



Runge–Kutta approximation for C_0 -semigroups in the graph norm with applications to time domain boundary integral equations

Alexander Rieder¹ · Francisco-Javier Sayas² · Jens Markus Melenk³ 

Received: 3 March 2020 / Accepted: 13 October 2020 / Published online: 24 November 2020
© The Author(s) 2020

Abstract

We consider the approximation of an abstract evolution problem with inhomogeneous side constraint using A -stable Runge–Kutta methods. We derive a priori estimates in norms other than the underlying Banach space. Most notably, we derive estimates in the graph norm of the generator. These results are used to study convolution quadrature based discretizations of a wave scattering and a heat conduction problem.

1 Introduction

Many time dependent partial differential equations can be conveniently described in the language of strongly continuous semigroups. In this language, these initial boundary value problems resemble systems of ordinary differential equations, which suggests that they are amendable to the standard discretization schemes of multistep or Runge–Kutta type. Unlike the ODE case, one needs to pay special attention to the boundary conditions imposed by the generator of the semigroup. This, in most cases, leads to a reduction of order phenomenon, meaning that the convergence rates are (mainly) determined by the stage order of the Runge–Kutta method instead of the classical order. The a priori convergence of Runge–Kutta methods for semigroups has been extensively studied in the literature. Starting with the early works [8, 9], it has been established that conditions of the form $u(t) \in \text{dom}(A^\mu)$, where A is the generator of the semigroup and dom denotes the domain of an operator, determine the convergence rates. In [25], this has been generalized to the case of non-integer $\mu \geq 1$ using the theory of interpolation spaces. Finally, in [1], the

This article is part of the topical collection “Waves 2019 – invited papers” edited by Manfred Kaltenbacher and Jens Markus Melenk.

✉ Jens Markus Melenk
melenk@tuwien.ac.at

Alexander Rieder
alexander.rieder@univie.ac.at

Extended author information available on the last page of the article

case of $\mu \in [0, 1]$ was addressed, which is the case needed for PDEs with inhomogeneous boundary conditions. We point out that in the case of analytic semigroups, Lubich and Ostermann [19] had already established convergence also for inhomogeneous boundary conditions. All of these works focus on establishing convergence rates with respect to the norm of the underlying Banach space. In many applications one needs to establish convergence with respect to other norms, for example, in order to be able to bound boundary traces of the solution. Most notably, one might be interested in convergence of $A_\star u$, where A_\star is an extension of the generator that disregards boundary conditions. If u is assumed to be in $\text{dom}(A)$, we get $A_\star u = Au$ and the convergence result can be easily established by using the fact that the time evolution commutes with the generator of the underlying semigroup (both in the continuous and discrete settings). If the boundary conditions are inhomogeneous, such a strategy cannot be pursued. It is the goal of this paper to establish convergence results for $A_\star u$ also for the case $u(t) \in \text{dom}(A^\mu)$ for $\mu \in [0, 1]$, again using the theory of interpolation spaces.

Similarly it is sometimes useful to compute discrete integrals of the time evolution by reusing the same Runge–Kutta method. Also in this case, we establish rigorous convergence rates.

Our interest in such estimates originally arose from the study of time domain boundary integral equations and their discretization using convolution quadrature (CQ). It has already been noticed in the early works (see e.g. [19]) that such discretizations have a strong relation to the Runge–Kutta approximation of the underlying semigroup. This approach of studying TDBIEs in a strictly time-domain way has recently garnered a lot of interest, see [3, 13, 15] and the monograph [31], as it potentially allows sharper bounds than the more standard Laplace domain based approach. Similar techniques have even been extended to the case of certain nonlinear problems in [4]. This paper can be seen as our latest addition to this effort. While the convergence rates provided by the Laplace-domain approach in [2] and the results in this current paper are essentially the same, the present new approach provides better insight into the dependence on the end-time of the computation (quadratic vs. general unknown polynomial behavior). This suggests that the present approach might be better suited for analyzing long term computations. It also fits more naturally with the time-domain analysis of the continuous problem and space discretization, as for example presented in [13].

The paper is structured as follows. Section 2 introduces the abstract setting and fixes notation, most notably for working with Runge–Kutta methods. Section 3 then contains the main estimates. Starting by summarizing known results from [1] in Sect. 3.1, we then formulate the main new results of this article in Sect. 3.2. After proving some preparatory lemmas related to Runge–Kutta methods in Sects. 4 and 5, we provide the proofs of the main estimates in Sect. 6. In Sect. 7, we show how our setting simplifies if we restrict our view to a subclass of admissible operators. In Sect. 8, to showcase how the theory developed in this paper is useful for this class of problems, we consider a simple exterior scattering problem in Sect. 8.3 and a heat transmission problem in Sect. 8.5. We note that Sect. 8.3 showcases the need for the bound on the discrete integral of the result, whereas Sect. 8.5 was chosen because, in order to bound the main quantity of interest on the boundary, we need to apply a trace theorem. This necessitates the use of the graph norm estimate.

2 Problem setting

We start by fixing the general setting used for the rest of the paper, first with respect to the equation to be solved and then with respect to its discretization.

2.1 Operator equation, functional calculus, and Sobolev towers

Assumption 2.1 We are given:

- (a) A closed linear operator $A_\star : \text{dom}(A_\star) \subset \mathcal{X} \rightarrow \mathcal{X}$ in a Banach space \mathcal{X} .
- (b) A bounded linear operator $B : \text{dom}(A_\star) \rightarrow \mathcal{M}$ with another Banach space \mathcal{M} .

We assume that $A := A_\star|_{\ker B}$ generates a C_0 -semigroup and that B admits a bounded right inverse \mathcal{E} such that $\text{range } \mathcal{E} \subset \ker(I - A_\star)$, where $I : \mathcal{X} \rightarrow \mathcal{X}$ is the identity operator.

We are given $u_0 \in \text{dom}(A)$ and data functions $F \in C^1([0, T], \mathcal{X})$, $\Xi \in C^1([0, T], \mathcal{M})$, and we consider the problem: find $u \in C^1([0, T], \mathcal{X})$ such that

$$\dot{u}(t) = A_\star u(t) + F(t), \quad t > 0, \tag{2.1a}$$

$$Bu(t) = \Xi(t), \quad t > 0, \tag{2.1b}$$

$$u(0) = u_0. \tag{2.1c}$$

For conditions on the well-posedness of this problem, see [13]. We start by recalling the following consequence of the Hille-Yosida theorem.

Proposition 2.1 ([26, Corollary 3.8]) *If A is the generator of a C_0 -semigroup on a Banach space \mathcal{X} , then there exist constants $\omega \geq 0$ and $M \geq 1$ such that the spectrum $\sigma(A)$ of A satisfies $\sigma(A) \subseteq \{z \in \mathbb{C} : \text{Re } z \leq \omega\}$ and the resolvent satisfies the estimates*

$$\|(A - zI)^{-1}\|_{\mathcal{X} \rightarrow \mathcal{X}} \leq \frac{M}{\text{Re } z - \omega} \quad \forall z \text{ s.t. } \text{Re } z > \omega. \tag{2.2}$$

When working with Runge–Kutta methods, it is useful to use a calculus that allows one to apply rational functions to (unbounded) operators, as long as the poles of the function are compatible with the spectrum of the operator.

Definition 2.2 (Rational functions of operators) *Let q be a rational function that is bounded at infinity. Let A be the set of poles of q , which we can write in the form (note that we allow for some of the factors in the numerator to be constant)*

$$q(z) = c_0 \prod_{i=1}^n \frac{c_i z - 1}{z - \lambda_i} = c_0 \prod_{i=1}^n \left(c_i + \frac{c_i \lambda_i - 1}{z - \lambda_i} \right).$$

If $A : \text{dom}(A) \subset \mathcal{X} \rightarrow \mathcal{X}$ is a linear operator such that $\sigma(A) \cap A = \emptyset$, we define

$$q(A) := c_0(c_1I + (c_1\lambda_1 - 1)(A - \lambda_1I)^{-1}) \cdots (c_nI + (c_n\lambda_n - 1)(A - \lambda_nI)^{-1}). \tag{2.3}$$

It is easy to see that different reorderings of the factors in the numerator and denominator of q produce the same result and that each factor in the definition of $q(A)$ is a bounded linear operator in \mathcal{X} since $\lambda_i \notin \sigma(A)$. The bounded linear operator $q(A) : \mathcal{X} \rightarrow \mathcal{X}$ satisfies

$$\|q(A)\|_{\mathcal{X} \rightarrow \mathcal{X}} \leq C_q \left(1 + \left(\max_{\lambda \in \Lambda} \|(A - \lambda I)^{-1}\|_{\mathcal{X} \rightarrow \mathcal{X}} \right)^n \right). \tag{2.4}$$

The error estimates of this paper use the theory of interpolation spaces. For Banach spaces $\mathcal{X}_1 \subset \mathcal{X}_0$ with continuous embedding and $\mu \in (0, 1)$, we define the space $[\mathcal{X}_0, \mathcal{X}_1]_{\mu, \infty}$ using real interpolation with the following norm:

$$\|u\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\mu, \infty}} := \operatorname{ess\,sup}_{t > 0} \left(t^{-\mu} \inf_{v \in \mathcal{X}_1} \left[\|u - v\|_{\mathcal{X}_0} + t\|v\|_{\mathcal{X}_1} \right] \right). \tag{2.5}$$

We will not go into details of the definitions and instead refer to [35, 36] or [22, Appendix B]. For simplicity of notation we often drop the second parameter ∞ and just write $[\mathcal{X}_0, \mathcal{X}_1]_{\mu}$.

The most important property is the following: a bounded linear operator $T : \mathcal{X}_0 \rightarrow \mathcal{Y}_0$ and $\mathcal{X}_1 \rightarrow \mathcal{Y}_1$ with $\mathcal{X}_1 \subseteq \mathcal{X}_0$ and $\mathcal{Y}_1 \subseteq \mathcal{Y}_0$ is also a bounded operator mapping $[\mathcal{X}_0, \mathcal{X}_1]_{\mu} \rightarrow [\mathcal{Y}_0, \mathcal{Y}_1]_{\mu}$ with the following norm bound

$$\|T\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\mu} \rightarrow [\mathcal{Y}_0, \mathcal{Y}_1]_{\mu}} \leq \|T\|_{\mathcal{X}_0 \rightarrow \mathcal{Y}_0}^{1-\mu} \|T\|_{\mathcal{X}_1 \rightarrow \mathcal{Y}_1}^{\mu}. \tag{2.6}$$

We also note that for $\mu_1 \leq \mu_2$, the spaces are nested, i.e., $[\mathcal{X}_0, \mathcal{X}_1]_{\mu_2} \subseteq [\mathcal{X}_0, \mathcal{X}_1]_{\mu_1}$ with continuous embedding. For notational convenience we write $[\mathcal{X}_0, \mathcal{X}_1]_0 := \mathcal{X}_0$ and $[\mathcal{X}_0, \mathcal{X}_1]_1 := \mathcal{X}_1$. We will be interested in a collection of spaces defined by interpolating the domains of the powers of the operator A . The details of this construction can be found, for example in [11].

Definition 2.3 (Sobolev towers) *Let A be a closed operator on a Banach space \mathcal{X} . For $\mu \in \mathbb{N}_0$, we define the following spaces $\mathcal{X}_0 := \operatorname{dom}(A^0) := \mathcal{X}$ and $\mathcal{X}_{\mu} := \operatorname{dom}(A^{\mu})$, equipped with the following norm*

$$\|u\|_{\mathcal{X}_{\mu}} := \sum_{j=0}^{\mu} \|A^j u\|_{\mathcal{X}}.$$

For $\mu \in [0, \infty)$, we define $\mathcal{X}_{\mu} := [\mathcal{X}_{[\mu]}, \mathcal{X}_{[\mu]+1}]_{\mu - [\mu]}$ by interpolation.

We sometimes consider $\operatorname{dom}(A)$ as a Banach space. It is to be understood carrying the graph norm, same as \mathcal{X}_1 .

2.2 Runge–Kutta approximation and discrete stage derivative

An m -stage Runge–Kutta method is given by its Butcher tableau, characterized by $Q \in \mathbb{R}^{m \times m}$ and $\mathbf{b}, \mathbf{c} \in \mathbb{R}^m$. The Runge–Kutta approximation of the problem (2.1a–2.1c) starts at

$u_0^k := u_0$ and then computes for $n \geq 0$ the stage vector $U_n^k \in \mathcal{X}^m$ and the step approximation $u_{n+1}^k \in \mathcal{X}$ by solving

$$U_n^k = \mathbf{1} u_n^k + k(Q \otimes A_\star)U_n^k + kQF(t_n + k \mathbf{c}), \tag{2.7a}$$

$$(\mathcal{I} \otimes B)U_n^k = \Xi(t_n + k \mathbf{c}), \tag{2.7b}$$

$$u_{n+1}^k = u_n^k + k(\mathbf{b}^\top \otimes A_\star)U_n^k + k\mathbf{b}^\top F(t_n + k \mathbf{c}). \tag{2.7c}$$

We have used the following notation (the spaces \mathcal{Y} and \mathcal{Z} are generic):

(a) For a function $G : [0, T] \rightarrow \mathcal{Y}$ we write

$$G(t_n + k\mathbf{c}) := (G(t_n + kc_1), \dots, G(t_n + kc_m))^\top \in \mathcal{Y}^m.$$

(b) For a matrix $\mathcal{S} \in \mathbb{R}^{m \times m}$ and an operator $C : \mathcal{Y} \rightarrow \mathcal{Z}$ we write

$$\mathcal{S} \otimes C := \begin{bmatrix} \mathcal{S}_{11}C & \cdots & \mathcal{S}_{1m}C \\ \vdots & & \vdots \\ \mathcal{S}_{m1}C & \cdots & \mathcal{S}_{mm}C \end{bmatrix} : \mathcal{Y}^m \rightarrow \mathcal{Z}^m.$$

(c) For the vector \mathbf{b} and an operator $C : \mathcal{Y} \rightarrow \mathcal{Z}$ we write

$$\mathbf{b}^\top \otimes C := [b_1C \ \cdots \ b_mC] : \mathcal{Y}^m \rightarrow \mathcal{Z}.$$

(d) \mathcal{I} is the $m \times m$ identity matrix, and $\mathbf{1} = (1, \dots, 1)^\top$.

(e) We admit shortened expressions such as

$$QF(t_n + k\mathbf{c}) := (Q \otimes I)F(t_n + k\mathbf{c}),$$

$$\mathbf{1} u := (\mathbf{1} \otimes I)u,$$

$$\mathbf{b}^\top F(t_n + k\mathbf{c}) := (\mathbf{b}^\top \otimes I)F(t_n + k\mathbf{c}).$$

The following lemma involving inversion of matrices of operators associated to an operator can be proved by taking the Jordan canonical form of the matrix \mathcal{S} .

Lemma 2.4 *If $A : \text{dom}(A) \subset \mathcal{X} \rightarrow \mathcal{X}$ is a linear operator on a Banach space \mathcal{X} and $\mathcal{S} \in \mathbb{C}^{m \times m}$ satisfies $\sigma(A) \cap \sigma(\mathcal{S}) = \emptyset$, then*

$$\mathcal{I} \otimes A - \mathcal{S} \otimes I : (\text{dom}(A))^m \rightarrow \mathcal{X}^m,$$

is invertible. Furthermore, there exists a constant $C_{\mathcal{S}}$, depending only on \mathcal{S} , such that

$$\|(\mathcal{I} \otimes A - \mathcal{S} \otimes I)^{-1}\|_{\mathcal{X}^m \rightarrow \mathcal{X}^m} \leq C_{\mathcal{S}} \left[1 + \max_{\mu \in \sigma(\mathcal{S})} \|(A - \mu I)^{-1}\|_{\mathcal{X} \rightarrow \mathcal{X}} \right]^m.$$

Under Assumption 2.1, the internal stage computation in the RK method can be decomposed in the following form:

$$Y_n^k := (\mathcal{I} \otimes \mathcal{E})\Xi(t_n + k \mathbf{c}), \tag{2.8a}$$

$$Z_n^k - k(\mathcal{Q} \otimes A)Z_n^k = \mathbf{1}u_n^k - Y_n^k + k\mathcal{Q}(Y_n^k + F(t_n + k \mathbf{c})), \tag{2.8b}$$

$$U_n^k := Y_n^k + Z_n^k. \tag{2.8c}$$

In (2.8b) we look for $Z_n^k \in (\text{dom}(A))^m$.

The stability function of the Runge–Kutta method is the rational function $r(z) := 1 + z\mathbf{b}^\top(I - z\mathcal{Q})^{-1}\mathbf{1}$. We will not consider the full class of Runge–Kutta methods, but will restrict our considerations to those satisfying the following Assumptions:

Assumption 2.II

- (i) The matrix \mathcal{Q} is invertible.
- (ii) The stability function r does not have poles in $\{z : \text{Re } z < 0\}$, and $|r(it)| \leq 1$ for all $t \in \mathbb{R}$ (i.e., the method is A -stable). Equivalently, $|r(z)| < 1$ for all z with negative real part.

We note that Assumption 2.II (i) implies that the following limit exists

$$\lim_{z \rightarrow \infty} r(z) = 1 - \mathbf{b}^\top \mathcal{Q}^{-1} \mathbf{1} =: r(\infty).$$

Assumption 2.II (ii) implies that

$$\sigma(\mathcal{Q}) \subset \mathbb{C}_+ := \{z \in \mathbb{C} : \text{Re } z > 0\},$$

and that r is a rational function with poles only in \mathbb{C}_+ and bounded at infinity.

The computation of the internal stages in the numerical approximation (2.7a–2.7c) requires the inversion of

$$\mathcal{I} \otimes I - k(\mathcal{Q} \otimes A) = (\mathcal{Q} \otimes I)(\mathcal{Q}^{-1} \otimes I - \mathcal{I} \otimes (kA)),$$

as can be seen from the equivalent form (2.8a–2.8c).

If A is the infinitesimal generator of a C_0 -semigroup and ω and M are given by Proposition 2.1 and if we choose (recall that $\sigma(\mathcal{Q}) \subset \mathbb{C}_+$)

$$k_0 < \omega^{-1}d_0, \quad d_0 := \min\{\text{Re } \lambda : \lambda \in \sigma(\mathcal{Q}^{-1})\}, \tag{2.9}$$

then the RK method can be applied for any $0 < k \leq k_0$. By Proposition 2.1 and Lemma 2.4, it follows that

$$\|(\mathcal{I} \otimes I - k(\mathcal{Q} \otimes A))^{-1}\|_{\mathcal{X}^m \rightarrow \mathcal{X}^m} \leq C_{\mathcal{Q}} \frac{M}{d_0 - k_0\omega}, \quad \forall k \leq k_0. \tag{2.10}$$

Using Definition 2.2, we can define $r(kA)$ for an RK method satisfying Assumption 2.II and $k \leq k_0$ satisfying (2.9). We then define

$$\rho_k(T) := \sup_{0 \leq nk \leq T} \|r(kA)^n\|_{\mathcal{X} \rightarrow \mathcal{X}}. \tag{2.11}$$

This quantity is relevant for the study of the error propagation in the Runge–Kutta method.

Given an RK method, we consider the following matrix-valued rational function

$$\delta(z) := \left(Q + \frac{z}{1-z} \mathbf{1b}^\top \right)^{-1} = Q^{-1} - \frac{z}{1-r(\infty)z} Q^{-1} \mathbf{1b}^\top Q^{-1}. \tag{2.12}$$

(The verification that these two formulas correspond to the same matrix is simple by using the Sherman–Morrison–Woodbury formula.) This matrix is related to the discrete differentiation process associated to an RK method satisfying Assumption 2.II: on the one hand $k^{-1}\delta(z)$ is the discrete symbol associated to the discrete operational calculus built with the RK method [19]; on the other hand, a direct interpretation of this symbol is possible using the Z-transformation (see [14, Sect. 6]). Given a sequence $U := \{U_n\}$ (tagged from $n \geq 0$) on a space, its Z-transform \mathcal{Z} is the formal series

$$\widehat{U}(z) := \sum_{n=0}^{\infty} U_n z^n.$$

For a detailed treatment on formal power series, see [12].

Definition 2.5 Let $U := \{U_n\}$ and $V := \{V_n\}$ be two sequences in \mathcal{X}^m and let \widehat{U} and \widehat{V} be their respective Z-transforms. If

$$k^{-1}\delta(z)\widehat{U}(z) = \widehat{V}(z),$$

we write

$$\partial^k U = V, \quad U = (\partial^k)^{-1} V.$$

The above definition is consistent with the RK discrete operational calculus of Lubich and Ostermann, see Sect. 8.1 and [19]. We now show an explicit form of the computation of ∂^k and its inverse.

Lemma 2.6 *If $U = \{U_n\}$ is a sequence in \mathcal{X}^m , then $X := (\partial^k)^{-1}U$ can be computed with the recurrence*

$$\begin{aligned} x_0 &:= 0, & X_n &:= \mathbf{1}x_n + kQ U_n, \\ x_{n+1} &:= x_n + k\mathbf{b}^\top U_n = r(\infty)x_n + \mathbf{b}^\top Q^{-1} X_n, \end{aligned} \tag{2.13}$$

and $V := \partial^k U$ can be computed with the inverse recurrence

$$\begin{aligned} u_0 &:= 0, & V_n &:= k^{-1}Q^{-1}(U_n - \mathbf{1}u_n), \\ u_{n+1} &:= u_n + k\mathbf{b}^\top V_n = r(\infty)u_n + \mathbf{b}^\top Q^{-1} U_n. \end{aligned} \tag{2.14}$$

Proof The proof of (2.13) is a simple exercise in Z-transforms, while (2.14) follows from (2.13) by writing U_n in terms of X_n (and changing names to the sequences). \square

The first result of Lemma 2.6 expresses the fact that if we apply the RK method to the equation

$$\dot{x}(t) = u(t), \quad x(0) = 0, \quad \text{i.e.,} \quad x(t) = \int_0^t u(\tau) d\tau,$$

and $X := \{X_n\}$ is the sequence of vectors of internal stages, then $X = (\partial^k)^{-1}U$, where $U_n := u(t_n + k\mathbf{c})$.

Finally we note that we call a Runge–Kutta method stiffly accurate, if it satisfies $\mathbf{b}^\top \mathcal{Q}^{-1} = \mathbf{e}_m^\top := (0, \dots, 0, 1)$. Stiffly accurate methods satisfy [we use that $\mathcal{Q}\mathbf{1} = \mathbf{c}$, see (5.1)]

$$c_m = \mathbf{b}^\top \mathcal{Q}^{-1} \mathbf{c} = \mathbf{b}^\top \mathcal{Q}^{-1} \mathcal{Q}\mathbf{1} = \mathbf{b}^\top \mathbf{1} = 1, \tag{2.15}$$

and $r(\infty) = 0$.

For stiffly accurate methods, taking the discrete derivative of a stage vector consisting of samples taken from a continuous function is particularly simple:

Lemma 2.7 *Let $t \mapsto F(t)$ be a continuous function with $F(0) = 0$. For stiffly accurate RK methods the sequence $G := \partial^k F$ with $F_n = F(t_n + k\mathbf{c})$ satisfies*

$$G_n = k^{-1} \mathcal{Q}^{-1} (F(t_n + k\mathbf{c}) - \mathbf{1}F(t_n)).$$

Proof For stiffly accurate methods we have $r(\infty) = 0$ and therefore

$$\delta(z) = \mathcal{Q}^{-1} - z\mathcal{Q}^{-1} \mathbf{1}\mathbf{b}^\top \mathcal{Q}^{-1} = \mathcal{Q}^{-1} - z\mathcal{Q}^{-1} \mathbf{1}\mathbf{e}_m^\top.$$

However, since $c_m = 1$, we have $\mathbf{e}_m^\top F(t_{n-1} + k\mathbf{c}) = F(t_{n-1} + kc_m) = F(t_n)$, which proves the result. □

We also make the following optional assumption, which allows us to increase the convergence order in some cases.

Assumption 2.III For all $t \in \mathbb{R}, t \neq 0$ the stability function satisfies $|r(it)| < 1$ and $r(\infty) < 1$.

3 Error estimates

We are now in a position to formulate the main results of this article and put them into context with previous results, most notably from [1].

To simplify notation, we will write for $v \in \mathcal{C}([0, T]; \mathcal{X}_\mu)$ with $\mu \geq 0$,

$$\|v\|_{T,\mu} := \max_{\tau \in [0, T]} \|v(\tau)\|_{\mathcal{X}_\mu}.$$

For functions $f : [0, T] \rightarrow \mathcal{Y}$, we will write $(\partial^{-1}f)(t) := \int_0^t f(\tau) d\tau$, where \mathcal{Y} denotes a generic Banach space.

3.1 The estimates of Alonso-Mallo and Palencia

The following two propositions summarize the results of Alonso-Mallo and Palencia [1], rewritten with the notation of the present paper. The ‘proofs’ which we provide clarify how notation needs to be adapted and how the hypotheses of the main results of [1] are satisfied in our context.

Proposition 3.1 ([1, Theorem 1]) *Let Assumption 2.I hold and assume that the exact solution u satisfies $u \in C^{p+1}([0, T], \mathcal{X}_\mu)$ for some $\mu \geq 0$. Let $\{u_n^k\}$ denote the Runge–Kutta approximation from (2.7a–2.7c). Then there exist constants $k_0 > 0$ and $C > 0$ such that for $0 < k \leq k_0$ and $0 < nk \leq T$ the following estimate holds:*

$$\|u(t_n) - u_n^k\|_{\mathcal{X}} \leq CT\rho_k(T)k^{\min\{q+\mu,p\}} \left(\sum_{\ell=q+1}^p \|u^{(\ell)}\|_{T,\mu} + \|u^{(p+1)}\|_{T,0} \right). \tag{3.1}$$

The constant C depends on the Runge–Kutta method, μ , and the constants M and ω from (2.2). The constant k_0 depends only on ω and the Runge–Kutta method.

Proof We only remark on the differences in notation. A different definition of interpolation spaces is given in [1], but the proof only relies on estimates of the form (2.6). The choice of k_0 follows from the fact that it is only needed to ensure that $(I - kQ \otimes A)$ is invertible, see (2.10). The assumption $\mu \leq p - q$ in [1, Theorem 1] can be replaced by using the rate $\min\{p, q + \mu\}$ in (3.1) as the spaces $\mathcal{X}_\mu \subseteq \mathcal{X}_{p-q}$ are nested for $\mu \geq p - q$. We also lowered the regularity requirements on the highest derivative compared to their stated result. The fact that this holds true follows from inspection of the proof. See also Lemma 5.9 for the key ingredient. \square

For certain Runge–Kutta methods, these estimates can be improved:

Proposition 3.2 ([1, Theorem 2]) *Let the assumptions of Proposition 3.1 hold and assume that, in addition, the RK method satisfies Assumption 2.III. Then there exist constants $k_0 > 0$, $C > 0$ such that for $0 < k \leq k_0$ and $0 < nk \leq T$ the following improved estimate holds:*

$$\|u(t_n) - u_n^k\|_{\mathcal{X}} \leq C(1 + T)\rho_k(T)k^{\min\{q+\mu+1,p\}} \sum_{\ell=q+1}^{p+1} \|u^{(\ell)}\|_{T,\mu}. \tag{3.2}$$

The constant C depends on the Runge–Kutta method, μ , and the constants M and ω from (2.2); k_0 depends only on the constant ω and the Runge–Kutta method.

Proof Again, this is just a reformulation of [1, Theorem 2]. We first note that, due to our assumption on $r(\infty)$, we are always in the case $m = 0$ of [1]. Since we assumed that on the imaginary axis $|r(it)| < 1$ for $0 \neq t \in \mathbb{R}$, we directly note that for sufficiently small k_0 , all the zeros of $r(z) - 1$ except $z = 0$ satisfy $\text{Re } z > k_0\omega$. By the resolvent bound (2.2) we can therefore estimate for $k \leq k_0$

$$\|(zI - kA)^{-1}\|_{\mathcal{X} \rightarrow \mathcal{X}} \leq \frac{M}{\operatorname{Re} z - k_0\omega}, \quad \text{if } \operatorname{Re} z \geq k_0\omega,$$

i.e., we have a uniform resolvent bound in the set $Z_{\alpha,\delta}$ in [1]. We also note that we reformulated the convergence rate such that we do not have the restriction $\mu \leq p - q - 1$, since the exceptional cases are already covered by Proposition 3.1. \square

Remark 3.3 The assumption $|r(z)| < 1$ for $\operatorname{Re}(z) \leq 0$ and $r(\infty) \neq 1$ is satisfied by the Radau IIA family of Runge–Kutta methods, but is violated by the Gauss methods, which satisfy $|r(z)| = 1$ on the imaginary axis.

3.2 New results in this article

In this section we present some a priori estimates for the convergence of Runge–Kutta methods when applied to the abstract problem (2.1a–2.1c). These can be seen as a continuation of [1] to the case where the boundary conditions are not given exactly but stem from computing discrete integrals and differentials using the same Runge–Kutta method.

Theorem 3.4 (Integrated estimate) *Let u solve (2.1a–2.1c) with $u_0 = 0$ and assume that for some $\mu \geq 0$ we have*

$$u \in \mathcal{C}^p([0, T]; \mathcal{X}_\mu), \quad \mathcal{E}\Xi, F \in \mathcal{C}^{p-1}([0, T]; \mathcal{X}_\mu) \cap \mathcal{C}^p([0, T]; \mathcal{X}_0).$$

Set $x := \partial^{-1}u$. Let $U^k = \{U_n^k\}$ and let $u^k = \{u_n^k\}$ be the discrete approximation given by (2.7a–2.7c) for a method satisfying Assumption 2.II. If $X^k := (\partial^k)^{-1}U^k$ and we define $x^k = \{x_n^k\}$ with the recurrence

$$x_0^k := 0, \quad x_{n+1}^k := r(\infty)x_n^k + \mathbf{b}^\top \mathcal{Q}^{-1}X_n^k,$$

then there exists a constant $k_0 > 0$ such that for all $k < k_0$ and $n \in \mathbb{N}$ with $nk \leq T$ the following estimate holds:

$$\begin{aligned} \|x(t_n) - x_n^k\|_{\mathcal{X}} &\leq \\ &CT\rho_k(T)k^{\min\{q+\mu+1,p\}} \left[\sum_{\ell=q}^{p-1} \left(\|u^{(\ell)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(\ell)}\|_{T,\mu} + \|F^{(\ell)}\|_{T,\mu} \right) \right. \\ &\quad \left. + \left(\|u^{(p)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(p)}\|_{T,0} + \|F^{(p)}\|_{T,0} \right) \right] \\ &+ CT^2\rho_k(T)k^p \left(\|\mathcal{E}\Xi^{(p)}\|_{T,0} + \|F^{(p)}\|_{T,0} \right). \end{aligned}$$

If Assumption 2.III holds and if we assume the stronger regularities

$$u \in \mathcal{C}^{p+1}([0, T]; \mathcal{X}_\mu), \quad F \in \mathcal{C}^p([0, T]; \mathcal{X}_\mu), \quad \mathcal{E}\Xi \in \mathcal{C}^p([0, T]; \mathcal{X}_\mu),$$

then

$$\begin{aligned} & \|x(t_n) - x_n^k\|_{\mathcal{X}} \leq \\ & C(1 + T)\rho_k(T)k^{\min\{q+\mu+2,p\}} \left[\sum_{\ell=q}^p \|u^{(\ell)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(\ell)}\|_{T,\mu} + \|F^{(\ell)}\|_{T,\mu} + \|u^{(p+1)}\|_{T,\mu} \right] \\ & + CT^2\rho_k(T)k^p \left(\|\mathcal{E}\Xi^{(p)}\|_{T,0} + \|F^{(p)}\|_{T,0} \right). \end{aligned}$$

The constant k_0 depends only on ω from (2.2) and the Runge–Kutta method. If $\omega = 0$ then k_0 can be chosen arbitrarily large. C depends on ω , M from (2.2), the Runge–Kutta method, and μ .

Theorem 3.5 (Differentiated estimate) *Let u solve (2.1a–2.1c) with $u_0 = 0$ and assume $\dot{u}(0) = 0$. Assume that for some $\mu \geq 0$ we have*

$$u \in \mathcal{C}^{p+1}([0, T]; \mathcal{X}_\mu) \cap \mathcal{C}^{p+2}([0, T]; \mathcal{X}_0), \quad \mathcal{E}\Xi, F \in \mathcal{C}^p([0, T]; \mathcal{X}_\mu) \cap \mathcal{C}^{p+1}([0, T]; \mathcal{X}_0),$$

and let $v := \dot{u}$. Let $U^k = \{U_n^k\}$ and $u^k = \{u_n^k\}$ be the discrete approximation given by (2.7a–2.7c) for a stiffly accurate method satisfying Assumption 2.II.

If $V^k := \partial^k U^k$ and $v_n^k = \mathbf{e}_m^\top v_{n-1}^k$, then there exists a constant $k_0 > 0$ such that for all $k < k_0$ and $n \geq 1$ such that $nk \leq T$ the following estimate holds:

$$\begin{aligned} \|v(t_n) - v_n^k\|_{\mathcal{X}} + \|A_\star(u(t_n) - u_n^k)\|_{\mathcal{X}} \leq & CT\rho_k(T)k^{\min\{q+\mu,p\}-1} \left(\sum_{\ell=q+1}^p (\|u^{(\ell+1)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(\ell)}\|_{T,\mu} + \|F^{(\ell)}\|_{T,\mu}) \right) \\ & + \|u^{(p+2)}\|_{T,0} + \|\mathcal{E}\Xi^{(p+1)}\|_{T,0} + \|F^{(p+1)}\|_{T,0}. \end{aligned}$$

If, in addition, the method satisfies Assumption 2.III and

$$u \in \mathcal{C}^{p+2}([0, T]; \mathcal{X}_\mu), \quad \mathcal{E}\Xi, F \in \mathcal{C}^{p+1}([0, T]; \mathcal{X}_\mu) \cap \mathcal{C}^{p+2}([0, T]; \mathcal{X}_0),$$

then

$$\begin{aligned} & \|v(t_n) - v_n^k\|_{\mathcal{X}} + \|A_\star(u(t_n) - u_n^k)\|_{\mathcal{X}} \\ & \leq C(1 + T)\rho_k(T)k^{\min\{q+\mu,p\}} \left(\sum_{\ell=q+1}^{p+1} (\|u^{(\ell)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(\ell)}\|_{T,\mu} + \|F^{(\ell)}\|_{T,\mu}) \right) \\ & \quad + \|u^{(p+2)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(p+2)}\|_{T,0} + \|F^{(p+2)}\|_{T,0}. \end{aligned}$$

The constant k_0 depends only on ω from (2.2) and the Runge–Kutta method. If $\omega = 0$, then k_0 can be chosen arbitrarily large. C depends on ω , M from (2.2), the Runge–Kutta method, and μ .

Remark 3.6 Most of the effort in proving the above theorem is done in order to obtain a convergence rate higher than q , even though the constraint in the stages is approximated only with order q . This is possible by exploiting the additional structure of the discretization error of the side constraint.

Remark 3.7 We formulated all our results for homogeneous initial conditions, since it is sufficient for our purposes in time domain BEM and convolution quadrature. It should be possible to generalize these results to the case of $u_0 \in \text{dom}(A^s)$ for sufficiently large $s \geq 1$

by considering the evolution of the semigroup with inhomogeneous side constraint but homogeneous initial condition and the semigroup of homogeneous constraint but inhomogeneous u_0 separately.

Remark 3.8 The loss of order by 1 in Theorem 3.5 compared to Propositions 3.1 and 3.2 is to be expected. Indeed, if we look at the case $u \in \text{dom}(A^\mu)$ for $\mu \geq 1$, this means $A_\star u \in \text{dom}(A^{\mu-1})$. Applying Proposition 3.2 to this semigroup then also gives a reduced order of $k^{\min(q+\mu,p)}$.

4 Some computations related to the main theorems

We will collect the sampled data and the stage and step parts of the solutions in four formal series

$$\widehat{F}^k(z) := \sum_{n=0}^{\infty} F(t_n + k\mathbf{c})z^n, \quad \widehat{\Xi}^k(z) := \sum_{n=0}^{\infty} \Xi(t_n + k\mathbf{c})z^n, \tag{4.1a}$$

$$\widehat{U}^k(z) := \sum_{n=0}^{\infty} U_n^k z^n, \quad \widehat{u}^k(z) := \sum_{n=0}^{\infty} u_n^k z^n. \tag{4.1b}$$

If the data functions are polynomially bounded in time, the series in (4.1a) are convergent (in \mathcal{X}^m and \mathcal{M}^m respectively) with at least unit radius of convergence. Because of the equivalent formulation of the numerical method in the form (2.8a–2.8c), and using (2.10), it follows that for $k \leq k_0$ [with k_0 chosen using (2.9)], the numerical solution is at least bounded in the form $\|U_n^k\|_{\mathcal{X}} \lesssim C^n$. Thus, the two series in (4.1b) also converge on a sufficiently small disk.

Proposition 4.1 *The sequences $\{U_n^k\}$ and $\{u_n^k\}$ satisfy equations (2.7a–2.7c) if and only if*

$$k^{-1}\delta(z)\widehat{U}^k(z) = (\mathcal{I} \otimes A_\star)\widehat{U}^k(z) + \widehat{F}^k(z) + \frac{k^{-1}}{1-r(\infty)z}Q^{-1}\mathbf{1}u_0, \tag{4.2a}$$

$$(\mathcal{I} \otimes B)\widehat{U}^k(z) = \widehat{\Xi}^k(z), \tag{4.2b}$$

$$\widehat{u}^k(z) = \frac{z}{1-r(\infty)z}\mathbf{b}^\top Q^{-1}\widehat{U}^k(z) + \frac{1}{1-r(\infty)z}u_0^k. \tag{4.2c}$$

Proof Let us start by proving a simple result: *the discrete equations (2.7a) and (2.7c) hold if and only if (2.7a) and*

$$u_{n+1}^k = r(\infty)u_n^k + \mathbf{b}^\top Q^{-1}U_n^k \tag{4.3}$$

hold. To see this, note that (2.7a) is equivalent to

$$\mathcal{Q}^{-1}(U_n^k - \mathbf{1}u_n^k) = k((\mathcal{I} \otimes A_\star)U_n^k + F(t_n + k\mathbf{c}))$$

and therefore (2.7a) and (2.7c) imply

$$\mathbf{b}^\top \mathcal{Q}^{-1}(U_n^k - \mathbf{1}u_n^k) = k((\mathbf{b}^\top \otimes A_\star)U_n^k + \mathbf{b}^\top F(t_n + k\mathbf{c})) = u_{n+1}^k - u_n^k,$$

or equivalently (4.3). The reciprocal statement is proved similarly. The recurrence (4.3) is equivalent to (4.2c). At the same time, the recurrence (2.7a) is equivalent to

$$k^{-1}\mathcal{Q}^{-1}(\widehat{U}^k(z) - \mathbf{1}\widehat{u}^k(z)) = (\mathcal{I} \otimes A_\star)\widehat{U}^k(z) + \widehat{F}^k(z). \tag{4.4}$$

After inserting (4.2c) into (4.4), the formula (4.2a) follows. □

Proposition 4.1 is a rephrasing of [28, Lemma 3.19], where the computation is also laid out in more detail. Note how equations (4.2a)–(4.2b) relate strongly to (2.1a–2.1c), with the discrete symbol $k^{-1}\delta(z)$ playing the role of the time derivative and

$$\frac{k^{-1}}{1 - r(\infty)z} \mathcal{Q}^{-1}\mathbf{1}$$

playing the role of a discrete Dirac delta at time $t = 0$.

Lemma 4.2 ([2, Lemma 2.6]) *If the matrix \mathcal{Q} of the RK method is invertible, then for $|z| < 1$*

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

In particular, if the Runge–Kutta method is A-stable (Assumption 2.II), then $\sigma(\delta(z)) \subset \mathbb{C}_+$.

We need a corollary to the previous result:

Corollary 4.3 *Let Assumption 2.II hold. Then, for all $r_0 < 1$, there exists a constant $d > 0$ such that for all $|z| < r_0$ there holds*

$$\sigma(\delta(z)) \subset \{w \in \mathbb{C}_+ : \text{Re}(w) > d\}.$$

Proof In view of of Lemma 4.2, since $\sigma(\mathcal{Q})$ is finite, independent of z , and contained in \mathbb{C}_+ , we are mainly concerned with the set $\{w \in \mathbb{C} : r(w)z = 1\}$. We first note that

$$\bigcup_{|z| \leq r_0} \{w \in \mathbb{C} : r(w)z = 1\} \subseteq \{w \in \mathbb{C} : |r(w)| \geq 1/r_0\}.$$

Second, we observe that by taking d_0 small enough, we can ensure that $w \mapsto r(w)$ is continuous for $\text{Re}(w) \leq d_0$ and thus

$$\{w \in \mathbb{C} : |r(w)| \geq 1/r_0\} \cap \{w \in \mathbb{C} : \text{Re}(w) \leq d_0\} = r|_{\{\text{Re}(w) \leq d_0\}}^{-1}([1/r_0, \infty))$$

is a closed set. Third, by considering the limit along the imaginary axis, we get

$$|r(\infty)| = \lim_{n \rightarrow \infty} |r(in)| \leq 1.$$

Thus, for $|w|$ sufficiently large, it holds that $|r(w)| \leq 1/r_0$.

Overall, we get that

$$\{w \in \mathbb{C} : |r(w)| \geq 1/r_0\} \cap \{w \in \mathbb{C} : \operatorname{Re}(w) \leq d_0\}$$

is a compact set with empty intersection with the imaginary axis. Thus, it must have a positive distance from it. These observations and Lemma 4.2 conclude the proof. \square

Lemma 4.4 *Let Assumptions 2.I and 2.II hold. For $r_0 < 1$, there exists $k_0 = k_0(\omega, r_0) > 0$ such that for all $k \leq k_0$ and $|z| \leq r_0$ the problem*

$$-k^{-1} \delta(z) \widehat{U} + (\mathcal{I} \otimes A_\star) \widehat{U} = \widehat{F}, \tag{4.5a}$$

$$(\mathcal{I} \otimes B) \widehat{U} = \widehat{E} \tag{4.5b}$$

has a unique solution for arbitrary $\widehat{F} \in \mathcal{X}^m$ and $\widehat{E} \in \mathcal{M}^m$. If $\omega = 0$ in Proposition 2.1, then there are no restrictions on k , and the results holds for all $|z| < 1$.

Proof Assume first that $\mathcal{S} \in \mathbb{C}^{m \times m}$ is such that $\sigma(\mathcal{S}) \subset \{z : \operatorname{Re} z > \omega\}$ and consider the problem

$$-(\mathcal{S} \otimes I) \widehat{U} + (\mathcal{I} \otimes A_\star) \widehat{U} = \widehat{F}, \tag{4.6a}$$

$$(\mathcal{I} \otimes B) \widehat{U} = \widehat{E}. \tag{4.6b}$$

Take first $\widehat{V} := (\mathcal{I} \otimes \mathcal{E}) \widehat{E}$ (where \mathcal{E} is the lifting operator of Assumption 2.I) and then seek $\widehat{W} \in (\operatorname{dom}(A))^m$ satisfying

$$-(\mathcal{S} \otimes I) \widehat{W} + (\mathcal{I} \otimes A) \widehat{W} = \widehat{F} + ((\mathcal{S} - \mathcal{I}) \otimes I) \widehat{V}.$$

This problem is uniquely solvable by Lemma 2.4, since $\sigma(A) \subset \{z : \operatorname{Re} z \leq \omega\}$ and therefore $\sigma(A) \cap \sigma(\mathcal{S}) = \emptyset$. We then define $\widehat{U} := \widehat{V} + \widehat{W}$, which solves (4.6a, 4.6b). To see uniqueness, one observes that the difference of two solutions of (4.6a, 4.6b) solves the homogeneous problem ($\widehat{E} = 0$ and $\widehat{F} = 0$) for which uniqueness was established in Lemma 2.4.

By Corollary 4.3, the union of the spectra of $\delta(z)$ for $|z| \leq r_0$ has a positive distance $d(r_0) > 0$ from the imaginary axis. If we take $k_0 < d(r_0)/\omega$, then $\sigma(k^{-1} \delta(z)) \subset \{s : \operatorname{Re} s > \omega\}$ for all $|z| \leq r_0$ and $k \leq k_0$. When $\omega = 0$, we can take any k_0 . By the previous considerations this implies unique solvability. \square

Proposition 4.5 *Let $U^k = \{U_n^k\}$ and $u^k = \{u_n^k\}$ be sequences satisfying (2.7a–2.7c) with $u_0^k = 0$. The sequence $V^k = \{V_n^k\} = \delta^k U^k$ satisfies*

$$V_n^k = \mathbf{1} v_n^k + k(\mathcal{Q} \otimes A_\star) V_n^k + k \mathcal{Q} G_n^k, \tag{4.7a}$$

$$(\mathcal{I} \otimes B)V_n^k = \Theta_n^k, \tag{4.7b}$$

$$v_{n+1}^k = r(\infty)v_n^k + \mathbf{b}^\top \mathcal{Q}^{-1}V_n^k, \tag{4.7c}$$

for data $v_0^k = 0$, $G^k = \{G_n^k\} := \partial^k\{F(t_n + k\mathbf{c})\}$, and $\Theta^k = \{\Theta_n^k\} := \partial^k\{\Xi(t_n + k\mathbf{c})\}$.
 Moreover,

$$(\mathcal{I} \otimes A_\star)U_n^k = V_n^k - F(t_n + k\mathbf{c}), \tag{4.8a}$$

$$V_n^k = k^{-1}\mathcal{Q}^{-1}(U_n^k - \mathbf{1}u_n^k). \tag{4.8b}$$

Proof Recall that Eqs. (2.7a) and (2.7c) are equivalent to (2.7a) and (4.3), as shown in the proof of Proposition 4.1. Moreover, the latter equations are equivalent to (4.2a–4.2c) in the Z -domain. In the present case we have $u_0 = 0$. For a given square matrix $\mathcal{P} \in \mathbb{C}^{m \times m}$ and an operator C , we have

$$(\mathcal{P} \otimes I)(\mathcal{I} \otimes C) = \mathcal{P} \otimes C = (\mathcal{I} \otimes C)(\mathcal{P} \otimes I),$$

which proves that

$$k^{-1}\delta(z)\widehat{V}^k(z) = (\mathcal{I} \otimes A_\star)\widehat{V}^k(z) + \widehat{G}^k(z), \tag{4.9a}$$

$$(\mathcal{I} \otimes B)\widehat{V}^k(z) = \widehat{\Theta}^k(z), \tag{4.9b}$$

$$\widehat{v}^k(z) = \frac{z}{1 - r(\infty)z} \mathbf{b}^\top \mathcal{Q}^{-1}\widehat{V}^k(z). \tag{4.9c}$$

By Proposition 4.1, Eq. (4.9a–4.9c) are equivalent to (4.7a–4.7c). Finally (4.8a) follows from (4.2a), while (4.8b) follows from (4.4) and (4.8a). \square

Proposition 4.6 Let $U^k = \{U_n^k\}$ and $u^k = \{u_n^k\}$ be sequences satisfying (2.7a–2.7c) with $u_0^k = 0$. The sequence $X^k = \{X_n^k\} = (\partial^k)^{-1}U^k$ satisfies

$$X_n^k = \mathbf{1}x_n^k + k(\mathcal{Q} \otimes A_\star)X_n^k + k\mathcal{Q}H_n^k, \tag{4.10a}$$

$$(\mathcal{I} \otimes B)X_n^k = \Gamma_n^k, \tag{4.10b}$$

$$x_{n+1}^k = r(\infty)x_n^k + \mathbf{b}^\top \mathcal{Q}^{-1}X_n^k \tag{4.10c}$$

$$= x_n^k + k(\mathbf{b}^\top \otimes A_\star)X_n^k + k\mathbf{b}^\top H_n^k, \tag{4.10d}$$

for data $x_0^k := 0$, $H^k = \{H_n^k\} := (\partial^k)^{-1}\{F(t_n + k\mathbf{c})\}$, $\Gamma^k = \{\Gamma_n^k\} := (\partial^k)^{-1}\{\Xi(t_n + k\mathbf{c})\}$.

Proof Follow the proof of Proposition 4.5. \square

5 Some Lemmas regarding Runge–Kutta methods

In order to shorten the statements of the results of this section, in all of them we will understand that:

- (1) We have an RK method with coefficients $\mathcal{Q}, \mathbf{b}, \mathbf{c}$ satisfying Assumption 2.11 (invertibility of \mathcal{Q} and A -stability). The method has classical order p and stage order q .
- (2) We have an operator A in \mathcal{X} that is the generator of a C_0 -semigroup, characterized by the quantities M and ω of Proposition 2.1. The associated Sobolev tower $\{\mathcal{X}_\mu\}$, obtained by interpolation of $\text{dom}(A^\mu)$ for positive integer values of μ , will also be used.

The following lemma will be used at a key point in the arguments below.

Lemma 5.1 *Let A be a linear operator in \mathcal{X} and q be a rational function bounded at infinity whose poles are outside $\sigma(A)$. The following properties hold:*

- (a) The operator $q(A)$ maps $\text{dom}(A^\ell)$ to $\text{dom}(A^\ell)$ for all ℓ .
- (b) If $0 \notin \sigma(A)$, and we define $p(z) := z^{-\ell}q(z)$, then $q(A) = p(A)A^\ell$ in $\text{dom}(A^\ell)$.

Proof To prove (a), show first by induction on ℓ that $(A - \lambda I)^{-1}$ maps $\text{dom}(A^\ell)$ into $\text{dom}(A^{\ell+1})$. Using this result for each of the factors in the definition (2.3) the result follows. To prove (b) note first that p is rational, bounded at infinity, and that $\sigma(A)$ does not intersect the set of poles of p . Using Definition 2.2, we have $p(A) = A^{-\ell}q(A) = q(A)A^{-\ell}$, and the result follows. □

We start by recalling some simple facts about RK methods that we will need in the sequel. Using the notation $\mathbf{c}^\ell := (c_1^\ell, \dots, c_m^\ell)^\top$, the following equalities (order conditions) hold (see e.g. [1, 25]):

$$\mathbf{c}^\ell = \ell \mathcal{Q} \mathbf{c}^{\ell-1}, \quad 0 \leq 1 \leq \ell \leq q, \tag{5.1a}$$

$$\mathbf{b}^\top \mathcal{Q}^j \mathbf{c}^\ell = \frac{\ell!}{(j + \ell + 1)!}, \quad 0 \leq j + \ell \leq p - 1. \tag{5.1b}$$

Therefore,

$$\mathbf{b}^\top \mathcal{Q}^j (\mathbf{c}^\ell - \ell \mathcal{Q} \mathbf{c}^{\ell-1}) = 0, \quad 0 \leq j \leq p - \ell - 1, \quad 1 \leq \ell \leq p \tag{5.2a}$$

$$\ell \mathbf{b}^\top \mathbf{c}^{\ell-1} = 1, \quad 1 \leq \ell \leq p - 1. \tag{5.2b}$$

For a stiffly accurate method we have (2.15) and therefore

$$\mathbf{b}^\top \mathcal{Q}^{-1} \mathbf{c}^\ell = c_m^\ell = 1 \quad \forall \ell \in \mathbb{N}_0. \tag{5.3}$$

The following result is well-known. We just summarize it for ease of reference later on.

Lemma 5.2 (Discrete antiderivative and RK quadrature) *Let $f : [0, T] \rightarrow \mathcal{X}$, $g := \partial^{-1}f$, $G^k = \{G_n^k\} = (\partial^k)^{-1} \{f(t_n + \mathbf{k}c)\}$ and $\{g_n^k\}$ be given by the recursion*

$$g_0^k := 0, \quad g_{n+1}^k := g_n^k + k\mathbf{b}^\top f(t_n + k\mathbf{c}).$$

For the errors $d_n^k := g(t_n) - g_n^k$, and for n such that $nk \leq T$, we have the estimates

$$\|d_n^k\|_{\mathcal{X}} \leq CTk^p \|f^{(p)}\|_{T,0}, \tag{5.4a}$$

$$\|d_n^k - d_{n-1}^k\|_{\mathcal{X}} \leq Ck^{p+1} \max_{t_{n-1} \leq t \leq t_n} \|f^{(p)}(t)\|_{\mathcal{X}}. \tag{5.4b}$$

Additionally, at the stage level we have

$$\|k\mathbf{b}^\top g(t_n + k\mathbf{c}) - k\mathbf{b}^\top G_n^k\|_{\mathcal{X}} \leq Ck^{p+1} (\|f^{(p-1)}\|_{T,0} + T\|f^{(p)}\|_{T,0}). \tag{5.4c}$$

Proof Follows from the fact that the Runge–Kutta method defines a quadrature formula of order p . □

5.1 Estimates on rational functions of the operator

The following results in this section are adaptations from [1]. While they focus on the case $\beta = 0$, we present the necessary generalizations to $\beta = \pm 1$. We will use the rational functions

$$r_{\ell,\beta}(z) := z\mathbf{b}^\top (I - z\mathcal{Q})^{-1} \mathcal{Q}^\beta (\mathbf{c}^\ell - \ell\mathcal{Q}\mathbf{c}^{\ell-1}), \quad \beta \in \{-1, 0, 1\}, \tag{5.5}$$

$$s_n(z) := \sum_{j=0}^n r(z)^j. \tag{5.6}$$

Note that these rational functions are bounded at infinity and that $r_{\ell,\beta}(0) = 0$. We will also use the vector-valued rational function

$$\mathbf{g}(z)^\top := z\mathbf{b}^\top (I - z\mathcal{Q})^{-1}, \tag{5.7}$$

and note that $\mathbf{g}(0) = \mathbf{0}$ and $r(z) = 1 + \mathbf{g}(z)^\top \mathbf{1}$.

Lemma 5.3 *The rational functions (5.5) satisfy*

$$r_{\ell,\beta}(z) = \mathcal{O}(|z|^{p+1-\ell-\beta}) \quad \text{as } |z| \rightarrow 0, \quad \ell \leq p, \quad \beta \in \{0, 1\}. \tag{5.8}$$

The estimate (5.8) is also valid for $\beta = -1$ if the method is stiffly accurate.

Proof For the case $\beta = 0$, the proof is given in [1, Lemma 5]. For the case $\beta = \pm 1$, an analogous proof can be brought to fruition. Namely, one can expand the inverse in the definition of $r_{\ell,\beta}$ into the Neumann series and apply the order conditions (5.2a, 5.2b). For $\beta = -1$, one has to use (5.3) for the leading term. □

Lemma 5.4 *If the RK method satisfies Assumption 2.III, then there exists a constant $k_0 > 0$ depending on the RK method and on ω such that for $\beta \in \{0, 1\}$, $\ell \leq p - \beta$, and all $0 < k \leq k_0$ and all n with $0 \leq nk \leq T$, we have the estimate*

$$\|s_n(kA)r_{\ell,\beta}(kA)\|_{\mathcal{X}_\mu \rightarrow \mathcal{X}} \leq C \rho_k(T) k^{\min\{\mu, p-\ell-\beta\}} \tag{5.9}$$

with $\rho_k(T)$ defined in (2.11). If $\ell = p$ and $\beta = 1$, the left-hand side of (5.9) is bounded by $C\rho_k(T)$. The constant $C > 0$ in (5.9) depends only on the Runge–Kutta method, M and ω , k_0 , ℓ , and μ , but is independent of n and k . If the Runge–Kutta method is stiffly accurate, then the estimate (5.9) also holds for $\beta = -1$. If $\omega = 0$, then k_0 can be chosen arbitrarily.

Proof We adapt the proof of [1, Lemma 6], which only covers the case $\beta = 0$. Consider first the case $p - \ell - \beta \geq 0$ and take any integer μ such that $0 \leq \mu \leq p - \ell - \beta$. Then

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

By Lemma 5.3, the rational function $r_{\ell,\beta}$ has a zero of order $p - \ell - \beta + 1$ at $z = 0$. The rational function $(r(z) - 1)z^\mu$ has a zero of order $\mu + 1 \leq p - \ell - \beta + 1$ at $z = 0$, and all other zeros are in \mathbb{C}_+ by A-stability and Assumption 2.III. This implies that the rational function $q_{\ell,\beta,\mu}$ has its poles in

$$A := \{z \neq 0 : r(z) = 1\} \cup \sigma(\mathcal{Q}^{-1}) \subset \mathbb{C}_+. \tag{5.10}$$

Therefore, for $k_0 > 0$ sufficiently small we get using (2.4):

$$\|q_{\ell,\beta,\mu}(kA)\|_{\mathcal{X} \rightarrow \mathcal{X}} \leq C \quad \forall 0 < k \leq k_0, \tag{5.11}$$

where C depends on M , ω , k_0 , and the RK method. By Lemma 5.1 we have

$$r_{\ell,\beta}(kA) \sum_{j=0}^n r(kA)^j x = k^\mu (r(kA)^{n+1} - I) q_{\ell,\beta,\mu}(kA) A^\mu x \quad \forall x \in \text{dom}(A^\mu), \quad k \leq k_0.$$

This, (5.11), and applying (2.11) to control $r(kA)^{n+1}$ by $\rho_k(T)$, proves (5.9) for integer $\mu \leq p - \ell - \beta$. For larger integer values of μ , the result does not need to be proved as the maximum rate is already attained. We just have to estimate the $\mathcal{X}_{p-\ell-\beta}$ norm by the stronger \mathcal{X}_μ norm. For real values of μ , we use interpolation.

We still need to prove the result when $p - \ell - \beta = -1$, which can only happen when $\ell = p$ and $\beta = 1$. We note that $r_{p,1}(0) = 0$ and we can therefore argue as in the previous case for $\mu = 0$. □

Lemma 5.5 *If the RK method satisfies Assumption 2.III and k_0 is the value given in Lemma 5.4, then*

$$\|s_n(kA)\mathbf{g}(kA)^\top\|_{\mathcal{X}^m \rightarrow \mathcal{X}} \leq C \rho_k(T), \tag{5.12}$$

for all $k \leq k_0$ and n such that $nk \leq T$.

Proof Since $\mathbf{g}(0) = \mathbf{0}$, we can adapt the proof of Lemma 5.4 to each of the components of the vector-valued function \mathbf{g} . The key step is to show that $\mathbf{h}(z)^\top := (r(z) - 1)^{-1}\mathbf{g}(z)$ is bounded at infinity and has all its poles in the set defined in (5.10) and therefore

$$\|\mathbf{h}(kA)^\top\|_{\mathcal{X}^m \rightarrow \mathcal{X}} \leq C \quad \forall k \leq k_0.$$

Since the operator $s_n(kA)\mathbf{g}(kA)^\top$ on the left-hand side of (5.12) can be rewritten as $(r(kA)^{n+1} - I)\mathbf{h}(kA)^\top$, the bound (5.12) follows readily. \square

When dealing with Runge–Kutta methods that do not satisfy the additional Assumption 2.III, we still have the following result:

Lemma 5.6 For $k_0 > 0$ taken as in Lemma 5.4, we can bound for all $k \leq k_0$

$$\|r_{\ell,\beta}(kA)\|_{\mathcal{X}_\mu \rightarrow \mathcal{X}} \leq Ck^{\min\{\mu, p+1-\ell-\beta\}} \tag{5.13}$$

for $\ell \leq p$, $\beta \in \{0, 1\}$, and $\mu \geq 0$. The constant C depends on M , ω , k_0 , μ , and the RK method. The estimate (5.13) also holds for $\beta = -1$ if the method is stiffly accurate. Additionally

$$\|\mathbf{g}(kA)^\top\|_{\mathcal{X}^m \rightarrow \mathcal{X}} \leq C, \quad \forall k \leq k_0. \tag{5.14}$$

Proof The argument to prove (5.13) is very similar to that of Lemma 5.4. By interpolation it is clear that we just need to prove the result for any integer μ satisfying $0 \leq \mu \leq p + 1 - \ell - \beta$. Consider then the rational function $q_{\ell,\beta,\mu}(z) := z^{-\mu}r_{\ell,\beta}(z)$, which is bounded at infinity and has all its poles in $\sigma(\mathcal{Q}^{-1})$ [see (5.10)]. We can then use the same argument to prove (5.11) for this redefined new function $q_{\ell,\beta,\mu}$. (Note that we do not use Assumption 2.III in this argument.) Using that $r_{\ell,\beta}(kA) = k^\mu q_{\ell,\beta,\mu}(kA)A^\mu$ in $\text{dom}(A^\mu)$, the result follows. Stiff accuracy of the method is used in the case $\beta = -1$ when we apply Lemma 5.3, dealing with the zeros of $r_{\ell,-1}$.

The proof of (5.14) is a similar adaptation of the proof of Lemma 5.5. \square

5.2 Estimates on discrete convolutions

The RK error will naturally induce several types of discrete convolutions that we will need to estimate separately. In all of them we will have the structure

$$\omega_0 = 0, \quad \omega_{n+1} := r(kA)\omega_n + k\eta_n, \quad n \geq 0. \tag{5.15}$$

We first deal with the simplest cases.

Lemma 5.7 For $nk \leq T$, the sequence defined by (5.15) can be bounded by

$$\|\omega_n\|_{\mathcal{X}} \leq nk\rho_k(T) \max_{j \leq n} \|\eta_j\|_{\mathcal{X}} \leq T\rho_k(T) \max_{j \leq n} \|\eta_j\|_{\mathcal{X}}.$$

If $\eta_n := \mathbf{g}(kA)^\top \xi_n$ for $\xi_n \in \mathcal{X}^m$, then

$$\|\omega_n\|_{\mathcal{X}} \leq CT\rho_k(T) \max_{j \leq n} \|\xi_j\|_{\mathcal{X}^m}.$$

Proof Follows by writing the recurrence (5.15) as a discrete convolution. □

The next estimate is related to the consistency error of the RK method in the sense of how the RK method approximates derivatives at the stage level. We introduce the operator

$$\mathbf{D}^k(y; t) := y(t + k\mathbf{c}) - y(t)\mathbf{1} - k\mathcal{Q}y(t + k\mathbf{c}). \tag{5.16}$$

The following well-known result about $\mathbf{D}^k(y; t)$ underlies the proofs of [1, Theorem 1] and [25, Theorem 2].

Lemma 5.8 *If $y \in \mathcal{C}^{p+1}([0, T]; \mathcal{X})$, then*

$$\mathbf{D}^k(y; t) = \sum_{j=q+1}^p \frac{k^j}{j!} (\mathbf{c}^j - j\mathcal{Q}\mathbf{c}^{j-1})y^{(j)}(t) + \mathbf{R}^k(t), \tag{5.17}$$

where

$$\|\mathbf{R}^k(t)\|_{\mathcal{X}^m} \leq Ck^{p+1} \max_{t \leq \tau \leq t+k} \|y^{(p+1)}(\tau)\|_{\mathcal{X}}. \tag{5.18}$$

Proof Follows easily from the Taylor expansion and the order conditions (5.1a, 5.1b). □

We are almost ready for the two main lemmas of this section, the first one without Assumption 2.III and the second one with it. These results and their proofs follow [1, Theorem 1 and 2], where only the case $\beta = 0$ is covered.

Lemma 5.9 *Let $y \in \mathcal{C}^{p+1}([0, T]; \mathcal{X}) \cap \mathcal{C}^p([0, T]; \mathcal{X}_\mu)$ for some $\mu \geq 0$. Let the sequence ω_n be defined by (5.15) with $\eta_n := k^{-1}\mathbf{g}(kA)^\top(k\mathcal{Q})^\beta \mathbf{D}(y; t_n)$. Then there exists a constant $k_0 > 0$ depending only on ω from (2.2) and the RK method such that for $k \leq k_0$, $\beta \in \{0, 1\}$ and for $nk \leq T$*

$$\|\omega_n\|_{\mathcal{X}} \leq CT\rho_k(T)k^{\min\{q+\mu+\beta, p+\beta, p\}} \left(\sum_{j=q+1}^p \|y^{(j)}\|_{T, \mu} + \|y^{(p+1)}\|_{T, 0} \right).$$

The estimate also holds for $\beta = -1$ if the method is stiffly accurate. If $\omega = 0$, then k_0 can be chosen arbitrarily.

Proof Introduce

$$e_\beta^k(t) := \mathbf{g}(kA)^\top(k\mathcal{Q})^\beta \mathbf{D}^k(y; t), \quad \beta \in \{-1, 0, 1\}, \tag{5.19}$$

and note $\eta_n = k^{-1}e_\beta^k(t_n)$. Using Lemmas 5.6 and 5.8 [recall the definition of $r_{j,\beta}$ in (5.5)], we can bound

$$\begin{aligned} \|e_\beta^k(t)\|_{\mathcal{X}} &\leq \sum_{j=q+1}^p \frac{k^{j+\beta}}{j!} \|r_{j,\beta}(kA)y^{(j)}(t)\|_{\mathcal{X}+k^\beta} \|g(kA)^\top Q^\beta R^k(t)\|_{\mathcal{X}} \\ &\lesssim \sum_{j=q+1}^p k^{\min\{\mu+j+\beta,p+1\}} \|y^{(j)}(t)\|_{\mathcal{X}_\mu} + k^\beta \|R^k(t)\|_{\mathcal{X}^m}. \end{aligned}$$

By (5.18), we then have

$$\|e_\beta^k(t)\|_{\mathcal{X}} \lesssim k^{1+\min\{q+\mu+\beta,p+\beta,p\}} \left(\sum_{j=q+1}^p \|y^{(j)}(t)\|_{\mathcal{X}_\mu} + \max_{t \leq \tau \leq t+k} \|y^{(p+1)}(\tau)\|_{\mathcal{X}} \right), \tag{5.20}$$

and the result then follows from Lemma 5.7. □

Lemma 5.10 *Let $y \in \mathcal{C}^{p+1}([0, T]; \mathcal{X}_\mu)$ for some $\mu \geq 0$. Let the RK method satisfy Assumption 2.III. Then there exists a constant $k_0 > 0$ depending only on ω from (2.2) and the RK method such that the sequence ω_n defined in Lemma 5.9 satisfies for $\beta \in \{0, 1\}$ and $k \leq k_0$*

$$\|\omega_n\|_{\mathcal{X}} \leq C(1 + T)\rho_k(T)k^{\min\{q+\mu+\beta+1,p\}} \left(\sum_{j=q+1}^{p+1} \|y^{(j)}\|_{T,\mu} \right).$$

If the method is stiffly accurate and $y \in \mathcal{C}^{p+2}([0, T]; \mathcal{X}) \cap \mathcal{C}^{p+1}([0, T]; \mathcal{X}_\mu)$, then for $\beta = -1$

$$\|\omega_n\|_{\mathcal{X}} \leq C(1 + T)\rho_k(T)k^{\min\{q+\mu,p\}} \left(\sum_{j=q+1}^{p+1} \|y^{(j)}\|_{T,\mu} + \|y^{(p+2)}\|_{T,0} \right).$$

If $\omega = 0$, then k_0 can be chosen arbitrarily.

Proof We will use the function e_β^k defined in (5.19) and Abel’s summation by parts:

$$\begin{aligned} \omega_n &= \sum_{j=0}^n r(kA)^{n-j} e_\beta^k(t_j) \\ &= s_n(kA)e_\beta^k(t_0) + \sum_{j=1}^n s_{n-j}(kA)(e_\beta^k(t_j) - e_\beta^k(t_{j-1})) + e_\beta^k(t_n), \end{aligned} \tag{5.21}$$

an expression involving the rational functions s_n defined in (5.6) (recall that $s_0 = 1$). We first apply Lemmas 5.4, 5.5 and 5.8 to estimate

$$\begin{aligned} \|s_n(kA)e_\beta^k(t)\|_{\mathcal{X}} &\leq \sum_{j=q+1}^p \frac{k^{j+\beta}}{j!} \|s_n(kA)r_{j,\beta}(kA)\|_{\mathcal{X}_\mu \rightarrow \mathcal{X}} \|y^{(j)}(t)\|_{\mathcal{X}_\mu} \\ &\quad + Ck^\beta \|s_n(kA)\mathbf{g}(kA)^\top\|_{\mathcal{X}^m \rightarrow \mathcal{X}} \|\mathbf{R}^k(t)\|_{\mathcal{X}^m} \\ &\lesssim \rho_k(T) \sum_{j=q+1}^p k^{\min\{j+\mu+\beta,p\}} \|y^{(j)}(t)\|_{\mathcal{X}_\mu} \\ &\quad + \rho_k(T)k^{\beta+p+1} \max_{t \leq \tau \leq t+k} \|y^{(p+1)}(\tau)\|_{\mathcal{X}} \\ &\lesssim \rho_k(T)k^{1+\min\{q+\mu+\beta,p+\beta,p-1\}} \left(\sum_{j=q+1}^p \|y^{(j)}(t)\|_{\mathcal{X}_\mu} \right. \\ &\quad \left. + \max_{t \leq \tau \leq t+k} \|y^{(p+1)}(\tau)\|_{\mathcal{X}} \right). \end{aligned}$$

Since

$$e_\beta^k(t) - e_\beta^k(t-k) = \mathbf{g}(kA)^\top (k\mathcal{Q})^\beta \mathbf{D}^k (y - y(\cdot - k); t),$$

and using that $\|y^{(j)}(t) - y^{(j)}(t-k)\|_{\mathcal{X}_\mu} \leq k \max_{t-k \leq \tau \leq t+k} \|y^{(j+1)}\|_{\mathcal{X}_\mu}$, a computation analogous to the above bound, but using $y - y(\cdot - k)$ as data implies

$$\begin{aligned} \|s_n(kA)(e_\beta^k(t) - e_\beta^k(t-k))\|_{\mathcal{X}} &\lesssim \rho_k(T)k^{1+\min\{q+1+\mu+\beta,p\}} \\ &\quad \left[\sum_{j=q+2}^{p+1} \max_{t-k \leq \tau \leq t} \|y^{(j)}(\tau)\|_{\mathcal{X}_\mu} + \max_{t-k \leq \tau \leq t+k} \|y^{(p+2)}(\tau)\|_{\mathcal{X}} \right], \end{aligned}$$

and therefore

$$\sum_{j=1}^n \|s_{n-j}(kA)(e_\beta^k(t_j) - e_\beta^k(t_{j-1}))\|_{\mathcal{X}} \lesssim \rho_k(T)t_n k^{\min\{q+1+\mu+\beta,p\}} \left(\sum_{j=q+2}^{p+1} \|y^{(j)}\|_{t_n, \mu} + \|y^{(p+2)}\|_{t_{n+1}, 0} \right).$$

Note that if $\beta \in \{0, 1\}$ we can make a simpler estimate for the term originating from \mathbf{R}^k , (i.e., the one containing the highest derivative) using less regularity for y by not taking advantage of the difference between $y^{(p+1)}(t_j)$ and $y^{(p+1)}(t_{j-1})$ and thus end up requiring less regularity. Using the estimate (5.20) for the last term in (5.21), we have thereby already derived estimates for all three terms in (5.21). \square

6 Proofs

The two different cases (with or without Assumption 2.III) will be collected by using the parameter

$$\alpha := \begin{cases} 1, & \text{if Assumption 2.III holds,} \\ 0, & \text{otherwise.} \end{cases} \tag{6.1}$$

6.1 Proof of Theorem 3.4

Recall that u solves (2.1) with $u(0) = 0$. The functions Ξ and F are the given boundary and volume data. If $\Gamma := \partial^{-1}\Xi$ and $H := \partial^{-1}F$, then $x = \partial^{-1}u$ solves

$$\dot{x}(t) = A_\star x(t) + H(t), \quad t > 0, \quad Bx(t) = \Gamma(t), \quad x(0) = 0. \tag{6.2}$$

On the other hand, $\{X_n^k\} = (\partial^k)^{-1}\{U_n^k\}$ solves by Proposition 4.6:

$$X_n^k = \mathbf{1}x_n^k + k(Q \otimes A_\star)X_n^k + kQH_n^k, \tag{6.3a}$$

$$(\mathcal{I} \otimes B)X_n^k = \Gamma_n^k, \tag{6.3b}$$

$$x_{n+1}^k = x_n^k + k(\mathbf{b}^\top \otimes A_\star)X_n^k + k\mathbf{b}^\top H_n^k. \tag{6.3c}$$

Before we can estimate the difference between the functions x and x_n^k , we need one final lemma.

Lemma 6.1 *Let x solve*

$$\dot{x}(t) = A_\star x(t) + H(t), \quad t > 0, \quad Bx(t) = \Gamma(t), \quad x(0) = 0. \tag{6.4}$$

Assume that for some $\mu \geq 0$ we have

$$x \in \mathcal{C}^{\mu+1}([0, T]; \mathcal{X}_\mu), \quad H \in \mathcal{C}^\mu([0, T]; \mathcal{X}_\mu), \quad \mathcal{E}\Gamma \in \mathcal{C}^\mu([0, T]; \mathcal{X}_\mu).$$

Then $x - \mathcal{E}\Gamma \in \mathcal{C}^\mu([0, T]; \mathcal{X}_{\mu+1})$.

Proof We set $y := x - \mathcal{E}\Gamma$. By assumption we have $y \in \mathcal{C}^\mu([0, T]; \mathcal{X}_\mu)$ and $B(x - \mathcal{E}\Gamma) = 0$. Since $x \in \text{dom}(A_\star)$ and $\text{range } \mathcal{E} \subset \text{dom}(A_\star)$ this implies $y(t) \in \text{dom}(A)$ for all $t \in [0, T]$. We further calculate using (6.4) and $\text{range } \mathcal{E} \subseteq \ker(I - A)_\star$:

$$Ay = A_\star x - A_\star \mathcal{E}\Gamma = \dot{x} - H - \mathcal{E}\dot{\Gamma}.$$

Each of the terms is assumed in $\mathcal{C}^\mu([0, T]; \mathcal{X}_\mu)$, thus $y \in \mathcal{C}^\mu([0, T]; \mathcal{X}_{\mu+1})$. □

We will need the sequences $\{\gamma_n^k\}$ and $\{h_n^k\}$ with the scalar parts of the computations of $\{\Gamma_n^k\}$ and $\{H_n^k\}$ respectively, namely (see Lemma 2.6),

$$\gamma_0^k := 0, \quad \gamma_n^k = \gamma_{n-1}^k + k\mathbf{b}^\top \Xi(t_n + k\mathbf{c}), \tag{6.5a}$$

$$h_0^k := 0, \quad h_n^k = h_{n-1}^k + k\mathbf{b}^\top F(t_n + k\mathbf{c}). \tag{6.5b}$$

We then consider

$$\Delta_n^k := (\mathcal{I} \otimes \mathcal{E})(\Gamma(t_n + k\mathbf{c}) - \Gamma_n^k), \quad \delta_n^k := \mathcal{E}(\Gamma(t_n) - \gamma_n^k).$$

Using (6.5a), the definition $\Gamma^k = (\partial^k)^{-1}\Xi$, and (2.13), we can write

$$\begin{aligned} \Delta_n^k - \mathbf{1}\delta_n^k &= (\mathcal{I} \otimes \mathcal{E})\Gamma(t_n + k\mathbf{c}) - \mathbf{1}\mathcal{E}\Gamma(t_n) - k\mathcal{Q} \otimes \mathcal{E}\dot{\Gamma}(t_n + k\mathbf{c}) \\ &= \mathbf{D}^k(\mathcal{E}\Gamma; t_n). \end{aligned} \tag{6.6}$$

Lemma 5.2 (take $f = \mathcal{E}\Xi$ for the first three inequalities and $f = F$ for the last one) proves that

$$\|\delta_n^k\|_{\mathcal{X}} \leq CTk^p \|\mathcal{E}\Xi^{(p)}\|_{T,0}, \tag{6.7a}$$

$$\|\delta_n^k - \delta_{n-1}^k\|_{\mathcal{X}} \leq Ck^{p+1} \|\mathcal{E}\Xi^{(p)}\|_{T,0}, \tag{6.7b}$$

$$\|\mathbf{k}\mathbf{b}^\top \Delta_n^k\|_{\mathcal{X}} \leq Ck^{p+1} (\|\mathcal{E}\Xi^{(p-1)}\|_{T,0} + T\|\mathcal{E}\Xi^{(p)}\|_{T,0}), \tag{6.7c}$$

$$\|H(t_n) - h_n^k\|_{\mathcal{X}} \leq CTk^p \|F^{(p)}\|_{T,0}. \tag{6.7d}$$

The error analysis is derived by tracking the evolution of the following differences

$$E_n^k := x(t_n + k\mathbf{c}) - X_n^k - \Delta_n^k \in (\text{dom}(A))^m, \quad e_n^k := x(t_n) - x_n^k - \delta_n^k,$$

(compare (6.2) and (6.3b) to see the vanishing boundary condition for E_n^k) and note that by (6.7a)

$$\|x(t_n) - x_n^k\|_{\mathcal{X}} \leq \|e_n^k\|_{\mathcal{X}} + CTk^p \|\mathcal{E}\Xi^{(p)}\|_{T,0},$$

which shows that we only need to estimate e_n^k to prove Theorem 3.4.

We start with the observation that x solves the following equation, as can be easily derived from Eq. (6.2):

$$\begin{aligned} x(t_n + k\mathbf{c}) &= \mathbf{1}x(t_n) + k\mathcal{Q} \otimes A_\star x(t_n + k\mathbf{c}) \\ &\quad + x(t_n + k\mathbf{c}) - k\mathcal{Q}\dot{x}(t_n + k\mathbf{c}) + k\mathcal{Q}H(t_n + k\mathbf{c}) - \mathbf{1}x(t_n) \\ &= \mathbf{1}x(t_n) + k\mathcal{Q} \otimes A_\star x(t_n + k\mathbf{c}) + k\mathcal{Q}H(t_n + k\mathbf{c}) + \mathbf{D}^k(x, t_n). \end{aligned} \tag{6.8}$$

Recalling that Assumption 2.I included the hypothesis $\text{range } \mathcal{E} \subset \ker(I - A_\star)$, we have $(\mathcal{Q} \otimes A_\star)\Delta_n^k = \mathcal{Q}\Delta_n^k$. Combining (6.8) and (6.3a), we get

$$E_n^k = \mathbf{1}e_n^k + k(\mathcal{Q} \otimes A)E_n^k + \mathbf{D}^k(x, t_n) - k\mathcal{Q}(H_n^k - H(t_n + k\mathbf{c})) + \mathbf{1}\delta_n^k - \Delta_n^k + k\mathcal{Q}\Delta_n^k.$$

Naive estimation of the terms $\mathbf{D}^k(x, t_n)$ and $\Delta_n^k - \mathbf{1}\delta_n^k$ would yield convergence rates similar to Propositions 3.1 and 3.2. In order to get an increased rate, as stated in Theorem 3.4, we combine these two terms using the function $Y(t) := x(t) - \mathcal{E}\Gamma(t)$. Lemma 6.1 and the assumptions of Theorem 3.4 ensure $Y \in \mathcal{C}^{p+\alpha}([0, T]; \mathcal{X}_{\mu+1}) \cap \mathcal{C}^{p+1}([0, T]; \mathcal{X})$.

We can thus further simplify

$$\begin{aligned} E_n^k &= \mathbf{1}e_n^k + k(\mathcal{Q} \otimes A)E_n^k + \mathbf{D}^k(x, t_n) - \mathbf{D}^k(\mathcal{E}\Gamma, t_n) \\ &\quad + k\mathcal{Q}\mathbf{D}^k(H, t_n) - k\mathcal{Q}\mathbf{1}(h_h^k - H(t_n)) + k\mathcal{Q}\Delta_n^k \\ &= \mathbf{1}e_n^k + k(\mathcal{Q} \otimes A)E_n^k + \mathbf{D}^k(Y, t_n) \\ &\quad + k\mathcal{Q}\mathbf{D}^k(H, t_n) - k\mathcal{Q}\mathbf{1}(h_h^k - H(t_n)) + k\mathcal{Q}\Delta_n^k. \end{aligned} \tag{6.9}$$

This then immediately gives (recall (5.7) for the definition of \mathbf{g})

$$k(\mathbf{b}^\top \otimes A)E_n^k = \mathbf{g}(kA)^\top [\mathbf{1}e_n^k + \mathbf{D}^k(Y, t) + k\mathbf{Q}\mathbf{D}^k(H, t_n) - k\mathbf{Q}\mathbf{1}(h_n^k - H(t_n)) + k\mathbf{Q}\Delta_n^k]. \tag{6.10}$$

It is easy to see from (6.2) that x satisfies

$$x(t_{n+1}) = x(t_n) + k\mathbf{b}^\top \otimes A_\star x(t_n + k\mathbf{c}) + [x(t_{n+1}) - x(t_n) - k\mathbf{b}^\top \dot{x}(t_n + k\mathbf{c}) + k\mathbf{b}^\top H(t_n + k\mathbf{c})].$$

Subtracting (6.3c) from this, inserting (6.10), using that $(\mathbf{b}^\top \otimes A_\star)\Delta_n^k = \mathbf{b}^\top \Delta_n^k$, and setting

$$\varphi_n^k := [x(t_{n+1}) - x(t_n) - k\mathbf{b}^\top \dot{x}(t_n + k\mathbf{c})] + k\mathbf{b}^\top (H(t_n + k\mathbf{c}) - H_n^k),$$

we have

$$\begin{aligned} e_{n+1}^k &= e_n^k + k(\mathbf{b}^\top \otimes A)E_n^k + k(\mathbf{b}^\top \otimes A_\star)\Delta_n^k + \delta_n^k - \delta_{n+1}^k + \varphi_n^k \\ &= r(kA)e_n^k + \mathbf{g}(kA)^\top (k\mathbf{Q})\Delta_n^k + \mathbf{g}(kA)^\top \mathbf{D}^k(Y, t_n) \\ &\quad + \mathbf{g}(kA)^\top (k\mathbf{Q})\mathbf{D}^k(H, t_n) - \mathbf{g}(kA)^\top (k\mathbf{Q})\mathbf{1}(h_n^k - H(t_n)) \\ &\quad + k\mathbf{b}^\top \Delta_n^k + \delta_n^k - \delta_{n+1}^k + \varphi_n^k. \end{aligned}$$

What is left is the careful combination of terms so that we can bound everything using Lemmas 5.7, 5.9, and 5.10 by writing

$$\begin{aligned} e_{n+1}^k - r(kA)e_n^k &= \mathbf{g}(kA)^\top (\mathbf{D}^k(Y; t_n)) + \mathbf{g}(kA)^\top (k\mathbf{Q})(\mathbf{D}^k(\mathcal{E}\Gamma + H; t_n)) \\ &\quad + \mathbf{g}(kA)^\top \mathbf{Q}\mathbf{1}k(\delta_n^k - (h_n^k - H(t_n))) \\ &\quad + k\mathbf{b}^\top \Delta_n^k + (\delta_n^k - \delta_{n+1}^k) + \varphi_n^k. \end{aligned}$$

Since the above recurrence defining $\{e_n^k\}$ is linear as a function of the right-hand side, we can estimate its norm by adding the effects of each of the terms. In the order in which they appear in the last expression, we use: Lemmas 5.9–5.10 with $\beta = 0$, but noting that $Y(t) \in \text{dom}(A^{\mu+1})$; Lemmas 5.9–5.10 with $\beta = 1$; Lemma 5.7 combined with (6.7a) and (6.7d); Lemma 5.7 combined with (6.7c) and (6.7b); for the first term of φ_n^k we use Lemma 5.7 combined with Lemma 5.2 with $f := \dot{x}$. Finally, for the second contribution to φ_n^k , we use (5.4c).

Combined, these results give

$$\begin{aligned} \|e_n^k\|_{\mathcal{X}} &\leq C(T + \alpha)\rho_k(T)k^{\min\{q+\mu+1+\alpha,p\}} \left(\sum_{j=q+1}^p \|Y^{(j)}\|_{T,\mu+1} + \|Y^{(p+1)}\|_{T,\alpha(\mu+1)} \right) \\ &\quad + C(T + \alpha)\rho_k(T)k^{\min\{q+\mu+1+\alpha,p\}} \left(\sum_{j=q+1}^p \|(\mathcal{E}\Gamma)^{(j)}\|_{T,\mu} + \|(\mathcal{E}\Gamma)^{(p+1)}\|_{T,\alpha\mu} \right) \\ &\quad + C(T + \alpha)\rho_k(T)k^{\min\{q+\mu+1+\alpha,p\}} \left(\sum_{j=q+1}^p \|H^{(j)}\|_{T,\mu} + \|H^{(p+1)}\|_{T,\alpha\mu} \right) \\ &\quad + CT^2\rho_k(T)k^{p+1}\|\mathcal{E}\Xi^{(p)}\|_{T,0} + CT^2\rho_k(T)k^{p+1}\|F^{(p)}\|_{T,0} \\ &\quad + CT\rho_k(T)k^{p+1}(\|\mathcal{E}\Xi^{(p-1)}\|_{T,0} + T\|\mathcal{E}\Xi^{(p)}\|_{T,0}) \\ &\quad + CT\rho_k(T)k^{p+1}(\|H^{(p-1)}\|_{T,0} + T\|H^{(p)}\|_{T,0}). \end{aligned}$$

If we apply Lemma 6.1 to bound the $\mathcal{X}_{\mu+1}$ -norm, we arrive at the stated estimate.

6.2 Proof of Theorem 3.5

This proof is very similar to the one of Theorem 3.4 but slightly simpler. We will point out the main steps of the proof. We first focus on showing the estimate for $v - v^k$. Note that we use the simple form of δ^k for stiffly accurate RK methods given in Lemma 2.7. We define $G := \dot{F}$ and $\Theta := \dot{\Xi}$ so that $v = \dot{u}$ satisfies

$$\dot{v}(t) = A_\star v(t) + G(t), \quad t > 0, \quad Bv(t) = \Theta(t), \quad v(0) = 0.$$

Its RK approximation

$$\tilde{V}_n^k = \mathbf{1}\tilde{v}_n^k + k(\mathcal{Q} \otimes A_\star)\tilde{V}_n^k + k\mathcal{Q}G(t_n + k\mathbf{c}), \tag{6.11a}$$

$$(\mathcal{I} \otimes B)\tilde{V}_n^k = \Theta(t_n + k\mathbf{c}), \tag{6.11b}$$

$$\tilde{v}_{n+1}^k = \tilde{v}_n^k + k(\mathbf{b}^\top \otimes A_\star)\tilde{V}_n^k + k\mathbf{b}^\top G(t_n + k\mathbf{c}), \tag{6.11c}$$

and $\{V_n^k\} = \delta^k\{U_n^k\}$ satisfies (see Proposition 4.5 and Lemma 2.7, where we use stiff accuracy of the RK scheme, and recall that $\{G_n^k\} = \delta^k\{F(t_n + k\mathbf{c})\}$ and $\{\Theta_n^k\} = \delta^k\{\Xi(t_n + k\mathbf{c})\}$)

$$V_n^k = \mathbf{1}v_n^k + k(\mathcal{Q} \otimes A_\star)V_n^k + k\mathcal{Q}G_n^k, \tag{6.12a}$$

$$(\mathcal{I} \otimes B)V_n^k = \Theta_n^k = k^{-1}\mathcal{Q}^{-1}(\Xi(t_n + k\mathbf{c}) - \mathbf{1}\Xi(t_n)), \tag{6.12b}$$

$$v_{n+1}^k = v_n^k + k(\mathbf{b}^\top \otimes A_\star)V_n^k + k\mathbf{b}^\top G_n^k. \tag{6.12c}$$

Let then

$$A_n^k := (\mathcal{I} \otimes \mathcal{E})(\Theta_n^k - \Theta(t_n + k\mathbf{c})) = k^{-1}\mathcal{Q}^{-1}\mathbf{D}^k(\mathcal{E}\Xi; t_n)$$

and [note (6.11b) and (6.11c)]

$$E_n^k := V_n^k - \tilde{V}_n^k - \Delta_n^k \in (\text{dom}(A))^m, \quad e_n^k := v_n^k - \tilde{v}_n^k.$$

By (6.11a) and (6.12a), using that $(\mathcal{Q} \otimes A_\star) \Delta_n^k = \mathcal{Q} \Delta_n^k$ (assumption on the lifting) and Lemma 2.7 to represent G_n^k , we have

$$k(\mathbf{b}^\top \otimes A)E_n^k = \mathbf{g}(kA)^\top (\mathbf{1}e_n^k - \Delta_n^k + k\mathcal{Q}\Delta_n^k + \mathbf{D}^k(F; t_n))$$

and therefore, from (6.11c) and (6.12c)

$$e_{n+1}^k = r(kA)e_n^k - \mathbf{g}(kA)^\top (k\mathcal{Q})^{-1} \mathbf{D}^k(\mathcal{E}\Xi; t_n) + \mathbf{g}(kA)^\top \mathbf{D}^k(\mathcal{E}\Xi + F; t_n) + k\mathbf{b}^\top \mathcal{Q}^{-1} \mathbf{D}^k(\mathcal{E}\Xi + F; t_n). \tag{6.13}$$

The final term can be shown to be of order $\mathcal{O}(k^{p+1})$ by combining (5.17) with (5.2b) and (5.3). We then use Lemmas 5.9 and 5.10 with $\beta = -1$ and $\beta = 0$ as well as Lemma 5.7 to bound

$$\begin{aligned} \|e_n^k\|_{\mathcal{X}} &\leq C(T + \alpha)\rho_k(T)k^{\alpha-1+\min\{q+\mu,p\}} \left(\sum_{j=q+1}^{p+\alpha} \|\mathcal{E}\Xi^{(j)}\|_{T,\mu} + \|\mathcal{E}\Xi^{(p+1+\alpha)}\|_{T,0} \right) \\ &\quad + C(T + \alpha)\rho_k(T)k^{\alpha-1+\min\{q+\mu,p\}} \left(\sum_{j=q+1}^{p+\alpha} \|F^{(j)}\|_{T,\mu} + \|F^{(p+1+\alpha)}\|_{T,0} \right). \end{aligned}$$

Finally Propositions 3.1 and 3.2 are used to bound

$$\|v(t_n) - \tilde{v}_n^k\|_{\mathcal{X}} \leq C(T + \alpha)\rho_k(T)k^{\min\{q+\mu+\alpha,p\}} \left(\sum_{\ell=q+2}^{p+1+\alpha} \|u^{(\ell)}\|_{T,\mu} + \|u^{(p+2)}\|_{T,0} \right). \tag{6.14}$$

The estimate involving $A_\star u$ can be proved as an easy corollary of the estimate on v . Since the last stage of a stiffly accurate method is the step, we have that (4.8a) implies that

$$A_\star u_n^k = v_n^k - F(t_n)$$

and therefore

$$A_\star u(t_n) - A_\star u_n^k = v(t_n) - v_n^k.$$

7 Maximal dissipative operators in Hilbert space

In this short section we summarize some results that show that the hypotheses on the abstract equation and its discretization are simpler for maximal dissipative operators on Hilbert spaces. These results are well-known and will be needed when applying the theory developed in the previous sections to some model problems in Sect. 8.

If A is maximal dissipative in the Hilbert space \mathcal{X} , i.e.,

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

and if $A - I : \text{dom}(A) \rightarrow \mathcal{X}$ is invertible with bounded inverse, then the constants in Proposition 2.1 can be chosen as $M = 1$ and $\omega = 0$. In this case A generates a contraction semigroup in \mathcal{H} . See [26, Sect. 1.4].

In particular, if the RK method satisfies Assumption 2.II and

$$\sigma(A) \subset \{z : \operatorname{Re} z \leq 0\}, \tag{7.1}$$

then the Eq. (2.7a–2.7c) [or equivalently (2.8a–2.8c)], defining the RK approximation of (2.1a–2.1c) are uniquely solvable for any $k > 0$ (apply Lemma 2.4 with $\mathcal{S} = k^{-1}\mathcal{Q}^{-1}$). The following lemma gives a bound for $\rho_k(T)$ in this specific setting.

Lemma 7.1 (Discrete stability) *Let A be a linear, maximal dissipative operator on a Hilbert space \mathcal{H} . For A -stable Runge–Kutta methods and arbitrary $k > 0$, we can bound*

$$\|r(kA)\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1, \tag{7.2}$$

and therefore $\rho_k(T) \leq 1$ for all k and $T > 0$.

Proof Let $c(z) := (z + 1)/(z - 1)$, and note that $c(A) = (A + I)(A - I)^{-1}$ is well defined and since

$$\|(A + I)x\|^2 - \|(A - I)x\|^2 = 4\operatorname{Re} \langle Ax, x \rangle \leq 0 \quad \forall x \in \operatorname{dom}(A),$$

it is clear that $\|c(A)\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1$. Consider now the rational function $q := r \circ c$. Since c maps $B(0; 1)$ bijectively into $\{z : \operatorname{Re} z < 0\}$ and r maps the latter set to $\overline{B(0; 1)}$ (this is A -stability), it follows that $q : B(0; 1) \rightarrow \overline{B(0; 1)}$. Since $\sigma(c(A)) \subset \overline{B(0; 1)}$ and $c(A)$ is bounded, we can define $q(c(A))$ and show (use a classical result of von Neumann [37, Section 4] or [30, Chapter XI, Section 154]) that $\|q(c(A))\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1$.

Finally, using that $c(c(z)) = z$ for all z , it follows that $r = q \circ c$. It is then an easy computation to prove that $r(A) = q(c(A))$. (We remark that this equality can also be proved using functional calculus.) \square

In Propositions 3.1 and 3.2, if A is maximal dissipative, k_0 can be chosen arbitrarily. In Lemma 5.4, if A is maximal dissipative, k_0 can be chosen arbitrarily.

8 Applications

In this section Ω is a bounded Lipschitz open set in \mathbb{R}^d ($d = 2$ or 3) with boundary Γ .

We use the usual (fractional) Sobolev spaces $H^s(\Omega)$ for $s \geq 0$ and introduce the space $H^1_A(\Omega) := \{u \in H^1(\Omega) : Au \in L^2(\Omega)\}$. On the boundary Γ , we also consider Sobolev spaces $H^s(\Gamma)$ and their duals $H^{-s}(\Gamma)$. Details can, for example be found in [22].

For the trace operators, we make the convention that the index $+$ relates to exterior and $-$ means the trace is taken from the interior of Ω . For example, the two bounded surjective trace operators $\gamma^\pm : H^1(\mathbb{R}^d \setminus \Gamma) \rightarrow H^{1/2}(\Gamma)$ denote the trace from $\mathbb{R}^d \setminus \overline{\Omega}$ and Ω , respectively, and we will denote $H^{-1/2}(\Gamma)$ for the dual of the trace space. The angled bracket $\langle \cdot, \cdot \rangle_\Gamma$ will be used for the $H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma)$ duality pairing and $(\cdot, \cdot)_{\mathbb{R}^d}$ will be used for the inner product in $L^2(\mathbb{R}^d)$ and $[L^2(\mathbb{R}^d)]^d$. We will also use the normal traces $\gamma^\pm_\nu : H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma) \rightarrow H^{-1/2}(\Gamma)$ and the normal derivative operators ∂^\pm_ν . Here we make the convention that the normal derivative points out of Ω for both interior and exterior trace.

We note that the applications in this section are chosen for their simplicity. More complicated applications, also involving full discretizations by convolution quadrature and boundary elements of systems of time domain boundary integral equations can be found in [29] and [27].

8.1 Boundary integral equations and convolution quadrature

In this section, we give a very brief introduction to boundary integral equations and their discretization using convolution quadrature. In that way, we can later easily state our methods for both the heat and wave equations in a concise and unified language. We present the result mostly formally, but note that they can be made rigorous under mild assumptions on the appearing functions. This theory can be found in most monographs on boundary element methods, see e.g. [22, 32, 33] or [31].

For $s \in \mathbb{C}_+$, we consider solutions $u \in H^1(\mathbb{R}^d \setminus \Gamma)$ to the Helmholtz equation

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

For this problem, the fundamental solution is given by

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

where $H_0^{(1)}$ denotes the first kind Hankel function of order 0. Using the representation formula, u can be rewritten using only its boundary data:

$$u(x) = S(s)[[\partial_\nu u]] - D(s)[[\gamma u]], \tag{8.1}$$

where the single layer and double layer potentials are given by

$$\begin{aligned} (S(s)\varphi)(x) &:= \int_\Gamma \Phi(x - y; s)\varphi(y) \, dy, \\ (D(s)\psi)(x) &:= \int_\Gamma \partial_{\nu(y)} \Phi(x - y; s)\psi(y) \, dy, \end{aligned}$$

and the expressions $[[\gamma u]] := \gamma^- u - \gamma^+ u$ and $[[\partial_\nu u]] := \partial_\nu^- u - \partial_\nu^+ u$ denote the jump of the trace of v and normal derivative across Γ .

We note that both $S(s)\lambda$ and $D(s)\psi$ solve the Helmholtz equation for any given densities $\lambda \in H^{-1/2}(\Gamma)$ and $\psi \in H^{1/2}(\Gamma)$.

We will need the following four boundary integral operators:

$$V(s) := \gamma^\pm S(s), \quad K(s) := \frac{1}{2}(\gamma^+ S(s) + \gamma^- S(s)), \tag{8.2}$$

$$K^t(s) := \frac{1}{2}(\partial_\nu^+ D(s) + \partial_\nu^- D(s)), \quad W(s) := -\partial_\nu^\pm D(s). \tag{8.3}$$

When solving problems in the time domain, we can leverage our knowledge of the Helmholtz equation using the Laplace transform \mathcal{L} . For an operator valued analytic function F with $\text{dom}(F) \supset \mathbb{C}_+$, we can then define the convolution operator $F(\partial) := \mathcal{L}^{-1} \circ F \circ \mathcal{L}$, where \mathcal{L} is the Laplace transform in the sense of causal distributions. (Precise definitions can be found in [31, Chapter 3] and [21]).

Given a Runge–Kutta method, it is then easy to define the convolution quadrature approximation to such operators, as was introduced in [19]. We just replace the Laplace transform by the Z -transform \mathcal{Z} and s with the function δ/k , i.e., we define:

$$F(\partial^k)g := \mathcal{L}^{-1} \left(F \left(\frac{\delta(\cdot)}{k} \right) \mathcal{L}[g] \right),$$

where g denotes a sequence in the shared domain of $F(s)$ and $k > 0$ denotes the step size. The matrix-valued function $z \mapsto F(\frac{\delta(z)}{k})$ is defined using the Riesz-Dunford calculus, but can be computed in practice by diagonalizing the argument.

Remark 8.1 We note that our use of the notation ∂^k and $(\partial^k)^{-1}$ is consistent with this definition by using the functions $F(s) := s$ and $F(s) := s^{-1}$.

8.2 An exotic transmission problem

In this section we show how to apply Theorems 3.4 and 3.5 to a transmission problem in free space associated to the infinitesimal generator of a group of isometries (both $\pm A$ are maximal dissipative) with some exotic transmission conditions which impose partial observation of a trace. In Sect. 8.3 we will explain how this problem is related to a boundary integral representation of a scattering problem and how the current results yield the analysis of a fully discrete method for that integral representation. We keep the presentation brief. For more details and exemplary applications we refer to [13].

Let Y_h be a closed subspace of $H^{1/2}(\Gamma)$ (in practice it will be finite-dimensional) and consider the spaces

$$\mathbf{H}(\text{div}, \mathbb{R}^d \setminus \Gamma) := \{ \mathbf{w} \in L^2(\mathbb{R}^d \setminus \Gamma)^d : \nabla \cdot \mathbf{w} \in L^2(\mathbb{R}^d \setminus \Gamma) \}, \tag{8.4a}$$

$$\mathbf{V}_h := \{ v \in H^1(\mathbb{R}^d \setminus \Gamma) : \llbracket \gamma v \rrbracket \in Y_h \}, \tag{8.4b}$$

$$\mathbf{W}_h := \{ \mathbf{w} \in \mathbf{H}(\text{div}, \mathbb{R}^d \setminus \Gamma) : \langle \gamma_v^- \mathbf{w}, \mu_h \rangle_\Gamma = 0 \quad \forall \mu_h \in Y_h \}, \tag{8.4c}$$

$$\mathbf{W}_h^0 := \mathbf{W}_h \cap \mathbf{H}(\text{div}, \mathbb{R}^d) \tag{8.4d}$$

$$= \{ \mathbf{w} \in \mathbf{H}(\text{div}, \mathbb{R}^d) : \langle \gamma_v^- \mathbf{w}, \mu_h \rangle_\Gamma = 0 \quad \forall \mu_h \in Y_h \}. \tag{8.4e}$$

The condition $\llbracket \gamma v \rrbracket \in Y_h$ is equivalent to

$$(\nabla \cdot \mathbf{w}, v)_{\mathbb{R}^d \setminus \Gamma} + (\mathbf{w}, \nabla v)_{\mathbb{R}^d} = 0 \quad \forall \mathbf{w} \in \mathbf{W}_h^0. \tag{8.5}$$

We then set

$$\mathcal{X} := L^2(\mathbb{R}^d \setminus \Gamma) \times L^2(\mathbb{R}^d \setminus \Gamma)^d, \quad \mathcal{V} := \mathbf{V}_h \times \mathbf{W}_h, \quad \mathcal{M} := \mathbf{H}^{-1/2}(\Gamma).$$

In \mathcal{X} we use the natural inner product, in \mathcal{V} we use the norm of $H^1(\mathbb{R}^d \setminus \Gamma) \times \mathbf{H}(\text{div}, \mathbb{R}^d \setminus \Gamma)$, and in \mathcal{M} we use the usual norm. We will define $A_\star : \text{dom}(A_\star) = \mathcal{V} \rightarrow \mathcal{X}$ and $B : \mathcal{V} \rightarrow \mathcal{M}$ by

$$A_\star(v, \mathbf{w}) := (\nabla \cdot \mathbf{w}, \nabla v), \quad B(v, \mathbf{w}) := \gamma_v^- v - \gamma_v^+ \mathbf{w},$$

understanding that A_\star can also be extended to $H^1(\mathbb{R}^d \setminus \Gamma) \times \mathbf{H}(\text{div}, \mathbb{R}^d \setminus \Gamma)$. As we did in

Assumption 2.1, we consider $\text{dom}(A) = \ker B = V_h \times W_h^0$ and define A as the restriction of A_\star to this subset.

Proposition 8.2 *The operators $\pm A$ are maximal dissipative.*

Proof The identity (8.5) shows that $\langle A(v, \mathbf{w}), (v, \mathbf{w}) \rangle_{\mathcal{X}} = 0$ for all $(v, \mathbf{w}) \in V_h \times W_h^0$. Given $(f, \mathbf{f}) \in \mathcal{X}$, solving the coercive problem

$$\text{find } v \in V_h : \quad (\nabla v, \nabla \tau)_{\mathbb{R}^d} + (v, \tau)_{\mathbb{R}^d} = (f, \tau)_{\mathbb{R}^d} - (\mathbf{f}, \nabla \tau)_{\mathbb{R}^d} \quad \forall \tau \in V_h,$$

and defining $\mathbf{w} = \nabla v + \mathbf{f}$, we have a pair $(v, \mathbf{w}) \in V_h \times W_h^0$ such that $(v, \mathbf{w}) - A(v, \mathbf{w}) = (f, \mathbf{f})$ and thus A is maximal dissipative. The proof that $-A$ is maximal dissipative is similar. (Note that this is a particular case of what appears in [13].) \square

We consider the standard problem (2.1a–2.1c) with vanishing initial conditions and data $F = 0$ and $\Xi = g : [0, \infty) \rightarrow L^2(\Gamma)$, namely, we look for $(v_h, \mathbf{w}_h) : [0, \infty) \rightarrow \text{dom}(A_\star)$ such that

$$(\dot{v}_h(t), \dot{\mathbf{w}}_h(t)) = (\nabla \cdot \mathbf{w}_h(t), \nabla v_h(t)) \quad \forall t > 0, \tag{8.6a}$$

$$\langle \gamma_v^+ \mathbf{w}_h(t) - \gamma_v^- \mathbf{w}^h(t), \mu \rangle_\Gamma = \langle g(t), \mu \rangle_\Gamma \quad \forall \mu \in Y_h, \quad \forall t > 0, \tag{8.6b}$$

$$(v_h(0), \mathbf{w}_h(0)) = (0, \mathbf{0}). \tag{8.6c}$$

Uniqueness of the solution to (8.6a–8.6c) follows from Proposition 8.2. We will handle existence of a solution below. The quantities of interest are $u_h := \partial^{-1} v_h$ and its Dirichlet trace $\psi_h := \llbracket \gamma u_h \rrbracket : [0, \infty) \rightarrow Y_h$.

Proposition 8.3 *There exists a linear bounded right inverse $\mathcal{E} : \mathcal{M} \rightarrow \text{dom}(A_\star)$ of B such that $\text{range } \mathcal{E} \subset \ker(I - A_\star)$. The norm of \mathcal{E} is independent of the space Y_h .*

Proof Given $\xi \in \mathcal{M} = H^{-1/2}(\Gamma)$, we solve the coercive problem

$$\text{find } v \in V_h : \quad (\nabla v, \nabla \tau)_{\mathbb{R}^d \setminus \Gamma} + (v, \tau)_{\mathbb{R}^d \setminus \Gamma} = \langle \xi, \gamma^+ \tau \rangle_\Gamma \quad \forall \tau \in V_h, \tag{8.7}$$

and we set $\mathbf{w} := \nabla v$.

This problem is equivalent to (note (8.5))

$$(v, \mathbf{w}) \in \text{dom}(A_\star), \quad (v, \mathbf{w}) = A_\star(v, \mathbf{w}), \quad B(v, \mathbf{w}) = \xi. \tag{8.8}$$

Since $|\langle \xi, \gamma^+ w \rangle_\Gamma| \lesssim \|\xi\|_{H^{-1/2}(\Gamma)} \|w\|_{H^1(\mathbb{R}^d \setminus \Gamma)}$ it follows that the norm of the solution operator for (8.7) is independent of the space Y_h . \square

Proposition 8.4 *The lifting \mathcal{E} from Proposition 8.3 is a bounded linear map $L^2(\Gamma) \rightarrow \mathcal{X}_{1/2} := [\mathcal{X}, \text{dom}(A)]_{1/2}$ with $\|\mathcal{E} \lambda\|_{\mathcal{X}_{1/2}} \leq C \|\lambda\|_{L^2(\Gamma)}$. The constant C depends only of Ω .*

Proof We will need spaces encoding homogeneous normal traces:

$$\begin{aligned} \mathbf{H}_0(\operatorname{div}, \Omega) &:= \{ \mathbf{w} \in \mathbf{H}(\operatorname{div}, \Omega) : \gamma_v^- \mathbf{w} = 0 \}, \\ \mathbf{H}_0(\operatorname{div}, \mathbb{R}^d \setminus \overline{\Omega}) &:= \{ \mathbf{w} \in \mathbf{H}(\operatorname{div}, \mathbb{R}^d \setminus \overline{\Omega}) : \gamma_v^+ \mathbf{w} = 0 \}. \end{aligned}$$

Let $\lambda \in L^2(\Gamma)$ be given. By applying Theorem A.4 to the exterior and setting $\tilde{\mathbf{w}} = 0$ inside, we can construct a function $\tilde{\mathbf{w}} \in \mathbf{H}(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)$ satisfying $\llbracket \gamma_v \tilde{\mathbf{w}} \rrbracket = \lambda$ and

$$\| \tilde{\mathbf{w}} \|_{\mathbf{H}(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)} + \| \tilde{\mathbf{w}} \|_{[L^2(\Omega), \mathbf{H}_0(\operatorname{div}, \mathbb{R}^d \setminus \overline{\Omega})]_{1/2}} \lesssim \| \lambda \|_{L^2(\Gamma)}. \tag{8.9}$$

Upon identifying the product of function spaces on Ω and $\mathbb{R}^d \setminus \overline{\Omega}$ with a function space on $\mathbb{R}^d \setminus \overline{\Gamma}$, we have

$$\mathbf{H}_0(\operatorname{div}, \Omega) \times \mathbf{H}_0(\operatorname{div}, \mathbb{R}^d \setminus \overline{\Omega}) \subseteq \mathbf{W}_h^0.$$

The product of interpolation spaces equals the interpolation of product spaces (cf. Lemma A.5); we can therefore also estimate:

$$\|(0, \tilde{\mathbf{w}})\|_{\mathcal{X}_{1/2}} \lesssim \| \tilde{\mathbf{w}} \|_{[L^2(\Omega), \mathbf{W}_h^0]_{1/2}} \lesssim \| \lambda \|_{L^2(\Gamma)}.$$

If we consider $(v, \mathbf{w}) := \mathcal{E}\lambda$, then $(v, \mathbf{w} - \tilde{\mathbf{w}}) \in \operatorname{dom}(A)$ by construction of the lifting. Thus we have

$$\begin{aligned} \|(v, \mathbf{w})\|_{\mathcal{X}_{1/2}} &\leq \|(v, \mathbf{w} - \tilde{\mathbf{w}})\|_{\mathcal{X}_{1/2}} + \|(0, \tilde{\mathbf{w}})\|_{\mathcal{X}_{1/2}} \\ &\leq \left(\|v\|_{H^1(\mathbb{R}^d \setminus \Gamma)} + \| \mathbf{w} - \tilde{\mathbf{w}} \|_{H(\operatorname{div}, \mathbb{R}^d \setminus \Gamma)} \right) + \|(0, \tilde{\mathbf{w}})\|_{\mathcal{X}_{1/2}}. \end{aligned}$$

The continuity of \mathcal{E} from Proposition 8.3 and (8.9) conclude the proof. □

Proposition 8.5 *If $g \in C^2([0, \infty); H^{-1/2}(\Gamma))$ satisfies $g(0) = \dot{g}(0) = 0$, then (8.6a–8.6c) has a unique strong solution.*

Proof Thanks to Propositions 8.2 and 8.3, this problem fits in the abstract framework described in [13], which proves existence and uniqueness of solution to (8.6a–8.6c). □

Propositions 8.2 and 8.3 have some consequences. First of all, Assumption 2.1 holds. Secondly, assuming $g(t) \in L^2(\Gamma)$, any solution to (2.1a–2.1c) with the above data ($F = 0, \Xi = g$) is in $\mathcal{X}_{1/2}$, and therefore, solutions to (8.6a–8.6c) take values in $\mathcal{X}_{1/2}$ as well. Finally, if $g \in C^s([0, \infty]; L^2(\Gamma))$ then $\mathcal{E}g \in C^s([0, \infty]; \mathcal{X}_{1/2})$.

We also need a regularity result that allows us to bound time derivatives of the solution in terms of the data. The continuity condition for the $(s + 2)$ -nd derivative of g in Proposition 8.6 can be relaxed to local integrability, but then the norms on the right-hand side of (8.10) have to be modified.

Proposition 8.6 *If $g \in C^{s+2}([0, \infty); L^2(\Gamma))$ satisfies $g^{(\ell)}(0) = 0$ for $\ell \leq s + 1$, then the unique solution to (8.6a–8.6c) satisfies*

- (a) $(v_h, \mathbf{w}_h) \in C^{s+1}([0, \infty); \mathcal{X})$,
- (b) $(v_h, \mathbf{w}_h) \in C^s([0, \infty); \mathcal{V})$ and $(v_h, \mathbf{w}_h) \in C^s([0, \infty); \mathcal{X}_{1/2})$,
- (c) For all $\ell \leq s$, there exists C , independent of the choice of Y_h , such that for all $t \geq 0$

$$\|(v_h^{(\ell)}(t), \mathbf{w}_h^{(\ell)}(t))\|_{\mathcal{X}_{1/2}} \leq C t \sum_{j=\ell}^{\ell+2} \max_{\tau \leq t} \|g^{(j)}(\tau)\|_{L^2(\Gamma)}. \tag{8.10}$$

Proof This result follows from [13, Theorem 3.1]. To see item (b), we note that (v_h, \mathbf{w}_h) is constructed by writing

$$(v_h(t), \mathbf{w}_h(t)) = (v_h^0(t), \mathbf{w}_h^0(t)) + \mathcal{E}g(t),$$

with $(v_h^0(t), \mathbf{w}_h^0(t)) \in \text{dom}(A)$. The statement then follows from Proposition 8.4. □

We now consider the RK approximation of (8.6a–8.6c) in a finite time interval $[0, T]$, which provides pairs of stage values $(V_{h,n}^k, \mathbf{W}_{h,n}^k) \in \mathcal{X}^m$ and step approximations $(v_{h,n}^k, \mathbf{w}_{h,n}^k) \in \mathcal{X}$. We then define

$$\{U_{h,n}^k\} = (\partial^k)^{-1}\{V_{h,n}^k\}, \quad u_{h,n}^k = r(\infty)u_{h,n}^k + \mathbf{b}^\top Q^{-1}U_{h,n}^k, \quad n \geq 0 \tag{8.11}$$

with $u_{h,0}^k = 0$ (see Lemma 2.6) and $\psi_{h,n}^k := \llbracket \gamma u_{h,n}^k \rrbracket$.

Proposition 8.7 *For sufficiently smooth g , with RK approximations using a method satisfying Assumption 2.II, and with α given by (6.1), for $nk \leq T$ we have the estimates*

$$\|u_h(t_n) - u_{h,n}^k\|_{L^2(\mathbb{R}^d \setminus \Gamma)} \leq C(T^2 + \alpha) k^{\min\{q+3/2+\alpha, p\}} \sum_{\ell=q}^{p+3} \max_{t \leq T} \|g^{(\ell)}(t)\|_{L^2(\Gamma)}, \tag{8.12}$$

and

$$\|u_h(t_n) - u_{h,n}^k\|_{H^1(\mathbb{R}^d \setminus \Gamma)} + \|\psi_h(t_n) - \psi_{h,n}^k\|_{H^{1/2}(\Gamma)} \leq C(T^2 + \alpha) k^{\min\{q+1/2+\alpha, p\}} \sum_{\ell=q}^{p+3} \max_{t \leq T} \|g^{(\ell)}(t)\|_{L^2(\Gamma)}. \tag{8.13}$$

The constants depend on Γ and the Runge–Kutta method, but do not depend on T or on the choice of Y_h .

Proof We will use Theorems 3.4 as well as Propositions 3.1 and 3.2. We note that $\rho_k(T) \leq 1$ by Lemma 7.1 and Proposition 8.2. Also, with the \mathcal{E} operator of Proposition 8.3, we have

$$\|\mathcal{E}\Xi^{(\ell)}\|_{\mathcal{X}_{1/2}} \leq C\|g^{(\ell)}\|_{L^2(\Gamma)}, \tag{8.14}$$

with C independent of Y_h . The bound (8.12) follows from Theorem 3.4, using (8.10) and (8.14) to estimate the right-hand side. The bound

$$\begin{aligned} \|\nabla u_h(t_n) - \nabla u_{h,n}^k\|_{L^2(\mathbb{R}^d \setminus \Gamma)} &= \|\mathbf{w}_h(t_n) - \mathbf{w}_{h,n}^k\|_{L^2(\mathbb{R}^d \setminus \Gamma)} \\ &\leq CT^2 k^{\min\{q+1/2+\alpha, p\}} \sum_{\ell=q+1}^{p+3} \max_{t \leq T} \|g^{(\ell)}\|_{H^{1/2}(\Gamma)} \end{aligned} \tag{8.15}$$

follows from Propositions 3.1 and 3.2, using (8.10) for the estimate in terms of the data. The $H^1(\mathbb{R}^d \setminus \Gamma)$ estimate (8.13) is then a direct consequence of (8.12) and (8.15). The estimate for $\psi_h - \psi_h^k$ follows from the standard trace theorem. \square

8.3 Scattering

As it does not incur much difficulty, we cover both the exterior scattering problem, which is an exterior Neumann problem, as well as the interior Neumann problem. In order to do so, we define the domain $\Omega^+ := \mathbb{R}^d \setminus \overline{\Omega}$ and $\Omega^- := \Omega$ and distinguish the problems by adding the superscripts + or - to the functions involved. We stay in the geometric setting of the previous section. Assume that $\mathbf{d} \in \mathbb{R}^d$ is a unit vector (direction of propagation) and that $c \in \mathbb{R}$ is such that $\Omega \subset \{\mathbf{x} \in \mathbb{R}^d : \mathbf{x} \cdot \mathbf{d} > c\}$. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a function such that $\phi(r) = 0$ for all $r \geq c$. The incident wave $u^{\text{inc}}(\mathbf{x}, t) := \phi(\mathbf{x} \cdot \mathbf{d} - t)$ propagates in the direction \mathbf{d} at unit speed and has not reached the scatterer given by Ω at time $t = 0$. The data for our problem will be the function $g : [0, T] \rightarrow L^2(\Gamma)$ given by $g(t) := -\partial_\nu^\pm u^{\text{inc}}(\cdot, t)$.

The problems under consideration are: Find $u^\pm : [0, T] \rightarrow H^1(\Omega^\pm)$ satisfying

$$\ddot{u}^\pm(t) = \Delta u^\pm(t), \quad u^\pm(0) = \dot{u}^\pm(0) = 0, \quad \partial_\nu^\pm u^\pm(t) = g(t),$$

so that $\partial_\nu^\pm(u^\pm + u^{\text{inc}}) = 0$. (Note that we can take the trace of the normal derivative of the incident wave, since it is locally smooth.) The exterior problem (posed on Ω^+) is the classical sound soft scattering problem of the incident wave u^{inc} .

A direct formulation for solving this problem is obtained by extended the function by zero to the complement of the domain of interest. That is, we solve:

$$\ddot{u}^\pm(t) = \Delta u^\pm(t) \text{ in } \Omega^\pm, \quad \llbracket \partial_\nu u^\pm(t) \rrbracket = g(t), \quad \partial_\nu^\mp u^\pm(t) = 0 \tag{8.16}$$

with $u^\pm(0) = \dot{u}^\pm(0) = 0$. By imposing some additional hypotheses on the growth of g (which is needed to have a well-defined distributional Laplace transform), we can represent the solution to (8.16) as $u^\pm = \mp S(\partial)g - D(\partial)\psi^\pm$, where $\psi^\pm := \llbracket \gamma u^\pm \rrbracket$. Note that, to be precise with the use of weak distributional definitions, all functions have to be extended by zero to $t < 0$ (we say that they are causal) and the time interval is extended to infinity.

Taking the trace in this representation formula, the solution of (8.16) can be found by solving an equation for ψ^\pm and then postprocessing with the potential operators:

$$W(\partial)\psi^\pm = (1/2 \pm K'(\partial))g, \quad u^\pm = \mp S(\partial)g - D(\partial)\psi^\pm, \tag{8.17}$$

and we still have that $\psi^\pm = \llbracket \gamma u^\pm \rrbracket$.

For simplicity of notation, we will skip the indices \pm for the different functions from now on. We can equivalently write (8.16) and the equivalent (8.17) by using the variables $v := \dot{u}$ and $\mathbf{w} := \nabla u$. We note that $u = \partial^{-1}v$ and $\psi = \partial^{-1}\llbracket \gamma v \rrbracket$. Here, (v, \mathbf{w}) solve (we restrict t to the interval $[0, T]$ again)

$$\dot{v}(t) = \nabla \cdot \mathbf{w}(t), \quad \dot{\mathbf{w}}(t) = \nabla v(t), \quad \llbracket \gamma_\nu \mathbf{w}(t) \rrbracket = g(t), \quad v(0) = 0, \quad \mathbf{w}(0) = \mathbf{0},$$

that is, (8.6a–8.6c) with $Y_h = H^{1/2}(\Gamma)$.

For the discretization, we consider a finite dimensional space Y_h and the Galerkin approximation to (8.17), so that we look for $\psi_h : \mathbb{R} \rightarrow X_h$ causal such that

$$\begin{aligned} \langle W(\partial) \psi_h, \mu \rangle_\Gamma &= \langle (1/2 \pm K^t(\partial))g, \mu \rangle_\Gamma \quad \forall \mu \in Y_h, \\ u_h &:= \mp S(\partial)g - D(\partial)\psi_h. \end{aligned} \tag{8.18}$$

The functions $v_h := \dot{u}_h$ and $\mathbf{w}_h := \nabla u_h$ satisfy (8.6a–8.6c). The difference between the solutions of (8.16) and (8.18) can be studied by comparing the solutions to (8.6a–8.6c) when $Y_h = H^{1/2}(\Gamma)$ and when Y_h is a finite dimensional space, see [13] for details. For our purposes, it is sufficient to note that we get quasi-optimal estimates for the discretization in space.

Discretization in time is performed by applying convolution quadrature to (8.18). The fully discrete solution reads

$$\begin{aligned} \langle W(\partial^k) \Psi_h, \mu \rangle_\Gamma &= \langle (1/2 \pm K^t(\partial^k))g, \mu \rangle_\Gamma \quad \forall \mu \in Y_h, \\ U_h &:= \mp S^t(\partial^k)g - D(\partial^k)\Psi_h. \end{aligned} \tag{8.19}$$

The approximations ψ_h^k and u_h^k are then computed by the usual post-processing, i.e.,

$$\begin{aligned} \psi_{h,0}^k &:= 0, & \psi_{h,n+1}^k &= r(\infty)\psi_{h,n}^k + \mathbf{b}^T \mathcal{Q}^{-1} \Psi_{h,n}^k, \\ u_{h,0}^k &:= 0, & u_{h,n+1}^k &= r(\infty)u_{h,n}^k + \mathbf{b}^T \mathcal{Q}^{-1} U_{h,n}^k. \end{aligned}$$

Lemma 8.8 *The sequences u_h^k and ψ_h^k computed via (8.19) coincide with the Runge–Kutta approximations to (8.6a–8.6c) and their traces respectively.*

Proof The details of the computation can be found in the appendix of [23]. The basic idea is to take the Z-transform and show that both approaches solve the matrix-valued Helmholtz problem (4.2a–4.2c). □

This gives the following immediate corollary, representing an *a priori* bound for the fully discrete method:

Corollary 8.9 *Let the assumptions of Proposition 8.7 hold. Then for u_h and ψ_h , approximated using convolution quadrature, we can estimate:*

$$\begin{aligned} \|u_h(t_n) - u_{h,n}^k\|_{H^1(\mathbb{R}^d \setminus \Gamma)} + \|\psi_h(t_n) - \psi_{h,n}^k\|_{H^{1/2}(\Gamma)} \\ \leq C(1 + T^2)k^{\min\{q+1/2+\alpha,p\}} \sum_{\ell=q}^{p+3} \max_{t \leq T} \|g^{(\ell)}(t)\|_{L^2(\Gamma)}. \end{aligned} \tag{8.20}$$

The constants depend on Γ and the Runge–Kutta method, but do not depend on T or on the choice of Y_h .

Remark 8.10 There is another approach for analyzing convolution quadrature methods, which is based on estimates in the Laplace domain. It can be shown that the Neumann-to-Dirichlet map, realized by the boundary integral Eq. (8.19), satisfies a bound of the form

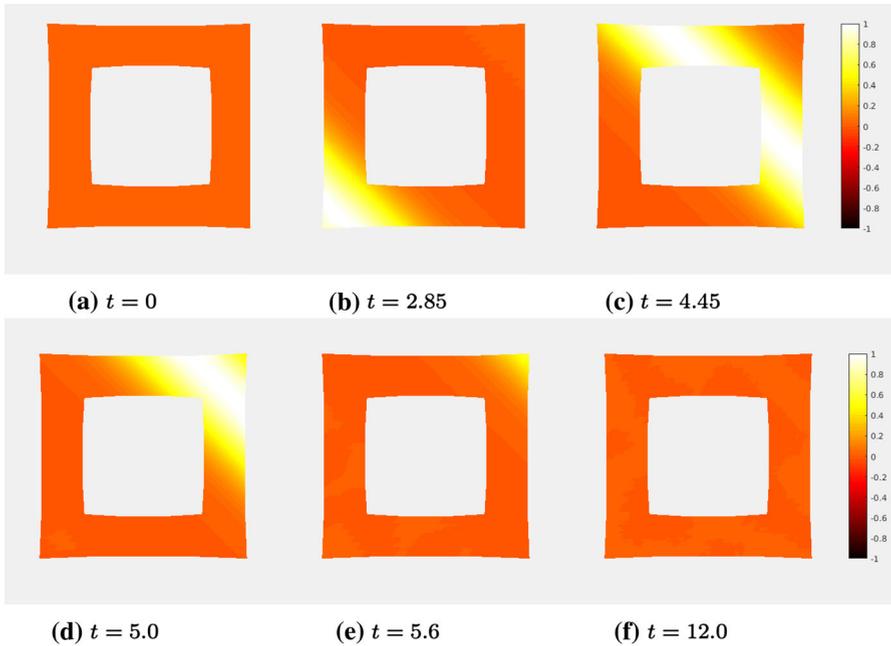


Fig. 1 Snapshots of the simulation at $t = 0, t = 2.85, t = 4.45, t = 5.0, t = 5.6, t = 12$

$$\|W(s)^{-1}(1/2 \pm K^t(s))\widehat{g}\|_{H^{1/2}(\Gamma)} \lesssim \frac{|s|}{\text{Re}(s)} \|\widehat{g}\|_{H^{-1/2}(\Gamma)}, \quad \forall \widehat{g} \in H^{-1/2}(\Gamma),$$

see [20, Appendix 2]. Applying the abstract theory of [2] then implies convergence rate $\min(q + 1, p)$ for the boundary data ψ_h . Modifying their proof, one can also get for $\widehat{g} \in L^2(\Gamma)$ that

$$\|W(s)^{-1}(1/2 \pm K^t(s))\widehat{g}\|_{H^{1/2}(\Gamma)} \lesssim \frac{|s|^{1/2}}{\text{Re}(s)} \|\widehat{g}\|_{L^2(\Gamma)},$$

which would yield the same convergence rate as Corollary 8.20, but without insight into the dependence on the end-time T . ■

8.4 Numerical example

We solve (8.19) on a “hollow square”, as depicted in Fig. 1, and focus on the interior Neumann problem, i.e. computing $\psi^- =: \psi$. The geometry was chosen to be non-convex and not simply connected, in order to test if the rate observed is a general result, or if our estimates might prove sharp in some situation.

We prescribe the exact solution as a traveling wave, given by

$$u(\mathbf{x}, t) := \phi(\mathbf{x} \cdot \mathbf{d} - t),$$

$$\phi(s) := \cos(\pi s/2) \exp(-4(s_0 - s)^2).$$

$s_0 := 4$ is chosen so that $\phi(0)$ is sufficiently small in the domain. We set $\mathbf{d} := [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]^\top$ and solve up to an end time of $T = 12$. An approximation of the $H^{1/2}$ -error is computed via

$$\left\langle W(1) \left(\psi_{h,n}^k - \Pi_{L^2} \psi(t_n) \right), \psi_{h,n}^k - \Pi_{L^2} \psi(t_n) \right\rangle_T,$$

i.e., we compare to the L^2 -projection of the exact solution. Since we are interested in the convergence rate with respect to the timestep size k , we consider a fixed, but sufficiently fine mesh.

We used 3 and 5 stage Radau IIA methods, with orders (q, p) of $(3, 5)$ and $(5, 9)$, respectively (see [16] for their definition). While their strong damping properties are not advantageous for wave propagation problems, they nevertheless are the standard method used with convolution quadrature. This is in part due to the fact that the standard theory (see, e.g., [2]) makes some assumptions not satisfied by the Gauss methods. A more detailed analysis of the dissipation and dispersion of the Radau methods was performed in [7, Sect. 4.3], showing that higher order Runge–Kutta methods possess favorable properties compared to their low order brethren.

Our theory predicts convergence rates of 4.5 and 6.5. In Fig. 2, we observe a rate that is closer to 5 and 8. This means that (just like the standard Laplace-domain estimates) our estimates do not appear to be sharp in this case. Further investigations into the cause of this phenomenon are required. Results trying to explain this phenomenon, initially prompted by the work on this article, can be found in [24] but with a different model problem.

8.5 The heat equation

In this section, as an example where our estimates turn out to be sharp, we consider a heat conduction problem and will apply Theorem 3.5 to get convergence of the boundary trace. The physical situation is a body $\Omega \subset \mathbb{R}^d$ that is held at a given temperature distribution and radiates heat into a medium $\Omega^+ := \mathbb{R}^d \setminus \overline{\Omega}$. We make the simplifying assumption that at $t = 0$ the temperature is 0. Since the problem is posed on an unbounded domain, it is a

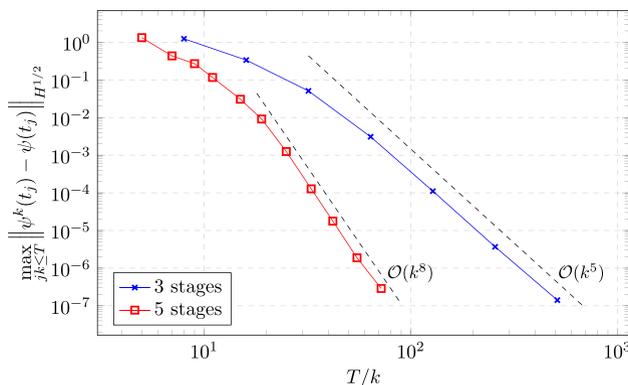


Fig. 2 Performance of Radau IIA methods for the wave equation, cf. Sect. 8.4

good candidate for boundary integral equations, while being simple enough to showcase our more general results. We only briefly give the mathematical setting. More details and a more involved physical example can be found in [27]. The setting is as follows: find $u : \mathbb{R}_+ \rightarrow H^1_{\Delta}(\Omega^+)$ such that

$$\dot{u} = \Delta u \quad \text{in } \mathbb{R}^d \setminus \overline{\Omega}, \tag{8.21a}$$

$$u(t)|_{\Gamma} = g(t) \quad \text{on } \Gamma := \partial\Omega, \tag{8.21b}$$

$$u(0) = 0 \quad \text{in } \mathbb{R}^d \setminus \overline{\Omega}. \tag{8.21c}$$

It is well-known that Δ with homogeneous Dirichlet boundary conditions generates an analytic semigroup (see e.g. [26, Sect. 7.2]) on $L^2(\mathbb{R}^d \setminus \overline{\Omega})$. The rest of our assumptions are also easily checked. We summarize:

- (i) $\text{dom}(A_{\star}) = \{u \in H^1(\mathbb{R}^d \setminus \overline{\Omega}) : \Delta u \in L^2(\mathbb{R}^d \setminus \overline{\Omega})\}$,
- (ii) $B : H^1(\mathbb{R}^d \setminus \overline{\Omega}) \rightarrow H^{1/2}(\Gamma) =: \mathcal{M}$, $Bv := \gamma^+v$ (using the standard trace operator).

In order to derive the boundary integral formulation, we take the Laplace transform of (8.21a), giving for $\kappa := \sqrt{s}$:

$$-\Delta \widehat{u}(s) + \kappa^2 \widehat{u}(s) = 0,$$

which is Helmholtz’s equation for a complex wave number κ . We make an ansatz of the form $\widehat{u} = S(\kappa)\widehat{\lambda}$ for some unknown density $\widehat{\lambda}$, which can be determined by applying the trace operator, giving the equation $V(\kappa)\widehat{\lambda} = \mathcal{L}(g)$.

Transforming back, and using the definition $V_{\kappa}(s) := V(\sqrt{s})$, we get the formulation:

$$[V_{\kappa}(\partial)\lambda](t) = g(t) \quad \forall t > 0.$$

The solution u can then be recovered by computing $u = S_{\kappa}(\partial)$, where $S_{\kappa}(s) := S(\sqrt{s})$.

The discrete version of this is then given by solving

$$V_{\kappa}(\partial^k)A^k = g. \tag{8.22}$$

It can be shown that plugging the discrete solution into the representation formula $U^k := S_{\kappa}(\partial^k)A^k$ gives back the Runge–Kutta approximation of (8.21a–8.21c). The approximations at the endpoints $t_n = nk$, denoted by λ^k and u^k respectively can be computed by the usual post-processing. We refer to the appendix of [23] for an analogous computation in the context of the Schrödinger equation, which easily transfers to our situation. For simplicity, we do not consider any discretization in space. A Galerkin approach could easily be included into the analysis, analogously to Sect. 8.2.

From the definition $A := A_{\star}|_{\ker(B)}$ we get $\text{dom}(A) := \{u \in H^1(\mathbb{R}^d \setminus \overline{\Omega}) : \Delta u \in L^2(\mathbb{R}^d \setminus \overline{\Omega}), \gamma^+u = 0\}$. We need the following analog of Proposition 8.4:

Proposition 8.11 $\text{dom}(A_{\star}) \subseteq [L^2(\mathbb{R}^d \setminus \overline{\Omega}), \text{dom}(A)]_{\mu, \infty}$ for $\mu \in [0, 1/4]$.

Proof It is easy to see that $H^2_0(\mathbb{R}^d \setminus \overline{\Omega}) \subseteq \text{dom}(A)$.

Using the Besov spaces introduced in Appendix A, we can write, if $\mu \leq 1/4$:

$$\begin{aligned}
 H^1(\mathbb{R}^d \setminus \overline{\Omega}) &\subseteq B_{2,1}^{2\mu}(\mathbb{R}^d \setminus \overline{\Omega}) \stackrel{Thm. A.1}{\subseteq} \widetilde{B}_{2,\infty}^{2\mu}(\mathbb{R}^d \setminus \overline{\Omega}) \\
 &= [L^2(\mathbb{R}^d \setminus \overline{\Omega}), H_0^1(\mathbb{R}^d \setminus \overline{\Omega})]_{2\mu,\infty} = [L^2(\mathbb{R}^d \setminus \overline{\Omega}), H_0^2(\mathbb{R}^d \setminus \overline{\Omega})]_{\mu,\infty},
 \end{aligned}$$

where in the last step, we used [22, Theorem B.9]. □

The convergence of our numerical method can then be analyzed quite easily using Proposition 3.2 and Theorem 3.5.

Theorem 8.12 *Let p and q denote the classical and stage order of the Runge–Kutta method used. Let $g \in \mathcal{C}^{p+3}([0, T], H^{1/2}(\Gamma))$ with $g^{(j)}(0) = 0$ for $j = 0, \dots, p + 2$. Set $\alpha := 1$ if the Runge–Kutta method satisfies Assumption 2.III and $\alpha := 0$ otherwise. Then the following estimate holds for the post-processed approximation:*

$$\|u^k(t_n) - u(t_n)\|_{L^2(\mathbb{R}^d \setminus \overline{\Omega})} \leq C(1 + T^2)k^{\min(q+\alpha+1/4,p)} \sum_{\ell=q+1}^{p+2} \max_{\tau \leq t_n} \|g^{(\ell)}\tau\|_{H^{1/2}(\Gamma)}. \tag{8.23}$$

Assume that the Runge–Kutta method used for discretization is stiffly accurate. Then the following estimates hold for the H^1 -norm:

$$\|u^k(t_n) - u(t_n)\|_{H^1(\mathbb{R}^d \setminus \overline{\Omega})} \leq C(1 + T^2)k^{r_1} \sum_{\ell=q+1}^{p+3} \max_{\tau \leq t_n} \|g^{(\ell)}\tau\|_{H^{1/2}(\Gamma)}, \tag{8.24}$$

with

$$r_1 := \begin{cases} q + \alpha - 1/4 & \text{for } q < p - 1, \\ q - 1/4 & \text{for } q = p - 1 \text{ and } \alpha = 0, \\ q + 5/8 & \text{for } q = p - 1 \text{ and } \alpha = 1, \\ q + \frac{\alpha - 1}{2} & \text{for } q = p. \end{cases}$$

For the density, we get:

$$\|\lambda^k(t_n) - \lambda(t_n)\|_{H^{-1/2}(\Gamma)} \leq C(1 + T^2)k^{r_\lambda} \sum_{\ell=q}^{p+1} \max_{\tau \leq t_n} \|g^{(\ell)}\tau\|_{H^{1/2}(\Gamma)}, \tag{8.25}$$

where the rate r_λ is given by:

$$r_\lambda := \begin{cases} q + \alpha - 1/2 & \text{for } q < p - 1, \\ q - 1/2 & \text{for } q = p - 1 \text{ and } \alpha = 0, \\ q + \frac{7}{16} & \text{for } q = p - 1 \text{ and } \alpha = 1, \\ q + \frac{3}{4}(\alpha - 1) & \text{for } q = p. \end{cases}$$

Proof We first note that we can control the derivatives $u^{(\ell)}$ by the data. This can be done completely analogous to Proposition 8.6 by the techniques of [13]. The estimates read:

$$\|u^{(\ell)}(t)\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})} \leq Ct \sum_{j=\ell}^{\ell+1} \max_{\tau \leq t} \|g^{(j)}(\tau)\|_{H^{1/2}(\Gamma)}.$$

For simplicity of notation, we only consider the case $q < p - 1$. All the other cases follow analogously but giving different rates when applying the abstract theory. By Proposition 8.11, we can apply Propositions 3.1 or 3.2 with $\mu = 1/4$, depending on whether we are in the setting $\alpha = 0$ or $\alpha = 1$. This gives estimate (8.23).

Applying Theorem 3.5, we get the following convergence in the graph norm of A_\star :

$$\|\Delta u^k(t_n) - \Delta u(t_n)\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})} \leq C(1 + T^2)k^{q+\alpha-1+1/4} \sum_{\ell=q}^{p+3} \max_{\tau \leq t_n} \|g^{(\ell)}\tau\|_{H^{1/2}(\Gamma)}. \tag{8.26}$$

Since for stiffly accurate RK-methods, u^k also satisfies the boundary conditions (it is the just last entry of the stage vector) we get from the Dirichlet-boundary conditions that $\gamma^+ u(t_n) = g(t_n) = \gamma^+ u^k(t_n)$. Therefore, integration by parts and the Cauchy-Schwarz inequality give:

$$\begin{aligned} \|\nabla u^k(t_n) - \nabla u(t_n)\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})}^2 &= -(\Delta u^k(t_n) - \Delta u(t_n), u^k(t_n) - u(t_n))_{L^2(\mathbb{R}^d \setminus \bar{\Omega})} \\ &\leq \|\Delta u^k(t_n) - \Delta u(t_n)\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})} \|u^k(t_n) - u(t_n)\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})}. \end{aligned}$$

Estimate (8.24) then follows from (8.23) and (8.26). For the estimate (8.25) of the density, we fix $\xi \in H^{1/2}(\Gamma)$, and let v denote a lifting to $H^1(\mathbb{R}^d)$. We calculate

$$\begin{aligned} \langle \lambda - \lambda^k, \xi \rangle_\Gamma &= (-\Delta u + \Delta u^k, v)_{L^2(\Omega)} + (\nabla u - \nabla u^k, \nabla v)_{L^2(\mathbb{R}^d \setminus \bar{\Omega})} \\ &\leq (k^{1/2} \|\Delta u - \Delta u^k\|_{L^2(\Omega)} + \|\nabla u - \nabla u^k\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})}) \\ &\quad \times (k^{-1/2} \|v\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\mathbb{R}^d \setminus \bar{\Omega})}). \end{aligned}$$

We are still free to choose the precise lifting v . Doing so as in [31, Proposition 2.5.1], we get

$$\inf\{k^{-1/2} \|v\|_{L^2(\mathbb{R}^d)} + \|\nabla v\|_{L^2(\mathbb{R}^d)} : v \in H^1(\mathbb{R}^d), \gamma v = \xi\} \lesssim \max\{1, k^{-1/4}\} \|\xi\|_{H^{1/2}(\Gamma)}.$$

The result then follows from the previous estimates. □

Remark 8.13 Note that in the cases $q = p - 1$ with $\alpha = 1$ and $q = p$ with $\alpha = 0$, the rates r_1 and r_λ in Theorem 8.12 are sharp from what can be extracted from Theorem 3.5 and Propositions 3.1 and 3.2. Nevertheless, we expect it to be possible to extract better rates from a more explicit investigation of these limiting cases.

8.5.1 Numerical example

In order to demonstrate that the estimate (8.25) is sharp, we consider a simple model problem. Following [34], we take Ω to be the unit sphere and consider a right-hand side $g(x, t)$ of the form

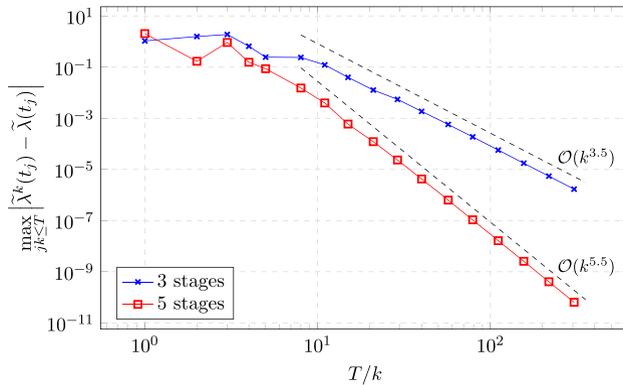


Fig. 3 Convergence for the density $\tilde{\lambda}$ for the heat conduction problem (cf. Sect. 8.5), comparing Radau IIA methods

$$g(x, t) := \psi(t)Y_n^m(x),$$

where Y_n^m is the spherical harmonic of degree n and order m . It is well-known that the spherical harmonics are eigenfunctions of the pertinent boundary integral operators. Most notably for us, we have

$$V(s)Y_n^m = \mu_n(s)Y_n^m \quad \text{with} \quad \mu_n(s) := -s j_n(is) h_n^{(1)}(is),$$

where j_n denotes the spherical Bessel functions and $h_n^{(1)}$ is the spherical Hankel function of the first kind. Due to this relation, solving (8.22) becomes a purely one dimensional problem, i.e., we can write $\lambda(x, t) = \tilde{\lambda}(t)Y_n^m(x)$ and the solution can be easily computed to very high accuracy. For our experiments we chose $n = 2$.

We compare the 3-stage and 5-stage Radau IIA methods (see [16] for their definitions). These methods have stage orders 3 and 5 respectively and both are stiffly accurate and satisfy Assumption 2.III. We therefore expect convergence rates for the density λ of order 3.5 and 5.5. Since the exact solution is not available, we compute the difference to an approximation with step size $k/4$ and use this as an approximation to the discretization error. The results can be seen in Fig. 3. We observe that the results are in good agreement with our predictions.

Acknowledgements Work of A.R. was financially supported by the Austrian Science Fund (FWF) through the doctoral school “Dissipation and Dispersion in Nonlinear PDEs” (project W1245) and the research program “Taming complexity in partial differential systems” (Grant SFB F65). FJS is partially supported by NSF-DMS grant 1818867. Part of this work was developed while FJS was a Visiting Professor at TU Wien.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Funding Open access funding provided by Austrian Science Fund (FWF).

A Interpolation of Sobolev spaces

In this appendix we prove that in Lipschitz domains and for certain parameters μ , the spaces $[L^2(\Omega), H_0^1(\Omega)]_\mu$ contain functions with non-vanishing boundary conditions. Such estimates are the main ingredient when determining the convergence rate of Runge–Kutta methods using the theory developed in the previous sections. For $\mu < 1/2$, it is well-known that the fractional Sobolev spaces $H^\mu(\Omega) = [L^2(\Omega), H^1(\Omega)]_{\mu,2}$ and $\tilde{H}^\mu(\Omega) = [L^2(\Omega), H_0^1(\Omega)]_{\mu,2}$ coincide (see e.g. [22, Theorem 3.40] together with the results in [22, Appendix B] to identify the Sobolev spaces with the interpolation space). We prove that when interpolating using the index ∞ , the critical value $\mu = 1/2$ is also admissible, provided that some further regularity is provided.

In order to state our result, we need additional notation, notably we define interpolation spaces for $q \in [1, \infty)$ as

$$\|u\|_{[\mathcal{X}_0, \mathcal{X}_1]_{\mu,q}}^q := \int_0^\infty t^{-\mu} \left[\inf_{v \in \mathcal{X}_1} \|u - v\|_{\mathcal{X}_0} + t\|v\|_{\mathcal{X}_1} \right]^q \frac{dt}{t}, \tag{A.1}$$

and introduce the following Besov spaces:

$$B_{2,q}^\mu(\Omega) := [L^2(\Omega), H^1(\Omega)]_{\mu,q} \quad \text{and} \quad \tilde{B}_{2,q}^\mu(\Omega) := [L^2(\Omega), H_0^1(\Omega)]_{\mu,q}. \tag{A.2}$$

For $t > 0$, we define the strip

$$\Omega_t := \{x \in \Omega : \text{dist}(x, \partial\Omega) < t\}, \tag{A.3}$$

which will play an important role in the following proofs.

Theorem A.1 *Let Ω be either a bounded Lipschitz domain or the complement of a bounded Lipschitz domain. Fix $\mu \in (0, 1/2]$. Then*

$$B_{2,1}^\mu(\Omega) \subseteq \tilde{B}_{2,\infty}^\mu(\Omega)$$

with equivalent norms. The implied constant depends on Ω and μ .

Proof For simplicity, assume that Ω is bounded. We focus on the case $\mu = 1/2$, since the case $0 \leq \mu < 1/2$ follows from $H^\mu(\Omega) = \tilde{H}^\mu(\Omega)$ (see [22, Theorem 3.40(i) and Theorem 3.33]) combined with the embeddings

$$B_{2,1}^\mu(\Omega) \subset B_{2,2}^\mu(\Omega) = H^\mu(\Omega) \subset \tilde{H}^\mu(\Omega) \subset \tilde{B}_{2,\infty}^\mu(\Omega).$$

Consider $u \in B_{2,1}^{1/2}(\Omega)$. For each $t > 0$, we select $v(t) \in H^1(\Omega)$ almost realizing the infimum appearing in the interpolation norm, i.e.,

$$\|u - v(t)\|_{L^2(\Omega)} + t\|v(t)\|_{H^1(\Omega)} \leq 2 \inf_{w \in H^1(\Omega)} \left(\|u - w\|_{L^2(\Omega)} + t\|w\|_{H^1(\Omega)} \right).$$

By [6, Lemma], the following estimate holds for all $t > 0$:

$$\|v(t)\|_{B_{2,1}^\mu(\Omega)} \leq 3\|u\|_{B_{2,1}^{1/2}(\Omega)}.$$

We also note that

$$\|u - v(t)\|_{L^2(\Omega)} + t\|v(t)\|_{H^1(\Omega)} \lesssim t^{1/2}\|u\|_{[L^2(\Omega), H^1(\Omega)]_{1/2, \infty}} \lesssim t^{1/2}\|u\|_{B_{2, \infty}^{1/2}(\Omega)} \lesssim t^{1/2}\|u\|_{B_{2,1}^{1/2}(\Omega)}.$$

We consider a smooth cutoff function $\chi_t : \Omega \rightarrow [0, 1]$ satisfying:

$$\chi_t(x) \equiv 0 \text{ on } \Omega_t, \quad \chi_t(x) \equiv 1 \text{ on } \Omega \setminus \Omega_{2t} \quad \text{and} \quad \|\nabla \chi_t\|_{L^\infty} \lesssim t^{-1}. \tag{A.4}$$

We then define $\tilde{v}(t) := \chi_t v(t) \in H_0^1(\Omega)$ and calculate:

$$\begin{aligned} \|u - \tilde{v}(t)\|_{L^2(\Omega)} &\leq \|u - v(t)\|_{L^2(\Omega)} + \|(1 - \chi_t)v(t)\|_{L^2(\Omega_{2t})} \\ &\lesssim \|u - v(t)\|_{L^2(\Omega)} + t^{1/2}\|v(t)\|_{B_{2,1}^{1/2}(\Omega)} \end{aligned}$$

where we used the fact that $1 - \chi_t$ vanishes on $\Omega \setminus \Omega_{2t}$ and applied [18, Lemma 2.1] to estimate the L^2 -norm there.

Similarly,

$$\begin{aligned} t\|\tilde{v}(t)\|_{H^1(\Omega)} &\lesssim t\|v\|_{H^1(\Omega)} + t\|(\nabla \chi_t)v(t)\|_{L^2(\Omega)} \lesssim t\|v(t)\|_{H^1(\Omega)} + \|v(t)\|_{L^2(\Omega_{2t})} \\ &\lesssim t\|v(t)\|_{H^1(\Omega)} + t^{1/2}\|v(t)\|_{B_{2,1}^{1/2}(\Omega)}. \end{aligned}$$

For the interpolation norm, we therefore get

$$\begin{aligned} \|u\|_{\tilde{B}_{2, \infty}^{1/2}(\Omega)} &\lesssim \text{ess sup}_{t > 0} \left[t^{-1/2} (\|u - v(t)\|_{L^2(\Omega)} + t\|v(t)\|_{H^1(\Omega)} + t^{1/2}\|v(t)\|_{B_{2,1}^{1/2}(\Omega)}) \right] \\ &\lesssim \|v(t)\|_{B_{2,1}^{1/2}(\Omega)} + \|u\|_{[L^2(\Omega), H^1(\Omega)]_{\mu, \infty}} \lesssim \|u\|_{B_{2,1}^{1/2}(\Omega)} + \|u\|_{B_{2, \infty}^{1/2}(\Omega)} \\ &\lesssim \|u\|_{B_{2,1}^{1/2}(\Omega)}. \end{aligned}$$

If Ω is the exterior of a bounded Lipschitz domain, the proof applies almost verbatim as all important steps can be localized to a neighborhood of the boundary. □

Remark A.2 The use of the second parameter ∞ in the interpolation norm is crucial for Theorem A.1 to hold in the case $\mu = 1/2$. For L^2 -based interpolation it is well-known that the interpolation space $[L^2(\Omega), H_0^1(\Omega)]_{1/2, 2}$ is the Lions-Magenes space $H_{00}^{1/2}(\Omega)$, see [35, Chapter 33], which is distinct from $H^{1/2}(\Omega)$.

When considering the Neumann problem in Sect. 8.3, we need to devise a lifting to a vector field with a given normal jump in L^2 . In general, such liftings do not have $B_{2,1}^{1/2}$ -regularity. Thus Theorem A.1 is not applicable. Instead, we have a modified construction.

Lemma A.3 *Let Ω be a bounded Lipschitz domain or the exterior of a bounded Lipschitz domain with boundary $\Gamma := \partial\Omega$. For $C > 0, c > 0$ fixed with c sufficiently small, define the non-tangential maximal function*

$$N(\nabla u)(x) := \sup_{y \in \Theta(x)} |\nabla u(y)|, \text{ where } \Theta(x) := \{y \in \Omega : |x - y| \leq \max(c, C\text{dist}(y, \Gamma))\}.$$

Let $u \in H^1(\Omega)$ be harmonic and satisfy $N(\nabla u) \in L^2(\Gamma)$.

Then for $t > 0$ we can bound the L^2 norm on strips Ω_t by

$$\|\nabla u\|_{L^2(\Omega_t)} \lesssim t^{1/2} \|N(\nabla u)\|_{L^2(\Gamma)}. \tag{A.5}$$

Proof We focus on a single chart in the parametrization of (a vicinity of) Γ . Let $\mathcal{O} \subseteq \Omega$ and $\mathcal{D} \subseteq \mathbb{R}^{d-1}$ be open, $\mathbf{r} \in \mathbb{R}^n$, $\varphi : \mathcal{D} \rightarrow \mathbb{R}$, $y_0 : \mathcal{D} \rightarrow \mathbb{R}$ such that we can write

$$\Omega_t \cap \mathcal{O} = \{(x, \varphi(x) + y\mathbf{r}) : x \in \mathcal{D}, \text{ and } y \in (0, y_0(x))\}.$$

By the Lipschitz assumption, we note that $y_0(x) \lesssim Ct$. Following the considerations in [10, Appendix A.4], one can see that as long as C in the definition of Θ is taken sufficiently large, we have that for all $x \in \mathcal{D}$

$$\{(x, \varphi(x) + y\mathbf{r}) : y \in (0, y_0(x))\} \subseteq \Theta((x, \varphi(x))).$$

We calculate

$$\begin{aligned} \|u\|_{L^2(\Omega_t \cap \mathcal{O})}^2 &= \int_{x \in \mathcal{D}} \int_{y=0}^{y_0(x)} |\nabla u(x, \varphi(x) + y\mathbf{r})|^2 dy dx \\ &\lesssim \int_{x \in \mathcal{D}} \int_{y=0}^{y_0(x)} (N(\nabla u)(x))^2 dy dx \\ &\leq t \int_{x \in \mathcal{D}} (N(\nabla u)(x))^2 dx \leq t \|N(\nabla u)\|_{L^2(\Gamma)}^2. \end{aligned} \tag{A.6}$$

Repeating the same calculation for all boxes needed to parametrize a neighborhood of Γ then concludes the proof. □

Theorem A.4 Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain or the exterior of a bounded Lipschitz domain and write $\mathbf{H}_0(\text{div}, \Omega) := \{\mathbf{w} \in \mathbf{H}(\text{div}, \Omega) : \gamma_\nu \mathbf{w} = 0\}$.

For every $g \in L^2(\Omega)$, there exists a function $\mathbf{w} \in \mathbf{H}(\text{div}, \Omega)$ such that

$$\gamma_\nu \mathbf{w} = g \quad \text{and} \quad \|\nabla \cdot \mathbf{w}\|_{H^1(\Omega)} + \|\mathbf{w}\|_{[L^2(\Omega), \mathbf{H}_0(\text{div}, \Omega)]_{1/2, \infty}} \lesssim \|g\|_{L^2(\Omega)} \tag{A.7}$$

with an implied constant depending only on Ω .

Proof For simplicity, assume that Ω is bounded. By performing an appropriate cutoff away from $\partial\Omega$, all arguments can be localized.

Step 1: Consider the case $\int_\Gamma g = 0$. Let u be the solution of the Neumann problem

$$\Delta u = 0 \text{ in } \Omega, \quad \partial_\nu u = g \text{ on } \partial\Omega, \quad \int_\Omega u = 0.$$

In addition to $u \in H^1(\Omega)$, by [17] (see also [10, Theorem A.6]), this harmonic function u also satisfies

$$\|N(\nabla u)\|_{L^2(\Gamma)} \leq \|g\|_{L^2(\Gamma)}.$$

For fixed $t > 0$ we again select a smooth cutoff function χ_t satisfying (A.4). We set $w := \nabla u$ and calculate using Lemma A.3:

$$\begin{aligned} \|w\|_{[L^2(\Omega), \mathbf{H}_0(\text{div}, \Omega)]_{1/2, \infty}} &\lesssim \text{esssup}_{t \geq 0} \left(t^{-1/2} [\|(1 - \chi_t)w\|_{L^2(\Omega)} + t\|\chi_t w\|_{\mathbf{H}(\text{div}, \Omega)}] \right) \\ &\lesssim \text{esssup}_{t \geq 0} t^{-1/2} \|\nabla u\|_{L^2(\Omega_{2r})} \lesssim \|N(\nabla u)\|_{L^2(\Gamma)} \lesssim \|g\|_{L^2(\Gamma)}. \end{aligned}$$

Step 2: In the case $\int_{\Gamma} g \neq 0$, the harmonic Neumann problem does not have a solution. Instead, we define u as the solution to

$$-\Delta u + u = 0 \text{ in } \Omega, \quad \partial_\nu u = g \text{ on } \partial\Omega,$$

and again set $w := \nabla u$. By construction we have $\nabla \cdot w = -u \in H^1(\Omega)$.

We decompose $u = u_0 + u_1$, where u_0 solves the full-space problem

$$-\Delta u_0 = -u \text{ in } \mathbb{R}^d$$

where u was extended by 0 outside of Ω . As $u \in L^2(\mathbb{R}^d)$, standard regularity theory gives $u_0 \in H^2(B)$ on any ball B and in particular $u_0 \in H^2(\Omega)$. In turn this yields $\partial_\nu u_0 \in L^2(\partial\Omega)$. By construction u_1 then solves

$$-\Delta u_1 = 0 \text{ in } \Omega, \quad \partial_\nu u_1 = g - \partial_\nu u_0 \text{ on } \partial\Omega.$$

As $g, \partial_\nu u_0 \in L^2(\partial\Omega)$ and $\int_{\Gamma} g - \partial_\nu u_0 = 0$, we can apply Step 1 to get

$$\nabla u_1 \in [L^2(\Omega), \mathbf{H}_0(\text{div}, \Omega)]_{1/2, \infty}.$$

Since $\nabla u_0 \in (H^1(\Omega))^d \subseteq (B_{2,1}^{1/2}(\Omega))^d$ we can apply Theorem A.1 to get

$$u_0 \in [(L^2(\Omega))^d, (H_0^1(\Omega))^d]_{1/2, \infty} \subset [L^2(\Omega), \mathbf{H}_0(\text{div}, \Omega)]_{1/2, \infty}$$

to conclude the proof of (A.7). □

The following lemma appears to be known in the community, see e.g. [5, Section 3.13, Exercise 4], but in order to be able to rigorously cite, we provide a short proof.

Lemma A.5 Let $X := (X_1, \dots, X_N)$ and $Y := (Y_1, \dots, Y_N)$, where X_j, Y_j are Banach spaces with continuous embedding $Y_j \subset X_j$, and the product space carries any l^p -norm. Fix $q \in [1, \infty]$ and $\theta \in (0, 1)$. Then, the product of the interpolation spaces coincides with the interpolation of the product spaces. Namely, the following estimate holds for all $x := (x_1, \dots, x_N) \in [X, Y]_{\theta, q}$:

$$N^{-1} \sum_{j=1}^N \|x_j\|_{[X_j, Y_j]_{\theta, q}} \leq \|x\|_{[X, Y]_{\theta, q}} \leq \sum_{j=1}^N \|x_j\|_{[X_j, Y_j]_{\theta, q}}.$$

Proof For $j \in \{1, \dots, N\}$, consider the operators S_j and T_j defined as

$$\begin{aligned}
 T_j : X &\rightarrow X_j & \text{and} & & S_j : X_j &\rightarrow X \\
 (x_1, \dots, x_N) &\mapsto x_j & & & x_j &\mapsto (0, \dots, x_j, \dots, 0).
 \end{aligned}$$

It is easy to see using the interpolation estimate (2.6) that

$$\|T_j x\|_{[X_j, Y_j]_{\theta, q}} \leq \|x\|_{[X, Y]_{\theta, q}} \quad \text{and} \quad \|S_j x_j\|_{[X, Y]_{\theta, q}} \leq \|x_j\|_{[X_j, Y_j]_{\theta, q}}.$$

We therefore calculate:

$$\sum_{j=1}^N \|x_j\|_{[X_j, Y_j]_{\theta, q}} = \sum_{j=1}^N \|T_j x\|_{[X_j, Y_j]_{\theta, q}} \leq \sum_{j=1}^N \|x\|_{[X, Y]_{\theta, q}} \leq N \|x\|_{[X, Y]_{\theta, q}}.$$

For the opposite direction, we observe that

$$\|x\|_{[X, Y]_{\theta, q}} = \left\| \sum_{j=1}^N S_j x_j \right\|_{[X, Y]_{\theta, q}} \leq \sum_{j=1}^N \|S_j x_j\|_{[X, Y]_{\theta, q}} \leq \sum_{j=1}^N \|x_j\|_{[X_j, Y_j]_{\theta, q}}. \quad \square$$

References

1. Alonso-Mallo, I., Palencia, C.: Optimal orders of convergence for Runge–Kutta methods and linear, initial boundary value problems. *Appl. Numer. Math.* **44**(1–2), 1–19 (2003)
2. Banjai, L., Lubich, Ch., Melenk, J.M.: Runge–Kutta convolution quadrature for operators arising in wave propagation. *Numer. Math.* **119**(1), 1–20 (2011)
3. Banjai, L., Laliena, A.R., Sayas, F.-J.: Fully discrete Kirchhoff formulas with CQ-BEM. *IMA J. Numer. Anal.* **35**(2), 859–884 (2015)
4. Banjai, L., Rieder, A.: Convolution quadrature for the wave equation with a nonlinear impedance boundary condition. *Math. Comput.* **87**(312), 1783–1819 (2018)
5. Bergh, J., Löfström, J.: *Interpolation Spaces. An Introduction.* Grundlehren der Mathematischen Wissenschaften, vol. 223. Springer, Berlin, New York (1976)
6. Bramble, J.H., Scott, R.: Simultaneous approximation in scales of Banach spaces. *Math. Comput.* **32**(144), 947–954 (1978)
7. Banjai, L., Schanz, M.: Wave propagation problems treated with convolution quadrature and BEM. In: Schanz, M. (ed.) *Fast Boundary Element Methods in Engineering and Industrial Applications: Lecture Notes in Applied and Computational Mechanics*, vol. 63, pp. 145–184. Springer, Heidelberg (2012)
8. Brenner, Ph, Thomée, V.: On rational approximations of semigroups. *SIAM J. Numer. Anal.* **16**(4), 683–694 (1979)
9. Crouzeix, M.: Sur les méthodes de Runge-Kutta pour l’approximation des problèmes d’évolution. In: *Computing methods in applied sciences and engineering (Second international symposium, Versailles, 1975), Part 1.* Lecture notes in Econom. and Math. systems, Vol. 134. Springer, Berlin, pp. 206–223 (1976)
10. Chandler-Wilde, S.N., Graham, I.G., Langdon, S., Spence, E.A.: Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering. *Acta Numer.* **21**, 89–305 (2012)
11. Engel, K.-J., Nagel, R.: *A Short Course on Operator Semigroups.* Springer, New York (2006)
12. Henrici, P.: *Applied and Computational Complex Analysis: Power Series—Integration—Conformal Mapping—Location of Zeros*, vol. 1. Wiley, New York (1988). (Reprint of the 1974 original)
13. Hassell, M.E., Qiu, T., Sánchez-Vizuet, T., Sayas, F.-J.: A new and improved analysis of the time domain boundary integral operators for the acoustic wave equation. *J. Integral Equ. Appl.* **29**(1), 107–136 (2017)
14. Hassell, M., Sayas, F.-J.: Convolution quadrature for wave simulations. *Numerical Simulation in Physics and Engineering: SEMA SIMAI*, vol. 9, pp. 71–159. Springer, Cham (2016)
15. Hassell, M.E., Sayas, F.-J.: A fully discrete BEM-FEM scheme for transient acoustic waves. *Comput. Methods Appl. Mech. Eng.* **309**, 106–130 (2016)

16. Hairer, E., Wanner, G.: Solving ordinary differential equations: II. Springer Series in Computational Mathematics: Stiff and differential-algebraic problems, vol. 14, 2nd edn. Springer, Berlin (2010)
17. Jerison, D.S., Kenig, C.E.: The Neumann problem on Lipschitz domains. *Bull. Am. Math. Soc.* **4**(2), 203–207 (1981)
18. Li, J., Melenk, J.M., Wohlmuth, B., Zou, J.: Optimal a priori estimates for higher order finite elements for elliptic interface problems. *Appl. Numer. Math.* **60**(1–2), 19–37 (2010)
19. Lubich, Ch., Ostermann, A.: Runge–Kutta methods for parabolic equations and convolution quadrature. *Math. Comp.* **60**(201), 105–131 (1993)
20. Laliena, A.R., Sayas, F.-J.: Theoretical aspects of the application of convolution quadrature to scattering of acoustic waves. *Numer. Math.* **112**(4), 637–678 (2009)
21. Lubich, Ch.: On the multistep time discretization of linear initial-boundary value problems and their boundary integral equations. *Numer. Math.* **67**(3), 365–389 (1994)
22. McLean, W.: *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge University Press, Cambridge (2000)
23. Melenk, J.M., Rieder, A.: Runge–Kutta convolution quadrature and FEM-BEM coupling for the time-dependent linear Schrödinger equation. *J. Integral Equ. Appl.* **29**(1), 189–250 (2017)
24. Melenk, J.M., Rieder, A.: On superconvergence of Runge–Kutta convolution quadrature for the wave equation. (2019). [arXiv:1904.00847](https://arxiv.org/abs/1904.00847) (in press in *Numer. Math.*)
25. Ostermann, A., Roche, M.: Runge–Kutta methods for partial differential equations and fractional orders of convergence. *Math. Comp.* **59**(200), 403–420 (1992)
26. Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences, vol. 44. Springer, New York (1983)
27. Qiu, T., Rieder, A., Sayas, F.-J., Zhang, S.: Time-domain boundary integral equation modeling of heat transmission problems. *Numer. Math.* **143**(1), 223–259 (2019)
28. Rieder, A.: *Convolution quadrature and boundary element methods in wave propagation : a time domain point of view*. PhD thesis, Technische Universität Wien (2017)
29. Rieder, A., Sayas, F.-J., Melenk, J. M.: Time domain boundary integral equations and convolution quadrature for scattering by composite media (2020). [arXiv:2010.14162](https://arxiv.org/abs/2010.14162)
30. Riesz, F., Nagy, S.-B.: *Functional analysis*. Dover Books on Advanced Mathematics. Dover Publications Inc, New York (1990)
31. Sayas, F.-J.: Retarded potentials and time domain boundary integral equations. In: Sayas, F. (ed.) *Springer Series in Computational Mathematics: a Road Map*, vol. 50. Springer, Cham (2016)
32. Sauter, S.A., Schwab, Ch.: *Boundary element methods*. Springer Series in Computational Mathematics: Translated and Expanded from the 2004 German Original, vol. 39. Springer, Berlin (2011)
33. Steinbach, O.: *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary elements*, Translated from the 2003 German Original. Springer, New York (2008)
34. Sauter, S., Veit, A.: Retarded boundary integral equations on the sphere: exact and numerical solution. *IMA J. Numer. Anal.* **34**(2), 675–699 (2014)
35. Tartar, L.: *An Introduction to Sobolev Spaces and Interpolation spaces: Lecture Notes of the Unione Matematica Italiana*, vol. 3. Springer, Berlin (2007)
36. Triebel, H.: *Interpolation theory, function spaces, differential operators*, 2nd edn. Johann Ambrosius Barth, Heidelberg (1995)
37. von Neumann, J.: Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. *Math. Nachr.* **4**, 258–281 (1951)

Affiliations

Alexander Rieder¹ · Francisco-Javier Sayas² · Jens Markus Melenk³ 

¹ Fakultät für Mathematik, Universität Wien, Vienna, Austria

² Department of Mathematical Sciences, University of Delaware, Newark, DE, USA

³ Institut für Analysis und Scientific Computing, Technische Universität Wien, Vienna, Austria