ON OPTIMAL ORDER ERROR ESTIMATES FOR THE NONLINEAR SCHRÖDINGER EQUATION

OHANNES KARAKASHIAN, GEORGIOS D. AKRIVIS, AND VASSILIOS A. DOUGALIS

Abstract. Implicit Runge–Kutta methods in time are used in conjunction with the Galerkin method in space to generate stable and accurate approximations to solutions of the nonlinear (cubic) Schrödinger equation. The temporal component of the discretization error is shown to decrease at the classical rates in some important special cases.

1. Introduction

In this paper we consider the following initial boundary value problem for the cubic Schrödinger equation: Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with boundary $\partial\Omega$. We seek a complex-valued function u satisfying

$$\begin{cases} u_t = \mathrm{i}\Delta u + \mathrm{i}\lambda |u|^2 u, & (x,t) \in \bar{\Omega} \times [0,t^*], \\ u = 0, & (x,t) \in \partial\Omega \times [0,t^*], \\ u(x,0) = u^0(x), & x \in \bar{\Omega}, \end{cases} \tag{1.1}$$

where λ is a nonzero real number and u^0 is a given complex-valued function on $\bar{\Omega}$. We assume that the data of (1.1) are such that the problem possesses a unique classical solution in $C^\mu(\bar{\Omega} \times [0,t^*])$, where μ is sufficiently large for the approximation results that will be proved in the sequel. We refer the reader to the surveys [[12] and [13] for an overview of the physical significance and the mathematical theory of the nonlinear Schrödinger equation.

We shall approximate the solution of (1.1) using Galerkin finite element type methods in space and suitable implicit Runge–Kutta (IRK) schemes for time-stepping. In another paper [2], to which we also refer the reader for references to previous work on the numerical solution of (1.1), we study Galerkin methods of second-order temporal accuracy and address issues of their efficient implementation. The emphasis in the paper at hand is on higher order IRK methods; in particular, we shall prove L^2 error

---

1991 *Mathematics Subject Classification.* 65M60.

Key words and phrases. nonlinear Schrödinger equation, Galerkin methods, implicit Runge–Kutta schemes, order reduction.

The work of G.D.A. and V.A.D. was supported by the Institute of Applied and Computational Mathematics of the Research Center of Crete-FORTH and by the Science Alliance program of the University of Tennessee.

The work of O.K. was supported by Air Force Office of Scientific Research grant 88-0019 and by the Science Alliance program of the University of Tennessee.

estimates whose spatial component is of optimal rate of convergence and whose temporal component decreases at the optimal (classical) rates in some important special cases.

Specifically, for suitable classes of IRK methods, we shall show that

$$\max_{0\le n\le N} \|u_h^n - u(t^n)\|_{L^2(\Omega)} \le c(k^\sigma + h^r), \tag{1.2}$$

where $\sigma = \min\{p+3, \nu\}$ if Ω is a general domain with piecewise smooth curved boundary and $\sigma = \nu$ if Ω is any polyhedral domain (or any finite interval if $d = 1$). Here h is the space discretization parameter, r is the optimal spatial rate of convergence in L^2 (cf. §2), k is the time step, $t^n = nk, 0 \le n \le N, N = t^*/k, u$ is the solution of (1.1) (assumed to be sufficiently smooth up to the boundary), u_h^n its fully discrete approximation at time t^n, and c is a constant depending on u and the data of (1.1) but independent of h and k. The time-stepping is effected through a $q-$stage IRK method with (classical) order of accuracy $\nu; p$ is the stage order (cf. (B), (C) in §2.3). For (1.2) to hold, the Galerkin subspaces and the IRK scheme must satisfy a series of standard properties (cf. §2), u_h^0 should be chosen so that $\|u_h^0 - u^0\|_{L^2(\Omega)} = O(h^r)$, $d < 2r$, and the (weak) mesh condition $k = O(h^{d/2\sigma})$ as $h \to 0$ should be satisfied.

It is well known that approximating smooth solutions of initial and boundary value problems for some partial differential equations (PDEs) by high order Runge–Kutta methods results sometimes in observed temporal rates of convergence lower than the (classical) order ν (cf., e.g., [6], [4], [9]). From (1.2) we may infer that, under our hypotheses, no reduction of the order ν occurs for the problem (1.1) if Ω is a polyhedral domain or if $p + 3 \ge \nu$, as would be the case, for example, with practically important schemes such as the conservative $q-$stage Gauss–Legendre collocation type methods $(p = q, \nu = 2q)$ with up to three stages, and the two-, respectively, three-stage optimal order diagonally implicit RK (DIRK) schemes, for which $p = 1, \nu = 3$ and $p = 1, \nu = 4$, respectively.

The proof of (1.2) relies on constructing suitable smooth approximations $u^{n,j}$ to the values $u(t^{n,j}), 1 \le j \le q$, of the solution u of (1.1) at the intermediate time levels $t^{n,j}$ of the Runge–Kutta scheme (cf. §2); the $u^{n,j}$ are combined to produce a smooth approximation u^{n+1} of $u(t^{n+1})$ to which u_h^{n+1} is then compared. The $u^{n,j}$ are expressed as polynomials of k (that may be viewed as extensions of Taylor expansions of $u(t)$ about $t^{n,j}$) in the form $u^{n,j} = \sum_{\ell=0}^{\sigma} k^\ell \alpha_{j\ell}$, where the $\alpha_{j\ell}$ are smooth functions on $\bar\Omega$ and depend on the solution of (1.1) and the Runge–Kutta method. They occur naturally in analyzing the consistency of the scheme and satisfy a crucial cancellation property (cf. (4.8)). The heart of the proof is then checking that the $\alpha_{j\ell}$'s vanish on $\partial\Omega$, a fact that allows approximating $u^{n,j}$ to optimal order in space by elements of the Galerkin subspace. The technique of constructing expansions in powers of k at the intermediate time levels has its origins in [10] but it was elaborated fully in [11] in the context of the initial and periodic boundary value problem for the Korteweg–de Vries equation. In the latter paper, the $u^{n,j}$ were defined implicitly as the intermediate stages that would be obtained in the process of applying the Runge–Kutta method to $u(t^n)$. In so doing, a host of difficulties had to be handled. In particular, the existence of the $u^{n,j}$ had to be established. Moreover, it was shown that the $u^{n,j}$ were as smooth as $u(t^n)$ and that appropriate high order Sobolev norms of theirs were bounded in terms of corresponding norms of $u(t^n)$ with constants free of any dependence on k. In the paper at hand, the $u^{n,j}$ are defined explicitly and thus the issue of existence is bypassed. On the other hand, the intermediate equations are not satisfied exactly, but to within an error of $O(k^{\sigma+1})$, which is perfectly acceptable for the purpose of error estimation. In summary, the approach followed here, although similar in spirit to the one adopted in [11], is considerably simpler.

Using a different technique, an error estimate analogous to (1.2) was shown in [1] in the context of a linear Schrödinger equation on a general domain Ω with a time-dependent potential (replace in (1.1) $\mathrm{i}\lambda|u|^2u$ by $\beta(x,t)u$) with an exponent of k equal to $\min\{q+2,2q\}$ for the $q-$stage Gauss–Legendre methods. In the specific nonlinear autonomous case at hand, the nonlinearity $|u|^2u$ affords proving that solutions of (1.1), that are smooth up to the boundary of $\Omega\times(0,t^*)$, satisfy $\Delta^m u=0$ on $\partial\Omega\times[0,t^*]$, for m high enough depending on Ω. This is the important observation that subsequently allows proving that $\alpha_{j\ell}=0$ for $0\le\ell\le\sigma$ on $\partial\Omega$, which in turn implies (1.2). In fact, for $\nu>p+3$, by introducing local coordinates, it is possible to verify, in the case of a nonpolyhedral domain with piecewise smooth boundary, that the functions $\alpha_{j,p+4}$ do not vanish on $\partial\Omega$ for arbitrary solutions of (1.1) that are smooth up to the boundary and for arbitrary schemes within the class of IRK methods under consideration. Hence, our technique of investigating whether temporal order reduction occurs for (1.1) encounters a barrier at $\sigma=p+3$ if the boundary of Ω is curved. To ascertain whether this barrier is real or merely an artifact of our particular proof one should perhaps resort to numerical experiments on plane domains. Such experiments will not be easy to design and perform because of the nonlinearity of the problem, the high degree of accuracy required of the spatial and temporal discretizations, and the fact that there is no order reduction if the domain is polygonal. As was stated previously, our proof depends on the assumption that the solution of (1.1) is smooth up to the boundary. This assumption may not be realistic for arbitrary polyhedral domains, for which singularities due to the corners cannot be ruled out. However, there are important special cases ($d=1$, special solutions on rectangles in the presence of symmetries, etc.) for which such smoothness is expected under reasonable smoothness and compatibility conditions on u^0.

The plan of the paper is as follows. In §2 we introduce notation, list our assumptions for the Galerkin subspaces and the IRK methods, and construct the fully discrete approximations u_h^n. In §3 we discuss briefly the existence, L^2-boundedness and uniqueness of u_h^n. The main results of the paper that lead to the proof of (1.2) are to be found in §4. In §5 we briefly indicate how the techniques and results of the present paper extend to some related PDEs.

In a subsequent paper we shall construct and analyze efficient implementations of Newton's method for solving the nonlinear systems produced by the IRK schemes at each time step and show results of relevant numerical experiments. This has been carried out in the context of $O(k^2)$ schemes in [2].

2. Preliminaries

2.1. **Some function spaces.** For $1 \le p \le \infty, L^p = L^p(\Omega)$ will denote the Banach space of (classes of) complex-valued measurable functions defined on Ω, equipped with the norm

$$\|v\|_{L^p} = \left(\int_\Omega |v|^p \, dx \right)^{1/p}, \quad 1 \le p < \infty,$$
$$\|v\|_{L^\infty} = \operatorname{ess\,sup}_{x\in\Omega} |v(x)|.$$

In particular, for $p = 2$, the space L^2 has the inner product $(u, v) = \int_\Omega u\bar{v}\, dx$. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d), \alpha_i \ge 0$, denote a *multi-integer*, and let $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d$.

For $m \ge 0$ integer, H^m will denote the Hilbert space of (classes of) complex-valued measurable functions which, together with their (distributional) derivatives of order up to m, are in L^2, i.e.,

$$H^m = \{v : D^\alpha v \in L^2, |\alpha| \le m\}.$$

These spaces are equipped with the norms $\|v\|_{H^m} = \left\{ \sum_{|\alpha|\le m} \|D^\alpha v\|_{L^2}^2 \right\}^{1/2}$. To simplify notation, we shall denote the norm on $L^2 = H^0$ by $\|\cdot\|$. We let $H_0^1 = \{v \in H^1 : v = 0 \text{ on } \partial\Omega\}$.

We shall also use spaces of continuously differentiable functions. With Q denoting a bounded domain in $\mathbb{R}^d, C^m(Q)$ is the usual space of complex-valued functions v defined on Q which, together with their partial derivatives $D^\alpha v$ of order $|\alpha| \le m$, are continuous on Q. Similarly, $C^m(\bar{Q})$ is the space of functions v in $C^m(Q)$ for which $D^\alpha v$ is bounded and uniformly continuous on Q for $|\alpha| \le m$. $C^m(\bar{Q})$ is a Banach space with norm

$$\|v\|_{C^m(\bar{Q})} = \max_{|\alpha|\le m} \sup_{x\in\bar{Q}} |D^\alpha v(x)|.$$

2.2. **The approximating spaces.** For integer $r \ge 2$ and $0 < h < 1, S_h^r \subset H^1 \cap C^0(\bar{\Omega})$ will represent an approximating finite-dimensional space of functions. Such spaces typically consist of piecewise polynomial functions of degree $\le r - 1$ defined on a suitable partition of Ω.

We assume that these spaces possess good approximation properties; indeed, that there exists a constant c independent of h such that for each $v \in H^r \cap H_0^1$, there exists $\chi \in S_h^r$ such that

$$\|v - \chi\| \le ch^r \|v\|_r, \tag{2.1}$$

and if in addition $v \in C^2(\bar{\Omega})$, then

$$\|v - \chi\|_{L^\infty} \le ch^2 \|v\|_{C^2(\bar{\Omega})}. \tag{2.2}$$

We shall assume that the elements of S_h^r satisfy the following *inverse* inequality

$$\|\chi\|_{L^\infty} \le ch^{-d/2} \|\chi\|. \tag{2.3}$$

Let $V = S_h^r + (H^2 \cap H_0^1)$. We assume the existence of a family of sesquilinear forms $B_h^r : V \times V \to \mathbb{C}$ with the following properties:

$$B_h^r(v, v) \ \text{ is real for } v \in V, \tag{2.4a}$$

$$B_h^r(v, v) \geq c\|v\|_1^2 \ \text{ for } c > 0, \forall v \in S_h^r, \tag{2.4b}$$

$$B_h^r(v, \chi) = -(\Delta v, \chi) \quad \forall \chi \in S_h^r, \quad v \in H^2 \cap H_0^1. \tag{2.4c}$$

With B_h^r we associate an *elliptic projection* operator $P_E : H^2 \cap H_0^1 \to S_h^r$ by

$$B_h^r(P_E v, \chi) = B_h^r(v, \chi) = -(\Delta v, \chi) \quad \forall \chi \in S_h^r. \tag{2.5}$$

We assume that for some constant c independent of h

$$\|P_E v - v\| \leq c h^r \|v\|_r \qquad \forall v \in H^r \cap H_0^1. \tag{2.6}$$

The most well-known family of such sesquilinear forms is provided by the so-called standard Galerkin method. In this case, $S_h^r \subset H_0^1$ and

$$B_h^r(v, w) = \int_\Omega \nabla v \cdot \nabla \bar{w} \, dx.$$

There are several other examples of finite element formulations which generate forms B_h^r satisfying the requisite properties. Among these we may cite two methods of Nitsche that use subspaces $S_h^r \subset H^1$ which do not satisfy the homogeneous Dirichlet boundary conditions and the Lagrange multiplier method of Babuska [3].

2.3. **The implicit Runge–Kutta methods.** For $q \geq 1$ integer, a $q-$stage IRK method is specified by a set of constants arranged in *tableau* form

$$\begin{array}{ccc|c} a_{11} & \dots & a_{1q} & \tau_1 \\ \vdots & & \vdots & \vdots \\ a_{q1} & \dots & a_{qq} & \tau_q \\ \hline b_1 & \dots & b_q & \end{array}.$$

Given the initial value problem

$$y = f(t, y), \quad 0 < t \leq t^*, \quad y(0) = y^0, \tag{2.7}$$

IRK methods generate approximations y^n to $y(t^n), 0 \leq n \leq N$, where $k = t^*/N$ is the temporal stepsize and $t^n = nk$, as follows. Let

$$y^{n+1} = y^n + k \sum_{j=1}^{q} b_j f(t^{n,j}, y^{n,j}), \tag{2.8}$$

where $t^{n,j} = t^n + k\tau_j$ and where the *intermediate stages* $y^{n,j}$ are given by the system of coupled equations

$$y^{n,j} = y^n + k \sum_{m=1}^{q} a_{jm} f(t^{n,m}, y^{n,m}), \quad j = 1, \dots, q. \tag{2.9}$$

We shall assume that these methods (constants) satisfy certain *stability* and *consistency* conditions [5], [7], [9]. Indeed it will be require that

$$\text{(S)}\qquad \begin{array}{l} b_i \geq 0, \quad i = 1, \dots, q. \\ \text{The } q \times q \text{ array } m_{ij} = a_{ij}b_i + a_{ji}b_j - b_ib_j \text{ is positive semidefinite.} \end{array}$$

The above condition, known as *algebraic stability*, is stronger than that of A–stability.

The consistency conditions are given by the following *simplifying assumptions*:

$$\text{(B)}\qquad \sum_{j=1}^{q} b_j \tau_j^\ell = \frac{1}{\ell+1}, \quad \ell = 0, \dots, \nu - 1,$$

$$\text{(C)}\qquad \sum_{j=1}^{q} a_{ij} \tau_j^\ell = \frac{\tau_i^{\ell+1}}{\ell+1}, \quad i = 1, \dots, q, \ \ell = 0, \dots, p - 1,$$

$$\text{(D)}\qquad \sum_{i=1}^{q} a_{ij} \tau_i^\ell b_i = \frac{b_j}{\ell+1}(1 - \tau_j^{\ell+1}), \quad j = 1, \dots, q, \ \ell = 0, \dots, \rho - 1,$$

for some integers $\nu \geq 1, p \geq 1, \rho \geq 1$. We assume that

$$\nu \leq \rho + p + 1, \tag{2.10a}$$

$$\nu \leq 2p + 2. \tag{2.10b}$$

We shall refer to p and ν as the *stage order* and *classical order*, respectively. It is well known [5], that the simplifying assumptions (B), (C), and (D) together with (2.10a) and (2.10b) imply that the method is consistent of order ν.

The existence of the numerical approximations is obtained by assuming the following positivity property (cf. [8]):

$$\text{(P)}\qquad \begin{array}{l} \text{The matrix } A = (a_{ij}) \text{ is invertible and there exists a positive diagonal} \\ \text{matrix } D \text{ such that } x^T C x > 0, \ \forall x \in \mathbb{R}^q, x \neq 0, \text{ where } C = DA^{-1}D^{-1}. \end{array}$$

We next give examples of some families of IRK methods that satisfy these properties and the tableaus of their two- and three-stage members.

(i) *Gauss–Legendre methods.* These methods form a particularly interesting class in that the matrix M in (S) vanishes identically, a fact that has important implications such as L^2-conservativeness of the schemes and mild growth of the discretization error. For this class, $\nu = 2q, p = \rho = q$ [9, p. 71]. For (S) see [9, p. 101], and for (P) see [9, p. 157].

$$\begin{array}{cc|c} \frac{1}{4} & \frac{1}{4} - \frac{1}{2\sqrt{3}} & \frac{1}{2} - \frac{1}{2\sqrt{3}} \\ \frac{1}{4} + \frac{1}{2\sqrt{3}} & \frac{1}{4} & \frac{1}{2} + \frac{1}{2\sqrt{3}} \\ \hline \frac{1}{2} & \frac{1}{2} & \end{array}\,, \qquad \begin{array}{ccc|c} \frac{5}{36} & \frac{80-24\sqrt{15}}{360} & \frac{50-12\sqrt{15}}{360} & \frac{1}{2} - \frac{\sqrt{15}}{10} \\ \frac{50+15\sqrt{15}}{360} & \frac{2}{9} & \frac{50-15\sqrt{15}}{360} & \frac{1}{2} \\ \frac{50+12\sqrt{15}}{360} & \frac{80+24\sqrt{15}}{360} & \frac{5}{36} & \frac{1}{2} + \frac{\sqrt{15}}{10} \\ \hline \frac{5}{18} & \frac{8}{18} & \frac{5}{18} & \end{array}$$

(ii) *Radau IIA methods.* These methods are characterized by $\tau_q = 1$. Also, $\nu = 2q-1, p = q, \rho = q-1$ (cf. [9, p. 71]). For (S) (P), see [9, p. 101, 164], respectively.

$$\begin{array}{cc|c} \frac{5}{12} & -\frac{1}{12} & \frac{1}{3} \\ \frac{3}{4} & \frac{1}{4} & 1 \\ \hline \frac{3}{4} & \frac{1}{4} & \end{array}, \qquad \begin{array}{ccc|c} \frac{88-7\sqrt{6}}{360} & \frac{296-169\sqrt{6}}{1800} & \frac{-2+3\sqrt{6}}{225} & \frac{4-\sqrt{6}}{10} \\ \frac{296+169\sqrt{6}}{1800} & \frac{88+7\sqrt{6}}{360} & \frac{-2-3\sqrt{6}}{225} & \frac{4+\sqrt{6}}{10} \\ \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} & 1 \\ \hline \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} & \end{array}$$

Both of the families above are infinite, in the sense that arbitrarily high order methods can be constructed.

(iii) *Two- and three-stage optimal order* DIRK *methods.*

$$\begin{array}{cc|c} \frac{1}{2}+\frac{1}{2\sqrt{3}} & 0 & \frac{1}{2}+\frac{1}{2\sqrt{3}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{2}+\frac{1}{2\sqrt{3}} & \frac{1}{2}-\frac{1}{2\sqrt{3}} \\ \hline \frac{1}{2} & \frac{1}{2} & \end{array}, \qquad \begin{array}{ccc|c} \gamma & 0 & 0 & \gamma \\ \frac{1}{2}-\gamma & \gamma & 0 & \frac{1}{2} \\ 2\gamma & 1-4\gamma & \gamma & 1-\gamma \\ \hline \frac{1}{24(\frac{1}{2}-\gamma)^2} & 1-\frac{1}{12(\frac{1}{2}-\gamma)^2} & \frac{1}{24(\frac{1}{2}-\gamma)^2} & \end{array}$$

Here $\gamma = 1/2 + 1/\sqrt{3}\cos\pi/18$ is the largest root of $24x^3 - 36x^2 + 12x - 1 = 0$. For the two-stage method $\nu = 3, p = \rho = 1$. For the three-stage method $\nu = 4, p = \rho = 1$; hence (2.10a) is not satisfied. This will necessitate a slight modification in the estimation of the local truncation error. For (S) see [9, p. 121]. (P) also holds for both methods but an existence proof may be given that does not use (P).

2.4. **The fully discrete approximations.** Following (2.8) and (2.9), we define the fully discrete approximations $\{u_h^n\}_{n=0}^N$ recursively as follows: Let $\pi_h u^0$ be any conveniently chosen element of S_h^r, e.g., L^2−projection, interpolant, etc., that is optimally close to u^0 in the sense that

$$\|u^0 - \pi_h u^0\| \le ch^r. \tag{2.11}$$

We set $u_h^0 = \pi_h u^0$ and for $n = 0, \dots, N-1, \chi \in S_h^r$,

$$(u_h^{n+1}, \chi) = (u_h^n, \chi) + k\sum_{j=1}^q b_j\big\{-\mathrm{i}B_h^r(u_h^{n,j}, \chi) + \mathrm{i}\lambda(|u_h^{n,j}|^2 u_h^{n,j}, \chi)\big\}, \tag{2.12}$$

where the intermediate stages $\{u_h^{n,j}\}_{j=1}^q$ satisfy the system of equations

$$\begin{gathered}(u_h^{n,j}, \chi) = (u_h^n, \chi) + k\sum_{m=1}^q a_{jm}\big\{-\mathrm{i}B_h^r(u_h^{n,m}, \chi) + \mathrm{i}\lambda(|u_h^{n,m}|^2 u_h^{n,m}, \chi)\big\}, \\ \forall\chi \in S_h^r, \quad j = 1, \dots, q.\end{gathered} \tag{2.13}$$

Since A is nonsingular we may also write (2.12), setting $\boldsymbol{e} = (1, \dots, 1)^T \in \mathbb{R}^q$, as

$$u_h^{n+1} = (1 - b^T A^{-1}\boldsymbol{e})u_h^n + b^T A^{-1}\mathcal{U}_h^n, \quad \mathcal{U}_h^n = (u_h^{n,1}, \dots, u_h^{n,q})^T. \tag{2.14}$$

For ease of notation we introduce the maps $\varDelta_h : S_h^r \to S_h^r$ and $g_h : S_h^r \to S_h^r$ defined by

$$(\varDelta_h w, \chi) = -B_h^r(w, \chi) \quad \forall \chi \in S_h^r, \tag{2.15}$$

$$(g_h(w), \chi) - (|w|^2 w, \chi) \quad \forall \chi \in S_h^r. \tag{2.16}$$

That these maps are well defined follows from the Riesz representation theorem. Letting $f_h = \mathrm{i}(\varDelta_h + \lambda g_h)$, we may write (2.12) and (2.13) as

$$u_h^{n+1} = u_h^n + k \sum_{j=1}^{q} b_j f_h(u_h^{n,j}), \tag{2.17}$$

$$u_h^{n,j} = u_h^n + k \sum_{m=1}^{q} a_{jm} f_h(u_h^{n,m}), \quad j = 1, \dots, q. \tag{2.18}$$

The map $g(z) = |z|^2 z$ is locally Lipschitz. It will prove very convenient to associate with g a Lipschitz map $\tilde{g} : \mathbb{C} \to \mathbb{C}$ as follows: With u denoting the solution of (1.1), let $M(u) = \sup_{(x,t)\in\bar{\varOmega}\times[0,t^*]} |u(x,t)|$ and

$$\tilde{g}(z) = \begin{cases} |z|^2 z, & |z| \le 2M(u), \\ 4M^2(u) z, & |z| \ge 2M(u). \end{cases} \tag{2.19}$$

In case $\max_{1\le i\le q} \tau_i > 1$, we replace t^* in the definition of $M(u)$ by $t^* - k + k \max_{1\le i\le q} \tau_i$, assuming that the solution possesses an extension out of the temporal interval under consideration.

The easy proof of the following result is left to the reader.

Lemma 2.1. *The map $\tilde{g}$ is Lipschitz continuous with Lipschitz constant $L = 12M^2(u)$.* □

The map $\tilde{g}$ induces a map $\tilde{g}_h : S_h^r \to S_h^r$ via

$$(\tilde{g}_h(w), \chi) = (\tilde{g}(w), \chi) \quad \forall \chi \in S_h^r. \tag{2.20}$$

Letting $\tilde{f}_h = \mathrm{i}(\varDelta_h + \lambda \tilde{g}_h)$, in analogy with (2.17) and (2.18), we define the auxiliary functions $\{\tilde{u}_h^n\}_{n=0}^N$ by $\tilde{u}_h^0 = u_h^0 = \pi_h u^0$, and, for $n = 0, \dots, N-1$, by

$$\tilde{u}_h^{n+1} = \tilde{u}_h^n + k \sum_{j=1}^{q} b_j \tilde{f}_h(\tilde{u}_h^{n,j}), \tag{2.21}$$

$$\tilde{u}_h^{n,j} = \tilde{u}_h^n + k \sum_{m=1}^{q} a_{jm} \tilde{f}_h(\tilde{u}_h^{n,m}), \quad j = 1, \dots, q. \tag{2.22}$$

3. Existence, L^2 bounds and uniqueness of the fully discrete approximations

The main tool in obtaining existence is the following well-known version of the Brouwer fixed point theorem (cf. [2]).

Lemma 3.1. *Let $(H,(\cdot,\cdot)_H)$ be a finite-dimensional inner product space. Let $g:H\to H$ be continuous and assume that for some $\alpha>0, \mathrm{Re}(g(z),z)_H\ge 0$ for every $z\in H$ with $\|z\|_H=\alpha$. Then there exists $z^*\in H$ with $\|z^*\|_H\le\alpha$ such that $g(z^*)=0$.* □

We shall also use the following lemma.

Lemma 3.2. *Assume that the IRK method satisfies* (P) *and let $C=DA^{-1}D^{-1}$ and $D=\mathrm{diag}\{d_1,\dots,d_q\}$. Let $\{\varphi^j\}_{j=1}^q\in L^2$. Then, for some positive constant c depending only on the IRK method*

$$\mathrm{Re}\sum_{j,m=1}^{q} c_{jm}d_jd_m(\varphi^m,\varphi^j)\ge c\sum_{j=1}^{q}\|\varphi^j\|^2. \tag{3.1}$$

Proof. Let $\varphi^j=\varphi_1^j+\mathrm{i}\varphi_2^j$ where φ_1^j and φ_2^j are real-valued. Then

$$\begin{aligned}\sum_{j,m=1}^{q} c_{jm}d_jd_m(\varphi^m,\varphi^j)&=\int_\Omega\sum_{j,m=1}^{q} c_{jm}d_jd_m(\varphi_1^m\varphi_1^j+\varphi_2^m\varphi_2^j)\,dx\\&\quad+\mathrm{i}\int_\Omega\sum_{j,m=1}^{q} c_{jm}d_jd_m(\varphi_2^m\varphi_1^j-\varphi_1^m\varphi_2^j)\,dx.\end{aligned}\tag{3.2}$$

The conclusion now follows from (P). □

Proposition 3.1. *Assume that* (P) *holds. Then, systems* (2.18) *and* (2.22) *have solutions.*

Proof. Let $H=(S_h^r)^q$ and equip it with the usual inner product $(\Phi,\Psi)_H=\sum_{i=1}^q(\varphi_i,\psi_i)$ and corresponding norm $\|\Phi\|_H^2=(\Phi,\Phi)_H$. Define $\mathcal{G}=(g_1,\dots,g_q)^T:H\to H$ by

$$g_j(z_1,\dots,z_q)=\sum_{m=1}^{q} c_{jm}d_jd_m(z_m-u_h^n)-kd_j^2\tilde f_h(z_j),\quad j=1,\dots,q.$$

Multiplying the jth equation with $\bar z_j$, integrating and summing, we get

$$(\mathcal{G}(Z),Z)=\sum_{j,m=1}^{q} c_{jm}d_jd_m\{(z_m,z_j)-(u_h^n,z_j)\}-k\sum_{j=1}^{q}d_j^2\big(\tilde f_h(z_j),z_j\big).$$

Using (3.1) and the fact that $(\tilde f_h(z),z)$ is imaginary, we get

$$\mathrm{Re}(\mathcal{G}(Z),Z)\ge\|Z\|_H\{c_1\|Z\|_H-c_2\|u_h^n\|\}$$

for some positive constants c_1,c_2. It follows that $\mathrm{Re}(\mathcal{G}(Z),Z)>0$ on the sphere of radius $1+c_2\|u_h^n\|/c_1$ in H. By Lemma 3.1, there exists $Z=(z_1,\dots,z_q)^T$ satisfying $\mathcal{G}(Z)=0$. That Z is indeed a solution to (2.22) is readily seen. As for (2.18), it is sufficient to note that we have used only the properties that $\tilde f_h$ is continuous and that $(\tilde f_h(z),z)$ is imaginary. □

As far the boundedness of the solution is concerned we have the following result, a consequence of algebraic stability.

Proposition 3.2. *Assume that* (2.18) *has a solution and that* (S) *holds. Then*

$$\|u_h^{n+1}\| \le \|u_h^n\|. \tag{3.3}$$

For the Gauss–Legendre methods (3.3) *holds as an equality, i.e., these methods are* L^2-*conservative.*

Proof. Take the L^2-inner product of u_h^{n+1} given by (2.17) by itself, denoting $f_h^j = f_h(u^{n,j})$, and obtain

$$\|u_h^{n+1}\|^2 = \|u_h^n\|^2 + k\sum_{j=1}^q b_j\big[(f_h^j,u_h^n)+(u_h^n,f_h^j)\big] + k^2\sum_{j,s=1}^q b_jb_s(f_h^j,f_h^s).$$

Using then (2.18) in the right-hand side of the above, we see that since $(f_h^j, u_h^{n,j})$ is imaginary,

$$\|u_h^{n+1}\|^2 = \|u_h^n\|^2 - k^2\sum_{j,s=1}^q m_{js}\,\mathrm{Re}(f_h^j,f_h^s);$$

the result now follows from (S). For the Gauss–Legendre methods we have already noted that $m_{js}=0$. □

Using the fact that $\tilde g$ is Lipschitz, we can prove in a straightforward way that the solution of (2.22) is unique. More specifically, we can show that there exists $k_0 > 0$ that depends on the Lipschitz constant L and the IRK method, such that (2.22) admits a unique solution provided $k \le k_0$.

We thus have a local uniqueness result for solutions of (2.18) in the sense that two solutions whose components are in the ball $\{v \in S_h^r : \|v\|_{L^\infty} \le 2M(u)\}$ are identical. In fact using the embedding of H^1 in L^4, it is possible to prove that for $d = 1, 2,$ or 3, solutions with components in $\{v \in S_h^r : \|v\|_{L^4} \le K\}$ for some $K > 0$ are unique provided $k \le k_0$ for some k_0 that depends on K and the IRK method.

For a detailed study of uniqueness in the context of single-step schemes of second-order temporal accuracy, we refer the reader to [2].

4. Error estimates

Given $n, 0 \le n \le N-1$, let $u = u(\cdot, t^n)$. Let the functions $\alpha_{j\ell}$ be defined by

$$\alpha_{j0} = u, \quad j=1,\dots,q,$$
$$\alpha_{j,\ell+1} = \mathrm{i}\sum_{s=1}^q a_{js}\Big\{\Delta\alpha_{s\ell} + \lambda \sum_{\substack{|m|=\ell\\ m=(m_1,m_2,m_3)}} \alpha_{sm_1}\alpha_{sm_2}\bar\alpha_{sm_3}\Big\}, j=1,\dots,q, \ell=0,\dots,\nu-1. \tag{4.1}$$

We shall first establish a series of results involving these functions. First some useful identities.

Lemma 4.1. *Assume that* (C) *holds together with* (2.10b)*. Denoting the vector* $(\alpha_{1\ell}, \dots, \alpha_{q\ell})^T$ *by* α_ℓ*, with* $D_t^\ell u = (\partial^\ell/\partial t^\ell)u(x,t)\big|_{t=t^n}$ *and* $T = \operatorname{diag}\{\tau_1, \dots, \tau_q\}$*, we have*

$$\alpha_\ell = D_t^\ell u\,\frac{T^\ell \boldsymbol{e}}{\ell!}, \quad \ell = 0, \dots, p \ \textit{if} \ \ p \le \nu, \tag{4.2}$$

$$\alpha_{p+1} = D_t^{p+1} u\,\frac{AT^p \boldsymbol{e}}{p!} \quad \textit{if} \ \ p \le \nu - 1, \tag{4.3}$$

$$\begin{aligned} \alpha_{\ell+1} = {} & D_t^{\ell+1} u\,\frac{AT^\ell \boldsymbol{e}}{\ell!} + \mathrm{i}A\varDelta\Big[\alpha_\ell - D_t^\ell u\,\frac{T^\ell \boldsymbol{e}}{\ell!}\Big] \\ & + 2\mathrm{i}\lambda A \sum_{m_1=p+1}^{\ell} \frac{D_t^{\ell-m_1}|u|^2}{(\ell-m_1)!}\Big[T^{\ell-m_1}\alpha_{m_1} - D_t^{m_1} u\,\frac{T^\ell \boldsymbol{e}}{m_1!}\Big] \\ & + \mathrm{i}\lambda A \sum_{m_3=p+1}^{\ell} \frac{D_t^{\ell-m_3}u^2}{(\ell-m_3)!}\Big[T^{\ell-m_3}\bar\alpha_{m_3} - D_t^{m_3}\bar u\,\frac{T^\ell \boldsymbol{e}}{m_3!}\Big], \\ & \ell = p+1, \dots, \nu-1 \ \textit{if} \ \ p \le \nu - 2. \end{aligned} \tag{4.4}$$

Proof. The case $\ell = 0$ is obvious. So assume that (4.2) holds for all indices up to some ℓ, where $0 \le \ell \le p-1$. Then using (1.1), (4.1), and (C)

$$\begin{aligned} \alpha_{j,\ell+1} &= \sum_{s=1}^q a_{js}\Big\{\mathrm{i}\varDelta\frac{\tau_s^\ell}{\ell!}D_t^\ell u + \mathrm{i}\lambda \sum_{|m|=\ell} \frac{\tau_s^{m_1}}{m_1!}D_t^{m_1}u\,\frac{\tau_s^{m_2}}{m_2!}D_t^{m_2}u\,\frac{\tau_s^{m_3}}{m_3!}D_t^{m_3}\bar u\Big\} \\ &= \sum_{s=1}^q a_{js}\big\{\mathrm{i}\varDelta\frac{\tau_s^\ell}{\ell!}D_t^\ell u + \mathrm{i}\lambda\frac{\tau_s^\ell}{\ell!}D_t^\ell\big(|u|^2u\big)\big\} \\ &= \frac{1}{\ell!}\Big(\sum_{s=1}^q a_{js}\tau_s^\ell\Big)\mathrm{i}D_t^\ell\{\varDelta u + \lambda|u|^2u\} \\ &= \frac{\tau_j^{\ell+1}}{(\ell+1)!}D_t^{\ell+1}u. \end{aligned}$$

Formula (4.3) can be established in an analogous manner. To establish (4.4), for a given $\ell, p+1 \le \ell \le \nu-1$, we let

$$\begin{aligned} M^\ell &= \{m = (m_1, m_2, m_3) : 0 \le m_i \le \ell,\ 1 \le i \le 3,\ |m| = \ell\}, \\ M_0^\ell &= \{m \in M^\ell : 0 \le m_i \le p,\ i = 1, 2, 3\}, \\ M_i^\ell &= \{m \in M^\ell : p+1 \le m_i \le \ell\}, \quad i = 1, 2, 3. \end{aligned}$$

It follows from (2.10b) that $M_i^\ell \cap M_j^\ell = \emptyset, i \neq j,\ i,j = 0,1,2,3,$ and that $M^\ell = \cup_{i=0}^3 M_i^\ell$. Using these facts

$$
\begin{aligned}
\alpha_{j,\ell+1} = \sum_{s=1}^q a_{js} \Big\{ & \mathrm{i}\Delta \frac{\tau_s^\ell}{\ell!} D_t^\ell u + \mathrm{i}\Delta \big[\alpha_{s\ell} - \frac{\tau_s^\ell}{\ell!} D_t^\ell u\big] \\
& + \mathrm{i}\lambda \sum_{m \in M_0^\ell} \frac{\tau_s^{m_1}}{m_1!} D_t^{m_1} u \, \frac{\tau_s^{m_2}}{m_2!} D_t^{m_2} u \, \frac{\tau_s^{m_3}}{m_3!} D_t^{m_3} \bar{u} \\
& + 2\mathrm{i}\lambda \sum_{m_1=p+1}^{\ell} \alpha_{sm_1} \sum_{m_2+m_3=\ell-m_1} \frac{\tau_s^{m_2}}{m_2!} D_t^{m_2} u \, \frac{\tau_s^{m_3}}{m_3!} D_t^{m_3} \bar{u} \\
& + \mathrm{i}\lambda \sum_{m_3=p+1}^{\ell} \bar{\alpha}_{sm_3} \sum_{m_1+m_2=\ell-m_3} \frac{\tau_s^{m_1}}{m_1!} D_t^{m_1} u \, \frac{\tau_s^{m_2}}{m_2!} D_t^{m_2} \bar{u} \Big\}.
\end{aligned} \tag{4.5}
$$

Now note that

$$\sum_{m_2+m_3=\ell-m_1} \frac{\tau_s^{m_2}}{m_2!} D_t^{m_2} u \, \frac{\tau_s^{m_3}}{m_3!} D_t^{m_3} \bar{u} = \frac{\tau_s^{\ell-m_1}}{(\ell-m_1)!} D_t^{\ell-m_1} |u|^2.$$

Using this and a similar identity for the last term in (4.5), we get

$$
\begin{aligned}
\alpha_{j,\ell+1} = \sum_{s=1}^q a_{js} \Big\{ & \mathrm{i}\Delta \frac{\tau_s^\ell}{\ell!} D_t^\ell u + \mathrm{i}\Delta \big[\alpha_{s\ell} - \frac{\tau_s^\ell}{\ell!} D_t^\ell u\big] \\
& + \mathrm{i}\lambda \sum_{|m|=\ell} \frac{\tau_s^{m_1}}{m_1!} D_t^{m_1} u \, \frac{\tau_s^{m_2}}{m_2!} D_t^{m_2} u \, \frac{\tau_s^{m_3}}{m_3!} D_t^{m_3} \bar{u} \\
& + 2\mathrm{i}\lambda \sum_{m_1=p+1}^{\ell} \big[\alpha_{sm_1} - \frac{\tau_s^{m_1}}{m_1!} D_t^{m_1} u\big] \frac{\tau_s^{\ell-m_1}}{(\ell-m_1)!} D_t^{\ell-m_1} |u|^2 \\
& + \mathrm{i}\lambda \sum_{m_3=p+1}^{\ell} \big[\bar{\alpha}_{sm_3} - \frac{\tau_s^{m_3}}{m_3!} D_t^{m_3} \bar{u}\big] \frac{\tau_s^{\ell-m_3}}{(\ell-m_3)!} D_t^{\ell-m_3} u^2 \Big\}.
\end{aligned}
$$

This is the componentwise form of (4.4). □

In the sequel we shall assume with no loss of generality that $p \leq \nu - 2$.

Lemma 4.2. *Assume that* (B), (C), *and* (D) *hold together with* (2.10a) *and* (2.10b). *Then, for each $\ell = 0, \dots, \nu-1$, using the notation of Lemma 4.1 we have*

$$b^T T^s \alpha_\ell = \frac{D_t^\ell u}{\ell!(s+\ell+1)}, \tag{4.6}$$

for every nonnegative integer s such that $s + \ell \leq \nu - 1$.

Proof. Assume $0 \leq \ell \leq p, 0 \leq s \leq \nu - 1$ with $s + \ell \leq \nu - 1$. From (4.2) and (B)

$$b^T T^s \alpha_\ell = b^T T^s \frac{T^\ell \boldsymbol{e}}{\ell!} D_t^\ell u = b^T \frac{T^{s+\ell} \boldsymbol{e}}{\ell!} D_t^\ell u = \frac{D_t^\ell u}{\ell!(s+\ell+1)}.$$

Now let $\ell = p+1$ and $0 \le s \le \nu - 1$ with $s+p+1 \le \nu - 1$. It follows from the inequalities $\nu \le \rho + p + 1$ and $s+p+1 \le \nu - 1$ that $s \le \rho - 1$; hence using (4.3), (B), and (D), we obtain

$$
\begin{aligned}
b^T T^s \alpha_{p+1} &= \frac{b^T T^s A T^p \boldsymbol{e}}{p!}\, D_t^{p+1} u \\
&= \frac{1}{(s+1)p!} b^T (1 - T^{s+1}) T^p \boldsymbol{e}\, D_t^{p+1} u \\
&= \frac{1}{(s+1)p!} \{ b^T T^p \boldsymbol{e} - T^{s+1+p} \boldsymbol{e} \} D_t^{p+1} u \\
&= \frac{1}{(s+1)p!} \Big\{ \frac{1}{p+1} - \frac{1}{s+p+2} \Big\} D_t^{p+1} u \\
&= \frac{D_t^{p+1} u}{(p+1)!(s+p+2)}.
\end{aligned}
$$

We now complete the proof using an induction argument; assume that (4.6) holds up to some ℓ with $p+1 \le \ell \le \nu - 2$. From (4.4) we have for $s = 0, \dots, \nu - 1$ with $s + \ell + 1 \le \nu - 1$,

$$
\begin{aligned}
b^T T^s \alpha_{\ell+1} = {} & \frac{b^T T^s A T^\ell \boldsymbol{e}}{\ell!} D_t^{\ell+1} u + \mathrm{i} \varDelta b^T T^s A \Big[\alpha_\ell - \frac{T^\ell \boldsymbol{e}}{\ell!}\, D_t^\ell u \Big] \\
& + 2\mathrm{i}\lambda \sum_{m_1 = p+1}^{\ell} \frac{D_t^{\ell - m_1} |u|^2}{(\ell - m_1)!} b^T T^s A \Big[T^{\ell - m_1} \alpha_{m_1} - D_t^{m_1} u\, \frac{T^\ell \boldsymbol{e}}{m_1!} \Big] \\
& + \mathrm{i}\lambda \sum_{m_3 = p+1}^{\ell} \frac{D_t^{\ell - m_3} u^2}{(\ell - m_3)!} b^T T^s A \Big[T^{\ell - m_3} \bar{\alpha}_{m_3} - D_t^{m_3} \bar{u}\, \frac{T^\ell \boldsymbol{e}}{m_3!} \Big].
\end{aligned}
\tag{4.7}
$$

As before, $s \le \rho - 1$; hence using (B) and (D),

$$
\begin{aligned}
b^T T^s A T^\ell \boldsymbol{e} &= \frac{1}{s+1} b^T (1 - T^{s+1}) T^\ell \boldsymbol{e} \\
&= \frac{1}{s+1} \{ b^T T^\ell \boldsymbol{e} - b^T T^{s+1+\ell} \boldsymbol{e} \} \\
&= \frac{1}{(\ell+1)(s+\ell+2)}.
\end{aligned}
$$

Therefore,

$$
\frac{b^T T^s A T^\ell \boldsymbol{e}}{\ell!} D_t^{\ell+1} u = \frac{1}{(\ell+1)!(s+\ell+2)} D_t^{\ell+1} u.
$$

Thus, to conclude the proof, it is sufficient to show that the last three terms in (4.7) are zero. This may be done in a similar way for all terms, so we consider here, as an example, only the second one. Indeed, from (B) and (D), for $p+1 \le m_1 \le \ell$, using the induction hypothesis we have

$$
\begin{aligned}
b^T T^s A T^{\ell-m_1}\alpha_{m_1} &= \frac{1}{s+1} b^T(1-T^{s+1})T^{\ell-m_1}\alpha_{m_1}\\
&= \frac{1}{s+1}\{b^T T^{\ell-m_1} - b^T T^{s+\ell-m_1+1}\}\alpha_{m_1}\\
&= \frac{1}{(s+1)}\Big\{\frac{1}{m_1!(\ell+1)} - \frac{1}{m_1!(s+\ell+2)}\Big\} D_t^{m_1}u\\
&= \frac{D_t^{m_1}u}{m_1!(\ell+1)(s+\ell+2)}.
\end{aligned}
$$

On the other hand, we have shown above that $b^T T^s A T^\ell \boldsymbol{e} = 1/(\ell+1)(s+\ell+2)$, so indeed the second term is zero. □

The key identity (4.6) is used to establish the following important cancellation property involving the α's.

Corollary 4.1. *Assume that* (B), (C) *and* (D) *hold together with* (2.10a) *and* (2.10b), *or that the RK method is the three-stage DIRK. Then*

$$
b^T A^{-1}\alpha_\ell = \frac{D_t^\ell u}{\ell!}, \quad \ell = 1, \dots, \nu. \tag{4.8}
$$

Proof. For $\ell = 1, \dots, p$ using (B), (C), and (4.2), we see that

$$
\begin{aligned}
b^T A^{-1}\alpha_\ell &= b^T A^{-1}\frac{T^\ell \boldsymbol{e}}{\ell!}\, D_t^\ell u\\
&= b^T A^{-1}\frac{A T^{\ell-1}\boldsymbol{e}}{(\ell-1)!}\, D_t^\ell u\\
&= \frac{D_t^\ell u}{\ell!}.
\end{aligned}
$$

For $\ell = p+1$, from (4.3) and (B),

$$
b^T A^{-1}\alpha_{p+1} = \frac{b^T T^p \boldsymbol{e}}{p!} D_t^{p+1}u = \frac{D_t^{p+1}u}{(p+1)!}.
$$

For $\ell = p+1, \dots, \nu-1$, from (4.4) we obtain

$$
\begin{aligned}
b^T A^{-1}\alpha_{\ell+1} = {}& \frac{b^T T^\ell \boldsymbol{e}}{\ell!} D_t^{\ell+1}u + \mathrm{i}\varDelta b^T\Big[\alpha_\ell - \frac{T^\ell \boldsymbol{e}}{\ell!}\, D_t^\ell u\Big]\\
&+ 2\mathrm{i}\lambda \sum_{m_1=p+1}^{\ell} \frac{D_t^{\ell-m_1}|u|^2}{(\ell-m_1)!} b^T\Big[T^{\ell-m_1}\alpha_{m_1} - D_t^{m_1}u\,\frac{T^\ell \boldsymbol{e}}{m_1!}\Big]\\
&+ \mathrm{i}\lambda \sum_{m_3=p+1}^{\ell} \frac{D_t^{\ell-m_3}u^2}{(\ell-m_3)!} b^T\Big[T^{\ell-m_3}\bar{\alpha}_{m_3} - D_t^{m_3}\bar{u}\,\frac{T^\ell \boldsymbol{e}}{m_3!}\Big].
\end{aligned}
$$

Using (B), the first term on the right-hand side gives $D_t^{\ell+1}u/(\ell+1)!$. On the other hand, the second, third, and fourth terms are zero by virtue of (4.6) and (B).

We finally consider briefly the special case of the three-stage DIRK method, for which, as we recall, $2p+2=4$ but $\rho+p+1=3$. We need only to verify that $b^TA^{-1}\alpha_4 = D_t^4u/24$. That this identity indeed holds can be seen by straightforward albeit long calculations and Lemma 4.1 in conjunction with the following three identities:

$$b^TTAT\boldsymbol{e} = \frac{1}{8}, \quad b^TAT^2\boldsymbol{e} = \frac{1}{12}, \quad b^TA^2T\boldsymbol{e} = \frac{1}{24}. \qquad \square$$

It follows from (4.1) that $\alpha_\ell, \ell=0,\dots,\nu-1$, are smooth if u, the solution of (1.1), is sufficiently smooth. Also, it will be very important for the analysis to follow that the α_ℓ's inherit the homogeneous Dirichlet boundary conditions of u, since we shall need to approximate them optimally by elements of S_h^r. Specifically, we have the following result.

Proposition 4.1. *Assume that u, the solution of* (1.1), *is in $C^\mu(\bar\Omega\times[0,t^*])$ for μ sufficiently large. Then, for each $n, 0\le n\le N-1$,*

(i) $\alpha_\ell\big|_{\partial\Omega}=0, \ell=0,\dots,\min\{p+3,\nu\}$. In fact, if $\nu>p+3, \alpha_{p+4}\big|_{\partial\Omega}\neq 0$ in general.

(ii) If Ω is a polyhedral domain or $d=1$, then

$$\alpha_\ell\big|_{\partial\Omega}=0, \quad \ell=0,\dots,\nu.$$

Proof. To establish (i) suppose, e.g., $p+3\le\nu$. Let $Q=\partial\Omega\times[0,t^*]$. Since $u\big|_Q = 0, \partial_t^j u\big|_Q=0, j=0,1,\dots$. Using the PDE in (1.1), we see that $\Delta u\big|_Q=0$, and hence $\Delta\partial_t^j u\big|_Q=0, j=0,1,\dots$. Since now

$$\begin{aligned}
\mathrm{i}\Delta^2 u &= \Delta u_t - \mathrm{i}\lambda\Delta(|u|^2u)\\
&= \Delta u_t - \mathrm{i}\lambda\big\{|u|^2\Delta u + u\Delta(|u|^2) + 2\bar u\nabla u\cdot\nabla u + 2u|\nabla u|^2\big\},
\end{aligned}$$

we see that $\Delta^2 u\big|_Q=0$, and hence $\Delta^2\partial_t^j u\big|_Q=0, j=0,1,\dots$.

We have thus shown that $\Delta^s\partial_t^j u\big|_Q=0, j=0,1,\dots, s=0,1,2$. Then, for each $n, 0\le n\le N-1$, it follows from (4.2) and (4.3) that $\Delta^s\alpha_\ell\big|_{\partial\Omega}=0, \ell=0,\dots,p+1, s=0,1,2$. Hence, by (4.1), it follows at once that $\alpha_{p+2}\big|_{\partial\Omega}=0$. Furthermore, from (4.1),

$$\begin{aligned}
\Delta\alpha_{j,p+2} &= \mathrm{i}\sum_{s=1}^q a_{js}\Big\{\Delta^2\alpha_{s,p+1} + \lambda\sum_{|m|=p+1}\Delta(\alpha_{sm_1}\alpha_{sm_2}\bar\alpha_{sm_3})\Big\}\\
&= \mathrm{i}\sum_{s=1}^q a_{js}\Big\{\Delta^2\alpha_{s,p+1} + \lambda\sum_{|m|=p+1}(\alpha_{sm_1}\alpha_{sm_2}\Delta\bar\alpha_{sm_3} + \alpha_{sm_2}\bar\alpha_{sm_3}\Delta\alpha_{sm_1} + \alpha_{sm_1}\bar\alpha_{sm_3}\Delta\alpha_{sm_2}\\
&\quad + 2\alpha_{sm_1}\nabla\alpha_{sm_2}\cdot\nabla\bar\alpha_{sm_3} + 2\alpha_{sm_2}\nabla\alpha_{sm_1}\cdot\nabla\bar\alpha_{sm_3} + 2\bar\alpha_{sm_3}\nabla\alpha_{sm_1}\cdot\nabla\alpha_{sm_2})\Big\},\\
&\qquad j=1,\dots,q.
\end{aligned}$$

We conclude that $\Delta\alpha_{p+2}\big|_{\partial\Omega}=0$, and hence by (4.1) that $\alpha_{p+3}\big|_{\partial\Omega}=0$.

A finer analysis shows in fact that the functions $\alpha_{j,p+4}$ (suppose $\nu>p+3$) do not vanish on $\partial\Omega$ for arbitrary solutions of (1.1) that are smooth up to $\partial\Omega$ and arbitrary IRK schemes that satisfy our stated assumptions. Indeed, suppose $\partial\Omega$ is the union of the closures of a finite number of open, smooth, connected subsets such as Γ in Figure 4.1.

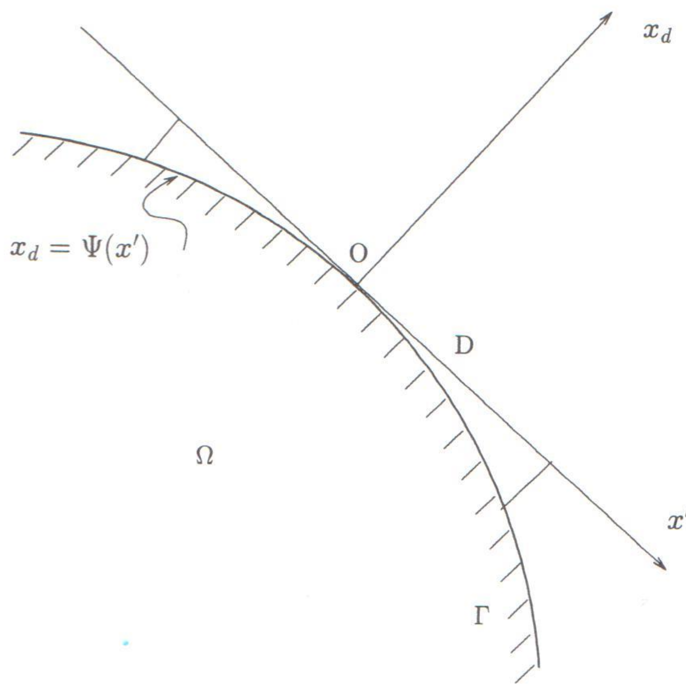

FIGURE 4.1.

Fix a point $x^0 \in \Gamma$ and consider a Cartesian coordinate system centered at x^0 with the x_d axis coinciding with the outer normal to $\partial\Omega$ at x^0. Since the PDE under consideration is invariant under translation and rotation and since the functions $\alpha_{j\ell}$ are also invariant under such changes, we may with no loss of generality assume that $x^0 = 0$ and that the coordinate system $(x_1, \dots, x_d)$, with respect to which (1.1) and the functions $\alpha_{j\ell}$ have been stated, is the one shown in Figure 4.1.

Again with no loss of generality assume that for some $\alpha > 0, \Gamma = \partial\Omega \cap \{x \in \mathbb{R}^d : |x| < \alpha\}$ and that Γ can be projected one-to-one along the Ox_d axis onto a region D of the tangent hyperplane $x_d = 0$. Assume that the equation of Γ is given by $x_d = \psi(x')$, for $x' = (x_1, \dots, x_{d-1}) \in D$, where $\psi \in C^j(\bar{D})$ for some j sufficiently large and $\psi(0) = 0, \partial\psi/\partial x_i(0) = 0, 1 \le i \le d-1$. Make now the change of variables $x \mapsto y$, defined by

$$\begin{aligned} y_i &= x_i, \qquad 1 \le i \le d-1, \\ y_d &= x_d - \psi(x'), \end{aligned}$$

and assume that Ω is such that $\Omega \cap \{x \in \mathbb{R}^d : |x| < \alpha\}$ is mapped one-to-one onto an open connected subset ω of the $y_d < 0$ halfspace in the $y-$space. Under this change of variables, Γ is flattened and mapped one-to-one onto D (which remains invariant) on the $y_d = 0$ hyperplane.

Under this change of independent variables, the PDE and the boundary condition in (1.1) are transformed into a local problem on $\bar{\omega}$ with homogeneous Dirichlet boundary condition on D. In addition, the problem of computing the boundary values on D of the transformed $\alpha_{j\ell}$ simplifies considerably. After a long computation (the details of which we omit) in the $y-$variables and transformation back to the $x-$variables, we obtain the formula

$$\alpha_{p+4}(0) = \Lambda \Delta\psi(0) A^3 \big(A T^p \boldsymbol{e} - \frac{T^{p+1}\boldsymbol{e}}{p+1} \big), \tag{4.9}$$

where

$$\Lambda = \Lambda(u,p) = -24\mathrm{i}\lambda \sum_{m \in M_0^{p+1}} \frac{1}{m!} D_t^{m_1}(\partial_d u) D_t^{m_2}(\partial_d u) D_t^{m_3}(\partial_d \bar{u}) \big|_{x=0}$$

and $\Delta\psi = \sum_{i=1}^{d-1} \partial_i^2 \psi$. (Here and in the sequel, we put $\partial_i v = -(\partial/\partial x_i)v, 1 \le i \le d$; the multi-integer set M_0^{p+1} was defined in the course of the proof of Lemma 4.1.)

For an arbitrary solution of (1.1) and arbitrary p, Λ is nonzero. (Take, e.g., p odd, $d = 2$, and solutions of (1.1) of the form $\mathrm{e}^{\mathrm{i}\beta t}\Phi(x_1, x_2)$.) In addition, we see by property (C) that $AT^p\boldsymbol{e} - T^{p+1}\boldsymbol{e}/(p+1)$ is nonzero. Hence $\alpha_{p+4}(0)$ is zero only when $\Delta\psi(0)$ vanishes, which can only be true (due to the arbitrary choice of x^0) if $\psi = 0$, i.e., when $\Gamma = D$. This leads us to consider, therefore, polyhedral domains for which we shall establish Proposition 4.1(ii) by a direct proof.

In the polyhedral case, no change of variables is required of course, and we simply let D be a $(d-1)-$dimensional face of $\bar{\Omega}$, on which we again orient the coordinate system so that the axis Ox_d is perpendicular to D. We let $Q' = D \times [0, t^*]$.

First we shall prove by induction that

$$\partial_d^{2\ell} u \big|_{Q'} = 0, \quad \ell = 0, \dots, \nu. \tag{4.10}$$

Since u vanishes on Q', (4.10) holds for $\ell = 0$. Assume that it is true up to some $\ell, 0 \le \ell \le \nu - 1$; we shall prove that $\partial_d^{2(\ell+1)} u\big|_{Q'} = 0$. From (1.1) we have

$$\begin{aligned} \partial_d^{2(\ell+1)} u &= \partial_d^{2\ell} \Big(\Delta u - \sum_{j=1}^{d-1} \partial_j^2 u \Big) \\ &= -\mathrm{i}\partial_d^{2\ell} u_t - \lambda \partial_d^{2\ell}(|u|^2 u) - \sum_{j=1}^{d-1} \partial_d^{2\ell}(\partial_j^2 u). \end{aligned}$$

Using the induction hypothesis, we see now that $\partial_d^{2\ell} u_t\big|_{Q'} = 0$, and

$$\sum_{j=1}^{d-1} \partial_d^{2\ell}(\partial_j^2 u)\big|_{Q'} = \sum_{j=1}^{d-1} \partial_j^2(\partial_d^{2\ell} u)\big|_{Q'} = 0.$$

By Leibniz's rule

$$\partial_d^{2\ell}(|u|^2 u) = \sum_{s_1+s_2+s_3=2\ell} \frac{(2\ell)!}{s_1! s_2! s_3!} \partial_d^{s_1} u \, \partial_d^{s_2} u \, \partial_d^{s_3} \bar{u}.$$

Since 2ℓ is even, at least one component s_i of each multi-integer $s=(s_1,s_2,s_3)$ with $|s|=2\ell$ is even. Therefore, by the induction hypothesis, the corresponding terms on the right-hand side of the sum above vanish on Q'. We conclude that $\partial_d^{2\ell}(|u|^2u)\big|_{Q'}=0$, which completes the inductive step for the proof of (4.10).

We can now prove that

$$\forall \ell,\quad 0\le \ell\le \nu,\quad \partial_d^{2\ell_1}\alpha_\ell\big|_D=0,\quad \forall \ell_1,\quad 0\le \ell_1\le \nu-\ell, \tag{4.11}$$

from which Proposition 4.1(ii) follows by taking $\ell_1=0$. It is obvious by (4.10) that (4.11) holds for $\ell=0$. Assume that it is valid up to some $\ell, 0\le \ell\le \nu-1$; we shall establish that it holds for $\ell+1$ as well. To this end, let ℓ_1 be an integer such that $0\le \ell_1\le \nu-(\ell+1)$. Then, for $1\le j\le q$,

$$\partial_d^{2\ell_1}\alpha_{j,\ell+1}=\mathrm{i}\sum_{s=1}^q a_{js}\Big(\partial_d^{2\ell_1}\Delta\alpha_{s,\ell}+\lambda\sum_{m_1+m_1+m_3=\ell}\partial_d^{2\ell_1}(\alpha_{sm_1}\alpha_{sm_2}\bar\alpha_{sm_3})\Big).$$

Now since $\partial_d^{2\ell_1}\Delta\alpha_\ell=\big(\partial_d^{2(\ell_1+1)}+\sum_{i=1}^{d-1}\partial_d^{2\ell_1}\partial_i^2\big)\alpha_\ell$ and by the induction hypothesis $\ell_1+1\le\nu-\ell, \partial_d^{2(\ell_1+1)}\alpha_\ell\big|_D=0, \partial_i^2(\partial_d^{2\ell_1}\alpha_\ell)\big|_D=0, 1\le i\le d-1$, we conclude, for $1\le s\le q$, that $\partial_d^{2\ell_1}\Delta\alpha_{s,\ell}\big|_D=0$.

On the other hand, for any $s, 1\le s\le q$, and multi-integer $m=(m_1,m_2,m_3)$ with $|m|=\ell$,

$$\partial_d^{2\ell_1}(\alpha_{sm_1}\alpha_{sm_2}\bar\alpha_{sm_3})=\sum_{\xi_1+\xi_2+\xi_3=2\ell_1}\frac{(2\ell_1)!}{\xi_1!\xi_2!\xi_3!}\partial_d^{\xi_1}\alpha_{sm_1}\,\partial_d^{\xi_2}\alpha_{sm_2}\,\partial_d^{\xi_3}\bar\alpha_{sm_3}.$$

Given $\xi=(\xi_1,\xi_2,\xi_3), |\xi|=2\ell_1$, at least one of its components, say ξ_1, is even. Since $m_1\le\ell, \xi_1\le 2\ell_1$, it follows that $\xi_1/2+m_1\le \ell_1+\ell\le\nu-1$. The induction hypothesis gives then that $\partial_d^{\xi_1}\alpha_{sm_1}\big|_D=0$. We conclude therefore that $\partial_d^{2\ell_1}\alpha_{j,\ell+1}\big|_D=0$ and complete the inductive step for the proof of (4.11). Needless to say, the proof is valid for $d=1$ as well. □

Henceforth the integer σ will be given by

$$\sigma=\begin{cases}\nu & \text{if } \varOmega \text{ is polyhedral or } d=1,\\ \min\{p+3,\nu\} & \text{otherwise.}\end{cases}$$

Now, given $n, 0\le n\le N-1$, define the (pseudo-)intermediate stages $u^{n,j}$ by

$$u^{n,j}=\sum_{\ell=0}^{\sigma}k^\ell\alpha_{j\ell},\quad j=1,\dots,q, \tag{4.12}$$

and u^{n+1} by

$$u^{n+1}=u(t^n)+b^TA^{-1}\big(\mathcal{U}^n-u(t^n)\boldsymbol{e}\big),\quad \text{with } \mathcal{U}^n=(u^{n,1},\dots,u^{n,q})^T. \tag{4.13}$$

Our temporal consistency result (the proof of which provides the motivation for the definition of the $\alpha_{j\ell}$) now follows. In Proposition 4.2 and in the sequel we shall denote by c, c_j, etc., generic positive constants independent of h and k but possibly depending on the solution and the data of (1.1).

Proposition 4.2. *Let the truncation errors $e^{n,j}, e^{n+1}$ be given by*

$$u^{n,j} = u(t^n) + k \sum_{s=1}^{q} a_{js} \big\{ \mathrm{i}\Delta u^{n,s} + \mathrm{i}\lambda |u^{n,s}|^2 u^{n,s} \big\} + e^{n,j}, \quad j = 1, \dots, q, \tag{4.14}$$

$$u^{n+1} = u(t^n) + k \sum_{j=1}^{q} b_j \big\{ \mathrm{i}\Delta u^{n,j} + \mathrm{i}\lambda |u^{n,j}|^2 u^{n,j} \big\} + e^{n+1}. \tag{4.15}$$

Then, under the hypotheses of Corollary 4.1 and Proposition 4.1, we have

$$\sum_{j=1}^{q} \|e^{n,j}\|_m + \|e^{n+1}\|_m \le c_m k^{\sigma+1}, \quad m = 0, 1, \dots, \tag{4.16}$$

and

$$\|u(t^{n+1}) - u^{n+1}\|_m \le c_m k^{\sigma+1}, \quad m = 0, 1, \dots . \tag{4.17}$$

Proof. We have

$$\begin{aligned} e^{n,j} &= \sum_{\ell=0}^{\sigma} k^\ell \alpha_{j\ell} - u(t^n) - k \sum_{s=1}^{q} a_{js} \Big\{ \mathrm{i}\Delta \sum_{\ell=0}^{\sigma} k^\ell \alpha_{s\ell} + \mathrm{i}\lambda \Big| \sum_{\ell=0}^{\sigma} k^\ell \alpha_{s\ell} \Big|^2 \sum_{\ell=0}^{\sigma} k^\ell \alpha_{s\ell} \Big\} \\ &= \sum_{\ell=1}^{\sigma} k^\ell \alpha_{j\ell} - k \sum_{s=1}^{q} a_{js} \Big\{ \mathrm{i}\Delta \sum_{\ell=0}^{\sigma} k^\ell \alpha_{s\ell} + \mathrm{i}\lambda \sum_{\ell=0}^{\sigma-1} k^\ell \sum_{|m|=\ell} \alpha_{sm_1} \alpha_{sm_2} \bar{\alpha}_{sm_3} \\ &\qquad\qquad + \mathrm{i}\lambda \sum_{\ell=\sigma}^{3\sigma} k^\ell \sum_{|m|=\ell} \delta_m \alpha_{sm_1} \alpha_{sm_2} \bar{\alpha}_{sm_3} \Big\}, \end{aligned}$$

where the constants δ_m are zero or 1. Now, using the defining relations (4.1),

$$e^{n,j} = -\sum_{s=1}^{q} a_{js} \big\{ \mathrm{i}k^{\sigma+1} \Delta \alpha_{s\sigma} + \mathrm{i}\lambda \sum_{\ell=\sigma}^{3\sigma} k^{\ell+1} \sum_{|m|=\ell} \delta_m \alpha_{sm_1} \alpha_{sm_2} \bar{\alpha}_{sm_3} \big\}.$$

Hence $e^{n,j}$ satisfies the bound (4.16). As for e^{n+1}, using (4.13) and (4.14) in (4.15),

$$\begin{aligned} e^{n+1} &= u^{n+1} - u(t^n) - k \sum_{j=1}^{q} b_j \big\{ \mathrm{i}\Delta u^{n,j} + \mathrm{i}\lambda |u^{n,j}|^2 u^{n,j} \big\} \\ &= u^{n+1} - u(t^n) - \sum_{j=1}^{q} b_j \sum_{s=1}^{q} a_{js}^{-1} \{ u^{n,s} - u(t^n) - e^{n,s} \} \\ &= \sum_{j,s=1}^{q} b_j a_{js}^{-1} e^{n,s}, \end{aligned}$$

where $a_{js}^{-1} = (A^{-1})_{js}$; (4.16) follows. Now using (4.13) and (4.8),

$$\begin{aligned} u^{n+1} &= u(t^n) + b^T A^{-1} \sum_{\ell=1}^{\sigma} k^\ell \alpha_\ell \\ &= \sum_{\ell=0}^{\sigma} k^\ell \frac{D_t^\ell u(t^n)}{\ell!}. \end{aligned}$$

By Taylors theorem (4.17) follows. □

Remark 4.1. From the definition of $u^{n,j}$ and $\alpha_{j\ell}$, it follows that, for k sufficiently small, $u^{n,j}$ will satisfy $\|u^{n,j}\|_{L^\infty} \le 2M(u)$. Hence the conclusion of Proposition 4.2 also holds if we replace$|u^{n,j}|^2 u^{n,j}$ by $\tilde{g}(u^{n,j})$ in (4.14) and (4.15).

Henceforth we shall let $\omega^{n,j}, \omega(t^n)$ and ω^{n+1} stand for $P_E u^{n,j}, P_E u(t^n)$ and $P_E u^{n+1}$, respectively. The following lemma, which establishes the spatial and temporal consistency of the methods, uses the results of Proposition 4.2.

Lemma 4.3. *Let $\eta^{n,j}, j = 1, \dots, q, \eta^{n+1}$ in S_h^r be given by*

$$\omega^{n,j} = \omega(t^n) + k \sum_{s=1}^{q} a_{js} \tilde{f}_h(\omega^{n,s}) + \eta^{n,j}, \quad j = 1, \dots, q, \tag{4.18}$$

$$\omega^{n+1} = \omega(t^n) + k \sum_{j=1}^{q} b_j \tilde{f}_h(\omega^{n,j}) + \eta^{n+1}. \tag{4.19}$$

Then, under the hypotheses of Proposition 4.1, we have

$$\sum_{j=1}^{q} \|\eta^{n,j}\| \le ck\{h^r + k^\sigma\}, \tag{4.20}$$

$$\|\eta^{n+1}\| \le ck\{h^r + k^\sigma\}. \tag{4.21}$$

Proof. Using (2.15), (2.5), and (2.20) we have for $\chi \in S_h^r$

$$(\eta^{n,j}, \chi) = (\omega^{n,j} - \omega(t^n), \chi) - k \sum_{s=1}^{q} a_{js} \big\{ \mathrm{i}(\Delta u^{n,s}, \chi) + \mathrm{i}\lambda(\tilde{g}(\omega^{n,s}), \chi) \big\}.$$

Thus, using (4.14), from Remark 4.1, it follows that

$$\begin{aligned} (\eta^{n,j}, \chi) &= ([\omega^{n,j} - u^{n,j}] - [\omega(t^n) - u(t^n)], \chi) \\ &\quad - \mathrm{i}k\lambda \sum_{s=1}^{q} a_{js} \big(\tilde{g}(\omega^{n,s}) - \tilde{g}(u^{n,s}), \chi \big) + (e^{n,j}, \chi). \end{aligned} \tag{4.22}$$

Now $\omega^{n,j} - u^{n,j} - [\omega(t^n) - u(t^n)] = (P_E - I) \sum_{\ell=1}^{\sigma} \alpha_{j\ell}$; hence from (2.6) and Proposition 4.1,

$$\|\omega^{n,j} - u^{n,j} - [\omega(t^n) - u(t^n)]\| \le ckh^r. \tag{4.23}$$

Furthermore, since the map $\tilde{g}$ is Lipschitz,

$$\|\tilde{g}(\omega^{n,s}) - \tilde{g}(u^{n,s})\| \le c\|\omega^{n,s} - u^{n,s}\| \le ch^r, \quad 1 \le s \le q. \tag{4.24}$$

Letting $\chi = \eta^{n,j}$ in (4.22), it follows from (4.23) and (4.24) that

$$\|\eta^{n,j}\| \le ckh^r + \|e^{n,j}\|, \quad j = 1, \dots, q.$$

Hence, (4.20) follows from (4.16). Now using (4.18) and (4.19)

$$\begin{aligned}\eta^{n+1} &= \omega^{n+1} - \omega(t^n) - k \sum_{j=1}^{q} b_j \tilde{f}_h(\omega^{n,j}) \\ &= \omega^{n+1} - \omega(t^n) - \sum_{j,s=1}^{q} b_j a_{js}^{-1} \{\omega^{n,s} - \omega(t^n) - \eta^{n,s}\} \\ &= \sum_{j,s=1}^{q} b_j a_{js}^{-1} \eta^{n,s};\end{aligned}$$

(4.21) now follows from (4.20). ☐

It will also be necessary to compare $\omega^{n,j}, \omega^{n+1}$ with an exact solution of a Runge–Kutta step of the form (2.21), (2.22) that has $\omega(t^n)$ as initial value.

Lemma 4.4. *Assume that the IRK method satisfies* (P). *Then, under the hypotheses of Proposition 4.1, there exist* $v^{n,j}, j = 1, \dots, q, v^{n+1}$ *in* S_h^r *satisfying*

$$v^{n,j} = \omega(t^n) + k \sum_{s=1}^{q} a_{js} \tilde{f}_h(v^{n,s}), \quad j = 1, \dots, q, \tag{4.25}$$

$$v^{n+1} = \omega(t^n) + k \sum_{j=1}^{q} b_j \tilde{f}_h(v^{n,j}). \tag{4.26}$$

Furthermore,

$$\|\omega^{n,j} - v^{n,j}\| \le ck\{h^r + k^\sigma\}, \quad j = 1, \dots, q, \tag{4.27}$$

$$\|\omega^{n+1} - v^{n+1}\| \le ck\{h^r + k^\sigma\}. \tag{4.28}$$

Proof. The existence of $v^{n,j}$ and hence that of v^{n+1} follows at once from Proposition 3.1. Letting $\zeta^{n,j} = \omega^{n,j} - v^{n,j}$, from (4.25) and (4.18) we obtain

$$\sum_{j,s=1}^{q} c_{js} d_j d_s (\zeta^{n,s}, \zeta^{n,j}) = k \sum_{j=1}^{q} d_j^2 (\tilde{f}_h(\omega^{n,j}) - \tilde{f}_h(v^{n,j}), \zeta^{n,j}) + \sum_{s,j=1}^{q} a_{js}^{-1} d_j^2 (\eta^{n,s}, \zeta^{n,j}).$$

Since $\tilde{g}$ is Lipschitz, using (3.1) we see that

$$\sum_{j=1}^{q} \|\zeta^{n,j}\|^2 \le ck \sum_{j=1}^{q} \|\zeta^{n,j}\|^2 + c \sum_{j,s=1}^{q} \|\eta^{n,s}\| \, \|\zeta^{n,j}\|.$$

Using the Cauchy–Schwarz inequality and (4.20), we get (4.27) for k sufficiently small. Now from (4.18), (4.19), (4.25), and (4.26) we obtain

$$\omega^{n+1} - v^{n+1} = \sum_{j,s=1}^{q} b_j a_{js}^{-1} (\omega^{n,s} - v^{n,s} - \eta^{n,s}) + \eta^{n+1}.$$

Inequality (4.28) now follows from this, (4.27) and (4.21). ☐

We now prove our main *stability* result.

Lemma 4.5. *Let* $\tilde u_h^{n+1}, v^{n+1}$ *be given by* (2.21) *and* (4.26)*, respectively, and the IRK method satisfy* (S)*. Then*

$$\|\tilde u_h^{n+1} - v^{n+1}\| \le (1+ck)\|\tilde u_h^n - \omega(t^n)\|. \tag{4.29}$$

Proof. Let $\varepsilon^{n,j} = v^{n,j} - \tilde u_h^{n,j}$ and $\delta \tilde f_h^j = \tilde f_h(v^{n,j}) - \tilde f_h(\tilde u_h^{n,j})$. Subtracting (4.26) from (2.21) and taking inner products, it follows that

$$\|v^{n+1} - \tilde u_h^{n+1}\|^2 = \|\omega(t^n) - \tilde u_h^n\|^2 + k\sum_{j=1}^q b_j(\delta\tilde f_h^j, \omega(t^n) - \tilde u_h^n)$$
$$+ k\sum_{j=1}^q b_j(\omega(t^n) - \tilde u_h^n, \delta\tilde f_h^j) + k^2\sum_{j,s=1}^q b_j b_s(\delta\tilde f_h^j, \delta\tilde f_h^s).$$

Subtracting (2.22) from (4.25), we obtain $\omega(t^n) - \tilde u_h^n = \varepsilon^{n,j} - k\sum_{s=1}^q a_{js}\delta\tilde f_h^s$. Therefore,

$$\|v^{n+1} - \tilde u_h^{n+1}\|^2 = \|\omega(t^n) - \tilde u_h^n\|^2 + 2k\,\mathrm{Re}\sum_{j=1}^q b_j(\delta\tilde f_h^j, \varepsilon^{n,j}) - k^2\sum_{j,s=1}^q b_j a_{js}(\delta\tilde f_h^j, \delta\tilde f_h^s)$$
$$- k^2\sum_{j,s=1}^q b_j a_{js}(\delta\tilde f_h^s, \delta\tilde f_h^j) + k^2\sum_{j,s=1}^q b_j b_s(\delta\tilde f_h^j, \delta\tilde f_h^s)$$
$$= \|\omega(t^n) - \tilde u_h^n\|^2 + 2k\,\mathrm{Re}\sum_{j=1}^q b_j(\delta\tilde f_h^j, \varepsilon^{n,j}) - k^2\sum_{j,s=1}^q m_{js}\,\mathrm{Re}(\delta\tilde f_h^j, \delta\tilde f_h^s),$$

where $\{m_{js}\}$ is the array pertaining to (S); given that it is positive semidefinite, we get

$$\|v^{n+1} - \tilde u_h^{n+1}\|^2 \le \|\omega(t^n) - \tilde u_h^n\|^2 + 2k\,\mathrm{Re}\sum_{j=1}^q b_j(\delta\tilde f_h^j, \varepsilon^{n,j}).$$

On the other hand, since $B_h^r(\varepsilon^{n,j}, \varepsilon^{n,j})$ is real and $\tilde g$ is Lipschitz, we obtain

$$\|v^{n+1} - \tilde u_h^{n+1}\|^2 \le \|\omega(t^n) - \tilde u_h^n\|^2 + ck\sum_{j=1}^q \|\varepsilon^{n,j}\|^2.$$

Now using (2.22), (4.25), and (P) together with previously used techniques,

$$\sum_{j=1}^q \|\varepsilon^{n,j}\|^2 \le c\|\omega(t^n) - \tilde u_h^n\|^2. \tag{4.30}$$

Using this in the above yields (4.29). ☐

We are now ready to state and prove the main convergence result.

Theorem 4.1. *Under our hypotheses on the IRK methods and if* $u \in C^\mu(\bar\Omega \times [0, t^*])$ *for* μ *sufficiently large, there exists a constant* c *independent of* k *and* h *such that*

$$\max_{0\le n\le N} \|\tilde u_h^n - u(t^n)\| \le c\{h^r + k^\sigma\}. \tag{4.31}$$

Assume that $d < 2r$ and that $k = o(h^{d/2\sigma})$ as $h \to 0$. Then, for h sufficiently small,

$$\max_{0\le n\le N} \|u_h^n - u(t^n)\| \le c\{h^r + k^\sigma\}. \tag{4.32}$$

Proof. Let $\omega(t^{n+1}) = P^E u(t^{n+1})$. We will first show that

$$\|\omega(t^{n+1}) - v^{n+1}\| \le ck\{h^r + k^\sigma\}. \tag{4.33}$$

Since $\omega(t^{n+1}) - v^{n+1} = \omega(t^{n+1}) - \omega^{n+1} + \omega^{n+1} - v^{n+1}$, it is sufficient, in view of (4.28), to show that

$$\|\omega(t^{n+1}) - \omega^{n+1}\| \le ck\{h^r + k^\sigma\}. \tag{4.34}$$

Writing $\omega(t^{n+1}) - \omega^{n+1} = (I - P_E)\big(u^{n+1} - u(t^{n+1})\big) + u(t^{n+1}) - u^{n+1}$, and recalling that u^{n+1} is smooth and $u^{n+1}\big|_{\partial\Omega} = 0$, we see that (4.34) follows from (2.6) and (4.17). Now from (4.33) and (4.29) we get

$$\begin{aligned}\|\omega(t^{n+1}) - \tilde u_h^{n+1}\| &\le \|\omega(t^{n+1}) - v^{n+1}\| + \|v^{n+1} - \tilde u_h^{n+1}\| \\ &\le ck\{h^r + k^\sigma\} + (1 + ck)\|\omega(t^n) - \tilde u_h^n\|.\end{aligned}$$

From this it follows that

$$\|\omega(t^n) - \tilde u_h^n\| \le c\mathrm{e}^{ct^n}\big\{\|\omega(0) - \tilde u_h^0\| + h^r + k^\sigma\big\}, \quad n = 0, \dots, N. \tag{4.35}$$

Hence, (4.35), (2.6), and (2.11) establish (4.31), since

$$\begin{aligned}\|u(t^n) - \tilde u_h^n\| &\le \|u(t^n) - \omega(t^n)\| + \|\omega(t^n) - \tilde u_h^n\| \\ &\le c\{h^r + k^\sigma\}.\end{aligned}$$

Now, using (4.30), (4.35), (4.27), and the definition of $\omega^{n,j}$, we have for all $j, 1 \le j \le q$,

$$\begin{aligned}\|\tilde u_h^{n,j} - u^{n,j}\| &\le \|\tilde u_h^{n,j} - v^{n,j}\| + \|v^{n,j} - \omega^{n,j}\| + \|\omega^{n,j} - u^{n,j}\| \\ &\le c(h^r + k^\sigma).\end{aligned} \tag{4.36}$$

Fix j. Then, letting χ be an element of S_h^r that satisfies (2.1) and (2.2) for $u^{n,j}$, we have by (2.3) and (4.36),

$$\begin{aligned}\|\tilde u_h^{n,j} - u^{n,j}\|_{L^\infty} &\le \|\tilde u_h^{n,j} - \chi\|_{L^\infty} + \|\chi - u^{n,j}\|_{L^\infty} \\ &\le ch^{-d/2}\|\tilde u_h^{n,j} - \chi\| + \|\chi - u^{n,j}\|_{L^\infty} \\ &\le ch^{-d/2}\big(\|\tilde u_h^{n,j} - u^{n,j}\| + \|u^{n,j} - \chi\|\big) + \|u^{n,j} - \chi\|_{L^\infty} \\ &\le ch^{-d/2}(h^r + k^\sigma) + h^2.\end{aligned}$$

It follows by our hypotheses that given $\varepsilon > 0$, there exists h_0 such that for $h \le h_0$, $\max_{j,n} \|\tilde u_h^{n,j} - u^{n,j}\|_{L^\infty} \le \varepsilon$. Hence for ε, k sufficiently small

$$\begin{aligned}\max_{j,n} \|\tilde u_h^{n,j}\|_{L^\infty} &\le \max_{j,n} \|\tilde u_h^{n,j} - u^{n,j}\|_{L^\infty} + \max_{j,n} \|u^{n,j}\|_{L^\infty} \\ &\le \varepsilon + \max_n \|u(t^n)\|_{L^\infty} + ck \le 2M(u).\end{aligned}$$

By local uniqueness we have therefore that $u_h^n = \tilde u_h^n$ and (4.32) follows. ∎

5. Remarks

In this section we shall briefly indicate how the techniques and results of the present paper extend to cases of some related PDEs. Our objective is to use these extensions to illuminate the capabilities and the limitations of our approach to the consistency of the fully discrete schemes using the functions $\alpha_{j\ell}$.

Our results can be easily extended first to analogous to (1.1) initial and boundary value problems for the nonlinear Schrödinger equation with a general power nonlinearity [12], obtained by replacing the nonlinear term $|u|^2u$ in the PDE in (1.1) by $|u|^{2\beta}u$, $\beta \geq 1$ integer. The results of §3 carry over immediately while the functions $\alpha_{j\ell}$ must now be redefined in the obvious fashion to correspond to the new nonlinearity. It may be easily established that the α_ℓ vanish on $\partial\Omega$ for $0 \leq \ell \leq \sigma$, where σ is now given by

$$\sigma = \begin{cases} \nu & \text{if } \Omega \text{ is polyhedral or } d = 1, \\ \min\{p + \beta + 2, \nu\} & \text{otherwise.} \end{cases} \tag{5.1}$$

The key observation that gets us up to the classical order in the polyhedral case is that in applying Leibniz's rule for evaluating $\partial_d^{2\ell}(|u|^{2\beta}u)$ in the course of the proof of the analogue of (4.10), the sum is taken over all multi-integers $s = (s_1, \ldots, s_{2\beta+1})$ [i.e., with an *odd* number of components (as in the cubic case)] that satisfy $|s| = 2\ell$. Hence, for each such s, at least one of the s_i is even and the corresponding term in the sum vanishes on $\partial\Omega$, thus allowing the inductive step to be completed (similarly with the analogue of ((4.11)). The rest of the results of §4 are easily established mutatis mutandis; in particular, the optimal order error estimate (4.32) is valid with σ given by (5.1).

The results above should be contrasted with those that may be obtained in the case of the analogous to (1.1) initial and boundary value problem for the *semilinear heat equation* with a power nonlinearity given by

$$u_t = \Delta u + \lambda u^\gamma. \tag{5.2}$$

The solution u is now real valued of course, λ is a constant, and $\gamma \geq 2$ is an integer. For the purposes of the consistency proof, $\alpha_{j\ell}$ can be again easily constructed to correspond to (5.2) and a straightforward computation as in the proof of Proposition 4.1(i) yields now that on a general domain the $\alpha_{j\ell}$ vanish on $\partial\Omega$ for $0 \leq \ell \leq \min\{p + [\gamma + 3/2], \nu\}$. However, if Ω is polyhedral, we may go up to ν mimicking the proof of Proposition 4.1(ii), only when γ is odd. In summary, in the case of (5.2) we may show that $\alpha_{j\ell}\big|_{\partial\Omega} = 0$ for $0 \leq \ell \leq \sigma$, where now

$$\sigma = \begin{cases} \nu & \text{if } \Omega \text{ is polyhedral (or } d = 1) \text{ and } \gamma \text{ is odd,} \\ \min\{p + [\frac{\gamma + 3}{2}], \nu\} & \text{otherwise.} \end{cases} \tag{5.3}$$

There is another difference between the error analysis for the cubic Schrödinger equation and that appropriate to (5.2). Unless λ is negative and γ is odd, it is no longer possible to establish a priori existence and boundedness (cf. Propositions 3.1 and 3.2) of the fully discrete approximations $u_h^{n,i}, u_h^n$, solutions of the analogues of

(2.17) and (2.18). However, we easily obtain existence and L^2-boundedness for the $\tilde{u}_h^n$ that satisfy the analogues of (2.21) and (2.22) defined by constructing in the standard way a globally Lipschitz map $\tilde{g}$ associated with $g(u) = u^\gamma$. In addition, local uniqueness again holds in the sense that two fully discrete solutions with components in the ball $\{v \in S_h^r : \|v\|_{L^\infty} \leq cM(u)\}$ (where $\tilde{g}(v)$ and $g(v)$ coincide for $|v| < cM(u)$) are identical. Mimicking the analysis of §4 we may at the end establish, as in the proof of Theorem 4.1, that $\max_{n,j} \|\tilde{u}_h^{n,j}\|_{L^\infty} \leq cM(u)$ and conclude by local uniqueness that $u_h^n = \tilde{u}_h^n$; an error estimate of the form (4.31), where σ is given by (5.3), follows.

References

1. G. D. Akrivis and V. A. Dougalis, *On a class of conservative, highly accurate Galerkin methods for the Schrödinger equation*, RAIRO Modél. Math. Anal. Numér. **25** (1991) 643–670.
2. G. D. Akrivis, V. A. Dougalis, and O. A. Karakashian, *On fully discrete Galerkin methods of second-order temporal accuracy for the nonlinear Schrödinger equation*, Numer. Math. **59** (1991) 31–53.
3. J. H. Bramble and J. E. Osborn, *Rate of convergence estimates for nonselfadjoint eigenvalue approximations*, Math. Comp. **27** (1973) 525–549.
4. P. Brenner, M. Crouzeix, and V. Thomée, *Single step methods for inhomogeneous linear differential equations in Banach space*, RAIRO Anal. Numér. **16** (1982) 5–26.
5. J. C. Butcher, *Implicit Runge–Kutta processes*, Math. Comp. **18** (1964) 50–64.
6. M. Crouzeix, *Sur l'approximation des équations différentielles opérationelles linéaires par des méthodes de Runge–Kutta*, Thèse, Université Paris VI, 1975.
7. M. Crouzeix, *Sur la B-stabilité des méthodes de Runge–Kutta*, Numer. Math. **32** (1979) 75–82.
8. M. Crouzeix, W. H. Hundsdorfer, and M. N. Spijker, *On the existence of solutions to the algebraic equations in implicit Runge–Kutta methods*, BIT **23** (1983) 84–91.
9. K. Dekker and J. G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, Amsterdam, 1984.
10. V. A. Dougalis and O. A. Karakashian, *On some high order accurate fully discrete Galerkin methods for the Korteweg–de Vries equation*, Math. Comp. **45** (1985) 329–345.
11. O. Karakashian and W. McKinney, *On optimal high order in time approximations for the Korteweg–de Vries equation*, Math. Comp. **55** (1990) 473–496.
12. J. J. Rasmussen and K. Rypdal, *Blow-up in Nonlinear Schrödinger equations*–I: *A general review*, Phys. Scripta **33** (1986) 481–497.
13. W. A. Strauss, *The nonlinear Schrödinger equation*, In: Contemporary Developments in Continuum Mechanics and Partial Differential Equations, G. M. de la Penha and L. A. J. Medeiros, eds., pp. 452–465. North–Holland, New York, 1978.

Mathematics Department, University of Tennessee, Knoxville, Tennessee 37996-1300, USA

Mathematics Department, University of Crete, 714 09 Heraklion, Crete, Greece

Mathematics Department, National Technical University, Zographou, 157 80 Athens, Greece