Numer. Math. 23, 289—303 (1975)

Error Estimates for a Galerkin Method for a Class of Model Equations for Long Waves

Lars Wahlbin

Received December 11, 1973

Abstract. The Galerkin method, together with a second order time discretization, is applied to the periodic initial value problem for

$$\frac{\partial}{\partial t}\left(u-(a(x)\,u_x)_x\right)+\left(f(x,u)\right)_x=0.$$

Here $f(x,\cdot)$ may be highly nonlinear, but a certain cancellation effect is assumed for $\int f(x,u)_x u$. Optimal order error estimates in L_2, H_1, and L_∞ are derived for a general class of piecewise polynomial spaces.

1. Introduction

Consider the intial value problem for $u=u(x,t)$,

$$\frac{\partial}{\partial t}\left(u-(a(x)\,u_x)_x\right)+\left(f(x,u)\right)_x=0, \qquad x\in R,\ t>0, \tag{1.1.a}$$

$$u(x,0)=u_0(x). \tag{1.1.b}$$

We assume that $a(\cdot)$, $f(\cdot,y)$, and $u_0(\cdot)$ are periodic of period 1, and that there exists a positive constant c such that

$$a(x)\geqq c, \qquad x\in R. \tag{1.2}$$

In Section 2, Assumptions 2.1–2.3, we list further conditions on a, f, and u_0. These conditions will in particular ensure the existence of a periodic solution of (1.1) which is sufficiently smooth for our analysis, cf. Theorem 1.1 and Lemma 2.1.

We seek an approximate solution of (1.1) in a piecewise polynomial spline space $S_m^\mu(\Delta)$, defined as follows: Let $0=x_0<x_1<\cdots<x_n=1$ be a partition of $[0,1]$, and let Δ denote such a partition extended periodically to the real line. Let m and μ be integers with $0\leqq m<\mu-1$, and

$$\begin{aligned} S_m^\mu(\Delta)=\{f\in C^m(R)\colon f \text{ periodic, and } f \text{ is a polynomial}\\ \text{of degree}\leqq\mu-1 \text{ on each subinterval of } \Delta\}. \end{aligned} \tag{1.3}$$

The parameter μ will be fixed for the rest of this paper.

Let

$$h=\max_{i=1,\ldots,N}(x_i-x_{i-1}).$$

If the functions a and f are complicated, we use an interpolation process

$$v\curvearrowright I(v)\in S_m^\mu(\Delta)$$

into a piecewise polynomial space $S_m^\mu(\Delta)$, not necessarily the same as that where the solution is sought, to evaluate certain integrals below. Assumptions concerning this interpolation process are made in Section 2. These assumptions are restrictive, but not impractical.

We apply a Galerkin procedure, combined with a second order discretization procedure in the time variable, to Eq. (1.1.a) in weak form,

$$A(u_t, v) + F(u, v) = 0, \qquad v \in H_1, \tag{1.1.a)$'$}$$

where

$$A(w, v) = \int_0^1 wv + a(x) w_x v_x, \tag{1.4}$$

and

$$F(w, v) = -\int_0^1 f(x, w) v_x. \tag{1.5}$$

Let k denote steplength in time. We seek $U^n \in S_m^\mu(\Delta)$, $\big(U^n(\cdot) \sim u(\cdot, nk)\big)$, by the rule

$$\tilde{A}\left(\frac{U^{n+1} - U^{n-1}}{2k}, V\right) + \tilde{F}(U^n, V) = 0, \qquad V \in S_m^\mu(\Delta). \tag{1.6}$$

Here

$$\tilde{A}(W, V) = \int_0^1 WV + I(a) W_x V_x, \tag{1.7}$$

and

$$\tilde{F}(W, V) = -\int_0^1 I\big(f(x, W)\big) V_x. \tag{1.8}$$

It is assumed that U^0 and U^1 are given in $S_m^\mu(\Delta)$.

We note that the integrals occurring in (1.6) involve only piecewise polynomials. In our analysis we assume that the integrals are evaluated exactly, and that the resulting system of linear equations for the coefficients of U^{n+1} in a suitable basis for $S_m^\mu(\Delta)$ is solved exactly at each step.

We have the following main result. For notation, see (2.1), (2.2), and (2.4) for the function spaces involved, and (4.1) for $E(f)$, the elliptic projection of f along A.

Theorem 1.1. Assume that Assumptions 2.1, 2.2, and 2.3 hold. Let $T > 0$ be given, and U^0, U^1 such that $\|U^0 - U^1\|_1 = o(1)$ as $h, k \to 0$.

Then there exist constants

$$h_0 = h_0[T, \|U^0 - U^1\|_1, \|U^0\|_1, \|U^1\|_1]$$

and

$$c = c[T, \|U^0\|_1, \|U^1\|_1, \|u\|_{W_0^\infty(H_\mu, T)}, \|u\|_{W_3^\infty(H_1, T)}]$$

such that for $h \leq h_0$, $nk \leq T$,

$$\|U^n - u(\cdot, nk)\|_i \leq c\{\|U^0 - E(u_0)\|_1 + \|U^1 - E(u(\cdot, k))\|_1 + k^2 + h^{\mu - i}\},$$

$i = 0, 1$.

A method of calculating U^0 and U^1 is given in Section 5. In the special case of Hermite cubics on a uniform mesh, an easier way of calculating U^0, (the evaluation of U^1 requires no additional coding once (1.6) is coded), is given in the Appendix, (joint work with Jim Douglas Jr. and Todd Dupont). In Section 6 is derived an optimal order error estimate in the maximum norm.

The methods used in this paper are similar to those used for parabolic problems in e.g. Wheeler [15]. Galerkin methods have been applied to equations similar to (1.1.a), so called pseudoparabolic equations, in Ford [8], where an H_1 estimate was derived, and in Nassif [12]. The present paper allows stronger nonlinearities than those treated in [8]; in [12] only linear problems were considered.

The idea of using interpolants to evaluate integrals is due to Douglas-Dupont [4]. Instead of this, numerical quadrature could be applied, cf. e.g. Fix [7].

We conclude this introduction with a numerical example. Consider the equation (cf. [1])

$$\frac{\partial}{\partial t}(u-\alpha u_{xx})+(u+\beta u^2)_x=0$$

with α and β positive constants. For $\alpha=\beta=0.1$, for instance, this equation has a 7-periodic travelling wave solution $u(x,t)=v_0(x-st)$ where $s=1.093857\ldots$ and

$$v_0(y)=1.331645\ldots+w(0.492141\ldots\cdot y)^2$$

with w the canoidal function $w(z)=cn(z, 0.314542\ldots)$. The notation for the canoidal function is as in [9] (and the numerical procedures of that paper were used in the computations reported below). For the existence of such a travelling wave solution, and for the evaluation of the different parameters involved, see [10, Art. 253]. Taking $u(x,0)=v_0(x)$ the procedure (1.6) was applied in the following slightly varied form ($\langle f,g\rangle=\int_0^7 f(x)g(x)\,dx$ in this example)

$$\left\langle\frac{U^{n+1}-U^{n-1}}{2k},V\right\rangle+\alpha\left\langle\frac{(U^{n+1}-U^{n-1})}{2k}x,V_x\right\rangle$$
$$-\theta\left\langle\frac{U^{n+1}+U^{n-1}}{2},V_x\right\rangle+(\theta-1)\langle U^n,V_x\rangle-\beta\langle(U^n)^2,V_x\rangle=0$$

where $0\leqq\theta\leqq 1$. The reason for this variation is as follows: as α and β tend to zero, the method (1.6), i.e., the above with $\theta=0$, reduces to

$$\left\langle\frac{U^{n+1}-U^{n-1}}{2k},V\right\rangle-\langle U^n,V_x\rangle=0$$

which is not necessarily stable. On the other hand, if θ tends to 1 as α and β tend to zero, the limiting scheme is stable. Hence, for small α and β, the varied procedure with $\theta>0$ may conceivably have better stability properties than (1.6). The analysis for the varied scheme is almost the same as for (1.6), and the results are the same.

The space of Hermite cubics $S_1^4(\Delta)$ with Δ a uniform partition of $[0,7]$ into M subintervals of equal length was used, and the integrations involved performed exactly. The parameter θ was taken as $1/3$, U^0 was the Hermite interpolant of v_0 (cf. the Appendix) and U^1 was evaluated by the obvious modification of (5.2) below. We give a few results of the numerical calculations. The solution was calculated for $0 \leqq t \leqq 40$, and to obtain an approximation to the error in the maximum norm, the error was evaluated at $t=40$ for $x=$ (integer) $\cdot$ 0.1.

Approximate error in the maximum norm for $t=40$

$M\backslash k$	0.05	0.025	0.0125	0.00625	0.003125
4	—	9.1×10^{-3}	9.0×10^{-3}	9.0×10^{-3}	9.0×10^{-3}
8	1.8×10^{-3}	9.3×10^{-4}	7.2×10^{-4}	7.6×10^{-4}	7.7×10^{-4}
16	1.3×10^{-3}	3.5×10^{-4}	1.1×10^{-4}	6.3×10^{-5}	5.5×10^{-5}

Letting $\varepsilon(k, M)$ denote the computed maximal error at $t=40$, we have

$$\log_2(\varepsilon(0.025,4)/\varepsilon(0.00625,8)) \sim 3.6,$$
$$\log_2(\varepsilon(0.0125,4)/\varepsilon(0.003125,8)) \sim 3.5,$$
$$\log_2(\varepsilon(0.05,8)/\varepsilon(0.0125,16)) \sim 4.0,$$
$$\log_2(\varepsilon(0.025,8)/\varepsilon(0.00625,16)) \sim 3.9,$$

and

$$\log_2(\varepsilon(0.0125,8)/\varepsilon(0.003125,16)) \sim 3.7.$$

The results to be established below (see in particular the Appendix) show that for M sufficiently large the error in the maximum norm is less than $C(k^2+M^{-4})$.

The author thanks J. Bona, J. Douglas Jr. and T. Dupont for useful conversations concerning the content of this paper.

2. Notation, General Assumptions, and Preliminary Lemmas

Notation. The letters c, C, and G will denote constants, not necessarily the same at each occurrence unless subindiced. Square brackets will be used to indicate the essential dependence for these constants, e.g. $C[T]$.

Let $H_j(W_j^\infty)$ denote the real Sobolev space of functions which have j derivatives in L_2 locally (in $\mathscr{C}$), and are periodic of period 1. Let

$$\langle w, v\rangle = \int_0^1 w v$$

and

$$\|v\|_j \equiv \|v\|_{H_j} = \left(\sum_{0 \leqq i \leqq j} \left\langle \frac{d^i v}{dx^i}, \frac{d^i v}{dx^i} \right\rangle\right)^{\frac{1}{2}}, \tag{2.1}$$

$$\|v\|_{W^\infty} = \sum_{0 \leqq i \leqq j} \sup_{x \in R} \left| \frac{d^i v}{dx^i}(x) \right|. \tag{2.2}$$

We note that there exists a constant C such that for $v \in H_1$,

$$\|v\|_{W_0^\infty} \leqq C \|v\|_1. \tag{2.3}$$

For $g = g(x, t)$, let $W_j^\infty(H_i, T)$ be the closure of smooth functions in t into H_i in the norm

$$\|g\|_{W_j^\infty(H_i, T)} = \sum_{0 \leqq l \leqq j} \sup_{0<t<T} \left\| \frac{\partial^l}{\partial t^l} g(\cdot, T) \right\|_{H_i}. \tag{2.4}$$

For k given, put

$$g^n(\cdot) = g(\cdot, nk),$$

and

$$\delta_k g^n = \frac{g^{n+1} - g^{n-1}}{2k}.$$

Also, with $\tilde{A}$ as in (1.7), let

$$\|v\|_{\tilde{A}} = \tilde{A}(v, v)^{\frac{1}{2}}. \tag{2.5}$$

Assumptions for the Problem (1.1)

Recall that μ is fixed, cf. (1.3).

Assumption 2.1. The function $a(x) \in W_\mu^\infty$, and there exists a positive constant c such that

$$a(x) \geqq c, \qquad x \in R. \tag{1.2}$$

Assumption 2.2. Given an integer $0 \leqq \beta \leqq \mu$, $\frac{\partial^\beta}{\partial y^\beta} f(\cdot, y) \in W_{\mu-\beta}^\infty$, and there exists a continuous function $G[\beta, y]$ such that

$$\left\| \frac{\partial^\beta f}{\partial y^\beta}(\cdot, y) \right\|_{W_{\mu-\beta}^\infty} \leqq G[\beta, y].$$

Furthermore, with $F(w, v) = -\int_0^1 f(x, W) V_x$, there exists a constant C such that for $v \in H_1$,

$$|F(v, v)| \leqq C(1 + \|v\|_{H_1}^2). \tag{2.6}$$

Assumption 2.3. $u_0 \in H_\mu$.

In connection with Assumption 2.2, we note that since

$$f(x, w(x)) = w(x) \int_0^1 f_y'(x, w(x)\tau)\, d\tau + f(x, 0),$$

$F(w, v)$ is well-defined for $w, v \in H_1$, by (2.3). The condition (2.6) does not follow from the other assumptions.

Lemma 2.1. Let $T > 0$ be given, and assume that Assumptions 2.1–2.3 hold. Then the problem (1.1) has a periodic solution u in $W_0^\infty(H_\mu, T) \cap W_3^\infty(H_1, T)$.

A proof, using the approach of converting (1.1.a) to an integral equation, will appear in [2]. The lemma can also be proved using the method of Faedo-Galerkin, cf. [11, Chapitre 1]. The crucial assumption is (2.6), which allows us to obtain a priori bounds in L_∞, and thus handle the nonlinear term.

Assumptions and Results for the Interpolation Process

We begin by stating exactly the approximation theoretic properties of $S_m^\mu(\Delta)$ (and $S_{\bar{m}}^\mu(\Delta)$) that we need.

Lemma 2.2. There exists a constant C independent of Δ, such that given $v \in H_\nu$, $1 \leqq \nu \leqq \mu$, $(v \in W_\mu^\infty)$, there exists $X \in S_m^\mu(\Delta)$ such that

$$\|v - X\|_j \leqq C h^{\nu - j} \|v\|_\nu, \qquad j = 0, 1, \tag{2.7}$$

$$\|v - X\|_{W_0^\infty} \leqq C h^\mu \|v\|_{W_\mu^\infty}. \tag{2.8}$$

Proof. The result (2.8) is proved in [3]. The same construction of X yields (2.7) for $\nu=\mu$. For $1\leqq\nu<\mu$, (2.7) follows by mollifying v, and constructing X for the mollified function.

For the spline space $S_m^\mu(\Delta)$ and its associated linear interpolation process I, we assume that I acts as the identity on $S_m^\mu(\Delta)$, and that there exists a constant C such that

$$\|I(v)\|_{W^\infty}\leqq C\|v\|_{W_0^\infty}, \tag{2.9}$$

$$\|I(v)\|_0\leqq C(\|v\|_0+h\|v\|_1). \tag{2.10}$$

In practice, these conditions restrict the interpolation process to using only values of the function v, i.e., not values of derivatives. For simplicity, we consider the same interpolation process for evaluating A and F. The restrictions (of which only (2.10) is essential since (2.9) in practice "follows" from (2.10)) pertain to the interpolation process used for evaluating F.

Lemma 2.3. There exists a constant C such that

$$\|I(v)-v\|_0\leqq Ch^\nu\|v\|_\nu,\qquad 1\leqq\nu\leqq\mu, \tag{2.11}$$

$$\|I(v)-v\|_{W_0^\infty}\leqq Ch^\mu\|v\|_{W_\mu^\infty}. \tag{2.12}$$

Proof. Let X be as in (2.7). Since $I(X)=X$, we obtain from (2.10),

$$\begin{aligned}\|I(v)-v\|_0&=\|I(v-X)+(X-v)\|_0\\&\leqq C(\|v-X\|_0+h\|v-X\|_1)\leqq Ch^\nu\|v\|_1.\end{aligned}$$

This proves (2.11); (2.12) is proved similarly using (2.8) and (2.9).

The next four lemmas give the properties of $\tilde{A}$ and $\tilde{F}$ that will be used later.

Lemma 2.4. There exists a constant C such that for v, $w\in H_1$,

$$|\tilde{A}(w,v)-A(w,v)|\leqq Ch^\mu\|w\|_1\|v\|_\nu.$$

Proof. This is immediate from the definitions (1.4) and (1.7), using (2.12) and Assumption 2.1.

Lemma 2.5. There exist positive constants c_1 and h_1 such that for $h\leqq h_1$, and $v\in H_1$,

$$c_1^{-1}\|v\|_1\leqq\|v\|_{\tilde{A}}\leqq c_1\|v\|_1.$$

Proof. This follows from (2.5), Lemma 2.4 and Assumption 2.1.

Lemma 2.6. There exists a function G_1 such that

$$|F(w,v)-\tilde{F}(w,v)|\leqq G_1[\|w\|_\nu]\,h^\nu\|v\|_1,\qquad 1\leqq\nu\leqq\mu.$$

Proof. We have by (2.11)

$$\begin{aligned}|F(w,v)-\tilde{F}(w,v)|&=\Big|\int_0^1\big(f(x,w)-I(f(x,w))\big)v_x\Big|\\&\leqq\|f(x,w)-I(f(x,w))\|_0\|v\|_1\\&\leqq Ch^\nu\|f(x,w)\|_\nu\|v\|_1.\end{aligned}$$

Using (2.3) and Assumption 2.2, we find that $\|f(x,w)\|_\nu$ can be bounded in terms of $\|w\|_\nu$.

Lemma 2.7. There exists a function G_2 such that

$$|\widetilde{F}(w_1,v)-\widetilde{F}(w_2,v)| \leqq G_2[\|w_1\|_1, \|w_2\|_1]\,\|v\|_1(\|w_1-w_2\|_0+h\|w_1-w_2\|_1).$$

Proof. We have

$$\begin{aligned}|\widetilde{F}(w_1,v)-\widetilde{F}(w_2,v)| &= \Big|\int_0^1 I\big(f(x,w_1)-f(x,w_2)\big)v_x\Big| \\ &\leqq C\big(\|f(x,w_1)-f(x,w_2)\|_0+h\|f(x,w_1)-f(x,w_2)\|_1\big)\|v\|_1.\end{aligned}$$

Since

$$f(x,w_1)-f(x,w_2)=(w_1-w_2)\int_0^1 f'_y\big(x,w_1+\tau(w_2-w_1)\big)\,d\tau,$$

the desired estimate obtains.

3. Stability in H_1 of the Galerkin Process

Throughout this section it is understood that $h\leqq h_1$, so that the conclusion of Lemma 2.5 holds. We shall show that the H_1-norm of the approximate solution is bounded as h, k tend to zero.

Let

$$e^j=U^j-U^{j-1},\qquad j=1,2,\ldots$$

We start with a preliminary result:

Lemma 3.1. Let c_1 be as in Lemma 2.5, and $G_2[\cdot,\cdot]$ as in Lemma 2.7. There exists a function $c_2[\cdot]$ such that

$$\|e^j\|_1\leqq c_1(\|e^1\|_{\tilde{A}}+kc_2[\|U^1\|_1])\cdot\prod_{l=1}^{j-1}(1+2kc_1^2G_2[\|U^l\|_1,\|U^{l-1}\|_1]).$$

Proof. Let, with c_2 defined below,

$$a_1=\|e^1\|_{\tilde{A}}+kc_2[\|U^1\|_1], \tag{3.1}$$

$$b_l=2kc_1^2G_2[\|U^l\|_1,\|U^{l-1}\|_1]. \tag{3.2}$$

Consider first the case $j=2$ ($j=1$ is trivial). Let $n=1$ and $V=e^2-e^1$ in (1.6), and note that $U^2-U^0=e^2+e^1$. Using Lemma 2.7 with $w_2=0$ it follows that

$$\frac{\|e^2\|_{\tilde{A}}^2-\|e^1\|_{\tilde{A}}^2}{2k}=-\widetilde{F}(U^1,e^2-e^1)\leqq G_2[\|U^1\|_1,0]\,\|U^1\|_1(\|e^2\|_1+\|e^1\|_1).$$

Hence, with this defining c_2,

$$\|e^2\|_{\tilde{A}}\leqq\|e^1\|_{\tilde{A}}+kc_2[\|U^1\|_1].$$

Assume now inductively that

$$\|e^j\|_{\tilde{A}}\leqq a_1\prod_{l=1}^{j-1}(1+b_l). \tag{3.3}$$

Write (1.6) at two adjacent time levels, and put $V = e^{j+1} - e^{j-1}$. We then obtain

$$\frac{\|e^{j+1}\|_{\tilde{A}}^2 - \|e^{j-1}\|_{\tilde{A}}^2}{2k} = -\left(\tilde{F}(U^j, e^{j+1} - e^{j-1}) - \tilde{F}(U^{j-1}, e^{j+1} - e^{j-1})\right)$$
$$\leqq c_1 G_2[\|U^j\|_1, \|U^{j-1}\|_1]\, \|e^j\|_1 (\|e^{j+1}\|_{\tilde{A}} + \|e^{j-1}\|_{\tilde{A}}),$$

or

$$\|e^{j+1}\|_{\tilde{A}} \leqq \|e^{j-1}\|_{\tilde{A}} + b_j \|e^j\|_{\tilde{A}}.$$

By the induction hypothesis (3.3) we then have

$$\|e^{j+1}\|_{\tilde{A}} \leqq a_1 \prod_{l=1}^{j-1} (1 + b_l) \left(\frac{1}{1 + b_{j-1}} + b_j\right) \leqq a_1 \prod_{l=1}^{j} (1 + b_l),$$

and (3.3) is proven. The result of the lemma is now immediate.

We can now prove the stability result:

Lemma 3.2. Let $T > 0$ be given, and assume that $\|e^1\|_1 = o(1)$ as $h, k \to 0$. Then there exist h_0 and c_3,

$$h_0 = h_0[T, \|e^1\|_1, \|U^0\|_1, \|U^1\|_1],$$
$$c_3 = c_3[T, \|U^0\|_1, \|U^1\|_1]$$

such that for $h \leqq h_0$, $nk \leqq T$,

$$\|U^n\|_1 \leqq c_3.$$

Proof. We first note, taking $\nu = 1$, $w = v = U^n$ in Lemma 2.6, and using (2.6) that

$$|\tilde{F}(U^n, U^n)| \leqq C(1 + \|U^n\|_{\tilde{A}}^2) + hG_1[\|U^n\|_1]\, \|U^n\|_1.$$

Let $V = U^{n+1} + U^{n-1}$ in (1.6). Then

$$\frac{\|U^{n+1}\|_{\tilde{A}}^2 - \|U^{n-1}\|_{\tilde{A}}^2}{2k} \leqq |\tilde{F}(U^n, U^{n+1} + U^{n-1})|$$
$$\leqq |\tilde{F}(U^n, U^{n+1} - U^n - (U^n - U^{n-1}))|$$
$$+ c(1 + \|U^n\|_{\tilde{A}}^2) + 2hG_1[\|U^n\|_1]\, \|U^n\|_1.$$

By Lemma 2.6 with $w_2 = 0$,

$$|\tilde{F}(U^n, e^{n+1} - e^n)| \leqq G_2[\|U^n\|_1, 0] \cdot \|U^n\|_1 (\|e^{n+1}\|_1 + \|e^n\|_1).$$

Hence we obtain, this defining G_3 and c_4,

$$\frac{\|U^{n+1}\|_{\tilde{A}}^2 - \|U^{n-1}\|_{\tilde{A}}^2}{2k} \leqq G_3[\|U^n\|_1] (\|e^{n+1}\|_1 + \|e^n\|_1 + h) + c_4(1 + \|U^n\|_{\tilde{A}}^2),$$

or, in the notation of (3.1), (3.2), by Lemma 3.1,

$$\frac{\|U^{n+1}\|_{\tilde{A}}^2 - \|U^{n-1}\|_{\tilde{A}}^2}{2k} \leqq G_3[\|U^n\|_1] \left(2a_1 \prod_{j=1}^{n} (1 + b_j) + h\right) + c_4(1 + \|U^n\|_{\tilde{A}}^2). \tag{3.4}$$

With $c_5 k < 1$, c_5 to be specified later, and $c_6 = \max(\|U^0\|_{\tilde{A}}, \|U^1\|_{\tilde{A}})$, we assume by induction that

$$\|U^j\|_{\tilde{A}} \leqq c_6 \exp(c_5 j k), \qquad j \leqq n. \tag{3.5}$$

Then

$$\prod_{j=1}^{n} (1+b_i) \leqq G_4[T, c_5, c_6], \qquad nk \leqq T.$$

Inserting into (3.4) we obtain, assuming $c_6 \geqq 1$,

$$\begin{aligned}\|U^{n+1}\|_{\tilde{A}}^2 &\leqq \|U^{n-1}\|_{\tilde{A}}^2 + 2kG_3(2G_4 a_1 + h) + 2kc_4(1 + \|U^n\|_{\tilde{A}}^2) \\ &\leqq c_6^2 \exp(2nkc_5) \cdot (e^{-2c_5 k} + 2kG_3(2G_4 a_1 + h) + 4kc_4) \\ &\leqq c_6^2 \exp(2nkc_5) \cdot \Big(1 + 2kc_5 \Big\{2\frac{c_4}{c_5} + \frac{1}{c_5}[c_6 \exp(c_5 T)] \cdot \\ &\qquad \cdot (2G_4[T, c_5, c_6] \cdot a_1 + h)\Big\}\Big).\end{aligned}$$

Hence, taking e.g., c_5 such that $2\frac{c_4}{c_5} = \frac{1}{2}$, the induction step will work provided h and $a_1 = \|e^1\|_{\tilde{A}} + kc_2[\|U^1\|_1]$ are small enough, i.e. provided h and k are small enough. For fixed T, the restriction on k is automatically removed.

This proves (3.5), and concludes the proof of the lemma.

4. Error Estimates in H_1 and H_0

Let $Z = Z(x,t)$ denote the elliptic projection of $u(x,t)$ along A, $Z(\cdot,t) = E(u(\cdot,t)) \in S_m^\mu(\Delta)$, defined by

$$A(Z, V) = A(u, V), \qquad V \in S_m^\mu(\Delta). \tag{4.1}$$

Here u is the solution to (1.1). Let

$$\xi^n = U^n - Z^n,$$
$$\eta^n = u^n - Z^n.$$

We note that by Taylor expansion around $t = nk$ we have

$$\delta_k\big(u^n - (a(x)u_x^n)_x\big) + \big(f(x, u^n)\big)_x = R_n, \tag{4.2}$$

where

$$|\langle R_n, v\rangle| \leqq ck^2 \|w\|_{W_3^\infty(H_1, T)} \|v\|_1. \tag{4.3}$$

The following lemma summarizes the results we need for the elliptic projection:

Lemma 4.1. There exists a constant c such that for $nk \leqq T$,

$$\sup_{0 \leqq j \leqq n} \|\eta^j\|_i \leqq ch^{\mu - i} \|u\|_{W_0^\infty(H_\mu, T)}, \qquad i = 0, 1, \tag{4.4}$$

$$\sup_{0 \leqq j \leqq n} \|Z^j\|_1 \leqq c\|u\|_{W_0^\infty(H_1, T)}, \tag{4.5}$$

$$\sup_{0 \leqq j \leqq n} \|\delta_k Z^j\|_1 \leqq c\|w\|_{W_1^\infty(H_1, T)}. \tag{4.6}$$

Proof. See e.g. [13, 14].

We now start to derive the error estimates. We obtain from (1.6) and (4.2),

$$\begin{aligned}\tilde{A}(\delta_k \xi^n, V) &= \tilde{A}(\delta_k U^n, V) - \tilde{A}(\delta_k Z^n, V)\\
&= -\tilde{F}(U^n, V) - A(\delta_k u^n, V) + A(\delta_k Z^n, V) - \tilde{A}(\delta_k Z^n, V)\\
&= -\tilde{F}(U^n, V) + F(u^n, V) + A(\delta_k Z^n, V) - \tilde{A}(\delta_k Z^n, V) + \langle R_n, V\rangle\\
&= \{-\tilde{F}(U^n, V) + \tilde{F}(u^n, V)\} + \{F(u^n, V) - \tilde{F}(u^n, V)\}\\
&\quad + \{A(\delta_k Z^n, V) - \tilde{A}(\delta_k Z^n, V)\} + \langle R_n, V\rangle.\end{aligned}$$

Let $V = \xi^{n+1} + \xi^{n-1}$. We use Lemma 2.7 to estimate the first term on the right, Lemma 2.6 with $\nu = \mu$ for the second, Lemma 2.4 and (4.6) for the third, and (4.3) for the last term. We obtain

$$\begin{aligned}\frac{\|\xi^{n+1}\|_{\tilde{A}}^2 - \|\xi^{n-1}\|_{\tilde{A}}^2}{2k} &\leq \{G_2[\|U^n\|_1, \|u^n\|_1](\|U^n - u^n\|_0 + h\|U^n - u^n\|_1)\\
&\quad + G_1[\|u^n\|_\mu] h^\mu + c h^\mu \|u\|_{W_1^\infty(H_1, T)}\\
&\quad + c k^2 \|u\|_{W_3^\infty(H_1, T)}\}(\|\xi^{n+1}\|_1 + \|\xi^{n-1}\|_1),\end{aligned}$$

or, using that $U^n - u^n = \xi^n - \eta^n$ and (4.4),

$$\begin{aligned}\frac{\|\xi^{n+1}\|_{\tilde{A}} - \|\xi^{n-1}\|_{\tilde{A}}}{2k} &\leq c\{\|\xi^n\|_0 + h\|\xi^n\|_1 + \|\eta^n\|_0 + h\|\eta^n\|_1\\
&\quad + h^\mu(G_1[\|u^n\|_\mu] + \|u\|_{W_1^\infty(H_1, T)}) + k^2\|u\|_{W_3^\infty(H_1, T)}\}\\
&\leq \frac{c_7}{2}\{\|\xi^n\|_{\tilde{A}} + h^\mu + k^2\},\end{aligned}$$

$$c_7 = c_7[\sup_{jk \leq T} \|U^j\|_1, \|u\|_{W_0^\infty(H_\mu, T)}, \|u\|_{W_3^\infty(H_1, T)}].$$

Hence

$$\|\xi^{n+1}\|_{\tilde{A}} + \|\xi^n\|_{\tilde{A}} \leq (1 + c_7 k)(\|\xi^n\|_{\tilde{A}} + \|\xi^{n-1}\|_{\tilde{A}}) + c_7 k(h^\mu + k^2)$$

and it follows that

$$\|\xi^n\|_1 \leq c_8(\|\xi^0\|_1 + \|\xi^1\|_1 + k^2 + h^\mu), \qquad nk \leq T, \tag{4.7}$$

$$c_8 = c_8[T, \sup_{jk \leq T} \|U^j\|_1, \|u\|_{W_0^\infty(H_\mu, T)}, \|u\|_{W_3^\infty(H_1, T)}]. \tag{4.8}$$

By (4.4) we obtain

$$\|U^n - u^n\|_i \leq c\{\|\xi^0\|_1 + \|\xi^1\|_1 + k^2 + h^{\mu - i}\}, \qquad nk \leq T,\ i = 0, 1 \tag{4.9}$$

where the constant has the same dependence as in (4.8).

Proof of Theorem 1.1. By Lemma 3.2, U^n is bounded in H_1 for $h \leq h_0$, $nk \leq T$, the bound depending on $\|U^0\|_1$ and $\|U^1\|_1$. Hence the theorem follows from (4.9).

5. A Method of Calculating U^0 and U^1

We shall give a general method for finding U^0 and U^1, close to the elliptic projections of u_0 and $u(\cdot, k)$ respectively.

In addition to the general assumptions of Section 2, we assume in this section that

$$u_0 \in H_{\mu+1}.$$

Let U^0 be given by the rule

$$\tilde{A}(U^0, V) = \int_0^1 I(u_0) V + \int_0^1 I\big(a(u_0)_x\big) V_x, \qquad V \in S_m^\mu(\Delta), \tag{5.1}$$

and U^1 by

$$\tilde{A}\left(\frac{U^1 - U^0}{k}, V\right) + \tilde{F}(U^0, V) = 0, \qquad V \in S_m^\mu(\Delta). \tag{5.2}$$

Using the same techniques as in Section 4, it is easy to prove that

$$\|\xi^0\|_1 \leqq c h^\mu \|u_0\|_{\mu+1},$$

where c only depends on a, and that

$$\|\xi^1\|_1 \leqq c(h^\mu + k^2) + \|\xi^0\|_1, \tag{5.3}$$

where $c = c[\|U^0\|_1, \|u_0\|_\mu, \|u\|_{W_2^\infty(H_1, k)}]$.

Taking $V = U^1 - U^0$ in (5.2) we see that

$$\|U^1 - U^0\|_1 \leqq c[\|U^0\|_1]\, k.$$

Combining those results with Theorem 1.1 the following theorem obtains:

Theorem 5.1. Assume that Assumptions 2.1 and 2.2. hold, and that $u_0 \in H_{\mu+1}$. Let $T > 0$ be given, and let U^0 and U^1 be given by (5.1) and (5.2).

Then there exist constants h_0 and c,

$$h_0 = h_0[T, \|u_0\|_{\mu+1}]$$

and

$$c = c[T, \|u_0\|_{\mu+1}, \|u\|_{W_0^\infty(H_\mu, T)}, \|u\|_{W_3^\infty(H_1, T)}]$$

such that for $h \leqq h_0$, $nk \leqq T$,

$$\|U^n - u(\cdot, nk)\|_i \leqq c(k^2 + h^{\mu - i}), \qquad i = 0, 1.$$

6. Optimal Order Error Estimates in Maximum Norm

We assume in this section that the partitions Δ considered are quasiuniform, i.e., that there exists a constant c such that

$$h \leqq c \min_{i=1,\dots,N} (x_i - x_{i-1}). \tag{6.1}$$

We have the following result for the error in the elliptic projection, cf. (4.1).

Lemma 6.1. Assume that (6.1) holds. Then there exists a constant C such that

$$\|v - E(v)\|_{W_0^\infty} \leqq C h_\mu \|v\|_{W_\mu^\infty}.$$

Proof. The proof of this will appear in [5]. For $m = 0$, i.e., continuous piecewise polynomials, the result holds without the assumption of quasiuniformicity, see [16].

Let

$$\|u\|_{W_0^\infty(W_\mu^\infty, T)} = \sup_{0<t\leq T} \|u(\cdot, t)\|_{W_\mu^\infty},$$

and let $W_0^\infty(W_\mu^\infty, T)$ denote the corresponding space. We have the following variant of Theorem 5.1.

Theorem 6.1. Assume that the hypotheses of Theorem 5.1 hold, that (6.1) holds, and that the solution of (1.1) belongs to $W_0^\infty(W_\mu^\infty, T)$.

Then there exist constants h_0 and c,

$$c = c[T, \|u_0\|_{\mu+1}, \|u\|_{W_0^\infty(W_\mu^\infty, T)}, \|u\|_{W_3^\infty(H_1, T)}]$$

such that for $h \leq h_0$, $nk \leq T$,

$$\|U^n - u(\cdot, nk)\|_{W_0^\infty} \leq c(k^2 + h^\mu).$$

Proof. By (4.7) and (2.3) we have

$$\|\xi^n\|_{W_0^\infty} \leq c(k^2 + h^\mu).$$

Hence Lemma 6.1 gives

$$\|U^n - u^n\|_{W_0^\infty} \leq \|u^n - E(u^n)\|_{W_0^\infty} + \|\xi^n\|_{W_0^\infty} \leq c(k^2 + h^\mu).$$

Remark 6.1. The solution will belong to $W_0^\infty(W_\mu^\infty, T)$ if it belongs to $W_0^\infty(H_{\mu+1}, T)$. Using Lemma 2.1, conditions which guarantee this are easily found in terms of a, f, and u_0.

Appendix

Joint work with Jim Douglas Jr. and Todd Dupont

We shall prove that in the case of Hermite cubics on a uniform mesh, the Hermite interpolant is close to the elliptic projection. Hence if U^0 is chosen as the Hermite interpolant of u_0, and U^1 according to (5.2), the optimal orders of convergence in H_0, H_1 and L_∞ are retained provided u_0 is smooth enough (see (A.1), (5.3), and the proof of Theorem 6.1).

Consider $S_1^4(\Delta)$, with $\Delta = \{x_i\}_{i\in Z}$, $x_i = ih$, $h^{-1} = N \in Z$. We introduce a basis $\{V_i, S_i\}$, $i = 1, \ldots, N$, for this space; V_i and S_i are the periodic extensions of $\overline{V}_i$ and $\overline{S}_i$, where $\overline{V}_i(x) = V(xh^{-1} - i)$, $\overline{S}_i(x) = S(xh^{-1} - i)$, and

$$V(y) = \begin{cases} 0, \quad y \geq 1, \quad \text{or} & y \leq -1, \\ (y-1)^2(2y+1), & 0 \leq y \leq 1, \\ (y+1)^2(-2y+1), & -1 \leq y \leq 0, \end{cases}$$

$$S(y) = \begin{cases} 0, \quad y \geq 1, \quad \text{or} \quad y \leq -1, \\ y(y-1)^2, & 0 \leq y \leq 1, \\ y(y+1)^2, & -1 \leq y \leq 0. \end{cases}$$

The Hermite interpolant of a function $u(x)$ is

$$I(u) = \sum_{i=1}^{N} u(ih) V_i + u'(ih) h S_i.$$

We shall prove that

$$\|I(u)-E(u)\|_1 \leqq ch^4 \|u\|_5. \tag{A.1}$$

It suffices to consider the case when

$$A(w,v)=\int_0^1 wv+w_x v_x,$$

see e.g. [5, Section 3]. With $z=I(u)-E(u)$ we then have

$$\|z\|_1^2=\int_0^1 (I(u)-u)z+\int_0^1 (I(u)-u)_x z_x. \tag{A.2}$$

It is well known that $\|I(u)-u\|_0 \leqq ch^4\|u\|_4$, and hence

$$\left|\int_1^0 (I(u)-u)z\right| \leqq ch^4 \|u\|_4 \|z\|_0 \leqq \tfrac{1}{10}\|z\|_0^2+ch^8\|u\|_4^2. \tag{A.3}$$

To handle the second term in (A.2), we use the fact (see e.g. [6, (3.48)]) that on (x_i, x_{i+1}) we have

$$I(u)-u=\tfrac{1}{2}(L_i^+ + L_{i+1}^- + R_i^+ + R_{i+1}^-), \tag{A.4}$$

where

$$L_i^+(x)=\frac{1}{4!}u^{(4)}(x_i)(x-x_i)^2(x_{i+1}-x)^2,$$

$$L_{i+1}^-(x)=\frac{1}{4!}u^{(4)}(x_{i+1})(x-x_i)^2(x_{i+1}-x)^2,$$

$$R_i^+(x)=\int_{x_i}^{x_{i+1}} r_i^+(x,t)\left(u^{(4)}(t)-u^{(4)}(x_i)\right)dt,$$

$$R_{i+1}^-(x)=\int_{x_i}^{x_{i+1}} r_{i+1}^-(x,t)\left(u^{(4)}(t)-u^{(4)}(x_{i+1})\right)dt,$$

and

$$\left|\frac{\partial}{\partial x} r_i^{\pm}(x,t)\right| \leqq ch^2. \tag{A.5}$$

Let

$$z=\sum_{i=1}^N a_i V_i + b_i S_i.$$

We have

$$\int_0^1 (I(u)-u)_x z_x=\tfrac{1}{2}\sum_{i=0}^{N-1}\int_{x_i}^{x_{i+1}} (L_i^+ + L_{i+1}^- + R_i^+ + R_{i+1}^-)_x z_x.$$

Using (A.5) it follows that

$$\left|\tfrac{1}{2}\sum_{i=0}^{N-1}\int_{x_i}^{x_{i+1}} (R_i^+ + R_{i+1}^-)_x z_x\right| \leqq \tfrac{1}{10}\|z\|_1^2+ch^8\|u\|_5^2. \tag{A.6}$$

By computation we find that

$$\int_{x_i}^{x_{i+1}} (L_i^{\pm})_x (V_j)_x=0, \qquad \text{any } j,$$

and, with subscripts interpreted modulo N if necessary,

$$\int_{x_{i-1}}^{x_i} (L_i^-)_x (S_i)_x + \int_{x_i}^{x_{i+1}} (L_i^+)_x (S_i)_x = 0.$$

Inserting (A.3) and (A.6) into (A.2) we hence obtain

$$\begin{aligned} \|z\|_1^2 &\leq c h^8 \|u\|_5^2 + c \left| \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \left(b_i (L_{i+1}^-)_x (S_i)_x + b_{i+1} (L_i^+)_x (S_{i+1})_x\right) \right| \\ &= c h^8 \|u\|_5^2 + c h^3 \left| \sum_{i=0}^{N-1} b_i \left(u^{(4)}(x_{i+1}) - u^{(4)}(x_{i-1})\right) \right|. \end{aligned} \tag{A.7}$$

We next note that

$$|u^{(4)}(x_{i+1}) - u^{(4)}(x_{i-1})| \leq c h^{\frac{1}{2}} \left(\int_{x_{i-1}}^{x_{i+1}} |u^{(5)}(t)|^2 dt \right)^{\frac{1}{2}},$$

and that, due to the fact that z is a piecewise polynomial,

$$|z'(x_i)|^2 \leq c h^{-1} \int_{x_i}^{x_{i+1}} |z'|^2.$$

Since $b_i = h z'(x_i)$, we obtain from (A.7),

$$\|z\|_1^2 \leq c h^8 \|u\|_5^2 + \frac{1}{10} \|z\|_1^2$$

which proves (A.1).

Bibliography

1. Benjamin, T. B., Bona, J. L., Mahoney, J. J.: Model equations for long waves in nonlinear dispersive systems. Philos. Trans. R. Soc. London, A **272**, 47–78 (1972)
2. Benjamin, T. B., Bona, J. L.: Model equations for long waves in nonlinear dispersive systems II. To appear.
3. de Boor, C.: On local spline approximation by moments. J. Math. Mech. **17**, 729–735 (1968)
4. Douglas, J., Dupont, T.: The effect of interpolating the coefficients in non-linear parabolic Galerkin procedures. To appear
5. Douglas, J., Dupont, T., Wahlbin, L.: Optimal L_∞ error estimates for Galerkin approximations to solutions of two point boundary value problems. To appear in Math. Comp.
6. Dupont, T.: Galerkin methods for first order hyperbolics: an example. SIAM J. Numer. Anal. **10**, 890–899 (1973)
7. Fix, G.: Effects of quadrature errors in finite element approximation of steady state, eigenvalue and parabolic problems. Proceedings of a Symposium on the mathematical foundations of the finite element method at the University of Maryland, June 1972. New York: Academic Press 1972
8. Ford, W.: Semidiscrete Galerkin approximations to non-linear pseudoparabolic partial differential equations. To appear in Aequationes Math.
9. Hofsommer, D. J., van de Riet, R. P.: On the numerical calculation of elliptic integrals of the first and second kind and the elliptic functions of Jacobi. Numer. Math. **5**, 291–302 (1963)
10. Lamb, H.: Hydrodynamics, 6th edition. New York: Dover Publications 1945
11. Lions, J. L.: Quelques méthodes de résolution des problèmes aux limites non linéaires. Paris: Dunod 1969
12. Nassif, N. R.: Thesis, Harvard University, 1972

13. Nitsche, J.: Ein Kriterium für die Quasioptimalität des Ritzschen Verfahrens. Numer. Math. **11**, 346–348 (1968)
14. Nitsche, J.: Verfahren von Ritz und Spline Interpolation bei Sturm-Liouville Randwertproblemen. Numer. Math. **13**, 260–265 (1969)
15. Wheeler, M. F.: A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations. SIAM J. Numer. Anal. **10**, 723–759 (1973)
16. Wheeler, M. F.: An optimal L_∞ error estimate for Galerkin approximations to the solution of two point boundary value problems. SIAM J. Numer. Anal. **10**, 914–917 (1973)

Lars Wahlbin
Dept. of Mathematics
White Hall
Cornell University
Ithaca, N.Y. 14850/U.S.A.