INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS VOL. 21, 273–294 (1995)

ON TAYLOR WEAK STATEMENT FINITE ELEMENT METHODS FOR COMPUTATIONAL FLUID DYNAMICS

D. J. CHAFFIN AND A. J. BAKER*

Department of Engineering Science and Mechanics, 310 Perkins Hall, University of Tennessee, Knoxville, TN 37996-2030, U.S.A.

SUMMARY

A Taylor series augmentation of a weak statement (a 'Taylor weak statement' or 'Taylor–Galerkin' method) is used to systematically reduce the dispersion error in a finite element approximation of the one-dimensional transient advection equation. A frequency analysis is applied to determine the phase velocity of semi-implicit linear, quadratic and cubic basis one-dimensional finite element methods and of several comparative finite difference/finite volume algorithms. The finite element methods analysed include both Galerkin and Taylor weak statements. The frequency analysis is used to obtain an improved linear basis Taylor weak statement finite element algorithm. Solutions are reported for verification problems in one and two dimensions and are compared with finite volume solutions. The improved finite element algorithms have sufficient phase accuracy to achieve highly accurate linear transient solutions with little or no artificial diffusion.

KEY WORDS: finite element; finite volume; Taylor weak statement; Taylor–Galerkin method; phase velocity

1. INTRODUCTION

Application of the Galerkin finite element method (FEM) to parabolic and hyperbolic differential equations has presented difficulties with the control of dispersion error. Dispersion error results from shorter-wavelength solution components travelling at the wrong speed, usually too slowly. Waves travelling at the wrong speed eventually appear in the wrong place as extraneous short waves and can lead to instability in a non-linear problem statement.

The 'upwind' finite volume method has been extensively applied either to reduce dispersion error or to artificially diffuse the resulting short waves. The interpolation is biased for greater contribution from the direction of the velocity. It originated with the donor cell method of Courant *et al.*[1] in 1952 and was applied to the FEM in the 1970s as the Petrov–Galerkin methods of Christie *et al.*[2,3] and Heinrich *et al.*[4,5] Early Petrov–Galerkin methods suffered from excess diffusion of the solution and later work has been oriented towards reducing the excess diffusion, including the SUPG method of Brooks and Hughes[6] in 1980 and the methods of Dick[7] in 1983, Westerink and Shea[8] in 1989, Bouloutas and Celia[9] in 1991 and Konda *et al.*[10] in 1992. Similar work in the finite volume method is typified by the QUICK methods of Leonard[11] in 1979 and Leonard and Mokhtari[12] in 1992.

Many upwind methods still cause excess diffusion and some also cause an undesirable increase in the width of the matrix stencil, thus increasing the computational effort. An alternative approach which avoids these problems originated for the finite volume method with Lax and Wendroff[13] in 1960. The

* Author to whom correspondence should be addressed.

CCC 0271–2091/95/040273–22

Received 18 March 1994
Revised 1 November 1994

Lax–Wendroff method uses the governing equation in a semi-discrete method to cancel error terms in time and space. It was applied to a finite volume method by Stone and Brian[14] in 1963 and Harten and Tal-Ezer[15] in 1981. Similar corrections were made to a finite element method by van Genuchten and Gray[16] in 1978 and became generally known as the Taylor–Galerkin method of Donea[17] in 1984. The method was extended as the Taylor weak statement of Baker and Kim[18] in 1987.

The frequency analysis of von Neumann and Richtmeyer[19] in 1950 was originally derived to determine stability and was extended by Stone and Brian[14] to determine the phase velocity and dispersion error for a finite difference method. The frequency analysis was applied to various forms of the FEM by Vichnevetsky and De Schutter[20] in 1975, Gray and Pinder[21] in 1976, Raymond and Garder[22] in 1976, Gresho *et al.*[23] in 1978, Chin *et al.*[24] in 1979, Wait and Mitchell[25] and Cathers and O'Connor[26] in 1985, Baker and Kim[18] and Gresho and Lee[27] in 1987 and Donea *et al.*[28] in 1987. Vichnevetsky and De Schutter,[20] Raymond and Garder[22] and Baker and Kim[18] each attempted some optimization of parameters to reduce dispersion error in linear basis finite elements. Vichnevetsky and De Schutter,[20] Gray and Pinder,[21] Gresho *et al.*,[23] Cathers and O'Connor,[26] Wait and Mitchell[25] and Gresho and Lee[27] each applied various analyses to quadratic basis finite elements. Vichnevetsky and De Schutter[20] and Gresho *et al.*[23] analysed Hermite cubic basis finite elements.

In this paper the frequency anlaysis is extended to fully discrete Taylor quadratic basis and Galerkin Lagrange cubic basis finite elements in one dimension and is compared with known results for several finite difference and linear basis finite element methods. In addition, a multiple-step Taylor linear basis finite element method is optimized for phase accuracy. The resulting algorithms are tested on verification problems in one and two dimensions.

2. TAYLOR WEAK STATEMENT

Considering approximate solutions to the one-dimensional unsteady linear advection–diffusion equation,

$$L(q) = \frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} - \varepsilon\frac{\partial^2 q}{\partial x^2} = 0. \tag{1}$$

Assuming sufficient continuity and semi-discretizing in time, two forward Taylor series are used, the first for a full time step of Δ_t and the second for a partial time step of $(1-\alpha)\Delta_t$:

$$q^{n+1} = q^n + \Delta_t\frac{\partial q^n}{\partial t} + \frac{\Delta_t^2}{2}\frac{\partial^2 q^n}{\partial t^2} + \frac{\Delta_t^3}{6}\frac{\partial^3 q^n}{\partial t^3} + \frac{\Delta_t^4}{24}\frac{\partial^4 q^n}{\partial t^4} + \mathcal{O}(\Delta_t^5), \tag{2}$$

$$q^{n+1-\alpha} = q^n + (1-\alpha)\Delta_t\frac{\partial q^n}{\partial t} + \frac{(1-\alpha)^2\Delta_t^2}{2}\frac{\partial^2 q^n}{\partial t^2} + \frac{(1-\alpha)^3\Delta_t^3}{6}\frac{\partial^3 q^n}{\partial t^3} + \frac{(1-\alpha)^4\Delta_t^4}{24}\frac{\partial^4 q^n}{\partial t^4} + \mathcal{O}(\Delta_t^5). \tag{3}$$

The corresponding backward series are also used:

$$q^n = q^{n+1} - \Delta_t\frac{\partial q^{n+1}}{\partial t} + \frac{\Delta_t^2}{2}\frac{\partial^2 q^{n+1}}{\partial t^2} - \frac{\Delta_t^3}{6}\frac{\partial^3 q^{n+1}}{\partial t^3} + \frac{\Delta_t^4}{24}\frac{\partial^4 q^{n+1}}{\partial t^4} - \mathcal{O}(\Delta_t^5), \tag{4}$$

$$q^{n+1-\alpha} = q^{n+1} - \alpha\Delta_t\frac{\partial q^{n+1}}{\partial t} + \frac{\alpha^2\Delta_t^2}{2}\frac{\partial^2 q^{n+1}}{\partial t^2} - \frac{\alpha^3\Delta_t^3}{6}\frac{\partial^3 q^{n+1}}{\partial t^3} + \frac{\alpha^4\Delta_t^4}{24}\frac{\partial^4 q^{n+1}}{\partial t^4} - \mathcal{O}(\Delta_t^5). \tag{5}$$

Eliminating third-order terms between the four series and substituting $\theta = (\alpha + 1)/3$ for $\frac{1}{3} < \theta < \frac{2}{3}$,

$$q^{n+1} - q^n - \Delta_t\left(\theta \frac{\partial q^{n+1}}{\partial t} + (1-\theta)\frac{\partial q^n}{\partial t}\right) = -\frac{\Delta_t^2}{6}\left((3\theta - 1)\frac{\partial^2 q^{n+1}}{\partial t^2} + (3\theta - 2)\frac{\partial^2 q^n}{\partial t^2}\right) - \frac{\Delta_t^4}{72}\left(-(3\theta - 1)^2 \frac{\partial^4 q^{n+1}}{\partial t^4} + (3\theta - 2)^2 \frac{\partial^4 q^n}{\partial t^4}\right) + \mathcal{O}(\Delta_t^5). \tag{6}$$

The left-hand side of (6) is the standard θ-implicit method approximation to (1) and the right-hand side is a correction resulting from the replacement of lower-order truncation errors. With $\theta = \frac{1}{3}, \frac{1}{2}$, or $\frac{2}{3}$ the resulting correction is equivalent to various Padé approximations. Assuming sufficient continuity, the time derivatives in (6) may be replaced with space derivatives using (1) in the manner of Lax and Wendroff:[13]

$$\frac{\partial q}{\partial t} = \left(-u\frac{\partial}{\partial x} + \varepsilon \frac{\partial^2}{\partial x^2}\right) q, \tag{7}$$

$$\frac{\partial^2 q}{\partial t^2} = \left(-u\frac{\partial}{\partial x} + \varepsilon \frac{\partial^2}{\partial x^2}\right)\left(-u\frac{\partial}{\partial x} + \varepsilon \frac{\partial^2}{\partial x^2}\right) q = \left(u^2 - 2u\varepsilon \frac{\partial}{\partial x} + \varepsilon^2 \frac{\partial^2}{\partial x^2}\right)\frac{\partial^2 q}{\partial x^2}. \tag{8}$$

Additional terms would be required in (7) and (8) for the non-linear case of $u = q$, not considered here. Substituting in (6) and neglecting terms of $\mathcal{O}(\varepsilon^2, \Delta_t^4)$,

$$q^{n+1} - q^n + u\Delta_t\left(\theta \frac{\partial q^{n+1}}{\partial x} + (1-\theta)\frac{\partial q^n}{\partial x}\right) = -\frac{u^2\Delta_t^2}{6}\left((3\theta - 1)\frac{\partial^2 q^{n+1}}{\partial x^2} + (3\theta - 2)\frac{\partial^2 q^n}{\partial x^2}\right) + \frac{u\varepsilon\Delta_t^2}{3}\left((3\theta - 1)\frac{\partial^3 q^{n+1}}{\partial x^3} + (3\theta - 2)\frac{\partial^3 q^n}{\partial x^3}\right) + \mathcal{O}(\Delta_t^4). \tag{9}$$

For $\theta = \frac{1}{2}$

$$q^{n+1} - q^n + u\frac{\Delta_t}{2}\left(\frac{\partial q^{n+1}}{\partial x} + \frac{\partial q^n}{\partial x}\right) = -\frac{u^2\Delta_t^2}{12}\left(\frac{\partial^2 q^{n+1}}{\partial x^2} - \frac{\partial^2 q^n}{\partial x^2}\right) + \frac{u\varepsilon\Delta_t^2}{6}\left(\frac{\partial^3 q^{n+1}}{\partial x^3} - \frac{\partial^3 q^n}{\partial x^3}\right) + \mathcal{O}(\Delta_t^4). \tag{10}$$

In this paper the trapezoidal rule of $\theta = \frac{1}{2}$ is used to define a non-diffusive algorithm concentrating only on reduction of dispersion error. An artitificial diffusion mechanism may be necessary for a non-linear problem statement.

Assuming the existence of appropriate initial and boundary conditions and writing (10) as a weak statement on region Ω with test function ϕ_i,

$$\int_\Omega \phi_i \left\{(q^{n+1} - q^n) + \frac{\Delta_t}{2}\left[\left(u - \varepsilon\frac{\partial}{\partial x}\right)\frac{\partial}{\partial x}(q^{n+1} + q^n)\right] + \frac{\Delta_t^2}{12}\left[\left(u^2 - 2u\varepsilon\frac{\partial}{\partial x}\right)\frac{\partial^2}{\partial x^2}(q^{n+1} - q^n)\right]\right\} d\Omega = 0. \tag{11}$$

The Galerkin method uses the trial function ϕ_j, identical with ϕ_i where $i = j$, to support a semi-discrete

approximation q^h over J nodes:

$$q^h(x,t) = \sum_{j=1}^{J} \phi_j(x) Q_j(t) \approx q(x,t). \tag{12}$$

Fully discrete approximations are defined at adjacent time levels as

$$Q_j^n = Q_j(t), \qquad Q_j^{n+1} = Q_j(t + \Delta_t). \tag{13}$$

Substituting (12) and (13) in (11), assuming ϕ is once-differentiable, using Green's theorem and neglecting boundary integrals,

$$\sum_{j=1}^{J} \int_\Omega \left[\left(\phi_i \phi_j - \frac{u^2 \Delta_t^2}{12} \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} + \frac{u \varepsilon \Delta_t^2}{6} \frac{\partial \phi_i}{\partial x} \frac{\partial^2 \phi_j}{\partial x^2} \right) (Q_j^{n+1} - Q_j^n) \right.$$
$$\left. + \left(\frac{u \Delta_t}{2} \phi_i \frac{\partial \phi_j}{\partial x} + \frac{\varepsilon \Delta_t}{2} \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} \right) (Q_j^{n+1} + Q_j^n) \right] d\Omega = 0 \quad (i = 1, \ldots, J). \tag{14}$$

Defining global integrals as follows allows a compact matrix notation shown at time level n:

$$\begin{aligned} \mathbf{M}\mathcal{Q}^n &\equiv \sum_{j=1}^{J} \int_\Omega \phi_i \phi_j \, d\Omega \, Q_j^n, & \mathbf{U}\mathcal{Q}^n &\equiv \sum_{j=1}^{J} \Delta_x \int_\Omega \phi_i \frac{\partial \phi_j}{\partial x} d\Omega \, Q_j^n \\ \mathbf{K}\mathcal{Q}^n &\equiv \sum_{j=1}^{J} \Delta_x^2 \int_\Omega \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} d\Omega \, Q_j^n, & \mathbf{E}\mathcal{Q}^n &\equiv \sum_{j=1}^{J} \Delta_x^3 \int_\Omega \frac{\partial \phi_i}{\partial x} \frac{\partial^2 \phi_j}{\partial x^2} d\Omega \, Q_j^n. \end{aligned} \tag{15}$$

Defining the Courant number $C \equiv u\Delta_t/\Delta_x$ and a dimensionless diffusion coefficient $\varepsilon^* \equiv \varepsilon/(u^2\Delta_t)$ and substituting in (14) yields a matrix statement

$$\left(\mathbf{M} - \frac{C^2}{12}\mathbf{K} + \frac{\varepsilon^* C^3}{6}\mathbf{E} \right) (\mathcal{Q}^{n+1} - \mathcal{Q}^n) + \left(\frac{C}{2}\mathbf{U} + \frac{\varepsilon^* C^2}{2}\mathbf{K} \right) (\mathcal{Q}^{n+1} + \mathcal{Q}^n) = \mathbf{0}. \tag{16}$$

Separating $\mathcal{Q}^{n+1}$ and $\mathcal{Q}^n$ yields

$$\left(\mathbf{M} + \frac{C}{2}\mathbf{U} - (1 - 6\varepsilon^*)\frac{C^2}{12}\mathbf{K} + \frac{\varepsilon^* C^3}{6}\mathbf{E} \right) \mathcal{Q}^{n+1} = \left(\mathbf{M} - \frac{C}{2}\mathbf{U} - (1 + 6\varepsilon^*)\frac{C^2}{12}\mathbf{K} + \frac{\varepsilon^* C^3}{6}\mathbf{E} \right) \mathcal{Q}^n. \tag{17}$$

The global matrix integrals **M**, **U**, **K**, and **E** are normally evaluated using a finite element method of the desired basis degree. In the FEM the test functions ϕ are supported (non-zero) only locally on each element and the global matrix integrals are assembled from the local evaluations. Using a linear basis finite element method, (17) is equivalent, when $\varepsilon^* = 0$, to the fourth-order linear basis FEM of van Genuchten and Gray[16] and the Crank–Nicolson Taylor–Galerkin method of Donea.[17] Similar forms may also be obtained from the non-diffusive Petrov–Galerkin methods of Dick[7] and Bouloutas and Celia,[9] from the finite difference method of Harten and Tal-Ezer[15] or from the Taylor weak statement (TWS) finite element method of Baker and Kim.[18] Rearranging Baker and Kim's equation (30),

$$[\mathbf{M} + (\theta - \alpha)C\mathbf{U} + (\gamma + \beta\theta)C^2\mathbf{K} + \mu\theta C^3\mathbf{E}]\mathcal{Q}^{n+1}$$
$$= \{\mathbf{M} - (1 - \theta + \alpha)C\mathbf{U} + [\gamma - \beta(1 - \theta)]C^2\mathbf{K} - \mu(1 - \theta)C^3\mathbf{E}\}\mathcal{Q}^n. \tag{18}$$

Setting $\theta = \frac{1}{2}$ and comparing with (17) defines the constants in the TWS form:

$$\alpha = 0, \qquad \beta = \varepsilon^*, \qquad \gamma = -\tfrac{1}{12}, \qquad \mu = \pm\varepsilon^*/3. \tag{19}$$

Baker and Kim defined the physical diffusion $\varepsilon^* = 0$ and defined the constant β to represent a variable artificial diffusion coefficient, used here to be added to the physical diffusion ε^*. Baker and Kim's second-order dispersion correction γ is also used here, but the third-order term containing μ and $\mathbf{E}$ will be dropped, since the focus is on the case of zero or very small physical diffusion ε^* and since retaining the third-derivative approximation produces an undesirable increase in matrix bandwidth for the linear basis finite element method.

Substituting the redefined constants in (17) with $\theta = \frac{1}{2}$ yields a modified form of the Taylor weak statement:

$$\left[\mathbf{M} + \frac{C}{2}\mathbf{U} + \left(\gamma + \frac{\beta + \varepsilon^*}{2}\right)C^2\mathbf{K}\right]\boldsymbol{Q}^{n+1} = \left[\mathbf{M} - \frac{C}{2}\mathbf{U} + \left(\gamma - \frac{\beta + \varepsilon^*}{2}\right)C^2\mathbf{K}\right]\boldsymbol{Q}^{n}. \tag{20}$$

3. FREQUENCY ANALYSIS

The familiar 'von Neumann' frequency analysis is readily applied to a compact discrete approximation of a one-dimensional differential equation. Although the method was originally used to determine stability, the concentration here is on phase accuracy. Physical and artificial diffusion will be neglected in this section ($\varepsilon = \varepsilon^* = \beta = 0$), since small diffusion terms have little effect on phase velocity. Considering the kth single mode of the Fourier series expansion of a spatially semi-discrete approximation, designed q_k^{h}, at time t and location $x = j\Delta_x$,

$$q_k^{\mathrm{h}}(j\Delta_x, t) = \exp\{i\omega_k[j\Delta_x - U_k(\omega_k)t]\}. \tag{21}$$

Only a single generic Fourier mode need be considered, since (1) and its approximations are linear and boundary conditions are assumed periodic. Omitting the subscript k, the mode q^{h} travels at a phase velocity U which is generally a complex function of ω. Each mode of a similar expansion of the analytical solution, q, has a real constant velocity u. The phase velocity of the semi-discrete approximation is $U(\omega)$ and a relative phase velocity, ideally equal to unity, is defined as

$$\Phi^{\mathrm{h}} = \frac{U(\omega)}{u}. \tag{22}$$

Some authors use phase error $(U(\omega) - u)$ or relative phase error $(\Phi^{\mathrm{h}} - 1)$, both ideally zero.

The amplification factor G^{h} of the fully discrete approximation is defined as its ratio at adjacent time levels t and $t + \Delta_t$:

$$G^{\mathrm{h}} = \frac{q^{\mathrm{h}}(j\Delta_x, t + \Delta_t)}{q^{\mathrm{h}}(j\Delta_x, t)} = \frac{Q_j^{n+1}}{Q_j^n} = \exp[-i\omega U(\omega)\Delta_t]. \tag{23}$$

Using a tridiagonal (linear basis) method with constant nodal spacing as an example, the approximate solution at an interior node is written as a recursion relation between times $n\Delta_t$ and $(n+1)\Delta_t$ and nodes $(j \pm 1)\Delta_x$:

$$a_{j-1}Q_{j-1}^{n+1} + a_jQ_j^{n+1} + a_{j+1}Q_{j+1}^{n+1} = b_{j-1}Q_{j-1}^{n} + b_jQ_j^{n} + b_{j+1}Q_{j+1}^{n}. \tag{24}$$

Substituting a dimensionless wave number m for $\omega\Delta_x$, assuming U and Δ_x constant over the stencil and substituting (21) at the same time level yields

$$\begin{aligned} Q_{j-1} &= Q(x - \Delta_x) = Q_j \exp(-i\omega\Delta_x) = Q_j\, \mathrm{e}^{-im}, \\ Q_{j+1} &= Q(x + \Delta_x) = Q_j \exp(+i\omega\Delta_x) = Q_j\, \mathrm{e}^{+im}. \end{aligned} \tag{25}$$

Substituting (22), (24) and (25) into (23) yields a computable form for the amplification factor in terms

of constants, the wave number m and the relative phase velocity:

$$G^{\mathrm{h}} = \frac{b_{j-1}\,\mathrm{e}^{-\mathrm{i}m} + b_j + b_{j+1}\,\mathrm{e}^{+\mathrm{i}m}}{a_{j-1}\,\mathrm{e}^{-\mathrm{i}m} + a_j + a_{j+1}\,\mathrm{e}^{+\mathrm{i}m}} = \exp\left(-\frac{\mathrm{i}mU(m)\Delta_t}{\Delta_x}\right) = \exp(-\mathrm{i}mC\Phi^{\mathrm{h}}). \tag{26}$$

The amplification factor is a complex-value scalar and stability requires that its modulus $|G^{\mathrm{h}}| \leq 1$. The relative phase velocity is calculated from (26) using (22):

$$\Phi^{\mathrm{h}} = \frac{\tan^{-1}[\mathrm{Im}(G^{\mathrm{h}})/\mathrm{Re}(G^{\mathrm{h}})]}{-mC}. \tag{27}$$

The amplification factor and phase velocity were calculated as functions of m and C using the symbolic computer language Macsyma.[29]

The procedure differs slightly for quadratic and cubic basis finite elements, which produce varying matrix stencils for element vertex nodes and element interior nodes. Following Wait and Mitchell,[25] the equivalent finite difference stencil is arranged as an $n \times n$ matrix for n-degree basis FEM. The n eigenvalues of the matrix become the values of G^{h}. One eigenvalue is the principal value, generally the one having positive relative phase velocity.

4. CALCULATION OF FREQUENCY ANALYSIS

4.1. One-dimensional linear basis finite elements

The well-known frequency analysis for the linear basis (superscript 'l') FEM is presented here for comparison. With element integrals assembled on a uniform grid to an equivalent finite difference stencil and using (25) to eliminate $Q_{j\pm1}$, the global matrix integrals from (15) become

$$\mathbf{M}^{\mathrm{l}}\boldsymbol{Q} = S_e(\mathbf{M}_e^{\mathrm{l}})\boldsymbol{Q} = S_e\left(\frac{\Delta_x}{6}\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}\right)\boldsymbol{Q} = \frac{\Delta_x}{6}[1 \quad 4 \quad 1]\begin{Bmatrix}\mathrm{e}^{-\mathrm{i}m}\\ 1\\ \mathrm{e}^{+\mathrm{i}m}\end{Bmatrix}Q_j, \tag{28}$$

$$\mathbf{U}^{\mathrm{l}}\boldsymbol{Q} = S_e(\mathbf{U}_e^{\mathrm{l}})\boldsymbol{Q} = S_e\left(\frac{\Delta_x}{2}\begin{bmatrix}-1 & 1\\ -1 & 1\end{bmatrix}\right)\boldsymbol{Q} = \frac{\Delta_x}{6}[-1 \quad 0 \quad 1]\begin{Bmatrix}\mathrm{e}^{-\mathrm{i}m}\\ 1\\ \mathrm{e}^{+\mathrm{i}m}\end{Bmatrix}Q_j, \tag{29}$$

$$\mathbf{K}^{\mathrm{l}}\boldsymbol{Q} = S_e(\mathbf{K}_e^{\mathrm{l}})\boldsymbol{Q} = S_e\left(\frac{\Delta_x^2}{\Delta_e}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\right)\boldsymbol{Q} = \Delta_x[-1 \quad 2 \quad -1]\begin{Bmatrix}\mathrm{e}^{-\mathrm{i}m}\\ 1\\ \mathrm{e}^{+\mathrm{i}m}\end{Bmatrix}Q_j, \tag{30}$$

The assembled matrices $\mathbf{M}^{\mathrm{l}}, \mathbf{U}^{\mathrm{l}}$ and $\mathbf{K}^{\mathrm{l}}$ are scaled finite element equivalents of the finite difference operators $I + \delta^2/6, \delta$ and $-\delta^2$ respectively. Substituting (28)–(30) into (20) and (26) with $\varepsilon^* = 0$ (zero physical diffusion) and $\theta = \frac{1}{2}$ yields the amplification factor

$$G^{\mathrm{h}} = \frac{\left[\frac{1}{6} + \frac{C}{4} - \left(\gamma - \frac{\beta}{2}\right)C^2\right]\mathrm{e}^{-\mathrm{i}m} + \left[\frac{4}{6} + 2\left(\gamma - \frac{\beta}{2}\right)C^2\right] + \left[\frac{1}{6} - \frac{C}{4} - \left(\gamma - \frac{\beta}{2}\right)C^2\right]\mathrm{e}^{+\mathrm{i}m}}{\left[\frac{1}{6} - \frac{C}{4} - \left(\gamma + \frac{\beta}{2}\right)C^2\right]\mathrm{e}^{-\mathrm{i}m} + \left[\frac{4}{6} + 2\left(\gamma + \frac{\beta}{2}\right)C^2\right] + \left[\frac{1}{6} + \frac{C}{4} - \left(\gamma + \frac{\beta}{2}\right)C^2\right]\mathrm{e}^{+\mathrm{i}m}}. \tag{31}$$

With $\beta = 0$ the phase velocity is

$$\Phi^{\mathrm{h}} = -\frac{2}{mC}\tan^{-1}\left(\frac{3C\,\sin(m)}{(12\gamma C^2 - 2)\cos(m) - (12\gamma C^2 + 4)}\right). \tag{32}$$

4.2. One-dimensional quadratic basis finite elements

Assembled Lagrange quadratic basis (superscript 'q') element integrals on a uniform grid are shown below. The assembled stencils are shown on the right-hand sides, with the upper row in the two-row matrices representing an element vertex node and the lower row a mid-element node:

$$\mathbf{M}^{\mathrm{q}}\boldsymbol{Q} = S_e(\mathbf{M}_e^{\mathrm{q}})\boldsymbol{Q} = S_e\left(\frac{\Delta_e}{30}\begin{bmatrix}4 & 2 & 1\\ 2 & 16 & 2\\ 1 & 2 & 4\end{bmatrix}\right)\boldsymbol{Q} = \frac{\Delta_x}{15}\begin{bmatrix}-1 & 2 & 8 & 2 & -1\\ 0 & 0 & 2 & 16 & 2\end{bmatrix}\begin{Bmatrix}\mathrm{e}^{-2im}\\ \mathrm{e}^{-im}\\ 1\\ \mathrm{e}^{+im}\\ \mathrm{e}^{+2im}\end{Bmatrix}Q_j, \tag{33}$$

$$\mathbf{U}^{\mathrm{q}}\boldsymbol{Q} = S_e(\mathbf{U}_e^{\mathrm{q}})\boldsymbol{Q} = S_e\left(\frac{\Delta_x}{6}\begin{bmatrix}-3 & 4 & -1\\ 4 & 0 & 4\\ 1 & -4 & 3\end{bmatrix}\right)\boldsymbol{Q} = \frac{\Delta_x}{6}\begin{bmatrix}1 & -4 & 0 & 4 & -1\\ 0 & 0 & -4 & 0 & 4\end{bmatrix}\begin{Bmatrix}\mathrm{e}^{-2im}\\ \mathrm{e}^{-im}\\ 1\\ \mathrm{e}^{+im}\\ \mathrm{e}^{+2im}\end{Bmatrix}Q_j, \tag{34}$$

$$\mathbf{K}^{\mathrm{q}}\boldsymbol{Q} = S_e(\mathbf{K}_e^{\mathrm{q}})\boldsymbol{Q} = S_e\left(\frac{\Delta_x^2}{3\Delta_e}\begin{bmatrix}7 & -8 & 1\\ -8 & 16 & -8\\ 1 & -8 & 7\end{bmatrix}\right)\boldsymbol{Q} = \frac{\Delta_x}{6}\begin{bmatrix}1 & -8 & 14 & -8 & 1\\ 0 & 0 & -8 & 16 & -8\end{bmatrix}\begin{Bmatrix}\mathrm{e}^{-2im}\\ \mathrm{e}^{-im}\\ 1\\ \mathrm{e}^{+im}\\ \mathrm{e}^{+2im}\end{Bmatrix}Q_j, \tag{35}$$

In the form used by Wait and Mitchell[25] to compute the amplification factor, the matrix stencils become

$$\mathbf{M}^{\mathrm{q}} = \frac{1}{15}\begin{bmatrix}8 - (\mathrm{e}^{+2im} + \mathrm{e}^{-2im}) & 2(\mathrm{e}^{+im} + \mathrm{e}^{-im})\\ 2(\mathrm{e}^{+im} + \mathrm{e}^{-im}) & 16\end{bmatrix}, \tag{36}$$

$$\mathbf{U}^{\mathrm{q}} = \frac{1}{6}\begin{bmatrix}-(\mathrm{e}^{+2im} - \mathrm{e}^{-2im}) & 4(\mathrm{e}^{+im} - \mathrm{e}^{-im})\\ 4(\mathrm{e}^{+im} - \mathrm{e}^{-im}) & 0\end{bmatrix}, \tag{37}$$

$$\mathbf{K}^{\mathrm{q}} = \frac{1}{6}\begin{bmatrix}14 + (\mathrm{e}^{+2im} + \mathrm{e}^{-2im}) & -8(\mathrm{e}^{+im} + \mathrm{e}^{-im})\\ -8(\mathrm{e}^{+im} + \mathrm{e}^{-im}) & 16\end{bmatrix}. \tag{38}$$

The amplification factor becomes a complex-valued 2 × 2 matrix with entries of (36)–(38):

$$G^{\mathrm{h}} = \left(\mathbf{M}^{\mathrm{q}} + \frac{C}{2}\mathbf{U}^{\mathrm{q}} + \gamma C^2\mathbf{K}^{\mathrm{q}}\right)^{-1}\left(\mathbf{M}^{\mathrm{q}} - \frac{C}{2}\mathbf{U}^{\mathrm{q}} + \gamma C^2\mathbf{K}^{\mathrm{q}}\right). \tag{39}$$

With $\beta = 0$ the phase velocity is

$$\Phi^{\mathrm{h}} = \frac{2}{mC}\tan^{-1} \times \left(\frac{C\sin(m)\left[\!\left[\sqrt{\{(2+5\gamma C^2)[18+45\gamma C^2+(2+15\gamma C^2)\sin^2(m)]\}} \pm (4-15\gamma C^2)\cos(m)\right]\!\right]}{4[(1-4\gamma C^2+15\gamma^2 C^4)\sin^2(m)+(1+15\gamma C^2)]}\right), \tag{40}$$

with the principal root having the negative sign. Gresho and Lee's[27] relation between semi-discrete and trapezoidal rule phase velocity was used to simplify (32) and (40).

4.3. *One-dimensional cubic basis finite elements*

The analysis for the cubic basis (superscript 'c') FEM is similar to that for the quadratic basis, except that Lagrange elements span four nodes with three different stencils, shown as the rows of the matrices on the right-hand sides below:

$$\mathbf{M}^{\mathrm{c}}\boldsymbol{Q} = S_e\left(\frac{\Delta_e}{1680}\begin{bmatrix} 128 & 99 & -36 & 19 \\ 99 & 648 & -81 & -36 \\ -36 & -81 & 648 & 99 \\ 19 & -36 & 99 & 128 \end{bmatrix}\right)\boldsymbol{Q}$$

$$= \frac{\Delta_x}{560}\begin{bmatrix} 19 & -36 & 99 & 256 & 99 & -36 & 19 \\ 0 & 0 & 0 & 99 & 648 & -81 & -36 \\ 0 & 0 & 0 & -36 & -81 & 648 & 99 \end{bmatrix}\begin{Bmatrix} \mathrm{e}^{-3im} \\ \mathrm{e}^{-2im} \\ \mathrm{e}^{-im} \\ 1 \\ \mathrm{e}^{+im} \\ \mathrm{e}^{+2im} \\ \mathrm{e}^{+3im} \end{Bmatrix} Q_j, \tag{41}$$

$$\mathbf{U}^{\mathrm{c}}\boldsymbol{Q} = S_e\left(\frac{\Delta_x}{80}\begin{bmatrix} -40 & 57 & -24 & 7 \\ -57 & 0 & 81 & 24 \\ 24 & -81 & 0 & 57 \\ -7 & -24 & -57 & 40 \end{bmatrix}\right)\boldsymbol{Q}$$

$$= \frac{\Delta_x}{80}\begin{bmatrix} -7 & 24 & -57 & 0 & 57 & -24 & 7 \\ 0 & 0 & 0 & -57 & 0 & 81 & -24 \\ 0 & 0 & 0 & 24 & -81 & 0 & 57 \end{bmatrix}\begin{Bmatrix} \mathrm{e}^{-3im} \\ \mathrm{e}^{-2im} \\ \mathrm{e}^{-im} \\ 1 \\ \mathrm{e}^{+im} \\ \mathrm{e}^{+2im} \\ \mathrm{e}^{+3im} \end{Bmatrix} Q_j, \tag{42}$$

$$\mathbf{K}^{c}\mathbf{Q} = S_e \left(\frac{\Delta_x^2}{40\Delta_e} \begin{bmatrix} 148 & -189 & 54 & -13 \\ -189 & 432 & -297 & 54 \\ 54 & -297 & 432 & -189 \\ -13 & 54 & -189 & 148 \end{bmatrix} \right) \mathbf{Q}$$

$$= \frac{\Delta_x}{120} \begin{bmatrix} -13 & 54 & -189 & 296 & -189 & 54 & -13 \\ 0 & 0 & 0 & -189 & 432 & -297 & 54 \\ 0 & 0 & 0 & -297 & 54 & 432 & -189 \end{bmatrix} \begin{Bmatrix} e^{-3im} \\ e^{-2im} \\ e^{-im} \\ 1 \\ e^{+im} \\ e^{+2im} \\ e^{+3im} \end{Bmatrix} Q_j, \tag{43}$$

The matrix form of the amplification factor is similar to that for the quadratic basis finite element, equation (39), except that the matrix stencils become 3×3 matrices:

$$\mathbf{M}^{c} = \frac{\Delta_x}{560} \begin{bmatrix} 256 + 19(e^{+3im} + e^{-3im}) & 99\, e^{+im} - 36\, e^{-2im} & -36\, e^{+2im} + 99\, e^{-im} \\ -36\, e^{+2im} + 99\, e^{-im} & 648 & -81\, e^{+im} \\ 99\, e^{+im} - 36\, e^{-2im} & -81\, e^{-im} & 648 \end{bmatrix}, \tag{44}$$

$$\mathbf{U}^{c} = \frac{\Delta_x}{80} \begin{bmatrix} 7(e^{+3im} - e^{-3im}) & 57\, e^{+im} + 24\, e^{-2im} & -24\, e^{+2im} - 57\, e^{-im} \\ -24\, e^{+2im} - 57\, e^{-im} & 0 & 81\, e^{+im} \\ 57\, e^{+im} + 24\, e^{-2im} & -81\, e^{-im} & 0 \end{bmatrix}, \tag{45}$$

$$\mathbf{K}^{c} = \frac{\Delta_x}{120} \begin{bmatrix} 296 - 13(e^{+3im} + e^{-3im}) & -189\, e^{+im} + 54\, e^{-2im} & 54\, e^{+2im} - 189\, e^{-im} \\ 54\, e^{+2im} - 189\, e^{-im} & 432 & -297\, e^{+im} \\ -189\, e^{+im} + 54\, e^{-2im} & -297\, e^{-im} & 432 \end{bmatrix}. \tag{46}$$

We did not obtain a sufficiently simplified closed form for the phase velocity and it was calculated at discrete values of C and m.

4.4. Finite difference and finite volume methods

Several finite difference and finite volume methods are included for comparison. Substituting a Taylor weak statement correction of $\gamma = 1/6C^2$ in (20) removes the contribution of the interpolation or 'mass' matrix $\mathbf{M}^{l}$, replacing it by a scaled identity matrix and producing a method equivalent to a finite volume method on a uniform mesh. Improvements to finite volume methods are usually stated in terms of modifications to the convection operator, which in its unmodified form (δ) is identical with the assembled linear basis finite element uniform mesh convection matrix ($\mathbf{U}^{l}$). The resulting scheme is the well-known Crank–Nicolson finite difference method with amplification factor

$$G^{h} = \frac{\frac{C}{4}e^{-im} + 1 - \frac{C}{4}e^{+im}}{-\frac{C}{4}e^{-im} + 1 + \frac{C}{4}e^{+im}}. \tag{47}$$

The group of QUICK finite volume methods has been widely used. The QUICK methods are intended to improve standard finite volume methods by using various forms of higher-order approximations to the convection term. The third-order upwind QUICK approximation[11] for the first derivative, designated here as δ_{Q3}, results in the following finite difference stencil with positive velocity on a uniform 1D mesh:

$$\delta_{Q3}(Q) = \frac{1}{8}Q_{j-2} - \frac{7}{8}Q_{j-1} + \frac{3}{8}Q_{j+1} + \frac{3}{8}Q_{j+2}. \tag{48}$$

The resulting amplification factor is

$$G^h = \frac{-\frac{C}{16}e^{-2im} + \frac{7C}{16}e^{-im} + \left(1 - \frac{3C}{16}\right) - \frac{3C}{16}e^{+im}}{\frac{C}{16}e^{-2im} - \frac{7C}{16}e^{-im} + \left(1 + \frac{3C}{16}\right) + \frac{3C}{16}e^{+im}}. \tag{49}$$

The fifth-order upwind QUICK approximation[12] for the first derivative, designated here as δ_{Q5}, results in the following finite difference stencil with positive velocity on a uniform 1D mesh:

$$\delta_{Q5}(q) = -\frac{9}{384}Q_{j-3} + \frac{77}{384}Q_{j-2} - \frac{346}{384}Q_{j-1} + \frac{90}{384}Q_j + \frac{211}{384}Q_{j+1} - \frac{23}{384}Q_{j+2}. \tag{50}$$

Small multidimensional terms are neglected in (50). The resulting amplification factor is

$$G^h = \frac{\frac{9C}{768}e^{-3im} - \frac{77C}{768}e^{-2im} + \frac{346C}{768}e^{-im} + \left(1 - \frac{90C}{768}\right) - \frac{211C}{768}e^{+im} + \frac{23C}{768}e^{+2im}}{-\frac{9C}{768}e^{-3im} + \frac{77C}{768}e^{-2im} - \frac{346C}{768}e^{-im} + \left(1 + \frac{90C}{768}\right) + \frac{211C}{768}e^{+im} - \frac{23C}{768}e^{+2im}}. \tag{51}$$

5. RESULTS OF FREQUENCY ANALYSIS FOR EXAMPLE METHODS

The relative velocity Φ^h and amplification factor modulus $|G^h|$ were calculated for the example methods at several Courant numbers C between 0 and 1 and wave numbers m between 0 and π. The trapezoidal rule of $\theta = \frac{1}{2}$ was used for time integration for each method. Phase velocity results are shown in Figure 1(a) for a Courant number of $\frac{1}{8}$.

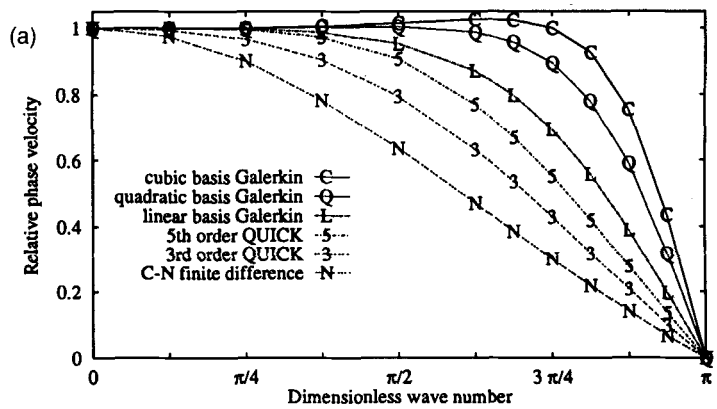

Figure 1(a). Phase velocity of various algorithms at $C = \frac{1}{8}$

Figure 1(a) shows a pattern that is repeated at larger Courant numbers, in which the Crank–Nicolson finite difference method has the poorest phase velocity of the tested algorithms. Third-order QUICK is improved over Crank–Nicolson and fifth-order QUICK is improved over third-order. Linear basis Galerkin is better in phase velocity than each finite volume method, quadratic basis Galerkin is better than linear and cubic basis Galerkin is better than quadratic. The relative ranking between the six non-Taylor algorithms is the same at each Courant number. At this small Courant number the TWS correction γC^2 is insignificant and the TWS methods (not shown) become essentially identical with the corresponding Galerkin methods.

Additional results are shown in Figures 1(b) and 1(c) for Courant numbers of $\frac{1}{2}$ and 1 respectively. The trend in Figures 1(b) and 1(c) is similar to that in Figure 1(a) except that the three TWS algorithms become more accurate than the corresponding Galerkin algorithms as the Courant number approaches unity. The other six algorithms become less accurate as the Courant number increases. At $C = 1$ all three TWS algorithms become very accurate in phase velocity for almost the entire range of wave number. The TWS algorithms were roughtly optimized for phase velocity by varying the constant γ as $-1/12$ for linear basis, $-1/15$ for quadratic basis and $-1/18$ for cubic basis. The discrete time Taylor series of (10) predicted that the value should be $-1/12$.

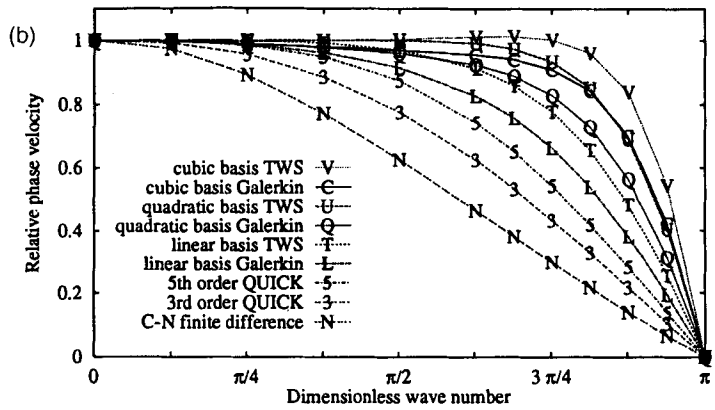

Figure 1(b). Phase velocity of various algorithms at $C = \frac{1}{2}$

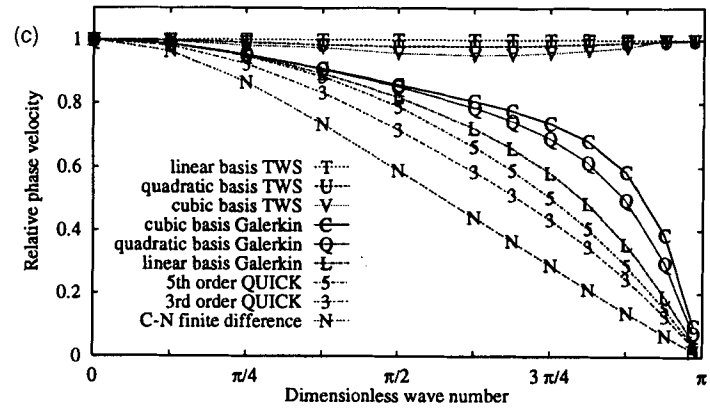

Figure 1(c). Phase velocity of various algorithms at $C = 1$

The Crank–Nicolson finite difference algorithm and each of the six finite element algorithms are non-diffusive and have an amplfication factor modulus identically equal to unity. The amplification factor modulus is shown in Figure 2 at $C = \frac{1}{2}$ for the QUICK methods and for a comparative linear upwind finite difference method, using $\gamma = 1/6C^2$ in (20) to replicate a finite difference method and an artifical diffusion term $\beta = 1/2C$ to eliminate the downstream (e^{+im}) term in the numerator of (31). Also shown for comparison are dissipative linear basis Galerking finite element methods with $\beta = \frac{1}{8}, \frac{1}{6}$, and $\frac{1}{3}$, which produce a level of short-wavelength dissipation comparable with that of the third-order and fifth-order QUICK methods and the linear upwind method respectively.

Figure 2 confirms that the QUICK methods are dissipative, since each modulus is uniformly less than unity, but are less dissipative than the linear upwind finite difference method. Dissipation at large wave numbers approaching π is considered advantageous for stability, since $2\Delta_x$ waves are damped accordingly. The linear upwind method also has significant dissipation at wave numbers as small as $\pi/4$, indicating that every solution component is heavily damped and accuracy is compromised. The trend is similar at other Courant numbers (results not shown). Each of the listed dissipative algorithms becomes more dissipative at higher Courant number.

6. FIFTH-ORDER-ACCURATE LINEAR BASIS TWS FINITE ELEMENT METHOD

The phase velocity results in Figures 1(a)–1(c) show that the linear basis TWS correction of $\gamma = -1/12$ onto C^2 is most effective for $C = 1$, at which value the method has the correct phase velocity over most of the range of wave number m. The TWS correction is less effective at smaller Courant numbers. This is illustrated by a Taylor expansion of the amplification factor G^h in powers of the wave number m about $m = 0$:

$$G^h = 1 - imC - \frac{(mC)^2}{2} + \frac{i(mC)^3(1+4\gamma)}{4} + \frac{(mC)^4(1+8\gamma)}{8} - \frac{im^5C[45C^4 + 60\gamma C^2(1-9C^2) + 720\gamma^2C^4 - 4]}{720} - \frac{m^6C^2[45C^4 + 120\gamma C^2(1-6C^2) + 2160\gamma^2C^4 - 8]}{1440} + \cdots . \tag{52}$$

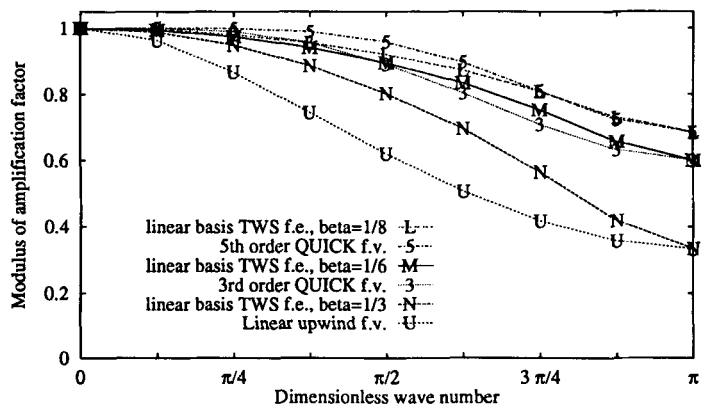

Figure 2. Amplification factor modulus of dissipative algorithms at $C = \frac{1}{2}$

This is compared, following the procedure of Baker and Kim,[18] with a similar expansion of the analytical amplification factor G:

$$G = e^{-imC} = 1 - imC - \frac{(mC)^2}{2} + \frac{i(mC)^3}{6} + \frac{(mC)^4}{24} - \frac{i(mC)^5}{120} - \frac{(mC)^6}{720} + \frac{i(mC)^7}{5040} + \frac{(mC)^8}{40320} - \frac{i(mC)^9}{362880} - \frac{(mC)^{10}}{3628800} + \cdots . \tag{53}$$

The analytical series (53) converges slowly for large wave numbers and Courant numbers, since the eighth-order term has a magnitude of about 0·23 and the 10th-order term a magnitude of about 0·03 for $m = \pi$ and $C = 1$. An accurate approximation at a Courant number near unity would then require about 10–12 orders of accuracy. Comparing equations (52) and (53), the Galerkin linear basis method with $\gamma = 0$ is second-order-accurate in wave number. The linear basis TWS method with $\gamma = -1/12$ is fourth-order-accurate in wave number, except for the special case of $C = 1$ which is at least 12th-order-accurate.

Phase accuracy can be improved by adapating an idea from the early work of Stone and Brian.[14] A two-step method is defined, one step with phase velocity too high and the other step with phase velocity too low, such that the mean phase velocity is approximately correct for an even number of steps. Defining a variable TWS correction term γ to be used at alternating time steps allows equations (52) and (53) to be equated to order five, yielding

$$\gamma = -\frac{1}{12}\left[1 \pm \frac{1}{C^2}\left(\frac{(4 - 5C^2 + C^4)}{5}\right)^{1/2}\right]. \tag{54}$$

The sixth-order term, also quadratic in γ, cannot be satisfied simultaneously in this form, since it requires

$$\gamma = -\frac{1}{12}\left[1 \pm \frac{1}{C^2}\left(\frac{2(4 - 5C^2 + C^4)}{15}\right)^{1/2}\right]. \tag{55}$$

For $C = \frac{1}{2}$ the solution of the fifth-order equation (54) is

$$\gamma = -\frac{1}{12}(1 \pm 3). \tag{56}$$

This method will be referred to here as the 'fifth-order' linear basis TWS method, not to be confused with the differently measured order of the 'fifth-order QUICK' method. The relative phase velocity of the individual steps of the fifth-order method is shown in Figure 3.

The phase velocity of the 'odd' steps in Figure 3 is higher than unity, decreasing with Courant number. The phase velocity of the 'even' steps is smaller than unity, increasing with Courant number, such that the mean phase velocity is nearly unity. Resulting mean phase velocities are shown in Figure 4 and compared with those of the fourth-order TWS method.

Figure 4 shows that the fifth-order mean phase velocities are slightly high for $C < \frac{1}{2}$, correct for $C = \frac{1}{2}$ and slightly low for $C > \frac{1}{2}$. The fourth-order TWS phase velocity is lower than the fifth-order at each Courant number.

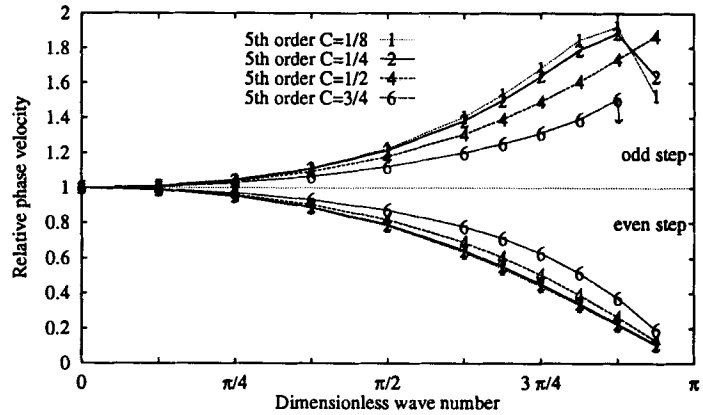

Figure 3. Individual step phase velocities of two-step linear basis algorithms

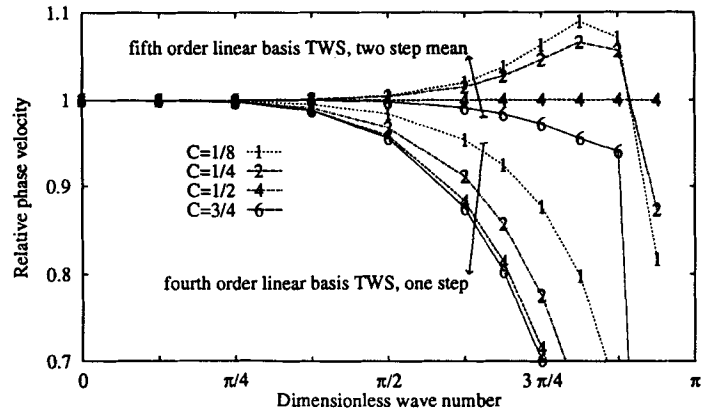

Figure 4. Phase velocity comparison between linear basis fifth-order mean and fourth-order

7. LINEAR BASIS TWS FINITE ELEMENT METHOD WITH APPROXIMATELY MINIMIZED PHASE ERROR

The preceding phase velocity results (and verification problem results to be discussed) show that the fifth-order method, while improved over the standard fourth-order method, is not optimal except at $C = \frac{1}{2}$. Higher-order matching of the amplification factor expansion does not appear practical for the 10 or so significant terms at Courant number near unity. The fifth-order method has the additional disadvantage of lacking diagonal dominance for $C > \frac{1}{2}$.

This section shows an alternative development of minimizing the error in phase velocity rather than exactly matching a low-order expansion of the amplification factor. Again we use a multiple-step method, with two steps for $C < \frac{1}{2}$ and four steps for $C > \frac{1}{2}$. The phase error from (32) does not appear to have a known analytical integral over m, so we numerically integrate the phase error over m from 0 to π for many possible values of γ and combine the steps so as to minimize the phase error in some manner. Following the accuracy criterion of Vichnevetsky and De Schutter,[20] we arbitrarily pick the numerical solution (γ_1, γ_2) which maintains 1% phase accuracy over the widest possible range of m.

This range turns out to be from $m = 0$ to about $3\pi/4$ for most values of C. The individual steps are combined as (γ_1, γ_2) for two-step solutions ($C < \frac{1}{2}$) and $(\gamma_1, \gamma_1, \gamma_1, \gamma_2)$ for four-step solutions ($C > \frac{1}{2}$). The selected values of γ at various Courant numbers are then curve fitted to Courant number for $C \leq \frac{1}{2}$ as

$$(\gamma_1, \gamma_2) = -\frac{1}{12} \pm \frac{0{\cdot}07025 - 0{\cdot}02609C^2 - 0{\cdot}00982C^3}{C^2} \tag{57}$$

and for $C > \frac{1}{2}$ as

$$\gamma_1 = -\min\left(\frac{1}{12C^2}, \frac{2C+1}{30C^2}\right), \tag{58}$$

with, for $\frac{1}{2} \leq C \leq \frac{3}{4}$,

$$\gamma_2 = -0{\cdot}061 + 0{\cdot}793(1 - C) - 0{\cdot}792(1 - C)^2 \tag{59}$$

and, for $\frac{3}{4} \leq C \leq 1$,

$$\gamma_2 = -\frac{1}{12} + 0{\cdot}3427(1 - C) + 1{\cdot}3364(1 - C)^2. \tag{60}$$

This approximate procedure is not represented to be the 'best' solution of this type, but as a 'good' solution for phase accuracy. The resulting phase velocity, the mean of two or four steps, is shown in Figure 5 for various Courant numbers. Note the highly expanded ordinate scale.

The preceding figures show that the approximately optimized method maintains better than 1% accuracy in phase velocity up to wave numbers of at least $3\pi/4$ for the Courant numbers shown. The method is superior in phase accuracy to the fourth-order and fifth-order linear basis TWS methods.

8. ONE-DIMENSIONAL VERIFICATION PROBLEMS

Several of the derived algorithms are verified for one-dimensional linear diffusive and non-diffusive transient advection. The non-diffusive verification problem is advection of a two-node wide square wave, with the analytical solution translation of the initial conditions, selected as

$$q(t = 0{\cdot}2) = \begin{cases} 0, & 0 \leq x < 0{\cdot}12, \\ 1, & 0{\cdot}12 \leq x < 0{\cdot}16, \\ 0, & 0{\cdot}16 \leq x \leq 0{\cdot}96. \end{cases} \tag{61}$$

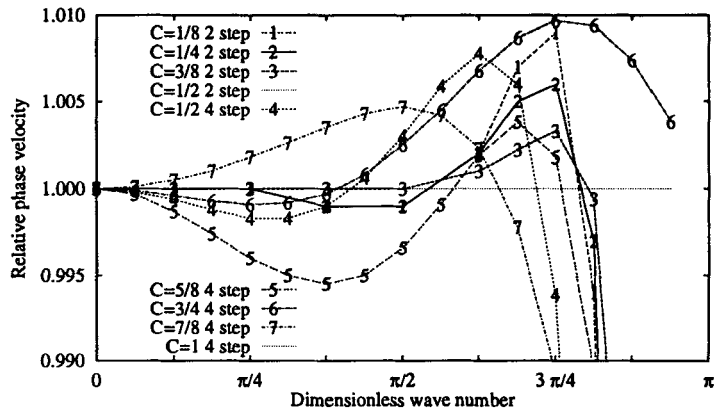

Figure 5. Phase velocity of approximately optimized method at various Courant numbers

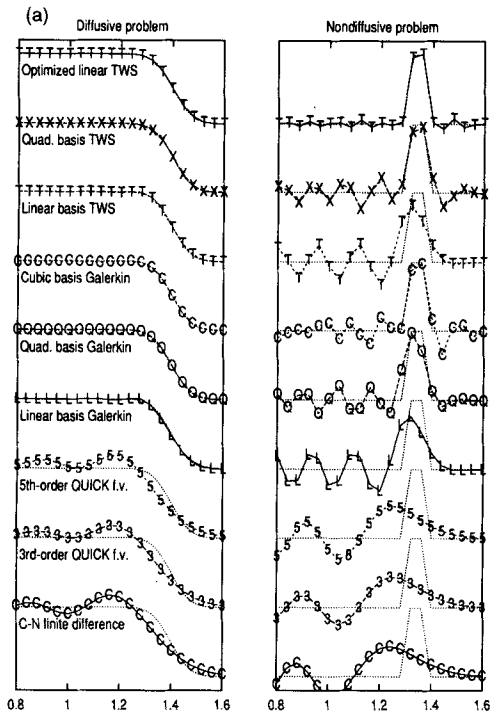

Figure 6(a). 1D verification problems, $C = \frac{1}{4}$

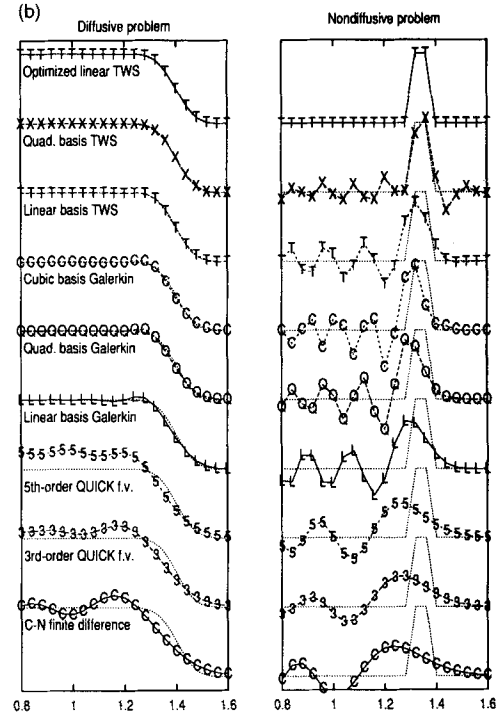

Figure 6(b). 1D verification problems, $C = \frac{1}{2}$

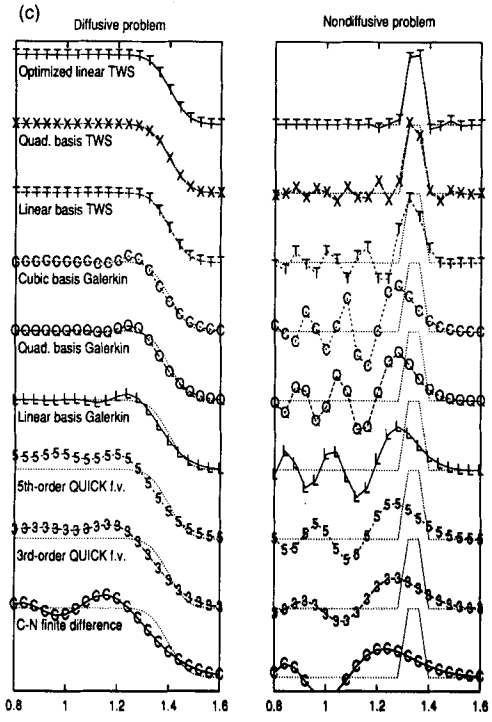

Figure 6(c). 1D verification problems, $C = \frac{3}{4}$

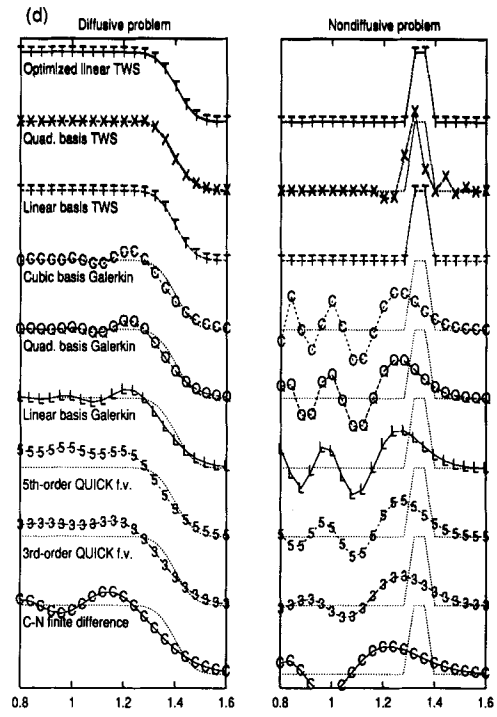

Figure 6(d). 1D verification problems, $C = 1$

The diffusive verification problem uses a known analytical solution of (1) with the Peclet number Pe set at 800:

$$q = \frac{e^{xPe}}{2}\left\{1 - \text{erf}\left[\left(\frac{Pe}{t}\right)^{1/2}\frac{x+t}{2}\right]\right\} + \left\{1 - \text{erf}\left[\left(\frac{Pe}{t}\right)^{1/2}\frac{x-t}{2}\right]\right\}. \tag{62}$$

Each problem was simulated for $0{\cdot}2 \leq t \leq 1{\cdot}4$ on a 40-node uniform mesh with $\Delta_x = 0{\cdot}04$ and time step varied to produce different Courant numbers. The mesh was extended to 42 nodes for cubic elements. The analytical solution was used as initial condition for the diffusive problem and the non-dissipative trapezoidal rule of $\theta = \frac{1}{2}$ was used for each algorithm. The boundary conditions were Dirichlet at $x = 0$ and Neumann at $x = L$. The Courant numbers were $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ and 1, with results shown in Figures 6(a)–6(d) respectively.

Figure 6(a) at $C = \frac{1}{4}$ shows that each finite element method produces relatively good solutions for the diffusive problem. The cubic basis Galerkin, quadratic basis TWS and optimized linear basis TWS methods produce relatively good solutions for the non-diffusive problem, with slight to moderate error waves both leading and lagging. The fifth-order TWS method produces slight leading error waves. Each of the finite difference/finite volume methods produces considerable lagging error waves for the diffusive problem and worse lagging error waves for the non-diffusive problem. The QUICK methods are slightly improved in phase lag over unmodified finite difference.

Figure 6(b) at $C = \frac{1}{2}$ shows much the same results. The finite difference/finite volume results are of similar poor quality as those for $C = \frac{1}{4}$. Each finite element solution is acceptable for the diffusive problem. For the non-diffusive problem each Galerkin method produces large lagging error waves. The linear basis TWS solution lags somewhat less than linear basis Galerkin, the quadratic basis TWS solution has moderate leading and lagging error and the fifth-order and optimized linear basis TWS solutions are nearly nodally exact.

The trends continue in Figure 6(c) at $C = \frac{3}{4}$. At this higher Courant number each Galerkin solution has noticeable phase lag for the diffusive problem, but each TWS method has a relatively good solution. For the non-diffusive problem each Galerkin method has large phase lag, with quadratic basis slightly preferred. Standard linear basis TWS, fifth-order TWS and quadratic basis Galerkin solutions produce moderate error waves. The optimized linear basis TWS solution is not nodally accurate, as occurred at $C = \frac{1}{2}$, but is superior to the other solutions.

In Figure 6(d) at $C = 1$ only the TWS methods produced reasonably good solutions to either diffusive or non-diffusive problems. The standard linear basis TWS and the fifth-order and optimized linear basis TWS methods produce nodally identical solutions to the non-diffusive problem, while the quadratic basis TWS solution produces moderate leading and lagging error waves.

9. TWO-DIMENSIONAL NON-DIFFUSIVE VERIFICATION PROBLEM

A classical two-dimensional verification problem simulates a half-wave cosine hill distribution on a rotational flow field (a 'rotating cone'). The analytical solution is circular translation of the non-zero initial condition. The centre of the cone was located initially at $r = 5L/16$ on a square grid of 32×32 bilinear or 16×16 biquadratic basis finite elements. The cosine hill distribution peak was 100 and spanned six linear elements in each Cartesian direction, or five non-zero nodes and two zero nodes. The simulation duration was one rotation or $20\pi/C$ time steps for Courant number at the cone peak. Boundary conditions were Dirichlet fixed at zero on inflow boundaries and Neumann on outflow boundaries. The time-split matrix factorization algorithm[30] was used to produce a quasi-one-dimensional alternating direction solution process. The trapezoidal rule of $\theta = \frac{1}{2}$ was used for each algorithm.

Two-dimensional simulation errors are summarized in Figure 7. Each group of four bars represents results at $C = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ and 1. The maximum relative error in minimum value, or magnitude of the largest negative wave, is shown in Figure 7(a). Third-order QUICK and the four TWS algorithms are the most accurate in this measure.

The maximum relative error in maximum value, or deviation of the largest wave from the analytical solution, is shown in Figure 7(b). Quadratic Galerkin and the four TWS algorithms are the most accurate in this measure.

The mean RMS error is also shown for the 1021 unconstrained nodes in Figure 7(c). The four TWS algorithms are the most accurate in this measure.

Selected simulation results are also illustrated in Figure 8 for $C = \frac{1}{2}$, with Figure 8(a) showing the initial condition and identical analytical solution after one rotation. The Crank–Nicolson (C–N) finite difference solution in Figure 8(b) shows large trailing error waves with significant peak reductions and peak lag of several nodes.

The QUICK solutions in Figures 8(c) and 8(d) are improved over Crank–Nicolson but show considerable phase lag and excess diffusion, particularly for the third-order QUICK method.

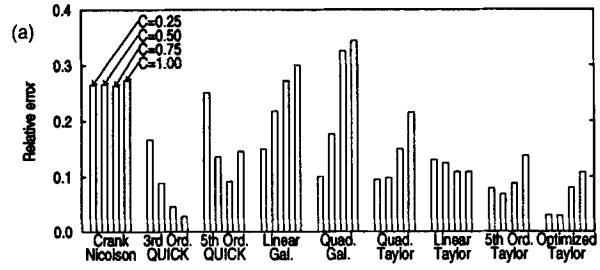

Figure 7(a). Error in minimum value for two-dimensional verification problem

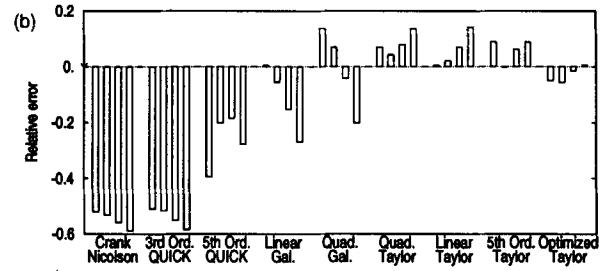

Figure 7(b). Error in maximum value for two-dimensional verification problem

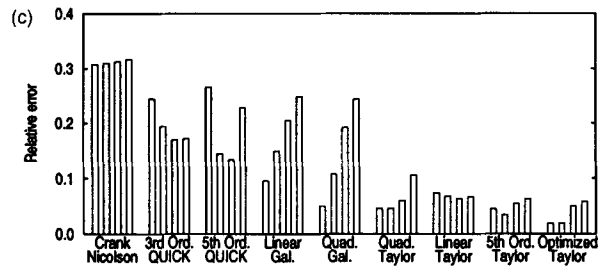

Figure 7(c). RMS error for two-dimensional verification problem

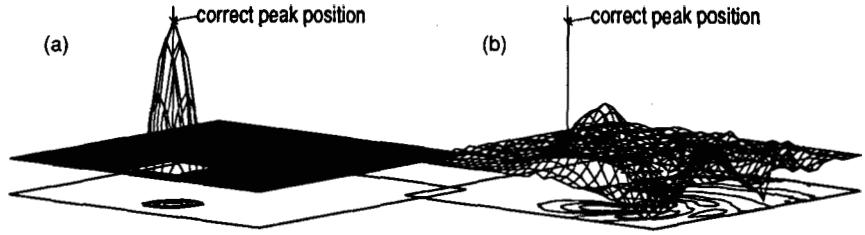

Figure 8(a). Analytical solution

Figure 8(b). C–N finite difference, $C = \frac{1}{2}$

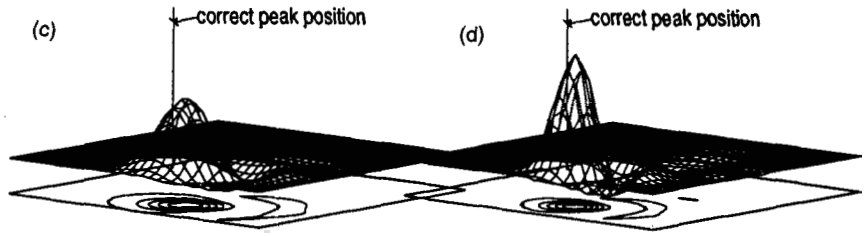

Figure 8(c).Third-order QUICK, $C = \frac{1}{2}$

Figure 8(d). Fifth-order QUICK, $C = \frac{1}{2}$

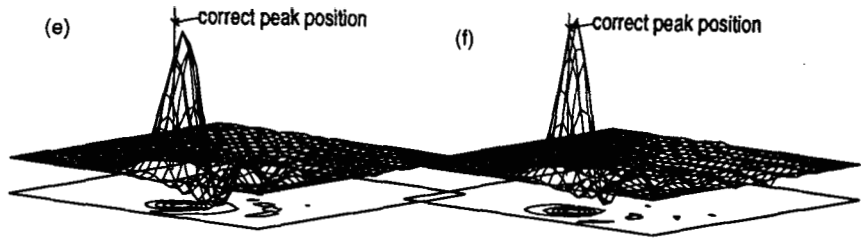

Figure 8(e). Linear basis Galerkin, $C = \frac{1}{2}$

Figure 8(f). Quadratic basis Galerkin, $C = \frac{1}{2}$

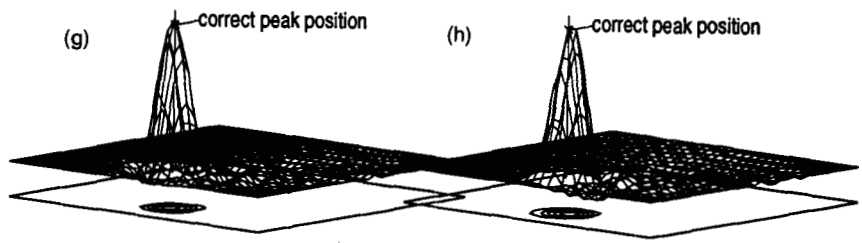

Figure 8(g). Quadratic basis TWS, $C = \frac{1}{2}$

Figure 8(h). Linear basis TWS, $C = \frac{1}{2}$

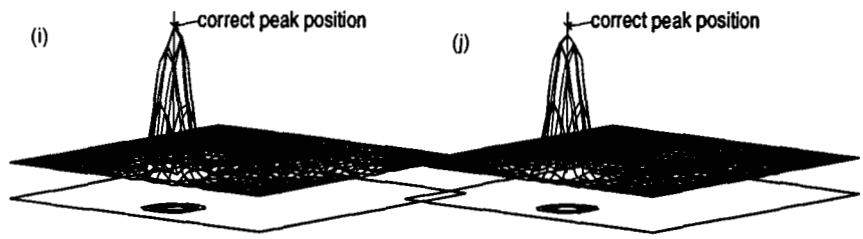

Figure 8(i) Fifth-order TWS solution, $C = \frac{1}{2}$

Figure 8(j) Optimized TWS solution, $C = \frac{1}{2}$

The linear basis Galerkin solutions such as Figure 8(e) show considerable error waves, the quality deteriorating with increasing Courant number. Nodal peaks are lagging by one to two nodes. Quadratic basis Galerkin solutions such as Figure 8(f) are slightly improved over linear basis Galerkin.

Quadratic basis and linear basis TWS solutions such as Figures 8(g) and 8(h) are considerably improved over corresponding Galerkin solutions, with little phase lag and small error waves.

Fifth-order TWS solutions such as Figure 8(i) show further improvement, particularly at $C = \frac{1}{2}$. Optimized linear basis TWS solutions such as Figure 8(j) are the best of the various algorithms at this Courant number. Since the Courant number is distributed, the fifth-order and optimized TWS solutions are not equal as they are in one dimension at $C = \frac{1}{2}$.

Figure 9 shows selected solutions at $C = \frac{3}{4}$. Again the linear basis TWS solution (Figure 9(b) is improved over linear basis Galerkin (Figure 9(a)).

The multiple-step solutions of fifth-order TWS (Figure 9(c)) and optimized TWS (Figure 9(d)) are again improved over standard Galerkin and TWS solutions.

10. CONCLUSIONS

A phase velocity analysis for one-dimensional transient convection was performed for linear, quadratic and cubic basis Galerkin and TWS finite element methods as well as for finite difference and QUICK finite volume methods.

Each Galerkin and TWS method was more phase-accurate than each finite difference or finite volume method at each Courant number up to unity. At small Courant numbers the cubic basis Galerkin method was the most accurate. As the Courant number approached unity, each TWS method became substantially more accurate than the corresponding Galerkin method.

The trapezoidal rule linear basis TWS method was found to be fourth-order-accurate in wave number. A two-step linear basis TWS method was formulated to be fifth-order-accurate in wave number. Another two- to four-step linear basis TWS method was formulated to approximately optimize phase accuracy.

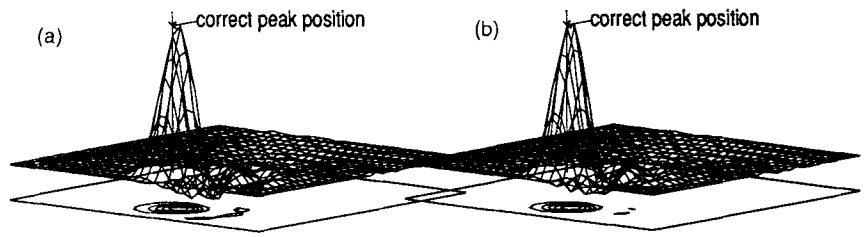

Figure 9(a). Linear basis Galerkin, $C = \frac{3}{4}$

Figure 9(b). Linear basis TWS, $C = \frac{3}{4}$

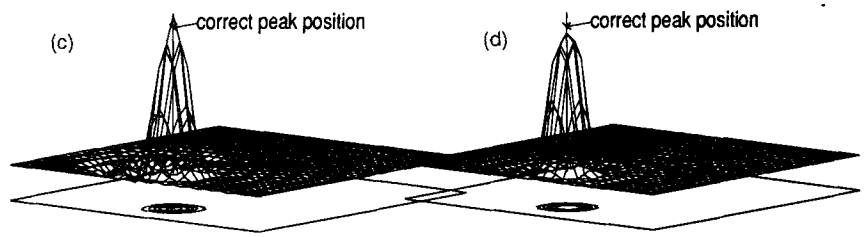

Figure 9(c). Fifth-order TWS solution, $C = \frac{3}{4}$

Figure 9(d). Optimized TWS solution, $C = \frac{3}{4}$

The various methods were tested on one-dimensional verification problems with small or zero diffusion and on a two-dimensional verification problem with zero diffusion. The relative performance of the methods on the verification problems correlated very well with the relative phase accuracy of the methods. The optimized linear basis TWS and cubic basis Galerkin methods were the most accurate on verification problems at small Courant number, while the several linear basis TWS methods were the most accurate at Courant number near unity. The optimized linear basis TWS and quadratic basis TWS methods had the best overall accuracy for the range of Courant numbers from zero to unity.

Results of the verification problems also showed that the most phase-accurate methods were capable of producing accurate solutions to linear transient hyperbolic problems without any artificial diffusion or peak limiting. The tested methods may still require artificial diffusion for non-linear problems, but should require much less diffusion than would a standard method. In their useful range of Courant number less than unity the best of the tested methods rival or surpass the solution quality of much more complex methods.

REFERENCES

1. R. Courant, E. Isaacson and M. Rees, 'On the solution of nonlinear hyperbolic differential equations', *Commun. Pure and Appl. Math.*, **5**, 243–255 (1952).
2. I. Christie, D. F. Griffiths, A. R. Mitchell and O. C. Zienkiewicz, 'Finite element methods for second order differential equations with significant first derivative', *Int. j. numer. methods eng.*, **10**, 1389–1396 (1976).
3. I. Christie and A. R. Mitchell, 'Upwinding of high order Galerkin methods in conduction–convection problems', *Int. j. numer. methods eng.*, **12**, 1764–1771 (1978).
4. J. C. Heinrich, P. S. Huyakorn, O. C. Zienkiewicz and A. R. Mitchell, 'An "upwind" finite element scheme for two-dimensional convective transport equation', *Int. j. numer. methods eng.*, **11**, 131–143 (1977).
5. J. C. Heinrich and O. C. Zienkiewicz, 'Quadratic finite element schemes for two-dimensional convective-transport problems', *Int. j. numer. methods eng.*, **11**, 1831–1844 (1977).
6. A. N. Brooks and T. J. R. Hughes, 'Streamline upwind Petrov–Galerkin methods for advection dominated flows', *Proc. Third Int. Conf. on Finite Element Methods in Fluid Flow*, Banff, 1980.
7. E. Dick, 'Accurate Petrov–Galerkin methods for transient convective diffusion problems', *Int. j. numer. methods eng.*, **19**, 1425–1433 (1983).
8. J. J. Westerink and D. Shea, 'Consistent higher degree Petrov–Galerkin methods for the solution of the transient convection–diffusion equation', *Int. j. numer. methods eng.*, **28**, 1077–1101 (1989).
9. E. T. Bouloutas and M. A. Celia, 'An improved cubic Petrov–Galerkin method for simulation of transient advection–diffusion processes in rectangularly decomposable domains', *Comput. Methods Appl. Mech. Eng.*, **92**, 289–308 (1991).
10. N. Kondo, N. Tosaka and T. Nishimura, 'Computation of the incompressible viscous flows by the third-order upwind finite element method', *Int. j. numer. methods fluids*, **15**, 1013–1024 (1992).
11. B. P. Leonard, 'A stable and accurate convective modelling procedure based on quadratic upstream interpolation', *Comput. Methods Appl. Mech. Eng.*, **19**, 59–98 (1979).
12. B. P. Leonard and S. Mokhtari, 'ULTRA-SHARP solution of the Smith–Hutton problem', *Technical Memorandum 105435*, NASA Lewis Research Center, Cleveland, OH, 1992.
13. P. Lax and B. Wendroff, 'Systems of conservation laws', *Commun. Pure Appl. Math.*, **8**, 217–237 (1960).
14. H. L. Stone and P. L. T. Brian, 'Numerical solution of convective transport problems', *AIChE J.*, **9**, 681–688 (1963).
15. A. Harten and H. Tal-Ezer, 'On a fourth order accurate implicit finite difference scheme for a hyperbolic conservation laws: I. Nostiff strongly dynamic problems', *Math. Comput.*, **36**, 353–373 (1981).
16. M. Th. van Genuchten and W. G. Gray, 'Analysis of some dispersion corrected numerical schemes for solution of the transport equation', *Int. j. numer. methods. eng.*, **12**, 387–404 (1978).
17. J. Donea, 'A Taylor–Galerkin method for convective transport problems', *Int. j. numer. methods eng.*, **20**, 101–119 (1984).
18. 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).
19. J. von Neumann and R. D. Richtmeyer, 'A method for the numerical calculation of hydrodynamic shocks', *J. Appl. Phys.*, **21**, 232–237 (1950).
20. R. Vichnevetsky and F. De Schutter, 'A frequency analysis of finite difference and finite-element methods for initial value problems', in *Advances in Computer Methods for Partial Differential Equations*, AICA, Ghent, 1975.
21. W. G. Gray and G. F. Pinder, 'An analysis of the numerical solution of the transport equation', *Water Resources Res.*, **12**, 547–555 (1976).
22. W. H. Raymond and A. Garder, 'Selective damping in a Galerkin method for solving wave problems with variable grids', *Mon. Weather Rev.*, **104**, 1583–1590 (1976).

23. P. M. Gresho, R. L. Lee and R. L. Sani, 'Advection-dominated flows, with emphasis on the consequences of mass lumping', in *Finite Elements in Fluids*, Vol. 3, Wiley, New York, 1978.
24. R. C. Y. Chin, G. W. Hedstrom and K. E. Karlsson, 'A simplified Galerkin method for hyperbolic equations', *Math. Comput.*, **33**, 647–658 (1979).
25. R. Wait and A. R. Mitchell, *Finite Element Analysis and Applications*, Wiley, New York, 1985.
26. B. Cathers and B. A. O'Connor, 'The group velocity of some numerical schemes', *Int. j. numer. methods fluids*, **5**, 201–224 (1985).
27. P. M. Gresho and R. L. Lee, 'Comments on "The group velocity of some numerical schemes"', *Int. j. numer. methods fluids*, **7**, 1357–1362 (1987).
28. J. Donea, L. Quartapelle and V. Selmin, 'An analysis of time discretization in the finite element solution of hyperbolic problems', *J. Comput. Phys.*, **70**, 463–499 (1987).
29. *Macsyma Reference Manual, Version 13*, Symbolics, Inc., Cambridge, MA, 1988.
30. A. J. Baker, *Finite Element Computational Fluid Dynamics*, Hemisphere, New York, 1983.