ISSN 0918-6654

Japan Society of CFD / CFD Journal

e-mail: jscfd@mail1.shocklab.mech.tuat.ac.jp
anonymous ftp: www.shocklab.mech.tuat.ac.jp
http://www.shocklab.mech.tuat.ac.jp

Managing Editor: Koichi Oshima
2-7-7 Sekimachi-Kita, Nerima, Tokyo 177-0051, Japan
e-mail: oshima@mail1.shocklab.mech.tuat.ac.jp

Dear Contributors to ISCFD99 Proceedings:

We apologize such long delay of publication of this Proceedings, which is caused by accidental damage of our network system and unexpectedly complicated software system for formatting the figures.

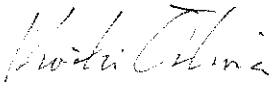
Now we are close to the final stage, and here present draft of the first part.

This proceedings are published as I and II;

Proceedings I: Invited and organized session papers, Special Issue of vol.9 no.1 pp.437-618
Proceedings II: Contributed papers, Special number of vol.9 pp.1-647

Here is the total list of papers (Contents of both Proceedings) and draft of Proceedings I.
The authors of Proceedings I are cordially requested to have proof reading of your papers.
Those of Proceedings II, please wait for a while.

Since these are quite bulky, we like to ask each author to pay the handling fee of 50US$ for one set of them. This is contrary to our original saying, but this is not a purchase act but only for handling.



Computational Fluid Dynamics JOURNAL vol.9 no.1-I April 2001 (pp.494–505)
Special Issue The Proceedings of ISCFD'99 held in Bremen on September 5 to 10, 1999

MACCORMACK'S METHOD FOR ADVECTION-REACTION EQUATIONS *

David F. GRIFFITHS † Desmond J. HIGHAM ‡

Abstract

MacCormack's method is an explicit, second order finite difference scheme that is widely used in the solution of hyperbolic problems. Here, we consider MacCormack's method applied to the linear advection equation with nonlinear source term. Various features of the method are analysed. First, we show that the conventional implementation is not stable for Courant numbers close to one unless a small time-step is used. A simple modification, based on source term averaging, is shown to remove this defect. We then examine spurious fixed points that are inherited from the underlying Runge–Kutta method. Next we consider adapting the time-step as a means of improving the efficiency of the method. Theoretical analysis based on the method of modified equations is combined with numerical tests on a travelling wave problem in order to give a feel for how the time-step should be refined. An adaptive approach based on temporal local error control is shown to have serious drawbacks. Much better performance is obtained with a modified error measure that takes account of immanent spatial errors.

Key Words: adaptivity, finite difference, modified equations, nonlinearity, source term, travelling wave, spuriosity.

1 INTRODUCTION

Our model problem is the partial differential equation (PDE)

$$u_t + au_x = g(u), \quad x, t > 0, \tag{1}$$

with solution $u(x,t)$, where $a > 0$ is constant and the source term g is generally nonlinear.

We are concerned with finite difference schemes that produce approximations $U_j^n \approx u(j\Delta x, n\Delta t)$ on a mesh $\{(j\Delta x, n\Delta t)\}_{j,n\geq 0}$. The schemes that we analyse come in two modes. In both cases we have

$$U_j^{n+1} = \tfrac{1}{2}(Z_j^{n+1} + W_j^{n+1}). \tag{2}$$

In forward-backward (FB) mode, the intermediate values Z_j^{n+1} and W_j^{n+1} are defined as

$$\begin{aligned} Z_j^{n+1} &= U_j^n - c(U_{j+1}^n - U_j^n) \\ &\quad + \Delta t\Big[\phi g(U_{j+1}^n) + (1-\phi)g(U_j^n)\Big] \qquad (3) \\ W_j^{n+1} &= U_j^n - c(Z_j^{n+1} - Z_{j-1}^{n+1}) \\ &\quad + \Delta t\Big[(1-\phi)g(Z_j^{n+1}) + \phi g(Z_{j-1}^{n+1})\Big] \qquad (4) \end{aligned}$$

where $c := a\Delta t/\Delta x$ is the Courant number and ϕ is a free parameter. In backward-forward (BF) mode, the intermediate values become

$$\begin{aligned} Z_j^{n+1} &= U_j^n - c(U_j^n - U_{j-1}^n) \\ &\quad + \Delta t\Big[(1-\phi)g(U_j^n) + \phi g(U_{j-1}^n)\Big], \qquad (5) \\ W_j^{n+1} &= U_j^n - c(Z_{j+1}^{n+1} - Z_j^{n+1}) \\ &\quad + \Delta t\Big[\phi g(Z_{j+1}^{n+1}) + (1-\phi)g(Z_j^{n+1})\Big]. \qquad (6) \end{aligned}$$

In the absence of a source term ($g(u) \equiv 0$) these methods reduce to conventional implementations of MacCormack's method for hyperbolic equations [10, 12].

* This manuscript appears as Technical Report NA/192, University of Dundee and University of Strathclyde Mathematics Research Report 25 (1999). The authors were supported by the Engineering and Physical Sciences Research Council of the UK under grant GR/K80228.

† Department of Mathematics, University of Dundee, Dundee, DD1 4HN, Scotland. Email:dfg@mcs.dund.ac.uk
‡ Department of Mathematics, University of Strathclyde, Glasgow, G1 1XH, Scotland. Email:d.j.higham@strath.ac.uk

The parameter ϕ introduces spatial averaging of the source term in an upwind/downwind manner consistent with the approximation of the advection term. The choice $\phi = 0$ corresponds to sampling the source term at a single point. We will show in section 2 that such a scheme is non-optimal in that it is not stable (in a von Neumann sense) for all Courant numbers in the range $0 < c \leq 1$. For this reason we will also consider the choice $\phi = \frac{1}{2}$, which maintains the order of accuracy and will be seen to improve the stability properties.

We concentrate on four particular schemes, which will be referred to as

FBpoint: (2), (3), (4) with $\phi = 0$
(forward-backward with point source)
FBave: (2), (3), (4) with $\phi = \frac{1}{2}$
(forward-backward with averaged source)
BFpoint: (2), (5), (6) with $\phi = 0$
(backward-forward with point source)
BFave: (2), (5), (6) with $\phi = \frac{1}{2}$
(backward-forward with averaged source)

We also note that a finite volume approach together with a natural treatment of the integration of the source term over each control volume in space yields a method with $\phi = \frac{1}{2}$ [7].

We consider the periodic initial value problem (PIVP) and the initial boundary value problem (IBVP) for (1). In the PIVP case, we are given $u(x,0)$ for $0 \leq x < 1$ and u is assumed to be periodic in space: $u(1+x,t) = u(x,t)$. In the IBVP case we are given $u(0,t)$ for $t > 0$ and $u(x,0)$ for x in the spatial interval of interest.

A finite difference scheme for (1) based on first order upwinding in space and timestepping with a general 2-stage, second order, explicit Runge–Kutta (RK) method was analysed in [2] with an emphasis on spurious solutions. In this work we study a higher order finite difference scheme and consider a range of topics that relate to long time behaviour; specifically, von Neumann stability, spurious solutions, travelling wave solutions and adaptivity.

In section 2 we perform a von Neumann stability analysis that justifies the choice $\phi = \frac{1}{2}$. Section 3 looks briefly at the existence of spurious fixed points. Related work in [2] studied an upwind difference scheme—the MacCormack type schemes analysed here may be regarded as using the $\theta = \frac{1}{2}$ timestepping method of [2] with more accurate spatial differencing. Previous work [1], devoted to ordinary differential equations (ODEs), showed that spurious behaviour of RK methods is essentially precluded by standard time-step adaptivity. In sections 4 and 5 we extend these ideas to hyperbolic PDEs of the type (1). We begin, in section 4, with analysis devoted to travelling wave solutions and show that in these circumstances it is possible to determine an optimal Courant number c and, thereby, an optimal time-step that depends on the spatial grid size Δx, parameters of the PDE and the exponential decay of the initial data. Then, in section 5, we consider ways of automatically choosing the time-step that will lead to the optimal value being achieved. We argue that a simple ODE approach cannot be successful, so we develop and test a more promising strategy that takes account of inherent spatial errors. This approach can be used more widely for other PDEs and other methods.

We note that a brief summary of the material in section 3 was given in the manuscript [3].

2 LINEAR STABILITY

In this section we examine the stability of the finite difference schemes with linear source terms by taking $g(u) = bu$. We assume that $b < 0$ so that $u = 0$ is a stable fixed point of the underlying continuous problem (1). Inserting $U_j^n = \xi^n e^{ij\theta}$ into the resulting difference equations we obtain an expression for ξ which is required to satisfy von Neumann's criterion for stability, namely $|\xi| \leq 1$ for all θ [10, 12]. Both FBpoint and BFpoint schemes lead to the same condition

$$16\left(c^2 - (r-1)^2\right)c^2 s^2 + 8r(r-2)c^2 s + r(r-2)\left((r-1)^2 + 3\right) \leq 0, \qquad (7)$$

where $r := -b\Delta t$ and $s := \sin^2 \frac{1}{2}\theta$ ranges through the interval $[0,1]$. This leads to the stability conditions

$$0 \leq r \leq 2, \qquad c^2 \leq \frac{1}{4}\left((r-1)^2 + 3\right)$$

and the corresponding region of the r-c plane is shown in Figure 1. Note that when $r = 0$ this condition reduces to $0 \leq c \leq 1$, which is the well-known stability constraint for the Lax-Wendroff scheme on the linear advection equation with no source.

We note that, when using finite difference schemes of this form, it is typical to refine Δx and Δt while keeping c fixed. It follows that there is a serious practical drawback resulting from the fact that the stability region in Figure 1 is not convex with respect to horizontal lines. If we choose values of r and c in the stability region close to $r = 2$ and $c = 1$, say $r = 1.95$ and $c = 0.95$, then fixing c and refining the grid takes the scheme *outside* the stability region, before eventually returning it to stability.

For the BFave and FBave schemes, von Neumann's criterion for stability leads to

$$\left((r-2)^2 - 4c^2\right)\left(r^2 - 4c^2\right)s^2 - 2r(r-2)\left((r-1)^2 - 4c^2 + 1\right)s + r(r-2)\left((r-1)^2 + 3\right) \leq 0,$$

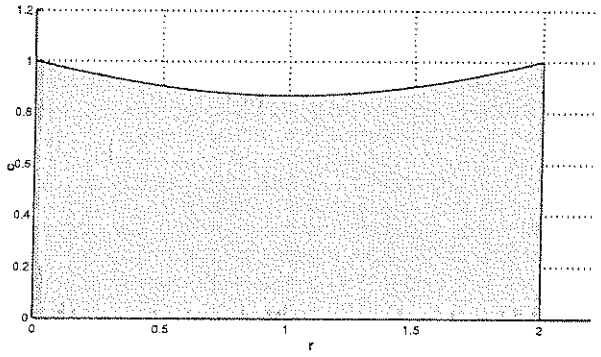


Fig.1: Stability region for BFpoint and FBpoint.

for all $0 \leq s \leq 1$. It is readily shown that this is equivalent to the conditions $0 \leq r \leq 2$ and $0 \leq c \leq 1$. We see that averaging the source term enlarges the stability region and makes it convex. This ensures that stability cannot be lost when the grid is refined with c fixed. Hence, the new BFave and FBave schemes introduced here have a significant advantage over the point source versions in terms of linear stability.

3 SPURIOUS SOLUTIONS FOR THE PERIODIC INITIAL VALUE PROBLEM

In this section we concentrate on the PIVP. We note that if $g(u^\star) = 0$ then (1) has a spatially uniform fixed point (SUFP) $u(x,t) \equiv u^\star$. The finite difference schemes have a corresponding SUFP $U_j^n \equiv u^\star$. If $g(u^\star) = 0$ and $g'(u^\star) < 0$, so that the SUFP of (1) is linearly stable, then the von Neumann analysis in section 2 determines the linear stability of the corresponding discrete SUFP—take $r = -g'(u^\star)\Delta t$ in Figure 1.

We also note that if the numerical solution is spatially uniform, then the finite difference schemes collapse to the improved Euler method for ODEs, which is known to admit spurious fixed points. Hence, for a given source term, all four schemes may produce a spurious SUFP; that is, a solution $U_j^n \equiv u^\star$ with $g(u^\star) \neq 0$. The linear stability of such solutions is analysed by first linearizing the finite difference equations around $U_j^n = u^\star$ and then applying von Neumann's method to the resulting constant coefficient equations. The results of the analysis are the same regardless of whether BF or FB modes are used but they do depend on whether or not the source term is averaged.

To illustrate this effect, we consider the logistic source term

$$g(u) = \alpha u(1-u), \quad \alpha > 0 \text{ constant.} \tag{8}$$

In this case, it is well known (see, for example, [2, 5]) that the improved Euler method admits the spurious fixed points

$$u_\pm^\star := \frac{2 + r \pm \sqrt{r^2 - 4}}{2r}, \quad \text{for } r > 2, \tag{9}$$

where $r := \alpha\Delta t$. These spurious fixed points are linearly stable for $2 < r < \sqrt{8}$.

On examining the linear stability of the corresponding SUFPs for each of the four MacCormack type schemes, we find it convenient to work in the transformed parameter space c-d, where $d := \sqrt{r^2 - 4}$. The $u_-^\star$ branch for the FBpoint and BFpoint schemes is found to be stable provided that

$$[4(c+d-1)cs + d(d-2)][4(c+d+1)cs + d(d+2)] < 0 \tag{10}$$

holds for all $s \in [0,1]$ (the variable s is defined following (7)). The individual factors on the left of (10) must have a constant sign for $s \in [0,1]$ and this leads to the condition $c < \frac{1}{2}(2-d)$. For the $u_+^\star$ branch we replace d by $-d$ in (10) and find that the FBpoint and BFpoint schemes are stable when $c < \frac{1}{2}d$. Figure 2 shows the regions of stability in the r-c plane for the true fixed point $u^\star = 1$ and the spurious SUFPs $u_\pm^\star$ of the FBpoint and BFpoint schemes.

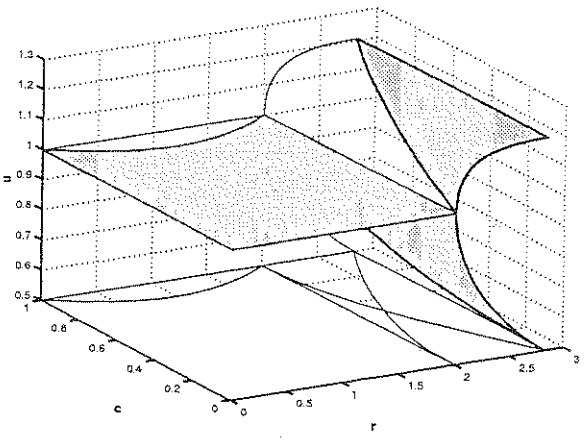


Fig.2: Stability region for spatially uniform fixed points with logistic source and point evaluation of the source (BFpoint and FBpoint).

For the source term averaged schemes, FBave and BFave, we find that the stability inequality for the

lower branch $u^\star_-$ corresponding to (10) is

$$-[d(2+d)+s(2c+d)(2c-d-2)][d(2-d) \\ -s(2c-d+2)(2c+d)] < 0. \quad (11)$$

The first bracketed factor on the left changes sign for $s \in (0,1)$ with $0 < c < 1$ and $d > 0$, whereas the second factor is of constant sign. This means that the fixed point is stable to perturbations of certain frequencies but is unstable to others and is therefore unstable for $0 < c < 1$ and $d > 0$ (it is, however, stable in the ODE case; this apparent paradox arises because we then have $c = 0$ and the first factor in (11) has its zero at $s = 1$).

To address the stability of the upper branch $u^\star_+$ we again replace d by $-d$ in (11) and find that it is stable for $0 < c < 1$ and $2 < r < \sqrt{8}$. Figure 3 shows the regions of stability for the true fixed point $u^\star = 1$ and the spurious SUFFPs $u^\star_\pm$ of the FBave and BFpoint schemes.

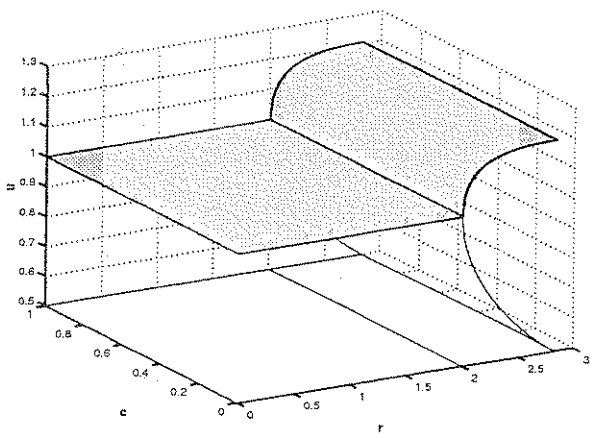


Fig.3: Stability region for spatially uniform fixed points of BFave and FBave with logistic source.

The von Neumann analysis in section 2 implies that source term averaging has the desirable effect of improving the linear stability of true SUFFPs. Comparing Figures 2 and 3 we see that, with the logistic source term, averaging also has the negative effect of stabilizing the upper spurious branch for all $0 < c < 1$ at the $r = 2$ boundary.

Qualitatively similar results are valid for small perturbations of spurious SUFFPs $u^\star$ for more general nonlinear source terms provided that $g'(u^\star) \neq 0$ (see [5, Section 3] for the ODE case). Note that with the logistic source term (8), the improved Euler method does not admit spurious fixed points for sufficiently small Δt; namely $\Delta t < 2/\alpha$. However, with other nonlinear source terms it is possible for spurious fixed points to persist for arbitrarily small Δt. For example, with $g(u) = u|u|$ we find spurious fixed points $u^\star_\pm = \pm 1/\Delta t$ for all $\Delta t > 0$. However, under mild assumptions, it follows from [6] that any spurious fixed point that exists for arbitrarily small Δt must blow up as $\Delta t \to 0$, and, from a result of [2, Lemma 5], it may be deduced that such a fixed point must be unstable for sufficiently small Δt. For the four MacCormack-type schemes, a continuity argument may be applied to prove that the corresponding spurious SUFP for the PIVP must be unstable for sufficiently small Δt. See [2, Theorem 6] for details of a similar continuity argument.

We conclude this section by showing the results of some numerical experiments on the BFpoint scheme that illustrate the analysis. (Similar experiments were performed in [3] for the FBpoint scheme.) We consider the PIVP with $a = 1$ using the logistic source term (8). We choose parameters

$$\Delta x = \frac{1}{36}, \quad c := \frac{\Delta x}{\Delta t} = 0.3, \quad r := \alpha \Delta t = 2.5$$

so that $\alpha = 300$. These give $u^\star_+ = 1.2$ and $u^\star_- = 0.6$. For this value of r the upper branch $u^\star_+$ is stable for $0 \leq c \leq 0.75$ and the lower branch $u^\star_-$ is stable for $0 \leq c \leq 0.25$; see, Figure 2. Hence, in these tests $u^\star_+$ is stable and $u^\star_-$ is unstable.

To begin, we take initial data of the form

$$u(x,0) = u^\star_+ + \gamma \sin(2\pi x), \quad (12)$$

where γ is a fixed parameter. We observed experimentally that this initial data lies in the basin of attraction of the stable spurious SUFP $u^\star_+$ when $|\gamma| <\approx 0.15$. The upper left-hand picture in Figure 4 shows the numerical solution obtained with $\gamma = 0.1$. We see that the spurious level of 1.2 is rapidly attained.

The upper right-hand picture in Figure 4 relates to the unstable lower branch. We take initial data

$$u(x,0) = u^\star_- + \gamma \sin(2\pi x),$$

with $\gamma = 0.001$. We see that the numerical solution is initially attracted to the fixed point $u^\star_-$, but at later time develops a high frequency spatial oscillation. A study of the stability inequality (10) reveals that the first factor changes sign at $s = \frac{25}{32}$ from which we deduce that the fixed point is unstable to high frequency perturbations. When the solution develops a high frequency component, due to rounding errors for example, these modes grow to pollute the solution. We note that nonlinear effects prevent the solution from blowing up.

A related type of behaviour is illustrated in the lower pictures in Figure 4. For the lower left picture the initial data is given by (12) with $\gamma = 0.2$ and we see that some initial data has initially been drawn towards the solution of the previous case (upper right,

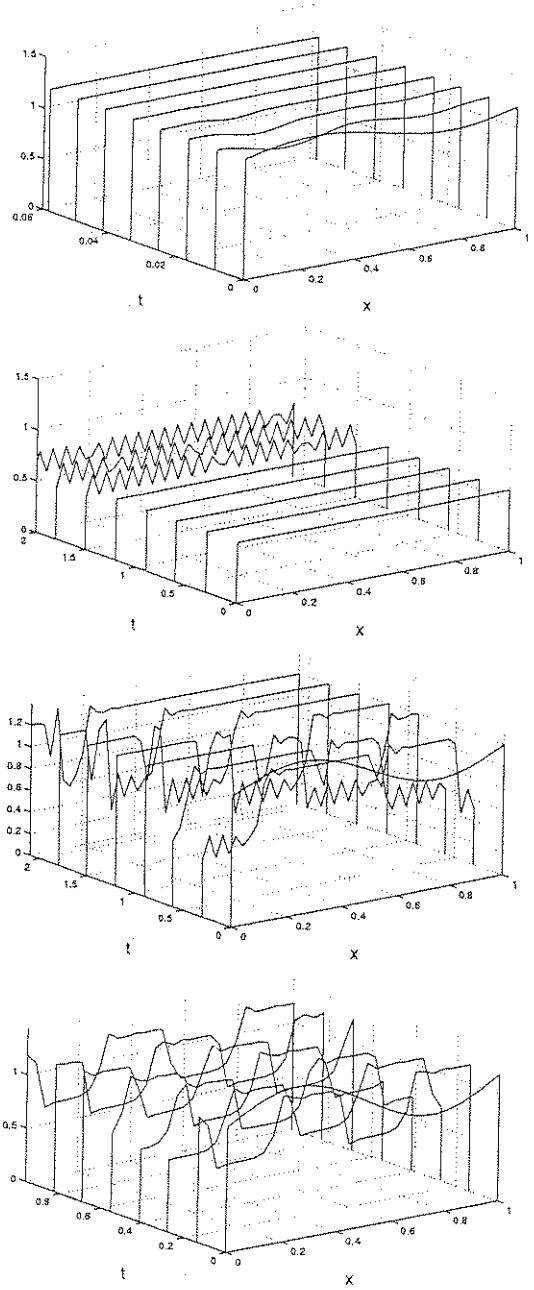


Fig.4: Spurious solutions with BFpoint on the PIVP.

oscillating around the unstable fixed point $u_-^\star$) while the majority of the initial data evolves towards $u_+^\star$.

For the solution shown in the lower right of Figure 4 the grid data are changed to

$$\Delta x = \frac{1}{27}, \quad c := \frac{\Delta x}{\Delta t} = 0.2, \quad r := \alpha \Delta t = 2.22$$

from which it may be verified that both $u_-^\star$ and $u_+^\star$ are stable. The initial data is given by (12) with $\gamma = 0.25$. The stability of each of the spurious SUFPs leads to smooth sections between "jumps" in contrast to the oscillatory behaviour around the unstable fixed point observed in the previous case. An investigation into this type of jump solution can be found in [2].

4 TRAVELLING WAVE PROBLEMS

We now consider wave-like solutions to the PDE (1). Our interest in these types of solution derives from the fact that we can perform numerical experiments with our MacCormack schemes and identify optimal time-steps Δt that minimize the phase error for given wave speeds and given grid sizes Δx. This information will then be used to test ideas on adaptive time stepping that are described in the next section.

4.1 Travelling Wave Solutions

The PDE (1) has travelling wave solutions $u(x,t) = \Phi(x - st)$ if Φ satisfies the ODE

$$\Phi' = \frac{1}{a-s} g(\Phi). \tag{13}$$

Thus, if the problem has a travelling wave solution for one particular speed $s \neq a$, it will have a travelling wave solution for every choice of wave speed, the profiles in each case differing only in the scaling of the independent variable. Waves of speed $s = a$, the characteristic speed, correspond to profiles Φ satisfying $g(\Phi) = 0$. These piecewise constant solutions and their corresponding numerical solutions have been studied in detail by LeVeque and Yee [9].

We are interested only in solutions that have bounded values at infinity and these values correspond to zeros of g. For example, if $g(u_-) = g(u_+) = 0$, where $u_- < u_+$ are simple zeros and $g(u)$ has constant sign for $u_- < u < u_+$, then profiles of the type illustrated in Figure 5 are possible depending on the sign of $g(u)/(a-s)$ for $u_- < u < u_+$. The profiles become steeper as s approaches a. We focus on the case shown in Figure 5(ii), corresponding to $g(u)/(a-s) < 0$ and $s > a > 0$, so that the wave travels to the right. With these assumptions $u = u_+$ and $u = u_-$ are, respectively, stable and unstable fixed points of equation (13).

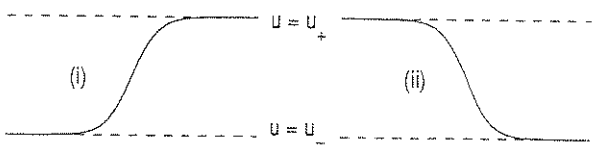


Fig.5: Travelling wave profiles for solutions Φ of equation (13). Case (i) when $g(u)/(a-s) > 0$ on $u_- < u < u_+$ and case (ii) when $g(u)/(a-s) < 0$.

By linearizing this ODE around the fixed points we find solutions of the form

$$\Phi(\xi) = u_- + A\exp(\frac{g'(u_-)}{a-s}\xi) \tag{14}$$

and

$$\Phi(\xi) = u_+ + B\exp(\frac{g'(u_+)}{a-s}\xi), \tag{15}$$

which give the asymptotic nature of the solutions at $\xi = +\infty$ and $\xi = -\infty$, respectively (A and B are arbitrary constants). Our assumptions regarding case (ii) of Figure 5 imply that $g'(u_-) > 0$ and $g'(u_+) < 0$.

These asymptotic results provide information on the development of travelling wave solutions when using more general initial data than $u(x,0) = \Phi(x)$. All initial data satisfying $u_- < u(x,0) < u_+$ will evolve to $u(x,t) = u_+$ as $t \to \infty$. If, additionally, $u(x,t)$ is monotonically decreasing then a wave-like solution will develop and the profile of this solution for large x will depend on precisely how $u(x,0)$ tends to zero as $x \to \infty$. Restricting ourselves to exponential decay, so that

$$u(x,0) \sim \exp(-\lambda x), \qquad \lambda > 0, \tag{16}$$

for large x, then, following Murray [11, Chapter 11], we find that

$$u(x,t) \sim \Phi(x - st),$$

where the speed s is deduced by comparing the decay rates in (14) and (16). This leads to

$$s = a + \frac{g'(u_-)}{\lambda}. \tag{17}$$

The development of solutions with other types of decay at infinity, such as algebraic decay, is more complex and will not be considered here.

4.2 Numerical Solution of Travelling Wave Problems

Our aim in this subsection is to show that optimal grid sizes can be identified when MacCormack schemes are applied to travelling wave problems associated with (1). Our analysis is based on the method of modified equations (see [4, 13]), which first requires the truncation errors of the methods to be computed.

In order to facilitate writing the rather cumbersome expressions for the truncation errors we define the nonlinear differential operator

$$\mathcal{F}(u) \equiv u_t + au_x - g(u)$$

and the linear differential operator

$$\mathcal{L}_u v \equiv v_t - av_x + g'(u)v,$$

where the subscript u emphasizes that the coefficients depend on the function $u(x,t)$. After tedious calculations we find that the leading terms in the truncation errors, T_{FB} and T_{BF}, for FB and BF, can be written as

$$\begin{aligned} T_{\mathsf{FB}} &= \left(1 + \tfrac{1}{2}\Delta t\mathcal{L}_u\right)\mathcal{F}(u) \\ &+ \tfrac{1}{2}\Delta x^2\left[\tfrac{1}{3}au_{xxx} - \phi g'(u)u_{xx} - \tfrac{1}{4}(2\phi + c + c^2)g''(u)u_x^2\right] \\ &+ \tfrac{1}{2}\Delta x\Delta t(c+\phi)g''(u)g(u)u_x \\ &+ \Delta t^2\left[\tfrac{1}{6}u_{ttt} - \tfrac{1}{4}g''(u)g^2(u)\right] + \ldots \end{aligned} \tag{18}$$

and

$$T_{\mathsf{BF}} = T_{\mathsf{FB}} + \Delta x\Delta t\left[\tfrac{1}{2}g''(u)u_x - \phi g(u)\right]u_x + \ldots, \tag{19}$$

where all functions are evaluated at (x,t) and the ellipses denote higher order terms. We note that both T_{FB} and T_{BF} are second order when evaluated at the solution $u(x,t)$ of (1), confirming that the methods are second order accurate.

The essence of the method of modified equations is that, for small grid sizes, the partial differential equation obtained by neglecting high order terms in the local truncation error provides a better approximation of the behaviour of the numerical scheme than does the underlying PDE which is the limiting case $\Delta x = \Delta t = 0$. For the FB method, the modified equation we obtain is given by $T_{\mathsf{FB}} = 0$ when the higher order terms (cubic or higher in Δt and Δx) are neglected. This leads to a rather complicated nonlinear PDE that is third order in space and time. We seek travelling wave solutions of this equation in the manner described in the previous section. The equation is linearized about a fixed point $u^\star$ to give

$$\begin{aligned} v_t + av_x = g'(u^\star)v - \tfrac{1}{2}\Delta x^2\left[\tfrac{1}{3}av_{xxx} - \phi g'(u^\star)v_{xx}\right] \\ - \tfrac{1}{6}\Delta t^2 v_{ttt} \end{aligned}$$

and, with $v(x,t) = \exp(-\lambda(x - s_\Delta t))$, we find that the numerical wave speed, s_Δ, satisfies the equation

$$(s_\Delta - a)\lambda - g'(u^\star) = \tfrac{1}{6}h^2\lambda^2\left[\lambda(a - c^2 s_\Delta^3/a^2) + 3\phi g'(u^\star)\right].$$

When $u^\star = u_-^\star$, we can ensure that the numerical wave speed s_Δ coincides with the exact wave speed s given by (17) if the Courant number is chosen so that

$$c^2 = \left(\frac{a}{s}\right)^3\left[3\phi\frac{s}{a} + 1 - 3\phi\right]. \tag{20}$$

Using the relationship (19) we find that the linearized PDE obtained for BF mode is identical to (20). Consequently, the numerical and exact wave speeds are

the same when the Courant number is given by (20). Thus, the optimal Courant numbers are

$$c^2 = \begin{cases} \frac{1}{2}\left(\frac{a}{s}\right)^3\left[3\frac{s}{a}-1\right] & \text{for FBave and BFave : } \phi = \frac{1}{2}, \\ \left(\frac{a}{s}\right)^3 & \text{for FBpoint and BFpoint : } \phi = 0. \end{cases} \tag{21}$$

These results are consistent with the fact that, in the absence of source terms ($g(u) = 0$, so that $s = a$), the optimal Courant number is $c = 1$, since, with constant advective speed, the exact solution is reproduced.

We now present some numerical results that corroborate these findings. The logistic nonlinearity (8) is used and we solve the IBVP on the domain $0 < t \leq 0.2$, $0 < x \leq L$ with initial and boundary data supplied by the travelling wave solution Φ:

$$u(x,0) = \Phi(x - x_0), \qquad u(0,t) = \Phi(-st - x_0),$$

where s is given by (see (17)) $s = a + \alpha/\lambda$. At the right boundary point a first order upwind scheme is employed. In our experiments we have taken $a = 1$, $L = 2.3$ and $x_0 = 0.3$. Results are presented for the parameter sets

Parameter Set 1:	$\alpha = 50,$	$\lambda = 50$	$(s = 2),$
Parameter Set 2:	$\alpha = 100,$	$\lambda = 100$	$(s = 2),$
Parameter Set 3:	$\alpha = 10,$	$\lambda = 50$	$(s = 6),$

where the mesh spacings are determined from given values of r and c by $\Delta t = r/\alpha$ and $\Delta x = \Delta t/c$. The computational domain is allowed to extend, if necessary, beyond $x = L$ by one mesh point. The error in each computed solution is measured in the maximum norm

$$\|u - U\|_\infty \equiv \max_{0<n\Delta t\leq T} \max_{0<j\Delta x\leq L} \left|u(j\Delta x, n\Delta t) - U_j^n\right|, \tag{22}$$

though the conclusions are not strongly dependent on the choice of norm.

The rectangle $(0,2] \times (0,1]$ in (r,c) is discretized by a fine grid (in practice consisting of 32×45 points) and the error $\|u - U\|_\infty$ in the numerical solution is computed for each of the four methods, for each the parameter sets and for each value of (r,c). It is then possible to draw curves of equal accuracy in parameter space, that is, points in the (r,c) plane where

$$\|u - U\|_\infty = 10^{-k/2}, \tag{23}$$

for $k = 2, 3, \ldots, 7$, corresponding roughly to $k/2$ digits of accuracy. The solutions take the form of travelling waves of the type illustrated in Figure 5 having unit amplitude. Thus, values of k smaller than $k = 1$ correspond to grossly inaccurate numerical solutions. A typical situation is illustrated in Figure 6, where the solid curve (computed with BFpoint, parameter set 1) connects points in parameter space which lead to numerical solutions of equal accuracy, in this case $k = 2$ in (23). The computational cost for each problem is proportional to the number of grid points: Cost $\propto (\Delta x \Delta t)^{-1}$, that is

$$\text{Cost} \propto \frac{c}{r^2}.$$

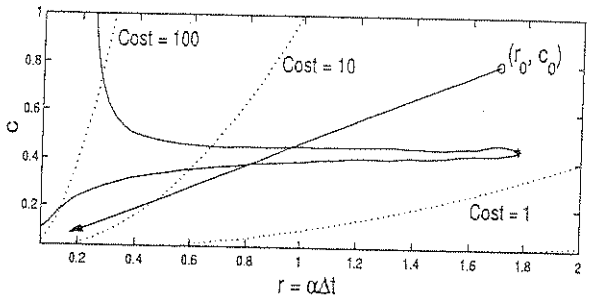


Fig.6: Illustration of curve connecting points of equal accuracy (solid curve), curves of equal cost (dotted) and the refinement path (arrow) resulting from reducing Δt while keeping the spatial grid size Δx fixed.

The (parabolic) curves of level cost are shown dotted in Figure 6 in nominal units. In order to achieve a given accuracy at minimum cost, grids should be chosen to correspond to (r,c) values near the "tip" of the level accuracy curves, shown by an asterisk in Figure 6.

The level accuracy curves for FB methods are shown in Figure 7 and those for BF methods in Figure 8. In each of the four sets of figures it is seen that, for a given level of accuracy (contour), there is a clearly defined Courant number that allows the largest value of Δt to be chosen. This value agrees closely with the optimal c predicted by (20) and (21), which is marked in the plots with a broken horizontal line. It is evident from these results that BF methods are more accurate than FB methods on the same grids. Averaging the source term invokes a larger optimal Courant number and greater computational cost, though it does lead to greater accuracy in BF methods.

5 ADAPTIVE TIME STEPPING

We now look at strategies for dynamically selecting a suitable time-step sequence $\{\Delta t^n\}$, where $\Delta t^n := t^{n+1} - t^n$ and t^n is the nth discrete time level. Our purpose is twofold. First, we wish to ascertain whether adaptive time stepping can prevent the occurrence of spurious solutions of the type discussed in Section 3.

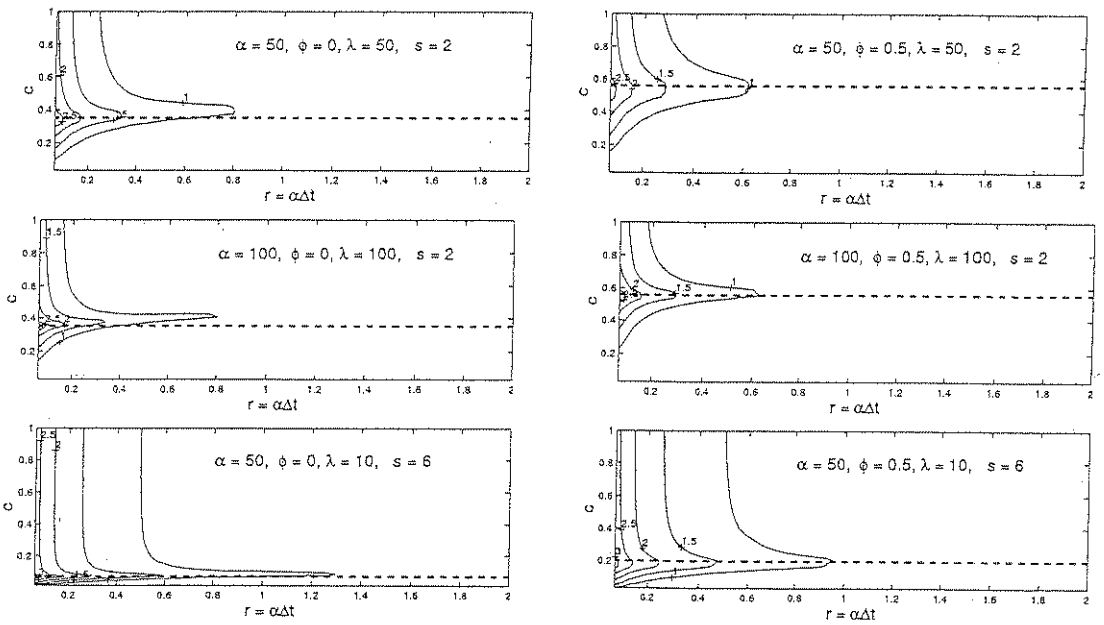


Fig.7: Accuracy contours of FB method with FBpoint on the left and FBave on the right for the three parameter sets.

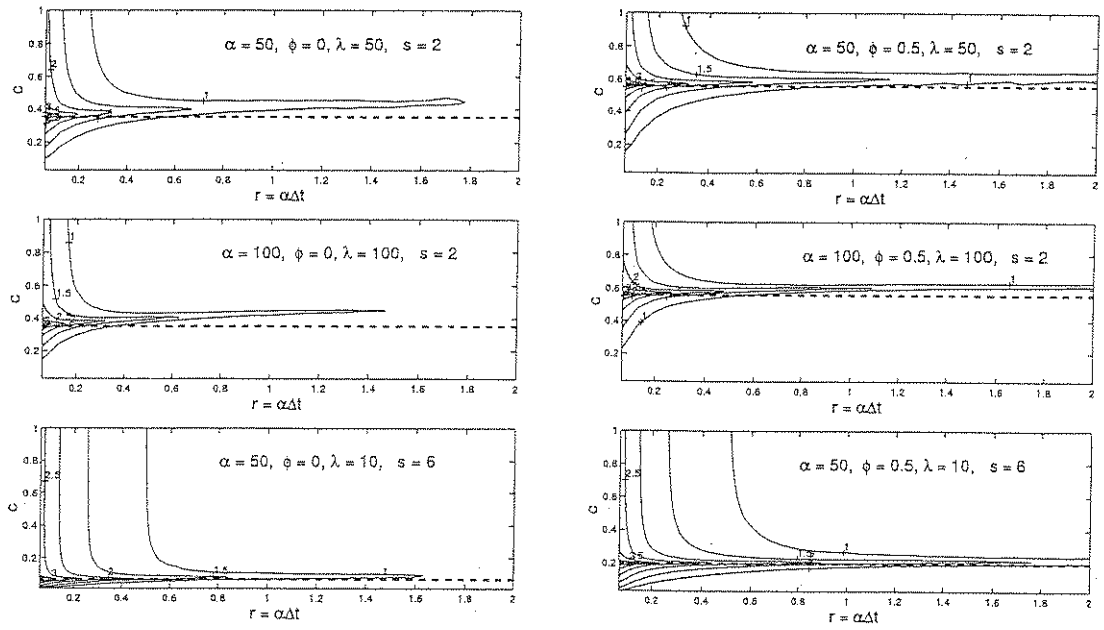


Fig.8: Accuracy contours of BF method with BFpoint on the left and BFave on the right for the three parameter sets.

Second, we wish to consider whether adaptive methods can generate near-optimal grids for travelling wave solutions. We assume throughout this section that the spatial grid remains fixed. Thus, since both c and r are proportional to Δt, the possible refinement paths in the (r, c) plane are straight lines through the origin (illustrated by the arrow in Figure 6). We focus on ways of estimating the local error and give only a brief discussion of implementation issues.

5.1 Richardson Extrapolation

As a starting point, we consider the idea of local error control based on Richardson extrapolation. In a conventional ODE setting the solution U^{n+1} computed using a time-step Δt^n is compared with that using two steps of size $\frac{1}{2}\Delta t^n$ [8, Section 5.10]. The difference is a measure of the local error and the next time-step is increased or decreased according to whether this difference is smaller or larger than some user specified tolerance. For ODEs, this approach is covered by the analysis in [1] (since Richardson extrapolation is a special case of using an RK pair). It follows that spurious solutions $u^\star$ of the type discussed in Section 3 are eliminated for small tolerances. However, Richardson extrapolation is less successful at identifying optimal time-steps. This is due to the fact that it depends on the idea that the local error is proportional to a power of Δt (the $(p+1)$st power for a pth order method), which implies that smaller time-steps lead to smaller (global) errors. For our MacCormack schemes the truncation errors are given by expressions (18) or (19) (with $\mathcal{F}(u) = 0$), depending on the mode. The leading terms are quadratic expressions in Δt (for a fixed Δx) which, in general cannot be driven to zero by reducing Δt to zero. Both the analysis and the numerical results of the preceding section suggest that a time-step Δt may be found for each travelling wave solution which balances (the leading) spatial and temporal errors and we now turn to two possible ways in which this value may be sought.

5.2 An Embedded Method

One of the most popular techniques for controlling the local error in ODE codes is to compute solutions by two different RK methods at each time step. Their difference can then be used to measure the local error and to adjust Δt^n. The two methods are generally of different orders and, for reasons of economy, are embedded so that they share the same intermediate stage values [8, Section 5.10].

Our MacCormack schemes (2)–(6) each use two stages (denoted by Z and W) in which the first stage, Z, is a first order method. Hence, in the spirit outlined above, the difference

$$e_{n+1}(\Delta t) = \max_{0<j\Delta x\le L} \left|U_j^{n+1} - Z_j^{n+1}\right| \qquad (24)$$

may therefore be used as a basis for error control (the maximum norm in space could be replaced by any suitable norm). This process actually estimates the local error in the Z-values, giving a so-called "extrapolation method".

This approach is not successful in FB mode due, we believe, to the fact that the Z-values are generated by an unstable, downwind method. We therefore present analysis and results only for BF mode.

We computed asymptotic estimates in the manner of (18) and (19) for the difference $U_j^{n+1} - Z_j^{n+1}$. Linearizing in the neighbourhood of the leading edge of a travelling wave, we may deduce a value for c that minimizes the local error estimate. By this process we obtain

$$\begin{aligned} c^2 &= \left(\frac{a}{s}\right)^2 \left(2(\frac{s}{a}-1)\phi + 1\right), \\ &= \begin{cases} \frac{a}{s}, & \text{for } \phi = \frac{1}{2}, \\ \left(\frac{a}{s}\right)^2, & \text{for } \phi = 0. \end{cases} \end{aligned}$$

These values do not agree with those given by (21) and hence this process cannot be expected to predict the optimal time-step for travelling waves. Contours of the quantity $\max_{0<n\Delta t<T} |e_{n+1}|$, the estimate of the accuracy, are shown in Figure 9.

Comparing with corresponding contours of the actual global error shown in Figure 8 it is seen that the predicted and actual optimal Courant numbers are not too dissimilar, leading us to conclude that this technique may have some practical use. The agreement between predicted and actual optimal Courant numbers is much closer with BFave than with BFpoint.

It is also observed from Figure 9 that the estimate of the error increases with Δt and indicates that the numerical solution has no correct digits as r approaches $r = 2$, where spurious bifurcations occur. This suggests that the presence of spurious solutions may be detected by monitoring the difference (24).

5.3 A Residual Based Estimator

Our third method of error estimation is based on the idea of using two separate schemes, one of which is of MacCormack type. We choose the second scheme to be the Box scheme which we express as being the solution of the nonlinear equations $R(U^{n+1}) = 0$, where the value of R at the grid point (x_j, t^{n+1}) is defined by [12]

$$\begin{aligned} R(U^{n+1})\big|_j =& \tfrac{1}{2}\left(U_{j+1}^{n+1} + U_j^{n+1} - U_{j+1}^n - U_j^n\right) \\ &+ \tfrac{1}{2}c\left(U_{j+1}^{n+1} - U_j^{n+1} + U_{j+1}^n - U_j^n\right) \\ &- \tfrac{1}{4}\Delta t\left(g(U_{j+1}^{n+1}) + g(U_j^{n+1}) + g(U_{j+1}^n) + g(U_j^n)\right). \end{aligned} \qquad (25)$$

This is an implicit scheme which, like the MacCormack schemes, is a second order accurate approximation of the PDE (1). The need to solve systems of nonlinear algebraic equations at each time step can be avoided by employing the scheme in an non-standard manner. Rather than solving (25) at each time step and using the difference between the resulting solution and that of the MacCormack schemes as an error estimator, we substitute the MacCormack solution into

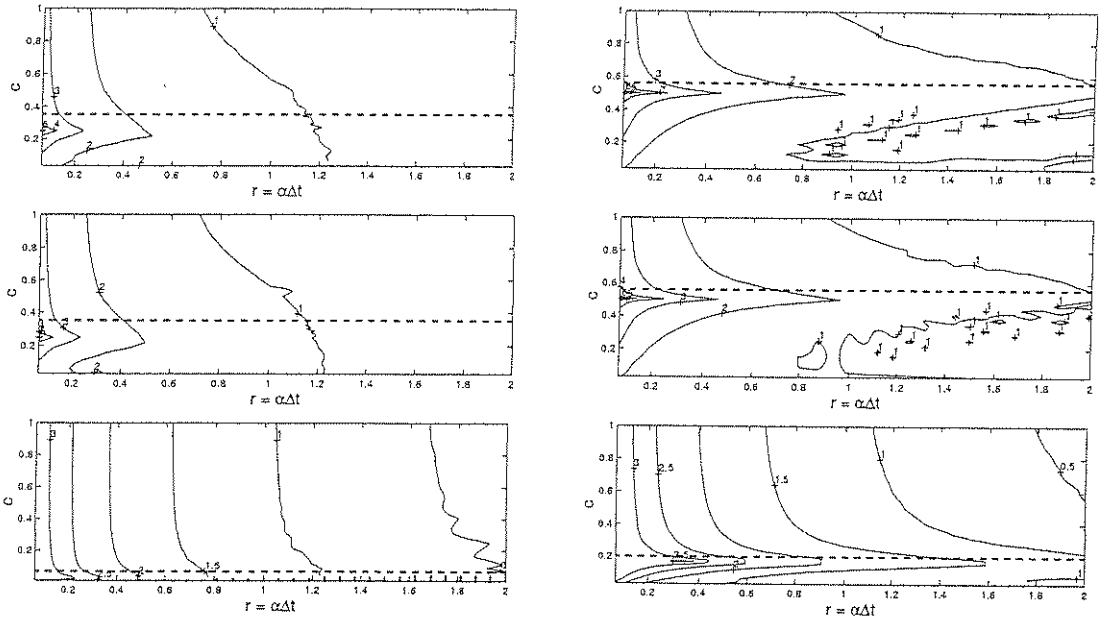


Fig.9: Predicted levels of local error using first-order embedding for the BF method with BFpoint on the left and BFave on the right for the three parameter sets (see Figure 8). The broken lines denote the optimal values of c given by (21).

(25) and use the residual R as the estimate of local error.

Our first observation is that inserting $U_j^n \equiv u^\star$ in (25) gives $g(u^\star) = 0$, so spurious fixed points of the type discussed in Section 3 are disallowed by the Box scheme. Hence, for small tolerances this approach will eliminate such spurious solutions.

An asymptotic estimate of R may be obtained by substituting for U_j^{n+1} and U_{j+1}^{n+1} from (2) and either (3-4) (FB mode) or (5 -6) (BF mode) and then developing the result in Taylor series in the usual manner. This leads to

$$\begin{aligned} R = &\frac{1}{4}\Big[\,\Delta t^2 (g')^2 (3u_x - g) + a\Delta t\Delta x(2\phi + c) g g'' u_x \\ &- \Delta x^2\left[(1-c^2)u_{xxx} + (c + 2c^2 - 2\phi) g''(u_x)^2 \right. \\ &\left. + (3c^2 - 2\phi g' u_{xx})\right]\Big], \end{aligned}$$

where we have ignored higher order terms. Then, for the leading edge of a travelling wave, that is writing $u = u^\star + \exp(-\lambda x)$ near a fixed point $u^\star$ and considering large values of x, we find that the leading term in R vanishes when

$$\begin{aligned} c^2 &= \left(\frac{a}{s}\right)^3 \left(2\phi(\frac{s}{a} - 1) + 1\right), \\ &= \begin{cases} \left(\frac{a}{s}\right)^2 & \text{for } \phi = \frac{1}{2}, \\ \left(\frac{a}{s}\right)^3 & \text{for } \phi = 0. \end{cases} \end{aligned} \tag{26}$$

The same result is also obtained for FB mode. Comparing the values given by (26) with the optimal Courant numbers in (21) it is seen that the value predicted by (26) provides the correct estimate when $\phi = 0$ (point evaluation of the source term) but not when $\phi = \frac{1}{2}$. This is confirmed in Figures 10 and 11 where the results of contours of the norm of the residual $\|R\|_\infty$ are shown for the travelling wave solutions introduced in Section 4.2. The optimal Courant number in each case is shown by the broken lines. For the results of the FBpoint methods shown on the left in Figure 10 the optimal Courant number is discernible though the minimum of the estimator may not be sufficiently clear that it may be exploited for purposes of time-step selection.

The contours of the norm of the residual for BF methods are shown in Figure 11. The optimal Courant numbers predicted by BFpoint for each parameter set are in close agreement with the values based on global errors (broken lines) as expected from our asymptotic result (26). What is less expected is the close agreement for the BFave method. This is perhaps fortuitous and comes about because the numerical values produced by the formulae (21) and (26) are quite close for the wave speeds $s/a = 2, 6$ used in these experiments.

In this section we have seen examples of ways of estimating the local error in MacCormack schemes, the most successful of these being the residual based es-

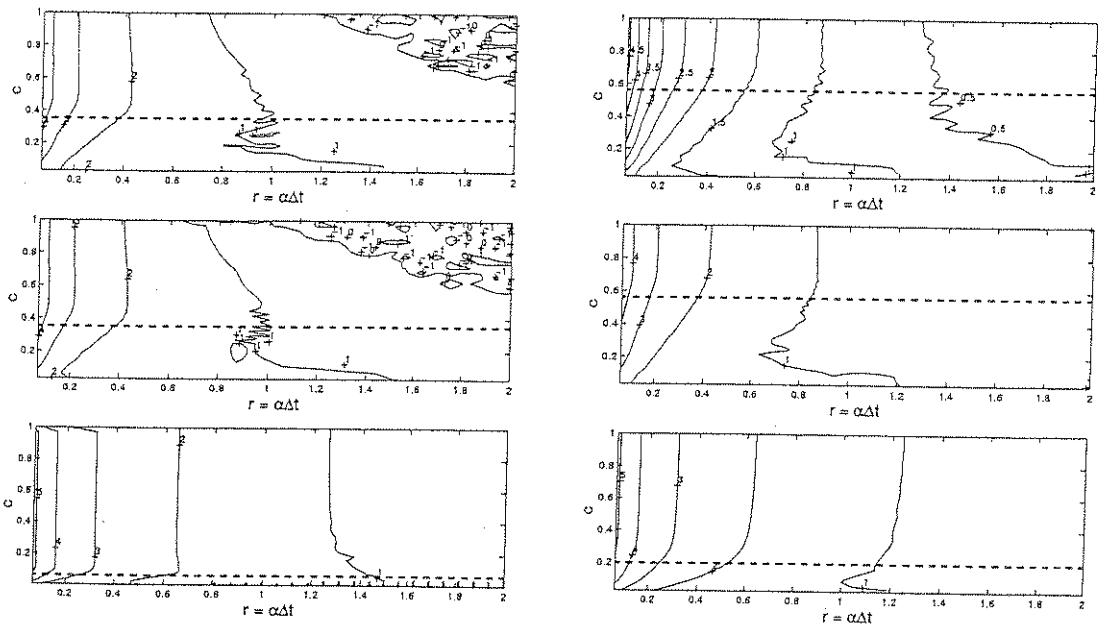


Fig.10: Predicted levels of local error using the norm of the residual in the Box scheme for the FB method with FBpoint on the left and FBave on the right for the three parameter sets (see Figure 7). The broken lines denote the optimal values of c given by (21).

timator for BFave. Exploitation of these estimates in the PDE setting are more complex than in comparable ODE settings since for ODEs it is only necessary to reduce the time-step in order to reduce both local and global errors (provided that the time-step is sufficiently small to lie within the asymptotic region of small step sizes). For MacCormack schemes the leading terms in the local and global errors are quadratic polynomials in both Δt and Δx and the optimal time-step has to be selected so as to minimize the norm of the estimate. No problems are foreseen in devising suitable algorithms for minimizing this norm for relatively small time-steps ($\alpha\Delta t < 1$, say, in the context of our travelling waves) since it is evident from our results that the norm is a smooth function in such regions. It is not so clear how to proceed when the time-steps are so large that the norm of the estimate is no longer smooth. We leave this issue for future work.

6 SUMMARY

We have addressed several issues concerning the use of MacCormack's finite difference scheme for advection-reaction equations. First, we showed that linear stability is improved by averaging the source term in a manner that respects the symmetry of the spatial differencing. We then looked at spurious solutions that are inherited from the underlying RK method. Stable, spatially uniform spurious steady states are admitted on the periodic problem, but these must blow up and lose stability as the mesh is refined. It is also possible for the numerical solution to evolve into a spatially discontinuous pattern that jumps between spurious levels.

We then considered an adaptive version of MacCormack's scheme. Analysis and testing on travelling wave problems supported the use of time-step control. In contrast with the case of embedded RK pairs for ODE problems, use of the stage values of the scheme to form an error estimate that represents the difference of MacCormack's scheme and one that is first order in time and space was unsuccessful. Hence, we developed a time-step selection strategy that attempts to balance the temporal and spatial error contributions based on the residual in the Box scheme. In this manner we regard the fixed spatial mesh as determining the accuracy of the scheme, and we aim to make the time-step as large as possible without significantly degrading the error. The results presented have shown the feasibility of the ideas developed here but further work is necessary in order to create practical implementations. There are many open questions in the area of error control for evolutionary partial differential equations relating to, for example, higher space dimensions, moving meshes, equidistribution and preservation of invariants. However, the idea of balancing errors in time and space is clearly a key element in all these areas.

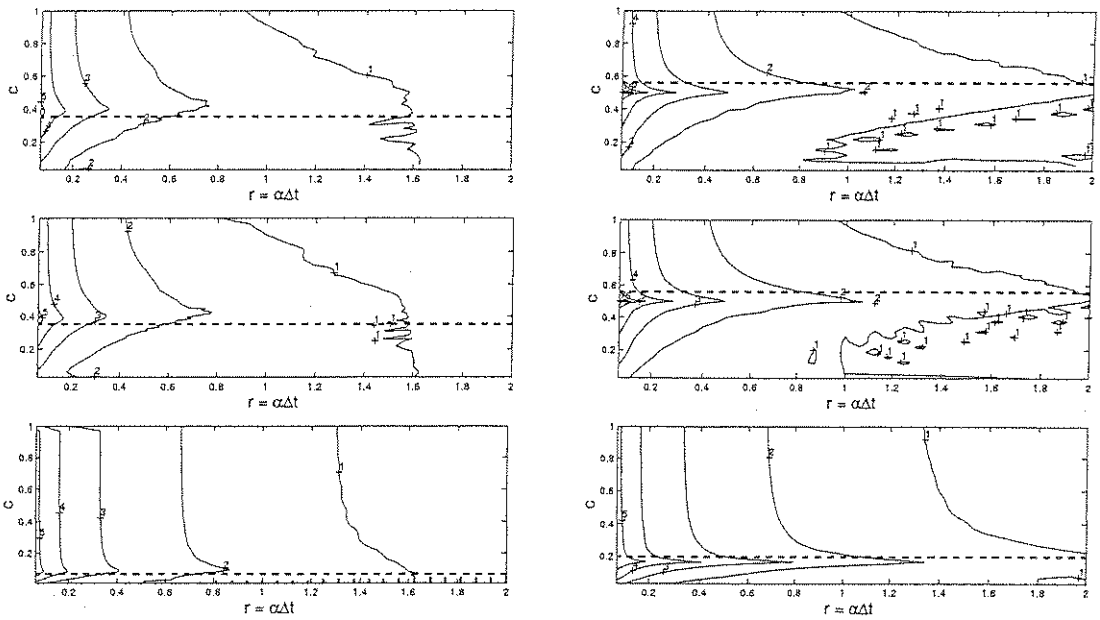


Fig.11: Predicted levels of local error using the norm of the residual in the Box scheme for the BF method with BFpoint on the left and BFave on the right for the three parameter sets (see Figure 8). The broken lines denote the optimal values of c given by (21).

Acknowledgement

We thank Mark Aves, who was involved in some of the initial computational experiments.

REFERENCES

[1] M. A. Aves, D. F. Griffiths, and D. J. Higham, *Does error control suppress spuriosity?*, SIAM J. Numer. Anal., 34 (1997), pp. 756–778.

[2] ———, *Runge-Kutta solutions of a hyperbolic conservation law with source term*, Tech. Rep. 6, Department of Mathematics, University of Strathclyde, 1998. To appear in SIAM J. Sci. Comp.

[3] D. F. Griffiths and D. J. Higham, *Runge-Kutta and MacCormack Dynamics*, Tech. Rep. 18, Department of Mathematics, University of Strathclyde, 1999. To appear in the proceedings of the 8th International Symposium on Computational Fluid Dynamics, 5–10th September, 1999, Bremen, Germany.

[4] D. F. Griffiths and J. M. Sanz-Serna, *On the scope of the method of modified equations*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 994–1008.

[5] D. F. Griffiths, P. K. Sweby, and H. C. Yee, *On spurious asymptotic numerical solutions of explicit Runge-Kutta methods*, IMA J. Numer. Anal., 12 (1992), pp. 319–338.

[6] A. R. Humphries, *Spurious solutions of numerical methods for initial value problems*, IMA J. Numer. Anal., 13 (1993), pp. 263–290.

[7] B. Koren, *A robust upwind discretization method for advection, diffusion and source terms*, in Numerical Methods for Advection-Diffusion Problems. Notes on Numerical Fluid Mechanics, 45, C. B. Vreugdenhil and B. Koren, eds., Vieweg, 1993, pp. 117–138.

[8] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*, Wiley, 1991.

[9] R. J. Leveque and H. C. Yee, *A study of numerical methods for hyperbolic conservation laws with stiff source terms*, J. Comp. Phys., 86 (1990), pp. 187–210.

[10] A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*, Wiley, 1985.

[11] J. D. Murray, *Mathematical Biology*, vol. 19 of Biomathematics, Springer-Verlag, 1989.

[12] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth and Brooks/Cole, 1989.

[13] R. F. Warming and B. J. Hyett, *The modified equation approach to the stability and accuracy analysis of finite difference methods*, J. Comput. Phys., 14 (1974), pp. 159–179.