

An implicit finite difference approximation to the one-dimensional transport equation

Frank G. Polanco* Robert L. May†

(Received 7 August 2000)

Abstract

The properties of an implicit three-level finite difference scheme are investigated. The modified equivalent partial differential equation is used to determine the speed and amplitude characteristics of the scheme, and these are used to determine the optimal value of the

* Aeronautical and Maritime Research Laboratory, GPO Box 4331, Melbourne, Victoria 3000, AUSTRALIA. <mailto:Frank.Polanco@dsto.defence.gov.au>

† Department of Mathematics, RMIT University, GPO Box 2476V, Melbourne, Victoria 3000, AUSTRALIA. <mailto:rob@rmit.edu.au>

⁰See <http://anziamj.austms.org.au/V42/CTAC99/Pola> for this article and ancillary services, © Austral. Mathematical Soc. 2000. Published 27 Nov 2000.

weighting between a central difference and an upwind biased difference. Results of numerical experiments that confirm that the predicted value of the weight does minimise the error are presented.

Contents

1	Introduction	C1180
2	Finite Difference Approximation	C1181
3	Modified Equivalent PDE	C1183
4	Numerical Experiments	C1187
5	Conclusion	C1197
	References	C1198

1 Introduction

Convection and diffusion occur simultaneously in many physical situations, so it is not surprising that the accurate modelling of these processes has been the subject of much research over the last few decades [1]. We consider

a finite difference (FD) approximation to the one-dimensional transport (or convection-diffusion) equation

$$\frac{\partial T}{\partial t} + \mathcal{C} \frac{\partial T}{\partial x} = \mathcal{D} \frac{\partial^2 T}{\partial x^2}, \quad (1)$$

where t is time, x is space, \mathcal{C} and \mathcal{D} are the convective velocity and the diffusive coefficient respectively, and T is some scalar quantity (temperature for example). We restrict the convective velocity to be non-negative ($\mathcal{C} \geq 0$), and obviously the diffusion coefficient is also non-negative ($\mathcal{D} \geq 0$).

2 Finite Difference Approximation

The FD scheme is implicit and second order in space and time; the diffusive term is approximated by a central difference and the convective term by a linear combination between a central and an upwind biased difference. A three-level approximation employing a variable time step is used for the temporal derivative. The FD approximation is

$$\left[\frac{(1 + 2\kappa)T_j^{(n)} - (1 + \kappa)^2 T_j^{(n-1)} + \kappa^2 T_j^{(n-2)}}{(1 + \kappa)\Delta t^{(n)}} \right] + \mathcal{C} \left[(1 - \omega) \left(\frac{T_{j+1}^{(n)} - T_{j-1}^{(n)}}{2\Delta x} \right) + \omega \left(\frac{3T_j^{(n)} - 4T_{j-1}^{(n)} + T_{j-2}^{(n)}}{2\Delta x} \right) \right] = \mathcal{D} \left(\frac{T_{j+1}^{(n)} - 2T_j^{(n)} + T_{j-1}^{(n)}}{(\Delta x)^2} \right), \quad (2)$$

where $T_j^{(n)}$ is the approximation of T at time t_n and $x = x_j$, $t_n = t_{n-1} + \Delta t^{(n)}$, $x_j = x_{j-1} + \Delta x$ and $\Delta t^{(n)} = \kappa^{(n)} \Delta t^{(n-1)}$. The superscript on κ is not shown; we understand that $\kappa = \kappa^{(n)}$. Note that $\kappa < 1$ and $\kappa > 1$ respectively yield a decrease or an increase in the time step. The weight ω , which controls the amount of upwind bias, is assumed to be non-negative. Choosing $\omega = 0$ or $\omega = 1$ approximates the convective term by a central difference or an upwind biased difference respectively.

The FD scheme blends an upwind biased scheme with a central scheme. Thus it may be of advantage in problems with time-dependent convective (or diffusive) terms, where the flow varies from being convection dominated to diffusive dominated. A reduction in the number of nodes needed to evaluate the FD scheme was the motivation behind deriving Equation (2). The above implicit scheme requires only $3n + 3$ nodes in n -dimensions compared with a Crank-Nicolson type of scheme (using the same spatial discretisation as Equation (2)) which would require $6n + 2$ nodes. The variable time step overcomes the limitation of some three-level schemes with fixed time steps.

Setting $\kappa = 1$ (constant time steps) gives the backward differentiation formula (which is second order in time), while setting $\kappa = 0$ gives a scheme that is first order in time. As the finite difference approximation involves three time levels if $\kappa \neq 0$, it can not be used for the first time step. Any two-level scheme could be used, but all results that follow were computed using the difference equation with $\kappa = 0$ for the first time step.

3 Modified Equivalent PDE

We use the modified equivalent partial differential equation (MEPDE) to investigate the numerical diffusive and convective characteristics of the FD scheme. The MEPDE technique was first outlined by Warming and Hyett [5], and later improved Noye and Hayman [3]. The MEPDE technique involves re-writing the Taylor series expansion of the FD scheme (or the equivalent PDE) with all temporal derivatives eliminated using the equivalent PDE. The equivalent PDE of the FD scheme given by Equation (2) is

$$\frac{\partial T}{\partial t} + \mathcal{C} \frac{\partial T}{\partial x} - \mathcal{D} \frac{\partial^2 T}{\partial x^2} + \text{TE} = 0,$$

where TE is the truncation error of the FD scheme. The truncation error from the equivalent PDE is given by

$$\begin{aligned} \text{TE} = \frac{\Delta x^2}{12} \left[2\mathcal{C}(1 - 3\omega) \frac{\partial^3 T}{\partial x^3} - \mathcal{D} \frac{\partial^4 T}{\partial x^4} \right] - \frac{\Delta t^2}{6} \frac{(1 + \kappa)}{\kappa} \frac{\partial^3 T}{\partial t^3} \\ + O(\Delta x)^3 + O(\Delta t)^3. \end{aligned}$$

Note that the above expression is only valid for $\kappa \neq 0$, since $\kappa = 0$ results in a first order approximation in time. It would appear that taking $\omega = \frac{1}{3}$ would give the highest accuracy as this makes the approximation to the convective term third order. However, we shall see that this is not the best choice.

Using the above equivalent PDE to eliminate temporal derivatives yields

the truncation error from the MEPDE

$$\begin{aligned} \text{TE} = \frac{\Delta t^2}{6} \frac{(1 + \kappa)}{\kappa} & \left[\mathcal{C}^3 \frac{\partial^3 T}{\partial x^3} - 3\mathcal{C}^2 \mathcal{D} \frac{\partial^4 T}{\partial x^4} + 3\mathcal{C} \mathcal{D}^2 \frac{\partial^5 T}{\partial x^5} - \mathcal{D}^3 \frac{\partial^6 T}{\partial x^6} \right] \\ & + \frac{\Delta x^2}{12} \left[2\mathcal{C}(1 - 3\omega) \frac{\partial^3 T}{\partial x^3} - \mathcal{D} \frac{\partial^4 T}{\partial x^4} \right] \\ & + O(\Delta x)^3 + O(\Delta t)^3 + O(\Delta x)^2 O(\Delta t)^2. \quad (3) \end{aligned}$$

The above expression may be written in the more compact form

$$\begin{aligned} \text{TE} = \left\{ \frac{\Delta t^2}{6} \frac{(1 + \kappa)}{\kappa} \left(\mathcal{C} - \mathcal{D} \frac{\partial}{\partial x} \right)^3 + \frac{\Delta x^2}{12} \left[2(1 - 3\omega)\mathcal{C} - \mathcal{D} \frac{\partial}{\partial x} \right] \right\} \frac{\partial^3 T}{\partial x^3} \\ + O(\Delta x)^3 + O(\Delta t)^3 + O(\Delta x)^2 O(\Delta t)^2. \end{aligned}$$

Warming and Hyett [5] were the first authors to note the relation between the MEPDE's truncation error and the artificial damping and phase shifting properties of the numerical solution. Noye [2, pp. 242–247] formalised the relation and later improved the presentation [4, pp. 236–237] that we now follow.

Let the truncation of the MEPDE be given by

$$\text{TE} = \sum_{k=3}^{\infty} \mathcal{C}(\Delta x)^{k-1} \frac{\eta_k}{k!} \frac{\partial^k T}{\partial x^k}. \quad (4)$$

The above series begins at $k = 3$ since the scheme we have developed is second order in time and space (unless $\kappa = 0$).

Consider the amplitude and speed of an infinite travelling wave of unit amplitude propagating according to both the transport equation (1), which we term the true solution, and FD equation (2), which we term the numerical solution. The ratio of the numerical to true wave amplitudes and the ratio of the numerical to true wave speed after one period are collectively known as the wave propagation characteristics, and denoted by ζ and μ respectively. Noye [4] provides the following two expressions

$$\zeta = \exp \left[2\pi \sum_{k=4(2)\infty} (-1)^{(k/2)+1} \left(\frac{2\pi}{N_\lambda} \right)^{k-1} \frac{\eta_k}{k!} \right]$$

and

$$\mu = 1 + \sum_{k=3(2)\infty} (-1)^{(k-1)/2} \left(\frac{2\pi}{N_\lambda} \right)^{k-1} \frac{\eta_k}{k!},$$

which relate the truncation error coefficients to the wave amplitude and speed propagation characteristics respectively. The parameter $N_\lambda = \lambda/\Delta x$ defines the number of grid spacings in one wavelength λ .

Using the truncation error given by Equation (3), the truncation error

coefficients η (as defined by Equation (4)) are

$$\begin{aligned} \frac{\eta_3}{3!} &= [1 - 3\omega + c^2(1 + \kappa)/\kappa]/6, & \frac{\eta_4}{4!} &= -[s/c + 6cs(1 + \kappa)/\kappa]/12, \\ \frac{\eta_5}{5!} &= s^2(1 + \kappa)/(2\kappa), & \text{and} & \quad \frac{\eta_6}{6!} &= -s^3(1 + \kappa)/(6c\kappa), \end{aligned}$$

where the Courant number, diffusion number, and Peclet number are defined as

$$c = \mathcal{C} \frac{\Delta t}{\Delta x}, \quad s = \mathcal{D} \frac{\Delta t}{\Delta x^2}, \quad \text{and} \quad \text{Pe} = \frac{c}{s} = \frac{\mathcal{C}}{\mathcal{D}} \Delta x$$

respectively.

Then the wave amplitude characteristic is

$$\zeta = \exp \left\{ \frac{\pi}{6} \left(\frac{2\pi}{N_\lambda} \right)^3 \frac{s}{c} \left[1 + 6 \left(\frac{1 + \kappa}{\kappa} \right) c^2 - 2 \left(\frac{2\pi}{N_\lambda} \right)^2 \left(\frac{1 + \kappa}{\kappa} \right) s^2 + \dots \right] \right\},$$

while the wave speed characteristic is

$$\mu = 1 - \frac{1}{6} \left(\frac{2\pi}{N_\lambda} \right)^2 \left[1 - 3\omega + \left(\frac{1 + \kappa}{\kappa} \right) c^2 - 3 \left(\frac{2\pi}{N_\lambda} \right)^2 \left(\frac{1 + \kappa}{\kappa} \right) s^2 + \dots \right].$$

As the time step ratio κ increases, both the wave amplitude and speed characteristics improve marginally. However, as κ increases so does the re-

sulting time step and associated truncation error. The amplitude characteristic does not depend on ω , but choosing

$$\omega = \omega_{\text{opt}} = \frac{1}{3} \left[1 + \left(\frac{1 + \kappa}{\kappa} \right) c^2 \right], \quad (5)$$

results in cancellation of the first 3 terms in the square brackets in the expression for the wave speed characteristic. For small c the optimal weight ω_{opt} is approximately $\frac{1}{3}$. As expected, for convection dominated flows (c large) the wave speed characteristics improve as ω increases, since more weighting is given to the upwind biased difference than the central difference.

4 Numerical Experiments

It is easily verified that

$$T(x, t) = \frac{0.1}{\sqrt{0.01 + 4\mathcal{D}t}} \exp \left[-\frac{(x - \frac{1}{2} - \mathcal{C}t)^2}{0.01 + 4\mathcal{D}t} \right] \quad (6)$$

is an exact solution of the transport equation. At $t = 0$ it gives a narrow Gaussian pulse of unit height centred on $x = \frac{1}{2}$ (see Figure 1). This solution is used to examine the accuracy of the FD approximation (2).

As time passes the Gaussian pulse moves to the right with velocity \mathcal{C} , at the same time becoming shorter and broader due to diffusion. Without loss

of generality we take $\mathcal{C} = 1$, so the pulse will be centred on $x = 3.5$ at time $t = 3$. Thus an interval of at least $[0, 5]$ would be required for large values of \mathcal{D} . The FD approximation (2) is implicit, and requires that a system of linear equations is solved for each time step. To avoid the large interval and consequent large number of linear equations, a unit interval is used with the cyclic boundary condition

$$T(1, t) = T(0, t).$$

Thus the pulse enters the left boundary as it leaves the right boundary, as if the unit interval was bent into a circle. An alternative interpretation is that we start with an infinite train of pulses spaced a unit distance apart, and we watch through the “window” $[0, 1]$ as they move to the right.

The use of the cyclic boundary condition does reduce the size of the system of linear equations to be solved at each time step, but unfortunately the coefficient matrix is no longer banded. In addition to the band of width 4, it has 4 extra non-zero elements; 1 in the lower left corner and 3 in the top right corner (in the uppermost 2 diagonals). An LU factorisation of the coefficient matrix is used to solve the linear equations, but instead of there being non-zero elements in only 3 diagonals of L and 2 diagonals of U , the “extra” elements require that there are also non-zero elements in the last row of L and last two columns of U . An algorithm has been constructed so that only the non-zero elements of L and U are calculated. The factorisation must be found at each step unless $\kappa = 1$ for consecutive steps.

The initial values $T_j^{(0)} = T(x_j, 0)$ are obtained from Equation (6). As

mentioned earlier, the first step is carried out using the FD approximation with $\kappa = 0$. Subsequent steps are calculated using a fixed value of κ (usually 1), and at specified times the root mean square error is calculated.

A grid with $\Delta x = \frac{1}{40}$ and a convective velocity $\mathcal{C} = 1$ was used for all calculations. Unless stated otherwise, the time step was kept constant ($\kappa = 1$). The Courant number c is determined by the choice of the time step Δt , and the diffusion number s (and hence the Peclet number) determined by the value of \mathcal{D} .

Figure 1 shows a comparison of the exact solution and the numerical solution. The values $\Delta t = 0.01$ and $\mathcal{D} = 0.008$ were used, resulting in a Courant number of $c = 0.4$ and a Peclet number of $\text{Pe} = 3.125$. The value of ω given by Equation (5) (that is ω_{opt}) was used for the calculation.

Initially the Gaussian pulse is at the centre of the interval. At $t = 0.5$ the pulse has moved a distance of 0.5, so half of the pulse has disappeared out the right side of the interval and re-entered the left side of the interval. An alternative, and perhaps simpler interpretation, is that we can see the left half of the pulse at the right end of the interval and can see the right half of the pulse which follows at the left. Due to the diffusive effects, their maxima are about 0.6 and they are correspondingly broader (the integral of T over the interval should be conserved). At $t = 1$ the pulses have diffused further and moved a distance of 1, so the following pulse is centred on $x = \frac{1}{2}$. The final plot is at $t = 3$; the pulse that was the third behind the initial pulse is centred on $x = \frac{1}{2}$, and due to the large amount of diffusion the pulses have

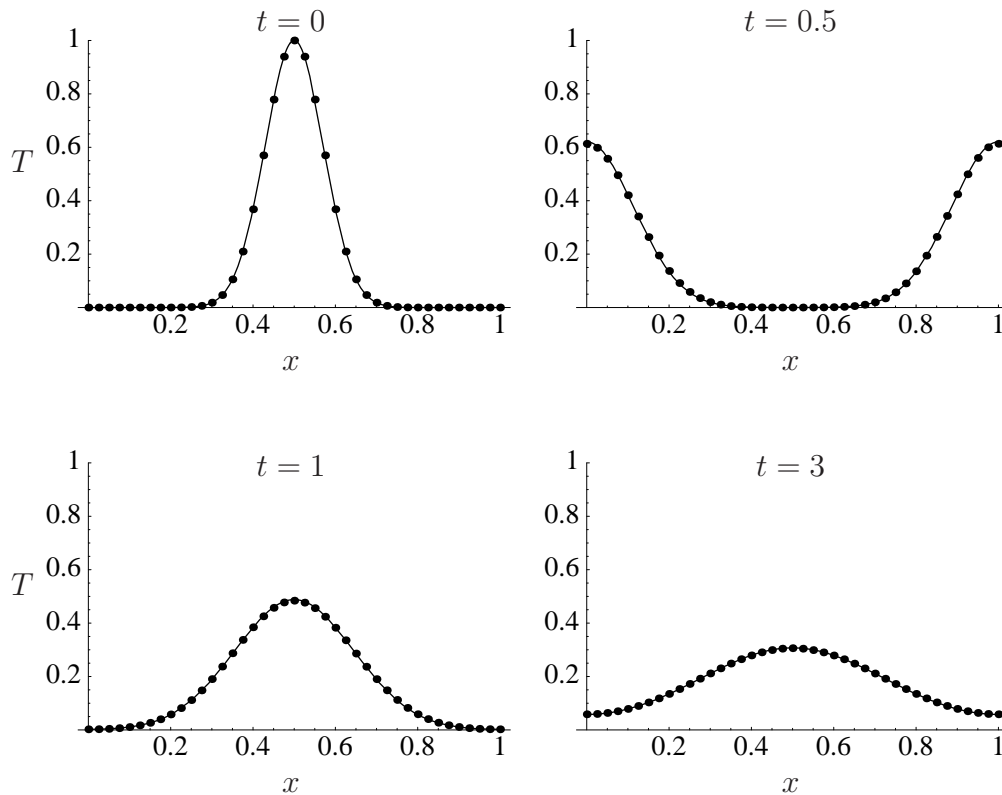


FIGURE 1: The exact solution (solid line) and numerical solution (dots) for $c = 0.4$ and $Pe = 3.125$ at the indicated times. The numerical solution was calculated using $\omega = \omega_{\text{opt}} = 0.44$.

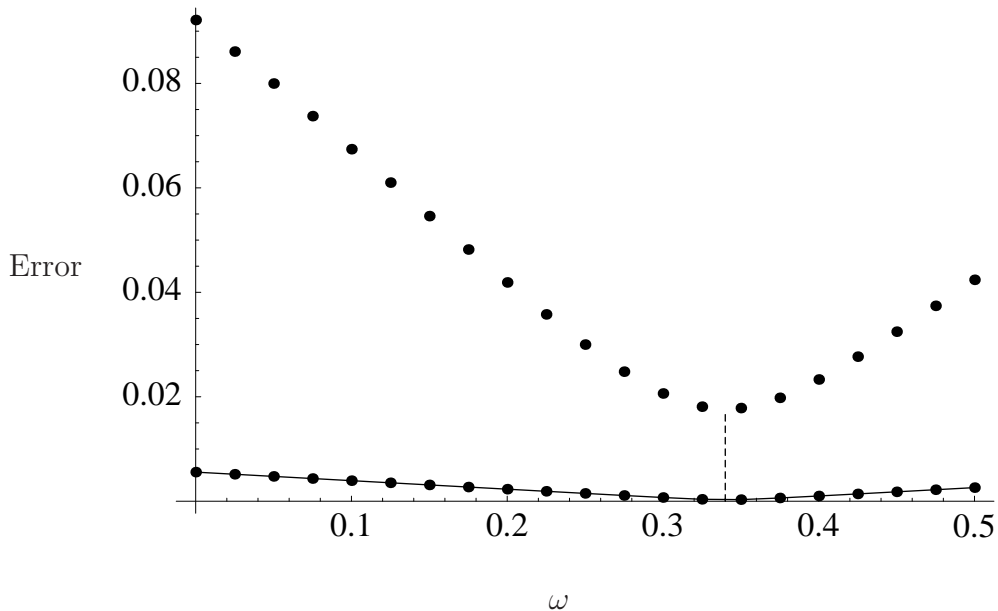


FIGURE 2: Error dependence on ω for $c = 0.1$ using $Pe = 25$ and $Pe = 2.5$ (dots and line-plus-dots respectively). The value of ω_{opt} is indicated by the dashed line.

“coalesced”. The numerical solution is clearly in excellent agreement with the exact solution at all times.

To determine the best value of the weight ω the numerical solution and the norm of its error were calculated with $\Delta t = 0.0025$ ($c = 0.1$) for $\omega = 0(0.025)0.5$. Two values of \mathcal{D} , corresponding to $\text{Pe} = 2.5$ and $\text{Pe} = 25$, were used. The results are shown in Figure 2. For both values of the Peclet number, the minimum error occurs for a value of ω extremely close to the value of ω_{opt} defined by Equation (5). The same outcome was produced when time steps of $\Delta t = 0.01$ and $\Delta t = 0.03$ ($c = 0.4$ and $c = 1.2$) were used (see Figures 3 and 4). Note that in the latter case the optimal weight is greater than 1.

It is apparent from Figures 2, 3 and 4 that the error is much smaller for the smaller Peclet number for any value of ω . This is not surprising as increasing the rate of diffusion lowers the Peclet number. Comparing Figures 2, 3 and 4 shows that the errors corresponding to the optimal value of ω increase with the Courant number. However, the results are still quite acceptable for $c = 1.2$, particularly for the smaller Peclet number. It should be noted that the best way to improve the accuracy for given values of \mathcal{C} and \mathcal{D} is to reduce Δx . Although this increases the Courant number, it reduces the value of the Peclet number and the Peclet number has a stronger influence on the error.

It was observed that the numerical wave speed was too low if $\omega < \omega_{\text{opt}}$, and too high if $\omega > \omega_{\text{opt}}$. This can be seen in Figure 5, where the solution at

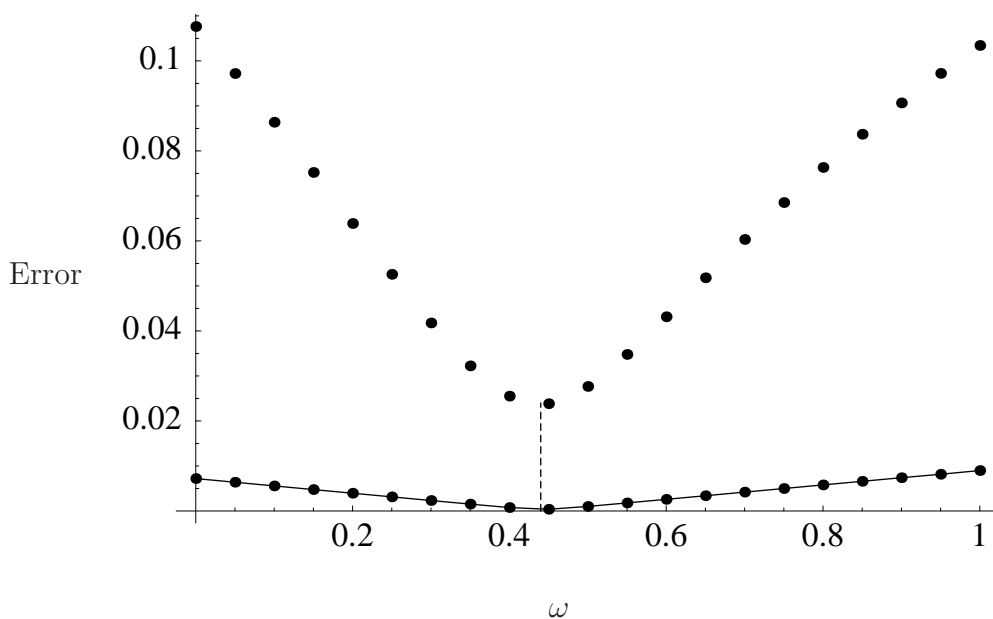


FIGURE 3: Error dependence on ω for $c = 0.4$ using $Pe = 25$ and $Pe = 2.5$ (dots and line-plus-dots respectively). The value of ω_{opt} is indicated by the dashed line.

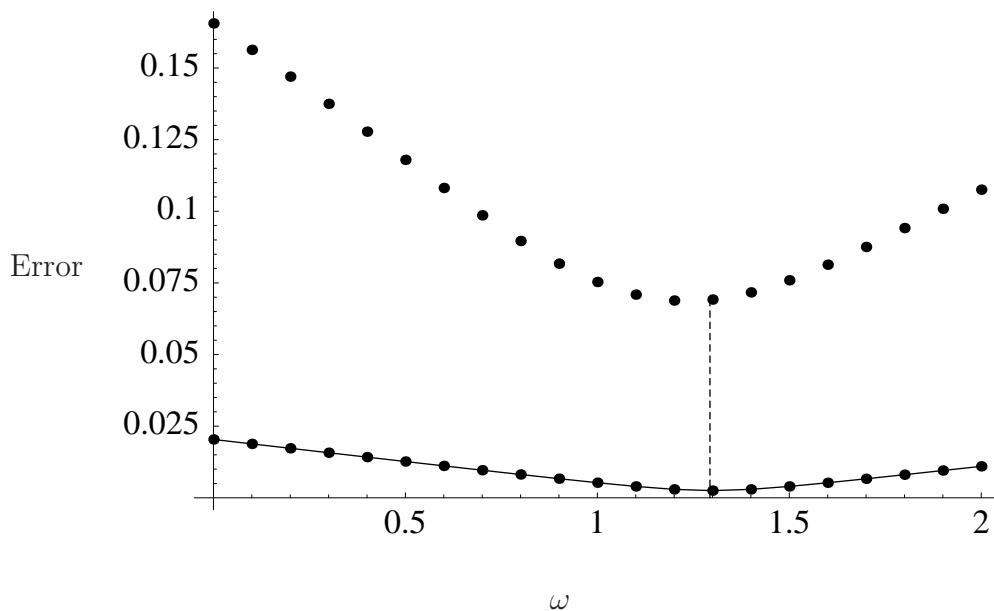


FIGURE 4: Error dependence on ω for $c = 1.2$ using $Pe = 25$ and $Pe = 2.5$ (dots and line-plus-dots respectively). The value of ω_{opt} is indicated by the dashed line.

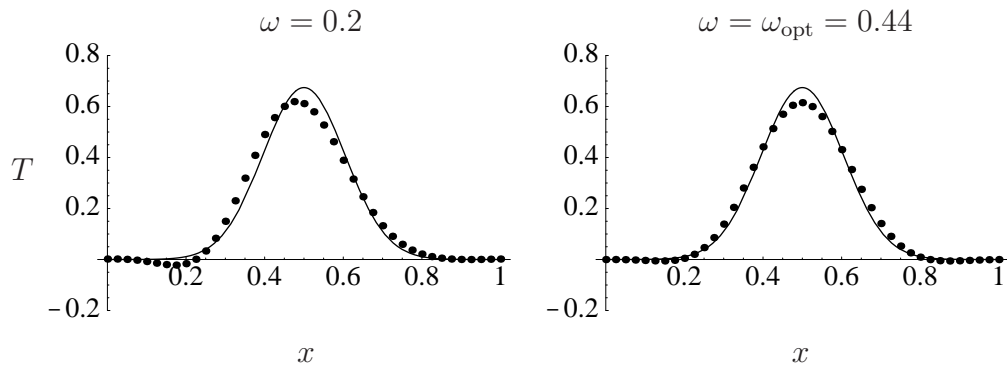


FIGURE 5: The exact solution (solid line) and numerical solution (dots) at time $t = 3$ for $c = 0.4$ and $\text{Pe} = 25$. The numerical solution was calculated using two values of ω .

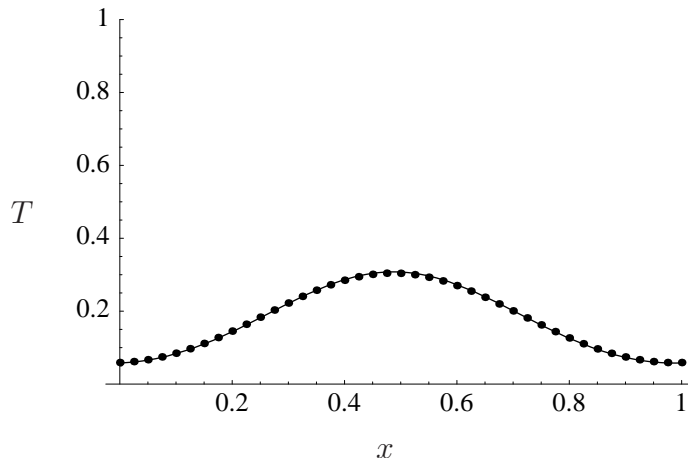


FIGURE 6: The numerical solution for for T at $Pe = 3.125$ calculated with an increasing time step ($\kappa = 1.01$). The optimal value of ω was used at each step.

$t = 3$ has been calculated using two values of ω . A time step of $\Delta t = 0.01$ and diffusion coefficient $\mathcal{D} = 0.001$ were used, giving the Courant number of $c = 0.4$ and Peclet number $Pe = 25$. The numerical solution calculated with $\omega = 0.2$ clearly lags the exact solution; the numerical wave speed is too low. When the optimal value $\omega = 0.44$ is used, the numerical wave speed is reasonably accurate. However, in both cases the peak value is too low; the numerical solution has suffered from slightly too much diffusion.

The numerical solution shown in Figure 6 was calculated using a variable time step. The initial time step was $\Delta t = 0.01$, and at each time step it was increased by the factor $\kappa = 1.01$ resulting in $t = 3$ being reached in 139 steps with a final time step of 0.0395. The Peclet number was 3.125, and the Courant number varied from 0.4 to about 1.58. The optimal value of ω was calculated at each time step. As expected, the solution is not as accurate as the one obtained with a constant time step of 0.01 (see Figure 1); the error is approximately 3 times greater. It is also more accurate to take the same number of steps of constant length, but in practice one would only use a value of $\kappa > 1$ for a few steps.

5 Conclusion

We have investigated the accuracy of an implicit FD approximation of the transport equation and found that the highest accuracy is obtained not by choosing the weight that maximises the order of the approximation to the convective term, but by choosing the value that makes the numerical wave speed as close as possible to the true wave speed. Although we have only considered one spatial dimension and a constant convective velocity, the method could be adapted to two or more dimensions and a variable velocity by calculating the optimal value of the weight at each grid point.

References

- [1] K. W. Morton. *Numerical Solution of Convection-Diffusion Problems*. Chapman & Hall, 1996. [C1180](#)
- [2] J. Noye. Finite Difference Techniques for Partial Differential Equations. *Computational Techniques for Differential Equations*, pages 95–354. North-Holland, 1984. [C1184](#)
- [3] J. Noye and K. Hayman. Accurate Finite Difference Methods for Solving the Advection-Diffusion Equation. In J. Noye and R. May, editors, *Computational Techniques and Applications: CTAC-85*, pages 137–158. Elsevier Science, 1986. [C1183](#)
- [4] J. Noye. Finite Difference Methods for Solving the One-Dimensional Transport Equation. *Numerical Modelling: Applications to Marine Systems*, pages 231–256. North-Holland, 1987. [C1184](#), [C1185](#)
- [5] R. Warming and B. Hyett. The Modified Equation Approach to the Stability and Accuracy of Finite-Difference Methods. *J. Computational Physics*, 14:159–179, 1974. [C1183](#), [C1184](#)