INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 20, 99–113 (1994)

COMPARISON OF FINITE ELEMENT METHODS FOR THE ST. VENANT EQUATIONS

F. E. HICKS AND P. M. STEFFLER

Department of Civil Engineering, University of Alberta, Edmonton, Alta, Canada T6G 2G7

SUMMARY

Finite element schemes for hyperbolic systems are applied to the St. Venant equations for one-dimensional, unsteady, open channel flow. The comparative performances of the characteristic–dissipative–Galerkin, Taylor–Galerkin and least squares finite element schemes are assessed by means of linear Fourier analysis and solution of idealized non-linear wave propagation problems. Of particular interest is the behaviour of these schemes for the regressive wave component in both subcritical and supercritical flows. To assess the quality of the basic solution, the methods are compared without any additional artificial diffusion or shock-capturing formulations. The balanced treatment of both wave components in the characteristic–dissipative–Galerkin method is illustrated. Also, the method displays little sensitivity to parameter variations. The Taylor–Galerkin scheme provides good solutions, although oscillations due to wave dispersion and minimal diffusion of the regressive wave are displayed. Also, this method is somewhat sensitive to the time step increment. The least squares method is considered unsuitable for unsteady, open channel flow problems owing to its inability to propagate a regressive wave in a supercritical flow.

KEY WORDS St. Venant equations Hyperbolic system Characteristic–dissipative–Galerkin Taylor–Galerkin Least squares finite element

INTRODUCTION

The St. Venant equations governing one-dimensional, unsteady, open channel flow form a hyperbolic system with two distinct disturbance propagation velocities. A number of finite element schemes have been proposed for these equations and general hyperbolic systems. While much of the research in this area is currently directed to the solution of the two-dimensional flow equations, investigation of the one-dimensional approximation remains a valid and important interest. For example, flood forecasting and flood plain delineation, the design of flood protection works, culverts, spillways and diversion canals and the assessment of the impact of dam failure or a sudden ice jam release all require knowledge of the water surface elevation, discharge and velocity in an open channel. In addition, from a numerical perspective, analytical solutions for idealized one-dimensional problems provide the means to compare the performances of various numerical schemes.

Three finite element schemes specifically designed for hyperbolic systems are investigated in this study. The first is the characteristic–dissipative–Galerkin (CDG) method,[1] an SU/PG-based[2] scheme in which upwind weighted test functions are used to introduce selective dissipation based on the characteristic velocities of the propagating disturbances. The second is the Taylor–Galerkin (TG) method,[3] an explicit formulation analogous to the familiar Lax–Wendroff finite difference scheme.[4] The third is the least squares (LS) method,[5,6] in which the residual is minimized in terms of least squares.

These three methods are compared by means of linear Fourier analysis and solution of idealized wave propagation problems. Of particular interest in these tests was the behaviour of the regressive

CCC 0271–2091/95/020099–15

Received 22 July 1993
Revised 13 June 1994

component in both subcritical and supercritical flows. To provide a complete comparison of these finite element schemes, their ability to handle a mild shock propagation problem is also presented. The methods are compared without any additional artificial diffusion or shock-capturing formulations. Such devices are necessary in practice, but it is desirable to minimize their use by improving the quality of the basic solution.[7]

ST. VENANT EQUATIONS

The system of equations that describes one-dimensional, unsteady flow in an open channel represents conservation of mass and momentum. Neglecting lateral (tributary) inflow, for a prismatic, wide rectangular channel they may be written as

$$\frac{\partial\{\phi\}}{\partial t}+\frac{\partial\{F\}}{\partial x}+\{f\}=\{0\}, \tag{1}$$

where

$$\{\phi\}\equiv\begin{Bmatrix}A\\Q\end{Bmatrix},\qquad \{F\}\equiv\begin{Bmatrix}Q\\UQ+(gH/2)A\end{Bmatrix},\qquad \{f\}\equiv\begin{Bmatrix}0\\-gA(S_o-S_f)\end{Bmatrix}, \tag{2}$$

A is the flow area, Q is the discharge, U is the cross-sectionally averaged longitudinal velocity ($U = Q/A$), H is the flow depth ($H = A/B$, where B is the channel width), S_0 is the longitudinal bed slope, S_f is the longitudinal friction slope, g is the acceleration due to gravity and x and t are the longinitudinal distance and temporal co-ordinates respectively. Since this study focuses on the treatment of the dynamic terms, horizontal frictionless flow ($\{f\} = \{0\}$) is considered. This also facilitates the determination of analytical solutions to non-linear problems, providing for a rigorous comparison of the numerical schemes.

A non-conservation form of this system may also be considered, namely

$$\frac{\partial\{\phi\}}{\partial t}+[A]\frac{\partial\{\phi\}}{\partial x}=\{0\}, \tag{3}$$

where

$$[A]\equiv\frac{\partial\{F\}}{\partial\{\phi\}}=\begin{bmatrix}0 & 1\\ c^2-U^2 & 2U\end{bmatrix}, \tag{4}$$

in which $c=\sqrt{(gH)}$ represents the propagation velocity of a small disturbance in still water.

FINITE ELEMENT METHODS

With the exception of the LS scheme, the finite element equations were derived using the Galerkin weighted residual method. The simplest implementation, commonly known as the Bubnov–Galerkin (BG) method, is analogous to centred finite differences. In the BG method the test functions are simply set equal to the basis functions. In open channel flow applications the Bubnov–Galerkin formulation has been shown to be useful for modelling relatively flat waves, but it performs poorly in the vicinity of steep gradients in the solution.[8] Instabilities result and the solution deteriorates rapidly. Nevertheless, the BG method (formulated with a Crank–Nicolson time discretization) is included for comparison.

Implementation

All four of the finite element schemes presented were fundamentally based on the conservation form of the equations. This implementation has two advantages. First, integration by parts over each element generates flux terms which are cancelled out upon assembly and enforcement of conservation of mass and momentum across element boundaries.[1,9] At the upstream and downstream ends of the channel these flux terms represent natural boundary conditions. Second, a conservation formulation involves a linear form of the continuity equation, leaving only one non-linear equation in the system.

Upwinding terms in the CDG, TG and LS schemes were based on the non-conservation, non-symmetric form represented by the system in (3). A Newton–Raphson iteration was used to solve the resulting non-linear systems. The solution was considered to have converged when the magnitude of the correction vector divided by the magnitude of the solution vector was less than 10^{-5}.

Linear elements were used in the formulation of all four schemes. To quantify the flux terms in the momentum equation for the numerical integration, values of A and Q were first evaluated at each integration point by interpolation of the nodal values and then used to calculate the flux terms as defined in (2).

Characteristic–dissipative–Galerkin

The characteristic–dissipative–Galerkin (CDG) scheme[1,9] was derived following Hughes and co-workers.[2,10,11] The significant difference lies in the form of the equations modelled, since Hughes *et al.*[10] formulate the scheme on symmetric (non-conservation) hyperbolic systems. A symmetric system equivalent to (3) can be formed, e.g.

$$\frac{\partial}{\partial t}\begin{Bmatrix} 2c \\ U \end{Bmatrix} + \begin{bmatrix} U & c \\ c & U \end{bmatrix} \frac{\partial}{\partial x}\begin{Bmatrix} 2c \\ U \end{Bmatrix} = \{0\} \tag{5}$$

However, since numerical tests of formulations based on (5) did not reveal any noticeable advantage, the non-symmetric conservation form was retained. The only difference is that the inverse of the eigenvector matrix is required in the eigenvector expansion of the non-symmetric convection matrix.

Adaptation of this concept to the problem of open channel flow is defined by

$$[\mathbf{A}] = [\mathbf{M}][\Lambda][\mathbf{M}]^{-1} = \begin{bmatrix} 1/2c & -1/2c \\ (\mathrm{U}+\mathrm{c})/2c & -(\mathrm{U}-\mathrm{c})/2c \end{bmatrix} \begin{bmatrix} \mathrm{U}+\mathrm{c} & 0 \\ 0 & \mathrm{U}-\mathrm{c} \end{bmatrix} \begin{bmatrix} -(\mathrm{U}-\mathrm{c}) & 1 \\ -(\mathrm{U}+\mathrm{c}) & 1 \end{bmatrix}. \tag{6}$$

Implementation is equivalent to a Bubnov–Galerkin formulation of the extended system

$$\underbrace{\frac{\partial\{\phi\}}{\partial t} + \frac{\partial\{F\}}{\partial x}}_{\textit{original system}} - \underbrace{\omega\frac{\Delta x}{2}[\mathbf{W}]\frac{\partial}{\partial x}\left(\frac{\partial\{\phi\}}{\partial t} + [\mathbf{A}]\frac{\partial\{\phi\}}{\partial x}\right)}_{\textit{upwinding terms}} = \{0\}, \tag{7}$$

where

$$[\mathbf{W}] \equiv \frac{[\mathbf{A}]}{|[\mathbf{A}]|} \equiv \frac{1}{2c}\begin{bmatrix} \dfrac{c^2-U^2}{|U+c|-|U-c|} & \dfrac{U+c}{|U+c|} - \dfrac{U-c}{|U-c|} \\ (c^2-U^2)\left(\dfrac{U+c}{|U+c|} - \dfrac{U-c}{|U-c|}\right) & \dfrac{(U+c)^2}{|U+c|} - \dfrac{(U-c)^2}{|U-c|} \end{bmatrix}, \tag{8}$$

in which ω is an 'upwinding coefficient' while the matrix [W] controls the distribution of the upwinding. It was found that $\omega = 0{\cdot}5$ optimized the phase accuracy in the linear case.[1,9] At $\omega = 0{\cdot}25$ the amplitude accuracy improved slightly while the reduction in phase accuracy was only marginal.[1,9] Because of the limited effect of varying ω in the linear case, a constant value of $\omega = 0{\cdot}5$ was used for this investigation. A Crank–Nicolson time discretization was used for this scheme.

The CDG scheme may also be considered a refinement to the dissipative–Galerkin scheme originally introduced by Katopodes[8] for solving these one-dimensional, open channel flow equations. This refinement accounts for the characteristic velocities of both progressive and regressive disturbances in the determination of the upwinding matrix. These characteristic velocities also have important physical significance to the problem of unsteady, open channel flow.

Taylor–Galerkin

The Taylor–Galerkin (TG) scheme,[3,12–15] also known as the characteristic–Galerkin scheme,[15] is an *explicit* formulation in which the unknown variables are written in terms of a Taylor series expansion in time. The governing equations are then used to devise a time-independent formulation. This exploitation of the time series is used to introduce numerical diffusion.[12] The implementation of the Taylor–Galerkin scheme is equivalent to a Bubnov–Galerkin formulation of the equation

$$\left(\frac{\partial\{\phi\}}{\partial t}+\frac{\partial\{F\}}{\partial x}\right)-\Delta t\frac{\partial}{\partial x}\left([\mathbf{A}][\mathbf{A}]\frac{\partial\{\phi\}}{\partial x}\right)=\{0\}. \tag{9}$$

In the Taylor–Galerkin scheme the size of the diffusion term is controlled by the time step increment. Although this scheme is termed explicit, the *consistent*[3] formulation used here reduces the stability limit of the Taylor–Galerkin method, such that the Courant number C is[12,14–16]

$$C=\frac{|U\pm c|\Delta t}{\Delta x}\leq\frac{1}{\sqrt{3}}\approx 0.577, \tag{10}$$

and requires the solution of a system of equations at each time step. However, unlike any of the other methods considered, this system is linear.

Least squares

The least squares (LS) method[5,6] is an implicit scheme in which the residual is minimized in terms of least squares rather than by the Galerkin method described above. Implementation is equivalent to a Bubnov–Galerkin formulation of the system

$$\left(\frac{\partial\{\phi\}}{\partial t}+\frac{\partial\{F\}}{\partial x}\right)-\Delta t\,\theta\frac{\partial[\mathbf{A}]}{\partial x}\left(\frac{\partial\{\phi\}}{\partial t}+[\mathbf{A}]\frac{\partial\{\phi\}}{\partial x}\right)-\Delta t\,\theta\frac{\partial}{\partial x}\left[[\mathbf{A}]^{T}\left(\frac{\partial\{\phi\}}{\partial t}+[\mathbf{A}]\frac{\partial\{\phi\}}{\partial x}\right)\right]=\{0\}, \tag{11}$$

where θ defines the level of implicitness ($\theta = 0{\cdot}5$, representing a semi-implicit formulation, was actually used for all tests). Here, as in the Taylor–Galerkin method, the amount of numerical diffusion added depends on Δt. Another similarity with the Taylor–Galerkin method is the necessity to discretize in time first, then in space.

FOURIER STABILITY ANALYSIS

A linear stability analysis was used to examine the amplification and phase characteristics of the various numerical methods investigated in this study for the linearized equations. The type of analysis used here may be described as a Fourier (or von Neumann) type.[17] Such an analysis is a valuable tool not only as a basis for the comparison of various numerical schemes but also to determine the

appropriate discretization of a problem for a particular method. Fourier analyses are available for all three numerical schemes,[1,5,14] though not for this system in the case of the LS method.

Equation (8) may be linearized and non-dimensionalized to obtain

$$\frac{\partial\{\phi_*\}}{\partial t_*} + \begin{bmatrix} 0 & 1 \\ 1 - F_o^2 & 2F_o \end{bmatrix} \frac{\partial\{\phi_*\}}{\partial x_*} = \{0\}, \tag{12}$$

where ϕ_*, t_* and x_* are defined by

$$x_* = \frac{x}{L}, \qquad t_* = \frac{tU_o}{L}, \qquad \{\phi_*\} = \begin{Bmatrix} H/H_o \\ UH/(U_oH_o) \end{Bmatrix}. \tag{13}$$

Here L represents a length scale and H_o and U_o are the uniform flow depth and velocity respectively. F_o is the Froude number of the uniform flow, defined by

$$F_o \equiv \frac{U_o}{\sqrt{(gH_o)}}. \tag{14}$$

Each Fourier component of the solution may assume the form[17]

$$\{\phi_*\}^{n+1} = \{\Phi\}^{n+1}\exp(\mathrm{i}k x_*), \tag{15}$$

where $\{\Phi\}^{n+1}$ represents the solution amplitude at time $n + 1$ and

$$\mathrm{i} \equiv \sqrt{(-1)}, \qquad k \equiv \frac{2\pi}{\zeta}, \qquad N \equiv \frac{\zeta}{\Delta x}, \tag{16}$$

in which k is the component wave number, ζ is the wavelength and N is the number of discretization intervals per wavelength. The complex 'amplification matrix' $[G]$, defined such that

$$\{\Phi\}^{n+1} = [G]\{\Phi\}^n, \tag{17}$$

describes both the amplification and phase characteristics of the numerical method. The two complex eigenvalues of this matrix, corresponding to the regressive and progressive waves, can be written in the form

$$re^{\mathrm{i}\Theta}. \tag{18}$$

Here r represents the magnification factor or the 'algorithmic damping' coefficient and Θ indicates the numerical phase shift, i.e.

$$\Theta \equiv k\Delta x_*, \tag{19}$$

where Δx_* is the distance moved by the wave. The 'relative celerity' of the progressive and regressive waves is defined by

$$\text{relative celerity} \equiv \frac{\Theta}{\Theta_{\text{anal.}}} = \Theta \Big/ \left(\mp \frac{2\pi}{N} \frac{(U \pm c)\Delta t_*}{\Delta x_*} \right). \tag{20}$$

Figures 1 and 2 present the results of the linear stability analyses for all four finite element methods for Courant numbers

$$C \equiv \frac{(1 + 1/F_o)\Delta t_*}{\Delta x_*} \tag{21}$$

of 0·25 and 0·5 respectively. The amplitude and phase accuracies are presented for both progressive waves (independent of Froude number) and regressive waves (F_o = 0·5 and 2·0) in each case.

Ideally, it is desirable that a numerical scheme display the ability to provide selective damping of high-frequency disturbances while exhibiting little or no algorithmic damping of longer waves of physical importance. Given that some of the most interesting disturbances generated in open channel

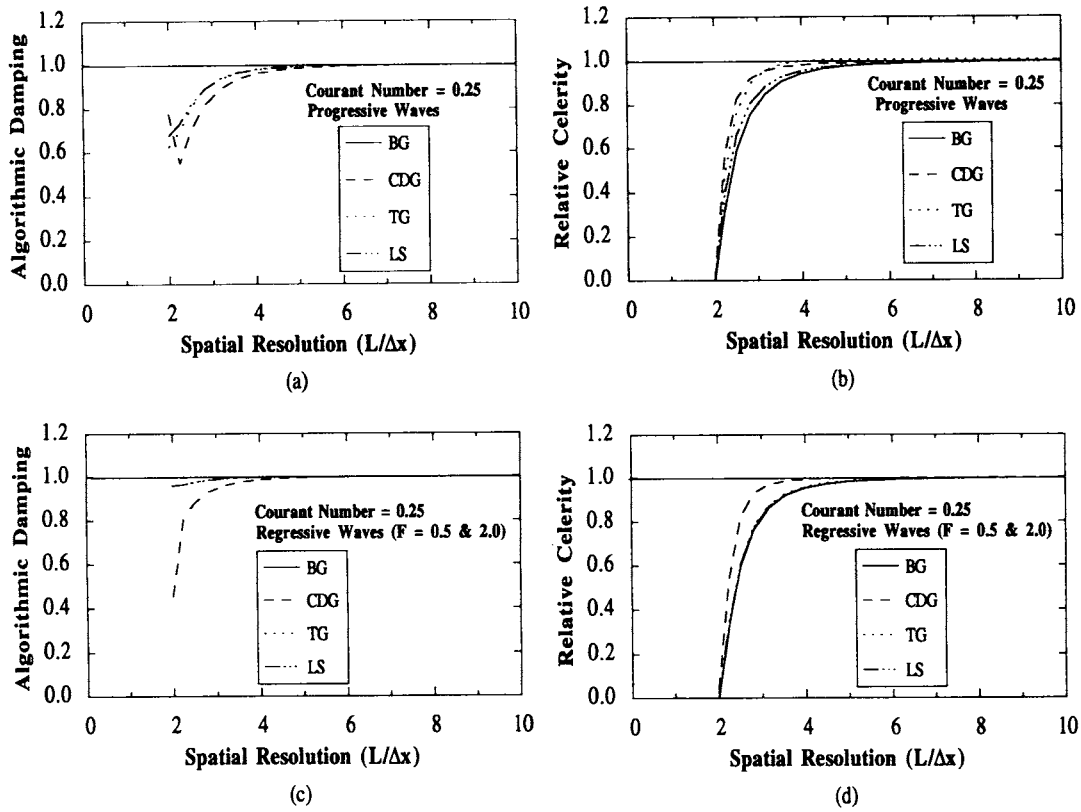

Figure 1. Comparison of amplitude and distance errors of BG, CDG, TG and LS methods for a Courant number of 0·25

flows occur over very short distances (i.e. shocks such as bores and surges), a steep increase in algorithmic damping from a very low value at a spatial resolution of 2·0 to a value of 1·0 at all higher spatial resolutions is extremely desirable. The Bubnov–Galerkin formulation, of course, is entirely non-dissipative (algorithmic damping equal to 1·0 at all values of spatial resolution).

Figure 1 (Courant number 0·25) shows that all three upwinding schemes (CDG, TG and LS) provide selective damping of progressive disturbances. The TG and LS schemes display very little damping of regressive waves at this Courant number (note that the TG and LS plots coincide in this graph). CDG is more dissipative of progressive waves and much more dissipative of regressive waves. The phase accuracy of CDG is also superior to that of the other three methods for regressive disturbances, while TG and CDG are comparable and marginally superior to LS and BG in phase accuracy for progressive waves.

As Figure 2 illustrates, when the Courant number is increased to 0·5, all three upwinding methods illustrate greater damping of progressive disturbances. For regressive waves the TG and LS methods are slightly more dissipative of the shorter wavelengths than when the Courant number is 0·25, while the performance of the CDG scheme is again superior. An examination of the phase accuracy illustrates that the CDG and LS schemes are clearly superior to TG for progressive waves, while TG and CDG are comparable and marginally superior to LS in phase accuracy for regressive waves. A comparison between Figures 1 and 2 illustrates that although TG clearly performs better at $C = 0{\cdot}25$ and LS at $C = 0{\cdot}50$, CDG is relatively insensitive to Courant number, performing well over this range.

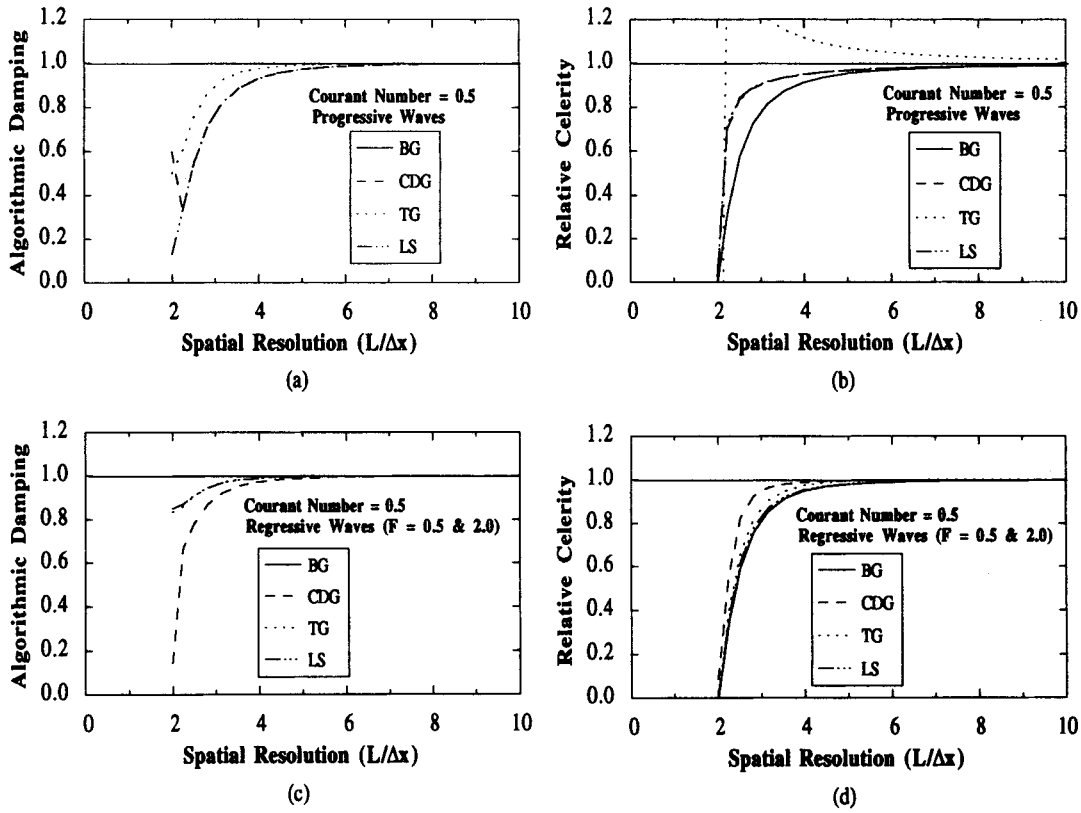

Figure 2. Comparison of amplitude and distance errors of BG, CDG, TG and LS methods for a Courant number of 0·50

NUMERICAL TESTS

Disturbance propagation

To compare these finite element methods for the full non-linear problem, three test problems were examined, each involving the propagation of both a progressive and a regressive disturbance:

(1) with ambient fluid stationary (Froude number 0·0)
(2) a subcritical flow (Froude number 0·5)
(3) a supercritical flow (Froude number 2·0).

In all three cases the geometry consisted of a unit-width section of a horizontal, frictionless channel; 49 linear elements each 10 m long were used, for a total of 50 nodes. The boundary conditions for the subcritical flow runs were given as discharge at the upstream and downstream ends, while for the supercritical flow both discharge and area were specified at the upstream end. To facilitate the computation of exact solutions for these nonlinear problems, the initial conditions were set as two identical disturbances, one progressive and one regressive. The initial depth H (m)

was set to

$$\begin{aligned}
0 \leq x \leq 155, &\quad H = 3{\cdot}5, \\
165 \leq x \leq 215, &\quad H = 3{\cdot}5 + 0{\cdot}02(x - 155), \\
225 \leq x \leq 245, &\quad H = 4{\cdot}7 - 0{\cdot}04(x - 215), \\
245 \leq x \leq 255, &\quad H = 3{\cdot}5, \\
265 \leq x \leq 285, &\quad H = 3{\cdot}5 + 0{\cdot}04(x - 255), \\
285 \leq x \leq 345, &\quad H = 3{\cdot}5 - 0{\cdot}02(x - 285), \\
355 \leq x \leq 495, &\quad H = 3{\cdot}5.
\end{aligned} \tag{22}$$

The initial discharge per unit width, UH ($m^3\ s^{-1}\ m^{-1}$), was specified based upon the Froude number selected for the test, calculated to make each wave propagate entirely in a single direction:

$$\begin{aligned}
(UH)_{\text{progressive}} &= H[\boldsymbol{F}\sqrt{(gH_o)} + 2\sqrt{(gH)} - 2\sqrt{(gH_o)}], \\
(UH)_{\text{regressive}} &= H[\boldsymbol{F}\sqrt{(gH_o)} - 2\sqrt{(gH)} + 2\sqrt{(gH_o)}],
\end{aligned} \tag{23}$$

where H_o is the depth of the undisturbed flow (3·5 m). All four schemes were run at a Courant number of 0·25, based on the propagation velocity of the progressive wave.

Figures 3–5 illustrate the initial conditions (I.C.), exact solution and numerical results for the three tests ($\boldsymbol{F}$ = 0·0, 0·5 and 2·0) at t = 17·34, 12·95 and 7·70 s respectively. For the first two (subcritical) tests the progressive disturbance had travelled exactly 150 m. For the supercritical flow test the progressive disturbance had travelled 156·8 m. Note that in the supercritical test both the regressive and progressive waves moved with the flow.

In the first test ($\boldsymbol{F}$ = 0·0) the BG scheme increased the wave peaks by 1·02%. For the second and third tests ($\boldsymbol{F}$ = 0·5 and 2·0) the BG scheme reduced the progressive peaks by 1·43% and 3·04%

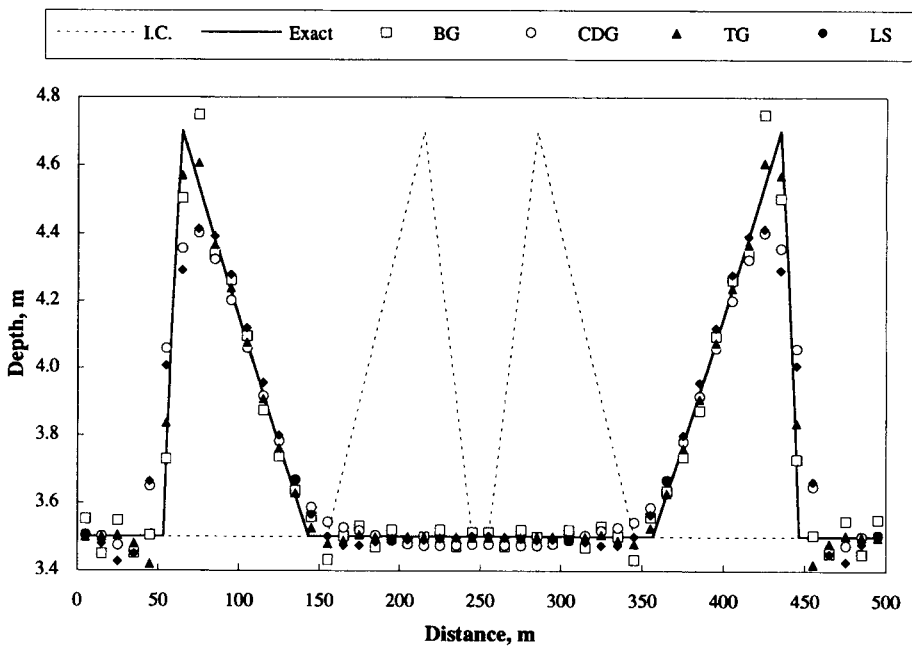

Figure 3. Non-linear propagation of progressive and regressive disturbances ($\boldsymbol{F}$ = 0·0)

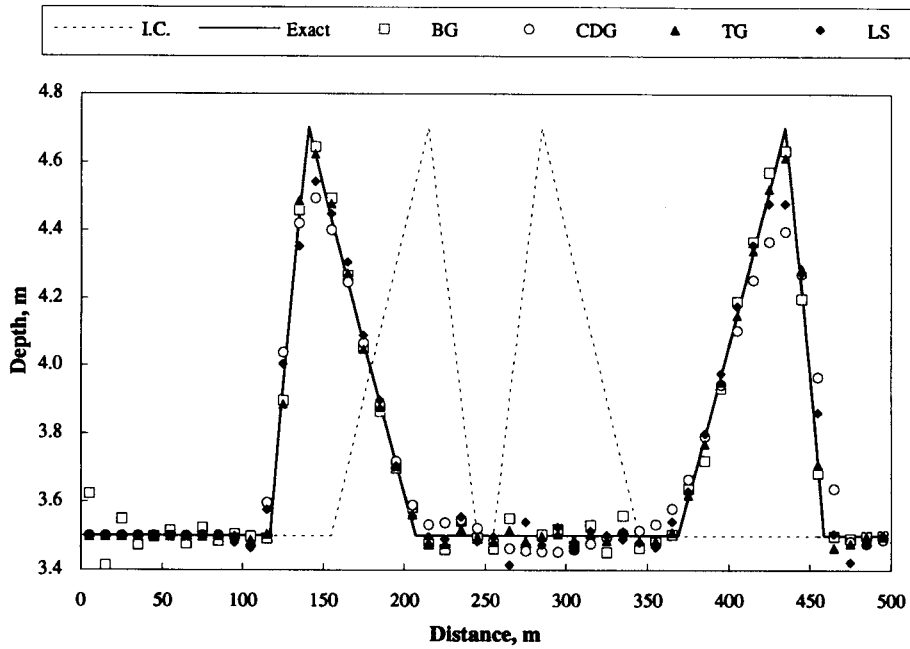

Figure 4. Non-linear propagation of progressive and regressive disturbances ($F = 0{\cdot}5$)

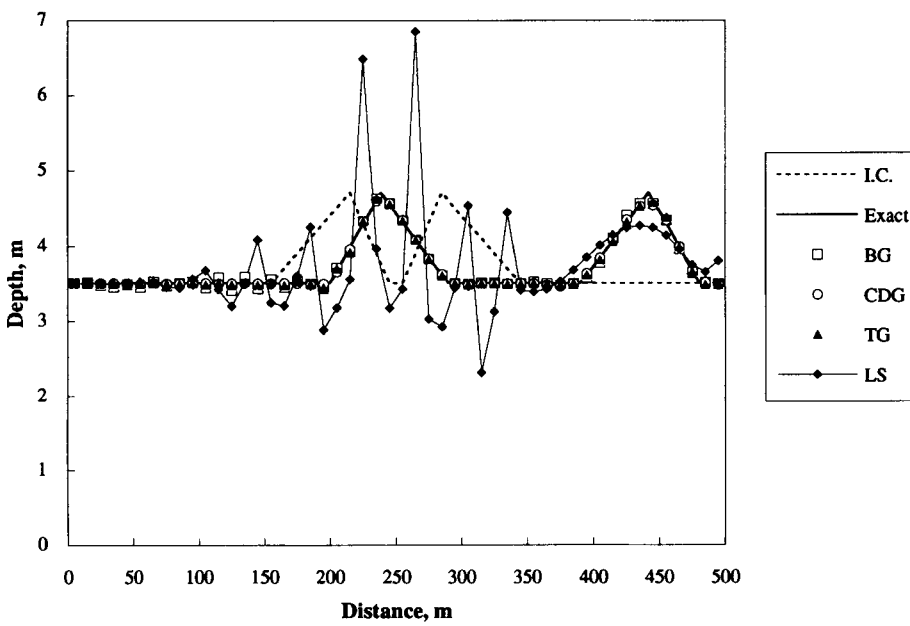

Figure 5. Non-linear propagation of progressive and regressive disturbances ($F = 2{\cdot}0$)

respectively and the regressive peaks were reduced by 1·19% and 1·67% respectively. As predicted by the linear stability analysis, these errors in peak magnitude were not due to algorithmic damping or amplification but rather wave dispersion as high-frequency components were propagated too slowly. This can be seen in the figures, where trailing disturbances were accumulating behind the waves.

The TG scheme produced peaks 1·96% below the exact solution in the $F = 0{\cdot}0$ test. For $F = 0{\cdot}5$ and 2·0 the progressive peaks were underestimated by 1·85% and 2·58% respectively and the regressive peaks by 1·63% and 1·61% respectively. As predicted by the linear stability analysis, the diminishment of the peak was due to both algorithmic damping and wave dispersion as high-frequency components were propagated too slowly. Algorithmic damping would be expected to be small for regressive waves (where the linear stability analysis displayed little diffusion) and this is seen in the figures as well, where the predicted regressive wave peaks were larger than the corresponding progressive wave peaks for $F = 0{\cdot}5$ and 2·0.

The dissipative nature of the CDG scheme was clearly evident in the results for all three tests. With $F = 0{\cdot}0$ the computed peaks were 6·32% below the exact solution. At $F = 0{\cdot}5$ the progressive and regressive wave peaks were 6·42% and 4·37% below the exact solution respectively and at $F = 2{\cdot}0$ they were 3·44% and 2·24% below respectively. The solutions were smoother than those for the other three methods, with very small undershoots. As expected from the linear stability analysis, the diminishment of the peak was mainly due to algorithmic damping. It is interesting to note the improved performance of this method in terms of amplitude accuracy with increasing Froude number, while both BG and TG exhibited the reverse trend.

The results for the LS scheme were between those of TG and CDG. From the linear stability analysis it was seen that the algorithmic damping expected would be very much like that expected for the TG scheme. However, in terms of the phase accuracy for progressive waves the linear stability analysis predicted that LS would be inferior to TG. As noted in the solution, the LS method illustrated both wave diffusion (though not to the extent of CDG) and wave dispersion, as seen in the oscillations (though not to the extent of BG or TG), particularly following the regressive wave. At $F = 0{\cdot}0$ the computed peaks were 6·08% below the exact solution. For $F = 0{\cdot}5$ the progressive and regressive wave peaks were 4·69% and 3·36% below the exact solution respectively. As Figure 5 shows, although this method produced stable results for the progressive wave in the supercritical test, oscillations in the regressive wave destroyed the solution.

The ability of the three formulations—conservation (non-symmetric), non-conservation (non-symmetric) and symmetric (non-conservation)—to conserve mass and momentum was examined using the CDG scheme on these wave propagation problems.[9] The wave propagation tests (for Froude numbers of 0·0, 0·5 and 2·0) were run for 60 time steps at a Courant number of 0·5. These tests were then repeated using only one wave to reveal any possible cancelling of errors. All three formulations were observed to conserve both mass and momentum to at least six significant figures for these tests.

Dam break analysis

Although the propagation of simple disturbances provides valuable insight to the comparative behaviours of these four FEM schemes, practical applications dictate a need for the ability to propagate shocks accurately as well. To examine the comparative abilities of these methods, a classic dam break test was employed. The test problem, reproduced from Reference 18, simulated the instantaneous failure of a dam in a horizontal, frictionless channel. In this case the channel geometry consisted of a unit-width section; 80 linear elements each of length 25 m were used in the analysis, for a total of 81 nodes. The dam itself was approximated over a single element 25 m long. The discharge was initially set to zero at all nodes, while through the upstream half of the domain the initial flow depth was set to

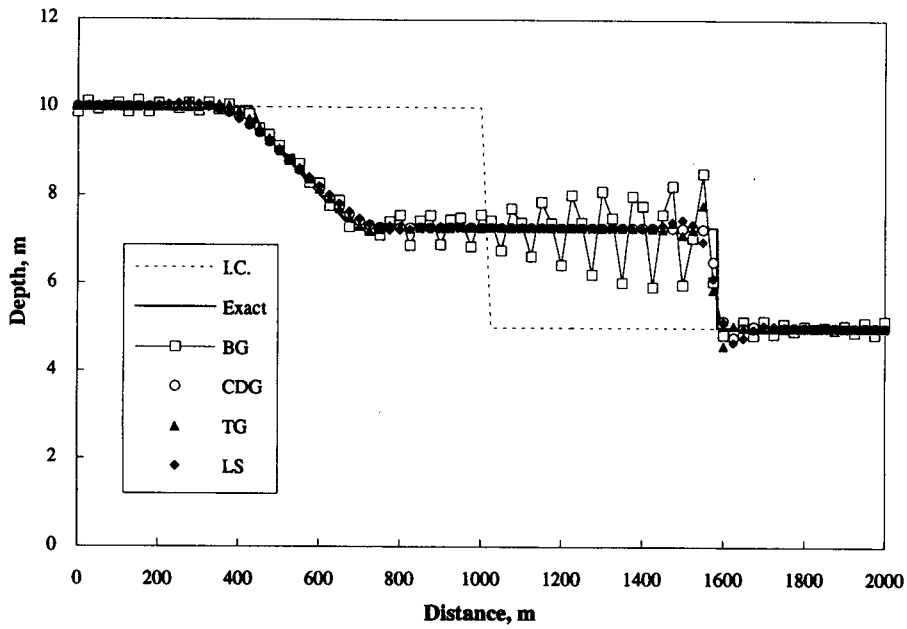

Figure 6. Dam break test

10 m. For the downstream half the flow depth was set to 5 m. The boundary conditions provided at both ends specified a discharge of zero.

Figure 6 illustrates the initial depth profile (I.C.), exact solution and numerical results for this test. All schemes were run at a time step increment of 0·625 s corresponding to a Courant number of 0·23 (based on the propagation velocity of the progressive wave). As in the original test, results are presented after 60 s.

Severe oscillations were observed in the BG solution, particularly trailing the progressive shock. For the TG scheme, although the surge was distributed over just two elements, localized oscillations just upstream and downstream of the jump were observed. With the CDG scheme the shock was distributed over two elements as in the TG scheme, though leading and trailing disturbances were noticeably smaller. Again the performance of the LS scheme was between those of TG and CDG. For the receding wave the BG solution exhibited small oscillations. Of the three upwinding schemes, the TG solution was slightly better than the CDG and LS solutions for this regressive disturbance.

This was a mild shock problem, with the ratio of progressive disturbance height to undisturbed flow depth only 0·45. Additional tests were performed for a more severe (supercritical) dam break problem in which this ratio was 5·20. The CDG method produced results comparable with those for the mild shock problem for $C = 0{\cdot}26$. TG was also successful in obtaining a comparable quality solution, but only when the Courant number was reduced to 0·13. The LS method was also found to be unstable at $C = 0{\cdot}26$.

The ability of the three formulations—conservation (non-symmetric), non-conservation (non-symmetric) and symmetric (non-conservation)—to conserve mass and momentum was examined using the CDG scheme for these dam break problems as well[9] and the results are summarized in Table I. The conservation formulation was found to conserve both mass and momentum to at least six significant figures. The non-conservation formulation was able to conserve mass to six significant figures as well, but not momentum. This was particularly true for the supercritical test, where a momentum error in

Table I. Comparison of conservation properties of the three formulations for the dam break test

Test description	Courant number	Error in mass conservation (%)			Error in momentum conservation (%)		
		Conservation (non-symmetric)	Non-conservation (non-symmetric)	Non-conservation (symmetric)	Conservation (non-symmetric)	Non-conservation (non-symmetric)	Non-conservation (symmetric)
Subcritical dam break	0·47	0·000	0·000	−0·080	0·000	0·006	−1·044
Supercritical dam break	0·26	0·000	0·000	−1·060	0·000	2·277	−9·509

excess of 2% was noted after 96 time steps. The symmetric formulation performed poorly in this test, with errors in both mass and momentum conservation. Again results were worse for the supercritical problem, where errors of the order of 1% and 10% for mass and momentum respectively were observed.

Other numerical tests

These finite element schemes have also been compared in a number of other tests,[9] including steady flow tests involving hydraulic jumps and transitions from mild to steep channels and unsteady tests involving the propagation of waves through hydraulic jumps. The comparative behaviours displayed by these finite element schemes for the numerical tests presented here are typical of their performance on these additional tests.

CONCLUSIONS

In the practical application of any numerical method to unsteady, open channel flow problems, situations requiring the addition of artificial diffusion are going to arise. Therefore, in judging the comparative performances of the characteristic–dissipative–Galerkin (CDG), Taylor–Galerkin (TG) and least squares (LS) finite element schemes, the type and amount of artificial diffusion required will have important implications for the final solution quality.

Because of the balanced treatment of both wave components, the artificial diffusion required for the CDG method will only be small and localized. In some cases the amount of diffusion introduced by the CDG scheme may be more than necessary. This method displayed little sensitivity to variation in the time step increment.

The type and amount of artificial diffusion needed for the TG scheme will be relatively more important to the final solution. Although good solutions can be obtained for relatively small computational cost, this method is somewhat sensitive to the time step increment.

The LS method is considered unsuitable for unsteady, open channel flow problems owing to its inability to propagate a regressive wave in a supercritical flow.

ACKNOWLEDGEMENTS

This research was financially supported through scholarships to the first author from the Natural Sciences and Engineering Research Council of Canada, the Alberta Heritage Scholarship Foundation, the University of Alberta, the North American Life Assurance Co. with the Canadian Council of Professional Engineers, and the Province of Alberta. All are most gratefully acknowledged.

APPENDIX: NOMENCLATURE

A channel cross-sectional area perpendicular to the flow
$\mathbf{A}$ convection matrix
B channel width
c celerity
C Courant number
F flux vector
F_0 uniform flow Froude number
f source term
g acceleration due to gravity
G amplification factor in stability analysis
H flow depth

H_o uniform flow depth
k wave number
L length scale
N number of nodes per wavelength
Q discharge
r algorithmic damping
S_f friction slope
S_o bed slope
t time co-ordinate
t_* non-dimensional time co-ordinate
U cross-sectionally averaged longitudinal velocity
U_o uniform flow velocity
x longitudinal co-ordinate
x_* non-dimensional longitudinal co-ordinate
$\mathbf{W}$ upwinding matrix

Greek letters
ζ disturbance wavelength
θ implicitness
Θ phase shift angle
μ_i characteristic velocities
Λ eigenvalue matrix
ϕ vector of unknowns
ϕ_* non-dimensional solution vector
ω upwinding coefficient

REFERENCES

1. F. E. Hicks and P. M. Steffler, 'A characteristic–dissipative–Galerkin scheme for open channel flow', *ASCE J. Hydraul. Eng.*, **118**, 337–352 (1992).
2. T. J. R. Hughes and M. Mallet, 'A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective–diffusive systems', *Comput. Methods Appl. Mech. Eng.*, **58**, 305–328 (1986).
3. J. Donea, 'A Taylor–Galerkin method for convective transport problems', *Int. j. numer. methods eng.*, **20**, 101–119 (1984).
4. R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial Value Problems*, 2nd edn, Interscience, New York, 1967.
5. G. F. Carey and B. N. Jiang, 'Least-squares finite element for first-order hyperbolic systems', *Int. j. numer. methods eng.*, **26**, 81–93 (1988).
6. B. N. Jiang and G. F. Carey, 'A stable least-squares finite element method for non-linear hyperbolic systems', *Int. j. numer. methods fluids*, **8**, 933–942 (1988).
7. S. T. Zalesak and R. Löhner, 'Minimizing numerical dissipation in modern shock-capturing schemes', in T. J. Chung and R. Karr (eds), *Proc. Int. Symp. on Finite Element Methods in Flow Problems*, University of Alabama at Huntsville Press, Huntsville, AL, 1989, pp. 1205–1211.
8. N. D. Katopodes, 'A dissipative Galerkin scheme for open-channel flow', *ASCE J. Hydraul. Eng.*, **110**, 450–466 (1984).
9. F. E. Hicks and P. M. Steffler, 'Finite element modeling of open channel flow,' *Water Resources Engineering Rep. 90-6*, Civil Engineering Department, University of Alberta, Edmonton, 1990.
10. T. J. R. Hughes, L. P. Franca and M. Mallet, 'A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier–Stokes equations and the second law of thermodynamics', *Comput. Methods Appl. Mech. Eng.*, **54**, 223–234 (1986).
11. T. J. R. Hughes, M. Mallet and A. Mizukami, 'A new finite element formulation for computational fluid dynamics: II. Beyond SUPG', *Comput. Methods Appl. Mech. Eng.*, **54**, 341–355 (1986).
12. V. Selmin, J. Donea and L. Quartapelle, 'Finite element methods for non-linear advection', *Comput. Methods Appl. Mech. Eng.*, **52**, 817–845 (1985).
13. J. Peraire, O. C. Zienkiewicz and K. Morgan, 'Shallow water problems: a general explicit formulation', *Int. j. numer. methods eng.*, **22**, 547–574 (1986).
14. N. D. Katopodes and C. T. Wu, 'Explicit computation of discontinuous channel flow', *ASCE J. Hydraul. Eng.*, **112**, 456–475 (1986).
15. A. J. Baker and J. W. Kim, 'A Taylor weak-statement algorithm for hyperbolic conservation laws', *Int. j. numer. methods fluids*, **7**, 489–520 (1987).

16. J. H. W. Lee, J. Peraire and O. C. Zienkiewicz, 'The characteristic–Galerkin method for advection-dominated problems—an assessment', *Comput. Methods Appl. Mech. Eng.*, **61**, 359–369 (1987).
17. P. J. Roache, *Computational Fluid Dynamics*, Hermosa, Albuquerque, NM, 1972.
18. R. J. Fennema and M. H. Chaudhry, 'Simulation of one-dimensional dam-break flows', *Int. J. Hydraul. Res.*, **25**, 41–51 (1987).