To the Graduate Council:
I am submitting herewith a dissertation written by Hyun Young Lee entitled "Galerkin/Runge-Kutta Discretization for the Nonlinear Schrödinger Equation." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in mathematics.

Ohanes Karakashian

Ohanes A. Karakashian, Major professor

We have read this dissertation and recommend its acceptance:

Tsoukiffi...

Steven M. Serbin

D. Samuel Jordan

Accepted for the council:

C. W. Minkel

Vice Provost
and Dean of the Graduate School

GALERKIN/RUNGE-KUTTA DISCRETIZATION FOR THE NONLINEAR SCHRÖDINGER EQUATION

A Dissertation

Presented for the

Doctor of Philosophy

degree

The University of Tennessee, Knoxville

Hyun Young Lee

December 1991

To my parents

ACKNOWLEDGEMENT

I express my sincere appreciation to my advisor, Dr. Ohannes A Karakashian. With his consultation, guidance, criticism and encouragement during my work on this dissertation, my research, and other studies, I was able to achieve results otherwise impossible.

I also must acknowledge with thanks Professors G.Baker, L. Bales, T. Janik, B. Fitzpatrick, and S. Serbin. Through their courses and seminars, all have played significant roles in my mathematical education.

In addition to those mentioned, I want to thank Professors S. Jordan, S. Serbin and L. Tsoukalas for serving on committees and making valuable suggestions and corrections.

Finally, I wish to acknowledge Professor Ki Sik Ha of Busan National University in Korea. He is responsible for my becoming a mathematician.

ABSTRACT

A class of fully discrete high order Galerkin Runge-Kutta methods are constructed and analyzed for the nonlinear Schrödinger equation.

Optimal order error estimates are established for the 0-boundary and periodic boundary value problems, and several computational results such as the order of the temporal accuracy, preservation of two invariants, various kinds of errors are presented.

Furthermore, it is noted that these methods allow computations to be performed in parallel so that the final execution time can be reduced to that of a low order method.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

In this dissertation, we consider a numerical method of Galerkin-finite element type to approximate the solution of the following initial-boundary value problem for the Nonlinear Schrödinger Equation (NLS). Let $\Omega \subset R^d$, d=1,2 or 3, be a bounded domain, with sufficiently smooth boundary $\partial\Omega$, and let $0<T<\infty$ be given. We seek a complex-valued function u(x,t) satisfying

$$
(1.1)\qquad \begin{aligned}
&u_t = -iLu + f(u), && (x,t) \in \overline{\Omega}\times[0,T],\\
&u = 0, && (x,t) \in \partial\Omega\times[0,T],\\
&u(x,0) = u^0(x), && x \in \overline{\Omega},
\end{aligned}
$$

where $-Lu = \sum_{j=1}^{d} \partial_{x_j}^2 u$ and u^0 is a given complex-valued function on $\overline{\Omega}$. The initial data u^0 is assumed to be both sufficiently smooth and compatible, and precise hypotheses on the required smoothness of the solution u and f(u) are made as needed.

In the following, we recall that the existence and uniqueness of the solutions of initial value problem for NLS equation with $f(u) = i\lambda|u|^{p-1}u$ depend on p, d and λ. Consider the following problem:

$$
(1.1')\qquad \begin{aligned}
&u_t = -iLu + i\lambda|u|^{p-1}u,\\
&u(x,0) = u^0(x), \qquad x \in R^d.
\end{aligned}
$$

Theorem 1.1 ([24])

(i) Let $\lambda > 0$ and $p \geq 1+\frac{4}{d}$. Let $u^0(x)$ satisfy the condition

$$E = \int \left(\frac{1}{2}|\nabla u^0(x)|^2 - \frac{\lambda}{p+1}|u^0(x)|^{p+1}\right)dx < 0;$$

then no smooth solution of (1.1') can exist for all t.

(ii) If $\lambda > 0$, assume $p < 1+\frac{4}{d}$. If $\lambda < 0$ and $d > 2$, assume $p < 1+ \frac{4}{d-2}$. If $u^0(x) \in H^1(R^d)$, there exists a unique solution of (1.1') for all t which is a bounded, continuous function of t with values in $H^1(R^d)$.

For details about the physical significance and various properties of existence and uniqueness of the NLS equation, see [7],[20],[21] and [22].

During the past ten years, numerous research papers have been published in the numerical analysis and scientific computing literature on various methods of numerical approximations of finite difference(cf. e.g. [11],[22],[23], [29] and their references), finite element (cf. e.g. [13], [14],[22],[28] and their references), spectral (cf. e.g. [25] and references there) or more specific method (cf. e.g. [29] and references there)) for NLS equations.

In this work, we shall approximate the solution of (1.1) using a Galerkin-finite element type method for the spatial discretization, and implicit Runge-Kutta (IRK) methods for the time stepping. It is well known (cf. [6],[9],[10]) that the order reduction of convergence occurs for the IRK method implemented to an initial-boundary value partial differential equation, especially with some nonlinear terms. Recently, in [1], G. Akrivis, V. Dougalis and O. Karakashian constructed and analyzed Galerkin methods of second-order temporal accuracy and provided their efficient implementations. And also in [17], O. Karakashian, G. Akrivis, V. Dougalis proved the existence, uniqueness, and convergence of Galerkin-IRK methods for the cubic Schrödinger equation (i.e. with $f(u)=i\lambda|u|^2u$, $\lambda\in R\backslash\{0\}$) and they showed that no reduction of order occurs if Ω is a polyhedral domain (or Ω is any finite interval if $d=1$), or if the IRK method satisfies some condition (i.e. $p+3 \geq \nu$ where p is the stage order and ν is the optimal order), which is the case with practically important schemes such as Gauss-Legendre methods with up to 3 stages and 2 or 3-stage optimal-order Diagonally Implicit RK schemes.

In this work, we shall obtain the optimal order accuracy in the L^2-norm for both spatial and temporal approximations by introducing an extrapolation scheme to deal with the nonlinear term f(u). The idea of using extrapolation for nonlinear terms is well-known and has been used extensively for nonlinear problems. For example, in [3], G. Baker, V. Dougalis and O. Karakashian developed an extrapolation type multistep method for the semilinear hyperbolic and parabolic equations. And also in [18], S. Keeling introduced an extrapolation type Implicit Runge-Kutta method for the semilinear, linear with time dependent coefficients, and quasilinear parabolic equations.

What follows in the remaining sections of this chapter is first, a general notation, preliminaries and a development of the spatial and temporal discretizations considered here. Then, at the end of the introduction, results are stated precisely. In the remaining part of this dissertation, error estimates are established for the principal scheme and the starting scheme. Finally there is a chapter on computational aspects in which the results of numerical experiments are presented.

NOTATIONS AND PRELIMINARIES

Now for $1\le p\le\infty$, $L^p = L^p(\Omega)$ will denote the Banach space of complex valued measurable functions defined on Ω, equipped with the norm

$$\|v\|_{L^p} = \left(\int_\Omega |v(x)|^p dx\right)^{\frac{1}{p}}, \qquad 1\le p<\infty$$

$$\|v\|_{L^\infty} = \operatorname*{ess\,sup}_{x\in\Omega} |v(x)|, \qquad p=\infty$$

(In the sequel, we use $\|\cdot\|$ instead of $\|\cdot\|_{L^2}$)

Let $\alpha = (\alpha_1,\alpha_2,\dots,\alpha_N)$, $\alpha_i\ge 0$, denote a multiinteger, and let $|\alpha| = \alpha_1+\alpha_2+\cdots+\alpha_N$. For an integer $s \ge 0$ and $1\le p<\infty$, $W^{s,p}$ will denote the Banach space of complex valued measurable functions which, together with their distributional derivatives of order up to s are in L^p, i.e.

$$W^{s,p} = \{v: D^\alpha v \in L^p, \ |\alpha| \le s\}.$$

These spaces are equipped with the norms

$$\|v\|_{W^{s,p}} = \{ \sum_{|\alpha|\le s} \|D^\alpha v\|^p_{L^p(\Omega)} \}^{1/p}, \qquad 1\le p<\infty,$$

$$\|v\|_{W^{s,\infty}} = \max_{|\alpha|\le s} \|D^\alpha v\|_{L^\infty}, \qquad p = \infty.$$

In addition, let H_0^1 be the subspace of H^1 consisting of functions vanishing on $\partial\Omega$ in the sense of trace. We shall use the symbol H^s for $W^{s,2}$, and $\|\cdot\|_s$ instead of $\|\cdot\|_{s,2}$. Further, $\|\cdot\|_{s,\infty}$ represents the norm on $L^\infty([0,T],H^s)$.

Let L be extended to have domain $H^2 \cap H_0^1$. Then L is L_2 - selfadjoint and for every nonnegative integer s, it is bounded from $H^{s+2} \cap H_0^1$ into H^s. Furthermore, introducing the solution operator T for the elliptic problem

$$Lv = w \text{ in } \Omega,$$

$$v = 0 \text{ on } \partial\Omega,$$

as $Tw \equiv v$, it is well known ([12]) that for every nonnegative integer s, T is bounded from H^s into $H^{s+2} \cap H_0^1$. Also the solution operator is positive definite and self-adjoint on L^2.

SPATIAL DISCRETIZATION

In terms of the solution operator, (1.1) can be written as:

$$\partial_t Tu = -iu + Tf, \qquad (1.2)$$

$$u(0) = u^0.$$

Let $\{S_h\}$ be a family of finite dimensional subspaces of H_0^1. We assume that a corresponding family of operators $\{T_h\}_{0<h<1}$ is given satisfying:

(i) T_h: $L^2 \to S_h$ is selfadjoint, positive semidefinite on L^2, and positive definite on S_h.

(ii) There is an integer $r \geq 2$ such that:

$$\|(T-T_h)v\| \leq Ch^s\|v\|_{s-2} \quad \forall v \in H^{s-2}, 2 \leq s \leq r. \qquad (1.3)$$

As an example of S_h and T_h satisfying (i) and (ii), one can choose S_h as a subspace consisting of continuous piecewise polynomials of degree $\leq$ r-1 vanishing on the boundary and define T_h: $L^2 \to S_h$ such that

$(\nabla T_h v, \nabla\chi) = (v, \chi)$, $\forall\chi \in S_h$. (for details, see [8])

Now problem (1.2) has the following semidiscrete formulation. Find u_h: $[0,T] \to S_h$ such that

$$(1.4)\qquad \partial_t T_h u_h = -\mathrm{i}u_h + T_h f_h,$$

$$u_h(0) = u_h^0.$$

Throughout this work, we shall assume that $S_h \subset L^\infty$ and

$$(1.5)\qquad \|\chi\|_{L^\infty} \le Ch^{-\frac{N}{2}}\|\chi\|, \quad \forall\chi \in S_h.$$

According to the properties described above, the restriction of T_h to S_h is invertible and its inverse is henceforth denoted by L_h. Then L_h satisfies the following properties:

$(L_h v, v)$ is real for $v \in S_h$ and $(L_h v, v) \ge C\|v\|_1^2 \quad \forall v \in S_h$ for some constant C.

Denoting the elliptic projection operator as $P_E \equiv T_h L$, it follows from (1.3) that

$$(1.6)\qquad \|(I-P_E)v\| \le Ch^s\|v\|_s, \quad \forall v \in H^s \cap H_0^1, 2\le s\le r.$$

Also, it can be shown that $L_h T_h \equiv P_0$ is the orthogonal projection of L^2 onto S_h. Then since $I-P_0$ is majorized by $I-P_E$ in L^2, it follows from (1.6) that

$$(1.7)\qquad \|(I-P_0)v\| \le Ch^s\|v\|_s, \quad \forall v \in H^s \cap H_0^1, 2\le s\le r.$$

In addition to (1.6), with $w(t) \equiv P_E u(t)$, we assume (see [27]) that

$$(1.8) \qquad \sup_{0\le t\le T} \|u(t)-w(t)\|_{L^\infty} \equiv r_0(h) \to 0 \quad \text{as } h \to 0.$$

Now problem (1.4) takes the following form.

Find u_h: $[0,T] \to S_h$ satisfying:

$$(1.9) \qquad \begin{aligned} \partial_t u_h &= -iL_h u_h + P_0 f_h \\ u_h(0) &= u_h^0 \end{aligned}$$

In the present work, semidiscrete approximations are not analyzed. Instead, (1.9) serves only as a starting point for fully discrete approximations.

TEMPORAL DISCRETIZATION

For the temporal approximation of the solution to (1.9) implicit Runge-Kutta methods are now introduced. For $q \geq 1$ integer, a q-stage IRK method is characterized by a set of constants arranged in the following tableau form:

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

It is convenient to make the following definitions:

$$A \equiv (a_{i,j})_{1\le i,j\le q}, \quad T \equiv \underset{q\times q}{\mathrm{diag}} \{\tau_i\}, \quad b^T \equiv \langle b_1, b_2, \cdots, b_q\rangle$$

$$e^T \equiv \langle 1, 1, \cdots, 1\rangle \in R^q.$$

Given the initial value problem

$$(1.10)\qquad y' = f(t,y), \qquad 0 \le t \le T,$$
$$y(0) = y^0,$$

IRK methods can be applied to generate approximations to $y(t^n)$ where $k = \frac{T}{N}$ is the temporal stepsize and $t^n = nk$, as follows.

Let

$$(1.11)\qquad y^{n+1} = y^n + k\sum_{j=1}^{q} b_j f(t^{n,j}, y^{n,j})$$

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

$$(1.12)\qquad y^{n,j} = y^n + k\sum_{m=1}^{q} a_{j,m} f(t^{n,m}, y^{n,m}), \quad j = 1,2,\cdots,q.$$

To show the full order accuracy, we assume that

$$(1.13)\qquad l!\,b^T A^{l-1} e = 1, \qquad 1 \le l \le \nu.$$

To illuminate condition (1.13), let (1.10) have $y^0 = 1$ and $f(t,y) = -iy$ so that $y(t) = e^{-it}$. Then from (1.11) and (1.12), $y^n = r(ik)^n$ where $r(z)$ is a rational approximation to the exponential function e^{-z} given by:

$$(1.14)\qquad r(z) \equiv 1 - zb^T(I + zA)^{-1}e.$$

Expanding this expression shows that r(z) is a ν-th order approximation to the exponential e^{-z} if and only if (1.13) holds. Next with regard to stability, assume that

$$(1.15) \qquad |r(ix)| \le 1 \qquad x \ge 0,$$

(I-stable RK methods introduced by G. Baker and J. Bramble in [2] satisfy (1.15))

or

$$(1.15') \qquad \sup_{z \ge z_0} |r(iz)| < 1 \qquad \forall z_0 > 0$$

All A-stable RK methods satisfy (1.15) and 2,3-stage optimal order diagonally implicit RK schemes and 2,3-stage Radau IA and IIA methods satisfy (1.15').(see [10])

Note that the spectrum of A, $\sigma(A)$, consists of the negative reciprocals of the poles of r(z). It is assumed throughout this work that

$$(1.16) \qquad \sigma(A) \subset \{ z \in C \mid Re z \ne 0 \}.$$

But we point out that the condition needed for $\sigma(A)$ for the parabolic problem is $\sigma(A) \subset \{z \in C \mid z \ne 0 \text{ and } Re z \ge 0\}$ (see,[3],[18]). And also (1.16) is weaker than the following condition (p) needed in [17] to show the existence of the numerical approximation:

(p) The matrix $A = (a_{ij})$ is invertible and there exists a positive diagonal matrix D such that $x^T C x > 0 \ \forall x \in R^q$,

$x \neq 0$ where $C = DA^{-1}D^{-1}$.

Throughout this work we assume that there exist constants ρ, $C_\rho>0$ such that, the following local Lipschitz property holds:

$$(1.17)\qquad |u-v|\le\rho \;\Rightarrow\; |f(u)-f(v)|\le C_\rho|u-v|.$$

Returning to the temporal discretization of (1.9), let a q-stage IRK method of order $\nu \geq 1$ be given. Then, the constants $[\alpha_{j,m}]_{1\le j\le q}^{0\le m\le\nu-1}$ are well-defined by

$$(1.18)\qquad \sum_{m=0}^{\nu-1}\alpha_{j,m}m^{l} = (-1)^{l}l!\,\hat{e}_j^{T}A^{l}e \qquad \begin{array}{l} 0\le l\le\nu-1 \\ 1\le j\le q, \quad (m^{l}|_{m=l=0}=1) \end{array}$$

since the $\nu\times\nu$ Vandermonde matrix $(m^{l})_{0\le m,l\le\nu-1}$ is nonsingular. Next with $Nk = T$, and $t^n = nk$ for $\nu-1\le n\le N-1$, suppose the approximations $[U_h^m]_{m=0}^{n} \subset S_h$ are given where $U_h^m \cong u^m$ and $u^m \equiv u(x,t^m)$. Now with $\mathbf{L}^2 \equiv [L^2]^q$, define the extrapolation operators ϵ^n: $[L^2]^\nu \to \mathbf{L}^2$ to have components

$$(1.19)\qquad \epsilon_j^n f_h \equiv \sum_{m=0}^{\nu-1}\alpha_{j,m}f(U_h^{n-m})$$

Also, with $\mathbf{S}_h \equiv [S_h]^q$, define $\mathcal{L}_h$: $\mathbf{S}_h \to \mathbf{S}_h$ and $\mathcal{P}_0$: $\mathbf{L}^2 \to \mathbf{S}_h$ by

$$\mathcal{L}_h \equiv \underset{q\times q}{\mathrm{diag}}\{L_h\} \text{ and } \mathcal{P}_0 \equiv \underset{q\times q}{\mathrm{diag}}\{P_0\}$$

Finally, let $U_h^{n+1} \cong u^{n+1}$ be given by what will thereafter be called the principal scheme:

$$(1.20)\qquad \begin{aligned} \overline{U}_h^n &= eU_h^n - ikA\mathcal{L}_h\overline{U}_h^n + kA\wp_0\epsilon^n f_h \\ U_h^{n+1} &= U_h^n - ikb^T\mathcal{L}_h\overline{U}_h^n + kb^T\wp_0\epsilon^n f_h. \end{aligned}$$

where the existence of $\overline{U}_h^n \in \mathbf{S}_h$ depends on the invertibility of $[I + ikA\mathcal{L}_h]$. Note that (1.20) is only partially patterned after (1.11) and (1.12)-i.e., the intermediate stages $\overline{U}_h^n$ are not fully implicit, and extrapolation circumvents the solution of a nonlinear system of algebraic equations for each n.

Since the extrapolation for the principal scheme uses previously computed approximations, a starting procedure is required to generate $[U_h^m]_{m=0}^{\nu-1}$. Hence for $1\le s\le \nu-1$, let the constants $[\alpha_{j,m}^{n,s}]_{0\le m\le s-1,1\le j\le q}^{0\le n\le s-1}$ be determined by

$$(1.21)\qquad \sum_{m=0}^{s-1} \alpha_{j,m}^{n,s}(m-n)^l = l!\,\hat{e}_j^T A^l e \qquad 0\le l\le s-1,\; 1\le j\le q. \; \left((m-n)^l\big|_{m-n=l=0}\equiv 1\right)$$

Then with $0\le n\le s-1$ and $m_j \equiv j$, $1\le j<s$, $m_j \equiv n$, $j = s$, suppose the approximations $[U_h^{m,j}]_{0\le m\le m_j}^{1\le j\le s} \subset S_h$ are given, where $U_h^{m,j} \approx u^m$. Next, define the extrapolation operators $\epsilon^{n,s}$: $[L^2]^s \to L^2$ to have components

$$(1.22)\qquad \epsilon_j^{n,s} f_h \equiv \sum_{m=0}^{n} \alpha_{j,m}^{n,s} f(U_h^{m,s}) + \sum_{m=n+1}^{s-1} \alpha_{j,m}^{n,s} f(U_h^{m,s-1}),\quad 0\le n\le s-2,$$

$$\epsilon_j^{n,s} f_h \equiv \sum_{m=0}^{s-1} \alpha_{j,m}^{n,s} f(U_h^{m,s}),\qquad n=s-1\le \nu-2.$$

Finally, let $U_h^{n+1,s} \approx u^{n+1}$ be given by what is henceforth called the starting scheme:

$$(1.23)\qquad \overline{U}_h^{n,s} = eU_h^{n,s} - ikA\mathcal{L}_h\overline{U}_h^{n,s} + kAP_0\epsilon^{n,s}f_h$$

$$U_h^{n+1,s} = U_h^{n,s} - ikb^T\mathcal{L}_h\overline{U}_h^{n,s} + kb^TP_0\epsilon^{n,s}f_h$$

where $\overline{U}_h^{n,s} \in \mathbf{S}_h$ and $\epsilon^{n,s}f_h$, respectively, are well-defined provided $[I + ikA\mathcal{L}_h]$ is invertible .

With regard to the initial data $U_h^{0,j} \equiv U_h^0$ $j\ge 0$, it is sufficient to take as U_h^0 any element of S_h which is optimally close to u^0 in L^2, i.e.,

$$\|u^0 - U_h^0\| \le Ch^r$$

In particular, we shall work with

$$(1.24)\qquad U_h^0 = P_0 u^0 .$$

Under rather general conditions, optimal order convergence will be proved for (1.20) and (1.23). A description of the results now follows. In chapter 2, the error committed by (1.20) for the approximation of the solution to (1.1) is shown to be of optimal order in L^2, i.e.,

$$\max_{0\le n\le N} \|U_h^n - u^n\| \le C(h^r + k^\nu), \tag{1.25}$$

under the condition that $h^{-\frac{N}{2}}(h^r+k^\nu)+r_0(h)$ is sufficiently small. In chapter 3, first ν starting approximations are provided using (1.23), under the condition that $h^{-\frac{N}{2}}(h^r+k^\nu)+r_0(h)$ is sufficiently small. In (1.25) and in the sequel, c, C, etc, denote positive generic constants, not necessarily the same at any two different places, which are independent of the discretization parameters h and the time step k, unless the contrary is explicitly indicated.

CHAPTER II

THE PRINCIPAL SCHEME

Since the extrapolation operator $\epsilon^n f_h$ uses previous values of U_h^n, the existence of the sequence $[U_h^n]_{n=0}^N$ hinges on the invertibility of the operator $I + ikA\mathcal{L}_h$. The following result establishes this together with a needed stability estimate.

Lemma 2.1 Assume that (1.16) holds. Then $[I + ikA\mathcal{L}_h]$ is invertible and

$$\|(k\mathcal{L}_h)^\theta [I+ikA\mathcal{L}_h]^{-1}\chi\| \le C\|\chi\|, \qquad \forall\chi \in \mathbf{S}_h,\ 0\le\theta\le 1. \tag{2.1}$$

Proof Let $S^{-1}\Lambda S$ be the Jordan decomposition of A, with Λ an upper triangular matrix, then $I+ikA\mathcal{L}_h = S^{-1}(I+ik\Lambda\mathcal{L}_h)S$. Let Λ_s be the $i_s \times i_s$ block of Λ corresponding to the eigenvalue λ_s, and consider the diagonal operator $\mathcal{L}_h^s = \mathrm{diag}\,\{L_h, L_h, \cdots, L_h\}$ on $(S_h)^{i_s}$. We can show that $I+ikA\mathcal{L}_h$ is invertible in the following way. Assume that $\phi+ik\lambda_s L_h\phi = 0$. Then $(\phi,\phi)+ik\lambda_s(L_h\phi,\phi) = 0$ implies $\phi = 0$ since $(L_h\phi,\phi)>0$ if $\phi\neq 0$ and $\mathrm{Re}\lambda_s\neq 0$. It follows that $I+ikA\mathcal{L}_h$ is invertible. Moreover,

$$(2.2)\quad ((I+ik\Lambda_s\mathcal{L}_h^s)^{-1})_{m,l} = \begin{cases} 0, & m>l \\ (-ikL_h)^{l-m}(I+ik\lambda_s L_h)^{-(l-m+1)}, & 1\le m\le l\le i_s. \end{cases}$$

Set $g(z) = (kz)^{l-m+\theta}(1+ik\lambda_s z)^{-(l-m+1)}$. Then obviously $g(z)$ is continuous everywhere except at $z=\frac{-1}{ik\lambda_s} \notin R$. And also we can show that there exists a constant C independent of k such that

$$\sup_{z\in\sigma(L_h)}|g(z)| \le C.$$

By a spectral argument (see Theorem 9.6.3 [15]), it now follows from (2.2), that $\|(k\mathcal{L}_h^s)^{\theta}(I+ik\Lambda_s\mathcal{L}_h^s)^{-1}\|\le C$. This gives (2.1).

Now, sufficiently accurate starting approximations are assumed given and for $\nu-1\le n\le N-1$, an error equation relating $U_h^n-w^n$ to $U_h^{n+1}-w^{n+1}$ is presented. For the sequel, we make the following definitions:

$$\xi^n\equiv U_h^n-w^n,\quad \eta^n\equiv u^n-w^n,\quad f^n\equiv f(u^n),$$

$$\overline{u}^n\equiv\sum_{l=0}^{\nu} A^l e\partial_t^l u^n k^l,\quad \epsilon_j^n f\equiv\sum_{m=0}^{\nu-1}\alpha_{j,m}f^{n-m},$$

$$f_h^n\equiv f(U_h^n),\quad \mathcal{L}\equiv\underset{q\times q}{\mathrm{diag}}\{L\},\quad \mathcal{P}_E\equiv\underset{q\times q}{\mathrm{diag}}\{P_E\},$$

$$\gamma_h\equiv I-ikb^T\mathcal{L}_h(I+ikA\mathcal{L}_h)^{-1}e.$$

After some calculations, the following error equation is obtained:

$$
\begin{aligned}
(2.3)\qquad \xi^{n+1} \equiv\ & \gamma_h\xi^n \\
& + ikb^T\mathcal{L}_h\,[I+ikA\mathcal{L}_h]^{-1}\,\rho_0\{\,\overline{u}^n-eu^n+ikA\mathcal{L}\overline{u}^n-kA\epsilon^n f\,\} \\
& + ikb^T\mathcal{L}_h\,[I+ikA\mathcal{L}_h]^{-1}\,(\rho_E-\rho_0)\,(\overline{u}^n-eu^n) \\
& + (P_0-P_E)\,(u^{n+1}-u^n) \\
& - P_0\{\,u^{n+1}-u^n+kb^Ti\mathcal{L}\overline{u}^n-kb^T\epsilon^n f\,\} \\
& + kb^T\,[I+ikA\mathcal{L}_h]^{-1}\,\rho_0\,[\,\epsilon^n f_h-\epsilon^n f\,] \\
\equiv\ & \gamma_h\xi^n + \sum_{l=1}^{4}\Psi_l^n + \Phi^n, \qquad\qquad \nu-1\le n\le N-1.
\end{aligned}
$$

Stability follows with a spectral argument, and additional details connected with the following proposition are provided by Crouzeix [9].

Proposition 2.1 If the rational function (1.14) satisfies (1.15), then there exists a constant $\tilde{C}\le 0$, such that

$$
(2.4)\qquad \|\gamma_h\chi\| \le (1+\tilde{C}k)\,\|\chi\|, \qquad \forall\chi\in S_h.
$$

If (1.14) satisfies (1.15'), then $\tilde{C}<0$.

The following two well known discrete Gronwall type lemmas are useful in the convergence proof.

Lemma 2.2 (Gronwall) Suppose that nonnegative values $[X^n]_{n=0}^{n^*}$ are given for which there exist quantities Q_1, $Q_2 \geq 0$ such that for some integer $n_0 > 0$ and $k > 0$

$$X^{n+1} \leq Q_1 + Q_2 k \sum_{m=0}^{n} X^m, \qquad n_0 \leq n \leq n^*-1.$$

Then,

$$X^{n+1} \leq (1+Q_2 k)^{n-n_0}\{Q_1 + Q_2 k \sum_{m=0}^{n_0} x^m\}, \qquad n_0 \leq n \leq n^*-1.$$

Proof See [18].

Lemma 2.3 Suppose that nonnegative values $[X^n]_{n=0}^{n^*}$ are given for which there exist quantities Q_1, $Q_2 \geq 0$ such that for some integer $n_0 \geq 0$, some constant $\tilde{C} < -Q_2(n_0+1)$ and $k>0$

$$X^{n+1} \leq (1+\tilde{C}k)X^n + Q_1 k + Q_2 k \sum_{m=0}^{n_0} X^{n-m}, \quad n_0 \leq n \leq n^*-1.$$

Then there exists a constant $\hat{C} \in (\tilde{C}, 0)$ such that the following holds:

$$X^{n+1} \leq (1+\hat{C}k)^{n+1-n_0}\{X^{n_0} + |\tilde{C}-\hat{C}| k \sum_{m=0}^{n_0-1} X^m\} + Q_1 |\hat{C}|^{-1}\{1-(1+\hat{C}k)^{n+1-n_0}\} \quad n_0 \leq n \leq n^*-1$$

Proof See [18].

Next, the order of consistency is established in the following.

Proposition 2.2 The terms $(\Psi_l^n)_{l=1}^4$ of (2.3) satisfy

$$(2.5)\qquad \sum_{l=1}^{4}\|\Psi_l^n\| \le Ck(h^r+k^\nu)\{\|\partial_t^{\nu+1}u\|_{0,\infty}+\|\partial_t^\nu u\|_{2,\infty}+\|\partial_t u\|_{r,\infty}\}$$

Also, if $[U_h^{n-m}]_{m=0}^{\nu-1}$ are given and satisfy

$$(2.6)\qquad \max_{0\le m\le \nu-1}\|U_h^{n-m}-u^{n-m}\|_{L^\infty} \le \rho,$$

then Φ^n of (2.3) satisfies

$$(2.7)\qquad \|\Phi^n\| \le Ckh^r\|u\|_{r,\infty} + CC_\rho k\sum_{m=0}^{\nu-1}\|\xi^{n-m}\|$$

Proof The terms of (2.3) are considered in the order in which they appear. Since for $0\le l\le\nu$, $\partial_t^l u\in H^2\cap H_0^1$, by (1.1)

$$\begin{aligned}\bar{u}^n-eu^n+ikA\mathcal{L}\bar{u}^n &= \sum_{l=1}^{\nu}A^l e\partial_t^l u^n k^l + \sum_{l=0}^{\nu}A^{l+1}e[iL\partial_t^l u^n]k^{l+1}\\ &= -A^{\nu+1}e\partial_t^{\nu+1}u^n k^{\nu+1} + kA[\sum_{l=0}^{\nu}A^l e\partial_t^l f^n k^l].\end{aligned}$$

By (1.18)

$$(2.8)\qquad \epsilon^n f = \sum_{l=0}^{\nu-1}A^l e\partial_t^l f^n k^l + \mathrm{E}$$

where E has components

$$\mathrm{E}_j = \frac{1}{(\nu-1)!}\sum_{m=0}^{\nu-1}\alpha_{j,m}\int_{t^n}^{t^{n-m}}(t^{n-m}-t)^{\nu-1}\partial_t^\nu f(u(t))\,dt.$$

Now with (2.1) and (1.1), it follows that

$$(2.9) \qquad \|\Psi_1^n\| \le Ck^{\nu+1}\{\|\partial_t^{\nu+1}u\|_{0,\infty} + \|\partial_t^{\nu}f\|_{0,\infty}\}$$

$$\le Ck^{\nu+1}\{\|\partial_t^{\nu+1}u\|_{0,\infty} + \|\partial_t^{\nu}u\|_{2,\infty}\}.$$

Next, as with (1.18), the constants $[\beta_{j,m}]_{1\le j\le q}^{0\le m\le \nu}$ are well-defined by

$$\sum_{m=0}^{\nu}\beta_{j,m}m^{l} = l!\,\nu^{l}\hat{e}_j^{T}A^{l}e(1-\delta_{l0}), \qquad 0\le l\le \nu,$$

$$1\le j\le q. \qquad (m^{l}|_{m=l=0}=1)$$

where δ_{ij} is the Kronecker delta. Hence, define the extrapolation operator χ^n to have components

$$\chi_j^n \equiv \sum_{m=0}^{\nu}\beta_{j,m}u(t^{n}+\frac{mk}{\nu}), \quad 1\le j\le q,$$

so that for $0\le p\le \nu$, with ill-defined sums understood to be zero,

$$\chi^n u = \sum_{l=1}^{p}A^{l}e\partial_t^{l}u^{n}k^{l} + F^{p},$$

where F^p has components

$$F_j^p = \frac{1}{p!}\sum_{m=0}^{\nu}\beta_{j,m}\int_{t^n}^{t^{n+\frac{mk}{\nu}}}(t^{n}+\frac{mk}{\nu}-t)^{p}\partial_t^{p+1}u(t)\,dt.$$

Now note that Ψ_2^n is given by:

$$\Psi_2^n = ikb^{T}[I+ikA\mathcal{L}_h]^{-1}\rho_0\mathcal{L}[\overline{u}^{n}-eu^{n}-\chi^{n}u]$$

$$-ikb^T\mathcal{L}_h\,[I+ikA\mathcal{L}_h]^{-1}\{\rho_0\,[\overline{u}^n-eu^n-\chi^n u] \;+\; [\rho_0-\rho_E]\,\chi^n u\}.$$

With (2.1), (1.6) and (1.7), it follows that

$$\|\Psi_2^n\| \le Ck\;\|\mathcal{L}[A^\nu e\partial_t^\nu u^n k^\nu - F^{\nu-1}]\| + C\|F^\nu\| + C\|[\rho_0-\rho_E]F^0\| \tag{2.10}$$

$$\le Ck(k^\nu+h^r)\{\|\partial_t^\nu u\|_{2,\infty} + \|\partial_t^{\nu+1}u\|_{0,\infty} + \|\partial_t u\|_{r,\infty}\}$$

Now since

$$\Psi_3^n = \int_{t^n}^{t^{n+1}} (P_0-P_E)\,\partial_t u(t)\,dt,$$

with (1.6) and (1.7), it follows that

$$\|\Psi_3^n\| \le Ckh^r\|\partial_t u\|_{r,\infty}. \tag{2.11}$$

Next, using (1.1), (2.8) and (1.13), the following is obtained:

$$\begin{aligned}\Psi_4^n &= -P_0\{G + \sum_{l=0}^{\nu-1}\partial_t^{l+1}u^n\frac{k^{l+1}}{(l+1)!} + \sum_{l=0}^{\nu} b^T A^l e[iL\partial_t^l u^n]k^{l+1} \\ &\qquad -\sum_{l=0}^{\nu-1} b^T A^l e\partial_t^l f^n k^{l+1} - kb^T E\} \\ &= -P_0\{G + b^T A^\nu e[iL\partial_t^\nu u^n k^{\nu+1}] - kb^T E\}\end{aligned}$$

where

$$G = \frac{1}{\nu!}\int_{t^n}^{t^{n+1}}(t^{n+1}-t)^\nu\partial_t^{\nu+1}u(t)\,dt.$$

Then from (1.1)

(2.12) $$\|\psi_4^n\| \le Ck^{\nu+1}\{\|\partial_t^{\nu+1}u\|_{0,\infty} + \|\partial_t^{\nu}u\|_{2,\infty} + \|\partial_t^{\nu}f\|_{0,\infty}\}$$

$$\le Ck^{\nu+1}\{\|\partial_t^{\nu+1}u\|_{0,\infty} + \|\partial_t^{\nu}u\|_{2,\infty}\}.$$

Now, (2.5) follows after combining (2.9)-(2.12). Finally, because of (2.6), (1.17) and (2.1)

$$\begin{aligned}\|\phi^n\| &\le Ck\sum_{m=0}^{\nu-1}\|f_h^{n-m}-f^{n-m}\| \\ &\le CC_\rho k\sum_{m=0}^{\nu-1}\|\xi^{n-m}-\eta^{n-m}\|.\end{aligned}$$

and (2.7) follows from (1.6).

Now set

$$N(u) \equiv \|u\|_{r,\infty} + \|\partial_t u\|_{r,\infty} + \sum_{l=1}^{\nu}\|\partial_t^l u\|_{2,\infty} + \|\partial_t^{\nu+1}u\|_{0,\infty}.$$

Then (1.25) is finally established in the following.

Theorem 2.1 Assume that (1.15) (or (1.15')) and (1.16) hold. Suppose $[U_h^m]_{m=0}^{\nu-1}$ are given satisfying

(2.13) $$\max_{0\le m\le \nu-1}\|U_h^m-W^m\| \le C(u)(h^r+k^\nu).$$

Then, provided $h^{-\frac{N}{2}}(h^r+k^\nu)+r_0(h)$ is sufficiently small, $[U_h^n]_{n=\nu}^N$ are well-defined by (1.20) and

(2.14) $$\max_{0\le n\le N}\|U_h^n-u^n\| \le C(u)(h^r+k^\nu)$$

(C(u) depends exponentially upon T and C_ρ, unless $\tilde{C}$ of (2.4) is sufficiently negative and C_ρ of (1.17) is small enough).

Proof Set $\theta(h,k) \equiv h^{-\frac{N}{2}}(h^r+k^\nu)+r_0(h)$. It is first established that for $\theta(h,k)$ small enough

(2.15) $$\max_{0\le m\le N} \|U_h^m-u^m\|_{L^\infty} \le \rho .$$

Note that by (1.5), (2.13) and (1.8), for $\theta(h,k)$ small enough,

$$\begin{aligned}\|U_h^m-u^m\|_{L^\infty} &\le Ch^{-\frac{N}{2}}\|\xi^m\| + \|\eta^m\|_{L^\infty}\\ &\le Ch^{-\frac{N}{2}}(h^r+k^\nu) + r_0(h)\\ &\le C\theta(h,k)\\ &\le \rho, \qquad 0\le m\le \nu-1 .\end{aligned}$$

Now suppose that for some h, k, sufficiently small, there exists an n = n(h,k) such that $\nu-1\le n\le N-1$ and

(2.16) $$\max_{0\le l\le n} \|U_h^l-u^l\|_{L^\infty} \le \rho ,$$

while

(2.17) $$\|U_h^{n+1}-u^{n+1}\|_{L^\infty} > \rho .$$

Given (2.16), inequalities (2.4), (2.5) and (2.7) can be combined with the error equation (2.3) to obtain

(2.18) $$\|\xi^{l+1}\| \le (1+\tilde{C}k)\|\xi^l\| + C_1k(h^r+k^\nu)N(u) + C_2C_\rho k\sum_{m=0}^{\nu-1}\|\xi^{l-m}\|$$

$$(2.18)\qquad \|\xi^{l+1}\| \le (1+\tilde{C}k)\|\xi^{l}\| + C_1 k(h^{r}+k^{\nu})N(u) + C_2 C_\rho k\sum_{m=0}^{\nu-1}\|\xi^{l-m}\|$$
$$\nu-1\le l\le n$$

for some constants $C_1, C_2 > 0$. By Lemma 2.3, if $\tilde{C}+C_2C_\rho\nu < 0$, there exists a constant $\hat{C}\in(\tilde{C},0)$ such that

$$(2.19)\qquad \|\xi^{n+1}\| \le (1+\hat{C}k)^{n-\nu+2}\{\|\xi^{\nu-1}\| + |\tilde{C}-\hat{C}|k\sum_{m=0}^{\nu-2}\|\xi^{m}\|\}$$
$$+ C_1(h^{r}+k^{\nu})N(u)|\hat{C}|^{-1}\{1-(1+\hat{C}k)^{n-\nu+2}\}$$

Now by (1.5), (1.8), (2.13) and (2.19), for $\theta(h,k)$ sufficiently small

$$\|U_h^{n+1}-u^{n+1}\|_{L^\infty} \le Ch^{-\frac{N}{2}}\|\xi^{n+1}\| + \|\eta^{n+1}\|_{L^\infty}$$
$$\le C\theta(h,k) \le \rho.$$

This contradicts (2.17) and hence, (2.15) is established if $\tilde{C}+C_2C_\rho\nu < 0$. In fact (2.14) follows from (2.13) and (2.19) after using (1.6).

Now suppose that $C_3\equiv\tilde{C}+C_2C_\rho\nu \ge 0$. After summing (2.18) over $\nu-1\le l\le n$, it follows that

$$\|\xi^{n+1}\| \le \|\xi^{\nu-1}\| + C_1T(h^{r}+k^{\nu})N(u) + C_3k\sum_{m=0}^{n}\|\xi^{m}\|.$$

Hence by Lemma 2.2,

$$(2.20)\qquad \|\xi^{n+1}\| \le (1+C_3k)^{n-\nu+1}\{\|\xi^{\nu-1}\| + C_1T(h^{r}+k^{\nu})N(u) + C_3k\sum_{m=0}^{\nu-1}\|\xi^{m}\|\}$$

Now by (1.5), (2.20), (2.13) and (1.8), for
$\theta(h,k)$ sufficiently small

$$\|U_h^{n+1}-u^{n+1}\|_{L^\infty} \le Ch^{-\frac{N}{2}}\|\xi^{n+1}\| + \|\eta^{n+1}\|_{L^\infty}$$

$$\le C\theta(h,k) \le \rho .$$

This contradicts (2.17) and hence (2.15) is established. Finally, (2.14) follows from (2.13) and (2.20) after using (1.6).

CHAPTER III

THE STARTING SCHEME

In this section, (2.13) is established. For the sequel, define

$$\xi^{n,s} \equiv U_h^{n,s}-w^n, \quad \overline{u}^{n,s} \equiv \sum_{l=0}^{s} A^l e\partial_t^l u^n k^l$$
$$\epsilon_j^{n,s}f \equiv \sum_{m=0}^{s-1} \alpha_{j,m}^{n,s} f^m, \quad f_h^{n,s} \equiv f(U_h^{n,s}) .$$

The following error equation is established just as with (2.3)

$$\begin{aligned}(3.1)\quad \xi^{n+1,s} \equiv\ & \gamma_h\xi^{n,s} \\ & + ikb^T\mathcal{L}_h[I+ikA\mathcal{L}_h]^{-1}\rho_0\{\overline{u}^{n,s}-eu^n+ikA\mathcal{L}\overline{u}^{n,s}-kA\epsilon^{n,s}f\} \\ & + ikb^T\mathcal{L}_h[I+ikA\mathcal{L}_h]^{-1}(\rho_E-\rho_0)(\overline{u}^{n,s}-eu^n) \\ & +(P_0-P_E)(u^{n+1}-u^n) \\ & - P_0(u^{n+1}-u^n+kb^Ti\mathcal{L}\overline{u}^{n,s}-kb^T\epsilon^{n,s}f) \\ & + kb^T[I+ikA\mathcal{L}_h]^{-1}\rho_0\{\epsilon^{n,s}f_h-\epsilon^{n,s}f\} \\ \equiv\ & \gamma_h\xi^{n,s} + \sum_{l=1}^{4}\psi_l^{n,s} + \phi^{n,s} \\ \equiv\ & \gamma_h\xi^{n,s} + \psi^{n,s} + \phi^{n,s}.\end{aligned}$$

Proposition 3.1 For $1\le s\le \nu-1$, $\psi^{n,s}$ of (3.1) satisfies

$$(3.2)\quad \|\psi^{n,s}\| \le Ck(h^r+k^s)\{\|\partial_t u\|_{r,\infty} + \|\partial_t^s u\|_{2,\infty} + \|\partial_t^{s+1} u\|_{0,\infty}\}.$$

Also, if for $1\le s\le \nu-1$ and $0\le n\le s-1$, the $[U_h^{m,j}]_{0\le m\le m_j}^{1\le j\le s}$ (where $m_j\equiv$

j if $1\le j<s$, $m_j \equiv n$ if $j=s$) are given satisfying:

$$(3.3) \qquad \|U_h^{m,j}-u^m\|_{L^\infty} \le \rho, \quad 1\le j\le s, \quad 0\le m\le m_j,$$

then $\phi^{n,s}$ of (3.1) satisfies

$$(3.4) \qquad \|\phi^{n,s}\| \le Ckh^r\|u\|_{r,\infty} + CC_\rho k\sum_{m=0}^{n}\|\xi^{m,s}\| + CC_\rho k\sum_{m=n+1}^{s-1}\|\xi^{m,s-1}\|.$$

Proof The terms of (3.1) are considered in the order in which they appear. Since for $0\le l\le s$, $\partial_t^l u \in H^2\cap H_0^1$, by (1.1),

$$\begin{aligned}\overline{u}^{n,s}-eu^n+ikA\mathcal{L}\overline{u}^{n,s} &= \sum_{l=1}^{s}A^l e\partial_t^l u^n k^l + \sum_{l=0}^{s}A^{l+1}e[iL\partial_t^l u^n]k^{l+1}\\ &= -A^{s+1}e\partial_t^{s+1}u^n k^{s+1} + kA[\sum_{l=0}^{s}A^l e\partial_t^l f^n k^l].\end{aligned}$$

By (1.21),

$$(3.5) \qquad \epsilon^{n,s}f = \sum_{l=0}^{s-1}A^l e\partial_t^l f^n k^l + \mathrm{E},$$

where E has components

$$\mathrm{E}_j = \frac{1}{(s-1)!}\sum_{m=o}^{s-1}\alpha_{j,m}^{n,s}\int_{t^n}^{t^m}(t^m-t)^{s-1}\partial_t^s f(u(t))\,dt.$$

Now with (2.1) and (1.1), it follows that

$$\begin{aligned}(3.6) \qquad \|\psi_1^{n,s}\| &\le Ck^{s+1}\{\|\partial_t^{s+1}u\|_{0,\infty} + \|\partial_t^s f\|_{0,\infty}\}\\ &\le Ck^{s+1}\{\|\partial_t^{s+1}u\|_{0,\infty} + \|\partial_t^s u\|_{2,\infty}\}.\end{aligned}$$

The proof for $\psi_2^{n,s}$ is identical to that for ψ_2^n of (2.3) except that v is replaced by s. Therefore, the following holds:

$$(3.7) \qquad \|\psi_2^{n,s}\| \leq Ck(k^s+h^r)\{\|\partial_t^s u\|_{2,\infty} + \|\partial_t^{s+1}u\|_{0,\infty} + \|\partial_t u\|_{r,\infty}\}.$$

Since $\psi_3^{n,s} = \psi_3^n$, from (2.11)

$$(3.8) \qquad \|\psi_3^{n,s}\| \leq Ckh^r\|\partial_t u\|_{r,\infty}.$$

Next, using (1.1), (3.5) and (1.13), the following is obtained:

$$\begin{aligned}\psi_4^{n,s} &= -P_0\{G + \sum_{l=0}^{s-1}\partial_t^{l+1}u^n\frac{k^{l+1}}{(l+1)!} + \sum_{l=0}^{s} b^T A^l e[iL\partial_t^l u^n]k^{l+1} \\ &\qquad - \sum_{l=0}^{s-1} b^T A^l e\partial_t^l f^n k^{l+1} - kb^T E\} \\ &= -P_0\{G + b^T A^s e[iL\partial_t^s u^n k^{s+1}] - kb^T E\},\end{aligned}$$

where $G = \dfrac{1}{s!}\displaystyle\int_{t^n}^{t^{n+1}}(t^{n+1}-t)^s\partial_t^{s+1}u(t)\,dt.$

Then, from (1.1),

$$\begin{aligned}(3.9) \quad \|\psi_4^{n,s}\| &\leq Ck^{s+1}\{\|\partial_t^{s+1}u\|_{0,\infty} + \|\partial_t^s u\|_{0,\infty} + \|\partial_t^s f\|_{0,\infty}\} \\ &\leq Ck^{s+1}\{\|\partial_t^{s+1}u\|_{0,\infty} + \|\partial_t^s u\|_{2,\infty}\}.\end{aligned}$$

Now (3.2) follows after combining (3.6)-(3.9).

Finally, because of (3.3), (1.17) and (2.1),

$$
\begin{aligned}
\|\phi^{n,s}\| &\le Ck\sum_{m=0}^{S-1}\|f_h^m - f^m\| \\
&\le Ck\{\sum_{m=0}^{n}\|f_h^{m,s}-f^m\| + \sum_{m=n+1}^{s-1}\|f_h^{m,s-1}-f^m\|\} \\
&\le CkC_\rho\{\sum_{m=0}^{n}\|U_h^{m,s}-u^m\| + \sum_{m=n+1}^{s-1}\|U_h^{m,s-1}-u^m\|\} \\
&= CC_\rho k\{\sum_{m=0}^{n}\|\xi^{m,s}-\eta^m\| + \sum_{m=n+1}^{s-1}\|\xi^{m,s-1}-\eta^m\|\} \\
&\le Ckh^r\|u\|_{r,\infty} + CC_\rho k\sum_{m=0}^{n}\|\xi^{m,s}\| + CC_\rho k\sum_{m=n+1}^{s-1}\|\xi^{m,s-1}\|.
\end{aligned}
$$

The following lemma is useful in the proof of the theorem which follows.

Lemma 3.1 Let $i\ge 1$, $\mu\ge 2$, $0\le n\le \min(i-1,\mu-2)$ and take $\mu_i \equiv \min(\mu,i)$. Suppose that nonnegative values $[X_j^l]_{0\le l\le l_j}^{1\le j\le i}$ ($l_j \equiv \min(j,\mu-1)$, $1\le j<i$, $l_j \equiv n$, $j=i$) and $[\epsilon_j]_{j=1}^{i}$ are given for which there exist quantities Q_1, Q_2, Q_3, $\delta\ge 0$ such that for $k>0$

$$
X_j^{l+1} \le (1+\delta)X_j^l + Q_1 k\epsilon_j + Q_2 k\sum_{m=0}^{l} X_j^m + Q_3 k\sum_{m=l+1}^{\mu_j-1} X_{j-1}^m\ , \quad 1\le j\le i,\ 0\le l\le l_j
$$

(ill-defined sums understood to be zero). Then with $Q_4\equiv Q_1(\mu-1)$, $Q_5\equiv Q_2(\mu-1)$ and $Q_6\equiv Q_3(\mu-1)\mu$, the following holds:

$$
X_i^{n+1} \le (1+\delta+Q_5 k)^{(\mu-1)(i-1)(n+1)}\sum_{j=0}^{i-1}\{X_{i-j}^0 + Q_4 k\epsilon_{i-j}\}(Q_6 k)^j.
$$

Proof See [18].

Theorem 3.1 Assume that the conditions of Lemma 2.1 and

Proposition 2.1 hold and $h^{-\frac{N}{2}}(h^r+k^2)+r_0(h)$ is sufficiently small. Then $[U_h^{n,s}]_{0\le n\le s}^{1\le s\le \nu-1}$ are well defined by (1.23) and (1.24) and the following holds:

$$(3.10)\qquad \max_{0\le n\le s} \|U_h^{n,s}-w^n\| \le C(h^r+k^{s+1})N(u), \qquad 1\le s\le \nu-1.$$

Therefore, (2.13) follows with $U_h^n \equiv U_h^{n,\nu-1}$, $\quad 0\le n\le \nu-1$.

Proof Set $\theta \equiv h^{-\frac{N}{2}}(h^r+k^2)+r_0(h)$. It is first established that for θ small enough

$$(3.11)\qquad \|U_h^{m,j}-u^m\|_{L^\infty} \le \rho,\ 1\le j\le \nu-1,\ 0\le m\le j.$$

Note by (1.8), (1.5) and (1.24), for θ small enough,

$$\begin{aligned}\|U_h^{0,j}-u^0\|_{L^\infty} &\le Ch^{-\frac{N}{2}}\|\xi^0\| + \|\eta^0\|_{L^\infty}\\ &\le C\theta \le \rho.\end{aligned}$$

Now suppose that for each h and k, there exist an $s = s(h,k)$ and $n = n(h,k)$ such that $1\le s\le \nu-1$, $0\le n\le s-1$ and with $l_j \equiv j$, $1\le j<s$, $l_j \equiv n$, $j=s$,

$$(3.12)\qquad \|U_h^{l,j}-u^l\|_{L^\infty} \le \rho,\ 1\le j\le s,\ 0\le l\le l_j,$$

while

(3.13) $\quad \|U_h^{n+1,s}-u^{n+1}\|_{L^\infty} > \rho$.

Now by proposition (3.1), there exist constants C_1, C_2, $C_3 \geq 0$ such that for $1 \leq j \leq s$, $0 \leq l \leq l_j$,

$$(3.14) \quad \|\xi^{l+1,j}\| \leq \|\xi^{l,j}\| + C_1 k\epsilon_j N(u) + C_2 k\sum_{m=0}^{l} \|\xi^{m,j}\| + C_3 k \sum_{m=l+1}^{j-1} \|\xi^{m,j-1}\|$$

where $\epsilon_j = h^r + k^j$. By Lemma 3.1, (3.14) leads to the following with $C_4 = C_1(\nu-1)$, $C_5 = C_2(\nu-1)$, $C_6 = C_3(\nu-1)\nu$:

$$\|\xi^{n+1,s}\| \leq (1+C_5 k)^{(\nu-1)(s-1)(n+1)} \sum_{j=0}^{s-1} \{ \|\xi^0\| + C_4 k\epsilon_{s-j} N(u) \} (C_6 k)^j.$$

Since $k^{j+1}\epsilon_{s-j} \leq h^r + k^{s+1}$, with (1.24) it follows that

(3.15) $\quad \|\xi^{n+1,s}\| \leq C(h^r + k^{s+1})N(u)$.

Hence from (1.8), (1.5) and (3.15), for sufficiently small Θ

$$\begin{aligned} \|U_h^{n+1,s}-u^{n+1}\|_{L^\infty} &\leq Ch^{-\frac{N}{2}}\|\xi^{n+1,s}\| + \|\eta^{n+1}\|_{L^\infty} \\ &\leq C\Theta \leq \rho. \end{aligned}$$

This contradicts (3.13), so (3.11) is established. In fact, (3.15) gives (3.10). Therefore (2.13) follows with $U_h^n \equiv U_h^{n,\nu-1}$.

Remark We can extend the results of Chapters 2,3 to initial, periodic-boundary value problems for the NLS equation, $u_t = -i\beta Lu + f(u)$, where β is a nonzero real number.

The U_h^n will now be sought in a finite dimensional subspace S_h of H^1 endowed with the approximation property

$$\inf_{\chi \in S_h} (\|v-\chi\| + h\|v-\chi\|_1) \leq Ch^s\|v\|_s, \quad v \in H^s, \ 1\leq s\leq r. \tag{3.16}$$

Define T_h: $L^2 \to S_h$ such that

$$(\nabla T_h w, \nabla\chi) + (T_h w, \chi) = (w, \chi) \qquad \forall\chi \in S_h;$$

then T_h is positive definite on S_h. And, also we can define the inverse of T_h on S_h as $T_h^{-1} = -\Delta_h + I$, where the discrete Laplacian Δ_h: $S_h \to S_h$ is now given by

$$(\Delta_h w, \chi) = -(\nabla w, \nabla\chi), \quad \forall\chi \in S_h.$$

Define an elliptic projection operator P_E: $H^1 \to S_h$ such that

$$(\nabla P_E v, \nabla\chi) + (P_E v, \chi) = (\nabla v, \nabla\chi) + (v, \chi), \ \forall\chi \in S_h. \tag{3.17}$$

Then it is well known that (1.6) holds, i.e.,

$$\|(I-P_E)v\| \leq Ch^s\|v\|_s, \quad \forall v \in H^s, \ 2\leq s\leq r.$$

And we have obviously the following:

(i) $((-\Delta_h+I)v, v)$ is real, $\forall v \in S_h$,

(ii) $P_E = T_h(L + I)$ and $((-\Delta_h+I)v, v) \geq C\|v\|_1^2$, $\forall v \in S_h$, for some constant C,

(iii) $P_0 = (-\Delta_h+I)T_h$.

For the periodic boundary-value problem, we can easily prove that all the requirements and theorems in Chapter 1,2 and 3 can be established by replacing Lu by Lu + u, f(u) by iβu +f(u), L_h by $(-\Delta_h+I)$ and -i(or i) by -βi(or βi).

CHAPTER IV

COMPUTATIONAL ASPECTS

In this section, we describe the result of numerical experiments designed to measure the performance of the numerical methods introduced in Chapter 1.

Consider the following problem:

$$(4.1a) \qquad u_t = -i\frac{1}{40}u_{xx} - 128000i|u|^2u, \qquad (x,t)\in R\times[0,T],$$

$$u(x,0) = \frac{1}{40}e^{-i(80x+\frac{\pi}{2})}sech(40(x-0.5)), \qquad x\in R,$$

whose solution is

$$(4.1b) \qquad u(x,t) = \frac{1}{40}e^{-i(80(x-4t)+200t+\frac{\pi}{2})}sech(40(x-4t-0.5)).$$

The modulus of the initial condition is a solitary wave of amplitude $\frac{1}{40}$, centered at x = 0.5 and of support essentially in [0.1,0.9]. Although (4.1b) is a solution to the pure initial-value problem (4.1a) on the whole real line, it may be considered as a good approximation to the solution corresponding to the periodic boundary conditions since the tails of the solution decay exponentially. For t > 0, this moves to the right without change of form with speed 4 and has completed $\frac{1}{2}$ revolution at t = 0.125.

It is straightforward to see that the L^2 norm of the solution u(.,t) = u(t) of (4.1) is an invariant of the equation, i.e., that

$$(4.2) \qquad \|u(t)\| = \|u(0)\|, \quad 0 \le t \le T.$$

To see this, multiply (4.1a) by $\bar{u}$ and take real parts. In addition, we obtain the second invariant,

$$(4.3) \qquad \frac{1}{40}\|u_x(\cdot,t)\|^2 - 64000\|u(\cdot,t)\|_{L^4}^4$$

$$= \frac{1}{40}\|u_x(\cdot,0)\|^2 - 64000\|u(\cdot,0)\|_{L^4}^4, \quad 0 \le t \le T,$$

by multiplying (4.1a) by $\bar{u}_t$ and taking imaginary parts. We use the schemes (1.20) and (1.23) in their complex formulation, adapted of course to equation (4.1a) which has $-\frac{1}{40}u_{xx}$ instead of u_{xx}. As S_h we can take, for example, the space of smooth periodic splines of order r on a uniform mesh with mesh length h.

For the time stepping procedure, we use two or three stage Gauss Legendre method with constants a_{ij}, b_i, τ_i, $1 \le i,j \le q$, (where q=2 or q=3) given by

$$\begin{array}{cc|c}
\frac{1}{4} & \frac{1}{4}-\frac{1}{2\sqrt{3}} & \frac{1}{2}-\frac{1}{2\sqrt{3}} \\
\frac{1}{4}+\frac{1}{2\sqrt{3}} & \frac{1}{4} & \frac{1}{2}+\frac{1}{2\sqrt{3}} \\
\hline
\frac{1}{2} & \frac{1}{2} &
\end{array}$$

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

The 2 stage and 3 stage Gauss-Legendre method satisfies (1.13) with $v = 4$ and 6, respectively. Since these two methods satisfy (1.15) but not (1.15'), (see [10]), the constant C in Theorem 2.1 depends exponentially on T as we mentioned before. For both methods, A can be transformed into $A = S^{-1}\Lambda S$ where Λ is a complex-valued diagonal matrix $\Lambda = \operatorname*{diag}_{1\le i\le q} \{\lambda_i\}$.

Then the linear system from (1.20) or (1.23) can be written as

$$[I+ik\Lambda L]S\Phi = S\Psi$$

which decouples to equations of the form

$$(4.4) \qquad [I+ik\lambda_i L]\phi_i = \Psi_i, \quad i = 1,\cdots,q \text{ (where q=2 or 3)}$$

All of the experiments reported in the sequel were performed in double precision using complex arithmetic on the SPARC station 1 (manufactured by the Sun Microsystems Inc.)at the University of Tennessee at Knoxville.

To solve the linear system (4.4), which is the main task of this computational work, we use the following implementation introduced in [4]. Since the periodic boundary condition yields the upper-right corner which is an $r \times r$ upper triangular matrix and the lower-left corner which is an $r \times r$ lower triangular matrix, the system in (4.4) may be written in the form

$$\mathfrak{F}\hat{z} = \hat{g} \tag{4.5}$$

where $\mathfrak{F} = \mathfrak{F}_b + \mathfrak{F}_c$, $\mathfrak{F}_b$ is a band matrix with band width $2r-1$, and $\mathfrak{F}_c$ consists only of the upper-right and lower-left corners of $\mathfrak{F}$. To solve (4.5), the following steps are effective.

(a) Factor $\mathfrak{F}_b$ into upper and lower triangular banded matrices using a standard banded factoring routine. (In our computational code we use ZGBFA subroutine of LINPACK.)

(b) To solve $\mathfrak{F}_b x = Y$, we use ZGBSL subroutine of LINPACK.

(c) Compute the $2r-2$ N-vectors $\tilde{e}^j$, $j = 1,\cdots,r-1,N-r+2,\cdots,N$ that satisfy $\mathfrak{F}_b\tilde{e}^j = \hat{e}^j$, where the i-th component of $\hat{e}^j$ is δ_{ij}, the Kronecker δ-function.

(d) Evaluate the $N \times N$ matrix C whose j-th column is $\hat{e}^j + \mathfrak{F}_c\tilde{e}^j$, if $j = 1,2,\cdots,r-1,N-r+2,\cdots,N$, and whose j-th column is zero, otherwise. Note that only the four $(r-1)\times(r-1)$ corners of C are nonzero. The matrix C is compressed into a (2r-2) $\times$ (2r-2) array and factored as a product of upper and lower triangular matrices.

(e) The factored form of the compressed matrix C may be used to evaluate the 2r-2 nonzero entries of the N-vector $\hat{\lambda} = (\lambda_1, \lambda_2, \cdots, 0, \cdots, 0, \lambda_{N-r+2}, \cdots, \lambda_N)$ that satisfies $C\hat{\lambda} = -\mathfrak{J}_c \hat{z}_b$.

(f) Finally, the solution $\hat{z}$ of (4.5) is computed via the formula

$$\hat{z} = \hat{z}_b + \sum_{j=1}^{r-1} \lambda_j \hat{e}^j + \sum_{j=N-r+2}^{N} \lambda_j \hat{e}^j .$$

Now we report the outcome of a number of numerical experiments that were performed on (4.1a). First, the convergence rates of the scheme in the temporal variable were investigated for r = 3,4,6. The measures of error used were the L^2-norm given by

$$E(t) = \|U^n - u(\cdot, t^n)\|$$

and the normalized L^2-norm given by

$$N(t) = \frac{\|U^n - u(\cdot, t^n)\|}{\|u_0\|} \quad \text{if } t = nk,\ n=1,2,\cdots.$$

The experimental determination of the temporal accuracy is a somewhat delicate matter because long runs with very small values of h are extremely expensive, both in terms of run time and storage. To avoid this, we used the following very efficient device introduced in [5]. For a fixed h, we make a reference calculation with a sufficiently small $k = k_{ref}$.

For the same value h, we then defined a modified error associated to values of k that are larger than k_{ref}, namely,

$$E^*(t) = \|U^n(h,k) - U^m(h,k_{ref})\|,$$

$$N^*(t) = \frac{\|U^n(h,k) - U^m(h,k_{ref})\|}{\|u_0\|},$$

where $t = nk = mk_{ref}$.

Each of $E^*(t)$ and $N^*(t)$ can be considered as a pure temporal error because subtracting $U^m(h,k_{ref})$ from $U^n(h,k)$ essentially cancels the spatial error inherent in the latter approximation.

The results of these comparisons are shown in Tables 1-4, which refer to splines of order 3, 4 and 6. The expected temporal rate of convergence, ν = 4,6 respectively, emerges clearly from these experiments for the 2 stage or the 3 stage Gauss Legendre method, respectively. Therefore, for the values k where the temporal rate of convergence is close to the theoretical value ν, the temporal error can be expressed as

$$E^*(t) = C_2 k^\nu \text{ and } N^*(t) = C_3 k^\nu.$$

For the 2 stage Gauss Legendre method, from Tables 1 and 2, $C_2 \cong 1.4 \times 10^8$ ($C_3 \cong 2.7 \times 10^{10}$) emerges as a value that fits the data independently of h and the order of splines used.

For the 3 stage Gauss Legendre method, from Tables 3 and 4, $C_2 \cong 3.1 \times 10^{13}$ ($C_3 \cong 5.7 \times 10^{15}$) fits the data again

Table 1: Temporal rate of convergence at T=0.125 with cubic splines, 2 stage Gauss-Legendre method.

$h^{-1} = 200$

N	$\frac{1}{N}h^{-1}$	E(T)	$E^*(T)$	Rate
100	2	0.7826(-3)	0.7826(-3)	4.67
200	1	0.3074(-4)	0.3074(-4)	4.45
300	2/3	0.5056(-5)	0.5056(-5)	4.26
400	1/2	0.1486(-5)	0.1485(-5)	4.14
500	2/5	0.5906(-6)	0.5902(-6)	4.07
600	1/3	0.2816(-6)	0.2812(-6)	4.01
700	2/7	0.1516(-6)	0.1516(-6)	
REF				
5000	1/25	0.7796(-8)		

$h^{-1} = 300$

N	$\frac{1}{N}h^{-1}$	N(T)	$N^*(T)$	Rate
150	2	0.2067(-1)	0.2067(-1)	4.51
300	1	0.9085(-3)	0.9044(-3)	4.23
450	2/3	0.1688(-3)	0.1628(-3)	4.08
600	1/2	0.5777(-4)	0.5030(-4)	4.02
750	2/5	0.2969(-4)	0.2050(-4)	4.00
900	1/3	0.2070(-4)	0.9896(-5)	
REF				
7500	1/25	0.1420(-4)		

Table 2: Temporal rate of convergence at T=0.125 with quadratic splines, 2 stage Gauss-Legendre method

$h^{-1} = 200$

N	$\frac{1}{N}h^{-1}$	N(T)	$N^*(T)$	Rate
100	2	0.1383	0.1387	4.66
200	1	0.9985(-2)	0.5494(-2)	4.46
300	2/3	0.7041(-2)	0.8990(-3)	4.26
400	1/2	0.6646(-2)	0.2642(-3)	4.14
500	2/5	0.6050(-2)	0.1050(-3)	4.07
600	1/3	0.6485(-2)	0.5004(-4)	4.03
700	2/7	0.6465(-2)	0.2689(-4)	
REF				
5000	1/25	0.6441(-2)		

$h^{-1} = 300$

N	$\frac{1}{N}h^{-1}$	E(T)	$E^*(T)$	Rate
150	2	0.1168(-3)	0.1154(-3)	4.51
300	1	0.1068(-4)	0.5051(-5)	4.23
450	2/3	0.7520(-5)	0.9095(-6)	4.08
600	1/2	0.7016(-5)	0.2810(-6)	4.02
750	2/5	0.6872(-5)	0.1145(-6)	
REF				
7500	1/25	0.6766(-5)		

Table 3: Temporal rates of convergence at T=0.125 with quintic spline, 3 stage Gauss-Legendre method.

$h^{-1} = 100$

N	$\frac{1}{N}h^{-1}$	E(T)	$E^*(T)$	Rate
200	1/2	0.6461(-5)	0.5122(-5)	7.77
300	1/3	0.3593(-5)	0.2198(-6)	6.46
400	1/4	0.3606(-5)	0.3431(-7)	6.34
500	1/5	0.3613(-5)	0.8343(-8)	6.24
600	1/6	0.3615(-5)	0.2676(-8)	6.16
700	1/7	0.3616(-5)	0.1035(-8)	6.11
800	1/8	0.3616(-5)	0.4579(-9)	6.08
900	1/9	0.3616(-5)	0.2238(-9)	
REF				
6000	1/60	0.3616(-5)		

$h^{-1} = 150$

N	$\frac{1}{N}h^{-1}$	E(T)	$E^*(T)$	Rate
300	1/2	0.3946(-6)	0.3949(-6)	7.54
500	3/10	0.6603(-7)	0.8376(-8)	6.25
600	1/4	0.6607(-7)	0.2679(-8)	6.17
700	3/14	0.6620(-7)	0.1035(-8)	6.12
800	3/16	0.6627(-7)	0.4572(-9)	6.08
900	1/6	0.6630(-7)	0.2234(-9)	6.05
1000	3/20	0.6631(-7)	0.1181(-9)	
REF				
7000	3/140	0.6632(-7)		

Table 4: Temporal rates of convergence at T=0.125 with cubic spline, 3 stage Gauss-Legendre method

$h^{-1} = 100$

N	$\frac{1}{N}h^{-1}$	N(T)	$N^*(T)$	Rate
200	1/2	0.1288(-1)	0.4523(-2)	11.8
300	1/3	0.1206(-1)	0.3810(-4)	6.41
400	1/4	0.1207(-1)	0.6019(-5)	6.29
500	1/5	0.1207(-1)	0.1478(-5)	6.20
600	1/6	0.1208(-1)	0.4773(-6)	6.13
700	1/7	0.1208(-1)	0.1855(-6)	6.09
800	1/8	0.1208(-1)	0.8226(-7)	6.06
900	1/9	0.1208(-1)	0.4027(-7)	
REF				
6000	1/60	0.1208(-1)		

$h^{-1} = 150$

N	$\frac{1}{N}h^{-1}$	N(T)	$N^*(T)$	Rate
300	1/2	0.2393(-2)	0.2295(-2)	14.34
500	3/10	0.6904(-3)	0.1515(-5)	6.28
600	1/4	0.6904(-3)	0.4821(-6)	6.16
700	3/14	0.6911(-3)	0.1866(-6)	6.10
800	3/16	0.6912(-3)	0.8258(-7)	6.07
900	1/6	0.6912(-3)	0.4040(-7)	6.05
1000	3/20	0.6913(-3)	0.2135(-7)	
REF				
7000	3/200	0.6913(-3)		

independently of h and the order of splines used.

The constants C_2, C_3 are rather big because the temporal rate depends on $\|\partial_t^{\nu+1}u\|_{0,\infty} \cong (160)^{\nu+1}$ and $\|\partial_t^{\nu}u\|_{2,\infty} \cong (160)^{\nu}(80)^2$.

From Tables 1-4, we can conclude that our computational implementation follows the results of Chapters 2,3 very well.

The next set of experiments reported here feature computing the approximate solution of (4.1a) up to T = 1.0 in order to study the behavior of various kinds of errors over a longer temporal interval. In addition to the L_2-error E(t), we also provide L_2-error of modulus, L_2-based shape error and phase error.

The L_2-error of modulus M(t) at t is defined by

$M(t) = [\int_0^1 (|u(x,t)|-|U^n(x)|)^2 dx]^{\frac{1}{2}}$ where nk = t.

The shape error S^n is defined for each time step n = $0,1,2,\cdots,J$ as follows:

Fix n and consider the quantity

$$\mathscr{E}(\tau) = [\int_0^1 (|u(x,\tau)|-|U^n(x)|)^2 dx]^{\frac{1}{2}}$$

where $u(x,\tau)$ is given in (4.1b) and U^n is the computed solution at the n-th time step. Let $\tau^* = \tau^*(n)$ denote the value of τ near nk where $\mathscr{E}(\tau)$ takes its minimum value. If $|U^n|$ resembles a solitary wave in shape, it follows that τ^* is well-defined. Then S^n measures how far the computed solution differs from the original solution as regards to its shape.

The phase error P^n at any time step n with $0 \le n \le J$ is defined to be $|nk - \tau^*(n)|$. This quantity measures the error in the position at which the wave is located.

In Table 5, errors are shown at times t = 0.2,0.4,0.6,0.8 and 1.0 for the 2 stage Gauss-Legendre method with parameters r=4, h=1/100 and N=1500 (k=1/1500) on the temporal interval [0,1].

Table 5: The growth of errors over time interval [0,1] with k=1/1500, h=1/100

t	E(t)	Modulus Error	Shape Error	Phase Error
0.2	0.1391(-3)	0.1217(-3)	0.1515(-4)	0.2340(-3)
0.4	0.3925(-3)	0.3911(-3)	0.1597(-4)	0.7573(-3)
0.6	0.8365(-3)	0.8126(-3)	0.2000(-4)	0.1576(-2)
0.8	0.1477(-2)	0.1381(-2)	0.2632(-4)	0.2694(-2)
1.0	0.2299(-2)	0.2084(-2)	0.4369(-4)	0.4122(-2)

In Figure 1, we show E(t) and the phase error as functions of time for the data in Table 5.

As we expected, E(t) and the phase error exhibit exponential growth approximately. And M(t) exhibits similar to that of E(t), but the shape error stays small.

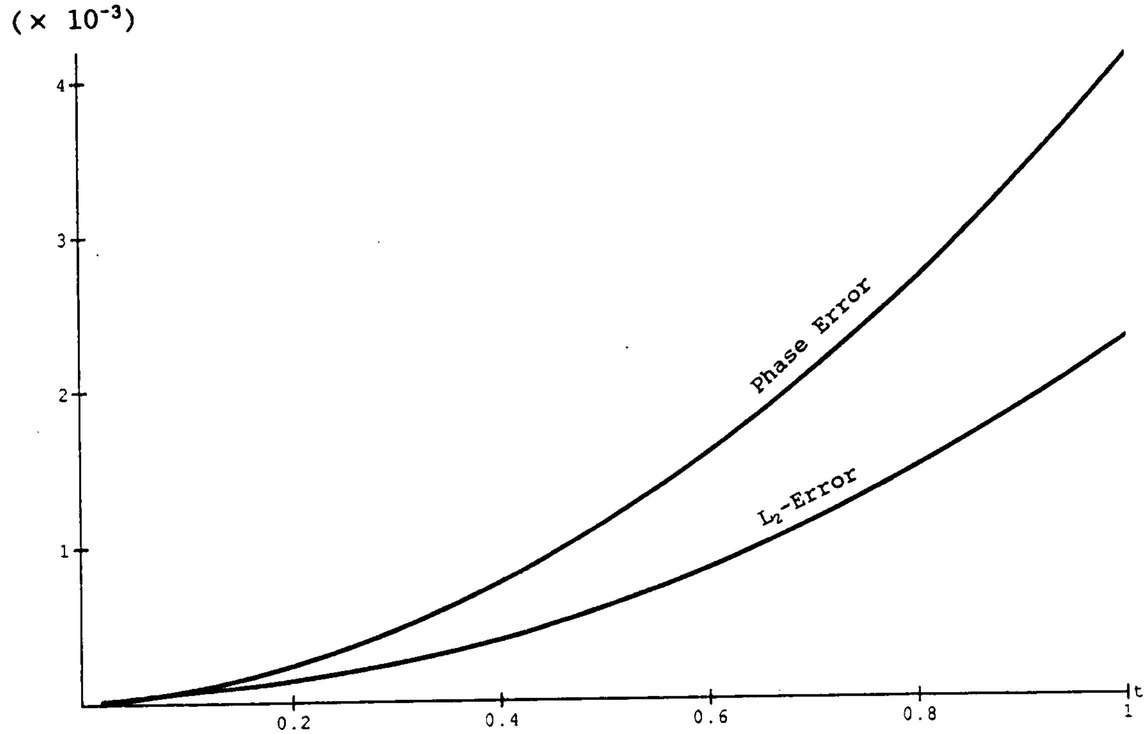

Figure 1. The growth of L_2-Error and Phase Error with h=1/100, k=1/1500 and r=4.

2-stage Gauss-Legendre Method.

Table 6: The growth of errors over time interval [0,1] with $k=1/8000$, $h=1/150$

t	E(T)	Modulus Error	Shape Error	Phase Error
0.2	0.5724(-5)	0.2246(-5)	0.1197(-5)	0.3667(-5)
0.4	0.1136(-4)	0.4015(-5)	0.1191(-5)	0.7417(-5)
0.6	0.1681(-4)	0.5928(-5)	0.1118(-5)	0.1125(-4)
0.8	0.2227(-4)	0.7894(-5)	0.1138(-5)	0.1517(-4)
1.0	0.2785(-4)	0.9949(-5)	0.1199(-5)	0.1917(-4)

Table 6 and Figure 2 show the analogous data for 2-stage Gauss Legendre method with $r=4, h=1/150$ and $N=8000$ $(k=1/8000)$. From Table 6 and Figure 2, we can conclude that $E(t)$, $M(t)$ and the phase error do not grow exponentially but grow linearly if we choose h,k sufficiently small.

Attention is now turned to the conservation of two invariants defined by (4.2) and (4.3). We monitored the quantities I_1^n, I_2^n given by

$$(4.6) \qquad I_1^n = \mathrm{I}_1(t^n) = \|U^n\|,$$

$$I_2^n = \mathrm{I}_2(t^n) = \frac{1}{40}\|U_x^n\|^2 - 64000\|U^n\|_{L^4}^4.$$

We computed with periodic splines of order $r = 4$ or 6 and evaluated the integrals in the inner products using the $2r+1$ point Gauss rule on each subinterval of the spatial

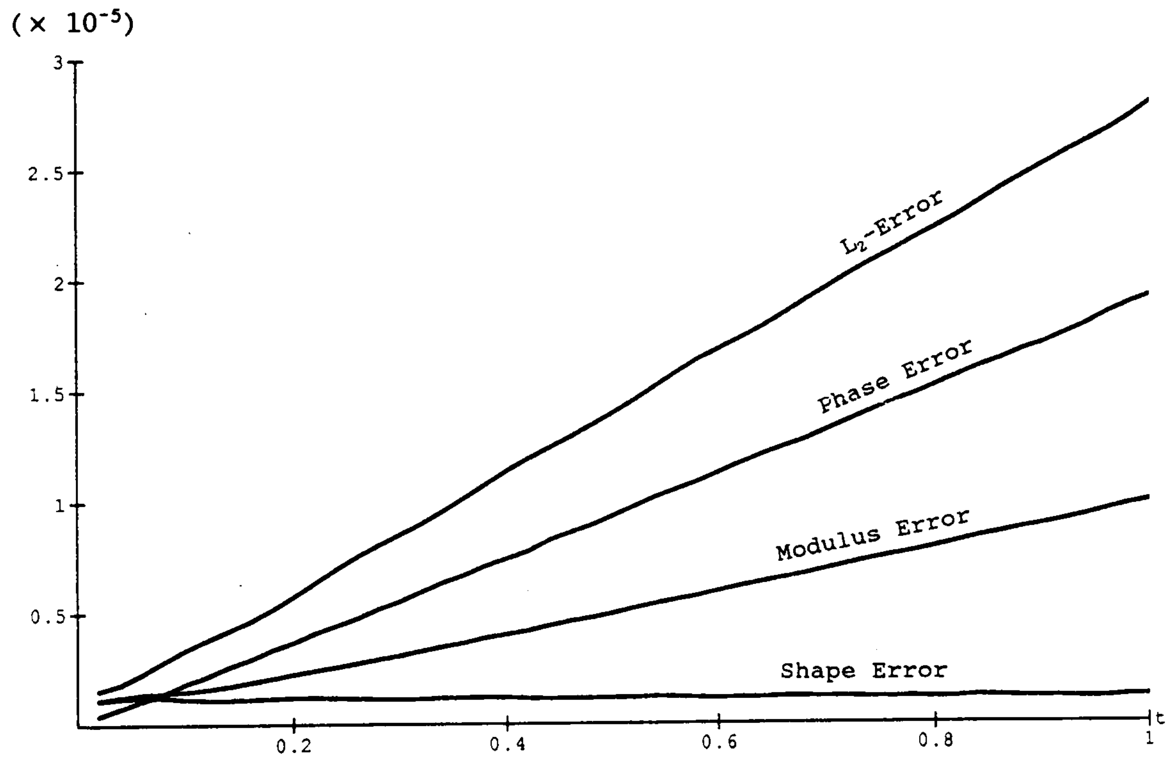

Figure 2. The growth of L_2-Error, Modulus Error, Phase Error and Shape Error with h=1/150, k=1/8000 and r=4. 2-stage Gauss-Legendre Method

discretization so that the nonlinear term is computed exactly.

Table 7: $I_1(t^n)$ and $I_2(t^n)$, 2 stage Gauss-Legendre method with cubic splines.

t^n	h=1/100	N=600	h=1/150	N=600
	$I_1(t^n)$	$I_2(t^n)$	$I_1(t^n)$	$I_2(t^n)$
0.0	0.00559	0.00229	0.0055916	0.0022916
0.125	0.00560	0.00231	0.0055913	0.0022944
0.25	0.00561	0.00235	0.0055924	0.0022970
0.375	0.00561	0.00235	0.0055934	0.0022997
0.5	0.00562	0.00237	0.0055945	0.0023024

Table 8: $I_1(t^n)$ and $I_2(t^n)$, 3 stage Gauss Legendre method with quintic splines

t^n	h=1/100	N=900	h=1/100	N=1350
	$I_1(t^n)$	$I_2(t^n)$	$I_1(t^n)$	$I_2(t^n)$
0.0	0.0055901	0.0022916	0.005590169	0.00229169
0.125	0.0055901	0.0022914	0.005590164	0.00229166
0.25	0.0055900	0.0022912	0.005590159	0.00229164
0.375	0.0055899	0.0022910	0.005590159	0.00229164
0.5	0.0055898	0.0022907	0.005590149	0.00229161

In Tables 7 and 8, we show I_1 and I_2 at t=0,0.125,0.25,0.375 and 0.5 with values of h equal to 1/100,1/150 and N equal to 600,900,1350. These Tables indicate that the code preserves well these two invariants.

Finally we close this chapter by giving some computational results which show that there is no effect when approximating the pure initial value problem (4.1a) by a periodic boundary value problem.

Table 9: The comparison of 0-boundary method and periodic boundary method

periodic method			0-boundary method		
h	N	E(T)	h	N	E(T)
1/100	100	0.7468(-3)	1.5/150	100	0.7468(-3)
1/200	200	0.3092(-4)	1.5/300	200	0.3092(-4)
1/300	300	0.5079(-5)	1.5/450	300	0.5079(-5)
1/400	400	0.1490(-5)	1.5/600	400	0.1490(-5)

To apply the spatial and temporal discretization introduced in chapter 1 with the pure initial value problem at T=0.125, we need to choose the space interval [0.0,1.5] so that the support of u(x,t), $0 \le t \le T=0.125$, stays inside [0.0,1.5]. And, as S_h, we can take for example the space of splines of order r with a 0-boundary condition on a uniform mesh with mesh length h.

By comparing two results in Table 9, to get the same order of error, the storage, work space required for 0-boundary method are 1+(speed)T times the ones of the boundary method. One can conclude that the modification of initial value problem to periodic boundary value problem is very efficient in terms of storage requirement and run time, especially when T is big.

List of References

References

[1] 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.

[2] G.A. Baker and J.H. Bramble, "Semidiscrete and single step fully discrete approximations for second order hyperbolic equation", Rapport Interne 22 (1977),'Ecole Polytechnique Centre de Math. Appliquées, Paris.

[3] G.A. Baker, V.A. Dougalis and O.A. Karakashian, "On multistep-Galerkin discretizations of semilinear hyperbolic and parabolic equations", Nonlinear Analysis, Theory, Methods & Applications, Vol.4. 1980, 579-597.

[4] J.L. Bona, V.A. Dougalis and O.A. Karakashian, "Fully discrete Galerkin methods for the Korteweg-De-Vries equation", Comp. & Maths with Appl., 124(1986), 859-884.

[5] J.L. Bona, V.A. Dougalis, O.A. Karakashian and W.R. Mckinney, "Conservative, high-order schemes for the generalized Korteweg-De-Vries equation", submitted.

[6] P. Brenner, M. Crouzeix, and V. Thomée,"Single step methods for inhomogeneous linear differential equation in Banach space", RAIRO Anal. Num., 16 1982, 5-26.

[7] H. Brezis and T. Gallouet, "Nonlinear Schrödinger evolution equations", Nonlinear Analysis, V.4 1980, 677-681.

[8] P.C. Ciarlet, "The finite element methods for elliptic problems", North-Holland, New York, 1987.

[9] M. Crouzeix, "Sur l'approximation des équation différentielles opérationelles linéaires par des méthodes de Runge-Kutta", Thèse, Université Paris VI 1975.

[10] K. Dekker and J.G. Verwer, "Stability of Runge-Kutta methods for stiff nonlinear differential equations, North-Holland, Amsterdam, 1984.

[11] M. Delfour, M. Fortin and G. Payre, "Finite-difference solutions of a nonlinear Schrödinger equation", J. Comp. Phys., V.44,1981, 277-288.

[12] A. Friedman, "Partial Differential Equations", Holt, Rinehart and Winston, New York, 1969.

[13] D.F. Griffiths, A.R. Mitchell, and J.Ll. Morris, "A numerical study of nonlinear Schrödinger equation", Comp. Methods Appl. Mech. Engrg., V.45, 1984, pp 177-215

[14] B.M. Herbst, J.Ll. Morris and A.R. Mitchell, "Numerical experience with the nonlinear Schrödinger equation", J. Comp. Phys., V.60, 1985, 282-305.

[15] V. Hutson, J.S. Pym, "Applications of functional analysis and operator theory", London, 1980.

[16] O.A. Karakashian, "On Runge-Kutta methods for parabolic problems with time dependent coefficients", Math. of Comp. V.47, 175, 1986, 77-101.

[17] O.A. Karakashian, G.D. Akrivis, V.A. Dougalis, "On optimal order error estimates for the nonlinear Schrödinger equation", submitted.

[18] S.L. Keeling, "Galerkin/Runge-Kutta discretizations for parabolic partial differential equations", Ph.D. Dissertation, University of Tennessee, 1986.

[19] B. LeMesurier, G. Papanicolaou, C. Sulem and P.L. Sulem, "The focusing singularity of the nonlinear Schrödinger equation", in Directions in Partial Differential Equations (M.G. Crandall, P.H. Rabinowitz and R.E. Turner, eds.), Academic Press, New York 1987, 159-201.

[20] A. Pazy, "Semigroups of linear operators and application to partial differential equations", Springer-Verlag, New York 1983.

[21] J.J. Rasmussen and K. Rypdal, "Blow-up in nonlinear Schrödinger equations-I. A general review", Phys. Scr., V.33, 1986, 481-497.

[22] J.M. Sanz-Serna, "Methods for the numerical solution of the nonlinear Schrödinger equation", Math. Comp., V.43, 1984, 21-27.

[23] J.M. Sanz-Serna and J.G. Verwer, "Conservative and nonconservative schemes for the solution of the nonlinear Schrödinger equation", IMA J. Numer. Anal., V.6, 1986, 25-42.

[24] W.A. Strauss, "The nonlinear Schrödinger equation", in Contemporary Developments in Continuum Mechanics and Partial Differential Equations (G.M. dela Penha and L.A.J. Medeiros, eds.), North-Holland, New York 1978, 452-465.

[25] P.L. Sulem, C. Sulem and A. Patera, "Numerical simulation of singular solutions to the two dimensional cubic Schrödinger equation", Comm. Pure Appl. Math., V.37, 1984, 755-778.

[26] T.R. Taha and M.J. Ablowitz, "Analytical and numerical aspects of certain nonlinear evolution equations II. Numerical, nonlinear Schrödinger equation", J. Comp. Phys., V.55, 1984, 203-230.

[27] V. Thomée, "Galerkin finite element methods for parabolic problems", lecture notes in Math. 1054, Springer-Verlag, New York, 1984.

[28] Y. Tourigny and J.Ll. Morris, "An investigation into the effect of product approximation in the numerical solution of the cubic nonlinear Schrödinger equation", J. Comp. Phys., V.76, 1988, 103-130.

[29] J.A.C. Weideman and B.M. Herbst, "Split-step methods for the solution of the nonlinear Schrödinger equation", SIAM J. Numer. Anal., V.23, 1986, 485-507.

VITA

Hyun Y. Lee was born in Busan, Korea on September 28, 1958. She received a Bachelor of Science degree in Mathemetics (1980) and a Master's degree in Mathematics from Busan National University (1982). In 1984, She began to study at The University of Tennessee, Knoxville, where she was awarded the Doctor of philosophy degree in Mathematics in December 1991.