A-STABLE RUNGE–KUTTA METHODS FOR SEMILINEAR EVOLUTION EQUATIONS

MARCEL OLIVER AND CLAUDIA WULFF

Abstract. We consider semilinear evolution equations for which the linear part generates a strongly continuous semigroup and the nonlinear part is sufficiently smooth on a scale of Hilbert spaces. In this setting, we prove the existence of solutions which are temporally smooth in the norm of the lowest rung of the scale for an open set of initial data on the highest rung of the scale. Under the same assumptions, we prove that a class of implicit, A-stable Runge–Kutta semidiscretizations in time of such equations are smooth as maps from open subsets of the highest rung into the lowest rung of the scale. Under the additional assumption that the linear part of the evolution equation is normal or sectorial, we prove full order convergence of the semidiscretization in time for initial data on open sets. Our results apply, in particular, to the semilinear wave equation and to the nonlinear Schrödinger equation.

Contents

1. Introduction

We study numerical schemes for evolution equations on Hilbert spaces by first looking at the properties of a semidiscretization in time only; discretization in space is then treated as a perturbation within the Hilbert space setting. This

Date: August 10, 2012.
Key words and phrases. Semilinear evolution equation, smoothness of semiflow, A-stable Runge Kutta semidiscretization in time.

approach was introduced by Rothe [28]. When successful, results so obtained are naturally uniform in the spatial discretization parameter. In contrast, when first discretizing in space, the resulting finite dimensional system of ordinary differential equations (ODEs) can be treated with ODE-based techniques which avoids the difficulty arising from the analysis of equations on infinite-dimensional spaces, but where uniformity in the spatial mesh size is not immediate.

In this paper, we consider semilinear evolution equations of the form

$$\partial_t U = AU + B(U)$$

on a Hilbert space $\mathcal{Y}$. The linear operator A is assumed to generate a strongly continuous, not necessarily analytic semigroup and B is a bounded nonlinear operator on $\mathcal{Y}$. The examples we have in mind are semilinear Hamiltonian evolution equations such as the semilinear wave equation or the nonlinear Schrödinger equation with periodic, homogeneous Dirichlet, or homogeneous Neumann boundary conditions, or on the line. However, the results in this paper do not depend on a Hamiltonian structure.

We analyze the differentiability properties in initial value and time step of the semiflow of the evolution equation and of a large class of A-stable Runge–Kutta methods, including the Gauss–Legendre methods, when applied to the evolution equation. To be able to differentiate the semiflow and the numerical method we formulate conditions that guarantee uniformity of the time-interval of existence (for the semiflow) and the maximum step size (for the numerical methods) over bounded sets of parameters. We present two versions of such uniformity results: Whenever existence can be achieved, uniformity holds on sufficiently small balls of initial data; we will label results of this type by "local version." Assuming more regularity for the initial data, we also obtain results which are uniform on bounded open sets so long as B is well-defined and bounded. We will label results of this type by "uniform version."

Note that differentiation in time results in multiplication with the unbounded operator A and is only well-defined when considered as a map from a subset of $D(A)$ to $\mathcal{Y}$; this is easily seen by taking $B \equiv 0$ and differentiating the exact semiflow $\mathrm{e}^{tA}U^0$. To be able to differentiate repeatedly in time we assume that B is $\mathcal{C}^{N-k}$ as map from some open set $\mathcal{D}_k \subset \mathcal{Y}_k \equiv D(A^k)$ to $\mathcal{Y}_k$ for $k = 0, \dots, K$ and $N > K$. Whether or not this condition is satisfied depends on the given evolution equation, and in particular on the boundary conditions; it is satisfied for the equations mentioned above in the case of periodic boundary conditions and smooth nonlinearities. We also give examples of PDEs with Neumann boundary conditions, Dirichlet boundary conditions, and on the line where this condition is true. We then prove that the semiflow of the evolution equation and the numerical method are of class $\mathcal{C}^K$ jointly in time (resp. step size) and initial data when considered as a map from $\mathcal{D}_K$ to $\mathcal{Y}$. Both results require carefully tracking the domains of definition of B. Moreover, under the additional assumption that A is normal (or, more generally, normal up to a perturbation which is a bounded linear operator on each of the $\mathcal{Y}_k$) or that A is sectorial, we show convergence of the semidiscretization in time at its full order p provided $K = p$ and for initial data $U^0 \in \mathcal{D}_{K+1}$.

The exact solution $U(t)$ of the semilinear evolution equation is obtained as a fixed point of a contraction map. Similarly, the Runge–Kutta methods we consider are implicit as they are functions of the Runge–Kutta stage vectors, which in turn are obtained as fixed points of contraction maps. As for the exact solution, differentiation in the step size of the Runge–Kutta method results in multiplication by the unbounded operator A. Hence, these derivatives are also only well-defined on the scale of Hilbert spaces $\mathcal{Y}_k$. An additional difficulty arises from the fact that the semiflow, the numerical method, the contraction maps for semiflow and stage vectors, and their derivatives with respect to the initial data are only strongly continuous in the time-like parameter, but not continuous in the operator norm. Hence, these maps do not fit into the usual setting of contraction mapping theorems with parameters. We therefore address these issues by providing an abstract theory for the differentiability properties of fixed points of contraction mappings on a scale of Banach spaces. This theory provides a unified framework for the time-continuous and time-semidiscrete case.

Let us mention some related results in the literature. Le Roux [23] studies convergence results for strongly A-stable approximations $\mathsf{S}(hA)$ of holomorphic semigroups e^{hA} on Hilbert spaces, an example of which are strongly A-stable Runge–Kutta methods applied to linear parabolic systems. Palencia [26] and Crouzeix *et al.* [12] study stability of A-acceptable rational approximations $\mathsf{S}(hA)$ of holomorphic semigroups e^{hA} on Banach spaces; they show that when $\operatorname{Re}(\operatorname{spec} A) \leq \omega$ for some $\omega > 0$, then $\|\mathsf{S}^n(hA)\| \leq \Theta_{\mathsf{S}}\, \mathrm{e}^{\omega_{\mathsf{S}} nh}$ for some $\omega_{\mathsf{S}} > 0$, $\Theta_{\mathsf{S}} > 0$, and all $n \in \mathbb{N}$. Lubich and Ostermann [24] prove convergence results for Runge–Kutta methods applied to semilinear parabolic equations on Banach spaces, cf. [11]. Variable step size schemes applied to fully nonlinear parabolic problems have been studied in [16]. González and Palencia [17] study stability of A-stable Runge–Kutta methods in the initial value, as we do, but they study quasilinear parabolic problems and do not consider the differentiability properties of the solution. Akrivis and Crouzeix [3] discuss multistep semidiscretizations in time for parabolic problems; see references therein for further related work.

In [21], quoted above, Hersh and Kato prove convergence of A-acceptable rational approximations $\mathsf{S}(hA)$ of non-analytic C_0-semigroups e^{hA} for smooth initial data. Brenner and Thomée [6] show that A-acceptable rational approximations $\mathsf{S}(hA)$ of non-analytic C_0-semigroups e^{hA} with $\operatorname{Re}(\operatorname{spec} A) \leq 0$ in general grow like $\|\mathsf{S}^n(hA)\| = O(n^{1/2})$ and study fractional order convergence for non-smooth initial data of linear evolution equations, see also [22]; for extensions to variable step size, see [4]. Brenner *et al.* [7] study convergence of rational approximations of inhomogeneous linear differential equations on Banach spaces, assuming stability of the approximation. Colin *et al.* [8, 9] study modified Crank–Nicolson semidiscretizations in time of nonlinear Schrödinger equations and Zakharov wave equations. They prove convergence as $h \to 0$, but do not analyze the order of convergence.

In this paper we have a related, but different objective. Similarly as in [6, 21, 22] we consider semidiscretizations in time of evolution equations which are not parabolic, i.e., equations whose linear part does not generate an analytic semigroup. But whereas the main issue in [6, 21, 22] is that the spectral theorem is not available for the linear operator A so that stability estimates of the form $\|\mathsf{S}^n(hA)\| \leq \Theta_{\mathsf{S}}\, \mathrm{e}^{\omega_{\mathsf{S}} nh}$, as required for the standard convergence analysis, are not available, we assume here, like [7], that this estimate holds true, e.g. due to normality of A on the Hilbert space $\mathcal{Y}$. Our focus is rather on semilinear problems as were considered by Lubich and Ostermann [24] in the parabolic case; the class of Runge–Kutta schemes considered here is the same as in their work. However, while [24, 11] assume the existence of a temporally smooth solution $U(t)$ of the semilinear evolution equation (or a perturbation of it) to obtain higher order convergence, we provide a detailed analysis under which conditions this assumption holds true.

Our conditions on B yield, in particular, $\mathcal{C}^K$ smoothness of the semiflow jointly in time and in the initial data for initial values in an open set of a Hilbert space $\mathcal{Y}_K$. If the conditions on B are not satisfied, the set of initial values of temporally smooth solutions is generally a complicated set which is characterized by nonlinear conditions; hence such initial data are in general difficult to prepare numerically. We illustrate this for the semilinear wave equation with generic nonlinearity and Dirichlet boundary conditions; see Section 2.5.3. Under the same conditions on B, differentiability of the numerical method in the step size h and in the initial data holds on open sets. This allows us to prove convergence of the numerical method without additional stage order conditions as have been assumed in [24]. Moreover, we obtain full order convergence, whereas e.g. the convergence results for semilinear PDEs of [24] provide an order of convergence that is determined by the stage order and that, in general, is smaller than the order of the method.

Lubich and Ostermann [24], in the parabolic setting, also obtain convergence results when only a perturbation of the continuous solution $U(t)$ is temporally smooth, and their estimate of the trajectory error then also depends on this perturbation error. In practice, such a perturbation would typically be a space discretization; if the continuous solution lacks temporal smoothness, the assumption of a temporally smooth solution of a perturbation tending to zero with the step size h typically imposes mesh conditions that exclude order p convergence of the semidiscretization in time.

In contrast to [6, 24], our interest in this paper is not on fractional order convergence for non-smooth initial data. Rather, since we are interested in obtaining higher order differentiability of the numerical method in the time step, we restrict attention to regular initial data $U^0 \in \mathcal{Y}_{K+1}$; in particular, we assume enough regularity to have full order of convergence, i.e., $K \geq p$ where p is the order of the numerical method. Our convergence result extends the corresponding classical result for linear evolution equations of Hersh and Kato [21] to nonlinear systems.

There has been a lot of recent activity in the application of split step time-semidiscretizations of nonlinear Schrödinger and wave equations: Besse *et al.* [5] and Lubich [25] study convergence of split step time-semidiscretizations for nonlinear Schrödinger equations; also see [18] for a general framework in the linear case and more references, and [15, 14] for long-time preservation of actions of nonlinear Schrödinger equations under split step time-semidiscretizations. While splitting methods are very effective for simulating evolution equations for which the linear evolution e^{tA} can easily be computed explicitly, Runge–Kutta methods are still a good choice when an eigendecomposition of A is not available, as for example for the semilinear wave equation in an inhomogeneous medium; see Section 2.5.5. Moreover, the simplest example of a Gauss–Legendre Runge-Kutta method, the implicit mid point rule, appears to have some advantage over split step time-semidiscretizations for the computation of wave trains for nonlinear Schrödinger equations [20, 32] because the latter introduce an artificial instability while the former reproduces recurrences well.

In this paper, we shall hence restrict our attention to Runge–Kutta methods which have a long history as robust and effective time integrators for both ODEs and PDEs; see, e.g., [29, 30]. Gauss–Legendre Runge–Kutta methods, in particular, have attracted attention as they are symplectic and yield multisymplectic space-time schemes for PDEs [2].

While we restrict attention to evolution equations on Hilbert spaces $\mathcal{Y}$, our results on differentiability of the semiflow and of the numerical method also hold true when $\mathcal{Y}$ is a Banach space. However, the stability condition $\|\mathsf{S}^n(hA)\| \leq \Theta_{\mathsf{S}}\, \mathrm{e}^{n\omega_{\mathsf{S}} h}$, which we need for our convergence result, is quite restrictive in the Banach space setting, as discussed above.

The paper is organized as follows. In Section 2, we introduce the class of semilinear evolution equations considered, and study the differentiability properties of the semiflow of these evolution equations. We also present a general result on the differentiability of superposition operators. We then show how the semilinear wave equation and the nonlinear Schrödinger equation with different types of boundary conditions fit into this framework. In Section 3, we derive corresponding statements on the well-posedness, differentiability properties, and convergence of A-stable Runge–Kutta methods when applied to such evolution equations. In Appendix A, we present a number of technical results, most notably a contraction mapping theorem on a scale of Banach spaces, which are needed in the main body of the paper.

2. Semilinear evolution equations

In this section, we set up the framework for a class of semilinear evolution equations whose time discretization we analyze subsequently. After introducing some notation (Section 2.1) and setting up the general functional framework in Sections 2.2, we provide a setting in which the semiflow is differentiable with respect to the initial data as well as time (Section 2.3). In many examples, the nonlinearities are superposition operators of nonlinear functions; we collect their fundamental properties in Section 2.4. These results enable us to fit our two main examples, the semilinear wave equation (Section 2.5) and the nonlinear Schrödinger equation (Section 2.6), into the abstract framework.

2.1. **Some notation.** Let $\mathcal{Y}$ be a Banach space. We write

$$\mathcal{B}_R^{\mathcal{Y}}(U^0) = \{U \in \mathcal{Y} \colon \|U - U^0\|_{\mathcal{Y}} \leq R\}$$

to denote the closed ball of radius R around $U^0 \in \mathcal{Y}$. (If no confusion about the space is possible, we may drop the superscript, or write $\mathcal{B}_R(U_0) \subset \mathcal{Y}$ instead of $\mathcal{B}_R^{\mathcal{Y}}(U_0)$.) Let $\mathcal{D} \subset \mathcal{Y}$ be open. We define

$$\mathcal{D}^{-\delta} = \{U \in \mathcal{D} \colon \operatorname{dist}_{\mathcal{Y}}(U, \partial\mathcal{D}) > \delta\}\,, \tag{2.1}$$

where $\operatorname{dist}_{\mathcal{Y}}(U, \mathcal{D}) = \inf_{W \in \mathcal{D}} \|U - W\|_{\mathcal{Y}}$ denotes the distance between a point $U \in \mathcal{Y}$ and the set $\mathcal{D} \subset \mathcal{Y}$ measured in the $\mathcal{Y}$-norm.

For Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, and $j \in \mathbb{N}_0$, we write $\mathcal{E}^j(\mathcal{Y}, \mathcal{X})$ to denote the vector space of j-multilinear bounded mappings from $\mathcal{Y}$ to $\mathcal{X}$; we set $\mathcal{E}^j(\mathcal{X}) \equiv \mathcal{E}^j(\mathcal{X}, \mathcal{X})$.

For Banach spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, and open subsets $\mathcal{U} \subset \mathcal{X}$, $\mathcal{V} \subset \mathcal{Y}$, and $\mathcal{W} \subset \mathcal{Z}$, we write

$$F \in \mathcal{C}^{(\underline{m},n)}(\mathcal{U} \times \mathcal{V}; \mathcal{W})$$

to denote a continuous function $F\colon \mathcal{U}\times\mathcal{V}\to\mathcal{W}$ whose partial Fréchet derivatives $\mathrm{D}_X^i\mathrm{D}_Y^j F(X,Y)$ exist and are such that the maps

$$(X,Y,X_1,\dots,X_i)\mapsto \mathrm{D}_X^i\mathrm{D}_Y^j F(X,Y)(X_1,\dots,X_i) \tag{2.2}$$

are continuous from $\mathcal{U}\times\mathcal{V}\times\mathcal{X}^i$ into $\mathcal{E}^j(\mathcal{Y},\mathcal{Z})$ for $i=0,\dots,m$ and $j=0,\dots,n$. In particular, all directional derivatives are continuous. We write

$$F\in\mathcal{C}_{\mathrm{b}}^{(\underline{m},n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})$$

if, in addition, the partial Fréchet derivatives are bounded and the maps (2.2) extend continuously to the boundary. (The latter is important as we will apply the contraction mapping theorem to maps in such classes.) As usual, we write

$$F\in\mathcal{C}^{(m,n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})$$

to denote that the partial Fréchet derivatives up to order (m,n) exist and are continuous in the norm topology; we write $\mathcal{C}_{\mathrm{b}}^{(m,n)}$ if these derivatives are, in addition, bounded and extend continuously to the boundary. If any of the sets is not open, we define

$$\mathcal{C}^{(m,n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})\equiv\mathcal{C}^{(m,n)}(\operatorname{int}(\mathcal{U})\times\operatorname{int}(\mathcal{V});\operatorname{int}(\mathcal{W}))\,,$$

where $\operatorname{int}(\mathcal{U})$ denotes the interior of $\mathcal{U}$, with analogous notation for the $\mathcal{C}_{\mathrm{b}}$-spaces. The spaces $\mathcal{C}^m(\mathcal{U};\mathcal{W})$ and $\mathcal{C}_{\mathrm{b}}^m(\mathcal{U};\mathcal{W})$ are defined likewise.

Note that $\mathcal{C}^{(\underline{m},n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})=\mathcal{C}^{(m,n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})$ only if $\mathcal{X}$ is finite-dimensional. In general,

$$\mathcal{C}_{\mathrm{b}}^{(m,n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})\supset\mathcal{C}_{\mathrm{b}}^{(\underline{m+1},n)}(\mathcal{U}\times\mathcal{V};\mathcal{W})\cap\mathcal{C}_{\mathrm{b}}^{(\underline{m},n+1)}(\mathcal{U}\times\mathcal{V};\mathcal{W}) \tag{2.3a}$$

because any differentiable function is continuous. Moreover,

$$\mathcal{C}_{\mathrm{b}}^{(\underline{0},k)}(\mathcal{U}\times\mathcal{V};\mathcal{W})=\mathcal{C}_{\mathrm{b}}^{(0,k)}(\mathcal{U}\times\mathcal{V};\mathcal{W})\,. \tag{2.3b}$$

In the above, $\mathcal{V}$ will typically be some interval of time.

2.2. **General setting.** We consider semilinear evolution equations on a Hilbert space $\mathcal{Y}$,

$$\partial_t U = F(U) = AU + B(U)\,, \tag{2.4}$$

where $U\colon[0,T]\to\mathcal{Y}$. Equation (2.4) formally looks like an ODE, but will be thought of as being posed on an infinite-dimensional function space $\mathcal{Y}$.

Our main examples are the following.

Example 2.1 (Semilinear wave equation). For the semilinear wave equation

$$\partial_{tt}u = \partial_{xx}u - f(u)\,, \tag{2.5}$$

we write $v=\partial_t u$ and $U=(u,v)^T$ which, for t fixed, shall be an element of a Hilbert space $\mathcal{Y}$ to be specified later, so that

$$A=\begin{pmatrix}0 & \mathrm{id}\\ \partial_x^2 & 0\end{pmatrix}\quad\text{and}\quad B(U)=\begin{pmatrix}0\\ -f(u)\end{pmatrix}\,. \tag{2.6}$$

Example 2.2 (Nonlinear Schrödinger equation). For the nonlinear Schrödinger equation

$$\mathrm{i}\,\partial_t u = -\partial_{xx}u + \partial_{\overline{u}}V(u,\overline{u})\,, \tag{2.7}$$

we set $U\equiv u$, so that

$$A=\mathrm{i}\,\partial_x^2\quad\text{and}\quad B(U)=-\mathrm{i}\,\partial_{\overline{u}}V(u,\overline{u})\,. \tag{2.8}$$

In the following, we introduce the framework in which we obtain smooth solutions of (2.4). Later, in Sections 2.5 and 2.6, we show how the semilinear wave equation and the nonlinear Schrödinger equation as formally introduced above fit into this framework. It is well known that the following conditions imply the existence of a semiflow of (2.4).

(A0) A is a closed, densely defined linear operator on $\mathcal{Y}$ and generates a C_0-semigroup on $\mathcal{Y}$.

(B0) $B\colon \mathcal{D} \to \mathcal{Y}$ is Lipschitz on some open set $\mathcal{D} \subset \mathcal{Y}$.

For the definition of strongly continuous semigroups (C_0-semigroups) and detailed proofs, see, e.g., [27]. For our purposes, the main points can be summarized as follows.

Condition (A0) implies, in particular, that there exist constants ω and Θ such that for every $t \geq 0$

$$\|\mathrm{e}^{tA}\| \leq \Theta\, \mathrm{e}^{\omega t} \tag{2.9}$$

with $\operatorname{Re}(\operatorname{spec} A) \leq \omega$. Moreover, for every $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$,

$$\|(\lambda - A)^{-1}\| \leq \frac{\Theta}{\operatorname{Re}\lambda - \omega}\,. \tag{2.10}$$

After reformulating (2.4) in its *mild formulation*

$$U(t) = \mathrm{e}^{tA}U^0 + \int_0^t \mathrm{e}^{(t-s)A}\, B(U(s))\,\mathrm{d}s\,, \tag{2.11}$$

the contraction mapping theorem applies and we obtain local-in-time well-posedness of our abstract semilinear evolution equation.

Let Φ^t denote the semiflow of (2.4), i.e. the map $U^0 \mapsto \Phi^t(U^0)$ such that $U(t) = \Phi^t(U^0)$ satisfies (2.11) with $U(0) = U^0$. We sometimes write $\Phi(U^0, t)$ in place of Φ^t. When $U^0 \in D(A)$, then $t \mapsto \Phi^t(U^0)$ is differentiable.

2.3. **Regularity of the semiflow.** When $B = 0$, then $t \mapsto \Phi^t(U)$ is k-times differentiable as a map from $D(A^k)$ to $\mathcal{Y}$ for every $k \in \mathbb{N}$. In this section, we extend this result to semilinear evolution equations under suitable assumptions on the nonlinearity B and provide bounds on the derivatives.

For $k \in \mathbb{N}_0$, we define

$$\mathcal{Y}_k = D(A^k)$$

endowed with the inner product

$$\langle U_1, U_2\rangle_{\mathcal{Y}_k} = \langle AU_1, AU_2\rangle_{\mathcal{Y}_{k-1}} + \langle U_1, U_2\rangle_{\mathcal{Y}_{k-1}}\,. \tag{2.12}$$

Then

$$\|A\|_{\mathcal{Y}_{\ell+1}\to\mathcal{Y}_\ell} \leq 1 \quad \text{and} \quad \|U\|_{\mathcal{Y}_\ell} \leq \|U\|_{\mathcal{Y}_{\ell+1}} \tag{2.13}$$

for all $U \in \mathcal{Y}_{\ell+1}$.

Given $\delta > 0$ and a hierarchy of open sets $\mathcal{D}_\ell \subset \mathcal{Y}_\ell$ for $\ell = 0, \dots, L$ for $L \in \mathbb{N}$ with $\mathcal{D}_0 \equiv \mathcal{D}$, we define $\mathcal{D}_0^{-\delta} \equiv \mathcal{D}^{-\delta}$ as in (2.1) and, for $\ell = 1, \dots, L$,

$$\mathcal{D}_\ell^{-\delta} \equiv \{U \in \mathcal{D}_\ell\colon\ \operatorname{dist}_{\mathcal{Y}_\ell}(U, \partial\mathcal{D}_\ell) > \delta\}\,.$$

Then, by construction and due to (2.13), $\mathcal{B}_\delta^{\mathcal{Y}_\ell}(U) \subset \mathcal{D}_\ell$ for all $U \in \mathcal{D}_\ell^{-\delta}$ and $\ell = 0, \dots, L$.

Let $\mathcal{Y}_1$ be a Banach space continuously embedded into the Banach space $\mathcal{Y}$. Then $\mathcal{D}_1 \subset \mathcal{Y}_1$ is called a δ_*-nested subset of $\mathcal{D} \subset \mathcal{Y}$ if $\mathcal{D}_1^{-\delta} \subset \mathcal{D}^{-\delta}$ for all $\delta \in [0, \delta_*]$. Furthermore we say that the family $\mathcal{D}_0, \dots, \mathcal{D}_L$ is δ_*-nested if $\mathcal{D}_\ell^{-\delta} \subset \mathcal{D}_{\ell-1}^{-\delta}$ for all

$\delta \in [0, \delta_*]$ with $\delta_* > 0$ and $\ell = 1, \dots, L$. For example, the family $\mathcal{D}_k = \mathcal{B}_R^{\mathcal{Y}_k}(U^0)$ is δ_*-nested for every $\delta_* \in (0, R)$ and $U^0 \in \mathcal{Y}_L$. However, an arbitrary nested family $\mathcal{D}_\ell \subset \mathcal{Y}_\ell$ may not be δ_*-nested for any $\delta_* > 0$.

To state a differentiability result for higher time derivatives, we need the following specific assumptions on the regularity of B on the scale $\mathcal{Y}_j$. The same assumptions will also be required for the convergence analysis of A-stable Runge–Kutta schemes in Section 3.

(B1) There exist $K \in \mathbb{N}_0$, $N \in \mathbb{N}$ with $N > K$, and a sequence of δ_*-nested $\mathcal{Y}_k$-bounded and open sets $\mathcal{D}_k$ such that $B \in \mathcal{C}_{\mathrm{b}}^{N-k}(\mathcal{D}_k; \mathcal{Y}_k)$ for $k = 0, \dots, K$.

We denote the bounds on $B \colon \mathcal{D}_k \to \mathcal{Y}_k$ and its derivatives by constants M_k, M'_k, etc., for $k = 0, \dots, K$, and identify $M = M_0$, $M' = M'_0$, and $\mathcal{D} = \mathcal{D}_0$. In addition to the domains $\mathcal{D}_0, \dots, \mathcal{D}_K$ defined in this assumption, we will sometimes need to refer to $\mathcal{D}_{K+1}$, which may be any δ_*-nested subset of $\mathcal{D}_K$ which is bounded and open in $\mathcal{Y}_{K+1}$.

We note that nonlinear continuous operators $B \in \mathcal{C}(\mathcal{D}; \mathcal{Y})$ do not generally map closed bounded sets into closed bounded sets, see Remark 2.3 below. However, the boundedness requirements can always be met on balls inside the domain of B, i.e., if B is $\mathcal{C}^n$ from some open set $\mathcal{D} \subset \mathcal{Z}$ to $\mathcal{Z}$ then, by continuity, for every $U^0 \in \mathcal{D}$ there is some $R > 0$ such that $B \colon \mathcal{B}_R(U^0) \subset \mathcal{D} \to \mathcal{Z}$ and its derivatives are uniformly bounded so that $B \in \mathcal{C}_{\mathrm{b}}^n(\mathcal{B}_R(U^0); \mathcal{Y})$.

Remark 2.3. The existence of continuous unbounded nonlinear functionals on an infinite-dimensional Banach space $\mathcal{X}$ can be seen by the following construction. It is a standard result that there exists a sequence $x_j \in \mathcal{X}$ such that $\|x_j\| = 1$ and $\|x_j - x_k\| \geq 3/4$; on a Hilbert space, an orthonormal basis will do. Now let $h_j \in \mathcal{C}(\mathcal{X}, \mathbb{R})$ have support on $\mathcal{B}_{1/4}^{\mathcal{X}}(x_j)$ with $h_j(x_j) = 1$. Then F defined by

$$F(x) = \sum_{j=0}^{\infty} j \, h_j(x)$$

satisfies $F \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, since we have $h_j(x) = 0$ for all but at most one j. But F does not map the closed bounded set $\mathcal{B}_1^{\mathcal{X}}(0)$ into a bounded set.

Superposition operators of smooth functions $f \colon D \subset \mathbb{R}^d \to \mathbb{R}^m$ on Sobolev spaces as occur in Examples 2.1 and 2.2 above are bounded. Indeed, for superposition operators we can construct δ_*-nested domains such that condition (B1) holds, see Theorem 2.12 and Sections 2.5 and 2.6 below.

Under assumptions (A0) and (B1), the semiflow Φ^t of (2.4) exists on each $\mathcal{Y}_k$. In the following, we show that a time derivative of order ℓ maps ℓ rungs down this scale of Hilbert spaces.

Theorem 2.4 (Regularity of the semiflow, local version)**.** *Assume* (A0) *and* (B1). *Choose* $R \in (0, \delta_*]$ *such that* $\mathcal{D}_K^{-R} \neq \emptyset$ *and pick* $U^0 \in \mathcal{D}_K^{-R}$. *Let* $R_* = R/(2\Theta)$ *with* Θ *from* (2.9). *Then there is* $T_* = T_*(R, U^0) > 0$ *such that the semiflow* $(U, t) \mapsto \Phi^t(U)$ *of* (2.4) *satisfies*

$$\Phi \in \bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(B_{R_*}^{\mathcal{Y}_K}(U^0) \times [0, T_*]; \mathcal{B}_R^{\mathcal{Y}_{k-\ell}}(U^0)) \,. \tag{2.14a}$$

In particular,

$$\Phi \in \mathcal{C}_{\mathrm{b}}^K(B_{R_*}^{\mathcal{Y}_K}(U^0) \times [0, T_*]; \mathcal{B}_R^{\mathcal{Y}}(U^0)) \,. \tag{2.14b}$$

The bounds on Φ and T_* depend only on the bounds afforded by (B1)*,* (2.9)*, R, and U^0.*

Proof. Writing $t = \tau T$ for some fixed $T > 0$, we see that a solution to the mild formulation (2.11) is a fixed point of the map

$$\Pi(W; U, T)(\tau) = \mathrm{e}^{\tau T A} U + T \int_0^\tau \mathrm{e}^{(\tau-\sigma)TA} \, B(W(\sigma)) \, \mathrm{d}\sigma \,. \tag{2.15}$$

This reformulation is useful because we want to quote the contraction mapping theorem on a scale of Banach spaces, Theorem A.9, to prove the differentiability properties of Φ as claimed. We work on the scale $\mathcal{Z}_j = \mathcal{C}_{\mathrm{b}}([0,1]; \mathcal{Y}_j)$ and seek a fixed point of Π in $\mathcal{W}_j = \mathcal{C}_{\mathrm{b}}([0,1]; \mathcal{B}_R^{\mathcal{Y}_j}(U^0))$ for $j = 0, \dots, K$, with parameter sets $\mathcal{U} \equiv \operatorname{int}(\mathcal{B}_{R_*}^{\mathcal{Y}_K}(U^0)) \subset \mathcal{X} = \mathcal{Y}_K$ and $\mathcal{I} = (0, T_*)$. Clearly, Π maps $\mathcal{W}_j \times \mathcal{U} \times \mathcal{I}$ into $\mathcal{Z}_j$. To bound the range of Π, we estimate, for $j = 0, \dots, K$,

$$\begin{aligned} &\|\Pi(W; U, T) - U^0\|_{\mathcal{Y}_j} \\ &\quad \le \|\mathrm{e}^{\tau T A} U^0 - U^0\|_{\mathcal{Y}_j} + \|\mathrm{e}^{\tau T A}(U - U^0)\|_{\mathcal{Y}_j} + T \int_0^\tau \|\mathrm{e}^{(\tau-\sigma)TA} \, B(W(\sigma))\|_{\mathcal{Y}_j} \, \mathrm{d}\sigma \\ &\quad \le \|\mathrm{e}^{\tau T A} U^0 - U^0\|_{\mathcal{Y}_j} + \Theta \, \mathrm{e}^{\omega T} \, R_* + T \, \Theta \, \mathrm{e}^{\omega T} \, M_j \,. \end{aligned} \tag{2.16}$$

With the choice $R_* = R/2\Theta$, we observe that for sufficiently small T_* and all $T \in [0, T_*]$ the right hand side can be made less than R for $j = 0, \dots, K$ independent of $\tau \in [0,1]$. We can thus take the supremum over $\tau \in [0,1]$, which altogether proves that Π maps $\mathcal{W}_j \times \mathcal{U} \times \mathcal{I}$ into $\mathcal{W}_j$. Condition (i) of Theorem A.9 then follows from our assumptions on A and B.

Similarly, we estimate

$$\|\mathrm{D}_W \Pi(W; U, T)\|_{\mathcal{E}(\mathcal{C}_{\mathrm{b}}([0,1];\mathcal{Y}_j))} \le T \, \Theta \, \mathrm{e}^{\omega T} \, M_j' \,, \tag{2.17}$$

so that Π is a uniform contraction for all $U \in \mathcal{U}$, $W \in \mathcal{W}$, and $T \in \mathcal{I} = (0, T_*)$ with a possibly smaller value of T_*. Here we used that B is at least $\mathcal{C}^1$ on the highest rung of the scale due to the requirement that $N > K$ in (B1). Hence, condition (ii) of Theorem A.9 is verified.

Theorem A.9 then implies that the fixed point W of Π satisfies

$$W \in \bigcap_{\substack{j+k \le N \\ \ell \le k \le K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(B_{R_*}^{\mathcal{Y}_K}(U^0) \times [0, T_*]; \mathcal{W}_{k-\ell}) \,.$$

To infer (2.14a), we recall that $\Phi^{\tau T}(U) = W(U,T)(\tau)$; hence $\partial_U^m \partial_T^n \Phi^T(U) = \partial_U^m \partial_T^n W(U,T)(1)$. Finally, (2.14b) follows from Lemma A.2. $\square$

Remark 2.5*.* With the choice of norm (2.12), the fundamental estimates in this paper which carry named constants, in particular Θ in (2.20) and Λ, c_{S} in Lemma 3.11, are the same on all $\mathcal{Y}_k$ for $k \in \mathbb{N}_0$ as these constants are norms of operators like e^{tA} which commute with A. Thus, if $\mathcal{Y}_k$ for $k \in \mathbb{N}$ were endowed with a different, but equivalent set of norms, these and consequent constants would need to be adopted and possibly become dependent on the rung.

Theorem 2.4 does not guarantee that the time of existence of the solution can be chosen uniformly over $\mathcal{D}$ or even over $\mathcal{D}^{-\delta}$ for some $\delta > 0$. The following theorem shows that such uniformity can be obtained along with improved regularity over bounded domains other than balls on the expense of requiring the initial data to lie in a set one step up the scale.

In the following, define

$$R_{K+1} = \sup_{U \in \mathcal{D}_{K+1}^{-\delta}} \|U\|_{\mathcal{Y}_{K+1}} . \tag{2.18}$$

Theorem 2.6 (Regularity of the semiflow, uniform version)**.** *Assume* (A0) *and* (B1)*. Choose $\delta \in (0, \delta_*]$ small enough such that $\mathcal{D}_{K+1}^{-\delta} \neq \emptyset$. Then there exists $T_* = T_*(\delta) > 0$ such that the semiflow $(U, t) \mapsto \Phi^t(U)$ of* (2.4) *satisfies* (2.14) *with uniform bounds for all $U^0 \in \mathcal{D}_{K+1}^{-\delta}$, with $R = \delta$, and such that*

$$\Phi \in \bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K+1}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0, T_*]; \mathcal{Y}_{k-\ell}) . \tag{2.19a}$$

In particular, when $N > K + 1$,

$$\Phi \in \mathcal{C}_{\mathrm{b}}^{K+1}(\mathcal{D}_{K+1}^{-\delta} \times [0, T_*]; \mathcal{D}) . \tag{2.19b}$$

The bounds on Φ and T_* depend only on δ and on the bounds afforded by (B1)*,* (2.18)*, and* (2.9)*.*

Proof. We apply Theorem 2.4 for each $U^0 \in \mathcal{D}_{K+1}^{-\delta}$ with $R = \delta$. We note that in the proof of Theorem 2.4, even in the case $K = 0$, the guaranteed time of existence T_* cannot be chosen uniformly for $U^0 \in \mathcal{D}^{-\delta}$ because the first term on the right of (2.16) cannot be made uniformly small. However, we may alternatively estimate, using (2.13) and (2.18), that for $j = 0, \ldots, K$

$$\|\mathrm{e}^{\tau T A} U^0 - U^0\|_{\mathcal{Y}_j} \leq T \max_{t \in [0,T]} \|A \mathrm{e}^{tA} U^0\|_{\mathcal{Y}_j} \leq T \; \Theta \, \mathrm{e}^{\omega T} \, R_{1+j} . \tag{2.20}$$

Inserting this estimate into (2.16), we see that we can choose $T_* > 0$ small enough such that $\Pi(\,\cdot\,; U, T)$ maps $\mathcal{W}_j = \mathcal{B}_R^{\mathcal{Z}_j}(U^0)$ into itself for all $U^0 \in \mathcal{D}_{K+1}^{-\delta}$ and $T \in [0, T_*]$. Following the proof of Theorem 2.4, we find that (2.14a) holds with uniform bounds for all $U^0 \in \mathcal{D}_{K+1}^{-\delta}$ when $R = \delta$, thereby implying

$$\Phi \in \bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0, T_*]; \mathcal{D}_{k-\ell}) \tag{2.21}$$

with bounds which only depend on the bounds afforded by (B1), (2.9), (2.18), and on δ. Next, we prove that Φ maps into a space one step up the scale, namely

$$A\Phi \in \bigcap_{\substack{j+k \leq N-1 \\ \ell \leq k \leq K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0, T_*]; \mathcal{Y}_{k-\ell}) . \tag{2.22}$$

Note that by [27, Theorem 6.1.5], a mild solution $U(t)$ of (2.4) satisfies $U(t) \in D(A)$ if $U(0) \in D(A)$ and $B \in \mathcal{C}^1(\mathcal{D}, \mathcal{Y})$; thus, the formal identity $\mathrm{d}W(\tau)/\mathrm{d}\tau = T\,(AW(\tau T) + B(W(\tau)))$ for $W(\tau) = U(\tau T)$ holds true. Hence, by applying A to the fixed point equation (2.15) and integrating by parts, we find

$$\begin{aligned} AW(\tau) &= \mathrm{e}^{\tau TA} AU + T \int_0^\tau A\mathrm{e}^{(\tau-\sigma)TA} B(W(\sigma))\,\mathrm{d}\sigma \\ &= \mathrm{e}^{\tau TA}\,(AU + B(U)) - B(W(\tau)) \\ &\quad + T\int_0^\tau \mathrm{e}^{(\tau-\sigma)TA}\,\mathrm{D}B(W(\sigma))(AW(\sigma) + B(W(\sigma)))\,\mathrm{d}\sigma\,. \end{aligned}$$

This is a linear fixed point equation

$$\begin{aligned} \tilde W = \tilde\Pi(\tilde W, U, T)(\tau) &= \mathrm{e}^{\tau TA}\,(AU + B(U)) - B(W(\tau)) \\ &\quad + T\int_0^\tau \mathrm{e}^{(\tau-\sigma)TA}\,\mathrm{D}B(W(\sigma))(\tilde W(\sigma) + B(W(\sigma)))\,\mathrm{d}\sigma \end{aligned} \tag{2.23}$$

for $\tilde W(U,T) = AW(U,T)$. We consider the fixed point equation (2.23) for $\tilde W = AW$ with $\mathcal{W}_j = \mathcal{B}^r_{\mathcal{Z}_j}(0)$ for $j = 0, \dots, K$ with $r > 0$ big enough such that $\tilde\Pi$ maps each $\mathcal{W}_j$ into itself. Applying Lemma A.6 (chain rule on the scale of Banach spaces) and Lemma A.7 to the right hand side of the fixed point equation (2.23), we verify once more the assumptions of Theorem A.9 with N replaced by $N-1$. This yields (2.22).

It remains to be shown that we can translate improved spatial regularity into differentiability in time by invoking the semilinear evolution equation (2.4). Due to (2.21), Lemma A.6 implies that $B \circ \Phi$ is in the same class (2.22) as $A\Phi$ and, since $\partial_t \Phi = A\Phi + B \circ \Phi$, so is $\partial_t\Phi$. Combining this result, (2.21), and (2.22) via Lemma A.4 implies (2.19a).

Finally, (2.19b) follows from Lemma A.2 with K replaced by $K+1$. □

Remark 2.7. It is worth noting that, even though we find that Φ^t maps into $\mathcal{Y}_{K+1}$, the proof, being based on the fixed point problem (2.23), requires B to be defined only up to rung K. The same pattern occurs when studying the Runge–Kutta numerical time-h maps in Section 3.

Remark 2.8 (Image of semiflow). The proof of Theorem 2.6 shows that, actually,

$$\Phi \in \bigcap_{\substack{j+k\le N \\ \ell\le k\le K+1 \\ (k,\ell)\ne(K+1,0)}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}^{-\delta}_{K+1} \times [0,T_*]; \mathcal{D}_{k-\ell})\,.$$

Remark 2.9. If $A = -A^*$ is skew-symmetric on the Hilbert space $\mathcal{Y}$, as for the nonlinear Schrödinger equation (see Section 2.6), then A is normal, $\mathrm{i}A$ is self-adjoint and, by Stone's theorem, generates a unitary group e^{tA}. More generally, if A is skew-symmetric up to a perturbation which is bounded on all $\mathcal{Y}_k$, as for the semilinear wave equation (see Section 2.5), then A generates a C_0 group e^{tA} on each $\mathcal{Y}_k$. In both cases, (2.9) and (2.10) may be replaced by the following statement: There exist a constant ω with $|\mathrm{Re}(\operatorname{spec} A)| \le \omega$ and a constant Θ such that for every $t \in \mathbb{R}$ and for every $\lambda \in \mathbb{C}$ with $|\mathrm{Re}\,\lambda| > \omega$

$$\|\mathrm{e}^{tA}\| \le \Theta\,\mathrm{e}^{\omega|t|}, \quad \|(\lambda - A)^{-1}\| \le \frac{\Theta}{|\mathrm{Re}\,\lambda| - \omega}\,,$$

see [27]. Then the semiflow Φ^t is also a flow with interval of existence $[-T_*, T_*]$ and regularity as specified in Theorem 2.4 and Theorem 2.6.

2.4. **Superposition operators.** To study the well-posedness of evolution equations such as (2.5) and (2.7), we need to consider superposition operators $f\colon \mathcal{G}_\ell \subset \mathcal{H}_\ell \to \mathcal{H}_\ell$ of functions $f\colon G \subset \mathbb{R}^d \to \mathbb{R}^d$. This concept is widely used; see, e.g., [19, 27] for specific examples. In this section, we characterize superposition operators in sufficient generality for later use.

Let $I = [a,b] \subset \mathbb{R}$ be a bounded closed interval. We write $\mathcal{H}_\ell(I;\mathbb{R}^d)$ to denote the Sobolev space of functions $u\colon I \to \mathbb{R}^d$ whose weak derivatives up to order ℓ are contained in $\mathcal{L}_2(I;\mathbb{R}^d)$.

Lemma 2.10 ([1])**.** *The space $\mathcal{H}_\ell(I;\mathbb{R})$ is a topological algebra for every $\ell > 1/2$. Specifically, there exists a constant $c = c(\ell)$ such that for every $u, v \in \mathcal{H}_\ell(I;\mathbb{R})$ the product $uv \in \mathcal{H}_\ell(I;\mathbb{R})$ satisfies*

$$\|uv\|_{\mathcal{H}_\ell(I,\mathbb{R})} \le c\,\|u\|_{\mathcal{H}_\ell(I;\mathbb{R})}\,\|v\|_{\mathcal{H}_\ell(I;\mathbb{R})}\,. \tag{2.24}$$

Armed with this result, we can characterize more general superposition operators where a function $f\colon G \to \mathbb{R}^m$ for some open $G \subset \mathbb{R}^d$ induces a mapping $u \mapsto f(u)$ between function spaces. The kth derivative of f as a function on $\mathbb{R}^d$ is a k-linear map on $\mathbb{R}^d$. As such, it induces a k-linear superposition operator between function spaces. *A priori*, it is not clear whether the kth Fréchet derivative of the superposition operator of f equals the superposition operator of the kth derivative of f on $\mathbb{R}^d$. The following lemma and theorem provide a setting in which this is true, so that we use the symbol $\mathrm{D}^k f$ for both these objects.

Lemma 2.11. *Let $G \subset \mathbb{R}^d$ be open, let $f \in \mathcal{C}_{\mathrm{b}}^N(G;\mathbb{R}^m)$ for some $N \in \mathbb{N}_0$, and set*

$$\mathcal{G} = \{u \in \mathcal{C}(I;\mathbb{R}^d) \colon u(I) \subset G\}\,.$$

Then $f \in \mathcal{C}_{\mathrm{b}}^N(\mathcal{G},\mathcal{C}(I;\mathbb{R}^m))$ and the derivatives of f as an operator from $\mathcal{C}(I;\mathbb{R}^d)$ to $\mathcal{C}(I;\mathbb{R}^m)$ are the superposition operators of the derivatives of f as a function on $\mathbb{R}^d$.

Proof. We proceed iteratively for $n = 0, \dots, N$. The Taylor theorem with integral remainder asserts that for fixed $z_0 \in G$

$$\left| f(z) - \sum_{i=0}^n \frac{\mathrm{D}^i f(z_0)}{i!}\,(z - z_0)^i \right| \le \rho(z_0, z)\,|z - z_0|^n \tag{2.25}$$

(when $d > 1$, $\mathrm{D}^i f$ is an i-linear map acting on the tensor product $(z - z_0)^i$), where

$$\rho(z_0, z) = \frac{1}{n!}\max_{\theta \in [0,1]} |\mathrm{D}^n f(z_0 + \theta(z - z_0)) - \mathrm{D}^n f(z_0)|$$

is continuous in $z_0, z \in G$ and uniformly continuous for $z_0, z \in K$ whenever $K \subset G$ is compact.

We now fix $u_0 \in \mathcal{G}$ and let $u \in \mathcal{G}$. Clearly, $u_0(I)$ and $u(I)$ are compact subsets of G, so that, setting $z_0 = u_0(x)$ and $z = u(x)$ in (2.25), we may take the supremum over $x \in I$, thereby obtaining

$$\left\| f(u) - \sum_{i=0}^n \frac{\mathrm{D}^i f(u_0)}{i!}\,(u - u_0)^i \right\|_{\mathcal{C}(I;\mathbb{R}^m)} \le \|\rho(u_0, u)\|_{\mathcal{C}(I;\mathbb{R}^m)}\,\|u - u_0\|_{\mathcal{C}(I;\mathbb{R}^d)}^n\,.$$

Since $\rho(u_0, u_0) = 0$, this proves that $f \in \mathcal{C}^n(\mathcal{G};\mathcal{C}(I;\mathbb{R}^m))$ and identifies the Fréchet derivative of order n as the superposition operator of the derivative of order n on $\mathbb{R}^d$.

Since $f \in \mathcal{C}_{\mathrm{b}}(G, \mathbb{R}^m)$, the set $f(\mathcal{G})$ is a bounded subset of $\mathcal{C}(I; \mathbb{R}^m)$. Moreover f extends continuously to the boundary of $\mathcal{G}$ since $f\colon G \to \mathbb{R}^m$ does.

To prove boundedness of $\mathrm{D}^k f$ as a map from $\mathcal{G}$ to $\mathcal{E}^k(\mathcal{C}(I;\mathbb{R}^d), \mathcal{C}(I;\mathbb{R}^m))$ for $k = 1, \dots, N$, we employ its identification with the superposition operator of the k-linear map $\mathrm{D}^k f$ on $\mathbb{R}^d$ and estimate

$$\|\mathrm{D}^k f(u)(u_1, \dots, u_k)\|_{\mathcal{C}(I;\mathbb{R}^m)} \le c \, \|\mathrm{D}^k f(u)\|_{\mathcal{C}(I;\mathbb{R}^{md^k})} \prod_{i=1}^{k} \|u_i\|_{\mathcal{C}(I;\mathbb{R}^d)} \tag{2.26}$$

for some $c > 0$, noting that $\mathrm{D}^k f \in \mathcal{C}_{\mathrm{b}}(\mathcal{G}; \mathcal{C}(I; \mathbb{R}^{md^k}))$ by the argument for the case $k = 0$. □

The corresponding result on the Sobolev scale is as follows.

Theorem 2.12. *Let $f \in \mathcal{C}_{\mathrm{b}}^N(G; \mathbb{R}^m)$ for some $N \in \mathbb{N}_0$ and open set $G \subset \mathbb{R}^d$. For each $\ell = 1, \dots, N$, let $\mathcal{G}_\ell$ denote an $\mathcal{H}_\ell$-bounded and open subset of $\mathcal{G} \cap \mathcal{H}_\ell(I; \mathbb{R}^d)$ with $\mathcal{G}$ as in Lemma 2.11. Then*

$$f \in \mathcal{C}_{\mathrm{b}}^N(\mathcal{G}_1; \mathcal{L}_2(I; \mathbb{R}^m)) \cap \bigcap_{\substack{k+\ell \le N \\ \ell \ge 1}} \mathcal{C}_{\mathrm{b}}^k(\mathcal{G}_\ell; \mathcal{H}_\ell(I; \mathbb{R}^m)) \, .$$

The derivatives of f as an operator on $\mathcal{H}_\ell$ are the superposition operators of the derivatives of f as a function from $\mathbb{R}^d$ to $\mathbb{R}^m$.

Proof. The statement $f \in \mathcal{C}_{\mathrm{b}}^N(\mathcal{G}_1; \mathcal{L}_2(I))$ is a direct consequence of Lemma 2.11 and the continuity of the embeddings $\mathcal{H}_1(I; \mathbb{R}^m) \subset \mathcal{C}(I; \mathbb{R}^m) \subset \mathcal{L}_2(I; \mathbb{R}^m)$. (The first inclusion is due to the Sobolev embedding theorem.)

Next, we show that $f \in \mathcal{C}_{\mathrm{b}}^\ell(G; \mathbb{R}^m)$ is bounded as an operator from $\mathcal{G}_\ell$ to $\mathcal{H}_\ell(I; \mathbb{R}^m)$ for $\ell = 1, \dots, N$. We proceed inductively in ℓ. Since, for some $C_\ell > 0$,

$$\|w\|_{\mathcal{H}_\ell} \le C_\ell \left(\|w\|_{\mathcal{H}_{\ell-1}} + \|w_x\|_{\mathcal{H}_{\ell-1}} \right) \tag{2.27}$$

for $w \in \mathcal{H}_\ell(I)$, the inductive step is achieved by taking $w = f(u)$ and showing that $\|\partial_x f(u)\|_{\mathcal{H}_{\ell-1}}$ is bounded over $u \in \mathcal{G}_\ell$. Indeed, when $\ell = 1$, $\|f(u)\|_{\mathcal{L}_2}$ is uniformly bounded in $u \in \mathcal{G}_1$ by the argument above, and there is a constant $c_1 > 0$ such that

$$\|\partial_x f(u)\|_{\mathcal{L}_2(I;\mathbb{R}^m)} \le c_1 \, \|\mathrm{D} f(u)\|_{\mathcal{C}(I;\mathbb{R}^{dm})} \, \|u_x\|_{\mathcal{L}_2(I;\mathbb{R}^d)}$$

is uniformly bounded for $u \in \mathcal{G}_1$ by Lemma 2.11. We conclude that f is bounded as map from $\mathcal{G}_1$ to $\mathcal{H}_1(I; \mathbb{R}^m)$. When $\ell \ge 2$, applying the algebra inequality (2.24) component-wise, we estimate

$$\|\partial_x f(u)\|_{\mathcal{H}_{\ell-1}(I;\mathbb{R}^m)} \le c_2 \, \|\mathrm{D} f(u)\|_{\mathcal{H}_{\ell-1}(I;\mathbb{R}^{dm})} \, \|u_x\|_{\mathcal{H}_{\ell-1}(I;\mathbb{R}^d)} \, ,$$

for some constant $c_2 > 0$, where the right side is uniformly bounded for $u \in \mathcal{G}_\ell$ since $\|\mathrm{D} f(u)\|_{\mathcal{H}_{\ell-1}(I,\mathbb{R}^{dm})}$ is uniformly bounded for $u \in \mathcal{G}_{\ell-1}$ by induction hypothesis. Thus, by (2.27) with $w = f(u)$, using the induction hypothesis once more, we obtain boundedness of $f\colon \mathcal{G}_\ell \to \mathcal{H}_\ell(I; \mathbb{R}^m)$.

To prove continuity and continuous differentiability of $f\colon \mathcal{G}_\ell \to \mathcal{H}_\ell$, we introduce, for $k = 0, \dots, N-1$,

$$F_k(u_0, u) = f(u) - \sum_{i=0}^{k} \frac{\mathrm{D}^i f(u_0)}{i!} \, (u - u_0)^i$$

and write $\partial_x F_k(u_0, u)$ in the form

$$\partial_x F_k = \left[\mathrm{D}f(u) - \sum_{i=0}^{k} \frac{\mathrm{D}^{i+1} f(u_0)}{i!} \, (u-u_0)^i\right] \partial_x u + \frac{\mathrm{D}^{k+1} f(u_0)}{k!} \, (u-u_0)^k \, \partial_x (u-u_0) \, .$$

When $\ell = 1$, we estimate for every $k = 0, \dots, N-1$, using Lemma 2.11 and the Sobolev embedding theorem again, that, for $u \in \mathcal{G}_\ell$,

$$\begin{aligned} \|\partial_x F_k\|_{\mathcal{H}_{\ell-1}(I;\mathbb{R}^m)} &\leq c_3 \left\| \mathrm{D}f(u) - \sum_{i=0}^{k} \frac{\mathrm{D}^{i+1} f(u_0)}{i!} \, (u-u_0)^i \right\|_{\mathcal{C}(I)} \|u\|_{\mathcal{H}_\ell} \\ &\quad + c_3 \, \frac{1}{k!} \, \|\mathrm{D}^{k+1} f(u_0)\|_{\mathcal{C}_{\mathrm{b}}(I)} \, \|u-u_0\|_{\mathcal{C}(I)}^k \, \|u-u_0\|_{\mathcal{H}_\ell(I)} \\ &\leq \sigma(u_0,u) \, \|u-u_0\|_{\mathcal{H}_\ell(I)}^k \end{aligned} \tag{2.28}$$

for some $\sigma \in \mathcal{C}(\mathcal{G}_\ell \times \mathcal{G}_\ell; \mathbb{R}_0^+)$ with $\sigma(u_0, u) = 0$ and some constant $c_3 > 0$. Moreover, since $f \in \mathcal{C}^N(\mathcal{G}_1; \mathcal{L}_2)$, there exists a function $\omega \in \mathcal{C}(\mathcal{G}_1 \times \mathcal{G}_1; \mathbb{R}_0^+)$ with $\omega(u_0, u) = 0$ such that $\|F_k\|_{\mathcal{L}_2} \leq \omega(u_0, u) \, \|u - u_0\|_{\mathcal{H}_1}^k$. Hence, (2.27) with $w = F_k(u_0, u)$ implies $f \in \mathcal{C}^{N-1}(\mathcal{G}_1; \mathcal{H}_1)$.

When $\ell \geq 2$, we obtain, by applying (2.24) recursively and component-wise to the second term of $\partial_x F_k$, an estimate as on the first and second line of (2.28) with $\mathcal{H}_{\ell-1}(I)$ in place of $\mathcal{C}(I)$ for every $k = 0, \dots, N-\ell$. Applying the induction hypothesis to both f and $\mathrm{D}f$ shows, as before, that $f \in \mathcal{C}^{N-\ell}(\mathcal{G}_\ell; \mathcal{H}_\ell)$ and that its derivatives are the superposition operators of the derivatives of f as a function on $\mathbb{R}^d$.

Due to this identification, we can prove boundedness of $\mathrm{D}^k f$ as a map from $\mathcal{G}_\ell$ to $\mathcal{E}^k(\mathcal{H}_\ell(I;\mathbb{R}^d), \mathcal{H}_\ell(I;\mathbb{R}^m))$ by applying (2.24) recursively and component-wise to $\mathrm{D}^k f(u)(u_1, \dots, u_k)$. In this way we obtain an estimate of the form (2.26) with $\mathcal{H}_\ell$ in place of $\mathcal{C}_{\mathrm{b}}$. The bound is then achieved by noting that $\mathrm{D}^k f \colon \mathcal{G}_\ell \to \mathcal{H}_\ell(I;\mathbb{R}^{md^k})$ is a bounded operator by the argument provided earlier in this proof for $k + \ell \leq N$.

Finally, we need to show that $\mathrm{D}^k f \colon \mathcal{G}_\ell \to \mathcal{E}^k(\mathcal{H}_\ell(I;\mathbb{R}^d), \mathcal{H}_\ell(I;\mathbb{R}^m))$ extends continuously to the boundary of $\mathcal{G}_\ell$ when $k + \ell \leq N$. For $k = 0$ this follows recursively from (2.28) and $f \in \mathcal{C}_{\mathrm{b}}^N(\mathcal{G}; \mathcal{C}(I))$ as above. Applying this result to $\mathrm{D}^k f \colon \mathcal{G}_\ell \to \mathcal{H}_\ell(I;\mathbb{R}^{md^k})$ and using once again the identification of derivatives of the superposition operator with the superposition operators of the derivatives, we complete the proof. □

2.5. **Example: the semilinear wave equation.** In the case of the semilinear wave equation (2.5), the operators A and B are given by (2.6).

2.5.1. *Periodic boundary conditions.* Since the Laplacian is diagonal in the Fourier representation, it is easy to see that the spectrum of A is given by $\operatorname{spec} A = \{\mathrm{i}k \colon k \in \mathbb{Z}\}$ and that the group generated by $\mathbb{Q}_0 A$ is unitary on any

$$\mathcal{Y}_\ell = \mathcal{H}_{\ell+1}(I;\mathbb{R}) \times \mathcal{H}_\ell(I;\mathbb{R}) \quad \text{for } \ell \in \mathbb{N}_0 \, .$$

Here $\mathbb{P}_0$ is the spectral projection associated with eigenvalue 0 and $\mathbb{Q}_0 = \mathrm{id} - \mathbb{P}_0$. Hence, A generates a C_0-group on $\mathcal{Y}_\ell$ and assumption (A0) is met. The full group e^{tA}, however, is not unitary due to the secular term from the Jordan block of A when restricted to $\mathbb{P}_0 \mathcal{Y}_\ell$.

Assume that the nonlinearity f of the semilinear wave equation (2.5) satisfies $f \in \mathcal{C}_{\mathrm{b}}^N(G;\mathbb{R})$ for some $N \in \mathbb{N}$ and some open set $G \subset \mathbb{R}$, and let $\mathcal{D} = \mathcal{D}_u \times \mathcal{D}_v$ where $\mathcal{D}_u$ is the set $\mathcal{G}_1$ from Theorem 2.12 and $\mathcal{D}_v$ denotes an open bounded subset of $\mathcal{L}_2(I)$. Then, by Theorem 2.12, the nonlinearity B satisfies assumption (B1) on the scale defined above with $K < N$ if we recursively define $\mathcal{D}_k = \mathcal{D}_{k-1}^{-\delta_*} \cap \operatorname{int}(\mathcal{B}_R^{\mathcal{Y}_k}(0))$ for some $R > 0$ with $\mathcal{D} \subset \mathcal{B}_R^{\mathcal{Y}}(0)$ and choose $\delta_* > 0$ small enough to ensure that all $\mathcal{D}_k$ are non-empty. Hence, Theorems 2.4 and 2.6 give regularity of the flow of the semilinear wave equation on the scale $\mathcal{Y}_k$ defined above.

2.5.2. *Neumann boundary conditions.* In the case of Neumann boundary conditions on $I = [0,\pi]$, we set $\mathcal{Y} = \mathcal{H}_1(I,\mathbb{R}) \times \mathcal{L}_2(I,\mathbb{R})$ as before; the operator A then has the same spectrum and $\mathrm{e}^{t\mathbb{Q}_0 A}$ is again unitary. In this case, $\mathcal{Y}_k = \mathcal{H}_{k+1}^{\mathrm{nb}}(I,\mathbb{R}) \times \mathcal{H}_k^{\mathrm{nb}}(I,\mathbb{R})$ with

$$\mathcal{H}_k^{\mathrm{nb}}(I,\mathbb{R}) = \{u \in \mathcal{H}_k(I,\mathbb{R})\colon u^{(2j+1)}(0) = u^{(2j+1)}(\pi) = 0 \text{ for } j = 0,\dots,\lfloor k/2 \rfloor - 1\}\,.$$

When $G \subset \mathbb{R}$ is open and $f \in \mathcal{C}_{\mathrm{b}}^N(G;\mathbb{R})$, assumption (B0) holds as before on the open bounded set $\mathcal{D} \subset \mathcal{Y}$ defined above. We claim that (B1) also holds for $K < N$. To prove the claim, we must show that f maps $\mathcal{H}_{k+1}^{\mathrm{nb}}(I,\mathbb{R}) \cap \mathcal{D}_u$ into $\mathcal{H}_k^{\mathrm{nb}}(I,\mathbb{R})$ for $k = 0,\dots,K$. When $k = 1$, no boundary conditions need to be checked. When $k = 2$, we observe that $(\partial_x f(u))(x) = f'(u(x))\, u_x(x) = 0$ for $x = 0,\pi$ and $u \in \mathcal{H}_2^{\mathrm{nb}} \cap \mathcal{D}_u$, so $f(u) \in \mathcal{H}_1^{\mathrm{nb}}$. Further, when $k = 3,\dots,K$, all terms in the sum obtained from computing $\partial_x^{2j+1} f(u)$ contain at least one odd derivative of u of order at most $2j+1$, so the boundary conditions remain satisfied.

2.5.3. *Dirichlet boundary conditions.* When endowed with Dirichlet boundary conditions, A generates a unitary semigroup. We take $I = [0,\pi]$ as before and set $\mathcal{Y}_k = \mathcal{H}_{k+1}^0(I,\mathbb{R}) \times \mathcal{H}_k^0(I,\mathbb{R})$, where

$$\mathcal{H}_k^0(I,\mathbb{R}) = \{u \in \mathcal{H}_k(I,\mathbb{R})\colon u^{(2j)}(0) = u^{(2j)}(\pi) = 0 \text{ for } j \in \mathbb{N}_0 \text{ with } 2j \le k-1\}\,.$$

Let $G \subset \mathbb{R}$ be open with $0 \in G$ and let $f \in \mathcal{C}_{\mathrm{b}}^N(G,\mathbb{R})$ as before. Then condition (B0) is satisfied with $\mathcal{D} = \mathcal{D}_u \times \mathcal{D}_v$, as before. Condition (B1) is satisfied if $f^{(2j)}(0) = 0$ for $0 \le 2j \le K-1$.

When f does not satisfy these boundary conditions, *necessary* conditions for the existence of time derivatives take a complicated structure. To see this, it suffices to consider differentiability at $t = 0$. For $U'(0)$ to exist, we have the obvious requirement that $v(0,0) = v(0,\pi) = 0$. For $U''(0)$ to exist, the non-homogeneous boundary condition $\partial_x^2 u(0,0) = \partial_x^2 u(0,\pi) = -f(0)$ needs to be satisfied. For $U'''(0)$ to exist, $\partial_x^2 v(0,0) = \partial_x^2 v(0,\pi) = 0$ must hold. Finally, for $U^{(4)}(0)$ to exist, a straightforward computation shows that $\partial_x^4 u(0,x) + f''(0)\, u_x^2(0,x) = f'(0)\, f(0)$ must hold at $x = 0,\pi$. This nonlinear boundary condition is difficult to handle, and in this situation the space of initial conditions allowing temporally smooth solutions is not an open set in a suitable Hilbert space. Therefore, we restrict our attention to nonlinearities B of the semilinear evolution equation (2.4) which satisfy condition (B1).

2.5.4. *The semilinear wave equation on the line.* When $I = \mathbb{R}$, we take $\mathcal{Y} = \mathcal{H}_1(\mathbb{R}) \times \mathcal{L}_2(\mathbb{R})$. Using the Fourier transform, we verify that e^{tA} is unitary on $\mathcal{Y}$. Lemma 2.10 remains valid with $I = \mathbb{R}$, but the assertions of Theorem 2.12 only hold true provided $0 \in G$ and $f(0) = 0$. For example, when f is a polynomial without constant term and $\mathcal{D}_k = \mathcal{B}_R^{\mathcal{Y}_k}(0)$ for some $R > 0$, then B satisfies condition (B1) and Theorem 2.4 applies.

2.5.5. *A semilinear wave equation in an inhomogeneous material.* Instead of (2.5), let us consider the non-constant coefficient semilinear wave equation

$$\partial_{tt} u = \partial_x (a \, \partial_x u) + b \, u + f(u)$$

where $a, b \in \mathcal{C}_{\mathrm{b}}^N(I; \mathbb{R})$ with $a(x) > 0$ and $b(x) \leq 0$ for $x \in I$. For periodic boundary conditions and on the line, the setting and conclusions of Sections 2.5.1 and 2.5.4 apply. For Dirichlet boundary conditions on $I = [0, \pi]$, the spaces $\mathcal{Y}_k$ also carry over from Section 2.5.3 and it is straightforward to verify that (B1) is satisfied with $K = 4$ provided $f(0) = f''(0) = 0$ and $N > K$.

We remark that the semilinear wave equation in inhomogeneous media can, in principle, be solved numerically by splitting methods (see the introduction for references). Here, however, splitting methods lose their advantage because the explicit computation of e^{tA} is expensive for operators with non-constant coefficients.

2.6. **Example: the nonlinear Schrödinger equation.** We first consider periodic boundary conditions. In this case, the Laplacian is diagonal in the Fourier representation with eigenvalues $-k^2$ and A generates a unitary group on $\mathcal{L}_2(I; \mathbb{C})$ and, more generally, on $\mathcal{H}_\ell(I; \mathbb{C})$ with $\ell \in \mathbb{N}_0$.

In the notation of Section 2.2, we choose $\mathcal{Y}_\ell = \mathcal{H}_{2\ell+1}(I; \mathbb{C})$. Then (A0) is satisfied. If the potential $V(u, \overline{u})$ satisfies $V \in \mathcal{C}_{\mathrm{b}}^{K+2+N}(G; \mathbb{R})$ with $K < N$ for some open subset $G \subset \mathbb{R}^2 \equiv \mathbb{C}$ then, by Theorem 2.12, the nonlinearity B defined in (2.8) satisfies assumption (B1) with $\mathcal{D} = \mathcal{G}_1$ from Theorem 2.12 and $\mathcal{D}_k$ defined recursively as for the semilinear wave equation (Section 2.5.1). Therefore, Theorem 2.4 and Remark 2.9 assert the existence of a flow Φ on $\mathcal{Y}$ and specify its regularity.

In the case of Neumann boundary conditions, we choose $\mathcal{Y}_\ell = \mathcal{H}_{2\ell+1}^{\mathrm{nb}}(I; \mathbb{C})$ with $I = [0, \pi]$ and $\mathcal{H}_{2\ell+1}^{\mathrm{nb}}$ defined in Section 2.5.2, so that (B1) is satisfied.

In the case of Dirichlet boundary conditions, we choose $\mathcal{Y}_\ell = \mathcal{H}_{2\ell+1}^0(I; \mathbb{C})$ as defined in Section 2.5.3. Then (B1) is satisfied for any $V(u, \bar{u}) = v(|u|^2)$ where $v \in \mathcal{C}_{\mathrm{b}}^{K+2+N}(\mathbb{R}_0^+; \mathbb{R})$, in particular for the standard case where $V(u) = |u|^4/2$.

On the line, $\mathcal{Y}_\ell = \mathcal{H}_{2\ell+1}(\mathbb{R}; \mathbb{C})$ and condition (B1) is satisfied if, for example, $V(u)$ is a polynomial in $u_1 = \mathrm{Re}(u)$ and $u_2 = \mathrm{Im}(u)$ with no linear term, so that $f(0) = \partial_{\bar{u}} V(0,0) = 0$.

Remark 2.13. While the setup in this section concern PDEs in one spatial dimension, our results on superposition operators can be extended to "nice" n-dimensional spatial domains since Lemma 2.10 holds on $\mathcal{H}_\ell(\Omega, \mathbb{R}^d)$ for $\ell > n/2$ [1], when, e.g., $\Omega \subset \mathbb{R}^n$ is a domain with smooth boundary or $\Omega = \mathbb{R}^n$. So we could also consider the nonlinear Schrödinger equation on $\mathbb{R}^2$ and $\mathbb{R}^3$.

Remark 2.14 (Inhomogeneous boundary conditions)*.* We can treat inhomogeneous time-independent mixed linear boundary conditions of the form $\mathrm{BC}(U) = g$ for the above examples by solving $Av = 0$, $\mathrm{BC}(v) = g$ and then applying a Runge–Kutta method to $U - v$. This is equivalent to applying a Runge-Kutta method to U with boundary conditions $\mathrm{BC}(U) = g$, cf. the discussion in [24].

3. A-stable Runge–Kutta methods on Hilbert spaces

In this section, we first prove an abstract convergence result for discretizations of evolution equations on Hilbert spaces. Then, in Section 3.2, we introduce a class of A-stable Runge–Kutta methods which are well defined when applied to the semilinear PDE (2.4) under assumptions (A0) and (B0). In Section 3.3, we study the regularity of A-stable Runge–Kutta methods under the additional condition (B1) and finally apply the abstract convergence result to those schemes.

3.1. **An abstract convergence theorem on Hilbert spaces.** In this section we prove an abstract convergence result for evolution equations on Hilbert spaces, Theorem 3.1. Although this theorem is modeled after the basic local convergence result for ODEs and there are a lot of results on the convergence of time discretizations of specific PDEs in the literature, see Section 1, we are not aware of any result that is as general as this theorem.

In the classical setting of ordinary differential equations $\dot{y} = f(y)$, a one-step method $y^{n+1} = \psi^h(y^n)$ is of order p if, formally, $y(h) - \psi^h(y^0) = O(h^{p+1})$. In other words, the local error is controlled by the Taylor integral remainder of order $p+1$. It is then easy to show that the method is globally convergent of order p; see, e.g., [13].

The situation is more subtle in the case of a differential equation

$$\dot{U} = F(U) \tag{3.1}$$

on a Hilbert space $\mathcal{X}$: First, it is not clear whether the time-h map Ψ^h associated with a given one-step method applied to (3.1) is well defined as map from an open subset of $\mathcal{X}$ to itself. It depends on the equation and on the chosen one-step method, and typically fails for explicit Runge–Kutta methods. Second, even if $U \mapsto \Psi^h(U)$ is well defined and continuous, its derivatives with respect to h will usually fail to be defined on the same set. Thus, in order to control the Taylor remainder $U(h) - \Psi^h(U^0)$ in the case of a discretization of a PDE (3.1), we must consider the remainder as a map from a space $\mathcal{Z}$ of high regularity into a space $\mathcal{X}$ of low regularity. In this setting, the usual proof that consistent one-step methods are convergent applies under the following assumptions.

Let $\mathcal{X}$ and $\mathcal{Z} \subset \mathcal{X}$ be Hilbert spaces, where $\mathcal{Z}$ is continuously embedded in $\mathcal{X}$ and let Ψ^h be a one-step discretization of (3.1) which is of classical order p. Assume there exist sets $\mathcal{D}_\mathcal{X} \subset \mathcal{X}$ and $\mathcal{D}_\mathcal{Z} \subset \mathcal{Z}$ such that $\mathcal{D}_\mathcal{X}$ is open in $\mathcal{X}$, $\mathcal{D}_\mathcal{Z} \subset \mathcal{D}_\mathcal{X}$, and there exist constants $h_* > 0$, $\Theta_* > 0$, such that the following hold.

(C1) For fixed $h \in [0, h_*]$, the map $U \mapsto \Psi^h(U)$ is $\mathcal{C}^1(\mathcal{D}_\mathcal{X}; \mathcal{X})$. Moreover, there exists a possibly h-dependent norm $\|\cdot\|_{\mathcal{X},h}$ on $\mathcal{X}$ with

$$\|U\|_\mathcal{X} \leq \|U\|_{\mathcal{X},h} \leq \Theta_* \|U\|_\mathcal{X} \tag{3.2}$$

for all $U \in \mathcal{X}$ and $h \in [0, h_*]$ such that

$$\sup_{U \in \mathcal{D}_\mathcal{X}} \|\mathrm{D}\Psi^h(U)\|_{\mathcal{E}(\mathcal{X}),h} = 1 + O(h) \tag{3.3}$$

for all $h \in [0, h_*]$. Here, $\|\cdot\|_{\mathcal{E}(\mathcal{X}),h}$ denotes the operator norm induced by $\|\cdot\|_{\mathcal{X},h}$.

(C2) For fixed $U \in \mathcal{D}_{\mathcal{Z}}$, the map $h \mapsto \Psi^h(U)$ is in $\mathcal{C}^{p+1}([0,h_*];\mathcal{X})$, and

$$\sup_{\substack{U\in\mathcal{D}_{\mathcal{Z}}\\ h\in[0,h_*]}} \|\partial_h^{p+1}\Psi^h(U)\|_{\mathcal{X}} < \infty. \tag{3.4}$$

Condition (C1) can be seen as a stability condition, whereas condition (C2) ensures consistency.

Theorem 3.1. *In the setting above, fix $U^0 \in \mathcal{D}_{\mathcal{Z}}$ and suppose that there exists a solution*

$$U \in \mathcal{C}([0,T];\mathcal{D}_{\mathcal{Z}}) \cap \mathcal{C}^{p+1}([0,T];\mathcal{D}_{\mathcal{X}}) \tag{3.5}$$

to the initial value problem (3.1) for some $T>0$ with $U(0)=U^0$. Let Ψ^h be a one-step discretization of (3.1) of order $p \geq 1$; let $U^m = (\Psi^h)^m(U^0)$ denote the associated numerical solution.

Then there exist constants $h_*>0$, c_1, and c_2, depending only on T, the norm of U in $\mathcal{C}^{p+1}([0,T];\mathcal{X})$, $\operatorname{dist}_{\mathcal{X}}(\{U(t)\colon t\in[0,T]\},\partial\mathcal{D}_{\mathcal{X}})$, and on the constants from (3.2), (3.3) and (3.4), such that for every $h \in [0,h_*]$,

$$\|U^m - U(mh)\|_{\mathcal{X}} \leq c_2\,\mathrm{e}^{c_1 mh}\,h^p$$

so long as $mh \leq T$.

Proof. Since $\mathcal{D}_{\mathcal{X}}$ is open, there is some $\delta>0$ such that $\mathcal{B}_\delta^{\mathcal{X}}(U(t)) \subset \mathcal{D}_{\mathcal{X}}$ for each $t\in[0,T]$. Setting

$$E_m = \|U^m - U(mh)\|_{\mathcal{X},h}\,,$$

we estimate, with $\Phi^t(U(s)) = U(t+s)$,

$$\begin{aligned}
E_{m+1} &\leq \|\Psi^h(U^m) - \Psi^h(U(mh))\|_{\mathcal{X},h} + \|\Psi^h(U(mh)) - \Phi^h(U(mh))\|_{\mathcal{X},h} \\
&\leq \sup_{\theta\in[0,1]} \|\mathrm{D}\Psi^h(U(mh)+\theta\,(U^m-U(mh)))\|_{\mathcal{E}(\mathcal{X}),h}\,E_m \\
&\quad + \frac{\Theta_*\,h^{p+1}}{(p+1)!}\sup_{s\in[0,h]}\Big(\|\partial_s^{p+1}\Psi^s(U(mh))\|_{\mathcal{X}} + \|\partial_s^{p+1}\Phi^s(U(mh))\|_{\mathcal{X}}\Big) \\
&\leq \sup_{U\in\mathcal{D}_{\mathcal{X}}} \|\mathrm{D}\Psi^h(U)\|_{\mathcal{E}(\mathcal{X}),h}\,E_m \\
&\quad + \frac{\Theta_*\,h^{p+1}}{(p+1)!}\sup_{t\in[0,T]}\Big(\sup_{h\in[0,h_*]}\|\partial_h^{p+1}\Psi^h(U(t))\|_{\mathcal{X}} + \|\partial_t^{p+1}U(t)\|_{\mathcal{X}}\Big) \\
&\leq (1+c_1\,h)\,E_m + c_3\,h^{p+1}\,.
\end{aligned}$$

The suprema in the estimate above are finite due to (3.3), (3.4) and (3.5), respectively, so long as $E_m<\delta$ since then, due to (3.2), $U(mh)+\theta\,(U^m-U(mh)) \in \mathcal{D}_{\mathcal{X}}$.

Thus, since $E_0=0$,

$$E_m \leq c_3\,h^{p+1}\,\frac{(1+h\,c_1)^m-1}{h\,c_1} \leq \frac{c_3}{c_1}\left(1+\frac{mh\,c_1}{m}\right)^m h^p \leq c_2\,\mathrm{e}^{c_1 mh}\,h^p\,.$$

Thus, we can choose h_* small enough such that $E_m<\delta$ for all $m \leq T/h_*$. This concludes the proof. □

Remark 3.2. The proof of Theorem 3.1 does not use any Hilbert space structure, so that the result holds true when $\mathcal{X}$ and $\mathcal{Z}$ are Banach spaces. However, condition (C1) is rather restrictive on general Banach spaces, see Remark 3.7 below and the discussion in the introduction.

3.2. **Regularity of A-stable Runge–Kutta discretizations.** Applying an s-stage Runge–Kutta method to the semilinear evolution equation (2.4), we obtain

$$W = U^0 \, \mathbb{1} + h \, \mathsf{a} \left(AW + B(W) \right), \tag{3.6a}$$

$$U^1 = U^0 + h \, \mathsf{b}^T \left(AW + B(W) \right). \tag{3.6b}$$

We write, with $U \in \mathcal{Y}$,

$$\mathbb{1} U = \begin{pmatrix} U \\ \vdots \\ U \end{pmatrix} \in \mathcal{Y}^s \,, \quad W = \begin{pmatrix} W^1 \\ \vdots \\ W^s \end{pmatrix} \,, \quad B(W) = \begin{pmatrix} B(W^1) \\ \vdots \\ B(W^s) \end{pmatrix} \,,$$

where $W^1, \dots, W^s$ are the stages of the Runge–Kutta method,

$$(\mathsf{a} W)^i = \sum_{j=1}^s \mathsf{a}_{ij} \, W^j \,, \qquad \mathsf{b}^T W = \sum_{j=1}^s \mathsf{b}_j \, W^j \,,$$

and A acts diagonally on the stages, i.e., $(AW)^i = AW^i$ for $i = 1, \dots, s$.

Written this way, it is not transparent that, under certain conditions, this class of methods results in a well defined numerical time-h map Ψ^h on a Hilbert space $\mathcal{Y}$. A more suitable form is achieved by rewriting (3.6a) as

$$W = \Pi(W; U, h) \equiv (\mathrm{id} - h\mathsf{a}A)^{-1} \left(\mathbb{1} U + h \mathsf{a} B(W) \right). \tag{3.7}$$

Noting that

$$(\mathrm{id} - h\mathsf{a}A)^{-1} = \mathrm{id} + h\mathsf{a}A \, (\mathrm{id} - h\mathsf{a}A)^{-1} \tag{3.8}$$

and inserting (3.7) into (3.6b), we obtain

$$\begin{aligned} \Psi^h(U) &= U + h\mathsf{b}^T \left(AW(U,h) + B(W(U,h)) \right) \\ &= \mathsf{S}(hA)U + h\mathsf{b}^T \, (\mathrm{id} - h\mathsf{a}A)^{-1} \, B(W(U,h)) \,, \end{aligned} \tag{3.9}$$

where S is the so-called *stability function*

$$\mathsf{S}(z) = 1 + z\mathsf{b}^T \, (\mathrm{id} - z\mathsf{a})^{-1} \, \mathbb{1} \,. \tag{3.10}$$

We now make a number of assumptions on the method and its interaction with the linear operator A. First, we assume that the method is A-stable in the sense of [24]. Setting $\mathbb{C}^- = \{ z \in \mathbb{C} \colon \operatorname{Re} z \leq 0 \}$, the conditions are as follows.

(RK1) The stability function (3.10) is bounded with $|\mathsf{S}(z)| \leq 1$ for all $z \in \mathbb{C}^-$.
(RK2) The $s \times s$ matrices $\mathrm{id} - z\mathsf{a}$ are invertible for all $z \in \mathbb{C}^-$.

Sometimes, we will also assume that a is invertible.

Remark 3.3. The matrix $\mathrm{id} - z\mathsf{a}$ is invertible for all $z \in \mathbb{C}^-$ if and only if a has no eigenvalues in $\mathbb{C}^- \setminus \{0\}$. Its inverse is then bounded uniformly for $z \in \mathbb{C}^-$ by a constant $\Lambda \geq 1$ (insert, in particular, $z = 0$).

Remark 3.4. In general, Runge–Kutta methods are called $A(\theta)$-stable for some $\theta \in [0, \pi/2]$ if $|\mathsf{S}(z)| \leq 1$ for all $z \in \mathbb{C}$ with $|\arg(-z)| \leq \theta$; see, e.g., [13]. A definition of $A(\theta)$-stability that requires, in addition, invertibility of $\mathrm{id} - z\mathsf{a}$ was introduced by Lubich and Ostermann [24] in the context of parabolic equations; their results also depend, to a large extent, on the invertibility of a. Thus, our assumptions can be described as $A(\theta)$-stability for $\theta = \pi/2$ in the sense of [24]. Note that the requirement $\theta = \pi/2$ arises as we include operators A which are not necessarily sectorial, but whose spectrum may, for example, contain a strip about the imaginary axis, cf. Sections 2.5 and 2.6.

Example 3.5. The implicit midpoint rule has stability function $\mathsf{S}(z) = (1+z/2)/(1-z/2)$, $s = 1$, $\mathsf{a}_{11} = \frac{1}{2}$, and $\mathsf{b}_1 = 1$. Conditions (RK1) and (RK2) are readily verified; moreover, a is invertible.

Lemma 3.6. *Gauss–Legendre Runge–Kutta methods satisfy* (RK1) *and* (RK2) *with* a *invertible.*

Proof. Condition (RK1) is the classical notion of A-stability; it is proved for Gauss–Legendre methods in [13, Theorem 6.44], for example.

To verify condition (RK2), write $\mathsf{S}(z) = \mathsf{P}(z)/\mathsf{Q}(z)$ as the quotient of polynomials P and Q with no common roots. We claim that $\mathsf{Q}(z) = \det(\mathrm{id} - z\mathsf{a})$. To see this, note first that $\det(\mathrm{id} - z\mathsf{a})$ arises naturally as the common denominator when solving for the terms of an explicit rational expansion of $(\mathrm{id} - z\mathsf{a})^{-1}$ by Cramer's rule; see the proof of [13, Lemma 6.30]. The claim follows if we can show that the numerator does not have any factor in common with $\det(\mathrm{id} - z\mathsf{a})$. Since $p = 2s$ for Gauss–Legendre methods [13, Theorem 6.43], $\deg \mathsf{Q} \leq s$ and $\deg \mathsf{P} \leq s$ for s-stage implicit Runge–Kutta methods [13, Lemma 6.30] and, generally, $p \leq \deg \mathsf{P} + \deg \mathsf{Q}$ [13, Lemma 6.4], we conclude that $\deg \mathsf{P} = \deg \mathsf{Q} = s$ so that indeed $\mathsf{Q}(z) = \det(\mathrm{id} - z\mathsf{a})$.

Since, by (RK1), the rational function S is nonsingular on $\mathbb{C}^-$, all eigenvalues of a must lie outside of $\mathbb{C}^- \setminus \{0\}$. This proves invertibility of $\mathrm{id} - z\mathsf{a}$ on $\mathbb{C}^-$, cf. Remark 3.3. Finally, since $\mathsf{Q}(z) = \det(\mathrm{id} - z\mathsf{a})$ has degree s, a must also be nonsingular. □

For the convergence analysis in Section 3.3, we need the following additional assumption on the operator A and on the scheme.

(A1) Assumption (A0) holds, and there exist constants $\omega_{\mathsf{S}}, \Theta_{\mathsf{S}}, h_* > 0$ such that for all $h \in [0, h_*]$ and $n \in \mathbb{N}$,

$$\|\mathsf{S}^n(hA)\|_{\mathcal{Y} \to \mathcal{Y}} \leq \Theta_{\mathsf{S}} \, \mathrm{e}^{\omega_{\mathsf{S}} n h} \,. \tag{3.11}$$

If assumption (A1) holds, we define, for $U \in \mathcal{Y}$,

$$\|U\|_{\mathcal{Y},h} \equiv \sup_{n \in \mathbb{N}_0} \mathrm{e}^{-n\omega_{\mathsf{S}} h} \, \|\mathsf{S}^n(hA)U\|_{\mathcal{Y}} \,. \tag{3.12}$$

Then $\|\cdot\|_{\mathcal{Y},h}$ is equivalent to the $\mathcal{Y}$-norm in the sense of (3.2) with $\Theta_* = \Theta_{\mathsf{S}}$. Moreover, there is some $\sigma > 0$ such that

$$\|\mathsf{S}(hA)\|_{\mathcal{E}(\mathcal{Y}),h} \leq \mathrm{e}^{\omega_{\mathsf{S}} h} \leq 1 + \sigma \, h \tag{3.13}$$

for $h \in [0, h_*]$.

Remark 3.7. When an A-stable Runge–Kutta is applied to discretize a general C_0-semigroup e^{tA} on a Banach space, estimate (3.11) is in general false. A counterexample is the implicit midpoint rule applied to $A = \partial_x$ on $\mathcal{L}_1(\mathbb{R})$ [21]. When A is a sectorial operator, then (3.11) is satisfied [26].

Remark 3.8. In the time-continuous case discussed in Section 2, the estimate corresponding to (3.11) is (2.9). Note that, by replacing the $\mathcal{Y}$-norm with the equivalent norm $\|U\| = \sup_{t \geq 0} \mathrm{e}^{-\omega t} \, \|\mathrm{e}^{tA} U\|_{\mathcal{Y}}$ the constant Θ in (2.9) becomes 1, analogous to (3.13).

We state the following sufficient condition for (A1), which is often satisfied in applications.

(A2) Assumption (A0) holds, $\mathcal{Y}$ is a Hilbert space, and $A = A_{\rm n} + A_{\rm b}$ with $A_{\rm n}$ normal and $A_{\rm b}$ bounded as a linear operator on each $\mathcal{Y}_0, \ldots, \mathcal{Y}_K$.

(Recall that an operator A is normal if it is closed and $AA^* = A^*A$.) Condition (A2) implies that the non-normal part $A_{\rm b}$ of A can be included with B as it satisfies the sufficient condition (B1). Note that $A_{\rm b}$ is a bounded linear operator on each $\mathcal{Y}_k$ if, for example, $A_{\rm b} = \mathbb{P}A$ and $A_{\rm n} = \mathbb{Q}A$ is normal, where $\mathbb{P}$ is a spectral projector of A onto a finite dimensional subspace and $\mathbb{Q} = \mathrm{id} - \mathbb{P}$.

Remark 3.9. In the case of the semilinear wave equation, see Section 2.5, assumption (A2) is satisfied with $A_{\rm b} = \mathbb{P}_0 A$, where $\mathbb{P}_0$ denotes the spectral projection corresponding to the eigenvalue 0 of A. In the case of the nonlinear Schrödinger equation, see Section 2.6, the operator A is normal, so that (A2) holds trivially.

Lemma 3.10. *Assume that* (RK1) *and* (RK2) *hold and that A satisfies conditions* (A2)*. Then* (3.11) *is satisfied with* $\Theta_{\mathsf{S}} = 1$.

Before we can prove Lemma 3.10, we need some technical estimates on the operators which appear on the right of Eq. (3.9) and Eq. (3.7). In the following, we denote s copies of $\mathcal{Y}$ by $\mathcal{Y}^s$ and use the norm

$$\|W\|_{\mathcal{Y}^s} = \max_{j=1,\ldots,s} \|W^j\|_{\mathcal{Y}} \,.$$

Lemma 3.11. *Assume* (RK2) *and* (A0)*. Then, for $h_* > 0$ small enough, there exist $\Lambda \geq 1$ and $c_{\mathsf{S}} \geq 1$ such that*

$$\|(\mathrm{id} - h\mathsf{a}A)^{-1}\|_{\mathcal{Y}^s \to \mathcal{Y}^s} \leq \Lambda \tag{3.14a}$$

and

$$\|h\mathsf{a}A(\mathrm{id} - h\mathsf{a}A)^{-1}\|_{\mathcal{Y}^s \to \mathcal{Y}^s} \leq 1 + \Lambda \tag{3.14b}$$

for all $h \in [0, h_*]$. Moreover, for any $\ell, n, \in \mathbb{N}_0$,

$$(W,h) \mapsto (\mathrm{id} - h\mathsf{a}A)^{-1} W \text{ is a map of class } \mathcal{C}_{\rm b}^{(\underline{n},\ell)}(\mathcal{Y}_\ell^s \times [0,h_*]; \mathcal{Y}^s)\,, \tag{3.15a}$$

$$(W,h) \mapsto h\mathsf{a}A(\mathrm{id} - h\mathsf{a}A)^{-1} W \text{ is a map of class } \mathcal{C}_{\rm b}^{(\underline{n},\ell)}(\mathcal{Y}_\ell^s \times [0,h_*]; \mathcal{Y}^s)\,, \tag{3.15b}$$

and

$$(W,h) \to h(\mathrm{id} - h\mathsf{a}A)^{-1} W \text{ is a map of class } \mathcal{C}_{\rm b}^{(\underline{n},\ell+1)}(\mathcal{Y}_\ell^s \times [0,h_*]; \mathcal{Y}^s)\,. \tag{3.15c}$$

Remark 3.12. Estimates of the form (3.14) were proved in [24] under the assumption that A is sectorial.

Proof. Transforming a into Jordan normal form, we see that there exists a constant $c = c(\mathsf{a})$ such that

$$\|(\mathrm{id} - h\mathsf{a}A)^{-1}\|_{\mathcal{Y}^s \to \mathcal{Y}^s} \leq c \max_{i=1,\ldots,k} \|(\mathrm{id} - h\lambda_i A)^{-1}\|_{\mathcal{Y} \to \mathcal{Y}}^{m_i}$$

where $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of a with algebraic multiplicities $m_1, \ldots, m_k$.

Hence, let λ be one of the eigenvalues of a; we know that $\operatorname{Re}\lambda > 0$ due to assumption (RK2) and Remark 3.3. Referring to (2.10), we estimate, for $h \geq 0$,

$$\|(\mathrm{id} - h\lambda A)^{-1}\|_{\mathcal{Y}\to\mathcal{Y}} \leq \frac{1}{|h\lambda|} \frac{\Theta}{\operatorname{Re}\frac{1}{h\lambda} - \omega} = \frac{\Theta}{\frac{|\lambda|}{\operatorname{Re}\lambda} - |h\lambda|\,\omega} .$$

Thus, the right hand bound is positive and finite for all $h \in [0, h_*]$ provided that $h_* > 0$ is small enough. This proves estimate (3.14a). Due to identity (3.8), estimate (3.14b) follows immediately.

To prove continuity of the map $(\mathrm{id} - h\mathsf{a}A)^{-1}W \colon [0, h_*] \to W$ for fixed $W \in \mathcal{Y}^s$, we proceed as follows. Let $\varepsilon > 0$. Then for every $W_1 \in \mathcal{Y}_1^s$, $h, h' \in [0, h_*]$,

$$\begin{aligned}
&\|(\mathrm{id} - h\mathsf{a}A)^{-1}W - (\mathrm{id} - h'\mathsf{a}A)^{-1}W\|_{\mathcal{Y}^s} \\
&\quad \leq \|((\mathrm{id} - h\mathsf{a}A)^{-1} - (\mathrm{id} - h'\mathsf{a}A)^{-1})W_1\|_{\mathcal{Y}^s} + \|(\mathrm{id} - h\mathsf{a}A)^{-1}(W - W_1)\|_{\mathcal{Y}^s} \\
&\qquad + \|(\mathrm{id} - h'\mathsf{a}A)^{-1}(W - W_1)\|_{\mathcal{Y}^s} \\
&\quad \leq \|((\mathrm{id} - h\mathsf{a}A)^{-1} - (\mathrm{id} - h'\mathsf{a}A)^{-1})W_1\|_{\mathcal{Y}^s} + 2\Lambda\,\|W - W_1\|_{\mathcal{Y}} ,
\end{aligned} \tag{3.16}$$

where the second inequality is based on (3.14a). Now, since A is assumed to be densely defined and $\mathcal{Y}_1 = D(A)$, we can choose W_1 so close to W that the last term on the right is less than $\varepsilon/2$. Then, since $W_1 \in \mathcal{Y}_1^s$, there exists a $\delta = \delta(W_1)$ such that the first term on the right is less than $\varepsilon/2$ whenever $|h - h'| < \delta$. This proves continuity of $h \mapsto (\mathrm{id} - h\mathsf{a}A)^{-1}W$ on the interval $[0, h_*]$.

To complete the proof of (3.15), we must compute the h-derivatives of the map (3.15a). Once we have shown (3.15a), estimate (3.15b) follows immediately via (3.8). First,

$$\partial_h^\ell (\mathrm{id} - h\mathsf{a}A)^{-1} = \ell!\,(\mathsf{a}A)^\ell\,(\mathrm{id} - h\mathsf{a}A)^{-\ell-1} .$$

Using estimates (2.13) and (3.14a), and noting the continuity of $h \mapsto (\mathrm{id} - h\mathsf{a}A)^{-1}W$ proved above, (3.15a) follows. Finally, noting that

$$\partial_h[h(\mathrm{id} - h\mathsf{a}A)^{-1}] = (\mathrm{id} - h\mathsf{a}A)^{-1} + h\mathsf{a}A(\mathrm{id} - h\mathsf{a}A)^{-2} = (\mathrm{id} - h\mathsf{a}A)^{-2} ,$$

we obtain

$$\partial_h^\ell[h(\mathrm{id} - h\mathsf{a}A)^{-1}] = \partial_h^{\ell-1}(\mathrm{id} - h\mathsf{a}A)^{-2} = \ell!\,(\mathsf{a}A)^{\ell-1}\,(\mathrm{id} - h\mathsf{a}A)^{-\ell-1} ,$$

which implies (3.15c). □

Proof of Lemma 3.10. Recall that $\operatorname{Re}(\operatorname{spec} A) \leq \omega$ for some $\omega > 0$ so that the spectrum of $A - \omega$ is contained in $\mathbb{C}^-$. Moreover, by assumption (A2) we can split A into a normal part A_{n} and a bounded part A_{b}. Now decompose $A = A_1 + A_2$ with $A_1 = A_{\mathrm{n}} - \omega$ and $A_2 = A_{\mathrm{b}} + \omega$. We now apply the Runge–Kutta scheme to the linear problem $\partial_t U = AU$ in two different ways. First, we take the full A and $B \equiv 0$; second we take A replaced by A_1 and $B(U) = A_2 U$. Since the respective numerical time-h maps given by (3.9) must be the same, we obtain the identity

$$\mathsf{S}(hA) = \mathsf{S}(hA_1) + h\mathsf{b}^T\,(\mathrm{id} - h\mathsf{a}A_1)^{-1}\,A_2 W \tag{3.17a}$$

where

$$W = (\mathrm{id} - h\mathsf{a}A_1)^{-1}\,(\mathbb{1} + h\mathsf{a}A_2 W) . \tag{3.17b}$$

Rewrite (3.17b) as $W = MW + G$ with

$$M = (\mathrm{id} - h\mathsf{a}A_1)^{-1}\,h\mathsf{a}A_2 \qquad \text{and} \qquad G = (\mathrm{id} - h\mathsf{a}A_1)^{-1}\,\mathbb{1} .$$

Since A_2 is bounded and, by Lemma 3.11, $(\mathrm{id} - h\mathsf{a}A_1)^{-1}$ is uniformly bounded for $h \in [0, h_*]$, the matrix M has norm smaller than one for $h \in [0, h_*]$ with some possibly smaller $h_* > 0$. Consequently, we can solve for $W = (\mathrm{id} - M)^{-1}G$, whence the second term on the right of (3.17a) is $O(h)$ in the norm of $\mathcal{E}(\mathcal{Y})$.

As A_1 is normal with $\operatorname{spec} A_1 \subset \mathbb{C}^-$, we have, referring to (RK1),

$$\|\mathsf{S}(hA_1)\|_{\mathcal{Y}\to\mathcal{Y}} \le \sup_{\lambda \in \operatorname{spec} A_1} |\mathsf{S}(h\,\lambda)| \le 1\,.$$

Altogether, this proves that there exists $\sigma > 0$ such that $\|\mathsf{S}(hA)\|_{\mathcal{Y}\to\mathcal{Y}} \le 1 + \sigma\, h$ for $h \in [0, h_*]$. This in turn implies (3.11) with $\Theta_{\mathsf{S}} = 1$. □

Next, we describe the differentiability properties of $\mathsf{S}(hA)$ which will be needed later on.

Lemma 3.13. *Assume* (RK2), (A0), *and either that the Runge–Kutta matrix* a *is invertible or that* (A1) *holds. Then there exist* $h_* > 0$ *and* $c_{\mathsf{S}} \ge 1$ *such that for all* $h \in [0, h_*]$,

$$\|\mathsf{S}(hA)\|_{\mathcal{Y}\to\mathcal{Y}} \le c_{\mathsf{S}} \tag{3.18}$$

and, for all $\ell, n \in \mathbb{N}_0$,

$$(U, h) \mapsto \mathsf{S}(hA)U \text{ is a map of class } \mathcal{C}_{\mathrm{b}}^{(\underline{n},\ell)}(\mathcal{Y}_\ell \times [0, h_*]; \mathcal{Y})\,. \tag{3.19}$$

Proof. First, (3.18) clearly holds when (A1) holds. To prove (3.18) when a is invertible, we estimate, using (3.10) and (3.14b),

$$\|\mathsf{S}(hA)\|_{\mathcal{Y}\to\mathcal{Y}} \le 1 + s\,\|\mathsf{b}\|\,\|\mathsf{a}^{-1}\|\,(1+\Lambda) \equiv c_{\mathsf{S}}\,.$$

Next, we show that $\mathsf{S}(hA)U \colon [0, h_*] \to \mathcal{Y}$ is continuous for every $U \in \mathcal{Y}$ as in the proof of Lemma 3.11, replacing $(\mathrm{id} - h\mathsf{a}A)^{-1}$ by $\mathsf{S}(hA)$, Λ by c_{S}, and $\mathcal{Y}^s$ by $\mathcal{Y}$ in (3.16). This proves (3.19) for $\ell = 0$.

Finally, to prove (3.19) for $\ell \in \mathbb{N}$, we note that, due to (3.15c), the map $(W, h) \mapsto hA(\mathrm{id} - h\mathsf{a}A)^{-1}W$ is of class $\mathcal{C}_{\mathrm{b}}^{(\underline{n},\ell)}(\mathcal{Y}_\ell^s \times [0, h_*]; \mathcal{Y}^s)$; the claim then follows directly from the definition of S in (3.10). □

In Theorem 2.4, we studied differentiability in time of the semiflow Φ^t of (2.4). An analogous result holds for differentiability of the discretization Ψ^h of (2.4) in the step size h.

Theorem 3.14 (Existence and regularity of numerical method, local version)**.** *Assume that the semilinear evolution equation* (2.4) *satisfies conditions* (A0) *and* (B1), *and apply a Runge–Kutta method subject to condition* (RK2) *to it. Moreover, assume that* (A1) *holds or that the Runge–Kutta matrix* a *is invertible. Choose* $R \in (0, \delta_*]$ *such that* $\mathcal{D}_K^{-R} \neq \emptyset$ *and pick* $U^0 \in \mathcal{D}_K^{-R}$. *Let* $R_* = R/(2\max\{c_{\mathsf{S}}, \Lambda\})$ *with* c_{S} *from* (3.18) *and* Λ *from* (3.14). *Then, for sufficiently small* $h_* > 0$, *there exists a unique stage vector* W *and numerical time-h map* $\Psi(U, h) = \Psi^h(U)$ *which satisfy*

$$W^i, \Psi \in \bigcap_{\substack{j+k \le N \\ \ell \le k \le K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{B}_{R_*}^{\mathcal{Y}_K}(U^0) \times [0, h_*]; \mathcal{B}_R^{\mathcal{Y}_{k-\ell}}(U^0)) \tag{3.20}$$

for $i = 1, \ldots, s$. *In particular,*

$$W^i, \Psi \in \mathcal{C}_{\mathrm{b}}^K(B_{R_*}^{\mathcal{Y}_K}(U^0) \times [0, h_*]; \mathcal{B}_R^{\mathcal{Y}}(U^0)) \tag{3.21}$$

for $i = 1, \dots, s$. The bounds on W, Ψ and h_* depend only on the bounds afforded by (B1), (3.14), (3.18), *on the coefficients of the method, R, and U^0. If, in addition,* (A1) *holds, then there exists a constant σ_Ψ, such that for $h \in [0, h_*]$ with a possibly smaller choice of $h_* > 0$,*

$$\sup_{U \in B^{\mathcal{Y}}_{R_*}(U^0)} \|\mathrm{D}\Psi^h(U)\|_{\mathcal{E}(\mathcal{Y}),h} \leq 1 + \sigma_\Psi h \,, \tag{3.22}$$

where the norm on the left is defined by (3.12) *and h_* and σ_Ψ depend only on the above quantities and on the constants in* (A1)*.*

Note that statement (3.21) for Ψ is analogous to (2.14a) for the semiflow Φ.

Proof. We apply the contraction mapping theorem on a scale of Banach spaces, Theorem A.9, to the map from (3.7),

$$\Pi(W; U, h) \equiv (\mathrm{id} - h\mathsf{a}A)^{-1}\, \mathbb{1}U + h\mathsf{a}\,(\mathrm{id} - h\mathsf{a}A)^{-1}\, B(W)$$

with $u = U$, $w = W$, and $\mu = h$ on the scale $\mathcal{Z}_j = \mathcal{Y}^s_j$ for $j = 0, \dots K$. We further identify $\mathcal{X} = \mathcal{Y}_K$, $\mathcal{W}_j = \mathcal{B}_R(\mathbb{1}U^0) \subset \mathcal{Y}^s_j$, $\mathcal{I} = (0, h_*)$, and $\mathcal{U} = \operatorname{int} \mathcal{B}_{R_*}(U^0) \subset \mathcal{Y}_K$. To verify condition (i) of Theorem A.9, we note that Eq. (3.15a) of Lemma 3.11 asserts that the map $(U, h) \mapsto (\mathrm{id} - h\mathsf{a}A)^{-1}\, \mathbb{1}U$ is, in particular, of class

$$\bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K}} \mathcal{C}^{(\underline{j},\ell)}_{\mathrm{b}}(\mathcal{D}_K \times [0, h_*]; \mathcal{Y}^s_{k-\ell}) \,.$$

The differentiability assumptions on B from (B1) are precisely such that the map $(W, h) \mapsto h\mathsf{a}\,(\mathrm{id} - h\mathsf{a}A)^{-1}\, B(W)$ is of the same class.

First, we show that $\Pi(\,\cdot\,; U, h)$ maps $\mathcal{W}_j$, $j = 0, \dots, K$, into itself for fixed $U \in \mathcal{U}$ and $h \in [0, h_*]$ with appropriate $h_* > 0$. We begin by taking h_* as in Lemma 3.11 and estimate, for $W \in \mathcal{W}_j$,

$$\begin{aligned} \|\Pi(W; U, h) - \mathbb{1}U^0\|_{\mathcal{Z}_j} &\leq \|(\mathrm{id} - (\mathrm{id} - h\mathsf{a}A)^{-1})\mathbb{1}U^0\|_{\mathcal{Z}_j} \\ &\quad + \|(\mathrm{id} - h\mathsf{a}A)^{-1}\|_{\mathcal{Y}^s_j \to \mathcal{Z}_j}\, \|U - U^0\|_{\mathcal{Y}_j} \\ &\quad + h\, \|(\mathrm{id} - h\mathsf{a}A)^{-1}\mathsf{a}\|_{\mathcal{Y}^s_j \to \mathcal{Z}_j}\, \|B(W)\|_{\mathcal{Z}_j} \\ &\leq \|(\mathrm{id} - (\mathrm{id} - h\mathsf{a}A)^{-1})\mathbb{1}U^0\|_{\mathcal{Z}_j} + \Lambda\, R_* + h\, \Lambda\, \|\mathsf{a}\|\, M_j \,, \end{aligned} \tag{3.23}$$

where, in the last step, we have used (3.14a) from Lemma 3.11 and M_j is the bound on B from condition (B1). Since, again by Lemma 3.11, the map $h \mapsto (\mathrm{id} - h\mathsf{a}A)^{-1}W$ is continuous on each $\mathcal{Z}_j$, we can possibly shrink h_* such that the right hand side of (3.23) is less than R. This proves that $\Pi(\,\cdot\,; U, h)$ maps $\mathcal{W}_j$ into itself and implies condition (i) of Theorem A.9.

Next, for $j = 0, \dots, K$,

$$\begin{aligned} \|\mathrm{D}_W \Pi(W; U, h)\|_{\mathcal{Z}_j \to \mathcal{Z}_j} &\leq h\, \|(\mathrm{id} - h\mathsf{a}A)^{-1}\mathsf{a}\|_{\mathcal{Z}_j \to \mathcal{Z}_j}\, \|\mathrm{D}B(W)\|_{\mathcal{Z}_j \to \mathcal{Z}_j} \\ &\leq h\, \Lambda\, \|\mathsf{a}\|\, M'_j \,. \end{aligned} \tag{3.24}$$

Thus, by possibly shrinking h_* again, the right hand bound can be made less than 1. This proves that $\Pi(\,\cdot\,; U, h)$ is a contraction on $\mathcal{B}^{\mathcal{Y}^s}_R(\mathbb{1}U^0)$ uniformly for $U \in \mathcal{B}^{\mathcal{Y}}_{R_*}(U^0)$ and $h \in [0, h_*]$. Here we used that B is at least $\mathcal{C}^1$ on the highest rung $\mathcal{Y}_K$ of the scale since, in condition (B1), we require $N > K$. This verifies condition (ii) of Theorem A.9.

Theorem A.9 then applies and asserts the existence of a fixed point

$$W \in \bigcap_{\substack{j+k\leq N\\ \ell\leq k\leq K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(B_{R_*}^{\mathcal{Y}_K}(U^0)\times(0,h_*);\mathcal{B}_R^{\mathcal{Y}_{k-\ell}^s}(U^0\mathbb{1}))\,.$$

Assertion (3.21) for the W^i then follows from Lemma A.2.

To prove the corresponding estimates for Ψ^h, note that by (3.9), condition (B1), Lemma 3.11, and Lemma 3.13 we can adapt $h_* > 0$ such that for $j = 0,\dots,K$,

$$\begin{aligned}\|\Psi^h(U)-U^0\|_{\mathcal{Y}_j} &\leq \|\mathsf{S}(hA)(U-U^0)\|_{\mathcal{Y}_j} + \|\mathsf{S}(hA)U^0-U^0\|_{\mathcal{Y}_j} + h\|b\|\Lambda M_j\\ &\leq R/2 + \|\mathsf{S}(hA)U^0-U^0\|_{\mathcal{Y}_j} + h\|b\|\Lambda M_j \leq R\,.\end{aligned}$$

Further, the first term of (3.9) is of class (3.21) by Lemma 3.13. For the second term of (3.9), we note that the map Σ defined as

$$\Sigma(W,U,h) = h(\mathrm{id}-h\mathsf{a}A)^{-1}B(W)\,,$$

satisfies

$$\Sigma \in \bigcap_{\substack{i+j+k\leq N\\ \ell\leq k\leq K}} \mathcal{C}_{\mathrm{b}}^{(\underline{i},\underline{j},\ell)}((\mathcal{D}_k)^s\times\mathcal{Y}_K\times\mathcal{I};\mathcal{Y}_{k-\ell}^s)\,.$$

Lemma A.6 then implies (3.21) for Ψ. Assertion (3.21) for Ψ then follows from Lemma A.2.

Finally, differentiating (3.9) and taking the operator norm on $\mathcal{Y}$, we obtain

$$\begin{aligned}&\|\mathrm{D}\Psi^h(U)\|_{\mathcal{E}(\mathcal{Y}),h} \leq \|\mathsf{S}(hA)\|_{\mathcal{E}(\mathcal{Y}),h}\\ &\qquad + hs\,\Theta_{\mathsf{S}}\,\|\mathsf{b}\|\,\|(\mathrm{id}-h\mathsf{a}A)^{-1}\|_{\mathcal{E}(\mathcal{Y}^s)}\,\|\mathrm{D}B(W)\|_{\mathcal{E}(\mathcal{Y}^s)}\,\|\mathrm{D}_U W(U,h)\|_{\mathcal{E}(\mathcal{Y},\mathcal{Y}^s)}\\ &\leq (1+\sigma\,h) + h\,s\,\Theta_{\mathsf{S}}\,\|\mathsf{b}\|\,\Lambda\,M_0'\,\|W\|_{\mathcal{C}_{\mathrm{b}}^{(\underline{1},0)}(\mathcal{B}_{R_*}^{\mathcal{Y}}(U^0)\times[0,h_*];\mathcal{Y}^s)}\\ &\equiv 1+\sigma_\Psi\,h\,,\end{aligned}$$

where we use (3.2), (3.13) and (3.14a), and refer to (3.21) for the bound on W. This proves (3.22). □

While this theorem gives an existence and regularity result for the numerical time-h map Ψ^h, it does not yield control over the maximum step size h_* when we want to define Ψ^h on a general open bounded domain. We address this issue in the following theorem which is the discrete time analogue of Theorem 2.6.

Theorem 3.15 (Existence and regularity of numerical method, uniform version)**.** *Let the semilinear evolution equation* (2.4) *satisfy conditions* (A0) *and* (B1) *and apply a Runge–Kutta method subject to condition* (RK2) *to it. Moreover, assume* (A1) *or that the Runge–Kutta matrix* a *is invertible. Choose* $\delta\in(0,\delta_*]$ *small enough such that* $\mathcal{D}_{K+1}^{-\delta}$ *is non-empty. Then there exists* $h_*>0$ *such that* (3.20) *and* (3.21) *and, under assumption* (A1), (3.22) *hold with bounds uniform for* $U^0\in\mathcal{D}_{K+1}^{-\delta}$ *with* $R=\delta$. *Moreover, the stage vector* $W(U,h)$ *satisfies*

$$W \in \bigcap_{\substack{j+k\leq N\\ \ell\leq k\leq K+1}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1}^{-\delta}\times[0,h_*];\mathcal{Y}_{k-\ell}^s) \tag{3.25a}$$

and, if a *is invertible, the numerical time-h map* $\Psi(U,h) = \Psi^h(U)$ *satisfies*

$$\Psi \in \bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K+1}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0,h_*]; \mathcal{Y}_{k-\ell}) \,. \tag{3.25b}$$

In particular, when $N > K+1$*,*

$$W^j \in \mathcal{C}_{\mathrm{b}}^{K+1}(\mathcal{D}_{K+1}^{-\delta} \times [0,h_*]; \mathcal{D}), \quad j = 1, \ldots, s, \tag{3.26a}$$

and, if addition a *is invertible,*

$$\Psi \in \mathcal{C}_{\mathrm{b}}^{K+1}(\mathcal{D}_{K+1}^{-\delta} \times [0,h_*]; \mathcal{D}) \,. \tag{3.26b}$$

The bounds on W*,* Ψ *and* h_* *depend only on the bounds afforded by* (B1)*,* (2.18)*,* (3.14)*,* (3.18)*, on the coefficients of the method, and on* δ*.*

Proof. Let $R = \delta$. We apply Theorem 3.14 for each $U^0 \in \mathcal{D}_{K+1}^{-\delta}$. Note that for $j = 0, \ldots, K$,

$$\begin{aligned} \|(\mathrm{id} - (\mathrm{id} - h\mathsf{a}A)^{-1}) \mathbb{1} U^0\|_{\mathcal{Y}_j^s} &\leq h \max_{s \in [0,h]} \|\mathsf{a}A(\mathrm{id} - s\mathsf{a}A)^{-2} \mathbb{1} U^0\|_{\mathcal{Y}_j^s} \\ &\leq h \, \Lambda^2 \, \|\mathsf{a}\| \, R_{K+1} \,. \end{aligned}$$

Inserting this estimate into (3.23), we see that we can choose $h_* > 0$ small enough such that $\Pi(\,\cdot\,; U, h)$ maps $\mathcal{W}_j(U^0) \equiv \mathcal{B}_R(\mathbb{1}U^0) \subset \mathcal{Y}_j^s$ into itself for $j = 0, \ldots, K$, and, from (3.24), such that Π is a contraction on $\mathcal{W}_j(U^0)$ uniformly for $U^0 \in \mathcal{D}_{K+1}^{-\delta}$, $U \in \mathcal{B}_{R_*}^{\mathcal{Y}_{K+1}}(U^0)$, and $h \in [0,h_*]$, where $R_* = R/(2\max\{c_{\mathsf{S}}, \Lambda\})$. As in the proof of Theorem 3.14, we find that (3.20), (3.21) and (3.22) hold with uniform bounds in $U^0 \in \mathcal{D}_{K+1}^{-\delta}$, and that

$$W^i, \Psi \in \bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0,h_*]; \mathcal{D}_{k-\ell}) \tag{3.27}$$

for $i = 1, \ldots, s$.

To prove that W actually maps into a space one step up the scale, we show that

$$AW^i \in \bigcap_{\substack{j+k \leq N \\ \ell \leq k \leq K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0,h_*]; \mathcal{Y}_{k-\ell}) \tag{3.28}$$

for $i = 1, \ldots, s$. We apply A to (3.7), so that

$$AW = A(\mathrm{id} - h\mathsf{a}A)^{-1} \mathbb{1} U + h\mathsf{a}A(\mathrm{id} - h\mathsf{a}A)^{-1} B(W) \,. \tag{3.29}$$

The first term of (3.29) is of class (3.28) by Lemma 3.11. For the second term, we note that, by (B1) and (3.15b),

$$\Sigma(W,U,h) = h\mathsf{a}A(\mathrm{id} - h\mathsf{a}A)^{-1} \, B(W)$$

is of class

$$\Sigma \in \bigcap_{\substack{i+j+k \leq N \\ \ell \leq k \leq K}} \mathcal{C}_{\mathrm{b}}^{(\underline{i},\underline{j},\ell)}((\mathcal{D}_k)^s \times \mathcal{Y}_{K+1} \times \mathcal{I}; \mathcal{Y}_{k-\ell}^s) \,,$$

so that $(\Sigma \circ W)^i$ is of class (3.28) for $i = 1, \ldots, s$ by Lemma A.6. This proves (3.28).

To prove that, for a invertible, $A\Psi$ is also of class (3.28), we proceed analogously. Applying A to (3.9), we obtain

$$A\Psi^h(U) = \mathsf{S}(hA)AU + h\mathsf{b}^T A(\mathrm{id} - h\mathsf{a}A)^{-1} B(W(U,h)) \,. \tag{3.30}$$

The first term on the right of (3.30) is of class (3.28) by Lemma 3.11. We already proved above that $V = \Sigma \circ W$ is of class (3.28). As a is invertible, $\mathsf{b}^T \mathsf{a}^{-1} V$ and, hence, (3.30) are of class (3.28).

Next, we show improved regularity of W and Ψ with respect to the step size in the same way as in the proof of Theorem 2.6. Namely, we prove that

$$\partial_h W^i, \partial_h \Psi \in \bigcap_{\substack{j+k \le N-1 \\ \ell \le k \le K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1} \times [0, h_*]; \mathcal{Y}_{k-\ell}) \, .$$

Consider the $(K+1)$-scale of Banach spaces $\mathcal{Z}_j = \mathcal{Y}_j^s$ for $j = 0, \dots, K$ and $\mathcal{Z}_{K+1} = \mathcal{Y}_K^s$ with $\mathcal{W}_j = \mathcal{D}_j^s$ for $j = 0, \dots, K$ and $\mathcal{W}_{K+1} = \mathcal{D}_K^s$. Set $\mathcal{U} = \mathcal{D}_{K+1}^{-\delta}$, $\mathcal{X} = \mathcal{Y}_{K+1}$, and $\mathcal{I} = (0, h_*)$. Due to (3.15c) and (B1), the map Π from (3.7) satisfies the assumptions of Theorem A.9 in this setting. This shows that $\partial_h W$ is of the above class, and proves, with (3.27), (3.28) and Lemma A.3 claim (3.25a) for the stage vector W. Then (3.15c), Lemma 3.13, (3.25a), Lemma A.6 and Lemma A.7 applied to

$$\begin{aligned} \partial_h \Psi^h(U) &= \partial_h \mathsf{S}(hA) U + \mathsf{b}^T (\mathrm{id} - h\mathsf{a}A)^{-2} \, B(W(U,h)) \\ &\quad + h\mathsf{b}^T (\mathrm{id} - h\mathsf{a}A)^{-1} \, \mathrm{D}B(W(U,h)) \, \partial_h W(U,h) \, , \end{aligned}$$

imply that $\partial_h \Psi$ is of the same class as $\partial_h W^i$. When a is invertible, then, using that $A\Psi$ is of class (3.28) and using Lemma A.3 as before, claim (3.25b) follows.

Statements (3.26a) and (3.26b) are, as before, a consequence of Lemma A.2. $\square$

Remark 3.16. We actually showed in Theorem 3.15 that for a invertible

$$W^i, \Psi \in \bigcap_{\substack{j+k \le N \\ \ell \le k \le K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0, h_*]; \mathcal{Y}_{k-\ell+1}),$$

$i = 1, \dots, s$, i.e. W^i and Ψ have slightly higher regularity in U^0 than the semiflow Φ^t.

Remark 3.17 (Image of the numerical method and stage vector). Analogous to the situation for the semiflow noted in Remark 2.8, the proof of Theorem 3.15 actually shows that, when a is invertible,

$$\Psi, W^j \in \bigcap_{\substack{j+k \le N \\ \ell \le k \le K+1 \\ (k,\ell) \ne (K+1,0)}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}(\mathcal{D}_{K+1}^{-\delta} \times [0, T_*]; \mathcal{D}_{k-\ell})$$

for $j = 1, \dots, s$.

Remark 3.18. The proof of Theorem 3.15 shows that, when (A0), (A1), (B1) and (RK2) hold, but a is not assumed invertible, we still have

$$\Psi \in \bigcap_{\substack{j+k \le N-1 \\ k \le K}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},k+1)}(\mathcal{D}_{K+1}^{-\delta} \times [0, T_*]; \mathcal{D}).$$

Remark 3.19. If A generates a group rather than a semigroup, we may assume (A1) for $h \in [-h_*, h_*]$ for some $h_* > 0$. Then Theorems 3.14 and 3.15 hold with $h \in [-h_*, h_*]$ (for some, possibly, smaller choice of $h_* > 0$). Moreover, we can then also weaken the requirement in (RK1) to $|\mathsf{S}(z)| \le 1$ for $z \in \mathrm{i}\mathbb{R}$ and still show that (A2) implies (A1) for $h \in [-h_*, h_*]$. In this setting, the proof of Lemma 3.10

proceeds by recalling that, according to Remark 2.9, there exists $\omega > 0$ such that $|\mathrm{Re}(\operatorname{spec} A)| \leq \omega$. Hence, we can decompose $\mathbb{Q}A$ into a skew-symmetric operator $A_1 = \mathrm{Im}(A_{\mathrm{n}})$ and a bounded operator $A_2 = A_{\mathrm{b}} + \mathrm{Re}(A_{\mathrm{n}})$.

3.3. **Convergence analysis of A-stable Runge-Kutta methods.** In this section we present a convergence analysis of A-stable Runge–Kutta methods applied to semilinear evolution equation (2.4). The main difficulty is to prove differentiability in the step size h of the implicitly defined Runge–Kutta methods as maps from a space of functions with higher regularity to a space with lower regularity.

Theorem 3.20 (Convergence)**.** *Apply a Runge–Kutta method of classical order p subject to conditions* (RK2) *and* (A1) *to the semilinear evolution equation* (2.4)*. Assume further that* (B1) *holds with $K \geq p$. Pick $\delta \in (0, \delta_*]$ such that $\mathcal{D}^{-\delta}_{p+1}$ is non-empty and $T > 0$. Then there exist positive constants h_*, c_1, and c_2 which only depend on the bounds afforded by* (B1) *and* (A1), (3.14)*, on the coefficients of the method, and on δ, such that for every U^0 with*

$$\{\Phi^t(U^0) \colon t \in [0,T]\} \subset \mathcal{D}^{-\delta}_{p+1} \tag{3.31}$$

and for every $h \in [0, h_*]$, the numerical solution $(\Psi^h)^m(U^0)$ lies in $\mathcal{D}$ and satisfies

$$\|(\Psi^h)^m(U^0) - \Phi^{mh}(U^0)\|_{\mathcal{Y}} \leq c_2 \, \mathrm{e}^{c_1 mh} \, h^p$$

so long as $mh \leq T$.

Proof. We invoke Theorem 3.1 with $\mathcal{Z} = \mathcal{Y}_{p+1}$, $\mathcal{D}_{\mathcal{Z}} = \mathcal{D}^{-\delta}_{p+1}$, $\mathcal{X} = \mathcal{Y}$,

$$\mathcal{D}_{\mathcal{X}} = \bigcup_{U \in \mathcal{D}^{-\delta}_{p+1}} \mathcal{B}^{\mathcal{Y}}_R(U) \subset \mathcal{D}$$

where $R = \delta$, and note that $\operatorname{dist}_{\mathcal{X}}(\{U(t) \colon t \in [0,T]\}, \partial\mathcal{D}_{\mathcal{X}}) \geq \delta$. To verify the assumptions of the theorem, we first note that local existence and regularity of a solution to the evolution equation (2.4) in the appropriate spaces is always guaranteed by Theorem 2.6. In particular, for initial data U^0 such that (3.31) holds, we also have $U \in \mathcal{C}([0,T]; \mathcal{D}^{-\delta}_{p+1}) \cap \mathcal{C}^{p+1}([0,T]; \mathcal{D}^{-\delta})$ with uniform bounds in the norms of both spaces. Conditions (C1) and (C2) follow from Theorem 3.15 and Remark 3.18. □

Remark 3.21*.* As explained in Sections 2.5 and 2.6, the semilinear wave equation and the nonlinear Schrödinger equation satisfy the assumptions of Theorems 3.14–3.20 provided the nonlinearity is sufficiently smooth.

In the following corollary we prove the convergence of the U-derivatives of the numerical solution.

Corollary 3.22 (Convergence of derivatives)**.** *In the setting of Theorem 3.20 there exist positive constants h_*, c_1, and c_2 which only depend on the bounds afforded by* (B1) *and* (A1), (3.14)*, on the coefficients of the method, and on δ, such that for every U^0 satisfying* (3.31) *and for every $h \in [0, h_*]$*

$$\|\mathrm{D}^j_U (\Psi^h)^m(U^0) - \mathrm{D}^j_U \Phi^{mh}(U^0)\|_{\mathcal{E}^j(\mathcal{Y}_{p+1}, \mathcal{Y})} \leq c_2 \, \mathrm{e}^{c_1 mh} \, h^p$$

for $j \leq N - p - 1$ so long as $mh \leq T$.

Proof. We proceed by induction over j. The case $j=0$ is already asserted by Theorem 3.20. When $j>0$, we note that $\tilde{U}(t) \equiv (U(t), W(t)) \equiv (\Phi^t(U^0), \mathrm{D}\Phi^t(U^0)W^0)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{U}(t) = \tilde{A}\tilde{U} + \tilde{B}(\tilde{U}) \tag{3.32}$$

where

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}, \qquad \tilde{B}(\tilde{U}) = \begin{pmatrix} B(U) \\ \mathrm{D}B(U)W \end{pmatrix},$$

and we take $W^0 \in \mathcal{B}_k \equiv \operatorname{int}(\mathcal{B}_1^{\mathcal{Y}_k}(0))$. Similarly, the Runge–Kutta method applied to (3.32) satisfies

$$\tilde{\Psi}^h(\tilde{U}^0) = \begin{pmatrix} \Psi^h(U^0) \\ \mathrm{D}\Psi^h(U^0)W^0 \end{pmatrix} \quad \text{where} \quad \tilde{U}^0 = \begin{pmatrix} U^0 \\ W^0 \end{pmatrix}.$$

Eq. (3.32) and the Runge–Kutta method applied to it again satisfy (A1), (B1) with N replaced by $N-1$ and $\mathcal{D}_k$ replaced by $\mathcal{D}_k \times \mathcal{B}_k$ for $k = 0, \dots, K$ and (RK2). We can therefore apply the induction hypothesis to the extended system so long as $j + p \le N - 1$. $\square$

Appendix A. Contraction mappings on a scale of Banach spaces

In the appendix we present a contraction mapping theorem on a scale of Banach spaces, our main technical tool. Our results are more general than precursor versions in [31, 33]. The proofs are technically involved for two reasons. First, there is some combinatorial complexity in the estimates due to the implicitness of the fixed point of the contraction map. For this reason we decided to derive estimates in all required norms at once. Second, the maps we consider have derivatives with respect to the parameters that are only strongly continuous, but not continuous in the operator norm. This precludes a straightforward induction argument. What we find is that this weaker notion of continuity is entirely sufficient, but requires some extra care and notational effort.

For $K \in \mathbb{N}_0$, let $\mathcal{Z} = \mathcal{Z}_0 \supset \mathcal{Z}_1 \supset \ldots \supset \mathcal{Z}_K$ be a scale of Banach spaces, each continuously embedded in its predecessor, and let $\mathcal{V}_j, \mathcal{W}_j \subset \mathcal{Z}_j$ be nested sequences of sets. Let $\mathcal{X}$ be a Banach space, and let $\mathcal{U} \subset \mathcal{X}$ and $\mathcal{I} \subset \mathbb{R}$ be open. We note that all results in this section easily extend to the case where $\mathcal{I}$ is an open subset of $\mathbb{R}^p$. Without loss of generality, we may assume that $\|w\|_{\mathcal{Z}_j} \le \|w\|_{\mathcal{Z}_{j+1}}$ for all $w \in \mathcal{Z}_{j+1}$. (If this is not the case, we inductively equip $\mathcal{Z}_{j+1}$ with the equivalent norm $\|\cdot\|_{\mathcal{Z}_{j+1}} + \|\cdot\|_{\mathcal{Z}_j}$.)

We use the following additional integer indices. The minimal regularity we guarantee for the image space of the function considered is the regularity of the *lowest scale index* L of the image, the *loss index* S indicates how many rungs on the scale the range of a function is down relative to its domain, and N denotes the maximal regularity of the function. We assume $0 \le L \le K - S \le N - S$. Taking the dependence on parameters into account, we work with the family of spaces

$$\mathcal{C}_{N,K,L,S}(\{\mathcal{V}_j\}, \mathcal{U}, \mathcal{I}; \{\mathcal{W}_j\}) = \bigcap_{\substack{i+j+k \le N-S \\ L+\ell \le k \le K-S}} \mathcal{C}_{\mathrm{b}}^{(\underline{i}, \underline{j}, \ell)}(\mathcal{V}_{k+S} \times \mathcal{U} \times \mathcal{I}; \mathcal{W}_{k-\ell}),$$

endowed with norm

$$\|\Pi\|_{N,K,L,S} = \max_{\substack{i+j+k \le N-S \\ L+\ell \le k \le K-S}} \|\mathrm{D}_w^i \mathrm{D}_u^j \partial_\mu^\ell \Pi\|_{\mathcal{L}_\infty(\mathcal{V}_{k+S} \times \mathcal{U} \times \mathcal{I}; \mathcal{E}^i(\mathcal{Z}_{k+S}, \mathcal{E}^j(\mathcal{X}; \mathcal{Z}_{k-\ell})))}$$

for $0 \le L \le K - S \le N - S$, and abbreviate

$$\mathcal{C}_{N,K,L}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) = \mathcal{C}_{N,K,L,0}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) \,,$$
$$\mathcal{C}_{N,K}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) = \mathcal{C}_{N,K,0,0}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$$

with corresponding norms

$$|\!|\!|\Pi|\!|\!|_{N,K,L} = |\!|\!|\Pi|\!|\!|_{N,K,L,0} \,,$$
$$|\!|\!|\Pi|\!|\!|_{N,K} = |\!|\!|\Pi|\!|\!|_{N,K,0,0} \,.$$

Note that any function of class $\mathcal{C}_{N,K,L,S}$ has a maximal number of $N - L - S$ derivatives in its first and second argument on the lowest admissible domain scale $\mathcal{Z}_{L+S}$.

Furthermore, let

$$\mathcal{C}_{N,K,L}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) = \bigcap_{\substack{j+k\le N\\ L+\ell\le k\le K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(\mathcal{U}\times\mathcal{I};\mathcal{W}_{k-\ell}) \,,$$

endowed with norm

$$\|w\|_{N,K,L} = \max_{\substack{j+k\le N\\ L+\ell\le k\le K}} \|\mathrm{D}_u^j \partial_\mu^\ell w\|_{\mathcal{L}_\infty(\mathcal{U}\times\mathcal{I};\mathcal{E}^j(\mathcal{X};\mathcal{Z}_{k-\ell}))} \tag{A.1}$$

for $0 \le L \le K \le N$, where we abbreviate

$$\mathcal{C}_{N,K}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) = \mathcal{C}_{N,K,0}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$$

with corresponding norm

$$\|w\|_{N,K} = \|w\|_{N,K,0} \,.$$

For future reference, we note the following.

Remark A.1. When a map $\Pi \in \mathcal{C}_{N,K,L,S}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$ does not depend on w, it can be interpreted as an element from $\mathcal{C}_{N,K,L}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$ where

$$|\!|\!|\Pi|\!|\!|_{N,K,L,S} = \|\Pi\|_{N-S,K-S,L} \,.$$

We simply write $\mathcal{C}_{N,K,L,S}$ and $\mathcal{C}_{N,K,L}$ when the arguments are unambiguous. We also write

$$\partial_\mu\Pi(w(u,\mu);u,\mu) = \partial_\mu\Pi(w;u,\mu)\big|_{w=w(u,\mu)} = (\partial_\mu\Pi\circ w)(u,\mu)$$

to denote partial μ-derivatives vs. $\mathrm{D}_\mu(\Pi(w(u,\mu),u,\mu))$ to denote full μ-derivatives.

We begin with four short technical lemmas. The first specifies the relation between the spaces $\mathcal{C}_{N,K}$ and $\mathcal{C}^K$.

Lemma A.2. *If $N > K$ then, with $\mathcal{W} \equiv \mathcal{W}_0$,*

$$\mathcal{C}_{N,K}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) \subset \mathcal{C}_{\mathrm{b}}^K(\mathcal{U}\times\mathcal{I};\mathcal{W}) \,.$$

Proof. Let $w \in \mathcal{C}_{N,K}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$. Fixing $\ell = k$ in the definition of $\mathcal{C}_{N,K}$ and recalling (2.3b), i.e., strong and uniform continuity coincide if no derivative in u is taken, we find that

$$w \in \mathcal{C}_{\mathrm{b}}^{(0,K)}(\mathcal{U}\times\mathcal{I};\mathcal{W}) \cap \bigcap_{\substack{j+\ell\le K+1\\ \ell\le K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(\mathcal{U}\times\mathcal{I};\mathcal{W}) \,.$$

The claimed uniform continuity then holds because of (2.3a). □

The following lemma captures the essence of the inductive step in N as needed in the main results which follow.

Lemma A.3. *If $w \in \mathcal{C}_{N,K,L}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$ and the map $(u,\tilde{u},\mu) \mapsto \mathrm{D}_u w(u,\mu)\tilde{u}$ is of class $\mathcal{C}_{N,K,L}(\mathcal{U}\times\mathcal{B}_1^{\mathcal{X}}(0),\mathcal{I};\{\mathcal{Z}_j\})$, then $w \in \mathcal{C}_{N+1,K,L}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$ and*

$$\|w\|_{N+1,K,L} \leq \sup_{\|\tilde{u}\|_{\mathcal{X}}\leq 1} \|\mathrm{D}_u w\,\tilde{u}\|_{N,K,L} + \|w\|_{N,K,L}\,.$$

Proof. The claim is a direct consequence of the partitioning of the index set in the definition of the $(N+1,K,L)$-norm, see (A.1), into

$$\{0 \leq j+k \leq N+1\} = \{0 \leq j+k \leq N\} \cup \{0 \leq \tilde{j}+k \leq N\}$$

where $\tilde{j} = j-1$, using the definition of the operator norm,

$$\|T\|_{\mathcal{E}(\mathcal{X},\mathcal{Y})} = \sup_{\|x\|_{\mathcal{X}}=1} \|Tx\|_{\mathcal{Y}}\,,$$

and the definition of the (N,K,L)-norm (A.1). □

The next lemma captures the essence of the inductive step in K. Namely, a scale of length $K+1$ can be broken up into two scales which have only length K, plus a trivial remaining bit.

Lemma A.4. *When $N > K$, $w \in \mathcal{C}_{N,K+1,L+1}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) \cap \mathcal{C}_{N,L,L}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$, and $\partial_\mu w \in \mathcal{C}_{N-1,K,L}(\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})$, then $w \in \mathcal{C}_{N,K+1,L}(\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})$ and*

$$\|w\|_{N,K+1,L} \leq \|w\|_{N,K+1,L+1} + \|w\|_{N,L,L} + \|\partial_\mu w\|_{N-1,K,L}\,.$$

Proof. Translating the scale, i.e., setting $\tilde{\mathcal{Z}}_j = \mathcal{Z}_{j+L}$, $\tilde{K} = K-L$, and $\tilde{N} = N-L$, we can reduce to the case $L=0$. Since

$$\{0 \leq \ell \leq k \leq K+1\} = \{0 \leq \ell < k \leq K+1\} \cup \{1 \leq \ell \leq k \leq K+1\} \cup \{k=\ell=0\}$$

and $\partial_\mu w \in \mathcal{C}_{N-1,K}$ if and only if

$$w \in \bigcap_{\substack{j+k\leq N-1\\ \ell\leq k\leq K}} \mathcal{C}_{\mathrm{b}}^{(j,\ell+1)}(\mathcal{U}\times\mathcal{I};\mathcal{W}_{k-\ell}) = \bigcap_{\substack{j+k\leq N\\ 1\leq\ell\leq k\leq K+1}} \mathcal{C}_{\mathrm{b}}^{(j,\ell)}(\mathcal{U}\times\mathcal{I};\mathcal{W}_{k-\ell})\,,$$

the claim follows directly from definition of $\mathcal{C}_{N,K,L}$ and its norm (A.1). □

Finally, we prove that the space $\mathcal{C}_{N,K,0,S}$ can be expressed in terms of $\mathcal{C}_{N,K}$-type spaces with domains defined on a scale.

Lemma A.5. *We have*

$$\bigcap_{S\leq\kappa\leq K} \mathcal{C}_{N-S,\kappa-S,L}(\mathcal{V}_\kappa\times\mathcal{U},\mathcal{I};\{\mathcal{W}_j\}) = \mathcal{C}_{N,K,L,S}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})\,,$$

and

$$\|\Pi\|_{\mathcal{C}_{N,K,L,S}(\{\mathcal{V}_j\},\mathcal{U};\mathcal{I};\{\mathcal{W}_j\})} \sim \max_{S\leq\kappa\leq K} \|\Pi\|_{\mathcal{C}_{N-S,\kappa-S,L}(\mathcal{V}_\kappa\times\mathcal{U},\mathcal{I};\{\mathcal{W}_j\})}\,,$$

where $\sim$ denotes that left hand and right hand sides provide equivalent norms on $\mathcal{C}_{N,K,L,S}$.

Proof. Translating the scale, i.e., setting $\tilde{\mathcal{Z}}_j = \mathcal{Z}_{j+L}$, $\tilde{K} = K - L$, and $\tilde{N} = N - L$, we can reduce to the case $L = 0$. Next, we identify

$$\begin{aligned}
\bigcap_{S\le\kappa\le K} \mathcal{C}_{N-S,\kappa-S}(\mathcal{V}_\kappa \times \mathcal{U}; \mathcal{I}, \{\mathcal{W}_j\}) &= \bigcap_{\substack{S\le\kappa\le K\\ j+k\le N-S\\ \ell\le k\le\kappa-S}} \mathcal{C}_{\mathrm{b}}^{(\underline{j},\ell)}((\mathcal{V}_\kappa \times \mathcal{U}) \times \mathcal{I}; \mathcal{W}_{k-\ell}) \\
&= \bigcap_{\substack{S\le\kappa\le K\\ i+j+k\le N-S\\ \ell\le k\le\kappa-S}} \mathcal{C}_{\mathrm{b}}^{(\underline{i},\underline{j},\ell)}(\mathcal{V}_\kappa \times \mathcal{U} \times \mathcal{I}; \mathcal{W}_{k-\ell}) \\
&= \bigcap_{\substack{0\le\tilde{k}\le K-S\\ i+j+k\le N-S\\ \ell\le k\le\tilde{k}}} \mathcal{C}_{\mathrm{b}}^{(\underline{i},\underline{j},\ell)}(\mathcal{V}_{\tilde{k}+S} \times \mathcal{U} \times \mathcal{I}; \mathcal{W}_{k-\ell}) \\
&= \bigcap_{\substack{i+j+k\le N-S\\ \ell\le k\le K-S}} \mathcal{C}_{\mathrm{b}}^{(\underline{i},\underline{j},\ell)}(\mathcal{V}_{k+S} \times \mathcal{U} \times \mathcal{I}; \mathcal{W}_{k-\ell}) \,,
\end{aligned}$$

which equals $\mathcal{C}_{N,K,0,S}$. Noting that

$$\begin{aligned}
&\max_{\substack{S\le\kappa\le K\\ j+k\le N-S\\ \ell\le k\le\kappa-S}} \|\mathrm{D}^j_{(w,u)} \partial^\ell_\mu \Pi\|_{\mathcal{L}_\infty(\mathcal{V}_\kappa\times\mathcal{U}\times\mathcal{I};\mathcal{E}^j(\mathcal{Z}_\kappa\times\mathcal{X};\mathcal{Z}_{k-\ell}))} \\
&\qquad \sim \max_{\substack{S\le\kappa\le K\\ i+j+k\le N-S\\ \ell\le k\le\kappa-S}} \|\mathrm{D}^i_w \mathrm{D}^j_u \partial^\ell_\mu \Pi\|_{\mathcal{L}_\infty(\mathcal{V}_\kappa\times\mathcal{U}\times\mathcal{I};\mathcal{E}^i(\mathcal{Z}_\kappa;\mathcal{E}^j(\mathcal{X};\mathcal{Z}_{k-\ell})))} \,,
\end{aligned}$$

the statement about the norms follows analogously. □

The next lemma will be our main tool for obtaining estimates on the scale of Banach spaces for compositions of maps of the form

$$(\Pi \circ w)(u,\mu) \equiv \Pi(w(u,\mu); u, \mu) \,.$$

The essence of the result is very natural: When the outer function Π loses S rungs on the scale, the inner function w must have minimal regularity $L = S$ and the composition maps at best into rung $K - S$.

The main difficulty in the proof of this lemma and of the subsequent results is that the maps considered lose smoothness when derivatives in μ are taken. In particular, these derivatives are only strongly continuous with respect to the parameters u and μ and, in our infinite-dimensional setting, are discontinuous with respect to u and μ in the operator norm. As a result, in the proofs below the induction hypothesis cannot be applied to the derivatives in a straightforward way.

Lemma A.6 (Chain rule on a scale of Banach spaces)**.** *Let* $\Pi = \Pi(w; u, \mu)$ *and* $w = w(u,\mu)$ *satisfy*

$$\Pi \in \mathcal{C}_{N,K,L,S}(\{\mathcal{W}_j\}, \mathcal{U}, \mathcal{I}; \{\mathcal{Z}_j\}) \quad \text{and} \quad w \in \mathcal{C}_{N,K,S+L}(\mathcal{U}, \mathcal{I}; \{\mathcal{W}_j\}) \,.$$

Then $\Pi \circ w \in \mathcal{C}_{N-S,K-S,L}(\mathcal{U}, \mathcal{I}; \{\mathcal{Z}_j\})$ *and* $\|\Pi \circ w\|_{N-S,K-S,L}$ *can be bounded by a polynomial with non-negative coefficients in* $\|\!|\Pi|\!\|_{N,K,L,S}$ *and* $\|w\|_{N,K,S+L}$.

Proof. Translating the scale, i.e., setting $\tilde{\mathcal{Z}}_j = \mathcal{Z}_{j+L}$, $\tilde{K} = K - L$, and $\tilde{N} = N - L$, we can reduce to the case $L = 0$ as in the proof of Lemma A.4. We proceed by induction in N and K as follows.

For $N = K = S$, we have $\Pi \in \mathcal{C}_{\mathrm{b}}(\mathcal{W}_S \times \mathcal{U} \times \mathcal{I}; \mathcal{Z}_0)$, $w \in \mathcal{C}_{\mathrm{b}}(\mathcal{U} \times \mathcal{I}; \mathcal{W}_S)$, hence $\Pi \circ w \in \mathcal{C}_{\mathrm{b}}(\mathcal{U}, \mathcal{I}; \mathcal{Z}_0)$ with bound $\|\!|\Pi|\!\|_{S,S,0,S}$.

Let us now increment N holding K and S fixed. Let $\mathcal{B} \equiv \mathcal{B}_1^{\mathcal{X}}(0)$. We claim that the map

$$(u, \tilde{u}, \mu) \mapsto \mathrm{D}_u(\Pi \circ w)\,\tilde{u} \quad \text{is of class} \quad \mathcal{C}_{N-S,K-S}(\mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{Z}_j\}) \tag{A.2}$$

with a bound which is a polynomial in $\|\!|\Pi|\!\|_{N+1,K,0,S}$ and $\|w\|_{N+1,K,S}$. The inductive step is achieved by Lemma A.3 which then asserts that $\Pi \circ w \in \mathcal{C}_{N+1-S,K-S}$ with its norm bounded as required.

To prove this claim, let $\tilde{u} \in \mathcal{B}$, write

$$\mathrm{D}_u(\Pi \circ w)(u,\mu)\,\tilde{u} = \partial_u \Pi(w(u,\mu); u, \mu)\,\tilde{u} + \partial_w \Pi(w(u,\mu); u,\mu)\,\mathrm{D}_u w(u,\mu)\,\tilde{u}\,, \tag{A.3}$$

and consider each term on the right of (A.3) separately. For the first term on the right, set $\hat{u} \equiv (u, \tilde{u}) \in \mathcal{U} \times \mathcal{B} \equiv \hat{\mathcal{U}}$ and define

$$\Pi_1(w; \hat{u}, \mu) = \partial_u \Pi(w; u, \mu)\,\tilde{u}\,. \tag{A.4}$$

By assumption, this map is of class $\mathcal{C}_{N,K,0,S}(\{\mathcal{W}_j\}, \mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{Z}_j\})$. The induction hypothesis, applied to the maps Π_1 and w, then asserts that

$$\Pi_1 \circ w \in \mathcal{C}_{N-S,K-S}(\mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{Z}_j\}) \tag{A.5}$$

and that its $\mathcal{C}_{N-S,K-S}$-norm is bounded by a polynomial with non-negative coefficients in $\|\!|\Pi|\!\|_{N+1,K,0,S} \geq \|\!|\Pi_1|\!\|_{N,K,0,S}$ and $\|w\|_{N,K,S}$.

For the second term on the right of (A.3), we must proceed in stages. Fix $r = \|w\|_{N+1,K,S}$ and let $\mathcal{V}_\kappa = \mathcal{B}_r^{\mathcal{Z}_\kappa}(0)$. For $\kappa = S, \ldots, K$ and $(u, \hat{w}) \in \mathcal{U} \times \mathcal{V}_\kappa$, we set

$$\Pi_2(w; (u, \hat{w}), \mu) = \mathrm{D}_w \Pi(w, u, \mu)\,\hat{w}\,.$$

By assumption, this map is of class $\mathcal{C}_{N,\kappa,0,S}(\{\mathcal{W}_j\}, \mathcal{U} \times \mathcal{V}_\kappa, \mathcal{I}; \{\mathcal{Z}_j\})$. The induction hypothesis, applied to the maps Π_2 and w, then asserts that

$$\Pi_2 \circ w \in \mathcal{C}_{N-S,\kappa-S}(\mathcal{U} \times \mathcal{V}_\kappa, \mathcal{I}; \{\mathcal{Z}_j\}) \tag{A.6}$$

and that its $\mathcal{C}_{N-S,\kappa-S}$-norm is bounded by a polynomial in $\|w\|_{N,K,S} \geq \|w\|_{N,\kappa,S}$ and

$$\|\!|\Pi|\!\|_{N+1,K,0,S} \sup_{\hat{w} \in \mathcal{V}_\kappa} \|\hat{w}\|_{\mathcal{Z}_\kappa} \geq \|\!|\Pi_2|\!\|_{\mathcal{C}_{N,\kappa,0,S}(\{\mathcal{W}_j\}, \mathcal{U} \times \mathcal{V}_\kappa, \mathcal{I}; \{\mathcal{Z}_j\})}\,.$$

We now consider the composition $\Pi_2 \circ w$ as a map

$$\hat{\Pi}(\hat{w}; u, \mu) = \partial_w \Pi(w(u,\mu); u, \mu)\,\hat{w}\,.$$

Recalling that (A.6) applies for all $\kappa = S, \ldots, K$, we can apply Lemma A.5 to obtain that

$$\hat{\Pi} \in \mathcal{C}_{N,K,0,S}(\{\mathcal{V}_j\}, \mathcal{U}, \mathcal{I}; \{\mathcal{Z}_j\})$$

and that its norm is bounded by a polynomial in $r\,\|\!|\Pi|\!\|_{N+1,K,0,S}$ and $\|w\|_{N,K,S}$. (This is summarized in Lemma A.7 for later use.)

Now consider $\hat{\Pi}$ as a function of $\hat{w}$, $\hat{u} = (u, \tilde{u}) \in \mathcal{U} \times \mathcal{B}$, and μ. Since

$$\|\mathrm{D}_u w(u,\mu)\,\tilde{u}\|_{\mathcal{C}_{\mathrm{b}}(\mathcal{U} \times \mathcal{I}; \mathcal{Z}_j)} \leq \|w\|_{N+1,K,S}\,\|\tilde{u}\|_{\mathcal{X}} \leq r$$

for $j = S, \ldots, K$, the function $\hat{w}(\hat{u}, \mu) = \mathrm{D}_u w(u,\mu)\,\tilde{u}$ is of class $\mathcal{C}_{N,K,S}(\mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{V}_j\})$. Applying the induction hypothesis to $\hat{\Pi}$ and $\hat{w}$, we conclude that $\hat{\Pi} \circ \hat{w}$ or, written explicitly, the map

$$((u, \tilde{u}), \mu) \mapsto \partial_w \Pi(w(u,\mu); u, \mu)\,\mathrm{D}_u w(u,\mu)\,\tilde{u}$$

is of class $\mathcal{C}_{N-S,K-S}(\mathcal{U}\times\mathcal{B},\mathcal{I};\{\mathcal{Z}_j\})$, with norm bounded by an increasing polynomial in $\|\Pi\|_{N+1,K,0,S}$ and $\|w\|_{N+1,K,S}\geq\|\hat{w}\|_{N,K,S}$. Due to (A.3), (A.4), and (A.5), this also holds for the map $((u,\tilde{u}),\mu)\mapsto \mathrm{D}_u\Pi(w(u,\mu);u,\mu)\,\tilde{u}$, thus proves our claim (A.2); the inductive step in N is complete.

Next, we increment $K-S$ keeping N fixed. Here the inductive step will be achieved by Lemma A.4; we must hence verify its assumptions. First, applying the induction hypothesis on the scale $\tilde{\mathcal{Z}}_j=\mathcal{Z}_{j+1}$ with $j=0,\dots,K$, we infer that

$$\Pi\circ w\in\mathcal{C}_{N-1-S,K-S}(\mathcal{U},\mathcal{I};\{\tilde{\mathcal{Z}}_j\})=\mathcal{C}_{N-S,K+1-S,1}(\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})$$

with the corresponding norm bounded by a polynomial with non-negative coefficients in $\|\Pi\|_{N,K+1,0,S}\geq\|\Pi\|_{N,K+1,1,S+1}$ and $\|w\|_{N,K+1,S+1}$. Second, by the induction hypothesis applied on the trivial scale,

$$\Pi\circ w\in\mathcal{C}_{N-S,0}(\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})\,,$$

with the corresponding norm bounded by a polynomial with non-negative coefficients in $\|\Pi\|_{N,K+1,0,S}\geq\|\Pi\|_{N,0,0,S,0}$ and $\|w\|_{N,K+1,S}\geq\|w\|_{N,0,S}$. Third, we claim that

$$\mathrm{D}_\mu(\Pi\circ w)\in\mathcal{C}_{N-1-S,K-S}(\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})\,,\tag{A.7}$$

with the corresponding norm bounded by a polynomial with non-negative coefficients in $\|\Pi\|_{N,K+1,0,S}$ and $\|w\|_{N,K+1,S}$. Then Lemma A.4 applied to $\Pi\circ w$ where N and K there correspond to $N-S$ and $K-S$ here proves that $\Pi\circ w\in\mathcal{C}_{N-S,K+1-S}$ with the required bound on its norm; this concludes the inductive step.

It remains to prove claim (A.7). Following the steps in the estimate of $\mathrm{D}_u(\Pi\circ w)$ above, we write

$$\mathrm{D}_\mu(\Pi\circ w)(u,\mu)=\partial_\mu\Pi(w(u,\mu);u,\mu)+\partial_w\Pi(w(u,\mu);u,\mu)\,\mathrm{D}_\mu w\tag{A.8}$$

and consider each term on the right of (A.8) separately. For the first term, note that the assumption on Π implies, in particular, that $\partial_\mu\Pi\in\mathcal{C}_{N,K+1,0,S+1}$ and that, by assumption, $w\in\mathcal{C}_{N,K+1,S+1}$. Since $K-S$ is not increased, the induction hypothesis applies to this pair of maps and yields

$$\partial_\mu\Pi\circ w\in\mathcal{C}_{N-S-1,K-S}(\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})\tag{A.9}$$

with a polynomial bound in $\|\Pi\|_{N,K+1,0,S}\geq\|\partial_\mu\Pi\|_{N,K+1,0,S+1}$ and $\|w\|_{N,K+1,S}\geq\|w\|_{N,K+1,S+1}$.

For the second term on the right of (A.8), fix $r=\|w\|_{N,K+1,S}$ and let $\mathcal{V}_j=\mathcal{B}_r^{\mathcal{Z}_j}(0)$ for $j=S,\dots,K+1$. We saw above that the map $\hat{\Pi}$ from (A.11) is of class $\mathcal{C}_{N-1,K,0,S}(\{\mathcal{V}_j\},\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})$ with norm bounded as specified in Lemma A.7. The assumption on w and the definition of r above imply, moreover, that

$$\mathrm{D}_\mu w\in\mathcal{C}_{N-1,K,S}(\mathcal{U},\mathcal{I};\{\mathcal{V}_j\})\,.$$

Thus, the induction hypothesis applied once more to this pair of maps yields

$$\hat{\Pi}\circ\partial_\mu w=\mathrm{D}_w(\Pi\circ w)\mathrm{D}_\mu w\in\mathcal{C}_{N-1-S,K-S}(\mathcal{U},\mathcal{I};\{\mathcal{Z}_j\})\tag{A.10}$$

with a polynomial bound in $\|\Pi\|_{N,K+1,0,S}$ and $\|w\|_{N,K+1,S}\geq\|\partial_\mu w\|_{N,K,S}$. Together, (A.9) and (A.10) imply (A.7) with the required bound. □

In the proof of Lemma A.6, we implicitly proved the following result which we state here for later reference.

Lemma A.7. *Let* Π *and* w *satisfy the conditions of Lemma A.6 with* $L = 0$*; let* $r > 0$ *and* $\mathcal{V}_j = \mathcal{B}_r^{\mathcal{Z}_j}(0)$ *for* $j = 0, \dots, K$*. Then*

$$\hat{\Pi}(\hat{w}; u, \mu) \equiv \mathrm{D}_w \Pi(w(u, \mu); u, \mu)\, \hat{w} \tag{A.11}$$

satisfies

$$\hat{\Pi} \in \mathcal{C}_{N-1,K,0,S}(\{\mathcal{V}_j\}, \mathcal{U}, \mathcal{I}; \{\mathcal{Z}_j\})$$

with a polynomial bound in $\|w\|_{N-1,K,S}$ *and* $r\, \|\!|\Pi|\!\|_{N,K,0,S}$*.*

Remark A.8. The Faà di Bruno formula (see, e.g., [10]) can be used to compute the derivatives of compositions of functions explicitly. However, it does not remove the need to estimate complete $\mathcal{C}_{N,K}$ norms. Thus, an inductive argument seems to be the most manageable way of writing out a proof.

We now proceed to the crucial contraction mapping theorem for maps $\Pi(\,\cdot\,; u, \mu)$ of class $\mathcal{C}_{N,K}$.

Theorem A.9 (Contraction mappings on a scale of Banach spaces)**.** *For* $N, K \in \mathbb{N}_0$ *with* $N \geq K$*, let* $\mathcal{Z} = \mathcal{Z}_0 \supset \mathcal{Z}_1 \supset \ldots \supset \mathcal{Z}_K$ *be a scale of Banach spaces, each continuously embedded in its predecessor, let* $\mathcal{W}_j \subset \mathcal{Z}_j$ *be a nested sequence of closures of open sets, let* $\mathcal{X}$ *be a Banach space, and let* $\mathcal{U} \subset \mathcal{X}$ *and* $\mathcal{I} \subset \mathbb{R}$ *be open. Let* $(w, u, \mu) \mapsto \Pi(w; u, \mu)$ *be a nonlinear map such that*

(i) $\Pi \in \mathcal{C}_{N,K}(\{\mathcal{W}_j\}, \mathcal{U}, \mathcal{I}; \{\mathcal{W}_j\})$*;*
(ii) $w \mapsto \Pi(w; u, \mu)$ *is a contraction on* $\mathcal{W}_j$ *with contraction constant* $c'_j < 1$ *uniformly for all* $u \in \mathcal{U}$*,* $\mu \in \mathcal{I}$*, and* $j = 0, \dots, K$*.*

Then the fixed point equation $\Pi(w; u, \mu) = w$ *has a unique solution*

$$w \in \mathcal{C}_{N,K}(\mathcal{U}, \mathcal{I}; \{\mathcal{W}_j\})$$

and $\|w\|_{N,K}$ *is bounded by a function which is a polynomial with non-negative coefficients in* $\|\!|\Pi|\!\|_{N,K}$ *and* $(1 - c'_j)^{-1}$*.*

Similar theorems were proved in [31] for the case $K = 1$, $\mathcal{U} = \emptyset$ and in [33] for the case $N = K \in \mathbb{N}$, $\mathcal{U} = \emptyset$. Due to Lemma A.2, the theorem as stated here implies, in particular, that $w \in \mathcal{C}^K_{\mathrm{b}}(\mathcal{U} \times \mathcal{I}; \mathcal{W})$. This simple statement on $\mathcal{C}^K$ differentiability is reminiscent of the standard form of the contraction mapping theorem with parameters as, for example, stated in [19, p. 13].

Proof of Theorem A.9. The argument is once more an induction in N and K, following the combinatorial pattern of the proof of Lemma A.6. For $N = K = 0$, the regular contraction mapping theorem with parameters asserts that $w \in \mathcal{C}(\mathcal{U} \times \mathcal{I}; \mathcal{W})$. Moreover, Π is a contraction uniformly for $(u, \mu) \in \mathrm{cl}(\mathcal{U}) \times \mathrm{cl}(\mathcal{I})$ so that Π has a unique fixed point $w(u_*, \mu_*)$ also for (u_*, μ_*) on the boundary of $\mathcal{U} \times \mathcal{I}$. From this, a straightforward estimate yields continuity of w up to the boundary; thus, $w \in \mathcal{C}_{0,0}(\mathcal{U}, \mathcal{I}; \{\mathcal{W}_0\})$.

Assume now that the conclusion of the theorem holds for fixed K and $N \geq K$. We first employ Lemma A.3 to show that the conclusion also holds when we increment N, holding K fixed.

As in the proof of Lemma A.6, we set $\mathcal{B} \equiv \mathcal{B}_1^{\mathcal{X}}(0)$ and let $(u, \tilde{u}) \in \mathcal{U} \times \mathcal{B} \equiv \tilde{\mathcal{U}}$. Differentiating the fixed point equation $w = \Pi \circ w$ with respect to u, we find that

$\mathrm{D}_u w\, \tilde u$ formally solves the fixed point equation $\tilde w = \tilde\Pi(\tilde w; (u, \tilde u), \mu)$, where

$$\begin{aligned}\tilde\Pi(\tilde w; (u, \tilde u), \mu) &= \partial_w \Pi(w(u,\mu); u, \mu)\, \tilde w + \partial_u \Pi(w(u,\mu); u, \mu)\, \tilde u \\ &\equiv \hat\Pi(\tilde w; u, \mu) + \partial_u \Pi(w(u,\mu); u, \mu)\, \tilde u\,.\end{aligned}$$

Using the chain rule Lemma A.6 and Lemma A.7, we infer that

$$\tilde\Pi \in \mathcal{C}_{N,K}(\{\mathcal{V}_j\}, \mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{Z}_j\})$$

with $\mathcal{V}_j = \mathcal{B}_r^{\mathcal{Z}_j}(0)$ for $j = 0, \dots, K$ and arbitrary $r > 0$. Here, we must prove in addition that $\tilde\Pi$ maps each of the $\mathcal{V}_0, \dots, \mathcal{V}_K$ into itself. Indeed, a direct estimate shows that it suffices to take

$$r = \|\Pi\|_{N+1,K,S} \max_{j=0,\dots,K} \frac{1}{1 - c_j'} \geq \max_{j=0,\dots,K} \frac{\|\partial_u \Pi \circ w\|_{\mathcal{L}_\infty(\mathcal{U}\times\mathcal{I}; \mathcal{E}(\mathcal{X}, \mathcal{Z}_j))}}{1 - c_j'}\,.$$

The induction hypothesis then applies to $\tilde\Pi \in \mathcal{C}_{N,K}(\{\mathcal{V}_j\}, \mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{V}_j\})$, yielding the existence of a fixed point $\tilde w \in \mathcal{C}_{N,K}(\mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{V}_j\})$.

It remains to be shown that the formal identity $\tilde w = \mathrm{D}_u w\, \tilde u$ holds true on each $\mathcal{Z}_j$ for $j = 0, \dots, K$. This, however, follows by [33, Theorem 4.8] (see also the proof of [31, Theorem 3]) applied to the one-parameter family of maps $(w; \nu) \mapsto \Pi(w; u + \nu \tilde u, \mu)$ for fixed $\mu \in \mathcal{I}$, $u \in \mathcal{U}$, and $\tilde u \in \mathcal{B}$ on the scale $\{\tilde{\mathcal{Z}}_0, \tilde{\mathcal{Z}}_1\} = \{\mathcal{Z}_j, \mathcal{Z}_j\}$ for each $j = 0, \dots, K$.

Altogether, since $w \in \mathcal{C}_{N,K}(\mathcal{U}, \mathcal{I}; \{\mathcal{W}_j\})$, Lemma A.3 applies and yields gives $w \in \mathcal{C}_{N+1,K}(\mathcal{U}, \mathcal{I}; \{\mathcal{W}_j\})$; the inductive step in N is complete.

Next, we increment $K < N$ holding N fixed. For this, we use Lemma A.4. First, we note that assumptions (i) and (ii) hold on the K-step scale $\mathcal{Z}_1 \supset \cdots \supset \mathcal{Z}_{K+1}$ so that the induction hypothesis applies; we find that

$$w \in \mathcal{C}_{N,K+1,1}(\mathcal{U} \times \mathcal{I}; \{\mathcal{W}_j\})\,.$$

Second, by the induction hypothesis applied on the trivial scale, $w \in \mathcal{C}_{N,0}$. Third, differentiating the fixed point equation $w = \Pi \circ w$ with respect to μ, we obtain that $\partial_\mu w$ formally solves the fixed point equation $\tilde w = \tilde\Pi(\tilde w; u, \mu)$, where

$$\begin{aligned}\tilde\Pi(\tilde w; u, \mu) &= \partial_w \Pi(w(u,\mu); u, \mu)\, \tilde w + \partial_\mu \Pi(w(u,\mu); u, \mu) \\ &\equiv \hat\Pi(\tilde w; u, \mu) + \partial_\mu \Pi(w(u,\mu); u, \mu)\,.\end{aligned}$$

Since, by assumption, $w \in \mathcal{C}_{N,K}$, we infer from Lemma A.6 and Lemma A.7 that

$$\tilde\Pi \in \mathcal{C}_{N-1,K}(\{\mathcal{V}_j\}, \mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{Z}_j\})\,.$$

Here, we need in addition that $\tilde\Pi$ maps each $\mathcal{V}_0, \dots, \mathcal{V}_K$ into itself. This is satisfied whenever

$$r = \|\Pi\|_{N,K+1,S} \max_{j=0,\dots,K} \frac{1}{1 - c_j'} \geq \max_{j=0,\dots,K} \frac{\|\partial_\mu \Pi \circ w\|_{\mathcal{L}_\infty(\mathcal{U}\times\mathcal{I}; \mathcal{Z}_j)}}{1 - c_j'}\,.$$

The induction hypothesis then applies to $\tilde\Pi \in \mathcal{C}_{N-1,K}(\{\mathcal{V}_j\}, \mathcal{U}\times\mathcal{B}, \mathcal{I}; \{\mathcal{V}_j\})$, yielding the existence of a fixed point $\tilde w \in \mathcal{C}_{N-1,K}(\mathcal{U} \times \mathcal{B}, \mathcal{I}; \{\mathcal{V}_j\})$. By [33, Theorem 4.8] (see also the proof of [31, Theorem 3]), applied to $(w; \mu) \mapsto \Pi(w; u, \mu)$ for each fixed $u \in \mathcal{U}$ on the two-step scale $\{\tilde{\mathcal{Z}}_0, \tilde{\mathcal{Z}}_1\} \equiv \{\mathcal{Z}_j, \mathcal{Z}_{j+1}\}$ for each $j = 0, \dots, K$, we ensure that the formal identity $\tilde w = \partial_\mu w$ holds true across the scale $\mathcal{Z}_0, \dots \mathcal{Z}_K$. We conclude that $\partial_\mu w \in \mathcal{C}_{N-1,K}$.

Altogether, Lemma A.4 applies and yields $w \in \mathcal{C}_{N,K+1}$; the inductive step in K is now complete. We note that the required polynomial bounds are obtained, as before, by carefully tracking all the bounds in the respective norms through the argument. We omit all detail. □

Acknowledgments

C.W. thanks the Courant Institute of Mathematical Sciences for their hospitality during the preparation of parts of the manuscript, and acknowledges funding by the Nuffield Foundation, by the Leverhulme Foundation and by EPSRC grant EP/D063906/1. M.O. was visiting the Courant Institute of Mathematical Sciences supported by a Max–Kade Fellowship when part of this work was done, and further acknowledges support through the ESF network Harmonic and Complex Analysis and Applications (HCAA).

References

[1] Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. 2nd ed., Elsevier, Oxford (2003)
[2] Bridges, T., Reich, S.: Numerical Methods for Hamiltonian PDEs. J. Phys. A: Math. Gen. 39, 5287–5320 (2006)
[3] Akrivis, G., Crouzeix, M.: Linearly implicit methods for nonlinear parabolic equations. Math. Comp. 73, 613–635 (2004)
[4] Bakaev, N., Ostermann, A.: Long-term stability of variable stepsize approximations of semigroups. Math. Comp. 71, 1545–1567 (2002)
[5] Besse, C., Bidégaray, B., Descombes, S.: Order estimates in time of splitting methods for the nonlinear Schrödinger equation. SIAM J. Numer. Anal. 40, 26–40 (2002)
[6] Brenner, P., Thomée, V.: On rational approximations of semigroups. SIAM J. Numer. Anal. 16, 683–694 (1979)
[7] Brenner, P., Crouzeix, M., Thomée, V.: Single step methods for inhomogeneous linear differential equations in Banach space. RAIRO Modél. Math. Anal. Numér. 16, 5–26 (1982)
[8] Colin, T., Fabrie, P.: Semidiscretization in time for nonlinear Schrödinger-waves equations. Discret. Contin. Dyn. S. 4, 277–288 (1998)
[9] Colin, T., Ebrard, G., Gallice, G.: Semi-discretization in time for nonlinear Zakharov waves equations. Discret. Contin. Dyn. S. Ser. B 11, 263–282 (2009)
[10] Constantine, G.M., Savits, T.H.: A multivariate Faà di Bruno formula with applications. Trans. Am. Math. Soc. 348, 503–520 (1996)
[11] Crouzeix, M., Thomée, V.: On the discretization in time of semilinear parabolic equations with nonsmooth initial data. Math. Comp. 49, 359–377 (1987)
[12] Crouzeix, M., Larsson, S., Piskarev, S., Thomée, V.: The stability of rational approximations of analytic semigroups. BIT 33, 74–84 (1993)
[13] Deuflhard, P., Bornemann, F.: Scientific Computing with Ordinary Differential Equations. Springer, New York (2002)
[14] Faou, E., Grébert, B., Paturel, E.: Birkhoff normal form for splitting methods applied to semilinear Hamiltonian PDEs. Part II: Abstract splitting. Numer. Math. 114, 459–490 (2010)
[15] Gauckler, L., Lubich, C.: Splitting integrators for nonlinear Schrödinger equations over long times. Found. Comput. Math. 10, 275–302 (2010)
[16] González, C., Ostermann, A., Palencia, C., Thalhammer, M.: Backward Euler discretization of fully nonlinear parabolic problems. Math. Comp. 71, 125–145 (2002)
[17] González, C., Palencia, C.: Stability of Runge–Kutta methods for quasilinear parabolic problems. Math. Comp. 69, 609–628 (2000)
[18] Hansen, E.; Ostermann, A.: Exponential splitting for unbounded operators. Math. Comp. 78, 1485–1496 (2009)
[19] Henry, D.: Geometric Theory of Semilinear Parabolic Equations. Springer, New York (1983)
[20] Herbst, B.M., Mitchell, A.R., Weideman, J.A.C.: On the stability of the nonlinear Schrödinger equation. J. Comp. Phys. 60, 263–281 (1985)
[21] Hersh, R., Kato, T.: High-accuracy stable difference schemes for well-posed initial value problems. SIAM J. Numer. Anal. 16, 670–682 (1979)

[22] Kovacs, M.: On the convergence of rational approximations of semigroups on intermediate spaces. Math. Comp. 76, 273–286 (2007)
[23] Le Roux, M.N.: Semidiscretizations in time for parabolic problems. Math. Comp. 33, 919–931 (1979)
[24] Lubich, C., Ostermann, A.: Runge–Kutta methods for parabolic equations and convolution quadrature. Math. Comp. 60, 105–131 (1993)
[25] Lubich, C.: On splitting methods for Schrödinger–Poisson and cubic nonlinear Schrödinger equations. Math. Comp. 77, 2141–2153 (2008)
[26] Palencia, C.: A stability result for sectorial operators in Banach spaces. SIAM J. Numer. Anal. 30, 1373–1384 (1993)
[27] Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer, New York (1983)
[28] Rothe, E.: Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben. Math. Ann. 102, 650–670 (1930)
[29] Smith, G. D.: Numerical Solution of Partial Differential Equations: Finite Difference Methods. 3rd ed., Oxford University Press (1986)
[30] Tannehill, J. C., Anderson D.A., Pletche, R.H.: Computational fluid mechanics and heat transfer. Taylor and Francis, Philadelphia (1997)
[31] Vanderbauwhede, A., van Gils, S.A.: Center manifolds and contractions on a scale of Banach spaces. J. Funct. Anal. 72, 209–224 (1987)
[32] Weideman, J.A.C., Herbst, B.M.: Split-step methods for the solution of the nonlinear Schrödinger equation. SIAM J. Num. Anal. 23, 485–507 (1986)
[33] Wulff, C.: Transition from relative equilibria to relative periodic orbits. Doc. Math. 5, 227–274 (2000)

(M. Oliver) School of Engineering and Science, Jacobs University, 28759 Bremen, Germany

E-mail address: `oliver@member.ams.org`

(C. Wulff) Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK

E-mail address: `c.wulff@surrey.ac.uk`