Temporal Discretization of Constrained Partial Differential Equations

vorgelegt von
Master of Science Technomathematik
Christoph Zimmer
ORCID: 0000-0002-4589-9000

an der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Wilhelm Stannat
Gutachter:	Prof. Dr. Volker Mehrmann
Gutachter:	Prof. Dr. Christopher Beattie
Gutachterin:	Prof. Dr. Caren Tischendorf

Tag der wissenschaftlichen Aussprache: 20. April 2021

Berlin 2021

Abstract

This thesis is devoted to the application and analysis of time integration schemes for differential-algebraic equations (DAEs) stated in (abstract) Banach spaces. The existence, uniqueness, and regularity of solutions of these so-called operator DAEs are analyzed with the help of temporal discretization methods. The convergence behavior of the time-discrete approximations and their convergence orders are addressed as well.

Besides being of interest as a generalization of the concept of DAEs to the infinite-dimensional setting, operator DAEs are an abstract approach for the analysis of constrained partial differential equations (PDEs) in their weak form. The constraints on the solution of the PDEs are possibly given by spatial differential operators like the divergence-free condition on the velocity field in the incompressible Navier-Stokes equations. Examples of constrained PDEs appear in all kinds of physical fields such as fluid dynamics, thermodynamics, electrodynamics, mechanics, chemical kinetics, as well as in multi-physical applications where different physical domains are coupled.

The first main results of this thesis cover the existence, uniqueness, and regularity of solutions of semi-linear, semi-explicit operator DAEs. In this analysis, the challenges known for DAEs and PDEs have to be tackled simultaneously. These include a limited set of feasible initial values, requirements on the temporal and spatial regularity of the data, and a high sensitivity to perturbations. For operator DAEs with time-independent operators, continuity results for the solutions in the data are used to extend well-known existence, uniqueness, and regularity results to systems with less regular or state-dependent right-hand sides. Similar results for operator DAEs with time-dependent operators are derived by studying the convergence of time-discrete solutions obtained by the implicit Euler method. In this study, time-varying inner products as well as time-dependent kernels of the constraints operators complicate the analysis.

As the second main topic, the convergence of the temporal discretization of semi-explicit operator DAEs by implicit, algebraically stable Runge-Kutta methods and explicit exponential integrators is analyzed. As expected from the theory of DAEs and PDEs, the convergence properties depend strongly on the assumed temporal and spatial regularity of the data, vary for the single variables, and differ from finite-dimensional systems. For Runge-Kutta schemes, a regularization is introduced and the strong convergence of the time-discrete approximations under minimal regularity assumptions is proven. A convergence order of $\mathsf{q}+1$ and of $\mathsf{q}+1/2$ is shown for the state and the Lagrange multiplier, respectively. Here, q denotes the stage order of the Runge-Kutta scheme. For explicit exponential integrators, order conditions for methods up to order three are derived for the state of semi-linear operator DAEs. In addition, an approximation of the Lagrange multiplier is introduced whose convergence order is reduced by half an order. For both classes of integration schemes, sufficient conditions are formulated which increase the convergence order. The results are supported by numerical examples.

Zusammenfassung

Diese Arbeit befasst sich mit der Anwendung von Zeitintegrationsverfahren auf differentiell-algebraische Gleichungen (DAEs), welche in (abstrakten) Banachräumen gestellt sind. Die Existenz, Eindeutigkeit sowie die Glattheit der Lösungen von diesen sogenannten Operator-DAEs werden mit Hilfe von Einschrittverfahren analysiert. Sowohl das Konvergenzverhalten der zeitdiskreten Approximationen als auch deren Konvergenzordnung sind ebenfalls Untersuchungsschwerpunkt.

Operator-DAEs stellen eine Verallgemeinerung des DAE-Begriffs auf Systeme in unendlich-dimensionalen Vektorräumen dar. Dabei lassen sich mit ihnen partielle Differentialgleichungen mit Nebenbedingung (PDAEs) untersuchen, die in ihrer schwachen Formulierung gestellt sind. Diese Nebenbedingungen an die Lösung von den partiellen Differentialgleichungen (PDEs) sind möglicherweise durch Differentialoperatoren gegeben. Zum Beispiel wird in den inkompressiblen Navier-Stokes-Gleichungen an das Geschwindigkeitsfeld gefordert, dass dessen Divergenz verschwindet. PDAEs treten in vielen Anwendungsbereiche auf, wie zum Beispiel in der Fluiddynamik, Thermodynamik, Elektrodynamik, Mechanik, chemischen Kinetik sowie in multiphysikalischen Anwendungen, in denen verschiedene physikalische Domänen miteinander gekoppelt werden.

Die Existenz, Eindeutig und Glattheit von Lösungen von semi-linearen, semi-expliziten Operator-DAEs sind die ersten Untersuchungsschwerpunkte dieser Arbeit. Dabei müssen die Herausforderungen, die von DAEs und PDEs bekannt sind, gleichzeitig gemeistert werden. Diese Herausforderungen umfassen unter anderem eine eingeschränkte Menge zulässiger Anfangswerte, Glattheitsanforderungen an die Daten sowohl in der Zeit als auch im Ort, sowie eine starke Störungsempfindlichkeit. Für Operator-DAEs mit zeitunabhängigen Operatoren wird die stetige Abhängigkeit der Lösungen von den Daten genutzt, um wohlbekannte Existenz-, Eindeutigkeits- und Glattheitsresultate auf Systeme mit schwächeren Voraussetzungen an die rechten Seiten oder mit zustandsabhängigen rechten Seiten zu erweitern. Ähnliche Ergebnisse für Operator-DAEs mit zeitabhängigen Operatoren werden mittels einer Konvergenzuntersuchung von zeitdiskreten Lösungen bewiesen. Die zeitdiskreten Lösungen entspringen dabei der Diskretisierung durch das implizite Euler Verfahren. Zeitvariante Skalarprodukte sowie zeitabhängige Kerne des Operators, der die algebraischen Nebenbedingungen stellt, erschweren dabei die Untersuchung.

Die Konvergenz der zeitlichen Diskretisierung von semi-expliziten Operator-DAEs durch implizite, algebraisch stabile Runge-Kutta-Verfahren sowie durch explizite, exponentielle Integratoren wird im zweiten Hauptteil dieser Arbeit betrachtet. Wie von der Theorie der DAEs und PDEs zu erwarten ist, hängen die Konvergenzeigenschaften stark von der zeitlichen und örtlichen Glattheit der Daten ab, variieren für die einzelnen Zustandsvariablen und unterscheiden sich im Vergleich zu endlich-dimensionalen Systemen. Für die Runge-Kutta-Methoden wird eine Regularisierung eingeführt und die starke Konvergenz der zeitdiskreten Approximationen unter minimalen Annahmen an die Daten bewiesen. Es wird gezeigt, dass die Konvergenzordnung gleich $\mathsf{q} + 1$ für die Zustandsvariable ist und für den Lagrange-Multiplikator $\mathsf{q} + 1/2$ entspricht. Dabei bezeichnet q die Stufenordnung des Runge-Kutta Verfahrens. Für die Anwendung von exponentiellen Integratoren auf semi-lineare Operator-DAEs werden die Ordnungsbedingungen bis zur dritten Ordnung hergeleitet. Zusätzlich wird eine Approximation des Lagrange-Multiplikators eingeführt, dessen Konvergenzordnung um eine halbe Ordnung reduziert ist. Für beide Klassen von Integrationsverfahren werden hinreichende Bedingungen formuliert, die die Konvergenzordnung verbessern. Numerische Beispiele illustrieren die Ergebnisse.

Declaration of Personal Contribution

Several results of this thesis are already published in preprints, journal papers, or book chapters. A short overview, an indication where in this thesis they are used, and the contribution of the co-authors Dr. Robert Altmann, Universität Augsburg, and Arbi Moses Badlyan, Weierstraß-Institut für Angewandte Analysis und Stochastik, are given as follows.

- [AltZ18a] R. Altmann and C. Zimmer. "On the smoothing property of linear delay partial differential equations". *J. Math. Anal. Appl.* (2): 916–934 (2018)

  This paper analyzes the temporal smoothing of the solution of PDEs with a delay term. A result of this paper is mentioned in Subsection 4.3.2.2.

  Robert Altmann suggested the topic of this paper. Theorem 4.22 is the only contribution of this paper to this thesis and was originally proven by Christoph Zimmer as [AltZ18a, Th. 3.5].

- [AltZ18b] R. Altmann and C. Zimmer. "Runge-Kutta methods for linear semi-explicit operator differential-algebraic equations". *Math. Comp.* 87 (309): 149–174 (2018)

  This paper considers implicit, algebraically stable, L-stable Runge-Kutta methods applied to operator DAEs and their qualitative convergence. Lemma 3.5 as well as Sections 8.1 and 8.3 and of this thesis are essential copies of [AltZ18b, Lem. 2.4, Sec. 3 & 5]. Sections 5.1 and 8.2 are revised and extended versions of [AltZ18b, Sec. 2.1 & Sec. 4].

  Robert Altmann suggested the topic and supervised this paper. He proved the statements in [AltZ18b, Sec. 3], which are mentioned in Section 8.1. In particular, he suggested a proof of Lemma 8.2, which is given as a revised version in this thesis. Lemma 3.5 and the parts, on which the results of Sections 5.1, 8.2 and 8.3 are based, were originally elaborated by Christoph Zimmer.

- [AltZ18c] R. Altmann and C. Zimmer. *Time discretization schemes for hyperbolic systems on networks by ε-expansion.* ArXiv e-print 1810.04278. 2018

  This preprint studies singular perturbations of constrained hyperbolic PDEs with a small parameter ε such that setting ε to zero leads to constrained parabolic PDEs. It exploits the singular perturbation to derive temporal integration schemes for constrained hyperbolic PDEs. The results include convergence orders for the implicit Euler scheme as well as for algebraically and L-stable Runge-Kutta methods. The results in Section 8.5 are based on and extend [AltZ18c, Sec. 6].

  Robert Altmann suggested the topic of this paper. Theorem 8.37 and Lemma 8.41 of this thesis are based on [AltZ18c, Sec. 6], in the sense that the preprint considers slightly different systems but the ideas are transferred and extended. Christoph Zimmer originally elaborated [AltZ18c, Sec. 6].

- [MosZ18] A. Moses Badlyan and C. Zimmer. *Operator-GENERIC Formulation of Thermodynamics of Irreversible Processes.* ArXiv e-print 1807.09822. 2018

  This preprint extends the idea of GENERIC to a variational setting and illustrates this approach on the Navier-Stokes equations with reactive fluids. It is mentioned as an example in Section 2.2.

  Arbi Moses Badlyan suggested the topic of this preprint. Its results were worked out in a close cooperation between the two authors.

- [AltZ19] R. Altmann and C. Zimmer. "Time discretization of nonlinear hyperbolic systems on networks". *PAMM* 19 (1): e201900057 (2019)

  This conference proceeding extends the ideas of [AltZ18c] to a nonlinear constrained hyperbolic PDE. The results of this proceeding are not used in this thesis.

  The theoretical work as well as the numerical evaluation of [AltZ19] was done in a close cooperation between Robert Altmann and Christoph Zimmer.

- [AltZ20] R. Altmann and C. Zimmer. "Exponential integrators for semi-linear parabolic problems with linear constraints". In: *Progress in Differential-Algebraic Equations II*. ed. by T. Reis, S. Grundel, and S. Schöps. Springer International Publishing, Cham, 2020, pp. 137–164

  This book chapter analyzes exponential integrators applied to semi-linear operator DAEs. Sections 9.1, 9.2, and 9.4.2 of this thesis are essential copies of [AltZ20, Sec. 3, 4, & 5.2] with minor changes. Sections 5.2, 6.4, and 9.4.1 are revised and extended version of [AltZ20, Sec. 2.1, 2.3, & 5.1].

  Robert Altmann suggested the topic of this book chapter. He elaborated [AltZ20, Sec. 3.3 & Th. 4.1], which did not find their way into this thesis. The numerical experiment in Subsection 9.4.2 has been implemented mainly by him. The original proof of Corollary 3.9 was found by Robert Altmann. The application of the exponential Euler scheme and the exponential Runge scheme to constrained PDEs in Subsections 9.1.1 and 9.2.1, respectively, was worked out in a close cooperation between the two authors. The remaining results of the book chapter were originally proven by Christoph Zimmer.

Acknowledgements

First and foremost, my sincere thanks go to my supervisors Volker Mehrmann and Christopher Beattie for all their support during the last years and the provided freedom in the choice of my research directions.
I acknowledge funding from the Deutsche Forschungsgemeinschaft through the Collaborative Research Center 910 *Control of self-organizing nonlinear systems: Theoretical methods and concepts of application* (project number 163436311) and from the Einstein Foundation Berlin through the project *Model reduction for complex transport-dominated phenomena and reactive flows*.
My thanks go to my coauthors Robert Altmann and Arbi Moses Badlyan for pleasant and fruitful cooperation. You both gave me in your own ways the incentive to look deeper into nice mathematical problems and to learn something new every time. These learned tricks can be found all over this thesis.
I am very grateful to my dear friends Robert Altmann, Daniel Bankmann, Marine Froidevaux, Charlotte Schrape, Philipp Schulze, and Benjamin Unger for the discussions we had and for proofreading of this thesis. I wrote this thesis in a language which I believed to be English. You are the reason that someone else could believe that too.
Furthermore, I would like to thank my colleagues for the cheerful and positive atmosphere at TU Berlin. I will miss the chats in the hallway or in the cake and soup breaks. I am particularly grateful to Lena, Marine, and Matthew for turning the offices we shared into such joyful places.
At last, I would like to thank my family and my friends for their constant support and loving care.

Contents

1. Introduction

Modern modular modeling packages, such as MODELICA[1], MATLAB/SIMULINK[2], and SIMPACK[3], allows one to describe complex dynamical systems by simply interconnecting smaller models, facilitating rapid model development. The dynamical submodels together with interconnection constraints, e.g., the Kirchhoff's circuit laws [Kir45], lead to systems, which contain differential equations and algebraic constraints, so-called *differential-algebraic equations* (DAEs). This modeling approach simplifies the interconnection processes and preserves the system's sparsity but comes at the cost of analytical and numerical difficulties; see [AscP98; BreCP96; HaiLR89; HaiW96; KunM06; LamMT13; Sim13].

If the submodels are given by *partial differential equations* (PDEs), then the resulting coupled systems are mixtures of DAEs and PDEs, so-called *constrained PDEs*. Typical examples are flexible multibody systems [Sim00; Sim13], circuit networks [Tis96; Tis03], or the gas transfer in pipeline networks [EggKL+18; GruJH+14; JanT14]. For the latter two, the network structure describes the interconnection where transmission lines or the propagation of pressure waves are modeled by PDEs on the single edges [MagWT+00]. Outside of the interconnection context, constrained PDEs appear in the description of fluid flow problems. Often, one assumes that a fluid is incompressible leading to a divergence-free velocity field as constraint [EmmM13; Tem77]. Furthermore, constrained PDEs are used for the analysis of PDEs with nontrivial boundary conditions like moving or dynamical boundary conditions [Alt14; Alt19; HinPU+09].

It is well-known from the theory of DAEs that the combined presence of differential and algebraic equations comes with several difficulties, e.g., initial values and solutions restricted to manifolds as well as regularity conditions on inhomogeneities [KunM06; LamMT13; Rhe84]. In the temporal discretization of DAEs, these difficulties translate into high sensitivity to perturbations, reduction of convergence order, or even loss of convergence [HaiLR89; HaiW96; Pet82]. Here, constraints, which are only apparent after manipulations with differentiation, must be treated with special care [HaiW96; KunM06]. On the other hand, problems occurring in the analysis and simulation of PDEs include restrictions on the spatial regularity of initial values and inhomogeneities as well as reduced temporal convergence order in contrast to systems of finite dimensions [HocO10; OstR92; Tar06; Zei90a]. Since constrained PDEs generalize the concept of DAEs and PDEs, they suffer from all the difficulties mentioned above [Alt15; Deb04; EmmM13; LamMT13].

For the numerical simulation of constrained PDEs, one typically discretizes them first in space and then in time or the other way around. The first approach, the so-called *methods of lines* [Sch91], leads to finite-dimensional DAEs [Tem77; Tis03; Wei97]. Following the *Rothe method* [Rot30], i.e., discretizing the constrained PDEs first in time, produces sequences of stationary but infinite-dimensional problems; see [Alt15] and for the particular case of the incompressible Navier-Stokes equations [Emm01]. This ansatz simplifies the analysis of adaptive strategies in space [SchB98] and of temporal error bounds which are independent of the spatial mesh width [HocO10]. These mesh-independent bounds are vital for the simulation of infinite-dimensional systems, since the temporal convergence order differs in general for finite versus infinite-dimensional systems [HocO10; OstR92]. If the spatial mesh gets finer, the temporal convergence behavior can fade to the one for the infinite-dimensional system; cf. [ProR74]. Besides its application in the field of numeric,

[1] www.modelica.org
[2] www.mathworks.com
[3] www.simpack.com

the Rothe method is also applied for the investigation of solutions of dynamical systems [Emm04; Rou13].

The Thesis

One aim of this thesis is the analysis of a specific class of constrained PDEs. The existence of solutions of constrained PDEs with time-dependent coefficients is studied with the Rothe method using the implicit Euler scheme. Uniqueness and regularity results of the solution are also stated. Furthermore, the convergence behavior and the convergence order are investigated for the time-discrete approximation of the constrained PDEs using implicit, algebraically stable Runge-Kutta methods and explicit exponential integrators. In the study of well-posedness and the related convergence analysis, great importance is attached not only to the state u but also to the Lagrange multiplier λ; cf. (1.1). Here, the Lagrange multiplier measures an abstract force which is exerted on the solution of the PDE such that it satisfies the constraints [Bra07].

Operator Differential-Algebraic Equations

While a well-developed theory for the solvability of DAEs and PDEs exists in the literature, see the exemplary works, [HaiLR89; HaiW96; KunM06; LamMT13] and [Bra07; DauL92; Wlo87; Zei90a], respectively, the theory for constrained PDEs, in contrast, is quite limited. Results are available only for specific classes and mostly for systems with time-independent coefficients [Alt15; Deb04; EmmM13; Hei14; LamMT13; LucSE99; Mar97]. To receive statements on constrained PDEs, one promising approach is to mathematically interpret them as DAEs in Banach spaces, also referred to as *operator DAEs* [Alt15; EmmM13]. In this thesis, we consider semi-linear operator DAEs with semi-explicit structure, i.e., systems of the form

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) + \big(\mathcal{A} - \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}\big)u \; - \; \mathcal{B}^*\lambda = f, \tag{1.1a}$$

$$\mathcal{B}u \qquad\quad = g \tag{1.1b}$$

with linear operators $\mathcal{M}$, $\mathcal{A}$, $\mathcal{B}$ and right-hand sides $f = f(u)$ and g defined on appropriate Hilbert spaces. Here, the time derivatives in (1.1a) should be understood in a distributional sense. The well-posedness of operator DAEs of the form (1.1) with time-independent operators and state-independent right-hand sides are well-studied; see [FavY99; Rei06; Sho10] for a semigroup ansatz and [Alt15; EmmM13; Hei14; Zim15] for a variational approach as well as [DauL93; Tar06; Tem77] in the context of fluid dynamics. On the other hand, a rigorous analysis for the operator DAEs (1.1) with time-dependent operators is missing. In particular, for systems with a time-dependent operator $\mathcal{B}$ there are only results known if the kernel of $\mathcal{B}$ is time-independent [AltH18], or for a specific choice of $\mathcal{B}$ [Alt14]. For general operator DAEs (1.1) with state-dependent right-hand sides, no results are known outside of the context of fluid dynamics [Tar06; Tem77].

Time-Integration Schemes

As for the well-posedness problem, numerical integration schemes are well-studied for DAEs and PDEs, see [AscP98; BreCP96; HaiLR89; HaiW96; KunM06; LamMT13] and [Emm05; HocO10; HunV03; LubO95a; LubO95b; Tho06], respectively, and the references therein. All these time-stepping methods must respect the infinitely stiff nature of the problem arising either due to the algebraic constraints or due to the spatial differential operators of DAEs and PDEs, respectively, i.e., certain associated (operator) spectra have accumulation points at $-\infty$ [BreCP96; OstR92]. Classical examples of such a family of methods are implicit Runge-Kutta schemes. Runge-Kutta methods have been analyzed in the early work of Euler in the 18^{th} century and are among the best understood time-stepping methods [GonO99; HaiNW93; HaiW96]. Exponential integrators, on the other hand, are time-stepping methods based on the possibility to solve the linear part of semi-linear systems in an exact manner [HocO10]. As a result, large time steps are possible even for highly stiff systems, like stiff ordinary differential equations, PDEs, and DAEs of (differentiation) index one [Cer60; HocLS98; HocO10; Law67]. In particular, these systems can be discretized by explicit exponential integrators. For these integrators, all root-finding problems are linear and the approximation requires in total a priori known number of evaluations of the nonlinear right-hand side.

The literature on the temporal discretization of operator DAEs of the form (1.1) is quite limited. For linear systems only the implicit Euler method [Alt15], splitting schemes [AltO17], and discontinuous Galerkin methods [VouR19] have been studied. The analysis of the nonlinear case, on the other hand, has been restricted to the incompressible Navier-Stokes equations, where results are known for the implicit Euler scheme, the two-step BDF method [Emm00; Emm01] and exponential integrators (without a convergence analysis) [EdwTF+94; KooBG18; New03]. However, studies on the convergence behavior of the temporal discretization of (1.1) by general Runge-Kutta methods and exponential integrators are not available.

It is worthy to mention that a standard spatial discretization of (1.1) by finite elements leads to an index-2 DAE [Alt15]. This indicates that the temporally discretized operator DAE is sensitive to perturbations of the discrete right-hand side g [Alt15; HaiLR89]. Therefore, the high index must be considered in the construction of time-stepping methods by regularization techniques [AltH15; HaiW96; VouR19], also called index reduction in the DAE case, or by exploiting the structure of the system [AscP98].

Organization of This Thesis

This thesis is divided into three parts. In Part A, we introduce the essential mathematical concepts needed in this thesis. As mentioned above, operator DAEs extend the framework of DAEs to infinite-dimensional systems. Therefore, we properly define DAEs and the differentiation index. For the infinite-dimensional part, we recall basic functional analytic concepts like Gelfand triples, inf-sup stability, as well as real- and Banach space-valued functions and their generalized derivatives. These concepts allow us to consider infinite-dimensional dynamical systems, including integral equations as well as differential equations with or without constraints. Part A closes with the introduction of one-step methods, which are later applied to operator DAEs, namely Runge-Kutta methods and exponential integrators.

The existence of solutions for the operator DAE (1.1) is subject of Part B. At first, we study systems of the form (1.1) with time-independent operators. We generalize known results by abstract linear extensions and make statements on the existence, uniqueness, as well as the regularity of solutions. This allows us to consider semi-linear systems where the right-hand side f depends not only on time but also on the state u. Afterwards, we study operator DAEs (1.1) with time-dependent operators. The analysis is split into three separate steps. In the first step we restrict our analysis to time-dependent operators $\mathcal{M}$ and $\mathcal{A}$. In the second step we allow $\mathcal{A}$ and $\mathcal{B}$ to change over time. Both cases are analyzed by a discretization with the implicit Euler scheme. The investigation leads to Hilbert spaces equipped with a time-dependent inner product induced by the operator $\mathcal{M}$. We investigate whether functions with a generalized derivative in these Hilbert spaces have a continuous representative. In the case of a non-constant operator $\mathcal{B}$, we study time-dependent direct sums in Hilbert spaces. For this, we analyze a differential equation whose solution tracks the kernel of $\mathcal{B}$ over time. In the last step, we discuss the uniqueness of solutions and combine the results to operator DAEs (1.1) with time-dependent operators $\mathcal{M}$, $\mathcal{A}$, $\mathcal{B}$, and state-dependent right-hand side f.

Part C is devoted to the temporal discretization of the operator DAE (1.1). We first consider implicit, algebraically stable Runge-Kutta methods. Here, a regularization is introduced that maintains the saddle-point structure of the original problem (1.1). Starting with the implicit Euler scheme, we study the convergence of the time-discrete solution of the regularized system under minimal assumptions on the data. We generalize the obtained results to algebraically and L-stable Runge-Kutta methods and discuss the need of L-stability. The convergence order is studied afterwards under the assumption of more regular solutions. In particular, we analyze the decrease of the convergence rate of the Lagrange multiplier λ. The results are illustrated by means of two numerical examples.

In the second half of Part C, we investigate semi-explicit time-stepping schemes for semi-linear operator DAEs of the form (1.1), which are based on the idea of explicit exponential integrators. The associated algorithms are illustrated for the exponential Euler scheme and the exponential Runge schemes. We discuss how the solution can be approximated by linear, stationary and linear, transient saddle-point problems with homogeneous right-hand sides by using the structure of (1.1). Order conditions up to the order three are studied for explicit exponential integrators as well as the positive effect of spatially more regular right-hand sides. Since only the state u is approximated by the exponential integrators, we discuss the approximation of the Lagrange multiplier λ by a single additional saddle-point problem. Finally, we make comments on efficient computation and present numerical experiments for semi-linear systems, illustrating the convergence results obtained.

Part A.

Preliminaries

The analytic and numerical treatment of constrained partial differential equations combines the difficulties of partial differential equations (PDEs) and differential-algebraic equations (DAEs). Thus, we need the knowledge of different mathematical disciplines to get to the heart of constrained PDEs and their temporal discretization. In this part we provide the essential concepts.

In general, a system of equations, which combines differential and algebraic equations, suffers from hidden constraints, consistency requirements for the initial conditions, and unexpected regularity requirements [KunM06, Part I]. To understand these difficulties, we briefly review the theory of finite-dimensional DAEs in Chapter 2. We recall the concept of the differentiation index for linear DAEs and introduce so-called port-Hamiltonian descriptor systems as a special class of controlled DAEs. On the other hand, the analysis of the infinite-dimensional behavior of (constrained) PDEs requires several functional analytic concepts such as Gelfand triples, Sobolev spaces, and Bochner spaces. We introduce these among basic features of operators and Banach spaces as well as frequently used inequalities in Chapter 3. In Chapter 4 we consider dynamic infinite-dimensional equations. Since the operators in these equations are possibly time-dependent, we consider Nemytskiĭ operators. Afterwards we consider integral equations as well as differential equations with bounded and unbounded operators. We complete the chapter with a short introduction of operator DAEs, which is the abstract framework we use for the analysis of constrained PDEs and their temporal discretization. Later in Part C, we study Runge-Kutta methods and exponential integrators as numerical integration schemes for operator DAEs. We introduce these two families of time-integration schemes in Chapter 5.

Cor. 3.9 is a copy of [AltZ20, Lem. 3.1] and was shown by Robert Altmann. The author of this thesis originally proved Lem. 3.5 and Th. 4.22 as [AltZ18b, Lem. 2.4] and [AltZ18a, Th. 3.5].

Notation In the whole thesis we use $\mathbb{R}$ for the set of real numbers, and $\mathbb{R}_{\geq 0}$ ($\mathbb{R}_{>0}$) for its subset of non-negative (positive) real numbers. The set of non-negative (positive) integers is denoted by $\mathbb{N}_0$ ($\mathbb{N}$). For the restriction of $f\colon X \to Y$ to a subset $Z \subset X$, we write $f\big|_Z\colon Z \subset X \to Y$. For a linear map A, its kernel and its image are denoted by $\ker A$ and $\operatorname{im} A$, respectively.

2. Linear Differential-Algebraic Equations

The most general form of a *differential-algebraic equation* (*DAE*) is $F(t, x(t), \dot{x}(t)) = 0$ [KunM06, p. 7 f.] with a partial derivative $\frac{\partial F}{\partial \dot{x}}$ which possibly loses rank along a solution. Note that the definition includes both under- and overdetermined systems. Typical examples for underdetermined DAEs are control problems; see e.g. [PolW98]. Therein, state feedback or output control can be used to get square systems, i.e., systems with the same number of equations and variables, without changing some internal properties. For more details see [KunM06, Sec. 4.4]. On the other hand, overdetermined DAEs with non-contradicting equations contain redundancies under some technical assumptions [KunM06, p. 207 ff.]. These redundant equations can simply be removed without altering the solution set. Thus, we assume that the systems are square. For the analysis of these DAEs one usually linearizes around a trajectory; see e.g. [KunM06, Ch. 4]. This leads to DAEs of the form

$$E(t)\dot{x}(t) + A(t)x(t) = f(t). \tag{2.1}$$

Such linear DAEs are studied in this chapter.

We consider the DAE (2.1) on the compact interval $[0, T]$, $T > 0$. The state is given by $x\colon [0, T] \to \mathbb{R}^{n_x}$. We assume that the matrix-valued functions $E, A\colon [0, T] \to \mathbb{R}^{n_x \times n_x}$ as well as the right-hand side $f\colon [0, T] \to \mathbb{R}^{n_x}$ are sufficiently regular. Note that, if E is differentiable, we can rewrite (2.1) with a leading $\frac{\mathrm{d}}{\mathrm{d}t}(Ex)$. However, in contrast to ordinary differential equations (ODEs) the matrix-valued function E is in general not pointwise invertible. This forces the state to stay on a time-dependent manifold [Rhe84]. In particular, for an initial condition

$$x(0) = x_0, \tag{2.2}$$

the initial value $x_0 \in \mathbb{R}^{n_x}$ is restricted to be on this manifold at $t = 0$.

We call x a *solution of* (2.1) *with initial condition* (2.2) if $x \in C^1(0, T; \mathbb{R}^{n_x})$, the DAE (2.1) is pointwise satisfied, and (2.2) is fulfilled as well. An initial value x_0 is called *consistent* with (2.1), if the associated initial value problem has at least one solution.

A special class of linear DAEs, which often appears in this thesis are linear semi-explicit DAEs of the form

$$M(t)\dot{x}(t) + A(t)x(t) - B^T(t)\lambda(t) = f(t), \tag{2.3a}$$

$$B(t)x(t) = g(t). \tag{2.3b}$$

The desired solution is $(x, \lambda)\colon [0, T] \to \mathbb{R}^{n_x} \times \mathbb{R}^{n_\lambda}$, $n_\lambda \le n_x$, where we refer to λ as *Lagrange multiplier* and to x as state. The right-hand sides f and g map into $\mathbb{R}^{n_x}$ and $\mathbb{R}^{n_\lambda}$, respectively, and the matrix-valued functions are well-sized.

Lemma 2.1 (Cf. [Zim15, Sec. 2.5.1]). *Let $A, M \in C^1([0, T], \mathbb{R}^{n_x \times n_x})$, and $B \in C^2([0, T], \mathbb{R}^{n_\lambda \times n_x})$. Suppose that M is pointwise invertible and $BM^{-1}B^T$ as well. Assume that $f \in C^1([0, T], \mathbb{R}^{n_x})$ and $g \in C^2([0, T], \mathbb{R}^{n_\lambda})$. Then the DAE (2.3) has a unique solution for every consistent initial value x_0, λ_0, i.e., $B(0)x_0 = g(0)$ and $(BM^{-1}B^T)(0)\lambda_0 = (BM^{-1}A - \dot{B})x_0 + (\dot{g} - BM^{-1}f)(0)$. In particular, the Lagrange multiplier λ is completely determined by*

$$(BM^{-1}B^T)(t)\lambda(t) = (\dot{g} - \dot{B}x + BM^{-1}Ax - BM^{-1}f)(t). \tag{2.4}$$

Equation (2.4) is called a *hidden constraint* of the DAE (2.3) [KunM06, p. 201], since it is not apparent in (2.3). Note that the initial value of the Lagrange multiplier λ is completely determined by x_0 and the right-hand sides. Thus, we only have freedom in the choice of the initial value x_0. More precisely, only the part in the kernel of $B(0)$ is not predetermined by the constraint (2.3b).

The differentiability of all matrices and of the right-hand sides can be reduced by one in Lemma 2.1 if we expect from a solution that λ is only continuous. These assumptions can be even more weakened if we consider solutions in Lebesgue spaces [Han89; KunMS+06].

The following two subsections introduce the differentiation index and port-Hamiltonian DAEs.

2.1. The Differentiation Index

The indices of a DAE give, on one hand, an insight in the analytic properties of the DAE, e.g., the needed regularity of the data to be solvable. On the other hand, they play a fundamental role in the numerical treatment of DAEs, see Section 5.1, and can be seen as a measure of difficulty of the numerically solving [Meh15, p. 677]. We introduce for the DAE (2.1) the most common used index of a DAE, the differentiation index.

Definition 2.2 (Differentiation Index; [KunM06, Def. 3.37])**.** For the DAE (2.1) let the *inflated pair* $\mathbf{M}_\ell, \mathbf{N}_\ell \colon [0,T] \to \mathbb{R}^{(\ell+1)n_x \times (\ell+1)n_x}$ be defined via

$$\mathbf{M}_\ell := \begin{bmatrix} E & & & & \\ \dot{E}+A & E & & & \\ \ddot{E}+2\dot{A} & 2\dot{E}+A & E & & \\ \vdots & & \ddots & \ddots & \\ E^{(\ell)}+\ell A^{(\ell-1)} & \cdots & \cdots & \ell\dot{E}+A & E \end{bmatrix}, \qquad \mathbf{N}_\ell := \begin{bmatrix} A & 0 & \cdots & 0 \\ \dot{A} & 0 & \cdots & 0 \\ \ddot{A} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ A^{(\ell)} & 0 & \cdots & 0 \end{bmatrix},$$

$\ell = 0, 1, \dots$ The *differentiation index* $i_\nu \in \mathbb{N}_0$ of (2.1) is the smallest number, if it exists, for which $\mathbf{M}_{i_\nu}$ has constant rank and is *pointwise* 1*-full*, i.e., for every $t \in [0,T]$ there exists a nonsingular matrix $\mathbf{R}_{i_\nu}(t) \in \mathbb{R}^{(i_\nu+1)n_x \times (i_\nu+1)n_x}$ such that

$$\mathbf{R}_{i_\nu}(t)\mathbf{M}_{i_\nu}(t) = \begin{bmatrix} I_{n_x} & 0 \\ 0 & * \end{bmatrix}.$$

Here, I_{n_x} denotes the identity matrix in $\mathbb{R}^{n_x \times n_x}$ and $*$ a non-specified part.

Remark 2.3. Under the assumption of Lemma 2.1 on the matrix-valued functions M and B the DAE (2.3) is of differentiation index $i_\nu = 2$ [Hei14, Prop. 2.6 & Rem. 2.9]. In this thesis we also consider semi-linear DAEs of the form (2.1) where f depends on u. This, however, does not change the differentiation index under the assumptions made on M and B even if we have to consider nonlinear DAEs; cf. [KunM06, Lem. 4.8 & p. 195 f.].

The differentiation index is the minimal number of differentiations such that by algebraic manipulation the so-called *underlying ordinary differential equation (ODE)* can be extracted. The underlying ODE and the DAE (2.1) then share the same solutions for the set of consistent initial values [KunM06, p. 115]. Note that, for the extraction of the underlying ODE the functions E, A, and f have to be i_ν times differentiable. This is an upper bound on the necessary regularity in order that (2.1) is solvable. The actual differentiability that is needed, can be smaller for parts of E, A, and f; see Lemma 2.1.

Besides the differentiation index, there exist other index concepts for DAEs. For over- and underdetermined DAEs the *strangeness-index* is introduced in [KunM06, Ch. 3]. Beside control problems, underdetermined DAEs appear in the spatial discretization of constrained PDEs where the discretization should reflect that some infinite-dimensional variables are underdetermined like the pressure in the incompressible Navier-Stokes equations [Wei97]. The *perturbation index* [HaiLR89, p. 459 ff.] measures the effect of perturbations of right-hand side f on the solution of (2.1) and was extended to constrained PDEs in [AngR07; CamM99; LucSE99; RanA05]. However, it lacks a rigorous definition for DAEs in general Banach spaces. The *tractability index* [LamMT13, Part I] is connected to the structural decoupling of DAEs into so-called inherent ODEs and a triangular subsystem where the solutions must be less regular. It is based on projections and can be extended to DAEs in abstract spaces [Tis03]. Unfortunately, this index is not applicable for the abstract DAEs considered here, since the embedding used here of the Banach space into its dual is not surjective; see Section 3.1. For more index concepts we refer to [Meh15] and the references therein. For the applications considered in this thesis, all these concepts are essentially equivalent.

In the following, whenever we refer to the *index*, this should be understood as the *differentiation index*.

2.2. Port-Hamiltonian Differential-Algebraic Equations

In almost all applications, DAEs model physical systems. If these systems are network-based models where the submodels are interconnected through exchange of energy, the DAEs can be expressed as port-Hamiltonian differential-algebraic equations (pHDAEs). In this thesis, we concentrate on linear pHDAEs of the form

$$E\dot{x} = (J - R)x + Gw, \tag{2.5a}$$

$$y = G^T x. \tag{2.5b}$$

The matrix-valued functions satisfy $E \in C^1([0,T], \mathbb{R}^{n_x \times n_x})$ and $J, R \in C([0,T], \mathbb{R}^{n_x \times n_x})$ and $G \in C([0,T], \mathbb{R}^{n_x \times n_u})$. We assume that E and R are pointwise symmetric positive semidefinite and that $J + J^T = -\dot{E}$. The system's internal energy is typically described by its quadratic *Hamiltonian*

$$H\colon C^1([0,T], \mathbb{R}^{n_x}) \to C^1([0,T], \mathbb{R}), x \mapsto H(x) := \frac{1}{2} x^T E x.$$

The so-called *(external) port variables* $w, y\colon [0,T] \to \mathbb{R}^{n_u}$ describe the system interaction with the environment, in the sense that every solution of the pHDAE (2.5) satisfies

$$H(x(t)) - H(x_0) = \int_0^t -x^T R x + w^T y \,\mathrm{d}s \le \int_0^t w^T y \,\mathrm{d}s, \tag{2.6}$$

cf. [BeaMX+18, Th. 15]. Therefore, the change of the internal energy is bounded by the *supplied power* $w^T y$. The inequality (2.6) is called *dissipation inequality*. In a system theoretical language, this proves that system (2.5) is passive with the Hamiltonian H as storage function, [SchJ14, Ch. 7], as well as stable if $E(t)$ is uniformly positive definite and $w = 0$ [MehM19, p. 6864].

The type of pHDAEs described in (2.5) is a special case of the port-Hamiltonian descriptor systems in [BeaMX+18, Rem. 14]. For more general systems and more details on pHDAEs we refer to [BeaMX+18; MehM19; SchJ14; SchM18] and the references therein.

Remark 2.4. There are a lot of attempts to generalize the port-Hamiltonian (pH) framework to systems of infinite dimensions. An extension of pH systems to systems with distributed parameters based on differential forms can be found in [SchJ14, Ch. 14]. In [JacZ12] pH systems for linear dynamic PDEs on a one-dimensional spatial domain are introduced and analyzed with methods of semigroups. The authors of [MehM19] extended pHDAE to systems with non-quadratic Hamiltonians. This extension can also be considered in infinite-dimensions if the solution is smooth enough. Nonlinear, unconstrained, infinite-dimensional pH systems without inputs are introduced in [Egg19] with variational methods. In [MosZ18] the authors extend the idea of unconstrained pH systems to infinite-dimensional systems where the internal energy as well as the system's entropy are the quantities considered, and such systems are studied with variational methods.

3. Functional Analytic Tools

For the analysis of constrained PDEs, we need several functional analytic concepts. Starting with Section 3.1, we introduce basic definitions and notations of Banach spaces and their operators. This includes fundamental definitions such as right-inverses of inf-sup stable operators, Gelfand triples, as well as weak and weak* convergence. After a summary of some frequently used inequalities in Section 3.2, we introduce a weaker differentiation concept for real-valued functions in Section 3.3. These generalized derivatives are connected to the weak formulation of (constrained) PDEs. The weak formulation is the abstract framework, in which we analyze constrained PDEs. In particular, this leads to functions with images in Banach spaces. These so-called abstract functions are introduced in Section 3.4. A measure and an integral for these functions are defined. Based on this we introduce generalized derivatives for abstract functions.

3.1. Banach and Hilbert Spaces and Their Operators

In the whole section, $\mathcal{X}$ and $\mathcal{Y}$ are real Banach spaces and $\mathcal{H}$ is a real Hilbert space. In the following, the given definitions and results are taken from [Alt16] if no other reference is given.

Operators and the Dual Space A map $\mathcal{A}\colon \mathcal{X} \to \mathcal{Y}$ is called an *operator*. A linear operator is *bounded* or *continuous* if a constant $c \in \mathbb{R}_{\geq 0}$ exists such that $\|\mathcal{A}x\|_\mathcal{Y} \leq c\,\|x\|_\mathcal{X}$ for all $x \in \mathcal{X}$. The set $\mathscr{L}(\mathcal{X},\mathcal{Y})$ of all bounded linear operators $\mathcal{A}\colon \mathcal{X} \to \mathcal{Y}$ is a Banach space with respect to the norm

$$\|\mathcal{A}\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \coloneqq \sup_{x\in\mathcal{X}\setminus\{0\}} \frac{\|\mathcal{A}x\|_\mathcal{Y}}{\|x\|_\mathcal{X}}.$$

In the following, we also denote the operator norm of $\mathcal{A}$ by $C_\mathcal{A}$. We set $\mathscr{L}(\mathcal{X}) \coloneqq \mathscr{L}(\mathcal{X},\mathcal{X})$. The space $\mathcal{X}^* \coloneqq \mathscr{L}(\mathcal{X},\mathbb{R})$ is denoted as the *dual space of* $\mathcal{X}$. The bilinear map $\langle\,\cdot\,,\cdot\,\rangle_{\mathcal{X}^*,\mathcal{X}}\colon \mathcal{X}^* \times \mathcal{X} \to \mathbb{R}$, $(f,x) \mapsto f(x)$ is called a *duality pairing*. In the following, we omit the subscripted specification if the spaces are clear from the context. For a subset $\mathcal{X}_1 \subset \mathcal{X}$, the *annihilator of* $\mathcal{X}_1$ is defined as the subspace

$$\mathcal{X}_1^0 \coloneqq \big\{ f \in \mathcal{X}^* \mid \langle f, x\rangle = 0 \text{ for all } x \in \mathcal{X}_1 \big\} \subset \mathcal{X}^*.$$

Lemma 3.1 (Continuous Linear Extension; see e.g. [Alt16, p. 160 f.])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. Suppose that $\mathcal{Z} \subset \mathcal{X}$ is a dense subspace of $\mathcal{X}$. Let $\mathcal{A}\colon \mathcal{Z} \to \mathcal{Y}$ be linear and bounded, where $\mathcal{Z}$ is equipped with $\|\cdot\|_\mathcal{X}$. Then there exist a unique* extension $\overline{\mathcal{A}} \in \mathscr{L}(\mathcal{X},\mathcal{Y})$ *of $\mathcal{A}$ with $\overline{\mathcal{A}}\big|_\mathcal{Z} = \mathcal{A}$.*

Projections and Embeddings An operator $\mathcal{P} \in \mathscr{L}(\mathcal{X})$ is called a *projection onto* $\mathcal{X}_1 \subset \mathcal{X}$ if $\mathcal{P}^2 = \mathcal{P}$ and $\operatorname{im}\mathcal{P} = \mathcal{X}_1$. According to the *closed complement theorem*, see e.g. [Alt16, Th. 9.15], a closed subspace $\mathcal{X}_1$ of $\mathcal{X}$ has a closed complement $\mathcal{X}_2$ if and only if there exists a projection $\mathcal{P} \in \mathscr{L}(\mathcal{X})$ onto $\mathcal{X}_1$ with $\ker\mathcal{P} = \operatorname{im}(\mathrm{id}_\mathcal{X} - \mathcal{P}) = \mathcal{X}_2$. For a Hilbert space $\mathcal{H}$, a projection $\mathcal{P} \in \mathscr{L}(\mathcal{H})$ is called *orthogonal* if $\ker\mathcal{P} = (\operatorname{im}\mathcal{P})^\perp$. Here, $S^\perp \coloneqq \{h \in \mathcal{H} \,|\, (h,s)_\mathcal{H} = 0 \text{ for all } s \in S\} \subset \mathcal{H}$ is the *orthogonal complement* of the set $S \subset \mathcal{H}$. The property $\ker\mathcal{P} = (\operatorname{im}\mathcal{P})^\perp$ is equivalent to $(\mathcal{P}h_1, h_2)_\mathcal{H} = (h_1, \mathcal{P}h_2)_\mathcal{H}$ for all $h_1, h_2 \in \mathcal{H}$.

We say $\mathcal{X}$ is *continuously embedded* in $\mathcal{Y}$, if an injective mapping $\iota \in \mathscr{L}(\mathcal{X},\mathcal{Y})$ exists. We write $\mathcal{X} \hookrightarrow \mathcal{Y}$ and denote $C_{\mathcal{X}\hookrightarrow\mathcal{Y}} \coloneqq C_\iota$. An element $x \in \mathcal{X}$ is identified as an element in $\mathcal{Y}$ via ιx. We omit the embedding ι in ιx if it is clear from the context. The embedding is *dense*, denoted as $\mathcal{X} \overset{\mathrm{d}}{\hookrightarrow} \mathcal{Y}$ if $\operatorname{im}\iota$ is dense in $\mathcal{Y}$. If an isometric isomorphism from $\mathcal{X}$ to $\mathcal{Y}$ exists, we write $\mathcal{X} \cong \mathcal{Y}$.

Lemma 3.2 ([BerL76, Lem. 2.3.1 & Th. 2.7.1])**.** *Let the Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ be contained in a bigger linear Hausdorff space. Then*

$$\mathcal{X} \cap \mathcal{Y} \quad \text{and} \quad \mathcal{X} + \mathcal{Y} := \{a = x + y \,|\, x \in \mathcal{X}, y \in \mathcal{Y}\}$$

are complete with respect to the norms

$$\|a\|_{\mathcal{X}\cap\mathcal{Y}} := \max(\|a\|_{\mathcal{X}}, \|a\|_{\mathcal{Y}}) \quad \text{and} \quad \|a\|_{\mathcal{X}+\mathcal{Y}} := \inf\{\|x\|_{\mathcal{X}} + \|y\|_{\mathcal{Y}} \,|\, a = x + y, x \in \mathcal{X}, y \in \mathcal{Y}\},$$

respectively. If $\mathcal{X} \cap \mathcal{Y}$ is dense in $\mathcal{X}$ and $\mathcal{Y}$, then we have $(\mathcal{X} \cap \mathcal{Y})^* \cong \mathcal{X}^* + \mathcal{Y}^*$ and $(\mathcal{X} + \mathcal{Y})^* \cong \mathcal{X}^* \cap \mathcal{Y}^*$.

Reflexivity, Riesz Isomorphism, and Adjoint Operator We call $\mathcal{X}$ *reflexive* if the embedding $\mathcal{X} \hookrightarrow (\mathcal{X}^*)^*$ given by $x \mapsto \langle \cdot, x \rangle_{\mathcal{X}^*,\mathcal{X}}$ is surjective. In particular, we obtain $\mathcal{X} \cong (\mathcal{X}^*)^*$. By the following Riesz Representation Theorem 3.3 the identification $\mathcal{H} \cong \mathcal{H}^*$ holds. Especially, every Hilbert space is reflexive and $\mathcal{H}^*$ is a Hilbert space with inner product $(f_1, f_2)_{\mathcal{H}^*} := (\mathcal{R}_{\mathcal{H}}^{-1} f_1, \mathcal{R}_{\mathcal{H}}^{-1} f_2)_{\mathcal{H}}$.

Theorem 3.3 (Riesz Representation Theorem & Riesz Isomorphism; see e.g. [Alt16, Th. 6.1])**.** *Let $\mathcal{H}$ be a Hilbert space. Then the linear, bounded operator $\mathcal{R}_{\mathcal{H}} \colon \mathcal{H} \to \mathcal{H}^*, h \mapsto (h, \cdot)_{\mathcal{H}}$ is an isometric isomorphism.*

For every operator $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ there exists a unique *adjoint operator* $\mathcal{A}^* \in \mathscr{L}(\mathcal{Y}^*, \mathcal{X}^*)$ such that $\langle f, \mathcal{A}x \rangle_{\mathcal{Y}^*,\mathcal{Y}} = \langle \mathcal{A}^* f, x \rangle_{\mathcal{X}^*,\mathcal{X}}$ for all $x \in \mathcal{X}$, $f \in \mathcal{Y}^*$. The map $\mathcal{A} \mapsto \mathcal{A}^*$ defines an isometric embedding from $\mathscr{L}(\mathcal{X}, \mathcal{Y})$ to $\mathscr{L}(\mathcal{Y}^*, \mathcal{X}^*)$. If $\mathcal{Y}$ is reflexive the adjoint operator $\mathcal{A}^*$ of $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{Y}^*)$ can be identified with an element of $\mathscr{L}(\mathcal{Y}, \mathcal{X}^*)$. An operator $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ with reflexive $\mathcal{X}$ is *self-adjoint* if $\mathcal{A}^* = \mathcal{A}$ and *skew-adjoint* if $\mathcal{A}^* = -\mathcal{A}$. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces, then every $\mathcal{A} \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ has a unique *Hilbert-adjoint* $\mathcal{A}^{\mathrm{H}} \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_1)$ such that $(\mathcal{A}h_1, h_2)_{\mathcal{H}_2} = (\mathcal{A}^{\mathrm{H}} h_2, h_1)_{\mathcal{H}_1}$ for all $h_i \in \mathcal{H}_i$, $i = 1, 2$. The Hilbert adjoint is given by $\mathcal{A}^{\mathrm{H}} = \mathcal{R}_{\mathcal{H}_1}^{-1} \mathcal{A}^* \mathcal{R}_{\mathcal{H}_2}$.

Elliptic and Inf-Sup Stable Operators An operator $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ is *elliptic*, if there exists a constant $\mu_{\mathcal{A}} \in \mathbb{R}_{>0}$ with

$$\langle \mathcal{A}x, x \rangle_{\mathcal{X}^*,\mathcal{X}} \geq \mu_{\mathcal{A}} \|x\|_{\mathcal{X}}^2.$$

We call $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ elliptic on a subspace $\mathcal{X}_1$ of $\mathcal{X}$ if its restriction $\mathcal{A}\big|_{\mathcal{X}_1} \colon \mathcal{X}_1 \to \mathcal{X}^* \subset \mathcal{X}_1^*$ is elliptic. If $\mathcal{X}$ is reflexive, then $(\cdot, \cdot)_{\mathcal{X}} := \frac{1}{2}\langle (\mathcal{A} + \mathcal{A}^*) \cdot, \cdot \rangle$ defines an inner product for every elliptic operator $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ and $\mathcal{X}$ is a Hilbert space with this inner product. The induced norm is denoted by $\| \cdot \|_{\mathcal{A}}$.

Theorem 3.4 (Lax-Milgram Theorem; see e.g. [Alt16, Th. 6.2])**.** *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{A} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ be elliptic. Then $\mathcal{A}$ has an inverse $\mathcal{A}^{-1} \in \mathscr{L}(\mathcal{H}^*, \mathcal{H})$ with operator norm $C_{\mathcal{A}^{-1}} \leq \mu_{\mathcal{A}}^{-1}$.*

Lemma 3.5. *Let $\mathcal{H}_1$ be a closed subspace of the Hilbert space $\mathcal{H}$. Suppose that $\mathcal{A} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ is elliptic on $\mathcal{H}_1$ and define*

$$\mathcal{H}_2 := \{h \in \mathcal{H} \,|\, \mathcal{A}h \in \mathcal{H}_1^0\}. \tag{3.1}$$

Then $\mathcal{H}_2$ is a closed subspace of $\mathcal{H}$ and we have the direct sum $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$.

Proof. By the linearity and continuity of $\mathcal{A}$ it follows that $\mathcal{H}_2$ is a closed subspace of $\mathcal{H}$. The definition of $\mathcal{H}_2$ and the ellipticity of $\mathcal{A}$ implies $0 = \mu_{\mathcal{A}}^{-1}\langle \mathcal{A}h, h \rangle \geq \|h\|_{\mathcal{H}}$ for every $h \in \mathcal{H}_1 \cap \mathcal{H}_2$ and therefore $\mathcal{H}_1 \cap \mathcal{H}_2 = \{0\}$. It remains to show $\mathcal{H} \subseteq \mathcal{H}_1 \oplus \mathcal{H}_2$. Let $h \in \mathcal{H}$ be given. By the Lax-Milgram Theorem 3.4 there exists a $h_1 \in \mathcal{H}_1$ with $\mathcal{A}h_1 = \mathcal{A}h$ in $\mathcal{H}_1^*$. We define $h_2 := h - h_1$ and observe $\mathcal{A}h_2 = \mathcal{A}h - \mathcal{A}h_1 \in \mathcal{H}_1^0$. Thus, $h_2 \in \mathcal{H}_2$ and $\mathcal{H} \subseteq \mathcal{H}_1 \oplus \mathcal{H}_2$. □

A generalization of ellipticity is inf-sup stability. An operator $\mathcal{B} \in \mathscr{L}(\mathcal{X}, \mathcal{Y}^*)$ fulfills such an *inf-sup* or *Ladyzhenskaya–Babuška–Brezzi (LBB) condition* if a constant $\beta \in \mathbb{R}_{>0}$ exists such that

$$\inf_{y \in \mathcal{Y}\setminus\{0\}} \sup_{x \in \mathcal{X}\setminus\{0\}} \frac{\langle \mathcal{B}x, y\rangle}{\|x\|_{\mathcal{X}}\|y\|_{\mathcal{Y}}} \geq \beta. \tag{3.2}$$

In this thesis we consider inf-sup stable operators from a Hilbert space into another. In the following we give implications of this assumptions.

Lemma 3.6 ([Bra07, Lem. III.4.2] & [Zim15, Rem. 3.5])**.** *Let $\mathcal{V}$ and $\mathcal{Q}$ be real Hilbert spaces. Assume that $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable. Suppose that $\mathcal{V}^0_{\mathrm{ker}}$ is the annihilator of $\mathcal{V}_{\mathrm{ker}} \coloneqq \ker \mathcal{B}$ and $\mathcal{V}_{\mathrm{c}}$ is a closed subspace of $\mathcal{V}$ and a complement of $\mathcal{V}_{\mathrm{ker}}$.*

Then the restricted operators $\mathcal{B}\colon \mathcal{V}_{\mathrm{c}} \to \mathcal{Q}^*$ and $\mathcal{B}^*\colon \mathcal{Q} \to \mathcal{V}^0_{\mathrm{ker}}$ are isomorphisms with bounded inverses. In particular, $\|\mathcal{B}^* q\|_{\mathcal{V}^*} \geq \beta\|q\|_{\mathcal{Q}}$ for all $q \in \mathcal{Q}$.

Definition 3.7 (Right Inverse of $\mathcal{B}$ and Left Inverse of $\mathcal{B}^*$)**.** Let the assumptions of Lemma 3.6 be satisfied. We call the inverse of $\mathcal{B} \in \mathscr{L}(\mathcal{V}_{\mathrm{c}}, \mathcal{Q}^*)$ from Lemma 3.6 a *right inverse of* $\mathcal{B}$ and denote it by $\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}} \in \mathscr{L}(\mathcal{Q}^*, \mathcal{V}_{\mathrm{c}}) \subset \mathscr{L}(\mathcal{Q}^*, \mathcal{V})$. If $\mathcal{V}_{\mathrm{c}}$ is defined as in equation (3.1) with respect to $\mathcal{V}^0_{\mathrm{ker}}$ and an operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$, which is elliptic on $\mathcal{V}_{\mathrm{ker}}$, we write $\mathcal{B}^-_{\mathcal{A}} \coloneqq \mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}$. If $\mathcal{V}_{\mathrm{c}} = \mathcal{V}^{\perp}_{\mathrm{ker}}$ we use $\mathcal{B}^-_{\perp}$ as well. Further, the *left inverse* $\mathcal{B}^{-*}_{\mathit{left}}$ *of* $\mathcal{B}^*$ is the inverse of $\mathcal{B}^* \in \mathscr{L}(\mathcal{Q}, \mathcal{V}^0_{\mathrm{ker}})$ as defined in Lemma 3.6.

The notation $\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}$ and $\mathcal{B}^{-*}_{\mathrm{left}}$ as right and left inverse, respectively, is well-defined since by their definitions we have $\mathcal{B}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}} = \mathrm{id}_{\mathcal{Q}^*}$ and $\mathcal{B}^{-*}_{\mathrm{left}}\mathcal{B}^* = \mathrm{id}_{\mathcal{Q}}$. In this thesis we write $\mathcal{B}^-$ and skip the index if the space $\mathcal{V}_{\mathrm{c}}$ or the operator $\mathcal{A}$ can be chosen arbitrarily. We now show how $\mathcal{B}^-_{\mathcal{A}}$ can be computed.

Theorem 3.8 (Stationary Saddle Point Problem; [Bra07, Th. III.4.3])**.** *Suppose that the assumptions of Lemma 3.6 are satisfied and the operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is elliptic on $\mathcal{V}_{\mathrm{ker}}$. Then for every $f \in \mathcal{V}^*$ and $g \in \mathcal{Q}^*$ the system*

$$\mathcal{A}u - \mathcal{B}^*\lambda = f \qquad \text{in } \mathcal{V}^*, \tag{3.3a}$$

$$\mathcal{B}u \qquad\;\; = g \qquad \text{in } \mathcal{Q}^* \tag{3.3b}$$

has a unique solution $(u, \lambda) \in \mathcal{V} \times \mathcal{Q}$, which depends linearly and continuously on (f, g) with bounds

$$\|u\|_{\mathcal{V}} \leq \frac{1}{\mu_{\mathcal{A}}}\|f\|_{\mathcal{V}^*} + \frac{1}{\beta}\Big(1 + \frac{C_{\mathcal{A}}}{\mu_{\mathcal{A}}}\Big)\|g\|_{\mathcal{Q}^*}, \qquad \|\lambda\|_{\mathcal{Q}} \leq \frac{1}{\beta}\Big(1 + \frac{C_{\mathcal{A}}}{\mu_{\mathcal{A}}}\Big)\|f\|_{\mathcal{V}^*} + \frac{C_{\mathcal{A}}}{\beta^2}\Big(1 + \frac{C_{\mathcal{A}}}{\mu_{\mathcal{A}}}\Big)\|g\|_{\mathcal{Q}^*}.$$

Corollary 3.9 ([AltZ20, Lem. 3.1])**.** *Let the assumptions of Theorem 3.8 be satisfied. Suppose that $\mathcal{V}_{\mathrm{c}}$ is defined as*

$$\mathcal{V}_{\mathrm{c}} \coloneqq \{v \in \mathcal{V} \mid \mathcal{A}v \in \mathcal{V}^0_{\mathrm{ker}}\} \tag{3.4}$$

and $\mathcal{B}^-_{\mathcal{A}}$ be the right inverse of $\mathcal{B}$. Then, for every $g \in \mathcal{Q}^*$ the term $\mathcal{B}^-_{\mathcal{A}}g \in \mathcal{V}_{\mathrm{c}}$ is the part u of the solution (u, λ) of (3.3) with the right-hand side $(0, g)$.

Remark 3.10. If in addition to the assumptions of Corollary 3.9 the operator $\mathcal{A}$ is elliptic on whole $\mathcal{V}$, then $\mathcal{B}\mathcal{A}^{-1}\mathcal{B}^* \in \mathscr{L}(\mathcal{Q}, \mathcal{Q}^*)$ is elliptic as well, since

$$\langle \mathcal{B}\mathcal{A}^{-1}\mathcal{B}^* q, q\rangle = \langle \mathcal{B}^* q, \mathcal{A}^{-1}\mathcal{B}^* q\rangle \overset{(4.5)}{\geq} \tfrac{\mu_{\mathcal{A}}}{C^2_{\mathcal{A}}}\|\mathcal{B}^* q\|^2_{\mathcal{V}^*} \overset{\text{Lem. 3.6}}{\geq} \tfrac{\mu_{\mathcal{A}}\beta^2}{C^2_{\mathcal{A}}}\|q\|^2_{\mathcal{Q}}$$

for every $q \in \mathcal{Q}$. In particular, it follows by Corollary 3.9 that $\mathcal{B}^-_{\mathcal{A}} = \mathcal{A}^{-1}\mathcal{B}^*(\mathcal{B}\mathcal{A}^{-1}\mathcal{B}^*)^{-1}$.

Remark 3.11. By the estimates in Theorem 3.8 the operator norm of $\mathcal{B}^-_{\mathcal{A}}$ is bounded by $\frac{1}{\beta}(1 + \frac{C_{\mathcal{A}}}{\mu_{\mathcal{A}}})$. In particular, the calculation of $\mathcal{B}^-_{\mathcal{A}}g$ may be ill-conditioned if β is small and the bound is sharp.

Note that, for $\mathcal{V}_\mathrm{c} = \mathcal{V}_\mathrm{ker}^\perp$ the estimate $\|\mathcal{B}_\perp^-\|_{\mathscr{L}(\mathcal{Q}^*,\mathcal{V})} = \beta^{-1}$ holds if the inf-sup condition (3.2) is satisfied with an equal sign [Bra07, Lem. III.4.2.ii].

Separability and Gelfand Triple The space $\mathcal{X}$ is called *separable* if $\mathcal{X}$ contains a countable dense subset. Every subspace of a separable space is separable. The space $\mathcal{X}$ is separable and reflexive if and only if $\mathcal{X}^*$ is so [Bre10, Cor. 3.27]. Let a separable and reflexive Banach space $\mathcal{V}$ be densely embedded in a Hilbert space $\mathcal{H}$, then $\mathcal{H}$ is separable and $\mathcal{H}^*$ is densely embedded in $\mathcal{V}^*$, [Zei90a, p. 417]. The embedding $\mathcal{H}^* \hookrightarrow \mathcal{V}^*$ is simply given by $\langle f, v \rangle_{\mathcal{V}^*,\mathcal{V}} = \langle f, v \rangle_{\mathcal{H}^*,\mathcal{H}}$ for $f \in \mathcal{H}^*$.

Definition 3.12 (Gelfand Triple; [Zei90a, Def. 23.11])**.** We say a separable, reflexive Banach space $\mathcal{V}$ and a Hilbert space $\mathcal{H}$ form a *Gelfand triple* or *evolution triple* if

$$\mathcal{V} \overset{\mathrm{d}}{\hookrightarrow} \mathcal{H} \cong \mathcal{H}^* \overset{\mathrm{d}}{\hookrightarrow} \mathcal{V}^*. \tag{3.5}$$

We use the shorthand $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ for the Gelfand triple (3.5). The Hilbert space $\mathcal{H}$ is called *pivot space*.

The Gelfand triple $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ defines a continuous embedding of $\mathcal{V}$ in $\mathcal{V}^*$ with $v \mapsto \langle v, \cdot \rangle_{\mathcal{V}^*,\mathcal{V}} = (v, \cdot)_\mathcal{H}$. Note that, if $\mathcal{V}$ is a Hilbert space this is a second embedding of $\mathcal{V}$ in $\mathcal{V}^*$ besides the Riesz isomorphism. Then both embeddings do not coincide in general [Bre10, p. 137]. We also have to pay attention to the identification of $\mathcal{H}^*$ with $\mathcal{H}$, which can lead to inaccuracy, in particular in the context of differential equations; see [Sim10] and [Tar06, p. 121]. Therefore, we distinguish between these two spaces if it is necessary.

Given a Gelfand triple $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$, the operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ satisfies a *Gårding inequality* if two constants $\mu_\mathcal{A} \in \mathbb{R}_{>0}$ and $\kappa_\mathcal{A} \in \mathbb{R}$ exist such that

$$\langle \mathcal{A}v, v \rangle_{\mathcal{V}^*,\mathcal{V}} \geq \mu_\mathcal{A} \|v\|_\mathcal{V}^2 - \kappa_\mathcal{A} \|v\|_\mathcal{H}^2 \tag{3.6}$$

for all $v \in \mathcal{V}$. In particular, the operator $\mathcal{A} + \kappa_\mathcal{A} \,\mathrm{id}_\mathcal{H} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is elliptic and thus $\mathcal{V}$ a Hilbert space.

Strong, Weak, and Weak* Convergence A sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ *converges (strongly)* in $\mathcal{X}$ if it converges to $x \in \mathcal{X}$ with respect to the norm in $\mathcal{X}$. We write $x_n \to x$ as $n \to \infty$. The sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ is called *weakly convergent* in $\mathcal{X}$ if a *weak limit* $x \in \mathcal{X}$ exists such that $\lim_{n\to\infty} \langle f, x_n \rangle_{\mathcal{X}^*,\mathcal{X}} = \langle f, x \rangle_{\mathcal{X}^*,\mathcal{X}}$ for every $f \in \mathcal{X}^*$. We write $x_n \rightharpoonup x$ as $n \to \infty$. We say $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{X}^*$ *converges weakly** in $\mathcal{X}^*$ to the *weak* limit* $f \in \mathcal{X}^*$ and write $f_n \overset{*}{\rightharpoonup} f$ as $n \to \infty$ if $\lim_{n\to\infty} \langle f_n, x \rangle_{\mathcal{X}^*,\mathcal{X}} = \langle f, x \rangle_{\mathcal{X}^*,\mathcal{X}}$ for all $x \in \mathcal{X}$.

The weak (weak*) limit is unique, every weakly (weakly*) convergent sequence is bounded, and strong convergence implies weak (weak*) convergence. If $\mathcal{X}$ is separable, then every bounded sequence in $\mathcal{X}^*$ has a weak* convergence subsequence. If $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ is bounded and $\mathcal{X}$ is reflexive, then the sequence has a weakly convergent subsequence [Bre10, Th. 3.18]. Let $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ be weakly and $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{X}^*$ strongly convergent (or respectively, strongly and weakly* convergent) with limits $x \in \mathcal{X}$ and $f \in \mathcal{X}^*$, respectively. Then $\lim_{n\to\infty} \langle f_n, x_n \rangle_{\mathcal{X}^*,\mathcal{X}} = \langle f, x \rangle_{\mathcal{X}^*,\mathcal{X}}$. For every $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ the weak convergence $x_n \rightharpoonup x$ in $\mathcal{X}$ implies $\mathcal{A}x_n \rightharpoonup \mathcal{A}x$ in $\mathcal{Y}$ as $n \to \infty$ [Zei90a, Prop. 21.28]. In finite dimension strong, weak, and weak* convergence coincide.

3.2. Frequently Used Inequalities

In this section we summarize inequalities, which we frequently use to bound solutions of operator equations as well as to estimate the errors of their approximations. We introduce therefore the

notation
$$[1,\infty] := [1,\infty) \cup \{\infty\}.$$
Furthermore, $q \in [1,\infty]$ is called the *conjugated index* of $p \in [1,\infty]$ or we say p, q are *conjugated indices*, if $\frac{1}{p} + \frac{1}{q} = 1$, where we define
$$\frac{1}{\infty} := 0.$$

Lemma 3.13 (Young's Inequality; see e.g. [Emm04, Th. A.1.1 & A.1.4])**.** *Let $a, b \geq 0$ and $p, q \in (1,\infty)$ be conjugated indices. Then the inequality*
$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \tag{3.7}$$
holds. Furthermore, for $p = q = 2$ and every $\varepsilon > 0$ we have the estimate
$$ab \leq \frac{a^2}{2\varepsilon} + \frac{\varepsilon b^2}{2}. \tag{3.8}$$

Lemma 3.14 (Hölder's Inequality; see e.g. [Emm04, Th. A.1.7])**.** *Let $u \in L^p(a,b)$ and $v \in L^q(a,b)$ with $p, q \in [1,\infty]$ be conjugated indices. Then we have $uv \in L^1(a,b)$ with the estimate $\|uv\|_{L^1(a,b)} \leq \|u\|_{L^p(a,b)}\|v\|_{L^q(a,b)}$.*

The next inequality estimates a function if its current value is bounded by the integral over the previous values.

Lemma 3.15 (Gronwall's Lemma; see e.g. [Pac98, p.13 f.])**.** *Suppose that $u, f, g, h\colon [a,b] \to \mathbb{R}$ are Lebesgue-measurable and nonnegative as well as that hu, hf, and hg are Lebesgue-integrable. Let the inequality*
$$u(t) \leq f(t) + g(t)\int_a^t h(s)u(s)\,\mathrm{d}s$$
be satisfied for almost all $t \in [a,b]$. Then the estimate
$$u(t) \leq f(t) + g(t)\int_a^t f(s)h(s)\exp\Big(\int_s^t h(\eta)g(\eta)\,\mathrm{d}\eta\Big)\,\mathrm{d}s \tag{3.9}$$
holds almost everywhere.

Remark 3.16 (Gronwall's Lemma with Nondecreasing f)*.* If in addition to Lemma 3.15 the function f is continuous and nondecreasing, and $g(t) \equiv 1$, then the inequality (3.9) implies, cf. [Emm04, Lem. 7.3.1], the estimate
$$u(t) \leq f(t)\exp\Big(\int_a^t h(s)\,\mathrm{d}s\Big). \tag{3.10}$$

In this thesis, we sometimes estimate two functions $f_1, f_2\colon \mathcal{X} \to \mathbb{R}_{\geq 0}$ with respect to each other where scalar factors are not of interest. We therefore write $f_1(x) \lesssim f_2(x)$ as shorthand for the inequality $f_1(x) \leq C f_2(x)$ with a constant $C \in \mathbb{R}_{>0}$ independent of $x \in \mathcal{X}$.

3.3. Sobolev Spaces

The modern theory of partial differential equations is based on the concepts of generalized derivatives; see e.g. [Wlo87]. In the introduction of these derivatives, we use notions as measurable and integrable in the sense of Lebesgue from measure theory.

In the whole section $\Omega \subset \mathbb{R}^d$ is a *domain*, i.e., an open, simply connected, and bounded set. We say that Ω is a *Lipschitz domain* or has a *Lipschitz boundary* $\partial\Omega$ [Alt16, Sec. A8.2], if for every $\xi \in \partial\Omega$ there exists a neighborhood $U(\xi) \subset \mathbb{R}^d$, such that $\partial\Omega \cap U(\xi)$ is a graph of a Lipschitz continuous function with $\Omega \cap U(\xi)$ lying on one side of this graph. The closure of Ω is denoted by $\overline{\Omega}$.

Let $C_c^\infty(\Omega)$ be the set of infinitely often differentiable functions with compact support. We call a sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subset C_c^\infty(\Omega)$ *convergent in* $C_c^\infty(\Omega)$ to $\varphi \in C_c^\infty(\Omega)$ if the support of every φ_n, $n = 1, 2, \ldots$, is contained in a compact subset of Ω and $\varphi_n^{(k)}$ converges uniformly to $\varphi^{(k)}$ for every $k \in \mathbb{N}_0$. This definition allows us to introduce distributions.

Definition 3.17 (Distribution; [Rou13, p. 10])**.** A linear function $\Psi\colon C_c^\infty(\Omega) \to \mathbb{R}$ is called a *distribution* if for every convergent sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subset C_c^\infty(\Omega)$ in $C_c^\infty(\Omega)$ with limit φ it follows $\lim_{n\to\infty} \Psi(\varphi_n) = \Psi(\varphi)$.

Every local integrable function $u \in L^1_{\text{loc}}(\Omega)$ defines a distribution by $\Psi_u(\varphi) = \int_\Omega u(\xi)\varphi(\xi)\,\mathrm{d}\xi$ [Alt16, Ex. 5.18(2)]. The mapping $u \mapsto \Psi_u$ is linear, one-to-one, and u is reconstructable from Ψ_u [Alt16, Lem. 5.16]. A distribution Ψ is called *regular* if $u \in L^1_{\text{loc}}(\Omega)$ exists such that $\Psi = \Psi_u$. The set of regular distributions is a proper subspace [AdaF03, Rem. 1.59]. However, for a regular distribution we identify Ψ_u with u. Since the distributions are based on smooth functions we can define derivatives of arbitrary order for a function $u \in L^1_{\text{loc}}(\Omega)$ in the distributional sense.

Definition 3.18 (Distributional and Generalized Derivative; [Rou13, p. 15])**.** Let Ω be a domain and $u \in L^1_{\text{loc}}(\Omega)$ be given. We define its *distributional derivative* $D^\alpha u$ with multi-index $\alpha \in \mathbb{N}_0^d$ and $|\alpha| := \sum_{k=1}^d \alpha_k$ as the distribution

$$D^\alpha u(\varphi) = (-1)^{|\alpha|} \int_\Omega u(\xi) D^\alpha \varphi(\xi)\,\mathrm{d}\xi.$$

Here, D^α is short for $\frac{\partial^{|\alpha|}}{\partial_{\xi_1}^{\alpha_1} \ldots \partial_{\xi_d}^{\alpha_d}}$. If $D^\alpha u$ is regular, we call $D^\alpha u \in L^1_{\text{loc}}(\Omega)$ (after the reconstruction) a *generalized derivative of* u of type D^α.

Remark 3.19. For $d = 1$ we write $\partial_{\underbrace{\xi \ldots \xi}_{k \text{ times}}}$ instead of $D^{(k)}$, $k \in \mathbb{N}$.

The generalized derivative is unique in $L^1_{\text{loc}}(\Omega)$ and if $u \in C^m(\Omega)$ then every classical partial derivative coincides with the associated generalized derivative [Zei90a, Prop. 21.3]. For e_i as the ith canonical unit vector in $\mathbb{R}^d$, $i = 1, \ldots, d$, the d-tuple $\nabla u := (D^{e_1} u, \ldots, D^{e_d} u)$ is called the *gradient* of u and $\operatorname{div} u := \sum_{i=1}^d D^{e_i} u_i$ the *divergence* of $u = [u_1, \ldots, u_d]^T$. For functions with integrable generalized derivatives we introduce the following spaces.

Definition 3.20 (Sobolev Space $W^{k,p}(\Omega)$; [Rou13, p. 15])**.** Let Ω be a domain. For $k \in \mathbb{N}$ and $p \in [1, \infty]$ we define the *Sobolev space*

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega) \text{ for all } |\alpha| \le k\}.$$

Furthermore, $W^{0,p}(\Omega) := L^p(\Omega)$, $H^k(\Omega) := W^{k,2}(\Omega)$, and $W^{k,p}(\Omega; \mathbb{R}^m) := [W^{k,p}(\Omega)]^m$.

The Sobolev spaces $W^{k,p}(\Omega)$ equipped with the norm

$$\|u\|^p_{W^{k,p}(\Omega)} := \sum_{|\alpha|\le k} \|D^\alpha u\|^p_{L^p(\Omega)} \text{ for } p < \infty \quad \text{and} \quad \|u\|_{W^{k,\infty}(\Omega)} := \max_{|\alpha|\le k} \|D^\alpha u\|_{L^\infty(\Omega)} \tag{3.11}$$

are Banach spaces [Zei90a, p. 237 f.]. Especially, $H^k(\Omega)$ is a Hilbert space with the inner product

$$(u, v)_{H^k(\Omega)} = \sum_{|\alpha|\le k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}. \tag{3.12}$$

Along with its completeness and the embedding $W^{k,p}(\Omega) \hookrightarrow W^{k',p'}(\Omega)$ for $k' \leq k$, $p' \leq p$, the Sobolev space $W^{k,p}(\Omega)$ inherits further properties from the associated Lebesgue space.

Theorem 3.21 (Separability and Reflexivity; [Alt16, Ex. 4.18(6) & 8.11(3)])**.** *The space $W^{k,p}(\Omega)$ is separable if $p \in [1,\infty)$ and reflexive if $p \in (1,\infty)$.*

Theorem 3.22 ([AdaF03, Th. 3.17, 3.22, & p. 83 f.])**.** *Let $k \in \mathbb{N}_0$ and $p \in [1,\infty)$. Then the space $W^{k,p}(\Omega) \cap C^\infty(\Omega)$ is dense in $W^{k,p}(\Omega)$. In particular, $C^\infty(\Omega)$ can be replaced by $C^k(\Omega)$. If $\Omega \subset \mathbb{R}^d$ is, in addition, a Lipschitz domain, then $C^k(\overline{\Omega})$ and $C^\infty(\overline{\Omega})$ are dense subspaces of $W^{k,p}(\Omega)$.*

The conclusions of Theorem 3.22 do not hold for $p = \infty$ [AdaF03, Ex. 3.18]. However, for $p < \infty$ Theorem 3.22 implies an alternative definition of $W^{k,p}(\Omega)$ as the closure of $W^{k,p}(\Omega) \cap C^\infty(\Omega)$ with respect to the norms in (3.11). The closure of $C_c^\infty(\Omega)$ is of special interest. In the following we use the notation $\operatorname{clos}_{\|\cdot\|}$ as closure with respect to the norm $\|\cdot\|$.

Definition 3.23 (Sobolev Space $W_0^{k,p}(\Omega)$; [Rou13, p. 18])**.** Let Ω be a domain, $k \in \mathbb{N}_0$, and $p \in [1,\infty)$. We define $W_0^{k,p}(\Omega) := \operatorname{clos}_{\|\cdot\|_{W^{k,p}(\Omega)}} C_c^\infty(\Omega)$ and $H_0^k(\Omega) := W_0^{k,2}(\Omega)$.

The Sobolev space $W_0^{k,p}(\Omega)$ is a proper closed subspace of $W^{k,p}(\Omega)$ if $k > 0$ and Ω has positive measure; see Lemma 3.25. The non-negative function

$$\|u\|^p_{W_0^{k,p}(\Omega)} = \sum_{|\alpha|=k} \|D^\alpha u\|^p_{L^p(\Omega)} \tag{3.13}$$

defines an equivalent norm to (3.11) in $W_0^{k,p}(\Omega)$ [GajGZ74, p. 31]. In particular, summing only over all multi-indices with $|\alpha| = k$ in (3.12) defines an inner product in $H_0^k(\Omega)$. The dual space of $W_0^{k,p}(\Omega)$ is denoted by $W^{-k,q}(\Omega)$ with p, q being conjugated indices [GajGZ74, Ch. II, Def. 1.20]. Analogously, one defines $H^{-k}(\Omega) := [H_0^k(\Omega)]^*$. The Sobolev spaces with negative k can be understood as spaces of special distributions [GajGZ74, Ch. II, Lem. 1.37 f.].

For a finer investigation, it is useful to consider Sobolev spaces with non-integer k, so-called *fractional Sobolev spaces*; see [KufJF77, Ch. 6 & 8] and [AdaF03, Ch. 7]. One approach is the *Sobolev-Slobodeckiĭ spaces* [KufJF77, Def. 6.8.2], which are defined for $k \in \mathbb{R}_{>0} \setminus \mathbb{N}$ and $p \in [1,\infty)$ by

$$W^{k,p}(\Omega) := \left\{ u \in W^{\lfloor k \rfloor,p}(\Omega) \;\middle|\; \sum_{|\alpha|=\lfloor k \rfloor} \int_\Omega \int_\Omega \frac{|D^\alpha u(\xi) - D^\alpha u(\eta)|^p}{\|\xi - \eta\|_{\mathbb{R}^d}^{d+p(k-\lfloor k \rfloor)}} \,\mathrm{d}\xi \,\mathrm{d}\eta < \infty \right\}.$$

Here, $\lfloor k \rfloor \in \mathbb{N}_0$ denotes the integer part of k, i.e., $k - 1 < \lfloor k \rfloor \leq k$. The Sobolev-Slobodeckiĭ spaces are Banach spaces with the norm

$$\|u\|^p_{W^{k,p}(\Omega)} = \|u\|^p_{W^{\lfloor k \rfloor,p}(\Omega)} + \sum_{|\alpha|=\lfloor k \rfloor} \int_\Omega \int_\Omega \frac{|D^\alpha u(\xi) - D^\alpha u(\eta)|^p}{\|\xi - \eta\|_{\mathbb{R}^d}^{d+p(k-\lfloor k \rfloor)}} \,\mathrm{d}\xi \,\mathrm{d}\eta; \tag{3.14}$$

see [KufJF77, Th. 6.8.4]. Analogously to the Sobolev spaces with non-negative integers, the closure of $C_c^\infty(\Omega)$ with respect to the norm (3.14) is denoted by $W_0^{k,p}(\Omega)$. We set $W^{-k,q}(\Omega) := \left[W_0^{k,p}(\Omega)\right]^*$ with $p, q \in (1,\infty)$ being conjugated indices. The Sobolev-Slobodeckiĭ spaces with $p = 2$ are also denoted by $H^k(\Omega)$ and $H_0^k(\Omega)$, respectively.

For boundary value problems in a domain Ω with a Lipschitz boundary $\partial\Omega$ it is reasonable to ask if, and in which sense, generalized differentiable functions can be restricted on the boundary $\partial\Omega$. Note that the boundary $\partial\Omega$ has measure zero in $\mathbb{R}^d$ and $W^{k,p}(\Omega)$ is not embedded in $C(\overline{\Omega})$ in general [AdaF03, Ex. 4.43]. However, for every function u from the dense subset $C^1(\overline{\Omega}) \subset W^{1,p}(\Omega)$, $p \in [1,\infty)$, see Theorem 3.22, its restriction to $\partial\Omega$ is well-defined. With Lemma 3.1 this implies the following.

Theorem 3.24 (Trace Theorem; see e.g. [KufJF77, Th. 6.8.13 & Th. 6.9.2] & [Rou13, Th. 1.23])**.** *Let $p \in [1, \infty)$ and Ω be a Lipschitz domain. Then there exists exactly one continuous linear operator $\operatorname{tr}\colon W^{1,p}(\Omega) \to W^{1-1/p,p}(\partial\Omega)$ with $\operatorname{tr} u = u\big|_{\partial\Omega}$ for all $u \in C^1(\overline{\Omega})$. The mapping tr has a right-inverse in $\mathscr{L}(W^{1-1/p,p}(\partial\Omega), W^{1,p}(\Omega))$ for $p > 1$.*

The space $W^{1-1/p,p}(\partial\Omega)$ in Theorem 3.24 should be understood as a Sobolev space on a manifold defined by local charts; see [KufJF77, Def. 6.8.6]. The operator tr is called the *trace operator*. We also write $u\big|_{\partial\Omega}$ instead of $\operatorname{tr} u$ for $u \in W^{1,p}(\Omega)$ or omit tr if its clear that u is restricted to the boundary $\partial\Omega$.

Lemma 3.25 ([Alt16, Lem. A8.10])**.** *Let Ω be a Lipschitz domain and $p \in [1, \infty)$. Then the kernel of the trace operator with domain $W^{1,p}(\Omega)$ is $W^{1,p}_0(\Omega)$.*

For a relatively open subset Γ of $\partial\Omega$ with positive surface measure we define $\operatorname{tr}_\Gamma\colon u \mapsto (\operatorname{tr} u)\big|_\Gamma$. In this thesis we also consider the spaces

$$W^{1,p}_\Gamma(\Omega) := \ker \operatorname{tr}_\Gamma \subset W^{1,p}(\Omega).$$

Equipped with the norm (3.13), they are Banach spaces for $p \in [1, \infty)$. The space $H^1_\Gamma(\Omega) := W^{1,2}_\Gamma(\Omega)$ is a Hilbert space with the inner product defined as for $H^1_0(\Omega)$; cf. [Rou13, Th. 1.32]. If Ω and Γ are Lipschitz domains, then $\operatorname{tr}_\Gamma \in \mathscr{L}(W^{1,p}(\Omega), W^{1-1/p,p}(\Gamma))$ has a bounded right-inverse $\operatorname{tr}_\Gamma^- \in \mathscr{L}(W^{1-1/p,p}(\Gamma), W^{1,p}(\Omega))$ [Wil19, Th. 4.2.4]. In particular,

$$H^{1/2}(\Gamma) = W^{1/2,2}(\Gamma)$$

is a Hilbert space with the inner product $(w_1, w_2)_{H^{1/2}(\Gamma)} := (\operatorname{tr}_\Gamma^- w_1, \operatorname{tr}_\Gamma^- w_2)_{H^1(\Omega)}$. Its induced norm is equivalent to $\|\cdot\|_{W^{1/2,2}(\Gamma)}$; cf. [BreF91, p. 90]. Thus, $\operatorname{tr}_\Gamma \in \mathscr{L}(H^1(\Omega), H^{1/2}(\Gamma))$ is inf-sup stable. However, since $C^1(\Gamma) \hookrightarrow H^{1/2}(\Gamma)$ is dense in $L^2(\Gamma)$ [KufJF77, Rem. 6.8.3], $L^2(\Gamma)$ is densely embedded in $H^{-1/2}(\Gamma) := [H^{1/2}(\Gamma)]^*$ [Zei90a, Prop. 21.35(e)]. Especially, $H^{-1/2}(\Gamma)$ is separable and reflexive and so $H^{1/2}(\Gamma)$ is as well. Note that, we set $H^{-1/2}(\Gamma)$ as the dual space of both $H^{1/2}(\Gamma)$ and $H^{1/2}_0(\Gamma)$. This, however, is well-defined since the spaces coincide [LioM72, Ch. 1, Th. 11.1].

In general, a function in $H^{1/2}(\Gamma)$ cannot be extended by zero outside of $\Gamma \subset \partial\Omega$ to a function in $H^{1/2}(\partial\Omega)$ [BreF91, Rem. III.1.2]. The set of functions, where the extension is in $H^{1/2}(\partial\Omega)$, is denoted as $H^{1/2}_{00}(\Gamma)$. This space is not closed in $H^{1/2}(\Gamma)$ [LioM72, Ch.1, Rem. 11.4 & Th. 11.7]. However, the existence of a right-inverse tr^- of the trace operator tr implies

$$H^{1/2}_{00}(\Gamma) = \{\operatorname{tr}_\Gamma u \mid u \in H^1_{\partial\Omega\setminus\Gamma}(\Omega)\} \subset H^{1/2}(\Gamma).$$

We define the inner product

$$(w_1, w_2)_{H^{1/2}_{00}(\Gamma)} := (\operatorname{tr}^- w_1, \operatorname{tr}^- w_2)_{H^1_{\partial\Omega\setminus\Gamma}(\Omega)}, \tag{3.15}$$

where we extend w_i, $i = 1, 2$, by zero outside of $\Gamma \subset \partial\Omega$. Thus, $H^{1/2}_{00}(\Gamma)$ is a Hilbert space. Its dual space is denoted by $H^{-1/2}_{00}(\Gamma)$.

3.4. Spaces of Abstract Functions

The analysis of (constrained) PDEs leads to time-dependent equations of *abstract functions*, i.e., functions of the form

$$u\colon [a, b] \to \mathcal{X} \tag{3.16}$$

with a real Banach space $\mathcal{X}$ and $-\infty < a < b < \infty$. The space of all m-times continuously differentiable abstract functions is denoted by $C^m([a,b],\mathcal{X})$. Equipped with the norm

$$\|u\|_{C^m([a,b],\mathcal{X})} \coloneqq \sum_{k=0}^{m} \max_{t\in[a,b]} \|u^{(k)}(t)\|_{\mathcal{X}}$$

it is a Banach space, [Zei90a, Prop. 23.2(a)]. Here, $u^{(k)}$ denotes the kth derivative of u. We also write $C([a,b],\mathcal{X}) \coloneqq C^0([a,b],\mathcal{X})$ for the space of continuous abstract functions. The set of infinitely often differentiable functions with compact support is denoted by $C_c^\infty(0,T;\mathcal{X})$.

As for Sobolev spaces, one can investigate generalized derivatives for abstract functions. Before doing so we introduce the Bochner-integral.

3.4.1. Bochner Spaces

The measurability and integrability of abstract functions are strongly connected to the Lebesgue-measure and integral. As for the Lebesgue-integral, we call an abstract function $u\colon [a,b] \to \mathcal{X}$ a *simple function* if finitely many $x_k \in \mathcal{X}$ and pairwise disjoint Lebesgue-measurable sets $E_k \subset [a,b]$, $k = 1,\dots,m$, exist such that $u(t) = \sum_{k=1}^m x_k \chi_{E_k}(t)$. The map χ_{E_k} is the characteristic function of E_k. The Bochner-integral of a simple function is defined by

$$\int_a^b u(t)\,\mathrm{d}t = \int_a^b \sum_{k=1}^{m} x_k \chi_{E_k}(t)\,\mathrm{d}t \coloneqq \sum_{k=1}^{m} x_k \mu(E_k) \in \mathcal{X}$$

with the Lebesgue-measure $\mu(E_k)$ of E_k.

Definition 3.26 (Bochner-Measurable; [KufJF77, p. 107])**.** An abstract function $u\colon [a,b] \to \mathcal{X}$ is called *Bochner-measurable* if a sequence $\{u_n\}_{n\in\mathbb{N}}$ of simple functions exists such that $u_n(t) \to u(t)$ in $\mathcal{X}$ as $n \to \infty$ at almost every time-point $t \in [a,b]$.

For separable $\mathcal{X}$ it is enough for the definition of Bochner-measurable functions that the convergence in Definition 3.26 is weak [Yos80, p. 131]. If a sequence $\{u_n\}_{n\in\mathbb{N}}$ of Bochner-measurable functions converges weakly at almost every time-point to u in $\mathcal{X}$ with $\mathcal{X}$ separable, then u is Bochner-measurable [Emm04, Cor. 7.1.5]. Furthermore, $t \mapsto \|u(t)\|$ is Lebesgue-measurable if u is Bochner-measurable [Emm04, Lem. 7.1.10]. This allows the following definition of Bochner-integrable functions and the Bochner-integral.

Definition 3.27 (Bochner-Integrable and Integral; [KufJF77, Def. 2.19.6])**.** Let $u\colon [a,b] \to \mathcal{X}$ be Bochner-measurable and the sequence $\{u_n\}_{n\in\mathbb{N}}$ of simple functions converges at almost every time-point $t \in [a,b]$ to u in $\mathcal{X}$. We call u *Bochner-integrable* if for every $\varepsilon > 0$ a $N(\varepsilon) \in \mathbb{N}$ exists, such that $\int_a^b \|u_n(t) - u_m(t)\|_{\mathcal{X}}\,\mathrm{d}t < \varepsilon$ for all $n, m > N(\varepsilon)$. The *Bochner-integral* of u then is defined by

$$\int_a^b u(t)\,\mathrm{d}t \coloneqq \lim_{n\to\infty} \int_a^b u_n(t)\,\mathrm{d}t \in \mathcal{X}.$$

The convergence and the independence of the Bochner-integral on the chosen sequence is shown in [KufJF77, Rem. 2.19.7]. Examples of Bochner-integrable functions are obviously simple functions but also continuous abstract functions [GajGZ74, Ch. IV, Th. 1.9]. Furthermore, the Bochner-integral is linear by its definition. We summarize properties of Bochner-integrable functions, which we use frequently.

Theorem 3.28 (Bochner's Theorem; see e.g. [KufJF77, Th. 2.19.8])**.** *Let $u\colon [a,b] \to \mathcal{X}$ be Bochner-measurable. Then u is Bochner-integrable if and only if $[a,b] \to \|u(\cdot)\|_{\mathcal{X}}$ is Lebesgue-integrable.*

Lemma 3.29 ([KufJF77, Cor. 2.19.9]). *Let $u\colon [a,b] \to \mathcal{X}$ be Bochner-integrable. Then the estimate $\| \int_a^b u(t)\,\mathrm{d}t \|_{\mathcal{X}} \leq \int_a^b \|u(t)\|_{\mathcal{X}}\,\mathrm{d}t$ holds.*

Lemma 3.30 ([KufJF77, Cor. 2.19.11]). *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $\mathcal{A} \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$. Suppose that $u\colon [a,b] \to \mathcal{X}$ is Bochner-integrable. Then $\mathcal{A}u\colon [a,b] \to \mathcal{Y}, t \mapsto (\mathcal{A}u)(t) := \mathcal{A}u(t)$ is Bochner-integrable and for all $a \leq t_1 < t_2 \leq b$ we have*

$$\mathcal{A} \int_{t_1}^{t_2} u(t)\,\mathrm{d}t = \int_{t_1}^{t_2} \mathcal{A}u(t)\,\mathrm{d}t.$$

With Theorem 3.28 we can define Bochner-spaces for equivalence classes of Bochner-integrable functions where the equivalence relation is given by $u_1 \sim u_2$ if $\|u_1 - u_2\|_{\mathcal{X}} = 0$ in $L^1(a,b)$.

Definition 3.31 (Bochner Space $L^p(a,b;\mathcal{X})$; [Emm04, Def. 7.1.22]). For $p \in [1,\infty]$ the *Bochner space* $L^p(a,b;\mathcal{X})$ contains the equivalence classes of Bochner-integrable functions with

$$\|u\|_{L^p(a,b;\mathcal{X})} := \big\| \, \|u(\cdot)\|_{\mathcal{X}} \, \big\|_{L^p(a,b)} < \infty. \tag{3.17}$$

Furthermore, we write $u \in L^1_{\mathrm{loc}}(a,b;\mathcal{X})$ if $u \in L^1(a',b';\mathcal{X})$ for every $a < a' < b' < b$.

In this thesis we do not distinguish between u and its equivalence class. In following we summarize some results about Bochner-spaces.

Theorem 3.32 ([Zei90a, Prop. 23.2] & [Emm04, Th. 7.1.23]). *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $p \in [1,\infty]$, and $m \in \mathbb{N}_0$. Suppose that $L^p(a,b;\mathcal{X})$ is equipped with the norm (3.17). Then the following holds.*

i) *$L^p(a,b;\mathcal{X})$ is a Banach space.*

ii) *If $p \neq \infty$, then $C^m([a,b],\mathcal{X}) \overset{d}{\hookrightarrow} L^p(a,b;\mathcal{X})$ and $C^\infty([a,b],\mathcal{X})$ is dense in $L^p(a,b;\mathcal{X})$.*

iii) *$C([a,b],\mathcal{X})$ is continuously embedded in $L^\infty(a,b;\mathcal{X})$.*

iv) *Let $u \in L^p(a,b;\mathcal{X})$ and $f \in L^q(a,b;\mathcal{X}^*)$ with $q \in [1,\infty]$ as the conjugated index of p, then $t \mapsto \langle f(t), u(t) \rangle_{\mathcal{X}^*,\mathcal{X}} \in L^1(a,b)$ and the* Hölder's inequality

$$\int_a^b \langle f(t), u(t) \rangle_{\mathcal{X}^*,\mathcal{X}}\,\mathrm{d}t \leq \|f\|_{L^q(a,b;\mathcal{X}^*)} \|u\|_{L^p(a,b;\mathcal{X})}$$

holds.

v) *If $\mathcal{X}$ is reflexive or $\mathcal{X}^*$ is separable, then $L^p(a,b;\mathcal{X})$ is reflexive for $p \in (1,\infty)$ and $L^p(a,b;\mathcal{X}^*)$ is isometric isomorphic to $(L^q(a,b;\mathcal{X}))^*$ with $p \in (1,\infty]$, $q \in [1,\infty)$ conjugated indices.*

vi) *$L^p(a,b;\mathcal{X})$ is separable, if $\mathcal{X}$ is separable and $p \neq \infty$.*

vii) *If $\mathcal{X}$ is a Hilbert spaces, then $L^2(a,b;\mathcal{X})$ is a Hilbert space with inner product*

$$(u,v)_{L^2(a,b;\mathcal{X})} = \int_a^b (u(t),v(t))_{\mathcal{X}}\,\mathrm{d}t.$$

viii) *If $\mathcal{X} \hookrightarrow \mathcal{Y}$, then $L^p(a,b;\mathcal{X}) \hookrightarrow L^q(a,b;\mathcal{Y})$ for $1 \leq q \leq p \leq \infty$.*

Remark 3.33. If $\mathcal{X} \hookrightarrow \mathcal{Y}$ is dense, then the embedding in viii) is dense; cf. [Zei90a, p. 442].

An additional property of Bochner spaces, which we need in the Parts B and C, is given in the following lemma.

Lemma 3.34 (Convergence of Piecewise Mean; cf. [Tem77, Ch. III, Lem. 4.9]). *Let $u \in L^p(0,T;\mathcal{X})$ with $p \in [1,\infty)$. Define $\tau := T/N$ with $N \in \mathbb{N}$, $t_n := \tau n$ with $n = 0,\dots,N$, and $u_\tau\colon [0,T] \to \mathcal{X}$ with*

$$u_\tau\big|_{(t_{n-1},t_n]} \equiv \frac{1}{\tau} \int_{t_{n-1}}^{t_n} u(t)\,\mathrm{d}t.$$

Then $u_\tau \in L^p(0,T;\mathcal{X})$ and u_τ converges strongly to u in $L^p(0,T;\mathcal{X})$ as N tends to infinity.

3.4.2. Sobolev-Bochner Spaces

The generalized derivative for abstract functions is analogously defined to the real-valued case in Section 3.3. Therefore, we consider for the open interval (a,b) the function space $C_c^\infty(a,b) := C_c^\infty((a,b))$. We define *distributions with values in $\mathcal{X}$* [GajGZ74, Sec. IV, Def. 1.9 & Rem. 1.12] as the subset of linear functions $\Psi\colon C_c^\infty(a,b) \to \mathcal{X}$, where $\{\Psi(\varphi_n)\}_{n\in\mathbb{N}}$ has the weak limit $\Psi(\varphi)$ in $\mathcal{X}$ for every convergence sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ in $C_c^\infty(a,b)$ with limit φ. A distribution Ψ with values in $\mathcal{X}$ is called *regular* if $u \in L^1_{\mathrm{loc}}(a,b;\mathcal{X})$ exist such that $\Psi_u(\varphi) := \int_a^b u(t)\varphi(t)\,\mathrm{d}t = \Psi(\varphi)$ for all $\varphi \in C_c^\infty(a,b)$. The mapping $u \mapsto \Psi_u$ is one-to-one and u is reconstructable from Ψ_u [GajGZ74, Sec. IV Lem. 1.7].

Definition 3.35 (Derivative of Vector-Valued Functions; [Rou13, p.201])**.** For a given abstract function $u \in L^1_{\mathrm{loc}}(a,b;\mathcal{X})$ we define its kth *distributional derivative* $\frac{\mathrm{d}^k u}{\mathrm{d}t^k}$, $k \in \mathbb{N}_0$, as the distribution

$$\frac{\mathrm{d}^k u}{\mathrm{d}t^k}(\varphi) = (-1)^k \int_a^b u(t)\varphi^{(k)}(t)\,\mathrm{d}t$$

with values in $\mathcal{X}$. If $\frac{\mathrm{d}^k u}{\mathrm{d}t^k}$ is regular we call $\frac{\mathrm{d}^k}{\mathrm{d}t^k}u \in L^1_{\mathrm{loc}}(a,b;\mathcal{X})$ (after the reconstruction) the *kth generalized derivative* of u. We also write $\dot{u} := \frac{\mathrm{d}}{\mathrm{d}t}u$, $\ddot{u} := \frac{\mathrm{d}^2}{\mathrm{d}t^2}u$, and $u^{(k)} := \frac{\mathrm{d}^k u}{\mathrm{d}t^k}$.

The kth generalized derivative is unique and coincides with the kth classical derivative of u, if it exists [Zei90a, Ex. 23.16 & Prop. 23.18]. The following theorem gives alternative definitions of a generalized derivative in $L^1(a,b;\mathcal{X})$. For functions in $L^1_{\mathrm{loc}}(a,b;\mathcal{X})$ the statement should be read as for all a', b' with $a < a' < b' < b$.

Theorem 3.36 ([Emm04, Th. 8.1.5])**.** *Let $u, v \in L^1(a,b;\mathcal{X})$. Then the following are equivalent.*

i) The function v is the generalized derivative of u, i.e., $\dot{u} = v$.

ii) There exists $u_0 \in \mathcal{X}$ such that $u(t) = u_0 + \int_a^t v(s)\,\mathrm{d}s$ for almost all $t \in [a,b]$.

iii) For every $f \in \mathcal{X}^*$, $t \mapsto \langle f, u(t) \rangle$ has the generalized derivative $\frac{\mathrm{d}}{\mathrm{d}t}\langle f, u \rangle = \langle f, v \rangle$.

As in Section 3.3 we consider spaces of functions with generalized derivatives in $L^p(a,b;\mathcal{X})$.

Definition 3.37 (Sobolev-Bochner Space $W^{k,p}(a,b;\mathcal{X})$)**.** For a Banach space $\mathcal{X}$ we define the *Sobolev-Bochner space* $W^{k,p}(a,b;\mathcal{X})$, $k \in \mathbb{N}_0$, $p \in [1,\infty]$, as

$$W^{k,p}(a,b;\mathcal{X}) := \{u \in L^p(a,b;\mathcal{X}) \,|\, u^{(\ell)} \in L^p(a,b;\mathcal{X}), \ell = 1,\dots,k\}.$$

Furthermore, $H^k(a,b;\mathcal{X})$ denotes $W^{k,2}(a,b;\mathcal{X})$.

We equip the Sobolev-Bochner space $W^{k,p}(a,b;\mathcal{X})$ with the norm

$$\|u\|^p_{W^{k,p}(a,b;\mathcal{X})} = \sum_{\ell=0}^k \|u^{(\ell)}\|^p_{L^p(a,b;\mathcal{X})} \text{ for } p<\infty \quad \text{and} \quad \|u\|_{W^{k,\infty}(a,b;\mathcal{X})} = \max_{\ell=0,\dots,k} \|u^{(\ell)}\|_{L^\infty(a,b;\mathcal{X})}. \tag{3.18}$$

By Theorem 3.32 the space $W^{k,p}(a,b;\mathcal{X})$ is a Banach space. In particular, $H^k(a,b;\mathcal{X})$ is a Hilbert space, if $\mathcal{X}$ is one. The inner product is given by

$$(u,v)_{H^k(a,b;\mathcal{X})} = \sum_{\ell=0}^k \int_a^b (u^{(\ell)}(t), v^{(\ell)}(t))_{\mathcal{X}}\,\mathrm{d}t.$$

Theorem 3.38 ([Rou13, Lem. 7.1])**.** *Suppose that $\mathcal{X}$ is a Banach space and $p \in [1, \infty]$. Then $W^{1,p}(a, b; \mathcal{X}) \hookrightarrow C([a, b], \mathcal{X})$ holds with*

$$C_{W^{1,p}(a,b;\mathcal{X})\hookrightarrow C([a,b],\mathcal{X})} \leq 2^{1-1/p} \max\left((b-a)^{-1/p}, 2(b-a)^{1-1/p}\right).$$

Remark 3.39. If $b - a \geq 1/2$, then the embedding constant in Theorem 3.38 is bounded by 2, independent of p.

By [Wlo87, Th. 25.3] an embedding $\mathcal{X} \hookrightarrow \mathcal{Y}$ implies that distributions with values in $\mathcal{X}$ can be included in the distributions with values in $\mathcal{Y}$. Therefore, an abstract function with values in $\mathcal{X}$ may have a generalized derivative in $L^1_{\text{loc}}(a, b; \mathcal{Y})$. We define the Sobolev Bochner space

$$W^{1,p}(a, b; \mathcal{X}, \mathcal{Y}) := \{u \in L^p(a, b; \mathcal{X}) \mid \dot{u} \in L^p(a, b; \mathcal{Y})\}$$

with $p \in [1, \infty]$. By [Rou13, p. 201], this space is complete with respect to the norm

$$\|u\|_{W^{1,p}(a,b;\mathcal{X},\mathcal{Y})} = \begin{cases} \left(\|u\|^p_{L^p(a,b;\mathcal{X})} + \|\dot{u}\|^p_{L^p(a,b;\mathcal{Y})}\right)^{1/p}, & \text{if } p \in [1, \infty), \\ \max\left(\|u\|_{L^\infty(a,b;\mathcal{X})}, \|\dot{u}\|_{L^\infty(a,b;\mathcal{Y})}\right), & \text{if } p = \infty. \end{cases}$$

Of special interest in this thesis is the space $W^{1,2}(a, b; \mathcal{V}, \mathcal{V}^*)$, where the embedding $\mathcal{V} \hookrightarrow \mathcal{V}^*$ is given by a Gelfand triple $\mathcal{V}, \mathcal{H}, \mathcal{V}^*$. The space $W^{1,2}(a, b; \mathcal{V}, \mathcal{V}^*)$ then is a Hilbert space if $\mathcal{V}$ is one [Emm04, p. 211]. Further, $W^{1,2}(a, b; \mathcal{V}, \mathcal{V}^*)$ can be continuously embedded into the space of continuous functions with images in $\mathcal{H}$ and for every $u, v \in W^{1,2}(a, b; \mathcal{V}, \mathcal{V}^*)$ the equality

$$(u(t_2), v(t_2))_\mathcal{H} - (u(t_1), v(t_1))_\mathcal{H} = \int_{t_1}^{t_2} \langle\, \dot{u}(s), v(s)\, \rangle_{\mathcal{V}^*,\mathcal{V}} + \langle\, \dot{v}(s), u(s)\, \rangle_{\mathcal{V}^*,\mathcal{V}}\, \mathrm{d}s \tag{3.19}$$

holds for almost every $a \leq t_1 < t_2 \leq b$, [Zei90a, Prop. 23.23]. Since $W^{1,2}(a, b; \mathcal{V}, \mathcal{V}^*) \hookrightarrow L^2(a, b; \mathcal{V}) \cap C([a, b], \mathcal{H})$, we can weaken the assumptions on the derivative of u by $\dot{u} \in [L^2(a, b; \mathcal{V}) \cap C([a, b], \mathcal{H})]^* \subset L^2(a, b; \mathcal{V}^*) + L^1(a, b; \mathcal{H}^*)$, see Lemma 3.2, such that the duality pairs in (3.19) are still well-defined. So, we consider the bigger space

$$W_1(a, b; \mathcal{V}, \mathcal{V}^*) := \{u \in L^2(a, b; \mathcal{V}) \mid \dot{u} \in L^2(a, b; \mathcal{V}^*) + L^1(a, b; \mathcal{H}^*)\}. \tag{3.20}$$

Theorem 3.40 (Continuous Embedding of $W_1(a, b; \mathcal{V}, \mathcal{V}^*)$; [DauL92, p. 521] & [Tar06, p.114 f.])**.** *Let the Banach space $\mathcal{V}$ be reflexive, and the Hilbert space $\mathcal{H}$ be separable. Assume that the Gelfand-triple $\mathcal{V}, \mathcal{H}, \mathcal{V}^*$ is given. Then $W_1(a, b; \mathcal{V}, \mathcal{V}^*)$ equipped with the norm*

$$\|u\|_{W_1(a,b;\mathcal{V},\mathcal{V}^*)} := \|u\|_{L^2(a,b;\mathcal{V})} + \|\dot{u}\|_{L^2(a,b;\mathcal{V}^*)+L^1(a,b;\mathcal{H}^*)},$$

is a Banach space. This space is continuously embedded in $C([a, b], \mathcal{H})$. For every $u, v \in W_1(a, b; \mathcal{V}, \mathcal{V}^*)$ and almost every $a \leq t_1 < t_2 \leq b$ we have

$$(u(t_2), v(t_2))_\mathcal{H} - (u(t_1), v(t_1))_\mathcal{H} = \int_{t_1}^{t_2} \langle\, \dot{u}(s), v(s)\, \rangle + \langle\, \dot{v}(s), u(s)\, \rangle\, \mathrm{d}s. \tag{3.21}$$

In particular, $\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)\|^2_\mathcal{H} = 2\langle\, \dot{u}(t), u(t)\, \rangle$ for every $u \in W_1(a, b; \mathcal{V}, \mathcal{V}^*)$.

Remark 3.41. For $u, v \in W_1(a, b; \mathcal{V}, \mathcal{V}^*) \hookrightarrow L^2(a, b; \mathcal{V}) \cap C([a, b], \mathcal{H})$ the duality pairings in Theorem 3.40 read as $\langle\, \dot{u}, v\, \rangle = \langle\, \dot{u}_1, v\, \rangle_{\mathcal{V}^*,\mathcal{V}} + \langle\, \dot{u}_2, v\, \rangle_{\mathcal{H}^*,\mathcal{H}}$ with $\dot{u} = \dot{u}_1 + \dot{u}_2$, $\dot{u}_1 \in L^2(a, b; \mathcal{V}^*)$ and $\dot{u}_2 \in L^1(a, b; \mathcal{H}^*)$.

4. Dynamic Abstract Equations

In this chapter we consider dynamic equations with abstract functions as solution. This includes (constrained) PDEs in their strong and weak form.

Before we study dynamic equations in Banach spaces we investigate the properties of the abstract function $t \mapsto a(t, u(t))$ as the image of a function $u\colon [0,T] \to \mathcal{X}$ under a map $a\colon [0,T] \times \mathcal{X} \to \mathcal{Y}$ with $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. This is done in Section 4.1. In particular, we analyze measurability, integrability, and differentiability. Afterwards in Section 4.2 we consider integral equations as first example of dynamic abstract equations. Besides being of interest on their own, integral equations appear as reformulations of operator differential equations with bounded operators. Operator differential equations with bounded and unbounded operators are subject of Section 4.3. We finish this chapter with a short introduction to operator differential-algebraic equations in Section 4.4.

4.1. Nemytskiĭ Mappings

For the analysis of dynamic equations in Banach spaces we consider the map

$$a\colon [0,T] \times \mathcal{X} \to \mathcal{Y}. \tag{4.1}$$

With the dynamic equations in mind, the map a can be a right-hand side, e.g., f in the operator differential-algebraic equation (1.1), or is defined via $a(t,x) = \mathcal{A}(t)x$ with a one-parameter family of linear operators $\mathcal{A}(t)\colon \mathcal{X} \to \mathcal{Y}$. The *Nemytskiĭ mapping* $\mathcal{N}_a$ [Rou13, p. 19] is the extension of function a to abstract functions with range in $\mathcal{X}$, i.e.,

$$\mathcal{N}_a\colon ([0,T] \to \mathcal{X}) \to ([0,T] \to \mathcal{Y}); \quad u(\cdot) \mapsto a(\cdot, u(\cdot)).$$

In the following we collect conditions on a, such that its corresponding Nemystkiĭ map $\mathcal{N}_a$ maps Bochner-measurable (Bochner-integrable) to Bochner-measurable (Bochner-integrable) functions.

Definition 4.1 (Carathéodory Conditions; [GolKT92, p. 128])**.** A map $a\colon [0,T] \times \mathcal{X} \to \mathcal{Y}$ is said to satisfy the *Carathéodory conditions* or is a *Carathéodory mapping*, if

i.) $x \mapsto a(t,x)$ is a continuous function for almost all $t \in [0,T]$,

ii.) $t \mapsto a(t,x)$ is a Bochner-measurable function for all $x \in \mathcal{X}$.

Lemma 4.2 ([GolKT92, Rem. 1, Th. 1(ii)&(iv), 4, & 5])**.** *Let the map $a\colon [0,T] \times \mathcal{X} \to \mathcal{Y}$ be given. Then $\mathcal{N}_a$ maps Bochner-measurable functions $u\colon [0,T] \to \mathcal{X}$ to Bochner-measurable functions $\mathcal{N}_a u\colon [0,T] \to \mathcal{Y}$, if a satisfies the Carathéodory conditions.*

If in addition, for every $c > 0$ a function $k_a(\,\cdot\,; c) \in L^q(0,T)$, $q \in [1,\infty]$, $k_a \geq 0$, exists such that

$$\|a(t,x)\|_{\mathcal{Y}} \leq k_a(t;c) \tag{4.2}$$

is satisfied for all $x \in \mathcal{X}$ with $\|x\|_{\mathcal{X}} \leq c$ at almost every time-point $t \in [0,T]$, then $\mathcal{N}_a$ maps continuously $L^\infty(0,T;\mathcal{X})$ to $L^q(0,T;\mathcal{Y})$.

The condition (4.2) is necessary if $\mathcal{X}$ is separable, [GolKT92, Th. 3]. For sufficient and necessary conditions that $\mathcal{N}_a$ maps $L^p(0,T;\mathcal{X})$ into $L^q(0,T;\mathcal{Y})$ with arbitrary $p,q \in [1,\infty]$ we refer to [GolKT92, Ch. 2]. In the remainder of this thesis, we do not distinguish between the map a and its corresponding Nemytskiĭ mapping $\mathcal{N}_a$.

Let us now consider the special case $a(t,x) \equiv \mathcal{A}(t)x$ with a one-parameter family of bounded linear operators $\mathcal{A}(t) \in \mathscr{L}(\mathcal{X},\mathcal{Y})$. Since $\mathcal{A}(t)\colon \mathcal{X} \to \mathcal{Y}$ is linear and continuous for every t, the first Carathéodory condition is satisfied for $a(t,x)$. For the second condition we introduce different notions of measurability for operator-valued functions.

Definition 4.3 (Measurable Operator-Valued Functions; [HilP57, p. 74])**.** An operator-valued function $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y})$ is said to be

- *uniformly measurable*, if $t \mapsto \mathcal{A}(t)$ is Bochner-measurable in $\mathscr{L}(\mathcal{X},\mathcal{Y})$.
- *strongly measurable*, if $t \mapsto \mathcal{A}(t)x$ is Bochner-measurable in $\mathcal{Y}$ for every $x \in \mathcal{X}$.
- *weakly measurable*, if $t \mapsto \langle f, \mathcal{A}(t)x \rangle_{\mathcal{Y}^*,\mathcal{Y}}$ is Lebesgue-measurable for every $x \in \mathcal{X}$, $f \in \mathcal{Y}^*$.

Example 4.4.

i) The operator-valued function $\delta\colon [0,T] \to \mathscr{L}(C([0,T],\mathbb{R}),\mathbb{R})$, $t \mapsto \delta_t$ with $\delta_t[f] := f(t)$ is strongly measurable, but is not uniformly measurable [BlaN10, p. 66, Ex. 2].

ii) A weakly but not strongly measurable function is given by $[0,T] \to \mathscr{L}(\mathbb{R}, L^\infty(0,T)) \cong L^\infty(0,T)$, $t \mapsto \chi_{[0,t]}(\cdot)$ with $\chi_{[0,t]}$ as the characteristic function of $[0,t]$; see [Edg77, p. 672].

By Definition 4.1 the map $a(t,x) = \mathcal{A}(t)x$ is a Carathéodory mapping if and only if $\mathcal{A}\colon [0,T] \to \mathscr{L}(X,Y)$ is strongly measurable. In the following theorem we give a connection between the different definitions of measurability in Definition 4.3.

Theorem 4.5 ([HilP57, Th. 3.5.5])**.** *Let $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y})$. Then $\mathcal{A}$ is strongly measurable if and only if $\mathcal{A}$ is weakly measurable and $t \mapsto \mathcal{A}(t)x$ is almost separably-valued in $\mathcal{Y}$ for every $x \in \mathcal{X}$, i.e., there exists a null set $E_0 \subset [0,T]$ such that the image of $\mathcal{A}(\cdot)x$ restricted to $[0,T]\setminus E_0$ is separable in $\mathcal{Y}$. Furthermore, $\mathcal{A}$ is uniformly measurable if and only if $\mathcal{A}$ is weakly measurable and almost separably-valued in $\mathscr{L}(\mathcal{X},\mathcal{Y})$.*

Remark 4.6.

i) Since $\mathscr{L}(\mathbb{R},\mathcal{Y}) \cong \mathcal{Y}$ and $\mathscr{L}(\mathcal{X},\mathbb{R}) = \mathcal{X}^*$, the space $\mathscr{L}(\mathcal{X},\mathcal{Y})$ is separable, if $\mathcal{X}$ or $\mathcal{Y}$ is finite-dimensional and $\mathcal{Y}$ or $\mathcal{X}^*$ is separable, respectively. In particular, the different concepts of measurability in Definition 4.3 then coincide by Theorem 4.5.

ii) For infinite-dimensional $\mathcal{X}, \mathcal{Y}$ the space $\mathscr{L}(\mathcal{X},\mathcal{Y})$ is not separable, even if $\mathcal{X}$ and $\mathcal{Y}$ are separable Hilbert spaces. To verify this, let x_i, y_j be orthonormal bases of $\mathcal{X}$, $\mathcal{Y}$, respectively, and $\ell^\infty := \{\{\alpha_n\}_{n\in\mathbb{N}} \subset \mathbb{R} \,|\, \sup_{n\in\mathbb{N}} |\alpha_n| < \infty\}$ be the set of bounded sequences in $\mathbb{R}$. We consider the linear, bounded, injective map $\mathcal{G}\colon \ell^\infty \to \mathscr{L}(\mathcal{X},\mathcal{Y}); (\alpha_1,\alpha_2,\dots) \mapsto \sum_{i=1}^\infty \alpha_i (x_i,\cdot)_{\mathcal{X}} y_i$. One can show that the image of $\mathcal{G}$ is closed in $\mathscr{L}(\mathcal{X},\mathcal{Y})$ and that $\mathcal{G}$ has a bounded left inverse. Since ℓ^∞ is not separable [Alt16, Ex. 4.18(2)], the image of $\mathcal{G}$ is a non-separable subspace of $\mathscr{L}(\mathcal{X},\mathcal{Y})$. By [Alt16, Lem. 4.17(2)] this proves that $\mathscr{L}(\mathcal{X},\mathcal{Y})$ is not separable.

The weak formulation of a (constrained) PDE is sometimes written with a time-dependent bounded bilinear form $a_{\mathrm{bi}}\colon [0,T] \times \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$; see e.g. [Emm04, Sec. 8.2]. By [Zei90a, Prop. 21.31(a)] there is a pointwise bijective connection between the bilinear form a_{bi} and $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y}^*)$ with $a_{\mathrm{bi}}(t,x,y) = \langle \mathcal{A}(t)x, y \rangle_{\mathcal{Y}^*,\mathcal{Y}}$ for all t,x,y. If $\mathcal{Y}^*$ is separable, then it is enough to consider $\mathcal{A}$ by Theorem 4.5. The results of this section can then be translated to the bilinear form a_{bi}.

Let us consider a pointwise linear $\mathcal{A}$ that satisfies the uniform boundedness condition (4.2). We define $k_a := k_a(\,\cdot\,;1)$. Then the estimate

$$\|\mathcal{A}(t)x\|_{\mathcal{Y}} = \left\|\mathcal{A}(t)\frac{x}{\|x\|_{\mathcal{X}}}\right\|_{\mathcal{Y}} \|x\|_{\mathcal{X}} \overset{(4.2)}{\leq} k_a(t)\|x\|_{\mathcal{X}} \tag{4.3}$$

holds for all $x \in \mathcal{X} \setminus \{0\}$. For separable $\mathcal{X}$ an implication is given in the following lemma.

Lemma 4.7 ([BlaN10, Cor. 2.3 & Prop. 3.7])**.** *Suppose $\mathcal{X}$ is separable. Let $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y})$ be strongly measurable and satisfy the boundedness condition (4.3) for $k_a \in L^p(0,T)$, $p \in [1,\infty]$. Then*

$t \mapsto \|\mathcal{A}(t)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})}$ is an element of $L^p(0,T)$ with $\|\mathcal{A}(\cdot)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \leq k_a$ at almost every time-point $t \in [0,T]$.

As next step we consider the set of operator-valued functions, which satisfy (4.3). Since it is enough for the analysis in this thesis, we assume that $\mathcal{X}$ is separable. We follow the notation of [BlaN10].

Definition 4.8 (Space $L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$)**.** Let $\mathcal{X},\mathcal{Y}$ be Banach spaces and $\mathcal{X}$ be separable. The abstract function $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y})$ belongs to the space $L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$, $p \in [1,\infty]$, if $\mathcal{A}$ is strongly measurable, $t \mapsto \|\mathcal{A}(t)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})}$ is measurable, and

$$\|\mathcal{A}\|_{L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]} := \big\| \, \|\mathcal{A}(\cdot)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \, \big\|_{L^p(0,T)} < \infty. \tag{4.4}$$

The space $L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$ is a Banach space with the norm defined in (4.4) [BlaN10, p. 74 & Prop. 3.7]. By the definition of the spaces, it follows $L^{p'}[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})] \hookrightarrow L^p[a,b;\mathscr{L}(\mathcal{X},\mathcal{Y})]$ for $1 \leq p \leq p' \leq \infty$ and $0 \leq a < b \leq T$. Furthermore, one should not confuse the space $L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$ with the Bochner space $L^p(0,T;\mathscr{L}(\mathcal{X},\mathcal{Y}))$. For example, for δ from Example 4.4.i) we have $\|\delta(t)\|_{\mathscr{L}(C([0,T],\mathbb{R}),\mathbb{R})} \equiv 1$. Especially, δ is an element of $L^\infty[0,T;\mathscr{L}(C([0,T],\mathbb{R}),\mathbb{R})]$ but not of $L^\infty(0,T;\mathscr{L}(C([0,T],\mathbb{R}),\mathbb{R}))$, since δ is not uniformly measurable. In general, $L^p(0,T;\mathscr{L}(\mathcal{X},\mathcal{Y}))$ is a proper subspace of $L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$; see [BlaN10, p. 75]. Under the conditions of Remark 4.6.i) both spaces coincide.

Next we summarize some properties of functions in $L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$.

Lemma 4.9 ([BlaN10, Th. 3.6] & [GolKT92, Th. 1(ii) & (iv)])**.** *Let $\mathcal{A} \in L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$, $p \in [1,\infty]$. Then the corresponding Nemytskiĭ mapping is an element of $\mathscr{L}(L^q(0,T;\mathcal{X}), L^r(0,T;\mathcal{Y}))$ with $q,r \in [1,\infty]$ and $\frac{1}{r} - \frac{1}{q} = \frac{1}{p}$.*

Lemma 4.10. *Let $\mathcal{X}_1,\mathcal{X}_2,\mathcal{X}_3$ be Banach spaces. Assume $\mathcal{A}_i \in L^{p_i}[0,T;\mathscr{L}(\mathcal{X}_i,\mathcal{X}_{i+1})]$ with $p_i \in [1,\infty]$, $i = 1,2$. Suppose $\frac{1}{q} = \frac{1}{p_1} + \frac{1}{p_2}$ with $q \in [1,\infty]$. Then the pointwise composition $(\mathcal{A}_2\mathcal{A}_1)(t) = \mathcal{A}_2(t)\mathcal{A}_1(t)$ is an element of $L^q[0,T;\mathscr{L}(\mathcal{X}_1,\mathcal{X}_3)]$.*

Proof. Since $\mathcal{A}_1$ and $\mathcal{A}_2$ are pointwise linear, so is $\mathcal{A}_2\mathcal{A}_1$. Let $x \in \mathcal{X}_1$ be arbitrary and consider the constant function $u_x\colon [0,T] \to \mathcal{X}_1; t \mapsto x$. Then $u_x \in L^\infty(0,T;\mathcal{X}_1)$ and by Lemma 4.9 we obtain $\mathcal{A}_1 u_x \in L^{p_1}(0,T;\mathcal{X}_2)$ and $\mathcal{A}_2\mathcal{A}_1 u_x \in L^q(0,T;\mathcal{X}_3)$. In particular, $t \mapsto \mathcal{A}_2(t)\mathcal{A}_1(t)u_x(t) = \mathcal{A}_2(t)\mathcal{A}_1(t)x$ is Bochner-measurable in $\mathcal{X}_3$ and hence $\mathcal{A}_2\mathcal{A}_1$ is strongly measurable. Furthermore, $k_{\mathcal{A}_2\mathcal{A}_1} := \|\mathcal{A}_2(\cdot)\|_{\mathscr{L}(\mathcal{X}_2,\mathcal{X}_3)}\|\mathcal{A}_1(\cdot)\|_{\mathscr{L}(\mathcal{X}_1,\mathcal{X}_2)}$ is an element of $L^q(0,T)$ as product of a L^{p_1} and a L^{p_2}-function [Bre10, Ch. 4, Rem. 2] and satisfies (4.3) for $\mathcal{A}_2\mathcal{A}_1$. The assertion then follows by Lemma 4.7. □

For the later analysis we need a concept of derivatives for strongly measurable operator-valued functions.

Definition 4.11 (Derivative of Operator-Valued Functions)**.** Let $\mathcal{X},\mathcal{Y}$ be Banach spaces and $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y})$ be strongly measurable. Assume that $t \mapsto \mathcal{A}(t)x$ has a kth generalized derivative, $k \in \mathbb{N}_0$, for every $x \in \mathcal{X}$. Then we define the *kth derivative $\mathcal{A}^{(k)}\colon [0,T] \to \mathscr{L}(\mathcal{X},\mathcal{Y})$ of $\mathcal{A}$* by $\mathcal{A}^{(k)}(t)x := \frac{\mathrm{d}^k}{\mathrm{d}t^k}(\mathcal{A}(t)x)$.

Lemma 4.12. *The kth derivative $\mathcal{A}^{(k)}$ of $\mathcal{A}$ defined as in Definition 4.11 is unique. If $\mathcal{A}$ has a kth generalized derivative $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathcal{A}$ in the sense of Definition 3.35, then $\mathcal{A}^{(k)} = \frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathcal{A}$.*

Proof. This immediately follows by considering $t \mapsto \mathcal{A}(t)x$ for fixed but arbitrary $x \in \mathcal{X}$. □

Justified by Lemma 4.12 we also use the notation $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathcal{A} := \mathcal{A}^{(k)}$, $\mathcal{A}^{(0)} := \mathcal{A}$, $\dot{\mathcal{A}} := \mathcal{A}^{(1)}$, and $\ddot{\mathcal{A}} := \mathcal{A}^{(2)}$. Analogously to the Sobolev-Bochner spaces we introduce a space of strongly measurable operator-valued functions with derivatives.

Definition 4.13 (Space $W^{k,p}[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$)**.** Let $\mathcal{X},\mathcal{Y}$ be Banach spaces. We say that $\mathcal{A}\colon [0,T]\to \mathscr{L}(\mathcal{X},\mathcal{Y})$ is an element of $W^{k,p}[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$, $k\in\mathbb{N}_0$ and $p\in[1,\infty]$, if $\mathcal{A}^{(i)}\in L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$ for all $i=0,\dots,k$. Further, we define $H^k[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})] := W^{k,2}[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$.

By the Definition 4.11 of the derivative, the operator-valued function $\mathcal{A}$ inherits properties of the generalized differentiable function $\mathcal{A}(\cdot)x$.

Lemma 4.14. *Let $\mathcal{A}\in W^{k,p}[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$, $k\in N$, $p\in[1,\infty]$. Then $\mathcal{A}$ has a $(k-1)$ times continuously differentiable representative.*

Proof. For $k=1$, the assertion holds, since $\|\dot{\mathcal{A}}(\cdot)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})}\in L^p(0,T)$ and thus the term

$$\sup_{x\in\mathcal{X}\setminus\{0\}}\frac{\|\mathcal{A}(t)x-\mathcal{A}(s)x\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}} = \sup_{x\in\mathcal{X}\setminus\{0\}}\frac{1}{\|x\|_{\mathcal{X}}}\left\|\int_s^t \dot{\mathcal{A}}(\eta)x\,\mathrm{d}\eta\right\|_{\mathcal{Y}} \overset{(4.3)}{\leq} \int_{\min(s,t)}^{\max(s,t)} \|\dot{\mathcal{A}}(\eta)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})}\,\mathrm{d}\eta$$

vanishes as $s\to t$ for almost all $s,t\in[0,T]$ by [KufJF77, Cor. 2.19.10]. Here, we used Theorem 3.36.ii). The cases $k>1$ follow by induction. □

Lemma 4.15. *Let $\mathcal{X}_1,\mathcal{X}_2,\mathcal{X}_3$ be Banach spaces. Suppose $\mathcal{A}_i\in W^{k,p}[0,T;\mathscr{L}(\mathcal{X}_i,\mathcal{X}_{i+1})]$, $i=1,2$, and $u\in W^{k,p}(0,T;\mathcal{X}_1)$, $k\in\mathbb{N}$. Then $\mathcal{A}_1u\in W^{k,p}(0,T;\mathcal{X}_2)$ and $\mathcal{A}_2\mathcal{A}_1\in W^{k,p}[0,T;\mathscr{L}(\mathcal{X}_1,\mathcal{X}_3)]$ with*

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}(\mathcal{A}_1u)=\sum_{\ell=0}^{k}\binom{k}{\ell}\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\mathcal{A}_1\frac{\mathrm{d}^{k-\ell}}{\mathrm{d}t^{k-\ell}}u \quad \text{and} \quad \frac{\mathrm{d}^k}{\mathrm{d}t^k}(\mathcal{A}_2\mathcal{A}_1)=\sum_{\ell=0}^{k}\binom{k}{\ell}\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\mathcal{A}_2\frac{\mathrm{d}^{k-\ell}}{\mathrm{d}t^{k-\ell}}\mathcal{A}_1.$$

Proof. The assertion on $\mathcal{A}_1u$ follows the lines of [Wlo87, Th. 27.1] where the case $p=\infty$ is shown. The assertion on $\mathcal{A}_2\mathcal{A}_1$ is a consequence of the first part of this lemma, Lemma 4.10, and 4.14. □

We finish this section by investigating pointwise elliptic operator-valued functions.

Definition 4.16 (Uniformly Elliptic and Uniform Gårding Inequality)**.** Let $\mathcal{V}$ be a Hilbert space and $\mathcal{A}\colon[0,T]\to\mathscr{L}(\mathcal{V},\mathcal{V}^*)$ be weakly measurable. We call $\mathcal{A}$ *uniformly elliptic on* $\mathcal{V}$, if for all $v\in\mathcal{V}$ the inequality $\langle\,\mathcal{A}(t)v,v\,\rangle\geq\mu_{\mathcal{A}}\|v\|_{\mathcal{V}}^2$ holds at almost every time-point $t\in[0,T]$ with a constant $\mu_{\mathcal{A}}>0$. Given a Gelfand triple $\mathcal{V},\mathcal{H},\mathcal{V}^*$, we say that $\mathcal{A}$ satisfies *uniformly a Gårding inequality on* $\mathcal{V}$, if a constant $\kappa_{\mathcal{A}}\in\mathbb{R}$ exist, such that $\mathcal{A}+\kappa_{\mathcal{A}}\,\mathrm{id}_{\mathcal{H}}$ is uniformly elliptic.

Lemma 4.17. *Let $\mathcal{A}\in W^{1,p}[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ be uniformly elliptic. Then $\mathcal{A}$ has a uniformly elliptic pointwise inverse $\mathcal{A}^{-1}\in W^{1,p}[0,T;\mathscr{L}(\mathcal{V}^*,\mathcal{V})]$ with derivative $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{-1}=-\mathcal{A}^{-1}\dot{\mathcal{A}}\mathcal{A}^{-1}$.*

Proof. Since $\mathcal{A}$ has a continuous representative by Lemma 4.14, we define pointwise $\mathcal{A}^{-1}(t):=(\mathcal{A}(t))^{-1}$. The inverse then satisfies the pointwise estimate $\|\mathcal{A}^{-1}(t)\|_{\mathscr{L}(\mathcal{V}^*,\mathcal{V})}\leq\frac{1}{\mu_{\mathcal{A}}}$; cf. [Alt16, p. 166]. Furthermore, for every $s,t\in[0,T]$ we have

$$\begin{aligned}\|\mathcal{A}^{-1}(t)-\mathcal{A}^{-1}(s)\|_{\mathscr{L}(\mathcal{V}^*,\mathcal{V})} &= \|\mathcal{A}^{-1}(s)\mathcal{A}(s)\mathcal{A}^{-1}(t)-\mathcal{A}^{-1}(s)\mathcal{A}(t)\mathcal{A}^{-1}(t)\|_{\mathscr{L}(\mathcal{V}^*,\mathcal{V})}\\ &\leq \|\mathcal{A}^{-1}(s)\|_{\mathscr{L}(\mathcal{V}^*,\mathcal{V})}\|\mathcal{A}(s)-\mathcal{A}(t)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)}\|\mathcal{A}^{-1}(t)\|_{\mathscr{L}(\mathcal{V}^*,\mathcal{V})}.\end{aligned}$$

This shows the continuity of $\mathcal{A}^{-1}$ and in particular $\mathcal{A}^{-1}\in L^\infty[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$. For its derivative we note that $0=\frac{\mathrm{d}}{\mathrm{d}t}x=\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{A}\mathcal{A}^{-1}x)=\dot{\mathcal{A}}\mathcal{A}^{-1}x+\mathcal{A}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{A}^{-1}x)$ for all $x\in\mathcal{X}$. Therefore, $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{-1}=-\mathcal{A}^{-1}\dot{\mathcal{A}}\mathcal{A}^{-1}\in L^p[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$ holds by Lemma 4.10. To show that $\mathcal{A}^{-1}$ is uniformly elliptic we observe $0<\mu_{\mathcal{A}}\leq\|\mathcal{A}(t)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)}\leq\|\mathcal{A}\|_{C([0,T],\mathscr{L}(\mathcal{V},\mathcal{V}^*))}$. Therefore, the estimate

$$\langle\, f,\mathcal{A}^{-1}(t)f\,\rangle=\langle\,\mathcal{A}(t)\mathcal{A}^{-1}(t)f,\mathcal{A}^{-1}(t)f\,\rangle\geq\mu_{\mathcal{A}}\|\mathcal{A}^{-1}(t)f\|_{\mathcal{V}}^2\geq\frac{\mu_{\mathcal{A}}}{\|\mathcal{A}\|_{C([0,T],\mathscr{L}(\mathcal{V},\mathcal{V}^*))}^2}\|\mathcal{A}(t)\mathcal{A}^{-1}(t)f\|_{\mathcal{V}^*}^2 \tag{4.5}$$

is well-defined for every $f\in\mathcal{V}^*$ and at almost every time-point $t\in[0,T]$. □

4.2. Volterra Integral Equations of Second Kind

As the first example of a dynamic equation in a Banach space we investigate the *Volterra integral equation (of second kind)*

$$u(t) = f(t) + \mathcal{A}_1(t) \int_0^t \mathcal{A}_2(s)u(s)\,\mathrm{d}s \tag{4.6}$$

on a bounded time interval $[0,T]$. We call a Bochner-measurable function $u\colon [0,T] \to \mathcal{X}$ a *solution of* (4.6), if $t \to \mathcal{A}_2(t)u(t)$ is Bochner-integrable on $[0,T]$ and (4.6) is satisfied at almost every time-point $t \in [0,T]$. In the monographs [GriLS90; Lin85] integral equations in finite dimensional spaces $\mathcal{X}, \mathcal{Y}$ are analyzed. For studies in infinite dimension we refer to [GajGZ74, Ch. V § 1] and [Zei90b, Ch. 28]. However, these references do not include integral equations of the form (4.6). In Subsection 4.3.1 we use them to investigate operator differential equations and in Chapter 7 to make statements about the Lagrange multiplier of operator differential-algebraic equations.

Let $p \in [1,\infty]$ and $\mathcal{X}$, $\mathcal{Y}$ be separable Banach spaces. We assume $f \in L^p(0,T;\mathcal{X})$. The linear, time-dependent operators $\mathcal{A}_1$ and $\mathcal{A}_2$ are elements of $L^p[0,T;\mathscr{L}(\mathcal{Y},\mathcal{X})]$ and $L^q[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$, respectively, with p and q being conjugated indices. For the analysis we write (4.6) in the compact form

$$\boldsymbol{\Phi}_{[a,b]}\colon ([a,b] \to \mathcal{X}) \to ([a,b] \to \mathcal{X}), \quad u \mapsto f + \boldsymbol{\mathcal{A}}_{[a,b]}u \tag{4.7}$$

with $0 \le a < b \le T$ and the integral part

$$(\boldsymbol{\mathcal{A}}_{[a,b]}u)(t) = \mathcal{A}_1(t) \int_a^t \mathcal{A}_2(s)u(s)\,\mathrm{d}s \tag{4.8}$$

for almost every $t \in [a,b]$. With the definition of $\boldsymbol{\Phi}_{[a,b]}$ the stated integral equation (4.6) becomes the fixed point problem $u = \boldsymbol{\Phi}_{[0,T]}u$. To investigate the fixed point problem, we summarize some properties of $\boldsymbol{\Phi}_{[a,b]}$ and $\boldsymbol{\mathcal{A}}_{[a,b]}$ in the following lemma.

Lemma 4.18. *Assume that $f \in L^p(0,T;\mathcal{X})$, $\mathcal{A}_1 \in L^p[0,T;\mathscr{L}(\mathcal{Y},\mathcal{X})]$, and $\mathcal{A}_2 \in L^q[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]$ are given with conjugated indices p and q. Let $0 \le a < b \le T$ and $\boldsymbol{\Phi}_{[a,b]}$ as well as $\boldsymbol{\mathcal{A}}_{[a,b]}$ be defined as in (4.7) and (4.8), respectively. Then $\boldsymbol{\Phi}_{[a,b]}$ maps $L^p(a,b;\mathcal{X})$ Lipschitz continuously into itself and $\boldsymbol{\mathcal{A}}_{[a,b]}$ is an element of $\mathscr{L}(L^p(a,b;\mathcal{X}))$ with*

$$\big\|\boldsymbol{\mathcal{A}}_{[a,b]}\big\|_{\mathscr{L}(L^p(a,b;\mathcal{X}))} \le \|\mathcal{A}_1\|_{L^p[a,b;\mathscr{L}(\mathcal{Y},\mathcal{X})]}\|\mathcal{A}_2\|_{L^q[a,b;\mathscr{L}(\mathcal{X},\mathcal{Y})]}. \tag{4.9}$$

Proof. We only have to prove $\boldsymbol{\mathcal{A}}_{[a,b]} \in \mathscr{L}(L^p(a,b;\mathcal{X}))$. The stated properties of $\boldsymbol{\Phi}_{[a,b]}$ then follow immediately with $f \in L^p(0,T;\mathcal{X})$. The linearity of $\boldsymbol{\mathcal{A}}_{[a,b]}$ follows by the linearity of $\mathcal{A}_1$, $\mathcal{A}_2$, and of the integral. For $p < \infty$ and $u \in L^p(a,b;\mathcal{X})$, we get the estimate

$$\begin{aligned}\big\|\boldsymbol{\mathcal{A}}_{[a,b]}u\big\|^p_{L^p(a,b;\mathcal{X})} &\le \int_a^b \left(\|\mathcal{A}_1(t)\|_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \int_a^t \|\mathcal{A}_2(s)\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})}\|u(s)\|_{\mathcal{X}}\,\mathrm{d}s\right)^p \mathrm{d}t \\ &\le \|\mathcal{A}_1\|^p_{L^p[a,b;\mathscr{L}(\mathcal{Y},\mathcal{X})]}\|\mathcal{A}_2\|^p_{L^q[a,b;\mathscr{L}(\mathcal{X},\mathcal{Y})]}\|u\|^p_{L^p(a,b,\mathcal{X})},\end{aligned}$$

where we used Hölder's inequality in the last line. Analogously, one proves (4.9) for $p = \infty$. □

With Lemma 4.18 we can prove the solvability of the integral equation (4.6).

Theorem 4.19. *Let the assumption of Lemma 4.18 be satisfied. Then the Volterra integral equation (4.6) has a unique solution $x \in L^p(0,T;\mathcal{X})$, which depends linearly and continuously on f.*

Proof. With the bound (4.9) of $\boldsymbol{\mathcal{A}}_{[a,b]}$, we find a partition of $[0,T]$, $0 = t_0 < t_1 < \ldots < t_N = T$ with $\big\|\boldsymbol{\mathcal{A}}_{[t_{i-1},t_i]}\big\|_{\mathscr{L}(L^p(t_{i-1},t_i;\mathcal{X}))} \le \frac{1}{2}$. Therefore, $\boldsymbol{\Phi}_{[t_{i-1},t_i]}$ is a contraction in $L^p(t_{i-1},t_i;\mathcal{X})$, i.e., a Lipschitz continuous function from $L^p(t_{i-1}, t_i; \mathcal{X})$ into itself with a Lipschitz constant smaller than one. By the Banach fixed-point theorem [Zei86, Th. 1.A], the map $\boldsymbol{\Phi}_{[0,t_1]}$ has a unique fixed-point $u_1 \in L^p(0, t_1; \mathcal{X})$. In particular, $u_1 = \lim_{n\to\infty} \boldsymbol{\Phi}^n_{[0,t_1]} f\big|_{[0,t_1]} = \sum_{k=0}^{\infty} \boldsymbol{\mathcal{A}}^k_{[0,t_1]} f\big|_{[0,t_1]}$. Therefore, we have

$$\|u_1\|_{L^p(0,t_1;\mathcal{X})} \leq \sum_{k=0}^{\infty} \|\boldsymbol{\mathcal{A}}_{[0,t_1]}\|^k_{\mathscr{L}(L^p(0,t_1;\mathcal{X}))} \|f\|_{L^p(0,t_1;\mathcal{X})} \leq 2\|f\|_{L^p(0,t_1;\mathcal{X})}$$

Iteratively one shows the existence of a unique solution u_i on $[t_{i-1}, t_i]$ by replacing f in the definition of $\boldsymbol{\Phi}_{[t_{i-1},t_i]}$ by $f\big|_{[t_{i-1},t_i]} + \sum_{k=1}^{i-1} \mathcal{A}_1(\cdot) \int_{t_{k-1}}^{t_k} \mathcal{A}_2(s) u_k(s) \,\mathrm{d}s \in L^p(t_{i-1}, t_i; \mathcal{X})$. A solution of (4.6) then is given by $u \in L^p(0, T; \mathcal{X})$ with $u|_{[t_{i-1},t_i]} = u_i$. The solution u is unique, since all $u_1, \ldots, u_N$ are.

The linearity of $f \mapsto u$ follows immediately from the linearity of $\boldsymbol{\mathcal{A}}_{[0,T]}$. Furthermore, by (4.6) and Gronwall's Lemma 3.15 the estimate

$$\|u(t)\|_{\mathcal{X}} \leq \|f(t)\|_{\mathcal{X}} + \|\mathcal{A}_1(t)\|_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \int_0^t \|f\|_{\mathcal{X}} \|\mathcal{A}_2\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \exp\left(\int_s^t \|\mathcal{A}_1\|_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \|\mathcal{A}_2\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \,\mathrm{d}\eta \right) \mathrm{d}s$$

is satisfied at almost every time-point $t \in [0, T]$. For $p = 1$ or $p = \infty$ it follows that $\|u\|_{L^p(0,T;\mathcal{X})}$ is bounded by

$$\Big(1 + \|\mathcal{A}_1\|_{L^p[0,T;\mathscr{L}(\mathcal{Y},\mathcal{X})]} \|\mathcal{A}_2\|_{L^q[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]} \exp\big(\|\mathcal{A}_1\|_{L^p[0,T;\mathscr{L}(\mathcal{Y},\mathcal{X})]} \|\mathcal{A}_2\|_{L^q[0,T;\mathscr{L}(\mathcal{X},\mathcal{Y})]}\big)\Big) \|f\|_{L^p(0,T;\mathcal{X})} \tag{4.10}$$

from above. On the other hand, for $p \in (1, \infty)$ we obtain

$$\begin{aligned}
&\|u(t)\|^p_{\mathcal{X}} - 2^{p-1}\|f(t)\|^p_{\mathcal{X}} \\
&\leq 2^{p-1} \|\mathcal{A}_1(t)\|^p_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \left(\int_0^t \|f\|_{\mathcal{X}} \|\mathcal{A}_2\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \exp\left(\int_s^t \|\mathcal{A}_1\|_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \|\mathcal{A}_2\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \,\mathrm{d}\eta \right) \mathrm{d}s \right)^p \\
&\leq 2^{p-1} \|\mathcal{A}_1(t)\|^p_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \int_0^t \|f\|^p_{\mathcal{X}} \,\mathrm{d}s \left(\int_0^t \|\mathcal{A}_2\|^q_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \exp\left(q \int_s^t \|\mathcal{A}_1\|_{\mathscr{L}(\mathcal{Y},\mathcal{X})} \|\mathcal{A}_2\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \,\mathrm{d}\eta \right) \mathrm{d}s \right)^{\frac{p}{q}}
\end{aligned}$$

by $(a+b)^p \leq 2^{p-1}(a^p + b^p)$ for all $a, b \geq 0$, cf. [Emm04, Th. A.1.5], and Hölder's inequality. Replacing the limits s and t by 0 and T, respectively, and integrating from 0 to T leads to the bound (4.10) for $\|u\|_{L^p(0,T;\mathcal{X})}$ with the prefactor $\sqrt[q]{2}$. □

4.3. Operator Differential Equations

We now investigate linear differential equations for abstract functions in Banach spaces,

$$\dot{u}(t) + \mathcal{A}(t)u(t) = f(t). \tag{4.11}$$

Systems of the form (4.11) are called *operator differential equations* (*operator ODEs*), *abstract ODEs*, *Cauchy problems*, or *evolution equations* in the literature. In contrast to ordinary differential equations, operator ODEs are stated in general Banach spaces $\mathcal{X}$ – not necessarily finite-dimensional. We assume that $\mathcal{X}$ is reflexive and separable.

In our analysis we consider (4.11) on a bounded time interval $[0, T]$. The right-hand side f maps $[0, T]$ to $\mathcal{X}$ and the operator $\mathcal{A}$ maps $[0, T] \times \mathcal{X}$ to $\mathcal{X}$. In the following we distinguish between the case that the linear operator $\mathcal{A}(t)\colon \mathcal{X} \to \mathcal{X}$ is bounded for almost every $t \in [0, T]$ or not. This means, the first Carathéodory condition is either satisfied or not; see Definition 4.1. We always assume that the second Carathéodory condition is fulfilled. However, the two different cases lead to different solution concepts as well as different functional analytic approaches.

4.3.1. Bounded Operators

In this subsection we study the operator ODE (4.11) with $\mathcal{A}(t) \in \mathscr{L}(\mathcal{X})$ for almost all $t \in [0,T]$. As for ordinary differential equations, we need in addition to (4.11) an initial value $u_0 \in \mathcal{X}$. A *solution of the operator ODE* (4.11) *with an initial value* $u_0 \in \mathcal{X}$ is an abstract function $u\colon [0,T] \to \mathcal{X}$, which solves the Volterra integral equation

$$u(t) = u_0 + \int_0^t f(s) - \mathcal{A}(s)u(s)\,\mathrm{d}s. \tag{4.12}$$

An existence analysis for finite-dimensional $\mathcal{X}$ can be found in [Hal80, Ch. I.5]. In [Zei86, Ch. 3] systems with continuous data and general Banach spaces $\mathcal{X}$ are studied. For general Bochner-integrable data we get the following result.

Theorem 4.20. *Let $f \in L^p(0,T;\mathcal{X})$, $\mathcal{A} \in L^p[0,T;\mathscr{L}(\mathcal{X})]$, and $u_0 \in \mathcal{X}$ be given. Then the operator ODE* (4.11) *has a unique solution $u \in W^{1,p}(0,T;\mathcal{X})$ with $u(0) = u_0$. The solution map $(f,u_0) \mapsto u$ is linear and bounded.*

Proof. By Theorem 4.19, equation (4.12) has a unique solution $u \in L^\infty(0,T;\mathcal{X})$ which depends linearly and continuously on (f,u_0). Its generalized derivative is $f - \mathcal{A}u \in L^p(0,T;\mathcal{X})$ by Theorem 3.36. □

4.3.2. Unbounded Operators

Let us now consider the operator ODE (4.11) with an unbounded time-independent operator $\mathcal{A}$. Such systems appear, for example, in the investigation of PDEs. A classical prototype is the heat equation, which describes the flow of heat in a homogeneous and isotropic medium in a domain $\Omega \subset \mathbb{R}^d$. With homogeneous Dirichlet boundary conditions this parabolic PDE is given by

$$\dot{u}(\xi,t) - \Delta u(\xi,t) = f(\xi,t) \qquad \text{in } \Omega \times (0,T], \tag{4.13a}$$

$$u(\xi,t) = 0 \qquad \text{on } \partial\Omega \times (0,T]. \tag{4.13b}$$

Note that the *Laplace operator* $\Delta = \operatorname{div}\nabla$ is unbounded as an operator from $L^2(\Omega)$ into itself. The investigation of the solvability of the heat equation leads to the methods of semigroups or variational methods. These approaches are the topics of the following two subsections.

4.3.2.1. The Method of Semigroups

We recall the basic definition and properties of semigroups for a Banach space $\mathcal{X}$. The following definitions and results are taken from [Paz83, Sec. 1.1 f., & 4.2 f.], if no other references are given.

A one-parameter family $S(t) \in \mathscr{L}(\mathcal{X})$, $t \in \mathbb{R}_{\geq 0}$, is called a *semigroup* if $S(0) = \mathrm{id}_{\mathcal{X}}$ and $S(t_1)S(t_2) = S(t_1 + t_2)$ for every $t_1, t_2 \geq 0$. A semigroup is *strongly continuous* if $S(t)x$ is right-continuous at $t = 0$ for every $x \in \mathcal{X}$. Right-continuity at $t = 0$ implies the continuity of the map $t \mapsto S(t)x$ on $\mathbb{R}_{\geq 0}$ for every $x \in \mathcal{X}$ and the existence of constants $m \geq 1$, $\omega \in \mathbb{R}$ such that

$$\|S(t)\|_{\mathscr{L}(\mathcal{X})} \leq m e^{\omega t}. \tag{4.14}$$

The linear operator A defined by

$$Ax := \lim_{t \to 0^+} \tfrac{S(t)x - x}{t} \quad \text{for } x \in D(A) := \left\{ x \in \mathcal{X} \;\middle|\; \lim_{t\to 0^+} \tfrac{S(t)x-x}{t} \text{ exists} \right\} \subset \mathcal{X}$$

is called the *infinitesimal generator* of the semigroup $S(t)$ with domain $D(A)$. If $S(t)$ is strongly continuous, then $D(A)$ is dense in $\mathcal{X}$ and $S(t)$ is uniquely determined by A.

Example 4.21.

i) For a bounded operator $\mathcal{A} \in \mathscr{L}(\mathcal{X})$ the function $t \mapsto e^{t\mathcal{A}} := \sum_{k=0}^{\infty}(t\mathcal{A})^k/k!$ is well-defined and a strongly continuous semigroup with infinitesimal generator $\mathcal{A}$ [EngN00, Ch. I, Prop. 3.5].

ii) The Laplace operator Δ generates a strongly continuous semigroup $S(t) \in \mathscr{L}(L^2(\Omega))$ with $D(\Delta) = \{u \in H_0^1(\Omega) \,|\, \Delta u \in L^2(\Omega)\}$ for every domain $\Omega \subset \mathbb{R}^d$ [DauL92, p. 328 f.], and with $D(\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$ if $\partial\Omega$ is smooth enough; cf. [Alt16, Rem. 6.9]. In particular, for the unit interval $\Omega = (0,1)$, the identity

$$\big(S(t)u\big)(\xi) = 2\sum_{k=1}^{\infty} e^{-k^2\pi^2 t}\sin(k\pi\xi)\int_0^1 \sin(k\pi\eta)u(\eta)\,\mathrm{d}\eta$$

holds [Olv14, p. 124 ff.].

Motivated by Example 4.21.i) we denote the strongly continuous semigroup $S(t)$ with infinitesimal generator A by e^{tA}. Let us now consider the initial value problem

$$\dot{u}(t) = Au(t) + f(t), \tag{4.15a}$$
$$u(0) = u_0, \tag{4.15b}$$

on $t \in [0,T]$. We call $u\colon [0,T] \to \mathcal{X}$ a *strong solution of* (4.15) if $u \in W^{1,1}(0,T;\mathcal{X})$, $u(0) = u_0$ and (4.15a) is satisfied at almost every time-point $t \in [0,T]$. Inspired by the well-known *variation-of-constants formula*, see [Hal80, Eq. (4.14)], we weaken the assumptions on a solution and say that $u\colon [0,T] \to \mathcal{X}$ is a *mild solution of* (4.15) if

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}f(s)\,\mathrm{d}s \tag{4.16}$$

at almost every time-point $t \in [0,T]$. If $f \in L^1(0,T;\mathcal{X})$ and $u_0 \in \mathcal{X}$, then the initial value problem (4.15) has a unique weak solution. The weak solution is a strong solution, if $u_0 \in D(A)$ and $f \in W^{1,1}(0,T;\mathcal{X})$ or f is Lipschitz continuous for reflexive $\mathcal{X}$.

We call a strongly continuous semigroup e^{tA} *analytic*, [EngN00, Def. 4.5 & Th. 4.6], if a constant $M > 0$ exists such that

$$\|tAe^{tA}\|_{\mathscr{L}(\mathcal{X})} \leq M \tag{4.17}$$

for all $t > 0$. The Laplace operator Δ in Example 4.21.ii) generates an analytic semigroup, [DauL92, p. 412 f.]. The property (4.17) implies that every mild solution u becomes more regular over time in the sense that if f is *Hölder continuous* [Paz83, p. 112], then so is $Au, \dot{u} \in C([\varepsilon,T],\mathcal{X})$, $\varepsilon > 0$, with the same exponent. This property is sometimes called *parabolic smoothing*. In particular, the parabolic smoothing holds for a vanishing right-hand side.

For an analysis of evolution equations (4.11) with time-dependent A and an approach close to semigroups, we refer to [Paz83, Ch. 5].

4.3.2.2. Variational Method

Consider again the heat equation with homogeneous Dirichlet boundary condition (4.13) in a Lipschitz domain. We follow the steps of [Zei90a, Sec. 23.1] and multiply (4.13) by $\varphi \in C_c^{\infty}(\Omega)$, integrate over the domain, and use Green's formula [Rou13, Eq. (1.54)] to derive its *weak formulation*

$$\int_\Omega f(\xi,t)\varphi(\xi)\,\mathrm{d}\xi = \int_\Omega \dot{u}(\xi,t)\varphi(\xi) - \Delta u(\xi,t)\varphi(\xi)\,\mathrm{d}\xi = \int_\Omega \dot{u}(\xi,t)\varphi(\xi) + \nabla u(\xi,t)\cdot\nabla\varphi(\xi)\,\mathrm{d}\xi. \tag{4.18}$$

For the functional analytic treatment of (4.18) we separate the spatial and temporal variable. This allows us to write (4.18) as

$$(\dot{u}(t), \varphi)_{L^2(\Omega)} + a(u(t), \varphi) = (f(t), \varphi)_{L^2(\Omega)} \tag{4.19}$$

with $u\colon [0,T] \to H_0^1(\Omega)$ and the bilinear form

$$a(v_1, v_2) := \int_\Omega \nabla v_1(\xi) \cdot \nabla v_2(\xi)\,\mathrm{d}\xi.$$

Note that we choose $H_0^1(\Omega)$ as the range of u to incorporate the boundary condition (4.13b). The real bilinear form a is bounded in $H_0^1(\Omega)$. By [Zei90a, Prop. 21.31(a)] there is a one-to-one correspondence between a and the operator $\mathcal{A} \in \mathscr{L}(H_0^1(\Omega), H^{-1}(\Omega))$ with $\langle \mathcal{A}v_1, v_2 \rangle = a(v_1, v_2)$. Furthermore, the assumptions on $\dot{u}$ and f can be weakened to be elements of $H^{-1}(\Omega)$. The weak formulation (4.18) then can be written as an operator ODE (4.11) stated in $H^{-1}(\Omega)$, where we have used that $C_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$. Note that $\mathcal{A}$ is unbounded as an operator from $L^2(\Omega)$ into itself.

The derivation of the weak formulation (4.11) requires, in a natural way, the three spaces $H_0^1(\Omega)$, $L^2(\Omega)$, and $H^{-1}(\Omega)$, which form a Gelfand triple [Zei90a, Ex. 23.12]. In general, we study the operator ODE (4.11) with respect to a general Gelfand triple $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$. In addition to the operator ODE (4.11) we have the initial value

$$u(0) = u_0 \in \mathcal{H}. \tag{4.20}$$

We call an abstract function $u\colon [0,T] \to \mathcal{V}$ a *weak solution of the operator ODE* (4.11) *with initial value* u_0 [LioM72, p. 239] if the equality

$$\int_0^T -(u,v)_\mathcal{H}\,\dot{\varphi} + \langle \mathcal{A}, v \rangle_{\mathcal{V}^*,\mathcal{V}}\,\varphi\,\mathrm{d}t = (u_0, v)_\mathcal{H}\,\varphi(0) + \int_0^T \langle f, v \rangle_{\mathcal{V}^*,\mathcal{V}}\,\varphi\,\mathrm{d}t \tag{4.21}$$

is satisfied for every $v \in \mathcal{V}$ and $\varphi \in C^\infty([0,T])$ with $\varphi(T) = 0$.

Note that a *weak solution* of (4.11) is called a *very weak solution* in the context of the associated PDE [Rou13, p. 215]. Furthermore, the term *weak solution* is also used in the context of semigroups and coincides with the *mild solution*; cf. [Paz83, p. 258 f.]. On the one hand, the *mild solution* is a weaker solution concept than the *weak solution* in the variational setting. On the other hand, the existence of a Gelfand triple $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ allows more freedom in the choice of the right-hand side and is connected to the Galerkin method and the finite element method [Zei90a, p. 405].

Theorem 4.22 (Theorem of Lions-Tartar I)**.** *Let $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ satisfy a Gårding inequality* (3.6)*. Assume that $u_0 \in \mathcal{H}$ and $f = f_1 + f_2$ with $f_1 \in L^2(0,T;\mathcal{V}^*)$ and $f_2 \in L^1(0,T;\mathcal{H}^*)$. Then the operator ODE* (4.11) *has a unique weak solution $u \in W_1(0,T;\mathcal{V},\mathcal{V}^*)$, i.e.,*

$$u \in C([0,T],\mathcal{H}) \cap L^2(0,T;\mathcal{V}) \quad \text{ and } \quad \dot{u} \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*),$$

with $u(0) = u_0$. Furthermore, the map $(f, u_0) \mapsto u$ is linear and bounded with

$$\|u(t)\|_\mathcal{H}^2 + \mu_\mathcal{A} \int_0^t \|u(s)\|_\mathcal{V}^2\,\mathrm{d}s \le e^{2\kappa_\mathcal{A} t}\Big[\Big(\|u_0\|_\mathcal{H}^2 + \int_0^t \tfrac{1}{\mu_\mathcal{A}} \|f_1(s)\|_{\mathcal{V}^*}^2\,\mathrm{d}s\Big)^{1/2} + \int_0^t e^{-\kappa_\mathcal{A} s} \|f_2(s)\|_{\mathcal{H}^*}\,\mathrm{d}s\Big]^2. \tag{4.22}$$

Proof. The existence of a unique solution and the continuous dependence on the data is proven in [Tar06, Lem. 19.1] and [DauL92, Ch. XVIII. §3.5]. For the proof of the estimate (4.22), we test the operator ODE (4.11) with the solution u and integrate over $[0,t]$. Together with (3.6) this yields

$$\|u(t)\|_{\mathcal{H}}^2 + \mu_{\mathcal{A}} \int_0^t \|u\|_{\mathcal{V}}^2 \,\mathrm{d}s \leq \|u_0\|_{\mathcal{H}}^2 + \int_0^t \Big(\tfrac{1}{\mu_{\mathcal{A}}} \|f_1\|_{\mathcal{V}^*}^2 + 2\kappa_{\mathcal{A}} \|u\|_{\mathcal{H}}^2 + 2\|f_2\|_{\mathcal{H}^*} \|u\|_{\mathcal{H}} \Big) \,\mathrm{d}s. \tag{4.23}$$

We set $\varphi(t) := \|u(t)\|_{\mathcal{H}}^2$. Then for every $s \in [0,t]$ the estimate (4.23) implies

$$\varphi(s) \leq \psi_\varepsilon(s) := c_\varepsilon + \int_0^s \Big(2\kappa_{\mathcal{A}} \varphi(\eta) + 2\|f_2(\eta)\|_{\mathcal{H}^*} \sqrt{\varphi(\eta)} \Big) \,\mathrm{d}\eta \tag{4.24}$$

with the time-independent constant $c_\varepsilon := \varepsilon + \|u_0\|_{\mathcal{H}}^2 + \int_0^t \frac{1}{\mu_{\mathcal{A}}} \|f_1(\eta)\|_{\mathcal{V}^*}^2 \,\mathrm{d}\eta > 0$ and an arbitrary constant $\varepsilon > 0$. Thus, the function ψ_ε is positive and $\sqrt{\psi_\varepsilon}$ as well as its derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} \sqrt{\psi_\varepsilon} = \frac{\dot{\psi}_\varepsilon}{2\sqrt{\psi_\varepsilon}} = \frac{1}{\sqrt{\psi_\varepsilon}} \big(\kappa_{\mathcal{A}} \varphi + \|f_2\|_{\mathcal{H}} \sqrt{\varphi} \big) \leq \kappa_{\mathcal{A}} \sqrt{\psi_\varepsilon} + \|f_2\|_{\mathcal{H}^*}$$

are well-defined. A differential version of Gronwall's lemma [Emm04, Lem. 7.3.2] implies

$$\|u(t)\|_{\mathcal{H}}^2 + \mu_{\mathcal{A}} \int_0^t \|u(s)\|_{\mathcal{V}}^2 \,\mathrm{d}s \overset{(4.24)}{\leq} \psi_\varepsilon(t) \leq \bigg(\sqrt{c_\varepsilon} + \int_0^t e^{-\kappa_{\mathcal{A}} \eta} \,\|f_2(s)\|_{\mathcal{H}^*} \,\mathrm{d}s \bigg)^2 e^{2\kappa_{\mathcal{A}} t}.$$

Letting ε go to zero proves the estimate (4.22). □

Remark 4.23. The assumption on the right-hand side can be weakened, such that a third summand $f_3 \in L^2(0,T;\mathcal{V})$ with $f(t) \equiv f_1(t) + f_2(t) + \frac{\mathrm{d}}{\mathrm{d}t} t f_3(t)$ exists; cf. [LioM72, Ch. 3 Th. 4.4]. The unique solution u then has only a distributional derivative in general.

Remark 4.24. Under the assumptions of Theorem 4.22 on $\mathcal{A}$ the operator $-\mathcal{A}$ is an infinitesimal generator of an analytic semigroup in $\mathcal{H}$ [DauL92, Ch. XVII.A. §.3 Th. 3 & §.6 Prop. 3]. In particular, estimate (4.14) is satisfied with $m = 1$ and $\omega = \kappa_{\mathcal{A}}$; see [DauL92, p. 413].

The weak formulation of a parabolic PDE with time and state-independent coefficients satisfies the assumptions of Theorem 4.22 [Zei90a, Ch. 23]. Therefore, we will denote the operator ODE (4.11) also as a parabolic PDE under the assumptions of Theorem 4.22.

To get more regular solutions we can assume that the right-hand side is generalized differentiable, see [Wlo87, Th. 27.2] and Section 6.2, or that $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ splits into $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ being self-adjoint and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$. Note that, if $\mathcal{A}$ satisfies a Gårding inequality (3.6), then $\mathcal{A}_1$ is elliptic on $\mathcal{V}$ without loss of generality. This can be seen as follows: With the Gårding constants $\mu_{\mathcal{A}}$ and $\kappa_{\mathcal{A}}$ of $\mathcal{A}$ and $C_{\mathcal{A}_2}$ the continuity constant of $\mathcal{A}_2$ from an arbitrary splitting of $\mathcal{A}$, we set $\widetilde{\mathcal{A}}_1 := \mathcal{A}_1 + \Big(\kappa_{\mathcal{A}} + \frac{C_{\mathcal{A}_2}^2}{2\mu_{\mathcal{A}}} \Big) \mathrm{id}_{\mathcal{H}}$ and $\widetilde{\mathcal{A}}_2 := \mathcal{A}_2 - \Big(\kappa_{\mathcal{A}} + \frac{C_{\mathcal{A}_2}^2}{2\mu_{\mathcal{A}}} \Big) \mathrm{id}_{\mathcal{H}}$. Then the sum of the self-adjoint $\widetilde{\mathcal{A}}_1 \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ and $\widetilde{\mathcal{A}}_2 \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$ still equals $\mathcal{A}$. Furthermore, $\widetilde{\mathcal{A}}_1$ satisfies

$$\langle \widetilde{\mathcal{A}}_1 v, v \rangle \geq \mu_{\mathcal{A}} \|v\|_{\mathcal{V}}^2 - \kappa_{\mathcal{A}} \|v\|_{\mathcal{H}}^2 - C_{\mathcal{A}_2} \|v\|_{\mathcal{V}} \|v\|_{\mathcal{H}} + \bigg(\kappa_{\mathcal{A}} + \frac{C_{\mathcal{A}_2}^2}{2\mu_{\mathcal{A}}} \bigg) \|v\|_{\mathcal{H}}^2 \overset{(3.8)}{\geq} \frac{\mu_{\mathcal{A}}}{2} \|v\|_{\mathcal{V}}^2$$

for all $v \in \mathcal{V}$. Hence, throughout this thesis we assume that, given the splitting $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ as described above, $\mathcal{A}_1$ is elliptic on $\mathcal{V}$. As a result, $\mathcal{A}_1$ induces an equivalent norm in $\mathcal{V}$, i.e.,

$$\mu_{\mathcal{A}_1} \|v\|_{\mathcal{V}}^2 \leq \|v\|_{\mathcal{A}_1}^2 \leq C_{\mathcal{A}_1} \|v\|_{\mathcal{V}}^2. \tag{4.25}$$

Theorem 4.25 (Theorem of Lions-Tartar II; [Tar06, Ch. 21] & [Zim15, Th. 3.16])**.** *Suppose that $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ satisfies a Gårding inequality (3.6) and $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ being self-adjoint and elliptic and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$. Assume that $u_0 \in \mathcal{V}$ and $f \in L^2(0,T;\mathcal{H}^*)$.*

Then the operator ODE (4.11) *has a unique weak solution* $u \in H^1(0,T;\mathcal{H}) \cap C([0,T],\mathcal{V})$ *with* $\mathcal{A}u \in L^2(0,T;\mathcal{H}^*) \cap C([0,T],\mathcal{V}^*)$ *and* $u(0) = u_0$*. The map* $(f,u_0) \mapsto u$ *is linear and bounded with*

$$\int_0^T \|\dot u(s)\|_{\mathcal{H}}^2 \,\mathrm{d}s + \|u(t)\|_{\mathcal{A}_1}^2 \le \exp\Big(2\frac{C_{\mathcal{A}_2}^2}{\mu_{\mathcal{A}_1}}t\Big)\Big[\|u_0\|_{\mathcal{A}_1}^2 + 2\int_0^t \|f(s)\|_{\mathcal{H}^*}^2 \,\mathrm{d}s\Big]. \tag{4.26}$$

Furthermore, the estimate (4.26) *implies that* $\mathcal{A}u \in L^2(0,T;\mathcal{H}^*)$ *is bounded in terms of the data.*

Remark 4.26. Alternatively to (4.26), one can bound the solution in Theorem 4.25 via

$$\int_0^t \|\dot u(s)\|_{\mathcal{H}}^2 \,\mathrm{d}s + \|u(t)\|_{\mathcal{A}_1}^2 \le \|u_0\|_{\mathcal{A}_1}^2 + 4\frac{C_{\mathcal{A}_2}}{\mu_{\mathcal{A}}} e^{2\kappa_{\mathcal{A}} t}\|u_0\|_{\mathcal{H}}^2 + \Big(2 + \frac{2C_{\mathcal{A}_2}^2}{\kappa_{\mathcal{A}}\mu_{\mathcal{A}}}(e^{2\kappa_{\mathcal{A}} t} - 1)\Big)\int_0^t \|f(s)\|_{\mathcal{H}^*}^2 \,\mathrm{d}s. \tag{4.27}$$

The estimate is still valid for an elliptic $\mathcal{A}$ by setting $\kappa_{\mathcal{A}} = 0$ with $\frac{\exp(2\kappa_{\mathcal{A}} t)-1}{\kappa_{\mathcal{A}}}|_{\kappa_{\mathcal{A}}=0} := 2t$.

For the analysis of operator ODEs with a time-dependent operator $\mathcal{A}$ by variational methods we refer to [DauL92, Ch. XVIII. § 3] and Section 7.1.

4.4. Operator Differential-Algebraic Equations

As DAEs are roughly speaking ODEs with additional algebraic restrictions, so-called *constrained PDEs* or *partial differential-algebraic equations* (*PDAE*) are dynamic PDEs with additional constraints. These constraints are possibly given by spatial differential operators. The incompressible Navier-Stokes equations [Tem77, p. 280], for example, force its velocity field u to be divergence-free, i.e., $\operatorname{div} u = 0$. The temporal derivative of the pressure does not appear at all.

As the weak formulation of PDEs leads to operator ODEs, see Subsection 4.3.2, PDAEs becomes to so-called *abstract DAEs*, often referred as *operator DAEs*. In general, operator DAEs are operator ODEs with additional constraints. These constraints force the solution to lie on a manifold of a Banach space, e.g., the set of functions with vanishing divergence for the incompressible Navier-Stokes equations. On the other hand, operator DAEs can also be characterized by the fact that their spatial Galerkin discretization leads to a DAE; see [Alt15, Sec. 4.3].

Operator DAEs with time-independent operators are well-studied; see [FavY99; Rei06; Sho10] for a semigroup ansatz and [Alt15; EmmM13; Hei14; Zim15] for a variational approach as well as [DauL93; Tar06; Tem77] in the context of fluid dynamics. In this thesis we investigate semi-explicit operator DAEs of the form

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) + (\mathcal{A} - \tfrac12\dot{\mathcal{M}})u \;-\; \mathcal{B}^*\lambda = f \qquad \text{in } \mathcal{V}^*, \tag{4.28a}$$

$$\mathcal{B}u \qquad = g \qquad \text{in } \mathcal{Q}^* \tag{4.28b}$$

on a bounded time interval $[0,T]$, $T > 0$. The spaces $\mathcal{V}$ and $\mathcal{Q}$ are assumed to be reflexive and separable Banach spaces. We suppose that a separable Hilbert space $\mathcal{H}$ exists, such that $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ forms a Gelfand triple. Operator DAEs of the form (4.28) can be derived by considering operator ODEs of the form $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) + (\mathcal{A} - \frac12\dot{\mathcal{M}})u = f$ in $\mathcal{V}^*$ and introducing the constraint (4.28b) by the Lagrange multiplier method [Ste08, Ch. 4.1.2]; see also [Alt15, Ch. 6]. Anyway, the operators are possibly time-dependent with $\mathcal{M}(t) \in \mathscr{L}(\mathcal{H},\mathcal{H}^*)$, $\mathcal{A}(t) \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$, and $\mathcal{B}(t) \in \mathscr{L}(\mathcal{V},\mathcal{Q}^*)$. The right-hand sides f and g are pointwise elements of $\mathcal{V}^*$ and $\mathcal{Q}^*$, respectively, at almost every time-point $t \in [0,T]$. Both functions are in general time-dependent, and f may also depend on the solution u itself.

In the Parts B and C we assume that $\mathcal{A}$ (uniformly) satisfies a Gårding inequality (3.6) on the pointwise kernel of $\mathcal{B}$. Referring to the discussion of parabolic PDEs in Subsection 4.3.2.2, we will then say that the operator DAE (4.28) is of *parabolic type* or is a *constrained, parabolic PDE.*

In addition to the operator DAE (4.28) we assume that an initial condition

$$u(0) = u_0 \in \mathcal{H} \tag{4.29}$$

is given. From the theory of DAEs it is known that by the constraint (4.28b) the initial value cannot be arbitrary and must be consistent in some sense. However, since the domain of $\mathcal{B}(0)$ is $\mathcal{V}$, an evaluation of $\mathcal{B}(0)$ at the initial value $u_0 \in \mathcal{H}$ is not well-defined in general. In Part B we characterize the set of consistent initial values, which incorporate $u_0 \in \mathcal{H}$ with a formal consistency condition $\mathcal{B}(0)u_0 = g(0)$. This is also discussed in [EmmM13, Rem. 3.1] and [AltH18, Cor. 3.5]. A solution (u, λ) then should satisfy the operator DAE (4.28) and the initial condition (4.29) in a distributional sense; cf. [EmmM13, p. 462] and (4.21).

Definition 4.27 (Solution of Operator DAE (4.28))**.** We call a tuple (u, λ) of abstract functions $u\colon [0,T] \to \mathcal{V}$ and $\lambda\colon [0,T] \to \mathcal{Q}$ a *(weak) solution of the operator DAE* (4.28) *with initial condition* (4.29) if an abstract function $\Lambda_{\mathcal{B}^*}\colon [0,T] \to \mathcal{V}^*$ with distributional derivative $\frac{\mathrm{d}}{\mathrm{d}t}\Lambda_{\mathcal{B}^*} = \mathcal{B}^*\lambda$ exists such that the identity

$$\int_0^T -\langle \mathcal{M}u, v\rangle_{\mathcal{H}^*,\mathcal{H}}\,\dot\varphi + \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u, v\rangle_{\mathcal{V}^*,\mathcal{V}}\,\varphi + \langle \Lambda_{\mathcal{B}^*}, v\rangle_{\mathcal{V}^*,\mathcal{V}}\,\dot\varphi + \langle \mathcal{B}u, q\rangle_{\mathcal{Q}^*,\mathcal{Q}}\,\varphi\,\mathrm{d}t$$
$$= \langle \mathcal{M}(0)u_0, v\rangle_{\mathcal{H}^*,\mathcal{H}}\,\varphi(0) + \int_0^T \langle f, v\rangle_{\mathcal{V}^*,\mathcal{V}}\,\varphi + \langle g, q\rangle_{\mathcal{Q}^*,\mathcal{Q}}\,\varphi\,\mathrm{d}t \tag{4.30}$$

holds for every $v \in \mathcal{V}$, $q \in \mathcal{Q}$, and $\varphi \in C^\infty([0,T])$ with $\varphi(T) = 0$. All integrals in (4.30) are assumed to be well-defined.

Note that, neither the derivative $\dot u$ nor the Lagrange multiplier λ are required in (4.30). They exist in general only as distributional derivatives [EmmM13, Sec. 3.2.1].

Remark 4.28 (Superposition Principle)*.* Since the operators are pointwise linear the solutions of the operator DAE (4.28) are linear in the data (f, g, u_0), if f and g are independent of u and λ.

Remark 4.29*.* The operator DAE (4.28) can formally be rewritten as

$$\mathcal{M}\dot u + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u - \mathcal{B}^*\lambda = f \quad \text{in } \mathcal{V}^*, \tag{4.31a}$$
$$\mathcal{B}u = g \quad \text{in } \mathcal{Q}^*. \tag{4.31b}$$

For a uniformly elliptic $\mathcal{A}$, system (4.31) is in the form of a pHDAE (2.5) in infinite dimensions with the operator-valued functions

$$\mathcal{E} = \begin{bmatrix} \mathcal{M} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{R} = \begin{bmatrix} \frac{1}{2}(\mathcal{A} + \mathcal{A}^*) & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{J} = \begin{bmatrix} \frac{1}{2}(\mathcal{A}^* - \mathcal{A} - \dot{\mathcal{M}}) & \mathcal{B}^* \\ -\mathcal{B} & 0 \end{bmatrix}$$

and the control given by the right-hand sides f, g. The system (4.31) also fits in the framework of [MehM19, Def. 1] with the same $\mathcal{R}$ and $\mathcal{J}$ but without $\frac{1}{2}\dot{\mathcal{M}}$. The so-called effort function and time-flow function in the definition of pHDAE in [MehM19] is given by the state itself and by $(\frac{1}{2}\dot{\mathcal{M}}u, 0)$, respectively.

5. Temporal Discretization

In this chapter, we introduce the two temporal integration schemes, which we are focusing on in this thesis. The two families of iterative one-step methods are implicit Runge-Kutta methods and explicit exponential integrators. Runge-Kutta methods are recapped with its application to DAEs in Section 5.1. In Section 5.2 we study exponential integrators. Exponential integrators are used to solve dynamical systems with a nonlinear right-hand side. The nonlinear part is approximated by polynomials and the resulting system is solved exactly. These integrators are introduced by their application to ODEs and PDEs.

In this thesis we focus on a uniform partition of the interval $[0, T]$ with step size $\tau = T/N$. For a given dynamical system with solution x, e.g., an ODE, a DAE, or a PDE, a numerical integration scheme is said to be *convergent of order* p if x_n as the approximation of x at timepoint $t_n = n\tau$, $n = 1, \dots, N$, satisfies

$$\|x_n - x(t_n)\| \lesssim \tau^{\mathsf{p}} + \text{h.o.t.}$$

The acronym h.o.t. stands for *higher order terms* and summarizes finitely many positive terms with prefactors of τ raised to powers greater than p. The constant is independent of τ but may depend on the system's data and the interval length T.

5.1. Runge-Kutta Methods for DAEs

A Runge-Kutta method is defined by the *Butcher tableau*

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^T \end{array} \tag{5.1}$$

with $\mathbf{b}, \mathbf{c} \in \mathbb{R}^s$ and $\mathbf{A} \in \mathbb{R}^{s\times s}$; cf. [KunM06, p. 225]. Here, s denotes the number of stages. We say the Runge-Kutta method (5.1) has *classical order* p if it is convergent of order p when applied to ODEs with smooth right-hand sides. It is well-known, see e.g. [HaiNW93, Ch. II, Th. 7.4], that a Runge-Kutta method, which satisfies the assumptions

$$\begin{aligned} \sum_{i=1}^{s} \mathbf{b}_i \mathbf{c}_i^{k-1} &= \frac{1}{k}, && k = 1, \dots, \mathsf{p}, \\ \sum_{j=1}^{s} \mathbf{A}_{ij} \mathbf{c}_j^{k-1} &= \frac{\mathbf{c}_i^k}{k}, && i = 1, \dots, s,\ k = 1, \dots, \mathsf{q}, \\ \sum_{i=1}^{s} \mathbf{b}_i \mathbf{c}_i^{k-1} \mathbf{A}_{ij} &= \frac{\mathbf{b}_j}{k}(1 - \mathbf{c}_j^k), && j = 1, \dots, s,\ k = 1, \dots, \mathsf{r}, \end{aligned}$$

is of classical order p if $\mathsf{p} \le \mathsf{q} + \mathsf{r} + 1$ and $\mathsf{p} \le 2\mathsf{q} + 2$. The integer q is called the *stage order* of the Runge-Kutta method.

For the numerical treatment of DAEs it is necessary that $\mathbf{A}$ is invertible [KunM06, p. 256], which implies that the method is implicit. Therefore, the limit of the *stability function* $R(z) :=$

$1 + z\mathbf{b}^T(I_s - z\mathbf{A})^{-1}\mathbb{1}_s$ [HaiW96, p. 40] as $z \to \infty$ is well-defined,

$$R(\infty) := \lim_{z\to\infty} 1 + z\mathbf{b}^T(I_s - z\mathbf{A})^{-1}\mathbb{1}_s = 1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s.$$

Here, $\mathbb{1}_s$ is given by

$$\mathbb{1}_s := \big[1, \ldots, 1\big]^T \in \mathbb{R}^s.$$

Consider an initial value problem for a linear DAE with constant coefficients

$$E\dot{x}(t) + Ax(t) = f(t), \qquad x(0) = x_0. \tag{5.2}$$

We assume that the matrix pair $(E, A) \in \mathbb{R}^{n_x\times n_x} \times \mathbb{R}^{n_x\times n_x}$ is regular, i.e., $\lambda \mapsto \det(\lambda E - A) \not\equiv 0$, $x_0 \in \mathbb{R}^{n_x}$ is consistent, the right-hand side $f\colon [0, T] \to \mathbb{R}^{n_x}$ is sufficiently regular, and a unique solution $x \in C^1([0, T], \mathbb{R}^{n_x})$ of (5.2) is given. With the *Kronecker product* [KunM06, p. 220] given by $\otimes$, an implicit Runge-Kutta method with constant step size τ applied to DAE (5.2) leads to the iteration scheme

$$x_n = (1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)x_{n-1} + (\mathbf{b}^T\mathbf{A}^{-1} \otimes I_{n_x})\boldsymbol{x}_n, \tag{5.3a}$$

$$\frac{1}{\tau}(\mathbf{A}^{-1} \otimes E)(\boldsymbol{x}_n - \mathbb{1}_s \otimes x_{n-1}) + (I_s \otimes A)\boldsymbol{x}_n = \boldsymbol{f}_n. \tag{5.3b}$$

The vector $x_n \in \mathbb{R}^{n_x}$ is an approximation of x at $t_n = n\tau$ and $\boldsymbol{x}_n \in \mathbb{R}^{s\cdot n_x}$ are the so-called *internal stages*. The right-hand side is defined by $\boldsymbol{f}_n := \big[f(t_{n-1} + \tau\mathbf{c}_1)^T, \ldots, f(t_{n-1} + \tau\mathbf{c}_s)^T\big]^T \in \mathbb{R}^{s\cdot n_x}$.

The convergence order of the Runge-Kutta method (5.1) applied to the DAE (5.2) not only depends on the method itself but also on the index of the DAE. It is well-known that a high index, i.e., $i_\nu > 1$, can reduce the convergence order or leads in the extreme to divergence; see e.g. [KunM06, Th. 5.12], [HaiW96, p. 504], [HaiLR89, p. 18 f.] or for a special class of PDAEs [DebS05, Th. 7]. Therefore, one usually reduces the index of the considered DAE beforehand by constructing a DAE of lower index with the same solution space; see [KunM06, Ch. 6] and [HaiW96, Sec. VII.2]. Note that, the naive approach of differentiating the constraints leads to drift-off phenomena [HaiW96, p. 468 f.]. Appropriate index reduction actually reveals and introduces hidden constraints, e.g., equation (2.4) for DAE (2.3), and thus increases the number of algebraic equations instead of differential equations; cf. e.g. [KunM06, Ex. 6.17]. However, if, after a possible index reduction, the DAE (5.2) is of index one, then there exist by [HaiW96, p 378] invertible matrices $P, Q \in \mathbb{R}^{n_x\times n_x}$ such that

$$\begin{bmatrix}\tilde{E}_{11} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}_1\\ \frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}_2\end{bmatrix} + \begin{bmatrix}\tilde{A}_{11} & \tilde{A}_{12}\\ \tilde{A}_{21} & \tilde{A}_{22}\end{bmatrix}\begin{bmatrix}\tilde{x}_1\\ \tilde{x}_2\end{bmatrix} = PEQ\,\tfrac{\mathrm{d}}{\mathrm{d}t}(Q^{-1}x) + PAQ\,Q^{-1}x = Pf = \begin{bmatrix}\tilde{f}_1\\ \tilde{f}_2\end{bmatrix} \tag{5.4}$$

with $Q^{-1}x = [\tilde{x}_1^T\ \tilde{x}_2^T]^T$ and $\tilde{E}_{11}, \tilde{A}_{22}$ be square and regular.

An important class of Runge-Kutta schemes in the numerical treatment of DAEs and operator DAEs are stiffly accurate methods.

Definition 5.1 (Stiffly Accurate; [KunM06, p. 231])**.** A Runge-Kutta scheme with s stages and Butcher tableau $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$ is called *stiffly accurate* if $\mathbf{b}$ satisfies $\mathbf{b}^T = e_s^T\mathbf{A}$ with $e_s = [0, \ldots, 0, 1]^T \in \mathbb{R}^s$.

Example 5.2. A stiffly accurate Runge-Kutta method of second classical order and first stage order with two stages is defined by the Butcher tableau

$$\mathbf{A} = \begin{bmatrix}-3.25 & 6.25\\ -0.25 & 1.25\end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix}-0.25\\ 1.25\end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix}3\\ 1\end{bmatrix}.$$

Another important class for (operator) DAEs are L-stable methods. For their definition, we call a Runge-Kutta scheme *A-stable* [HaiW96, Ch. IV, Def. 3.3], if $|R(z)| \leq 1$ holds for every complex number $z = a + ib$ with $a \leq 0$.

Definition 5.3 (L-Stable; [HaiW96, Ch. IV, Def. 3.7])**.** An A-stable Runge-Kutta method given by the Butcher tableau (5.1) is *L-stable*, if $R(\infty) = 1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s = 0$.

Every A-stable, stiffly accurate scheme is L-stable, since

$$R(\infty) = 1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s = 1 - e_s^T\mathbb{1}_s = 0.$$

As a consequence, stiff components of the dynamical system are numerically damped out fast, [HaiW96, p. 44]. In particular, $R(\infty) = 0$ implies that $(1-\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)x_{n-1}$ vanishes in equation (5.3a). Therefore, the values of the previous step are not needed for variables, which are in the kernel of E. As an example, $\tilde{x}_{2,0}$ is not needed for the numerical integration of the index-1 DAE (5.4). Furthermore, for stiffly accurate Runge-Kutta methods the approximation x_n is given by the last n_x entries of $\boldsymbol{x}_n$. Thus, the numerical approximation $\tilde{x}_{1,n}$, $\tilde{x}_{2,n}$ of (5.4) satisfies all algebraic constraints, i.e., $\tilde{A}_{21}\tilde{x}_{1,n} + \tilde{A}_{22}\tilde{x}_{2,n} = \tilde{f}_2(t_n)$, for every $n = 1, \dots, N$, if $\mathbf{c}_s = \sum_{j=1}^{s}\mathbf{A}_{sj} = \sum_{j=1}^{s}\mathbf{b}_j = 1$.

We can now formulate a convergence result for the implicit Runge-Kutta methods applied to linear DAEs of index 1.

Theorem 5.4 (Convergence Order for DAEs; [HaiW96, p. 380])**.** *Suppose that the Runge-Kutta method* (5.1) *has classical order* p*, stage order* q*, and an invertible matrix* $\mathbf{A}$*. Consider the DAE* (5.4) *with square and invertible matrices* $\tilde{E}_{11}, \tilde{A}_{22}$*. Assume that the initial value is consistent. Then the global error satisfies*

$$\|\tilde{x}_{1,n} - \tilde{x}_1(t_n)\| \lesssim \tau^{\mathsf{p}} + h.o.t. \quad \text{and} \quad \|\tilde{x}_{2,n} - \tilde{x}_2(t_n)\| \lesssim \tau^{\mathsf{k}} + h.o.t.$$

for $n = 1, \dots, N$*, where*
- *i)* $\mathsf{k} = \mathsf{p}$ *if the method is stiffly accurate,*
- *ii)* $\mathsf{k} = \min(\mathsf{p}, \mathsf{q}+1)$ *if* $R(\infty) \in [-1, 1)$*,*
- *iii)* $\mathsf{k} = \min(\mathsf{p}-1, \mathsf{q})$ *if* $R(\infty) = 1$*.*

The method is not convergent if $|R(\infty)| > 1$*.*

In Chapter 8 we study the discretization of the operator DAE (4.28) with Runge-Kutta methods. There, we assume that the operator $\mathcal{A}$ of (4.28) is elliptic on a subspace. It is crucial that the ellipticity is preserved for the time discretized problem; see [LubO95b] for operator ODEs and Example 8.20 for operator DAEs. Therefore, we consider a subclass of Runge-Kutta methods in this thesis.

Definition 5.5 (Algebraically Stable; [HaiW96, Ch. IV, Def. 12.5])**.** A Runge-Kutta scheme with Butcher tableau $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$ is called *algebraically stable* if $\mathbf{b}$ has only non-negative entries and $\mathbf{BA} + \mathbf{A}^T\mathbf{B} - \mathbf{bb}^T$ is positive semidefinite with the diagonal matrix $\mathbf{B} \in \mathbb{R}^{s\times s}$ given by $\mathbf{B}_{ii} = \mathbf{b}_i$.

Theorem 5.6 ([HaiW96, Ch. IV, Th. 12.11])**.** *Every algebraically stable Runge-Kutta method is A-stable.*

For the error analysis in Chapter 8 we make the following observation.

Lemma 5.7. *Suppose the Runge-Kutta method with Butcher tableau* (5.1) *is algebraically stable and* $\mathbf{A}$ *is invertible. Let* $x_0 \in \mathbb{R}$ *and* $\boldsymbol{x} \in \mathbb{R}^s$ *be arbitrary and* $x_1 \coloneqq (1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)x_0 + \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{x}$*. Then we have*

$$2\boldsymbol{x}^T\mathbf{B}\mathbf{A}^{-1}(\boldsymbol{x} - x_0\mathbb{1}_s) \geq x_1^2 - x_0^2. \tag{5.5}$$

Proof. Consider the matrix $\mathbf{M} := \mathbf{B}\mathbf{A} + \mathbf{A}^T\mathbf{B} - \mathbf{b}\mathbf{b}^T \in \mathbb{R}^{s\times s}$. Since the Runge-Kutta method is algebraically stable, $\mathbf{M}$ is positive semidefinite by definition and it follows

$$\begin{aligned}
&2\boldsymbol{x}^T\mathbf{B}\mathbf{A}^{-1}(\boldsymbol{x} - x_0\mathbb{1}_s) - x_1^2 + x_0^2 \\
=&\, 2\boldsymbol{x}^T\mathbf{B}\mathbf{A}^{-1}(\boldsymbol{x} - x_0\mathbb{1}_s) + 2\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s x_0^2 - (\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)^2 x_0^2 \\
&\qquad - 2x_0(1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{x} - (\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{x})^2 \\
=&\, \boldsymbol{x}^T\mathbf{A}^{-T}\mathbf{M}\mathbf{A}^{-1}\boldsymbol{x} - 2\boldsymbol{x}^T\mathbf{A}^{-T}\mathbf{M}\mathbf{A}^{-1}\mathbb{1}_s x_0 + \mathbb{1}_s^T\mathbf{A}^{-T}\mathbf{M}\mathbf{A}^{-1}\mathbb{1}_s x_0^2 \\
=&\, (\boldsymbol{x} - \mathbb{1}_s x_0)^T\mathbf{A}^{-T}\mathbf{M}\mathbf{A}^{-1}(\boldsymbol{x} - \mathbb{1}_s x_0) \geq 0.
\end{aligned}$$

□

5.2. An Introduction to Exponential Integrators

In this section we introduce exponential integrators and recall their basic properties when applied to ODEs and PDEs of parabolic type. For more details we refer to [StrWP12, Ch. 11] and [HocO10].

5.2.1. Exponential Integrators for Ordinary Differential Equations

Exponential integrators were first introduced by Certaine in [Cer60] for the simulation of systems of the form

$$\dot{x}(t) + Ax(t) = f(t, x(t)), \quad x_0 \in \mathbb{R}^{n_x}. \tag{5.6}$$

For the matrix $A \in \mathbb{R}^{n_x\times n_x}$ it is assumed that it possesses eigenvalues with large negative real part. Such systems occur for example in the discretization of semi-linear parabolic PDEs [HaiW96, Sec. IV.1]. Note that, because of their relatively small linear stability domain, most explicit methods require very small step sizes for the simulation of (5.6); see [HaiW96, Sec. IV.2 & V.1] and [HunV03, Ch. II, Sec. 1.4]. On the other hand, implicit methods are not preferable if the evaluation of the nonlinearity f is expensive.

Exponential integrators are based on the *variation-of-constants formula* of the solution of (5.6), see e.g. [Hal80, Eq. (4.14)], given by

$$x(t) = e^{-tA}x_0 + \int_0^t e^{-(t-s)A} f(s, x(s))\,\mathrm{d}s. \tag{5.7}$$

For the construction of exponential integrators for (5.6) we consider the recursively defined *ϕ-functions* [StrWP12, Ch. 11.1]

$$\phi_0(z) := e^{z} := \sum_{\ell=0}^{\infty} \frac{z^\ell}{\ell!}, \qquad \phi_k(z) := \frac{\phi_{k-1}(z) - \phi_{k-1}(0)}{z} = \sum_{\ell=0}^{\infty} \frac{z^\ell}{(\ell+k)!}, \tag{5.8}$$

$k = 1, 2, \ldots$ Note that $\phi_k(A)$ is well-defined for every $k \in \mathbb{N}_0$ and matrix $A \in \mathbb{R}^{n_x\times n_x}$; see [GolV96, Th. 11.2.3]. The importance of the ϕ-functions comes from the fact that

$$\phi_k(z) = \int_0^1 e^{(1-s)z} \frac{s^{k-1}}{(k-1)!}\,\mathrm{d}s,$$

$k \geq 1$. As a consequence, the exact solution given in (5.7) with initial value $x_0 \in \mathbb{R}^{n_x}$ and polynomial right-hand side $f = \sum_{k=0}^{r} \frac{f_k}{k!} t^k$, $r \in \mathbb{N}_0$, with coefficients $f_k \in \mathbb{R}^{n_x}$, $k = 0, \ldots, r$, can be expressed in terms of ϕ_k. More precisely, the solution of (5.6) then is

$$x(t) = e^{-tA}x_0 + \int_0^t e^{-(t-s)A} \sum_{k=0}^{r} \frac{f_k}{k!} s^k\,\mathrm{d}s = \phi_0(-tA)\,x_0 + \sum_{k=0}^{r} \phi_{k+1}(-tA)\,f_k\,t^{k+1}. \tag{5.9}$$

Remark 5.8. By the linearity of formula (5.9) in x_0 and f_k, the function $t \mapsto \phi_0(-tA)\,x_0$ is the solution of (5.6) with initial value x_0 and homogeneous right-hand side, whereas $t \mapsto t^k\phi_k(-tA)\,f_k$ is the solution of (5.6) with vanishing initial value and monomial right-hand side $\frac{f_k}{(k-1)!}\,t^{k-1}$, $k \in \mathbb{N}$.

The main idea of exponential integrators is to replace the nonlinearity f in (5.6) by polynomials designed from the previous approximation and internal stages. For example, considering the interpolation polynomial of degree 0 by evaluating the nonlinearity only in the previous step, we obtain the *exponential Euler scheme*,

$$x_{n+1} = e^{-\tau A}x_n + \tau\phi_1(-\tau A)f(t_n, x_n).$$

For a general exponential integrator, the solution $x(t_{n+1})$ of (5.6) is approximated by

$$x_{n+1} = e^{-\tau A}x_n + \tau\sum_{i=1}^{s} b_i(-\tau A)f(t_n + c_i\tau, X_{n,i}) \tag{5.10a}$$

with *notes* $c_i \in \mathbb{R}_{\geq 0}$. For $i = 1, \dots, s$ the *internal stage* $X_{n,i}$ denotes an approximation of $x(t_n + c_i\tau)$ and is given by

$$X_{n,i} = e^{-c_i\tau A}x_n + \tau\sum_{j=1}^{s} a_{i,j}(-\tau A)f(t_n + c_j\tau, X_{n,j}). \tag{5.10b}$$

In this thesis, we assume that the functions $b_i(z)$ and $a_{i,j}(z)$ are linear combinations of $\phi_k(z)$ and

$$\phi_{k,i}(z) := \phi_k(c_i z), \tag{5.11}$$

$k = 1, 2, \dots$, $i = 1, \dots, s$. Like a Runge-Kutta method, the exponential integrator defined in (5.10) is specified by the *Butcher tableau*

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A}(z) \\ \hline & \mathbf{b}^T(z) \end{array} = \begin{array}{c|ccc} c_1 & a_{1,1}(z) & \dots & a_{1,s}(z) \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s,1}(z) & \dots & a_{s,s}(z) \\ \hline & b_1(z) & \dots & b_s(z) \end{array}. \tag{5.12}$$

Note that $\mathbf{A}$ and $\mathbf{b}$ in (5.12) are function-valued.

In this thesis, we focus on explicit exponential integrators, i.e., $c_1 = 0$ and $a_{i,j} = 0$ if $i \leq j \leq s$. They have the advantages that no nonlinear root finding problems have to be solved and the number of evaluations of the nonlinearity f is known a priori.

For order conditions of explicit exponential integrators applied to ODEs we refer to [StrWP12, Ch. 11.2].

5.2.2. Exponential Integrators for Partial Differential Equations

As a preparation for our investigation of exponential integrators for operator DAEs in Chapter 9, we consider semi-linear systems

$$\dot{u}(t) + \mathcal{A}u(t) = f(t, u(t)) \tag{5.13}$$

with $\mathcal{A}$ mapping linearly from an infinite dimensional Banach space $\mathcal{X}$ into itself. In Chapter 9 the operator $\mathcal{A}$ is usually a (unbounded) differential operator. For a bounded operator $\mathcal{A}\colon \mathcal{X} \to \mathcal{X}$ and $k \in \mathbb{N}_0$ the notation $\phi_k(t\mathcal{A})$ given as an infinite sum (5.8) is well-defined [Alt16, Th. 5.9]. Therefore, the derivation and interpretation of exponential integrators for bounded operators is similar to those in Subsection 5.2.1. If $-\mathcal{A}\colon D(-\mathcal{A}) \subset \mathcal{X} \to \mathcal{X}$ is an unbounded differential operator, which generates a strongly continuous semigroup $e^{-t\mathcal{A}}$, we use the interpretation as in Remark 5.8.

Definition 5.9 (ϕ-Functions for Unbounded Operators)**.** Let the linear operator $-\mathcal{A}\colon D(-\mathcal{A}) \subset \mathcal{X} \to \mathcal{X}$ generate a strongly continuous semigroup $e^{-t\mathcal{A}}$. We define the operator $\phi_0(-t\mathcal{A})\colon \mathcal{X} \to \mathcal{X}$ as the semigroup $e^{-t\mathcal{A}}$ for nonnegative times $t \geq 0$.

For $k \in \mathbb{N}$ we set $\phi_k(-t\mathcal{A})\colon \mathcal{X} \to \mathcal{X}$ to $\mathrm{id}/k!$, if $t = 0$, and for $t > 0$ as the map from $f_k \in \mathcal{X}$ to $u(t)\, t^{-k}$ where $u(t)$ is the mild solution of (5.13) with initial value $u_0 = 0$ and monomial right-hand side $t \to \frac{f_k}{(k-1)!}\, t^{k-1}$.

We obtain the following major property for the corresponding ϕ-functions.

Theorem 5.10 (Properties of the ϕ-Functions; cf. [HocO10, Lem. 2.4])**.** *Assume that the linear operator $-\mathcal{A}$ is the infinitesimal generator of a strongly continuous semigroup $e^{-t\mathcal{A}}$. Then the operators $\phi_k(-t\mathcal{A})$ from Definition 5.9 are elements of $\mathscr{L}(\mathcal{X})$ for all $k = 0, 1, \dots$ and $t \geq 0$.*

With the reinterpretation of the ϕ-functions, the solution formula for bounded operators (5.9) directly translates to the operator ODE (4.11) with an unbounded $\mathcal{A}$; cf. [HocO10]. Furthermore, the approximation u_{n+1} and the internal stages $U_{n,i}$ similar determined as in (5.10a) and (5.10b) are well-defined.

For our investigation in Chapter 9 we take a deeper look at the norm of $\phi_k(-t\mathcal{A})$ in Theorem 5.10. For this, let $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ be a Gelfand triple and $\mathcal{A}$ be an element of $\mathscr{L}(\mathcal{V}, \mathcal{V}^*)$. We suppose that the operator $\mathcal{A}$ satisfies a splitting as in Theorem 4.25. Further, we assume that $\mathcal{A}$ is elliptic. This is reasonable, since by the assumed splitting $\mathcal{A}$ satisfies a Gårding inequality (3.6); see page 31. Therefore, we can add the term $\kappa_{\mathcal{A}} u$ to both sides of the semi-linear operator ODE (5.13) such that the new operator $\mathcal{A} + \kappa_{\mathcal{A}}\,\mathrm{id}_{\mathcal{H}}$ is elliptic. Note that, $-\mathcal{A}$ then generates an analytic semigroup in $\mathcal{H}$, see Remark 4.24, and the function $\phi_k(-t\mathcal{A})\colon \mathcal{H} \to \mathcal{H}$ are well-defined. In the estimate of the norm of these functions, we distinguish between the domains $\mathcal{H} \cong \mathcal{H}^*$ and $\mathcal{V}$ as well as between the codomains $\mathcal{H}$ and $\mathcal{V}$. Here, we use the parabolic smoothing of analytic semigroup $e^{-t\mathcal{A}}$; see Subsection 4.3.2.1.

Lemma 5.11. *Let $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ be elliptic with a self-adjoint and elliptic $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V}, \mathcal{H}^*)$. Then for every $k \in \mathbb{N}$ and $t \geq 0$ the estimates*

a) $\|e^{-t\mathcal{A}}\|_{\mathscr{L}(\mathcal{H})} \lesssim 1$, *b)* $\|e^{-t\mathcal{A}}\|_{\mathscr{L}(\mathcal{V})} \lesssim 1$, *c)* $\|\sqrt{t}e^{-t\mathcal{A}}\|_{\mathscr{L}(\mathcal{H},\mathcal{V})} \lesssim 1 + \sqrt{t}$,

d) $\|\phi_k(-t\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*,\mathcal{H})} \lesssim 1$, *e)* $\|\phi_k(-t\mathcal{A})\|_{\mathscr{L}(\mathcal{V})} \lesssim 1$, *f)* $\|\sqrt{t}\phi_k(-t\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*,\mathcal{V})} \lesssim 1 + \sqrt{t}$

hold with generic constants independent of t.

Proof. By Definition 5.9 of the ϕ-functions, all estimates hold for $t = 0$ and we can consider $t > 0$.

Estimates for $e^{-t\mathcal{A}}$: By the definition of ϕ_0, the function $u(t) = e^{-t\mathcal{A}}u_0 = \phi_0(-t\mathcal{A})u_0$ describes the solution of

$$\dot{u}(t) + \mathcal{A}u(t) = 0 \quad \text{in } \mathcal{V}^*, \qquad u(0) = u_0. \tag{5.14}$$

Estimate a) then follows directly from (4.22) by setting $\kappa_{\mathcal{A}}$ to zero, since

$$\max_{u_0 \in \mathcal{H}} \frac{\|e^{-t\mathcal{A}}u_0\|_{\mathcal{H}}^2}{\|u_0\|_{\mathcal{H}}^2} = \max_{u_0 \in \mathcal{H}} \frac{\|u(t)\|_{\mathcal{H}}^2}{\|u_0\|_{\mathcal{H}}^2} \overset{(4.22)}{\leq} \max_{u_0 \in \mathcal{H}} \frac{\|u_0\|_{\mathcal{H}}^2}{\|u_0\|_{\mathcal{H}}^2} = 1.$$

Analogously, one proves inequality b) with (4.27) and $\kappa_{\mathcal{A}} = 0$. For the case c), i.e., $\phi_0(-t\mathcal{A})\colon \mathcal{H} \to \mathcal{V}$, we observe $\sqrt{t}u \in C([0,T], \mathcal{V})$, $\sqrt{t}\dot{u} \in L^2(0,T;\mathcal{H})$, and $\sqrt{t}\mathcal{A}_1 u \in L^2(0,T;\mathcal{H}^*)$ by [Tar06, Lem. 21.1]. In particular,

$$(\sqrt{t}\dot{u}, \sqrt{t}\dot{u})_{\mathcal{H}} + \langle \sqrt{t}\mathcal{A}_1 u, \sqrt{t}\dot{u} \rangle_{\mathcal{H}^*,\mathcal{H}} = -\langle \sqrt{t}\mathcal{A}_2 u, \sqrt{t}\dot{u} \rangle_{\mathcal{H}^*,\mathcal{H}} \tag{5.15}$$

is well-defined in $L^1(0,T)$. By [Zim15, Th. 3.20], this equality implies the estimate

$$\begin{aligned}
\|\sqrt{t}u(t)\|^2_{\mathcal{A}_1} &= \lim_{t_0\to 0+} \|\sqrt{t}u(t)\|^2_{\mathcal{A}_1} - \|\sqrt{t_0}u(t_0)\|^2_{\mathcal{A}_1} \\
&= \lim_{t_0\to 0+} \int_{t_0}^t 2\langle \sqrt{s}\mathcal{A}_1 u(s), \sqrt{s}\dot{u}(s) \rangle_{\mathcal{H}^*,\mathcal{H}} + \langle \mathcal{A}_1 u(s), u(s) \rangle_{\mathcal{V}^*,\mathcal{V}} \,\mathrm{d}s \\
&\overset{(5.15)}{\leq} \int_0^t C^2_{\mathcal{A}_2} s\|u(s)\|^2_{\mathcal{V}} + C_{\mathcal{A}_1}\|u(s)\|^2_{\mathcal{V}} \,\mathrm{d}s \\
&\overset{(4.22)}{\leq} \frac{C^2_{\mathcal{A}_2} t + C_{\mathcal{A}_1}}{\mu_{\mathcal{A}}} \|u_0\|^2_{\mathcal{H}},
\end{aligned}$$

which shows c) with (4.25) and $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$ [Emm04, Cor. A.1.2].

Estimates for $\phi_k(-t\mathcal{A})$, $k \geq 1$: We fix $t > 0$ and consider $u(t) = \phi_k(-t\mathcal{A})f_k = t^k\phi_k(-t\mathcal{A})[f_k t^{-k}]$ as the solution of the associated operator ODE

$$\dot{u}(s) + \mathcal{A}u(s) = F_k(s) := s^{k-1}\frac{f_k}{t^k(k-1)!} \quad \text{in } \mathcal{V}^*, \qquad u(0) = 0 \tag{5.16}$$

at the time point $s = t$. Thus, by Remark 4.26 and $\kappa_{\mathcal{A}} = 0$ we get

$$\|u(t)\|^2_{\mathcal{A}_1} \overset{(4.27)}{\leq} \left(2 + t\frac{2C^2_{\mathcal{A}_2}}{\mu_{\mathcal{A}}}\right) \int_0^t \|F_k(s)\|^2_{\mathcal{H}^*} \,\mathrm{d}s = \left(\frac{2}{t} + \frac{2C^2_{\mathcal{A}_2}}{\mu_{\mathcal{A}}}\right) \frac{\|f_k\|^2_{\mathcal{H}^*}}{((k-1)!)^2}.$$

This bound gives us the estimate f). Analogously, the bound d) is a direct consequence of Theorem 4.22.

For ϕ_k as a map from $\mathcal{V}$ to $\mathcal{V}$ we note that $\dot{u}$ satisfies the formal derivative of the operator ODE (5.16) with initial value $\dot{u}_0 = F_k(0) \in \mathcal{V} \hookrightarrow \mathcal{H}$ [Emm04, Th. 8.5.1]. Hence, $\dot{u} \in L^2(0,T;\mathcal{V})$ and $\mathcal{A}u \in H^1(0,T;\mathcal{H}^*)$ holds by Theorem 4.22 and 4.25. Therefore, we can test (5.16) with $\mathcal{A}\dot{u} \in L^2(0,T;\mathcal{H}^*) \hookrightarrow L^2(0,T;\mathcal{V}^*)$ by using that $\mathcal{V}$ is densely embedded in $\mathcal{H}^*$. Then integration over time leads to the estimate

$$\mu_{\mathcal{A}} \int_0^t \|\dot{u}(s)\|^2_{\mathcal{V}} \,\mathrm{d}s + \|\mathcal{A}u(t)\|^2_{\mathcal{H}^*} \leq \frac{C^2_{\mathcal{A}}}{\mu_{\mathcal{A}}} \int_0^t \|f_k(s)\|^2_{\mathcal{V}} \,\mathrm{d}s = \frac{1}{t}\frac{C^2_{\mathcal{A}}}{\mu_{\mathcal{A}}} \frac{\|f_k\|^2_{\mathcal{V}}}{((k-1)!)^2}. \tag{5.17}$$

Note that this inequality bounds $\dot{u}$ in $L^2(0,T;\mathcal{V})$. Since $u(0) = 0$, we obtain $\|u(t)\|^2_{\mathcal{V}} \leq t\int_0^t \|\dot{u}\|^2_{\mathcal{V}} \,\mathrm{d}s$ by the Cauchy-Schwarz inequality. This estimate, together with (5.17), imply the estimate e). $\square$

Remark 5.12. In Chapter 9 we consider the case that $\mathcal{V}$ is a closed subspace of $\mathcal{V}_{\text{sup}}$ and that an operator $\mathcal{A}_{\text{sup}} \in \mathscr{L}(\mathcal{V}_{\text{sup}}, \mathcal{V}^*_{\text{sup}})$ exists with $\mathcal{A}_{\text{sup}}\big|_{\mathcal{V}} = \mathcal{A}$. With these additional conditions one can enlarge the domain of estimate e) in Lemma 5.11 to $\mathcal{V}_{\text{sup}}$, i.e., $\|\phi_k(-t\mathcal{A})\|_{\mathscr{L}(\mathcal{V}_{\text{sup}},\mathcal{V})} \lesssim 1$ for $k \geq 1$. The associated proof follows the lines of Lemma 5.11 and estimate (5.17), where we use that $\dot{u} \in W^{1,2}(0,T;\mathcal{V},\mathcal{V}^*) \hookrightarrow L^2(0,T;\mathcal{V}) \subset L^2(0,T;\mathcal{V}_{\text{sup}})$ implies $\mathcal{A}\dot{u} \in L^2(0,T;\mathcal{V}^*_{\text{sup}}) \cap L^2(0,T;\mathcal{H}^*)$.

A proof of Lemma 5.11 based on the theory of semigroups can be found in [HocO05a, Ch. 3 f.]. Under this consideration, the estimates for the operators from $\mathcal{H}^*$ to $\mathcal{V}$ in Lemma 5.11 and Remark 5.12 are not obvious. Anyway, in addition to the estimates above, we need bounds of the operator norm of $\phi_k(-t\mathcal{A})$ as function with codomain $D(-\mathcal{A})$.

Lemma 5.13. *Let the assumptions of Lemma 5.11 be satisfied. Then for every $k \in \mathbb{N}$ and $t \geq 0$ the estimates $\|t\mathcal{A}e^{-t\mathcal{A}}\|_{\mathscr{L}(\mathcal{H},\mathcal{H}^*)} \lesssim 1$ and $\|t\mathcal{A}\phi_k(-t\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*)} \lesssim 1$ hold. Furthermore, we have $\|\sqrt{t}\mathcal{A}e^{-t\mathcal{A}}\|_{\mathscr{L}(\mathcal{V},\mathcal{H}^*)} \lesssim 1 + \sqrt{t}$. All generic constants are independent of t.*

Proof. The estimates for $t = 0$ follow by Definition 5.9. Thus, we can consider $t > 0$.

Estimates for $e^{-t\mathcal{A}}$: Let $u(t) = e^{-t\mathcal{A}}u_0$ be the solution of (5.14) with initial value $u_0 \in \mathcal{H}$. By [Emm04, Th. 8.5.3 f.] we have that $v(t) := t\dot{u}(t) \in W^{1,2}(0,T;\mathcal{V},\mathcal{V}^*)$ solves $\dot{v} + \mathcal{A}v = \dot{u}$ with vanishing initial value. With the associated operator ODE (5.14) this leads to

$$\|t\mathcal{A}u(t)\|^2_{\mathcal{H}^*} \overset{(5.14)}{=} \|t\dot{u}(t)\|^2_{\mathcal{H}} = \|v(t)\|^2_{\mathcal{H}} \overset{(4.22)}{\leq} \frac{1}{\mu_{\mathcal{A}}}\int_0^t \|\dot{u}(s)\|^2_{\mathcal{V}^*}\,\mathrm{d}s \overset{(5.14)}{=} \frac{C^2_{\mathcal{A}}}{\mu_{\mathcal{A}}}\int_0^t \|u(s)\|^2_{\mathcal{V}}\,\mathrm{d}s \overset{(4.22)}{\leq} \frac{C^2_{\mathcal{A}}}{\mu^2_{\mathcal{A}}}\|u_0\|^2_{\mathcal{H}}.$$

For $u_0 \in \mathcal{V}$, we consider a sequence $\{u_{0,n}\}_{n\in\mathbb{N}} \subset \mathcal{V}$ with $\mathcal{A}u_{0,n} \in \mathcal{H}^*$ and $u_{0,n} \to u_0$ in $\mathcal{V}$ as $n \to \infty$. Such a sequence exists, since $\mathcal{A}$ is elliptic and we find arbitrarily close elements in $\mathcal{H}^*$ for $\mathcal{A}u_0 \in \mathcal{V}^*$. Let u_n the solution of (5.14) with initial value $u_{0,n}$ and $v_n := t\dot{u}_n$. Then the estimate

$$\int_0^t \|\dot{v}_n(s)\|^2_{\mathcal{H}}\,\mathrm{d}s \overset{(4.27)}{\leq} \Big(2 + 4\frac{C^2_{\mathcal{A}_2}}{\mu_{\mathcal{A}}}t\Big)\int_0^t \|\dot{u}_n\|^2_{\mathcal{H}}\,\mathrm{d}s \overset{(4.27)}{\leq} C_{\mathcal{A}_1}\Big(2 + 4\frac{C^2_{\mathcal{A}_2}}{\mu_{\mathcal{A}}}t\Big)\|u_{0,n}\|^2_{\mathcal{V}} \tag{5.18}$$

holds. This implies a bound for

$$t\ddot{u}_n(t) = \dot{v}_n(t) - \dot{u}_n(t) \tag{5.19}$$

in $L^2(0,T;\mathcal{H})$. Furthermore, $\dot{u}_n$ satisfies the formal derivative of operator ODE (5.14) with initial value $\dot{u}_{0,n} = -\mathcal{A}u_{0,n} \in \mathcal{H}^* \cong \mathcal{H}$ [Emm04, Th. 8.5.1]. This implies $\sqrt{t}\dot{u}_n(t) \in C([0,T],\mathcal{V})$ and $\sqrt{t}\ddot{u}_n(t) \in L^2(0,T;\mathcal{H})$ by [Tar06, Lem. 21.1] and the estimate

$$\begin{aligned}
\|\sqrt{t}\dot{u}_n(t)\|^2_{\mathcal{H}} &= \lim_{t_0\to 0+} \|\sqrt{t}\dot{u}_n(t)\|^2_{\mathcal{H}} - \|\sqrt{t_0}\dot{u}_n(t_0)\|^2_{\mathcal{H}} \\
&= \lim_{t_0\to 0+} \int_{t_0}^t \|\dot{u}_n(s)\|^2_{\mathcal{H}} + 2(\sqrt{s}\ddot{u}_n(s), \sqrt{s}\dot{u}_n(s))_{\mathcal{H}}\,\mathrm{d}s \\
&= \int_0^t \|\dot{u}_n(s)\|^2_{\mathcal{H}} + 2(s\ddot{u}_n(s), \dot{u}_n(s))_{\mathcal{H}}\,\mathrm{d}s \\
&\overset{(5.19)}{=} \int_0^t 2(\dot{v}_n(s), \dot{u}_n(s))_{\mathcal{H}} - \|\dot{u}_n(s)\|^2_{\mathcal{H}}\,\mathrm{d}s \\
&\overset{(3.8)}{\leq} \int_0^t \|\dot{v}_n(s)\|^2_{\mathcal{H}}\,\mathrm{d}s \overset{(5.18)}{\leq} C_{\mathcal{A}_1}\Big(2 + 4\frac{C^2_{\mathcal{A}_2}}{\mu_{\mathcal{A}}}t\Big)\|u_{0,n}\|^2_{\mathcal{V}}
\end{aligned} \tag{5.20}$$

holds. With the same steps, estimate (5.20) shows that $\{\sqrt{t}\dot{u}_n(t)\}_{n\in\mathbb{N}} \subset C([0,T],\mathcal{H})$ is a Cauchy sequence. Its limit is given by $\sqrt{t}\dot{u}$, since $\dot{u}_n$ converges to $\dot{u}$ in $L^2(0,T;\mathcal{H})$ as $n \to \infty$ by Theorem 4.25. The estimate for $\sqrt{t}\mathcal{A}e^{-t\mathcal{A}}u_0 = \sqrt{t}\mathcal{A}u(t) \in \mathcal{H}^*$ follows then by the limit of (5.20) for $n \to \infty$ and $\sqrt{t}\mathcal{A}u(t) = -\sqrt{t}\dot{u}(t) \in C([0,T],\mathcal{H}^*)$ by (5.14).

Estimates for $\phi_k(-t\mathcal{A})$, $k \geq 1$: Let $t > 0$ be fixed. We consider $u(t) = \phi_k(-t\mathcal{A})f_k$ as the solution of the operator ODE (5.16) at $s = t$. By [Emm04, Th. 8.5.1] the derivative of u satisfies the operator ODE (4.11) with right-hand side $\frac{\mathrm{d}}{\mathrm{d}t}F_k \in L^2(0,T;\mathcal{H}^*)$ and initial value $\dot{u}_0 = F_k(0) \in \mathcal{H}^* \cong \mathcal{H}$. Therefore, the estimate

$$\|\mathcal{A}u(t)\|_{\mathcal{H}^*} \overset{(5.16)}{=} \|\dot{u}(t)\|_{\mathcal{H}} + \|F_k(t)\|_{\mathcal{H}^*} \overset{(4.22)}{\leq} \|F_k(0)\|_{\mathcal{H}^*} + \int_0^t \|\dot{F}_k(s)\|_{\mathcal{H}^*}\,\mathrm{d}s + \|F_k(t)\|_{\mathcal{H}^*} = \frac{1}{t}\frac{2\|f_k\|_{\mathcal{H}^*}}{(k-1)!}$$

holds. This finishes the proof. □

Part B.

Solutions of Operator Differential-Algebraic Equations

In this part we analyze the existence and uniqueness of solutions of operator DAEs as well as their regularity. Here, we focus on semi-explicit systems of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u - \mathcal{B}^*\lambda = f \qquad \text{in } \mathcal{V}^*, \tag{B.1a}$$

$$\mathcal{B}u = g \qquad \text{in } \mathcal{Q}^* \tag{B.1b}$$

with solutions in the sense of Definition 4.27. As mentioned in Section 4.4, operator DAEs of the form (B.1) are used to describe the weak formulation of constrained PDEs in an abstract fashion. This type of equation appears among others in the field of fluid dynamics, thermodynamics, electrodynamics, and chemical kinetics; see the examples in Chapter 6 and 7.

In Chapter 6 we investigate operator DAEs with time-independent operators $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{M}$. We mainly extend known results, e.g., existence, uniqueness, and regularity of solutions, to systems with more general right-hand sides. These results are used for the temporal discretization of (B.1) in Part C. For the proofs of the extension results we use the continuity of the solution with respect to the data. Furthermore, we specify the assumptions on the operators and on the right-hand sides as well as characterize consistent initial values. These assumptions are transferred to systems with time-dependent operators in Chapter 7. For the study of such operator DAEs we semi-discretize the system in time. We use the implicit Euler scheme to get existence, uniqueness, and regularity results by considering sequences of stationary solutions. Thereby, we investigate Gelfand triples where the pivot space has a time-dependent inner product. We analyze the effect of this nonconstant inner product on the embedding of generalized differentiable functions into the space of continuous functions. Furthermore, time-dependent splittings of Hilbert spaces are studied as well.

Theorem 6.15, which can be found as Theorem 2.7 in [AltZ20], was originally proved by the author of this thesis. All remaining results are unpublished work.

6. Systems with Time-Independent Operators

Semi-explicit operator DAEs with time-independent operators of the form

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u(t)) + \mathcal{A}u(t) - \mathcal{B}^*\lambda(t) = f(t, u(t)) \qquad \text{in } \mathcal{V}^*, \tag{6.1a}$$

$$\mathcal{B}u(t) = g(t) \qquad \text{in } \mathcal{Q}^*, \tag{6.1b}$$

with the initial condition

$$u(0) = u_0 \tag{6.1c}$$

are well-studied in the special case where the right hand side f is independent of u; see the references in Section 4.4. In this chapter we summarize and extend these results. In particular, we investigate the existence and uniqueness of solutions as well as their regularity and their continuous dependence on the data.

We assume that $\mathcal{V}$ and $\mathcal{Q}$ are real separable Hilbert spaces. The operators in (6.1) are all time-independent, linear, and continuous, i.e.,

$$\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*), \qquad \mathcal{M} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*), \qquad \mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*).$$

Here, $\mathcal{H}$ denotes an additional real Hilbert space such that $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ form a Gelfand triple. Since $\mathcal{B}$ is continuous, its kernel is a closed subspace of $\mathcal{V}$. We denote this subspace and its closure in $\mathcal{H}$ by

$$\mathcal{V}_{\ker} := \ker \mathcal{B} \quad \text{and} \quad \mathcal{H}_{\ker} := \operatorname{clos}_{\|\cdot\|_{\mathcal{H}}} \mathcal{V}_{\ker},$$

respectively. By the assumption on $\mathcal{V}$ and $\mathcal{H}$ these spaces then form another Gelfand triple

$$\mathcal{V}_{\ker} \hookrightarrow \mathcal{H}_{\ker} \cong \mathcal{H}_{\ker}^* \hookrightarrow \mathcal{V}_{\ker}^*.$$

Example 6.1 (Unsteady Stokes Equations)**.** The weak formulation of the linear unsteady Stokes equations with homogeneous Dirichlet boundary conditions can be written as an operator DAE of the form (6.1). The unsteady Stokes equation characterizes the evolution of a Newtonian fluid and is the linearization of incompressible Navier–Stokes equations around a vanishing velocity field [Tem77, Ch. III, § 1]. The state u describes the velocity field of the fluid, whereas λ relates to a relative pressure, which is assumed to have zero mean. For the application of the Stokes equation, we consider the Hilbert spaces

$$\mathcal{V} := [H_0^1(\Omega)]^d, \qquad \mathcal{H} := [L^2(\Omega)]^d, \qquad \mathcal{Q} := L^2(\Omega)/\mathbb{R} := \big\{p \in L^2(\Omega) \mid \textstyle\int_\Omega p \,\mathrm{d}\xi = 0\big\}.$$

Here, $\Omega \subset \mathbb{R}^d$ denotes a bounded computational domain with Lipschitz boundary.

The operator $\mathcal{A}\colon \mathcal{V} \to \mathcal{V}^*$ corresponds to the Laplace operator in the weak formulation and is defined by

$$\langle \mathcal{A}u, v\rangle := \mu \sum_{i=1}^d \int_\Omega \nabla u_i \cdot \nabla v_i \,\mathrm{d}\xi$$

with the dynamic viscosity $\mu \in \mathbb{R}_{>0}$. The operator $\mathcal{B}\colon \mathcal{V} \to \mathcal{Q}^*$ denotes the inf-sup stable divergence operator [Bra07, Ch. III, Rem. 6.5]. Its dual $\mathcal{B}^*\colon \mathcal{Q} \to \mathcal{V}^*$ is the gradient operator. The operator $\mathcal{M}$ is induced by the componentwise inner product of $L^2(\Omega)$ multiplied by the constant density $\rho > 0$.

The space $\mathcal{V}_{\text{ker}}$ are the divergence-free functions of $[H_0^1(\Omega)]^d$. Its closure $\mathcal{H}_{\text{ker}}$ is the subset of functions in $[L^2(\Omega)]^d$ whose trace in normal direction and whose divergence vanish in the distributional sense [Tem77, Ch. 1, Th 1.4].

Example 6.2 (Heat Equation with Boundary Control)**.** The constraint (6.1b) may also be used for boundary control [HinPU+09]. As a prototype we consider the heat equation (4.13a) with inhomogeneous Dirichlet boundary conditions. The spaces then are given by $\mathcal{V} := H^1(\Omega)$, $\mathcal{H} := L^2(\Omega)$, and $\mathcal{Q} := H^{-1/2}(\Omega)$ with a Lipschitz domain $\Omega \subset \mathbb{R}^d$. The operator $\mathcal{B}$ is the trace operator, $\mathcal{A}$ is the weak Laplacian as in Subsection 4.3.2.2, and $\mathcal{M}$ is induced by the inner product of $L^2(\Omega)$. Since $C_c^\infty(\Omega)$ is dense in $\mathcal{H} = L^2(\Omega)$, see [AdaF03, Cor. 2.39], the closure of $\mathcal{V}_{\text{ker}} = H_0^1(\Omega)$ equals $\mathcal{H}$ itself.

In Section 6.1 we specify the assumptions on the operators and define consistent initial values. Furthermore, we prove the existence and uniqueness of solutions of operator DAE (6.1) with right-hand sides, which are independent of the solution. The regularity of these solutions is topic of Section 6.2. In Section 6.3 we discuss whether and in which sense a controlled operator DAE implies a dissipation inequality. Finally, in 6.4 we extend the results on the existence and uniqueness to semi-linear systems, where f depends on time and on the solution u itself.

6.1. Existence and Uniqueness

In this section we study the solvability of the operator DAE (6.1) with a right-hand side $f = f(t)$, which is not dependent on the state u. We assume that $\mathcal{B}$ satisfies an inf-sup condition of the form (3.2). By Lemma 3.6 the operator $\mathcal{B}$ then has a right inverse $\mathcal{B}^-_{\mathcal{V}_c}$ for every closed complement $\mathcal{V}_c$ of $\mathcal{V}_{\text{ker}}$ in $\mathcal{V}$. In [AltH18, Th. 3.4 & Cor. 3.5] the authors prove under some additional assumptions that for a fixed $\mathcal{V}_c$ the condition

$$u_0 \in \mathcal{H}_{\text{ker}} + \mathcal{B}^-_{\mathcal{V}_c} g(0) := \{h_{\text{ker}} + \mathcal{B}^-_{\mathcal{V}_c} g(0) \mid h_{\text{ker}} \in \mathcal{H}_{\text{ker}}\} \subset \mathcal{H} \tag{6.2}$$

on the initial value u_0 and right-hand side g is necessary and sufficient for the existence of a solution. For a given right-hand side g we state the following lemma for the initial value u_0.

Lemma 6.3. *Let $\mathcal{V}_c^1$ and $\mathcal{V}_c^2$ be complements of $\mathcal{V}_{\text{ker}}$ in $\mathcal{V}$ and $g_0 \in \mathcal{Q}^*$. Suppose that $h_1 + \mathcal{B}^-_{\mathcal{V}_c^1} g_0 = h_2 + \mathcal{B}^-_{\mathcal{V}_c^2} g_0$ is satisfied. Then $h_1 \in \mathcal{H}_{\text{ker}}$ if and only if $h_2 \in \mathcal{H}_{\text{ker}}$.*

Proof. Since $h_1 - h_2 = \mathcal{B}^-_{\mathcal{V}_c^2} g_0 - \mathcal{B}^-_{\mathcal{V}_c^1} g_0 \in \mathcal{V}$ with $\mathcal{B}(h_1 - h_2) = g_0 - g_0 = 0$, we obtain $h_1 - h_2 \in \mathcal{V}_{\text{ker}} \hookrightarrow \mathcal{H}_{\text{ker}}$. □

Given the right-hand side g of (6.1b), Lemma 6.3 implies that the specific choice of $\mathcal{V}_c$ is irrelevant in (6.2). Therefore, we do not fix the complement $\mathcal{V}_c$ in (6.2) and write in the following

$$u_0 \in \mathcal{H}_{\text{ker}} + \mathcal{B}^- g(0) \subset \mathcal{H}. \tag{6.3}$$

Following [EmmM13] we call an initial value u_0 *consistent* with respect to the operator DAE (6.1) if u_0 satisfies (6.3). If in addition $u_0 \in \mathcal{V}$ holds, then (6.3) should be understood as $u_0 \in \mathcal{V}_{\text{ker}} + \mathcal{B}^- g(0)$ and therefore $\mathcal{B}u_0 = g(0)$.

Remark 6.4. Since the spaces $\mathcal{H}$ and $\mathcal{H}_{\text{ker}}$ in Example 6.2 are equal, every initial value $u_0 \in \mathcal{H}$ is consistent for the boundary controlled heat equation.

Remark 6.5. With Lemma 9.22 a consistent initial value u_0 can be characterized as an element of the closure of $\mathcal{V}$ with respect to the norm $(\|\cdot\|_{\mathcal{H}}^2 + \|\mathcal{B}\cdot\|_{\mathcal{Q}^*}^2)^{1/2}$. The initial value then satisfies $\overline{\mathcal{B}}u_0 = g(0)$ where $\overline{\mathcal{B}}$ is the extension of $\mathcal{B}$ to this closure of $\mathcal{V}$; see Subsection 9.3.2.2.

For the assumptions on the operators $\mathcal{B}$ and $\mathcal{M}$ let us recall the finite-dimensional DAE (2.3). To be uniquely solvable it is necessary that the matrix B in (2.3) has full (column) rank; see Lemma 2.1. This is equivalent to matrix B fulfilling an inf-sup stability condition of the form (3.2), since $\beta > 0$ is a lower bound for the smallest singular value [BreF91, Prop. II.3.1]. Hence, it is natural to assume that the operator $\mathcal{B}$ in the operator DAE (6.1) satisfies an inf-sup condition. Furthermore, if B has full rank in finite-dimensional DAE (2.3), then a sufficient condition for a unique solution of (2.3) is by Lemma 2.1 that M is symmetric positive definite. Therefore, we assume for the operator DAE (6.1) that $\mathcal{M} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ is self-adjoint and elliptic.

Remark 6.6. If $\mathcal{M} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ is self-adjoint and elliptic, then the bilinear form $\langle \mathcal{M}\,\cdot, \cdot\rangle_{\mathcal{H}^*,\mathcal{H}}$ is an inner product in $\mathcal{H}$. Its induced norm $\|\cdot\|_{\mathcal{M}}$ is equivalent to $\|\cdot\|_{\mathcal{H}}$. Therefore, $\mathcal{H}$ equipped with the inner product $\langle \mathcal{M}\,\cdot, \cdot\rangle_{\mathcal{H}^*,\mathcal{H}}$ is a Hilbert space denoted by $(\mathcal{H}, \|\cdot\|_{\mathcal{M}})$. The three spaces $\mathcal{V}, (\mathcal{H}, \|\cdot\|_{\mathcal{M}}), \mathcal{V}^*$ then form a Gelfand triple. The Riesz isomorphism in $(\mathcal{H}, \|\cdot\|_{\mathcal{M}})$ is the operator $\mathcal{M}$; see Theorem 3.3.

The existence of a right inverse of $\mathcal{B}$ by Lemma 3.6 implies that the dynamics of u in the complement $\mathcal{V}_{\mathrm{c}}$ is completely given by the constraint (6.1b). The remaining part of u, which maps into $\mathcal{V}_{\ker}$, has to be determined by the differential equation (6.1a). By the Theorem of Lions-Tartar 4.22 the operator $\mathcal{A}$ must satisfy a Gårding inequality (3.6) on $\mathcal{V}_{\ker}$. With these assumptions on the operators we can prove the existence and uniqueness of a solution of (6.1). To do so, we assume that the right-hand sides f and g satisfy assumptions adapted from Theorem 4.22.

Theorem 6.7 (Solutions of Operator DAEs)**.** *Suppose that $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ satisfies a Gårding inequality on $\mathcal{V}_{\ker}$, $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable, and $\mathcal{M} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ is self-adjoint and elliptic. Let $f = f_1 + f_2$ with $f_1 \in L^2(0,T;\mathcal{V}^*)$ and $f_2 \in L^1(0,T;\mathcal{H}^*)$ and $g \in W^{1,1}(0,T;\mathcal{Q}^*)$ be given. Assume $u_0 \in \mathcal{H}_{\ker} + \mathcal{B}^- g(0)$. Then the operator DAE (6.1) has a unique solution (u, λ), which satisfies*

$$u \in L^2(0,T;\mathcal{V}) \cap C([0,T],\mathcal{H}), \qquad \lambda = \tfrac{\mathrm{d}}{\mathrm{d}t}\Lambda \text{ for a } \Lambda \in C([0,T],\mathcal{Q}) \text{ with } \Lambda(0) = 0,$$
$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) \in L^2(0,T;\mathcal{V}_{\ker}^*) + L^1(0,T;\mathcal{H}_{\ker}^*), \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - \mathcal{B}^*\lambda \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*),$$

and the initial condition $u(0) = u_0$. The mapping of the data (f, g, u_0) to (u, Λ) is linear and continuous.

Proof. By Remark 6.6 we can assume without loss of generality that $\mathcal{M} = \mathcal{R}_{\mathcal{H}}$. Let $u_{\mathrm{c}} := \mathcal{B}^- g \in W^{1,1}(0,T;\mathcal{V})$, cf. [EmmM13, p. 463], and $u_{\ker} := u - u_{\mathrm{c}}$. Then (u, λ) is a solution of the operator DAE (6.1) if and only if $(u_{\ker}, \lambda)$ is a solution of the operator DAE (6.1) with right-hand sides $f_{\ker} = f - \mathcal{A}u_{\mathrm{c}} - \dot{u}_{\mathrm{c}}$, $g_{\ker} = 0$ and initial value $u_{\ker,0} := u_0 - \mathcal{B}^- g(0)$. Theorem 3.3 from [EmmM13] proves the existence, the uniqueness, and the continuous dependence on the data $(f_{\ker}, 0, u_{\ker,0})$ of $(u_{\ker}, \Lambda)$. The assertion of this theorem follows then by $u = u_{\mathrm{c}} + u_{\ker}$. □

6.2. Regularity of Solutions

Theorem 6.7 shows the existence of a unique solution of the operator DAE (6.1), where the Lagrange multiplier λ exists only in a distributional sense. For a regular λ, Lemma 3.6 and (6.1a) imply that the existence of a Bochner-integrable Lagrange multiplier is equivalent to $\mathcal{M}u$ having a generalized derivative with images in $\mathcal{V}^*$; cf. [Zim15, Sec. 3.1.2]. Note that the norm of $\mathcal{V}_{\ker}^*$ is weaker than the norm of $\mathcal{V}^*$.

In this section we investigate conditions such that $\dot u \in L^2(0,T;\mathcal{H})$, which implies $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) = \mathcal{M}\dot u \in L^2(0,T;\mathcal{H}^*)$. For this we assume that the right-hand sides f and g are more regular as stated in Theorem 6.7 or that $\mathcal{A}$ can be split similarly to the assumption in Theorem 4.25. We discuss these two approaches in the following two paragraphs. In both cases we need that the initial value u_0 is an element of $\mathcal{V}$ and satisfies the consistency condition $\mathcal{B}u_0 = g(0)$.

Regular Right-Hand Sides As a first approach we formally differentiate the operator DAE (6.1) and consider the needed assumptions on the data for a solution. Obviously, the right-hand sides must be one times more differentiable than stated in Theorem 6.7. The possible initial value w_0 of $\dot u$ has to satisfy $w_0 \in \mathcal{H}_{\mathrm{ker}} + \mathcal{B}^-\dot g(0)$ by the algebraic constraint (6.1b) and

$$\langle \mathcal{M}w_0, v_{\mathrm{ker}} \rangle_{\mathcal{H}^*,\mathcal{H}} = \langle f(0) - \mathcal{A}u_0, v_{\mathrm{ker}} \rangle_{\mathcal{V}^*,\mathcal{V}} \tag{6.4}$$

for all $v_{\mathrm{ker}} \in \mathcal{V}_{\mathrm{ker}}$ by the differential equation (6.1a); cf. [Wlo87, Sec. 27] and [Zim15, Sec. 3.1.2.1].

Theorem 6.8 (Regular Solutions of Operator DAEs I)**.** *Let the assumptions of Theorem 6.7 on the operators be satisfied. Assume that $f \in H^1(0,T;\mathcal{V}^*) + W^{1,1}(0,T;\mathcal{H}^*)$, $g \in W^{2,1}(0,T;\mathcal{Q}^*)$, and $u_0 \in \mathcal{V}$ with $\mathcal{B}u_0 = g(0)$. Suppose that a $w_0 \in \mathcal{H}_{\mathrm{ker}} + \mathcal{B}^-\dot g(0)$ exists, which satisfies (6.4). Then the solution of the operator DAE (6.1) satisfies*

$$u \in H^1(0,T;\mathcal{V}) \cap C^1([0,T],\mathcal{H}), \qquad \lambda \in C([0,T],\mathcal{Q}).$$

The solution (u,λ) depends linearly and continuously on the data (f,g,u_0,w_0).

Proof. Let (u,λ) and (w,μ) be the solution of (6.1) with data (f,g,u_0) and $(\dot f,\dot g,w_0)$, respectively. We prove $\dot u = w$. For this we use an arbitrary direct sum $\mathcal{V} = \mathcal{V}_{\mathrm{ker}} \oplus \mathcal{V}_{\mathrm{c}}$ with associated splittings $u = u_{\mathrm{ker}} + u_{\mathrm{c}}$ and $w = w_{\mathrm{ker}} + w_{\mathrm{c}}$. We note $\dot u_{\mathrm{c}} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}} g) = \mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\dot g = w_{\mathrm{c}}$. For the part in $\mathcal{V}_{\mathrm{ker}}$ we consider the function

$$v_{\mathrm{ker}}(t) := u_{\mathrm{ker}}(0) + \int_0^t w_{\mathrm{ker}}(s)\,\mathrm{d}s - u_{\mathrm{ker}}(t). \tag{6.5}$$

Then v_{ker} vanishes at the initial time-point $t=0$ and satisfies the operator ODE

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}v_{\mathrm{ker}}) + \mathcal{A}v_{\mathrm{ker}} &\overset{(6.5)}{=} \mathcal{M}w_{\mathrm{ker}} - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{ker}}) + \mathcal{A}\Big(u_{\mathrm{ker}}(0) + \int_0^t w_{\mathrm{ker}}\,\mathrm{d}s - u_{\mathrm{ker}}\Big) \\
&\overset{(6.1a)}{=} \mathcal{A}u_{\mathrm{ker}}(0) + \mathcal{M}w_{\mathrm{ker},0} + \int_0^t \dot f - \mathcal{A}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\dot g - \mathcal{M}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\ddot g\,\mathrm{d}s - f + \mathcal{A}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}} g + \mathcal{M}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\dot g \\
&= \mathcal{A}u_{\mathrm{ker}}(0) + \mathcal{M}w_{\mathrm{ker},0} - f(0) + \mathcal{A}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}} g(0) + \mathcal{M}\mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\dot g(0) \overset{(6.4)}{=} 0
\end{aligned}$$

in $\mathcal{V}^*_{\mathrm{ker}}$. This equivalent to v_{ker} fulfilling the operator DAE (6.1) with homogeneous data. So, $v_{\mathrm{ker}} = 0$ by Theorem 6.7. Theorem 3.36 then implies $\dot u_{\mathrm{ker}} = w_{\mathrm{ker}}$ in $\mathcal{V}^*$ and therefore $\dot u = w$. The continuous dependence of $u = u_0 + \int_0^\cdot w(s)\,\mathrm{d}s$ on the data follows by the continuity of $(\dot f,\dot g,w_0) \mapsto w$; see Theorem 6.7. The assertion on λ follows by

$$\lambda \overset{(6.1a)}{=} \mathcal{B}^{-*}_{\mathrm{left}}\big(f - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - \mathcal{A}u\big) = \mathcal{B}^{-*}_{\mathrm{left}}(f - \mathcal{M}w - \mathcal{A}u). \qquad \square$$

Operator $\mathcal{A}$ with a Special Splitting As a second approach for a more regular solution of (6.1) we adapt Theorem 4.25 to operator DAEs. For this we assume that $\mathcal{A}$ is the sum of an elliptic, self-adjoint operator and an operator with codomain $\mathcal{H}^*$. Since the part u_{c} is determined by $\mathcal{B}^-g$ and only u_{ker} by an operator ODE in $\mathcal{V}^*_{\mathrm{ker}}$, see e.g. [EmmM13, Th. 3.3], we can restrict the assumption on $\mathcal{A}$ to the space $\mathcal{V}_{\mathrm{ker}}$.

Theorem 6.9 (Regular Solutions of Operator DAEs II; [Tar06, Ch. 21] & [Zim15, Sec. 3.1.2.2]). *In addition to the assumptions of Theorem 6.7 on the operators, let $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ be self-adjoint and elliptic on $\mathcal{V}_{\text{ker}}$ and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V}, \mathcal{H}^*)$. Assume that $f \in L^2(0,T;\mathcal{H}^*)$, $g \in H^1(0,T;\mathcal{Q}^*)$, and $u_0 \in \mathcal{V}$ with $\mathcal{B}u_0 = g(0)$. Then the operator DAE (6.1) with initial value u_0 has a unique solution*

$$u \in C([0,T],\mathcal{V}) \cap H^1(0,T;\mathcal{H}), \qquad \lambda \in L^2(0,T;\mathcal{Q})$$

with $u(0) = u_0$. The solution depends linearly and continuously on the data.

Remark 6.10. The assumptions of Theorem 6.9 can be weakened in the sense that $f \in L^2(0,T;\mathcal{H}^*) + W^{1,1}(0,T;\mathcal{V}^*)$ and an operator $\mathcal{A}_3 \in \mathscr{L}(\mathcal{H}, \mathcal{V}^*)$ exists such that $\mathcal{A} = \sum_{i=1}^{3} \mathcal{A}_i$. The operators $\mathcal{A}_1$ and $\mathcal{A}_2$ still satisfy the assumption of Theorem 6.9. The spaces of the unique solution u and λ stay the same and the mapping $(f, g, u_0) \mapsto (u, \lambda)$ is linear and continuous. For an associated proof, one considers two sequences $\{f_n\}_{n\in\mathbb{N}} \subset H^1(0,T;\mathcal{H}^*)$ and $\{u_{0,n}\} \subset \mathcal{V}$. The sequences are constructed such that $f_n \to f$ in $L^2(0,T;\mathcal{H}^*) + W^{1,1}(0,T;\mathcal{V}^*)$, cf. Remark 3.33, as well as $u_{0,n} \to u_0$ in $\mathcal{V}$, $\mathcal{B}u_{0,n} = g(0)$, and $\mathcal{A}u_{0,n} \in \mathcal{H}^*_{\text{ker}}$; cf. Lemma 6.11. Note that the statements i) $\mathcal{A}v \in \mathcal{H}^*_{\text{ker}}$, ii) $(\kappa\,\text{id} + \mathcal{A})v \in \mathcal{H}^*_{\text{ker}}$ for all $\kappa \in \mathbb{R}$, and iii) $(\kappa\,\text{id} + \mathcal{A})v \in \mathcal{H}^*_{\text{ker}}$ for a $\kappa \in \mathbb{R}$ are equivalent. The assertion follows then by the limit behavior of the sequence of solutions (u_n, λ_n) and Theorem 6.8, 6.9, and Lemma 3.1.

6.3. Dissipation Inequality

In this section we analyze the operator DAE (6.1) in a port-Hamiltonian setting similar to the one we introduced for descriptor systems in Section 2.2. Here, the control is given by the functions $w_i\colon [0,T] \to \mathcal{W}_i$ with reflexive Banach spaces $\mathcal{W}_i$, $i = 1, 2$. Given the operators $\mathcal{D}_1 \in \mathscr{L}(\mathcal{W}_1, \mathcal{V}^*)$ and $\mathcal{D}_2 \in \mathscr{L}(\mathcal{W}_2, \mathcal{Q}^*)$ we consider the controlled operator DAE

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) + (\mathcal{J} + \mathcal{R})u - \mathcal{B}^*\lambda &= \mathcal{D}_1\omega_1 && \text{in } \mathcal{V}^*, && (6.6a)\\ \mathcal{B}u &= \mathcal{D}_2\omega_2 && \text{in } \mathcal{Q}^*, && (6.6b)\\ \mathcal{D}_1^* u &= y_1 && \text{in } \mathcal{W}_1^*, && (6.6c)\\ \mathcal{D}_2^*\lambda &= y_2 && \text{in } \mathcal{W}_2^*. && (6.6d) \end{aligned}$$

The additional functions $y_1\colon [0,T] \to \mathcal{V}$ and $y_2\colon [0,T] \to \mathcal{Q}$ are the system's outputs. We assume that the operator $\mathcal{J} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is skew-adjoint and that $\mathcal{R} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is self-adjoint, elliptic on $\mathcal{V}_{\text{ker}}$, and satisfies $\langle \mathcal{R}v, v \rangle \geq 0$ for all $v \in \mathcal{V}$.

In the setting of the operator DAE (6.1) we have $\mathcal{A} = \mathcal{J} + \mathcal{R}$, $f = \mathcal{D}_1\omega_1$, and $g = \mathcal{D}_2\omega_2$. By Theorem 6.7 the controlled operator DAE (6.6) then has a unique solution (u, λ) for every $\omega_1 \in L^2(0,T;\mathcal{W}_1)$, $\omega_2 \in H^1(0,T;\mathcal{W}_2)$ and consistent initial value $u_0 \in \mathcal{H}_{\text{ker}} + \mathcal{B}^-\mathcal{D}_2\omega_2(0)$. For this solution we want to study a dissipation inequality similar to (2.6) with respect to the Hamiltonian defined by $(u, \lambda) \mapsto \frac{1}{2}\|u\|^2_{\mathcal{M}}$. In the case of the operator DAE (6.1), the right-hand side of the dissipation inequality (2.6) is not well-defined, since λ and therefore also y_2 exist only in a distributional sense. To interpret the integral of $\langle y_2, w_2 \rangle$ in (2.6) we consider a sequence of more regular solutions. To do so, we need the following lemma.

Lemma 6.11. *Assume that $\omega_1 \in L^2(0,T;\mathcal{W}_1)$, $\omega_2 \in H^1(0,T;\mathcal{W}_2)$, and $u_0 \in \mathcal{H}_{\text{ker}} + \mathcal{B}^-\mathcal{D}_2\omega_2(0) \subset \mathcal{H}$ are given. Let $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ be elliptic on $\mathcal{V}_{\text{ker}}$.*

Then there exist sequences $\{\omega_{1,n}\}_{n\in\mathbb{N}} \subset C_c^\infty([0,T],\mathcal{W}_1)$, $\{\omega_{2,n}\}_{n\in\mathbb{N}} \subset C^\infty([0,T],\mathcal{W}_2)$ with $\omega_{2,n}(0) = \omega_2(0)$ and $\dot\omega_{2,n} \in C_c^\infty([0,T],\mathcal{W}_2)$, as well as $\{u_{0,n}\}_{n\in\mathbb{N}} \subset \mathcal{V}$ with $\mathcal{B}u_{0,n} = \mathcal{D}_2\omega_2(0)$, and $\mathcal{A}u_{0,n} \in \mathcal{H}^*_{\text{ker}}$, such that

$$\lim_{n\to\infty} (\omega_{1,n}, \omega_{2,n}, u_{0,n}) = (\omega_1, \omega_2, u_0) \quad \text{in } L^2(0,T;\mathcal{W}_1) \times H^1(0,T;\mathcal{W}_2) \times \mathcal{H}.$$

Proof. Since C_c^∞ is dense in L^2, cf. [GajGZ74, Ch. II, § 1, Lem. 1.20 & Ch. IV, § 1, Lem. 1.13], we find sequences $\{\omega_{1,n}\}_{n\in\mathbb{N}} \subset C_C^\infty([0,T],\mathcal{W}_1)$ and $\{\omega_{2,n}'\}_{n\in\mathbb{N}} \subset C_C^\infty([0,T],\mathcal{W}_2)$ with

$$\|\omega_1 - \omega_{1,n}\|_{L^2(0,T;\mathcal{W}_1)} < \tfrac{1}{n} \quad \text{and} \quad \|\dot\omega_2 - \omega_{2,n}'\|_{L^2(0,T;\mathcal{W}_2)} < \tfrac{1}{n}.$$

We set $\omega_{2,n}(t) := \omega_2(0) + \int_0^t \omega_{2,n}' \,\mathrm{d}s$. Then the inequality $\|\omega_2 - \omega_{2,n}\|_{H^1(0,T;\mathcal{W}_2)} < \frac{1}{n}\sqrt{1+\frac{T^2}{2}}$ holds.

For the construction of $u_{0,n}$, we define $u_{0,\mathrm{ker}} := u_0 - \mathcal{B}^-\mathcal{D}_2\omega_2(0) \in \mathcal{H}_{\mathrm{ker}}$. Since $\mathcal{V}_{\mathrm{ker}}$ is dense in $\mathcal{H}_{\mathrm{ker}}$ and $\mathcal{H}_{\mathrm{ker}}^*$ in $\mathcal{V}_{\mathrm{ker}}^*$, we find $u_{0,\mathrm{ker},n}' \in \mathcal{V}_{\mathrm{ker}}$ and $f_n \in \mathcal{H}_{\mathrm{ker}}^*$ with

$$\|u_{0,\mathrm{ker}} - u_{0,\mathrm{ker},n}'\|_{\mathcal{H}_{\mathrm{ker}}} < \tfrac{1}{n} \quad \text{and} \quad \|f_n - \mathcal{A}u_{0,\mathrm{ker},n}'\|_{\mathcal{V}_{\mathrm{ker}}^*} < \tfrac{1}{n}.$$

Finally, $u_{0,n} := \mathcal{A}|_{\mathcal{V}_{\mathrm{ker}}}^{-1} f_n + \mathcal{B}^-\mathcal{D}_2\omega_2(0) \in \mathcal{V}$ satisfies the identities $\mathcal{B}u_{0,n} = \mathcal{D}_2\omega_2(0)$, $\mathcal{A}u_{0,n} = f_n$ in $\mathcal{H}_{\mathrm{ker}}^*$, as well as (with $\mu_\mathcal{A}$ as the ellipticity constant of $\mathcal{A}|_{\mathcal{V}_{\mathrm{ker}}}$) the estimate

$$\begin{aligned}\|u_0 - u_{0,n}\|_\mathcal{H} &= \big\|u_{0,\mathrm{ker}} - \mathcal{A}|_{\mathcal{V}_{\mathrm{ker}}}^{-1} f_n\big\|_{\mathcal{H}_{\mathrm{ker}}} \\ &\le \|u_{0,\mathrm{ker}} - u_{0,\mathrm{ker},n}'\|_{\mathcal{H}_{\mathrm{ker}}} + \frac{C_{\mathcal{V}_{\mathrm{ker}}\hookrightarrow\mathcal{H}_{\mathrm{ker}}}}{\mu_\mathcal{A}}\|f_n - \mathcal{A}u_{0,\mathrm{ker},n}'\|_{\mathcal{V}_{\mathrm{ker}}^*} \\ &\le \Big(1 + \frac{C_{\mathcal{V}_{\mathrm{ker}}\hookrightarrow\mathcal{H}_{\mathrm{ker}}}}{\mu_\mathcal{A}}\Big)\frac{1}{n}.\end{aligned}$$

□

Lemma 6.12 (Dissipation Inequality). *Let $\mathcal{B}$ be inf-sup stable, $\mathcal{M} \in \mathscr{L}(\mathcal{H},\mathcal{H}^*)$ be self-adjoint as well as elliptic. Assume that $\mathcal{J} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ is skew-adjoint, i.e., $\mathcal{J} = -\mathcal{J}^*$, $\mathcal{R} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ with $\mathcal{R} = \mathcal{R}^*$ is elliptic on $\mathcal{V}_{\mathrm{ker}}$, and satisfies $\langle \mathcal{R}v, v\rangle \ge 0$ for all $v \in \mathcal{V}$. Suppose that $\omega_1 \in L^2(0,T;\mathcal{W}_1)$, $\omega_2 \in H^1(0,T;\mathcal{W}_2)$, and $u_0 \in \mathcal{H}_{\mathrm{ker}} + \mathcal{B}^-\mathcal{D}_2\omega_2(0)$ are given. Let (u,λ) be the solution of (6.6) given by Theorem 6.7 and $Y_2 := \mathcal{D}_2^*\Lambda$. Define*

$$\int_0^t \langle y_2, \omega_2\rangle \,\mathrm{d}s := -\int_0^t \langle Y_2, \dot\omega_2\rangle \,\mathrm{d}s + \langle Y_2(t), \omega_2(t)\rangle. \tag{6.7}$$

Then we have

$$\tfrac12\|u(t)\|_\mathcal{M}^2 - \tfrac12\|u_0\|_\mathcal{M}^2 = \int_0^t -\langle \mathcal{R}u, u\rangle + \langle y_1, \omega_1\rangle + \langle y_2, \omega_2\rangle \,\mathrm{d}s \le \int_0^t \langle y_1, \omega_1\rangle + \langle y_2, \omega_2\rangle \,\mathrm{d}s. \tag{6.8}$$

Proof. Consider the sequences $\{\omega_{1,n}\}_{n\in\mathbb{N}} \subset C_c^\infty([0,T],\mathcal{W}_1)$, $\{\omega_{2,n}\}_{n\in\mathbb{N}} \subset C^\infty([0,T],\mathcal{W}_2)$, and $\{u_{0,n}\}_{n\in\mathbb{N}} \subset \mathcal{V}$ as stated in Lemma 6.11 with $\mathcal{A} = \mathcal{J} + \mathcal{R}$. Then for every $n \in \mathbb{N}$ we have a solution $(u_n, \lambda_n, y_{1,n}, y_{2,n})$ of (6.6) with $\lambda_n \in C([0,T],\mathcal{Q})$ and $y_{2,n} \in C([0,T],\mathcal{W}_2)$ by Theorem 6.8. Since the solution of (6.6) is continuous with respect to the data, u_n converges to u in $L^2(0,T;\mathcal{V}) \cap C([0,T],\mathcal{H})$ and $\int_0^\cdot \lambda_n \,\mathrm{d}s$ to Λ in $C([0,T],\mathcal{Q})$. In particular, we have

$$y_{1,n} = \mathcal{D}_1^* u_n \to \mathcal{D}_1^* u = y_1 \quad \text{and} \quad Y_{2,n} = \mathcal{D}_2^* \int_0^\cdot \lambda_n(s)\,\mathrm{d}s \to \mathcal{D}_2^*\Lambda = Y_2$$

in $L^2(0,T;\mathcal{W}_1^*)$ and in $C([0,T],\mathcal{W}_2^*)$ as $n \to \infty$, respectively. Therefore, we get

$$\begin{aligned}\tfrac12\|u(t)\|_\mathcal{M}^2 - \tfrac12\|u_0\|_\mathcal{M}^2 &= \lim_{n\to\infty} \tfrac12\|u_n(t)\|_\mathcal{M}^2 - \tfrac12\|u_{0,n}\|_\mathcal{M}^2 \\ &\overset{(6.6a)}{=} \lim_{n\to\infty} \int_0^t \langle -\mathcal{R}u_n + \mathcal{B}^*\lambda_n + \mathcal{D}_1\omega_{1,n}, u_n\rangle \,\mathrm{d}s\end{aligned}$$

$$\overset{(6.6c)}{=} \lim_{n\to\infty} \int_0^t \langle -\mathcal{R}u_n, u_n \rangle + \langle y_{1,n}, \omega_{1,n} \rangle + \langle \mathcal{D}_2\omega_{2,n}, \lambda_n \rangle \,\mathrm{d}s$$
$$\overset{(6.6d)}{=} \lim_{n\to\infty} \int_0^t \langle -\mathcal{R}u_n, u_n \rangle + \langle y_{1,n}, \omega_{1,n} \rangle + \langle y_{2,n}, \omega_{2,n} \rangle \,\mathrm{d}s$$
$$= \lim_{n\to\infty} \int_0^t \langle -\mathcal{R}u_n, u_n \rangle + \langle y_{1,n}, \omega_{1,n} \rangle - \langle Y_{2,n}, \dot\omega_{2,n} \rangle \,\mathrm{d}s + \langle Y_{2,n}(t), \omega_{2,n}(t) \rangle$$
$$= \int_0^t \langle -\mathcal{R}u, u \rangle + \langle y_1, \omega_1 \rangle - \langle Y_2, \dot\omega_2 \rangle \,\mathrm{d}s + \langle Y_2(t), \omega_2(t) \rangle. \qquad \square$$

6.4. semi-linear Systems

This section is devoted to the extension of our analysis of the semi-explicit linear operator DAE (6.1) to semi-linear systems. The investigated system has still linear constraints, but a nonlinearity appears in the low-order terms of the dynamic equation. Thus, we consider the following semi-linear operator DAE: find $u\colon [0,T] \to \mathcal{V}$ and $\lambda\colon [0,T] \to \mathcal{Q}$ such that

$$\dot u(t) + \mathcal{A}u(t) - \mathcal{B}^*\lambda(t) = f(t,u) \qquad \text{in } \mathcal{V}^*, \tag{6.9a}$$
$$\mathcal{B}u(t) = g(t) \qquad \text{in } \mathcal{Q}^*. \tag{6.9b}$$

Example 6.13 (Dynamical Boundary Conditions)**.** The (weak) formulation of semi-linear parabolic equations with dynamical boundary conditions [SprW10] fits into the given framework. As a prototype consider

$$\dot u - \Delta u = f_\Omega \qquad \text{in } \Omega, \tag{6.10a}$$
$$\mu\dot u - \beta\Delta_\Gamma u + \partial_n u + \alpha u = f_\Gamma \qquad \text{on } \Gamma \subset \partial\Omega, \tag{6.10b}$$
$$u = 0 \qquad \text{on } \partial\Omega \setminus \Gamma \tag{6.10c}$$

with constants $\mu, \kappa > 0$, $\beta \ge 0$, the Laplace-Beltrami operator Δ_Γ, see [GilT01, Ch. 16.1], and the normal derivative ∂_n on Γ. Equation (6.10b) describes a heat source on the boundary Γ; see [Gol06]. Then the reformulation of (6.10) as a coupled system with the dummy variable $p = u|_\Gamma$ has in its weak form the operator DAE structure (6.9); cf. [Alt19]. We choose $\mathcal{H} := L^2(\Omega) \times L^2(\Gamma)$, $\mathcal{Q} := H_{00}^{-1/2}(\Gamma)$, and $\mathcal{V}$ depending on β as $\mathcal{V} := H^1(\Omega) \times H_{00}^{1/2}(\Gamma)$ for $\beta = 0$ or $\mathcal{V} := H^1(\Omega) \times H_0^1(\Gamma)$ if $\beta > 0$. The constraint is given by $0 = \mathcal{B}(u,p) = u|_\Gamma - p$, where $\mathcal{B}$ is inf-sup stable [Alt19, Lem. 2 & 5]. The operators $\mathcal{A}$ and $\mathcal{M}$ are the weak versions of the associated operators.

We emphasize that PDEs with dynamical boundary conditions may have nonlinear reaction terms also on the boundary. Examples include the Allen-Cahn equation [ColF15b; GalG08], the Cahn-Hillard equation [ColF15a; Gal07], and the Caginalp equation [ChiFP06].

Example 6.14 (semi-linear Reaction–Diffusion–Advection Equations)**.** The weak formulations of semi-linear reaction–diffusion–advection equations with boundary control can be modeled as semi-linear operator DAEs (6.9); cf. Example 6.2. Examples include the conduction of heat with chemical reactions [CarJ96, Sec. 1.6.II], the flow of electrons and holes in semiconductors [Van50, Sec. 2.2], and chemical reactions in a catalyst pellet [Gav68, Sec. 2.1]. For more examples see [Hen81, Ch. 2] and the references therein.

In this section we are interested in solutions u, which are continuous with images in $\mathcal{V}$. Following Theorem 6.9, we restrict the analysis of (6.9) to operator DAEs with an operator $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ where $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is self-adjoint and elliptic on $\mathcal{V}_{\ker}$ and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V}, \mathcal{H}^*)$. Furthermore, we choose $\mathcal{M}$ to be the Riesz isomorphism in $\mathcal{H}$. This can be extended to a general self-adjoint, elliptic operator $\mathcal{M} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ by interpreting $\mathcal{M}$ as the underlying inner product in $\mathcal{H}$. The operator $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable.

The nonlinearity f maps from $[0,T] \times \mathcal{V}$ to $\mathcal{H}^*$. In order to transfer the results of Theorem 6.9 to the semi-linear operator DAE (6.9) we need to consider the Nemytskiĭ map $\mathcal{N}_f$ induced by f, which maps abstract measurable functions $u\colon [0,T] \to \mathcal{V}$ to abstract measurable functions $f(\,\cdot\,, u(\,\cdot\,))\colon [0,T] \to \mathcal{H}^*$. For this, we need the classical Carathéodory conditions of Definition 4.1. Furthermore, we need a growth condition such that $\mathcal{N}_f$ maps $C([0,T], \mathcal{V})$ to $L^2(0,T;\mathcal{H}^*)$; cf. Theorem 6.9. We assume in the following that there exists a function $k \in L^2(0,T)$ such that

$$\|f(t,v)\|_{\mathcal{H}^*} \leq k(t)(1 + \|v\|_{\mathcal{V}}) \tag{6.11}$$

for all $v \in \mathcal{V}$ and almost all $t \in [0,T]$. We emphasize that (6.11) is sufficient but not necessary for f to induce a Nemytskiĭ map; cf. [GolKT92, Th. 1(ii)] and Remark 6.17. A last crucial point for the existence of a solution is that f is locally Lipschitz continuous. This means that for every $v \in \mathcal{V}$ an open ball $B_r(v) \subseteq \mathcal{V}$ with center v and radius $r = r(v) > 0$ as well as a constant $L = L(v) \geq 0$ exist, such that

$$\|f(t,v_1) - f(t,v_2)\|_{\mathcal{H}^*} \leq L\, \|v_1 - v_2\|_{\mathcal{V}} \tag{6.12}$$

for all $v_1, v_2 \in B_r(v)$ and almost every $t \in [0,T]$. We use these conditions to prove the existence and uniqueness of a global solution of (6.9).

Theorem 6.15 (Existence of Solutions for Semi-Linear Operator DAEs)**.** *Assume that $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ can be split into $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ self-adjoint and elliptic on $\mathcal{V}_{\text{ker}}$ and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V}, \mathcal{H}^*)$. Let $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ be inf-sup stable. Further, suppose that $g \in H^1(0,T;\mathcal{Q}^*)$ and that $f\colon [0,T] \times \mathcal{V} \to \mathcal{H}^*$ satisfies the Carathéodory conditions as in Definition 4.1, the growth condition* (6.11)*, and is locally Lipschitz continuous* (6.12)*. Then for every consistent initial value $u_0 \in \mathcal{V}$, i.e., $\mathcal{B}u_0 = g(0)$, the semi-linear operator DAE* (6.9) *has a unique solution*

$$u \in C([0,T], \mathcal{V}) \cap H^1(0,T;\mathcal{H}), \qquad \lambda \in L^2(0,T;\mathcal{Q})$$

with $u(0) = u_0$.

Proof. Without loss of generality, we assume that $\mathcal{A} = \mathcal{A}_1$ and $\mathcal{A}_2 = 0$. Otherwise, we redefine $f(t,v) \leftarrow f(t,v) - \mathcal{A}_2 v$, leading to an update of the involved constant $L \leftarrow L + C_{\mathcal{A}_2}$ and $k(t) \leftarrow k(t) + C_{\mathcal{A}_2}$ but leaving the radius r of the local Lipschitz condition (6.12) unchanged.

To prove the assertion, we follow the steps of [Paz83, Ch. 6.3]. Let $t' \in (0,T]$ be arbitrary but fixed. With (6.11) we note that the Nemyskii map induced by f maps $C([0,t'], \mathcal{V})$ into $L^2(0,t';\mathcal{H}^*)$; see Lemma 4.2. Therefore, the solution map $\mathcal{S}_{[0,t']}\colon C([0,t'], \mathcal{V}) \to C([0,t'], \mathcal{V})$, which maps $y \in C([0,t'], \mathcal{V})$ to the solution of

$$\dot{u}(t) + \mathcal{A}u(t) - \mathcal{B}^*\lambda(t) = f(t, y(t)) \qquad \text{in } \mathcal{V}^*, \tag{6.13a}$$

$$\mathcal{B}u(t) = g(t) \qquad \text{in } \mathcal{Q}^* \tag{6.13b}$$

with initial value u_0, is well-defined by Theorem 6.9. To find a solution to (6.9) we have to look for a fixed point of $\mathcal{S}_{[0,t']}$ and show that the interval of existence $[0,t']$ can be extended to $[0,T]$.

Let $\tilde{u} \in C([0,T], \mathcal{V})$ be the solution of the operator DAE (6.9) for $f \equiv 0$ and initial value u_0. With $r = r(u_0)$, $L = L(u_0)$, and the ellipticity constant $\mu_{\mathcal{A}}$ of $\mathcal{A}$ restricted to $\mathcal{V}_{\text{ker}}$ we choose $t_1 \in (0,T]$ such that

$$\|\tilde{u}(t) - u_0\|_{\mathcal{V}} \leq \frac{r}{2}, \tag{6.14a}$$

$$\int_0^t |k|^2 \,\mathrm{d}s \leq \frac{\mu_{\mathcal{A}} r^2}{4\,(1 + r + \|u_0\|_{\mathcal{V}})^2}, \tag{6.14b}$$

$$L^2 t_1 \ < \ \mu_{\mathcal{A}}, \tag{6.14c}$$

$$\int_0^t \tfrac{3}{\mu_{\mathcal{A}}}\, |k|^2 (1 + \|\tilde u\|^2_{\mathcal{V}}) \,\mathrm{d}s \ \le \ \frac{r^2}{4} \cdot \exp\big(-\tfrac{3}{\mu_{\mathcal{A}}} \int_0^t |k|^2 \,\mathrm{d}s \big) \tag{6.14d}$$

for all $t \in [0, t_1]$. This is well-defined, since $\tilde u - u_0$ and the integrals in (6.14b) and (6.14d) are continuous functions in t, which vanish for $t = 0$. We define

$$D := \big\{ y \in C([0,t_1], \mathcal{V}) \,|\, \|y - \tilde u\|_{C([0,t_1],\mathcal{V})} \le r/2 \big\}$$

and consider $y_1, y_2 \in D$. By (6.14a) we have $\|y_i - u_0\|_{C([0,t_1],\mathcal{V})} \le r$. Using that $\tilde u$ and $\mathcal{S}_{[0,t_1]} y_i$ satisfy the constraint (6.13b), we obtain the estimate

$$\begin{aligned} \mu_{\mathcal{A}} \,\|(\mathcal{S}_{[0,t_1]} y_i - \tilde u)(t)\|^2_{\mathcal{V}} &\overset{(4.26)}{\le} \int_0^t \|f(s, y_i(s))\|^2_{\mathcal{H}^*} \,\mathrm{d}s \overset{(6.11)}{\le} \int_0^t |k(s)|^2 \big(1 + \|y_i(s) - u_0\|_{\mathcal{V}} + \|u_0\|_{\mathcal{V}}\big)^2 \,\mathrm{d}s \\ &\le \ \big(1 + r + \|u_0\|_{\mathcal{V}}\big)^2 \int_0^t |k(s)|^2 \,\mathrm{d}s, \end{aligned} \tag{6.15}$$

which implies with (6.14b) that $\mathcal{S}_{[0,t_1]}$ maps D into itself. Further, we have

$$\mu_{\mathcal{A}} \,\|(\mathcal{S}_{[0,t_1]} y_1 - \mathcal{S}_{[0,t_1]} y_2)(t)\|^2_{\mathcal{V}} \overset{(4.26)}{\le} \int_0^t \|f(s, y_1(s)) - f(s, y_2(s))\|^2_{\mathcal{H}^*} \,\mathrm{d}s \overset{(6.12)}{\le} L^2 t_1 \|y_1 - y_2\|^2_{C([0,t_1],\mathcal{V})}$$

for all $t \le t_1$. This, together with the previous estimates (6.14c) and (6.15) shows that $\mathcal{S}_{[0,t_1]}$ is a contraction on D, i.e., a Lipschitz continuous function from D into itself with a Lipschitz constant smaller than one. Hence, there exists a unique fixed point $u \in D \subset C([0,t_1], \mathcal{V})$ of $\mathcal{S}_{[0,t_1]}$ by the Banach fixed point theorem [Zei86, Th. 1.A]. On the other hand, for every fixed point $u^\star = \mathcal{S}_{[0,t_1]} u^\star$ in $C([0,t_1], \mathcal{V})$, we have the estimate

$$\mu_{\mathcal{A}} \,\|(u^\star - \tilde u)(t)\|^2_{\mathcal{V}} = \mu_{\mathcal{A}} \,\|(\mathcal{S}_{[0,t_1]} u^\star - \tilde u)(t)\|^2_{\mathcal{V}} \le \int_0^t |k(s)|^2 \big(1 + \|(u^\star - \tilde u)(s)\|_{\mathcal{V}} + \|\tilde u(s)\|_{\mathcal{V}}\big)^2 \,\mathrm{d}s.$$

Using $(a+b+c)^2 \le 3\,(a^2 + b^2 + c^2)$ and Gronwall's inequality (3.10) it follows

$$\|(u^\star - \tilde u)(t)\|^2_{\mathcal{V}} \le \int_0^t \tfrac{3}{\mu_{\mathcal{A}}}\, |k(s)|^2 \big(1 + \|\tilde u(s)\|^2_{\mathcal{V}}\big) \,\mathrm{d}s \cdot \exp\Big(\tfrac{3}{\mu_{\mathcal{A}}} \int_0^t |k(s)|^2 \,\mathrm{d}s \Big) \tag{6.16}$$

for every $t \le t_1$. Because of (6.14d), this shows that $u^\star$ is an element of D and thus, $u^\star = u$.

By considering problem (6.9) iteratively from $[t_{i-1}, T]$, $t_0 := 0$, to $[t_i, T]$ with consistent initial value $u_0 = u(t_i)$, we can extend u uniquely on an interval $\mathbb{I}$ with $u \in C(\mathbb{I}; \mathcal{V})$ and $u = \mathcal{S}_{[0,t']} u$ for every $t' \in \mathbb{I}$. Note that either $\mathbb{I} = [0, T]$ or $\mathbb{I} = [0, T_{\mathrm{f}})$ with $T_{\mathrm{f}} \le T$. The second case is only possible if $\|u(t)\|_{\mathcal{V}}$ is unbounded near T_{f}, otherwise we can extend u to T_{f} and start at T_{f} again; see also [Zim15, Th. 3.20]. But, since the estimate (6.16) also holds for $u = u^\star$ and $t < T_{\mathrm{f}}$, we have that $\lim_{t \to T_{\mathrm{f}}^-} \|u(t)\|_{\mathcal{V}} \le \lim_{t \to T_{\mathrm{f}}^-} \|u(t) - \tilde u(t)\|_{\mathcal{V}} + \|\tilde u(t)\|_{\mathcal{V}}$ is bounded. Therefore, $u = \mathcal{S}_{[0,T]} u \in C([0,T], \mathcal{V})$. Finally, the stated spaces for u and λ follow by Theorem 6.9 with right-hand sides $f = f(\,\cdot\,, u(\,\cdot\,))$, g, and initial value $u_0 \in \mathcal{V}$. □

Remark 6.16. In the proof of Theorem 6.15 we follow the steps of [Paz83, Ch. 6.3]. The assumptions considered in [Paz83, Ch. 6.3], however, are stronger than the one in Theorem 6.15. If these additional assumptions are satisfied, then the existence and uniqueness of a solution to (6.9) follows directly by Theorem 6.9, [Paz83, Ch. 6, Th. 3.1 & 3.3], and the fact that every self-adjoint, elliptic operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ has a unique invertible square root $\mathcal{A}^{1/2} \in \mathscr{L}(\mathcal{V}, \mathcal{H})$ with

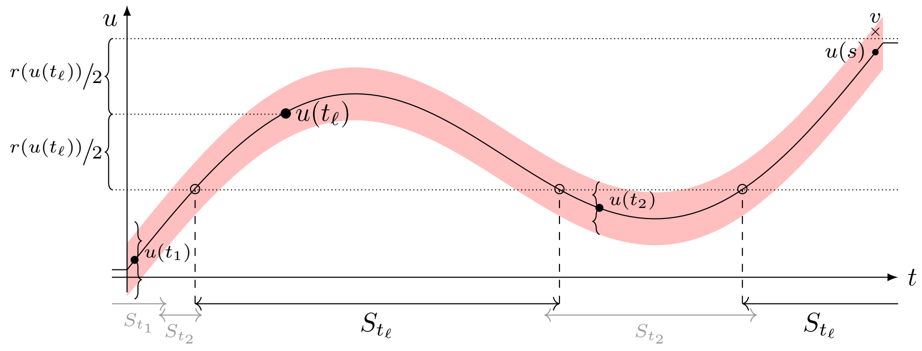

Figure 6.1.: Illustration of the proof of Lemma 6.18, where $\ell = 3$ and $r_u = r(u(t_2))$.

$\langle \mathcal{A}v_1, v_2\rangle = (\mathcal{A}^{1/2}v_1, \mathcal{A}^{1/2}v_2)_{\mathcal{H}}$ for all $v_1, v_2 \in \mathcal{V}$. This can be proven by interpreting $-\mathcal{A}$ as an (unbounded) operator $\mathrm{A}\colon D(\mathrm{A}) \subset \mathcal{H} \to \mathcal{H}$ with domain $D(\mathrm{A}) := \{\mathcal{A}^{-1}h \,|\, h \in \mathcal{H}^* \cong \mathcal{H}\} \subset \mathcal{V} \hookrightarrow \mathcal{H}$ and the results of [BirS87, Ch. 6, Th. 4 & Ch. 10, Th. 1] and [Paz83, Ch. 2, Th. 6.8].

Remark 6.17. It is possible to weaken the assumption (6.11) in Theorem 6.15 to $\|f(t,v)\|_{\mathcal{H}^*} \leq k(t)(1 + \|v\|_{\mathcal{V}}^p)$ for an arbitrary $p > 1$. Under this assumption one can show the existence of a unique solution of (6.9), which may only exist locally.

As a next step we show that the unique solution of (6.9) depends continuously on the initial value u_0 and on the right-hand side g. For the associated proof we need the following lemma.

Lemma 6.18. *Let f satisfy the assumptions of Theorem 6.15 and let $u \in C([0,T], \mathcal{V})$ be arbitrary. Then there exists a radius $r_u > 0$ and a Lipschitz constant $L_u \in [0, \infty)$, both depending on the function u, such that (6.12) holds for all $v_1, v_2 \in B_{r_u}(u(s))$ with $L = L_u$ and arbitrary $s \in [0,T]$.*

Proof. The proof is constructive and its main idea is illustrated in Figure 6.1.

We extend u to a function on $\mathbb{R}$ by setting $u(t)$ to $u(0)$ if $t < 0$ and to $u(T)$ if $t > T$. Furthermore, we define for every $t \in [0,T]$ the set

$$S_t := \left\{ s \in \mathbb{R} \,\middle|\, \|u(t) - u(s)\|_{\mathcal{V}} < \frac{r(u(t))}{2} \right\} \subseteq \mathbb{R}$$

with the local radius $r > 0$. Since $t \in S_t$ and S_t is open by the continuity of u, the set $\{S_t\}_{t\in[0,T]}$ is a open cover of the compact interval $[0,T]$. Thus, there exist finitely many time points $t_i \in [0,T]$, $i = 1, \ldots, N$, such that $\bigcup_{i=1}^N S_{t_i} \supseteq [0,T]$. We define

$$r_u := \min_{i=1,\ldots,N} \frac{r(u(t_i))}{2} \qquad \text{and} \qquad L_u := \max_{i=1,\ldots,N} L(u(t_i)).$$

Let now $s \in [0,T]$ be arbitrary. By the construction of the sets S_{t_i} there exists an $\ell \in \{1, \ldots, N\}$ with $s \in S_{t_\ell}$. Then, for every $v \in B_{r_u}(u(s))$ we have

$$\|v - u(t_\ell)\|_{\mathcal{V}} \leq \|v - u(s)\|_{\mathcal{V}} + \|u(s) - u(t_\ell)\|_{\mathcal{V}} < r_u + \frac{r(u(t_\ell))}{2} \leq r(u(t_\ell)),$$

which shows $B_{r_u}(u(s)) \subseteq B_{r(u(t_\ell))}(u(t_\ell))$. Finally, for every $v_1, v_2 \in B_{r_u}(u(s))$ this implies

$$\|f(t,v_1) - f(t,v_2)\|_{\mathcal{H}^*} \overset{(6.12)}{\leq} L(u(t_\ell))\, \|v_1 - v_2\|_{\mathcal{V}} \leq L_u\, \|v_1 - v_2\|_{\mathcal{V}}. \qquad \square$$

The existence of the uniform radius r_u and Lipschitz constant L_u is one of the key ingredients in the proof of the convergence order of the exponential integrator schemes in Chapter 9. It is also helpful to prove the continuity of the mapping from the data to the solution of the semi-linear operator DAE (6.9).

Theorem 6.19 (Continuous Dependence on Data for Semi-Linear Operator DAEs)**.** *Let the assumption of Theorem 6.15 be satisfied. Then the mapping from the data $u_0 \in \mathcal{V}$ and $g \in H^1(0,T;\mathcal{Q}^*)$ with $\mathcal{B}u_0 = g(0)$ to the solution (u,λ) of (6.9) is continuous.*

Proof. Let r_u and L_u be the uniform radius and Lipschitz constant of Lemma 6.18. We consider an arbitrary sequence $\{(u_{0,n}, g_n)\}_{n\in\mathbb{N}} \subseteq \mathcal{V} \times H^1(0,T;\mathcal{Q}^*)$ of consistent initial values and right-hand sides, which converges to (u_0, g) in $\mathcal{V} \times H^1(0,T;\mathcal{Q}^*)$. The associated solutions of the operator DAE (6.9) are (u_n, λ_n) and (u, λ), respectively. We show $\Delta u_n := u - u_n \to 0$ in $C([0,T],\mathcal{V})$ as $n \to \infty$. This is sufficient, since for big enough $n \in \mathbb{N}$ we have $\|\Delta u_n\|_{C([0,T],\mathcal{V})} \le r_u$ such that

$$\int_0^T \|f(t,u(t)) - f(t,u_n(t))\|^2_{\mathcal{H}^*}\,\mathrm{d}t \le \int_0^T L_u^2 \|\Delta u_n(t)\|^2_{\mathcal{V}}\,\mathrm{d}t \le L_u^2 T \|\Delta u_n\|^2_{C([0,T],\mathcal{V})} \to 0$$

as $n \to \infty$. The continuity follows then by Theorem 6.9.

To prove $\Delta u_n \to 0$, we define for arbitrary but fixed n the strictly monotonically increasing function $R_n \in C([0,T])$ by

$$\begin{aligned} R_n(t) = \|\mathcal{B}^-_{\mathcal{A}_1}\Delta g_n\|_{C([0,t],\mathcal{V})} + \Bigg[\frac{C_{\mathcal{A}_1}}{\mu_{\mathcal{A}_1}}\Big(\|\Delta u_{0,n}\|_{\mathcal{V}} + \|\mathcal{B}^-_{\mathcal{A}_1}\Delta g_n(0)\|_{\mathcal{V}}\Big)^2 \\ + \frac{3}{\mu_{\mathcal{A}_1}}\int_0^t (L_u + C_{\mathcal{A}_2})^2 \|\mathcal{B}^-_{\mathcal{A}_1}\Delta g_n(s)\|^2_{\mathcal{V}} + \|\mathcal{B}^-_{\mathcal{A}_1}\tfrac{\mathrm{d}}{\mathrm{d}s}\Delta g_n(s)\|^2_{\mathcal{H}^*}\,\mathrm{d}s\Bigg]^{1/2} \exp\left(\frac{3(L_u + C_{\mathcal{A}_2})^2}{2\mu_{\mathcal{A}_1}}t\right), \end{aligned}$$

where $\Delta u_{0,n}$ and Δg_n are defined analogously to Δu_n. Since $\Delta u_{0,n}$ and Δg_n are zero-sequences in $\mathcal{V}$ and $H^1(0,T;\mathcal{Q}^*)$, respectively, $R_n(T)$ vanishes as $n \to \infty$. Therefore, we find an $N \in \mathbb{N}$ such that $R_n(T) < r_u$ for all $n \ge N$. We want to prove $\|\Delta u_n(t)\|_{\mathcal{V}} < r_u$ for all $n \ge N$. Let us assume that this is not true for an $n \ge N$. Then by the continuity of Δu_n there exists a smallest $t^* \in [0,T]$ such that $\|\Delta u_n(t^*)\|_{\mathcal{V}} = r_u$. On the one hand, $t^* \ne 0$ holds by $R_n(0) \le R_n(T) < r_u$ and $\mu_{\mathcal{A}_1} \le C_{\mathcal{A}_1}$, and therefore the difference of the initial values $\Delta u_{0,n}$ is smaller than r_u. On the other hand, one has

$$\begin{aligned} &\mu_{\mathcal{A}_1}\|\Delta u_{n,\ker}(t)\|^2_{\mathcal{V}} \\ &\overset{(4.26)}{\le} C_{\mathcal{A}_1}\|\Delta u_{n,\ker}(0)\|^2_{\mathcal{V}} + \int_0^t \|f(s,u(s)) - f(s,u_n(s)) - \mathcal{A}_2\Delta u_n - \mathcal{B}^-_{\mathcal{A}_1}\tfrac{\mathrm{d}}{\mathrm{d}s}\Delta g_n(s)\|^2_{\mathcal{H}^*}\,\mathrm{d}s \\ &\overset{(6.12)}{\le} C_{\mathcal{A}_1}(\|\Delta u_{0,n}\|_{\mathcal{V}} + \|\mathcal{B}^-_{\mathcal{A}_1}\Delta g_n(0)\|_{\mathcal{V}})^2 \\ &\qquad + 3\int_0^t (L_u + C_{\mathcal{A}_2})^2\|\Delta u_{n,\ker}(s)\|^2_{\mathcal{V}} + (L_u + C_{\mathcal{A}_2})^2\|\mathcal{B}^-_{\mathcal{A}_1}\Delta g_n(s)\|^2_{\mathcal{V}} + \|\mathcal{B}^-_{\mathcal{A}_1}\tfrac{\mathrm{d}}{\mathrm{d}s}\Delta g_n(s)\|^2_{\mathcal{H}^*}\,\mathrm{d}s \end{aligned}$$

for all $t \in [0,t^*]$. By Gronwall's inequality (3.10), it follows

$$r_u = \|\Delta u_n(t^*)\|_{\mathcal{V}} \le \|\mathcal{B}^-_{\mathcal{A}_1} g(t^*)\|_{\mathcal{V}} + \|\Delta u_{n,\ker}(t^*)\|_{\mathcal{V}} \le R_n(t^*) \le R_n(T) < r_u.$$

This is a contradiction. Therefore, $\|\Delta u_n(t)\|_{\mathcal{V}} < r_u$ for all $t \in [0,T]$ and $n \ge N$. With the same estimate as before, one has $\|\Delta u_n(t)\|_{\mathcal{V}} \le R_n(t)$. The proof is finished by taking the limit

$$\lim_{n\to 0}\max_{t\in[0,T]}\|\Delta u_n(t)\|_{\mathcal{V}} \le \lim_{n\to 0}\max_{t\in[0,T]} R_n(t) \le \lim_{n\to 0} R_n(T) = 0. \qquad \square$$

7. Systems with Time-Dependent Operators

In this chapter we generalize the existence, uniqueness, and regularity results of the previous Chapter 6, to the case where the operators $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{B}$ in the operator DAE (B.1) depend on time. We therefore consider the operator DAE

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(t)u(t)) + (\mathcal{A}(t) - \tfrac{1}{2}\dot{\mathcal{M}}(t))u(t) \;-\; \mathcal{B}^*(t)\lambda(t) &= f(t) \qquad \text{in } \mathcal{V}^*, && (7.1\text{a}) \\ \mathcal{B}(t)u(t) \qquad\qquad &= g(t) \qquad \text{in } \mathcal{Q}^* && (7.1\text{b}) \end{aligned}$$

with operator-valued functions $\mathcal{M}\colon [0,T] \to \mathscr{L}(\mathcal{H},\mathcal{H}^*)$, $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{V},\mathcal{V}^*)$, and $\mathcal{B}\colon [0,T] \to \mathscr{L}(\mathcal{V},\mathcal{Q}^*)$. The initial value is given by u_0. The definition of a solution is given in Definition 4.27.

In the analysis of the operator DAE (7.1) we restrict ourselves to $\mathcal{A} \in L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$. The operator-valued functions $\mathcal{B}$ and $\mathcal{M}$ have derivatives in the sense of Definition 4.11. In the following, we assume that $\mathcal{B} \in H^1[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$ and $\mathcal{M} \in H^1[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$. The function $\mathcal{B}$ satisfies uniformly an inf-sup condition of the form (3.2). Inspired by the results of Chapter 6 we assume in the following that the operator $\mathcal{A}$ satisfies *uniformly a Gårding inequality* (3.6) *on* $\ker\mathcal{B}(t)$, i.e., the inequality

$$\langle\, \mathcal{A}(t)v_{\ker}, v_{\ker}\,\rangle \geq \mu_{\mathcal{A}}\|v_{\ker}\|^2_{\mathcal{V}} - \kappa_{\mathcal{A}}\|v_{\ker}\|^2_{\mathcal{H}} \tag{7.2}$$

holds for every $v_{\ker} \in \ker\mathcal{B}(t)$ at almost every time-point $t \in [0,T]$. Note that the kernel of $\mathcal{B}$ may depend on time, whereas $\mu_{\mathcal{A}} > 0$ and $\kappa_{\mathcal{A}} \in \mathbb{R}$ are constant. In addition, we assume that $\mathcal{M}$ is uniformly elliptic on $\mathcal{H}$.

Remark 7.1. Without loss of generality we can assume that $\mathcal{A}$ is *uniformly elliptic on* $\ker\mathcal{B}(t)$, i.e., $\kappa_{\mathcal{A}} = 0$ in (7.2). To verify the validity of this assumption, suppose that $\mathcal{A}$ only satisfies uniformly a Gårding inequality on $\ker\mathcal{B}(t)$ and that $\mathcal{M}$ is uniformly elliptic on $\mathcal{H}$. Then (u,λ) is a solution of the operator DAE (7.1) if and only if the transformed tuple $(\widetilde{u}(t), \widetilde{\lambda}(t)) = \exp(-\frac{\kappa_{\mathcal{A}}}{\mu_{\mathcal{M}}}t)(u(t),\lambda(t))$ fulfills

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}\widetilde{u}) \;+\; \big(\mathcal{A} + \tfrac{\kappa_{\mathcal{A}}}{\mu_{\mathcal{M}}}\mathcal{M} - \tfrac{1}{2}\dot{\mathcal{M}}\big)\widetilde{u} - \mathcal{B}^*\widetilde{\lambda} &= e^{-\frac{\kappa_{\mathcal{A}}}{\mu_{\mathcal{M}}}t} f \qquad \text{in } \mathcal{V}^*, \\ \mathcal{B}\widetilde{u} \qquad &= e^{-\frac{\kappa_{\mathcal{A}}}{\mu_{\mathcal{M}}}t} g \qquad \text{in } \mathcal{Q}^* \end{aligned}$$

with initial value $\widetilde{u}_0 = u_0$. The operator $\widetilde{\mathcal{A}}(t) \equiv \mathcal{A}(t) + \frac{\kappa_{\mathcal{A}}}{\mu_{\mathcal{M}}}\mathcal{M}(t) \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ is uniformly elliptic on $\ker\mathcal{B}(t)$, since for all $v_{\ker} \in \ker\mathcal{B}(t)$ and at almost every time-point $t \in [0,T]$ it satisfies

$$\langle\, \widetilde{\mathcal{A}}(t)v_{\ker}, v_{\ker}\,\rangle \geq \mu_{\mathcal{A}}\|v_{\ker}\|^2_{\mathcal{V}} - \kappa_{\mathcal{A}}\|v_{\ker}\|^2_{\mathcal{H}} + \tfrac{\kappa_{\mathcal{A}}}{\mu_{\mathcal{M}}}\mu_{\mathcal{M}}\|v_{\ker}\|^2_{\mathcal{H}} = \mu_{\mathcal{A}}\|v_{\ker}\|^2_{\mathcal{V}}.$$

In comparison to Chapter 6, we make the more restrictive assumptions $f \in L^2(0,T;\mathcal{V}^*)$ and $g \in H^1(0,T;\mathcal{V}^*)$ on the right-hand sides of the operator DAE (7.1). The case $f \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$ and $g \in W^{1,1}(0,T;\mathcal{V}^*)$ similar to Chapter 6 is discussed in the separate Remarks 7.15, 7.46, and in Theorem 7.21. In any case, the initial value u_0 is consistent, i.e., $u_0 \in \operatorname{clos}_{\mathcal{H}}(\ker\mathcal{B}(0)) + \mathcal{B}^-(0)g(0) \subset \mathcal{H}$. In particular, for time-independent $\mathcal{B}$ the consistency condition reads $u_0 \in \mathcal{H}_{\ker} + \mathcal{B}^- g(0)$ as usual. We summarize these assumptions.

Assumption 7.2 (Operators, Right-Hand Sides, and Initial Value of Operator DAE (7.1)).

i) The operator-valued function $\mathcal{B} \in H^1[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$ is uniformly inf-sup stable.

ii) We assume that $\mathcal{M} \in H^1[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$ is uniformly elliptic with constant $\mu_{\mathcal{M}} > 0$ and that $\mathcal{A}(t)$ with $\mathcal{A} \in L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ is elliptic on $\ker \mathcal{B}(t)$ with a uniform constant $\mu_{\mathcal{A}} > 0$ at almost every time-point $t \in [0,T]$.

iii) The right-hand sides satisfy $f \in L^2(0,T;\mathcal{V}^*)$ and $g \in H^1(0,T;\mathcal{Q}^*)$.

iv) The initial value u_0 fulfills $u_0 \in \operatorname{clos}_{\mathcal{H}}(\ker \mathcal{B}(0)) + \mathcal{B}^-(0)g(0) \subset \mathcal{H}$.

Example 7.3 (Linearized Navier–Stokes Equations)**.** The linearization of the incompressible Navier-Stokes equations around a prescribed vector field $v_\infty\colon [0,T] \times \Omega \to \mathbb{R}^d$ in its weak form is an operator DAE of the form (7.1); cf. [EmmM13, Eq. (3)]. The spaces $\mathcal{V}$, $\mathcal{H}$, and $\mathcal{Q}$ as well as the time-independent operators $\mathcal{M}$ and $\mathcal{B}$ are the same as for the unsteady Stokes equation; see Example 6.1. For regular v_∞ the operator $\mathcal{A}\colon [0,T] \to \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ is given by

$$\langle \mathcal{A}(t)u, v\rangle := \sum_{i=1}^{d} \int_\Omega \mu \nabla u_i \cdot \nabla v_i + v_i(v_\infty(t) \cdot \nabla u_i) + v_i(u \cdot \nabla v_{\infty,i}(t)) \,\mathrm{d}x.$$

Example 7.4 (Dynamical Boundary Condition with Non-Constant Relaxation Time)**.** The weak formulation of the heat equation with dynamic boundary conditions and a non-constant relaxation time $\sigma(t)$ on the boundary leads to an operator DAE (7.1) with a time-dependent operator $\mathcal{M}(t) = \sigma(t)\mathcal{R}_{L^2(\partial\Omega)}$; see [KovL17, Sec. 2.2.2] and Example 6.13.

Example 7.5. Consider the PDE

$$\dot{u}(\xi,t) + \partial_{\xi\xi\xi\xi}u(\xi,t) = f(\xi,t) \quad \text{in } \Omega \times (0,T] \tag{7.3}$$

with the domain $\Omega = (0,1)$ and vanishing Dirichlet and Neumann boundary conditions, i.e., $u(0,t) = u(1,t) = 0$ and $\partial_\xi u(0,t) = \partial_\xi u(1,t) = 0$. In addition, u should satisfy $u(\Phi(t),t) = g(t)$ with functions $\Phi, g \in H^1(0,T)$, where $0 < \Phi(t) < 1$. By applying Green's formula [Rou13, Eq. (1.54)] twice and the Lagrange multiplier method [Ste08, Ch. 4.2.1] we derive the weak formulation of (7.3), which is in the from of the operator DAE (7.1). The associated spaces are $\mathcal{V} = \mathcal{H}^2_0(0,1)$, $\mathcal{H} = L^2(0,1)$, and $\mathcal{Q} = \mathbb{R}$. The operator $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ given by $\langle\, \mathcal{A}u, v\,\rangle = (u,v)_{H^2_0(0,1)} = \big(\partial_{\xi\xi}u, \partial_{\xi\xi}v\big)_{L^2(0,1)}$ is elliptic, see Section 3.3, and $\mathcal{M} = \mathcal{R}_{L^2(0,1)}$. Furthermore, $\mathcal{B} \in H^1[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$ with $\mathcal{B}(t)v = v(\Phi(t))$ and derivative $\dot{\mathcal{B}}(t)v = \partial_\xi v(\Phi(t))\dot{\Phi}(t)$ is well-defined by [Bre10, Th. 8.8] and the assumptions on Φ. For its uniform inf-sup stability we set $v_t(\xi) = \frac{\xi^2(1-\xi)^2}{\Phi^2(t)(1-\Phi(t))^2} \in H^2_0(0,1)$ for fixed $t \in [0,T]$. For every $\alpha \in \mathbb{R} \setminus \{0\}$ we then have

$$\sup_{v \in H^2_0(0,1)\setminus\{0\}} \frac{\alpha \cdot \mathcal{B}(t)v}{\|v\|_{H^2_0(0,1)}|\alpha|} \geq \frac{\alpha^2}{\|\alpha v_t\|_{H^2_0(0,1)}|\alpha|} \geq \frac{\sqrt{5}}{2} \min_{t\in[0,T]} \Phi^2(t)(1-\Phi(t))^2 > 0.$$

Note that Φ is a continuous function [Bre10, Th. 8.8] with $0 < \Phi(t) < 1$ by assumption.

Linear time-dependent operators $\mathcal{B}$ occur by the linearization of PDEs with nonlinear constraints, e.g., the nonlinear boundary condition in the Stefan problem [DiPVY15, Sec. 2]. However, for systems of the form (7.1) with a time-dependent operator $\mathcal{B}$ there are only few results known. The paper [AltH18] is devoted to systems with $\mathcal{M}$ as the constant Riesz isomorphism in $\mathcal{H}$ and a time-dependent $\mathcal{B}$ with a constant kernel. Hyperbolic PDEs with moving Dirichlet boundary conditions are studied in [Alt14]. Here, roughly speaking, the operator $\mathcal{B}(t)$ is given by the trace operator restricted to a time-dependent part of the boundary $\Gamma(t) \subset \partial\Omega$. This operator, however, does not satisfy Assumption 7.2.i).

The analysis of operator DAEs (7.1) where only $\mathcal{A}$ is time-dependent is a straightforward generalization of the results of Chapter 6. We consider this case together with nonautonomous $\mathcal{M}$ in

Section 7.1. A time-dependent $\mathcal{M}$ can be interpreted as a nonautonomous inner product in $\mathcal{H}$; cf. Remark 6.6. In Section 7.2 we analyze operator DAEs with a time-dependent $\mathcal{B}$. In particular, we investigate nonautonomous splittings of $\mathcal{V}$ in Section 7.2.1. Finally in Section 7.3, we combine the results of the sections 7.1 and 7.2 and make statements on operator DAEs of the form (7.1) where all operators, namely $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{B}$, are dependent on time.

7.1. Time-Dependent Inner Products of the Pivot Space and Operators $\mathcal{A}$

In this section we investigate the operator DAE

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(t)u(t)) + (\mathcal{A}(t) - \tfrac{1}{2}\dot{\mathcal{M}}(t))u(t) - \mathcal{B}^*\lambda(t) = f(t) \quad \text{in } \mathcal{V}^*, \tag{7.4a}$$
$$\mathcal{B}u(t) = g(t) \quad \text{in } \mathcal{Q}^*. \tag{7.4b}$$

We emphasize that $\mathcal{A}$ and $\mathcal{M}$ are time-dependent operators, but $\mathcal{B}$ is constant in time. We assume that $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable and that Assumptions 7.2.ii)–iv) are satisfied. The existence of solutions of (7.4) to a given consistent initial value

$$u_0 \in \mathcal{H}_{\mathrm{ker}} + \mathcal{B}^- g(0)$$

is analyzed in the first part 7.1.1 of this section. For the analysis we discretize the operator DAE (7.4) in time by the implicit Euler method. We consider the weak/weak* limits of sequences of time-discrete solution given by the Euler scheme. The uniqueness of the solution is topic of Subsection 7.1.2. For this, we need to generalize the continuous embedding of $W^{1,2}(a, b; \mathcal{V}, \mathcal{V}^*)$ in $C([0, T], \mathcal{H})$, see Theorem 3.40, to the case of time-dependent inner products in $\mathcal{H}$. This is investigated in 7.1.2.1. Solutions under weaker assumptions on the right-hand sides f and g are analyzed in Subsection 7.1.3. There, we use continuity results of the solution on the data. This section finishes with comments on the regularity of the solution in 7.1.4, where we consider analogous cases to the ones in Section 6.2.

7.1.1. Existence Results

Let us consider the operator DAE (7.4). As in the case of autonomous operators in Chapter 6, we split the possible solution u into two parts and set $u = u_{\mathrm{ker}} + u_{\mathrm{c}}$. The first part u_{ker} is a function mapping into $\mathcal{V}_{\mathrm{ker}} = \ker \mathcal{B}$ and u_{c} satisfies $\mathcal{B}u_{\mathrm{c}} = g$ for almost all $t \in [0, T]$. There are many possible choices for u_{c}, e.g.,

$$u_{\mathrm{c}} = \mathcal{B}_\perp^- g \in H^1(0, T; \mathcal{V}_{\mathrm{ker}}^\perp); \tag{7.5}$$

see Definition 3.7 and Assumption 7.2.iii). Later in Subsection 7.1.4 we choose u_{c} depending on $\mathcal{A}$ in order to improve the regularity of the solution u as in Section 6.2. This, however, requires a smoother $\mathcal{A}$ in time. Therefore, we use (7.5) for now.

Using the splitting $u = u_{\mathrm{ker}} + u_{\mathrm{c}} = u_{\mathrm{ker}} + \mathcal{B}_\perp^- g$ in (7.4), we get for u_{ker} the operator DAE

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(t)u_{\mathrm{ker}}(t)) + (\mathcal{A}(t) - \tfrac{1}{2}\dot{\mathcal{M}}(t))u_{\mathrm{ker}}(t) - \mathcal{B}^*\lambda(t) = f_{\mathrm{ker}}(t) \quad \text{in } \mathcal{V}^*, \tag{7.6a}$$
$$\mathcal{B}u_{\mathrm{ker}}(t) = 0 \quad \text{in } \mathcal{Q}^* \tag{7.6b}$$

with the right-hand side

$$f_{\mathrm{ker}} := f - (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} - \mathcal{M}\dot{u}_{\mathrm{c}}. \tag{7.7}$$

By Lemma 4.15 and Assumptions 7.2.ii) and iii) the function f_{ker} is an element of $L^2(0, T; \mathcal{V}^*)$. In the following we study the operator DAE (7.6). The associated initial value $u_{\mathrm{ker},0} = u_0 - \mathcal{B}_\perp^- g(0)$ is an element of $\mathcal{H}_{\mathrm{ker}}$ by Assumption 7.2.iv). Note that we could consider equation (7.6a) tested only with functions in $\mathcal{V}_{\text{ker}}$. This would lead to an operator ODE with a time-dependent inner product of the pivot space $\mathcal{H}_{\text{ker}}$; cf. Remark 6.6. However, having the analysis of operator DAEs with time-dependent $\mathcal{B}$ in Section 7.2 in mind we directly consider system (7.6).

7.1.1.1. Rothe Method

To analyze the existence of solutions $(u_{\text{ker}}, \lambda)$ of (7.6), we use the *Rothe method* [Rot30] consisting in semi-discretization in time. We take an equidistant partition of the interval $[0, T]$ with step size $\tau = T/N > 0$, $N \in \mathbb{N}$, and discretize the operator DAE (7.6) formally with a Runge-Kutta method. This leads to stationary problems for the discrete time points $t_n \coloneqq n\tau$, $n = 1, \dots, N$. We work with an arbitrary monotonically decreasing zero-sequence of step sizes $\{\tau_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_{>0}$. In the following we omit the index k and write $\tau \to 0$ instead of $\lim_{k\to\infty} \tau_k = 0$.

In the whole of Chapter 7 we use for the Rothe method the implicit Euler scheme. Applied to (7.6) this leads to a stationary saddle point problem for every discrete time point t_n given by

$$D_\tau(\mathcal{M}u_{\text{ker}})_n + (\mathcal{A}_n - \tfrac{1}{2}D_\tau\mathcal{M}_n)u_{\text{ker},n} - \mathcal{B}^*\lambda_n = f_{\text{ker},n} \qquad \text{in } \mathcal{V}^*, \tag{7.8a}$$
$$\mathcal{B}u_{\text{ker},n} = 0 \qquad \text{in } \mathcal{Q}^* \tag{7.8b}$$

with $n = 1, \dots, N$. The terms $D_\tau(\mathcal{M}u_{\text{ker}})_n$ and $D_\tau\mathcal{M}_n$ denote the discrete derivatives

$$D_\tau(\mathcal{M}u_{\text{ker}})_n = \frac{\mathcal{M}_n u_{\text{ker},n} - \mathcal{M}_{n-1}u_{\text{ker},n-1}}{\tau} \quad \text{and} \quad D_\tau\mathcal{M}_n = \frac{\mathcal{M}_n - \mathcal{M}_{n-1}}{\tau},$$

respectively, with the pointwise evaluation $\mathcal{M}_n \coloneqq \mathcal{M}(t_n)$, $n = 0, \dots, N$. This is well-defined by Assumption 7.2.ii) and Lemma 4.14. Since $\mathcal{A} \in L^\infty[0, T; \mathscr{L}(\mathcal{V}, \mathcal{V}^*)]$ and $f_{\text{ker}} \in L^2(0, T; \mathcal{V}^*)$ are not continuous in general, they cannot be pointwise evaluated. Instead we use their mean over the interval $[t_{n-1}, t_n]$, i.e.,

$$f_{\text{ker},n} \coloneqq \frac{1}{\tau}\int_{t_{n-1}}^{t_n} f_{\text{ker}}(s)\,\mathrm{d}s \in \mathcal{V}^* \quad \text{and} \quad \mathcal{A}_n v \coloneqq \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \mathcal{A}(s)v\,\mathrm{d}s \tag{7.9}$$

for all $v \in \mathcal{V}$. Note that $\mathcal{A}_n \colon \mathcal{V} \to \mathcal{V}^*$ is linear by the pointwise linearity of $\mathcal{A}$ and of the integral, and bounded by

$$\|\mathcal{A}_n v\|_{\mathcal{V}^*} \le \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \|\mathcal{A}(s)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)}\|v\|_{\mathcal{V}}\,\mathrm{d}s \le \|\mathcal{A}\|_{L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]}\|v\|_{\mathcal{V}}. \tag{7.10}$$

Furthermore, the operator $\mathcal{A}_n$ is still elliptic on $\mathcal{V}_{\text{ker}}$, since for every $v_{\text{ker}} \in \mathcal{V}_{\text{ker}}$ we have

$$\langle \mathcal{A}_n v_{\text{ker}}, v_{\text{ker}} \rangle = \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \langle \mathcal{A}(s)v_{\text{ker}}, v_{\text{ker}} \rangle\,\mathrm{d}s \ge \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \mu_{\mathcal{A}}\|v_{\text{ker}}\|_{\mathcal{V}}^2\,\mathrm{d}s = \mu_{\mathcal{A}}\|v_{\text{ker}}\|_{\mathcal{V}}^2. \tag{7.11}$$

Remark 7.6. Under spatial discretization the operator DAE (7.6) becomes an index-2 DAE; see Chapter 2 and [Alt15, Sec. 8.2]. This leads in general to numerical difficulties for the temporally discretized DAE; see Section 5.1 for the finite-dimensional case as well as [Alt15, Sec. 6.1.3] for the temporally semi-discretized operator DAE (7.4). Under the assumption of consistent initial values, these difficulties, however, come from the inexact approximation of the hidden constraints; cf. [HaiLR89, p. 33]. Since the right-hand side of the constraint (7.6b) is homogeneous, its numerical derivative is exact, and no instabilities occur. A possible treatment of operator DAEs with inhomogeneous constraints is discussed in Section 8.1.

Remark 7.7. By the definition of $D_\tau \mathcal{M}_n$, for every $h \in \mathcal{H}$ we have

$$D_\tau \mathcal{M}_n h = \frac{1}{\tau}\big(\mathcal{M}(t_n)h - \mathcal{M}(t_{n-1})h\big) = \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \dot{\mathcal{M}}(s)h \,\mathrm{d}s. \tag{7.12}$$

The existence and uniqueness of a solution $(u_{\mathrm{ker},n}, \lambda_n)$ of (7.8) is discussed in the following lemma.

Lemma 7.8 (Solvability of Stationary System)**.** *Let $u_{\mathrm{ker},n-1} \in \mathcal{H}_{\mathrm{ker}}$ and $f_{\mathrm{ker},n} \in \mathcal{V}^*$ be given. Assume that $\mathcal{M}_n, \mathcal{M}_{n-1} \in \mathscr{L}(\mathcal{H}, \mathcal{H}^*)$ are elliptic. Suppose that $\mathcal{A}_n \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is elliptic on $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable. Then the stationary saddle point problem (7.8) has a unique solution $(u_{\mathrm{ker},n}, \lambda_n) \in \mathcal{V}_{\mathrm{ker}} \times \mathcal{Q}$, which depends linearly and continuously on $f_{\mathrm{ker},n}$ and $u_{\mathrm{ker},n-1}$.*

Proof. Since $\mathcal{A}_n$ is elliptic on $\mathcal{V}_{\mathrm{ker}}$ so is $\frac{1}{\tau}\mathcal{M}_n + \mathcal{A}_n - \frac{1}{2}D_\tau\mathcal{M}_n = \mathcal{A}_n + \frac{1}{2\tau}(\mathcal{M}_n + \mathcal{M}_{n-1})$. The assertion then follows by Theorem 3.8 and equation (7.8b). ☐

By Lemma 7.8 the stationary saddle point problem (7.8) generates for a given initial value $u_{\mathrm{ker},0} \in \mathcal{H}_{\mathrm{ker}}$ a sequence of solutions $u_{\mathrm{ker},n} \in \mathcal{V}_{\mathrm{ker}}$ and $\lambda_n \in \mathcal{Q}$ for the time points t_n, $n = 1, \dots, N$. In the following lemma we bound the elements of this sequence in terms of f_{ker} and $u_{\mathrm{ker},0}$.

Theorem 7.9 (Bound for the Discrete Solutions)**.** *Let $u_{\mathrm{ker},0} \in \mathcal{H}_{\mathrm{ker}}$ be given. Suppose that $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable and that Assumptions 7.2.ii) and iii) are satisfied. Let f_{ker} be defined by (7.7). Let $\mathcal{M}_n := \mathcal{M}(t_n)$ as well as $\mathcal{A}_n$ and $f_{\mathrm{ker},n}$ be defined as in (7.9). Then the sequentially defined unique solution $(u_{\mathrm{ker},n}, \lambda_n) \in \mathcal{V}_{\mathrm{ker}} \times \mathcal{Q}$, $n = 1, \dots, N$, of (7.6) satisfies*

$$\|u_{\mathrm{ker},n}\|^2_{\mathcal{M}_n} + \sum_{k=1}^n \|u_{\mathrm{ker},k} - u_{\mathrm{ker},k-1}\|^2_{\mathcal{M}_{k-1}} + \mu_{\mathcal{A}} \sum_{k=1}^n \tau \|u_{\mathrm{ker},k}\|^2_{\mathcal{V}} \le M^2(u_{\mathrm{ker},0}, f_{\mathrm{ker}}) \tag{7.13}$$

with the constant $M(u_{\mathrm{ker},0}, f_{\mathrm{ker}}) = \sqrt{\|u_{\mathrm{ker},0}\|^2_{\mathcal{M}_0} + \frac{1}{\mu_{\mathcal{A}}}\|f_{\mathrm{ker}}\|^2_{L^2(0,T;\mathcal{V}^*)}}$. Furthermore, the discrete derivative $D_\tau(\mathcal{M}u_{\mathrm{ker}})_n$ satisfies

$$\begin{aligned}&\sum_{k=1}^n \tau \|D_\tau(\mathcal{M}u_{\mathrm{ker}})_k\|^2_{\mathcal{V}^*_{\mathrm{ker}}} \\ &\quad \le 3\Big(\mu_{\mathcal{A}} + \frac{1}{\mu_{\mathcal{A}}}\|\mathcal{A}\|^2_{L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]} + \frac{C^2_{\mathcal{H}^*\hookrightarrow\mathcal{V}^*}}{4\mu_{\mathcal{M}}}\|\dot{\mathcal{M}}\|^2_{L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]}\Big) M^2(u_{\mathrm{ker},0}, f_{\mathrm{ker}}).\end{aligned} \tag{7.14}$$

Proof. As first step we note that the equality

$$\begin{aligned}&2\langle \mathcal{M}_n h - \mathcal{M}_{n-1}\bar{h}, h\rangle - \langle (\mathcal{M}_n - \mathcal{M}_{n-1})h, h\rangle \\ &= \langle \mathcal{M}_n h, h\rangle - \langle \mathcal{M}_{n-1}\bar{h}, \bar{h}\rangle + \langle \mathcal{M}_{n-1}(h-\bar{h}), h-\bar{h}\rangle \\ &= \|h\|^2_{\mathcal{M}_n} - \|\bar{h}\|^2_{\mathcal{M}_{n-1}} + \|h-\bar{h}\|^2_{\mathcal{M}_{n-1}}\end{aligned} \tag{7.15}$$

holds for every $h, \bar{h} \in \mathcal{H}$. Testing (7.6a) with $\tau u_{\mathrm{ker},n}$ then leads to

$$\begin{aligned}&\frac{1}{2}\big(\|u_{\mathrm{ker},n}\|^2_{\mathcal{M}_n} - \|u_{\mathrm{ker},n-1}\|^2_{\mathcal{M}_{n-1}}\big) + \frac{1}{2}\|u_{\mathrm{ker},n} - u_{\mathrm{ker},n-1}\|^2_{\mathcal{M}_{n-1}} + \tau\mu_{\mathcal{A}}\|u_{\mathrm{ker},n}\|^2_{\mathcal{V}} \\ &\overset{(7.15)}{\le} \tau\langle D_\tau(\mathcal{M}u_{\mathrm{ker}})_n + (\mathcal{A}_n - \tfrac{1}{2}D_\tau\mathcal{M}_n)u_{\mathrm{ker},n}, u_{\mathrm{ker},n}\rangle \\ &\overset{(7.6a)}{=} \tau\langle f_{\mathrm{ker},n}, u_{\mathrm{ker},n}\rangle \overset{(3.8)}{\le} \frac{\tau}{2\mu_{\mathcal{A}}}\|f_{\mathrm{ker},n}\|^2_{\mathcal{V}^*} + \tau\frac{\mu_{\mathcal{A}}}{2}\|u_{\mathrm{ker},n}\|^2_{\mathcal{V}}.\end{aligned}$$

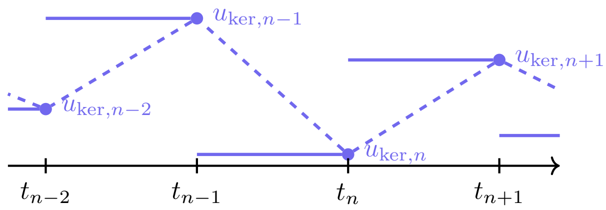

Figure 7.1.: Piecewise functions $u_{\ker,\tau}$ (solid) and $\widehat{u}_{\ker,\tau}$ (dashed) defined by $u_{\ker,n}$, $n = 0, \dots, N$.

Summing this inequality from $k = 1$ to n implies (7.13), since by the Cauchy-Schwarz inequality we have

$$\tau \sum_{k=1}^{n} \|f_{\ker,k}\|_{\mathcal{V}^*}^2 \overset{(7.9)}{=} \frac{1}{\tau} \sum_{k=1}^{n} \| \int_{t_{k-1}}^{t_k} f_{\ker}(s)\, \mathrm{d}s\|_{\mathcal{V}^*}^2 \leq \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|f_{\ker}(s)\|_{\mathcal{V}^*}^2\, \mathrm{d}s \leq \int_0^T \|f_{\ker}(s)\|_{\mathcal{V}^*}^2\, \mathrm{d}s.$$

Further, we note that $\mathcal{B}^*\lambda_n \in \mathcal{V}_{\ker}^0$ vanishes if we consider (7.6a) as an equation in $\mathcal{V}_{\ker}^*$. Inequality (7.14) then follows by (7.13), Young's inequality (3.8), and

$$\begin{aligned} &\|D_\tau(\mathcal{M}u_{\ker})_n\|_{\mathcal{V}_{\ker}^*} \\ \leq\ &\|f_{\ker,n}\|_{\mathcal{V}^*} + \|(\mathcal{A}_n - \tfrac{1}{2}D_\tau\mathcal{M}_n)u_{\ker,n}\|_{\mathcal{V}^*} \\ \overset{(7.12)}{=}\ &\|f_{\ker,n}\|_{\mathcal{V}^*} + \frac{1}{\tau}\Big\| \int_{t_{n-1}}^{t_n} (\mathcal{A}(s) - \tfrac{1}{2}\dot{\mathcal{M}}(s))u_{\ker,n}\, \mathrm{d}s\Big\|_{\mathcal{V}^*} \\ \leq\ &\|f_{\ker,n}\|_{\mathcal{V}^*} + \|\mathcal{A}\|_{L^\infty[t_{n-1},t_n;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]}\|u_{\ker,n}\|_{\mathcal{V}} + \frac{C_{\mathcal{H}^*\hookrightarrow\mathcal{V}^*}}{2\sqrt{\tau}}\|\dot{\mathcal{M}}\|_{L^2[t_{n-1},t_n;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]}\|u_{\ker,n}\|_{\mathcal{H}}. \end{aligned} \tag{7.16}$$ □

Remark 7.10. Equality (7.15) also holds for arbitrary self-adjoint, elliptic operators $\mathcal{K}_1, \mathcal{K}_2 \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ instead of $\mathcal{M}_n, \mathcal{M}_{n-1}$. Especially, for $\mathcal{K} = \mathcal{K}_1 = \mathcal{K}_2$ and $x, \bar{x} \in \mathcal{X}$ equation (7.15) becomes

$$2\langle\, \mathcal{K}(x - \bar{x}), x\,\rangle = \|x\|_{\mathcal{K}}^2 - \|\bar{x}\|_{\mathcal{K}}^2 + \|x - \bar{x}\|_{\mathcal{K}}^2. \tag{7.17}$$

7.1.1.2. Convergence of Temporal Discrete Solutions

Given the sequence of the discrete solutions $(u_{\ker,n}, \lambda_n) \in \mathcal{V}_{\ker} \times \mathcal{Q}$, $n = 1, \dots, N$ of (7.8) from Lemma 7.9, we build functions over the whole interval $[0, T]$ and show that they converge to a solution of (7.4). More precisely, we define the piecewise constant function $u_{\ker,\tau}\colon [0, T] \to \mathcal{V}_{\ker}$ and the piecewise linear function $\widehat{u}_{\ker,\tau}\colon [0, T] \to \mathcal{H}_{\ker}$ by

$$u_{\ker,\tau}(t) := \begin{cases} u_{\ker,1}, & \text{if } t = 0, \\ u_{\ker,n}, & \text{if } t \in (t_{n-1}, t_n], \end{cases} \qquad \widehat{u}_{\ker,\tau}(t) := \begin{cases} u_{\ker,0}, & \text{if } t = 0, \\ u_{\ker,n} + D_\tau u_{\ker,n}(t - t_n), & \text{if } t \in (t_{n-1}, t_n]. \end{cases} \tag{7.18}$$

A sketch of $u_{\ker,\tau}$ and $\widehat{u}_{\ker,\tau}$ is given in Figure 7.1. Analogously, we define the piecewise constant function λ_τ for the Lagrange multiplier λ via λ_n. The starting value $\lambda_\tau(0)$ can be chosen arbitrarily, since we will study the convergence in spaces of Bochner-integrable functions and the initial time-point is a null set. In the same manner we define piecewise constant functions associated to the right-hand side $f_{\ker}$ and the operator $\mathcal{A}$ denoted by $f_{\ker,\tau}$ and $\mathcal{A}_\tau$, respectively. At last, we define the piecewise linear functions

$$\widehat{\mathcal{M}}_\tau(t) := \begin{cases} \mathcal{M}_0, & \text{if } t = 0, \\ \mathcal{M}_n + D_\tau \mathcal{M}_n(t - t_n), & \text{if } t \in (t_{n-1}, t_n], \end{cases} \tag{7.19a}$$

$$\widehat{\mathcal{M}u}_{\text{ker},\tau}(t) := \begin{cases} \mathcal{M}_0 u_{\text{ker},0}, & \text{if } t = 0, \\ \mathcal{M}_n u_{\text{ker},n} + D_\tau(\mathcal{M}u_{\text{ker}})_n(t - t_n), & \text{if } t \in (t_{n-1}, t_n]. \end{cases} \tag{7.19b}$$

Note that, in the interest of readability we write $\widehat{\mathcal{M}u}_{\text{ker},\tau}$ for the piecewise linear function defined by $\mathcal{M}_n u_{\text{ker},n}$ instead of $(\widehat{\mathcal{M}u_{\text{ker}}})_\tau$. Furthermore, we denote the generalized time derivative of $\widehat{\mathcal{M}u}_{\text{ker},\tau}$ by $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\text{ker},\tau}$, which is piecewise constant with value $D_\tau(\mathcal{M}u_{\text{ker}})_n$ at $(t_{n-1}, t_n]$, $n = 1, \dots, N$. Analogously, we use the notation $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_\tau$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}$. With this the temporally discretized system (7.8) can be reformulated as

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\text{ker},\tau} + (\mathcal{A}_\tau - \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_\tau)u_{\text{ker},\tau} - \mathcal{B}^*\lambda_\tau &= f_{\text{ker},\tau} \quad \text{in } \mathcal{V}^*, && (7.20\text{a}) \\ \mathcal{B}u_{\text{ker},\tau} &= 0 \quad \text{in } \mathcal{Q}^*. && (7.20\text{b}) \end{aligned}$$

The main goal of this section is to prove that the piecewise defined functions converge to a solution of the operator DAE (7.6) as $\tau \to 0$. At first we show that $u_{\text{ker},\tau}$ and $\widehat{u}_{\text{ker},\tau}$ have the same weak limit as τ tends to zero.

Lemma 7.11. *Let the assumptions of Theorem 7.9 be satisfied. Then there exists a function $u_{\text{ker}} \in L^2(0,T;\mathcal{V}_{\text{ker}}) \cap L^\infty(0,T;\mathcal{H}_{\text{ker}})$ and a subsequence τ' of τ such that as $\tau' \to 0$ we have*

$$u_{\text{ker},\tau'} \rightharpoonup u_{\text{ker}} \quad \text{in } L^2(0,T;\mathcal{V}), \qquad u_{\text{ker},\tau'}, \widehat{u}_{\text{ker},\tau'} \overset{*}{\rightharpoonup} u_{\text{ker}} \quad \text{in } L^\infty(0,T;\mathcal{H}), \tag{7.21a}$$

$$\widehat{\mathcal{M}}_{\tau'}u_{\text{ker},\tau'}, \widehat{\mathcal{M}}_{\tau'}\widehat{u}_{\text{ker},\tau'}, \widehat{\mathcal{M}u}_{\text{ker},\tau'} \overset{*}{\rightharpoonup} \mathcal{M}u_{\text{ker}} \quad \text{in } L^\infty(0,T;\mathcal{H}^*). \tag{7.21b}$$

Proof. By Theorem 7.9 the function $u_{\text{ker},\tau}$ is bounded by $c\,M := c\,M(u_{\text{ker},0}, f_{\text{ker}})$ in $L^\infty(0,T;\mathcal{H})$ and $L^2(0,T;\mathcal{V})$ with $c^2 = 1/\mu_\mathcal{M}$ and $c^2 = 1/\mu_\mathcal{A}$, respectively, independently by τ. For a bound for the piecewise linear function $\widehat{u}_{\text{ker},\tau}$, we note

$$\sqrt{\mu_\mathcal{M}}\|\widehat{u}_{\text{ker},\tau}(t)\|_\mathcal{H} = \tfrac{\sqrt{\mu_\mathcal{M}}}{\tau}\|u_{\text{ker},n-1}(t_n - t) + u_{\text{ker},n}(t - t_{n-1})\|_\mathcal{H} \le \|u_{\text{ker},n-1}\|_{\mathcal{M}_{n-1}} + \|u_{\text{ker},n}\|_{\mathcal{M}_n} \le 2M$$

in $[t_{n-1}, t_n]$. Therefore, the weak limits (7.21a) follow by the same arguments as in [Emm04, Th. 8.3.8]. Since $\mathcal{V}_{\text{ker}}$ is closed in $\mathcal{V}$ and $u_{\text{ker},\tau}(t) \in \mathcal{V}_{\text{ker}}$ for almost all $t \in [0,T]$, the limit u_{ker} is pointwise an element of $\mathcal{V}_{\text{ker}}$ at almost every time-point [Trö10, Th. 2.11].

Since the operator $\mathcal{M}$ can be identified as an operator from $\mathscr{L}(L^p(0,T;\mathcal{H}), L^p(0,T;\mathcal{H}^*))$, $p \in [1,\infty]$, by Lemma 4.9, the functions $\mathcal{M}u_{\text{ker},\tau'}, \mathcal{M}\widehat{u}_{\text{ker},\tau'}$ converge in a weak* sense to $\mathcal{M}u_{\text{ker}}$ in $L^\infty(0,T;\mathcal{H}^*)$ as $\tau' \to 0$; cf. [Zei90a, Prop. 21.28]. By the continuity of $\mathcal{M}$, see Lemma 4.14, we have $\widehat{\mathcal{M}}_\tau \to \mathcal{M}$ in $C([0,T], \mathscr{L}(\mathcal{H}, \mathcal{H}^*))$ as $\tau \to 0$. This implies

$$\widehat{\mathcal{M}}_{\tau'}u_{\text{ker},\tau'} = (\widehat{\mathcal{M}}_{\tau'} - \mathcal{M})u_{\text{ker},\tau'} + \mathcal{M}u_{\text{ker},\tau'} \overset{*}{\rightharpoonup} \mathcal{M}u_{\text{ker}} \quad \text{in } L^\infty(0,T;\mathcal{H}^*)$$

as $\tau' \to 0$ and analogously for $\widehat{\mathcal{M}}_{\tau'}\widehat{u}_{\text{ker},\tau'}$. For $\widehat{\mathcal{M}u}_{\text{ker},\tau'}$ we observe

$$\begin{aligned} &\mu_\mathcal{M} \max_{t\in[0,T]} \|\widehat{\mathcal{M}u}_{\text{ker},\tau}(t) - \mathcal{M}(t)\widehat{u}_{\text{ker},\tau}(t)\|^2_{\mathcal{H}^*} \\ &= \mu_\mathcal{M} \max_{n=1,\dots,N} \max_{t\in[t_{n-1},t_n]} \Big\|(\mathcal{M}_n - \mathcal{M}(t))u_{\text{ker},n}\tfrac{(t-t_{n-1})}{\tau} + (\mathcal{M}_{n-1} - \mathcal{M}(t))u_{\text{ker},n-1}\tfrac{(t_n - t)}{\tau}\Big\|^2_{\mathcal{H}^*} \end{aligned}$$

$$\leq 2\mu_{\mathcal{M}} \max_{n=1,\ldots,N} \max_{t\in[t_{n-1},t_n]} \Big\| \int_t^{t_n} \dot{\mathcal{M}}(s) u_{\mathrm{ker},n} \,\mathrm{d}s \Big\|_{\mathcal{H}^*}^2 \tfrac{(t-t_{n-1})^2}{\tau^2} + \Big\| \int_{t_{n-1}}^{t} \dot{\mathcal{M}}(s) u_{\mathrm{ker},n-1} \,\mathrm{d}s \Big\|_{\mathcal{H}^*}^2 \tfrac{(t_n-t)^2}{\tau^2}$$

$$\overset{(7.13)}{\leq} \tau \frac{8}{27} \int_0^T \|\dot{\mathcal{M}}\|_{\mathscr{L}(\mathcal{H},\mathcal{H}^*)}^2 \,\mathrm{d}s\, M^2(u_{\mathrm{ker},0}, f_{\mathrm{ker}}) \to 0$$

as $\tau \to 0$. Therefore, $\widehat{\mathcal{M}u}_{\mathrm{ker},\tau'}$ and $\mathcal{M}\widehat{u}_{\mathrm{ker},\tau'}$ have the same weak* limit in $L^\infty(0,T;\mathcal{H}^*)$. □

In the next step of the proof that the piecewisely defined functions converge to a solution of (7.6) we investigate the limiting behavior of the derivative $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau}$.

Lemma 7.12. *Suppose that the assumptions of Theorem 7.9 are satisfied. Let u_{ker} be the function and τ' be the subsequence of τ introduced in Lemma 7.11. Then $\mathcal{M}u_{\mathrm{ker}} \in H^1(0,T;\mathcal{V}_{\mathrm{ker}}^*) \cap L^\infty(0,T;\mathcal{H}^*)$ with $(\mathcal{M}u_{\mathrm{ker}})(0) = \mathcal{M}(0)u_{\mathrm{ker},0}$ in $\mathcal{V}_{\mathrm{ker}}^*$ holds. Further, as $\tau' \to 0$ we have*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau'} \rightharpoonup \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{ker}}) \quad \text{in } L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*).$$

Proof. By Theorem 7.9 the estimate

$$\|\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau}\|_{L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*)}^2 = \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \|D_\tau(\mathcal{M}u_{\mathrm{ker}})_k\|_{\mathcal{V}_{\mathrm{ker}}^*}^2 \,\mathrm{d}s \overset{(7.14)}{\lesssim} M^2(u_{\mathrm{ker},0}, f_{\mathrm{ker}})$$

holds with a constant independent of τ. Therefore, there exists a subsequence τ'' of τ' such that $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau''}$ converges weakly to a w in $L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*)$. Proposition 23.19 in [Zei90a] then shows that $\mathcal{M}u_{\mathrm{ker}}$ has a generalized derivative in $L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*)$ given by w. Since the derivative is unique, $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau'}$ converges weakly to $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{ker}})$ for the whole sequence τ' [GajGZ74, Ch. 1, Lem. 5.4]. In addition, for every $v_{\mathrm{ker}} \in \mathcal{V}_{\mathrm{ker}}$ the equality

$$\begin{aligned}
\langle (\mathcal{M}u_{\mathrm{ker}})(0), Tv_{\mathrm{ker}} \rangle &= \int_0^T -\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathcal{M}u_{\mathrm{ker}}, (T-t)v_{\mathrm{ker}} \rangle \,\mathrm{d}t \\
&= \int_0^T \langle \mathcal{M}u_{\mathrm{ker}}, v_{\mathrm{ker}} \rangle - \langle \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{ker}}), (T-t)v_{\mathrm{ker}} \rangle \,\mathrm{d}t \\
&= \lim_{\tau'\to 0} \int_0^T \langle \widehat{\mathcal{M}u}_{\mathrm{ker},\tau'}, v_{\mathrm{ker}} \rangle - \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau'}, (T-t)v_{\mathrm{ker}} \rangle \,\mathrm{d}t \\
&= \lim_{\tau'\to 0} \langle \mathcal{M}_0 u_{\mathrm{ker},0}, Tv_{\mathrm{ker}} \rangle = \langle \mathcal{M}(0)u_{\mathrm{ker},0}, Tv_{\mathrm{ker}} \rangle
\end{aligned}$$

holds. Since $T \neq 0$, this shows $(\mathcal{M}u_{\mathrm{ker}})(0) = \mathcal{M}(0)u_{\mathrm{ker},0}$ in $\mathcal{V}_{\mathrm{ker}}^*$. □

As our last preparatory step, we consider the approximations of the operators $\mathcal{A}$ and $\dot{\mathcal{M}}$.

Lemma 7.13. *Let the assumptions of Theorem 7.9 be satisfied. Suppose that u_{ker} is the same function and τ' is the same subsequence of τ as in Lemma 7.12. Then as $\tau' \to 0$, we have*

$$\mathcal{A}_{\tau'} u_{\mathrm{ker},\tau'} \rightharpoonup \mathcal{A}u_{\mathrm{ker}} \quad \text{in } L^2(0,T;\mathcal{V}^*) \quad \text{and} \quad \big(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_{\tau'}\big)u_{\mathrm{ker},\tau'} \rightharpoonup \dot{\mathcal{M}}u_{\mathrm{ker}} \quad \text{in } L^2(0,T;\mathcal{H}^*).$$

Proof. Note that both limits are well-defined by Lemma 4.9.

Let us first consider $\mathcal{A}_\tau u_{\mathrm{ker},\tau}$. By the estimates (7.10) and (7.13) the function $\mathcal{A}_\tau u_{\mathrm{ker},\tau}$ is bounded in $L^2(0,T;\mathcal{V}^*)$ independently of τ. With [Yos80, Sec. V.1, Th. 3] it then is enough to show $\langle \mathcal{A}_{\tau'} u_{\mathrm{ker},\tau'}, \varphi \rangle \to \langle \mathcal{A}u_{\mathrm{ker}}, \varphi \rangle$ as $\tau' \to 0$ for every element φ of a dense subset of $L^2(0,T;\mathcal{V})$. We consider the set of polynomials $\varphi\colon [0,T] \to \mathcal{V}$, $\varphi(t) = \sum_{k=0}^{r} v_k t^k$ with $r \in \mathbb{N}_0$ and $v_k \in \mathcal{V}$ for $k = 0,\ldots,r$, which is dense in $L^2(0,T;\mathcal{V})$; see [Zei90a, Prop. 23.2.d]. Let φ be an arbitrary polynomial of degree r with monomial terms $\varphi_k(t) := v_k t^k$, $v_k \in \mathcal{V}$ and $k = 0, \dots, r$. Then the identity

$$\langle \mathcal{A}_\tau u_{\mathrm{ker},\tau}, \varphi_k \rangle = \int_0^T \langle \mathcal{A}_\tau(s) u_{\mathrm{ker},\tau}(s), v s^k \rangle \,\mathrm{d}s = \int_0^T \langle \mathcal{A}_\tau^*(s) v s^k, u_{\mathrm{ker},\tau}(s) \rangle \,\mathrm{d}s = \langle \mathcal{A}_\tau^* \varphi_k, u_{\mathrm{ker},\tau} \rangle$$

holds for every φ_k. We note that $\mathcal{A}_\tau^* v_k = (\mathcal{A}^* v_k)_\tau$ is the piecewise constant function corresponding to $\mathcal{A}^* v_k \in L^2(0,T;\mathcal{V}^*)$, analogously to the way $f_{\mathrm{ker},\tau}$ corresponds to f_{ker}; see (7.9). By Lemma 3.34, the function $\mathcal{A}_\tau^* v_k = (\mathcal{A}^* v_k)_\tau$ converges strongly to $\mathcal{A}^* v_k$ in $L^2(0,T;\mathcal{V}^*)$ as $\tau \to 0$. Hence, $\lim_{\tau\to 0} \mathcal{A}_\tau^* \varphi_k = \mathcal{A}^* \varphi_k$ in $L^2(0,T;\mathcal{V}^*)$. Since $\mathcal{A}_{\tau'}^* \varphi_k$ converges strongly and $u_{\mathrm{ker},\tau'}$ weakly as $\tau' \to 0$, the limit

$$\langle \mathcal{A}_{\tau'} u_{\mathrm{ker},\tau'}, \varphi \rangle = \sum_{k=0}^r \langle \mathcal{A}_{\tau'} u_{\mathrm{ker},\tau'}, \varphi_k \rangle = \sum_{k=0}^r \langle \mathcal{A}_{\tau'}^* \varphi_k, u_{\mathrm{ker},\tau'} \rangle \to \sum_{k=0}^r \langle \mathcal{A}^* \varphi_k, u_{\mathrm{ker}} \rangle = \langle \mathcal{A} u_{\mathrm{ker}}, \varphi \rangle$$

holds as $\tau' \to 0$. This implies the asserted weak convergence of $\mathcal{A}_{\tau'} u_{\mathrm{ker},\tau'}$.

Analogously, one proves the weak limit of $(\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_\tau) u_{\mathrm{ker},\tau}$, where its boundedness is given by

$$\begin{aligned}\int_0^T \|(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_\tau) u_{\mathrm{ker},\tau}\|_{\mathcal{H}^*}^2 \,\mathrm{d}s &\overset{(7.12)}{=} \frac{1}{\tau^2} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \| \int_{t_{n-1}}^{t_n} \dot{\mathcal{M}} u_{\mathrm{ker},n} \,\mathrm{d}s \|_{\mathcal{H}^*}^2 \,\mathrm{d}t \\ &\le \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\dot{\mathcal{M}}\|_{\mathscr{L}(\mathcal{H},\mathcal{H}^*)}^2 \,\mathrm{d}s \|u_{\mathrm{ker},n}\|_{\mathcal{H}}^2 \\ &\overset{(7.13)}{\le} \frac{1}{\mu_{\mathcal{M}}} \|\dot{\mathcal{M}}\|_{L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]}^2 M^2(u_{\mathrm{ker},0}, f_{\mathrm{ker}}). \end{aligned}$$ □

We can now prove the existence of a solution of the operator DAE (7.4).

Theorem 7.14 (Existence of a Solution). *Let $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ satisfy an inf-sup condition (3.2) and let Assumptions 7.2.ii) and 7.2.iii) be fulfilled. Suppose that Assumption 7.2.iv) is satisfied, i.e., $u_0 \in \mathcal{H}_{\mathrm{ker}} + \mathcal{B}^- g(0)$. Then the operator DAE (7.4) has at least one solution (u, λ), which satisfies*

a) $u \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$, *b)* $\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda$ *for an* $\Lambda \in L^\infty(0,T;\mathcal{Q})$,

c) $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) \in L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*)$, *d)* $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - \mathcal{B}^*\lambda \in L^2(0,T;\mathcal{V}^*)$.

Proof. Let $u_{\mathrm{c}} \in H^1(0,T;\mathcal{V}_{\mathrm{ker}}^\perp)$ be defined as in (7.5) and $u_{\mathrm{ker}} \in L^2(0,T;\mathcal{V}_{\mathrm{ker}}) \cap L^\infty(0,T;\mathcal{H}_{\mathrm{ker}})$ be the weak limit mentioned in Lemma 7.11 with initial value $u_{\mathrm{ker},0} = u_0 - u_{\mathrm{c}}(0) \in \mathcal{H}_{\mathrm{ker}}$. Then $u := u_{\mathrm{c}} + u_{\mathrm{ker}}$ satisfies a) and c) by the embeddings $\mathcal{V} \hookrightarrow \mathcal{H}$, $\mathcal{H}^* \hookrightarrow \mathcal{V}_{\mathrm{ker}}^*$, Theorem 3.38, Lemmas 4.15, 7.11, and 7.12.

For the Lagrange multiplier λ we investigate λ_τ. We consider the integration

$$\mathcal{I} \colon L^1(0,T;\mathcal{X}) \to L^\infty(0,T;\mathcal{X}),\ x \mapsto \Big(t \mapsto \int_0^t x(s) \,\mathrm{d}s\Big). \tag{7.22}$$

Since $\mathcal{I}$ is linear and bounded we have the weak* limit

$$\begin{aligned}\mathcal{I}(\mathcal{B}^*\lambda_{\tau'}) &\overset{(7.20a)}{=} \widehat{\mathcal{M}} u_{\mathrm{ker},\tau'} - \mathcal{M}(0) u_{\mathrm{ker},0} + \mathcal{I}(\mathcal{A}_{\tau'} u_{\mathrm{ker},\tau'}) - \tfrac{1}{2}\mathcal{I}(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_{\tau'} u_{\mathrm{ker},\tau'}) - \mathcal{I}(f_{\mathrm{ker},\tau'}) \\ &\overset{*}{\rightharpoonup} \mathcal{M} u_{\mathrm{ker}} - \mathcal{M}(0) u_{\mathrm{ker},0} + \mathcal{I}((\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}}) u_{\mathrm{ker}}) - \mathcal{I}(f_{\mathrm{ker}}) \\ &\overset{(7.7)}{=} \mathcal{M} u - \mathcal{M}(0) u_0 + \mathcal{I}((\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}}) u) - \mathcal{I}(f) =: \Lambda_{\mathcal{B}^*}\end{aligned}$$

in $L^\infty(0,T;\mathcal{V}^*)$ as $\tau' \to 0$. Here, we have used Lemmas 7.11, 7.13, and $f_{\mathrm{ker},\tau} \to f_{\mathrm{ker}}$ in $L^2(0,T;\mathcal{V}^*)$ as $\tau \to 0$ by Lemma 3.34. By Lemma 3.30 the identity $\mathcal{I}(\mathcal{B}^*\lambda_{\tau'}) = \mathcal{B}^*\mathcal{I}\lambda_{\tau'}$ holds, which implies $\Lambda_{\mathcal{B}^*}(t) \in \mathcal{V}^0_{\mathrm{ker}}$ for almost every t [Trö10, Th. 2.11]. Therefore, $\Lambda := \mathcal{B}^{-*}_{\mathrm{left}}\Lambda_{\mathcal{B}^*} \in L^\infty(0,T;\mathcal{Q})$ is well-defined with $\mathcal{B}^{-*}_{\mathrm{left}}$ as the left-inverse of $\mathcal{B}^*$; see Lemma 3.6. In particular, $\mathcal{I}\lambda_{\tau'} \overset{*}{\rightharpoonup} \Lambda$ in $L^\infty(0,T;\mathcal{Q})$ as $\tau' \to 0$.

We have to show that (u,λ) with $\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda$ is a solution of (7.4). For this, we note that by (7.5), (7.7), and (7.20a), for every $v \in \mathcal{V}$, $\varphi \in C^\infty([0,T])$ with $\varphi(T) = 0$ we have

$$
\begin{aligned}
0 &= \int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}u}_{\mathrm{ker},\tau'} + \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{c}}) + (\mathcal{A}_{\tau'} - \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_{\tau'})u_{\mathrm{ker},\tau'} + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}}, v \rangle \varphi \\
&\qquad\qquad + \langle -\mathcal{B}^*\lambda_\tau - f_{\mathrm{ker},\tau'} + f_{\mathrm{ker}} - f, v \rangle \varphi \,\mathrm{d}s \\
&= \int_0^T \langle (\mathcal{A}_{\tau'} - \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{M}}_{\tau'})u_{\mathrm{ker},\tau'} + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} - f_{\mathrm{ker},\tau'} + f_{\mathrm{ker}} - f, v \rangle \varphi \\
&\qquad\qquad - \langle \widehat{\mathcal{M}u}_{\mathrm{ker},\tau'} + \mathcal{M}u_{\mathrm{c}} - \mathcal{B}^*\mathcal{I}\lambda_{\tau'}, v \rangle \dot{\varphi}\,\mathrm{d}s + \langle \mathcal{M}_0 u_{\mathrm{ker},0} + \mathcal{M}(0)u_{\mathrm{c}}(0), v \rangle \varphi(0) \\
&\to \int_0^T \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u - f, v \rangle \varphi - \langle \mathcal{M}u - \mathcal{B}^*\Lambda, v \rangle \dot{\varphi}\,\mathrm{d}s + \langle \mathcal{M}(0)u_0, v \rangle \varphi(0)
\end{aligned}
\tag{7.23}
$$

as $\tau' \to 0$ using Lemmas 3.34, 7.11, and 7.13. For every $q \in \mathcal{Q}$ and φ as for (7.23) we observe

$$
\int_0^T \langle \mathcal{B}u, q \rangle \varphi \,\mathrm{d}s = \int_0^T \langle \mathcal{B}u_{\mathrm{ker}} + \mathcal{B}u_{\mathrm{c}}, q \rangle \varphi \,\mathrm{d}s = \int_0^T \langle \mathcal{B}u_{\mathrm{c}}, q \rangle \varphi \,\mathrm{d}s \overset{(7.5)}{=} \int_0^T \langle g, q \rangle \varphi \,\mathrm{d}s \tag{7.24}
$$

by Lemma 7.11. The equations (7.23) and (7.24) prove that (u,λ) with λ as the distributional derivative of Λ is a solution of the operator DAE (7.4). In particular, b) is satisfied by the choice of λ.

Finally, equation (7.23) with $\varphi \in C^\infty_{\mathrm{c}}(0,T)$ implies that $\mathcal{M}u - \mathcal{B}^*\widetilde{\lambda}$ has a generalized derivative in $L^2(0,T;\mathcal{V}^*)$. This proves assertion d). □

Remark 7.15. The existence of a solution can be proven under the weaker assumptions $\mathcal{M} \in W^{1,1}[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$, $f \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$, and $g \in W^{1,1}(0,T;\mathcal{Q}^*)$. The spaces of c) and d) then change to $L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}) + L^1(0,T;\mathcal{H}^*_{\mathrm{ker}})$ and $L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$, respectively; see also Theorem 7.21.

7.1.2. Uniqueness Results

Theorem 7.14 proves the existence of a solution (u,λ) of the operator DAE (7.4). In this subsection we investigate its uniqueness. In the case where operator $\mathcal{M}$ is time-independent, the usual proof is based on the embedding of generalized differentiable functions in the space of continuous functions; see e.g. [DauL92, Sec. XVIII.3, Th. 1] and [DauL93, Sec. XIX.2, Th. 1]. Therefore, we generalize $W^{1,2}(0,T;\mathcal{V},\mathcal{V}^*) \hookrightarrow C([0,T],\mathcal{H})$ for a time-dependent inner product of the pivot space $\mathcal{H}$ induced by $\mathcal{M}$ in Subsection 7.1.2.1. With this generalization we prove the uniqueness of the solution (u,λ) of the operator DAE (7.4) in Subsection 7.1.2.2.

7.1.2.1. $W^{1,2}$-Functions with Nonautonomous Inner Product

Since the operator $\mathcal{M}$ is time-dependent in (7.4), the inner product $m(t,h_1,h_2) = \langle \mathcal{M}(t)h_1, h_2 \rangle_{\mathcal{H}^*,\mathcal{H}}$ of the pivot space $\mathcal{H}$ is time-dependent as well. In order to analyze the uniqueness of the solution, we need an embedding similar to $W^{1,2}(0,T;\mathcal{V},\mathcal{V}^*) \hookrightarrow C([0,T],\mathcal{H})$ with a nonconstant inner product in $\mathcal{H}$. We slightly generalize a result from [Str66].

Theorem 7.16. *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $\mathcal{X}$ be reflexive. Assume that $\mathcal{X}$ and $\mathcal{Y}$ are contained in a linear space and that $\mathcal{X} \cap \mathcal{Y}$ is dense in $\mathcal{X}$ and $\mathcal{Y}$. Suppose that a Banach space $\mathcal{Z}$ exists such that $\mathcal{Z}$ and $\mathcal{Z}^*$ are dense in $L^1(0,T;\mathcal{Y})$ and $L^1(0,T;\mathcal{Y}^*)$, respectively, with continuous inclusions. Let for every $f \in \mathcal{Z}^*$ and $z \in \mathcal{Z}$, the duality pairing $\langle f, z \rangle_{\mathcal{Y}^*,\mathcal{Y}}$ be integrable. Suppose that $\mathcal{Z}$ is closed under multiplications with real-valued step functions and that translations in t are continuous for elements in $\mathcal{Z}$, where the functions are extended by zero outside of $[0,T]$.*

Let $\mathcal{K} \in H^1[0,T;\mathscr{L}(\mathcal{X},\mathcal{X}^*)]$ be pointwise self-adjoint and uniformly elliptic. Suppose that $w \in L^\infty(0,T;\mathcal{X})$ fulfills $\dot{w} \in \mathcal{Z}$ and $\mathcal{K}w \in \mathcal{Z}^*$. Then $w \in C([0,T],\mathcal{X})$ and for all $t \in [0,T]$ we have

$$\langle \mathcal{K}w, w \rangle_{\mathcal{X}^*,\mathcal{X}}\big|_0^t = \int_0^t \langle \dot{\mathcal{K}}(s)w(s), w(s) \rangle_{\mathcal{X}^*,\mathcal{X}} + 2\langle \mathcal{K}(s)w(s), \dot{w}(s) \rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}s. \tag{7.25}$$

Proof. The case with $\mathcal{K} \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{X},\mathcal{X}^*)]$ is proven in [Str66, Th. 3.1 & Th. 3.2]. An adaptation of the proof of [Str66, Th. 3.1] shows the assertion. Here, we use that for every $p \in [1,\infty)$ and $w \in L^p(0,T;\mathcal{X})$ (extended by zero outside of $[0,T]$) the sequence of convolutions $\{\varphi_\varepsilon * w\}_{\varepsilon\in\mathbb{R}_{>0}}$ with mollifiers $\{\varphi_\varepsilon\}_{\varepsilon\in\mathbb{R}_{>0}}$, i.e., nonnegative $C_c^\infty(\mathbb{R})$-functions with $\operatorname{supp}(\varphi_\varepsilon) \subset [-\varepsilon,\varepsilon]$ and $\int_{\mathbb{R}} \varphi_\varepsilon \, \mathrm{d}t = 1$, converges strongly to w in $L^p(0,T;\mathcal{X})$ as $\varepsilon \to 0$ [KufJF77, Sec. 2.5]. □

Remark 7.17. If two function w_1, w_2 satisfy the assumptions of Theorem 7.16, then they satisfy

$$\begin{aligned} 2\langle \mathcal{K}w_1, w_2 \rangle\big|_0^t &= \big(\langle \mathcal{K}(w_1+w_2), w_1+w_2 \rangle - \langle \mathcal{K}w_1, w_1 \rangle - \langle \mathcal{K}w_2, w_2 \rangle\big)\big|_0^t \\ &\overset{(7.25)}{=} \int_0^t \langle \dot{\mathcal{K}}w_1, w_2 \rangle + \langle \dot{\mathcal{K}}w_2, w_1 \rangle + 2\langle \mathcal{K}w_1, \dot{w}_2 \rangle + 2\langle \mathcal{K}w_2, \dot{w}_1 \rangle \, \mathrm{d}s. \end{aligned}$$

Theorem 7.16 implies the generalization of $W^{1,2}(0,T;\mathcal{V},\mathcal{V}^*) \hookrightarrow C([0,T],\mathcal{H})$.

Theorem 7.18. *Let $\mathcal{V}, \mathcal{H}, \mathcal{V}^*$ be a Gelfand triple. Suppose that $\mathcal{M} \in H^1[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$ is pointwise self-adjoint and uniformly elliptic. Assume that $u \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$ fulfills $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) \in L^2(0,T;\mathcal{V}^*)$. Then $u \in C([0,T],\mathcal{H})$ and for all $t \in [0,T]$ we have*

$$\langle \mathcal{M}(t)u(t), u(t) \rangle - \langle \mathcal{M}(0)u(0), u(0) \rangle = \int_0^t 2\langle \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(s)u(s)), u(s) \rangle - \langle \dot{\mathcal{M}}(s)u(s), u(s) \rangle \, \mathrm{d}s.$$

Proof. With the notation of Theorem 7.16 we choose $\mathcal{X} = \mathcal{H}^*$, $\mathcal{Y} = \mathcal{V}^*$, $\mathcal{K} = \mathcal{M}^{-1}$, $\mathcal{Z} = L^2(0,T;\mathcal{V}^*)$, and $w = \mathcal{M}u \in L^\infty(0,T;\mathcal{H}^*)$. The continuity under translation for every element in $\mathcal{Z}$ is proven in [GajGZ74, Ch. 4, Lem. 1.5] and $C([0,T],\mathcal{V}^*) \hookrightarrow \mathcal{Z}$ is dense in $L^1(0,T;\mathcal{V}^*) = L^1(0,T;\mathcal{Y})$ by Theorem 3.32. Analogously, the embedding $\mathcal{Z}^* \hookrightarrow L^1(0,T;\mathcal{V}) = L^1(0,T;\mathcal{Y})$ holds. The assumptions of Theorem 7.16 are then fulfilled by $\mathcal{K}w = \mathcal{M}^{-1}\mathcal{M}u = u$ and since $\mathcal{M}^{-1} \in H^1[0,T;\mathscr{L}(\mathcal{H}^*,\mathcal{H})]$ is uniformly elliptic by Lemma 4.17. Therefore, $w = \mathcal{M}u$ is continuous with images in $\mathcal{H}^*$ by Theorem 7.16. This implies $u = \mathcal{M}^{-1}w \in C([0,T],\mathcal{H})$ and

$$\begin{aligned} \langle \mathcal{M}u, u \rangle\big|_0^t &\overset{(7.25)}{=} \int_0^t \langle \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}^{-1}(s))\mathcal{M}(s)u(s), \mathcal{M}(s)u(s) \rangle + 2\langle \mathcal{M}^{-1}(s)\mathcal{M}(s)u(s), \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(s)u(s)) \rangle \, \mathrm{d}s \\ &= \int_0^t -\langle \dot{\mathcal{M}}(s)u(s), u(s) \rangle + 2\langle u(s), \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(s)u(s)) \rangle \, \mathrm{d}s, \end{aligned}$$

where we used Lemma 4.17 and $\mathcal{M}^* = \mathcal{M}$. □

7.1.2.2. Uniqueness of the Solution and Continuity in the Data

With Theorem 7.18 we are able to prove the uniqueness of the solution of the operator DAE (7.4). For the estimate of the solution we use the shortened notation $\|\cdot\|_C$ for $\|\cdot\|_{C([0,T],\mathcal{X})}$, where the

Banach space $\mathcal{X}$ is clear by the argument. Analogously, we use a shortened notation for the norms of $H^1(0,T;\mathcal{X})$, $L^\infty[0,T;\mathscr{L}(\mathcal{X},\mathcal{X}^*)]$, and $L^2[0,T;\mathscr{L}(\mathcal{X},\mathcal{X}^*)]$.

Theorem 7.19 (Uniqueness of the Solution)**.** *Let the assumptions of Theorem 7.14 be satisfied. Then there exists only one solution (u,λ) of the operator DAE (7.4), which fulfills the condition a) and b) in Theorem 7.14. In addition to the stated spaces in Theorem 7.14, the solution satisfies $u \in C([0,T],\mathcal{H})$ with $u(0) = u_0$ and $\Lambda \in C([0,T],\mathcal{Q})$ with $\Lambda(0) = 0$ where $\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda$ in a distributional sense. The bounds*

$$\|u\|^2_{L^2(0,T;\mathcal{V})} \le \frac{2\|\mathcal{M}\|_C}{\mu_\mathcal{A}}\|u_0\|^2_\mathcal{H} + \frac{4}{\mu_\mathcal{A}^2}\|f\|^2_{L^2} + \frac{1}{\beta^2}\Big(1 + \frac{4\widetilde{C}}{\mu_\mathcal{A}}\Big)\|g\|^2_{H^1} \tag{7.26a}$$

$$\|u\|^2_{C([0,T],\mathcal{H})} \le \frac{4\|\mathcal{M}\|_C}{\mu_\mathcal{M}}\|u_0\|^2_\mathcal{H} + \frac{8}{\mu_\mathcal{A}\mu_\mathcal{M}}\|f\|^2_{L^2} + \frac{1}{\beta^2}\Big(\frac{4C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\max(1,4T^2)}{T} + \frac{8\widetilde{C}}{\mu_\mathcal{M}}\Big)\|g\|^2_{H^1} \tag{7.26b}$$

$$\begin{aligned}\|\Lambda\|^2_{C([0,T],\mathcal{Q})} \le \frac{5}{\beta^2}\Big(&\|\mathcal{M}\|^2_C C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\big(\|u\|^2_{C([0,T],\mathcal{H})} + \|u_0\|^2_\mathcal{H}\big) + T\big(\|\mathcal{A}\|^2_{L^\infty}\|u\|_{L^2(0,T;\mathcal{V})} \\ &+ \tfrac{1}{4}C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\|\dot{\mathcal{M}}\|^2_{L^2}\|u\|^2_{C([0,T],\mathcal{H})} + \|f\|^2_{L^2}\big)\Big)\end{aligned} \tag{7.26c}$$

hold with the constant $\widetilde{C} := \frac{\|\mathcal{A}\|^2_{L^\infty}}{\mu_\mathcal{A}} + \frac{\|\mathcal{M}\|^2_C C^4_{\mathcal{V}\hookrightarrow\mathcal{H}}}{\mu_\mathcal{A}} + \Big(\|\mathcal{M}\|_C + \frac{\|\dot{\mathcal{M}}\|^2_{L^2} C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}}{2\mu_\mathcal{A}}\Big)\frac{C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\max(1,4T^2)}{T}$*. In particular, the solution operator*

$$\begin{aligned}\mathcal{S}\colon \{(f,g,u_0) \in L^2(0,T;\mathcal{V}^*) \times H^1(0,T;\mathcal{Q}^*) \times \mathcal{H} \,|\, u_0 - \mathcal{B}^- g(0) \in \mathcal{H}_\mathrm{ker}\} \\ \to L^2(0,T;\mathcal{V}) \cap C([0,T],\mathcal{H}) \times C([0,T],\mathcal{Q}), \quad (f,g,u_0) \mapsto (u,\Lambda)\end{aligned}$$

is linear and continuous.

Proof. Theorem 7.14 proves the existence of at least one solution (u,λ). Let now (u,λ) be an arbitrary solution, which satisfies condition a) of Theorem 7.14, i.e., $u \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$. We split u into $u_\mathrm{ker} \in L^2(0,T;\mathcal{V}_\mathrm{ker})$ and $u_\mathrm{c} \in L^2(0,T;\mathcal{V}^\perp_\mathrm{ker})$. This is well-defined, since $\mathcal{V}_\mathrm{ker}$ is a closed subspace of $\mathcal{V}$. The algebraic constraint (7.4b) then implies $u_\mathrm{c} = \mathcal{B}^-_\perp\mathcal{B}u = \mathcal{B}^-_\perp g$. Especially, u_c is an element of $H^1(0,T;\mathcal{V}^\perp_\mathrm{ker}) \hookrightarrow C([0,T],\mathcal{H})$ with bounds

$$\|u_\mathrm{c}^{(k)}\|_{L^2(0,T;\mathcal{V})} \le \frac{1}{\beta}\|g^{(k)}\|_{L^2(0,T;\mathcal{V})}, \quad \|u_\mathrm{c}\|_{C([0,T],\mathcal{H})} \le \frac{\sqrt{2}C_{\mathcal{V}\hookrightarrow\mathcal{H}}}{\beta}\frac{\max(1,2T)}{\sqrt{T}}\|g\|_{H^1(0,T;\mathcal{V})}, \tag{7.27}$$

$k = 0,1$, by Theorem 3.38 and [Bra07, Lem. III.4.2.b]. The part $u_\mathrm{ker} \in L^2(0,T;\mathcal{V}_\mathrm{ker}) \cap L^\infty(0,T;\mathcal{H}_\mathrm{ker})$ satisfies by Definition 4.27 of a solution the equality

$$\begin{aligned}0 &= \int_0^T \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u - f, v_\mathrm{ker}\rangle\varphi - \langle \mathcal{M}u, v_\mathrm{ker}\rangle\dot\varphi\,\mathrm{d}s \\ &= \int_0^T \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u - f + \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_\mathrm{c}), v_\mathrm{ker}\rangle\varphi - \langle \mathcal{M}u_\mathrm{ker}, v_\mathrm{ker}\rangle\dot\varphi\,\mathrm{d}s\end{aligned} \tag{7.28}$$

for every $v_\mathrm{ker} \in \mathcal{V}_\mathrm{ker}$ and $\varphi \in C^\infty_\mathrm{c}(0,T)$, where we used

$$\int_0^T \langle \Lambda_{\mathcal{B}^*}, v_\mathrm{ker}\rangle\dot\varphi\,\mathrm{d}s = -\int_0^T \langle \mathcal{B}^*\lambda, v_\mathrm{ker}\rangle\varphi\,\mathrm{d}s = -\int_0^T \langle \mathcal{B}v_\mathrm{ker}, \lambda\rangle\varphi\,\mathrm{d}s = \int_0^T \langle \mathcal{B}v_\mathrm{ker}, \Lambda\rangle\dot\varphi\,\mathrm{d}s = 0. \tag{7.29}$$

Equation (7.28) implies that $\mathcal{M}u_\mathrm{ker} \in L^2(0,T;\mathcal{H}^*_\mathrm{ker})$ has the derivative $f - (\mathcal{A} - \frac{1}{2}\dot{\mathcal{M}})u - \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_\mathrm{c})$ in $L^2(0,T;\mathcal{V}^*_\mathrm{ker})$. Therefore, the assumptions of Theorem 7.18 are satisfied for u_ker with the Gelfand triple $\mathcal{V}_{\text{ker}}, \mathcal{H}_{\text{ker}}, \mathcal{V}_{\text{ker}}^*$. Theorem 7.18 then implies $u_{\text{ker}} \in C([0,T], \mathcal{V}_{\text{ker}})$ and

$$\begin{aligned}
&\tfrac{1}{2}\|u_{\text{ker}}(t)\|^2_{\mathcal{M}(t)} - \tfrac{1}{2}\|u_{\text{ker}}(0)\|^2_{\mathcal{M}(0)} + \mu_{\mathcal{A}} \int_0^t \|u_{\text{ker}}\|^2_{\mathcal{V}} \,\mathrm{d}s \\
\leq\; & \int_0^t \langle \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\text{ker}}) + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\text{ker}}, u_{\text{ker}} \rangle \,\mathrm{d}s \\
\overset{(7.4a)}{=}\; & \int_0^t \langle f - (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\text{c}} - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\text{c}}), u_{\text{ker}} \rangle \,\mathrm{d}s \\
\overset{(3.8)}{\leq}\; & \int_0^t \frac{1}{2\mu_{\mathcal{A}}}\|f - (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})\mathcal{B}^- g - \mathcal{M}\mathcal{B}^-\dot{g}\|^2_{\mathcal{V}^*} + \frac{\mu_{\mathcal{A}}}{2}\|u_{\text{ker}}\|^2_{\mathcal{V}} \,\mathrm{d}s.
\end{aligned} \tag{7.30}$$

Further, the identity $u_{\text{ker}}(0) = u_0 - \mathcal{B}_{\perp}^- g(0) \in \mathcal{H}_{\text{ker}}$ holds, since for every $v_{\text{ker}} \in \mathcal{V}_{\text{ker}}$ and $\varphi \in C^\infty([0,T])$ with $\varphi(T) = 0$ we have

$$\begin{aligned}
0 &\overset{(4.30)}{=} \int_0^T \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u - f, v_{\text{ker}} \rangle \varphi - \langle \mathcal{M}u, v_{\text{ker}} \rangle \dot{\varphi} \,\mathrm{d}s + \langle \mathcal{M}(0)u_0, v_{\text{ker}} \rangle \varphi(0) \\
&= \int_0^T \langle \underbrace{(\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u - f + \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u)}_{=0 \text{ in } \mathcal{V}_{\text{ker}}^*}, v_{\text{ker}} \rangle \varphi \,\mathrm{d}s + \langle \mathcal{M}(0)(u_0 - u_{\text{c}}(0) - u_{\text{ker}}(0)), v_{\text{ker}} \rangle \varphi(0).
\end{aligned}$$

The estimates (7.26a) and (7.26b) then follow from $\|u_{\text{ker}}(0)\|_{\mathcal{H}} \leq \|u_0\|_{\mathcal{H}} + \|\mathcal{B}_{\perp}^- g(0)\|_{\mathcal{H}}$, (7.27), and (7.30). In particular, these bounds imply $u = 0$ for a solution for the operator DAE (7.4) with homogeneous initial value and right-hand sides. The uniqueness of the solution u then follows by the superposition principle 4.28.

For the Lagrange multiplier λ we use the integration operator $\mathcal{I}$ as in (7.22). By (4.30) the equality

$$0 = \int_0^T \langle \mathcal{I}f - \mathcal{I}((\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u) - \mathcal{M}u + \Lambda_{\mathcal{B}^*}, v \rangle \dot{\varphi} \,\mathrm{d}s - \langle \mathcal{M}(0)u_0, v \rangle \varphi(0)$$

holds for every $v \in \mathcal{V}$ and $\varphi \in C^\infty([0,T])$ with $\varphi(T) = 0$. This implies that a constant $c \in \mathcal{V}^*$ exists such that $c = \mathcal{I}f - \mathcal{I}((\mathcal{A} - \frac{1}{2}\dot{\mathcal{M}})u) - \mathcal{M}u + \Lambda_{\mathcal{B}^*}$ at almost every time-point $t \in [0,T]$; see [Emm04, Cor. 8.1.4]. By choosing $\varphi(t) = (T-t)/T$ we obtain $\langle \mathcal{M}(0)u_0, v \rangle = \int_0^T \langle c, v \rangle \dot{\varphi} \,\mathrm{d}s = -\langle c, v \rangle$. Together this implies

$$\Lambda_{\mathcal{B}^*} = \mathcal{M}u - \mathcal{M}(0)u_0 + \mathcal{I}((\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u) - \mathcal{I}f \in C([0,T], \mathcal{V}^*) \tag{7.31}$$

with $\Lambda_{\mathcal{B}^*}(0) = 0$. Note that this shows the uniqueness of $\Lambda_{\mathcal{B}^*}$. By (7.29) we have $\Lambda = \mathcal{B}_{\text{left}}^{-*}\Lambda_{\mathcal{B}^*}$. Therefore, Λ is unique, continuous, and vanishes at the initial time-point, and its distributional derivative λ is unique as well. Estimate (7.26c) follows from (7.31).

The estimates (7.26) show the boundedness of $\mathcal{S}$. Its linearity follows by Remark 4.28. □

Remark 7.20. Since the solution of the operator DAE (7.4) is unique, every sequence – not only its subsequence – converges entirely to its weak/weak* limit described in Lemmas 7.11–7.13 [GajGZ74, Ch. I, Lem. 5.4].

7.1.3. Generalizations

In Chapter 6 we have proved the existence of a solution of the operator DAE (6.1) with time-independent operators for right-hand sides $f \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$ and $g \in W^{1,1}(0,T;\mathcal{Q}^*)$. This holds also if $\mathcal{M}$ and $\mathcal{A}$ are nonconstant.

Theorem 7.21 (Existence and Uniqueness of Solutions II). *Let $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ satisfy an inf-sup condition (3.2), Assumption 7.2.ii) be fulfilled, and $f \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$ and $g \in W^{1,1}(0,T;\mathcal{Q}^*)$. Suppose that u_0 is consistent in the sense of Assumption 7.2.iv). Then the operator DAE (7.4) has a unique solution (u, λ), which satisfies*

a) $u \in L^2(0,T;\mathcal{V}) \cap C([0,T],\mathcal{H})$, b) $\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda$ *for an* $\Lambda \in C([0,T],\mathcal{Q})$,

c) $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) \in L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}) + L^1(0,T;\mathcal{H}^*_{\mathrm{ker}})$, d) $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - \mathcal{B}^*\lambda \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$.

Further, the solution fulfills $u(0) = u_0$, $\Lambda(0) = 0$ and depends linearly and continuously on the data (f, g, u_0).

Proof. As in proof of Theorems 7.14 and 7.19, we define $u_{\mathrm{c}} \coloneqq \mathcal{B}^-_\perp g \in W^{1,1}(0,T;\mathcal{V}_{\mathrm{c}})$ and consider $u_{\mathrm{ker}} = u - u_{\mathrm{c}}$ with a possible solution u. We note that (u,λ) solves the operator DAE (7.4) with initial value u_0 and right-hand sides f, g, if and only if $(u_{\mathrm{ker}},\lambda)$ solves the operator DAE (7.6) with initial value $u_{\mathrm{ker},0} \coloneqq u_0 - u_{\mathrm{c}}(0) \in \mathcal{H}_{\mathrm{ker}}$ and right-hand side f_{ker} as defined in equation (7.7).

By assumptions, there exists $f_1 \in L^2(0,T;\mathcal{V}^*)$ and $f_2 \in L^1(0,T;\mathcal{H}^*)$ such that $f = f_1 + f_2$. Therefore, we can split f_{ker} into

$$f^{[1]}_{\mathrm{ker}} \coloneqq f_1 - \mathcal{A}\mathcal{B}^-_\perp g - \tfrac{1}{2}\dot{\mathcal{M}}\mathcal{B}^-_\perp g \in L^2(0,T;\mathcal{V}^*) \quad \text{and} \quad f^{[2]}_{\mathrm{ker}} \coloneqq f_2 - \mathcal{M}\mathcal{B}^-_\perp \dot g \in L^1(0,T;\mathcal{H}^*).$$

By Theorem 7.14 and 7.19 there exists a unique solution $(u^{[1]}_{\mathrm{ker}}, \lambda^{[1]})$ to the data $(f^{[1]}_{\mathrm{ker}}, 0, u_{\mathrm{ker},0})$ of (7.4), which satisfies a)–d).

By the superposition principle, see Remark 4.28, it is enough to prove the assertion for the operator DAE (7.6) with right-hand sides $f_{\mathrm{ker}} = f^{[2]}_{\mathrm{ker}}$, $g = 0$, and a vanishing initial value. Then by Theorem 3.32.ii) there exists a sequence $\{f^{[2]}_{\mathrm{ker},n}\}_{n\in\mathbb{N}} \subset C([0,T],\mathcal{H}^*)$, which converges to $f^{[2]}_{\mathrm{ker}}$ in $L^1(0,T;\mathcal{H}^*)$ as $n \to \infty$. By Theorem 3.32.viii), 7.14, and 7.19 the operator DAE (7.6) has a unique solution $(u^{[2]}_{\mathrm{ker},n}, \lambda^{[2]}_n)$ for every $f^{[2]}_{\mathrm{ker},n}$, which satisfies

$$\frac{\mu_\mathcal{M}}{2}\|u^{[2]}_{\mathrm{ker},n}(t)\|^2_\mathcal{H} + \mu_\mathcal{A}\int_0^t \|u^{[2]}_{\mathrm{ker},n}\|^2_\mathcal{V}\,\mathrm{d}s \le \int_0^t \langle f^{[2]}_{\mathrm{ker},n}, u^{[2]}_{\mathrm{ker},n}\rangle\,\mathrm{d}s \le \int_0^t \|f^{[2]}_{\mathrm{ker},n}\|_{\mathcal{H}^*}\|u^{[2]}_{\mathrm{ker},n}\|_\mathcal{H}\,\mathrm{d}s;$$

cf. estimate (7.30). Following the lines of Theorem 4.22 one has

$$\mu_\mathcal{M}\|u^{[2]}_{\mathrm{ker},n}\|^2_{C([0,T],\mathcal{H})} + \mu_\mathcal{A}\|u^{[2]}_{\mathrm{ker},n}\|^2_{L^2(0,T;\mathcal{V})} \le \tfrac{1}{\mu_\mathcal{M}}\|f^{[2]}_{\mathrm{ker},n}\|^2_{L^1(0,T;\mathcal{H}^*)}.$$

Similarly as in the proof of Theorem 7.19, one shows that $\Lambda^{[2]}_n$ is bounded as in the estimate (7.26c) with $C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\|f^{[2]}_{\mathrm{ker},n}\|^2_{L^1(0,T;\mathcal{H}^*)}$ instead of $T\|f^{[2]}_{\mathrm{ker},n}\|^2_{L^2(0,T;\mathcal{V}^*)}$. Note that the estimates for $u^{[2]}_{\mathrm{ker},n}$ and $\Lambda^{[2]}_n$ also hold for the differences $u^{[2]}_{\mathrm{ker},n} - u^{[2]}_{\mathrm{ker},m}$ and $\Lambda^{[2]}_n - \Lambda^{[2]}_m$, respectively, with right-hand side $f^{[2]}_{\mathrm{ker},n} - f^{[2]}_{\mathrm{ker},m}$. Therefore, $\{u^{[2]}_{\mathrm{ker},n}\}_{n\in\mathbb{N}}$ and $\{\Lambda^{[2]}_n\}_{n\in\mathbb{N}}$ are Cauchy sequences in $L^2(0,T;\mathcal{V}_{\mathrm{ker}}) \cap C([0,T],\mathcal{H}_{\mathrm{ker}})$ and $C([0,T],\mathcal{Q})$, respectively. Let $u^{[2]}_{\mathrm{ker}}$ and $\Lambda^{[2]}$ be their limits. Then we have $u^{[2]}_{\mathrm{ker}}(0) = \lim_{n\to\infty} u^{[2]}_{\mathrm{ker},n}(0) = 0$, $\Lambda^{[2]}(0) = \lim_{n\to\infty}\Lambda^{[2]}_n(0) = 0$, and the equality

$$\begin{aligned} 0 &= \lim_{n\to\infty}\int_0^T \langle \mathcal{B}^*\Lambda^{[2]}_n - \mathcal{M}u^{[2]}_{\mathrm{ker},n}, v\rangle\dot\varphi + \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u^{[2]}_{\mathrm{ker},n} - f^{[2]}_{\mathrm{ker},n}, v\rangle\varphi + \langle \mathcal{B}u^{[2]}_{\mathrm{ker},n}, q\rangle\varphi\,\mathrm{d}t \\ &= \int_0^T \langle \mathcal{B}^*\Lambda^{[2]} - \mathcal{M}u^{[2]}_{\mathrm{ker}}, v\rangle\dot\varphi + \langle (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u^{[2]}_{\mathrm{ker}} - f^{[2]}_{\mathrm{ker}}, v\rangle\varphi + \langle \mathcal{B}u^{[2]}_{\mathrm{ker}}, q\rangle\varphi\,\mathrm{d}t \end{aligned}$$

is satisfied for every $v \in \mathcal{V}$, $q \in \mathcal{Q}$, and $\varphi \in C^\infty([0,T])$ with $\varphi(T) = 0$. This shows that $(u^{[2]}_{\mathrm{ker}}, \lambda^{[2]})$

solves the operator DAE (7.4) with data $(f_{\ker}^{[2]}, 0, 0)$ where $\lambda^{[2]} = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda^{[2]}$ in the distributional sense. The estimates for $u_{\ker,n}^{[2]}$ and $\Lambda_n^{[2]}$ also hold in the limit. By the superposition principle this shows the uniqueness of $(u_{\ker}^{[2]}, \lambda^{[2]})$. The assertions c) and d) follow by (7.6a). □

For later analysis we need another generalization where we introduce an additional operator.

Lemma 7.22. *Let the assumptions of Theorem 7.21 be satisfied. Suppose that the operator-valued function $\mathcal{A}_{\mathcal{H}} \in L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$ fulfills $\langle\, \mathcal{A}_{\mathcal{H}}(t)h, h\,\rangle \geq \kappa\|h\|_{\mathcal{H}}^2$ at almost every time-point $t \in [0,T]$ with a uniform constant $\kappa \in \mathbb{R}$. Then the operator DAE*

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(t)u(t)) + (\mathcal{A}(t) + \mathcal{A}_{\mathcal{H}}(t) - \tfrac{1}{2}\dot{\mathcal{M}}(t))u(t) \;-\; \mathcal{B}^*\lambda(t) &= f(t) \qquad \text{in } \mathcal{V}^*, \\ \mathcal{B}u(t) &= g(t) \qquad \text{in } \mathcal{Q}^* \end{aligned}$$

has a unique solution (u, λ), which satisfies the assertions a)–d) and the initial conditions from Theorem 7.21. The solution map $(f, g, u_0) \mapsto (u, \Lambda)$ is linear and continuous.

Proof. With the trick of Remark 7.1 we can assume $\kappa = 0$ without loss of generality. Then the proof follows the lines of Theorems 7.14, 7.19, and 7.21. Therein, the temporal discretization of $\mathcal{A}_{\mathcal{H}}$ is given by its means, cf. (7.9), and the bound (7.14) of $D_\tau(\mathcal{M}u_{\ker})_n$ has the additional term $\frac{C^2_{\mathcal{H}^*\hookrightarrow\mathcal{V}^*}}{4\mu_{\mathcal{M}}}\|\mathcal{A}_{\mathcal{H}}\|^2_{L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]}M^2(u_{\ker,0}, f_{\ker})$; cf. estimate (7.16). □

7.1.4. Regularity of Solutions

In comparison to Theorem 7.14, Theorem 7.21 proves the existence and uniqueness of a solution under mildly weaker assumptions on the right-hand side. In this subsection we consider the regularity of the solution under stronger assumptions on the data and the operators. For this we adapt the two cases from Section 6.2.

For the first approach we look at the operator DAE and the initial value, which are formally satisfied by the generalized time derivative of (u, λ). Instead of the operator DAE (7.4) we therefore consider its reformulation (4.31). If we denote the formal derivative of (u, λ) by (w, μ), this ansatz leads to the operator DAE

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}(t)w(t)) + (\mathcal{A}(t) + \tfrac{1}{2}\dot{\mathcal{M}}(t))w(t) - \mathcal{B}^*\mu(t) &= \dot{f}(t) - (\dot{\mathcal{A}}(t) + \tfrac{1}{2}\ddot{\mathcal{M}}(t))u(t) \qquad \text{in } \mathcal{V}^*, \qquad (7.32\text{a}) \\ \mathcal{B}w(t) &= \dot{g}(t) \qquad \text{in } \mathcal{Q}^*. \qquad (7.32\text{b}) \end{aligned}$$

By condition (7.32b) the initial value w_0 of w must be an element of $\mathcal{H}_{\ker} + \mathcal{B}^-\dot{g}(0) \subset \mathcal{H}$. Furthermore, since w_0 is formally the value of $\dot{u}$ at $t = 0$, the initial value w_0 has to satisfy

$$\langle\, \mathcal{M}(0)w_0, v_{\ker}\,\rangle_{\mathcal{H}^*,\mathcal{H}} = \langle\, f(0) - (\mathcal{A}(0) + \tfrac{1}{2}\dot{\mathcal{M}}(0))u_0, v_{\ker}\,\rangle_{\mathcal{V}^*,\mathcal{V}} \qquad (7.33)$$

for all $v_{\ker} \in \mathcal{V}_{\ker}$; cf. Section 6.2.

Theorem 7.23 (Regularity of Solutions I)**.** *Let $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ satisfy an inf-sup condition (3.2), $\mathcal{A} \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ be uniformly elliptic on $\mathcal{V}_{\ker}$, and $\mathcal{M} \in H^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$ be pointwise self-adjoint and uniformly elliptic. Suppose that the right-hand sides fulfill $f \in H^1(0,T;\mathcal{V}^*) + W^{1,1}(0,T;\mathcal{H}^*)$ and $g \in W^{2,1}(0,T;\mathcal{Q}^*)$. Assume that $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}u_0 = g(0)$, and that a $w_0 \in \mathcal{H}_{\ker} + \mathcal{B}^-\dot{g}(0)$ exists which satisfies (7.33). Then the solution of the operator DAE (7.4) fulfills*

$$u \in H^1(0,T;\mathcal{V}) \cap C^1([0,T],\mathcal{H}), \qquad \lambda \in C([0,T],\mathcal{Q}).$$

The solution (u, λ) depends linearly and continuously on the data (f, g, u_0, w_0).

Proof. We note that $\mathcal{A} + \frac{3}{2}\dot{\mathcal{M}}$ satisfies uniformly a Gårding inequality on $\mathcal{V}_{\mathrm{ker}}$, since

$$\langle (\mathcal{A}(t) + \tfrac{3}{2}\dot{\mathcal{M}}(t))v_{\mathrm{ker}}, v_{\mathrm{ker}} \rangle \geq \mu_{\mathcal{A}} \|v_{\mathrm{ker}}\|_{\mathcal{V}}^2 - \tfrac{3}{2}\|\dot{\mathcal{M}}\|_{C([0,T],\mathscr{L}(\mathcal{H},\mathcal{H}^*))}\|v_{\mathrm{ker}}\|_{\mathcal{H}}^2$$

for all $v_{\mathrm{ker}} \in \mathcal{V}_{\mathrm{ker}}$ at almost every time-point $t \in [0,T]$. Further, $(\dot{\mathcal{A}} + \frac{1}{2}\ddot{\mathcal{M}})u$ is an element of $L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$ by Theorem 7.21 and Lemma 4.9. By Remark 7.1 and Theorem 7.21 the operator DAE (7.4) with initial value w_0 has a unique solution (w,μ) satisfying a) and d) from Theorem 7.21.

We follow the steps of the proof of Theorem 6.8 and show $\dot{u} = w$. For this we split $u = u_{\mathrm{ker}} + u_{\mathrm{c}} = u_{\mathrm{ker}} + \mathcal{B}_{\mathcal{V}_{\mathrm{c}}}^{-} g$ and $w = w_{\mathrm{ker}} + w_{\mathrm{c}} = w_{\mathrm{ker}} + \mathcal{B}_{\mathcal{V}_{\mathrm{c}}}^{-}\dot{g}$ with an arbitrary direct sum $\mathcal{V} = \mathcal{V}_{\mathrm{ker}} \oplus \mathcal{V}_{\mathrm{c}}$. Then $\dot{u}_{\mathrm{c}} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{B}_{\mathcal{V}_{\mathrm{c}}}^{-} g) = \mathcal{B}_{\mathcal{V}_{\mathrm{c}}}^{-}\dot{g} = w_{\mathrm{c}}$. For the part in $\mathcal{V}_{\mathrm{ker}}$ we consider

$$v_{\mathrm{ker}} := u_{\mathrm{ker},0} + \int_0^{\cdot} w_{\mathrm{ker}} \,\mathrm{d}s - u_{\mathrm{ker}} \in L^2(0,T;\mathcal{V}_{\mathrm{ker}}) \cap C([0,T],\mathcal{H}_{\mathrm{ker}}) \tag{7.34}$$

with $u_{\mathrm{ker},0} = u_0 - \mathcal{B}_{\mathcal{V}_{\mathrm{c}}}^{-} g(0) \in \mathcal{V}_{\mathrm{ker}}$. With $w_{\mathrm{ker},0} = w_0 - \mathcal{B}_{\mathcal{V}_{\mathrm{c}}}^{-}\dot{g}(0) \in \mathcal{H}_{\mathrm{ker}}$ the following equation holds in $\mathcal{V}_{\mathrm{ker}}^*$:

$$\begin{aligned}
&\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}v_{\mathrm{ker}}) + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})v_{\mathrm{ker}} \\
\overset{(7.34)}&{=} \dot{\mathcal{M}}u_{\mathrm{ker},0} + \dot{\mathcal{M}}\int_0^{\cdot} w_{\mathrm{ker}}\,\mathrm{d}s + \mathcal{M}w_{\mathrm{ker}} - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{ker}}) + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})(u_{\mathrm{ker},0} + \int_0^{\cdot} w_{\mathrm{ker}}\,\mathrm{d}s - u_{\mathrm{ker}}) \\
&= (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})\Big(u_{\mathrm{ker},0} + \int_0^{\cdot} w_{\mathrm{ker}}\,\mathrm{d}s\Big) + \mathcal{M}(0)w_{\mathrm{ker},0} + \int_0^{\cdot} \tfrac{\mathrm{d}}{\mathrm{d}s}(\mathcal{M}w_{\mathrm{ker}})\,\mathrm{d}s - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{ker}}) - (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{ker}} \\
\overset{(7.32)}&{=} (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{ker},0} + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})\int_0^{\cdot} w_{\mathrm{ker}}\,\mathrm{d}s + \mathcal{M}(0)w_{\mathrm{ker},0} \\
&\quad + \int_0^{\cdot} \dot{f} - (\dot{\mathcal{A}} + \tfrac{1}{2}\ddot{\mathcal{M}})u - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}w_{\mathrm{c}}) - (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})w\,\mathrm{d}s - f + \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{c}}) + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} \\
\overset{(7.33)}&{=} (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{ker},0} - (\mathcal{A}(0) + \tfrac{1}{2}\dot{\mathcal{M}}(0))u_0 + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})\int_0^{\cdot} w_{\mathrm{ker}}\,\mathrm{d}s \\
&\quad - \int_0^{\cdot} (\dot{\mathcal{A}} + \tfrac{1}{2}\ddot{\mathcal{M}})u + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})w\,\mathrm{d}s - \mathcal{M}w_{\mathrm{c}} + \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u_{\mathrm{c}}) + (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} \\
&= (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{ker},0} - (\mathcal{A}(0) + \tfrac{1}{2}\dot{\mathcal{M}}(0))u_{\mathrm{ker},0} + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})\int_0^{\cdot} w_{\mathrm{ker}}\,\mathrm{d}s \\
&\quad - \int_0^{\cdot} (\dot{\mathcal{A}} + \tfrac{1}{2}\ddot{\mathcal{M}})u + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})w\,\mathrm{d}s + (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} - (\mathcal{A}(0) + \tfrac{1}{2}\dot{\mathcal{M}}(0))u_{\mathrm{c}}(0) \\
&= \int_0^{\cdot} (\dot{\mathcal{A}} + \tfrac{1}{2}\ddot{\mathcal{M}})(u_{\mathrm{ker},0} + \int_0^{s} w_{\mathrm{ker}}\,\mathrm{d}\eta - u) - (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})w_{\mathrm{c}} + \tfrac{\mathrm{d}}{\mathrm{d}s}((\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}})\,\mathrm{d}s \\
\overset{(7.34)}&{=} \int_0^{\cdot} (\dot{\mathcal{A}} + \tfrac{1}{2}\ddot{\mathcal{M}})v_{\mathrm{ker}}\,\mathrm{d}s.
\end{aligned}$$

Therefore, v_{ker} satisfies the operator DAE (7.6) with right-hand side $\int_0^{\cdot}(\dot{\mathcal{A}} + \frac{1}{2}\ddot{\mathcal{M}})v_{\mathrm{ker}}\,\mathrm{d}s$ and initial value $v_{\mathrm{ker}}(0) = 0$, which directly follows from (7.34). Especially, we get the bound

$$\begin{aligned}
&\|v_{\mathrm{ker}}\|_{C([0,t],\mathcal{H})}^2 + \int_0^t \|v_{\mathrm{ker}}\|_{\mathcal{V}}^2\,\mathrm{d}s \\
\overset{(7.26)}&{\leq} \frac{8\mu_{\mathcal{A}} + 4\mu_{\mathcal{M}}}{\mu_{\mathcal{A}}^2\mu_{\mathcal{M}}} \int_0^t \Big\| \int_0^s (\dot{\mathcal{A}} + \tfrac{1}{2}\ddot{\mathcal{M}})v_{\mathrm{ker}}\,\mathrm{d}\eta \Big\|_{\mathcal{V}^*}^2 \mathrm{d}s
\end{aligned}$$

$$\leq \frac{8\mu_{\mathcal{A}} + 4\mu_{\mathcal{M}}}{\mu_{\mathcal{A}}^2 \mu_{\mathcal{M}}} \int_0^t 2s \int_0^s \|\dot{\mathcal{A}} v_{\ker}\|_{\mathcal{V}^*}^2 \,\mathrm{d}\eta + \frac{1}{2} \Big\| \int_0^s \ddot{\mathcal{M}} v_{\ker} \,\mathrm{d}\eta \Big\|_{\mathcal{V}^*}^2 \,\mathrm{d}s$$
$$\leq \frac{8\mu_{\mathcal{A}} + 4\mu_{\mathcal{M}}}{\mu_{\mathcal{A}}^2 \mu_{\mathcal{M}}} \max\Big(2T\|\dot{\mathcal{A}}\|_{L^\infty}^2, \frac{C_{\mathcal{V}\hookrightarrow\mathcal{H}}^2}{2}\|\ddot{\mathcal{M}}\|_{L^1}^2\Big) \int_0^t \|v_{\ker}\|_{C([0,s],\mathcal{H})}^2 + \int_0^s \|v_{\ker}\|_{\mathcal{V}}^2 \,\mathrm{d}\eta \,\mathrm{d}s$$

with the abbreviations L^∞ and L^1 for $L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ and $L^1[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$, respectively. Gronwall's Lemma 3.15 then implies $\|v_{\ker}\|_{C([0,t],\mathcal{H})}^2 + \int_0^t \|v_{\ker}\|_{\mathcal{V}}^2 \,\mathrm{d}s \leq 0$ for all $t \in [0,T]$. Therefore, $v_{\ker} = 0$, which proves $\dot{u}_{\ker} = w_{\ker}$ by (7.34) and so $\dot{u} = w$. The continuous dependence of $u(t) = u_0 + \int_0^t w(s)\,\mathrm{d}s$ on the data follows by the continuity of $(\dot{f} - (\dot{\mathcal{A}} + \frac{1}{2}\ddot{\mathcal{M}})u, \dot{g}, w_0) \mapsto w$; see Remark 7.1 and Theorem 7.21. Finally, the assertion on λ holds because

$$\lambda \overset{(7.4a)}{=} \mathcal{B}_{\text{left}}^{-*}(f - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - (\mathcal{A} - \tfrac{1}{2}\dot{\mathcal{M}})u) = \mathcal{B}_{\text{left}}^{-*}(f - (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u - \mathcal{M}w) \in C([0,T],\mathcal{Q}). \qquad \square$$

For the second approach we assume that $\mathcal{A}$ can be split into time-dependent operators $\mathcal{A}_1$ and $\mathcal{A}_2$ similar as in Theorem 6.9. The associated proof needs a specific splitting of the solution $u = u_{\ker} + u_{\mathrm{c}}$ such that $u_{\ker}(t) \in \mathcal{V}_{\ker}$ and $\mathcal{A}_1(t)u_{\mathrm{c}}(t) \in \mathcal{V}_{\ker}^0$ at almost every time-point $t \in [0,T]$. This kind of time-dependent splittings of $\mathcal{V}$ are studied in Subsection 7.2.1. In particular, we prove there that u_{c} inherits the smoothness of $\mathcal{A}_1$, g, and $\mathcal{B}$. Therefore, $\mathcal{A}_1$ must be differentiable.

Theorem 7.24 (Regularity of Solutions II). *Let $\mathcal{B} \in \mathscr{L}(\mathcal{V},\mathcal{Q}^*)$ satisfy an inf-sup condition (3.2) and Assumptions 7.2.ii), 7.2.iii) as well as $f \in L^2(0,T;\mathcal{H}^*)$ be fulfilled. Suppose that $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ where $\mathcal{A}_1 \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ is pointwise self-adjoint and uniformly elliptic on $\mathcal{V}_{\ker}$ and $\mathcal{A}_2 \in L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{H}^*)]$. Assume that $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}u_0 = g(0)$. Then the solution of the operator DAE (7.4) satisfies*

$$u \in C([0,T],\mathcal{V}) \cap H^1(0,T;\mathcal{H}), \qquad \lambda \in L^2(0,T;\mathcal{Q}).$$

The map $(f, g, u_0) \mapsto (u, \lambda)$ is linear and continuous.

Proof. In order to avoid confusion with the temporal discretization, we use in this proof the notation $\mathcal{A}^{[1]}$ and $\mathcal{A}^{[2]}$ instead of $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively. We split the solution u into $u_{\ker} + u_{\mathrm{c}}$. For the part u_{c} we consider the time-dependent saddle point problem

$$\mathcal{A}^{[1]}(t)u_{\mathrm{c}}(t) - \mathcal{B}^*\nu(t) = 0 \quad \text{in } \mathcal{V}^*, \tag{7.35a}$$
$$\mathcal{B}u_{\mathrm{c}}(t) = g(t) \quad \text{in } \mathcal{Q}^*. \tag{7.35b}$$

Theorem 7.27 and Lemma 4.15 then guarantee a unique solution $u_{\mathrm{c}} \in H^1(0,T;\mathcal{V})$ with

$$\|u_c\|_{H^1(0,T;\mathcal{V})}^2 \lesssim \|g\|_{H^1(0,T;\mathcal{Q}^*)}^2.$$

Here, the constant dependent on $\mathcal{A}^{[1]}$ and $\mathcal{B}$. By (7.4) and (7.35b) the part $u_{\ker}$ satisfies the operator DAE (7.6) with right-hand side $f_{\ker}$ as in (7.7). Note that we have

$$f_{\ker} = f - (\mathcal{A} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} - \mathcal{M}\dot{u}_{\mathrm{c}} = f - (\mathcal{A}^{[2]} + \tfrac{1}{2}\dot{\mathcal{M}})u_{\mathrm{c}} - \mathcal{M}\dot{u}_{\mathrm{c}} \in L^2(0,T;\mathcal{H}_{\ker}^*),$$

since the equality $\mathcal{A}^{[1]}(t)u_{\mathrm{c}}(t) = \mathcal{B}^*\nu(t) \in \mathcal{V}_{\ker}^0$ is satisfied by (7.35a) at almost every time-point $t \in [0,T]$. To prove the assertion we consider again the discretization (7.8) of (7.6). Instead of $\mathcal{A}_n$ as integral mean, cf. (7.9), we define $\mathcal{A}_n := \mathcal{A}_n^{[1]} + \mathcal{A}_n^{[2]}$ where $\mathcal{A}_n^{[1]} := \mathcal{A}^{[1]}(t_n)$ and $\mathcal{A}_n^{[2]}$ is the mean of $\mathcal{A}^{[2]}$ over $[t_{n-1}, t_n]$. Note that by the continuity of $\mathcal{A}^{[1]}$ the operator $\mathcal{A}_n$ satisfies

$$\langle \mathcal{A}_n v_{\ker}, v_{\ker} \rangle \overset{(7.11)}{\geq} \mu_{\mathcal{A}}\|v_{\ker}\|_{\mathcal{V}}^2 + \tfrac{1}{\tau}\langle (\textstyle\int_{t_{n-1}}^{t_n} \mathcal{A}^{[1]}(t_n) - \mathcal{A}^{[1]}(s)\,\mathrm{d}s) v_{\ker}, v_{\ker} \rangle \geq \tfrac{\mu_{\mathcal{A}}}{2}\|v_{\ker}\|_{\mathcal{V}}^2$$

for all $v_{\ker} \in \mathcal{V}_{\ker}$, if τ is small enough. Furthermore, by the continuity of $\mathcal{A}^{[1]}$ the piecewise constant operator-valued function $\mathcal{A}^{[1]}_\tau$ defined by $\mathcal{A}^{[1]}_\tau(t) \equiv \mathcal{A}^{[1]}_n$ on $[t_{n-1}, t_n)$ converges strongly to $\mathcal{A}^{[1]}$ in $L^\infty(0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*))$ as $\tau \to 0$. Therefore, the results of Subsection 7.1.1 still hold for τ small enough and $\mu_{\mathcal{A}}/2$ instead of $\mu_{\mathcal{A}}$.

Since $u_{\ker,0} = u_0 - u_c(0) \in \mathcal{V}_{\ker}$, the discrete difference $D_\tau u_{\ker,n} := \frac{1}{\tau}(u_{\ker,n} - u_{\ker,n-1})$ is an element of $\mathcal{V}_{\ker}$ by Lemma 7.8 for every $n = 1, \dots, N$. Therefore, we can test equation (7.8a) with $\tau D_\tau u_{\ker,n} = u_{\ker,n} - u_{\ker,n-1}$, which leads to

$$
\begin{aligned}
&2\tau \sum_{k=1}^n \|D_\tau u_{\ker,k}\|^2_{\mathcal{M}_{k-1}} + \|u_{\ker,n}\|^2_{\mathcal{A}^{[1]}_n} - \|u_{\ker,0}\|^2_{\mathcal{A}^{[1]}_1} + \tau \sum_{k=1}^n \|u_{\ker,k} - u_{\ker,k-1}\|^2_{\mathcal{A}^{[1]}_k} \\
&\qquad + \sum_{k=1}^n \tau\langle D_\tau \mathcal{M}_k u_{\ker,k}, D_\tau u_{\ker,k}\rangle + \sum_{k=1}^{n-1} \|u_{\ker,k}\|^2_{\mathcal{A}^{[1]}_k} - \|u_{\ker,k}\|^2_{\mathcal{A}^{[1]}_{k+1}} \\
&\overset{(7.17)}{=} 2\sum_{k=1}^n \langle \mathcal{M}_{k-1} D_\tau u_{\ker,k} + \tfrac12 D_\tau \mathcal{M}_k u_{\ker,k} + \mathcal{A}^{[1]}_k u_{\ker,k}, \tau D_\tau u_{\ker,k}\rangle \\
&= 2\sum_{k=1}^n \langle D_\tau (\mathcal{M} u_{\ker})_k - \tfrac12 D_\tau \mathcal{M}_k u_{\ker,k} + \mathcal{A}^{[1]}_k u_{\ker,k}, \tau D_\tau u_{\ker,k}\rangle \\
&\overset{(7.8a)}{=} 2\sum_{k=1}^n \tau\langle f_{\ker,\tau} - \mathcal{A}^{[2]}_k u_{\ker,k}, D_\tau u_{\ker,k}\rangle \\
&\overset{(3.8)}{\le} \sum_{k=1}^n \frac{4\tau}{\mu_{\mathcal{M}}} \|f_{\ker,k}\|^2_{\mathcal{H}^*_{\ker}} + \sum_{k=1}^n \frac{4\tau}{\mu_{\mathcal{M}}} \|\mathcal{A}^{[2]}_k u_{\ker,k}\|^2_{\mathcal{H}^*} + \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} \\
&\overset{(7.9)}{\le} \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \frac{4}{\mu_{\mathcal{M}}} \|f_{\ker}\|_{\mathcal{H}^*_{\ker}} \,\mathrm{d}s + \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \frac{4}{\mu_{\mathcal{M}}} \|\mathcal{A}^{[2]}_k u_{\ker,k}\|^2_{\mathcal{H}^*} \,\mathrm{d}s + \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} \qquad (7.36) \\
&\le \frac{4}{\mu_{\mathcal{M}}} \int_0^T \|f_{\ker}\|^2_{\mathcal{H}^*_{\ker}} \,\mathrm{d}s + \frac{4}{\mu_{\mathcal{M}}} \|\mathcal{A}^{[2]}\|^2_{L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{H}^*)]} \sum_{k=1}^n \tau \|u_{\ker,k}\|^2_{\mathcal{V}} + \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}}.
\end{aligned}
$$

Further, we observe with $M := M(u_{\ker,0}, f_{\ker})$ from Theorem 7.9 that the first sum in the second line of (7.36) is bounded by

$$
\begin{aligned}
&-\sum_{k=1}^n \tau\langle D_\tau \mathcal{M}_k u_{\ker,k}, D_\tau u_{\ker,k}\rangle \\
&\overset{(3.8)}{\le} \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} + \frac{\tau}{2\mu_{\mathcal{M}}} \|D_\tau \mathcal{M}_k u_{\ker,k}\|^2_{\mathcal{H}^*} \\
&\overset{(7.12)}{\le} \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} + \frac{1}{2\mu_{\mathcal{M}}\tau} \Big\| \int_{t_{k-1}}^{t_k} \dot{\mathcal{M}} \,\mathrm{d}s \Big\|^2_{\mathscr{L}(\mathcal{H},\mathcal{H}^*)} \|u_{\ker,k}\|^2_{\mathcal{H}} \\
&\le \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} + \frac{1}{2\mu_{\mathcal{M}}} \int_0^{t_n} \|\dot{\mathcal{M}}\|^2_{\mathscr{L}(\mathcal{H},\mathcal{H}^*)} \,\mathrm{d}s \max_{k=1,\dots,n} \|u_{\ker,k}\|^2_{\mathcal{H}} \\
&\overset{(7.13)}{\le} \sum_{k=1}^n \frac{\tau\mu_{\mathcal{M}}}{2} \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} + \frac{1}{2\mu^2_{\mathcal{M}}} \|\dot{\mathcal{M}}\|^2_{L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]} M^2 \qquad (7.37)
\end{aligned}
$$

and the second sum is bounded by

$$\begin{aligned}\sum_{k=1}^{n-1} \|u_{\ker,k}\|^2_{\mathcal{A}^{[1]}_{k+1}} - \|u_{\ker,k}\|^2_{\mathcal{A}^{[1]}_{k}} &= \sum_{k=1}^{n-1} \langle\, (\mathcal{A}^{[1]}(t_{k+1}) - \mathcal{A}^{[1]}(t_k)) u_{\ker,k}, u_{\ker,k} \,\rangle \\ &\overset{(7.12)}{\leq} \sum_{k=1}^{n-1} \tau \|\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{[1]}\|_{L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]} \|u_{\ker,k}\|^2_{\mathcal{V}} \\ &\overset{(7.13)}{\leq} \frac{2}{\mu_{\mathcal{A}}} \|\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{[1]}\|_{L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]} M^2. \end{aligned} \tag{7.38}$$

Theorem 7.9, the uniform ellipticity of $\mathcal{M}$ and $\mathcal{A}^{[1]}$, and the estimates (7.36)–(7.38) then imply

$$\sum_{k=1}^{n} \tau \|D_\tau u_{\ker,k}\|^2_{\mathcal{H}} + \|u_{\ker,n}\|^2_{\mathcal{V}} + \sum_{k=1}^{n} \tau \|u_{\ker,k} - u_{\ker,k-1}\|^2_{\mathcal{V}} \lesssim \|u_{\ker,0}\|^2_{\mathcal{V}} + \int_0^T \|f_{\ker}\|^2_{\mathcal{H}^*_{\ker}} \,\mathrm{d}s + M^2 \tag{7.39}$$

for all $n = 1, \dots, N$ independently of τ, if τ is small enough. In particular, this shows that $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\ker,\tau}$ and $u_{\ker,\tau}$ are bounded in $L^2(0,T;\mathcal{H})$ and, respectively, $L^\infty(0,T;\mathcal{V})$ independent of τ and therefore weakly and, respectively, weakly* convergent. By Lemma 7.11, an argument similar as in the proof of Lemma 7.12, and the uniqueness of the solution $u_{\ker}$ we get

$$u_{\ker,\tau} \overset{*}{\rightharpoonup} u_{\ker} \quad \text{in } L^\infty(0,T;\mathcal{V}), \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\ker,\tau} \rightharpoonup \dot{u}_{\ker} \quad \text{in } L^2(0,T;\mathcal{H})$$

as $\tau \to 0$. Theorem 7.16 with the choices $\mathcal{X} = \mathcal{V}_{\ker}$, $\mathcal{Y} = \mathcal{H}_{\ker}$, $\mathcal{K} = \mathcal{A}$, $\mathcal{Z} = L^2(0,T;\mathcal{V}_{\ker})$, and $u_{\ker} = w$ then implies $u_{\ker} \in C([0,T], \mathcal{V}_{\ker})$ and

$$\begin{aligned} &\mu_{\mathcal{M}} \int_0^t \|\dot{u}_{\ker}\|^2_{\mathcal{H}} \,\mathrm{d}s + \mu_{\mathcal{A}^{[1]}} \|u_{\ker}(t)\|^2_{\mathcal{V}} \\ &\leq \int_0^t 2\langle\, \mathcal{M}\dot{u}_{\ker}, \dot{u}_{\ker} \,\rangle \,\mathrm{d}s + \langle\, \mathcal{A}^{[1]}(t) u_{\ker}(t), u_{\ker}(t) \,\rangle - \mu_{\mathcal{M}} \int_0^t \|\dot{u}_{\ker}\|^2_{\mathcal{H}} \,\mathrm{d}s \\ &\overset{(7.25)}{=} \|u_{\ker,0}\|^2_{\mathcal{A}^{[1]}(0)} + 2\int_0^t \langle\, \mathcal{M}\dot{u}_{\ker} + \mathcal{A}^{[1]} u_{\ker}, \dot{u}_{\ker} \,\rangle + \langle\, \tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{[1]} u_{\ker}, u_{\ker} \,\rangle \,\mathrm{d}s - \mu_{\mathcal{M}} \int_0^t \|\dot{u}_{\ker}\|^2_{\mathcal{H}} \,\mathrm{d}s \\ &\overset{(7.6a)}{=} \|u_{\ker,0}\|^2_{\mathcal{A}^{[1]}(0)} + 2\int_0^t \langle\, f_{\ker} - (\mathcal{A}^{[2]} - \tfrac12 \dot{\mathcal{M}}) u_{\ker}, \dot{u}_{\ker} \,\rangle + \langle\, \tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{[1]} u_{\ker}, u_{\ker} \,\rangle \,\mathrm{d}s - \mu_{\mathcal{M}} \int_0^t \|\dot{u}_{\ker}\|^2_{\mathcal{H}} \,\mathrm{d}s \\ &\overset{(3.8)}{\leq} \|u_{\ker,0}\|^2_{\mathcal{A}^{[1]}(0)} + \frac{3}{\mu_{\mathcal{M}}} \int_0^t \|f_{\ker}\|^2_{\mathcal{H}^*_{\ker}} \,\mathrm{d}s + \frac{3}{4\mu_{\mathcal{M}}} \|\dot{\mathcal{M}}\|^2_{L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]} \max_{s\in[0,t]} \|u_{\ker}\|^2_{\mathcal{H}} \\ &\qquad + \Big(\frac{3}{\mu_{\mathcal{M}}} \|\mathcal{A}^{[2]}\|^2_{L^\infty[0,t;\mathscr{L}(\mathcal{V},\mathcal{H}^*)]} + \|\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}^{[1]}\|_{L^\infty[0,t;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]} \Big) \int_0^t \|u_{\ker}\|^2_{\mathcal{V}} \,\mathrm{d}s. \end{aligned}$$

Since the right-hand side of the previous estimate is bounded by Theorem 7.19, the linear mapping $(f_{\ker}, u_{\ker,0}) \mapsto u_{\ker}$ is continuous and so is $(f, g, u_0) \mapsto u$ as well. The assertion on λ follows by

$$\lambda \overset{(7.4a)}{=} \mathcal{B}^{-*}_{\text{left}}(f - \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - (\mathcal{A} - \tfrac12 \dot{\mathcal{M}})u) = \mathcal{B}^{-*}_{\text{left}}(f - \mathcal{M}\dot{u} - (\mathcal{A} + \tfrac12 \dot{\mathcal{M}})u) \in L^2(0,T;\mathcal{Q}). \qquad \square$$

Remark 7.25. The assumption of Theorem 7.24 can be weakened to $f = f_1 + f_2$ with $f_1 \in L^2(0,T;\mathcal{H}^*)$ and $f_2 \in W^{1,1}(0,T;\mathcal{V}^*)$. The idea of the associated proof is the same as in Remark 6.10.

Remark 7.26. The results of Theorem 7.24 can be extended to semi-linear operator DAEs (7.4) with a nonlinear $f\colon [0,T] \times \mathcal{V} \to \mathcal{H}^*$. The right-hand side f should satisfy the same assumption as in

Section 6.4; see page 51. The existence and uniqueness of the solution as well as the continuous dependence on the data (g, u_0) follows along the lines of Theorem 6.15 and 6.19.

7.2. Constraints with Time-Dependent Operators

After considering in Section 7.1 operator DAEs with a constant operator $\mathcal{B}$, we now investigate systems with a time-dependent $\mathcal{B}\colon [0,T] \to \mathscr{L}(\mathcal{V},\mathcal{Q}^*)$. We assume that $\mathcal{B} \in H^1[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$ is uniformly inf-sup stable, i.e., Assumption 7.2.i) is satisfied. To concentrate on the difficulties, which come with a time-dependent operator $\mathcal{B}$, we restrict the analysis to operator DAEs (7.1) with $\mathcal{M}(t) \equiv \mathcal{R}_{\mathcal{H}}$, i.e.,

$$\dot{u}(t) + \mathcal{A}(t)u(t) - \mathcal{B}^*(t)\lambda(t) = f(t) \qquad \text{in } \mathcal{V}^*, \tag{7.40a}$$

$$\mathcal{B}(t)u(t) = g(t) \qquad \text{in } \mathcal{Q}^*. \tag{7.40b}$$

All results obtained in this section hold also for the operator DAEs (7.1) if $\mathcal{M}$ satisfies Assumption 7.2.ii); see Section 7.3.

Since $\mathcal{B}$ is time-dependent, its pointwise kernel is in general not constant. In particular, the usual splitting $\mathcal{V} = \mathcal{V}_{\text{ker}} \oplus \mathcal{V}_{\text{c}}$ with $\mathcal{V}_{\text{ker}} = \ker \mathcal{B}$ changes over time. Therefore, we investigate time-dependent splittings of $\mathcal{V}$ in Subsection 7.2.1. Especially, we define an operator-valued function $\mathcal{W}\colon [0,T] \to \mathscr{L}(\mathcal{V})$, which maps $\ker \mathcal{B}(0)$ to $\ker \mathcal{B}(t)$. With this map we derive the existence of a solution of the operator DAE (7.40) via a temporal discretization by the implicit Euler scheme in Subsection 7.2.2. Since $\mathcal{V}_{\text{ker}}$ is time-dependent, Theorem 7.18, which implied the uniqueness of the solution in Section 7.1, is not applicable for the operator DAE (7.40). In Subsection 7.2.3 we discuss additional assumptions which guarantee unique solutions.

7.2.1. Dynamical Splitting

In this subsection we consider splittings of $\mathcal{V} = \mathcal{V}_{\text{ker}} \oplus \mathcal{V}_{\text{c}}$ where $\mathcal{V}_{\text{ker}}$ and $\mathcal{V}_{\text{c}}$ depend on time. We assume that the space $\mathcal{V}_{\text{ker}}$ is the pointwise kernel of a uniform inf-sup stable $\mathcal{B} \in W^{1,p}[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$, $p \in [1,\infty]$. Note that the analysis of (7.40) requires only the case $p = 2$. However, the investigation for general p needs no additional effort.

By the assumptions on $\mathcal{B}$ and Lemma 4.14, the operator-valued function $\mathcal{B}$ has a continuous representative. Therefore,

$$\mathcal{V}_{\text{ker}}(t) := \ker \mathcal{B}(t) \tag{7.41}$$

is well-defined at almost every time-point $t \in [0,T]$. Its pointwise dual space and orthogonal complement are denoted by

$$\mathcal{V}^*_{\text{ker}}(t) := (\mathcal{V}_{\text{ker}}(t))^* \quad \text{and} \quad \mathcal{V}^{\perp}_{\text{ker}}(t) := (\mathcal{V}_{\text{ker}}(t))^{\perp},$$

respectively. In addition to the operator $\mathcal{B}$, we assume that an operator $\mathcal{A} \in W^{1,p}[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ exists which is uniformly eliptic on $\mathcal{V}_{\text{ker}}$. This implies that

$$\mathcal{V}_{\text{c}}(t) := \{v \in \mathcal{V} \mid \langle \mathcal{A}(t)v, v_{\text{ker}} \rangle = 0 \text{ for all } v_{\text{ker}} \in \mathcal{V}_{\text{ker}}(t)\} \tag{7.42}$$

is well-defined at almost every time-point $t \in [0,T]$. Lemma 3.5 then shows $\mathcal{V} = \mathcal{V}_{\text{ker}}(t) \oplus \mathcal{V}_{\text{c}}(t)$ at almost every time-point $t \in [0,T]$. Moreover, the mappings $t \mapsto \mathcal{V}_{\text{ker}}(t)$ and $t \mapsto \mathcal{V}_{\text{c}}(t)$ inherit in some sense the regularity of $\mathcal{A}$ and $\mathcal{B}$; see Theorem 7.31.

In this subsection, we will show that the assumptions made on the operators $\mathcal{A}$ and $\mathcal{B}$ imply the existence of a pointwise right-inverse $\mathcal{B}^-_{\mathcal{A}}$ of $\mathcal{B}$ with $\operatorname{im} \mathcal{B}^-_{\mathcal{A}}(t) = \mathcal{V}_{\text{c}}(t)$ for almost all $t \in [0,T]$. Such a right-inverse was already used in the proof of Theorem 7.24. Furthermore, we analyze how $\mathcal{V}_{\text{ker}}$

and $\mathcal{V}_c$ change with respect to time. The starting point of our investigation is the saddle-point problem induced by $\mathcal{A}$ and $\mathcal{B}$.

Theorem 7.27 (Time-Dependent Saddle Point Problem)**.** *Let $p \in [1, \infty]$. Assume that $\mathcal{B} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)]$, $p \in [1, \infty]$, satisfies uniformly an inf-sup condition (3.2). Suppose that $\mathcal{V}_{\ker}(t)$ is defined as in (7.41), $\mathcal{A} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V}, \mathcal{V}^*)]$, and $\mathcal{A}(t)$ is elliptic on $\mathcal{V}_{\ker}(t)$ at almost every time-point $t \in [0, T]$ with a uniform ellipticity constant. Then the saddle point problem*

$$\mathcal{A}(t)x_t - \mathcal{B}^*(t)\nu_t = f \qquad \text{in } \mathcal{V}^*, \tag{7.43a}$$
$$\mathcal{B}(t)x_t = g \qquad \text{in } \mathcal{Q}^* \tag{7.43b}$$

is uniquely solvable at almost every time-point $t \in [0, T]$ for every $f \in \mathcal{V}^*$ and $g \in \mathcal{Q}^*$. Its pointwise solution operator $\mathcal{S}$ with $\mathcal{S}(t)\colon \mathcal{V}^* \times \mathcal{Q}^* \to \mathcal{V} \times \mathcal{Q}, (f, g) \mapsto (x_t, \nu_t)$ is an element of $W^{1,p}[0, T; \mathscr{L}(\mathcal{V}^* \times \mathcal{Q}^*, \mathcal{V} \times \mathcal{Q})]$.

Proof. By Lemma 4.14 we can consider the continuous representatives of $\mathcal{A}$ and $\mathcal{B}$. Then the pointwise linear solution operator $\mathcal{S}$ of (7.43) is well-defined and the solution (x_t, ν_t) satisfies $\|x_t\|_{\mathcal{V}}, \|\nu_t\|_{\mathcal{Q}} \lesssim \|f\|_{\mathcal{V}^*} + \|g\|_{\mathcal{Q}^*}$ with a constant which can be bounded independently of t; see Theorem 3.8. Let $(x_s, \nu_s), (x_t, \nu_t)$ be the solutions of (7.43) for arbitrary but fixed f, g at the time-points $s, t \in [0, T]$, respectively. The differences $\Delta x := x_t - x_s$ and $\Delta\nu := \nu_t - \nu_s$ then solve

$$\begin{aligned} \mathcal{A}(t)\Delta x - \mathcal{B}^*(t)\Delta\nu &= (\mathcal{A}(s) - \mathcal{A}(t))x_s - (\mathcal{B}^*(s) - \mathcal{B}^*(t))\nu_s \qquad &&\text{in } \mathcal{V}^*,\\ \mathcal{B}(t)\Delta x &= (\mathcal{B}(s) - \mathcal{B}(t))x_s &&\text{in } \mathcal{Q}^*. \end{aligned}$$

Therefore, we get the estimate

$$\begin{aligned} \|(\mathcal{S}(t) - \mathcal{S}(s))(f, g)\|_{\mathcal{V}\times\mathcal{Q}} &\lesssim \|\Delta x\|_{\mathcal{V}} + \|\Delta\nu\|_{\mathcal{Q}}\\ &\lesssim \|\mathcal{A}(t) - \mathcal{A}(s)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)}\|x_s\|_{\mathcal{V}} + \|\mathcal{B}(t) - \mathcal{B}(s)\|_{\mathscr{L}(\mathcal{V},\mathcal{Q}^*)}\big(\|x_s\|_{\mathcal{V}} + \|\nu_s\|_{\mathcal{Q}}\big)\\ &\lesssim \big(\|\mathcal{A}(t) - \mathcal{A}(s)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)} + \|\mathcal{B}(t) - \mathcal{B}(s)\|_{\mathscr{L}(\mathcal{V},\mathcal{Q}^*)}\big)\big(\|f\|_{\mathcal{V}^*} + \|g\|_{\mathcal{Q}^*}\big). \end{aligned}$$

Considering the supremum on the unit circle of $\mathcal{V}^* \times \mathcal{Q}^*$ shows $\mathcal{S} \in C([0, T], \mathscr{L}(\mathcal{V}^* \times \mathcal{Q}^*, \mathcal{V} \times \mathcal{Q}))$, because $\mathcal{A}$ and $\mathcal{B}$ are continuous. In particular, $x\colon t \mapsto x_t$ and $\nu\colon t \mapsto \nu_t$ are continuous functions with images in $\mathcal{V}$ and $\mathcal{Q}$, respectively. Furthermore, their formal derivatives satisfy

$$\begin{aligned} \mathcal{A}(t)\dot{x}(t) - \mathcal{B}^*(t)\dot{\nu}(t) &= -\dot{\mathcal{A}}(t)x(t) + \dot{\mathcal{B}}^*(t)\nu(t) \qquad &&\text{in } \mathcal{V}^*,\\ \mathcal{B}(t)\dot{x}(t) &= -\dot{\mathcal{B}}(t)x(t) &&\text{in } \mathcal{Q}^*, \end{aligned}$$

by (7.43) and Lemma 4.15. The right-hand sides are elements of $L^p(0, T; \mathcal{V}^*)$ and $L^p(0, T; \mathcal{Q}^*)$, respectively, and therefore $(\dot{x}, \dot{\nu}) \in L^p(0, T; \mathcal{V} \times \mathcal{Q})$ by Lemma 4.9, since $\mathcal{S}$ is continuous. Therefore, $\dot{\mathcal{S}}$ is strongly measurable with bound

$$\begin{aligned} \|\dot{\mathcal{S}}(t)(f, g)\|_{\mathcal{V}\times\mathcal{Q}} \lesssim \|\dot{x}(t)\| + \|\dot{\nu}(t)\|_{\mathcal{Q}} &\lesssim \|\dot{\mathcal{A}}(t)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)}\|x(t)\|_{\mathcal{V}} + \|\dot{\mathcal{B}}(t)\|_{\mathscr{L}(\mathcal{V},\mathcal{Q}^*)}\big(\|x(t)\|_{\mathcal{V}} + \|\nu(t)\|_{\mathcal{Q}}\big)\\ &\lesssim \big(\|\dot{\mathcal{A}}(t)\|_{\mathscr{L}(\mathcal{V},\mathcal{V}^*)} + \|\dot{\mathcal{B}}(t)\|_{\mathscr{L}(\mathcal{V},\mathcal{Q}^*)}\big)\big(\|f\|_{\mathcal{V}^*} + \|g\|_{\mathcal{Q}^*}\big). \end{aligned}$$

In particular, $\dot{\mathcal{S}}$ satisfies the boundedness condition (4.3). Lemma 4.7 finishes the proof. □

With the saddle-point problem (7.43) we calculate a right-inverse of $\mathcal{B}$ with images in $\mathcal{V}_c$.

Lemma 7.28 (Time-Dependent Right-Inverse)**.** *Let the assumptions of Theorem 7.27 be satisfied. Suppose that $\mathcal{V}_{\ker}$ and $\mathcal{V}_c$ are defined as in (7.41) and (7.42), respectively.*

Then $\mathcal{B}$ has a pointwise right-inverse $\mathcal{B}^-_{\mathcal{A}} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{Q}^*, \mathcal{V})]$ with $\mathcal{B}^-_{\mathcal{A}}(t)g \in \mathcal{V}_c(t)$ for every $g \in \mathcal{Q}^*$ at almost every time-point $t \in [0, T]$. Furthermore, the time-dependent projection $\mathcal{P}_{\mathcal{A}}\colon [0, T] \to$

$\mathscr{L}(\mathcal{V})$ with $\operatorname{im}\mathcal{P}_{\mathcal{A}}(t) = \mathcal{V}_{\mathrm{ker}}(t)$ and $\ker\mathcal{P}_{\mathcal{A}}(t) = \mathcal{V}_{\mathrm{c}}(t)$ at almost every time-point $t \in [0,T]$ is an element of $W^{1,p}[0,T;\mathscr{L}(\mathcal{V})]$.

Proof. By Corollary 3.9 we have that $\mathcal{B}^-_{\mathcal{A}}(t)g$ is the restriction of $\mathcal{S}(0,g)$ from Theorem 7.27 on x_t. Therefore, $\mathcal{B}^-_{\mathcal{A}} \in W^{1,p}[0,T;\mathscr{L}(\mathcal{Q}^*,\mathcal{V})]$ follows by Theorem 7.27 and Lemma 4.15. The pointwise projection with the stated image and kernel is given by

$$\mathcal{P}_{\mathcal{A}}(t) \coloneqq \operatorname{id}_{\mathcal{V}} - \mathcal{B}^-_{\mathcal{A}}(t)\mathcal{B}(t), \tag{7.44}$$

see e.g. [AltH18, Lem. 3.1 & Rem. 3.3]. Lemma 4.15 then implies $\mathcal{P}_{\mathcal{A}} \in W^{1,p}[0,T;\mathscr{L}(\mathcal{V})]$. □

Remark 7.29. If, in addition to the assumptions of Lemma 7.28, the operator-valued function $\mathcal{A}$ is uniformly elliptic on whole $\mathcal{V}$, then the right-inverse is given by $\mathcal{B}^-_{\mathcal{A}}(t) \equiv \mathcal{A}^{-1}(t)\mathcal{B}^*(t)(\mathcal{B}(t)\mathcal{A}^{-1}(t)\mathcal{B}^*(t))^{-1}$; see Remark 3.10. Especially, by Lemma 4.15 and Lemma 4.17 its derivative is given by

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{B}^-_{\mathcal{A}} &= -\mathcal{A}^{-1}\dot{\mathcal{A}}\mathcal{B}^-_{\mathcal{A}} + \mathcal{A}^{-1}\dot{\mathcal{B}}^*(\mathcal{B}\mathcal{A}^{-1}\mathcal{B}^*)^{-1} - \mathcal{B}^-_{\mathcal{A}}\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{B}\mathcal{A}^{-1}\mathcal{B}^*)(\mathcal{B}\mathcal{A}^{-1}\mathcal{B}^*)^{-1}\\ &= -\mathcal{B}^-_{\mathcal{A}}\dot{\mathcal{B}}\mathcal{B}^-_{\mathcal{A}} + (\operatorname{id}_{\mathcal{V}} - \mathcal{B}^-_{\mathcal{A}}\mathcal{B})\mathcal{A}^{-1}\dot{\mathcal{B}}^*(\mathcal{B}\mathcal{A}^{-1}\mathcal{B}^*)^{-1} - (\operatorname{id}_{\mathcal{V}} - \mathcal{B}^-_{\mathcal{A}}\mathcal{B})\mathcal{A}^{-1}\dot{\mathcal{A}}\mathcal{B}^-_{\mathcal{A}}.\end{aligned}$$

In Subsection 7.2.2 we transform the operator DAE (7.40) to a system where the kernel of $\mathcal{B}$ is constant. Especially, the derivative of $\mathcal{B}^-_{\mathcal{A}}$ becomes simpler, if the spaces $\mathcal{V}_{\mathrm{ker}} = \ker\mathcal{B}$ and $\mathcal{V}_{\mathrm{c}}$ are time-independent.

Lemma 7.30. *Suppose that the assumptions of Lemma 7.28 are satisfied. In addition, let $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{V}_{\mathrm{c}}$ be time-independent. Then the generalized derivative of $\mathcal{B}^-_{\mathcal{A}}$ is given by $-\mathcal{B}^-_{\mathcal{A}}\dot{\mathcal{B}}\mathcal{B}^-_{\mathcal{A}}$.*

Proof. Let $v \in \mathcal{V}$ be arbitrary. Since $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{V}_{\mathrm{c}}$ are constant and $v = v_{\mathrm{ker}} + v_{\mathrm{c}}$ with $v_{\mathrm{ker}} \in \mathcal{V}_{\mathrm{ker}}$, $v_{\mathrm{c}} \in \mathcal{V}_{\mathrm{c}}$ is unique, $v_{\mathrm{ker}} = \mathcal{P}_{\mathcal{A}}(t)v$ is time-independent. Therefore, $\dot{\mathcal{P}}_{\mathcal{A}} = 0$ and

$$0 = -\dot{\mathcal{P}}_{\mathcal{A}}\mathcal{B}^-_{\mathcal{A}} \overset{(7.44)}{=} \big(\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{B}^-_{\mathcal{A}})\mathcal{B} + \mathcal{B}^-_{\mathcal{A}}\dot{\mathcal{B}}\big)\mathcal{B}^-_{\mathcal{A}} = \tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{B}^-_{\mathcal{A}} + \mathcal{B}^-_{\mathcal{A}}\dot{\mathcal{B}}\mathcal{B}^-_{\mathcal{A}}. \qquad \square$$

We now show the existence of a differentiable operator-valued function, which describe the evolution of the spaces $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{V}_{\mathrm{c}}$ over time.

Theorem 7.31 (Tracking of $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{V}_{\mathrm{c}}$ over Time)**.** *Let the assumption of Theorem 7.27 be satisfied. Then there exists an operator-valued function $\mathcal{W} \in W^{1,p}[0,T;\mathscr{L}(\mathcal{V})]$, which is pointwise bijective at almost every time-point $t \in [0,T]$ and satisfies*

$$\mathcal{W}(t)\mathcal{V}_{\mathrm{ker}}(0) \coloneqq \operatorname{im}\mathcal{W}(t)\big|_{\mathcal{V}_{\mathrm{ker}}(0)} = \mathcal{V}_{\mathrm{ker}}(t) \quad \text{and} \quad \mathcal{W}(t)\mathcal{V}_{\mathrm{c}}(0) \coloneqq \operatorname{im}\mathcal{W}(t)\big|_{\mathcal{V}_{\mathrm{c}}(0)} = \mathcal{V}_{\mathrm{c}}(t).$$

Further, the pointwise inverse $\mathcal{W}^{-1}$ is an element of $W^{1,p}[0,T;\mathscr{L}(\mathcal{V})]$.

Proof. The idea of the proof is based on [KunM06, p. 80]. Let $v_0 \in \mathcal{V}$ be arbitrary and $v \in W^{1,p}(0,T;\mathcal{V})$ be the unique solution of the operator ODE

$$\dot{v}(t) = (\dot{\mathcal{P}}_{\mathcal{A}}\mathcal{P}_{\mathcal{A}} - \mathcal{P}_{\mathcal{A}}\dot{\mathcal{P}}_{\mathcal{A}})(t)v(t), \qquad v(0) = v_0; \tag{7.45}$$

see Theorem 4.20 and Lemma 7.28. We define the operator-valued function $\mathcal{W}$ via

$$\mathcal{W}\colon [0,T] \to \mathscr{L}(\mathcal{V}) \quad \text{with} \quad \mathcal{W}(t)v_0 = v(t). \tag{7.46}$$

By $v \in W^{1,p}(0,T;\mathcal{V})$ and (7.45) the function $\mathcal{W}$ is an element of $W^{1,p}[0,T;\mathscr{L}(\mathcal{V})]$ with derivative $\dot{\mathcal{W}} = (\dot{\mathcal{P}}_{\mathcal{A}}\mathcal{P}_{\mathcal{A}} - \mathcal{P}_{\mathcal{A}}\dot{\mathcal{P}}_{\mathcal{A}})\mathcal{W}$. We show that $\mathcal{W}(t)$ is bijective. Note that this is obvious for $\mathcal{W}(0) = \operatorname{id}_{\mathcal{V}}$. For an arbitrary but fixed $t' \in (0,T]$, we consider the uniquely solvable operator ODE

$$\tfrac{\mathrm{d}}{\mathrm{d}s}w(s) = -(\dot{\mathcal{P}}_{\mathcal{A}}\mathcal{P}_{\mathcal{A}} - \mathcal{P}_{\mathcal{A}}\dot{\mathcal{P}}_{\mathcal{A}})(t'-s)w(s), \qquad w(0) = w_0 \in \mathcal{V}, \tag{7.47}$$

$s \in [0, t']$ and define $\widetilde{\mathcal{W}}_{t'} \in \mathscr{L}(\mathcal{V})$ via $\widetilde{\mathcal{W}}_{t'} w_0 = w(t')$. Then $t \mapsto w(t'-t)$ with $t \in [0, t']$ satisfies

$$\tfrac{\mathrm{d}}{\mathrm{d}t} w(t'-t) = -\tfrac{\mathrm{d}}{\mathrm{d}s} w(t'-t) \overset{(7.47)}{=} (\dot{\mathcal{P}}_\mathcal{A}\mathcal{P}_\mathcal{A} - \mathcal{P}_\mathcal{A}\dot{\mathcal{P}}_\mathcal{A})(t) w(t'-t).$$

Thus, it solves (7.45) with initial value $w(t')$. Since the solution of (7.45) is unique, we have $w_0 = w(t'-t') = \mathcal{W}(t')w(t') = \mathcal{W}(t')\widetilde{\mathcal{W}}_{t'} w_0$ for every $w_0 \in \mathcal{V}$. Analogously, one shows $v_0 = \widetilde{\mathcal{W}}_{t'}\mathcal{W}(t)v_0$ for arbitrary $v_0 \in \mathcal{V}$. Therefore, $\mathcal{W}(t')\widetilde{\mathcal{W}}_{t'} = \widetilde{\mathcal{W}}_{t'}\mathcal{W}(t') = \mathrm{id}_\mathcal{V}$. Since $t' \in (0, T]$ was arbitrary this proves that $\mathcal{W}$ is pointwise bijective with pointwise inverse $\mathcal{W}^{-1}(t) := (\mathcal{W}(t))^{-1} = \widetilde{\mathcal{W}}_t$ for $t > 0$ and $\mathcal{W}^{-1}(0) = \mathrm{id}_\mathcal{V}$. By the operator ODE (7.47) and Theorem 4.20 the pointwise inverse $\mathcal{W}^{-1}(t)$ is bounded independently of t in terms of $\mathcal{P}_\mathcal{A}$. The assertion $\mathcal{W}^{-1} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V})]$ then follows by the steps of the proof of Lemma 4.17.

For the tracking of $\mathcal{V}_{\ker}(t)$ we note $\mathcal{P}_\mathcal{A}\dot{\mathcal{P}}_\mathcal{A}\mathcal{P}_\mathcal{A} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{P}_\mathcal{A}^3 - \mathcal{P}_\mathcal{A}^2) = \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{P}_\mathcal{A} - \mathcal{P}_\mathcal{A}) = 0$. In particular, this implies

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{P}_\mathcal{A}\mathcal{W}) = \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{P}_\mathcal{A}^2\mathcal{W}) = (\dot{\mathcal{P}}_\mathcal{A}\mathcal{P}_\mathcal{A} + \mathcal{P}_\mathcal{A}\dot{\mathcal{P}}_\mathcal{A})\mathcal{W} + \mathcal{P}_\mathcal{A}\dot{\mathcal{W}} \overset{(7.45)}{=} \dot{\mathcal{P}}_\mathcal{A}\mathcal{P}_\mathcal{A}\mathcal{W} = (\dot{\mathcal{P}}_\mathcal{A}\mathcal{P}_\mathcal{A} - \mathcal{P}_\mathcal{A}\dot{\mathcal{P}}_\mathcal{A})\mathcal{P}_\mathcal{A}\mathcal{W}.$$

Therefore, for every $v_0 \in \mathcal{V}$ the function $t \mapsto \mathcal{P}_\mathcal{A}(t)\mathcal{W}(t)v_0$ solves (7.45) with the initial value $\mathcal{P}_\mathcal{A}(0)\mathcal{W}(0)v_0 = \mathcal{P}_\mathcal{A}(0)v_0$. The same holds for $t \mapsto \mathcal{W}(t)\mathcal{P}_\mathcal{A}(0)v_0$. Since the solution of (7.45) is unique, we have $\mathcal{P}_\mathcal{A}(t)\mathcal{W}(t) = \mathcal{W}(t)\mathcal{P}_\mathcal{A}(0)$ and therefore

$$\mathcal{V}_{\ker}(t) = \operatorname{im}\mathcal{P}_\mathcal{A}(t) = \operatorname{im}\mathcal{P}_\mathcal{A}(t)\mathcal{W}(t) = \operatorname{im}\mathcal{W}(t)\mathcal{P}_\mathcal{A}(0) = \mathcal{W}(t)\mathcal{V}_{\ker}(0).$$

Analogously, one proves $\mathcal{V}_\mathrm{c}(t) = \operatorname{im}\mathrm{id}_\mathcal{V} - \mathcal{P}_\mathcal{A}(t) = \mathcal{W}(t)\mathcal{V}_\mathrm{c}(0)$. □

Remark 7.32. The function $\mathcal{W}$ is the fundamental solution of the operator ODE (7.45).

For the analysis of the operator DAE (7.40) we use primarily the pointwise orthogonal decomposition

$$\mathcal{V} = \mathcal{V}_{\ker}(t) \oplus \mathcal{V}_{\ker}^\perp(t).$$

The associated pointwise right inverse $\mathcal{B}_\perp^-$ of $\mathcal{B}$ is then defined by the saddle point problem (7.43) where $\mathcal{A}$ is the time-independent Riesz isomorphism $\mathcal{R}_\mathcal{V} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$. Since $\mathcal{R}_\mathcal{V}$ is constant, $\mathcal{B} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)]$ is sufficient for the existence of the right inverse $\mathcal{B}_\perp^- \in W^{1,p}[0, T; \mathscr{L}(\mathcal{Q}^*, \mathcal{V})]$ with derivative

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{B}_\perp^- = -\mathcal{B}_\perp^-\dot{\mathcal{B}}\mathcal{B}_\perp^- + (\mathrm{id}_\mathcal{V} - \mathcal{B}_\perp^-\mathcal{B})\mathcal{R}_\mathcal{V}^{-1}(\mathcal{B}_\perp^-\dot{\mathcal{B}})^*\mathcal{R}_\mathcal{V}\mathcal{B}_\perp^-, \tag{7.48}$$

see Lemma 7.28 and Remark 7.29. For the derivation of (7.48) we have used that $\mathcal{R}_\mathcal{V}$ is self-adjoint by the symmetry of the inner product $(\cdot,\cdot)_\mathcal{V}$. Formula (7.48) then is the infinite dimensional variant of [GolP73, Eq. (4.12)]. Anyway, the pointwise orthogonal projection (7.44) is denoted by $\mathcal{P}_\perp \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V})]$. Since $\mathcal{P}_\perp(t)$ is an orthogonal projection, $\mathcal{W}(t)$ from Theorem 7.31 is unitary.

Lemma 7.33. *Let $\mathcal{B} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)]$, $p \in [1, \infty]$, satisfy uniformly an inf-sup condition* (3.2). *Let $\mathcal{W}$ be the fundamental solution of the operator ODE* (7.45) *with the orthogonal projection $\mathcal{P}_\perp$. Then $\mathcal{W} \in W^{1,p}[0, T; \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)]$ satisfies $\mathcal{W}(t)\mathcal{V}_{\ker}(0) = \mathcal{V}_{\ker}(t)$ and $\mathcal{W}(t)\mathcal{V}_{\ker}^\perp(0) = \mathcal{V}_{\ker}^\perp(t)$. Furthermore, $\mathcal{W}$ is pointwise unitary and its pointwise inverse is given by its pointwise Hilbert space-adjoint $\mathcal{W}^\mathrm{H} = \mathcal{R}_\mathcal{V}^{-1}\mathcal{W}^*\mathcal{R}_\mathcal{V}$.*

Proof. The regularity and the tracking property for the spaces $\mathcal{V}_{\ker}(t)$ and $\mathcal{V}_{\ker}^\perp(t)$ follow by the proof of Theorem 7.31. For the inverse of $\mathcal{W}$, we note that the equality $\mathcal{P}_\perp = \mathcal{P}_\perp^\mathrm{H} = \mathcal{R}_\mathcal{V}^{-1}\mathcal{P}_\perp^*\mathcal{R}_\mathcal{V}$ holds by [Alt16, p. 302], since $\mathcal{P}_\perp$ is pointwise an orthogonal projection. This then implies

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}^\mathrm{H}\mathcal{W}) = \mathcal{R}_\mathcal{V}^{-1}\mathcal{W}^*(\mathcal{P}_\perp^*\dot{\mathcal{P}}_\perp^* - \dot{\mathcal{P}}_\perp^*\mathcal{P}_\perp^*)\mathcal{R}_\mathcal{V}\mathcal{W} + \mathcal{W}^\mathrm{H}\dot{\mathcal{W}} = \mathcal{W}^\mathrm{H}(\mathcal{P}_\perp\dot{\mathcal{P}}_\perp - \dot{\mathcal{P}}_\perp\mathcal{P}_\perp)\mathcal{W} + \mathcal{W}^\mathrm{H}\dot{\mathcal{W}} \overset{(7.45)}{=} 0.$$

Therefore, $\mathcal{W}^{\mathrm{H}}(t)\mathcal{W}(t) = \mathcal{W}^{\mathrm{H}}(0)\mathcal{W}(0) = \mathcal{R}_{\mathcal{V}}^{-1}\mathcal{R}_{\mathcal{V}} = \mathrm{id}_{\mathcal{V}}$ and for every $v \in \mathcal{V}$ we obtain

$$\|\mathcal{W}(t)v\|_{\mathcal{V}}^2 = (\mathcal{W}(t)v, \mathcal{W}(t)v)_{\mathcal{V}} = (\mathcal{W}^{\mathrm{H}}(t)\mathcal{W}(t)v, v)_{\mathcal{V}} = (v, v)_{\mathcal{V}} = \|v\|_{\mathcal{V}}^2. \qquad \square$$

7.2.2. Existence Results

In this subsection we analyze the existence of solutions of the operator DAE (7.40) under Assumptions 7.2.i)–iv). For this we construct a solution by using the time-dependent splitting

$$\mathcal{V} = \mathcal{V}_{\mathrm{ker}}(t) \oplus \mathcal{V}_{\mathrm{ker}}^{\perp}(t)$$

and similar steps as in Subsection 7.1.1. The pointwise closure of $\mathcal{V}_{\mathrm{ker}}$ in $\mathcal{H}$ is denoted by $\mathcal{H}_{\mathrm{ker}}$, i.e.,

$$\mathcal{H}_{\mathrm{ker}}(t) := \mathrm{clos}_{\|\cdot\|_{\mathcal{H}}} \mathcal{V}_{\mathrm{ker}}(t).$$

Let $\mathcal{W}$ be the operator-valued function defined in Lemma 7.33. By Assumption 7.2.i) the map $\mathcal{W}$ is pointwise unitary and an element of $H^1[0,T;\mathscr{L}(\mathcal{V})]$. We set

$$\widetilde{\mathcal{B}} := \mathcal{B}\mathcal{W}, \qquad \widetilde{\mathcal{A}} := \mathcal{W}^*\mathcal{A}\mathcal{W}, \qquad \widetilde{f} := \mathcal{W}^* f. \tag{7.49}$$

For a possible solution u of (7.40) we define $\widetilde{u} = \mathcal{W}^{-1}u = \mathcal{W}^{\mathrm{H}}u$ such that the tuple $(\widetilde{u}, \lambda)$ solves

$$\mathcal{W}^*(t)\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}(t)\widetilde{u}(t)) + \widetilde{\mathcal{A}}(t)\widetilde{u}(t) - \widetilde{\mathcal{B}}^*(t)\lambda(t) = \widetilde{f}(t) \qquad \text{in } \mathcal{V}^*, \tag{7.50a}$$

$$\widetilde{\mathcal{B}}(t)\widetilde{u}(t) = g(t) \qquad \text{in } \mathcal{Q}^*, \tag{7.50b}$$

if and only if (u, λ) solves the operator DAE (7.40). For the initial value we choose $\widetilde{u}_0 = u_0 \in \mathcal{H}$. This is well-defined, since $\mathcal{W}(0) = \mathrm{id}_{\mathcal{V}} \in \mathscr{L}(\mathcal{V})$ can be extended to $\mathrm{id}_{\mathcal{H}}$.

By Theorem 7.31 and Lemma 4.15, the operator-valued function $\widetilde{\mathcal{B}} = \mathcal{B}\mathcal{W}$ is an element of $H^1[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$. Its pointwise kernel is time-invariant, since $\ker\widetilde{\mathcal{B}}(t) = \mathcal{W}^{-1}(t)\mathcal{V}_{\mathrm{ker}}(t) = \mathcal{V}_{\mathrm{ker}}(0)$, and $\widetilde{\mathcal{B}}$ is uniformly inf-sup stable by

$$\inf_{q\in\mathcal{Q}\setminus\{0\}} \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle\widetilde{\mathcal{B}}(t)v, q\rangle}{\|v\|_{\mathcal{V}}\|q\|_{\mathcal{Q}}} = \inf_{q\in\mathcal{Q}\setminus\{0\}} \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle\mathcal{B}(t)\mathcal{W}(t)v, q\rangle}{\|\mathcal{W}(t)v\|_{\mathcal{V}}\|q\|_{\mathcal{Q}}} = \inf_{q\in\mathcal{Q}\setminus\{0\}} \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle\mathcal{B}(t)v, q\rangle}{\|v\|_{\mathcal{V}}\|q\|_{\mathcal{Q}}} \geq \beta > 0,$$

where we have used that $\mathcal{W}$ is pointwise unitary. Lemma 7.28 then implies that $\widetilde{\mathcal{B}}$ has a right-inverse $\widetilde{\mathcal{B}}_{\perp}^{-} \in H^1[0,T;\mathscr{L}(\mathcal{Q}^*,\mathcal{V})]$ with $\mathrm{im}\,\widetilde{\mathcal{B}}_{\perp}^{-}(t) = \mathcal{V}_{\mathrm{ker}}^{\perp}(0)$. In particular, one shows with Theorem 7.31 that

$$\mathcal{W}\widetilde{\mathcal{B}}_{\perp}^{-} = \mathcal{B}_{\perp}^{-}. \tag{7.51}$$

Remark 7.34. Because of the time-invariance of the kernel of $\widetilde{\mathcal{B}}$, the DAE formulation with the transformation $\mathcal{W}$ is more favorable than the one without. For finite dimensional DAEs of the form (2.3) with a matrix-valued function B stability properties may not be preserved under temporal discretization – even after an index reduction – if the kernel of B changes too fast for the temporal step size [KunM07]. A possibility to overcome this issue is a smooth transformation, which eliminates the time-changing kernel [KunM07, p. 409]. This is in our case the purpose of $\mathcal{W}$.

Following the ideas of Subsection 7.1.1 we split $\widetilde{u}$ into $\widetilde{u}_{\mathrm{ker}} + \widetilde{u}_{\mathrm{c}}$ with $\widetilde{u}_{\mathrm{ker}}\colon [0,T] \to \mathcal{V}_{\mathrm{ker}}(0)$ and $\widetilde{u}_{\mathrm{c}}\colon [0,T] \to \mathcal{V}_{\mathrm{ker}}^{\perp}(0)$. The function $\widetilde{u}_{\mathrm{c}}$ is defined via

$$\widetilde{u}_{\mathrm{c}} := \widetilde{\mathcal{B}}_{\perp}^{-} g \in H^1(0,T;\mathcal{V}_{\mathrm{c}}(0)) \subset H^1(0,T;\mathcal{V}), \tag{7.52}$$

where its generalized derivative is given by $\dot{\widetilde{u}}_{\mathrm{c}} = \frac{\mathrm{d}}{\mathrm{d}t}(\widetilde{\mathcal{B}}_{\perp}^{-})g + \widetilde{\mathcal{B}}_{\perp}^{-}\dot{g}$; see Lemma 4.15. For the remaining part of the solution of (7.50), i.e., $\widetilde{u}_{\mathrm{ker}} = \widetilde{u} - \widetilde{u}_{\mathrm{c}}$ and λ, we consider the operator DAE

$$\begin{aligned} \mathcal{W}^*(t)\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}(t)\widetilde{u}_{\mathrm{ker}}(t)) + \widetilde{\mathcal{A}}(t)\widetilde{u}_{\mathrm{ker}}(t) - \widetilde{\mathcal{B}}^*(t)\lambda(t) &= \widetilde{f}_{\mathrm{ker}}(t) \qquad \text{in } \mathcal{V}^*, && (7.53a)\\ \widetilde{\mathcal{B}}(t)\widetilde{u}_{\mathrm{ker}}(t) &= 0 \qquad \text{in } \mathcal{Q}^*, && (7.53b)\end{aligned}$$

with the right-hand side

$$\widetilde{f}_{\mathrm{ker}} := \widetilde{f} - \widetilde{\mathcal{A}}\widetilde{u}_{\mathrm{c}} - \mathcal{W}^*\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}\widetilde{u}_{\mathrm{c}}) \in L^2(0,T;\mathcal{V}^*). \tag{7.54}$$

The associated initial value is given as $\widetilde{u}_{\mathrm{ker},0} = \widetilde{u}_0 - \widetilde{\mathcal{B}}_\perp^-(0)g(0) = u_0 - \mathcal{B}_\perp^-(0)g(0) \in \mathcal{H}_{\mathrm{ker}}(0)$ by Assumption 7.2.iv).

Remark 7.35. Instead of (7.50) one can also investigate directly the operator DAE (7.40). The possible solution u then can be split into $u_c = \mathcal{B}_\perp^- g$ and u_{ker} as a solution of the operator DAE (7.6), where $\mathcal{B}$ is also time-dependent. In this case, a restriction of the differential equation (7.6a) to test functions with images in $\mathcal{V}_{\mathrm{ker}}$ leads to an operator ODE with an evolving test space, since $\mathcal{V}_{\mathrm{ker}}$ is time-dependent. Such a type of differential equation also occurs for parabolic PDEs with time-dependent domains; see e.g. [AlpES15; BonG01].

For the construction of a solution $(\widetilde{u}_{\mathrm{ker}}, \lambda)$ we discretize (7.53) with the implicit Euler scheme. The pointwise evaluation of the operator $\widetilde{\mathcal{B}}$ is denoted by $\widetilde{\mathcal{B}}_n := \widetilde{\mathcal{B}}(t_n)$ and analogously for $\mathcal{W}_n$, $n = 0, \dots, N$. The values $\widetilde{f}_{\mathrm{ker},n}$ and $\widetilde{\mathcal{A}}_n$ are defined by the integral means of $\widetilde{f}_{\mathrm{ker}}$ and $\widetilde{\mathcal{A}}$ over the interval $[t_{n-1}, t_n]$, respectively; cf. (7.9). The time-discrete system then reads

$$\begin{aligned} \mathcal{W}_n^* D_\tau(\mathcal{W}\widetilde{u}_{\mathrm{ker}})_n + \widetilde{\mathcal{A}}_n\widetilde{u}_{\mathrm{ker},n} - \widetilde{\mathcal{B}}_n^*\lambda_n &= \widetilde{f}_{\mathrm{ker},n} \qquad \text{in } \mathcal{V}^*, && (7.55a)\\ \widetilde{\mathcal{B}}_n\widetilde{u}_{\mathrm{ker},n} &= 0 \qquad \text{in } \mathcal{Q}^*, && (7.55b)\end{aligned}$$

with the discrete derivative $D_\tau(\mathcal{W}\widetilde{u}_{\mathrm{ker}})_n = (\mathcal{W}_n\widetilde{u}_{\mathrm{ker},n} - \mathcal{W}_{n-1}\widetilde{u}_{\mathrm{ker},n-1})/\tau$ and $n = 1, \dots, N$. Following Subsection 7.1.1, we show that the system (7.55) is uniquely solvable and bound its solutions.

Lemma 7.36. *Let Assumptions 7.2.i)–iv) be satisfied. Suppose that $\mathcal{W}$ is defined as in Lemma 7.33, $\widetilde{\mathcal{B}}$, $\widetilde{\mathcal{A}}$ as in (7.49), and $\widetilde{f}_{\mathrm{ker}}$ as in (7.54). Assume that $\mathcal{W}_n$ and $\widetilde{\mathcal{B}}_n$ are defined by pointwise evaluation as well as $\widetilde{f}_{\mathrm{ker},n}$ and $\widetilde{\mathcal{A}}_n$ by integral means. Let $\widetilde{u}_{\mathrm{ker},0} := u_0 - \mathcal{B}_\perp^-(0)g(0) \in \mathcal{H}_{\mathrm{ker}}(0)$.*

Then there exists a unique sequence of solutions $\{(\widetilde{u}_{\mathrm{ker},n}, \lambda_n)\}_{n=1,\dots,N} \subset \mathcal{V}_{\mathrm{ker}}(0) \times \mathcal{Q}$ of (7.55), which satisfies the bounds

$$\|\mathcal{W}_n\widetilde{u}_{\mathrm{ker},n}\|_{\mathcal{H}}^2 + \sum_{k=1}^n \|\mathcal{W}_k\widetilde{u}_{\mathrm{ker},k} - \mathcal{W}_{k-1}\widetilde{u}_{\mathrm{ker},k-1}\|_{\mathcal{H}}^2 + \mu_{\mathcal{A}}\sum_{k=1}^n \tau\|\widetilde{u}_{\mathrm{ker},k}\|_{\mathcal{V}}^2 \le M^2(\widetilde{u}_{\mathrm{ker},0}, \widetilde{f}_{\mathrm{ker}}), \tag{7.56a}$$

$$\sum_{k=1}^n \tau\|\mathcal{W}_k^* D_\tau(\mathcal{W}\widetilde{u}_{\mathrm{ker}})_k\|_{\mathcal{V}_{\mathrm{ker}}^*(0)}^2 \le 2\big(1 + \tfrac{1}{\mu_{\mathcal{A}}}\|\mathcal{A}\|_{L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]}^2\big)M^2(\widetilde{u}_{\mathrm{ker},0}, \widetilde{f}_{\mathrm{ker}}) \tag{7.56b}$$

with the constant $M(\widetilde{u}_{\mathrm{ker},0}, \widetilde{f}_{\mathrm{ker}}) := \sqrt{\|\widetilde{u}_{\mathrm{ker},0}\|_{\mathcal{H}}^2 + \tfrac{1}{\mu_{\mathcal{A}}}\int_0^T \|\widetilde{f}_{\mathrm{ker}}\|_{\mathcal{V}^*}^2 \,\mathrm{d}s}$.

Proof. Because $\mathcal{W}_0 = \mathcal{W}(0) = \mathrm{id}_{\mathcal{V}}$ can be extended to $\mathrm{id}_{\mathcal{H}}$, the expression $\mathcal{W}_0\widetilde{u}_{\mathrm{ker},0} = \widetilde{u}_{\mathrm{ker},0}$ is well-defined for $\widetilde{u}_{\mathrm{ker},0} \in \mathcal{H}_{\mathrm{ker}}(0)$. Furthermore, Assumption 7.2.ii), $\mathcal{W}(t)\mathcal{V}_{\mathrm{ker}}(0) = \mathcal{V}_{\mathrm{ker}}(t)$, and the pointwise unitarity of $\mathcal{W}$ imply

$$\begin{aligned} \tfrac{1}{\tau}(\mathcal{W}_n v_{\mathrm{ker}}, \mathcal{W}_n v_{\mathrm{ker}}) + \langle \widetilde{\mathcal{A}}_n v_{\mathrm{ker}}, v_{\mathrm{ker}} \rangle &= \frac{1}{\tau}\|\mathcal{W}_n v_{\mathrm{ker}}\|_{\mathcal{H}}^2 + \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \langle \mathcal{A}(s)\mathcal{W}(s)v_{\mathrm{ker}}, \mathcal{W}(s)v_{\mathrm{ker}} \rangle \,\mathrm{d}s \\ &\ge \frac{1}{\tau}\|\mathcal{W}_n v_{\mathrm{ker}}\|_{\mathcal{H}}^2 + \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \mu_{\mathcal{A}}\|\mathcal{W}(s)v_{\mathrm{ker}}\|_{\mathcal{V}}^2 \,\mathrm{d}s \end{aligned}$$

$$= \mu_{\mathcal{A}} \|v_{\ker}\|_{\mathcal{V}}^2 + \frac{1}{\tau} \|\mathcal{W}_n v_{\ker}\|_{\mathcal{H}}^2 \geq \mu_{\mathcal{A}} \|v_{\ker}\|_{\mathcal{V}}^2$$

for every $v_{\ker} \in \mathcal{V}_{\ker}(0)$ and $n = 1, \ldots, N$. The existence and the uniqueness of the sequence of solution then follows iteratively by Theorem 3.8. In particular, $\widetilde{u}_{\ker,n} \in \ker \widetilde{\mathcal{B}}_n = \mathcal{V}_{\ker}(0)$ holds by (7.55b), $n = 1, \ldots, N$. Following the lines of Theorem 7.9, testing (7.55a) with $\tau \widetilde{u}_{\ker,n} \in \mathcal{V}_{\ker}(0)$ implies (7.56a) and $\mathcal{W}_n^* D_\tau(\mathcal{W}\widetilde{u}_{\ker})_n = \widetilde{f}_{\ker,n} - \widetilde{\mathcal{A}}_n \widetilde{u}_{\ker,n}$ in $\mathcal{V}_{\ker}(0)^*$ leads to the estimate (7.56b). □

Remark 7.37. Since $\mathcal{W}(t)$ is unitary and $\mathcal{W}(0) = \mathrm{id}_{\mathcal{V}}$, we have $M(\widetilde{u}_{\ker,0}, \widetilde{f}_{\ker}) = M(u_{\ker,0}, f_{\ker})$.

With the solution sequence of Lemma 7.36 we define time-continuous functions. The piecewise constant function $\widetilde{u}_{\ker,\tau}$ is defined as $u_{\ker,\tau}$ in (7.18) and the piecewise linear function $\widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau}$ as $\widehat{\mathcal{M}u}_{\ker,\tau}$ in (7.19b), i.e.,

$$\widetilde{u}_{\ker,\tau}(t) := \begin{cases} \widetilde{u}_{\ker,1}, & \text{if } t = 0, \\ \widetilde{u}_{\ker,n}, & \text{if } t \in (t_{n-1}, t_n], \end{cases}$$

$$\widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau}(t) := \begin{cases} \mathcal{W}_0 \widetilde{u}_{\ker,0}, & \text{if } t = 0, \\ \mathcal{W}_n \widetilde{u}_{\ker,n} + D_\tau(\mathcal{W}\widetilde{u}_{\ker})_n (t - t_n), & \text{if } t \in (t_{n-1}, t_n]. \end{cases}$$

We write again $\widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau}$ instead of $(\widehat{\mathcal{W}\widetilde{u}_{\ker}})_\tau$ in interest of readability. Note that $\mathcal{W}_0 \widetilde{u}_{\ker,0} := \widetilde{u}_{\ker,0} \in \mathcal{H}_{\ker}(0)$. Analogously to the piecewise constant functions in Subsection 7.1.1.2, we define λ_τ, $\widetilde{f}_{\ker,\tau}$, $\widetilde{\mathcal{A}}_\tau$, $\widetilde{\mathcal{B}}_\tau$, and $\mathcal{W}_\tau$. With these definitions, the system (7.55) can be rewritten as

$$\mathcal{W}_\tau^* \tfrac{\mathrm{d}}{\mathrm{d}t} \widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau} + \widetilde{\mathcal{A}}_\tau \widetilde{u}_{\ker,\tau} - \widetilde{\mathcal{B}}_\tau^* \lambda_\tau = \widetilde{f}_{\ker,\tau} \quad \text{in } \mathcal{V}^*, \tag{7.57a}$$

$$\widetilde{\mathcal{B}}_\tau \widetilde{u}_{\ker,\tau} = 0 \quad \text{in } \mathcal{Q}^*. \tag{7.57b}$$

Similarly to the bounds (7.13) and (7.14), the inequalities (7.56a) and (7.56b) imply weak and weak* convergence. We analyze them in the following lemmas.

Lemma 7.38. *Let the assumptions of Lemma 7.36 be satisfied. Then there exists a function $\widetilde{u}_{\ker} \in L^2(0,T;\mathcal{V}_{\ker}(0))$ with $\mathcal{W}\widetilde{u}_{\ker} \in L^\infty(0,T;\mathcal{H})$ and a subsequence of τ denoted by τ', such that*

$$\widetilde{u}_{\ker,\tau'} \rightharpoonup \widetilde{u}_{\ker} \quad \text{in } L^2(0,T;\mathcal{V}), \qquad \widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau'} \rightharpoonup \mathcal{W}\widetilde{u}_{\ker} \quad \text{in } L^2(0,T;\mathcal{H}),$$

$$\mathcal{W}_{\tau'}\widetilde{u}_{\ker,\tau'}, \widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau'} \overset{*}{\rightharpoonup} \mathcal{W}\widetilde{u}_{\ker} \quad \text{in } L^\infty(0,T;\mathcal{H}), \qquad \widetilde{\mathcal{A}}_{\tau'}\widetilde{u}_{\ker,\tau'} \rightharpoonup \widetilde{\mathcal{A}}\widetilde{u}_{\ker} \quad \text{in } L^2(0,T;\mathcal{V}^*)$$

as $\tau' \to 0$.

Proof. By the estimate (7.56a) the function $\widetilde{u}_{\ker,\tau}$ is bounded by $\mu_{\mathcal{A}}^{-1/2} M(\widetilde{u}_{\ker,0}, \widetilde{f}_{\ker})$ in $L^2(0,T;\mathcal{V})$ independently of τ. Therefore, there exists a subsequence τ' of τ such that $\widetilde{u}_{\ker,\tau'}$ converges weakly to a function $\widetilde{u}_{\ker}$ in $L^2(0,T;\mathcal{V})$ as $\tau' \to 0$. By the arguments of the proof of Lemma 7.11 we have $\widetilde{u}_{\ker} \in L^2(0,T;\mathcal{V}_{\ker}(0))$ and $\mathcal{W}_{\tau'}\widetilde{u}_{\ker,\tau'} \rightharpoonup \mathcal{W}\widetilde{u}_{\ker}$ in $L^2(0,T;\mathcal{V})$ as $\tau' \to 0$; cf. (7.11). On the other hand, the estimate (7.56a) shows also that the sequence $\mathcal{W}_\tau \widetilde{u}_{\ker,\tau}$ is bounded in $L^\infty(0,T;\mathcal{H})$ independently of τ. By [Emm04, Cor. 8.1.11] there exists then a subsequence of τ' denoted by τ' as well, such that $\mathcal{W}_{\tau'}\widetilde{u}_{\ker,\tau'} \overset{*}{\rightharpoonup} \mathcal{W}\widetilde{u}_{\ker}$ in $L^\infty(0,T;\mathcal{H})$ as $\tau' \to 0$.

For the piecewise linear function $\widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau}$ we note

$$\left\|\mathcal{W}_\tau \widetilde{u}_{\ker,\tau} - \widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau}\right\|_{L^2(0,T;\mathcal{H})}^2 = \sum_{n=1}^N \|D_\tau(\mathcal{W}\widetilde{u}_{\ker})_n\|_{\mathcal{H}}^2 \int_{t_{n-1}}^{t_n} (t - t_n)^2 \,\mathrm{d}t \overset{(7.56a)}{\leq} \frac{\tau}{3} M^2(\widetilde{u}_{\ker,0}, \widetilde{f}_{\ker}) \to 0$$

as $\tau \to 0$. This shows that $\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau'}$ converges weakly in $L^2(0,T;\mathcal{H})$ as $\tau' \to 0$. Since $L^2(0,T;\mathcal{H}^*)$ is dense in $L^1(0,T;\mathcal{H}^*)$ and $\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}$ is bounded independently of τ in $L^\infty(0,T;\mathcal{H})$ by

$$\begin{aligned}\|\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}\|^2_{L^\infty(0,T;\mathcal{H})} &= \max_{n=1,\dots,N}\max_{t\in[t_{n-1},t_n]} \|\mathcal{W}_n\widetilde{u}_{\mathrm{ker},n} + D_\tau(\mathcal{W}\widetilde{u}_{\mathrm{ker}})_n(t-t_n)\|^2_{\mathcal{H}} \\ &\le \max_{n=1,\dots,N} 2\|\mathcal{W}_n\widetilde{u}_{\mathrm{ker},n}\|^2_{\mathcal{H}} + 2\tau^2\|D_\tau(\mathcal{W}\widetilde{u}_{\mathrm{ker}})_n\|^2_{\mathcal{H}} \overset{(7.56a)}{\le} 2M(\widetilde{u}_{\mathrm{ker},0},\widetilde{f}_{\mathrm{ker}}),\end{aligned}$$

its weak* convergence in $L^\infty(0,T;\mathcal{H})$ follows by [Yos80, Sec. V.1, Th. 3]. Finally, the weak convergence of $\widetilde{\mathcal{A}}_{\tau'}\widetilde{u}_{\mathrm{ker},\tau'}$ can be proven with the same steps as in the proof of Lemma 7.13. □

Lemma 7.39. *Let the assumption of Lemma 7.36 be satisfied. Suppose that the function $\widetilde{u}_{\mathrm{ker}}$ and the subsequence τ' of τ are the same as in Lemma 7.38. Then the function $\mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}$ has a derivative in $L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}(0))$ with*

$$\mathcal{W}^*_\tau \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau'} \rightharpoonup \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}) - \dot{\mathcal{W}}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}} \quad \text{in } L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}(0))$$

as $\tau' \to 0$. Furthermore, $(\mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}})(0) = \widetilde{u}_{\mathrm{ker},0}$ holds in $\mathcal{V}^*_{\mathrm{ker}}(0)$.

Proof. By the estimate (7.56b) the function $\mathcal{W}^*_\tau \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}$ is bounded in $L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}(0))$ independently of τ. Therefore, there exists a subsequence τ'' of τ' such that $\mathcal{W}^*_{\tau''}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}$ converges weakly to a $\widetilde{w}$ in $L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}(0))$ as $\tau'' \to 0$. For the generalized derivative of $\mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}$, we prove

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}) = \widetilde{w} + \dot{\mathcal{W}}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}} \in L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}(0)). \tag{7.58}$$

Note that the right hand side of (7.58) is well-defined by $\mathcal{W}\widetilde{u}_{\mathrm{ker}} \in L^\infty(0,T;\mathcal{H}) \hookrightarrow L^\infty(0,T;\mathcal{V}^*)$ and Lemma 4.9. Let now $v_{\mathrm{ker}} \in \mathcal{V}_{\mathrm{ker}}(0)$ and $\varphi \in C^\infty_c(0,T)$ be arbitrary. On the one hand it then holds

$$\begin{aligned}&\int_0^T \langle (\mathcal{W}^*_{\tau''} - \mathcal{W}^*)\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, v_{\mathrm{ker}}\rangle\varphi\,\mathrm{d}s \\ &= \int_0^T \langle \mathcal{W}^*_{\tau''}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, v_{\mathrm{ker}}\rangle\varphi + \langle \dot{\mathcal{W}}^*\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, v_{\mathrm{ker}}\rangle\varphi + \langle \mathcal{W}^*\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, v_{\mathrm{ker}}\rangle\dot{\varphi}\,\mathrm{d}s \\ &= \int_0^T \langle \mathcal{W}^*_{\tau''}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, v_{\mathrm{ker}}\rangle\varphi + (\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, \dot{\mathcal{W}}v_{\mathrm{ker}})_{\mathcal{H}}\varphi + (\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau''}, \mathcal{W}v_{\mathrm{ker}})_{\mathcal{H}}\dot{\varphi}\,\mathrm{d}s \\ &\to \int_0^T \langle \widetilde{w} + \dot{\mathcal{W}}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}, v_{\mathrm{ker}}\rangle\varphi\,\mathrm{d}s + \int_0^T \langle \mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}, v_{\mathrm{ker}}\rangle\dot{\varphi}\,\mathrm{d}s\end{aligned} \tag{7.59}$$

as $\tau'' \to 0$ by the weak convergences of Lemma 7.38. On the other hand, we have

$$\begin{aligned}&\|\langle (\mathcal{W}^*_\tau - \mathcal{W}^*)\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}, v_{\mathrm{ker}}\rangle\varphi\|^2_{L^1(0,T)} \\ &\le C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\|\varphi\|^2_{C(0,T)}\int_0^T \|\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}\|^2_{\mathcal{H}}\,\mathrm{d}s \int_0^T \|(\mathcal{W}_\tau - \mathcal{W})v_{\mathrm{ker}}\|^2_{\mathcal{V}}\,\mathrm{d}s \\ &= C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\|\varphi\|^2_{C(0,T)}\sum_{n=1}^N \tau\|D_\tau(\mathcal{W}\widetilde{u}_{\mathrm{ker}})_n\|^2_{\mathcal{H}}\sum_{n=1}^N\int_{t_{n-1}}^{t_n}\Big\|\int_s^{t_n}\dot{\mathcal{W}}v_{\mathrm{ker}}\,\mathrm{d}\eta\Big\|^2_{\mathcal{V}}\,\mathrm{d}s \\ &\overset{(7.56a)}{\le} \tau^{-1}M(\widetilde{u}_{\mathrm{ker},0},\widetilde{f}_{\mathrm{ker}})C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\|\varphi\|^2_{C(0,T)}\|v_{\mathrm{ker}}\|^2_{\mathcal{V}}\sum_{n=1}^N\int_{t_{n-1}}^{t_n}(t_n - s)\int_s^{t_n}\|\dot{\mathcal{W}}\|^2_{\mathscr{L}(\mathcal{V})}\,\mathrm{d}\eta\,\mathrm{d}s \\ &\le \frac{\tau}{2}M(\widetilde{u}_{\mathrm{ker},0},\widetilde{f}_{\mathrm{ker}})C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}\|\varphi\|^2_{C(0,T)}\|v_{\mathrm{ker}}\|^2_{\mathcal{V}}\int_0^T\|\dot{\mathcal{W}}\|^2_{\mathscr{L}(\mathcal{V})}\,\mathrm{d}s \to 0\end{aligned} \tag{7.60}$$

as $\tau \to 0$. Then (7.59) and (7.60) prove

$$0 = \int_0^T \langle\, \widetilde{w} + \dot{\mathcal{W}}^* \mathcal{W}\widetilde{u}_{\text{ker}}, v_{\text{ker}} \,\rangle \varphi \,\mathrm{d}s + \int_0^T \langle\, \mathcal{W}^*\mathcal{W}\widetilde{u}_{\text{ker}}, v_{\text{ker}} \,\rangle \dot{\varphi} \,\mathrm{d}s,$$

which is equivalent to (7.58). The uniqueness of the derivative proves the weak convergence for the whole sequence τ' [GajGZ74, Ch. 1, Lem. 5.4]. Finally, the assertion on the initial value is proven by the arguments of Lemma 7.12 and the fact that for every $v_{\text{ker}} \in \mathcal{V}_{\text{ker}}(0)$ we have

$$\begin{aligned}
&\langle\, (\mathcal{W}^*\mathcal{W}\widetilde{u}_{\text{ker}})(0), T v_{\text{ker}} \,\rangle \\
&\overset{(7.58)}{=} \int_0^T \langle\, \mathcal{W}^*\mathcal{W}\widetilde{u}_{\text{ker}}, v_{\text{ker}} \,\rangle - \langle\, \widetilde{w} - \dot{\mathcal{W}}^*\mathcal{W}\widetilde{u}_{\text{ker}}, (T-t)v_{\text{ker}} \,\rangle \,\mathrm{d}t \\
&= \lim_{\tau' \to 0} \int_0^T \langle\, \mathcal{W}^*\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau'}, v_{\text{ker}} \,\rangle - \langle\, \mathcal{W}_{\tau'}^* \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau'} - \dot{\mathcal{W}}^*\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau'}, (T-t)v_{\text{ker}} \,\rangle \,\mathrm{d}t \\
&\overset{(7.60)}{=} \lim_{\tau' \to 0} \int_0^T -\tfrac{\mathrm{d}}{\mathrm{d}t}\langle\, \widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau'}, (T-t)\mathcal{W}v_{\text{ker}} \,\rangle \,\mathrm{d}t = \langle\, \widetilde{u}_{\text{ker},0}, T v_{\text{ker}} \,\rangle. \qquad \square
\end{aligned}$$

Remark 7.40. By (7.58) it formally holds $\widetilde{w} = \mathcal{W}^* \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}\widetilde{u}_{\text{ker}})$. Since $(\mathcal{W}u_{\text{ker}})(t) \in \mathcal{V}_{\text{ker}}(t)$, its derivative would satisfy $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}\widetilde{u}_{\text{ker}})(t) \in \mathcal{V}_{\text{ker}}^*(t)$ at almost every time-point $t \in [0,T]$. Note that there is in general no Banach space, which contains $\mathcal{V}_{\text{ker}}^*(t)$ for all $t \in [0,T]$. However, one can prove at least that $w := \mathcal{P}_\perp^*(\mathcal{W}^*)^{-1}\widetilde{w} \in L^2(0,T;\mathcal{V}^*)$ satisfies $\|w(t)\|_{\mathcal{V}^*} = \|w(t)\|_{\mathcal{V}_{\text{ker}}^*(t)}$ as well as $\int_0^T \langle\, w, v_{\text{ker}} \,\rangle \,\mathrm{d}s = \int_0^T \langle\, \mathcal{W}\widetilde{u}_{\text{ker}}, \dot{v}_{\text{ker}} \,\rangle \,\mathrm{d}s$ for every $v_{\text{ker}} \in H^1(0,T;\mathcal{V})$ with $v_{\text{ker}}(0) = v_{\text{ker}}(T) = 0$ and $v_{\text{ker}}(t) \in \mathcal{V}_{\text{ker}}(t)$.

With the two previous lemmas we can now prove the existence of a solution of the operator DAE (7.50) and therefor of (7.40) as well.

Theorem 7.41 (Existence of Solutions)**.** *Suppose that $\mathcal{M} = \mathcal{R}_\mathcal{V}$ and that Assumptions 7.2.i)–iii) are satisfied. Let Assumption 7.2.iv) be fulfilled, i.e., $u_0 \in \mathcal{H}_{\text{ker}}(0) + \mathcal{B}^-(0)g(0)$. Then the operator DAE (7.40) has at least one solution (u, λ) with*

- *a)* $u \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$,
- *b)* $\mathcal{B}^*\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda_{\mathcal{B}^*}$ *for an* $\Lambda_{\mathcal{B}^*} \in L^\infty(0,T;\mathcal{V}^*)$,
- *c)* $\mathcal{P}_\perp^*\dot{u} \in L^2(0,T;\mathcal{V}^*)$,
- *d)* $\dot{u} - \mathcal{B}^*\lambda \in L^2(0,T;\mathcal{V}^*)$.

Proof. We define $\widetilde{u}_{\text{c}}$ as in (7.52) and $\widetilde{u}_{\text{ker}}$ as in Lemmas 7.38 and 7.39 with the initial value $\widetilde{u}_{\text{ker},0} = u_0 - \widetilde{u}_{\text{c}}(0) = u_0 - \mathcal{B}_\perp^-(0)g(0) \in \mathcal{H}_{\text{ker}}(0)$. Then $u := \mathcal{W}\widetilde{u}_{\text{c}} + \mathcal{W}\widetilde{u}_{\text{ker}}$ satisfies a) and c) by Lemma 7.28, 7.38, Theorem 7.31, and Remark 7.40.

For the part $\Lambda_{\mathcal{B}^*}$ of the solution we define pointwisely $(\mathcal{W}_\tau^*)^{-1}(t) := (\mathcal{W}_\tau^*(t))^{-1}$ and note

$$\big\|(\mathcal{W}_\tau^*)^{-1}(t)\big\|_{\mathscr{L}(\mathcal{V}^*)} = \big\|\big(\mathcal{W}_\tau^{\text{H}}\big)^*(t)\big\|_{\mathscr{L}(\mathcal{V}^*)} = \big\|\mathcal{W}_\tau^{\text{H}}(t)\big\|_{\mathscr{L}(\mathcal{V})} = \big\|\mathcal{W}_\tau(t)\big\|_{\mathscr{L}(\mathcal{V})} = 1$$

by Lemma 7.33 at almost every time-point $t \in [0,T]$. With the integration operator $\mathcal{I}$ from (7.22) the term

$$\begin{aligned}
\mathcal{I}((\mathcal{W}_\tau^*)^{-1}\widetilde{\mathcal{B}}_\tau^*\lambda_\tau) &\overset{(7.53a)}{=} \mathcal{I}(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau}) + \mathcal{I}((\mathcal{W}_\tau^*)^{-1}\widetilde{\mathcal{A}}_\tau\widetilde{u}_{\text{ker},\tau}) - \mathcal{I}((\mathcal{W}_\tau^*)^{-1}\widetilde{f}_{\text{ker},\tau}) \\
&= \widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau} - u_{\text{ker},0} + \mathcal{I}((\mathcal{W}_\tau^*)^{-1}\widetilde{\mathcal{A}}_\tau\widetilde{u}_{\text{ker},\tau}) - \mathcal{I}((\mathcal{W}_\tau^*)^{-1}\widetilde{f}_{\text{ker},\tau})
\end{aligned}$$

is bounded in $L^\infty(0,T;\mathcal{V}^*)$ independently of τ; see Lemma 7.36. Therefore, by Lemma 7.38, the continuity of $\mathcal{W}^{-1}$, see Theorem 7.31, and the strong convergence of $\widetilde{f}_{\text{ker},\tau}$ to $\widetilde{f}_{\text{ker}}$, see Lemma 3.34, we have

$$\begin{aligned}\mathcal{I}((\mathcal{W}^*_{\tau'})^{-1}\widetilde{\mathcal{B}}^*_{\tau'}\lambda_{\tau'}) &\overset{*}{\rightharpoonup} \mathcal{W}\widetilde{u}_{\text{ker}} - u_{\text{ker},0} + \mathcal{I}((\mathcal{W}^*)^{-1}\widetilde{\mathcal{A}}\widetilde{u}_{\text{ker}}) - \mathcal{I}((\mathcal{W}^*)^{-1}\widetilde{f}_{\text{ker}})\\ &\overset{(7.54)}{=} u - u_0 + \mathcal{I}((\mathcal{W}^*)^{-1}\widetilde{\mathcal{A}}\widetilde{u}) - \mathcal{I}((\mathcal{W}^*)^{-1}\widetilde{f})\\ &\overset{(7.49)}{=} u - u_0 + \mathcal{I}(\mathcal{A}u) - \mathcal{I}(f) =: \Lambda_{\mathcal{B}^*}\end{aligned}$$

in $L^\infty(0,T;\mathcal{V}^*)$ as $\tau'\to 0$. Here, τ' denotes the subsequence of τ from Lemma 7.38. With equivalent arguments one shows for arbitrary but fixed $v\in\mathcal{V}$ and $\varphi\in C^\infty([0,T])$ with $\varphi(T)=0$ the limit

$$\begin{aligned}0 &\overset{(7.57a)}{=} \int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau'} + (\mathcal{W}^*_{\tau'})^{-1}\widetilde{\mathcal{A}}_{\tau'}\widetilde{u}_{\text{ker},\tau'} - (\mathcal{W}^*_{\tau'})^{-1}\widetilde{\mathcal{B}}^*_{\tau'}\lambda_{\tau'} - (\mathcal{W}^*_{\tau'})^{-1}\widetilde{f}_{\text{ker},\tau'}, v\rangle\varphi\,\mathrm{d}s\\ &\to \int_0^T \langle \mathcal{A}u - f, v\rangle\varphi - \langle u - \Lambda_{\mathcal{B}^*}, v\rangle\dot\varphi\,\mathrm{d}s + \langle u_0, v\rangle\varphi(0)\end{aligned}$$

as $\tau'\to 0$. Furthermore, u fulfills (7.40b) since $\mathcal{B}u = \widetilde{\mathcal{B}}\widetilde{u}_{\text{c}} + \widetilde{\mathcal{B}}\widetilde{u}_{\text{ker}} = \widetilde{\mathcal{B}}\widetilde{\mathcal{B}}^-_\perp g = g$ by (7.49), (7.52), and Lemma 7.38. Therefore, (u,λ) is a solution in the sense of Definition 4.27. The remaining assertions b) and d) follow with the arguments of the proof of Theorem 7.14. □

For a time-independent operator $\mathcal{B}$, Theorem 7.14 shows that not only $\mathcal{B}^*\lambda$ but also λ has a regular primitive. In the associated proof we used that $\mathcal{B}^{-*}_{\text{left}}$ and integration commute; see Lemma 3.30. Since $\mathcal{B}^{-*}_{\text{left}}(t) := (\mathcal{B}(t))^{-*}_{\text{left}}$ is now time-dependent the proof of the existence of a primitive of λ is more delicate.

Theorem 7.42 (Existence of a Regular Primitive of λ)**.** *Suppose that the assumptions of Theorem 7.41 are satisfied. Then the Lagrange multiplier λ has a regular primitive $\Lambda\in L^\infty(0,T;\mathcal{Q})$.*

Proof. We introduce the piecewise functions

$$w_{\text{ker},\tau}(t) := \begin{cases}\mathcal{W}_n\widetilde{u}_{\text{ker},n}, & \text{if } t\in[t_n,t_{n+1})\\ \mathcal{W}_N\widetilde{u}_{\text{ker},N}, & \text{if } t=T\end{cases},\quad \widehat{\mathcal{W}}_\tau(t) := \begin{cases}\mathcal{W}_0, & \text{if } t=0\\ \mathcal{W}_n + \frac{\mathcal{W}_n-\mathcal{W}_{n-1}}{\tau}(t-t_n), & \text{if } t\in(t_{n-1},t_n]\end{cases};$$

cf. (7.19a). By the estimate (7.56a) the function $w_{\text{ker},\tau}$ is bounded in $L^\infty(0,T;\mathcal{H})$ independently of τ. Its weak* limit for the subsequence τ' of Lemma 7.38 is given by $\widetilde{u}_{\text{ker}}$, since

$$\int_0^T \|\mathcal{W}_\tau\widetilde{u}_{\text{ker},\tau} - w_{\text{ker},\tau}\|^2_{\mathcal{H}}\,\mathrm{d}s = \sum_{n=1}^N \tau\|\mathcal{W}_n\widetilde{u}_{\text{ker},n} - \mathcal{W}_{n-1}\widetilde{u}_{\text{ker},n-1}\|^2_{\mathcal{H}} \overset{(7.56a)}{\le} \tau M^2(\widetilde{u}_{\text{ker},0},\widetilde{f}_{\text{ker}}) \to 0$$

as $\tau\to 0$. Furthermore, following the lines of Lemma 7.13 one shows $(\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}}^*_{\tau'})w_{\text{ker},\tau'} \rightharpoonup \dot{\mathcal{W}}^*\mathcal{W}\widetilde{u}_{\text{ker}}$ in $L^2(0,T;\mathcal{V}^*)$ as $\tau'\to 0$.

With the introduced piecewise functions we note that for every $t\in(t_{n-1},t_n]\subset[0,T]$ we have

$$\begin{aligned}\int_0^t \mathcal{W}^*_\tau \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau} &= \sum_{k=1}^{n-1}\mathcal{W}^*_k\int_{t_{k-1}}^{t_k}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau}\,\mathrm{d}s + \mathcal{W}^*_n\int_{t_{n-1}}^t \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau}\,\mathrm{d}s\\ &= \sum_{k=1}^{n-1}\mathcal{W}^*_k(\mathcal{W}_k\widetilde{u}_{\text{ker},k} - \mathcal{W}_{k-1}\widetilde{u}_{\text{ker},k-1}) - \mathcal{W}^*_n\mathcal{W}_{n-1}\widetilde{u}_{\text{ker},n-1} + \mathcal{W}^*_\tau(t)\widehat{\mathcal{W}\widetilde{u}}_{\text{ker},\tau}(t)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n-1} (\mathcal{W}_k^* - \mathcal{W}_{k+1}^*)\mathcal{W}_k \widetilde{u}_{\mathrm{ker},k} + \mathcal{W}_\tau^*(t)\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}(t) - \widetilde{u}_{\mathrm{ker},0} \\
&= \int_{t_{n-1}}^t (\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}}_\tau^*) w_{\mathrm{ker},\tau}\,\mathrm{d}s - \int_0^t (\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}}_\tau^*) w_{\mathrm{ker},\tau}\,\mathrm{d}s + \mathcal{W}_\tau^*(t)\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau}(t) - \widetilde{u}_{\mathrm{ker},0}. \quad (7.61)
\end{aligned}
$$

The first integral of the right-hand side of (7.61) vanishes uniformly in $L^\infty(0,T;\mathcal{V}^*)$ as $\tau \to 0$ because

$$
\begin{aligned}
\int_{t_{n-1}}^t \|(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}}_\tau^*) w_{\mathrm{ker},\tau}\|_{\mathcal{V}^*}\,\mathrm{d}s \quad &\le \quad \int_{t_{n-1}}^t \left\| \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \dot{\mathcal{W}}^*\,\mathrm{d}\eta \right\|_{\mathscr{L}(\mathcal{V}^*)} \mathrm{d}s\, \|\mathcal{W}_{n-1}\widetilde{u}_{\mathrm{ker},n-1}\|_{\mathcal{V}^*} \\
&\overset{(7.56a)}{\le} \frac{C_{\mathcal{V}\hookrightarrow\mathcal{H}}}{\tau} \int_{t_{n-1}}^t \int_{t_{n-1}}^{t_n} \|\dot{\mathcal{W}}\|_{\mathscr{L}(\mathcal{V})}\,\mathrm{d}\eta\,\mathrm{d}s M(\widetilde{u}_{\mathrm{ker},0}, \widetilde{f}_{\mathrm{ker}}) \\
&\le \quad \sqrt{\tau}\, C_{\mathcal{V}\hookrightarrow\mathcal{H}} \|\dot{\mathcal{W}}\|_{L^2[0,T;\mathscr{L}(\mathcal{V})]} M(\widetilde{u}_{\mathrm{ker},0}, \widetilde{f}_{\mathrm{ker}}).
\end{aligned}
$$

Then, by (7.61) and an adaptation of Theorem 7.41 we obtain

$$
\begin{aligned}
\mathcal{I}(\widetilde{\mathcal{B}}_{\tau'}\lambda_{\tau'}) &\overset{(7.57a)}{=} \mathcal{I}(\mathcal{W}_{\tau'}^* \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\mathrm{ker},\tau'}) + \mathcal{I}(\widetilde{\mathcal{A}}_{\tau'}\widetilde{u}_{\mathrm{ker},\tau'}) - \mathcal{I}(\widetilde{f}_{\mathrm{ker},\tau'}) \\
&\overset{*}{\rightharpoonup} \mathcal{W}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}} - \widetilde{u}_{\mathrm{ker},0} - \mathcal{I}(\dot{\mathcal{W}}^*\mathcal{W}\widetilde{u}_{\mathrm{ker}}) + \mathcal{I}(\widetilde{\mathcal{A}}\widetilde{u}_{\mathrm{ker}}) - \mathcal{I}(\widetilde{f}_{\mathrm{ker}}) \\
&\overset{(7.54)}{=} \mathcal{W}^*u - u_0 - \mathcal{I}(\dot{\mathcal{W}}^*u) + \mathcal{I}(\mathcal{W}^*\mathcal{A}u) - \mathcal{I}(\mathcal{W}^*f) =: \Lambda_{\widetilde{\mathcal{B}}^*}
\end{aligned}
$$

in $L^\infty(0,T;\mathcal{V}^*)$ as $\tau' \to 0$. By (7.50a) the function $\Lambda_{\widetilde{\mathcal{B}}^*} \in L^\infty(0,T;\mathcal{V}^*)$ has the distributional derivative $\widetilde{\mathcal{B}}^*\lambda$. Furthermore, by Lemma 3.30 the weak* convergence

$$
0 = \mathcal{I}(\mathcal{P}_\perp^*(0)\widetilde{\mathcal{B}}_{\tau'}^*\lambda_{\tau'}) = \mathcal{P}_\perp^*(0)\mathcal{I}(\widetilde{\mathcal{B}}_{\tau'}^*\lambda_{\tau'}) \overset{*}{\rightharpoonup} \mathcal{P}_\perp^*(0)\Lambda_{\widetilde{\mathcal{B}}^*} \quad \text{in } L^\infty(0,T;\mathcal{V}^*)
$$

holds as $\tau' \to 0$. This implies that $\Lambda_{\widetilde{\mathcal{B}}^*}$ is pointwise an element of the annihilator of $\mathcal{V}_{\mathrm{ker}}(0)$. Therefore, the Volterra integral equation

$$
\Lambda(t) = \widetilde{\mathcal{B}}_{\mathrm{left}}^{-*}(t)\Lambda_{\widetilde{\mathcal{B}}^*}(t) + \widetilde{\mathcal{B}}_{\mathrm{left}}^{-*}(t)\int_0^t \tfrac{\mathrm{d}}{\mathrm{d}s}\widetilde{\mathcal{B}}^*(s)\Lambda(s)\,\mathrm{d}s \qquad \text{in } \mathcal{Q} \tag{7.62}
$$

is well-defined. Note that $\widetilde{\mathcal{B}}$ maps into $\mathcal{V}_{\mathrm{ker}}^0(0) := (\mathcal{V}_{\mathrm{ker}}(0))^0$ and so does $\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathcal{B}}$ as well. Furthermore, one can show that for an arbitrary $f \in \mathcal{V}_{\mathrm{ker}}^0(0)$ the term $\widetilde{\mathcal{B}}_{\mathrm{left}}^{-*}(t)f$ is given by the partial solution ν_t of the saddle-point problem (7.43) with operators $\mathcal{R}_\mathcal{V}$ and $\widetilde{\mathcal{B}}(t)$ as well as right-hand side $(-f,0)$. Thus, $\widetilde{\mathcal{B}}_{\mathrm{left}}^{-*}$ is an element of $H^1[0,T;\mathscr{L}(\mathcal{V}_{\mathrm{ker}}^0(0),\mathcal{Q})] \hookrightarrow C([0,T],\mathscr{L}(\mathcal{V}_{\mathrm{ker}}^0(0),\mathcal{Q}))$; cf. Lemma 7.28.

The Volterra integral equation (7.62) has a unique solution $\Lambda \in L^\infty(0,T;\mathcal{Q})$ by Theorem 4.19. Let now $\varphi_{\mathcal{Q}^*} \in C_c^\infty(0,T;\mathcal{Q}^*)$ be arbitrary. Then there exists a $\varphi_\mathcal{V} \in H^1(0,T;\mathcal{V})$ with $\varphi_{\mathcal{Q}^*} = \widetilde{\mathcal{B}}\varphi_\mathcal{V}$, $\operatorname{supp}\varphi_\mathcal{V} = \operatorname{supp}\varphi_{\mathcal{Q}^*} \subset (0,T)$, and we get

$$
\begin{aligned}
\int_0^T \langle \dot{\varphi}_{\mathcal{Q}^*}, \Lambda \rangle\,\mathrm{d}t &= \int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathcal{B}}\varphi_\mathcal{V} + \widetilde{\mathcal{B}}\dot{\varphi}_\mathcal{V}, \Lambda \rangle\,\mathrm{d}t = \int_0^T \langle \widetilde{\mathcal{B}}^*\Lambda - \textstyle\int_0^t \tfrac{\mathrm{d}}{\mathrm{d}s}\widetilde{\mathcal{B}}^*\Lambda\,\mathrm{d}s, \dot{\varphi}_\mathcal{V} \rangle\,\mathrm{d}t \\
&= \int_0^T \langle \Lambda_{\widetilde{\mathcal{B}}^*}, \dot{\varphi}_\mathcal{V} \rangle\,\mathrm{d}t = -\int_0^T \langle \widetilde{\mathcal{B}}^*\lambda, \varphi_\mathcal{V} \rangle\,\mathrm{d}t = -\int_0^T \langle \varphi_{\mathcal{Q}^*}, \lambda \rangle\,\mathrm{d}t.
\end{aligned}
$$

This shows that λ is a distributional derivative of Λ. □

7.2.3. Discussion of Uniqueness

In Subsection 7.1.2 we reduced the proof of the uniqueness of the whole solution u to the uniqueness of the part of u, which maps into $\mathcal{V}_{\text{ker}}$. As mentioned in Remark 7.35, the analysis of $u_{\text{ker}} = \mathcal{P}_\perp u = \mathcal{P}_\perp \mathcal{W}\widetilde{u} = \mathcal{W}\widetilde{u}_{\text{ker}}$ is connected to parabolic PDEs on evolving domains. Therefore, we can use the arguments of the theory of PDEs on evolving domains for the uniqueness of solutions of the operator DAE (7.40). Anyway, the uniqueness of solutions is in general unsolved [VouR18, Rem. 2.4]. The open problem is the density of smooth function in the solution space; cf. [VouR18, Rem. 3.1]. For our purpose this is equivalent to the question: Given the Hilbert space

$$W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot)) := \{u \in L^2(0,T;\mathcal{V}) \,|\, \mathcal{P}_\perp u = u,\ \mathcal{P}^*_\perp u \in H^1(0,T;\mathcal{V}^*),\ \dot{\mathcal{P}}^*_\perp u \in L^2(0,T;\mathcal{V}^*)\}$$

equipped with the norm

$$\|u\|^2_{W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot))} = \|u\|^2_{L^2(0,T;\mathcal{V})} + \|\mathcal{P}^*_\perp u\|^2_{H^1(0,T;\mathcal{V}^*)} + \|\dot{\mathcal{P}}^*_\perp u\|^2_{L^2(0,T;\mathcal{V}^*)},$$

is the closure

$$U_{\text{ker}} := \operatorname{clos}_{\|\cdot\|_{W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot))}} \{u \in W^{1,2}(0,T;\mathcal{V},\mathcal{H}) \,|\, \mathcal{P}_\perp u = u\}$$

equal to the whole space $W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot))$?

Remark 7.43. By Lemma 7.38 and Remark 7.40, the function $u_{\text{ker}} = \mathcal{W}\widetilde{u}_{\text{ker}}$ is an element of the space $W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot))$, where $\widetilde{u}_{\text{ker}}$ is the constructed solution of (7.53) from Subsection 7.2.2.

The usual techniques to prove $U_{\text{ker}} = W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot))$ are not applicable, since the space $W^{1,2}(0,T;\mathcal{V}_{\text{ker}}(\cdot),\mathcal{V}^*_{\text{ker}}(\cdot))$ is not closed under translation in time; cf. [Wlo87, p. 393 ff.] or [Str66, Th. 3.1]. One can show at least that, if the solution is an element of U_{ker}, then it is unique. However, it is unclear, if every solution is an element of U_{ker}; cf. [VouR18, Ch. 3].

We finish this section with two cases with unique solutions. The first result is inspired by the results of [AlpES15].

Theorem 7.44 (Uniqueness of Solutions I)**.** *Let the assumptions of Theorem 7.41 be satisfied. Suppose that there exists an operator-valued function $\mathcal{W}$, which satisfies the assertion of Theorem 7.31 and can be pointwise extended to an invertible operator in $\mathscr{L}(\mathcal{H})$ at almost every time-point $t \in [0,T]$, such that $\mathcal{W} \in H^1[0,T;\mathscr{L}(\mathcal{H})]$ with a pointwise inverse $\mathcal{W}^{-1} \in H^1[0,T;\mathscr{L}(\mathcal{H})]$.*

Then there is only one solution (u,λ) of operator DAE (7.40), which satisfies a) and b) from Theorem 7.41. Furthermore, the solution satisfies $u \in C([0,T],\mathcal{H})$ and $\Lambda \in C([0,T],\mathcal{Q})$ from Theorem 7.42 with $u(0) = u_0$ and $\Lambda(0) = 0$. The solution map $(f,g,u_0) \mapsto (u,\Lambda)$ is linear and continuous.

Proof. By the assumptions on $\mathcal{W}$ the operator $\widetilde{\mathcal{M}} := \mathcal{W}^*\mathcal{W} \in H^1[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$ is uniformly elliptic and $\mathcal{J} := \frac{1}{2}(\mathcal{W}^*\dot{\mathcal{W}} - \dot{\mathcal{W}}^*\mathcal{W}) \in L^2[0,T;\mathscr{L}(\mathcal{H},\mathcal{H}^*)]$ is pointwise skew-adjoint. With the defined operators we can rewrite the operator DAE (7.50) as

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\widetilde{\mathcal{M}}(t)\widetilde{u}) + (\widetilde{\mathcal{A}}(t) + \mathcal{J}(t) - \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathcal{M}}(t))\widetilde{u} - \widetilde{\mathcal{B}}^*(t)\lambda = \widetilde{f}(t) \quad \text{in } \mathcal{V}^*, \tag{7.63a}$$

$$\widetilde{\mathcal{B}}(t)\widetilde{u} = g(t) \quad \text{in } \mathcal{Q}^*. \tag{7.63b}$$

As usual we split $\widetilde{u}$ into $\widetilde{u}_{\text{c}}$ given by (7.52) and the remainder $\widetilde{u}_{\text{ker}}$ with images in $\mathcal{V}_{\text{ker}}(0)$. Lemma 7.22 then proves the uniqueness of $\widetilde{u}_{\text{ker}} \in L^2(0,T;\mathcal{V}_{\text{ker}}(0)) \cap C([0,T],\mathcal{H}_{\text{ker}}(0))$ and therefore of $\widetilde{u}$ as well. Note that the time-dependence of $\widetilde{\mathcal{B}}$ does not affect the result of Lemma 7.22, since its kernel is time-independent. Because every solution u of (7.40) implies a solution $\widetilde{u} = \mathcal{W}^{-1}u$ of (7.63) and vice versa, u is unique and continuous with image in $\mathcal{H}$ as well. With $\|u(t)\|_{\mathcal{X}} = \|\mathcal{W}(t)\widetilde{u}(t)\|_{\mathcal{X}} \lesssim \|\widetilde{u}(t)\|_{\mathcal{X}}$,

$\mathcal{X} \in \{\mathcal{H}, \mathcal{V}\}$ by the assumptions on $\mathcal{W}$, the remaining assertions follow by the steps of Theorem 7.19 and 7.42. □

Remark 7.45. If the kernel of $\mathcal{B}$ is time-independent, then we can choose $\mathcal{W} = \mathrm{id}_{\mathcal{V}}$ and the assumptions on $\mathcal{W}$ in Theorem 7.50 are satisfied.

Remark 7.46. The assumptions on the right-hand sides in Theorem 7.44 can be weakened such that g is only an element of $W^{1,1}(0,T;\mathcal{Q}^*)$ and f can be split into $f = f_1 + f_2$, where $f_1 \in L^2(0,T;\mathcal{V}^*)$ and $f_2 \in L^1(0,T;\mathcal{H}^*)$. Then there exists a unique solution (u, λ), which satisfies the conditions a), b), and d) of Theorem 7.21. The associated proof follows the lines of Theorem 7.21 with the results of Theorems 7.41 and 7.44.

As the second case with a unique solution, we transfer the results of Theorem 7.24 to operator DAE (7.50) with a time-dependent operator $\mathcal{B}$.

Theorem 7.47 (Uniqueness of Solutions II)**.** *In addition to the assumptions of Theorem 7.41, suppose that $\mathcal{B} \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$ and $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ pointwise self-adjoint and uniformly elliptic on $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{A}_2 \in L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{H}^*)]$. Assume that $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}(0)u_0 = g(0)$, and that f is an element of $L^2(0,T;\mathcal{H}^*)$.*

Then there exists a unique solution $u \in C([0,T],\mathcal{V}) \cap H^1(0,T;\mathcal{H})$ and $\lambda \in L^2(0,T;\mathcal{Q})$ with $u(0) = u_0$, which depends linearly and continuously on the data f, g, and u_0.

Proof. We follow the steps of Theorem 7.24. For this let us denote $\mathcal{A}_1$ and $\mathcal{A}_2$ by $\mathcal{A}^{[1]}$ and $\mathcal{A}^{[2]}$, respectively. We split u into u_{c} and u_{ker} where $u_{\mathrm{c}} := \mathcal{B}^-_{\mathcal{A}^{[1]}} g$. By Lemma 7.28 the function $u_{\mathrm{c}} \in H^1(0,T;\mathcal{V})$ satisfies $\mathcal{A}^{[1]}(t)u_{\mathrm{c}}(t) \in \mathcal{V}^0_{\mathrm{ker}}(t)$ at almost every time-point $t \in [0,T]$. The functions u_{ker} and λ are given by a possible solution $(\widetilde{u}_{\mathrm{ker}}, \lambda)$ of the operator DAE (7.53) with $\mathcal{W}$ from Lemma 7.33, where $u_{\mathrm{ker}} = \mathcal{W}\widetilde{u}_{\mathrm{ker}}$. The right-hand side $\widetilde{f}_{\mathrm{ker}}$ is given by $\mathcal{W}^*(f^{[1]}_{\mathrm{ker}} + f^{[2]}_{\mathrm{ker}})$ with

$$f^{[1]}_{\mathrm{ker}} := \mathcal{A}^{[1]} u_{\mathrm{c}} \in W^{1,\infty}(0,T;\mathcal{V}^*) \qquad f^{[2]}_{\mathrm{ker}} := f - \mathcal{A}^{[2]} u_{\mathrm{c}} - \dot{u}_{\mathrm{c}} \in L^2(0,T;\mathcal{H}^*)$$

and the initial value $\widetilde{u}_{\mathrm{ker},0} = \mathcal{W}(0)u_{\mathrm{ker},0} = u_{\mathrm{ker},0} = u_0 - \mathcal{B}^-_{\mathcal{A}^{[1]}}(0)g(0) \in \mathcal{V}_{\mathrm{ker}}(0)$.

Let us define $\mathcal{W}_n \in \mathscr{L}(\mathcal{V})$, $f^{[1]}_{\mathrm{ker},n} \in \mathcal{V}^*$, and $\widetilde{\mathcal{A}}^{[1]}_n \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ as the evaluation of $\mathcal{W}$, $f^{[1]}_{\mathrm{ker}}$, and $\mathcal{W}^*\mathcal{A}^{[1]}\mathcal{W}$ at the time-point t_n, respectively. Further, $f^{[2]}_{\mathrm{ker},n} \in \mathcal{H}^*$ and $(\mathcal{A}^{[2]}\mathcal{W})_n \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$ are set to the integral means of $f^{[2]}_{\mathrm{ker}}$ and $\mathcal{A}^{[2]}\mathcal{W}$, respectively; cf. (7.9). We then consider the temporal discretization (7.55) of (7.53) with the discrete operator $\widetilde{\mathcal{A}}_n := \widetilde{\mathcal{A}}^{[1]}_n + \mathcal{W}^*_n(\mathcal{A}^{[2]}\mathcal{W})_n$ and the right-hand side $\widetilde{f}_{\mathrm{ker},n} := \mathcal{W}^*_n(f^{[1]}_{\mathrm{ker},n} + f^{[2]}_{\mathrm{ker},n})$. With the steps of Theorem 7.24 one proves that the statements of Subsection 7.2.2 are still valid with $\frac{\mu_{\mathcal{A}}}{2}$ instead of $\mu_{\mathcal{A}}$ in Lemma 7.36 if τ is small enough. On the other hand, since $\mathcal{A}^{[1]}u_c \in \mathcal{V}^0_{\mathrm{ker}}(\cdot)$ implies $\mathcal{W}^* f^{[1]}_{\mathrm{ker}} \in \mathcal{V}^0_{\mathrm{ker}}(0)$, testing (7.55) with $\tau D_\tau \widetilde{u}_{\mathrm{ker},n} := \widetilde{u}_{\mathrm{ker},n} - \widetilde{u}_{\mathrm{ker},n-1} \in \mathcal{V}_{\mathrm{ker}}(0)$ leads to

$$\tau\langle D_\tau(\mathcal{W}\widetilde{u})_{\mathrm{ker},n}, \mathcal{W}_n D_\tau \widetilde{u}_{\mathrm{ker},n} \rangle + \tau\langle \widetilde{\mathcal{A}}^{[1]}_n \widetilde{u}_{\mathrm{ker},n}, D_\tau \widetilde{u}_{\mathrm{ker},n} \rangle = \langle f^2_{\mathrm{ker},n} - (\mathcal{A}^{[2]}\mathcal{W})_n \widetilde{u}_{\mathrm{ker},n}, \mathcal{W}_n D_\tau \widetilde{u}_{\mathrm{ker},n} \rangle. \tag{7.64}$$

Note that $\mathcal{W} \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V})]$ by Lemma 7.33. Therefore, equation (7.64) implies the estimate

$$\begin{aligned}
&\tau\|D_\tau(\mathcal{W}\widetilde{u})_{\mathrm{ker},n}\|^2_{\mathcal{H}} + \langle \widetilde{\mathcal{A}}^{[1]}_n \widetilde{u}_{\mathrm{ker},n}, \widetilde{u}_{\mathrm{ker},n} - \widetilde{u}_{\mathrm{ker},n-1} \rangle \\
&\overset{(7.64)}{=} \tau\langle f^{[2]}_{\mathrm{ker},n} - (\mathcal{A}^{[2]}\mathcal{W})_n \widetilde{u}_{\mathrm{ker},n}, D_\tau(\mathcal{W}\widetilde{u})_{\mathrm{ker},n} \rangle \\
&\qquad - \tau\Big\langle f^{[2]}_{\mathrm{ker},n} - (\mathcal{A}^{[2]}\mathcal{W})_n \widetilde{u}_{\mathrm{ker},n} - D_\tau(\mathcal{W}\widetilde{u})_{\mathrm{ker},n}, \frac{\mathcal{W}_n - \mathcal{W}_{n-1}}{\tau} \widetilde{u}_{\mathrm{ker},n-1} \Big\rangle \\
&\overset{(3.8)}{\leq} \frac{3\tau}{2}\|f^{[2]}_{\mathrm{ker},n} - (\mathcal{A}^{[2]}\mathcal{W})_n \widetilde{u}_{\mathrm{ker},n}\|^2_{\mathcal{H}^*} + \frac{3C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}}{2\tau}\|(\mathcal{W}_n - \mathcal{W}_{n-1})\widetilde{u}_{\mathrm{ker},n-1}\|^2_{\mathcal{V}} + \frac{\tau}{2}\|D_\tau(\mathcal{W}\widetilde{u})_{\mathrm{ker},n}\|^2_{\mathcal{H}}
\end{aligned}$$

$$\overset{(7.12)}{\leq} \frac{3\tau}{2}\|f^{[2]}_{\ker,n} - (\mathcal{A}^{[2]}\mathcal{W})_n\widetilde{u}_{\ker,n}\|^2_{\mathcal{H}^*} + \tau\frac{3C^2_{\mathcal{V}\hookrightarrow\mathcal{H}}}{2}\|\dot{\mathcal{W}}\|^2_{L^\infty[0,T;\mathscr{L}(\mathcal{V})]}\|\widetilde{u}_{\ker,n-1}\|^2_{\mathcal{V}} + \frac{\tau}{2}\|D_\tau(\mathcal{W}\widetilde{u})_{\ker,n}\|^2_{\mathcal{H}}.$$

With estimates analogous to (7.36)–(7.38) we get the inequality

$$\sum_{k=1}^{n}\tau\|D_\tau(\mathcal{W}\widetilde{u})_{\ker,k}\|^2_{\mathcal{H}} + \|\widetilde{u}_{\ker,n}\|^2_{\mathcal{V}} + \sum_{k=1}^{n}\tau\|\widetilde{u}_{\ker,k} - \widetilde{u}_{\ker,k-1}\|^2_{\mathcal{V}} \lesssim \|u_{\ker,0}\|^2_{\mathcal{V}} + \int_0^T \|f^{[2]}_{\ker}\|^2_{\mathcal{H}^*}\,\mathrm{d}s + M^2$$

for every $n = 1,\dots,N$ and $M = M(\widetilde{u}_{\ker,0}, \widetilde{f}_{\ker})$ from Lemma 7.36. This shows that $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{W}\widetilde{u}}_{\ker,\tau}$ is bounded in $L^2(0,T;\mathcal{H})$ and $\widetilde{u}_{\ker,\tau}$ in $L^\infty(0,T;\mathcal{V})$ both independently of τ.

By [Zei90a, Prop. 23.19] and the continuity of $\mathcal{W}$, we have $u_{\ker} = \mathcal{W}\widetilde{u}_{\ker} \in L^\infty(0,T;\mathcal{V})\cap H^1(0,T;\mathcal{H})$ with the limit $\widetilde{u}_{\ker}$ from Lemmas 7.38 and 7.39. This allows us to test the operator DAE (7.53) with $u_{\ker}$ and $\dot{u}_{\ker}$, which leads to $u_{\ker} \in C([0,T],\mathcal{V})$ and the continuity with respect to $(f_{\ker}, u_{\ker,0})$; see Theorems 7.19 and 7.24. The assertion on u follows by the properties of $u_{\ker}$.

Finally, the Lagrange multiplier λ and the solution u are the unique solution of the time-dependent saddle point problem (7.43) with right-hand sides $f - \dot{u} \in L^2(0,T;\mathcal{H}^*)$ and $g \in H^1(0,T;\mathcal{Q}^*) \hookrightarrow L^2(0,T;\mathcal{Q}^*)$. The properties of λ then follow by Theorem 7.27 and Lemma 4.9. □

Remark 7.48. Example 7.5 is uniquely solvable by Theorem 7.47, if $u_0 \in H^2_0(0,1)$ is consistent, $f \in L^2(0,T;L^2(0,1))$, and $\Phi \in W^{1,\infty}(0,T)$.

7.3. Main Results

In the previous two sections 7.1 and 7.2 we considered only cases where $\mathcal{B}$ or $\mathcal{M}$ were constant in time. In this section, we finally investigate operator DAEs of the form (7.1), where all operators, namely $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{B}$, are simultaneously time-dependent. The results of this section are combinations of the ideas of the previous two sections. Since the associated proofs are rather technical and mostly done in the sections 7.1 and 7.2, we will only sketch them.

Let us start with the existence of solutions of the operator DAE (7.1).

Theorem 7.49 (Existence of Solutions of Operator DAEs with Time-Dependent Operators)**.** *Let Assumption 7.2 be satisfied. Then the operator DAE (7.1) has at least one solution (u,λ) which fulfills*

a) $u \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$,

b) $\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda$ *for an* $\Lambda \in L^\infty(0,T;\mathcal{Q})$,

c) $\mathcal{P}^*_\perp \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) \in L^2(0,T;\mathcal{V}^*)$,

d) $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u) - \mathcal{B}^*\lambda \in L^2(0,T;\mathcal{V}^*)$.

Proof. Let $\mathcal{W}$ be defined as in Lemma 7.33. For the proof we split $\widetilde{u} = \mathcal{W}^{-1}u$ into $\widetilde{u}_c$ given by (7.52) and the remainder $\widetilde{u}_{\ker}$ with images in $\mathcal{V}_{\ker}(0)$. For $\widetilde{u}_{\ker}$, we define $\widetilde{f}_{\ker} := \widetilde{f} - \widetilde{\mathcal{A}}\widetilde{u}_c - \mathcal{W}^*\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}\mathcal{W}\widetilde{u}_c) \in L^2(0,T;\mathcal{V}^*)$ and consider, similar to (7.8) and (7.55), the temporally discretized system

$$\begin{aligned}
\mathcal{W}^*_n D_\tau(\mathcal{M}\mathcal{W}\widetilde{u}_{\ker})_n + (\widetilde{\mathcal{A}}_n - \tfrac{1}{2}D_\tau\widetilde{\mathcal{M}}_n)\widetilde{u}_{\ker,n} - \widetilde{\mathcal{B}}^*_n\lambda_n &= \widetilde{f}_{\ker,n} &&\text{in } \mathcal{V}^*,\\
\widetilde{\mathcal{B}}_n\widetilde{u}_{\ker,n} &= 0 &&\text{in } \mathcal{Q}^*.
\end{aligned}$$

Here, we set $\widetilde{\mathcal{M}}_n := \mathcal{W}(t_n)^*\mathcal{M}(t_n)\mathcal{W}(t_n)$. The discrete derivatives $D_\tau(\mathcal{M}\mathcal{W}\widetilde{u}_{\ker})_n$ and $D_\tau\widetilde{\mathcal{M}}_n$ are given by

$$D_\tau(\mathcal{M}\mathcal{W}\widetilde{u}_{\ker})_n = \frac{\mathcal{M}_n\mathcal{W}_n\widetilde{u}_{\ker,n} - \mathcal{M}_{n-1}\mathcal{W}_{n-1}\widetilde{u}_{\ker,n-1}}{\tau} \quad\text{and}\quad D_\tau\widetilde{\mathcal{M}}_n = \frac{\widetilde{\mathcal{M}}_n - \widetilde{\mathcal{M}}_{n-1}}{\tau}.$$

Then the assertions a), c), and d) follow along the lines of Theorem 7.14 and 7.41, where $\mathcal{B}^*\lambda = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda_{\mathcal{B}^*}$ also holds with a $\Lambda_{\mathcal{B}^*} \in L^\infty(0,T;\mathcal{V}^*)$. We note that $\mathcal{W}^*\mathcal{M}\mathcal{W}\widetilde{u}_{\ker} \in H^1(0,T;\mathcal{V}^*_{\ker}(0))$ is satisfied by arguments similar to those of Lemma 7.39. Finally, assertion b) follows as in Theorem 7.42. □

For the operator DAE (7.1) with a constant operator $\mathcal{B}$ we proved the uniqueness of solutions in Theorem 7.19. For a time-dependent operator $\mathcal{B}$, we mentioned in Subsection 7.2.3, that the question of unique solutions is related to the unsolved problem of unique solutions of PDEs on evolving domains. We proved the uniqueness for two special cases. These results are translatable to systems with a nonconstant operator $\mathcal{M}$.

Theorem 7.50 (Uniqueness of Solutions of Operator DAEs with Time-Dependent Operators I)**.** *Assume that g is an element of $W^{1,1}(0,T;\mathcal{Q}^*)$ and that f can be split into $f = f_1 + f_2$, where $f_1 \in L^2(0,T;\mathcal{V}^*)$ and $f_2 \in L^1(0,T;\mathcal{H}^*)$. Let Assumptions 7.2.i), ii), and iv) be satisfied. Suppose that there exists an operator-valued function $\mathcal{W}$, which fulfills the assumptions of Theorem 7.44.*

Then there exists only one solution (u,λ) of the operator DAE (7.1), which satisfies a) and b) from Theorem 7.49. Furthermore, the solution satisfies $u \in C([0,T],\mathcal{H})$ and $\Lambda \in C([0,T],\mathcal{Q})$ with $u(0) = u_0$ and $\Lambda(0) = 0$. The map from the data (f,g,u_0) to (u,Λ) is linear and continuous.

Proof. We set $\widetilde{\mathcal{M}} = \mathcal{W}^*\mathcal{M}\mathcal{W}$ and $\mathcal{J} = \frac{1}{2}(\mathcal{W}^*\mathcal{M}\dot{\mathcal{W}} - \dot{\mathcal{W}}^*\mathcal{M}\mathcal{W})$. Then the assertions follow with the arguments of Theorem 7.44 and Remark 7.46. □

Theorem 7.51 (Uniqueness of Solutions of Operator DAEs with Time-Dependent Operators II)**.** *In addition to Assumption 7.2, suppose that the operator $\mathcal{B}$ satisfies $\mathcal{B} \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V},\mathcal{Q}^*)]$. Let $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ be satisfied, where $\mathcal{A}_1 \in W^{1,\infty}[0,T;\mathscr{L}(\mathcal{V},\mathcal{V}^*)]$ is pointwise self-adjoint and uniformly elliptic on $\mathcal{V}_{\ker}$ and $\mathcal{A}_2 \in L^\infty[0,T;\mathscr{L}(\mathcal{V},\mathcal{H}^*)]$. Assume that $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}(0)u_0 = g(0)$, and that f is an element of $L^2(0,T;\mathcal{H}^*)$.*

Then there exists a unique solution $u \in C([0,T],\mathcal{V}) \cap H^1(0,T;\mathcal{H})$ and $\lambda \in L^2(0,T;\mathcal{Q})$ with $u(0) = u_0$ of the operator DAE (7.1). The solution depends linearly and continuously on the data f, g, and u_0.

Proof. The proof follows the lines of Theorems 7.24 and 7.47. □

Remark 7.52. Theorem 7.51 can be extended to semi-linear operator DAEs of the form (7.1) with non-linear right-hand sides $f\colon [0,T] \times \mathcal{V} \to \mathcal{H}^*$ satisfying the assumptions of Section 6.4. This can be proven like Theorems 6.15 and 6.19.

Part C.

Temporal Discretization of Operator Differential-Algebraic Equations

In this part we study time-stepping methods for constrained PDEs. For their analysis, again we use the framework of operator DAE. In this thesis, we focus on the systems of Chapter 6 with time-independent operators. We directly discretize the systems in time without a spatial discretization beforehand, i.e., we use the Rothe method. This leads to infinite-dimensional stationary systems for every time step; see e.g. [Alt15, Sec. 10.2]. In particular, for each individual time step the stationary problem then can be discretized individually in space [SchB98, p. 137]. This allows adaptive methods for every discrete time-point, whose advantages were studied for linear PDEs in [Bor90; Bor91; Bor92; SchB98]. Furthermore, the Rothe method allows to derive temporal error bounds which are independent of the mesh width of spatial discretization [HocO10, Ch. 2 & p. 212]. This is vital for spatially discretized systems, since the temporal convergence order of finite and infinite-dimensional systems may differ; see e.g. [OstR92, Th. 2], [DebS05, Th. 7], and [HocO05a, Sec. 4.3]. If the spatial mesh gets finer the order of the infinite system dominates the convergence of the finite dimensional discretization; cf. [ProR74]. Note that, the (discrete) norms of $\mathcal{V}$ and $\mathcal{H}$ are equivalent under a fixed spatial discretization. Anyway, Figure C.1 illustrates a case where the convergence order for the finite dimensional, spatially discretized system is two, for the infinite-dimensional original system it is one and a half, and the convergence rate fades for large numbers of degrees of freedom.

We analyze Runge-Kutta methods and exponential integrators in this thesis. Note that a time-integration scheme for DAEs and parabolic PDEs must respect their infinitely stiff nature, in the sense that both classes of differential equations can be approximated by a sequence of ODEs, where every ODE in the sequence is stiffer than its predecessor; see [BreCP96, Sec. 4.5], [Zei90a, Sec. 23.7], and [OstR92, p. 404]. Because of the stiff behavior of the problem, we consider implicit, algebraically

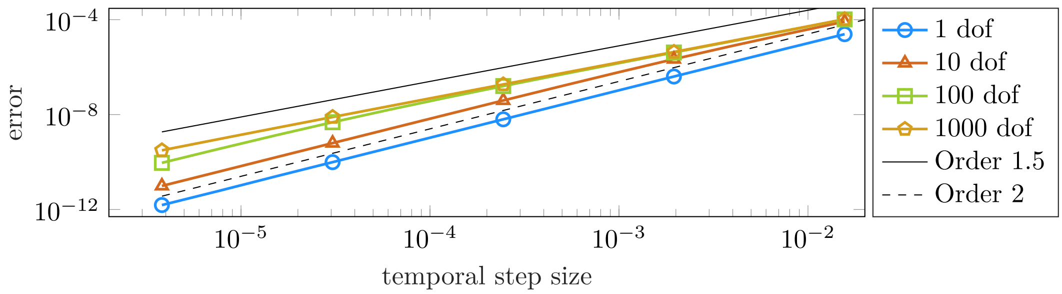

Figure C.1.: Temporal error of the (constrained) PDE (9.50) temporally discretized by scheme (9.10) for different spatial degrees of freedom (dof).

stable Runge-Kutta methods; see [HaiW96, Sec. IV.12] and [HunV03, Ch. II, Sec. 1.4]. On the other hand, exponential integrators are always suited for stiff semi-linear problems, since they are based on the exact solution of the linear part, which is responsible for the stiffness [HocO10]. In particular, this allows large time steps for *explicit* exponential integrators [HocLS98, p. 1572]. These integrators have the advantages of explicit methods like limited number of evaluations of the nonlinear inhomogeneity and no nonlinear root-finding problem.

This part is organized as follows. In Chapter 8 we temporally discretize the operator DAEs from Sections 6.1 and 6.2 with Runge-Kutta methods. Here, the operators $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{B}$ are time-independent and the right-hand sides f, g only depend on time. Since the operator DAEs are sensitive to perturbations of g, we introduce a regularization which maintains the saddle-point structure of the original systems. We temporally discretize the regularized system by implicit, algebraically stable Runge-Kutta schemes, starting with the implicit Euler method. The convergence of the time-discrete solution to the solution of the original operator DAE is studied under minimal assumptions on the data. Afterwards, we analyze the convergence order and discuss it by means of numerical examples. Chapter 9 is devoted to the application of explicit exponential integrators to the operator DAE from Section 6.4, i.e., the right-hand side f depends on time and the state u. By exploiting the structure of the operator DAE, semi-explicit integration schemes are constructed based on the exponential Euler and Runge methods. We study their convergence orders and derive the order conditions for schemes with an order up to three. Afterwards, we consider the temporal discretization of the Lagrange multiplier λ. Finally, the performance of the numerical schemes is presented, where an efficient solver for intermediate transient problems is used.

Sections 8.1 and 8.3 are essential copies of [AltZ18b, Sec. 3 & 5]. Lemmas 8.2, 8.3, and 8.5 were originally proven by Robert Altmann. We omit the proofs of the latter two lemmas and give an alternative proof of the first one, which is closer to the definition of the differentiation index. Section 8.2 extends the results of [AltZ18b, Sec. 4] to less regular right-hand sides. The results in Section 8.5 are based on and extend [AltZ18c, Sec. 6]. The author of this thesis originally elaborated [AltZ18b, Sec. 4 f.] and [AltZ18c, Sec. 6]. All remaining results of Chapter 8 are unpublished.

Sections 9.1, 9.2, and Subsection 9.4.2 are essentially copies of [AltZ20, Sec. 2.4, 3 f., & 5.2]. The schemes of the exponential integrators were elaborated in close cooperation between the two authors of the book chapter. The algorithms in Subsections 9.1.1 and 9.2.1 were developed by Robert Altmann. Furthermore, Robert Altmann originally implemented the numerical example in Subsection 9.4.2. The convergence analysis in Subsections 9.1.2 and 9.2.2 was originally elaborated by the author of this thesis. The same holds for the results of [AltZ20, Sec. 5.1] which are extended in this thesis in Subsection 9.4.1. All remaining results of Chapter 9 are unpublished.

8. Runge-Kutta Methods

In this chapter we investigate Runge-Kutta methods as time integration schemes for operator DAEs of the form

$$\dot{u}(t) + \mathcal{A}u(t) - \mathcal{B}^*\lambda(t) = f(t) \qquad \text{in } \mathcal{V}^*, \tag{8.1a}$$

$$\mathcal{B}u(t) \qquad\qquad\qquad = g(t) \qquad \text{in } \mathcal{Q}^*. \tag{8.1b}$$

The spaces $\mathcal{V}$ and $\mathcal{Q}$ are separable Hilbert spaces. We assume that a third Hilbert space $\mathcal{H}$ exists such that $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ is a Gelfand triple. All operators are time-independent. The existence, uniqueness, and regularity of solutions of (8.1) were studied in Sections 6.1 and 6.2. In addition to the assumptions of these sections, we suppose that the operator $\mathcal{A}$ is elliptic on $\mathcal{V}_{\mathrm{ker}}$. This is not necessary for the existence of an approximation for small enough step sizes τ, but will imply strong convergence of the approximations to the solution. We summarize the assumptions in the following.

Assumption 8.1 (Operators, Right-Hand Sides, and Initial Value of Operator DAE (8.1)).

i) The operator $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ is inf-sup stable and the operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is elliptic on $\mathcal{V}_{\mathrm{ker}} := \ker \mathcal{B}$ with an ellipticity constant $\mu_{\mathcal{A}} > 0$.
ii) The right-hand sides satisfy $f \in L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*)$ and $g \in W^{1,1}(0,T;\mathcal{Q}^*)$
iii) The initial value u_0 fulfills $u_0 \in \mathcal{H}_{\mathrm{ker}} + \mathcal{B}^- g(0) \subset \mathcal{H}$.

The application of Runge-Kutta schemes to parabolic PDEs and operator ODEs without a spatial discretization is well-studied; see for example [Emm04; Rou13, Ch. 8 each] for the study of existence of solutions and [Cro75; EmmT10; GonO99; LubO93; LubO95b; OstR92] for convergence orders. In contrast, only few results are known for constrained PDEs and operator DAEs. In [Emm00; Emm01] the temporal discretization of the unsteady Stokes equation and the incompressible Navier-Stokes equations by the implicit Euler scheme is studied. The author of [Emm00; Emm01] proves convergence orders and studies by the nature of the considered constrained PDEs only systems with constraints with homogeneous right-hand sides. The convergence of the temporally discretized solutions of the operator DAE (8.1) given by the implicit Euler scheme is studied in [Alt15, Ch. 10]. There the author focuses on the weak convergence of the approximation, as we did in Subsections 7.1.1 and 7.2.2, and on the influence of perturbations. The convergence order of Runge-Kutta schemes applied to a class of PDAEs with d-dimensional boxes as spatial domains is investigated in [Deb04; DebS05]. This class excludes systems, which are considered here, and vice versa.

In [Alt15, Sec. 8.2] it is shown that the spatial discretization of the operator DAE (8.1) by the *mixed Galerkin method* [Bra07, p. 134 ff.] leads to a DAE of index 2. Therefore, this DAE under temporal discretization with Runge-Kutta methods may suffer from numerical instabilities. In particular, perturbations of the discretized right-hand side g in (8.1b) may be amplified by the inverse of the step size, i.e., τ^{-1}; see [HaiLR89, Th. 4.2] for the finite-dimensional case and [Alt15, Rem. 10.12] for the infinite-dimensional one with the implicit Euler scheme. Therefore, we consider a regularization of the operator DAE (8.1) in Section 8.1. The resulting operator DAE has the same solution, but the finite-dimensional DAE, which we get by Galerkin discretization, is of index 1. In Section 8.2 we consider the discretization of the regularized operator DAE satisfying Assumption 8.1 with the implicit Euler scheme. Under slightly stronger assumptions we analyze algebraically L-stable Runge-Kutta methods in Section 8.3. The results, which include the weak and strong convergence of the discrete solution, are extended to algebraically stable Runge-Kutta methods with $|R(\infty)| \leq 1$ in Section 8.4. The aim of Sections 8.2–8.4 is the analysis of the qualitative behavior of the convergence under minimal assumptions on the data and on the Runge-Kutta methods. The convergence analysis does not require any additional regularity of the exact solution. For regular solutions the actual convergence order of the Runge-Kutta methods is considered in Section 8.5. Note that, higher regularity is equivalent to compatibility conditions on the data, which can be restrictive; see [EmmT10, p. 786] and [Tem77]. However, in addition to the convergence rate we investigate the error under perturbation of the data in Section 8.5. Finally, the performance of the numerical schemes is presented in Section 8.6.

8.1. Regularization

As mentioned in the introduction of Chapter 8, the spatial discretization of the linear semi-explicit operator DAE (8.1) leads to a DAE of index 2. Therefore, the temporal discretization of the operator DAE by Runge-Kutta methods is highly sensitive to perturbations.

Motivated by the Gear-Gupta-Leimkuhler formulation for multibody systems [GeaGL85], we extend the system (8.1) by an additional Lagrange multiplier $\gamma\colon [0,T] \to \mathcal{Q}$. With this, we enforce the system to satisfy additionally the hidden constraint, i.e., the derivative of constraint (8.1b). The proposed regularization makes the system more robust and achieves that a spatial discretization leads to a DAE of index one rather than index two.

8.1.1. Finite-Dimensional Case

Consider the DAE, which results from a spatial discretization of system (8.1) by mixed finite elements. With the symmetric positive definite mass matrix $M \in \mathbb{R}^{n_x \times n_x}$ as discretized version of $(\cdot,\cdot)_{\mathcal{H}}$, the matrix $A \in \mathbb{R}^{n_x \times n_x}$ as discrete version of $\mathcal{A}$, and the constraint matrix $B \in \mathbb{R}^{n_\mu \times n_x}$, which we assume to have full row rank, the DAE has the form

$$M\dot{x}(t) + Ax(t) - B^T\mu(t) = d(t), \tag{8.2a}$$
$$Bx(t) = h(t). \tag{8.2b}$$

Here, $x\colon [0,T] \to \mathbb{R}^{n_x}$ denotes the coefficient vector associated with a basis of the finite element space, which approximates the solution u. The vector-valued function $\mu\colon [0,T] \to \mathbb{R}^{n_\mu}$ corresponds to the Lagrange multiplier λ in the continuous setting, as well as $d\colon [0,T] \to \mathbb{R}^{n_x}$ and $h\colon [0,T] \to \mathbb{R}^{n_\mu}$ to the right-hand sides f and g, respectively. The initial condition is given by $x(0) = x_0$ with x_0 being the discrete version of u_0.

By Remark 2.3 the DAE (8.2) is of index 2 by the assumptions made. As mentioned in the beginning of Section 8.1, we reduce the index, and thus regularize the system equations, by adding the derivative of the constraint (8.2b) and an additional Lagrange multiplier $\gamma\colon [0,T] \to \mathbb{R}^{n_\mu}$. With some regular matrix $C \in \mathbb{R}^{n_\mu \times n_\mu}$, the extended DAE reads

$$M\dot{x}(t) + Ax(t) - B^T\mu(t) - B^T\gamma(t) = d(t), \tag{8.3a}$$
$$Bx(t) + C\gamma(t) = h(t), \tag{8.3b}$$
$$B\dot{x}(t) = \dot{h}(t). \tag{8.3c}$$

Here, the Lagrange multiplier γ measures the difference between the right-hand side of (8.3b) and the primitive of the right-hand side of (8.3c). However, we show that the system (8.3) has the same solution as DAE (8.2) but a lower index.

Lemma 8.2. *Let $M \in \mathbb{R}^{n_x \times n_x}$ be symmetric positive definite, $B \in \mathbb{R}^{n_\mu \times n_x}$ have full row rank, and $C \in \mathbb{R}^{n_\mu \times n_\mu}$ be invertible. Then the DAE (8.3) is of index one.*

Proof. We consider the matrix $\mathbf{M}_1 \in \mathbb{R}^{(2n_x+4n_\mu)\times(2n_x+4n_\mu)}$ from the inflated pair of DAE (8.3), see Definition 2.2, and the matrix $\mathbf{P}_1 \in \mathbb{R}^{(n_x+2n_\mu)\times(2n_x+4n_\mu)}$ given by

$$\mathbf{M}_1 = \left[\begin{array}{ccc|ccc} M & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ B & 0 & 0 & 0 & 0 & 0 \\ \hline A & -B^T & -B^T & M & 0 & 0 \\ B & 0 & C & 0 & 0 & 0 \\ 0 & 0 & 0 & B & 0 & 0 \end{array}\right], \quad \mathbf{P}_1 = \left[\begin{array}{ccc|ccc} I_{n_x} & 0 & 0 & 0 & 0 & 0 \\ BM^{-1}AM^{-1} & 0 & 0 & -BM^{-1} & 0 & I_{n_\mu} \\ 0 & 0 & -I_{n_\mu} & 0 & I_{n_\mu} & 0 \end{array}\right].$$

The matrix $\mathbf{P}_1$ has full row rank. Therefore, it can be extended to an invertible matrix $\widetilde{\mathbf{R}}_1 = [\mathbf{P}_1^T \,|\, *]^T$ of the same size as $\mathbf{M}_1$ with a non-specified part $*$. Then the equality

$$\widetilde{\mathbf{R}}_1\mathbf{M}_1 = \left[\begin{array}{ccc|c} M & 0 & 0 & 0 \\ 0 & BM^{-1}B^T & BM^{-1}B^T & 0 \\ 0 & 0 & C & 0 \\ \hline * & * & * & * \end{array}\right] \in \mathbb{R}^{(2n_x+4n_\mu)\times(2n_x+4n_\mu)}$$

holds. By the assumptions on M, B, and C the top left $(n_x + 2n_\mu) \times (n_x + 2n_\mu)$-block is invertible. Therefore, $\widetilde{\mathbf{R}}_1\mathbf{M}_1$ is 1-full and so is $\mathbf{M}_1$. This shows the assertion by Definition 2.2. □

Lemma 8.3 ([AltZ18b, Lem 3.3])**.** *Suppose that the assumptions of Lemma 8.2 are satisfied. For consistent initial data $x_0 \in \mathbb{R}^{n_x}$, i.e., $Bx_0 = g(0)$, the DAEs (8.2) and (8.3) are equivalent in the following sense. A solution pair (x, μ) of system (8.2) implies the solution $(x, \mu, 0)$ of (8.3). On the other hand, a solution (x, μ, γ) of the DAE (8.3) satisfies $\gamma = 0$ and (x, μ) solves (8.2).*

Remark 8.4. An alternative strategy to reduce the index is the singular perturbation approach of replacing the constraint (8.2b) by $Bx + \varepsilon C\mu = g$. The parameter $\varepsilon > 0$ is assumed to be small and the matrix $C \in \mathbb{R}^{n_\mu \times n_\mu}$ is invertible, e.g., the identity matrix or the discretized version of $(\cdot,\cdot)_\mathcal{Q}$. This is known as *penalty method* or *pressure penalization* in the field of fluid dynamics [HeiV95; She95]. The difference between the original solution x and the solution of the regularized system is in the range of $\sqrt{\varepsilon}$ [She95, Th. 3.1]. For the solution μ, however, the error, which is mainly located close to the boundary $\partial\Omega$ [Ran93, p. 207], is for $t\mu(t)$ and smooth data of order $\mathcal{O}(\varepsilon)$ and divergent in general [She95, Lem. 3.1 & Th. 3.1]. Furthermore, for a small parameter ε the temporally discretized system is ill-conditioned [BenH15, p. 13 ff.] and a wise choice of ε depends strongly on the used temporal (and spatial) discretization scheme [She95, Rem. 5.2].

8.1.2. Infinite-Dimensional Case

The regularization procedure from the previous subsection motivates to apply the same ideas also to the operator DAE (8.2). This leads to an extended system of the form: find $u\colon [0,T] \to \mathcal{V}$, $\lambda\colon [0,T] \to \mathcal{Q}$, and $\gamma\colon [0,T] \to \mathcal{Q}$ such that for almost every time-point $t \in [0,T]$ the system

$$\dot{u}(t) + \mathcal{A}u(t) - \mathcal{B}^*\lambda(t) - \mathcal{B}^*\gamma(t) = f(t) \quad \text{in } \mathcal{V}^*, \tag{8.4a}$$
$$\mathcal{B}u(t) + \mathcal{C}\gamma(t) = g(t) \quad \text{in } \mathcal{Q}^*, \tag{8.4b}$$
$$\tfrac{\mathrm{d}}{\mathrm{d}t}(\mathcal{B}u(t)) = \dot{g}(t) \quad \text{in } \mathcal{Q}^*, \tag{8.4c}$$

and the consistent initial condition $u(0) = u_0$ hold. The right-hand sides still satisfy Assumption 8.1.ii), whereas the linear operator $\mathcal{C}\colon \mathcal{Q} \to \mathcal{Q}^*$ is assumed to be elliptic and bounded.

From the construction of the operator DAE (8.4) and the results of the previous subsection, we already know that a spatial discretization leads to a DAE of the form (8.3) and thus, is of index 1. It remains to show the equivalence of the original and extended operator DAE.

Lemma 8.5 (Equivalence of Operator DAEs; [AltZ18b, Lem 3.6])**.** *Let Assumption 8.1 be satisfied and let $\mathcal{C} \in \mathscr{L}(\mathcal{Q}, \mathcal{Q}^*)$ be elliptic. Then the operator DAEs (8.1) and (8.4) are equivalent in the following sense. Every solution (u, λ) of (8.1) implies a solution $(u, \lambda, 0)$ of the operator DAE (8.4). On the other hand, if (u, λ, γ) solves the extended system, then $\gamma \equiv 0$ and (u, λ) is a solution of system (8.1).*

Remark 8.6. By Lemma 9.22 there exists an intermediate Hilbert space $\mathcal{H}_{\text{ker}} \oplus \mathcal{V}_{\text{c}}$ between $\mathcal{V}$ and $\mathcal{H}$. The operator $\mathcal{B}$ can be extended to this space with the new kernel $\mathcal{H}_{\text{ker}}$. Therefore, we define $\mathcal{B}h := 0$ for $h \in \mathcal{H}_{\text{ker}}$ with $\mathcal{B}$ read as its operator extension. For more details see Subsection 9.3.2.2.

Remark 8.7. By Remark 8.6 a consistent initial value $u_0 = \mathcal{H}_{\text{ker}} + \mathcal{B}^- g(0)$ implies $\gamma_0 := \gamma(0) = 0$.

Finally, we make some comments on an alternative regularization approach.

Remark 8.8. A regularization approach of the operator DAE (8.1) based on dummy variables is introduced in [AltH15]. For this, the variable u is decomposed similarly as in Part B into a function u_{ker} with image in $\mathcal{V}_{\text{ker}}$ and u_{c} with range in a complementary space $\mathcal{V}_{\text{c}}$ of $\mathcal{V}_{\text{ker}}$. By introducing the new variable $w_{\text{c}} := \dot{u}_{\text{c}} \in L^2(0, T; \mathcal{V}_{\text{c}})$ we get the regularized operator DAE

$$\begin{aligned} \dot{u}_{\text{ker}}(t) + w_{\text{c}}(t) + \mathcal{A}u_{\text{ker}}(t) + \mathcal{A}u_{\text{c}}(t) - \mathcal{B}^*\lambda(t) &= f(t) \quad \text{in } \mathcal{V}^*, && (8.5a)\\ \mathcal{B}u_{\text{c}}(t) &= g(t) \quad \text{in } \mathcal{Q}^*, && (8.5b)\\ \mathcal{B}w_{\text{c}}(t) &= \dot{g}(t) \quad \text{in } \mathcal{Q}^*. && (8.5c) \end{aligned}$$

The results of the whole chapter can be easily extended to this operator DAE.

Remark 8.9. In contrast to the regularization (8.5), the system (8.4) preserves the saddle point structure of the original operator DAE (8.1). The temporal discretization by algebraically stable Runge-Kutta methods can be modified such that for every time step a stationary saddle point problem has to be solved; see system (8.6) and Remark 8.24. Therefore, efficient solution algorithms for (generalized) saddle point problems are applicable; see [BanWY90; BenGL05] and the references therein. Furthermore, the spatial discretization of the operator DAE (8.5) needs a splitting of the finite dimensional subspace V_h of $\mathcal{V}$ with $V_h = V_{h,1} \oplus V_{h,2}$ such that $\dim V_{h,1} = n_x - n_\mu$, $\dim V_{h,2} = n_\mu$, and the columns of matrix $B \in \mathbb{R}^{n_\mu \times n_x}$ corresponding to $V_{h,2}$ compose an invertible matrix [AltH15, Ch. 3]. Such a splitting is not needed for the operator DAE (8.4), since the operator DAE does not rely on a splitting of $\mathcal{V}$.

8.2. Implicit Euler Method

As a first step towards the convergence of Runge-Kutta schemes, we prove in this section the convergence of the implicit Euler method. For this, we follow the steps of Subsections 7.1.1 and 7.2.2. We show first that the semi-discrete system has a unique solution for every time step. With these approximations, we construct global approximations of the solution of system (8.4) on $[0, T]$ and investigate their convergence behavior.

8.2.1. Temporal Discretization

We formally apply the implicit Euler scheme to the operator DAE (8.4) as in Subsection 7.1.1. We consider a uniform partition of the interval $[0, T]$ with step size $\tau = T/N$, $N \in \mathbb{N}$. The stationary system, which has to be solved for each time step $t_n = \tau n$, $n = 1, \ldots, N$, is given by

$$\begin{aligned} D_\tau u_n + \mathcal{A}u_n - \mathcal{B}^*\lambda_n - \mathcal{B}^*\gamma_n &= f_n \quad \text{in } \mathcal{V}^*, && (8.6a)\\ \mathcal{B}u_n + \mathcal{C}\gamma_n &= g_n \quad \text{in } \mathcal{Q}^*, && (8.6b)\\ \mathcal{B}D_\tau u_n &= \dot{g}_n \quad \text{in } \mathcal{Q}^*. && (8.6c) \end{aligned}$$

Here, D_τ denotes again the discrete derivative, defined by $D_\tau u_n := (u_n - u_{n-1})/\tau$. For $n = 1$, equation (8.6c) includes the term $\mathcal{B}u_0$. Assuming u_0 to be consistent, we have $\mathcal{B}u_0 = g(0)$ in the sense of Remark 8.6. Note that system (8.6) gives an implicit formula for u_n, λ_n, and γ_n in terms of a given approximation u_{n-1}. There is no dependence on previous approximations of λ and γ.

Since the right-hand sides are assumed to be Bochner functions by Assumption 8.1.ii), function evaluations are typically not defined. Thus, only for $g \in W^{1,1}(0,T;\mathcal{Q}) \hookrightarrow C([0,T],\mathcal{Q})$ we may define $g_n := g(t_n)$, whereas we define f_n and $\dot{g}_n$ by their means; cf. (7.9). These approximations are of first order but may be replaced by any other approximation, especially for more regular data f and g. Nevertheless, we require certain convergence properties, which we summarize in the following assumption.

Assumption 8.10. Suppose that $f_n \in \mathcal{V}^*$, $g_n \in \mathcal{Q}^*$, and $\dot{g}_n \in \mathcal{Q}^*$ are given, $n = 1, \dots, N$. Let $f_\tau \colon [0,T] \to \mathcal{V}^*$ denote the piecewise constant function with $f_\tau(t) := f_n$ for $t \in (t_{n-1}, t_n]$ and $f(0) := f_1$. Analogously, we define the piecewise constant functions g_τ and $\dot{g}_\tau$ via g_n and $\dot{g}_n$, respectively. We assume that f_τ, g_τ, and $\dot{g}_\tau$ converge for $\tau \to 0$ in the strong sense, i.e.,

$$f_\tau \to f \text{ in } L^2(0,T;\mathcal{V}^*) + L^1(0,T;\mathcal{H}^*), \qquad g_\tau \to g \text{ in } L^\infty(0,T;\mathcal{Q}^*), \qquad \dot{g}_\tau \to \dot{g} \text{ in } L^1(0,T;\mathcal{Q}^*).$$

Remark 8.11. We emphasize that $\dot{g}_\tau$ is not the derivative of g_τ in the notation of Assumption 8.10.

The discretization of the right-hand sides above from this page fulfills Assumption 8.10; cf. Lemma 3.34. For any discretization satisfying Assumption 8.10, system (8.6) is well-defined. It remains to check the solvability of this system.

Lemma 8.12 (Solvability of the Time-Discrete System)**.** *Suppose that Assumption 8.1.i) on the operators is satisfied and that $\mathcal{C} \in \mathscr{L}(\mathcal{Q},\mathcal{Q}^*)$ is elliptic. Let u_{n-1} be an element of $\mathcal{H}_{\text{ker}} + \mathcal{V}_{\text{c}}$ such that the operator $\mathcal{B}$ is applicable, $n = 1, \dots, N$. Assume that the right-hand sides of (8.6) satisfy $f_n \in \mathcal{V}^*$, $g_n \in \mathcal{Q}^*$, and $\dot{g}_n \in \mathcal{Q}^*$. Then system (8.6) has a unique solution $(u_n, \lambda_n, \gamma_n) \in \mathcal{V} \times \mathcal{Q} \times \mathcal{Q}$.*

Proof. By (8.6b) and (8.6c) the equality $\mathcal{C}\gamma_n = g_n - \tau\dot{g}_n - \mathcal{B}u_{n-1}$ holds. This equation has a unique solution γ_n by the Lax-Milgram Theorem 3.4. Therefore, we consider the system given by the equations (8.6a) and (8.6c). This reduced problem and thus also system (8.6) have a unique solution by Theorem 3.8. □

8.2.2. Convergence Results

Due to Lemma 8.12, system (8.6) provides the discrete approximations u_n, λ_n, and γ_n at any time point t_n for a given consistent initial value u_0. With these, we define global approximations of the weak solution u on the interval $[0,T]$ similar to (7.18). More precisely, we define $u_\tau, \widehat{u}_\tau \colon [0,T] \to \mathcal{H}_{\text{ker}} + \mathcal{V}_{\text{c}}$ by

$$u_\tau(t) := \begin{cases} u_0, & \text{if } t = 0 \\ u_n, & \text{if } t \in (t_{n-1}, t_n] \end{cases}, \qquad \widehat{u}_\tau(t) := \begin{cases} u_0, & \text{if } t = 0 \\ u_n + D_\tau u_n (t - t_n), & \text{if } t \in (t_{n-1}, t_n] \end{cases}. \tag{8.7}$$

Analogously, we define piecewise constant approximations of the Lagrange multipliers λ and γ, which we denote by λ_τ and γ_τ, respectively. As starting value we set $\gamma_\tau(0) := \gamma_0$ and $\lambda_\tau(0)$ arbitrarily. By $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$ we denote the generalized time derivative of $\widehat{u}_\tau$, which is piecewise constant with values $D_\tau u_n$. By the global approximations, the temporally discretized system (8.6) can be reformulated as

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau + \mathcal{A}u_\tau - \mathcal{B}^*\lambda_\tau - \mathcal{B}^*\gamma_\tau = f_\tau \quad \text{in } \mathcal{V}^*, \tag{8.8a}$$

$$\mathcal{B}u_\tau + \mathcal{C}\gamma_\tau = g_\tau \quad \text{in } \mathcal{Q}^*, \tag{8.8b}$$

$$\mathcal{B}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau\big) = \dot{g}_\tau \quad \text{in } \mathcal{Q}^*. \tag{8.8c}$$

We assume that the discrete right-hand sides f_τ, g_τ, and $\dot g_\tau$ satisfy Assumption 8.10. For the study of the piecewise approximations, we state a discrete version of Gronwall's lemma.

Lemma 8.13. *Let $a, x_1, b_1, x_2, b_2, \ldots \in \mathbb{R}_{\geq 0}$ be given with $x_n^2 \leq a + \sum_{i=1}^n b_i x_i$. Then the inequality $x_n \leq \sqrt{a} + \sum_{i=1}^n b_i$ holds for all $n = 1, 2, \ldots$*

Proof. By the assumptions on x_n, the non-negative root of $y \mapsto y^2 - b_n y - a - \sum_{i=1}^{n-1} b_i x_i$ is an upper bound of x_n, i.e.,

$$x_n \leq \frac{b_n}{2} + \Big(\frac{b_n^2}{4} + a + \sum_{i=1}^{n-1} b_i x_i\Big)^{1/2}. \tag{8.9}$$

The proof of this lemma is inductive. The base case $x_1 \leq \frac{b_1}{2} + \big(\frac{b_1^2}{4} + a\big)^{1/2} \leq \sqrt{a} + b_1$ holds by inequality (8.9) and [Emm04, Cor. A.1.2], where the induction step is given by

$$\begin{aligned} x_n &\overset{(8.9)}{\leq} \frac{b_n}{2} + \Big(\frac{b_n^2}{4} + a + \sum_{i=1}^{n-1} b_i x_i\Big)^{1/2} \\ &\overset{\text{IH}}{\leq} \frac{b_n}{2} + \Big(\frac{b_n^2}{4} + a + \sum_{i=1}^{n-1} b_i \Big(\sum_{j=1}^{i} b_j + \sqrt{a}\Big)\Big)^{1/2} \\ &\leq \frac{b_n}{2} + \Big(\Big(\frac{b_n}{2} + \sum_{i=1}^{n-1} b_i\Big)^2 + 2\Big(\frac{b_n}{2} + \sum_{i=1}^{n-1} b_i\Big)\sqrt{a} + a\Big)^{1/2} = \sum_{i=1}^{n} b_i + \sqrt{a}. \end{aligned}$$

□

We analyze the convergence behavior of the introduced approximations in the following theorem.

Theorem 8.14 (Convergence of the Implicit Euler Scheme)**.** *Let Assumption 8.1 be satisfied and $\mathcal{C} \in \mathscr{L}(\mathcal{Q}, \mathcal{Q}^*)$ be elliptic. Suppose that $(u, \lambda, 0)$ is the solution of the operator DAE (8.4). If the approximations of the right-hand sides f_τ, g_τ, and $\dot g_\tau$ fulfill Assumption 8.10, then*

$$\begin{aligned} u_\tau &\to u \quad \text{in } L^2(0,T;\mathcal{V}), & \widehat{u}_\tau &\to u \quad \text{in } L^2(0,T;\mathcal{H}), \\ \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau &\to \dot u \quad \text{in } L^2(0,T;\mathcal{V}^*_{\ker}) + L^1(0,T;\mathcal{H}^*_{\ker}), & \gamma_\tau &\to 0 \quad \text{in } L^\infty(0,T;\mathcal{Q}) \end{aligned}$$

as $\tau \to 0$. Furthermore, the primitive of λ_τ, namely $\Lambda_\tau(t) := \int_0^t \lambda_\tau(s)\,\mathrm{d}s$, converges strongly to Λ in $L^2(0,T;\mathcal{Q})$ with Λ defined as in Theorem 6.7.

Proof. In the first step we show the convergence of the Lagrange multiplier γ_τ. With this, we show the weak and afterwards even the strong convergence of u_τ and the derivative of $\widehat{u}_\tau$. Finally, we prove the assertions for $\widehat{u}_\tau$ and λ_τ.

Step 1 (Convergence of γ_τ): With the consistent initial value u_0, equation (8.6b), and a successive application of equation (8.6c), we obtain

$$\mathcal{C}\gamma_n = g_n - g(0) - \tau\sum_{i=1}^n \dot g_i = \int_0^{t_n} \dot g(t)\,\mathrm{d}t - \sum_{i=1}^n \tau \dot g_i + g_n - g(t_n) = \int_0^{t_n} \dot g(t) - \dot g_\tau(t)\,\mathrm{d}t + g_n - g(t_n). \tag{8.10}$$

Since $\mathcal{C}$ is elliptic and bounded, using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|\gamma_\tau\|_{L^\infty(0,T;\mathcal{Q})} = \max_{n=1,\ldots,N} \|\gamma_n\|_{\mathcal{Q}} &\overset{(8.10)}{\lesssim} \max_{n=1,\ldots,N} \Big\| \int_0^{t_n} \dot g(t) - \dot g_\tau(t)\,\mathrm{d}t + g_n - g(t_n)\Big\|_{\mathcal{Q}^*} \\ &\leq \|\dot g - \dot g_\tau\|_{L^1(0,T;\mathcal{Q}^*)} + \|g - g_\tau\|_{L^\infty(0,T;\mathcal{Q}^*)}. \end{aligned} \tag{8.11}$$

Thus, by Assumption 8.10 it follows that $\|\gamma_\tau\|_{L^\infty(0,T;\mathcal{Q})} \to 0$.

Step 2 (Weak convergence of u_τ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$): We use the splitting $\mathcal{V} = \mathcal{V}_{\mathrm{ker}} \oplus \mathcal{V}_{\mathrm{c}}$ with $\mathcal{V}_{\mathrm{c}} = \{v \in \mathcal{V} \,|\, \mathcal{A}v \in \mathcal{V}^0_{\mathrm{ker}}\}$ as discussed in Section 3.1. We decompose u_n and $D_\tau u_n$ for $n = 1, \dots, N$ as well as the initial value $u_0 = u_{\mathrm{ker},0} + u_{\mathrm{c},0}$ with $u_{\mathrm{ker},0} \in \mathcal{H}_{\mathrm{ker}}$ and $u_{\mathrm{c},0} = \mathcal{B}^-_{\mathcal{A}} g(0) \in \mathcal{V}_{\mathrm{c}}$. We also split the global approximations of u into

$$u_\tau = u_{\mathrm{ker},\tau} + u_{\mathrm{c},\tau}, \qquad \widehat{u}_\tau = \widehat{u}_{\mathrm{ker},\tau} + \widehat{u}_{\mathrm{c},\tau}, \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau = \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau} + \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{c},\tau}. \tag{8.12}$$

The exact solution u is decomposed into $u_{\mathrm{ker}}\colon [0,T] \to \mathcal{V}_{\mathrm{ker}}$ and $u_{\mathrm{c}}\colon [0,T] \to \mathcal{V}_{\mathrm{c}}$. Equation (8.8), Assumption 8.10, and the convergence of γ_τ imply

$$u_{\mathrm{c},\tau} = \mathcal{B}^-_{\mathcal{A}}\mathcal{B}u_\tau = \mathcal{B}^-_{\mathcal{A}}(g_\tau - \mathcal{C}\gamma_\tau) \to \mathcal{B}^-_{\mathcal{A}} g = u_{\mathrm{c}} \qquad \text{in } L^\infty(0,T;\mathcal{V}_{\mathrm{c}}), \tag{8.13a}$$

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{c},\tau} = \mathcal{B}^-_{\mathcal{A}}\mathcal{B}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau\big) = \mathcal{B}^-_{\mathcal{A}}\dot{g}_\tau \to \mathcal{B}^-_{\mathcal{A}}\dot{g} = \dot{u}_{\mathrm{c}} \qquad \text{in } L^1(0,T;\mathcal{V}_{\mathrm{c}}). \tag{8.13b}$$

The linearity of the discrete derivative yields $(D_\tau u_n)_{\mathrm{ker}} = D_\tau u_{\mathrm{ker},n}$. Thus, we can rewrite equation (8.6a) as

$$D_\tau u_{\mathrm{ker},n} + \mathcal{A}u_{\mathrm{ker},n} - \mathcal{B}^*\lambda_n - \mathcal{B}^*\gamma_n = f_n - D_\tau u_{\mathrm{c},n} - \mathcal{A}u_{\mathrm{c},n} \quad \text{in } \mathcal{V}^*. \tag{8.14}$$

Since $D_\tau u_{\mathrm{ker},n} \in \mathcal{V}_{\mathrm{ker}} = \ker \mathcal{B}$ for $n > 1$ and $D_\tau u_{\mathrm{ker},1} \in \mathcal{H}_{\mathrm{ker}}$, we conclude with $u_{\mathrm{c},n} \in \mathcal{V}_{\mathrm{c}}$ that $u_{\mathrm{ker},n}$ is fully determined by

$$D_\tau u_{\mathrm{ker},n} + \mathcal{A}u_{\mathrm{ker},n} = f_n - D_\tau u_{\mathrm{c},n} \quad \text{in } \mathcal{V}^*_{\mathrm{ker}}, \tag{8.15}$$

where $\mathcal{A}u_{\mathrm{c},n}$ vanishes by the definition of $\mathcal{V}_{\mathrm{c}}$. Note that equation (8.15) may also be written in its continuous form

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau} + \mathcal{A}u_{\mathrm{ker},\tau} = f_\tau - \mathcal{B}^-_{\mathcal{A}}\dot{g}_\tau \quad \text{in } \mathcal{V}^*_{\mathrm{ker}}. \tag{8.16}$$

By the assumption on the right-hand side f there exists $f^{[1]} \in L^2(0,T;\mathcal{V}^*)$ and $f^{[2]} \in L^1(0,T;\mathcal{H}^*)$ with $f = f^{[1]} + f^{[2]}$. Testing (8.15) with $u_{\mathrm{ker},n}$ and summing over $i = 1, \dots, n$ leads to

$$\begin{aligned} \|u_{\mathrm{ker},n}\|^2_{\mathcal{H}} &+ \sum_{i=1}^n \|u_{\mathrm{ker},i} - u_{\mathrm{ker},i-1}\|^2_{\mathcal{H}} + \tau\mu_{\mathcal{A}} \sum_{i=1}^n \|u_{\mathrm{ker},i}\|^2_{\mathcal{V}} \\ &\le \|u_{\mathrm{ker},0}\|^2_{\mathcal{H}} + \frac{\tau}{\mu_{\mathcal{A}}} \sum_{i=1}^n \big\|f_i^{[1]}\big\|^2_{\mathcal{V}^*} + 2\tau \sum_{i=1}^n \big\|f_i^{[2]} - \mathcal{B}^-_{\mathcal{A}}\dot{g}_i\big\|_{\mathcal{H}^*} \|u_{\mathrm{ker},n}\|_{\mathcal{H}} \end{aligned} \tag{8.17}$$

cf. Theorem 7.9. Then Lemma 8.13 with $x_i := \|u_{\mathrm{ker},i}\|_{\mathcal{H}}$, $a := \|u_{\mathrm{ker},0}\|^2_{\mathcal{H}} + \frac{\tau}{\mu_{\mathcal{A}}} \sum_{k=1}^n \|f_k^{[1]}\|^2_{\mathcal{V}^*}$, and $b_i := 2\tau\big\|f_i^{[2]} - \mathcal{B}^-_{\mathcal{A}}\dot{g}_i\big\|_{\mathcal{H}^*}$, $i = 1, \dots, n$, implies

$$\|u_{\mathrm{ker},n}\|_{\mathcal{H}} \le \bigg(\|u_{\mathrm{ker},0}\|^2_{\mathcal{H}} + \frac{\tau}{\mu_{\mathcal{A}}} \sum_{i=1}^n \big\|f_i^{[1]}\big\|^2_{\mathcal{V}^*}\bigg)^{1/2} + 2\tau \sum_{i=1}^n \big\|f_i^{[2]} - \mathcal{B}^-_{\mathcal{A}}\dot{g}_i\big\|_{\mathcal{H}^*}. \tag{8.18}$$

By a combination of the inequalities (8.17) and (8.18) the estimate

$$\begin{aligned} &\|u_{\mathrm{ker},n}\|^2_{\mathcal{H}} + \sum_{i=1}^n \|u_{\mathrm{ker},i} - u_{\mathrm{ker},i-1}\|^2_{\mathcal{H}} + \tau\mu_{\mathcal{A}} \sum_{i=1}^n \|u_{\mathrm{ker},i}\|^2_{\mathcal{V}} \\ &\le \bigg[\bigg(\|u_{\mathrm{ker},0}\|^2_{\mathcal{H}} + \frac{\tau}{\mu_{\mathcal{A}}} \sum_{i=1}^n \big\|f_i^{[1]}\big\|^2_{\mathcal{V}^*}\bigg)^{1/2} + 2\tau \sum_{i=1}^n \big\|f_i^{[2]} - \mathcal{B}^-_{\mathcal{A}}\dot{g}_i\big\|_{\mathcal{H}^*}\bigg]^2 \end{aligned}$$

$$= \Big[\Big(\|u_{\text{ker},0}\|_{\mathcal{H}}^2 + \frac{1}{\mu_{\mathcal{A}}} \int_0^T \|f_\tau^{[1]}\|_{\mathcal{V}^*}^2 \,\mathrm{d}t \Big)^{1/2} + 2 \int_0^T \|f_\tau^{[2]} - \mathcal{B}_{\mathcal{A}}^- \dot{g}_\tau\|_{\mathcal{H}^*} \,\mathrm{d}t \Big]^2 \tag{8.19}$$

holds. Note that the right-hand side of (8.19) is bounded independently of τ by Assumption 8.10. By the arguments of Lemma 7.11, there exists a subsequence τ' of τ and a function $w \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$ such that w is the weak limit of $u_{\text{ker},\tau'}$ in $L^2(0,T;\mathcal{V})$ and the weak* limit of $u_{\text{ker},\tau'}$ and $\widehat{u}_{\text{ker},\tau'}$ in $L^\infty(0,T;\mathcal{H})$ as $\tau' \to 0$. The convergence of $u_{\text{ker},\tau'}$, $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{c},\tau}$, and f_τ then implies

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau'} \overset{(8.16)}{=} f_{\tau'} - \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{c},\tau'} - \mathcal{A}u_{\text{ker},\tau'} \rightharpoonup f - \dot{u}_{\text{c}} - \mathcal{A}w =: w_{\text{d}} \quad \text{in } L^2(0,T;\mathcal{V}_{\text{ker}}^*) + L^1(0,T;\mathcal{H}_{\text{ker}}^*).$$

Following the lines of Lemma 7.12 and Theorem 7.14 one shows $w_{\text{d}} = \dot{w}$ and that w is a solution of the operator DAE

$$\dot{u}_{\text{ker}} + \mathcal{A}u_{\text{ker}} = f - \dot{u}_{\text{c}} \quad \text{in } \mathcal{V}_{\text{ker}}^* \tag{8.20}$$

with initial value $w(0) = u_{\text{ker},0}$. Since the solution is unique by Theorem 4.22, the identity $w = u_{\text{ker}}$ holds and the piecewise functions converge weakly/weakly* for the whole sequence τ [GajGZ74, Ch. 1, Lem. 5.4]. In combination with (8.13) this shows the weak (respectively weak*) convergence of u_τ in $L^2(0,T;\mathcal{V})$ and of $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$ in $L^2(0,T;\mathcal{V}_{\text{ker}}^*) + L^1(0,T;\mathcal{H}_{\text{ker}}^*)$.

For the following analysis of the strong convergence we note that $u_{\text{ker},N}$ is bounded in $\mathcal{H}$ by (8.19). Its weak limit is given by $u_{\text{ker}}(T)$ in $\mathcal{H}$ as $N \to \infty$ or $\tau \to 0$, respectively, since for every $v_{\text{ker}} \in \mathcal{V}_{\text{ker}}$ we have

$$\begin{aligned} \lim_{N\to\infty} (u_{\text{ker},N}, Tv_{\text{ker}})_{\mathcal{H}} &\overset{(3.21)}{=} \lim_{\tau\to 0} \int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}, tv_{\text{ker}} \rangle + (\widehat{u}_{\text{ker},\tau}, v_{\text{ker}})_{\mathcal{H}} \,\mathrm{d}t \\ &= \int_0^T \langle \dot{u}_{\text{ker}}, tv_{\text{ker}} \rangle + (u_{\text{ker}}, v_{\text{ker}})_{\mathcal{H}} \,\mathrm{d}t \overset{(3.21)}{=} (u_{\text{ker}}(T), Tv_{\text{ker}})_{\mathcal{H}}. \end{aligned}$$

Step 3 (Strong convergence of u_τ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$): It remains to prove that the sequences $u_{\text{ker},\tau}$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}$ converge strongly. For this, we note that equation (8.16) leads to the estimate

$$\begin{aligned} \|u_{\text{ker},\tau} - u_{\text{ker}}\|_{L^2(0,T;\mathcal{V})}^2 &\lesssim \int_0^T \langle \mathcal{A}u_{\text{ker},\tau} - \mathcal{A}u_{\text{ker}}, u_{\text{ker},\tau} - u_{\text{ker}} \rangle \,\mathrm{d}t \\ &= -\int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau},\; u_{\text{ker},\tau} - u_{\text{ker}} \rangle \,\mathrm{d}t + \int_0^T \langle \dot{u}_{\text{ker}},\; u_{\text{ker},\tau} - u_{\text{ker}} \rangle \,\mathrm{d}t \\ &\quad + \int_0^T \langle f_\tau - f + \mathcal{B}_{\mathcal{A}}^-(\dot{g}_\tau - \dot{g}),\; u_{\text{ker},\tau} - u_{\text{ker}} \rangle \,\mathrm{d}t. \end{aligned} \tag{8.21}$$

The second integral on the right-hand side of (8.21) converges to zero because of the weak and weak* convergence of $u_{\text{ker},\tau}$ to u_{ker} in $L^2(0,T;\mathcal{V})$ and $L^\infty(0,T;\mathcal{H})$, respectively, and the third integral because of the assumption on the right-hand sides and the boundedness of $u_{\text{ker},\tau}$ and u_{ker} in $L^2(0,T;\mathcal{V}_{\text{ker}}) \cap L^\infty(0,T;\mathcal{H}_{\text{ker}})$; cf. (8.21) and Theorem 4.22. For the first integral we use

$$\int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}, u_{\text{ker},\tau} \rangle \,\mathrm{d}t = \sum_{n=1}^N \langle u_{\text{ker},n} - u_{\text{ker},n-1}, u_{\text{ker},n} \rangle \overset{(7.17)}{\geq} \frac{1}{2}\|u_{\text{ker},N}\|_{\mathcal{H}}^2 - \frac{1}{2}\|u_{\text{ker},0}\|_{\mathcal{H}}^2.$$

Then, by the weak convergence of $u_{\text{ker},N}$ and [Alt16, Rem. 8.3.4] the limit

$$\liminf_{\tau\to 0} \int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}, u_{\text{ker},\tau} \rangle \,\mathrm{d}t \geq \frac{1}{2}\|u_{\text{ker}}(T)\|_{\mathcal{H}}^2 - \frac{1}{2}\|u_{\text{ker}}(0)\|_{\mathcal{H}}^2 \overset{(3.21)}{=} \int_0^T \langle \dot{u}_{\text{ker}}, u_{\text{ker}} \rangle \,\mathrm{d}t \tag{8.22}$$

holds. With this inequality and the weak convergence of $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau}$, the estimate (8.21) implies

$$\begin{aligned} 0 &\leq \limsup_{\tau\to 0} \|u_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\|^2_{L^2(0,T;\mathcal{V})} \\ &\overset{(8.21)}{\lesssim} \limsup_{\tau\to 0} \int_0^T \big\langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau},\, u_{\mathrm{ker}} \big\rangle \,\mathrm{d}t - \liminf_{\tau\to 0} \int_0^T \big\langle \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau},\, u_{\mathrm{ker},\tau} \big\rangle \,\mathrm{d}t \\ &\leq \int_0^T \big\langle \dot{u}_{\mathrm{ker}},\, u_{\mathrm{ker}} \big\rangle \,\mathrm{d}t - \int_0^T \big\langle \dot{u}_{\mathrm{ker}},\, u_{\mathrm{ker}} \big\rangle \,\mathrm{d}t = 0. \end{aligned}$$

This shows the strong convergence $u_{\mathrm{ker},\tau} \to u_{\mathrm{ker}}$ as well as

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau} = f_\tau - \mathcal{B}^-_{\mathcal{A}}\dot{g}_\tau - \mathcal{A}u_{\mathrm{ker},\tau} \;\to\; f - \mathcal{B}^-_{\mathcal{A}}\dot{g} - \mathcal{A}u_{\mathrm{ker}} = \dot{u}_{\mathrm{ker}} \quad \text{in } L^2(0,T;\mathcal{V}^*_{\mathrm{ker}}) + L^1(0,T;\mathcal{H}^*_{\mathrm{ker}}).$$

By the triangle inequality we obtain the claimed convergence of $u_\tau = u_{\mathrm{ker},\tau} + u_{\mathrm{c},\tau}$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau = \frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau} + \frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{c},\tau}$.

Step 4 (Convergence of $\widehat{u}_\tau$): We observe that by Theorem 3.36 and $\widehat{u}_{\mathrm{c},\tau}(0) = u_{\mathrm{c}}(0)$ the limit

$$\lim_{\tau\to 0} \|\widehat{u}_{\mathrm{c},\tau} - u_{\mathrm{c}}\|_{L^\infty(0,T;\mathcal{V})} \leq \lim_{\tau\to 0} \int_0^T \big\| \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{c},\tau} - \dot{u}_{\mathrm{c}} \big\|_{\mathcal{V}} \,\mathrm{d}t \overset{(8.13b)}{=} 0 \tag{8.23}$$

holds. For the convergence of $\widehat{u}_{\mathrm{ker},\tau}$ we note that

$$\lim_{\tau\to 0} \|\widehat{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker},\tau}\|^2_{L^2(0,T;\mathcal{H})} = \lim_{\tau\to 0} \frac{\tau}{3} \sum_{n=1}^N \|u_{\mathrm{ker},n} - u_{\mathrm{ker},n-1}\|^2_{\mathcal{H}} \leq \lim_{\tau\to 0} \frac{\tau}{3} M^2(u_{\mathrm{ker},0}, f_\tau, \dot{g}_\tau) = 0,$$

where $M^2(u_{\mathrm{ker},0}, f_\tau, \dot{g}_\tau) \geq 0$ is right-hand side of (8.19), which can be bounded independently of τ. Thus, $\widehat{u}_{\mathrm{ker},\tau}$ and $u_{\mathrm{ker},\tau}$ have the same limit u_{ker} in $L^2(0,T;\mathcal{H})$, which implies the strong convergence $\widehat{u}_\tau \to u$ in $L^2(0,T;\mathcal{H})$.

Step 5 (Convergence of λ_τ): Let Λ_τ, U_τ, Γ_τ, and F_τ denote the primitives of λ_τ, u_τ, γ_τ, and f_τ, respectively, which vanish at $t = 0$. An integration of equation (8.8a) then leads to

$$\mathcal{B}^*\Lambda_\tau = \widehat{u}_\tau + \mathcal{A}U_\tau - \mathcal{B}^*\Gamma_\tau - F_\tau - u_0 \qquad \text{in } AC([0,T],\mathcal{V}^*), \tag{8.24}$$

where $AC([0,T],\mathcal{V}^*)$ denotes the space of absolutely continuous functions with values in $\mathcal{V}^*$. The inf-sup condition of $\mathcal{B}$ implies

$$\beta\|\Lambda_\tau(t)\|_{\mathcal{Q}} \leq \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle \mathcal{B}v, \Lambda_\tau(t)\rangle}{\|v\|_{\mathcal{V}}} \overset{(8.24)}{\lesssim} \|u_0\|_{\mathcal{H}} + \|\widehat{u}_\tau(t)\|_{\mathcal{H}} + \|U_\tau(t)\|_{\mathcal{V}} + \|\Gamma_\tau(t)\|_{\mathcal{Q}} + \|F_\tau(t)\|_{\mathcal{V}^*}$$

and thus, the bound

$$\|\Lambda_\tau\|^2_{L^2(0,T;Q)} \lesssim \|\widehat{u}_\tau\|^2_{L^2(0,T;\mathcal{H})} + \|u_0\|^2_{\mathcal{H}} + \|u_\tau\|^2_{L^2(0,T;\mathcal{V})} + \|\gamma_\tau\|^2_{L^2(0,T;\mathcal{Q})} + \|f_\tau\|^2_{L^2(0,T;\mathcal{V}^*)+L^1(0,T;\mathcal{H}^*)}$$

holds. Inserting $\Lambda_{\tau_1} - \Lambda_{\tau_2}$ instead of Λ_τ for two different time step sizes τ_1, τ_2, we obtain that Λ_τ is a Cauchy sequence in $L^2(0,T;\mathcal{Q})$. Thus, there exists a unique limit $\widetilde{\lambda} \in L^2(0,T;\mathcal{Q})$. Finally, a comparison of (8.1a) and the limit of (8.24) shows $\widetilde{\lambda} = \Lambda$. □

Remark 8.15. Since u_τ and $\widehat{u}_\tau$ converge strongly in $L^2(0,T;\mathcal{H})$ and are bounded in $L^\infty(0,T;\mathcal{H})$ by (8.19), both sequences converge strongly to u in $L^p(0,T;\mathcal{H})$ for every $p \in [1,\infty)$ [Emm04, Rem. 8.1.13]. Analogously, the limit $\Lambda_\tau \to \Lambda$ holds in $L^p(0,T;\mathcal{Q})$ for every $p \in [1,\infty)$.

8.2.3. Convergence Results for More Regular Data

By Theorem 6.7 the Lagrange multiplier λ exists only in a distributional sense under the assumptions of Theorem 8.14. Therefore, we cannot expect better convergence for λ as shown in Theorem 8.14. In Section 6.2, however, we studied conditions such that λ is a Bochner-integrable function. In this subsection, we consider these additional assumptions to prove a convergence result for λ_τ and improve the one for $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$ to the convergence in the more restrictive space $L^2(0,T;\mathcal{H})$ than the one in Theorem 8.14.

Theorem 8.16 (Convergence for More Regular Data)**.** *In addition to the assumptions of Theorem 8.14 suppose that $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}u_0 = g(0)$, and one of the following conditions holds:*

i) The right-hand sides f and g are elements of $L^2(0,T;\mathcal{H}^*)$ and $H^1(0,T;\mathcal{Q}^*)$, respectively. The approximations f_τ and $\dot{g}_\tau$ satisfy Assumption 8.10 in $L^2(0,T;\mathcal{H}^*)$ and $L^2(0,T;\mathcal{Q}^*)$, respectively. Furthermore, there exist an operator $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$, which is self-adjoint and elliptic on $\mathcal{V}_{\text{ker}}$, and an operator $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$ such that $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$.

ii) The right-hand sides satisfy $f \in H^1(0,T;\mathcal{V}^*) + W^{1,1}(0,T;\mathcal{H}^*)$ and $g \in W^{2,1}(0,T;\mathcal{Q}^*)$. The approximations f_τ, g_τ, and $\dot{g}_\tau$ are given by pointwise function evaluations of f, g, and $\dot{g}$, respectively. Furthermore, the compatibility condition $f(0) - \mathcal{A}u_0 \in \mathcal{H}^*_{\text{ker}}$ is fulfilled.

Then the piecewise constant approximations $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$ and λ_τ satisfy

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau \to \dot{u} \quad \text{in } L^2(0,T;\mathcal{H}), \qquad\qquad \lambda_\tau \to \lambda \quad \text{in } L^2(0,T;\mathcal{Q}).$$

Proof. If $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$ strongly converges to $\dot{u}$ in $L^2(0,T;\mathcal{H}) \hookrightarrow L^2(0,T;\mathcal{V}^*)$, then the convergence of λ_τ follows immediately by

$$\lambda_\tau = \mathcal{B}^{-*}_{\text{left}}(-f_\tau + \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau + \mathcal{A}u_\tau - \mathcal{B}^*\gamma_\tau) \to \mathcal{B}^{-*}_{\text{left}}(-f + \dot{u} + \mathcal{A}u) = \lambda \quad \text{in } L^2(0,T;\mathcal{Q}).$$

The convergence of u_τ and γ_τ is shown in Theorem 8.14. Thus, it is enough to prove $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau \to \dot{u}$ in $L^2(0,T;\mathcal{H})$. We therefore split u_n and $D_\tau u_n$ into their components in $\mathcal{V}_{\text{ker}}$ and $\mathcal{V}_{\text{c}}$ as in the proof of Theorem 8.14. In both cases, i.e., i) and ii), one proves $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{c},\tau} \to \dot{u}_{\text{c}}$ in $L^2(0,T;\mathcal{V}_{\text{c}}) \hookrightarrow L^2(0,T;\mathcal{V})$ analogously to equation (8.13b). It remains to prove the convergence of $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}$.

Proof for condition i): Following the lines of Theorem 7.24 shows that $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}$ is bounded in $L^2(0,T;\mathcal{H})$ and has the weak limit $\dot{u}_{\text{ker}}$. In particular, this together with (8.16) and $\mathcal{A}_2 u_{\text{ker},\tau} \to \mathcal{A}_2 u_{\text{ker}}$ in $L^2(0,T;\mathcal{H}^*)$ by the strong convergence of $u_{\text{ker},\tau}$ imply $\mathcal{A}_1 u_{\text{ker},\tau} \rightharpoonup \mathcal{A}_1 u_{\text{ker}}$ in $L^2(0,T;\mathcal{H}^*_{\text{ker}})$.
By the weak convergence of $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau}$ we note that the second and third integral of the right-hand side in

$$\begin{aligned}
\big\|\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau} - \dot{u}_{\text{ker}}\big\|^2_{L^2(0,T;\mathcal{H})} &= \int_0^T \big(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau} - \dot{u}_{\text{ker}}, \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau} - \dot{u}_{\text{ker}}\big)_{\mathcal{H}} \,\mathrm{d}t \\
&= -\int_0^T \langle \mathcal{A}_1 u_{\text{ker},\tau}, \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau} - \dot{u}_{\text{ker}} \rangle \,\mathrm{d}t + \int_0^T \langle \mathcal{A}_1 u_{\text{ker}}, \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau} - \dot{u}_{\text{ker}} \rangle \,\mathrm{d}t \\
&\quad + \int_0^T \langle f_\tau - f + \mathcal{B}^-_{\mathcal{A}}(\dot{g}_\tau - \dot{g}) - \mathcal{A}_2(u_{\text{ker},\tau} - u_{\text{ker}})), \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\text{ker},\tau} - \dot{u}_{\text{ker}} \rangle \,\mathrm{d}t
\end{aligned}$$

vanish as $\tau \to 0$. The first integral is a zero sequence by similar arguments as in the proof of Theorem 8.14 Step 3. Here, we used [Zim15, Th. 3.20] and the weak convergence of $u_{\text{ker},N}$ to $u_{\text{ker}}(T)$ in $\mathcal{V}_{\text{ker}}$, which follows by $u_{\text{ker},N} \rightharpoonup u_{\text{ker}}(T)$ in $\mathcal{H}_{\text{ker}}$ and an estimate similar to (7.39).

Proof for condition ii): We note that $D_\tau u_{\text{ker},n} \in \mathcal{V}_{\text{ker}}$ satisfies the discrete system

$$D_\tau D_\tau u_{\text{ker},n} + \mathcal{A} D_\tau u_{\text{ker},n} = D_\tau f_n - D_\tau \mathcal{B}^-_{\mathcal{A}} \dot{g}_n \qquad \text{in } \mathcal{V}^*_{\text{ker}}, \tag{8.25}$$

cf. equation (8.15). Here, we define $D_\tau u_{\ker,0} := w_{\ker,0}$ with $w_{\ker,0}$ as the representation of $f(0) - \mathcal{A}u_0 - \mathcal{B}_\mathcal{A}^- \dot{g}(0) \in \mathcal{H}_{\ker}^*$ in $\mathcal{H}_{\ker}$. The system (8.25) can be seen as the approximation of the operator ODE (8.20) with right-hand sides $\dot{f} - \mathcal{B}_\mathcal{A}^- \ddot{g}$ and initial value $w_{\ker,0}$. Note that the solution of this operator ODE is $\dot{u}_{\ker}$ by [Wlo87, Th. 27.2]. Furthermore, we have

$$D_\tau f_n = \frac{1}{\tau}\big(f(t_n) - f(t_{n-1})\big) = \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \dot{f}\,\mathrm{d}t.$$

Thus, the piecewise constant function $(D_\tau f)_\tau$ defined as in Assumption 8.10 converges to $\dot{f}$ in $L^2(0,T;\mathcal{V}^*)+L^1(0,T;\mathcal{H}^*)$ by Lemma 3.34. Analogously we have $(D_\tau \dot{g})_\tau \to \ddot{g}$ in $L^1(0,T;\mathcal{Q}^*)$. Finally, $\lim_{\tau\to 0} \frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\ker,\tau} = \dot{u}_{\ker}$ in $L^2(0,T;\mathcal{V}_{\ker}) \hookrightarrow L^2(0,T;\mathcal{V})$ follows by the arguments of Theorem 8.14 Step 2 and 3. □

Remark 8.17. Under the assumptions of Theorem 8.16.i) one can prove that u_τ and $\widehat{u}_\tau$ converge strongly to u in $L^p(0,T;\mathcal{V})$ for $p \in [1,\infty)$ and weakly* for $p = \infty$.

The conditions i) and ii) in Theorem 8.16 are the assumptions of Theorem 6.9 and 6.8, respectively, where the discretized right-hand sides converge in the associated spaces. In Section 8.5 we prove for the implicit Euler method and each of these conditions that, in contrast to general algebraically stable Runge-Kutta methods, the convergence order of the Lagrange multiplier λ is not reduced for infinite-dimensional DAEs, if the solution is regular enough.

8.3. Algebraically and L-Stable Runge-Kutta Methods

In this section, we analyze the convergence of a special class of Runge-Kutta schemes applied to the operator DAEs (8.4). Note that in general for operator ODEs/DAEs, an implicit Runge-Kutta scheme may not even provide a unique approximation, which then leads to unbounded solutions and thus, to divergence; see Example 8.20. Thus, we first give sufficient conditions on the approximation scheme, which guarantee a unique solution in every time step. Afterwards we study the convergence behavior of the discrete solution.

We consider an s-stage Runge-Kutta scheme given by the Butcher tableau $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$. As mentioned in Section 5.1, we assume $\mathbf{A}$ to be regular and $R(\infty) = 1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s = 0$ with $\mathbb{1}_s := [1, \ldots, 1]^T \in \mathbb{R}^s$. In this case, the approximations of λ and γ are independent of their approximations from the previous time step.

For the application of the Runge-Kutta method to the operator DAE (8.4), we need the spaces $\mathcal{V}$, $\mathcal{H}$, and $\mathcal{Q}$ in s components. This is necessary in order to define generalized state vectors. Therefore, we introduce

$$\mathcal{V}_s := \mathcal{V}^s, \quad \mathcal{V}_{\ker,s} := (\mathcal{V}_{\ker})^s, \quad \mathcal{V}_{\mathrm{c},s} := (\mathcal{V}_\mathrm{c})^s, \quad \mathcal{H}_s := \mathcal{H}^s, \quad \mathcal{H}_{\ker,s} := \mathcal{H}_{\ker}^s, \quad \mathcal{Q}_s := \mathcal{Q}^s.$$

equipped with the associated norms $\|\boldsymbol{x}\|_{\mathcal{X}^s} := \big(\sum_{i=1}^s \|\boldsymbol{x}_i\|_\mathcal{X}^2\big)^{1/2}$. Accordingly, we define their dual spaces $\mathcal{V}_s^*$, $\mathcal{V}_{\ker,s}^*$, $\mathcal{H}_s^*$, $\mathcal{H}_{\ker,s}^*$, and $\mathcal{Q}_s^*$.

8.3.1. Temporal Discretization

Similar to the finite-dimensional case u_n, λ_n, and γ_n are approximations of u, λ, and γ at time $t_n = n\tau$, respectively. We introduce the internal stages

$$\boldsymbol{u}_n = \begin{bmatrix} \boldsymbol{u}_{n,1} \\ \vdots \\ \boldsymbol{u}_{n,s}\end{bmatrix} \in \mathcal{V}_s, \qquad \boldsymbol{\lambda}_n = \begin{bmatrix} \boldsymbol{\lambda}_{n,1} \\ \vdots \\ \boldsymbol{\lambda}_{n,s}\end{bmatrix} \in \mathcal{Q}_s, \qquad \boldsymbol{\gamma}_n = \begin{bmatrix} \gamma_{n,1} \\ \vdots \\ \gamma_{n,s}\end{bmatrix} \in \mathcal{Q}_s.$$

These stage vectors call for corresponding operators such as $\mathcal{A}_s\colon \mathcal{V}_s \to \mathcal{V}_s^*$, which is induced by $\mathcal{A}$ via a componentwise application. In the sequel, we do not distinguish between these two operators such that for $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}_s$ we write

$$\langle \mathcal{A}\boldsymbol{u}, \boldsymbol{v}\rangle \coloneqq \langle \mathcal{A}_s\boldsymbol{u}, \boldsymbol{v}\rangle \coloneqq \sum_{i=1}^{s} \langle \mathcal{A}\boldsymbol{u}_i, \boldsymbol{v}_i\rangle.$$

In a corresponding manner, the operators $\mathcal{B}$ and $\mathcal{C}$ can be applied componentwise to elements with s components.

Finally, we denote for an arbitrary matrix $M \in \mathbb{R}^{m\times s}$ and an element $\boldsymbol{x} \in \mathcal{X}^s$ by $M\boldsymbol{x} \in \mathcal{X}^m$ the formal matrix-vector multiplication $(M\boldsymbol{x})_k \coloneqq \sum_{i=1}^{s} M_{k,i}\boldsymbol{x}_i \in \mathcal{X}$ for $k = 1, \dots, m$.

Lemma 8.18. *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed spaces. Consider a matrix $M \in \mathbb{R}^{s\times s}$ and a linear operator $\mathcal{K}\colon \mathcal{X} \to \mathcal{Y}^*$, which induces a linear operator $\mathcal{K}\colon \mathcal{X}^s \to (\mathcal{Y}^s)^*$ by a componentwise application. Then $\langle \mathcal{K}M\boldsymbol{x}, \boldsymbol{y}\rangle = \langle M\mathcal{K}\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \mathcal{K}\boldsymbol{x}, M^T\boldsymbol{y}\rangle$ holds for all $\boldsymbol{x} \in \mathcal{X}^s$ and $\boldsymbol{y} \in \mathcal{Y}^s$.*

Proof. The result follows by a simple calculation,

$$\langle \mathcal{K}M\boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{k=1}^{s} \Big\langle \mathcal{K}\sum_{i=1}^{s} M_{k,i}\boldsymbol{x}_i, \boldsymbol{y}_k\Big\rangle = \sum_{k,i=1}^{s} M_{k,i}\langle \mathcal{K}\boldsymbol{x}_i, \boldsymbol{y}_k\rangle = \sum_{i=1}^{s} \Big\langle \mathcal{K}\boldsymbol{x}_i, \sum_{k=1}^{s} M_{k,i}\boldsymbol{y}_k\Big\rangle = \langle \mathcal{K}\boldsymbol{x}, M^T\boldsymbol{y}\rangle. \quad \square$$

The approximations of the right-hand sides f and g need to be extended for elements with s components as well. For this, we introduce $\boldsymbol{f}_n \in \mathcal{V}_s^*$ and $\boldsymbol{g}_n, \dot{\boldsymbol{g}}_n \in \mathcal{Q}_s^*$, $n = 1, \dots, N$. As in Subsection 8.2.1, the specific definition of $\boldsymbol{f}_n$, $\boldsymbol{g}_n$, and $\dot{\boldsymbol{g}}_n$ is not of importance as long as it satisfies an analog of Assumption 8.10. Anyway, we emphasize that it is not possible to estimate the internal stages $\boldsymbol{u}_n$, $n = 1, \dots, N$ under Assumption 8.1 in the time-discrete counterpart of $L^\infty(0,T;\mathcal{H}_s)$ [EmmT10, p. 793]. Therefore, we have to strengthen the assumptions on the right-hand sides in comparison to Assumption 8.10. We demand $f \in L^2(0,T;\mathcal{V}^*)$ and $g \in H^1(0,T;\mathcal{Q}^*)$. We now state the assumptions on the approximations of f, g and $\dot{g}$.

Assumption 8.19. Let $\boldsymbol{f}_n \in \mathcal{V}_s^*$, $\boldsymbol{g}_n \in \mathcal{Q}_s^*$, and $\dot{\boldsymbol{g}}_n \in \mathcal{Q}_s^*$ be given for $n = 1, \dots, N$. The functions $\boldsymbol{f}_\tau$, $\boldsymbol{g}_\tau$, and $\dot{\boldsymbol{g}}_\tau$ denote the piecewise constant functions defined on $[0,T]$ with

$$\boldsymbol{f}_\tau(t)|_{(t_{n-1},t_n]} \equiv \boldsymbol{f}_n, \qquad \boldsymbol{g}_\tau(t)|_{(t_{n-1},t_n]} \equiv \boldsymbol{g}_n, \qquad \dot{\boldsymbol{g}}_\tau(t)|_{(t_{n-1},t_n]} \equiv \dot{\boldsymbol{g}}_n,$$

for $n = 1, \dots, N$ and with a continuous extension at $t = 0$. We assume that for $\tau \to 0$ we have

$$\boldsymbol{f}_\tau \to f\mathbb{1}_s \text{ in } L^2(0,T;\mathcal{V}_s^*), \qquad \boldsymbol{g}_\tau \to g\mathbb{1}_s \text{ in } L^\infty(0,T;\mathcal{Q}_s^*), \qquad \dot{\boldsymbol{g}}_\tau \to \dot{g}\mathbb{1}_s \text{ in } L^2(0,T;\mathcal{Q}_s^*).$$

An example, which satisfies Assumption 8.19, is given by $\boldsymbol{f}_n \coloneqq f_n\mathbb{1}_s$, $\boldsymbol{g}_n \coloneqq g_n\mathbb{1}_s$, and $\dot{\boldsymbol{g}}_n \coloneqq \dot{g}_n\mathbb{1}_s$, $n = 1, \dots, N$, if f_n, g_n, and $\dot{g}_n$ fulfill Assumption 8.10 for $L^2(0,T;\mathcal{V}^*)$, $L^\infty(0,T;\mathcal{Q}^*)$, and $L^2(0,T;\mathcal{Q}^*)$, respectively. Given the Butcher tableau (5.1), we could define $\boldsymbol{g}_n$ by the componentwise function evaluation $\boldsymbol{g}_{n,\ell} \coloneqq g(t_{n-1} + \mathbf{c}_\ell\tau)$. Also, this approach satisfies Assumption 8.19, since g is uniformly continuous on $[0,T]$. In any case, we are able to prove the convergence to the solution of the operator DAE (8.1). Recall that we aim for convergence behavior under minimal assumptions on the data in this section. The convergence order is studied in Section 8.5.

With the introduced notation, the temporal discretization of system (8.4) yields the time-discrete problem

$$u_n = \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{u}_n, \qquad \lambda_n = \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\lambda}_n, \qquad \gamma_n = \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\gamma}_n, \tag{8.26}$$

where $\boldsymbol{u}_n$, $\boldsymbol{\lambda}_n$, and $\boldsymbol{\gamma}_n$ satisfy the operator equation

$$\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_n + \mathcal{A}\boldsymbol{u}_n - \mathcal{B}^*\boldsymbol{\lambda}_n - \mathcal{B}^*\boldsymbol{\gamma}_n = \boldsymbol{f}_n \quad \text{in } \mathcal{V}_s^*, \tag{8.27a}$$

$$\mathcal{B}\boldsymbol{u}_n + \mathcal{C}\boldsymbol{\gamma}_n = \boldsymbol{g}_n \quad \text{in } \mathcal{Q}_s^*, \tag{8.27b}$$

$$\mathcal{B}\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_n = \dot{\boldsymbol{g}}_n \quad \text{in } \mathcal{Q}_s^*. \tag{8.27c}$$

The discrete derivative $\boldsymbol{D}_\tau\boldsymbol{u}_n$ is given by $(\boldsymbol{u}_n - u_{n-1}\mathbb{1}_s)/\tau$.

Unfortunately, $\boldsymbol{u}_n$, $\boldsymbol{\lambda}_n$, and $\boldsymbol{\gamma}_n$ are not bounded in terms of the right-hand sides for all Runge-Kutta schemes, even for an arbitrarily small step size τ as we show by means of the following example.

Example 8.20. Consider the discretization (8.27) with vanishing right-hand sides and $u_0 = 0$. Furthermore, we assume that the operator $\mathcal{A}$ is self-adjoint and that $\mathcal{V}$ is compactly embedded in $\mathcal{H}$; see [Alt16, Def. 10.1] for a definition. We show that the discrete solution given by the 2-stage stiffly accurate Runge-Kutta scheme from Example 5.2 may be non-zero no matter how small τ is chosen and thus, not stable with respect to the initial value and the right-hand sides. For this, we note that $\mathbf{A}^{-1}$ has a negative eigenvalue $\alpha \in \mathbb{R}$ with eigenvector $\boldsymbol{w} \in \mathbb{R}^2$, which satisfies $\mathbf{b}^T\boldsymbol{w} \neq 0$.

Since $\langle \mathcal{A}\,\cdot, \cdot\rangle$ defines an elliptic, bounded, and symmetric bilinear form on $\mathcal{V}_{\text{ker}}$, there exist countably many eigenpairs $(\mu_k, v_k) \in \mathbb{R} \times \mathcal{V}_{\text{ker}}$ of the infinite-dimensional eigenvalue problem $\mu v = \mathcal{A}v$ in $\mathcal{V}_{\text{ker}}^*$. More precisely, all μ_k are positive and tend to infinity as $k \to \infty$ and v_k are normalized for all $k \in \mathbb{N}$ [Mic62, Ch. 4.34]. Let $\varepsilon > 0$ be arbitrarily small and choose k large enough such that $\tau := |\alpha|/\mu_k < \varepsilon$. We set $\boldsymbol{u} := v_k\boldsymbol{w} \in \mathcal{V}_{\text{ker},s}$. The given eigenvalue problem implies $(\mathbf{A}^{-1} + \tau\mathcal{A})\boldsymbol{u} \in \mathcal{V}_{\text{ker},s}^0$ such that there exists a unique $\boldsymbol{\lambda}$ with

$$\mathcal{B}^*\boldsymbol{\lambda} = \big(\tau^{-1}\mathbf{A}^{-1} + \mathcal{A}\big)\boldsymbol{u} \quad \text{in } \mathcal{V}_s^*.$$

Thus, the tuple $(\boldsymbol{u}, \boldsymbol{\lambda}, 0)$ satisfies system (8.27) and we obtain as approximation in the first time step

$$u_1 = \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{u} = \alpha\mathbf{b}^T\boldsymbol{w}v_k \neq 0.$$

In summary, one step of the given Runge-Kutta scheme with step size τ yields an approximation, which is unbounded with respect to the data.

Example 8.20 shows that it is not sufficient to require that the discretization scheme satisfies $R(\infty) = 0$. We introduce a class of Runge-Kutta methods, which provide a unique and bounded solution for every discrete time point. For this, we state further assumptions on the Runge-Kutta scheme.

Assumption 8.21. The Runge-Kutta method (5.1) is algebraically stable, i.e., the matrix $\mathbf{BA} + \mathbf{A}^T\mathbf{B} - \mathbf{bb}^T$ is positive semidefinite with the diagonal matrix $\mathbf{B}_{ii} = \mathbf{b}_i$, and L-stable, i.e., $R(\infty) = 0$. All weights $\mathbf{b}_i$ are positive and its classical order is at least one, i.e., $\sum_{i=1}^s \mathbf{b}_i = \mathbb{1}_s^T\mathbf{b} = 1$.

Example 8.22. Radau IA, Radau IIA, Lobatto IIIC [HaiW96, p. 72 ff.], and Lobatto IIID [NørW81, p. 205] methods satisfy Assumption 8.21; see [HaiW96, Ch. IV, Pro. 3.8, Th. 12.7 & 12.9] and [Jay15, p. 822].

With the given assumptions on the discretization scheme, we are able to show the unique solvability for every time step.

Lemma 8.23 (Solvability of the Time-Discrete System)**.** *Consider $u_{n-1} \in \mathcal{H}_{\text{ker}} + \mathcal{V}_{\text{c}}$, $n \in \{1, \dots, N\}$, and right-hand sides $\boldsymbol{f}_n \in \mathcal{V}_s^*$ and $\boldsymbol{g}_n, \dot{\boldsymbol{g}}_n \in \mathcal{Q}_s^*$. Suppose that Assumption 8.1.i) on the operators is fulfilled and $\mathcal{C} \in \mathscr{L}(\mathcal{Q}, \mathcal{Q}^*)$ is elliptic. If the Runge-Kutta method satisfies Assumption 8.21, then system (8.27) has a unique solution of internal stages $(\boldsymbol{u}_n, \boldsymbol{\lambda}_n, \boldsymbol{\gamma}_n) \in \mathcal{V}_s \times \mathcal{Q}_s \times \mathcal{Q}_s$ and thus, there exists a unique approximation $(u_n, \lambda_n, \gamma_n) \in \mathcal{V} \times \mathcal{Q} \times \mathcal{Q}$.*

Proof. Since $\mathbf{M} := \mathbf{BA} + \mathbf{A}^T\mathbf{B} - \mathbf{bb}^T$ is positive semidefinite by Assumption 8.21, the inequality

$$\boldsymbol{x}^T\mathbf{BA}^{-1}\boldsymbol{x} = \frac{1}{2}(\mathbf{A}^{-1}\boldsymbol{x})^T[\mathbf{BA} + \mathbf{A}^T\mathbf{B}](\mathbf{A}^{-1}\boldsymbol{x}) \geq \frac{1}{2}(\mathbf{A}^{-1}\boldsymbol{x})^T\mathbf{M}(\mathbf{A}^{-1}\boldsymbol{x}) \geq 0$$

is satisfied for arbitrary $\boldsymbol{x} \in \mathbb{R}^s$ and consequently $\mathbf{BA}^{-1}$ is also positive semidefinite. If we multiply the equations (8.27a) and (8.27b) by $\mathbf{B}$ and (8.27c) by $\mathbf{BA}$, then it results in the system

$$\begin{aligned}
\mathbf{BA}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_n + \mathcal{A}\mathbf{B}\boldsymbol{u}_n - \mathcal{B}^*\mathbf{B}\boldsymbol{\lambda}_n - \mathcal{B}^*\mathbf{B}\boldsymbol{\gamma}_n &= \mathbf{B}\boldsymbol{f}_n \quad &&\text{in } \mathcal{V}_s^*, &&(8.28a)\\
\mathcal{B}\mathbf{B}\boldsymbol{u}_n + \mathcal{C}\mathbf{B}\boldsymbol{\gamma}_n &= \mathbf{B}\boldsymbol{g}_n \quad &&\text{in } \mathcal{Q}_s^*, &&(8.28b)\\
\mathcal{B}\mathbf{B}\boldsymbol{D}_\tau\boldsymbol{u}_n &= \mathbf{BA}\dot{\boldsymbol{g}}_n \quad &&\text{in } \mathcal{Q}_s^*. &&(8.28c)
\end{aligned}$$

Note that we have used $\mathbf{B}\mathcal{A} = \mathcal{A}\mathbf{B}$ as well as similar results for the other operators. Let $\mathbf{B}^{1/2}$ be the diagonal matrix with $\mathbf{B}_{ii}^{1/2} = \sqrt{\mathbf{b}_i}$. Since

$$\langle \mathbf{BA}^{-1}\boldsymbol{u}_{\ker} + \tau\mathcal{A}\mathbf{B}\boldsymbol{u}_{\ker}, \boldsymbol{u}_{\ker}\rangle \geq \langle \tau\mathcal{A}\mathbf{B}\boldsymbol{u}_{\ker}, \boldsymbol{u}_{\ker}\rangle \geq \tau\mu_{\mathcal{A}}\|\mathbf{B}^{1/2}\boldsymbol{u}_{\ker}\|_{\mathcal{V}_s}^2 \geq \tau\mu_{\mathcal{A}} \min_{i=1,\dots,s} \mathbf{b}_i\|\boldsymbol{u}_{\ker}\|_{\mathcal{V}_s}^2$$

for all $\boldsymbol{u}_{\ker} \in \mathcal{V}_{\ker,s}$, the operator $\mathbf{BA}^{-1} + \tau\mathcal{A}\mathbf{B}$ is elliptic. The solvability then follows by the invertibility of $\mathbf{B}$ and a similar argument as in the implicit Euler case in Lemma 8.12. □

Remark 8.24. System (8.28) preserves the saddle point structure of the time-continuous operator DAE (8.4). In particular, under spatial discretization and a rearrangement the system (8.28) reads

$$\left[\begin{array}{cc|c}
\frac{1}{\tau}(\mathbf{BA}^{-1} \otimes M) + (\mathbf{B} \otimes A) & -\mathbf{B} \otimes B^T & -\mathbf{B} \otimes B^T\\
\mathbf{B} \otimes B & \mathbf{B} \otimes C & 0\\ \hline
\mathbf{B} \otimes B & 0 & 0
\end{array}\right]
\begin{bmatrix} \boldsymbol{x}_n \\ \gamma_n \\ \boldsymbol{\mu}_n \end{bmatrix}
=
\begin{bmatrix}
(\mathbf{B} \otimes I_{n_x})\boldsymbol{d}_n + \frac{1}{\tau}(\mathbf{BA}^{-1} \otimes M)\boldsymbol{x}_{n-1}\\
(\mathbf{B} \otimes I_{n_\mu})\boldsymbol{h}_n\\
\tau(\mathbf{BA} \otimes I_{n_\mu})\dot{\boldsymbol{h}}_n + (\mathbf{B} \otimes B)\boldsymbol{x}_{n-1}
\end{bmatrix}$$

with the notation of Subsection 8.1.1. Here, C denotes the discrete version of the elliptic operator $\mathcal{C}$. We emphasis that $\mathbf{B}$ is a diagonal matrix and that the top left $s(n_x + n_\mu) \times s(n_x + n_\mu)$-block of the iteration matrix is positive definite.

Before we investigate the convergence of the Runge-Kutta schemes applied to operator DAEs, we summarize results on the convergence for unconstrained operator equations.

8.3.2. Convergence Results for Operator Differential Equations

Let us consider a linear operator ODE

$$\dot{v}(t) + \mathcal{A}v(t) = f(t) \quad \text{in } \mathcal{V}^* \tag{8.29}$$

with initial condition $v(0) = v_0 \in \mathcal{H}$, right-hand side $f \in L^2(0,T;\mathcal{V}^*)$, and elliptic operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$. These assumptions guarantee a unique solution by Theorem 4.22. The following convergence analysis is based on the paper [EmmT10], where the authors investigate the behavior of stiffly accurate and algebraically stable Runge-Kutta schemes applied to the evolution problem (8.29). They assume that the schemes are of at least first order and all entries of $\mathbf{b}$ are positive. Note that such methods fulfill Assumption 8.21.

Lemma 8.25. *Let the Runge-Kutta method with Butcher tableau* $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$ *satisfy Assumption 8.21. Suppose that* $\mathcal{K} \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ *is elliptic and self-adjoint. Then for every* $x_0 \in \mathcal{X}$ *and* $\boldsymbol{x} \in \mathcal{X}^s$ *we have*

$$2\langle \mathcal{K}\boldsymbol{x}, \mathbf{BA}^{-1}(\boldsymbol{x} - x_0\mathbb{1}_s)\rangle \geq \langle \mathcal{K}\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{x}, \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{x}\rangle - \langle \mathcal{K}x_0, x_0\rangle. \tag{8.30}$$

Proof. The proof follows the lines of Lemma 5.7. □

In the following theorem, we consider the temporal discretization of the operator ODE (8.29) by the Runge-Kutta method (5.1). The associated time-discrete system is given by

$$v_n = \mathbf{b}^T \mathbf{A}^{-1} \boldsymbol{v}_n, \tag{8.31a}$$
$$\mathbf{A}^{-1} \boldsymbol{D}_\tau \boldsymbol{v}_n + \mathcal{A} \boldsymbol{v}_n = \boldsymbol{f}_n \tag{8.31b}$$

with the discrete derivative $\boldsymbol{D}_\tau \boldsymbol{v}_n$ defined as in (8.27).

Theorem 8.26. *Consider the operator ODE (8.29) with $f \in L^2(0,T;\mathcal{V}^*)$, initial data $v_0 \in \mathcal{H}$, and an elliptic operator $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$. The corresponding exact solution is denoted by v. Assume that in (8.31) the step size τ is uniform and that the Runge-Kutta method satisfies Assumption 8.21. Suppose that the piecewise constant function $\boldsymbol{f}_\tau \in L^2(0,T;\mathcal{V}_s^*)$ defined by $\boldsymbol{f}_\tau(t) = \boldsymbol{f}_n$ for $t \in (t_{n-1}, t_n]$ satisfies $\boldsymbol{f}_\tau \to f \mathbb{1}_s$ in $L^2(0,T;\mathcal{V}_s^*)$ as $\tau \to 0$.*

Then there exists a unique solution $v_n \in \mathcal{V}$ and $\boldsymbol{v}_n \in \mathcal{V}_s$ of system (8.31) for every time step $n = 1, \dots, N$. Furthermore, the piecewise constant functions $\boldsymbol{v}_\tau$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{v}_\tau$ associated to $\boldsymbol{v}_n$ and $D_\tau v_n$ for $n = 1, \dots, N$, respectively, cf. Subsection 7.1.1.2, and v_N are weakly convergent in the sense

$$\boldsymbol{v}_\tau \rightharpoonup v \mathbb{1}_s \text{ in } L^2(0,T;\mathcal{V}_s), \qquad v_N \rightharpoonup v(T) \text{ in } \mathcal{H}, \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{v}_\tau \rightharpoonup \dot{v} \text{ in } L^2(0,T;\mathcal{V}^*).$$

Proof. By the same arguments as in the proof of Lemma 8.23 one shows that $\mathbf{B}\mathbf{A}^{-1} + \tau \mathcal{A}\mathbf{B}$ is elliptic and bounded. The existence of a unique solution of (8.31) then follows by the Lax-Milgram Theorem 3.4. With $\mathbf{b}^T \mathbf{A}^{-1} \boldsymbol{v}_n = v_n$ and estimate (8.30) one proves the stated convergence behavior by an adaptation of the proof of [EmmT10, Th. 5.1 & Rem. 5.3]. □

8.3.3. Convergence Results for Operator Differential-Algebraic Equations

In this section, we investigate the convergence behavior of the semi-discretized system (8.27). For this, we recall the piecewise constant and piecewise linear approximations u_τ, $\widehat{u}_\tau$, $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$, λ_τ, and γ_τ from Subsection 8.2.2. For the internal stages we introduce accordingly

$$\begin{aligned} \boldsymbol{u}_\tau(t) &:= \begin{cases} u_0 \mathbb{1}_s, & \text{if } t = 0 \\ \boldsymbol{u}_n, & \text{if } t \in (t_{n-1}, t_n] \end{cases}, \qquad & \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_\tau(t) &:= \begin{cases} 0, & \text{if } t = 0 \\ \boldsymbol{D}_\tau \boldsymbol{u}_n, & \text{if } t \in (t_{n-1}, t_n] \end{cases}, \\ \boldsymbol{\lambda}_\tau(t) &:= \boldsymbol{\lambda}_n, \ \text{if } t \in (t_{n-1}, t_n], & \boldsymbol{\gamma}_\tau(t) &:= \boldsymbol{\gamma}_n, \ \text{if } t \in (t_{n-1}, t_n]. \end{aligned} \tag{8.32}$$

The values for $\boldsymbol{\lambda}_\tau$ and $\boldsymbol{\gamma}_\tau$ at time $t = 0$ can be chosen arbitrarily.

Theorem 8.27 (Convergence of L-Stable Runge-Kutta Schemes)**.** *Let the approximations $\boldsymbol{f}_\tau$, $\boldsymbol{g}_\tau$, and $\dot{\boldsymbol{g}}_\tau$ of the right-hand sides $f \in L^2(0,T;\mathcal{V}^*)$, $g \in H^1(0,T;\mathcal{Q}^*)$ satisfy Assumption 8.19. Suppose that Assumptions 8.1.i), iii) on the operators and the initial value are fulfilled. The corresponding solution of the operator DAE (8.1) is denoted by (u, λ). Then, every Runge-Kutta scheme, which satisfies Assumption 8.21 yields the convergence results*

$$\begin{aligned} u_\tau &\to u \quad \text{in } L^2(0,T;\mathcal{V}), & \widehat{u}_\tau &\to u \quad \text{in } L^2(0,T;\mathcal{H}), \\ \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau &\to \dot{u} \quad \text{in } L^2(0,T;\mathcal{V}_{\ker}^*), & \gamma_\tau &\to 0 \quad \text{in } L^\infty(0,T;\mathcal{Q}) \end{aligned}$$

for $\tau \to 0$. Furthermore, $\int_0^\cdot \mathbf{b}^T \boldsymbol{\lambda}_\tau(s)\,\mathrm{d}s$ converges strongly to Λ in $L^2(0,T;\mathcal{Q})$.

Proof. We follow the steps of the proof of Theorem 8.14.

Step 1 (Convergence of γ_τ): With (8.27b), a successive application of (8.27c), and $\mathbf{b}^T \mathbf{A}^{-1} \boldsymbol{D}_\tau \boldsymbol{u}_n =$

$D_\tau u_n$ we obtain

$$
\begin{aligned}
\mathcal{C}\boldsymbol{\gamma}_n &= \boldsymbol{g}_n - \tau \mathbf{A}\dot{\boldsymbol{g}}_n - \tau \sum_{i=1}^{n-1} \mathbf{b}^T \dot{\boldsymbol{g}}_i \mathbb{1}_s - \mathcal{B}u_0 \mathbb{1}_s \\
&= \tau\Big(\mathbf{b}^T \dot{\boldsymbol{g}}_n \mathbb{1}_s - \mathbf{A}\dot{\boldsymbol{g}}_n\Big) + \Big(\int_0^{t_n} \dot{g}(t) - \mathbf{b}^T \dot{\boldsymbol{g}}_\tau(t)\,\mathrm{d}t\Big)\mathbb{1}_s + \boldsymbol{g}_n - g(t_n)\mathbb{1}_s. \qquad (8.33)
\end{aligned}
$$

Furthermore, with $\mathbf{b}^T \mathbf{A}^{-1} \mathbb{1}_s = 1$ the equality

$$
\mathcal{C}\gamma_n = \mathcal{C}\mathbf{b}^T \mathbf{A}^{-1} \gamma_n = \int_0^{t_n} \dot{g}(t) - \mathbf{b}^T \dot{\boldsymbol{g}}_\tau(t)\,\mathrm{d}t + \mathbf{b}^T \mathbf{A}^{-1}\big(\boldsymbol{g}_n - g(t_n)\mathbb{1}_s\big)
$$

holds. Similar as in the proof of Theorem 8.14, Assumption 8.19, the Cauchy-Schwarz inequality, and $\mathbf{b}^T \mathbb{1}_s = 1$ imply

$$
\lim_{\tau\to 0} \|\gamma_\tau\|_{L^\infty(0,T;\mathcal{Q})} \lesssim \lim_{\tau\to 0} \sqrt{T}\|\dot{g}\mathbb{1}_s - \dot{\boldsymbol{g}}_\tau\|_{L^2(0,T;\mathcal{Q}_s^*)} + \|g\mathbb{1}_s - \boldsymbol{g}_\tau\|_{L^\infty(0,T;\mathcal{Q}_s^*)} = 0. \qquad (8.34)
$$

Given equation (8.33), Assumption 8.19 also implies $\boldsymbol{\gamma}_\tau \to 0$ in $L^2(0,T;\mathcal{Q}_s)$ by the estimate

$$
\|\boldsymbol{\gamma}_\tau\|^2_{L^2(0,T;\mathcal{Q}_s)} \lesssim \tau\,\|\dot{\boldsymbol{g}}_\tau\|^2_{L^2(0,T;\mathcal{Q}_s^*)} + T^2\,\|\dot{g}\mathbb{1}_s - \dot{\boldsymbol{g}}_\tau\|^2_{L^2(0,T;\mathcal{Q}_s^*)} + T\,\|g\mathbb{1}_s - \boldsymbol{g}_\tau\|^2_{L^\infty(0,T;\mathcal{Q}_s^*)}.
$$

Step 2 (Weak Convergence of u_τ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$): Note that the splitting $\mathcal{V} = \mathcal{V}_{\mathrm{ker}} \oplus \mathcal{V}_{\mathrm{c}}$ implies the splitting $\mathcal{V}_s = \mathcal{V}_{\mathrm{ker},s} \oplus \mathcal{V}_{\mathrm{c},s}$. With this, we obtain

$$
\boldsymbol{u}_n = \boldsymbol{u}_{\mathrm{ker},n} + \boldsymbol{u}_{\mathrm{c},n}, \qquad \boldsymbol{D}_\tau \boldsymbol{u}_n = \boldsymbol{D}_\tau \boldsymbol{u}_{\mathrm{ker},n} + \boldsymbol{D}_\tau \boldsymbol{u}_{\mathrm{c},n}.
$$

Analogously, we split the global approximations into

$$
\boldsymbol{u}_\tau = \boldsymbol{u}_{\mathrm{ker},\tau} + \boldsymbol{u}_{\mathrm{c},\tau}, \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_\tau = \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} + \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{c},\tau}.
$$

Thus, formula (8.27b) yields

$$
\boldsymbol{u}_{\mathrm{c},\tau} = \mathcal{B}_{\mathcal{A}}^{-}\boldsymbol{g}_\tau + \mathcal{B}_{\mathcal{A}}^{-}\mathcal{C}\boldsymbol{\gamma}_\tau \;\to\; \mathcal{B}_{\mathcal{A}}^{-} g\mathbb{1}_s = u_{\mathrm{c}}\mathbb{1}_s \quad \text{in } L^2(0,T;\mathcal{V}_s),
$$

which implies $u_{\mathrm{c},\tau} \to u_{\mathrm{c}}$ and respectively by (8.27c) and $\mathbf{b}^T \mathbb{1}_s = 1$,

$$
\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{c},\tau} = \mathbf{b}^T \mathbf{A}^{-1}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{c},\tau}\big) = \mathcal{B}_{\mathcal{A}}^{-}\mathbf{b}^T \dot{\boldsymbol{g}}_\tau \;\to\; \mathcal{B}_{\mathcal{A}}^{-}\mathbf{b}^T \dot{g}\mathbb{1}_s = \mathcal{B}_{\mathcal{A}}^{-}\dot{g} = \dot{u}_{\mathrm{c}} \quad \text{in } L^2(0,T;\mathcal{V}).
$$

By a combination of the equations (8.27a), (8.27c) and a restriction of the test functions to $\mathcal{V}_{\mathrm{ker},s}$, we obtain

$$
\mathbf{A}^{-1}\boldsymbol{D}_\tau \boldsymbol{u}_{\mathrm{ker},n} + \mathcal{A}\boldsymbol{u}_{\mathrm{ker},n} = \boldsymbol{f}_n - \mathbf{A}^{-1}\boldsymbol{D}_\tau \boldsymbol{u}_{\mathrm{c},n} = \boldsymbol{f}_n - \mathcal{B}_{\mathcal{A}}^{-}\dot{\boldsymbol{g}}_n \quad \text{in } \mathcal{V}^*_{\mathrm{ker},s}. \qquad (8.35)
$$

Note that (8.35) equals the Runge-Kutta approximation of the operator ODE (8.20). With the initial value $u_{\mathrm{ker},0} \in \mathcal{H}_{\mathrm{ker}}$, the conditions of Theorem 8.26 are satisfied. Thus, $\boldsymbol{u}_{\mathrm{ker},\tau}$ converges weakly to $u_{\mathrm{ker}}\mathbb{1}_s$ in $L^2(0,T;\mathcal{V}_{\mathrm{ker},s})$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau}$ converges weakly to $\dot{u}_{\mathrm{ker}}$ in $L^2(0,T;\mathcal{V}^*_{\mathrm{ker}})$ as $\tau \to 0$.

Step 3 (Strong Convergence of u_τ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$): For the strong convergence we note that by equation (8.35) we have

$$
\begin{aligned}
&\|\boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\|^2_{L^2(0,T;\mathcal{V}_s)} \\
&\quad\lesssim \int_0^T \big\langle \mathcal{A}\mathbf{B}\big(\boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big),\, \boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t
\end{aligned}
$$

$$= -\int_0^T \big\langle \mathbf{BA}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}, \boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t + \int_0^T \big\langle \mathbf{B}\dot{u}_{\mathrm{ker}}\mathbb{1}_s,\, \boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t$$
$$+ \int_0^T \big\langle \mathbf{B}(\boldsymbol{f}_\tau - f\mathbb{1}_s) - \mathcal{B}_\mathcal{A}^-\mathbf{B}(\dot{\boldsymbol{g}}_\tau - \dot{g}\mathbb{1}_s),\, \boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t, \tag{8.36}$$

since $\mathcal{A}$ is elliptic and all $\mathbf{b}_i$ are positive. As for the implicit Euler method, we only need to analyze the first integral, since the remaining terms vanish as $\tau \to 0$ by the weak convergence of $\boldsymbol{u}_{\mathrm{ker},\tau}$ and Assumption 8.19. By Lemma 8.25 we obtain

$$\int_0^T \big\langle \mathbf{BA}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}, \boldsymbol{u}_{\mathrm{ker},\tau}\big\rangle\,\mathrm{d}t = \sum_{n=1}^N \tau\,\big\langle \mathbf{BA}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_{\mathrm{ker},n},\ \boldsymbol{u}_{\mathrm{ker},n}\big\rangle \geq \frac{1}{2}\|u_{\mathrm{ker},N}\|^2_\mathcal{H} - \frac{1}{2}\|u_{\mathrm{ker},0}\|^2_\mathcal{H}.$$

From Theorem 8.26 we know $u_{\mathrm{ker},N} \rightharpoonup u_{\mathrm{ker}}(T)$. As for the implicit Euler method this limit implies

$$\liminf_{\tau\to 0} \int_0^T \big\langle \mathbf{BA}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}, \boldsymbol{u}_{\mathrm{ker},\tau}\big\rangle\,\mathrm{d}t \geq \int_0^T \big\langle \dot{u}_{\mathrm{ker}}, u_{\mathrm{ker}}\big\rangle\,\mathrm{d}t.$$

Further, by the convergence results for $\boldsymbol{u}_\tau$, $\boldsymbol{f}_\tau$, $\dot{\boldsymbol{g}}_\tau$, equation (8.35), and $\mathbb{1}_s^T\mathbf{B}\mathbb{1}_s = 1$ we get

$$\int_0^T \big\langle \mathbf{BA}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}, u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t \to \int_0^T \big\langle \mathbf{B}\dot{u}_{\mathrm{ker}}\mathbb{1}_s, u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t = \int_0^T \langle \dot{u}_{\mathrm{ker}}, u_{\mathrm{ker}}\rangle\,\mathrm{d}t.$$

As in the proof of Theorem 8.14 we conclude with (8.36) that $\boldsymbol{u}_{\mathrm{ker},\tau} \to u_{\mathrm{ker}}\mathbb{1}_s$ in $L^2(0,T;\mathcal{V}_{\mathrm{ker},s})$. A direct implication is given by

$$u_{\mathrm{ker},\tau} = \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{u}_{\mathrm{ker},\tau} \ \to\ \mathbf{b}^T\mathbf{A}^{-1}u_{\mathrm{ker}}\mathbb{1}_s = u_{\mathrm{ker}} \qquad \text{in } L^2(0,T;\mathcal{V}_{\mathrm{ker}}) \subset L^2(0,T;\mathcal{V}).$$

Furthermore, we obtain the convergence of $\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau}$ in $L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*)$ by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau} = \mathbf{b}^T\mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} = \mathbf{b}^T(\boldsymbol{f}_\tau - \mathcal{B}_\mathcal{A}^-\dot{\boldsymbol{g}}_\tau - \mathcal{A}\boldsymbol{u}_{\mathrm{ker},\tau}) \ \to\ \mathbf{b}^T(f - \mathcal{B}_\mathcal{A}^-\dot{g} - \mathcal{A}u_{\mathrm{ker}})\mathbb{1}_s = \dot{u}_{\mathrm{ker}}.$$

By the proven convergences of their parts in $\mathcal{V}_\mathrm{c}$ and $\mathcal{V}_{\mathrm{ker}}$, u_τ and $\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$ converge strongly to u and $\dot{u}$.

Step 4 (Convergence of $\widehat{u}_\tau$): For the convergence of $\widehat{u}_{\mathrm{c},\tau} \to u_\mathrm{c}$ we argue as in the proof of Theorem 8.14. For $\widehat{u}_{\mathrm{ker},\tau}$ we observe

$$\begin{aligned}
&\frac{1}{2\tau}\Big(\|u_{\mathrm{ker},n}\|^2_\mathcal{H} - \|u_{\mathrm{ker},n-1}\|^2_\mathcal{H} + \|u_{\mathrm{ker},n} - u_{\mathrm{ker},n-1}\|^2_\mathcal{H}\Big)\\
&\overset{(7.17)}{=} \langle D_\tau u_{\mathrm{ker},n}, u_{\mathrm{ker},n}\rangle\\
&= \langle \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_{\mathrm{ker},n}, u_{\mathrm{ker},n}\rangle\\
&\overset{(8.35)}{\lesssim} \|\boldsymbol{f}_n\|^2_{\mathcal{V}_s^*} + \|\mathcal{B}_\mathcal{A}^-\boldsymbol{g}_n\|^2_{\mathcal{H}_s^*} + \|\boldsymbol{u}_{\mathrm{ker},n}\|^2_{\mathcal{V}_s} + \|u_{\mathrm{ker},n}\|^2_\mathcal{V}.
\end{aligned}$$

This estimate together with the telescope sum $\sum_{n=1}^N \big(\|u_{\mathrm{ker},n}\|^2_\mathcal{H} - \|u_{\mathrm{ker},n-1}\|^2_\mathcal{H}\big) = \|u_{\mathrm{ker},N}\|^2_\mathcal{H} - \|u_{\mathrm{ker},0}\|^2_\mathcal{H}$ yields

$$\begin{aligned}
\|\widehat{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker},\tau}\|^2_{L^2(0,T;\mathcal{H})} &= \frac{\tau}{3}\sum_{n=1}^N \big\|u_{\mathrm{ker},n} - u_{\mathrm{ker},n-1}\big\|^2_\mathcal{H}\\
&\lesssim \tau\bigg(\|u_{\mathrm{ker},0}\|^2_\mathcal{H} + \tau\sum_{n=1}^N \|\boldsymbol{f}_n\|^2_{\mathcal{V}_s^*} + \|\mathcal{B}_\mathcal{A}^-\boldsymbol{g}_n\|^2_{\mathcal{H}_s^*} + \|\boldsymbol{u}_{\mathrm{ker},n}\|^2_{\mathcal{V}_s} + \|u_{\mathrm{ker},n}\|^2_\mathcal{V}\bigg).
\end{aligned}$$

Note that the terms in parentheses are bounded independently of τ, since the right-hand sides are bounded by Assumption 8.19 and $u_{\ker,\tau}$ as well as $\boldsymbol{u}_{\ker,\tau}$ are convergent sequences. Thus, $\widehat{u}_{\ker,\tau}$ and $u_{\ker,\tau}$ have the same limit $u_{\ker}$ in $L^2(0,T;\mathcal{H})$, which implies the strong convergence $\widehat{u}_\tau \to u$ in $L^2(0,T;\mathcal{H})$.

Step 5 (Convergence of $\mathbf{b}^T\boldsymbol{\lambda}_\tau$): For the proof of the distributional convergence of $\mathbf{b}^T\boldsymbol{\lambda}_\tau$ we introduce primitives for the expansions of the stages $\boldsymbol{u}_\tau$, $\boldsymbol{\lambda}_\tau$, $\boldsymbol{\gamma}_\tau$, and for $\boldsymbol{f}_\tau$. We use capital letters for the absolutely continuous primitives, which vanish at $t=0$. Then we have

$$\mathcal{B}^*\mathbf{b}^T\boldsymbol{\Lambda}_\tau(t) = \mathcal{B}^*\int_0^t \mathbf{b}^T\boldsymbol{\lambda}_\tau(s)\,\mathrm{d}s = \widehat{u}_\tau(t) + \mathcal{A}\mathbf{b}^T\boldsymbol{U}_\tau(t) - \mathcal{B}^*\mathbf{b}^T\boldsymbol{\Gamma}_\tau(t) - \mathbf{b}^T\boldsymbol{F}_\tau(t) - u_0 \tag{8.37}$$

in $AC([0,T],\mathcal{V}^*)$, which follows from equation (8.27a). Then an argument as in the proof of Theorem 8.14 yields that $\mathbf{b}^T\boldsymbol{\Lambda}_\tau$ converges to Λ in $L^2(0,T;\mathcal{Q})$. □

Remark 8.28. In Theorem 8.27 we show the convergence of $\mathbf{b}^T\boldsymbol{\Lambda}_\tau$. For a proof of $\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\Lambda}_\tau = \Lambda_\tau \rightharpoonup \Lambda$ in $L^2(0,T;\mathcal{Q})$ we would need a result of the form

$$\mathbf{A}\mathbb{1}_s u_{\ker,0} + \int_0^{\cdot} \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\ker,\tau}\,\mathrm{d}s \rightharpoonup \mathbf{A}\mathbb{1}_s u_{\ker} \quad \text{in } L^2(0,T;\mathcal{V}_s^*).$$

With this, we could consider $\mathcal{B}^*\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\Lambda}_\tau$ similarly as in equation (8.37).

Remark 8.29. The proof of Theorem 8.27 also shows the strong convergence of the continuous representation of the internal stages $\boldsymbol{u}_\tau$ to $u\mathbb{1}_s$ in $L^2(0,T;\mathcal{V}_s)$. By (8.27a) and (8.27c), this implies $\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_\tau = \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\ker,\tau} + \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{c},\tau} \to \mathbf{A}\mathbb{1}_s\dot{u}_{\ker} + \mathbf{A}\mathbb{1}_s\dot{u}_{\mathrm{c}} = \mathbf{A}\mathbb{1}_s\dot{u}$ in $L^2(0,T;\mathcal{V}^*_{\ker,s})$.

As for the implicit Euler scheme, we can prove the convergence of the Lagrange multiplier λ if we assume additional regularity of the right-hand side f and the initial data.

Theorem 8.30 (Convergence with More Regular Data)**.** *In addition to the assumptions of Theorem 8.27, consider an initial value $u_0 \in \mathcal{V}$ with $\mathcal{B}u_0 = g(0)$ and a right-hand side $f \in L^2(0,T;\mathcal{H}^*)$. Furthermore, let the approximation $\boldsymbol{f}_n$ satisfy Assumption 8.19 in $L^2(0,T;\mathcal{H}_s^*)$. Assume that an operator $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$, which is self-adjoint and elliptic on $\mathcal{V}_{\ker}$, and an operator $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$ exist such that $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$. Then the approximations satisfy*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau \to \dot{u} \quad \text{in } L^2(0,T;\mathcal{H}), \qquad\qquad \lambda_\tau \to \lambda \quad \text{in } L^2(0,T;\mathcal{Q}).$$

Proof. We follow the ideas of the proofs of Theorems 8.16 and 8.27. Due to the splitting $\mathcal{V} = \mathcal{V}_{\ker} \oplus \mathcal{V}_{\mathrm{c}}$ and the strong convergence

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{c},\tau} = \mathbf{b}^T\mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{c},\tau} \;\to\; \mathbf{b}^T\mathbf{A}^{-1}\dot{u}_{\mathrm{c}}\mathbb{1}_s = \dot{u}_{\mathrm{c}} \quad \text{in } L^2(0,T;\mathcal{V}_{\mathrm{c}}) \hookrightarrow L^2(0,T;\mathcal{H}),$$

cf. the proof of Theorem 8.27, it is sufficient to consider the remaining part $\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\ker,\tau}$. We first show its weak and afterwards its strong convergence. For this, we test equation (8.35) by $\mathbf{B}\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_{\ker,n} \in \mathcal{V}_{\ker,s}$. Lemma 8.25 with $\mathcal{K} = \mathcal{A}_1$ yields

$$\begin{aligned} c\|\boldsymbol{D}_\tau\boldsymbol{u}_{\ker,n}\|^2_{\mathcal{H}_s} &+ \frac{1}{2\tau}\big(\langle\mathcal{A}_1 u_{\ker,n}, u_{\ker,n}\rangle - \langle\mathcal{A}_1 u_{\ker,n-1}, u_{\ker,n-1}\rangle\big) \\ &\le \langle\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_{\ker,n} + \mathcal{A}_1\boldsymbol{u}_{\ker,n}, \mathbf{B}\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_{\ker,n}\rangle = \langle\boldsymbol{f}_n + \mathcal{B}^-_{\mathcal{A}}\boldsymbol{g}_n - \mathcal{A}_2\boldsymbol{u}_{\ker,n}, \mathbf{B}\mathbf{A}^{-1}\boldsymbol{D}_\tau\boldsymbol{u}_{\ker,n}\rangle. \end{aligned}$$

Here, $c>0$ denotes the smallest eigenvalue of $\mathbf{A}^{-T}\mathbf{B}\mathbf{A}^{-1}$. As in the proof of Theorem 8.16, a multiplication by τ and a summation over all time steps leads to the estimate

$$\int_0^T \|\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}\|^2_{\mathcal{H}_s}\,\mathrm{d}t + \|u_{\mathrm{ker},N}\|^2_{\mathcal{A}_1} \lesssim \|u_{0,\mathrm{ker}}\|^2_{\mathcal{V}} + \sum_{n=1}^{N} \tau\|\mathbf{A}^{-T}\mathbf{B}(\boldsymbol{f}_n + \mathcal{B}^-_{\mathcal{A}}\dot{\boldsymbol{g}}_n - \mathcal{A}_2\boldsymbol{u}_{\mathrm{ker},n})\|^2_{\mathcal{H}_s^*}. \tag{8.38}$$

Since the right-hand side of (8.38) is bounded, so is $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}$ in $L^2(0,T;\mathcal{H}_s)$ and $u_{\mathrm{ker},N}$ in $\mathcal{V}_{\mathrm{ker}}$. By Remark 8.29 and Theorem 8.26, the weak limit of the whole sequence is given by $\mathbf{A}\dot{u}_{\mathrm{ker}}\mathbb{1}_s$ and $u_{\mathrm{ker}}(T)$, respectively. Thus, the strong convergence of $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} \to \mathbf{A}\dot{u}_{\mathrm{ker}}\mathbb{1}_s$ in $L^2(0,T;\mathcal{H}_s)$ follows by the estimate

$$\begin{aligned}
&\|\mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} - \dot{u}_{\mathrm{ker}}\mathbb{1}_s\|^2_{L^2(0,T;\mathcal{H}_s)}\\
&\lesssim \int_0^T \big(\mathbf{B}\big(\mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} - \dot{u}_{\mathrm{ker}}\mathbb{1}_s\big),\, \mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} - \dot{u}_{\mathrm{ker}}\mathbb{1}_s\big)_{\mathcal{H}_s}\,\mathrm{d}t\\
&= -\int_0^T \big\langle \mathcal{A}_1\mathbf{B}\mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}, \boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t + \int_0^T \big\langle \mathbf{B}\dot{u}_{\mathrm{ker}}\mathbb{1}_s,\, \mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} - \dot{u}_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t\\
&\quad + \int_0^T \big\langle \mathbf{B}\big(\boldsymbol{f}_\tau - f\mathbb{1}_s\big) - \mathcal{B}^-_{\mathcal{A}}\mathbf{B}\big(\dot{\boldsymbol{g}}_\tau - \dot{g}\mathbb{1}_s\big) - \mathcal{A}_2\mathbf{B}\big(\boldsymbol{u}_{\mathrm{ker},\tau} - u_{\mathrm{ker}}\mathbb{1}_s\big),\, \mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} - \dot{u}_{\mathrm{ker}}\mathbb{1}_s\big\rangle\,\mathrm{d}t
\end{aligned}$$

and arguments similar to the ones in Theorem 8.27 Step 3. This shows the strong convergence of $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathrm{ker},\tau} = \mathbf{b}^T\mathbf{A}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau}$ in $L^2(0,T;\mathcal{H})$. On the other hand, with the continuity of the operators

$$\begin{aligned}
\mathcal{B}^*\lambda_\tau &= \mathbf{b}^T\mathbf{A}^{-1}\left(\mathbf{A}^{-1}\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_{\mathrm{ker},\tau} + \mathcal{B}^-_{\mathcal{A}}\dot{\boldsymbol{g}}_\tau - \boldsymbol{f}_\tau\right) + \mathcal{A}u_\tau - \mathcal{B}^*\gamma_\tau\\
&\to \mathbf{b}^T\mathbf{A}^{-1}\left(\dot{u}_{\mathrm{ker}} + \mathcal{B}^-_{\mathcal{A}}\dot{g} - f\right)\mathbb{1}_s + \mathcal{A}u = \dot{u} - f + \mathcal{A}u \qquad \text{in } L^2(0,T;\mathcal{V}^*)
\end{aligned}$$

holds. As in the proof of Theorem 8.16, this results in the claimed convergence of λ_τ. □

Remark 8.31. The condition in Assumption 8.21 that the scheme has to be algebraically stable may be weakened. It is sufficient if a positive definite matrix $M \in \mathbb{R}^{s\times s}$ exists such that $\mathbf{M} := M\mathbf{A} + \mathbf{A}^TM^T - \mathbf{b}\mathbf{b}^T$ is positive semidefinite and $M^T\mathbb{1}_s = \mathbf{b}$.

8.4. Comments on Non-L-Stable Methods

Many proofs in the previous section 8.3 did not use the L-stability of the Runge-Kutta scheme. For this reason, we want to discuss in this section what happens if we drop the L-stability from Assumption 8.21. The other assumptions on the discretization scheme are still satisfied such that the considered Runge-Kutta methods fulfill the following assumptions.

Assumption 8.32. The Runge-Kutta method (5.1) is algebraically stable and $R(\infty) \in [-1,1]$. Furthermore, all weights $\mathbf{b}_i$ are positive with $\sum_{i=1}^s \mathbf{b}_i = \mathbb{1}_s^T\mathbf{b} = 1$.

Example 8.33. The Gauss–Legendre methods [HaiW96, p. 71 f.] satisfy Assumption 8.32; see [HaiNW93, Th. II.16.5] and [HaiW96, Th. IV.12.7]. In particular, these methods satisfy $R(\infty) = -1$ for odd and $R(\infty) = 1$ for even stage numbers [HaiW96, p. 227].

Since $R(\infty) = 1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s$ must not vanish under the weakened assumption 8.32, we need an approximation of the previous time step of the Lagrange multipliers λ_{n-1} and γ_{n-1} to calculate λ_n and γ_n, respectively; see Section 5.1. In particular, we need initial values for λ and γ. By Remark 8.7 we set $\gamma_0 := 0$. However, a well-defined initial value λ_0 would require continuous right-hand sides f, $\dot{g}$ and a regular u_0 with $\mathcal{A}u_0 \in \mathcal{H}^*_{\mathrm{ker}}$; cf. Lemma 9.23. Thus, we omit the calculation of λ_n and only set

$$u_n = R(\infty)u_{n-1} + \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{u}_n, \qquad \gamma_n = R(\infty)\gamma_{n-1} + \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\gamma}_n. \tag{8.39}$$

The internal stages are given by the system (8.27) as for the L-stable schemes, i.e., by

$$\begin{aligned} \mathbf{A}^{-1}\boldsymbol{D}_\tau \boldsymbol{u}_n + \mathcal{A}\boldsymbol{u}_n - \mathcal{B}^*\boldsymbol{\lambda}_n - \mathcal{B}^*\boldsymbol{\gamma}_n &= \boldsymbol{f}_n \quad \text{in } \mathcal{V}_s^*, && \text{(8.40a)}\\ \mathcal{B}\boldsymbol{u}_n + \mathcal{C}\boldsymbol{\gamma}_n &= \boldsymbol{g}_n \quad \text{in } \mathcal{Q}_s^*, && \text{(8.40b)}\\ \mathcal{B}\mathbf{A}^{-1}\boldsymbol{D}_\tau \boldsymbol{u}_n &= \dot{\boldsymbol{g}}_n \quad \text{in } \mathcal{Q}_s^*. && \text{(8.40c)} \end{aligned}$$

Lemma 8.23 guarantees the existence of unique internal stages $(\boldsymbol{u}_n, \boldsymbol{\lambda}_n, \boldsymbol{\gamma}_n) \in \mathcal{V}_s \times \mathcal{Q}_s \times \mathcal{Q}_s$. We point out that in contrast to the L-Stable methods, the approximation u_n is only an element of $\mathcal{H}$, if $u_0 \notin \mathcal{V}$. Thus, only the internal stage $\boldsymbol{u}_n$ has images in $\mathcal{V}$.

Before we investigate the Runge-Kutta methods for operator DAEs we study their application to the operator ODE (8.29). The temporal discretization then reads

$$\begin{aligned} v_{n+1} &= (1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)v_n + \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{v}_n, && \text{(8.41a)}\\ \mathbf{A}^{-1}\boldsymbol{D}_\tau \boldsymbol{v}_n + \mathcal{A}\boldsymbol{v}_n &= \boldsymbol{f}_n. && \text{(8.41b)} \end{aligned}$$

Lemma 8.34. *Let the assumption of Theorem 8.26 on f, v_0, and $\mathcal{A}$ be satisfied. Suppose that a Runge-Kutta method is given, which satisfies Assumption 8.32. Let the operator ODE (8.29) be discretized by (8.41) on $[0,T]$ with constant step size $\tau = T/N$, $N \in \mathbb{N}$. Assume that the piecewise constant function $\boldsymbol{f}_\tau$ defined by $\boldsymbol{f}_\tau|_{(t_{n-1},t_n]} = f_n$ converges strongly to $f\mathbb{1}_s$ in $L^2(0,T;\mathcal{V}_s^*)$.*

Then there exists a unique solution $v_n \in \mathcal{H}$ and $\boldsymbol{v}_n \in \mathcal{V}_s$ of system (8.41) for every time step $n = 1,\dots,N$. Furthermore, the piecewise constant functions v_τ, $\boldsymbol{v}_\tau$, and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{v}_\tau$ given by v_n, $\boldsymbol{v}_n$, and $D_\tau v_n$ at $(t_{n-1},t_n]$, $n = 1,\dots,N$, respectively, satisfy

$$\begin{aligned} v_\tau &\rightharpoonup v \quad \text{in } L^2(0,T;\mathcal{H}), & \boldsymbol{v}_\tau &\rightharpoonup v\mathbb{1}_s \quad \text{in } L^2(0,T;\mathcal{V}_s),\\ \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{v}_\tau &\rightharpoonup \dot{v} \quad \text{in } L^2(0,T;\mathcal{V}^*), & v_\tau(T) &\rightharpoonup v(T) \quad \text{in } \mathcal{H} \end{aligned}$$

as $\tau \to 0$, where v is the solution of the operator ODE (8.29).

Proof. We note that the proof of Theorem 8.26 only uses the algebraic stability of the Runge-Kutta method and the classical order of at least one. Thus, the assertion follows the lines of Theorem 8.26, where the approximation v_n is given by (8.41a) rather than by $\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{v}_n$. □

For the study of Runge-Kutta methods applied to the operator DAE (8.4) we define the piecewise functions u_τ, $\boldsymbol{u}_\tau$, $\widehat{u}_\tau$, $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau$, $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{u}}_\tau$, $\boldsymbol{\lambda}_\tau$, γ_τ, and $\boldsymbol{\gamma}_\tau$ as in Subsections 8.2.2 and 8.3.3.

We start by discussing the convergence behavior of the piecewise constant functions associated to γ. For $\boldsymbol{\gamma}_\tau$ we note that equation (8.33) still holds and therefore the arguments $\boldsymbol{\gamma}_\tau$ tending to zero in $L^2(0,T;\mathcal{Q}_s)$ are still valid. Furthermore, $\gamma_0 = 0$ implies

$$\begin{aligned} \mathcal{C}\gamma_n &\overset{(8.39)}{=} R^n(\infty)\mathcal{C}\gamma_0 + \sum_{k=0}^{n-1} R^k(\infty)\mathbf{b}^T\mathbf{A}^{-1}\mathcal{C}\boldsymbol{\gamma}_{n-k} && \text{(8.42)}\\ &\overset{(8.33)}{=} \sum_{k=0}^{n-1} R^k(\infty)\Big(\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s \int_0^{t_{n-k}} \dot{g} - \mathbf{b}^T\dot{\boldsymbol{g}}_\tau \,\mathrm{d}t + \mathbf{b}^T\mathbf{A}^{-1}\big(\boldsymbol{g}_{n-k} - g(t_{n-k})\mathbb{1}_s\big) - \tau R(\infty)\mathbf{b}^T\dot{\boldsymbol{g}}_{n-k}\Big). \end{aligned}$$

Thus, for $R(\infty) \in (-1,1) \setminus \{0\}$ we observe by the geometric series $\sum_{k=0}^\infty |R^k(\infty)| = \frac{1}{1-|R(\infty)|}$ and $|R(\infty)| < 1$ that

$$\|\gamma_n\|_{\mathcal{Q}} \overset{(8.42)}{\lesssim} \sum_{k=0}^{n-1} |R^k(\infty)|\Big(\int_0^{t_n} \|\dot{g} - \mathbf{b}^T\dot{\boldsymbol{g}}_\tau\|_{\mathcal{Q}^*}\,\mathrm{d}t + \max_{i=1,\dots,n} \|\boldsymbol{g}_i - g(t_i)\mathbb{1}_s\| + \int_{t_{n-k-1}}^{t_{n-k}} \|\mathbf{b}^T\dot{\boldsymbol{g}}_\tau\|_{\mathcal{Q}^*}\,\mathrm{d}t\Big)$$

$$\lesssim \sqrt{T}\|\dot{g} - \mathbf{b}^T\dot{\boldsymbol{g}}_\tau\|_{L^2(0,T;\mathcal{Q}^*)} + \|\boldsymbol{g}_\tau - g\mathbb{1}_s\|_{L^\infty(0,T;\mathcal{Q}_s^*)} + \max_{i=1,\dots,N}\int_{t_{i-1}}^{t_i}\|\dot{g}\|_{\mathcal{Q}^*}\,\mathrm{d}t.$$

Since the right-hand side is independent of n and vanishes as $\tau \to 0$ by [KufJF77, Cor. 2.19.10] and Assumption 8.19, we have $\lim_{\tau\to 0}\gamma_\tau = 0$ in $L^\infty(0,T;\mathcal{Q})$.

For $R(\infty) = -1$ and even n we have

$$\begin{aligned}
\mathcal{C}\gamma_n &\overset{(8.42)}{=} \sum_{k=1}^{n}(-1)^k\Big(2\int_0^{t_k}\dot{g} - \mathbf{b}^T\dot{\boldsymbol{g}}_\tau\,\mathrm{d}t + \mathbf{b}^T\mathbf{A}^{-1}\big(\boldsymbol{g}_k - g(t_k)\mathbb{1}_s\big) + \tau\mathbf{b}^T\dot{\boldsymbol{g}}_k\Big)\\
&= \sum_{k=1}^{n/2} 2\int_{t_{2k-1}}^{t_{2k}}\dot{g} - \mathbf{b}^T\dot{\boldsymbol{g}}_\tau\,\mathrm{d}t + \mathbf{b}^T\mathbf{A}^{-1}\big(\boldsymbol{g}_{2k} - \boldsymbol{g}_{2k-1}\big) - 2(g(t_{2k}) - g(t_{2k-1})) + \tau\mathbf{b}^T\big(\dot{\boldsymbol{g}}_{2k} - \dot{\boldsymbol{g}}_{2k-1}\big)\\
&= \sum_{k=1}^{n/2}\mathbf{b}^T\mathbf{A}^{-1}\big(\boldsymbol{g}_{2k} - \boldsymbol{g}_{2k-1}\big) - \tau\mathbf{b}^T\big(\dot{\boldsymbol{g}}_{2k} + \dot{\boldsymbol{g}}_{2k-1}\big).
\end{aligned} \tag{8.43}$$

This expression does not vanish uniformly for every $\boldsymbol{g}_\tau$ and $\dot{\boldsymbol{g}}_\tau$, which satisfy Assumption 8.19. A counterexample is given by $g(t) = t^2$ with approximations $\boldsymbol{g}_k = t_k^2\mathbb{1}_s$, and $\dot{\boldsymbol{g}}_k = 2t_k\mathbb{1}_s$ if k is even and $\dot{\boldsymbol{g}}_k = 2t_{k-1}\mathbb{1}_s$ otherwise, $k = 1,\dots,N$. Since we want to consider Runge-Kutta methods and data under minimal assumptions, we choose a specific approximation. We set

$$\boldsymbol{g}_n \coloneqq g(t_n)\mathbb{1}_s + \tau\mathbf{A}\dot{\boldsymbol{g}}_n, \qquad n = 1,\dots,N. \tag{8.44}$$

Then $\boldsymbol{g}_n$, $n = 1,\dots,N$, fulfills Assumption 8.19 by the uniformly continuity of g and [KufJF77, Cor. 2.19.10]. Furthermore, for the limit of γ_τ we consider for every arbitrary but fixed $\varepsilon > 0$ a function $\varphi_\varepsilon \in C([0,T],\mathcal{Q}^*)$ with $\|\dot{g} - \varphi_\varepsilon\|_{L^2(0,T;\mathcal{Q}^*)} < \varepsilon$; see Theorem 3.32.ii). Then we observe that

$$\begin{aligned}
&\max_{n=2,4,\dots,N}\|\gamma_n\|_{\mathcal{Q}}\\
&\overset{(8.43)}{\lesssim} \max_{n=2,4,\dots,N}\sum_{k=1}^{n/2}\|\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s\big(g(t_{2k}) - g(t_{2k-1})\big) - \tau 2\mathbf{b}^T\dot{\boldsymbol{g}}_{2k-1}\|_{\mathcal{Q}^*}\\
&= \max_{n=2,4,\dots,N} 2\sum_{k=1}^{n/2}\left\|\int_{t_{2k-1}}^{t_{2k}}\dot{g}\,\mathrm{d}t - \int_{t_{2k-2}}^{t_{2k-1}}\dot{g}\,\mathrm{d}t + \mathbf{b}^T\int_{t_{2k-2}}^{t_{2k-1}}\dot{g}\mathbb{1}_s - \dot{\boldsymbol{g}}_\tau\,\mathrm{d}t\right\|_{\mathcal{Q}^*}\\
&\lesssim \max_{n=2,4,\dots,N} 2\sum_{k=1}^{n/2}\int_{t_{2k-2}}^{t_{2k}}\|\dot{g} - \varphi_\varepsilon\|_{\mathcal{Q}^*} + \|\dot{g}\mathbb{1}_s - \dot{\boldsymbol{g}}_\tau\|_{\mathcal{Q}_s^*}\,\mathrm{d}t + \int_{t_{2k-1}}^{t_{2k}}\|\varphi_\varepsilon(t) - \varphi_\varepsilon(t-\tau)\|_{\mathcal{Q}^*}\,\mathrm{d}t\\
&\le 2\int_0^T\|\dot{g} - \varphi_\varepsilon\|_{\mathcal{Q}^*} + \|\dot{g}\mathbb{1}_s - \dot{\boldsymbol{g}}_\tau\|_{\mathcal{Q}_s^*}\,\mathrm{d}t + T\max_{t\in[\tau,T]}\|\varphi_\varepsilon(t) - \varphi_\varepsilon(t-\tau)\|_{\mathcal{Q}^*}\\
&\to 2\sqrt{T}\varepsilon
\end{aligned}$$

as $\tau \to 0$. Here, we used that φ_ε is uniformly continuous on the interval $[0,T]$. In comparison to even n, for odd n the sum (8.43) has an additional summand, which vanishes as $\tau \to 0$ by Assumption 8.19. Since $\varepsilon > 0$ was arbitrary, this shows $\|\gamma_\tau\|_{L^\infty(0,T;\mathcal{Q}^*)} \to 0$ for $R(\infty) = -1$.

Finally, for $R(\infty) = 1$ we have $C\gamma_n = \sum_{k=1}^n\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{g}_k - \tau\mathbf{b}^T\dot{\boldsymbol{g}}_k$ by $\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s = 1 - R(\infty) = 0$ and (8.42). This term does not vanish in general. A counterexample is the same as for the case $R(\infty) = -1$. However, by the approximation (8.44) we have $C\gamma_n = 0$ and therefore $\gamma_\tau = 0$.

With the convergence of γ_τ we can now investigate the approximations of the state u. As a start, it follows $\boldsymbol{u}_{c,\tau} \to u_c\mathbb{1}_s$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{c,\tau} \to \dot{u}_c$ in the associated L^2-spaces by the convergence of $\boldsymbol{\gamma}_\tau$ and arguments similar to Theorem 8.27 Step 2. For the piecewise linear approximations we observe

$\widehat{u}_{c,\tau} \to u_c$ in $L^\infty(0,T;\mathcal{V})$ by (8.23) and $u_{c,\tau} \to u_c$ in $L^2(0,T;\mathcal{V})$ by

$$\lim_{\tau\to 0} \|\widehat{u}_{c,\tau} - u_{c,\tau}\|^2_{L^2(0,T;\mathcal{V})} = \sum_{n=1}^{N} \frac{\tau^3}{3} \|\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{D}_\tau \boldsymbol{u}_{c,n}\|^2_{\mathcal{V}} \overset{(8.41)}{=} \lim_{\tau\to 0} \frac{\tau^2}{3} \int_0^T \|\mathbf{b}^T\mathcal{B}^-_{\mathcal{A}}\dot{\boldsymbol{g}}_\tau\|^2_{\mathcal{Q}^*_s} \,\mathrm{d}t = 0.$$

For the part with images in $\mathcal{V}_{\ker}$, we test equation (8.40a) by $\mathbf{B}\boldsymbol{u}_{\ker,n}$. Then we get by Lemma 8.25 the inequality $\|u_{\ker,n}\|^2_{\mathcal{H}} + \tau\mu_{\mathcal{A}} \sum_{i=1}^n \|\boldsymbol{u}_{\ker,i}\|^2_{\mathcal{V}_s} \leq \|u_{\ker,0}\|^2_{\mathcal{H}} + \tau \sum_{i=1}^n \|\boldsymbol{f}_i\|^2_{\mathcal{V}^*_s} + \|\mathcal{B}^-_{\mathcal{A}}\dot{\boldsymbol{g}}_i\|^2_{\mathcal{H}^*_s}$ for every $n = 1,\dots,N$. Thus, we can use the arguments of Theorem 8.27 Step 3 to prove $\boldsymbol{u}_{\ker,\tau} \to u_{\ker}\mathbb{1}_s$ in $L^2(0,T;\mathcal{V}_s)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\ker,\tau} \to \dot{u}_{\ker}$ in $L^2(0,T;\mathcal{V}^*_{\ker})$. In addition, we have

$$\begin{aligned}\int_0^T \|\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s u_{\ker,\tau} - \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{u}_{\ker,\tau}\|^2_{\mathcal{H}} &= \tau\sum_{n=1}^N \|\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s u_{\ker,n} - \mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{u}_{\ker,n}\|^2_{\mathcal{H}} \\ &\overset{(8.39)}{=} (1-\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s)^2\tau\sum_{n=1}^N \|u_{\ker,n} - u_{\ker,n-1}\|^2_{\mathcal{H}},\end{aligned}$$

where the right-hand side vanishes as $\tau \to 0$ by Theorem 8.27 Step 4. Note that this implies $u_{\ker,\tau} \to u_{\ker}$ in $L^2(0,T;\mathcal{H})$ by the strong convergence of $\boldsymbol{u}_{\ker,\tau}$, if $\mathbf{b}^T\mathbf{A}^{-1}\mathbb{1}_s \neq 0$, i.e., $R(\infty) \in [-1,1)$.

Finally, Theorem 8.27 Step 4 and Step 5 show $\widehat{u}_{\ker,\tau} \to u$ in $L^2(0,T;\mathcal{H})$ and $\mathbf{b}^T \int_0^\cdot \boldsymbol{\lambda}_\tau \,\mathrm{d}s \to \Lambda$ in $L^2(0,T;\mathcal{Q})$ for $R(\infty) \in [-1,1)$. For $R(\infty) = 1$ the functions $u_{\ker,\tau}$, $\widehat{u}_{\ker,\tau}$, and $\mathbf{b}^T \int_0^\cdot \boldsymbol{\lambda}_\tau \,\mathrm{d}s$ converge weakly by Lemma 8.34 and the steps of the proof of Theorem 8.27.

We summarize our observations in the following Theorem.

Theorem 8.35 (Convergence of Runge-Kutta Schemes). *Let the approximations $\boldsymbol{f}_\tau$, $\boldsymbol{g}_\tau$, and $\dot{\boldsymbol{g}}_\tau$ of the right-hand sides $f \in L^2(0,T;\mathcal{V}^*)$, $g \in H^1(0,T;\mathcal{Q}^*)$ satisfy Assumption 8.19. Suppose that Assumptions 8.1.i), iii) on the operators and the initial value are fulfilled. The corresponding solution of the operator DAE (8.4) is denoted by $(u,\lambda,0)$. Suppose that the Runge-Kutta scheme satisfies Assumption 8.32.*

If $R(\infty) \in (-1,1)$, then we have

$$u_\tau, \widehat{u}_\tau \to u \quad \text{in } L^2(0,T;\mathcal{H}), \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau \to \dot{u} \quad \text{in } L^2(0,T;\mathcal{V}^*_{\ker}), \qquad \gamma_\tau \to 0 \quad \text{in } L^\infty(0,T;\mathcal{Q})$$

for $\tau \to 0$. The function $\int_0^\cdot \mathbf{b}^T\boldsymbol{\lambda}_\tau(s)\,\mathrm{d}s$ converges strongly to Λ in $L^2(0,T;\mathcal{Q})$. If we replace $\boldsymbol{g}_n$ by (8.44), then for $R(\infty) = -1$ the statements still hold and for $R(\infty) = 1$ the convergences of u_τ, $\widehat{u}_\tau$, and $\int_0^\cdot \mathbf{b}^T\boldsymbol{\lambda}_\tau \,\mathrm{d}s$ are weak.

Proof. The assertions are proven by the previous discussion in this section 8.4. □

Remark 8.36. If the data is more regular in the sense of Theorem 8.30, then we have

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau \to \dot{u} \quad \text{in } L^2(0,T;\mathcal{H}), \qquad \lambda_\tau \to \lambda \quad \text{in } L^2(0,T;\mathcal{Q})$$

under the assumptions of Theorem 8.35. This follows along the lines of Theorem 8.30.

8.5. Convergence Order

In the previous sections 8.2–8.4 we studied the qualitative behavior of the convergence of the discrete solution. In this section we analyze the convergence order of the Runge-Kutta methods applied to the operator DAE (8.4). For this we combine convergence results for operator ODEs [LubO95b, Ch. 1] and for DAEs; see Theorem 5.4. Following [LubO95b] we restrict our investigation to regular solutions and Runge-Kutta methods which satisfy Assumption 8.32 and $|R(\infty)| < 1$.

Before we formulate the main result of this section, we recall the order conditions for the Runge-Kutta method (5.1), namely

$$\sum_{i=1}^{s} \mathbf{b}_i \mathbf{c}_i^{k-1} = \frac{1}{k}, \qquad k = 1, \dots, \mathsf{p}, \tag{8.45a}$$

$$\sum_{j=1}^{s} \mathbf{A}_{ij} \mathbf{c}_j^{k-1} = \frac{\mathbf{c}_i^k}{k}, \quad i = 1, \dots, s, \; k = 1, \dots, \mathsf{q}; \tag{8.45b}$$

see Section 5.1. Note that every Runge-Kutta method of classical order p satisfies (8.45a) [HaiNW93, Th. II.2.13 & p. 208]. We can now state the first theorem on the convergence order.

Theorem 8.37 (Convergence Order for Runge-Kutta Methods)**.** *Suppose that the Runge-Kutta scheme* (5.1) *satisfies Assumption 8.32, fulfills* $|R(\infty)| < 1$*, has stage order* q*, and has classical order* $\mathsf{p} \geq \mathsf{q}+1$*. Let Assumption 8.1.i) on the operators be fulfilled. Suppose that* $u_0 \in \mathcal{V}$ *is consistent, i.e.,* $\mathcal{B}u_0 = g(0)$*. Assume that the right-hand sides* f *and* g *as well as the solution* (u, λ) *of the operator DAE* (8.1) *are sufficiently regular. In particular, let* $\lambda_0 := \lambda(0) \in \mathcal{Q}$ *for non-L-stable Runge-Kutta methods be well-defined. Furthermore, suppose that the solution and the right-hand sides are extendable outside of* $[0,T]$ *by maintaining its regularity. Assume that the norm (in the associated space of regular functions) of the respective extension is bounded by a multiple of the norm of the associated function on* $[0,T]$*.*

Then the approximations $\{(u_n, \lambda_n, \gamma_n)\}_{n=1,\dots,N} \subset \mathcal{V} \times \mathcal{Q} \times \mathcal{Q}$ *given by the temporal discretization* (8.39), (8.40) *of the operator DAE* (8.4) *with* $\boldsymbol{f}_{n,i} = f(t_n + \tau \mathbf{c}_i)$*,* $\boldsymbol{g}_{n,i} = g(t_n + \tau \mathbf{c}_i)$*, and* $\dot{\boldsymbol{g}}_{n,i} = \dot{g}(t_n + \tau \mathbf{c}_i)$ *satisfy*

$$\max_{n=1,\dots,N} \|u_n - u(t_n)\|_{\mathcal{H}}^2 + \tau \sum_{n=1}^{N} \|u_n - u(t_n)\|_{\mathcal{V}}^2$$
$$\lesssim \tau^{2\mathsf{q}+2} \int_0^T \|u^{(\mathsf{q}+1)}\|_{\mathcal{V}}^2 + \|u_{\ker}^{(\mathsf{q}+2)}\|_{\mathcal{V}_{\ker}^*}^2 + \|g^{(\mathsf{q}+2)}\|_{\mathcal{Q}^*}^2 \,\mathrm{d}t, \tag{8.46a}$$

$$\tau \sum_{n=1}^{N} \|\lambda_n - \lambda(t_n)\|_{\mathcal{Q}}^2 \lesssim \tau^{2\mathsf{q}+1} \int_0^T \|u^{(\mathsf{q}+1)}\|_{\mathcal{V}}^2 + \|\lambda^{(\mathsf{q}+1)}\|_{\mathcal{Q}}^2 + \|u_{\ker}^{(\mathsf{q}+2)}\|_{\mathcal{V}_{\ker}^*}^2 + \|g^{(\mathsf{q}+2)}\|_{\mathcal{Q}^*}^2 \,\mathrm{d}t, \tag{8.46b}$$

$$\tau \sum_{n=1}^{N} \|\gamma_n\|_{\mathcal{Q}}^2 \lesssim \tau^{2\mathsf{k}} \int_0^T \|g^{(\mathsf{k})}\|_{\mathcal{Q}^*}^2 + \|g^{(\mathsf{k}+1)}\|_{\mathcal{Q}^*}^2 \,\mathrm{d}t. \tag{8.46c}$$

The integer k *in* (8.46c) *is given by* p *if the Runge-Kutta method is stiffly accurate and by* $\mathsf{q}+1$ *otherwise. All constants which are suppressed by* $\lesssim$ *depend on the coefficients* $\mathbf{A}$*,* $\mathbf{b}$*, and* $\mathbf{c}$ *of the Runge-Kutta method, on the operators* $\mathcal{A}$ *and* $\mathcal{B}$*, on* T*, as well as the factor of the possible extension.*

Proof. For the sake of brevity, we introduce the errors

$$\Delta u_n := u_n - u(t_n), \qquad \Delta \boldsymbol{u}_n := \boldsymbol{u}_n - \begin{bmatrix} u(t_{n-1} + \tau \mathbf{c}_1) \\ \vdots \\ u(t_{n-1} + \tau \mathbf{c}_s) \end{bmatrix}, \qquad \Delta \dot{\boldsymbol{u}}_n := \mathbf{A}^{-1} \boldsymbol{D}_\tau \boldsymbol{u}_n - \begin{bmatrix} \dot{u}(t_{n-1} + \tau \mathbf{c}_1) \\ \vdots \\ \dot{u}(t_{n-1} + \tau \mathbf{c}_s) \end{bmatrix}.$$

Analogously, we introduce the error for $u_{\ker}$, u_{c}, and λ, where we use again the direct sum $\mathcal{V} = \mathcal{V}_{\ker} \oplus \mathcal{V}_{\mathrm{c}}$ with $\mathcal{V}_{\mathrm{c}} := \{v \in \mathcal{V} \,|\, \mathcal{A}v \in \mathcal{V}_{\ker}^0\}$. Note that $\Delta\gamma_n = \gamma_n$ holds by $\gamma = 0$. Furthermore, we assume $\mathbf{c}_i \in [0,1]$, $i = 1, \dots, s$, for the simplicity of this proof. For Runge-Kutta methods with a $\mathbf{c}_i \notin [0,1]$ one uses the assumed extension of the solution and the right-hand sides.

The proof is split into four parts. We start by estimating the error for u_{c} and γ. Afterwards we bound $\Delta u_{\ker,n}$. We finish the proof with the error $\Delta\lambda_n$.

Step 1 (Error for u_c): For the part in $\mathcal{V}_c$, we obtain by $\mathcal{B}u_{c,0} = g(0)$ and several Taylor expansions

$$\begin{aligned}\mathcal{B}\Delta u_{c,n} &\overset{(8.39)}{=} \mathcal{B}u_{c,n-1} + \tau\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{D}_\tau\mathcal{B}\boldsymbol{u}_{c,n} - g(t_n)\\ &\overset{(8.40c)}{=} \mathcal{B}\Delta u_{c,n-1} + \tau\sum_{i=1}^{s}\mathbf{b}_i\dot{g}(t_{n-1}+\tau\mathbf{c}_i) - \big(g(t_n) - g(t_{n-1})\big)\\ &= \mathcal{B}\Delta u_{c,n-1} + \sum_{k=0}^{\overline{\mathsf{p}}-1}\tau^{k+1}\Big(\sum_{i=1}^{s}\frac{\mathbf{b}_i\mathbf{c}_i^k}{k!} - \frac{1}{(k+1)!}\Big)g^{(k+1)}(t_{n-1}) + R_n\\ &\overset{(8.45a)}{=} \mathcal{B}\Delta u_{c,n-1} + R_n \end{aligned} \tag{8.47}$$

for every $\overline{\mathsf{p}} = 1, \dots, \mathsf{p}$ with the remainder

$$R_n = \tau\sum_{i=1}^{s}\mathbf{b}_i\int_{t_{n-1}}^{t_{n-1}+\tau\mathbf{c}_i}\frac{(t_{n-1}+\tau\mathbf{c}_i - t)^{\overline{\mathsf{p}}-1}}{(\overline{\mathsf{p}}-1)!}g^{(\overline{\mathsf{p}}+1)}(t)\,\mathrm{d}t - \int_{t_{n-1}}^{t_n}\frac{(t_n - t)^{\overline{\mathsf{p}}}}{\overline{\mathsf{p}}!}g^{(\overline{\mathsf{p}}+1)}(t)\,\mathrm{d}t.$$

Note that $0 \le \mathbf{c}_i \le 1$ for all $i = 1, \dots, s$ and a successive application of (8.47) implies the estimate

$$\|\Delta u_{c,n}\|_{\mathcal{V}}^2 \lesssim \Big\|\sum_{i=1}^{n}R_i\Big\|_{\mathcal{Q}^*}^2 \le N\sum_{i=1}^{N}\|R_i\|_{\mathcal{Q}^*}^2 \lesssim \tau^{2\overline{\mathsf{p}}}\int_0^T\|g^{(\overline{\mathsf{p}}+1)}\|_{\mathcal{Q}^*}^2\,\mathrm{d}t. \tag{8.48}$$

We emphasize that the included constant only depends on the Runge-Kutta method, $\mathcal{B}_{\mathcal{A}}^-$, and T. Furthermore, with analogous arguments we have for the internal stage $\boldsymbol{u}_{c,n}$ that

$$\begin{aligned}\mathcal{B}\Delta\boldsymbol{u}_{c,n,i} &\overset{(8.40c)}{=} \mathcal{B}u_{c,n} + \tau\sum_{j=1}^{s}\mathbf{A}_{ij}\dot{g}(t_{n-1}+\tau\mathbf{c}_j) - g(t_{n-1}+\tau\mathbf{c}_i)\\ &= \mathcal{B}\Delta u_{c,n-1} + \sum_{k=0}^{\mathsf{q}-1}\tau^{k+1}\Big(\sum_{j=1}^{s}\frac{\mathbf{A}_{ij}\mathbf{c}_j^k}{k!} - \frac{\mathbf{c}_i^{k+1}}{(k+1)!}\Big)g^{(k+1)}(t_{n-1}) + \boldsymbol{R}_{n,i}\\ &\overset{(8.45b)}{=} \mathcal{B}\Delta u_{c,n-1} + \boldsymbol{R}_{n,i},\end{aligned}$$

where the correction term $\boldsymbol{R}_n \in \mathcal{Q}_s^*$ satisfies $\|\boldsymbol{R}_n\|_{\mathcal{Q}_s^*}^2 \lesssim \tau^{2\mathsf{q}+1}\int_{t_{n-1}}^{t_n}\|g^{(\mathsf{q}+1)}\|_{\mathcal{Q}^*}^2\,\mathrm{d}t$. This equality and inequality (8.48) with $\overline{\mathsf{p}} = \mathsf{q}+1$ lead to the bound

$$\tau\sum_{n=1}^{N}\|\Delta\boldsymbol{u}_{c,n}\|_{\mathcal{V}_s}^2 \lesssim \tau^{2\mathsf{q}+2}\int_0^T\|g^{(\mathsf{q}+1)}\|_{\mathcal{Q}^*}^2 + \|g^{(\mathsf{q}+2)}\|_{\mathcal{Q}^*}^2\,\mathrm{d}t. \tag{8.49}$$

Step 2 (Error for γ): We observe $\boldsymbol{\gamma}_n = -\mathcal{C}^{-1}\mathcal{B}\Delta\boldsymbol{u}_{c,n}$ by (8.40b) and (8.4b). This implies (8.46c) for stiffly accurate methods, since $\gamma_n = \boldsymbol{\gamma}_{n,s} = -\mathcal{C}^{-1}\mathcal{B}\Delta u_{c,n}$. For non-stiffly-accurate schemes we note that by Young's inequality (3.8) with $\varepsilon = \frac{|R(\infty)|}{1-R^2(\infty)} > 0$ the estimate

$$\|\gamma_n\|_{\mathcal{Q}}^2 \le \frac{1+R^2(\infty)}{2}\|\gamma_{n-1}\|_{\mathcal{Q}}^2 + \frac{2-R^2(\infty)}{2-2R^2(\infty)}\|\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\gamma}_n\|_{\mathcal{Q}}^2$$

is satisfied. This inequality implies (8.46c) by (8.49), since $\frac{1}{2}(1+R^2(\infty))<1$ holds and thus

$$\tau\sum_{n=1}^{N}\|\gamma_n\|_{\mathcal{Q}}^2\lesssim\tau\sum_{n=1}^{N}\|\mathbf{b}^T\mathbf{A}^{-1}\boldsymbol{\gamma}_n\|_{\mathcal{Q}}^2\lesssim\tau\sum_{n=1}^{N}\|\Delta\boldsymbol{u}_{\mathrm{c},n}\|_{\mathcal{V}_s}^2.$$

Step 3 (Error for u_{ker}): As in the proof of Theorem 8.27, we use that $u_{\mathrm{ker},n}$ and $\boldsymbol{u}_{\mathrm{ker},n}$ are also the solutions of the temporal discretization of the operator ODE (8.20) by the Runge-Kutta scheme (5.1). By [LubO95b, Th. 1.1] we then have

$$\max_{n=1,\dots,N}\|\Delta u_{\mathrm{ker},n}\|_{\mathcal{H}_{\mathrm{ker}}}^2+\tau\sum_{n=1}^{N}\|\Delta u_{\mathrm{ker},n}\|_{\mathcal{V}_{\mathrm{ker}}}^2\lesssim\tau^{2\mathsf{q}+2}\int_0^T\|u_{\mathrm{ker}}^{(\mathsf{q}+1)}\|_{\mathcal{V}_{\mathrm{ker}}}^2+\|u_{\mathrm{ker}}^{(\mathsf{q}+2)}\|_{\mathcal{V}_{\mathrm{ker}}^*}^2\,\mathrm{d}t. \tag{8.50}$$

This estimate and (8.48) imply the error bound (8.46a). In addition to (8.50), the proof of [LubO95b, Th. 1.1] shows for the internal stages $\boldsymbol{u}_{\mathrm{ker},n}$ that

$$\tau\sum_{n=1}^{N}\|\Delta\boldsymbol{u}_{\mathrm{ker},n}\|_{\mathcal{V}_s}^2\lesssim\tau^{2\mathsf{q}+2}\int_0^T\|u_{\mathrm{ker}}^{(\mathsf{q}+1)}\|_{\mathcal{V}}^2+\|u_{\mathrm{ker}}^{(\mathsf{q}+2)}\|_{\mathcal{V}_{\mathrm{ker}}^*}^2\,\mathrm{d}t. \tag{8.51}$$

Furthermore, on the one hand the authors in [LubO95b, p. 606] prove $\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}=\tau^{-1}\mathbf{A}^{-1}(\Delta\boldsymbol{u}_{\mathrm{ker},n}-\Delta u_{\mathrm{ker},n-1}\mathbb{1}_s+\widetilde{\boldsymbol{R}}_n)$ where the remainder $\widetilde{\boldsymbol{R}}_n$ fulfills an estimate analogous to (8.51). On the other hand, we have $\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}=-\mathcal{A}\Delta\boldsymbol{u}_{\mathrm{ker},n}$ in $\mathcal{V}_{\mathrm{ker},s}^*$. Together these equations for $\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}$ lead to

$$\begin{aligned}\tau\sum_{n=1}^{N}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|_{\mathcal{V}_s^*}^2&\lesssim\tau\sum_{n=1}^{N}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|_{\mathcal{H}_s}^2\\&\overset{(3.8)}{\le}\tau\sum_{n=1}^{N}\frac{1}{2\tau}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|_{\mathcal{V}_{\mathrm{ker},s}^*}^2+\frac{\tau}{2}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|_{\mathcal{V}_{\mathrm{ker},s}}^2\\&\lesssim\tau\sum_{n=1}^{N}\frac{1}{2\tau}\|\Delta\boldsymbol{u}_{\mathrm{ker},n}\|_{\mathcal{V}_s}^2+\frac{1}{2\tau}\big(\|\Delta\boldsymbol{u}_{\mathrm{ker},n}\|_{\mathcal{V}_s}^2+\|\Delta u_{\mathrm{ker},n-1}\|_{\mathcal{V}}^2+\|\widetilde{\boldsymbol{R}}_n\|_{\mathcal{V}_s}^2\big)\\&\lesssim\tau^{2\mathsf{q}+1}\int_0^T\|u_{\mathrm{ker}}^{(\mathsf{q}+1)}\|_{\mathcal{V}}^2+\|u_{\mathrm{ker}}^{(\mathsf{q}+2)}\|_{\mathcal{V}_{\mathrm{ker}}^*}^2\,\mathrm{d}t.\end{aligned} \tag{8.52}$$

In the last line we used (8.50), (8.51), and the estimate for $\widetilde{\boldsymbol{R}}_n$.

Step 4 (Error for λ): We start by observing for every $k=1,\dots,\mathsf{q}$ that

$$\sum_{i=1}^{s}\sum_{j=1}^{s}\mathbf{b}_i(\mathbf{A}^{-1})_{ij}\mathbf{c}_j^k\overset{(8.45b)}{=}k\sum_{i=1}^{s}\mathbf{b}_i\mathbf{c}_i^{k-1}\overset{(8.45a)}{=}1$$

holds. Using this equality and several Taylor expansions one shows $\Delta\lambda_n=R(\infty)\Delta\lambda_{n-1}+\mathbf{b}^T\mathbf{A}^{-1}\Delta\boldsymbol{\lambda}_n+\widehat{R}_n$, where the remainder satisfies $\|\widehat{R}_n\|_{\mathcal{Q}}^2\lesssim\tau^{2\mathsf{q}+1}\int_{t_{n-1}}^{t_n}\|\lambda^{(\mathsf{q}+1)}\|_{\mathcal{Q}}^2\,\mathrm{d}t$. Thus, we get the estimate

$$\tau\sum_{n=1}^{N}\|\Delta\lambda_n\|_{\mathcal{Q}}^2\lesssim\tau\sum_{n=1}^{N}\|\mathbf{b}^T\mathbf{A}^{-1}\Delta\boldsymbol{\lambda}_n\|_{\mathcal{Q}}^2+\|\widehat{R}_n\|_{\mathcal{Q}}^2\lesssim\tau\sum_{n=1}^{N}\|\mathbf{b}^T\mathbf{A}^{-1}\Delta\boldsymbol{\lambda}_n\|_{\mathcal{Q}}^2+\tau^{2\mathsf{q}+2}\int_0^T\|\lambda^{(\mathsf{q}+1)}\|_{\mathcal{Q}}^2\,\mathrm{d}t \tag{8.53}$$

by following Step 2. For the first term of the right-hand side of (8.53), we note that the error of the internal stages $\boldsymbol{\lambda}_n$ satisfies $\mathcal{B}^*\Delta\boldsymbol{\lambda}_n = \Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n} + \mathcal{A}\Delta\boldsymbol{u}_n + \mathcal{B}^*\mathcal{C}^{-1}\mathcal{B}\Delta\boldsymbol{u}_{\mathrm{c},n}$ in $\mathcal{V}^*$ by (8.40) and (8.4). Finally, inequality (8.46b) follows by the estimates (8.49) and (8.51)–(8.53). □

Remark 8.38. Theorem 1.1 in [LubO95b], which states error bounds for operator ODEs and which we used in the proof of Theorem 8.37, is proven by energy estimates. Using techniques similar to the method of semigroups the error estimates can be sharpened; see [GonO99; LubO93; LubO95b, Ch. 3 each]. These results would also refine (8.46a). However, this approach requires a deeper knowledge of the underlying constrained PDE, since the convergence order then also depends on the spatial regularity of the solution as well as on the type of boundary condition; see [LubO93, p. 116 f.] and [LubO95b, p. 616 f.]. For a certain class of constrained PDEs this was investigated in [Deb04; DebS05].

Remark 8.39. The convergence rate for the Lagrange multiplier λ in Theorem 8.37 can be refined to $\mathsf{q} + \theta$ if the interpolation space $\left[\mathcal{V}_{\mathrm{ker}}, \mathcal{V}_{\mathrm{ker}}^*\right]_\theta$ is embeddable into $\mathcal{V}^*$. Here, we use the notation from [LioM72, Ch. 1, Sec. 2], where the theory of interpolation between separable Hilbert spaces is discussed in detail. For the interpolation between general Banach spaces we refer to [BerL76]. However, by the assumed embedding we have

$$\begin{aligned}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|_{\mathcal{V}_s^*}^2 \lesssim \|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|^2_{[\mathcal{V}_{\mathrm{ker},s},\mathcal{V}^*_{\mathrm{ker},s}]_\theta} &\lesssim \|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|^{2-2\theta}_{\mathcal{V}_{\mathrm{ker},s}} \|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|^{2\theta}_{\mathcal{V}^*_{\mathrm{ker},s}} \\ &\lesssim \tau^{2\theta}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|^2_{\mathcal{V}_{\mathrm{ker},s}} + \tau^{2\theta-2}\|\Delta\dot{\boldsymbol{u}}_{\mathrm{ker},n}\|^2_{\mathcal{V}^*_{\mathrm{ker},s}},\end{aligned} \tag{8.54}$$

where we used [LioM72, Ch. 1, Prop. 2.3] and [Emm04, Th. A.1.4]. All constants are independent of τ. By the estimate (8.54) we can prove the refined convergence order $\mathsf{q} + \theta$ following the steps of (8.52) and Theorem 8.37 Step 4.

We point out that we showed the specific case $\theta = 1/2$ in Theorem 8.37, since the embedding $\left[\mathcal{V}_{\mathrm{ker}}, \mathcal{V}_{\mathrm{ker}}^*\right]_{1/2} = \mathcal{H}_{\mathrm{ker}} \subset \mathcal{H} \hookrightarrow \mathcal{V}^*$ holds by [LioM72, Ch. 1, Prop. 2.1(a)].

Remark 8.40. For more regular data one can estimate the errors of u_{c} and γ by applying Theorem 5.4 to the system (8.4b–c). As a matter of fact, Theorem 5.4 can be extended to this infinite-dimensional system since the operators $\mathcal{B}|_{\mathcal{V}_{\mathrm{c}}} \in \mathscr{L}(\mathcal{V}_{\mathrm{c}}, \mathcal{Q}^*)$ and $\mathcal{C} \in \mathscr{L}(\mathcal{Q}, \mathcal{Q}^*)$ have bounded inverses.

In contrast to u and γ the convergence rate for the Lagrange multiplier λ is half an order smaller. This is the case, since in general we can estimate $\dot{u}_{\mathrm{ker}}$ only in $\mathcal{V}_{\mathrm{ker}}^*$ rather than in the more restrictive space $\mathcal{V}^*$; see the proof of Theorem 8.37 and Remark 8.39. For the implicit Euler method, however, all approximations have the same convergence order.

Lemma 8.41 (Convergence Order of the Implicit Euler Scheme)**.** *In addition to the assumptions of Theorem 8.37 let $u_0 \in \mathcal{V}$ fulfill $f(0) - \mathcal{A}u_0 - \mathcal{B}^-\dot{g}(0) \in \mathcal{H}_{\mathrm{ker}}^*$. Then the discretization $\{(u_n, \lambda_n, \gamma_n)\}_{n=1,\dots,N}$ given by the implicit Euler scheme (8.6) satisfies the error bounds (8.46) with τ^2, i.e., the convergence order is one for every state variable.*

Proof. The convergence orders for u and γ follow along the lines of Theorem 8.37; see also [AltZ18c, Lem 6.4]. For $\Delta\lambda_n$ it is sufficient to show that the error $\Delta\dot{u}_{\mathrm{ker},n}$ is of first order; see (8.52) and the proof of Theorem 8.37. By (8.25) and the proof of Theorem 6.8 the error $\Delta\dot{u}_{\mathrm{ker},n}$ satisfies

$$D_\tau\Delta\dot{u}_{\mathrm{ker},n} + \mathcal{A}\Delta\dot{u}_{\mathrm{ker},n} = \mathcal{A}\big(D_\tau u_{\mathrm{ker}}(t_n) - \dot{u}_{\mathrm{ker}}(t_n)\big) = \int_{t_{n-1}}^{t_n} \frac{t - t_{n-1}}{\tau}\mathcal{A}\ddot{u}_{\mathrm{ker}}(t)\,\mathrm{d}t \quad \text{in } \mathcal{V}_{\mathrm{ker}}^* \tag{8.55}$$

with $D_\tau u_{\mathrm{ker},0} = \dot{u}_{\mathrm{ker}}(0) = f(0) - \mathcal{A}u_0 - \mathcal{B}^-\dot{g}(0) \in \mathcal{H}_{\mathrm{ker}}^* \cong \mathcal{H}_{\mathrm{ker}}$. By testing (8.55) with $\tau\Delta\dot{u}_{\mathrm{ker},n}$, we get similarly to the derivation of (7.13) the estimate

$$\max_{n=1,\dots,N}\|\Delta\dot{u}_{\mathrm{ker},n}\|^2_{\mathcal{H}} + \tau\sum_{n=1}^N\|\Delta\dot{u}_{\mathrm{ker},n}\|^2_{\mathcal{V}} \lesssim \sum_{n=1}^N\bigg\|\int_{t_{n-1}}^{t_n}\frac{t-t_{n-1}}{\tau}\mathcal{A}\ddot{u}_{\mathrm{ker}}(t)\,\mathrm{d}t\bigg\|^2_{\mathcal{V}^*_{\mathrm{ker}}} \lesssim \tau^2\int_0^T\|\ddot{u}_{\mathrm{ker}}\|^2_{\mathcal{V}}\,\mathrm{d}t.$$

□

According to Lemma 8.41, the state u and the Lagrange multiplier λ have the same convergence rate if we use the implicit Euler scheme. To get the same result for general Runge-Kutta methods, we assume the splitting $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ of Theorem 8.30. This additional assumption also improves the convergence of u in the time-discrete counterpart of $C([0,T],\mathcal{V})$.

Theorem 8.42 (Convergence Order for More Regular Data)**.** *In addition to the assumptions of Theorem 8.37, let an operator $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ exist, which is self-adjoint, elliptic on $\mathcal{V}_{\text{ker}}$, such that $\mathcal{A}_2 = \mathcal{A} - \mathcal{A}_1 \in \mathscr{L}(\mathcal{V},\mathcal{H}^*)$. Assume that f is an element of $L^2(0,T;\mathcal{H}^*)$. Suppose that the solution u and the right-hand side g are regular enough such that*

$$C := \int_0^T \|\mathcal{A}u_{\text{ker}}^{(\mathsf{q}+1)}\|^2_{\mathcal{H}^*_{\text{ker}}} + \|u_{\text{ker}}^{(\mathsf{q}+2)}\|^2_{\mathcal{H}_{\text{ker}}} + \|u^{(\mathsf{q}+1)}\|^2_{\mathcal{V}} + \|g^{(\mathsf{q}+2)}\|^2_{\mathcal{Q}^*}\,\mathrm{d}t$$

is well-defined with $u_{\text{ker}} = u - \mathcal{B}^-_{\mathcal{A}}\mathcal{B}u$. Then we have

$$\max_{n=1,\dots,N} \|u_n - u(t_n)\|^2_{\mathcal{V}} \lesssim \tau^{2\mathsf{q}+2}\,C, \tag{8.56a}$$

$$\tau\sum_{n=1}^N \|\lambda_n - \lambda(t_n)\|^2_{\mathcal{Q}} \lesssim \tau^{2\mathsf{q}+2}\Big(C + \int_0^T \|\lambda^{(\mathsf{q}+1)}\|^2_{\mathcal{Q}}\,\mathrm{d}t\Big). \tag{8.56b}$$

Proof. We use the notation and the splitting $u = u_{\text{ker}} + u_{\text{c}}$ of the proof of Theorem 8.37.

Starting with the approximation of the single parts of u, we have that the estimate (8.56a) for $\Delta u_{\text{c},n}$ follows by (8.48). For the part with images in $\mathcal{V}_{\text{ker}}$ we investigate the error $\mathcal{A}_1\Delta \boldsymbol{u}_{\text{ker},n}$ in the time-discrete counterpart of $L^2(0,T;\mathcal{H}^*_{\text{ker},s})$ in addition to $\Delta u_{\text{ker},n}$. Note that Theorem 4.25 and equation (8.35) guarantee that $\mathcal{A}_1 u$ and $\mathcal{A}_1\boldsymbol{u}_{\text{ker},\tau}$ are L^2-functions with images in $\mathcal{H}^*_{\text{ker}}$ and $\mathcal{H}^*_{\text{ker},s}$, respectively. Therefore, we have

$$\Delta\dot{\boldsymbol{u}}_{\text{ker},n} + \mathcal{A}_1\Delta\boldsymbol{u}_{\text{ker},n} = -\mathcal{A}_2\Delta\boldsymbol{u}_{\text{ker},n} \quad \text{in } \mathcal{H}^*_{\text{ker}} \tag{8.57}$$

by (8.20) and (8.35). Using (8.57) and $\|v_{\text{ker}}\|^2_{\mathcal{A}_1} \le \|\mathcal{A}_1 v_{\text{ker}}\|_{\mathcal{H}^*_{\text{ker}}}\|v_{\text{ker}}\|_{\mathcal{H}_{\text{ker}}}$ for every $v_{\text{ker}} \in \mathcal{V}_{\text{ker}}$ with $\mathcal{A}_1 v_{\text{ker}} \in \mathcal{H}^*_{\text{ker}}$ by the Gelfand triple $\mathcal{V}_{\text{ker}}$, $\mathcal{H}_{\text{ker}}$, $\mathcal{V}^*_{\text{ker}}$, one proves

$$\begin{aligned}&\|\Delta u_{\text{ker},n}\|^2_{\mathcal{A}_1} + \tau\sum_{i=1}^n \|\mathcal{A}_1\Delta\boldsymbol{u}_{\text{ker},i}\|^2_{\mathcal{H}^*_{\text{ker},s}}\\ &\quad\lesssim \|\Delta u_{\text{ker},0}\|^2_{\mathcal{A}_1} + \tau^{2\mathsf{q}+2}\int_0^T \|\mathcal{A}u_{\text{ker}}^{(\mathsf{q}+1)}\|^2_{\mathcal{H}^*_{\text{ker}}} + \|u_{\text{ker}}^{(\mathsf{q}+2)}\|^2_{\mathcal{H}_{\text{ker}}}\,\mathrm{d}t + \tau\sum_{i=1}^n \|\mathcal{A}_2\Delta\boldsymbol{u}_{\text{ker},i}\|^2_{\mathcal{H}^*_{\text{ker},s}}\end{aligned} \tag{8.58}$$

by an adaptation of the proof of [LubO95b, Th. 1.1]. In contrast to [LubO95b, Th. 1.1] which estimates the error $\Delta u_{\text{ker},n}$ in $\mathcal{H}_{\text{ker}}$, we consider $\Delta u_{\text{ker},n}$ in $(\mathcal{V}_{\text{ker}}, \|\cdot\|_{\mathcal{A}_1})$. However, the inequalities (8.51) and (8.58) imply the error estimate (8.56a) for u_{ker}. Thus, (8.56a) follows by the triangle inequality and the single error bounds for u_{c} and u_{ker}.

For the Lagrange multiplier λ we note that (8.51), (8.57), and (8.58) imply that $\tau\sum_{n=1}^N \|\Delta\dot{\boldsymbol{u}}_{\text{ker},n}\|^2_{\mathcal{H}_s}$ has convergence rate $2\mathsf{q}+2$. Thus, the error bound (8.56b) for $\Delta\lambda_n$ follows by Step 4 in the proof of Theorem 8.37. □

Remark 8.43. By piecewise polynomial interpolations based on the approximations u_n, λ_n, and γ_n one can construct abstract functions; cf. [AltZ18c, p. 23 f.]. These functions then satisfy the error estimates in Theorems 8.37 and 8.42 as well as in Lemma 8.41 in the norms of time-continuous functions, i.e., $\max_{n=1,\dots,N}\|(\cdot)_n\|^2_{\mathcal{X}}$ and $\tau\sum_{n=1}^N\|(\cdot)_n\|^2_{\mathcal{X}}$ become the norm in $C([0,T],\mathcal{X})$ and $L^2(0,T;\mathcal{X})$, respectively; cf. [AltZ18c, Th. 6.6].

For finite-dimensional DAEs, Theorem 5.4 provides a higher convergence order than Theorem 8.37. If we consider the spatial discretization (8.3) of the operator DAE (8.4), it may be more reasonable to compare the *fully-discrete* solution, i.e., the spatially and temporally discretized solution, to the solution of the operator DAE (8.4) rather than to the solution of (8.3). If the spatial discretization is based on *conforming finite elements* [Bra07, p. 60] then the fully-discrete solution can be rewritten as the space-continuous functions $(\overline{u}_n, \overline{\lambda}_n, \overline{\gamma}_n) \in \mathcal{V} \times \mathcal{Q} \times \mathcal{Q}$. By the arguments of [Alt15, Sec. 10.4.2], these time-discrete solutions satisfy

$$\overline{u}_n = R(\infty)\overline{u}_{n-1} + \mathbf{b}^T \mathbf{A}^{-1} \overline{\boldsymbol{u}}_n, \qquad \overline{\gamma}_n = R(\infty)\overline{\gamma}_{n-1} + \mathbf{b}^T \mathbf{A}^{-1} \overline{\boldsymbol{\gamma}}_n, \tag{8.59a}$$

with the internal stages $(\overline{\boldsymbol{u}}_n, \overline{\boldsymbol{\lambda}}_n, \overline{\boldsymbol{\gamma}}_n) \in \mathcal{V}_s \times \mathcal{Q}_s \times \mathcal{Q}_s$ given by

$$\mathbf{A}^{-1} \boldsymbol{D}_\tau \overline{\boldsymbol{u}}_n + \mathcal{A}\overline{\boldsymbol{u}}_n - \mathcal{B}^* \overline{\boldsymbol{\lambda}}_n - \mathcal{B}^* \overline{\boldsymbol{\gamma}}_n = \boldsymbol{f}_n + \boldsymbol{\delta}_n \quad \text{in } \mathcal{V}_s^*, \tag{8.59b}$$

$$\mathcal{B}\overline{\boldsymbol{u}}_n + \mathcal{C}\overline{\boldsymbol{\gamma}}_n = \boldsymbol{g}_n + \boldsymbol{\theta}_n \quad \text{in } \mathcal{Q}_s^*, \tag{8.59c}$$

$$\mathcal{B}\mathbf{A}^{-1} \boldsymbol{D}_\tau \overline{\boldsymbol{u}}_n = \dot{\boldsymbol{g}}_n + \boldsymbol{\xi}_n \quad \text{in } \mathcal{Q}_s^*. \tag{8.59d}$$

The new initial values are $\overline{u}_0 = u_0 + e_0$ and $\overline{\gamma}_0 = -\mathcal{C}^{-1}\mathcal{B}e_0$. This allows us to interpret the spatial truncation error e_0 and the residuals $\boldsymbol{\delta}_n$, $\boldsymbol{\theta}_n$, and $\boldsymbol{\xi}_n$ as perturbations of the system (8.40). The following lemma addresses the influence of perturbations of (8.40) onto its solutions.

Lemma 8.44 (Error under Perturbations)**.** *Let the assumptions of Theorem 8.37 be satisfied. Suppose that the initial value u_0 is perturbed by $e_0 = e_{\ker,0} + e_{c,0} \in \mathcal{V}_{\ker} \oplus \mathcal{V}_c = \mathcal{V}$. Assume that the perturbations $\boldsymbol{\delta}_n \in \mathcal{V}_s^*$, $\boldsymbol{\theta}_n \in \mathcal{Q}_s^*$, and $\boldsymbol{\xi}_n \in \mathcal{Q}_s^*$ of the right-hand sides $\boldsymbol{f}_n$, $\boldsymbol{g}_n$, and $\dot{\boldsymbol{g}}_n$, respectively, are given, $n = 1, \dots, N$. Then the discrete solutions $\overline{u}_n$ and $\overline{\gamma}_n$ of the perturbed system (8.59) satisfy*

$$\max_{n=1,\dots,N} \|\overline{u}_n - u(t_n)\|^2_{\mathcal{H}} + \tau \sum_{n=1}^N \|\overline{u}_n - u(t_n)\|^2_{\mathcal{V}}$$
$$\lesssim E_u + \|e_0\|^2_{\mathcal{V}} + \tau |R(\infty)|\, \|e_{\ker,0}\|^2_{\mathcal{V}} + \tau \sum_{n=1}^N \|\boldsymbol{\delta}_n\|^2_{\mathcal{V}_s^*} + \|\boldsymbol{\xi}_n\|^2_{\mathcal{Q}_s^*},$$
$$\tau \sum_{n=1}^N \|\overline{\gamma}_n\|^2_{\mathcal{Q}} \lesssim E_\gamma + \|e_{c,0}\|^2_{\mathcal{V}} + \tau \sum_{n=1}^N \|\boldsymbol{\theta}_n\|^2_{\mathcal{Q}_s^*} + \|\boldsymbol{\xi}_n\|^2_{\mathcal{Q}_s^*},$$

where E_u and E_γ denote the right-hand side of (8.46a) and (8.46c), respectively.

Proof. The proof follows along the lines of Theorem 8.37 using [LubO95b, p. 605, Remark (c)]. □

Remark 8.45. Similar to Lemma 8.44 one can estimate the error in Theorem 8.42 under perturbations.

Remark 8.46. By the loss of half an order for the Lagrange multiplier λ in (8.46b), an estimate of $\tau \sum_{n=1}^N \|\overline{\lambda}_n - \lambda(t_n)\|^2_{\mathcal{Q}}$ would include the terms $\tau^{-1}\|e_{\ker,0}\|^2_{\mathcal{V}}$ and $\sum_{n=1}^N \|\boldsymbol{\delta}_n\|^2_{\mathcal{V}_s^*}$.

Finally, we make some remarks on the temporal discretization of operator DAEs with time-dependent operators.

Remark 8.47.

i) Theorem 8.37 as well as Lemmas 8.41 and 8.44 still hold for operator DAEs of the form (8.1) with a time-dependent operator $\mathcal{A} \in C([0,T], \mathscr{L}(\mathcal{V}, \mathcal{V}^*))$, which satisfies uniformly a Gårding inequality on $\ker \mathcal{B}$. The term $\mathcal{A}\boldsymbol{u}_n$ is then replaced by $\mathcal{A}_n \boldsymbol{u}_n := [\mathcal{A}(t_{n-1}+\tau \mathbf{c}_1)\boldsymbol{u}_{n,1}, \dots, \mathcal{A}(t_{n-1}+\tau \mathbf{c}_s)\boldsymbol{u}_{n,s}]^T \in \mathcal{V}_s^*$ in the discretized systems (8.6a) and (8.40a). For the associated proofs one adapts the results of the first chapter in [LubO95b] to operator ODEs with an operator $\mathcal{A} \colon [0,T] \to \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$, which satisfies uniformly a Gårding inequality.

ii) An extension to systems with a time-dependent, uniformly inf-sup stable operator $\mathcal{B} \in C^1([0,T], \mathscr{L}(\mathcal{V}, \mathcal{Q}^*))$, which has a time-independent kernel, is also possible. The discretization of (8.4b) and (8.4c) are then given by $\mathcal{B}_n \boldsymbol{u}_n + \mathcal{C}\boldsymbol{\gamma}_n = \boldsymbol{g}_n$ and $\mathcal{B}_n \mathbf{A}^{-1} \boldsymbol{D}_\tau \boldsymbol{u}_n + \dot{\mathcal{B}}_n \boldsymbol{u}_n = \dot{\boldsymbol{g}}_n$, respectively. The terms $\mathcal{B}_n, \dot{\mathcal{B}}_n \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ are defined similar to $\mathcal{A}_n$.
If $\mathcal{B}$ and g are regular enough as well as τ is small enough, then $\Delta u_{\mathrm{c},n}$ and $\Delta\gamma_n$ can be bounded by Theorem 5.4. Estimates under perturbations can be derived by [HaiW96, Th. 3.6]. For the remaining steps of the associated proofs one uses $\Delta \boldsymbol{u}_{\mathrm{c},n} = \Delta u_{\mathrm{c},n-1} \mathbb{1}_s + \tau \mathbf{A} \mathcal{B}^-_{n,\mathcal{A}} \dot{\mathcal{B}}_n \Delta \boldsymbol{u}_{\mathrm{c},n} + \check{\boldsymbol{R}}_n$. Note that $\tau \|\mathbf{A} \mathcal{B}^-_{n,\mathcal{A}} \dot{\mathcal{B}}_n\|_{\mathscr{L}(\mathcal{V}_s)} < 1$ holds if τ is small enough and that the correction term $\check{\boldsymbol{R}}_n \in \mathcal{V}_s$ satisfies $\|\check{\boldsymbol{R}}_n\|^2_{\mathcal{V}_s} \lesssim \tau^{2\mathsf{q}+1} \int_{t_{n-1}}^{t_n} \|u_{\mathrm{c}}^{(\mathsf{q}+1)}\|^2_{\mathcal{V}^*} \,\mathrm{d}t$. Finally, the additional term $\Delta \dot{\boldsymbol{u}}_{\mathrm{c},n} = -\mathcal{B}^-_{n,\mathcal{A}} \dot{\mathcal{B}}_n \Delta \boldsymbol{u}_{\mathrm{c},n}$ in the error estimates for $u_{\ker}$ can be treated as a perturbation using Lemma 8.44.

8.6. Numerical Examples

We illustrate the performance of Runge-Kutta schemes for two operator DAEs in this final section of Chapter 8. The first example is a so-called lid-driven cavity modeled by the unsteady Stokes equation. As the second example, we consider a synthetic model to investigate the impact of the spatial regularity on the convergence order. The associated simulation code can be found in [Zim20].

8.6.1. Stokes Problem – Lid-Driven Cavity

As a first example we consider a fluid described by the unsteady Stokes equation, see Example 6.1, inside a square cavity. The cavity has three rigid walls with no-slip conditions and a fourth side where a moving lid enforces a velocity field. The governing equations are given by

$$\begin{aligned} \dot u(\xi,t) - \mu\Delta u(\xi,t) - \nabla p(\xi,t) &= 0 && \text{in } \Omega \times (0,T], && (8.60a)\\ \operatorname{div} u(\xi,t) &= 0 && \text{in } \Omega \times (0,T], && (8.60b)\\ u(\xi,t) &= 0 && \text{on } \Gamma_1 \times (0,T], && (8.60c)\\ u(\xi,t) &= g(\xi,t) && \text{on } \Gamma_2 \times (0,T]. && (8.60d) \end{aligned}$$

The state $u\colon \Omega \times [0,T] \to \mathbb{R}^2$ is the velocity field and $p\colon \Omega \times [0,T] \to \mathbb{R}$ is the pressure. In our example the domain is the unit square $\Omega = (0,1)^2$. The subset Γ_1 and Γ_2 of the boundary $\partial\Omega$ are defined by $\Gamma_1 \coloneqq \{0,1\} \times [0,1] \cup (0,1) \times \{0\}$ and $\Gamma_2 \coloneqq (0,1) \times \{1\}$.

Following [Zim15, Sec. 4.3.1] the weak formulation of the constrained PDE (8.60) has the form of the operator DAE (8.1). We incorporate the no-slip conditions (8.60c) into the space $\mathcal{V}$. The divergence condition (8.60b) is stated explicitly as a constraint in the weak formulation. For the boundary condition (8.60d) we choose $g = [g_1\ 0]^T$ and formulate $u_1|_{\Gamma_2} = g_1$ as a constraint whereas $u_2|_{\Gamma_2} = 0$ is integrated into the space $\mathcal{V}$. Note that, u_1 vanishes at Γ_1 by (8.60c). Thus, the right-hand side g_1 must be extendable to a function of $H^{1/2}(\partial\Omega)$ by setting g_1 to zero on Γ_1, i.e., $g_1 \in H^{1/2}_{00}(\Gamma_2)$. Taking all these aspects into consideration, cf. also Example 6.1, we have

$$\mathcal{V} = H^1_{\Gamma_1}(\Omega) \times H^1_0(\Omega), \qquad \mathcal{H} = L^2(\Omega) \times L^2(\Omega), \qquad \mathcal{Q} = (L^2(\Omega) \setminus \mathbb{R}) \times H^{-1/2}_{00}(\Gamma_2).$$

For the regularization from Section 8.1 the operator $\mathcal{C} \in \mathscr{L}(\mathcal{Q}, \mathcal{Q}^*)$ is chosen as

$$\mathcal{C} = \begin{bmatrix} \mathcal{C}_1 & 0 \\ 0 & \mathcal{C}_2 \end{bmatrix},$$

where the operators $\mathcal{C}_1 \in \mathscr{L}(L^2(\Omega) \setminus \mathbb{R}, (L^2(\Omega) \setminus \mathbb{R})^*)$ and $\mathcal{C}_2 \in \mathscr{L}(H^{-1/2}_{00}(\Gamma_2), H^{1/2}_{00}(\Gamma_2))$ are elliptic.

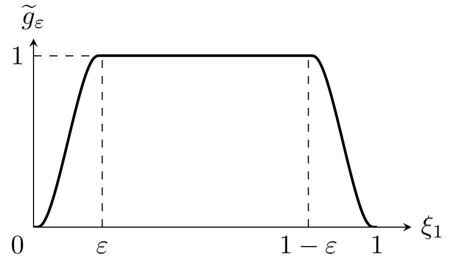

Figure 8.1.: Sketch of the function $\widetilde{g}_\varepsilon$.

We choose $\mathcal{C}_1 = \mathcal{R}_{L^2(\Omega)}$, which is well-defined since $L^2(\Omega) \setminus \mathbb{R}$ is a closed subspace of $L^2(\Omega)$ [Zim15, Rem. 2.12]. For $\mathcal{C}_2$ we note that there exists a bounded linear map $H_{00}^{-1/2}(\Gamma_2) \to H^1_{\Gamma_1}(\Omega)$; $\omega \mapsto u_\omega$ where u_ω is the unique solution of

$$\int_\Omega \nabla u_\omega \cdot \nabla v \, \mathrm{d}\xi = \langle \omega, \operatorname{tr}_{\Gamma_2} v \rangle_{H_{00}^{-1/2}(\Gamma_2), H_{00}^{1/2}(\Gamma_2)} \qquad \text{for all } v \in H^1_{\Gamma_1}(\Omega); \tag{8.61}$$

see [Sch98, Sec. 1.4.3]. We set $\mathcal{C}_2\omega = \operatorname{tr}_{\Gamma_2} u_\omega \in H_{00}^{1/2}(\Gamma_2)$. By the right-inverse tr^- of the trace operator, see Theorem 3.24, the inner product (3.15) of $H_{00}^{1/2}(\Gamma_2)$, and $r := \mathcal{R}^{-1}_{H_{00}^{1/2}(\Gamma_2)}\omega$ the estimate

$$\|\omega\|^2_{H_{00}^{-1/2}(\Gamma_2)} = \langle \omega, r \rangle = \int_\Omega \nabla u_\omega \cdot \nabla \operatorname{tr}^- r \, \mathrm{d}\xi \le \|u_\omega\|_{H^1_{\Gamma_1}(\Omega)} \|\operatorname{tr}^- r\|_{H^1_{\Gamma_1}(\Omega)} = \|u_\omega\|_{H^1_{\Gamma_1}(\Omega)} \|\omega\|_{H_{00}^{-1/2}(\Gamma_2)}$$

holds. Therefore, $\mathcal{C}_2$ is elliptic by $\langle \omega, \mathcal{C}_2\omega \rangle = \|u_\omega\|^2_{H^1_{\Gamma_1}(\Omega)} \ge \|\omega\|^2_{H_{00}^{-1/2}(\Gamma_2)}$.

Before we come to the simulation, we have to discuss the right-hand side $g = [g_1 \, 0]^T$. We point out that the moving lid would imply that g_1 is constant on Γ_2, which is not an $H_{00}^{1/2}(\Gamma_2)$-function unless g_1 vanishes everywhere [Sch98, Ex. 1.38]. Therefore, we translate smoothly the enforced velocity field to zero near the corners of $\Gamma_2 \subset \partial\Omega$. In particular, we will use $g_1(\xi_1, t) = \overline{g}_1(t)\, \widetilde{g}_{1/64}(\xi_1)$ for the simulation, where $\overline{g}_1(t)$ passes slowly from zero to one and $\widetilde{g}_{1/64}(\xi_1)$ is given by the parameterized function

$$\widetilde{g}_\varepsilon(\xi_1) := (\varphi_{\varepsilon/2} * \chi_{[\varepsilon/2, 1-\varepsilon/2]})(\xi_1) := \int_\mathbb{R} \varphi_{\varepsilon/2}(\eta) \chi_{[\varepsilon/2, 1-\varepsilon/2]}(\xi_1 - \eta) \, \mathrm{d}\eta = \int_{\varepsilon/2}^{1-\varepsilon/2} \varphi_{\varepsilon/2}(\xi_1 - \eta) \, \mathrm{d}\eta$$

with choice $\varepsilon = 1/64$. In the definition of $\widetilde{g}_\varepsilon(\xi_1)$ the parameter ε satisfies $0 < \varepsilon \le 0.5$, χ_E denotes the characteristic function of the set E, and φ_ε is given by $\varphi_\varepsilon(\eta) = c\varepsilon^{-1} \exp(\varepsilon^2/(\eta^2 - \varepsilon^2))\chi_{(-\varepsilon,\varepsilon)}(\eta)$, where the constant c is chosen such that $\int_\mathbb{R} \varphi_1 \, \mathrm{d}\eta = 1$. The function $\widetilde{g}_\varepsilon$ is depicted in Figure 8.1. However, by construction the function $\widetilde{g}_\varepsilon \in C^\infty([0,1])$ is non-negative, constantly one at $[\varepsilon, 1-\varepsilon]$, and satisfies $\widetilde{g}_\varepsilon^{(k)}(0) = \widetilde{g}_\varepsilon^{(k)}(1) = 0$ for all $k \in \mathbb{N}_0$; see e.g. [Zei90a, Sec. 18.14].

Numerical Simulation For the calculation we choose as the dynamic viscosity $\mu = 1/400$ and vanishing initial values. The final time-point is $T = 10$ and the non-zero entry g_1 of the right-hand side g is given by $g_1(\xi_1, t) = \big(1 - (\frac{t}{10} - 1)^8\big)\, \widetilde{g}_{1/64}(\xi_1)$. An illustration of the associated solution is given in Figure 8.2.

As spatial discretization we use quadratic finite elements for the velocity field, which incorporate the homogeneous boundary conditions. The pressure is discretized by linear finite elements which vanish at $\xi = (0,0)$. In a post-processing step the zero mean condition is realized. The additional

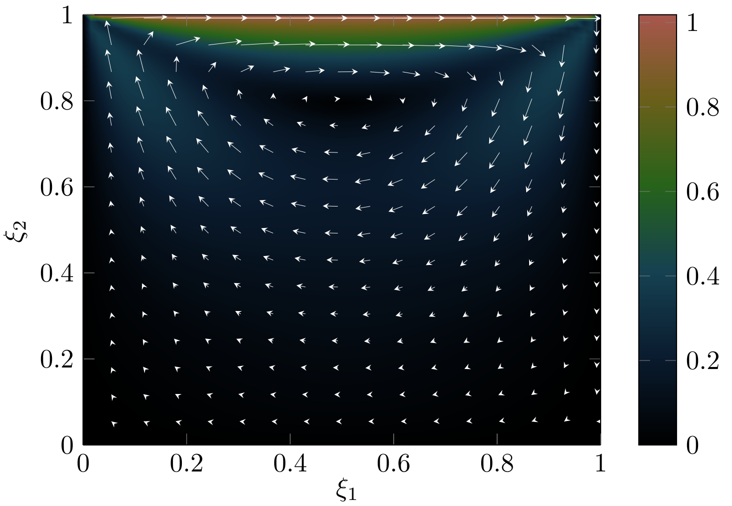

Figure 8.2.: Illustration of the velocity field u at final time $t = 10$.

Lagrange multiplier for the constraint (8.60d) is approximated by piecewise constant finite elements. As mesh we use a uniform criss-cross triangulation with spatial mesh size h. For more details see [AltH15, Sec. 4], [Ver84], and [Zim15, Sec. 4.3.2]. The implementation of the triangulation is based on the MATLAB software package AFEM [CarGK+10].

We implemented the Radau IIA schemes with one, two, and three stages for the temporal discretization; see e.g. [HaiW96, p. 74, Tab. 5.5 & 5.6]. These methods are stiffly accurate, and have classical order $\mathsf{p} = 2s - 1$ and stage order $\mathsf{q} = s$ [HaiW96, p. 72 ff. & p. 227]. The reference solution is determined by the three-stage method with step size $\tau = 5 \cdot 2^{-12}$. Figure 8.3 displays the approximation error of the velocity field u and of the pressure p for different spatial and temporal mesh sizes. We observe for the velocity field u that every graph for the one- and two-stage methods has optimal slope $2s - 1 = \min(\mathsf{p}, \mathsf{q} + 1)$ independently of the mesh size. The convergence order for the scheme with three stages is 4.75. This is less than the classical order but still 0.75 better than predicted by Theorem 8.37. This observed order, however, is the optimal rate for PDEs with homogeneous Dirichlet boundary conditions [LubO95b, p. 616, Ex. (i)]. Since $\mathcal{V}_{\text{ker}} = \ker \mathcal{B}$ is the space of trace- and divergence-free functions, we expect that the convergence order of u_{ker} and therefore of u is limited by $\mathsf{q} + 1.75$. Under the assumption of smooth solutions, this explains the observed convergence rates.

Furthermore, we observe in Figure 8.3 that the convergence orders for the pressure p are the same as for the state u. Lemma 8.41 and Theorem 8.42 predict these rates for the one- and two-stage method. For the three-stage method we note that Theorem 8.42 only proves the same convergence rate $\mathsf{q} + 1$ for u and p, not the improved rate of $\mathsf{q} + 1.75$. However, we expect that an adaptation of [LubO95b, Th. 3.3] for an error estimate of $\tau \sum_{n=1}^{N} \|\mathcal{A}\Delta u_{\text{ker},n}\|^2_{\mathcal{H}^*_{\text{ker}}}$ combined with the stiff accuracy of the Radau IIA methods proves the observed convergence order; cf. Subsection 9.3.2.2 and Theorem 9.24 in particular. As for finite-dimensional DAEs, see Theorem 5.4, the stiff accuracy is expected to be the key point in the proof.

8.6.2. A Synthetic Example

In the previous numerical example, we saw an improvement of the convergence order, which could be explained by solutions which are smooth in space and time. In this subsection we study the effect of the spatial regularity of the solution from (8.1) on the convergence rate. For this, we consider as

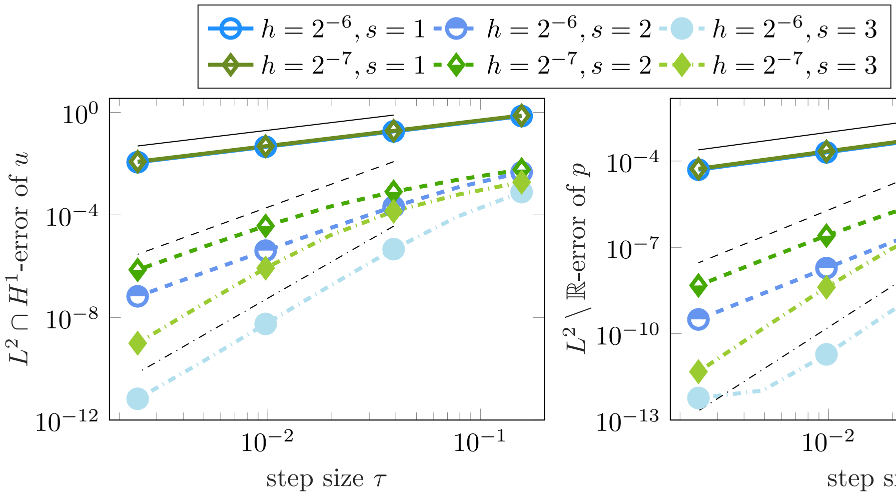

Figure 8.3.: Convergence history for the error of the velocity field u in the time-discrete counterpart of $C([0,10],[L^2(\Omega)]^2) \cap L^2(0,10;H^1_{\Gamma_2}(\Omega) \times H^1_0(\Omega))$ (left) and for the error of the pressure p in the time-discrete counterpart of $L^2(0,10;L^2(\Omega) \setminus \mathbb{R})$ (right). For comparison, we added lines with slope one (solid), three (dotted), and 4.75 (dash-dotted).

the second numerical example the parameterized PDE

$$\dot{u}_\alpha(\xi,t) - \partial_{\xi\xi} u_\alpha(\xi,t) - \partial_{\xi\xi} v_\alpha(\xi,t) = \quad -\pi e^{-\pi t} + \sum_{k=1}^{\infty} \frac{2\pi^2 k^2 - \pi k^2}{k^\alpha} e^{-k^2 \pi t} \sin(k\pi\xi) \tag{8.62a}$$

$$\dot{v}_\alpha(\xi,t) + \quad u_\alpha(\xi,t) - \partial_{\xi\xi} v_\alpha(\xi,t) = (1-\pi)e^{-\pi t} + \sum_{k=1}^{\infty} \frac{\pi^2 k^2 - \pi k^2 + 1}{k^\alpha} e^{-k^2 \pi t} \sin(k\pi\xi) \tag{8.62b}$$

$$u_\alpha(0,t) = u_\alpha(1,t) = v_\alpha(0,t) = v_\alpha(1,t) = e^{-\pi t} \tag{8.62c}$$

with the unit interval as spatial domain and $T = 1$. We choose the initial value $(u_{\alpha;0}, v_{\alpha;0})$ such that the solution of (8.62) is given by

$$u_\alpha(\xi,t) = v_\alpha(\xi,t) = e^{-\pi t} + \sum_{k=1}^{\infty} \frac{e^{-k^2\pi t}}{k^\alpha} \sin(k\pi\xi). \tag{8.63}$$

The parameter α will be used to vary the spatial regularity of the solution.

Following Example 6.2, we incorporate the boundary conditions (8.62c) by means of the trace operator as constraints such that the weak formulation of the PDE (8.62) is an operator DAE of the form (8.1). Thereby, we introduce the Lagrange multiplier $\lambda_\alpha \in \mathcal{Q}$. The spaces are given by $\mathcal{V} = [H^1(0,1)]^2$, $\mathcal{V}_{\text{ker}} = [H^1_0(0,1)]^2$, and $\mathcal{Q} = \mathbb{R}^4$. The part $(u_{\alpha;\text{ker}}, v_{\alpha;\text{ker}})$ of the solution (u_α, v_α) with images in $\mathcal{V}_{\text{ker}}$ is given by the infinite sum in the solution (8.63). We emphasize that the right-hand side $g(t) = e^{-\pi t}\mathbb{1}_4$ is infinitely differentiable and the solution satisfies

$$u^{(k)}_{\alpha;\text{ker}} = v^{(k)}_{\alpha;\text{ker}} \in L^2(0,T;H_0^{\alpha-2k+1/2-\varepsilon}(0,1)) \quad \text{and} \quad u^{(k+1)}_{\alpha;\text{ker}} = v^{(k+1)}_{\alpha;\text{ker}} \in L^2(0,T;H^{\alpha-2k-3/2-\varepsilon}(0,1)) \tag{8.64}$$

for $k \in \mathbb{N}_0$ with $\alpha + \frac{1}{2} \geq 2k > \alpha - \frac{3}{2}$ and an arbitrary $\varepsilon > 0$.

Table 8.1.: Properties of the Radau IA, Radau IIA, Lobatto IIIC, Lobatto IIID, and SDIRK Cro method with three stages; [HaiW96, p. 77 & p. 100] and [Jay15, p. 822].

	Radau IA	Radau IIA	Lobatto IIIC	Lobatto IIID	SDIRK Cro
Classical order	5	5	4	4	4
Stage order	2	3	2	1	1
$R(\infty)$	0	0	0	0	≈ -0.63
Algebraically stable	✓	✓	✓	✓	✓
Stiffly accurate		✓	✓		

For the simulation of the operator DAE associated to the PDE (8.62) we choose $\mathcal{C}$ as the identity matrix in $\mathbb{R}^{4\times 4}$ in the regularization described in Section 8.1. The spatial discretization is given by spectral finite elements with 400 degrees of freedom enriched by linear polynomials, i.e., we approximate $\mathcal{V}$ by the space

$$\mathcal{V}_{400} = \operatorname{span}\{1, \xi, \sin(\pi\xi), \dots, \sin(400\pi\xi)\} \subset \mathcal{V}.$$

As temporal integration schemes we use the Radau IA, Radau IIA, Lobatto IIIC, and Lobatto IIID methods with three stages each, cf. [HaiW96, Ch. IV, Tab. 5.4, 5.6, & 5.11] and [Jay15, Tab. 5]. In addition to these L-stable schemes we implemented the singly diagonally implicit method from Crouzeix (SDIRK Cro) with three stages; see [HaiW96, p. 100]. Some properties of these methods are collected in Table 8.1. Note that the node $\mathbf{c}_3 \approx -6.858 \cdot 10^{-2}$ of the SDIRK Cro scheme is negative and the right-hand side of (8.62) is unbounded for $t < 0$. Therefore, we choose $t = \tau$ as the initial time for this method.

Figure 8.4 illustrates the experimental convergence order for the state (u_α, v_α) and for the Lagrange multiplier λ_α as a function of the parameter α. We remind that the spatial regularity depends strongly on α; see (8.64). The depicted convergence rates are approximated by the means of the slopes of the logarithmic errors against the logarithmic step sizes. In these calculations we reject outliers which may occur in transient phases or by errors which are close to machine precision. We note that the convergence rates for (u_α, v_α) and λ_α are mostly better than predicted by Theorem 8.37. In particular, the convergence rate for (u_α, v_α) is not limited by $\mathsf{q}+1$ or by $\mathsf{q}+1.75$ for the stiffly accurate methods. This behavior can be explained by [LubO93, p. 116], because the boundary of $\Omega = (0,1)$ consists of two separated points and the solution $(u_{\alpha;\ker}, v_{\alpha;\ker})$ vanishes on this boundary. For α which is infinitesimal bigger than $2\ell + 2.5$, $\ell \in \mathbb{N}_0$, the convergence rates of $(u_{\alpha;\ker}, v_{\alpha;\ker})$ are $\min(\ell+1, \mathsf{p})$ as anticipated by Theorem 8.37. These values of α are special in the sense that by (8.64) the solutions $u^{(\ell+1)}_{\alpha;\ker} = v^{(\ell+1)}_{\alpha;\ker}$ are contained in $W^{1,2}(0,T; H^1_0(0,1), H^{-1}(0,1))$ for every $\alpha > 2\ell + 2.5$ and are not contained in this space for $\alpha \le 2\ell + 2.5$.

For λ_α we note that after the convergence order becomes stationary it is equal to $\mathsf{q}+1$ for the non-stiffly-accurate methods and it is similar to the order of (u_α, v_α) for the stiffly accurate ones. As for the numerical example in Subsection 8.6.1 we expect that this can be proven by error estimates for $\Delta u_{\ker,n}$ and $\Delta \boldsymbol{u}_{\ker,n}$ for spaces which are more regular than $H^1_0(0,1)$. However, before it transitions into the constant phase the convergence order for λ_α is around 0.25 smaller than the convergence rate of $(u_{\alpha;\ker}, v_{\alpha;\ker})$. This follows by Remark 8.39 and

$$[H^1_0(0,1), H^{-1}(0,1)]_{3/4-\varepsilon} = H^{-1/2+\varepsilon}(0,1) = \left[H_0^{1/2-\varepsilon}(0,1)\right]^* = \left[H^{1/2-\varepsilon}(0,1)\right]^* \hookrightarrow \left[H^1(0,1)\right]^*$$

with an arbitrary $\varepsilon > 0$, where we used [LioM72, Ch. 1, Th. 11.1 & 12.3].

Beside the convergence of (u_α, v_α) and λ_α, the results of this numerical example show that the convergence rate for the additional Lagrange multiplier γ_α is constant over α and as predicted in Theorem 8.37. Thus, we omit the corresponding plot here.

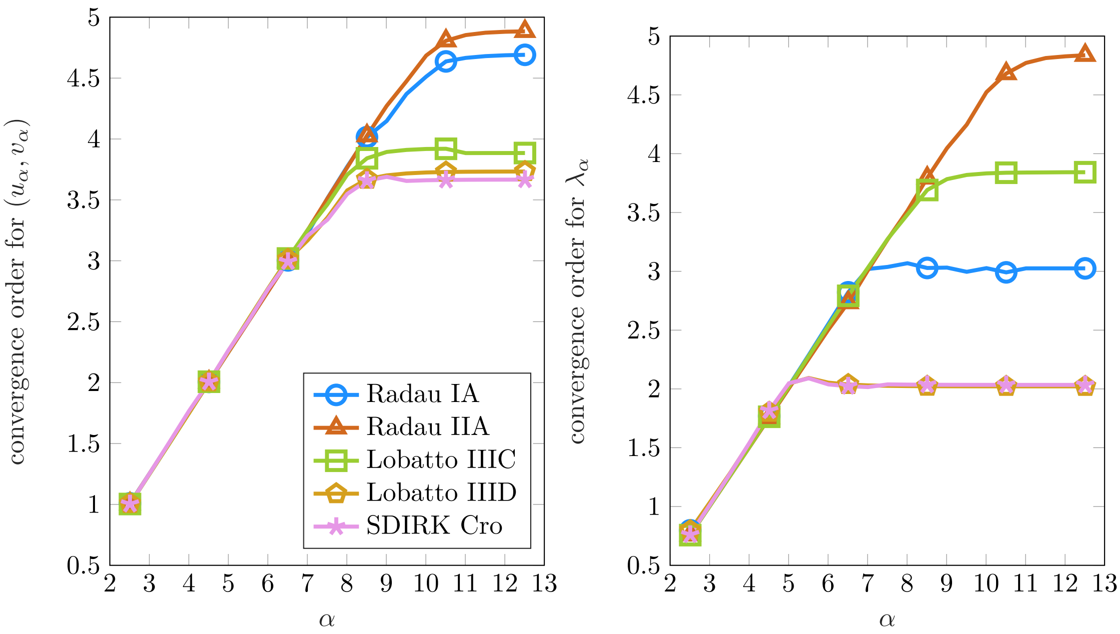

Figure 8.4.: The convergence order for (u_α, v_α) and λ_α against the spatial regularity of the solution characterized by the parameter α.

Finally, it is worth to mention that for the approximation of (u_α, v_α), λ_α, and γ_α the convergence rates of the non-L-stable SDIRK Cro method are similar to those of the L-stable Lobatto IIID method, which has the same classical and stage order.

9. Exponential Integrators

In this chapter we extend the idea of exponential integrators introduced in Section 5.2 to semi-linear operator DAEs of the form

$$\dot{u}(t) + \mathcal{A}u(t) - \mathcal{B}^*\lambda(t) = f(t, u(t)) \qquad \text{in } \mathcal{V}^*, \tag{9.1a}$$

$$\mathcal{B}u(t) = g(t) \qquad \text{in } \mathcal{Q}^*. \tag{9.1b}$$

The spaces $\mathcal{V}$ and $\mathcal{Q}$ are separable Hilbert spaces. We assume a third Hilbert space $\mathcal{H}$ exists such that $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$ is a Gelfand triple. The operators $\mathcal{A}$ and $\mathcal{B}$ are time-independent. In the interest of readability, we assume in addition to the assumptions of Subsection 6.4, which give us unique solutions, see Theorem 6.15, that $\mathcal{A}$ is elliptic on $\mathcal{V}_{\ker}$. Anyway, the results of this chapter are still valid if $\mathcal{A}$ satisfies only a Gårding inequality (3.6) on $\mathcal{V}_{\ker}$. In this case, we add to $\mathcal{A}$ the term $\kappa_{\mathcal{A}} \operatorname{id}_{\mathcal{H}}$ and accordingly to the nonlinearity f, i.e., we redefine $f(t,u) \leftarrow f(t,u) + \kappa_{\mathcal{A}} u$, such that $\mathcal{A} + \kappa_{\mathcal{A}} \operatorname{id}_{\mathcal{H}}$ is elliptic on $\mathcal{V}_{\ker}$. The assumption on the operators and the right-hand sides are summarized in the following.

Assumption 9.1 (Operator $\mathcal{B}$). The operator $\mathcal{B} \in \mathscr{L}(\mathcal{V}, \mathcal{Q}^*)$ satisfies the inf-sup condition (3.2).

Assumption 9.2 (Operator $\mathcal{A}$).

i) The operator $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ has the form $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ being self-adjoint and $\mathcal{A}_2 \in \mathscr{L}(\mathcal{V}, \mathcal{H}^*)$.

ii) We assume that $\mathcal{A}$ is elliptic on $\mathcal{V}_{\ker} := \ker \mathcal{B}$.

Assumption 9.3 (Right-Hand Sides f and g).

i) The right-hand side $f\colon [0,T] \times \mathcal{V} \to \mathcal{H}^*$ fulfills the Carathéodory conditions, see Definition 4.1, and is bounded via $\|f(t,v)\|_{\mathcal{H}^*} \leq k(t)(1 + \|v\|_{\mathcal{V}})$ with $k \in L^2(0,T)$ for all $v \in \mathcal{V}$ and almost all $t \in [0,T]$. Furthermore, f is locally Lipschitz continuous in the second argument; see (6.12).

ii) The function g is an element of $H^1(0,T;\mathcal{Q}^*)$.

For semi-linear ODEs and parabolic PDEs exponential integrators are well-studied. This includes explicit and implicit exponential Runge-Kutta methods [Cer60; CoxM02; HocO05a; HocO05b; Law67], exponential Runge-Kutta methods of high order [LuaO14a; LuaO14b], exponential Rosenbrock-type methods [HocOS09], and multistep exponential integrators [CalP06]. To the best of the author's knowledge, exponential integrators for constrained PDEs have only be studied for the incompressible Navier-Stokes equations [EdwTF+94; KooBG18; New03], where the convergence order is not investigated. In [HocLS98], the authors successfully apply exponential integrators to finite-dimensional index-1 DAEs. We emphasize that a standard spatial discretization of the operator DAE (9.1) by finite elements leads to a DAE of index two; see Remark 2.3 and Section 8.1.1. In this thesis, we consider *explicit* exponential integrators for the operator DAE (9.1). The approach for the explicit schemes, however, can be translated to the general exponential integrators like the methods mentioned above.

In the linear case, the solution of (9.1) can be expressed by the variation-of-constants formula; cf. [EmmM13, Sec. 3.2.2]. In the semi-linear case, we consider the term $f(t,u)$ as a right-hand side, which leads to an implicit formula only. This, however, is still of value for the numerical analysis of time integration schemes.

The solution formula is based on the decomposition $u = u_{\ker} + u_{\mathrm{c}}$ with $u_{\ker}\colon [0,T] \to \mathcal{V}_{\ker}$ and $u_{\mathrm{c}}\colon [0,T] \to \mathcal{V}_{\mathrm{c}}$ with $\mathcal{V}_{\mathrm{c}} := \{v \in \mathcal{V} \,|\, \mathcal{A}v \in \mathcal{V}_{\ker}^0\}$. The latter is fully determined by the constraint (9.1b), namely $u_{\mathrm{c}}(t) = \mathcal{B}_{\mathcal{A}}^- g(t) \in \mathcal{V}_{\mathrm{c}}$. For $u_{\ker}$ we consider the restriction of (9.1a) to the test space $\mathcal{V}_{\text{ker}}$. Since the Lagrange multiplier disappears in this case, we obtain

$$\dot{u}_{\text{ker}}(t) + \mathcal{A}_{\text{ker}} u_{\text{ker}}(t) = f(t, u_{\text{ker}} + u_{\text{c}}) - \dot{u}_{\text{c}}(t) \qquad \text{in } \mathcal{V}_{\text{ker}}^*, \tag{9.2}$$

with

$$\mathcal{A}_{\text{ker}} := \mathcal{A}|_{\mathcal{V}_{\text{ker}}} \colon \mathcal{V}_{\text{ker}} \to \mathcal{V}_{\text{ker}}^*.$$

Note that we use here the fact that functionals in $\mathcal{V}^*$ define functionals in $\mathcal{V}_{\text{ker}}^*$ simply through the restriction to $\mathcal{V}_{\text{ker}}$. The term $\mathcal{A}u_{\text{c}}$ disappears under test functions in $\mathcal{V}_{\text{ker}}$ due to the definition of $\mathcal{V}_{\text{c}}$. If this orthogonality is not respected within the implementation of numerical simulations, then this term needs to be reconsidered.

The solution to (9.2) can be obtained by an application of the variation-of-constants formula. Therefore, we note that $\mathcal{A}_{\text{ker}}$ is an elliptic operator by Assumptions 9.2.ii). This in turn implies that $-\mathcal{A}_{\text{ker}}$ generates an analytic semigroup on $\mathcal{H}_{\text{ker}}$; see Remark 4.24. Since the semigroup can only be applied to functions in $\mathcal{H}_{\text{ker}} \cong \mathcal{H}_{\text{ker}}^*$, we introduce the operator

$$\iota_{\text{ker}} \colon \mathcal{H}^* \to \mathcal{H}_{\text{ker}}^* \cong \mathcal{H}_{\text{ker}}.$$

This operator is again based on a simple restriction of test functions and leads to the solution formula

$$\begin{aligned} u(t) &= u_{\text{c}}(t) + u_{\text{ker}}(t) \\ &= \mathcal{B}_{\mathcal{A}}^- g(t) + e^{-t\mathcal{A}_{\text{ker}}} u_{\text{ker}}(0) + \int_0^t e^{-(t-s)\mathcal{A}_{\text{ker}}} \iota_{\text{ker}} \big[f(s, u_{\text{ker}}(s) + u_{\text{c}}(s)) - \dot{u}_{\text{c}}(s) \big] \,\mathrm{d}s; \end{aligned}$$

cf. [EmmM13, Sec. 3.2.2] and (4.16). Assuming a partition of the time interval $[0, T]$ by $0 = t_0 < t_1 < \cdots < t_N = T$, we can write the solution formula in the form

$$\begin{aligned} &u(t_{n+1}) - \mathcal{B}_{\mathcal{A}}^- g_{n+1} \\ &\quad = e^{-(t_{n+1}-t_n)\mathcal{A}_{\text{ker}}} \big[u(t_n) - \mathcal{B}_{\mathcal{A}}^- g_n \big] + \int_{t_n}^{t_{n+1}} e^{-(t_{n+1}-s)\mathcal{A}_{\text{ker}}} \iota_{\text{ker}} \big[f(s, u(s)) - \dot{u}_{\text{c}}(s) \big] \,\mathrm{d}s. \end{aligned} \tag{9.3}$$

Note that we use here the abbreviation $g_n := g(t_n)$. In the following sections we construct explicit exponential integrators for constrained semi-linear systems of the form (9.1) by approximating the integral in (9.3). In Sections 9.1 and 9.2 we consider schemes based on the exponential Euler and Runge methods, respectively. The order conditions for schemes up to an order of three are studied in Section 9.3. At first, we consider the approximation of operator ODEs in Subsection 9.3.1. In Subsection 9.3.2 we translate the results to the temporal discretization of operator DAE (9.1). Here, we also approximate the Lagrange multiplier λ. Comments on the efficient computation and numerical experiments for semi-linear parabolic systems illustrating the obtained convergence results are presented in Section 9.4. We recall that we consider in this thesis only explicit integrators. We start with a first-order scheme based on the exponential Euler method.

9.1. The Exponential Euler Scheme

The idea of exponential integrators is to approximate the integral term in (9.3) by an appropriate quadrature rule. Following the construction for PDEs [HocO10], we consider the function evaluation at the beginning of the interval. This then leads to the scheme

$$\begin{aligned} u_{n+1} &= \mathcal{B}_{\mathcal{A}}^- g_{n+1} + e^{-\tau\mathcal{A}_{\text{ker}}} \big[u_n - \mathcal{B}_{\mathcal{A}}^- g_n \big] + \int_0^\tau e^{-(\tau-s)\mathcal{A}_{\text{ker}}} \iota_{\text{ker}} \big[f(t_n, u_n) - \dot{u}_{\text{c}}(t_n) \big] \,\mathrm{d}s \\ &= \mathcal{B}_{\mathcal{A}}^- g_{n+1} + \phi_0(-\tau\mathcal{A}_{\text{ker}}) \big(u_n - \mathcal{B}_{\mathcal{A}}^- g_n \big) + \tau\phi_1(-\tau\mathcal{A}_{\text{ker}}) \big(\iota_{\text{ker}} \big[f(t_n, u_n) - \mathcal{B}_{\mathcal{A}}^- \dot{g}_n \big] \big) \end{aligned} \tag{9.4}$$

with ϕ_0 and ϕ_1 defined as in Section 5.2. As usual, u_n denotes the approximation of $u(t_n)$. Further, for simplicity we restrict ourselves to a uniform partition of $[0,T]$ with step size τ. Assuming that the resulting approximation satisfies the constraint in every step, we have $u_n - \mathcal{B}_{\mathcal{A}}^- g_n \in \mathcal{V}_{\text{ker}} \hookrightarrow \mathcal{H}_{\text{ker}}$ such that the semigroup $e^{-\tau\mathcal{A}_{\text{ker}}}$ is applicable.

The derived formula (9.4) is beneficial for the numerical analysis, but its practical utility depends on the evaluation of the ϕ-functions. In Subsection 9.4.1 we explain a direct implementation. If $e^{-\tau\mathcal{A}_{\text{ker}}}$ is simple to evaluate, the method described in the following section is more favorable.

9.1.1. An Algorithm

Since the evaluation of the ϕ-functions with the operator $\mathcal{A}_{\text{ker}}$ is not straightforward, we reformulate the method in terms of saddle point problems. Furthermore, we need evaluations of $\mathcal{B}_{\mathcal{A}}^-$ applied to the right-hand side g or its time derivative.

Corollary 3.9 shows that these evaluations of $\mathcal{B}_{\mathcal{A}}^-$ can be replaced by the solution of a saddle point problem. As a reminder, for given $g_n \in \mathcal{Q}^*$, the vector $u_{n,\text{c}} = \mathcal{B}_{\mathcal{A}}^- g_n \in \mathcal{V}_\text{c} \subseteq \mathcal{V}$ is equal to the partial solution of

$$\begin{aligned} \mathcal{A}u_{n,\text{c}} - \mathcal{B}^*\nu &= 0 \qquad \text{in } \mathcal{V}^*, && (9.5\text{a})\\ \mathcal{B}u_{n,\text{c}} &= g_n \qquad \text{in } \mathcal{Q}^*. && (9.5\text{b}) \end{aligned}$$

Analogously one calculates $\dot u_{n,\text{c}} = \dot u_\text{c}(t_n) = \mathcal{B}_{\mathcal{A}}^- \dot g_n$. The Lagrange multiplier ν in (9.5) is not of particular interest and simply serves as a dummy variable. For a saddle point problem, which can be used to approximate the Lagrange multiplier $\lambda(t_n)$ for a given $u_n \approx u(t_n)$, we refer to Subsection 9.3.2.2.

Being able to compute $\mathcal{B}_{\mathcal{A}}^- g_n$ and $\mathcal{B}_{\mathcal{A}}^- \dot g_n$, we are now interested in the solution of problems involving the operator $\mathcal{A}_{\text{ker}}$. This is helpful for the reformulation of the exponential Euler method (9.4). We introduce the auxiliary variable $w_n \in \mathcal{V}_{\text{ker}}$ as the solution of

$$\mathcal{A}_{\text{ker}} w_n = f(t_n, u_n) - \dot u_\text{c}(t_n) = f(t_n, u_n) - \mathcal{B}_{\mathcal{A}}^- \dot g_n \qquad \text{in } \mathcal{V}_{\text{ker}}^*.$$

Then w_n is again equivalent to a partial solution of a stationary saddle point problem, namely

$$\begin{aligned} \mathcal{A}w_n - \mathcal{B}^*\nu_n &= f(t_n, u_n) - \mathcal{B}_{\mathcal{A}}^- \dot g_n \qquad \text{in } \mathcal{V}^*, && (9.6\text{a})\\ \mathcal{B}w_n &= 0 \qquad \text{in } \mathcal{Q}^*. && (9.6\text{b}) \end{aligned}$$

As for (9.5), the Lagrange multiplier is only introduced for a proper formulation and not of particular interest in the following. The unique solvability of system (9.6) follows by Theorem 3.8, since $f(t_n, u_n) - \mathcal{B}_{\mathcal{A}}^- \dot g_n \in \mathcal{V}^*$. In order to rewrite (9.4), we further note that the recursion formula (5.8) for ϕ_1 implies

$$\tau\phi_1(-\tau\mathcal{A}_{\text{ker}})h = -\big[\phi_0(-\tau\mathcal{A}_{\text{ker}}) - \text{id}\,\big]\mathcal{A}_{\text{ker}}^{-1} h$$

for all $h \in \mathcal{H}_{\text{ker}}^* \cong \mathcal{H}_{\text{ker}}$. Recall that $\mathcal{A}_{\text{ker}}$ is indeed invertible due to Assumption 9.2.ii). Thus, the exponential Euler scheme can be rewritten as

$$u_{n+1} = \mathcal{B}_{\mathcal{A}}^- g_{n+1} + \phi_0(-\tau\mathcal{A}_{\text{ker}})\big(u_n - \mathcal{B}_{\mathcal{A}}^- g_n - w_n\big) + w_n.$$

Finally, we need a way to compute the action of $\phi_0(-\tau\mathcal{A}_{\text{ker}})$. For this, we consider the underlying operator DAE formulation,

$$\begin{aligned} \dot z(t) + \mathcal{A}z(t) - \mathcal{B}^*\mu(t) &= 0 \qquad \text{in } \mathcal{V}^*, && (9.7\text{a})\\ \mathcal{B}z(t) &= 0 \qquad \text{in } \mathcal{Q}^* && (9.7\text{b}) \end{aligned}$$

with initial condition $z(t_n) = u_n - \mathcal{B}_{\mathcal{A}}^- g_n - w_n$. The resulting method then reads $u_{n+1} = \mathcal{B}_{\mathcal{A}}^- g_{n+1} + z(t_{n+1}) + w_n$. Thus, the exponential Euler scheme given in (9.4) can be computed by solving a number of saddle point problems. We summarize the necessary steps in Algorithm 1.

Algorithm 1 Exponential Euler Scheme for Operator DAE (9.1)

1: **Input**: step size τ, consistent initial data $u_0 \in \mathcal{V}$, right-hand sides f, g with $g_n := g(t_n)$ and $\dot{g}_n := \dot{g}(t_n)$
2: **for** $n = 0$ **to** $N-1$ **do**
3: compute $\mathcal{B}_{\mathcal{A}}^- g_n$, $\mathcal{B}_{\mathcal{A}}^- g_{n+1}$, and $\mathcal{B}_{\mathcal{A}}^- \dot{g}_n$ by (9.5)
4: compute w_n by (9.6)
5: compute z as solution of (9.7) on $[t_n, t_{n+1}]$ with initial data $u_n - \mathcal{B}_{\mathcal{A}}^- g_n - w_n$
6: set $u_{n+1} = \mathcal{B}_{\mathcal{A}}^- g_{n+1} + z(t_{n+1}) + w_n$
7: **end for**

Remark 9.4. One step of the exponential Euler scheme consists of the solution of three (in the first step four) stationary – twice (9.5) and once (9.6) – and a single transient saddle point problem (9.7), including only one evaluation of the nonlinear function f in total. We emphasize that all these systems are linear such that no Newton iteration is necessary in the solution process. Furthermore, the time-dependent system (9.7) is homogeneous such that it can be solved without the need of a regularization; cf. Remark 7.6.

9.1.2. Convergence Analysis

In this section we analyze the convergence order of the exponential Euler method for constrained PDEs of parabolic type (9.1). For the unconstrained case it is well-known that the convergence order is one, see e.g. [HocO05a, Sec. 4.2]. Since our approach is based on the unconstrained PDE (9.2) of the dynamical part in $\mathcal{V}_{\text{ker}}$, we expect the same order for the solution of Algorithm 1. For the associated proof we assume that the approximation u_n lies within a strip of radius r around u, where f is locally Lipschitz continuous with constant $L > 0$. By Lemma 6.18 there exists such a uniform radius and local Lipschitz constant. Furthermore, a sufficiently small step size τ guarantees that u_n stays within this strip around u, since the solution z of (9.7) and $\mathcal{B}_{\mathcal{A}}^- g$ are continuous by Theorem 6.9 and 3.38.

Theorem 9.5 (Exponential Euler)**.** *Suppose that Assumptions 9.1–9.3 are fulfilled and $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}u_0 = g(0)$. Further, let the step size τ be sufficiently small such that the derived approximation u_n lies within a strip along u, in which f is locally Lipschitz continuous with a uniform constant $L > 0$. For the right-hand side of the constraint we assume $g \in H^2(0, T; \mathcal{Q}^*)$. If the exact solution of (9.1) satisfies $\frac{\mathrm{d}}{\mathrm{d}t} f(\cdot, u(\cdot)) \in L^2(0, T; \mathcal{H}^*)$, then the approximation u_n obtained by the exponential Euler scheme of Algorithm 1 satisfies*

$$\|u_n - u(t_n)\|_{\mathcal{V}}^2 \;\lesssim\; \tau^2 \int_0^{t_n} \|\tfrac{\mathrm{d}}{\mathrm{d}t} f(t, u(t))\|_{\mathcal{H}^*}^2 + \|\mathcal{B}_{\mathcal{A}}^- \ddot{g}(t)\|_{\mathcal{H}^*}^2 \,\mathrm{d}t.$$

The involved constant only depends on t_n, L, and the operator $\mathcal{A}$.

Proof. With the constant w_n and function z from (9.6) and (9.7), respectively, we define $v(t) := z(t) + w_n + \mathcal{B}_{\mathcal{A}}^- g(t)$ for $t \in [t_n, t_{n+1}]$, $n = 0, \dots, N-1$. This function satisfies

$$v(t_n) = z(t_n) + w_n + \mathcal{B}_{\mathcal{A}}^- g_n = u_n \quad \text{and} \quad v(t_{n+1}) = z(t_{n+1}) + w_n + \mathcal{B}_{\mathcal{A}}^- g_{n+1} = u_{n+1}.$$

Furthermore, since $\dot v(t) = \dot z(t) + \mathcal{B}_\mathcal{A}^-\dot g(t)$, the function v solves the operator DAE

$$\begin{aligned} \dot v(t) + \mathcal{A}v(t) - \mathcal{B}^*\gamma(t) &= f(t_n,u_n) + \mathcal{B}_\mathcal{A}^-(\dot g(t) - \dot g_n) \qquad &\text{in } \mathcal{V}^*,\\ \mathcal{B}v(t) &= g(t) &\text{in } \mathcal{Q}^* \end{aligned}$$

on $[t_n, t_{n+1}]$, $n = 0,\dots,N-1$ with initial value $v(t_0) = u_0$. To shorten notation we define $\Delta u \coloneqq u - v$ and $\Delta\lambda \coloneqq \lambda - \gamma$, which satisfy

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\Delta u + \mathcal{A}_1\Delta u - \mathcal{B}^*\Delta\lambda &= f(\cdot,u(\cdot)) - f(t_n,u_n) - \mathcal{A}_2\Delta u - \mathcal{B}_\mathcal{A}^-\big(\dot g - \dot g_n\big) \qquad &\text{in } \mathcal{V}^*,\\ \mathcal{B}\Delta u &= 0 &\text{in } \mathcal{Q}^* \end{aligned}$$

on each interval $[t_n, t_{n+1}]$ with initial value $\Delta u(t_0) = 0$ if $n = 0$ and $\Delta u(t_n) = u(t_n) - u_n$ otherwise. In the following, we derive estimates of Δu on all sub-intervals. For this, we assume without loss of generality that the operator $\mathcal{A}_1$ is elliptic on $\mathcal{V}_\mathrm{ker}$ with ellipticity constant $\mu_{\mathcal{A}_1}$; see page 31. Starting with $n = 0$, by Theorem 4.25 with $t = t_1 = \tau$ and $\omega \coloneqq 2C_{\mathcal{A}_2}^2\mu_{\mathcal{A}_1}^{-1}$ we obtain the bound

$$\begin{aligned} \|u(t_1) - u_1\|_{\mathcal{A}_1}^2 &\overset{(4.26)}{\le} 2\,\exp(\omega\tau)\int_0^\tau \Big\| \int_0^s \tfrac{\mathrm{d}}{\mathrm{d}\eta} f(\eta,u(\eta)) - \mathcal{B}_\mathcal{A}^-\ddot g(\eta)\,\mathrm{d}\eta\Big\|_{\mathcal{H}^*}^2 \,\mathrm{d}s\\ &\le\; 2\,\exp(\omega\tau)\int_0^\tau s\int_0^s \|\tfrac{\mathrm{d}}{\mathrm{d}\eta} f(\eta,u(\eta)) - \mathcal{B}_\mathcal{A}^-\ddot g(\eta)\|_{\mathcal{H}^*}^2\,\mathrm{d}\eta\,\mathrm{d}s\\ &\le\; 2\,\exp(\omega\tau)\tau^2 \underbrace{\int_0^\tau \|\tfrac{\mathrm{d}}{\mathrm{d}s} f(s,u(s))\|_{\mathcal{H}^*}^2 + \|\mathcal{B}_\mathcal{A}^-\ddot g(s)\|_{\mathcal{H}^*}^2\,\mathrm{d}s}_{=:\,\mathcal{I}(\frac{\mathrm{d}}{\mathrm{d}t}f,\,\ddot g,\,0,\,t_1)}. \end{aligned} \tag{9.8}$$

With the uniform Lipschitz constant L we have for $n \ge 1$ that

$$\begin{aligned} &\int_{t_n}^{t_{n+1}} \|f(s,u(s)) - f(t_n,u_n)\|_{\mathcal{H}^*}^2\,\mathrm{d}s\\ &\qquad\le 2\int_{t_n}^{t_{n+1}} \|f(t_n,u(t_n)) - f(t_n,u_n)\|_{\mathcal{H}^*}^2 + \|f(s,u(s)) - f(t_n,u(t_n))\|_{\mathcal{H}^*}^2\,\mathrm{d}s\\ &\qquad\le 2\,\tau\tfrac{L^2}{\mu_{\mathcal{A}_1}}\,\|u(t_n) - u_n\|_{\mathcal{A}_1}^2 + 2\int_{t_n}^{t_{n+1}}(s - t_n)\int_{t_n}^s \|\tfrac{\mathrm{d}}{\mathrm{d}\eta} f(\eta,u(\eta))\|_{\mathcal{H}^*}^2\,\mathrm{d}\eta\,\mathrm{d}s. \end{aligned}$$

This estimate together with Young's inequality (3.8) lead to the bound

$$\|u(t_{n+1}) - u_{n+1}\|_{\mathcal{A}_1}^2 \le \exp(\omega\tau)\Big[\big(1 + 3\,\tau\tfrac{L^2}{\mu_{\mathcal{A}_1}}\big)\|u(t_n) - u_n\|_{\mathcal{A}_1}^2 + 3\,\tau^2\,\mathcal{I}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}f,\ddot g,t_n,t_{n+1}\big)\Big] \tag{9.9}$$

similarly as in (9.8). With $(1 + x) \le \exp(x)$, estimate (9.8), and an iterative application of the estimate (9.9) we get

$$\begin{aligned} \|u(t_{n+1}) - u_{n+1}\|_{\mathcal{A}_1}^2 &\le \tau^2\;3\sum_{k=0}^n \exp(\omega\tau)^{n+1-k}\big(1 + 3\,\tau\tfrac{L^2}{\mu_{\mathcal{A}_1}}\big)^{n-k}\,\mathcal{I}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}f,\ddot g,t_k,t_{k+1}\big)\\ &\le \tau^2\;3\,\exp(\omega\,t_{n+1})\exp\big(3\tfrac{L^2}{\mu_{\mathcal{A}_1}}t_n\big)\,\mathcal{I}\big(\tfrac{\mathrm{d}}{\mathrm{d}t}f,\ddot g,0,t_{n+1}\big) \end{aligned}$$

for all $n = 0,\dots,N-1$. The stated estimate finally follows by the equivalence of $\|\cdot\|_\mathcal{V}$ and $\|\cdot\|_{\mathcal{A}_1}$ on $\mathcal{V}_\mathrm{ker}$; see (4.25). □

Remark 9.6. Since $\phi_0(-\mathcal{A}t)$ is calculated exactly, the assumption of Theorem 9.5 on the step size τ depends not on the operator $\mathcal{A}$ but only on the nonlinearity f such that the approximation stays inside the strip along u. Thus, the step size does not depend on the stiffness of the system and is still allowed to be large.

Remark 9.7. In the case of a self-adjoint operator $\mathcal{A}$, i.e., $\mathcal{A}_2 = 0$, the convergence result can also be proven by the restriction to test functions in $\mathcal{V}_{\ker}$ and the application of corresponding results for the unconstrained case, namely [HocO10, Th. 2.14]. This requires similar assumptions but with $\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\cdot, u(\cdot)) \in L^\infty(0,T;\mathcal{H}^*)$. Anyway, we like to emphasize that this procedure is also applicable if $\mathcal{A}_2 \neq 0$ by moving $\mathcal{A}_2$ into the nonlinearity f. This, however, slightly changes the proposed scheme, since then only $\mathcal{A}_2 u_n$ enters the approximation instead of $\mathcal{A}_2 u(t)$. In practical applications this would also need to find the symmetric part of the differential operator $\mathcal{A}$, which is still elliptic on $\mathcal{V}_{\ker}$.

9.2. Two-Stage Methods

This section is devoted to the construction of an explicit exponential integrator with two stages and an order of one and a half for constrained parabolic systems. In particular, we aim to transfer the method given in [StrWP12, Exp. 11.2.2] to the operator DAE (9.1). This explicit exponential integrator is described by its Butcher tableau

$$\begin{array}{c|cc} 0 & & \\ 1 & \phi_1 & \\ \hline & \phi_1 - \phi_2 & \phi_2 \end{array} \tag{9.10}$$

and is a special case of the exponential Runge methods from Subsection 9.2.3. In the unconstrained case, i.e., for $\dot{v} + \mathcal{A}_{\ker} v = \tilde{f}(t,v)$ in $\mathcal{V}_{\ker}^*$, one step of this method is defined through

$$v_{n+1}^{\mathrm{Eul}} := \phi_0(-\tau\mathcal{A}_{\ker}) v_n + \tau\phi_1(-\tau\mathcal{A}_{\ker}) \tilde{f}(t_n, v_n), \tag{9.11a}$$
$$v_{n+1} := v_{n+1}^{\mathrm{Eul}} + \tau\phi_2(-\tau\mathcal{A}_{\ker}) \big[\tilde{f}(t_{n+1}, v_{n+1}^{\mathrm{Eul}}) - \tilde{f}(t_n, v_n)\big]. \tag{9.11b}$$

Similarly as for the exponential Euler method, we define a number of auxiliary problems in order to obtain an applicable method for parabolic systems with constraints.

9.2.1. An Algorithm

We translate the numerical scheme (9.11) to the constrained case. Let u_n denote the given approximation of $u(t_n)$. Then the first step is to perform one step of the exponential Euler method, cf. Algorithm 1, leading to u_{n+1}^{Eul}. Second, we compute w_n' as the solution of the stationary problem

$$\begin{aligned} \mathcal{A} w_n' - \mathcal{B}^* \nu_n' &= f(t_{n+1}, u_{n+1}^{\mathrm{Eul}}) - \mathcal{B}_{\mathcal{A}}^- \dot{g}_{n+1} - f(t_n, u_n) + \mathcal{B}_{\mathcal{A}}^- \dot{g}_n \quad && \text{in } \mathcal{V}*, && (9.12a) \\ \mathcal{B} w_n' &= 0 && \text{in } \mathcal{Q}* && (9.12b) \end{aligned}$$

and w_n'' as the solution of

$$\begin{aligned} \mathcal{A} w_n'' - \mathcal{B}^* \nu_n'' &= \tfrac{1}{\tau} w_n' \quad && \text{in } \mathcal{V}*, && (9.13a) \\ \mathcal{B} w_n'' &= 0 && \text{in } \mathcal{Q}*. && (9.13b) \end{aligned}$$

Note that, due to the recursion formula (5.8), w_n' and w_n'' satisfy the identity

$$\begin{aligned} \tau\phi_2(-\tau\mathcal{A}_{\ker})\, \iota_{\ker} \big[f(t_{n+1}, u_{n+1}^{\mathrm{Eul}}) - \mathcal{B}_{\mathcal{A}}^- \dot{g}_{n+1} - f(t_n, u_n) + \mathcal{B}_{\mathcal{A}}^- \dot{g}_n\big] &= -\phi_1(-\tau\mathcal{A}_{\ker}) w_n' + w_n' \\ &= \phi_0(-\tau\mathcal{A}_{\ker}) w_n'' - w_n'' + w_n'. \end{aligned}$$

It remains to compute $\phi_0(-\tau\mathcal{A}_{\text{ker}})w_n''$ and thus, to consider the homogeneous system (9.7) on the time interval $[t_n, t_{n+1}]$ with initial value $z(t_n) = w_n''$. The solution z at time t_{n+1} then defines the new approximation by

$$u_{n+1} := u_{n+1}^{\text{Eul}} + z(t_{n+1}) - w_n'' + w_n'.$$

Note that the consistency is already guaranteed by the exponential Euler step, which yields $\mathcal{B}u_{n+1} = \mathcal{B}u_{n+1}^{\text{Eul}} = g_{n+1}$. The resulting exponential integrator is summarized in Algorithm 2.

Algorithm 2 Exponential Integrator (9.10) for Operator DAE (9.1)

1: **Input**: step size τ, consistent initial data $u_0 \in \mathcal{V}$, right-hand sides f, g with $\dot{g}_n := \dot{g}(t_n)$
2: **for** $n = 0$ **to** $N-1$ **do**
3: compute with Algorithm 1 one step of the exponential Euler method for u_n leading to u_{n+1}^{Eul}
4: compute $\mathcal{B}_{\mathcal{A}}^{-}\dot{g}_n$ and $\mathcal{B}_{\mathcal{A}}^{-}\dot{g}_{n+1}$ by (9.5)
5: compute w_n' by (9.12)
6: compute w_n'' by (9.13)
7: compute z as solution of (9.7) on $[t_n, t_{n+1}]$ with initial condition $z(t_n) = w_n''$
8: set $u_{n+1} = u_{n+1}^{\text{Eul}} + z(t_{n+1}) - w_n'' + w_n'$
9: **end for**

9.2.2. Convergence Analysis

In this subsection we investigate the convergence order of Algorithm 2 when applied to operator DAEs of the form (9.1). For PDEs it is well-known that the exponential integrator given by the Butcher tableau (9.10) has a convergence order of one and a half if we assume $\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\cdot, u(\cdot)) \in L^\infty(0,T;\mathcal{H}^*)$; cf. [HocO05a, Th. 4.3]. This carries over to the operator DAE case.

Theorem 9.8 (Two-Stage Method)**.** *Suppose that Assumptions 9.1–9.3 are fulfilled and $u_0 \in \mathcal{V}$ is consistent, i.e., $\mathcal{B}u_0 = g(0)$. Let the step size τ be sufficiently small such that the discrete solution u_n and the internal stage u_n^{Eul} lie in a strip along u, where f is locally Lipschitz continuous with a uniform constant $L > 0$. Further assume $g \in H^3(0,T;\mathcal{Q}^*)$. If the exact solution of (9.1) satisfies $f(\,\cdot\,, u(\,\cdot\,)) \in H^2(0,T;\mathcal{H}^*)$, then the approximation u_n obtained by Algorithm 2 satisfies the error bound*

$$\|u_n - u(t_n)\|_{\mathcal{V}}^2 \lesssim \int_0^{t_n} \tau^3\big(\|\tfrac{\mathrm{d}}{\mathrm{d}t} f(t,u(t))\|_{\mathcal{H}^*}^2 + \|\mathcal{B}_{\mathcal{A}}^{-}\ddot{g}(t)\|_{\mathcal{H}^*}^2\big) + \tau^4\big(\|\tfrac{\mathrm{d}^2}{\mathrm{d}t^2} f(t,u(t))\|_{\mathcal{H}^*}^2 + \|\mathcal{B}_{\mathcal{A}}^{-}\tfrac{\mathrm{d}^3}{\mathrm{d}t^3} g(t)\|_{\mathcal{H}^*}^2\big)\,\mathrm{d}t.$$

The involved constant only depends on t_n, L, and the operator $\mathcal{A}$.

Proof. Let v^{Eul} be the function constructed in the proof of Theorem 9.5, which satisfies $v^{\text{Eul}}(t_n) = u_n$ and $v^{\text{Eul}}(t_{n+1}) = u_{n+1}^{\text{Eul}}$, and set $v(t) := v^{\text{Eul}}(t) + z(t) - w_n'' + \frac{t-t_n}{\tau} w_n'$ with z as in Step 7 in Algorithm 2. This function satisfies

$$v(t_n) = v^{\text{Eul}}(t_n) = u_n, \qquad v(t_{n+1}) = v^{\text{Eul}}(t_{n+1}) + z(t_{n+1}) - w_n'' + w_n' = u_{n+1}.$$

Note that the estimates (9.8) and (9.9) are still valid if one replaces u_{n+1} by $v^{\text{Eul}}(t_{n+1})$ on the left-hand side of these estimates. As in the proof of Theorem 9.5, we can interpret v as the solution of an operator DAE on $[t_n, t_{n+1}]$. The corresponding right-hand sides are then given by

$$f(t_n, u_n) + \tfrac{t-t_n}{\tau}\big(f(t_{n+1}, v^{\text{Eul}}(t_{n+1})) - f(t_n, u_n)\big) + \mathcal{B}_{\mathcal{A}}^{-}\big(\dot{g}(t) - \dot{g}_n - \tfrac{t-t_n}{\tau}(\dot{g}_{n+1} - \dot{g}_n)\big)$$

for the dynamic equation and $g(t)$ for the constraint. By Young's inequality, Theorem 4.25, and an error bound of the right-hand side by Taylor expansions we then get

$$\begin{aligned} \|u(t_{n+1}) - u_{n+1}\|^2_{\mathcal{A}_1} \leq \exp(\omega\tau) \Big[&\big(1 + 4\tau \tfrac{L^2}{\mu_{\mathcal{A}_1}}\big) \|u(t_n) - u_n\|^2_{\mathcal{A}_1} + 4\,\tau \tfrac{L^2}{\mu} \|(u - v^{\mathrm{Eul}})(t_{n+1})\|^2_{\mathcal{A}_1} \\ &+ \tau^4 \int_{t_n}^{t_{n+1}} \tfrac{2}{15} \|\tfrac{\mathrm{d}^2}{\mathrm{d}t^2} f(t, u(t))\|^2_{\mathcal{H}} + \tfrac{2}{45} \|\mathcal{B}^-_{\mathcal{A}} \tfrac{\mathrm{d}^3}{\mathrm{d}t^3} g(t)\|^2_{\mathcal{H}} \,\mathrm{d}t \Big] \end{aligned}$$

with $\omega = 2C^2_{\mathcal{A}_2}\mu^{-1}_{\mathcal{A}_1}$. The stated error bound then follows by an iterative application of the previous estimate together with the estimates (9.8), (9.9) and the norm equivalence of $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{A}_1}$. □

In Subsection 9.3.2 we prove that the convergence order can improve to two if $t \mapsto f(t, u(t))$ and its derivatives map into $\mathcal{V}$. The authors of [HocO05a] analyze the PDE case where $t \mapsto f(t, u(t))$ maps into a interpolation space $[\mathcal{V}, \mathcal{H}]_\theta$, $\theta \in [0, 1]$; see [LioM72, Ch. 1, Sec. 2] for an introduction of interpolation spaces. The convergence rate then is $\frac{4-\theta}{2}$. We close this section with remarks on alternative two-stage schemes.

9.2.3. Exponential Runge Methods

The analyzed scheme (9.10) is a special case of a one-parameter family of exponential Runge-Kutta methods described by the tableau

$$\begin{array}{c|cc} 0 & & \\ c_2 & c_2\phi_{1,2} & \\ \hline & \phi_1 - \frac{1}{c_2}\phi_2 & \frac{1}{c_2}\phi_2 \end{array} \tag{9.14}$$

with positive parameter $c_2 > 0$; cf. [HocO10]. Here, $\phi_{1,2}(\cdot)$ is defined by $\phi_1(c_2\,\cdot)$; see (5.11). The members of the family are called *exponential Runge methods*. Note that, for $c_2 > 1$ we must assume that f and $t \mapsto f(t, u(t))$ have regular extension for $t > T$; cf. Theorem 8.37. However, we regain (9.10) for $c_2 = 1$. For $c_2 \neq 1$, the resulting scheme for constrained systems calls for two additional saddle point problems in order to compute $\mathcal{B}^-_{\mathcal{A}} g(t_n + c_2\tau)$ and $\mathcal{B}^-_{\mathcal{A}} \dot{g}(t_n + c_2\tau)$. This then leads to an exponential integrator summarized in Algorithm 3 with the abbreviations

$$g_{n,2} := g(t_n + c_2\tau), \qquad \dot{g}_{n,2} := \dot{g}(t_n + c_2\tau), \qquad t_{n,2} := t_n + c_2\tau.$$

We emphasize that the convergence result of Theorem 9.8 transfers to this family of integrators.

Algorithm 3 Exponential Runge Integrators for Operator DAE (9.1)

Input: step size τ, consistent initial data $u_0 \in \mathcal{V}$, right-hand sides f, g with $g_n := g(t_n)$, $g_{n,2} := g(t_n + c_2\tau)$, $\dot{g}_n := \dot{g}(t_n)$, $\dot{g}_{n,2} := \dot{g}(t_n + c_2\tau)$
for $n = 0$ **to** $N - 1$ **do**
  compute $\mathcal{B}^-_{\mathcal{A}} g_n$, $\mathcal{B}^-_{\mathcal{A}} g_{n,2}$, $\mathcal{B}^-_{\mathcal{A}} g_{n+1}$, $\mathcal{B}^-_{\mathcal{A}} \dot{g}_n$, $\mathcal{B}^-_{\mathcal{A}} \dot{g}_{n,2}$, and $\mathcal{B}^-_{\mathcal{A}} \dot{g}_{n+1}$ by (9.5)
  compute w_n by (9.6)
  solve (9.7) on $[t_n, t_{n,2}]$ with initial condition $z(t_n) = u_n - \mathcal{B}^-_{\mathcal{A}} g_n - w_n$
  set $u_{n,2} = z(t_{n,2}) + w_n + \mathcal{B}^-_{\mathcal{A}} g_{n,2}$
  compute w'_n by (9.12) with right-hand side $\frac{1}{c_2}\big(f(t_{n,2}, u_{n,2}) - f(t_n, u_n) - \mathcal{B}^-_{\mathcal{A}} \dot{g}_{n,2} + \mathcal{B}^-_{\mathcal{A}} \dot{g}_n\big)$
  compute w''_n by (9.13)
  solve (9.7) on $[t_n, t_{n+1}]$ with initial condition $z(t_n) = u_n - \mathcal{B}^-_{\mathcal{A}} g_n - w_n + w''_n$
  set $u_{n+1} = z(t_{n+1}) + w_n + w'_n - w''_n + \mathcal{B}^-_{\mathcal{A}} g_{n+1}$.
end for

9.3. Order Conditions for Schemes of Order up to Three

In the previous two sections we constructed explicit exponential integrators of convergence order up to one and a half. The associated proofs needed the local Lipschitz continuity of the nonlinearity but nothing more. In [HocO05b] the authors prove under the same condition the existence of so called *exponential Runge-Kutta method of collocation type* of arbitrary order. However, the investigated methods are of collocation type and therefore implicit [AscP98, p. 101]. In this thesis we are interested in explicit exponential integrators. In this case, we have to assume more regularity of f and $t \mapsto f(t, u(t))$ if we want to overcome the bound of one and a half as convergence order. For semi-linear PDEs with a nonlinearity with domain $[0, T] \times \mathcal{H}$ the necessity of this additional assumption is investigated in [HocO05a, Sec. 4.4].

Assumption 9.9. Let the function $t \mapsto f(t, u(t))$ be sufficiently many times differentiable with images in $\mathcal{H}^*$. Suppose f is sufficiently many times Fréchet-differentiable with uniform bounded derivatives in a strip along the solution u.

Before we can analyze the order of exponential integrators of higher order for operator DAEs we have to investigate the unconstrained case.

9.3.1. Operator Differential Equations

In this subsection we consider the semi-linear operator ODE

$$\dot{u}(t) + \mathcal{A}u(t) = f(t, u(t)) \quad \text{ in } \mathcal{V}^* \tag{9.15}$$

and its temporal discretization by an arbitrary explicit exponential integrator. The approximation u_n and the internal stages $U_{n,i}$ are determined by (5.10), i.e.,

$$u_{n+1} = e^{-\tau A} u_n + \tau \sum_{i=1}^{s} b_i(-\tau A) f(t_n + c_i\tau, U_{n,i}) \tag{9.16a}$$

$$U_{n,i} = e^{-c_i \tau A} u_n + \tau \sum_{j=1}^{i-1} a_{i,j}(-\tau A) f(t_n + c_j\tau, U_{n,j}), \tag{9.16b}$$

$i = 1, \dots, s$. Here, $c_i \in \mathbb{R}_{\geq 0}$ and the functions b_i, $a_{i,j}$ are given by the Butcher tableau (5.12). For the sake of simplicity, we assume in Section 9.3 that $c_i \leq 1$. Otherwise, we have to extend the function f for $t > T$.

The main goal is to estimate the global error

$$e_n := u_n - u(t_n). \tag{9.17}$$

For our investigation we combine the main ideas of [HocO05a] and [LuaO14a]. We want to point out that the authors of [HocO05a; LuaO14a] study the order condition of exponential integrators with convergence order higher than two for the special case that the non-linearity f maps from $[0, T] \times \mathcal{H}$ to $\mathcal{H}$ and where the error is measured in the $\mathcal{H}$-norm. We overcome this restriction in our investigation.

In the following, we make use of the local approximation

$$\widehat{u}_{n+1} = e^{-\tau \mathcal{A}} u(t_n) + \tau \sum_{i=1}^{s} b_i(-\tau \mathcal{A}) f(t_n + c_i\tau, \widehat{U}_{n,i}), \tag{9.18}$$

i.e., instead of making one step with u_n as starting value we use $u(t_n)$. The associated internal stages $\widehat{U}_{n,i}$ are analogously defined by (9.16b). In addition to the global errors (9.17), we introduce the differences

$$\widetilde{e}_{n+1} := u_{n+1} - \widehat{u}_{n+1}, \qquad \widehat{e}_{n+1} := \widehat{u}_{n+1} - u(t_{n+1}), \qquad \widehat{E}_{n,i} := \widehat{U}_{n,i} - u(t_n + c_i\tau). \tag{9.19}$$

Since $e_{n+1} = \widetilde{e}_{n+1} + \widehat{e}_{n+1}$, we can treat the two parts $\widetilde{e}_{n+1}$ and $\widehat{e}_{n+1}$ separately. Furthermore, we introduce for a shorter notation

$$f_u(t) = f(t, u(t)).$$

Let us start with the local error $\widehat{e}_{n+1}$. As a first step, we represent the exact solution by the variation-of-constants formula

$$u(t_{n+1}) = e^{-\tau\mathcal{A}} u(t_n) + \int_0^\tau e^{-(\tau-s)\mathcal{A}} f_u(t_n + s)\,\mathrm{d}s.$$

A Taylor expansion of $f_u(t_n + s)$ around t_n gives us the representation

$$u(t_{n+1}) = e^{-\tau\mathcal{A}} u(t_n) + \sum_{j=1}^{q} \tau^j \phi_j(-\tau\mathcal{A}) f_u^{(j-1)}(t_n) + \int_0^\tau e^{-(\tau-s)\mathcal{A}} \int_0^s \frac{(s-\eta)^{q-1}}{(q-1)!} f_u^{(q)}(t_n+\eta)\,\mathrm{d}\eta\,\mathrm{d}s. \tag{9.20}$$

On the other hand, if we insert the analytic solution u in our numerical scheme and use again a Taylor expansion we get with a defect R_{n+1} the identity

$$\begin{aligned} &u(t_{n+1}) \\ &= e^{-\tau\mathcal{A}} u(t_n) + \tau \sum_{i=1}^{s} b_i(-\tau\mathcal{A}) f_u(t_n + c_i\tau) + R_{n+1} \qquad (9.21) \\ &= e^{-\tau\mathcal{A}} u(t_n) + \sum_{i=1}^{s} b_i(-\tau\mathcal{A}) \left[\sum_{j=1}^{q} \tau^j \frac{c_i^{j-1}}{(j-1)!} f_u^{(j-1)}(t_n) + \tau \int_0^{c_i\tau} \frac{(c_i\tau - s)^{q-1}}{(q-1)!} f_u^{(q)}(t_n+s)\,\mathrm{d}s \right] + R_{n+1}. \end{aligned}$$

With the ω-functions defined as

$$\omega_j(-\tau\mathcal{A}) := \phi_j(-\tau\mathcal{A}) - \sum_{i=1}^{s} b_i(-\tau\mathcal{A}) \frac{c_i^{j-1}}{(j-1)!}, \tag{9.22}$$

$j = 1, 2, \ldots$, the expansions (9.20) and (9.21) of $u(t_{n+1})$ leads to the representation of the defect as

$$R_{n+1} = \sum_{j=1}^{q} \tau^j \omega_j(-\tau\mathcal{A}) f_u^{(j-1)}(t_n) + R_{n+1}^{[q]}.$$

The remaining integrals $R_{n+1}^{[q]}$ are

$$R_{n+1}^{[q]} = \int_0^\tau e^{-(\tau-s)\mathcal{A}} \int_0^s \frac{(s-\eta)^{q-1}}{(q-1)!} f_u^{(q)}(t_n+\eta)\,\mathrm{d}\eta\,\mathrm{d}s - \tau \sum_{i=1}^{s} b_i(-\tau\mathcal{A}) \int_0^{c_i\tau} \frac{(c_i\tau - s)^{q-1}}{(q-1)!} f_u^{(q)}(t_n+s)\,\mathrm{d}s.$$

Analogously we can rewrite the internal stages by

$$u(t_n + c_i\tau) = e^{-c_i\tau\mathcal{A}} u(t_n) + \tau \sum_{j=1}^{i-1} a_{i,j}(-\tau\mathcal{A}) f_u(t_n + c_j\tau) + \boldsymbol{R}_{n,i}, \tag{9.23}$$

where we expand the defect $\boldsymbol{R}_{n,i}$ as $\sum_{k=1}^{r} \tau^k \psi_{k,i}(-\tau\mathcal{A}) f_u^{(k-1)}(t_n) + \boldsymbol{R}_{n,i}^{[r]}$. The function $\psi_{k,i}$ are defined as

$$\psi_{k,i}(-\tau\mathcal{A}) \coloneqq \phi_k(-c_i\tau\mathcal{A})c_i^k - \sum_{\ell=1}^{i-1} a_{i,\ell}(-\tau\mathcal{A})\frac{c_\ell^{k-1}}{(k-1)!} \tag{9.24}$$

and the remainder $\boldsymbol{R}_{n,i}^{[r]}$ is given by

$$\begin{aligned}\boldsymbol{R}_{n,i}^{[r]} = \int_0^{c_i\tau} e^{-(c_i\tau-s)\mathcal{A}} \int_0^s \frac{(s-\eta)^{r-1}}{(r-1)!} f_u^{(r)}(t_n+\eta)\,\mathrm{d}\eta\,\mathrm{d}s \\ - \tau\sum_{j=1}^{i-1} a_{i,j}(-\tau\mathcal{A}) \int_0^{c_j\tau} \frac{(c_j\tau-s)^{r-1}}{(r-1)!} f_u^{(r)}(t_n+s)\,\mathrm{d}s.\end{aligned} \tag{9.25}$$

We determine bounds for the two remainders $R_{n+1}^{[r]}$ and $\boldsymbol{R}_{n,i}^{[r]}$, which we use later in this subsection for the order conditions. Note that, we use the assumption $c_i \leq 1$ in the following lemma.

Lemma 9.10. *Let $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ be elliptic and satisfy Assumption 9.2.i). Suppose that the right-hand side $f_u^{(r)}$ is an element of $L^\infty(0,T;\mathcal{H}^*)$. Then the remainders $R_{n+1}^{[r]}$ and $\boldsymbol{R}_{n,i}^{[r]}$ satisfy*

$$\big\|\boldsymbol{R}_{n,i}^{[r]}\big\|_{\mathcal{H}} \lesssim \tau^{r+1} \operatorname*{ess\,sup}_{\zeta\in[0,1]} \big\|f_u^{(r)}(t_n+\tau\zeta)\big\|_{\mathcal{H}^*}, \tag{9.26}$$

$$\sqrt{\tau}\big\|\boldsymbol{R}_{n,i}^{[r]}\big\|_{\mathcal{V}} \lesssim (1+\sqrt{\tau})\tau^{r+1} \operatorname*{ess\,sup}_{\zeta\in[0,1]} \big\|f_u^{(r)}(t_n+\tau\zeta)\big\|_{\mathcal{H}^*}, \tag{9.27}$$

$$\Big\|\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} R_{j+1}^{[r]}\Big\|_{\mathcal{V}} \lesssim \tau^r\,(\sqrt{t_{n+1}}+t_{n+1}) \operatorname*{ess\,sup}_{t\in[0,t_{n+1}]} \big\|f_u^{(r)}(t)\big\|_{\mathcal{H}^*}. \tag{9.28}$$

If $f_u^{(r)} \in L^\infty(0,T;\mathcal{V})$, then (9.27) improves to $\big\|\boldsymbol{R}_{n,i}^{[r]}\big\|_{\mathcal{V}} \lesssim \tau^{r+1}\operatorname{ess\,sup}_{\zeta\in[0,1]}\big\|f_u^{(r)}(t_n+\tau\zeta)\big\|_{\mathcal{V}}$.

Proof. For the estimates of $\boldsymbol{R}_{n,i}^{[r]}$ we note that the first term in (9.25) can be estimated with the same tricks as in the proof of Lemma 5.11. For this, one considers the underlying operator ODE with homogeneous initial value and right-hand side $\int_0^t \frac{(t-s)^{r-1}}{(r-1)!} f_u^{(r)}(t_n+s)\,\mathrm{d}s$. The bounds for the second term in the expression (9.25) of $\boldsymbol{R}_{n,i}^{[r]}$ is a consequence of Lemma 5.11.d)–f). Analogously one shows that the remainder $R_{n+1}^{[r]}$ satisfies upper bounds analogous to (9.26) and (9.27). With Lemma 5.11, we get

$$\begin{aligned}\Big\|\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} R_{j+1}^{[r]}\Big\|_{\mathcal{V}} &\leq \sum_{j=0}^{n-1} t_{n-j}^{-1/2}\big\|t_{n-j}^{1/2} e^{-\tau(n-j)\mathcal{A}}\big\|_{\mathscr{L}(\mathcal{H},\mathcal{V})}\big\|R_{j+1}^{[r]}\big\|_{\mathcal{H}} + \big\|R_{n+1}^{[r]}\big\|_{\mathcal{V}} \\ &\lesssim \Big(\tau\sum_{j=0}^{n-1} t_{n-j}^{-1/2} + \sqrt{\tau} + t_{n+1}\Big)\cdot\tau^r \operatorname*{ess\,sup}_{t\in[0,t_{n+1}]}\big\|f_u^{(r)}(t)\big\|_{\mathcal{H}^*}.\end{aligned}$$

For the term in the parentheses we note that $t_{j+1} \leq 2t_j$, $j = 1,\dots,n$. This and the monotonicity of $t \mapsto 1/\sqrt{t}$ lead to

$$\frac{1}{\sqrt{2}}\Big(\sqrt{\tau}+\tau\sum_{j=0}^{n-1} t_{n-j}^{-1/2}\Big) \leq \sqrt{\tau}+\tau\sum_{j=0}^{n-1} t_{n+1-j}^{-1/2} = \tau\sum_{j=1}^{n+1} t_j^{-1/2} \leq \int_0^{t_{n+1}} \frac{1}{\sqrt{s}}\,\mathrm{d}s = 2\sqrt{t_{n+1}}. \tag{9.29}$$

□

With the expression (9.18) of the local approximation $\widehat{u}_{n+1}$ and the expansion (9.21) of the solution u_{n+1} with defect R_{n+1} we get the expression for the local error

$$\widehat{e}_{n+1} = \widehat{u}_{n+1} - u(t_{n+1}) = \tau \sum_{i=1}^{s} b_i(-\tau\mathcal{A})(f(t_n + c_i\tau, \widehat{U}_{n,i}) - f_u(t_n + c_i\tau)) - R_{n+1}. \tag{9.30}$$

To obtain a formulation of the difference of the nonlinearity in (9.30) we can use the local error of the internal stages $\widehat{E}_{n,i}$ and the expansion

$$f(t_n + c_i\tau, \widehat{U}_{n,i}) - f_u(t_n + c_i\tau) = J_n\widehat{E}_{n,i} + \tau c_i K_n \widehat{E}_{n,i} + \widehat{Q}_{n,i}, \tag{9.31}$$

with the Fréchet derivatives

$$J_n = \frac{\partial}{\partial u} f(t_n, u(t_n)) \quad \text{and} \quad K_n = \frac{\partial^2}{\partial t \partial u} f(t_n, u(t_n)).$$

If $\widehat{E}_{n,i}$ is small enough, then the linear operators J_n and K_n map continuously from $\mathcal{V}$ to $\mathcal{H}^*$ and the defect $\widehat{Q}_{n,i} \in \mathcal{H}^*$ is bounded by $\|\widehat{Q}_{n,i}\|_{\mathcal{H}^*} \lesssim (c_i^2\tau^2 + \|\widehat{E}_{n,i}\|_{\mathcal{V}})\|\widehat{E}_{n,i}\|_{\mathcal{V}}$ by Assumption 9.9; cf. [HocO05a, Lem. 4.4]. As the last preparation step, we introduce some notation, such that we can describe the local error of the whole step and of the internal stages in a compact way. Therefore, we introduce the expressions

$$\widehat{\mathbf{E}}_n \coloneqq \begin{bmatrix} \widehat{E}_{n,1} \\ \vdots \\ \widehat{E}_{n,s} \end{bmatrix} \in \mathcal{V}_s, \qquad \boldsymbol{R}_n \coloneqq \begin{bmatrix} \boldsymbol{R}_{n,1} \\ \vdots \\ \boldsymbol{R}_{n,s} \end{bmatrix} \in \mathcal{V}_s, \qquad \widehat{\mathbf{Q}}_n \coloneqq \begin{bmatrix} \widehat{Q}_{n,1} \\ \vdots \\ \widehat{Q}_{n,s} \end{bmatrix} \in \mathcal{H}_s.$$

The notation $\mathbf{A}$ and $\mathbf{b}$ are short notations for $\mathbf{A}(-\tau\mathcal{A})$ and $\mathbf{b}(-\tau\mathcal{A})$, respectively, from the Butcher tableau (5.12) and $\mathbf{J}_n$ and $\mathbf{K}_n^{\mathbf{c}}$ are the block operators

$$\mathbf{J}_n \coloneqq \begin{bmatrix} J_n & & \\ & \ddots & \\ & & J_n \end{bmatrix} \in \mathscr{L}(\mathcal{V}_s, \mathcal{H}_s^*), \qquad \mathbf{K}_n^{\mathbf{c}} \coloneqq \begin{bmatrix} c_1 K_n & & \\ & \ddots & \\ & & c_s K_n \end{bmatrix} \in \mathscr{L}(\mathcal{V}_s, \mathcal{H}_s^*).$$

This allows us to express the local error $\widehat{e}_{n+1} = \widehat{u}_{n+1} - u(t_{n+1})$ in a compact manner.

Lemma 9.11. *Let Assumption 9.9 be satisfied. Suppose that $\mathcal{A} \in \mathscr{L}(\mathcal{V}, \mathcal{V}^*)$ is elliptic and satisfies Assumption 9.2.i). Let the step size τ be sufficiently small such that the local approximation $\widehat{u}_{n+1}$ and the local internal stages $\widehat{U}_{n,i}$ lie in a strip along u, where f is locally Lipschitz continuous with a uniform constant $L \geq 0$. Then the local error $\widehat{e}_{n+1}$ defined as in (9.19) can be expressed as*

$$\widehat{e}_{n+1} = \tau\widehat{S}_n - \tau\mathbf{b}^T \sum_{k=0}^{s-1} ((\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\tau\mathbf{A})^k (\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\boldsymbol{R}_n - R_{n+1},$$

where $\widehat{S}_n$ is an element of $\mathcal{V}$ with bounds

$$\sqrt{\tau}\|\widehat{S}_n\|_{\mathcal{V}} + \|\widehat{S}_n\|_{\mathcal{H}} \lesssim (\tau^2 + \|\boldsymbol{R}_n\|_{\mathcal{V}_s})\|\boldsymbol{R}_n\|_{\mathcal{V}_s} + h.o.t. \tag{9.32}$$

The higher order terms are all bounded, positive, and terms with higher powers of τ.

Proof. To show the assertion we make use of the representation (9.30) of the local error $\widehat{e}_{n+1}$. Note that $\widehat{e}_{n+1}$ is a function of the local error of the internal stages $\widehat{E}_{n,i}$, $i = 1, \dots, s$, by the presence of the nonlinearity. By the definition of $\widehat{U}_{n,i}$ and the expression (9.23) we have

$$\widehat{E}_{n,i} = \tau \sum_{j=2}^{i-1} a_{i,j}(-\tau\mathcal{A})(f(t_n + c_i\tau, \widehat{U}_{n,j}) - f_u(t_n + c_i\tau)) - \boldsymbol{R}_{n,i}.$$

Expanding the nonlinearity f by (9.31), the error $\widehat{E}_{n,i}$ is implicitly given by

$$\widehat{\mathbf{E}}_n = \tau\mathbf{A}(\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\widehat{\mathbf{E}}_n + \tau\mathbf{A}\mathbf{Q}_n - \boldsymbol{R}_n.$$

Since the exponential integrator scheme is explicit, the block operators $\sqrt{\tau}\mathbf{A}\mathbf{J}_n, \sqrt{\tau}\mathbf{A}\mathbf{K}_n^{\mathbf{c}} \in \mathscr{L}(\mathcal{V}_s)$ are strictly lower triangular. By a successive application of the last representation of $\widehat{\mathbf{E}}_n$ we get

$$\widehat{\mathbf{E}}_n = \sum_{k=0}^{s-1} (\tau\mathbf{A}(\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}}))^k (\tau\mathbf{A}\mathbf{Q}_n - \boldsymbol{R}_n). \tag{9.33}$$

On the other hand, we can use the local Lipschitzity of f, Lemma 5.11, and that the method is explicit, which implies $\widetilde{E}_{n,1} = \widehat{e}_n = 0$, such that

$$\|\widehat{E}_{n,i}\|_\mathcal{V} \le \sum_{j=2}^{i-1} \sqrt{\tau}L\|\sqrt{\tau}a_{i,j}(-\tau\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*,\mathcal{V})}\|\widehat{E}_{n,j}\|_\mathcal{V} + \|\boldsymbol{R}_{n,i}\|_\mathcal{V} \lesssim \sum_{j=2}^{i-1}(\sqrt{\tau}+\tau)L\|\widehat{E}_{n,j}\|_\mathcal{V} + \|\boldsymbol{R}_{n,i}\|_\mathcal{V}.$$

Here, we used that f has images in $\mathcal{H}^*$. With this bound and an induction argument over $i = 1, \dots, s$, one shows $\|\widehat{E}_{n,i}\|_\mathcal{V} \lesssim \|\boldsymbol{R}_{n,i}\|_\mathcal{V} + \text{h.o.t.}$, where the higher order terms are all positive with higher powers of τ. This then implies

$$\|\widehat{Q}_{n,i}\|_{\mathcal{H}^*} \lesssim (c_i^2\tau^2 + \|\boldsymbol{R}_{n,i}\|_\mathcal{V})\|\boldsymbol{R}_{n,i}\|_\mathcal{V} + \text{h.o.t.} \tag{9.34}$$

Let us come back to local error $\widehat{e}_{n+1}$. With the formulation (9.30) and with the reformulation of the local internal error (9.33) we have the identity

$$\begin{aligned}
\widehat{e}_{n+1} &\overset{(9.30)}{=} \tau\sum_{i=1}^{s} b_i(-\tau\mathcal{A})(f(t_n + c_i\tau, \widehat{U}_{n,i}) - f_u(t_n + c_i\tau)) - R_{n+1} \\
&\overset{(9.31)}{=} \tau\mathbf{b}^T((\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\widehat{\mathbf{E}}_n + \mathbf{Q}_n) - R_{n+1} \\
&\overset{(9.33)}{=} \tau\sum_{k=0}^{s-1}\mathbf{b}^T((\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\tau\mathbf{A})^k\mathbf{Q}_n - \tau\mathbf{b}^T\sum_{k=0}^{s-1}((\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\tau\mathbf{A})^k(\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\boldsymbol{R}_n - R_{n+1}.
\end{aligned}$$

We define $\tau\widehat{S}_n$ as the first sum of the right-hand side. Lemma 5.11 and the estimate (9.34) then yield the desired statement. □

For the difference $\widetilde{e}_{n+1} = u_{n+1} - \widehat{u}_{n+1}$ between the global approximation u_{n+1} and the local one $\widehat{u}_{n+1}$ we get a similar result.

Lemma 9.12. *Let the difference $\widetilde{e}_{n+1}$ be given as in (9.19) and e_n be the global error (9.17). Suppose that $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ is elliptic and satisfies Assumption 9.2.i). Further, let the step size τ be sufficiently small such that the derived local and global approximation and their associated internal stages lie within a strip along u in which f is locally Lipschitz continuous with a uniform constant $L > 0$. Then $\widetilde{e}_{n+1}$ satisfies*

$$\widetilde{e}_{n+1} = e^{-\tau\mathcal{A}}e_n + \tau\widetilde{S}_n$$

with an element $\widetilde{S}_n \in \mathcal{V}$*, which fulfills*

$$\|\widetilde{S}_n\|_{\mathcal{H}} \lesssim \|e_n\|_{\mathcal{V}} + h.o.t. \quad \text{and} \quad \sqrt{\tau}\|\widetilde{S}_n\|_{\mathcal{V}} \lesssim (1+\sqrt{\tau})(\|e_n\|_{\mathcal{V}} + h.o.t.). \tag{9.35}$$

Here, the higher order terms are all positive and of the form $\tau^{p/2}\,\|e_n\|_{\mathcal{V}}$*,* $p > 1$*.*

Proof. Following the steps of the proof of Lemma 9.11, we get for the difference of the internal stages $\widetilde{E}_{n,i} = U_{n,i} - \widehat{U}_{n,i}$, $i = 1,\dots,s$, the estimate

$$\|\widetilde{E}_{n,i}\|_{\mathcal{V}} \lesssim \|e_n\|_{\mathcal{V}} + L\sum_{j=1}^{i-1}\|\tau a_{i,j}(-\tau\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*,\mathcal{V})}\|\widetilde{E}_{n,j}\|_{\mathcal{V}} \lesssim \|e_n\|_{\mathcal{V}} + (\sqrt{\tau}+\tau)L\sum_{j=1}^{i-1}\|\widetilde{E}_{n,j}\|_{\mathcal{V}}.$$

Note that in contrast to the estimate of $\|\widehat{E}_{n,i}\|_{\mathcal{V}}$ the difference of the first terms in the definition of internal stages, i.e., $e^{-c_i\tau\mathcal{A}}(u_n - u(t_n))$, does not vanish. Anyway, the bound of $\|\widetilde{E}_{n,i}\|_{\mathcal{V}}$ and an induction argument over $i = 1,\dots,s$ proves $\|\widetilde{E}_{n,i}\|_{\mathcal{V}} \lesssim \|e_n\|_{\mathcal{V}} + \text{h.o.t.}$

Inserting the definition of global u_{n+1} and local approximations $\widehat{u}_{n+1}$, i.e., (9.16a) and (9.18), respectively, into $\widetilde{e}_{n+1} = u_{n+1} - \widehat{u}_{n+1}$ we get the identity

$$\widetilde{e}_{n+1} = e^{-\tau\mathcal{A}}(u_n - u(t_n)) + \tau\sum_{i=1}^{s} b_i(-\tau\mathcal{A})(f(t_n + c_i\tau, U_{n,i}) - f(t_n + c_i\tau, \widehat{U}_{n,i})) = e^{-\tau\mathcal{A}}e_n + \tau\widetilde{S}_n.$$

The local Lipschitzity of f and the estimate of $\|\widetilde{E}_{n,i}\|_{\mathcal{V}}$ imply

$$\|\widetilde{S}_n\|_{\mathcal{X}} \le L\sum_{i=1}^{s}\|b_i(-\tau\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*,\mathcal{X})}\|\widetilde{E}_{n,i}\|_{\mathcal{V}} \lesssim L\sum_{i=1}^{s}\|b_i(-\tau\mathcal{A})\|_{\mathscr{L}(\mathcal{H}^*,\mathcal{X})}(\|e_n\|_{\mathcal{V}} + \text{h.o.t.})$$

with $\mathcal{X} \in \{\mathcal{H},\mathcal{V}\}$. Then the assertion follows by Lemma 5.11. □

An important consequence of the previous lemmas 9.11 and 9.12 is the representation of the global error e_{n+1} as

$$\begin{aligned}
e_{n+1} =& \widetilde{e}_{n+1} + \widehat{e}_{n+1}\\
=& e^{-\tau\mathcal{A}}e_n + \tau\widetilde{S}_n + \tau\widehat{S}_n - \tau\mathbf{b}^T\sum_{k=0}^{s-1}((\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\tau\mathbf{A})^k(\mathbf{J}_n + \tau\mathbf{K}_n^{\mathbf{c}})\boldsymbol{R}_n - R_{n+1}\\
=& \underbrace{e^{-\tau(n+1)\mathcal{A}}e_0}_{(9.36a)} + \underbrace{\tau\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}}\widetilde{S}_j}_{(9.36b)} + \underbrace{\tau\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}}\widehat{S}_j}_{(9.36c)}\\
&\underbrace{-\tau\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}}\mathbf{b}^T\sum_{k=0}^{s-1}((\mathbf{J}_j + \tau\mathbf{K}_j^{\mathbf{c}})\tau\mathbf{A})^k(\mathbf{J}_j + \tau\mathbf{K}_j^{\mathbf{c}})\boldsymbol{R}_j}_{(9.36d)} - \underbrace{\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}}R_{j+1}}_{(9.36e)}
\end{aligned} \tag{9.36}$$

where we applied the second equality successive to get the third one. To get the order conditions for our exponential integrator we assume that the initial value is correct, i.e., $e_0 = 0$, and we expand the defects $\boldsymbol{R}_j$ and R_{j+1}. The resulting order conditions are summarized in Table 9.1. In the following associated theorem $\lceil p\rceil \in \mathbb{N}$ denotes for $p \in \mathbb{R}_{>0}$ the unique integer, which satisfies $p \le \lceil p\rceil < p+1$.

Theorem 9.13 (Order Conditions)**.** *Let Assumption 9.9 on the nonlinearity* f *be satisfied. Suppose that* $\mathcal{A} \in \mathscr{L}(\mathcal{V},\mathcal{V}^*)$ *is elliptic and satisfies Assumption 9.2.i). Assume that for* $1 \le \mathsf{p} \le 3$ *the order*

Table 9.1.: Order conditions for exponential integrators. The function ω_i and $\psi_{k,i}$ are defined at (9.22) and (9.24), respectively. Conditions with $\mathcal{J}$ or $\mathcal{K}$ should hold for all operators $\mathcal{J}, \mathcal{K} \in \mathscr{L}(\mathcal{V}, \mathcal{H}^*)$.

No.	Order	Order condition
1	1	$\omega_1(-\tau\mathcal{A}) = 0$
2	1	$\psi_{1,i}(-\tau\mathcal{A}) = 0, \quad i = 1, \dots, s$
3	$3/2$	$\omega_2(-\tau\mathcal{A}) = 0$
4	2	$\sum_{i=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}\psi_{2,i}(-\tau\mathcal{A}) = 0$
5	$5/2$	$\omega_3(-\tau\mathcal{A}) = 0$
6	$5/2$	$\sum_{i,k=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}a_{i,k}(-\tau\mathcal{A})\mathcal{J}\psi_{2,k}(-\tau\mathcal{A}) = 0$
7	3	$\sum_{i=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}\psi_{3,i}(-\tau\mathcal{A}) = 0$
8	3	$\sum_{i,k,\ell=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}a_{i,k}(-\tau\mathcal{A})\mathcal{J}a_{k,\ell}(-\tau\mathcal{A})\mathcal{J}\psi_{2,\ell}(-\tau\mathcal{A}) = 0$
9	3	$\sum_{i=1}^{s} b_i(-\tau\mathcal{A})c_i\mathcal{K}\psi_{2,i}(-\tau\mathcal{A}) = 0$

conditions in Table 9.1 hold up to order p. *Let* τ *be fixed and sufficiently small such that the discrete solution* u_n *and its internal stages* $U_{n,i}$, $i = 1, \dots, s$, *lie in a strip along* u, *where* f *is locally Lipschitz continuous with a uniform constant* $L > 0$. *Then we have*

$$\|u_n - u(t_n)\|_{\mathcal{V}} \lesssim \tau^{\mathsf{p}} + h.o.t.,$$

$n = 1, \dots, N$, *where the constant depends on* t_n, L, $\mathcal{A}$, *and* $\|f(\cdot, u(\cdot))\|_{W^{q,\infty}(0,t_n,\mathcal{H})}$ *with* $q = \lceil \mathsf{p} \rceil$.

Proof. Theorem 4.3 in [HocO05a] shows that the conditions 1, 2, and 3 in Table 9.1 are sufficient for an exponential integrator to be of order one and a half. With the same steps of the proof one shows that the order condition No. 1 and 2 imply at least first order of the method.

Therefore, we may assume that the order conditions No. 1, 2, and 3 are fulfilled. In particular, a consequence of condition 2, i.e., $\psi_{1,i}(-\tau\mathcal{A}) = 0$, $i = 1, \dots, s$, is by Lemma 9.10 that

$$\|\boldsymbol{R}_{n,i}\|_{\mathcal{V}} = \|\boldsymbol{R}_{n,i}^{[1]}\|_{\mathcal{V}} \lesssim \tau^{3/2} \|\tfrac{\mathrm{d}}{\mathrm{d}t} f_u\|_{L^\infty(t_n,t_{n+1};\mathcal{H}^*)}. \tag{9.37}$$

We now consider the representation (9.36) of the global error and estimate the summands (9.36a) to (9.36e). Note that the error (9.36) is for the $(n+1)$st step. Anyway, (9.36a) vanishes since we use the exact initial value, i.e., $e_0 = 0$. By Lemma 5.11 and 9.12 we can bound (9.36b) by

$$\begin{aligned}
\Big\| \tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \widetilde{S}_j \Big\|_{\mathcal{V}} &\lesssim \tau \sum_{j=1}^{n-1} (t_{n-j}^{-1/2} + 1)\|\widetilde{S}_j\|_{\mathcal{H}} + \tau \|\widetilde{S}_n\|_{\mathcal{V}} \\
&\overset{(9.35)}{\lesssim} \tau \sum_{j=1}^{n-1} (t_{n-j}^{-1/2} + 1)(\|e_j\|_{\mathcal{V}} + \text{h.o.t.}) + (\sqrt{\tau} + \tau)(\|e_n\|_{\mathcal{V}} + \text{h.o.t.}) \\
&\leq \sqrt{2}(1 + \sqrt{t_{n+1}})\, \tau \sum_{j=1}^{n} t_{n+1-j}^{-1/2} (\|e_j\|_{\mathcal{V}} + \text{h.o.t.}),
\end{aligned}$$

where we used $\widetilde{S}_0 = 0$ and $t_{j+1} \leq 2t_j$, $j = 1, \dots, n$. With the same argument for (9.36c) we get

$$\Big\| \tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \widehat{S}_j \Big\|_{\mathcal{V}} \overset{(9.32)}{\lesssim} \sqrt{2}\tau \sum_{j=0}^{n} (t_{n+1-j}^{-1/2} + 1)(\tau^2 + \|\boldsymbol{R}_j\|_{\mathcal{V}_s})\|\boldsymbol{R}_j\|_{\mathcal{V}_s} + \text{h.o.t.}$$

$$\overset{(9.29)}{\leq} \sqrt{2}(2\sqrt{t_{n+1}} + t_{n+1}) \max_{j=0,\dots,n} (\tau^2 + \|\boldsymbol{R}_j\|_{\mathcal{V}_s})\|\boldsymbol{R}_j\|_{\mathcal{V}_s} + \text{h.o.t.}$$

$$\overset{(9.37)}{\leq} \tau^3\sqrt{2}(2\sqrt{t_{n+1}} + t_{n+1})\|\tfrac{\mathrm{d}}{\mathrm{d}t} f_u\|^2_{L^\infty(0,t_{n+1};\mathcal{H}^*)} + \text{h.o.t.}$$

For the term (9.36d) we introduce the vectors $\boldsymbol{\psi}_k = [\psi_{k,1} \dots \psi_{k,s}]^T \in \mathcal{V}_s$ with $\psi_{k,i}$ from (9.24), $k = 2, 3$. We then bound (9.36d) via

$$\begin{aligned}
&\Big\|\tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \mathbf{b}^T \sum_{k=0}^{s-1} ((\mathbf{J}_j + \tau \mathbf{K}_j^{\mathbf{c}})\tau \mathbf{A})^k (\mathbf{J}_j + \tau \mathbf{K}_j^{\mathbf{c}})\boldsymbol{R}_j\Big\|_{\mathcal{V}} \\
&\leq \tau^{3/2}\Big\|\tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \mathbf{b}^T \mathbf{J}_j \sqrt{\tau}\boldsymbol{\psi}_2 \tfrac{\mathrm{d}}{\mathrm{d}t} f_u(t_j)\Big\|_{\mathcal{V}} + \tau^2\Big\|\tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \mathbf{b}^T \mathbf{J}_j \sqrt{\tau}\mathbf{A}\mathbf{J}_j\sqrt{\tau}\boldsymbol{\psi}_2 \tfrac{\mathrm{d}}{\mathrm{d}t} f_u(t_j)\Big\|_{\mathcal{V}} \\
&\quad + \tau^{5/2}\Big\|\tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \mathbf{b}^T \big[\mathbf{J}_j\sqrt{\tau}\boldsymbol{\psi}_3 \tfrac{\mathrm{d}^2}{\mathrm{d}t^2} f_u(t_j) + \mathbf{K}_j^{\mathbf{c}}\sqrt{\tau}\boldsymbol{\psi}_2 \tfrac{\mathrm{d}}{\mathrm{d}t} f_u(t_j) + \mathbf{J}_j\sqrt{\tau}(\mathbf{A}\mathbf{J}_j\sqrt{\tau})^2\boldsymbol{\psi}_2 \tfrac{\mathrm{d}}{\mathrm{d}t} f_u(t_j)\big]\Big\|_{\mathcal{V}} \\
&\quad + \tau^3 \widehat{R}_n.
\end{aligned} \tag{9.38}$$

By the same steps as in the proof of Lemma 9.10 one shows

$$\Big\|\tau \sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} \mathbf{b}^T \mathbf{J}_j \sqrt{\tau}\boldsymbol{\psi}_2 \tfrac{\mathrm{d}}{\mathrm{d}t} f_u(t_j)\Big\|_{\mathcal{V}} \lesssim (\sqrt{t_{n+1}} + t_{n+1})\|f_u\|_{W^{q,\infty}(0,t_{n+1};\mathcal{H}^*)} + \text{h.o.t.}$$

with $q = 2$. The higher order terms are terms with positive integer powers of $\sqrt{\tau}$. The next two summands of the right-hand side of (9.38) can be bounded analogously. The needed regularity q then is one order higher as in the summands itself, i.e., for the second summand $q = 2$ and for the third $q = 3$. The remainder $\widehat{R}_n$ can be estimated by $\widehat{R}_n \lesssim (\sqrt{t_{n+1}} + t_{n+1})\|f_u\|_{W^{3,\infty}(0,t_{n+1};\mathcal{H}^*)} + \text{h.o.t.}$ For (9.36e) we have by Lemma 9.10 the estimate

$$\Big\|\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} R_{j+1}\Big\|_{\mathcal{V}} = \Big\|\sum_{j=0}^{n} e^{-\tau(n-j)\mathcal{A}} R_{j+1}^{[r]}\Big\|_{\mathcal{V}} \overset{(9.28)}{\lesssim} \tau^r\,(\sqrt{t_{n+1}} + t_{n+1})\|f_u^{(r)}\|_{L^\infty(0,t_{n+1};\mathcal{H}^*)}$$

with $r = 3$ if $\omega_3 = 0$ and $r = 2$, otherwise. Finally, the assertion follows directly by the discrete version of Gronwall's lemma from [HocO05b, Ch. 4, Lem 4.]. ☐

Remark 9.14. The application of the discrete version of Gronwall's lemma in the last step of the proof of Theorem 9.13 introduces a constant depending on e^{Lt_n}. This constant is hidden in the prefactor in the error estimate. For its determination one possibly could use the exponential decay of the semigroup $e^{-t\mathcal{A}}$. For the simplest case of an operator DAE (9.1) with a self-adjoint, elliptic $\mathcal{A}$, i.e., $\mathcal{A} = \mathcal{A}_1$, discretized with the exponential Euler the error becomes

$$\|u_n - u(t_n)\|^2_{\mathcal{V}} \leq \tau^2 \frac{2}{LC_{\mathcal{V}\hookrightarrow\mathcal{H}}} \frac{e^{\beta t_n} - 1}{\beta} \frac{\sqrt{1+\varepsilon}}{\varepsilon} \|\tfrac{\mathrm{d}}{\mathrm{d}t} f_u\|^2_{L^\infty(0,t_n;\mathcal{H}^*)}, \tag{9.39}$$

if τ is sufficiently small. The sufficiently small upper bound for τ depends on $\varepsilon > 0$ and $\beta = 2\sqrt{1+\varepsilon}LC^{-1}_{\mathcal{V}\hookrightarrow\mathcal{H}} - 2\mu_{\mathcal{A}}C^{-2}_{\mathcal{V}\hookrightarrow\mathcal{H}}$. The associated proof is similar to the one of Theorem 9.5, where we consider the operator $\mathcal{A} - \kappa\,\mathrm{id}$, $\kappa \leq \mu_{\mathcal{A}}C^{-2}_{\mathcal{V}\hookrightarrow\mathcal{H}}$, and optimize over κ afterwards.

Table 9.1 summarizes the order conditions for the case that $f(t, u(t)) \in \mathcal{H}^*$. If the solution u behaves well such that $f(t, u(t))$ and its time derivatives are elements of $\mathcal{V}$ as well as that the Fréchet derivatives $J_n, K_n \in \mathscr{L}(\mathcal{V})$, then the convergence order can increase. For example, condition 1 in

Table 9.2.: Order conditions for exponential integrators if $f(t, u(t))$ and their derivatives have images in $\mathcal{V}$. The function ω_i and $\psi_{k,i}$ are defined at (9.22) and (9.24), respectively. Conditions with $\mathcal{J}$ or $\mathcal{K}$ should hold for all operators $\mathcal{J}, \mathcal{K} \in \mathscr{L}(\mathcal{V})$.

No.	Order	Order condition
1	1	$\omega_1(-\tau\mathcal{A}) = 0$
2	2	$\psi_{1,i}(-\tau\mathcal{A}) = 0, \quad i = 1, \ldots, s$
3	2	$\omega_2(-\tau\mathcal{A}) = 0$
4	3	$\sum_{i=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}\psi_{2,i}(-\tau\mathcal{A}) = 0$
5	3	$\omega_3(-\tau\mathcal{A}) = 0$
6	4	$\sum_{i,k=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}a_{i,k}(-\tau\mathcal{A})\mathcal{J}\psi_{2,k}(-\tau\mathcal{A}) = 0$
7	4	$\sum_{i=1}^{s} b_i(-\tau\mathcal{A})\mathcal{J}\psi_{3,i}(-\tau\mathcal{A}) = 0$
8	4	$\sum_{i=1}^{s} b_i(-\tau\mathcal{A})c_i\mathcal{K}\psi_{2,i}(-\tau\mathcal{A}) = 0$
9	4	$\omega_4(-\tau\mathcal{A}) = 0$

Table 9.1 then is already sufficient for a first order method and the conditions 1 to 5 in Table 9.1 would lead to third order. The order conditions for methods up to fourth order then are the same as in [HocO05a, Tab. 4.1]. They are summarized in Table 9.2. The associated proof is equivalent to the one of Theorem 9.13, where one uses Lemma 9.10 to improve the estimate of the remainder $\boldsymbol{R}_{n,i}^{[1]}$.

The used techniques in this section can only show the conditions for methods of order up to three and four for $f(\cdot, u(\cdot))$ having codomain $\mathcal{H}^*$ and $\mathcal{V}$, respectively. By also considering the bilinear form $\frac{\partial^2 f}{\partial u^2}$, the authors of [LuaO14a] derived the conditions of methods up to fifth order for an error in the $\mathcal{H}$-norm if the nonlinearity has domain $\mathcal{H}$. It is possible to adapt this approach to get higher order methods for our purpose. This, however, is not investigated in this thesis, since the construction would rely not only on the smoothness of nonlinearity $f(t, u(t))$ like before but also of the initial value u_0 and the solution u itself; cf. [LuaO14b]. Even in the linear case of operator DAEs theses smoothness conditions can be very restrictive and hardly practical; cf. [Tem82].

The first order of the one stage exponential Euler of Section 9.1 and order $3/2$ of exponential Runge methods from (9.14) are presented in Table 9.1. Note that there does not exist an exponential integrator of second order with two stages since in this case the second and fourth condition in Table 9.1 contradict each other. In [HocO05a, Sec. 5.2] the so-called *exponential Heun methods* are introduced. This three-parameter family of three-stage integrators is given by

$$\begin{array}{c|ccc} 0 & & & \\ c_2 & c_2\phi_{1,2} & & \\ c_3 & c_3\phi_{1,3} - \gamma c_2\phi_{2,2} - \frac{c_3^2}{c_2}\phi_{2,3} & \gamma c_2\phi_{2,2} + \frac{c_3^2}{c_2}\phi_{2,3} & \\ \hline & \phi_1 - \frac{1+\gamma}{\gamma c_2 + c_3}\phi_2 & \frac{\gamma}{\gamma c_2 + c_3}\phi_2 & \frac{1}{\gamma c_2 + c_3}\phi_2 \end{array} \tag{9.40}$$

with $c_2, c_3, \gamma c_2 + c_3 \neq 0$ (and $c_2, c_3 \leq 1$). Every scheme of this family satisfies the first four conditions in Table 9.1 and has therefore convergence order two. The specific choice $c_2 = c_3 = 1$ leads to a minimal number of evaluations of the right-hand side. The authors of [HocO05a] were actually interested in a three-stage method, which fulfills the first five conditions in Table 9.1 and showed that such a method does not exist. With a long calculation one can also prove that there is even no four-stage method that satisfies the first six conditions in Table 9.1. Therefore, a four-stage method of order $5/2$ is not possible. A five-stage method is presented in [HocO05a, Eq. (5.19)].

This subsection is concluded with an investigation of the exponential integrators under perturbations of the right-hand side and the initial value. Recall, that this may then be interpreted as the error of a spatial discretization; cf. p. 118. For the spatially discretized system the nonlinearity $f(t,u(t))$ might be further approximated by model reduction techniques like the discrete empirical interpolation method [ChaS10]. This introduces additional perturbations of the original right-hand side f.

Lemma 9.15 (Error under Perturbations)**.** *Suppose the assumption of Theorem 9.13 are satisfied. Assume that there is an error $e_0 \in \mathcal{V}$ in the initial value and the evaluation of the nonlinearity $f(t_n + c_i\tau, U_{n,i})$ is perturbed by $\delta_{n,i} \in \mathcal{H}^*$, $n = 0, \dots, N-1$, $i = 1, \dots, s$. Let all perturbations be small enough such that the discrete solution and its internal stages lie in a strip along u, where f is locally Lipschitz continuous with a uniform constant. Then the error between the solution and its numerical approximation is bounded by*

$$\|u_n - u(t_n)\|_{\mathcal{V}} \lesssim \tau^{\mathsf{p}} + \Big(1 + \sqrt{\tfrac{\tau}{n}}\Big)\|e_0\|_{\mathcal{V}} + \tau \sum_{j=0}^{n-1} (t_{n-j}^{-1/2} + 1) \sum_{i=1}^{s} \|\delta_{j,i}\|_{\mathcal{H}^*} + h.o.t.,$$

$n = 0, \dots, N$, with positive higher order terms.

Proof. We review the proof of Theorem 9.13. At first, we note that the statement of Lemma 9.11 does not change under perturbations. For Lemma 9.12 we adapt the estimate for the internal stages under sufficiently small perturbations, which results in $\|\widetilde{E}_{n,i}\|_{\mathcal{V}} \lesssim \|e_n\|_{\mathcal{V}} + (\sqrt{\tau} + \tau) \sum_{k=1}^{i-1} \big(L\|\widetilde{E}_{n,k}\|_{\mathcal{V}} + \|\delta_{n,k}\|_{\mathcal{H}^*}\big)$. By an induction argument we get $\|\widetilde{E}_{n,i}\|_{\mathcal{V}} \lesssim \|e_n\|_{\mathcal{V}} + (\sqrt{\tau} + \tau) \sum_{k=1}^{i-1} \|\delta_{n,k}\|_{\mathcal{H}^*} + \text{h.o.t.}$ With the steps of the proof of Lemma 9.12 one shows that $\widetilde{S}_n \in \mathcal{V}$ under small perturbations satisfies the bounds (9.35) with $\|e_n\|_{\mathcal{V}} + \sum_{i=1}^{s} \|\delta_{n,i}\|_{\mathcal{H}^*}$ instead of $\|e_n\|_{\mathcal{V}}$. The assertion then follows by an adaptation of the proof of Theorem 9.13, where one considers $e^{-(n+1)\tau\mathcal{A}}e_0$ and $\widetilde{S}_0$ in (9.36) and uses $\tau t_n^{-1/2} = \sqrt{\tau/n}$. □

Remark 9.16. As in Remark 9.14 the estimate under perturbations can be improved, if one takes the exponential decay of the semigroup $e^{-t\mathcal{A}}$ into account. For the exponential Euler with $\mathcal{A}_2 = 0$ the error of a perturbed solver can be bounded for small τ and small enough perturbations by

$$\begin{aligned}\frac{\sqrt{1+\varepsilon}LC_{\mathcal{V}\hookrightarrow\mathcal{H}}}{2}\|u_n - u(t_n)\|_{\mathcal{V}}^2 &\le \tau^2\, \frac{e^{\beta t_n} - 1}{\beta}\Big(2 + \frac{2}{\varepsilon}\Big)\, \|\tfrac{\mathrm{d}}{\mathrm{d}t} f_u\|_{L^\infty(0,t_n;\mathcal{H}^*)}^2 \\ &\quad + e^{\beta t_n}\Big(C_{\mathcal{A}} + \frac{\sqrt{1+\varepsilon}LC_{\mathcal{V}\hookrightarrow\mathcal{H}}}{2}\Big)\|e_0\|_{\mathcal{V}}^2 + \tau\Big(1 + \frac{2}{\varepsilon}\Big)\sum_{j=0}^{n-1} e^{\beta t_{n-j-1}}\|\delta_{j,1}\|_{\mathcal{H}^*}^2,\end{aligned}$$

with β as in (9.39). In particular, β is negative if $\mu_{\mathcal{A}} > LC_{\mathcal{V}\hookrightarrow\mathcal{H}}$ and $\varepsilon > 0$ is small enough.

9.3.2. Systems with Linear Constrains

We now return to the analysis of the semi-linear operator DAE (9.1). We explained the steps to get an approximation of the solution u with exponential integrators in the beginning of this chapter 9. The main idea is to calculate the part of the solution in $\{v \in \mathcal{V} \,|\, \mathcal{A}v \in \mathcal{V}_{\text{ker}}^0\}$ by the stationary saddle point problem (9.5) and apply the exponential integrator to the operator ODE (9.2) to approximate the part in $\mathcal{V}_{\text{ker}}$. If the explicit exponential integrator is given by the Butcher tableau (5.12), then the approximation reads

$$\begin{aligned}u_{n+1} = \mathcal{B}_{\mathcal{A}}^- g(t_{n+1}) + e^{-\tau\mathcal{A}_{\text{ker}}}(u_n - \mathcal{B}_{\mathcal{A}}^- g(t_n)) \\ + \tau \sum_{i=1}^{s} b_i(-\tau\mathcal{A}_{\text{ker}})\iota_{\text{ker}}\big[f(t_n + c_i\tau, U_{n,i}) - \mathcal{B}_{\mathcal{A}}^- \dot{g}(t_n + c_i\tau)\big]\end{aligned} \qquad (9.41a)$$

with $n = 0, \dots, N-1$. The internal stages $U_{n,i}$, $i = 1, \dots, s$ are given by

$$U_{n,i} = \mathcal{B}_{\mathcal{A}}^{-} g(t_n + \tau c_i) + e^{-\tau c_i \mathcal{A}_{\text{ker}}}(u_n - \mathcal{B}_{\mathcal{A}}^{-} g(t_n)) \\ + \tau \sum_{j=1}^{i-1} a_{i,j}(-\tau \mathcal{A}_{\text{ker}}) \iota_{\text{ker}} \big[f(t_n + c_i \tau, U_{n,i}) - \mathcal{B}_{\mathcal{A}}^{-} \dot{g}(t_n + c_i \tau) \big]. \tag{9.41b}$$

In Subsection 9.3.2.1 we investigate the convergence order of this time-stepping method. Note that, scheme (9.41) only approximates the state u. An approximation of the Lagrange multiplier λ is constructed and analyzed in Subsection 9.3.2.2.

9.3.2.1. Error Analysis

The practical calculation of the internal states (9.41b) and the whole step (9.41a) can be done by solving some stationary saddle point problems as (9.5) or (9.6) and transient ones with homogenous right-hand side; cf. Subsection 9.1.1 and 9.2.1. Alternatively, one also can determine the approximation by transient saddle point problems with a polynomial right-hand side, which reduces the number of stationary saddle point problems to solve; see Subsection 9.4.1. However, the practical computation does not change the convergence order.

Theorem 9.17 (Error Estimate for Operator DAEs)**.** *Let Assumptions 9.1 and 9.2 on $\mathcal{B}$ and $\mathcal{A}$ as well as Assumptions 9.3 and 9.9 on the nonlinearity f and right-hand side g be fulfilled. The initial value $u_0 \in \mathcal{V}$ be consistent, i.e., $\mathcal{B}u_0 = g(0)$. Assume that the order conditions up to order p, $1 \le \mathsf{p} \le 3$, of Table 9.1 are satisfied. Let $f(\cdot, u(\cdot)) \in W^{q,\infty}(0,T;\mathcal{H}^*)$ and $g \in W^{q+1,\infty}(0,T;\mathcal{Q}^*)$ with u be the solution of (9.1) and $q = \lceil \mathsf{p} \rceil$. Suppose that the step size τ is fixed and small enough such that the numerical solution and the associated internal stages lie in a strip along u, where f is locally Lipschitz continuous with a uniform constant $L > 0$. Then the numerical error is bounded by*

$$\|u_n - u(t_n)\|_{\mathcal{V}} \lesssim \tau^{\mathsf{p}} + h.o.t.,$$

$n = 1, \dots, N$, where the constant depends on t_n, the Lipschitz constant L, the operators $\mathcal{A}$ and $\mathcal{B}$, as well as the norms of right-hand sides $\|f(\cdot, u(\cdot))\|_{W^{q,\infty}(0,t_n,\mathcal{H}^*)}$ and $\|g\|_{W^{q+1,\infty}(0,t_n,\mathcal{Q}^*)}$.

Proof. Since $u_{c,n} = \mathcal{B}_{\mathcal{A}}^{-} g(t_n) = u_c(t_n)$ and analogously $U_{c,n,i} = u_c(t_n + \tau c_i)$, the assertion follows by Theorem 9.13 and the approximation of u_{ker} by the operator ODE (9.2). □

In Subsection 9.3.1 we mentioned that the convergence order for PDEs can increase if $f(t, u(t)) \in \mathcal{V}$ and $J_n, K_n \in \mathscr{L}(\mathcal{V})$. This still holds for the operator DAE case. The arguments are the same where we additionally use Remark 5.12. For completeness, we summarize this result in the following lemma.

Lemma 9.18. *In addition to the assumptions of Theorem 9.17 suppose that the function $t \mapsto f(t, u(t))$ and its time derivatives have images in $\mathcal{V}$ as well as the Fréchet derivatives J_n, K_n are element of $\mathscr{L}(\mathcal{V})$, $n = 0, \dots, N-1$. Then the order conditions for methods up to fourth order are given in Table 9.2.*

In comparison to the unconstrained case, the error analysis under perturbations becomes more delicate for operator DAEs. The source of possible perturbations doubles, since not only the initial value, and the right-hand side f could be perturbed, but also the right-hand side g of the constraint (9.1b) and its derivative $\dot{g}$.

Lemma 9.19 (Error Estimate for Operator DAEs under Perturbations)**.** *Let the assumptions of Theorem 9.17 be satisfied. In addition, let the initial value u_0 be perturbed by $e_0 \in \mathcal{V}$, $g(t_n + c_i \tau)$ by $\theta_{n,i} \in \mathcal{Q}^*$, $\dot{g}(t_n + c_i \tau)$ by $\xi_{n,i} \in \mathcal{Q}^*$, and the evaluation of the nonlinearity $f(t_n + c_i \tau, U_{n,i})$ by $\delta_{n,i} \in \mathcal{H}^*$. Suppose that the perturbations are consistent, i.e., $\mathcal{B}e_0 = \theta_{0,1}$, and sufficiently small such* *that u_n and $U_{n,i}$, $i = 1, \dots, s$, lie in a strip along u, where f is Lipschitz continuous with a uniform Lipschitz constant. Then the error can be estimated by*

$$\|u_n - u(t_n)\|_{\mathcal{V}} \lesssim \tau^{\mathsf{p}} + \Big(1 + \sqrt{\tfrac{\tau}{n}}\Big) \|e_0 - \mathcal{B}_{\mathcal{A}}^- \theta_{0,1}\|_{\mathcal{V}} + \|\mathcal{B}_{\mathcal{A}}^- \theta_{n,1}\|_{\mathcal{V}} \\ + \tau \sum_{j=1}^{n-1} (t_{n-j}^{-1/2} + 1) \sum_{i=1}^{s} \big(\|\delta_{j,i} - \mathcal{B}_{\mathcal{A}}^- \xi_{j,i}\|_{\mathcal{H}^*} + \|\mathcal{B}_{\mathcal{A}}^- \theta_{j,i}\|_{\mathcal{V}} \big) + h.o.t.,$$

$n = 0, \dots, N$, with positive higher order terms.

Proof. We split the solution in its parts in $\mathcal{V}_{\mathrm{c}}$ and $\mathcal{V}_{\mathrm{ker}}$. Then the error of the approximation in $\mathcal{V}_{\mathrm{c}}$ is given by $u_{\mathrm{c},n} - u_{\mathrm{c}}(t_n) = \mathcal{B}_{\mathcal{A}}^-(g_{t_n} + \theta_{n,1} - g_{t_n}) = \mathcal{B}_{\mathcal{A}}^- \theta_{n,1}$, where we used $c_1 = 0$. Analogously, one determines the error of $U_{\mathrm{c},n,i}$ as $\mathcal{B}_{\mathcal{A}}^- \theta_{n,i}$. For the part in $\mathcal{V}_{\mathrm{ker}}$ we consider the approximation of the solution of the operator DAE (9.2) under perturbations. Note that the initial value is perturbed by $u_{\mathrm{ker},0} - u_{\mathrm{ker}}(0) = u_0 - u_{\mathrm{c},0} - u(0) + u_{\mathrm{c}}(0) = e_0 - \mathcal{B}_{\mathcal{A}}^- \theta_{0,1}$ and the right-hand side associated with $U_{\mathrm{ker},n,i}$ by

$$\eta_{n,i} = f(t_n + c_i\tau, U_{\mathrm{ker},n,i} + U_{\mathrm{c},n,i}) + \delta_{n,i} - \mathcal{B}_{\mathcal{A}}^- \xi_{n,i} - f(t_n + c_i\tau, U_{\mathrm{ker},n,i} + u_{\mathrm{c}}(t_n + c_i\tau)) \in \mathcal{H}^*$$

with $\|\eta_{n,i}\|_{\mathcal{H}^*} \le L\|\mathcal{B}_{\mathcal{A}}^- \theta_{n,i}\|_{\mathcal{V}} + \|\delta_{n,i} - \mathcal{B}_{\mathcal{A}}^- \xi_{n,i}\|_{\mathcal{H}^*}$. The assertion then follows by Lemma 9.15. □

Remark 9.20. Under the conditions of Remarks 9.14 and 9.16 the error estimate for the exponential Euler method under perturbations can be improved for sufficiently small τ to

$$\|u_n - u(t_n)\|_{\mathcal{V}}^2 \\ \le \tau^2 \frac{e^{\beta t_n} - 1}{\beta} \frac{12\sqrt{1+\varepsilon}}{\varepsilon L C_{\mathcal{V}\hookrightarrow\mathcal{H}}} \|\tfrac{\mathrm{d}}{\mathrm{d}t} f - \mathcal{B}_{\mathcal{A}}^- \ddot{g}\|_{L^\infty(0,t_n;\mathcal{H}^*)}^2 + e^{\beta t_n} \Big(\frac{4C_{\mathcal{A}}}{\sqrt{1+\varepsilon L C_{\mathcal{V}\hookrightarrow\mathcal{H}}}} + 2 \Big) \|e_0 - \mathcal{B}_{\mathcal{A}}^- \theta_{0,1}\|_{\mathcal{V}}^2 \\ + 2\|\mathcal{B}_{\mathcal{A}}^- \theta_{n,1}\|_{\mathcal{V}}^2 + \tau \frac{12\sqrt{1+\varepsilon}}{\varepsilon L C_{\mathcal{V}\hookrightarrow\mathcal{H}}} \sum_{j=0}^{n-1} e^{\beta t_{n-j-1}} \big[\|\delta_{j,1} - \mathcal{B}_{\mathcal{A}}^- \xi_{j,1}\|_{\mathcal{H}^*}^2 + L^2 \|\mathcal{B}_{\mathcal{A}}^- \theta_{j,1}\|_{\mathcal{V}}^2 \big].$$

9.3.2.2. Approximation of the Lagrange Multiplier

To this point, we have only investigated the difference between the solution $u(t_n)$ and its approximation u_n. In contrast to the Runge-Kutta methods in Chapter 8, the exponential integrators do not calculate an approximation of the Lagrange multiplier λ, only of u. In this subsection we show that λ can be approximate at t_n by only using the already calculated u_n and the saddle point structure. For this task we introduce the norm

$$\|v\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}^2 \coloneqq \|v\|_{\mathcal{H}}^2 + \|\mathcal{B}v\|_{\mathcal{Q}^*}^2$$

on $\mathcal{V}$. Since $\mathcal{H}$ and $\mathcal{Q}^*$ are Hilbert spaces, the norm $\|\cdot\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}$ is induced by an inner product. The space $\mathcal{V}$ with this norm is in general only a pre-Hilbert space; see Example 9.21. The closure of $\mathcal{V}$ with respect to $\|\cdot\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}$ is denoted by $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$, i.e.,

$$[\mathcal{V},\mathcal{H}]_{\mathcal{B}} \coloneqq \operatorname{clos}_{\|\cdot\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}} \mathcal{V}. \tag{9.42}$$

Example 9.21 (The Space $H^1(\mathrm{div};\Omega)$)**.** We consider the spaces $\mathcal{V} \coloneqq [H^1(\Omega)]^d$ and $\mathcal{H} \coloneqq [L^2(\Omega)]^d$ with a Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$. The operator $\mathcal{B}$ denoted the weak divergence. Then $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ matches the space of all $[L^2(\Omega)]^d$-functions with a distributional divergence in $L^2(\Omega)$, i.e., $H^1(\mathrm{div};\Omega)$; see [DauL90, p. 203 ff.].

The space $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ can be interpreted as the largest space, which is (densely) embedded in $\mathcal{H}$ and where the operator $\mathcal{B}$ is well-defined. In the following lemma we summarize properties of $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$.

Lemma 9.22 (Properties of $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$)**.** *Let $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ be defined as in (9.42). Suppose that $v_n \to 0$ in $\mathcal{H}$ and $\mathcal{B}v_n \to g$ in $\mathcal{Q}^*$ for a sequence $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{V}$ imply $g = 0$. Then the following holds.*

i) The space $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ with the norm $\|\cdot\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}$ is a Hilbert space and $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}, \mathcal{H}, [\mathcal{V},\mathcal{H}]_{\mathcal{B}}^*$ forms a Gelfand triple. The space $\mathcal{V}$ is densely embedded into $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ and the operator $\mathcal{B}$ has a unique extension $\overline{\mathcal{B}} \in \mathscr{L}([\mathcal{V},\mathcal{H}]_{\mathcal{B}}, \mathcal{Q}^*)$.

ii) The space $\ker\overline{\mathcal{B}}$ is isometric isomorphic to $\mathcal{H}_{\ker}$. For every closed complement $\mathcal{V}_{\mathrm{c}} \subset \mathcal{V}$ with $\mathcal{V} = \mathcal{V}_{\ker} \oplus \mathcal{V}_{\mathrm{c}}$ its embedding in $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ is also close in $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ and

$$[\mathcal{V},\mathcal{H}]_{\mathcal{B}} \cong \mathcal{H}_{\ker} \oplus \mathcal{V}_{\mathrm{c}}.$$

iii) For every $f \in [\mathcal{V},\mathcal{H}]_{\mathcal{B}}^*$ and $g \in \mathcal{Q}^*$ the saddle point problem

$$\begin{aligned} \mathcal{M}u - \overline{\mathcal{B}}^*\lambda &= f \qquad \text{in } [\mathcal{V},\mathcal{H}]_{\mathcal{B}}^*, \\ \overline{\mathcal{B}}u \qquad &= g \qquad \text{in } \mathcal{Q}^* \end{aligned}$$

with $\mathcal{M} = \mathcal{R}_{\mathcal{H}}$, has a unique solution $(u_{\mathcal{B}},\lambda) \in [\mathcal{V},\mathcal{H}]_{\mathcal{B}} \times \mathcal{Q}$. The solution depends linearly and continuously on f and g.

Proof. Item i): This follows directly from the definition of $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$, $\|\cdot\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}$, and Lemma 3.1.

Item ii): Let us consider $h \in \mathcal{H}_{\ker}$. By the definition of $\mathcal{H}_{\ker}$ there exists a sequence $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{V}_{\ker}$ with $v_n \to h$ in $\mathcal{H}$ as $n \to \infty$. Since $\mathcal{B}v_n = 0$, $n \in \mathbb{N}$, the sequence $\{v_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$, too. The limit of v_n is h in $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ as well, because $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ and $\mathcal{H}_{\ker}$ are both embedded in $\mathcal{H}$. Furthermore, the equality $\overline{\mathcal{B}}h = \lim_{n\to\infty} \mathcal{B}v_n = 0$ holds. Hence, $\mathcal{H}_{\ker} \hookrightarrow \ker\overline{\mathcal{B}}$. On the other hand, for every $v \in \ker\overline{\mathcal{B}}$ there exists $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{V}$ with $0 \leftarrow \|v - v_n\|^2_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}} = \|v - v_n\|^2_{\mathcal{H}} + \|\mathcal{B}v_n\|^2_{\mathcal{Q}^*}$ as $n \to \infty$. To show $v \in \mathcal{H}_{\ker}$, we have to proof that v can be approximated in $\mathcal{H}$ by a sequence in $\mathcal{V}_{\ker}$. Therefore, let $\mathcal{P}$ be the orthogonal projection of $\mathcal{V}$ onto $\mathcal{V}_{\ker}$. Then $\mathcal{P}v_n \in \mathcal{V}_{\ker}$ converges to v in $\mathcal{H}_{\ker} \subset \mathcal{H}$, since

$$\|v - \mathcal{P}v_n\|^2_{\mathcal{H}} \lesssim \|v - v_n\|^2_{\mathcal{H}} + \|(\mathrm{id} - \mathcal{P})v_n\|^2_{\mathcal{V}} \lesssim \|v - v_n\|^2_{\mathcal{H}} + \|\mathcal{B}(\mathrm{id} - \mathcal{P})v_n\|^2_{\mathcal{Q}^*} = \|v - v_n\|^2_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}$$

holds. Here, we used in the second estimate Lemma 3.6. This proves $\mathcal{H}_{\ker} \cong \ker\overline{\mathcal{B}}$.

Let now $\mathcal{V}_{\mathrm{c}}$ be defined as in ii) and $\{v_{\mathrm{c},n}\}_{n\in\mathbb{N}} \subset \mathcal{V}_{\mathrm{c}} \hookrightarrow [\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ be a Cauchy sequence in $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$. In particular, $\|\mathcal{B}v_{\mathrm{c},n} - \mathcal{B}v_{\mathrm{c},m}\|_{\mathcal{Q}^*}$ is a Cauchy sequence, which implies by Lemma 3.6 that $v_{\mathrm{c},n}$ converges in $\mathcal{V}_{\mathrm{c}} \subset \mathcal{V}$. Therefore, $\mathcal{V}_{\mathrm{c}}$ is close in $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$. For the direct sum we note that $\mathcal{V}_{\mathrm{c}} + \mathcal{H}_{\ker} \hookrightarrow [\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ is well-defined since $\mathcal{V}_{\mathrm{c}}$ and $\mathcal{H}_{\ker} \cong \ker\overline{\mathcal{B}}$ can be embedded into $[\mathcal{V},\mathcal{H}]_{\mathcal{B}}$. Let $v \in \mathcal{V}_{\mathrm{c}} \cap \mathcal{H}_{\ker}$. Then v is an element of $\mathcal{V}_{\mathrm{c}}$ with $0 = \overline{\mathcal{B}}v = \mathcal{B}v$ since $v \in \mathcal{H}_{\ker}$. By Lemma 3.6 this implies $v = 0$. Therefore, the sum is direct. Let $v_{\mathcal{B}} \in [\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ be arbitrary. Then $v_{\mathcal{B}} = \mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\overline{\mathcal{B}}v_{\mathcal{B}} + (\mathrm{id} - \mathcal{B}^-_{\mathcal{V}_{\mathrm{c}}}\overline{\mathcal{B}})v_{\mathcal{B}}$ holds, where the first summand is an element of $\mathcal{V}_{\mathrm{c}}$ and the second of $\ker\overline{\mathcal{B}}$. Thus, we have $[\mathcal{V},\mathcal{H}]_{\mathcal{B}} \cong \mathcal{V}_{\mathrm{c}} \oplus \mathcal{H}_{\ker}$.

Item iii): We note that $\overline{\mathcal{B}}$ is inf-sup stable by

$$\sup_{v\in[\mathcal{V},\mathcal{H}]_{\mathcal{B}}\setminus\{0\}} \frac{\langle\overline{\mathcal{B}}v,q\rangle}{\|v\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}} \geq \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle\mathcal{B}v,q\rangle}{\|v\|_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}} \geq \frac{1}{C_{\mathcal{V}\hookrightarrow[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}} \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle\mathcal{B}v,q\rangle}{\|v\|_{\mathcal{V}}} \geq \frac{\beta}{C_{\mathcal{V}\hookrightarrow[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}}\|q\|_{\mathcal{Q}}$$

for every $q \in \mathcal{Q}$. Furthermore, the operator $\mathcal{M}$ is elliptic on $\ker\overline{\mathcal{B}} \subset [\mathcal{V},\mathcal{H}]_{\mathcal{B}}$ since for every $v_{\mathcal{B}} \in \ker\overline{\mathcal{B}}$ we have

$$\langle\,\mathcal{M}v_{\mathcal{B}}, v_{\mathcal{B}}\,\rangle = \|v_{\mathcal{B}}\|^2_{\mathcal{H}} = \|v_{\mathcal{B}}\|^2_{\mathcal{H}} + \|\overline{\mathcal{B}}v_{\mathcal{B}}\|^2_{\mathcal{Q}^*} = \|v_{\mathcal{B}}\|^2_{[\mathcal{V},\mathcal{H}]_{\mathcal{B}}}.$$

The properties of the solution of the saddle point problem follows with Theorem 3.8. □

The main idea for the approximation of the Lagrange multiplier is to solve

$$\begin{aligned} \mathcal{M}\dot{u}_n - \overline{\mathcal{B}}^*\lambda_n &= f(t_n, u_n) - \mathcal{A}u_n && \text{in } [\mathcal{V},\mathcal{H}]_\mathcal{B}^*, && (9.43a)\\ \overline{\mathcal{B}}\dot{u}_n &= \dot{g}(t_n) && \text{in } \mathcal{Q}^*, && (9.43b)\end{aligned}$$

with $n = 1,\dots,N$. This problem is associated to the semi-linear operator DAE (9.1) evaluated at the time-point t_n. Anyway, for the analysis so far u_n was considered as an element of $\mathcal{V}$, such that $\mathcal{A}u_n \in \mathcal{V}^*$. But the space $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$ is more restrictive than $\mathcal{V}^*$, since $\mathcal{V}$ has a stronger norm than $[\mathcal{V},\mathcal{H}]_\mathcal{B}$. The same holds for $u(t_n)$, since Theorem 6.15 only predicts $\mathcal{A}u \in C([0,T],\mathcal{V}^*)$. We actually need that $\mathcal{A}u$ is continuous in t_n with images in $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$. We recall that the dynamics in the complement of $\mathcal{V}_\text{ker} \subset \mathcal{V}$ is determined by (9.1b). Therefore, we have to assume that $\mathcal{A}u_{\text{c},n} = \mathcal{A}\mathcal{B}_\mathcal{A}^- g(t_n) = \mathcal{A}u_\text{c}(t_n)$ is an element of $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$. For the part in $\mathcal{V}_\text{ker}$ itself we can use the parabolic smoothing to prove the additional regularity. The needed assumption on the temporal regularity of g and $f(\cdot,u(\cdot))$ are the same as for the convergence of the exponential Euler in Theorem 9.5.

Lemma 9.23 (Smoothness of $u_\text{ker}(t_n)$ and $u_{\text{ker},n}$)**.** *Let Assumptions 9.1–9.3, 9.9, as well as the assumptions of Lemma 9.22. In addition, suppose $f(\cdot,u(\cdot)) \in H^1(0,T;\mathcal{H}^*)$, $g \in H^2(0,T;\mathcal{Q}^*)$, and that $u_0 \in \mathcal{V}$ is consistent. Let u_ker be the solution of (9.2) with initial value $u_{\text{ker},0} = u_0 - \mathcal{B}_\mathcal{A}^- g(0)$ and $u_{\text{ker},n} = u_n - \mathcal{B}_\mathcal{A}^- g(t_n)$ its approximation given by (9.41). Then $\mathcal{A}u_\text{ker}(t_n)$ and $\mathcal{A}u_{\text{ker},n}$ can be extended to elements in $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$ for $n = 1,\dots,N$.*

Proof. By the assumptions, the right-hand side of (9.2) is in $H^1(0,T;\mathcal{H}_\text{ker}^*)$. Therefore, by [Emm04, Th. 8.5.3] we have $t\dot{u}_\text{ker}(t) \in W^{1,2}(0,T;\mathcal{V}_\text{ker},\mathcal{V}_\text{ker}^*)$ and thus $t\mathcal{A}u_\text{ker}(t) = tf(t,u(t)) - t\mathcal{B}_\mathcal{A}^- g(t) - t\dot{u}_\text{ker}(t) \in \mathcal{H}_\text{ker}^* \cap \mathcal{V}^*$ at almost every time-point $t \in (0,T]$. Let $\mathcal{V}_\text{c} := \{v \in \mathcal{V} \,|\, \mathcal{A}^* v \in \mathcal{V}_\text{ker}^0\}$ be defined. Note that, in contrast to our usual choice of $\mathcal{V}_\text{c}$, cf. (3.4), we used the adjoint operator of $\mathcal{A}$. Anyway, the space $\mathcal{V}_\text{c}$ satisfies the condition of Lemma 9.22.ii) by Lemma 3.5 such that $[\mathcal{V},\mathcal{H}]_\mathcal{B} \cong \mathcal{V}_\text{c} \oplus \mathcal{H}_\text{ker}$ holds. Let $t > 0$ and $v = v_\text{c} + h_\text{ker} \in [\mathcal{V},\mathcal{H}]_\mathcal{B}$ be arbitrary with $v_\text{c} \in \mathcal{V}_\text{c}$, $h_\text{ker} \in \mathcal{H}_\text{ker}$. Then

$$\begin{aligned}\langle \mathcal{A}u_\text{ker}(t), v\rangle &= \langle \mathcal{A}^* v_\text{c}, u_\text{ker}(t)\rangle + (\mathcal{A}u_\text{ker}(t), h_\text{ker})\\ &= (\mathcal{A}u_\text{ker}(t), h_\text{ker}) \le \|\mathcal{A}u_\text{ker}(t)\|_{\mathcal{H}_\text{ker}^*}\|h_\text{ker}\|_\mathcal{H} \lesssim \|\mathcal{A}u_\text{ker}(t)\|_{\mathcal{H}_\text{ker}^*}\|v\|_{[\mathcal{V},\mathcal{H}]_\mathcal{B}}\end{aligned}$$

is fulfilled, where we used [BowK14, Th. 4.42] for the last inequality. Therefore, $\mathcal{A}u_\text{ker}(t) \in \mathcal{H}_\text{ker}^*$ can be extended to an element of $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$.

For the temporal approximation $u_{\text{ker},n}$ we note that by its definition $u_{\text{ker},n}$ is the solution of an operator ODE on $(t_{n-1},t_n]$ at the final time point t_n. In particular, the right-hand side of the operator ODE is a polynomial with images in $\mathcal{H}_\text{ker}^*$. With the arguments as for u one then shows that $(t_n - t_{n-1})\mathcal{A}u_{\text{ker},n}$ can be extended to an element of $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$. □

Lemma 9.23 and the additional assumption $\mathcal{A}\mathcal{B}_\mathcal{A}^- g(t_n) \in [\mathcal{V},\mathcal{H}]_\mathcal{B}^*$ guarantees that the saddle point problem (9.43) is well-defined. We only consider $n \ge 1$, since the parabolic smoothing applies only on positive times. If one is interested in λ at the initial time point, i.e., $n = 0$, we have to make additional assumptions on the initial value u_0. The value $\lambda(0)$ and λ_0 are then identical determined by the data u_0, $g(0)$, and $\dot{g}(0)$ as well as (9.43). Therefore, the error would be zero. The difference between $\lambda(t_n)$ and λ_n can be bounded as follows.

Theorem 9.24 (Convergence Order for the Lagrange Multiplier)**.** *Let the assumptions of Theorem 9.17 and Lemma 9.22 be satisfied. Suppose $\mathcal{A}\mathcal{B}_\mathcal{A}^- g$ is a continuous function with images in $[\mathcal{V},\mathcal{H}]_\mathcal{B}^*$. Then $\lambda(t_n) \in \mathcal{Q}$ and $\lambda_n \in \mathcal{Q}$ are well-defined by (9.1) and (9.43), respectively. Further, their difference is bounded by*

$$\|\lambda_n - \lambda(t_n)\|_\mathcal{Q} \lesssim \tau^{\mathsf{p}-\frac{1}{2}} + h.o.t.,$$

$n = 1,\dots,N$, where the constant depends on t_n, L, operators $\mathcal{A}$ and $\mathcal{B}$, $\|f(\cdot,u(\cdot))\|_{W^{q,\infty}(0,t_n,\mathcal{H}^*)}$, as well as $\|g\|_{W^{q+1,\infty}(0,t_n,\mathcal{Q}^*)}$ with $q = \lceil \mathsf{p} \rceil$.

Proof. By Lemmas 9.22 and 9.23 the solutions $\lambda(t_n), \lambda_n \in \mathcal{Q}$ are well-defined. For the difference between them, we consider the saddle point problem

$$\begin{aligned} \mathcal{M}\Delta\dot{u}_n - \overline{\mathcal{B}}^*\Delta\lambda_n &= f(t_n,u_n) - f(t_n,u(t_n)) - \mathcal{A}(u_n - u(t_n)) \qquad &&\text{in } [\mathcal{V},\mathcal{H}]^*_{\mathcal{B}},\\ \overline{\mathcal{B}}\Delta\dot{u}_n &= 0 &&\text{in } \mathcal{Q}^*, \end{aligned}$$

with $n = 1,\dots,N$. It is easy to check that this saddle point problem is well-defined by $\mathcal{H}^* \hookrightarrow [\mathcal{V},\mathcal{H}]^*_{\mathcal{B}}$ and that its unique partial solution is $\Delta\lambda_n = \lambda_n - \lambda(t_n) \in \mathcal{Q}$. By Lemma 9.22 and the embedding $\mathcal{H}^* \hookrightarrow [\mathcal{V},\mathcal{H}]^*_{\mathcal{B}}$ we have

$$\|\lambda_n - \lambda(t_n)\|_{\mathcal{Q}} \lesssim \|f(t_n,u_n) - f(t_n,u(t_n)) - \mathcal{A}(u_n - u(t_n))\|_{[\mathcal{V},\mathcal{H}]^*_{\mathcal{B}}} \leq L\|e_n\|_{\mathcal{V}} + \|\mathcal{A}e_n\|_{[\mathcal{V},\mathcal{H}]^*_{\mathcal{B}}}.$$

To bound the first term, we can use Theorem 9.17. For the second term we note that $u(t_n)$, u_n and $U_{n,i}$ are all consistent. Thus the error e_n is given by the difference of $u_{\mathrm{ker}}(t_n)$ and $u_{\mathrm{ker},n}$. With the argument of Lemma 9.23, an index shift, and (9.36) it is enough to consider the error

$$\begin{aligned} \|\mathcal{A}e_{n+1}\|_{[\mathcal{V},\mathcal{H}]^*_{\mathcal{B}}} &\lesssim \|\mathcal{A}_{\mathrm{ker}}e_{n+1}\|_{\mathcal{H}^*_{\mathrm{ker}}}\\ &\leq \|\mathcal{A}_{\mathrm{ker}}e^{-\tau\mathcal{A}_{\mathrm{ker}}}e_n\|_{\mathcal{H}^*_{\mathrm{ker}}} + \tau\|\mathcal{A}_{\mathrm{ker}}\widetilde{S}_n\|_{\mathcal{H}^*_{\mathrm{ker}}} + \tau\|\mathcal{A}_{\mathrm{ker}}\widehat{S}_n\|_{\mathcal{H}^*_{\mathrm{ker}}} \\ &\quad + \tau\Big\|\mathcal{A}_{\mathrm{ker}}\mathbf{b}^T\sum_{k=0}^{s-1}((\mathbf{J}_n + \tau\mathbf{K}^{\mathbf{c}}_n)\tau\mathbf{A})^k(\mathbf{J}_n + \tau\mathbf{K}^{\mathbf{c}}_n)\boldsymbol{R}_n\Big\|_{\mathcal{H}^*_{\mathrm{ker}}} + \|\mathcal{A}_{\mathrm{ker}}R_{n+1}\|_{\mathcal{H}^*_{\mathrm{ker}}}. \end{aligned} \tag{9.44}$$

By Lemma 5.13 and a revision of Lemmas 9.11, 9.12, we get $\|\mathcal{A}_{\mathrm{ker}}e^{-\tau\mathcal{A}_{\mathrm{ker}}}e_n\|_{\mathcal{H}^*_{\mathrm{ker}}} \lesssim (\tau^{-\frac{1}{2}} + 1)\|e_n\|_{\mathcal{V}}$ as well as the bounds

$$\tau\|\mathcal{A}_{\mathrm{ker}}\widetilde{S}_n\|_{\mathcal{H}^*_{\mathrm{ker}}} \lesssim \|e_n\|_{\mathcal{V}} \quad \text{and} \quad \tau\|\mathcal{A}_{\mathrm{ker}}\widehat{S}_n\|_{\mathcal{H}^*_{\mathrm{ker}}} \lesssim (\tau^2 + \|\boldsymbol{R}_n\|_{\mathcal{V}_s})\|\boldsymbol{R}_n\|_{\mathcal{V}_s} + \text{h.o.t.}$$

for the second and third term of (9.44), respectively. The fourth term of (9.44) can be expanded and bounded like the fourth term of (9.36); cf. (9.38). The estimates then are equivalent where we bound $\|\tau\mathcal{A}_{\mathrm{ker}}b_i(-\tau\mathcal{A}_{\mathrm{ker}})\|_{\mathscr{L}(\mathcal{H}_{\mathrm{ker}},\mathcal{H}^*_{\mathrm{ker}})}$ with Lemma 5.13. By a reconsideration of Lemma 9.10 with the help of Lemma 5.13, one proves that $\|\mathcal{A}_{\mathrm{ker}}R^{[r]}_{n+1}\|_{\mathcal{H}^*_{\mathrm{ker}}}$ is bounded by $\tau^r C\operatorname{ess\,sup}_{t\in[t_n,t_{n+1}]}\|f^{(r)}_u(t)\|_{\mathcal{H}^*_{\mathrm{ker}}}$ from above. These estimates of the terms of (9.44) together with Theorem 9.17 lead to the desired bound. □

Remark 9.25. The proven convergence order in Theorem 9.24 is rather pessimistic and has space for improvements. If, for example, $\mathcal{A}_2$ and f maps (locally) Lipschitz continuously into an interpolation space between $\mathcal{H}$ and $\mathcal{V}$, i.e., into $[\mathcal{V},\mathcal{H}]_\theta$, $\theta \in (0,1)$, then the error for λ is again of order p. This can be proven similarly to [HocO05b] with fractional powers of operators.

9.4. Numerical Examples

In this final section of Chapter 9 we illustrate the performance of the introduced time integration schemes for three numerical examples. The first example is a heat equation with nonlinear dynamic boundary conditions. In the second experiment, we consider a constrained PDE where the smoothness of the right-hand side is minimal such that the convergence orders are the same as predicted in Theorem 9.17. At last, we look at a toy model with a non-homogeneous constraint. The simulation code of all these problems can be found in [Zim20].

Since exponential integrators for (operator) DAEs are based on the exact solution of (operator) DAE systems with polynomial right-hand sides, we first discuss the efficient solution of such systems.

9.4.1. Efficient Solution of Differential-Algebraic Equations in Saddle-Point Form

This subsection is devoted to the efficient calculation of $u_{\ker,n+1}$ and $U_{\ker,n,i}$ in (9.41). This is equivalent to solving the operator DAE (9.1) with a polynomial right-hand side f of degree p and $g = 0$. In particular, this includes the case of homogeneous right-hand side, which is needed in Algorithm 1 and 2. Given a spatial discretization, e.g., by a finite element method, the operator DAE turns into the DAE

$$M\dot{x}(t) + Ax(t) - B^T\lambda(t) = \sum_{k=1}^{p} \frac{f_k}{(k-1)!} t^{k-1}, \tag{9.45a}$$

$$Bx(t) = 0 \tag{9.45b}$$

with consistent initial value $x(0) = x_0$. The matrices $M, A \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_\lambda \times n_x}$ with $n_\lambda \leq n_x$ are possibly sparse. Here, the mass matrix M is symmetric, positive definite and B has full rank. The goal is to find an efficient method to calculate the solution x at a specific time point $t \in [0, T]$.

Let us first recall the ODE case with homogeneous right-hand side, $\dot{x}(t) + Ax(t) = 0$, where we update $M^{-1}A \to A$. There exist various methods to approximate the solution $x(t) = e^{-tA}x_0$ of this linear ODE with initial condition $x(0) = x_0$. For an overview see [MolV78]. These apporaches include Krylov subspace methods based on the Krylov space $\mathcal{K}_n(A, x_0) := \operatorname{span}\{x_0, Ax_0, \dots, A^{n-1}x_0\}$ [EieE06; HocL97; NieW12; Saa92] but also methods based on polynomial interpolation of $e^{-tA}x_0$ [CalKO+16; CalO09; CalVB04]. All these methods have in common that they never use the explicit representation of A, only its action onto a vector. In particular we never have to work with the possibly full matrix $M^{-1}A$ but with the two sparse matrices A and M. These algorithms for evaluating exponential functions can be used for ODEs with polynomial right-hand sides.

Lemma 9.26 ([Al-H11, Ch. 2])**.** *Let $A \in \mathbb{R}^{n_x \times n_x}$, $F = [f_1, \dots, f_p] \in \mathbb{R}^{n_x \times p}$, $x_0 \in \mathbb{R}^{n_x}$ be given and ϕ_k defined as in (5.8). Suppose*

$$e_1 := \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T \in \mathbb{R}^p \quad \text{and} \quad N := \begin{bmatrix} 0 & 0 \\ I_{p-1} & 0 \end{bmatrix} \in \mathbb{R}^{p \times p}$$

are defined and $t \in \mathbb{R}$. Then we have

$$\begin{bmatrix} I_{n_x} & 0_{n_x \times p} \end{bmatrix} \exp\left(t \begin{bmatrix} A & F \\ 0 & N \end{bmatrix} \right) \begin{bmatrix} x_0 \\ e_1 \end{bmatrix} = \exp(tA)x_0 + \sum_{k=1}^{p} t^k \phi_k(tA) f_k.$$

With Lemma 9.26 and the mentioned algorithms for the calculation of $e^{tA}x_0$ we can efficiently solve linear ODEs with polynomial right-hand side. We return to the DAE (9.45). By [EmmM13, Th. 2.4 & Eq. (18)] there exists matrices $X, Y \in \mathbb{R}^{n_x \times n_x}$ such that its unique solution x with consistent initial value $x_0 \in \ker B$ is given by

$$x(t) = e^{tX}x_0 + \int_0^t e^{(t-s)X} Y \sum_{k=1}^{p} \frac{f_k}{(k-1)!} s^{k-1} \, \mathrm{d}s = e^{tX}x_0 + \sum_{k=1}^{p} t^k \phi_k(tX) Y f_k. \tag{9.46}$$

To efficiently solve (9.46) it is left to identify $X \in \mathbb{R}^{n_x \times n_x}$ and $Y \in \mathbb{R}^{n_x \times n_x}$.

Lemma 9.27. *Let $w \in \mathbb{R}^{n_x}$, $A, M \in \mathbb{R}^{n_x \times n_x}$, and $B \in \mathbb{R}^{n_\lambda \times n_x}$ be arbitrary, with M symmetric positive definite and let B have full row rank. Suppose that the linear matrix $Y \in \mathbb{R}^{n_x \times n_x}$ is given* *by* $Y: w \mapsto y$ *where* $y \in \ker B$ *is the unique solution of*

$$\begin{aligned} My - B^T\mu &= w, && (9.47a)\\ By &= 0. && (9.47b) \end{aligned}$$

Define $X := -YA$*. Then* x *given as in* (9.46) *solves the DAE* (9.45) *with initial value* $x_0 \in \ker B$*.*

Proof. By the linearity of (9.45) and (9.46) we can investigate the ϕ_k-functions, $k = 0, \dots, p$, individually. Let us start with $e^{tX}x_0 = \phi_0(tX)x_0$. By (9.47b) and $x_0 \in \ker B$ we have $Be^{tX}x_0 = Be^{-tYA}x_0 \equiv 0$. On the other hand, $\frac{\mathrm{d}}{\mathrm{d}t}e^{tX}x_0 = Xe^{tX}x_0 = -YAe^{tX}x_0$ holds, which means by (9.47a) that for every $t \geq 0$ a $\mu(t)$ exists with $M\frac{\mathrm{d}}{\mathrm{d}t}e^{tX}x_0 - B^T\mu(t) = -Ae^{tX}x_0$. Thus $x(t) = e^{tX}x_0$ and $\lambda(t) = \mu(t)$ solves the DAE (9.45) with homogeneous right-hand sides and $x(0) = e^{0X}x_0 = x_0$. Analogously to ϕ_0 one shows with (5.8) that $Bt^k\phi_k(tX)Yf_k \equiv 0$ and $t^k\phi_k(tX)Yf_k|_{t=0} = 0$, $k \geq 1$. Finally, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}t^k\phi_k(tX)Yf_k \overset{(5.8)}{=} \sum_{\ell=0}^{\infty}\frac{t^{\ell+k-1}}{(\ell+k-1)!}(-YA)^\ell Yf_k \overset{(5.8)}{=} -YAt^k\phi_k(tX)Yf_k + Y\frac{t^{k-1}}{(k-1)!}f_k.$$

Therefore, $t^k\phi_k(tX)Yf_k$ solves (9.45) with the monomial right-hand side $t^{k-1}f_k/(k-1)!$. □

With the previous two lemmas we can approximate the solution of (9.45) by the algorithms for $e^{tA}x_0$. We want to point out that we only work with the sparse saddle-point problem (9.47), not with X or Y. For solving the saddle-point problems there exists a zoo of effective algorithms; see [BanWY90; BenGL05] and the references therein.

Remark 9.28. Since the saddle point problem (9.47) must be solved in the algorithms for the exponential function several times, see Lemmas 9.26 and 9.27, the numerical solution $\tilde{x}$ of (9.45) may not satisfy the constraint (9.45b) due to round-off errors. To prevent a drift-off for longtime simulations, one can project $\tilde{x}$ onto $\ker B$ by solving an additional saddle point problem (9.47) with right-hand side $M\tilde{x}$.

9.4.2. Nonlinear Dynamic Boundary Conditions

In this first experiment we revisit Example 6.13 and consider the linear heat equation with nonlinear dynamic boundary conditions; cf. (6.10). More precisely, we consider the system

$$\begin{aligned} \dot{u} - \kappa\,\Delta u &= 0 && \text{in } \Omega := (0,1)^2, && (9.48a)\\ \dot{u} + \partial_n u + \alpha\, u &= f_{\Gamma_{\mathrm{dyn}}}(t,u) && \text{on } \Gamma_{\mathrm{dyn}} := (0,1)\times\{0\}, && (9.48b)\\ u &= 0 && \text{on } \Gamma_D := \partial\Omega\setminus\Gamma_{\mathrm{dyn}}, && (9.48c) \end{aligned}$$

with $\alpha = 1$, $\kappa = 0.02$, and the nonlinearity $f_{\Gamma_{\mathrm{dyn}}}(t,u)(\xi_1) = -3\cos(2\pi t)\sin(2\pi\xi_1) - u^3(\xi_1)$. As initial condition we set $u(0) = u_0 = \sin(\pi\xi_1)\cos(5\pi\xi_2/2)$. Following [Alt19], the weak formulation of (9.48) can be written in the form

$$\begin{aligned} \begin{bmatrix}\dot{u}\\ \dot{p}\end{bmatrix} + \begin{bmatrix}\mathcal{K} & \\ & \alpha\end{bmatrix}\begin{bmatrix}u\\ p\end{bmatrix} - \mathcal{B}^*\lambda &= \begin{bmatrix}0\\ f_{\Gamma_{\mathrm{dyn}}}(t,p)\end{bmatrix} && \text{in } \mathcal{V}^*, && (9.49a)\\ \mathcal{B}\begin{bmatrix}u\\ p\end{bmatrix} &= 0 && \text{in } \mathcal{Q}^* && (9.49b) \end{aligned}$$

with spaces $\mathcal{V} = H^1_{\Gamma_D}(\Omega) \times H^{1/2}_{00}(\Gamma_{\mathrm{dyn}})$, $\mathcal{H} = L^2(\Omega)\times L^2(\Gamma_{\mathrm{dyn}})$, $\mathcal{Q} = H^{-1/2}_{00}(\Gamma_{\mathrm{dyn}})$, the weak Laplace operator $\mathcal{K}$, and operator $\mathcal{B}(u,p) = u\big|_{\Gamma_{\mathrm{dyn}}} - p$. Here, p denotes a dummy variable modeling the dynamics on the boundary Γ_{dyn}. The constraint (9.49b) couples the two variables u and p. This

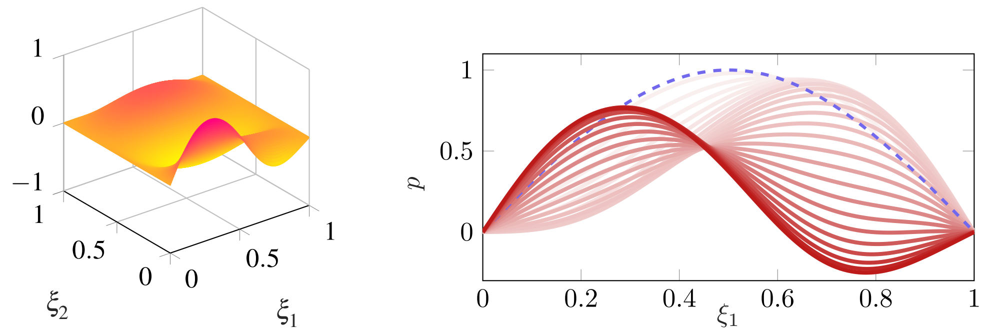

Figure 9.1.: Illustration of the solution (u, p). The left figure shows u at time $t = 0.7$, whereas the right figure includes several snapshots of p in the time interval $[0, 0.7]$. The graph of p becomes darker over time and the dashed line shows the initial value of p.

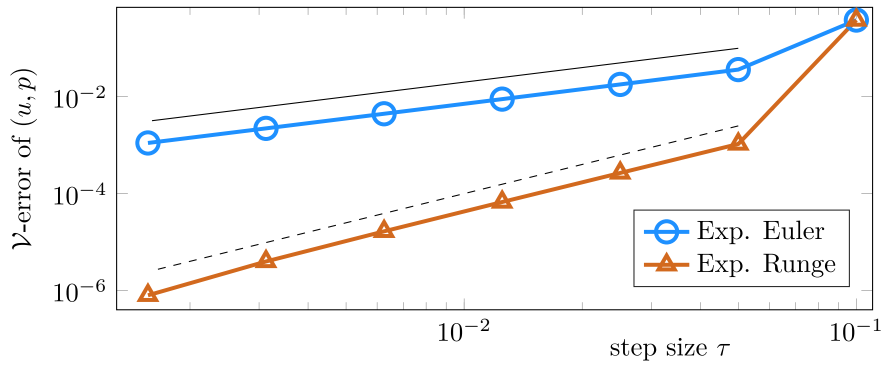

Figure 9.2.: Convergence history for the error in (u, p), measured in the $\mathcal{V}$-norm. The solid line shows first and the dashed line second order rate.

example fits into the framework of the considered semi-linear operator DAEs (9.1) with $g = 0$. Further, the nonlinearity satisfies the assumptions of the convergence results in Theorems 9.5 and 9.8 due to the Sobolev embedding $H^{1/2}_{00}(\Gamma_{\text{dyn}}) \subset H^{1/2}(\Gamma_{\text{dyn}}) \hookrightarrow L^6(\Gamma_{\text{dyn}})$ [Rou13, Cor. 1.22 & p. 18].

For the simulation we consider a spatial discretization by bilinear finite elements on a uniform mesh with mesh size $h = 1/128$. The associated matrices are determined by the Matlab software package AFEM [CarGK+10]. As time-stepping schemes we use the exponential Euler method (9.4) and the exponential Runge method (9.10) given by Algorithm 1 on page 128 and by Algorithm 2 on page 131, respectively. The simulated time horizon is $[0, 0.7]$. For the calculation of the exponential functions we use the Matlab routine phipm from [NieW12], which allows function handles as input. The initial value of p is chosen in a consistent manner, i.e., $p_0 = u_0|_{\Gamma_{\text{dyn}}}$. An illustration of the dynamics is given in Figure 9.1.

The convergence results of the exponential Euler scheme of Section 9.1 and the exponential Runge method introduced in Section 9.2 are displayed in Figure 9.2 and show first and second order convergence, respectively. These convergence orders are predicted by Theorem 9.5 and Lemma 9.18, respectively, where we expect the solution p to be as smooth such that an abstract function $\widetilde{p} \in L^2(0, T; H^{1/2}_{00}(\Gamma_{\text{dyn}}))$ exists, where $(f_{\Gamma_{\text{dyn}}}(t, p(t)), w)_{L^2(\Gamma_{\text{dyn}})} = (\widetilde{p}(t), w)_{L^2(\Gamma_{\text{dyn}})}$ holds for all $w \in H^{1/2}_{00}(\Gamma_{\text{dyn}})$.

Finally, we note that the computations remain stable for very coarse step sizes τ, since the possible temporal step size is not limited through the stiffness of the operator $\mathcal{A}$.

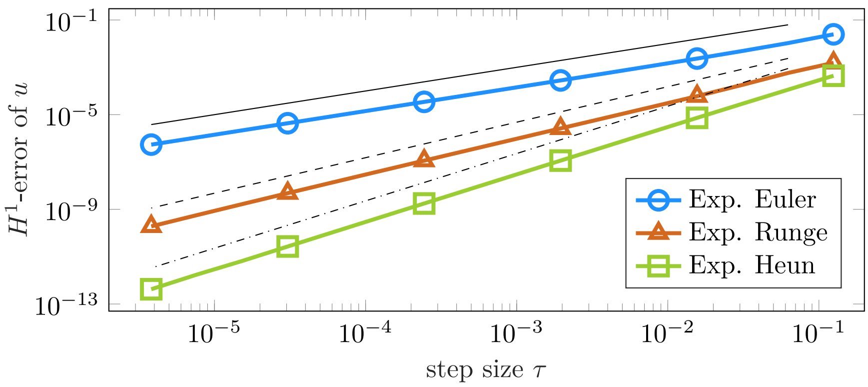

Figure 9.3.: Convergence history for the error in u, measured in the $H^1(0,1)$-norm. For comparison we added lines of different slopes: slope 1 is represented by the solid line, 1.5 by the dashed line, and 2 by the dash-dotted line.

9.4.3. Slow Convergence

In the first numerical example in Subsection 9.4.2 we saw that the order of the two-stage method of Section 9.2 can increase if $f(t,u(t))$ is regular enough in time and space. In this subsection we investigate the convergence order under less regular data. For this, let $\gamma(\xi) := -2\sum_{k=1}^{\infty}\frac{\sin(k\pi\xi)}{k^{0.51}} \in L^2(0,1)$ be defined. We choose Φ such that

$$u(\xi,t) = \sum_{k=1}^{\infty}\Big(e^{-(\pi k)^2 t} + \frac{\pi k}{(\pi k)^2-1}\big[e^{-t} - e^{-(\pi k)^2 t}\big]\Big)\frac{\sin(k\pi\xi)}{k^{1.51}} \tag{9.50}$$

is the solution of the PDE

$$\begin{aligned} \dot u(\xi,t) - \partial_{\xi\xi}u(\xi,t) &= (\mathcal{D}^{0.49}u(\cdot,t))^2\gamma(\xi) + \Phi(\xi,t) && \text{in } \Omega = (0,1),\\ u(\xi,t) &= 0 && \text{on } \partial\Omega = \{0,1\}, \end{aligned}$$

with the bounded functional $\mathcal{D}^\alpha\colon v \mapsto \sqrt{2}\sum_{k=1}^{\infty}(2k-1)^{\alpha-1}(\partial_\xi v, \cos((2k-1)\pi\,\cdot))_{L^2(0,1)}$, $\alpha < 0.5$, from $H^1(0,1)$ to $\mathbb{R}$. The solution (9.50) and the right-hand side are constructed such that the convergence rates of the exponential integrators do not improve. Note that the whole right-hand side of the PDE is equal to $\pi e^{-t}\sum_{k=1}^{\infty}\frac{\sin(k\pi\xi)}{k^{0.51}}$ for the specific choice of the solution (9.50). Therefore, it is an element of $C^\infty([0,T], L^2(0,1)) \hookrightarrow L^2(0,T;L^2(0,1))$ and by Theorem 4.25 the solution u satisfies $u \in H^1(0,T;L^2(0,1)) \cap C([0,T],H^1(0,1))$.

For the simulation we rewrite the problem as a constrained PDE by introducing the boundary condition as constraint; cf. Example 6.2. As spatial discretization we use spectral finite elements with 1000 degrees of freedom enriched with linear polynomials; cf. Subsection 8.6.2. For the temporal discretization we predetermine the precise value of $\phi_k(-\tau M_{\text{ker}}^{-1}A_{\text{ker}})$, $k=0,1,2$. Here, M_{ker} and A_{ker} are the discrete versions of $\mathcal{M} = \mathcal{R}_{L^2(0,1)}$ and $\mathcal{A} = \mathcal{R}_{H_0^1(0,1)}$, respectively, restricted to the kernel of the (discrete) trace operator, i.e., restricted to the spectral finite elements $\operatorname{span}\{\sin(\pi\xi),\dots,\sin(1000\pi\xi)\}$. The determination of $\phi_k(-\tau M_{\text{ker}}^{-1}A_{\text{ker}})$ then is reasonable since M_{ker} and A_{ker} are diagonal matrices. Figure 9.3 illustrates the convergence rate for the exponential Euler method (9.4), the exponential Runge scheme (9.10), and exponential Heun method (9.40) with $c_2 = c_3 = 1$. The numerical convergence rates coincide with the proven ones of Theorem 9.13.

Table 9.3 summarizes the errors of the associated Lagrange multiplier $\lambda = [\,\partial_x u|_{x=0},\, -\partial_x u|_{x=1}\,]^T$. Since $\gamma \in [H^1(0,1), L^2(0,1)]_{0.99}$ by [LioM72, Ch. 1 Th. 9.1 & Rem. 10.5], Remark 9.25 predicts that

Table 9.3.: Convergence history for the error in λ measured in the Euclidean norm.

	Exp. Euler		Exp. Runge		Exp. Heun	
	Error	Rate	Error	Rate	Error	Rate
$\tau = 2^{-3}$	$1.098 \cdot 10^{-1}$		$7.535 \cdot 10^{-3}$		$1.232 \cdot 10^{-3}$	
$\tau = 2^{-6}$	$1.019 \cdot 10^{-2}$	1.14	$4.095 \cdot 10^{-4}$	1.40	$2.720 \cdot 10^{-5}$	1.83
$\tau = 2^{-9}$	$1.119 \cdot 10^{-3}$	1.06	$1.478 \cdot 10^{-5}$	1.60	$4.237 \cdot 10^{-7}$	2.00
$\tau = 2^{-12}$	$1.286 \cdot 10^{-4}$	1.04	$5.555 \cdot 10^{-7}$	1.58	$6.288 \cdot 10^{-9}$	2.02
$\tau = 2^{-15}$	$1.525 \cdot 10^{-5}$	1.03	$2.143 \cdot 10^{-8}$	1.57	$9.453 \cdot 10^{-11}$	2.02
$\tau = 2^{-18}$	$1.845 \cdot 10^{-6}$	1.02	$8.034 \cdot 10^{-10}$	1.58	$1.380 \cdot 10^{-12}$	2.03

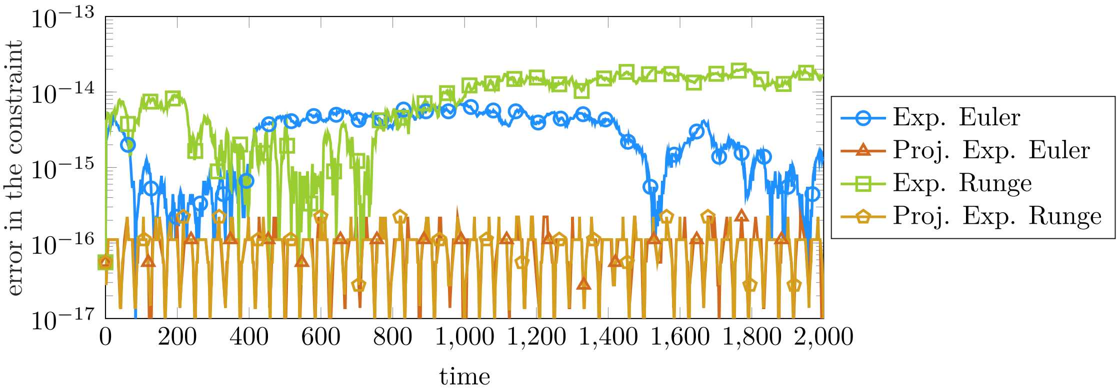

Figure 9.4.: Error in the constraint over time.

the convergence rate of u and λ are the same. The calculated rates in Table 9.3 confirm this.

9.4.4. Non-Homogeneous Constraints

As last example we consider a finite-dimensional DAE (2.3) with

$$A = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}, \quad M = \frac{1}{6}\begin{bmatrix} 4 & 1 & & \\ 1 & 4 & 1 & \\ & 1 & 4 & 1 \\ & & 1 & 4 \end{bmatrix}, \quad B^T = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad x_0 = \begin{bmatrix} -0.8 \\ 0.6 \\ 0.4 \\ -0.2 \end{bmatrix}$$

and right-hand sides $f_i(t, x) = e^{-t}\cos(x_i)$, $i = 1, \dots, 4$, and $g(t) = \sin(\pi t)$. In contrast to the previous two examples, the constraint is not homogeneous. For the simulations we use as temporal step size $\tau = 2^{-9}$ and the method described in Subsection 9.4.1. The error of the approximation in the constraint is illustrated in Figure 9.4. One observes that the error of the exponential Euler method mostly is smaller than of the exponential Runge method – even if both are of moderate size. A possible explanation is that the exponential Runge method needs twice as many evaluations of the ϕ-functions per time step such that the drift-off is stronger; see Remark 9.28. In addition to the exponential Euler and Runge methods, we modified these schemes such that $u_{\ker,n}$ is projected onto $\ker B$ after every time step. The associated methods have the prefix *Proj.* in Figure 9.4. These methods do not suffer from drift-offs and the errors of the constraint are in the range of double precision $\mathcal{O}(10^{-16})$.

The convergence orders of the methods with and without the projection steps are the same, i.e., first and second order. We omit the corresponding error plots here.

10. Summary and Outlook

In this thesis we have analyzed the existence, uniqueness, and regularity of solutions of the semi-linear, constrained PDE in the framework of the operator DAE (1.1); see Part B. Furthermore, we have investigated their temporal discretization by implicit, algebraically stable Runge-Kutta methods and explicit exponential integrators in Part C.

In Part A, we provided the essential mathematical concepts for the analysis of constrained PDEs. In particular, we introduced the abstract framework of operator DAEs, in which we considered the constrained PDEs. While most results in this part are well-known, we also derived novel contributions. These included the analysis of strongly measurable operator-valued functions and their generalized derivatives in Section 4.1. We proved the existence and uniqueness of solutions of Volterra integral equations with values in abstract Banach spaces; Theorem 4.19. In Lemmas 5.11 and 5.13, we estimated the norm of the $\phi_k(-t\mathcal{A})$-functions as maps from $\mathcal{V}$ and $\mathcal{H} \cong \mathcal{H}^*$ into the spaces $\mathcal{V}$, $\mathcal{H}$, and the domain of $\mathcal{A}$.

Part B was devoted to the analysis of solutions of the operator DAE (1.1). Starting with systems with time-independent operators, we extended well-known existence, uniqueness, and regularity results in Sections 6.1 and 6.2. Here, the right-hand side f splits into parts with images in $\mathcal{V}^*$ and $\mathcal{H}^*$ and the right-hand side g and its derivative are L^1-functions. This can be seen as an abstract linear extension of the solution operator. We showed that controlled operator DAEs of the form (6.6) satisfy a dissipation inequality in a distributional sense, see Lemma 6.12, and proved the continuous dependence of the unique solutions of systems with a state-dependent f on the data u_0 and g in Section 6.4. Afterwards, we analyzed the solution of operator DAEs with time-dependent operators in Chapter 7. Since the kernel of $\mathcal{B}$ is time-dependent, we introduced the operator ODE (7.45), whose fundamental solution $\mathcal{W}$ tracks the kernel of $\mathcal{B}$ over time; see Theorem 7.31. This allowed us to restate the system as the operator DAE (7.50) with an operator $\widetilde{\mathcal{B}} = \mathcal{B}\mathcal{W}$ with a constant kernel. The derivatives $\dot{u}$ and $\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}u)$ become $\mathcal{W}^* \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{W}u)$ and $\mathcal{W}^* \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{M}\mathcal{W}u)$, respectively. For the existence results, we temporally discretized the (restated) system by the implicit Euler scheme. We showed that weak/weak* limits of the time-discrete solutions solve the (restated) operator DAE, see Sections 7.1.1 and 7.2.2. The existence theorem of the Lagrange multiplier λ in a distributional sense was derived by an Volterra integral equation; see Theorem 7.42. The uniqueness of the solution (u, λ) was proven under additional assumptions on $\mathcal{W}$ or $\mathcal{A}$ in Subsection 7.2.3 and Section 7.3. For the first case, we used that $u \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{H})$ with $\frac{\mathrm{d}}{\mathrm{d}t}(\widetilde{\mathcal{M}}u) \in L^2(0,T;\mathcal{V}^*)$ and a uniformly elliptic operator $\widetilde{\mathcal{M}}\colon [0,T] \to \mathscr{L}(\mathcal{H},\mathcal{H}^*)$ has a continuous representative. The details are presented in Theorem 7.18.

Table 10.1 summarizes the main results obtained in Part B on the solutions of the operator DAE (1.1). In addition, we made comments on possible extensions of the solvability results in Remarks 6.10, 7.25, 7.26, and 7.52. The latter two, in particular, consider operator DAEs with time-dependent operators and a state-dependent right-hand side $f = f(\,\cdot\,, u(\,\cdot\,))$.

The temporal discretization of the operator DAE (1.1) was analyzed in Part C. Here, the operators $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{B}$ are time-independent. We focused on the saddle-point structure of the operator DAE (1.1) and chose time-stepping schemes that exploit this structure. For the Runge-Kutta methods we introduced the regularization (8.4), which is still in saddle-point form and where a spatial discretization leads directly to a DAE of index one. This implies that the system is more stable than the original formulation although the solution set remains unchanged; see Lemma 8.5 for

Table 10.1.: Main results on the solutions of the operator DAE (1.1) in Part B.

	Existence of solutions	Uniqueness of solutions	Continuous sol. operator	More regular solutions
Constant operators…	Th. 6.7	Th. 6.7	Th. 6.7	Th. 6.8, 6.9
…with state-dependent f	Th. 6.15	Th. 6.15	Th. 6.19	
Time-dependent $\mathcal{M}$ and $\mathcal{A}$	Th. 7.14, 7.21	Th. 7.19, 7.21	Th. 7.19, 7.21	Th. 7.23, 7.24
Time-dependent $\mathcal{A}$ and $\mathcal{B}$	Th. 7.41, 7.42	Th. 7.44, 7.47	Th. 7.44, 7.47	
Time-dependent $\mathcal{M}$, $\mathcal{A}$, and $\mathcal{B}$	Th. 7.49	Th. 7.50, 7.51	Th. 7.50, 7.51	

details. Under the minimal assumptions on the data of Section 6.1, we showed that the sequence of stationary solutions given by the regularized operator DAE under the implicit Euler method converge strongly to the solution of the original operator DAE; Theorem 8.14. We proved strong convergence in more restrictive spaces under the assumptions for more regular solutions; see Section 6.2. Similar results were proven for algebraically stable, L-stable Runge-Kutta schemes in Theorems 8.27 and 8.30. Here, we considered $f \in L^2(0, T; \mathcal{V}^*)$, since estimates of the internal stages in the discrete counterpart of $C([0, T], \mathcal{H}_s)$ are not possible. Theorem 8.35 and Remark 8.36 extend the results to non-L-stable methods. For schemes with $R(\infty) = \pm 1$, we introduced a special discretization of the right-hand side g. The convergence order was analyzed in Section 8.5. In Theorem 8.37 we showed for algebraically stable Runge-Kutta methods with $R(\infty) \in (-1, 1)$ a convergence rate of $\mathsf{q} + 1$ for the state u and $\mathsf{q} + 1/2$ for the Lagrange multiplier λ, where q denotes the stage order. For the Euler scheme or for systems with an operator $\mathcal{A}$ satisfying the usual splitting $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$, the convergence order of λ can be improved to $\mathsf{q} + 1$; Lemma 8.41 and Theorem 8.42. In addition, we made comments on time-dependent operators $\mathcal{A}$ and $\mathcal{B}$ as well as on perturbed data, where, for example, a spatially discretized operator DAE can be interpreted as a perturbed infinite-dimensional system. The theory of Chapter 8 was verified by numerical experiments.

In Chapter 9, we derived a novel class of time integration schemes for semi-linear operator DAEs. For this, we combined explicit exponential integrators for the dynamical part of the system with (stationary) saddle point problems for the part of the solution with images in $\mathcal{V}_c$. For the methods based on the exponential Euler and the exponential Runge methods we have proven a convergence order of one and one and a half, respectively, under minimal assumptions on the right-hand side $f = f(\,\cdot\,, u(\cdot))$; Theorem 9.5 and 9.8. Under regularity assumptions on f with respect to the second arguments, we derived order conditions for the approximation of the state u up to order three or four, depending on the image of f in Subsection 9.3.2.1. The Lagrange multiplier λ was approximated by solving an additional saddle-point problem, which leads to a decrease of half an order; Theorem 9.24. Numerical experiments were presented to validate the theoretical results.

Although operator DAEs are powerful in the analysis of constrained PDEs, their mathematical understanding is still far from complete. Linked to this thesis, nonlinear constrained PDEs, e.g., the Stefan problem [DiPVY15], can be studied by linearization, which leads to systems with time-dependent operators. Here, the assumption that the operator $\mathcal{B}$ and its derivative have pointwise the same domain, is maybe too restrictive. In modeling moving boundary conditions [Alt14], for example, the derivative $\dot{\mathcal{B}}$ asks for more regular functions than the operator $\mathcal{B}$ itself. These differences are subject of future works. Also, extensions of the results of this thesis to constrained hyperbolic PDEs in saddle-point form with time- and state-dependent operators are conceivable.

Investigations of other temporal discretization schemes like multistep and Rosenbrock methods can be considered in the future. With the ideas presented in this thesis, algorithms based on other exponential integrators besides explicit ones can be constructed. For explicit exponential integrators, weakly and very weakly satisfied order conditions, see [HocO05a; HocO10], may improve the predicted convergence order. For example, for the PDE with dynamic boundary conditions in Subsection 9.4.2 the exponential Heun method is of third order [Wie20]. Future studies may show that the observed convergence rate is due to weakened order conditions. Explicit exponential integrators for constrained PDEs with time-dependent coefficients can be considered by extending the ideas of [HipHO12; HipHO14]. Further convergence results for Runge-Kutta methods and systems with rough right-hand side would be interesting as well. Here, the parabolic smoothing and energy estimates can be used as in [Emm00]. For time-dependent operators $\mathcal{M}$, $\mathcal{B}$, and $\mathcal{A}$ one can possibly adapt the results of [GonO99]. The effect of a first-order approximation of the operator-valued function $\mathcal{W}$ as in [KunM07] for a time-dependent operator $\mathcal{B}$ can be subject to future work. In the context of longtime simulations, the preservation of the dissipation inequality for constrained PDEs is of importance under temporal discretization. Similar to [Egg19; HaiW96; MehM19], one can consider discontinuous Galerkin time-stepping, partitioned Runge-Kutta, or collocation methods. Here, a regularization for the constrained PDEs is desirable on the infinite-dimensional level, that maintains the port-Hamiltonian structure. The temporal discretization of constrained hyperbolic PDEs in saddle-point form is another possible research topic. Here, schemes like the Gautschi-type methods [HocL99], which deploy the possible second temporal derivatives, are favorable.

Finally, most results in this thesis rely on the saddle-point structure of the problem. It would be of interest whether statements can be translated to constrained PDEs of more general form.

Bibliography

[AdaF03] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. Second edition. Elsevier/Academic Press, Amsterdam, 2003.

[Al-H11] A. H. Al-Mohy and N. J. Higham. "Computing the action of the matrix exponential, with an application to exponential integrators". *SIAM J. Sci. Comput.* 33 (2): 488–511 (2011).

[AlpES15] A. Alphonse, C. M. Elliott, and B. Stinner. "An abstract framework for parabolic PDEs on evolving spaces". *Port. Math.* 72 (1): 1–46 (2015).

[Alt16] H. W. Alt. *Linear Functional Analysis: An Application-Oriented Introduction*. Springer-Verlag, London, 2016.

[Alt14] R. Altmann. "Moving Dirichlet boundary conditions". *ESAIM: M2AN* 48 (6): 1859–1876 (2014).

[Alt15] R. Altmann. "Regularization and Simulation of Constrained Partial Differential Equations". PhD thesis. Technische Universität Berlin, 2015.

[Alt19] R. Altmann. "A PDAE formulation of parabolic problems with dynamic boundary conditions". *Appl. Math. Lett.* 90: 202–208 (2019).

[AltH15] R. Altmann and J. Heiland. "Finite element decomposition and minimal extension for flow equations". *ESAIM Math. Model. Numer. Anal.* 49 (5): 1489–1509 (2015).

[AltH18] R. Altmann and J. Heiland. "Regularization and Rothe discretization of semi-explicit operator DAEs". *Int. J. Numer. Anal. Model.* 15 (3): 452–478 (2018).

[AltO17] R. Altmann and A. Ostermann. "Splitting methods for constrained diffusion-reaction systems". *Comput. Math. Appl.* 74 (5): 962–976 (2017).

[AltZ18a] R. Altmann and C. Zimmer. "On the smoothing property of linear delay partial differential equations". *J. Math. Anal. Appl.* (2): 916–934 (2018).

[AltZ18b] R. Altmann and C. Zimmer. "Runge-Kutta methods for linear semi-explicit operator differential-algebraic equations". *Math. Comp.* 87 (309): 149–174 (2018).

[AltZ18c] R. Altmann and C. Zimmer. *Time discretization schemes for hyperbolic systems on networks by ε-expansion*. ArXiv e-print 1810.04278. 2018.

[AltZ19] R. Altmann and C. Zimmer. "Time discretization of nonlinear hyperbolic systems on networks". *PAMM* 19 (1): e201900057 (2019).

[AltZ20] R. Altmann and C. Zimmer. "Exponential integrators for semi-linear parabolic problems with linear constraints". In: *Progress in Differential-Algebraic Equations II*. Ed. by T. Reis, S. Grundel, and S. Schöps. Springer International Publishing, Cham, 2020, pp. 137–164.

[AngR07] L. Angermann and R. Rang. "Perturbation index of linear partial differential-algebraic equations with a hyperbolic part". *Centr. Eur. J. Math.* 5: 19–49 (2007).

[AscP98] U. M. Ascher and L. R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1998.

[BanWY90] R. E. Bank, B. D. Welfert, and H. Yserentant. “A class of iterative methods for solving saddle point problems”. *Numer. Math.* 56 (7): 645–666 (1990).

[BeaMX+18] C. Beattie, V. Mehrmann, H. Xu, and H. Zwart. “Linear port-Hamiltonian descriptor systems”. *Math. Control Signals Syst.* 30 (4): 17 (2018).

[BenH15] P. Benner and J. Heiland. “Time-dependent Dirichlet conditions in finite element discretizations”. *ScienceOpen Research*: (2015).

[BenGL05] M. Benzi, G. H. Golub, and J. Liesen. “Numerical solution of saddle point problems”. *Acta Numer.* 14: 1–137 (2005).

[BerL76] J. Bergh and J. Löfström. *Interpolation Spaces: An Introduction.* Springer-Verlag, Berlin, 1976.

[BirS87] M. S. Birman and M. Z. Solomjak. *Spectral Theory of Self-Adjoint Operators in Hilbert Space.* D. Reidel Publishing Company, Dordrecht, 1987.

[BlaN10] O. Blasco and J. van Neerven. “Spaces of operator-valued functions measurable with respect to the strong operator topology”. In: *Vector Measures, Integration and Related Topics.* Ed. by G. P. Curbera, G. Mockenhaupt, and W. J. Ricker. Birkhäuser, Basel, 2010, pp. 65–78.

[BonG01] S. Bonaccorsi and G. Guatteri. “A variational approach to evolution problems with variable domains”. *J. Differential Equations* 175 (1): 51–70 (2001).

[Bor90] F. A. Bornemann. “An adaptive multilevel approach to parabolic equations. I. General theory and 1D-implementation”. *IMPACT Comput. Sci. Engrg.* 2: 279–317 (1990).

[Bor91] F. A. Bornemann. “An adaptive multilevel approach to parabolic equations. II. Variable-order time discretization based on a multiplicative error correction”. *IMPACT Comput. Sci. Engrg.* 3: 93–122 (1991).

[Bor92] F. A. Bornemann. “An adaptive multilevel approach to parabolic equations. III. 2D error estimation and multilevel preconditioning”. *IMPACT Comput. Sci. Engrg.* 4: 1–45 (1992).

[BowK14] A. Bowers and N. J. Kalton. *An Introductory Course in Functional Analysis.* Springer, New York, 2014.

[Bra07] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics.* Third edition. Cambridge University Press, New York, 2007.

[BreCP96] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1996.

[Bre10] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, New York, 2010.

[BreF91] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, New York, 1991.

[CalKO+16] M. Caliari, P. Kandolf, A. Ostermann, and S. Rainer. “The Leja method revisited: Backward error analysis for the matrix exponential”. *SIAM J. Sci. Comput.* 38 (3): A1639–A1661 (2016).

[CalO09] M. Caliari and A. Ostermann. “Implementation of exponential Rosenbrock-type integrators”. *Appl. Numer. Math.* 59 (3–4): 568–581 (2009).

[CalVB04] M. Caliari, M. Vianello, and L. Bergamaschi. “Interpolating discrete advection-diffusion propagators at Leja sequences”. *J. Comput. Appl. Math.* 172 (1): 79–99 (2004).

[CalP06] M. P. Calvo and C. Palencia. "A class of explicit multistep exponential integrators for semilinear problems". *Numer. Math.* 102 (3): 367–381 (2006).

[CamM99] S. L. Campbell and W. Marszalek. "The index of an infinite dimensional implicit system". *Math. Comput. Model. Dyn. Sys.* 5 (1): 18–42 (1999).

[CarJ96] H. S. Carslaw and J. C. Jaeger. *Conduction of Heat in Solids.* Second edition. Clarendon Press, Oxford, 1996.

[CarGK+10] C. Carstensen, J. Gedicke, L. Kern, J. Neumann, H. Rabus, and M. Rozova. *AFEM.* Software package and documentation. Humboldt Universität zu Berlin, 2010.

[Cer60] J. Certaine. "The solution of ordinary differential equations with large time constants". In: *Mathematical Methods for Digital Computers.* Ed. by A. Ralston and H. S. Wilf. Wiley, New York, 1960, pp. 128–132.

[ChaS10] S. Chaturantabut and D. Sorensen. "Nonlinear model reduction via discrete empirical interpolation". *SIAM J. Sci. Comput.* 32 (5): 2737–2764 (2010).

[ChiFP06] R. Chill, E. Fašangová, and Prüss. "Convergence to steady state of solutions of the Cahn-Hilliard and Caginalp equations with dynamic boundary conditions". *Math. Nachr.* 279 (13–14): 1448–1462 (2006).

[ColF15a] P. Colli and T. Fukao. "Cahn-Hilliard equation with dynamic boundary conditions and mass constraint on the boundary". *J. Math. Anal. Appl.* 429: 1190–1213 (2015).

[ColF15b] P. Colli and T. Fukao. "The Allen-Cahn equation with dynamic boundary conditions and mass constraints". *Math. Methods Appl. Sci.* 38: 3950–3967 (2015).

[CoxM02] S. M. Cox and P. C. Matthews. "Exponential time differencing for stiff systems". *J. Comput. Phys.* 176 (2): 430–455 (2002).

[Cro75] M. Crouzeix. "Sur l'approximation des équations différentielles opérationelles linéaires par des méthodes de Runge-Kutta". PhD thesis. Université Paris VI, Paris, 1975.

[DauL90] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 3, Spectral Theory and Applications.* Springer-Verlag, Berlin, 1990.

[DauL92] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5, Evolution Problems. I.* Springer-Verlag, Berlin, 1992.

[DauL93] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 6, Evolution Problems. II.* Springer-Verlag, Berlin, 1993.

[Deb04] K. Debrabant. "Numerische Behandlung Linearer und Semilinearer Partieller Differentiell-Algebraischer Systeme mit Runge-Kutta-Methoden". PhD thesis. Martin-Luther-Universität Halle-Wittenberg, 2004.

[DebS05] K. Debrabant and K. Strehmel. "Convergence of Runge-Kutta methods applied to linear partial differential-algebraic equations". *Appl. Numer. Math.* 53 (2–4): 213–229 (2005).

[DiPVY15] D. A. Di Pietro, M. Vohralík, and S. Yousef. "Adaptive regularization, linearization, and discretization and a posteriori error control for the two-phase Stefan problem". *Math. Comp.* 84 (291): 153–186 (2015).

[Edg77] G. A. Edgar. "Measurability in a Banach Space". *Indiana Univ. Math. J.* 26 (4): 663–677 (1977).

[EdwTF+94] W. S. Edwards, L. S. Tuckerman, R. A. Friesner, and D. C. Sorensen. “Krylov methods for the incompressible Navier-Stokes equations”. *J. Comput. Phys.* 110 (1): 82–102 (1994).

[Egg19] H. Egger. “Structure preserving approximation of dissipative evolution problems”. *Numer. Math.*: 85–106 (2019).

[EggKL+18] H. Egger, T. Kugler, B. Liljegren-Sailer, N. Marheineke, and V. Mehrmann. “On structure-preserving model reduction for damped wave propagation in transport networks”. *SIAM J. Sci. Comput.* 40: A331–A365 (2018).

[EieE06] M. Eiermann and O. G. Ernst. “A restarted Krylov subspace method for the evaluation of matrix functions”. *SIAM J. Numer. Anal.* 44 (6): 2481–2504 (2006).

[Emm00] E. Emmrich. *Quantitative error estimates for the implicit Euler scheme for linear evolutionary problems with rough data.* Preprint No. 671. Germany: Fachbereich Mathematik, TU Berlin, 2000.

[Emm01] E. Emmrich. *Analysis von Zeitdiskretisierungen des inkompressiblen Navier-Stokes-Problems.* Cuvillier, Göttingen, 2001.

[Emm04] E. Emmrich. *Gewöhnliche und Operator-Differentialgleichungen: Eine Integrierte Einführung in Randwertprobleme und Evolutionsgleichnugen für Studierende.* Vieweg, Wiesbaden, 2004.

[Emm05] E. Emmrich. “Stability and error of the variable two-step BDF for semilinear parabolic problems”. *J. Appl. Math. and Comput.* 19 (1–2): 33–55 (2005).

[EmmM13] E. Emmrich and V. Mehrmann. “Operator differential-algebraic equations arising in fluid dynamics”. *Comput. Methods Appl. Math.* 13 (4): 443–470 (2013).

[EmmT10] E. Emmrich and M. Thalhammer. “Stiffly accurate Runge-Kutta methods for nonlinear evolution problems governed by a monotone operator”. *Math. Comp.* 79 (270): 785–806 (2010).

[EngN00] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations.* Springer-Verlag, New York, 2000.

[FavY99] A. Favini and A. Yagi. *Degenerate Differential Equations in Banach Spaces.* Springer-Verlag, Boca Raton, 1999.

[GajGZ74] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferential-Gleichungen.* Akademie-Verlag, Berlin, 1974.

[Gal07] C. G. Gal. “Global well-posedness for the non-isothermal Cahn-Hilliard equation with dynamic boundary conditions”. *Adv. Differential Equ.* 12 (11): 1241–1274 (2007).

[GalG08] C. G. Gal and M. Grasselli. “The non-isothermal Allen-Cahn equation with dynamic boundary conditions”. *Discrete Cont. Dyn. Syst. Ser. A* 22: 1009–1040 (2008).

[Gav68] G. R. Gavalas. *Nonlinear Differential Equations of Chemically Reacting Systems.* Springer-Verlag, Berlin, 1968.

[GeaGL85] C. W. Gear, G. K. Gupta, and B. Leimkuhler. “Automatic integration of Euler-Lagrange equations with constraints”. *J. Comput. Appl. Math.* 12–13: 77–90 (1985).

[GilT01] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer-Verlag, Berlin, 2001.

[GolKT92] H. Goldberg, W. Kampowsky, and F. Tröltzsch. “On Nemytskij operators in L_p-spaces of abstract functions”. *Math. Nachr.* 155: 127–140 (1992).

[Gol06] G. R. Goldstein. “Derivation and physical interpretation of general boundary conditions”. *Adv. Differential Equ.* 11 (4): 457–480 (2006).

[GolP73] G. H. Golub and V. Pereyra. “The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate”. *SIAM J. Numer. Anal.* 10: 413–432 (1973).

[GolV96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Third edition. Johns Hopkins University Press, Baltimore, MD, 1996.

[GonO99] C. González and A. Ostermann. “Optimal convergence results for Runge-Kutta discretizations of linear nonautonomous parabolic problems”. *BIT* 39: 79–95 (1999).

[GriLS90] G. Gripenberg, S.-O. Londen, and O. Staffans. *Volterra Integral and Functional Equations*. Cambridge University Press, Cambridge, 1990.

[GruJH+14] S. Grundel, L. Jansen, N. Hornung, T. Clees, C. Tischendorf, and P. Benner. “Model order reduction of differential algebraic equations arising from the simulation of gas transport networks.” In: *Progress in Differential-Algebraic Equations. Deskriptor 2013. Proceedings of the 8th Workshop on Coupled Descriptor Systems, Geseke, Germany, March 4–7, 2013.* Springer, Berlin, 2014, pp. 183–205.

[HaiLR89] E. Hairer, C. Lubich, and M. Roche. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Springer-Verlag, Berlin, 1989.

[HaiNW93] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Second revised edition. Springer-Verlag, Berlin, 1993.

[HaiW96] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. Second revised edition. Springer-Verlag, Berlin, 1996.

[Hal80] J. K. Hale. *Ordinary Differential Equations*. Second edition. Robert E. Krieger Publishing Company, Inc., Malabar, 1980.

[Han89] M. Hanke. “Linear differential-algebraic equations in spaces of integrable functions”. *J. Differ. Equations* 79 (1): 14–30 (1989).

[Hei14] J. Heiland. “Decoupling and Optimization of Differential-Algebraic Equations with Application in Flow Control”. PhD thesis. Technische Universität Berlin, 2014.

[HeiV95] J. C. Heinrich and C. A. Vionnet. “The penalty method for the Navier-Stokes equations”. *Arch. Comput. Method E.* 2 (2): 51–65 (1995).

[Hen81] D. Henry. *Geometric Theory of Semilinear Parabolic Equations*. Springer, Berlin, 1981.

[HilP57] E. Hille and R. S. Phillips. *Functional Analysis and Semi-Groups*. American Mathematical Society (AMS), Providence, 1957.

[HinPU+09] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer Netherland, Heidelberg, 2009.

[HipHO12] D. A. Hipp, M. Hochbruck, and A. Ostermann. “Exponential integrators for parabolic problems with time dependent coefficients”. *Oberwolfach Rep.* 9: 3602–3606 (2012).

[HipHO14] D. A. Hipp, M. Hochbruck, and A. Ostermann. “An exponential integrator for non-autonomous parabolic problems”. *Electron. Trans. Numer. Anal.* 41: 497–511 (2014).

[HocL97] M. Hochbruck and C. Lubich. “On Krylov subspace approximations to the matrix exponential operator”. *SIAM J. Numer. Anal.* 34 (5): 1911–1925 (1997).

[HocL99] M. Hochbruck and C. Lubich. “A Gautschi-type method for oscillatory second-order differential equations”. *Numer. Math.* 83: 403–426 (1999).

[HocLS98] M. Hochbruck, C. Lubich, and H. Selhofer. “Exponential integrators for large systems of differential equations”. *SIAM J. Sci. Comput.* 19 (5): 1552–1574 (1998).

[HocO05a] M. Hochbruck and A. Ostermann. “Explicit exponential Runge–Kutta methods for semilinear parabolic problems”. *SIAM J. Numer. Anal.* 43 (3): 1069–1090 (2005).

[HocO05b] M. Hochbruck and A. Ostermann. “Exponential Runge–Kutta methods for parabolic problems”. *Appl. Numer. Math.* 53 (2): 323–339 (2005).

[HocO10] M. Hochbruck and A. Ostermann. “Exponential integrators”. *Acta Numer.* 19: 209–286 (2010).

[HocOS09] M. Hochbruck, A. Ostermann, and J. Schweitzer. “Exponential Rosenbrock-type methods”. *SIAM J. Numer. Anal.* 47 (1): 786–803 (2009).

[HunV03] W. Hundsdorfer and J. Verwer. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations.* Springer-Verlag, Berlin, 2003.

[JacZ12] B. Jacob and H. Zwart. *Linear Port-Hamiltonian Systems on Infinite-Dimensional Spaces.* Birkhäuser, Basel, 2012.

[JanT14] L. Jansen and C. Tischendorf. “A unified (P)DAE modeling approach for flow networks.” In: *Progress in Differential-Algebraic Equations. Deskriptor 2013. Proceedings of the 8th Workshop on Coupled Descriptor Systems, Geseke, Germany, March 4–7, 2013.* Springer, Berlin, 2014, pp. 127–151.

[Jay15] L. O. Jay. “Lobatto Methods”. In: *Encyclopedia of Applied and Computational Mathematics.* Ed. by B. Engquist. Springer-Verlag, Berlin, 2015, pp. 817–826.

[Kir45] S. Kirchhoff. “Ueber den Durchgang eines elektrischen Stromes durch eine Ebene, insbesondere durch eine kreisförmige”. *Ann. Phys.* 140: 497–514 (1845).

[KooBG18] G. L. Kooij, M. A. Botchev, and B. J. Geurts. “An exponential time integrator for the incompressible Navier-Stokes equation”. *SIAM J. Sci. Comput.* 40 (3): B684–B705 (2018).

[KovL17] B. Kovács and C. Lubich. “Numerical analysis of parabolic problems with dynamic boundary conditions”. *IMA J. Numer. Anal.* 37 (1): 1–39 (2017).

[KufJF77] A. Kufner, O. John, and S. Fučík. *Function Spaces.* Noordhoff International Publishing, Leyden, and Academia, Prague, 1977.

[KunM06] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations: Analysis and Numerical Solution.* European Mathematical Society (EMS), Zürich, 2006.

[KunM07] P. Kunkel and V. Mehrmann. “Stability properties of differential-algebraic equations and spin-stabilized discretizations”. *Electron. Trans. Numer. Anal.* 26: 385–420 (2007).

[KunMS+06] P. Kunkel, V. Mehrmann, M. Schmidt, I. Seufer, and A. Steinbrecher. *Weak formulations of linear differential-algebraic systems.* Preprint 16-2006. Institut für Mathematik, TU Berlin, 2006.

[LamMT13] R. Lamour, R. März, and C. Tischendorf. *Differential-Algebraic Equations: A Projector Based Analysis.* Springer-Verlag, Heidelberg, 2013.

[Law67] J. Lawson. “Generalized Runge-Kutta processes for stable systems with large Lipschitz constants”. *SIAM J. Numer. Anal.* 4 (3): 372–380 (1967).

[Lin85] P. Linz. *Analytical and Numerical Methods for Volterra Equations.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1985.

[LioM72] J.-L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications. Vol. I.* Springer-Verlag, New York, 1972.

[LuaO14a] V. T. Luan and A. Ostermann. “Explicit exponential Runge–Kutta methods of high order for parabolic problems”. *J. Comput. Appl. Math.* 256: 168–179 (2014).

[LuaO14b] V. T. Luan and A. Ostermann. “Stiff order conditions for exponential Runge–Kutta methods of order five”. In: *Modeling, Simulation and Optimization of Complex Processes - HPSC 2012, Proceedings of the Fifth International Conference on High Performance Scientific Computing, March 5-9, 2012, Hanoi, Vietnam.* Springer, Cham, 2014, pp. 133–143.

[LubO93] C. Lubich and A. Ostermann. “Runge-Kutta methods for parabolic equations and convolution quadrature”. *Math. Comp.* 60 (201): 105–131 (1993).

[LubO95a] C. Lubich and A. Ostermann. “Linearly implicit time discretization of non-linear parabolic equations”. *IMA J. Numer. Anal.* 15 (4): 555–583 (1995).

[LubO95b] C. Lubich and A. Ostermann. “Runge-Kutta approximation of quasi-linear parabolic equations”. *Math. Comp.* 64 (210): 601–627 (1995).

[LucSE99] W. Lucht, K. Strehmel, and C. Eichler-Liebenow. “Indexes and special discretization methods for linear partial differential algebraic equations”. *BIT* 39 (3): 484–512 (1999).

[MagWT+00] P. C. Magnusson, A. Weisshaar, V. K. Tripathi, and G. C. Alexander. *Transmission Lines and Wave Propagation.* Fourth edition. CRC Press, Boca Raton, 2000.

[Mar97] W. Marszalek. “Analysis of Partial Differential Algebraic Equations”. PhD thesis. North Carolina State University, 1997.

[Meh15] V. Mehrmann. “Index concepts for differential-algebraic equations”. In: *Encyclopedia of Applied and Computational Mathematics.* Ed. by B. Engquist. Springer-Verlag, Berlin, 2015, pp. 676–681.

[MehM19] V. Mehrmann and R. Morandin. “Structure-preserving discretization for port-Hamiltonian descriptor systems”. In: *IEEE 58th Conference on Decision and Control (CDC) 2019.* Nice, France, 2019, pp. 6863–6868.

[Mic62] S. G. Michlin. *Variationsmethoden der Mathematischen Physik.* Akademie-Verlag, Berlin, 1962.

[MolV78] C. Moler and C. Van Loan. “Nineteen dubious ways to compute the exponential of a matrix”. *SIAM Rev.* 20 (4): 801–836 (1978).

[MosZ18] A. Moses Badlyan and C. Zimmer. *Operator-GENERIC Formulation of Thermodynamics of Irreversible Processes.* ArXiv e-print 1807.09822. 2018.

[New03] C. K. Newman. “Exponential Integrators for the Incompressible Navier–Stokes Equations”. PhD thesis. Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 2003.

[NieW12] J. Niesen and W. M. Wright. “Algorithm 919: a Krylov subspace algorithm for evaluating the ϕ-functions appearing in exponential integrators”. *ACM Trans. Math. Software* 38 (3): Art. 22 (2012).

[NørW81] S. P. Nørsett and G. Wanner. “Perturbed collocation and Runge-Kutta methods”. *Numer. Math.* 38: 193–208 (1981).

[Olv14] P. J. Olver. *Introduction to Partial Differential Equations.* Springer International Publishing, Basel, 2014.

[OstR92] A. Ostermann and M. Roche. “Runge-Kutta methods for partial differential equations and fractional orders of convergence”. *Math. Comp.* 59 (200): 403–420 (1992).

[Pac98] B. G. Pachpatte. *Inequalities for Differential and Integral Equations.* Academia Press, San Diego, 1998.

[Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.

[Pet82] L. Petzold. “Differential/Algebraic Equations are not ODE’s”. *SIAM J. Sci. and Stat. Comput.* 3 (3): 367–384 (1982).

[PolW98] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioural Approach*. Springer-Verlag, New York, 1998.

[ProR74] A. Prothero and A. Robinson. “On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations”. *Math. Comp.* 28 (125): 145–162 (1974).

[RanA05] J. Rang and L. Angermann. “Perturbation index of linear partial differential-algebraic equations”. *Appl. Numer. Math.* 53 (2): 437–456 (2005).

[Ran93] R. Rannacher. “On the numerical solution of the incompressible Navier-Stokes equations”. *Z. Angew. Math. Mech. (ZAMM)* 73 (9): 203–216 (1993).

[Rei06] T. Reis. “Systems Theoretic Aspects of PDAEs and Applications to Electrical Circuits”. PhD thesis. Technische Universität Kaiserslautern, 2006.

[Rhe84] W. C. Rheinboldt. “Differential-algebraic systems as differential equations on manifolds”. *Math. Comp.* 43: 473–482 (1984).

[Rot30] E. Rothe. “Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben”. *Math. Ann.* 102 (1): 650–670 (1930).

[Rou13] T. Roubíček. *Nonlinear Partial Differential Equations with Applications*. Second edition. Birkhäuser, Basel, 2013.

[Saa92] Y. Saad. “Analysis of some Krylov subspace approximations to the matrix exponential operator”. *SIAM J. Numer. Anal.* 29 (1): 209–228 (1992).

[SchJ14] A. van der Schaft and D. Jeltsema. “Port-Hamiltonian Systems Theory: An Introductory Overview”. *Foundations and Trends® in Systems and Control* 1 (2–3): 173–378 (2014).

[SchM18] A. van der Schaft and B. Maschke. “Generalized port-Hamiltonian DAE systems”. *Systems & Control Letters* 121: 31–37 (2018).

[SchB98] M. Schemann and F. A. Bornemann. “An adaptive Rothe method for the wave equation”. *Comput. Vis. Sci.* 1 (3): 137–144 (1998).

[Sch91] W. E. Schiesser. *The Numerical Method of Lines: Integration of Partial Differential Equations*. Academia Press, San Diego, 1991.

[Sch98] C. Schwab. *p- and hp-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*. Clarendon Press, Oxford, 1998.

[She95] J. Shen. “On error estimates of the penalty method for unsteady Navier-Stokes equations.” *SIAM J. Numer. Anal.* 32 (2): 386–403 (1995).

[Sho10] R. E. Showalter. “Nonlinear degenerate evolution equations in mixed formulation”. *SIAM J. Math. Anal.* 42 (5): 2114–2131 (2010).

[Sim00] B. Simeon. *Numerische Simulation Gekoppelter Systeme von Partiellen und Differential-Algebraischen Gleichungen der Mehrkörperdynamik*. VDI Verlag, Düsseldorf, 2000.

[Sim13] B. Simeon. *Computational Flexible Multibody Dynamics: A Differential-Algebraic Approach*. Differential-Algebraic Equations Forum. Springer-Verlag, Berlin, 2013.

[Sim10] J. Simon. “On the identification $H = H'$ in the Lions theorem and a related inaccuracy”. *Ricerche mat.* 59: 245–255 (2010).

[SprW10] J. Sprekels and H. Wu. “A note on parabolic equation with nonlinear dynamical boundary condition”. *Nonlinear Anal.* 72 (6): 3028–3048 (2010).

[Ste08] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary Elements.* Springer, New York, 2008.

[Str66] W. A. Strauss. “On continuity of functions with values in various Banach spaces”. *Pacific J. Math.* 19: 543–551 (1966).

[StrWP12] K. Strehmel, R. Weiner, and H. Podhaisky. *Numerik Gewöhnlicher Differentialgleichungen: Nichtsteife, Steife und Differential-Algebraische Gleichungen.* Vieweg+Teubner Verlag, Wiesbaden, 2012.

[Tar06] L. Tartar. *An Introduction to Navier-Stokes Equation and Oceanography.* Springer-Verlag, Berlin, 2006.

[Tem77] R. Temam. *Navier-Stokes Equations. Theory and Numerical Analysis.* North-Holland, Amsterdam, 1977.

[Tem82] R. Temam. “Behaviour at time $t = 0$ of the solutions of semi-linear evolution equations”. *J. Differ. Equ.* 43 (1): 73–92 (1982).

[Tho06] V. Thomeé. *Galerkin Finite Element Methods for Parabolic Problems.* Second edition. Springer-Verlag, Berlin, 2006.

[Tis96] C. Tischendorf. “Solution of Index-2 Differential Algebraic Equations and Its Application in Circuit Simulation”. PhD thesis. Humboldt-Universität zu Berlin, 1996.

[Tis03] C. Tischendorf. “Coupled Systems of Differential Algebraic and Partial Differential Equations in Circuit and Device Simulation: Modeling and Numerical Analysis”. Habilitationsschrift. Humboldt-Universität zu Berlin, 2003.

[Trö10] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications.* American Mathematical Society, Providence, Rhode Island, 2010.

[Van50] W. Van Roosbroeck. “Theory of the flow of electrons and holes in germanium and other semiconductors”. *Bell Syst. Tech. J.* 29 (4): 560–607 (1950).

[Ver84] R. Verfürth. “Error estimates for a mixed finite element approximation of the Stokes equations”. *RAIRO Anal. Numér.* 18 (2): 175–182 (1984).

[VouR18] I. Voulis and A. Reusken. “A time dependent Stokes interface problem: well-posedness and space-time finite element discretization”. *ESAIM: M2AN* 52 (6): 2187–2213 (2018).

[VouR19] I. Voulis and A. Reusken. “Discontinuous Galerkin time discretization methods for parabolic problems with linear constraints”. *J. Numer. Math.* 27 (3): 155–182 (2019).

[Wei97] J. Weickert. “Applications of the Theory of Differential-Algebraic Equations to Partial Differential Equations of Fluid Dynamics”. PhD thesis. Technische Universität Chemnitz-Zwickau, Chemnitz, 1997.

[Wie20] J. Wiedemann. “Higher-order exponential integrators for constrained semi-linear parabolic problems”. *GAMMAS* 2 (1): 14–20 (2020).

[Wil19] U. Wilbrandt. *Stokes-Darcy Equations: Analytic and Numerical Analysis.* Birkhäuser/Springer, Cham, 2019.

[Wlo87] J. Wloka. *Partial Differential Equations.* Cambridge University Press, Cambridge, 1987.

[Yos80] K. Yosida. *Functional Analysis.* Sixth edition. Springer-Verlag, Berlin, 1980.

[Zei86] E. Zeidler. *Nonlinear Functional Analysis and Its Applications I: Fixed-Point Theorems.* Springer-Verlag, New York, 1986.

[Zei90a] E. Zeidler. *Nonlinear Functional Analysis and Its Applications II/A: Linear Monotone Operators*. Springer-Verlag, New York, 1990.

[Zei90b] E. Zeidler. *Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators*. Springer-Verlag, New York, 1990.

[Zim15] C. Zimmer. "Theorie und Numerik von Linearen Operator-Differentiell-Algebraischen Gleichungen mit Zeitverzögertem Term". Master's thesis. Technische Universität Berlin, 2015.

[Zim20] C. Zimmer. *Simulation code for the thesis "Temporal Discretization of Constrained Partial Differential Equations"*. Zenodo. 2020. URL: `https://doi.org/10.5281/zenodo.4321809`.

Index