CONTROL OF THE DISSIPATIVITY OF LAX-WENDROFF TYPE METHODS FOR FIRST ORDER SYSTEMS OF HYPERBOLIC EQUATIONS

by

Tony F. C. Chan and Joseph Oliger

STAN-CS-77-604
MARCH 1977

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

CONTROL OF THE DISSIPATIVITY OF LAX-WENDROFF TYPE METHODS FOR FIRST ORDER SYSTEMS OF HYPERBOLIC EQUATIONS

Tony F. C. Chan and Joseph Oliger*

Computer Science Department
Stanford University
Stanford, California 94305

ABSTRACT

Lax-Wendroff methods for hyperbolic systems have two characteristics which are sometimes troublesome. They are sometimes too dissipative--they may smooth the solution excessively--and their dissipative behavior does not affect all modes of the solution equally. Both of these difficulties can be remedied by adding properly chosen accretive terms. We develop modifications of the Lax-Wendroff method which equilibrate the dissipativity over the fundamental modes of the solution and allow the magnitude of the dissipation to be controlled. We show that these methods are stable for the mixed initial boundary value problem and develop analogous formulations for the two-step Lax-Wendroff and MacCormack methods.

* This work has been supported in part by the National Science Foundation under Grant DCR75-13497 and the Office of Naval Research under Contract N00014-75-C-1132.

1. Introduction

We shall consider approximations for hyperbolic systems of partial differential equations. We begin by considering the Cauchy problem for

$$(1.1) \qquad u_t = Au_x , \quad -\infty < x < \infty , \quad t \geq 0$$

with initial data

$$(1.2) \qquad u(x,0) = u_o(x) , \quad -\infty < x < \infty ,$$

where u is a vector of length n and A is an $n \times n$ matrix with real eigenvalues. We assume that A has a complete set of eigenvectors and can therefore be transformed to diagonal form. We will denote the eigenvalues of A by $\mu_1, \ldots, \mu_n$. If A is a function of x,t we assume that this transformation can be done smoothly.

We shall discuss the well-known Lax-Wendroff method, several of its varients, and modifications thereof. Discussions of these methods and modifications of them which improve their phase errors and stability regions have been carried out by Turkel [9], Gottlieb and Turkel [2], and Eilon, Gottlieb and Zwas [1]. We are going to discuss the dissipative properties of these approximations and modifications of them which improve their dissipative properties. We will also comment on the combination of our modifications with some of those of the previously mentioned authors. Recently, Turkel [10] has discussed a hybrid leap-frog--Lax Wendroff method which is less dissipative than the Lax-Wendroff method but, like the Lax-Wendroff method, has dissipation which varies radically for the various modes of the solution.

To make this more precise we now introduce some notation and definitions which are discussed in Richtmyer and Morton [8] and Kreiss and Oliger [6].

In order to approximate (1.1) we introduce a grid function $v_\nu(t) = v(x_\nu,t)$, $x_\nu = \nu h$, $h > 0$, $\nu = 0, \pm 1, \pm 2, \ldots$ and $t = 0, k, 2k, \ldots, k > 0$. We write our approximations in the form

$$v_\nu(t+k) = \sum_{j=0}^{p} Q_j v_\nu(t-jk) \tag{1.3}$$

where

$$Q_j = \sum_{\ell=-\infty}^{\infty} A_\ell E^\ell \quad \text{and} \quad Ev_\nu = v_{\nu+1}.$$

Associated with (1.3) is the characteristic equation

$$G\varphi \equiv (\varkappa^{p+1} - \sum_{j=0}^{p} \hat{Q}_j \varkappa^j)\varphi = 0 \tag{1.4}$$

where $\hat{Q}_j = \sum_{\ell=-\infty}^{\infty} A_\ell e^{i\ell\xi}$, $\xi = \omega h$. Equation (1.4) is obtained from (1.3) by letting $v(x,t) = \varkappa^{t/k} e^{i\omega x}\varphi(\omega)$. G, as defined by (1.4), is often called the <u>symbol</u> of (1.3).

<u>Definition 1.1.</u> The approximation (1.3) is said to be accurate of order (q_1,q_2) for solutions u of (1.1) if there is a function $C(t)$ which

is bounded on every finite interval $[0,T]$ such that, for all sufficiently small h and k,

$$(1.5)\qquad \|u(x,t+k) - \sum^{p} Q_j u(x,t-jk)\|_{\ell_2} \leq kC(t)(h^{q_1}+k^{q_2}) .$$

If $q_1 = q_2 = q$ we say the approximation (1.3) is accurate of order q.

We can now introduce Kreiss' definition of dissipativity [6, 8].

Definition 1.2. If the solutions $\varkappa_j$ of (1.5) satisfy

$$(1.6)\qquad |\varkappa_j| \leq 1 - \delta|\xi|^{2r} \text{ for } 0 \leq |\xi| \leq \pi$$

for some $\delta > 0$ and natural number r, then the approximation (1.3) is said to be dissipative of order 2r.

One important consequence of dissipativity is the theorem of Kreiss and Parlett [6, 8].

Theorem 1.1. Let (1.3) be accurate of order 2m-2 or 2m-1 and dissipative of order 2m, then (1.3) is strictly stable.

Recall that strict stability implies that the ℓ_2-norm of the approximation, as a function of t, does not grow faster than the L_2-norm of the solution.

The Lax-Wendroff approximation for (1.1) is

$$(1.7)\qquad v_\nu(t+k) = v_\nu(t) + kAD_0 v_\nu(t) + \frac{k^2}{2} A^2 D_+ D_- v_\nu(t)$$

where

$$D_o v_\nu(t) = (2h)^{-1}(v_{\nu+1}(t)-v_{\nu-1}(t))$$

$$D_+ v_\nu(t) = h^{-1}(v_{\nu+1}(t)-v_\nu(t))$$

$$D_- v_\nu(t) = h^{-1}(v_\nu(t)-v_{\nu-1}(t)) .$$

This method is well known to be accurate of order 2 and dissipative of order 4 if $\lambda = k/h$ satisfies $0 < \lambda \max_j |\mu_j| < 1$. The characteristic equation (1.5) can be written as

$$(1.8) \qquad \varkappa_j = 1 + i\lambda\mu_j \sin \xi - 2\lambda^2\mu_j^2 \sin^2(\xi/2) , \qquad j = 1,2,\ldots,n$$

so

$$(1.9) \qquad |\varkappa_j|^2 = 1 - 4\lambda^2\mu_j^2(1-\lambda^2\mu_j^2) \sin^4\xi/2 \equiv 1 - m_j\alpha^4$$

where we let $\alpha = \sin \xi/2$ and $m_j = 4\lambda^2\mu_j^2(1-\lambda^2\mu_j^2)$. This equality yields an inequality of the form (1.6) if $0 < \lambda \max_j |\mu_j| < 1$ but it is more convenient to leave it in this form for our purposes. Note that the m_j are functions of the μ_j and that the "amount of dissipation", the amount $|\varkappa_j|$ differs from 1, is dependent on the eigenvalues of A. The dissipation is greatest for intermediate values of $|\mu_j|$ and least for the smallest and greatest values of $|\mu_j|$. See Fig. 1 where we plot $|\varkappa_j|$ as a function of $\beta_j = \lambda\mu_j$ for several values of ξ. The fact that the dissipation vanishes for all α if $\lambda\mu_j = 0$ is often troublesome in nonlinear calculations. The onset of nonlinear instability is often attributed to these facts, this has been discussed by Richtmyer and Morton [8] beginning on page 334, and by Turkel [9]. It is also often true that the fastest moving waves associated with the largest values of $|\mu_j|$ are often of minimal interest, contain the greatest observational error, and are contaminated by the most computational error. In such situations it would be appropriate to dissipate these modes most rapidly rather than least rapidly. Further, there is generally no reason why intermediate modes should be singled out for most rapid destruction. It is this aspect of Lax-Wendroff methods that we shall discuss. Our subsequent modifications will be directed to the equilibration of dissipativity over the modes (values of μ_j) or, alternatively, to produce decay of $|\varkappa_j|$ as a function of $|\mu_j|$.

Before proceeding we shall introduce the two-step Lax-Wendroff method and the MacCormack method which are also dissipative and behave similarly. The <u>two-step Lax-Wendroff method</u> can be written as

$$(1.10)\qquad \begin{aligned} \bar{v}_\nu &= \frac{v_{\nu+1/2}(t) + v_{\nu-1/2}(t)}{2} + \frac{k}{2}AD_o\left(\frac{h}{2}\right)v_\nu(t) \\ v_\nu(t+k) &= v_\nu(t) + kAD_o\left(\frac{h}{2}\right)\bar{v}_\nu \end{aligned}$$

where $D_o(\frac{h}{2})v_\nu(t) = h^{-1}(v_{\nu+1/2}(t) - v_{\nu-1/2}(t))$.

Our introduction of indices of the form $\nu \pm 1/2$ deviates from our earlier definition of the grid function but the meaning should be clear. The characteristic equation for (1.10) is just (1.8) in the linear case we are considering and the equality (1.9) holds in this case too. In this situation (1.10) is simply a rearrangement of (1.7). However, our modifications of (1.10) will not be the same as those of (1.7) since we may want to take advantage of the separate steps and implement them in a two-step manner. The MacCormack method can be written

(1.11)
$$\bar{v}_\nu = v_\nu(t) + kAD_+ v_\nu(t)$$
$$v_\nu(t+k) = \frac{v_\nu(t) + \bar{v}_\nu}{2} + \frac{k}{2} AD_- \bar{v}_\nu$$

where the $\bar{v}_\nu$ are intermediate values. The characteristic equation can be written

(1.12)
$$\varkappa_j = 1/2 + T + 2i\lambda\mu_j\sqrt{1-\alpha^2}\, T + 2\lambda\mu_j\alpha^2 T$$

where $T = 1/2 + i\lambda\mu_j\sqrt{1-\alpha^2}\,\alpha - \lambda\mu_j\alpha^2$.

$\varkappa_j$ can again be seen to satisfy (1.8) and (1.9). Again, in this simple situation, this is a rearrangement of (1.7) but our modifications will again be different since we will implement them in a two-step manner.

2. The Modified Methods and Their Properties

A nondissipative method like the leap-frog method,

$$v_\nu(t+k) = v_\nu(t-k) + 2kAD_0 v_\nu(t) ,$$

can be modified to yield a dissipative method by adding dissipative terms, e.g.,

$$v_\nu(t+k) = v_\nu(t-k) + 2kAD_0 v_\nu(t) - \epsilon \frac{h^4}{16} (D_+D_-)^2 v_\nu(t-k)$$

is accurate of order 2 and dissipative of order 4 for $0 < \epsilon < 1$, $|\lambda| \le 1 - \epsilon$ [6]. The eigenvalues of the symbol of this method satisfy

$$|\kappa_j|^2 = 1 - \epsilon \sin^4(\frac{\xi}{2})$$

so the amount of dissipation, the magnitude of δ in (1.6), can be controlled by varying ϵ.

We can similarly modify the Lax-Wendroff methods by adding terms-- we can reduce the amount of dissipation by adding accretive terms. Such terms must be of the order of the truncation error of the method so that they do not constitute a modification of the differential equation and do not affect the order of accuracy and the rate of convergence of the method.

We first consider a modification of the Lax-Wendroff method (1.7). Let M_1, M_2 and M_3 be arbitrary matrices which are diagonalizable by the same transformation which diagonalizes A. We consider

$$v_\nu(t+k) = (I + kAD_0 + \frac{k^2}{2} A^2 D_+ D\ ) v_\nu(t) +$$

(2.1)

$$(M_1 \frac{h^4}{16} (D_+ D\ )^2 + M_2 \frac{h^3}{8} D_+^2 D + M_3 \frac{h^3}{8} D_-^2 D_+) v_\nu(t) \ .$$

The symbol of (2.1) is

$$G = 1 + 2i\lambda A\alpha\sqrt{1-\alpha^2} - 2\lambda^2 A^2 \alpha^2 + M_1 \alpha^4$$

(2.2)

$$+ M_2(-i\sqrt{1-\alpha^2}\,\alpha^3 + \alpha^4) + M_3(-i\sqrt{1-\alpha^2}\,\alpha^3 - \alpha^4)$$

where $\alpha = \sin \xi/2$ and $A = k/h$. We also have

$$\mathrm{Re}(G) = 1 - 2\lambda^2 A^2 \alpha^2 + (M_1 + M_2 - M_3)\alpha^4 \ ,$$

$$\mathrm{Im}(G) = \sqrt{1-\alpha^2}\,\alpha(2\lambda A - (M_2 + M_3)\alpha^2) \ ,$$

and

$$|\kappa_j|^2 = 1 + [4\lambda^4 \mu_j^4 - 4\lambda^2 \mu_j^2 \ (-4m_{2j} - 4m_{3j})\lambda\mu_j$$

$$+ 2m_{1j} - 2m_{2j} - 2m_{3j}]\alpha^4 +$$

(2.3)

$$[(-4m_{1j} - 4m_{2j} + 4m_{3j})\lambda^2\mu_j^2 + (4m_{2j} + 4m_{3j})\lambda\mu_j$$

$$+ m_{2j}^2 + 2m_{3j}m_{2j} + m_{3j}^2]\alpha^6 +$$

$$[m_{1j}^2 + (2m_{2j} - 2m_{3j})m_{1j} - 4m_{3j}m_{2j}]\alpha^8$$

where the $m_{\ell j}$'s are the diagonal entries which result when M_ℓ is transformed to diagonal form with A.

We now define the phase error per time step, E, of an approximation as (see [9]):

E = (approximations phase speed - solutions 'phase speed) × k

Then it follows that [9]

(2.4) $$E = \arctan((\mathrm{Im}\ G)(\mathrm{Re}\ G)^{-1}) - \lambda A\xi .$$

The components of $E = \mathrm{diag}(e_1,\ldots,e_n)$ for our diagonalized system are

(2.5) $$e_j = \frac{1}{24}(4\lambda^3\mu_j^3 - 4\lambda\mu_j - 3m_{2j} - 3m_{3j})\xi^3 + O(\xi^5) .$$

We now consider two specific modifications. Let $M_2 = M_3 = 0$. If we take

(2.6) $$2M_1 = [-\epsilon I + 4\lambda^2 A^2(I-\lambda^2A^2)]$$

then (2.3) becomes

(2.7) $$|\varkappa_j|^2 = 1 - [\epsilon + 4m_{1j}\alpha^2\beta_j^2 - m_{1j}^2\alpha^4]\alpha^4$$

where

$$m_{1j} = -\epsilon + 4\beta_j^2(1-\beta_j)^2 \quad \text{and} \quad \beta_j = \lambda\mu_j .$$

We have thus cancelled out the μ_j dependence of the coefficient of α^4. We demonstrate the effect of this in Figures 2a-2d where we plot $|\kappa_j|$ as a function of β_j for several values of ξ and ϵ. For smaller values of ϵ the dissipation is reduced and nearly constant for a considerably larger neighborhood of $\beta = 0$. For larger values of ϵ the dissipation does not vanish for all ξ in the neighborhood of $\beta = 0$.

We next consider a modification which introduces a quadratic decay in $|\kappa_j|$ as a function of μ_j. We take

(2.8) $$2M_1 = [-\delta\lambda^2A^2 + 4\lambda^2A^2(I-\lambda^2A^2)]$$

Equation (2.3) is now

(2.9) $$|\kappa_j|^2 = 1 - [\delta\beta_j^2+4m_{1j}\alpha^2\beta_j^2-m_{1j}^2\alpha^4]\alpha^4$$

where

$$m_{1j} = -\delta\beta_j^2 + 4\beta_j^2(1-\beta_j^2) .$$

We demonstrate the effect of this modification in Figures 3a-3d where we plot $|\kappa_j|$ as a function of β_j for several values of ξ and δ as before.

The stability of the modified methods given in (2.6) and (2.9) follows from our general results in Section 3. We will state the particular form those results take for these methods here. The modification (2.6) is strictly stable if

(2.10) $$0 < \epsilon < 4(1-|\beta_{max}|^4)$$

where β_{max} is the value of $\beta_j = \lambda\mu_j$ with largest magnitude. The modification (2.9) is strictly stable if

$$(2.11) \qquad 0 < \delta < \frac{4(1-|\beta_{max}|^4)}{|\beta_{max}|^2} .$$

We could have introduced a term $-\delta\lambda A$ in (2.9) instead of the term $-\delta\lambda^2 A^2$ and this would result in a linear decay of $|\varkappa_j|^2$ with respect to β_j in the α^4 term. However, the stability analysis corresponding to (2.11) would yield $0 < \delta\beta_j < 4(1-\beta_j^4)$ for $j = 1,2,\ldots,n$ so that δ would have to be chosen to agree in sign with β_j, i.e., δI would need to be replaced by a matrix which, after transformation to diagonal form, would yield a matrix with entries $\pm\,\delta$ in appropriate places to match up with the μ_j. If all the μ_j are of the same sign this 'poses no problem; however, if they are of both signs this does not yield a practical procedure in general.

We now note that the first term of our expression (2.5) for the phase error is not affected by our choice of Ml, so M_2 and M_3 can be chosen as in Gottlieb and Turkel [2] to reduce the phase error to $O(\xi^5)$. However, these modifications do affect our expression for $|\varkappa_j|^2$ given in (2.3) and the results (2.10) and (2.11) no longer hold. (2.3) must be reexamined when nonzero M_2 and M_3 are used to establish the stability bounds for ϵ, δ, $|\beta_{max}|$, etc.

We now consider modifications of the two-step Lax-Wendroff method. We begin by noting that the modifications of the Lax-Wendroff method which we have already discussed can all be used in the second step of the two-step method as given in (1.10). Since the symbol is unchanged all our previous results hold for such modifications. We next consider modifications to the first step of (1.10) as given by

$$\bar{v}_\nu = \frac{v_{\nu+1/2}(t) + v_{\nu-1/2}(t)}{2} + \frac{k}{2} D_0(\tfrac{h}{2})v_\nu(t) - \frac{1}{16} M_1 h^3 D_0^3(\tfrac{h}{2})v_\nu(t)$$

(2.12)

$$v_\nu(t+k) = v_\nu(t) + kAD_0(\tfrac{h}{2})\bar{v}_\nu + \frac{1}{16} M_2 h^4 D_0^4(\tfrac{h}{2})v_\nu(t)$$

The symbol for this method is

(2.13)

$$G = 1 + 2\lambda A i\alpha[\sqrt{1-\alpha^2} + i(\lambda A\alpha + M_1 \tfrac{\alpha^3}{2})] + M_2\alpha^4 ,$$

$$\mathrm{Re}(G) = 1 - 2\lambda^2 A^2 \alpha^2 - M_1 \lambda A \alpha^4 + M_2 \alpha^4 ,$$

$$\mathrm{Im}(G) = 2\lambda A\alpha\sqrt{1-\alpha^2} ,$$

and

(2.14)

$$|\kappa_j|^2 = 1 + [4\lambda^4\mu_j^4 - 4\lambda^2\mu_j^2 - 2m_{1j}\lambda\mu_j + 2m_{2j}]\alpha^4$$
$$- [-4m_{1j}\lambda^3\mu_j^3 + 4m_{2j}\lambda^2\mu_j^2]\alpha^6$$
$$- [-m_{1j}^2 A^2 \mu_j + 2m_{2j}m_{1j}\lambda\mu_j - m_{2j}^2]\alpha^8 .$$

The phase error is now given by

(2.15)

$$e_j = \frac{1}{6}(\lambda^3\mu_j^3 - \lambda\mu_j)\xi^3 + O(\xi^5) .$$

We again note that e_j is not changed through terms of order ξ^3

and that the *phase error modification schemes of Gottlieb and Turkel can again be applied with our modifications in a straightforward manner.

If we choose

$$2M_1 = 4\lambda A(I-\lambda^2A^2) + \delta\lambda A \tag{2.16}$$

and $M_2 = 0$ we have

$$|\kappa_j|^2 = 1 - \{\delta\beta_j^2 - 4m_{1j}\beta_j^3\alpha^2 - m_{1j}^2\beta_j\alpha^4\}\alpha^4$$

(2.17)

with

$$2m_{1j} = \delta\beta_j - 4\beta_j(1-\beta_j^2) \quad .$$

Notice that in this case we have a quadratic decay of $|\kappa_j|$ with β_j in the α^4 term. Thus, this method is similar to our earlier modification (2.8). In fact, (2.17) is identical to (2.9). Cur remarks following (2.9) about linear decay modifications also apply here.

The stability of the modification (2.16) again follows from our general results in Section 3. This method is stable if the inequality (2.11) is satisfied.

We now turn to modifications of the MacCormack scheme. For the same reason as before for the two-step Lax-Wendroff method, we only consider modifications to the first step of (1.11). Modifications to the second step will be the same as those for Lax-Wendroff. We consider

$$\bar{v}_\nu = [I + kAD_+ + M_1 h^3 D_+^2 D_-]v_\nu(t) ,$$

(2.18)

$$v_\nu(t+k) = \frac{v_\nu(t) + \bar{v}_\nu}{2} + \frac{k}{2} A D_- \bar{v}_\nu .$$

The symbol for this method is

(2.19) $$G = \frac{1}{2} + T + 2i\lambda A\alpha\sqrt{1-\alpha^2}\, T + 2\lambda A\alpha^2 T$$

where

$$T = \frac{1}{2} + i\lambda A\sqrt{1-\alpha^2} - \lambda A\alpha^2 - 4iM_1\alpha^3\sqrt{1-\alpha^2} + 4M_1\alpha^4 ,$$

$$Re(G) = 1 + \alpha^2(-2\lambda^2A^2)+\alpha^4(4M_1 + 8\lambda AM_1) ,$$

$$Im(G) = \alpha\sqrt{1-\alpha^2}\,[2\lambda A + \alpha^2(-4M_1)] ,$$

and

(2.20)

$$\begin{aligned} |\varkappa_j|^2 = 1 &+ [4\lambda^4\mu_j^4 - 4\lambda^2\mu_j^2 + 8m_{1j}]\alpha^4 \\ &+ [-32m_{1j}\lambda^3\mu_j^3 - 16m_{1j}\lambda^2\mu_j^2 + 16m_{1j}\lambda\mu_j + 16m_{1j}^2]\alpha^6 \\ &+ [64m_{1j}^2\lambda^2\mu_j^2 + 64m_{1j}^2\lambda\mu_j]\alpha^8 \end{aligned}$$

The phase error is now given by

(2.21) $$e_j = \frac{1}{6}(\lambda^3\mu_j^3 - \lambda\mu_j - 3m_{1j})\xi^3 + O(\xi^5) .$$

Notice that the $O(\xi^3)$ term of the phase error is affected by our choice of M_1 this time. Simultaneous modifications to improve the phase error are not so easy to carry out for this method.

If we choose

(2.22) $$8M_1 = 4\lambda^2 A^2(I-\lambda^2 A^2) - \epsilon I$$

we have

$$|\kappa_j|^2 = 1 - \epsilon\alpha^4 + O(\alpha^6)$$

We illustrate the effect of this modification in Figures 4a-4d where we again plot $|\kappa_j|$ as a function of β_j for selected values of ξ and ϵ.

We can also choose

(2.23) $$8M_1 = 4\lambda^2 A^2(I-\lambda^2 A^2) - \delta\lambda^2 A^2$$

to introduce a quadratic decay so that

$$|\kappa_j|^2 = 1 - \delta\beta_j^2\alpha^4 + O(\alpha^6) \quad .$$

We have also plotted this in Figures 5a-5d.

The situation here is unlike that arising from our previous modifications. Previously the stability limit for $|\beta_{max}|$ has increased as ϵ decreased to zero and, in fact, approached 1 as ϵ tended to zero. Here the stability limit tends to zero as ϵ tends to zero. From the plots in Figures 4a-4d it seems that we should use ϵ somewhere between 0.1 and 2 in order to have a reasonable stability limit for $|\beta_{max}|$. The same comment applies to the MacCormack scheme with quadratic decay.

The case shown in Fig. 4b seems to achieve our goal of equilibration quite well but increases the dissipation.

In summary we remark that modifications of the second step analogous to those used with Lax-Wendroff are much more successful for both the two-step Lax-Wendroff and the MacCormack methods.

3. Stability of the Modified Methods

First of all, note that all the unmodified methods we have considered, namely, the one-step Lax-Wendroff, the two-step Lax-Wendroff and the MacCormack schemes, have the same κ_j and are all dissipative of order 4 (and therefore strictly stable by Theorem 1.1) if $|\beta_{max}| < 1$. For each of our modified schemes, we can write the modified κ_j^2 in the following form:

$$\kappa_j^2 = 1 - \alpha^4 F(\beta_j, m_{1j}, m_{2j}, m_{3j}, \alpha^2) \tag{3.1}$$

By Theorem 1.1, our modified schemes will be strictly stable if, for each j,

$$F(\beta_j, m_{1j}, m_{2j}, m_{3j}, \alpha^2) > 0 \quad \text{for} \quad \alpha^2 \in [0,1] \ . \tag{3.2}$$

Since $\max_{0 \le |\xi| \le \pi} \frac{|\xi|}{|\sin \frac{\xi}{2}|} = \pi$, (3.2) will imply (1.6).

Usually, the $m_{\ell j}$'s are functions of β_j. Hence (3.2) defines a stability limit for β_j. It is difficult to determine this stability limit for general functions $m_{\ell j}(\beta_j)$. However, for specific functions $m_{\ell j}(\beta_j)$, this stability limit can always be determined numerically as in our plots of $|\kappa_j|$.

Consider the modified one-step Lax-Wendroff scheme (2.1) with $M_2 = M_3 = 0$. The following theorem gives the conditions which the $m_{\ell j}$'s will have to satisfy in order to guarantee stability.

Theorem 3.1. Consider the scheme (2.1) with $M_2 = M_3 = 0$.

(3.3) If $|\beta_{max}| < 1$ and $-2(1-\beta_j^2) < m_{1j} < 2\beta_j^2(1-\beta_j^2)$ for all j ,

then the resulting scheme is dissipative of order 4 and hence strictly stable by Theorem 1.1.

Proof: Condition (3.2) implies that the following has to be satisfied:

$$(3.4) \qquad F \equiv -m_{1j}^2\alpha^4 + m_{1j}(4\beta_j^2\alpha^2-2) + 4\beta_j^2(1-\beta_j^2) > 0 \qquad \text{for all } j .$$

Since F is a quadratic in m_{1j} with negative leading coefficients, (3.4) will be satisfied if (see Fig. 6), for every j,

(i) F has real roots, say $m_L(\alpha^2)$ and $m_S(\alpha^2)$ for the larger and the smaller root, respectively (considered as functions of α^2), and

$$(3.5) \quad \text{(ii)} \quad m_S(\alpha^2) < m_{1j} < m_L(\alpha^2) \quad \text{for all } \alpha^2\in[0,1].$$

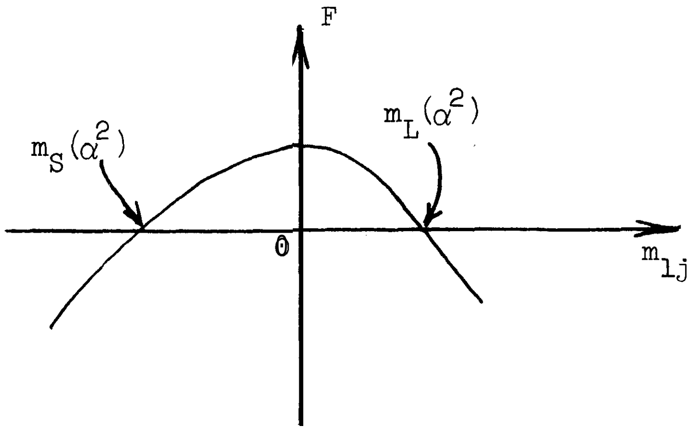

Fig. b

Now

$$m_L(\alpha^2) = [-1 + 2\beta_j^2\alpha^2 + \sqrt{1 + 4\alpha^4\beta_j^2 - 4\alpha^2\beta_j^2}\,]/\alpha^4 , \quad \text{if} \quad \alpha^2 \neq 0$$

$$(3.6) \qquad - 2\beta_j^2(1-\beta_j^2) \qquad \text{if} \quad \alpha^2 = 0$$

$$m_S(\alpha^2) = [-1 + 2\beta_j^2\alpha^2 - \sqrt{1 + 4\alpha^4\beta_j^2 - 4\alpha^2\beta_j^2}\,]/\alpha^4 , \quad \text{if} \quad \alpha^2 \neq 0$$

$$= \text{undefined} \qquad \text{if} \quad \alpha^2 = 0$$

Hence the roots will be real if

$$1 + 4\alpha^4\beta_j^2 - 4\alpha^2\beta_j^2 > 0 \qquad \text{for all } j \text{ and for all } \alpha^2 \epsilon [0,1]$$

or if

$$\beta_j^2 < \frac{1}{4\alpha^2(1-\alpha^2)} \qquad \text{for all } j \text{ and for all } \alpha^2 \epsilon [0,1]$$

or if

$$(3.7) \qquad \beta_j^2 < \min_{\alpha^2 \leq 1} \frac{1}{4\alpha^2(1-\alpha^2)} = 1 \qquad \text{for all } j ,$$

i.e., if $|\beta_{max}| < 1$.

Note that (3.7) is just the unmodified stability limit. Now, by straightforward differentiation, we can show that

$$(3.8) \qquad \frac{dm_S(\alpha^2)}{d\alpha^2} \geq 0 \text{ and } \frac{dm_L(\alpha^2)}{d\alpha^2} \geq 0 \text{ for } \alpha^2 \epsilon [0,1] \text{ if } |\beta_{max}| \text{ cl } .$$

Hence, (3.5) will be satisfied if

$$-2(1-\beta_j^2) \equiv m_S(1) < m_{1j} < m,(0) \equiv 2\beta_j^2(1-\beta_j^2) \quad \text{fbr all } j$$

which is (3.3). □

Corollary 3.1: The modification (2.6) is strictly stable if (2.10) is satisfied.

Proof: By the previous theorem, the following has to be satisfied:

$$-2(1-\beta_j^2) < -\frac{\epsilon}{2} + 2\beta_j^2(1-\beta_j^2) < 2\beta_j^2(1-\beta_j^2) \text{ for all } j$$

which reduces to

$$4(1-\beta_j^4) > \epsilon > 0 \text{ for all } j \tag{3.9}$$

Note that (3.9) automatically implies $|\beta_{max}| < 1$. Also, (2.10) easily follows from (3.9). □

Corollary 3.2: The modification (2.9) is strictly stable if (2.11) is satisfied.

Proof: Follows immediately from Theorem 3.1. □

Next, consider the two-step Lax-Wendroff modification (2.12) with $M_2 = 0$. The F we obtain in this case is very similar to that which we obtained for the Lax-Wendroff method, (3.4). We only need to replace m_{1j} in (3.4) by $m_{1j}\beta_j$ to obtain the correct F. Thus, we easily obtain the following theorem.

Theorem 3.2: Consider the scheme (2.12) with $M_2 = 0$. If

$$|\beta_{max}| < 1$$

(3.10)

$$\text{and} \quad -2\beta_j^2(1-\beta_j^2) < m_{1j}\beta_j < 2(1-\beta_j^2) \quad \text{for all } j ,$$

then the resulting scheme is dissipative of order 4 and hence strictly stable by Theorem 1.1.

Corollary 3.3: The modification (2.16) is strictly stable if (2.11) is satisfied.

A stability analysis is more difficult for modifications of the first step of the MacCormack method. No condition analogous to the condition (3.8) holds in this case. We have not been able to obtain clean conditions like those in Theorems 3.1 and 3.2 for the MacCormack scheme. However, given a specific M_1 one can easily determine the stability interval by examining $|\kappa_j|$ as a function of β_j.

We finally remark that the previous sufficient conditions for stability are all necessary if we allow equality. This easily follows from the von Neumann necessary condition [6,8].

4. The Initial Boundarv Value Problem

We now consider the problem of approximating equation (1.1) on a bounded x- interval $a \leq x \leq b$. Since the modified methods we have discussed have larger stencils (involve more neighboring points) than their unmodified counterparts there are more points at the ends of the interval [a,b] where these approximations cannot be used than there are with the original methods. However, this problem is easily avoided.

Stable approximations for the initial boundary-value problem for the Lax-Wendroff, two-step Lax-Wendroff, and MacCormack methods are discussed by Gustaffson, et. al. [4] and by Gottlieb and Turkel [3]. We will base our methods on these.

Assumptions

We assume that boundary conditions are given at the points a and b which yield a well posed problem for (1.1), see Kreiss [5]. We further assume that stable approximations for this problem are known for the underlying method that we are modifying, see [3, 4] for candidates, and finally that the mesh ratio λ and modification parameters ϵ, 6, etc., are chosen so that both the modified and unmodified methods are stable for the Cauchy problem.

We now form our approximations for the initial boundary-value problem by coupling the unmodified method with its stable boundary conditions to the modified method in the neighborhood of the boundary points a and b in the manner discussed by Oliger [7].

The Methods

We will use the desired modified method at all interior net points where it can be used, we then drop the modification (set $M_1 = M_2 = M_3 = 0$) and use the underlying method at all those points in the neighborhood of a and b where it can be used (at most one or two points at each end), the remaining points (only a and b) are then treated using the stable boundary approximation.

Theorem 4.1. The methods proposed above are stable in the sense of Gustaffson, et. al. [4] (definition 3.3) if our assumptions hold.

Proof: This result follows immediately from Theorem 2.4 of [7]. □

ACKNOWLEDGEMENT

We want to acknowledge our use of the MACSYMA symbol manipulation system developed by the Mathlab group at MIT and supported by the Defense Advanced Research Projects Agency work order 2095 and by the Office of Naval Research Contract N00014-75-C-0661. We also want to thank Eric Grosse for his help with the graphics.

REFERENCES

1. B. Eilon, D. Gottlieb and G. Zwas, "Numerical stabilizers and computing time for second-order accurate schemes," J. of Comp. Phys., v. 9, 1972, pp. 387-397.

2. D. Gottlieb and E. Turkel, "Phase error and stability of second order methods for hyperbolic problems. II," J. of Comp., v. 15, 1974, pp. 251-265.

3. D. Gottlieb and E. Turkel, "Boundary conditions for multistep finite difference methods for time dependent equations," ICASE Report No. 76-30, Sept. 1976.

4. B. Gustafsson, H.-O. Kreiss, and A. Sundström, "Stability theorey of difference approximations for mixed initial boundary-value problems. II," Math. Comp., v. 26, 1972, pp. 649-686.

5. H.-O. Kreiss, "Initial boundary value problems for hyperbolic equations," Comm. Pure and Appl. Math., v. 23, 1970, pp. 277-298.

6. H.-O. Kreiss and J. Oliger, "Methods for the approximate solution of time dependent problems," GARP Publications series No. 10, World Meteorological Organization, Geneva, 1973.

7. J. Oliger, "Hybrid difference methods for the initial boundary-value problem for hyperbolic equations," Math. Comp., v. 30, 1976, pp. 724-738.

8. R. D. Richtmyer and K. W. Morton, Difference methods for initial-value problems, 2nd ed., Interscience Publishers, New York, 1967.

9. E. Turkel, "Phase error and stability of second order methods for hyperbolic problems. I," J. of Comp. Phys., V. 15, 1974, pp. 226-250.

10. E. Turkel, "Composite methods for hyperbolic equations," ICASE Report No. 75-19, Sept. 1975.

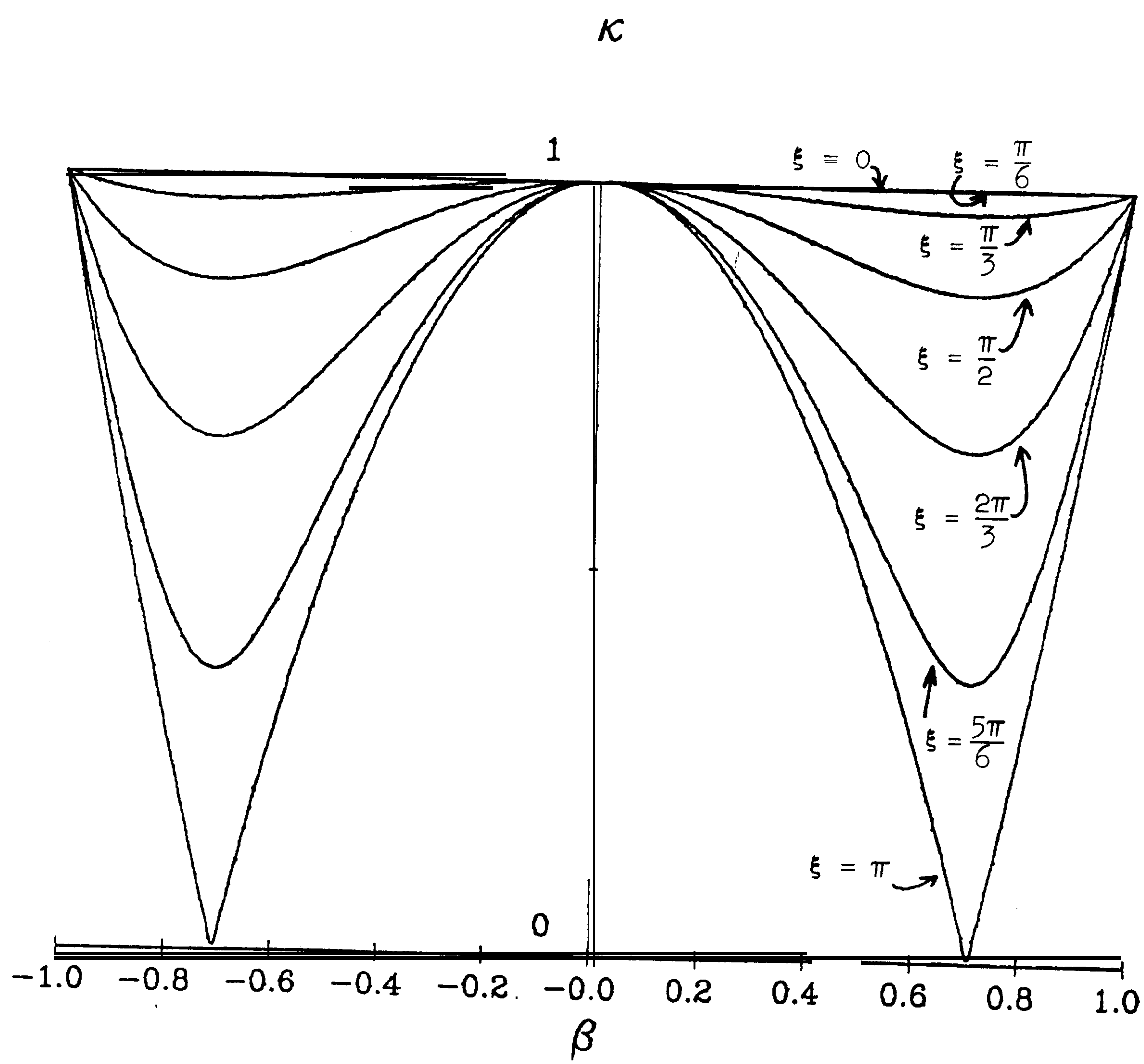

Fig 1 Lax-Wendroff

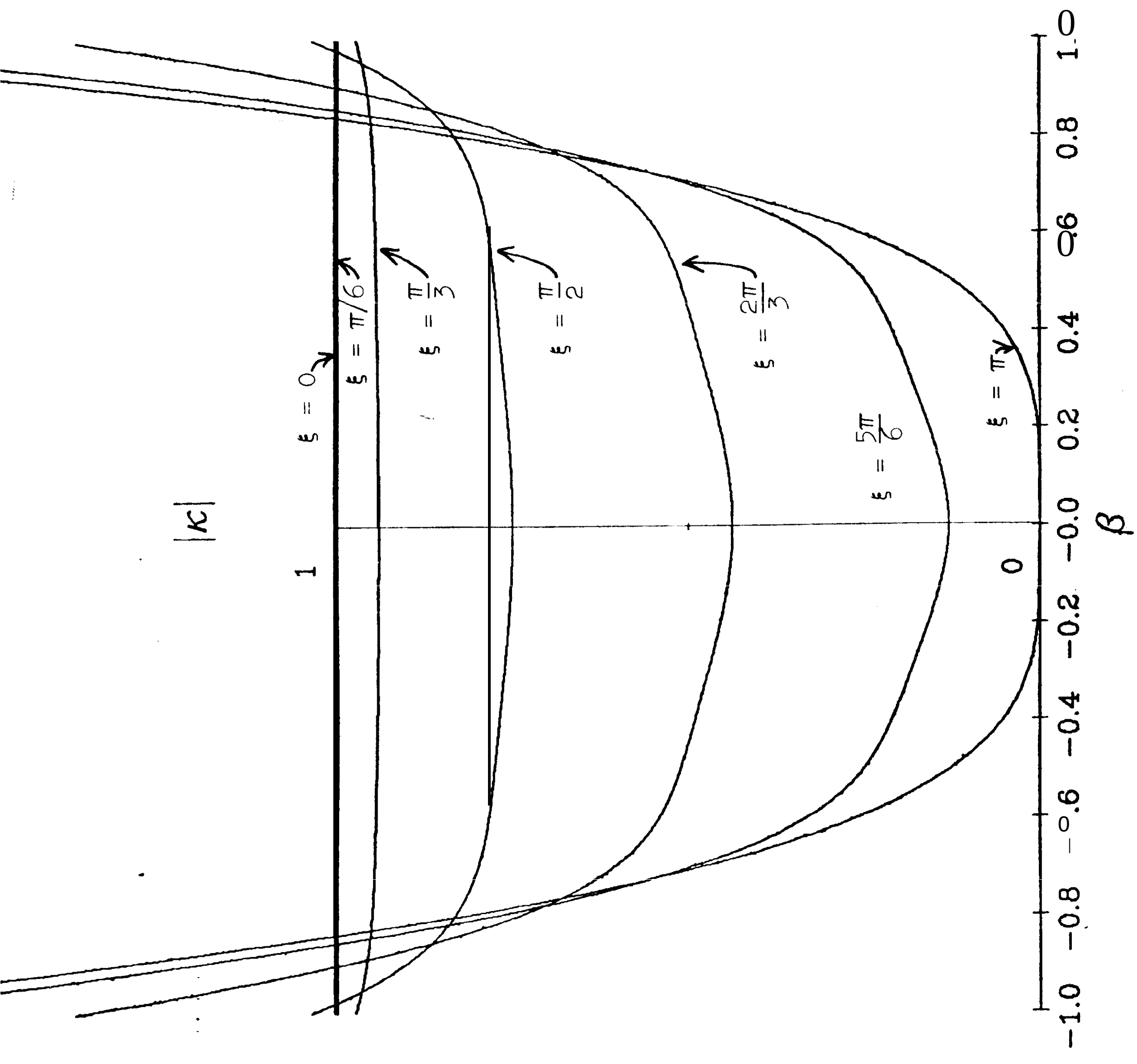

Fig. 2a Modified Lax-Wendroff $\epsilon = 2\ 0$

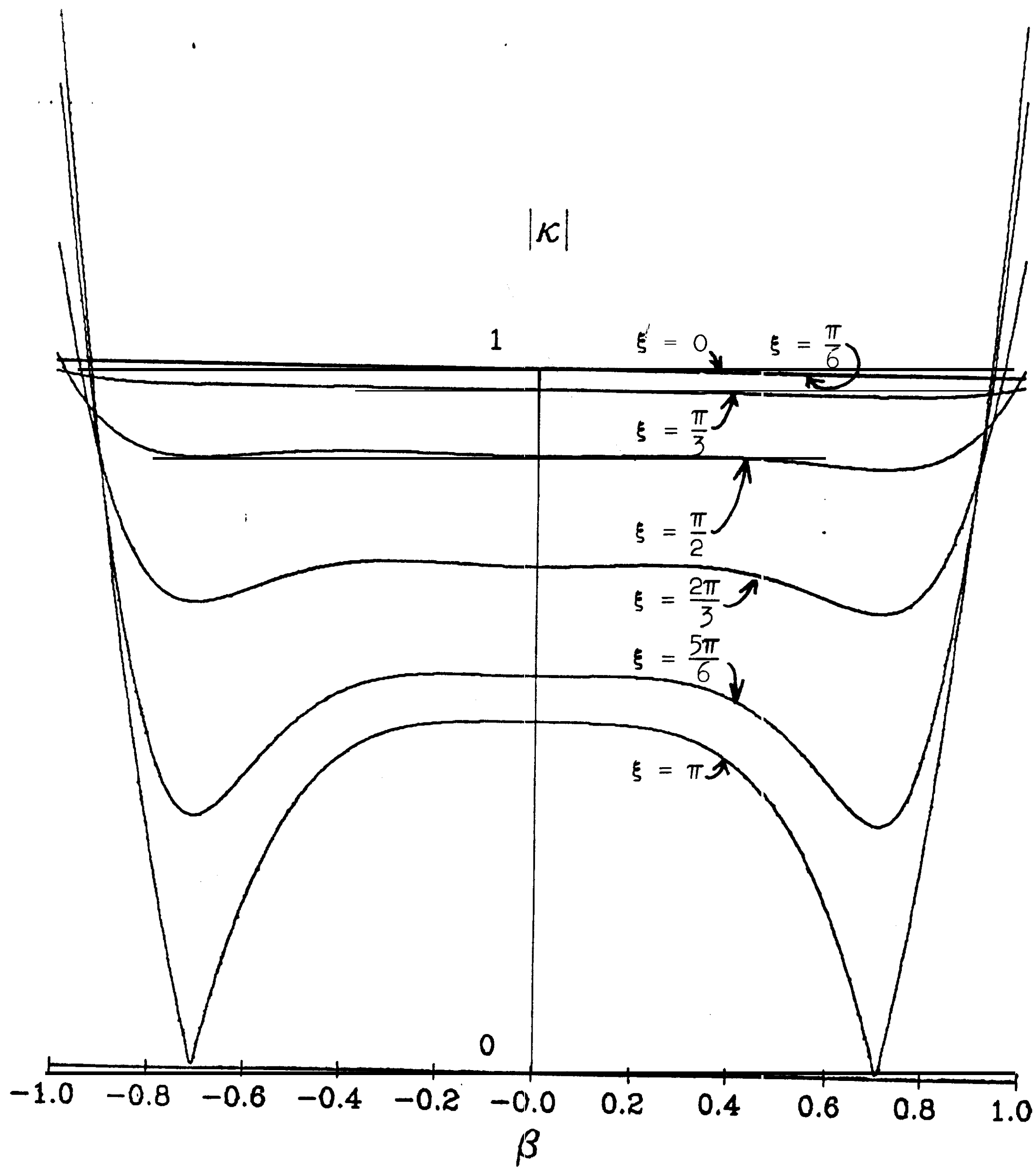

Fig. 2b Modified Lax-Wendroff. $\epsilon = 1.0$

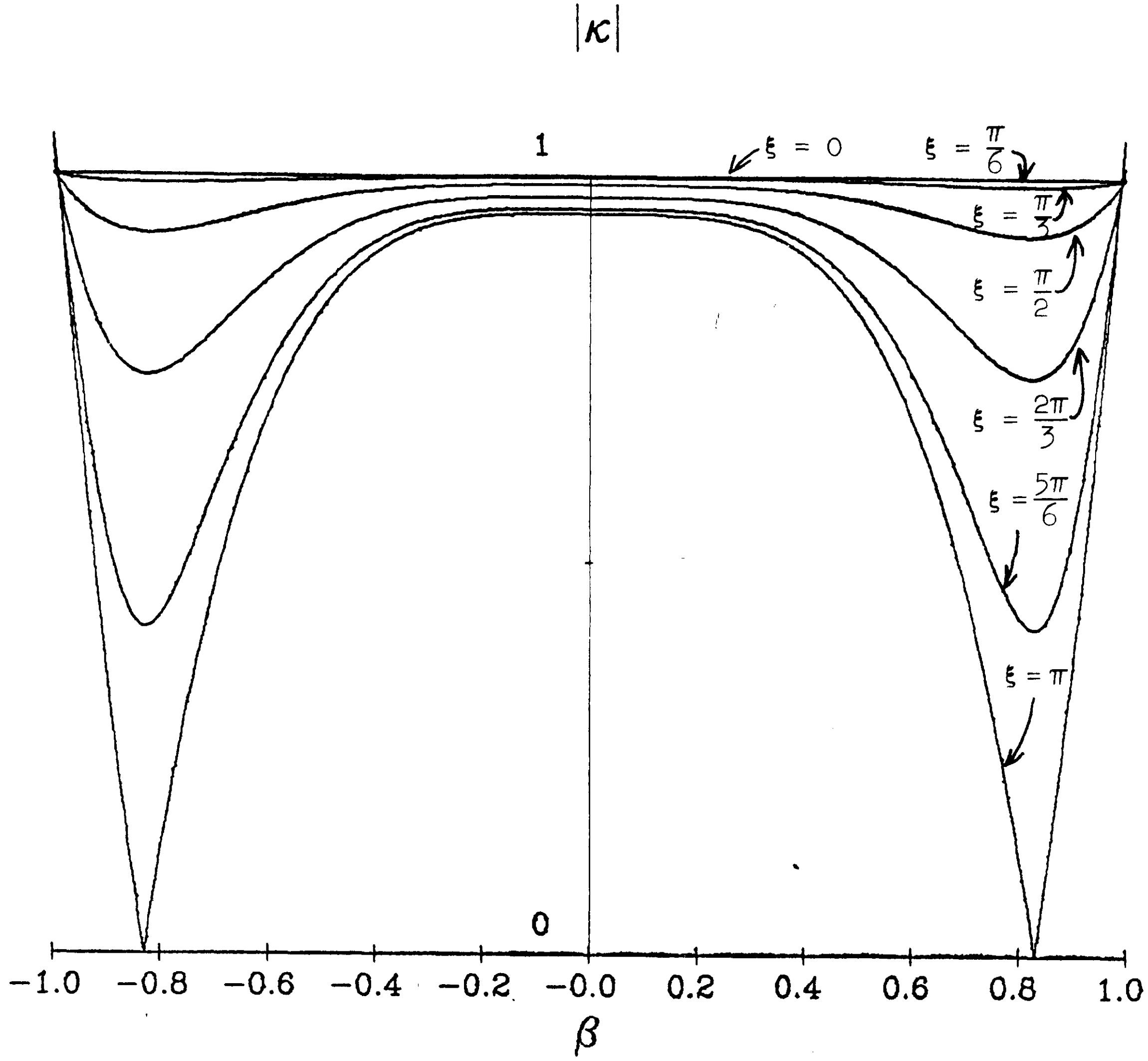

Fig. 2c Modified Lax-Wendroff 0.1

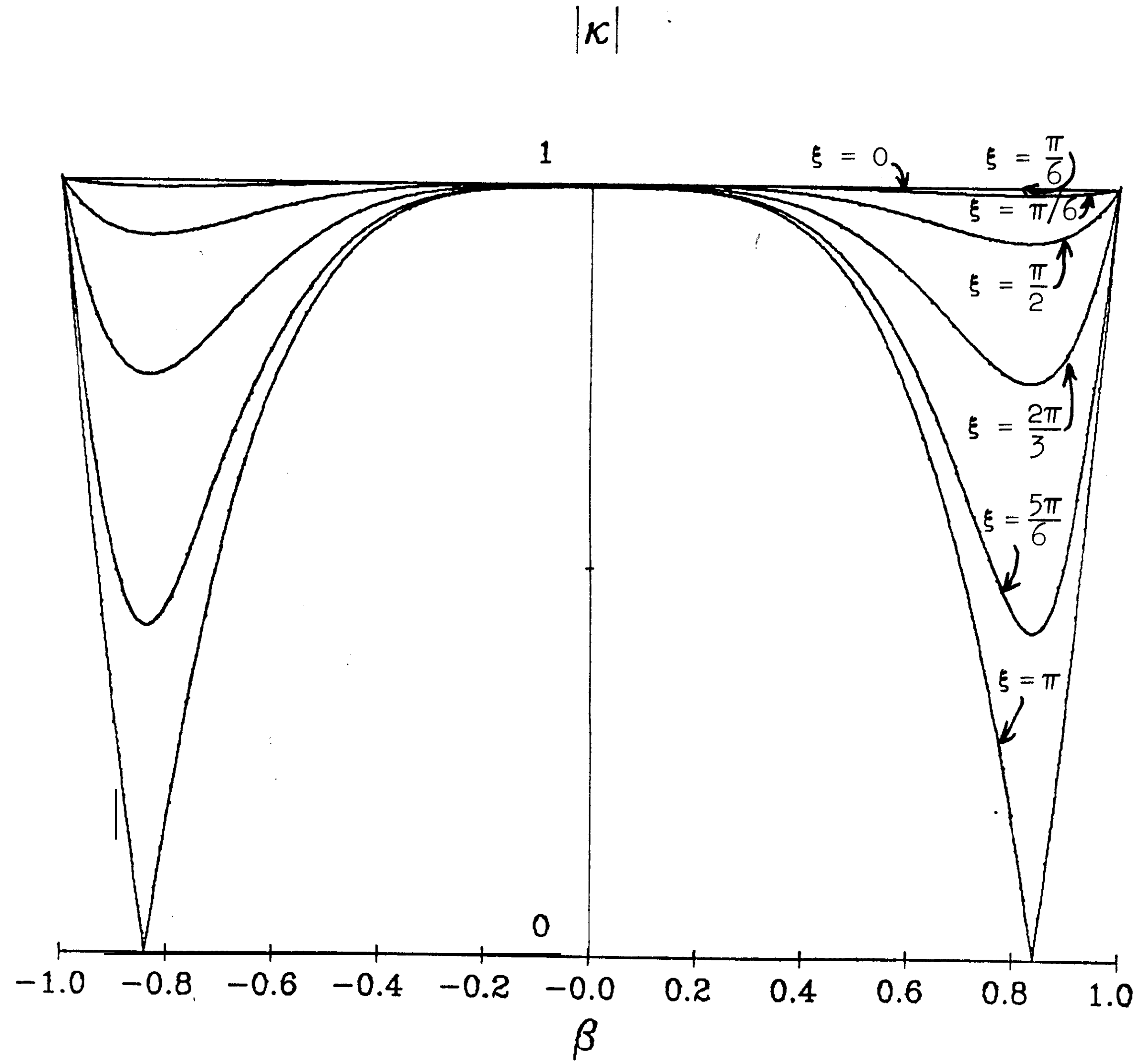

Fig. 2d Modified Lax-Wendroff. ϵ ∘ ∘1

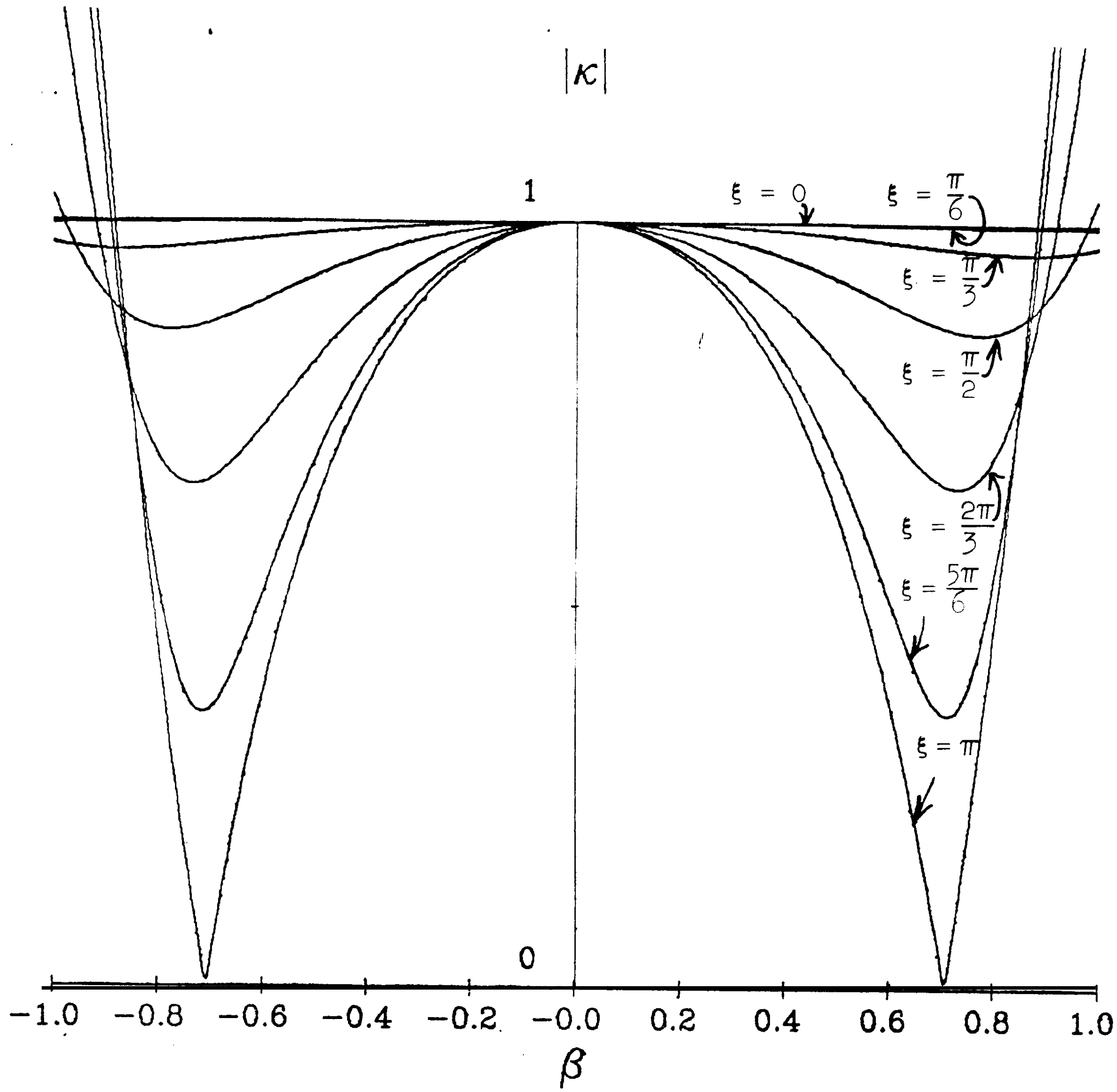

Fig. 3a Modified Lax-Wendroff with quadratic decay. $\delta = 2°$

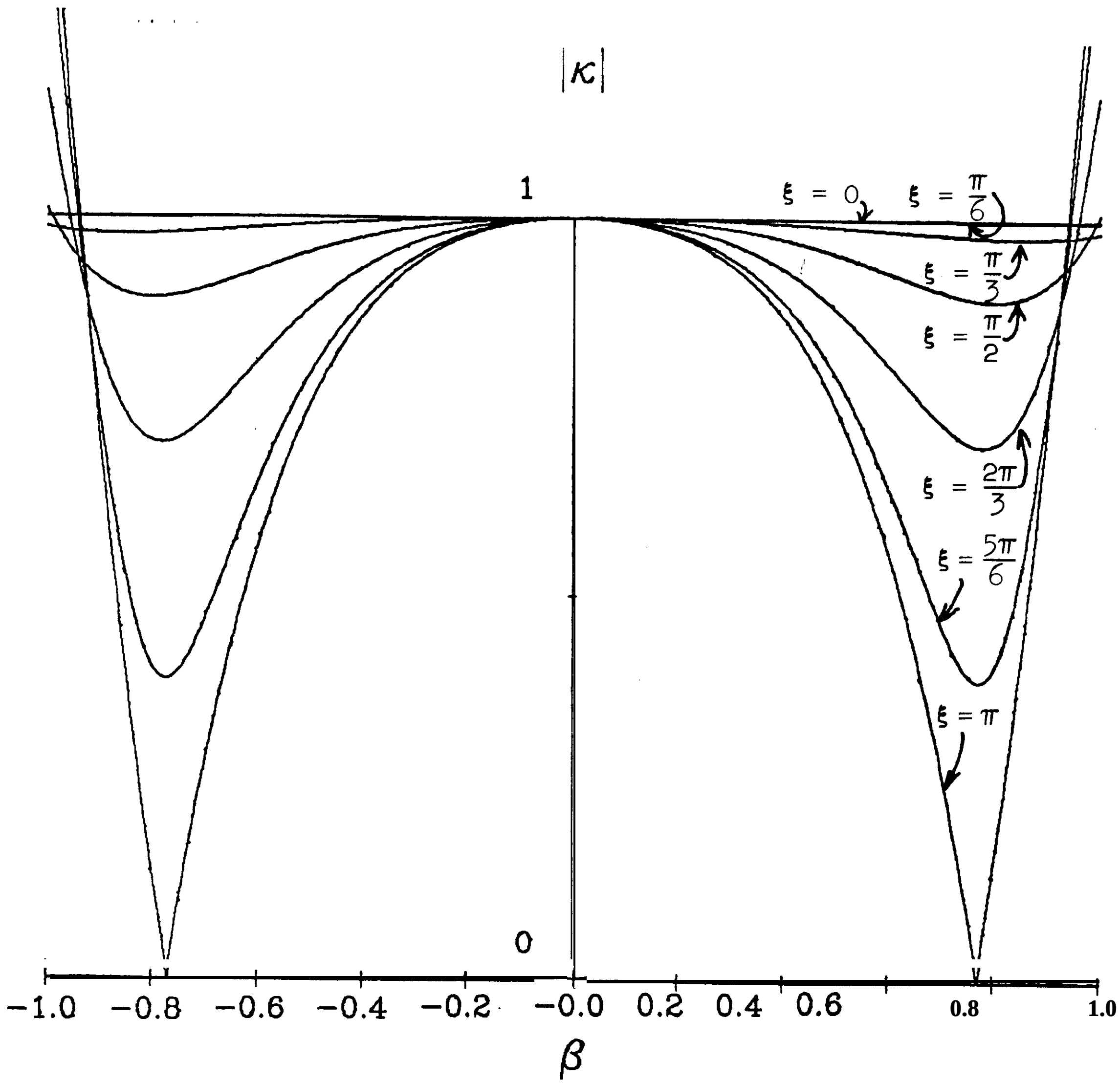

Fig. 3b Modified Lax-Wendroff with quadratic decay. $\delta = 1.0$

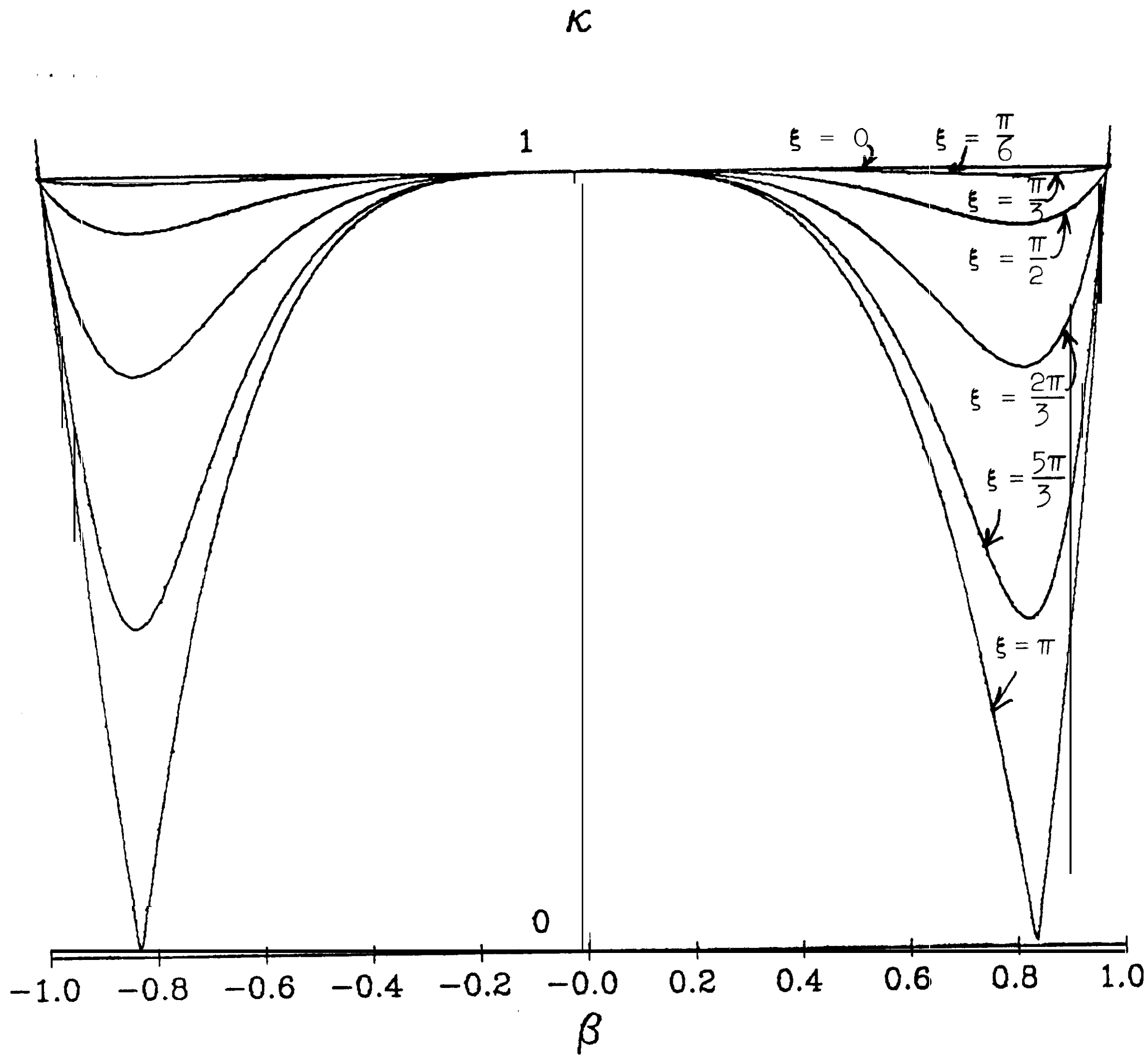

Fig. 3c Modified Lax-Wendroff with quadratic decay. $\delta = 0.1$

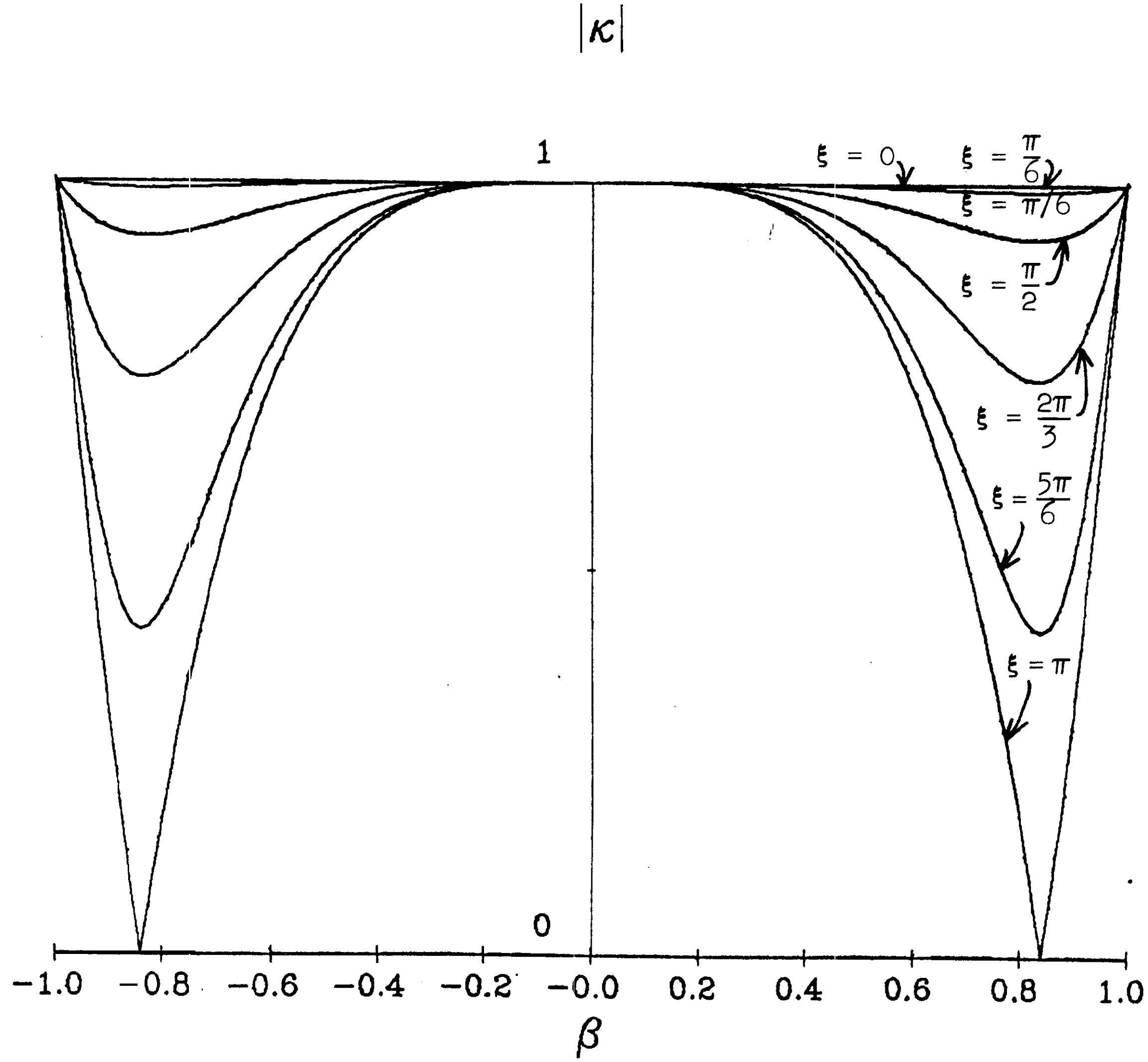

Fig. 3d Modified Lax-Wendroff with quadratic decay $\delta = 0{\cdot}1$

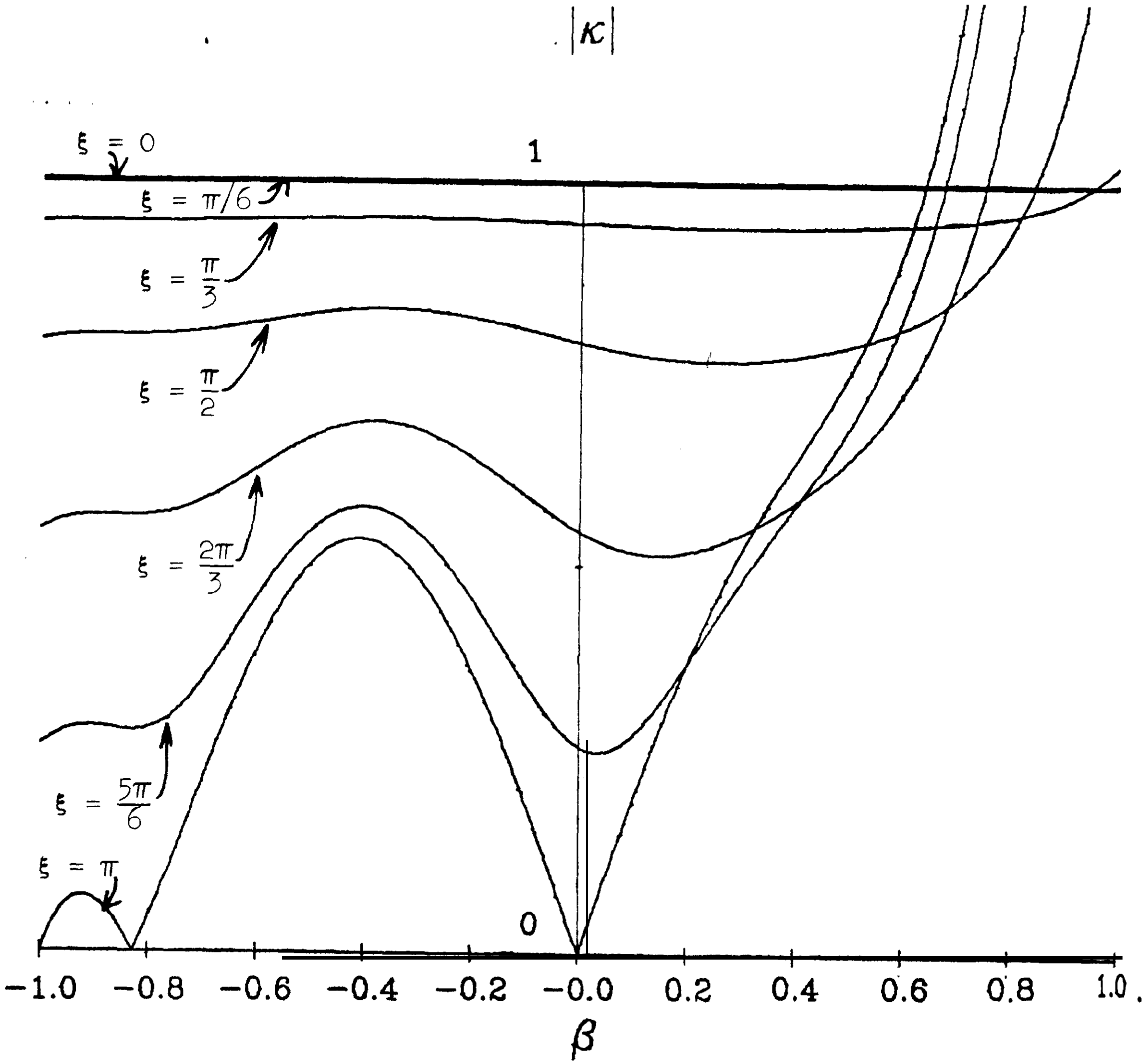

Fig. 4a MacCormack. $\epsilon = 2.0$

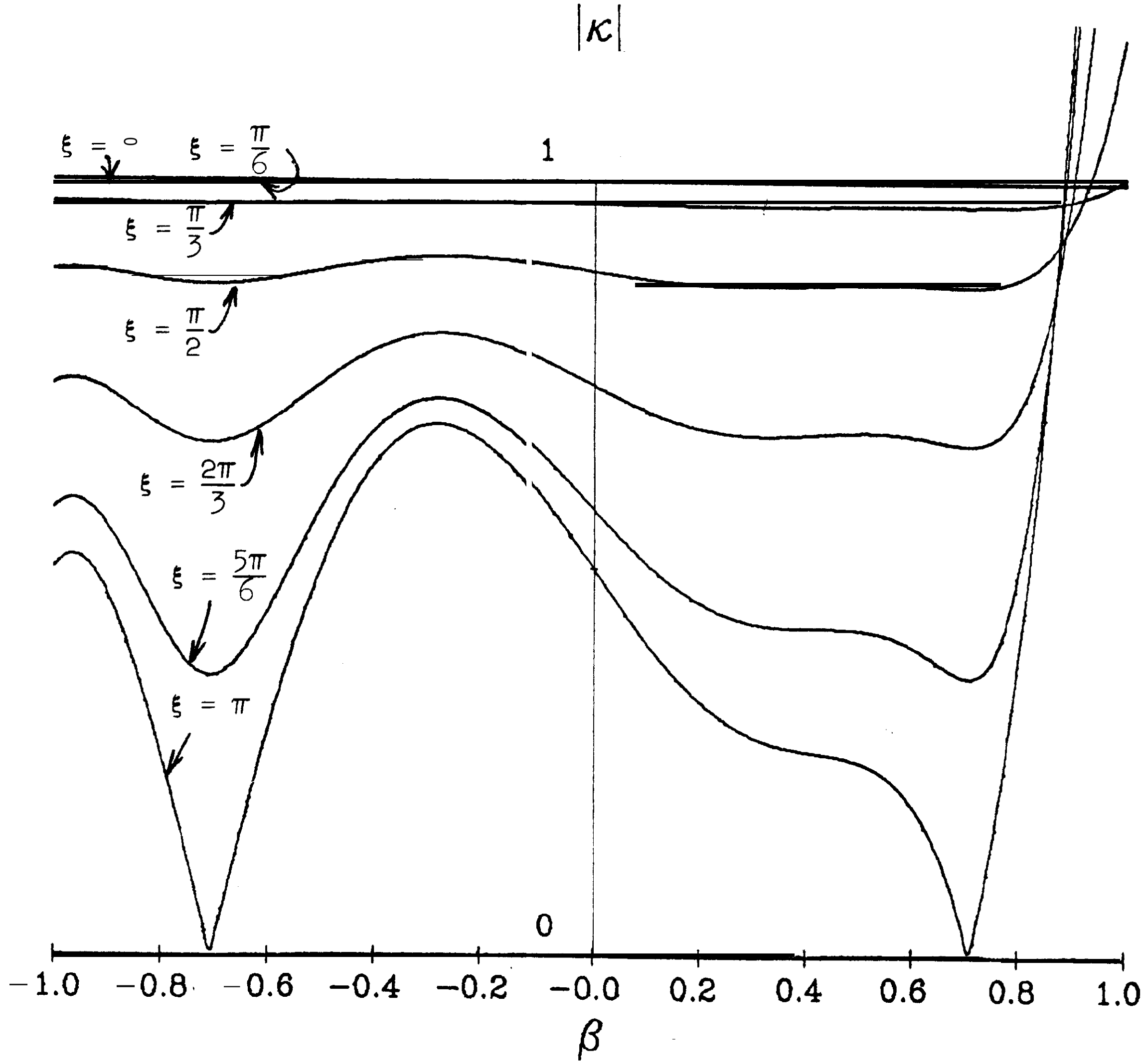

Fig. 4b MacCormack. $\epsilon = 1.0$

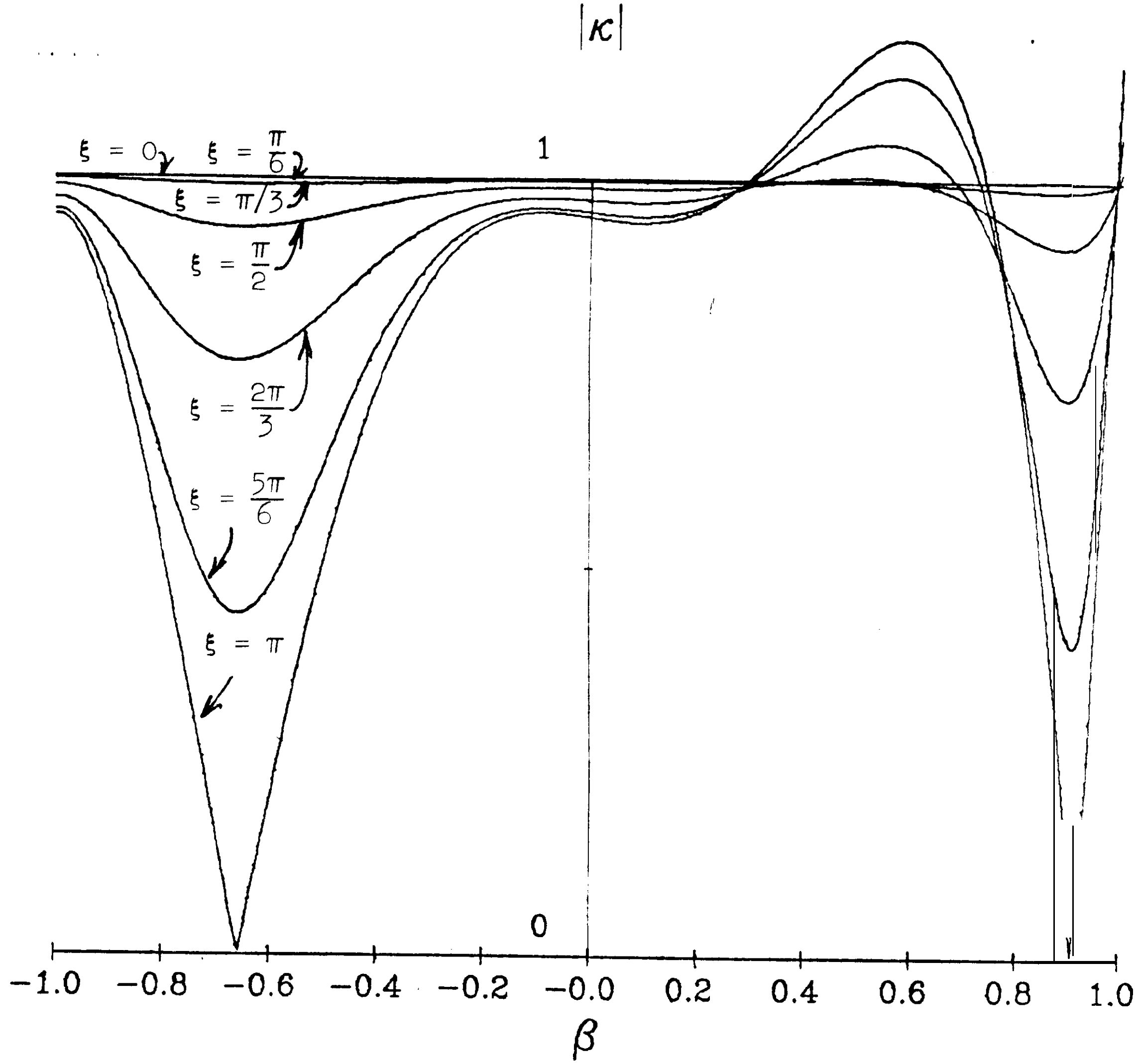

Fig. 4c MacCormack ϵ 0.1

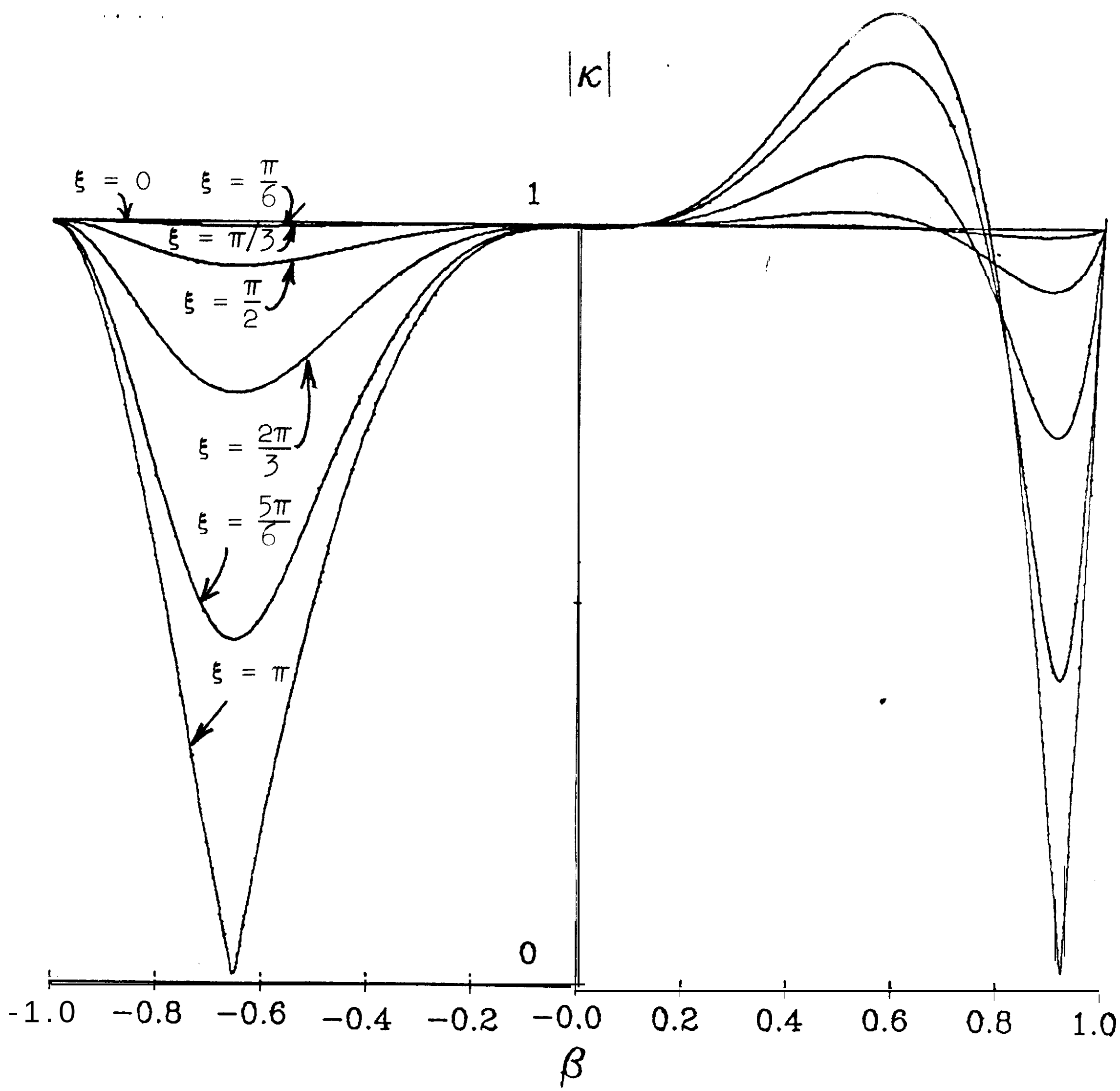

Fig. 4d MacCormack. ⊢ = 0.01

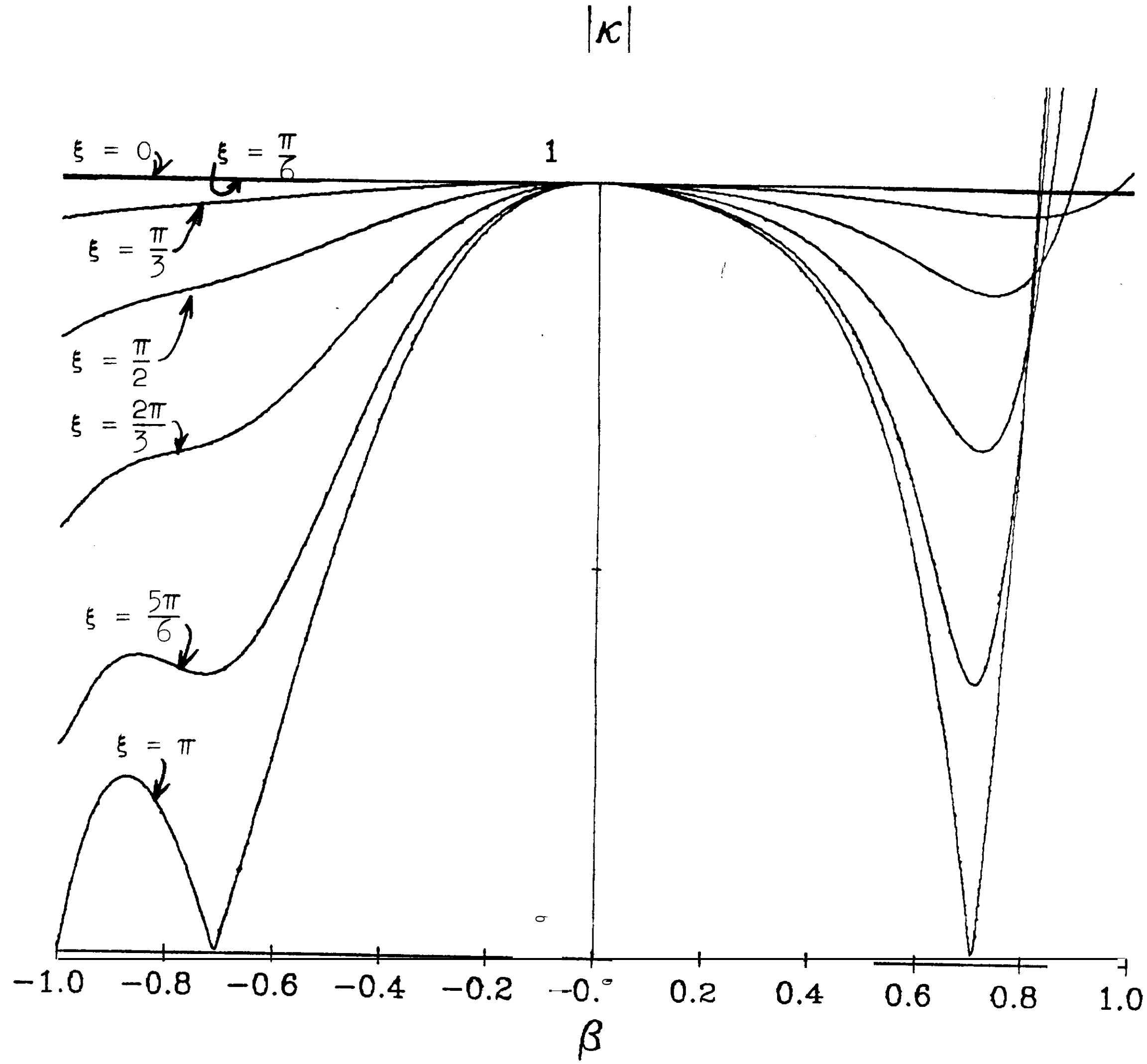

Fig. 5a MacCormack with quadratic decay δ 2.0

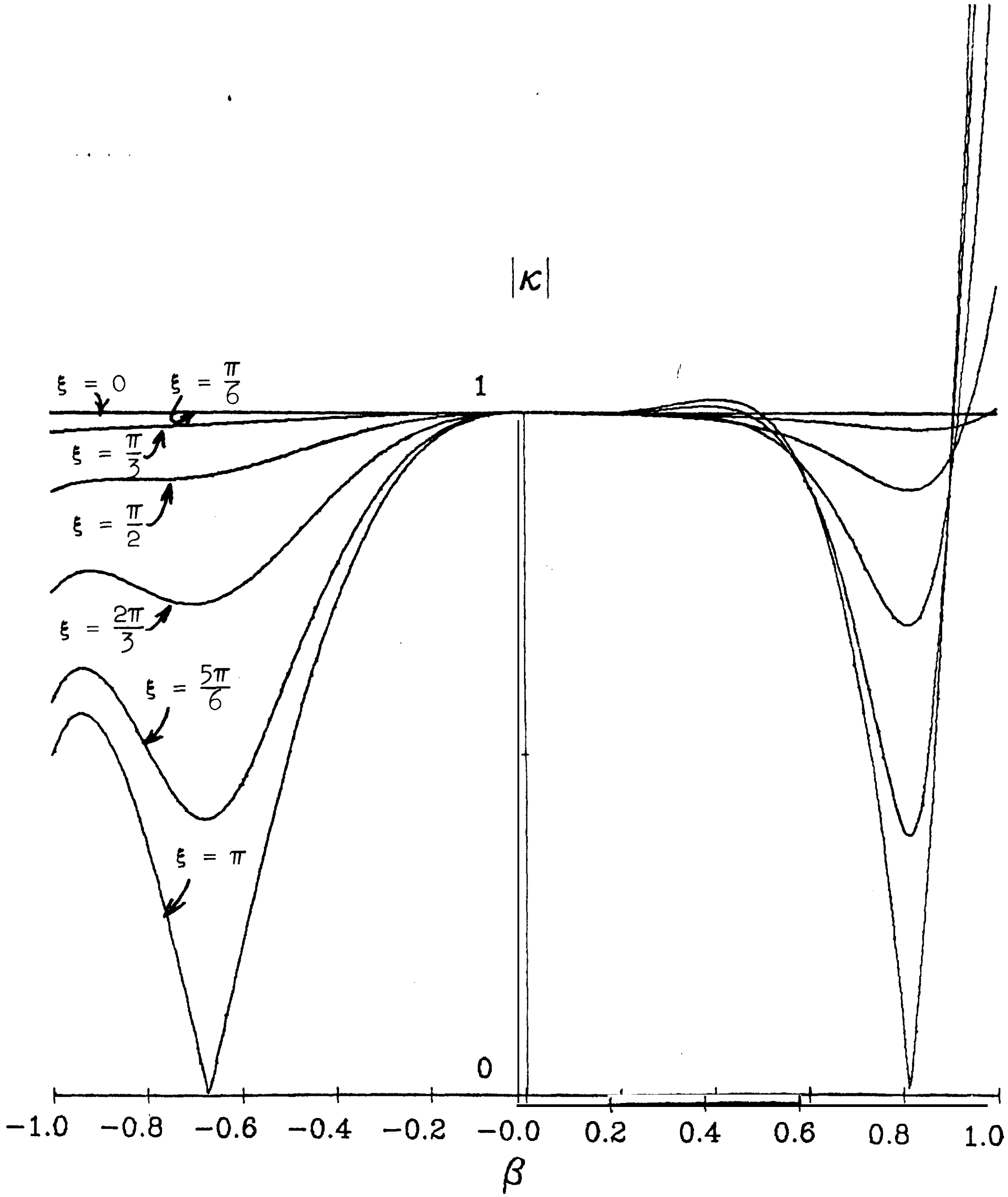

Fig. 5b MacCormack with quadratic decay. $\delta = 1.0$

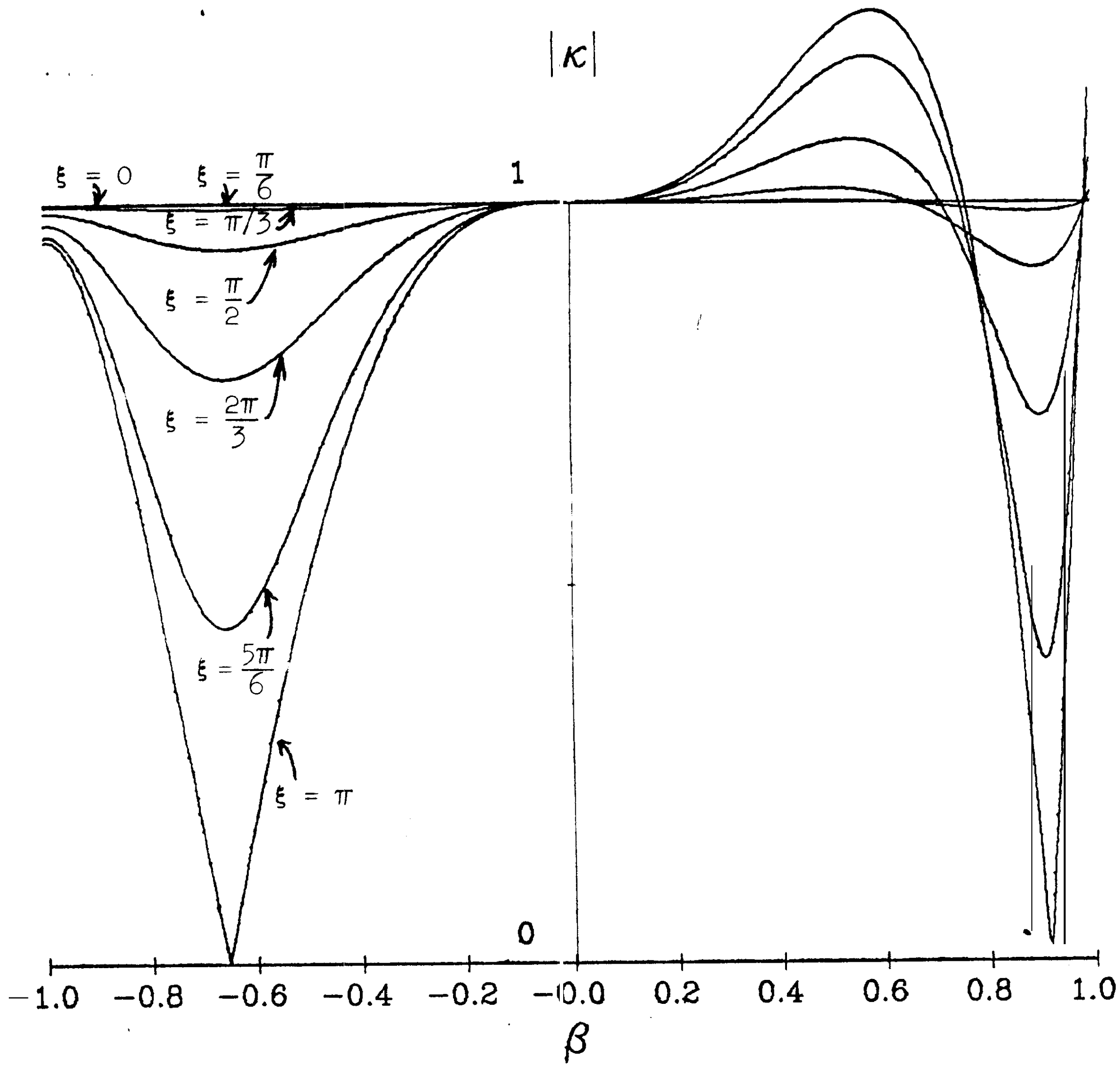

Fig 5c MacCormack with quadratic decay. δ ° 1

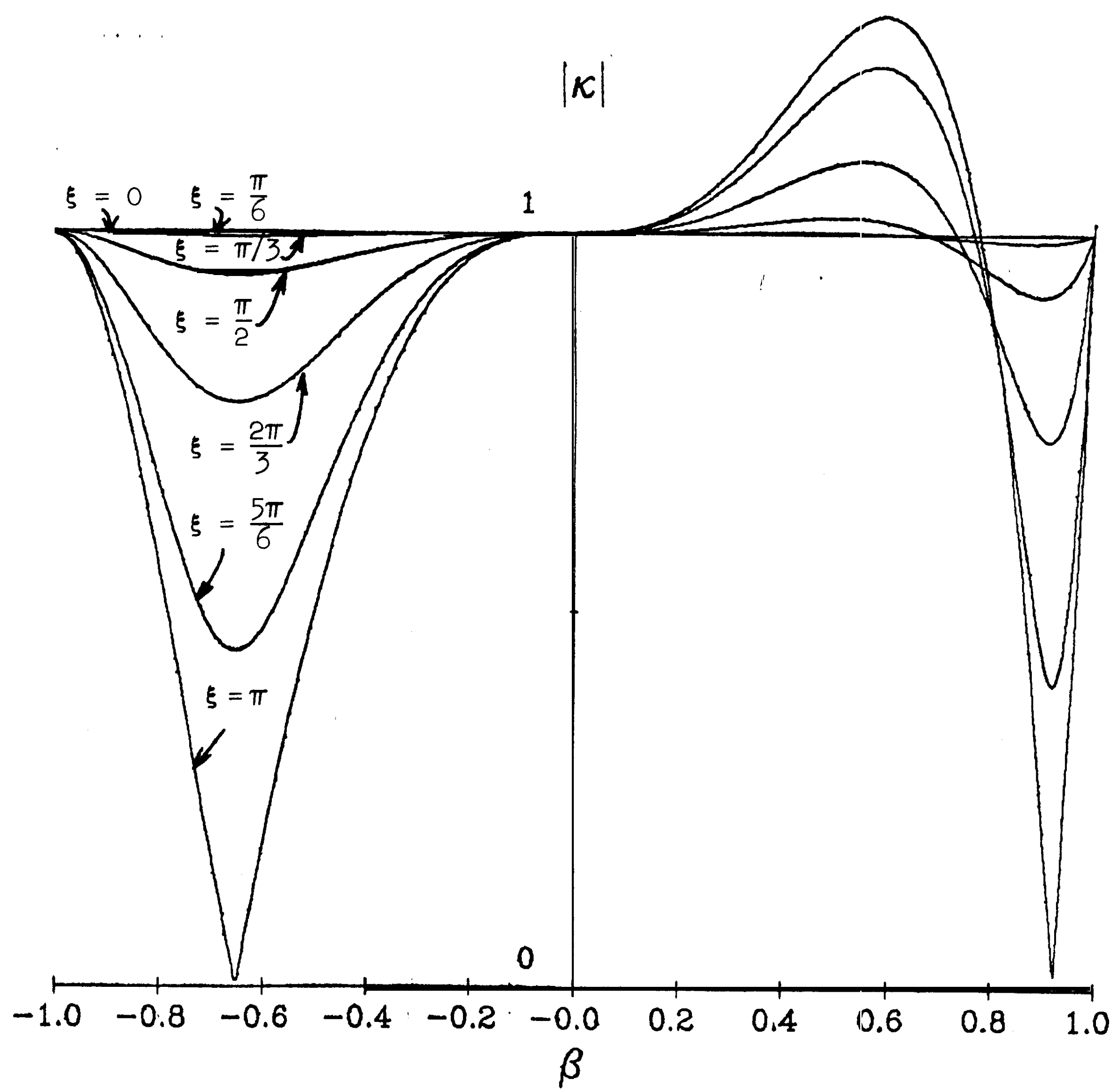

Fig. 5d MacCormack with quadratic decay. = 0.°1