*IMA Journal of Numerical Analysis* (2021) **00**, 1–36
https://doi.org/10.1093/imanum/drab060



# A high-order fully discrete scheme for the Korteweg–de Vries equation with a time-stepping procedure of Runge–Kutta-composition type

VASSILIOS A. DOUGALIS
*Mathematics Department, University of Athens, 15784 Zographou, Greece, and Institute of Applied and Computational Mathematics, Foundation for Research and Technology Hellas (IACM-FORTH), 71110 Heraklion, Greece*

AND

ÁNGEL DURÁN
*Applied Mathematics Department, University of Valladolid, and Institute of Mathematics of the University of Valladolid (IMUVA), Paseo de Belén S/N, 47011 Valladolid, Spain*
*Corresponding author: angeldm@uva.es

[Received on 8 September 2020; revised on 3 July 2021]

We consider the periodic initial-value problem for the Korteweg–de Vries equation that we discretize in space by a spectral Fourier–Galerkin method and in time by an implicit, high-order, Runge–Kutta scheme of composition type based on the implicit midpoint rule. We prove $L^2$ error estimates for the resulting semidiscrete and the fully discrete approximations. Some numerical experiments illustrate the results.

*Keywords*: Korteweg–de Vries equation; spectral method; Runge–Kutta composition methods; error estimates.

## 1. Introduction

In this paper we consider the periodic initial-value problem (IVP) for the Korteweg–de Vries (KdV) equation

$$\begin{aligned} &u_t + uu_x + u_{xxx} = 0, \quad x \in [-\pi, \pi],\ 0 \leqslant t \leqslant T, \\ &u(x,0) = u_0(x), \quad x \in [-\pi, \pi], \end{aligned} \tag{1.1}$$

where $u_0$ is a smooth, $2\pi$-periodic, real-valued function. The KdV is one of the simplest nonlinear partial differential equations (PDE) modelling one-way propagation in one space dimension of long waves in which the nonlinear term (here given by $uu_x$) and the linear dispersive term (modelled by $u_{xxx}$) are suitably balanced. It has been studied extensively and has a rich mathematical theory. In the case of (1.1) it is well known for example, cf. e.g. Bona *et al.* (1995), Bona & Smith (1975), Temam (1969), that if $u_0 \in H^\mu$ for $\mu \geqslant 2$, where $H^\mu$ is the $L^2$-based Sobolev space of order $\mu$ of periodic functions on $[-\pi, \pi]$, then for any $T > 0$, (1.1) has a unique solution in $C(0,T;H^\mu)$, which also belongs to $C^k(0,T;H^{\mu-3k})$ for $k \leqslant \left[\frac{\mu+2}{3}\right]$. (Here $C(0,T;X)$ is the space of continuous maps $u : [0,T] \to X$, where $X$ is a Banach space and $C^k(0,T;X)$ is the space of $X$-valued functions defined on $[0,T]$ that are $k$ times continuously differentiable.) For the modern literature on the well-posedness of the IVP for the KdV, we refer the reader to Killip & Vişan (2019) and Kappeler & Topalov (2006), and their references.

We will analyze a high-order fully discrete, conservative numerical method for (1.1). The scheme consists of a spectral Fourier–Galerkin discretization of the PDE in the spatial variable, coupled with a high-order, diagonally implicit Runge–Kutta (RK) time-stepping scheme of *composition type* based on the implicit midpoint rule (IMR). Although the analysis is done in the case of the model problem (1.1) the main ideas and techniques behind the derivation of the error estimates may be used to establish analogous results for more complicated, $L^2$-conservative periodic IVPs for one-way nonlinear dispersive wave PDEs with more general nonlinearities and linear dispersive terms and may also prove useful in analyzing temporal discretizations by more general composition-type RK schemes. As far as we know this is the first work in which fully discrete schemes for a PDE, with time stepping of RK-composition type, have been analyzed.

Among the many available $L^2$-conservative spatial discretizations for (1.1) (cf. e.g. the references of Baker *et al.*, 1983, and Bona *et al.*, 2013) we chose, for reasons of simplicity, the spectral Fourier–Galerkin method. This semidiscretization conserves the first three invariants of the KdV and is straightforward to analyze; for rigorous error estimates for the semidiscrete problem see e.g. Deng & Ma (2009), Kalisch (2005), Maday & Quarteroni (1988) and their references. In the first two of these papers one may find proofs of $L^2$ error bounds of spectral accuracy, whose rates of convergence depend on the smoothness of the initial value $u_0$. Specifically, if $N$ is the order of the trigonometric polynomials used in the Fourier basis, it is shown in Maday & Quarteroni (1988) by an energy method that if $u_0 \in H^\mu, \mu \geqslant 2$, then the $L^2$ error of the semidiscrete problem is $\mathcal{O}(N^{1-\mu})$. In Deng & Ma (2009) the estimate is improved to $\mathcal{O}(N^{-\mu})$ if $\mu \geqslant 3$, in fact, for the generalized KdV equation. In order to obtain this optimal-order result, Deng & Ma (2009) compare the semidiscrete approximation to the third-order projection of Wahlbin (1974), and the proof is accordingly more complicated. In Kalisch (2005) a result of a different kind is proved: specifically, if $u_0$ is analytic in a strip about the real axis, then the $L^2$ error bound is $\mathcal{O}(e^{-\sigma N})$, where $\sigma$ is a constant depending on $T$; the proof relies on analyticity results in Bona & Grujic (2003).

In this paper, since we will be primarily concerned with establishing error estimates for our fully discrete scheme, we give in Section 3 a simplified proof of the error of the semidiscretization with an $L^2$ error bound of $\mathcal{O}(N^{1-\mu})$, provided $\mu \geqslant 2$; the method of proof differs from that of Maday & Quarteroni (1988). An important property of the semidiscrete spectral approximation is that its temporal derivatives are bounded, uniformly with respect to $N$, in the Sobolev space norms, provided $u_0$ is sufficiently smooth; cf. Proposition 3.2. This property considerably simplifies estimating the errors of the full discretization.

An efficient time-stepping procedure for a conservative spatially discrete method for (1.1), such as the one considered here, should be chosen so that the resulting fully discrete scheme has the following properties:

- It is $L^2$-conservative, preferably symplectic: these properties will give the scheme the chance to simulate accurately properties of the solution of the KdV that depend on the balance of dispersive and nonlinear terms, such as the propagation of solitary waves with constant speed and shape and their asymptotic stability properties, for example, the resolution of general initial profiles into sequences of solitary waves plus dispersive tails, their interactions, etc. A dissipative scheme will not reproduce accurately such properties as time increases.

- It is convergent, at most under a weak mesh condition.

- It is of high temporal accuracy, in order to take advantage of the high accuracy in space.
- It may be easily implemented.

The class of implicit RK methods includes schemes that fulfill the above requirements. An example is the family of Gauss–Legendre collocation schemes. It is well known, cf. e.g. Hairer *et al.* (2004) and its references, that the $q$-stage Gauss–Legendre scheme has order of accuracy equal to $2q$, is B-stable and is symplectic. These schemes have been used for the temporal discretization of many nonlinear dispersive wave PDEs that give rise to stiff semidiscrete systems. Their convergence was analyzed in Bona *et al.* (1995) in the case of the periodic IVP for the generalized KdV equation, discretized in space by the Galerkin finite element method with smooth periodic splines.

In the paper at hand for the temporal discretization we will use implicit, symplectic RK schemes of composition type, whose general step is constructed as the composition of $s$ steps, of length $b_ik$, $1 \leqslant i \leqslant s$ (where $k$ is the basic time step), of the IMR; cf. e.g. Frutos & Sanz-Serna (1992), Hairer *et al.* (2004), Sanz-Serna & Abia (1991), Sanz-Serna & Calvo (1994), Yoshida (1990). For general RK-composition methods we refer the reader to Hairer *et al.* (2004) and its references. The particular scheme corresponding to $s = 3$, of fourth-order temporal accuracy, was used in Frutos & Sanz-Serna (1992), to integrate the IVP (1.1) for the KdV, discretized in space by finite element and spectral methods. It was also used in Dougalis *et al.* (2019) (see also the arXiv version of the paper) for long-time computations in a study of the evolution and stability of solitary waves of the generalized Benjamin equation (see also Section 7 in the sequel), discretized in space by a spectral method. It should be pointed out that the schemes in this class are not A-stable, since some of the $b_i$ are not positive and the attendant rational approximations to $e^z$ have poles in the left half of the complex plane. However, for a conservative problem like (1.1), the scheme, being symplectic, is unconditionally $L^2$-conservative and convergent under a weak mesh condition, as will be proved in Theorem 5.4 in this paper. Let us also remark that symplectic schemes have other well-known properties related to their long-time fidelity to solutions of a problem like (1.1). For example, since the spectral semidiscretization of (1.1), when implemented in the Fourier collocation form, leads to a Hamiltonian system of ordinary differential equations (ODEs) for the semidiscrete solution at the collocation points (the proof for the KdV case follows along similar lines to those in Cano, 2006 for the nonlinear wave equation and the nonlinear Schrödinger equation); the property of symplecticity (Sanz-Serna & Calvo 1994; Hairer *et al.*, 2004) ensures a good conservation of the Hamiltonian.

In Section 4.1 we review the error estimate for the fully discrete IMR-spectral scheme, while in Section 4.2 we present the RK-composition scheme under study in the context of ODEs. In Section 4.3 we consider the general $s$-stage fully discrete scheme and establish the existence of its solutions, its $L^2$-conservation property and state, under general hypotheses, a result on the uniqueness of solutions. In Section 4.4 we study the local temporal error of the time-stepping scheme with $s = 3$ stages (of fourth order of accuracy), applied to the semidiscrete system. Assuming that the solution of (1.1) is sufficiently smooth and that $k = \mathcal{O}(N^{-1})$ we prove in Proposition 4.4 that the local temporal error is $\mathcal{O}(k^5)$ in $L^2$. The result is achieved by computing the asymptotic expansions in powers of $k$, up to $\mathcal{O}(k^5)$ terms, of the intermediate steps of the local error about the points $\tau^{n,i} = t^n + (b_1 + \cdots + b_i)k$ in terms of the semidiscrete solution and its partial derivatives. We compute the coefficients of these asymptotic expansions, estimate their residuals and substitute them into the final stage of the local error equation, whereupon, after cancellation, there emerges the $\mathcal{O}(k^5)$ local error. Thus, the overall plan of the proof resembles that adopted in the case of other implicit, high-order RK schemes for the KdV and its generalized version in Bona *et al.* (1995), Dougalis & Karakashian (1985) and Karakashian & McKinney (1990) for the nonlinear Schrödinger equation in Karakashian *et al.* (1993) and for the explicit, $(4,4)$ 'classical' RK scheme for the system of shallow water equations in Antonopoulos *et al.* (2020). With the exception of Karakashian & McKinney (1990), where only the temporal discretization of the PDE was considered, in the other papers cited above the spatial discretization was effected by Galerkin finite element methods and the stages of the local temporal error were computed in terms of continuous in time finite element approximations of the solution of the PDE, such as the quasi-interpolant, the elliptic projection and the $L^2$ projection. Here, the use of the semidiscrete spectral approximation itself for this purpose simplifies the analysis; however, many technical difficulties remain and they are resolved in the course of the proof of Proposition 4.4.

In Section 5 we revert to the general $s$-stage temporal discretization scheme and, under the hypotheses that the solution of (1.1) is in $H^\mu$ for $t \in [0,T]$ for $\mu$ sufficiently large, and that the local temporal error is $\mathcal{O}(k^{\alpha+1})$ in $L^2$, we prove that there exists a constant $C$ such that if $kN \leqslant C$, the fully discrete scheme has a unique solution whose maximum $L^2$ error over $[0,T]$ has a bound of $\mathcal{O}(k^\alpha + N^{1-\mu})$. Therefore, the RK scheme with $s=3$ stages, whose local temporal error was analyzed in Section 4.4, leads to a fully discrete method with an $L^2$ error bound of $\mathcal{O}(k^4 + N^{1-\mu})$. In a remark at the end of Section 5 we discuss the convergence of a simple iterative scheme approximating the nonlinear system of equations that must be solved at each IMR stage of the RK time-stepping scheme.

In Section 6 we illustrate the convergence results and other geometric properties (in the sense of Hairer *et al.*, 2004; Sanz-Serna & Calvo, 1994) of the scheme, implemented in spectral collocation form, with some numerical experiments of simulations of solitary-wave solutions of the KdV equation. In a final Section 7 we summarize the results of the paper and indicate how they may be extended e.g. to the case of the generalized Benjamin equation, solved numerically with the present scheme in Dougalis *et al.* (2019).

As was already mentioned we will denote by $H^\mu$, for real $\mu \geqslant 0$, the $L^2$-based Sobolev space of order $\mu$ consisting of periodic functions on $(-\pi,\pi)$. For $g \in H^\mu$ its norm is given by

$$||g||_\mu = \left(\sum_{k\in\mathbb{Z}} (1+k^2)^\mu |\widehat{g}(k)|^2\right)^{1/2},$$

where $\widehat{g}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx} g(x)\,\mathrm{d}x$ is the $k$th Fourier coefficient of $g$. For $1 \leqslant p \leqslant \infty$ we denote by $W_p^\mu = W_p^\mu(-\pi,\pi)$ the real Sobolev space of periodic functions on $(-\pi,\pi)$ and denote its norm by $||\cdot||_{\mu,p}$, while $|\cdot|_\infty$ will stand for the norm of $L^\infty(-\pi,\pi)$. Finally, the inner product in $L^2 = L^2(-\pi,\pi)$ will be defined by $(u,v) = \int_{-\pi}^{\pi} u(x)\overline{v(x)}\,\mathrm{d}x$, and $||\cdot||$ will denote the induced $L^2$ norm.

## 2. Semidiscretization and preliminaries

Let $N \geqslant 1$ be an integer and consider the finite-dimensional space $S_N$ defined by

$$S_N = \mathrm{span}\{e^{ikx},\ k \text{ integer},\ -N \leqslant k \leqslant N\}.$$

Let $P_N$ denote the $L^2$ projection operator onto $S_N$ defined for $v \in L^2$ by

$$P_N v = \sum_{|k|\leqslant N} \widehat{v}_k e^{ikx},$$

where $\widehat{v}_k = \widehat{v}(k)$ is the $k$th Fourier coefficient of $v$. We note some well-known properties of $P_N$ that will be used throughout the paper. It is obvious that $P_N$ commutes with the differentiation operator $\partial_x$. We also recall that for $v \in L^2$ and $\chi \in S_N$,

$$(P_N v, \chi) = (v, \chi). \tag{2.1}$$

Moreover, cf. Mercier (1989), given integers $0 \leqslant j \leqslant \mu$, there exists a constant $C$, independent of $N$, such that for any $v \in H^\mu$,

$$||v - P_N v||_j \leqslant CN^{j-\mu}||v||_\mu, \quad \mu \geqslant 0, \tag{2.2}$$

$$|v - P_N v|_\infty \leqslant CN^{1/2-\mu}||v||_\mu, \quad \mu \geqslant 1. \tag{2.3}$$

In addition, the following inverse inequalities hold on $S_N$. Given $0 \leqslant j \leqslant \mu$ there exists a constant $C_0$ independent of $N$, such that for all $\psi \in S_N$,

$$||\psi||_\mu \leqslant C_0 N^{\mu-j}||\psi||_j, \quad ||\psi||_{\mu,\infty} \leqslant C_0 N^{1/2+\mu-j}||\psi||_j. \tag{2.4}$$

The semidiscrete Fourier–Galerkin approximation to the solution of (1.1) is a real-valued map $u^N$ : $[0, T] \to S_N$ such that, for all $\chi \in S_N$,

$$\begin{aligned} &\left(u^N_t + u^N u^N_x + u^N_{xxx}, \chi\right) = 0, \quad 0 \leqslant t \leqslant T, \\ &u^N(x, 0) = P_N u_0(x). \end{aligned} \tag{2.5}$$

It is straightforward to see that while the solution of (2.5) exists, it satisfies

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\pi}^{\pi} u^N \mathrm{d}x = 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\pi}^{\pi} (u^N)^2 \mathrm{d}x = 0, \end{aligned} \tag{2.6}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\pi}^{\pi} \left((u^N_x)^2 - \frac{1}{3}(u^N)^3\right) \mathrm{d}x = 0. \tag{2.7}$$

In particular, while $u^N$ exists, we have

$$||u^N(t)|| = ||u^N(0)||, \tag{2.8}$$

from which, from standard ODE theory, we see that $u^N(t)$ exists uniquely for all $t > 0$ and, in particular, satisfies (2.8) and the other conservation laws for $0 \leqslant t \leqslant T$.

**3. Convergence of the semidiscretization**

THEOREM 3.1 Let $u^N$ be the solution of (2.5) and suppose that $u$, the solution of (1.1), belongs to $H^\mu$, $\mu \geqslant 2$, for $t \in [0, T]$. Then for some constant $C$ independent of $N$ it holds that

$$\max_{0\leqslant t\leqslant T} ||u^N - u|| \leqslant \frac{C}{N^{\mu-1}}. \tag{3.1}$$

*Proof.* We write $u^N - u = \theta + \rho$, where $\theta = u^N - P_N u \in S_N$, and $\rho = P_N u - u$. We then have for $\chi \in S_N, 0 \leqslant t \leqslant T$, in view of (2.5), (1.1), (2.1),

$$(\theta_t, \chi) + (\theta_{xxx}, \chi) = (u_t^N + u_{xxx}^N, \chi) - (P_N(u_t + u_{xxx}), \chi) = (uu_x - u^N u_x^N, \chi).$$

Since

$$\begin{aligned} uu_x - u^N u_x^N &= uu_x - (u + \theta + \rho)(u_x + \theta_x + \rho_x) \\ &= -u\theta_x - u\rho_x - u_x\theta - \theta\theta_x - \rho_x\theta - u_x\rho - \rho\theta_x - \rho\rho_x, \end{aligned}$$

we have for $\chi \in S_N$,

$$\begin{aligned} (\theta_t, \chi) + (\theta_{xxx}, \chi) = -\big( &(u\theta_x, \chi) + (u\rho_x, \chi) + (u_x\theta, \chi) + (\theta\theta_x, \chi) \\ &+ (\rho_x\theta, \chi) + (u_x\rho, \chi) + (\rho\theta_x, \chi) + (\rho\rho_x, \chi)\big). \end{aligned}$$

Putting $\chi = \theta$ in the above and using integration by parts and periodicity we obtain for $0 \leqslant t \leqslant T$,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}||\theta||^2 = -\left(\frac{1}{2}(u_x, \theta^2) + (u\rho_x, \theta) + \frac{1}{2}(\rho_x, \theta^2) + (u_x\rho, \theta) + (\rho\rho_x, \theta)\right). \tag{3.2}$$

Now we estimate the various inner products in the right-hand side of the above, taking into account that $u \in H^\mu$, $\mu \geqslant 2$. We first have

$$|(u_x, \theta^2)| \leqslant |u_x|_\infty ||\theta||^2 \leqslant C||\theta||^2. \tag{3.3}$$

(Here and in the sequel $C$ will denote a generic constant independent of the discretization parameters.) By (2.2),

$$|(u\rho_x, \theta)| \leqslant |u|_\infty ||\rho_x|| ||\theta|| \leqslant CN^{1-\mu}||\theta||. \tag{3.4}$$

Using the inequality $|v_x|_\infty \leqslant C||v_x||^{1/2}||v_{xx}||^{1/2}$, valid in $H^2$, we see, in view of (2.2), since $\mu \geqslant 2$,

$$|(\rho_x, \theta^2)| \leqslant |\rho_x|_\infty ||\theta||^2 \leqslant CN^{\frac{3}{2}-\mu}||\theta||^2 \leqslant C||\theta||^2. \tag{3.5}$$

Also, by (2.2),

$$|(u_x\rho, \theta)| \leqslant |u_x|_\infty ||\rho|| ||\theta|| \leqslant CN^{-\mu}||\theta||. \tag{3.6}$$

And, as above,

$$|(\rho\rho_x, \theta)| \leqslant |\rho|_\infty ||\rho_x|| ||\theta|| \leqslant CN^{\frac{3}{2}-2\mu}||\theta|| \leqslant CN^{-\mu}||\theta||. \quad (3.7)$$

We conclude by (3.2)–(3.7) that

$$\frac{\mathrm{d}}{\mathrm{d}t}||\theta||^2 \leqslant C\left(N^{2(1-\mu)} + ||\theta||^2\right), \quad 0 \leqslant t \leqslant T,$$

from which, by Gronwall's lemma, since $\theta(0) = 0$, we get

$$\max_{0\leqslant t\leqslant T} ||\theta|| \leqslant CN^{1-\mu},$$

and (3.1) follows, in view of (2.2). □

For the purposes of estimating the error of the temporal discretization of (2.5) we note the following boundedness result for the semidiscrete approximation $u^N$, which is a consequence of the error estimate (3.1).

PROPOSITION 3.2 Let $u^N$ be the solution of (2.5) and suppose that the solution $u$ of (1.1) belongs to $H^\mu$ for $t \in [0, T]$. Then, given non-negative integers $j$ and $l$, and provided $\mu \geqslant \max\{2, 3j + l + 1\}$, there exists a constant $C$ independent of $N$ such that

$$\max_{0\leqslant t\leqslant T} ||\partial_t^j u^N||_l \leqslant C. \quad (3.8)$$

*Proof.* Using (2.2), (2.4) and (3.1), provided $\mu \geqslant 2$, we have

$$\begin{aligned}
||u^N||_l &\leqslant ||u - P_N u||_l + ||P_N u - u^N||_l + ||u||_l \\
&\leqslant CN^{l-\mu}||u||_\mu + CN^l\left(||P_N u - u|| + ||u - u^N||\right) + ||u||_l \\
&\leqslant CN^{l-\mu}||u||_\mu + CN^{l+1-\mu}||u||_\mu + ||u||_l.
\end{aligned}$$

Therefore, if $\mu \geqslant \max\{2, l + 1\}$, it holds that

$$\max_{0\leqslant t\leqslant T} ||u^N||_l \leqslant C. \quad (3.9)$$

Since $P_N$ is the $L^2$ projection and $u^N_{xxx} \in S_N$ we have from (2.5),

$$\partial_t u^N = -u^N_{xxx} - P_N(u^N u^N_x). \quad (3.10)$$

Hence

$$||\partial_t u^N||_l \leqslant ||u^N||_{l+3} + ||P_N(u^N u^N_x)||_l \leqslant ||u^N||_{l+3} + ||u^N u^N_x||_l.$$

From Sobolev's theorem and the fact that $H^l$ is an algebra for $l \geqslant 1$ we conclude from the above that

$$||\partial_t u^N||_l \leqslant ||u^N||_{l+3} + C||u^N||_{l+1}^2.$$

Therefore, in view of (3.9) and if $\mu \geqslant l + 4$, we have

$$||\partial_t u^N||_l \leqslant C. \tag{3.11}$$

Finally, differentiating (3.10) $j-1$ times with respect to $t$ and repeatedly using (3.9) and (3.11), we obtain (3.8). □

## 4. Full discretization by an RK method of composition type

As mentioned in the introduction we will discretize in time the IVP for the system of ODEs (ODE IVP) represented by (2.5), using an implicit $s$-stage RK-composition scheme based on the IMR. In this section, after briefly reviewing the IMR, we will present the time-stepping RK method to be analyzed, study the existence of solutions of the resulting fully discrete scheme and its $L^2$-conservation property, present a preliminary uniqueness of solutions result and, for $s = 3$, prove an $L^2$ estimate of its local temporal error.

### 4.1 *Fully discrete scheme with IMR time stepping*

A simple time-stepping method that may be used to discretize the ODE IVP (2.5) in $t$ is the IMR, which, in the case of the autonomous ODE system $\dot{y} = \phi(y)$, is the single-step scheme

$$y^{n+1} - y^n = k\phi(y^{n+1/2}),$$

where $k$ here and in the sequel will denote the (uniform) time step, $y^n$ is the approximation of $y(t^n)$, $t^n = nk$ and $y^{n+1/2} = \frac{1}{2}(y^{n+1} + y^n)$. In the case of the IVP (2.5), assuming that $T = Mk$ where $M$ is an integer, the scheme is the following: we seek $U^n \in S_N$ for $n = 0, \dots, M$, satisfying for each $\chi \in S_N$,

$$\begin{aligned}(U^{n+1} - U^n, \chi) &= k\left(-(U^{n+1/2})_{xxx} - f(U^{n+1/2})_x, \chi\right), \\ U^0 &= P_N u_0,\end{aligned} \tag{4.1}$$

where here and in the sequel we put $f(v) = v^2/2$ and $U^{n+1/2} = \frac{U^{n+1}+U^n}{2}$. For the Fourier coefficients $\widehat{U^n}(j), -N \leqslant j \leqslant N$, of $U^n$ we may write

$$\begin{aligned}\frac{\widehat{U^{n+1}}(j) - \widehat{U^n}(j)}{k} &= i\left(j^3\,\widehat{U^{n+1/2}}(j) - j\widehat{f(U^{n+1/2})}(j)\right), \\ \widehat{U^0}(j) = \widehat{u}_0(j), &\quad -N \leqslant j \leqslant N,\end{aligned}$$

where $\widehat{f(U^{n+1/2})}(j)$ denotes the $j$th Fourier coefficient of $f(U^{n+1/2})$.

It is easy to see, by writing, for each $n$, equations (4.1) in fixed-point form and applying a variant of Brouwer's fixed point theorem, that, given $U^n \in S_N$, there exists a solution $U^{n+1} \in S_N$ of the nonlinear system of equations represented by (4.1). Putting $\chi = U^n + U^{n+1}$ and using periodicity one may also obtain that the method is $L^2$-conservative, i.e. that

$$||U^n|| = ||U^0||, \quad 0 \leqslant n \leqslant M. \tag{4.2}$$

By comparing $U^n$ with $u^N(t^n)$, where $u^N$ is the solution of (2.5), using (3.1) and (3.8), one may derive in a straightforward way the following error estimate for $U^n$.

PROPOSITION 4.1 Suppose that $u$, the solution of (1.1), belongs to $H^\mu$ for $t \in [0, T]$, where $\mu \geqslant 10$. Then there exists a constant $\alpha > 0$, such that if $k \leqslant \frac{\alpha}{N}$, there exists a unique solution $\{U^n\}_{n=0}^M$ of (4.1) satisfying

$$\max_{0 \leqslant n \leqslant M} ||U^n - u(t^n)|| \leqslant C(k^2 + N^{1-\mu}), \tag{4.3}$$

where $C$ is a constant independent of $N$ and $k$.

As the error analysis of the IMR fully discrete scheme (4.1) may be viewed as a special case of the convergence proof for the general fully discrete scheme to be considered in the sequel we will not present the proof of Proposition 4.1 here.

### 4.2 *An RK composition method*

We will consider an RK method with $s$ stages for the autonomous ODE system $\dot{y} = \phi(y)$, whose Butcher tableau is of the form

$$\begin{array}{c|c} & a_{ij} \\ \hline & b_i \end{array} = \begin{array}{c|ccccc} & b_1/2 & & & & \\ & b_1 & b_2/2 & & & \\ & b_1 & b_2 & \ddots & & \\ & \vdots & \vdots & & \ddots & \\ & b_1 & b_2 & \cdots & \cdots & b_s/2 \\ \hline & b_1 & b_2 & \cdots & \cdots & b_s \end{array}, \tag{4.4}$$

where the $b_i$ are nonzero real numbers. As has been pointed out in Sanz-Serna & Abia (1991), all symplectic (canonical) RK schemes, i.e. those satisfying $b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, i \leqslant i,j \leqslant s$, with lower triangular matrix $a_{ij}$ (i.e. all diagonally implicit symplectic schemes) are of the form (4.4).

It is well known, cf. e.g. Yoshida (1990), Frutos & Sanz-Serna (1992), Sanz-Serna & Calvo (1994), Hairer *et al.* (2004) and their references, that the RK scheme corresponding to the tableau (4.4) is of *composition* type, since it may be constructed as the composition of $s$ steps of the IMR with step sizes

$b_1k, b_2k, \ldots, b_sk$, i.e. in the case of $\dot{y} = \phi(y)$ it is equivalent to the scheme

$$\begin{aligned} y^{n,1} &= y^n + b_1 k\phi\left(\frac{y^n + y^{n,1}}{2}\right), \\ y^{n,j} &= y^{n,j-1} + b_j k\phi\left(\frac{y^{n,j-1} + y^{n,j}}{2}\right), \ 2 \leqslant j \leqslant s, \\ y^{n+1} &= y^{n,s}. \end{aligned} \tag{4.5}$$

For example, a method mentioned in the above references and used in Frutos & Sanz-Serna (1992) for the temporal discretization of the KdV equation, corresponds to $s = 3$ and

$$b_1 = (2 + 2^{1/3} + 2^{-1/3})/3 = \frac{1}{2 - 2^{1/3}} \cong 1.351, \quad b_2 = 1 - 2b_1 \cong -1.702, \quad b_3 = b_1, \tag{4.6}$$

has order of accuracy $p = 4$ and is symmetric since $b_3 = b_1$. This scheme may be generalized using Yoshida's approach (Yoshida, 1990), which yields recursively symplectic symmetric methods (in our case taking the IMR as the base scheme) as follows. Let $\psi_k^{[2]}$ be the mapping that effects the step $n \mapsto n + 1$ of the IMR with step size $k$. Then the method with $s = 3$ may be viewed, in the notation of Hairer *et al.* (2004), as the composition

$$\psi_k^{[4]} = \psi_{b_3k}^{[2]} \circ \psi_{b_2k}^{[2]} \circ \psi_{b_1k}^{[2]}.$$

From $\psi_k^{[4]}$ one gets the sixth-order accurate symmetric method with $s = 3^2$,

$$\psi_k^{[6]} = \psi_{\gamma_3k}^{[4]} \circ \psi_{\gamma_2k}^{[4]} \circ \psi_{\gamma_1k}^{[4]},$$

where $\gamma_1 = \gamma_3 = \frac{1}{2-2^{1/5}}, \gamma_2 = 1 - 2\gamma_1$. In general, given the method $\psi_k^{[2r]}$ with $s = 3^{r-1}$ stages and order of accuracy $2r$, one may construct a symmetric scheme with $s = 3^r$ stages and order of accuracy $2r + 2$ by the formula

$$\psi_k^{[2r+2]} = \psi_{\delta_{3,r}k}^{[2r]} \circ \psi_{\delta_{2,r}k}^{[2r]} \circ \psi_{\delta_{1,r}k}^{[2r]},$$

where $\delta_{1,r} = \delta_{3,r} = \frac{1}{2-2^{\frac{1}{2r+1}}}, \delta_{2,r} = 1 - 2\delta_{1,r}$.

Some properties of the resulting family of $s$-stage methods are summarized below.

(i) The number of stages is $s = 3^{p-1}$ and the order of accuracy of the method is $2p$.

(ii) $\sum_{i=1}^{s} b_i = 1, \quad \sum_{i=1}^{s} b_i^j = 0$, with $s = 3^{p-1}$ and $j = 3, 5, \ldots, 2p - 1$; cf. Yoshida (1990).

(iii) The methods are symplectic and symmetric. Thus, when applied to an ODE system with a Hamiltonian structure, they will preserve important properties of the system and behave well in long-time computations (Frutos & Sanz-Serna 1992; Sanz-Serna & Calvo 1994; Hairer *et al.* 2004; Cano 2006).

(iv) Since some of the $b_i$ are negative, cf. e.g. (4.6) and property (ii) above, these schemes are not A-stable. They are however absolutely stable in a strip of finite width in $\mathrm{Re}(z) \leqslant 0$ including the imaginary axis, and therefore it is expected that a step-size restriction will be needed for stability in the case of dissipative problems.

(v) The implementation of the schemes is straightforward as, for each time step, it requires solving $s$ nonlinear systems of the size of the ODE system, as is evident from e.g. (4.5).

### 4.3 *The fully discrete scheme. $L^2$-conservation, existence and uniqueness of solutions*

The high-accuracy, straightforward manner of implementation, and the good stability properties (in the case of conservative, stiff ODE systems) of the family of RK composition methods given by (4.4) (equivalently by (4.5)), make them a good choice as time-stepping schemes for the semidiscrete IVP (2.5).

As mentioned already, the fully discrete scheme to be fully analyzed in the sequel is obtained by applying the $s$-stage RK composition method (4.4) or (4.5) to the semidiscrete problem (2.5) when $s = 3$ and the coefficients $b_i$ are given by (4.6). However, with the exception of the estimation of the local temporal error in Section 4.4, the rest of the proof of convergence holds for the general $s$-stage scheme and therefore we will treat the general case and specialize to $s = 3$ when needed. To simplify notation we let $F : S_N \to S_N$ be the nonlinear map defined for $v \in S_N$ by the equation

$$(F(v), \chi) = (-v_{xxx} - P_N f(v)_x, \chi) \quad \forall \chi \in S_N,$$

where $f(v) = v^2/2$, or equivalently by

$$F(v) = -v_{xxx} - P_N f(v)_x.$$

Note that, by periodicity,

$$(F(v), v) = 0 \quad \forall v \in S_N, \tag{4.7}$$

and that the semidiscrete IVP (2.5) may be written as

$$\begin{aligned} u_t^N &= F(u^N), \quad 0 \leqslant t \leqslant T, \\ u^N(0) &= P_N u_0. \end{aligned} \tag{4.8}$$

Using the notation introduced for the temporal discretization in Section 4.1 we write the RK scheme (4.4) applied to (4.8) as follows. For $0 \leqslant n \leqslant M$ we seek $U^n \in S_N$, approximating $u^N(t^n)$, and

$U^{n,i} \in S_N$, $1 \leqslant i \leqslant s$, such that for $0 \leqslant n \leqslant M-1$,

$$U^{n,i} = U^n + \frac{b_i k}{2} F(U^{n,i}) + k \sum_{j=1}^{i-1} b_j F(U^{n,j}), \quad 1 \leqslant i \leqslant s,$$

$$U^{n+1} = U^n + k \sum_{i=1}^{s} b_i F(U^{n,i}), \tag{4.9}$$

and $U^0 = P_N u_0$. By eliminating recursively the intermediate nonlinear terms and defining $\mu_{ij} = 2(-1)^{i+j+1}$, $1 \leqslant j < i \leqslant s$ it is easy to check that scheme (4.9) may be equivalently stated for $0 \leqslant n \leqslant M-1$ as

$$U^{n,i} = (-1)^{i+1} U^n + \frac{b_i k}{2} F(U^{n,i}) + \sum_{j=1}^{i-1} \mu_{ij} U^{n,j}, \quad 1 \leqslant i \leqslant s,$$

$$U^{n+1} = (-1)^s U^n + 2 \sum_{j=1}^{s} (-1)^{s-j} U^{n,j}, \tag{4.10}$$

and $U^0 = P_N u_0$. As already mentioned, scheme (4.9) is also equivalent to the following IMR-type formulation (cf. (4.5)) in which, given $U^n \in S_N, 0 \leqslant n \leqslant M-1$, $Y^{n,i} \in S_N, i \leqslant i \leqslant s$ and $U^{n+1} \in S_N$ are computed by the formulas

$$Y^{n,1} = U^n + kb_1 F\left(\frac{Y^{n,1} + U^n}{2}\right),$$

$$Y^{n,i} = Y^{n,i-1} + kb_i F\left(\frac{Y^{n,i} + Y^{n,i-1}}{2}\right), \quad 2 \leqslant i \leqslant s,$$

$$U^{n+1} = Y^{n,s}, \tag{4.11}$$

and $U^0 = P_N u_0$. Note that the intermediate approximations $Y^{n,i}$ of (4.11) are related to the $U^{n,i}$ of (4.9) or (4.10) by the formulas

$$Y^{n,i} = 2U^{n,i} - Y^{n,i-1}, \ 2 \leqslant i \leqslant s, \quad Y^{n,1} = 2U^{n,1} - U^n.$$

Any one of the formulations (4.9)–(4.11) may be used to study the properties and the convergence of the fully discrete scheme. We will mainly use (4.11), which brings out the fact that the scheme is an $s$-stage composition method with IMR as its base scheme.

As is expected by the symplecticity of the RK method (4.4), (4.5), the fully discrete schemes (4.9)–(4.11) are $L^2$-conservative. Taking, for example, (4.11) and supposing that given $U^n \in S_N$ it has a solution $Y^{n,i} \in S_N$, $1 \leqslant i \leqslant s$, then, if $i \geqslant 2$,

$$(Y^{n,i} - Y^{n,i-1}, Y^{n,i} + Y^{n,i-1}) = 2kb_i \left(F\left(\frac{Y^{n,i} + Y^{n,i-1}}{2}\right), \frac{Y^{n,i} + Y^{n,i-1}}{2}\right),$$

which, in view of (4.7), yields $||Y^{n,i-1}|| = ||Y^{n,i}||$. The same argument works for $i = 1$ if we put $Y^{n,0} = U^n$ and yields $||Y^{n,1}|| = ||U^n||$. Therefore, $||U^{n+1}|| = ||Y^{n,i}|| = ||U^n||, 1 \leqslant i \leqslant s$, and overall

$$||U^n|| = ||U^0||, \ 0 \leqslant n \leqslant M,$$

provided $U^n, 1 \leqslant n \leqslant M$ exist. The existence of solutions may be proved by a variant of Brouwer's fixed-point theorem. We use (4.11) again.

PROPOSITION 4.2 Given $U^n \in S_N$ there are $Y^{n,i} \in S_N, 1 \leqslant i \leqslant s$ and $U^{n+1}$ in $S_N$ satisfying (4.11).

*Proof.* Putting $Z = \frac{Y^{n,1}+U^n}{2}$, we write the first equation in (4.11) in the form $Z - U^n = \frac{kb_1}{2}F(Z)$. Hence, if we define $G : S_N \to S_N$ for $v \in S_N$ as $G(v) = v - U^n - \frac{kb_1}{2}F(v)$, for $\chi \in S_N$ we have $(G(v), \chi) = (v - U^n, \chi) - \frac{kb_1}{2}(F(v), \chi)$. Taking $\chi = v$ we get, in view of (4.7), $(G(v), v) = ||v||^2 - (U^n, v) \geqslant ||v|| \, (||v|| - ||U^n||)$. Therefore, if $||v|| = ||U^n||$, then $(G(v), v) \geqslant 0$. By the definition of $F$ and the inverse inequalities (2.4) it follows that $F$, and hence $G$, is continuous on $S_N$. By a well-known variant of Brouwer's fixed-point theorem (see e.g. Bona *et al.*, 1995, Lemma 3.1) there exists $Z \in S_N$ with $||Z|| = ||U^n||$, such that $G(Z) = 0$, i.e. $Z - U^n = \frac{kb_1}{2}F(Z)$, and the existence of $Y^{n,1}$ follows. (For $Y^{n,1}$ we know *a priori* that $||Y^{n,1}|| = ||U^n||$.) In an analogous way we may prove recursively the existence of $Y^{n,i}, 2 \leqslant i \leqslant s$, satisfying (4.11). □

The uniqueness of solutions of the nonlinear systems represented by the nonlinear equations in (4.11) will be shown in the course of the proof of convergence of the fully discrete scheme in Section 5. The following lemma establishes uniqueness under a condition that will be verified in Section 5.

LEMMA 4.3 Suppose that $U^n$ and $Y^{n,i}, 1 \leqslant i \leqslant s$ are solutions of (4.11) satisfying $|U^n|_\infty \leqslant R$, $|Y^{n,i}|_\infty \leqslant R, 1 \leqslant i \leqslant s$ for some constant $R$. Then the $Y^{n,i}, 1 \leqslant i \leqslant s$ are unique, provided that

$$\frac{k}{2} \max_{1 \leqslant i \leqslant s} |b_i| C_0 N R < 1,$$

where $C_0$ is the constant in the inverse properties (2.4).

*Proof.* We prove the uniqueness of $Y^{n,1}$; that of $Y^{n,i}, i \geqslant 2$ follows by a similar argument. Suppose $Z_1, Z_2 \in S_N$ are two solutions of the first equation in (4.11). (Note that $||Z_1|| = ||Z_2|| = ||U^n||$.) Then $Z_1 - Z_2 = kb_1\big(F\big(\frac{Z_1+U^n}{2}\big) - F\big(\frac{Z_2+U^n}{2}\big)\big)$. Taking the inner product of both sides of this equation with $Z_1 - Z_2$ and using periodicity, the Cauchy–Schwarz inequality and (2.4) we have if $Z_1 \neq Z_2$ that

$$||Z_1 - Z_2|| \leqslant k|b_1|C_0 N \left\| f\left(\frac{Z_1 + U^n}{2}\right) - f\left(\frac{Z_2 + U^n}{2}\right) \right\|.$$

By the definition of $f$ and the hypothesis of the lemma we see that

$$\left\| f\left(\frac{Z_1 + U^n}{2}\right) - f\left(\frac{Z_2 + U^n}{2}\right) \right\| \leqslant \frac{||Z_1 - Z_2||}{2} \frac{1}{4} |Z_1 + U^n + Z_2 + U^n|_\infty \leqslant \frac{1}{2} ||Z_1 - Z_2|| R.$$

These two inequalities imply that $1 \leqslant \frac{1}{2}k|b_1|C_0NR$, which contradicts the other hypothesis of the lemma. Therefore $Z_1 = Z_2$. □

### 4.4 *Local temporal error of the fully discrete scheme for $s = 3$*

In this section we suppose that $s = 3$ and that the coefficients $b_i$ are given by (4.6). The local temporal error of the resulting scheme (4.11) is defined in terms of the semidiscrete approximation $u^N$. For this purpose we let for $0 \leqslant n \leqslant M$, $V^n = u^N(t^n)$, and $V^{n,i} \in S_N$ for $0 \leqslant i \leqslant 3, 0 \leqslant n \leqslant M-1$, be given by

$$V^{n,0} = V^n,$$
$$V^{n,i} = V^{n,i-1} + kb_iF\left(\frac{V^{n,i} + V^{n,i-1}}{2}\right), \quad 1 \leqslant i \leqslant 3. \tag{4.12}$$

The local temporal error $\theta^n \in S_N$, $0 \leqslant n \leqslant M-1$ is then

$$\theta^n = V^{n+1} - V^{n,3} \equiv u^N(t^{n+1}) - V^{n,3}. \tag{4.13}$$

Obviously, cf. Section 4.3, the $V^{n,i}$ exist and satisfy the $L^2$-conservation laws

$$||V^{n,i}|| = ||V^n|| = ||u^N(t^n)|| = ||u^N(0)||.$$

The consistency of the scheme is established in the following.

Proposition 4.4 Let $V^{n,i}$ and $\theta^n$ be defined by (4.12) and (4.13) and suppose the $b_i$ are given by (4.6). Let $u$, the solution of (1.1), belong to $H^\mu$ for $0 \leqslant t \leqslant T$ and let $\mu$ be sufficiently large. Suppose there exists a constant $C_1$ such that $kN \leqslant C_1$. Then, for $k$ sufficiently small, there exists a constant $C$ independent of $k$ and $N$ such that

$$\max_{0\leqslant n\leqslant M-1} ||\theta^n|| \leqslant Ck^5.$$

*Proof.* The plan of the proof is first to obtain asymptotic expressions of the $V^{n,i}$, $i = 1, 2$ of the form

$$V^{n,1} = u^N(\tau^{n,1}) + A_1k^3 + A_2k^4 + e^{n,1}, \tag{4.14}$$

$$V^{n,2} = u^N(\tau^{n,2}) + B_1k^3 + B_2k^4 + e^{n,2}, \tag{4.15}$$

where $\tau^{n,1} = t^n + kb_1, \tau^{n,2} = t^n + k(b_1 + b_2), A_i, B_i, e^{n,i} \in S_N$ and $||e^{n,i}|| \leqslant Ck^5, i = 1, 2$; then we show that

$$V^{n,3} = u^N(t^{n+1}) + e^{n,3}, \tag{4.16}$$

where $||e^{n,3}|| \leqslant Ck^5$, implying that $||\theta^n|| \leqslant Ck^5$. These estimates will be uniformly valid in $n$. The coefficients $A_i$ and $B_i$ will be $\mathcal{O}(1)$. Here, and in the sequel, $C$ will denote generic constants independent of $k$ and $N$. As the necessary computations require a lot of algebra we present here the main steps of the proof in outline; the interested reader may find the requisite details in the arXiv version of the paper Dougalis & Durán (2020).

(i) Asymptotic expansion of $V^{n,1}$. Determination of the coefficients $A_1, A_2$.

From (4.12) it follows for $i = 1$,

$$V^{n,1} = u^N - kb_1 \left(\partial_x^3 \left(\frac{V^{n,1} + u^N}{2}\right)\right) - \frac{kb_1}{4} P_N \left((V^{n,1} + u^N)(V^{n,1} + u^N)_x\right). \tag{4.17}$$

In (4.17) and in the sequel we put $u^N = u^N(t^n)$. Similarly, we will suppress the argument $t^n$ from derivatives of $u^N$, i.e. write $u_x^N = u_x^N(t^n)$, $u_t^N = u_t^N(t^n)$, etc. For the intermediate times $\tau^{n,i}$ we will write in full $u^N(\tau^{n,i})$, etc.

Inserting in (4.17) the assumed expression (4.14) for $V^{n,1}$ and expanding in Taylor series about $t^n$ gives

$$\begin{aligned}
&u^N + kb_1 u_t^N + \frac{k^2 b_1^2}{2} u_{tt}^N + \frac{k^3 b_1^3}{6} u_{ttt}^N + \frac{k^4 b_1^4}{24} \partial_t^4 u^N + \rho_1 + A_1 k^3 + A_2 k^4 + e^{n,1} \\
&= u^N - kb_1 \partial_x^3 \left( u^N + \frac{kb_1}{2} u_t^N + \frac{k^2 b_1^2}{4} u_{tt}^N + \frac{k^3 b_1^3}{12} u_{ttt}^N + \rho_2 \right) \\
&\quad - \frac{kb_1}{2} \left( k^3 \partial_x^3 A_1 + k^4 \partial_x^3 A_2 + \partial_x^3 e^{n,1} \right) \\
&\quad - \frac{kb_1}{4} P_N \Big( (u^N + u^N(\tau^{n,1}))(u^N + u^N(\tau^{n,1}))_x + k^3((u^N + u^N(\tau^{n,1}))A_1)_x \\
&\quad + k^4((u^N + u^N(\tau^{n,1}))A_2)_x + k^6 A_1 A_{1,x} + k^7 (A_1 A_2)_x + k^8 A_2 A_{2,x} + \mathcal{A}(e^{n,1}) \Big),
\end{aligned} \tag{4.18}$$

where the residuals $\rho_1, \rho_2 \in S_N$ satisfy

$$||\rho_1||_j \leqslant Ck^5 \max_t ||\partial_t^5 u^N||_j, \quad ||\rho_2||_j \leqslant Ck^4 \max_t ||\partial_t^4 u^N||_j \tag{4.19}$$

and

$$\mathcal{A}(e^{n,1}) = ((u^N + u^N(\tau^{n,1}))e^{n,1})_x + k^3(A_1 e^{n,1})_x + k^4(A_2 e^{n,1})_x + e^{n,1} e_x^{n,1}. \tag{4.20}$$

Now equate the terms of equal powers of $k$ in the left- and right-hand sides of (4.18), after expanding the $u^N(\tau^{n,1})$ terms about $t^n$. The $\mathcal{O}(1), \mathcal{O}(k), \mathcal{O}(k^2)$ terms give identities, as may be seen by (3.10), and differentiating (3.10) with respect to $t$. Equating $\mathcal{O}(k^3)$ terms gives

$$\frac{k^3 b_1^3}{6} \partial_t^3 u^N + k^3 A_1 = -\frac{k^3 b_1^3}{4} \partial_x^3 u_{tt}^N - kb_1 T_2, \tag{4.21}$$

where

$$T_2 = \frac{k^2 b_1^2}{4} P_N \left( (u^N u_x^N)_{tt} - u_t^N u_{xt}^N \right).$$

Substituting this into (4.21) and using the equation that results by differentiating (3.10) twice with respect to $t$ we obtain

$$A_1 = b_1^3 \left( \frac{1}{12} u_{ttt}^N + \frac{1}{4} P_N \left( u_t^N u_{xt}^N \right) \right). \tag{4.22}$$

Equating $\mathcal{O}(k^4)$ *terms* yields

$$\frac{k^4 b_1^4}{24} \partial_t^4 u^N + k^4 A_2 = -\frac{k^4 b_1^4}{12} \partial_t^3 \partial_x^3 u^N - \frac{k^4 b_1}{2} \partial_x^3 A_1 - k b_1 T_3 - \frac{k^4 b_1}{2} S_0, \tag{4.23}$$

where

$$-k b_1 T_3 = -\frac{k^4 b_1^4}{12} P_N \left( (u^N u_x^N)_{ttt} - \frac{3}{2} \left( u_t^N u_{xtt}^N + u_{tt}^N u_{xt}^N \right) \right),$$

and, in view of (4.22),

$$-\frac{k^4 b_1}{2} S_0 = -\frac{k^4 b_1^4}{2} \left( \frac{1}{12} P_N \left( u^N u_{ttt}^N \right)_x + \frac{1}{4} P_N \left( u^N P_N \left( u_t^N u_{xt}^N \right) \right)_x \right).$$

Therefore, by (4.22), (4.23) and the above, using Leibniz's formula for $(v v_x)_{ttt}$ and replacing the linear term $\partial_t^3(-\partial_x^3 u^N)$ by $\partial_t^4 u^N + P_N \partial_t^3 (u^N u_x^N)$ in view of (3.10), we obtain, after some algebra, that

$$A_2 = \frac{b_1^4}{12} \partial_t^4 u^N + \frac{b_1^4}{8} \left( -P_N \partial_x^3 (u_t^N u_{xt}^N) + 2 P_N (u_t^N u_{xtt}^N + u_{tt}^N u_{xt}^N) - P_N \left( u^N P_N \left( u_t^N u_{xt}^N \right) \right)_x \right). \tag{4.24}$$

(ii) Estimation of $e^{n,1}$.

Having determined $A_1$ and $A_2$ we now equate the $\mathcal{O}(k^5)$ (and higher-order) terms in (4.18) in order to find an equation for $e^{n,1}$. This gives

$$e^{n,1} + \frac{k b_1}{2} \partial_x^3 e^{n,1} = \Gamma_1 - \frac{k b_1}{4} P_N \mathcal{A}(e^{n,1}), \tag{4.25}$$

where

$$\Gamma_1 = -\rho_1 - kb_1\partial_x^3\rho_2 - \frac{k^5 b_1}{2}\partial_x^3 A_2 - kb_1 P_N\left[\left(\frac{u^N + u^N(\tau^{n,1})}{2}\right)\left(\frac{u^N + u^N(\tau^{n,1})}{2}\right)_x\right]\Big|_{\mathcal{O}(k^4)}$$
$$- \frac{k^4 b_1}{4} P_N\left[\left(\frac{u^N + u^N(\tau^{n,1})}{2})A_1\right)\right]_x\Big|_{\mathcal{O}(k)} - \frac{k^5 b_1}{4} P_N\left[\left(\frac{u^N + u^N(\tau^{n,1})}{2}\right)A_2\right]_x$$
$$- \frac{kb_1}{4} P_N\left(k^6 A_1 A_{1,x} + k^7 (A_1 A_2)_x + k^8 A_2 A_{2,x}\right), \tag{4.26}$$

where $\rho_1, \rho_2$ satisfy (4.19) and $\mathcal{A}(e^{n,1})$ is defined in (4.20). The two terms denoted above as $\cdots\big|_{\mathcal{O}(k^4)}, \cdots\big|_{\mathcal{O}(k)}$ will include Taylor remainders of the indicated order. We will prove below that for $\mu$ sufficiently large there is a constant $C$, independent of $N$ and $k$, such that

$$||\Gamma_1|| \leqslant Ck^5. \tag{4.27}$$

Assuming for the moment the validity of (4.27), and taking inner products of both sides of (4.25) with $e^{n,1} \in S_N$, we have, using integration by parts, periodicity and (4.20), that

$$||e^{n,1}||^2 = (\Gamma_1, e^{n,1}) - \frac{kb_1}{8}((u^N + u^N(\tau^{n,1}))_x e^{n,1}, e^{n,1}) - \frac{k^4 b_1}{8}(A_{1,x}e^{n,1}, e^{n,1}) - \frac{k^5 b_1}{8}(A_{2,x}e^{n,1}, e^{n,1}).$$

Therefore,

$$||e^{n,1}||^2 \leqslant ||\Gamma_1|| ||e^{n,1}|| + Ck|(u^N + u^N(\tau^{n,1}))_x|_\infty ||e^{n,1}||^2 + Ck^4|A_{1,x}|_\infty ||e^{n,1}||^2 + Ck^5|A_{2,x}|_\infty ||e^{n,1}||^2.$$

Using Proposition 3.2 the fact that $|P_N v|_\infty \leqslant C||v||_1$, which follows from (2.3) and Sobolev's theorem, and (4.22), (4.24), we see that for $0 \leqslant t \leqslant T$,

$$|(u^N + u^N(\tau^{n,1}))_x|_\infty \leqslant C \max_t ||u^N||_2 \leqslant C \text{ for } \mu \geqslant 3, \tag{4.28}$$

$$|A_{1,x}|_\infty \leqslant C \text{ for } \mu \geqslant 12, \qquad |A_{2,x}|_\infty \leqslant C \text{ for } \mu \geqslant 15. \tag{4.29}$$

Therefore, using (4.22), (4.24), (4.27)–(4.29) we see for $k$ sufficiently small that

$$||e^{n,1}|| \leqslant Ck^5 \text{ for } \mu \geqslant 15. \tag{4.30}$$

Now, by (4.19), (4.26), (4.22), (4.24), Proposition 3.2 and estimates like (4.28)–(4.29), we may see that (4.27) holds for $\mu \geqslant 16$.

We note, for future use, that (4.25) gives

$$k||\partial_x^3 e^{n,1}|| \leqslant C||\Gamma_1|| + Ck||\mathcal{A}(e^{n,1})|| + C||e^{n,1}||.$$

Therefore, by (4.20), (4.28), (4.29), (4.30) and (2.4), we have, provided $k = \mathcal{O}(N^{-1})$, that

$$k||\partial_x^3 e^{n,1}|| \leqslant Ck^5 \text{ if } \mu \geqslant 16, k = \mathcal{O}(N^{-1}). \tag{4.31}$$

For estimates (4.30) and (4.31) we tracked, as an example, lower bounds of $\mu$ so that the constants involved are bounded. In the sequel we will just assume that $\mu$ is ‘sufficiently large’. Sufficient lower bounds of $\mu$ can always be retrieved if needed.

(iii) Asymptotic expansion of $V^{n,2}$. Determination of the coefficients $B_1, B_2$.

Using the ansatz (4.15) we now evaluate the $\mathcal{O}(1)$ quantities $B_1$ and $B_2$. From (4.12) for $i = 2$ it follows that

$$V^{n,2} = V^{n,1} - kb_2\left(\partial_x^3\left(\frac{V^{n,1}+V^{n,2}}{2}\right)\right) - \frac{kb_2}{4}P_N\left((V^{n,1}+V^{n,2})(V^{n,1}+V^{n,2})_x\right). \tag{4.32}$$

Since in our case $b_1 + b_2 = 1 - b_1 \cong -0.351$, in the first step of the fully discrete scheme $\tau^{0,2} = k(b_1 + b_2)$ will be negative. In addition, since $b_1 > 1$, for $n = M - 1$ $\tau^{n,1}$ will exceed $T$. Using the reversibility for $t < 0$ of the KdV it is easy to see that $u^N(t)$ is defined for $t \in [-k, 0]$ and satisfies the semidiscrete equations (2.5) in $[-k, 0]$. Obviously, we may also extend the well-posedness of (1.1) and the validity of (2.5) up to $t = T + k$, as we have tacitly assumed in parts (i) and (ii) of the proof already. Hence, the error estimate (3.1) and the boundedness estimate (3.8) are valid with the maximum taken over $[-k, T + k]$ now. In the sequel we will accordingly not specify the range of the subscripted max in time, as such estimates are obviously valid in the relevant intervals of $t$.

Inserting in (4.32) the assumed expression (4.15) for $V^{n,2}$, using (4.14), and expanding (in the linear terms) in Taylor series about $t^n$ gives

$$\begin{aligned}
&u^N + k(b_1+b_2)u_t^N + \frac{k^2(b_1+b_2)^2}{2}u_{tt}^N + \frac{k^3(b_1+b_2)^3}{6}u_{ttt}^N + \frac{k^4(b_1+b_2)^4}{24}\partial_t^4 u^N + \rho_3 + B_1k^3 + B_2k^4 + e^{n,2}\\
&= u^N + kb_1u_t^N + \frac{k^2b_1^2}{2}u_{tt}^N + \frac{k^3b_1^3}{6}u_{ttt}^N + \frac{k^4b_1^4}{24}\partial_t^4u^N + \rho_1 + e^{n,1} + A_1k^3 + A_2k^4\\
&\quad - \frac{kb_2}{2}\partial_x^3\Big(2u^N + k(2b_1+b_2)u_t^N + \frac{k^2}{2}(b_1^2 + (b_1+b_2)^2)u_{tt}^N + \frac{k^3}{6}(b_1^3 + (b_1+b_2)^3)u_{ttt}^N + \rho_4 + A_1k^3\\
&\quad + A_2k^4 + B_1k^3 + B_2k^4 + e^{n,1} + e^{n,2}\Big) - \frac{kb_2}{4}P_N\Big((u^N(\tau^{n,1}) + u^N(\tau^{n,2}))(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))_x\\
&\quad + k^3\left((A_1+B_1)(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))\right)_x + k^4\left((A_2+B_2)(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))\right)_x\\
&\quad + k^6(A_1+B_1)(A_1+B_1)_x + k^7\left((A_1+B_1)(A_2+B_2)\right)_x + k^8(A_2+B_2)(A_2+B_2)_x + \mathcal{B}(e^{n,1}, e^{n,2})\Big),
\end{aligned} \tag{4.33}$$

where the residuals $\rho_3$ and $\rho_4 \in S_N$ satisfy

$$||\rho_3||_j \leqslant Ck^5 \max_t ||\partial_t^5 u^N||_j, \quad ||\rho_4||_j \leqslant Ck^4 \max_t ||\partial_t^4 u^N||_j, \tag{4.34}$$

and $\mathcal{B}(e^{n,1}, e^{n,2})$ is given by

$$\mathcal{B}(e^{n,1}, e^{n,2}) = \Big((u^N(\tau^{n,1}) + u^N(\tau^{n,2}))(e^{n,1} + e^{n,2})\Big)_x + k^3\Big((A_1 + B_1)(e^{n,1} + e^{n,2})\Big)_x$$
$$+ k^4\Big((A_2 + B_2)(e^{n,1} + e^{n,2})\Big)_x + (e^{n,1} + e^{n,2})(e^{n,1} + e^{n,2})_x. \quad (4.35)$$

We now equate, as before, terms of the same power of $k$ in both sides of (4.33). (For this purpose we will need to expand some $u^N(\tau^{n,i})$ terms in the right-hand side of (4.33) in Taylor series about $t = t^n$.) It is straightforward to see that we get identities by equating the $\mathcal{O}(1), \mathcal{O}(k)$ and $\mathcal{O}(k^2)$ terms in both sides of (4.33), as may be seen by (3.10) and differentiating (3.10) once with respect to $t$. Now equating $\mathcal{O}(k^3)$ terms in (4.33) gives

$$\frac{k^3}{6}(b_1 + b_2)^3 u^N_{ttt} + k^3 B_1 = \frac{k^3}{6} b_1^3 u^N_{ttt} + k^3 A_1 - \frac{k^3}{4} b_2(b_1^2 + (b_1 + b_2)^2)\partial_x^3 u^N_{tt} - \frac{kb_2}{4}\Delta_2, \quad (4.36)$$

where

$$\Delta_2 = P_N\left((u^N(\tau^{n,1}) + u^N(\tau^{n,2}))(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))_x\Big|_{\mathcal{O}(k^2)}\right)$$
$$= k^2 P_N\Big((b_1^2 + (b_1 + b_2)^2)(u^N u^N_{ttx} + u^N_x u^N_{tt}) + (2b_1 + b_2)^2 u^N_t u^N_{tx}\Big).$$

Differentiating (3.10) twice with respect to $t$, using Leibniz's rule for derivatives of products in the resulting equation, and using the facts that $b_1 = b_3, b_1^3 + b_2^3 + b_3^3 = 0$, from which $b_2^3 = -2b_1^3$, we see after some algebra and using (4.22) that

$$B_1 = -A_1. \quad (4.37)$$

From (4.33), equating $\mathcal{O}(k^4)$ terms, using appropriate Taylor expansions, and the fact that $A_1 + B_1 = 0$, we obtain

$$B_2 = \frac{1}{24}(b_1^4 - (b_1 + b_2)^4)\partial_t^4 u^N + A_2 - \frac{b_2}{12}(b_1^3 + (b_1 + b_2)^3)\partial_x^3 u^N_{ttt} - \frac{b_2}{4}\Delta_3, \quad (4.38)$$

where

$$\Delta_3 = P_N\Big(\frac{1}{3}(b_1^3 + (b_1 + b_2)^3)(u^N u^N_{tttx} + u^N_{ttt} u^N_x) + \frac{1}{2}(b_1^3 + b_1^2(b_1 + b_2) + b_1(b_1 + b_2)^2$$
$$+ (b_1 + b_2)^3)(u^N u^N_{ttx} + u^N_{tt} u^N_{tx})\Big). \quad (4.39)$$

Differentiating (3.10) three times with respect to $t$ and using Leibniz's rule on the nonlinear terms, and taking into account the relations $2b_1 + b_2 = 1, 2b_1^3 + b_2^3 = 0$, we see, after a number of algebraic computations, that $B_2$ is given by the formula

$$B_2 = A_2 - \frac{b_1^3}{12}\partial_t^4 u^N + \frac{b_2^3}{8} P_N\left(u_t^N u_{xtt}^N + u_{tt}^N u_{xt}^N\right). \tag{4.40}$$

(iv) Estimation of $e^{n,2}$.

Having determined the $\mathcal{O}(1)$ quantities $B_1, B_2 \in S_N$ we equate now the $\mathcal{O}(k^5)$ and higher-order terms in (4.33) in order to find an equation for the residual $e^{n,2}$. This yields (if we use the fact that $A_1 + B_1 = 0$)

$$e^{n,2} + \frac{kb_2}{2}\partial_x^3 e^{n,2} = \Gamma_2 + e^{n,1} - \frac{kb_2}{2}\partial_x^3 e^{n,1} - \frac{kb_2}{4} P_N(\mathcal{B}(e^{n,1}, e^{n,2})), \tag{4.41}$$

where

$$\begin{aligned}\Gamma_2 = {} & \rho_1 - \rho_3 - \frac{kb_2}{4}\partial_x^3 \rho_4 - \frac{k^5 b_2}{2}\partial_x^3 (A_2 + B_2)\\ & - \frac{kb_2}{4} P_N\Big[(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))_x\Big|_{\mathcal{O}(k^4)}\\ & + k^4\left((A_2 + B_2)(u^N(\tau^{n,1}) + u^N(\tau^{n,2}))\Big|_{\mathcal{O}(k)}\right)_x + k^8 (A_2 + B_2)(A_2 + B_2)_x\Big].\end{aligned} \tag{4.42}$$

We will prove below that for $\mu$ sufficiently large there is a constant $C$, independent of $N$ and $k$, such that

$$||\Gamma_2|| \leqslant Ck^5. \tag{4.43}$$

Assuming for the time being the validity of (4.43) and taking inner products in (4.41) with $e^{n,2}$ we have, using (4.35), periodicity and the fact that $A_1 + B_1 = 0$, that

$$\begin{aligned}||e^{n,2}||^2 = {} & (\Gamma_2, e^{n,2}) + (e^{n,1}, e^{n,2}) - \frac{kb_2}{2}(\partial_x^3 e^{n,1}, e^{n,2}) - \underbrace{\frac{kb_2}{4}\left(((u^N(\tau^{n,1}) + u^N(\tau^{n,2}))(e^{n,1} + e^{n,2}))_x, e^{n,2}\right)}_{\mathrm{I}}\\ & - \underbrace{\frac{k^5 b_2}{4}\left(((A_2 + B_2)(e^{n,1} + e^{n,2}))_x, e^{n,2}\right)}_{\mathrm{II}} - \underbrace{\frac{kb_2}{4}\left((e^{n,1} + e^{n,2})(e^{n,1} + e^{n,2})_x, e^{n,2}\right)}_{\mathrm{III}}.\end{aligned} \tag{4.44}$$

For term I above, using integration by parts, and taking into account (3.8), (2.4), (4.30) and the fact that $kN = \mathcal{O}(1)$, we obtain, for $\mu$ sufficiently large,

$$|I| \leqslant Ck^5||e^{n,2}|| + Ck||e^{n,2}||^2.$$

To estimate II, note that by (4.24), (4.29), (4.40) and (3.8), we have, for $\mu$ sufficiently large, $||A_2 + B_2||_{1,\infty} \leqslant C$. Therefore, using integration by parts, (4.30), (2.4), the fact that $kN = \mathcal{O}(1)$ and taking $\mu$ sufficiently large, we see that

$$|II| \leqslant Ck^9||e^{n,2}|| + Ck^5||e^{n,2}||^2.$$

Finally, for term III, using integration by parts, (4.30), (2.4) and that $kN = \mathcal{O}(1)$, we get

$$|III| \leqslant Ck^{9.5}||e^{n,2}|| + Ck^{4.5}||e^{n,2}||^2.$$

Using these estimates in (4.44) we obtain by (4.30), (4.31), (4.43), the fact that $kN = \mathcal{O}(1)$, taking $\mu$ sufficiently large and $k$ sufficiently small, that

$$||e^{n,2}|| \leqslant Ck^5. \tag{4.45}$$

As for $e^{n,1}$ it turns out that we will need in the sequel an optimal-order estimate for the term $k||\partial_x^3 e^{n,2}||$ under no prohibitive stability assumptions. For this purpose, note that for $\mu$ sufficiently large, we have from (4.35), using (4.30), (4.45), (2.4) and $kN = \mathcal{O}(1)$, that $k||\mathcal{B}(e^{n,1}, e^{n,2})|| \leqslant Ck^5$. It follows from (4.41), (4.43), (4.30), (4.45), (4.31) that

$$k||\partial_x^3 e^{n,2}|| \leqslant Ck^5. \tag{4.46}$$

Note finally that (4.43) follows if we use Taylor expansions to the required order, take $\mu$ sufficiently large and use (4.19), (4.34), (4.24), (4.40) in conjunction with (3.8).

(v) Final consistency step: verify that (4.16) holds with $e^{n,3}$ satisfying $||e^{n,3}|| \leqslant Ck^5$.

In this final step we let $e^{n,3} \in S_N$ be defined by (4.16), find a suitable equation for $e^{n,3}$ (as we did for $e^{n,1}$ and $e^{n,2}$) and prove that $||e^{n,3}|| \leqslant Ck^5$. For this purpose we substitute (4.16) in the equation for $V^{n,3}$ in (4.12) and prove that $||e^{n,3}|| \leqslant Ck^5$, using the expansion (4.15) for $V^{n,2}$ and the estimates that we have for $B_i, e^{n,2}$.

Substituting $V^{n,3}$ from (4.16) in (4.12) and using (4.15) and Taylor expansions about $t^n$ in the linear terms, we get

$$\begin{aligned}
&u^N + ku_t^N + \frac{k^2}{2}u_{tt}^N + \frac{k^3}{6}u_{ttt}^N + \frac{k^4}{24}\partial_t^4 u^N + \rho_5 + e^{n,3} \\
&= u^N + k(b_1+b_2)u_t^N + \frac{k^2}{2}(b_1+b_2)^2 u_{tt}^N + \frac{k^3}{6}(b_1+b_2)^3 u_{ttt}^N + \frac{k^4}{24}(b_1+b_2)^4 \partial_t^4 u^N + \rho_6 + B_1k^3 + B_2k^4 \\
&\quad + e^{n,2} - \frac{kb_3}{2}\partial_x^3\Big(2u^N + k(b_1+b_2+1)u_t^N + \frac{k^2}{2}((b_1+b_2)^2+1)u_{tt}^N + \frac{k^3}{6}((b_1+b_2)^3+1)u_{ttt}^N + \rho_7\Big) \\
&\quad - \frac{kb_3}{2}\partial_x^3 e^{n,3} - \frac{kb_3}{2}\partial_x^3(B_1k^3 + B_2k^4 + e^{n,2}) - \frac{kb_3}{4}P_N\Big((u^N(\tau^{n,2}) + u^N(t^{n+1}))(u^N(\tau^{n,2}) + u^N(t^{n+1}))_x \\
&\quad + k^3\Big((u^N(\tau^{n,2}) + u^N(t^{n+1}))B_1\Big)_x + k^4\Big((u^N(\tau^{n,2}) + u^N(t^{n+1}))B_2\Big)_x + k^6 B_1 B_{1x} \\
&\quad + k^7(B_1B_2)_x + k^8(B_2B_{2x}) + \mathcal{E}(e^{n,2}, e^{n,3})\Big),
\end{aligned} \tag{4.47}$$

where

$$\mathcal{E}(e^{n,2}, e^{n,3}) = \Big( (u^N(\tau^{n,2}) + u^N(t^{n+1}))(e^{n,2} + e^{n,3}) \Big)_x + k^3 \Big( B_1(e^{n,2} + e^{n,3}) \Big)_x$$
$$+ k^4 \Big( B_2(e^{n,2} + e^{n,3}) \Big)_x + (e^{n,2} + e^{n,3})(e^{n,2} + e^{n,3})_x \tag{4.48}$$

and where the residuals $\rho_5, \rho_6, \rho_7 \in S_N$ satisfy

$$||\rho_5|| + ||\rho_6|| \leqslant Ck^5 \max_t ||\partial_t^5 u^N||, \quad ||\rho_7||_j \leqslant Ck^4 \max_t ||\partial_t^4 u^N||_j. \tag{4.49}$$

Now we equate equal-power terms in (4.47). It is evident that the $\mathcal{O}(1)$ terms give an identity. It is also straightforward to see that we get identities for the $\mathcal{O}(k)$ and $\mathcal{O}(k^2)$ terms, using the facts that $b_1 + b_2 + b_3 = 1, b_1 = b_3$, and (3.10) and its temporal derivative, respectively. We now proceed to prove that we have identities for the $\mathcal{O}(k^3)$ and $\mathcal{O}(k^4)$ terms as well.

In order to show the identity of the $\mathcal{O}(k^3)$ terms we note that from (4.47), using (4.37), (4.22) and a Taylor expansion up to $\mathcal{O}(k^2)$ terms in the first nonlinear term in the right-hand side of (4.47), we have to check whether

$$\frac{1}{6}u^N_{ttt} = \frac{1}{6}(b_1 + b_2)^3 u^N_{ttt} - \frac{b_1^3}{12}u^N_{ttt} - \frac{b_1^3}{4}P_N(u^N_t u^N_{xt}) - \frac{b_3}{4}((b_1 + b_2)^2 + 1)\partial_x^3 u^N_{tt}$$
$$- \frac{b_3}{4}P_N\Big( ((b_1 + b_2)^2 + 1)(u^N u^N_{xtt} + u^N_{tt} u^N_x) + (b_1 + b_2 + 1)^2 u^N_t u^N_{tx} \Big).$$

If we differentiate (3.10) twice with respect to $t$, use Leibniz's rule on the nonlinear terms and insert the resulting expression for the term $\partial_x^3 u^N_{tt}$ in the right-hand side of the above we may verify, using the relations $b_1 = b_3, b_1 + b_2 + b_3 = 1$ and some algebra, that it is indeed an identity.

In order to show the identity of the $\mathcal{O}(k^4)$ terms we note, using (4.40), (4.24), (4.37) and Taylor expansions in the first and second nonlinear terms on the right-hand side of (4.47), that we must verify whether

$$\frac{1}{24}\partial_t^4 u^N = \frac{1}{24}(b_1 + b_2)^4 \partial_t^4 u^N - \frac{1}{12}b_1^3 \partial_t^4 u^N + \frac{b_1^4}{12}\partial_t^4 u^N$$
$$+ \frac{b_1^4}{8}\Big( - P_N \partial_x^3 (u^N_t u^N_{xt}) + 2P_N(u^N_t u^N_{xtt} + u^N_{tt} u^N_{xt})$$
$$- P_N \Big( u^N P_N(u^N_t u^N_{xt}) )_x \Big) \Big) + \frac{b_2^3}{8}P_N(u^N_t u^N_{ttx} + u^N_{tt} u^N_{tx})$$

$$
\begin{aligned}
&- \frac{b_3}{12}((b_1+b_2)^3+1)\partial_x^3 u^N_{ttt} + \frac{b_3 b_1^3}{2}\left(\frac{1}{12}\partial_x^3\partial_t^3 u^N + \frac{1}{4}\partial_x^3 P_N(u^N_t u^N_{xt})\right)\\
&- \frac{b_3}{4}P_N\left(\frac{1}{3}((b_1+b_2)^3+1)(u^N u^N_{xttt} + u^N_{ttt}u^N_x)\right.\\
&+ \frac{1}{2}\left(1+(b_1+b_2)+(b_1+b_2)^2+(b_1+b_2)^3\right)(u^N_t u^N_{xtt} + u^N_{tt}u^N_{xt})\Bigg)\\
&+ \frac{b_3}{2}P_N\left(\frac{b_1^3}{12}(u^N u^N_{ttt})_x + \frac{b_1^3}{4}\left(u^N P_N(u^N_t u^N_{xt})\right)_x\right).
\end{aligned}
\tag{4.50}
$$

Let $L$ be the sum of the linear terms and $N_1$ the sum of the nonlinear terms on the right-hand side of (4.50). Then $L = \gamma_1\partial_t^4 u^N + \gamma_2\partial_x^3 u^N_{ttt}$, where $\gamma_1 = \frac{1}{24}(b_1+b_2)^4 - \frac{b_1^3}{12} + \frac{b_1^4}{12}$, $\gamma_2 = -\frac{b_3}{12}((b_1+b_2)^3+1) + \frac{b_1^3 b_3}{24}$. Now we differentiate (3.10) three times with respect to $t$ and use the result in the definition of $L$. Taking into account that $b_1 = b_3, b_2 = 1-2b_1$ we see, after some algebra, that $\gamma_1 - \gamma_2 = \frac{1}{24}$, which implies that $L = \frac{1}{24}\partial_t^4 u^N - \gamma_2 P_N(u^N u^N_x)_{ttt}$, i.e. that $L$ has now acquired a nonlinear term as a result of eliminating $\partial_x^3 u^N_{ttt}$. Hence the coefficients of $\partial_t^4 u^N$ on the two sides of (4.50) match, and, therefore, in order to show that (4.50) holds, we have to check the validity of the identity

$$
N_1 - \gamma_2 P_N(u^N u^N_x)_{ttt} = 0. \tag{4.51}
$$

From (4.50) we have $N_1 - \gamma_2 P_N(u^N u^N_x)_{ttt} = P_N\mathcal{G}$, where

$$
\begin{aligned}
\mathcal{G} = &-\frac{b_1^4}{8}\partial_x^3(u^N_t u^N_{xt}) + \frac{b_1^4}{4}(u^N_t u^N_{xtt} + u^N_{tt}u^N_{xt}) - \frac{b_1^4}{8}\left(u^N P_N(u^N_t u^N_{xt})\right)_x\\
&+ \frac{b_2^3}{8}(u^N_t u^N_{xtt} + u^N_{tt}u^N_{xt}) + \frac{b_1^3 b_3}{8}\partial_x^3(u^N_t u^N_{xt}) - \frac{b_3}{12}((b_1+b_2)^3+1)(u^N u^N_{xttt} + u^N_{ttt}u^N_x)\\
&- \frac{b_3}{8}\left(1+(b_1+b_2)+(b_1+b_2)^2+(b_1+b_2)^3\right)(u^N_t u^N_{xtt} + u^N_{tt}u^N_{xt})\\
&+ \frac{b_1^3 b_3}{24}(u^N u^N_{ttt})_x + \frac{b_1^3 b_3}{8}\left(u^N P_N(u^N_t u^N_{xt})\right)_x - \gamma_2(u^N u^N_x)_{ttt}.
\end{aligned}
$$

Since $b_1 = b_3$ we see that the strange terms $\partial_x^3(u^N_t u^N_{xt})$ and $\left(u^N P_N(u^N_t u^N_{xt})\right)_x$ cancel, and we are left (after some algebra and the application of Leibniz's rule) simply with $\mathcal{G} = \gamma_3(u^N u^N_{xttt} + u^N_{ttt}u^N_x) + \gamma_4(u^N_t u^N_{xtt} + u^N_{tt}u^N_{xt})$, where

$$
\begin{aligned}
\gamma_3 &= -\frac{b_3}{12}((b_1+b_2)^3+1) + \frac{b_1^3 b_3}{24} - \gamma_2,\\
\gamma_4 &= \frac{b_1^4}{4} + \frac{b_2^3}{8} - \frac{b_3}{8}\left(1+(b_1+b_2)+(b_1+b_2)^2+(b_1+b_2)^3\right) - 3\gamma_2.
\end{aligned}
$$

From the definition of $\gamma_2$, we see that $\gamma_3 = 0$. Finally, after some algebra and using the facts $b_1 = b_3, b_1 + b_2 + b_3 = 1$ and that $x = b_1$ satisfies the cubic equation $x^3 - 2x^2 + x - 1/6 = 0$, we get $\gamma_4 = 0$. Therefore, $\mathcal{G} = 0$ and (4.51), and consequently (4.50), hold.

We now embark upon finding an equation for $e^{n,3}$ from the remaining terms in (4.47). We recall that we have used Taylor expansions in the first two terms of the nonlinear $-\frac{kb_3}{4}P_N(\cdots)$ term on the right-hand side of (4.47) and now we have to put in the residuals. In this way (4.47) becomes

$$e^{n,3} + \frac{kb_3}{2}\partial_x^3 e^{n,3} = \varGamma_3 + e^{n,2} - \frac{kb_3}{2}\partial_x^3 e^{n,2} - \frac{kb_3}{4}P_N\mathcal{E}(e^{n,2}, e^{n,3}), \tag{4.52}$$

where

$$\varGamma_3 = -\rho_5 + \rho_6 - \frac{kb_3}{2}\partial_x^3\rho_7 - \frac{k^5 b_3}{2}\partial_x^3 B_2 - \frac{kb_3}{4}P_N\Big(\rho_8 + k^3(\rho_9 B_1)_x + k^4\Big((u^N(\tau^{n,2}) + u^N(t^{n+1}))B_2\Big)_x + k^6 B_1 B_{1x} + k^7(B_1B_2)_x + k^8(B_2B_{2x})\Big), \tag{4.53}$$

and where the quantities $\rho_8, \rho_9$ satisfy, in view of (3.8),

$$||\rho_8|| \leqslant Ck^4, \quad ||\rho_9||_1 \leqslant Ck, \tag{4.54}$$

for $\mu$ sufficiently large. (Recall that we have estimates for $\rho_5, \rho_6, \rho_7$, cf. (4.49), and that $\mathcal{E}$ was defined in (4.48).)

We will prove below that for $\mu$ sufficiently large there is a constant $C$, independent of $k$ and $N$, such that

$$||\varGamma_3|| \leqslant Ck^5. \tag{4.55}$$

Assuming (4.55) for the time being and taking $L^2$ inner products with $e^{n,3} \in S_N$ in (4.52) we see, using integration by parts, that, in view of (4.48),

$$\begin{aligned}
||e^{n,3}||^2 = {} & (\varGamma_3, e^{n,3}) + (e^{n,2}, e^{n,3}) - \frac{kb_3}{2}(\partial_x^3 e^{n,2}, e^{n,3}) \\
& \underbrace{- \frac{kb_3}{4}\Big(\Big((u^N(\tau^{n,2}) + u^N(t^{n+1}))(e^{n,2} + e^{n,3})\Big)_x, e^{n,3}\Big)}_{\text{I}} \\
& \underbrace{- \frac{k^4 b_3}{4}\Big(\Big(B_1(e^{n,2} + e^{n,3})\Big)_x, e^{n,3}\Big)}_{\text{II}} \underbrace{- \frac{k^5 b_3}{4}\Big(\Big(B_2(e^{n,2} + e^{n,3})\Big)_x, e^{n,3}\Big)}_{\text{III}} \\
& \underbrace{- \frac{kb_3}{4}\Big((e^{n,2} + e^{n,3})(e^{n,2} + e^{n,3})_x, e^{n,3}\Big)}_{\text{IV}}.
\end{aligned} \tag{4.56}$$

Using integration by parts, (2.4), (4.29), (4.37), (4.40), (4.41), (4.45), our hypothesis that $kN = \mathcal{O}(1)$ and taking $\mu$ sufficiently large we may prove that

$$|\mathrm{I}| + |\mathrm{II}| + |\mathrm{III}| + |\mathrm{IV}| \leqslant Ck^5||e^{n,3}|| + Ck||e^{n,3}||^2.$$

Therefore, by the above estimate, (4.56), (4.55), (4.45), (4.46), our hypothesis that $kN = \mathcal{O}(1)$ and taking $\mu$ sufficiently large and $k$ sufficiently small we finally obtain

$$||e^{n,3}|| \leqslant Ck^5. \tag{4.57}$$

To conclude the proof we have to check (4.55). This is not hard to verify, in view of (4.53), (4.49), (4.24), (4.40), (4.37), (3.8), assuming as usual that $\mu$ is sufficiently large. Therefore, by (4.57) and (4.16), the proof of Proposition 4.4 is now complete. □

## 5. Error estimate for the fully discrete scheme

In this section we will consider again the fully discrete scheme given by (4.9) or (4.11) and corresponding to the temporal discretization of (2.5) by the general IMR-based, $s$-stage RK-composition method given by (4.4) or (4.5) and prove, under certain conditions on the discretization parameters and provided the solution of (1.1) belongs to $H^\mu$ for $\mu$ sufficiently large, that it has a unique solution $U^n$ satisfying the $L^2$ error estimate

$$\max_{0\leqslant n\leqslant M} ||U^n - u(t^n)|| \leqslant C\left(k^\alpha + N^{1-\mu}\right),$$

where $\alpha$ is a positive integer, such that the local temporal error, defined by an analogous formula to (4.13), is $\mathcal{O}(k^{\alpha+1})$ in $L^2$. (In Section 4.4 we considered the special case corresponding to $s = 3$ and constants $b_i$ given by (4.6) and proved that for that scheme $\alpha$ was equal to 4 provided $\mu$ was sufficiently large and $kN = \mathcal{O}(1)$.)

We first establish notation and present a summary of the main steps of the proof. For convenience in referencing we rewrite scheme (4.11) here. We seek $Y^{n,i}, 0 \leqslant i \leqslant s$ and $U^n$, $0 \leqslant n \leqslant M$ in $S_N$, such that for $0 \leqslant n \leqslant M-1$,

$$Y^{n,0} = U^n,\ Y^{n,i} = Y^{n,i-1} + kb_iF\left(\frac{Y^{n,i} + Y^{n,i-1}}{2}\right),\ 1 \leqslant i \leqslant s,\ U^{n+1} = Y^{n,s}, \tag{5.1}$$

and $U^0 = P_N u_0$. (Recall that for $v \in S_N, F(v) = -v_{xxx} - P_N f(v)_x$, where $f(v) = v^2/2$.) As in Section 4.4 we define the local temporal error of scheme (5.1) in terms of the semidiscrete approximation $u^N$. For this purpose we write for $0 \leqslant n \leqslant M$, $V^n = u^N(t^n)$ and define $V^{n,i} \in S_N$ for $0 \leqslant i \leqslant s, 0 \leqslant n \leqslant M-1$, by

$$V^{n,0} = V^n, \quad V^{n,i} = V^{n,i-1} + kb_iF\left(\frac{V^{n,i} + V^{n,i-1}}{2}\right), \quad 1 \leqslant i \leqslant s, \tag{5.2}$$

and the local temporal error $\theta^n \in S_N, 0 \leqslant n \leqslant M-1$ as

$$\theta^n = V^{n+1} - V^{n,s} \equiv u^N(t^{n+1}) - V^{n,s}. \tag{5.3}$$

(We use the same notation for $V^n, V^{n,i}, \theta^n$ as in Section 4.4 as no confusion will arise.) For the local error we will assume that

$$\max_{0\leqslant n\leqslant M-1} ||\theta^n|| \leqslant Ck^{\alpha+1}. \tag{5.4}$$

We let $\epsilon^n = V^n - U^n \equiv u^N(t^n) - U^n$. Our aim will be to prove that $\max_n ||\epsilon^n|| = \mathcal{O}(k^\alpha)$, which, together with (3.1), will give the desired error estimate

$$\max_{0\leqslant n\leqslant M} ||U^n - u(t^n)|| \leqslant C\left(k^\alpha + N^{1-\mu}\right).$$

We also let $\epsilon^{n,i} = V^{n,i} - Y^{n,i}, 1 \leqslant i \leqslant s$, and note, in view of (5.3), (5.1), that $\epsilon^{n+1} = V^{n+1} - U^{n+1} = \theta^n + V^{n,s} - Y^{n,s} = \theta^n + \epsilon^{n,s}$. Obviously, cf. Section 4.3, the $Y^{n,i}$ and $V^{n,i}$ exist and satisfy for all $n$ and $1 \leqslant i \leqslant s$ the $L^2$-conservation laws

$$||Y^{n,i}|| = ||U^n|| = ||U^0||, \quad ||V^{n,i}|| = ||V^n|| = ||u^N(t^n)|| = ||u^N(0)||. \tag{5.5}$$

In order to bound the $\epsilon^{n,i}$ and $\epsilon^n$ in $L^2$, the $L^2$ bounds of $V^{n,i}$ in (5.5) are not enough. So we first establish in Lemma 5.1 a bound for $||V^{n,i}||_{1,\infty}$ uniformly in $n$ and $i$. The proof of Lemma 5.3 follows easily; in it we show that $\max_n ||\epsilon^n|| \leqslant Ck^\alpha$ after establishing estimates of the form

$$\max_i ||\epsilon^{n,i}|| \leqslant (1 + Ck)||\epsilon^n||.$$

Finally, in Theorem 5.4, we prove the uniqueness of the fully discrete approximations $U^n, Y^{n,i}$ and the final error estimate.

Lemma 5.1 Let $V^{n,i}$ be defined by (5.2). Suppose that $\mu$ is sufficiently large, $k$ is sufficiently small and that $k = \mathcal{O}(N^{-1/2})$. Then

$$\max_{i,n} ||V^{n,i}||_{1,\infty} \leqslant C. \tag{5.6}$$

Remark 5.2 Here and in the sequel we let $\tau^{n,i} = t^n + k(b_1 + b_2 + \cdots + b_i), 1 \leqslant i \leqslant s$, so that $\tau^{n,s} = t^{n+1}$. Since some of the $b_i$ may be negative, and some $\tau^{n,i}$ may exceed $t^{n+1}$, as was remarked in the course of the proof of Proposition 4.4, it may be necessary to extend the well-posedness of (1.1) and the validity of (2.5) in temporal intervals of the form $[-l_1k, T + l_2k]$ for small non-negative integers $l_1, l_2$. In such temporal intervals the bounds in (3.1) and (3.8) obviously hold.

*Proof.* We break the proof into three steps for ease of reading it.

(i) First prove that $\max_n ||V^{n,1}||_{1,\infty} \leqslant C$.

We will show that $V^{n,1}$ is close to $u^N(\tau^{n,1})$, specifically to $\mathcal{O}(k^3)$ in $L^2$, and then use (2.4) and (3.8) to prove the desired bound. For this purpose we first need the following consistency result for one step of length $kb_1$ for the IMR scheme. (As before we denote the values of $u^N$ and its derivatives at $t^n$ simply by $u^N, u_t^N$, etc.)

Define $\zeta^{n,1} \in S_N$ by the equation

$$u^N(\tau^{n,1}) = u^N + kb_1 F\left(\frac{u^N(\tau^{n,1}) + u^N}{2}\right) + \zeta^{n,1}. \tag{5.7}$$

Then, it may be easily seen, cf. e.g. Dougalis & Durán (2020), that for $\mu$ sufficiently large,

$$\max_n ||\zeta^{n,1}|| \leqslant Ck^3. \tag{5.8}$$

We now proceed to bound $V^{n,1}$ in the $||\cdot||_{1,\infty}$ norm. By (5.2) for $i = 1$ and (5.7) (since $V^n = u^N$), we obtain

$$V^{n,1} - u^N(\tau^{n,1}) = kb_1\left(F\left(\frac{V^{n,1} + u^N}{2}\right) - F\left(\frac{u^N(\tau^{n,1}) + u^N}{2}\right)\right) - \zeta^{n,1}.$$

Therefore, by integration by parts, we see that

$$\begin{aligned} ||V^{n,1} - u^N(\tau^{n,1})||^2 = kb_1 &\left(f\left(\frac{V^{n,1} + u^N}{2}\right)_x - f\left(\frac{u^N(\tau^{n,1}) + u^N}{2}\right)_x, V^{n,1} - u^N(\tau^{n,1})\right) \\ &- \left(\zeta^{n,1}, V^{n,1} - u^N(\tau^{n,1})\right). \end{aligned} \tag{5.9}$$

Now, by integration by parts, and (3.8), for $\mu$ sufficiently large, we see that

$$\begin{aligned} &\left|\left(f\left(\frac{V^{n,1} + u^N}{2}\right)_x - f\left(\frac{u^N(\tau^{n,1}) + u^N}{2}\right)_x, V^{n,1} - u^N(\tau^{n,1})\right)\right| \\ &= \left|\left(f\left(\frac{u^N(\tau^{n,1}) + u^N}{2} + \frac{V^{n,1} - u^N(\tau^{n,1})}{2}\right)_x - f\left(\frac{u^N(\tau^{n,1}) + u^N}{2}\right)_x, V^{n,1} - u^N(\tau^{n,1})\right)\right| \\ &= \left|\left(\left[\left(\frac{u^N(\tau^{n,1}) + u^N}{2}\right)\left(\frac{V^{n,1} - u^N(\tau^{n,1})}{2}\right)\right]_x, V^{n,1} - u^N(\tau^{n,1})\right)\right| \\ &\leqslant C||u^N(\tau^{n,1}) + u^N||_{1,\infty}||V^{n,1} - u^N(\tau^{n,1})||^2 \leqslant C||V^{n,1} - u^N(\tau^{n,1})||^2. \end{aligned}$$

We conclude by (5.9), and (5.8), for $k$ sufficiently small, that

$$||V^{n,1} - u^N(\tau^{n,1})|| \leqslant Ck^3. \tag{5.10}$$

Therefore, by the above, (2.4) and (3.8), for $\mu$ sufficiently large we get

$$||V^{n,1}||_{1,\infty} \leqslant ||V^{n,1} - u^N(\tau^{n,1})||_{1,\infty} + ||u^N(\tau^{n,1})||_{1,\infty} \leqslant Ck^3N^{3/2} + C \leqslant C, \tag{5.11}$$

using the mesh condition $k = \mathcal{O}\left(N^{-1/2}\right)$.

(ii) Now prove that $\max_n ||V^{n,2}||_{1,\infty} \leqslant C$.

We will follow the same general plan as in (i). We let $\zeta^{n,2}$ be the local temporal error of the scheme during the substep $\tau^{n,1} \mapsto \tau^{n,2}$, i.e. define it by the equation

$$u^N(\tau^{n,2}) = u^N(\tau^{n,1}) + kb_2 F\left(\frac{u^N(\tau^{n,2}) + u^N(\tau^{n,1})}{2}\right) + \zeta^{n,2}. \tag{5.12}$$

Then, we may prove as in (i), *mutatis mutandis* that

$$\max_n ||\zeta^{n,2}|| \leqslant Ck^3. \tag{5.13}$$

By (5.2) for $i = 2$ and (5.12) we have

$$\begin{aligned} V^{n,2} - u^N(\tau^{n,2}) &= V^{n,1} - u^N(\tau^{n,1}) \\ &+ kb_2\left(F\left(\frac{V^{n,2}+V^{n,1}}{2}\right) - F\left(\frac{u^N(\tau^{n,2}) + u^N(\tau^{n,1})}{2}\right)\right) - \zeta^{n,2}. \end{aligned} \tag{5.14}$$

In order to simplify the algebra a little, we define $\chi_j \in S_N, 1 \leqslant j \leqslant 4$ as $\chi_1 = V^{n,2} - u^N(\tau^{n,2})$, $\chi_2 = V^{n,1} - u^N(\tau^{n,1}),\ \chi_3 = (V^{n,2} + V^{n,1})/2,\ \chi_4 = (u^N(\tau^{n,2}) + u^N(\tau^{n,1}))/2$. Then (5.14) is written as

$$\chi_1 - \chi_2 = kb_2\left(F(\chi_3) - F(\chi_4)\right) - \zeta^{n,2}.$$

Take $L^2$ inner products in the above with $\frac{\chi_1+\chi_2}{2}$, noting that $\frac{\chi_1+\chi_2}{2} = \chi_3 - \chi_4$ and using integration by parts, and get

$$\frac{1}{2}\left(||\chi_1||^2 - ||\chi_2||^2\right) = kb_2\left(f\left(\chi_4 + \frac{\chi_1+\chi_2}{2}\right)_x - f(\chi_4)_x, \frac{\chi_1+\chi_2}{2}\right) - \left(\zeta^{n,2}, \frac{\chi_1+\chi_2}{2}\right). \tag{5.15}$$

Now, by integration by parts and (3.8), for $\mu$ sufficiently large we see that

$$\begin{aligned} \left|\left(f\left(\chi_4 + \frac{\chi_1+\chi_2}{2}\right)_x - f(\chi_4)_x, \frac{\chi_1+\chi_2}{2}\right)\right| &= \left|\left(\left(\chi_4\left(\frac{\chi_1+\chi_2}{2}\right)\right)_x, \frac{\chi_1+\chi_2}{2}\right)\right| \\ &\leqslant C||\chi_4||_{1,\infty}||\chi_1+\chi_2||^2 \leqslant C||\chi_1+\chi_2||^2, \end{aligned}$$

and (5.13), (5.15) yield $\frac{1}{2}\left(||\chi_1||^2 - ||\chi_2||^2\right) \leqslant Ck\left(||\chi_1|| + ||\chi_2||\right)^2 + Ck^3\left(||\chi_1|| + ||\chi_2||\right)$, i.e. $||\chi_1|| - ||\chi_2|| \leqslant Ck\left(||\chi_1|| + ||\chi_2||\right) + Ck^3$, from which, if we recall the definition of $\chi_1$ and $\chi_2$, it follows for $k$ sufficiently small, that $||V^{n,2} - u^N(\tau^{n,2})|| \leqslant C||V^{n,1} - u^N(\tau^{n,1})|| + Ck^3$. Therefore, by (5.10),

$$||V^{n,2} - u^N(\tau^{n,2})|| \leqslant Ck^3, \tag{5.16}$$

from which, as in the derivation of (5.11), we get, for $\mu$ sufficiently large, since $k = \mathcal{O}(N^{-1/2})$, that

$$||V^{n,2}||_{1,\infty} \leqslant C.$$

Hence, the proof of (ii) is complete.

(iii) Prove that $||V^{n,i}||_{1,\infty} \leqslant C, 3 \leqslant i \leqslant s$.

The bounds $||V^{n,i} - u^N(\tau^{n,i})|| \leqslant Ck^3$, implying that $||V^{n,i}||_{1,\infty} \leqslant C, 3 \leqslant i \leqslant s$, are obtained entirely analogously, as in step (ii) above, and their proof is omitted. We conclude that (5.6) holds. □

LEMMA 5.3 Let $\epsilon^n = V^n - U^n$, where $V^n = u^N$, and $U^n$ is the fully discrete approximation, defined by (5.1). We assume the smoothness of $u$ and the mesh condition stated in Lemma 5.1 and we suppose that the temporal local error estimate (5.4) holds. Then

$$\max_n ||\epsilon^n|| \leqslant Ck^\alpha. \tag{5.17}$$

*Proof.* We use throughout the notation introduced at the beginning of the section. We first estimate $\epsilon^{n,1} = V^{n,1} - Y^{n,1}$ in terms of $\epsilon^n$. Since

$$\epsilon^{n,1} - \epsilon^n = kb_1 \left( F\left(\frac{V^{n,1} + V^n}{2}\right) - F\left(\frac{V^{n,1} + V^n}{2} - \frac{\epsilon^{n,1} + \epsilon^n}{2}\right)\right),$$

taking $L^2$ inner products in this equation with $\frac{\epsilon^{n,1}+\epsilon^n}{2}$ we obtain, by integration by parts,

$$\frac{1}{2}\left(||\epsilon^{n,1}||^2 - ||\epsilon^n||^2\right) = -kb_1 \left( f\left(\frac{V^{n,1} + V^n}{2} - \frac{\epsilon^{n,1} + \epsilon^n}{2}\right)_x - f\left(\frac{V^{n,1} + V^n}{2}\right)_x, \frac{\epsilon^{n,1} + \epsilon^n}{2}\right).$$

Therefore, using integration by parts again, we get $||\epsilon^{n,1}||^2 - ||\epsilon^n||^2 \leqslant Ck||V^{n,1} + V^n||_{1,\infty}||\epsilon^{n,1} + \epsilon^n||^2$, from which, taking into account (5.6) and (3.8), it follows that, for all $n$, $||\epsilon^{n,1}|| - ||\epsilon^n|| \leqslant Ck\left(||\epsilon^{n,1}|| + ||\epsilon^n||\right)$. Hence, for $k$ sufficiently small, for all $n$ it holds that

$$||\epsilon^{n,1}|| \leqslant (1 + Ck)||\epsilon^n||. \tag{5.18}$$

We get similarly that

$$\max_i ||\epsilon^{n,i}|| \leqslant (1 + Ck)||\epsilon^n||. \tag{5.19}$$

This may be seen as follows: since in view of (5.6), as previously, it holds that $||\epsilon^{n,2}|| - ||\epsilon^{n,1}|| \leqslant Ck\left(||\epsilon^{n,2}|| + ||\epsilon^{n,1}||\right)$, we obtain by (5.19) that $||\epsilon^{n,2}|| \leqslant (1+Ck)||\epsilon^n||$. The general case (5.19) follows inductively.

Recall by (5.1) and (5.3) that $\epsilon^{n+1} = V^{n+1} - U^{n+1} = V^{n+1} - Y^{n,s} = V^{n,s} - Y^{n,s} + \theta^n = \epsilon^{n,s} + \theta^n$. Therefore, by (5.19), for all $n$ we have

$$||\epsilon^{n+1}|| \leqslant (1 + Ck)||\epsilon^n|| + ||\theta^n||,$$

from which, by the discrete Gronwall inequality, since $\epsilon^0 = 0$, and the hypothesis (5.4), we conclude that (5.17) holds. □

We now state and prove the main error estimate for our fully discrete method.

THEOREM 5.4 Suppose that $\mu$ is sufficiently large, that (5.4) holds for some $\alpha \geqslant 1$ and that $kN$ is sufficiently small. Then the fully discrete scheme (5.1) has for all $n$ a unique solution $U^n$ such that

$$\max_n ||U^n - u(t^n)|| \leqslant C(k^\alpha + N^{1-\mu}). \tag{5.20}$$

*Proof.* Let $U^n$ be a solution of (5.1). Then

$$||U^n - u(t^n)|| \leqslant ||U^n - u^N(t^n)|| + ||u^N(t^n) - u(t^n)|| = ||\epsilon^n|| + ||u^N(t^n) - u(t^n)||,$$

and (5.20) follows from (5.18) and (3.1).

In order to prove the uniqueness of $U^n$ we have to verify the hypotheses of Lemma 4.3. Note that it follows from (5.17), (3.8), (2.4) and our mesh condition that for all $n$, $|U^n|_\infty \leqslant |\epsilon^n|_\infty + |u^N|_\infty \leqslant Ck^\alpha N^{1/2} + C \leqslant R_1$, for some constant $R_1$, independent of $k, N$. In addition, for all $i$ and $n$, by (2.4), (5.6), (5.19), (5.17), and our mesh condition, we see that

$$|Y^{n,i}|_\infty \leqslant |\epsilon^{n,i}|_\infty + |V^{n,i}|_\infty \leqslant CN^{1/2}||\epsilon^{n,i}|| + C \leqslant CN^{1/2}||\epsilon^n|| + C \leqslant CN^{1/2}k^\alpha + C \leqslant R_2$$

for some constant $R_2$ independent of $k$ and $N$. If $R$ is taken as $\max(R_1, R_2)$, by Lemma 4.3 we have uniqueness of $U^n = Y^{n,s}$ if $kN$ is sufficiently small as we have assumed. □

REMARK 5.5 Since the fully discrete scheme (5.1) is written as a sequence of IMR steps, its implementation is quite straightforward, as the attendant nonlinear systems are decoupled and may each be solved by an iterative scheme.

Indeed, suppose that, for some $n$ and $i \geqslant 1$, $Y^{n,i-1}$ is known. Then if $Z^* = \frac{1}{2}\left(Y^{n,i-1} + Y^{n,i}\right)$ it follows that $Z^* \in S_N$ satisfies

$$Z^* = Y^{n,i-1} + \frac{kb_i}{2}F(Z^*). \tag{5.21}$$

Suppose that the hypotheses of Theorem 5.4 hold. Then $Z^*$ is unique, and if it is known, $Y^{n,i}$ may be computed as $Y^{n,i} = 2Z^* - Y^{n,i-1}$.

In order to approximate $Z^*$, consider the following simple iterative scheme. For $\nu = 0, 1, 2, \ldots$, seek $Z_\nu \in S_N$, such that

$$\begin{aligned} Z_0 &= Y^{n,i-1}, \\ \left(I + \frac{kb_i}{2}\partial_x^3\right) Z_{\nu+1} &= Y^{n,i-1} - \frac{kb_i}{2} f(Z_\nu)_x, \quad \nu = 0, 1, 2, \ldots. \end{aligned} \tag{5.22}$$

Given $Z_\nu$, the next iteration $Z_{\nu+1}$ satisfies a linear system of equations. The associated homogeneous system clearly has only the trivial solution; hence $Z_{\nu+1}$ is uniquely defined and its Fourier coefficients may be readily computed. One may prove by a straightforward argument, see Dougalis & Durán (2020) for details, that $Z_\nu$ converges to $Z^*$ as $\nu \to \infty$ in $L^2$, and that if $\nu = \mathcal{O}(|\log k|)$ the error $||Z^* - Z_\nu||$ may be bounded by a constant times a sufficiently large power of $k$.

## 6. Numerical experiments

In this section we present the results of numerical experiments we performed in order to illustrate the convergence and stability properties and check the efficiency of the fully discrete scheme that corresponds to the parameters (4.6). The experiments concern long-time computations since we would like to emphasize that the high order of accuracy, the stability and symplectic character of the time-stepping scheme make it suitable for accurately approximating long-time properties of solutions of equations such as the KdV.

To this end we study the accuracy of the scheme in approximating the solitary-wave solution of the IVP of the KdV, given by $u_c(x,t) = \phi_c(x-ct)$, where $c > 0$ and

$$\phi_c(z) = 3c\,\mathrm{sech}^2\left(\frac{\sqrt{c}}{2}z\right). \tag{6.1}$$

It is well known that the solitary wave (6.1) is the uniform limit of periodic solutions of cnoidal type of the KdV. This fact and the exponential decay of (6.1) as $|z| \to \infty$ allow one to perform the numerical approximation by integrating the periodic IVP of the KdV on a long enough interval $(-L, L)$ with (6.1) as initial condition.

The spectral method for the spatial discretization is implemented in collocation form; cf. e.g. Maday & Quarteroni (1988). For $T > 0$ and an integer $N \geqslant 1$ the semidiscrete solution is defined as a mapping $u_h : [0,T] \to S_N$ satisfying the KdV at a uniform grid of spatial collocation points $x_j = -L + jh, j = 0, \ldots, N-1$, where $h = 2L/N$. The approximation $u_h$ is represented by the nodal values

$$U_h(t) = (u_h(x_0,t), \ldots, u_h(x_{N-1},t))^{\mathrm{T}},$$

where the vector $U_h$ satisfies the semidiscrete system

$$\frac{\mathrm{d}}{\mathrm{d}t}U_h + D_N\left(\frac{U_h^2}{2}\right) + D_N^3 U_h = 0,\ 0 \leqslant t \leqslant T, \tag{6.2}$$

where $D_N$ denotes the $N \times N$ Fourier pseudospectral differentiation matrix (scaled to the interval $[-L, L]$) and the product in the nonlinear term in (6.2) is understood in the Hadamard sense.

We mention now some properties of the ODE system (6.2) that will be used in the experiments in the sequel. Similar arguments to those in Cano (2006) prove that (6.2) is Hamiltonian with respect to the symplectic structure given by $-D_N$, and the Hamiltonian is given by

$$H_h(U) = ||D_N U^n||_N^2 - \frac{1}{3}\sum_{j=0}^{N-1}(U_j^n)^3, \tag{6.3}$$

TABLE 1 *$L^2$ and $L^\infty$ errors and temporal convergence rates. Solitary-wave solution (6.1) with $c = 1$, $T = 10^4$, $N = 4096$*

| $k$ | $L^\infty$ error | Rate | $L^2$ error | Rate |
|---|---|---|---|---|
| $2.5 \times 10^{-2}$ | $2.4281 \times 10^{-4}$ | | $4.6054 \times 10^{-4}$ | |
| $1.25 \times 10^{-2}$ | $1.5115 \times 10^{-5}$ | 4.008 | $2.8661 \times 10^{-5}$ | 4.009 |
| $6.25 \times 10^{-3}$ | $9.3657 \times 10^{-7}$ | 4.017 | $1.7761 \times 10^{-6}$ | 4.017 |

for $U = (U_0, \ldots, U_{N-1})^{\mathrm{T}}$ and where $||\cdot||_N$ stands for the discrete $L^2$ norm in $\mathbb{R}^N$. Similarly, the quantity

$$I_h(U) = ||U||_N^2 \tag{6.4}$$

is preserved in time by those solutions $U_h(t)$ of (6.2) satisfying the symmetry condition

$$DU_h(t) = U_h(t), \quad t \geqslant 0,$$

where if $U = (U_0, \ldots, U_{N-1})^{\mathrm{T}}$ then $DU = (U_{N-1}, \ldots, U_0)^{\mathrm{T}}$; cf. Cano (2006).

The ODE system (6.2) is represented in the Fourier space (in order to make use of fast Fourier transform techniques for computing $D_N$), and integrated numerically in time by using the RK composition method (4.4) corresponding to $s = 3$ and parameters (4.6). For the experiments below we took $L = 64$ and $N = 4096$. These values ensured that errors due to the truncation to the interval $[-L, L]$ and to the spatial approximation are negligible. We implemented the fully discrete method by using the iterative scheme (5.22) for the intermediate stages (4.11). The experiments did not require more than two iterations per stage. Errors were measured with the discrete $L^2$ norm $||\cdot||_N$, defined above, and the $L^\infty$ norm

$$||U||_\infty = \max_{0 \leqslant j \leqslant N-1} |U_j|, \; U = (U_0, \ldots, U_{N-1})^{\mathrm{T}}.$$

We first compare the numerical solution at final time $T = 10^4$ with the exact solution $u_c$ of speed $c = 1$ for several time-step sizes. The errors in the $L^2$ and $L^\infty$ norms are displayed in Table 1, which shows, as expected, the fourth order of convergence of the temporal discretization.

In order to illustrate the benefits of the geometric numerical integration resulting from the additional properties of the scheme (cf. the introduction and references therein) we include here some numerical experiments of long-time simulation of the propagation of the solitary wave (6.1).

Figure 1 shows, in logarithmic scale, the temporal behaviour of the error in the $L^2$ norm for several values of the time step. The order of convergence can be checked in the figure by comparing the distance between lines corresponding to consecutive time steps. The observed linear in time growth of the error for larger values of $t$ is expected for this type of conservative scheme; cf. e.g. de Frutos & Sanz-Serna (1997).

The preservation of invariants is illustrated in Fig. 2. If $U^n = (U_0^n, \ldots, U_{N-1}^n)^{\mathrm{T}}$ denotes the approximation to the solution of (6.2) at $t = t_n = nk$, Fig. 2 shows the temporal behaviour of the errors $I(U^n) - I(U^0), H(U^n) - H(U^0)$, where

$$I(U) = hI_h(U), \quad H(U) = hH_h(U),$$

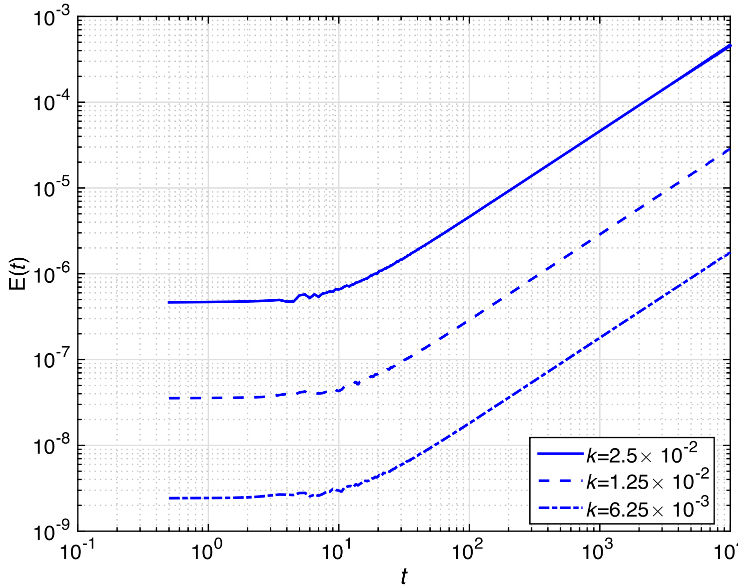


FIG. 1. $L^2$ Error w.r.t the solitary wave (6.1) as function of time (log-log scale).

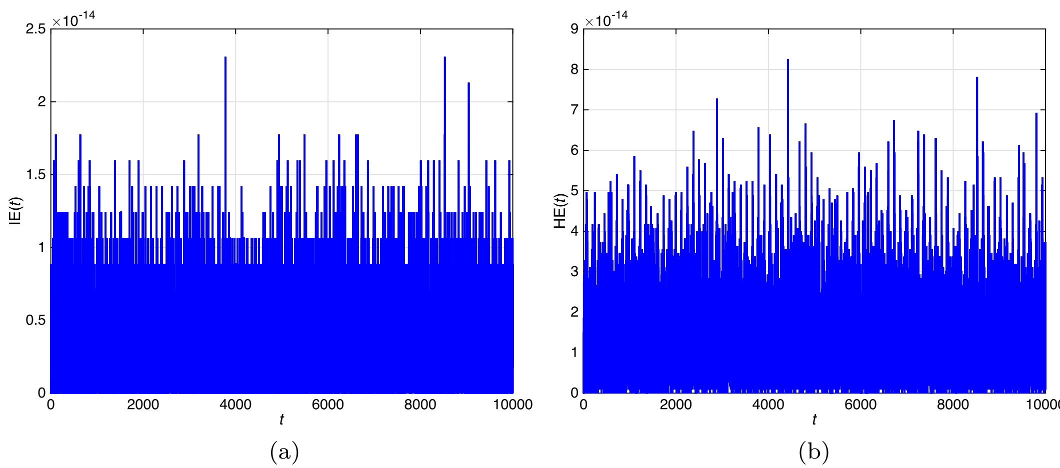


FIG. 2. Approximation to the solitary wave (6.1). Errors (a) $I(U^n) - I(U^0)$ and (b) $H(U^n) - H(U^0)$ as functions of time; $k = 1.25 \times 10^{-2}$.

and $I_h, H_h$ are given by (6.4) and (6.3), respectively. The preservation of $H$ is a consequence of the symplectic character of the time integration and the interpretation of the solitary wave solution (6.1) as relative equilibrium; cf. Cano (2006); Frutos & Sanz-Serna (1997).

The conservative character of the scheme also influences the approximation of the amplitude and phase of the solitary wave. Figure 3 shows the temporal evolution of the error in the amplitude and in the phase, computed in the usual way; cf. e.g. Dougalis *et al.* (2019). The growth of the error of the speed is similar to that of the amplitude shown in Fig. 3(a). On the other hand, the linear growth of the

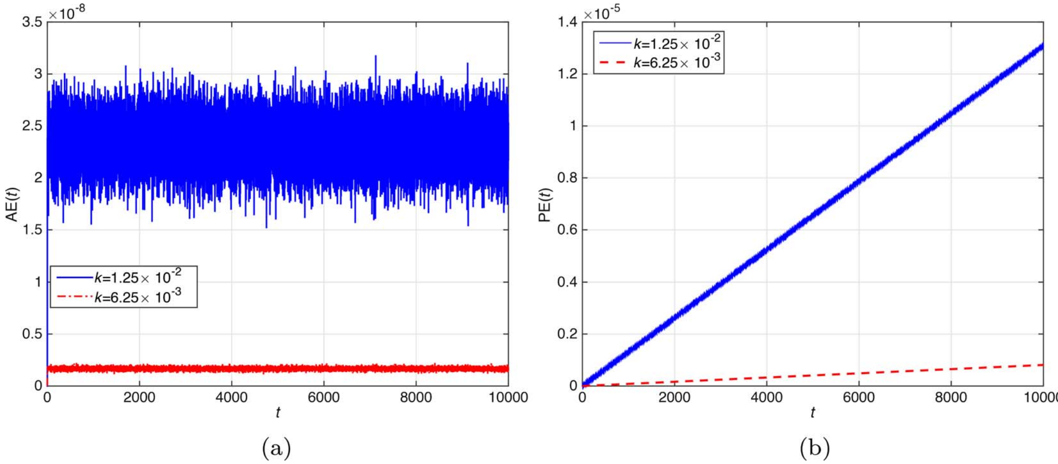


FIG. 3. Approximation to the solitary wave (6.1). Temporal evolution of the error in amplitude (a) and phase (b).

phase error, observed in Fig. 3(b), is responsible for similar behaviour of the leading term of the global error shown in Fig. 1.

It is to be noted that taking a larger time-step size $k$ for a fixed $N$ just increases the phase error in amplitude as is expected from the $L^2$-conservation of the scheme. The restriction on $kN$ mentioned in the hypotheses of the convergence Theorem 5.4 seems to be a theoretical artefact of the uniqueness proof (see Lemma 4.3) and not a stability restriction.

## 7. Conclusions and extensions

In this paper we analyzed a high-order accurate fully discrete scheme for the periodic IVP for the KdV equation. The problem was discretized in space by the standard Fourier–Galerkin spectral method. For the temporal discretization we used a diagonally implicit RK scheme of composition type with $s$ stages; cf. Hairer *et al.* (2004), Yoshida (1990), effected by $s$ steps of the IMR method. This type of scheme is not A-stable, but they are symplectic; hence they are unconditionally $L^2$-conservative for the periodic IVP and semidiscretization at hand. They are also easy to implement. We proved that the local temporal error of the scheme with $s = 3$ stages applied to the semidiscrete equations is $\mathcal{O}(k^5)$ in $L^2$, where $k$ is the time step, under the hypothesis that the solution of the periodic IVP belongs to the periodic Sobolev space $H^\mu$ for $\mu$ sufficiently large and that $k = \mathcal{O}(N^{-1})$, where $N$ is the order of the trigonometric polynomials used in the semidiscretization. We also proved that if $kN$ is sufficiently small the fully discrete scheme has a unique solution and satisfies an $L^2$ error estimate of $\mathcal{O}(k^\alpha + N^{1-\mu})$, provided the local temporal error is $\mathcal{O}(k^{\alpha+1})$ in $L^2$ and if $\mu$ is sufficiently large. So, for the particular scheme with $s = 3$ stages (first used in computations for solving the KdV in Frutos & Sanz-Serna, 1992), the resulting error estimate is $\mathcal{O}(k^4 + N^{1-\mu})$. These results are illustrated with some numerical experiments involving the simulation of the solitary wave solution of the KdV equation.

The RK scheme considered in this paper may be used to discretize, in the temporal variable, other conservative IVPs for PDEs that model one-way propagation of nonlinear dispersive waves. For example, in Dougalis *et al.* (2019), the three-stage, fourth-order accurate scheme, coupled with a spectral discretization in space, was used for approximating the solution of the periodic IVP for the generalized

Benjamin equation. This nonlinear PDE, introduced in Chen & Bona (1998) as a generalization of the KdV, the Benjamin–Ono and the Benjamin equation, is of the form

$$u_t - \mathcal{L}u_x + f(u)_x = 0, \tag{7.1}$$

where $\mathcal{L}$ is the linear, nonlocal, pseudodifferential operator with Fourier symbol

$$\widehat{\mathcal{L}u}(\xi) = l(\xi)\widehat{u}(\xi) = (\delta|\xi|^{2m} - \gamma|\xi|^{2r})\widehat{u}(\xi), \quad \xi \in \mathbb{R},$$

where $m \geqslant 1$ is an integer, $0 \leqslant r < m, \gamma \geqslant 0, \delta > 0, \widehat{u}(\xi)$ denotes the Fourier transform of $u$ at $\xi$ and the nonlinear term $f$ is given by

$$f(u) = \frac{u^{q+1}}{q+1},$$

with $q \geqslant 1$ integer. The Cauchy problem for (7.1) has been shown to be locally well posed in $H^s(\mathbb{R})$ for $s \geqslant 1$; see e.g. Linares & Scialom (2005), and globally well posed if $q = 2$ or 3. The equation possesses solitary-wave solutions, cf. Chen & Bona (1998); the numerical study in Dougalis *et al.* (2019) was focused on describing their generation, interactions and stability. In the case of the periodic IVP for (7.1) the standard Fourier–Galerkin semidiscrete approximation $u^N$ may be shown to possess an $L^2$ error estimate of the form $||u - u^N|| \leqslant CN^{1-\mu}$ if $u \in H^\mu, \mu \geqslant 5/2$, and satisfy (3.8) if $\mu$ is sufficiently large. If now we define, for $v \in S_N, F(v) = \mathcal{L}v_x - P_N f(v)_x$, so that $(F(v), v) = 0$ for $v \in S_N$, it may be seen that the proof of existence of solutions and of the $L^2$-conservation property of the fully discrete scheme proceed as in Section 4.3. The study of the local temporal error of the scheme with $s = 3$ stages may be along the lines of Proposition 4.4. An analog of Theorem 5.4 holds as well. It may be proved that the solution of the $s$-stage, fully discrete scheme is unique and satisfies (5.20) *mutatis mutandis*, under the assumptions that the solution of the periodic IVP is sufficiently smooth, and that $kN$ is sufficiently small if $q = 1, 2$ or 3, and $kN^{\frac{q-1}{2}}$ is sufficiently small if $q \geqslant 4$. The general plan of the proof is that of Theorem 5.4 but, as expected, considerable technical complications enter the picture due to the generalized nonlinear term.


## Acknowledgements

The authors would like to acknowledge travel support, which made this collaboration possible, from the Institute of Applied and Computational Mathematics of the Foundation for Research and Technology Hellas (IACM-FORTH) and the Institute of Mathematics of the University of Valladolid (IMUVA).

## Funding

Spanish Ministerio de Ciencia e Innovación (PID2020-113554GB-100); Junta de Castilla y León (VA193P20 to A.D.); Fondos Europeos de Desarrollo Regional FEDER (EU) (VA193P20 to A.D.).


## References


Antonopoulos, D. C., Dougalis, V. A. & Kounadis, G. (2020) On the standard Galerkin method with explicit RK4 time stepping for the shallow water equations. *IMA J. Numer. Anal.*, **40**, 2415–2449.

Baker, G. A., Dougalis, V. A. & Karakashian, O. A. (1983) Convergence of Galerkin approximations for the Korteweg–de Vries equation. *Math. Comp.*, **40**, 419–433.

Bona, J. L., Chen, H., Karakashian, O. A. & Xing, Y. (2013) Conservative, discontinuous Galerkin methods for the generalized Korteweg–de Vries equation. *Math. Comp.*, **82**, 1401–1432.

Bona, J. L., Dougalis, V. A., Karakashian, O. A. & McKinney, W. R. (1995) Conservative, high-order numerical schemes for the generalized Korteweg–de Vries equation. *Phil. Trans. R. Soc. London A*, **351**, 107–164.

Bona, J. L. & Grujic, Z. (2003) Spatial analyticity for nonlinear waves. *Math. Models Methods Appl. Sci.*, **13**, 1–15.

Bona, J. L. & Smith, R. (1975) The initial-value problem for the Korteweg–de Vries equation. *Phil. Trans. Roy. Soc. London A*, **278**, 555–601.

Cano, B. (2006) Conserved quantities of some Hamiltonian wave equations after full discretization. *Numer. Math.*, **103**, 197–223.

Chen, H. & Bona, J. L. (1998) Existence and asymptotic properties of solitary-wave solutions of Benjamin-type equations. *Adv. Diff. Eq.*, **3**, 51–84.

de Frutos, J. & Sanz-Serna, J. M. (1992) An easily implementable fourth-order method for the time integration of wave problems. *J. Comput. Phys.*, **103**, 160–168.

de Frutos, J. & Sanz-Serna, J. M. (1997) Accuracy and conservation properties in numerical integration: the case of the Korteweg–de Vries equation. *Numer. Math.*, **75**, 421–445.

Deng, Z.-G. & Ma, H.-P. (2009) Optimal error estimates for Fourier spectral approximation of the generalized KdV equation. *Appl. Math. Mech.-Engl. Ed.*, **30**, 29–38.

Dougalis, V. A. & Durán, A. (2020). Notes on a high order fully discrete scheme for the Korteweg–de Vries equation with a time stepping procedure of Runge–Kutta composition type. Available at http://arxiv.org/abs/2005.12955.

Dougalis, V. A., Durán, A. & Mitsotakis, D. E. (2019) Numerical approximation to Benjamin type equations. Generation and stability of solitary waves. *Wave Motion*, **85**, 34–56.

Dougalis, V. A. & Karakashian, O. A. (1985) On some high-order accurate fully discrete Galerkin methods for the Korteweg–de Vries equation. *Math. Comp.*, **45**, 329–345.

Hairer, E., Lubich, C. & Wanner, G. (2004) *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*. New York–Heidelberg–Berlin: Springer.

Kalisch, H. (2005) Rapid convergence of a Galerkin projection of the KdV equation. *C. R. Acad. Sci. Paris, Ser. I*, **341**, 457–460.

Kappeler, T. & Topalov, P. (2006) Global well-posedness of KdV in $H^{-1}$ $(T, \mathbb{R})$. *Duke Math. J.*, **135**, 327–360.

Karakashian, O. A., Akrivis, G. D. & Dougalis, V. A. (1993) On optimal-order error estimates for the nonlinear Schrödinger equation. *SIAM J. Numer. Anal.*, **30**, 377–400.

Karakashian, O. A. & McKinney, W. (1990) On optimal high-order in time approximations for the Korteweg–de Vries equation. *Math. Comp.*, **55**, 473–496.

Killip, R. & Vişan, V. (2019) KdV is well-posed in $\mathrm{H}^{-1}$. *Ann. Math.*, **190**, 249–305.

Linares, F. & Scialom, M. (2005) On generalized Benjamin type equations. *Disc. Cont. Dyn. Sys.*, **12**, 161–174.

Maday, Y. & Quarteroni, A. (1988) Error analysis for spectral approximation of the Korteweg–de Vries equation. *RAIRO Model. Math. Analyse Numér.*, **22**, 499–529.

Mercier, B. (1989) *An Introduction to the Numerical Analysis of Spectral Methods*. Lecture Notes in Physics, vol. 310. Berlin: Springer.

Sanz-Serna, J. M. & Abia, L. M. (1991) Order conditions for canonical Runge–Kutta schemes. *SIAM J. Numer. Anal.*, **28**, 1081–1096.

Sanz-Serna, J. M. & Calvo, M. P. (1994) *Numerical Hamiltonian Problems*. London: Chapman and Hall.

Temam, R. (1969) Sur un problème non linéaire. *J. Math. Pure. Appl.*, **48**, 159–172.

Wahlbin, L. B. (1974) A dissipative Galerkin method for the numerical solution of first order hyperbolic equations. *Mathematical Aspects of Finite Elements in Partial Differential Equations* (C. de Boor ed). New York: Academic Press, pp. 147–169.

Yoshida, H. (1990) Construction of higher order symplectic integrators. *Phys. Lett. A*, **150**, 262–268.