IMA Journal of Numerical Analysis (1981) **1**, 215–221

Convergence of Methods for the Numerical Solution of the Korteweg–de Vries Equation

KUO PEN-YU†

Département Mathématiques, Collège de France, Paris, France

AND

J. M. SANZ-SERNA

Departamento de Matematicas,
Facultad de Ciencias, Valladolid, Spain

[Received 18 March 1980]

We prove that a family of methods for the numerical solution of the Korteweg–de Vries equation is convergent. This family includes as particular cases some known finite difference and finite element schemes. It is also found that the stability properties of the methods vary significantly with the treatment of the non-linear term.

1. Introduction

THE NUMERICAL SOLUTION of the Korteweg–de Vries (KdV) equation has been the subject of many papers over the last few years. Zabusky & Kruskal (1965) employed a second order accurate leap-frog scheme. Dissipative difference methods were considered by Vliegenthart (1971) and the Hopscotch technique by Greig & Morris (1976). More recently Kuo Pen-Yu (1978) has studied a family of explicit difference methods. Galerkin methods for the KdV equation were analysed by Wahlbin (1974) and implemented by Alexander & Morris (1979). Sanz-Serna & Christie (1979) have introduced a modified Petrov–Galerkin technique. Finite Fourier transform methods have also been applied to the KdV equation (Canosa & Gazdag, 1977).

In this paper we obtain results concerning the convergence and stability of a family of methods which includes as particular cases some of the schemes above. (Note that most of the quoted papers restrict their study of stability to the *linearized* KdV equation.) Our analysis is based on a perturbation technique which is also capable of handling other non-linear partial differential equations (Kuo Pen-Yu, 1965, 1977, 1978, 1979*a*, *b*, 1980*a*, *b*, *c*, *d*).

For simplicity we consider the periodic initial-value problem

$$\begin{aligned} Lu(x,t) &= u_t + uu_x + u_{xxx} = f(x,t), \\ u(x,0) &= u_0, \quad -\infty < x < \infty, \end{aligned} \tag{1.1}$$

where u_0, f have period 1 with respect to x and satisfy adequate smoothness requirements (cf. Sjöberg, 1970). The study of the Cauchy problem is completely analogous and will not be considered here. The problem (1.1) will be discretized in

† On leave from Shanghai University of Science and Technology, Shanghai, China.

0272–4979/81/020215+07 $02.00/0

space but not in time, so as to approximate it by a system of ordinary differential equations. Our notations are as follows. The net spacing is called h, with $h = 1/N$, N integer; $V_i(t)$, or simply V_i, denote the value $V(ih, t)$ of a mesh function. We set

$$(V, W) = \sum_{i=1}^{N} hV_iW_i, \tag{1.2}$$

$$\|V\|^2 = (V, V), \tag{1.3}$$

$$V_{\hat{x},i} = (V_{i+1} - V_{i-1})/2h, \tag{1.4}$$

$$V_{x,i} = (V_{i+1} - V_i)/h, \tag{1.5}$$

$$V_{\bar{x},i} = (V_i - V_{i-1})/h, \tag{1.6}$$

$$V_{\underset{\sim}{x},i} = (V_{i+2} - V_{i-2})/4h, \tag{1.7}$$

For each operator introduced there are two possible replacements for the term uu_x, as this can also be written in the form $(u^2)_x/2$. In fact we consider a blend between those two alternatives and introduce, for $0 \leqslant \alpha \leqslant 1$, the operators

$$J^{(\alpha)}(V, W)_i = \alpha W_i V_{\hat{x},i} + \frac{1-\alpha}{2}(WV)_{\hat{x},i}, \tag{1.8}$$

$$I^{(\alpha)}(V, W)_i = \alpha W_i V_{\underset{\sim}{x},i} + \frac{1-\alpha}{2}(WV)_{\underset{\sim}{x},i}. \tag{1.9}$$

It is well known that for $\alpha = \frac{1}{3}$, $J^{(\alpha)}(U, U)$, $I^{(\alpha)}(U, U)$ are, respectively, the replacements obtained for uu_x, when the Galerkin method based on piecewise linear interpolants with element size h, $2h$ is used.

Furthermore, if β is real, set

$$K^{(\alpha,\beta)}(V, W) = \beta J^{(\alpha)}(V, W) + (1-\beta) I^{(\alpha)}(V, W). \tag{1.10}$$

In order to accommodate the possibility of a mass matrix, we define

$$MV_i = \sum_{j=-2}^{2} M_j V_{i+j}, \tag{1.11}$$

where M_j are real numbers such that their sum is unity, $M_j = M_{-j}$, $j = 1, 2$ and there exist constants C_0, C_1, independent of h, with

$$C_0\|V\|^2 \leqslant (MV, V) \leqslant C_1\|V\|^2,$$

for arbitrary V. Hypotheses of this sort are naturally fulfilled by the mass matrices which arise in the finite element method.

With these notations we consider the following approximation to (1.1)

$$\begin{aligned} (L_hU)_i &= M\left(\frac{\partial U}{\partial t}\right)_i + K^{(\alpha,\beta)}(U, U)_i + U_{xx\hat{x},i} = f_i, \\ U_i(0) &= u_{0,i}. \end{aligned} \tag{1.12}$$

There is no substantial change in the subsequent analysis if the right-hand side of (1.12) is replaced by Mf_i.

2. Results

Consider, for the time being, h fixed, and suppose that in (1.12). f, u_0 are perturbed and become $f+\tilde{f}$, $u_0+\tilde{u}_0$, respectively. Let $\tilde{U}$ denote the resulting perturbation in the solution U. Fix an interval $(0, T)$ in the t-axis, where $0 < T < \infty$ and assume that $\tilde{f}$, $\tilde{u}_0$ are such that

$$\tilde{\rho} = ||\tilde{u}_0||^2 + \int_0^T ||\tilde{f}(\tau)||^2 \, d\tau, \tag{2.1}$$

is finite. Then:

THEOREM 1. *There exists a constant* $C > 0$, *which depends only on* U, *such that for* $0 < t < T$

$$||\tilde{U}(t)||^2 \leqslant \begin{cases} C\tilde{\rho}\, e^{Ct}, & \text{if } \alpha = \frac{1}{3}, \\ \dfrac{C\tilde{\rho}\, e^{Ct}}{1+C\tilde{\rho}h^{-3}(1-e^{Ct})}, & \text{if } \alpha \neq \frac{1}{3} \text{ and } 1+C\tilde{\rho}h^{-3}(1-e^{Ct}) > 0. \end{cases} \tag{2.2}$$

Remarks. It should be emphasized that when $\alpha \neq \frac{1}{3}$, the bound on the perturbation only applies provided that $\tilde{\rho}$ is suitably small. This phenomenon is typical of non-linear stability analyses (Stetter, 1973) and has no counterpart in the Lax–Richtmyer linear theory (Richtmyer & Morton, 1967). A similar remark applies to the fact of the constant C being dependent on the solution.

Proof. C will denote a positive constant which depends only on U, but is not necessarily the same at each occurrence.

The identity

$$(VW)_{\hat{x},i} = V_i W_{\hat{x},i} + V_{\hat{x},i} W_i + (h/2)(V_{x,i} W_{x,i} - V_{\bar{x},i} W_{\bar{x},i}) \tag{2.3}$$

is required later. We drop the superscripts α, β.

Taking into account that the operator (1.10) is bilinear, we find the equation

$$M\left(\frac{\partial \tilde{U}}{\partial t}\right)_i + K(\tilde{U}, \tilde{U})_i + K(\tilde{U}, U)_i + K(U, \tilde{U})_i + \tilde{U}_{x\bar{x}\hat{x},i} = \tilde{f}_i, \tag{2.4}$$

which is satisfied by $\tilde{U}$. Now multiply (2.4) by $\tilde{U}_i$ and sum to get

$$\frac{1}{2}\frac{d}{dt}(M\tilde{U}, \tilde{U}) + (\tilde{U}, K(\tilde{U}, \tilde{U})) + (\tilde{U}, K(\tilde{U}, U)) + (\tilde{U}, K(U, \tilde{U})) = (\tilde{U}, \tilde{f}), \tag{2.5}$$

since the term involving $\tilde{U}_{x\bar{x}\hat{x}}$ is easily seen to vanish.

For any value of α, the identity (2.3) can be used to yield

$$\begin{aligned}(\tilde{U}, J(\tilde{U}, U)) &= \alpha(\tilde{U}, U\tilde{U}_{\hat{x}}) + \frac{1-\alpha}{2}(\tilde{U}, (U\tilde{U})_{\hat{x}}) \\ &= \alpha(\tilde{U}, U\tilde{U}_{\hat{x}}) + \frac{1-\alpha}{2}(\tilde{U}, U\tilde{U}_{\hat{x}}) + \frac{1-\alpha}{2}(\tilde{U}, U_{\hat{x}}\tilde{U}) + \\ &\quad \frac{h(1-\alpha)}{4}(\tilde{U}, U_x \tilde{U}_x) - \frac{h(1-\alpha)}{4}(\tilde{U}, U_{\bar{x}}\tilde{U}_{\bar{x}}).\end{aligned} \tag{2.6}$$

Periodicity enables us to prove

$$(\tilde{U}, J(\tilde{U}, U)) = -\alpha((U\tilde{U})_{\hat{x}}, \tilde{U}) - \frac{1-\alpha}{2}(\tilde{U}_{\hat{x}}, U\tilde{U}), \tag{2.7}$$

and a new application of (2.3) shows that the right-hand side of (2.7) is

$$-\alpha(U\tilde{U}_{\hat{x}}, \tilde{U}) - \alpha(U_{\hat{x}}\tilde{U}, \tilde{U}) - \frac{\alpha h}{2}(U_x\tilde{U}_x, \tilde{U}) + \frac{\alpha h}{2}(U_{\bar{x}}\tilde{U}_{\bar{x}}, \tilde{U}) - \frac{1-\alpha}{2}(\tilde{U}_{\hat{x}}, U\tilde{U}). \tag{2.8}$$

We take the arithmetic mean of the expressions (2.6) and (2.8) and get

$$(\tilde{U}, J(\tilde{U}, U)) = \frac{1-3\alpha}{4}(U_{\hat{x}}, \tilde{U}^2) + \frac{h(1-3\alpha)}{8}(U_x\tilde{U}_x, \tilde{U}) - \frac{h(1-3\alpha)}{8}(U_{\bar{x}}U_{\bar{x}}, \tilde{U}), \tag{2.9}$$

whence

$$|(\tilde{U}, J(\tilde{U}, U))| \leqslant C\|\tilde{U}\|^2. \tag{2.10}$$

The same procedure would be valid for the operator I and so

$$|(\tilde{U}, K(\tilde{U}, U))| \leqslant C\|\tilde{U}\|^2 \tag{2.11}$$

holds for arbitrary α. Analogously

$$|(\tilde{U}, K(U, \tilde{U}))| \leqslant C\|\tilde{U}\|^2. \tag{2.12}$$

Next

$$\begin{aligned} |(\tilde{U}, K(\tilde{U}, \tilde{U}))| &\leqslant C(\|\tilde{U}\|^2 + \|\tilde{U}\tilde{U}_{\bar{x}}\|^2 + \|\tilde{U}\tilde{U}_x\|^2) \\ &\leqslant C(\|\tilde{U}\|^2 + h^{-3}\|\tilde{U}\|^4), \end{aligned} \tag{2.13}$$

and

$$|(\tilde{U}, \tilde{f})| \leqslant (\|\tilde{U}\|^2 + \|\tilde{f}\|^2)/2. \tag{2.14}$$

When $\alpha = \frac{1}{3}$, the bounds (2.11), (2.13) can be improved, as in this case the identities

$$(\tilde{U}, K(\tilde{U}, \tilde{U})) = (\tilde{U}, K(\tilde{U}, U)) = 0 \tag{2.15}$$

are verified in a straightforward manner.

Set $\varepsilon(\alpha) = 0$ if $\alpha = \frac{1}{3}$, $\varepsilon(\alpha) = 1$ if $\alpha \neq \frac{1}{3}$. Then substitution of (2.11), (2.12), (2.13), (2.14) [or (2.15)] in (2.5) yields

$$\frac{d}{dt}(M\tilde{U}, \tilde{U}) \leqslant C(\|\tilde{U}\|^2 + \varepsilon(\alpha)h^{-3}\|\tilde{U}\|^4 + \|\tilde{f}\|^2). \tag{2.16}$$

Integration and the properties of M give

$$\|\tilde{U}(t)\|^2 \leqslant C(\|\tilde{u}_0\|^2 + \int_0^t \|\tilde{f}(\tau)\|^2\, d\tau) + C\int_0^t (\|\tilde{U}(\tau)\|^2 + \varepsilon(\alpha)h^{-3}\|\tilde{U}(\tau)\|^4)\, d\tau, \tag{2.17}$$

and *a fortiori*

$$\|\tilde{U}(t)\|^2 \leqslant C\tilde{\rho} + C\int_0^t (\|\tilde{U}(\tau)\|^2 + \varepsilon(\alpha)h^{-3}\|\tilde{U}(\tau)\|^4)\, d\tau.$$

The theory of integral inequalities (Hartman, 1964, pp. 24–29) shows that $\|U(t)\|^2$ can be bounded by the solution of the initial value problem

$$\dot{v} = C[v + \varepsilon(\alpha)h^{-3}v^2], \qquad v(0) = C\tilde{\rho}, \tag{2.18}$$

in the interval of existence of v. An elementary integration yields

$$v(t) = \frac{C\tilde{\rho}\, e^{Ct}}{1+\varepsilon(\alpha)h^{-3}C\tilde{\rho}(1-e^{Ct})}, \tag{2.19}$$

which is defined provided that its denominator does not vanish. Now the proof is complete.

As a first application of the theorem we study the propagation, as t increases, of a perturbation in the initial data. With $\tilde{f} \equiv 0$ and $T = \infty$, we obtain, if $\alpha = \frac{1}{3}$

$$\|\tilde{U}(t)\|^2 \leqslant C\|\tilde{u}_0\|^2\, e^{Ct}, \quad t > 0,$$

while if $\alpha \neq \frac{1}{3}$ the bound (2.19) becomes infinite when t is so large that the denominator vanishes. This seems to agree with the results of Fornberg (1973), who proved that when the periodic problem for the equation

$$u_t + uu_x = 0 \tag{2.20}$$

is solved by a method which replaces uu_x by $J^{(\alpha)}(U, U)$, $\alpha \neq \frac{1}{3}$, then it is possible to introduce perturbations $\tilde{u}_0$ of the initial condition $u_0 \equiv 0$, which make the numerical solution to blow-up in finite time. [Note that the proof of Theorem 1 is valid for Equation (2.20).]

Let us turn to the question of convergence. We introduce the order of formal approximation p of L_h to L, by

$$Lu - L_h u = O(h^p), \quad h \to 0, \tag{2.21}$$

then:

THEOREM 2. *Assume that the solution u of* (1.1) *is smooth and that either $\alpha = \frac{1}{3}$ or $\alpha \neq \frac{1}{3}$ and $p > \frac{3}{2}$. Then for each fixed t, $0 < t < \infty$, $U(t)$ converges to $u(t)$, with*

$$\|u(t) - U(t)\| = O(h^p), \quad \textit{as } h \to 0.$$

Proof. Set $g_h = L_h u$. Then, according to (2.21), U solves the perturbed problem

$$L_h U = f = Lu = g_h + O(h^p).$$

For each fixed h, Theorem 1 supplies a bound for the norm of the resulting perturbation $u - U$, in such a way that the corresponding constant C depends only on the restriction of u to the given mesh. Furthermore, inspection of the proof of Theorem 1 reveals that the same value of C will do for all such restrictions, provided that u is smooth. Thus C depends on u but is completely independent of h, and the proof follows easily. [Note that when $\alpha \neq \frac{1}{3}$, the condition $p > \frac{3}{2}$ ensures that for h small $1 + C\tilde{\rho}h^{-3}(1 - e^{Ct}) > 0$, making the application of Theorem 1 possible.]

3. Concluding Remarks

When $\alpha = \frac{1}{3}$, $\beta = 1$, $M_0 = 1$, and (1.12) is discretized in time by means of the mid-point rule (leap-frog), the resulting scheme reduces to that introduced by Zabusky & Kruskal (1965), with $p = 2$. Other choices of β are useful when $\alpha = \frac{1}{3}$. For instance, if $\beta = 2$, then the replacement for uu_x is fourth order accurate.

Sanz-Serna & Christie (1979) have suggested a fourth order accurate method based on the Petrov–Galerkin finite element procedure. Their method is recovered from (1.12) when $\alpha = 0$, $\beta = \frac{5}{6}$, $M_{-2} = M_2 = \frac{1}{120}$, $M_{-1} = M_1 = \frac{13}{60}$, $M_0 = \frac{33}{60}$. It should be emphasized that the introduction of the mass matrix, while allowing a higher order of accuracy, results in an implicit scheme. For the KdV equation, where by stability requirements finite difference schemes have the time stepsize k restricted to be $O(h^3)$, implicitness is not to be particularly feared.

Numerical performances of these schemes are reported by Sanz-Serna & Christie (1979), and show, in agreement with Theorem 2, that methods with higher order of local accuracy lead to lower values of the global error $u - U$. For instance, in the study of the propagation of a single solution, comparable errors are obtained between the Zabusky–Kruskal method with $h = 0{\cdot}01$, $k = 0{\cdot}0005$ and the Sanz-Serna–Christie scheme with $h = 0{\cdot}033$, $k = 0{\cdot}01$. (The trapezoidal rule was employed to advance in time the latter.) Thus the use of fourth order methods seems advantageous. On the other hand our analysis shows that schemes with $\alpha \neq \frac{1}{3}$ could suffer from poor stability properties, although it is fair to say that the present authors have not come across any cases of dramatic growth of the computational errors.

References

ALEXANDER, M. E. & MORRIS, J. LL. 1979 Galerkin methods applied to some modest equations for non-linear dispersive waves. *J. comput. Phys.* **30**, 428–451.

CANOSA, J. & GAZDAG, J. 1977 The Korteweg–de Vries–Burgers equation. *J. comput. Phys.* **23**, 393–403.

FORNBERG, B. 1973 On the instability of leap-frog and Crank–Nicolson approximations of a non-linear partial differential equation. *Maths Comput.* **27**, 45–57.

GREIG, I. & MORRIS, J. LL. 1976 A Hopscotch method for the numerical solution of the Korteweg–de Vries equation. *J. comput. Phys.* **20**, 67–80.

HARTMAN, P. 1964 *Ordinary Differential Equations*. London: John Wiley.

PEN-YU, K. 1965 A class of difference scheme for two-dimensional vorticity equation of viscous fluid. NR, Mathematical department of Shanghai University of Science and Technology: see also 1974 *Acta math. sin.* **17**, 242–258.

PEN-YU, K. 1977 Numerical method of incompressible viscous flow. *Scientia sin.* **20**, 287–304.

PEN-YU, K. 1978 Numerical solution of soliton. *Kexuo Tongba* **23**, 592–597.

PEN-YU, K. 1979*a* Numerical solution of two dimensional vorticity equation. *Acta mech. sin.* **2**, 129–147.

PEN-YU, K. 1979*b* The conservative scheme for atmospheric circulation equations and error estimation of global prediction. *Scientia sin. mathematical collection* (I), 39–52.

PEN-YU, K. 1980*a* Numerical solution of three dimensional vorticity equation. *Mathematica numerica sinica* **2**, 307–318.

PEN-YU, K. 1980*b* Numerical solution of Navier–Stokes equation. *Scientia sin.* (to appear).

PEN-YU, K. 1980*c* Numerical solution of Low Mach number flow. *Proc. Num. Maths Chinese Universities* **2**, 64–77.

PEN-YU, K. 1980*d* Résolution numérique de fluide compressible. *C. r. hebd. Séanc. Acad. Sci., Paris, Série A* **291**, 167–170.

RICHTMYER, R. D. & MORTON, K. W. 1967 *Difference Methods for Initial-value Problems*. New York: Interscience.

SANZ-SERNA, J. M. & CHRISTIE, I. 1979 Petrov–Galerkin methods for non-linear dispersive waves. *University of Dundee Report NA*/34. Also accepted for publication in *Journal of Computational Physics*.

SJÖBERG, A. 1970 On the Korteweg–de Vries equation, existence and uniqueness. *Journal of Mathematical Analysis and Applications* **29,** 569–579.

STETTER, H. J. 1973 *Analysis of Discretization Methods for Ordinary Differential Equations.* Berlin: Springer.

VLIEGENTHART, A. C. 1971 On finite-difference methods for the Korteweg–de Vries equation. *Journal of Engineering Mathematics* **5,** 137–155.

WAHLBIN, L. 1974 A dissipative Galerkin method for the numerical solution of first order hyperbolic equations. In *Mathematical Aspects of Finite Elements* (C. de Boor, Ed.). New York: Academic Press.

ZABUSKY, N. J. & KRUSKAL, M. D. 1965 Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Physics Review Letters* **15,** 240–243.