\
&= \mathbf{p}_k - \kappa M(T)\left(M(T)\mathbf{p}_k - M(T)\mathbf{p}\right) \\
&= \mathbf{p}_k - \kappa M(T)M(T)\left(\mathbf{p}_k - \mathbf{p}\right).
\end{aligned}$$

Therefore,

$$\mathbf{p}_{k+1} - \mathbf{p} = \left(I - \kappa M(T)M(T)\right)\left(\mathbf{p}_k - \mathbf{p}\right).$$

This is a Landweber-Fridman iteration scheme for solving the operator equation (8.17). The standard proof of convergence for Landweber's iterations is given for $T_2 = T_1^*$ and $T_1 \in \mathcal{L}(X, Y)$ with X and Y abstract Hilbert spaces in [52, Theorem 6.1]. This theorem is based on a more general version of the proof of convergence for two not self-adjoint operators $T_1, T_2 \in \mathcal{L}(X, Y)$, which is given in [180, Theorem 3]. This implies thanks to the assumption $0 < \kappa < \|M(T)\|^{-2}$ that the sequence $\mathbf{p}_k$ converges to $\mathbf{p}$ in $\mathbf{L}^2(\Omega)$ for an arbitrary $\mathbf{p}_0 \in \mathbf{L}^2(\Omega)$. Inequality (8.13) implies that $\mathbf{v}_k \to \mathbf{v}$ and $\zeta_k \to \zeta$ in $\mathrm{C}\left([0,T], \mathbf{H}_0^1(\Omega)\right)$. Furthermore, $\partial_t\mathbf{v}_k \to \partial_t\mathbf{v}$ in $\mathrm{L}^2\left((0,T), \mathbf{H}_0^1(\Omega)\right)$ and $\partial_t\zeta_k \to \partial_t\zeta$ in $\mathrm{L}^2\left((0,T), \mathrm{H}_0^1(\Omega)\right)$. □

8.2.2 Stopping criterion

Reconsider the algorithm given in Subsection 8.2.1. The displacement at the final time is measured to obtain a solution to problem (8.1)-(8.2)-(8.3). There is noise present in each practical experiment. Hence, there is some error considered in the additional measurement (8.4), i.e.

$$\|\boldsymbol{\xi}_T - \boldsymbol{\xi}_T^e\| \leqslant e, \tag{8.18}$$

with $e > 0$. This implies that also $\widetilde{\boldsymbol{\xi}}_T$ is perturbed, see (8.16). The perturbed field is designated analogously by $\widetilde{\boldsymbol{\xi}}_T^e$. The functions $\mathbf{p}_k^e, \mathbf{v}_k^e$ and ζ_k^e are obtained by using the algorithm with no noise on the initial data (8.2). Note that this latter is a good assumption since the inverse problem is stable with respect to small perturbations in the initial data.

The absolute $\mathbf{L}^2$-error between this final measurement $\widetilde{\boldsymbol{\xi}}_T^e$ and the k-th approximation $\mathbf{v}_k^e(\cdot, t)$ at $t = T$ is denoted by

$$E_{k,\mathbf{u}_T} = \left\| \mathbf{v}_k^e(\cdot, T) - \widetilde{\boldsymbol{\xi}}_T^e \right\|. \tag{8.19}$$

and depends on the noise level e and the choice of the relaxation parameter κ. Given the noise level e, the discrepancy principle [163] can be used to obtain a stopping criterion for the algorithm [52, Proposition 6.4]. This principle suggests to finish the iterations at the lowest index $k = k(e, \kappa)$ for which

$$E_{k,\mathbf{u}_T} \leqslant e.$$

Remark 8.2.2. *The iteration index k takes the role of the regularization parameter α in the Tikhonov regularization and the stopping rule plays the role of the parameter selection method.*

Remark 8.2.3. *The algorithm is a special case of the gradient (steepest) descent method.*

8.3 Numerical experiment

In the numerical experiments, it is assumed that the density, the thermal coefficient, the Lamé parameters and the coupling coefficient are normalized to one, i.e. $\alpha = \beta = \gamma = \rho = 1$. The 1D linear model of type-I thermoelasticity is considered, which reads as: find $\langle u, \theta, p\rangle$ such that ($x \in (0, L)$ and $t \in (0, T]$)

$$\left\{ \begin{array}{rl} u_{tt}(x,t) + u_t(x,t) - u_{xx}(x,t) + \theta_x(x,t) & = p(x) + r(x,t), \\ \theta_t(x,t) - \theta_{xx}(x,t) + u_{xt}(x,t) & = h(x,t), \\ u(0,t) = u(L,t) = \theta(0,t) = \theta(L,t) & = 0, \\ u(x,0) = \overline{u}_0(x),\ u_t(x,0) = \overline{u}_1(x),\ \theta(x,0) & = \overline{\theta}_0(x) \end{array} \right. \tag{8.20}$$

and such that the final overdetermination condition

$$u(x,T) = \xi_T(x), \qquad x \in (0, L) \tag{8.21}$$

is satisfied. The solution to problem (8.20)-(8.21) is recovered by applying the algorithm proposed in Subsection 8.2.1. In all the experiments, it is assumed that $L = T = 1$. The forward mixed problems in this procedure are discretized in time according to the backward Euler method. The time step for the equidistant time partitioning is chosen to be 0.001. At each time step, the resulting elliptic mixed problems are solved numerically by the finite element method using first order (P1-FEM) Lagrange polynomials for the space discretization.

In each experiment, the exact solution for p is compared with the numerical solution $p_{\tilde{k}}$ obtained when the algorithm finishes after a finite number of $\tilde{k}$ iterations. The index $\tilde{k}$ is the lowest index $k = k(e, \kappa)$ for which $E_{k,\mathbf{u}_T} \leqslant e$ or is the maximum number of iterations when this number is reached. In both experiments, the maximum number of iterations equals 10000.

8.3.1 Experiment 1

The exact solution $\langle u, \theta, p \rangle$ to problem (8.20)-(8.21) is prescribed as follows

$$\begin{aligned} u(x,t) &= (1+t)^2 x(x-1), \\ \theta(x,t) &= (1+t)x(1-x), \\ p(x) &= x(x-1). \end{aligned} \tag{8.22}$$

Some simple calculations with the use of this exact solution give the exact data for the numerical experiment

$$\begin{aligned} r(x,t) &= 2\,tx^2 - 2\,t^2 - 4\,tx + 3\,x^2 - 3\,t - 5\,x - 1 \\ h(x,t) &= 4\,tx - x^2 + 5\,x \\ \xi_1(x) &= 4x(x-1) \\ \overline{u}_0(x) &= x(x-1) \\ \overline{u}_1(x) &= 2x(x-1) \\ \overline{\theta}_0(x) &= x(1-x). \end{aligned}$$

For the space discretization, a fixed uniform mesh consisting of 50 intervals is used. In this experiment, an uncorrelated noise is added to the additional condition (8.21) in order to simulate the errors present in real measurements. The noise is generated randomly with given magnitude $\tilde{e} = 1\%, 3\%$ and 5% and the resulting final measurement is denoted by $\xi_T^e(x)$, see also Section 8.2.2. This gives for (8.18) that

$$\|\xi_T - \xi_T^e\| \approx e(\tilde{e}) = \begin{cases} 0.0047 & \tilde{e} = 1\% \\ 0.0148 & \tilde{e} = 3\% \\ 0.0222 & \tilde{e} = 5\%. \end{cases} \tag{8.23}$$

According to the discrepancy principle, the algorithm is finished at the lowest index $k = k(e, \kappa)$ such that (8.19) is satisfied, see Table 8.1. The obtained results,

$\kappa \setminus \tilde{e}$	1%	3%	5%
1	151	108	107
10	14	10	10
50	3	2	2

Table 8.1: The stopping iteration number $k = k(e, \kappa)$ for Experiment 1 given by (8.19)*, with $e(\tilde{e})$ given by* (8.23)*.*

see Figure 8.1, are in accordance with the numerical experiments performed for the heat conduction equation in [175]. As κ or $\tilde{e}$ increases, the attainability of the stopping criterion (8.19) becomes faster. In this experiment, the numerical solution $p_{\tilde{k}}$, with $\tilde{k}$ the stopping index, approximates the unknown source better for larger values of κ. Even for a large amount of noise (5%), an accurate approximation for the source is obtained. Note that the algorithm is divergent for $\kappa > 50$.

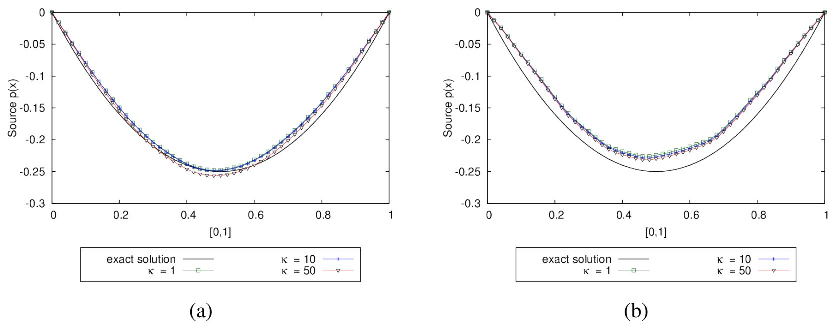


Figure 8.1: The exact solution (8.24) *and the numerical solution for the source* (8.22) *for $\tilde{e} = 1\%$ and $\tilde{e} = 5\%$ for different values of κ.*

8.3.2 Experiment 2

In this experiment a discontinuous source given by

$$p(x) = \begin{cases} 0 & \text{for } 0 \leqslant x < \frac{1}{3} \\ 1 & \text{for } \frac{1}{3} \leqslant x < \frac{2}{3} \\ 0 & \text{for } \frac{2}{3} < x \leqslant 1 \end{cases} \tag{8.24}$$

is reconstructed. Since the direct problems in the algorithm do not have an analytical solution for the given p, the data (8.21) is obtained by solving the direct problem using the FEM for $r = h = \overline{u}_0 = \overline{u}_1 = \overline{\theta}_0 = 0$. Now, a fixed uniform mesh consisting of 100 intervals is used for the space discretization. Only a randomly generated noise with magnitude $\tilde{e} = 1\%$ is considered on the resulting final measurement. In this experiment, $e(\tilde{e}) = 0.0048$. The stopping index is equal to 229 and 1145 for $\kappa = 50$ and $\kappa = 10$, respectively. There are more than 10000 iterations needed for $\kappa = 1$. The numerical solution for the source in the case that $\kappa = 10$ is given in Figure 8.2. The approximation is in accordance with the experiment performed in [175] with the same unknown source for only the heat conduction equation. Note that there is a large number of iterations needed to obtain this solution. In the light of this, other stopping criteria were also considered, but they need the same amount of iterations to obtain the same accuracy.

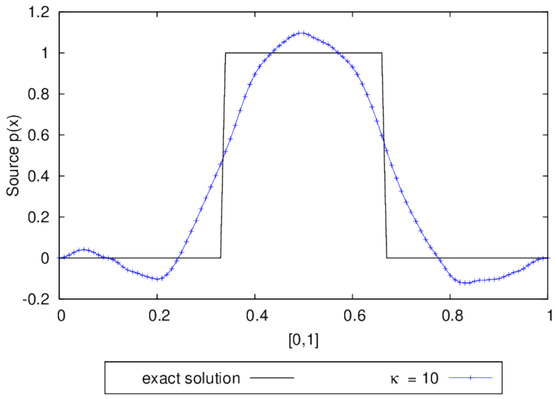


Figure 8.2: The exact solution (8.24) *and the numerical solution for the source* (8.24) *for* $\tilde{e} = 1\%$ *and* $\kappa = 10$.

8.4 Conclusion

The determination of a space-dependent vector source in a thermoelastic system of type-I, type-II and type-III has been studied using information from a supplementary measurement at the final time. The uniqueness of a solution to the inverse problem has been proved using a variational approach when a damping term $\mathbf{g}(\partial_t u)$ is added in the hyperbolic equation of the classic thermoelasticity system (7.1). The main assumption is that $\mathbf{g}$ is componentwise strictly monotone increasing.

Landweber's regularization method has been applied to cope with the ill-posedness of the inverse problem. In the case that the damping term is linear, a stable iterative algorithm has been proposed to recover the unknown source. This method is based on a sequence of well-posed direct problems that are numerically solved at each iteration step by using the finite element method. The instability of this inverse source problem has been overcome by stopping the iterations using the discrepancy principle of Morozov. The convergence of the algorithm has been illustrated by numerical experiments.

9

Recovery of a time-dependent source in 1D thermoelastic systems

This chapter is based on the article [181], which is submitted to *Inverse Problems in Science and Engineering*.

In this chapter, an inverse problem of determining a time-dependent heat source in a thermoelastic system of type-III using an additional global measurement is studied. Up to now, to the best of the author's knowledge, there are no papers dealing with this topic.

An isotropic and homogeneous thermoelastic body occupying a one-dimensional slab of length L is considered, i.e. $\Omega = (0, L) \subset \mathbb{R}$. Let $Q_T = \Omega \times (0, T)$ for a given final time $T > 0$. The convolution product of a kernel k and a function v is denoted with the sign '$*$', i.e.

$$(k * v)(x,t) := \int_0^t k(t-s)v(x,s)\mathrm{d}s, \qquad (x,t) \in Q_T.$$

For simplicity, the following notations for spacial and time derivatives for functions depending both on time and space variables are used:

$$\begin{array}{rlrl} v'(x,t) & := \frac{\partial v}{\partial x}(x,t), & v''(x,t) & := \frac{\partial^2 v}{\partial x^2}(x,t), \\ \dot{v}(x,t) & := \frac{\partial v}{\partial t}(x,t), & \ddot{v}(x,t) & := \frac{\partial^2 v}{\partial t^2}(x,t). \end{array}$$

The following thermoelastic system of type-III describing the elastic and thermal behaviour in Ω is discussed: find a triple $\langle u, \theta, h\rangle$ such that

$$\begin{cases} \ddot{u} - \alpha u'' + \gamma\theta' &= r & \text{in } Q_T, \\ \dot{\theta} - \rho\theta'' - k * \theta'' + \gamma\dot{u}' &= h(t)f(x) + s & \text{in } Q_T, \\ u'(0,t) = u'(L,t) = \theta(0,t) = \theta(L,t) &= 0 & \text{in } (0,T], \\ u(x,0) = u_0(x),\ \dot{u}(x,0) = \dot{u}_0(x),\ \theta(x,0) &= \theta_0(x) & \text{in } (0,L). \end{cases} \tag{9.1}$$

Here, u and θ denote respectively the displacement and the temperature difference from the reference value (in Kelvin) of the solid elastic material at the location x and time t. The Lamé parameter α, the coupling (absorbing) coefficient γ and the thermal coefficient ρ are assumed to be positive constants because the medium is supposed to be isotropic homogeneous. The kernel function (also called relaxation function) $k \in \mathrm{C}([0,T])$ is decaying to zero as the time goes to infinity.

Due to a lack of information, the space average of the temperature is measured to recover the unknown source $h(t)$, i.e.

$$\int_0^L \theta(x,t)\mathrm{d}x = m(t), \quad t \in [0,T]. \tag{9.2}$$

The added value of this research consists of the global (in time) solvability of problem (9.1)-(9.2) and of the designed numerical scheme for computations. The way of retrieving the triplet $\langle u, \theta, h\rangle$ is not by the minimization of a cost functional (which is typical for IPs), but by the semidiscretization in time by Rothe's method.

In [42, 182], the reconstruction of an unknown time-dependent source term in a semilinear parabolic problem is studied. The same subject for a damped wave equation is addressed in [183]. In [184, 185], the authors investigated the identification of a solely time-dependent memory kernel in a semilinear integrodifferential parabolic problem. In these papers, the inverse problems are reformulated into an appropriate direct formulation by using an additional measurement. A numerical scheme is developed to recover the unknowns. The same technique is used in this contribution. Nevertheless, the analysis (uniqueness of a solution, a priori estimates, convergence of the numerical scheme) is more complicated due to the coupling of both PDEs under consideration, even in this one-dimensional case.

The remainder of this chapter is organized as follows. In Section 9.1, the inverse problem is re-casted into a direct problem by using the additional measurement. A suitable variational formulation is deduced. The uniqueness of a solution is studied in Section 9.2. Section 9.3 deals with the time discretization. The convergence of the proposed numerical scheme to the unique weak solution of problem (9.1)-(9.2) is shown. Finally, in Section 9.4, numerical experiments support the theoretically obtained results.

Remark 9.0.1. *The results in this chapter stay true for type-I thermoelasticity ($\rho \neq 0, k = 0$). However, they are not valid for type-II thermoelasticity ($\rho =$*

$0, k \neq 0$*). The results in this chapter are not valid if the boundary conditions are switched:*

$$u(0,t) = u(L,t) = \theta'(0,t) = \theta'(L,t) = 0 \quad \textit{in } (0,T].$$

9.1 Reformulation of the inverse problem to a direct problem

The idea is to recast the inverse problem into a direct coupled problem by eliminating the unknown source function h using the additional measurement (9.2).

First, the second equation in (9.1) is integrated over Ω. This gives an expression for the unknown function h in terms of the unknown u and θ, i.e.

$$h(t) = \frac{\dot{m}(t) - \rho \int_0^L \theta''(t) - \left(k * \int_0^L \theta''\right)(t) + \gamma \int_0^L \dot{u}'(t) - \int_0^L s(t)}{\int_0^L f} \in \mathbb{R}, \tag{9.3}$$

$t \in (0,T]$, if $\int_0^L f \neq 0$. Next, this expression for h is substituted in (9.1). Using Green's formulas, the following coupled variational formulation for (9.1)-(9.2) is obtained: find $\langle u(t), \theta(t) \rangle \in \mathrm{H}^1(\Omega) \times \mathrm{H}^1_0(\Omega)$ such that

$$(\ddot{u}(t), \phi) + \alpha\,(u'(t), \phi') + \gamma\,(\theta'(t), \phi) = (r(t), \phi) \tag{9.4}$$

and

$$\begin{aligned}\left(\dot{\theta}(t), \psi\right) &+ \rho\,(\theta'(t), \psi') + (k * \theta'(t), \psi') - \gamma\,(\dot{u}(t), \psi') \\ &= \frac{\dot{m}(t) - \rho \int_0^L \theta''(t) - \left(k * \int_0^L \theta''\right)(t) + \gamma \int_0^L \dot{u}'(t) - \int_0^L s(t)}{\int_0^L f}\,(f, \psi) \\ &\qquad + (s(t), \psi)\end{aligned} \tag{9.5}$$

for a.a. $t \in [0,T]$ and for all $\phi \in \mathrm{H}^1(\Omega)$ and $\psi \in \mathrm{H}^1_0(\Omega)$.

The following Hilbert spaces for u and θ are used in the analysis

$$\begin{aligned} V_u &= \mathrm{H}^2(\Omega), & V_u^* &= \mathrm{H}^2(\Omega)^*, \\ V_\theta &= \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega), & V_\theta^* &= \left(\mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)\right)^*. \end{aligned}$$

These spaces are endowed with the norms

$$\|\phi\|^2_{V_u} = \|\phi\|^2_{\mathrm{H}^1(\Omega)} + \|\phi''\|^2$$

and

$$\|\psi\|^2_{V_\theta} = \|\psi\|^2_{\mathrm{H}^1(\Omega)} + \|\psi''\|^2 \equiv \|\psi''\|^2.$$

The equivalence of the norms $\|\psi\|^2_{\mathrm{H}^1(\Omega)} + \|\psi''\|^2$ and $\|\psi''\|^2$ in V_θ is given in Theorem 2.9.32.

9.2 Uniqueness

In this section, the uniqueness of the solution to the ISP (9.1)-(9.2) is proved under the assumption that the variational problem has a solution.

Theorem 9.2.1 (Uniqueness). *Assume that* $m \in \mathrm{C}([0,T])$, $k \in \mathrm{C}([0,T])$, $f \in \mathrm{H}^1(\Omega)$ *and* $\int_0^L f \neq 0$. *Then there exists at most one triple*

$$\langle u, \theta, h\rangle \in \left[\mathrm{C}\left([0,T], \mathrm{H}_0^1(\Omega)\right) \cap \mathrm{L}^2\left((0,T), V_u\right)\right] \times \mathrm{L}^2\left((0,T), V_\theta\right) \times \mathrm{L}^2(0,T)$$

solving problem (9.1)-(9.2).

Proof. Suppose that there are two solutions $\langle u_1, \theta_1, h_1\rangle$ and $\langle u_2, \theta_2, h_2\rangle$ to (9.1)-(9.2). Then $u := u_1 - u_2$, $\theta := \theta_1 - \theta_2$ and $h := h_1 - h_2$ satisfy (9.1)-(9.2) with $m = r = s = u_0 = \dot{u}_0 = \theta_0 = 0$. We cannot prove the uniqueness of a solution by using the classical weak formulation (9.4) – (9.5) since higher regularity is necessary to hold the second space derivative of θ under control in (9.3). From this point of view, the first equation of (9.1) is multiplied with $\dot{u}''(x,t)$ and the second equation with $\theta''(x,t)$. This choice of test functions imply also that the coupling term is cancelled out later in the proof. The resulting equations are integrated over the domain. Then, using Green's formulas and (9.3), we get that

$$(\ddot{u}'(t), \dot{u}'(t)) + \alpha\left(u''(t), \dot{u}''(t)\right) + \gamma\left(\theta''(t), \dot{u}'(t)\right) = 0 \tag{9.6}$$

and

$$\begin{aligned}\left(\dot{\theta}'(t), \theta'(t)\right) + \rho\left\|\theta''(t)\right\|^2 + \left(\left(k * \theta''\right)(t), \theta''(t)\right) - \gamma\left(\dot{u}'(t), \theta''(t)\right)&\\ = \frac{\rho\int_0^L \theta''(t) + \left(k * \int_0^L \theta''\right)(t) - \gamma\int_0^L \dot{u}'(t)}{\int_0^L f}\left(f, \theta''(t)\right).&\end{aligned}$$

We integrate both equations in time over $(0,\eta) \subset (0,T)$ and add them up to obtain

$$\begin{aligned}&\frac{\left\|\dot{u}'(\eta)\right\|^2}{2} + \alpha\frac{\left\|u''(\eta)\right\|^2}{2} + \frac{\left\|\theta'(\eta)\right\|^2}{2}\\ &\qquad + \rho\int_0^\eta \left\|\theta''(t)\right\|^2 \mathrm{d}t + \int_0^\eta \left(\left(k * \theta''\right)(t), \theta''(t)\right) \mathrm{d}t\\ &= \int_0^\eta \frac{\rho\int_0^L \theta''(t) + \left(k * \int_0^L \theta''\right)(t) - \gamma\int_0^L \dot{u}'(t)}{\int_0^L f}\left(f, \theta''(t)\right) \mathrm{d}t.\end{aligned} \tag{9.7}$$

Using Nečas inequality (2.12) in 1D

$$|z(0)| + |z(L)| \leqslant \varepsilon\left\|z'\right\| + C_\varepsilon\left\|z\right\|, \quad z \in \mathrm{H}^1(\Omega), \quad 0 < \varepsilon < \varepsilon_0,$$

the RHS of (9.7) can be estimated as

$$\begin{aligned}
&\left|\int_0^\eta \frac{\rho\int_0^L \theta''(t) + \left(k * \int_0^L \theta''\right)(t) - \gamma\int_0^L \dot{u}'(t)}{\int_0^L f}\,(f,\theta''(t))\,\mathrm{d}t\right| \\
&\lesssim \int_0^\eta \left(\|\theta''(t)\|_{\mathrm{L}^1(\Omega)} + \int_0^t \|\theta''(s)\|_{\mathrm{L}^1(\Omega)}\,\mathrm{d}s + \|\dot{u}'(t)\|_{\mathrm{L}^1(\Omega)}\right)|(f,\theta''(t))|\,\mathrm{d}t \\
&\lesssim \int_0^\eta \left(\|\theta''(t)\| + \int_0^t \|\theta''\| + \|\dot{u}'(t)\|\right)\left|-(f',\theta'(t)) + f(x)\theta'(x,t)|_0^L\right|\mathrm{d}t \\
&\lesssim \int_0^\eta \left(\|\theta''(t)\| + \int_0^t \|\theta''\| + \|\dot{u}'(t)\|\right)(\varepsilon_1\|\theta''(t)\| + C_{\varepsilon_1}\|\theta'(t)\|)\,\mathrm{d}t \\
&\lesssim (\varepsilon_1 + \varepsilon_2 C_{\varepsilon_1})\int_0^\eta \|\theta''(t)\|^2\,\mathrm{d}t + (C_{\varepsilon_1}C_{\varepsilon_2} + C_{\varepsilon_1})\int_0^\eta \|\theta'(t)\|^2\,\mathrm{d}t \\
&\quad + (\varepsilon_1 + C_{\varepsilon_1})\int_0^\eta\left(\int_0^t \|\theta''(s)\|^2\,\mathrm{d}s\right)\mathrm{d}t + (\varepsilon_1 + C_{\varepsilon_1})\int_0^\eta \|\dot{u}'(t)\|^2\,\mathrm{d}t,
\end{aligned}$$

if $k \in \mathrm{C}([0,T])$ and $f \in \mathrm{H}^1(\Omega)$. The last term in the LHS can be estimated as

$$\begin{aligned}
&\left|\int_0^\eta \left(\int_0^t k(t-s)\theta''(s)\mathrm{d}s, \theta''(t)\right)\mathrm{d}t\right| \\
&\quad\leqslant\; C_{\varepsilon_3}\int_0^\eta \left\|\int_0^t k(t-s)\theta''(s)\mathrm{d}s\right\|^2\mathrm{d}t + \varepsilon_3\int_0^\eta \|\theta''(t)\|^2\,\mathrm{d}t \\
&\quad\leqslant\; C_{\varepsilon_3}\int_0^\eta\left(\int_0^t \|\theta''(s)\|^2\,\mathrm{d}s\right)\mathrm{d}t + \varepsilon_3\int_0^\eta \|\theta''(t)\|^2\,\mathrm{d}t.
\end{aligned}$$

Now, we first fix ε_1 and ε_3, and then ε_2 such that $\varepsilon_1 + \varepsilon_2 C_{\varepsilon_1} + \varepsilon_3$ is sufficiently small to simplify equation (9.7) to

$$\begin{aligned}
&\|\dot{u}'(\eta)\|^2 + \|u''(\eta)\|^2 + \|\theta'(\eta)\|^2 + \int_0^\eta \|\theta''\|^2 \\
&\qquad\qquad \lesssim \int_0^\eta \left[\|\theta'\|^2 + \left(\int_0^t \|\theta''\|^2\right) + \|\dot{u}'\|^2\right].
\end{aligned}$$

An application of Grönwall's lemma implies that

$$\|\dot{u}'(\eta)\|^2 + \|u''(\eta)\|^2 + \|\theta'(\eta)\|^2 + \int_0^\eta \|\theta''\|^2 = 0. \tag{9.8}$$

We see that $\theta' = 0$ in Q_T. Due to the homogeneous boundary condition for θ, we get that

$$\theta = 0 \text{ in } Q_T.$$

From equation (9.8), we also have that $\|u''(\eta)\| = 0$ in $(0,T)$ or $u'(\cdot,\eta)$ is constant in Ω for every $\eta \in [0,T]$. Employing the boundary conditions of u, we get that

$u'(x,\eta) = 0$ with $(x,\eta) \in Q_T$. This implies that u is constant in Q_T. This fact, together with the initial condition $u_0 = 0$ in Ω, gives that $u = 0$ in Q_T. Now, the uniqueness of h follows immediately from (9.3). $\square$

9.3 Existence of a solution

To address the existence of a solution to (9.1)-(9.2), a semidiscretization in time is employed. This discretization is based on Rothe's method, see Section 2.12. The interval $[0,T]$ is divided into $n \in \mathbb{N}$ equidistant subintervals $[t_{i-1}, t_i]$ with time step $\tau = \frac{T}{n} < 1$, thus $t_i = i\tau, i = 0, \ldots, n$. With the standard notation for the discretized fields for any function z

$$z_i \approx z(t_i) \quad \text{and} \quad \delta z_i = \frac{z_i - z_{i-1}}{\tau},$$

the following linear recurrent scheme is proposed to approximate the original problem for $i = 1, \ldots, n$ ($\phi \in \mathrm{H}^1(\Omega)$ and $\psi \in \mathrm{H}^1_0(\Omega)$):

$$\left(\delta^2 u_i, \phi\right) + \alpha\left(u_i', \phi'\right) + \gamma\left(\theta_i', \phi\right) = \left(r_i, \phi\right), \tag{9.9}$$

$$\begin{aligned}\left(\delta\theta_i, \psi\right) + \rho\left(\theta_i', \psi'\right) + \left(\sum_{l=1}^{i} k_l \theta_{i-l}'\tau, \psi'\right) \\ -\gamma\left(\delta u_i, \psi'\right) = h_{i-1}\left(f, \psi\right) + \left(s_i, \psi\right),\end{aligned} \tag{9.10}$$

$$u_0 = u_0, \quad \delta u_0 = \dot{u}_0, \quad \theta_0 = \theta_0 \tag{9.11}$$

with

$$h_i := \frac{m'(t_{i+1}) - \rho\int_0^L \theta_i'' - \sum_{l=1}^{i+1} \tau k_l \int_0^L \theta_{i+1-l}'' + \gamma\int_0^L \delta u_i' - \int_0^L s_{i+1}}{\int_0^L f}. \tag{9.12}$$

Note that $\int_0^L \theta_i'' = \theta_i'|_0^L$ and that for a given $i \in \{1, \ldots, n\}$, first equation (9.12) and next problem (9.9)-(9.10)-(9.11) are solved. Then, i is increased to $i+1$. This is equivalent with solving for any $i \in \{1, \ldots, n\}$:

$$a\left(\begin{pmatrix} u_i \\ \theta_i \end{pmatrix}, \begin{pmatrix} \phi \\ \psi \end{pmatrix}\right) = F_i\begin{pmatrix} \phi \\ \psi \end{pmatrix}, \quad u_0 = u_0, \quad \delta u_0 = \dot{u}_0, \quad \theta_0 = \theta_0$$

with

$$a\left(\begin{pmatrix} u_i \\ \theta_i \end{pmatrix}, \begin{pmatrix} \phi \\ \psi \end{pmatrix}\right) := \frac{L_1}{\tau} + L_2, \quad F_i\begin{pmatrix} \phi \\ \psi \end{pmatrix} := \frac{R_1}{\tau} + R_2,$$

and

$$\begin{aligned}
L_1 &:= \frac{1}{\tau^2}(u_i,\phi) + \alpha\,(u_i',\phi') + \gamma\,(\theta_i',\phi) \\
&= (r_i,\phi) + \frac{1}{\tau^2}(u_{i-1},\phi) + \frac{1}{\tau}(\delta u_{i-1},\phi) =: R_1, \qquad (9.13)\\
L_2 &:= \frac{1}{\tau}(\theta_i,\psi) + \rho\,(\theta_i',\psi') - \frac{\gamma}{\tau}(u_i,\psi') \\
&= h_{i-1}(f,\psi) + (s_i,\psi) - \left(\sum_{l=1}^{i} k_l\theta_{i-l}'\tau,\psi'\right) \\
&\quad + \frac{1}{\tau}(\theta_{i-1},\psi) - \frac{\gamma}{\tau}(u_{i-1},\psi') =: R_2. \qquad (9.14)
\end{aligned}$$

In the following theorem, the existence and uniqueness of a solution on a single time step is proved.

Theorem 9.3.1 (Single time step)**.** *Suppose that* $k \in \mathrm{C}([0,T])$*,* $m \in \mathrm{C}^1([0,T])$*,* $r, s \in \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$ *and* $f \in \mathrm{H}^1(\Omega)$ *with* $\int_0^L f \neq 0$*. Moreover, assume that* $\theta_0 \in \mathrm{H}^2(\Omega)$*,* $u_0 \in \mathrm{H}^1(\Omega)$ *and* $\dot u_0 \in \mathrm{H}^1(\Omega)$*. Then there exists a unique triple* $(u_i,\theta_i,h_i) \in V_u \times V_\theta \times \mathbb{R}$ *solving* (9.9)-(9.10)-(9.11) *for* $i = 1,\ldots,n$ *and for any* $\tau > 0$*.*

Proof. The bilinear form a is coercive and continuous on $\mathrm{H}^1(\Omega)\times\mathrm{H}^1_0(\Omega)$. Note that

$$|h_0| \lesssim \left| m'(t_1) - \rho\int_0^L \theta_0'' - \tau k_1 \int_0^L \theta_0'' + \gamma\int_0^L \dot u_0' - \int_0^L s_1 \right| \lesssim 1$$

if $\theta_0 \in \mathrm{H}^2(\Omega)$ and $\dot u_0 \in \mathrm{H}^1(\Omega)$. Then $F_1 \in \left(\mathrm{H}^1(\Omega)\times\mathrm{H}^1_0(\Omega)\right)^*$ if also $u_0 \in \mathrm{L}^2(\Omega)$. From the Lax-Milgram lemma 2.11.1, we obtain the existence and uniqueness of a solution $(u_1,\theta_1) \in \mathrm{H}^1(\Omega)\times\mathrm{H}^1_0(\Omega)$ to (9.13)-(9.14) or equivalently to (9.9)-(9.10). Now, we can apply Green's theorem in a backward way to equations (9.13)-(9.14) for $i = 1$ to obtain for any $\phi \in \mathrm{H}^1(\Omega)$ and $\psi \in \mathrm{H}^1_0(\Omega)$ that

$$\begin{aligned}
-\alpha\,(u_1'',\phi) &= (r_1,\phi) + \frac{1}{\tau^2}(u_0,\phi) + \frac{1}{\tau}(\dot u_0,\phi) - \frac{1}{\tau^2}(u_1,\phi) - \gamma\,(\theta_1',\phi)\,, \\
-\rho\,(\theta_1'',\psi) &= h_0\,(f,\psi) + (s_1,\psi) + (k_1\theta_0''\tau,\psi) \\
&\quad + \frac{1}{\tau}(\theta_0,\psi) + \frac{\gamma}{\tau}(u_0',\psi) - \frac{1}{\tau}(\theta_1,\psi) - \frac{\gamma}{\tau}(u_1',\psi)\,.
\end{aligned}$$

The term $-\alpha u_1''$ has to be understood in the sense of duality, as a functional on $\mathrm{H}^1(\Omega)$. In the same way, the term $-\rho\theta_1''$ can be seen as a functional on $\mathrm{H}^1_0(\Omega)$.

The RHSs of the first and second equation are a linear and bounded functional on $\mathrm{H}^1(\Omega)$ and on $\mathrm{H}^1_0(\Omega)$ respectively. Thus both RHSs can be extended to a functional on $\mathrm{L}^2(\Omega)$ with the same norm by the density of $\mathrm{H}_0^1(\Omega)$ and $\mathrm{H}^1(\Omega)$ in $\mathrm{L}^2(\Omega)$ and applying the Hahn-Banach theorem. Therefore,

$$\begin{aligned}
-\alpha u_1'' &= r_1 + \frac{1}{\tau^2}u_0 + \frac{1}{\tau}\dot{u}_0 - \frac{1}{\tau^2}u_1 - \gamma\theta_1' \in \mathrm{L}^2(\Omega),\\
-\rho\theta_1'' &= h_0 f + s_1 + k_1\theta_0''\tau + \frac{1}{\tau}\theta_0 + \frac{\gamma}{\tau}u_0' - \frac{1}{\tau}\theta_1 - \frac{\gamma}{\tau}u_1' \in \mathrm{L}^2(\Omega).
\end{aligned}$$

Hence, $(u_1,\theta_1) \in V_u \times V_\theta$. Now, we can go to the following time steps and prove in an analogous way that also $(u_i,\theta_i) \in V_u \times V_\theta$ for $i = 2,\ldots,n$. The uniqueness of $h_i \in \mathbb{R}$ follows from the uniqueness of $(u_i,\theta_i) \in V_u \times V_\theta$, $i = 1,\ldots,n$. $\square$

Lemma 9.3.1. *Let the assumptions of Theorem 9.3.1 be fulfilled. Then there exist positive constants C and τ_0 such that for $0 < \tau < \tau_0$ we have*

$$\begin{aligned}
&\max_{1\leqslant j\leqslant n}\left\{\|\delta u_j\|^2 + \left\|u_j'\right\|^2 + \|\theta_j\|^2\right\} + \sum_{i=1}^n \|\delta u_i - \delta u_{i-1}\|^2\\
&+\sum_{i=1}^n \left\|u_i' - u_{i-1}'\right\|^2 + \sum_{i=1}^n \|\theta_i - \theta_{i-1}\|^2 + \sum_{i=1}^n \|\theta_i'\|^2\tau \leqslant C\left(1 + \sum_{i=1}^n |h_{i-1}|^2\tau\right).
\end{aligned}$$

Proof. We set $\phi = \delta u_i\tau$ and $\psi = \theta_i\tau$ in (9.9)-(9.10) and sum both equations up for $i = 1,\ldots,j$ with $1 \leqslant j \leqslant n$. Then, we add both resulting equations up. In this way, the coupling term is cancelled out and we obtain that

$$\begin{aligned}
&\sum_{i=1}^j \left(\delta^2 u_i, \delta u_i\right)\tau + \alpha\sum_{i=1}^j \left(u_i', \delta u_i'\right)\tau + \sum_{i=1}^j \left(\delta\theta_i, \theta_i\right)\tau + \rho\sum_{i=1}^j \|\theta_i'\|^2\tau\\
&+\sum_{i=1}^j\left(\sum_{l=1}^i k_l\theta_{i-l}'\tau, \theta_i'\right)\tau = \sum_{i=1}^j \left(r_i, \delta u_i\right)\tau + \sum_{i=1}^j h_{i-1}\left(f,\theta_i\right)\tau + \sum_{i=1}^j \left(s_i,\theta_i\right)\tau.
\end{aligned}$$

We use Abel's summation rule for the first three terms on the LHS:

$$\begin{aligned}
2\sum_{i=1}^j \left(\delta^2 u_i, \delta u_i\right)\tau &= \|\delta u_j\|^2 - \|\dot{u}_0\|^2 + \sum_{i=1}^j \|\delta u_i - \delta u_{i-1}\|^2,\\
2\sum_{i=1}^j \left(u_i', \delta u_i'\right)\tau &= \left\|u_j'\right\|^2 - \|u_0'\|^2 + \sum_{i=1}^j \left\|u_i' - u_{i-1}'\right\|^2,\\
2\sum_{i=1}^j \left(\delta\theta_i, \theta_i\right)\tau &= \|\theta_j\|^2 - \|\theta_0\|^2 + \sum_{i=1}^j \|\theta_i - \theta_{i-1}\|^2.
\end{aligned}$$

Using Hölder's inequality, the last term in the LHS is bounded by

$$\begin{aligned}\left|\sum_{i=1}^{j}\left(\sum_{l=1}^{i}k_l\theta'_{i-l}\tau,\theta'_i\right)\tau\right| &\leqslant C_\varepsilon\sum_{i=1}^{j}\left\|\sum_{l=1}^{i}k_l\theta'_{i-l}\tau\right\|^2\tau+\varepsilon\sum_{i=1}^{j}\left\|\theta'_i\right\|^2\tau\\ &\leqslant C_\varepsilon\sum_{i=1}^{j}\left(\sum_{l=1}^{i}\left\|\theta'_{i-l}\right\|^2\tau\right)\tau+\varepsilon\sum_{i=1}^{j}\left\|\theta'_i\right\|^2\tau\\ &\leqslant C_\varepsilon\sum_{i=1}^{j}\left(\sum_{l=0}^{i-1}\left\|\theta'_{l}\right\|^2\tau\right)\tau+\varepsilon\sum_{i=1}^{j}\left\|\theta'_i\right\|^2\tau.\end{aligned}$$

The first and third terms in the RHS can be estimated in a classical way by using the Cauchy and Young inequalities. Analogously, for the second term in the RHS, we get

$$\left|\sum_{i=1}^{j}h_{i-1}\left(f,\theta_i\right)\tau\right|\lesssim\sum_{i=1}^{j}|h_{i-1}|^2\tau+\sum_{i=1}^{j}\left\|\theta_i\right\|^2\tau.$$

Collecting all the results above, fixing ε small enough and applying Grönwall's argument concludes the proof. □

Lemma 9.3.2. *Let the assumptions of Theorem 9.3.1 be fulfilled. Moreover, assume that* $r\in\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$. *Then there exist positive constants* C *and* τ_0 *such that for* $0<\tau<\tau_0$, *it holds that*

$$\begin{aligned}\max_{1\leqslant i\leqslant n}\left\{\left\|\delta u'_j\right\|^2+\left\|\theta'_j\right\|^2+\left\|u''_j\right\|^2\right\}+\sum_{i=1}^{n}\left\|\theta''_i\right\|^2\tau\\ +\sum_{i=1}^{n}\left\|\delta u'_i-\delta u'_{i-1}\right\|^2+\sum_{i=1}^{n}\left\|\theta'_i-\theta'_{i-1}\right\|^2\\ +\sum_{i=1}^{n}\left\|u''_i-u''_{i-1}\right\|^2\leqslant C\left(1+\sum_{i=1}^{n}|h_{i-1}|^2\tau\right).\end{aligned}$$

Proof. The starting point is the strong form of (9.9)-(9.10) in $\mathrm{L}^2(\Omega)$:

$$\delta^2u_i-\alpha u''_i+\gamma\theta'_i=r_i, \tag{9.15}$$

$$\delta\theta_i-\rho\theta''_i-k*\theta''_i+\gamma\delta u'_i=h_{i-1}f+s_i. \tag{9.16}$$

We multiply the first equation with $-\delta u''_i\tau$ and the second equation with $-\theta''_i\tau$. Then, we integrate both equations over Ω and sum the resulting equations up for $i=1,\ldots,j$ ($1\leqslant j\leqslant n$) to obtain

$$\begin{aligned}-\sum_{i=1}^{j}\left(\delta^2u_i,\delta u''_i\right)\tau+\alpha\sum_{i=1}^{j}\left(u''_i,\delta u''_i\right)\tau\\ -\gamma\sum_{i=1}^{j}\left(\theta'_i,\delta u''_i\right)\tau=-\sum_{i=1}^{j}\left(r_i,\delta u''_i\right)\tau\end{aligned} \tag{9.17}$$

and

$$-\sum_{i=1}^{j}(\delta\theta_i,\theta_i'')\,\tau+\rho\sum_{i=1}^{j}(\theta_i'',\theta_i'')\,\tau+\sum_{i=1}^{j}\left(\sum_{l=1}^{i}k_l\theta_{i-l}''\tau,\theta_i''\right)\tau$$
$$-\gamma\sum_{i=1}^{j}(\delta u_i',\theta_i'')\,\tau=-\sum_{i=1}^{j}(h_{i-1}f,\theta_i'')\,\tau-\sum_{i=1}^{j}(s_i,\theta_i'')\,\tau. \quad (9.18)$$

Note that due to the boundary conditions, we may write

$$\sum_{i=1}^{j}(\delta u_i',\theta_i'')\,\tau=-\sum_{i=1}^{j}(\theta_i',\delta u_i'')\,\tau.$$

Then, adding up relations (9.17) and (9.18) implies

$$-\sum_{i=1}^{j}\left(\delta^2u_i,\delta u_i''\right)\tau+\alpha\sum_{i=1}^{j}(u_i'',\delta u_i'')\,\tau-\sum_{i=1}^{j}(\delta\theta_i,\theta_i'')\,\tau+\rho\sum_{i=1}^{j}\|\theta_i''\|^2\,\tau$$
$$=-\sum_{i=1}^{j}(r_i,\delta u_i'')\,\tau-\sum_{i=1}^{j}(h_{i-1}f,\theta_i'')\,\tau$$
$$-\sum_{i=1}^{j}(s_i,\theta_i'')\,\tau-\sum_{i=1}^{j}(k*\theta_i'',\theta_i'')\,\tau. \quad (9.19)$$

We apply the integration by parts formula and Abel's summation rule on the first and third terms of equation (9.19), i.e.

$$-\sum_{i=1}^{j}\left(\delta^2u_i,\delta u_i''\right)\tau=\sum_{i=1}^{j}\left(\delta^2u_i',\delta u_i'\right)\tau$$
$$=\frac{\left\|\delta u_j'\right\|^2}{2}-\frac{\|\dot{u}_0'\|^2}{2}+\tfrac{1}{2}\sum_{i=1}^{j}\left\|\delta u_i'-\delta u_{i-1}'\right\|^2,$$
$$-\sum_{i=1}^{j}(\delta\theta_i,\theta_i'')\,\tau=\sum_{i=1}^{j}(\delta\theta_i',\theta_i')\,\tau=\frac{\left\|\theta_j'\right\|^2}{2}-\frac{\|\theta_0'\|^2}{2}+\tfrac{1}{2}\sum_{i=1}^{j}\left\|\theta_i'-\theta_{i-1}'\right\|^2.$$

The second term equals

$$\sum_{i=1}^{j}(u_i'',\delta u_i'')\,\tau=\frac{\left\|u_j''\right\|^2}{2}-\frac{\|u_0''\|^2}{2}+\tfrac{1}{2}\sum_{i=1}^{j}\left\|u_i''-u_{i-1}''\right\|^2.$$

Now, we can estimate the remaining terms in (9.19) by using the Cauchy and Young inequalities as in Lemma 9.3.1. We only point out the first two terms in the

RHS of (9.19). From the assumptions $r \in \mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$, $f \in \mathrm{H}^1(\Omega)$ and the Nečas inequality, we deduce that

$$-\sum_{i=1}^{j}(r_i,\delta u_i'')\,\tau = \sum_{i=1}^{j}(r_i',\delta u_i')\,\tau \lesssim \sum_{i=1}^{j}\|r_i'\|^2\,\tau + \sum_{i=1}^{j}\|\delta u_i'\|^2\,\tau$$

and

$$\left|\sum_{i=1}^{j}(h_{i-1}f,\theta_i'')\,\tau\right| \leqslant C_\varepsilon\sum_{i=1}^{j}|h_{i-1}|^2\tau + \varepsilon\sum_{i=1}^{j}\|\theta_i''\|^2\,\tau.$$

The rest of the proof is straightforward. □

Lemma 9.3.3. *Let the assumptions of Lemma 9.3.2 be satisfied. Then there exist positive constants C and τ_0 such that for $0<\tau<\tau_0$ it holds that*

$$\begin{aligned}&\max_{1\leqslant j\leqslant n}\left\{\|\delta u_j\|^2 + \left\|u_j'\right\|^2 + \|\theta_j\|^2 + \left\|\delta u_j'\right\|^2 + \left\|\theta_j'\right\|^2 + \left\|u_j''\right\|^2\right\}\\ &+\sum_{i=1}^{n}\|\delta u_i-\delta u_{i-1}\|^2 + \sum_{i=1}^{n}\left\|u_i'-u_{i-1}'\right\|^2 + \sum_{i=1}^{n}\|\theta_i-\theta_{i-1}\|^2 + \sum_{i=1}^{n}\|\theta_i''\|^2\,\tau\\ &\qquad+\sum_{i=1}^{n}\left\|\delta u_i'-\delta u_{i-1}'\right\|^2 + \sum_{i=1}^{n}\left\|\theta_i'-\theta_{i-1}'\right\|^2 + \sum_{i=1}^{n}\left\|u_i''-u_{i-1}''\right\|^2 \leqslant C.\end{aligned}$$

and

$$\sum_{i=0}^{n-1}|h_i|^2\tau \leqslant C.$$

Proof. Putting the results of Lemma 9.3.1 and 9.3.2 together, we get that

$$\begin{aligned}&\|\delta u_j\|^2 + \left\|u_j'\right\|^2 + \|\theta_j\|^2 + \left\|\delta u_j'\right\|^2 + \left\|\theta_j'\right\|^2 + \left\|u_j''\right\|^2\\ &+\sum_{i=1}^{j}\|\delta u_i-\delta u_{i-1}\|^2 + \sum_{i=1}^{j}\left\|u_i'-u_{i-1}'\right\|^2 + \sum_{i=1}^{j}\|\theta_i-\theta_{i-1}\|^2 + \sum_{i=1}^{j}\|\theta_i'\|^2\,\tau\\ &+\sum_{i=1}^{j}\|\theta_i''\|^2\,\tau + \sum_{i=1}^{j}\left\|\delta u_i'-\delta u_{i-1}'\right\|^2 + \sum_{i=1}^{j}\left\|\theta_i'-\theta_{i-1}'\right\|^2 + \sum_{i=1}^{j}\left\|u_i''-u_{i-1}''\right\|^2\\ &\leqslant C_\varepsilon + C_\varepsilon\sum_{i=1}^{j}|h_{i-1}|^2\tau + C_\varepsilon\sum_{i=1}^{j}\|\theta_i\|^2\,\tau + C\sum_{i=1}^{j}\|\delta u_i\|^2\,\tau + C\sum_{i=1}^{j}\|\delta u_i'\|^2\,\tau\\ &\qquad+ C_\varepsilon\sum_{i=1}^{j}\|\theta_i'\|^2\,\tau + \varepsilon\sum_{i=1}^{j}\|\theta_i''\|^2\,\tau + C_\varepsilon\sum_{i=1}^{j}\left(\sum_{l=0}^{i-1}\|\theta_l'\|^2\,\tau\right)\tau.\end{aligned}$$

From equation (9.12), using the Nečas inequality (2.12), it is possible to deduce that

$$\begin{aligned}|h_{i-1}| &\leqslant |m'(t_i)| + \rho \left| \theta'_{i-1}\Big|_0^L \right| + \sum_{l=1}^{i} \tau |k_l| \left\| \theta''_{i-l} \right\|_{\mathrm{L}^1(\Omega)} \\ &\quad + \gamma \left\| \delta u'_{i-1} \right\|_{\mathrm{L}^1(\Omega)} + \| s_i \|_{\mathrm{L}^1(\Omega)} \\ &\leqslant C + \varepsilon \left\| \theta''_{i-1} \right\| + C_\varepsilon \left\| \theta'_{i-1} \right\| + C \sum_{l=0}^{i-1} \| \theta''_l \| \tau + C \left\| \delta u'_{i-1} \right\| + C \| s_i \| .\end{aligned}$$

Therefore, using Hölder's inequality, we get that

$$\begin{aligned}\sum_{i=1}^{j} |h_{i-1}|^2 \tau &\leqslant C + \varepsilon \sum_{i=1}^{j} \left\| \theta''_{i-1} \right\|^2 \tau + C_\varepsilon \sum_{i=1}^{j} \left\| \theta'_{i-1} \right\|^2 \tau \\ &\quad + C \sum_{i=1}^{j} \left(\sum_{l=0}^{i-1} \| \theta''_l \|^2 \tau \right) \tau + C \sum_{i=1}^{j} \left\| \delta u'_{i-1} \right\|^2 \tau + C \sum_{i=1}^{j} \| s_i \|^2 \tau. \end{aligned} \tag{9.20}$$

Collecting the previous estimates, fixing ε sufficiently small and applying Grönwall's argument, we get the estimates for u_i and θ_i. The estimate for h_i follows from equation (9.20). □

Lemma 9.3.4. *Let the assumptions of Lemma 9.3.2 be fulfilled. Then there exist positive constants C and τ_0 such that*

$$\max_{1 \leqslant j \leqslant n} \| u_j \| \leqslant C$$

for $0 < \tau < \tau_0$.

Proof. This follows easily from $u_j = u_0 + \sum_{i=1}^{j} \delta u_i \tau$ together with Lemma 9.3.3. □

Lemma 9.3.5. *Let the assumptions of Lemma 9.3.2 be fulfilled. Then there exist positive constants C and τ_0 such that*

$$\max_{1 \leqslant i \leqslant n} \left\| \delta^2 u_i \right\| \leqslant C \quad \textit{and} \quad \sum_{i=1}^{n} \| \delta \theta_i \|^2 \tau \leqslant C$$

for $0 < \tau < \tau_0$.

Proof. We multiply equation (9.15) with $\delta^2 u_i$ and (9.16) with $\delta \theta_i$. Then, we integrate the result over the domain Ω. For $i = 1, \ldots, n$, it holds that

$$\begin{aligned}\left\| \delta^2 u_i \right\|^2 &= \left(r_i, \delta^2 u_i \right) + \alpha \left(u''_i, \delta^2 u_i \right) - \gamma \left(\theta'_i, \delta^2 u_i \right), \\ \| \delta \theta_i \|^2 &= h_{i-1} \left(f, \delta \theta_i \right) + s_i + \rho \left(\theta''_i, \delta \theta_i \right) + \left(k * \theta''_i, \delta \theta_i \right) - \gamma \left(\delta u'_i, \delta \theta_i \right).\end{aligned}$$

Employing Young's inequality and Lemma 9.3.3, we get the asked estimates.

We can also use the variational formulation to prove this theorem. We point this out for the second estimate. It holds that $\|\delta\theta_i\| = \sup_{\|\psi\|\leqslant 1} (\delta\theta_i, \psi)$ for $i = 1, \ldots, n$. Using the variational formulation (9.10), we have that

$$(\delta\theta_i, \psi) = h_{i-1}\,(f, \psi) + (s_i, \psi) + \rho\,(\theta_i'', \psi) + \left(\sum_{l=1}^{i} k_l \theta_{i-l}''\tau, \psi\right) - \gamma\,(\delta u_i', \psi)\,.$$

Therefore, we deduce that

$$\begin{aligned}\|\delta\theta_i\| &\lesssim \sup_{\|\psi\|\leqslant 1} \left(1 + |h_{i-1}| + \|\theta_i''\| + \sum_{l=1}^{i} \left\|\theta_{i-l}''\right\| \tau + \|\delta u_i'\|\right) \|\psi\| \\ &\lesssim 1 + |h_{i-1}| + \|\theta_i''\| + \sum_{l=1}^{i} \left\|\theta_{i-l}''\right\| \tau + \|\delta u_i'\|\,.\end{aligned}$$

Using Hölder's inequality and Lemma 9.3.3, we get that

$$\|\delta\theta_i\| \lesssim 1 + |h_{i-1}| + \|\theta_i''\|\,. \tag{9.21}$$

We multiply the square of (9.21) with τ and sum the result up for $i = 1, \ldots, n$. Again, using Lemma 9.3.3, we obtain that

$$\sum_{i=1}^{n} \|\delta\theta_i\|^2 \tau \lesssim \sum_{i=1}^{n} \tau + \sum_{i=1}^{n} |h_{i-1}|^2 \tau + \sum_{i=1}^{n} \|\theta_i''\|^2 \tau \leqslant C.$$

□

The following piecewise linear in time functions $u_n : [0,T] \to \mathrm{L}^2(\Omega)$ and $v_n : [0,T] \to \mathrm{L}^2(\Omega)$

$$\begin{array}{llll} u_n(0) & = u_0, & & \\ u_n(t) & = u_{i-1} + (t - t_{i-1})\delta u_i & \text{for } t \in (t_{i-1}, t_i], & i = 1, \ldots, n; \\ v_n(0) & = \dot{u}_0, & & \\ v_n(t) & = \delta u_{i-1} + (t - t_{i-1})\delta^2 u_i & \text{for } t \in (t_{i-1}, t_i], & i = 1, \ldots, n \end{array}$$

and the piecewise constant in time functions $\overline{u}_n : [0,T] \to \mathrm{L}^2(\Omega)$ and $\overline{v}_n : [0,T] \to \mathrm{L}^2(\Omega)$

$$\begin{array}{llllll} \overline{u}_n(0) & = u_0, & \overline{u}_n(t) & = u_i, & \text{for } t \in (t_{i-1}, t_i], & i = 1, \ldots, n; \\ \overline{v}_n(0) & = \dot{u}_0, & \overline{v}_n(t) & = \delta u_i, & \text{for } t \in (t_{i-1}, t_i], & i = 1, \ldots, n \end{array}$$

are introduced. Similarly, the functions $\theta_n, \overline{\theta}_n, \overline{k}_n, \overline{r}_n, \overline{h}_n, \overline{m}_n$ and $\overline{\dot{m}}_n$ are defined. Note that $\overline{v}_n = \dot{u}_n$. Introduce also the notation $\lceil\cdot\rceil_\tau$ defined by $\lceil t\rceil_\tau = i$ when $t \in (t_{i-1}, t_i]$. Using these so-called Rothe functions, the variational formulation

(9.9–9.12) can be rewritten for all $\phi \in \mathrm{H}^1(\Omega)$ and $\psi \in \mathrm{H}^1_0(\Omega)$ and a.a. $t \in (0,T)$ as

$$(\dot{v}_n(t), \phi) + \alpha\left(\overline{u}_n'(t), \phi'\right) + \gamma\left(\overline{\theta}_n'(t), \phi\right) = \left(\overline{r}_n(t), \phi\right), \tag{9.22}$$

$$\left(\dot{\theta}_n(t), \psi\right) + \rho\left(\overline{\theta}_n'(t), \psi'\right) + \left(\sum_{l=1}^{\lceil t \rceil_\tau} \overline{k}_n(t_l)\overline{\theta}_n'(t-t_l)\tau, \psi'\right) \\ - \gamma\left(\dot{u}_n(t), \psi'\right) = \overline{h}_n(t-\tau)\left(f, \psi\right) + \left(\overline{s}_n(t), \psi\right), \tag{9.23}$$

with

$$\left(\int_0^L f\right)\overline{h}_n(t) = \overline{m}_n(t+\tau) - \rho\int_0^L \overline{\theta}_n''(t) \\ - \sum_{l=1}^{\lceil t+\tau \rceil_\tau} \tau\overline{k}_n(t_l)\int_0^L \overline{\theta}_n''(t+\tau-t_l) + \gamma\int_0^L \dot{u}_n'(t) - \int_0^L \overline{s}_n(t+\tau). \tag{9.24}$$

The a priori estimates in Lemma 9.3.3 and 9.3.4 in the new notations read as

$$\max_{t\in[0,T]}\left\{\left\|\overline{u}_n(t)\right\|^2_{\mathrm{H}^2(\Omega)} + \left\|\dot{u}_n(t)\right\|^2_{\mathrm{H}^1(\Omega)} + \left\|\overline{\theta}_n(t)\right\|^2_{\mathrm{H}^1(\Omega)} + \left\|\dot{v}_n(t)\right\|^2\right\} \\ + \int_0^T \left\|\overline{\theta}_n''(t)\right\|^2 \mathrm{d}t + \int_0^T \left\|\dot{\theta}_n(t)\right\|^2 \mathrm{d}t + \sum_{i=1}^n \left\|\int_{t_{i-1}}^{t_i} \dot{u}_n(t)\,\mathrm{d}t\right\|^2_{\mathrm{H}^2(\Omega)} \\ + \sum_{i=1}^n \left\|\int_{t_{i-1}}^{t_i} \dot{\theta}_n(t)\,\mathrm{d}t\right\|^2_{\mathrm{H}^1(\Omega)} + \sum_{i=1}^n \left\|\int_{t_{i-1}}^{t_i} \dot{v}_n(t)\,\mathrm{d}t\right\|^2_{\mathrm{H}^1(\Omega)} \leqslant C$$

and

$$\int_0^T \left|\overline{h}_n(t-\tau)\right|^2 \mathrm{d}t \leqslant C. \tag{9.25}$$

Now, the convergence of the sequences $\{u_n\}$, $\{\overline{u}_n\}$, $\{\theta_n\}$ and $\{\overline{\theta}_n\}$ to the unique weak solution of (9.1)-(9.2) is proved as $\tau \to 0$ or $n \to \infty$.

Theorem 9.3.2 (Existence)**.** *Suppose that the conditions of Lemma 9.3.2 are fulfilled. Then there exists a triplet $\langle u, \theta, h\rangle$ such that*

(i) $u_n \to u$ *in* $\mathrm{C}\left([0,T], \mathrm{H}^1(\Omega)\right)$, $\overline{u}_n \to u$ *in* $\mathrm{L}^2\left((0,T), \mathrm{H}^1(\Omega)\right)$ *and* $\dot{u}_n \rightharpoonup \dot{u}$ *in* $\mathrm{L}^2\left((0,T), \mathrm{H}^1(\Omega)\right)$*;*

(ii) $\theta_n, \overline{\theta}_n \to \theta$ *in* $\mathrm{L}^2\left((0,T), \mathrm{H}^1(\Omega)\right)$ *and* $\dot{\theta}_n \rightharpoonup \dot{\theta}$ *in* $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$*;*

(iii) $v_n \rightharpoonup \dot{u}$ *in* $\mathrm{L}^2\left((0,T), \mathrm{H}^1(\Omega)\right)$ *and* $\dot{v}_n \rightharpoonup \ddot{u}$ *in* $\mathrm{L}^2\left((0,T), \mathrm{L}^2(\Omega)\right)$*;*

(iv) $\overline{h}_n(t-\tau) \rightharpoonup z$ *in* $\mathrm{L}^2(0,T)$*;*

(v) $\overline{\theta}''_n \rightharpoonup \theta''$ *in* $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$*;*

(vi) $\int_0^\eta \overline{\theta}'_n(0,t-\tau)\,\mathrm{d}t \to \int_0^\eta \theta'(0,t)\,\mathrm{d}t$ *and* $\int_0^\eta \theta'_n(L,t-\tau)\,\mathrm{d}t \to \int_0^\eta \theta'(L,t)\,\mathrm{d}t$*,* $\forall \eta \in [0,T]$*;*

(vii) $\int_0^\eta \int_0^L \overline{\theta}''_n(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t \to \int_0^\eta \int_0^L \theta''_n(x,t)\,\mathrm{d}x\,\mathrm{d}t$*,* $\forall \eta \in [0,T]$*;*

(viii) $\int_0^\eta \int_0^L \dot{u}'_n(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t \to \int_0^\eta \int_0^L \dot{u}'(x,t)\,\mathrm{d}x\,\mathrm{d}t$*,* $\forall \eta \in [0,T]$*;*

(ix) $\int_0^\eta \overline{h}_n(t-\tau)\,\mathrm{d}t \to \int_0^\eta h(t)\,\mathrm{d}t$*,* $\forall \eta \in [0,T]$*, i.e.* $z = h$ *in* $\mathrm{L}^2(0,T)$*;*

(x) $\langle u,\theta,h\rangle \in \left[\mathrm{C}\left([0,T],\mathrm{C}(\overline{\Omega})\right) \cap \mathrm{L}^\infty\left((0,T),\mathrm{C}^1(\overline{\Omega})\right)\right] \times \left[\mathrm{C}\left([0,T],\mathrm{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T),\mathrm{C}^1(\overline{\Omega})\right)\right] \times \mathrm{L}^2(0,T)$ *is a weak solution to* (9.1)-(9.2).

Proof. The proof of convergence is split up into several steps.

(i) Thanks to Theorem 2.9.23, we have that

$$\mathrm{H}^2(\Omega) \hookrightarrow\hookrightarrow \mathrm{H}^1(\Omega).$$

Due to

$$\max_{t\in[0,T]} \left\{ \|\overline{u}_n(t)\|^2_{\mathrm{H}^2(\Omega)} + \|\dot{u}_n(t)\|^2_{\mathrm{H}^1(\Omega)} \right\} \leqslant C,$$

we have that the conditions of Lemma 2.12.3 are satisfied. Therefore, there exists $u \in \mathrm{C}\left([0,T],\mathrm{H}^1(\Omega)\right) \cap \mathrm{L}^\infty\left((0,T),\mathrm{H}^2(\Omega)\right)$ and a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$\begin{cases} u_{n_k} \to u, & \text{in} \quad \mathrm{C}\left([0,T],\mathrm{H}^1(\Omega)\right), \\ u_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{H}^2(\Omega) \text{ for all } t \in [0,T], \\ \overline{u}_{n_k}(t) \rightharpoonup u(t), & \text{in} \quad \mathrm{H}^2(\Omega) \text{ for all } t \in [0,T], \\ \dot{u}_{n_k} \rightharpoonup \dot{u}, & \text{in} \quad \mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right). \end{cases} \tag{9.26a}$$

We denote this subsequence again with $\{u_n\}$ to skip double indices. Moreover, $u : [0,T] \to \mathrm{H}^1(\Omega)$ is Lipschitz continuous. For every $t \in (t_{i-1},t_i]$, it holds that

$$\|u_n(t) - \overline{u}_n(t)\|^2_{\mathrm{H}^1(\Omega)} = \|(t-t_i)\delta u_i\|^2_{\mathrm{H}^1(\Omega)} \leqslant \tau^2 \|\dot{u}_n(t)\|^2_{\mathrm{H}^1(\Omega)} \lesssim \tau^2,$$

i.e. $\{u_n\}$ and $\{\overline{u}_n\}$ have the same limit in $\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$.

(ii) The Rellich-Kondrachov theorem 2.9.22 implies that

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

We can use the generalized Aubin-Lions lemma 2.12.4 because θ_n and $\overline{\theta}_n \in \mathrm{L}^2\left((0,T),\mathrm{H}^2(\Omega)\right)$, and $\dot{\theta}_n \in \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$. Therefore, there exists a function $\theta \in \mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$ and a subsequence $\{\theta_{n_k}\}_{k\in\mathbb{N}}$ of $\{\theta_n\}_{n\in\mathbb{N}}$ such that

$$\begin{cases} \theta_{n_k} \to \theta, & \text{in} \quad \mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right), & (9.27\text{a}) \\ \theta_{n_k} \rightharpoonup \theta, & \text{in} \quad \mathrm{L}^2\left((0,T),\mathrm{H}^2(\Omega)\right), & (9.27\text{b}) \\ \dot{\theta}_{n_k} \rightharpoonup \dot{\theta}, & \text{in} \quad \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right), \end{cases}$$

which we denote again by $\{\theta_n\}$ for ease of reading. Applying Lemma 2.9.5(i), we get $\theta \in \mathrm{C}\left([0,T],\mathrm{L}^2(\Omega)\right)$ because $\theta \in \mathrm{L}^2\left((0,T),\mathrm{H}^2(\Omega)\right)$ and $\dot{\theta} \in \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$. Applying Lemma 9.3.3, we obtain that

$$\int_0^T \left\|\theta_n(t) - \overline{\theta}_n(t)\right\|^2_{\mathrm{H}^1(\Omega)} \mathrm{d}t \leqslant \tau \sum_{i=1}^n \|\theta_i - \theta_{i-1}\|^2_{\mathrm{H}^1(\Omega)} \lesssim \tau,$$

i.e. $\{\theta_n\}$ and $\{\overline{\theta}_n\}$ have the same limit in $\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$.

(iii) From (i), we immediately get that $\overline{v}_n = \dot{u}_n \rightharpoonup \dot{u}$ in $\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$. Employing Lemma 9.3.3, we see that

$$\int_0^T \|v_n(t) - \overline{v}_n(t)\|^2_{\mathrm{H}^1(\Omega)} \mathrm{d}t \leqslant \tau \sum_{i=1}^n \|\delta u_i - \delta u_{i-1}\|^2_{\mathrm{H}^1(\Omega)} \lesssim \tau,$$

i.e. $\{v_n\}$ and $\{\overline{v}_n\}$ have the same limit in $\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$. From the boundedness of $\dot{v}_n$ in the reflexive space $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$ follows that $\dot{v}_n \rightharpoonup \ddot{u}$ in $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$.

(iv) This follows from (9.25) and the reflexivity of $\mathrm{L}^2(0,T)$.

(v) From $\int_0^T \left\|\overline{\theta}''_n(t)\right\|^2 \mathrm{d}t \leqslant C$ and the reflexivity of $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$, we get that $\overline{\theta}''_n \rightharpoonup w$ in $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$. Using the Green theorem and (ii), we obtain for all $\phi \in \mathrm{C}_0^\infty(\overline{\Omega})$ that

$$\begin{array}{ccccc} \displaystyle\int_0^\eta \left(\overline{\theta}''_n(t),\phi\right)\mathrm{d}t & = & \displaystyle -\int_0^\eta \left(\overline{\theta}'_n(t),\phi'\right) & & \\ \downarrow & & \downarrow & & \\ \displaystyle\int_0^\eta (w(t),\phi)\,\mathrm{d}t & & \displaystyle -\int_0^\eta (\theta',\phi') & = & \displaystyle\int_0^\eta (\theta'',\phi)\,. \end{array}$$

From the density argument $\overline{\mathrm{C}_0^\infty(\overline{\Omega})} = \mathrm{L}^2(\Omega)$, it follows that $\int_0^\eta (w(t),\phi)\,\mathrm{d}t = \int_0^\eta (\theta'',\phi)$ for all $\phi \in \mathrm{L}^2(\Omega)$. Therefore, $w = \theta''$ in $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$.

(vi) Using the Nečas inequality, $\int_0^T \left\|\overline{\theta}_n''(t)\right\|^2 \mathrm{d}t \leqslant C$ and (9.27b), we obtain for every $\eta \in [0,T]$ that

$$\begin{aligned}
&\int_0^\eta \left|\overline{\theta}_n'(0,t-\tau) - \theta'(0,t)\right|^2 \mathrm{d}t \\
&\leqslant \quad \varepsilon \int_0^\eta \left\|\overline{\theta}_n''(t-\tau) - \theta''(t)\right\|^2 \mathrm{d}t + C_\varepsilon \int_0^\eta \left\|\overline{\theta}_n'(t-\tau) - \theta'(t)\right\|^2 \mathrm{d}t \\
&\leqslant \quad \varepsilon + C_\varepsilon \int_0^\eta \left\|\overline{\theta}_n'(t-\tau) \pm \overline{\theta}_n'(t) - \theta'(t)\right\|^2 \mathrm{d}t. \qquad (9.28)
\end{aligned}$$

As before, we have that

$$\int_0^\eta \left\|\overline{\theta}_n'(t-\tau) - \overline{\theta}_n'(t)\right\|^2 \mathrm{d}t \leqslant \tau \sum_{i=1}^n \left\|\theta_i' - \theta_{i-1}'\right\|^2 \lesssim \tau.$$

Employing (9.27a), we get that

$$\int_0^\eta \left\|\overline{\theta}_n'(t) - \theta'(t)\right\|^2 \mathrm{d}t \to 0 \quad \text{as} \quad n \to \infty.$$

Passing to the limit for $\tau \to 0$ in (9.28), it holds that

$$\lim_{\tau\to 0} \int_0^\eta \left|\overline{\theta}_n'(0,t-\tau) - \theta'(0,t)\right|^2 \mathrm{d}t \leqslant \varepsilon,$$

which is valid for any small $\varepsilon > 0$. Hence,

$$\lim_{\tau\to 0} \int_0^\eta \left|\overline{\theta}_n'(0,t-\tau) - \theta'(0,t)\right|^2 \mathrm{d}t = 0.$$

Analogously, one can prove that

$$\lim_{\tau\to 0} \int_0^\eta \left|\overline{\theta}_n'(L,t-\tau) - \theta'(L,t)\right|^2 \mathrm{d}t = 0.$$

(vii) We may write

$$\int_0^\eta \int_0^L \overline{\theta}_n''(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t = \int_0^\eta \left(\overline{\theta}_n'(L,t-\tau) - \overline{\theta}_n'(0,t-\tau)\right) \mathrm{d}t.$$

Using (vi), we obtain for $\tau \to 0$ that

$$\begin{aligned}
&\int_0^\eta \int_0^L \overline{\theta}_n''(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad \to \int_0^\eta \left(\theta'(L,t) - \theta'(0,t)\right) \mathrm{d}t = \int_0^\eta \int_0^L \theta''(x,t)\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

(viii) We easily see that

$$\begin{aligned}\left|\int_0^\eta \int_0^L \dot{u}_n'(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t - \int_0^\eta \int_0^L \dot{u}'(x,t)\,\mathrm{d}x\,\mathrm{d}t\right| \\ \leqslant \left|\int_0^\eta \int_0^L \dot{u}_n'(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t - \int_0^\eta \int_0^L \dot{u}_n'(x,t)\,\mathrm{d}x\,\mathrm{d}t\right| \\ + \left|\int_0^\eta \int_0^L \dot{u}_n'(x,t)\,\mathrm{d}x\,\mathrm{d}t - \int_0^\eta \int_0^L \dot{u}'(x,t)\,\mathrm{d}x\,\mathrm{d}t\right|.\end{aligned}$$

Both terms in the RHS of this inequality converge to zero as $\tau \to 0$. For the first term, we have by Lemma 9.3.3 that

$$\begin{aligned}\left|\int_0^\eta \left(\dot{u}_n'(x,t-\tau) - \dot{u}_n'(x,t), 1\right)\mathrm{d}x\,\mathrm{d}t\right| &\leqslant \left|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \left(\delta u_{i-1}' - \delta u_i', 1\right)\right| \\ &\lesssim \sqrt{\tau}\sqrt{\sum_{i=1}^n \left\|\delta u_{i-1}' - \delta u_i'\right\|^2} \lesssim \sqrt{\tau}.\end{aligned}$$

The second term converges to zero by (9.26a).

(ix) We integrate (9.24) at $t-\tau$ in time over $(0,\eta) \subset (0,T)$ and obtain

$$\begin{aligned}\left(\int_0^L f\right)\int_0^\eta \overline{h}_n(t-\tau)\,\mathrm{d}t = \int_0^\eta \overline{\dot{m}}_n(t)\,\mathrm{d}t - \rho\int_0^\eta\int_0^L \overline{\theta}_n''(x,t-\tau)\,\mathrm{d}x\,\mathrm{d}t \\ - \int_0^\eta\left(\sum_{l=1}^{\lceil t\rceil_\tau} \tau\overline{k}_n(t_l)\int_0^L \overline{\theta}_n''(x,t-t_l)\,\mathrm{d}x\right)\mathrm{d}t \\ + \gamma\int_0^\eta\left(\int_0^L \dot{u}_n'(x,t-\tau)\,\mathrm{d}x\right)\mathrm{d}t - \int_0^\eta\left(\int_0^L \overline{s}_n(x,t)\,\mathrm{d}x\right)\mathrm{d}t. \quad (9.29)\end{aligned}$$

It is clear that $\overline{\dot{m}}_n \to \dot{m}$ in $\mathrm{L}^2([0,T])$, $\overline{k}_n \to k$ in $\mathrm{L}^2([0,T])$, $\overline{s}_n \to s$ in $\mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$ and $r_n \to r$ in $\mathrm{L}^2\left((0,T),\mathrm{H}^1(\Omega)\right)$ because m, k, s and r are prescribed data functions. We pass to the limit for $\tau \to 0$ in (9.29). Note that

$$\begin{aligned}&\sum_{l=1}^{\lceil t\rceil_\tau} \tau\overline{k}_n(t_l)\int_0^L \overline{\theta}_n''(x,t-t_l) \\ &= \left(\overline{k}_n * \int_0^L \overline{\theta}_n''\right)(t) + \int_t^{\tau\lceil t\rceil_\tau} \overline{k}_n(s)\int_0^L \overline{\theta}_n''(x,t-s)\,\mathrm{d}x\ \mathrm{d}s \\ &= \left(\overline{k}_n * \int_0^L \overline{\theta}_n''\right)(t) + \mathcal{O}(\tau).\end{aligned}$$

Employing (iv), (v), (vii) and (viii), we obtain

$$\left(\int_0^L f\right)\int_0^\eta z(t) = \int_0^\eta \dot{m}(t) - \rho\int_0^\eta\int_0^L \theta''(x,t)$$
$$-\int_0^\eta\left(k*\int_0^L\theta''\right)(t) + \gamma\int_0^\eta\int_0^L \dot{u}'(x,t) - \int_0^\eta\int_0^L s(x,t). \quad (9.30)$$

By (9.3), we get that $\int_0^\eta z(t) = \int_0^\eta h(t)$. Finally, we differentiate (9.30) with respect to η and we arrive at (9.3).

(x) Note that

$$\sum_{l=1}^{\lceil t\rceil_\tau}\overline{k}_n(t_l)\overline{\theta}_n'(t-t_l)\tau = \left(\overline{k}_n * \overline{\theta}_n'\right)(t) + \int_t^{\tau\lceil t\rceil_\tau}\overline{k}_n(s)\overline{\theta}_n'(t-s)\ \mathrm{d}s$$
$$= \left(\overline{k}_n * \overline{\theta}_n'\right)(t) + \mathcal{O}(\tau).$$

We integrate (9.22) and (9.23) in time and pass to the limit for $\tau \to 0$ using (i), (ii), (iii) and (ix). We differentiate the result with respect to the time variable to arrive at (9.4)-(9.5). The convergences of Rothe's functions towards the weak solution have been shown for a subsequence. However, taking into account Theorem 9.2.1, the whole Rothe's sequence converges towards the solution. Note that by [39, Theorem 3.5], $\mathrm{H}^1(\Omega) \subset \mathrm{C}(\overline{\Omega})$ and $\mathrm{H}^2(\Omega) \subset \mathrm{C}^1(\overline{\Omega})$. □

9.4 Numerical examples

The aim of the simulations is to demonstrate the proposed numerical scheme from the previous section for the recovery of the unknown heat source $h(t)$ in the case of type-I thermoelasticity ($\rho \neq 0, k = 0$) with $\rho = \gamma = \alpha = 1$.

In each experiment, the domain $\Omega = (0,1)$ and the final time $T = 1$. The number of time discretization interval is chosen to be 2^6, 2^7 and 2^9. At each time-step, the resulting elliptic problems are solved numerically by the finite element method (FEM) using first order (P1-FEM) Lagrange polynomials for the space discretization for which a fixed uniform mesh of 2000 subintervals is used. The displacement and temperature are the same in each experiment:

$$u_{\mathrm{ex}}(x,t) = \left(t^2+t+1\right)(1+\cos(\pi x))$$

and

$$\theta_{\mathrm{ex}}(x,t) = 1 + 2\left(t^2+1\right)x(1-x).$$

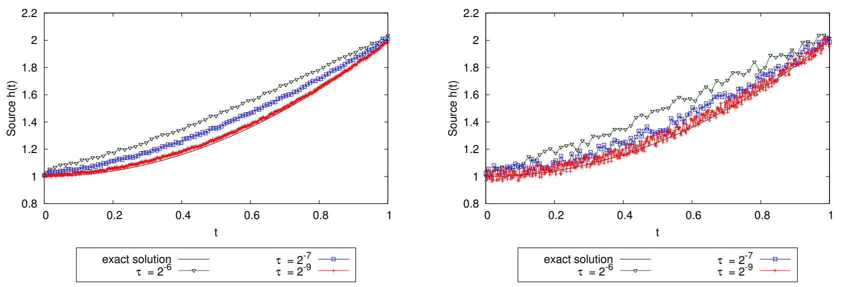
The additional measurement is given by

$$m(t) = \frac{4}{3} + \frac{1}{3}t^2.$$

An uncorrelated noise is added to $m'(t) = \frac{2}{3}t$ in order to simulate the errors present in real measurements. This noise is generated randomly with given magnitude of 1% and 5%. Two numerical experiments are developped depending on the choice of the unknown heat source

$$h_{\mathrm{ex}}(t) = 1 + t^2 \quad \text{or} \quad h_{\mathrm{ex}}(t) = \sin(2\pi t).$$

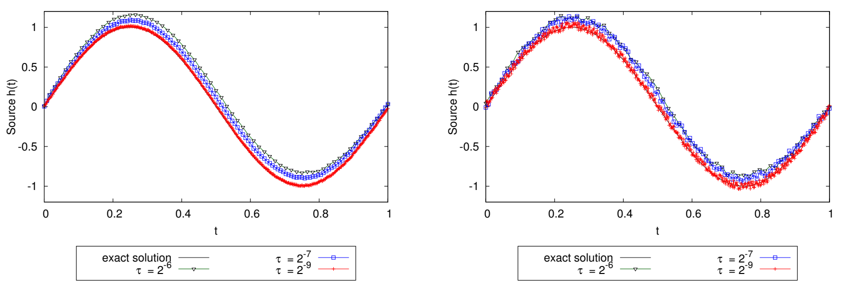
The results of these experiments are depicted in Figures 9.1–9.2. The experiments show that good approximations are obtained when the timestep is small enough even if the noise is larger.



(a) Exact solution $h_{\mathrm{ex}}(t) = 1 + t^2$ and its numerical approximations (noise with magnitude 1%)

(b) Exact solution $h_{\mathrm{ex}}(t) = 1 + t^2$ and its numerical approximations (noise with magnitude 5%)

Figure 9.1: Unknown heat source in thermoelasticity: results of Experiment 1



(a) Exact solution $h_{\mathrm{ex}}(t) = \sin(2\pi t)$ and its numerical approximations (noise with magnitude 1%)

(b) Exact solution $h_{\mathrm{ex}}(t) = \sin(2\pi t)$ and its numerical approximations (noise with magnitude 5%)

Figure 9.2: Unknown heat source in thermoelasticity: results of Experiment 2

9.5 Conclusion

A thermoelastic system of type-III in 1D with an unknown solely time-dependent heat source has been considered. The missing heat source has been recovered from a measurement in time of the average temperature inside the body. Using this observation, the inverse problem has been reformulated in a direct setting. The existence and uniqueness of a weak solution has been addressed and a new numerical algorithm based on Rothe's method has been designed. The convergence of the numerical scheme is demonstrated by means of some numerical experiments.

10

Conclusions and perspectives for further research

The present dissertation is a study on numerical techniques for partial differential equations arising in superconductivity and in thermoelasticity. This chapter summarizes and discusses the results presented in the previous chapters. In Chapter 1, two general goals were formulated. Section 10.1 and Section 10.2 discuss and evaluate Goal I, respectively Goal II of this study. Also future research directions are included.

10.1 Goal I

The first general goal of this study was to present mathematical models for nonlocal superconductivity and to analyse these models using Rothe's method.

In Chapter 3, three new macroscopic models for nonlocal superconductivity have been introduced: a parabolic and hyperbolic model for type-I superconductivity and a model for an intermediate state between type-I and type-II superconductivity. These models have the magnetic field as unknown function and they have been studied in the first part of this dissertation.

A vectorial nonlocal linear parabolic problem (4.1) with applications in superconductors of type-I has been studied in Chapter 4. This model was obtained from the eddy current version of the Maxwell equations, the two-fluid model of London and London, and the nonlocal representation of the superconductive current by Eringen. The nonlocal term in this model is given by a space convolution with a singular kernel. The problem is standard, except of the appearance of the convolution term. The existence of a unique solution has been proved using Rothe's method.

Two time-discrete numerical schemes based on backward Euler method have been developed to approximate the solution to the parabolic problem. In the first scheme, the convolution has been taken implicitly (from the actual time step). In the second one, the convolution has been taken explicitly (from the previous time step). This second scheme was considered because it is easier to implement than the first one and it gives the same order of convergence.

For both schemes, error estimates for the time discretization have been obtained using a priori estimates, which are based on Grönwall's argument. The convergence rates are of (optimal) order $\mathcal{O}(\tau) = e^{CT}\tau$ in the space $\mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T), \mathbf{H}(\mathbf{curl};\Omega)\right)$ under appropriate conditions, where τ is the discretization parameter. To get rid of the exponential (in time) character of this constant, a new convolution kernel has been derived under the assumption that the normal component of the unknown vector field equals zero on the boundary of the superconductor. With the aid of the additional assumption, it has been demonstrated that under higher regularity the solution of the original model satisfies a simpler problem which is easier to implement. Both time-discrete schemes stayed valid. One major advantage of working with this simpler model is the positive definiteness of the kernel. Using this property, better error estimates of order $\mathcal{O}(\tau) = C\tau$ have been obtained for the implicit scheme.

A numerical experiment for the semi-implicit scheme supported the obtained theoretical results. The time-discrete problems have been solved using the finite element method. The convolution integral has been approximated by a space-discrete convolution in such a way that the singularity in the kernel has been avoided. Also convergence of a fully discrete finite element scheme to a solution of the problem has been shown. In a similarly way to the time-discrete schemes, it has been demonstrated how to improve the error estimates under higher regularity.

An analogue working scheme has been followed for the hyperbolic model. This model was derived from the full Maxwell equations instead of the eddy current version. The well-posedness of this model has been addressed in Chapter 5.

Two time-discrete schemes (based on an explicit and implicit handling of the convolution term) have been established. The error estimates have been derived for both schemes. As in the parabolic case, the solution of the original model satisfied a simpler problem under the assumption that the normal component of the unknown vector field equals zero on the boundary of the superconductor. No better error estimates for the time discretization have been obtained despite the positive definiteness of the kernel. This model has not yet been computationally implemented.

In Chapter 6, a vectorial nonlocal nonlinear parabolic problem for an intermediate state between type-I and type-II superconductivity has been analysed. This model was obtained from the eddy current version of the Maxwell equations, the two-fluid model of London and London, the nonlocal representation (by a space convolution with a singular kernel) of the superconductive current by Eringen for type-I superconductivity and the power law by Rhyner for type-II superconductivity. A semi-implicit time-discrete scheme based on the backward Euler method in which the convolution is taken explicitly has been developed. The well-posedness of the problem has been shown under low regularity assumptions and suboptimal error estimates have been derived for the time-discretization. No numerical experiments have been presented.

An interesting area for future research is the further implementation of the different models (with correct physical parameters) and their related schemes. In particular for the parabolic model, one can focus on the comparison of the error estimates for both schemes. Possibly, also a comparison can be made between the results of the nonlocal parabolic and hyperbolic model. Moreover, the validity of the proposed models should be checked experimentally (in particular for the intermediate model) and the results should be compared with available results from physics. This would give more support to the findings in this study.

10.2 Goal II

The second goal of this study was to recover unknown sources in thermoelastic systems from additional data.

A thermoelastic system consists of two equations that are coupled: a parabolic (heat) equation and a vectorial hyperbolic equation for the displacement. Green and Naghdi developed three theories of thermoelasticity, each with a corresponding system: type-I, type-II and type-III thermoelasticity. In the analysis presented in this thesis, an isotropic and homogeneous thermoelastic body has been considered. Two inverse source problems for thermoelasticity have been discussed.

In Chapter 8, the determination of a space-dependent vector source in a thermoelastic system of type-I, type-II and type-III has been studied using information from a final in time measurement of the displacement. Using a variational approach, the uniqueness of a solution to the inverse problem has been proved when a damping term $\mathbf{g}(\partial_t u)$ is added in the hyperbolic equation of the classic thermoelasticity system. The main assumption was that $\mathbf{g}$ is componentwise strictly monotone increasing.

Landweber's regularization method has been applied to cope with the ill-posedness of the inverse problem. In the case that the damping term is linear, a stable iterative algorithm has been proposed to recover the unknown source. This algorithm is based on a sequence of well-posed (direct) problems, which are solved at each iteration step using the finite element method. The instability has been overcome by stopping the iterations at the first iteration for which the discrepancy principle of Morozov is satisfied. The convergence of the algorithm has been illustrated by numerical experiments.

In Chapter 9, a classic thermoelastic system of type-III in 1D with an unknown solely time-dependent heat source has been considered. The missing heat source has been recovered from a measurement in time of the average temperature inside the rod. The proposed technique differs from the standard approaches for inverse source problems. Using the measurement, the inverse problem has been reformulated in a direct setting. This is a new technique, which can also be applied to inverse problems in other settings. The existence and uniqueness of a weak solution has been addressed and a new numerical algorithm based on Rothe's method is designed. The results are also valid for type-I thermoelasticity. The convergence of the numerical scheme is demonstrated by means of some numerical experiments.

Before closing this study, some suggestions for further research are offered. In the case of the unknown space-dependent vector source, future research can concern an extension of the results to general thermoelastic systems or to more general (Robin) boundary conditions. Moreover, the main trick in obtaining a unique solution fails when applying the same technique if a space-dependent heat source is unknown. This gives rise to a research question in that direction: to establish a numerical scheme to recover a solely space-dependent heat source in thermoelastic systems.

In the case of the unknown time-dependent heat source, the focus can be put on the multi-dimensional case and on the recovery of a solely time-dependent vector source. Another research direction is the recovery of the solely time-dependent convolution kernel in a thermoelastic system of type-III.

In both cases, it can be interesting to consider Tikhonov regularization to retrieve the unknown sources such that the results obtained with the different methods can be compared. The existence and uniqueness questions can be repeated when considering other additional measurements in relation with the unknown sources.

A
Mathematical background

A.1 Some proofs

Lemma A.1.1. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$*, it holds that*

(i)

$$|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{y}|_{\mathrm{e}}^2 + \mathbf{x} \cdot \mathbf{y} \geqslant C_* \, |\mathbf{x} - \mathbf{y}|_{\mathrm{e}}^2 , \qquad C_* \in \left[-\frac{1}{2}, \frac{1}{4} \right],$$

(ii)

$$\left(|\mathbf{x}|_{\mathrm{e}}^{\beta-1} \, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1} \, \mathbf{y} \right) \cdot (\mathbf{x} - \mathbf{y}) \geqslant \frac{1}{4 \cdot 12^{\frac{\beta+1}{2}}} \, |\mathbf{x} - \mathbf{y}|_{\mathrm{e}}^{\beta+1}, \quad \beta \in [1, +\infty),$$

(iii)

$$\left(|\mathbf{x}|_{\mathrm{e}}^{\beta-1} \, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1} \, \mathbf{y} \right) \cdot \left(|\mathbf{x}|_{\mathrm{e}}^{\alpha-1} \, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\alpha-1} \, \mathbf{y} \right) \geqslant \frac{4\alpha\beta}{(\alpha+\beta)^2} \left(|\mathbf{x}|_{\mathrm{e}}^{\frac{\alpha+\beta}{2}} - |\mathbf{y}|_{\mathrm{e}}^{\frac{\alpha+\beta}{2}} \right)^2 \geqslant 0,$$

for $\alpha, \beta \in [0, +\infty)$*,*

(iv) there exists a positive constant C *such that*

$$\left| |\mathbf{x}|_{\mathrm{e}}^{\beta-1} \, \mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1} \, \mathbf{y} \right|_{\mathrm{e}} \leqslant C \, |\mathbf{x} - \mathbf{y}|_{\mathrm{e}}^{\beta}, \qquad \beta \in (0, 1].$$

Proof. (i) For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, it holds that

$$|\mathbf{x} + \mathbf{y}|_{\mathrm{e}}^2 \geqslant 0.$$

Therefore,

$$|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{y}|_{\mathrm{e}}^2 \geqslant -2\mathbf{x}\cdot\mathbf{y}.$$

Moreover, for $p \geqslant q \geqslant 0$, we get that

$$p\left[|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{y}|_{\mathrm{e}}^2\right] \geqslant -2q\mathbf{x}\cdot\mathbf{y}. \tag{A.1}$$

The asked lower bound can be rewritten in this form. Indeed,

$$|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{y}|_{\mathrm{e}}^2 + \mathbf{x}\cdot\mathbf{y} \geqslant C_* |\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^2 = C_*\left[|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{x}|_{\mathrm{e}}^2 - 2\mathbf{x}\cdot\mathbf{y}\right]$$

is equivalent with

$$(1-C_*)\left[|\mathbf{x}|_{\mathrm{e}}^2 + |\mathbf{y}|_{\mathrm{e}}^2\right] \geqslant -(1+2C_*)\mathbf{x}\cdot\mathbf{y}.$$

Employing (A.1), this is only valid if $1 - C_* \geqslant \frac{1}{2} + C_* \geqslant 0$ or $C_* \in [-\frac{1}{2}, \frac{1}{4}]$.

(ii) The proof is the corrected version of [186, Lemma 4.4]. Take $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Denote $\mathbf{x}\cdot\mathbf{y}$ as $\langle \mathbf{x}, \mathbf{y}\rangle$. Introducing a parameter $s \in [0,1]$ gives that

$$\begin{aligned}
&\langle |\mathbf{x}|_{\mathrm{e}}^{\beta-1}\mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1}\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle \\
&= \langle \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\left(|s\mathbf{x} + (1-s)\mathbf{y}|_{\mathrm{e}}^{\beta-1}\left[s\mathbf{x} + (1-s)\mathbf{y}\right]\right)\,\mathrm{d}s, \mathbf{x}-\mathbf{y}\rangle \\
&= (\beta-1)\int_0^1 |s\mathbf{x} + (1-s)\mathbf{y}|_{\mathrm{e}}^{\beta-3}\left(\langle s\mathbf{x} + (1-s)\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle\right)^2\,\mathrm{d}s \\
&\qquad + \int_0^1 |s\mathbf{x} + (1-s)\mathbf{y}|_{\mathrm{e}}^{\beta-1}|\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^2\,\mathrm{d}s.
\end{aligned}$$

Due to $\beta \geqslant 1$, we have that

$$\langle |\mathbf{x}|_{\mathrm{e}}^{\beta-1}\mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1}\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle \geqslant |\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^2 \int_0^1 |\mathbf{x} - (1-s)(\mathbf{x}-\mathbf{y})|_{\mathrm{e}}^{\beta-1}\,\mathrm{d}s.$$

Now, a distinction is made between $|\mathbf{x}|_{\mathrm{e}} \geqslant |\mathbf{x}-\mathbf{y}|_{\mathrm{e}}$ and $|\mathbf{x}|_{\mathrm{e}} < |\mathbf{x}-\mathbf{y}|_{\mathrm{e}}$. Firstly, when $|\mathbf{x}|_{\mathrm{e}} \geqslant |\mathbf{x}-\mathbf{y}|_{\mathrm{e}}$, then

$$\begin{aligned}
|\mathbf{x} - (1-s)(\mathbf{x}-\mathbf{y})|_{\mathrm{e}} &\geqslant \left||\mathbf{x}|_{\mathrm{e}} - (1-s)|\mathbf{x}-\mathbf{y}|_{\mathrm{e}}\right| \\
&\geqslant s|\mathbf{x}-\mathbf{y}|_{\mathrm{e}}.
\end{aligned}$$

We obtain that

$$\begin{aligned}
&\langle |\mathbf{x}|_{\mathrm{e}}^{\beta-1}\mathbf{x} - |\mathbf{y}|_{\mathrm{e}}^{\beta-1}\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle \\
&\qquad \geqslant |\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^2 \int_0^1 s^{\beta-1}|\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^{\beta-1}\,\mathrm{d}s = \frac{1}{\beta}|\mathbf{x}-\mathbf{y}|_{\mathrm{e}}^{\beta+1}.
\end{aligned}$$

Secondly, when $|\mathbf{x}|_e < |\mathbf{x}-\mathbf{y}|_e$, then using Jensen's inequality (see Lemma 2.3.10), we get that

$$\begin{aligned}
\langle |\mathbf{x}|_e^{\beta-1}\mathbf{x} - |\mathbf{y}|_e^{\beta-1}\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle &\geqslant |\mathbf{x}-\mathbf{y}|_e^2 \int_0^1 \frac{|\mathbf{x}-(1-s)(\mathbf{x}-\mathbf{y})|_e^{\beta+1}}{|\mathbf{x}-(1-s)(\mathbf{x}-\mathbf{y})|_e^2}\,\mathrm{d}s \\
&\geqslant |\mathbf{x}-\mathbf{y}|_e^2 \int_0^1 \frac{\left(|\mathbf{x}-(1-s)(\mathbf{x}-\mathbf{y})|_e^2\right)^{\frac{\beta+1}{2}}}{(2-s)^2\,|\mathbf{x}-\mathbf{y}|_e^2}\,\mathrm{d}s \\
&\geqslant \frac{1}{4}\left(\int_0^1 |s\mathbf{x}+(1-s)\mathbf{y}|_e^2\,\mathrm{d}s\right)^{\frac{\beta+1}{2}} \\
&\geqslant \frac{1}{4}\frac{1}{3^{\frac{\beta+1}{2}}}\left(|\mathbf{x}|_e^2 + \langle \mathbf{x},\mathbf{y}\rangle + |\mathbf{y}|_e^2\right)^{\frac{\beta+1}{2}} \\
&\overset{(*)}{\geqslant} \frac{1}{4\cdot 12^{\frac{\beta+1}{2}}}\,|\mathbf{x}-\mathbf{y}|_e^{\beta+1},
\end{aligned}$$

where in the step $(*)$ part (i) of the lemma is used with $C_* = \frac{1}{4}$. The proof concludes by observing that $\frac{1}{4\cdot 12^{\frac{\beta+1}{2}}} \leqslant \frac{1}{\beta}$ for $\beta > 0$.

(iii) This follows form the fourth inequality in Lemma 2.3.1:

$$\begin{aligned}
&\left(|\mathbf{x}|_e^{\beta-1}\mathbf{x} - |\mathbf{y}|_e^{\beta-1}\mathbf{y}\right)\cdot\left(|\mathbf{x}|_e^{\alpha-1}\mathbf{x} - |\mathbf{y}|_e^{\alpha-1}\mathbf{y}\right) \\
&= \quad |\mathbf{x}|_e^{\alpha+\beta} + |\mathbf{y}|_e^{\alpha+\beta} - |\mathbf{x}|_e^{\beta-1}|\mathbf{y}|_e^{\alpha-1}\mathbf{x}\cdot\mathbf{y} - |\mathbf{y}|_e^{\beta-1}|\mathbf{x}|_e^{\alpha-1}\mathbf{x}\cdot\mathbf{y} \\
&\geqslant \quad |\mathbf{x}|_e^{\alpha+\beta} + |\mathbf{y}|_e^{\alpha+\beta} - |\mathbf{x}|_e^{\beta}|\mathbf{y}|_e^{\alpha} - |\mathbf{y}|_e^{\beta}|\mathbf{x}|_e^{\alpha} \\
&= \quad \left(|\mathbf{x}|_e^{\beta} - |\mathbf{y}|_e^{\beta}\right)\left(|\mathbf{x}|_e^{\alpha} - |\mathbf{y}|_e^{\alpha}\right) \\
&\geqslant \quad \frac{4\alpha\beta}{(\alpha+\beta)^2}\left(|\mathbf{x}|_e^{\frac{\alpha+\beta}{2}} - |\mathbf{y}|_e^{\frac{\alpha+\beta}{2}}\right)^2.
\end{aligned}$$

(iv) The proof can be found in [187, Lemma 6.4]. □

Lemma A.1.2. *Let Ω be a nonempty bounded set in $\mathbb{R}^d, d \in \mathbb{N}$, and let $1 \leqslant p < \infty$. Suppose that*

$$u_n \to u \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad n \to \infty.$$

(i) If $h : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, then

$$h(u_n) \to h(u) \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad n \to \infty.$$

(ii) If $h : \mathbb{R} \to \mathbb{R}$ is continuous and satisfies the growth condition $|h(s)| \leqslant C_0(1+s)$ for all $s \in \mathbb{R}$ with $C_0 > 0$, then there exists a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that

$$h(u_{n_k}) \to h(u) \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad k \to \infty.$$

(iii) If $h : \mathbb{R} \to \mathbb{R}$ is continuous and linear (thus bounded), then there exists a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that

$$h(u_{n_k}) \to h(u) \quad \textit{in } \mathrm{L}^p(\Omega) \quad \textit{as} \quad k \to \infty.$$

Proof. From the convergence $u_n \to u$ in $\mathrm{L}^p(\Omega)$, it follows (see Example 2.7.8) that there exists a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}$ such that

$$u_{n_k}(\mathbf{x}) \to u(\mathbf{x}) \qquad \text{as } k \to \infty \qquad \text{for a.a.} \quad \mathbf{x} \in \Omega. \tag{A.2}$$

Moreover, there exists a function $v \in \mathrm{L}^p(\Omega)$ such that

$$|u_{n_k}(\mathbf{x})| \leqslant v(\mathbf{x}) \qquad \text{for all } n_k \quad \text{and a.a.} \quad \mathbf{x} \in \Omega.$$

(i) If h is Lipschitz continuous, then

$$\begin{aligned}
\lim_{n\to\infty} \|h(u_n) - h(u)\|^p_{\mathrm{L}^p(\Omega)} &= \lim_{n\to\infty} \int_\Omega |h(u_n(\mathbf{x})) - h(u(\mathbf{x}))|^p \ \mathrm{d}\mathbf{x} \\
&\leqslant \lim_{n\to\infty} L_0^p \int_\Omega |u_n(\mathbf{x}) - u(\mathbf{x})|^p \ \mathrm{d}\mathbf{x} \\
&= L_0^p \lim_{n\to\infty} \|u_n - u\|^p_{\mathrm{L}^p(\Omega)} \\
&= 0.
\end{aligned}$$

(ii) If h is continuous, then the convergence result follows from

$$\begin{aligned}
\lim_{n\to\infty} \|h(u_{n_k}) - h(u)\|^p_{\mathrm{L}^p(\Omega)} &= \lim_{n\to\infty} \int_\Omega |h(u_{n_k}(\mathbf{x})) - h(u(\mathbf{x}))|^p \ \mathrm{d}\mathbf{x} \\
&\overset{(*)}{=} \int_\Omega \lim_{n\to\infty} |h(u_{n_k}(\mathbf{x})) - h(u(\mathbf{x}))|^p \ \mathrm{d}\mathbf{x} \\
&= \int_\Omega 0 \ \mathrm{d}\mathbf{x} \\
&= 0.
\end{aligned}$$

For $(*)$, we need the Lebesgue dominated convergence theorem 2.9.12 with $M = \Omega$, $Y = \mathbb{R}$ and $f_{n_k}(\mathbf{x}) = |h(u_{n_k}(\mathbf{x})) - h(u(\mathbf{x}))|^p$. By (A.2), the continuity of h and $x \mapsto |x|^p$, we get that $f_{n_k}(\mathbf{x})$ converges for a.a. $\mathbf{x} \in \Omega$ to 0. Therefore, the second condition is already satisfied. For the first condition, we get by the triangle inequality and the Growth condition on h that for almost all $x \in \Omega$ hold

$$\begin{aligned}
|f_{n_k}(\mathbf{x})| = |h(u_{n_k}(\mathbf{x})) - h(u(\mathbf{x}))|^p &\leqslant (|h(u_{n_k}(\mathbf{x}))| + |h(u(\mathbf{x}))|)^p \\
&\leqslant (C_0(|u_{n_k}(\mathbf{x})| + 1) + C_0(|u(\mathbf{x})| + 1))^p \\
&\leqslant C_0^p(|v(\mathbf{x})| + |u(\mathbf{x})| + 2)^p.
\end{aligned}$$

Define $g : \Omega \to \mathbb{R}$ by $g(\mathbf{x}) = C_0^p(|v(\mathbf{x})| + |u(\mathbf{x})| + 2)^p$. Then, we get by Jensen inequality (see Lemma 2.3.11) that

$$\begin{aligned}
\int_\Omega g(\mathbf{x})\ \mathrm{d}\mathbf{x} &= C_0^p \int_\Omega (|v(\mathbf{x})| + |u(\mathbf{x})| + 2)^p\ \mathrm{d}\mathbf{x} \\
&\leqslant \frac{C_0^p}{3}\left[\int_\Omega |v(\mathbf{x})|^p\ \mathrm{d}\mathbf{x} + \int_\Omega |u(\mathbf{x})|^p\ \mathrm{d}\mathbf{x} + \int_\Omega 2^p\ \mathrm{d}\mathbf{x}\right] \\
&\leqslant C.
\end{aligned}$$

Hence, the first condition is also satisfied and we can apply the Lebesgue dominated convergence theorem.

(iii) If h is linear and continuous, then h is bounded and no growth condition is necessary to obtain the convergence via the Lebesgue dominated convergence theorem. We follow the same lines as in (ii). Suppose that $|h(s)| \leqslant C_1$ for all $s \in \mathbb{R}$. Then, the sequence f_{n_k} is uniformly bounded, i.e. $|f_{n_k}(\mathbf{x})| \leqslant C_1^p$ for a.a. $\mathbf{x}$ and for all n_k. Define now $g : \Omega \to \mathbb{R}$ by $g(\mathbf{x}) = C_1^p$. Then the sequence is dominated by g. Furthermore, g is integrable since it is a constant function on a set of finite measure. □

Lemma A.1.3. *Let V, Y and W be Banach spaces, V be separable and reflexive,*

$$V \hookrightarrow\hookrightarrow Y \qquad \textit{and} \qquad Y \hookrightarrow W.$$

Then $\mathrm{W}^{1,2,2}([0,T]; V, W) \hookrightarrow\hookrightarrow \mathrm{L}^2((0,T), Y)$. For every bounded sequence $\{u_n\}$ in $\mathrm{W}^{1,2,2}([0,T]; V, W)$ there exists a function $u \in \mathrm{L}^2((0,T), Y)$ and a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}}$ of $\{u_n\}_{n\in\mathbb{N}}$ such that

$$\begin{cases} u_{n_k} \to u, & \textit{in} \quad \mathrm{L}^2((0,T), Y), \\ u_{n_k} \rightharpoonup u, & \textit{in} \quad \mathrm{L}^2((0,T), V). \end{cases}$$

Moreover,

$$u \in \mathrm{C}([0,T], Y) \quad \textit{and} \quad \partial_t u_{n_k} \rightharpoonup \partial_t u, \quad \textit{in} \quad \mathrm{L}^2((0,T), W)$$

if also the following evolution triple is satisfied

$$V \subset V_1 \hookrightarrow Y \cong Y^* \hookrightarrow W = V_1^* \subset V^*, \tag{A.4}$$

i.e. W has a predual, with
- *V_1 a reflexive and separable Banach space;*
- *Y Hilbert;*
- *V_1 dense in Y.*

Proof. The compact embedding of $\mathrm{W}^{1,2,2}([0,T]; V, W)$ in $\mathrm{L}^2((0,T), Y)$ is shown in [38, Lemma 7.7]. In the proof of Lemma 7.7 in [38], the weak convergence of $\{\partial_t u_{n_k}\}$ is not studied in detail. Now, let us consider the prescribed situation in

(A.4). The duality pairing between $W = V_1^*$ and V_1 can be seen as a continuous extension of the inner product on Y, i.e.

$$\langle u, v\rangle_{W\times V_1} = (u, v)_Y, \qquad \forall u \in Y \text{ and } v \in V_1.$$

The spaces W and V^* are separable and reflexive because V_1, respectively V is separable and reflexive. Therefore, also the spaces $\mathrm{L}^2((0,T),W)$, $\mathrm{L}^2((0,T),V_1)$ and $\mathrm{L}^2((0,T),V^*)$ are reflexive. By Lemma 2.9.5(iii), we have for each $u \in \mathrm{W}^{1,2,2}([0,T];V_1,W)$, the following generalized integration by parts formula with $0 \leqslant t_1 \leqslant t_2 \leqslant T$:

$$\begin{aligned}(u(t_2), v(t_2))_Y &- (u(t_1), v(t_1))_Y \\ &= \int_{t_1}^{t_2} \left\langle \frac{\mathrm{d}u(t)}{\mathrm{d}t}, v(t)\right\rangle_{W\times V_1} \mathrm{d}t + \int_{t_1}^{t_2} \left\langle u(t), \frac{\mathrm{d}v(t)}{\mathrm{d}t}\right\rangle_{W\times V_1} \mathrm{d}t. \end{aligned} \tag{A.5}$$

Note that also $u_{n_k} \rightharpoonup u$ in $\mathrm{L}^2((0,T),V_1)$ due to the reflexivity of this space, i.e. $u_{n_k}(t) \rightharpoonup u(t)$ in V_1 for a.a. $t \in [0,T]$. The sequence $\{u_n\}$ is bounded in $\mathrm{W}^{1,2,2}([0,T];V,W)$. This implies that $\{\partial_t u_{n_k}\}$ is bounded in $\mathrm{L}^2((0,T),W)$. By the reflexivity of this space, there exists a subsequence $\{\partial_t u_{n_{kl}}\}$ such that

$$\partial_t u_{n_{kl}} \rightharpoonup z \quad \text{in } \mathrm{L}^2((0,T),W).$$

Thanks to Theorem 2.9.11, we immediately get that $z = \partial_t u \in \mathrm{L}^2((0,T),W)$. Nonetheless, it is an interesting exercise to make the proof. For this, we take v time independent in (A.5), i.e. put $v = \varphi \in V_1$. Then, the following diagram is valid

$$\begin{array}{ccc} \left(u_{n_{kl}}(t_2), \varphi\right)_Y - \left(u_{n_{kl}}(t_1), \varphi\right)_Y & = & \displaystyle\int_{t_1}^{t_2} \left\langle \frac{\mathrm{d}u_{n_{kl}}(t)}{\mathrm{d}t}, \varphi\right\rangle_{W\times V_1} \mathrm{d}t \\ \downarrow & & \downarrow \\ (u(t_2), \varphi)_Y - (u(t_1), \varphi)_Y & = & \displaystyle\int_{t_1}^{t_2} \langle z, \varphi\rangle_{W\times V_1} \mathrm{d}t, \end{array} \tag{A.6}$$

for all $\varphi \in V_1$ and for a.a. $0 \leqslant t_1 \leqslant t_2 \leqslant T$. First, we prove that this relation is valid for every time. From the result of (A.6) follows for all $\varphi \in V_1$ that

$$(u(t_2), \varphi)_Y - (u(t_1), \varphi)_Y \leqslant \|\varphi\|_{V_1} \sqrt{\int_{t_1}^{t_2} \|z(t)\|_W^2 \,\mathrm{d}t}\sqrt{t_2 - t_1} \lesssim \sqrt{t_2 - t_1},$$

i.e.

$$\lim_{t_2 \to t_1} (u(t_2), \varphi)_Y = (u(t_1), \varphi)_Y, \quad \forall \varphi \in V_1.$$

This means that $u(t_2) \rightharpoonup u(t_1)$ as $t_2 \to t_1$. Consider $\varphi = u(t_2)$ and $\varphi = u(t_1)$ in (A.6). Adding up the resulting equations gives

$$\|u(t_2)\|_Y^2 - \|u(t_1)\|_Y^2 = \int_{t_1}^{t_2} \langle z, u(t_2) + u(t_1)\rangle_{W\times V_1} \mathrm{d}t.$$

This implies that $\|u(t_2)\|_Y \to \|u(t_1)\|_Y$, hence by Lemma 2.4.19, we get $u(t_2) \to u(t_1)$ in Y. Therefore, $u \in \mathrm{C}([0,T],Y)$ and (A.6) is valid for every $0 \leqslant t_1 \leqslant t_2 \leqslant T$. From relation (A.6), we get that

$$\left(\frac{u(t+h)-u(t)}{h},\varphi\right)_Y = \frac{1}{h}\int_t^{t+h} \langle z(s),\varphi\rangle_{W\times V_1}\,\mathrm{d}s$$

and

$$\begin{aligned}
\left\|\frac{u(t+h)-u(t)}{h}-z(t)\right\|_W &= \sup_{\substack{\varphi\in V_1\\ \|\varphi\|_{V_1}=1}} \left\langle \frac{u(t+h)-u(t)}{h}-z(t),\varphi\right\rangle_{W\times V_1} \\
&= \sup_{\substack{\varphi\in V_1\\ \|\varphi\|_{V_1}=1}} \frac{1}{h}\int_t^{t+h} \langle z(s)-z(t),\varphi\rangle_{W\times V_1}\,\mathrm{d}s \\
&\leqslant \sup_{\substack{\varphi\in V_1\\ \|\varphi\|_{V_1}=1}} \frac{1}{h}\int_t^{t+h} \|z(s)-z(t)\|_W\,\|\varphi\|_{V_1}\,\mathrm{d}s \\
&= \frac{1}{h}\int_t^{t+h} \|z(s)-z(t)\|_W\,\mathrm{d}s \\
&\leqslant \sqrt{\frac{1}{h}\int_t^{t+h} \|z(s)-z(t)\|_W^2\,\mathrm{d}s}. \qquad \text{(A.7)}
\end{aligned}$$

By Theorem 2.9.10(ii), there exists an element $y \in \mathrm{C}\left([0,T],W\right)$ such that

$$\int_0^T \|y(t)-z(t)\|^2\,\mathrm{d}t < h^2. \qquad \text{(A.8)}$$

Using this function y, we easily see that

$$\begin{aligned}
\frac{1}{h}\int_t^{t+h} \|z(s)-z(t)\|_W^2\,\mathrm{d}s &\leqslant \frac{1}{h}\int_t^{t+h} \|z(s)-y(s)\|_W^2\,\mathrm{d}s \\
&\quad + \frac{1}{h}\int_t^{t+h} \|y(s)-y(t)\|_W^2\,\mathrm{d}s + \frac{1}{h}\int_t^{t+h} \|y(t)-z(t)\|_W^2\,\mathrm{d}s.
\end{aligned}$$

By the first mean value theorem for integration 2.2.6, there exists $\xi \in [t,t+h]$ such that

$$\begin{aligned}
\frac{1}{h}\int_t^{t+h} \|z(s)-z(t)\|_W^2\,\mathrm{d}s &\leqslant \frac{1}{h}\int_t^{t+h} \|z(s)-y(s)\|_W^2\,\mathrm{d}s \\
&\quad + \|y(\xi)-y(t)\|_W^2 + \|y(t)-z(t)\|_W^2.
\end{aligned}$$

By the density argument (A.8), we have that

$$\frac{1}{h}\int_t^{t+h} \|z(s)-z(t)\|_W^2\,\mathrm{d}s \leqslant h + \|y(\xi)-y(t)\|_W^2 + \|y(t)-z(t)\|_W^2.$$

Therefore, by the continuity of y in the time variable and $\xi \to t$ as $h \to 0$, we obtain that

$$\begin{aligned}\int_0^T \|\partial_t u(t) - z(t)\|_W^2 \,\mathrm{d}t &= \lim_{h\to 0}\int_0^T \left\|\frac{u(t+h)-u(t)}{h} - z(t)\right\|_W^2 \mathrm{d}t \\ &\overset{(A.7)}{\leqslant} \lim_{h\to 0}\int_0^T \left[h + \|y(\xi)-y(t)\|_W^2 + \|y(t)-z(t)\|_W^2\right] \mathrm{d}t \\ &\overset{(A.8)}{\leqslant} \lim_{h\to 0}\left[hT + \int_0^T \|y(\xi)-y(t)\|_W^2 \,\mathrm{d}t + h^2\right] \\ &= 0.\end{aligned}$$

This implies that $z = \partial_t u \in \mathrm{L}^2\left((0,T),W\right)$. The weak convergence of the whole sequence $\{\partial_t u_{n_k}\}$ to $\{\partial_t u\}$ follows from Lemma 2.4.20 because this result can be obtained for each subsequence of $\{\partial_t u_{n_k}\}$. $\square$

Lemma A.1.4.

(i) Assume that $\mathbf{r} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$*,* $h \in \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$*,* $\overline{\mathbf{u}}_0 \in \mathbf{H}^1(\Omega)$*,* $\overline{\mathbf{u}}_1 \in \mathbf{L}^2(\Omega)$*,* $\overline{\theta}_0 \in \mathrm{L}^2(\Omega)$*,* $\mathbf{g}(\mathbf{0}) = \mathbf{0}$*,* $\mathbf{g}' > \mathbf{0}$ *and* $|\mathbf{g}(s)| \leqslant C(1+|s|)$ *a.e. in* $\mathbb{R}$*. Then* (8.15)*, has a unique solution* $\langle \mathbf{u}, \theta\rangle$ *such that*

$$\begin{aligned}\mathbf{u} &\in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T),\mathbf{H}_0^1(\Omega)\right),\\ \partial_t \mathbf{u} &\in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right), \partial_{tt}\mathbf{u} \in \mathrm{L}^2\left((0,T),\mathbf{H}_0^1(\Omega)^*\right),\\ \theta &\in \mathrm{C}\left([0,T],\mathrm{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T),\mathrm{H}_0^1(\Omega)\right),\\ \partial_t\theta &\in \mathrm{L}^2\left((0,T),\mathrm{H}_0^1(\Omega)^*\right).\end{aligned}$$

(ii) Assume that $\mathbf{r}(0) \in \mathbf{L}^2(\Omega)$*,* $h(0) \in \mathrm{L}^2(\Omega)$*,* $\partial_t\mathbf{r} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$*,* $\partial_t h \in \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$*,* $\overline{\mathbf{u}}_0 \in \mathbf{H}^2(\Omega)\cap\mathbf{H}_0^1(\Omega)$*,* $\overline{\mathbf{u}}_1 \in \mathbf{H}^1(\Omega)$*,* $\overline{\theta}_0 \in \mathrm{H}^2(\Omega)\cap\mathrm{H}_0^1(\Omega)$*,* $\mathbf{g}(\mathbf{0}) = \mathbf{0}$ *and* $\mathbf{0} < \mathbf{g}'(s) \leqslant \mathbf{C}$ *a.e. in* $\mathbb{R}$*. Then* (8.15)*, has a unique solution* $\langle \mathbf{u}, \theta\rangle$ *such that*

$$\begin{aligned}\mathbf{u} &\in \mathrm{C}\left([0,T],\mathbf{H}_0^1(\Omega)\right),\\ \partial_t\mathbf{u} &\in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right) \cap \mathrm{L}^2\left((0,T),\mathbf{H}_0^1(\Omega)\right),\\ \partial_{tt}\mathbf{u} &\in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right),\\ \theta &\in \mathrm{C}\left([0,T],\mathrm{H}_0^1(\Omega)\right), \partial_t\theta \in \mathrm{L}^2\left((0,T),\mathrm{H}_0^1(\Omega)\right).\end{aligned}$$

In the special situation that $\overline{\mathbf{u}}_0 = \mathbf{0}$*,* $\overline{\mathbf{u}}_1 = \mathbf{0}$*,* $\overline{\theta}_0 = 0$*,* $h = 0$ *and* $\mathbf{r} = \mathbf{r}(\mathbf{x})$*, the following estimate is valid*

$$\begin{aligned}\max_{t\in[0,T]} &\left\{\|\partial_{tt}\mathbf{u}(t)\|^2 + \|\nabla\partial_t\mathbf{u}(t)\|^2 + \|\nabla\cdot\partial_t\mathbf{u}(t)\|^2\right.\\ &\left.+\|\partial_t\theta(t)\|^2 + \|\nabla\theta(t)\|^2\right\} + \int_0^T \|\nabla\partial_t\theta(s)\|^2 \,\mathrm{d}s \leqslant C\,\|\mathbf{r}\|^2.\end{aligned}$$

Proof. In this proof, a priori estimates are derived under the assumption that there exists a solution to (8.5)-(8.6) with $\mathbf{p} = \mathbf{0}$. This gives insight into the space where the solution belongs. Then, Rothe's method can be applied to establish the existence of a solution.

(i) Firstly, we choose $\boldsymbol{\varphi} = \partial_t \mathbf{u}(t)$ and $\psi = \theta(t)$ in (8.5)-(8.6) with $\mathbf{p} = \mathbf{0}$ and integrate in time over $t \in (0, \eta) \subset (0, T)$. We obtain that

$$\frac{\|\partial_t \mathbf{u}(\eta)\|^2}{2} + \int_0^\eta \left(\mathbf{g}\left(\partial_t \mathbf{u}\right), \partial_t \mathbf{u}\right) + \alpha \frac{\|\nabla \mathbf{u}(\eta)\|^2}{2} + \beta \frac{\|\nabla \cdot \mathbf{u}(\eta)\|^2}{2} \\ + \gamma \int_0^\eta (\nabla \theta, \partial_t \mathbf{u}) = \int_0^\eta (\mathbf{r}, \partial_t \mathbf{u}) + \frac{\|\overline{\mathbf{u}}_1\|^2}{2} + \alpha \frac{\|\nabla \overline{\mathbf{u}}_0\|^2}{2} + \beta \frac{\|\nabla \cdot \overline{\mathbf{u}}_0\|^2}{2}$$

and

$$\frac{\|\theta(\eta)\|^2}{2} + \rho \int_0^\eta \|\nabla \theta\|^2 + \int_0^\eta (k * \nabla \theta, \nabla \theta) \\ - \gamma \int_0^\eta (\partial_t \mathbf{u}, \nabla \theta) = \int_0^\eta (h, \theta) + \frac{\|\overline{\theta}_0\|^2}{2}.$$

Adding both equations together gives

$$\frac{\|\partial_t \mathbf{u}(\eta)\|^2}{2} + \int_0^\eta \left(\mathbf{g}\left(\partial_t \mathbf{u}\right), \partial_t \mathbf{u}\right) + \alpha \frac{\|\nabla \mathbf{u}(\eta)\|^2}{2} + \beta \frac{\|\nabla \cdot \mathbf{u}(\eta)\|^2}{2} + \frac{\|\theta(\eta)\|^2}{2} \\ + \rho \int_0^\eta \|\nabla \theta\|^2 + \int_0^\eta (k * \nabla \theta, \nabla \theta) = \int_0^\eta (\mathbf{r}, \partial_t \mathbf{u}) + \frac{\|\overline{\mathbf{u}}_1\|^2}{2} \\ + \alpha \frac{\|\nabla \overline{\mathbf{u}}_0\|^2}{2} + \beta \frac{\|\nabla \cdot \overline{\mathbf{u}}_0\|^2}{2} + \int_0^\eta (h, \theta) + \frac{\|\overline{\theta}_0\|^2}{2}.$$

If we assume that $\mathbf{g}$ is componentwise strictly monotone increasing, then

$$\int_0^\eta \left(\mathbf{g}\left(\partial_t \mathbf{u}\right), \partial_t \mathbf{u}\right) \geqslant 0.$$

The strongly positive-definiteness of k implies that

$$\int_0^\eta (k * \nabla \theta, \nabla \theta) \geqslant 0.$$

Employing the Cauchy, Young and Friedrichs inequalities give

$$\int_0^\eta (h, \theta) \leqslant \int_0^\eta \frac{\|h\|^2}{2} + \int_0^\eta \frac{\|\theta\|^2}{2}$$

and

$$\int_0^\eta (\mathbf{r}, \partial_t \mathbf{u}) \leqslant \int_0^\eta \frac{\|\mathbf{r}\|^2}{2} + \int_0^\eta \frac{\|\partial_t \mathbf{u}\|^2}{2}.$$

We arrive at

$$\begin{aligned}\frac{\|\partial_t\mathbf{u}(\eta)\|^2}{2} &+ \alpha\frac{\|\nabla\mathbf{u}(\eta)\|^2}{2} + \beta\frac{\|\nabla\cdot\mathbf{u}(\eta)\|^2}{2} + \frac{\|\theta(\eta)\|^2}{2} \\ &+ \rho\int_0^\eta \|\nabla\theta\|^2 \leqslant \int_0^\eta \frac{\|\mathbf{r}\|^2}{2} + \int_0^\eta \frac{\|\partial_t\mathbf{u}\|^2}{2} + \frac{\|\overline{\mathbf{u}}_1\|^2}{2} \\ &+ \alpha\frac{\|\nabla\overline{\mathbf{u}}_0\|^2}{2} + \beta\frac{\|\nabla\cdot\overline{\mathbf{u}}_0\|^2}{2} + \int_0^\eta \frac{\|h\|^2}{2} + \int_0^\eta \frac{\|\theta\|^2}{2} + \frac{\|\overline{\theta}_0\|^2}{2}.\end{aligned}$$

Assuming that $\mathbf{r} \in \mathrm{L}^2\left((0,T),\mathbf{L}^2(\Omega)\right)$, $\overline{\mathbf{u}}_1 \in \mathbf{L}^2(\Omega)$, $\overline{\mathbf{u}}_0 \in \mathbf{H}_0^1(\Omega)$, $\overline{\theta}_0 \in \mathrm{L}^2(\Omega)$ and $h \in \mathrm{L}^2\left((0,T),\mathrm{L}^2(\Omega)\right)$, an application of Grönwall's argument gives

$$\max_{t\in[0,T]}\left\{\|\partial_t\mathbf{u}(t)\|^2 + \|\nabla\mathbf{u}(t)\|^2 + \|\nabla\cdot\mathbf{u}(t)\|^2 + \|\theta(t)\|^2\right\} + \int_0^T \|\nabla\theta\|^2 \leqslant C.$$

Note that for a.a. $t \in [0,T]$ it holds that

$$\|(k*\nabla\theta)(t)\| \leqslant (|k| * \|\nabla\theta\|)(t) \leqslant \sqrt{\int_0^t |k(s)|^2 \,\mathrm{d}s}\sqrt{\int_0^t \|\nabla\theta(s)\|^2 \,\mathrm{d}s}.$$

From the a priori estimate, it follows that

$$\int_0^T \|\partial_{tt}\mathbf{u}(s)\|^2_{\mathbf{H}_0^1(\Omega)^*} \,\mathrm{d}s \leqslant C \quad \text{and} \quad \int_0^T \|\partial_t\theta(s)\|^2_{\mathrm{H}_0^1(\Omega)^*} \,\mathrm{d}s \leqslant C.$$

Consider the following evolution triple of spaces

$$\mathrm{H}_0^1(\Omega) \hookrightarrow \mathrm{L}^2(\Omega) \cong \mathrm{L}^2(\Omega)^* \hookrightarrow \mathrm{H}_0^1(\Omega)^*.$$

Applying Lemma 2.9.5(i) and (iii), we get that

$$\mathbf{u} \in \mathrm{C}\left([0,T],\mathbf{L}^2(\Omega)\right) \quad \text{and} \quad \theta \in \mathrm{C}\left([0,T],\mathrm{L}^2(\Omega)\right).$$

(ii) For the second estimate, we start with taking the time derivative of both equations in (8.5)-(8.6). Note that from the differentiation under the integral sign formula (see Theorem 2.2.7) follows that

$$\partial_t(k*\nabla\theta)(t) = k(0)\nabla\theta(t) + (k'*\nabla\theta)(t).$$

Then, we choose $\boldsymbol{\varphi} = \partial_{tt}\mathbf{u}(t)$ and $\psi = \partial_t\theta(t)$ in the result and integrate in time over $t \in (0,\eta) \subset (0,T)$. We get that

$$\begin{aligned}&\frac{\|\partial_{tt}\mathbf{u}(\eta)\|^2}{2} + \int_0^\eta \left(\mathbf{g}'\left(\partial_t\mathbf{u}\right)\partial_{tt}\mathbf{u}, \partial_{tt}\mathbf{u}\right) + \alpha\frac{\|\nabla\partial_t\mathbf{u}(\eta)\|^2}{2} + \beta\frac{\|\nabla\cdot\partial_t\mathbf{u}(\eta)\|^2}{2} \\ &+\gamma\int_0^\eta (\nabla\partial_t\theta, \partial_{tt}\mathbf{u}) = \int_0^\eta (\partial_t\mathbf{r}, \partial_{tt}\mathbf{u}) + \frac{\|\partial_{tt}\mathbf{u}(0)\|^2}{2} + \alpha\frac{\|\nabla\overline{\mathbf{u}}_1\|^2}{2} + \beta\frac{\|\nabla\cdot\overline{\mathbf{u}}_1\|^2}{2}\end{aligned}$$

and

$$\frac{\|\partial_t\theta(\eta)\|^2}{2} + \rho\int_0^\eta \|\nabla\partial_t\theta\|^2 + k(0)\frac{\|\nabla\theta(\eta)\|^2}{2} + \int_0^\eta (k' * \nabla\theta, \nabla\partial_t\theta)$$
$$- \gamma\int_0^\eta (\partial_{tt}\mathbf{u}, \nabla\partial_t\theta) = \int_0^\eta (\partial_t h, \partial_t\theta) + \frac{\|\partial_t\theta(0)\|^2}{2} + k(0)\frac{\|\nabla\overline{\theta}_0\|^2}{2}.$$

We add both equations together

$$\begin{aligned}
&\frac{\|\partial_{tt}\mathbf{u}(\eta)\|^2}{2} + \alpha\frac{\|\nabla\partial_t\mathbf{u}(\eta)\|^2}{2} + \beta\frac{\|\nabla\cdot\partial_t\mathbf{u}(\eta)\|^2}{2} \\
&\qquad + \frac{\|\partial_t\theta(\eta)\|^2}{2} + \rho\int_0^\eta \|\nabla\partial_t\theta\|^2 + k(0)\frac{\|\nabla\theta(\eta)\|^2}{2} \\
&= \int_0^\eta (\partial_t\mathbf{r}, \partial_{tt}\mathbf{u}) - \int_0^\eta (\mathbf{g}'(\partial_t\mathbf{u})\,\partial_{tt}\mathbf{u}, \partial_{tt}\mathbf{u}) + \frac{\|\partial_{tt}\mathbf{u}(0)\|^2}{2} \\
&\qquad + \alpha\frac{\|\nabla\overline{\mathbf{u}}_1\|^2}{2} + \beta\frac{\|\nabla\cdot\overline{\mathbf{u}}_1\|^2}{2} - \int_0^\eta (k' * \nabla\theta, \nabla\partial_t\theta) \\
&\qquad\qquad + \int_0^\eta (\partial_t h, \partial_t\theta) + \frac{\|\partial_t\theta(0)\|^2}{2} + k(0)\frac{\|\nabla\overline{\theta}_0\|^2}{2}.
\end{aligned}$$

Due to the definition of k, we have that $k(0) \in (0, +\infty)$. Note that $\partial_{tt}\mathbf{u}(0)$ and $\partial_t\theta(0)$ are not well-defined. Knowing that $\mathbf{r}(0) \in \mathbf{L}^2(\Omega)$, $h(0) \in \mathrm{L}^2(\Omega)$, $\mathbf{u}_0 \in \mathbf{H}^2(\Omega) \cap \mathbf{H}^1_0(\Omega)$, $\mathbf{u}_1 \in \mathbf{H}^1(\Omega)$, $\theta_0 \in \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$, we can define the following compatibility conditions

$$\partial_{tt}\mathbf{u}(0) = \mathbf{r}(0) - \mathbf{g}(\overline{\mathbf{u}}_1) + \alpha\Delta\overline{\mathbf{u}}_0 + \beta\nabla(\nabla\cdot\overline{\mathbf{u}}_0) - \gamma\nabla\overline{\theta}_0,$$
$$\partial_t\theta(0) = h(0) + \rho\Delta\overline{\theta}_0 + k * \Delta\overline{\theta}_0 - \gamma\nabla\cdot\overline{\mathbf{u}}_1,$$

i.e.

$$\|\partial_{tt}\mathbf{u}(0)\| \lesssim 1 \quad \text{and} \quad \|\partial_t\theta(0)\| \lesssim 1.$$

Thanks to (i), it holds that

$$\begin{aligned}
&\left|\int_0^\eta (k' * \nabla\theta, \nabla\partial_t\theta)\right| \\
&\leqslant C_\varepsilon\int_0^\eta \left\|\int_0^t k'(t-s)\nabla\theta(s)\,\mathrm{d}s\right\|^2 \mathrm{d}t + \varepsilon\int_0^\eta \|\nabla\partial_t\theta(t)\|^2\,\mathrm{d}t \\
&\leqslant C_\varepsilon\int_0^\eta \left(\int_0^t |k'(t-s)|\,\|\nabla\theta(s)\|\,\mathrm{d}s\right)^2 \mathrm{d}t + \varepsilon\int_0^\eta \|\nabla\partial_t\theta(t)\|^2\,\mathrm{d}t \\
&\leqslant C_\varepsilon\int_0^\eta \left(\int_0^t |k'(t-s)|^2\,\mathrm{d}s\right)\left(\int_0^t \|\nabla\theta(s)\|^2\,\mathrm{d}s\right)\mathrm{d}t \\
&\qquad + \varepsilon\int_0^\eta \|\nabla\partial_t\theta(t)\|^2\,\mathrm{d}t \\
&\leqslant C_\varepsilon + \varepsilon\int_0^\eta \|\nabla\partial_t\theta(t)\|^2\,\mathrm{d}t.
\end{aligned}$$

Fixing ε sufficiently small, an application of Grönwall's inequality gives

$$\max_{t\in[0,T]} \left\{ \|\partial_{tt}\mathbf{u}(t)\|^2 + \|\nabla\partial_t\mathbf{u}(t)\|^2 + \|\nabla\cdot\partial_t\mathbf{u}(t)\|^2 \right.$$
$$\left. + \|\partial_t\theta(t)\|^2 + \|\nabla\theta(t)\|^2 \right\} + \int_0^T \|\nabla\partial_t\theta\|^2 \leqslant C.$$

Applying Lemma 2.9.5(i), we obtain that

$$\mathbf{u} \in \mathrm{C}\left([0,T], \mathbf{H}_0^1(\Omega)\right), \quad \partial_t\mathbf{u} \in \mathrm{C}\left([0,T], \mathbf{L}^2(\Omega)\right)$$

and

$$\theta \in \mathrm{C}\left([0,T], \mathrm{H}_0^1(\Omega)\right).$$

□

A.2 Crank-Nicolson scheme: example

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz continuous boundary, $d \in \mathbb{N}$. Consider the following heat problem

$$\begin{cases} \partial_t u(t,\boldsymbol{x}) - \Delta u(t,\boldsymbol{x}) = f(\boldsymbol{x}) & \text{in } (0,T]\times\Omega, \\ u(t,\boldsymbol{x}) = 0 & \text{on } (0,T]\times\partial\Omega, \\ u(0,\boldsymbol{x}) = u_0(\boldsymbol{x}) & \text{in } \Omega. \end{cases}$$

First, the time domain is discretized n equal subintervals $[t_{i-1}, t_i]$ with length $\tau = T/n$ and approximate the problem using the trapezoidal rule (Crank-Nicolson) instead of backward Euler method. Use the notations

$$u(t_i) \approx u_i, \quad \partial_t u(t_i) \approx \delta u_i = \frac{u_i - u_{i-1}}{\tau}.$$

The sequence ($i = 1, \dots, n$) of time-discrete problems is given by

$$\delta u_i - \tfrac{1}{2}(\Delta u_i + \Delta u_{i-1}) = f$$

or

$$u_i - \frac{\tau}{2}\Delta u_i = \tau f + u_{i-1} + \frac{\tau}{2}\Delta u_{i-1}$$

with $u_i = 0$ on $\partial\Omega$. The variational formulation is given by

$$(\delta u_i, \phi) + \tfrac{1}{2}\left(\nabla u_i + \nabla u_{i-1}, \nabla\phi\right) = (f, \phi) \tag{A.9}$$

for all $\phi \in \mathrm{H}_0^1(\Omega)$. If $u_0 \in \mathrm{H}^1(\Omega)$ and $f \in \mathrm{L}^2(\Omega)$, then by the Lax-Milgram lemma, there exists a unique $u_i \in \mathrm{H}_0^1(\Omega)$.

Choose $\phi = (u_i + u_{i-1})\tau$ in (A.9) and sum up for $i = 1, \ldots, j$. We get that

$$\sum_{i=1}^{j} (\delta u_i, (u_i + u_{i-1})\tau) + \tfrac{1}{2} \sum_{i=1}^{j} \|\nabla u_i + \nabla u_{i-1}\|^2 \tau = \sum_{i=1}^{j} (f, (u_i + u_{i-1})\tau)\,.$$

We have that

$$\sum_{i=1}^{j} (\delta u_i, (u_i + u_{i-1})\tau) = \sum_{i=1}^{j} (u_i - u_{i-1}, u_i + u_{i-1}) = \|u_j\|^2 - \|u_0\|^2$$

and by Friedrich's inequality that

$$\begin{aligned} &\left| \sum_{i=1}^{j} (f, (u_i + u_{i-1})\tau) \right| \\ &\leqslant \quad C_\varepsilon \sum_{i=1}^{j} \|f\|^2 \tau + \varepsilon \sum_{i=1}^{j} \|\nabla u_i + \nabla u_{i-1}\|^2 \tau \\ &\leqslant \quad C_\varepsilon + \varepsilon \sum_{i=1}^{j} \|\nabla u_i + \nabla u_{i-1}\|^2 \tau. \end{aligned}$$

Fixing $\varepsilon > 0$ sufficiently small, we obtain that

$$\|u_j\|^2 + \sum_{i=1}^{j} \|\nabla u_i + \nabla u_{i-1}\|^2 \tau \leqslant C.$$

Put $\phi = \delta u_i \tau$ in (A.9) and sum up for $i = 1, \ldots, j$. We get that

$$\sum_{i=1}^{j} \|\delta u_i\|^2 \tau + \tfrac{1}{2} \sum_{i=1}^{j} (\nabla u_i + \nabla u_{i-1}, \nabla \delta u_i)\, \tau = \sum_{i=1}^{j} (f, \delta u_i)\, \tau.$$

For the RHS, we have that

$$\left| \sum_{i=1}^{j} (f, \delta u_i)\, \tau \right| \leqslant C_\varepsilon + \varepsilon \sum_{i=1}^{j} \|\delta u_i\|^2 \tau.$$

The second term in the LHS simplifies to

$$\tfrac{1}{2} \sum_{i=1}^{j} (\nabla u_i + \nabla u_{i-1}, \nabla u_i - \nabla u_{i-1}) = \tfrac{1}{2} \|\nabla u_j\|^2 - \tfrac{1}{2} \|\nabla u_0\|^2\,.$$

Fixing ε sufficiently small, we get that

$$\sum_{i=1}^{j} \|\delta u_i\|^2 \tau + \|\nabla u_j\|^2 \leqslant C, \tag{A.10}$$

if $u_0 \in \mathrm{H}^1(\Omega)$.

Put $\phi = \delta u_i \tau$ and sum up for $i = 1, \dots, j$. We obtain that

$$\sum_{i=1}^{j} \|\delta u_i\|^2 \tau + \tfrac{1}{2} \sum_{i=1}^{j} (\nabla u_i + \nabla u_{i-1}, \nabla \delta u_i)\, \tau = \sum_{i=1}^{j} (f, \delta u_i)\, \tau.$$

Using (A.10), the RHS can be estimated as

$$\left| \sum_{i=1}^{j} (f, \delta u_i)\, \tau \right| \leqslant C.$$

The second term in the LHS can be splitted by the trick $\pm \nabla u_{i-1}$ as follows

$$\begin{aligned} &\tfrac{1}{2} \sum_{i=1}^{j} (\nabla u_i \pm \nabla u_{i-1} + \nabla u_{i-1}, \nabla \delta u_i)\, \tau \\ &\qquad = \tfrac{1}{2} \sum_{i=1}^{j} (\nabla \delta u_i \tau, \nabla \delta u_i)\, \tau + \sum_{i=1}^{j} (\nabla u_{i-1}, \nabla \delta u_i)\, \tau =: S_1 + S_2. \end{aligned}$$

We immediately see that

$$S_1 = \tfrac{1}{2} \sum_{i=1}^{j} \|\nabla \delta u_i \tau\|^2 = \tfrac{1}{2} \sum_{i=1}^{j} \|\nabla u_i - \nabla u_{i-1}\|^2$$

and by (A.10) we derive that

$$\begin{aligned} S_2 &= \sum_{i=1}^{j} \left[(\nabla u_{i-1}, \nabla u_i) - \|\nabla u_{i-1}\|^2 \right] \\ &\leqslant \sum_{i=1}^{j} \left[\frac{\|\nabla u_{i-1}\|^2}{2} + \frac{\|\nabla u_i\|^2}{2} - \|\nabla u_{i-1}\|^2 \right] \\ &= \sum_{i=1}^{j} \left[\tfrac{1}{2} \|\nabla u_i\|^2 - \tfrac{1}{2} \|\nabla u_{i-1}\|^2 \right] \\ &= \tfrac{1}{2} \|\nabla u_j\|^2 - \tfrac{1}{2} \|\nabla u_0\|^2 \\ &\leqslant C. \end{aligned}$$

Combining the previous results give

$$\sum_{i=1}^{j} \|\delta u_i\|^2 \tau + \sum_{i=1}^{j} \|\nabla u_i - \nabla u_{i-1}\|^2 \leqslant C.$$

Therefore, we obtained the same estimates as in the case of using backward Euler method if $u_0 \in \mathrm{H}^1(\Omega)$. This implies that the convergence results stay valid and that there exists a unique weak solution to the problem under consideration by using the Crank-Nicolson scheme.

Abbreviations

a.a.	almost all
a.e.	almost everywhere
BC	boundary condition
BVP	boundary value problem
EV	eigenvalue
FEM	finite element method
IBVP	initial and boundary value problem
IC	initial condition
iff	if and only if
inf	infimum
IP	inverse problem
ISP	inverse source problem
LHS	left-hand side
lim inf	limit inferior
PDE	partial differential equation
RHS	right-hand side
sup	supremum

Index

References

[1] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. Providence, RI: American Mathematical Society, USA, 1998.

[2] M. Fabrizio and A. Morro. *Electromagnetism of Continuous Media*. Oxford University Press Inc., New York, 2003.

[3] J. Nečas. *Les méthodes directes en théorie des équations elliptiques*. Academia, Prague, 1967.

[4] M. M. Vainberg. *Variational method and method of monotone operators in the theory of nonlinear equations. Translated from Russian by A. Libin. Translation edited by D. Louvish.* A Halsted Press Book. New York-Toronto: John Wiley & Sons; Jerusalem- London: Israel Program for Scientific Translations. xi, 356 p. (1973), 1973.

[5] K. Yosida. *Functional analysis*. Classics in Mathematics. Springer Science & Business Media, New York, 1980.

[6] E. Zeidler. *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*. Springer-Verlag, 1986.

[7] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*. Springer, 1990.

[8] H. Brezis and F. Browder. *Partial Differential Equations in the 20th Century*. Advances in Mathematics, 135(1):76–144, 1998.

[9] Y. Eidelman, V. D. Milman, and A. Tsolomitis. *Functional Analysis: An Introduction*. Graduate studies in mathematics. American Mathematical Society, 2004.

[10] P. Šolín. *Partial Differential Equations and the Finite Element Method*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2005.

[11] O. Steinbach. *Numerical approximation methods for elliptic boundary value problems. Finite and boundary elements. Transl. from the 2003 German original*. New York, NY: Springer, 2008.

[12] S. C. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2008.

[13] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext (Berlin. Print). Springer, 2010.

[14] P. G. Ciarlet. *Linear and Nonlinear Functional Analysis with Applications:*. Applied mathematics. Society for Industrial and Applied Mathematics, 2013.

[15] O. A. Nielsen. *An Introduction to Integration and Measure Theory*. Wiley-Interscience and Canadian Mathematics Series of Monographs and Texts. Wiley, 1997.

[16] V. I. Bogachev. *Measure Theory*. Number v. 1 in Measure Theory. Springer, 2007.

[17] J. L. W. V. Jensen. *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*. Acta Mathematica, 30(1):175–193, 1906.

[18] Z. Pavić, J. Pečarić, and I. Perić. *Integral, discrete and functional variants of Jensen's inequality.* J. Math. Inequal., 5(2):253–264, 2011.

[19] T. H. Grönwall. *Note on the Derivatives with Respect to a Parameter of the Solutions of a System of Differential Equations*. Annals of Mathematics, 20(4):292–296, July 1919.

[20] D. Bainov and P. Simeonov. *Integral inequalities and applications*. Mathematics and Its Applications. East European Series. 57. Dordrecht, Kluwer Academic Publishers, 1992.

[21] C. Corduneanu. *Principles of differential and integral equations*. Chelsea Scientific Books. Allyn and Bacon, 1971.

[22] M. Lees. *A priori estimates for the solutions of difference approximations to parabolic partial differential equations*. Duke Mathematical Journal, 27(3):297–311, 09 1960. Duke University Press.

[23] R. B. Guenther and J. W. Lee. *Partial Differential Equations of Mathematical Physics and Integral Equations*. Dover books on mathematics. Dover Publications, United States of America, 1988.

[24] A. Kufner, O. John, and S. Fučík. *Function Spaces*. Monographs and textbooks on mechanics of solids and fluids. Noordhoff International Publishing, Leyden, 1977.

[25] F. Riesz and B. Sz.-Nagy. *Functional analysis*. Dover Books on Advanced Mathematics. Dover Publications, Inc., New York, 1990. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.

[26] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Verlag, 2006.

[27] E. Kreyszig. *Introductory Functional Analysis With Applications*. Wiley Classics Library. John Wiley & Sons, 1978.

[28] P. Grisvard. *Elliptic problems in nonsmooth domains*. Monographs and Studies in Mathematics, 24. Pitman Advanced Publishing Program. Boston-London-Melbourne: Pitman Publishing Inc. XIV, 410 p. (1985), 1985.

[29] J. Nečas. *Direct Methods in the Theory of Elliptic Equations*. Springer Monographs in Mathematics, 2011.

[30] W. Rudin. *Principles of mathematical analysis. 3rd ed.* International Series in Pure and Applied Mathematics. Düsseldorf etc.: McGraw-Hill Book Company. X, 342 p. DM 47.80, 1976.

[31] C. B. Jr . Morrey. *Functions of several variables and absolute continuity, II*. Duke Math. J., 6(1):187–215, 1940.

[32] E. Gagliardo. *Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili*. Rendiconti del Seminario Matematico della Universit di Padova, 27:284–305, 1957.

[33] S. L. Sobolev. *Applications of functional analysis in mathematical physics*. Translations of mathematical monographs. American Mathematical Society, 1963.

[34] N. G. Meyers and J. Serrin. $H = W$. Proc. Nat. Acad. Sci. U.S.A., 51:1055–1056, 1964.

[35] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Mathematische Nachrichten, 67(22):iv–iv, 1975.

[36] J. Kačur. *Method of Rothe in evolution equations*, volume 80 of *Teubner Texte zur Mathematik*. Teubner, Leipzig, 1985.

[37] A. J. Kurdila and M. Zabarankin. *Convex Functional Analysis*. Systems & Control: Foundations & Applications. Springer Science & Business Media, 2005.

[38] T. Roubíček. *Nonlinear partial differential equations with applications*, volume 153 of *ISNM*. Birkhäuser Verlag, Basel Boston Berlin, 2005.

[39] P. Monk. *Finite Element Methods for Maxwell's Equations*. Oxford University Press Inc., New York, 2003.

[40] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000.

[41] C. Baiocchi and A. Capelo. *Variational and quasivariational inequalities: applications to free boundary problems*. A Wiley-Interscience publication. Wiley, 1984.

[42] M. Grimmonprez and M. Slodička. *Reconstruction of an unknown source parameter in a semilinear parabolic problem*. Journal of Computational and Applied Mathematics, 289:331 – 345, 2015. Sixth International Conference on Advanced Computational Methods in Engineering (ACOMEN 2014).

[43] Y. V. Egorov, R. Cooke, and A. Shubin. *Foundations of the Classical Theory of Partial Differential Equations*. Number v. 1 in Classics in mathematics. Springer Berlin Heidelberg, 1998.

[44] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault. *Vector potentials in three-dimensional non-smooth domains*. Mathematical Methods in the Applied Sciences, 21(9):823–864, 1998.

[45] J. Saranen. *On a inequality of Friedrichs*. Mathematica Scandinavica, 51:310–322, 1982.

[46] G. Duvaut and J.-L. Lions. *Inéquations en thermoélasticité et magnéto-hydrodynamique*. Arch. Rational Mech. Anal., 46:241–279, 1972.

[47] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations. Theory and algorithms. (Extended version of the 1979 publ.)*. Springer Series in Computational Mathematics, 5. Berlin etc.: Springer-Verlag. X, 374 p. DM 198.00 (1986)., 1986.

[48] M. Costabel. *A remark on the regularity of solutions of Maxwell's equations on Lipschitz domains*. Math. Methods Appl. Sci., 12(4):365–368, 1990.

[49] G. Palatucci, O. Savin, and E. Valdinoci. *Local and global minimizers for a variational energy involving a fractional norm*. Annali di Matematica Pura ed Applicata, 192(4):673–718, 2013.

[50] Y. Pinchover and J. Rubinstein. *An Introduction to Partial Differential Equations*. Number v. 10 in An introduction to partial differential equations. Cambridge University Press, 2005.

[51] J. Hadamard. *Sur les problèmes aux dérivées partielles et leur signification physique*. Princeton University Bulletin, 13:49–52, 1902.

[52] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*, volume 375 of *Mathematics and Its Applications*. Kluwer Academic Publishers, Dordrecht, 1996.

[53] V. Isakov. *Inverse source problems*. Providence, RI: American Mathematical Society, 1990.

[54] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[55] G. J. Minty. *On a monotonicity method for the solution of nonlinear equations in Banach spaces*. Proceedings of the National Academy of Sciences of the United States of America, 50(6):1038–1041, December 1963.

[56] F. E. Browder. *Nonlinear Elliptic Boundary Value Problems. II*. Transactions of the American Mathematical Society, 117:530–550, May 1965.

[57] M. Slodička and S. Dehilis. *A nonlinear parabolic equation with a nonlocal boundary term*. Journal of Computational and Applied Mathematics, 233(12):3130–3138, 2010. Finite Element Methods in Engineering and Science (FEMTEC 2009).

[58] E. Rothe. *Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben*. Mathematische Annalen, 102(1):650–670, 1930.

[59] K. Rektorys. *The method of discretization in time and partial differential equations. Transl. from the Czech by the author.* Mathematics and Its Applications (East European Series), Vol. 4. Dordrecht - Boston - London: D. Reidel Publishing Company; Prague: SNTL - Publishers of Technical Literature, 1982.

[60] P. Brenner and V. Thomée. *On Rational Approximations of Semigroups*. SIAM Journal on Numerical Analysis, 16(4):683–694, 1979.

[61] A. Kufner. *Weighted Sobolev spaces*. Teubner-Texte zur Mathematik. Teubner, 1981.

[62] J.-P. Aubin. *Un théorème de compacité*. C. R. Acad. Sci. Paris, 256:5042–5044, 1963.

[63] J. L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Etudes mathématiques. Paris: Dunod; Paris: Gauthier-Villars. XX, 554 p, 1969.

[64] J. Cea. *Approximation variationnelle des problèmes aux limites*. Annales de l'institut Fourier, 14(2):345–444, 1964.

[65] H. Kardestuncer and D. H. Norrie, editors. *Finite Element Handbook*. McGraw-Hill, Inc., New York, NY, USA, 1987.

[66] H. Edelsbrunner. *Geometry and Topology for Mesh Generation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2001.

[67] J. C. Nedelec. *Mixed finite elements in* $\mathbb{R}^3$. Numerische Mathematik, 35(3):315–341, 1980.

[68] C. Paniconi and M. Putti. *A comparison of Picard and Newton iteration in the numerical solution of multidimensional variably saturated flow problems*. Water Resources Research, 30(12):3357–3374, 1994.

[69] A. Logg and G. N. Wells. *DOLFIN: Automated Finite Element Computing*. ACM Transactions on Mathematical Software, 37(2), 2010.

[70] A. Logg, G.N. Wells, and J. Hake. *DOLFIN: a C++/Python Finite Element Library*, chapter 10. Springer, 2012.

[71] A. Logg, K-A. Mardal, G. N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.

[72] H. Kamerlingh Onnes. *Further experiments with liquid helium. C. On the change of electric resistance of pure metals at very low temperatures etc. IV. The resistance of pure mercury at helium temperatures*. pages 1274–1276, Amsterdam, 1911. KNAW. 13 II.

[73] H. Kamerlingh Onnes. *Further experiments with Liquid Helium. D. On the change of Electrical Resistance of Pure Metals at very low Temperatures, etc. V. The Disappearance of the resistance of mercury*. pages 113–115, Amsterdam, 1911. KNAW. 13 II.

[74] H. Kamerlingh Onnes. *Further experiments with Liquid Helium G. On the electrical resistance of Pure Metals etc. VI. On the Sudden Change in the Rate at which the Resistance of Mercury Disappears*. pages 818–821, Amsterdam, 1912. KNAW. 14 II.

[75] M. Tinkham. *Introduction to Superconductivity*. McGraw-Hill Inc., New York, 2d ed. edition, 1996.

[76] J. F. Annett. *Superconductivity, Superfluids and Condensates*. Oxford Master Series in Physics. OUP Oxford, 2004.

[77] M. Sjöström. *Hysteresis Modelling of High Temperature Superconductors*. PhD thesis, Swiss Federal Institute of Technology of Lausanne, 2011.

[78] B. Xu. *Magnetic, Calorimetric and Electronic Properties of Vortex States in (3D) Mesoscopic Superconductors*. PhD thesis, University of Antwerp, 2011.

[79] S. J. Chapman. *Macroscopic Models of Superconductivity*. PhD thesis, St. Catherines College, Oxford, 1991.

[80] E. Janiková. *On numerical methods for diffusion of electric fields in type-II superconductors*. PhD thesis, Ghent University, 2008.

[81] W. Meissner and R. Ochsenfeld. *Ein neuer Effekt bei Eintritt der Supraleitfähigkeit*. Naturwissenschaften, 21(44):787–788, 1933.

[82] C. J. Gorter and H. Casimir. *Zur Thermodynamik des Supraleitenden Zustandes*. In Archives du Muse Teyler, pages 55–60. Springer Netherlands, 1935.

[83] S. Fujita and G. Salvador. *Theory of high temperature superconductivity*. Dordrecht: Kluwer academic, 2001.

[84] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. *Microscopic Theory of Superconductivity*. Phys. Rev., 106:162–164, Apr 1957.

[85] J. G. Bednorz and K. A. Müller. *Possible high T_c superconductivity in the Ba-La-Cu-O system*. Zeitschrift fur Physik B Condensed Matter, 64:189–193, June 1986.

[86] F. London and H. London. *The Electromagnetic Equations of the Supraconductor*. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 149(866):71–88, 1935.

[87] A. B. Pippard. *An Experimental and Theoretical Study of the Relation between Magnetic Field and Current in a Superconductor*. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 216(1127):547–568, 1953.

[88] J. M. Ziman. *Principles of the Theory of Solids*. Cambridge University Press, second edition, 1972. Cambridge Books Online.

[89] A. C. Eringen. *Electrodynamics of memory-dependent nonlocal elastic continua*. J. Math. Phys., 25:3235–3249, 1984.

[90] A. Suter, E. Morenzoni, N. Garifianov, R. Khasanov, E. Kirk, H. Luetkens, T. Prokscha, and M. Horisberger. *Observation of nonexponential magnetic penetration profiles in the Meissner state: A manifestation of nonlocal effects in superconductors*. Phys. Rev. B, 72:024506, July 2005.

[91] V. F. Kozhevnikov, C. V. Giuraniuc, M. J. Van Bael, K. Temst, C. Van Haesendonck, T. M. Mishonov, T. Charlton, R. M. Dalgliesh, Yu. N. Khaidukov, Yu. V. Nikitenko, V. L. Aksenov, V. N. Gladilin, V. M. Fomin, J. T. Devreese, and J. O. Indekeu. *Evidence for nonmonotonic magnetic field penetration in a type-I superconductor*. Phys. Rev. B, 78:012502, July 2008.

[92] T. Van Cauteren and J. O. Indekeu. *Non-local electrodynamics effect on surface superconductivity*. EPL (Europhysics Letters), 94(4):46001, 2011.

[93] V. Kozhevnikov, A. Suter, H. Fritzsche, V. Gladilin, A. Volodin, T. Moorkens, M. Trekels, J. Cuppens, B. M. Wojek, T. Prokscha, E. Morenzoni, G. J. Nieuwenhuys, M. J. Van Bael, K. Temst, C. Van Haesendonck, and J. O. Indekeu. *Nonlocal effect and dimensions of Cooper pairs measured by low-energy muons and polarized neutrons in type-I superconductors*. Phys. Rev. B, 87:104508, Mar 2013.

[94] C. P. Bean. *Magnetization of High-Field Superconductors*. Rev. Mod. Phys., 36:31–39, Jan 1964.

[95] A. Bossavit. *Numerical modelling of superconductors in three dimensions: a model and a finite element method*. Magnetics, IEEE Transactions on, 30(5):3363–3366, Sep 1994.

[96] L. Prigozhin. *The Bean Model in Superconductivity: Variational Formulation and Numerical Solution*. Journal of Computational Physics, 129(1):190– 200, 1996.

[97] J. W. Barrett and L. Prigozhin. *Bean's critical-state model as the $p \to \infty$ limit of an evolutionary p-Laplacian equation*. Nonlinear Anal. Theory Methods Appl, 6:977–993, 2000.

[98] H.-M. Yin, B. Li, and J. Zou. *A degenerate evolution system modeling Bean's critical-state type-II superconductors*. Discrete and Continuous Dynamical Systems, 8:781–794, 2002.

[99] L. Prigozhin and V. Sokolovsky. *AC Losses in Type-II Superconductors Induced by Nonuniform Fluctuations of External Magnetic Field*. IEEE Transactions on Applied Superconductivity, 14(1):69–81, 2004.

[100] W. Wei and H.-M. Yin. *Numerical solutions to Bean's critical-state model for type-II superconductors*. International Journal of Numerical Analysis and Modeling, 2(4):479–488, 2005.

[101] I. D. Mayergoyz. *Nonlinear Diffusion of Electromagnetic Fields: With Applications to Eddy Currents and Superconductivity*. Electromagnetism. Elsevier Science, United States of America, 1998.

[102] J. Rhyner. *Magnetic properties and AC-losses of superconductors with power law currentvoltage characteristics*. Physica C: Superconductivity, 212(3-4):292 – 300, 1993.

[103] E. H. Brandt. *Superconductors of finite thickness in a perpendicular magnetic field: Strips and slabs*. Phys. Rev. B, 54:4246–4264, Aug 1996.

[104] A. Gurevich and M. Friesen. *Nonlinear transport current flow in superconductors with planar obstacles*. Phys. Rev. B, 62:4004–4025, Aug 2000.

[105] C. M. Elliott and Y. Kashima. *On macroscopic critical state models of superconductivity in 3D*. In Proceedings of Czech-Japanese Seminar in Applied Mathematics 2005, Kuju, Japan, September 15–18, 2005., pages 39–55. Fukuoka: Kyushu University, The 21st Century COE Program "DMHF", 2006.

[106] C. M. Elliott and Y. Kashima. *A finite-element analysis of critical-state models for type-II superconductivity in 3D*. IMA J. Numer. Anal., 27(2):293–331, 2007.

[107] M. Slodička and E. Janiková. *Convergence of the backward Euler method for type-II superconductors*. Journal of Mathematical Analysis and Applications, 342(2):1026 – 1037, 2008.

[108] E. Babaev and M. Speight. *Semi-Meissner state and neither type-I nor type-II superconductivity in multicomponent superconductors*. Phys. Rev. B, 72:180502, Nov 2005.

[109] J. Nagamatsu, N. Nakagawa, T. Muranaka, Y. Zenitani, and J. Akimitsu. *Superconductivity at 39K in magnesium diboride*. Nature, 410:63–64, 2001.

[110] V. Moshchalkov, M. Menghini, T. Nishio, Q. H. Chen, A. V. Silhanek, V. H. Dao, L. F. Chibotaru, N. D. Zhigadlo, and J. Karpinski. *Type-1.5 Superconductivity*. Phys. Rev. Lett., 102:117001, Mar 2009.

[111] T. Nishio, V. H. Dao, Q. Chen, L. F. Chibotaru, K. Kadowaki, and V. V. Moshchalkov. *Scanning SQUID microscopy of vortex clusters in multiband superconductors*. Phys. Rev. B, 81:020506, Jan 2010.

[112] J. Carlström, E. Babaev, and M. Speight. *Type-1.5 superconductivity in multiband systems: Effects of interband couplings*. Phys. Rev. B, 83:174509, May 2011.

[113] M. Silaev and E. Babaev. *Microscopic theory of type-1.5 superconductivity in multiband systems*. Phys. Rev. B, 84:094515, Sep 2011.

[114] E. Babaev, J. Carlström, J. Garaud, M. Silaev, and J. M. Speight. *Type-1.5 superconductivity in multiband systems: Magnetic response, broken symmetries and microscopic theory - A brief overview*. Physica C: Superconductivity, 479(0):2 – 14, 2012. Proceedings of {VORTEX} {VII} Conference.

[115] C. N. Varney, K. A. H. Sellin, Q.-Z. Wang, H. Fangohr, and E. Babaev. *Hierarchical structure formation in layered superconducting systems with multi-scale inter-vortex interactions*. Journal of Physics: Condensed Matter, 25(41):415702, 2013.

[116] J. Carlström, J. Garaud, and E. Babaev. *Semi-Meissner state and nonpairwise intervortex interactions in type-1.5 superconductors*. Phys. Rev. B, 84:134515, Oct 2011.

[117] M. Silaev and E. Babaev. *Microscopic derivation of two-component Ginzburg-Landau model and conditions of its applicability in two-band systems*. Phys. Rev. B, 85:134514, Apr 2012.

[118] M. Slodička and K. Van Bockstal. *A nonlocal parabolic model for type-I superconductors*. Numerical Methods for Partial Differential Equations, 30(6):1821–1853, 2014.

[119] K. Van Bockstal and M. Slodička. *Error estimates for the full discretization of a nonlocal parabolic model for type-I superconductors*. Journal of Computational and Applied Mathematics, 275(0):516–526, 2015.

[120] A. Alonso and A. Valli. *An optimal domain decomposition preconditioner for low-frequency time-harmonic Maxwell equations*. Math. Comp, 68:607–631, 1999.

[121] H. Ammari, A. Buffa, and J.-C Nédélec. *A Justification of Eddy Currents Model for the Maxwell Equations*. SIAM J. Appl. Math., 60(5):1805–1823, May 2000.

[122] F. M. Odeh. *On the integrodifferential equations of the nonlocal theory of superconductivity*. J. Math. Phys., 5(8):1168–1182, 1964.

[123] A. Bossavit. *Computational electromagnetism. Variational formulations, complementarity, edge elements.* Orlando, FL: Academic Press, 1998.

[124] A. Visintin. *Electromagnetic processes in doubly-nonlinear composites.* Commun. Partial Differ. Equations, 33(5):808–841, 2008.

[125] T. Roubíček and G. Tomassetti. *Ferromagnets with eddy currents and pinning effects: their thermodynamics and analysis.* Math. Models Methods Appl. Sci., 21(1):29–55, 2011.

[126] J. Mercer. *Functions of Positive and Negative Type, and their Connection with the Theory of Integral Equations*. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 209(441-458):415–446, 1909.

[127] G. Fasshauer. *Solving differential equations with radial basis functions: multilevel methods and smoothing*. Advances in Computational Mathematics, 11:139–159, 1999. 10.1023/A:1018919824891.

[128] I. J. Schoenberg. *Metric spaces and completely monotone functions*. Ann. of Math., 39(4):811–841, 1938.

[129] C. Berg. *Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity*. In Mateu, J. and Porcu, E., Positive Definite Functions: from Schoenberg to Space-Time Challenges. Department of Mathematics, Universitat Jaume I, Castellon, Spain, 2008.

[130] E. W. Cheney and W. A. Light. *A Course in Approximation Theory*, volume 101 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 2000.

[131] R. A. Adams. *Sobolev Spaces*, volume 65 of *Pure and Applied Mathematics*. Academic Press, New York, 1975.

[132] K. Van Bockstal and M. Slodička. *The well-posedness of a nonlocal hyperbolic model for type-I superconductors*. Journal of Mathematical Analysis and Applications, 421(1):697 – 717, 2015.

[133] P. Monk. *Analysis of a Finite Element Method for Maxwells Equations*. SIAM Journal on Numerical Analysis, 29(3):714–729, 1992.

[134] F. Assous, P. Degond, E. Heintze, P.-A. Raviart, and J. Segre. *On a Finite-Element Method for Solving the Three-Dimensional Maxwell Equations*. Journal of Computational Physics, 109(2):222–237, 1993.

[135] Ch. G. Makridakis and P. Monk. *Time-discrete finite element schemes for Maxwell's equations*. RAIRO Modél. Math. Anal. Numér., 29(2):171–197, 1995.

[136] P. Ciarlet, Jr and J. Zou. *Fully discrete finite element approaches for time-dependent Maxwell's equations*. Numerische Mathematik, 82(2):193–219, 1999.

[137] K. Van Bockstal and M. Slodička. *A macroscopic model for an intermediate state between type-I and type-II superconductivity*. Numerical Methods for Partial Differential Equations, 31(5):1551–1567, 2015.

[138] M. Slodička. *Nonlinear Diffusion in type-II Superconductors*. J. Comput. Appl. Math., 215(2):568–576, May 2008.

[139] M. Slodička and V. Zemanová. *Time-discretization scheme for quasi-static Maxwell's equations with a non-linear boundary condition.* J. Comput. Appl. Math., 216(2):514–522, 2008.

[140] M. A. Biot. *Theory of stress-strain relations in anisotropic viscoelasticity and relaxation phenomena.* J. Appl. Phys., 25:1385–1391, 1954.

[141] M. A. Biot. *Variational principles in irreversible thermodynamics with application to viscoelasticity.* Phys. Rev., II. Ser., 97:1463–1469, 1955.

[142] M. A. Biot. *Thermoelasticity and Irreversible Thermodynamics*. Journal of Applied Physics, 27(3):240–253, 1956.

[143] C. M. Dafermos. *On the existence and the asymptotic stability of solutions to the equations of linear thermoelasticity*. Arch. Rational Mech. Anal., 29:241–271, 1968.

[144] M. E Gurtin. *The linear theory of elasticity*. Encyclopedia of Physics, Vol. VIa/2: Mechanics of Solids. Springer-Verlag, Berlin, 1972.

[145] R. E. Nickell and J. L. Sackman. *Variational principles for linear coupled thermoelasticity*. Quart. Appl. Math., 26:11–26, 1968.

[146] A. Bermúdez and J. M. Viaño. *Étude de deux schémas numériques pour les équations de la thermoélasticité*. RAIRO Anal. Numér., 17(2):121–136, 1983.

[147] S. I. Chou and C. C. Wang. *Estimates of error in finite element approximate solutions to problems in linear thermoelasticity. I. Computationally coupled numerical schemes*. Arch. Rational Mech. Anal., 77(3):263–299, 1981.

[148] R. E. Nickell and J. L. Sackman. *Approximate solutions in linear, coupled thermoelasticity.* J. Appl. Mech., 35:255–266, 1968.

[149] A. Ženíšek. *Finite element methods for coupled thermoelasticity and coupled consolidation of clay*. RAIRO Anal. Numér., 18(2):183–205, 1984.

[150] J. Kačur and A. Ženíšek. *Analysis of approximate solutions of coupled dynamical thermoelasticity and related problems*. Apl. Mat., 31(3):190–223, 1986.

[151] M. Slodička. *Application of Rothe's method to evolution integrodifferential systems.* Commentat. Math. Univ. Carol., 30(1):57–70, 1989.

[152] M. Slodička. *Smoothing effect and regularity for evolution integrodifferential systems.* Commentat. Math. Univ. Carol., 30(2):303–316, 1989.

[153] A. E. Green and P. M. Naghdi. *A re-examination of the basic postulates of themomechanics*. Proc. Royal Society London A, 432(1885):171–194, 1991.

[154] L. H. Fatori and J. E. Muñoz Rivera. *Energy Decay for Hyperbolic Thermoelastic Systems of Memory Type*. Q. Appl. Math., LXI(3):441–458, September 2001.

[155] L. H. Fatori, E. Lueders, and J. E. Muñoz Rivera. *Tramsmission problem for hyperbolic thermoelastic systems*. Journal of Thermal Stresses, 26(7):739–763, 2003.

[156] S. Bargmann and P. Steinmann. *Theoretical and computational aspects of non-classical thermoelasticity*. Computer Methods in Applied Mechanics and Engineering, 196(1–3):516 – 527, 2006.

[157] J.-M. Wang and B.-Z. Guo. *On dynamic behavior of a hyperbolic thermoelastic system with memory type in terms of eigenfrequencies*. In American Control Conference, Minneapolis, MN, 2006.

[158] O. Weinmann. *The equations of thermoelasticity with time-dependent coefficients*. Journal of Mathematical Analysis and Applications, 350(1):81 – 99, 2009.

[159] J. E. Muñoz Rivera and Y. Qin. *Global Existence and Exponential Stability in One-dimensional Nonlinear Thermoelasticity with Thermal Memory*. Nonlinear Anal., 51(1):11–32, October 2002.

[160] A. G. Megrabov. *Forward and Inverse Problems for Hyperbolic, Elliptic and Mixed Type Equa*. 2003.

[161] A. N. Tikhonov. *On the stability of inverse problems.* C. R. (Dokl.) Acad. Sci. URSS, n. Ser., 39:176–179, 1943.

[162] A. N. Tikhonov. *Regularization of incorrectly posed problems.* Sov. Math., Dokl., 4:1624–1627, 1963.

[163] V. A. Morozov. *On the solution of functional equations by the method of regularization*. Soviet Math. Dokl., 7:414–417, 1966.

[164] C. Kravaris and J. H. Seinfeld. *Identification of spatially varying parameters in distributed parameter systems by discrete regularization*. Journal of Mathematical Analysis and Applications, 119(1-2):128–152, 1986.

[165] M. Bellassoued and M. Yamamoto. *Carleman estimates and an inverse heat source problem for the thermoelasticity system*. Inverse Problems, 27(1):015006, 2011.

[166] B. Wu and J. Liu. *Determination of an unknown source for a thermoelastic system with a memory effect*. Inverse Problems, 28(9):095012, 2012.

[167] B. Wu, J. Yu, and Z. Wang. *Uniqueness and stability of an inverse kernel problem for type-III thermoelasticity*. Journal of Mathematical Analysis and Applications, 402(1):242 – 254, 2013.

[168] K. Van Bockstal and M. Slodička. *Recovery of a space-dependent vector source in thermoelastic systems*. Inverse Problems in Science and Engineering, 23(6):956–968, 2015.

[169] Y. Qin. *Nonlinear parabolic-hyperbolic coupled systems and their attractors*. Basel: Birkhäuser, 2008.

[170] R. C. MacCamy and J. S. W. Wong. *Stability Theorems for Some Functional Equations*. Transactions of the American Mathematical Society, 164:1–37, 1972.

[171] M. Kirane and N.-E. Tatar. *A Nonexistence Result to a Cauchy Problem in Nonlinear One Dimensional Thermoelasticity*. Journal of Mathematical Analysis and Applications, 254(1):71 – 86, 2001.

[172] J. C. Oliveira and R. C. Charão. *Stabilization of a locally damped thermoelastic system.* Comput. Appl. Math., 27(3):319–357, 2008.

[173] L. Landweber. *An Iteration Formula for Fredholm Integral Equations of the First Kind*. American Journal of Mathematics, 73(3):615–624, 1951.

[174] V. M. Fridman. *Method of successive approximations for a Fredholm integral equation of the 1st kind*. Uspekhi Mat. Nauk, 11(1(67)):233–234, 1956.

[175] T. Johansson and D. Lesnic. *Determination of a Spacewise Dependent Heat Source*. J. Comput. Appl. Math., 209(1):66–80, December 2007.

[176] S. D'haeyer, B. T. Johansson, and M. Slodička. *Reconstruction of a spacewise-dependent heat source in a time-dependent heat diffusion process*. IMA Journal of Applied Mathematics, 79(1):33–53, 2014.

[177] X. Zhang and E. Zuazua. *Decay of solutions of the system of thermoelasticity of type-III*. Communications in Contemporary Mathematics, 05(01):25–83, 2003.

[178] J. L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications*, volume 181 of *Non-homogeneous Boundary Value Problems and Applications*. Springer Berlin Heidelberg, 1972.

[179] B. T. Johansson and D. Lesnic. *A procedure for determining a spacewise dependent heat source and the initial temperature*. Applicable Analysis, 87(3):265–276, 2008.

[180] M. Slodička and V. Melicher. *An iterative algorithm for a Cauchy problem in eddy-current modelling*. Appl. Math. Comput., 217(1):237–246, 2010.

[181] K. Van Bockstal and M. Slodička. *Recovery of a time-dependent heat source in one-dimensional thermoelasticity of type-III*. Inverse Problems in Science and Engineering (submitted), 2015.

[182] M. Slodička. *A parabolic inverse source problem with a dynamical boundary condition*. Applied Mathematics and Computation, 256(0):529–539, 2015.

[183] L. Šeliga and M. Slodička. *An inverse source problem for a damped wave equation with memory*. Journal of Inverse and Ill-posed Problems, 0:–, 2015.

[184] R. H. De Staelen and M. Slodička. *Reconstruction of a convolution kernel in a semilinear parabolic problem based on a global measurement*. Nonlinear Analysis: Theory, Methods & Applications, 112(0):43 – 57, 2015.

[185] R. H. De Staelen, K. Van Bockstal, and M. Slodička. *Error analysis in the reconstruction of a convolution kernel in a semilinear parabolic problem with integral overdetermination*. Journal of Computational and Applied Mathematics, 275(0):382 – 391, 2015.

[186] E. DiBenedetto. *Degenerate Parabolic Equations*. Springer-Verlag, New York, 1993.

[187] V. Vrábel' and M. Slodička. *An eddy current problem with a nonlinear evolution boundary condition*. Journal of Mathematical Analysis and Applications, 387(1):267 – 283, 2012.