JOURNAL OF COMPUTATIONAL PHYSICS 39, 46–71 (1981)

Analysis of an Upstream Weighted Collocation Approximation to the Transport Equation

ALLEN SHAPIRO* AND GEORGE F. PINDER†

Princeton University, Princeton, New Jersey 08544

Received July 31, 1979; revised November 5, 1979

The numerical behavior of a modified orthogonal collocation method, as applied to the transport equations, can be examined through the use of a Fourier series analysis. The necessity of such a study becomes apparent in the analysis of several techniques which emulate classical upstream weighting schemes. These techniques are employed in orthogonal collocation and other numerical methods as a means of handling parabolic partial differential equations with significant first-order terms. Divergent behavior can be shown to exist in one upstream weighting method applied to orthogonal collocation.

INTRODUCTION

Under certain simplifying assumptions mass transport in one space dimension can be described by the relationship

$$L(c) = \frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} - D\frac{\partial^2 c}{\partial x^2} = 0, \tag{1}$$

where the velocity $u(L/T)$ and the dispersion coefficient $D(L^2/T)$ are assumed to be specified constant coefficients. It is well documented that numerical solutions to this equation are characterized by oscillations when the convective term is dominant (see, for example, Strang and Fix [4, p. 253] or Gray and Pinder [1]). As a means of alleviating this numerical difficulty at the cost of smearing the solution profile, numerically symmetric techniques are modified to account for the fact that the convective process has a directional or one-sided nature. One such modification, known as upstream weighting, is well known when applied to finite difference schemes and has recently been extended to Galerkin finite element formulations (see, for example, Huyakorn [2]). Pinder and Shapiro [3] applied a similar methodology to the orthogonal collocation method through the use of an asymmetric basis function. The propogation characteristics of these modified schemes have not been previously examined using a Fourier series analysis.

* Research assistant, Dept. of Civil Engineering.
† Professor of Civil Engineering.

0021-9991/81/010046-26$02.00/0

In this paper a Fourier series analysis of the orthogonal collocation method is presented. Finite difference and Galerkin finite element formulations are also analyzed with and without upstream weighting. In addition to information concerning stability and convergence the Fourier series analysis provides quantitative insight into the problems associated with the solution of parabolic partial differential equations with significant first-order spatial derivatives. It also illustrates the role of upstream weighting in each of the methods utilized. The method of analysis presented in this paper is similar to that employed by Gray and Pinder [1] in their investigation of finite difference and Galerkin finite element approximations to the mass transport equation.

Fourier Series Analysis

The Fourier series method of analysis employs a series of trigonometric functions to examine the translational and dissipative characteristics of numerical approximations to partial differential equations.

Analytical Representation

The solution to the analytical operator (Eq. (1)) is assumed to have the following form:

$$c(x, t) = \sum_{n=-\infty}^{\infty} F_n \exp[\hat{\imath}\beta_n t + \hat{\imath}\sigma_n x], \tag{2}$$

where the F_n coefficients are dependent upon the initial conditions imposed on Eq. (1), $\hat{\imath} = (-1)^{1/2}$, and β_n and σ_n are respectively the temporal and spatial frequencies of the nth component of the series. More commonly, σ_n is referred to as the wave number and defined as

$$\sigma_n = \frac{2\pi}{L_n}, \tag{3}$$

where L_n is the wavelength of the nth component of the series. Figure 1 graphically displays two conventional means of depicting the propagation of a given wavelength. Associated points in each diagram are labeled showing the correspondence that exists between each representation.

The linearity of Eq. (1) and the theory of superposition allows the eigenfunction relationship to be determined by an arbitrary term in the solution expansion, i.e.,

$$c \sim c_n = F_n \exp[\hat{\imath}\beta_n t + \hat{\imath}\sigma_n x]. \tag{4}$$

Substitution of Eq. (4) into (1) yields the analytical expression describing the temporal frequency as a function of the wave number or spatial frequency, viz.,

$$\beta_n = \sigma_n(\hat{\imath}D\sigma_n - u). \tag{5}$$

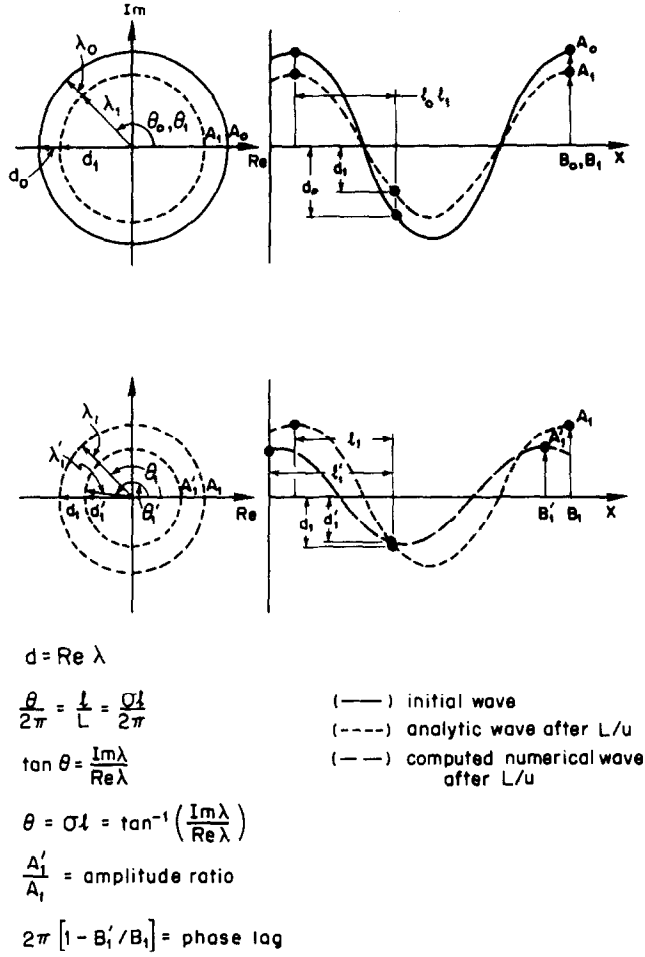

FIG. 1. Diagrammatic representation of a typical waveform in Fourier analysis in real space and in the complex plane; u is the wave celerity.

Therefore, a term in the general solution to Eq. (1) has the form

$$c \sim c_n = F_n \exp[-D\sigma_n^2 t] \exp[i\sigma_n(x - ut)]. \tag{6}$$

The amplitude of the nth component is described by the first two terms and the translation by the final component of Eq. (6).

Numerical Representation of Orthogonal Collocation

A numerical solution may be expressed by a Fourier series expansion in a similar way to that described above. However, instead of considering the analytical operator,

attention is now focused on a discretized form of Eq. (1). The collocation approximation is formulated by using the weighted residual method, viz.,

$$\int_I R \cdot w_i(x)\, dx = 0, \qquad i = 1, 2, \ldots, N, \tag{7}$$

where the weighting function w_i is selected as the Dirac delta function evaluated at the collocation points, i.e.,

$$w_i(x) = \delta(x - x_i) \tag{8}$$

and the residual R is defined as

$$R = L(\hat{c}), \tag{9}$$

where $\hat{c}$ is a trial function. When these collocation points are chosen to be the Gauss points of each spatial increment the method is called orthogonal collocation. Because of the properties of the Dirac delta function, Eq. (7) can be written simply as

$$L(\hat{c})_i = \left[\frac{\partial \hat{c}}{\partial t} + u\frac{\partial \hat{c}}{\partial x} - D\frac{\partial^2 \hat{c}}{\partial x^2}\right]_i = 0, \qquad i = 1, 2, \ldots, N. \tag{10}$$

Modified Hermite cubic polynomials are designated as the basis for the trial function, $\hat{c}$, providing two unknown coefficients (c and $\partial c/\partial x$) at each nodal location; thus

$$\hat{c} = \sum_{j=1}^{M} c_j(t)\, \phi_{0j}(x) + \frac{\partial c_j}{\partial x}(t)\, \phi_{1j}(x), \tag{11}$$

where

$$\phi_{0j} = \tfrac{1}{4}(\eta + \eta_j)^2(\eta - 2\eta_j) - \alpha_0 \eta_j(\eta^2 - 1)^2, \tag{12a}$$

$$\phi_{1j} = \frac{\Delta x}{8}(\eta + \eta_j)^2(\eta - \eta_j) - \alpha_1 \eta_j \Delta x(\eta^2 - 1)^2 \tag{12b}$$

with

$$\eta_j = \pm 1, \qquad \Delta x = x_2^e - x_1^e, \qquad \text{and} \qquad \eta = 2\left(\frac{x - x_1^e}{\Delta x}\right) - 1.$$

The coefficients α_0 and α_1 generate asymmetry in the basis functions. Although this method of applying the asymmetric basis function to each term in the differential operator is most consistent with the method of weighted residuals, it is shown in subsequent sections that such an approach leads to a numerical scheme which diverges from analytical behavior. A further modification of the collocation scheme will be discussed which eliminates this problem and allows the asymmetry in the basis function to alleviate some of the difficulties associated with the numerical solution to convection dominated forms of Eq. (1).

Combination of Eqs. (7) through (11) yields the following set of ordinary differential equations:

$$\sum_{j=1}^{M} \left\{ \frac{dc_j}{dt}\phi_{0j} + \frac{d}{dt}\left(\frac{\partial c_j}{\partial x}\right)\phi_{1j} + u\left(c_j\frac{d\phi_{0j}}{dx} + \frac{\partial c_j}{\partial x}\frac{d\phi_{1j}}{dx}\right) - D\left(c_j\frac{d^2\phi_{0j}}{dx^2} + \frac{\partial c_j}{\partial x}\frac{d^2\phi_{1j}}{dx^2}\right)\right\}\Bigg|_{x_i} = 0, \qquad i = 1, 2,..., N. \tag{13}$$

A finite difference approximation in time may be used to arrive at the final form of the approximation,

$$\sum_{j=1}^{M} \left\{ \left(\frac{c_j^{t+\Delta t} - c_j^t}{\Delta t}\right)\phi_{0j} + \left(\frac{\partial c_j^{t+\Delta t}/\partial x - \partial c_j^t/\partial x}{\Delta t}\right)\phi_{1j} + \gamma\left[u\left(c_j^{t+\Delta t}\frac{d\phi_{0j}}{dx} + \frac{\partial c_j^{t+\Delta t}}{\partial x}\frac{d\phi_{1j}}{dx}\right) - D\left(c_j^{t+\Delta t}\frac{d^2\phi_{0j}}{dx^2} + \frac{\partial c_j^{t+\Delta t}}{\partial x}\frac{d^2\phi_{1j}}{dx^2}\right)\right] + (1-\gamma)\left[u\left(c_j^t\frac{d\phi_{0j}}{dx} + \frac{\partial c_j^t}{\partial x}\frac{d\phi_{1j}}{dx}\right) - D\left(c_j^t\frac{d^2\phi_{0j}}{dx^2} + \frac{\partial c_j^t}{\partial x}\frac{d^2\phi_{1j}}{dx^2}\right)\right]\right\}\Bigg|_{x_i} = 0, \qquad i = 1, 2,..., N. \tag{14}$$

The parameter γ is used to position the spatial derivative in the time domain, e.g., when $\gamma = 1$, a backward difference scheme is generated.

It is now assumed that the solution to (14) can be written in a form analogous to the analytical case, Eq. (4):

$$c_j(t) \sim \sum_{n=-p}^{p} F_n \exp[\hat{i}\beta_n' t + \hat{i}\sigma_n(j\Delta x)] \tag{15}$$

and

$$\frac{\partial c_j}{\partial x}(t) \sim \sum_{n=-p}^{p} S_n \exp[\hat{i}\beta_n' t + \hat{i}\sigma_n(j\Delta x)], \tag{16}$$

where β_n' is the temporal frequency of the numerical scheme and p corresponds to the integer number associated with the smallest wavelength that can be propagated by the numerical approximation. The numerical representation having a finite number of Fourier series components is a result of the spatial discretization in the numerical scheme. Thus the range of spatial frequencies is limited by the size of Δx. If one is to consider only the nodal values c_j, wavelengths smaller than $2\Delta x$ cannot be resolved. Notice in Fig. 2a that the $1\Delta x$ wavelength cannot be defined. However, in dealing with the Hermite basis function and interpolating the solution between nodes a cubic polynomial results which allows the resolution of a $1\Delta x$ wavelength. Figure 2b gives a graphical representation of this phenomena.

Further examination of Eqs. (15) and (16) indicates that the numerical

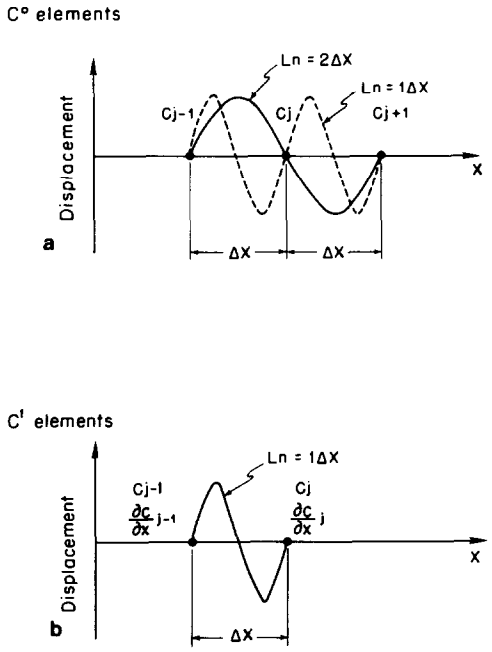

FIG. 2. (a) Resolution of minimum wavelengths for C^0 continuous elements. (b) Resolution of minimum wavelengths for C^1 continuous elements.

representation of the function and its slope are independent in the Fourier series analysis; that is to say, the slope is not merely the spatial differentiation of Eq. (15). This is a result of Hermite cubic polynomials being chosen as the basis for the numerical solution. Values of c and $\partial c/\partial x$ are generated from Eq. (14) at given nodal locations with no correlation being implied between the two. Furthermore, it is assumed in the series representations given above that the behavior of the temporal frequency, β'_n, for c_j and $\partial c_j/\partial x$ is the same. If one was to examine through a Fourier series analysis the accuracy of the spatial discretization only, thus dealing with Eq. (13), this approximation would be exact. Introducing discrete values of c and $\partial c/\partial x$ in time causes an associated error which is taken to be no larger than the error corresponding to the approximation of the temporal derivative.

As in the case of the analytical operator, it is necessary to consider only one arbitrary term in each of the above expressions, viz.,

$$c_j \sim c_{j_n} = F_n \exp[\hat{\imath}\beta'_n t + \hat{\imath}\sigma_n(j\Delta x)], \tag{17}$$

$$\frac{\partial c_j}{\partial x} \sim \frac{\partial c}{\partial x} j_n = S_n \exp[\hat{\imath}\beta'_n t + \hat{\imath}\sigma_n(j\Delta x)]. \tag{18}$$

Substituting these expressions into the collocation equations for a typical spatial increment and solving for the temporal frequency of the numerical solution one obtains

$$\beta_n' = \beta_n'(\mathscr{D}, \mathscr{U}, \sigma_n, \gamma, \alpha_0, \alpha_1), \tag{19}$$

where

$$\mathscr{D} = \frac{D\,\Delta t}{\Delta x^2} \qquad \text{and} \qquad \mathscr{U} = \frac{u\,\Delta t}{\Delta x}.$$

The functional form of Eqs. (19) and (5) are not the same due to the discretization in time and space of the numerical scheme. Therefore, differences in the dissipative and translational properties of the components of the numerical and analytical solutions should be expected (see Fig. 1).

Comparison of Numerical and Analytical Behavior

Similarities in the series expansions for the analytical and numerical solutions are now considered. The coefficients describing the initial conditions for both the numerical and analytical expressions are the same; therefore, at $t = 0$ the analytical solution is exactly represented to the degree of accuracy provided by the truncated numerical series. However, after an elapsed time of Δt, typical terms in the analytical and numerical solutions have the following form:

$$c^{t+\Delta t} \sim F_n \exp[\hat{\imath}\beta_n(t + \Delta t) + \hat{\imath}\sigma_n x] = F_n \exp[\hat{\imath}\beta_n t + \hat{\imath}\sigma_n x] \exp[\hat{\imath}\beta_n \Delta t]$$
$$= c^t \exp[\hat{\imath}\beta_n \Delta t] = c^t \lambda_n, \tag{20}$$

$$c_j^{t+\Delta t} \sim F_n \exp[\hat{\imath}\beta_n'(t + \Delta t) + \hat{\imath}\sigma_n(j\,\Delta x)] = c_j^t \exp[\hat{\imath}\beta_n' \Delta t] = c_j^t \lambda_n', \tag{21}$$

where λ_n and λ_n' are designated as analytical and numerical eigenvalues, respectively. The magnitude of the eigenvalue can be interpreted as the ratio of the amplitude of the nth Fourier component after an elapsed time of Δt to its amplitude at the beginning of the time step. After the nth component has propagated one complete wavelength this dissipative ratio takes the following form:

$$|\exp(\hat{\imath}\beta_n' N_n \Delta t)| = |\exp(\hat{\imath}\beta_n' \Delta t)|^{N_n} = |\lambda_n'|^{N_n}, \tag{22}$$

where N_n is the number of time steps necessary to propagate the nth component through one wavelength,

$$N_n = \frac{L_n}{u\,\Delta t} = \frac{L_n}{\Delta x \mathscr{U}}. \tag{23}$$

In view of Eqs. (21) and (22) the stability of the numerical scheme is guaranteed when the magnitude of λ_n' is bounded from above by unity for all n.

A corresponding analytical representation is obtained through the combination of Eq. (5) with the definition of the analytical eigenvalue (Eq. (20)). Thus after the propagation of one wavelength we have

$$|\exp(i\beta_n N_n \Delta t)| = |\exp(-D\sigma_n^2 N_n \Delta t)\exp(-i\sigma_n u N_n \Delta t)|. \tag{24}$$

Because the magnitude of $\exp(-i\sigma_n u N_n \Delta t)$ is unity Eq. (24) reduces to

$$|\exp(i\beta_n N_n \Delta t)| = (\exp(-D\sigma_n^2 \Delta t))^{N_n} = |\lambda_n|^{N_n} \tag{25}$$

which upon substituting the definition for the wave number simplifies to

$$|\lambda_n|^{N_n} = \left(\exp\left(-4\pi^2 \mathscr{D}\left(\frac{\Delta x}{L_n}\right)^2\right)\right)^{N_n}. \tag{26}$$

A comparison of the dissipative character in the analytical and numerical solutions can be considered as one measure of the accuracy of the numerical approximation. Hence, the amplitude ratio is defined as

$$R_n = \left(\frac{|\lambda_n'|}{|\lambda_n|}\right)^{N_n} = \left(\frac{|\lambda_n'|}{\exp(-4\pi^2 \mathscr{D}(\Delta x/L_n)^2)}\right)^{N_n}. \tag{27}$$

If the nth component of the numerical series solution exactly represents the analytic behavior, the above ratio would have a value of unity. Values less than one indicate that after the propagation of one wavelength the amplitude of the numerical wave is less than anticipated by the analytical results. The opposite is implied for values greater than one.

The translational property of the Fourier series component is characterized by the phase angle which may be expressed in terms of the real and imaginary parts of the eigenvalue. The nth component of the numerical Fourier series expansion has a phase angle of θ_n' after one time step has elapsed, where

$$\theta_n' = \tan^{-1}\left(\frac{I_m \lambda_n'}{R_e \lambda_n'}\right). \tag{28}$$

Figure 1 gives a graphical representation of this definition. After this component has propagated through one wavelength the value of the numerical phase angle is $N_n\theta_n'$. However, after the propagation of one wavelength analytically, the Fourier component is translated exactly through 2π radians. The difference between these two values constitutes a phase lag or lead and thus a second measure of the relative accuracy of the numerical scheme,

$$\Xi_n = \theta_n' N_n - 2\pi. \tag{29}$$

Analysis of Orthogonal Collocation

The stability of the classical orthogonal collocation method can be demonstrated for several values of $\mathscr{D}$ using a plot of the magniudes of the numerical eigenvalues (see Fig. 3). In this particular case a time weighting coefficient of $\gamma = 0.5$ and a dimensionless velocity value of $\mathscr{U} = 0.738$ is used. This scheme is shown to be stable for all values of Δx because movement along the abscissa can be interpreted in two ways. Larger values of $L_n/\Delta x$ are generated by either

(1) considering an entire spectrum of values for the physical wavelength (L_n) and one value of Δx, or

(2) decreasing the values of Δx for a given L_n with the understanding that $\mathscr{D}$ and $\mathscr{U}$ remain constant.

The latter implies that as Δx decreases, corresponding changes in Δt, D, or u must also be considered to accommodate unchanging values of the dimensionless parameters $\mathscr{D}$ and $\mathscr{U}$.

The amplitude ratio and phase lag plots for this formulation are displayed in Figs. 4 and 5, respectively. The values of γ and $\mathscr{U}$ previously mentioned are employed for various values of $\mathscr{D}$. Clearly, as the convective term becomes dominant, depicted by decreasing values of $\mathscr{D}$, the behavior of the amplitude ratio over the entire spectrum of wavelengths more closely reflects the analytical properties. However, the opposite is true with respect to the translational properties as illustrated in the phase-lag diagram. Furthermore, both properties display convergence to analytical behavior for large values of $L_n/\Delta x$.

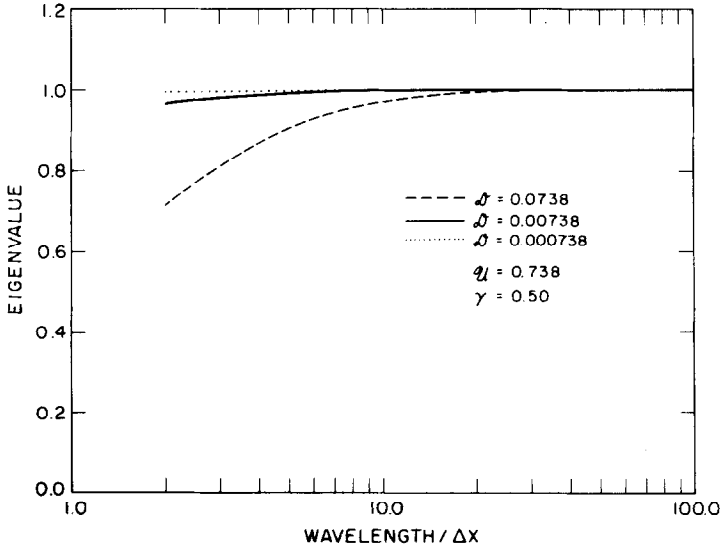

FIG. 3. Magnitude of numerical eigenvalues for orthogonal collocation (standard formulation).

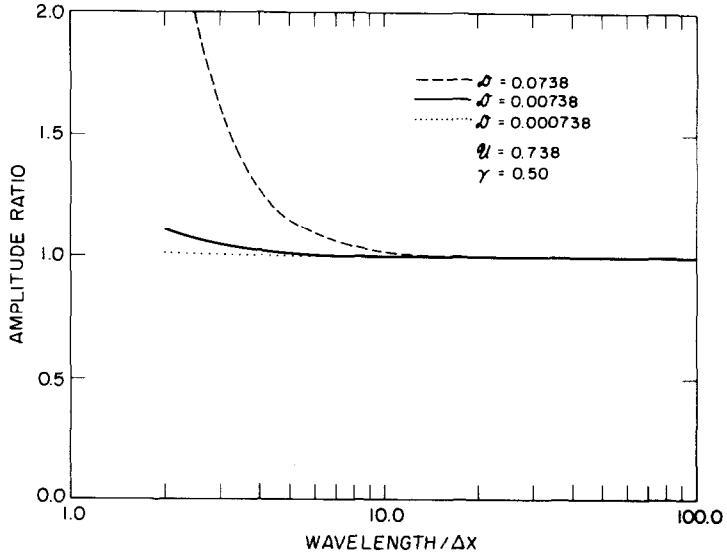

FIG. 4. Amplitude ratio for orthogonal collocation (standard formulation).

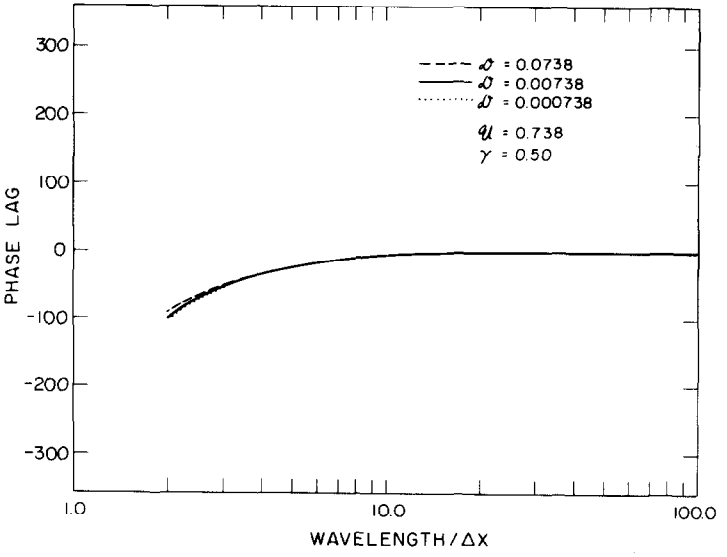

FIG. 5. Phase lag for orthogonal collocation (standard formulation).

To determine the impact of phase lag and amplitude modification on the numerical solution the series approximation to Eq. (15) is evaluated using appropriate initial conditions and numerical eigenvalues. Because the most difficult test of a numerical scheme is the propagation of an initially sharp front we use this initial condition which can be written (Gray and Pinder [1])

$$c(x) = 1 - H(x), \qquad t = 0, \tag{30}$$

where $H(x)$ is the unit step function. The F_n coefficients in (15) are therefore defined by

$$F_n = \frac{\hat{t}}{2n\pi}\left[1 - e^{i n\pi}\right]. \tag{31}$$

The solutions obtained for $\gamma = 0.5$, $\mathscr{U} = 0.738$, and several values of $\mathscr{D}$ are presented in Fig. 6. It is apparent that as the convective term becomes dominant the solution is characterized by large oscillations. Gray and Pinder [1] recognized this phenomena to be the result of the numerical dispersion of smaller wavelength harmonics.

Returning to Fig. 4 one notes that the greatest deviation from the analytical behavior is associated with the solution which displays the least oscillations ($\mathscr{D} = 0.0738$ in Fig. 6). To understand this apparent contradictory behavior it is necessary to examine individual components of the Fourier series. This information is presented in Fig. 7 for the Fourier component corresponding to the $4\Delta x$ wavelength.

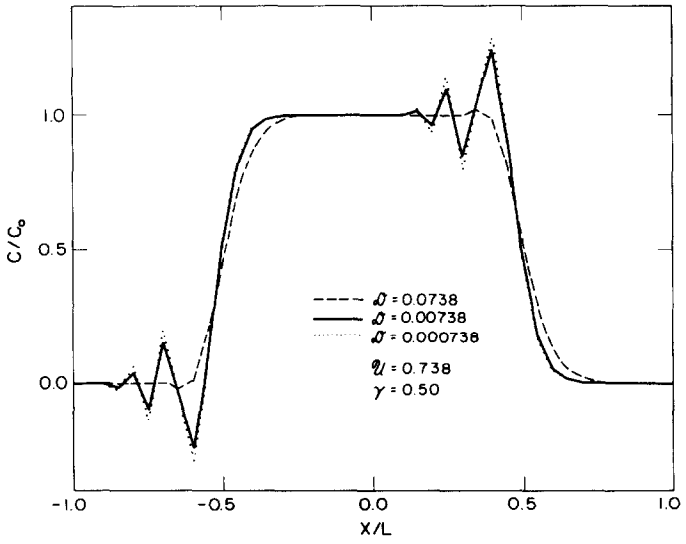

FIG. 6. Numerical solution to a parabolic partial differential equation formulated using numerical eigenvalues.

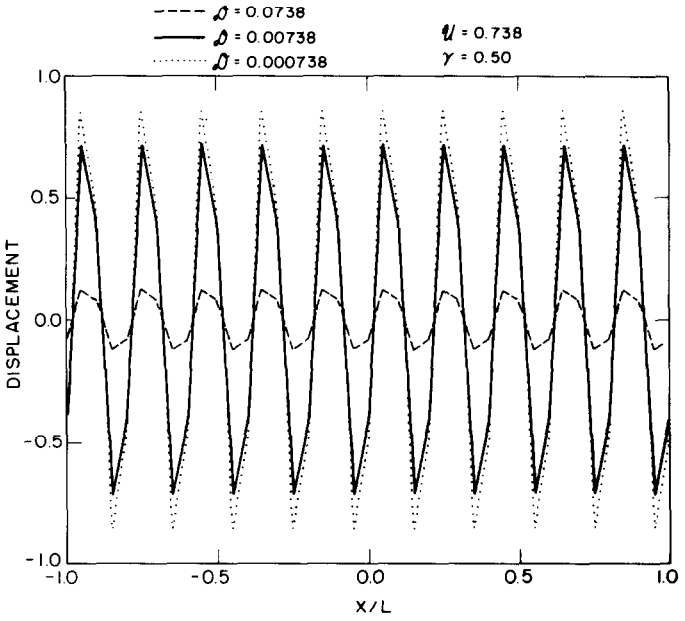

FIG. 7. Contribution of a small wavelength Fourier component in the numerical solution of a parabolic partial differential equation.

Here the waveform associated with each of the solutions observed in Fig. 6 is plotted. Clearly, in convection dominated equations the smaller wavelengths play a more important role in the definition of the solution. Figures of this type will be considered later when various numerical schemes are examined.

While the standard collocation method is stable and convergent, it fails to provide acceptable engineering solutions to parabolic partial differential equations with significant first-order terms. To alleviate these difficulties, Pinder and Shapiro [3] formulated a collocation scheme using asymmetric basis functions. Initially it appeared most consistent to apply the asymmetric basis function to each term in the governing equation. For the temporal and spatial increments considered there was an excellent match between this formulation and the analytical solution. The undesirable oscillations were removed with only a modest smearing of the solution profile. A Fourier series analysis of this scheme shows that it satisfies the stability criteria (see Fig. 8). However, examination of Figs. 9 and 10 demonstrates that amplitude ratio and phase-lag behavior are divergent for this modification in the collocation method. That is, the analytical and numerical properties diverge as $L_n/\Delta x$ increases.

The dashed curves in Figs. 8–10 are results from a second collocation scheme proposed as a means of effectively handling Eq. (1) when there is a significant first-order spatial term. In this formulation the asymmetric basis function is upwinded

convection is not a symmetric process. Figures 9 and 10 show this scheme to be

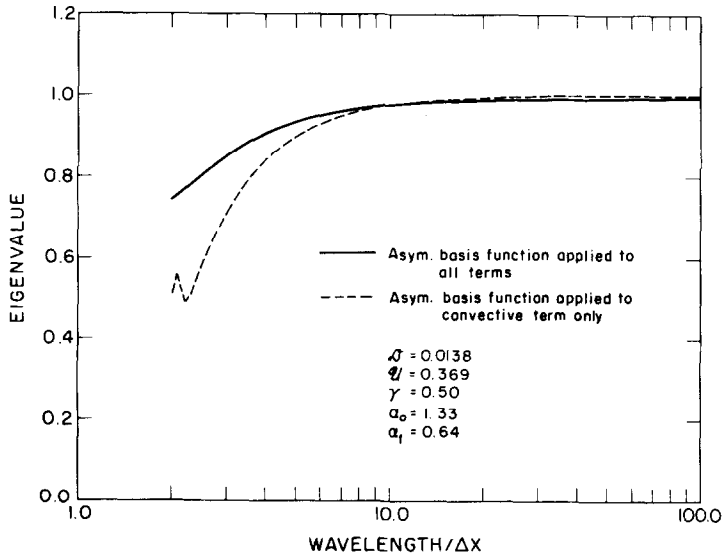

FIG. 8. Magnitude of numerical eigenvalues for two upstream weighted orthogonal collocation schemes.

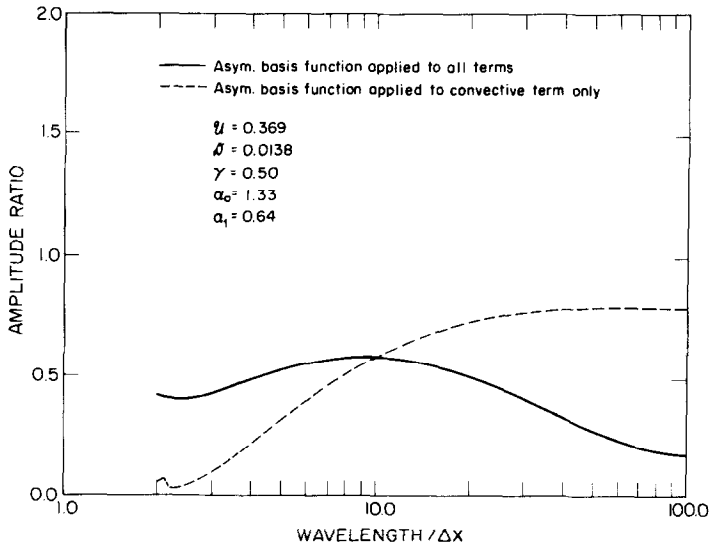

FIG. 9. Amplitude ratio for two upstream weighted orthogonal collocation schemes.

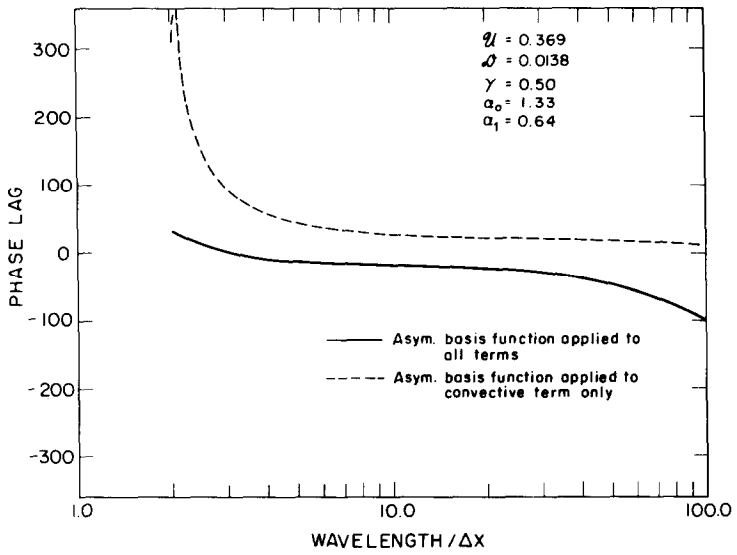

FIG. 10. Phase lag for two upstream weighted orthogonal collocation schemes.

convergent in comparison to the prior formulation when the same value of the weighting parameters (α_0 and α_1) are employed. Furthermore, all combinations of $\mathscr{D}$, $\mathscr{U}$, α_0, and α_1 proved to satisfy convergence and stability for this upstream weighted collocation technique. As a result of changing the formulation of the numerical scheme, the values of the weighting coefficients necessary to provide a satisfactory comparison with the analytical solution also change. Figure 11 shows that this scheme removes oscillations at the expense of smearing the solution profile.

Analysis of Upstream Weighting in Several Numerical Schemes

The behavior of standard finite element, finite difference, and collocation methods in the solution of parobolic partial differential equations with significant first-order terms is displayed in Fig. 12. The parameters employed in each of the numerical techniques are $\gamma = 0.5$, $\mathscr{D} = 0.000738$, and $\mathscr{U} = 0.738$. These solutions are clearly unacceptable and upstream weighting schemes have been developed to damp the spurious oscillations. At this point a Fourier analysis is employed to establish the impact on the numerical solution for each of the upstream weighting techniques, the results of which are displayed in Fig. 13.

Orthogonal Collocation

Standard and optimal upstream weighted (second formulation above) collocation solutions are compared in Figs. 14–18. The stability of both the standard and modified schemes is recognized in the eigenvalue plot (see Fig. 14). The amplitude

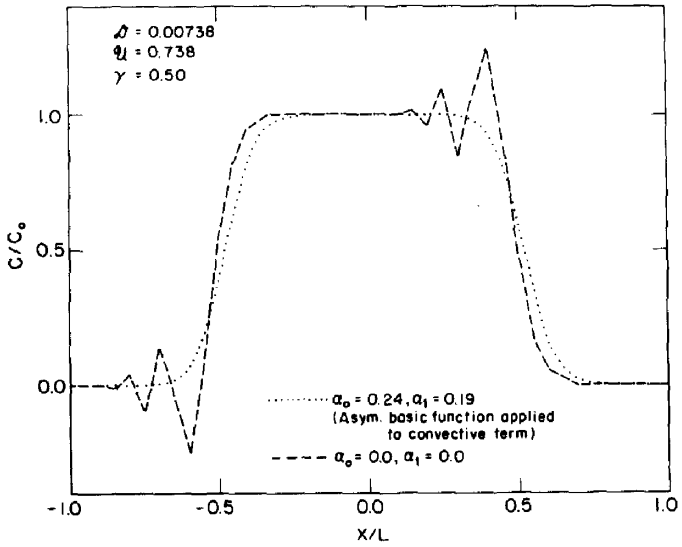

FIG. 11. Solution to a parabolic partial differential equation formulated using numerical eigenvalues from an orthogonal collocation scheme.

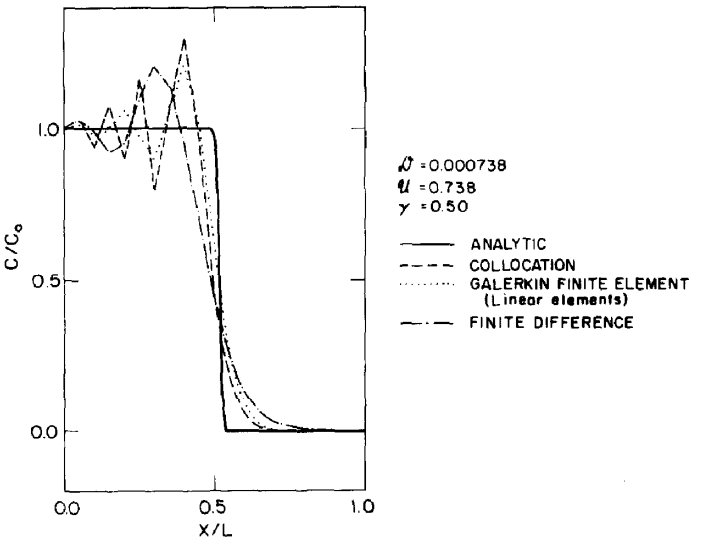

FIG. 12. Solution to a convection dominated parabolic partial differential equation using several standard numerical schemes.

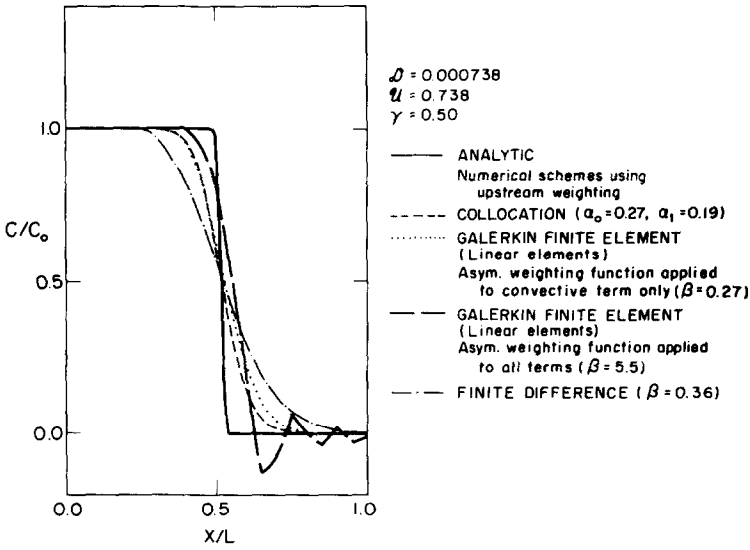

FIG. 13. Solution to a convection dominated parabolic partial differential equation using several upstream weighted numerical techniques.

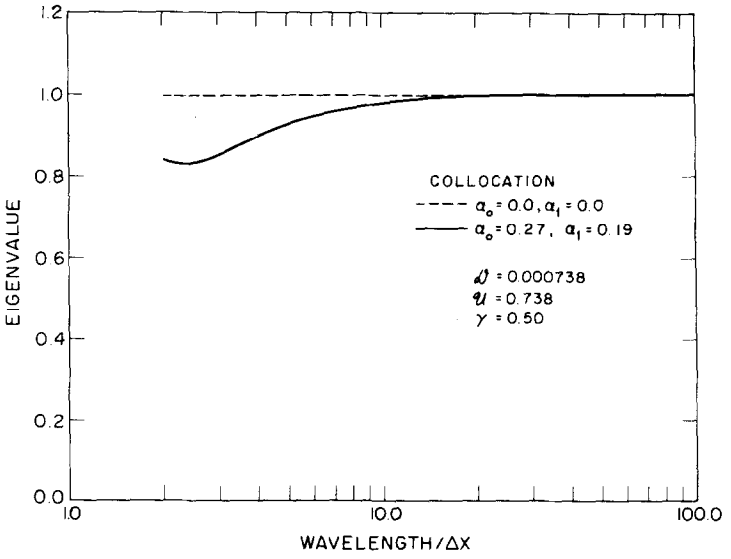

FIG. 14. Magnitude of numerical eigenvalues for upstream weighted and standard collocation schemes.

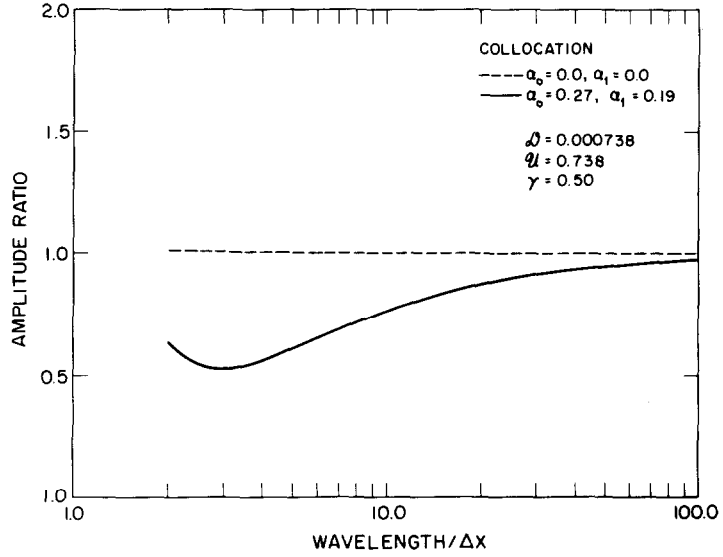

FIG. 15. Amplitude ratio for upstream weighted and standard collocation schemes.

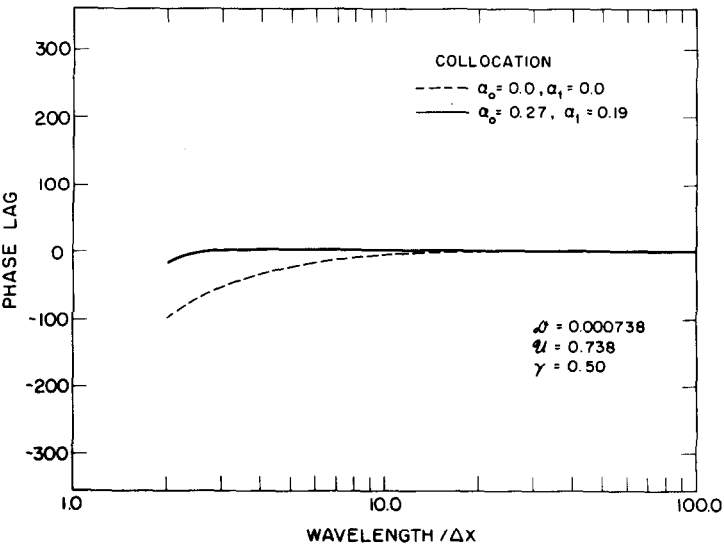

FIG. 16. Phase lag for upstream weighted and standard collocation schemes.

ratio and phase-lag diagrams, Figs. 15 and 16, respectively, explain the oscillations present in the standard formulation. The amplitude of the Fourier components of the numerical scheme exactly represents the analytical behavior. However, as a result of the small wavelengths being out of phase, numerical difficulties in the standard formulation arise. Notice also that the phase lag is associated with oscillations preceding the sharp front as displayed in Fig. 12. The translational properties of the numerical scheme are, therefore, slower than the predicted analytical behavior. Modifying the collocation scheme through the use of an asymmetric basis function in the convective term alters two characteristics of the numerical scheme,

(1) elimination of the phase lag associated with the small wavelengths,

(2) a degree of damping over the entire range of wavelengths.

In Fig. 17 a comparison of the contribution of the $4\Delta x$ wavelength to the series solution for the anlytical, standard, and modified schemes is made. The amplitude of the wave associated with the standard scheme is exactly that of the analytic. However, a phase lag is shown to exist. Modifying the scheme removes the phase lag at the cost of altering the amplitude. A similar comparison can be made with respect to a larger wavelength, as displayed in Fig. 18 for the $10\Delta x$ wave. The phase properties of the standard and modified schemes closely represent the analytical behavior. However, as forecast in Figs. 15 and 16 the amplitude of the modified

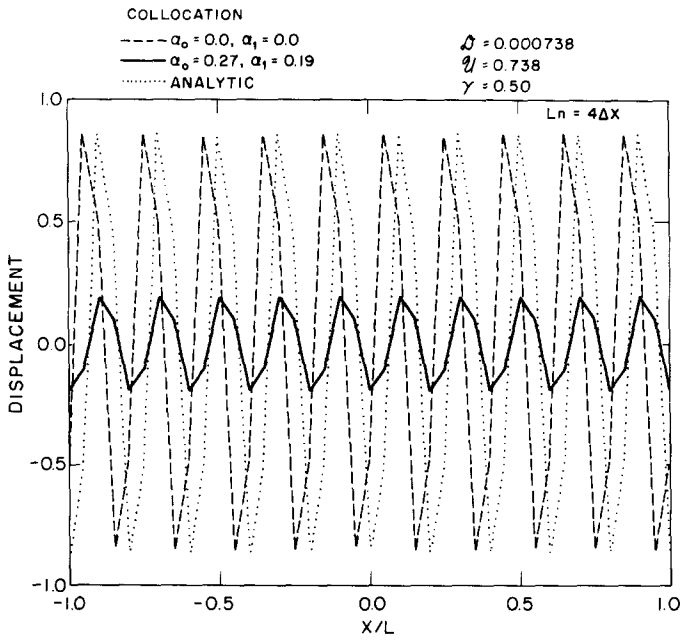

FIG. 17. Contribution of the $4\Delta x$ wave in the orthogonal collocation solution of a convection dominated parabolic partial differential equation.

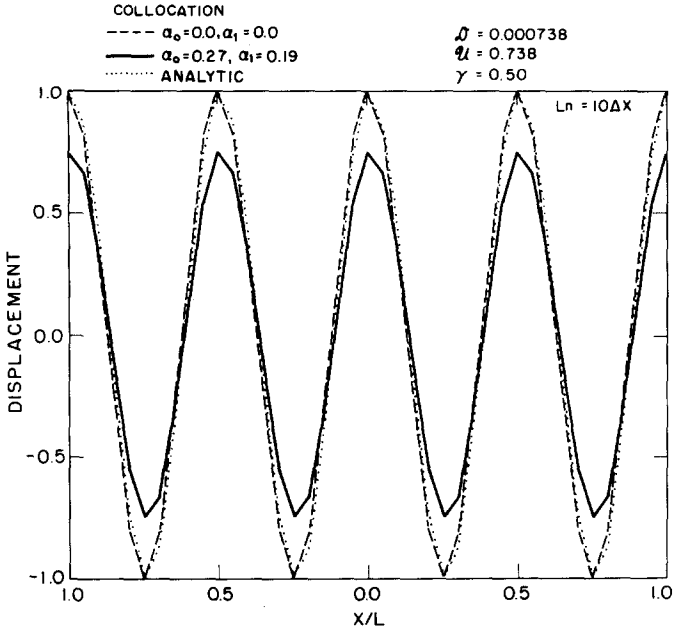

FIG. 18. Contribution of the $10\,\Delta x$ wave in the orthogonal collocation solution of a convection dominated parabolic partial differential equation.

scheme is again reduced. It is the damping of the amplitude which causes the characteristic smearing of the sharp front normally associated with upstream weighting techniques. This phenomenon gives the appearance of additional dispersion in the physical problem.

Finite Difference

An upstream weighted finite difference scheme is formulated using a weighted average of forward and backward difference approximations to the first-order spatial derivative, i.e.,

$$\frac{\partial c_i}{\partial x} = \left(\frac{1}{2} - \beta\right)\left(\frac{c_{i+1} - c_i}{\Delta x}\right) + \left(\frac{1}{2} + \beta\right)\left(\frac{c_i - c_{i-1}}{\Delta x}\right). \tag{32}$$

As in the orthogonal collocation scheme, the degree of weighting is dependent upon the magnitude of the convective and dispersive terms. The value $\beta = 0.36$ provided a best fit with the analytical solution subject to the constraint of no oscillations. The stability of both the standard and modified schemes is apparent from Fig. 19. Examination of the amplitude ratio and phase diagrams (Figs. 20 and 21) once again reveals the cause of the oscillations in the standard finite difference scheme. The small wavelengths which play a significant role in the definition of the solution

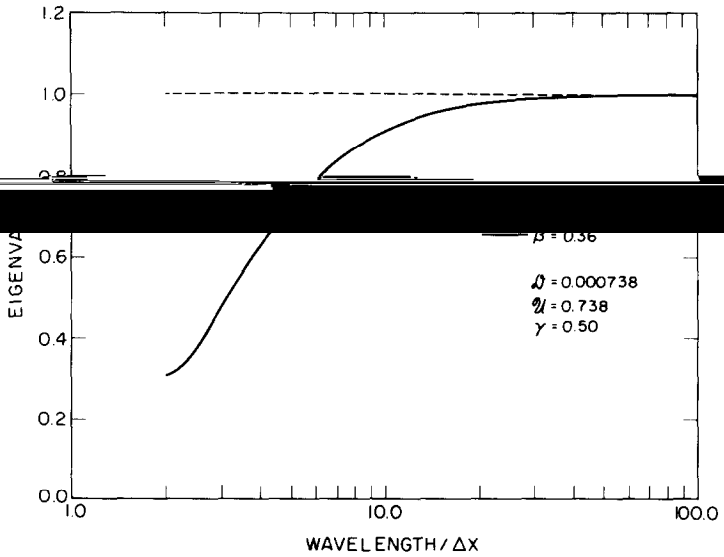

FIG. 19. Magnitude of numerical eigenvalues for standard and upstream weighted finite difference schemes.

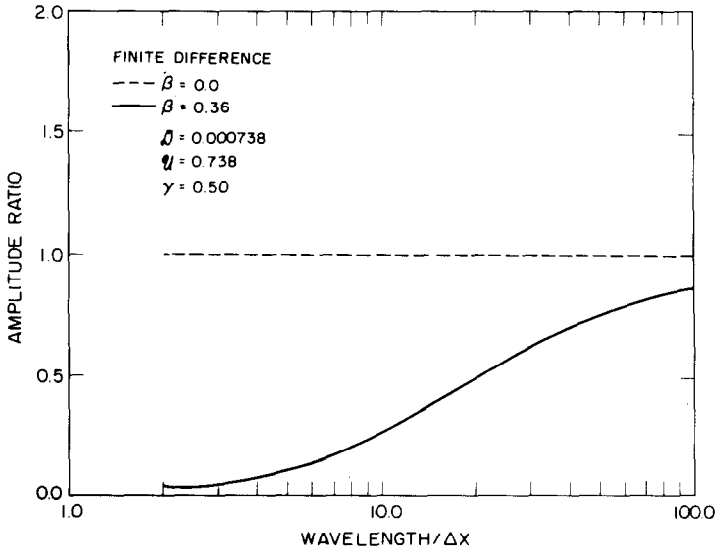

FIG. 20. Amplitude ratio for standard and upstream weighted finite difference schemes.

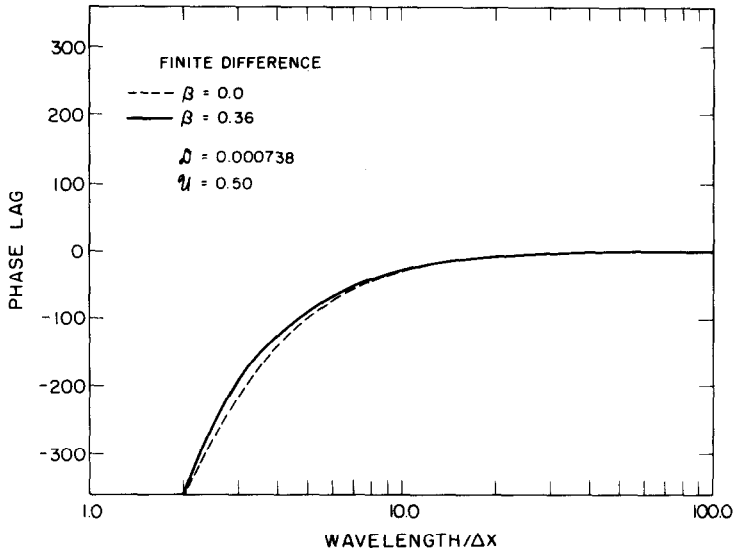

FIG. 21. Phase lag for standard and upstream weighted finite difference schemes.

exhibit a phase lag. Modifying the scheme does not adjust this error. Therefore, upstream weighting in the finite difference formulation merely damps those waves which are out of phase in order to achieve a smooth solution. Figures 20 and 21 also demonstrate that the modified scheme is convergent.

Finite Element

Two upstream weighted Galerkin finite element schemes are analyzed. Each is formulated using an asymmetric weighting function. This is in contrast to the collocation scheme which employs an asymmetric basis function. The modified weighting function was developed by Huyakorn [2] through the addition of a quadratic term to a linear chapeau function. The most consistent approach is to apply the weighting function to each term in the residual, viz.,

$$\int_l R \cdot w_i(x)\,dx = \int_l \left(\frac{\partial \hat{c}}{\partial t} + u\frac{\partial \hat{c}}{\partial x} - D\frac{\partial^2 \hat{c}}{\partial x^2}\right) \bar{\phi}_i(x)\,dx = 0, \qquad i = 1, 2, \ldots, N, \quad (33)$$

where

$$\hat{c} = \sum_{j=1}^{N} c_j(t)\,\phi_j(x)$$

and

$$\phi_j = \begin{cases} (x - x_{j-1})/\Delta x, & x_{j-1} \leqslant x \leqslant x_j, \\ (x_{j+1} - x)/\Delta x, & x_j \leqslant x \leqslant x_{j+1}. \end{cases} \quad (34)$$

The modified weighting function is defined as

$$\bar{\phi}_i = \begin{cases} \dfrac{(x - x_{i-1})}{\Delta x} - \dfrac{3x}{\Delta x^2}(x - \Delta x)\beta, & x_{i-1} \leqslant x \leqslant x_i, \\ \dfrac{(x_{i+1} - x)}{\Delta x} + \dfrac{3x}{\Delta x^2}(x - \Delta x)\beta, & x_i \leqslant x \leqslant x_{i+1}, \end{cases} \tag{35}$$

where the value of β is dependent on the magnitude of the convective and dispersive terms. A second approach is also analyzed. This involves associating the modified weighting function with only the convective term

$$\int_I R \cdot w_i(x)\, dx = \int_I \left(\frac{\partial \hat{c}}{\partial t}\phi_i + u\frac{\partial \hat{c}}{\partial x}\bar{\phi}_i - D\frac{\partial^2 \hat{c}}{\partial x^2}\phi_i \right) dx = 0, \qquad i = 1, 2, \ldots, N. \tag{36}$$

The stability of the standard formulation and the two modified schemes is displayed in Fig. 22. The amplitude ratio and phase diagrams for the three finite element schemes are shown in Figs. 23 and 24, respectively. As in the previous numerical techniques, the difficulties associated with the standard formulation are due to the small wavelengths being out of phase and their amplitude not being sufficiently damped.

Upstream weighting plays a different role in each of the modified finite element techniques. When the asymmetric weighting function is applied to all terms in the residual the method fails to give an acceptable solution as shown in Fig. 13. A value

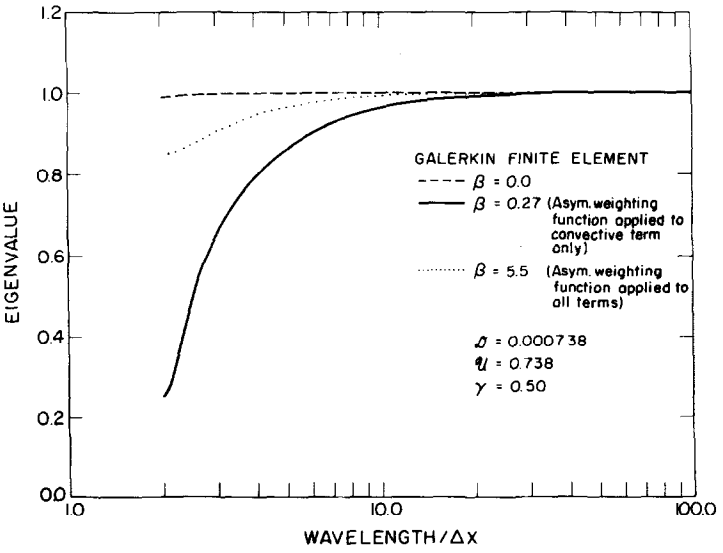

FIG. 22. Magnitude of numerical eigenvalues for standard and upstream weighted finite element schemes.

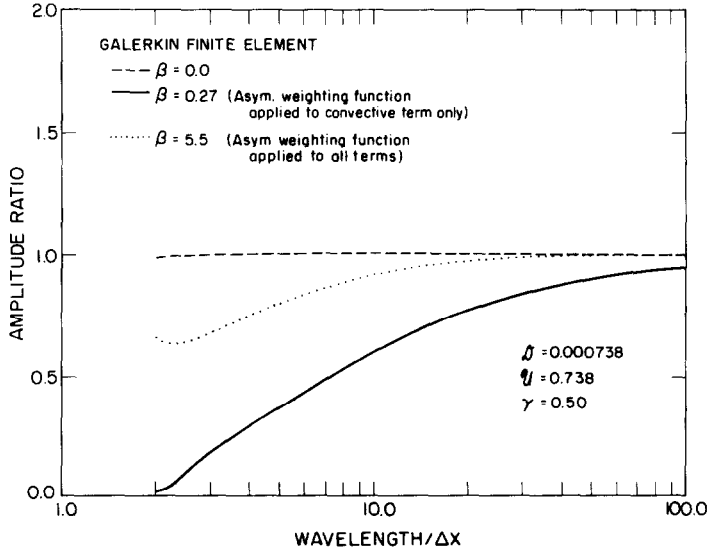

FIG. 23. Amplitude ratio for standard and upstream weighted finite element schemes.

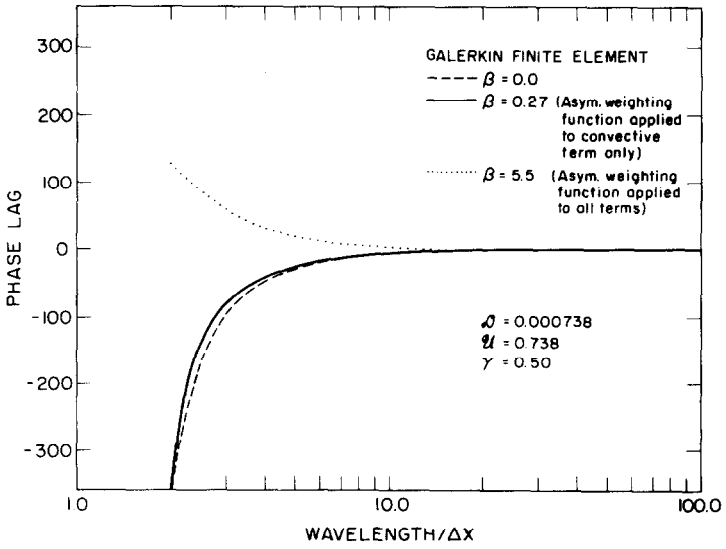

FIG. 24. Phase lag for standard and upstream weighted finite element schemes.

of $\beta = 5.5$ is necesssary to remove the oscillations preceding the sharp front. However, in doing this, oscillations are generated on the downstream side of the solution profile. This is predicted by the phase diagram where phase accelerations are forecast for small wavelengths. The fact that this method fails to damp out these wavelengths results in their appearance as oscillations. Even though unacceptable results are obtained, this numerical scheme is convergent for the parameters used.
The second upstream weighted finite element scheme behaves in a similar manner to that of the finite difference scheme previously described. The phase lag associated with the small wavelengths of the standard formulations is not altered. However, these wavelengths are damped, thereby generating a smooth solution. A value of $\beta = 0.27$ was necessary to obtain this result.

Conclusions

(1) A Fourier series analysis of modified numerical schemes is one method of analyzing stability and convergence. The Fourier series analysis also provides information which gives insight into the numerical difficulties of a given numerical scheme.

(2) The modified orthogonal collocation scheme using asymmetric basis functions alleviates numerical difficulties associated with parabolic partial differential equations having significant first-order terms. Oscillations are removed at the cost of moderately smearing the solution profile. Other modified numerical schemes provide similar results.

(3) The role upstream weighting plays in alleviating numerical difficulties is different for each numerical scheme. The modified collocation method provides extremely accurate translational properties as a means of eliminating numerical oscillations. Upstream weighted finite difference and finite element schemes rely primarily on damping the amplitude of the small wavelength harmonics which are out of phase. This is equivalent to adding additional dispersion to the physical problem.

APPENDIX: Notation

c	analytical solution to differential operator
$\hat{c}$	numerical approximation to the solution of differential operator
D	dispersion coefficient
$\mathscr{D}$	dimensionless dispersion
F_n	Fourier coefficient representing the initial conditions for the function c
$\hat{\imath}$	$=(-1)^{1/2}$
I_m	imaginary part of a complex number
L	differential operator
L_n	wavelength

N_n number of time steps necessary to propagate a Fourier component through one wavelength
R residual of the differential operator
R_e real part of a complex number
S_n Fourier coefficient representing the initial conditions for the slope of the function c
t time coordinate
u velocity
$\mathscr{U}$ dimensionless velocity
w_i weighting function in method of weighted residuals
x spatial coordinate
α_0, α_1 upstream weighting coefficients for the basis functions in the orthogonal collocation scheme
β upstream weighting coefficient for finite element and finite difference schemes
β_n analytical temporal frequency of Fourier component
β'_n numerically generated temporal frequency of Fourier component
γ positions the discretized spatial derivative in the time domain
Δt time increment
Δx spatial increment
δ Dirac delta function
θ'_n numerically generated phase angle of Fourier component after one time increment
λ_n analytical eigenvalue
λ'_n numerically generated eigenvalue
Ξ phase lag between numerical and analytical solutions
σ_n wave number
ϕ_{0j}, ϕ_{1j} basis functions in orthogonal collocation
ϕ_j basis function in finite element scheme
ϕ_i weighting function in finite element scheme

Superscripts

t evaluation at known time
$t + \Delta t$ evaluation at unknown time

Subscripts

j nodal index
i weighting function index

ACKNOWLEDGMENTS

The authors gratefully acknowledge the suggestions and clarifying remarks of Dr. W. G. Gray of Princeton University. This work was supported by the University of California through sub-contract No. P.O. 3143202 under Energy Research and Development Administration Contract No. W-7405-ENG. 48.

REFERENCES

1. W. G. GRAY AND G. F. PINDER, *Water Resources Res.* **12** (1976), 547–555.
2. P. S. HUYAKORN, *Appl. Math. Modelling* **1** (1977), 187–195.
3. G. F. PINDER AND A. SHAPIRO, *Water Resources Res.* **15** (1979), 1177–1182.
4. G. STRANG AND G. F. FIX, "*An Analysis of the Finite Element Method*," Prentice–Hall, Englewood Cliffs, N. J., 1973.