Karlsruhe Institute of Technology

wave phenomena
analysis and numerics

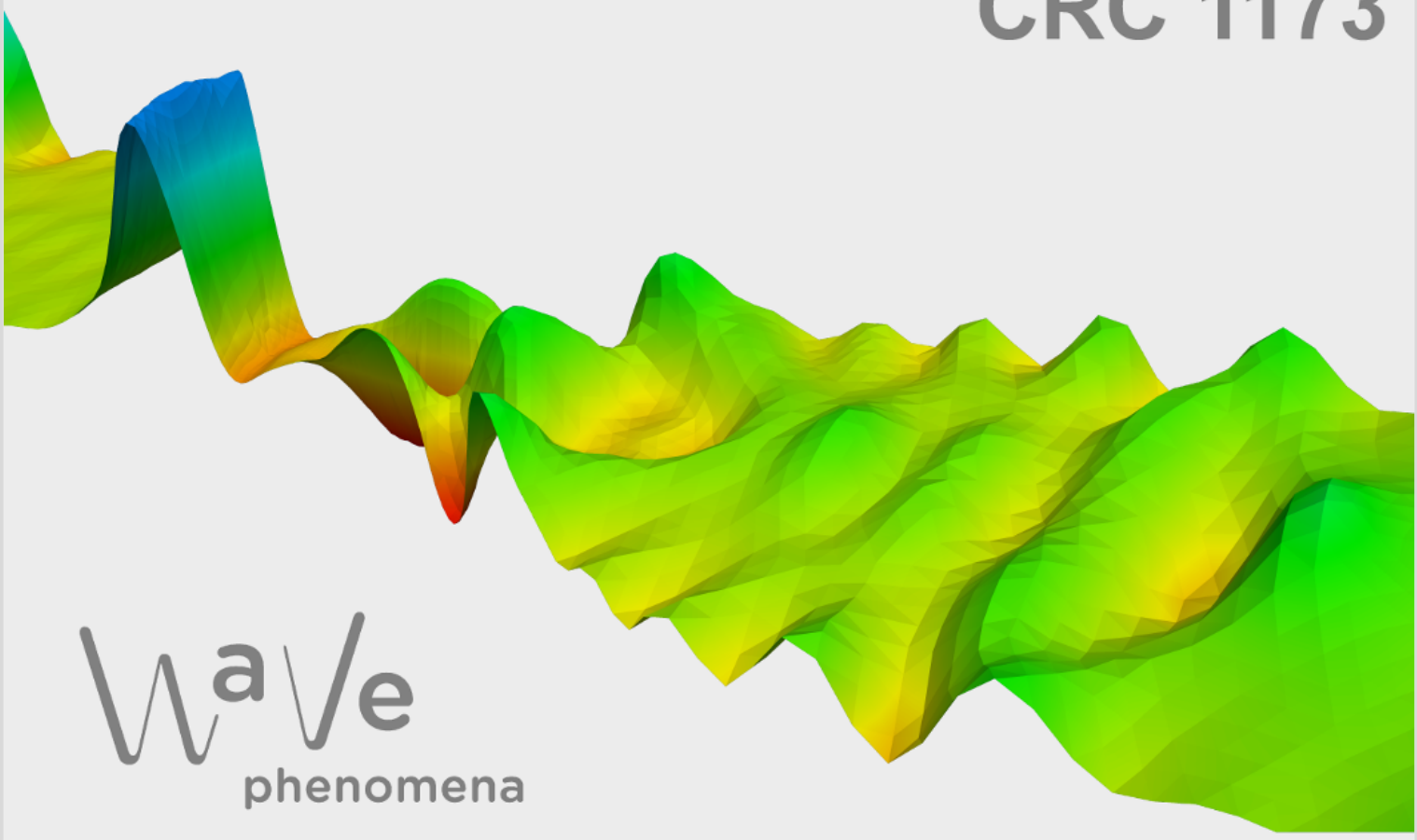
Gautschi-type and implicit-explicit integrators for constrained wave equations

Robert Altmann, Benjamin Dörich, Christoph Zimmer

CRC Preprint 2025/21, May 2025

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities



Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN





Funded by



ISSN 2365-662X

GAUTSCHI-TYPE AND IMPLICIT–EXPLICIT INTEGRATORS FOR CONSTRAINED WAVE EQUATIONS

R. ALTMANN[†], B. DÖRICH[*], C. ZIMMER[‡]

Abstract. This paper deals with the construction and analysis of two integrators for (semi-linear) second-order partial differential–algebraic equations of semi-explicit type. More precisely, we consider an implicit–explicit Crank–Nicolson scheme as well as an exponential integrator of Gautschi type. For this, well-known wave integrators for unconstrained systems are combined with techniques known from the field of differential–algebraic equations. The result are efficient time stepping schemes, which are provable of second order. Moreover, we discuss the practical implementation of the Gautschi-type method, which involves the solution of certain saddle point problems. The theoretical results are verified by numerical experiments for the the wave equation with kinetic boundary conditions.

Key words. PDAE, Gautschi-type integrator, implicit–explicit Crank–Nicolson scheme, time discretization

AMS subject classifications. 65L80, **65M12**, **65J15**

1. Introduction

Partial differential equations of second order are widely used to model wave-type phenomena. They often appear in a semi-linear manner with an unbounded linear part and a moderate nonlinearity. In this paper, we consider wave-type equations

$$\ddot{u}(t) + \mathcal{D}\dot{u}(t) + \mathcal{A}u(t) = f(t, u) \tag{1.1}$$

with initial conditions $u(0) = u^0$, $\dot{u}(0) = w^0$ which underlie an additional (linear) constraint of the form $\mathcal{B}u(t) = g(t)$. Such systems can be considered as differential–algebraic equations (DAEs) in Banach spaces, also called *partial differential–algebraic equations* (PDAEs); see [LMT13, Alt15] for an introduction. Second-order PDAEs of wave type appear, e.g., in the field of elastic multibody dynamics [Sim98, Sim13], in problems on wave propagation with non-standard boundary conditions [Alt23], or, more generally, if two second-order systems are coupled.

For the time discretization of (1.1) (without a constraint) the two main classes of methods are given by explicit and implicit schemes. The advantage of an explicit time-stepping lies in the easy implementation and possibility for parallelization. However, when combined with a spatial discretization small time steps have to be chosen. On the other hand, implicit methods are unconditionally stable and thus allow for large time steps. Here, however, a nonlinear system of equations has to be solved in each time step. Thus, it is desirable to use a scheme that takes the best from both methods and obtain a scheme which is unconditionally stable in the linear part, and treats the nonlinear term explicitly. In this work, we consider two types of methods, namely an implicit–explicit Crank–Nicolson scheme (IMEX) as well as an exponential integrator of Gautschi type.

Date: May 30, 2025.

The idea of IMEX schemes is to split the equation in a linear unbounded part that is treated implicitly and a (possibly nonlinear) part which is treated explicitly, under the condition that the scheme is stable under a step-size restriction which only depends on the Lipschitz condition of the explicit part and preserves the order of convergence of the underlying schemes. For IMEX schemes, there is a well-developed theory which covers IMEX multistep schemes [ARW95, FHV97, ACM99, HR07] as well as IMEX Runge–Kutta methods [ARS97, Bos07]. More recently, an IMEX scheme was proposed which exploits the structure of the second-order system by a combination of the leapfrog method with the Crank–Nicolson method [HL21]. This scheme will be the basis of our IMEX method for the constrained problem. Further, there is the so-called Crank–Nicolson–leapfrog scheme proposed in [LT12, LLT16], which uses the leapfrog method for the first-order system, and is stable only when the explicitly treated part is skew-symmetric.

In the undamped case $\mathcal{D} = 0$, also Gautschi-type integrators provide a powerful tool for the time integration of wave-type equations. These methods follow the methodology of exponential integrators in the parabolic setting [HO10]. Hence, the idea is to integrate the linear (stiff) part in an exact manner, which allows large time steps that are not restricted by the frequencies of the differential operator. For second-order wave-type systems (without a constraint), Gautschi-type integrators were developed in [Gau61, HL99, Gri02]. For this, one considers the first-order formulation and applies the variation-of-constants formula. The combination of the resulting formula for times $t+\tau$ and $t-\tau$ and an approximation of the integral term by the simplest left/right rectangle rule finally yields an efficient two-step time integration scheme. In the same way as exponential integrators rely on an efficient computation of the matrix exponential applied to a vector [MV78], Gautschi-type integrators are based on the cosine of a matrix (applied to a vector). Also this can be realized efficiently [HAM10], e.g., by using Krylov subspace methods [EE06, CKN24].

In this paper, we aim to translate these two concepts to wave-type equations (1.1) under a linear constraint. For the construction of our novel PDAE integrators, we decompose the solution as in [EM13]. This means that we consider the part of the solution, which lies in the kernel of the constraint operator, separately. For the IMEX scheme, we use the half-step formulation from [HL21] and impose the linear constraints only on the time derivative of the solution such that the solution u satisfies the constraint exactly. This leads to a method which requires (after the spatial discretization) the solution of two (linear) saddle point problems in each time step, one including the mass and one the stiffness matrix.

For the Gautschi-type integrators, we again consider only the part of the solution in the kernel, which together with techniques developed for constrained parabolic systems in [AZ20], then leads to the desired explicit integration scheme. To prove (mesh-independent) convergence of this novel integrator, the theory of [HL99, Gri02] for the unconstrained case is combined with the already mentioned PDAE techniques. In view of the efficiency of the practical implementation, the critical part is the computation of $\cos(\tau\Omega_{\ker})v$ for some vector v. Here, $\Omega_{\ker}$ denotes the root of the differential operator $\mathcal{A}$ (or its discrete counterpart) restricted to the kernel of the constraint operator $\mathcal{B}$. For this, we propose an Arnoldi method, which does not require the computation of $\Omega_{\ker}$; cf. [Gri05] for the unconstrained case. Moreover, by particularly constructed saddle point problems, the computations can be restricted to computations on the kernel of the constraint operator. At this point, we would like to emphasize that Gautschi-type integrators are tailored to problems with vanishing damping, i.e., $\mathcal{D} = 0$. In order to include damping in an

exponential manner, one could also use general exponential methods of Runge–Kutta or multistep type; cf. the review [HO10]. This, however, is beyond the scope of this work.

The paper is organized as follows. In Section 2 we introduce the system equations in detail and discuss its unique solvability. Afterwards, we turn to the numerical treatment of theses systems. First, an implicit–explicit version of the Crank–Nicolson scheme is introduced and analyzed in Section 3. Here, a detailed inspection of all the error contributions allows us to prove an error equation which can be resolved by the techniques developed for the unconstrained IMEX scheme. It follows the discussion of Gautschi-type integrators for damping-free systems in Section 4. Besides proving second-order accuracy of the method, we also provide comments on the practical implementation of the algorithm. Finally, we present a numerical example in Section 5 and conclude in Section 6.

2. Wave-type Systems with Linear Constraints

This section is devoted to the introduction of semi-linear PDAEs of wave type. Hence, we consider systems of the form (1.1) which underlie an additional (linear) constraint.

2.1. **Problem formulation.**

In this subsection, we introduce a class of hyperbolic wave-type systems, which are restricted by a linear constraint. Allowing (sufficiently smooth) nonlinearities in the low-order terms of the dynamic equation, we end up with the following PDAE: find abstract functions $u\colon [0,T] \to \mathcal{V}$ and $\lambda\colon [0,T] \to \mathcal{Q}$ such that

$$\ddot{u}(t) + \mathcal{D}\dot{u}(t) + \mathcal{A}u(t) + \mathcal{B}^*\lambda(t) = f(t,u) \qquad \text{in } \mathcal{V}^*, \tag{2.1a}$$

$$\mathcal{B}u(t) = g(t) \qquad \text{in } \mathcal{Q}^*. \tag{2.1b}$$

Therein, the spaces $\mathcal{V}$ and $\mathcal{Q}$ are assumed to be Hilbert spaces with respective duals $\mathcal{V}^*$ and $\mathcal{Q}^*$. Further, the space $\mathcal{V}$ is part of a Gelfand triple $\mathcal{V}$, $\mathcal{H}$, $\mathcal{V}^*$, i.e., we assume a continuous (and dense) embedding $\mathcal{V} \hookrightarrow \mathcal{H}$, which in turn implies a corresponding embedding of the duals [Zei90, Ch. 23.4].

For second-order systems in the weak form, one typically seeks a solution with the first and second derivative taking values in $\mathcal{H}$ and $\mathcal{V}^*$, respectively. Hence, initial data is expected to be of the form $u(0) = u^0 \in \mathcal{V}$, $\dot{u}(0) = w^0 \in \mathcal{H}$. Note, however, that the constraint (2.1b) typically implies certain consistency conditions on the initial data; cf. [Alt15, Rem. 7.4]. At this point, it is reasonable to assume $u^0, w^0 \in \mathcal{V}$ with $\mathcal{B}u^0 = g(0)$ and $\mathcal{B}w^0 = \dot{g}(0)$.

For the linear operators $\mathcal{A} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$, $\mathcal{B} \in \mathcal{L}(\mathcal{V}, \mathcal{Q}^*)$, and $\mathcal{D} \in \mathcal{L}(\mathcal{H}, \mathcal{V}^*)$ appearing in (2.1), we make the following assumptions.

Assumption 2.1 (Constraint operator $\mathcal{B}$). The operator $\mathcal{B}\colon \mathcal{V} \to \mathcal{Q}^*$ is linear, continuous, and satisfies an inf-sup condition, i.e., there exists a constant $\beta > 0$ such that

$$\inf_{q \in \mathcal{Q}\setminus\{0\}} \sup_{v \in \mathcal{V}\setminus\{0\}} \frac{\langle \mathcal{B}v, q\rangle}{\|v\|_{\mathcal{V}}\|q\|_{\mathcal{Q}}} \geq \beta.$$

Assumption 2.2 (Differential operator $\mathcal{A}$). The operator $\mathcal{A}\colon \mathcal{V} \to \mathcal{V}^*$ is linear, continuous, and elliptic on $\mathcal{V}_{\ker} \coloneqq \ker \mathcal{B}$, i.e., on the kernel of the constraint operator. Furthermore, we assume that we can decompose the operator into $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$, with $\mathcal{A}_1$ being self-adjoint and $\mathcal{A}_2 \in \mathcal{L}(\mathcal{V}, \mathcal{H}^*)$. In particular, we may assume that $\mathcal{A}_1$ is elliptic on $\mathcal{V}_{\ker}$ as well; cf. [AZ20].

Assumption 2.3 (Damping operator $\mathcal{D}$). The operator $\mathcal{D}\colon \mathcal{H} \to \mathcal{V}^*$ is linear and continuous. Moreover, its restriction to functions in $\mathcal{V}$ satisfies $\mathcal{D}\big|_{\mathcal{V}}\colon \mathcal{V} \to \mathcal{H}^*$, is again continuous, and satisfies the monotonicity condition

$$(\mathcal{D}u, u)_{\mathcal{H}} \geq 0 \quad \text{for all} \quad u \in \mathcal{V}. \tag{2.2}$$

Remark 2.4. The previous assumptions can be generalized to the case where $\mathcal{A}$ only satisfies a Gårding inequality, see [AZ20], or $\mathcal{D}$ satisfies a lower bound of the type $(\mathcal{D}u, u)_{\mathcal{H}} \geq -c_{\mathcal{D}}\|u\|_{\mathcal{H}}^2$.

As discussed in [AZ20], Assumption 2.1 implies the existence of a continuous right-inverse $\mathcal{B}^-\colon \mathcal{Q}^* \to \mathcal{V}$, which leads to the decomposition

$$\mathcal{V} = \mathcal{V}_{\ker} \oplus \mathcal{V}_{\mathrm{c}} \qquad \text{with} \qquad \mathcal{V}_{\ker} = \ker \mathcal{B}, \quad \mathcal{V}_{\mathrm{c}} = \operatorname{im} \mathcal{B}^-. \tag{2.3}$$

Following [AZ18], we define the complementary space $\mathcal{V}_{\mathrm{c}}$ as

$$\mathcal{V}_{\mathrm{c}} \coloneqq \big\{ v \in \mathcal{V} \,|\, \mathcal{A}v \in \mathcal{V}_{\ker}^0 \big\} = \big\{ v \in \mathcal{V} \,|\, \langle \mathcal{A}v, w\rangle = 0 \text{ for all } w \in \mathcal{V}_{\ker} \big\}.$$

We also introduce the restriction of the differential operator to the kernel of $\mathcal{B}$, namely

$$\mathcal{A}_{\ker} \coloneqq \mathcal{A}|_{\mathcal{V}_{\ker}}\colon \mathcal{V}_{\ker} \to \mathcal{V}_{\ker}^* \coloneqq (\mathcal{V}_{\ker})^*.$$

Here, we use the fact that functionals in $\mathcal{V}^*$ define functionals in $\mathcal{V}_{\ker}^*$ simply through the restriction to $\mathcal{V}_{\ker}$. Note that $\mathcal{A}_{\ker}$ is elliptic due to Assumption 2.2. Hence, $-\mathcal{A}_{\ker}$ generates an analytic semigroup on the closure of $\mathcal{V}_{\ker}$ in the $\mathcal{H}$-norm, i.e., on $\mathcal{H}_{\ker} \coloneqq \operatorname{clos}_{\mathcal{H}} \mathcal{V}_{\ker}$; cf. [Paz83, Ch. 7, Th. 2.7].

Similarly, we introduce the restriction

$$\mathcal{D}_{\ker} \coloneqq \mathcal{D}|_{\mathcal{H}_{\ker}}\colon \mathcal{H}_{\ker} \to \mathcal{V}_{\ker}^*,$$

which satisfies $\mathcal{D}_{\ker}\big|_{\mathcal{V}_{\ker}}\colon \mathcal{V}_{\ker} \to \mathcal{H}_{\ker}^*$ according to Assumption 2.3.

Remark 2.5. With the space $\mathcal{H}_{\ker}$ at hand, the assumptions on the initial data may be relaxed to $w^0 \in \mathcal{H}_{\ker} + \mathcal{B}^-\dot{g}(0)$, which is still consistent with the constraint.

Finally, we summarize the assumptions on the right-hand sides. Within the constraint, we assume a mapping $g\colon [0,T] \to \mathcal{Q}^*$. Moreover, the nonlinearity satisfies $f\colon [0,T] \times \mathcal{H} \to \mathcal{H}^*$. In addition, we require the following Lipschitz bounds on f, depending on the method we apply.

Assumption 2.6 (Nonlinearity f). One of the following two conditions is satisfied:
(a) There is a constant $L_{\mathcal{H}}$ such that for all $v_1, v_2 \in \mathcal{H}$ and $t \geq 0$ it holds

$$\|f(t, v_1) - f(t, v_2)\|_{\mathcal{H}^*} \leq L_{\mathcal{H}} \, \|v_1 - v_2\|_{\mathcal{H}}.$$

(b) There is a constant $L_{\mathcal{V}}(\rho)$ such that for all $v_1, v_2 \in \mathcal{V}$ with $\|v_i\|_{\mathcal{V}} \leq \rho$ and $t \geq 0$ it holds

$$\|f(t, v_1) - f(t, v_2)\|_{\mathcal{H}^*} \leq L_{\mathcal{V}}(\rho) \, \|v_1 - v_2\|_{\mathcal{V}}.$$

Remark 2.7. Within Assumption 2.6, condition (a) implies condition (b).

Example 2.8. Consider the wave equation with additional convection and damping terms,

$$\ddot{u} + (\alpha + \beta \cdot \nabla)\, \dot{u} - \Delta u + \bar{v} \cdot \nabla u = f(u) \qquad \text{in } \Omega \subseteq \mathbb{R}^3$$

with $\alpha, \bar{v} \in L^\infty(\Omega)$ and $\beta \in H(\operatorname{div}, \Omega) \cap L^\infty(\Omega)$ satisfying $\alpha - \frac{1}{2}\operatorname{div}\beta \geq 0$ almost everywhere, and homogeneous Dirichlet boundary conditions. We set $\mathcal{V} = H_0^1(\Omega)$ and $\mathcal{H} = L^2(\Omega)$. Then the operator $\mathcal{A}$ is given by

$$\langle \mathcal{A}u, w\rangle = \int_\Omega \nabla u \cdot \nabla w + \bar{v} \cdot \nabla u \, w \,\mathrm{d}x =: \langle \mathcal{A}_1 u, w\rangle + \langle \mathcal{A}_2 u, w\rangle.$$

Here, $\mathcal{A}_1$ corresponds to the Laplacian, which is elliptic and self-adjoint, whereas $\mathcal{A}_2$ models the convection term. Note that $\mathcal{A}_2$ is an element of $\mathcal{L}(\mathcal{V}, \mathcal{H}^*)$, since we do not need to differentiate the test function. For the damping operator, we consider $\mathcal{D}\big|_{\mathcal{V}} = \alpha + \beta \cdot \nabla$. Integration by parts gives

$$(\mathcal{D}u, u)_{\mathcal{H}} = \int_\Omega (\alpha - \tfrac{1}{2}\operatorname{div}\beta)\,|u|^2 \,\mathrm{d}x \geq 0. \tag{2.4}$$

Similarly, integration by parts gives that for $u \in \mathcal{H}$ and $v \in \mathcal{V}$ it holds that

$$\langle \mathcal{D}u, v\rangle_{\mathcal{V}^*,\mathcal{V}} = \int_\Omega (\alpha - \operatorname{div}\beta)\, uv - u\beta \cdot \nabla v \,\mathrm{d}x. \tag{2.5}$$

For $f(z) = \sin(z)$, Assumption 2.6 (a) is satisfied. For $f(z) = |z|^{p-1}z$, $1 \leq p \leq 3$, one can verify Assumption 2.6 (b) by Sobolev's embedding.

2.2. **Existence of solutions.** Before introducing numerical schemes for constrained systems of wave type, we need to discuss the existence of solutions to (2.1). Throughout this section, we denote the well-known Sobolev-Bochner spaces for a Banach space X by $L^2(0,T;X)$ and $H^1(0,T;X)$; cf. [Zei90, Ch. 23]. Moreover, $C([0,T],X)$ denotes the space of continuous functions in $[0,T]$ with values in X. As a first step, we give a classical existence result for the linear case without constraints. Here, linear means that the right-hand side f is independent of u.

Lemma 2.9 (Existence of solutions, linear and unconstrained case)**.** *Let Assumption 2.3 hold and consider $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ with $\mathcal{A}_1 \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ being elliptic with constant μ_1 and self-adjoint and $\mathcal{A}_2 \in \mathcal{L}(\mathcal{V}, \mathcal{H}^*)$ with continuity constant C_2. Further, suppose that $u^0 \in \mathcal{V}$, $w^0 \in \mathcal{H}$, and $f \in L^2(0,T;\mathcal{H}^*)$ are given. Then the linear equation*

$$\ddot{u}(t) + \mathcal{D}\dot{u}(t) + \mathcal{A}u(t) = f(t) \qquad \text{in } \mathcal{V}^*$$

with initial conditions $u(0) = u^0$ and $\dot{u}(0) = w^0$ has a unique solution

$$u \in C([0,T],\mathcal{V}), \qquad \dot{u} \in C([0,T],\mathcal{H}), \qquad \ddot{u} \in L^2(0,T;\mathcal{V}^*).$$

This solution satisfies the estimate

$$\|\dot{u}(t)\|_{\mathcal{H}}^2 + \|u(t)\|_{\mathcal{A}_1}^2 \leq e^{\kappa t}\left(\sqrt{\|w^0\|_{\mathcal{H}}^2 + \|u^0\|_{\mathcal{A}_1}^2} + \int_0^t e^{-\kappa s}\,\|f(s)\|_{\mathcal{H}^*}\,\mathrm{d}s\right)^2 \tag{2.6}$$

with exponent $\kappa := C_2/\sqrt{\mu_1}$ and the problem-dependent norm $\|\bullet\|_{\mathcal{A}_1}^2 := \langle \mathcal{A}_1 \bullet, \bullet\rangle$.

Proof. By an adaption of [Zei90, Cor. 24.3], the problem has a unique solution u within the given spaces. It remains to show the stability estimate (2.6). For this, note that u satisfies the system

$$\ddot{u}(t) + \mathcal{D}\dot{u}(t) + \mathcal{A}_1 u(t) = \tilde{f}(t) := f(t) - \mathcal{A}_2 u(t)$$

with the same initial values and right-hand side $\tilde{f} = f - \mathcal{A}_2 u \in L^2(0,T;\mathcal{H}^*)$. Along the same lines as in [LM72, Ch. 3, Lem 8.3] and Young's inequality, we conclude that

$$\begin{aligned}\|\dot{u}(t)\|_{\mathcal{H}}^2 + \|u(t)\|_{\mathcal{A}_1}^2 &\leq \|w^0\|_{\mathcal{H}}^2 + \|u^0\|_{\mathcal{A}_1}^2 + 2\int_0^t \langle f - \mathcal{A}_2 u, \dot{u}\rangle \,\mathrm{d}s \\ &\leq \|w^0\|_{\mathcal{H}}^2 + \|u^0\|_{\mathcal{A}_1}^2 + \kappa \int_0^t \|\dot{u}\|_{\mathcal{H}}^2 + \|u\|_{\mathcal{A}_1}^2 \,\mathrm{d}s + 2\int_0^t \|f\|_{\mathcal{H}^*}\|\dot{u}\|_{\mathcal{H}} \,\mathrm{d}s.\end{aligned}$$

Estimate (2.6) then follows along the lines of [Zim21, Th. 4.22]. By the linearity of the problem, this estimate also proves the uniqueness of the solution. □

We return to the constrained case, i.e., to the PDAE (2.1).

Lemma 2.10 (Existence of solutions, linear PDAE case)**.** *Consider Assumptions 2.1, 2.2, and 2.3 as well as right-hand sides $f \in L^2(0,T;\mathcal{H}^*)$ and $g \in H^2(0,T;\mathcal{Q}^*)$. Further assume consistent initial data, namely*

$$u^0 \in \mathcal{B}^- g(0) + \mathcal{V}_{\mathrm{ker}}, \qquad w^0 \in \mathcal{B}^- \dot{g}(0) + \mathcal{H}_{\mathrm{ker}}.$$

Then there exists a unique solution to (2.1) *with $u(0) = u^0$, $\dot{u}(0) = w^0$, satisfying*

$$u \in C([0,T],\mathcal{V}), \qquad \dot{u} \in C([0,T],\mathcal{H}), \qquad \ddot{u} \in L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*),$$

as well as a stability estimate of the form

$$\|\dot{u}(t)\|_{\mathcal{H}} + \|u(t)\|_{\mathcal{V}} \lesssim \|w^0\|_{\mathcal{H}} + \|u^0\|_{\mathcal{V}} + \int_0^t \|f(s)\|_{\mathcal{H}^*} \,\mathrm{d}s + \|g\|_{H^2(0,t;\mathcal{Q}^*)}. \tag{2.7}$$

Proof. Considering the decomposition $u = u_{\mathrm{ker}} + u_{\mathrm{c}}$ given by $\mathcal{V} = \mathcal{V}_{\mathrm{ker}} \oplus \mathcal{V}_{\mathrm{c}}$, we directly obtain $u_{\mathrm{c}}(t) = \mathcal{B}^- g(t) \in \mathcal{V}_{\mathrm{c}}$. For u_{ker}, we restrict (2.1a) to test functions in the kernel space $\mathcal{V}_{\mathrm{ker}}$. Due to the definition of $\mathcal{V}_{\mathrm{ker}}$, the Lagrange multiplier vanishes for such test functions and we get

$$\ddot{u}_{\mathrm{ker}} + \mathcal{D}\dot{u}_{\mathrm{ker}} + \mathcal{A}_{\mathrm{ker}} u_{\mathrm{ker}} = \tilde{f}(t) \coloneqq f(t) - \mathcal{D}\dot{u}_{\mathrm{c}} - \ddot{u}_{\mathrm{c}} = f(t) - \mathcal{D}\mathcal{B}^- \dot{g}(t) - \mathcal{B}^- \ddot{g}(t) \quad \text{in } \mathcal{V}_{\mathrm{ker}}^*.$$

The right-hand side satisfies $\tilde{f} \in L^2(0,T;\mathcal{H}^*)$. Moreover, the corresponding initial data is given by $u_{\mathrm{ker}}(0) = u^0 - \mathcal{B}^- g(0) \in \mathcal{V}_{\mathrm{ker}}$ and $\dot{u}_{\mathrm{ker}}(0) = w^0 - \mathcal{B}^- \dot{g}(0) \in \mathcal{H}_{\mathrm{ker}}$. Since $\mathcal{V}_{\mathrm{ker}}$, $\mathcal{H}_{\mathrm{ker}}$, $\mathcal{V}_{\mathrm{ker}}^*$ defines a Gelfand triple, we can apply Lemma 2.9, leading to a unique solution

$$u_{\mathrm{ker}} \in C([0,T],\mathcal{V}_{\mathrm{ker}}), \qquad \dot{u}_{\mathrm{ker}} \in C([0,T],\mathcal{H}_{\mathrm{ker}}), \qquad \ddot{u}_{\mathrm{ker}} \in L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*)$$

and the stated stability estimate. □

Finally, we transfer the existence results to the semi-linear case.

Theorem 2.11 (Existence of solutions)**.** *Consider Assumptions 2.1, 2.2, 2.3, and 2.6 as well as a right-hand side $g \in H^2(0,T;\mathcal{Q}^*)$. Further, assume consistent initial data*

$$u^0 \in \mathcal{B}^- g(0) + \mathcal{V}_{\mathrm{ker}}, \qquad w^0 \in \mathcal{B}^- \dot{g}(0) + \mathcal{H}_{\mathrm{ker}}.$$

Then there exists a $T > 0$ and a unique solution to (2.1) *on $[0,T]$ with $u(0) = u^0$, $\dot{u}(0) = w^0$ satisfying*

$$u \in C([0,T],\mathcal{V}), \qquad \dot{u} \in C([0,T],\mathcal{H}), \qquad \ddot{u} \in L^2(0,T;\mathcal{V}_{\mathrm{ker}}^*).$$

Moreover, we have the stability estimate that $\|\dot{u}(t)\|_{\mathcal{H}}^2 + \|u(t)\|_{\mathcal{V}}^2$ stays bounded.

Proof. Again, the decomposition $u = u_{\ker} + u_{\mathrm{c}}$ yields that we only have to study the part in the kernel and we proceed as in Lemma 2.10. We employ estimate (2.7) to set up a fixed-point argument with the nonlinearity

$$\widehat{f}(\,\cdot\,, u_{\ker}) \coloneqq f(\,\cdot\,, u_{\ker} + \mathcal{B}^{-} g) - \mathcal{D}\mathcal{B}^{-}\dot{g}(t) - \mathcal{B}^{-}\ddot{g}(t),$$

which satisfies for $v_1, v_2 \in \mathcal{V}_{\ker}$ the Lipschitz condition

$$\big\|\widehat{f}(t, v_1) - \widehat{f}(t, v_2)\big\|_{\mathcal{H}} = \big\|f(t, v_1 + \mathcal{B}^{-} g(t)) - f(t, v_2 + \mathcal{B}^{-} g(t))\big\|_{\mathcal{H}} \leq L_{\mathcal{V}} \, \|v_1 - v_2\|_{\mathcal{V}}.$$

This then yields the claimed well-posedness. □

2.3. **Spatial discretization.** In this paper, we focus on the temporal discretization and only shortly discuss the spatial discretization of the PDAE (2.1) using a (conforming) finite element approximation. Such a discretization yields a symmetric, positive definite mass matrix $M \in \mathbb{R}^{n,n}$, a stiffness matrix $A \in \mathbb{R}^{n,n}$ corresponding to the operator $\mathcal{A}$, a damping matrix $D \in \mathbb{R}^{n,n}$ corresponding to the operator $\mathcal{D}$, and a constraint matrix $B \in \mathbb{R}^{m,n}$ with $m \leq n$, which we assume to be of full (row) rank. According to the assumptions on $\mathcal{A}$ from Assumption 2.2, we further assume that the stiffness matrix A is positive on the kernel of B. The resulting semi-discrete system then reads

$$M\ddot{x}(t) + D\dot{x}(t) + Ax(t) + B^{T}\lambda(t) = f(t, x(t)), \tag{2.8a}$$

$$Bx(t) = g(t), \tag{2.8b}$$

where the right-hand sides are now functions $f\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $g\colon [0,T] \to \mathbb{R}^m$. Moreover, we assume consistent initial values, i.e., $x(0) = x^0$, $\dot{x}(0) = y^0$ with $Bx^0 = g(0)$ and $By^0 = \dot{g}(0)$.

Remark 2.12. Due to the semi-explicit structure of semi-discrete system, one can easily see that the stated assumptions imply that the DAE (2.8) is of index 3; cf. [HW96, Ch. VII.1].

3. An implicit–explicit Crank–Nicolson Scheme

This section is devoted to the extension of the IMEX scheme introduced in [HL21] to constrained systems. For this, we consider a uniform partition of the time interval $[0,T]$ into $0 = t_0 < t_1 < \cdots < t_N = T$ with constant time step size τ, i.e., $t_n = n\,\tau$. Furthermore, we denote by u^n the computed approximation of u at time t_n and abbreviate function evaluations by $g^n \coloneqq g(t_n)$. Throughout, we use the notation

$$a \lesssim b,$$

if there is a constant $C > 0$ independent of the step size τ such that $a \leq Cb$.

3.1. **Review of the unconstrained case.** We first recall the situation for the unconstrained case. For this, we consider a system of the form

$$\ddot{u}(t) + \mathcal{D}\dot{u}(t) + \mathcal{A}u(t) = f(t, u)$$

with initial conditions $u(0) = u^0$, $\dot{u}(0) = w^0$. A Crank–Nicolson discretization, where the nonlinear term is treated by the left and right rectangle rule (see [HL21] for more details) yields the IMEX scheme

$$\big(\mathrm{id} + \tfrac{\tau}{2}\mathcal{D} + \tfrac{\tau^2}{4}\mathcal{A}\big)\, w^{n+1/2} = w^n - \tfrac{\tau}{2}\mathcal{A}u^n + \tfrac{\tau}{2} f^n, \tag{3.1a}$$

$$u^{n+1} = u^n + \tau w^{n+1/2}, \tag{3.1b}$$

$$w^{n+1} = -w^n + 2w^{n+1/2} + \tfrac{\tau}{2}\big(f^{n+1} - f^n\big), \tag{3.1c}$$

where $f^n = f(t_n, u^n)$. We would like to emphasize that this scheme is *explicit* in the nonlinearity such that no nonlinear solver is needed. Moreover, it has been shown to remain second-order accuracy.

Theorem 3.1 (Second-order convergence, unconstrained case [HL21, Th. 2.9])**.** *Consider initial data u^0 in the domain of $\mathcal{A}$ and $w^0 \in \mathcal{V}$ as well as the unique solution u satisfying*

$$u \in C^3([0,T],\mathcal{V}) \cap C^4([0,T],\mathcal{H}), \qquad f(\,\cdot\,,u(\cdot)) \in C^2([0,T],\mathcal{H}).$$

Let Assumption 2.6 (b) hold. Then the discrete solution obtained by the IMEX scheme (3.1) *satisfies the convergence result*

$$\|u(t_n) - u^n\|_{\mathcal{V}} + \|\dot{u}(t_n) - w^n\|_{\mathcal{H}} \lesssim \tau^2.$$

Let us already mention that in the case of $\mathcal{D} = 0$, one can derive an equivalent two-step formulation, which eliminates w^n, see Lemma 3.3 below.

3.2. **An implicit–explicit scheme of second order.** We now turn to the PDAE system (2.1) with consistent initial data. Following the approach of the previous subsection, we propose to compute in each time step the following three substeps. First, compute the pair $(w^{n+1/2}, \lambda^{n+1}) \in \mathcal{V} \times \mathcal{Q}$ via

$$\big(\mathrm{id} + \tfrac{\tau}{2}\mathcal{D} + \tfrac{\tau^2}{4}\mathcal{A}\big)\, w^{n+1/2} + \tfrac{\tau}{2}\mathcal{B}^*\lambda^{n+1} = w^n - \tfrac{\tau}{2}\mathcal{A}u^n + \tfrac{\tau}{2}f^n, \tag{3.2a}$$

$$\mathcal{B}w^{n+1/2} = \tfrac{1}{\tau}\big(g^{n+1} - g^n\big). \tag{3.2b}$$

Second, we obtain $u^{n+1} \in \mathcal{V}$ simply by

$$u^{n+1} = u^n + \tau w^{n+1/2}. \tag{3.2c}$$

Finally, we consider the system

$$w^{n+1} + \mathcal{B}^*\mu^{n+1} = 2w^{n+1/2} - w^n + \tfrac{\tau}{2}\big(f^{n+1} - f^n\big), \tag{3.2d}$$

$$\mathcal{B}w^{n+1} = \tfrac{1}{\tau}\big(g^{n+3/2} - g^{n+1/2}\big). \tag{3.2e}$$

Here, the existence of a unique solution $w^{n+1} \in \mathcal{H}_{\ker} + \mathcal{V}_{\mathrm{c}}$ is only guaranteed if $\mathcal{H}_{\ker}$ is a proper subspace of $\mathcal{H}$; cf. [Zim21, Lem. 9.22]. In this case, the expression $\mathcal{B}w^{n+1}$ is indeed well-defined, even though $w^{n+1} \notin \mathcal{V}$. In the following, we assume the unique solvability of (3.2d)-(3.2e) such that the IMEX scheme (3.2) is well-defined overall.

Remark 3.2. One can easily see by mathematical induction that u^{n+1} is consistent, i.e.,

$$\mathcal{B}u^{n+1} = \mathcal{B}u^n + \tau\mathcal{B}w^{n+1/2} = \mathcal{B}u^n + g^{n+1} - g^n = g^{n+1},$$

as long as u^0 is consistent.

Lemma 3.3. *In the damping-free case, i.e., $\mathcal{D} = 0$, scheme* (3.2) *is equivalent to the two-step formulation*

$$u^{n+1} - 2u^n + u^{n-1} + \mathcal{B}^*\widetilde{\lambda}^{n+1} = -\tfrac{\tau^2}{4}\mathcal{A}\big(u^{n+1} + 2u^n + u^{n-1}\big) + \tau^2 f^n, \tag{3.3a}$$

$$\mathcal{B}u^{n+1} = g^{n+1}, \tag{3.3b}$$

if the first step is computed via

$$u^1 + \mathcal{B}^*\widetilde{\lambda}^1 = u^0 + \tau w^0 - \tfrac{\tau^2}{4}\mathcal{A}\big(u^0 + u^1\big) + \tfrac{\tau^2}{2}f^0, \tag{3.3c}$$

$$\mathcal{B}u^1 = g^1. \tag{3.3d}$$

Proof. We rewrite (3.2a) as

$$w^{n+1/2} - w^n = -\frac{\tau^2}{4}\mathcal{A}\, w^{n+1/2} - \frac{\tau}{2}\,\mathcal{A}u^n + \frac{\tau}{2}\,f^n - \frac{\tau}{2}\,\mathcal{B}^*\lambda^{n+1}$$

and insert this in (3.2d) to obtain

$$w^{n+1} - w^{n+1/2} = -\frac{\tau^2}{4}\mathcal{A}\, w^{n+1/2} - \frac{\tau}{2}\,\mathcal{A}u^n + \frac{\tau}{2}\,f^{n+1} - \mathcal{B}^*\big(\tfrac{\tau}{2}\,\lambda^{n+1} + \mu^{n+1}\big).$$

Using (3.2c), then gives

$$\begin{aligned}
&u^{n+1} - 2u^n + u^{n-1}\\
&\quad = \tau\left(w^{n+1/2} - w^n\right) + \tau\left(w^n - w^{n-1/2}\right)\\
&\quad = -\frac{\tau^2}{4}\mathcal{A}\left(\tau w^{n+1/2} + \tau w^{n-1/2}\right) - \frac{\tau}{2}\,\mathcal{A}\big(u^n + u^{n-1}\big) + \tau f^n - \tau\mathcal{B}^*\big(\tfrac{\tau}{2}\,\lambda^{n+1} + \tfrac{\tau}{2}\,\lambda^n + \mu^n\big)\\
&\quad = -\frac{\tau^2}{4}\mathcal{A}\left(u^{n+1} + 2u^n + u^{n-1}\right) + \tau f^n - \tau\mathcal{B}^*\big(\tfrac{\tau}{2}\,\lambda^{n+1} + \tfrac{\tau}{2}\,\lambda^n + \mu^n\big).
\end{aligned}$$

This immediately yields the equivalence for the part in $\mathcal{V}_{\mathrm{ker}}$. On the other hand, the constraints ensure equivalence for the parts in $\mathcal{V}_{\mathrm{c}}$, which gives the first part. For the initial step the very same considerations give the claim. □

Remark 3.4. To the best of our knowledge, the damped case cannot be formulated as a two-step scheme, since one cannot eliminate w^n anymore. However, in the damping-free case, formulation (3.3) is computationally much more attractive. Moreover, the two-step formulation does not need the assumption that $\mathcal{H}_{\mathrm{ker}}$ is a proper subspace of $\mathcal{H}$.

3.3. **Convergence analysis.** As first main result of this paper, we turn to the error analysis and show second-order convergence of the proposed scheme (3.2). We emphasize that the following error bounds is valid for both the damped and damping-free case, i.e., for both schemes (3.2) and (3.3).

Theorem 3.5 (Second-order convergence, damped IMEX method)**.** *Consider Assumptions 2.1, 2.2, and 2.3 as well as consistent initial data $u^0 = u^0_{ker} + \mathcal{B}^- g(0)$ with u^0_{ker} being an element of the domain of $\mathcal{A}_{\mathrm{ker}}$ and $w^0 = w^0_{ker} + \mathcal{B}^-\dot{g}(0)$ with $w^0_{ker} \in \mathcal{V}_{\mathrm{ker}}$. Let the solution u to* (2.1) *and the right-hand sides satisfy*

$$u \in C^3([0,T],\mathcal{V}) \cap C^4([0,T],\mathcal{H}), \quad f(\,\cdot\,,u(\cdot)) \in C^2([0,T],\mathcal{H}), \quad g \in C^4([0,T],\mathcal{Q}^*).$$

Then the scheme (3.2) *is convergent of second order with the error estimate*

$$\|u(t_n) - u^n\|_{\mathcal{V}} + \|\dot{u}(t_n) - w^n\|_{\mathcal{H}} \lesssim \tau^2,$$

where the hidden constant is independent of τ and n.

To prove second-order convergence, we use once more the decomposition $\mathcal{V} = \mathcal{V}_{\mathrm{ker}} \oplus \mathcal{V}_{\mathrm{c}}$ from (2.3) as discussed in Section 2.1. Due to Remark 3.2 and the assumed consistency of the initial data, we only have to bound the errors in $\mathcal{V}_{\mathrm{ker}}$ and $\mathcal{H}_{\mathrm{ker}}$,

$$u(t_n) - u^n = \big(u_{\mathrm{ker}}(t_n) - u^n_{\mathrm{ker}}\big) + \big(u_{\mathrm{c}}(t_n) - u^n_{\mathrm{c}}\big) = u_{\mathrm{ker}}(t_n) - u^n_{\mathrm{ker}}.$$

Similarly, we have by (3.2e),

$$\begin{aligned}
w(t_n) - w^n &= \big(w_{\mathrm{ker}}(t_n) - w^n_{\mathrm{ker}}\big) + \big(w_{\mathrm{c}}(t_n) - w^n_{\mathrm{c}}\big)\\
&= w_{\mathrm{ker}}(t_n) - w^n_{\mathrm{ker}} + \mathcal{B}^-\big(g^n - \tfrac{1}{\tau}\,(g^{n+1/2} - g^{n-1/2})\big)
\end{aligned}$$

$$= w_{\mathrm{ker}}(t_n) - w_{\mathrm{ker}}^n + \mathcal{O}(\tau^2).$$

We aim to show a similar error representation for the parts in the kernel of $\mathcal{B}$ as in [HL21]. For technical reasons, we introduce the following projection operator.

Lemma 3.6. *Define the orthogonal projection* $\pi_{ker}\colon \mathcal{H} \to \mathcal{H}_{\mathrm{ker}} \subseteq \mathcal{V}_{\mathrm{ker}}^*$ *via*

$$(\pi_{ker} v, \varphi_{ker})_{\mathcal{H}} = (v, \varphi_{ker})_{\mathcal{H}} \quad \text{for all} \quad \varphi_{ker} \in \mathcal{H}_{\mathrm{ker}}.$$

Then it holds $\|\pi_{ker} v\|_{\mathcal{H}} \le \|v\|_{\mathcal{H}}$ *and, since* $\mathcal{V}_{\mathrm{ker}} \subseteq \mathcal{H}_{\mathrm{ker}}$*, it particularly holds*

$$(\pi_{ker} v, \varphi_{ker})_{\mathcal{H}} = (v, \varphi_{ker})_{\mathcal{H}} \quad \text{for all} \quad \varphi_{ker} \in \mathcal{V}_{\mathrm{ker}}.$$

With $\mathcal{D}_{\mathrm{ker}}$ from Section 2.1, we introduce the operators $\mathcal{R}_\pm$ and $\mathcal{R}$ by

$$\mathcal{R}_\pm := \begin{bmatrix} \mathrm{id} & \\ & \mathrm{id} \end{bmatrix} \pm \frac{\tau}{2} \begin{bmatrix} 0 & -\mathrm{id} \\ \mathcal{A}_{\mathrm{ker}} & \mathcal{D}_{\mathrm{ker}} \end{bmatrix} : \mathcal{V}_{\mathrm{ker}} \times \mathcal{H}_{\mathrm{ker}} \to \mathcal{H}_{\mathrm{ker}} \times \mathcal{V}_{\mathrm{ker}}^*, \tag{3.4a}$$

$$\mathcal{R} := \mathcal{R}_+^{-1}\mathcal{R}_- : \mathcal{V}_{\mathrm{ker}} \times \mathcal{H}_{\mathrm{ker}} \to \mathcal{V}_{\mathrm{ker}} \times \mathcal{H}_{\mathrm{ker}}. \tag{3.4b}$$

The essential properties of these operators are collected in the following lemma.

Lemma 3.7. *The operators* $\mathcal{R}_\pm$ *and* $\mathcal{R}$ *from* (3.4) *satisfy the bounds*

$$\|\mathcal{R}\|_{\mathcal{V}_{\mathrm{ker}}\times\mathcal{H}_{\mathrm{ker}} \to \mathcal{V}_{\mathrm{ker}}\times\mathcal{H}_{\mathrm{ker}}},\ \|\mathcal{R}_+^{-1}\|_{\mathcal{H}_{\mathrm{ker}}\times\mathcal{V}_{\mathrm{ker}}^* \to \mathcal{V}_{\mathrm{ker}}\times\mathcal{H}_{\mathrm{ker}}} \ \le\ e^{c\tau}$$

for some constant $c > 0$*, independent of* τ*. Further, it holds for all* $x \in \mathcal{H}_{\mathrm{ker}}$ *the identity*

$$\tau \mathcal{R}_+^{-1} \begin{bmatrix} -x \\ \mathcal{D}_{\mathrm{ker}} x \end{bmatrix} = (\mathrm{id} - \mathcal{R}) \begin{bmatrix} 0 \\ x \end{bmatrix}$$

and, thus, the bound

$$\left\| \mathcal{R}_+^{-1} \begin{bmatrix} -x \\ \mathcal{D}_{\mathrm{ker}} x \end{bmatrix} \right\|_{\mathcal{V}_{\mathrm{ker}}\times\mathcal{H}_{\mathrm{ker}}} \le \frac{1 + e^{c\tau}}{\tau} \|x\|_{\mathcal{H}_{\mathrm{ker}}} \le \frac{2e^{c\tau}}{\tau} \|x\|_{\mathcal{H}_{\mathrm{ker}}}.$$

Proof. The bounds on $\mathcal{R}, \mathcal{R}_+^{-1}$ follow from [HL21, Lem. 2.4]. From the definition of $\mathcal{R}_\pm$, we have

$$(\mathcal{R}_+ - \mathcal{R}_-) \begin{bmatrix} 0 \\ x \end{bmatrix} = \tau \begin{bmatrix} -x \\ \mathcal{D}_{\mathrm{ker}} x \end{bmatrix},$$

and multiplying with $\mathcal{R}_+^{-1}$ gives the claim. □

With these operators, we can state an error representation of the IMEX method (3.2).

Lemma 3.8. *Let the assumptions of Theorem 3.5 hold. Then the errors in the kernel*

$$e_u^n := u_{ker}(t_n) - u_{ker}^n, \qquad e_w^n := w_{ker}(t_n) - w_{ker}^n \tag{3.5}$$

satisfy the relation

(3.6)

$$\mathcal{R}_+ \begin{bmatrix} e_u^{n+1} \\ e_w^{n+1} \end{bmatrix} = \mathcal{R}_- \begin{bmatrix} e_u^n \\ e_w^n \end{bmatrix} + \begin{bmatrix} 0 \\ \pi_{ker}\, S_2^n \end{bmatrix} + \begin{bmatrix} \pi_{ker}\, S_1^n \\ -\mathcal{D}_{\mathrm{ker}} \pi_{ker}\, S_1^n \end{bmatrix} + \begin{bmatrix} \pi_{ker}\, \delta_u^n \\ \pi_{ker}\, \delta_w^n \end{bmatrix} + \begin{bmatrix} -\pi_{ker}\,(\delta_1^n + \delta_2^n) \\ \mathcal{D}_{\mathrm{ker}} \pi_{ker}\,(\delta_1^n + \delta_2^n) \end{bmatrix},$$

where we have the two defects δ_u^n and δ_w^n, which essentially come from the trapezoidal rule and satisfy $\|\delta_u^n\|_{\mathcal{V}_{\ker}} + \|\delta_w^n\|_{\mathcal{H}_{\ker}} \leq C\tau^3$. For $\widehat{f}^n = f(t_n, u(t_n))$ and $f^n = f(t_n, u^n)$, the other terms are given by

$$S_1^n = \frac{\tau^2}{4}\left(\widehat{f}^n - \widehat{f}^{n+1}\right) - \frac{\tau^2}{4}\left(f^n - f^{n+1}\right), \qquad S_2^n = \frac{\tau}{2}\left(\widehat{f}^{n+1} + \widehat{f}^n\right) - \frac{\tau}{2}\left(f^{n+1} + f^n\right),$$
$$\delta_1^n = \frac{\tau^2}{4}\left(\widehat{f}^n - \widehat{f}^{n+1}\right), \qquad \delta_2^n = \frac{\tau}{2}\left(w_c^{n+1} + w_c^n\right) - \tau w_c^{n+1/2}.$$

Before proofing this lemma, we briefly sketch the proof of our main result.

Proof of Theorem 3.5. Resolving the error recursion in Lemma 3.8, and using the fact that $(\mathrm{id} - \mathcal{R})$ allows for summation by parts, the proofs of Theorems 2.9 and 3.3 in [HL21] also apply to the present case. □

Hence, it only remains to prove the error recursion. Here, we have to take into account all the additional contributions arising from the constraint.

Proof of Lemma 3.8. The proof is divided into three parts. First, we derive a suitable formulation for the approximations $u_{\ker}^n$ and $w_{\ker}^n$. In the second step, we derive a similar form for the exact solution. Lastly, we combine the representations to obtain the claimed formulation.

(a) From (3.2) we obtain two representations of $w^{n+1/2}$, namely

$$w^{n+1/2} = \frac{1}{2}\left(w^{n+1} + w^n\right) + \frac{\tau}{4}\left(f^n - f^{n+1}\right) + \frac{1}{2}\mathcal{B}^*\mu^{n+1}$$

as well as

$$w^{n+1/2} = -\frac{\tau}{2}\mathcal{D}w^{n+1/2} - \frac{\tau}{4}\mathcal{A}\,\tau w^{n+1/2} + w^n - \frac{\tau}{2}\mathcal{A}u^n - \frac{\tau}{2}\mathcal{B}^*\lambda^{n+1} + \frac{\tau}{2}f^n.$$

Note that both equations are stated in $\mathcal{V}^*$. From the first equation, we conclude that

$$u^{n+1} = u^n + \tau w^{n+1/2} = u^n + \frac{\tau}{2}\left(w^{n+1} + w^n\right) + \frac{\tau^2}{4}\left(f^n - f^{n+1}\right) + \frac{\tau}{2}\mathcal{B}^*\mu^{n+1}.$$

Moreover, the two representations of $w^{n+1/2}$ imply that

$$\begin{aligned}\frac{1}{2}\left(w^{n+1} + w^n\right) + \frac{\tau}{4}\left(f^n - f^{n+1}\right) + \frac{1}{2}\mathcal{B}^*\mu^{n+1}\\ = -\frac{\tau}{2}\mathcal{D}w^{n+1/2} - \frac{\tau}{4}\mathcal{A}\big(u^{n+1} - u^n\big) + w^n - \frac{\tau}{2}\mathcal{A}u^n - \frac{\tau}{2}\mathcal{B}^*\lambda^{n+1} + \frac{\tau}{2}f^n.\end{aligned}$$

We rearrange this to

$$w^{n+1} - w^n = -\tau\mathcal{D}w^{n+1/2} - \frac{\tau}{2}\mathcal{A}\big(u^{n+1} + u^n\big) + \frac{\tau}{2}\left(f^{n+1} + f^n\right) - \tau\mathcal{B}^*\lambda^{n+1} - \mathcal{B}^*\mu^{n+1}.$$

Thus, we have established

$$\begin{aligned}u^{n+1} &= u^n + \frac{\tau}{2}\left(w^{n+1} + w^n\right) + \frac{\tau^2}{4}\left(f^n - f^{n+1}\right) + \frac{\tau}{2}\mathcal{B}^*\mu^{n+1},\\ w^{n+1} &= w^n - \frac{\tau}{2}\mathcal{A}\big(u^{n+1} + u^n\big) + \frac{\tau}{2}\big(f^{n+1} + f^n\big) - \tau\mathcal{D}w^{n+1/2} - \tau\mathcal{B}^*\lambda^{n+1} - \mathcal{B}^*\mu^{n+1}.\end{aligned} \tag{3.7}$$

Restricting (3.2d) to test functions in $\mathcal{V}_{\ker}$, we obtain in $\mathcal{V}_{\ker}^*$ the equation

$$w_{\ker}^{n+1/2} = \frac{1}{2}\left(w_{\ker}^{n+1} + w_{\ker}^n\right) + \frac{\tau}{4}\pi_{\ker}\big(f^n - f^{n+1}\big) + \frac{1}{2}\pi_{\ker}\big(w_{\mathrm{c}}^{n+1} + w_{\mathrm{c}}^n\big) - \pi_{\ker}w_{\mathrm{c}}^{n+1/2}.$$

Similarly, testing (3.7) with functions in $\mathcal{V}_{\rm ker}$, it holds in $\mathcal{V}_{\rm ker}^*$,

$$\begin{aligned}
u_{\rm ker}^{n+1} &= u_{\rm ker}^n + \frac{\tau}{2}\left(w_{\rm ker}^{n+1} + w_{\rm ker}^n\right) + \frac{\tau^2}{4}\,\pi_{\rm ker}\big(f^n - f^{n+1}\big) + \pi_{\rm ker} r_1,\\
w_{\rm ker}^{n+1} &= w_{\rm ker}^n - \frac{\tau}{2}\,\mathcal{A}_{\rm ker}\left(u_{\rm ker}^{n+1} + u_{\rm ker}^n\right) + \frac{\tau}{2}\,\pi_{\rm ker}\big(f^{n+1} + f^n\big) - \tau\mathcal{D}_{\rm ker} w_{\rm ker}^{n+1/2} + \pi_{\rm ker} r_2\\
&= w_{\rm ker}^n - \frac{\tau}{2}\,\mathcal{A}_{\rm ker}\left(u_{\rm ker}^{n+1} + u_{\rm ker}^n\right) - \frac{\tau}{2}\,\mathcal{D}_{\rm ker}\big(w_{\rm ker}^{n+1} + w_{\rm ker}^n\big) + \frac{\tau}{2}\,\pi_{\rm ker}\big(f^{n+1} + f^n\big)\\
&\qquad - \frac{\tau^2}{4}\,\mathcal{D}_{\rm ker}\,\pi_{\rm ker}\big(f^n - f^{n+1}\big) + \pi_{\rm ker} r_2 - \mathcal{D}_{\rm ker}\,\pi_{\rm ker}\delta_2^n
\end{aligned}$$

with remainders

$$r_1 = u_{\rm c}^n - u_{\rm c}^{n+1} + \frac{\tau}{2}\left(w_{\rm c}^{n+1} + w_{\rm c}^n\right), \qquad r_2 = w_{\rm c}^n - w_{\rm c}^{n+1} - \tau\mathcal{D} w_{\rm c}^{n+1/2}.$$

Using that $u_{\rm c}^n - u_{\rm c}^{n+1} = \mathcal{B}^-\big(g^n - g^{n+1}\big) = -\tau w_{\rm c}^{n+1/2}$, we get $r_1 = \delta_2^n$.

(b) Noting that the exact solution satisfies

$$\dot u(t) = w(t), \qquad \dot w(t) = -\mathcal{D} w(t) - \mathcal{A} u(t) + f(t, u(t)) + \mathcal{B}^*\lambda(t),$$

we conclude together with $\dot u_{\rm c}(t) = w_{\rm c}(t)$ that

$$\dot u_{\rm ker}(t) = w_{\rm ker}(t), \quad \dot w_{\rm ker}(t) = -\mathcal{D}_{\rm ker} w_{\rm ker}(t) - \mathcal{A}_{\rm ker} u_{\rm ker}(t) + \pi_{\rm ker} f(t, u(t)) - \mathcal{D} w_{\rm c}(t) - \ddot u_{\rm c}(t).$$

We introduce the notation $\widehat u_{\rm ker}^n = u_{\rm ker}(t_n)$, $\widehat u_{\rm c}^n = u_{\rm c}(t_n)$ and further denote with δ_u^n, $\delta_{w,1}^n$ the defects from the trapezoidal rule which satisfy $\|\delta_u^n\|_{\mathcal{V}} + \|\pi_{\rm ker}\delta_{w,1}^n\|_{\mathcal{H}} \le C\tau^3$. We can thus conclude that the exact solution satisfies

$$\begin{aligned}
\widehat u_{\rm ker}^{n+1} &= \widehat u_{\rm ker}^n + \frac{\tau}{2}\left(\widehat w_{\rm ker}^{n+1} + \widehat w_{\rm ker}^n\right) + \frac{\tau^2}{4}\pi_{\rm ker}\big(\widehat f^n - \widehat f^{n+1}\big) + \pi_{\rm ker}\delta_u^n - \pi_{\rm ker}\delta_1^n\\
\widehat w_{\rm ker}^{n+1} &= \widehat w_{\rm ker}^n - \frac{\tau}{2}\,\mathcal{A}_{\rm ker}\left(\widehat u_{\rm ker}^{n+1} + \widehat u_{\rm ker}^n\right) - \frac{\tau}{2}\,\mathcal{D}_{\rm ker}\big(\widehat w_{\rm ker}^{n+1} + \widehat w_{\rm ker}^n\big) + \frac{\tau}{2}\,\pi_{\rm ker}\big(\widehat f^{n+1} + \widehat f^n\big)\\
&\quad - \frac{\tau}{2}\,\mathcal{D}\big(w_{\rm c}^{n+1} + w_{\rm c}^n\big) - \frac{\tau}{2}\big(\ddot u_{\rm c}^{n+1} + \ddot u_{\rm c}^n\big) - \frac{\tau^2}{4}\mathcal{D}_{\rm ker}\pi_{\rm ker}\big(\widehat f^n - \widehat f^{n+1}\big) + \pi_{\rm ker}\delta_{w,1}^n + \mathcal{D}_{\rm ker}\pi_{\rm ker}\delta_1^n,
\end{aligned}$$

where the defect δ_1^n is defined in the assertion of Lemma 3.8 and is artificially added to compensate the difference of $\widehat f^n$.

(c) Recalling the definitions of e_u^n and e_w^n, we obtain with parts (a) and (b),

$$\begin{aligned}
e_u^{n+1} &= e_u^n + \frac{\tau}{2}\big(e_w^{n+1} + e_w^n\big) + \pi_{\rm ker} S_1^n + \pi_{\rm ker}\delta_u^n - \pi_{\rm ker}\big(\delta_1^n + \delta_2^n\big),\\
e_w^{n+1} &= e_w^n - \frac{\tau}{2}\,\mathcal{A}_{\rm ker}\left(e_u^{n+1} + e_u^n\right) - \frac{\tau}{2}\,\mathcal{D}_{\rm ker}\big(e_w^{n+1} + e_w^n\big) + \pi_{\rm ker} S_2^n - \mathcal{D}_{\rm ker}\pi_{\rm ker} S_1^n\\
&\quad - \frac{\tau}{2}\,\mathcal{D}\big(w_{\rm c}^{n+1} + w_{\rm c}^n\big) - \frac{\tau}{2}\big(\ddot u_{\rm c}^{n+1} + \ddot u_{\rm c}^n\big) + \pi_{\rm ker}\delta_{w,1}^n + \mathcal{D}_{\rm ker}\pi_{\rm ker}\delta_1^n - \pi_{\rm ker} r_2 + \mathcal{D}_{\rm ker}\pi_{\rm ker}\delta_2^n\\
&= e_w^n - \frac{\tau}{2}\,\mathcal{A}_{\rm ker}\left(e_u^{n+1} + e_u^n\right) - \frac{\tau}{2}\,\mathcal{D}_{\rm ker}\big(e_w^{n+1} + e_w^n\big) + \pi_{\rm ker} S_2^n - \mathcal{D}_{\rm ker}\pi_{\rm ker} S_1^n\\
&\quad + \pi_{\rm ker}\delta_{w,1}^n + \pi_{\rm ker}\delta_{w,2}^n + \mathcal{D}_{\rm ker}\pi_{\rm ker}\big(\delta_1^n + \delta_2^n\big)
\end{aligned}$$

in $\mathcal{V}_{\rm ker}^*$, where $\delta_{w,2}^n$ satisfies

$$\begin{aligned}
\|\delta_{w,2}^n\|_{\mathcal{H}} &= \Big\|\frac{\tau}{2}\,\mathcal{D}\big(w_{\rm c}^{n+1} + w_{\rm c}^n\big) + \frac{\tau}{2}\big(\ddot u_{\rm c}^{n+1} + \ddot u_{\rm c}^n\big) + r_2\Big\|_{\mathcal{H}}\\
&= \Big\|\frac{\tau}{2}\,\mathcal{D}\big(w_{\rm c}^{n+1} + w_{\rm c}^n\big) + \frac{\tau}{2}\big(\ddot u_{\rm c}^{n+1} + \ddot u_{\rm c}^n\big) - \big(w_{\rm c}^{n+1} - w_{\rm c}^n + \tau\mathcal{D} w_{\rm c}^{n+1/2}\big)\Big\|_{\mathcal{H}}
\end{aligned}$$

$$= \left\| \tau \mathcal{D}\big(w_{\mathrm{c}}^{n+1/2} - \tfrac{1}{2}\,(w_{\mathrm{c}}^{n+1} + w_{\mathrm{c}}^{n})\big) - \big(w_{\mathrm{c}}^{n+1} - w_{\mathrm{c}}^{n} - \tfrac{\tau}{2}\,(\ddot{u}_{\mathrm{c}}^{n+1} + \ddot{u}_{\mathrm{c}}^{n})\big) \right\|_{\mathcal{H}}$$
$$\leq C\,\tau^3.$$

We finally define $\delta_w^n = \pi_{\ker}(\delta_{w,1}^n + \delta_{w,2}^n)$ and use the definition of $\mathcal{R}_\pm$ to obtain the relation (3.6). □

4. A Gautschi-type Integrator

This section is devoted to the extension of Gautschi-type integrators to constrained systems. For this, we restrict ourselves to the damping-free free case, i.e., we set $\mathcal{D} = 0$. Moreover, we consider the stronger assumption on the nonlinearity, i.e., Assumption 2.6 (a).

As in the previous section, we consider a uniform partition of $[0, T]$ with constant time step size τ and discrete time points $t_n = n\,\tau$.

4.1. **Review of the unconstrained case.** We recall the concept of Gautschi-type integrators for wave-type systems with bounded operators as well as their extension to unbounded operators. For this, an interpretation of trigonometric functions applied to an operator is needed.

Bounded operators. We recall solution strategies for conservative second-order systems as discussed in [HL99, Gri02]. Considering the corresponding first-order formulation, we consider the equations

$$\dot{u}(t) - w(t) = 0,$$
$$\dot{w}(t) + \mathcal{A}u(t) = f(t, u)$$

with a bounded operator $\mathcal{A}$, which also has a bounded inverse. In terms of operator matrices, this system reads

$$\begin{bmatrix} \dot{u}(t) \\ \dot{w}(t) \end{bmatrix} + \begin{bmatrix} 0 & -\mathrm{id} \\ \mathcal{A} & 0 \end{bmatrix} \begin{bmatrix} u(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} 0 \\ f(t, u) \end{bmatrix}, \tag{4.1}$$

which we abbreviate by $\dot{z} + \mathcal{Z}z = \tilde{f}$ for the unknown $z := [u; w]$, where we use the standard Matlab notation for the concatenation of vectors. The initial data is given by $z(0) = [u^0; w^0]$. By an application of the variation-of-constants formula [Paz83, Ch. 4], we conclude

$$z(t) = e^{-t\mathcal{Z}} z(0) + \int_0^t e^{-(t-s)\mathcal{Z}} \tilde{f}(s, z)\,\mathrm{d}s.$$

Under the assumption that $\mathcal{A}$ is bounded, the exponential is well-defined and has the form

$$e^{-t\mathcal{Z}} = \begin{bmatrix} \cos(t\Omega) & \Omega^{-1}\sin(t\Omega) \\ -\Omega\,\sin(t\Omega) & \cos(t\Omega) \end{bmatrix} = \begin{bmatrix} \cos(t\Omega) & t\,\operatorname{sinc}(t\Omega) \\ -t\mathcal{A}\,\operatorname{sinc}(t\Omega) & \cos(t\Omega) \end{bmatrix} \tag{4.2}$$

with $\Omega := \mathcal{A}^{1/2}$; cf. [Gri02, Sect. 3.1]. Although written in terms of the square root of $\mathcal{A}$, the functions sinc and cos given by the Taylor series

$$\operatorname{sinc} x = \sum_{k=0}^{\infty} \frac{(-x^2)^k}{(2k+1)!}, \qquad \cos x = \sum_{k=0}^{\infty} \frac{(-x^2)^k}{(2k)!} \tag{4.3}$$

only depend on even powers of the input such that $\operatorname{sinc}(t\Omega)$ and $\cos(t\Omega)$ are well-defined even if Ω is not. By formula (4.2), the solution u being equal to the first component of z can be written as

$$u(t) = \cos(t\Omega)\, u^0 + t\, \operatorname{sinc}(t\Omega)\, w^0 + \int_0^t (t-s)\, \operatorname{sinc}((t-s)\Omega)\, f(s,u)\, \mathrm{d}s. \tag{4.4}$$

For further calculations, we commemorate the identity

$$\frac{1-\cos x}{x^2} = \int_0^1 x^{-1} \sin\big((1-\eta)\, x\big)\, \mathrm{d}\eta = \int_0^1 (1-\eta)\, \operatorname{sinc}\big((1-\eta)\, x\big)\, \mathrm{d}\eta.$$

Inserting $x = \tau\, \Omega$ and using a substitution to solve the integral, we get

$$\mathcal{A}^{-1}\big[1-\cos(\tau\Omega)\big] = \int_0^1 \tau^2\, (1-\eta)\, \operatorname{sinc}\big((1-\eta)\, \tau\Omega\big)\, \mathrm{d}\eta = \int_0^\tau (\tau - s)\, \operatorname{sinc}\big((\tau-s)\, \Omega\big)\, \mathrm{d}s.$$

With the solution formula (4.4), we conclude for the solution of (4.1) with a constant right-hand side, i.e., for $f(t,u) \equiv f$, that

$$\begin{aligned} u(t+\tau) + u(t-\tau) &= 2\, \cos(\tau\Omega)\, u(t) + 2 \int_0^\tau (\tau - s)\, \operatorname{sinc}((\tau-s)\, \Omega)\, f\, \mathrm{d}s \\ &= 2\, \cos(\tau\Omega)\, u(t) + 2\, \mathcal{A}^{-1}\big[1-\cos(\tau\Omega)\big]\, f \\ &= 2\, \cos(\tau\Omega)\big[u(t) - \mathcal{A}^{-1} f\big] + 2\, \mathcal{A}^{-1} f. \end{aligned}$$

In the semi-linear case, we may approximate the integral term by the simplest integration scheme, i.e., we replace $f(t,u)$ by $f(t_n, u^n)$ with $u^n \approx u(t_n)$. Inserting this in the previous calculation yields the integration scheme

$$u^{n+1} = 2\, \cos(\tau\Omega)\big[u^n - \mathcal{A}^{-1} f(t_n, u^n)\big] + 2\, \mathcal{A}^{-1} f(t_n, u^n) - u^{n-1}. \tag{4.5}$$

Following [HL99], this may also be written in the alternative form

$$u^{n+1} - 2u^n + u^{n-1} = 2\, \mathcal{A}^{-1}\big(1 - \cos(\tau\Omega)\big)\, \big[f(t_n, u^n) - \mathcal{A} u^n\big]$$

or, due to $2\,(1-\cos(x))/x^2 = \operatorname{sinc}^2(\frac{x}{2})$, as

$$u^{n+1} - 2u^n + u^{n-1} = \tau^2 \operatorname{sinc}^2(\tfrac{\tau}{2}\Omega)\, \big[f(t_n, u^n) - \mathcal{A} u^n\big].$$

Extension to unbounded operators. So far, we have assumed that the operator $\mathcal{A}$ is bounded. Hence, the exponential as well as $\cos(t\Omega)$ are well-defined, e.g., using the Taylor series. In the here considered PDAE setting, $\mathcal{A}$ usually denotes an unbounded differential operator such that a different approach is needed. For this, consider the situation as introduced in Section 2.1 with the Hilbert spaces $\mathcal{V}, \mathcal{H}, \mathcal{Q}$ and an operator $\mathcal{A}$ as in Lemma 2.9.

In order to make the numerical scheme (4.5) also valid for unbounded operators, we need an interpretation of the trigonometric terms. For this, we assume for a minute that $\mathcal{A}$ is elliptic. In this case, we can interpret $\cos(t\Omega)\, v$ as the solution of the linear and homogeneous equation

$$\ddot{u}(t) + \mathcal{A} u(t) = 0, \qquad u(0) = v,\ \dot{u}(0) = 0. \tag{4.6}$$

Moreover, we can understand $t\, \operatorname{sinc}(t\Omega)\, v$ as the solution of

$$\ddot{u}(t) + \mathcal{A} u(t) = 0, \qquad u(0) = 0,\ \dot{u}(0) = v. \tag{4.7}$$

Note that both systems are well-posed by Lemma 2.9.

4.2. **An implicit solution formula.** We come back to the constrained case and consider the PDAE (2.1). In the linear case, i.e., if $f(t,u) \equiv f(t)$, we can explicitly write down the solution using the decomposition of $\mathcal{V}$ from (2.3), the first-order formulation, and the variation-of-constants formula; cf. [EM13]. This solution formula is also applicable to the general semi-linear case, leading to an implicit formula. As in the unconstrained case, quadrature rules then yield numerical schemes.

Following (2.3), we decompose the solution into $u = u_{\ker} + u_{\mathrm{c}}$ with $u_{\ker}\colon [0,T] \to \mathcal{V}_{\ker}$ and $u_{\mathrm{c}} = \mathcal{B}^- g(t) \in \mathcal{V}_{\mathrm{c}}\colon [0,T] \to \mathcal{V}_{\mathrm{c}}$. Then the restriction of (2.1a) to test functions in $\mathcal{V}_{\ker}$ yields

$$\ddot{u}_{\ker} + \mathcal{A}_{\ker} u_{\ker} = f(t, u_{\ker} + u_{\mathrm{c}}) - \ddot{u}_{\mathrm{c}} \qquad \text{in } \mathcal{V}_{\ker}^*. \tag{4.8}$$

Note that there is no term $\mathcal{A}u_{\mathrm{c}}$ on the right-hand side due to the particular definition of $\mathcal{V}_{\mathrm{c}}$. Since this is an unconstrained system with an elliptic differential operator $\mathcal{A}_{\ker}$, the variation-of-constants formula can be applied to the corresponding first-order formulation. With the operator $\pi_{\ker}\colon \mathcal{H} \to \mathcal{H}_{\ker}$ introduced in Lemma 3.6, which is simply based on the restriction of test functions, we conclude that

$$\begin{aligned} u(t) &= u_{\mathrm{c}}(t) + u_{\ker}(t) \\ &= \mathcal{B}^- g(t) + \cos(t\Omega_{\ker}) u_{\ker}(0) + t\,\operatorname{sinc}(t\Omega_{\ker}) \dot{u}_{\ker}(0) \\ &\qquad + \int_0^t (t-s)\, \operatorname{sinc}((t-s)\,\Omega_{\ker})\, \pi_{\ker}\big[f(s, u_{\ker}(s) + \mathcal{B}^- g(s)) - \mathcal{B}^- \ddot{g}(s)\big]\, \mathrm{d}s. \end{aligned} \tag{4.9}$$

Here, we use the notation $u_{\ker}(0) \coloneqq u^0 - \mathcal{B}^- g(0)$, $\dot{u}_{\ker}(0) \coloneqq w^0 - \mathcal{B}^- \dot{g}(0)$ and $\Omega_{\ker} \coloneqq \mathcal{A}_{\ker}^{1/2}$. Note that this solution formula is *implicit* in the sense that the solution also appears on the right-hand side within the integral term. Nevertheless, this representation is the basis for the derivation of a Gautschi-type integrator as we will see in the following.

4.3. **A Gautschi-type integration scheme.** Based on the solution formula (4.9) we are now able to construct a Gautschi-type integrator. For this, we fix the nonlinear term in the integral of (4.9) by its value on the left limit, i.e., by $f(t_n, u^n) - \mathcal{B}^- \ddot{g}^n$. With

$$b^n \coloneqq \mathcal{A}_{\ker}^{-1}\, \pi_{\ker}\big[f(t_n, u^n) - \mathcal{B}^- \ddot{g}^n\big],$$

we obtain by a calculation similarly as in Section 4.1 that

$$u^{n+1} = \mathcal{B}^- g^{n+1} + \cos(\tau\Omega_{\ker})(u^n - \mathcal{B}^- g^n) + \tau\,\operatorname{sinc}(\tau\Omega_{\ker})(\dot{u}^n - \mathcal{B}^- \dot{g}^n) - \cos(\tau\Omega_{\ker})\, b^n + b^n.$$

Considering the analog formula for u^{n-1}, we obtain the scheme

$$u^{n+1} = \mathcal{B}^- g^{n+1} - (u^{n-1} - \mathcal{B}^- g^{n-1}) + 2\cos(\tau\Omega_{\ker})(u^n - \mathcal{B}^- g^n - b^n) + 2\, b^n. \tag{4.10}$$

Remark 4.1. The proposed numerical scheme (4.10) is in line with Section 4.1. This means that we recover the scheme (4.5) if we leave out the constraint.

We would like to emphasize that the Gautschi-type integrator (4.10) is a two-step scheme. Hence, we need to provide approximations u^0 and u^1 to apply the scheme. For u^0 we can directly insert $u(0)$. For the second value, we need to ensure $u^1 = u(\tau) + \mathcal{O}(\tau^3)$ in order to allow a second-order accuracy of the scheme, see Theorem 4.4 below. As an example, one may use a Taylor series approach in combination with (4.8), leading to

$$u^1 \coloneqq u^0 + \tau\, w^0 + \tfrac{1}{2}\tau^2 \big[f(0, u^0) - \mathcal{A}_{\ker}(u^0 - \mathcal{B}^- g^0)\big].$$

Note, however, that this requires a sufficiently regular u^0.

4.4. **Convergence analysis.** As preparation for the upcoming convergence proof, we summarize some facts on $\cos(t\Omega)$, $\operatorname{sinc}(t\Omega)\colon [0,\infty) \to \mathcal{L}(\mathcal{V})$. Recall that these time-dependent operators should be understood as part of the solution mapping, see (4.6) and (4.7).

Lemma 4.2. *Let $\mathcal{A} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ satisfy the assumptions of Lemma 2.9 and set $\Omega = \mathcal{A}^{1/2}$. Then the operators $\cos(t\Omega)$, $t\operatorname{sinc}(t\Omega)$, and $t\mathcal{A}\operatorname{sinc}(t\Omega)$ are extendable to $t \in \mathbb{R}$ and it holds that*

a) $\cos(t\Omega)$ *is bounded as operator from $\mathcal{V} \to \mathcal{V}$ and $\mathcal{H} \to \mathcal{H}$ for every $t \in \mathbb{R}$,*

b) $\cos((t_1+t_2)\,\Omega) = \cos(t_1\Omega)\cos(t_2\Omega) - t_1\operatorname{sinc}(t_1\Omega)\,t_2\mathcal{A}\operatorname{sinc}(t_2\Omega)$ *for every $t_1, t_2 \in \mathbb{R}$,*

c) $\mathrm{id} = \cos^2(t\Omega) + t\operatorname{sinc}(t\Omega)\,t\mathcal{A}\operatorname{sinc}(t\Omega)$ *for every $t \in \mathbb{R}$.*

Proof. a) For non-negative t and as an operator in $\mathcal{L}(\mathcal{V})$ the assertion follows immediately from estimate (2.6). In particular, we get the bound 1 in the symmetric case and $e^{\kappa t/2}$ in the general case with $\mathcal{A}_2 \neq 0$. For $u(t) = \cos(t\Omega)\,u^0$, the function $y(t) \coloneqq \int_0^t u(s)\,\mathrm{d}s$ solves (4.1) with homogeneous right-hand side and initial values $y(0) = 0$ and $\dot{y}(0) = u^0$. In particular, it holds that $\dot{y} = u$. Therefore, the assertion follows again from (2.6) by extending the operator from $\mathcal{L}(\mathcal{V}, \mathcal{H})$ to $\mathcal{L}(\mathcal{H})$. Finally, for negative t we notice that system (4.1) with homogeneous right-hand side is invariant under time reversion such that $\cos(-t\Omega) = \cos(t\Omega)$.

b) Using the semigroup (4.2) we have

$$\begin{aligned}\begin{bmatrix}\cos((t_1+t_2)\,\Omega)\,u^0\\ *\end{bmatrix} &= \left(e^{-t_1\mathcal{Z}}e^{-t_2\mathcal{Z}}\right)\begin{bmatrix}u^0\\0\end{bmatrix}\\ &= \begin{bmatrix}\cos(t_1\Omega)\cos(t_2\Omega)u^0 - t_1\operatorname{sinc}(t_1\Omega)\,t_2\mathcal{A}\operatorname{sinc}(t_2\Omega)u^0\\ *\end{bmatrix}\end{aligned}$$

for all $u^0 \in \mathcal{V}$. Analogously to a), this equation holds for every $u^0 \in \mathcal{H}$ as well.

c) Since (4.1) is invariant under time reversion, it holds that $t\mathcal{A}\operatorname{sinc}(t\Omega)$ is odd. Therefore, it holds that

$$\begin{bmatrix}u^0\\0\end{bmatrix} = (e^{t\mathcal{Z}}e^{-t\mathcal{Z}})\begin{bmatrix}u^0\\0\end{bmatrix} = \begin{bmatrix}\cos^2(t\Omega)u^0 + t\operatorname{sinc}(t\Omega)\,t\mathcal{A}\operatorname{sinc}(t\Omega)u^0\\0\end{bmatrix},$$

which completes the proof. □

Before we investigate the convergence order we state the following auxiliary lemma.

Lemma 4.3. *Let $t > 0$ and $\cos(t\Omega)$ be defined as in Lemma 4.2. Suppose that $X_0 \coloneqq \mathrm{id}$, $X_1 \coloneqq 2\cos(t\Omega)$, and*

$$X_n \coloneqq 2\cos(t\Omega)X_{n-1} - X_{n-2} \qquad \text{for } n \geq 2.$$

For $n \geq 0$ it then holds that

$$X_n = \begin{cases}\mathrm{id} + 2\sum_{k=1}^{n/2}\cos(2kt\Omega), & \text{if } n \text{ is even}\\ 2\sum_{k=1}^{(n+1)/2}\cos\big((2k-1)t\Omega\big), & \text{if } n \text{ is odd}\end{cases}. \tag{4.11}$$

Proof. The proof follows by mathematical induction. The assertion is clear for X_0 and X_1. If n is even, then we have

$$\begin{aligned}
X_n &= 2\cos(t\Omega)X_{n-1} - X_{n-2} \\
&= 4\cos(t\Omega)\sum_{k=1}^{n/2}\cos\big((2k-1)\,t\Omega\big) - \mathrm{id} - 2\sum_{k=1}^{n/2-1}\cos(2kt\Omega) \\
&\overset{\text{Lem. 4.2 } b)}{=} 2\sum_{k=2}^{n/2}\Big[\cos(t\Omega)\cos\big((2k-1)\,t\Omega\big) - \sin\big((2k-1)\,t\Omega\big)\sin(t\Omega)\Big] + 4\cos^2(t\Omega) - \mathrm{id} \\
&\overset{\text{Lem. 4.2 } b)}{=} 2\sum_{k=1}^{n/2}\cos(2kt\Omega) + 2\cos^2(t\Omega) + 2\sin^2(t\Omega) - \mathrm{id} \\
&\overset{\text{Lem. 4.2 } c)}{=} 2\sum_{k=1}^{n/2}\cos(2kt\Omega) + \mathrm{id}\,.
\end{aligned}$$

Note that we use $2kt = (2k+1)t + (-t)$ within the first application of Lemma 4.2 b) in combination with an index shift. Moreover, we would like to emphasize at this point that $\sin(t\Omega)$ may not be defined. However, due to $x\,\mathrm{sinc}(x) = \sin(x)$, the above expressions only contain Ω^2, which makes it well-defined. If, on the other hand, n is odd, then we get

$$\begin{aligned}
X_n &= 2\cos(t\Omega)X_{n-1} - X_{n-2} \\
&= 2\cos(t\Omega)\Big(\mathrm{id} + 2\sum_{k=1}^{(n-1)/2}\cos(2kt\Omega)\Big) - 2\sum_{k=1}^{(n-1)/2}\cos\big((2k-1)\,t\Omega\big) \\
&\overset{\text{Lem. 4.2 } b)}{=} 2\cos(t\Omega) + 2\sum_{k=1}^{(n-1)/2}\Big[\cos(t\Omega)\cos(2kt\Omega) - \sin(t\Omega)\sin(2kt\Omega)\Big] \\
&\overset{\text{Lem. 4.2 } b)}{=} 2\cos(t\Omega) + 2\sum_{k=1}^{(n-1)/2}\cos\big((2k+1)\,t\Omega\big) \\
&= 2\sum_{k=1}^{(n+1)/2}\cos\big((2k-1)\,t\Omega\big).
\end{aligned}$$

□

We now turn to the second main result of this paper, namely the proof of second-order convergence of the proposed two-step Gautschi-type integrator (4.10).

Theorem 4.4 (Second-order convergence, Gautschi-type integrator)**.** *Let Assumptions 2.1, 2.2, and 2.6 (a) be satisfied and assume consistent initial data $u^0 \in \mathcal{V}$, $w^0 \in \mathcal{H}$ leading to u^1 satisfying $\|u^1 - u(t_1)\|_{\mathcal{H}} \le \tau^3 E_{init}$. Further assume right-hand sides*

$$f(\,\cdot\,, u(\cdot)) \in C^2([0,T],\mathcal{H}), \qquad g \in C^4([0,T],\mathcal{Q}^*).$$

Then the Gautschi-type integrator (4.10) *is second-order accurate with the error estimate*

$$\|u^n - u(t_n)\|_{\mathcal{H}} \le \tau^2\, t_n\, e^{\kappa t_n}\cosh(t_n L^{1/2})\,\big(E_{init} + t_n C_{f,g}\big) + \mathcal{O}(\tau^3)$$

with

$$C_{f,g} \coloneqq \big\|\ddot{f}(\,\cdot\,, u)\,\big\|_{L^\infty(0,t_n;\mathcal{H}^*)} + C_{\mathcal{V}\to\mathcal{H}} C_{\mathcal{B}^-}\big\|g^{(4)}\big\|_{L^\infty(0,t_n;\mathcal{Q}^*)}$$

including the embedding constant $C_{\mathcal{V}\to\mathcal{H}}$, the continuity constant $C_{\mathcal{B}^-}$ of the operator $\mathcal{B}^-$, $\kappa = C_2/\sqrt{\mu_1}$ with the continuity constant of $\mathcal{A}_2$ and the ellipticity constant of $\mathcal{A}_1$, and $L \coloneqq \exp(\kappa\, t_n) L_{\mathcal{H}}$.

Proof. Considering (4.9) at times t_{n+1} and t_{n-1} (and omitting the restriction operator $\pi_{\ker}$ for simplicity) yields

$$\begin{aligned} u(t_{n+1}) + u(t_{n-1}) &= \mathcal{B}^- g(t_{n+1}) + \mathcal{B}^- g(t_{n-1}) + 2\cos(t\Omega_{\ker})\big[u(t_n) - \mathcal{B}^- g(t_n)\big] \\ &\quad + \int_0^\tau (\tau - s)\, \operatorname{sinc}((\tau - s)\,\Omega_{\ker}) \big[f(t_n + s, u(t_n + s)) - \mathcal{B}^- \ddot{g}(t_n + s) \\ &\qquad + f(t_n - s, u(t_n - s)) - \mathcal{B}^- \ddot{g}(t_n - s)\big]\, \mathrm{d}s. \end{aligned}$$

In combination with the approximate solution and formula (4.10), we obtain for the error $e^n \coloneqq u^n - u(t_n)$ the equation

$$\begin{aligned} e^{n+1} + e^{n-1} &= 2\cos(\tau\Omega_{\ker}) e^n + 2\left(1 - \cos(\tau\Omega_{\ker})\right)\mathcal{A}_{\ker}^{-1}\left[f(t_n, u^n) - \mathcal{B}^- \ddot{g}^n\right] \\ &\quad - \int_0^\tau (\tau - s)\, \operatorname{sinc}((\tau - s)\,\Omega_{\ker}) \big[f(t_n + s, u(t_n + s)) - \mathcal{B}^- \ddot{g}(t_n + s) \\ &\qquad + f(t_n - s, u(t_n - s)) - \mathcal{B}^- \ddot{g}(t_n - s)\big]\, \mathrm{d}s. \end{aligned}$$

Using the equality

$$\left(1 - \cos(\tau\Omega_{\ker})\right)\mathcal{A}_{\ker}^{-1} h_n = \int_0^\tau (\tau - s)\, \operatorname{sinc}\left((\tau - s)\,\Omega_{\ker}\right) h_n \,\mathrm{d}s$$

with $h_n = f(t_n, u(t_n)) - \mathcal{B}^- \ddot{g}^n$, we get

$$\begin{aligned} e^{n+1} + e^{n-1} &= 2\cos(\tau\Omega_{\ker}) e^n + 2\left(1 - \cos(\tau\Omega_{\ker})\right)\mathcal{A}_{\ker}^{-1}\left[f(t_n, u^n) - f(t_n, u(t_n))\right] \\ &\quad - \int_0^\tau (\tau - s)\, \operatorname{sinc}((\tau - s)\,\Omega_{\ker}) \big[f(t_n + s, u(t_n + s)) - \mathcal{B}^- \ddot{g}(t_n + s) \\ &\qquad - 2f(t_n, u(t_n)) + 2\mathcal{B}^- \ddot{g}(t_n) + f(t_n - s, u(t_n - s)) - \mathcal{B}^- \ddot{g}(t_n - s)\big]\, \mathrm{d}s. \end{aligned} \tag{4.12}$$

Note that the second summand of the right-hand side equals the solution of (4.8) with constant right-hand side $2f(t_n, u^n) - 2f(t_n, u(t_n)) \in \mathcal{H}^*$ and homogeneous initial values (and $g = 0$) evaluated at $t = \tau$. We denote this solution by $\widetilde{u}_n$ and the stability estimate from Lemma 2.9 implies

$$\left\|\dot{\widetilde{u}}_n(t)\right\|_{\mathcal{H}} \le 2\, e^{\frac{\kappa t}{2}} \int_0^t e^{-\kappa s} \left\|f(t_n, u^n) - f(t_n, u(t_n))\right\|_{\mathcal{H}^*} \mathrm{d}s.$$

This turns into

$$\begin{aligned} \|\widetilde{u}_n(\tau)\|_{\mathcal{H}} \le \int_0^\tau \|\dot{\widetilde{u}}_n(t)\|_{\mathcal{H}}\, \mathrm{d}t &\le \int_0^\tau 2\, e^{\frac{\kappa t}{2}} \int_0^t e^{-\kappa s} \left\|f(t_n, u^n) - f(t_n, u(t_n))\right\|_{\mathcal{H}^*} \mathrm{d}s\, \mathrm{d}t \\ &\le 2\, L_{\mathcal{H}} \|e_n\|_{\mathcal{H}} \int_0^\tau e^{\frac{\kappa t}{2}} \int_0^t e^{-\kappa s}\, \mathrm{d}s\, \mathrm{d}t \\ &\le \tau^2\, L_{\mathcal{H}}\, e^{\frac{\kappa\tau}{2}} \|e_n\|_{\mathcal{H}} \end{aligned}$$

with $L_{\mathcal{H}}$ denoting the Lipschitz constant of f from Assumption 2.6. Also the integral term in (4.12) equals the solution of (4.8) evaluated at $t = \tau$ for homogeneous initial data but now with right-hand side

$$f(t_n + t, u(t_n + t)) - 2f(t_n, u(t_n)) + f(t_n - t, u(t_n - t)) - \mathcal{B}^- \ddot{g}(t_n + t) + 2\mathcal{B}^- \ddot{g}(t_n) - \mathcal{B}^- \ddot{g}(t_n - t).$$

We denote this solution by $\widehat{u}_n$. Using a Taylor expansion, we obtain similarly as before

$$\begin{aligned}
\|\widehat{u}_n(\tau)\|_{\mathcal{H}} &\le \int_0^\tau \|\dot{\widehat{u}}_n(t)\|_{\mathcal{H}}\,\mathrm{d}t \\
&\le e^{\frac{\tau\kappa}{2}} \int_0^\tau \int_0^t \big\|f(t_n+s,u(t_n+s)) - 2f(t_n,u(t_n)) + f(t_n-s,u(t_n-s))\big\|_{\mathcal{H}^*} \\
&\qquad\qquad + \big\|\mathcal{B}^-\ddot{g}(t_n+s) - 2\mathcal{B}^-\ddot{g}(t_n) + \mathcal{B}^-\ddot{g}(t_n-s)\big\|_{\mathcal{H}^*}\,\mathrm{d}s\,\mathrm{d}t \\
&\le e^{\frac{\tau\kappa}{2}} \int_0^\tau \int_0^t \|s^2\ddot{f}(t_n,u(t_n))\|_{\mathcal{H}^*} + \|s^2\mathcal{B}^- g^{(4)}(t_n)\|_{\mathcal{H}^*} + \mathcal{O}(s^3)\,\mathrm{d}s\,\mathrm{d}t \\
&\le \frac{\tau^4}{12}\, e^{\frac{\tau\kappa}{2}} \Big(\big\|\ddot{f}(t_n,u(t_n))\big\|_{\mathcal{H}^*} + C_{\mathcal{V}\to\mathcal{H}} C_{\mathcal{B}^-} \|g^{(4)}(t_n)\|_{\mathcal{Q}^*} \Big) + \mathcal{O}(\tau^5).
\end{aligned}$$

To summarize, we have the recursion

$$e^{n+1} + e^{n-1} = 2\cos(\tau\Omega_{\mathrm{ker}})e^n + \widetilde{u}_n(\tau) + \widehat{u}_n(\tau).$$

By mathematical induction, one can show that this can be written in the explicit form

$$e^n = -X_{n-2}\,e_0 + X_{n-1}\,e_1 + \sum_{j=1}^{n-1} X_{n-1-j}\,\big(\widetilde{u}_j(\tau) - \widehat{u}_j(\tau)\big)$$

with the X_j from Lemma 4.3. With formula (4.11) and the boundedness of $\cos(\tau\Omega_{\mathrm{ker}})$ shown in Lemma 4.2, these X_j satisfy

$$\|X_j\|_* \coloneqq \|X_j\|_{\mathcal{L}(\mathcal{H},\mathcal{H})} \le (j+1)\,e^{2\frac{\kappa j\tau}{2}} = (j+1)\,e^{\kappa t_j}.$$

and accordingly $\|X_j\|_* \le j+1$ in the symmetric case. With $C_{f,g}$ as defined above, this then leads to

$$\begin{aligned}
\|e^n\|_{\mathcal{H}} &\le 0 + \|X_{n-1}\|_*\tau^3 E_{\mathrm{init}} + \sum_{j=1}^{n-1} \|X_{n-1-j}\|_*\big(\|\widetilde{u}_j(\tau)\|_{\mathcal{H}} + \|\widehat{u}_j(\tau)\|_{\mathcal{H}}\big) \\
&\le n\,\tau^3\,e^{\kappa t_n}E_{\mathrm{init}} + \sum_{j=1}^{n-1}(n-j)\,e^{\kappa t_{n-j}}\big(\|\widetilde{u}_j(\tau)\|_{\mathcal{H}} + \|\widehat{u}_j(\tau)\|_{\mathcal{H}}\big) \\
&\lesssim \tau^2\,t_n\,e^{\kappa t_n}\big(E_{\mathrm{init}} + t_n C_{f,g}\big) + \tau^2\,e^{\kappa t_n}L_{\mathcal{H}}\sum_{j=1}^{n-1}(n-j)\,\|e_j\|_{\mathcal{H}} + \mathcal{O}(\tau^3).
\end{aligned}$$

An application of [GASSS98, Lem. 2] finally gives the stated result. □

4.5. **Practical Implementation.** Before presenting numerical examples, we need to discuss how the introduced Gautschi-type integrator (4.10) can be implemented. We first perform a spatial discretization, e.g., by finite elements, leading to matrices A and B as discrete versions of $\mathcal{A}$ and $\mathcal{B}$, respectively. Moreover, we obtain a mass matrix M, which appears in front of the second-order term.

In the following, we explain how to compute the action of $\mathcal{B}^-$ and $\cos(\tau\Omega_{\mathrm{ker}})$ in the discrete setting. For this, Ω_{ker} needs to be replaced by the square root of some matrix A_{ker} (analogously to $\mathcal{A}_{\mathrm{ker}}$). Here, we consider A_{ker} as a mapping $\ker B \subseteq \mathbb{R}^n \to \ker B \subseteq \mathbb{R}^n$. Since we do not want to compute the kernel of B explicitly, we will only describe the action of A_{ker} applied to a vector in terms of an auxiliary saddle point problem rather than

assembling the operator itself. Moreover, we are concerned with the efficient evaluation of the cosine-term, for which we introduce an Arnoldi algorithm.

First, we consider the computation of $B^- g^n$ for some $g^n \in \mathbb{R}^m$. Following the construction in Section 2.1, B^- should map into (the discrete version of) $\mathcal{V}_c$, i.e., the result should be A-orthogonal to the kernel of B. To be precise, $B^- g^n$ should satisfy $w^T A(B^- g^n) = 0$ for all $w \in \ker B \subseteq \mathbb{R}^n$. To achieve this, we set $x = B^- g^n \in \mathbb{R}^n$ as the solution of the saddle point problem

$$Ax + B^T \mu = 0, \tag{4.13a}$$

$$Bx = g^n. \tag{4.13b}$$

Note that this system has a unique solution pair $(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^m$, since B is of full rank and A is assumed to be positive definite on $\ker B$. Obviously, equation (4.13b) guarantees that $Bx = BB^- g^n = g^n$. Moreover, the desired orthogonality condition is satisfied, since $w \in \ker B$ yields due to (4.13a),

$$w^T Ax = w^T(Ax + B^T \mu) = 0.$$

Second, we discuss the application of the operator $\mathcal{A}_{\ker}$ in the discrete setting. Given a vector $v \in \ker B \subseteq \mathbb{R}^n$, we obtain $x = A_{\ker} v$ as the solution of the saddle point problem

$$Mx + B^T \mu = Av, \tag{4.14a}$$

$$Bx = 0. \tag{4.14b}$$

Due to the assumed properties on M and B, this system is again uniquely solvable. In order to see that x is indeed the desired vector, first note that equation (4.14b) ensures that x is an element of $\ker B$. The first equation, on the other hand, corresponds to

$$(x, w)_{\mathcal{H}} = \langle \mathcal{A} v, w \rangle = \langle \mathcal{A}_{\ker} v, w \rangle, \qquad w \in \mathcal{V}_{\ker}$$

in the continuous setting. The application of the operator $\mathcal{A}_{\ker}$ is crucial for the computation of the cosine-term. For the intermediate step of computing b^n, we also need the inverse of $\mathcal{A}_{\ker}$. In the discrete setting, $b^n = A_{\ker}^{-1} v$ is given by the solution of

$$\begin{aligned} Ab^n + B^T \mu &= v, \\ Bb^n &= 0. \end{aligned}$$

Finally, we are concerned with the computation of $\cos(\tau \Omega_{\ker}) x_0$ for a given vector $x_0 \in \ker B \subseteq \mathbb{R}^n$. Since the Taylor series of $\cos(\tau \Omega_{\ker})$ only contains even powers of $\Omega_{\ker}$, we do not need to compute $\Omega_{\ker}$ itself. It suffices to compute $A_{\ker}$ or rather its action on a vector. Hence, we consider the Krylov subspace spanned by the vectors $A_{\ker}^k x_0$ for $k \geq 0$. Following the general idea of Krylov subspace methods (see, e.g., [Saa92, EE06]), we introduce

$$\mathcal{K}_r \coloneqq \mathcal{K}_r(A_{\ker}, x_0) \coloneqq \operatorname{span}\left\{x_0, A_{\ker} x_0, \dots, A_{\ker}^{r-1} x_0\right\}.$$

Obviously, this space does not rely on the computation of $\Omega_{\ker}$. Moreover, we would like to recall that $A_{\ker}$ is never computed explicitly. Instead, only the action of $A_{\ker}$ is computed via the saddle point problem (4.14). In a second step, we generate an orthogonal basis of $\mathcal{K}_r$ using the Arnoldi algorithm with $v_1 = x_0 / \|x_0\|$ as initial vector. Collecting these vectors in the isometric matrix $V_r \in \mathbb{R}^{n \times r}$, we obtain $V_r^T A V_r = H_r$ with an upper Hessenberg matrix $H_r \in \mathbb{R}^{r \times r}$. This then yields the approximation

$$\cos(\tau \Omega_{\ker}) x_0 \approx \|x_0\| \, V_r \cos(\tau H_r) e_1$$

with unit basis vector $e_1 \in \mathbb{R}^r$. For the computation of $\cos(\tau H_r)$, which is of moderate size, one may use methods described in [Hig08, Ch. 12]. For our numerical examples, we use the algorithm introduced in [AMHR15].

Remark 4.5. Using the Arnoldi method for the computation of the matrix cosine leaves open the question of stopping criteria. Here, one may apply the determination criterion presented in [HPS+15]. In the numerical experiment of Section 5, $r = 3$ was already sufficient.

5. Numerical Example

As numerical example, we consider the semi-linear wave equation with kinetic boundary conditions without damping. In the strong form, the system equations read

$$\ddot{u} - \Delta u + u = f_\Omega(t,u) \qquad \text{in } \Omega, \tag{5.1a}$$

$$\ddot{u} - \Delta_\Gamma u + \partial_n u = f_\Gamma(t,u) \qquad \text{on } \Gamma \coloneqq \partial\Omega. \tag{5.1b}$$

Here, Δ_Γ denotes the Laplace–Beltrami operator, i.e., the Laplacian along the boundary; cf. [GT01, Ch. 16.1]. For the inhomogeneities, we set

$$f_\Omega(t,u) = \sin(t), \qquad f_\Gamma(t,u) = -u^3 + u.$$

To show well-posedness of (5.1), the system can be written in an abstract form; see [HK20, HL20]. To obtain the PDAE structure (2.1), we consider an alternative formulation as introduced in [Alt19, Alt23]. For this, we introduce a second variable $p \coloneqq u|_\Gamma$, which is only defined on the boundary. With this, we can interpret (5.1a) as a dynamic equation for u and (5.1b) as a dynamic equation for p, which are coupled through the mentioned connection. We further define

$$\mathcal{V} \coloneqq H^1(\Omega) \times H^1(\Gamma), \qquad \mathcal{Q} \coloneqq H^{-1/2}(\Gamma).$$

As shown in [Alt23], the weak formulation of (5.1) can be written as

$$\begin{bmatrix} \ddot{u} \\ \ddot{p} \end{bmatrix} + \begin{bmatrix} \mathcal{A}_\Omega & \\ & \mathcal{A}_\Gamma \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} + \mathcal{B}^* \lambda = \begin{bmatrix} f_\Omega(t,u) \\ f_\Gamma(t,p) \end{bmatrix} \qquad \text{in } \mathcal{V}^*, \tag{5.2a}$$

$$\mathcal{B} \begin{bmatrix} u \\ p \end{bmatrix} = 0 \qquad \text{in } \mathcal{Q}^*, \tag{5.2b}$$

which has exactly the structure of (2.1). Therein, the constraint operator is given by $\mathcal{B}\colon \mathcal{V} \to \mathcal{Q}^* = H^{1/2}(\Gamma)$, $\mathcal{B}(u,p) \coloneqq p - \operatorname{trace} u$ with the usual trace operator. The differential operators are given by $\mathcal{A}_\Omega\colon H^1(\Omega) \to [H^1(\Omega)]^*$ and $\mathcal{A}_\Gamma\colon H^1(\Gamma) \to [H^1(\Gamma)]^*$,

$$\langle \mathcal{A}_\Omega u, v\rangle = \int_\Omega \nabla u \cdot \nabla v + uv \,\mathrm{d}x, \qquad \langle \mathcal{A}_\Gamma p, q\rangle = \int_\Gamma \nabla_\Gamma p \cdot \nabla_\Gamma q \,\mathrm{d}x.$$

Note that, in contrast to (5.1), system (5.2) contains two additional variables, namely p for the trace values and a Lagrange multiplier $\lambda\colon [0,T] \to \mathcal{Q}$ which enforces the constraint.

The numerical tests are carried out on the unit disc, i.e., we consider

$$\Omega = \big\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1\big\}.$$

The spatial meshes are generated by DistMesh [PS04] and we use piecewise linear finite element approximations. Moreover, we set $T = 1$ and, as initial condition,

$$u(0) = u^0 \coloneqq \exp\big(-20\,((x-1)^2 + y^2)\big), \qquad \dot{u}(0) = \dot{u}^0 \coloneqq 0.$$

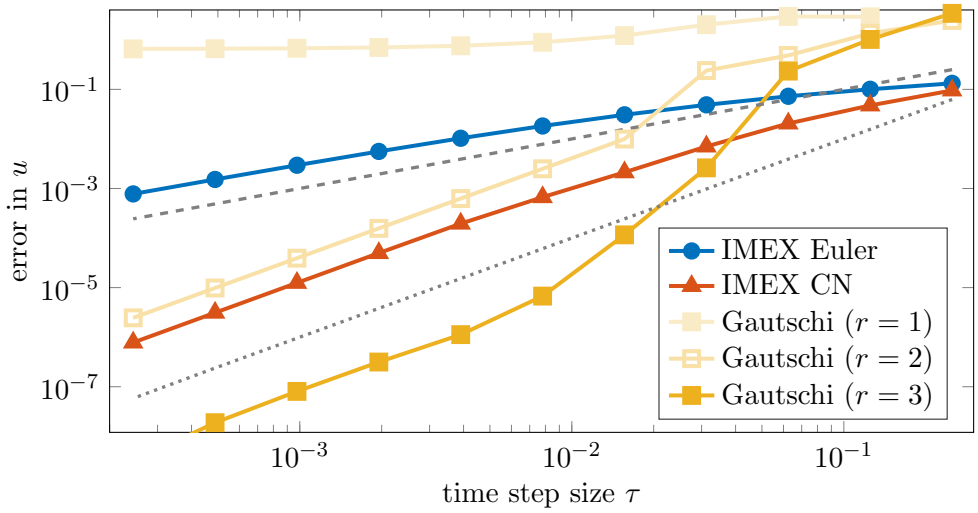


FIGURE 5.1. Convergence history for different time stepping schemes. Plot shows error in u, measured in the $L^\infty(0,T;L^2(\Omega))$-norm. The gray lines indicate orders 1 (dashed) and 2 (dotted).

Accordingly, $p(0)$ and $\dot{p}(0)$ are defined in a consistent manner. Within the experiment, we have used six spatial level refinements, leading to a mesh size of around $h \approx 0.0672$. The reference solution is computed on the same spatial grid by the Gautschi-type integrator with Krylov subspace dimension $r = 10$ and a time step size of $\tau_{\text{ref}} = 2^{-13}$.

The obtained convergence history is presented in Figure 5.1. The displayed lines show the errors in u measured in the $L^\infty(0,T;L^2(\Omega))$-norm. One can see the first-order convergence of the IMEX Euler scheme, i.e., of the implicit Euler scheme where the nonlinearity is treated explicitly, i.e.,

$$\begin{aligned} -\tau\, w^{n+1} \quad + u^{n+1} \qquad\qquad\quad &= u^n, \\ w^{n+1} + \tau \mathcal{A} u^{n+1} + \tau \mathcal{B}^* \lambda^{n+1} &= w^n + \tau f^n, \\ \mathcal{B} u^{n+1} \qquad\qquad\quad &= g^{n+1}. \end{aligned}$$

We further observe the proven second-order accuracy of the IMEX scheme (3.2). For the Gautschi-type integrator (4.10), we present three versions which differ by the dimension of the Krylov subspaces used for the computation of the matrix cosine. For dimension $r = 1$, the error stagnates but already for $r = 2$ the errors show second-order convergence. The results improve once more for $r = 3$ but remain unchanged for larger dimensions $r > 3$. If one considers the $L^\infty(0,T;H^1(\Omega))$-norm, in which the IMEX Crank–Nicolson error was estimated in Theorem 3.1, we observe the same orders but with larger error constants.

The code is available at `https://github.com/altmannr/codeForGautschiCNPaper`.

6. CONCLUSION

In this paper, we have introduced and analyzed two novel time integration schemes for semi-linear wave-type systems restricted by a linear constraint. More precisely, we have considered an implicit–explicit version of the Crank–Nicolson scheme as well as an

integrator of Gautschi-type. For the second-order convergence proofs, we combine well-known results for the dynamical part with certain saddle point problems in order to include the constraints.

We emphasize that the error analysis of the IMEX method can be easily adapted to the (original) Crank–Nicolson scheme. This, however, does not seem computationally attractive due to the nonlinear saddle problems that have to be solved in each time step. In addition, a combination of the two analyses seems to be a good starting point to derive error bounds also for more general exponential integrators applied to constrained systems.

Acknowledgments

R. Altmann acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the project 446856041. B. Dörich acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Project-ID 258734477 – SFB 1173. Finally, C. Zimmer acknowledges support by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 865751 – RandomMultiScales).

References

[ACM99] G. Akrivis, M. Crouzeix, and C. Makridakis. Implicit-explicit multistep methods for quasilinear parabolic equations. *Numer. Math.*, 82(4):521–541, 1999.

[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. *Applied Mathematics Letters*, 90:202–208, 2019.

[Alt23] R. Altmann. Splitting schemes for the semi-linear wave equation with dynamic boundary conditions. *Comput. Math. Appl.*, 151:12–20, 2023.

[AMHR15] A. H. Al-Mohy, N. J. Higham, and S. D. Relton. New algorithms for computing the matrix sine and cosine separately or simultaneously. *SIAM J. Sci. Comput.*, 37(1):A456–A487, 2015.

[ARS97] U. M. Ascher, S. J. Ruuth, and R. J. Spiteri. Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. *Appl. Numer. Math.*, 25(2-3):151–167, 1997.

[ARW95] U. M. Ascher, S. J. Ruuth, and B. T. R. Wetton. Implicit-explicit methods for time-dependent partial differential equations. *SIAM J. Numer. Anal.*, 32(3):797–823, 1995.

[AZ18] R. Altmann and C. Zimmer. Runge-Kutta methods for linear semi-explicit operator differential-algebraic equations. *Math. Comp.*, 87(309):149–174, 2018.

[AZ20] R. Altmann and C. Zimmer. Exponential integrators for semi-linear parabolic problems with linear constraints. In T. Reis, S. Grundel, and S. Schöps, editors, *Progress in Differential-Algebraic Equations II*, pages 137–164. Springer International Publishing, Cham, 2020.

[Bos07] S. Boscarino. Error analysis of IMEX Runge-Kutta methods derived from differential-algebraic systems. *SIAM J. Numer. Anal.*, 45(4):1600–1621, 2007.

[CKN24] A. Cortinovis, D. Kressner, and Y. Nakatsukasa. Speeding up Krylov subspace methods for computing $f(a)b$ via randomization. *SIAM J. Matrix Anal. Appl.*, 45(1):619–633, 2024.

[EE06] 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.

[EM13] E. Emmrich and V. Mehrmann. Operator differential-algebraic equations arising in fluid dynamics. *Comput. Methods Appl. Math.*, 13(4):443–470, 2013.

[FHV97] J. Frank, W. Hundsdorfer, and J. G. Verwer. On the stability of implicit-explicit linear multistep methods. *Appl. Numer. Math.*, 25(2-3):193–205, 1997.

[GASSS98] B. García-Archilla, J. M. Sanz-Serna, and R. D. Skeel. Long-time-step methods for oscillatory differential equations. *Siam J. Sci. Comput.*, 20(3):930–963, 1998.

[Gau61] W. Gautschi. Numerical integration of ordinary differential equations based on trigonometric polynomials. *Numer. Math.*, 3(1):381–397, 1961.

[Gri02] V. Grimm. *Exponentielle Integratoren als Lange-Zeitschritt-Verfahren für oszillatorische Differentialgleichungen zweiter Ordnung*. PhD thesis, Heinrich-Heine-Universität Düsseldorf, 2002.

[Gri05] V. Grimm. On error bounds for the Gautschi-type exponential integrator applied to oscillatory second-order differential equations. *Numer. Math.*, 100(1):71–89, 2005.

[GT01] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, Berlin, 2001.

[HAM10] N. J. Higham and A. H. Al-Mohy. Computing matrix functions. *Acta Numer.*, 19:159–208, 2010.

[Hig08] N. J. Higham. *Functions of Matrices*. Society for Industrial and Applied Mathematics, Philadelphia, 2008.

[HK20] D. Hipp and B. Kovács. Finite element error analysis of wave equations with dynamic boundary conditions: L^2 estimates. *IMA J. Numer. Anal.*, 41(1):638–728, 2020.

[HL99] M. Hochbruck and C. Lubich. A Gautschi-type method for oscillatory second-order differential equations. *Numer. Math.*, 83:403–426, 1999.

[HL20] M. Hochbruck and J. Leibold. Finite element discretization of semilinear acoustic wave equations with kinetic boundary conditions. *Electron. T. Numer. Ana.*, 53:522–540, 2020.

[HL21] M. Hochbruck and J. Leibold. An implicit-explicit time discretization scheme for second-order semilinear wave equations with application to dynamic boundary conditions. *Numer. Math.*, 147(4):869–899, 2021.

[HO10] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numer.*, 19:209–286, 2010.

[HPS+15] M. Hochbruck, T. Pažur, A. Schulz, E. Thawinan, and C. Wieners. Efficient time integration for discontinuous Galerkin approximations of linear wave equations. *Z. Angew. Math. Mech. (ZAMM)*, 95(3):237–259, 2015.

[HR07] W. Hundsdorfer and S. J. Ruuth. IMEX extensions of linear multistep methods with general monotonicity and boundedness properties. *J. Comput. Phys.*, 225(2):2016–2042, 2007.

[HW96] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. Springer-Verlag, Berlin, second edition, 1996.

[LLT16] W. Layton, Y. Li, and C. Trenchea. Recent developments in IMEX methods with time filters for systems of evolution equations. *J. Comput. Appl. Math.*, 299:50–67, 2016.

[LM72] J.-L. Lions and E. Magenes. *Non-homogeneous Boundary Value Problems and Applications. Vol. I*. Springer-Verlag, New York-Heidelberg, 1972.

[LMT13] R. Lamour, R. März, and C. Tischendorf. *Differential-Algebraic Equations: A Projector Based Analysis*. Springer-Verlag, Berlin, Heidelberg, 2013.

[LT12] W. Layton and C. Trenchea. Stability of two IMEX methods, CNLF and BDF2-AB2, for uncoupling systems of evolution equations. *Appl. Numer. Math.*, 62(2):112–120, 2012.

[MV78] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Rev.*, 20(4):801–836, 1978.

[Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.

[PS04] P.-O. Persson and G. Strang. A simple mesh generator in MATLAB. *SIAM Rev.*, 46:329–345, 2004.

[Saa92] Y. Saad. Analysis of some Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 29(1):209–228, 1992.

[Sim98] B. Simeon. DAEs and PDEs in elastic multibody systems. *Numer. Algorithms*, 19:235–246, 1998.

[Sim13] B. Simeon. *Computational flexible multibody dynamics. A differential-algebraic approach*. Differential-Algebraic Equations Forum. Springer-Verlag, Berlin, 2013.

[Zei90] E. Zeidler. *Nonlinear Functional Analysis and its Applications IIa: Linear Monotone Operators*. Springer-Verlag, New York, 1990.

[Zim21] C. Zimmer. *Temporal Discretization of Constrained Partial Differential Equations*. PhD thesis, Technische Universität Berlin, 2021.

† Institute of Analysis and Numerics, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany

* Institute for Applied and Numerical Mathematics, Karlsruhe Institute of Technology, Englerstrasse 2, 76149 Karlsruhe, Germany

‡ Department of Mathematics, University of Augsburg, Universitätsstr. 12a, 86159 Augsburg, Germany

Email address: robert.altmann@ovgu.de, benjamin.doerich@kit.edu, dr.christoph.zimmer@gmail.com