



TECHNISCHE
UNIVERSITÄT
WIEN

DISSERTATION

Convolution Quadrature and Boundary Element Methods in wave propagation: a time domain point of view

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften unter der Leitung von

Univ.-Prof. Jens Markus Melenk, PhD

E101 - Institut für Analysis und Scientific Computing, TU Wien

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Dipl.-Ing. Alexander Rieder

Matrikelnummer: 0827386

Plunkergasse 9/2

1150 Wien

Diese Dissertation haben begutachtet:

1. **Assoc. Prof. Dr. Lehel Banjai**
School of Mathematical and Computer Sciences, Heriot-Watt University
2. **Prof. Dr. Francisco-Javier Sayas**
Department of Mathematical Sciences, University of Delaware
3. **Univ.-Prof. Jens Markus Melenk, PhD**
Institut für Analysis und Scientific Computing, Technische Universität Wien

Wien, am 31. Mai 2017

Kurzfassung

In dieser Arbeit wird untersucht, inwieweit sich Wellenphänomene mittels Randintegralmethoden approximieren lassen. Derartige Methoden zeichnen sich dadurch aus, dass anstelle einer partiellen Differentialgleichung eine Integralgleichung am Rand des Gebiets betrachtet wird. Ein Vorteil dieser Formulierung ist, dass dadurch Streuprobleme auf unbeschränkten Gebieten numerisch behandelt werden können. Für stationäre Probleme sind Randelementmethoden bereits eine etablierte alternative zu klassischen Finite Element Methoden. Um Randintegralmethoden für Zeitabhängige Probleme anwendbar zu machen bietet sich die Methode der Faltungsquadratur von Lubich [Lub88a] an. Diese Methode besitzt günstige Stabilitätseigenschaften und besitzt ein Äquivalenzprinzip zur Approximation einer Halbgruppe mittels eines passenden Zeitschrittverfahrens. Dieses Prinzip wird in dieser Arbeit ausgenutzt um das betrachtete Diskretisierungsschema zu analysieren und unterscheidet sich von der klassischen Herangehensweise, welche die Konvergenz mittels Abschätzungen im Laplace-Bereich zeigt. Diese reine Zeitbereichsmethode hat den Vorteil, dass die so erlangten Abschätzungen meist schärfer sind und mit weniger Regularitätsannahmen auskommen.

In dieser Arbeit werden drei unterschiedliche Modellprobleme betrachtet: die zeitabhängige Schrödingergleichung in \mathbb{R}^d , diskretisiert mittels einer Kombination von Finiten- und Randelementen, ein nichtlineares Streuproblem im Außenraum, gegeben durch die lineare Wellengleichung mit nichtlinearer Randbedingung, und ein Streuproblem für Kompositmaterialien mit nicht-konstanter Wellenzahl. Für diese Modellprobleme werden Fragen zur Konvergenz und Stabilität beantwortet.

Numerische Simulationen untermauern die theoretischen Resultate.

Abstract

In this thesis, we consider different classes of time dependent wave propagation problems, and investigate whether they can be efficiently approximated using boundary integral methods. The idea of these methods is to replace partial differential equations with an integral equation on the boundary of the domain of interest. One of the main advantages of this approach is that problems posed on unbounded domains can be handled without further difficulties. For stationary problems boundary integral methods are well established as an alternative to more classical finite element based methods. In order to treat time dependent problems, one possibility is to apply Lubich's method of Convolution Quadrature [Lub88a]. This approach has many favorable properties, including an equivalence principle, which relates the CQ approximation to the approximation of the underlying semigroup with an appropriate time-stepping scheme. In this work, we exploit this equivalence to analyze the discretization schemes under consideration. Our approach differs from the more standard way of treating time domain boundary integral equations, which relies on estimate in the Laplace domain in order to infer convergence results. The pure time-domain approach has the benefit of yielding stronger estimates with fewer regularity assumptions than the Laplace domain counterpart.

In this thesis, we consider three different model problems, namely the time dependent Schrödinger equation posed in \mathbb{R}^d , treated by a coupling of Finite- and Boundary Element Methods, a nonlinear scattering problem in the exterior domain consisting of the linear wave equation augmented by a nonlinear impedance boundary condition, and a scattering problem by a composite material characterized by a non-constant wave number. For all of these model problems we answer questions regarding convergence and stability of the discretization scheme.

Numerical simulations support the theoretical findings.

Acknowledgements

I would like to thank the many people, without whom this thesis would not have been possible. First and foremost, I want to thank my advisor Prof. Jens Markus Melenk who has been supporting me since my undergraduate days. He started my interest in numerical analysis and over the years has been very generous and patient with sharing his knowledge of a wide range of the mathematical landscape.

I want to express my gratitude to Lehel Banjai and Francisco-Javier Sayas, not only for reviewing this thesis but also for inviting me to visit them and collaborate at their home institutions. The fruits of these collaborations can be seen in this thesis.

Thanks also to my colleagues at the Institute for Analysis and Scientific Computing, most notably my office mate Markus Faustmann who was always willing to help with the intricacies of boundary element methods and helped proofread this thesis. He also provided a great work atmosphere in the office. Thanks also to Bernhard Stiftner, who has been a great friend and colleague since the early days of our studies together. I would also like to thank Prof. Dirk Praetorius for including me in the social circles of his workgroup. I was able to share the experience of working on my Phd with many other current or former grad students, including Michael Feischl, Thomas Führer, Gregor Gantner, Alexander Haberl, Michele Ruggeri, Stefan Schimanko and Carl-Martin Pfeiler who also provided a great social balance to the time spent on mathematics.

Special thanks are also in order to Ursula Schweigler who has been very helpful in easing the bureaucracy of the university.

I also want to thank my parents for their support and encouragement over the years.

Last but not least, I would like to thank the Austrian Science Fund(FWF) and the Republic of Austria for the financial support, most notably during my stays abroad, and also thank all the members of the Doctoral School “Dissipation and dispersion in nonlinear PDEs” (FWF grant W1245).

Vienna, May 31st 2017

Alexander Rieder

Contents

1	Introduction	1
1.1	Structure of this Dissertation	2
2	Background	5
2.1	General notation	5
2.2	Semigroups	6
2.2.1	Linear Semigroups	6
2.2.2	Nonlinear Semigroups	12
2.3	Function spaces	12
2.3.1	Interpolation spaces	13
2.3.2	Sobolev spaces	14
2.4	Triangulations	19
2.5	Boundary integral equations for the Helmholtz equation	21
2.5.1	The Helmholtz equation with matrix-valued wave number	23
3	Time-stepping and Convolution Quadrature	27
3.1	Multistep methods	27
3.2	Runge-Kutta methods	28
3.3	Convolution Quadrature	30
3.3.1	Multistep method based CQ	32
3.3.2	Runge-Kutta based CQ	34
3.4	Approximation of Semigroups via Runge-Kutta methods	36
3.4.1	Rational functions of operators	37
3.4.2	Stability of Runge-Kutta time-stepping	41
3.4.3	Convergence of Runge-Kutta methods for semigroups	43
4	FEM-BEM coupling for the Schrödinger equation	59
4.1	Model problem and notation	59
4.2	The discretization scheme	60
4.3	An equivalent formulation as an exotic semigroup	62
4.4	Analysis of the discretization scheme	67
4.4.1	Convergence with respect to space discretization	69
4.4.2	Convergence of the fully discrete scheme	78
4.5	Numerical results	81
5	CQ for the wave equation with nonlinear impedance boundary condition	83
5.1	Model problem	83
5.2	Discretization using boundary integrals	84
5.3	Well posedness of the discretization scheme	86

5.4	Convergence analysis	88
5.4.1	A related semigroup - semi-discretization in space	91
5.4.2	Time discretization analysis	101
5.4.3	Convergence of the fully discrete scheme	104
5.5	Numerical results	111
6	Scattering by composite media	115
6.1	Model problem	115
6.2	A multiply overlapped wave problem	116
6.3	Time domain boundary integral equations	119
6.3.1	One possible construction for \mathcal{X}_h and \mathcal{Y}_h	122
6.4	Analysis of the transmission problem - the semigroup setting	124
6.4.1	Checking the semigroup requirements	125
6.4.2	Convergence of the space discretization	128
6.5	Runge-Kutta discretization	129
6.5.1	Determining the value of μ	132
6.6	Convergence of the fully discrete scheme	134
6.7	Numerical examples	136
	Bibliography	141

1 Introduction

Many phenomena in nature can be modeled by the propagation of waves, from water waves to electromagnetic scattering to the quantum wave functions on the microscopic scale. Mathematically, these phenomena can be modeled by partial differential equations of hyperbolic type. In this thesis, we are concerned with deriving and analyzing ways to approximately solve these partial differential equations in a robust and efficient way.

A particular feature of many of these problems is that they are posed on an unbounded domain, either the full space \mathbb{R}^d or the exterior of some object $\mathbb{R}^d \setminus \bar{\Omega}$. The most widely used discretization methods, namely Finite Element Methods (FEM) and Finite Difference (FD) discretizations, cannot handle such domains in a straightforward way. To overcome this limitation, oftentimes an artificial bounded domain is introduced on which the problem is discretized using these methods. Then, some transparent boundary conditions are imposed on the artificially introduced boundary, such as Perfectly Matched Layer and Infinite Element Methods.

Another possible approach, which has gained interest over the recent years, is to use boundary integral methods to replace the differential equation on the unbounded domain by an integral equation on the boundary. Due to this reformulation, it is possible to handle unbounded domains without introducing additional difficulties. If the resulting integral equations are discretized using an approach similar to the Finite Element Method, i.e., by approximating the solution via piecewise polynomials on a triangulation of the boundary, this is referred to as the Boundary Element Method (BEM).

For stationary problems, e.g., the Laplace or Helmholtz equations, boundary element methods are fairly widespread and presented in the monographs [SS11; HW08; Ste08]. The treatment of transient problems using this approach is not yet as common, an overview can be found in [Say16]. When discretizing time domain boundary integral equations, there are two main approaches: space-time Galerkin methods and Convolution Quadrature (CQ). When employing a space-time Galerkin scheme, one directly discretizes the retarded potentials associated with the wave equation. This introduces the need for accurately computing singular integrals on non-standard shapes (namely, triangles intersected with time cones), which makes the method difficult to implement in a stable way (early works on the mathematical basis for these methods are [BH86; BD86]). The other common approach, which is the one taken in this thesis, is to use Lubich's Convolution Quadrature, as introduced in [Lub88a; Lub88b]. This approach has the benefit that it can be implemented easily by reusing boundary element libraries developed for the Helmholtz equation. Convolution Quadrature comes in two flavors, based on multistep and Runge-Kutta time stepping schemes. While the multistep kind, introduced in the original works by Lubich, only allows order up to two, the Runge-Kutta Convolution Quadrature, introduced in [LO93], can be implemented for an arbitrary convergence order. While it is most common to use constant timestep size when using CQ (this will also be the class of methods considered in this

thesis), there have also been recent developments to allow varying timesteps in order to accommodate problems with non-smooth solutions, cf. [LS13; LS15; SS17] for generalized CQ based on the implicit Euler method and [LS16] for the Runge-Kutta based version.

The Convolution Quadrature method has many favorable stability properties, but the methods lead to non-local problems of convolution type. In order to efficiently solve such problems, there have been many efforts to develop fast algorithms see [Ban10; BS09; BK14; HKS09]).

The analysis for Convolution Quadrature based methods is most commonly carried out in the Laplace domain. There, frequency explicit bounds lead to convergence estimates in a very black-box kind of approach (see [Lub88a; Lub88b; BL11; BLM11] for the general theory). These kind of estimates in the Laplace domain were already developed for simple wave propagation problems in the works by Bamberger and Ha-Duong (cf. [BH86; BD86]), but a much larger class of problems was not unlocked until [LS09]. There, a new approach to formulate the spatially discrete problem using non-standard Hilbert spaces was developed.

When passing through the Laplace domain for the analysis of the discretization scheme, one loses certain information. Namely, the regularity assumptions on the exact solution are unnecessarily restrictive and the dependence of the discretization error on the time interval under consideration are somewhat opaque. In order to overcome these difficulties there has recently been interest in bypassing the Laplace domain for the analysis and carrying out all estimates directly in the time domain ([BLS15a; HS16; Has+15]; see also [MR17; BR17] which form part of this thesis). This will be the approach taken in this thesis.

The “strictly in the time domain” approach for the discretization relies on applying the “exotic Hilbert space” approach by Laliena-Sayas in the context of C_0 -semigroups for the discretization in space, and an equivalence principle between the Convolution Quadrature approximations and the multistep or Runge-Kutta approximations of said underlying semigroup. Estimates on the convergence of the method are then inferred by applying the general theory of such time-stepping methods, as developed in [BT79; BCT82; Cro76; AP03].

1.1 Structure of this Dissertation

In *Chapter 2*, we collect some of the definitions and results needed for the following chapters. Most notably, we present the theory of C_0 -semigroups in the linear and nonlinear case, as well as the most important results on Sobolev spaces, which will feature prominently throughout the rest of the thesis. We then present basic results on Finite- and Boundary Element Methods, and generalize some well known results about boundary integrals for the Helmholtz equation to the case of a class of Helmholtz-like systems.

Chapter 3 then presents the multistep and Runge-Kutta methods and introduces the Convolution Quadrature discretization for convolution integrals. We then go on to present results on the stability and approximation quality of Runge-Kutta methods when applied in a semigroup setting. Most notably, we prove some new results on the convergence rates when considering certain difference quotients and discrete integrals of RK-approximations, which allows for norm bounds other than those of the underlying Banach space of the semigroup.

In *Chapter 4* we apply the general methodology presented previously to a concrete problem, namely a discretization of the Schrödinger equation by a Runge-Kutta method in time and a coupling of a Finite Element Method with a Boundary Element discretization as transparent boundary condition. We then analyze the resulting fully discrete scheme in terms of stability and convergence.

Chapter 5 showcases that time domain boundary integral methods can be used to discretize certain nonlinear problems. We consider the (linear) wave equation in the exterior domain coupled with a nonlinear boundary condition of impedance type. We present several convergence results, differing in the assumptions made on the underlying problem. Most notably we show unconditional convergence, as well as full convergence rates if the exact solution of the problem is sufficiently smooth.

Finally, *Chapter 6* deals with a different kind of scattering problem, in which the scatterer consists of a composite material with different wave speeds. Unlike Chapter 4 we use a pure Boundary Element based method. In this case, we prove rigorous *a priori* estimates for the discretization as well.

2 Background

2.1 General notation

We start by introducing some general notation used throughout the thesis. We write $\mathbb{R}_+ := \{t \in \mathbb{R} : t > 0\}$ for the positive real numbers, and $\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}$ as well as $\mathbb{C}_- := \{z \in \mathbb{C} : \operatorname{Re}(z) < 0\}$ for the complex half spaces.

For Banach spaces \mathcal{X} and \mathcal{Y} we write $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ for the space of all bounded linear operators between \mathcal{X} and \mathcal{Y} together with the shorthand $\mathcal{B}(\mathcal{X}) := \mathcal{B}(\mathcal{X}, \mathcal{X})$. On $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ we consider the operator norm given by

$$\|T\|_{\mathcal{B}(\mathcal{X}, \mathcal{Y})} := \sup_{x \in \mathcal{X}} \frac{\|Tx\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}.$$

The topological dual space of \mathcal{X} will be denoted by \mathcal{X}' , and we define the duality bracket as

$$\langle x', x \rangle_{\mathcal{X}' \times \mathcal{X}} := x'(x) \quad \text{for } x' \in \mathcal{X}' \text{ and } x \in \mathcal{X}.$$

We will also write Id for a generic identity operator, where it should be clear from context which spaces are involved. If A is a linear operator (or matrix) on a space \mathcal{X} , we often write $A - \lambda := A - \lambda \operatorname{Id}$ and define

$$\rho(A) := \left\{ \lambda \in \mathbb{C} : (A - \lambda)^{-1} \text{ exists in } \mathcal{B}(\mathcal{X}) \right\}$$

for the resolvent set and $\sigma(A) := \mathbb{C} \setminus \rho(A)$ for the spectrum. The inner product on a space \mathcal{X} will be denoted by $\langle \cdot, \cdot \rangle_{\mathcal{X}}$, using the convention that it is linear in the first and antilinear in the second. For two spaces \mathcal{X} and \mathcal{Y} we write $\mathcal{X} \hookrightarrow \mathcal{Y}$ to mean $\mathcal{X} \subseteq \mathcal{Y}$ and $\operatorname{Id} : \mathcal{X} \rightarrow \mathcal{Y}$ is bounded.

For Banach spaces \mathcal{X}, \mathcal{Y} with $\mathcal{X} \subseteq \mathcal{Y}$ we define the annihilator space

$$\mathcal{X}^\circ := \left\{ y' \in \mathcal{Y}' : \langle y', x \rangle_{\mathcal{Y}' \times \mathcal{Y}} = 0 \quad \forall x \in \mathcal{X} \right\}. \quad (2.1)$$

We will also often write $\overline{\mathcal{X}}$ for the topological closure of a set in a larger space \mathcal{Y} which should be clear from context. If we want to emphasize which norm is used, we write $\operatorname{clos}(\mathcal{X}, \|\cdot\|_{\mathcal{Y}})$.

For operators A and B we write $A \subseteq B$ if $\operatorname{dom}(A) \subseteq \operatorname{dom}(B)$ and $Ax = Bx$ for $x \in \operatorname{dom}(A)$. Also we write \overline{A} for the operator obtained by taking the closure of the graph of A .

We introduce notation for the spaces of p -times continuously differentiable functions with values in a Banach space \mathcal{X} . For a bounded interval $I \subseteq \mathbb{R}$, we write $C^p(I, \mathcal{X})$, and equip this space with the norm

$$\|u\|_{C^p(I, \mathcal{X})} := \sum_{\ell=0}^p \sup_{\tau \in I} \|u^{(\ell)}(\tau)\|_{\mathcal{X}}.$$

In the same vein, we also introduce the space of essentially bounded functions $L^\infty(I, \mathcal{X})$ with the norm $\|u\|_{L^\infty(I, \mathcal{X})} := \text{esssup}_{t \in I} \|u(t)\|_{\mathcal{X}}$, where esssup denotes the supremum up to a set of measure zero.

Throughout this thesis, C will denote a generic constant greater than 0, which may be different in each instance but will not depend on any of the principal quantities of interest like mesh- or timestep size. We usually clarify the dependencies of the constant within the context. For two quantities a, b , we also write $a \lesssim b$ to mean $a \leq Cb$, as well as $a \sim b$ to mean $a \lesssim b \lesssim a$.

2.2 Semigroups

The phenomena considered in this dissertation can all be formulated using the theory of operator semigroups. In this chapter we present the most important definitions. We spend most of the time on the case of linear problems, the case of nonlinear semigroups will then be handled in Section 2.2.2. For convenience, we collect the main results used throughout this thesis. To keep the presentation succinct, we state all the results without proof and refer the reader to the relevant literature.

2.2.1 Linear Semigroups

The theory of linear, strongly continuous semigroups (or C_0 -semigroups) is well developed and can be found in most textbooks on functional analysis, see e.g. [Yos80] or, for a more detailed treatment, see [Paz83]. In this section all spaces can either be real or complex valued.

Definition 2.1. *Let \mathcal{X} denote a real or complex Banach space. A family of operators $T(t) \in \mathcal{B}(\mathcal{X})$ for $t \geq 0$ is called a linear C_0 -semigroup given that the following conditions hold:*

(i) $T(0) = \text{Id}$,

(ii) $T(s+t) = T(s)T(t)$ for $s, t \geq 0$,

(iii) $T(t)$ is strongly continuous at 0, i.e. $\|T(t)x - x\|_{\mathcal{X}} \xrightarrow{t \rightarrow 0^+} 0 \quad \forall x \in \mathcal{X}$.

Definition 2.2 (infinitesimal generator). *Let $(T(t))_{t \geq 0}$ be a C_0 semigroup. The linear operator A defined via*

$$\text{dom}(A) := \left\{ x \in \mathcal{X} : \lim_{t \rightarrow 0^+} \frac{T(t)x - x}{t} \text{ exists} \right\} \quad (2.2)$$

$$Ax := \lim_{t \rightarrow 0^+} \frac{T(t)x - x}{t} \quad \text{for } x \in \text{dom}(A) \quad (2.3)$$

is called the infinitesimal generator (or just generator) of T .

The following proposition is the main reason why we are interested in C_0 -semigroups, namely they can be used to solve initial value problems.

Proposition 2.3 ([Paz83, Chapter 4, Theorem 1.3]). *A semigroup $T(\cdot)$ is uniquely determined by its infinitesimal generator A . For all $x \in \mathcal{X}$, the map $t \mapsto T(t)x$ is continuous. For $x \in \text{dom}(A)$, the function $u(t) := T(t)x$ is continuously differentiable and solves the initial value problem*

$$\dot{u}(t) := \frac{d}{dt}u(t) = Au(t) \quad \forall t > 0, \quad (2.4a)$$

$$u(0) = x. \quad (2.4b)$$

This proposition motivates the notation $e^{tA} := T(t)$ for the semigroup which emphasizes the importance of the generator A and generalizes the usual matrix exponential.

When discussing the existence of solutions to evolution equations by using semigroup theory, it is necessary to determine whether a given operator A is the generator of a semigroup. The following proposition gives a characterization of all generators.

Proposition 2.4 (Hille-Yosida, see [Paz83, Chapter 1, Theorem 5.3]). *Let A be a linear (unbounded) operator on a Banach space \mathcal{X} . A is the generator of a C_0 -semigroup $T(\cdot)$ on \mathcal{X} if and only if:*

(i) A is closed and densely defined, i.e. $\bar{A} = A$ and $\overline{\text{dom}(A)} = \mathcal{X}$,

(ii) there exist constants $\omega \geq 0$, $M \geq 1$, such that the resolvent set $\rho(A)$ satisfies $\{\lambda : \text{Re}(\lambda) > \omega\} \subseteq \rho(A)$ and powers of the resolvent can be bounded by:

$$\|(A - \lambda)^{-n}\|_{\mathcal{B}(\mathcal{X})} \leq \frac{M}{(\text{Re}(\lambda) - \omega)^n} \quad \forall \text{Re}(\lambda) > \omega, \forall n \in \mathbb{N}. \quad (2.5)$$

The semigroup then satisfies the following a priori estimate:

$$\|T(t)\|_{\mathcal{B}(\mathcal{X})} \leq M e^{\omega t} \quad \forall t \geq 0. \quad (2.6)$$

Corollary 2.5. *Let A be the generator of a C_0 semigroup $T(\cdot)$ satisfying (2.5). For initial conditions $u_0 \in \text{dom}(A)$ we define $u(t) := T(t)u_0$. Then, the following estimate holds:*

$$\|\dot{u}\|_{\mathcal{X}} \leq M e^{\omega t} \|Au_0\|_{\mathcal{X}}.$$

Proof. It is easy to calculate that (see [Paz83, Chapter 1, Theorem 2.4, c])

$$\dot{u}(t) = AT(t)u_0 = T(t)Au_0.$$

The estimate follows from the operator bound on $T(t)$ in Proposition 2.4. □

While Proposition 2.4 gives necessary and sufficient conditions on generators of semigroups, they are often not easy to verify. We will instead make use of a different set of conditions, at the heart of which lies the following definition:

Definition 2.6. Let A be a linear (unbounded) operator on a Banach space \mathcal{X} . We say A is dissipative, iff for every $x \in \text{dom}(A)$ there exists a functional $x' \in \mathcal{X}'$ such that

$$\langle x', x \rangle_{\mathcal{X}' \times \mathcal{X}} = \|x\|_{\mathcal{X}}^2 = \|x'\|_{\mathcal{X}'}^2 \quad \text{and} \quad \text{Re} \langle x', Ax \rangle_{\mathcal{X}' \times \mathcal{X}} \leq 0.$$

If \mathcal{X} is a Hilbert space, this condition can be simplified to

$$\text{Re} \langle Ax, x \rangle_{\mathcal{X}} \leq 0 \quad \forall x \in \text{dom}(A).$$

We call A maximally dissipative, if $\text{range}(A - \lambda_0) = \mathcal{X}$ for some $\lambda_0 > 0$.

Proposition 2.7 (Lumer-Phillips, [Paz83, Chapter 1, Theorem 4.3]). Let A be a linear operator on a Banach space \mathcal{X} with dense domain. Then, the following statements hold:

- (i) If A is maximally dissipative, then A is the generator of a C_0 -semigroup of contractions $T(t)$, i.e., the semigroup satisfies $\|T(t)\|_{\mathcal{B}(\mathcal{X})} \leq 1 \quad \forall t \geq 0$.
- (ii) If $T(t)$ is a semigroup of contractions, then A is maximally dissipative.

This implies that in this case, the resolvent bound (2.5) becomes:

$$\|(A - \lambda)^{-1}\|_{\mathcal{B}(\mathcal{X})} \leq \frac{1}{\text{Re}(\lambda)} \quad \forall \text{Re}(\lambda) > 0. \quad (2.7)$$

One last existence theorem tells us when an operator A generates a group instead of a semigroup, i.e., we can also evaluate $T(-t)$ with $T(t)T(s) = T(s+t)$ for arbitrary $s, t \in \mathbb{R}$. We first recall the definition of symmetric and self-adjoint operators:

Definition 2.8 (see e.g. [Paz83, Chapter 1, Definition 10.7]). A linear operator A on a Hilbert space \mathcal{H} is called symmetric, if $\text{dom}(A)$ is dense and $A \subseteq A^*$ where A^* is defined via

$$\langle Ax, y \rangle_{\mathcal{H}} = \langle x, A^*y \rangle_{\mathcal{H}} \quad \forall x \in \text{dom}(A)$$

and $\text{dom}(A^*) := \{y \in \mathcal{H} : A^*y \text{ exists}\}$. We call the operator A self-adjoint if $A^* = A$, i.e. A is symmetric and $\text{dom}(A) = \text{dom}(A^*)$.

We call a linear operator U unitary if it is isometric and bijective, i.e., $\|Ux\|_{\mathcal{H}} = \|x\|_{\mathcal{H}}$ and U^{-1} exists.

Proposition 2.9 (Stone, [Paz83, Chapter 1, Theorem 10.8]). Let A be a linear, densely defined operator on a Hilbert space \mathcal{H} . Then, A is the generator of linear group of unitary operators if and only if iA is self-adjoint. This is also equivalent to the fact that $\pm A$ are both maximally dissipative.

Proof. The first part is Stone's theorem, [Paz83, Chapter 1, Theorem 10.8]. The second part seems to be well known, but is usually not stated explicitly. Therefore we sketch a proof for completeness. We show that generating a unitary C_0 -group is equivalent to $\pm A$ being maximally dissipative. For a group of unitary operators $T(t)$ it is easy to see that $T(t)$ and $T(-t)$ are C_0 -semigroups of contractions and the generator of $T(-t)$ is $-A$. Proposition 2.7 then implies that $\pm A$ are maximally dissipative. On the other hand, if $\pm A$ are both

maximally dissipative, they generate semigroups of contractions $T(t)$ and $S(t)$. We first show $S(t) = T(t)^{-1}$, which can be easily seen by $[T(t)S(t)]' = AT(t)S(t) - T(t)AS(t) = 0$, since semigroups commute with their generator. Since $S(0) = T(0) = \text{Id}$ this implies $T(t)S(t) = \text{Id}$. We define the function

$$U(t) := \begin{cases} T(t) & t \geq 0 \\ S(-t) & t \leq 0. \end{cases}$$

It is easy to check that $(U(t))_{t \in \mathbb{R}}$ is a group of operators (see also [Paz83, Chapter 1, Lemma 6.4]). To see that $U(t)$ is unitary we note that $T(t)$ and $T^{-1}(t) = S(t)$ are contractions by Proposition 2.7 thus $U(t)$ is isometric for $t \geq 0$ and analogously for $t \leq 0$. Since $U(t)$ is also invertible this implies that $U(t)$ is unitary. \square

Semigroup theory can also be used for analyzing PDEs with an inhomogeneous right-hand side. To do so, we use the following variation of the well known Duhamel formula, which can easily be checked:

Proposition 2.10 (Duhamel's formula, [Paz83, Chapter 4, Corollary 2.11]). *Let $T(t)$ be a C_0 -semigroup with infinitesimal generator A on a reflexive Banach space \mathcal{X} . Let the given right hand side $f \in C([0, \infty), \mathcal{X})$ be locally Lipschitz continuous and also assume $u_0 \in \text{dom}(A)$. Define*

$$u(t) := T(t)u_0 + \int_0^t T(t-\tau)f(\tau) d\tau, \quad (2.8)$$

with the integral to be understood in the sense of Riemann. Then u solves the problem

$$\frac{d}{dt}u(t) - Au(t) = f(t) \quad \forall t > 0, \quad (2.9)$$

$$u(0) = u_0. \quad (2.10)$$

For $f \in C^1((0, \infty), \mathcal{X})$, the assumption that \mathcal{X} is reflexive is not needed and the following a priori bounds holds:

$$\begin{aligned} \|u(t)\|_{\mathcal{X}} &\leq Me^{\omega t} \left[\|u_0\|_{\mathcal{X}} + \int_0^t \|f(\tau)\|_{\mathcal{X}} d\tau \right] \\ \|\dot{u}(t)\|_{\mathcal{X}} &\leq Me^{\omega t} \left[\|Au_0\|_{\mathcal{X}} + \|f(0)\|_{\mathcal{X}} + \int_0^t \|\dot{f}(\tau)\|_{\mathcal{X}} d\tau \right]. \end{aligned}$$

Proof. Existence is the content of [Paz83, Chapter 4, Corollary 2.10 and 2.11]. The a priori bound on u follows from the definition in (2.8) and (2.6). For the bound on the derivative, we use the representation

$$\dot{u} = T(t)Au_0 + T(t)f(0) + \int_0^t T(t-\tau)\dot{f}(\tau) d\tau$$

(the term $T(t)Au_0$ corresponds to the homogeneous semigroup, see Corollary 2.5; the inhomogeneous part follows by a change of variables and the fundamental theorem of calculus, see the proof of [Paz83, Chapter 4, Corollary 2.5]). The bound then follows from the operator bound on $T(t)$ in (2.6). \square

In order to account for non-homogeneous boundary conditions in a semigroup framework, we will work with a lifting operator. The construction is taken from [Has+15]. Since it highlights a common construction used throughout this thesis we include a short proof.

Proposition 2.11. *Let A_\star be a closed linear operator on a reflexive Banach space \mathcal{X} , and let $B : \text{dom}(A_\star) \rightarrow \mathcal{M}$ be a bounded linear operator, where \mathcal{M} is a normed space, and $\text{dom}(A_\star)$ carries the graph norm $\|x\|_{A_\star} := \|x\|_{\mathcal{X}} + \|A_\star x\|_{\mathcal{X}}$. We make the following assumptions:*

- (i) *the operator $A := A_\star|_{\ker(B)}$ generates a C_0 -semigroup,*
- (ii) *B is surjective and has a bounded right-inverse denoted by $\mathcal{E}_B : \mathcal{M} \rightarrow \text{dom}(A_\star)$ (again using the graph norm), which in addition satisfies $A_\star \circ \mathcal{E}_B = \mathcal{E}_B$,*
- (iii) *the data satisfies $f \in C^1([0, \infty), \mathcal{X})$, $\Xi \in C^2((0, \infty), \mathcal{M}) \cap C^0([0, \infty), \mathcal{M})$.*

Then, the initial value problem of finding $u \in C^1((0, \infty), \mathcal{X}) \cap C^0([0, \infty), \mathcal{X})$, such that

$$\dot{u}(t) = A_\star u(t) + f(t) \quad \forall t > 0, \quad (2.11a)$$

$$[Bu](t) = \Xi(t) \quad \text{and} \quad u(0) = u_0 \quad (2.11b)$$

has a unique solution for all $u_0 \in \text{dom}(A_\star)$ with $Bu_0 = \Xi(0)$.

Assume that the initial conditions satisfy $u_0 \in \text{dom}(A)$, $f(0) = 0$ as well as $\Xi(0) = \dot{\Xi}(0) = 0$. Then the solution satisfies the following a priori estimates:

$$\|u(t)\|_{\mathcal{X}} \lesssim e^{\omega t} \left[\|u_0\|_{\mathcal{X}} + \int_0^t \|\Xi(\tau)\|_{\mathcal{M}} + \|\dot{\Xi}(\tau)\|_{\mathcal{M}} + \|f(\tau)\|_{\mathcal{X}} d\tau \right] \quad (2.12a)$$

$$\|\dot{u}(t)\|_{\mathcal{X}} \lesssim e^{\omega t} \left[\|Au_0\|_{\mathcal{X}} + \int_0^t \|\dot{\Xi}(\tau)\|_{\mathcal{M}} + \|\ddot{\Xi}(\tau)\|_{\mathcal{M}} + \|\dot{f}(\tau)\|_{\mathcal{X}} d\tau \right]. \quad (2.12b)$$

Here, ω is the constant in (2.5), and the implied constants depend on M from (2.5) and the operator norm of \mathcal{E}_B .

Proof. By Proposition 2.10, there exists a solution to the problem with homogeneous boundary values

$$\begin{aligned} \dot{u}_{\text{hom}}(t) &= Au_{\text{hom}}(t) + f(t) + \mathcal{E}_B \Xi(t) - \mathcal{E}_B \dot{\Xi}(t) \quad \forall t > 0, \\ u_{\text{hom}}(0) &= u_0 - \mathcal{E}_B \Xi(0), \end{aligned}$$

(note that we have $Bu_0 - B\mathcal{E}_B \Xi(0) = Bu_0 - \Xi(0) = 0$ by assumption and therefore $u_{\text{hom}}(0) \in \text{dom}(A)$). By defining $u(t) := u_{\text{hom}}(t) + \mathcal{E}_B \Xi(t)$ for $t \geq 0$ we get that u solves:

$$\begin{aligned} \dot{u}(t) &= \dot{u}_{\text{hom}}(t) + \mathcal{E}_B \dot{\Xi}(t) = Au_{\text{hom}}(t) + \mathcal{E}_B \Xi(t) + f(t) \\ &= A_\star [u_{\text{hom}}(t) + \mathcal{E}_B \Xi(t)] + f(t) \\ &= A_\star u(t) + f(t). \end{aligned}$$

By definition u also satisfies $[Bu](t) = B[u_{\text{hom}}(t)] + B[\mathcal{E}_B \Xi(t)] = 0 + \Xi(t) = \Xi(t)$. To see uniqueness we note that the difference of two solutions satisfies the homogeneous equation and we can therefore apply the uniqueness of semigroups (Proposition 2.3).

The *a priori* bounds then follow from the bounds on u_{hom} in Proposition 2.10 and the continuity of \mathcal{E}_B , where the terms due to initial values of the right-hand side vanish by assumption. We absorb the term $\|\mathcal{E}_B[\Xi(t)]\|_{\mathcal{X}}$ in the integral of $\|\dot{\Xi}\|_{\mathcal{M}}$ for simpler presentation. \square

Remark 2.12. *In practice \mathcal{M} will be a Sobolev space on the boundary of a bounded domain (or a subspace thereof) and B will be some combination of trace operators (see Section 2.3.2).*

We end with a small lemma, relating the operator A_\star to the time derivative. This will be useful when estimating errors with regards to stronger spatial norms.

Lemma 2.13. *Let u denote the solution from Proposition 2.11. Assume that the right-hand side satisfies*

$$(i) \ f(t) \in \text{dom}(A_\star) \ \forall t \geq 0, \text{ and } A_\star f \in C^1([0, T], \mathcal{X}),$$

$$(ii) \ Bf \in C^2([0, T], \mathcal{M}),$$

$$(iii) \ \Xi \in C^3([0, T], \mathcal{M}),$$

$$(iv) \ A_\star u_0 \in \text{dom}(A_\star),$$

$$(v) \ Bu_0 = \Xi(0) = 0 \text{ and } B[A_\star u_0] = \dot{\Xi}(0) - Bf(0).$$

Then, the function $v(t) := A_\star u(t)$ solves:

$$\dot{v}(t) = A_\star v(t) + A_\star f(t), \tag{2.13a}$$

$$B[v(t)] = \dot{\Xi}(t) - B[f(t)] \quad \text{and} \quad v(0) = A_\star u_0. \tag{2.13b}$$

Proof. We define $w(t)$ as the solution of (2.13), which exists by Proposition 2.11. Define the function

$$y(t) := u_0 + \int_0^t w(\tau) + f(\tau) \, d\tau.$$

By construction, we have $\dot{y}(t) = w(t) + f(t)$. Thus, if we are able to show $y(t) = u(t)$ we get $v(t) = A_\star u(t) = \dot{y}(t) - f(t) = w(t)$. We show this by using the fact that solutions of (2.11) are unique. We calculate:

$$\begin{aligned} A_\star y(t) &= A_\star u_0 + \int_0^t A_\star w(\tau) + A_\star f(\tau) \, d\tau \\ &= A_\star u_0 + \int_0^t \dot{w}(\tau) \, d\tau = w(t) \\ &= \dot{y}(t) - f(t) \end{aligned}$$

(we can exchange the integral with the operator A_\star since the operator is assumed closed and both $w + f$ and $A_\star(w + f)$ are integrable by the *a priori* bounds (2.12)). For the constraint we get

$$B[y(t)] = Bu_0 + \int_0^t \dot{\Xi}(\tau) \, d\tau = \Xi(t)$$

since $Bu_0 = \Xi(0) = 0$ by assumption. Therefore, y solves (2.11), which gives $y = u$ and $w = A_\star u$. \square

2.2.2 Nonlinear Semigroups

When dealing with nonlinear evolution problems, one can hope to retain some of the language and results of Section 2.2.1. The presentation in this chapter mostly follows [Sho97]. We focus on the case that \mathcal{H} is a Hilbert space.

Following [Kōm67], we make the following generalization of a contraction semigroup:

Definition 2.14. *Let \mathcal{H} be a Hilbert space. A family of (nonlinear) operators $T(t) : \mathcal{H} \rightarrow \mathcal{H}$ for $t \geq 0$ is called a nonlinear contraction semigroup iff:*

- (i) For any fixed $t \geq 0$, $T(t)$ is a continuous (nonlinear) operator defined on \mathcal{H} into \mathcal{H} ,
- (ii) for any fixed $x \in \mathcal{H}$, the map $t \mapsto T(t)x$ is continuous in t ,
- (iii) $T(t+s) = T(t)T(s)$ for all $s, t \geq 0$ and $T(0) = \text{Id}$,
- (iv) $\|T(t)x - T(t)y\|_{\mathcal{H}} \leq \|x - y\|$ for every $x, y \in \mathcal{H}$ and $\forall t \geq 0$.

The nonlinear setting has an analog to the Lumer-Phillips theorem, using the following definition as starting point:

Definition 2.15. *Let \mathcal{H} be a Hilbert space and $A : \text{dom}(A) \subseteq \mathcal{H} \rightarrow \mathcal{H}$ be a (not necessarily linear or continuous) operator with domain $\text{dom}(A)$. We call A maximally monotone if it satisfies:*

- (i) $\langle Ax - Ay, x - y \rangle_{\mathcal{H}} \leq 0 \quad \forall x, y \in \text{dom } A$,
- (ii) $\text{range}(A - \text{Id}) = \mathcal{H}$.

Remark 2.16. *We follow the notation used in [Gra12]. Other authors, e.g. [Nev78] work with $-A$ instead. ■*

Proposition 2.17 (Kōmura-Kato, [Sho97, Proposition 3.1]). *Let A be a maximally monotone operator on a Hilbert space \mathcal{H} with domain $\text{dom}(A)$.*

For each $u_0 \in \text{dom}(A)$ there exists a unique absolutely continuous function $u : [0, \infty) \rightarrow \mathcal{H}$, which is differentiable almost everywhere and satisfies:

$$\dot{u} = Au \quad \text{and} \quad u(0) = u_0 \tag{2.14}$$

almost everywhere in t . In addition, u is Lipschitz continuous with

$$\|\dot{u}\|_{L^\infty((0, \infty); \mathcal{H})} \leq \|Au_0\|_{\mathcal{H}}$$

and $u(t) \in \text{dom}(A)$ for all $t \geq 0$. The family of operators $T(t) : \overline{\text{dom}(A)} \rightarrow \mathcal{H}$, defined as $T(t)u_0 := u(t)$, where $u(t)$ is the solution to (2.14), is called the nonlinear contraction semigroup generated by A .

2.3 Function spaces

In this thesis, we will need several classes of function spaces. The goal of this section is to introduce the spaces, collect their properties as needed for the subsequent results and present the notation used.

2.3.1 Interpolation spaces

In a lot of cases, the regularity of functions can be characterized as somewhere between two well-behaved function spaces. In order to formalize this idea, we use the concept of real interpolation of Banach spaces. See [Tar07; Tri95] or [McL00, Appendix B] for details and proofs.

Let \mathcal{X}_0 and \mathcal{X}_1 denote two Banach spaces with continuous embedding $\mathcal{X}_1 \subseteq \mathcal{X}_0$. For a parameter $\theta \in (0, 1)$ and $p \in [1, \infty]$, we define the K -functional and the interpolation norm as

$$\begin{aligned} K(t; u) &:= \inf_{v \in \mathcal{X}_1} (\|u - v\|_{\mathcal{X}_0} + t \|v\|_{\mathcal{X}_1}), \\ \|u\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\theta, p}} &:= \left(\int_0^\infty \left(t^{-\theta} K(t, u) \right)^p \frac{dt}{t} \right)^{1/p}, \quad \text{for } 1 \leq p < \infty, \\ \|u\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\theta, \infty}} &:= \text{esssup}_{t \in (0, \infty)} \left(t^{-\theta} K(t, u) \right), \end{aligned}$$

where esssup denotes the supremum up to a set of measure zero. We will mostly focus on the case of $p = 2$. For the cases $\theta = 0, 1$, we use the notational convention that $[\mathcal{X}_0, \mathcal{X}_1]_{\theta, p} := \mathcal{X}_\theta$.

The following proposition is the main reason, why we are interested in interpolation spaces:

Proposition 2.18 ([Tar07, Lemma 22.3]). *Consider two pairs of Banach spaces $\mathcal{X}_1 \subseteq \mathcal{X}_0$ and $\mathcal{Y}_1 \subseteq \mathcal{Y}_0$. Let $T : \mathcal{X}_0 \rightarrow \mathcal{Y}_0$ be a linear operator that is bounded for both pairs of spaces $\mathcal{X}_0 \rightarrow \mathcal{Y}_0$ and $\mathcal{X}_1 \rightarrow \mathcal{Y}_1$. Then, T is also a bounded operator mapping $[\mathcal{X}_0, \mathcal{X}_1]_{\theta, p} \rightarrow [\mathcal{Y}_0, \mathcal{Y}_1]_{\theta, p}$.*

The operator norm can be bounded by

$$\|T\|_{\mathcal{B}([\mathcal{X}_0, \mathcal{X}_1]_{\theta, p}, [\mathcal{Y}_0, \mathcal{Y}_1]_{\theta, p})} \leq \|T\|_{\mathcal{B}(\mathcal{X}_0, \mathcal{Y}_0)}^{1-\theta} \|T\|_{\mathcal{B}(\mathcal{X}_1, \mathcal{Y}_1)}^\theta \quad \forall \theta \in [0, 1]. \quad (2.15)$$

Another simple, but important property is the following:

Proposition 2.19 ([McL00, Lemma B.1]). *If $x \in \mathcal{X}_1$, then the interpolation norm can be estimated by*

$$\|x\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\theta, p}} \leq \|x\|_{\mathcal{X}_0}^{1-\theta} \|x\|_{\mathcal{X}_1}^\theta \quad \forall \theta \in [0, 1].$$

Most importantly, since $\mathcal{X}_1 \hookrightarrow \mathcal{X}_0$, we have:

$$\|x\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\theta, p}} \lesssim \|x\|_{\mathcal{X}_1} \quad \forall \theta \in [0, 1].$$

Often, it becomes necessary to interpolate pairs (or more generally tuples) of spaces. This can be done by the following proposition:

Proposition 2.20 ([Tri95, Sect. 1.18.1]). *Let $(\mathcal{X}_0^i)_{i=0}^N$ and $(\mathcal{X}_1^i)_{i=0}^N$ be tuples of Banach spaces satisfying $\mathcal{X}_1^i \subseteq \mathcal{X}_0^i$ with continuous embedding. Consider the product spaces $\mathcal{X}_0 :=$*

$\prod_{i=0}^N \mathcal{X}_0^i$ and $\mathcal{X}_1 := \prod_{i=0}^N \mathcal{X}_1^i$ with the norm $\|\cdot\|_{\mathcal{X}_j} := \left(\sum_{i=0}^N \|\cdot\|_{\mathcal{X}_j^i}^p \right)^{1/p}$ for $j = 0, 1$ and some parameter $p \in [1, \infty)$. Then,

$$[\mathcal{X}_0, \mathcal{X}_1]_{\theta, p} = \prod_{i=0}^N [\mathcal{X}_0^i, \mathcal{X}_1^i]_{\theta, p},$$

where the product is again equipped with the corresponding ℓ^p -norm and the equality is in the sense of equivalent norms.

2.3.2 Sobolev spaces

In this section, we introduce the usual Sobolev spaces, the main reference on the topic is [AF03], other books that give good introductions include [McL00] and [Tar07].

Sobolev spaces on Lipschitz domains

In this section $\Omega \subseteq \mathbb{R}^d$ denotes a Lipschitz domain. Its boundary will be denoted by $\Gamma := \partial\Omega$. We also introduce the space of smooth and compactly supported test functions. For an open set $\mathcal{O} \subseteq \mathbb{R}^d$ we denote them by $C_0^\infty(\mathcal{O})$.

We denote the usual Lebesgue spaces by $L^p(\Omega)$ for $p \in [1, \infty]$, where we will mostly work with complex valued functions. The norms are given by

$$\|u\|_{L^p(\Omega)} := \left(\int_{\Omega} |u|^p dx \right)^{1/p} \quad \text{for } 1 \leq p < \infty, \quad \text{and} \quad \|u\|_{L^\infty(\Omega)} := \text{esssup}_{\Omega}(|u|).$$

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, we denote the (classical) derivative by

$$D^\alpha u := \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}.$$

Derivatives will be understood in the weak sense, i.e. $D^\alpha u$ is defined as the locally integrable function satisfying

$$\int_{\Omega} D^\alpha u \varphi dx := (-1)^{|\alpha|} \int_{\Omega} u D^\alpha \varphi dx \quad \forall \varphi \in C_0^\infty(\Omega). \quad (2.16)$$

This definition of derivative leads to the scale Sobolev spaces for $p \in [1, \infty]$ and $m \in \mathbb{N}_0$ and their corresponding norms, defined via:

$$\begin{aligned} \|u\|_{W^{m,p}(\Omega)} &:= \sum_{\substack{\alpha \in \mathbb{N}_0 \\ |\alpha| \leq m}} \|D^\alpha u\|_{L^p(\Omega)}, \\ W^{m,p}(\Omega) &:= \left\{ u \in L^p(\Omega) : \|u\|_{W^{m,p}(\Omega)} < \infty \right\}. \end{aligned}$$

For $s \in \mathbb{R}_+$ with $s = m + r$ for $m \in \mathbb{N}_0$ and $r \in (0, 1)$ we define the fractional Sobolev spaces via interpolation:

$$W^{s,p}(\Omega) := [W^{m,p}(\Omega), W^{m+1,p}(\Omega)]_{r,p}.$$

We also need another family of Sobolev spaces, which encode homogeneous boundary conditions. For $s \geq 0$, we define:

$$\widetilde{W}^{s,p}(\Omega) := \overline{C_0^\infty(\Omega)}, \quad \text{closure with respect to the } W^{s,p}(\mathbb{R}^d)\text{-norm.}$$

For $s \in \mathbb{N}$, it is more common to use the notation $W_0^{s,p}(\Omega) := \widetilde{W}^{s,p}(\Omega)$. For negative s , we define the Sobolev spaces via duality:

$$W^{-s,p}(\Omega) := \left(\widetilde{W}^{s,p}(\Omega) \right)' \quad \text{and} \quad \widetilde{W}^{-s,p}(\Omega) := \left(W^{s,p}(\Omega) \right)'.$$

The most important case of Sobolev spaces is $p = 2$. Therefore we introduce the additional notation for $s \in \mathbb{R}$:

$$H^s(\Omega) := W^{s,2}(\Omega) \quad \text{and} \quad \widetilde{H}^s(\Omega) := \widetilde{W}^{s,2}(\Omega).$$

The spaces $H^s(\Omega)$ are Hilbert spaces. When working with the spaces $\widetilde{H}^s(\Omega)$, it is often convenient to characterize them as interpolation spaces:

Proposition 2.21 ([McL00, Theorems B.8 and B.9]). *Let $s_1, s_2 \in \mathbb{R}$, $\theta \in [0, 1]$. Then, the following equivalence holds:*

$$\begin{aligned} [H^{s_1}(\Omega), H^{s_2}(\Omega)]_{\theta,2} &= H^s(\Omega), \\ [\widetilde{H}^{s_1}(\Omega), \widetilde{H}^{s_2}(\Omega)]_{\theta,2} &= \widetilde{H}^s(\Omega) \end{aligned} \quad \text{with } s := (1 - \theta)s_1 + \theta s_2.$$

Proposition 2.22 ([McL00, Theorem 3.40(i)]). *For $0 \leq s < 1/2$, the two families of Sobolev spaces $H^s(\Omega)$ and $\widetilde{H}^s(\Omega)$ coincide, i.e.,*

$$H^s(\Omega) = \widetilde{H}^s(\Omega)$$

with equivalent norms. The implied constants depend only on Ω .

Proof. Theorem 3.40 (i) in [McL00] considers the slightly different family of spaces $H_0^s(\Omega)$ instead of $\widetilde{H}^s(\Omega)$. For $s \neq \frac{1}{2}, \frac{3}{2}, \dots$ they are shown to coincide with $\widetilde{H}^s(\Omega)$ in [McL00, Theorem 3.33]. \square

When working with Sobolev functions, it is useful to be able to extend them from the domain Ω to the full space. The fact that this can be done in a stable way is the content of the next proposition.

Proposition 2.23 (Stein extension operator, see [Ste70, Chap. VI.3]). *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain. Then, there exists a linear operator \mathcal{E} with the following properties:*

(i) *for $m \in \mathbb{N}_0$, $\mathcal{E} : H^m(\Omega) \rightarrow H^m(\mathbb{R}^d)$ is a bounded linear operator, satisfying*

$$\|\mathcal{E}u\|_{H^m(\mathbb{R}^d)} \leq C(m, \Omega) \|u\|_{H^m(\Omega)},$$

(ii) *$\mathcal{E}u$ is an extension of u , i.e. $\mathcal{E}u|_\Omega = u$.*

We will also need the space of functions whose divergence is in L^2 . We define the weak divergence analogously to (2.16) by $\int_{\Omega} \operatorname{div} \mathbf{v} \varphi \, dx = - \int_{\Omega} \mathbf{v} \cdot \nabla \varphi \, dx$ for all $\varphi \in C_0^\infty(\Omega)$. The corresponding function space is:

$$\|\mathbf{v}\|_{H(\operatorname{div}, \Omega)}^2 := \|\mathbf{v}\|_{[L^2(\Omega)]^d}^2 + \|\operatorname{div} \mathbf{v}\|_{L^2(\Omega)}^2, \quad (2.17)$$

$$H(\operatorname{div}, \Omega) := \left\{ \mathbf{v} \in [L^2(\Omega)]^d, \|\mathbf{v}\|_{H(\operatorname{div}, \Omega)} < \infty \right\}. \quad (2.18)$$

Similarly, we define the space of functions with (distributional) Laplacian in L^2 by:

$$\begin{aligned} \|u\|_{H_{\Delta}^1(\Omega)}^2 &:= \|u\|_{H^1(\Omega)}^2 + \|\Delta u\|_{L^2(\Omega)}^2, \\ H_{\Delta}^1(\Omega) &:= \left\{ u \in L^2(\Omega) : \|u\|_{H_{\Delta}^1(\Omega)} < \infty \right\}. \end{aligned}$$

It is easy to see that $H_{\Delta}^1(\Omega) = \{u \in H^1(\Omega) : \nabla u \in H(\operatorname{div}, \Omega)\}$.

Under certain conditions, we can trade in regularity for integrability. This is contained in the following proposition.

Proposition 2.24 (Sobolev embeddings, [AF03, Theorem 4.12]). *Let Ω be a bounded Lipschitz domain in \mathbb{R}^d . Let $m \in \mathbb{N}$ and $1 \leq p < \infty$. Then for $mp < d$:*

$$W^{m,p}(\Omega) \hookrightarrow L^q(\Omega) \quad \text{for } p \leq q \leq p^* := dp/(d - mp)$$

and in the case $mp = d$:

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

Sobolev spaces on the boundary and trace operators

In order to be able to treat boundary integral equations, we also need Sobolev spaces of functions supported on the boundary of a Lipschitz domain. Most of the spaces introduced in the previous setting have a natural correspondence on the boundary.

Let Ω denote a *bounded* Lipschitz domain, and $\Gamma := \partial\Omega$ be its boundary. We let $L^p(\Gamma)$ denote the usual Lebesgue space for $1 \leq p \leq \infty$, and define the space $H^1(\Gamma)$ via

$$\begin{aligned} \|u\|_{H^1(\Gamma)}^2 &:= \|u\|_{L^2(\Gamma)}^2 + \|\nabla_{\Gamma} u\|_{L^2(\Gamma)}^2, \\ H^1(\Gamma) &:= \left\{ u \in L^2(\Gamma) : \|u\|_{H^1(\Gamma)} < \infty \right\}, \end{aligned}$$

where ∇_{Γ} denotes the surface gradient. For $s \in (0, 1)$, we then define the fractional Sobolev spaces via interpolation as

$$H^s(\Gamma) := [L^2(\Gamma), H^1(\Gamma)]_{\theta, 2}.$$

Negative order Sobolev spaces are defined by duality $H^{-s}(\Gamma) := (H^s(\Gamma))'$. Since it is often convenient, we define the sesquilinear form $\langle \cdot, \cdot \rangle_{\Gamma}$ on $H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma)$ as the continuous extension of $\langle u, v \rangle_{\Gamma} := \int_{\Gamma} u \bar{v}$ for $u, v \in L^2(\Gamma)$.

Often, we need to distinguish between the interior and the exterior of the domain Ω . In order to have a unified notation we set $\Omega^- := \Omega$ and $\Omega^+ := \mathbb{R}^d \setminus \overline{\Omega}$. Many operators have a version defined on the interior and the exterior. As a notational convention we will indicate which domain an operator belongs to by the superscript \pm .

The following two propositions characterize the fractional Sobolev spaces on Γ as the trace spaces of functions on Ω .

Proposition 2.25. *Let Ω^\pm denote either Ω or the exterior $\mathbb{R}^d \setminus \overline{\Omega}$.*

(i) *For $s \in (1/2, 3/2)$ there exists a bounded linear operator $H^s(\Omega^\pm) \rightarrow H^{s-1/2}(\Gamma)$ such that:*

$$\gamma^\pm u := u|_\Gamma \quad \text{for } u \in C^\infty(\overline{\Omega}) \quad \text{and} \quad \|\gamma^\pm u\|_{H^{s-1/2}(\Gamma)} \leq C(s, \Omega^\pm) \|u\|_{H^s(\Omega^\pm)}.$$

(ii) *There exists a bounded linear operator $\gamma_\nu^\pm : H(\operatorname{div}, \Omega^\pm) \rightarrow H^{-1/2}(\Gamma)$ which satisfies*

$$\gamma_\nu^\pm \mathbf{v} := \mathbf{v} \cdot \nu|_\Gamma \quad \text{for } \mathbf{v} \in [C^\infty(\overline{\Omega})]^d \quad \text{and} \quad \|\gamma_\nu^\pm \mathbf{v}\|_{H^{-1/2}(\Gamma)} \leq C(\Omega^\pm) \|\mathbf{v}\|_{H(\operatorname{div}, \Omega^\pm)}$$

where ν denotes the normal vector pointing out of Ω .

Proof. Part (i) can be found in [Cos88b, Lemma 3.6]. Part (ii) is well known and can be found for example in [Tar07, Lemma 20.2]. \square

Proposition 2.26. *Let Ω^\pm denote either Ω or the exterior $\mathbb{R}^d \setminus \overline{\Omega}$.*

(i) *For $s \in (1/2, 3/2)$ the trace operator $\gamma^\pm : H^s(\Omega^\pm) \rightarrow H^{s-1/2}(\Gamma)$ is surjective and admits a bounded right-inverse \mathcal{E}_Γ^D , i.e.*

$$\gamma^\pm \circ \mathcal{E}_\Gamma^D = \operatorname{Id}, \quad \text{and} \quad \|\mathcal{E}_\Gamma^D u\|_{H^s(\Omega)} \leq C(s, \Omega^\pm) \|u\|_{H^{s-1/2}(\Gamma)}.$$

(ii) *For $s \in (0, 1/2]$ the normal trace operator has a bounded right inverse \mathcal{E}_Γ^N satisfying:*

$$\begin{aligned} \gamma_\nu^\pm \circ \mathcal{E}_\Gamma^N &= \operatorname{Id}, \\ \|\mathcal{E}_\Gamma^N u\|_{[H^s(\Omega^\pm)]^d} &\leq C(s, \Omega^\pm) \|u\|_{H^{-1/2+s}(\Gamma)}. \end{aligned}$$

For $s = 0$, the normal trace can be lifted into $H(\operatorname{div}, \Omega^\pm)$, i.e.

$$\|\mathcal{E}_\Gamma^N u\|_{H(\operatorname{div}, \Omega^\pm)} \leq C(\Omega^\pm) \|u\|_{H^{-1/2}(\Gamma)}.$$

Proof. For the proof of (i) see [Cos88b, Lemma 4.2]. The lifting from (ii) can be constructed by solving an elliptic Neumann problem

$$-\Delta \varphi + \varphi = 0 \quad \text{and} \quad \partial_n \varphi = u$$

and taking the gradient. This directly gives the case $s = 0$. The case $s = 1/2$ was proven in [JK81]. The intermediate cases can be seen by combining the mapping properties of the Neumann-to-Dirichlet map (see [Cos88b, Lemma 3.7] or [McL00, Theorem 4.24(ii)]) with the regularity properties of the Dirichlet mapping in (i). \square

The trace operator for $H(\text{div})$ functions allows us to define the normal derivative:

$$\partial_\nu^\pm : H_\Delta^1(\Omega^\pm) \rightarrow H^{-1/2}(\Gamma), \quad u \mapsto \gamma_\nu^\pm \nabla u. \quad (2.19)$$

When working with boundary integral equations, often times the quantity of interest is not the trace of a function but its jump across the interface. We introduce the jump operators for $s \in (1/2, 3/2)$

$$[\![\gamma \cdot]\!] : H^s(\mathbb{R}^d \setminus \Gamma) \rightarrow H^{s-1/2}(\Gamma), \quad u \mapsto \gamma^+ u - \gamma^- u, \quad (2.20a)$$

$$[\![\partial_\nu \cdot]\!] : H_\Delta^1(\mathbb{R}^d \setminus \Gamma) \rightarrow H^{-1/2}(\Gamma), \quad u \mapsto \partial_\nu^+ u - \partial_\nu^- u, \quad (2.20b)$$

$$[\![\gamma_\nu \cdot]\!] : H(\text{div}, \mathbb{R}^d \setminus \Gamma) \rightarrow H^{-1/2}(\Gamma), \quad u \mapsto \gamma_\nu^+ u - \gamma_\nu^- u, \quad (2.20c)$$

as well as the mean values:

$$\{\!\{ \gamma \cdot \}\!\} : H^s(\mathbb{R}^d \setminus \Gamma) \rightarrow H^{s-1/2}(\Gamma), \quad u \mapsto \frac{1}{2}(\gamma^+ u + \gamma^- u), \quad (2.21a)$$

$$\{\!\{ \partial_\nu \cdot \}\!\} : H_\Delta^1(\mathbb{R}^d \setminus \Gamma) \rightarrow H^{-1/2}(\Gamma), \quad u \mapsto \frac{1}{2}(\partial_\nu^+ u + \partial_\nu^- u), \quad (2.21b)$$

$$\{\!\{ \gamma_\nu \cdot \}\!\} : H(\text{div}, \mathbb{R}^d \setminus \Gamma) \rightarrow H^{-1/2}(\Gamma), \quad u \mapsto \frac{1}{2}(\gamma_\nu^+ u + \gamma_\nu^- u). \quad (2.21c)$$

Remark 2.27. *Oftentimes in the literature, the signs in the definitions of the jumps are reversed. In our applications we are often dealing with exterior problems, i.e., with functions vanishing in Ω^- and we therefore have $\gamma^+ u = [\![\gamma u]\!]$. When dealing with spatial semi-discretization this relation simplifies the signs in many of the formulas.*

For Sobolev indices higher than 1, the spaces $H^s(\Gamma)$ are only well defined for smooth geometries, as the definition would depend on the choice of coordinates. In order to prove *a priori* estimates, we instead consider the spaces of piecewise H^s -functions. For details on these spaces, we refer to [SS11, Definition 4.1.48].

Definition 2.28. *Let Γ be piecewise smooth, i.e., assume there exists an open partitioning $\mathcal{C} := \{\Gamma_i : 1 \leq i \leq q\}$ such that $\Gamma = \bigcup_{i=1}^q \overline{\Gamma}_i$. For $s > 1$, we define the space*

$$H_{\text{pw}}^s(\Gamma) := \{u \in H^1(\Gamma) : u|_{\Gamma_i} \in H^s(\Gamma_i) \ \forall \Gamma_i \in \mathcal{C}\} \quad (2.22)$$

equipped with the norm $\|u\|_{H_{\text{pw}}^s(\Gamma)}^2 := \sum_{\Gamma_i \in \mathcal{C}} \|u\|_{H^s(\Gamma_i)}^2$. For $s \leq 1$, we define $H_{\text{pw}}^s(\Gamma) := H^s(\Gamma)$ with the usual norm.

The Sobolev embeddings also transfer to the spaces on the boundary. We will need the following special case:

Proposition 2.29 (Sobolev embeddings, [Tar07, Lemma 32.1]). *Let Γ be the boundary of a bounded Lipschitz domain $\Omega \subseteq \mathbb{R}^d$. Then the following embeddings hold for $s < (d-1)/2$:*

$$H^s(\Gamma) \hookrightarrow L^{p(s)}(\Gamma) \quad \text{with} \quad \frac{1}{p(s)} = \frac{1}{2} - \frac{s}{d-1}. \quad (2.23)$$

For the space $H^{1/2}(\Gamma)$ this implies:

$$H^{1/2}(\Gamma) \hookrightarrow L^p(\Gamma) \quad \text{for} \quad \begin{cases} 1 \leq p < \infty & d = 2, \\ 1 \leq p \leq \frac{2d-2}{d-2} & d \geq 3. \end{cases} \quad (2.24)$$

Proposition 2.30 (Rellich, [McL00, Theorem 3.27]). *For $0 \leq t < s \leq 1$ the embedding*

$$H^s(\Gamma) \hookrightarrow H^t(\Gamma)$$

is compact.

2.4 Triangulations

In this work we consider discretizations of the space variables using Galerkin-type methods. In order to realize these, we need finite dimensional spaces $V_h \subseteq H^s(\Omega)$ with good approximation properties. While most of the results in this work are formulated for general (conforming) discretization spaces, in practice we will use standard finite element and boundary element spaces. The purpose of this section is to introduce these spaces together with the appropriate notation used throughout this thesis, as well as some projection and (quasi)-interpolation operators, which will provide concrete constructions for the abstract assumptions made when discussing the different discretizations for wave propagation problems. Throughout this section, let $\Omega \subseteq \mathbb{R}^d$ be a bounded Lipschitz polyhedron with boundary Γ .

Definition 2.31 (see [BS08, Definition 4.4.13]). *Let \mathcal{T}_h be a decomposition of $\Omega \subseteq \mathbb{R}^d$ into open d -simplices. We call \mathcal{T}_h a regular triangulation if it does not contain any hanging nodes (i.e. the intersection of elements $\bar{K} \cap \bar{K}'$ is either empty, a common vertex, side or face). We say \mathcal{T}_h is shape regular if the ratio $\text{diam}(K)/\rho(K)$, where $\rho(K)$ is the diameter of the largest ball contained in K , satisfies*

$$\max_{K \in \mathcal{T}_h} \left(\frac{\text{diam}(K)}{\rho(K)} \right) \leq \gamma$$

for a constant $\gamma > 0$ independent of the mesh size. We define the mesh size of \mathcal{T}_h as $h := \max_{K \in \mathcal{T}_h} (\text{diam}(K))$. The triangulation is called quasi-uniform, if there exists a constant $C > 0$ such that

$$\max_{K \in \mathcal{T}_h} (\text{diam}(K)) \leq C \min_{K \in \mathcal{T}_h} (\text{diam}(K)),$$

i.e., $h \sim \min_{K \in \mathcal{T}_h} (\text{diam}(K))$.

To each element $K \in \mathcal{T}_h$, we assign the (bijective) element map $F_K : \hat{K} \rightarrow K$ where \hat{K} denotes the reference element given by:

$$\hat{K} := \begin{cases} \{x \in \mathbb{R} : 0 \leq x \leq 1\} & \text{if } d = 1, \\ \{(x, y) \in \mathbb{R}^2 : 0 \leq x, y \leq 1 \wedge 0 \leq x + y \leq 1\} & \text{if } d = 2, \\ \{(x, y, z) \in \mathbb{R}^3 : 0 \leq x, y, z \leq 1 \wedge 0 \leq x + y + z \leq 1\} & \text{if } d = 3. \end{cases}$$

We are now able to define the spaces of piecewise polynomials:

Definition 2.32. Let \mathcal{T}_h be a regular and shape-regular triangulation of Ω . For $p \in \mathbb{N}_0$, denote by

$$\mathcal{P}_p(\widehat{K}) := \begin{cases} \text{span} \{x^i, 0 \leq i \leq p\} & d = 1, \\ \text{span} \{x^i y^j, 0 \leq i + j \leq p\} & d = 2, \\ \text{span} \{x^i y^j z^k, 0 \leq i + j + k \leq p\} & d = 3 \end{cases}$$

the space of polynomials of degree up to p on the reference element. Define the space of piecewise polynomials as

$$\mathcal{S}^{p,0}(\mathcal{T}_h) := \left\{ u \in L^\infty(\Omega) : u|_K \circ F_K \in \mathcal{P}_p(\widehat{K}) \quad \forall K \in \mathcal{T}_h \right\}, \quad (2.25)$$

$$\mathcal{S}^{p,1}(\mathcal{T}_h) := \mathcal{S}^{p,0}(\mathcal{T}_h) \cap C^0(\Omega). \quad (2.26)$$

For $d = 2, 3$, all the definitions above also transfer to the discretizations of the boundary in a natural way.

Definition 2.33 (see [SS11, Section 4.1.2]). Let \mathcal{T}_h^Γ denote a decomposition of Γ into line segments/triangles with element maps $F_K : \widehat{K} \rightarrow K$. The terms regular, shape-regular and uniform are defined analogously to the case of volume meshes (see Definition 2.31). We define the spaces

$$\mathcal{S}^{p,0}(\mathcal{T}_h^\Gamma) := \left\{ u \in L^\infty(\Gamma) : u|_K \circ F_K \in \mathcal{P}_p(\widehat{K}) \quad \forall K \in \mathcal{T}_h^\Gamma \right\}, \quad (2.27)$$

$$\mathcal{S}^{p,1}(\mathcal{T}_h^\Gamma) := \mathcal{S}^{p,0}(\mathcal{T}_h^\Gamma) \cap C^0(\Gamma). \quad (2.28)$$

The approximation properties of these spaces can be found in most literature on finite element methods, e.g. [BS08, Theorem 4.4.20]. We summarize them in the following proposition:

Proposition 2.34. Let \mathcal{T}_h be a quasi-uniform triangulation of a bounded Lipschitz polyhedron Ω . Then, the following estimates hold with constants depending on the shape-regularity constant γ , Ω , and p : For $m \geq 0$ the discontinuous spline spaces satisfy

$$\inf_{v_h \in \mathcal{S}^{p,0}(\mathcal{T}_h)} \|u - v_h\|_{L^2(\Omega)} \leq Ch^{\min(p+1, m)} \|u\|_{H^m(\Omega)} \quad \forall u \in H^m(\Omega).$$

In the case of continuous approximation spaces, we get for $u \in H^m(\Omega)$ with $m \geq 1$:

$$\inf_{v_h \in \mathcal{S}^{p,1}(\mathcal{T}_h)} \left(\|u - v_h\|_{L^2(\Omega)} + h \|\nabla u - \nabla v_h\|_{L^2(\Omega)} \right) \leq Ch^{\min(p+1, m)} \|u\|_{H^m(\Omega)}.$$

For a family of quasi-uniform triangulations $(\mathcal{T}_h)_{h \geq 0}$ such that the mesh size h goes to 0,

$$\bigcup_{h \geq 0} \mathcal{S}^{p,1}(\mathcal{T}_h) \subseteq H^s(\Omega) \text{ is dense for } s \leq 1,$$

$$\bigcup_{h \geq 0} \mathcal{S}^{p,0}(\mathcal{T}_h) \subseteq H^s(\Omega) \text{ is dense for } s \leq 0.$$

For approximations on the boundary Γ , a similar result can be shown, but one has to take into account that Sobolev spaces are not well defined for $s > 1$. Instead we have the following result:

Proposition 2.35 ([SS11, Theorem 4.3.20 and Theorem 4.1.51]). *Let Γ be the piecewise smooth boundary of a Lipschitz polyhedron and let \mathcal{T}_h^Γ be a regular and quasi-uniform triangulation of Γ : Then the following estimates hold:*

$$\begin{aligned} \inf_{v_h \in \mathcal{S}^{p,0}(\mathcal{T}_h^\Gamma)} \|u - v_h\|_{H^{-1/2}(\Gamma)} &\leq Ch^{\min(p+1,s)+1/2} \|u\|_{H_{\text{pw}}^s(\Gamma)} \quad \forall u \in H_{\text{pw}}^s(\Gamma), \quad s \geq -1/2, \\ \inf_{v_h \in \mathcal{S}^{p,1}(\mathcal{T}_h^\Gamma)} \|u - v_h\|_{H^{1/2}(\Gamma)} &\leq Ch^{\min(p+1,s)-1/2} \|u\|_{H_{\text{pw}}^s(\Gamma)} \quad \forall u \in H_{\text{pw}}^s(\Gamma), \quad s \geq 1/2. \end{aligned}$$

Sometimes, instead of relying on the abstract existence of approximating functions, we need to construct them using linear and stable operators. These can be constructed in multiple ways, see e.g. [KM15] for a very general construction which is robust in p . Other ways to construct such an operator include the construction by Scott and Zhang [SZ90] (we use the version as modified in [Aur+15, Lemma 3]) or (for low order) operators of Clément type (see [Clé75]), see also [GE16]. In the very simple case of quasi-uniform triangulations, the L^2 -projection also satisfies the necessary assumptions (see also [CT87; BY14] for other sufficient conditions).

Proposition 2.36. *Let Γ be the piecewise smooth boundary of a Lipschitz polyhedron and \mathcal{T}_h^Γ is a regular and quasi-uniform triangulation of Γ . Then, there exists an operator $J_h^\Gamma : L^2(\Gamma) \rightarrow \mathcal{S}^{p,1}(\mathcal{T}_h^\Gamma)$ with the following properties:*

- (i) $J_h^\Gamma : H^s(\Gamma) \rightarrow H^s(\Gamma)$ is linear and bounded for $s \in [0, 1]$,
- (ii) the following approximation property holds for $0 \leq t \leq 1$ and $s \geq t$:

$$\|u - J_h^\Gamma u\|_{H^t(\Gamma)} \leq Ch^{\min(p+1,s)-t} \|u\|_{H_{\text{pw}}^s(\Gamma)} \quad \forall u \in H_{\text{pw}}^s(\Gamma).$$

Analogously if \mathcal{T}_h is a regular and quasi-uniform triangulation of Ω , there exists an operator $J_h : L^2(\Omega) \rightarrow \mathcal{S}^{p,1}(\mathcal{T}_h)$ such that

- (i) $J_h : H^s(\Omega) \rightarrow H^s(\Omega)$ is linear and bounded for $s \in [0, 1]$,
- (ii) the following approximation property holds for $s \geq 1$:

$$\|u - J_h u\|_{L^2(\Omega)} + h \|\nabla(u - J_h u)\|_{L^2(\Omega)} \leq Ch^{\min(p+1,s)} \|u\|_{H^s(\Omega)} \quad \forall u \in H^s(\Omega).$$

2.5 Boundary integral equations for the Helmholtz equation

In this section, we develop the theory of boundary integral equations for the Helmholtz equation, which will form the basis for all the discretization schemes used in this thesis. The results in this section are all standard in the literature and can, for example, be

found in the monographs [McL00; SS11; Ste08; HW08]. For a more time-domain oriented introduction we refer to [Say16]. For $s \in \mathbb{C}$ with $\operatorname{Re}(s) > 0$, we consider solutions u of

$$-\Delta u + s^2 u = 0 \quad \text{in } \mathbb{R}^d \setminus \Gamma, \quad (2.29)$$

where Γ denotes the boundary of a bounded Lipschitz domain Ω .

We start with the simple fact that transmission problems are uniquely solvable:

Proposition 2.37. *For $s \in \mathbb{C}$ with $\operatorname{Re}(s) > 0$ and $\varphi \in H^{-1/2}(\Gamma)$, $\psi \in H^{1/2}(\Gamma)$, the problem of finding $u \in H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)$ such that*

$$-\Delta u + s^2 u = 0 \quad (2.30a)$$

$$[[\gamma u]] = \psi \quad \text{and} \quad [[\partial_{\nu} u]] = \varphi \quad (2.30b)$$

has a unique solution.

Proof. We only show uniqueness. Existence can, for example, be shown using the potentials in Proposition 2.38 or the Lax-Milgram theory together with appropriate boundary liftings.

By linearity it is sufficient to consider the homogeneous problem $\varphi = \psi = 0$ and show $u = 0$. We note that in $H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)$ the integration by parts formula holds. This can be seen by using the density of compactly supported test-functions in $H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)$. Multiplying (2.30a) by $\overline{s}u$, integrating and integration by parts gives:

$$\overline{s} \|\nabla u\|_{L^2(\mathbb{R}^d \setminus \Gamma)}^2 + |s|^2 s \|u\|_{L^2(\mathbb{R}^d \setminus \Gamma)}^2 = 0.$$

Since we assumed $\operatorname{Re}(s) > 0$ we can deduce $\|u\|_{H^1(\mathbb{R}^d \setminus \Gamma)} = 0$ by taking the real part. \square

We introduce the fundamental solution for the differential operator $-\Delta + s^2$:

$$\Phi(x; s) := \begin{cases} \frac{i}{4} H_0^{(1)}(is|x|) & \text{for } d = 2, \\ \frac{e^{-s|x|}}{4\pi|x|}, & \text{for } d = 3. \end{cases} \quad (2.31)$$

Here $H_0^{(1)}$ denotes the Hankel function of the first kind and order zero. For details, see [McL00, Chapter 9].

Proposition 2.38. *Define the single- and double-layer potentials for $x \in \mathbb{R}^d \setminus \Gamma$:*

$$(S(s)\varphi)(x) := \int_{\Gamma} \Phi(x-y; s)\varphi(y) dS(y), \quad (2.32a)$$

$$(D(s)\psi)(x) := \int_{\Gamma} \partial_{\nu(y)} \Phi(x-y; s)\psi(y) dS(y). \quad (2.32b)$$

If $u \in H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)$ solves (2.29) for $s \in \mathbb{C}_+$, it can be written using the representation formula

$$u(x) = -\left[S(s)[[\partial_{\nu} u]]\right](x) + \left[D(s)[[\gamma u]]\right](x) \quad \forall x \in \mathbb{R}^d \setminus \Gamma. \quad (2.33)$$

Conversely, if we define u via $u := -S(s)\varphi + D(s)\psi$ for some boundary data $\varphi \in H^{-1/2}(\Gamma)$ and $\psi \in H^{1/2}(\Gamma)$, then u solves (2.29) with $[[\partial_{\nu} u]] = \varphi$ and $[[\gamma u]] = \psi$.

Finally, we introduce the boundary integral operators corresponding to the potentials:

$$V(s) : H^{-1/2}(\Gamma) \rightarrow H^{1/2}(\Gamma), \quad V(s) := \gamma^\pm S(s), \quad \text{"single layer operator"}, \quad (2.34a)$$

$$K(s) : H^{1/2}(\Gamma) \rightarrow H^{1/2}(\Gamma), \quad K(s) := \{\{\gamma D(s)\}\}, \quad \text{"double layer operator"}, \quad (2.34b)$$

$$K^T(s) : H^{-1/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma), \quad K^T(s) := \{\{\partial_\nu S(s)\}\}, \quad \text{"adjoint double layer"}, \quad (2.34c)$$

$$W(s) : H^{1/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma), \quad W(s) := -\partial_\nu^\pm D(s), \quad \text{"hypersingular operator"}. \quad (2.34d)$$

In practice, these operators can be computed via explicit representation as integrals over the boundary Γ . For sufficiently smooth functions ψ , φ the following equations hold:

$$V(s)\varphi = \int_\Gamma \Phi(\cdot, y, s)\varphi(y) d\Gamma(y), \quad (2.35a)$$

$$K^T(s)\varphi = \int_\Gamma \partial_{\nu(\cdot)}\Phi(\cdot, y, s)\varphi(y) d\Gamma(y), \quad (2.35b)$$

$$K(s)\psi = \int_\Gamma \partial_{\nu(y)}\Phi(\cdot, y, s)\psi(y) d\Gamma(y), \quad (2.35c)$$

$$W(s)\psi = -\partial_\nu \int_\Gamma \partial_{\nu(y)}\Phi(\cdot, y, s)\psi(y) d\Gamma(y). \quad (2.35d)$$

Remark 2.39. *Since the fundamental solution depends analytically on the wave number for $s \in \mathbb{C}_+$, the dependence of the boundary integral operators on s is also analytic. This will be important, as it allows us to apply the Convolution Quadrature techniques, which will be introduced in Section 3.3, to these operators.*

2.5.1 The Helmholtz equation with matrix-valued wave number

In this section, we generalize the results from the scalar problem (2.29) to the system of equations $-\Delta U + B^2 U$ where $B \in \mathbb{C}^{m \times m}$ is a matrix and U is vector valued. This setting is common when considering Runge-Kutta type discretizations of wave propagation problems, see Section 3.2. Most of the results of the previous section have an analogous counterpart. We use the usual trace and jump operators also for vector valued functions, without notational distinction. They are meant to be applied component-wise.

In order to define potentials and boundary integral operators, we need to be able to apply the functions to a matrix. This is done using the following functional calculus:

Definition 2.40 (c.f. [Yos80, Chapter VIII.7], [GV13, Chapter 11]). *Let $F : G \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be a holomorphic function, which is defined on a domain $G \subseteq \mathbb{C}$, and \mathcal{X}, \mathcal{Y} be Banach spaces. Let B be a matrix with $\sigma(B) \subseteq G$. We then define $F(B)$ via the Riesz-Dunford functional calculus for holomorphic functions:*

$$F(B) := \frac{1}{2\pi i} \int_{\mathcal{C}} (B - \lambda)^{-1} \otimes F(\lambda) d\lambda,$$

where $\mathcal{C} \subset G$ is a closed path with winding number 1 encircling $\sigma(B)$. The operator \otimes

denotes the Kronecker product, i.e., for a matrix $A \in \mathbb{C}^{m \times m}$ we write

$$A \otimes F := \begin{pmatrix} a_{11}F & \dots & a_{1m}F \\ \vdots & \dots & \vdots \\ a_{m1}F & \dots & a_{mm}F \end{pmatrix}.$$

This defines the operator $F(B)$, mapping from the product space $[\mathcal{X}]^m$ to the product space $[\mathcal{Y}]^m$. For a fixed matrix $B \in \mathbb{C}^{m \times m}$, the mapping $F \mapsto F(B)$ is linear and an algebra homomorphism, i.e. $(FG)(B) = F(B)G(B)$.

Since we often need to apply the same operator A to vectors of functions, we use the notation $\underline{A} := [\text{Id} \otimes A] = \text{diag}(A, \dots, A)$ for the diagonal product operator, where we assume that the number of entries is clear from context.

We start with generalizing Proposition 2.37 in a straight-forward way:

Proposition 2.41. For $B \in \mathbb{C}^{m \times m}$ with $\sigma(B) \subseteq \mathbb{C}_+$, $\Phi \in [H^{-1/2}(\Gamma)]^m$, and $\Psi \in [H^{1/2}(\Gamma)]^m$, the transmission problem: find $U \in [H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)]^m$ such that

$$\begin{aligned} -\underline{\Delta}U + B^2U &= 0 & \text{in } \mathbb{R}^d \setminus \Gamma \\ \llbracket \gamma U \rrbracket &= \Psi, & \text{and } \llbracket \partial_{\nu} U \rrbracket &= \Phi \end{aligned}$$

has a unique solution.

Proof. We focus on uniqueness, existence follows from using potentials in Lemma 2.42. For a solution of the homogeneous problem, i.e. $\Phi = \Psi = 0$, we can write $B = XJX^{-1}$ in Jordan form. Most notably J is an upper triangular matrix. This implies $-\underline{\Delta}X^{-1}U + J^2X^{-1}U = 0$ with homogeneous jump conditions. Since the scalar problem has a unique solution by Proposition 2.37, we immediately see that the last component of $X^{-1}U$ must be 0. A backward substitution argument then gives, in each step using Proposition 2.37, that $X^{-1}U = 0$, which in turn means $U = 0$. \square

We now are able to generalize the representation formula (2.33):

Lemma 2.42. Let $U \in [H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)]^m$ solve the equation

$$-\underline{\Delta}U + B^2U = 0 \quad \text{in } \mathbb{R}^d \setminus \Gamma \tag{2.36}$$

for a matrix $B \in \mathbb{C}^{m \times m}$ with $\sigma(B) \subseteq \mathbb{C}_+$. Then U can be written as

$$U = -S(B) \llbracket \partial_{\nu} U \rrbracket + D(B) \llbracket \gamma U \rrbracket. \tag{2.37}$$

Proof. We show that the function $V := -S(B) \llbracket \partial_{\nu} U \rrbracket + D(B) \llbracket \gamma U \rrbracket$, defined analogously to (2.37), solves the transmission problem:

$$\begin{aligned} -\underline{\Delta}V + B^2V &= 0 & \text{in } \mathbb{R}^d \setminus \Gamma \\ \llbracket \gamma V \rrbracket &= \llbracket \gamma U \rrbracket & \text{and } \llbracket \partial_{\nu} V \rrbracket &= \llbracket \partial_{\nu} U \rrbracket. \end{aligned}$$

Since solutions to such problems are unique via Proposition 2.41 this is sufficient. The jump conditions follow directly from the definitions using the Riesz-Dunford calculus and

the jump properties of the scalar single- and double layer potentials. To see the differential equation, we first look at $S(B)$ and calculate:

$$\begin{aligned}
 -\underline{\Delta}S(B) \llbracket \partial_\nu U \rrbracket &= -\underline{\Delta} \left(\int_{\mathcal{C}} (B - \lambda)^{-1} \otimes S(\lambda) \llbracket \partial_\nu U \rrbracket d\lambda \right) \\
 &= - \int_{\mathcal{C}} (B - \lambda)^{-1} \otimes [-\Delta S(\lambda)] \llbracket \partial_\nu U \rrbracket d\lambda \\
 &= - \int_{\mathcal{C}} (B - \lambda)^{-1} \otimes [\lambda^2 S(\lambda)] \llbracket \partial_\nu U \rrbracket d\lambda \\
 &= -B^2 S(B) \llbracket \partial_\nu U \rrbracket,
 \end{aligned}$$

where in the last step we used that the functional calculus is an algebra homomorphism. The same calculation can be done for $D(B)$. \square

The representation formula can be used to derive boundary integral equations for the jumps or traces of solutions to (2.29). While there are many ways to write these equations, we focus on a version which allows us to easily derive an expression for the Dirichlet-to-Neumann map.

Proposition 2.43. *Let $U \in [H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)]^m$ solve (2.36). Then, the following identities hold on the boundary:*

$$\begin{pmatrix} \gamma^- U \\ \partial_\nu^+ U \end{pmatrix} = - \begin{pmatrix} \frac{1}{2} - K(B) & V(B) \\ W(B) & -\frac{1}{2} + K^T(B) \end{pmatrix} \begin{pmatrix} \llbracket \gamma U \rrbracket \\ \llbracket \partial_\nu U \rrbracket \end{pmatrix}. \quad (2.38)$$

If we consider solutions to the exterior problem, i.e. $-\underline{\Delta}U + B^2U = 0$ in $\mathbb{R}^d \setminus \bar{\Omega}$, the traces solve

$$\begin{pmatrix} 0 \\ -\partial_\nu^+ U \end{pmatrix} = \begin{pmatrix} \frac{1}{2} - K(B) & V(B) \\ W(B) & -\frac{1}{2} + K^T(B) \end{pmatrix} \begin{pmatrix} \gamma^+ U \\ \partial_\nu^+ U \end{pmatrix}. \quad (2.39)$$

Proof. Equation (2.38) follows from the representation formula and the definitions of the boundary integral operators. In order to derive (2.39), we extend U by 0 on Ω and apply (2.38). \square

3 Time-stepping and Convolution Quadrature

In this chapter, we introduce some of the most common ways of discretizing time evolution problems. We consider time-stepping based methods, which generate a sequence of approximations at discrete times t_n . Since we are interested in applying Convolution Quadrature based methods later, we restrict our considerations to the case where all t_n are multiples of a common factor, i.e., $t_n := k n$ for some parameter $k > 0$ which we call the timestep size. Most of the definitions and elementary results on these methods presented in this section are taken from the books [HNW93; HW10].

We will introduce the methods by first considering a simple ODE, namely, finding $u : [0, \infty) \rightarrow \mathbb{C}$ such that

$$\dot{u} = f(t, u) \quad \text{and} \quad u(0) := u_0, \quad (3.1)$$

for some initial condition $u_0 \in \mathbb{C}$ and given right-hand side $f : \mathbb{R}_+ \times \mathbb{C} \rightarrow \mathbb{C}$.

While all the algorithms generate a sequence of approximations at discrete times t_n , we usually assume that we are given initial conditions for $t \leq 0$ such that we may define approximations for all t by shifting the initial time t_0 such that t becomes a grid-point. This allows us to use function notation, e.g., $u^k(t)$ for our approximating sequence.

3.1 Multistep methods

A multistep method is determined by its number of steps m and by coefficients α_j and β_j , $j = 0, \dots, m$. The simplest cases are the explicit and implicit Euler methods, which approximate u by $u^k(t_n) := u^k(t_{n-1}) + k f(t_{n-1}, u^k(t_{n-1}))$ and $u^k(t_n) = u^k(t_{n-1}) + k f(t_n, u^k(t_n))$. Higher order methods are achieved by reusing more of the previously computed approximations. The general approximation scheme is given by the defining equation for $n \geq m$:

$$\frac{1}{k} \sum_{j=0}^m \alpha_j u^k(t_{n-j}) = \sum_{j=0}^m \beta_j f(t_{n-j}, u^k(t_{n-j})). \quad (3.2)$$

To close the system we assume that we are given $u^k(t_n) := u(t_n)$ for $n = 0, \dots, m-1$ (in practice when using multistep methods for ODEs, these are computed approximately using some other time-stepping method). The coefficients for the most common multistep methods are listed in Table 3.1. These and many other examples can be found in [HNW93]. The function $\delta(z)$ listed in the table is important in the context of Convolution Quadrature and will be introduced in (3.15).

We say a multistep method is of order p , if for arbitrary initial value problems (3.1) with sufficiently smooth exact solutions u , the discretization error can be bounded by

Method	m	$\alpha_0, \dots, \alpha_m$	β_0, \dots, β_m	$\delta(z)$
Implicit Euler/BDF1	1	1, -1	1, 0	$1 - z$
Trapezoidal rule	1	1, -1	$\frac{1}{2}, \frac{1}{2}$	$\frac{2(1-z)}{1+z}$
BDF2	2	$\frac{3}{2}, -2, \frac{1}{2}$	1, 0, 0	$\frac{1}{2}(1-z)(3-z)$

Table 3.1: Examples for commonly used A-stable multistep methods

$|u(t_n) - u^k(t_n)| = \mathcal{O}(k^p)$, where the implied constant is allowed to depend on u and f , but not on k .

An important class of multistep methods are the ones which preserve some qualitative properties of the exact solution in addition to the approximation properties. We only consider the following class of methods:

Definition 3.1. *A multistep method is called A-stable, if for linear ODEs of the form $f(t, y) := \lambda y$ with $\text{Re}(\lambda) \leq 0$, the approximations $u^k(t_n)$ computed via (3.2) are uniformly bounded with respect to n .*

The methods listed in Table 3.1 are all A-stable (see [HW10, Sect. V.1]). A strong limitation when using multistep methods for discretizing PDEs, especially using Convolution Quadrature, is the following result on the achievable order:

Proposition 3.2 (Dahlquist’s second barrier, [HW10, Sect. V.1, Theorem 1.6]). *An A-stable multistep method must be of order $p \leq 2$.*

The BDF1 and BDF2 methods have another important property which will prove useful when investigating nonlinear problems in Chapter 5.

Proposition 3.3. *The linear multistep methods BDF1 and BDF2 are G-stable. This means, there exists a matrix $G = (g_{ij})_{i,j=1,\dots,m}$ that is symmetric and positive definite such that*

$$\text{Re} \left\langle \sum_{j=0}^m \alpha_j u^{n-j}, u^n \right\rangle \geq \|U^n\|_G^2 - \|U^{n-1}\|_G^2,$$

where $U^n = (u^n, \dots, u^{n-m+1})^T$ and

$$\|U^n\|_G^2 = \sum_{i=1}^m \sum_{j=1}^m g_{ij} \langle u^{n-m+i}, u^{n-m+j} \rangle.$$

Proof. As BDF methods are equivalent to their corresponding one-leg methods, the result follows from [HW10, Chapter V.6, Theorem 6.7] and its proof. \square

3.2 Runge-Kutta methods

Proposition 3.2 limits the order achievable by A-stable multistep methods. Since we are interested in using higher order methods, we introduce a second kind of time-stepping

method, which will not suffer from such limitations. An m -stage Runge-Kutta method is characterized by a matrix $\mathcal{Q} \in \mathbb{R}^{m \times m}$ and two vectors $\mathbf{b}, \mathbf{c} \in \mathbb{R}^m$. The approximation to the problem (3.1) is then given by (in each step) first computing the stage vector $U^k(t_n)$ and then using this to compute the approximation $u^k(t_{n+1})$ via the following defining equations:

$$U^k(t_n) = u^k(t_n)\mathbf{1} + k\mathcal{Q}f\left(t_n + k\mathbf{c}, U^k(t_n)\right), \quad (3.3)$$

$$u^k(t_{n+1}) = u^k(t_n) + k\mathbf{b}^T f\left(t_n + k\mathbf{c}, U^k(t_n)\right), \quad (3.4)$$

together with the initial condition $u^k(0) := u_0$. Here $f(t_n + k\mathbf{c}, U^k(t_n))$ denotes the vector

$$f(t_n + k\mathbf{c}, U^k(t_n)) := \begin{pmatrix} f(t_n + k\mathbf{c}_1, U_1^k(t_n)) \\ \vdots \\ f(t_n + k\mathbf{c}_m, U_m^k(t_n)) \end{pmatrix},$$

and $\mathbf{1} := (1, \dots, 1)^T$ is the constant-one vector. Throughout this dissertation we require that \mathcal{Q} is invertible.

Completely analogous to the case of multistep methods, one can define the class of A-stable methods by: for ODEs of the form $f(t, y) := \lambda y$ with $\text{Re}(\lambda) \leq 0$, the approximations $u^k(t_n)$ computed via (3.2) are uniformly bounded with respect to n . In practice, it is more useful to use the following characterization:

Definition 3.4. A Runge-Kutta method is called A-stable, if for $z \in \mathbb{C}$ with $\text{Re}(z) \leq 0$, $(I - z\mathcal{Q})$ is invertible and the stability function

$$r(z) := 1 + z\mathbf{b}^T (I - z\mathcal{Q})^{-1} \mathbf{1} \quad (3.5)$$

satisfies $|r(z)| \leq 1$.

Corollary 3.5. For $z \in \mathbb{C}_-$, i.e., $\text{Re}(z) < 0$, the strict inequality holds:

$$|r(z)| < 1, \quad \forall \text{Re}(z) < 0.$$

Proof. The statement is well known, and follows from the maximum principle for holomorphic functions. \square

Another important class of methods, which is often convenient to work with has the property that the approximation $u^k(t_{n+1})$ coincides with the last entry of the stage vector $U^k(t_n)$. Algebraically they are characterized as follows:

Definition 3.6. We call a Runge-Kutta method stiffly accurate, if it satisfies $\mathbf{b}^T \mathcal{Q}^{-1} = (0, \dots, 1)$.

Sometimes we use the following reformulation of a Runge-Kutta step. Using the definition $r(\infty) := 1 - \mathbf{b}^T \mathcal{Q}^{-1} \mathbf{1}$, we can rewrite (3.3) as:

$$U^k(t_n) = u^k(t_n)\mathbf{1} + k\mathcal{Q}f\left(t_n + k\mathbf{c}, U^k(t_n)\right), \quad (3.6)$$

$$u^k(t_{n+1}) = r(\infty)u^k(t_n) + \mathbf{b}^T \mathcal{Q}^{-1} U^k(t_n). \quad (3.7)$$

When analyzing Runge-Kutta methods, the following order conditions play an important role:

Definition 3.7. We say the Runge-Kutta method defined by $\mathcal{Q}, \mathbf{b}, \mathbf{c}$ has stage order q and classical order p if the approximations satisfy:

$$\left| u(k\mathbf{c}) - U^k(k) \right| = \mathcal{O}(k^{q+1}) \quad \text{and} \quad \left| u(k) - u^k(k) \right| = \mathcal{O}(k^{p+1})$$

for arbitrary right-hand sides f in (3.1), as long as the exact solution is sufficiently smooth.

Using the notation $\mathbf{c}^\ell := (\mathbf{c}_1^\ell, \dots, \mathbf{c}_m^\ell)$ for the entry-wise powers, we use the following order conditions satisfied by such Runge-Kutta methods:

$$\mathbf{c}^\ell = \ell \mathcal{Q} \mathbf{c}^{\ell-1}, \quad 1 \leq \ell \leq q, \quad (3.8)$$

$$\mathbf{b}^T \mathcal{Q}^j \mathbf{c}^\ell = \frac{1}{(j+\ell+1)(j+\ell) \cdots (\ell+1)}, \quad 0 \leq j+\ell \leq p-1. \quad (3.9)$$

Remark 3.8. The fact that the definition of order implies the order conditions (3.8) and (3.9) is well known, see [AP03; OR92] and can be seen by applying the method to ODEs with right hand sides of the form $f(t, u) := u + g(t)$.

3.3 Convolution Quadrature

Convolution Quadrature (CQ) was introduced by Christian Lubich in [Lub88a; Lub88b] as a method for discretizing convolution integrals and fractional derivatives. It possesses very favorable stability properties due to an implicit regularization in time. For the purpose of using it to discretize boundary integral equations, its main advantage is that it can approximate convolutions without having to evaluate the convolution kernel in the time domain. Instead it is sufficient that the Laplace transform of the kernel is well-behaved and known explicitly. In our applications, these considerations in the Laplace domain will lead to boundary integral methods for the Helmholtz equation with complex wave numbers, as introduced in Section 2.5. Convolution Quadrature comes in two flavors corresponding to an underlying time discretization via multistep or Runge-Kutta methods. We will introduce the basic principles of these two methods together with some convenient operational calculus notation in the following sections.

The general premise is to approximate convolution integrals of the form

$$u(t) := \int_0^t \kappa(t-\tau)g(\tau) d\tau \quad (3.10)$$

for a given kernel function $\kappa : \mathbb{R}_+ \rightarrow \mathbb{C}$ and data $g : \mathbb{R} \rightarrow \mathbb{C}$. We make the additional assumption that g is a causal function, meaning that $g(t) = 0$ for $t \leq 0$ and assume that the Laplace transform $K(s) := \mathcal{L}[\kappa](s)$ of κ can be computed explicitly, where we define $\mathcal{L}[\kappa](s) := \int_0^\infty e^{-s\tau} \kappa(\tau) d\tau$. We can rewrite (3.10) as

$$u(t) := \mathcal{L}^{-1}(K(\cdot) \mathcal{L}[g]) \quad (3.11)$$

as long as the inverse Laplace transform exists (this imposes restrictions on the decay properties of the Laplace transform $\mathcal{L}[g]$ which we specify later on). The motivation for considering convolutions of the form (3.10) when interested in wave propagation problems

and the related semigroups can be seen from Duhamel's formula (2.8). Similarly, the boundary integral equations will be derived in convolution form from this principle.

In order to emphasize the importance of the Laplace transform $K(s)$ in (3.11) over the kernel function κ , we introduce an operational calculus notation which is common in the literature on Convolution Quadrature (see [Lub94]; note that the corresponding operational calculus dates back much further, e.g., [Gen83] and [Yos84]).

Definition 3.9. *Let $K(s) : \mathcal{X} \rightarrow \mathcal{Y}$ be a family of bounded linear operators between Banach spaces \mathcal{X} and \mathcal{Y} that is analytic for $\operatorname{Re}(s) > 0$. We define*

$$K(\partial_t)g := \mathcal{L}^{-1}(K(\cdot)\mathcal{L}[g]),$$

where $g \in \operatorname{dom}(K(\partial_t))$ is such that the inverse Laplace transform exists, and the expression above is well defined.

Remark 3.10. [Say16, Chapter 3] contains a very general treatment of when the inverse Laplace transform exists using the theory of distributions. If $\mathcal{L}[g]$ decays sufficiently fast, the inverse can be computed using the Bromwich integral:

$$K(\partial_t)g(t) = \frac{1}{2\pi i} \int_{\sigma+i\mathbb{R}} e^{st} K(s) \mathcal{L}[g](s) ds, \quad \sigma > 0.$$

Proposition 3.11. *This operational calculus has the following important properties:*

- (i) For kernels $K_1(s)$ and $K_2(s)$, we have $K_1(\partial_t)K_2(\partial_t) = (K_1K_2)(\partial_t)$.
- (ii) For $K(s) := s$, we have: $K(\partial_t)g(t) = g'(t) \quad \forall g \in C^1(\mathbb{R}^+)$, with $g(0) = 0$.
- (iii) For $K(s) := s^{-1}$ we have: $K(\partial_t)g(t) = \int_0^t g(\xi) d\xi \quad \forall g \in C(\mathbb{R}^+)$.

The last statement motivates the notation ∂_t^{-1} for the integral, which will be important when we introduce a corresponding discrete version.

Proof. Part (i) follows directly by inserting the definitions of $K_1(\partial_t)[K_2(\partial_t)g]$. Part (ii) follows from the properties of the Laplace transform, see, e.g., [Say16, Section 2.3]. The last part then directly follows from (i) and (ii). \square

As an application of this calculus, we look at the following time-domain analog of representation formula (2.33).

Proposition 3.12 (Kirchhoff's representation formula, [Say16, Proposition 3.5.1]). *Let Γ be the boundary of a bounded Lipschitz domain, and let $u \in C^2(\mathbb{R}, H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma))$ solve the wave equation $\ddot{u} = \Delta u$ in $\mathbb{R}^d \setminus \Gamma$. Assume that $u(t) = 0$ for $t \leq 0$ and that an a priori estimate $\|u(t)\|_{H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)} \leq C(1 + t^\ell)$ holds for all $t \geq 0$ and some $\ell \in \mathbb{N}_0$.*

Then, u can be written as

$$u = -S(\partial_t) \llbracket \partial_\nu u \rrbracket + D(\partial_t) \llbracket \gamma u \rrbracket, \quad (3.12)$$

where S and D denote the single and double layer potentials from (2.32).

3.3.1 Multistep method based CQ

The multistep based Convolution Quadrature was the original version introduced in [Lub88a; Lub88b]. A detailed introduction to multistep Convolution Quadrature can also be found in [Say16, Chapter 4]. An important tool when working with CQ is the so-called Z -transform. For a sequence $g := (g_n)_{n \in \mathbb{N}_0}$, it is defined by

$$\mathcal{Z}[g](z) := \sum_{n=0}^{\infty} g_n z^n. \quad (3.13)$$

(Note: we use the complex parameter z instead of s , as it does not directly correspond to the Laplace parameter s . Instead the correspondence is $z = e^s$). We also identify a function f with the sequence $(f(t_n))_{n \in \mathbb{N}}$ when writing $\mathcal{Z}[f]$. Let $(\alpha_j)_{j=0}^m, (\beta_j)_{j=0}^m$ be the coefficients of an m -step method as described in Section 3.1.

In order to motivate the definitions for approximating (3.10), we follow [Lub88a] and make a (in our case purely formal) calculation. Let $g : \mathbb{R} \rightarrow \mathbb{C}$ be a causal function such that (3.10) is well defined. Starting with (3.10), we use the Bromwich integral to write $\kappa = \mathcal{L}^{-1}[K]$, and calculate for $\sigma > 0$:

$$\begin{aligned} u(t) &= \int_0^t \left(\frac{1}{2\pi i} \int_{\sigma+i\mathbb{R}} K(s) e^{st} ds \right) g(t-\tau) d\tau \\ &= \frac{1}{2\pi i} \int_{\sigma+i\mathbb{R}} K(s) \int_0^t e^{st} g(t-\tau) d\tau ds \\ &= \frac{1}{2\pi i} \int_{\sigma+i\mathbb{R}} K(s) \int_0^t e^{s(t-\tau)} g(\tau) d\tau ds. \end{aligned}$$

The inner integral is the solution to the ODE $y' = sy + g$ with $y(0) = 0$ by Duhamel's formula, which we will approximate by a multistep method. Let $k > 0$ denote the timestep size used for the approximation. Since $g(t) = 0$ for $t \leq 0$, we can equivalently start the equation with $y(-mt) = 0$ in order to get the initial conditions for the multistep method $y(-jt) = 0$ for $j = 0, \dots, m$. Denoting the approximation at time step $t_n := nk$ (as defined in (3.2)) by the multistep method as $y^k(t_n; s)$ gives:

$$u(t_n) \approx \frac{1}{2\pi i} \int_{i\mathbb{R}} K(s) y^k(t_n; s) ds =: u^k(t_n).$$

In order to get an explicit representation of $u^k(t_n)$, we take the Z -transform to get:

$$\mathcal{Z}[u^k] = \frac{1}{2\pi i} \int_{i\mathbb{R}} K(s) \mathcal{Z}[y^k(\cdot; s)] ds. \quad (3.14)$$

We need an explicit representation of $\mathcal{Z}[y^k(\cdot; s)]$. This can be derived via the Z -transform

of (3.2):

$$\begin{aligned}
 \frac{1}{k} \sum_{n=0}^{\infty} z^n \left[\sum_{j=0}^m \alpha_j y^k(t_{n-j}; s) \right] &= \sum_{n=0}^{\infty} z^n \sum_{j=0}^m \beta_j \left[s y^k(t_{n-j}; s) + g(t_{n-j}) \right] && \iff \\
 \frac{1}{k} \sum_{j=0}^m \alpha_j z^j \left[\sum_{n=0}^{\infty} z^{n-j} y^k(t_{n-j}; s) \right] &= \sum_{j=0}^m \beta_j z^j \left[\sum_{n=0}^{\infty} z^{n-j} \left(s y^k(t_{n-j}; s) + g(t_{n-j}) \right) \right] && \iff \\
 \frac{1}{k} \sum_{j=0}^m \alpha_j z^j \mathcal{Z} \left[y^k(\cdot; s) \right] &= \sum_{j=0}^m \beta_j z^j \left(s \mathcal{Z} \left[y^k(\cdot; s) \right] + \mathcal{Z} [g] \right) && \iff \\
 \left[\frac{\delta(z)}{k} - s \right] \mathcal{Z} \left[y^k(\cdot; s) \right] &= \mathcal{Z} [g], &&
 \end{aligned}$$

with the generating function

$$\delta(z) := \frac{\sum_{j=0}^m \alpha_j z^j}{\sum_{j=0}^m \beta_j z^j}. \quad (3.15)$$

Inserting this into (3.14) and using the Cauchy integral theorem gives:

$$\mathcal{Z} \left[u^k \right] (z) = \frac{1}{2\pi i} \int_{\sigma+i\mathbb{R}} K(s) \left[\frac{\delta(z)}{k} - s \right]^{-1} \mathcal{Z} [g] ds = K \left(\frac{\delta(z)}{k} \right) \mathcal{Z} [g].$$

Taking the inverse of the Z -transform in this equation then motivates the following definition:

Definition 3.13. *Let $K(s) : \mathcal{X} \rightarrow \mathcal{Y}$ be a family of bounded linear operators between Banach spaces which is analytic in s for $\operatorname{Re}(s) > 0$. Let $g : \mathbb{R} \rightarrow \mathcal{X}$ be a causal function. Let $(\alpha_j)_{j=0}^m, (\beta_j)_{j=0}^m$ originate from an A -stable multistep method. Then, we define for $t \in \mathbb{R}$:*

$$\left[K(\partial_t^k) g \right] (t) := \sum_{j=0}^{\infty} W_j g(t - j k), \quad (3.16)$$

where the operators $W_j : \mathcal{X} \rightarrow \mathcal{Y}$ are defined as the coefficients in the power series

$$\sum_{j=0}^{\infty} W_j z^j = K \left(\frac{\delta(z)}{k} \right) \quad (3.17)$$

and $\delta(z)$ is defined in (3.15).

The convolution weights in (3.17) are well defined due to the following mapping property of δ and the analyticity assumption on $K(s)$ for $s \in \mathbb{C}_+$.

Proposition 3.14 ([HW10, Chapter V.1, Theorem 1.5]). *If a multistep method is A -stable, then its generating function $\delta(z)$ as defined in (3.15) has no poles on the open unit disk and satisfies $\operatorname{Re}(\delta(z)) > 0$ for $|z| < 1$.*

3.3.2 Runge-Kutta based CQ

In order to get methods of order higher than two, the Runge-Kutta based Convolution Quadrature method was introduced in [LO93]. The general derivation is similar to the multistep case, replacing the approximation of the integral $\int_0^t e^{s(t-\tau)}g(\tau) d\tau$ with a Runge-Kutta based time-stepping. The analysis based on Laplace domain estimates of $K(s)$ for hyperbolic problems has been successively developed in [BL11; BLM11]. We will not use those results as we instead take a pure time domain approach in the later chapters.

Convolution Quadrature is only well defined for a certain subclass of methods. We make the following restrictions:

Assumption 3.15. *The Runge-Kutta method given by $(\mathcal{Q}, \mathbf{b}, \mathbf{c})$, satisfies:*

- (i) *the method is A-stable,*
- (ii) *the matrix \mathcal{Q} is invertible.*

We again introduce the Convolution Quadrature approximation using an operational calculus notation:

Definition 3.16. *Assume that the RK-method defined by $\mathcal{Q}, \mathbf{b}, \mathbf{c}$ satisfies Assumption 3.15. Let $K(s) : \mathcal{X} \rightarrow \mathcal{Y}$ be a family of bounded linear operators between Banach spaces, which is analytic in s for $\operatorname{Re}(s) > 0$, and let $g : \mathbb{R} \rightarrow \mathcal{X}$ be a causal function. We define the function*

$$\delta(z) := \left(\mathcal{Q} + \frac{z}{1-z} \mathbf{1b}^T \right)^{-1}. \quad (3.18)$$

Using this, we define an analogous calculus to (3.16) via:

$$\left[K(\partial_t^k)g \right] (t) := \sum_{j=0}^{\infty} W_j g(t - jk + k\mathbf{c}),$$

and the matrix operators $W_j : [\mathcal{X}]^m \rightarrow [\mathcal{Y}]^m$ are defined as the coefficients in the power series

$$\sum_{j=0}^{\infty} W_j z^j = K \left(\frac{\delta(z)}{k} \right), \quad (3.19)$$

where $K(\delta(z)/k)$ is defined using the Riesz-Dunford calculus introduced in Definition 2.40.

We also use the same notation if $G : \mathbb{R} \rightarrow [\mathcal{X}]^m$ is already vector valued:

$$\left[K(\partial_t^k)G \right] (t) := \sum_{j=0}^{\infty} W_j G(t - jk).$$

Using a simple post-processing, we can compute an improved approximation. For a function $U : \mathbb{R}_+ \rightarrow \mathbb{C}^m$ we define

$$\mathbb{G}[U](t) := 0 \quad \text{for } t \leq 0, \quad (3.20)$$

$$\mathbb{G}[U](t) := r(\infty) \mathbb{G}[U](t - k) + \mathbf{b}^T \mathcal{Q}^{-1} U(t) \quad \text{for } t > 0, \quad (3.21)$$

and analogously for Banach space valued functions. We immediately note that for stiffly accurate methods this definition simplifies to projecting to the last component of U .

Remark 3.17. Note that for Runge-Kutta methods, $[K(\partial_t^k)g](t) \in [\mathcal{Y}]^m$ is defined to be vector valued. Sometimes in the literature, the notation $[K(\partial_t^k)g](t)$ is instead used for the post processing $\mathbb{G} \left[\sum_{j=0}^{\infty} W_j g(t - j k + k\mathbf{c}) \right]$.

In order to see that (3.19) is well defined, we need the following statement about the spectrum of $\delta(z)$:

Proposition 3.18 ([BLM11, Lemma 2.6]). Assume that the Runge-Kutta matrix defined by \mathcal{Q} , \mathbf{b} , \mathbf{c} satisfies Assumption 3.15. Then, for $|z| < 1$, the spectrum of $\delta(z)$ satisfies

$$\sigma(\delta(z)) \subseteq \sigma(A^{-1}) \cup \{w \in \mathbb{C} : r(w)z = 1\},$$

where $r(w)$ denotes the stability function from Definition 3.4. Hence, if the Runge-Kutta method is A -stable, then $\sigma(\delta(z))$ lies in the open right half-plane $\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}$.

While most of the classical analysis of Convolution Quadrature methods are based on estimates in the Laplace domain and the relation of the CQ-approximation to the approximation of a Duhamel integral ([Lub88a; Lub88b; BL11; LO93; BLM11]), we take a different approach based on studying the approximation properties directly in the time domain. This approach has recently gained interest and was pioneered in [BLS15a; MR17]. The basis of this approach is the following lemma, which tells us that the Convolution Quadrature approximation corresponds to the solution of a time-stepping scheme. This relationship was already noticed in the early works on the topic (e.g. [LO93]). We formulate the transformation in a general lemma:

Lemma 3.19. Let \mathcal{X} be a Banach space and A denotes a closed, not necessarily bounded, linear operator on \mathcal{X} . Let functions $y : \mathbb{R}_+ \rightarrow \mathcal{X}$, $Y : \mathbb{R}_+ \rightarrow [\mathcal{X}]^m$, and $F : \mathbb{R}_+ \rightarrow [\mathcal{X}]^m$ be given such that the following relations hold:

$$Y(t_n) = y(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A]Y(t_n) + k[\mathcal{Q} \otimes \operatorname{Id}]F(t_n), \quad (3.22)$$

$$y(t_{n+1}) = r(\infty)y(t_n) + \mathbf{b}^T \mathcal{Q}^{-1}Y(t_n). \quad (3.23)$$

Assume that the Z -transforms $\widehat{Y} := \mathcal{Z}[Y]$ and $\widehat{F} := \mathcal{Z}[F]$ exist for sufficiently small z as power series in $[\mathcal{X}]^m$. Then, the Z -transforms solve the problem

$$-\frac{\delta(z)}{k}\widehat{Y} + \underline{A}\widehat{Y} = \left[\frac{1}{1 - r(\infty)z} k^{-1} \mathcal{Q}^{-1} \mathbf{1} \right] y(0) - \widehat{F}, \quad (3.24)$$

where the matrix-valued function $\delta(z)$ is defined as in (3.18).

Proof. We begin by noting a different representation of $\delta(z)$, which is a simple consequence of the Sherman-Morrison formula. Namely, for $|z| < 1$ we have

$$\delta(z) = \mathcal{Q}^{-1} - \frac{z\mathcal{Q}^{-1}\mathbf{1}b^T\mathcal{Q}^{-1}}{1 - zr(\infty)}. \quad (3.25)$$

We consider the Z -transform of (3.23). Multiplying by z^n and summing over $n \in \mathbb{N}_0$ gives

$$z^{-1}(\widehat{y} - y(0)) = r(\infty)\widehat{y} + \mathbf{b}^T \mathcal{Q}^{-1}\widehat{Y}$$

or equivalently after some simple manipulations:

$$\widehat{y} = \frac{z}{1 - r(\infty)z} \left(\mathbf{b}^T \mathcal{Q}^{-1} \widehat{Y} + z^{-1} y(0) \right). \quad (3.26)$$

The Z -transform of (3.22) is more involved, as it involves an unbounded operator. We multiply by z^n and sum up to a fixed $N \in \mathbb{N}$. This gives (after rearranging for the terms involving A):

$$k [\mathcal{Q} \otimes A] \sum_{n=0}^N z^n Y(t_n) = \sum_{z=0}^N z^n Y(t_n) - \sum_{z=0}^N z^n y(t_n) \mathbf{1} - k [\mathcal{Q} \otimes \text{Id}] \sum_{n=0}^N z^n F(t_n).$$

Since A is closed, $k[\mathcal{Q} \otimes A]$ is also a closed operator. By assumption, the Z -transforms \widehat{Y} , and \widehat{F} exist, and so does \widehat{y} by (3.26). Therefore by the equality above, the limit $k[\mathcal{Q} \otimes A] \sum_{n=0}^N Y^n z^n$ also exists for $N \rightarrow \infty$. Due to the closedness of A the equality becomes:

$$k [\mathcal{Q} \otimes A] \widehat{Y} = \widehat{Y} - \widehat{y} \mathbf{1} - k [\mathcal{Q} \otimes \text{Id}] \widehat{F}.$$

Inserting (3.26) this becomes:

$$k [\mathcal{Q} \otimes A] \widehat{Y} = \left[\text{Id} - \frac{z}{1 - r(\infty)z} \mathbf{1} \mathbf{b}^T \mathcal{Q}^{-1} \right] \widehat{Y} + \frac{1}{1 - r(\infty)z} y(0) \mathbf{1} - k [\mathcal{Q} \otimes \text{Id}] \widehat{F}.$$

Applying \mathcal{Q}^{-1} and rearranging the terms then gives the stated result. \square

3.4 Approximation of Semigroups via Runge-Kutta methods

In this section, we look at how well a semigroup can be approximated by using a Runge-Kutta method. We perform our analysis in a general Banach or Hilbert space setting, which will allow us to apply the results to a variety of problems, see Chapters 4 and 6. We note that, since we are interested in time domain BEM based approximation schemes rather than simple time stepping, the spaces we encounter later on will be somewhat non-standard and infinite dimensional.

For maximal generality, we consider the setting of Proposition 2.11. Let A_\star be a closed linear operator on a Banach space \mathcal{X} . The operator $B : \text{dom } A_\star \rightarrow \mathcal{M}$ is assumed to be linear and bounded, such that $A := A_\star|_{\ker B}$ generates a C_0 -semigroup. Then, the Runge-Kutta approximation of the Problem (2.11) is given by:

$$U^k(t_n) = u^k(t_n) \mathbf{1} + k [\mathcal{Q} \otimes A_\star] U^k(t_n) + k [\mathcal{Q} \otimes I] F(t_n + k \mathbf{c}), \quad (3.27a)$$

$$\underline{B} [U^k(t_n)] = \Xi(t_n + k \mathbf{c}), \quad (3.27b)$$

$$u^k(t_{n+1}) = u^k(t_n) + k [\mathbf{b}^T \otimes A_\star] U^k(t_n) + k [\mathbf{b}^T \otimes I] F(t_n + k \mathbf{c}), \quad (3.27c)$$

for $n \in \mathbb{N}_0$ and $u^k(0) := u_0$.

3.4.1 Rational functions of operators

We often work with rational functions of unbounded operators, as they provide a nice way of transferring results from the scalar theory of Runge-Kutta methods to the semigroup setting. The definitions are slightly awkward since rational functions over \mathbb{C} commute, but when dealing with operators we must pay attention to the domains.

Lemma 3.20. *Let A be an (unbounded) operator on a Banach space \mathcal{X} . Consider a family $R := (\xi_j, s_j)_{j=0}^n$ such that $\xi_j \in \mathbb{C}$, $s_j \in \{-1, 1\}$ and $\xi_j \in \rho(A)$ if $s_j = -1$. We call such a family a rational expression of A and associate an operator and a rational function over \mathbb{C} to it by defining:*

$$R(A) := \prod_{j=0}^n (A - \xi_j)^{s_j}, \quad R(z) := \prod_{j=0}^n (z - \xi_j)^{s_j}.$$

Let $q(z) := \frac{p_1(z)}{p_2(z)}$ be a rational function defined by the two polynomials

$$p_1(z) := \prod_{j=0}^{n_1} (z - \mu_j) \quad \text{and} \quad p_2(z) := \prod_{j=0}^{n_2} (z - \lambda_j).$$

If we assume $\lambda_j \notin \sigma(A) \forall j = 0, \dots, n_2$, the function $q(z)$ induces a rational expression $Q := (\xi_j, s_j)_{j=0}^{n_1+n_2+1}$ defined as

$$(\xi_j, s_j) := \begin{cases} (\mu_j, 1) & \text{for } j = 0, \dots, n_1, \\ (\lambda_{j-n_1-1}, -1) & \text{for } j = n_1 + 1, \dots, n_1 + n_2 + 1. \end{cases}$$

By construction, we get $Q(z) = q(z)$. We write $q(A) := Q(A)$ using this associated rational expression.

The following statements hold:

- (i) $R(A)$ is well defined and a linear operator $\text{dom}(R(A)) \rightarrow \mathcal{X}$.
- (ii) For two rational expressions R, Q , for which the associated rational functions coincide, i.e. $R(z) = Q(z) \forall z \in \mathbb{C}$, the associated operators coincide on the shared domain, i.e., $R(A)x = Q(A)x$ for $x \in \text{dom}(R(A)) \cap \text{dom}(Q(A))$.
- (iii) If $\sum_{j=0}^n s_j \leq 0$, then $R(A)$ has a bounded extension $B \in \mathcal{B}(\mathcal{X})$ with $R(A) \subseteq B$. The norm $\|B\|_{\mathcal{B}(\mathcal{X})}$ depends on q and $\|(A - \xi_j)^{-1}\|_{\mathcal{B}(\mathcal{X})}$ for $j = 0, \dots, n$ with $s_j = -1$.
For rational functions this means that, if $\deg(p_1) \leq \deg(p_2)$, then $q(A)$ is bounded.
- (iv) For two rational expressions R and Q , the rational expression $R \cdot Q$, defined by concatenating the tuples, satisfies $[R \cdot Q] = R(z)Q(z)$ and $[R \cdot Q](A) = R(A)Q(A)$.
- (v) For P, Q rational expressions and $\mu \in \mathbb{C}$, the operator $B := P(A) + \mu Q(A)$ has an extension, which can be written as $R(A)$ for a new rational expression R .

Proof. Part (i) is easily seen since we assumed that $\xi_j \in \rho(A)$ for $s_j = -1$.

Ad (ii): We write $R := (\xi_j, s_j)_{j=0}^{n_1}$ and $Q := (\tilde{\xi}_j, \tilde{s}_j)_{j=0}^{n_2}$. We note that for $x \in \text{dom}(A)$, the factors satisfy $(A - \xi_i)^{-1}(A - \xi_j)x = (A - \xi_j)(A - \xi_i)^{-1}x$ for $\xi_i \in \rho(A)$, $\xi_j \in \mathbb{C}$. Thus we can reorder the terms in the product, such that those with $s_j = -1$ are on the right-most side of the expression without changing the value (but possibly extending the domain). We also note that since linear factors $(A - \xi_i)$ commute and so do the factors $(A - \xi_i)^{-1}$, we can permute these factors internally. By this insight, we observe that we may assume that R (and by analogous considerations Q) is in simplified form, meaning that for $i, j \in \{1, \dots, n_1\}$ with $\xi_i = \xi_j$ we always have $s_i = s_j$, as otherwise we may permute the factors to cancel out.

Since the associated rational functions coincide over \mathbb{C} and both expressions are in simplified form, the coefficients must be a permutation of each other, i.e., $\tilde{\xi}_j = \xi_{P(j)}$ and $\tilde{s}_j = s_{P(j)}$. By the previous considerations we can reorder the expression to be the same for $R(A)$ and $Q(A)$.

Ad (iii): We reorder the coefficients and set $Q := (\tilde{\xi}_i, \tilde{s}_i) = (\xi_{P(i)}, s_{P(i)})$ such that $\tilde{s}_i = 1$ for the first n_1 terms and $\tilde{s}_i = -1$ for the others. Since $\sum_{j=0}^n s_j \leq 0$, we get that $n_1 \leq n - n_1$. Therefore, we can group the operators into pairs and define the operator:

$$Q(A) = \left[\prod_{j=0}^{n_1} (A - \tilde{\xi}_j)(A - \tilde{\xi}_{j+n_1})^{-1} \right] \left[\prod_{j=n_1+1}^n (A - \tilde{\xi}_{j+n_1})^{-1} \right]$$

and its corresponding rational expression Q . Since the pairings $(A - \tilde{\xi}_j)(A - \tilde{\xi}_{j+n_1})^{-1}$ are bounded with an operator norm depending only on $\tilde{\xi}_j, \tilde{\xi}_{j+n_1}$ and $\left\| (A - \tilde{\xi}_{j+n_1})^{-1} \right\|_{\mathcal{B}(\mathcal{X})}$, this operator is bounded. Also, as it corresponds to a reordering of the coefficients $(\xi_j, s_j)_{j=0}^n$, it is an extension of $R(A)$ via (ii), which proves (iii).

Part (iv) follows trivially from the definitions.

Part (v): Since we're only talking about extensions, let w.l.o.g. $P(A) = p_1(A)[p_2(A)]^{-1}$ and $Q(A) = q_1(A)[q_2(A)]^{-1}$ with p_i, q_i polynomials. Then, since the bounded operators $[p_2(A)]^{-1}$ and $[q_2(A)]^{-1}$ commute, we can write for $x \in \text{dom}(P(A)) \cap \text{dom}(P(Q))$:

$$(P(A) + \mu Q(A))x = [p_1(A)q_2(A) + p_2(A)q_1(A)][q_1(A)q_2(A)]^{-1}x.$$

Since $p_1(A)q_2(A) + p_2(A)q_1(A)$ is a polynomial in A , it can be written using linear factors. Together with the factors of $[q_1(A)q_2(A)]^{-1}$ this gives a new rational expression R with $R(z) = P(z) + \mu Q(z)$ and $P(A) + \mu Q(A) \subseteq R(A)$. \square

When working with Runge-Kutta methods, it is often the case that we would like to use (3.27a) in order to express an explicit formula for the stage vectors $U^k(t_n)$. In order to do so, we need to invert a system of equations involving the operator A . The fact that this can be done is subject of the next lemma.

Lemma 3.21. *Let A be a linear (not necessarily bounded) operator on a Banach space \mathcal{X} and $M \subseteq \mathbb{C}^{m \times m}$ be a matrix. Assume $\sigma(A) \cap \sigma(M) = \emptyset$ and the resolvent satisfies $\left\| (A - \mu \text{Id})^{-1} \right\|_{\mathcal{B}(\mathcal{X})} \leq C_\rho$ for $\mu \in \sigma(M)$.*

Then, the matrix operator $(\underline{A} - M \otimes \text{Id})$ mapping $[\text{dom}(A)]^m \rightarrow [\mathcal{X}]^m$ is invertible and satisfies

$$\left\| (\underline{A} - M \otimes \text{Id})^{-1} \right\|_{\mathcal{B}([\mathcal{X}]^m)} \leq CC_\rho.$$

The constant $C > 0$ depends only on M . The inverse can be written as a rational expression $\left[(\underline{A} - M \otimes \text{Id})^{-1} \right]_{ij} = R_{ij}(A)$, $i, j = 1, \dots, m$.

Proof. We can bring M to Jordan form, and write $M = XJX^{-1}$ where J consists of Jordan-blocks and most notably is an upper triangular matrix with eigenvalues $(\mu_j)_{j=1}^m$ on its diagonal. It is easy to see that we can write the inverse as

$$(\underline{A} - M \otimes \text{Id})^{-1} = [X \otimes \text{Id}]^{-1} (\underline{A} - J \otimes \text{Id})^{-1} [X \otimes \text{Id}]$$

where the left-hand side exists if and only if the right-hand side does. Due to its triangular structure, $(\underline{A} - J \otimes \text{Id})^{-1}$ can be constructed by a backward substitution, each entry involving the solution of a problem $(\underline{A} - \mu_j)^{-1}$ and linear combinations of the entries computed before. Since the resolvents exist and are bounded by C_ρ , the full inverse can then easily be constructed, and the bound follows from the triangle inequality. The form using rational expressions follows directly from this construction. \square

An analogous result to Lemma 3.21 also holds for two different classes of problems which will feature occasionally in this thesis. We compile these results here for ease of presentation.

Lemma 3.22. *Let A be a linear (not necessarily bounded) operator on a Banach space \mathcal{X} and $M \subseteq \mathbb{C}^{m \times m}$ be a given matrix. Let $B : \text{dom}(A) \rightarrow \mathcal{Y}$ be a linear operator representing a side constraint such that for all $\lambda \in \sigma(M)$ and $f \in \mathcal{X}$, $g \in \mathcal{Y}$ there exists a unique $u \in \text{dom}(A)$, satisfying*

$$Au - \lambda u = f \quad \text{and} \quad Bu = g.$$

Assume that this solution u satisfies the estimate $\|u\|_{\mathcal{X}} \leq C_\lambda (\|f\|_{\mathcal{X}} + \|g\|_{\mathcal{Y}})$. Then, the matrix valued problem, given $F \in [\mathcal{X}]^m$ and $G \in [\mathcal{Y}]^m$, find $U \in [\text{dom}(A)]^m$ such that

$$\underline{A}U - MU = F \quad \text{and} \quad \underline{B}U = G$$

has a unique solution, which satisfies

$$\|U\|_{[\mathcal{X}]^m} \leq C \left(\|F\|_{[\mathcal{X}]^m} + \|G\|_{[\mathcal{Y}]^m} \right).$$

The constant $C > 0$ depends only on M and the constants C_λ for $\lambda \in \sigma(M)$.

Lemma 3.23. *Let $a(\cdot, \cdot)$ be a bilinear form on a Banach space \mathcal{Y} . Let \mathcal{X} be a Hilbert space with $\mathcal{Y} \subseteq \mathcal{X}$ with continuous embedding. Assume that for $\lambda \in \sigma(M)$ and $f \in \mathcal{Y}'$ the scalar problems of finding $u \in \mathcal{Y}$ such that*

$$a(u, v) - \lambda \langle u, v \rangle_{\mathcal{X}} = \langle f, v \rangle_{\mathcal{Y}' \times \mathcal{Y}} \quad \forall v \in \mathcal{Y}$$

have a unique solution with $\|u\|_1 \leq C_\lambda \|f\|_2$ for some norms $\|\cdot\|_1$ on \mathcal{Y} and $\|\cdot\|_2$ on \mathcal{Y}' .

Given $F \in [\mathcal{Y}']^m$, the problem find $U = (U_j)_{j=0}^m \in [\mathcal{Y}]^m$ such that

$$\sum_{j=0}^m a(U_j, V_j) - \langle MU, V \rangle_{[\mathcal{X}]^m} = \langle F, V \rangle_{[\mathcal{Y}' \times \mathcal{Y}]^m} \quad \forall V = (V_j)_{j=0}^m \in [\mathcal{Y}]^m$$

has a unique solution, which satisfies

$$\|U\|_1 \leq \left(\max_{\lambda \in \sigma(M)} C_\lambda \right) C_M \|F\|_2$$

with the product versions of the norms $\|\cdot\|_1$ and $\|\cdot\|_2$. The constant $C_M > 0$ depends only on M .

Proof of Lemmas 3.22 and 3.23. Follows verbatim to the proof of Lemma 3.21, as we only needed that, after transforming to Jordan form, the scalar problems had unique solutions. \square

Before we can prove the stability of Runge-Kutta methods, we need two results from functional analysis. We will make use of the spectral theorem for self-adjoint operators.

Proposition 3.24 (Spectral theorem, [Wer07, Satz VII.3.1, page 354] or [RS80, Theorem VIII.4]). *Let A be a (not necessarily bounded) self-adjoint operator on a Hilbert space \mathcal{H} . Then there exists a measure space $\langle \mathcal{O}, \mu \rangle$, a measurable function $F : \mathcal{O} \rightarrow \mathbb{R}$, and a unitary map $U : \mathcal{H} \rightarrow L^2(\mathcal{O}, d\mu)$ with the following properties:*

- (i) $x \in \text{dom}(A)$ if and only if $F \cdot Ux \in L^2(\mu)$,
- (ii) $(UAU^{-1}f)(z) = F(z)f(z) \quad \forall z \in \mathcal{O}$.

We use this result in the following version:

Corollary 3.25. *Let A be a (not necessarily bounded) self-adjoint operator on a Hilbert space \mathcal{H} . Then, for a rational function $q(z)$ with $|q(z)| = 1$ for $z \in \mathbb{R}$, the operator $q(A)$ is an isometry.*

Proof. It is easy to see using the notation from Proposition 3.24 that, since the poles of q are not on the real line, $(A - \mu)^{-1}$ corresponds to multiplying by $(F(z) - \mu)^{-1}$. By writing the linear factors in the definition of $q(A)$ as $U^{-1}(F(z) - \xi)^{s_j}U$, and cancelling the pairings $U^{-1}U$, we get $q(A) = U^{-1}q(F(\cdot))U$, where $q(F(\cdot))$ denotes the multiplication operator.

Since U is unitary, this allows us to calculate the norms:

$$\begin{aligned} \|q(A)x\|_{\mathcal{H}}^2 &= \|q(F(\cdot))[Ux](z)\|_{L^2(\mu)}^2 = \int |q(F(z))|^2 |[Ux](z)|^2 d\mu(z) \\ &= \int |[Ux](z)|^2 d\mu(z) \\ &= \|x\|_{\mathcal{H}}^2. \end{aligned} \quad \square$$

When dealing with non-self-adjoint operators we will use the following result:

Proposition 3.26 (von Neumann). *Let \mathcal{H} be a Hilbert space and $q(z)$ a rational function with $|q(z)| \leq 1$ for $|z| \leq 1$. Then, for all bounded linear operators $T \in \mathcal{B}(\mathcal{X})$:*

$$\|T\|_{\mathcal{B}(\mathcal{X})} \leq 1 \quad \text{implies} \quad \|q(T)\|_{\mathcal{B}(\mathcal{H})} \leq 1.$$

Proof. See [Neu51, Section 4] or [RS90, Chapter XI, Section 154]. \square

3.4.2 Stability of Runge-Kutta time-stepping

The previous preparatory results allow us to show the following stability estimate:

Lemma 3.27 (Discrete Stability). *Let A be a linear, maximally dissipative operator on a Hilbert space \mathcal{H} . Then, for A -stable Runge-Kutta methods and arbitrary $k > 0$, we can bound*

$$\|r(kA)\|_{\mathcal{B}(\mathcal{H})} \leq 1. \quad (3.28)$$

If in addition $-A$ is also maximally dissipative (or equivalently iA is self-adjoint), and the RK-method satisfies $|r(it)| = 1$ for $t \in \mathbb{R}$, then the operator is an isometry:

$$\|r(kA)u\|_{\mathcal{H}} = \|u\|_{\mathcal{H}} \quad \forall u \in \mathcal{H}. \quad (3.29)$$

Proof. In order to show (3.28) we make use of results and ideas in [Neu51]. The results there are formulated for bounded operators A , but we will follow the same techniques. We consider the operator $\Phi := (A + \text{Id})(A - \text{Id})^{-1}$ (the existence of the inverse is guaranteed by the maximality assumption, see (2.7)). Then, Φ is a contraction by the following calculation for $u \in \text{dom}(A)$:

$$\begin{aligned} \|(A + \text{Id})u\|_{\mathcal{H}}^2 &= \langle (A + \text{Id})u, (A + \text{Id})u \rangle_{\mathcal{H}} = \|Au\|_{\mathcal{H}}^2 + 2 \operatorname{Re} \langle Au, u \rangle_{\mathcal{H}} + \|u\|_{\mathcal{H}}^2 \\ &= \langle (A - \text{Id})u, (A - \text{Id})u \rangle_{\mathcal{H}} + 4 \operatorname{Re} \langle Au, u \rangle_{\mathcal{H}} \\ &\leq \|(A - \text{Id})u\|_{\mathcal{H}}^2, \end{aligned}$$

where in the last step we used the dissipativity, i.e. $4 \operatorname{Re} \langle Au, u \rangle_{\mathcal{H}} \leq 0$. Setting $u := (A - \text{Id})^{-1}\varphi$ then gives $\|\Phi\varphi\|_{\mathcal{H}} \leq \|\varphi\|_{\mathcal{H}}$ or $\|\Phi\|_{\mathcal{B}(\mathcal{H})} \leq 1$.

Looking at the map $\mathcal{C} : z \mapsto \frac{z+1}{z-1}$, it is easy to see that for $\operatorname{Re}(z) < 0$ it holds that $|\mathcal{C}(z)| \leq 1$ and $\mathcal{C}(\mathcal{C}(z)) = z$. Thus \mathcal{C} maps $B_1(0)$ to $\{z \in \mathbb{C} : \operatorname{Re}(z) < 0\}$.

Since $|r(z)| \leq 1$ for $\operatorname{Re}(z) \leq 0$ due to A -stability, this implies that the rational function $g(z) := r\left(\frac{z+1}{z-1}\right)$ satisfies $|g(z)| \leq 1$ for $|z| \leq 1$ (where we have implicitly removed the singularity at $z = 1$, since $|g(1)| = |r(\infty)| \leq 1$). By Proposition 3.26 this implies $\|g(\Phi)\|_{\mathcal{B}(\mathcal{X})} \leq 1$.

Since $g(\Phi)$ and $r(A)$ are two rational expressions of A , both defining operators defined on all of \mathcal{X} , for which the rational functions satisfy $g(\mathcal{C}(z)) = r(\mathcal{C} \circ \mathcal{C}(z)) = r(z)$ by definition, we get that $g(\Phi) = r(A)$ by Lemma 3.20(ii).

If $-A$ is also maximally dissipative, this implies that iA is self adjoint by Proposition 2.9. We can apply Corollary 3.25 to the function $q(z) := r(-iz)$, to estimate $\|u\|_{\mathcal{H}} = \|q(iA)u\|_{\mathcal{H}} = \|r(-i^2A)u\|_{\mathcal{H}} = \|r(A)u\|_{\mathcal{H}}$. \square

In the previous theorem it is crucial that \mathcal{H} is a Hilbert space. For general Banach spaces the estimate does not hold. Instead, the following estimate was shown by Brenner and Thomée:

Proposition 3.28 ([BT79]). *Let A be the generator of a C_0 -semigroup on a Banach space \mathcal{X} . Then, for $nk \leq T$, there exist constants C and κ such that the following estimate holds:*

$$\|r(kA)^n\|_{\mathcal{B}(\mathcal{X})} \leq C n^{1/2} e^{\kappa\omega T},$$

where κ and C depend on the a priori estimates from (2.5) but are independent of k and n .

Now that we have studied the operator $r(kA)$, we need to make the connection to the Runge-Kutta approximation. We start with the following simple result, showcasing the calculus of rational functions.

Lemma 3.29. *For homogeneous constraints, i.e. $\Xi \equiv 0$, and $F \equiv 0$, we can reformulate (3.27) as*

$$u^k(t_{n+1}) = r(kA)u^k(t_n), \quad (3.30)$$

where $r(z)$ is the stability function (3.5).

Proof. For $\Xi \equiv 0$ and $F \equiv 0$, the problem becomes finding $U^k(t_n) \in [\text{dom}(A)]^m$ such that:

$$U^k(t_n) = u^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A]U^k(t_n) \quad (3.31)$$

$$u^k(t_{n+1}) = u^k(t_n) + k[\mathbf{b}^T \otimes A]U^k(t_n). \quad (3.32)$$

Bringing $k[\mathcal{Q} \otimes A]U^k(t_n)$ to the left-hand side, we can write the first equation by Lemma 3.21 as

$$U^k(t_n) = [\text{Id} - k\mathcal{Q} \otimes A]^{-1} u^k(t_n)\mathbf{1}.$$

Inserting this equation into (3.32) allows us to eliminate the stage vector $U^k(t_n)$ and get

$$u^k(t_{n+1}) = \left(1 + k[\mathbf{b}^T \otimes A][\text{Id} - k\mathcal{Q} \otimes A]^{-1}\mathbf{1}\right) u^k(t_n).$$

It is easy to check that $[\mathbf{b}^T \otimes A][\text{Id} - k\mathcal{Q} \otimes A]^{-1}\mathbf{1}$ is a bounded operator corresponding to a rational expression of A . Therefore, by 3.20(ii) we can write this using our calculus for rational functions as (3.30). \square

The Runge-Kutta approximation has the following stability property:

Lemma 3.30. *Let \mathcal{X} be a Hilbert space and A a maximally dissipative operator. For $f \in L^\infty([0, T], \mathcal{X})$ and $F \in L^\infty([0, T], [\mathcal{X}]^m)$, let U^k and u^k be a perturbed Runge-Kutta approximation, i.e., assume they solve:*

$$\begin{aligned} U^k(t_n) &= u^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A]U^k(t_n) + k[\mathcal{Q} \otimes \text{Id}]F(t_n), \\ u^k(t_{n+1}) &= u^k(t_n) + k[\mathbf{b}^T \otimes A]U^k(t_n) + kf(t_n). \end{aligned}$$

Then, the following estimate holds:

$$\left\| u^k(t_n) \right\|_{\mathcal{X}} \leq \left\| u^k(0) \right\|_{\mathcal{X}} + Ck \sum_{j=0}^{n-1} \left[\|F(t_j)\|_{[\mathcal{X}]^m} + \|f(t_j)\|_{\mathcal{X}} \right]. \quad (3.33)$$

For the stages this implies:

$$\left\| U^k(t_n) \right\|_{[\mathcal{X}]^m} \leq C \left\| u^k(0) \right\|_{\mathcal{X}} + Ck \sum_{j=0}^n \left[\|F(t_j)\|_{[\mathcal{X}]^m} + \|f(t_j)\|_{\mathcal{X}} \right]. \quad (3.34)$$

The constant $C > 0$ depends only on the Runge-Kutta method.

Proof. Using an analogous reformulation to Lemma 3.29, we can write

$$u^k(t_{n+1}) = r(kA)u^k + k [\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1} k\mathcal{Q}F(t_n) + kf(t_n).$$

The operator $k [\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1}$ can be rewritten as

$$k [\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1} = [\mathbf{b}^T \mathcal{Q}^{-1} \otimes \text{Id}] - \mathbf{b}^T [\text{Id} - k\mathcal{Q} \otimes A]^{-1},$$

which is a bounded operator by Lemma 3.21. Estimate (3.33) then follows from (3.28) and the discrete Gronwall lemma (see [Tho06, Lemma 10.5]). Estimate (3.34) follows from Lemma 3.21 and the estimate on the end points $u^k(t_n)$. \square

3.4.3 Convergence of Runge-Kutta methods for semigroups

The approximation properties of Runge-Kutta methods for semigroups have been extensively studied in the literature. Early results by Crouzeix, Brenner, and Thomé considered the case of homogeneous problems [BT79; Cro76]. Some inhomogeneities were later allowed in [BCT82]. In order to obtain convergence rates which exceed the stage order, the authors focus on solutions satisfying $u(t) \in \text{dom}(A^\ell)$ for $\ell \in \mathbb{N}$, which is not suitable when dealing with inhomogeneous boundary conditions. Later in [OR92], it was shown that sometimes one achieves fractional orders of convergence, depending on the assumption $u(t) \in \text{dom}(A^\mu)$, but still based on the assumption $\mu \geq 1$. The case $\mu \in [0, 1]$ was later addressed in [AP03], in which the authors proved that it is possible to regain a fractional convergence rate, assuming that the exact solution is in some interpolation space between \mathcal{X} and $\text{dom}(A)$.

We use the following notation, generalizing the interpolation spaces from Section 2.3.1

Definition 3.31. For a closed operator A on a Banach space \mathcal{X} , we define the interpolation space for $\mu \in \mathbb{N}_0$ as

$$\mathcal{X}_\mu := [\mathcal{X}, \text{dom}(A)]_{\mu,2} := \text{dom}(A^\mu),$$

where $\text{dom}(A^\mu)$ is equipped with the graph norm $\|A^\mu u\| := \sum_{\ell=0}^{\mu} \|A^\ell u\|_{\mathcal{X}}$.

For $0 \leq \mu \notin \mathbb{N}_0$, we define the spaces by interpolation as:

$$\mathcal{X}_\mu := [\mathcal{X}, \text{dom}(A)]_{\mu,2} := [\text{dom}(A^{\mu_0}), \text{dom}(A^{\mu_0+1})]_{\mu-\mu_0,2},$$

where $\mu_0 := \lfloor \mu \rfloor$ is the integer part of μ .

For $k > 0$ and $T > 0$, we define the quantity $\rho_k(T)$, which gives an estimate on the error propagation due to the Runge-Kutta method by

$$\rho_k(T) := \sup_{0 \leq nk \leq T} \|r(kA)^n\|_{\mathcal{B}(\mathcal{X})}. \quad (3.35)$$

Most notably, in the case of contraction semigroups on a Hilbert space, we have $\rho_k(T) \leq 1$ via Lemma 3.27.

The following two Propositions summarize the results of [AP03] in our notation.

Proposition 3.32 ([AP03, Theorem 1]). *Let the assumptions of Proposition 2.11 hold. Assume that the exact solution $u \in C^{p+1}([0, T], \mathcal{X}_\mu)$ for $\mu \geq 0$. Let u^k denote the Runge-Kutta approximation from (3.27). Then, there exist constants $k_0 > 0$, $C > 0$, such that for $0 < k \leq k_0$ and $0 < nk \leq T$ the following estimate holds:*

$$\left\| u(t_n) - u^k(t_n) \right\|_{\mathcal{X}} \leq CT \rho_k(T) k^{\min(p, q + \mu)} \sum_{\ell=q+1}^{p+1} \max_{\tau \in [0, T]} \left\| u^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu}. \quad (3.36)$$

The constant C depends on the Runge-Kutta method, μ , and the constants M and ω from Proposition 2.4. The constant k_0 depends only on the constants M and ω . If A generates a semigroup of contractions, k_0 can be chosen arbitrarily.

Proof. We only make some small remarks on the differences in notation. [AP03] uses a different definition of interpolation spaces, but the proof only relies on estimates of the form (2.15). The choice of k_0 follows from the fact that it is only needed to ensure that $(I - k\mathcal{Q} \otimes A)$ is invertible. For contraction semigroups we have that $\sigma(kA) \subseteq \mathbb{C}_-$ and $\sigma(\mathcal{Q}) \subseteq \mathbb{C}_+$. By Proposition 3.21 this inverse exists for $k > 0$. The assumption $\mu \leq p - q$, the authors of [AP03] made in their formulation of Theorem 1, can be replaced by using the rate $\min(p, q + \mu)$ in (3.36) as the spaces $\mathcal{X}_\mu \subseteq \mathcal{X}_{p-q}$ are nested for $\mu \geq p - q$. \square

For a subset of Runge-Kutta methods, these estimates can be improved:

Proposition 3.33 ([AP03, Theorem 2]). *Let the assumptions of Proposition 3.32 hold, and assume that, in addition, the RK-method satisfies*

$$|r(z)| < 1 \quad \text{for } \operatorname{Re}(z) \leq 0 \wedge z \neq 0 \quad \text{and} \quad r(\infty) \neq 1. \quad (3.37)$$

Then, there exist constants $k_0 > 0$, $C > 0$, such that for $0 < k \leq k_0$ and $0 < nk \leq T$ the following improved estimate holds:

$$\left\| u(t_n) - u^k(t_n) \right\|_{\mathcal{X}} \leq CT \rho_k(T) k^{\min(p, q + 1 + \mu)} \sum_{\ell=q+1}^{p+1} \max_{\tau \in [0, T]} \left\| u^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu}. \quad (3.38)$$

The constant C depends on the Runge-Kutta method, μ , and the constants M and ω from Proposition 2.4. The constant k_0 depends only on the constants M and ω . If A generates a semigroup of contractions it can be chosen arbitrarily.

Proof. Again, this is just a slight simplification of [AP03, Theorem 2]. We first note that, due to our assumption on $|r(z)|$, we are always in the case $m = 0$. Since we assumed that on the imaginary axis $r(z)$ is not equal to 1, we directly note that for sufficiently small $k \leq k_0$, all the zeros of $r(z) - 1$ satisfy $\operatorname{Re}(z) > k\omega$. By the resolvent bound (2.5) we therefore can estimate $\|(zI - kA)^{-1}\|_{\mathcal{B}(\mathcal{X})} \leq \frac{M}{\operatorname{Re}(z) - k_0\omega}$ for $\operatorname{Re}(z) < k_0\omega$, i.e., we have a uniform resolvent bound in the set $Z_{\alpha, \delta}$ (using their notation). We also note that we reformulated the convergence rate such that we do not have the restriction $\mu \leq p - q - 1$, since the exceptional cases are already proved by Theorem 3.32. \square

Remark 3.34. *The assumption $|r(z)| < 1$ for $\operatorname{Re}(z) \leq 1$ and $r(\infty) \neq 1$ is satisfied by the RadauIIa family of Runge-Kutta methods, but is violated by the Gauss methods, which satisfy $|r(z)| = 1$ on the imaginary axis.*

We need the following simple lemma when working with Z-transforms of semigroups.

Lemma 3.35. *Let the assumptions of Proposition 2.11 hold, i.e., A_\star is a closed operator on a Banach space \mathcal{X} , $B : \operatorname{dom}(A_\star) \rightarrow \mathcal{Y}$ is a surjective bounded linear operator such that $A := A_\star|_{\ker(B)}$ generates a C_0 -semigroup. Then there exists a constant $k_0 > 0$, such that for all $\widehat{F} \in [\mathcal{X}]^m$, $\widehat{\Xi} \in [\mathcal{Y}]^m$ and $|z| < 1$, the problem of finding $\widehat{U} \in [\operatorname{dom}(A)]^m$ such that*

$$\begin{aligned} -\frac{\delta(z)}{k}\widehat{U} + \underline{A}_\star\widehat{U} &= \widehat{F}, \\ \underline{B}\widehat{U} &= \widehat{\Xi} \end{aligned}$$

has a unique solution for $k \leq k_0$. The constant k_0 depends on the constants in (2.5) and can be chosen arbitrarily for the case that A is maximally dissipative.

Proof. To see existence, let $\widehat{W} \in [\operatorname{dom}(A_\star)]^m$ satisfy $\underline{B}\widehat{W} = \widehat{\Xi}$ (B was assumed surjective). By setting $\widehat{U} := \widehat{U}_0 + \widehat{W}$, where \widehat{U}_0 solves

$$-\frac{\delta(z)}{k}\widehat{U}_0 + \underline{A}\widehat{U}_0 = \widehat{F} + \frac{\delta(z)}{k}\widehat{W} - \underline{A}\widehat{W},$$

it is sufficient to show that the problem with homogeneous side constraint, i.e., $\widehat{\Xi} = 0$ has a solution. Since $\sigma(\delta(z)) \subseteq \mathbb{C}_+$ (Proposition 3.18), and $\mu \in \rho(A)$ for $\operatorname{Re}(\mu) > \omega$, where ω is the constant from (2.5). For k sufficiently small, we can apply Lemma 3.21 to get existence.

To see uniqueness, consider the difference of two solutions $\widehat{U} - \widehat{V}$. Due to linearity the difference satisfies the homogeneous problem, i.e., $\widehat{\Xi} = 0$ and $\widehat{F} = 0$, and we can apply Lemma 3.21 again. \square

Since we are interested in using the Runge-Kutta theory for boundary integral equations, we often need estimates in norms which are stronger than the natural norm of the underlying space. The next lemma is the basis for obtaining such stronger estimates (see also Lemma 2.13 for a continuous-in-time version).

Lemma 3.36. *Let k be sufficiently small such that Lemma 3.35 holds. Let U^k, u^k solve (3.27) with homogeneous initial conditions $u^k(t) = 0$ for $t \leq 0$. Define $V^k := k^{-1}Q^{-1}(U^k - u^k\mathbf{1})$ and $v^k := \mathbb{G}[V^k]$. Then, the following statements hold:*

(i) V^k, v^k solve (3.27) with modified right-hand sides:

$$V^k(t_n) = v^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A_\star]V^k(t_n) + k[\mathcal{Q} \otimes \text{Id}]G(t_n) \quad (3.39a)$$

$$\underline{B}[V^k(t_n)] = \Theta(t_n), \quad (3.39b)$$

with $G := \partial_t^k F$, $\Theta := \partial_t^k \Xi$.

(ii) V^k can be related to $\underline{A}_\star U^k$ via

$$\underline{A}_\star U^k(t_n) = V^k(t_n) - F(t_n + k\mathbf{c}).$$

(iii) If the Runge-Kutta method is stiffly accurate, the inhomogeneities can be simplified to $G(t) = k^{-1}\mathcal{Q}^{-1}(F(t + k\mathbf{c}) - F(t)\mathbf{1})$ and $\Theta(t) := k^{-1}\mathcal{Q}^{-1}(\Xi(t + k\mathbf{c}) - \Xi(t))$.

Proof. Ad (i): We prove that V^k solves (3.39) by showing that the Z -transform of V^k coincides with the Z -transform of the solutions to (3.39), which will be denoted by Y^k . By Lemma 3.19 and the definition of G and Θ , the transformed variable $\widehat{Y}^k := \mathcal{Z}[Y^k]$ solves:

$$\begin{aligned} -\frac{\delta(z)}{k}\widehat{Y}^k + \underline{A}_\star \widehat{Y}^k &= -\frac{\delta(z)}{k}\widehat{F}, \\ \underline{B}[\widehat{Y}^k] &= \frac{\delta(z)}{k}\widehat{\Xi}. \end{aligned}$$

On the other hand, $\widehat{V}^k = k^{-1}\mathcal{Q}^{-1}(\widehat{U}^k - \widehat{u}^k\mathbf{1})$. Using (3.26), this becomes

$$\widehat{V}^k = k^{-1}\mathcal{Q}^{-1}\left(\widehat{U}^k - \frac{z}{1-r(\infty)z}(\mathbf{b}^T\mathcal{Q}^{-1}\widehat{U}^k)\mathbf{1}\right) = \frac{\delta(z)}{k}\widehat{U}^k$$

via (3.25). Therefore, \widehat{V}^k solves:

$$\begin{aligned} -\frac{\delta(z)}{k}\widehat{V}^k + \underline{A}_\star \widehat{V}^k &= -\left(\frac{\delta(z)}{k}\right)^2\widehat{U}^k + \frac{\delta(z)}{k}\underline{A}_\star \widehat{U}^k \\ &= -\frac{\delta(z)}{k}\widehat{F}, \end{aligned}$$

and analogously for the constraint. By the uniqueness of the transformed problem, proved in Lemma 3.35, this proves (i).

Ad (ii): Follows directly from equation (3.27a).

Ad (iii): For stiffly accurate methods, we have $r(\infty) = 0$. Thus, we calculate

$$\delta(z) = \mathcal{Q}^{-1} - z\mathcal{Q}^{-1}\mathbf{1}\mathbf{b}^T\mathcal{Q}^{-1},$$

or, in terms of F :

$$\begin{aligned} [\partial_t^k F](t_n) &= k^{-1}\mathcal{Q}^{-1}F(t_n + k\mathbf{c}) - \mathcal{Q}^{-1}\mathbf{1}\mathbf{b}^T\mathcal{Q}^{-1}F(t_{n-1} + k\mathbf{c}) \\ &= k^{-1}\mathcal{Q}^{-1}F(t_n + k\mathbf{c}) - \mathcal{Q}^{-1}\mathbf{1}F(t_n), \end{aligned}$$

since stiffly accurate methods satisfy $\mathbf{b}^T\mathcal{Q}^{-1} = (0, \dots, 1)$ and $\mathbf{c}_m = 1$. A completely analogous computation for $\partial_t^k \Xi$ concludes the proof. \square

We now consider what happens with the approximation properties when using the discrete derivative or integral, i.e., apply ∂_t^k and $(\partial_t^k)^{-1}$. We prove statements analogous to Proposition 3.32 and Proposition 3.33. Due to the technical nature of these results, we defer the proofs to the end of the section.

Theorem 3.37. *Let u solve (2.11) and let U^k, u^k solve (3.27) with homogeneous initial conditions $u(t) = u^k(t) = 0$ for $t \leq 0$. Define $x(t) := \partial_t^{-1}u(t)$. Let \mathcal{X}_μ denote the interpolation space $[\mathcal{X}, \text{dom}(A)]_{\mu,2}$ for $\mu \in [0, \infty)$. Assume that the exact solution satisfies $u \in C^p([0, T], \mathcal{X}_\mu)$ and that $\mathcal{E}_B[\Xi^{(\ell)}] \in C^0([0, T], \mathcal{X}_\mu)$ for $\ell = -1, \dots, p$ as well as that $F \in C^p([0, T], \mathcal{X}_\mu)$. In addition, assume that k is sufficiently small such that Lemma 3.35 applies.*

Consider $X^k := (\partial_t^k)^{-1}U^k$ and $x^k := \mathbb{G}[X^k]$. Then, the following statements hold:

(i) *The functions X^k, x^k satisfy the system:*

$$X^k(t_n) = x^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A_\star]X^k(t_n) + k[\mathcal{Q} \otimes \text{Id}]G(t_n) \quad (3.40a)$$

$$\underline{B}X^k(t_n) = \Gamma^k(t_n), \quad (3.40b)$$

$$x^k(t_{n+1}) := x^k(t_n) + k[\mathbf{b}^T \otimes A_\star]X^k(t_n) + k[\mathbf{b}^T \otimes \text{Id}]G(t_n), \quad (3.40c)$$

$$x^k(t) = 0 \quad \text{for } t \leq 0, \quad (3.40d)$$

where $\Gamma^k(t_n)$ is itself given by a Runge-Kutta integration:

$$\begin{aligned} \Gamma^k(t_n) &= \gamma^k(t_n)\mathbf{1} + k\mathcal{Q}\Xi(t_n + k\mathbf{c}), \\ \gamma^k(t_{n+1}) &= \gamma^k(t_n) + k\mathbf{b}^T\Xi(t_n + k\mathbf{c}), \\ \gamma^k(t) &= 0 \quad \text{for } t \leq 0, \end{aligned}$$

and the new right-hand side is defined by $G := (\partial_t^k)^{-1}F$.

(ii) *For the approximation of the integral, there exists a parameter $k_0 > 0$ such that the following error estimates holds for all $0 \leq k \leq k_0$ and $nk \leq T$:*

$$\begin{aligned} &\left\|x(t_n) - x^k(t_n)\right\|_{\mathcal{X}} \leq CT\rho_k(T)k^{\min(q+\mu,p)} \\ &\times \sum_{\ell=q+1}^{p+1} \left(\max_{\tau \in [0, T]} \left\|x^{(\ell)}(\tau)\right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0, T]} \left\|\mathcal{E}_B[\Xi^{(\ell-1)}](\tau)\right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0, T]} \left\|F^{(\ell-1)}(\tau)\right\|_{\mathcal{X}_\mu} \right) \end{aligned} \quad (3.41)$$

with a constant C , which depends only on the Runge-Kutta method, μ , and the constants M and ω from Proposition 2.4. If A generates a semigroup of contractions k_0 can arbitrarily be chosen.

If the Runge-Kutta method satisfies the additional assumption (3.37), then (3.41) can

be improved to:

$$\begin{aligned} & \left\| x(t_n) - x^k(t_n) \right\|_{\mathcal{X}} \leq CT\rho_k(T)k^{\min(q+1+\mu,p)} \\ & \times \sum_{\ell=q+1}^{p+1} \left(\max_{\tau \in [0,T]} \left\| u^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0,T]} \left\| \mathcal{E}_B \left[\Xi^{(\ell-1)} \right] (\tau) \right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0,T]} \left\| F^{(\ell-1)}(\tau) \right\|_{\mathcal{X}_\mu} \right). \end{aligned} \quad (3.42)$$

A similar estimate can be shown for the difference quotients $\partial_t^k U^k$. We focus on the stiffly accurate case.

Theorem 3.38. *Let u be the exact solution to (2.11) and let v^k and V^k be defined as in Lemma 3.36. Assume that the Runge-Kutta method is stiffly accurate, i.e. $\mathbf{b}^T \mathbf{Q}^{-1} = (0, \dots, 1)$.*

Define $v := \dot{u}$ and assume that the exact solution satisfies $v \in C^{p+1}([0, T], \mathcal{X}_\mu)$ and that $\mathcal{E}_B [\Xi^{(\ell)}] \in C^0([0, T], \mathcal{X}_\mu)$ for $\ell = 1, \dots, p+2$ as well as $F \in C^{p+2}([0, T], \mathcal{X}_\mu)$. Then the following error estimate holds:

$$\begin{aligned} & \left\| v(t) - v^k(t_n) \right\|_{\mathcal{X}} \leq CTk^{\min(q+\mu-1, p-1)} \\ & \times \sum_{\ell=q+1}^{p+1} \left(\max_{\tau \in [0,T]} \left\| v^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0,T]} \left\| \mathcal{E}_B \left[\Xi^{(\ell+1)}(\tau) \right] \right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0,T]} \left\| F^{(\ell+1)}(\tau) \right\|_{\mathcal{X}_\mu} \right). \end{aligned} \quad (3.43)$$

If the method also satisfies (3.37), the estimate can be improved to

$$\begin{aligned} & \left\| v(t) - v^k(t_n) \right\|_{\mathcal{X}} \leq CTk^{\min(q+\mu, p)} \\ & \times \sum_{\ell=q+1}^{p+1} \left(\max_{\tau \in [0,T]} \left\| v^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0,T]} \left\| \mathcal{E}_B \left[\Xi^{(\ell+1)}(\tau) \right] \right\|_{\mathcal{X}_\mu} + \max_{\tau \in [0,T]} \left\| F^{(\ell+1)}(\tau) \right\|_{\mathcal{X}_\mu} \right). \end{aligned} \quad (3.44)$$

Remark 3.39. *Most of the effort in proving the above theorem is in order to obtain a convergence rate higher than q , even though the constraint in the stages only approximate with order q .*

Some Lemmata regarding Runge-Kutta methods

In order to prove Theorems 3.37 and 3.38, we need some more technical lemmas, which summarize properties of some rational functions when applied to the generator of a semigroup.

Lemma 3.40. *Let A be the generator of a C_0 -semigroup on a Banach space \mathcal{X} , and let $\mathcal{X}_\mu := [\mathcal{X}, \text{dom}(A)]_{\mu, 2}$ for $\mu \geq 0$. Assume that the Runge-Kutta method given by \mathcal{Q} , \mathbf{b} , \mathbf{c} satisfies (3.37).*

We define the following rational functions for $\beta \in \{-1, 0, 1\}$:

$$\begin{aligned} g(z) &:= z \mathbf{b}^T [\text{Id} - z \mathcal{Q}]^{-1}, \\ r(z) &:= 1 + z \mathbf{b}^T [\text{Id} - z \mathcal{Q}]^{-1} \mathbf{1}, \\ r_{\ell, \beta}(z) &:= z \mathbf{b}^T [\text{Id} - z \mathcal{Q}]^{-1} \mathcal{Q}^\beta (\mathbf{c}^\ell - \ell \mathcal{Q} \mathbf{c}^{\ell-1}). \end{aligned}$$

Then, there exists a constant $k_0 > 0$, depending on the constants M and ω from (2.5), such that for $\ell \leq p$, $0 < k \leq k_0$ with $nk \leq T$ and $\beta = 0, 1$ the following estimates hold for all $x \in \mathcal{X}_\mu$ and $X \in [\mathcal{X}]^m$:

$$\left\| \sum_{j=0}^n [r(kA)]^j r_{\ell, \beta}(kA) x \right\|_{\mathcal{X}} \leq C \rho_k(T) k^{\min(\mu, p - \ell - \beta)} \|x\|_{\mathcal{X}_\mu}, \quad (3.45)$$

$$\left\| \sum_{j=0}^n [r(kA)]^j g(kA) X \right\|_{\mathcal{X}} \leq C \rho_k(T) \|X\|_{[\mathcal{X}]^m}. \quad (3.46)$$

The constant $C > 0$ depends only on the Runge-Kutta method, the constants M and ω from (2.5), k_0 , ℓ , and μ , but is independent of n , k , or x . If the Runge-Kutta method is stiffly accurate, estimate (3.45) also holds for $\beta = -1$. If A generates a C_0 -semigroup of contractions, k_0 can be chosen arbitrarily.

Proof. We modify the proof of [AP03, Lemma 6], which only covers the case $\beta = 0$. We first assume $p - \ell - \beta \geq 0$ and fix $\mu \in \mathbb{N}_0$ such that $\mu \leq p - \ell - \beta$ (for $\mu \geq p - \ell - \beta$ we can just redefine $\mu := p - \ell - \beta$ without changing the claimed estimate). Define the rational function

$$f_{\ell, \mu}^\beta(z) := \frac{r_{\ell, \beta}(z)}{(r(z) - 1) z^\mu}.$$

Since we assumed $r(\infty) \neq 1$, it is easy to see that $f_{\ell, \mu}^\beta$ is bounded for $z \rightarrow \infty$. By considering the RK-approximation of the ODE $y' = zy$, we get that $r(z)$ approximates e^z with order $p + 1$, and since $p \geq 1$ this means $r(z) - 1 = z + z^2/2 + \dots$ for $z \rightarrow 0$, and thus 0 is a simple root of the function $r(z) - 1$.

We recall the order conditions from (3.9):

$$\mathbf{b}^T \mathcal{Q}^j \mathbf{c}^\ell = \frac{1}{(j + \ell + 1)(j + \ell) \cdots (\ell + 1)}, \quad 0 \leq j + \ell \leq p - 1. \quad (3.47)$$

We now have to distinguish the different cases for $\beta = -1, 0, 1$. We prove that in all three cases we have that $r_{\ell, \beta}(z) = \mathcal{O}(z^{p+1-\ell-\beta})$ for $z \rightarrow 0$ (for the case $\mu = -1$ we need the additional assumption of stiff accuracy).

For $\beta = -1$ we first note that since we assumed the method to be stiffly accurate, we can calculate using $\mathbf{c} = \mathcal{Q}\mathbf{1}$ and the order conditions:

$$\mathbf{c}_m = \mathbf{b}^T \mathcal{Q}^{-1} \mathbf{c} = \mathbf{b}^T \mathcal{Q}^{-1} \mathcal{Q}\mathbf{1} = \mathbf{b}^T \mathbf{1} = 1.$$

Using the definition of $r_{\ell,-1}$ and the Neumann series for $(\text{Id} - z\mathcal{Q})^{-1}$ gives:

$$\begin{aligned} r_{\ell,-1}(z) &= z\mathbf{b}^T(\text{Id} - z\mathcal{Q})^{-1}\mathcal{Q}^{-1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}) = z\mathbf{b}^T \sum_{j=0}^{\infty} z^j \mathcal{Q}^j \mathcal{Q}^{-1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}) \\ &= z\mathbf{b}^T \mathcal{Q}^{-1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}) + z \sum_{j=1}^{\infty} z^j \mathbf{b}^T \mathcal{Q}^{j-1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}) \\ &= z \sum_{j=1}^{\infty} z^j \mathbf{b}^T \mathcal{Q}^{j-1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}), \end{aligned}$$

where in the last step, we used that $\mathbf{b}^T \mathcal{Q}^{-1} \mathbf{c}^\ell = c_m^\ell = 1$ by our assumptions together with $\mathbf{b}^T \mathbf{c}^{\ell-1} = \frac{1}{\ell}$ by the order condition to eliminate the first term. By the order conditions, the next $p - \ell$ terms also vanish, i.e., we can start the sum at $j = p + 1 - \ell$, which means that $r_{\ell,-1}(z) = \mathcal{O}(z^{p+2-\ell})$ for $z \rightarrow 0$.

For $\beta = 0$ an analogous computation, but using only the order conditions gives:

$$r_{\ell,0}(z) = z\mathbf{b}^T(\text{Id} - z\mathcal{Q})^{-1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}) = z \sum_{j=p-\ell}^{\infty} z^j \mathbf{b}^T \mathcal{Q}^j (\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}),$$

which proves that $r_{\ell,0}(z) = \mathcal{O}(z^{p+1-\ell})$ for $z \rightarrow 0$.

For $\beta = 1$, the order conditions lead to one less term vanishing, i.e., we have

$$r_{\ell,1}(z) := z\mathbf{b}^T(\text{Id} - z\mathcal{Q})^{-1}\mathcal{Q}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}) = z \sum_{j=p-1-\ell}^{\infty} z^j \mathbf{b}^T \mathcal{Q}^{j+1}(\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}),$$

which gives $r_{\ell,1}(z) = \mathcal{O}(z^{p-\ell})$.

Since we assumed $1 + \mu \leq p + 1 - \ell - \beta$, and $|r(z)| < 1$ for $z \in i\mathbb{R} \setminus \{0\}$, we get that $f_{\ell,\mu}^\beta$ has no pole in the closed left-half plane. For k_0 sufficiently small, we therefore get that $\sigma(kA) \subseteq \{z \in \mathbb{C} : \text{Re}(z) \leq k_0\omega\}$ and the set of poles of $f_{\ell,\mu}^\beta$ are uniformly separated w.r.t. k (for semigroups of contractions, we have $\omega = 0$ and therefore k_0 can be arbitrary). We apply Proposition 3.20(iii) to get that $f_{\ell,\mu}^\beta(kA)$ is a bounded linear operator with the constant depending only on $f_{\ell,\mu}^\beta$ and the distance of $\sigma(kA)$ to the poles of $f_{\ell,\mu}^\beta$ by (2.5).

Using $f_{\ell,\mu}^\beta$ allows us to write:

$$\sum_{j=0}^n r(z)^j r_{\ell,\beta}(z) z^{-\mu} = \frac{r(z)^{n+1} - 1}{(r(z) - 1)z^\mu} r_{\ell,\beta}(z) = (r(z)^{n+1} - 1) f_{\ell,\mu}^\beta(z).$$

Setting $z = kA$ gives for $x \in \mathcal{X}_\mu$ (implicitly using Lemma 3.20(ii)):

$$\sum_{j=0}^n r(kA)^j r_{\ell,\beta}(kA)x = (r(kA)^{n+1} - 1) f_{\ell,\mu}^\beta(kA) k^\mu A^\mu x.$$

Taking the norm, using the definition of $\rho_k(T)$ and interpolating between $[\mu]$ and $[\mu] + 1$, then concludes the proof for $r_{\ell,\beta}$ in the case $p - \ell - \beta \geq 0$.

An analogous argument proves the estimate involving g . The important part is that we get $g(z) = \mathcal{O}(z)$ for $z \rightarrow 0$ also in this case, as the function $(\text{Id} - z\mathcal{Q})^{-1} \rightarrow \text{Id}$. Since $z(\text{Id} - z\mathcal{Q})^{-1}$ can be bounded independently of z for $\text{Re}(z) \leq 0$, this implies that the function $\frac{g(z)}{r(z)^{-1}}$ also has no poles and is bounded in $\{z \in \mathbb{C} : \text{Re}(z) \leq 0\}$. The rest of the argument follows analogously.

It remains to show the estimate involving $r_{\ell,\beta}$ in the case $p - \ell < \beta$. Since $\ell \leq p$ this only happens for $\beta = 1$. Since we can write $r_{\ell,\beta}(z) = g(z)\mathcal{Q}^\beta (\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1})$, the stated estimate is weaker than the previous estimate involving just $g(kA)$. \square

When dealing with Runge-Kutta methods which do not satisfy the additional assumption (3.37), we still have the following result:

Lemma 3.41. *Let A be the generator of a C_0 -semigroup on a Banach space \mathcal{X} , and let $\mathcal{X}_\mu := [\mathcal{X}, \text{dom}(A)]_{\mu,2}$ for $\mu \geq 0$ as in Definition 3.31. Define $r_{\ell,\beta}$ as in Lemma 3.40.*

Then there exists a constant $k_0 > 0$, depending on the constants M and ω from (2.5), such that for $\ell \leq p$, $0 < k \leq k_0$ and $\mu = 0, 1$ the following estimates hold for all $x \in \mathcal{X}_\mu$:

$$\|r_{\ell,\beta}(kA)x\|_{\mathcal{X}} \leq C k^{\min(\mu, p+1-\ell-\beta)} \|x\|_{\mathcal{X}_\mu}. \quad (3.48)$$

The constant $C > 0$ depends only on the Runge-Kutta method, the constants M and ω from (2.5), k_0 , ℓ , and μ , but is independent of k or x .

If the Runge-Kutta method is stiffly accurate, (3.48) also holds for $\beta = -1$. If A generates a C_0 -semigroup of contractions, k_0 can be chosen arbitrarily.

Proof. The proof follows using similar techniques to Lemma 3.40. We fix $\mu \in \mathbb{N}_0$ such that $\mu \leq p + 1 - \ell - \beta$. We have already established in the proof of Lemma 3.40 that $r_{\ell,\beta} = \mathcal{O}(z^{p+1-\ell-\beta})$ for $z \rightarrow 0$ and that $r_{\ell,\beta}$ is bounded for $z \in \mathbb{C}_-$. Therefore, we can write $r_{\ell,\beta}(z) = q(z)z^\mu$, where $q(z)$ is a bounded rational function on $\overline{\mathbb{C}_-}$. For $k \leq k_0$ sufficiently small, the poles of q and the spectrum of kA are uniformly separated (for contraction semigroups the spectrum of A is in the left half plane so the condition is always satisfied). Using Lemma 3.20(ii), we get:

$$\|r_{\ell,\beta}(kA)x\|_{\mathcal{X}} \leq \|q(kA)(kA)^\mu x\|_{\mathcal{X}} \lesssim k^\mu \|A^\mu x\|_{\mathcal{X}}.$$

The case for non-integer μ follows again by interpolation. \square

Lemma 3.42. *Let A generate a C_0 -semigroup on a Banach space \mathcal{X} . Assume $u \in C^{p+1}([0, T], \mathcal{X}_\mu)$ for some $\mu \geq 0$. For $\beta \in \{-1, 0, 1\}$, we define the rational functions $g(z)$, $r(z)$ and $r_{\ell,\beta}(z)$ as in Lemma 3.40. We set $\alpha := 1$ if the Runge-Kutta method satisfies (3.37), i.e., $|r(z)| < 1$ for $z \in i\mathbb{R} \setminus \{0\}$ and $r(\infty) \neq 1$. Otherwise we set $\alpha := 0$.*

Define $e(t_n)$ as

$$e(t_n) := \sum_{j=0}^n [r(kA)]^{n-j} g(kA) (k\mathcal{Q})^\beta [u(t_j + k\mathbf{c}) - u(t_j)\mathbf{1} - k\mathcal{Q}\dot{u}(t_j + k\mathbf{c})].$$

Then, the following error estimates hold for $\beta \in \{0, 1\}$:

$$\|e(t_n)\|_{\mathcal{X}} \leq C \rho_k(T) T k^{\min(q+\alpha+\beta+\mu, p)} \sum_{\ell=0}^{p+1} \max_{\tau \in [0, T]} \|u^{(\ell)}(\tau)\|_{\mathcal{X}_\mu}.$$

If we assume that the Runge-Kutta method is stiffly accurate, i.e., $\mathbf{b}^T \mathcal{Q}^{-1} = (0, \dots, 1)$, then the following estimate holds for $\beta = -1$:

$$\|e(t_n)\|_{\mathcal{X}} \leq CT\rho_k(T)k^{\min(q+\mu-1, p-1)+\alpha} \sum_{\ell=0}^{p+1} \max_{\tau \in [0, T]} \|u^{(\ell)}(\tau)\|_{\mathcal{X}_\mu}.$$

Proof. We prove both estimates at the same time, the only place where the stiff accuracy enters is in the applicability of Lemma 3.40. The case $\beta = 0$ was shown in [AP03, Theorem 1 and 2], we modify their proof slightly to also account for $\beta = \pm 1$.

For ease of notation, we introduce a new symbol for the term in the rightmost bracket of the definition of $e(t_n)$ and write:

$$\Delta^k(t) := [u(t + k\mathbf{c}) - u(t)\mathbf{1} - k\mathcal{Q}\dot{u}(t + k\mathbf{c})]. \quad (3.49)$$

In the case $\alpha = 1$, i.e., the Runge-Kutta method satisfies (3.37), we sum by parts, writing $S_n(z) := \sum_{j=0}^n [r(z)]^j$. This gives:

$$\begin{aligned} e(t_n) &= S_{n-1}(kA)g(kA)(k\mathcal{Q})^\beta \Delta^k(t_0) + S_0(kA)g(kA)(k\mathcal{Q})^\beta \Delta^k(t_n) \\ &\quad + \sum_{j=1}^n S_{n-j}(kA)g(kA)(k\mathcal{Q})^\beta [\Delta^k(t_j) - \Delta^k(t_{j-1})] \\ &=: e_1(t_n) + \sum_{j=1}^n e_2^{(j)}(t_n). \end{aligned} \quad (3.50)$$

For fixed $j \in \mathbb{N}_0$, we write $u(t_j + t) = \pi_j(t) + \varphi_j(t)$ as its Taylor polynomial π_j of order p , centered at t_j , with remainder term φ_j . This means, we can write $\Delta^k(t_j)$ as:

$$\begin{aligned} \Delta^k(t_j) &= [u(t_j + k\mathbf{c}) - u(t_j)\mathbf{1} - k\mathcal{Q}\dot{u}(t_j + k\mathbf{c})] \\ &= \pi_j(k\mathbf{c}) + \varphi_j(k\mathbf{c}) - \pi_j(0)\mathbf{1} - k\mathcal{Q}\dot{\pi}_j(k\mathbf{c}) - k\mathcal{Q}\dot{\varphi}_j(k\mathbf{c}) \\ &= \sum_{\ell=1}^p k^\ell [\mathbf{c}^\ell - \mathcal{Q}\ell\mathbf{c}^{\ell-1}] u^{(\ell)}(t_j) + \varphi_j(k\mathbf{c}) - k\mathcal{Q}\dot{\varphi}_j(k\mathbf{c}) \\ &=: \sum_{\ell=1}^p k^\ell [\mathbf{c}^\ell - \mathcal{Q}\ell\mathbf{c}^{\ell-1}] u^{(\ell)}(t_j) + \theta(t_j) \end{aligned} \quad (3.51)$$

with $\theta(t_j) := \varphi_j(k\mathbf{c}) - k\mathcal{Q}\dot{\varphi}_j(k\mathbf{c})$ collecting the remainder terms. Due to the stage order condition (3.8), the first q terms of the above sum vanish, giving:

$$\Delta^k(t_j) = \sum_{\ell=q+1}^p k^\ell [\mathbf{c}^\ell - \mathcal{Q}\ell\mathbf{c}^{\ell-1}] u^{(\ell)}(t_j) + \theta(t_j).$$

From the integral representation of the remainder term φ_j , one can see that the following estimates hold for $0 \leq t \leq k$:

$$\|\varphi_j(t)\|_{\mathcal{X}} \lesssim \frac{k^{p+1}}{(p+1)!} \max_{\tau \in [t_j, t_{j+1}]} \|u^{(p+1)}(\tau)\|_{\mathcal{X}}, \quad \|\dot{\varphi}_j(t)\|_{\mathcal{X}} \lesssim \frac{k^p}{p!} \max_{\tau \in [t_j, t_{j+1}]} \|u^{(p+1)}(\tau)\|_{\mathcal{X}},$$

which implies $\|\theta(t_j)\|_{\mathcal{X}} = \mathcal{O}(k^{p+1})$.

Using the definition of the rational function $r_{\ell,\beta}(z)$, we can rewrite the error terms in (3.50) as

$$\begin{aligned} e_1(t_n) &= \sum_{\ell=q+1}^p \frac{k^{\ell+\beta}}{\ell!} S_{n-1}(kA) r_{\ell,\beta}(kA) u^{(\ell)}(t_0) \\ &\quad + S_{n-1}(kA) g(kA) (k\mathcal{Q})^\beta \theta(t_0) \\ e_2^{(j)}(t_n) &= \sum_{\ell=q+1}^p \frac{k^{\ell+\beta}}{\ell!} S_{n-1}(kA) r_{\ell,\beta}(kA) \left(u^{(\ell)}(t_j) - u^{(\ell)}(t_{j-1}) \right) \\ &\quad + S_{n-1}(kA) g(kA) (k\mathcal{Q})^\beta [\theta(t_j) - \theta(t_{j-1})]. \end{aligned}$$

We first look at the contributions $e_2^{(j)}(t_n)$. Using the mean value theorem, we can write for some $\xi \in [t_{j-1}, t_j]$:

$$\left\| u^{(\ell)}(t_j) - u^{(\ell)}(t_{j-1}) \right\|_{\mathcal{X}_\mu} = \left\| \int_{t_{j-1}}^{t_j} u^{(\ell+1)}(\xi) d\xi \right\|_{\mathcal{X}_\mu} \leq k \max_{\xi \in [t_{j-1}, t_j]} \left\| u^{(\ell+1)}(\xi) \right\|_{\mathcal{X}_\mu}.$$

Applying the results of Lemma 3.40, we can conclude that

$$\begin{aligned} \left\| e_2^{(j)}(t_n) \right\|_{\mathcal{X}} &\lesssim k^{1+\min(q+1+\beta+\mu,p)} \sum_{\ell=q+1}^p \max_{\tau \in [0, T]} \left\| u^{(\ell+1)}(\tau) \right\|_{\mathcal{X}_\mu} \\ &\quad + \left\| S_{n-1}(kA) g(kA) (k\mathcal{Q})^\beta [\theta(t_j) - \theta(t_{j-1})] \right\|. \end{aligned}$$

To estimate the last term, we recall the definition of $\theta(t_j)$ and reorder the terms to

$$\theta(t_j) - \theta(t_{j-1}) = (\varphi_j(k\mathbf{c}) - \varphi_{j-1}(k\mathbf{c})) - k\mathcal{Q}(\dot{\varphi}_j(k\mathbf{c}) - \dot{\varphi}_{j-1}(k\mathbf{c})).$$

We use the integral form of the remainder in Taylor's formula and calculate:

$$\varphi_j(t) - \varphi_{j-1}(t) = \int_0^t \frac{(t-\tau)^p}{p!} \left(u^{(p+1)}(t_j + \tau) - u^{(p+1)}(t_j + \tau - k) \right) d\tau$$

which gives $\|\varphi_j(t) - \varphi_{j-1}(t)\|_{\mathcal{X}} = \mathcal{O}(k^{p+2})$ by again writing the difference $u^{(p+1)}(\tau) - u^{(p+1)}(\tau - k)$ as an integral. Analogously, we can estimate $\dot{\varphi}_j(t) - \dot{\varphi}_{j-1}(t)$ to be $\mathcal{O}(k^{p+1})$. Combined, this gives the estimate $\theta(t_j) - \theta(t_{j-1}) = \mathcal{O}(k^{p+2})$. Together with the stability of $S_{n-1}(kA)g(kA)$ from Lemma 3.40 this gives:

$$\left\| e_2^{(j)}(t_n) \right\|_{\mathcal{X}} \leq k^{1+\min(q+1+\beta+\mu,p)} \sum_{\ell=q+1}^p \max_{\tau \in [0, T]} \left\| u^{(\ell+1)}(\tau) \right\|_{\mathcal{X}_\mu} + \mathcal{O}(k^{p+2+\beta}).$$

The term $e_1(t_n)$ can be estimated to be $\mathcal{O}(k^{\min(q+1+\beta+\mu,p)})$ using similar estimates but without the extra power of k due differences $u^{(\ell)}(t_j) - u^{(\ell)}(t_{j-1})$ etc. Inserting all these

estimates into (3.50) and using the estimate $nk \leq T$ gives the stated estimates for the case $\alpha = 1$.

In the case $\alpha = 0$, the proof is even simpler. We start by applying the triangle inequality and the definition of $\rho_k(T)$ to get:

$$\|e(t_n)\|_{\mathcal{X}} \leq \rho_k(T) \sum_{j=0}^n \left\| g(kA) (k\mathcal{Q})^\beta \Delta^k(t_j) \right\|_{\mathcal{X}}.$$

For fixed $j \in \mathbb{N}_0$, we write $\Delta^k(t_j)$ as its Taylor polynomial of order p (cf. (3.51)). Since the remainder term takes the form $k^\beta g(kA) \mathcal{Q}^\beta (\varphi(\mathbf{c}) - k\mathcal{Q}\dot{\varphi}(\mathbf{c}))$ and is of order $k^{p+1+\beta}$, we get by applying Lemma 3.41:

$$\begin{aligned} \left\| g(kA) (k\mathcal{Q})^\beta \Delta^k(t_j) \right\|_{\mathcal{X}} &\lesssim \sum_{\ell=q+1}^p k^{\ell+\beta} \left\| r_{\ell,\beta}(kA) u^{(\ell)}(t_j) \right\|_{\mathcal{X}} + \mathcal{O}(k^{p+1+\beta}) \\ &\lesssim \sum_{\ell=q+1}^p k^{\ell+\beta} k^{\min(\mu, p-\ell-\beta)} \left\| u^{(\ell)}(t_j) \right\|_{\mathcal{X}_\mu} + \mathcal{O}(k^{p+1+\beta}) \\ &\lesssim k^{\min(q+1+\beta+\mu, p+1)} \left\| u^{(\ell)}(t_j) \right\|_{\mathcal{X}_\mu} + \mathcal{O}(k^{p+1+\beta}). \end{aligned}$$

Summing over j then completes the proof. \square

Proofs of Theorem 3.37 and Theorem 3.38

We can now finally give the proofs of Theorem 3.37 and Theorem 3.38. We start with a remark on notation. Since the spectral properties are very important for the analysis and the spectrum depends on the constraint B , we emphasize the difference between functions in $\text{dom}(A)$ and $\text{dom}(A_\star)$ by consequently using the operator A for functions satisfying $Bu = 0$.

Proof of Theorem 3.37. Ad (i): For the sake of this proof, let X^k be defined as the abstract solution to problem (3.40). We note that, using the Z -transform, we can see that $\widehat{\Gamma} := \mathcal{Z}[\Gamma]$ solves the same equation as $k[\delta(z)]^{-1} \widehat{\Xi}(z)$, where $\widehat{\Xi} := \mathcal{Z}[\Xi]$. Taking the Z -transform of problem (3.40) using Lemma 3.19, we get that \widehat{X}^k and $k[\delta(z)]^{-1} \widehat{U}^k(z)$ solve the same boundary value problem in the frequency domain. Due to the uniqueness result of Lemma 3.35, X^k and $(\partial_t^k)^{-1} U^k$ must coincide.

Ad (ii): The proof has a similar structure as in [AP03, Theorems 1 and 2]. By applying Propositions 3.33 or 3.32, we can reduce the problem to an approximation by a perturbed constraint and right-hand side. To formalize this, let \tilde{x}^k and \tilde{X}^k denote the solutions to the problem (3.40) where the constraint is replaced by $\underline{B}\tilde{X}^k(t_n) = (\partial_t^{-1}\Xi)(t_n + k\mathbf{c})$, and the right hand side by $\tilde{G} := \partial_t^{-1}F$ using the exact integrals. Then, the difference to the exact solution can be bounded by Propositions 3.32/ 3.33 as

$$\left\| x(t) - \tilde{x}^k(t_n) \right\|_{\mathcal{X}} \leq C\rho_k(T)Tk^{\min(q+\alpha+\mu, p)} \sum_{\ell=0}^{p+1} \max_{\tau \in [0, T]} \left\| x^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu}, \quad (3.53)$$

where $\alpha = 0, 1$ depending on whether $r(z)$ satisfies the additional assumption (3.37).

For ease of presentation, we assume a homogeneous right hand side $F(t) = 0$, and only deal with the perturbed constraint. The error due to the difference of G and \tilde{G} can be treated using similar techniques, but is slightly simpler since no further corrections are needed to enforce that the error functions are in the domain of A . We define the following error terms:

$$E^k(t_n) := \tilde{X}^k(t_n) - X^k(t_n) - \Delta^k(t_n), \quad (3.54)$$

$$e^k(t_n) := \tilde{x}^k(t_n) - x^k(t_n) - \delta^k(t_n), \quad (3.55)$$

where

$$\Delta^k(t_n) := \mathcal{E}_B \left(B\tilde{X}^k(t_n) - BX^k(t_n) \right), \quad (3.56)$$

$$\delta^k(t_n) := \mathcal{E}_B \left(\partial_t^{-1} \Xi(t_n) - \gamma^k(t_n) \right) \quad (3.57)$$

are corrections to ensure $\underline{B}E^k = 0$, which means $E^k \in \text{dom}(A)$, and δ^k will be needed to ensure full convergence order.

Since the Runge-Kutta method induces a quadrature formula of order p (order $p + 1$ in each step, see (3.9) with $j = 0$) and the lifting \mathcal{E}_B is a bounded operator, we observe that $\|\tilde{x}^k(t_n) - x^k(t_n)\|_{\mathcal{X}} = \|e^k(t_n) + \delta^k(t_n)\|_{\mathcal{X}} \lesssim \|e^k(t)\|_{\mathcal{X}} + \mathcal{O}(k^p)$. This means, it is sufficient to bound $e^k(t_n)$.

From the definition of the error terms, we get the following error equations:

$$E^k(t_n) = e^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A]E^k(t_n) - \Delta^k(t_n) + \delta^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A_\star] \Delta^k(t_n), \quad (3.58)$$

$$e^k(t_{n+1}) = e^k(t_n) + k[\mathbf{b}^T \otimes A] E^k(t_n) - \delta^k(t_{n+1}) + \delta^k(t_n) + k[\mathbf{b}^T \otimes A_\star] \Delta^k(t_n). \quad (3.59)$$

We relabel the different error terms in order to consider them separately later on:

$$E^k(t_n) = e^k(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A]E^k(t_n) + \Theta_I^k(t_n), \quad (3.60)$$

$$e^k(t_{n+1}) = e^k(t_n) + k[\mathbf{b}^T \otimes A] E^k(t_n) + \Theta_{II}^k(t_n), \quad (3.61)$$

with

$$\Theta_I^k(t) := -\Delta^k(t) + \delta^k(t)\mathbf{1} + k[\mathcal{Q} \otimes A_\star] \Delta^k(t),$$

$$\Theta_{II}^k(t) := -\delta^k(t+k) + \delta^k(t) + k[\mathbf{b}^T \otimes A_\star] \Delta^k(t).$$

Inserting the equation for the stage vector (3.60) into (3.61), we get:

$$\begin{aligned} e^k(t_{n+1}) &= \left(\text{Id} + k[\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1} \mathbf{1} \right) e^k(t_n) \\ &\quad + k[\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1} \Theta_I^k(t_n) + \Theta_{II}^k(t_n) \\ &= r(kA)e^k(t_n) + k[\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1} \Theta_I^k(t_n) + \Theta_{II}^k(t_n), \end{aligned}$$

where we used that the operator $[\mathbf{b}^T \otimes A] [\text{Id} - k\mathcal{Q} \otimes A]^{-1}$ is bounded and thus we can use our calculus for rational functions without worrying about domains.

By unravelling the recursion and using the initial conditions $e^k(0) = 0$, we get using the triangle inequality and the definition of $\rho_k(T)$ from (3.35):

$$\begin{aligned} & \left\| e^k(t_{n+1}) \right\|_{\mathcal{X}} \\ & \leq \left\| \sum_{j=0}^n [r(kA)]^{n-j} [\mathbf{b}^T \otimes kA] [\text{Id} - k\mathcal{Q} \otimes A]^{-1} \Theta_{II}^k(t_j) \right\|_{\mathcal{X}} + \rho_k(T) \sum_{j=0}^n \left\| \Theta_{II}^k(t_j) \right\|_{\mathcal{X}}. \end{aligned} \quad (3.62)$$

We estimate the two terms separately, focusing on the term involving Θ_{II}^k first, splitting it up further into two contributions:

$$\Theta_{II}^k(t) = \Theta_{IIa}^k(t) + \Theta_{IIb}^k(t) := \left(-\delta^k(t+k) + \delta^k(t) \right) + \left(k[\mathbf{b}^T \otimes A_\star] \Delta^k(t) \right).$$

A direct estimate of the term Θ_{IIa}^k using the triangle inequality would give order $\mathcal{O}(k^p)$ as δ^k is the Runge-Kutta approximation of an integral. Instead, we rewrite the difference as:

$$\begin{aligned} \delta^k(t_{j+1}) - \delta^k(t_j) &= \mathcal{E}_B \left([\partial_t^{-1} \Xi](t_{j+1}) - \gamma^k(t_{j+1}) - [\partial_t^{-1} \Xi](t_j) + \gamma^k(t_j) \right) \\ &= \mathcal{E}_B \left([\partial_t^{-1} \Xi](t_{j+1}) - [\partial_t^{-1} \Xi](t_j) + k\mathbf{b}^T \Xi(t_j + k\mathbf{c}) \right), \end{aligned}$$

where in the last step we used the defining equation for $\gamma^k(t_{j+1})$. Since \mathbf{b} and \mathbf{c} induce a quadrature formula of order $p+1$, this term is also of order $p+1$.

For the term Θ_{IIb}^k , we first observe that, since $A_\star \mathcal{E}_B = \mathcal{E}_B$, we can rewrite it as

$$\Theta_{IIb}^k(t) = k[\mathbf{b}^T \otimes \text{Id}] \Delta^k(t),$$

or inserting the definition of $\Delta^k(t)$:

$$\begin{aligned} k[\mathbf{b}^T \otimes \text{Id}] \Delta^k(t_j) &= k\mathcal{E}_B[\mathbf{b}^T \otimes \text{Id}] \left([\partial_t^{-1} \Xi](t_j + k\mathbf{c}) - [\partial_t^{-1} \Xi](t_j) \mathbf{1} - k\mathcal{Q} \Xi(t_j + k\mathbf{c}) \right) \\ &\quad - k\mathcal{E}_B[\mathbf{b}^T \otimes \text{Id}] \left(\gamma^k(t_j) \mathbf{1} - [\partial_t^{-1} \Xi](t_j) \mathbf{1} \right). \end{aligned}$$

The last term is of order $\mathcal{O}(k^{p+1})$ by the approximation properties of γ^k . Next, we write $\partial_t^{-1} \Xi(t_j + t) =: \pi(t) + \varphi(t)$ as its Taylor polynomial of degree $p-1$ based in t_j with remainder term φ . This allows us to write:

$$\begin{aligned} & k[\mathbf{b}^T \otimes \text{Id}] \left(\partial_t^{-1} \Xi(t_j + k\mathbf{c}) - \partial_t^{-1} \Xi(t_j) \mathbf{1} - k\mathcal{Q} \Xi(t_j + k\mathbf{c}) \right) \\ &= k[\mathbf{b}^T \otimes \text{Id}] \left(\pi(k\mathbf{c}) + \varphi(k\mathbf{c}) - \pi(0) \mathbf{1} - k\mathcal{Q} \dot{\pi}(\mathbf{c}) - k\mathcal{Q} \dot{\varphi}(k\mathbf{c}) \right) \\ &= k[\mathbf{b}^T \otimes \text{Id}] \left(\varphi(k\mathbf{c}) - k\mathcal{Q} \dot{\varphi}(k\mathbf{c}) \right), \end{aligned}$$

where we used that $[\mathbf{b}^T \otimes \text{Id}] (\mathbf{c}^\ell - \ell \mathcal{Q} \mathbf{c}^{\ell-1})$ vanishes for $\ell \leq p-1$ due to the order conditions (3.9). The functions φ and $k\dot{\varphi}$ are of order $\mathcal{O}(k^p)$ and their norm can be controlled using $\Xi^{(p)}$. We conclude that $k[\mathbf{b}^T \otimes \text{Id}] \Delta^k(t_n)$ is of order $\mathcal{O}(k^{p+1})$. Overall, the previous calculations show that the error terms involving Θ_{II}^k all have order $\mathcal{O}(k^{p+1})$ and therefore do not impact the convergence rate, even when summed over all $j = 0, \dots, n$.

Next, we look the error terms involving Θ_I^k . Again, we split it into two contributions:

$$\Theta_I^k(t) =: \Theta_{Ia}^k(t) + \Theta_{Ib}^k(t) := [-\Delta^k(t) + \delta^k(t)\mathbf{1}] + k[\mathcal{Q} \otimes A_\star] \Delta^k(t).$$

Starting with Θ_{Ia}^k , and inserting the definition of $\Delta^k(t_n)$ and $\delta^k(t_n)$, we get

$$\begin{aligned} \Delta^k(t_n) - \delta^k(t_n)\mathbf{1} &= \mathcal{E}_B \left(\partial_t^{-1} \Xi(t_n + k\mathbf{c}) - \gamma^k(t_n)\mathbf{1} - k\mathcal{Q}\Xi(t_n + k\mathbf{c}) - \partial_t^{-1} \Xi(t_n)\mathbf{1} + \gamma^k(t_n)\mathbf{1} \right) \\ &= \mathcal{E}_B \left[\partial_t^{-1} \Xi(t_n + k\mathbf{c}) - (\partial_t^{-1} \Xi(t_n)\mathbf{1} + k\mathcal{Q}\Xi(t_n + k\mathbf{c})) \right]. \end{aligned}$$

This means that the term involving Θ_{Ia}^k in (3.62) structurally fits the setting of Lemma 3.42 for $u := \mathcal{E}_B [\partial_t^{-1} \Xi]$ and $\beta = 0$. Depending on whether $r(z)$ fulfills the additional assumption (3.37), this implies convergence rates $\mathcal{O}(k^{\min(q+\mu, p)})$ or $\mathcal{O}(k^{\min(q+1+\mu, p)})$.

For the term Θ_{Ib}^k , we note that due to the lifting property $A_\star \mathcal{E}_B = \mathcal{E}_B$ we can drop A_\star and get $k[\mathcal{Q} \otimes A_\star] \Delta^k(t_n) = k[\mathcal{Q} \otimes \text{Id}] \Delta^k(t_n)$. Using the definition of Δ^k , and inserting a term of the form $\partial_t^{-1} \Xi - \gamma^k$, we calculate

$$\begin{aligned} k[\mathcal{Q} \otimes \text{Id}] \Delta^k(t_n) &= k\mathcal{Q}\mathcal{E}_B \left(\partial_t^{-1} \Xi(t_n + k\mathbf{c}) - \gamma^k(t_n)\mathbf{1} - k\mathcal{Q}\Xi(t_n + k\mathbf{c}) \right) \\ &= k\mathcal{Q}\mathcal{E}_B \left(\partial_t^{-1} \Xi(t_n + k\mathbf{c}) - \partial_t^{-1} \Xi(t_n)\mathbf{1} - k\mathcal{Q}\Xi(t_n + k\mathbf{c}) \right) \\ &\quad + k\mathcal{Q}\mathcal{E}_B \left(\partial_t^{-1} \Xi(t_n) - \gamma^k(t_n) \right) \mathbf{1}. \end{aligned}$$

The last term is of order $\mathcal{O}(k^{p+1})$ as $\gamma^k(t_n)$ approximates the integral with order p . The first term is again similar to a Runge-Kutta consistency error term, and fits Lemma 3.42 with $\beta = 1$ (due to the additional power of $k\mathcal{Q}$), which implies convergence of order $\mathcal{O}(k^{\min(q+1+\mu+\alpha, p)})$ ($\alpha = 0$ or 1 depending on whether (3.37) holds). This concludes the proof. \square

Proof of Theorem 3.38. The proof is similar to the one of Theorem 3.37. Again, we reduce the problem to the case of perturbed data by using Propositions 3.32 or 3.33. Namely, we have the following estimate:

$$\left\| v(t) - \tilde{v}^k(t_n) \right\|_{\mathcal{X}} \leq CT\rho_k(T)k^{\min(q+\alpha+\mu, p)} \sum_{\ell=0}^{p+1} \max_{\tau \in [0, T]} \left\| v^{(\ell)}(\tau) \right\|_{\mathcal{X}_\mu},$$

where \tilde{v}^k together with the stages \tilde{V}^k solves (3.39) with the “exact” constraint $\underline{B}\tilde{V}^k(t) = \dot{\Xi}(t + k\mathbf{c})$ instead of the difference quotient and a right-hand side $\tilde{G}(t) := \dot{F}(t + k\mathbf{c})$.

What remains to estimate is the influence of the error term in the constraint and right-hand side. Again, we drop the right-hand side term as it can be dealt with using the same techniques, and focus on the boundary terms. We define

$$\begin{aligned} \Delta^k(t) &:= \mathcal{E}_B \left[k^{-1} \mathcal{Q}^{-1} (\Xi(t + k\mathbf{c}) - \Xi(t)\mathbf{1}) - \dot{\Xi}(t + k\mathbf{c}) \right] \\ &= k^{-1} \mathcal{Q}^{-1} \left[\Xi(t + k\mathbf{c}) - \Xi(t)\mathbf{1} - k\mathcal{Q}\dot{\Xi}(t + k\mathbf{c}) \right], \\ E^k(t) &:= V^k(t) - \tilde{V}^k(t) - \Delta^k(t), \\ e^k(t) &:= v^k(t) - \tilde{v}^k(t). \end{aligned}$$

We note that, by the lifting property, we have $\underline{B}E^k(t) = 0$ and thus $E^k(t) \in [\text{dom}(A)]^m$. The stages E^k satisfy the following equation:

$$E^k(t_n) = e^k(t_n)\mathbb{1} + k[\mathcal{Q} \otimes A]X^k(t_n) - \Delta^k(t_n) + k[\mathcal{Q} \otimes \text{Id}]\Delta^k(t_n),$$

where we used the lifting property $A_\star \mathcal{E}_B = \mathcal{E}_B$. Eliminating E^k and using the rational function calculus $r(kA)$, gives for the end point:

$$e^k(t_{n+1}) = r(kA)e^k(t_n) + k[\mathbf{b}^T \otimes A][\text{Id} - k\mathcal{Q} \otimes A]^{-1}(k\mathcal{Q} - \text{Id})\Delta^k(t_n).$$

We recursively plug in the representation of the truncation error, reorganize some terms, and use the homogeneous initial conditions to get the representation:

$$\begin{aligned} e^k(t_n) &= \sum_{j=0}^n [r(kA)]^{n-j} g(kA) [k\mathcal{Q} - \text{Id}] \Delta^k(t_n) \\ &= \sum_{j=0}^n r(kA)^{n-j} g(kA) k\mathcal{Q} \Delta^k(t_j) - \sum_{j=0}^n [r(kA)]^{n-j} g(kA) \Delta^k(t_j) \\ &=: \Delta_I(t_n) + \Delta_{II}(t_n), \end{aligned}$$

with $g(z) := z\mathbf{b}^T(\text{Id} - z\mathcal{Q})^{-1}$.

Both terms structurally fit the assumptions of Lemma 3.42: For Δ_I we use $\beta = 0$, which gives the rate $\mathcal{O}(k^{\min(q+\alpha+\mu, p)})$; For Δ_{II} , we use $\beta = -1$ which gives convergence rate $\mathcal{O}(k^{\min(q+\mu-1, p-1)+\alpha})$ (note the extra term $k^{-1}\mathcal{Q}^{-1}$ hidden in the definition of Δ^k). Thus, we conclude that overall we get the convergence rate $\mathcal{O}(k^{\min(q+\mu, p)})$ if $\alpha = 1$, i.e., the Runge-Kutta method satisfies (3.37), and $\mathcal{O}(k^{\min(q+\mu-1, p-1)})$ otherwise. This concludes the proof. \square

4 FEM-BEM coupling for the Schrödinger equation

In this section, we apply the boundary integral equation techniques to the Schrödinger equation. The Schrödinger equation is one of the main governing equations in quantum mechanics, and as such has a myriad of applications in the sciences. In its usual form, it is posed on the full space \mathbb{R}^d , which makes pure finite elements impractical. By making the assumption that the complicated dynamics of the system are localized in a bounded domain Ω^- , we are able to replace the approximation on the unbounded domain \mathbb{R}^d with a discretization on Ω^- augmented by transparent boundary conditions using the tools of boundary integral equations and convolution quadrature developed in the previous sections.

This discretization scheme leads to a coupling of finite element and boundary element methods (FEM-BEM coupling). Two classical procedures for such a coupling are the *symmetric coupling* introduced in [Cos88a] and [Han90], as well as the *Johnson-Nédélec coupling* [JN80]. We will focus on the symmetric approach.

A first numerical study of using Convolution Quadrature for the Schrödinger equation in 2D is given in [Sch02]. There, the author uses a Johnson-Nédélec based FEM-BEM coupling and convolution quadrature based on the trapezoidal rule for discretization and observes optimal convergence rates in time.

While boundary integral equations provide a convenient way of implementing transparent boundary conditions, it is in no way the only or even the most widespread approach. Other ways of representing these boundary conditions can be found under the names “Perfectly matched Layer” (PML) or “infinite elements”. A recent survey of the different approaches that may be taken for transparent boundary conditions for the Schrödinger equation is [Ant+08].

Most of the results in this section have already appeared as part of [MR17].

4.1 Model problem and notation

We consider the time-dependent Schrödinger equation in \mathbb{R}^d for $d = 2$ or $d = 3$. For a potential $\mathcal{V} : \mathbb{R}^d \rightarrow \mathbb{R}$, we seek u such that

$$\begin{aligned} i \frac{\partial}{\partial t} u &= -\Delta u + \mathcal{V} u \quad \text{in } \mathbb{R}^d, \\ u(0) &= u_0, \end{aligned}$$

where u_0 is a given initial condition.

By introducing the Hamilton operator

$$\begin{aligned} \mathbf{H} &: \text{dom}(\mathbf{H}) \subseteq L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d) \\ \mathbf{H}u &:= -\Delta u + \mathcal{V}(\cdot)u, \end{aligned}$$

this equation can be written more succinctly as

$$i\dot{u} = \mathbf{H}u \quad \text{and} \quad u(0) = u_0. \quad (4.1)$$

In order to be able to apply our discretization scheme, we make the following assumptions:

Assumption 4.1. *We consider the 2d and 3d case, i.e., $d = 2$ or 3 .*

- (i) $\mathcal{V} : \mathbb{R}^d \rightarrow \mathbb{R}$ is real valued,
- (ii) $\mathcal{V}(x) \equiv \mathcal{V}_0$ for $x \in \Omega^+ := \mathbb{R}^d \setminus \overline{\Omega^-}$, where Ω^- is a bounded Lipschitz domain,
- (iii) $u_0 \in \text{dom}(\mathbf{H})$ and $\text{supp } u_0 \subseteq \Omega^-$,
- (iv) the potential can be written as $\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2$ with $\mathcal{V}_1 \in L^\infty(\mathbb{R}^d)$ and $\mathcal{V}_2 \in L^2(\mathbb{R}^d)$,
- (v) \mathcal{V} is bounded from below, i.e. $\mathcal{V} \geq \mathcal{V}_-$ for a constant $\mathcal{V}_- \in \mathbb{R}$.

Remark 4.2. *We note that assumption (iv) is somewhat natural to make as it leads to a self-adjoint Hamilton operator. In the case $d = 3$, this assumption is, for example, satisfied by potentials of Coulomb type which behave like $|x|^{-1}$ for $|x| \rightarrow 0$. See also [RS75, Section X] for similar assumptions.*

In order to discretize this problem, we first perform a discretization in time by an A-stable Runge-Kutta method (see Section 3.2), derive an equivalent formulation using boundary integral equations, and then discretize those using a Galerkin method. We derive the equations in a mostly formal way, the construction will then later be made rigorous as part of Lemma 4.8.

4.2 The discretization scheme

The semi-discretization of (4.1) is given by

$$U^k(t_n) = u^k(t_n)\mathbb{1} - ik[\mathcal{Q} \otimes \mathbf{H}]U^k(t_n), \quad (4.2)$$

$$u^k(t_{n+1}) = u^k(t_n) - ik[\mathbf{b}^T \otimes \mathbf{H}]U^k(t_n) \quad (4.3)$$

and the initial value $u^k(0) := u_0$.

By Lemma 3.19, the Z -transform of the stage vectors $\widehat{U}^k := \mathcal{Z}[U^k]$ solves:

$$-\frac{\delta(z)}{k}\widehat{U}^k - i\mathbf{H}\widehat{U}^k = \frac{1}{1 - r(\infty)z}k^{-1}\mathcal{Q}^{-1}\mathbb{1}u_0.$$

Restricting this equation to $\mathbb{R}^d \setminus \overline{\Omega^-}$ and applying the assumption $\text{supp } u_0 \subseteq \Omega^-$ gives:

$$\frac{-i\delta(z)}{k} \widehat{U}^k - \underline{\Delta} \widehat{U}^k + \mathcal{V}_0 \widehat{U}^k = 0, \quad \text{in } \mathbb{R}^d \setminus \overline{\Omega^-}, \quad (4.4)$$

which corresponds to a Helmholtz equation with the matrix-valued wave number $B(z) := \sqrt{\frac{-i\delta(z)}{k} + \mathcal{V}_0} \text{Id}$. Here, we consider the branch of the square root satisfying $\text{Re}(z) \geq 0$ and use the Riesz-Dunford calculus (Definition 2.40) for the matrix square root. As the solution to a Helmholtz-type problem, we can now derive the boundary integral equations, which will serve as transparent boundary conditions. The Z -transform $\widehat{U}^k := \mathcal{Z}[U^k]$ satisfies the following boundary integral equations via Proposition 2.43 (we use (2.39) and the fact that $\gamma^+ \widehat{U}^k = \gamma^- \widehat{U}^k$ and $\partial_\nu^+ \widehat{U}^k = \partial_\nu^- \widehat{U}^k$):

$$\begin{pmatrix} \frac{1}{2} - K(B(z)) & V(B(z)) \\ W(B(z)) & -\frac{1}{2} + K^T(B(z)) \end{pmatrix} \begin{pmatrix} \gamma^- \widehat{U}^k \\ \partial_\nu^- \widehat{U}^k \end{pmatrix} = \begin{pmatrix} 0 \\ -\partial_\nu^- \widehat{U}^k \end{pmatrix}. \quad (4.5)$$

In order to simplify the notation, we introduce the new boundary integral operators

$$\begin{aligned} \widetilde{V}(z) &:= V\left(\sqrt{-iz + \mathcal{V}_0}\right), & \widetilde{K}(z) &:= K\left(\sqrt{-iz + \mathcal{V}_0}\right), \\ \widetilde{K}^T(z) &:= K^T\left(\sqrt{-iz + \mathcal{V}_0}\right), & \widetilde{W}(z) &:= W\left(\sqrt{-iz + \mathcal{V}_0}\right) \end{aligned}$$

and the modified potentials:

$$\widetilde{S}(z) := S\left(\sqrt{-iz + \mathcal{V}_0}\right), \quad \widetilde{D}(z) = D\left(\sqrt{-iz + \mathcal{V}_0}\right).$$

By using the inverse Z -transform and the discrete operational calculus, we can derive the following set of semi-discrete equations for U^k , introducing the new unknown $\Lambda^k := \partial_\nu^+ U^k$:

$$U^k(t_n) = u^k(t_n) \mathbf{1} - ik[\mathcal{Q} \otimes \mathbf{H}]U^k(t_n) \quad \text{in } \Omega^-, \quad (4.6a)$$

$$u^k(t_{n+1}) = r(\infty)u^k(t_n) + \mathbf{b}^T \mathcal{Q}^{-1}U^k(t_n), \quad (4.6b)$$

$$\begin{pmatrix} 0 \\ -\partial_\nu^- U^k \end{pmatrix} = \begin{pmatrix} \frac{1}{2} - \widetilde{K}(\partial_t^k) & \widetilde{V}(\partial_t^k) \\ \widetilde{W}(\partial_t^k) & -\frac{1}{2} + \widetilde{K}^T(\partial_t^k) \end{pmatrix} \begin{pmatrix} \gamma^- U^k \\ \Lambda^k \end{pmatrix}. \quad (4.6c)$$

For discretization, let $V_h \subseteq H^1(\Omega^-)$ and $X_h \subseteq H^{-1/2}(\Gamma)$ be closed (not necessarily finite dimensional) subspaces. In order to deal with unbounded potentials, we have to make an additional assumption:

Assumption 4.3. *Let one of the following assumptions hold:*

(i) V_h contains test functions, i.e., $C_0^\infty(\Omega^-) \subseteq V_h$, or

(ii) V_h is finite dimensional such that $V_h \subseteq L^\infty(\Omega^-)$ and admits an inverse estimate $\|\varphi\|_{L^\infty(\Omega^-)} \leq C(V_h) \|\varphi\|_{L^2(\Omega^-)} \forall \varphi \in V_h$. The constant may depend on anything except on φ . ■

Using the sesquilinear form

$$\mathcal{A}(U, W) := (U, W)_{L^2(\Omega^-)} + ik([\mathcal{Q} \otimes \nabla]U, \nabla W)_{L^2(\Omega^-)} + ik([\mathcal{Q} \otimes \mathcal{V}]U, W)_{L^2(\Omega^-)},$$

the discretized form of (4.6a) is then given by: Find $U^{k,h} \in [V_h]^m$, $\Lambda^{k,h} \in [X_h]^m$, such that for discrete times $t = t_n$:

$$\begin{aligned} \mathcal{A}(U^{k,h}, W_h) + ik \left\langle \mathcal{Q} \left[\widetilde{W}(\partial_t^k) \gamma^- U^{k,h} - [1/2 - \widetilde{K}^T(\partial_t^k)] \Lambda^{k,h} \right], \gamma^- W_h \right\rangle_{\Gamma} \\ = \left(u^{k,h} \mathbf{1}, W_h \right)_{L^2(\Omega^-)}, \end{aligned} \quad (4.7a)$$

and the boundary values solve

$$\left\langle \widetilde{V}(\partial_t^k) \Lambda^{k,h}, Q_h \right\rangle_{\Gamma} + \left\langle \left(1/2 - \widetilde{K}(\partial_t^k) \right) \gamma^- U^{k,h}, Q_h \right\rangle_{\Gamma} = 0 \quad (4.7b)$$

for all $W_h \in [V_h]^m$, $Q_h \in [X_h]^m$. The approximation at t_{n+1} is computed by

$$u^{k,h}(t_{n+1}) = u^{k,h}(t_n) + \mathbf{b}^T \mathcal{Q}^{-1} U^k(t_n),$$

and we assume we are given an approximation of the initial condition $u^{k,h}(0) \in V_h$.

4.3 An equivalent formulation as an exotic semigroup

The system (4.7) is not well suited for direct analysis, as it involves the non-local in time operators $\widetilde{V}(\partial_t^k)$, $\widetilde{K}(\partial_t^k)$, etc. Instead, we analyze a different system based on the exotic Hilbert spaces from [LS09]. A similar construction has recently been presented in [HS16] in the context of coupling FEM and BEM for the wave equation but focus on a trapezoidal rule time discretization.

Since we will be dealing a lot with pairs of spaces, we introduce $\mathcal{X}^0 := L^2(\Omega^-) \times L^2(\mathbb{R}^d \setminus \Gamma)$ and $\mathcal{X}^1 := H^1(\Omega^-) \times H^1(\mathbb{R}^d \setminus \Gamma)$, both equipped with the natural inner product. In order to make use of the abstract setting of Section 3.4, we introduce a new operator \mathbf{H}_h . It is defined in such a way that $A := -i\mathbf{H}_h$ generates a semigroup and the Runge-Kutta approximation of this semigroup coincides with solving (4.7).

Definition 4.4. *We introduce the spaces*

$$\mathcal{X}_h^1 := \left\{ (u_h, u_\star) \in V_h \times H^1(\mathbb{R}^d \setminus \Gamma), \gamma^- u_h = \llbracket \gamma u_\star \rrbracket \wedge \gamma^- u_\star \in X_h^0 \right\} \quad (4.8)$$

with the \mathcal{X}^1 norm and inner-product, as well as $\mathcal{X}_h^0 := \text{clos}(\mathcal{X}_h^1, \|\cdot\|_{\mathcal{X}^0})$ with the \mathcal{X}^0 product.

Let the operator $\mathbf{H}_h : \text{dom}(\mathbf{H}_h) \subseteq \mathcal{X}_h^0 \rightarrow \mathcal{X}_h^0$ be defined as $\mathbf{H}_h : (u_h, u_\star) \mapsto (x_h, x_\star)$, where $(x_h, x_\star) \in \mathcal{X}_h^0$ is such that for all $(w_h, w_\star) \in \mathcal{X}_h^1$

$$\begin{aligned} \langle x_h, w_h \rangle_{L^2(\Omega)} + \langle x_\star, w_\star \rangle_{L^2(\mathbb{R}^d)} \\ = \langle \nabla u_h, \nabla w_h \rangle_{L^2(\Omega)} + \langle \nabla u_\star, \nabla w_\star \rangle_{L^2(\mathbb{R}^d)} + \langle \mathcal{V} u_h, w_h \rangle_{L^2(\Omega)} + \langle \mathcal{V}_0 u_\star, w_\star \rangle_{L^2(\mathbb{R}^d)}. \end{aligned} \quad (4.9)$$

The domain is defined as

$$\text{dom}(\mathbf{H}_h) := \left\{ (u_h, u_\star) \in \mathcal{X}_h^1 : \exists (x_h, x_\star) \in \mathcal{X}_h^0, \text{ s.t. (4.9) holds} \right\}.$$

We also introduce the notation $\mathcal{V}(u, u_\star) := (\mathcal{V}u, \mathcal{V}_0 u_\star)$ for elements of \mathcal{X}^0 .

We note that \mathbf{H}_h is well defined since for two solutions $(x_h^1, x_\star^1), (x_h^2, x_\star^2)$ satisfying (4.9), the difference solves $0 = \langle x_h^1 - x_h^2, w_h \rangle_{\mathcal{X}^0} + \langle x_\star^1 - x_\star^2, w_\star \rangle_{\mathcal{X}^0}$. Since the inner product vanishes on a dense subset of \mathcal{X}_h^0 , we get $(x_h^1, x_\star^1) = (x_h^2, x_\star^2)$.

Remark 4.5. *The use of the space \mathcal{X}_h^0 is needed to treat the case $V_h := H^1(\Omega^-)$ and “ V_h is finite dimensional” in the same setting. In the former case, we get $\mathcal{X}_h^0 = \mathcal{X}^0$ since $C_0^\infty(\Omega^-) \times C_0^\infty(\mathbb{R}^d \setminus \Gamma)$ is dense. If V_h is finite dimensional, it becomes $\mathcal{X}_h^0 = V_h \times L^2(\mathbb{R}^d \setminus \Gamma)$ since V_h is closed and the restrictions on the traces are not seen by the L^2 -based norm. This also means that \mathbf{H}_h corresponds to either the classical weak Laplacian or a version of the discrete Laplacian defined using the usual stiffness and mass matrices.*

Using this definition, we can show the following consequence of Assumption 4.3.

Lemma 4.6. *For every $a > 0$ there exists a constant $b > 0$, such that*

$$\|\mathcal{V}u\|_{\mathcal{X}^0} \leq a \left\| \tilde{\mathbf{H}}_h u \right\|_{\mathcal{X}^0} + b \|u\|_{\mathcal{X}^0},$$

where $\tilde{\mathbf{H}}_h$ denotes the Hamilton operator corresponding to $\mathcal{V} = 0$, i.e., $\mathbf{H}_h - \mathcal{V}$.

Additionally, the bilinear form induced by \mathcal{V} is bounded in the following sense:

$$|\langle \mathcal{V}u, v \rangle_{\mathcal{X}^0}| \leq C \|u\|_{\mathcal{X}^1} \|v\|_{\mathcal{X}^1}. \quad (4.10)$$

Proof. If Assumption 4.3(i) holds, this is well known as a consequence of [RS75, Theorem X.15] and Assumption 4.1(iv) as $\tilde{\mathbf{H}}_h$ coincides with the classical weak definition of the Laplacian in each component.

If Assumption 4.3(ii) holds, we can use Assumption 4.1(iv) to estimate:

$$\|\mathcal{V}u\|_{\mathcal{X}^0} \leq \|\mathcal{V}_1 u\|_{\mathcal{X}^0} + \|\mathcal{V}_2 u\|_{\mathcal{X}^0} \leq \|\mathcal{V}_1\|_{L^\infty(\mathbb{R}^d)} \|u\|_{\mathcal{X}^0} + \|\mathcal{V}_2\|_{L^2(\mathbb{R}^d)} \|u\|_{L^\infty(\Omega^-)},$$

where we used that we may choose $\mathcal{V}_2 = 0$ on $\mathbb{R}^d \setminus \overline{\Omega^-}$ by absorbing it into \mathcal{V}_1 . Since we assumed the inverse estimate of Assumption 4.3(ii), this gives the estimate

$$\|\mathcal{V}u\|_{\mathcal{X}^0} \lesssim C(X_h) \|u\|_{\mathcal{X}^0},$$

i.e., we may even use $a = 0$.

To see (4.10), we estimate for the interior, unbounded contribution:

$$\begin{aligned} \left| \int_{\Omega^-} \mathcal{V}_2 u v \right| &\leq \left\| |\mathcal{V}_2|^{1/2} u \right\|_{L^2(\Omega^-)} \left\| |\mathcal{V}_2|^{1/2} v \right\|_{L^2(\Omega^-)} \leq \|\mathcal{V}_2\|_{L^2(\Omega^-)}^{1/2} \|u\|_{L^4(\Omega^-)} \|\mathcal{V}_2\|_{L^2(\Omega^-)}^{1/2} \|v\|_{L^4(\Omega^-)} \\ &\leq \|\mathcal{V}_2\|_{L^2(\Omega^-)} \|u\|_{H^1(\Omega^-)} \|v\|_{H^1(\Omega^-)}, \end{aligned}$$

where in the last step we used the Sobolev embedding $H^1(\Omega^-) \hookrightarrow L^4(\Omega^-)$ (for $d \leq 3$) from Proposition 2.24. The contributions of \mathcal{V}_1 and the exterior part of \mathcal{X}^0 can easily be bounded by the L^2 -norm as the corresponding potential is bounded. \square

Lemma 4.7. *The operator \mathbf{H}_h is self-adjoint on \mathcal{X}_h^0 and $-i\mathbf{H}_h$ generates a unitary C_0 -group.*

Proof. We start with the case $\mathcal{V} = 0$, i.e., we look at the operator $\tilde{\mathbf{H}}_h := \mathbf{H}_h - \mathcal{V}$. We use the characterization of Proposition 2.9 and set $A := -i\tilde{\mathbf{H}}_h$. To see that $\pm A$ is dissipative, we calculate for $u \in \text{dom}(A)$:

$$\text{Re} \langle \pm Au, u \rangle_{\mathcal{X}^0} = \pm \text{Im} \left\langle \tilde{\mathbf{H}}_h u, u \right\rangle_{\mathcal{X}^0} = \pm \text{Im} \langle \nabla u, \nabla u \rangle_{\mathcal{X}^0} = 0.$$

We now show that $\pm i\tilde{\mathbf{H}}_h - \text{Id}$ is invertible on \mathcal{X}_h^0 . Consider $f = (f_0, f_\star) \in \mathcal{X}_h^0$. The sesquilinear form

$$a(u, v) := \langle \pm i \nabla u, \nabla v \rangle_{\mathcal{X}^0} - \langle u, v \rangle_{\mathcal{X}^0}$$

is bounded and coercive on \mathcal{X}_h^1 . By the Lax-Milgram lemma, we can find $u = (u_h, u_\star) \in \mathcal{X}_h^1$, such that $a(u, v) = \langle f, v \rangle_{\mathcal{X}^0}$. We need to show that $u \in \text{dom}(A)$ and $\mp Au - u = f$. Defining $x := u + f \in \mathcal{X}_h^0$, we see that for $v \in \mathcal{X}_h^1$:

$$\begin{aligned} \langle x, v \rangle_{\mathcal{X}^0} &= \langle u, v \rangle_{\mathcal{X}^0} + \langle f, v \rangle_{\mathcal{X}^0} \\ &= \langle u, v \rangle_{\mathcal{X}^0} + a(u, v) \\ &= \langle u, v \rangle_{\mathcal{X}^0} + \langle \pm i \nabla u, \nabla v \rangle_{\mathcal{X}^0} - \langle u, v \rangle_{\mathcal{X}^0} \\ &= \langle \pm i \nabla u, \nabla v \rangle_{\mathcal{X}^0}. \end{aligned}$$

This means x satisfies the defining equation for $\pm i\tilde{\mathbf{H}}_h u$ (see (4.9)), and therefore $u \in \text{dom}(A) = \text{dom}(\tilde{\mathbf{H}}_h)$ and $\pm i\tilde{\mathbf{H}}_h u = x$.

To see that $\pm i\tilde{\mathbf{H}}_h$ generates a C_0 -semigroup, we need to show that $\text{dom}(\tilde{\mathbf{H}}_h)$ is dense in \mathcal{X}_h^0 . In the case of Assumption 4.3(i), we get that $C_0^\infty(\Omega^-) \times C_0^\infty(\mathbb{R}^d \setminus \Gamma) \subseteq \text{dom}(\tilde{\mathbf{H}}_h)$. This set is dense in $\mathcal{X}_h^0 = \mathcal{X}^0$. If V_h is finite dimensional, we get that functions of the form (v_h, v_\star) with $v_h \in V_h$, $v_\star|_{\Omega^-} \in C_0^\infty(\Omega^-)$, $v_\star|_{\Omega^+} \in C_0^\infty(\Omega^+)$ with $\gamma^+ v_\star = \gamma^- v_h$ are both, contained in $\text{dom}(\tilde{\mathbf{H}}_h)$ and dense in \mathcal{X}_h^0 (note that $\gamma^- v_\star = 0$ and $[[\gamma v_\star]] = \gamma^- v_h$ by construction). This concludes the proof without potentials, i.e., for $\tilde{\mathbf{H}}_h$.

In order to deal with the potential, we use [RS75, p. X.12], which states that if the estimate from Lemma 4.6 is satisfied and $\tilde{\mathbf{H}}_h$ is self adjoint, then $\mathbf{H}_h = \tilde{\mathbf{H}}_h + \mathcal{V}$ is also self-adjoint. The fact that A generates a unitary C_0 group is Stone's theorem (Proposition 2.9). \square

The next lemma forms the basis of our future analysis. It allows us to replace the boundary integrals from (4.7) with the Runge-Kutta time-stepping of a non-standard semigroup.

Lemma 4.8. *If $(U^{k,h}, \Lambda^{k,h})$ solves (4.7) and we define $U_\star^{k,h} := -\tilde{S}(\partial_t^k) \Lambda^{k,h} + \tilde{D}(\partial_t^k) \gamma^- U^{k,h}$, then $X^{k,h} := (U^{k,h}, U_\star^{k,h})$ solves*

$$X^{k,h}(t_n) = x^{k,h}(t_n) \mathbf{1} - ik[\mathcal{Q} \otimes \mathbf{H}_h] X^{k,h}(t_n), \quad (4.11a)$$

$$x^{k,h}(t_{n+1}) = x^{k,h}(t_n) + \mathbf{b}^T \mathcal{Q}^{-1} X^{k,h}(t_n). \quad (4.11b)$$

Conversely, if $X^{k,h} =: (U^{k,h}, U_\star^{k,h})$, $x^{k,h} =: (u^{k,h}, u_\star^{k,h})$ solves (4.11), then $(U^{k,h}, \Lambda^{k,h})$ with $\Lambda^{k,h} := \left[\partial_\nu U_\star^{k,h} \right]$ solves (4.7).

Before we can show the equivalence, we need the following small lemma, which we prove separately:

Lemma 4.9. *Let $\Psi \in L^\infty\left([0, T], [H^{1/2}(\Gamma)]^m\right)$ and $\Phi \in L^\infty\left([0, T], [H^{-1/2}(\Gamma)]^m\right)$ be given. Define the functions $U_\star := -\tilde{S}(\partial_t^k)\Phi + \tilde{D}(\partial_t^k)\Psi$ and $u_\star := \mathbb{G}[U_\star]$. Then these functions have the following properties:*

(i) $u_\star(t) = 0$ for $t \leq 0$,

(ii) $[\gamma U_\star] = \Psi$ and $[\partial_\nu U_\star] = \Phi$,

(iii) for $t > 0$, $U_\star(t) \in H_\Delta^1(\mathbb{R}^d \setminus \Gamma)$ solves:

$$U_\star(t) - ik\mathcal{Q}\Delta U_\star(t) + ik\mathcal{Q}\mathcal{V}_0 U_\star(t) = u_\star(t)\mathbf{1}, \quad (4.12)$$

(iv) the boundary traces of U_\star solve the following boundary integral equations:

$$\begin{pmatrix} -\partial_\nu^+ U_\star \\ -\gamma^- U_\star \end{pmatrix} = \begin{pmatrix} \tilde{W}(\partial_t^k) & -\frac{1}{2} + \tilde{K}^T(\partial_t^k) \\ \frac{1}{2} - \tilde{K} & \tilde{V}(\partial_t^k) \end{pmatrix} \begin{pmatrix} \Psi \\ \Phi \end{pmatrix}. \quad (4.13)$$

Proof. We will define a function X_\star such that (4.12) is satisfied, and then show that X_\star coincides to U_\star , as defined via the representation formula. To do so, we define a function $x_\star(t) := 0$ for $t \leq 0$, and then inductively define X_\star such that (4.12) is satisfied and $[\gamma X_\star(t)] = \Psi(t)$ as well as $[\partial_\nu X_\star(t)] = \Phi(t)$. By Proposition 3.22 this problem has a unique solution. Then, defining $x_\star(t+k) := r(\infty)x_\star(t) + \mathbf{b}^T \mathcal{Q}^{-1} X_\star(t)$ allows us to define $X_\star(t+k)$ etc. We also have the (crude) a priori estimates

$$\|X_\star(t)\|_{[H_\Delta^1(\mathbb{R}^d \setminus \Gamma)]^m} \leq C(k) \left(\|x_\star(t)\|_{H_\Delta^1(\mathbb{R}^d \setminus \Gamma)} + \|\Phi(t)\|_{[H^{-1/2}(\Gamma)]^m} + \|\Psi(t)\|_{[H^{1/2}(\Gamma)]^m} \right).$$

As such, the Z-transform $\hat{X}_\star := \mathcal{Z}[X_\star]$ exists, and solves by Lemma 3.19

$$-\Delta \hat{X}_\star + \left(-\frac{i\delta(\cdot)}{k} + \mathcal{V}_0 \text{Id} \right) \hat{X}_\star = 0.$$

The traces satisfy $[\gamma \hat{X}_\star] = \mathcal{Z}[\Psi]$ and $[\partial_\nu \hat{X}_\star] = \mathcal{Z}[\Phi]$. By the properties of the potentials in the Laplace domain $\mathcal{Z}[U_\star]$ solves the same transmission problem. The uniqueness result from Proposition 2.41 gives that $\hat{X}_\star = \mathcal{Z}[U_\star]$ and therefore $X_\star = U_\star$. The rest of the Lemma follows immediately by definition of the CQ-operators and the integral equations in the Laplace domain (Proposition 2.43). \square

We are now in a position to prove Lemma 4.8.

Proof of Lemma 4.8. Let $(U^{k,h}, \Lambda^{k,h})$ solve (4.7) and define $U_\star^{k,h} := -\tilde{S}(\partial_t^k)\Lambda^{k,h} + \tilde{D}(\partial_t^k)\gamma^- U^{k,h}$, $u_\star^{k,h} := \mathbb{G}[U_\star^{k,h}]$. We show that both $(U^{k,h}, U_\star^{k,h})$ and $X^{k,h}$ defined by (4.11) solve the following weak formulation for all $W := (Z_h, Z_\star) \in [\mathcal{X}_h^1]^m$ and times $t = t_n$:

$$\left\langle Y^{k,h}, W \right\rangle_{\mathcal{X}^0} + \left\langle ik[\mathcal{Q} \otimes \nabla] Y^{k,h}, \nabla W \right\rangle_{\mathcal{X}^0} + \left\langle ik[\mathcal{Q} \otimes \mathcal{V}] Y^{k,h}, W \right\rangle_{\mathcal{X}^0} = \left\langle y^{k,h} \mathbf{1}, W \right\rangle_{\mathcal{X}^0}, \quad (4.14)$$

where ∇ is applied to both components. For $Y^{k,h} \in \mathcal{X}_h^1$ the solution to this problem is unique by Lemma 3.23, and thus, if the solutions exist, they have to coincide.

We note that by (4.13) and (4.7b) we get that $\gamma^- U_\star^{k,h} \in [X_h^\circ]^m$. The fact that $Y^{k,h} := (U^{k,h}, U_\star^{k,h})$ then solves (4.14) follows from (4.7a) and integration by parts in (4.12). The boundary terms involving $W(\partial_t^k)$ and $1/2 - \tilde{K}^T(\partial_t^k)$ cancel due to (4.13). The boundary term $\langle \mathcal{Q} \llbracket \partial_\nu U_\star^{k,h} \rrbracket, \gamma^- W_\star \rangle_\Gamma$ vanishes since $\llbracket \partial_\nu U_\star^{k,h} \rrbracket = \Lambda^{k,h} \in [X_h]^m$ and $\gamma^- Z_\star \in [X_h^\circ]^m$.

The fact that $Y^{k,h} := X^{k,h}$ also solves (4.14) follows from (4.11), the definition of \mathbf{H}_h , and $X^{k,h} \in [\mathcal{X}_h^1]^m$ by assumption.

What is left to establish is that the original system (4.7) has a solution. From the abstract semigroup theory we get that $X^{k,h}$ exists and is unique. We show that $(U^{k,h}, \llbracket \partial_\nu U_\star^{k,h} \rrbracket)$ solves (4.7) by starting from (4.14). Testing with $W := (0, Z_\star)$ with $Z_\star \in [C_0^\infty(\mathbb{R}^d \setminus \Gamma)]^m$ gives the pointwise equality

$$U_\star^{k,h} - ik\mathcal{Q}\Delta U_\star^{k,h} + ik\mathcal{Q}\mathcal{V}_0 U_\star^{k,h} = u_\star^{k,h} \mathbf{1}. \quad (4.15)$$

Hence (4.14) becomes

$$\mathcal{A}(U^{k,h}, Z_h) - \langle ik\mathcal{Q}\partial_\nu^+ U_\star^{k,h}, \llbracket \gamma Z_\star \rrbracket \rangle_\Gamma + \langle i\mathcal{Q} \llbracket \partial_\nu U_\star^{k,h} \rrbracket, \gamma^- Z_\star \rangle_\Gamma = (u^{k,h} \mathbf{1}, Z_h)_{L^2(\Omega^-)}.$$

If we show that $\llbracket \partial_\nu U_\star^{k,h} \rrbracket \in [X_h]^m$ and $-\partial_\nu U_\star^{k,h} = \tilde{W}(\partial_t^k) \gamma^- U^{k,h} + (-1/2 + \tilde{K}^T(\partial_t^k)) \llbracket \partial_\nu U_\star^{k,h} \rrbracket$ this becomes (4.7a) since $\gamma^- Z_\star \in X_h^\circ$ and $\gamma^- Z_h = \llbracket \gamma Z_\star \rrbracket$ from the definition of \mathcal{X}_h^1 . To see $\llbracket \partial_\nu U_\star^{k,h} \rrbracket \in [X_h]^m$, we choose $\Xi \in [X_h^\circ]^m$ and $W := (0, \underline{\mathcal{E}}_\Gamma^D[\Xi])$ a lifting of Ξ . Integration by parts and the already established (4.15) gives

$$0 = \langle ik\mathcal{Q}\partial_\nu^+ U_\star^{k,h}, \Xi \rangle_\Gamma - \langle ik\mathcal{Q}\partial_\nu^- U_\star^{k,h}, \Xi \rangle_\Gamma,$$

i.e. $\llbracket \partial_\nu U_\star^{k,h} \rrbracket \in ([X_h^\circ]^m)^\circ = [X_h]^m$. The integral equations (4.7b) then follow by taking the Z -transform via Lemma 3.19 and Proposition 2.43. \square

What is left to do is to establish the connection between the original Schrödinger equation and the semigroup generated by $-i\mathbf{H}_h$. This is content of the following lemma:

Lemma 4.10. *Assume for the moment that $V_h := H^1(\Omega)$ and $X_h := H^{-1/2}(\Gamma)$. Then the following equivalences hold:*

- (i) *If $u \in H^2(\mathbb{R}^d)$ solves the Schrödinger equation (4.1), then $y := (v, v_\star)$ defined as $v := u|_{\Omega^-}$ and $v_\star := u|_{\Omega^+}$ in Ω^+ and $v_\star := 0$ in Ω^- solves the equation $iy = \mathbf{H}_h y$ and $y(0) = (u_0, 0)$.*
- (ii) *If $y := (v, v_\star)$ solves $iy = \mathbf{H}_h y$ and $y(0) = (u_0, 0)$, then $u := v$ in Ω^- and $u := v_\star$ in Ω^+ solves (4.1).*

Proof. Ad (i): Since u is in $H^2(\mathbb{R}^d)$ we get that $\partial_\nu^+ u = \partial_\nu^- u$. Inserting the definitions of (v, v_\star) , we get that $(-\Delta v + \mathcal{V}v, -\Delta v + \mathcal{V}_0 v)$ satisfies the defining equation for $\mathbf{H}_h(v, v_\star)$ in (4.9) via integration by parts. Since $\dot{y} = (\dot{u}|_\Omega, \dot{u}\mathbb{1}_{\Omega^+})$ where $\mathbb{1}_{\Omega^+}$ denotes the indicator function on Ω^+ . Therefore y is the solution to the semigroup.

Ad (i): Since $y \in \text{dom}(\mathbf{H}_h)$, we can insert test functions in $C_0^\infty(\Omega^-) \times C_0^\infty(\mathbb{R}^d \setminus \Gamma)$ in (4.9) and get that $\mathbf{H}_h y = (-\Delta v + \mathcal{V}v, -\Delta v_\star + \mathcal{V}_0 v_\star)$ using the weak Laplacian on both domains. From the definition of \mathcal{X}_h^1 we get $\gamma^- v = \gamma^+ v_\star$, and integration by parts implies that $\partial_\nu^- v = \partial_\nu^+ v_\star$, since the corresponding boundary term vanishes in the weak form. This means that u defined in the piecewise way is in $H^2(\mathbb{R}^d)$. The differential equation then follows from the fact that y is the semigroup solution. \square

In Lemma 4.10, we have seen that \mathbf{H} and \mathbf{H}_h are closely related in the case of no discretization in space. Namely, we have that $\mathbf{H}_h(u, u_\star) = (-\Delta v + \mathcal{V}v, -\Delta v + \mathcal{V}_0 v)$. Therefore we will extend the definition of \mathbf{H} to pairs of functions, i.e.,

$$\mathbf{H}(u, u_\star) := (-\Delta v + \mathcal{V}v, -\Delta v + \mathcal{V}_0 v), \quad (4.16)$$

with $\text{dom}(\mathbf{H}) = \text{dom}(\mathbf{H}_h)$ as defined in Definition 4.4 (using $V_h := H^1(\Omega)$ and $X_h := H^{-1/2}(\Gamma)$). Since the corresponding semigroups are closely related this should not cause confusion.

4.4 Analysis of the discretization scheme

This reformulation now allows us to make statements about the discretization by applying the abstract ‘‘Runge-Kutta for semigroups’’ theory. We begin with stability. It is well known that the Schrödinger equation conserves the L^2 norm and energy. While we cannot completely retain this property, we get the following similar result:

Theorem 4.11. *Let $(U^{k,h}, \Lambda^{k,h})$ solve (4.6a), define $U_\star^{k,h}$ via post-processing as in Lemma 4.8 and set $u_\star^{k,h} := \mathbb{G} [U_\star^{k,h}]$. Then the following estimates hold for all $n \in \mathbb{N}$:*

$$\left\| u^{k,h}(t_n) \right\|_{L^2(\Omega^-)}^2 + \left\| u_\star^{k,h}(t_n) \right\|_{L^2(\mathbb{R}^d)}^2 \leq \left\| u^{k,h}(0) \right\|_{L^2(\Omega^-)}^2. \quad (4.17)$$

For the energy $H(u) := \|\nabla u\|_{L^2(\mathbb{R}^d)}^2 + \|\mathcal{V}u\|_{L^2(\Omega^-)}^2$ we get for $\alpha \geq -\mathcal{V}_-$:

$$H(u^{k,h}(t_n)) + H(u_\star^{k,h}(t_n)) \leq H(u^{k,h}(0)) + \alpha \left\| u^{k,h}(0) \right\|_{L^2(\Omega^-)}^2. \quad (4.18)$$

If we assume that the Runge-Kutta method satisfies $|r(it)| = 1$ for $t \in \mathbb{R}$, we get conservation of mass and energy:

$$\left\| u^{k,h}(t_n) \right\|_{L^2(\Omega^-)}^2 + \left\| u_\star^{k,h}(t_n) \right\|_{L^2(\mathbb{R}^d)}^2 = \left\| u^{k,h}(0) \right\|_{L^2(\Omega^-)}^2, \quad (4.19)$$

$$H(u^{k,h}(t_n)) + H(u_\star^{k,h}(t_n)) = H(u^{k,h}(0)). \quad (4.20)$$

Proof. To see (4.17) and (4.19), we apply Lemma 3.27 to the reformulation (4.11).

For $\alpha \geq -\mathcal{V}_-$, the operator $B := \mathbf{H}_h + \alpha \text{Id}$ satisfies $\langle Bu, u \rangle_{\mathcal{X}^0} \geq 0$. Since it is also self-adjoint, we can use the spectral representation theorem to define $B^{1/2}$ with the following properties (see also [Ber68; Wou66] for elementary constructions of the operator square root):

- $B^{1/2}$ is self-adjoint
- $(B^{1/2})^2 = B$
- $\text{dom}(B^{1/2}) \supseteq \text{dom}(B)$
- $B^{1/2}$ commutes with B , i.e., $BB^{1/2} = B^{1/2}B$.

The last statement implies that $B^{1/2}$ commutes with \mathbf{H}_h , which in turn gives that

$$B^{1/2} [r(i\mathbf{H}_h)] x = r(i\mathbf{H}_h) B^{1/2} x \quad \text{for } x \in \text{dom}(\mathbf{H}_h).$$

This operator allows us to calculate, using the fact that $r(ki\mathbf{H}_h)$ is a contraction:

$$\begin{aligned} H \left(u^{k,h}(t_{n+1}) \right) + \alpha \left\| u^{k,h}(t_{n+1}) \right\|_{\mathcal{X}^0}^2 &= \left\langle B u^{k,h}(t_{n+1}), u^{k,h}(t_{n+1}) \right\rangle_{\mathcal{X}^0} \\ &= \left\| B^{1/2} u^{k,h}(t_{n+1}) \right\|_{\mathcal{X}^0}^2 \\ &= \left\| r(ik\mathbf{H}_h) B^{1/2} u^{k,h}(t_n) \right\|_{\mathcal{X}^0}^2 \\ &\leq \left\| B^{1/2} u^{k,h}(t_n) \right\|_{\mathcal{X}^0}^2 \\ &= H \left(u^{k,h}(t_n) \right) + \alpha \left\| u^{k,h}(t_n) \right\|_{\mathcal{X}^0}^2, \end{aligned}$$

we note that $u^{k,h}(t) \in \text{dom}(\mathbf{H}_h)$ for all $t \geq 0$. The conservation of H under the stricter assumptions follows along the same lines from the conservation of the \mathcal{X}^0 -norm, we just note that since the L^2 -norms are conserved, we can just subtract the terms $\alpha \left\| u^{k,h}(\cdot) \right\|_{\mathcal{X}^0}^2$ at the end. \square

Corollary 4.12. *For $m \in \mathbb{N}_0$, the discrete stability also holds for higher derivatives, i.e., assume that $x^{k,h}(0) \in \text{dom}(\mathbf{H}_h^m)$. Then we can estimate*

$$\begin{aligned} \left\| \mathbf{H}_h^m x^{k,h}(t_n) \right\|_{\mathcal{X}^0} &\leq \left\| \mathbf{H}_h^m x^{k,h}(0) \right\|_{\mathcal{X}^0}, \\ \left\| \mathbf{H}_h^m X^{k,h}(t_n) \right\|_{\mathcal{X}^0} &\leq C \left\| \mathbf{H}_h^m x^{k,h}(0) \right\|_{\mathcal{X}^0}. \end{aligned}$$

Proof. Applying \mathbf{H}_h commutes with the Runge-Kutta time-stepping. Therefore, the statement follows by applying the stability results from Lemma 3.30. \square

Remark 4.13. *The statements of Corollary 4.12 for $m > 1$ are mostly interesting for the semi-discrete case of $V_h = H^1(\Omega^-)$ and $X_h = H^{-1/2}(\Gamma)$, where they directly correspond to regularity assumptions on the initial data since $\mathbf{H}_h = \mathbf{H}$ is the classical (weak) Hamiltonian (see Lemma 4.21). \blacksquare*

4.4.1 Convergence with respect to space discretization

We now study the convergence of the approximation scheme. In order to do so, we consider two sequences of approximations X^k and $X^{k,h}$ of (4.11), where for X^k we use the full spaces $V_h := H^1(\Omega^-)$, $X_h = H^{-1/2}(\Gamma)$. We denote the corresponding Hamiltonian as defined in (4.9) by \mathbf{H} (or equivalently (4.16)). Since it corresponds to the pointwise Hamilton operator, where $-\Delta + \mathcal{V}$ is applied to u and $-\Delta + \mathcal{V}_0$ to u_* , this should not cause confusion.

Lemma 4.14. *We introduce a variation of the Ritz projector $\Pi_h : \text{dom}(\mathbf{H}) \subseteq \mathcal{X}^1 \rightarrow \mathcal{X}_h^1$ as*

$$\Pi_h := (\mathbf{H}_h - i)^{-1} \Pi_{\mathcal{X}_h^0} (\mathbf{H} - i)$$

where $\Pi_{\mathcal{X}_h^0} : \mathcal{X}^0 \rightarrow \mathcal{X}_h^0$ is the L^2 -projection. We write $\kappa_h := \Pi_{\mathcal{X}_h^0} - \Pi_h$.

Then the following error estimate holds:

$$\begin{aligned} \left\| x^k(t_n) - x^{k,h}(t_n) \right\|_{\mathcal{X}^0} &\leq \left\| \Pi_h x^k(0) - x^{k,h}(0) \right\|_{\mathcal{X}^0} + \left\| x^k(t_n) - \Pi_h x^k(t_n) \right\|_{\mathcal{X}^0} \\ &\quad + C k \sum_{j=0}^{n-1} \left\| \underline{\kappa}_h [X^k(t_j)] \right\|_{\mathcal{X}^0} + \left\| \underline{\kappa}_h [\underline{\mathbf{H}} X^k(t_j)] \right\|_{\mathcal{X}^0}. \end{aligned} \quad (4.21)$$

The constant $C > 0$ depends on \mathcal{Q} and \mathbf{b} .

Proof. We start by noting that, by construction, the operator Π_h satisfies

$$\mathbf{H}_h \Pi_h x = \Pi_{\mathcal{X}_h^0} [\mathbf{H}x - ix] + i \Pi_h x = \Pi_{\mathcal{X}_h^0} \mathbf{H}x - i \kappa_h [x] \quad \text{for } x \in \text{dom}(\mathbf{H}). \quad (4.22)$$

The function $\underline{\Pi}_h X^k$ solves (pointwise in time):

$$\underline{\Pi}_h X^k = \Pi_h x^k \mathbf{1} - ik [\mathcal{Q} \otimes \mathbf{H}_h] \underline{\Pi}_h X^k + \left[\underline{\Pi}_h X^k - \Pi_h x^k \mathbf{1} + ik [\mathcal{Q} \otimes \mathbf{H}_h \Pi_h] X^k \right].$$

We have a closer look at the consistency errors $\Delta^k := [\underline{\Pi}_h X^k - \Pi_h x^k \mathbf{1} + ik [\mathcal{Q} \otimes \mathbf{H}_h \Pi_h] X^k]$. By linearity and equation (4.11) in the semi-discrete setting, we have:

$$\begin{aligned} \Delta^k &= \left[\underline{\Pi}_h (X^k - x^k \mathbf{1}) + ik [\mathcal{Q} \otimes \mathbf{H}_h \Pi_h X^k] \right] = \left[\mathcal{Q} \otimes \Pi_h (-ik \underline{\mathbf{H}} X^k) + ik [\mathcal{Q} \otimes \mathbf{H}_h \Pi_h] X^k \right] \\ &= ik \left[[\mathcal{Q} \otimes \kappa_h] \underline{\mathbf{H}} X^k - i [\mathcal{Q} \otimes \kappa_h] X^k \right], \end{aligned} \quad (4.23)$$

where in the last step we used (4.22).

Therefore, the differences $E := X^{k,h} - \underline{\Pi}_h X^k$, $e := x^{k,h} - \Pi_h x^k$ solve:

$$\begin{aligned} E(t_n) &= e(t_n) \mathbf{1} - ik [\mathcal{Q} \otimes \mathbf{H}_h] E(t_n) + \Delta^k(t_n) \\ e(t_{n+1}) &= r(\infty) e(t_n) + \mathbf{b}^T \mathcal{Q}^{-1} E(t_n). \end{aligned}$$

The stability result of Lemma 3.30, together with the triangle inequality to estimate

$$\left\| x^k(t_n) - x^{k,h}(t_n) \right\|_{\mathcal{X}^0} \leq \left\| x^k(t_n) - \Pi_h x^k(t_n) \right\|_{\mathcal{X}^0} + \left\| \Pi_h x^k(t_n) - x^{k,h}(t_n) \right\|_{\mathcal{X}^0}$$

then implies (4.21). □

In order to estimate the approximation properties of Π_h , we have to get our hands dirty and see what it actually does in our concrete setting of spaces. This is done in the following lemma:

Lemma 4.15. *The operator Π_h has the following approximation property for all pairs $x = (u, u_\star) \in \text{dom}(\mathbf{H})$:*

$$\|x - \Pi_h x\|_{\mathcal{X}^1} \leq C \inf_{w_h \in \mathcal{X}_h^1} \|u - w_h\|_{\mathcal{X}^1} + C \inf_{\varphi_h \in X_h} \|[\partial_\nu u] - \varphi_h\|_{H^{-1/2}(\Gamma)}. \quad (4.24)$$

The constant $C > 0$ depends only on Ω^- and \mathcal{V}_- . This also immediately gives the estimate:

$$\|\kappa_h x\|_{\mathcal{X}^0} \leq C \inf_{w_h \in \mathcal{X}_h^1} \|u - w_h\|_{\mathcal{X}^1} + C \inf_{\varphi_h \in X_h} \|[\partial_\nu u] - \varphi_h\|_{H^{-1/2}(\Gamma)}. \quad (4.25)$$

Proof. By definition, $y := \Pi_h x$ solves for $w_h = (z_h, z_\star) \in \mathcal{X}_h^1$:

$$\begin{aligned} \langle (\mathbf{H}_h - i)y, w_h \rangle_{\mathcal{X}^0} &= \langle \Pi_{\mathcal{X}_h^0}(\mathbf{H} - i)x, w_h \rangle_{\mathcal{X}^0} \\ &= \langle (\mathbf{H} - i)x, w_h \rangle_{\mathcal{X}^0}. \end{aligned}$$

Using the definition of \mathbf{H}_h , and defining $a(x, w) := \langle \nabla x, \nabla w \rangle_{\mathcal{X}^0} + \langle (\mathcal{V} - i)x, \nabla w \rangle_{\mathcal{X}^0}$, this can be written as:

$$a(y, w_h) = \langle (\mathbf{H} - i)x, w_h \rangle_{\mathcal{X}^0}.$$

We note that w_h is not in the set of admissible test functions for the semi-discrete definition of \mathbf{H} due to the constraint $\gamma^- z_\star \notin [H^{-1/2}(\Gamma)]^\circ = \{0\}$. But, since $C_0^\infty(\mathbb{R}^d \setminus \Gamma) \subseteq \text{dom}(\mathbf{H})$, we get that the definition coincides with the weak Laplacian in both components together with the additional condition that $\partial_\nu^- u = \partial_\nu^+ u_\star$. Integration by parts then gives:

$$\langle (\mathbf{H} - i)x, w_h \rangle_{\mathcal{X}^0} = a(x, w_h) + \langle [\partial_\nu u_\star], \gamma^- z_\star \rangle_\Gamma.$$

This means, we can alternatively characterize the Ritz projector via

$$a(\Pi_h x, w_h) = a(x, w_h) + \langle [\partial_\nu u_\star], \gamma^- z_\star \rangle_\Gamma \quad \forall w_h = (z_h, z_\star) \in \mathcal{X}_h^1. \quad (4.26)$$

We now proceed similarly to the usual proof of Céa's Lemma. We write y_\star for the second component of y . For $w_h := (z_h, z_\star) \in \mathcal{X}_h^1$ and $\varphi_h \in X_h$ we calculate:

$$\begin{aligned} a(x - y, x - y) &= a(x - y, x - w_h) + a(x - y, w_h - y) \\ &= a(x - y, x - w_h) + \langle [\partial_\nu u_\star], \gamma^- z_\star - \gamma^- y_\star \rangle_\Gamma \\ &= a(x - y, x - w_h) + \langle [\partial_\nu u_\star] - \varphi_h, \gamma^- z_\star - \gamma^- y_\star \rangle_\Gamma, \end{aligned} \quad (4.27)$$

where in the last step we used that $\gamma^- z_\star, \gamma^- y_\star \in X_h^\circ$.

By using the fact that \mathcal{V} induces an \mathcal{X}^1 -bounded bilinear form by (4.10) and Young's inequality, this lets us bound

$$\begin{aligned} |a(x - y, x - y)| &\leq \|x - y\|_{\mathcal{X}^1} \|x - w_h\|_{\mathcal{X}^1} + \|[\partial_\nu u_\star] - \varphi_h\|_{H^{-1/2}(\Gamma)} \|w_h - z\|_{\mathcal{X}^1} \\ &\leq \varepsilon \|x - y\|_{\mathcal{X}^1}^2 + \varepsilon^{-1} \|[\partial_\nu u_\star] - \varphi_h\|_{H^{-1/2}(\Gamma)}^2 + 2\varepsilon^{-1} \|x - w_h\|_{\mathcal{X}^1}^2 + \varepsilon \|x - y\|_{\mathcal{X}^1}. \end{aligned}$$

From the real part of (4.27), we get:

$$\|x - y\|_{\mathcal{X}^0}^2 \leq 2\varepsilon \|x - y\|_{\mathcal{X}^1}^2 + \varepsilon^{-1} \|[\partial_\nu u_\star] - \varphi_h\|_{H^{-1/2}(\Gamma)}^2 + 2\varepsilon^{-1} \|x - w_h\|_{\mathcal{X}^1}^2.$$

Taking the imaginary part of (4.27) gives

$$\begin{aligned} \|\nabla x - \nabla y\|_{\mathcal{X}^0}^2 + \langle \mathcal{V}(x - y), x - y \rangle_{\mathcal{X}^0} \\ \lesssim 2\varepsilon \|x - y\|_{\mathcal{X}^1}^2 + \varepsilon^{-1} \|[\partial_\nu u_\star] - \varphi_h\|_{H^{-1/2}(\Gamma)}^2 + 2\varepsilon^{-1} \|x - w_h\|_{\mathcal{X}^1}^2. \end{aligned}$$

We add these two estimates and note that, since \mathcal{V} is bounded from below by \mathcal{V}_0 , we can choose ε sufficiently small to get (4.24). To see (4.25), we note that $\kappa_h = (\Pi_{\mathcal{X}_h^0} - \text{Id}) + (\text{Id} - \Pi_h)$. Since the L^2 -projection is optimal, we can absorb the error term corresponding to $\Pi_{\mathcal{X}_h^0} - \text{Id}$ into the $\text{Id} - \Pi_h$ term and apply (4.24). \square

Lemma 4.16. *The operator Π_h also has approximation properties with respect to the Hamiltonian, i.e. for all $x \in \text{dom}(\mathbf{H})$ we can estimate:*

$$\|\mathbf{H}x - \mathbf{H}_h \Pi_h x\|_{\mathcal{X}^0} \leq C \inf_{w_h \in \mathcal{X}_h^0} \|\mathbf{H}x - w_h\|_{\mathcal{X}^0} + C \|x - \Pi_h x\|_{\mathcal{X}^0}. \quad (4.28)$$

For the second component, the approximation is even better. If we write u_\star for the second component of $x \in \text{dom}(\mathbf{H})$ and y_\star for the second component of $\Pi_h x$, we get:

$$-\Delta y_\star + (\mathcal{V}_0 - i)y_\star = -\Delta u_\star + (\mathcal{V}_0 - i)u_\star \quad \text{in } \mathbb{R}^d \setminus \Gamma.$$

Proof. By definition, we have $\mathbf{H}_h \Pi_h x = \Pi_{\mathcal{X}_h^0} \mathbf{H}x - i\kappa_h[x]$. Therefore, we calculate:

$$\begin{aligned} \|\mathbf{H}x - \mathbf{H}_h \Pi_h x\|_{\mathcal{X}^0} &\leq \left\| \mathbf{H}x - \Pi_{\mathcal{X}_h^0} \mathbf{H}x \right\|_{\mathcal{X}^0} + \|\kappa_h[x]\|_{\mathcal{X}^0} \\ &\leq \left\| \mathbf{H}x - \Pi_{\mathcal{X}_h^0} \mathbf{H}x \right\|_{\mathcal{X}^0} + \left\| (\text{Id} - \Pi_{\mathcal{X}_h^0})x \right\|_{\mathcal{X}^0} + \|(\text{Id} - \Pi_h)x\|_{\mathcal{X}^0}. \end{aligned}$$

Since the L^2 projection is optimal, estimate (4.28) follows.

To see equality in the second component, we take test functions of the form $(0, w_\star)$ with $w_\star \in C_0^\infty(\mathbb{R}^d \setminus \Gamma)$. Inserting the definition of \mathbf{H} and \mathbf{H}_h gives:

$$\begin{aligned} (\nabla y_\star, \nabla w_\star)_{L^2(\mathbb{R}^d)} + ((\mathcal{V}_0 - i)y_\star, w_\star)_{L^2(\mathbb{R}^d)} \\ = (\nabla u_\star, \nabla w_\star)_{L^2(\mathbb{R}^d)} + ((\mathcal{V}_0 - i)u_\star, w_\star)_{L^2(\mathbb{R}^d)}, \end{aligned}$$

where we used the equality $\mathbf{H}_h \Pi_h x - i\Pi_h x = \Pi_{\mathcal{X}_h^0}(\mathbf{H}x - ix)$. By the definition of the weak Laplacian this concludes the proof. \square

Lemma 4.17. *The space \mathcal{X}_h^1 has the following approximation properties for all $x = (u, u_\star) \in \mathcal{X}^1$ with $\gamma^- u_\star = 0$ and $[\gamma u_\star] = \gamma^- u$:*

$$\begin{aligned} \inf_{w_h \in \mathcal{X}_h^1} \|x - w_h\|_{\mathcal{X}^0} &\leq C \inf_{v_h \in V_h} \|u - v_h\|_{L^2(\Omega^-)}, \\ \inf_{w_h \in \mathcal{X}_h^1} \|x - w_h\|_{\mathcal{X}^1} &\leq C \inf_{v_h \in V_h} \|u - v_h\|_{H^1(\Omega^-)}, \end{aligned}$$

where both constants depend on Ω^- . The function $w_h = (z_h, z_\star)$ in the infima can be chosen to satisfy $\gamma^- z_\star = 0$.

Proof. For $v_h \in V_h$, we define the function $w_h := (z_h, z_\star) \in \mathcal{X}_h^1$ as $z_h := v_h$, $z_\star := u_\star$ on Ω^- and $z_\star := u_\star + \delta_\star$ on Ω^+ with $\delta_\star := \mathcal{E}(v_h - u)$. Here, \mathcal{E} denotes the Stein extension operator to Ω^+ from Proposition 2.23. By definition, we have $z_h \in V_h$, and $\gamma^- z_\star = 0 \in X_h^0$. The Dirichlet jump satisfies:

$$\llbracket \gamma z_\star \rrbracket = \llbracket \gamma u_\star \rrbracket + \gamma^+ \delta_\star = \gamma^- u + \gamma^- v_h - \gamma^- u = \gamma^- v_h$$

due to the extension properties of \mathcal{E} . This means $w \in \mathcal{X}_h^1$. The L^2 - and H^1 -estimates on $x - w$ then follow immediately from the definitions and the stability of the Stein extension. \square

While L^2 is a natural setting for the Schrödinger equation, we are also interested in the convergence rate in stronger norms, most notably the convergence of the boundary variable $\Lambda^{k,h} \rightarrow \Lambda^k$.

Lemma 4.18. *Assume that the Runge-Kutta method satisfies $|r(\infty)| < 1$, and that the semi-discretization satisfies $U^k(t_n) \in \text{dom}(\mathbf{H}^2)$ for all $n \in \mathbb{N}_0$. Then the following error estimate holds, where the implied constant depends only on the Runge-Kutta method:*

$$\begin{aligned} & \left\| \mathbf{H}_h x^{k,h}(t_n) - \mathbf{H} x^k(t_n) \right\|_{\mathcal{X}^0} \\ & \lesssim \left\| \mathbf{H} x^k(t_n) - \mathbf{H}_h \Pi_h x^k(t_n) \right\|_{\mathcal{X}^0} \\ & \quad + \left\| \kappa_h [\mathbf{H} x^k(0)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [x^k(0)] \right\|_{\mathcal{X}^0} + \left\| \mathbf{H}_h \Pi_h x^k(0) - \mathbf{H}_h x^{k,h}(0) \right\|_{\mathcal{X}^0} \\ & \quad + k \sum_{j=0}^{n-1} \left(\left\| \kappa_h [\mathbf{H}^2 X^k(t_j)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [\mathbf{H} X^k(t_j)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [X^k(t_j)] \right\|_{\mathcal{X}^0} \right). \end{aligned}$$

We can also quantify the convergence in the H^1 -norm by:

$$\begin{aligned} & \left\| x^{k,h}(t_n) - x^k(t_n) \right\|_{\mathcal{X}^1} \\ & \lesssim \left\| x^k(t_n) - \Pi_h x^k(t_n) \right\|_{\mathcal{X}^1} \\ & \quad + \left\| \kappa_h [\mathbf{H} x^k(0)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [x^k(0)] \right\|_{\mathcal{X}^0} + \left\| \Pi_h x^k(0) - x^{k,h}(0) \right\|_{\mathcal{X}^1} \\ & \quad + k \sum_{j=0}^{n-1} \left(\left\| \kappa_h [\mathbf{H}^2 X^k(t_j)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [\mathbf{H} X^k(t_j)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [X^k(t_j)] \right\|_{\mathcal{X}^0} \right). \end{aligned}$$

Proof. For ease of presentation, we assume that $x^{k,h}(0) = \Pi_h x^k(0)$. The general statement can then be recovered by using the stability from Theorem 4.11 and Corollary 4.12 when perturbing the initial conditions.

As in the proof of Lemma 4.14, we start by again considering $E(t) := X^{k,h}(t) - \Pi_h X^k(t)$, $e := \mathbf{G}[E]$. This quantity satisfies (cf. Lemma 4.14)

$$\begin{aligned} E(t_n) &= e(t_n) \mathbf{1} - ik[\mathcal{Q} \otimes \mathbf{H}_h]E(t_n) + [\mathcal{Q} \otimes \text{Id}] \Delta^k(t_n) \\ e(t_{n+1}) &= r(\infty)e(t_n) + k \mathbf{b}^T \mathcal{Q}^{-1} E(t_n), \end{aligned}$$

where $\Delta^k := ik(\underline{\kappa}_h [\mathbf{H}X^{k,h}] - i\underline{\kappa}_h [X^{k,h}])$ is the consistency error from (4.23) (up to a factor \mathcal{Q} which is factored out to fit the setting of Lemma 3.36). By Lemma 3.36(ii), we can write $\mathbf{H}_h E(t_n) = k^{-1}\mathcal{Q}^{-1}(E(t_n) - e(t_n)\mathbf{1}) - \Delta(t_n)$, and by Lemma 3.36(i) and the discrete stability (Lemma 3.30) we can bound

$$\|\mathbf{H}_h E(t_n)\|_{\mathcal{X}^0} \leq C \sum_{j=0}^n \left\| \Theta^k(t_j) \right\|_{\mathcal{X}^0} \quad (4.29)$$

with $\Theta^k := [\partial_t^k \Delta^k]$. Next, we need to compute Θ^k explicitly. For simplicity, we focus on the first term $\Delta_1^k := ik\underline{\kappa}_h [\underline{\mathbf{H}}X^k]$ and set $\Theta_1^k := [\partial_t^k \Delta_1^k]$. Taking the Z -transform of X^k via Lemma 3.19, we get that

$$-\frac{\delta(z)}{k} \widehat{X}^k = i\mathbf{H}\widehat{X}^k + \frac{1}{1-r(\infty)z} k^{-1}\mathcal{Q}^{-1}\mathbf{1}x^k(0)$$

and by linearity of κ_h :

$$\begin{aligned} \mathcal{Z} \left[\Theta_1^k \right] &= \frac{\delta(z)}{k} \mathcal{Z} \left[\Delta_1^k \right] = ik\underline{\kappa}_h \left[\underline{\mathbf{H}} \frac{\delta(z)}{k} \widehat{X}^k \right] \\ &= k\underline{\kappa}_h \left(\underline{\mathbf{H}}^2 \widehat{X}^k \right) - \left[i \frac{1}{1-r(\infty)z} \kappa_h \left(\mathbf{H}x^k(0) \right) \right] \mathcal{Q}^{-1}\mathbf{1}. \end{aligned}$$

Taking the inverse Z -transform implies for Θ_1 by using the geometric series for $\frac{1}{1-r(\infty)z}$:

$$\Theta_1^k(t_j) = k\underline{\kappa}_h \left[\underline{\mathbf{H}}^2 X^k(t_j) \right] - i[r(\infty)]^j \kappa_h \left[\mathbf{H}x^k(0) \right] \mathcal{Q}^{-1}\mathbf{1}.$$

The first term has the right powers of k in order to not impact the convergence rates. For the second one, we use the geometric series and our assumption that $|r(\infty)| < 1$ to estimate:

$$\begin{aligned} \sum_{j=0}^n \left\| \Theta_1^k(t_j) \right\| &\leq Ck \sum_{j=0}^n \left\| \kappa_h \left(\underline{\mathbf{H}}^2 X^k(t_j) \right) \right\|_{\mathcal{X}^0} + C \sum_{j=0}^n |r(\infty)^j| \left\| \kappa_h [\mathbf{H}x^k(0)] \right\|_{\mathcal{X}^0} \\ &\leq Ck \sum_{j=0}^n \left\| \kappa_h \left(\underline{\mathbf{H}}^2 X^k(t_j) \right) \right\|_{\mathcal{X}^0} + C \frac{1-|r(\infty)|^n}{1-|r(\infty)|} \left\| \kappa_h \left[\mathbf{H}x^k(0) \right] \right\|_{\mathcal{X}^0}. \end{aligned}$$

A completely analogous computation gives the explicit representation of Θ_2^k as:

$$\Theta_2^k(t_j) = -k\underline{\kappa}_h \left(\underline{\mathbf{H}}X^k(t_j) \right) + i[r(\infty)]^j \kappa_h \left[x^k(0) \right] \mathcal{Q}^{-1}\mathbf{1},$$

which can be estimated along the same lines as Θ_1^k . Collecting the estimates thus far gives:

$$\begin{aligned} &\left\| \mathbf{H}_h X^{k,h}(t_n) - \mathbf{H}_h \Pi_h X^k(t_n) \right\|_{\mathcal{X}^0} \\ &\lesssim \left\| \kappa_h \mathbf{H}x^k(0) \right\|_{\mathcal{X}^0} + \left\| \kappa_h x^k(0) \right\|_{\mathcal{X}^0} \\ &\quad + k \sum_{j=0}^n \left(\left\| \kappa_h \left[\underline{\mathbf{H}}^2 X^k(t_j) \right] \right\|_{\mathcal{X}^0} + \left\| \kappa_h \left[\underline{\mathbf{H}}X^k(t_j) \right] \right\|_{\mathcal{X}^0} + \left\| \kappa_h \left[X^k(t_j) \right] \right\|_{\mathcal{X}^0} \right). \end{aligned} \quad (4.30)$$

Since $\mathbf{H}_h x^{k,h} - \mathbf{H}_h \Pi_h x^k = \mathbf{G} [\mathbf{H}_h X^{k,h} - \mathbf{H}_h \Pi_h X^k]$, we get inductively:

$$\left\| \mathbf{H}_h x^{k,h}(t_n) - \mathbf{H}_h \Pi_h x^k(t_n) \right\|_{\mathcal{X}^0} \lesssim \sum_{j=0}^n |r(\infty)|^{n-j} \left\| \mathbf{H}_h X^{k,h}(t_j) - \mathbf{H}_h \Pi_h X^{k,h}(t_j) \right\|_{\mathcal{X}^0}.$$

A triangle inequality to relate $\mathbf{H}x^k - \mathbf{H}_h x^k$ to $\mathbf{H}_h[x^{k,h} - \Pi_h x^k]$ and $\mathbf{H}x^k - \mathbf{H}_h \Pi_h x^k$ concludes the proof of the estimate for $\mathbf{H}x^{k,h}$. To get the estimate in the \mathcal{X}^1 -norm, we calculate for any function $x_h \in \text{dom}(\mathbf{H}_h)$ since \mathcal{V} is bounded from below:

$$\begin{aligned} \langle \nabla x_h, \nabla x_h \rangle_{\mathcal{X}^0} &= \langle \mathbf{H}_h x_h, x_h \rangle_{\mathcal{X}^0} - \langle \mathcal{V} x_h, x_h \rangle_{\mathcal{X}^0} \\ &\lesssim \|\mathbf{H}_h x_h\|_{\mathcal{X}^0} \|x_h\|_{\mathcal{X}^0} + \|x_h\|_{\mathcal{X}^0}^2 \\ &\lesssim \|\mathbf{H}_h x_h\|_{\mathcal{X}^0}^2 + \|x_h\|_{\mathcal{X}^0}^2. \end{aligned} \quad (4.31)$$

Setting $x_h := [x^{k,h}(t_n) - \Pi_h x^{k,h}(t_n)]$ then gives:

$$\left\| x^{k,h}(t_n) - \Pi_h x^{k,h}(t_n) \right\|_{\mathcal{X}^1} \lesssim \|\mathbf{H}_h x_h\|_{\mathcal{X}^0} + \|x_h\|_{\mathcal{X}^0}.$$

Therefore the estimate in the \mathcal{X}^1 -norm follows from the estimate in \mathcal{X}^0 and the estimate on $\mathbf{H}_h x^{k,h}$ together with the triangle estimate to get to $x^k - x^{k,h}$ instead of the Ritz-projector. \square

A refined L^2 -estimate on convex or smooth geometries

The convergence of our method relies on the approximation properties of the Ritz-projector. As seen in Lemma 4.15, the approximation with respect to the H^1 -based norm is quasi-optimal, but for the L^2 -norm we would hope to get an increased rate of convergence, as Lemma 4.15 tells us that the space \mathcal{X}_h^1 does indeed have such an improved approximation property. While we cannot obtain this rate in the fully general setting, the improved rate can indeed be proven by using the common Aubin-Nitsche trick under slight additional assumptions. This is the content of the following lemma:

Lemma 4.19. *Assume that Ω^- is convex or has a smooth boundary (such that a shift theorem holds for the homogeneous Dirichlet problem, see [Gri11]).*

Assume that the spaces V_h and X_h satisfy the following approximation property:

$$\inf_{v_h \in V_h} \|\psi - v_h\|_{H^1(\Omega^-)} + \inf_{x_h \in X_h} \|\lambda - x_h\|_{H^{-1/2}(\Gamma)} \leq C_{\text{approx}} h \left(\|\psi\|_{H^2(\Omega^-)} + \|\lambda\|_{H^{1/2}(\Gamma)} \right) \quad (4.32)$$

for all $(\psi, \lambda) \in H^2(\Omega^-) \times H^{1/2}(\Gamma)$.

Let $x =: (u, u^*) \in \text{dom}(\mathbf{H})$, in particular $u \in H_{\Delta}^1(\Omega^-)$, $u_* \in H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)$ and $\gamma^- u = \llbracket \gamma u^* \rrbracket$, as well as $\gamma^- u^* = 0$. Then the following error estimate holds for the Ritz projector Π_h :

$$\|x - \Pi_h x\|_{\mathcal{X}^0} \leq Ch \|x - \Pi_h x\|_{\mathcal{X}^1}.$$

The constant C depends only on C_{approx} and Γ .

Proof. We write $\Pi_h x =: (u_h, u_h^*)$ for the two components. Consider the solutions ψ_1, ψ_2 to the following two problems:

$$\begin{aligned} -\Delta\psi_1 + (\mathcal{V} - i)\psi_1 &= \begin{cases} u - u_h & \text{in } \Omega^- \\ u^* - u_h^* & \text{in } \Omega^+ \end{cases}, \\ \llbracket \gamma\psi_1 \rrbracket = \llbracket \partial_\nu\psi_1 \rrbracket &= 0, \\ -\Delta\psi_2 + (\mathcal{V}_0 - i)\psi_2 &= u^* - u_h^* \quad \text{in } \Omega^-, \\ \gamma^-\psi_2 &= 0. \end{aligned}$$

We note that ψ_1 is the solution to an elliptic full space problem. Using Fourier techniques (see e.g., [RS75, page 52]) we can estimate $\|\psi_1\|_{H^2(\mathbb{R}^d)} \lesssim \|-\Delta\psi_1\|_{L^2(\mathbb{R}^d)} + \|\psi_1\|_{H^1(\mathbb{R}^d)}$. Due to our assumptions on \mathcal{V} and Lemma 4.6, we can further estimate for constants $0 \leq a < 1/2$ and $b \geq 0$:

$$\begin{aligned} \|-\Delta\psi_1\|_{L^2(\mathbb{R}^d)} &\leq \|\mathbf{H}\psi_1\|_{L^2(\mathbb{R}^d)} + \|\mathcal{V}\psi_1\|_{L^2(\mathbb{R}^d)} \\ &\leq \|\mathbf{H}\psi_1\|_{L^2(\mathbb{R}^d)} + a\|-\Delta\psi_1\|_{L^2(\mathbb{R}^d)} + b\|\psi_1\|_{L^2(\mathbb{R}^d)}, \end{aligned}$$

which in turn gives due to the equation for $\mathbf{H}\psi_1$:

$$\begin{aligned} \|\psi_1\|_{H^2(\mathbb{R}^d)} &\lesssim \|\mathbf{H}\psi_1\|_{L^2(\mathbb{R}^d)} + \|\psi_1\|_{H^1(\mathbb{R}^d)} \\ &\lesssim \|x - \Pi_h x\|_{\mathcal{X}^0}. \end{aligned}$$

Since we assumed that a shift theorem holds for Ω^- , and we are working with the constant potential \mathcal{V}_0 , the same estimate holds for ψ_2 , i.e., $\|\psi_2\|_{H^2(\Omega^-)} \leq C\|x - \Pi_h x\|_{\mathcal{X}^0}$.

We rearrange the terms into

$$\begin{aligned} \psi &:= \psi_1|_{\Omega^-}, \\ \psi^* &:= \begin{cases} \psi_2 & \text{in } \Omega^- \\ \psi_1 & \text{in } \Omega^+ \end{cases}, \end{aligned}$$

and write $\chi := (\psi, \psi^*)$. Integration by parts then gives (again using the bilinear form $a(\cdot, \cdot)$ from the proof of Lemma 4.15:

$$\begin{aligned} \|u - u_h\|_{L^2(\Omega^-)}^2 + \|u^* - u_h^*\|_{L^2(\mathbb{R}^d)}^2 &= (-\Delta\psi + (\mathcal{V} - i)\psi, u - u_h)_{L^2(\Omega^-)} + (-\Delta\psi^* + (\mathcal{V}_0 - i)\psi^*, u^* - u_h^*)_{L^2(\mathbb{R}^d)} \\ &= a(\chi, x - \Pi_h x) - \langle \partial_\nu^-\psi, \gamma^-(u - u_h) \rangle_\Gamma \\ &\quad - \langle \partial_\nu^-\psi^*, \gamma^-(u^* - u_h^*) \rangle_\Gamma + \langle \partial_\nu^+\psi^*, \gamma^+(u^* - u_h^*) \rangle_\Gamma \\ &= a(\chi, x - \Pi_h x) - \langle \partial_\nu^-\psi, \gamma^-(u - u_h) \rangle_\Gamma + \langle \partial_\nu^+\psi^*, \llbracket \gamma(u^* - u_h^*) \rrbracket \rangle_\Gamma \\ &\quad + \langle \llbracket \partial_\nu\psi^* \rrbracket, \gamma^-(u^* - u_h^*) \rangle_\Gamma. \end{aligned}$$

Since $\partial_n^-\psi = \partial_n^-\psi_1 = \partial_n^+\psi_1 = \partial_n^+\psi^*$ and $\gamma^-(u - u_h) = \llbracket \gamma(u^* - u_h^*) \rrbracket$, this becomes:

$$\|u - u_h\|_{L^2(\Omega^-)}^2 + \|u^* - u_h^*\|_{L^2(\mathbb{R}^d)}^2 = a(\chi, x - \Pi_h x) + \langle \llbracket \partial_\nu\psi^* \rrbracket, \gamma^-(u^* - u_h^*) \rangle_\Gamma.$$

For $\chi_h := (\psi_h, \psi_h^*) \in \mathcal{X}_h^1$ with $\gamma^- \psi_h^* = 0$ and $\lambda_h, \mu_h \in X_h$ we can use the alternative characterization of the Ritz-projector (4.26) as well as the knowledge that $\gamma^- \psi_h = 0$ and $\gamma^-(u - u_h) \in X_h^\circ$, to get:

$$\begin{aligned} & \|u - u_h\|_{L^2(\Omega^-)}^2 + \|u^* - u_h^*\|_{L^2(\mathbb{R}^d)}^2 \\ &= a(\chi - \chi_h, x - \Pi_h x) + \langle [\partial_\nu \psi^*] - \mu_h, \gamma^-(u^* - u_h^*) \rangle_\Gamma \\ &\lesssim \left(\|\chi - \chi_h\|_{\mathcal{X}^1} + \|[\partial_\nu \psi^*] - \mu_h\|_{H^{-1/2}(\Gamma)} \right) \|x - \Pi_h x\|_{\mathcal{X}^1}. \end{aligned}$$

Here we used that the bilinear form $a(\cdot, \cdot)$ is bounded on \mathcal{X}^1 via (4.10). The best approximation property of \mathcal{X}_h^1 , derived in Lemma 4.15 and the assumed approximation property (4.32) (note that the restriction to discrete functions satisfying $\gamma^- \psi_h = 0$ does not impact the best approximation property, see Lemma 4.17) then give

$$\begin{aligned} \|u - u_h\|_{L^2(\Omega^-)}^2 + \|u^* - u_h^*\|_{L^2(\mathbb{R}^d)}^2 &\lesssim h \left[\|\psi\|_{H^2(\Omega^-)} + \|[\partial_\nu \psi^*]\|_{H^{1/2}(\Gamma)} \right] \|x - \Pi_h x\|_{\mathcal{X}^1} \\ &\lesssim h \|x - \Pi_h x\|_{\mathcal{X}^0} \|x - \Pi_h x\|_{\mathcal{X}^1}, \end{aligned}$$

where in the last step we used the regularity of (ψ, ψ^*) . \square

Looking back at Lemma 4.18, the argument only works for a certain class of methods, such as the Runge-Kutta, but does not hold for example for the Gauss methods. Since Gauss methods are of interest due to their better conservation of energy (see Theorem 4.11), we also show a variation of Lemma 4.14, which covers all A-stable methods but requires a condition on the time step size, the approximation quality of V_h as well as on the geometry of Ω^- .

Lemma 4.20. *Assume that the requirements of Lemma 4.19 hold, namely that Ω^- is convex or has smooth boundary and that the spaces V_h and X_h have the approximation property (4.32). Then the following error estimate holds, where the implied constants depends on the Runge-Kutta method, Ω and C_{approx} :*

$$\begin{aligned} \left\| \mathbf{H}x^k(t_n) - \mathbf{H}_h x^{k,h}(t_n) \right\|_{\mathcal{X}^0} &\lesssim \left\| \mathbf{H}x^{k,h}(0) - \mathbf{H}_h \Pi_h x^k(0) \right\|_{\mathcal{X}^0} + \left\| \mathbf{H}_h x^{k,h}(t_n) - \mathbf{H}_h \Pi_h x^k(t_n) \right\|_{\mathcal{X}^0} \\ &\quad + h \sum_{j=0}^{n-1} \left(\left\| \kappa_h [\mathbf{H}X^k(t_j)] \right\|_{\mathcal{X}^1} + \left\| \kappa_h [X^k(t_j)] \right\|_{\mathcal{X}^1} \right). \end{aligned}$$

Thus for $h \lesssim k$ we regain full order of convergence. In the \mathcal{X}^1 norm we get

$$\begin{aligned} \left\| x^k(t_n) - x^{k,h}(t_n) \right\|_{\mathcal{X}^1} &\lesssim \left\| x^k(t_n) - \Pi_h x^k(t_n) \right\|_{\mathcal{X}^1} + \left\| x^{k,h}(0) - \Pi_h x^k(0) \right\|_{\mathcal{X}^1} \\ &\quad + (hk + h\sqrt{k}) \sum_{j=0}^{n-1} \left(\left\| \kappa_h [\mathbf{H}X^k(t_j)] \right\|_{\mathcal{X}^1} + \left\| \kappa_h [X^k(t_j)] \right\|_{\mathcal{X}^1} \right), \end{aligned}$$

i.e., it suffices to assume $h \lesssim \sqrt{k}$ to retain full convergence rates.

Proof. As in the previous lemmas, we assume that $x^{k,h}(0) = \Pi_h x^k(0)$ and we consider the quantities $E(t) := X^{k,h}(t) - \Pi_h X^k(t)$, $e := \mathbf{G}[E]$. These functions satisfy (cf. Lemma 4.14)

$$\begin{aligned} E(t_n) &= e(t_n)\mathbf{1} - ik[\mathcal{Q} \otimes \mathbf{H}_h]E(t_n) + \Delta^k(t_n) \\ e(t_{n+1}) &= r(\infty)e(t_n) + k\mathbf{b}^T \mathcal{Q}^{-1}E(t_n), \end{aligned}$$

where $\Delta^k := ik \left[[\mathcal{Q} \otimes \kappa_h] \mathbf{H} X^{k,h} - i [\mathcal{Q} \otimes \kappa_h] X^{k,h} \right]$ is the consistency error from (4.23). Using the notation $g(-ik\mathbf{H}_h) := \mathbf{b}^T \mathcal{Q}^{-1} [\text{Id} + ik\mathcal{Q} \otimes \mathbf{H}_h]^{-1}$, and using the stability function $r(-ik\mathbf{H}_h)$ we get (cf. Lemma 3.29):

$$e(t_{n+1}) = r(-ik\mathbf{H}_h)e(t_n) + g(-ik\mathbf{H}_h)\Delta^k(t_n). \quad (4.33)$$

The operator $k\mathbf{H}_h g(-ik\mathbf{H}_h)$ satisfies:

$$\begin{aligned} k\mathbf{H}_h g(-ik\mathbf{H}_h) &= k\mathbf{b}^T \mathcal{Q}^{-1} \underline{\mathbf{H}}_h [\text{Id} + ik\mathcal{Q} \otimes \mathbf{H}_h]^{-1} \\ &= -i\mathbf{b}^T \mathcal{Q}^{-1} \mathcal{Q}^{-1} \left(\text{Id} - [\text{Id} + ik\mathcal{Q} \otimes \mathbf{H}_h]^{-1} \right). \end{aligned}$$

We can write $[\text{Id} + ik\mathcal{Q} \otimes \mathbf{H}_h] = -\mathcal{Q} [-ik\underline{\mathbf{H}}_h - \mathcal{Q}^{-1}]$. The spectrum of \mathcal{Q}^{-1} satisfies $\text{Re}(\lambda) > 0$ due to A-stability. From (2.7) we get that the resolvent of $-ik\underline{\mathbf{H}}_h$ is uniformly bounded on $\sigma(\mathcal{Q}^{-1})$ with respect to k . This means that $k\underline{\mathbf{H}}_h(g - ik\underline{\mathbf{H}}_h)$ is a bounded operator on \mathcal{X}_h^0 via Lemma 3.21. \mathbf{H}_h commutes with $r(-ik\mathbf{H}_h)$, and we get from (4.33) by applying \mathbf{H}_h to both sides and taking the norm:

$$\|\mathbf{H}_h e(t_{n+1})\|_{\mathcal{X}^0} \leq \|r(-ik\mathbf{H}_h)\mathbf{H}_h e(t_n)\|_{\mathcal{X}^0} + Ck^{-1} \left\| \Delta^k(t_n) \right\|_{\mathcal{X}^0}.$$

Using Lemma 4.19, we can estimate the \mathcal{X}^0 -norm of $\Delta^k(t_n)$ by

$$\begin{aligned} \left\| \Delta^k(t_n) \right\|_{\mathcal{X}^0} &\lesssim k \left(\left\| \kappa_h [\underline{\mathbf{H}} X^k(t_n)] \right\|_{\mathcal{X}^0} + \left\| \kappa_h [X^k(t_n)] \right\|_{\mathcal{X}^0} \right) \\ &\lesssim kh \left(\left\| (\text{Id} - \Pi_h) [\underline{\mathbf{H}} X^k(t_n)] \right\|_{\mathcal{X}^1} + \left\| (\text{Id} - \Pi_h) [X^k(t_n)] \right\|_{\mathcal{X}^1} \right). \end{aligned}$$

This proves the estimate for $\mathbf{H}x^k - \mathbf{H}_h x^{k,h}$. Using the same argument as in Lemma 4.18 (see (4.31)), we get for the \mathcal{X}^1 -norm:

$$\begin{aligned} \|\nabla e(t_{n+1})\|_{\mathcal{X}^0}^2 &\lesssim \|\mathbf{H}_h e(t_{n+1})\|_{\mathcal{X}^0} \|e(t_{n+1})\|_{\mathcal{X}^0} + \|e(t_{n+1})\|_{\mathcal{X}^0}^2 \\ &\lesssim k^{-1} \left(\sum_{j=0}^n \left\| \Delta^k(t_j) \right\|_{\mathcal{X}^0} \right)^2 + \left(\sum_{j=0}^n \left\| \Delta^k(t_j) \right\|_{\mathcal{X}^0} \right)^2 \\ &\lesssim (h^2 k + h^2 k^2) \left(\sum_{j=0}^n \left[\left\| (\text{Id} - \Pi_h) [\underline{\mathbf{H}} X^k(t_j)] \right\|_{\mathcal{X}^1} + \left\| (\text{Id} - \Pi_h) [X^k(t_j)] \right\|_{\mathcal{X}^1} \right] \right)^2. \end{aligned}$$

Taking the square root, we get the stated result, after again using the triangle inequality to estimate

$$\left\| x^{k,h} - x^k \right\|_{\mathcal{X}^1} \leq \|e\|_{\mathcal{X}^1} + \left\| \Pi_h x^k - x^k \right\|_{\mathcal{X}^1}. \quad \square$$

4.4.2 Convergence of the fully discrete scheme

In this section we investigate the convergence of the fully discrete scheme, i.e. Runge-Kutta in time and the FEM-BEM coupling in space. This is now a simple consequence of the previous section and the general Runge-Kutta theory for semigroups presented in Section 3.4.

We start with a small lemma, giving a sufficient condition for the regularity of the semi-discrete stage vector:

Lemma 4.21. *Let $u_0 \in H^s(\Omega^-)$ for $s \geq 2$ and assume that \mathcal{V} is sufficiently smooth. Then the following estimates holds:*

$$\left\| u^k(t_n) \right\|_{H^s(\mathbb{R}^d)} \leq C \|u_0\|_{H^s(\Omega^-)} \quad \text{and} \quad \left\| U^k(t_n) \right\|_{[H^s(\mathbb{R}^d)]^m} \leq C \|u_0\|_{H^s(\Omega^-)}.$$

The constant $C > 0$ depends on s and the potential \mathcal{V} .

Proof. We first show the case $s = 2\ell$ for $\ell \in \mathbb{N}$. We have already established the following estimates in Corollary 4.12:

$$\left\| \mathbf{H}^\ell u^k(t_n) \right\|_{L^2(\mathbb{R}^d)} \leq C \|u_0\|_{H^{2\ell}(\Omega^-)}, \quad \text{and} \quad \left\| \mathbf{H}^\ell U^k(t_n) \right\|_{[L^2(\mathbb{R}^d)]^m} \leq C \|u_0\|_{H^{2\ell}(\Omega^-)}.$$

Since we assumed that \mathcal{V} is smooth, we can estimate $\left\| \Delta^\ell u^k(t_n) \right\| \leq C \left\| \mathbf{H}^\ell u^k(t_n) \right\|_{L^2(\mathbb{R}^d)} +$ lower order terms, and using Fourier techniques (see e.g., [RS75, page 52]) this allows us to bound

$$\left\| u^k(t_n) \right\|_{H^{2\ell}(\mathbb{R}^d)} \leq C \left\| \mathbf{H}^\ell u^k(t_n) \right\|_{L^2(\mathbb{R}^d)} + \text{lower order terms.},$$

By induction in ℓ , we get the bound of the $H^{2\ell}(\mathbb{R}^d)$ norm. The general case follows by interpolation of the operator $T_n : u_0 \mapsto u^k(t_n)$. The proof for the stages follows along the same lines. \square

We make the following assumptions on the discretization spaces:

Assumption 4.22. *Let $q \in \mathbb{N}$ be such that the pair (V_h, X_h) has the following approximation properties:*

(i) *for all $u \in H^m(\Omega^-)$ where $m \geq 1$:*

$$\inf_{v_h \in V_h} \left(\|u - v_h\|_{L^2(\Omega^-)} + h \|\nabla u - \nabla v_h\|_{L^2(\Omega^-)} \right) \leq Ch^{\min(q+1, m)} \|u\|_{H^m(\Omega^-)},$$

(ii) *for all $\lambda \in H_{\text{pw}}^s(\Gamma)$ with $s \geq -1/2$:*

$$\inf_{\lambda_h \in X_h} \left(\|\lambda - \lambda_h\|_{H^{-1/2}(\Gamma)} \right) \leq Ch^{\min(q+1, s)+1/2} \|\lambda\|_{H_{\text{pw}}^s(\Gamma)}.$$

Remark 4.23. *For arbitrary $q > 1$, this assumption can be satisfied by choosing $V_h := \mathcal{S}^{q,1}(\mathcal{T}_h)$, $X_h := \mathcal{S}^{q-1,0}(\mathcal{T}_h^\Gamma)$ as defined in Section 2.4.*

Theorem 4.24. *Let $u_0 \in \text{dom}(\mathbf{H}^{p+1})$ where p denotes the order of the Runge-Kutta method employed. Assume that \mathcal{V} is sufficiently smooth, and assume that the approximation of the initial condition satisfies $\|u_0 - u^{k,h}(0)\|_{\mathcal{X}^0} \lesssim h^{q+1}$.*

Then the following estimate holds for $0 \leq nk \leq T$:

$$\|u^{k,h}(t_n) - u(t_n)\|_{L^2(\Omega^-)} + \|u_\star^{k,h}(t_n) - u(t_n)\|_{L^2(\Omega^+)} \leq C[u_0]T(k^p + h^q).$$

If Ω^- is convex or has a smooth boundary, the spatial rate can be improved to

$$\|u^{k,h}(t_n) - u(t_n)\|_{L^2(\Omega^-)} + \|u_\star^{k,h}(t_n) - u(t_n)\|_{L^2(\Omega^+)} \leq C[u_0]T(k^p + h^{q+1}). \quad (4.34)$$

Proof. We use the equivalent characterization of $(u^{k,h}, u_\star^{k,h})$ as the Runge-Kutta approximation $x^{k,h}(t_n)$ from Lemma 4.8. Additionally, we identify the exact solution with the pair $(u|_{\Omega^-}, u_\star)$ with $u_\star = u|_{\Omega^+}$ on Ω^+ and 0 on Ω^- , as was justified in Lemma 4.10. This gives:

$$\|u^{k,h}(t_n) - u(t_n)\|_{L^2(\Omega^-)} + \|u_\star^{k,h}(t_n) - u(t_n)\|_{L^2(\Omega^+)} \leq \|u(t_n) - x^{k,h}(t_n)\|_{\mathcal{X}^0}.$$

We insert the semi-discrete approximation $x^k(t_n)$ and estimate:

$$\|u(t_n) - x^{k,h}(t_n)\|_{\mathcal{X}^0} \leq \|u(t_n) - x^k(t_n)\|_{\mathcal{X}^0} + \|x^k(t_n) - x^{k,h}(t_n)\|_{\mathcal{X}^0}.$$

The first term is $\mathcal{O}(Tk^p)$ by Proposition 3.32. The second term can be estimated by Lemma 4.14, the best approximation property of the Ritz-projector in Lemma 4.15, and the assumption on the spaces of Assumption 4.22. The necessary regularity of the stages was already established in Lemma 4.21.

The improved estimate (4.34) follows along the same lines, using Lemma 4.19 in addition to Lemma 4.15. \square

Theorem 4.25. *Let $u_0 \in \text{dom}(\mathbf{H}^{p+2})$ where p denotes the order of the Runge-Kutta method employed. Assume that \mathcal{V} is sufficiently smooth and the Runge-Kutta method satisfies $|r(\infty)| < 1$. Assume that $\|u_0 - u^{k,h}(0)\|_{\mathcal{X}^1} \lesssim h^q$ and $\|\mathbf{H}u_0 - \mathbf{H}_h u^{k,h}(0)\|_{\mathcal{X}^0} \lesssim h^q$ (this is for example satisfied for the Ritz-projector). Define $\lambda^{h,k} := \mathbf{G}[\Lambda^{k,h}]$.*

Then the following estimates hold for $0 < nk \leq T$:

$$\begin{aligned} \|u^{k,h}(t_n) - u(t_n)\|_{H^1(\Omega^-)} + \|u_\star^{k,h}(t_n) - u(t_n)\|_{H^1(\Omega^+)} &\leq C[u_0]T(k^p + h^q), \\ \|\lambda^{h,k}(t_n) - \partial_\nu^+ u(t_n)\|_{H^{-1/2}(\Gamma)} &\leq C[u_0]T(k^p + h^q). \end{aligned}$$

Proof. We again split the error term into $\|u(t_n) - x^k(t_n)\|_{\mathcal{X}^1} + \|x^k(t_n) - x^{k,h}(t_n)\|_{\mathcal{X}^1}$. The error due to discretization in time can be estimated by $\|\mathbf{H}[u(t_n) - x^k(t_n)]\|_{\mathcal{X}^0}$ and $\|u(t_n) - x^{k,h}(t_n)\|_{\mathcal{X}^0}$ by integration by parts and the Cauchy-Schwarz inequality. Both of these terms are of order k^p by Proposition 3.32 (the operator \mathbf{H} commutes with the Runge-Kutta discretization). The H^1 -estimate in space follows from the \mathcal{X}^1 -estimate in Lemma 4.18 with the approximation properties in Lemma 4.15.

In order to see the estimate for the normal derivative, we first note that we can write

$$\begin{aligned} \left\| \lambda^{h,k}(t_n) - \partial_\nu^+ u(t_n) \right\|_{H^{-1/2}(\Gamma)} &= \left\| \left[\partial_\nu u_\star^{k,h}(t_n) \right] - \partial_\nu^+ u(t_n) \right\|_{H^{-1/2}(\Gamma)} \\ &\lesssim \left\| \mathbf{H}_h u_\star^{k,h}(t_n) - \mathbf{H}u(t_n) \right\|_{L^2(\mathbb{R}^d \setminus \Gamma)} + \left\| u_\star^{k,h}(t_n) - u(t_n) \right\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \end{aligned}$$

by using the trace theorem in $H_\Delta^1(\mathbb{R}^d \setminus \Gamma)$ and the fact that we can bound $-\Delta$ by \mathbf{H} and lower order terms. Thus what is left to do is estimate the convergence of the hamiltonians, i.e. of $\mathbf{H}_h x^{k,h} - \mathbf{H}u$. The proof works analogously to the proof of Theorem 4.24. Since \mathbf{H} commutes with Runge-Kutta approximation and the continuous time evolution, we can apply Proposition 3.32 to $\mathbf{H}u - \mathbf{H}u^k$ to get $\left\| \mathbf{H}u(t_n) - \mathbf{H}u^k(t_n) \right\|_{L^2(\mathbb{R}^d)} = \mathcal{O}(k^p)$. We have also already established the convergence of $\mathbf{H}_h x^{k,h} - \mathbf{H}x^k$ in Lemma 4.18. \square

If we drop the requirement $|r(\infty)| < 1$, we can still obtain the following result:

Theorem 4.26. *Let $u_0 \in \text{dom}(\mathbf{H}^{p+2})$, where p denotes the order of the Runge-Kutta method employed. Assume that \mathcal{V} is sufficiently smooth and Ω is convex or has smooth boundary, such that the assumptions of Lemma 4.19 are fulfilled. Define $\lambda^{h,k} := \mathbf{G}[\Lambda^{k,h}]$.*

Then the following estimate holds for $h \lesssim k^{1/2}$:

$$\left\| u^{k,h}(t_n) - u(t_n) \right\|_{H^1(\Omega^-)} + \left\| u_\star^{k,h}(t_n) - u(t_n) \right\|_{H^1(\Omega^+)} \leq C[u_0] T (k^p + h^q).$$

If $h \lesssim k$, we also get the full rate for the approximation of the normal derivative:

$$\left\| \lambda^{h,k}(t_n) - \partial_\nu^+ u(t_n) \right\|_{H^{-1/2}(\Gamma)} \leq C[u_0] T (k^p + h^q).$$

Proof. Follows along the same line as Theorem 4.25, except that instead of Lemma 4.18, we use Lemma 4.20. \square

For the case of an unbounded or non-smooth potential \mathcal{V} , we still get the rate $\mathcal{O}(k^p)$ in time (assuming $u_0 \in \text{dom}(\mathbf{H}^{p+2})$) and quasi-optimality in space (see Lemma 4.17), but the regularity of the semi-discrete stage vectors is less clear. As an exemplary result of what can be expected, we treat the case of “lowest order” FEM/BEM discretization.

Theorem 4.27. *Let $u_0 \in \text{dom}(\mathbf{H}^{p+1})$, where $p \geq 1$ denotes the order of the Runge-Kutta method employed, and let $V_h \times X_h$ satisfy Assumption 4.22 for $q \geq 1$. Then the following estimate holds for $0 \leq n k \leq T$:*

$$\left\| u^{k,h}(t_n) - u(t_n) \right\|_{L^2(\Omega^-)} + \left\| u_\star^{k,h}(t_n) - u(t_n) \right\|_{L^2(\Omega^+)} \leq C[u_0] T (k^p + h^1).$$

If Ω^- is convex or has a smooth boundary, the spatial rate can be improved to

$$\left\| u^{k,h}(t_n) - u(t_n) \right\|_{L^2(\Omega^-)} + \left\| u_\star^{k,h}(t_n) - u(t_n) \right\|_{L^2(\Omega^+)} \leq C[u_0] T (k^p + h^2). \quad (4.35)$$

Proof. Follows along the same lines as Theorem 4.24. We just note that the only place where the smoothness of \mathcal{V} was used, was in the estimates for $\|U^k\|_{H^s(\mathbb{R}^d)}$ and $\|\mathbf{H}U^k\|_{H^s(\mathbb{R}^d)}$. But for the case $s = 2$, this is not necessary, as for $u \in \text{dom}(\mathbf{H})$ we can bound using Lemma 4.6:

$$\begin{aligned} \|-\Delta u\|_{L^2(\mathbb{R}^d)} &\leq \|\mathbf{H}u\|_{L^2(\mathbb{R}^d)} + \|\mathcal{V}u\|_{L^2(\mathbb{R}^d)} \\ &\leq \|\mathbf{H}u\|_{L^2(\mathbb{R}^d)} + \frac{1}{2} \|-\Delta u\|_{L^2(\mathbb{R}^d)} + b \|u\|_{L^2(\mathbb{R}^d)} \end{aligned}$$

for some constant $b > 0$. Together with the general estimate

$$\|u\|_{H^2(\mathbb{R}^d)} \lesssim \|-\Delta u\|_{L^2(\mathbb{R}^d)} + \|u\|_{L^2(\mathbb{R}^d)}$$

this implies that we can bound both the H^2 -norm of the stages and of $\mathbf{H}U^k$ uniformly, as long as the initial condition is smooth enough. This in turn is sufficient to get the full convergence rate for lowest order discretization in space via Assumption 4.22. \square

4.5 Numerical results

In order to support our theoretical findings, we implemented the proposed method using a combination of the finite element software package NGSolve [Sch14] and the boundary element library BEM++ [Smi+15]. To compute the convolution quadrature contributions we used the algorithm presented in [Ban10], which avoids computing the convolution weights and instead replaces them with the trapezoidal discretization of Cauchy's integral formula computed on the fly. While this introduces an additional error that has not been accounted for in our analysis, this error can be reduced easily and should therefore not impact the convergence rate (cf. [BS09]).

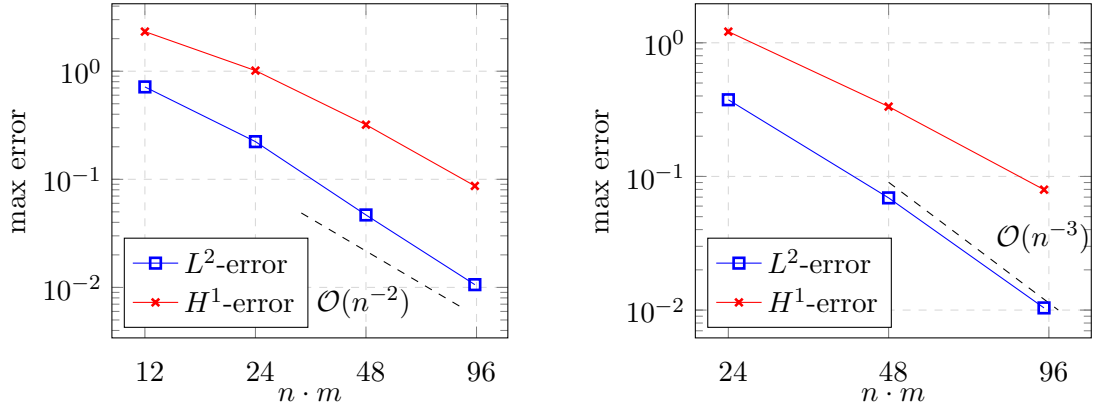
As a model problem we consider the free Schrödinger equation, i.e., $\mathcal{V} = 0$, in 3D. Following [Ant+08], given a point $x_c \in \mathbb{R}^3$ and a wave number $p_0 \in \mathbb{R}^3$, we set the initial condition as

$$u_0(x) := \sqrt[4]{\frac{2}{\pi}} e^{-|x-x_c|^2 + ip_0 \cdot (x-x_c)}.$$

For this initial condition, the exact solution is given by

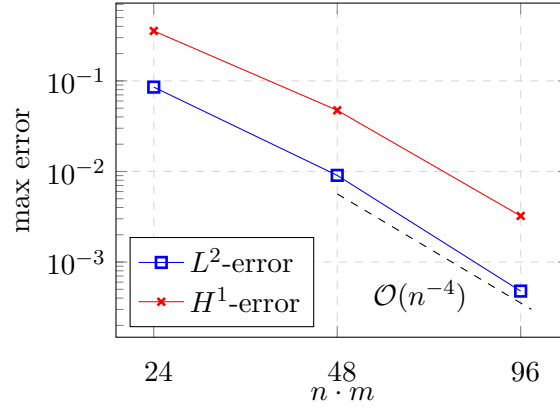
$$u_{ex}(x, t) = \sqrt[4]{\frac{2}{\pi}} \sqrt{\frac{i}{-4t+i}} \exp\left(\frac{-i|x-x_c|^2 - p_0 \cdot (x-x_c) + |p_0|^2 t}{-4t+i}\right).$$

For the computational domain we chose the cube $\Omega^- := [-4, 4]^3$. In order to be able to better distinguish convergence from artificial damping, we consider the case of two Gaussian beams. One is centered at $x_c^1 := (1, -1, 0)^T$ and has wave number $p_0^1 := (0, 0, 0)^T$, which means it will remain stationary and we expect to only see a dispersive effect. The second beam is centered at $(-1, 1, 0)^T$ and has a wave number $(1, 0, 0)^T$, which means the wave packet will travel out of the domain Ω^- . While this choice of initial conditions does not satisfy the assumption $\text{supp } u_0 \subset \Omega^-$, the decay of the functions is sufficiently fast that the truncation at the domain boundary becomes negligible.



(a) 1 stage Gauss (order 2)

(b) 2 stage RadauIIa (order 3)



(c) 3 stage RadauIIa (order 5)

Figure 4.1: Comparison of different Runge-Kutta methods for the Schrödinger equation

We look at the convergence rates for different Runge-Kutta methods, namely we compare a one-stage Gauss method to the 2 and 3 stage RadauIIa methods. For spatial discretization we use finite and boundary elements, such that the order is the same as for the time-stepping, i.e. we use $q = p$ for the FEM and $q = p - 1$ for the BEM part. We compare the maximum error for the FEM part in the L^2 - and the H^1 -norm, i.e., we compute $\max_{n=0,\dots,N} \left\| u^{k,h}(t_n) - u(t_n) \right\|_{L^2(\Omega)}$ and $\max_{n=0,\dots,N} \left\| u^{k,h}(t_n) - u(t_n) \right\|_{H^1(\Omega)}$. Figure 4.1 collects the results. For the Gauss and two-stage Radau method, we see the full classical order of 2 and 3 respectively as predicted. For the higher order RadauIIa method, we do not see the predicted rate, but this might just be pre-asymptotic behavior due to the limited number of timesteps. Nevertheless the high order method offers a beneficial ratio of number of operations to achieved accuracy.

5 CQ for the wave equation with nonlinear impedance boundary condition

In this chapter, we investigate whether boundary integral methods and Convolution Quadrature can be used when discretizing nonlinear problems. In order to keep the level of technicalities to a minimum, we only consider a rather simple model problem and restrict our methods to the case of BDF1 and BDF2 based Convolution Quadrature. The problem consists of the linear wave equation complemented with a nonlinear boundary condition. The motivation for the problem under consideration comes from nonlinear acoustic boundary conditions, as investigated in [LT93; Gra12], and boundary conditions in electromagnetism obtained by asymptotic approximation of thin layers of nonlinear materials [HJ02]. Another source of interesting nonlinear boundary conditions is [Ayg+04], which investigates the coupling of wave propagation with nonlinear circuits. Compared with these references, the boundary condition used in this chapter is simple, but already contains enough difficulties to warrant development of new tools of analysis, which hopefully can be extended to more involved situations. Many of the results of this section have appeared in [BR17] and are the result of collaboration with Lehel Banjai. This model problem was first suggested in [Ban15] and it has recently also been investigated using different techniques, namely a positivity preservation property of the underlying boundary integral operators [BL17], wherein the authors focus on a Runge-Kutta based discretization instead of multistep methods.

5.1 Model problem

We consider an exterior scattering problem, using the linear wave equation and a nonlinear boundary condition. Let $\Omega^- \subseteq \mathbb{R}^d$ be a bounded Lipschitz domain, and let $\Omega^+ := \mathbb{R}^d \setminus \overline{\Omega^-}$ be the exterior. We denote the boundary by $\Gamma := \partial\Omega^-$. Given a function $g : \mathbb{R} \rightarrow \mathbb{R}$ and a constant wave speed $c > 0$, we consider the model problem:

$$\frac{1}{c^2} \ddot{u}^{\text{tot}} = \Delta u^{\text{tot}} \quad \text{in } \Omega^+, \quad (5.1a)$$

$$\partial_n^+ u^{\text{tot}} = g(\dot{u}^{\text{tot}}) \quad \text{on } \Gamma, \quad (5.1b)$$

and use the initial condition $u^{\text{tot}}(t) = u^{\text{inc}}(t)$ for $t \leq 0$, where the incident wave u^{inc} itself satisfies the wave equation

$$\frac{1}{c^2} \ddot{u}^{\text{inc}} = \Delta u^{\text{inc}} \quad \text{in } \Omega^+.$$

We assume that at time $t = 0$ the incident wave has not reached the scatterer yet, i.e., we assume that for $t \leq 0$, $u^{\text{inc}}(\cdot, t)$ vanishes in a neighborhood of Ω^- . For notational convenience, we set $u^{\text{inc}}(x, t) := 0$ for $x \in \Omega^-$ and $\forall t \in \mathbb{R}$.

We make the following assumptions on the nonlinearity g :

Assumption 5.1. (i) $g \in C^1(\mathbb{R})$,

(ii) $g(0) = 0$,

(iii) $g(\mu)\mu \geq 0 \quad \forall \mu \in \mathbb{R}$,

(iv) $g'(\mu) \geq 0 \quad \forall \mu \in \mathbb{R}$,

(v) g satisfies the growth condition $|g(\mu)| \leq C(1 + |\mu|^p)$, where

$$\begin{cases} 1 < p < \infty & d = 2, \\ 1 < p \leq \frac{d}{d-2} & d \geq 3. \end{cases}$$

(vi) g is strictly monotone, i.e., there exists $\beta > 0$, such that

$$(g(\lambda) - g(\mu))(\lambda - \mu) \geq \beta |\lambda - \mu|^2 \quad \forall \lambda, \mu \in \mathbb{R}. \quad (5.2)$$

Remark 5.2. The growth condition (v) is chosen in a way to ensure that the operator $\eta \mapsto g(\eta)$ is bounded, i.e., the estimate $\|g(u)\|_{H^{-1/2}(\Omega)} \leq C \left(1 + \|u\|_{H^{1/2}(\Gamma)}^p\right)$ holds. This will be proved in Lemma 5.7. ■

Remark 5.3. We will show well-posedness of (5.1) under these conditions on g as part of Theorem 5.12. The well posedness of such problems has already been established with more general boundary conditions but under slightly stronger growth conditions in [LT93] and [Gra12]. ■

5.2 Discretization using boundary integrals

In order to discretize the problem using boundary integral techniques, it is convenient to work with homogeneous initial conditions. We therefore decompose the solution using the ansatz $u^{\text{tot}} = u^{\text{inc}} + u^{\text{scat}}$. Since u^{inc} satisfies the wave equation and $u^{\text{tot}}(t) = u^{\text{inc}}(t)$ for $t \leq 0$, the scattered field u^{scat} solves the following problem:

$$\frac{1}{c^2} \ddot{u}^{\text{scat}} = \Delta u^{\text{scat}} \quad \text{in } \Omega^+, \quad (5.3)$$

$$\partial_n^+ u^{\text{scat}} = g(\dot{u}^{\text{scat}} + \dot{u}^{\text{inc}}) - \partial_n^+ u^{\text{inc}} \quad \text{on } \Gamma, \quad (5.4)$$

$$u^{\text{scat}}(t) = 0 \quad \text{in } \mathbb{R}^d \text{ for all } t \leq 0. \quad (5.5)$$

In order to simplify the notation, we make the assumption $c = 1$.

To derive the boundary integral equations and their discretization, we will work mostly formally as we will not need the precise statements for the analysis. Since u^{scat} solves the wave equation with homogeneous initial conditions, we can apply Kirchhoff's representation formula (3.12) and write it as

$$u^{\text{scat}} = -S(\partial_t) \partial_n^+ u^{\text{scat}} + D(\partial_t) \gamma^+ u^{\text{scat}}.$$

We define the Calderón operators for $s \in \mathbb{C}_+$ as

$$B(s) := \begin{pmatrix} sV(s) & K(s) \\ -K^T(s) & s^{-1}W(s) \end{pmatrix}, \quad (5.6a)$$

$$B_{\text{imp}}(s) := B(s) + \begin{pmatrix} 0 & -\frac{1}{2}\text{Id} \\ \frac{1}{2}\text{Id} & 0 \end{pmatrix}. \quad (5.6b)$$

Taking the traces in the representation formula and inserting the boundary condition then gives an equivalence between the boundary integral equation

$$B_{\text{imp}}(\partial_t) \begin{pmatrix} \varphi \\ \psi \end{pmatrix} + \begin{pmatrix} 0 \\ g(\psi + \dot{u}^{\text{inc}}) \end{pmatrix} = \begin{pmatrix} 0 \\ -\partial_n^+ u^{\text{inc}} \end{pmatrix} \quad (5.7)$$

and the scattering problem (5.3), namely:

- (i) If u^{scat} solves (5.3), then (φ, ψ) , with $\varphi := -\partial_n^+ u^{\text{scat}}$ and $\psi := \gamma^+ \dot{u}^{\text{scat}}$, solves (5.7).
- (ii) If (φ, ψ) solves (5.7), then $u^{\text{scat}} := S(\partial_t)\varphi + \partial_t^{-1}D(\partial_t)\psi$ solves (5.3).

Remark 5.4. *We keep this equivalence statement purely formal and without detailed assumptions as we will not make use of the continuous boundary integral equations. Instead, we will prove a discrete version of this equivalence principle in Lemma 5.11.*

For discretization, let $X_h \subseteq H^{-1/2}(\Gamma)$ and $Y_h \subseteq H^{1/2}(\Gamma)$ be closed (not necessarily finite dimensional) subspaces and let $J_\Gamma^{Y_h} : H^{1/2}(\Gamma) \rightarrow Y_h$ be a stable linear operator with “good” approximation properties, see Proposition 2.36 for possible constructions. The detailed approximation requirements for the projection operator and the discrete spaces can be found in Assumption 5.32 or Lemma 5.35 respectively; in practice, we used a simple L^2 -projection (see [CT87; BY14] for sufficient conditions on the stability of the L^2 -projection).

In order to discretize in time, we use a multistep based Convolution Quadrature with step size $k > 0$, see Section 3.1, based on either the BDF1 or BDF2 method. We denote the corresponding coefficients by $(\alpha_j)_{j=0}^m$ (for BDF methods we have $\beta_0 = 1$ and $\beta_j = 0 \forall j > 0$, thus we drop these coefficients).

The discretized version of (5.7) then reads:

Problem 5.5. *Find functions φ^k and ψ^k , such that $\varphi^k(t) \in X_h$, $\psi^k(t) \in Y_h$ and*

$$\left\langle B_{\text{imp}}(\partial_t^k) \begin{pmatrix} \varphi^k \\ \psi^k \end{pmatrix}, \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle_\Gamma + \left\langle g(\psi^k + J_\Gamma^{Y_h} \dot{u}^{\text{inc}}), \eta \right\rangle_\Gamma = \left\langle -\partial_n^+ u^{\text{inc}}(t_n), \eta \right\rangle_\Gamma \quad (5.8)$$

for all $(\xi, \eta) \in X_h \times Y_h$ (equality to be understood as a function in $t \in [0, T]$). \square

Remark 5.6. *As always, we will only need to compute the solution to (5.8) for discrete times $t_n := nk$, using continuous time just simplifies the notation.*

Since we will be dealing with pairs $(\varphi, \psi) \in H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma)$ on a regular basis, we introduce the norm on the product space

$$\|(\varphi, \psi)\|^2 := \|\varphi\|_{H^{-1/2}(\Gamma)}^2 + \|\psi\|_{H^{1/2}(\Gamma)}^2. \quad (5.9)$$

5.3 Well posedness of the discretization scheme

We start by investigating under which conditions the discrete Problem 5.5 has a unique solution. In order to do so, we start with some basic properties of the operator induced by g and the operator $B_{\text{imp}}(s)$.

Lemma 5.7. *The operator $g : H^{1/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma)$, $\eta \mapsto \langle g(\eta), \cdot \rangle_{\Gamma}$ has the following properties:*

(i) *The operator g is a bounded (nonlinear) operator, satisfying*

$$\|g(\eta)\|_{H^{-1/2}(\Gamma)} \leq C \left(1 + \|\eta\|_{H^{1/2}(\Gamma)}^p\right),$$

where p is the bound from Assumption 5.1(v), and the constant $C > 0$ depends on Γ and g .

(ii) *The operator g is monotone in the following sense:*

$$\langle g(\eta_1) - g(\eta_2), \eta_1 - \eta_2 \rangle_{\Gamma} \geq 0 \quad \text{for all (real valued) } \eta_1, \eta_2 \in H^{1/2}(\Gamma).$$

(iii) *The map $\eta \mapsto g(\eta)$ is continuous as a map $H^{1/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma)$.*

Proof. Ad (i): We recall that the following Sobolev embeddings hold (Proposition 2.29)

$$H^{1/2}(\Gamma) \subseteq L^{p'}(\Gamma) \quad \begin{cases} \forall 1 \leq p' < \infty & \text{for } d = 2, \\ \forall 1 \leq p' \leq \frac{2d-2}{d-2} & \text{for } d \geq 3. \end{cases} \quad (5.10)$$

Using p from Assumption 5.1 (v), we fix p', q' such that $1/p' + 1/q' = 1$ and both p' and pq' are in the admissible range of the Sobolev embedding. The case $d = 2$ is clear. For $d \geq 3$ we use $p' = \frac{2d-2}{d-2}$ and $q' := \frac{2d-2}{d}$.

For $\eta, \xi \in H^{1/2}(\Gamma)$ we calculate:

$$\begin{aligned} \int_{\Gamma} g(\eta) \bar{\xi} &\leq \|g(\eta)\|_{L^{q'}(\Gamma)} \|\xi\|_{L^{p'}(\Gamma)} \lesssim \left(1 + \|\eta\|_{L^{q'p}(\Gamma)}^p\right) \|\xi\|_{H^{1/2}(\Gamma)} \\ &\lesssim \left(1 + \|\eta\|_{H^{1/2}(\Gamma)}^p\right) \|\xi\|_{H^{1/2}(\Gamma)}. \end{aligned}$$

Ad (ii): Given $\eta_1, \eta_2 \in H^{1/2}(\Gamma)$, we apply the mean value theorem to get:

$$\langle g(\eta_1) - g(\eta_2), \eta_1 - \eta_2 \rangle_{\Gamma} = \int_{\Gamma} g'(s(x)) (\eta_1(x) - \eta_2(x))^2 dx.$$

Since $g' \geq 0$ by Assumption 5.1(iv), the integral is non-negative.

Ad (iii): We take a sequence $\eta_h \rightarrow \eta$ in $H^{1/2}(\Gamma)$ and show $g(\eta_h) \rightarrow g(\eta)$ in $H^{-1/2}(\Gamma)$. We focus on the case $d \geq 3$, the case $d = 2$ is even simpler because all the Sobolev embeddings hold for arbitrary L^p -spaces. By the Sobolev embedding (Proposition 2.29, c.f. (5.10)), we can estimate:

$$\begin{aligned} \|g(\eta) - g(\eta_h)\|_{H^{-1/2}(\Gamma)} &= \sup_{0 \neq v \in H^{1/2}(\Gamma)} \frac{\langle g(\eta) - g(\eta_h), v \rangle_{\Gamma}}{\|v\|_{H^{1/2}(\Gamma)}} \\ &\lesssim \|g(\eta) - g(\eta_h)\|_{L^{q'}(\Gamma)}, \end{aligned}$$

with $q' := (2d - 2)/d$.

If $\eta_h \rightarrow \eta$ in $H^{1/2}(\Gamma)$, the Sobolev-embeddings give $\eta_h \rightarrow \eta$ in $L^q(\Gamma)$ for $q \leq \frac{2d-2}{d-2}$. By [Bre83, Theorem IV.9] (see also the proof of [Rud87, Theorem 3.11]), this implies that there exists a sub-sequence η_{h_j} , which converges pointwise almost everywhere and there exists a function $\zeta \in L^q(\Gamma)$ such that $|\eta_h| \leq \zeta$ almost everywhere.

From the growth conditions on g , we get that $|g(\eta_{h_j})| \leq C(1 + |\zeta|^p)$ for all $j \in \mathbb{N}$. The same calculation from the proof of part (i) gives that $(1 + |\zeta|^p) \in L^{q'}(\Gamma)$, thus we have an integrable upper bound.

Since g is continuous by assumption, $g(\eta_{h_j})$ converges to $g(\eta)$ pointwise almost everywhere. By the dominated convergence theorem this implies $\int_{\Gamma} |g(\eta_{h_j}) - g(\eta)|^{q'} \rightarrow 0$. The same argument can be applied to show that every sub-sequence of $g(\eta_h)$ has a sub-sequence that converges to $g(\eta)$ in $H^{-1/2}(\Gamma)$. This is sufficient to show that the whole sequence converges. \square

The operator $B_{\text{imp}}(s)$ is elliptic in the frequency domain:

Lemma 5.8. *There exists a constant $\beta > 0$ depending only on Γ , such that*

$$\operatorname{Re} \left\langle B_{\text{imp}}(s) \begin{pmatrix} \varphi \\ \psi \end{pmatrix}, \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \right\rangle_{\Gamma} \geq \beta \min(1, |s|^2) \frac{\operatorname{Re}(s)}{|s|^2} \|\varphi, \psi\|^2. \quad (5.11)$$

Proof. An analogous estimate to (5.11) for the operator $B(s)$ was shown in [BLS15b, Lemma 3.1]. Estimate (5.11) then follows directly since the difference $B_{\text{imp}}(s) - B(s)$ is skew-hermitean. \square

In order to prove the solvability of the nonlinear discrete problem (5.8), we need the following tool from functional analysis:

Proposition 5.9 (Browder and Minty, [Sho97, Chapter II, Theorem 2.2]). *Let \mathcal{X} be a real, separable and reflexive Banach space and let $A : \mathcal{X} \rightarrow \mathcal{X}'$ be a (not necessarily linear) bounded, continuous, coercive and monotone map from \mathcal{X} to its dual space. In other words, let A satisfy:*

- (i) $A : \mathcal{X} \rightarrow \mathcal{X}'$ is continuous,
- (ii) the set $A(M)$ is bounded in \mathcal{X}' for all bounded sets $M \subseteq \mathcal{X}$,
- (iii) $\lim_{\|u\|_{\mathcal{X}} \rightarrow \infty} \frac{\langle A(u), u \rangle_{\mathcal{X}' \times \mathcal{X}}}{\|u\|_{\mathcal{X}}} = \infty$,
- (iv) $\langle A(u) - A(v), u - v \rangle_{\mathcal{X}' \times \mathcal{X}} \geq 0$ for all $u, v \in \mathcal{X}$.

Then the variational equation

$$\langle A(u), v \rangle_{\mathcal{X}' \times \mathcal{X}} = \langle f, v \rangle_{\mathcal{X}' \times \mathcal{X}} \quad \forall v \in \mathcal{X},$$

has at least one solution for all $f \in \mathcal{X}'$. If the operator is strongly monotone, i.e., if there exists a constant $\beta > 0$ such that

$$\langle A(u) - A(v), u - v \rangle_{\mathcal{X}' \times \mathcal{X}} \geq \beta \|u - v\|_{\mathcal{X}}^2 \quad \text{for all } u, v \in \mathcal{X},$$

then the solution is unique.

Proof. The existence statement is just a slight reformulation of [Sho97, Theorem 2.2], based on some of the equivalences stated in the same chapter. Uniqueness can be shown by considering two solutions u and v and applying the strong monotonicity to conclude $\|u - v\|_{\mathcal{X}} = 0$. \square

These preparatory results allow us to now prove that indeed, Problem 5.5 is well posed.

Theorem 5.10. *Let $k > 0$ and $(X_h, Y_h) \subseteq H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma)$ be closed subspaces. Then the discrete system of equations (5.8) has a unique solution in the space $X_h \times Y_h$ for all times $t \in \mathbb{R}_+$.*

Proof. We show the result for discrete times $t_n = nk$, the general case follows by considering shifted initial conditions. We remark that due to the symmetry properties of the boundary integral operators, namely $\overline{B(s)} = B(\overline{s})$, the convolution weights are real-valued and we can restrict all our considerations on real valued functions. We prove this by induction on n . For $n = 0$ we are given the initial condition $\varphi^k = \psi^k = 0$. Now assume that we have solved (5.8) up to the $n - 1$ st step. We denote the operators from the definition of $B_{\text{imp}}(\partial_t^k)$ as B_j , $j \in \mathbb{N}_0$, dropping the subscript and set $\tilde{\psi} := \psi^k + J_{\Gamma}^{Y_h} \dot{u}^{\text{inc}}$ and bring all known terms to the right-hand side. Then, in the n -th step the equation reads

$$\left\langle B_0 \begin{pmatrix} \varphi^k(t_n) \\ \tilde{\psi}(t_n) \end{pmatrix}, \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle_{\Gamma} + \left\langle g(\tilde{\psi}(t_n)), \eta \right\rangle_{\Gamma} = \left\langle f^n, \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle_{\Gamma}, \quad (5.12)$$

$$\text{with } f^n := \begin{pmatrix} 0 \\ -\partial_n^+ u^{\text{inc}}(t_n) \end{pmatrix} - \sum_{j=0}^{n-1} B_{n-j} \begin{pmatrix} \varphi^k(t_j) \\ \psi^k(t_j) \end{pmatrix} + B_0 \begin{pmatrix} 0 \\ J_{\Gamma}^{Y_h} \dot{u}^{\text{inc}}(t_n) \end{pmatrix}.$$

The right-hand side is a continuous linear functional with respect to (ξ, η) due to the mapping properties of the operators B_j that are easily transferred from the frequency-domain versions (2.34); see [Lub94].

We want to apply Proposition 5.9, we note that the operator $B_0 : H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma) \rightarrow H^{1/2}(\Gamma) \times H^{-1/2}(\Gamma)$ is the leading term of a power series centered at zero. This means we can compute $B_0 = B_{\text{imp}}\left(\frac{\delta(0)}{k}\right)$. Lemma 5.8 and Proposition 3.14 therefore imply B_0 is elliptic. The non-linearity satisfies: $\langle g(\eta), \eta \rangle_{\Gamma} = \int_{\Gamma} g(\eta)\eta \geq 0$ by Assumption 5.1 (iii). This implies that the left-hand side in (5.12) is coercive.

Since B_0 is linear and elliptic and g is monotone via Lemma 5.7(ii), the left-hand side in (5.12) is strongly monotone. Boundedness has been proven in Lemma 5.7. The continuity is a consequence of the boundedness of B_0 and the continuity from Lemma 5.7(iii). \square

5.4 Convergence analysis

In this section, we will analyze the convergence behavior of the discrete solutions to the exact ones. This is done by making use of a correspondence between the convolution quadrature based method and the multistep approximation of a related semigroup. We start with showing that the CQ-approximations can be written as the solution of a sequence of multistep type problems.

Lemma 5.11. *Let $k > 0$, and let $X_h \subseteq H^{-1/2}(\Gamma)$, $Y_h \subseteq H^{1/2}(\Gamma)$ be closed subspaces. Define*

$$\mathcal{H}_h := \left\{ u \in H^1(\mathbb{R}^d \setminus \Gamma) : \llbracket \gamma u \rrbracket \in Y_h, \gamma^- u \in X_h^\circ \right\}. \quad (5.13)$$

For $n \in \mathbb{N}$, let $v_{ie}^{h,k}(t_n) \in \mathcal{H}_h$ and $\mathbf{w}_{ie}^{h,k}(t_n) \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)$ solve

$$\frac{1}{k} \sum_{j=0}^m \alpha_j \mathbf{w}_{ie}^{h,k}(t_{n-j}) = \nabla v_{ie}^{h,k}(t_n), \quad (5.14a)$$

$$\frac{1}{k} \sum_{j=0}^m \alpha_j v_{ie}^{h,k}(t_{n-j}) = \operatorname{div} \left(\mathbf{w}_{ie}^{h,k}(t_n) \right), \quad (5.14b)$$

$$\llbracket \gamma_\nu \mathbf{w}_{ie}^{h,k}(t_n) \rrbracket \in X_h, \quad (5.14c)$$

$$\gamma_\nu^+ \mathbf{w}_{ie}^{h,k}(t_n) - g \left(\llbracket \gamma v_{ie}^{h,k}(t_n) \rrbracket + J_\Gamma^{Y_h} \dot{u}^{inc}(t_n) \right) + \partial_n^+ \dot{u}^{inc}(t_n) \in X_h^\circ, \quad (5.14d)$$

where $t_n := nk$ and $\mathbf{w}_{ie}^{h,k}(t) = v_{ie}^{h,k}(t) := 0$ for $t \leq 0$. Then the following two statements hold:

(i) If φ^k and ψ^k solve (5.8) and we define $u_{ie}^{h,k} := S(\partial_t^k) \varphi^k + (\partial_t^k)^{-1} D(\partial_t^k) \psi^k$, then $\mathbf{w}_{ie}^{h,k} := \nabla u_{ie}^{h,k}$ and $v_{ie}^{h,k} := \partial_t^k u_{ie}^{h,k}$ solve (5.14).

(ii) If $\mathbf{w}_{ie}^{h,k}$ and $v_{ie}^{h,k}$ solve (5.14), the jumps $\varphi^k := -\llbracket \gamma_\nu \mathbf{w}_{ie}^{h,k} \rrbracket$, $\psi^k := \llbracket \gamma v_{ie}^{h,k} \rrbracket$ solve (5.8).

Note: the subindex “ie”, which stands for “integral equations”, is used to separate this sequence from the one obtained by applying the multistep method to the semigroup, as defined in (5.25).

Proof. We first note that (5.14) has a solution in \mathcal{H}_h .

We show this by induction on n . For $n \leq 0$ we set $u_{ie}^{h,k}(t_n) := v_{ie}^{h,k}(t_n) := 0$. For $n \in \mathbb{N}$, we consider the weak formulation, find $\mathbf{w}_{ie}^{h,k}(t_n) \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)$, $v_{ie}^{h,k}(t_n) \in \mathcal{H}_h$, such that

$$\frac{1}{k} \sum_{j=0}^m \alpha_j \mathbf{w}_{ie}^{h,k}(t_{n-j}) = \nabla v_{ie}^{h,k}(t_n), \quad (5.15a)$$

$$\begin{aligned} \left(\frac{1}{k} \sum_{j=0}^m \alpha_j v_{ie}^{h,k}(t_{n-j}), z_h \right)_{L^2(\mathbb{R}^d)} &= - \left(\mathbf{w}_{ie}^{h,k}(t_n), \nabla z_h \right)_{L^2(\mathbb{R}^d)} \\ &\quad - \left\langle g \left(\llbracket \gamma v_{ie}^{h,k} \rrbracket(t_n) + J_\Gamma^{Y_h} \dot{u}^{inc}(t_n) \right) - \partial_n^+ \dot{u}^{inc}(t_n), \llbracket \gamma z_h \rrbracket \right\rangle_\Gamma \end{aligned} \quad (5.15b)$$

for all $z_h \in \mathcal{H}_h$. Multiplying the first equation by k and collecting all the terms involving $\mathbf{w}_{ie}^{h,k}(t_j)$ for $j < n$ in $F_n \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)$, the condition becomes $\alpha_0 \mathbf{w}_{ie}^{h,k}(t_n) =$

$k\nabla v_{\text{ie}}^{h,k}(t_n) + F_n$. After inserting this identity into (5.15b) and combining all known terms into a new right-hand side $\tilde{F}^n \in \mathcal{H}'_h$, the second equation becomes

$$\frac{\alpha_0}{k} \left(v_{\text{ie}}^{h,k}(t_n), z_h \right)_{L^2(\mathbb{R}^d)} + \frac{k}{\alpha_0} \left(\nabla v_{\text{ie}}^{h,k}(t_n), \nabla z_h \right)_{L^2(\mathbb{R}^d)} + \left\langle g \left(\left[\gamma v_{\text{ie}}^{h,k}(t_n) \right] + J_{\Gamma}^{Y_h} \dot{u}^{inc}(t_n) \right), \left[\gamma z_h \right] \right\rangle_{\Gamma} = \left\langle \tilde{F}^n, z_h \right\rangle_{\mathcal{H}'_h \times \mathcal{H}_h}.$$

Since \mathcal{H}_h is a closed subspace of $H^1(\mathbb{R}^d \setminus \Gamma)$, this equation can be solved for all $n \in \mathbb{N}$ due to the monotonicity of the operators involved and the Browder-Minty theorem; see Proposition 5.9 and also the proof of Theorem 5.10 for how to treat the nonlinearity. Defining $\alpha_0 \mathbf{w}_{\text{ie}}^{h,k}(t_n) := k\nabla v_{\text{ie}}^{h,k}(t_n) + F_n$, we have found a solution to (5.15).

What still needs to be shown is that $\left[\gamma_{\nu} \mathbf{w}_{\text{ie}}^{h,k}(t_n) \right] \in X_h$. Note that it is sufficient to show $\left[\gamma_{\nu} \mathcal{Z}[\mathbf{w}_{\text{ie}}^{h,k}](z) \right] \in X_h$ for the Z-transformed variable, as we can then express $\left[\gamma_{\nu} \mathbf{w}_{\text{ie}}^{h,k}(t_n) \right]$ as a Cauchy integral in X_h . The details of this argument are given later.

It is easy to see that

$$\left\| \begin{pmatrix} \mathbf{w}_{\text{ie}}^{h,k}(t_n) \\ v_{\text{ie}}^{h,k}(t_n) \end{pmatrix} \right\|_{H(\text{div}, \mathbb{R}^d \setminus \Gamma) \times H^1(\mathbb{R}^d \setminus \Gamma)} \leq C(k) \sum_{j=0}^{n-1} \left\| \begin{pmatrix} \mathbf{w}_{\text{ie}}^{h,k}(t_j) \\ v_{\text{ie}}^{h,k}(t_j) \end{pmatrix} \right\|_{H(\text{div}, \mathbb{R}^d \setminus \Gamma) \times H^1(\mathbb{R}^d \setminus \Gamma)},$$

where the constant may depend on k , but not on $v_{\text{ie}}^{h,k}$, $\mathbf{w}_{\text{ie}}^{h,k}$ or n . This implies that the Z-transform $\widehat{\mathbf{w}}(z)$ is well defined for $|z|$ sufficiently small.

To simplify notation, define the function $G := g\left(\left[\gamma v_{\text{ie}}^{h,k}\right] + J_{\Gamma}^{Y_h} \dot{u}^{inc}\right) - \partial_n u^{inc}$. Taking the Z-transforms $\widehat{\mathbf{w}} := \mathcal{Z}\left[\mathbf{w}_{\text{ie}}^{h,k}\right]$ and $\widehat{v} := \mathcal{Z}\left[v_{\text{ie}}^{h,k}\right]$, a simple calculation shows that $\frac{\delta(z)}{k} \widehat{\mathbf{w}}(z) = \nabla \widehat{v}(z)$ and for $z_h \in \mathcal{H}_h$

$$\left(\left(\frac{\delta(z)}{k} \right)^2 \widehat{v}, z_h \right)_{L^2(\mathbb{R}^d)} + (\nabla \widehat{v}, \nabla z_h)_{L^2(\mathbb{R}^d)} + \left\langle \frac{\delta(z)}{k} \widehat{G}, \left[\gamma z_h \right] \right\rangle_{\Gamma} = 0.$$

For $z_h \in C_0^{\infty}(\mathbb{R}^d \setminus \Gamma)$, this implies

$$-\Delta \widehat{v}(z) + \left(\frac{\delta(z)}{k} \right)^2 \widehat{v}(z) = 0.$$

From $z_h \in \mathcal{H}_h$ with $z_h|_{\Omega^-} = 0$ we see $\partial_n^+ \widehat{v} - k^{-1} \delta(z) \widehat{G} \in Y_h^{\circ}$. Let $\xi \in X_h^{\circ}$ and choose z_h as a lifting of ξ to $H^1(\mathbb{R}^d \setminus \Gamma)$, i.e., $\gamma^+ z_h = \gamma^- z_h = \xi$. Then we get by integration by parts:

$$\langle -\partial_n^+ \widehat{v}, \xi \rangle_{\Gamma} + \langle \partial_n^- \widehat{v}, \xi \rangle_{\Gamma} = 0,$$

or $\left[\partial_{\nu} \widehat{v} \right] \in (X_h^{\circ})^{\circ} = X_h$, which in turn implies $\left[\gamma \widehat{\mathbf{w}} \right](z) = k \delta(z)^{-1} \left[\partial_{\nu} \widehat{v}(z) \right] \in X_h$. We can use the Cauchy-integral formula to write:

$$\left[\gamma_{\nu} \mathbf{w}_{\text{ie}}^{h,k} \right](t_n) = \frac{1}{2\pi i} \int_{\mathcal{C}} \left[\gamma_{\nu} \widehat{\mathbf{w}}(z) \right] z^{-n-1} dz,$$

where the contour $\mathcal{C} := \{z \in \mathbb{C} : |z| = \text{const}\}$ denotes a sufficiently small circle, such that all the Z -transforms exist. Since we have shown that $[[\gamma_\nu \widehat{\mathbf{w}}]] \in X_h$ and we assumed that X_h is a closed space, this implies $[[\gamma_\nu \mathbf{w}_{\text{ie}}^{h,k}(t_n)]] \in X_h$. Thus, we have shown the existence of a solution to (5.14).

We can now show the equivalence of (i) and (ii). We start by showing that the traces of the solutions to (5.14) solve the boundary integral equation. We have the following equation in the frequency domain:

$$-\Delta \widehat{v}(z) + \left(\frac{\delta(z)}{k}\right)^2 \widehat{v}(z) = 0,$$

$$\partial_n^+ \widehat{v}(z) - \frac{\delta(z)}{k} \widehat{G}(z) \in Y_h^\circ.$$

For simpler notation, we define $s_k := \frac{\delta(z)}{k}$. The representation formula then tells us that we can write $\widehat{v}(z) = -S(s_k) [[\gamma_\nu \widehat{v}(z)]] + D(s_k) [[\gamma \widehat{v}(z)]]$. Introducing the new traces

$$\widetilde{\psi}(z) := [[\gamma \widehat{v}(z)]] \quad \text{and} \quad \widetilde{\varphi}(z) := -\frac{1}{s_k} [[\partial_\nu \widehat{v}(z)]] = [[\gamma_\nu \widehat{w}(z)]]$$

and inserting these definitions into the representation formula gives

$$\widehat{v}(z) = s_k S(z) \widetilde{\varphi}(z) + D(z) \widetilde{\psi}(z).$$

Taking the interior trace γ^- and testing with a discrete function $\xi_h \in X_h$ gives

$$0 = \langle \gamma^- \widehat{v}, \xi_h \rangle_\Gamma = \langle s_k V(s_k) \widetilde{\varphi}, \xi_h \rangle_\Gamma + \left\langle (K(s_k) - 1/2) \widetilde{\psi}, \xi_h \right\rangle_\Gamma.$$

Analogously, by starting from the representation formula multiplied by s_k^{-1} , taking the exterior normal derivative ∂_n^+ , and testing with $\eta_h \in Y_h$, we obtain that

$$\left\langle \widehat{G}, \eta_h \right\rangle_\Gamma = \left\langle \frac{1}{s_k} \partial_n^+ \widehat{v}, \eta_h \right\rangle_\Gamma = \langle (1/2 - K^T(s_k)) \widetilde{\varphi}, \eta_h \rangle_\Gamma + \left\langle \frac{1}{s_k} W(s_k) \widetilde{\psi}, \eta_h \right\rangle_\Gamma.$$

Together, this is just the Z -transform of (5.8). By taking the inverse Z -transform, we conclude that the traces $[[\gamma v_{\text{ie}}^{h,k}]]$ and $[[\gamma_\nu \mathbf{w}_{\text{ie}}^{h,k}]]$ solve (5.8). By the uniqueness of the solution via Theorem 5.10, this implies $\varphi^k = -[[\gamma_\nu u_{\text{ie}}^{h,k}]]$ and $\psi^k = [[\gamma v_{\text{ie}}^{h,k}]]$, which then shows (ii).

For (i), we observe that due to the uniqueness of solutions to Helmholtz transmission problems $(\mathbf{w}_{\text{ie}}^{h,k}, v_{\text{ie}}^{h,k})$ defined via (5.14) and $(\mathbf{w}_{\text{ie}}^{h,k}, v_{\text{ie}}^{h,k})$ defined via potentials have the same Z -transform and therefore coincide also in the time domain. \square

5.4.1 A related semigroup - semi-discretization in space

In order to analyze (5.1) and its discretization by solving Problem 5.5, we reformulate the problem using a nonlinear semigroup framework as introduced in Section 2.2.2. To do so,

we switch to a first order formulation of (5.1) by introducing the new variables $v := \dot{u}$ and $\mathbf{w} := \nabla u$ to get:

$$\begin{pmatrix} \dot{\mathbf{w}} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} \nabla v \\ \operatorname{div} \mathbf{w} \end{pmatrix} \quad \text{and} \quad \gamma_\nu^+ \mathbf{w} = g(\gamma^+ v + \dot{u}^{\text{inc}}) - \partial_\nu^+ \dot{u}^{\text{inc}}.$$

The next theorem lays out the functional analytic setting in detail and shows existence of a continuous solution as well as a semi-discrete in space version.

Theorem 5.12. *Consider the space $\mathcal{X} := [L^2(\mathbb{R}^d)]^d \times L^2(\mathbb{R}^d)$ with the product norm and corresponding inner product and define the block operator*

$$A := \begin{pmatrix} 0 & \nabla \\ \operatorname{div} & 0 \end{pmatrix},$$

$$\operatorname{dom}(A) := \left\{ (\mathbf{w}, v) \in \mathcal{X} : \operatorname{div} \mathbf{w} \in L^2(\mathbb{R}^d \setminus \Gamma), v \in \mathcal{H}_h, \right. \\ \left. \llbracket \gamma_\nu \mathbf{w} \rrbracket \in X_h, \gamma_\nu^+ \mathbf{w} - g(\llbracket \gamma v \rrbracket) \in Y_h^\circ \right\}. \quad (5.16)$$

Then A is a maximally monotone operator on \mathcal{X} and generates a strongly continuous semi-group that solves

$$\begin{pmatrix} \dot{\mathbf{w}} \\ \dot{v} \end{pmatrix} = A \begin{pmatrix} \mathbf{w} \\ v \end{pmatrix}, \quad \mathbf{w}(0) = \mathbf{w}_0, \quad v(0) = v_0 \quad (5.17)$$

for all initial data $(\mathbf{w}_0, v_0) \in \operatorname{dom}(A)$. The solution satisfies:

- (i) $u(t) := u_0 + \partial_t^{-1} v = u_0 + \int_0^t v(\tau) d\tau$ solves the wave equation $\ddot{u} = \Delta u$ and satisfies $\dot{u} = v$. If $\nabla u_0 = \mathbf{w}_0$, then u satisfies $\nabla u(t) = \mathbf{w}(t)$ for $t \geq 0$. If $X_h = H^{-1/2}(\Gamma)$ and $Y_h = H^{1/2}(\Gamma)$, then u solves (5.1).
- (ii) $(\mathbf{w}(t), v(t)) \in \operatorname{dom}(A)$. If $u_0 \in \mathcal{H}_h$, then $u(t) \in \mathcal{H}_h$ and $v(t) \in \mathcal{H}_h$ for all $t > 0$.
- (iii) If $u_0 \in \mathcal{H}_h$, then $u \in C^{1,1}([0, \infty), L^2(\mathbb{R}^d)) \cap C^{0,1}([0, \infty), H^1(\mathbb{R}^d \setminus \Gamma))$,
- (iv) $\dot{u} \in L^\infty((0, \infty), H^1(\mathbb{R}^d \setminus \Gamma))$,
- (v) $\ddot{u} \in L^\infty((0, \infty), L^2(\mathbb{R}^d))$.

Proof. We first show monotonicity. Let $x_1 = (\mathbf{w}_1, v_1)$, $x_2 := (\mathbf{w}_2, v_2)$ be in $\operatorname{dom}(A)$. Then

$$\begin{aligned} \langle Ax_1 - Ax_2, x_1 - x_2 \rangle_{\mathcal{X}} &= \langle \nabla v_1 - \nabla v_2, \mathbf{w}_1 - \mathbf{w}_2 \rangle + \langle \operatorname{div} \mathbf{w}_1 - \operatorname{div} \mathbf{w}_2, v_1 - v_2 \rangle \\ &= \langle \gamma_\nu^- \mathbf{w}_1 - \gamma_\nu^- \mathbf{w}_2, \gamma^- v_1 - \gamma^- v_2 \rangle_\Gamma - \langle \gamma_\nu^+ \mathbf{w}_1 - \gamma_\nu^+ \mathbf{w}_2, \gamma^+ v_1 - \gamma^+ v_2 \rangle_\Gamma \\ &= - \langle \llbracket \gamma_\nu [\mathbf{w}_1 - \mathbf{w}_2] \rrbracket, \gamma^- [v_1 - v_2] \rangle_\Gamma - \langle \gamma_\nu^+ \mathbf{w}_1 - \gamma_\nu^+ \mathbf{w}_2, \llbracket \gamma (v_1 - v_2) \rrbracket \rangle_\Gamma \\ &= - \langle g(\llbracket \gamma v_1 \rrbracket) - g(\llbracket \gamma v_2 \rrbracket), \llbracket \gamma (v_1 - v_2) \rrbracket \rangle_\Gamma \\ &\leq 0, \end{aligned}$$

where in the last step, we used the definition of the domain of A , which contains the boundary conditions, and the fact that $[\![\gamma_\nu \mathbf{w}_j]\!] \in X_h$. The definition of \mathcal{H}_h from (5.13) gives that $[\![\gamma(v_1 - v_2)]\!] \in Y_h$.

Next we show $\text{range}(A - \text{Id}) = \mathcal{X}$, i.e., for $(\mathbf{x}, y) \in \mathcal{X}$ we have to find $U = (\mathbf{w}, v) \in \text{dom}(A)$, such that $AU - U = (\mathbf{x}, y)$. In order to do so, we first assume $\mathbf{x} \in H(\text{div}, \mathbb{R}^d \setminus \Gamma)$ (a dense subspace of $[L^2(\mathbb{R}^d)]^d$). From the first equation, we get $\nabla v - \mathbf{w} = \mathbf{x}$, or $\mathbf{w} = \nabla v - \mathbf{x}$, which makes the second equation: $\Delta v - v = y + \text{div } \mathbf{x}$. For the boundary conditions this gives us the requirements

$$\begin{aligned} \gamma^- v &\in X_h^\circ, & [\![\gamma v]\!] &\in Y_h, \\ [\![\gamma_\nu \mathbf{w}]\!] &\in X_h, & \partial_n^+ v - g([\![\gamma v]\!]) - \gamma_\nu^+ \mathbf{x} &\in Y_h^\circ. \end{aligned}$$

This can be solved analogously to the proof of Lemma 5.11. The weak formulation is: Find $v \in \mathcal{H}_h$, such that for all $\xi_h \in \mathcal{H}_h$

$$(v, \xi_h)_{L^2(\mathbb{R}^d)} + (\nabla v, \nabla \xi_h)_{L^2(\mathbb{R}^d)} + \langle g([\![\gamma v]\!]), [\![\gamma \xi_h]\!] \rangle_\Gamma = -(y, \xi_h)_{L^2(\mathbb{R}^d)} + (\mathbf{x}, \nabla \xi_h)_{L^2(\mathbb{R}^d)}.$$

Due to the monotonicity of the left-hand side, this problem has a solution via Proposition 5.9 (see also the proof of Lemma 5.11).

We set $\mathbf{w} = \nabla v - \mathbf{x}$. The fact that the conditions on $[\![\gamma_\nu \mathbf{w}]\!] \in X_h$ hold follows from the same argument as in Lemma 5.11, by using test functions satisfying $[\![\gamma \xi_h]\!] = 0$ and $\gamma^+ \xi_h \in X_h^\circ$ and the fact that $(X_h^\circ)^\circ = X_h$. We therefore have $(\mathbf{w}, v) \in \text{dom } A$.

For general $X := (\mathbf{x}, y) \in \mathcal{X}$, we argue via a density argument. Let $X_n := (\mathbf{x}_n, y_n)$ be a sequence in $\mathcal{X} \cap (H(\text{div}, \mathbb{R}^d \setminus \Gamma) \times L^2(\mathbb{R}^d))$, such that $X_n \rightarrow X$. Let $U_n := (u_n, v_n)$ be the respective solutions to $(A - \text{Id})U_n = X_n$. From the monotonicity of A , we easily see that for $n, m \in \mathbb{N}$: $\|U_n - U_m\|_{\mathcal{X}} \leq \|X_n - X_m\|_{\mathcal{X}}$, which means (U_n) is Cauchy and converges to some $U := (\mathbf{w}, v)$. From the first equation $\nabla v_n - \mathbf{w}_n = \mathbf{x}_n$ we get that $v_n \rightarrow v$ in $H^1(\mathbb{R}^d \setminus \Gamma)$. From the second equation $\text{div } \mathbf{w}_n - v_n = y_n$ we get $\text{div } \mathbf{w}_n \rightarrow \text{div } \mathbf{w}$ in $L^2(\mathbb{R}^d)$. Therefore, we have $\mathbf{w}_n \rightarrow \mathbf{w}$ in $H(\text{div}, \mathbb{R}^d \setminus \Gamma)$, which implies $[\![\gamma_\nu \mathbf{w}_n]\!] \rightarrow [\![\gamma_\nu \mathbf{w}]\!] \in X_h$. From Lemma 5.7(iii) we get $\langle g([\![\gamma v_n]\!]), \xi \rangle_\Gamma \rightarrow \langle g([\![\gamma v]\!]), \xi \rangle_\Gamma$, which implies $\gamma^+ \mathbf{w} - g([\![\gamma v]\!]) \in Y_h^\circ$. The other trace conditions follow from the H^1 -convergence of v_n . The existence of the semigroup then follows from the Kōmura-Kato theorem (Proposition 2.17).

The fact that we can recover a solution to the wave equation using the definition of $u(t)$ is straight forward, and $\nabla u = \mathbf{w}$ follows from $\dot{u} = v$ and $\dot{w} = \nabla v$ together with the initial condition. Since \mathcal{H}_h is a closed subspace and $u_0 = u(0) \in \mathcal{H}_h$, it follows that $u(t) \in \mathcal{H}_h$.

The regularity results follow from the regularity statements in Proposition 2.17. Since v is Lipschitz continuous with values in $L^2(\mathbb{R}^d)$ we get $u \in C^{1,1}((0, T), L^2(\mathbb{R}^d))$. From $\nabla u = \mathbf{w} + \nabla u_0 - \mathbf{w}_0$, we further get $u \in C^{0,1}((0, T), H^1(\mathbb{R}^d \setminus \Gamma))$ as long as $\nabla u_0 \in L^2(\mathbb{R}^d)$. \square

Approximation theory in \mathcal{H}_h and $\text{dom}(A)$

In this section, we collect some important properties of the space \mathcal{H}_h , namely how well it is able to approximate arbitrary functions in $H^1(\Omega^+)$ under different assumptions, and look at how these approximation properties influence the approximation properties in $\text{dom}(A)$. In order to do so, we introduce several projection/quasi-interpolation operators.

We start by defining an operator, which in some sense represents a “volume version” of $J_\Gamma^{Y_h}$; see Lemma 5.14 (ii).

Definition 5.13. Let $\mathcal{E}_\Gamma^D : H^{1/2}(\Gamma) \rightarrow H^1(\mathbb{R}^d \setminus \Gamma)$ denote the linear, continuous lifting operator, such that $\gamma^+ \mathcal{E}_\Gamma^D v = v$ and $\mathcal{E}_\Gamma^D v = 0$ in Ω^- . Then, we define the operator Π_0 as

$$\begin{aligned} \Pi_0 : \left\{ u \in H^1(\mathbb{R}^d \setminus \Gamma) : \gamma^- u = 0 \right\} &\rightarrow \left(\mathcal{H}_h, \|\cdot\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \right) \\ v &\mapsto v - \mathcal{E}_\Gamma^D \left(\left(\text{Id} - J_\Gamma^{Y_h} \right) \gamma^+ v \right) \quad \text{in } \mathbb{R}^d \setminus \Gamma. \end{aligned}$$

Recall that $J_\Gamma^{Y_h}$ denotes a linear, $H^{1/2}(\Gamma) \rightarrow (Y_h, \|\cdot\|_{1/2})$ stable operator.

In the next lemma, we collect some of the most important properties of Π_0 .

Lemma 5.14. *The following statements hold:*

- (i) if $J_\Gamma^{Y_h}$ is a projection, then Π_0 is a projection,
- (ii) $[\gamma \Pi_0 u] = J_\Gamma^{Y_h} [\gamma u]$,
- (iii) Π_0 is stable, with the constant depending only on Γ and $\left\| J_\Gamma^{Y_h} \right\|_{H^{1/2}(\Gamma) \rightarrow H^{1/2}(\Gamma)}$,
- (iv) Π_0 has the same approximation properties in the exterior domain as $J_\Gamma^{Y_h}$ on Γ , i.e.,

$$\|u - \Pi_0 u\|_{H^1(\Omega^+)} \leq C \left\| [\gamma u] - J_\Gamma^{Y_h} [\gamma u] \right\|_{H^{1/2}(\Gamma)}.$$

Proof. All the statements are immediate consequences of the definition and the continuity of \mathcal{E}_Γ^D (Proposition 2.26) and $J_\Gamma^{Y_h}$. \square

In the analysis of time-stepping schemes, it is important to construct an operator which is well behaved with respect to the operator A . This is achieved by constructing a Ritz-type projector, which for our functional-analytic setting takes the following form:

Lemma 5.15. Let $\alpha > 0$ be a fixed stabilization parameter. Define the Ritz-projector $\Pi_1 : H_\Delta^1(\mathbb{R}^d \setminus \Gamma) \rightarrow \mathcal{H}_h$, where $\Pi_1 u$ is the unique solution to

$$\begin{aligned} (\nabla \Pi_1 u, \nabla z_h)_{L^2(\mathbb{R}^d)} + \alpha (\Pi_1 u, z_h)_{L^2(\mathbb{R}^d)} \\ = (\nabla u, \nabla z_h)_{L^2(\mathbb{R}^d)} + \alpha (u, z_h)_{L^2(\mathbb{R}^d)} + \langle [\partial_\nu u], \gamma^- z_h \rangle_\Gamma \quad \forall z_h \in \mathcal{H}_h. \end{aligned} \quad (5.18)$$

The operator Π_1 has the following properties:

- (i) Π_1 is a stable projection onto the space $\mathcal{H}_h \cap \{u \in H_\Delta^1(\mathbb{R}^d \setminus \Gamma) : [\partial_\nu u] \in X_h\}$ with respect to the $H^1(\mathbb{R}^d \setminus \Gamma)$ -norm.
- (ii) Π_1 almost reproduces the exterior normal trace:

$$\langle \partial_n^+ \Pi_1 u, \xi \rangle_\Gamma = \langle \partial_n^+ u, \xi \rangle_\Gamma \quad \forall \xi \in Y_h.$$

(iii) Π_1 has the following approximation property for all $u \in H^1_\Delta(\mathbb{R}^d \setminus \Gamma)$:

$$\|(\text{Id} - \Pi_1)u\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \leq C \left(\inf_{u_h \in \mathcal{H}_h} \|u - u_h\|_{H^1(\mathbb{R}^d \setminus \Gamma)} + \inf_{x_h \in X_h} \|[\partial_\nu u] - x_h\|_{H^{-1/2}(\Gamma)} \right).$$

For $u \in H^1(\mathbb{R}^d \setminus \Gamma)$ with $\gamma^- u = 0$, this approximation problem can be reduced to the boundary spaces X_h, Y_h :

$$\|(\text{Id} - \Pi_1)u\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \leq C_1 \inf_{y_h \in Y_h} \|\gamma^+ u - y_h\|_{H^{1/2}(\Gamma)} + C_2 \inf_{x_h \in X_h} \|[\partial_\nu u] - x_h\|_{H^{-1/2}(\Gamma)}. \quad (5.19)$$

All the constants depend only on Γ and α .

Proof. The operator is well defined and stable as \mathcal{H}_h is a closed subspace of $H^1(\mathbb{R}^d \setminus \Gamma)$, the bilinear form used is elliptic and the right-hand side is a bounded linear functional.

In order to prove $[\partial_\nu \Pi_1 u] \in X_h$, we follow the same argument as in the proof of Lemma 5.11. First, we establish by using $z_h \in C_0^\infty(\mathbb{R}^d \setminus \Gamma)$ that $\Pi_1 u$ solves the PDE $-\Delta \Pi_1 u + \alpha \Pi_1 u = -\Delta u + \alpha u$. For $\xi \in X_h^\circ$ we obtain by using a global H^1 -lifting, i.e., such that $\gamma^+ z_h = \gamma^- z_h = \xi$, and using integration by parts:

$$\langle [\partial_\nu \Pi_1 u], \xi \rangle_\Gamma = \langle [\partial_\nu u], \xi \rangle_\Gamma - \langle [\partial_\nu u], \xi \rangle_\Gamma = 0.$$

This means $[\partial_\nu \Pi_1 u] \in (X_h^\circ)^\circ = X_h$ and $\text{range}(\Pi_1) \subseteq \mathcal{H}_h \cap \{u \in H^1_\Delta(\mathbb{R}^d \setminus \Gamma) : [\partial_\nu u] \in X_h\}$.

The fact that Π_1 reproduces the normal jump follows from testing with an arbitrary $z_h \in \mathcal{H}_h$ with $\gamma^- z_h = 0$ together with integration by parts and the PDE for $\Pi_1 u$.

To see that it is a projection, we note that for $u \in \mathcal{H}_h$ with $[\partial_\nu u] \in X_h$, the term $\langle [\partial_\nu u], \gamma^- z_h \rangle_\Gamma$ vanishes due to the requirement $\gamma^- z_h \in X_h^\circ$.

In order to prove the approximation property, we fix arbitrary $u_h \in \mathcal{H}_h$ and $x_h \in X_h$ and calculate:

$$\begin{aligned} & \|(\text{Id} - \Pi_1)u\|_{H^1(\mathbb{R}^d \setminus \Gamma)}^2 \\ & \lesssim (\nabla(\text{Id} - \Pi_1)u, \nabla(\text{Id} - \Pi_1)u)_{L^2(\mathbb{R}^d)} + \alpha ((\text{Id} - \Pi_1)u, (\text{Id} - \Pi_1)u)_{L^2(\mathbb{R}^d)} \\ & = (\nabla(\text{Id} - \Pi_1)u, \nabla u - \nabla u_h)_{L^2(\mathbb{R}^d)} + \alpha ((\text{Id} - \Pi_1)u, u - u_h)_{L^2(\mathbb{R}^d)} \\ & \quad + \langle [\partial_\nu u] - x_h, \gamma^- \Pi_1 u - \gamma^- u_h \rangle_\Gamma \\ & \lesssim \|(\text{Id} - \Pi_1)u\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \|u - u_h\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \\ & \quad + \|[\partial_\nu u] - x_h\|_{H^{-1/2}(\Gamma)} \left(\|u - \Pi_1 u\|_{H^1(\mathbb{R}^d \setminus \Gamma)} + \|u - u_h\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \right). \end{aligned}$$

Young's inequality concludes the proof.

For (5.19), we need to estimate $\inf_{u_h \in \mathcal{H}_h} \|u - u_h\|_{H^1(\mathbb{R}^d \setminus \Gamma)}$. We do this in a manner similar to what was done for the Schrödinger case in Lemma 4.17. Let $y_h \in Y_h$ be arbitrary and let θ be a continuous H^1 -lifting of $\gamma^+ u - y_h$ to Ω^+ . Define $u_h := u|_{\Omega^-}$ in Ω^- and $u_h := u|_{\Omega^+} - \theta$ in Ω^+ . Since we assumed $\gamma^- u = 0$, we get $[\gamma u_h] = y_h \in Y_h$ and therefore $u_h \in \mathcal{H}_h$. For the norm we estimate:

$$\|u - u_h\|_{H^1(\mathbb{R}^d \setminus \Gamma)} = \|\theta\|_{H^1(\Omega^+)} \leq C \|\gamma^+ u - y_h\|_{H^{1/2}(\Gamma)}$$

due to the continuity of the lifting operator (Proposition 2.26). \square

All the previous approximation operators require the function to have at least H^1 -regularity. Since we are interested in the case of no additional regularity assumptions on the exact solution, we also need an additional operator that is stable in an L^2 -setting. The construction is very similar to the one used in Lemma 4.15 for the Schrödinger setting. In order to be able to define such an operator, we need to make an additional assumption on Y_h .

Assumption 5.16. *For all $h > 0$, there exist spaces $Y_h^{\Omega^-} \subseteq H^1(\Omega^-)$, such that $Y_h \supseteq \gamma^- Y_h^{\Omega^-}$ and there exists a linear operator $J_{\Omega^-}^{Y_h} : L^2(\Omega^-) \rightarrow Y_h^{\Omega^-}$ with the following properties: $J_{\Omega^-}^{Y_h}$ is stable in the L^2 - and H^1 -norm and for $s \in \{0, 1\}$ satisfies the strong convergence*

$$\left\| u - J_{\Omega^-}^{Y_h} u \right\|_{H^s(\Omega^-)} \rightarrow 0 \quad \text{for } h \rightarrow 0, \quad \forall u \in H^s(\Omega^-).$$

This allows us to define our last approximation operator.

Lemma 5.17. *Let $\mathcal{E}^\pm : H^m(\Omega^\pm) \rightarrow H^m(\mathbb{R}^d)$ denote the Stein extension operator from Proposition 2.23, which is stable for all $m \in \mathbb{N}_0$. Then we define a new operator $\Pi_2 : L^2(\Omega^+) \rightarrow \mathcal{H}_h$ by*

$$u \mapsto \Pi_2 u := \begin{cases} u - \mathcal{E}^- \left(\left(\text{Id} - J_{\Omega^-}^{Y_h} \right) \mathcal{E}^+ u \right) & \text{in } \Omega^+ \\ 0 & \text{in } \Omega^- \end{cases},$$

i.e., in order to get a correction term similar to the one for Π_0 , instead of relying on traces and lifting, we extend the function to the interior, apply the approximation operator mapping to $Y_h^{\Omega^-}$ and extend it back outwards.

This operator has the following nice properties:

(i) Π_2 is stable in L^2 and H^1 ,

(ii) for $s \in [0, 1]$: $\|u - \Pi_2 u\|_{H^s(\Omega^+)} \lesssim \left\| \left(\text{Id} - J_{\Omega^-}^{Y_h} \right) \mathcal{E}^+ u \right\|_{H^s(\Omega^-)}$,

(iii) for all $u \in L^2(\Omega^+)$ and for $h \rightarrow 0$, $\Pi_2 u$ converges to u in the L^2 -norm without further regularity assumptions. For $u \in H^1(\Omega^+)$, the convergence is in the H^1 -norm.

Proof. In order to see that $\Pi_2 u \in \mathcal{H}_h$, we have to show $\gamma^+ \Pi_2 u \in Y_h$. We calculate:

$$\gamma^+ \Pi_2 u = \gamma^+ u - \gamma^+ u + \gamma^- J_{\Omega^-}^{Y_h} \mathcal{E}^+ u \in Y_h$$

due to the fact that \mathcal{E}^- reproduces the trace and the assumptions on $Y_h^{\Omega^-}$. For the approximation properties, we use the continuity of the Stein extension (see Proposition 2.23(i)) and estimate:

$$\|u - \Pi_2 u\|_{H^s(\mathbb{R}^d \setminus \Gamma)} = \left\| \mathcal{E}^- \left(\left(\text{Id} - J_{\Omega^-}^{Y_h} \right) \mathcal{E}^+ u \right) \right\|_{H^s(\mathbb{R}^d \setminus \Gamma)} \lesssim \left\| \left(\text{Id} - J_{\Omega^-}^{Y_h} \right) \mathcal{E}^+ u \right\|_{H^s(\mathbb{R}^d \setminus \Gamma)}.$$

The extension $\mathcal{E}^+ u$ has the same regularity as u , thus we end up with the correct convergence rates of $Y_h^{\Omega^-}$. \square

Remark 5.18. While Assumption 5.16 may look somewhat artificial, in most cases it is easily verified by constructing a “virtual triangulation” of Ω^- with piecewise polynomials in the spirit of a FEM-BEM coupling procedure as for example used in Chapter 4. The projection operator $J_{\Omega^-}^{Y_h}$ is then (for example) given by some stable quasi-interpolation operator (see 2.36). ■

Remark 5.19. We used a space on Ω^- for the quasi-interpolation step as it reflects the fact that usually a natural triangulation on Ω^- is given and the boundary mesh was generated by restricting a volume mesh. This choice is arbitrary and could for example be replaced by an artificial layer of triangles around Γ in Ω^+ . This would have allowed to drop the extension step to the interior. ■

To conclude this section, we look at some properties of the nonlinearity g , namely how approximations in η impact the convergence of $g(\eta)$.

Lemma 5.20. Let $\eta \in H^{1/2}(\Gamma)$ and $\eta_h \in H^{1/2}(\Gamma)$ be such that η_h converges to η weakly, i.e., for $\xi \in H^{-1/2}(\Gamma)$ it holds that

$$\langle \xi, \eta_h \rangle_{\Gamma} \rightarrow \langle \xi, \eta \rangle_{\Gamma} \quad \text{for } h \rightarrow 0;$$

we write $\eta_h \rightharpoonup \eta$. Then the following statements hold:

(i) $g(\eta_h) \rightharpoonup g(\eta)$ in $H^{-1/2}(\Gamma)$.

(ii) Assume that $\eta_h \rightarrow \eta$ in the $H^{1/2}(\Gamma)$ norm, and assume the stricter growth condition $|g(\mu)| \leq C(1 + |\mu|^p)$ with $p < \infty$ for $d = 2$ and $p \leq \frac{d-1}{d-2}$ for $d \geq 3$. Then the nonlinearity converges with respect to the $L^2(\Gamma)$ -norm as well:

$$\|g(\eta) - g(\eta_h)\|_{L^2(\Gamma)} \rightarrow 0.$$

(iii) If we assume $|g'(s)| \leq C_g(1 + |s|^q)$, where $q < \infty$ is arbitrary for $d = 2$, and $q \leq 1$ for $d = 3$, then the following estimates hold:

$$\begin{aligned} \|g(\eta) - g(\eta_h)\|_{L^2(\Gamma)} &\leq C(\eta) \|\eta - \eta_h\|_{L^{2+\varepsilon}(\Gamma)} && \text{for } d = 2, \forall \varepsilon > 0, \\ \|g(\eta) - g(\eta_h)\|_{L^2(\Gamma)} &\leq C(\eta) \|\eta - \eta_h\|_{H^{1/2}(\Gamma)}, \end{aligned}$$

where the constant $C(\eta)$ depends on η but does not depend on h .

(iv) For $\eta \in L^\infty(\Gamma)$ and $\|\eta_h - \eta\|_{L^\infty(\Gamma)} \leq C_\infty$ we have

$$\|g(\eta) - g(\eta_h)\|_{L^2(\Gamma)} \leq C(\|\eta\|_\infty, C_\infty) \|\eta - \eta_h\|_{L^2(\Gamma)}.$$

where $C(\|\eta\|_\infty, C_\infty)$ depends on η , C_∞ and g but not on h .

Proof. Ad (i): We focus on the case $d \geq 3$, the case $d = 2$ follows along the same lines but is simpler since the Sobolev embeddings hold for arbitrary $p \in [1, \infty)$. Since weakly convergent sequences are bounded (see [Yos80, Theorem 1(ii), Chapter V.1]), we can apply Lemma 5.7(i) to get that $g(\eta_h)$ is uniformly bounded in $H^{-1/2}(\Gamma)$. By the Eberlein-Šmulian

theorem, see [Yos80, page 141], this implies that there exists a subsequence $g(\eta_{h_j})$, $j \in \mathbb{N}$, which converges weakly to some limit $\tilde{g} \in H^{-1/2}(\Gamma)$. We need to identify the limit \tilde{g} as $g(\eta)$. By Rellich's theorem, the sequence η_h converges to η in $H^s(\Gamma)$ for $s < 1/2$, and using Sobolev embeddings we get that $\eta_{h_j} \rightarrow \eta$ in $L^{p'}(\Gamma)$ for $p' < \frac{2d-2}{d-2}$. [Bre83, Theorem IV.9] then gives (up to picking another subsequence) that $\eta_{h_j} \rightarrow \eta$ pointwise almost everywhere and there exists an upper bound $\zeta \in L^{p'}(\Gamma)$ such that $|\eta_{h_j}| \leq \zeta$ almost everywhere. By the growth condition on g , we get that $|g(\eta)| \leq C(1 + |\zeta|^p)$ and since $p \leq \frac{d}{d-2} \leq p'$, the function $1 + |\zeta|^p$ is integrable. By the continuity of g we also get $g(\eta_{h_j}) \rightarrow g(\eta)$ almost everywhere. For test functions $\phi \in C^\infty(\Gamma)$, we get:

$$\int_{\Gamma} g(\eta_{h_j})\phi \rightarrow \int_{\Gamma} g(\eta)\phi$$

by the dominated convergence theorem (note that ϕ is bounded). On the other hand, since $C^\infty(\Gamma) \subseteq H^{1/2}(\Gamma)$, we get $\langle g(\eta_{h_j}), \phi \rangle_{\Gamma} \rightarrow \langle \tilde{g}, \phi \rangle_{\Gamma}$ due to the weak convergence. Since $C^\infty(\Gamma)$ is dense in $H^{1/2}(\Gamma)$, we get $g(\eta) = \tilde{g}$. This proof can be repeated for every subsequence, thus the whole sequence must converge weakly to $g(\eta)$. Ad (ii): The proof follows along the same lines as in Lemma 5.7(iii). Instead of estimating the $H^{-1/2}$ -norm by the p' norm via the duality argument, we can directly work in L^2 . Due to our restrictions on g and the Sobolev embedding, we get an upper bound $C(1 + |\eta|^p)$ in $L^2(\Gamma)$, which allows us to apply the same argument as before to get convergence.

Ad (iii): Using the growth condition on g' we estimate for fixed $x \in \Gamma$:

$$\begin{aligned} |g(\eta(x)) - g(\eta_h(x))| &= \left| \int_{\eta_h(x)}^{\eta(x)} g'(\xi) d\xi \right| \leq |\eta_h(x) - \eta(x)| \sup_{\xi \in [\eta_h(x), \eta(x)]} |g'(\xi)| \\ &\leq |\eta_h(x) - \eta(x)| C(1 + \max(|\eta_h(x)|^q, |\eta(x)|^q)). \end{aligned} \quad (5.20)$$

In the case $d = 3$, we use the Cauchy-Schwarz inequality to estimate:

$$\begin{aligned} \|g(\eta) - g(\eta_h)\|_{L^2(\Gamma)}^2 &\lesssim \left\| (1 + \max(|\eta_h|^q, |\eta|^q))^2 \right\|_{L^2(\Gamma)} \left\| (\eta - \eta_h)^2 \right\|_{L^2(\Gamma)} \\ &\lesssim \left(1 + \|\eta_h\|_{L^{4q}(\Gamma)}^2 + \|\eta\|_{L^{4q}(\Gamma)}^2 \right) \|\eta - \eta_h\|_{L^4(\Gamma)}^2 \\ &\lesssim \left(1 + \|\eta_h\|_{H^{1/2}(\Gamma)}^2 + \|\eta\|_{H^{1/2}(\Gamma)}^2 \right) \|\eta - \eta_h\|_{H^{1/2}(\Gamma)}^2, \end{aligned}$$

where in the last step we used the Sobolev embedding. Since weakly convergent sequences are bounded, the first term can be uniformly bounded with respect to h , which shows (iii) for $d \geq 3$.

In the case $d = 2$, we have by Sobolev's embedding that $\|\max(|\eta_h|, |\eta|)\|_{L^{p'}(\Gamma)}$ can be bounded independently of h for arbitrary $p' > 1$. Using (5.20) to estimate the difference and applying Hölders inequality then proves (iii) in the case $d = 2$.

Ad (iv): Since g is assumed to be continuously differentiable, g' is bounded on compact subsets of \mathbb{R} . Arguing as before and using the bounds on η and η_h , the derivative $g'(\xi(x))$ is therefore uniformly bounded in this case. The statement then follows again by using Hölders inequality. \square

Convergence of the semi-discretization in space

We can now give an estimate of how the discretization in space, due to the choice of spaces X_h and Y_h and represented in the domain of the semigroup in Theorem 5.12, impacts the solution. In order to do so, it is easier to work in a weak formulation of the semigroup.

Lemma 5.21. *Let \mathcal{H}_h be defined as in Lemma 5.11. Then the semi-discrete solution to (5.17) denoted by $(\mathbf{w}^h, v^h)(t)$ solves*

$$\left(\dot{\mathbf{w}}^h(t), \mathbf{q}_h \right)_{[L^2(\mathbb{R}^d)]^d} = \left(\nabla v^h(t), \mathbf{q}_h \right)_{[L^2(\mathbb{R}^d)]^d}, \quad (5.21a)$$

$$\left(\dot{v}^h(t), z_h \right)_{L^2(\mathbb{R}^d)} = - \left(\mathbf{w}^h(t), \nabla z_h \right)_{[L^2(\mathbb{R}^d)]^d} - \left\langle g \left(\llbracket v^h \rrbracket \right), \llbracket \gamma z_h \rrbracket \right\rangle_{\Gamma} \quad (5.21b)$$

for all $(\mathbf{q}_h, z_h) \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma) \times \mathcal{H}_h$ and $t > 0$. The exact solution satisfies

$$\left(\dot{\mathbf{w}}(t), \mathbf{q}_h \right)_{[L^2(\mathbb{R}^d)]^d} = \left(\nabla v(t), \mathbf{q}_h \right)_{[L^2(\mathbb{R}^d)]^d}, \quad (5.22a)$$

$$\begin{aligned} \left(\dot{v}(t), z_h \right)_{L^2(\mathbb{R}^d)} = & - \left(\mathbf{w}(t), \nabla z_h \right)_{[L^2(\mathbb{R}^d)]^d} \\ & - \left\langle g(\llbracket \gamma v \rrbracket), \llbracket \gamma z_h \rrbracket \right\rangle_{\Gamma} - \left\langle \llbracket \gamma v \mathbf{w} \rrbracket, \gamma^- z_h \right\rangle_{\Gamma} \end{aligned} \quad (5.22b)$$

for all $(\mathbf{q}_h, z_h) \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma) \times \mathcal{H}_h$ and $t > 0$. We note that we also have the function $u(t) \in C^{1,1}(\mathbb{R}_+, H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma))$ such that $\nabla u = \mathbf{w}$ and $\dot{u} = v$.

Proof. This is a simple consequence of (5.14), the definition of $\operatorname{dom}(A)$ and integration by parts. The last term in (5.22b), which would not appear in a straight-forward weak formulation of the exterior scattering problem (5.1), is due to the fact that we replaced $\gamma^+ z_h$ with $\llbracket \gamma z_h \rrbracket$ in the boundary term containing the nonlinearity. In comparison to (5.21) with X_h and Y_h as the full space, the additional term is there because the condition $\gamma^- z_h \in X_h^{\circ}$, which would imply $\gamma^- z_h = 0$ for the full space case, is violated for our admissible test functions in (5.22). \square

Theorem 5.22. *Assume that there exists an L^2 -stable operator Π_2 , which takes values in \mathcal{H}_h with $\|v - \Pi_2 v\|_{L^2(\Omega_+)} \rightarrow 0$ for $h \rightarrow 0$ as described in Lemma 5.17.*

Introducing the error functions

$$\begin{aligned} \rho(t) &:= \begin{pmatrix} \rho_{\mathbf{w}}(t) \\ \rho_v(t) \end{pmatrix} := \begin{pmatrix} \mathbf{w} - \nabla \Pi_1 u \\ v - \Pi_2 v \end{pmatrix}, \\ \theta(t) &:= g(\llbracket \gamma v(t) \rrbracket) - g(\llbracket \gamma \Pi_2 v(t) \rrbracket), \\ \varepsilon(t) &:= u(t) - \Pi_1 u(t), \end{aligned}$$

the convergence can be quantified as

$$\begin{aligned} & \left\| v^h(t) - v(t) \right\|_{L^2(\mathbb{R}^d)}^2 + \left\| \mathbf{w}^h(t) - \mathbf{w}(t) \right\|_{[L^2(\mathbb{R}^d)]^d}^2 + \beta \int_0^t \left\| \llbracket \gamma v^h(\tau) \rrbracket - \llbracket \gamma v(\tau) \rrbracket \right\|_{L^2(\Gamma)}^2 d\tau \\ & \lesssim \left\| v^h(0) - v(0) \right\|_{L^2(\mathbb{R}^d)}^2 + \left\| \mathbf{w}^h(0) - \mathbf{w}(0) \right\|_{[L^2(\mathbb{R}^d)]^d}^2 \\ & \quad + T \int_0^t \left\| \dot{\rho}(\tau) \right\|_{\mathcal{X}}^2 + \left\| \nabla \rho_v(\tau) \right\|_{[L^2(\mathbb{R}^d)]^d}^2 + \left\| \varepsilon(\tau) \right\|_{L^2(\mathbb{R}^d)}^2 + \beta^{-1} \left\| \theta(\tau) \right\|_{L^2(\Gamma)}^2 d\tau. \end{aligned} \quad (5.23)$$

The implied constant depends only on the stabilization parameter α from (5.18).

If the operator $g : H^{1/2}(\Gamma) \rightarrow L^2(\Gamma)$ is continuous (see Lemma 5.20(ii) for a sufficient condition), then the right-hand side converges to zero for $h \rightarrow 0$.

Proof. The additional error function

$$e(t) := \begin{pmatrix} e_{\mathbf{w}}(t) \\ e_v(t) \end{pmatrix} := \begin{pmatrix} \mathbf{w}^h - \nabla \Pi_1 u \\ v^h - \Pi_2 v \end{pmatrix}$$

solves for fixed $t \geq 0$

$$(\dot{e}_{\mathbf{w}}, q_h)_{[L^2(\mathbb{R}^d)]^d} = (\nabla e_v, q_h)_{[L^2(\mathbb{R}^d)]^d} - (\dot{\rho}_{\mathbf{w}}, q_h)_{[L^2(\mathbb{R}^d)]^d} + (\nabla \rho_v, q_h)_{[L^2(\mathbb{R}^d)]^d}$$

for $w_h \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)$ and $z_h \in \mathcal{H}_h$ as well as

$$\begin{aligned} (\dot{e}_v, z_h)_{L^2(\mathbb{R}^d)} &= - (e_{\mathbf{w}}, \nabla z_h)_{[L^2(\mathbb{R}^d)]^d} - \left\langle g(\llbracket v^h \rrbracket) - g(\llbracket \Pi_2 v \rrbracket), \llbracket \gamma z_h \rrbracket \right\rangle_{\Gamma} \\ &\quad + (\dot{\rho}_v, z_h)_{L^2(\mathbb{R}^d)} + \alpha (\varepsilon, z_h)_{L^2(\mathbb{R}^d)} + \langle \theta, \llbracket \gamma z_h \rrbracket \rangle_{\Gamma}. \end{aligned}$$

Testing with $q_h := e_{\mathbf{w}}$ and $z_h := e_v$ and adding the two equations gives, using the strong monotonicity:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|e\|_{\mathcal{X}}^2 + \beta \|\llbracket \gamma e_v \rrbracket\|_{L^2(\Gamma)}^2 &\leq \langle \dot{e}, e \rangle_{\mathcal{X}} + \left\langle g(\llbracket v^h \rrbracket) - g(\Pi_2 \llbracket \gamma v \rrbracket), \llbracket \gamma e_v \rrbracket \right\rangle_{\Gamma} \\ &= - (\dot{\rho}_{\mathbf{w}}, e_{\mathbf{w}})_{[L^2(\mathbb{R}^d)]^d} + (\nabla \rho_v, e_{\mathbf{w}})_{L^2(\mathbb{R}^d)} \\ &\quad + (\dot{\rho}_v, e_v)_{L^2(\mathbb{R}^d)} + \alpha (\varepsilon, e_v)_{L^2(\mathbb{R}^d)} + \langle \theta, \llbracket \gamma e_v \rrbracket \rangle_{\Gamma} \\ &\leq \|\dot{\rho}\|_{\mathcal{X}} \|e\|_{\mathcal{X}} + \|\nabla \rho_v\|_{[L^2(\mathbb{R}^d)]^d} \|e_{\mathbf{w}}\|_{[L^2(\mathbb{R}^d)]^d} \\ &\quad + \alpha \|\varepsilon\|_{L^2(\mathbb{R}^d)} \|e_v\|_{L^2(\mathbb{R}^d)} + \|\theta\|_{L^2(\Gamma)} \|\llbracket \gamma e_v \rrbracket\|_{L^2(\Gamma)} \\ &\lesssim \left(\|\dot{\rho}\|_{\mathcal{X}} + \|\nabla \rho_v\|_{[L^2(\mathbb{R}^d)]^d} + \|\varepsilon\|_{L^2(\mathbb{R}^d)} \right) \|e\|_{\mathcal{X}} \\ &\quad + \|\theta\|_{L^2(\Gamma)} \|\llbracket \gamma e_v \rrbracket\|_{L^2(\Gamma)}. \end{aligned}$$

Young's inequality and integrating then gives:

$$\begin{aligned} \|e\|_{\mathcal{X}}^2 + \beta \int_0^t \|\llbracket \gamma e_v \rrbracket\|_{L^2(\Gamma)}^2 &\lesssim T \int_0^t \|\dot{\rho}(\tau)\|_{\mathcal{X}}^2 + \|\nabla \rho_v(\tau)\|_{[L^2(\mathbb{R}^d)]^d}^2 + \|\varepsilon(\tau)\|_{L^2(\mathbb{R}^d)}^2 + \|\theta\|_{L^2(\Gamma)}^2 \, d\tau \\ &\quad + T^{-1} \int_0^t \|e(\tau)\|_{\mathcal{X}}^2 + \|\llbracket \gamma e_v(\tau) \rrbracket\|_{L^2(\Gamma)}^2 \, d\tau. \end{aligned}$$

Gronwall's inequality ([Tes12, Lemma 2.7]) then lets us bound $\|e\|_{\mathcal{X}}^2 + \beta \int_0^t \|\llbracket \gamma e_v \rrbracket\|_{L^2(\Gamma)}^2$ by the right hand side of (5.23). The triangle inequality $\|(\mathbf{w}^h - \mathbf{w}, v^h - v)\|_{\mathcal{X}} \leq \|e\|_{\mathcal{X}} + \|\rho\|_{\mathcal{X}}$ then completes the proof of (5.23).

In order to see convergence, we need to investigate the different error contributions. By Theorem 5.12, we have $u, v \in L^\infty((0, T), H^1(\Omega))$. $\dot{\rho}$ measures the approximation of

$\nabla \dot{u} = \nabla v$ and \dot{v} . By the approximation properties of Π_1 in the H^1 -norm and Π_2 in the L^2 -norm (see Lemmas 5.15 and 5.17) we get convergence of $\|\dot{\rho}\|_{\mathcal{X}}$. The term $\|\varepsilon\|_{L^2(\mathbb{R}^d)}$ also converges due to the properties of the Ritz-projector. The term $\|\nabla \rho_v\|$ is also no problem as Π_2 has approximation properties in the full range of Sobolev spaces in $[0, 1]$. This means, as long as the nonlinear term converges, we obtain convergence of the fully discrete scheme. \square

Remark 5.23. *It might seem advantageous to use the Ritz projector Π_1 throughout the proof of Theorem 5.22 (or a Ritz-type operator adapted to $\begin{pmatrix} 0 & \nabla \\ \operatorname{div} & 0 \end{pmatrix}$) as this choice eliminates the term $\|\nabla \rho_v\|_{L^2(\mathbb{R}^d \setminus \Gamma)}$, but the Ritz projector is not defined for $\dot{v} \in L^2(\mathbb{R}^d)$. Thus, we have to either assume additional regularity or use the projector Π_2 . \blacksquare*

5.4.2 Time discretization analysis

The next step in analyzing the discretization quality of our numerical method is to investigate the error made due to the Convolution Quadrature treatment. Lemma 5.11 hints that it is sufficient to look at the multistep approximation of the related semigroup introduced in the previous section. This will be the topic of this section. We start with a general result from the literature regarding the approximation of semigroups by multistep methods.

Proposition 5.24. *Let A be a maximally monotone operator on a separable Hilbert space \mathcal{H} with domain $\operatorname{dom}(A) \subseteq \mathcal{H}$, and let u denote the semigroup solution from Proposition 2.17.*

For $k > 0$, we define the multistep approximation u^k by

$$\frac{1}{k} \sum_{j=0}^m \alpha_j u^k(t_{n-j}) = Au^k(t_n), \quad (5.24)$$

where we assumed $u^k(t) = u(t)$ for $t < mk$, and the coefficients α_j originate from the implicit Euler or the BDF2 method.

Then u^k is well-defined, i.e., (5.24) has a unique sequence of solutions, with $u^k(t_n) \in \operatorname{dom}(A)$. If $u^k(t) \in \operatorname{dom}(A)$ for $t < mk$, then the following estimate holds for $nk \leq T$:

$$\max_{nk \leq T} \left\| u(t_n) - u^k(t_n) \right\|_{\mathcal{H}} \leq C \|Au_0\|_{\mathcal{H}} \left[k + T^{1/2} k^{1/2} + \left(T + T^{1/2} \right) k^{1/3} \right].$$

Assume that $u \in C^{p+1}([0, T], \mathcal{H})$, where p is the order of the multistep method. Then

$$\max_{nk \leq T} \left\| u(t_n) - u^k(t_n) \right\|_{\mathcal{H}} \leq C(u) T k^p.$$

Here the constant $C(u)$ depends on u and its derivatives but not on k .

Proof. The general convergence result was shown by Nevanlinna in [Nev78, Corollary 1]. The improved convergence rate is shown in [HW10, Chapter V.8, Theorem 8.2] or follows directly by inserting exact solution into the discrete scheme, applying the stability theorem [Nev78, Theorem 1] and estimating the consistency error by Taylor's theorem. \square

Remark 5.25. We will use a shifted version of the previous proposition, where we assume $u^k(t) = u(t)$ for $t \leq 0$ and define all $u^k(t)$ via (5.24) for all $t = t_0 + kn$ with $t_0 \in (-mk, 0]$ and $n \in \mathbb{N}$. This does not impact the stated results. \blacksquare

We are now able to define the multistep approximation sequence $(u_{\text{sg}}^{h,k}, v_{\text{sg}}^{h,k}) \subseteq \text{dom}(A)$ of the (spatially discrete) semigroup from Theorem 5.12 as

$$\frac{1}{k} \sum_{j=0}^m \alpha_j \mathbf{w}_{\text{sg}}^{h,k}(t_{n-j}) = \nabla v_{\text{sg}}^{h,k}(t_n), \quad (5.25a)$$

$$\frac{1}{k} \sum_{j=0}^m \alpha_j v_{\text{sg}}^{h,k}(t_{n-j}) = \text{div } \mathbf{w}_{\text{sg}}^{h,k}(t_n), \quad (5.25b)$$

together with the initial conditions $u_{\text{sg}}^{h,k}(t) = u^{\text{inc}}(t)$, $v_{\text{sg}}^{h,k}(t) = \dot{u}^{\text{inc}}(t)$ for $t \leq 0$; see Proposition 5.24.

Comparing this definition to (5.14), we see that due to the way we dealt with u^{inc} , the approximation of the semigroup does not coincide with the approximation induced by the boundary integral equations. This discrepancy does not compromise the convergence rates, as is shown in the following lemma.

Lemma 5.26. Let $p > 0$ denote the order of the multistep method and assume

$$\begin{aligned} u_{\text{inc}} &\in C^\mu \left((0, T), H^1(\mathbb{R}^d \setminus \Gamma) \right), \\ \dot{u}_{\text{inc}} &\in C^\mu \left((0, T), L^2(\mathbb{R}^d \setminus \Gamma) \right) \end{aligned}$$

for $\mu > 1$. Using the operators Π_0 and Π_1 , as defined as in Section 5.4.1, we define the shifted version of $u_{\text{ie}}^{h,k}$ via $\tilde{u}_{\text{ie}}^{h,k} := u_{\text{ie}}^{h,k} + \Pi_1 u^{\text{inc}}$, $\tilde{v}_{\text{ie}}^{h,k} := \partial_t^k u_{\text{ie}}^{h,k} + \Pi_0 \dot{u}^{\text{inc}}$ and $\tilde{\mathbf{w}}_{\text{ie}}^{h,k} := \nabla \tilde{u}_{\text{ie}}^{h,k}$.

Then the following error estimate holds:

$$\begin{aligned} &\left(\left\| v_{\text{sg}}^{h,k}(t_n) - \tilde{v}_{\text{ie}}^{h,k}(t_n) \right\|_{L^2(\mathbb{R}^d)}^2 + \left\| \mathbf{w}_{\text{sg}}^{h,k} - \tilde{\mathbf{w}}_{\text{ie}}^{h,k} \right\|_{[L^2(\mathbb{R}^d)]^d}^2 \right)^{1/2} \\ &\lesssim k^{\min(p, \mu-1)} T \left\| (\nabla u^{\text{inc}}, \dot{u}^{\text{inc}}) \right\|_{C^\mu((0, T), \mathcal{X})} + k \sum_{j=0}^n \left\| \left(\text{Id} - J_\Gamma^{Y_h} \right) \gamma^+ \dot{u}^{\text{inc}}(t_j) \right\|_{H^{1/2}(\Gamma)} \\ &\quad + k \sum_{j=0}^n \inf_{x_h \in X_h} \left\| \partial_n^+ \dot{u}^{\text{inc}}(t_j) - x_h \right\|_{H^{-1/2}(\Gamma)} + k \sum_{j=0}^n \left\| \left(\text{Id} - J_\Gamma^{Y_h} \right) \gamma^+ \ddot{u}^{\text{inc}}(t_j) \right\|_{H^{1/2}(\Gamma)}. \end{aligned}$$

The same convergence rates hold if we use u^{inc} and \dot{u}^{inc} instead of the projected versions on the left hand side. Thus, in practice, we do not depend on the non-computable operators Π_0 and Π_1 .

Proof. Inserting the definition of $\tilde{\mathbf{w}}_{\text{ie}}^{h,k}$ and $\tilde{v}_{\text{ie}}^{h,k}$ into the multistep method, using (5.14), we

see that $(\tilde{\mathbf{w}}_{\text{ie}}^{h,k}, \tilde{v}_{\text{ie}}^{h,k})$ solves

$$\begin{aligned} \frac{1}{k} \sum_{j=0}^m \alpha_j \tilde{\mathbf{w}}_{\text{ie}}^{h,k}(t_{n-j}) &= \nabla \tilde{v}_{\text{ie}}^{h,k}(t_n) + \nabla \varepsilon(t_n), \\ \frac{1}{k} \sum_{j=0}^m \beta_j \tilde{v}_{\text{ie}}^{h,k}(t_{n-j}) &= \operatorname{div}(\tilde{\mathbf{w}}_{\text{ie}}^{h,k}(t_n)) + \theta(t_n) \end{aligned}$$

with right-hand sides

$$\begin{aligned} \varepsilon(t) &:= \frac{1}{k} \sum_{j=0}^m \alpha_j \Pi_1 u^{\text{inc}}(t - t_j) - \Pi_0 \dot{u}^{\text{inc}}(t), \\ \theta(t) &:= \frac{1}{k} \sum_{j=0}^m \alpha_j \Pi_0 \dot{u}^{\text{inc}}(t - t_j) - \Delta \Pi_1 u^{\text{inc}}(t). \end{aligned}$$

In Theorem 5.12 we have shown that A is maximally monotone. From the properties of Π_0 and Π_1 , we have $(\tilde{u}_{\text{ie}}^{h,k}(t_n), \tilde{v}_{\text{ie}}^{h,k}(t_n)) \in \operatorname{dom}(A)$ (Lemmas 5.14 and 5.15). This means, we can apply the stability estimate [Nev78, Theorem 1] to get for the differences:

$$\begin{aligned} &\left(\left\| v_{\text{sg}}^{h,k}(t_n) - \tilde{v}_{\text{ie}}^{h,k}(t_n) \right\|_{L^2(\mathbb{R}^d)}^2 + \left\| \mathbf{w}_{\text{sg}}^{h,k}(t_n) - \tilde{\mathbf{w}}_{\text{ie}}^{h,k}(t_n) \right\|_{[L^2(\mathbb{R}^d)]^d}^2 \right)^{1/2} \\ &\leq k \sum_{j=0}^n \left(\left\| \theta(t_j) \right\|_{L^2(\mathbb{R}^d)}^2 + \left\| \nabla \varepsilon(t_j) \right\|_{[L^2(\mathbb{R}^d)]^d}^2 \right)^{1/2}. \end{aligned}$$

It remains to estimate the error terms $\|\theta\|_{L^2(\mathbb{R}^d)}$ and $\|\nabla \varepsilon\|_{[L^2(\mathbb{R}^d)]^d}$. We start with ε and rewrite it as

$$\varepsilon(t) = \Pi_1 \left[\frac{1}{k} \sum_{j=0}^m \alpha_j u^{\text{inc}}(t - t_j) - \dot{u}^{\text{inc}}(t) \right] + (\Pi_0 - \Pi_1) [\dot{u}^{\text{inc}}(t)].$$

Due to the H^1 stability of Π_1 and the approximation properties of Π_0 and Π_1 this gives for the norm of the gradient:

$$\begin{aligned} \|\nabla \varepsilon(t)\|_{[L^2(\mathbb{R}^d)]^d} &\lesssim \left\| \frac{1}{k} \sum_{j=0}^m \alpha_j u^{\text{inc}}(t - t_j) - \dot{u}^{\text{inc}}(t) \right\|_{H^1(\mathbb{R}^d \setminus \Gamma)} \\ &\quad + \left\| (\operatorname{Id} - J_{\Gamma}^{Y_h}) \dot{u}^{\text{inc}}(t) \right\|_{H^{1/2}(\Gamma)} + \inf_{x_h \in X_h} \|\partial_n^+ \dot{u}^{\text{inc}}(t) - x_h\|_{H^{-1/2}(\Gamma)}. \end{aligned}$$

The first term is $\mathcal{O}(k^{\min(p, \mu-1)})$ as the consistency error of a p -th order multistep method.

A similar argument can be employed for θ ; noticing that $\Delta \Pi_1 u = \Delta u$ and arranging the terms as above gives

$$\|\theta(t)\|_{L^2(\mathbb{R}^d)} \lesssim \left\| \frac{1}{k} \sum_{j=0}^m \alpha_j \dot{u}^{\text{inc}}(t - t_j) - \Delta u^{\text{inc}}(t) \right\|_{L^2(\mathbb{R}^d)} + \|(\operatorname{Id} - \Pi_0) \Delta u^{\text{inc}}(t)\|_{L^2(\mathbb{R}^d)}.$$

Since we assumed that u^{inc} solves the wave equation, we can recognize the first term as another consistency error (this time for \ddot{u}), and replace Δu^{inc} with \ddot{u}^{inc} . This means, we can estimate:

$$\|\theta(t)\|_{L^2(\mathbb{R}^d)} \leq \mathcal{O}\left(k^{\min(p, \mu-1)}\right) + C \sum_{j=0}^m \left\| (\text{Id} - \Pi_0) \ddot{u}^{\text{inc}}(t_n) \right\|_{L^2(\mathbb{R}^d)},$$

which, together with the approximation properties of Π_0 and Π_1 , concludes the proof. \square

Summarizing the previous results, we get the following convergence for a pure time-discretization:

Theorem 5.27. *Assume for a moment that $X_h = H^{-1/2}(\Gamma)$ and $Y_h = H^{1/2}(\Gamma)$. The discrete solutions, obtained by $u_{ie}^{h,k} := S(\partial_t^k)\varphi^k + (\partial_t^k)^{-1}D(\partial_t^k)\psi^k$ converge to the exact solution u of (5.17) with the rate*

$$\begin{aligned} \max_{nk \leq T} \left\| \nabla u(t_n) - \nabla u_{ie}^{h,k}(t_n) - \nabla u^{\text{inc}}(t_n) \right\|_{[L^2(\mathbb{R}^d)]^d} &\leq C(u)T k^{1/3}, \\ \max_{nk \leq T} \left\| \dot{u}(t_n) - v_{ie}^{h,k}(t_n) - \dot{u}^{\text{inc}}(t_n) \right\|_{L^2(\mathbb{R}^d)} &\leq C(u)T k^{1/3}. \end{aligned}$$

If we assume that the exact solution satisfies $(\nabla u, \dot{u}) \in C^{p+1}([0, T], \mathcal{X})$, then we regain the full convergence rate of the multistep method

$$\begin{aligned} \max_{nk \leq T} \left\| \nabla u(t_n) - \nabla u_{ie}^{h,k}(t_n) - u^{\text{inc}}(t_n) \right\|_{[L^2(\mathbb{R}^d)]^d} &\leq C(u)T k^p, \\ \max_{nk \leq T} \left\| \dot{u}(t_n) - v_{ie}^{h,k}(t_n) - \dot{u}^{\text{inc}}(t_n) \right\|_{L^2(\mathbb{R}^d)} &\leq C(u)T k^p. \end{aligned}$$

For the fully discrete setting, the same rates in time hold, but with additional projection errors due to u^{inc} ; see Lemma 5.26. All the constants depend on the exact solution u and on u^{inc} but are independent of k .

Proof. This statement is easily obtained by combining Proposition 5.24 with Lemma 5.26. \square

5.4.3 Convergence of the fully discrete scheme

Finally, we are able to answer the question under which conditions the solutions to the fully discrete boundary integral equations (5.8) converge to the exact solutions. This will be answered in Corollary 5.28. We also look at the case of what happens if we make the additional assumption that the exact solution satisfies additional smoothness properties.

The non-smooth case

We start by investigating the case of no assumed regularity. This case can be handled easily by combining the estimates from the previous sections which immediately give a convergence result for the full discretization:

Corollary 5.28. *Assume that the incoming wave satisfies $(\nabla u^{inc}, \dot{u}^{inc}) \in C^\mu((0, T), \mathcal{X})$ for $\mu > 1$. Setting $\tilde{u}_{ie}^{h,k} := u_{ie}^{h,k} + u^{inc}$ and $\tilde{v}_{ie}^{h,k} := v_{ie}^{h,k} + \dot{u}^{inc}$, the discretization error can be quantified by:*

$$\begin{aligned}
 & \left\| \nabla \tilde{u}_{ie}^{h,k}(t_n) - \nabla u(t_n) \right\|_{[L^2(\mathbb{R}^d)]^d} + \left\| \tilde{v}_{ie}^{h,k}(t_n) - \dot{u}(t_n) \right\|_{L^2(\mathbb{R}^d)} \\
 \lesssim & T \left(\|Au(0)\|_{\mathcal{X}} + \|(\nabla u, \dot{u})\|_{C^\mu((0,T), \mathcal{X})} \right) k^{\min(\mu-1, 1/3)} \\
 & + k \sum_{j=0}^n \left[\left\| (\text{Id} - J_\Gamma^{Y_h}) \gamma^+ \dot{u}^{inc}(t_j) \right\|_{H^{1/2}(\Gamma)} + \inf_{x_h \in X_h} \left\| \partial_n^+ \dot{u}^{inc}(t_j) - x_h \right\|_{H^{-1/2}(\Gamma)} \right] \\
 & + k \sum_{j=0}^n \left[\left\| (\text{Id} - J_\Gamma^{Y_h}) \gamma^+ \ddot{u}^{inc}(t_j) \right\|_{H^{1/2}(\Gamma)} + T^{1/2} \left(\int_0^{t_n} \|\dot{\rho}(\tau)\|_{\mathcal{X}}^2 + \|\nabla \rho_v(\tau)\|_{L^2(\mathbb{R}^d)}^2 d\tau \right)^{1/2} \right. \\
 & \left. + T^{1/2} \left(\int_0^{t_n} \|u - \Pi_1 u\|_{L^2(\mathbb{R}^d)}^2 + \beta^{-1/2} \|\theta(\tau)\|_{L^2(\Gamma)}^2 d\tau \right)^{1/2} \right]
 \end{aligned}$$

where the error terms satisfy

$$\begin{aligned}
 \|\dot{\rho}(t)\|_{\mathcal{X}} & \lesssim \inf_{y_h \in Y_h} \|\dot{u} - y_h\|_{H^1(\Omega^-)} + \left\| \dot{u} - J_{\Omega^-}^{Y_h} \mathcal{E}^+ \dot{u} \right\|_{L^2(\Omega^-)} + \inf_{x_h \in X_h} \left\| \partial_\nu^+ \dot{u} - x_h \right\|_{H^{-1/2}(\Gamma)}, \\
 \|\nabla \rho_v(t)\|_{L^2(\mathbb{R}^d)} & \lesssim \left\| \dot{u} - J_{\Omega^-}^{Y_h} \mathcal{E}^+ \dot{u} \right\|_{H^1(\Omega^-)}, \\
 \|\theta(t)\|_{L^2(\Gamma)} & = \|g(\llbracket \gamma v \rrbracket) - g(\llbracket \gamma \Pi_2 v \rrbracket)\|_{L^2(\Gamma)}.
 \end{aligned}$$

Assuming $|g(\mu)| \lesssim \left(1 + |\mu|^{\frac{d-1}{d-2}}\right)$ for $d \geq 3$, this gives strong convergence

$$\begin{aligned}
 \nabla u_{ie}^{h,k} + \nabla u^{inc} & \rightarrow \nabla u \quad \text{in } L^\infty \left((0, T), [L^2(\mathbb{R}^d)]^d \right), \\
 \partial_t^k u_{ie}^{h,k} + \dot{u}^{inc} & \rightarrow \dot{u} \quad \text{in } L^\infty \left((0, T), L^2(\mathbb{R}^d) \right),
 \end{aligned}$$

with a rate in time of $k^{1/3}$ and quasi-optimality in space.

Proof. We just collect all the estimates from the previous sections. The stronger growth condition on g is needed to ensure that the error in the nonlinearity θ converges in $L^2(\Gamma)$ (see Lemma 5.20(ii)). \square

The smooth case

While the general convergence result of Corollary 5.28 is nice, it does not provide any insight into how fast the convergence with respect to the spatial discretization is, and in time we were only able to prove the reduced rate of $\mathcal{O}(k^{1/3})$. Both of these shortcomings can be overcome if we make further assumptions on the regularity of the exact solution. Namely we assume:

Assumption 5.29. *Assume that the exact solution of (5.1) has the following regularity properties:*

- (i) $u \in C^p((0, T), H^1(\Omega^+))$,
- (ii) $\dot{u} \in C^p((0, T), L^2(\Omega^+))$,
- (iii) $\gamma^+ u, \gamma^+ \dot{u} \in L^\infty((0, T), H^s(\Gamma))$,
- (iv) $\partial_n^+ u, \partial_n^+ \dot{u} \in L^\infty((0, T), H^{s-1}(\Gamma))$,
- (v) $\ddot{u} \in L^\infty((0, T), H^s(\Omega^-))$,
- (vi) $\gamma^+ \dot{u} \in L^\infty((0, T) \times \Gamma)$,

for some $s \geq 1/2$. Here p denotes the order of the multistep method that is used.

Remark 5.30. We need the strong requirement of $\gamma^+ \dot{u} \in L^\infty((0, T) \times \Gamma)$ in order to be able to apply Lemma 5.20 (iv). Alternatively, we can make stronger growth assumptions on g' instead and drop this requirement. ■

Since we only made assumptions on the exact solution u , instead of on the continuous-in-time/discrete in space solution of (5.17), we can not use the procedure of splitting the discretization steps into first in space then in time and analyze them separately as was done in 5.28. Instead, we will look at the discretization in space and time separately and repeat the argument of Theorem 5.22 in a time discrete setting. The main tool for this argument will be the G -stability of the linear multistep method used.

Lemma 5.31. Let $u_{sg}^{h,k}, v_{sg}^{h,k}$ denote the sequence of approximations obtained by applying the BDF1 or BDF2 method to (5.17), as defined in (5.25), and let u be the exact solution of (5.1) with $v := \dot{u}$. We will use the finite difference operator ∂_t^k , defined as

$$[\partial_t^k u](t) := \frac{1}{k} \sum_{j=0}^k \alpha_j u(t - t_j)$$

(This is consistent with Definition 3.13 for $K(s) = s$ since we are using a BDF method). We introduce the following error terms:

$$\begin{aligned} \Theta_I(t) &:= \Pi_1 \left([\partial_t^k u](t) - \dot{u}(t) \right), \\ \Theta_{II}(t) &:= \Pi_2 \left([\partial_t^k v](t) - \dot{v}(t) \right), \\ \Theta_{III}(t) &:= (\Pi_1 - \Pi_2) v(t), \\ \Theta_{IV}(t) &:= (\text{Id} - \Pi_2) \dot{v}(t), \\ \Theta_V(t) &:= g(\llbracket \gamma v(t) \rrbracket) - g(\llbracket \gamma \Pi_2 v(t) \rrbracket), \\ \Theta_{VI}(t) &:= (\text{Id} - \Pi_1) u(t). \end{aligned}$$

Then

$$\begin{aligned} & \left\| \begin{pmatrix} v_{sg}^{h,k}(t_n) - v(t_n) \\ \mathbf{w}_{sg}^{h,k}(t_n) - \nabla u(t_n) \end{pmatrix} \right\|_{\mathcal{X}^0}^2 + \beta k \sum_{j=0}^n \left\| \llbracket \gamma(v_{sg}^{h,k}(t_j) - v(t_j)) \rrbracket \right\|_{L^2(\Gamma)}^2 \\ & \lesssim k \sum_{j=0}^n \|\nabla \Theta_I(t_j)\|_{[L^2(\mathbb{R}^d)]^d}^2 + \|\Theta_{II}(t_j)\|_{L^2(\mathbb{R}^d)}^2 + \|\nabla \Theta_{III}(t_j)\|_{[L^2(\mathbb{R}^d)]^d}^2 \\ & \quad + k \sum_{j=0}^n \|\Theta_{IV}(t_j)\|_{L^2(\mathbb{R}^d)}^2 + \|\Theta_V(t_j)\|_{L^2(\Gamma)}^2 + \|\Theta_{VI}(t_j)\|_{L^2(\mathbb{R}^d)}^2. \end{aligned}$$

The implied constant depends only on the parameter α in the definition of Π_1 .

Proof. The proof is fairly similar to the ones of Theorem 5.22 and Lemma 5.26 and many similar terms appear. We define the additional error functions

$$e := \begin{pmatrix} e_{\mathbf{w}} \\ e_v \end{pmatrix} := \begin{pmatrix} \nabla \Pi_1 u - \mathbf{w}_{sg}^{h,k} \\ \Pi_2 v - v_{sg}^{h,k} \end{pmatrix}.$$

The overall strategy of the proof is to substitute e in the defining equation for the multistep method and compute the truncation terms. We then test with e in order to get discrete stability just as we did in Theorem 5.22. The proof becomes technical, due to the many different error terms that appear.

The error $e(t)$ solves the following equation for all fixed times $t > 0$ and for all $q_h \in H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)$ and $z_h \in \mathcal{H}_h$, see (5.22),

$$\begin{aligned} \left(\partial_t^k [e_{\mathbf{w}}], q_h \right)_{[L^2(\mathbb{R}^d)]^d} &= (\nabla e_v, q_h)_{[L^2(\mathbb{R}^d)]^d} + (\nabla (\Theta_I + \Theta_{III}), q_h)_{[L^2(\mathbb{R}^d)]^d} \\ \left(\partial_t^k [e_v], z_h \right)_{L^2(\mathbb{R}^d)} &= - (e_{\mathbf{w}}, \nabla z_h)_{[L^2(\mathbb{R}^d)]^d} - \left\langle g(\llbracket v_{sg}^{h,k} \rrbracket) - g(\llbracket \gamma \Pi_2 v \rrbracket), \llbracket \gamma z_h \rrbracket \right\rangle_{\Gamma} \\ & \quad + (\Theta_{II} + \Theta_{IV} + \alpha \Theta_{VI}, z_h)_{L^2(\mathbb{R}^d)} + \langle \Theta_V, \llbracket \gamma z_h \rrbracket \rangle_{\Gamma}. \end{aligned}$$

By testing with $q_h := \nabla e$, $z_h := e$, we get from the strict monotonicity of g :

$$\begin{aligned} \left\langle \partial_t^k [e], e \right\rangle_{\mathcal{X}} + \beta \|\llbracket \gamma e_v \rrbracket\|_{L^2(\Gamma)}^2 &\leq (\nabla (\Theta_I + \Theta_{III}), e_{\mathbf{w}})_{L^2(\mathbb{R}^d)} + \langle \Theta_V, \llbracket \gamma e_v \rrbracket \rangle_{\Gamma} \\ & \quad + (\Theta_{II} + \Theta_{IV} + \alpha \Theta_{VI}, e_v)_{L^2(\mathbb{R}^d)}. \end{aligned}$$

We write $E^n := (e(t_n), \dots, e(t_{n-m}))^T$ and use the G-stability of the BDF methods (Proposition 3.3) to obtain a lower bound on the left-hand side. Using the Cauchy-Schwarz and Young inequalities on the right-hand side then gives after multiplying by k :

$$\begin{aligned} \|E^n\|_G^2 - \|E^{n-1}\|_G^2 + \beta k \|\llbracket \gamma e_v(t_n) \rrbracket\|_{L^2(\Gamma)}^2 \\ \lesssim k \|\nabla \Theta_I + \nabla \Theta_{III}\|_{[L^2(\mathbb{R}^d)]^d}^2 + k \|\Theta_V\|_{L^2(\Gamma)}^2 + k \|\Theta_{II} + \Theta_{IV} + \alpha \Theta_{VI}\|_{L^2(\mathbb{R}^d)}^2. \end{aligned}$$

Summing over n and noting the equivalence of the G -induced-norm to the standard \mathbb{R}^m norm gives the stated result. \square

We close this section by writing down explicit convergence rates for the approximation instead of the best-approximation properties. In order to do so, we make the following assumptions on the spaces X_h and Y_h :

Assumption 5.32. *Assume that the spaces X_h , Y_h and the operator $J_\Gamma^{Y_h}$ satisfy the following approximation properties*

$$\inf_{x_h \in X_h} \|\varphi - x_h\|_{H^{-1/2}(\Gamma)} \leq Ch^{1/2+\min(q+1,s)} \|\varphi\|_{H_{pw}^s(\Gamma)} \quad \forall \varphi \in H_{pw}^s(\Gamma), \quad (5.26a)$$

$$\left\| \psi - J_\Gamma^{Y_h} \psi \right\|_{H^{1/2}(\Gamma)} \leq Ch^{\min(q+1,s)-1/2} \|\psi\|_{H_{pw}^s(\Gamma)} \quad \forall \psi \in H_{pw}^s(\Gamma), \quad (5.26b)$$

for parameters $h > 0$ and $q \in \mathbb{N}$, with constants that depend only on Γ and q . Assume further that the fictitious space $Y_h^{\Omega^-}$ and the operator $J_{\Omega^-}^{Y_h}$ from Assumption 5.16 also satisfy

$$\left\| u - J_{\Omega^-}^{Y_h} u \right\|_{L^2(\Omega^-)} \leq Ch^{\min(q+1,s)} \|u\|_{H^s(\Omega^-)} \quad \forall u \in H^s(\Omega^-). \quad (5.27)$$

Theorem 5.33. *Let Assumptions 5.29 and 5.32 be satisfied and assume that we use BDF1($p = 1$) or BDF2($p = 2$) discretization in time. Assume either $|g'(s)| \lesssim 1 + |s|^r$ with $r \in \mathbb{N}$ arbitrary for $d = 2$ or $r \leq 1$ for $d = 3$, or assume that $\|\Pi_2 \dot{u}\|_{L^\infty(\mathbb{R}^d)} \leq C$ is uniformly bounded w.r.t. h . Then the following convergence result holds:*

$$\left\| v_{sg}^{h,k}(t_n) - \dot{u}(t_n) \right\|_{L^2(\mathbb{R}^d)} + \left\| \mathbf{w}_{sg}^{h,k}(t_n) - \nabla u(t_n) \right\|_{[L^2(\mathbb{R}^d)]^d} = \mathcal{O} \left(T \left(h^{\min(q+1/2,s)} + k^p \right) \right).$$

The implied constant depends on Γ , g , α , the regularity of u as required in Assumptions 5.29, and the constants from Assumption 5.32.

Proof. We already have proven all the necessary estimates, and only need to combine them with our additional assumptions. We combine Lemma 5.31 with the approximation properties of the operators from Section 5.4.1. Θ_I and Θ_{II} are the local truncation errors of the multistep method and therefore $\mathcal{O}(k^p)$ (the operators Π_1 , Π_2 are stable). To estimate Θ_V , the assumptions on g or Π_2 are such that we can apply Lemma 5.20. \square

Remark 5.34. *The assumptions on the spaces X_h , Y_h are satisfied, if we use standard piecewise polynomials of degree q for Y_h and $q - 1$ for X_h on some triangulation \mathcal{T}_h of Γ (Proposition 2.35). In this case, the approximation property (5.27) holds via the construction from Lemma 5.17 using some standard quasi-interpolation operator from the FEM theory (see Proposition 2.36). The requirements on $\Pi_2 \dot{u}$ can be fulfilled in numerous ways, e.g., in 2D and 3D it can be shown by balancing approximation and inverse estimates, as is for example done in [Tho06, Lemma 13.3]. The same result could also be achieved by replacing Π_2 by some operator which also provides L^∞ -stability like nodal interpolation. \blacksquare*

The case of more general nonlinearities g

Up to now, we only considered nonlinearities, which were strictly monotone. This was needed in order to get control over the boundary values of v . Since many authors, e.g., [Gra12] treat this assumption as optional, we investigate in what sense we can recover the

convergence results of the previous sections. We first generalize Theorem 5.22 to the new setting, where the lack of strict monotonicity requires us to weaken the convergence and drop the explicit error estimates. Instead we get:

Lemma 5.35. *Assume the families of spaces $(X_h)_{h>0}$ and $(Y_h)_{h>0}$ are dense in $H^{-1/2}(\Gamma)$ and $H^{1/2}(\Gamma)$ respectively. Assume that g satisfies Assumption 5.1(i)–(v) but is not necessarily strictly monotone.*

Then, for $h \rightarrow 0$, the sequence of solutions $\mathbf{w}^h(t), v^h(t)$ of (5.21) converges weakly towards the solution of (5.22) for almost all $t \in (0, T)$.

Proof. We fix $t \in (0, T]$; all the arguments hold only almost everywhere w.r.t t , which is sufficient to obtain the stated result. Since $g(\llbracket \gamma v^h \rrbracket) \cdot \llbracket \gamma v^h \rrbracket \geq 0$, testing (5.21) with $q_h = \mathbf{w}^h(t)$ and $z_h = v^h(t)$ in (5.21) gives

$$\left\| \left(\mathbf{w}^h(t), v^h(t) \right) \right\|_{\mathcal{X}} \leq \left\| \left(\mathbf{w}^h(0), v^h(0) \right) \right\|_{\mathcal{X}}.$$

By the Eberlein–Šmulian theorem, see [Yos80, page 141], the sequence $(\mathbf{w}^h(t), v^h(t))$ has a weakly convergent sub-sequence; for ease of notation again denoted by $(\mathbf{w}^h(t), v^h(t))$, uniqueness of the solution will give convergence of the whole sequence anyway. We write $(\mathbf{w}(t), v(t))$ for the weak limit. Since the convergence is with respect to the spatial discretization, it is easy to see that $\dot{\mathbf{w}}^h \rightharpoonup \dot{\mathbf{w}}$ and $\dot{v}^h \rightharpoonup \dot{v}$.

As the incident wave vanishes at the scatterer for $t \leq 0$, we get for the initial condition $(u^{\text{inc}}(0), \dot{u}^{\text{inc}}(0)) \in \text{dom}(A)$. The a priori estimate in Proposition 2.17 then implies

$$\left\| \left(\dot{\mathbf{w}}^h(t), \dot{v}^h(t) \right) \right\|_{\mathcal{X}} \leq C(u^{\text{inc}}).$$

Since $\dot{\mathbf{w}}^h = \nabla v^h$, this implies that $\|v^h(t)\|_{H^1(\mathbb{R}^d \setminus \Gamma)}$ is uniformly bounded and therefore (up to another sub-sequence) the trace also converges: $\llbracket v^h(t) \rrbracket \rightharpoonup \llbracket \gamma v(t) \rrbracket$ in $H^{1/2}(\Gamma)$. It was already shown in Lemma 5.20(i) that this implies $g(\llbracket v^h(t) \rrbracket) \rightharpoonup g(\llbracket \gamma v(t) \rrbracket)$. Overall, we have shown that there exists a sub-sequence of (\mathbf{w}^h, v^h) , which converges weakly to a solution of (5.22). Since the same argument can be applied to each sub-sequence of (\mathbf{w}^h, v^h) and the limit is unique, due to the uniqueness from Proposition 2.17, we get that the full series converges weakly. \square

The convergence of the time-discretization in Proposition 5.24 does not depend on the strong monotonicity of g . This insight immediately gives the following corollary:

Corollary 5.36. *Assume the families of spaces $(X_h)_{h>0}$ and $(Y_h)_{h>0}$ are dense in $H^{-1/2}(\Gamma)$ and $H^{1/2}(\Gamma)$ respectively and assume that the operator $J_{\Gamma}^{Y_h}$ converges strongly to the identity, i.e., for all $y \in H^{1/2}(\Gamma)$ we have that $J_{\Gamma}^{Y_h} y \rightarrow y$ converges for $h \rightarrow 0$.*

Let $\tilde{u}_{ie}^{h,k}, \tilde{v}_{ie}^{h,k}$ again be defined using the representation formula and adding back u^{inc} (see Lemma 5.11; we note that this is defined for all $t \in \mathbb{R}$ by using the staggered initial conditions). Then these approximations converge weakly towards the solution of (5.1) for almost all $t \in (0, T)$ for $k \rightarrow 0$ and $h \rightarrow 0$.

Proof. Inspecting the proof of Lemma 5.26 and Theorem 5.27, we did not require g to be strictly monotone. This means that $\nabla \tilde{u}_{ie}^{h,k}$ and $\tilde{v}_{ie}^{h,k}$ converge (strongly) to \mathbf{w}^h and v^h respectively. The stated result then follows directly from Lemma 5.35. \square

Convergence of the boundary traces

In the previous sections, we have looked at convergence of the solutions obtained by applying the representation formula to the boundary data. Since it is hard to compute the norms of these functions on the unbounded domain Ω^+ in practice, we look at what these convergence results imply for the boundary traces themselves. To prove convergence of the traces, we first need the following simple result.

Lemma 5.37. *Let $f \in C^r((0, T), \mathcal{X})$, $\tilde{f} \in C((0, T), \mathcal{X})$ for some Banach space \mathcal{X} and $T \in \mathbb{R}_+$ with $0 = f(0) = f'(0) = \dots = f^{(r-1)}(0) = \tilde{f}(0)$ and $r \leq p$, where p is the order of the multistep method used. Then*

$$\max_{t \in [0, T]} \left\| \partial_t^{-1} f(t) - \left[\left(\partial_t^k \right)^{-1} \tilde{f} \right] (t) \right\|_{\mathcal{X}} \leq Ct \left[k^r \max_{\tau \in [0, t]} \|f^r(\tau)\|_{\mathcal{X}} + \max_{\tau \in [0, t]} \|f(\tau) - \tilde{f}(\tau)\|_{\mathcal{X}} \right].$$

Proof. We split the error into two terms by writing

$$\left\| \partial_t^{-1} f - \left(\partial_t^k \right)^{-1} \tilde{f} \right\|_{\mathcal{X}} \leq \left\| \partial_t^{-1} f - \left(\partial_t^k \right)^{-1} f \right\|_{\mathcal{X}} + \left\| \left(\partial_t^k \right)^{-1} (f - \tilde{f}) \right\|_{\mathcal{X}}.$$

The stated estimate then follows from the standard theory of convolution quadrature; see [Lub88a, Theorem 3.1], noting that ∂_t^{-1} is a sectorial operator. \square

This allows us to prove convergence estimates for ψ^k and φ^k .

Lemma 5.38. *Let u solve (6.3), write $u^{\text{scat}}(t) := u(t) - u^{\text{inc}}(t)$, and define the continuous traces $\psi(t) := \gamma^+ \dot{u}^{\text{scat}}(t)$ and $\varphi(t) := -\partial_n^+ u^{\text{scat}}$. Let $u_{\text{ie}}^{h,k}, v_{\text{ie}}^{h,k}$ solve (5.14) with corresponding traces $\varphi^k := -\llbracket \partial_\nu u_{\text{ie}}^{h,k} \rrbracket$, $\psi^k := \llbracket \gamma v_{\text{ie}}^{h,k} \rrbracket$ solving (5.8). Then*

$$\left\| \begin{pmatrix} \partial_t^{-1} \psi - \left[\left(\partial_t^k \right)^{-1} \psi^k \right] \\ \partial_t^{-1} \varphi - \left[\left(\partial_t^k \right)^{-1} \varphi^k \right] \end{pmatrix} \right\|_{\Gamma} \lesssim \left\| \nabla u^{\text{scat}} - \nabla u_{\text{ie}}^{h,k} \right\|_{L^2(\mathbb{R}^d)} + \left\| \dot{u}^{\text{scat}} - v_{\text{ie}}^{h,k} \right\|_{L^2(\mathbb{R}^d)} + \mathcal{O}(k^r)$$

holds for all $t \leq T$, with constants that depend only on Γ and the end time T , and where r is the smaller of the two regularity indices of $\psi(t)$ and $\varphi(t)$.

Proof. From the definition of ψ and the properties of the operational calculus, we have that

$$\left(\partial_t^k \right)^{-1} \psi_{\text{ie}} = \left(\partial_t^k \right)^{-1} \llbracket \gamma v_{\text{ie}}^{h,k} \rrbracket = \left(\partial_t^k \right)^{-1} \partial_t^k \llbracket \gamma u_{\text{ie}}^{h,k} \rrbracket = \llbracket \gamma u_{\text{ie}}^{h,k} \rrbracket$$

and analogously $\partial_t^{-1} \psi = \llbracket \gamma u^{\text{scat}} \rrbracket$. The standard trace theorem then gives the estimate

$$\left\| \partial_t^{-1} \psi(t_n) - \left(\partial_t^k \right)^{-1} \psi^k(t_n) \right\|_{H^{1/2}(\Gamma)} \leq C \left\| u^{\text{scat}} - u_{\text{ie}}^{h,k} \right\|_{H^1(\mathbb{R}^d \setminus \Gamma)}.$$

We can further estimate the L^2 contribution in the norm above by noting

$$\left\| u^{\text{scat}}(t_n) - u_{\text{ie}}^{h,k}(t_n) \right\|_{L^2(\mathbb{R}^d)} = \left\| \partial_t^{-1} \dot{u}^{\text{scat}}(t_n) - \left(\partial_t^k \right)^{-1} v_{\text{ie}}^{h,k}(t_n) \right\|_{L^2(\mathbb{R}^d)}$$

and applying Lemma 5.37.

Since $u_{\text{ie}}^{h,k}$ solves (5.14) and Δ is linear, we note that

$$\Delta \left(\partial_t^k \right)^{-1} u_{\text{ie}}^{h,k} = \left(\partial_t^k \right)^{-1} \Delta u_{\text{ie}}^{h,k} = \left(\partial_t^k \right)^{-1} \partial_t^k v_{\text{ie}}^{h,k} = v_{\text{ie}}^{h,k},$$

and analogously $\Delta \left(\partial_t^{-1} u^{\text{scat}} \right) = \dot{u}^{\text{scat}}$.

From the definition of φ and the stability of the normal trace operator in $H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)$ we have

$$\begin{aligned} & \left\| \partial_t^{-1} \varphi(t_n) - \left(\partial_t^k \right)^{-1} \varphi^k(t_n) \right\|_{H^{-1/2}(\Gamma)} \\ & \leq C \left\| \partial_t^{-1} u^{\text{scat}}(t_n) - \left(\partial_t^k \right)^{-1} u_{\text{ie}}^{h,k}(t_n) \right\|_{H_{\Delta}^1(\mathbb{R}^d \setminus \Gamma)} \\ & \leq C \left[\left\| \partial_t^{-1} u^{\text{scat}}(t_n) - \left(\partial_t^k \right)^{-1} u_{\text{ie}}^{h,k}(t_n) \right\|_{H^1(\mathbb{R}^d \setminus \Gamma)} + \left\| \dot{u}^{\text{scat}}(t_n) - v_{\text{ie}}^{h,k}(t_n) \right\|_{L^2(\mathbb{R}^d)} \right]. \end{aligned}$$

We apply Lemma 5.37 twice to estimate the H^1 -term of the integral by the L^2 -norm of the gradient and to estimate the L^2 -norm of the derivative up to higher order error terms. \square

5.5 Numerical results

We again implemented the methods proposed in this chapter in order to compare the performance of the method to the theoretical findings. Since the problem is nonlinear, we need to solve the system in a time-stepping fashion. This is achieved in an efficient way by the FFT-based algorithm in [Ban10]. For the boundary integral operators we rely on the BEM++ software library [Šmi+15].

A scalar example

We start by a simple example, which allows us to efficiently compute the approximation. We follow the ideas by [SV14] and consider the case of scattering by a spatially homogeneous incident wave by the unit sphere. Since the constant functions are eigenfunctions of the boundary operators, we can replace them by multiplication with the eigenvalue and the resulting scattered wave is also constant in space. This means that we can reduce the 3D problem to the scalar case. Since an explicit expression for the exact solution is not known, we compute the numerical approximation to a high degree of accuracy and compare it to the previous approximations, i.e., the reference solution was computed with at least twice the number of timesteps compared to the approximate solutions.

Example 5.39. Let $g(\mu) := \frac{1}{4}\mu + |\mu|\mu$ and consider the incident wave

$$u^{\text{inc}}(x, t) := \cos(2\pi t) e^{-(t-t_0)/\alpha},$$

where we set $t_0 := \pi/2$, $\alpha = 1/5$. In Figure 5.1, we see that the exact solution is not smooth. The first derivative has several kinks and the second derivative has singularities.

Nevertheless, the convergence graphs of Figure 5.2 show that the numerical method gives optimal convergence rates, not only in the integrated norm but also for ψ itself.

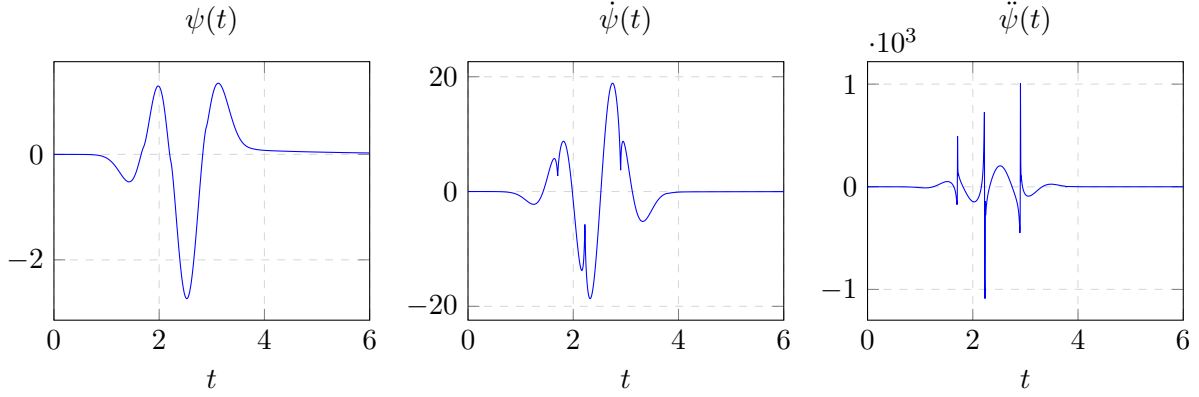


Figure 5.1: Exact solution to Example 5.39

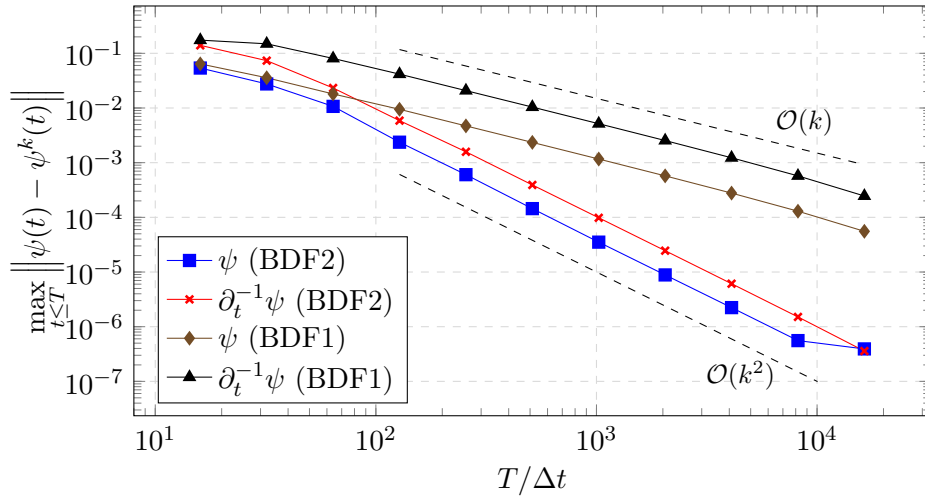


Figure 5.2: Convergence rates for Example 5.39

Scattering from a non-convex domain

In order to test our method on a more realistic problem, we consider the scattering from a three-dimensional non-convex domain, as was also used in [Ban10].

Example 5.40. We chose $g(\mu) := \frac{\mu}{4} + |\mu|\mu$, and $u^{inc}(x, t) := \phi(p \cdot x - t)$ with $p := (1, 0, 0)^T$ and

$$\phi(t) := \cos(\omega t) e^{-\frac{(t-t_0)^2}{\sigma^2}}.$$

We chose $t_0 := 3$, $\sigma := 0.5$ and $\omega := 4\pi$, computing up to an end-time $T = 10$. We applied a BDF2 method for the Convolution Quadrature and chose a fixed mesh consisting of 4074

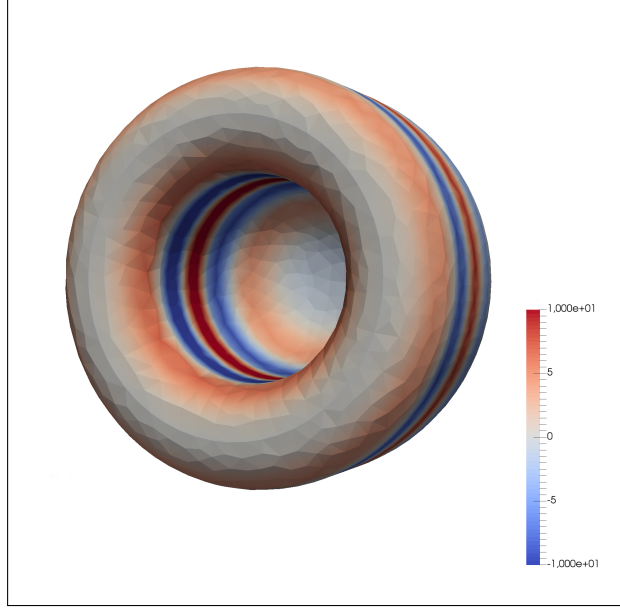


Figure 5.3: Geometry for Example 5.40

vertices and 8688 triangles. We used lowest order BEM spaces and computed the difference to the reference solution, which was computed by using 1024 timesteps.

In Figure 5.4 we see that in this case the method does not provide us with the optimal convergence rate, but instead the rate appears to be reduced to linear convergence.

Since it is difficult to determine the asymptotic behavior for the full 3D problem, we return to the case of scattering from a sphere. In order to capture the difficulty of the non-convex domain, we consider the case of a “completely trapping sphere”, i.e., we assume that the wave starts inside the sphere and has no way to escape. Mathematically, this means we solve the boundary integral equations

$$\begin{pmatrix} \partial_t^k V(\partial_t^k) & -K(\partial_t^k) - \frac{1}{2} \\ \frac{1}{2} + K^T(\partial_t^k) & (\partial_t^k)^{-1} W(\partial_t^k) \end{pmatrix} \begin{pmatrix} \varphi \\ \psi \end{pmatrix} + \begin{pmatrix} 0 \\ g(\psi + i^{inc}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We use a similar model problem to Example 5.40, namely $g(\mu) := \frac{\mu}{4} + |\mu| \mu$ and an incoming wave of the form $u^{inc}(x, t) := \phi(t)$, with the same parameters $t_0 := 3$, $\sigma = 0.5$ and $\omega = 4\pi$. Figure 5.5 shows that the BDF2 method no longer offers the full convergence rate of $\mathcal{O}(k^2)$.

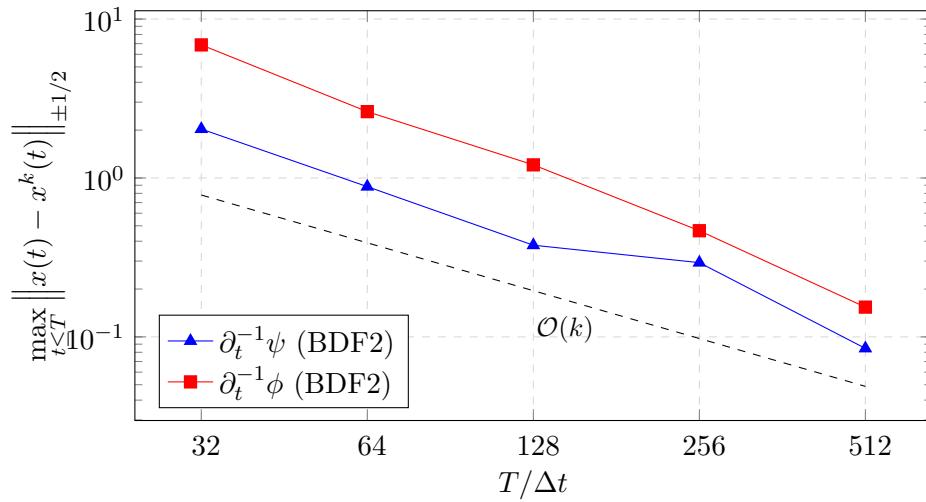


Figure 5.4: Convergence for Example 5.40

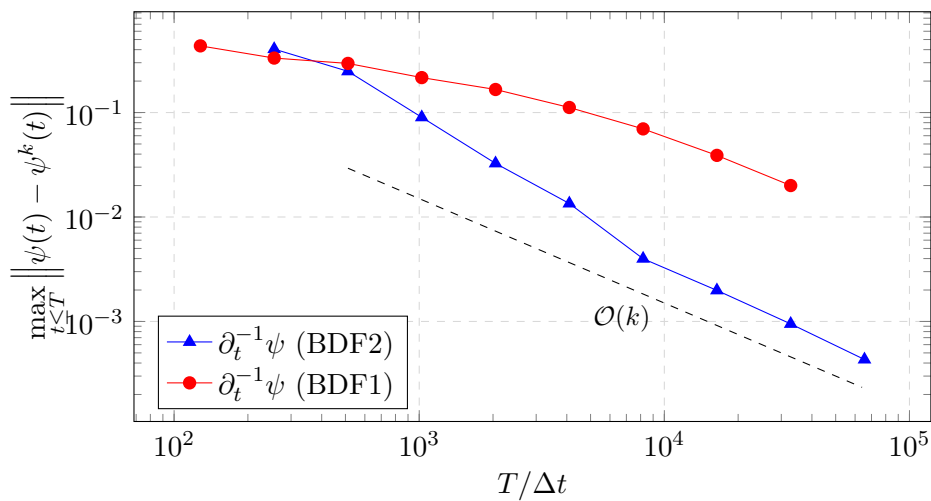


Figure 5.5: Convergence rates for the scattering by the completely trapping sphere

6 Scattering by composite media

In this chapter, we look at a scattering problem, where the scatterer is not made up of a homogeneous material, but instead consists of layers of materials with different wave propagation speeds. Mathematically, this means we deal with the wave equation with piecewise constant coefficients. One way to handle these difficulties would be to use a coupling of Finite Element and Boundary Element Methods, similar to what we did in Chapter 4 when discussing the Schrödinger equation. This approach was for example taken in [HS16]. An alternative view, which will be the approach taken in this chapter, is to do a pure Boundary Element formulation of the problem. This is achieved by considering separate scattering problems on subdomains, where the coefficient functions are constant and enforcing continuity conditions across subdomain boundaries. In the setting of time harmonic scattering, this formulation was pioneered in [CS85] for the case of two subdomains and in more generality in [Pet89]. There is a multitude of possible formulations for the boundary integral equations, based on both a single- and multitrace formulation and boundary integral equations of the first and second kind, see, e.g., [Cla11; CH13; HJ12]. We restrict our considerations to a singletrace approach of first kind.

In the case of time-domain scattering, the case of two domains and of “nested domains”, in which domains are layered within each other, not allowing multiple domains to touch in a single point was investigated in [QS16] and [Qiu16] respectively. Our work generalizes these results in a unified framework by working in abstract spaces, which reduces to the specific constructions of the Costabel-Stephan system in the two domain case. The results of this chapter are part of a joint work with Francisco-Javier Sayas. This chapter also serves as a showcase for the power of the results on general Runge-Kutta approximation of semigroups as presented in Section 3.4.

6.1 Model problem

The scattering problem we are interested in is given by Lipschitz domains $\Omega_1, \dots, \Omega_L$ for $L \in \mathbb{N}$, where we assume that the domains are pairwise disjoint and bounded. We define the exterior and the skeleton as

$$\Omega_0 := \mathbb{R}^d \setminus \bigcup_{\ell=0}^L \overline{\Omega_\ell} \quad \text{and} \quad \Gamma := \bigcup_{\ell=1}^L \partial\Omega_\ell.$$

The material properties of the scatterer, determined by the functions $\kappa, c: \mathbb{R}^d \rightarrow \mathbb{R}$, are assumed to be piecewise constant and positive, i.e., we write

$$\kappa|_{\Omega_\ell} \equiv \kappa_\ell > 0 \quad \text{and} \quad c|_{\Omega_\ell} \equiv c_\ell > 0, \quad \ell = 0, \dots, L,$$

for coefficients $\kappa_\ell, c_\ell \in \mathbb{R}$.

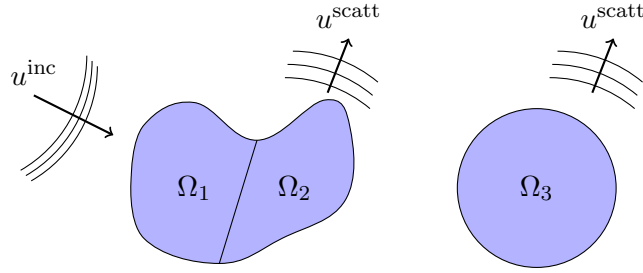


Figure 6.1: Example of geometric situation

We are interested in solutions of the wave equation $u^{\text{tot}} \in C^2(\mathbb{R}_+, H_{\text{loc}}^1(\mathbb{R}^d))$ satisfying:

$$c^{-2}\ddot{u}^{\text{tot}}(t) = \text{div}(\kappa\nabla u^{\text{tot}}(t)), \quad (6.1a)$$

$$u^{\text{tot}}(t) = u^{\text{inc}}(t) \quad \text{for } t \leq 0, \quad (6.1b)$$

where u^{inc} is a given initial condition. Just like in Chapter 5, we make the additional assumption that u^{inc} is an incoming wave, as formalized in the next assumption.

Assumption 6.1. *The incident wave u^{inc} satisfies:*

(i) u^{inc} solves the wave equation in the exterior:

$$c_0^{-2}\ddot{u} = \kappa_0\Delta u \quad \text{in } \Omega_0, \quad (6.2)$$

(ii) $\text{supp } u^{\text{inc}}(t) \subseteq \Omega_0$ for $t \leq 0$.

For notational convenience, we define $u^{\text{inc}}(x, t) := 0$ for $x \in \mathbb{R}^d \setminus \overline{\Omega_0}$. Just like in Chapter 5, we make the decompositional ansatz $u^{\text{tot}} = u + u^{\text{inc}}$. In order to get a convenient formulation for the continuity conditions across interfaces, let $B \subseteq \mathbb{R}^d$ be a sufficiently large ball containing the skeleton, i.e., assume that $\Omega_\ell \subset B$ for all $\ell = 1, \dots, L$. This gives us the model problem: find $u \in C^2(\mathbb{R}, H^1(\mathbb{R}^d \setminus \Gamma))$, such that:

$$\ddot{u}(t) = \text{div}(\kappa\nabla u(t)) \quad \text{in } \mathbb{R}^d \setminus \Gamma, \quad (6.3a)$$

$$u(t) - u^{\text{inc}}(t) \in H^1(B), \quad (6.3b)$$

$$\kappa\nabla(u(t) - u^{\text{inc}}(t)) \in H(\text{div}, B), \quad (6.3c)$$

and initial condition $u(t) = \dot{u}(t) = 0$ for $t \leq 0$. We are now interested in deriving a boundary integral based discretization scheme for (6.3).

6.2 A multiply overlapped wave problem

In this section, we reformulate (6.3) in a way that makes it more easily treatable by boundary integral methods. We will do this by splitting the equation (6.3) into subproblems on

the domains Ω_ℓ and imposing the jump conditions in a clever way. This means we will end up with $L+1$ fields u_0, \dots, u_L with $u_\ell \approx u|_{\Omega_\ell}$. In order to derive this formulation we need to introduce some additional notation, which mainly extends the definitions of Section 2.3.2 to the multi-domain case. For $\ell = 0, \dots, L$, we define the traces, jumps and means

$$\gamma_\ell^{\text{int}}, \gamma_\ell^{\text{ext}}, \llbracket \gamma_\ell \cdot \rrbracket, \{\{\gamma_\ell \cdot\}\} : H^1(\mathbb{R}^d \setminus \partial\Omega_\ell) \rightarrow H^{1/2}(\partial\Omega_\ell),$$

where the interior and exterior traces are self-explanatory (note that for the boundary of unbounded domain Ω_0 , the interior trace is taken from Ω_0 , i.e., it corresponds to the external trace of the scatterer) and

$$\llbracket \gamma_\ell u \rrbracket := \gamma_\ell^{\text{ext}} u - \gamma_\ell^{\text{int}} u, \quad \{\{\gamma_\ell u\}\} := \frac{1}{2}(\gamma_\ell^{\text{int}} u + \gamma_\ell^{\text{ext}} u).$$

Similarly, we introduce the normal traces

$$\gamma_{\nu,\ell}^{\text{int}}, \gamma_{\nu,\ell}^{\text{ext}}, \llbracket \gamma_{\nu,\ell} \cdot \rrbracket, \{\{\gamma_{\nu,\ell} \cdot\}\} : H(\text{div}, \mathbb{R}^d \setminus \partial\Omega_\ell) \rightarrow H^{-1/2}(\partial\Omega_\ell),$$

where we note that the normal is always taken to point out of the domain; most notably for Ω_0 it points into the scatterer.

The fields we consider live in a certain class of product spaces. We define:

$$\mathcal{H}^{\text{div}} := \prod_{\ell=0}^L H(\text{div}, \mathbb{R}^d \setminus \partial\Omega_\ell), \quad \mathcal{H}^1 := \prod_{\ell=0}^L H^1(\mathbb{R}^d \setminus \partial\Omega_\ell), \quad (6.4)$$

$$\mathcal{H}^{-1/2} := \prod_{\ell=0}^L H^{-1/2}(\partial\Omega_\ell), \quad \mathcal{H}^{1/2} := \prod_{\ell=0}^L H^{1/2}(\partial\Omega_\ell), \quad (6.5)$$

all endowed with the corresponding product norms. On these spaces, we define the diagonal trace operators

$$\begin{aligned} \gamma^{\text{int}}, \gamma^{\text{ext}}, \llbracket \gamma \cdot \rrbracket, \{\{\gamma \cdot\}\} &: \mathcal{H}^1 \rightarrow \mathcal{H}^{1/2}, \\ \gamma_\nu^{\text{int}}, \gamma_\nu^{\text{ext}}, \llbracket \gamma_\nu \cdot \rrbracket, \{\{\gamma_\nu \cdot\}\} &: \mathcal{H}^{\text{div}} \rightarrow \mathcal{H}^{-1/2}. \end{aligned}$$

We make the convention that $\langle \cdot, \cdot \rangle_\Gamma$ denotes the extension of the L^2 product to $\mathcal{H}^{-1/2} \times \mathcal{H}^{1/2}$, while $\langle \cdot, \cdot \rangle_\ell$ denotes the same on $H^{-1/2}(\partial\Omega_\ell) \times H^{1/2}(\partial\Omega_\ell)$. Since we will often be working with pairs of functions on the boundary, we define the product norm on $\mathcal{H}^{-1/2} \times \mathcal{H}^{1/2}$ by $\|(\boldsymbol{\lambda}, \boldsymbol{\psi})\|_\Gamma^2 := \|\boldsymbol{\lambda}\|_{\mathcal{H}^{-1/2}}^2 + \|\boldsymbol{\psi}\|_{\mathcal{H}^{1/2}}^2$.

In order to enforce continuity between the different components across the interfaces $\partial\Omega_\ell \cap \partial\Omega_k$, we follow ideas by [Pet89] (but use the notation by [CH13]) and introduce the single-trace spaces:

$$\begin{aligned} \mathcal{Y} &:= \left\{ (\gamma_\ell^{\text{int}} u)_{\ell=0}^L : u \in H^1(\mathbb{R}^d) \right\} \\ &= \left\{ \boldsymbol{\psi} \in \mathcal{H}^{1/2} : \exists u \in H^1(\mathbb{R}^d) : \boldsymbol{\psi} = \boldsymbol{\gamma}^{\text{int}} u \right\} \end{aligned} \quad (6.6a)$$

$$\begin{aligned} \mathcal{X} &:= \left\{ (\gamma_{\nu,\ell}^{\text{int}} \boldsymbol{v})_{\ell=0}^L : \boldsymbol{v} \in H(\text{div}, \mathbb{R}^d) \right\} \\ &= \left\{ \boldsymbol{\lambda} \in \mathcal{H}^{-1/2} : \exists \boldsymbol{v} \in H(\text{div}, \mathbb{R}^d) : \boldsymbol{\lambda} = \boldsymbol{\gamma}_\nu^{\text{int}} \boldsymbol{v} \right\}. \end{aligned} \quad (6.6b)$$

(Note: we have committed a slight abuse of notation by applying the diagonal trace operators to a single function, this is to be understood as the function copied $L + 1$ times.) Since the trace operators are bounded, it is easy to see that \mathcal{X} and \mathcal{Y} are closed subspaces of $\mathcal{H}^{-1/2}$ and $\mathcal{H}^{1/2}$ respectively.

In order to derive a discretization scheme with respect to the space variable, we consider closed subspaces $\mathcal{X}_h \subseteq \mathcal{X}$ and $\mathcal{Y}_h \subseteq \mathcal{Y}$. We will later present possible ways to construct such spaces in Section 6.3.1. The annihilator spaces take the following form

$$\begin{aligned}\mathcal{X}_h^\circ &= \left\{ \boldsymbol{\psi} \in \mathcal{H}^{1/2} : \langle \boldsymbol{\mu}, \boldsymbol{\psi} \rangle_\Gamma = 0 \ \forall \boldsymbol{\mu} \in \mathcal{X}_h \right\}, \\ \mathcal{Y}_h^\circ &= \left\{ \boldsymbol{\lambda} \in \mathcal{H}^{-1/2} : \langle \boldsymbol{\lambda}, \boldsymbol{\chi} \rangle_\Gamma = 0 \ \forall \boldsymbol{\chi} \in \mathcal{Y}_h \right\}.\end{aligned}$$

(Note that we are using the annihilators with respect to the full spaces $\mathcal{H}^{\pm 1/2}$ instead of the single trace spaces \mathcal{X} and \mathcal{Y}). In the case that $\mathcal{X}_h = \mathcal{X}$ and $\mathcal{Y}_h = \mathcal{Y}$, it can be shown (see [Cla11, Proposition 2.1]) that $\mathcal{X}^\circ = \mathcal{Y}$ and $\mathcal{Y}^\circ = \mathcal{X}$, most notably, we have

$$\langle \boldsymbol{\lambda}, \boldsymbol{\psi} \rangle_\Gamma = 0 \quad \text{for } \boldsymbol{\lambda} \in \mathcal{X} \text{ and } \boldsymbol{\psi} \in \mathcal{Y}. \quad (6.7)$$

When working with the spaces, we often need to change our point of view from the “global field satisfying jump conditions” to the “family of fields on each subdomain”. This is formalized in the next lemma:

Lemma 6.2 (Restricting and gluing). *Let $U = (u_\ell)_{\ell=0}^L \in \mathcal{H}^1$ satisfy $[\![\gamma U]\!] \in \mathcal{Y}$ and $\gamma^{\text{ext}} U \in \mathcal{Y}$. Then $u : \mathbb{R}^d \rightarrow \mathbb{R}$, defined by $u|_{\Omega_\ell} := u_\ell|_{\Omega_\ell}$ is in $H^1(\mathbb{R}^d)$. Similarly, for $\mathbf{V} = (\mathbf{v}_\ell)_{\ell=0}^L \in \mathcal{H}^{\text{div}}$ satisfying $[\![\gamma_\nu \mathbf{V}]\!] \in \mathcal{X}$ and $\gamma_\nu^{\text{ext}} \mathbf{V} \in \mathcal{X}$, the function $\mathbf{v} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ defined as $\mathbf{v}|_{\Omega_\ell} := \mathbf{v}_\ell|_{\Omega_\ell}$ is in $H(\text{div}, \mathbb{R}^d)$.*

Proof. The conditions imply $\gamma^{\text{int}} U \in \mathcal{Y}$ and $\gamma_\nu^{\text{int}} \mathbf{V} \in \mathcal{X}$. The statements then follow from the definition of \mathcal{X} , \mathcal{Y} as they enforce that the traces of the functions defined domain-wise match up. \square

Using these preparatory results, we can now present a new formulation of the model problem as a multiply overlapped transmission problem, which will be more amendable to discretization via boundary integral methods and the analysis thereof. (Note that the choice of transmission conditions is motivated by the Galerkin discretization of the underlying boundary integral equations. The reasons for these particular conditions will become apparent in Section 6.3.)

Problem 6.3. *Given boundary data $\boldsymbol{\beta}^0 : [0, \infty) \rightarrow \mathcal{H}^{1/2}$ and $\boldsymbol{\beta}^1 : [0, \infty) \rightarrow \mathcal{H}^{-1/2}$, find $U^h =: (u_\ell^h)_{\ell=0}^L : [0, \infty) \rightarrow \mathcal{H}^1$ and $\mathbf{V}^h =: (\mathbf{v}_\ell^h)_{\ell=0}^L : [0, \infty) \rightarrow \mathcal{H}^{\text{div}}$ such that:*

(i) *the individual fields satisfy the wave equation in its first order form:*

$$\dot{u}_\ell^h = c_\ell^2 \text{div}(\mathbf{v}_\ell^h), \quad \dot{\mathbf{v}}_\ell^h = \kappa_\ell \nabla u_\ell^h \quad \text{in } \mathbb{R}^d \setminus \partial\Omega_\ell \text{ for } \ell = 0, \dots, L, \quad (6.8a)$$

(ii) *the following trace relations hold for $t \geq 0$:*

$$[\![\gamma U^h]\!](t) + \boldsymbol{\beta}^0(t) \in \mathcal{Y}_h, \quad [\![\gamma_\nu \mathbf{V}^h]\!](t) + \boldsymbol{\beta}^1(t) \in \mathcal{X}_h, \quad (6.8b)$$

$$\gamma^{\text{ext}} U^h(t) \in \mathcal{X}_h^\circ, \quad \gamma_\nu^{\text{ext}} \mathbf{V}^h(t) \in \mathcal{Y}_h^\circ, \quad (6.8c)$$

(iii) the functions satisfy homogeneous initial conditions:

$$U^h(0) = 0 \quad \mathbf{V}^h(0) = 0. \quad (6.8d)$$

The relationship between (6.8) and (6.3) is laid out in the next lemma. As a short remark on notation, we write $\mathbb{1}_{\Omega_\ell}$ for the characteristic function of the set Ω_ℓ and $\partial_t^{-1}u := \int_0^t u(\tau) d\tau$ for the integral of a function u .

Lemma 6.4. *Let $\beta^0 := (\gamma_0^{\text{int}} u^{\text{inc}}, 0, \dots, 0)$ and $\beta^1 := (\kappa_0 \gamma_{\nu,0}^{\text{int}} \nabla \partial_t^{-1} u^{\text{inc}}, 0, \dots, 0)$, and take $\mathcal{X}_h = \mathcal{X}$ and $\mathcal{Y}_h = \mathcal{Y}$. Then the following equivalence holds:*

(i) *If (U^h, \mathbf{V}^h) solves Problem 6.3, then $u : [0, \infty) \rightarrow H^1(\mathbb{R}^d \setminus \Gamma)$ defined as $u(t)|_{\Omega_\ell} := u_\ell^h(t)|_{\Omega_\ell}$ solves (6.3).*

(ii) *If u solves (6.3), then the fields U^h, \mathbf{V}^h , defined via*

$$u_\ell^h := u \mathbb{1}_{\Omega_\ell} \quad \text{and} \quad v_\ell^h := \partial_t^{-1} \nabla u_\ell^h,$$

solve Problem 6.3.

Proof. Follows by inspection and Lemma 6.2. □

6.3 Time domain boundary integral equations

In this section, we relate Problem 6.3 to an equivalent system of boundary integral equations, which is better suited for practical computations. We will then discretize these using a Runge-Kutta convolution quadrature approach. The boundary integral equations also motivate the choice of transmission conditions for (6.8b) and (6.8c).

In order to adapt the results from Section 2.5, we need to introduce some further notation. For $\ell = 0, \dots, L$, we define the potentials $S_\ell(s)$ and D_ℓ as the single- and double layer potentials corresponding to the subdomain Ω_ℓ in the following modified form:

$$\begin{aligned} (S_\ell(s)\lambda)(\mathbf{x}) &:= \int_{\partial\Omega_\ell} \Phi(\mathbf{x} - \mathbf{y}; s/m_\ell) \lambda(\mathbf{y}) d\sigma(\mathbf{y}), \\ (D_\ell(s)\phi)(\mathbf{x}) &:= \int_{\partial\Omega_\ell} \partial_{\nu(\mathbf{y})} \Phi(\mathbf{x} - \mathbf{y}; s/m_\ell) \phi(\mathbf{y}) d\sigma(\mathbf{y}), \end{aligned}$$

where $m_\ell := c_\ell \sqrt{\kappa_\ell}$ is the effective wave speed on the subdomain. Using the operator calculus, these operators induce time-domain counterparts $S_\ell(\partial_t)$ and $D_\ell(\partial_t)$. We collect the operators on the subdomains into the diagonal operators

$$\begin{aligned} S(\partial_t)\boldsymbol{\lambda} &= S(\partial_t)(\lambda_\ell)_{\ell=0}^L := (S_\ell(\partial_t)\lambda_\ell)_{\ell=0}^L, \\ D(\partial_t)\boldsymbol{\phi} &= D(\partial_t)(\phi_\ell)_{\ell=0}^L := (D_\ell(\partial_t)\phi_\ell)_{\ell=0}^L, \end{aligned}$$

and collect those operators again into the Green representation operator

$$G(\partial_t) := [-S(\partial_t), D(\partial_t)].$$

Just like we did for trace and normal trace, we introduce the vector version of the normal derivative $\boldsymbol{\partial}_\nu$ and the corresponding interior and exterior traces, jumps and means.

We introduce the spaces

$$\mathcal{L}^2 := \prod_{\ell=0}^L [L^2(\mathbb{R}^d)], \quad \mathbf{L}^2 := \prod_{\ell=0}^L [L^2(\mathbb{R}^d)]^d.$$

It is convenient to introduce the component-wise differentiation operators $\nabla : \mathcal{H}^1 \rightarrow \mathcal{L}^2$ and $\text{div} : \mathcal{H}^{\text{div}} \rightarrow \mathbf{L}^2$, which are defined in the natural way. We also introduce the diagonal scaling operators $\mathbf{T}_{c^2} : \mathcal{L}^2 \rightarrow \mathcal{L}^2$ and $\mathbf{T}_\kappa : \mathbf{L}^2 \rightarrow \mathbf{L}^2$ defined by

$$\mathbf{T}_{c^2} (u_\ell)_{\ell=0}^L := (c_\ell^2 u_\ell)_{\ell=0}^L \quad \text{and} \quad \mathbf{T}_\kappa (\mathbf{v}_\ell)_{\ell=0}^L := (\kappa_\ell \mathbf{v}_\ell)_{\ell=0}^L.$$

We will now derive a system of continuous in time boundary integral equations. We proceed in a mostly formal way, as we do not need the equivalence principle on the continuous level. The equivalence can be made rigorous using the theory of Laplace transformable causal distributions, as is presented in [Say16]. We will instead prove the rigorous equivalence in the time-discrete setting later on in Section 6.5.

Solutions to the wave equation, given by

$$\dot{U} = \mathbf{T}_{c^2} \text{div}(\mathbf{V}), \quad \text{and} \quad \dot{\mathbf{V}} = \mathbf{T}_\kappa \nabla U,$$

can be written using Kirchhoffs formula (3.12) as

$$\begin{aligned} U &= -S(\partial_t) \llbracket \boldsymbol{\partial}_\nu U \rrbracket + D(\partial_t) \llbracket \gamma U \rrbracket = -S(\partial_t) \llbracket \gamma_\nu \mathbf{T}_\kappa^{-1} \dot{\mathbf{V}} \rrbracket + D(\partial_t) \llbracket \gamma U \rrbracket \\ &= G(\partial_t) Q_\kappa^{-1} \left(\llbracket \gamma_\nu \dot{\mathbf{V}} \rrbracket, \llbracket \gamma U \rrbracket \right)^T, \end{aligned}$$

with the diagonal scaling operator $Q_\kappa(\boldsymbol{\lambda}, \phi)^T := (\mathbf{T}_\kappa \boldsymbol{\lambda}, \phi)^T$. We define the Calderón operator as

$$C(\partial_t) := \begin{bmatrix} \{\{\boldsymbol{\partial}_\nu \cdot\}\} \\ \{\{\gamma \cdot\}\} \end{bmatrix} G(\partial_t) = \begin{bmatrix} K^T(\partial_t) & W(\partial_t) \\ V(\partial_t) & -K(\partial_t) \end{bmatrix},$$

(the operators $K^T(\partial_t)$ etc. are to be understood as the diagonal operators on each subdomain and using the modified wave number with the additional factor m_ℓ^{-1}). From this definition and the jump conditions of G , we immediately get

$$\begin{bmatrix} \boldsymbol{\partial}_\nu^{\text{ext}} \\ \gamma^{\text{ext}} \end{bmatrix} G(\partial_t) = C(\partial_t) - \frac{1}{2} \text{Id}. \quad (6.9)$$

We collect the right hand sides into the vector $\boldsymbol{\Theta} := (\dot{\boldsymbol{\beta}}^1, \boldsymbol{\beta}^0)^T$, and write for the boundary traces $\boldsymbol{\Lambda}^h := (\boldsymbol{\lambda}^h, \boldsymbol{\psi}^h)^T = \left(\llbracket \gamma_\nu \dot{\mathbf{V}} \rrbracket + \dot{\boldsymbol{\beta}}^1, \llbracket \gamma U \rrbracket + \boldsymbol{\beta}^0 \right)$.

Writing $U^h = G(\partial_t) Q_\kappa^{-1} \left(\llbracket \gamma_\nu \dot{\mathbf{V}} \rrbracket, \llbracket \gamma U \rrbracket \right)^T$, and taking the exterior normal derivative then immediately gives the following boundary integral equations if the jump relations (6.8b) and (6.8c) hold:

$$Q_\kappa \left(C(\partial_t) - \frac{1}{2} \text{Id} \right) Q_\kappa^{-1} \left(\boldsymbol{\Lambda}^h - \boldsymbol{\Theta} \right) \in \mathcal{Y}_h^\circ \times \mathcal{X}_h^\circ.$$

Since $\mathcal{X}_h \times \mathcal{Y}_h \subseteq \mathcal{Y}_h^\circ \times \mathcal{X}_h^\circ$ (see (6.7)), the statement still holds if we drop the factor $\frac{1}{2}\mathbf{\Lambda}^h$. On the other hand, one can prove using a similar approach that the following equivalence holds:

(i) If $\mathbf{\Lambda}^h : \mathbb{R} \rightarrow \mathcal{X}_h \times \mathcal{Y}_h$ solves

$$\left\langle \mathbf{Q}_\kappa \mathbf{C}(\partial_t) \mathbf{Q}_\kappa^{-1} \mathbf{\Lambda}^h, \underline{\boldsymbol{\eta}} \right\rangle_\Gamma = \left\langle \mathbf{Q}_\kappa \left(\mathbf{C}(\partial_t) - \frac{1}{2} \text{Id} \right) \mathbf{Q}_\kappa^{-1} \boldsymbol{\Theta}, \underline{\boldsymbol{\eta}} \right\rangle_\Gamma \quad (6.10)$$

for all $\underline{\boldsymbol{\eta}} \in \mathcal{X}_h \times \mathcal{Y}_h$, and $\mathbf{\Lambda}^h(t) = (0, 0)^T$ for $t \leq 0$, then

$$U^h := \mathbf{G}(\partial_t) \mathbf{Q}_\kappa^{-1} \left(\mathbf{\Lambda}^h - \boldsymbol{\Theta} \right) \quad \text{and} \quad \mathbf{V}^h := \mathbf{T}_\kappa \nabla \partial_t^{-1} U^h$$

solve Problem 6.3.

(ii) If (U^h, \mathbf{V}^h) solves Problem 6.3, then $\mathbf{\Lambda}^h := \left(\llbracket \gamma_\nu \dot{\mathbf{V}} \rrbracket + \dot{\boldsymbol{\beta}}^1, \llbracket \gamma U^h \rrbracket + \boldsymbol{\beta}^0 \right)$ solves (6.10).

This shows that the transmission conditions in Problem 6.3 correspond to a conforming Galerkin discretization using the spaces \mathcal{X}_h and \mathcal{Y}_h of the equivalent boundary integral equations. By using the formalism of convolution quadrature, we can immediately formulate the fully discrete boundary integral problem:

Problem 6.5 (Fully discrete formulation 1). *Given $\boldsymbol{\Theta} := (\dot{\boldsymbol{\beta}}^1, \boldsymbol{\beta}^0)^T$ with $\boldsymbol{\Theta}(t) = (0, 0)^T$ for $t \leq 0$, find $\underline{\boldsymbol{\Lambda}}^{h,k} : \mathbb{R} \rightarrow [\mathcal{X}_h \times \mathcal{Y}_h]^m$, such that $\underline{\boldsymbol{\Lambda}}^{h,k} = (0, 0)^T$ for $t \leq 0$ and*

$$\left\langle \mathbf{Q}_\kappa \mathbf{C}(\partial_t^k) \mathbf{Q}_\kappa^{-1} \underline{\boldsymbol{\Lambda}}^{h,k}, \underline{\boldsymbol{\eta}} \right\rangle_\Gamma = \left\langle \mathbf{Q}_\kappa \left(\mathbf{C}(\partial_t^k) - \frac{1}{2} \text{Id} \right) \mathbf{Q}_\kappa^{-1} \boldsymbol{\Theta}, \underline{\boldsymbol{\eta}} \right\rangle_\Gamma \quad \forall \underline{\boldsymbol{\eta}} \in [\mathcal{Y}_h \times \mathcal{X}_h]^m, \quad (6.11)$$

where equality is understood as a function in t . The approximation of the traces at time t is again given by $\boldsymbol{\lambda}^{h,k}(t) := \mathbf{G}[\underline{\boldsymbol{\Lambda}}^{h,k}]$ and $\boldsymbol{\psi}^{h,k}(t) := \mathbf{G}[\underline{\boldsymbol{\psi}}^{h,k}]$.

The analysis of an equivalent problem will reveal that it is advantageous to consider the following fully discrete problem:

Problem 6.6 (Fully discrete formulation 2). *Given $\boldsymbol{\Sigma} := (\dot{\boldsymbol{\beta}}^1, \dot{\boldsymbol{\beta}}^0)^T$ with $\boldsymbol{\Sigma}(t) = (0, 0)^T$ for $t \leq 0$, find $\underline{\boldsymbol{\Lambda}}^{h,k} : \mathbb{R} \rightarrow [\mathcal{X}_h \times \mathcal{Y}_h]^m$, such that $\underline{\boldsymbol{\Lambda}}^{h,k}(t) = (0, 0)^T$ for $t \leq 0$ and*

$$\left\langle \mathbf{Q}_\kappa \mathbf{C}(\partial_t^k) \mathbf{Q}_\kappa^{-1} \underline{\boldsymbol{\Lambda}}^{h,k}, \underline{\boldsymbol{\eta}} \right\rangle_\Gamma = \left\langle \mathbf{Q}_\kappa \left(\mathbf{C}(\partial_t^k) - \frac{1}{2} \text{Id} \right) \mathbf{Q}_\kappa^{-1} \mathfrak{J}(\partial_t^k) \boldsymbol{\Sigma}, \underline{\boldsymbol{\eta}} \right\rangle_\Gamma \quad \forall \underline{\boldsymbol{\eta}} \in [\mathcal{Y}_h \times \mathcal{X}_h]^m, \quad (6.12)$$

where equality is understood as a function in t and $\mathfrak{J}(s) := \text{diag}(1, s^{-1})$ denotes the discrete integral in the second argument. The approximation of the traces at time t is then given by $\boldsymbol{\lambda}^{h,k}(t) := \mathbf{G}[\underline{\boldsymbol{\Lambda}}^{h,k}]$ and $\boldsymbol{\psi}^{h,k}(t) := \mathbf{G}[\underline{\boldsymbol{\psi}}^{h,k}]$.

Theorem 6.7. *Problems 6.5 and 6.6 are well posed, i.e., they have a unique solution.*

Proof. We show the theorem for Problem 6.5. Since the left-hand sides are the same the statement for Problem 6.6 follows. By induction on n , the existence and uniqueness of solutions can be reduced to the question if

$$\left\langle \mathbb{Q}_\kappa \mathbb{C} \left(\frac{\delta(0)}{k} \right) \mathbb{Q}_\kappa^{-1} x, \boldsymbol{\eta} \right\rangle_\Gamma = \langle F, \boldsymbol{\eta} \rangle_\Gamma, \quad \forall \boldsymbol{\eta} \in [\mathcal{Y}_h \times \mathcal{X}_h]^m$$

can be solved for all right-hand sides $F \in [\mathcal{X}'_h \times \mathcal{Y}'_h]^m$ (see the proof of Lemma 5.10). We note that the bilinear form induced by the operator $\mathbb{Q}_\kappa \mathbb{C}(s) \mathbb{Q}_\kappa^{-1}$ is coercive on the product space $\mathcal{H}^{-1/2} \times \mathcal{H}^{1/2}$ for $\operatorname{Re}(s) > 0$. This can be seen since on each subdomain the operators are coercive by [BLS15b, Lemma 3.1] (cf. Lemma 5.8, the diagonal scaling by s and $s^{-1/2}$ does not impact the coercivity). Since (6.10) is just a restriction of this bilinear form, it is also coercive on $\mathcal{X}_h \times \mathcal{Y}_h$. This means the scalar version of (6.11) has a unique solution for $\operatorname{Re}(s) > 0$. Let $B(s) : \mathcal{X}_h \times \mathcal{Y}_h \rightarrow \mathcal{X}_h \times \mathcal{Y}_h$ denote the operator induced by this bilinear form. Then $B^{-1}(s)$ exists and is bounded for $\operatorname{Re}(s) > 0$ by the previous considerations. Since $\operatorname{Re}(\delta(0)) > 0$, we therefore get that $B^{-1} \left(\frac{\delta(0)}{k} \right)$ is well defined using the Riesz-Dunford calculus and is the inverse of $B \left(\frac{\delta(0)}{k} \right)$ due to the homomorphism property. \square

6.3.1 One possible construction for \mathcal{X}_h and \mathcal{Y}_h

In this section, we present one possible way to construct spaces \mathcal{X}_h and \mathcal{Y}_h , such that the resulting spaces have good approximation properties, and they can be easily implemented using existing boundary element technology. We assume that all domains Ω_ℓ are Lipschitz polyhedra. We separate Γ into a finite collection of (relatively) open, flat surfaces $\Gamma_1, \dots, \Gamma_M$ such that for all $\ell \in \{0, \dots, L\}$ there exists an index set $\mathcal{I}(\ell) \subseteq \{1, \dots, M\}$, such that

$$\partial\Omega_\ell = \bigcup_{i \in \mathcal{I}(\ell)} \overline{\Gamma}_i.$$

Let \mathcal{T}_h^Γ denote a regular and shape-regular triangulation of Γ (see Section 2.4) which respects this subdivision, i.e., for all $K \in \mathcal{T}_h^\Gamma$, either $K \subseteq \Gamma_i$ or $K \cap \Gamma_i = \emptyset$. This can be constructed by generating a volume mesh \mathcal{T}_h of $\overline{\Omega}_1 \cup \dots \cup \overline{\Omega}_K$, such that no tetrahedral element intersects Γ , and then setting $\mathcal{T}_h^\Gamma := \mathcal{T}_h|_\Gamma$ for the restriction.

On \mathcal{T}_h^Γ we use the usual spaces of piecewise polynomials $\mathcal{S}^{r,0}(\mathcal{T}_h^\Gamma)$ and $\mathcal{S}^{r+1,1}(\mathcal{T}_h^\Gamma)$ as defined in Definition 2.33. In order to define a conforming subspace of \mathcal{Y} , we define:

$$\mathcal{Y}_h := \left\{ (\psi_h|_{\partial\Omega_\ell})_{\ell=0}^L : \psi_h \in \mathcal{S}^{r+1,1}(\mathcal{T}_h^\Gamma) \right\}. \quad (6.13)$$

It is easy to prove that \mathcal{Y}_h and $\mathcal{S}^{r+1,1}(\mathcal{T}_h^\Gamma)$ are isomorphic and the corresponding map can easily be implemented in practice by exploiting the association of basis functions to the geometric quantities of the mesh.

In order to define \mathcal{X}_h , we need to deal with the orientation of the boundaries. For that purpose we introduce a sign function. For $\ell \in \{0, \dots, L\}$, we define a function $s_\ell : \Gamma \rightarrow$

$\{-1, 0, 1\}$, which is constant on each face Γ_i , such that $s_\ell := 0$ outside of $\partial\Omega_\ell$ and $|s_\ell| \equiv 1$ on $\partial\Omega_\ell$. We assume that common faces have opposite signs, i.e.,

$$s_\ell + s_k \equiv 0 \quad \text{on } \Gamma_i \text{ for all } i \in \mathcal{I}(\ell) \cap \mathcal{I}(k).$$

Such a function can be constructed by assigning a normal vector to each face Γ_i , and then write $s_\ell|_{\Gamma_i} := 1$ if the normal vector on $\partial\Omega_\ell$ matches the chosen one on Γ_i and $s_\ell|_{\Gamma_i} := -1$ otherwise. This allows us to define the conforming space $\mathcal{X}_h \subseteq \mathcal{X}$ by

$$\mathcal{X}_h := \left\{ (s_\ell \lambda_h|_{\partial\Omega_\ell})_{\ell=0}^L : \lambda_h \in \mathcal{S}^{r,0}(\mathcal{T}_h^\Gamma) \right\}. \quad (6.14)$$

Proposition 6.8. *For $r \in \mathbb{N}_0$, the spaces \mathcal{X}_h and \mathcal{Y}_h have the following approximation property:*

$$\inf_{\lambda^h \in \mathcal{X}_h} \left\| \boldsymbol{\lambda} - \boldsymbol{\lambda}^h \right\|_{\mathcal{H}^{-1/2}} \leq Ch^{r+3/2} \sum_{\ell=0}^L \|\lambda_\ell\|_{H_{\text{pw}}^{r+1}(\partial\Omega_\ell)}, \quad (6.15a)$$

$$\inf_{\boldsymbol{\psi}^h \in \mathcal{Y}_h} \left\| \boldsymbol{\psi} - \boldsymbol{\psi}^h \right\|_{\mathcal{H}^{1/2}} \leq Ch^{r+3/2} \sum_{\ell=0}^L \|\psi_\ell\|_{H_{\text{pw}}^{r+2}(\partial\Omega_\ell)} \quad (6.15b)$$

for all $\boldsymbol{\lambda} := (\lambda_\ell)_{\ell=0}^L \in \mathcal{X}$ with $\lambda_\ell \in H_{\text{pw}}^{r+1}(\partial\Omega_\ell)$ and all $\boldsymbol{\psi} := (\psi_\ell)_{\ell=0}^L \in \mathcal{Y}$ with $\psi_\ell \in H_{\text{pw}}^{r+2}(\partial\Omega_\ell)$ and with the additional restriction that the lifting $u \in H^1(\mathbb{R}^d)$ from (6.6a) is continuous on Γ .

Proof. We start with the estimate for $\boldsymbol{\lambda}$. For each $i \in \{0, \dots, M\}$, we pick a subdomain Ω_{ℓ_i} , such that $i \in \mathcal{I}(\ell_i)$, and set $\lambda^h|_{\Gamma_i} := s_{\ell_i} \Pi_i \lambda_\ell|_{\Gamma_i}$, where $\Pi_i \lambda_\ell \in \mathcal{S}^{m,0}(\mathcal{T}_h^\Gamma)$ is the orthogonal projection with respect to the L^2 -product on Γ_i . Since $\mathcal{S}^{m,0}(\mathcal{T}_h^\Gamma)$ is only required to be L^2 -conforming, this defines a function in \mathcal{X}_h via $\boldsymbol{\lambda}^h := (\lambda_\ell^h) := (s_\ell \lambda^h|_{\partial\Omega_\ell})_{\ell=0}^L$. We calculate for $\ell \in 0, \dots, L$:

$$\left\| \lambda_\ell - \lambda_\ell^h \right\|_{L^2(\partial\Omega_\ell)}^2 = \sum_{i \in \mathcal{I}(\ell)} \left\| \lambda_\ell - s_\ell \lambda^h \right\|_{L^2(\Gamma_i)}^2 = \sum_{i=1}^M \|s_{\ell_i} s_\ell \lambda_\ell - s_\ell \lambda_h\|_{L^2(\Gamma_i)}^2,$$

where in the last step we used that for $\boldsymbol{\lambda} \in \mathcal{H}^{-1/2}$, the components of two subdomains sharing a face Γ_i differ only by a sign, and $s_\ell = 0$ if the components do not share a face. By definition of λ_h we can therefore estimate

$$\begin{aligned} \left\| \lambda_\ell - \lambda_\ell^h \right\|_{L^2(\partial\Omega_\ell)} &\lesssim \sum_{i=1}^M \|\lambda_{\ell_i} - \Pi_i \lambda_{\ell_i}\|_{L^2(\Gamma_i)} \lesssim h^{r+1} \sum_{i=1}^M \|\lambda_{\ell_i}\|_{H^{r+1}(\Gamma_i)} \\ &\lesssim h^{r+1} \sum_{\ell=0}^L \|\lambda_\ell\|_{H_{\text{pw}}^{r+1}(\partial\Omega_\ell)}, \end{aligned}$$

where we used the approximation property of Proposition 2.35.

To get an estimate in the $\mathcal{H}^{-1/2}$ -norm, we calculate for $\ell \in \{0, \dots, L\}$, using the fact that $\lambda_\ell - \lambda_\ell^h$ is orthogonal to the piecewise polynomials on each face:

$$\begin{aligned} \left\| \lambda_\ell - \lambda_\ell^h \right\|_{H^{-1/2}(\partial\Omega_\ell)} &= \sup_{0 \neq \chi \in H^{1/2}(\partial\Omega_\ell)} \frac{\langle \lambda_\ell - \lambda_\ell^h, \chi \rangle_{\partial\Omega_\ell}}{\|\chi\|_{H^{1/2}(\partial\Omega_\ell)}} = \sup_{0 \neq \chi \in H^{1/2}(\partial\Omega_\ell)} \sum_{i \in \mathcal{I}(\ell)} \frac{\langle \lambda_\ell - \lambda_\ell^h, \chi \rangle_{\Gamma_i}}{\|\chi\|_{H^{1/2}(\partial\Omega_\ell)}} \\ &= \sup_{0 \neq \chi \in H^{1/2}(\partial\Omega_\ell)} \sum_{i \in \mathcal{I}(\ell)} \frac{\langle \lambda_\ell - \lambda_\ell^h, \chi - \Pi_i \chi \rangle_{\Gamma_i}}{\|\chi\|_{H^{1/2}(\partial\Omega_\ell)}} \\ &\lesssim \sqrt{h} \sup_{0 \neq \chi \in H^{1/2}(\partial\Omega_\ell)} \sum_{i \in \mathcal{I}(\ell)} \frac{\|\lambda_\ell - \lambda_\ell^h\|_{L^2(\Gamma_i)} \|\chi\|_{H^{1/2}(\Gamma_i)}}{\|\chi\|_{H^{1/2}(\partial\Omega_\ell)}} \\ &\lesssim \sqrt{h} \left\| \lambda_\ell - \lambda_\ell^h \right\|_{L^2(\partial\Omega_\ell)}. \end{aligned}$$

For estimating ψ , we note that our assumptions on the lifting u implies that the functions ψ_ℓ are continuous on $\partial\Omega_\ell$, most notably at the boundary of the facets. Therefore, we may employ a nodal interpolation operator $I_\ell : C(\partial\Omega_\ell) \rightarrow \mathcal{S}^{p+1,1}(\mathcal{T}_h^\Gamma|_{\partial\Omega_\ell})$ (where $\mathcal{T}_h^\Gamma|_{\partial\Omega_\ell}$ denotes the restriction of the triangulation to the subdomain $\partial\Omega_\ell$). It is well known that

$$\|\psi_\ell - I_\ell \psi_\ell\|_{H^{1/2}(\partial\Omega_\ell)} \lesssim h^{r+3/2} \|\psi_\ell\|_{H_{\text{pw}}^{r+2}(\partial\Omega_\ell)},$$

see [SS11, Theorem 4.3.22].

Since the functions ψ_ℓ, ψ_k are assumed to be traces of a continuous function u , they must coincide on $\partial\Omega_\ell \cap \partial\Omega_k$. This means that the interpolated functions $I_\ell \psi_\ell$ also coincide on $\partial\Omega_\ell \cap \partial\Omega_k$ or $(I_\ell \psi_\ell)_{\ell=0}^L \in \mathcal{X}_h$. \square

6.4 Analysis of the transmission problem - the semigroup setting

Now that we have derived the model problem and a suitable discretization scheme, we would like to analyze its properties. In order to do so, we fit the problem into the abstract semigroup framework from Section 2.2. We start with the case of continuous time.

While the spaces $\mathcal{H}^1, \mathcal{H}^{\text{div}}$ are a natural setting to formulate problem 6.3, they are not the right spaces if one wants to use semigroup theory. Just like in the previous chapters, it is advantageous to adopt an L^2 -based view. The spaces most convenient to recast Problem 6.3 into the semigroup framework, as introduced in Proposition 2.11, are defined as follows:

$$\mathbb{H} := \mathcal{L}^2 \times \mathbf{L}^2 \quad \text{and} \quad \mathbb{V} := \mathcal{H}^1 \times \mathcal{H}^{\text{div}},$$

where \mathbb{V} is equipped with the usual norm, and for \mathbb{H} we use the norm:

$$(U, \mathbf{V}) = ((u_\ell)_{\ell=0}^L, (\mathbf{v}_\ell)_{\ell=0}^L) \mapsto \|(U, \mathbf{V})\|_{\mathbb{H}}^2 := \sum_{\ell=0}^L c_\ell^{-2} \|u_\ell\|_{L^2(\mathbb{R}^d)}^2 + \sum_{\ell=0}^L \kappa_\ell^{-1} \|\mathbf{v}_\ell\|_{L^2(\mathbb{R}^d)}^2,$$

and the corresponding inner product. In order to enforce the boundary conditions of (6.8b) and (6.8c), we introduce the space $\mathbb{M} := (\mathcal{Y}_h^\circ)' \times (\mathcal{X}_h^\circ)' \times \mathcal{X}_h' \times \mathcal{Y}_h'$, where \mathcal{Y}_h° and \mathcal{X}_h are

equipped with the $\mathcal{H}^{-1/2}$ -norm, \mathcal{X}_h° and \mathcal{Y}_h with the $\mathcal{H}^{1/2}$ -norm and \mathbb{M} carries the product-dual norm of these spaces.

We define the operators $A_\star : \mathbb{V} \rightarrow \mathbb{H}$ and $B : \mathbb{V} \rightarrow \mathbb{M}$ by

$$A_\star \begin{pmatrix} U \\ \mathbf{V} \end{pmatrix} := \begin{pmatrix} T_{c^2} \operatorname{div}(\mathbf{V}) \\ T_\kappa \nabla U \end{pmatrix} \quad \text{and} \quad B \begin{pmatrix} U \\ \mathbf{V} \end{pmatrix} := \begin{pmatrix} \llbracket \gamma U \rrbracket|_{\mathcal{Y}_h^\circ} \\ \llbracket \gamma_\nu \mathbf{V} \rrbracket|_{\mathcal{X}_h^\circ} \\ \gamma^{\text{ext}} U|_{\mathcal{X}_h} \\ \gamma_\nu^{\text{ext}} \mathbf{V}|_{\mathcal{Y}_h} \end{pmatrix}.$$

Here, the restrictions are using the same notational convention as [Has+15]. Namely we regard the traces as functionals via the Riesz-map and restrict the functionals. This means we define:

$$\begin{aligned} \left\langle \llbracket \gamma U \rrbracket|_{\mathcal{Y}_h^\circ}, \boldsymbol{\mu} \right\rangle_{(\mathcal{Y}_h^\circ)' \times (\mathcal{Y}_h^\circ)} &:= \langle \boldsymbol{\mu}, \llbracket \gamma U \rrbracket \rangle_\Gamma & \forall \boldsymbol{\mu} \in \mathcal{Y}_h^\circ, \\ \left\langle \llbracket \gamma_\nu \mathbf{V} \rrbracket|_{\mathcal{X}_h^\circ}, \boldsymbol{\chi} \right\rangle_{(\mathcal{X}_h^\circ)' \times (\mathcal{X}_h^\circ)} &:= \langle \llbracket \gamma_\nu \mathbf{V} \rrbracket, \boldsymbol{\chi} \rangle_\Gamma & \forall \boldsymbol{\chi} \in \mathcal{X}_h^\circ, \\ \langle \gamma^{\text{ext}} U|_{\mathcal{X}_h}, \boldsymbol{\mu} \rangle_{\mathcal{X}_h' \times \mathcal{X}_h} &:= \langle \boldsymbol{\mu}, \gamma^{\text{ext}} U \rangle_\Gamma & \forall \boldsymbol{\mu} \in \mathcal{X}_h, \\ \langle \gamma_\nu^{\text{ext}} \mathbf{V}|_{\mathcal{Y}_h}, \boldsymbol{\chi} \rangle_{\mathcal{Y}_h' \times \mathcal{Y}_h} &:= \langle \gamma_\nu^{\text{ext}} \mathbf{V}, \boldsymbol{\chi} \rangle_\Gamma & \forall \boldsymbol{\chi} \in \mathcal{Y}_h. \end{aligned}$$

Using this notation, we can recast Problem 6.3 in the following form.

Problem 6.9. Given $\Xi(t) : [0, \infty) \rightarrow \mathbb{M}$, defined by

$$\Xi(t) := - \left(\boldsymbol{\beta}^0(t)|_{\mathcal{Y}_h^\circ}, \boldsymbol{\beta}^1(t)|_{\mathcal{X}_h^\circ}, 0, 0 \right),$$

find $\mathbf{X}^h := (U^h, \mathbf{V}^h) : [0, \infty) \rightarrow \mathbb{V}$ satisfying

$$\dot{\mathbf{X}}^h(t) = A_\star \mathbf{X}^h(t) \quad \text{and} \quad B \mathbf{X}^h(t) = \Xi(t) \quad \forall t \geq 0 \quad \text{with} \quad \mathbf{X}^h(0) = 0.$$

We remark that $\|\Xi(t)\|_{\mathbb{M}} \lesssim \|\boldsymbol{\beta}^0(t)\|_{\mathcal{H}^{1/2}} + \|\boldsymbol{\beta}^1(t)\|_{\mathcal{H}^{-1/2}}$. This problem fits the abstract framework of Proposition 2.11, we just have to check some conditions. This will be content of the next section.

6.4.1 Checking the semigroup requirements

In order to apply the abstract semigroup theory, we need to check some conditions. We start by characterizing the kernel of B .

Lemma 6.10. $(U, \mathbf{V}) \in \ker(B)$ is equivalent to the four conditions

$$\llbracket \gamma U \rrbracket \in \mathcal{Y}_h, \quad \llbracket \gamma_\nu \mathbf{V} \rrbracket \in \mathcal{X}_h, \quad (6.16a)$$

$$\gamma^{\text{ext}} U \in \mathcal{X}_h^\circ, \quad \gamma_\nu^{\text{ext}} \mathbf{V} \in \mathcal{Y}_h^\circ. \quad (6.16b)$$

Proof. If $\llbracket \gamma U \rrbracket \in \mathcal{Y}_h$, then the functional $\boldsymbol{\xi} \mapsto \langle \boldsymbol{\xi}, \llbracket \gamma U \rrbracket \rangle_\Gamma$ vanishes for $\boldsymbol{\xi} \in \mathcal{Y}_h^\circ$ by definition of the annihilator, thus $\llbracket \gamma U \rrbracket|_{\mathcal{Y}_h^\circ} = 0$. An analogous consideration shows that the conditions in (6.16) imply $(U, \mathbf{V}) \in \ker(B)$. On the other hand, if $(U, \mathbf{V}) \in \ker(B)$, we get for the first component $0 = \langle \boldsymbol{\xi}, \llbracket \gamma U \rrbracket \rangle_\Gamma$ for all $\boldsymbol{\xi} \in \mathcal{Y}_h^\circ$ or $\llbracket \gamma U \rrbracket \in (\mathcal{Y}_h^\circ)^\circ = \mathcal{Y}_h$, since we assumed \mathcal{Y}_h to be closed. \square

Lemma 6.10 also establishes that Problem 6.9 is indeed equivalent to Problem 6.3.

Lemma 6.11. A_\star is dissipative on the kernel of B , i.e.,

$$\langle A_\star \mathbf{X}, \mathbf{X} \rangle_{\mathcal{H}} = 0 \quad \forall \mathbf{X} \in \ker(B).$$

Proof. We write $\mathbf{X} =: (U, \mathbf{V})$ with $U =: (u_\ell)_{\ell=0}^L$ and $\mathbf{V} =: (\mathbf{v}_\ell)_{\ell=0}^L$ for the different fields. Integration by parts then gives:

$$\begin{aligned} \langle A_\star \mathbf{X}, \mathbf{X} \rangle_{\mathcal{H}} &= \sum_{\ell=0}^L c_\ell^{-2} \langle c_\ell^2 \operatorname{div} \mathbf{v}_\ell, u_\ell \rangle_{L^2(\mathbb{R}^d)} + \sum_{\ell=0}^L \kappa_\ell^{-1} \langle \kappa_\ell \nabla u_\ell, \mathbf{v}_\ell \rangle_{[L^2(\mathbb{R}^d)]^d} \\ &= \sum_{\ell=0}^L \langle \gamma_{\nu, \ell}^{\text{int}} \mathbf{v}_\ell, \gamma_\ell^{\text{int}} u_\ell \rangle_\ell - \langle \gamma_{\nu, \ell}^{\text{ext}} \mathbf{v}_\ell, \gamma_\ell^{\text{ext}} u_\ell \rangle_\ell \\ &= \langle \llbracket \gamma_\nu \mathbf{V} \rrbracket, \gamma^{\text{int}} U \rangle_\Gamma + \langle \gamma_\nu^{\text{ext}} \mathbf{V}, \llbracket \gamma U \rrbracket \rangle_\Gamma. \end{aligned} \quad (6.17)$$

By Lemma 6.10, the second term vanishes by definition of the polar sets.

To see that the first term also vanishes, we write $\gamma^{\text{int}} U = \gamma^{\text{ext}} U - \llbracket \gamma U \rrbracket$. Since $\mathcal{Y}_h \subseteq \mathcal{Y} \subseteq \mathcal{X}^\circ \subseteq \mathcal{X}_h^\circ$ by (6.7), we get $\llbracket \gamma U \rrbracket \in \mathcal{X}_h^\circ$. $\gamma^{\text{ext}} U \in \mathcal{X}_h^\circ$ was already established in (6.16). Since $\llbracket \gamma_\nu \mathbf{V} \rrbracket \in \mathcal{X}_h$, the corresponding term in (6.17) also vanishes. \square

Lemma 6.12. For all $(F, \mathbf{G}) \in \mathbb{H}$ and $\Xi \in \mathbb{M}$, there exists a unique $(U, \mathbf{V}) \in \mathbb{V}$ solving

$$(U, \mathbf{V}) = A_\star(U, \mathbf{V}) + (F, \mathbf{G}) \quad \text{and} \quad B(U, \mathbf{V}) = \Xi.$$

There exists a constant $C > 0$ depending only on the geometry and the physical parameters κ, c , such that

$$\|(U, \mathbf{V})\|_{\mathbb{V}} \leq C(\|(F, \mathbf{G})\|_{\mathbb{H}} + \|\Xi\|_{\mathbb{M}}).$$

In the special case $(F, \mathbf{G}) = (0, 0)$, we write $\mathcal{E}_B[\Xi]$ for the solution to $U = A_\star U$ and $B U = \Xi$.

Proof. We write $F = (f_\ell)_{\ell=0}^L$, $\mathbf{G} = (\mathbf{g}_\ell)_{\ell=0}^L$ and

$$\Xi = (\xi_1, \xi_2, \xi_3, \xi_4) \in (\mathcal{Y}_h^\circ)' \times (\mathcal{X}_h^\circ)' \times \mathcal{X}_h' \times \mathcal{Y}_h'.$$

We rewrite the equation as a second order elliptic system. To that end, we define

$$\begin{aligned} a(U, W) &:= \sum_{\ell=0}^L c_\ell^{-2} (u_\ell, w_\ell)_{L^2(\mathbb{R}^d)} + \sum_{\ell=0}^L \kappa_\ell (\nabla u_\ell, \nabla w_\ell)_{[L^2(\mathbb{R}^d \setminus \partial\Omega_\ell)]^d}, \\ b(W) &:= \sum_{\ell=0}^L c_\ell^{-2} (f_\ell, w_\ell)_{L^2(\mathbb{R}^d)} - \sum_{\ell=0}^L (\mathbf{g}_\ell, \nabla w_\ell)_{L^2(\mathbb{R}^d \setminus \partial\Omega_\ell)} \\ &\quad + \langle \xi_2, \gamma^{\text{int}} W \rangle_{(\mathcal{X}_h^\circ)' \times \mathcal{X}_h^\circ} + \langle \xi_4, \llbracket \gamma W \rrbracket \rangle_{\mathcal{Y}_h' \times \mathcal{Y}_h}, \end{aligned}$$

using the space

$$\mathcal{W} := \{W \in \mathcal{H}^1 : \llbracket \gamma W \rrbracket \in \mathcal{Y}_h, \gamma^{\text{ext}} W \in \mathcal{X}_h^\circ\} = \{W \in \mathcal{H}^1 : (W, 0) \in \ker B\}.$$

We now look for $U \in \mathcal{H}^1$ satisfying

$$a(U, W) = b(W) \quad \forall W \in \mathcal{W}, \quad (6.18a)$$

$$\llbracket \gamma U \rrbracket|_{\mathcal{Y}_h^\circ} = \xi_1, \quad \gamma^{\text{ext}} U|_{\mathcal{X}_h} = \xi_3. \quad (6.18b)$$

The bilinear form a can be checked easily to be bounded and coercive with respect to the \mathcal{H}^1 norm, with constants depending only on the physical coefficients. The linear form b is bounded by

$$\|b\|_{\mathcal{W}'} \leq C (\|(F, \mathbf{G})\|_{\mathbb{H}} + \|\Xi\|_{\mathbb{M}}).$$

In order to show that the inhomogeneous problem (6.18) has a solution, we need to show that the trace conditions have a bounded right-inverse. The general solution is then constructed by lifting the boundary data and solving a modified homogeneous problem via Lax-Milgram.

The map

$$\mathcal{H}^1 \ni U \longmapsto (\llbracket \gamma U \rrbracket, \gamma^{\text{ext}} U) \in \mathcal{H}^{1/2} \times \mathcal{H}^{1/2}$$

admits a bounded right-inverse by using the lifting operators from Proposition 2.26. The restriction map

$$\mathcal{H}^{1/2} \times \mathcal{H}^{1/2} = (\mathcal{H}^{-1/2})' \times (\mathcal{H}^{-1/2})' \rightarrow (\mathcal{Y}_h^\circ)' \times \mathcal{X}_h'$$

admits a norm-preserving right-inverse via Hahn-Banach's theorem [Rud91, Theorem 3.3]. This shows that the linear map that imposes the essential transmission conditions in (6.18b) admits a bounded right-inverse with bound independent of the choice of \mathcal{X}_h and \mathcal{Y}_h . Overall, this shows that (6.18) has a unique solution. Define $\mathbf{V} := \mathbf{T}_\kappa \nabla U + \mathbf{G}$. It is then straightforward to prove by using smooth test functions which on the skeleton, that $\mathbf{T}_{c^2} \operatorname{div} \mathbf{V} + F = U$ (in particular $\mathbf{V} \in \mathcal{H}^{\operatorname{div}}$). Integration by parts then gives for arbitrary $W \in \mathcal{W}$:

$$\langle \llbracket \gamma_\nu \mathbf{V} \rrbracket, \gamma^{\text{int}} W \rangle_\Gamma + \langle \gamma_\nu^{\text{ext}} \mathbf{V}, \llbracket \gamma W \rrbracket \rangle_\Gamma = \langle \xi_2, \gamma^{\text{int}} W \rangle_{(\mathcal{X}_h^\circ)' \times \mathcal{X}_h^\circ} + \langle \xi_4, \llbracket \gamma W \rrbracket \rangle_{\mathcal{Y}_h' \times \mathcal{Y}_h} \quad \forall W \in \mathcal{W}.$$

Using the fact that the map $\mathcal{W} \ni W \longmapsto (\gamma^{\text{int}} W, \llbracket \gamma W \rrbracket) \in \mathcal{X}_h^\circ \times \mathcal{Y}_h$ is surjective, the missing two conditions to obtain $B(U, \mathbf{V}) = \Xi$ are proved. The rest of the proof is straightforward. \square

The following lemma allows us to conveniently infer results about $-A_\star$ from their counterpart about A_\star :

Lemma 6.13. *The sign flipping operator $\Phi : \mathbb{H} \rightarrow \mathbb{H}$, $(U, \mathbf{V}) \mapsto (U, -\mathbf{V})$ is an isometric involution that leaves $\ker(B)$ invariant and satisfies $\Phi A_\star = -A_\star \Phi$.*

Proof. Follows directly from the definitions. \square

Theorem 6.14. *Problem 6.9 is well-posed, the operator $A := A_\star|_{\ker B}$ generates a unitary C_0 -group. For $(\beta^0, \beta^1) \in C^2([0, \infty), \mathcal{H}^{1/2} \times \mathcal{H}^{-1/2})$ with $\beta^0(0) = \dot{\beta}^0(0) = 0$ as well as $\beta^1(0) = \dot{\beta}^1(0) = 0$, the solution satisfies:*

$$\begin{aligned} \left\| (U^h(t), \mathbf{V}^h(t)) \right\|_{\mathbb{H}} &\leq C \int_0^t \left\| (\beta^1(\tau), \beta^0(\tau)) \right\|_{\Gamma} + \left\| (\dot{\beta}^1(\tau), \dot{\beta}^0(\tau)) \right\|_{\Gamma} d\tau, \\ \left\| (\dot{U}^h(t), \dot{\mathbf{V}}^h(t)) \right\|_{\mathbb{H}} &\leq C \int_0^t \left\| (\dot{\beta}^1(\tau), \dot{\beta}^0(\tau)) \right\|_{\Gamma} + \left\| (\ddot{\beta}^1(\tau), \ddot{\beta}^0(\tau)) \right\|_{\Gamma} d\tau. \end{aligned}$$

If in addition $(\beta^0, \beta^1) \in C^3([0, \infty), \mathcal{H}^{1/2} \times \mathcal{H}^{-1/2})$ and $(\ddot{\beta}^0(0), \ddot{\beta}^1(0)) = (0, 0)$, we can estimate:

$$\begin{aligned} \left\| (\dot{U}^h(t), \dot{\mathbf{V}}^h(t)) \right\|_{\mathbb{V}} &\leq C \int_0^t \left\| (\dot{\beta}^1(\tau), \dot{\beta}^0(\tau)) \right\|_{\Gamma} + \left\| (\ddot{\beta}^1(\tau), \ddot{\beta}^0(\tau)) \right\|_{\Gamma} d\tau \\ &\quad + C \int_0^t \left\| (\ddot{\beta}^1(\tau), \ddot{\beta}^0(\tau)) \right\|_{\Gamma} d\tau. \end{aligned} \quad (6.19)$$

Proof. By Lemma 6.11 and 6.12, the operator A is maximally dissipative. By using the sign flip, this implies that $-A$ is also maximally dissipative. To see that $\text{dom}(A)$ is dense in \mathbb{H} , we consider the set

$$\left[C_0^\infty(\mathbb{R}^d \setminus \Gamma) \right]^{L+1} \times \left(\left[C_0^\infty(\mathbb{R}^d \setminus \Gamma) \right]^d \right)^{L+1} \subseteq \text{dom}(A),$$

(the boundary conditions are trivially fulfilled). Since this set is dense in \mathbb{H} , we can apply Stone's theorem (Proposition 2.9), which then gives that A generates a unitary C_0 -group. The lifting operator $\mathcal{E}_B[\Xi]$ was already defined in Lemma 6.12 and is bounded with constants depending only on the geometry, κ , and c . Therefore, we are in the setting of Proposition 2.11, which gives existence, uniqueness, and the *a priori* bounds.

The bound of (6.19) follows by using Lemma 2.13, and using the *a priori* bounds for this initial value problem. \square

Remark 6.15. *Higher derivatives of (U^h, \mathbf{V}^h) can also be estimated by inductively using Lemma 2.13 and Theorem 6.14.*

6.4.2 Convergence of the space discretization

Now that we have done the preparatory work, we can analyze how the choice of \mathcal{X}_h and \mathcal{Y}_h impacts the convergence rates. We consider solutions $\mathbf{X} := (U, \mathbf{V})$ of Problem 6.3 in the case $\mathcal{X}_h = \mathcal{X}$ and $\mathcal{Y}_h = \mathcal{Y}$ (or equivalently the solution of (6.3)) and the solution $\mathbf{X}^h := (U^h, \mathbf{V}^h)$ for general conforming subspaces $\mathcal{X}_h \subseteq \mathcal{X}$ and $\mathcal{Y}_h \subseteq \mathcal{Y}$. The convergence rate is characterized in the following theorem:

Theorem 6.16. *Let $\Xi := -(\beta^0|_{\mathcal{Y}_h^\circ}, \beta^1|_{\mathcal{X}_h^\circ}, 0, 0) \in C^3(\mathbb{R}_+, \mathbb{M})$ be given, with $\Xi^{(j)}(0) = 0$ for $j \in \{0, 1, 2\}$. Let $\mathbf{X} := (U, \mathbf{V})$ denote the solution to Problem 6.3 for $\mathcal{X}_h = \mathcal{X}$ and $\mathcal{Y}_h = \mathcal{Y}$, and let $\mathbf{X}^h := (U^h, \mathbf{V}^h)$ be the solution for general conforming subspaces $\mathcal{X}_h \subseteq \mathcal{X}$ and $\mathcal{Y}_h \subseteq \mathcal{Y}$.*

Define the traces $\boldsymbol{\lambda} := \gamma_\nu^{\text{int}} \dot{\mathbf{V}} + \dot{\boldsymbol{\beta}}^1$, $\boldsymbol{\psi} := \gamma^{\text{int}} U + \boldsymbol{\beta}^0$, and set $\boldsymbol{\Lambda} := (\boldsymbol{\lambda}, \boldsymbol{\psi})$. Then the following estimates hold, using the best approximation operator $\Pi : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{X}_h \times \mathcal{Y}_h$:

$$\left\| \mathbf{X}(t) - \mathbf{X}^h(t) \right\|_{\mathbb{H}} \leq C \sum_{j=-1}^1 \int_0^t \left\| \boldsymbol{\Lambda}^{(j)}(\tau) - \Pi \boldsymbol{\Lambda}^{(j)}(\tau) \right\|_{\Gamma} d\tau, \quad (6.20a)$$

$$\left\| \mathbf{X}(t) - \mathbf{X}^h(t) \right\|_{\mathbb{V}} \leq C \sum_{j=-1}^2 \int_0^t \left\| \boldsymbol{\Lambda}^{(j)}(\tau) - \Pi \boldsymbol{\Lambda}^{(j)}(\tau) \right\|_{\Gamma} d\tau. \quad (6.20b)$$

For the discrete boundary traces $\boldsymbol{\lambda}^h := \gamma_\nu^{\text{int}} \dot{\mathbf{V}}^h + \dot{\boldsymbol{\beta}}^1$ and $\boldsymbol{\psi}^h := \gamma^{\text{int}} U + \boldsymbol{\beta}^0$, we get:

$$\left\| \boldsymbol{\psi}(t) - \boldsymbol{\psi}^h(t) \right\|_{\mathcal{H}^{-1/2}} \leq C \sum_{j=-1}^2 \int_0^t \left\| \boldsymbol{\Lambda}^{(j)}(\tau) - \Pi \boldsymbol{\Lambda}^{(j)}(\tau) \right\|_{\Gamma} d\tau, \quad (6.21a)$$

$$\left\| \boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^h(t) \right\|_{\mathcal{H}^{-1/2}} \leq C \sum_{j=0}^3 \int_0^t \left\| \boldsymbol{\Lambda}^{(j)}(\tau) - \Pi \boldsymbol{\Lambda}^{(j)}(\tau) \right\|_{\Gamma} d\tau. \quad (6.21b)$$

Proof. We consider the difference $\mathbf{E} := (\mathbf{E}_1, \mathbf{E}_2) := \mathbf{X} - \mathbf{X}^h$. This function solves the differential equation $\dot{\mathbf{E}} = A_\star \mathbf{E}$, and the transmission conditions satisfied by \mathbf{X}^h give the following transmission conditions for \mathbf{E} :

$$\begin{aligned} \llbracket \gamma \mathbf{E}_1 \rrbracket(t) - \llbracket \gamma U \rrbracket(t) - \boldsymbol{\beta}^0(t) &\in \mathcal{Y}_h, & \llbracket \gamma_\nu \mathbf{E}_2 \rrbracket(t) - \llbracket \gamma_\nu \mathbf{V} \rrbracket(t) - \boldsymbol{\beta}^1(t) &\in \mathcal{X}_h, \\ \gamma^{\text{ext}} \mathbf{E}_1(t) &\in \mathcal{X}_h^\circ, & \gamma_\nu^{\text{ext}} \mathbf{E}_2(t) &\in \mathcal{Y}_h^\circ \end{aligned}$$

for all $t \geq 0$. Secondly, we notice that these conditions are invariant under subtracting discrete functions, i.e., for $\boldsymbol{\chi}_h(t) \in \mathcal{Y}_h$, $\boldsymbol{\mu}_h(t) \in \mathcal{X}_h$, the following conditions are equivalent:

$$\begin{aligned} \llbracket \gamma \mathbf{E}_1 \rrbracket(t) - \boldsymbol{\psi}(t) + \boldsymbol{\chi}_h(t) &\in \mathcal{Y}_h, & \llbracket \gamma_\nu \mathbf{E}_2 \rrbracket(t) - \partial_t^{-1} \boldsymbol{\lambda}(t) + \boldsymbol{\mu}_h(t) &\in \mathcal{X}_h, \\ \gamma^{\text{ext}} \mathbf{E}_1(t) &\in \mathcal{X}_h^\circ, & \gamma_\nu^{\text{ext}} \mathbf{E}_2(t) &\in \mathcal{Y}_h^\circ, \end{aligned}$$

where we also inserted the definitions of $\boldsymbol{\psi}$ and $\boldsymbol{\lambda}$ to shorten notation. Using the best approximation operator Π , setting $(\boldsymbol{\chi}_h(t), \boldsymbol{\mu}_h(t)) := \Pi \boldsymbol{\Lambda}(t)$ and applying the stability estimate of Theorem 6.14 gives the estimate (6.20a). To get (6.20b), we use the estimate on $\dot{\mathbf{E}}$ from Theorem 6.14 and the equation $A_\star \mathbf{E} = \dot{\mathbf{E}}$ to bound the stronger norm. To bound the traces $\boldsymbol{\lambda}$ and $\boldsymbol{\psi}$, we use the bounds on \mathbf{E} (see also (6.19)) and the trace theorem from Proposition 2.25. \square

6.5 Runge-Kutta discretization

Since we have already established the semigroup setting, we can immediately write down the Runge-Kutta approximation $\underline{\mathbf{X}}^{h,k} := (\underline{\mathbf{U}}^{h,k}, \underline{\mathbf{V}}^{h,k})$ as

$$\underline{\mathbf{X}}^{h,k}(t_n) = \mathbf{X}^{h,k}(t_n) \mathbf{1} + k[\mathcal{Q} \otimes A_\star] \underline{\mathbf{X}}^{h,k}(t_n), \quad (6.22a)$$

$$\underline{B} \underline{\mathbf{X}}^{h,k}(t_n) = \left(\left(\partial_t^k \right)^{-1} \boldsymbol{\Sigma}(t_n), 0, 0 \right), \quad (6.22b)$$

$$\mathbf{X}^{h,k}(t_{n+1}) = R(\infty) \mathbf{X}^{h,k}(t_n) + \mathbf{b}^T \mathcal{Q}^{-1} \underline{\mathbf{X}}^{h,k}(t_n), \quad (6.22c)$$

and $\mathbf{X}^{h,k}(0) := 0$. Note that instead of using the boundary data $\Theta = (\beta_0, \dot{\beta}_1)$, we used the differentiated version $\Sigma := (\dot{\beta}^0, \dot{\beta}^1)$ as introduced in Problem 6.6. If instead of using $(\partial_t^k)^{-1}\Sigma$ we used $(\beta^0, (\partial_t^k)^{-1}\dot{\beta}^1)$, we would end up with a system equivalent to Problem 6.5.

Theorem 6.17. *The system (6.22) and the fully discrete system of integral equations (6.12) are equivalent in the following sense:*

(i) *If $\underline{\mathbf{X}}^{h,k}$ solves (6.22), then the traces $\underline{\lambda}^{h,k} := \partial_t^k \llbracket \gamma_\nu \underline{\mathbf{V}}^{h,k} \rrbracket + \dot{\beta}^1$ and $\underline{\psi}^{h,k} := \llbracket \gamma \underline{U}^{h,k} \rrbracket + (\partial_t^k)^{-1} \dot{\beta}^0$ solve Problem 6.6.*

(ii) *If $\underline{\Lambda}^{h,k} := (\underline{\lambda}^{h,k}, \underline{\psi}^{h,k})$ solves Problem 6.6, then*

$$\underline{U}^{h,k} := \mathbf{G}(\partial_t^k) \mathbf{Q}_\kappa^{-1} \left(\underline{\Lambda}^{h,k} - \mathfrak{J}(\partial_t^k) \Sigma \right), \quad \underline{\mathbf{V}}^{h,k} := \mathbf{T}_\kappa \nabla \left[(\partial_t^k)^{-1} \underline{U}^{h,k} \right], \quad (6.23)$$

$$U^{h,k} := \mathbf{G} \left[\underline{U}^{h,k} \right], \quad \mathbf{V}^{h,k} := \mathbf{G} \left[\underline{\mathbf{V}}^{h,k} \right] \quad (6.24)$$

solves (6.22).

Proof. We show that both approximations satisfy the following problem in the Z -domain:

$$A_* \widehat{\mathbf{Y}} - s_k^2 \widehat{\mathbf{Y}} = 0, \quad (6.25)$$

$$\llbracket \gamma \widehat{\mathbf{Y}}_1 \rrbracket + s_k^{-1} \mathcal{L} \left[\dot{\beta}^0 \right] \in [\mathcal{Y}_h]^m, \quad \llbracket \gamma_\nu \widehat{\mathbf{Y}}_2 \rrbracket + s_k^{-1} \mathcal{L} \left[\dot{\beta}^1 \right] \in [\mathcal{X}_h]^m, \quad (6.26)$$

$$\gamma^{\text{ext}} \widehat{\mathbf{Y}}_1 \in [\mathcal{X}_h^\circ]^m, \quad \gamma_\nu^{\text{ext}} \widehat{\mathbf{Y}}_2 \in [\mathcal{Y}_h^\circ]^m, \quad (6.27)$$

with $s_k := \frac{\delta(z)}{k} \in \mathbb{C}^{m \times m}$. Since the boundary conditions fit the semigroup setting, this problem has a unique solution via Lemma 3.35. The fact that the Z -transform of the Runge-Kutta approximation in (6.22) solves this problem is an easy consequence of Lemma 3.19 and the definition of $(\partial_t^k)^{-1}$.

By Theorem 6.14, Problem 6.6 has a unique solution. From the proof of the theorem it is also straight forward to see that we have some *a priori* estimate

$$\left\| \underline{\Lambda}^{h,k}(t_n) \right\|_\Gamma \leq C(k) \left(\max_{0 \leq \tau \leq t_n} \|\Sigma(\tau)\|_\Gamma + \max_{0 \leq \tau \leq t_{n-1}} \left\| \underline{\Lambda}^{h,k}(\tau) \right\|_\Gamma \right).$$

Thus, as long as $\Sigma(\tau)$ is uniformly bounded, the Z transform $\widehat{\underline{\Lambda}}^{h,k}(z) := \mathcal{L} \left[\underline{\Lambda}^{h,k} \right](z)$ exists (for $|z|$ sufficiently small). The requirement on Σ does not impact the result, as we may consider a modified problem where Σ is cut off for sufficiently large times without impacting the approximating sequences at finite times $t_n \leq T$.

Using the Z -transform for the potentials, we get that the Z -transform of the function $\underline{U}^{h,k}$ as defined in (6.23) satisfies

$$\begin{aligned} \widehat{\underline{U}}^{h,k}(z) &= \mathbf{G}(s_k) \mathbf{Q}_\kappa^{-1} \left(\widehat{\underline{\Lambda}}^{h,k} - \mathfrak{J}(s_k) \widehat{\Sigma} \right), \\ \widehat{\underline{\mathbf{V}}}^{h,k}(z) &= \mathbf{T}_\kappa \nabla \left[s_k^{-1} \widehat{\underline{U}}^{h,k} \right] \end{aligned}$$

with $\widehat{\underline{\Lambda}}^{h,k} := \mathcal{Z} [\underline{\Lambda}^{h,k}]$ and $\widehat{\underline{\Sigma}} := \mathcal{Z} [\underline{\Sigma}]$. Since the potentials solve the Helmholtz equation, we get that $\widehat{\underline{Y}} := \left(\widehat{\underline{U}}^{h,k}, \widehat{\underline{V}}^{h,k} \right)$ solves (6.25). We write $\widehat{\underline{\Lambda}}^{h,k} := \left(\widehat{\underline{\lambda}}^{h,k}, \widehat{\underline{\psi}}^{h,k} \right)$ for the two components. From the jump properties, we get $\llbracket \gamma \widehat{\underline{U}}^{h,k} \rrbracket = \widehat{\underline{\psi}}^{h,k} - s_k^{-1} \mathcal{Z} [\dot{\underline{\beta}}^0]$, as well as $\llbracket \gamma_\nu \widehat{\underline{V}}^{h,k} \rrbracket = s_k^{-1} \widehat{\underline{\lambda}}^{h,k} - s_k^{-1} \mathcal{Z} [\dot{\underline{\beta}}^1]$, and therefore (6.26) since $\widehat{\underline{\lambda}}^{h,k} \in \mathcal{X}_h$ and $\widehat{\underline{\psi}}^{h,k} \in \mathcal{Y}_h$.

To see the conditions on the exterior traces, we calculate using the frequency analogue of (6.9):

$$\begin{bmatrix} s_k \gamma_\nu^{\text{ext}} \widehat{\underline{V}}^{h,k} \\ \gamma^{\text{ext}} \widehat{\underline{U}}^{h,k} \end{bmatrix} = Q_\kappa \left(C(s_k) - \frac{1}{2} \text{Id} \right) Q_\kappa^{-1} \left(\widehat{\underline{\Lambda}}^{h,k} - \mathfrak{T}(s_k) \widehat{\underline{\Sigma}} \right),$$

which is in $[\mathcal{Y}_h^\circ \times \mathcal{X}_h^\circ]^m$ by (6.12). Since the multiplication with s_k in the first component leaves the space \mathcal{Y}_h° invariant, this concludes the proof. \square

Now that we have established the equivalence of the TDBIE formulation and the Runge-Kutta approximation, we can use the theory from Section 3.4 to easily derive error estimates. For notational convenience, we introduce the interpolation space $\mathbb{H}_\mu := [\mathbb{H}, \text{dom}(A)]_{\mu,2}$, where $\mu \geq 0$. For $p \in \mathbb{N}$, we write $\|\cdot\|_{p,T,\mathbb{H}_\mu} := \|\cdot\|_{C^p([0,T],\mathbb{H}_\mu)}$ and if $p = 0$, we also write $\|\cdot\|_{T,\mathbb{H}_\mu} := \|\cdot\|_{C([0,T],\mathbb{H}_\mu)}$.

Theorem 6.18. *Set $\underline{\Xi} := (\underline{\Sigma}, 0, 0)$. Let \mathbf{X}^h be the solution to Problem 6.3 and assume $\mathbf{X}^h \in C^{p+3}([0, T], \mathbb{H}_\mu)$ as well as $\mathcal{E}_B[\underline{\Xi}^{(\ell)}] \in C([0, T], \mathbb{H}_\mu)$ for $\ell = 0, \dots, p+3$. Set $\alpha := 1$ if the Runge-Kutta method satisfies (3.37) and $\alpha := 0$ otherwise. If $\mathbf{X}^{h,k}$ is the solution to (6.22), then the following error estimates hold for $0 < t_n \leq T$:*

$$\left\| \mathbf{X}^h(t_n) - \mathbf{X}^{h,k}(t_n) \right\|_{\mathbb{H}} \lesssim T k^{\min(q+\mu+\alpha,p)} \left[\left\| \mathbf{X}^h \right\|_{p+1,T,\mathbb{H}_\mu} + \sum_{\ell=0}^{p+1} \left\| \mathcal{E}_B [\underline{\Xi}^{(\ell-1)}] \right\|_{T,\mathbb{H}_\mu} \right], \quad (6.28a)$$

$$\left\| \mathbf{X}^h(t_n) - \mathbf{X}^{h,k}(t_n) \right\|_{\mathbb{V}} \lesssim T k^{\min(q+\mu+\alpha,p)} \left[\left\| \mathbf{X}^h \right\|_{p+2,T,\mathbb{H}_\mu} + \sum_{\ell=0}^{p+2} \left\| \mathcal{E}_B [\underline{\Xi}^{(\ell-1)}] \right\|_{T,\mathbb{H}_\mu} \right], \quad (6.28b)$$

where p and q denote the classical and stage order of the Runge-Kutta method employed. For the traces in (6.12), the following estimates can be shown:

$$\left\| \boldsymbol{\psi}^h(t_n) - \boldsymbol{\psi}^{h,k}(t_n) \right\|_{\mathcal{H}^{1/2}} \lesssim T k^{\min(q+\mu+\alpha,p)} \left[\left\| \mathbf{X}^h \right\|_{p+2,T,\mathbb{H}_\mu} + \sum_{\ell=0}^{p+2} \left\| \mathcal{E}_B [\underline{\Xi}^{(\ell-1)}] \right\|_{T,\mathbb{H}_\mu} \right]. \quad (6.28c)$$

If the Runge-Kutta method is stiffly accurate, we can also estimate $\boldsymbol{\lambda}^h$ by:

$$\left\| \boldsymbol{\lambda}^h(t_n) - \boldsymbol{\lambda}^{h,k}(t_n) \right\|_{\mathcal{H}^{-1/2}} \lesssim T k^{\min(q+\mu-1,p-1)+\alpha} \left[\left\| \mathbf{X}^h \right\|_{p+3,T,\mathbb{H}_\mu} + \sum_{\ell=0}^{p+3} \left\| \mathcal{E}_B [\underline{\Xi}^{(\ell-1)}] \right\|_{T,\mathbb{H}_\mu} \right]. \quad (6.28d)$$

The constants depend on the Runge-Kutta method and μ .

Proof. (6.28a) follows from Theorem 3.37, since we are in the situation of an integrated boundary condition. In order to get an estimate on $A_\star(\mathbf{X}^h - \underline{\mathbf{X}}^{h,k})$, we use Lemma 3.36 and Lemma 2.13. Namely, we get that $\underline{\mathbf{Y}}^{h,k} := A_\star \underline{\mathbf{X}}^{h,k}$ solves

$$\underline{\mathbf{Y}}^{h,k}(t_n) = \mathbf{Y}^{h,k}(t_n)\mathbf{1} + k[\mathcal{Q} \otimes A_\star]\underline{\mathbf{Y}}^{h,k}(t_n), \quad (6.29a)$$

$$B\underline{\mathbf{X}}^{h,k}(t_n) = \left(\partial_t^k(\partial_t^k)^{-1}\boldsymbol{\Sigma}(t_n), 0, 0\right) = (\boldsymbol{\Sigma}(t_n), 0, 0), \quad (6.29b)$$

$$\mathbf{Y}^{h,k}(t_{n+1}) = R(\infty)\mathbf{Y}^{h,k}(t_n) + \mathbf{b}^T \mathcal{Q}^{-1}\underline{\mathbf{Y}}^{h,k}(t_n), \quad (6.29c)$$

while $Y^h := A_\star \mathbf{X}^h$ solves $\dot{Y}^h = A_\star Y^h$ and $BY^h = (\boldsymbol{\Sigma}, 0, 0)$. Therefore we can apply Proposition 3.32 or 3.33 to get the estimate

$$\left\|A_\star \left[\mathbf{X}^h(t_n) - \underline{\mathbf{X}}^{h,k}(t_n)\right]\right\|_{\mathbb{H}} \lesssim Tk^{\min(q+\mu+\alpha,p)} \left[\left\|\mathbf{X}^h\right\|_{p+2,T,\mathbb{H}_\mu} + \sum_{\ell=0}^{p+2} \left\|\mathcal{E}_B \left[\boldsymbol{\Xi}^{(\ell-1)}\right]\right\|_{T,\mathbb{H}_\mu} \right].$$

Together with the \mathbb{H} -estimate we can estimate the \mathbb{V} -norm. The trace theorem then immediately gives (6.28c).

In order to estimate $\boldsymbol{\lambda}^h - \boldsymbol{\lambda}^{h,k}$ we need to control $\operatorname{div}(\dot{\mathbf{V}}^h) - \operatorname{div}(\mathbb{G}[\partial_t^k \underline{\mathbf{V}}^{h,k}])$. This can be estimated by applying Lemma 3.36 again and using Theorem 3.38 to estimate the differentiated error. \square

Remark 6.19. *This is the point, where it became advantageous to use Problem 6.6 instead of Problem 6.5 for the discretization. In Problem 6.5, we would already have a consistency error when estimating $\partial_t^k \underline{\mathbf{X}}^{h,k}$ in (6.29). Thus, we would have to use Theorem 3.38 instead of Proposition 3.33 and would lose a convergence rate of one for the \mathbb{V} -norm (and get no result for $\boldsymbol{\lambda}$).*

6.5.1 Determining the value of μ

The convergence rates in Theorem 6.18 depend on the condition $\mathcal{E}_B[\boldsymbol{\Xi}^{(\ell)}] \in C([0, T], \mathbb{H}_\mu)$ for some interpolation space \mathbb{H}_μ between \mathbb{H} and $\operatorname{dom}(A)$. While the regularity of $\mathcal{E}_B[\boldsymbol{\Xi}]$ is not an issue since it is in $\operatorname{dom}(A_\star)$ by construction, the boundary condition $B(\mathcal{E}_B[\boldsymbol{\Xi}]) = 0$ is violated, unless when dealing with the homogeneous problem $\boldsymbol{\Xi} = 0$. The goal of this section is to determine under what conditions we can expect convergence rates with $\mu > 0$.

In order to not get lost in notation, we often silently identify spaces of the form $\prod_{\ell=0}^L [H_\ell]^d$ with $\left[\prod_{\ell=0}^L H_\ell\right]^d$. This should not cause confusion within the context. We will need some Sobolev spaces in addition to $\mathcal{H}^{\pm 1/2}$ and \mathcal{H}^1 . For $s \in [0, 1]$, we write:

$$\begin{aligned} \mathcal{H}^s(\mathbb{R}^d \setminus \Gamma) &:= \prod_{\ell=0}^L H^s(\mathbb{R}^d \setminus \partial\Omega_\ell), & \tilde{\mathcal{H}}^s(\mathbb{R}^d \setminus \Gamma) &:= \prod_{\ell=0}^L \tilde{H}^s(\mathbb{R}^d \setminus \partial\Omega_\ell), \\ \mathcal{H}^s(\Gamma) &:= \prod_{\ell=0}^L H^s(\partial\Omega_\ell). \end{aligned}$$

In order to differentiate “volume” and “boundary” spaces, this means we will now add the parameters $\mathbb{R}^d \setminus \Gamma$ and Γ respectively, whereas up to now it was clear from the index of the space whether it was a boundary or volume space.

We start with the following lemma, relating the interpolation spaces \mathbb{H}_μ to Sobolev norms:

Lemma 6.20. *Let $0 \leq \mu < 1/2$. Then the following estimates holds for all $(u, \mathbf{V}) \in \mathcal{H}^1(\mathbb{R}^d \setminus \Gamma) \times [\mathcal{H}^\mu(\mathbb{R}^d \setminus \Gamma)]^d$:*

$$\|(u, \mathbf{V})\|_{\mathbb{H}_\mu} \lesssim \|u\|_{\mathcal{H}^1(\mathbb{R}^d \setminus \Gamma)} + \|\mathbf{V}\|_{[\mathcal{H}^\mu(\mathbb{R}^d \setminus \Gamma)]^d},$$

where $\mathbb{H}_\mu := [\mathbb{H}, \text{dom}(A)]_{\mu,2}$. The implied constant depends only on the geometries $\Omega_0, \dots, \Omega_\ell$ and μ .

Proof. It is easy to see that the space with zero traces satisfies:

$$\tilde{\mathbb{H}}_1 := \tilde{\mathcal{H}}^1(\mathbb{R}^d \setminus \Gamma) \times [\tilde{\mathcal{H}}^1(\mathbb{R}^d \setminus \Gamma)]^d \subseteq \ker(B) = \text{dom}(A).$$

Thus, by interpolating the identity operator we estimate, using Proposition 2.20 to split the norms of the two components:

$$\|(u, \mathbf{V})\|_{\mathbb{H}_\mu} \leq \|(u, \mathbf{V})\|_{[\mathbb{H}, \tilde{\mathbb{H}}_1]_{\mu,2}} \lesssim \|u\|_{[\mathcal{H}^0(\mathbb{R}^d \setminus \Gamma), \tilde{\mathcal{H}}^1(\mathbb{R}^d \setminus \Gamma)]_{\mu,2}} + \|\mathbf{V}\|_{[[\mathcal{H}^0(\mathbb{R}^d \setminus \Gamma), \tilde{\mathcal{H}}^1(\mathbb{R}^d \setminus \Gamma)]_{\mu,2}]^d}.$$

For $\mu < 1/2$, the interpolation spaces $\mathcal{H}^s(\mathbb{R}^d \setminus \Gamma)$ and $\tilde{\mathcal{H}}^s(\mathbb{R}^d \setminus \Gamma)$ coincide with equivalent norms by Proposition 2.22 (again using Proposition 2.20 to deal with the product spaces). Thus, we can further estimate:

$$\begin{aligned} \|(u, \mathbf{V})\|_{\mathbb{H}_\mu} &\lesssim \|u\|_{[\mathcal{H}^0(\mathbb{R}^d \setminus \Gamma), \mathcal{H}^1(\mathbb{R}^d \setminus \Gamma)]_{\mu,2}} + \|\mathbf{V}\|_{[[\mathcal{H}^0(\mathbb{R}^d \setminus \Gamma), \mathcal{H}^1(\mathbb{R}^d \setminus \Gamma)]_{\mu,2}]^d} \\ &\lesssim \|u\|_{\mathcal{H}^\mu(\mathbb{R}^d \setminus \Gamma)} + \|\mathbf{V}\|_{[\mathcal{H}^\mu(\mathbb{R}^d \setminus \Gamma)]^d}. \end{aligned} \quad \square$$

With this previous lemma, we have answered the question of admissible μ for the case of functions which have additional Sobolev regularity. As a final tool to show lower bounds for μ , we investigate the mapping properties of the lifting operator \mathcal{E}_B by splitting the result into a part in $\text{dom}(A)$ and a part with higher regularity.

Lemma 6.21. *Let $0 \leq \mu < 1/2$, and assume*

$$\Xi =: (\xi_1, \xi_2, \xi_3, \xi_4) \in \mathbb{M}_\mu := \mathcal{H}^{1/2+\mu}(\Gamma) \times \mathcal{H}^{-1/2+\mu}(\Gamma) \times \mathcal{H}^{1/2+\mu}(\Gamma) \times \mathcal{H}^{-1/2+\mu}(\Gamma).$$

We identify Ξ with its induced functional in $\mathbb{M} = (\mathcal{Y}'_h) \times (\mathcal{X}'_h) \times \mathcal{X}'_h \times \mathcal{Y}'_h$ via the Riesz- and restriction maps. Then $\mathcal{E}_B[\Xi]$ can be estimated in the following stronger norms:

$$\|\mathcal{E}_B[\Xi]\|_{\mathbb{H}_\mu} \leq C \|\Xi\|_{\mathbb{M}_\mu},$$

where C depends on the geometries $\Omega_0, \dots, \Omega_\ell$ and μ , but not on the spaces \mathcal{X}_h or \mathcal{Y}_h .

Proof. We write $\mathcal{E}_B[\Xi] =: (U, \mathbf{V})$ for the two components. Since the transmission conditions of $\text{dom}(A)$ are independent for U and \mathbf{V} , we can perform the interpolation of the spaces independently in each of the two components. We note that for each domain Ω_ℓ there exist continuous lifting operators $\tilde{\mathcal{E}}_\ell^D : [H^{1/2+\mu}(\partial\Omega_\ell)]^2 \rightarrow H^{1+\mu}(\mathbb{R}^d \setminus \partial\Omega_\ell)$ and $\tilde{\mathcal{E}}_\ell^N : [H^{-1/2+\mu}(\partial\Omega_\ell)]^2 \rightarrow [H^\mu(\mathbb{R}^d \setminus \partial\Omega_\ell)]^d$ satisfying the following lifting properties:

$$\begin{aligned} \llbracket \gamma_\ell \tilde{\mathcal{E}}_\ell^D(\eta, \chi) \rrbracket &= \eta, & \gamma_\ell^{\text{ext}} \tilde{\mathcal{E}}_\ell^D(\eta, \chi) &= \chi, \\ \llbracket \gamma_{\nu, \ell} \tilde{\mathcal{E}}_\ell^N(\lambda, \mu) \rrbracket &= \lambda, & \gamma_{\nu, \ell}^{\text{ext}} \tilde{\mathcal{E}}_\ell^N(\lambda, \mu) &= \mu. \end{aligned}$$

They can be constructed by the Dirichlet- and Neumann-liftings from Proposition 2.26. We collect the operators on each subdomain into the diagonal operators $\tilde{\mathcal{E}}^D$ and $\tilde{\mathcal{E}}^N$.

We define functions $\hat{U} := \tilde{\mathcal{E}}^D(\xi_1, \xi_3)$ and $\hat{\mathbf{V}} := \tilde{\mathcal{E}}^N(\xi_2, \xi_4)$. By construction, we have that $B(\hat{U}, \hat{\mathbf{V}}) = \Xi$, which implies by linearity $(U - \hat{U}, \mathbf{V} - \hat{\mathbf{V}}) \in \ker(B) = \text{dom}(A)$. We calculate, using the fact that the interpolation spaces are nested:

$$\begin{aligned} \|(U, \mathbf{V})\|_{\mathbb{H}_\mu} &\leq \|(U - \hat{U}, \mathbf{V} - \hat{\mathbf{V}})\|_{\mathbb{H}_\mu} + \|(\hat{U}, \hat{\mathbf{V}})\|_{\mathbb{H}_\mu} \\ &\lesssim \|(U - \hat{U}, \mathbf{V} - \hat{\mathbf{V}})\|_{\mathbb{H}_1} + \|(\hat{U}, \hat{\mathbf{V}})\|_{\mathbb{H}_\mu}. \end{aligned}$$

The first term can be estimated by the $\mathcal{H}^1(\mathbb{R}^d) \times \mathcal{H}^{\text{div}}$ -norm of (U, \mathbf{V}) and $(\hat{U}, \hat{\mathbf{V}})$, which are both bounded by the \mathbb{M}_0 norm of Ξ due to the continuity of \mathcal{E}_B and the lifting operators. To estimate $\|(\hat{U}, \hat{\mathbf{V}})\|_{\mathbb{H}_\mu}$, we apply Lemma 6.20 together with the stricter regularity assumptions and mapping properties of the liftings. \square

Corollary 6.22. *For $\mu \in [0, 1/2)$ and $m \in \mathbb{N}_0$, let $\gamma u^{\text{inc}} \in C^m([0, T], H^{1/2+\mu}(\partial\Omega_0))$ and $\partial_\nu u^{\text{inc}} \in C^{m-1}([0, T], H^{-1/2+\mu}(\partial\Omega_0))$, and assume that the solution to Problem 6.3 satisfies $\mathbf{X}^h \in C^m([0, T], \mathbb{V})$. Then \mathbf{X}^h is also in $C^m([0, T], \mathbb{H}_\mu)$.*

Proof. For $\Xi := -(\beta^0, \beta^1, 0, 0)$, and $\beta^0 := (\gamma u^{\text{inc}}, 0, \dots, 0)$, $\beta^1 := (\kappa_0 \partial_\nu \partial_t^{-1} u^{\text{inc}}, 0, \dots, 0)$ we write $\mathbf{X}^h = (\mathbf{X}^h - \mathcal{E}_B[\Xi]) + \mathcal{E}_B[\Xi]$. Due to the boundary conditions on \mathbf{X}^h we get $(\mathbf{X}^h(t) - \mathcal{E}_B[\Xi(t)]) \in \text{dom}(A)$ and we can estimate:

$$\begin{aligned} \left\| (\mathbf{X}^h(t) - \mathcal{E}_B[\Xi(t)]) \right\|_{\mathbb{H}_\mu} &\leq \left\| (\mathbf{X}^h(t) - \mathcal{E}_B[\Xi(t)]) \right\|_{\mathbb{H}} + \left\| A(\mathbf{X}^h(t) - \mathcal{E}_B[\Xi(t)]) \right\|_{\mathbb{H}} \\ &\leq \left\| \mathbf{X}^h(t) \right\|_{\mathbb{V}} + \|\Xi(t)\|_{\mathbb{M}}. \end{aligned}$$

The term $\|\mathcal{E}_B[\Xi(t)]\|_{\mathbb{H}_\mu}$ can be estimated by Lemma 6.21. \square

6.6 Convergence of the fully discrete scheme

In order to quantify the convergence rates of the full discretization, we make the following assumption on the spaces \mathcal{X}_h and \mathcal{Y}_h (see Section 6.3.1 for a way to satisfy this assumption).

Assumption 6.23. For a parameter $r \in \mathbb{N}_0$, the discrete spaces \mathcal{X}_h and \mathcal{Y}_h satisfy the following approximation property for all $\boldsymbol{\lambda} := (\lambda_\ell)_{\ell=0}^L \in \mathcal{X}$ with $\lambda_\ell \in H_{\text{pw}}^{r+1}(\partial\Omega_\ell)$ and all $\boldsymbol{\psi} := (\psi_\ell)_{\ell=0}^L \in \mathcal{Y}$ with $\psi_\ell \in H_{\text{pw}}^{r+2}(\partial\Omega_\ell)$, such that the lifting in (6.6a) is continuous on Γ :

$$\inf_{\boldsymbol{\lambda}_h \in \mathcal{X}_h} \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_h\|_{\mathcal{H}^{-1/2}} \leq Ch^{r+3/2} \sum_{\ell=0}^L \|\lambda_\ell\|_{H_{\text{pw}}^r(\partial\Omega_\ell)}, \quad (6.30a)$$

$$\inf_{\boldsymbol{\psi}_h \in \mathcal{Y}_h} \|\boldsymbol{\psi} - \boldsymbol{\psi}_h\|_{\mathcal{H}^{1/2}} \leq Ch^{r+3/2} \sum_{\ell=0}^L \|\psi_\ell\|_{H_{\text{pw}}^{r+1}(\partial\Omega_\ell)}, \quad (6.30b)$$

where the constant C depends on r , and the geometry but not on h , $\boldsymbol{\lambda}$ or $\boldsymbol{\psi}$.

Then the following theorem holds:

Theorem 6.24. For $\mu \in [0, 1/2)$, let $\gamma u^{\text{inc}} \in C^{p+3}([0, T], H^{1/2+\mu}(\partial\Omega_0))$ and $\partial_n u^{\text{inc}} \in C^{p+3}([0, T], H^{-1/2+\mu}(\partial\Omega_0))$. Assume that the traces of the exact solution satisfy $\lambda_\ell \in C^3([0, T], H_{\text{pw}}^{r+1}(\partial\Omega_\ell))$ and $\psi_\ell \in C^3([0, T], H_{\text{pw}}^{r+2}(\partial\Omega_\ell))$ for a parameter $r \in \mathbb{N}_0$. Also assume that $\psi^{(\ell)}$ admits a lifting to $H^1(\mathbb{R}^d)$ which is continuous on Γ for $\ell = 0, \dots, 3$ (see Assumption 6.23).

Let p denote the classical order of the Runge-Kutta method and q its stage order. Set $\alpha := 1$ if the method satisfies assumption (3.37) (i.e. $|r(z)| < 1$ for $0 \neq z \in i\mathbb{R}$ and $r(\infty) \neq 1$), and set $\alpha := 0$ otherwise. Let Assumption 6.23 be satisfied for \mathcal{X}_h and \mathcal{Y}_h for $r \in \mathbb{N}_0$.

Let $\mathbf{X}^{h,k}$, $\boldsymbol{\lambda}^{h,k}$ and $\boldsymbol{\psi}^{h,k}$ denote the solutions to Problem 6.6. Then the following estimates hold for $t_n = nk$ with $t_n \leq T$:

$$\left\| \mathbf{X}(t_n) - \mathbf{X}^{h,k}(t_n) \right\|_{\mathbf{V}} \leq C(\mathbf{X}) \left(h^{r+3/2} + k^{\min(q+\alpha+\mu, p)} \right), \quad (6.31a)$$

$$\left\| \boldsymbol{\psi}(t_n) - \boldsymbol{\psi}^{h,k}(t_n) \right\|_{\mathcal{H}^{1/2}} \leq C(\mathbf{X}) \left(h^{r+3/2} + k^{\min(q+\mu+\alpha, p)} \right). \quad (6.31b)$$

If the method is stiffly accurate, we get:

$$\left\| \boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^{h,k}(t_n) \right\|_{\mathcal{H}^{-1/2}} \leq C(\mathbf{X}) \left(h^{r+3/2} + k^{\min(q+\mu-1, p-1)+\alpha} \right). \quad (6.31c)$$

The constant $C(\mathbf{X})$ depends on the exact solution, the incoming wave, the geometry, T , the Runge-Kutta method, and the constants in Assumption 6.23, but is independent of h and k .

Proof. We estimate

$$\left\| \mathbf{X}(t_n) - \mathbf{X}^{h,k}(t_n) \right\|_{\mathbf{V}} \lesssim \left\| \mathbf{X}(t_n) - \mathbf{X}^h(t_n) \right\|_{\mathbf{V}} + \left\| \mathbf{X}^h(t_n) - \mathbf{X}^{h,k}(t_n) \right\|_{\mathbf{V}}.$$

The convergence of the semi-discretization in space is quasi-optimal by Theorem 6.16. The convergence with respect to time can be estimated by Theorem 6.18, where Lemma 6.21 and Corollary 6.22 tell us that we may use the value $\mu \geq 0$ determined by the regularity assumptions. The bounds on $\boldsymbol{\psi} - \boldsymbol{\psi}^{h,k}$ follows from the continuity of the trace operator. The trace on $\boldsymbol{\lambda} - \boldsymbol{\lambda}^{h,k}$ follows along the same lines but using the stronger bounds proved in Theorems 6.16 and 6.18. \square

6.7 Numerical examples

In order to verify our theoretical findings, we implemented the proposed schemes (6.11) and (6.12) using the 2D boundary element code developed by F.-J. Sayas and his work group at the University of Delaware. For the Convolution Quadrature, we used the algorithm presented in [BS09] as implemented in the deltaBEM package (see [Say17]). As the underlying geometry, we used a simple checkerboard consisting of 2×2 squares with differing wave numbers, see Figure 6.2.

In order to be able to quantify the convergence, we prescribe an exact solution in the following way: On each subdomain Ω_ℓ for $\ell = 1, \dots, L$, the solution u_ℓ is given as a plane wave with

$$u_\ell(\mathbf{x}, t) := G(\mathbf{d}_\ell \cdot \mathbf{x} - \kappa_\ell(t - t_{\text{lag}})), \quad \text{where} \quad G(t) := e^{-2t/\alpha} \sin(t).$$

Here, \mathbf{d}_ℓ is the direction the wave is traveling. We chose the following parameters for our example:

$$\mathbf{d}_\ell := \begin{cases} \frac{1}{\sqrt{2}}(1, -1)^T & \ell \text{ is even} \\ \frac{1}{\sqrt{2}}(1, 1)^T & \ell \text{ is odd} \end{cases},$$

$t_{\text{lag}} = 5/2$, and $\alpha := 1/4$. On the exterior, we chose $u_0 := 0$ in order to not have to worry about radiation conditions.

The boundary traces were chosen in the following way, using the function $\phi(t) := t^9 e^{-2t}$ to ensure homogeneous initial conditions:

$$\psi(\mathbf{x}, t) := \phi(t) \sin(x_1) \cos(x_2) \quad \text{and} \quad \lambda(\mathbf{x}, t) := \phi(t) \cos(x_1) \sin(x_2).$$

(These functions are to be understood as functions on the skeleton. λ is then built by restricting to the subdomains and multiplying with a sign function as is done in Section 6.3.1). The boundary data β^0 and β^1 were then calculated accordingly in order to yield these solutions.

Example 6.25. *In this example, we are interested in the convergence with regards to the time discretization. Therefore, we fix a fine uniform mesh with $h \sim 0.03125$ and use $r = 4$, i.e. quartic polynomials for the discontinuous space and quintic for the continuous splines. We applied a two-stage RadauIIa method which satisfies $q = 2$ and $p = 3$. By Theorem 6.24, we expect order $\mathcal{O}(k^3)$ for the Dirichlet trace and $\mathcal{O}(k^{2.5-\varepsilon})$ for the Neumann trace when using (6.12). As a comparison, we also compute the solutions using (6.11). Figure 6.3 shows the result. Most notably, it shows that when using (6.12), the Neumann trace outperforms our predictions and converges with the full classical order. We also see that using (6.11) gives a reduced order of 2 when approximating λ .*

Example 6.26. *We perform the same experiment as in Example 6.25, but use a 3-stage RadauIIa method. We expect orders $\mathcal{O}(k^{4.5-\varepsilon})$ and $\mathcal{O}(k^{3.5-\varepsilon})$ for the Dirichlet and Neumann traces respectively. Again the method (6.12) outperforms our expectations, giving the full classical order 5, while using (6.11) gives a reduced rate.*

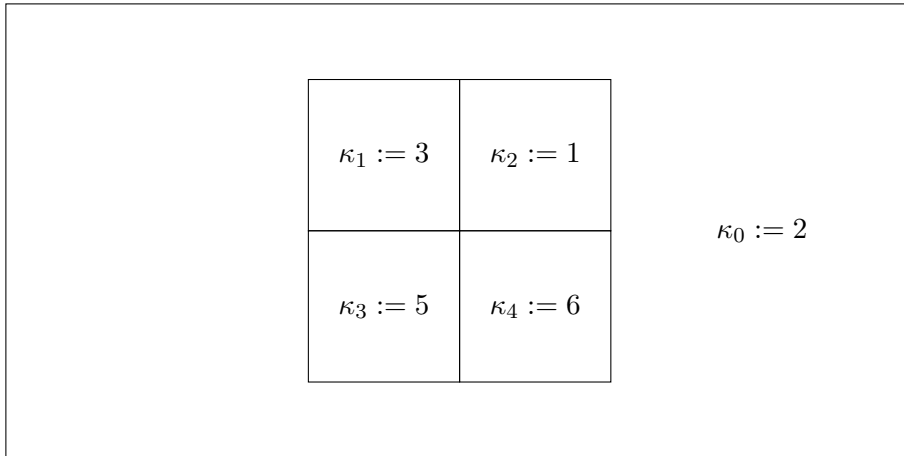


Figure 6.2: Example geometry and wave-numbers used throughout Section 6.7

We also look at the convergence rate with respect to the space discretization.

Example 6.27. *We use the same model problem as Example 6.25, but we fix the time discretization at $k \approx 0.015$ using a 3-stage RadauIIa method. We vary the approximation in space by performing successive uniform refinements of the grid, and compare different polynomial degrees $s = 0, \dots, 3$. Since it is easier to compute, we consider the L^2 -norm of the errors. As we see in Figure 6.5, we get the optimal convergence rates, up to an error of $\approx 10^{-6}$, at which point other error contributions prohibit further convergence.*

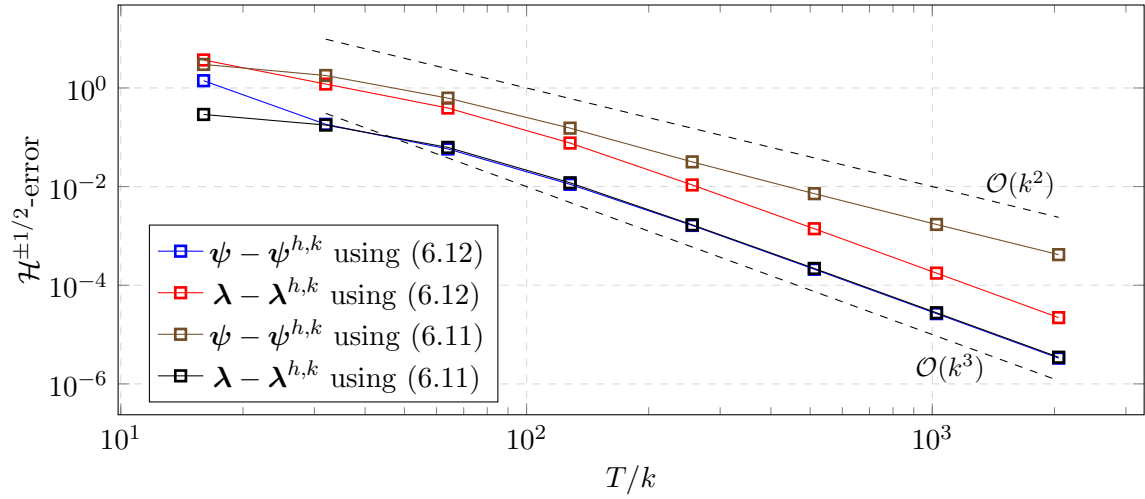


Figure 6.3: Convergence rates using a 2-stage RadauIIa method. Comparison of discretization schemes

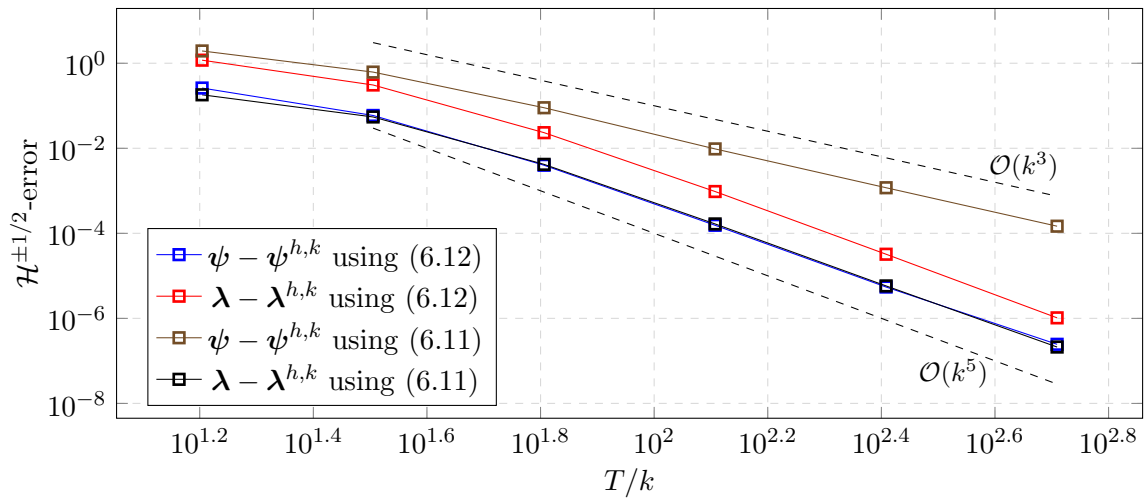


Figure 6.4: Convergence rates using a 3-stage RadauIIa method. Comparison of discretization schemes

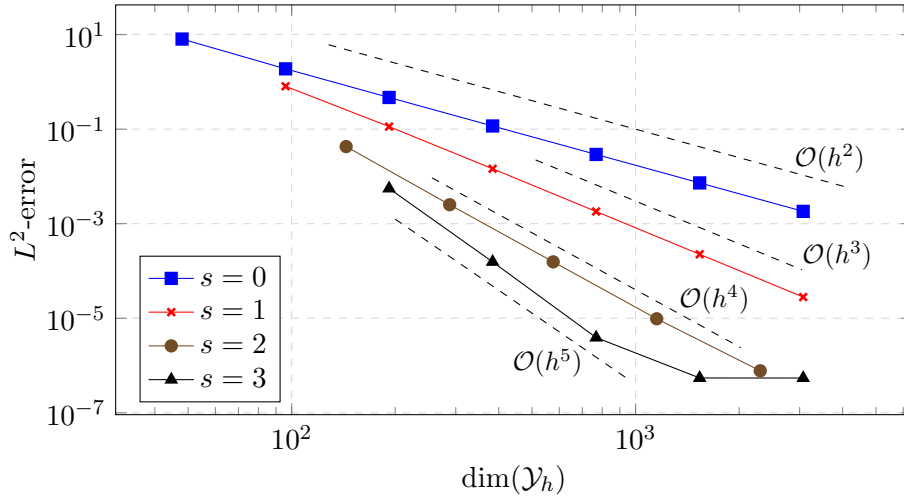
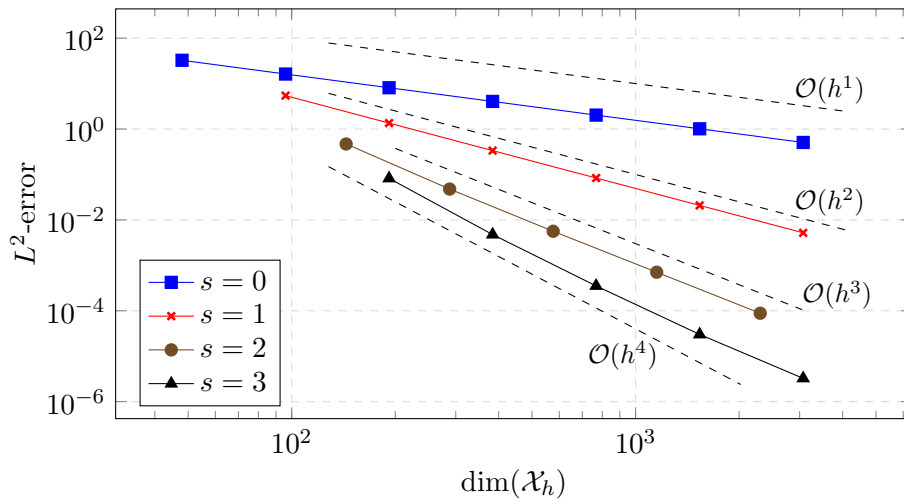
(a) Convergence $\psi - \psi^h$ (b) Convergence $\lambda - \lambda^h$

Figure 6.5: Convergence rates w.r.t. the spatial discretization

Bibliography

- [AF03] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*. Second. Vol. 140. Pure and Applied Mathematics (Amsterdam). Elsevier/Academic Press, Amsterdam, 2003, pp. xiv+305.
- [Ant+08] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, and A. Schädle. A review of transparent and artificial boundary conditions techniques for linear and non-linear Schrödinger equations. *Commun. Comput. Phys.* (4)4: (2008), pp. 729–796.
- [AP03] I. Alonso-Mallo and C. Palencia. Optimal orders of convergence for Runge-Kutta methods and linear, initial boundary value problems. *Appl. Numer. Math.* (44)1-2: (2003), pp. 1–19.
- [Aur+15] M. Aurada, M. Feischl, T. Führer, M. Karkulik, and D. Praetorius. Energy norm based error estimators for adaptive BEM for hypersingular integral equations. *Appl. Numer. Math.* (2015), published online first.
- [Ayg+04] K. Aygun, B. C. Fischer, J. Meng, B. Shanker, and E. Michielssen. A fast hybrid field-circuit simulator for transient analysis of microwave circuits. *IEEE Transactions on Microwave Theory and Techniques* (52)2: (Feb. 2004), pp. 573–583.
- [Ban10] L. Banjai. Multistep and multistage convolution quadrature for the wave equation: algorithms and experiments. *SIAM J. Sci. Comput.* (32)5: (2010), pp. 2964–2994.
- [Ban15] L. Banjai. Some new directions in the applications of time-domain boundary integral operators. *Proceedings of Waves, Karlsruhe, Germany* (2015), pp. 27–32.
- [BCT82] P. Brenner, M. Crouzeix, and V. Thomée. Single-step methods for inhomogeneous linear differential equations in Banach space. *RAIRO Anal. Numér.* (16)1: (1982), pp. 5–26.
- [BD86] A. Bamberger and T. H. Duong. Formulation variationnelle pour le calcul de la diffraction d’une onde acoustique par une surface rigide. *Math. Methods Appl. Sci.* (8)4: (1986), pp. 598–608.
- [Ber68] S. J. Bernau. The square root of a positive self-adjoint operator. *J. Austral. Math. Soc.* (8) (1968), pp. 17–36.
- [BH86] A. Bamberger and T. Ha-Duong. Formulation variationnelle espace-temps pour le calcul par potentiel retardé d’une onde acoustique. *Math. Meth. Appl. Sci.* (8) (1986), pp. 405–435.

- [BK14] L. Banjai and M. Kachanovska. Fast convolution quadrature for the wave equation in three dimensions. *J. Comput. Phys.* (279) (2014), pp. 103–126.
- [BL11] L. Banjai and C. Lubich. An error analysis of Runge-Kutta convolution quadrature. *BIT* (51)3: (2011), pp. 483–496.
- [BL17] L. Banjai and C. Lubich. Runge–Kutta convolution coercivity and its use for time-dependent boundary integral equations. *ArXiv e-prints* (Feb. 2017). arXiv: 1702.08385 [math.NA].
- [BLM11] L. Banjai, C. Lubich, and J. M. Melenk. Runge-Kutta convolution quadrature for operators arising in wave propagation. *Numer. Math.* (119)1: (2011), pp. 1–20.
- [BLS15a] L. Banjai, A. R. Laliena, and F.-J. Sayas. Fully discrete Kirchhoff formulas with CQ-BEM. *IMA J. Numer. Anal.* (35)2: (2015), pp. 859–884.
- [BLS15b] L. Banjai, C. Lubich, and F.-J. Sayas. Stable numerical coupling of exterior and interior problems for the wave equation. *Numer. Math.* (129)4: (2015), pp. 611–646.
- [BR17] L. Banjai and A. Rieder. Convolution quadrature for the wave equation with a nonlinear impedance boundary condition. *Math. Comp.* (2017). Forthcoming.
- [Bre83] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques appliquées pour la maîtrise. Théorie et applications. [Theory and applications]. Masson, Paris, 1983, pp. xiv+234.
- [BS08] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Third. Vol. 15. Texts in Applied Mathematics. Springer, New York, 2008, pp. xviii+397.
- [BS09] L. Banjai and S. Sauter. Rapid solution of the wave equation in unbounded domains. *SIAM J. Numer. Anal.* (47)1: (2008/09), pp. 227–249.
- [BT79] P. Brenner and V. Thomée. On rational approximations of semigroups. *SIAM J. Numer. Anal.* (16)4: (1979), pp. 683–694.
- [BY14] R. E. Bank and H. Yserentant. On the H^1 -stability of the L_2 -projection onto finite element spaces. *Numer. Math.* (126)2: (2014), pp. 361–381.
- [CH13] X. Claeys and R. Hiptmair. Multi-trace boundary integral formulation for acoustic scattering by composite structures. *Comm. Pure Appl. Math.* (66)8: (2013), pp. 1163–1201.
- [Cla11] X. Claeys. A single trace integral formulation of the second kind for acoustic scattering. In: *ETH, Seminar of Applied Mathematics Research*. 2011, pp. 2011–14.
- [Clé75] P. Clément. Approximation by finite element functions using local regularization. English. *Rev. Franc. Automat. Inform. Rech. Operat., R* (9)2: (1975), pp. 77–84.
- [Cos88a] M. Costabel. A symmetric method for the coupling of finite elements and boundary elements. In: *The mathematics of finite elements and applications, VI (Uxbridge, 1987)*. Academic Press, London, 1988, pp. 281–288.

-
- [Cos88b] M. Costabel. Boundary integral operators on Lipschitz domains: elementary results. *SIAM J. Math. Anal.* (19)3: (1988), pp. 613–626.
- [Cro76] M. Crouzeix. Sur les méthodes de Runge Kutta pour l’approximation des problèmes d’évolution. In: *Computing methods in applied sciences and engineering (Second Internat. Sympos., Versailles, 1975), Part 1*. Springer, Berlin, 1976, 206–223. Lecture Notes in Econom. and Math. Systems, Vol. 134.
- [CS85] M. Costabel and E. Stephan. A direct boundary integral equation method for transmission problems. *J. Math. Anal. Appl.* (106)2: (1985), pp. 367–413.
- [CT87] M. Crouzeix and V. Thomée. The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces. *Math. Comp.* (48)178: (1987), pp. 521–532.
- [GE16] J.-L. Guermond and A. Ern. Finite element quasi-interpolation and best approximation. *ESAIM: Mathematical Modelling and Numerical Analysis* (Oct. 2016).
- [Gen83] J. Geniusz-Mikusiński. Hypernumbers. I. Algebra. *Studia Math.* (77)1: (1983), pp. 3–16.
- [Gra12] P. J. Graber. Uniform boundary stabilization of a wave equation with nonlinear acoustic boundary conditions and nonlinear boundary damping. *J. Evol. Equ.* (12)1: (2012), pp. 141–164.
- [Gri11] P. Grisvard. *Elliptic problems in nonsmooth domains*. Vol. 69. Classics in Applied Mathematics. Reprint of the 1985 original, With a foreword by Susanne C. Brenner. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011, pp. xx+410.
- [GV13] G. H. Golub and C. F. Van Loan. *Matrix computations*. Fourth. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, 2013, pp. xiv+756.
- [Han90] H. D. Han. A new class of variational formulations for the coupling of finite and boundary element methods. *J. Comput. Math.* (8)3: (1990), pp. 223–232.
- [Has+15] M. Hassell, T. Qiu, T. Sanchez-Vizuet, and F.-J. Sayas. A new and improved analysis of the time domain boundary integral operators for acoustics. *ArXiv e-prints* (Dec. 2015). arXiv: 1512.02919 [math.NA].
- [HJ02] H. Haddar and P. Joly. Stability of thin layer approximation of electromagnetic waves scattering by linear and nonlinear coatings. *J. Comput. Appl. Math.* (143)2: (2002), pp. 201–236.
- [HJ12] R. Hiptmair and C. Jerez-Hanckes. Multiple traces boundary integral formulation for Helmholtz transmission problems. *Adv. Comput. Math.* (37)1: (2012), pp. 39–91.
- [HKS09] W. Hackbusch, W. Kress, and S. A. Sauter. Sparse convolution quadrature for time domain boundary integral formulations of the wave equation. *IMA J. Numer. Anal.* (29)1: (2009), pp. 158–179.

- [HNW93] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I*. Second. Vol. 8. Springer Series in Computational Mathematics. Nonstiff problems. Springer-Verlag, Berlin, 1993, pp. xvi+528.
- [HS16] M. E. Hassell and F.-J. Sayas. A fully discrete BEM-FEM scheme for transient acoustic waves. *Comput. Methods Appl. Mech. Engrg.* (309) (2016), pp. 106–130.
- [HW08] G. C. Hsiao and W. L. Wendland. *Boundary integral equations*. Vol. 164. Applied Mathematical Sciences. Springer-Verlag, Berlin, 2008, pp. xx+618.
- [HW10] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*. Vol. 14. Springer Series in Computational Mathematics. Stiff and differential-algebraic problems, Second revised edition, paperback. Springer-Verlag, Berlin, 2010, pp. xvi+614.
- [JK81] D. S. Jerison and C. E. Kenig. The Neumann problem on Lipschitz domains. *Bull. Amer. Math. Soc. (N.S.)* (4)2: (1981), pp. 203–207.
- [JN80] C. Johnson and J.-C. Nédélec. On the coupling of boundary integral and finite element methods. *Math. Comp.* (35)152: (1980), pp. 1063–1079.
- [KM15] M. Karkulik and J. M. Melenk. Local high-order regularization and applications to *hp*-methods. *Comput. Math. Appl.* (70)7: (2015), pp. 1606–1639.
- [Kōm67] Y. Kōmura. Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Japan* (19) (1967), pp. 493–507.
- [LO93] C. Lubich and A. Ostermann. Runge-Kutta methods for parabolic equations and convolution quadrature. *Math. Comp.* (60)201: (1993), pp. 105–131.
- [LS09] A. R. Laliena and F.-J. Sayas. Theoretical aspects of the application of convolution quadrature to scattering of acoustic waves. *Numer. Math.* (112)4: (2009), pp. 637–678.
- [LS13] M. Lopez-Fernandez and S. Sauter. Generalized convolution quadrature with variable time stepping. *IMA J. Numer. Anal.* (33)4: (2013), pp. 1156–1175.
- [LS15] M. Lopez-Fernandez and S. Sauter. Generalized convolution quadrature with variable time stepping. Part II: Algorithm and numerical results. *Appl. Numer. Math.* (94) (2015), pp. 88–105.
- [LS16] M. Lopez-Fernandez and S. Sauter. Generalized convolution quadrature based on Runge-Kutta methods. *Numer. Math.* (133)4: (2016), pp. 743–779.
- [LT93] I. Lasiecka and D. Tataru. Uniform boundary stabilization of semilinear wave equations with nonlinear boundary damping. *Differential Integral Equations* (6)3: (1993), pp. 507–533.
- [Lub88a] C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.* (52)2: (1988), pp. 129–145.
- [Lub88b] C. Lubich. Convolution quadrature and discretized operational calculus. II. *Numer. Math.* (52)4: (1988), pp. 413–425.

-
- [Lub94] C. Lubich. On the multistep time discretization of linear initial-boundary value problems and their boundary integral equations. *Numer. Math.* (67)3: (1994), pp. 365–389.
- [McL00] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000, pp. xiv+357.
- [MR17] J. M. Melenk and A. Rieder. Runge-Kutta convolution quadrature and FEM-BEM coupling for the time-dependent linear Schrödinger equation. *J. Integral Equations Appl.* (29)1: (2017), pp. 189–250.
- [Neu51] J. von Neumann. Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. *Math. Nachr.* (4) (1951), pp. 258–281.
- [Nev78] O. Nevanlinna. On the convergence of difference approximations to nonlinear contraction semigroups in Hilbert spaces. *Math. Comp.* (32)142: (1978), pp. 321–334.
- [OR92] A. Ostermann and M. Roche. Runge-Kutta methods for partial differential equations and fractional orders of convergence. *Math. Comp.* (59)200: (1992), pp. 403–420.
- [Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Vol. 44. Applied Mathematical Sciences. Springer-Verlag, New York, 1983, pp. viii+279.
- [Pet89] T. von Petersdorff. Boundary integral equations for mixed Dirichlet, Neumann and transmission problems. *Math. Methods Appl. Sci.* (11)2: (1989), pp. 185–213.
- [Qiu16] T. Qiu. Time domain boundary integral equation methods in acoustics, heat diffusion and electromagnetism. PhD thesis. University of Delaware, 2016.
- [QS16] T. Qiu and F.-J. Sayas. The Costabel-Stephan system of boundary integral equations in the time domain. *Math. Comp.* (85)301: (2016), pp. 2341–2364.
- [RS75] M. Reed and B. Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975, pp. xv+361.
- [RS80] M. Reed and B. Simon. *Methods of modern mathematical physics. I. Second. Functional analysis*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1980, pp. xv+400.
- [RS90] F. Riesz and B. Sz.-Nagy. *Functional analysis*. Dover Books on Advanced Mathematics. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original. Dover Publications, Inc., New York, 1990, pp. xii+504.
- [Rud87] W. Rudin. *Real and complex analysis*. Third. McGraw-Hill Book Co., New York, 1987, pp. xiv+416.
- [Rud91] W. Rudin. *Functional analysis*. Second. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991, pp. xviii+424.

- [Say16] F.-J. Sayas. *Retarded Potentials and Time Domain Boundary Integral Equations*. Vol. 50. Springer Series in Computational Mathematics. Springer, Heidelberg, 2016.
- [Say17] F.-J. Sayas. *deltaBEM package*. 2017. URL: www.math.udel.edu/~fjsayas/deltaBEM/index.html.
- [Sch02] A. Schädle. Non-reflecting boundary conditions for the two-dimensional Schrödinger equation. *Wave Motion* (35)2: (2002), pp. 181–188.
- [Sch14] J. Schöberl. C++11 Implementation of Finite Elements in NGSolve. In: *ASC Report 30/2014*. Institute for Analysis and Scientific Computing, Vienna University of Technology, 2014.
- [Sho97] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*. Vol. 49. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1997, pp. xiv+278.
- [Śmi+15] W. Śmigaj, T. Betcke, S. Arridge, J. Phillips, and M. Schweiger. Solving boundary integral problems with BEM++. *ACM Trans. Math. Software* (41)2: (2015), Art. 6, 40.
- [SS11] S. A. Sauter and C. Schwab. *Boundary element methods*. Vol. 39. Springer Series in Computational Mathematics. Translated and expanded from the 2004 German original. Springer-Verlag, Berlin, 2011, pp. xviii+561.
- [SS17] S. A. Sauter and M. Schanz. Convolution quadrature for the wave equation with impedance boundary conditions. *J. Comput. Phys.* (334) (2017), pp. 442–459.
- [Ste08] O. Steinbach. *Numerical approximation methods for elliptic boundary value problems*. Finite and boundary elements, Translated from the 2003 German original. Springer, New York, 2008, pp. xii+386.
- [Ste70] E. Stein. *Singular integrals and differentiability properties of functions*. Princeton University Press, 1970.
- [SV14] S. Sauter and A. Veit. Retarded boundary integral equations on the sphere: exact and numerical solution. *IMA J. Numer. Anal.* (34)2: (2014), pp. 675–699.
- [SZ90] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* (54)190: (1990), pp. 483–493.
- [Tar07] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*. Vol. 3. Lecture Notes of the Unione Matematica Italiana. Springer, Berlin; UMI, Bologna, 2007, pp. xxvi+218.
- [Tes12] G. Teschl. *Ordinary differential equations and dynamical systems*. Vol. 140. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012, pp. xii+356.
- [Tho06] V. Thomée. *Galerkin finite element methods for parabolic problems*. Second. Vol. 25. Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2006, pp. xii+370.

- [Tri95] H. Triebel. *Interpolation theory, function spaces, differential operators*. Second. Heidelberg: Johann Ambrosius Barth, 1995, p. 532.
- [Wer07] D. Werner. *Funktionalanalysis*. 6., korrigierte Auflage. Springer-Verlag, Berlin, 2007.
- [Wou66] A. Wouk. A note on square roots of positive operators. *SIAM Rev.* (8) (1966), pp. 100–102.
- [Yos80] K. Yosida. *Functional analysis*. Sixth. Vol. 123. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin-New York, 1980, pp. xii+501.
- [Yos84] K. Yosida. *Operational calculus*. Vol. 55. Applied Mathematical Sciences. A theory of hyperfunctions. Springer-Verlag, New York, 1984, pp. x+170.

Alexander Rieder

Date of birth 28.03.1989
Place of birth Bruneck
Nationality Italian
E-Mail alexander.rieder@tuwien.ac.at
Homepage <http://www.asc.tuwien.ac.at/~arieder>



Education

- Since 10/2013 **Doctoral program**, Dissipation and Dispersion in Nonlinear PDEs, *TU Wien*.
2001-2013 **Master studies**, Technische Mathematik, *TU Wien*.
2008-2011 **Bachelor studies**, Technische Mathematik, *TU Wien*.
2008 **School**, *Gewerbeoberschule J.Ph. Fallmerayer, Brixen (Italy)*.

Publications

1. L. Banjai and A. Rieder. Convolution quadrature for the wave equation with a non-linear impedance boundary condition. *Math. Comp.* (2017). Forthcoming.
2. T. Führer, J. M. Melenk, D. Praetorius, and A. Rieder. Optimal additive Schwarz methods for the *hp*-BEM: the hypersingular integral operator in 3D on locally refined meshes. *Comput. Math. Appl.* (70)7: (2015), pp. 1583–1605.
3. J. M. Melenk and A. Rieder. Runge-Kutta convolution quadrature and FEM-BEM coupling for the time-dependent linear Schrödinger equation. *J. Integral Equations Appl.* (29)1: (2017), pp. 189–250.