HIGH–ORDER FINITE ELEMENT METHODS FOR THE KURAMOTO–SIVASHINSKY EQUATION

GEORGIOS AKRIVIS

Abstract. We discretize the periodic initial–value problem for the Kuramoto–Sivashinsky equation by implicit Runge–Kutta methods in time combined with the Galerkin–finite element method in space. Optimal–order error estimates are established and the linearization of the schemes is also discussed.

Résumé. Nous considérons l'équation de Kuramoto–Sivashinsky munie de conditions aux limites périodiques et d'une donnée initiale. Nous l'approchons en utilisant une méthode d'éléments finis de type Galerkin pour la discrétisation en espace, et un schéma de Runge–Kutta implicite pour la discrétisation en temps. Nous obtenons des estimations d'erreur optimales et discutons de la linéarisation de cette méthode.

1. Introduction

In this paper we shall analyze high–order finite element approximations to the solution of the following periodic initial–value problem for the Kuramoto–Sivashinsky (KS) equation: For $t^*, \nu > 0$, we seek a real–valued function u defined on $\mathbb{R} \times [0, t^*]$, 1–periodic in the space variable and satisfying

$$u_t + uu_x + u_{xx} + \nu u_{xxxx} = 0 \qquad \text{in} \quad \mathbb{R} \times [0, t^*] \tag{1.1}$$

and

$$u(\cdot, 0) = u^0 \qquad \text{in} \quad \mathbb{R}, \tag{1.2}$$

where u^0 is a given 1–periodic function. An alternative form of the KS equation is obtained through the change of variables $v(x,t) := \sqrt{\nu}\, u(\sqrt{\nu}x, \nu t)$, namely

$$v_t + vv_x + v_{xx} + v_{xxxx} = 0 \qquad \text{in} \quad \mathbb{R} \times [0, \frac{t^*}{\nu}]. \tag{1.1$'$}$$

The function v is obviously periodic in the space variable with period $\frac{1}{\sqrt{\nu}}$. It is shown in [13] and in [22] that the periodic initial–value problem for the KS equation is well–posed; in particular for $u^0 \in H^2_{\rm per}$ there exists a unique solution u of (1.1)–(1.2), $u(\cdot,t) \in H^2_{\rm per}$, and $u(\cdot,t)$ depends continuously on the initial data. Here, for $m \in \mathbb{N}, H^m_{\rm per}$ denotes the periodic Sobolev space of order m, consisting of the 1–periodic elements of $H^m_{\rm loc}(\mathbb{R})$, and $\|\cdot\|_m$ is the norm over a period in $H^m_{\rm per}$. The inner product in $L^2(0,1)$ is denoted by $(\cdot,\cdot)$, and the induced norm by $\|\cdot\|$. In the sequel we assume existence of a solution u of (1.1)–(1.2), which is smooth enough for our purposes.

Work supported by the Institute of Applied and Computational Mathematics of the Research Center of Crete – FO.R.T.H..

The KS equation was independently derived by Kuramoto and Sivashinsky in the late 70's, and is related to turbulence phenomena in chemistry and combustion, cf. [12], [21]; it also arises in a variety of other physical problems such as plasma physics and two–phase flows in cylindrical geometries, cf. [18]. See also, for instance, [5], [7], [8], [11], [13], [14] and [19], for various interesting properties of the KS equation and for related computational work. We refer the reader to Temam [23] for an overview.

In [1] the discretization of (1.1)–(1.2) by a Crank–Nicolson finite difference scheme and a linearization thereof by Newton's method is studied. In [2] the semidiscretization of (1.1)–(1.2) by the standard Galerkin–finite element method as well as the discretization of the resulting initial–value problem by the Crank–Nicolson method is considered.

In this paper we analyze the discretization of (1.1)–(1.2) by implicit Runge–Kutta (RK) methods in time combined with the standard Galerkin–finite element method in space. For a suitable class of algebraically stable implicit RK methods we shall show

$$\max_{0\le n\le N} \|u(\cdot,t^n)-U^n\| \le c\,(k^\sigma+h^r) \tag{1.3}$$

where σ is the classical order of accuracy of the RK method and r is the optimal spatial rate of convergence in L^2; $k=t^*/N$ is the time step, $t^n:=nk, n=0,\dots,N$, and h is the spatial discretization parameter, $U^0,\dots,U^N$ are the RK approximations and U^0 is supposed to approximate u^0 to optimal order in L^2. Some mild mesh conditions are required for (1.3) to hold. A slight modification of the results of [3] yields linearizations of the RK schemes which preserve the overall accuracy of the methods.

Our approach is similar to the one in [10], [9] where optimal–order error estimates for the Korteweg–de Vries equation and the cubic Schrödinger equation, respectively, are derived. In contrast to [9], due to the spatial derivative in the nonlinearity of the KS equation, we use a *time–dependent* elliptic projection operator in order to obtain optimal–order estimates in the spatial discretization parameter h. The quasi–interpolant of u, cf. [24], can be used instead of the elliptic projection in the error analysis, if one is willing to restrict himself to smooth splines on uniform partitions, cf. [10]. Note however that the quasi–interpolant technique allows also $r=3$, i.e., quadratic splines, while in this paper we will assume $r\ge 4$.

An outline of the remaining part of the paper is as follows: In section 2 we introduce a time–dependent elliptic projection and derive some error estimates which play an important role in the sequel. Section 3 is devoted to the time–stepping by a suitable class of implicit RK schemes. Under some mild mesh restrictions, we establish optimal–order error estimates and prove uniqueness of the fully discrete approximations. In the last section we briefly discuss the linearization of the fully discrete methods by an explicit–implicit procedure which retains the order of convergence.

2. An elliptic projection

In this section we introduce a time–dependent elliptic projection operator and derive some estimates which will be useful in §3.

We shall discretize (1.1)–(1.2) in space by the standard Galerkin method. To this effect, let $0 = x_0 < x_1 < \cdots < x_J = 1$ be a partition of $[0,1]$, $h := \max_j (x_{j+1} - x_j)$ and $\underline{h} := \min_j (x_{j+1} - x_j)$. Setting $x_{jJ+s} := x_s,\ j \in \mathbb{Z},\ s = 0, \dots, J-1$, this partition is periodically extended to a partition of $\mathbb{R}$. For integer $r \geq 4$, let S_h^r denote a space of at least once continuously defferentiable, 1–periodic splines of degree $r-1$, in which approximations to the solution $u(\cdot, t)$ of (1.1)–(1.2) will be sought for $0 \leq t \leq t^*$. The following approximation property of the family $\{S_h^r\}_{0<h<1}$ is well known

$$\inf_{\chi \in S_h^r} \sum_{j=0}^{2} h^j \, \|v - \chi\|_j \leq c h^s \|v\|_s,\ v \in H^s_{\text{per}}, \quad 2 \leq s \leq r, \tag{2.1}$$

with a constant c independent of v and h, cf., e.g., [20, §8.1].

Motivated by the following variational formulation of the KS equation

$$(u_t, v) + (u u_x, v) - (u_x, v') + \nu(u_{xx}, v'') = 0 \quad \forall v \in H^2_{\text{per}}, \quad 0 \leq t \leq t^*, \tag{2.2}$$

we define the *semidiscrete approximation* $\tilde{u}_h(\cdot, t) \in S_h^r,\ \ 0 \leq t \leq t^*$, to u by

$$(\tilde{u}_{ht}, \chi) + (\tilde{u}_h \tilde{u}_{hx}, \chi) \ - \ (\tilde{u}_{hx}, \chi') \ + \ \nu(\tilde{u}_{hxx}, \chi'') = 0 \quad \forall \chi \in S_h^r, \quad 0 \leq t \leq t^*, \tag{2.3}$$

where $\tilde{u}_h(\cdot, 0) := u_h^0 \ \in S_h^r$, and u_h^0 is such that

$$\|u^0 - u_h^0\| \leq c h^r. \tag{2.4}$$

The semidiscrete approximation is uniquely defined and has the following properties

$$\|\tilde{u}_h(\cdot, t)\| \leq \|u_h^0\| \, e^{\frac{t}{4\nu}}\,, \quad 0 \leq t \leq t^*, \tag{2.5}$$

$$\|\tilde{u}_h(\cdot, t)\| \ \leq \ \|\tilde{u}_h(\cdot, s)\|, \quad 0 \leq s \leq t \leq t^* \quad \text{for} \quad \nu \geq \frac{1}{4\pi^2}, \tag{2.6}$$

$$\max_{0 \leq t \leq t^*} \|u(\cdot, t) - \tilde{u}_h(\cdot, t)\| \ \leq c h^r, \tag{2.7}$$

cf. [2]. The error estimate (2.7) can be derived by comparing $\tilde{u}_h(\cdot, t)$ to $\tilde{P}_E u(\cdot, t)$, where $\tilde{P}_E : H^2_{\text{per}} \to S_h^r$ is the *time–independent elliptic projection operator* defined by

$$\nu((v - \tilde{P}_E\, v)'', \chi'') - ((v - \tilde{P}_E v)', \chi') + \lambda(v - \tilde{P}_E v, \chi) = 0 \qquad \forall \chi \in S_h^r$$

with $\lambda > \frac{1}{2\nu}$, say. In this paper we will use the *time–dependent* elliptic projection operator $P_E(t) : H^2_{\text{per}} \to S_h^r,\ \ 0 \leq t \leq t^*$, defined by

$$\begin{aligned} &\nu\left((v - P_E(t)v)_{xx}, \chi''\right) - ((v - P_E(t)v)_x, \chi') \\ &\quad + (u(\cdot, t)(v - P_E(t)v)_x, \chi) \ + \ \lambda(v - P_E(t)v, \chi) = 0 \qquad \forall \chi \in S_h^r, \end{aligned} \tag{2.8}$$

with a sufficiently large constant λ, say

$$\lambda > \frac{1}{2} + \frac{1}{2\nu} \left[1 + \frac{1}{2} \max |u(x, t)|^2\right]^2.$$

This elliptic projection will play an important role in the next section in deriving optimal–order error estimates for fully discrete methods.

First, for the elliptic projection we have the following estimate

$$\sum_{j=0}^{2} h^j \, \|v - P_E(t)v\|_j \le ch^s \|v\|_s, \quad v \in H^s_{\text{per}}, \qquad 2 \le s \le r, \tag{2.9}$$

with a constant c independent of h, v and t. This estimate can be proved in the usual manner. For $v \in H^2_{\text{per}}$, obviously $\|v'\|^2 = -(v, v'')$, i.e.,

$$\|v'\|^2 \le \|v\| \, \|v''\|, \qquad v \in H^2_{\text{per}}. \tag{2.10}$$

Now, the bilinear form $a(t; \cdot, \cdot), 0 \le t \le t^*$,

$$a(t; v, w) := \nu(v'', w'') - (v', w') + (u(\cdot, t) \, v', w) + \lambda(v, w)$$

is obviously continuous in H^2_{per}, i.e.,

$$|a(t; v, w)| \le c_1 \|v\|_2 \, \|w\|_2, \qquad v, w \in H^2_{\text{per}}, \tag{2.11}$$

and the constant c_1 can be chosen independent of t. Further, using (2.10), the Cauchy–Schwarz and the arithmetic–geometric mean inequalities, we easily see that $a(t; \cdot, \cdot)$ is coercive in H^2_{per}, i.e.,

$$a(t; v, v) \ge c_2 \, \|v\|_2^2, \qquad v \in H^2_{\text{per}}, \tag{2.12}$$

again with a positive constant c_2 independent of t. Hence, the Lax–Milgram lemma yields in view of the approximation property (2.1)

$$\|v - P_E(t)v\|_2 \le c \, h^{s-2} \|v\|_s, \qquad v \in H^s_{\text{per}}, \qquad 2 \le s \le r, \tag{2.13}$$

with a constant c independent of t. Next, to estimate $\|v - P_E(t)v\|$, consider the auxiliary problem

$$a(t; w, \psi) = (v - P_E(t)v, w) \qquad \forall w \in H^2_{\text{per}},$$

cf. [16]. Then, for $\chi \in S^r_h$ we have

$$\|v - P_E(t)v\|^2 \; = \; a(t; v - P_E(t)v, \psi - \chi) \le c_1 \|\psi - \chi\|_2 \, \|v - P_E(t)v\|_2.$$

Now, the easily established regularity estimate $\|\psi\|_4 \le c\|v - P_E(t)v\|$, with a constant c independent of t, and (2.1) yield, since $r \ge 4$

$$\inf_{\chi \in S^r_h} \|\psi - \chi\|_2 \le c \, h^2 \|v - P_E(t)v\|,$$

and in view of (2.13) we obtain

$$\|v - P_E(t)v\| \le ch^s \|v\|_s, \quad v \in H^s_{\text{per}}, \quad 2 \le s \le r. \tag{2.14}$$

The estimate (2.9) follows now from (2.13), (2.14) and (2.10).

Consequently, setting $W(\cdot, t) := P_E(t) \, u(\cdot, t)$, we have the following estimate

$$\|u(\cdot, t) - W(\cdot, t)\| \; \le \; ch^r \, \|u(\cdot, t)\|_r \, , \qquad 0 \; \le t \le t^*. \tag{2.15}$$

Next, we want to estimate time derivatives of P_E. To this end we prove the following Lemma, cf. [4].

Lemma 2.1. *Let $v \in H^s_{per}$, $2 \le s \le r$. Then, with $P_E^{(j)}(t) := (\frac{d}{dt})^j P_E(t)$, $j \ge 0$,*

$$\|P_E^{(m)}(t)v\|_2 \le C(m)h^s\|v\|_s \quad \text{for} \quad 0 \le t \le t^* \quad \text{and} \quad m > 0. \tag{2.16}$$

Proof. Differentiating (2.8) m times with respect to t, we obtain

$$\begin{aligned} a(t; P_E^{(m)}(t)v, \chi) &= \left(\frac{\partial^m u}{\partial t^m}(\cdot, t)\ [v - P_E(t)v]_x, \chi\right) \\ &- \sum_{j=1}^{m-1} \binom{m}{j} \left(\frac{\partial^j u}{\partial t^j}(\cdot, t)\ [P_E^{(m-j)}(t)v]_x, \chi\right). \end{aligned} \tag{2.17}$$

Taking now $\chi = P_E^{(m)}(t)v$, using (2.12), integrating by parts the first term on the right–hand side of (2.17), applying the Cauchy–Schwarz inequality, and using (2.14), we can easily show inductively that (2.16) holds. □

Remark 2.1. Setting $W^{(m)}(\cdot, t) := (\frac{\partial}{\partial t})^m W(\cdot, t), u^{(m)}(\cdot, t) := (\frac{\partial}{\partial t})^m u(\cdot, t)$, we have

$$W^{(m)}(\cdot, t) = \sum_{j=0}^{m} \binom{m}{j} P_E^{(j)}(t)\ u^{(m-j)}(\cdot, t),$$

and, therefore,

$$\begin{aligned} \|u^{(m)}(\cdot, t) - W^{(m)}(\cdot, t)\| \ &\le\ \|u^{(m)}(\cdot, t) - P_E(t)\, u^{(m)}(\cdot, t)\| \\ &+ \sum_{j=1}^{m} \binom{m}{j} \|P_E^{(j)}(t)\, u^{(m-j)}(\cdot, t)\|. \end{aligned}$$

Hence, in view of (2.9), (2.16), we have

$$\|u^{(m)}(\cdot, t) - W^{(m)}(\cdot, t)\| \ \le\ Ch^r, \quad m \ge 0. \tag{2.18}$$

3. Runge–Kutta discretizations

In this section we discretize in time the semidiscrete problem by suitable implicit RK methods, and, under some mild mesh hypotheses, derive optimal–order error estimates and prove uniqueness of the RK approximations.

For $q \in \mathbb{N}$, a q–stage implicit RK method is specified by a set of constants arranged in tableau form

$$\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1q} & \tau_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qq} & \tau_q \\ \hline b_1 & b_2 & \dots & b_q & \end{array} =: \begin{array}{c|c} A & \tau \\ \hline b^T & \end{array}$$

We shall assume that these methods satisfy certain stability and consistency conditions. We start with the well–known algebraic stability condition

$$\text{(S)} \qquad \begin{cases} b_i \ge 0, \quad i = 1, \dots, q, \\ \text{the matrix } M,\ m_{ij} := a_{ij}b_i + a_{ji}b_j - b_ib_j, \text{is positive semidefinite.} \end{cases}$$

The consistency conditions are given by the simplifying assumptions

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

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

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

for some integers $\sigma, p, \varrho \geq 1$, where p and σ is the stage–order and the classical order, respectively. We will assume that

$$\sigma \leq \varrho + p + 1 \tag{3.1a}$$

$$\sigma \leq 2p + 2. \tag{3.1b}$$

The positivity property

$$\text{(P)}\qquad \begin{cases} A \text{ is invertible and there exists a diagonal matrix } D \text{ with positive} \\ \quad\text{diagonal elements such that } C := DA^{-1}D^{-1} \text{ is positive definite} \end{cases}$$

plays an important role in proving existence of the numerical approximations.

The Gauss–Legendre methods, the Radau IIA methods (with $\tau_q = 1$) and the two– and three–stage optimal–order diagonally implicit (DIRK) methods are examples of implicit RK methods which satisfy all these assumptions (except of the three–stage DIRK that does not satisfy (3.1a)), see [6]. These methods satisfy also the hypothesis

$$\text{(H)}\qquad r(p+1) \geq \sigma$$

which will be occasionally used in the sequel to avoid mesh conditions in the consistency proofs.

Let $F_h : S_h^r \to S_h^r$ be defined by

$$(F_h(v),\chi) = -(vv',\chi) + (v',\chi') - \nu(v'',\chi'') \quad \forall v,\chi \in S_h^r. \tag{3.2}$$

Then, (2.3) may be written in the form

$$\begin{cases} \tilde{u}_{ht} = F_h(\tilde{u}_h) \qquad 0 \leq t \leq t^* \\ \tilde{u}_h(\cdot,0) = u_h^0. \end{cases} \tag{3.3}$$

Let $N \in \mathbb{N}$, $k := t^*/N$, and $t^n := nk$, $n = 0,\dots,N$. The *RK approximations* $U^0,\dots,U^N \in S_h^r$ to $u^0,\dots,u^N, u^n := u(\cdot,t^n)$, are defined by $U^0 := u_h^0$, and

$$U^{n+1} = U^n + k\sum_{j=1}^{q} b_j F_h(U^{n,j}),\quad n=0,\dots,N-1, \tag{3.4}$$

where $U^{n,1}, \dots, U^{n,q} \in S_h^r$ are such that

$$U^{n,j} = U^n + k\sum_{i=1}^{q} a_{ji}F_h(U^{n,i}), \qquad j = 1, \dots, q, \ \ n = 0, \dots, N-1. \tag{3.5}$$

Note that (3.4) can also be written in the form

$$U^{n+1} = U^n + b^T A^{-1}(U^{n,1} - U^n, \dots, U^{n,q} - U^n)^T, \quad n = 0, \dots, N-1. \tag{3.4$'$}$$

Existence. For sufficiently small k (independent of h), the existence of $U^1, \dots, U^N \in S_h^r$ can be shown inductively via a well–known variant of the Brouwer fixed–point theorem. The proof proceeds along similar lines to analogous proofs in [10], [9], using (P), (2.10) and the arithmetic–geometric mean inequality, and will be omitted.

Remark 3.1. Let us note for later use that the same argument that shows existence of the RK approximations, allows us also to conclude that the homogeneous linear system

$$V^{n,j} - k\sum_{i=1}^{q} a_{ji}L_h V^{n,i} = 0, \qquad j = 1, \dots, q, \ \ n = 0, \dots, N-1,$$

where L_h denotes the linear part of F_h,

$$(L_h v, \chi) = (v', \chi') - \nu(v'', \chi'') \quad \forall v, \chi \in S_h^r,$$

has for $k < c\nu$, where c depends only on the specific RK scheme, only the trivial solution in $(S_h^r)^q$.

Error estimates. Given n, $0 \le n \le N-1$, let the 1–periodic functions $\alpha_{j\ell}$, $j = 1, \dots, q$, be recursively defined by

$$\begin{cases} \alpha_{j0} := u(\cdot, t^n) \\ \alpha_{j,\ell+1} := -\sum\limits_{i=1}^{q} a_{ji}\{\alpha''_{i\ell} + \nu\alpha''''_{i\ell} + \sum\limits_{m=0}^{\ell} \alpha_{im}\alpha'_{i,\ell-m}\}, \ \ell = 0, \dots, \sigma-1. \end{cases} \tag{3.6}$$

The following auxiliary results will be used to prove consistency, see Proposition 3.2 below. The proofs of Lemmata 3.1, 3.2 and of Corollary 3.1 are similar to analogous results in [10], [9] and are omitted.

Lemma 3.1. *Let* $\alpha_\ell := (\alpha_{1\ell}, \dots, \alpha_{q\ell})^T, D_t^\ell u := \frac{\partial^\ell u}{\partial t^\ell}(\cdot, t^n), u := u(\cdot, t^n), T := \operatorname{diag}\{\tau_1, \dots, \tau_q\}$, *and* $e := (1, \dots, 1)^T \in \mathbb{R}^q$. *Then, if* (C) *and* (3.1b) *hold, we have*

$$\alpha_\ell = \frac{1}{\ell!} D_t^\ell u \; T^\ell e, \quad \ell = 0, \dots, p \quad \text{if } p \le \sigma, \tag{3.7}$$

$$\alpha_{p+1} = \frac{1}{p!} D_t^{p+1} u \; AT^p e \quad \text{if } p \le \sigma - 1, \tag{3.8}$$

$$
(3.9)\quad \begin{cases} \alpha_{\ell+1} = \dfrac{1}{\ell!}D_t^{\ell+1}u\ AT^\ell e - A\big(\alpha_\ell'' - \dfrac{1}{\ell!}D_t^\ell u_{xx}\ T^\ell e\big) - \nu A\big(\alpha_\ell'''' - \dfrac{1}{\ell!}D_t^\ell u_{xxxx}\ T^\ell e\big) \\ \qquad - A\displaystyle\sum_{m=p+1}^{\ell} \frac{1}{(\ell-m)!}D_t^{\ell-m}u_x\big(T^{\ell-m}\alpha_m - \frac{1}{m!}D_t^m u\ T^\ell e\big) \\ \qquad - A\displaystyle\sum_{m=0}^{p}\frac{1}{m!}D_t^m u\big[T^m\alpha_{\ell-m}' - \frac{1}{(\ell-m)!}D_t^{\ell-m}u_x\ T^\ell e\big], \ell = p+1,\dots,\sigma-1. \end{cases}
$$

Lemma 3.2. *Assume that* (B), (C), (D) *and* (3.1) *hold. Then*

$$
(3.10)\qquad b^T T^s \alpha_\ell = \frac{1}{\ell!\,(\ell+s+1)}D_t^\ell u, \quad s,\ell = 0,\dots,\sigma-1, \quad s+\ell \le \sigma-1. \qquad \square
$$

Corollary 3.1. *Assume that* (B), (C), (D) *and* (3.1) *hold, or that the RK method is the* 3*–stage DIRK. Then*

$$
(3.11)\qquad b^T A^{-1}\alpha_\ell = \frac{1}{\ell!}D_t^\ell u, \quad \ell = 1,\dots,\sigma. \qquad \square
$$

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

$$
(3.12)\qquad u^{n,j} := \sum_{\ell=0}^{\sigma} k^\ell \alpha_{j\ell}, \quad j = 1,\dots,q,
$$

and $\tilde u^{n+1}$ by

$$
(3.13)\qquad \tilde u^{n+1} := u^n + b^T A^{-1}(u^{n,1} - u^n,\dots,u^{n,q} - u^n)^T,
$$

cf. (3.4′). Using (3.12) and (3.11), we have

$$
\tilde u^{n+1} = u^n + b^T A^{-1}\sum_{\ell=1}^{\sigma} k^\ell \alpha_\ell = \sum_{\ell=0}^{\sigma}\frac{k^\ell}{\ell!}D_t^\ell u(\cdot,t^n),
$$

and, consequently,

$$
(3.14)\qquad \|u^{n+1} - \tilde u^{n+1}\|_m \le c_m k^{\sigma+1}, \quad m = 0,1,2,\dots .
$$

The main step towards a consistency proof is undertaken in the following proposition. For the sake of brevity we set $F(v) := -(vv_x + v_{xx} + \nu v_{xxxx})$.

Proposition 3.1. *Let the truncation errors* $e^{n,j}, e^{n+1}$ *be given by*

$$
(3.15)\qquad u^{n,j} = u^n + k\sum_{i=1}^{q} a_{ji}F(u^{n,i}) + e^{n,j}, \quad j = 1,\dots,q,
$$

$$
(3.16)\qquad \tilde u^{n+1} = u^n + k\sum_{j=1}^{q} b_j F(u^{n,j}) + e^{n+1}.
$$

Then, under the hypotheses of Corollary 3.1*, we have*

$$
(3.17)\qquad \|e^{n+1}\|_m + \sum_{j=1}^{q}\|e^{n,j}\|_m \le c_m k^{\sigma+1}, \quad m = 0,1,2,\dots .
$$

Proof. We have

$$\begin{aligned} e^{n,j} &= \sum_{\ell=0}^{\sigma} k^\ell \alpha_{j\ell} - u^n + k\sum_{i=1}^{q} a_{ji}\big\{\sum_{\ell=0}^{\sigma} k^\ell(\alpha''_{i\ell} + \nu\alpha''''_{i\ell}) + (\sum_{\ell=0}^{\sigma} k^\ell \alpha_{i\ell})(\sum_{\ell=0}^{\sigma} k^\ell \alpha'_{i\ell})\big\} \\ &= \sum_{\ell=1}^{\sigma} k^\ell \alpha_{j\ell} + k\sum_{i=1}^{q} a_{ji}\big\{\sum_{\ell=0}^{\sigma-1} k^\ell(\alpha''_{i\ell} + \nu\alpha''''_{i\ell} + \sum_{m=0}^{\ell} \alpha_{im}\alpha'_{i,\ell-m})\big\} + \varepsilon^{n,j} \\ &= \sum_{\ell=0}^{\sigma-1} k^{\ell+1} \alpha_{j,\ell+1} + k\sum_{\ell=0}^{\sigma-1} k^\ell \big\{\sum_{i=1}^{q} a_{ji}(\alpha''_{i\ell} + \nu\alpha''''_{i\ell} + \sum_{m=0}^{\ell} \alpha_{im}\alpha'_{i,\ell-m})\big\} + \varepsilon^{n,j} \end{aligned}$$

with $\|\varepsilon^{n,j}\|_m = O(k^{\sigma+1})$. Using (3.6) we conclude $e^{n,j} = \varepsilon^{n,j}$, i.e.,

$$\sum_{j=1}^{q} \|e^{n,j}\|_m \le c_m k^{\sigma+1}. \tag{3.18}$$

Further, using (3.15), (3.13), we obtain

$$\begin{aligned} e^{n+1} &= \tilde{u}^{n+1} - u^n - k\sum_{j=1}^{q} b_j F(u^{n,j}) \\ &= \tilde{u}^{n+1} - u^n - \sum_{j=1}^{q} b_j \sum_{i=1}^{q} (A^{-1})_{ji}(u^{n,i} - u^n - e^{n,i}) \\ &= \sum_{i,j=1}^{q} b_j (A^{-1})_{ji} e^{n,i}, \end{aligned}$$

and (3.17) follows from (3.18). □

Henceforth we shall let $W^n := W(\cdot, t^n) = P_E(t^n)u(\cdot, t^n), t^{n,j} := t^n + k\tau_j, u_h^{n,j} := P_E(t^{n,j})u^{n,j}$ and $\tilde{u}_h^{n+1} := P_E(t^{n+1})\tilde{u}^{n+1}$, respectively, see (2.8).

Lemma 3.3. *Let $\eta^{n,j},\ j = 1, \dots, q, \eta^{n+1}$ in S_h^r be given by*

$$u_h^{n,j} = W^n + k\sum_{i=1}^{q} a_{ji} F_h(u_h^{n,i}) + \eta^{n,j},\ j = 1, \dots, q, \tag{3.19}$$

$$\tilde{u}_h^{n+1} = W^n + k\sum_{j=1}^{q} b_j F_h(u_h^{n,j}) + \eta^{n+1}. \tag{3.20}$$

Then, under the hypotheses of Corollary 3.1*, unconditionally if* (H) *is satisfied and for $k = O(h^{\frac{1}{p+1}})$ otherwise, we have*

$$\sum_{j=1}^{q} \|\eta^{n,j}\| \le c\, k(k^\sigma + h^r), \tag{3.21}$$

$$\|\eta^{n+1}\| \le c\, k(k^\sigma + h^r). \tag{3.22}$$

Proof. Subtracting (3.19) from (3.15) we get

$$\eta^{n,j} = e^{n,j} + [(u_h^{n,j} - u^{n,j}) - (W^n - u^n)] + k\sum_{i=1}^{q} a_{ji}[F(u^{n,i}) - F_h(u_h^{n,i})]. \tag{3.23}$$

Now, using (3.12), we have

$$(u_h^{n,j} - u^{n,j}) - (W^n - u^n) = (P_E(t^{n,j}) - P_E(t^n))u^n + \sum_{\ell=1}^{\sigma} k^\ell (P_E(t^{n,j})\alpha_{j\ell} - \alpha_{j\ell});$$

thus, by (2.16) and (2.9),

$$\|(u_h^{n,j} - u^{n,j}) - (W^n - u^n)\| \le c\, kh^r. \tag{3.24}$$

Further, let $\omega^{n,j} := P_0 F(u^{n,j}) - F_h(u_h^{n,j})$, P_0 being the L^2-orthogonal projection operator onto S_h^r. Then, with $\vartheta^{n,j} := u^{n,j} - u_h^{n,j}$,

$$\begin{aligned}
\|\omega^{n,j}\|^2 &= (F(u^{n,j}) - F_h(u_h^{n,j}), \omega^{n,j}) \\
&= -(u^{n,j}u_x^{n,j} - u_h^{n,j}u_{hx}^{n,j}, \omega^{n,j}) + (u_x^{n,j} - u_{hx}^{n,j}, \omega_x^{n,j}) - \nu(u_{xx}^{n,j} - u_{hxx}^{n,j}, \omega_{xx}^{n,j}) \\
&= -(u^{n,j}u_x^{n,j} - u_h^{n,j}u_{hx}^{n,j}, \omega^{n,j}) + (u(\cdot, t^{n,j})(u_x^{n,j} - u_{hx}^{n,j}), \omega^{n,j}) + \lambda(u^{n,j} - u_h^{n,j}, \omega^{n,j}) \\
&= -([u^{n,j} - u(\cdot, t^{n,j})]\vartheta_x^{n,j} + \vartheta^{n,j}u_{hx}^{n,i} - \lambda\vartheta^{n,j}, \omega^{n,j}).
\end{aligned}$$

Therefore, using (2.9), (3.12) and (3.7), we easily see that, with $p' := \min(p, \sigma)$,

$$\|\omega^{n,j}\| \le C(k^{p'+1}h^{r-1} + h^r). \tag{3.25}$$

Now by Young's inequality

$$k^{p+1}h^{r-1} \le C(k^{r(p+1)} + h^r). \tag{3.26}$$

Combining (3.23) with (3.17), (3.24)and (3.25), (3.26), we get (3.21). Now, using (3.19), (3.20),

$$\begin{aligned}
\eta^{n+1} &= \tilde{u}_h^{n+1} - W^n - k\sum_{j=1}^{q} b_j F_h(u_h^{n,j}) \\
&= \tilde{u}_h^{n+1} - W^n - \sum_{i,j=1}^{q} b_j (A^{-1})_{ji}(u_h^{n,i} - W^n - \eta^{n,i}),
\end{aligned}$$

and by (2.16)

$$\begin{aligned}
\eta^{n+1} =& P_E(t^n)(\tilde{u}^{n+1} - u^n - b^T A^{-1}(u^{n,1} - u^n, \dots, u^{n,q} - u^n)^T) \\
&+ \sum_{i,j=1}^{q} b_j (A^{-1})_{ji}\, \eta^{n,i} + \varepsilon^{n+1},
\end{aligned}$$

with $\|\varepsilon^{n+1}\| \le c\, kh^r$; thus, (3.22) follows from (3.13) and (3.21). $\square$

Lemma 3.4. *Assume that the implicit RK method satisfies* (P)*, k is sufficiently small, and $v^{n,1},\dots,v^{n,q},v^{n+1}\in S_h^r$ satisfy*

$$v^{n,j}=W^n+k\sum_{i=1}^{q}a_{ji}F_h(v^{n,i}),\quad j=1,\dots,q, \tag{3.27}$$

$$v^{n+1}=W^n+k\sum_{j=1}^{q}b_jF_h(v^{n,j}). \tag{3.28}$$

Then, under the hypotheses of Lemma 3.3,

$$\|u_h^{n,j}-v^{n,j}\|\le c\,k(k^\sigma+h^r),\quad j=1,\dots,q, \tag{3.29}$$

and

$$\|\tilde u_h^{n+1}-v^{n+1}\|\le c\,k(k^\sigma+h^r). \tag{3.30}$$

Proof. The existence of $v^{n,1},\dots,v^{n,q}$ for sufficiently small k can be shown in exactly the same way as the existence of $U^{n,1},\dots,U^{n,q}$ satisfying (3.5). Letting $\zeta^{n,j}:=u_h^{n,j}-v^{n,j}$, from (3.19), (3.27), we obtain

$$\sum_{j,i=1}^{q}c_{ji}d_id_j(\zeta^{n,i},\zeta^{n,j})=k\sum_{j=1}^{q}d_j^2(F_h(u_h^{n,j})-F_h(v^{n,j}),\ \zeta^{n,j})+\sum_{j,i=1}^{q}(A^{-1})_{ji}d_j^2(\eta^{n,i},\zeta^{n,j})=$$
$$k\sum_{j=1}^{q}d_j^2\{\|\zeta_x^{n,j}\|^2-\nu\|\zeta_{xx}^{n,j}\|^2\}-k\sum_{j=1}^{q}d_j^2(u_h^{n,j}u_{hx}^{n,j}-v^{n,j}v_x^{n,j},\zeta^{n,j})+\sum_{j,i=1}^{q}(A^{-1})_{ji}d_j^2(\eta^{n,i},\zeta^{n,j}).$$

Now

$$\begin{aligned}&-(u_h^{n,j}u_{hx}^{n,j}-v^{n,j}v_x^{n,j},\zeta^{n,j})=-(u_{hx}^{n,j}\zeta^{n,j},\zeta^{n,j})+(\zeta^{n,j}\zeta_x^{n,j},\zeta^{n,i})\\&-(u_h^{n,j}\zeta_x^{n,j},\zeta^{n,j})=(u_h^{n,j}\zeta^{n,j},\zeta_x^{n,j}).\end{aligned}$$

Hence, using (P), the Cauchy–Schwarz inequality, (2.10) and the arithmetic–geometric mean inequality,

$$c_1\sum_{j=1}^{q}\|\zeta^{n,j}\|^2\le c_2k\max_i\|u_h^{n,i}\|_{L^\infty}^2\sum_{j=1}^{q}\|\zeta^{n,j}\|^2+\sum_{j,i=1}^{q}(A^{-1})_{ji}d_j^2(\eta^{n,i},\zeta^{n,j}), \tag{3.31}$$

and (3.29) follows from (3.21) in view of (2.9), (3.12) and (3.7). Further, from (3.19), (3.20), (3.27), (3.28), we get

$$\tilde u_h^{n+1}-v^{n+1}=\sum_{j,i=1}^{q}b_j(A^{-1})_{ji}(u_h^{n,i}-v^{n,i}-\eta^{n,i})+\eta^{n+1}$$

and (3.30) follows from (3.29), (3.21) and (3.22). □

We are now ready to prove consistency.

Proposition 3.2. *Assume that u, the solution of* (1.1)–(1.2)*, is sufficiently smooth. Then, under the hypotheses of Lemma* 3.4*,*

$$\|v^{n+1} - W^{n+1}\| \le c\,k(k^\sigma + h^r). \tag{3.32}$$

Proof. Since $W^{n+1} - v^{n+1} = (W^{n+1} - \tilde{u}_h^{n+1}) + (\tilde{u}_h^{n+1} - v^{n+1})$, it suffices, in view of (3.30), to show that

$$\|W^{n+1} - \tilde{u}_h^{n+1}\| \le c\,k(k^\sigma + h^r).$$

Writing $W^{n+1} - \tilde{u}_h^{n+1} = \{(\tilde{u}^{n+1} - u^{n+1}) - P_E(t^{n+1})(\tilde{u}^{n+1} - u^{n+1})\} + (u^{n+1} - \tilde{u}^{n+1})$ the result follows from (3.14) and (2.9). □

In addition to our assumptions on S_h^r, we suppose in the sequel for the family of partitions that

$$\underline{h} \ge c\,h^{2\mu} \tag{3.33}$$

for some positive constants c and μ. It is well known that this hypothesis implies, cf. [15],

$$\|\chi\|_{L^\infty} \le c\,h^{-\mu}\,\|\chi\| \qquad \forall \chi \in S_h^r\,. \tag{3.34}$$

We next prove stability.

Proposition 3.3. *Let U^{n+1}, v^{n+1} satisfy* (3.4) *and* (3.28)*, respectively, and assume that the implicit RK method satisfies* (S) *and* (P)*. Then, for $k^{\sigma+1}h^{-\mu}$ and $kh^{r-\mu}$ bounded,*

$$\|U^{n+1} - v^{n+1}\| \le (1 + ck)\|U^n - W^n\|. \tag{3.35}$$

Proof. Let $\varepsilon^{n,j} := v^{n,j} - U^{n,j}$ and $\delta F_h^j := F_h(v^{n,j}) - F_h(U^{n,j})$. Subtracting (3.4) from (3.28), and taking inner products, we obtain

$$\|v^{n+1} - U^{n+1}\|^2 = \|W^n - U^n\|^2 + 2k\sum_{j=1}^{q} b_j(\delta F_h^j, W^n - U^n) + k^2\sum_{j,i=1}^{q} b_j b_i(\delta F_h^j, \delta F_h^i).$$

Subtracting (3.5) from (3.27) we have

$$W^n - U^n = \varepsilon^{n,j} - k\sum_{i=1}^{q} a_{ji}\delta F_h^i.$$

Therefore

$$\|v^{n+1} - U^{n+1}\|^2 = \|W^n - U^n\|^2 + 2k\sum_{j=1}^{q} b_j(\delta F_h^j, \varepsilon^{n,j}) - k^2\sum_{i,j=1}^{q} m_{ij}(\delta F_h^j, \delta F_h^i),$$

i.e., by (S),

$$\|v^{n+1} - U^{n+1}\|^2 \le \|W^n - U^n\|^2 + 2k\sum_{j=1}^{q} b_j(\delta F_h^j, \varepsilon^{n,j}). \tag{3.36}$$

Now

$$\begin{aligned}(\delta F_h^j, \varepsilon^{n,j}) &= \|\varepsilon_x^{n,j}\|^2 - \nu\|\varepsilon_{xx}^{n,j}\|^2 - (v^{n,j}v_x^{n,j} - U^{n,j}U_x^{n,j}, \varepsilon^{n,j})\\ &= \|\varepsilon_x^{n,j}\|^2 - \nu\|\varepsilon_{xx}^{n,j}\|^2 + (v^{n,j}\varepsilon^{n,j}, \varepsilon_x^{n,j})\\ &\le 2\|\varepsilon_x^{n,j}\|^2 - \nu\|\varepsilon_{xx}^{n,j}\|^2 + \frac{1}{4}\|v^{n,j}\|_{L^\infty}^2\|\varepsilon^{n,j}\|^2,\end{aligned}$$

and, using (2.10),

$$(\delta F_h^j, \varepsilon^{n,j}) \le (\frac{1}{4}\|v^{n,j}\|_{L^\infty}^2 + \frac{1}{\nu})\|\varepsilon^{n,j}\|^2.$$

Therefore, in view of (3.29) and the inverse inequality (3.34),

$$(\delta F_h^j, \varepsilon^{n,j}) \le c\,\|\varepsilon^{n,j}\|^2,$$

and (3.36) yields

$$\|v^{n+1} - U^{n+1}\|^2 \le \|W^n - U^n\|^2 + ck\sum_{j=1}^{q}\|\varepsilon^{n,j}\|^2.$$

Moreover, it can be easily seen that

$$\sum_{j=1}^{q}\|\varepsilon^{n,j}\|^2 \le c\,\|W^n - U^n\|^2, \tag{3.37}$$

cf. the proof of Lemma 3.4, and the result follows. □

Combining consistency and stability we can now easily prove convergence.

Theorem 3.1. *Assume that u, the solution of* (1.1)–(1.2), *is sufficiently smooth, $k^{\sigma+1}h^{-\mu}$ and $kh^{r-\mu}$ are bounded, and that* (3.33) *holds. In case* (H) *is not satisfied assume further $k = O(h^{\frac{1}{p+1}})$. Then, under our hypotheses on the implicit RK method, we have*

$$\max_{0\le n\le N}\|u^n - U^n\| \le c\,(k^\sigma + h^r). \tag{3.38}$$

Proof. From (3.32), (3.35) we obtain

$$\begin{aligned}\|W^{n+1} - U^{n+1}\| &\le \|W^{n+1} - v^{n+1}\| + \|v^{n+1} - U^{n+1}\|\\ &\le c\,k(k^\sigma + h^r) + (1+ck)\|W^n - U^n\|,\end{aligned}$$

and we easily conclude

$$\max_{0\le n\le N}\|W^n - U^n\| \le c\,(k^\sigma + h^r). \tag{3.39}$$

Now, the result follows from (3.39) and (2.15). □

Uniqueness. For fixed $n, 0 \le n \le N-1$, let $V^{n,1}, \dots, V^{n,q} \in S_h^r$ be such that

$$V^{n,j} = U^n + k\sum_{i=1}^{q} a_{ji}F_h(V^{n,i}), \quad j = 1, \dots, q. \tag{3.40}$$

Let $E^{n,j} := U^{n,j} - V^{n,j}$, $j = 1, \dots, q$. Subtracting (3.40) from (3.5), we obtain

$$E^{n,j} = k \sum_{i=1}^{q} a_{ji}[F_h(U^{n,i}) - F_h(V^{n,i})], \quad j = 1, \dots, q.$$

Proceeding as in the proof of Lemma 3.4, we see that for sufficiently small k

$$\sum_{j=1}^{q} \|E^{n,j}\|^2 \le ck \max_i \|U^{n,i}\|_{L^\infty}^2 \sum_{j=1}^{q} \|E^{n,j}\|^2. \tag{3.41}$$

Now, from (3.29), (3.37) and (3.39) we conclude

$$\|u_h^{n,i} - U^{n,i}\| \le c\,(k^\sigma + h^r), \quad i = 1, \dots, q,$$

i.e., using (3.34),

$$\|u_h^{n,i} - U^{n,i}\|_{L^\infty} \le c\, h^{-\mu}(k^\sigma + h^r), \quad i = 1, \dots, q.$$

Therefore, using (3.12), (3.6) and (2.9), we have

$$\max_{i,n} \|U^{n,i}\|_{L^\infty} \le C(1 + k^\sigma h^{-\mu} + h^{r-\mu}),$$

and (3.41) yields, under the hypotheses of Theorem 3.1, for $k^{2\sigma+1}h^{-2\mu}$ and $kh^{2r-2\mu}$ sufficiently small, $E^{n,j} = 0$, i.e., uniqueness of the RK approximations.

4. Solving the nonlinear systems arising in the Runge–Kutta methods

The implementation of the implicit RK method (3.4)–(3.5) requires solving the nonlinear system (3.5). The quite general theory of [3] for linearizing RK equations by Newton's method or by a modified Newton method applies to the KS equation as well and yields under some mild mesh conditions optimal–order error estimates for the resulting approximations provided that at least a specific number of iterations is performed at every time step and accurate starting values are used. In [3] a simpler iterative scheme of explicit–implicit type, explicit in the nonlinear and implicit in the linear part of the equation, for implementing RK methods is also analyzed. This result is not directly applicable in our case, since, due to the presence of a spatial derivative in the nonlinear part of the KS equation, the nonlinear part of F_h does not satisfy hypothesis $(H6)$ of [3] with $\gamma = 0$, an assumption for the analysis of the explicit–implicit scheme in [3], cf. [3, Theorem 6.1]. However, the particular form of the linear part of the KS equation allows us to modify the analysis slightly and prove that the explicit–implicit scheme retains the order of convergence of the RK methods shown in section 3.

Let $\varphi_h : S_h^r \to S_h^r$ denote the nonlinear part of F_h,

$$(\varphi_h(v), \chi) = -(vv', \chi) \quad \forall v, \chi \in S_h^r.$$

Note that $\varphi_h(v) = -P_0(vv')$, where P_0 denotes the L^2–orthogonal projection operator onto S_h^r. Separating linear and nonlinear parts in (3.5), we define $V_\ell^{n,j} \in S_h^r$ recursively by

$$V_{\ell+1}^{n,j} - k\sum_{i=1}^{q} a_{ji} L_h V_{\ell+1}^{n,i} = V^n + k\sum_{i=1}^{q} a_{ji}\varphi_h(V_\ell^{n,i}), \qquad j = 1,\dots,q, \tag{4.1}$$

$\ell = 0,\dots,\ell_n - 1$. Given $V_\ell^{n,i}, i = 1,\dots,q$, the $V_{\ell+1}^{n,j}, j = 1,\dots,q$, are, for sufficiently small k, well defined by (4.1), cf. Remark 3.1. The starting values are assumed given, and $\ell_n \geq 1$ is the number of iterations to be performed at step n. We then define V^{n+1} by

$$V^{n+1} := V^n + b^T A^{-1}(V_{\ell_n}^{n,1} - V^n, \dots, V_{\ell_n}^{n,q} - V^n)^T, \tag{4.2}$$

cf. (3.4′). Starting values $V_0^{n,i}$ may be generated by extrapolating previously computed values $V^n, V^{n-1},\dots$, according to

$$V_0^{n,i} := \sum_{j=0}^{p_n} \mu_{ij}^{p_n} V^{n-j}, \qquad i = 1,\dots,q, \quad n = 0,\dots,N-1, \tag{4.3}$$

where $p_n \leq n$ is a nonnegative integer, and where the extrapolation coefficients are generated as follows: For integer $\ell, 0 \leq \ell \leq n$, let $L_i^{\ell,n}, i = 0,\dots,\ell$, be the Lagrange polynomials of degree ℓ that satisfy $L_i^{\ell,n}(t^{n-j}) = \delta_{ij}, i, j = 0,\dots,\ell$. Then set

$$\mu_{ij}^{\ell} := L_j^{\ell,n}(t^n + k\tau_i), \qquad i = 1,\dots,q, \quad j = 0,\dots,\ell. \tag{4.4}$$

It is easily seen that for a smooth function y

$$\sum_{j=0}^{\ell} \mu_{ij}^{\ell} y(t^{\lambda-j}) = y(t^\lambda + k\tau_i) + O(k^{\ell+1}), \qquad i = 1,\dots,q, \quad \lambda \geq \ell. \tag{4.5}$$

Since the accuracy of the extrapolated values is limited by the number of available past data as well as by $p+1$ and σ, we shall take

$$p_n := \min(n, p, \sigma - 1). \tag{4.6}$$

It is easily seen that the Fréchet derivative $D\varphi_h(\omega)$ is given by $D\varphi_h(\omega)v = -P_0((\omega v)'), \omega, v \in S_h^r$. Hence, $(D\varphi_h(\omega)v, w) = (\omega v, w')$, i.e., for $K > 0$,

$$(D\varphi_h(\omega)v, w) \leq \frac{K^2}{4}\|v\|^2 + \|w'\|^2 \qquad \forall \omega, v, w \in S_h^r,\ \|\omega\|_{L^\infty} \leq K. \tag{4.7}$$

Further, obviously,

$$(L_h v, v) = \|v'\|^2 - \nu\|v''\|^2 \qquad \forall v \in S_h^r. \tag{4.8}$$

Theorem 4.1. *Assume that the hypotheses of Theorem 3.1 are satisfied and that we are given initial data* $V^0,\dots,V^{\tilde p}, \tilde p := \min(p, \sigma - 1)$, *in* S_h^r *satisfying*

$$\max_{0\leq j\leq \tilde p} \|u^j - V^j\| \leq c\,(k^\sigma + h^r). \tag{4.9}$$

Then for $r > \mu$ *and* $k^{\tilde p+1}h^{-\mu}$ *sufficiently small,* $\ell_n \geq 2(\sigma - \tilde p) + 1$, *and* $V^{\tilde p+1},\dots,V^N$ *given by* (4.1)–(4.2), *we have*

$$\max_{0\leq n\leq N} \|u^n - V^n\| \leq c\,(k^\sigma + h^r). \tag{4.10}$$

Proof. It follows from (2.15) and (4.9) that

$$\|W^n - V^n\| \le c_n\,(k^\sigma + h^r), \quad n = 0,\dots,\tilde p. \tag{4.11}$$

We shall prove inductively that

$$(I_i) \qquad \|W^n - V^n\| \le c_n\,(k^\sigma + h^r), \qquad n = \tilde p,\dots,N,$$

$$(I_{ii}) \qquad c_n = (1+\tilde c k)c_{n-1} + \tilde c k, \qquad n = \tilde p + 1,\dots,N,$$

where $\tilde c$ depends only on the implicit RK method and the constant c in (3.32), (3.35). It follows easily from (I_{ii}) that

$$c_n \le c^* := (c_{\tilde p} + 1)e^{\tilde c T}, \qquad n = \tilde p,\dots,N.$$

Let $V^{n,1},\dots,V^{n,q}$ and $\tilde V^{n+1}$ in S_h^r be such that for $n = 0,\dots,N-1$

$$V^{n,j} = V^n + k\sum_{i=1}^{q} a_{ji}F_h(V^{n,i}), \qquad j = 1,\dots,q, \tag{4.12}$$

$$\tilde V^{n+1} = V^n + b^T A^{-1}(V^{n,1} - V^n,\dots,V^{n,q} - V^n)^T. \tag{4.13}$$

Now assume that (I_i), (I_{ii}) hold up to some $n, \tilde p \le n \le N-1$. To extend these to $n+1$, we shall prove inductively that

$$(II_i) \qquad \max_{1\le i\le q} \|V_\ell^{n,i}\|_{L^\infty} \le K, \qquad \ell = 0,\dots,\ell_n,$$

$$(II_{ii}) \qquad \max_{1\le i\le q} \|V^{n,i} - V_\ell^{n,i}\| \le (cK\sqrt{k})^\ell \max_{1\le i\le q} \|V^{n,i} - V_0^{n,i}\|, \qquad \ell = 0,\dots,\ell_n,$$

where $K := 2\max\{|u(x,t)| : 0 \le x \le 1, 0 \le t \le t^*\}$. Next, we verify (II_i) for $\ell = 0$. Obviously

$$\max_{1\le i\le q} \|V^{n,i} - v^{n,i}\| \le c\,\|W^n - V^n\|,$$

see (3.37), and consequently by the induction hypothesis,

$$\max_{1\le i\le q} \|V^{n,i} - v^{n,i}\| \le cc^*\,(k^\sigma + h^r). \tag{4.14}$$

Further

$$V^{n,i} - W(\cdot,t^{n,i}) = (V^{n,i} - v^{n,i}) + (v^{n,i} - u_h^{n,i}) + (u_h^{n,i} - u^{n,i}) + (u^{n,i} - W(\cdot,t^{n,i})),$$

and, hence, in view of (4.14), (3.29), (2.9) and (3.7), (3.12), (3.6)

$$\max_{1\le i\le q} \|V^{n,i} - W(\cdot,t^{n,i})\| \le C\,(k^{\tilde p+1} + h^r). \tag{4.15}$$

Therefore, by (2.9) and (3.34), for k and h sufficiently small,

$$\max_{1\le i\le q} \|V^{n,i}\|_{L^\infty} \le \frac{2}{3}K. \tag{4.16}$$

Further

$$V^{n,i} - V_0^{n,i} = [V^{n,i} - W(\cdot,t^{n,i})] + [W(\cdot,t^{n,i}) - \sum_{j=0}^{\tilde p} \mu_{ij}^{\tilde p} W^{n-j}] + \sum_{j=0}^{\tilde p} \mu_{ij}^{\tilde p}(W^{n-j} - V^{n-j}),$$

and, using (4.15), (4.5), (2.18) and the induction hypothesis,

$$\max_{1\le i\le q} \|V^{n,i} - V_0^{n,i}\| \le Cc^* \, (k^{\tilde p+1} + h^r), \tag{4.17}$$

where C does not depend on k, h, n and the induction indices. Thus, by (4.16) and (3.34) we see that, under our hypotheses, for sufficiently small k and h, (II_i) holds for $\ell = 0$.

Now assume that (II_i) and (II_{ii}) hold up to some $\ell, \ell < \ell_n$. We shall next prove, for k sufficiently small, the estimate

$$\max_{1\le i\le q} \|V^{n,i} - V_{\ell+1}^{n,i}\| \le cK\sqrt{k} \max_{1\le i\le q} \|V^{n,i} - V_\ell^{n,i}\|, \tag{4.18}$$

for some constant c depending only on the RK method and ν. Indeed, from (4.1), (4.12), we obtain, for $j = 1, \dots, q$,

$$\begin{aligned} V^{n,j} - V_{\ell+1}^{n,j} - k\sum_{i=1}^{q} a_{ji} L_h(V^{n,i} - V_{\ell+1}^{n,i}) &= k\sum_{i=1}^{q} a_{ji}[\varphi_h(V^{n,i}) - \varphi_h(V_\ell^{n,i})] \\ &= k\sum_{i=1}^{q} a_{ji} \int_0^1 D\varphi_h(sV^{n,i} + (1-s)V_\ell^{n,i})(V^{n,i} - V_\ell^{n,i}) ds. \end{aligned}$$

Multiplying this system by $D^2 A^{-1}$, taking inner products, using (4.7), (4.16) and the induction hypothesis, for k and h sufficiently small, we obtain

$$\begin{aligned} \sum_{j,i=1}^{q} c_{ij}\, d_i\, d_j (V^{n,j} - V_{\ell+1}^{n,j}, V^{n,i} - V_{\ell+1}^{n,i}) - k \sum_{i=1}^{q} d_i^2 \, (L_h\, (V^{n,i} - V_{\ell+1}^{n,i}), V^{n,i} - V_{\ell+1}^{n,i}) \\ = k\sum_{i=1}^{q} d_i^2 \int_0^1 (D\varphi_h\, (sV^{n,i} + (1-s)V_\ell^{n,i})(V^{n,i} - V_\ell^{n,i}), V^{n,i} - V_{\ell+1}^{n,i}) ds \\ \le \frac{K^2}{4} k \sum_{i=1}^{q} d_i^2 \, \|V^{n,i} - V_\ell^{n,i}\|^2 + k\sum_{i=1}^{q} d_i^2 \, \|(V^{n,i} - V_{\ell+1}^{n,i})'\|^2, \end{aligned}$$

C and D being as in (P). Hence, in view of (P) and (4.8),

$$\begin{aligned} c_1 \sum_{i=1}^{q} \|V^{n,i} - V_{\ell+1}^{n,i}\|^2 \le k \sum_{i=1}^{q} d_i^2 \, [2\|(V^{n,i} - V_{\ell+1}^{n,i})'\|^2 - \nu\|(V^{n,i} - V_{\ell+1}^{n,i})''\|^2] \\ + \frac{K^2}{4} k \sum_{i=1}^{q} d_i^2 \, \|V^{n,i} - V_\ell^{n,i}\|^2, \end{aligned}$$

and (4.18) follows easily using (2.10). From (4.18) we conclude that (II_{ii}) holds for $\ell + 1$ as well. We next verify (II_i) for $\ell + 1$. From (II_{ii}) and (4.17),

$$\max_{1\le i\le q} \|V^{n,i} - V_{\ell+1}^{n,i}\| \le C(cK)^{\ell+1}(c^*\sqrt{k})(\sqrt{k})^\ell \, (k^{\tilde p+1} + h^r);$$

therefore, for k sufficiently small,

$$\max_{1\le i\le q} \|V^{n,i} - V_{\ell+1}^{n,i}\| \le (\sqrt{k})^\ell \, (k^{\tilde p+1} + h^r). \tag{4.19}$$

Using (4.16) and (3.34), we see that, for k and h sufficiently small, (II_i) is satisfied for $\ell+1$. This completes the secondary induction argument (II), and we return to the primary argument (I). For $\ell_n \geq 2(\sigma-\tilde{p})+1$ it follows from (4.2), (4.13) and (4.19)

$$\|V^{n+1}-\tilde{V}^{n+1}\| \leq C\,k(k^\sigma+h^r). \tag{4.20}$$

Now, $W^{n+1}-V^{n+1}=(W^{n+1}-v^{n+1})+(v^{n+1}-\tilde{V}^{n+1})+(\tilde{V}^{n+1}-V^{n+1})$, and thus in view of the consistency and the local stability of the RK method, cf. (3.32), (3.35), respectively, and (4.20)

$$\|W^{n+1}-V^{n+1}\| \leq [(1+\tilde{c}k)c_n+\tilde{c}k](k^\sigma+h^r).$$

This establishes both (I_i) and (I_{ii}). The estimate (4.10) follows now immediately using (2.9) and (4.9). □

Remark 4.1. The implementation of the method described in this section requires solving at every time step a number of $q\dim S_h^r \times q\dim S_h^r$ linear systems, see (4.1). These systems have the same matrix, and in some important cases, e.g. for Gauss–Legendre and for Radau IIA methods, can be decomposed into $q \quad \dim S_h^r \times \dim S_h^r$ systems independent of each other, which can be solved simultaneously on a computer with at least q processors, see [3] and the references therein. We refer the reader also to [3] for some techniques for generating initial data $V^0,\dots,V^{\tilde{p}}$ satisfying (4.9).

References

1. G. D. Akrivis, *Finite difference discretization of the Kuramoto–Sivashinsky equation*, Numer. Math. **63** (1992) 1–11.
2. G. D. Akrivis, *Finite element discretization of the Kuramoto–Sivashinsky equation*, Banach Center Publications (to appear).
3. G. Akrivis, V.A. Dougalis, O. Karakashian: *On the efficient implementation of implicit Runge–Kutta methods*, Submitted.
4. J.H. Bramble, P.H. Sammon, *Efficient higher order single step methods for parabolic problems. Part I*, Math. Comp. **35** (1980) 655–677.
5. P. Constantin, C. Foias, B. Nicolaenko, R. Temam, *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations*, New York: Springer–Verlag 1989.
6. K. Dekker, J.G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations*, Amsterdam: North–Holland 1984.
7. J.M. Hyman, B. Nicolaenko, *The Kuramoto–Sivashinsky equation: A bridge between PDE'S and dynamical systems*, Physica **18 D** (1986) 113–126.
8. J.M. Jolly, I.G. Kevrekidis, E.S. Titi, *Approximate inertial manifolds for the Kuramoto–Sivashinsky equation; analysis and computations*, Physica **44 D** (1990) 38–60.
9. O. Karakashian, G.D. Akrivis, V.A. Dougalis, *On optimal–order error estimates for the nonlinear Schrödinger equation*, SIAM J. Numer. Anal. **30** (1993) 377–400.
10. O. Karakashian, W. McKinney, *On optimal high order in time approximations for the Korteweg–de Vries equation*, Math. Comp. **55** (1990) 473–496.
11. I.G. Kevrekidis, B. Nicolaenko, J.C. Scovel, *Back in the saddle again; a computer assisted study of the Kuramoto–Sivashinsky equation*, SIAM J. Appl. Math. **50** (1990) 760–790.
12. Y. Kuramoto, *Diffusion induced chaos in reaction systems*, Progr. Theoret. Phys. Suppl. **64** (1978) 346–367.

13. B. Nicolaenko, B. Scheurer, *Remarks on the Kuramoto–Sivashinsky equation*, Physica **2 D** (1984) 391–395.
14. B. Nicolaenko, B. Scheurer, R. Temam, *Some global dynamical properties of the Kuramoto–Sivashinsky equation: Nonlinear stability and attractors*, Physica **16 D** (1985) 155–183.
15. J. Nitsche, *Umkehrsätze für Spline–Approximationen*, Compositio Mathematica **21** (1969) 400–416.
16. J. Nitsche, *Verfahren von Ritz und Spline–Interpolation bei Sturm–Liouville–Randwertproblemen*, Numer. Math. **13** (1969) 260–265.
17. R. Osserman, *The isoperimetric inequality*, Bulletin of the A.M.S. **84** (1978) 1182–1238.
18. D.T. Papageorgiou, C. Maldarelli, D.S. Rumschitzki, *Nonlinear interfacial stability of core–annular film flows*, Phys. Fluids **A2** (1990) 340–352.
19. D.T. Papageorgiou, Y.S. Smyrlis, *The route to chaos for the Kuramoto–Sivashinsky equation*, Theoret. Comput. Fluid Dynamics **3** (1991) 15–42.
20. L.L. Schumaker, *Spline Functions: Basic Theory*, New York: John Wiley and Sons, Inc. 1980.
21. G. Sivashinsky, *On flame propagation under conditions of stoichiometry*, SIAM J. Appl. Math. **39** (1980) 67–72.
22. E. Tadmor, *The well–posedness of the Kuramoto–Sivashinsky equation*, SIAM J. Math. Anal. **17** (1986) 884–893.
23. R. Temam, *Infinite–Dimensional Dynamical Systems in Mechanics and Physics*, New York: Springer–Verlag 1988.
24. V. Thomée, B. Wendroff, *Convergence estimates for Galerkin methods for variable coefficient initial value problems*, SIAM J. Numer. Anal. **11** (1974) 1059–1068.

Mathematics Department, University of Crete, 71409 Heraklion, Crete, Greece