NASA TM-86775

NASA Technical Memorandum 86775

Generalized Formulation of a Class of Explicit and Implicit TVD Schemes

H.C. Yee

July 1985

National Aeronautics and
Space Administration

NASA Technical Memorandum 86775

Generalized Formulation of a Class of Explicit and Implicit TVD Schemes

H. C. Yee, Ames Research Center, Moffett Field, California

July 1985

NASA
National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

N85-32830#

GENERALIZED FORMULATION OF A CLASS OF EXPLICIT AND IMPLICIT TVD SCHEMES

H.C. YEE[1]

NASA Ames Research Center
Moffett Field, CA., 94035, USA

Abstract

A one-parameter family of second-order explicit and implicit total variation diminishing (TVD) schemes is reformulated so that a simplier and wider group of limiters is included. The resulting scheme can be viewed as a symmetrical algorithm with a variety of numerical dissipation terms that are designed for weak solutions of hyperbolic problems. This is a generalization of Roe and Davis's recent works to a wider class of symmetric schemes other than Lax-Wendroff. The main properties of the present class of schemes are that they can be implicit, and when steady-state calculations are sought, the numerical solution is independent of the time step.

I. Introduction

The notion of total variation diminishing (TVD) schemes was introduced by Harten [1,2]. He derived a set of sufficient conditions which are very useful in checking or constructing second-order TVD schemes. The main mechanism that is currently in use for satisfying TVD sufficient conditions involves some kind of limiting procedure. There are generally two types of limiters: namely slope limiters [3] and flux limiters [4-6]. For a slope limiter one imposes constraints on the gradients of the dependent variables. In contrast, for a flux limiter one imposes constraints on the gradients of the flux functions. For constant coefficients, the two type of limiters are equivalent. The main property of a TVD scheme is that, unlike monotone schemes, it can be second-order accurate and is oscillation-free across discontinuities (when applied to nonlinear scalar hyperbolic conservation laws and constant coefficient hyperbolic systems). Sweby [5] and Roe [6] constructed a class of limiters as a function of the gradient ratio. Most of the current limiters in used are equivalent to members of this class.

Although TVD schemes are designed for transient applications, they also have been applied to steady-state problems [6-10]. It is well known that explicit methods are usually easier to program and often require less storage than implicit methods, but can suffer a loss of efficiency when the time step is restricted by stability rather than accuracy. It is also commonly known that it is not very useful to extend Lax-Wendroff-type schemes to implicit methods, since the resulting schemes are not suitable for steady-state calculations.

[1]Research Scientist, Computational Fluid Dynamics Branch.

This is due to the fact that the steady-state solution will depend on the time step. Roe has recently proposed a very enlightening generalized formulation of TVD Lax-Wendroff schemes [11]. Roe's results, in turn, is a generalization of Davis's work [12]. It was the investigation of these schemes which prompted the work of this paper. Their formulation has great potential for transient applications, but as it stands is not suitable for an extension to implicit methods.

The aim of this paper is to incorporate the results of Roe [11], with minor modification, to a one-parameter family of explicit and implicit TVD schemes [2,8-9] so that a wider group of limiters can be represented in a general but rather simple form which is at the same time suitable for steady-state applications. The final scheme can be interpreted as a three-point, spatially central difference explicit or implicit scheme which has a whole variety of more rational numerical dissipation terms than the classical way of handling shock-capturing algorithms. In other words, it is a non-upwind TVD scheme or a symmetric TVD scheme. The proposed scheme can be used for time-accurate or steady-state calculations. Moreover, Roe's formulation can be considered as a member of the explicit scheme for time-accurate calculations. By writing the scheme in terms of numerical fluxes with two input parameters (one for the choice of the time-differencing method and one for the option of choosing the Lax-Wendroff flux), a single computer program can be easily coded to include all of the schemes under discussion. Various limiters can be considered as external functions inside the computer program. Extension of the schemes to nonlinear scalar and system of hyperbolic conservation laws is discussed in detail. An equivalent representation of the proposed numerical dissipation terms which avoids extra logic in the computer implementation is also included in the appendix.

The effort involved in modifying some existing central difference computer codes for systems of hyperbolic conservation laws is fairly simple and straightforward. The proposed algorithms should have the potential of improving the robustness and accuracy of many practical physical and engineering applications. Only analytical formulations are presented here. Numerical testing and practical applications will be the subject of a forthcoming paper.

II. Preliminaries

In this section, a class of explicit and implicit TVD schemes [2] is reviewed. Harten's sufficient conditions for this class of schemes are also stated. This set of conditions is then ultilized in the subsequent sections to construct and reformulate the second-order explicit and implicit TVD schemes of Harten [1,2].

Consider the scalar hyperbolic conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \tag{2.1}$$

where f is the flux and $a(u) = \partial f / \partial u$ is the characteristic speed. Let u_j^n be the numerical solution of (2.1) at $x = j\Delta x$ and $t = n\Delta t$, with Δx the spatial mesh size and Δt the time step. Consider a one-parameter family of five-point difference schemes in conservation form

$$u_j^{n+1} + \lambda\theta(h_{j+\frac{1}{2}}^{n+1} - h_{j-\frac{1}{2}}^{n+1}) = u_j^n - \lambda(1-\theta)(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n), \tag{2.2}$$

where θ is a nonnegative parameter, $\lambda = \Delta t/\Delta x$, $h_{j\pm\frac{1}{2}}^n = h(u_{j\mp1}^n, u_j^n, u_{j\pm1}^n, u_{j\pm2}^n)$, and $h_{j\pm\frac{1}{2}}^{n+1} = h(u_{j\mp1}^{n+1}, u_j^{n+1}, u_{j\pm1}^{n+1}, u_{j\pm2}^{n+1})$. The function $h_{j+\frac{1}{2}}$ is commonly called a numerical flux function. Let

$$\bar{h}_{j+\frac{1}{2}} = (1-\theta)h_{j+\frac{1}{2}}^n + \theta h_{j+\frac{1}{2}}^{n+1} \tag{2.3}$$

be another numerical flux function. Then (2.2) can be rewritten as

$$u_j^{n+1} = u_j^n - \lambda(\bar{h}_{j+\frac{1}{2}} - \bar{h}_{j-\frac{1}{2}}). \tag{2.4}$$

This numerical flux is a function of eight variables, $\bar{h}_{j+\frac{1}{2}} = \bar{h}(u_{j-1}^n, u_j^n, u_{j+1}^n, u_{j+2}^n, u_{j-1}^{n+1} u_j^{n+1}, u_{j+1}^{n+1}, u_{j+2}^{n+1})$, and is consistent with the conservation law(2.1) in the following sense

$$\bar{h}(u,u,u,u,u,u,u,u) = f(u). \tag{2.5}$$

This one-parameter family of schemes contains implicit as well as explicit schemes. When $\theta = 0$, (2.2) is an explicit method. When $\theta \neq 0$, (2.2) is an implicit scheme. For example, if $\theta = 1/2$, the time-differencing is the trapezoidal formula, and if $\theta = 1$, the time-differencing is the backward Euler method. To simplify the notation, rewrite equation (2.2) as

$$L \cdot u^{n+1} = R \cdot u^n, \tag{2.6}$$

where L and R are the following finite-difference operators:

$$(L \cdot u)_j = u_j + \lambda\theta(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) \tag{2.7a}$$

$$(R \cdot u)_j = u_j - \lambda(1-\theta)(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}). \tag{2.7b}$$

The total variation of a mesh function u^n is defined to be

$$TV(u^n) = \sum_{j=-\infty}^{\infty} |u_{j+1}^n - u_j^n| = \sum_{j=-\infty}^{\infty} |\Delta_{j+\frac{1}{2}} u^n|, \tag{2.8}$$

where $\Delta_{j+\frac{1}{2}} u^n = u_{j+1} - u_j$. Here, the general notation convention

$$\Delta_{j+\frac{1}{2}} z = z_{j+1} - z_j \tag{2.9}$$

for any mesh function z is used. The numerical scheme (2.2) for an initial-value problem of (2.1) is said to be TVD if

$$TV(u^{n+1}) \leq TV(u^n). \tag{2.10}$$

The following sufficient conditions for (2.2) to be a TVD scheme are due to Harten [2]:

$$TV(R \cdot u^n) \leq TV(u^n) \tag{2.11a}$$

and

$$TV(L \cdot u^{n+1}) \geq TV(u^{n+1}). \tag{2.11b}$$

Assume the numerical flux h in (2.2) is Lipschitz continuous and (2.2) can be written as

$$u_j^{n+1} - \lambda\theta\left(\widetilde{C}_{j+\frac{1}{2}}^{-}\Delta_{j+\frac{1}{2}}u - \widetilde{C}_{j-\frac{1}{2}}^{+}\Delta_{j-\frac{1}{2}}u\right)^{n+1} = u_j^n + \lambda(1-\theta)\left(\widetilde{C}_{j+\frac{1}{2}}^{-}\Delta_{j+\frac{1}{2}}u - \widetilde{C}_{j-\frac{1}{2}}^{+}\Delta_{j-\frac{1}{2}}u\right)^{n}, \tag{2.12}$$

where $\widetilde{C}_{j\pm\frac{1}{2}}^{\mp} = \widetilde{C}^{\mp}(u_j, u_{j\pm 1}, u_{j\pm 2})$ or possibily $\widetilde{C}_{j\pm\frac{1}{2}}^{\mp} = \widetilde{C}^{\mp}(u_{j\mp 1}, u_j, u_{j\pm 1}, u_{j\pm 2})$ are some bounded functions. Then Harten further showed that sufficient conditions for (2.11) are

(a) if for all j

$$C_{j+\frac{1}{2}}^{\pm} = \lambda(1-\theta)\widetilde{C}_{j+\frac{1}{2}}^{\pm} \geq 0 \tag{2.13a}$$

$$C_{j+\frac{1}{2}}^{+} + C_{j+\frac{1}{2}}^{-} = \lambda(1-\theta)(\widetilde{C}_{j+\frac{1}{2}}^{+} + \widetilde{C}_{j+\frac{1}{2}}^{-}) \leq 1, \tag{2.13b}$$

and

(b) if for all j

$$-\infty < C \leq -\lambda\theta\widetilde{C}_{j+\frac{1}{2}}^{\pm} \leq 0 \tag{2.14}$$

for some finite C. Conditions (2.13) and (2.14) are very useful in guiding the construction of second-order accurate TVD schemes which do not exhibit the spurious oscillation associated with the more classical second-order schemes.

Harten [1-2], Yee et al. and Yee [7-10] investigated a particular form of $C^{\pm}$. They have shown in a variety of numerical tests that the scheme is quite useful for gas-dynamic calculations. The recent work of Roe [11] suggests a wider class of flux limiters for the Lax-Wendroff-type of TVD schemes which with a minor modification are found to have an immediate application to scheme (2.2). The details will be discussed in the next two sections.

III. A Generalized Formulation of a Class of Symmetric Schemes

In this section, Roe's formulation is reviewed. Then, with a minor modification, his numerical flux is shown to be applicable to a larger class of symmetric schemes. Sufficient conditions for this new class of schemes to be TVD are derived for both the constant coefficient and nonlinear scalar hyperbolic equations.

3.1 Roe's TVD Lax-Wendroff Schemes

Roe [11] has recently developed a generalized formulation of TVD Lax-Wendroff schemes. The form of the schemes is the usual Lax-Wendroff plus a general conservative dissipation term designed in such a way that the final scheme is TVD. For $\partial f/\partial u = a =$ constant, his scheme is written as

$$\begin{aligned} u_j^{n+1} = u_j^n &- \frac{1}{2}\nu(1+\nu)\Delta_{j-\frac{1}{2}}u - \frac{1}{2}\nu(1-\nu)\Delta_{j+\frac{1}{2}}u \\ &- \frac{1}{2}|\nu|(1-|\nu|)(1-Q_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u \\ &+ \frac{1}{2}|\nu|(1-|\nu|)(1-Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u. \end{aligned} \tag{3.1}$$

Here $\nu = a\lambda = a\Delta t/\Delta x$, the first two terms represent the usual Lax-Wendroff scheme, and the other two terms represent an additional conservative dissipation. The function $Q_{j+\frac{1}{2}}$ depends on three consecutive gradients $\Delta_{j-\frac{1}{2}}u$, $\Delta_{j+\frac{1}{2}}u$, $\Delta_{j+\frac{3}{2}}u$ and is of the form

$$Q_{j+\frac{1}{2}} = Q(r^-_{j+\frac{1}{2}}, r^+_{j+\frac{1}{2}}), \tag{3.2a}$$

where

$$r^-_{j+\frac{1}{2}} = \frac{\Delta_{j-\frac{1}{2}}u}{\Delta_{j+\frac{1}{2}}u}, \qquad r^+_{j+\frac{1}{2}} = \frac{\Delta_{j+\frac{3}{2}}u}{\Delta_{j+\frac{1}{2}}u}. \tag{3.2b}$$

Here $r^{\pm}$ are not defined if $\Delta_{j+\frac{1}{2}}u = 0$. To avoid this, an equivalent representation will be discussed in the appendix. If one assumes both Q and Q/r are always positive, then a set of sufficient conditions for (3.1) to be TVD is

$$Q_{j+\frac{1}{2}} < \frac{2}{1-|\nu|} \tag{3.3a}$$

$$\left(Q_{j+\frac{1}{2}} \Big/ r^-_{j+\frac{1}{2}}\right) < \frac{2}{|\nu|} \tag{3.3b}$$

$$\left(Q_{j+\frac{1}{2}} \Big/ r^+_{j+\frac{1}{2}}\right) < \frac{2}{|\nu|}. \tag{3.3c}$$

Two examples for the function Q are

$$Q(r^-, r^+) = \text{minmod}(1, r^-) + \text{minmod}(1, r^+) - 1, \tag{3.4}$$

and

$$Q(r^-, r^+) = \text{minmod}(1, r^-, r^+). \tag{3.5}$$

Normally the "minmod" function of two arguments is defined as

$$\text{minmod}(x,y) = \text{sgn}(x)\cdot\max\{0, \min[|x|, y\cdot\text{sgn}(x)]\}.$$

but within this context

$$\text{minmod}(1, r^{\pm}) = \begin{cases} \min(1, r^{\pm}) & r^{\pm} > 0 \\ 0 & r^{\pm} \le 0. \end{cases} \tag{3.6}$$

Other forms of $Q(r^-, r^+)$ are discussed in Sweby [5] and Roe [6].

Scheme (3.1) is a reformulation of Davis's work [12] in a way which is easier to analyze and includes a class of TVD schemes not observed by Davis. The numerical flux denoted by $h^{LW}_{j+\frac{1}{2}}$ for (3.1) is

$$h^{LW}_{j+\frac{1}{2}} = \frac{1}{2}\left\{a(u_{j+1}+u_j) - \left[\lambda a^2 Q_{j+\frac{1}{2}} + |a|(1-Q_{j+\frac{1}{2}})\right]\Delta_{j+\frac{1}{2}}u\right\}. \tag{3.7}$$

Scheme (3.1) is second-order accurate in time and space. Observe that by setting $\theta = 0$ in (2.2) and by using (3.7) as the numerical flux, the resulting scheme is (3.1).

3.2 Schemes for Linear Scalar Hyperbolic Equations

If one is to use (3.7) as the numerical flux for (2.2) with $\theta \neq 0$, then the resulting scheme is only useful for transient calculations. For steady-state applications, either one has to restrict the time step similar to the explicit method or the steady-state solution will depend on the time step. It is emphasized here that the dependence of the time step in steady-state solutions occurs even though the value of Δt is similar to an explicit method. In this case Δt is most often in the same order as Δx; thus the dependence in Δt is less severe. The term that causes this undesirable property is the one with coefficient a^2 in equation (3.7), since the λ appears in this term. Therefore, besides considering the use of (3.7) as the numerical flux for (2.2) when $\theta = 0$, the numerical flux (3.7) with $a^2 = 0$ is also considered; i.e., the numerical flux is of the form

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[a(u_{j+1}+u_j) - |a|(1-Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]. \tag{3.8}$$

Now the question is, will the new numerical flux (3.8) satisfy the sufficient conditions (2.11)? The answer is yes. It turns out that some of the Q functions that are suitable for the generalized TVD Lax-Wendroff scheme are also suitable for (3.8). The implication is that if one chose the proper Q function, the resulting scheme (2.2) together with (3.8) can be viewed as a symmetrical algorithm with a wide variety of numerical dissipation terms that satisfy the TVD property.

Now with the choice of (3.8), the corresponding $\widetilde{C}^{\pm}$ of equation (2.12) are

$$\widetilde{C}^{+}_{j-\frac{1}{2}} = a\left[1 - \frac{1}{2}Q_{j-\frac{1}{2}} + \frac{1}{2}\left(Q_{j+\frac{1}{2}} \Big/ r^{-}_{j+\frac{1}{2}}\right)\right], \qquad a > 0 \tag{3.9a}$$

$$\widetilde{C}^{-}_{j+\frac{1}{2}} = |a|\left[1 - \frac{1}{2}Q_{j+\frac{1}{2}} + \frac{1}{2}\left(Q_{j-\frac{1}{2}}\Big/ r^{+}_{j-\frac{1}{2}}\right)\right], \qquad a < 0. \tag{3.9b}$$

Therefore, sufficient conditions for this specific numerical flux function (3.8) to be TVD are

$$0 < \lambda(1-\theta)a\left[1 - \frac{1}{2}Q_{j-\frac{1}{2}} + \frac{1}{2}\left(Q_{j+\frac{1}{2}}\Big/ r^{-}_{j+\frac{1}{2}}\right)\right] < 1 \qquad a > 0 \tag{3.10a}$$

$$0 < \lambda(1-\theta)|a|\left[1 - \frac{1}{2}Q_{j+\frac{1}{2}} + \frac{1}{2}\left(Q_{j-\frac{1}{2}}\Big/ r^{+}_{j-\frac{1}{2}}\right)\right] < 1 \qquad a < 0 \tag{3.10b}$$

and

$$-\infty < -\lambda\theta\widetilde{C}^{\pm}_{j+\frac{1}{2}} \le 0. \tag{3.11}$$

For $0 \le \theta \le 1$ and $\nu \ne 0$, condition (3.10a) is satisfied if

$$Q_{j-\frac{1}{2}} - \left(Q_{j+\frac{1}{2}}\Big/ r^{-}_{j+\frac{1}{2}}\right) < 2 \tag{3.12a}$$

$$\left(Q_{j+\frac{1}{2}}\Big/ r^{-}_{j+\frac{1}{2}}\right) - Q_{j-\frac{1}{2}} < \frac{2}{\lambda(1-\theta)a} - 2 \tag{3.12b}$$

$$\lambda a < \frac{1}{1-\theta}, \tag{3.12c}$$

and condition (3.10b) is satisfied if

$$Q_{j+\frac{1}{2}} - \left(Q_{j-\frac{1}{2}}\Big/ r^{+}_{j-\frac{1}{2}}\right) < 2, \tag{3.12d}$$

$$\left(Q_{j-\frac{1}{2}}\Big/ r^{+}_{j-\frac{1}{2}}\right) - Q_{j+\frac{1}{2}} < \frac{2}{\lambda(1-\theta)|a|} - 2 \tag{3.12e}$$

$$\lambda|a| < \frac{1}{1-\theta}. \tag{3.12f}$$

Since $\lambda|a| \le \frac{1}{1-\theta}$, the term $\frac{2}{\lambda(1-\theta)|a|} - 2$ is always positive. Therefore the same assumption as Roe can be made; i.e., assume both Q and Q/r are always positive. Then all one has to do is devise a function Q such that

$$Q_{j+\frac{1}{2}} < 2 \tag{3.13a}$$

$$\left(Q_{j+\frac{1}{2}}\Big/ r^{-}_{j+\frac{1}{2}}\right) < \frac{2}{\lambda(1-\theta)|a|} - 2 \tag{3.13b}$$

$$\left(Q_{j+\frac{1}{2}}\Big/ r^{+}_{j+\frac{1}{2}}\right) < \frac{2}{\lambda(1-\theta)|a|} - 2 \tag{3.13c}$$

$$\lambda|a| < \frac{1}{1-\theta}. \tag{3.13d}$$

With the above choice of Q, the last sufficient condition (3.11) is immediately satisfied. Some readily available $Q_{j+\frac{1}{2}}$ functions can be found in references [5,11]. For instance, the two examples given in equations (3.4) and (3.5) satisfy condition (3.13).

For $\theta = 1/2$, scheme (2.2) together with (3.8) and (3.13) is second-order accurate in both space and time. The CFL-like restriction for (2.2) to be TVD in this case is 2. When $\theta = 1$, scheme (2.2) together with (3.8) and (3.13) is unconditionally TVD, but the resulting scheme is first-order in time and second-order in space. When $\theta = 0$, the scheme is explicit, and unlike Roe's schemes, is only first-order in time but second-order in space.

As noted before, the value $r^-_{j+\frac{1}{2}}$ (or $r^+_{j+\frac{1}{2}}$) is not defined if $\Delta_{j-\frac{1}{2}}u$ (or $\Delta_{j+\frac{3}{2}}u$) is finite but $\Delta_{j+\frac{1}{2}}u = 0$. For computer implementation purposes, it might be more convenient to define $Q_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}u = \widehat{Q}_{j+\frac{1}{2}}$, where $\widehat{Q}_{j+\frac{1}{2}}$ is a function of $\Delta_{j-\frac{1}{2}}u$, $\Delta_{j+\frac{1}{2}}u$, and $\Delta_{j+\frac{3}{2}}u$, but not a ratio of those gradients. For this formulation and its extension to the nonlinear case, see the appendix.

3.3 Linearized Version of the Proposed Scheme for Constant Coefficient Equations

For $\theta \neq 0$, scheme (2.2) is implicit. Moreover, this is a genuinely nonlinear scheme in the sense that the final algorithm is nonlinear even for the constant coefficient case. The value of u^{n+1} is obtained as the solution of a system of nonlinear algebraic equations. To solve this set of nonlinear equations noniteratively, a linearized version of (2.2) together with (3.8) is considered. Substituting (3.8) in (2.2), one obtains

$$\begin{aligned} u_j^{n+1} &+ \frac{\lambda\theta}{2}\left[au_{j+1} - |a|(1 - Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]^{n+1} \\ &- \frac{\lambda\theta}{2}\left[au_{j-1} - |a|(1 - Q_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u\right]^{n+1} = \text{RHS of (2.2)}. \end{aligned} \tag{3.14}$$

Here "RHS of (2.2)" means the right hand side of equation (2.2) with $h_{j+\frac{1}{2}}$ defined in (3.8). Locally linearizing the coefficients of $(\Delta_{j\pm\frac{1}{2}}u)^{n+1}$ in (3.14) by dropping the time index from $(n+1)$ to n, one gets

$$\begin{aligned} u_j^{n+1} + \frac{\lambda\theta}{2}\Big[au_{j+1}^{n+1} - au_{j-1}^{n+1} &- |a|(1 - Q^n_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u^{n+1} \\ &+ |a|(1 - Q^n_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u^{n+1}\Big] = \text{RHS of (2.2)}. \end{aligned} \tag{3.15}$$

Let $d_j = u_j^{n+1} - u_j^n$; i.e., the "delta" notation, equation (3.15) can be written as

$$e_1 d_{j-1} + e_2 d_j + e_3 d_{j+1} = -\lambda(h^n_{j+\frac{1}{2}} - h^n_{j-\frac{1}{2}}) \tag{3.16}$$

where

$$e_1 = \frac{\lambda\theta}{2}\left[-a - |a|(1 - Q_{j-\frac{1}{2}})\right]^n \tag{3.16b}$$

$$e_2 = 1 + \frac{\lambda\theta}{2}\left[|a|(1 - Q_{j-\frac{1}{2}}) + |a|(1 - Q_{j+\frac{1}{2}})\right]^n \tag{3.16c}$$

$$e_3 = \frac{\lambda\theta}{2}\left[a - |a|(1 - Q_{j+\frac{1}{2}})\right]^n. \tag{3.16d}$$

The linearized form (3.16) is a spatially five-point scheme and yet it is a tridiagonal system of linear equations. This is because at the $(n+1)$th time level, only three points are involved; i.e., $u^{n+1}_{j-1}, u^{n+1}_j, u^{n+1}_{j+1}$. Although the coefficients e_i involve five points, they are at the nth time level.

The form of $\widetilde{C}^{\pm}_{j\pm\frac{1}{2}}$ for (3.15) is the same as (3.9) except the time index for the $Q_{j\pm\frac{1}{2}}$ and $r^{\pm}_{j\mp\frac{1}{2}}$ is dropped from $(n+1)$ to n for the implicit operator. Thus the linearized form (3.16) is still TVD. It was found in reference [7] that when time-accurate TVD schemes are used as a relaxation method for steady-state calculations, the convergence rate is degraded if limiters are present on the implicit operator. For steady-state applications, one can obtain another TVD linearized form by setting $Q_{j\pm\frac{1}{2}} = 0$ in (3.16); i.e., by redefining (3.16) by

$$e_1 = \frac{\lambda\theta}{2}(-a - |a|) \tag{3.17a}$$

$$e_2 = 1 + \lambda\theta(|a|) \tag{3.17b}$$

$$e_3 = \frac{\lambda\theta}{2}(a - |a|). \tag{3.17c}$$

Scheme (3.16a) together with (3.17) is spatially first-order accurate for the implicit operator and spatially second-order accurate for the explicit operator. It can be shown that (3.16a) together with (3.17) is still TVD. Equation (3.17) is considered because no limiter is present for the implicit operator.

3.4 Scheme for Nonlinear Scalar Hyperbolic Conservation Laws

To extend the scheme to nonlinear scalar problems, one simply defines a local characteristic speed

$$a_{j+\frac{1}{2}} = \begin{cases} \Delta_{j+\frac{1}{2}} f / \Delta_{j+\frac{1}{2}} u & \Delta_{j+\frac{1}{2}} u \neq 0 \\ (\partial f/\partial u)\big|_{u_j} & \Delta_{j+\frac{1}{2}} u = 0 \end{cases} \tag{3.18}$$

and redefines the $r^{\pm}_{j+\frac{1}{2}}$ in (3.2b) as

$$r^{-}_{j+\frac{1}{2}} = \frac{|a_{j-\frac{1}{2}}|\Delta_{j-\frac{1}{2}}u}{|a_{j+\frac{1}{2}}|\Delta_{j+\frac{1}{2}}u}, \qquad r^{+}_{j+\frac{1}{2}} = \frac{|a_{j+\frac{3}{2}}|\Delta_{j+\frac{3}{2}}u}{|a_{j+\frac{1}{2}}|\Delta_{j+\frac{1}{2}}u}. \tag{3.19}$$

Unlike the constant coefficient case, $a_{j+\frac{1}{2}}$ and $a_{j-\frac{1}{2}}$ are not always of the same sign. After considering all the possible combinations of the signs of the $a_{j+\frac{1}{2}}$ and $a_{j-\frac{1}{2}}$, a set of sufficient conditions on Q still can be of similar form to (3.13) and is

$$Q_{j+\frac{1}{2}} < 2 \tag{3.20a}$$

$$\left(Q_{j+\frac{1}{2}} \Big/ r^{-}_{j+\frac{1}{2}}\right) < \frac{2}{\lambda(1-\theta)|a_{j-\frac{1}{2}}|} - 2 \tag{3.20b}$$

$$\left(Q_{j+\frac{1}{2}} \Big/ r^{+}_{j+\frac{1}{2}}\right) < \frac{2}{\lambda(1-\theta)|a_{j+\frac{3}{2}}|} - 2 \tag{3.20c}$$

$$\lambda|a_{j+\frac{1}{2}}| < \frac{1}{(1-\theta)}. \tag{3.20d}$$

The numerical flux for the nonlinear case is

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[(f_{j+1}+f_j) - |a_{j+\frac{1}{2}}|(1-Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]. \tag{3.21}$$

Observe that when $a_{j+\frac{1}{2}} = 0$, the scheme has zero dissipation. One way is to approximate $|a_{j+\frac{1}{2}}|$ by a Lipschitz continuous function [2]. For example, instead of using (3.21), one can use

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[(f_{j+1}+f_j) - \psi(a_{j+\frac{1}{2}})(1-Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]. \tag{3.22a}$$

Here ψ is a function of $a_{j+\frac{1}{2}}$ and is of the form

$$\psi(z) = \begin{cases} |z| & |z| \geq \epsilon \\ (z^2+\epsilon^2)/2\epsilon & |z| < \epsilon \end{cases} \tag{3.23}$$

or

$$\psi(z) = \begin{cases} |z| & |z| \geq \epsilon \\ \epsilon & |z| < \epsilon \end{cases}, \tag{3.24}$$

where ϵ is a positive small number [9].

3.5 Alternate Scheme for the Nonlinear Scalar Hyperbolic Problem

A simpler but less rigorous way of extending the constant coefficient case to the nonlinear case is to define a local characteristic speed $a_{j+\frac{1}{2}}$ and keep the restriction on Q the same as in (3.18) and (3.20), but use the $r^{\pm}_{j+\frac{1}{2}}$ in (3.2b) instead of (3.19). The alternate form requires less computation than the previous more rigorous approach. The relative advantage and disadvantage between these two forms remain to be shown. However, numerical experiments with two-dimensional Euler equations of gas dynamics [7,9] show that the alternate form gives a better shock resolution than the former one (3.18)-(3.22).

As a side remark, a case of Harten's second-order explicit TVD scheme is contained in the class of limiters of Sweby [5] and Roe [6] and is equivalent to a case of Roe's second-order scheme of reference [6]. See Sweby's original manuscript [13] instead of the published version [5] for details. The numerical experiments of Yee et al. and Yee [7-9] with Harten's second-order TVD scheme indicate that the alternate form is favored over the more rigorous approach (3.18)-(3.22). This indication is further endorsed by Davis's numerical experiments [12] with similar examples. Since different limiters have different effects which are highly problem-dependent on the resolution of the numerical solutions, all these possibilities in extending to nonlinear equations require more extensive numerical testing before a clearer picture can be drawn.

3.6 Linearized Version of the Proposed Implicit Scheme for Nonlinear Equations

For the nonlinear case, the situation is slightly more complicated since the characteristic speed $\partial f/\partial u$ is no longer a constant. Substituting (3.22) in (2.2), one obtains

$$\begin{aligned} u_j^{n+1} &+ \frac{\lambda\theta}{2}\left[f_{j+1} - \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]^{n+1} \\ &- \frac{\lambda\theta}{2}\left[f_{j-1} - \psi(a_{j-\frac{1}{2}})(1 - Q_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u\right]^{n+1} = \text{RHS of (2.2)}. \end{aligned} \tag{3.25}$$

Here "RHS of (2.2)" means the right hand side of equation (2.2) with $h_{j+\frac{1}{2}}$ defined in (3.22). Unlike the constant coefficient case, one also has to linearized $f^{n+1}_{j\pm 1}$, $\psi(a^{n+1}_{j\pm\frac{1}{2}})$, and $Q^{n+1}_{j\pm\frac{1}{2}}$. Following the same procedure as in [9], two linearized versions of (3.25) are considered.

Linearized Nonconservative Implicit Form

Add and substract f_j^{n+1} on the left-hand-side of (3.25) and using the relation (3.18), one can rewrite (3.25) as

$$\begin{aligned} u_j^{n+1} &+ \frac{\lambda\theta}{2}\left[a^{n+1}_{j+\frac{1}{2}} - \psi(a^{n+1}_{j+\frac{1}{2}})(1 - Q^{n+1}_{j+\frac{1}{2}})\right]\Delta_{j+\frac{1}{2}}u^{n+1} \\ &- \frac{\lambda\theta}{2}\left[-a^{n+1}_{j-\frac{1}{2}} - \psi(a^{n+1}_{j-\frac{1}{2}})(1 - Q^{n+1}_{j-\frac{1}{2}})\right]\Delta_{j-\frac{1}{2}}u^{n+1} = \text{RHS of (2.2)}. \end{aligned} \tag{3.26}$$

By dropping the time index of the coefficients of $\Delta_{j\pm\frac{1}{2}}u^{n+1}$ from $(n+1)$ to n and letting $d_j = u_j^{n+1} - u_j^n$, (3.26) becomes

$$\bar{e}_1 d_{j-1} + \bar{e}_2 d_j + \bar{e}_3 d_{j+1} = -\lambda(h^n_{j+\frac{1}{2}} - h^n_{j-\frac{1}{2}}) \tag{3.27a}$$

where

$$\bar{e}_1 = \lambda\theta B^- \tag{3.27b}$$

$$\bar{e}_2 = 1 - \lambda\theta\Big(B^- + B^+\Big) \tag{3.27c}$$

$$\bar{e}_3 = \lambda\theta B^+ \tag{3.27d}$$

and

$$B^{\pm} = \frac{1}{2}\left[\pm a_{j\pm\frac{1}{2}} - \psi(a_{j\pm\frac{1}{2}})(1 - Q_{j\pm\frac{1}{2}})\right]^n \tag{3.27e}$$

Again equation (3.27) is a five-point scheme, and yet the coefficient matrix associated with the d_j's is tridiagonal. With this linearization, the method is no longer conservative. Therefore (3.27) is only applicable for steady-state calculations. The form of $\widetilde{C}^{\pm}_{j\pm\frac{1}{2}}$ (which is the counterpart of (3.9) for the nonlinear case) is undisturbed, except the time index is dropped from $(n+1)$ to n for the implicit operator. Thus the linearized form (3.27) is still TVD. Again, a spatially first-order accurate implicit operator similar to (3.17) can be obtained for (3.27) by setting $B^{\pm} = \frac{1}{2}[\pm a_{j\pm\frac{1}{2}} - \psi(\pm a_{j\pm\frac{1}{2}})]^n$. Since the limiter does not appear on the left-hand-side, improvement in efficiency over (3.17) might be possible [7,9]. This reduced form is especially useful for multidimensional, nonlinear, hyperbolic conservation laws.

Linearized Conservative Implicit Form

One can obtain a linearized conservative implicit form by using a local Taylor expansion about u^n and expressing $f^{n+1} - f^n$ in the following form

$$f^{n+1}_j - f^n_j = a^n_j(u^{n+1}_j - u^n_j) + O(\Delta t^2), \tag{3.28}$$

where $a^n_j = (\partial f/\partial u)^n_j$. Applying the first-order approximation of (3.28) and locally linearizing the coefficients of $(\Delta_{j\pm\frac{1}{2}}u)^{n+1}$ in (3.25) by dropping the time index from $(n+1)$ to n, one gets

$$\begin{aligned} u^{n+1}_j + \frac{\lambda\theta}{2}\Big[a^n_{j+1}u^{n+1}_{j+1} - a^n_{j-1}u^{n+1}_{j-1} - \psi(a^n_{j+\frac{1}{2}})(1 - Q^n_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u^{n+1} \\ + \psi(a^n_{j-\frac{1}{2}})(1 - Q^n_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u^{n+1}\Big] = \text{RHS of (2.2)}. \end{aligned} \tag{3.29}$$

Letting $d_j = u^{n+1}_j - u^n_j$, equation (3.29) can be written as

$$e_1 d_{j-1} + e_2 d_j + e_3 d_{j+1} = -\lambda(h^n_{j+\frac{1}{2}} - h^n_{j-\frac{1}{2}}), \tag{3.30a}$$

where

$$e_1 = \frac{\lambda\theta}{2}\left[-a_{j-1} - \psi(a_{j-\frac{1}{2}})(1 - Q_{j-\frac{1}{2}})\right]^n \tag{3.30b}$$

$$e_2 = 1 + \frac{\lambda\theta}{2}\left[\psi(a_{j-\frac{1}{2}})(1 - Q_{j-\frac{1}{2}}) + \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\right]^n \tag{3.30c}$$

$$e_3 = \frac{\lambda\theta}{2}\left[a_{j+1} - \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\right]^n. \tag{3.30d}$$

The linearized form (3.30) is conservative and is a spatially five-point scheme with a tridiagonal system of linear equations. Scheme (3.30) is applicable to transient as well as steady-state calculation. But the form of $\widetilde{C}^{\pm}_{j\pm\frac{1}{2}}$ for (3.30) is no longer the same as its nonlinear counter part. As of this writing, the conservative linearized form (3.30) has not been proven to be TVD.

For steady-state application, one can use a spatially first-order implicit operator for (3.30) by simply setting all the $Q_{j\pm\frac{1}{2}} = 0$; i.e., redefine (3.30b)-(3.30d) as

$$e_1 = \frac{\lambda\theta}{2}\left[-a_{j-1} - \psi(a_{j-\frac{1}{2}})\right]^n \tag{3.31a}$$

$$e_2 = 1 + \frac{\lambda\theta}{2}\left[\psi(a_{j-\frac{1}{2}}) + \psi(a_{j+\frac{1}{2}})\right]^n \tag{3.31b}$$

$$e_3 = \frac{\lambda\theta}{2}\left[a_{j+1} - \psi(a_{j+\frac{1}{2}})\right]^n. \tag{3.31c}$$

In reference [7,9], this type of linearization proved to be very useful for two-dimensional steady-state airfoil calculations.

IV. Extension to Hyperbolic System of Conservation Laws

Extension of the scalar scheme (3.14), (3.27), or (3.30) to systems of conservation laws can be accomplished by defining at each point a "local" system of characteristic fields, and then applying the scheme to each of the m scalar characteristic equations. Here m is the dimension of the hyperbolic system. Extension of the scalar scheme to higher than one-dimensional systems of conservation laws (for practical calculations) can be accomplished by an alternating direction implicit (ADI) method similar to the one described in Yee et al. and Yee [7,9]. Only the one-dimensional case will be described here.

Formal Extension

Consider a system of hyperbolic conservation laws

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0. \tag{4.1}$$

Here U and $F(U)$ are column vectors of m components. Let $A = \partial F/\partial U$ and the eigenvalues of A be $(a^1, a^2, ..., a^m)$. Denote R (R^{-1}) as the matrices whose columns are right (left) eigenvectors of A (A^{-1}). Let $U_{j+\frac{1}{2}}$ denote some symmetric average of U_j and U_{j+1} (see references [1,7,9] for a formula). Let $a^l_{j+\frac{1}{2}}$, $R_{j+\frac{1}{2}}$, $R^{-1}_{j+\frac{1}{2}}$ denote the quantities a^l, R, R^{-1} evaluated at $U_{j+\frac{1}{2}}$. Define

$$\alpha_{j+\frac{1}{2}} = R^{-1}_{j+\frac{1}{2}} \Delta_{j+\frac{1}{2}} U \tag{4.2}$$

as the forward difference (or the jump) of the local characteristic variables. With the above notation, a one-parameter family of TVD schemes (2.2) in the system case can be written as

$$U_j^{n+1} + \lambda\theta(H^{n+1}_{j+\frac{1}{2}} - H^{n+1}_{j-\frac{1}{2}}) = U_j^n - \lambda(1-\theta)(H^n_{j+\frac{1}{2}} - H^n_{j-\frac{1}{2}}). \tag{4.3a}$$

The numerical flux function $H_{j+\frac{1}{2}}$ expressed as

$$H_{j+\frac{1}{2}} = \frac{1}{2}\left(F_j + F_{j+1} - R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\right), \tag{4.3b}$$

where the elements of the $\Phi_{j+\frac{1}{2}}$ denoted by $\phi^l_{j+\frac{1}{2}}, l = 1, ..., m$ are

$$\phi^l_{j+\frac{1}{2}} = \psi(a^l_{j+\frac{1}{2}})(1 - Q^l_{j+\frac{1}{2}})\alpha^l_{j+\frac{1}{2}}, \tag{4.3c}$$

with $\psi(z)$ defined in (3.23), and

$$Q^l_{j+\frac{1}{2}} = Q\left[(r^-_{j+\frac{1}{2}})^l, (r^+_{j+\frac{1}{2}})^l\right] \tag{4.3d}$$

$$(r^-_{j+\frac{1}{2}})^l = \frac{|a^l_{j-\frac{1}{2}}|\alpha^l_{j-\frac{1}{2}}}{|a^l_{j+\frac{1}{2}}|\alpha^l_{j+\frac{1}{2}}} \tag{4.4a}$$

$$(r^+_{j+\frac{1}{2}})^l = \frac{|a^l_{j+\frac{3}{2}}|\alpha^l_{j+\frac{3}{2}}}{|a^l_{j+\frac{1}{2}}|\alpha^l_{j+\frac{1}{2}}} \tag{4.4b}$$

or

$$(r^-_{j+\frac{1}{2}})^l = \frac{\alpha^l_{j-\frac{1}{2}}}{\alpha^l_{j+\frac{1}{2}}} \tag{4.4c}$$

$$(r^+_{j+\frac{1}{2}})^l = \frac{\alpha^l_{j+\frac{3}{2}}}{\alpha^l_{j+\frac{1}{2}}}. \tag{4.4d}$$

Here $\alpha^l_{j+\frac{1}{2}}$ are the elements of (4.2). The corresponding conservative linearized form (3.30) for the system case can be expressed as

$$E_1 D_{j-1} + E_2 D_j + E_3 D_{j+1} = -\lambda(H^n_{j+\frac{1}{2}} - H^n_{j-\frac{1}{2}}), \tag{4.5a}$$

where

$$E_1 = \frac{\lambda\theta}{2}\Big(-A_{j-1} - K_{j-\frac{1}{2}}\Big)^n \tag{4.5b}$$

$$E_2 = I + \frac{\lambda\theta}{2}\Big(K_{j-\frac{1}{2}} + K_{j+\frac{1}{2}}\Big)^n \tag{4.5c}$$

$$E_3 = \frac{\lambda\theta}{2}\Big(A_{j+1} - K_{j+\frac{1}{2}}\Big)^n, \tag{4.5d}$$

with

$$K_{j\pm\frac{1}{2}} = \Big(R^{-1}\Omega R\Big)^n_{j\pm\frac{1}{2}} \tag{4.5e}$$

and

$$\Omega_{j\pm\frac{1}{2}} = \mathrm{diag}\Big[\psi(a^l_{j\pm\frac{1}{2}})(1 - Q^l_{j\pm\frac{1}{2}})\Big], \tag{4.5f}$$

or for the first-order left-hand-side

$$\Omega_{j\pm\frac{1}{2}} = \mathrm{diag}\Big[\psi(a^l_{j\pm\frac{1}{2}})\Big]. \tag{4.5g}$$

Here $\mathrm{diag}(z^l)$ denotes a diagonal matrix with diagonal elements z^l. Aside from computing the right hand side, the rest of the arithmetic involved for equation (4.5) is two matrix multiplications and a block tridiagonal inversion. The value of $\Omega_{j+\frac{1}{2}}$ in (4.5f) or (4.5g) can be saved while calculating the right hand side. Similarly, one can express the nonconservative linearized form (3.27) for the system case.

As a side remark, with the same procedure Roe's numerical flux in system case can be written

$$H^{LW}_{j+\frac{1}{2}} = \frac{1}{2}\Big(F_j + F_{j+1} - R_{j+\frac{1}{2}}\Phi^{LW}_{j+\frac{1}{2}}\Big), \tag{4.6a}$$

where the elements of the $\Phi^{LW}_{j+\frac{1}{2}}$ denoted by $\Big(\phi^l_{j+\frac{1}{2}}\Big)^{LW}, l = 1, ..., m$ are

$$\Big(\phi^l_{j+\frac{1}{2}}\Big)^{LW} = \Big[\lambda(a^l_{j+\frac{1}{2}})^2 Q^l_{j+\frac{1}{2}} + |a^l_{j+\frac{1}{2}}|(1 - Q^l_{j+\frac{1}{2}})\Big]\,\alpha^l_{j+\frac{1}{2}}. \tag{4.6b}$$

Simplified Version

As one can see, the main work for scheme (4.3a) with numerical flux function (4.3b) or (4.6) is the term $R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}$. A similar situation is also true for the corresponding conservative or nonconservative linearized form. Since

$$R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}} = R_{j+\frac{1}{2}}\Omega_{j+\frac{1}{2}}\alpha_{j+\frac{1}{2}} = R_{j+\frac{1}{2}}\Omega_{j+\frac{1}{2}}R^{-1}_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}U$$
$$= R_{j+\frac{1}{2}}\text{diag}\left[\psi(a^l_{j+\frac{1}{2}})(1 - Q^l_{j+\frac{1}{2}})\right]R^{-1}_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}U, \quad (4.7)$$

if somehow one can simplify $R\Omega R^{-1}$ to be a diagonal matrix, then the current implicit scheme should be competitive in terms of operation count with the widely distributed codes such as ARC2D (version 150) of Pulliam and Steger [14] and FLO52R of Jameson et al. [15]. Both ARC2D and FLO52R use a spatially three-point central differencing scheme with identical numerical dissipation terms, but they use different time-stepping methods for steady-state calculations.

Two possible ways of simplifying (4.7) are one suggested by Davis [12], and one suggested by Roe [11]. Davis suggests approximating (4.7) with

$$R_{j+\frac{1}{2}}\Omega_{j+\frac{1}{2}}R^{-1}_{j+\frac{1}{2}} \approx \overline{\omega}(r^-_{j+\frac{1}{2}}, r^+_{j+\frac{1}{2}})\mathrm{I}, \quad (4.8)$$

where $\overline{\omega}_{j+\frac{1}{2}} = \overline{\omega}(r^-_{j+\frac{1}{2}}, r^+_{j+\frac{1}{2}})$ is a scalar function of $r^-_{j+\frac{1}{2}}$, $r^+_{j+\frac{1}{2}}$ and $a^l_{j+\frac{1}{2}}$. The symbol I is a $m \times m$ identity matrix. Adapted to the current scheme, $\overline{\omega}_{j+\frac{1}{2}}$ can be expressed as

$$\overline{\omega}_{j+\frac{1}{2}} = \psi(\overline{a}_{j+\frac{1}{2}})\left[1 - Q(r^-_{j+\frac{1}{2}}, r^+_{j+\frac{1}{2}})\right], \quad (4.9)$$

where

$$\overline{a}_{j+\frac{1}{2}} = \max_l |a^l_{j+\frac{1}{2}}|. \quad (4.10)$$

In order to not have to compute R and R^{-1}, $r^{\pm}$ have to be redefined such that they are functions of gradients of the original variables U instead of $\alpha_{j+\frac{1}{2}}$. Davis suggests using

$$r^-_{j+\frac{1}{2}} = \frac{(\Delta_{j-\frac{1}{2}}U, \Delta_{j+\frac{1}{2}}U)}{(\Delta_{j+\frac{1}{2}}U, \Delta_{j+\frac{1}{2}}U)} = \frac{\sum_l(\Delta_{j-\frac{1}{2}}u^l)(\Delta_{j+\frac{1}{2}}u^l)}{\sum_l(\Delta_{j+\frac{1}{2}}u^l)^2} \quad (4.11a)$$

$$r^+_{j+\frac{1}{2}} = \frac{(\Delta_{j+\frac{3}{2}}U, \Delta_{j+\frac{1}{2}}U)}{(\Delta_{j+\frac{1}{2}}U, \Delta_{j+\frac{1}{2}}U)} = \frac{\sum_l(\Delta_{j+\frac{3}{2}}u^l)(\Delta_{j+\frac{1}{2}}u^l)}{\sum_l(\Delta_{j+\frac{1}{2}}u^l)^2}, \quad (4.11b)$$

where $(\cdot,\cdot)$ denotes the usual inner product on the real m-dimensional vector space R^m.

Roe suggests that instead of using the fastest wave (4.10), one might consider the use of the strongest wave in the following sense

$$\bar{a}_{j+\frac{1}{2}} = \left.\frac{\sum_{l=1}^{m} a^l (\alpha^l)^2}{\sum_{l=1}^{m} (\alpha^l)^2}\right|_{j+\frac{1}{2}} \tag{4.12a}$$

or

$$\bar{a}_{j+\frac{1}{2}} = \left.\frac{\sum_{l=1}^{m} (a^l \alpha^l)^2}{\sum_{l=1}^{m} (\alpha^l)^2}\right|_{j+\frac{1}{2}}. \tag{4.12b}$$

For the one-dimensional Euler equation of gas dynamics (perfect gas), the a^l are simply

$$a^1 = u - c, \qquad a^2 = u, \qquad a^3 = u + c, \tag{4.13}$$

and a formula for the α^l are

$$\alpha^1_{j+\frac{1}{2}} = \frac{1}{2c^2_{j+\frac{1}{2}}}\left(\Delta_{j+\frac{1}{2}} p - \rho_{j+\frac{1}{2}} c_{j+\frac{1}{2}} \Delta_{j+\frac{1}{2}} u\right) \tag{4.14a}$$

$$\alpha^2_{j+\frac{1}{2}} = \frac{1}{c^2_{j+\frac{1}{2}}}\left(c^2_{j+\frac{1}{2}} \Delta_{j+\frac{1}{2}} \rho - \Delta_{j+\frac{1}{2}} p\right) \tag{4.14b}$$

$$\alpha^3_{j+\frac{1}{2}} = \frac{1}{2c^2_{j+\frac{1}{2}}}\left(\Delta_{j+\frac{1}{2}} p + \rho_{j+\frac{1}{2}} c_{j+\frac{1}{2}} \Delta_{j+\frac{1}{2}} u\right). \tag{4.14c}$$

Here u is the velocity, c is the sound speed, p is the pressure, and ρ is the density.

As for the $r^{\pm}$, Roe suggests the use of $\alpha_{j\pm\frac{1}{2}}$ instead of $\Delta_{j\pm\frac{1}{2}} U$ in (4.11). In this case, the operation count between the local characteristic variable approach (4.5) and Roe's suggestion might be very competitive, since in Roe's suggestion one has to compute all the $\alpha^l_{j+\frac{1}{2}}$ anyway. Therefore, the main difference is between computing $R_{j+\frac{1}{2}}\Omega_{j+\frac{1}{2}} = R_{j+\frac{1}{2}} \mathrm{diag}\left[\psi(a^l_{j\pm\frac{1}{2}})(1 - Q^l_{j\pm\frac{1}{2}})\right]$ together with (4.4), or computing equations (4.8)-(4.9) and equation (4.12a) or (4.12b) together with

$$r^-_{j+\frac{1}{2}} = \frac{(\alpha_{j-\frac{1}{2}}, \alpha_{j+\frac{1}{2}})}{(\alpha_{j+\frac{1}{2}}, \alpha_{j+\frac{1}{2}})} = \frac{\sum_l \alpha^l_{j-\frac{1}{2}} \alpha^l_{j+\frac{1}{2}}}{\sum_l (\alpha^l_{j+\frac{1}{2}})^2} \tag{4.15a}$$

$$r^+_{j+\frac{1}{2}} = \frac{(\alpha_{j+\frac{3}{2}}, \alpha_{j+\frac{1}{2}})}{(\alpha_{j+\frac{1}{2}}, \alpha_{j+\frac{1}{2}})} = \frac{\sum_l \alpha^l_{j+\frac{3}{2}} \alpha^l_{j+\frac{1}{2}}}{\sum_l (\alpha^l_{j+\frac{1}{2}})^2}. \tag{4.15b}$$

In summary, for the system case three approaches are suggested here. (a) The more systematic approach (4.5) (from here on referred to as the local characteristic approach), (b) Davis's approach, and (c) Roe's approach. Davis's suggestion is by far the simpliest to implement and requires the least operation count but is the least rigorous. In numerical experiments with a two-dimensional shock reflection problem and a circular arc airfoil problem, Davis's approach shows good potential. Roe's suggestion, in the author's opinion, seems more rigorious than Davis's, but requires a similar operation count to the current suggestion. An obvious advantage of the Davis and Roe approaches over the local characteristic approach is that one does not have to deal with the problem of singular values of $r^{\pm}$ and in turn $Q_{j\pm\frac{1}{2}}$ are rarely zero. An analytical, but not necessarily practical, advantage of the current aproach over the other two approaches is that (4.5) collaspes into the exact scalar scheme for the case $m = 1$. The implication is that if one locally freezes the coefficients in (4.5), then the resulting constant coefficient system is TVD and convergent subject to the CFL restriction of $\frac{1}{(1-\theta)}$. The proof of this statement is readily available in reference [2]. The total variation definition for a vector grid function for the system case is

$$\mathrm{TV}(U) = \sum_j \sum_{l=1}^{m} |\alpha^l_{j+\frac{1}{2}}| \tag{4.16}$$

V. Concluding Remarks

The present paper was inspired by the work of Roe [11] and Davis [12], and is based on the work of Harten [1-2] and of Harten and the author [7-10]. A one-parameter family of explicit and implicit TVD schemes is reformulated so that a wider group of limiters is included. The current class of schemes as well as Roe and Davis's can be classified as non-upwind TVD schemes or symmetric TVD schemes. The main advantages of the present class of schemes over the ones suggested by Osher and Chakravarthy [16], Roe, or Davis are that (a) a wider class of time-differencing is included, (b) the implicit scheme allows a natural linearized procedure for a noniterative implicit procedure, and thus might have a greater potential for practical applications, especially for "stiff" problems, and (c) when applied to steady-state calculations, the numerical solution is independent of the time step. Furthermore, Roe's formulation can be considered as a member of this family by simply setting $\theta = 0$ and using the numerical fluxes (3.7). Extension of this class of schemes and Roe's schemes to a system of equations is straightforward. One can define a general numerical flux function

$$H^G_{j+\frac{1}{2}} = \frac{1}{2}\left(F_j + F_{j+1} - R_{j+\frac{1}{2}}\Phi^G_{j+\frac{1}{2}}\right), \tag{5.1a}$$

where the elements of the $\Phi^G_{j+\frac{1}{2}}$ denoted by $\left(\phi^l_{j+\frac{1}{2}}\right)^G, l = 1, ..., m$ are

$$\left(\phi^l_{j+\frac{1}{2}}\right)^G = \left[\lambda\beta(a^l_{j+\frac{1}{2}})^2 Q_{j+\frac{1}{2}} + |a^l_{j+\frac{1}{2}}|(1 - Q^l_{j+\frac{1}{2}})\right]\alpha^l_{j+\frac{1}{2}}, \tag{5.1b}$$

with $\beta = 1$ or 0 for transient calculations, and $\beta = 0$ for steady-state calculations. Here when $\beta = 0$, (5.1) is (4.3b), and when $\beta = 1$, (5.1) is the Roe's Lax-Wendroff numerical flux (4.6). The suggestions of Roe and Davis to simplify the system case are also discussed. Numerical testing and practical applications will be carried out in a separate paper.

The results of Roe, Davis, and the present formulation provide a more rational way of supplying additional numerical dissipation terms to the commonly known schemes such as the Lax-Wendroff type and some spatially symmetrical explicit and implicit types of schemes. Here the amount of work required to modify existing computer codes with the suggested numerical dissipation terms varies from very minor changes to moderate yet straightforward computer programming. The potential of improving the robustness and accuracy of a wide variety of physical applications is worth the effort of further pursuing the implementation of these ideas into the many existing user-oriented computer codes.

Appendix

(Equivalent Representation for the Conservative Dissipation Term)

The terms $r^{\pm}_{j+\frac{1}{2}}$ in (3.2b) are not defined if $\Delta_{j-\frac{1}{2}}u$ and $\Delta_{j+\frac{3}{2}}u$ are finite and $\Delta_{j+\frac{1}{2}}u = 0$. To avoid the use of extra logic in a computer implementation, it might be better to rewrite the terms $Q_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}u$ in equations (3.1), (3.8), and thereafter in the form

$$Q_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}u = \widehat{Q}_{j+\frac{1}{2}}. \tag{A.1}$$

Linear Scalar Hyperbolic Equations

The form $\widehat{Q}_{j+\frac{1}{2}}$ is a function of $\Delta_{j-\frac{1}{2}}u$, $\Delta_{j+\frac{1}{2}}u$, and $\Delta_{j+\frac{3}{2}}u$, but not $r^{\pm}$; i.e.,

$$\widehat{Q}_{j+\frac{1}{2}} = \widehat{Q}(\Delta_{j-\frac{1}{2}}u, \Delta_{j+\frac{1}{2}}u, \Delta_{j+\frac{3}{2}}u). \tag{A.2}$$

The numerical flux $h_{j+\frac{1}{2}}$ in (3.8) can be rewritten as

$$h_{j+\frac{1}{2}} = \frac{1}{2}\Big[a(u_{j+1} + u_j) - |a|(\Delta_{j+\frac{1}{2}}u - \widehat{Q}_{j+\frac{1}{2}})\Big]. \tag{A.3}$$

The corresponding $\widetilde{C}^{\pm}$ in (3.9) becomes

$$\widetilde{C}^{+}_{j-\frac{1}{2}} = a\left(1 - \frac{1}{2}\frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u} + \frac{1}{2}\frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u}\right), \qquad a > 0 \tag{A.4a}$$

$$\widetilde{C}^{-}_{j+\frac{1}{2}} = |a|\left(1 - \frac{1}{2}\frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} + \frac{1}{2}\frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u}\right), \qquad a < 0. \tag{A.4b}$$

The sufficient conditions for TVD in terms of $\widehat{Q}$ are

$$\frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u} - \frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u} < 2 \tag{A.5a}$$

$$\frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u} - \frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u} < \frac{2}{\lambda(1-\theta)a} - 2 \tag{A.5b}$$

$$a < \frac{1}{1-\theta} \tag{A.5c}$$

for $a > 0$, and

$$\frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} - \frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} < 2 \tag{A.6a}$$

$$\frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} - \frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} < \frac{2}{\lambda(1-\theta)|a|} - 2 \tag{A.6b}$$

$$|a| < \frac{1}{1-\theta} \tag{A.6c}$$

for $a < 0$. Assume $\left(\widehat{Q}_{j+\frac{1}{2}}/\Delta_{j+\frac{1}{2}}u\right)$ and $\left(\widehat{Q}_{j\pm\frac{1}{2}}/\Delta_{j\mp\frac{1}{2}}u\right)$ are always positive. Then sufficient conditions for (A.5) and (A.6) are

$$\frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} < 2, \tag{A.7a}$$

$$\frac{\widehat{Q}_{j+\frac{1}{2}}}{\Delta_{j-\frac{1}{2}}u} < \frac{2}{\lambda(1-\theta)|a|} - 2 \tag{A.7b}$$

$$\frac{\widehat{Q}_{j-\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}u} < \frac{2}{\lambda(1-\theta)|a|} - 2. \tag{A.7c}$$

The expressions (3.4)-(3.5) now become

$$\begin{aligned}\widehat{Q}\left(\Delta_{j-\frac{1}{2}}u, \Delta_{j+\frac{1}{2}}u, \Delta_{j+\frac{3}{2}}u\right) &= \operatorname{minmod}\left(\Delta_{j+\frac{1}{2}}u, \Delta_{j-\frac{1}{2}}u\right) \\ &\quad + \operatorname{minmod}\left(\Delta_{j+\frac{1}{2}}u, \Delta_{j+\frac{3}{2}}u\right) - \Delta_{j+\frac{1}{2}}u,\end{aligned} \tag{A.8a}$$

and

$$\widehat{Q}\left(\Delta_{j-\frac{1}{2}}u, \Delta_{j+\frac{1}{2}}u, \Delta_{j+\frac{3}{2}}u\right) = \operatorname{minmod}\left(\Delta_{j-\frac{1}{2}}u, \Delta_{j+\frac{1}{2}}u, \Delta_{j+\frac{3}{2}}u\right). \tag{A.8b}$$

Here

$$\mathrm{minmod}(x,y) = \mathrm{sgn}(x) \cdot \max\{0, \min[|x|, y \cdot \mathrm{sgn}(x)]\}. \tag{A.9}$$

In general, the "minmod" function of a list of arguments is equal to the smallest number in absolute value if the list of arguments is of the same sign, or is equal to zero if any argument is negative.

Nonlinear Scalar Hyperbolic Conservation Laws

For nonlinear problems, one way is to replace all the a's in equation (A.3)-(A.7) by $a_{j\pm\frac{1}{2}}$ accordingly. The value of $a_{j+\frac{1}{2}}$ is defined in (3.14). For a more rigorous approach $\widehat{Q}$ should also be redefined as

$$\widehat{Q}_{j+\frac{1}{2}} = \widehat{Q}(|a_{j-\frac{1}{2}}|\Delta_{j-\frac{1}{2}}u, |a_{j+\frac{1}{2}}|\Delta_{j+\frac{1}{2}}u, |a_{j+\frac{3}{2}}|\Delta_{j+\frac{3}{2}}u), \tag{A.10}$$

and $\Delta_{j\pm\frac{1}{2}}u$ should be replaced by $|a_{j\pm\frac{1}{2}}|\Delta_{j\pm\frac{1}{2}}u$ whenever they appear in equations (A.4)-(A.9). Similarly, the system case can be rewritten in terms of the $\widehat{Q}$'s.

References

[1] A. Harten, A High Resolution Scheme for the Computation of Weak Solutions of Hyperbolic Conservation Laws, NYU Report, Oct., 1981, and J. Comp. Phys., **49**, 357-393 (1983).

[2] A. Harten, On a Class of High Resolution Total-Variation-Stable Finite-Difference Schemes, NYU Report, Oct., 1982; SIAM J. Num. Anal., **21**, 1-23 (1984).

[3] B. van Leer, Towards the Ultimate Conservation Difference Scheme II, Monotonicity and Conservation Combined in a Second Order Scheme, J. Comp. Phys. **14**, 361-370 (1974).

[4] J.P. Boris and D.L. Book, Flux Corrected Transport. I. SHASTA, A Fluid Transport Algorithm That Works, J. Comp. Phys., **11**, 38-69 (1973).

[5] P.K. Sweby, High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws, SIAM J. Num. Analy., **21**, 995-1011 (1984).

[6] P.L. Roe, Some Contributions to the Modelling of Discontinuous Flows , Proceedings of the AMS-SIAM Summer Seminar on Large-Scale Computation in Fluid Mechanics June 27-July 8, 1983, Lectures in Applied Mathematics, **22** (1985).

[7] H.C. Yee, R.F. Warming and A. Harten, Implicit Total Variation Diminishing (TVD) Schemes for Steady-State Calculations, AIAA Paper No. 83-1902, Proc. of the AIAA 6th Computational Fluid Dynamics Conference, Danvers, Mass., July, 1983, also in J. Comp. Phys., **57**, 327-360 (1985).

[8] H.C. Yee, R.F. Warming and A. Harten, Application of TVD Schemes for the Euler Equations of Gas Dynamics, Proceedings of the AMS-SIAM Summer Seminar on Large-Scale Computation in Fluid Mechanics June 27-July 8, 1983, Lectures in Applied Mathematics, **22** (1985).

[9] H.C. Yee, Linearized Form of Implicit TVD Schemes for the Multidimensional Euler and Navier-Stokes Equations, submitted to *Advances in Hyperbolic Partial Differential Equations*, a special issue of Intl. J. Comp. Math. Appl., April 1985.

[10] H.C. Yee and A. Harten, Implicit TVD Schemes for Hyperbolic Conservation Laws in Curvilinear Coordinates , AIAA Paper No. 85-1513-CP, Proc. of the AIAA 7th Computational Fluid Dynamics Conference, Cinn., Ohio, July 15-17, 1985.

[11] P.L. Roe, Generalized Formulation of TVD Lax-Wendroff Schemes, ICASE Report No. 84-53, October 1984.

[12] S.F. Davis, TVD Finite Difference Schemes and Artificial Viscosity, ICASE Report No. 84-20, June 1984.

[13] P.K. Sweby, High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws, U.C.L.A. Report, June 1983, Los Angeles, Calif.

[14] T.H. Pulliam and J. Steger, Recent Improvements in Efficiency, Accuracy and Convergence for Implicit Approximate Factorization Algorithms, AIAA Paper No. 85-0360 (1985).

[15] A. Jameson, W. Schmidt and E. Turkel, Numerical Solutions of the Euler Equations by Finite Volume Methods Using Runge-Kutta Time-Stepping Schemes, AIAA Paper No. 81-1259 (1981).

[16] S. Osher and S. Chakravarthy, High Resolution Schemes and the Entropy Condition, SIAM J. Num. Analy., **21**, 955-984 (1984).

1. Report No. NASA TM-86775	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle GENERALIZED FORMULATION OF A CLASS OF EXPLICIT AND IMPLICIT TVD SCHEMES		5. Report Date July 1985
		6. Performing Organization Code
7. Author(s) H. C. Yee		8. Performing Organization Report No. 85343
		10. Work Unit No. T-6461
9. Performing Organization Name and Address Ames Research Center Moffett Field, CA 94035		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Memorandum
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		14. Sponsoring Agency Code 505-31-01-01-00-21

15. Supplementary Notes

Point of Contact: H. C. Yee, Ames Research Center, MS 202A-1, Moffett Field, CA 94035 FTS 464-5548 or (415)694-5548

16. Abstract

A one-parameter family of second-order explicit and implicit total variation diminishing (TVD) schemes is reformulated so that a simpler and wider group of limiters is included. The resulting scheme can be viewed as a symmetrical algorithm with a variety of numerical dissipation terms that are designed for weak solutions of hyperbolic problems. This is a generalization of Roe and Davis's recent works to a wider class of symmetric schemes other than Lax-Wendroff. The main properties of the present class of schemes are that they can be implicit, and when steady-state calculations are sought, the numerical solution is independent of the time step.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Numerical method, Finite difference method, Computational fluid dynamics, System of hyperbolic conversation laws, Weak solutions, Shock capturing, Conservative differencing, TVD schemes	Unlimited Subject category - 64

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	25	A02

*For sale by the National Technical Information Service, Springfield, Virginia 22161

End of Document