

**NASA Contractor Report 172614**

**ICASE REPORT NO. 85-28**

NASA-CR-172614  
19850022413

# ICASE

ON THE CONVERGENCE OF DIFFERENCE APPROXIMATIONS TO SCALAR  
CONSERVATION LAWS

Stanley Osher  
Eitan Tadmor

Contract No. NAS1-17070  
May 1985

**LIBRARY COPY**

JUL 21 1985

LANGLEY RESEARCH CENTER  
LIBRARY, NASA  
HAMPTON, VIRGINIA

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING  
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the Universities Space Research Association



National Aeronautics and  
Space Administration

Langley Research Center  
Hampton, Virginia 23665



NF00736

ON THE CONVERGENCE OF DIFFERENCE APPROXIMATIONS  
TO SCALAR CONSERVATION LAWS

Stanley Osher<sup>\*</sup>  
University of California, Los Angeles

and

Eitan Tadmor<sup>\*\*</sup>  
Tel-Aviv University  
and  
Institute for Computer Applications in Science and Engineering

**Abstract**

We present a unified treatment of explicit in time, two level, second order resolution, total variation diminishing, approximations to scalar conservation laws. The schemes are assumed only to have conservation form and incremental form. We introduce a modified flux and a viscosity coefficient and obtain results in terms of the latter. The existence of a cell entropy inequality is discussed and such an equality for all entropies is shown to imply that the scheme is an E scheme on monotone (actually more general) data, hence at most only first order accurate in general. Convergence for TVD-SOR schemes approximating convex or concave conservation laws is shown by enforcing a single discrete entropy inequality.

---

<sup>\*</sup>Research supported in part by NSF Grant No. MCS82-00788, ARO Grant No. DAAG 29-82-0090, NASA Grant No. NAG 1-270, and NASA Consortium Agreement No. NCA2-1R 390-403.

<sup>\*\*</sup>Research supported in part by the National Aeronautics and Space Administration under NASA Contract No. NAS1-17070 while the author was in residence at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23665.

**This Page Intentionally Left Blank**

## TABLE OF CONTENTS

<b>Introduction.....</b>	<b>1</b>
 <b>PART I. TOTAL VARIATION DIMINISHING SCHEMES</b>	
1. The Modified Flux and the Numerical Viscosity Coefficient.....	5
2. Three-Point TVD Schemes.....	9
3. Second-Order Resolution (SOR) Schemes.....	11
4. Two Recipes for the Construction of SOR-TVD Schemes.....	16
 <b>PART II. ON A CELL ENTROPY INEQUALITY</b>	
5. Godunov's Scheme and its Cell Entropy Inequality.....	25
6. The Modified Flux Function.....	28
7. Necessity and Sufficiency of E-Fluxes.....	44
8. Second-Order Resolution and a Cell Entropy Inequality.....	52
 <b>References.....</b>	 <b>65</b>

## Introduction

Recently there has been an enormous amount of activity related to the construction and analysis of finite difference approximations which approximate nonlinear hyperbolic systems of conservation laws and which are supposed to have the following properties:

- (1) Limit solutions which satisfy a geometric and/or analytic entropy condition.
- (2) A bound on the variation of the approximate solutions at least in the scalar and linear systems case. This bound is such as to imply the absence of spurious oscillations in the approximate solutions.
- (3) At least second-order accuracy in regions of smoothness, except for certain isolated points as described below.

Some examples of the successful computational consequences of this activity can be found in the proceedings of the sixth AIAA Computational Fluid Dynamics Conference, and elsewhere see e.g., the bibliography in [21].

Some of the earliest work in the design of schemes having properties (2) and (3) above was done by Van Leer [27], [28]. There he introduced the concepts of flux limiters and higher order Riemann solvers. Recently Harten [10], [11] obtained conditions which he showed to be compatible with second-order accuracy, and which guarantee that a scalar one-dimensional scheme is TVD--total variation diminishing. He constructed a scheme having that property, and formally extended it to systems using a field-by-field limiter and Roe's decomposition [22].

We would also like to mention the work of Boris and Book [1] concerning FCT schemes. They also used flux limiters to suppress oscillations in their schemes.

Harten's construction in [10] was done first for a fully discrete, explicit in time approximation. P. Sweby [24] has investigated the properties of various limiters in this context.

We shall use the term "high resolution scheme" to mean a formal extension to systems via a field-by-field decomposition of a scalar, higher than first-order accurate, variation diminishing scheme. These schemes do not, in general satisfy the entropy condition e.g. expansion shocks exist as stable solutions of high resolution schemes based on Roe's (unmodified) scheme. In [20], Osher's decomposition and certain flux limiters were used to prove that limit solutions to a class of semi-discrete, time continuous high resolution schemes do satisfy the entropy condition for hyperbolic systems of conservation laws. Convergence of other classes of semi-discrete, time continuous, high resolution approximations to scalar convex conservation laws was proven in [19] and [20].

A systematic recipe for constructing semi-discrete high resolution schemes whose formal accuracy is higher than two (away from the isolated points) using a minimal band width, was presented in [21].

In the present paper we are considering two-level, (and for simplicity only), explicit finite difference approximations to a scalar conservative law having two properties:

- (1) Lax-Wendroff conservative form [15], and
- (2) Roe's incremental form [22].

In Section 1, we introduce a modified flux and a viscosity coefficient, both grid dependent quantities, with which we set up and advocate a third form of these schemes the viscosity form. A sufficient TVD criterion is then presented in terms of this form.

In Section 2 we consider schemes which have a 3-point stencil. In this case the modified flux coincides with the original (differential equation's) flux; accuracy is consequently limited to first-order.

Thus, in Section 3, we consider wider stencils. A comprehensive framework of second-order resolution (SOR), total variation diminishing (TVD) schemes is presented. The viscous form of the scheme again plays a key role here.

Two concrete examples for the construction of SOR-TVD schemes according to the above guidelines are detailed in section 4; both approaches use the viscous form of the underlying 3 point TVD schemes. Special attention is called to the second ("piecewise constant viscosity modification") recipe in Corollary (4.9), which results in an easily implemented and highly attractive two step formulation.

These four sections comprised Part I. In our attempt at a unified treatment of this subject, we derived (and in some cases rederived) some notable conclusions:

- (i) Three point TVD schemes are, at most, first-order accurate.
- (ii) Accuracy at nonsonic critical points is limited to first-order.
- (iii) Our piecewise linear flux correction (Theorem 4.3) extends Harten's recipe [10], under a CFL of limitation  $2/3$ , (Remark 4.9), in agreement with that found by Sweby in [24].
- (iv) Our recipe for converting three point TVD schemes to five point SOR-TVD schemes seems quite general and attractive from both a computational and a theoretical point of view. The latter point of view will be used in Part II.

The discussion concerning a cell entropy inequality begins in Section 5 using the canonical Godunov scheme. Section 6 contains the heart of our

discussion. The grid dependent modified flux of Section 1 is extended to a modified flux function defined on the intervals connecting the grid values. The definition involves a piecewise linear function (i.e., the "double wing" seen in figures (6.4)) and the remark made after the statement of Theorem (6.11), which takes into account the presence or absence of critical points.

The schemes under consideration are then expressed as convex combinations of Riemann problem solvers, i.e., of Godunov type schemes, using the above modified flux functions. We obtain an entropy in cell inequality involving a familiar (integral) residual term.

Section 7 discusses the significance of  $E$  fluxes (introduced in [18], c.f also [26]). We show that for any approximate entropy in cell inequality to be valid for all intervals, the underlying scheme must have an  $E$  flux hence be at most first-order accurate (Theorem (7.2)). Conversely, the existence of an  $E$  flux implies the nonpositivity of the residual term mentioned above, hence implies a general cell entropy inequality. To obtain convergence for SOR schemes we give up the requirement that all discrete entropy inequalities be valid. This limits our convergence proof to convex (or concave) conservation laws, as in [19], [20].

In Section 8 we treat SOR-TVD schemes satisfying a single quadratic cell entropy inequality. Here we use the viscosity modification recipe of Section 4. A cell entropy inequality follows by estimating the residual integral presented in Section 6. The rather delicate tuning of the modified flux so as to comply with the three criteria SOR, TVD, and entropy inequality, leads us to a wide class of explicit in time SOR-TVD schemes (again only for the convex or concave scalar case).

The main convergence results are contained in Theorems 8.4 and 8.5



## PART I. TOTAL VARIATION DIMINISHING SCHEMES

### 1. The Modified Flux and the Numerical Viscosity Coefficient

We study two-step difference schemes of the form

$$v_v(t+k) = H(v_{v-p}(t), \dots, v_{v+p}(t); f, \lambda), \quad (1.1)$$

serving as consistent approximations to the scalar conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} = 0. \quad (1.2)$$

Here,  $v_v(t) \equiv v(x_v, t)$  denotes the approximation value at the grid point  $(x_v \equiv v\Delta x, t)$ ,  $k \equiv \Delta t$ , and  $\Delta x$  are respectively, the temporal and spatial mesh size with fixed mesh ratio  $\lambda = k/\Delta x$ , and  $p$  a natural number.

We postulate

#### Assumption 1.1

(i) The scheme (1.1) admits a conservative form in the sense of Lax-Wendroff [15]

$$H(v_{v-p}, \dots, v_{v+p}; f, \lambda) = v_v - \lambda(h_{v+1/2} - h_{v-1/2}) \quad (1.3)$$

where  $h_{v+1/2}$  stands for the Lipschitz continuous numerical flux

$$h_{v+1/2} \equiv h(v_{v-p+1}, \dots, v_{v+p}; f, \lambda),$$

consistent with the differential one

$$h(w, w, \dots, w; f, \lambda) = f(w).$$

(ii) The scheme (1.1) can be also written in an incremental form

$$H(v_{v-p}, \dots, v_{v+p}; f, \lambda) = v_v + C_{v+1/2}^+ \Delta v_{v+1/2} - C_{v-1/2}^- \Delta v_{v-1/2}; \quad (1.4)$$

where we use the standard notation

$$\Delta v_{v+1/2} = v_{v+1} - v_v.$$

Equating the right hand sides of (1.3) and (1.4), rearranging and dividing by  $\lambda$ , we find

$$h_{v+1/2} + \frac{1}{\lambda} C_{v+1/2}^+ \Delta v_{v+1/2} = h_{v-1/2} + \frac{1}{\lambda} C_{v-1/2}^- \Delta v_{v-1/2}. \quad (1.5a)$$

We term the equated grid dependent quantities in (1.5a) as the modified flux associated with scheme (1.1):

$$\begin{aligned} g_v &= h_{v\pm 1/2} + \frac{1}{\lambda} C_{v\pm 1/2}^\pm \Delta v_{v\pm 1/2} \\ &= \frac{1}{2} [h_{v-1/2} + h_{v+1/2} + \frac{1}{\lambda} C_{v-1/2}^- \Delta v_{v-1/2} + \frac{1}{\lambda} C_{v+1/2}^+ \Delta v_{v+1/2}]. \end{aligned} \quad (1.5b)$$

As in [25, Section 2], we now use the consistency relation

$$C_{v+1/2}^- - C_{v+1/2}^+ = \lambda \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}}, \quad (1.6a)$$

to conclude that with a given modified flux,  $g_v$ , there is only one degree of freedom in setting up the recipe of the difference scheme (1.1): this will be manifested in terms of the quantity,

$$Q_{v+1/2} = C_{v+1/2}^- + C_{v+1/2}^+. \quad (1.6b)$$

Indeed, by averaging (1.3) and (1.4) we find

$$v_v(t+k) = v_v(t) - \frac{\lambda}{2} (h_{v+1/2} - h_{v-1/2}) + \frac{1}{2} (C_{v+1/2}^+ \Delta v_{v+1/2} - C_{v+1/2}^- \Delta v_{v-1/2});$$

adding and subtracting  $\frac{1}{2} (C_{v+1/2}^- \Delta v_{v+1/2} - C_{v-1/2}^+ \Delta v_{v-1/2})$  on the right-hand side, we arrive at

$$\begin{aligned} v_v(t+k) = v_v(t) - \frac{\lambda}{2} [ (h_{v+1/2} + \frac{1}{\lambda} C_{v+1/2}^- \Delta v_{v+1/2}) - (h_{v-1/2} + \frac{1}{\lambda} C_{v-1/2}^+ \Delta v_{v-1/2}) ] \\ + \frac{1}{2} [ (C_{v+1/2}^- + C_{v+1/2}^+) \Delta v_{v+1/2} - (C_{v-1/2}^- + C_{v-1/2}^+) \Delta v_{v-1/2} ]. \end{aligned}$$

In view of (1.5b) and (1.6b), the scheme (1.1) is thus finally recast into its viscous form

$$v_v(t+k) = v_v(t) - \frac{\lambda}{2} (g_{v+1} - g_{v-1}) + \frac{1}{2} [\Delta(Q_{v-1/2} \Delta v_{v-1/2}(t))], \quad (1.7a)$$

expressed in terms of the numerical viscosity coefficient [10,26]

$$Q_{v+1/2} = C_{v+1/2}^- + C_{v+1/2}^+ = \lambda \frac{g_v + g_{v+1} - 2h_{v+1/2}}{\Delta v_{v+1/2}}. \quad (1.7b)$$

Let  $TV[v(t)] \equiv \sum_v |v_{v+1}(t) - v_v(t)|$  denote the total variation of the computed solution at time level  $t$ ; the following lemma provides us with sufficient criterion for the scheme (1.1) to be total-variation diminishing (TVD), in the spirit of [10].

LEMMA 1.2 (Total-variation diminishing)

The scheme (1.1) has a diminishing total-variation

$$TV[v(t+k)] \leq TV[v(t)],$$

provided its numerical viscosity coefficient  $Q_{v+1/2}$ , satisfies

$$\lambda \left| \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2} \leq 1. \quad (1.8)$$

Proof. By averaging (1.6a) and (1.6b) we find

$$C_{v+1/2}^{\pm} = \frac{1}{2} \left( Q_{v+1/2} \mp \lambda \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right). \quad (1.9)$$

In view of these last relations, the inequalities (1.8) boil down to

$$C_{v+1/2}^{+} \geq 0, \quad C_{v+1/2}^{-} \geq 0, \quad 1 - C_{v+1/2}^{-} - C_{v+1/2}^{+} \geq 0, \quad (1.10)$$

and TVD follows along the lines of [10].

## 2. Three-Point TVD Schemes

In the case of 3-point schemes,  $p = 1$ , it was shown in [25; Lemma 2.1] that there exists one and only one incremental form (1.4), whose incremental coefficients are

$$C_{v+1/2}^+ = \lambda \frac{f_v - h_{v+1/2}}{\Delta v_{v+1/2}}, \quad C_{v+1/2}^- = \lambda \frac{f_{v+1} - h_{v+1/2}}{\Delta v_{v+1/2}}. \quad (2.1)$$

Inserted into (1.5b) we find that the modified flux in this case coincides with the original one,  $g_v = f_v$ , and the scheme is therefore completely determined by its numerical viscosity coefficient,  $Q_{v+1/2}$ ,

$$v_v(t+k) = v_v(t) - \frac{\lambda}{2} (f_{v+1} - f_{v-1}) + \frac{1}{2} [\Delta(Q_{v-1/2} \Delta v_{v-1/2})]. \quad (2.2)$$

Abbreviating  $a(w) = \frac{df}{dw}(w)$ , the following notation will be used throughout

$$a_{v+1/2} = \begin{cases} \frac{\Delta f_{v+1/2}}{\Delta v_{v+1/2}} & \Delta v_{v+1/2} \neq 0 \\ a \frac{v_v + v_{v+1}}{2} & \Delta v_{v+1/2} = 0. \end{cases} \quad (2.3)$$

The TVD constraint (1.8) now reads

$$\lambda |a_{v+1/2}| \leq Q_{v+1/2} \leq 1, \quad (2.4)$$

and the following list quotes the most frequently referred to schemes in this

3-point TVD category [6,7,16,22]:

$$Q_{v+1/2}^{LF} = 1, \quad (\text{Lax-Friedrichs scheme}) \quad (2.5a)$$

$$Q_{v+1/2}^{EO} = \frac{\lambda}{\Delta v_{v+1/2}} \int_{v_v}^{v_{v+1}} |a(v)| dv, \quad (\text{Engquist-Osher scheme}) \quad (2.5b)$$

$$Q_{v+1/2}^G = \lambda_{(v-v_v) \cdot (v-v_{v+1})}^{\text{Max}} \leq 0 \frac{f_v + f_{v+1} - 2f(v)}{\Delta v_{v+1/2}}, \quad (\text{Godunov scheme}) \quad (2.5c)$$

$$Q_{v+1/2}^R = \lambda |a_{v+1/2}|. \quad (\text{Roe-Murman scheme}) \quad (2.5d)$$

In fact, in the special case under consideration of 3-point schemes, condition (2.4) is necessary as well as sufficient for TVD, see [25, Corollary 2.3]. Hence, the following 3-point schemes are not TVD ones

$$Q_{v+1/2}^{LW} = (\lambda a_{v+1/2})^2, \quad (\text{Lax-Wendroff scheme}) \quad (2.6a)$$

$$Q_{v+1/2}^E = 0, \quad (\text{Forward-Euler scheme}). \quad (2.6b)$$

A special significance is attached to the Lax-Wendroff scheme

$$v_v(t+k) = v_v(t) - \frac{\lambda}{2} (f_{v+1} - f_{v-1}) + \frac{1}{2} [\Delta((\lambda a_{v-1/2})^2 \Delta v_{v-1/2})]; \quad (2.7)$$

when seeking second-order accurate schemes, the further limitation placed on  $Q_{v+1/2}$  in the 3-point case, singles out the Lax-Wendroff choice, (2.6a). In

view of the above, we therefore conclude that 3-point TVD schemes are at most first-order accurate [10, 18, 25].

At this point we are widening our discussion to include schemes whose stencil occupies more than three points: the further freedom in setting up the modified flux  $g_v$  in this case, would enable us to achieve higher (than one) degree of accuracy.

### 3. Second-Order Resolution Schemes

We start by rewriting the modified flux,  $g_v$ , in terms of the correction  $\tilde{g}_v$  to the original one,

$$g_v = f_v + \frac{1}{\lambda} \tilde{g}_v. \quad (3.1)$$

Inserted into (1.7a), our scheme now reads

$$\begin{aligned} v_v(t+k) &= v_v(t) - \frac{\lambda}{2} [(f_{v+1} + \frac{1}{\lambda} \tilde{g}_{v+1}) - (f_{v-1} + \frac{1}{\lambda} \tilde{g}_{v-1})] \\ &\quad + \frac{1}{2} [\Delta(Q_{v-1/2} \Delta v_{v-1/2})] \\ &= v_v(t) - \frac{\lambda}{2} (f_{v+1} - f_{v-1}) + \frac{1}{2} [\Delta(Q_{v-1/2} \Delta v_{v-1/2})], \end{aligned} \quad (3.2)$$

where  $Q_{v+1/2}$  stands for the modified viscosity coefficient given by

$$Q_{v+1/2} = Q_{v+1/2} - \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} = \lambda \left[ \frac{f_v + f_{v+1} - 2h_{v+1/2}}{\Delta v_{v+1/2}} \right]. \quad (3.3)$$

In other words, we rewrite our scheme w.r.t the original flux,  $f_v$ , modifying the viscosity coefficient instead.

Remark 3.1. The viscosity coefficient discussion in [10], corresponds to the modified one in (3.3); indeed, in the 3-point case treated in [25] where  $\tilde{g}_v = \lambda(g_v - f_v) = 0$ , it coincides with (regular) viscosity in (1.7b).

Comparison of (3.2) and the Lax-Wendroff scheme (2.7) leads us to the following second-order accuracy requirement

$$Q_{v+1/2} \equiv Q_{v+1/2} - \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} = \lambda^2(a^2)_{v+1/2}; \quad (3.4a)$$

the deliberately vague notation of  $\lambda^2(a^2)_{v+1/2}$  on the right, stands for  $Q_{v+1/2}^{LW} = \lambda^2(a_{v+1/2})^2$  modulo first-order errors, i.e.,

$$\lambda^2(a^2)_{v+1/2} \equiv \lambda^2(a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2}. \quad (3.4b)$$

Together with the TVD constraint (1.8), we finally arrive at a general description of second-order TVD schemes, which is summarized in the following

LEMMA 3.2 (Second-order TVD schemes).

Consider the difference scheme (1.1) written in its viscous form

$$v_v(t+k) = v_v(t) - \frac{\lambda}{2} (g_{v+1} - g_{v-1}) + \frac{1}{2} [\Delta(Q_{v-1/2} \Delta v_{v-1/2})], \quad (3.5a)$$

with modified flux,  $g_v$ , and viscosity coefficient,  $Q_{v+1/2}$ , given respectively by



$$g_v = f_v + \frac{1}{\lambda} \tilde{g}_v \quad (3.5b)$$

$$Q_{v+1/2} = \lambda^2 (a^2)_{v+1/2} + \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}}. \quad (3.5c)$$

The scheme (3.5) is second-order TVD, provided the correction terms  $\tilde{g}_v$  and  $\lambda^2 (a^2)_{v+1/2}$  on the right-hand side of (3.5b-c) satisfy the following two requirements of

(i) second-order accuracy:

$$\lambda^2 (a^2)_{v+1/2} = (\lambda a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2}; \quad (3.6a)$$

(ii) total-variation diminishing

$$\left| \lambda a_{v+1/2} + \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq \lambda^2 (a^2)_{v+1/2} + \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} \leq 1. \quad (3.6b)$$

As we shall see below, one cannot satisfy both requirements (3.6a) and (3.6b) at the nonsonic critical grid values  $v_v$  where  $\Delta v_{v-1/2} \cdot \Delta v_{v+1/2} < 0 \neq a(v_v)$ , and, therefore, second-order accuracy must be given up at these values. Difference schemes with (formal) second-order accuracy at all but those critical grid values are classified as having second-order resolution, after Harten [10]: thanks to the TVD property, no new such first-order accurate critical grid values are added during the computation (since the scheme is monotonicity preserving [10,25]), and the overall second-order accuracy does not seem to be degraded in this case, at least in the  $L^1$  norm.

There are various approaches to the construction of second-order resolution (abbreviated hereafter SOR) schemes, the main three of which are

the following:

(i) The MUSCL approach [27,19], directly modifies the first-order numerical flux  $h_{v+1/2}$ , using a Godunov-like solver for a limited slope piecewise linear grid data.

(ii) The modified equation approach [10,11] directly modifies the flux  $f(v_v)$ , based on considerations of the modified equation associated with a first-order scheme, so that the limited modified flux constructed, guarantees second-order TVD resolution.

(iii) The flux limiter approach [24] directly modifies the first-order incremental coefficients  $C_{v+1/2}^{\pm}$ , using a class of flux limiters which preserve TVD and maintain second-order resolution; in particular, such flux limiters can be chosen to interpret MUSCL and modified flux-type schemes due to Van Leer, Roe, Harten, Chakravarthy, Osher and others, see [24].

All the above approaches can be entertained of course, within the general framework provided in Lemma 3.2. A common "limiting" feature shared by all of these approaches can be directly derived from the TVD constraints in that lemma: the grid values  $\tilde{g}_v$  read in (3.5), are expected to form a first-order "smooth" grid function correction, at least at the generic noncritical zones; since  $(a^2)_{v+1/2}$  is determined up to first-order perturbations, see (3.6a), we can absorb such a perturbation,  $\pm \Delta \tilde{g}_{v+1/2} / \Delta v_{v+1/2} = O(|\Delta v|_{v+1/2})$ , into the middle term of (3.6b), and the right-hand inequality now reads

$$\lambda^2 (a^2)_{v+1/2} + 2 \frac{\tilde{g}_{v+1}}{\Delta v_{v+1/2}} \leq 1, \quad \lambda^2 (a^2)_{v+1/2} + 2 \frac{\tilde{g}_v}{\Delta v_{v+1/2}} \leq 1. \quad (3.7a)$$

The left-hand inequality in (3.6b) gives us

$$\lambda a_{v+1/2} - \lambda^2 (a^2)_{v+1/2} \leq 2 \frac{\tilde{g}_v}{\Delta v_{v+1/2}}, \quad -\lambda a_{v+1/2} - \lambda^2 (a^2)_{v+1/2} \leq 2 \frac{\tilde{g}_{v+1}}{\Delta v_{v+1/2}}. \quad (3.7b)$$

We shift indices in the first and fourth inequalities in (3.7). The resulting inequalities boil down to the following inequality to be satisfied by TVD schemes:

$$\frac{1}{2} (\pm \lambda a_{v\pm 1/2} - \lambda^2 (a^2)_{v\pm 1/2}) \leq \frac{\tilde{g}_v}{\Delta v_{v\pm 1/2}} \leq \frac{1}{2} (1 - \lambda^2 (a^2)_{v\pm 1/2}). \quad (3.7c)$$

Hence,  $\tilde{g}_v / \Delta v_{v\pm 1/2}$  must lie between the above two bounds, both determined up to first-order perturbations. In particular, in view of the second-order accuracy requirement in (3.6a), the sign of the parenthesis on the left of (3.7c) is determined by that of the first term,  $\pm \lambda a_{v\pm 1/2}$ ; if further, we are at a nonsonic value,  $a(v_v) \neq 0$ , then this first term cannot be absorbed as a first-order perturbation of the second one, and (3.7c) requires the middle terms  $\tilde{g}_v / \Delta v_{v\pm 1/2}$  to be positive. This implies that the second-order accuracy must be given up at the nonsonic critical values  $v_v$  where  
 $\Delta v_{v-1/2} \cdot \Delta v_{v+1/2} < 0 \neq a(v_v)$ , [25,20].

We now describe two specific recipes which convert arbitrary 3-point TVD schemes into second-order resolution ones, in the spirit of the above guidelines. Indeed, although the reasoning may be different, the various "conversion" recipes including the two below, end up with difference schemes which bear close similarities to each other, as dictated by the framework provided in Lemma 3.2.

#### 4. Two Recipes for the Construction of SOR-TVD Schemes

We start by identifying a 3-point TVD scheme by its numerical viscosity coefficient  $Q_{v+1/2} = Q(f; v_v, v_{v+1}, \lambda)$ ; it is considered here as a functional in the original flux  $f(\cdot)$  with the further dependence on the grid values  $v_v, v_{v+1}$  and the mesh ratio  $\lambda$ . In view of the necessity of the TVD constraint, (2.4), we may assume without loss of generality that the following functional inequality is obeyed:

$$\lambda \left| \frac{z(w_2) - z(w_1)}{w_2 - w_1} \right| \leq Q(z; w_1, w_2, \lambda). \quad (4.1)$$

Thus, the TVD requirement (2.4) is reduced in this case to the CFL-like condition read in the right of (2.4)

$$Q_{v+1/2}(f) \equiv Q(f; v_v, v_{v+1}, \lambda) \leq 1. \quad (4.2)$$

Remark 4.1. The above description seems to exclude several difference schemes, where it is actually the inequality on the left of (2.4),

$\lambda |\Delta f_{v+1/2} / \Delta v_{v+1/2}| \leq Q_{v+1/2}(f)$ , which leads to the CFL limitation, as in the case, for example, of Lax-Friedrichs scheme (2.5a) with,  $Q_{v+1/2}^{LF} = 1$ . Together with the TVD constraint (2.4), however, the Lax-Friedrichs scheme, for example, can be equivalently represented by a numerical viscosity coefficient  $Q_{v+1/2} = \text{Max}(1, \lambda |\Delta f_{v+1/2}| / |\Delta v_{v+1/2}|)$  which fits into our above interpretation. In general, an expansion of  $Q(\cdot; \cdot, \cdot, \lambda)$  in powers of  $\lambda$  is called for, but we shall not elaborate on that here.

To construct an SOR-TVD scheme, the two first-order correction terms  $\tilde{g}_v$  and  $\lambda^2 (a^2)_{v+1/2}$  are to be determined; see (3.5). Taking advantage of

relation (3.5c), we shall also make use of the further (first-order) flexibility in determining the viscosity coefficient,  $Q_{v+1/2}^{SOR}$ .

Let us denote

$$s_v = \frac{1}{2} (s_{v-1/2} + s_{v+1/2}), \quad s_{v+1/2} \equiv \text{sgn}(\Delta v_{v+1/2}), \quad (4.3a)$$

and, in view of (3.7c), set the flux correction to be

$$\tilde{g}_v = \frac{s_v}{2} \text{Min} ([Q_{v\pm 1/2}(f) - \lambda^2(a^2)_{v\pm 1/2}] \cdot |\Delta v_{v\pm 1/2}|). \quad (4.3b)$$

Here, the term  $\lambda^2(a^2)_{v+1/2}$  is chosen so that

$$\lambda^2(a^2)_{v+1/2} = (\lambda a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2}, \quad \lambda^2(a^2)_{v+1/2} \leq Q_{v+1/2}(f) \quad (4.3c)$$

with otherwise arbitrary first-order perturbation  $O(|\Delta v|)_{v+1/2}$ . The modified flux is then given by

$$g_v = f(v_v) + \frac{1}{\lambda} \tilde{g}_v, \quad (4.3d)$$

and we are considering the difference scheme whose viscous form is expressed in terms of that modified flux

$$v_v(t+k) = v_v(t) - \frac{\lambda}{2} (g_{v+1} - g_{v-1}) + \frac{1}{2} [\Delta(Q_{v-1/2}^{SOR} \Delta v_{v-1/2})]. \quad (4.3e)$$

We have

THEOREM 4.2.

The difference scheme (4.3) is SOR-TVD provided its viscosity coefficient,  $Q_{v+1/2}^{SOR}$ , satisfies

$$\left| \lambda a_{v+1/2} + \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2}^{SOR} \leq 1 \quad (4.4a)$$

$$|Q_{v+1/2}^{SOR} - Q_{v+1/2}(f)| = O(|\Delta v|)_{v+1/2}. \quad (4.4b)$$

Remark. We note that the definition of  $(a^2)_{v+1/2}$  in (4.3c) still allows the flexibility of (some limited amount of) first-order perturbation as long as  $Q_{v+1/2}(f) - \lambda^2(a^2)_{v+1/2}$  remains nonnegative; the latter guarantees that the sign of  $\tilde{g}_v$  agrees with  $s_{v\pm 1/2}$ .

Proof. The first condition, (4.4a) is nothing but the TVD requirement (3.6b); consult (1.8). Away from the critical values where  $s_v = 0$ , (4.4b) implies second-order accuracy; consult (3.4),

$$\begin{aligned} Q_{v+1/2}^{SOR} - \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} &= Q_{v+1/2}^{SOR} - [Q_{v+1/2}(f) - \lambda^2(a^2)_{v+1/2}] + O(|\Delta v|)_{v+1/2} \\ &= [Q_{v+1/2}^{SOR} - Q_{v+1/2}(f)] + \lambda^2(a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2} \\ &= \lambda^2(a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2}. \end{aligned}$$

Two concrete choices for SOR viscosity will now be discussed.

As a building block for the first, we introduce the piecewise linear flux correction

$$\tilde{g}_{v+1/2}(v) = \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} (v - v_v) + \tilde{g}_v \quad (4.5a)$$

and use the original 3-point viscosity functional, setting

$$Q_{v+1/2}^{SOR} = Q(f(\cdot) + \frac{1}{\lambda} \tilde{g}_{v+1/2}(\cdot); v_v, v_{v+1}, \lambda). \quad (4.5b)$$

Lipschitz continuity of the viscosity functional implies<sup>(1)</sup>

$$\left| Q_{v+1/2}^{SOR} - Q_{v+1/2}(f) \right| \leq L \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right|, \quad (4.6)$$

and we end up with

THEOREM 4.3. (Piecewise linear flux modification)

The difference scheme (4.3), (4.5) is SOR-TVD under the CFL-like condition.<sup>(1)</sup>

$$Q_{v+1/2}(f) \leq \frac{2}{2 + L}. \quad (4.7)$$

Proof. The SOR requirement in (4.4b) is fulfilled because of (4.6), whose right-hand side is of order  $O(|\Delta v|)_{v+1/2}$ ; the TVD requirement on the left of (4.4a) follows from (4.1), and we are left with the inequality on the right, requiring

---

(1) To be precise, the constant  $L$  stands for the Lipschitz constant times the (assumed finite) maximal bound of the ratios  $|Q(\tilde{g}_{v+1/2}(\cdot); v_v, v_{v+1}, \lambda)| / |\Delta \tilde{g}_{v+1/2} / \Delta v_{v+1/2}|$ .

$$Q_{v+1/2}^{\text{SOR}} \leq 1. \quad (4.8)$$

Indeed, in view of (4.6) we have

$$Q_{v+1/2}^{\text{SOR}} \leq Q_{v+1/2}(f) + L \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right|; \quad (4.9a)$$

since  $\tilde{g}_{v+1}$  and  $\tilde{g}_v$  cannot have different signs by the choice of  $s_v$  in (4.3a), we also find

$$\left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq \frac{\text{Max}(|\tilde{g}_v|, |\tilde{g}_{v+1}|)}{|\Delta v_{v+1/2}|} \leq \frac{1}{2} [Q_{v+1/2}(f) - \lambda^2(a^2)_{v+1/2}]. \quad (4.9b)$$

The last two inequalities yield the crude bound

$$Q_{v+1/2}^{\text{SOR}} \leq Q_{v+1/2}(f) + \frac{L}{2} [Q_{v+1/2}(f) - \lambda^2(a^2)_{v+1/2}] \leq (1 + \frac{L}{2}) Q_{v+1/2}(f) \quad (4.10)$$

and augmented with the assumed CFL-like condition (4.7), the TVD constraint (4.8) is now satisfied.

Remark 4.4. With the primary examples of Engquist-Osher, Godunov and Roe-Murman schemes in (2.5b - d), the Lipschitz constant  $L$  in (4.6) may be taken to be  $L = 1$ , and the CFL limit in (4.7) is found to equal (compare, e.g., [24, Section 3])

$$Q_{v+1/2}(f) \leq 2/3. \quad (4.11)$$

(Taking into account the negative term  $-\lambda^2(a^2)_{v+1/2}$  which was ignored in the crude bound of (4.10), the CFL limit may be further improved to be



$$Q_{v+1/2}(f) \leq \frac{2 + \lambda^2(a^2)_{v+1/2}}{3} .) \quad (4.12)$$

Remark 4.5. In the case of Lax-Friedrichs scheme, where  $Q_{v+1/2}^{LF} = 1$  violates (4.11), we may take  $L$  to be zero, leaving us with the original 3-point CFL limitation

$$\lambda |a_{v+1/2}| \leq 1. \quad (4.13)$$

Remark 4.6. Consider the special choice  $(a^2)_{v+1/2} = (a_{v+1/2})^2$  in (4.3b), which is admissible in view of (2.4). The piecewise linear flux modification in Theorem 4.3, extends Harten's recipe [10], which was restricted to viscosity functions of the form  $Q_{v+1/2} = Q(\lambda |a_{v+1/2}|)$ ; (hence, the Engquist-Osher and Godunov schemes, for example, were excluded from the discussion [10]).

We now turn to a second more attractive choice of SOR viscosity, inspired by Theorem 4.2 (see also [10, p. 368]). Here, we use a piecewise constant viscosity modification of the form

$$Q_{v+1/2}^{SOR} = Q_{v+1/2}(f) + \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right|, \quad (4.14)$$

which leads to

THEOREM 4.7. (Piecewise constant viscosity modification)

The difference scheme (4.3), (4.14) is SOR-TVD under the CFL-like condition

$$Q_{v+1/2}(f) \leq 2/3. \quad (4.15)$$

Proof. The SOR requirement in (4.4b) is directly verified by the very definition of  $Q_{v+1/2}^{\text{SOR}}$  in (4.14); the TVD requirement on the left of (4.4a) follows from (4.1),

$$\left| \lambda a_{v+1/2} + \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq \lambda \left| \frac{f(v_{v+1}) - f(v_v)}{v_{v+1} - v_v} \right| + \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2}^{\text{SOR}}, \quad (4.16)$$

and we are left with the inequality on the right, requiring

$$Q_{v+1/2}^{\text{SOR}} \leq 1. \quad (4.17)$$

Indeed, the same estimate used in Theorem 4.3 before, see (4.9b), together with the definition of  $Q_{v+1/2}^{\text{SOR}}$ , (4.14) gives us the pessimistic bound:

$$Q_{v+1/2}^{\text{SOR}} \leq Q_{v+1/2}(f) + \frac{1}{2} [Q_{v+1/2}(f) - \lambda^2 (a^2)_{v+1/2}] \leq \frac{3}{2} Q_{v+1/2}(f) \quad (4.18)$$

and (4.17) follows in view of the assumed CFL-like condition (4.15).

Remark 4.8. As before, by improving the crude bound of (4.18), the CFL-limit may be improved to be that of (4.12).

We note that the piecewise constant viscosity modification involves linear corrections of the numerical flux  $f_v$  in (4.3d) and the viscosity coefficient in (4.14); both are appearing linearly in the original 3-point scheme we are starting with; see (2.2). Hence, the resulting modification recipe does not change the underlying 3-point TVD code, but, rather, adds to it an antidiffusive term which boils down to

$$-\frac{1}{2} (\tilde{g}_{v+1} - \tilde{g}_{v-1}) + \frac{1}{2} \left( \Delta \left( \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \Delta v_{v+1/2} \right) \right). \quad (4.19)$$

To make our point more precise, we state as our final result of this first part, the following

**COROLLARY 4.9.** (Piecewise constant viscosity modification - revisited).

Consider the 3-point scheme

$$v_v(t+k) = H(v_{v-1}, v_v, v_{v+1}; f, \lambda) \equiv v_v - \lambda(h_{v+1/2} - h_{v-1/2}) \quad (4.20a)$$

and assume the TVD-like constraint, compare (2.4)

$$\lambda |a_{v+1/2}| \leq Q_{v+1/2}(f) \leq \frac{2}{3}, \quad Q_{v+1/2} = \lambda \frac{f_v + f_{v+1} - 2h_{v+1/2}}{\Delta v_{v+1/2}}, \quad (4.20b)$$

holds. Then the modified scheme

$$v_v(t+k) = H(v_{v-1}, v_v, v_{v+1}; f, \lambda) - \left[ \left( \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right)^- \Delta v_{v+1/2} + \left( \frac{\Delta \tilde{g}_{v-1/2}}{\Delta v_{v-1/2}} \right)^+ \Delta v_{v-1/2} \right] \quad (4.21a)$$

is SOR-TVD. Here,  $\tilde{g}_v$  is the flux correction, given by

$$\tilde{g}_v = \frac{s_v}{2} \text{Min} ([Q_{v\pm 1/2}(f) - \lambda^2(a^2)_{v\pm 1/2}] |\Delta v_{v\pm 1/2}|) \quad (4.21b)$$

where

$$\lambda^2(a^2)_{v+1/2} = \lambda^2(a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2} \leq Q_{v+1/2}(f). \quad (4.21c)$$

Alternatively, we can rewrite (4.21a) in the conservative form

$$\begin{aligned} v_v(t+k) &= H(v_{v-1}, v_v, v_{v+1}; f, \lambda) - [\tilde{h}_{v+1/2} - \tilde{h}_{v-1/2}] \\ &= v_v(t) - \lambda \left[ (h_{v+1/2} + \frac{1}{\lambda} \tilde{h}_{v+1/2}) - (h_{v-1/2} + \frac{1}{\lambda} \tilde{h}_{v-1/2}) \right]. \end{aligned} \quad (4.22a)$$

Here  $\tilde{h}_{v+1/2}$  is the numerical flux correction given by

$$\tilde{h}_{v+1/2} = \frac{1}{2} [\tilde{g}_v + \tilde{g}_{v+1} - s_{v+1/2} |\tilde{g}_{v+1} - \tilde{g}_v|]. \quad (4.22b)$$

Remark 4.10. The Lax-Friedrichs scheme violates the CFL limitation (4.20b). Nonetheless, an "antidiffusive" modification follows in this case directly from the flux modification of Theorem 4.3, with numerical flux correction

$$\tilde{h}_{v+1/2}^{LF} = \frac{1}{2} (\tilde{g}_v + \tilde{g}_{v+1}),$$

Remark 4.11. The conservative form of scheme (4.22) is identical with the corrective type of Harten's artificial compression method [8,9], the difference lying, of course, in the exact details of the flux correction used. In particular, we have an operator splitting which could be easily implemented as a two-step predictor-corrector method [8, Section 6].

## PART II. ON A CELL ENTROPY INEQUALITY

### 5. Godunov's Scheme and its Cell Entropy Inequality

An entropy pair  $(U, F)$  associated with the conservative model (1.2), consists of convex entropy function,  $U(\cdot)$ , augmented with an entropy flux,  $F(\cdot)$ , such that  $U'f' = F'$ . The requirement of having an entropy inequality of the form

$$\frac{\partial U}{\partial t}(u(x, t)) + \frac{\partial F}{\partial x}(u(x, t)) \leq 0 \quad (5.1)$$

for all entropy pairs, singles out the unique physically relevant (weak) solution of (1.2), e.g., [13,14]. Accordingly, we are seeking conditions which guarantee that the difference scheme (1.1) will satisfy a cell entropy inequality of the form [12]

$$U(v_v(t+k)) \leq U(v_v(t)) - \lambda(F_{v+1/2} - F_{v-1/2}); \quad (5.2a)$$

here,  $F_{v+1/2} = F(v_{v-q+1}, \dots, v_{v+q}; f, \lambda)$  is a numerical entropy flux, assumed to be consistent with the differential one<sup>(1)</sup>

$$F(w, w, \dots, w; f, \lambda) = F(w). \quad (5.2b)$$

If this holds for wide enough class of entropy functions, the cell entropy inequalities (5.2) will guarantee the uniqueness of limits of all converging

---

(1) Both the differential and the numerical entropy fluxes are denoted by  $F$ ; their distinction is made by the number of arguments.

subsequences of the total-variation bounded solutions, and there follows the convergence of the whole computed sequence to the unique physically relevant (weak) solution , e.g., [3, 12, 18, 23, 26].

As an example, we begin with the all important cell entropy inequality associated with Godunov's scheme: by averaging two Riemann solvers, one obtains Godunov's scheme whose numerical flux is given by [17,18]

$$h_{v+1/2}^G = s_{v+1/2} \underset{(v-v_v) \cdot (v-v_{v+1}) < 0}{\text{Min}} [s_{v+1/2} f(v)], \quad s_{v+1/2} \equiv \text{sgn}(\Delta v_{v+1/2}). \quad (5.3)$$

Assuming that the minimum on the right-hand side is attained at a (not necessarily unique) value,  $v_{v+1/2}^G = v^G(v_v, v_{v+1}, f)$ ,

$$s_{v+1/2} \underset{(v-v_v) \cdot (v-v_{v+1}) < 0}{\text{Min}} [s_{v+1/2} f(v)] = f(v_{v+1/2}^G), \quad (5.4a)$$

then the Godunov scheme

$$v(t+k) = H^G(v_{v-1}, v_v, v_{v+1}; f, \lambda) \quad (5.4b)$$

boils down to

$$H^G(v_{v-1}, v_v, v_{v+1}; f, \lambda) = v_v - \lambda [f(v_{v+1/2}^G) - f(v_{v-1/2}^G)]. \quad (5.4c)$$

Integrating over a typical cell the differential entropy inequality (5.1), then Jensen's inequality yields the following well-known lemma, e.g., [3,18,26].

LEMMA 5.1. (Cell entropy inequality of Godunov scheme.)

Consider the Godunov scheme (5.4) satisfying the CFL condition.

$$\lambda \max_v |f'(v)| \leq 1. \quad (1) \quad (5.5a)$$

Then, the following cell entropy inequality holds

$$U[H^G(v_{v-1}, v_v, v_{v+1}; f, \lambda)] \leq U(v_v) - \lambda [F(v_{v+1/2}^G) - F(v_{v-1/2}^G)] \quad (5.5b)$$

for all entropy pairs (U,F). Equivalently, (5.5b) can be rewritten

$$U[H^G(v_{v-1}, v_v, v_{v+1}; f, \lambda)] \leq U(v_v) - \int_{v_{v-1/2}^G}^{v_{v+1/2}^G} U'(w) f'(w) dw. \quad (5.5c)$$

As in [26], the cell entropy inequality associated with the Godunov scheme, will be used as a building block for studying such an inequality in conjunction with other TVD schemes. To this end, we shall make use of a modified flux function introduced below.

---

(1) The maximum is taken over all values  $v$  varying between  $v_{v-1}$ ,  $v_v$ , and  $v_{v+1}$ .

## 6. The Modified Flux Function

While studying total-variation diminishing schemes, only the modified flux grid values,  $g_v$ , played a role, see the TVD requirement (1.8); subsequently, a simple piecewise linear modified flux correction,  $\tilde{g}_{v+1/2}(v)$ , was introduced in (4.5a), in connection with the (first) recipe of SOR-TVD schemes.

Regarding the question of a cell entropy inequality which we now consider, a more sophisticated construction of a modified flux function, is required: let  $g^{(v+1/2)}(v)$  denote its restriction to the  $I_{v+1/2}$ -interval,  $I_{v+1/2} = \{v | (v-v_v)(v-v_{v+1}) \leq 0\}$ . The piecewise linear function  $g^{(v+1/2)}(v)$  connects the modified flux grid values  $g_v$  and  $g_{v+1}$  on both ends of the interval, through a constant numerical flux value,  $h_{v+1/2}$ ; it depends on two, yet to be determined parameters  $(s^+, s^-)$ ,  $s^\pm > 0$ , in the following fashion:

Setting the intermediate values

$$v_{v+1/2}^+ \equiv v_{v+1/2}^+(s^+) = v_v + \frac{1}{s^+} (g_v - h_{v+1/2}) \quad (6.1a)$$

$$v_{v+1/2}^- \equiv v_{v+1/2}^-(s^-) = v_{v+1} - \frac{1}{s^-} (g_{v+1} - h_{v+1/2}), \quad (6.1b)$$

we define, depending on whether  $v_v \leq v_{v+1}$  or  $v_{v+1} \leq v_v$ ,

$$g^{(v+1/2)}(v; s^+, s^-) = \begin{cases} g_v - s^+(v-v_v), & v_v \leq v \leq v_{v+1/2}^+ \\ h_{v+1/2}, & v_{v+1/2}^+ \leq v \leq v_{v+1/2}^- \\ g_{v+1} + s^-(v-v_{v+1}), & v_{v+1/2}^- \leq v \leq v_{v+1} \end{cases}, \quad v_v \leq v_{v+1} \quad (6.1c)$$



$$g^{(v+1/2)}(v; s^+, s^-) = \begin{bmatrix} g_{v+1} - s^-(v-v_{v+1}) & , & v_{v+1} \leq v \leq v_{v+1/2}^- \\ h_{v+1/2} & , & v_{v+1/2}^- \leq v \leq v_{v+1/2}^+ \\ g_v - s^+(v-v_v) & , & v_{v+1/2}^+ \leq v \leq v_v \end{bmatrix}, \quad v_{v+1} \leq v_v. \quad (6.1d)$$

We have

LEMMA 6.1.

Assume the TVD condition, (1.8), holds

$$\lambda \left| \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right| < Q_{v+1} \leq 1. \quad (6.2a)$$

Then  $g^{(v+1/2)}(v; s^+, s^-)$  given in (6.1) and satisfying

$$Q_{v+1/2} \leq \lambda s^\pm \leq 1, \quad (6.2b)$$

is a well-defined piecewise linear function in the  $I_{v+1/2}$  interval.

Verification. In view of (6.2a) we have, using (1.10),

$$c_{v+1/2}^+ \geq 0, \quad c_{v+1/2}^- \geq 0, \quad Q_{v+1/2} = c_{v+1/2}^- + c_{v+1/2}^+ \leq 1,$$

and the definition of  $g_v$  in (1.5b) yields

$$v_{v+1/2}^+ = v_v + \frac{1}{s^+} (g_v - h_{v+1/2}) = v_v + \frac{1}{\lambda s^+} c_{v+1/2}^+ \Delta v_{v+1/2} \quad (6.3a)$$

$$v_{v+1/2}^- = v_{v+1}^- - \frac{1}{s^-} (g_{v+1} - h_{v+1/2}) = v_{v+1} - \frac{1}{\lambda s^-} C_{v+1/2}^- \Delta v_{v+1/2}. \quad (6.3b)$$

Consider first the case where  $v_v \leq v_{v+1}$ : then by (6.2b),  $v_{v+1/2}^+$  is located to the left of  $v_{v+1/2}^-$ ,

$$v_{v+1/2}^- - v_{v+1/2}^+ \geq \Delta v_{v+1/2} - \frac{1}{Q_{v+1/2}} (C_{v+1/2}^- + C_{v+1/2}^+) \Delta v_{v+1/2} = 0,$$

and to the right of  $v_v$ ,

$$v_{v+1/2}^+ - v_v = \frac{1}{\lambda s^+} C_{v+1/2}^+ \Delta v_{v+1/2} \geq 0.$$

Taking into account the positivity of the incremental coefficients  $C_{v+1/2}^\pm$ , the following graph of  $g^{(v+1/2)}(v; s^+, s^-)$  is obtained in this case; see Figure (6.4a).

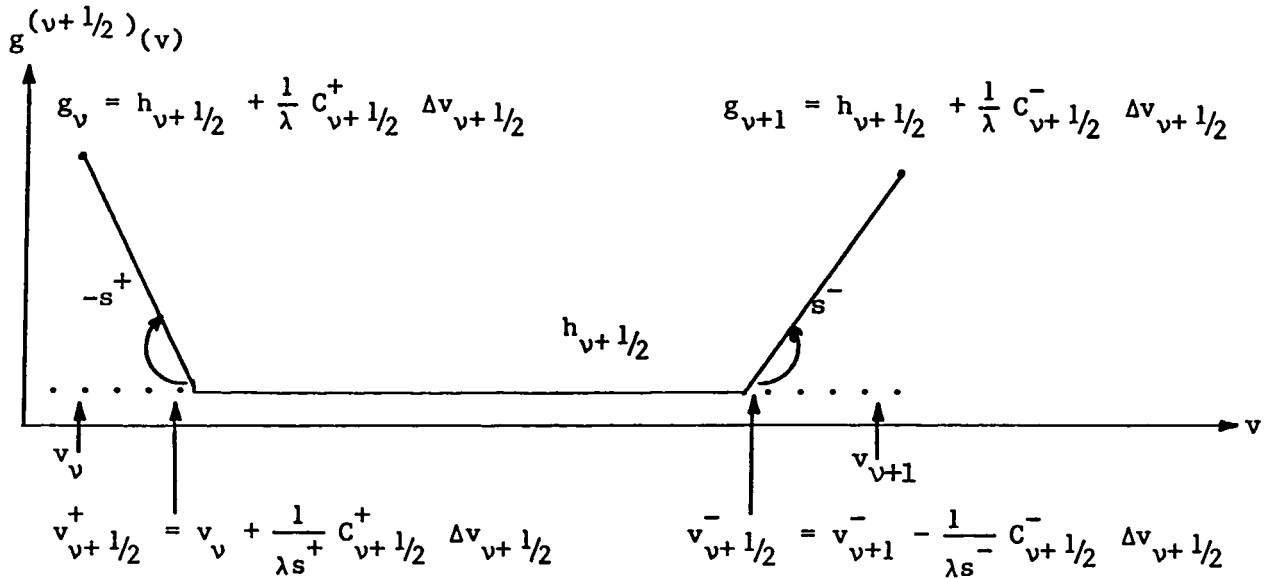


Figure 6.4a

A completely analogous situation occurs when  $v_{v+1} \leq v_v$ , where the graph of  $g^{(v+1/2)}(v)$  takes the form, see Figure (6.4b).

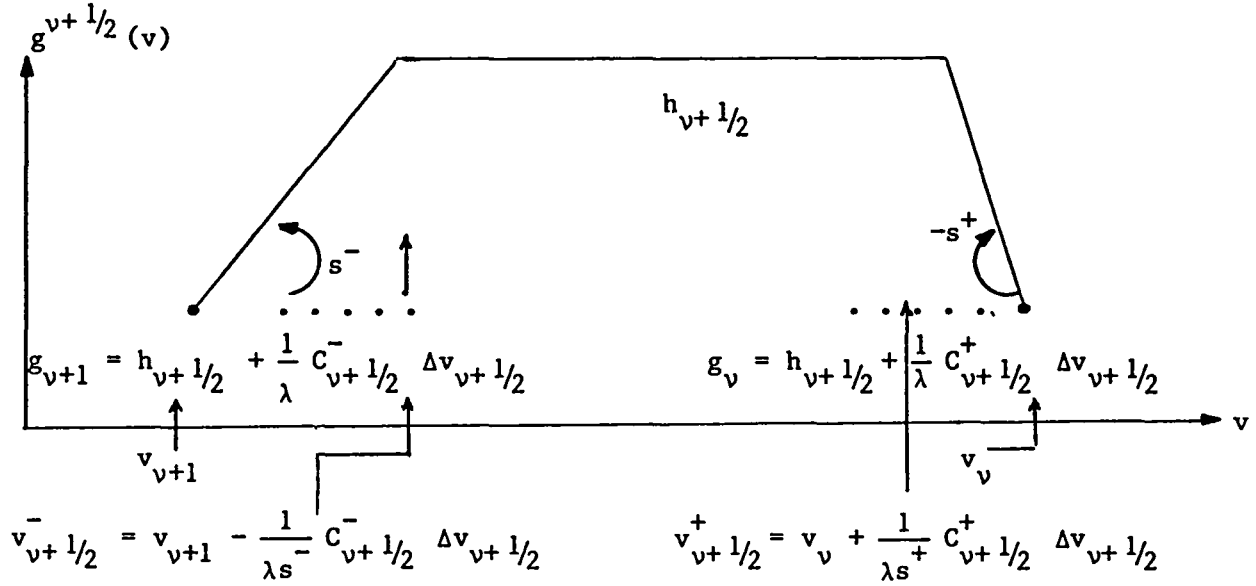


Figure 6.4b

Our next lemma is in the heart of the matter.

LEMMA 6.2.

Assume the TVD condition, compare (1.8),

$$\lambda \left| \frac{\Delta g_{v \pm 1/2}}{\Delta v_{v \pm 1/2}} \right| \leq Q_{v \pm 1/2} \leq \theta \leq 1$$

holds. Then we have

$$H^G[v_v, v_v, v_{v+1}; g^{(v+1/2)}(s^+ = s^- = \theta\lambda^{-1}), \theta^{-1}\lambda] = v_v - \theta^{-1}\lambda[h_{v+1/2} - g_v] \quad (6.6a)$$

$$H^G[v_{v-1}, v_v, v_v; g^{(v-1/2)}(s^+ = s^- = \theta\lambda^{-1}), \theta^{-1}\lambda] = v_v - \theta^{-1}\lambda[g_v - h_{v-1/2}]. \quad (6.6b)$$

Proof. We first note that according to Lemma 6.1,  $g^{(v\pm 1/2)}(s^+ = s^- = \theta\lambda^{-1})$  are well defined:  $(6.5)_\pm$  yields

$$\lambda \left| \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2} \leq \lambda s^\pm = \theta \leq 1.$$

By definition we have, see (5.4c),

$$\begin{aligned} H^G[v_v, v_v, v_{v+1}; g^{(v+1/2)}(s^+ = s^- = \theta\lambda^{-1}), \theta^{-1}\lambda] \\ = v_v - \theta^{-1}\lambda[g^{(v+1/2)}(v_{v+1/2}^G) - g^{(v-1/2)}(v_v)]; \end{aligned} \quad (6.7a)$$

a straightforward computation gives us, see (5.4a),

$$g^{(v+1/2)}(v_{v+1/2}^G) = s_{v+1/2} \min_{I_{v+1/2}} [s_{v+1/2} g^{(v+1/2)}(v)] = h_{v+1/2} \quad (6.7b)$$

while consistency implies

$$g^{(v+1/2)}(v_v) = g_v \quad (6.7c)$$

and (6.6a) follows. Similar arguments apply for (6.6b).

Remark 6.3. Godunov's numerical flux is determined by the differential one, computed at the intermediate value  $v_{v+1/2}^G$ ; see (5.4a). In our case, (6.7b), any intermediate value between  $v_{v+1/2}^+$  and  $v_{v+1/2}^-$  can be taken: for later purposes, we shall choose in particular

$$\begin{aligned} v_{v+1/2}^G(\theta) &= \frac{1}{2} [v_{v+1/2}^+(s^+ = \theta\lambda^{-1}) + v_{v+1/2}^-(s^- = \theta\lambda^{-1})] \\ &= \frac{v_v + v_{v+1}}{2} - \frac{\lambda}{2\theta} (g_{v+1} - g_v). \end{aligned} \quad (6.8)$$

Taking  $\theta = 1/2$  in Lemma 6.2 and averaging (6.6a) and (6.6b), we finally arrive at

THEOREM 6.4.

Consider the difference scheme (1.1) given in its viscous form (1.7a) and satisfying the TVD condition

$$\lambda \left| \frac{\Delta g_{v\pm 1/2}}{\Delta v_{v\pm 1/2}} \right| \leq Q_{v\pm 1/2} \leq 1/2. \quad (6.9)$$

We then have

$$\begin{aligned} v_v(t+k) &= \frac{1}{2} \left[ H^G[v_{v-1}, v_v, v_v; g^{(v-1/2)}(s^+ = s^- = \frac{1}{2\lambda}), 2\lambda] \right. \\ &\quad \left. + H^G[v_v, v_v, v_{v+1}; g^{(v+1/2)}(s^+ = s^- = \frac{1}{2\lambda}), 2\lambda] \right]. \end{aligned} \quad (6.10)$$

In other words, we have shown that any TVD satisfying difference scheme, (1.7a), (6.9), is given as an average of two Godunov solvers, thus refining a

similar (convex) decomposition introduced in [26, Theorem 5.1]; the current decomposition differs however, in its use of a modified flux function, rather than using the original differential one. In fact, away from the critical extremum values  $v_v$  where  $\Delta v_{v-1/2} \cdot \Delta v_{v+1/2} \leq 0$ , we can do even better: under appropriate TVD condition, any difference scheme coincides with one (rather than the average of two) Godunov solver. This is the content of our next

THEOREM 6.5.

Consider the difference scheme (1.1) given in its viscous form (1.7a) and satisfying the TVD condition

$$\lambda \left| \frac{\Delta g_{v \pm 1/2}}{\Delta v_{v \pm 1/2}} \right| \leq Q_{v \pm 1/2} \leq 1. \quad (6.11)$$

Then, in the neighborhood of noncritical values  $v_v$  such that

$\Delta v_{v-1/2} \cdot \Delta v_{v+1/2} \geq 0$ , we have

$$v_v(t+k) = H^G[v_{v-1}, v_v, v_{v+1}; g(s^+ = s^- = \frac{1}{\lambda}), \lambda]. \quad (6.12a)$$

Remark 6.6. The modified flux function  $g(v) = g(v; s^+, s^-)$  appearing in (6.12a), is uniquely defined in this case, composed of its restriction to the consecutive intervals  $I_{v \pm 1/2}$ , see Figures (6.13),

$$g(v; s^+ = s^- = \frac{1}{\lambda}) = \begin{cases} g^{(v-1/2)}(v; s^+ = s^- = \frac{1}{\lambda}), & v \in I_{v+1/2} \\ g^{(v+1/2)}(v, s^+ = s^- = \frac{1}{\lambda}), & v \in I_{v+1/2} \end{cases}. \quad (6.12b)$$

$$g(v; s^+ = s^- = \frac{1}{\lambda})$$

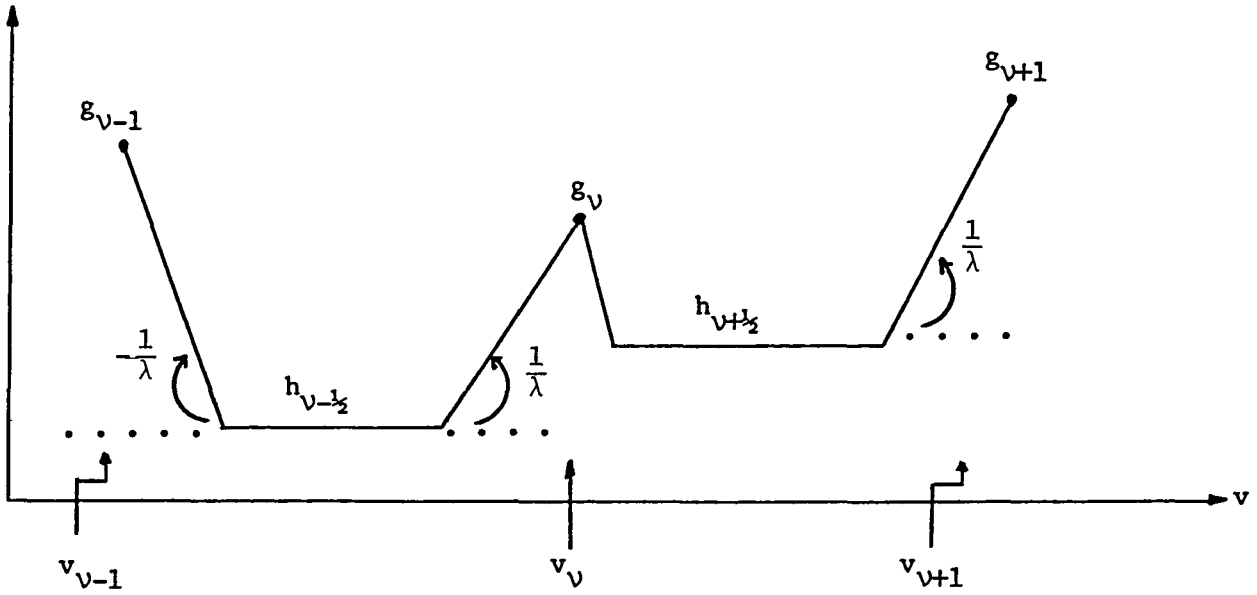


Figure (6.13a)

$$g(v; s^- = s^- = \frac{1}{\lambda})$$

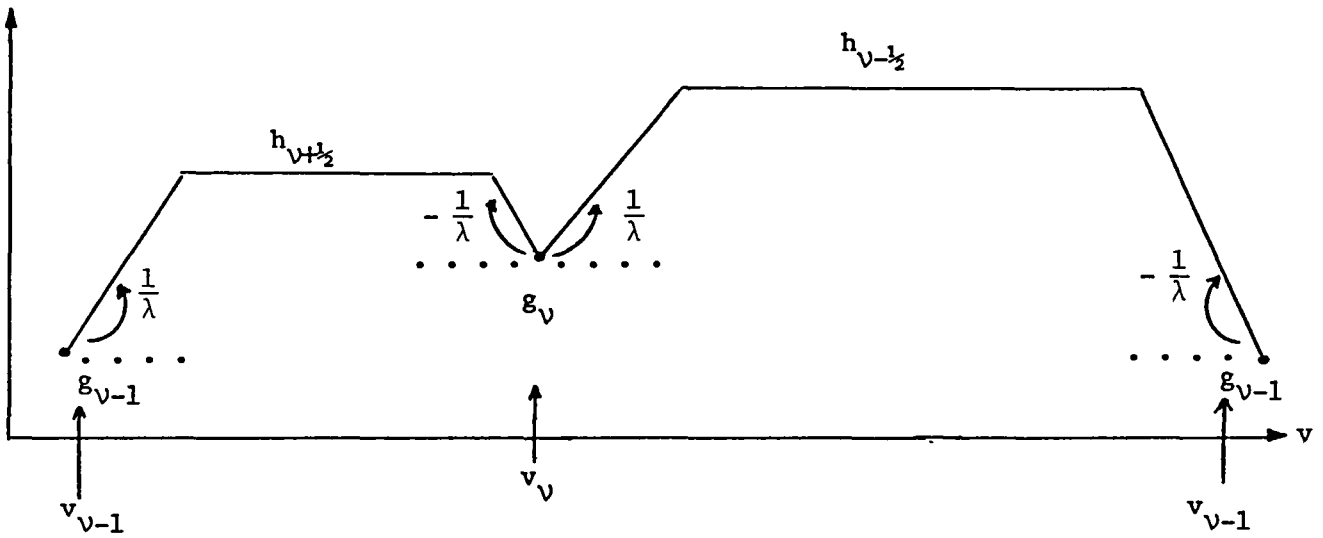


Figure (6.13b)

Verification of Theorem 6.5 is straightforward: by (6.11) we may employ Lemma 6.2 with  $\theta = 1$  in  $I_{v \pm 1/2}$ ; using (5.4c) and (6.7b) the result (6.12) follows.

We next use the cell entropy inequality associated with Godunov's scheme, see Lemma 5.1, to infer such an inequality for the difference schemes under consideration. We start with

LEMMA 6.7.

Consider the TVD scheme (1.7a), (6.9). Then for all entropy pairs (U, F) we have

$$\begin{aligned} U[v_v(t+k)] &\leq U[v_v(t)] - \lambda(F_{v+1/2} - F_{v-1/2}) \\ &\quad + \lambda \int_{v_{v-1/2}^G}^{v_v} U''(v) [g^{(v-1/2)}(v; s^+ = s^- = \frac{1}{2\lambda}) - f(v)] dv \\ &\quad + \lambda \int_v^{v_{v+1/2}^G} U''(v) [g^{(v+1/2)}(v; s^+ = s^- = \frac{1}{2\lambda}) - f(v)] dv, \end{aligned} \quad (6.14a)$$

with a consistent numerical entropy flux given by

$$F_{v+1/2} = F(v_{v+1/2}^G) - U'(v_{v+1/2}^G) \cdot [h_{v+1/2} - f(v_{v+1/2})]. \quad (6.14b)$$

Remark. Here, the intermediate Godunov value  $v_{v+1/2}^G$  may be taken as, see (6.8),

$$v_{v+1/2}^G = v_{v+1/2}^G(\theta = 1/2) = \frac{v_v + v_{v+1}}{2} - \frac{\lambda}{4} (g_{v+1} - g_v). \quad (6.14c)$$

Proof. The maximal slope of  $g^{(v+1/2)}(v; s^+ = s^- = \frac{1}{2\lambda})$  is  $1/2 \lambda$ ,



$$2\lambda \max_{v \in I_{v+1/2}} \left| g^{(v \pm 1/2)^-}(v; s^+ = s^- = \frac{1}{2\lambda}) \right| \leq 1.$$

Therefore, we may apply Lemma 5.1, obtaining, see (4.5c),

$$\begin{aligned} & U \left[ H^G[v_{v-1}, v_v, v_v; g^{(v-1/2)^-}(v; s^+ = s^- = \frac{1}{2\lambda}), 2\lambda] \right] \\ & \leq U(v_v) - 2\lambda \int_{v_v^G - 1/2}^{v_v} U^-(v) g^{(v-1/2)^-}(v; s^+ = s^- = \frac{1}{2\lambda}) dv \end{aligned} \quad (6.15a)$$

$$\begin{aligned} & U \left[ H^G[v_v, v_v, v_{v+1}; g^{(v+1/2)^-}(v; s^+ = s^- = \frac{1}{2\lambda}), 2\lambda] \right] \\ & \leq U(v_v) - 2\lambda \int_{v_v}^{v_v^G + 1/2} U^-(v) g^{(v+1/2)^-}(v; s^+ = s^- = \frac{1}{2\lambda}) dv. \end{aligned} \quad (6.15b)$$

Adding and subtracting  $2\lambda[F(v_v) - F(v_{v-1/2}^G)] \equiv 2\lambda \int_{v_{v-1/2}^G}^{v_v} U^- f^-$  to the right-hand side of (6.15a), we find after integration by parts

$$\begin{aligned} & U \left[ H^G[v_{v-1}, v_v, v_v; g^{(v-1/2)^-}(v; s^+ = s^- = \frac{1}{2\lambda}), 2\lambda] \right] \\ & \leq U(v_v) - 2\lambda \int_{v_v^G - 1/2}^{v_v} U^-(g^{(v-1/2)^-} - f)^- - 2\lambda[F(v_v) - F(v_{v+1/2}^G)] \\ & = U(v_v) - 2\lambda \cdot U^-(v) \cdot (g^{(v-1/2)^-}(v) - f(v)) \Big|_{v=v_v^G - 1/2}^{v=v_v} - 2\lambda[F(v_v) - F(v_{v-1/2}^G)] \\ & \quad + 2\lambda \int_{v_v^G - 1/2}^{v_v} U''[g^{(v-1/2)^-} - f]. \end{aligned} \quad (6.16a)$$

In a similar manner, (6.15b) yields

$$\begin{aligned}
 & U \left[ H^G[v_v, v_v, v_{v+1}; g^{(v+1/2)}(v; s^+ = s^- = \frac{1}{2\lambda}), 2\lambda] \right] \\
 & \leq U(v_v) - 2\lambda \cdot U'(v) \cdot (g^{(v+1/2)}(v) - f(v)) \Big|_{v=v_v}^{v=v_v^G+1/2} - 2\lambda [F(v_{v+1}^G) - F(v_v)] \\
 & \quad + 2\lambda \int_{v_v}^{v_{v+1}^G+1/2} U''[g^{(v+1/2)} - f]. \tag{6.16b}
 \end{aligned}$$

By convexity, the average of the two terms on the left of (6.16a) and (6.16b), dominates the entropy value  $U(\cdot)$ , computed at the average of the corresponding two Godunov solvers; invoking Theorem 6.4 we arrive at

$$\begin{aligned}
 & U[v_v(t+k)] \leq U[v_v(t)] \\
 & \quad - \lambda \Delta \left[ F(v_{v-1/2}^G) + U'(v_{v-1/2}^G) \cdot [g^{(v-1/2)}(v_{v-1/2}^G) - f(v_{v-1/2}^G)] \right] \\
 & \quad + \lambda \int_{v_{v-1/2}^G}^{v_v} U''(g^{(v-1/2)} - f) + \lambda \int_{v_v}^{v_{v+1}^G+1/2} U''(g^{(v+1/2)} - f), \tag{6.17}
 \end{aligned}$$

and the result (6.14) follows, noting that  $g^{(v\pm 1/2)}(v_{v\pm 1/2}^G) = h_{v\pm 1/2}$ ; see (6.7b).

In the same way as the last lemma followed from Theorem 6.4, Theorem 6.5 gives us

LEMMA 6.8.

Consider the TVD scheme (1.7a), (6.11) in the neighborhood of a non-critical value  $v_v$  where  $\Delta v_{v-1/2} \cdot \Delta v_{v+1/2} \geq 0$ . Then for all entropy pairs  $(U, F)$  we have

$$\begin{aligned} U[v_v(t+k)] \leq & U[v_v(t)] - \lambda(F_{v+1/2} - F_{v-1/2}) \\ & + \lambda \int_{v_{v-1/2}^G}^{v_v} U''(v) [g^{(v-1/2)}(v; s^+ = s^- = \frac{1}{\lambda}) - f(v)] dv \\ & + \lambda \int_{v_v}^{v_{v+1/2}^G} U''(v) [g^{(v+1/2)}(v; s^+ = s^- = \frac{1}{\lambda}) - f(v)] dv. \quad (6.18) \end{aligned}$$

Remark 6.9. The numerical entropy flux,  $F_{v+1/2}$ , in (6.18) is determined by Godunov's intermediate value  $v_{v+1/2}^G$ . It coincides with the one given in (6.14b) and (6.14c): indeed,  $v_{v+1/2}^G = v_{v+1/2}^G(\theta = 1/2)$  is located in between  $v_{v+1/2}^+(s^+ = \lambda^{-1})$  and  $v_{v+1/2}^-(s^- = \lambda^{-1})$ . The proof is omitted.

The only difference in Lemma 6.8 is in the slopes,  $s^\pm$ , involved in the last two integrals on the right of (6.18): at a non-critical value, they are twice then otherwise allowed by Lemma 6.7; see (6.14a). Making use of  $|s_v|$  to distinguish between these two cases, see (4.3a), the corresponding two estimates can be unified as follows.

THEOREM 6.10.

Consider the difference scheme (1.7a) satisfying the TVD condition

$$\lambda \left| \frac{\Delta g_{v \pm 1/2}}{\Delta v_{v \pm 1/2}} \right| \leq Q_{v \pm 1/2} \leq \frac{1 + |s_v|}{2}. \quad (6.19)$$

Then, for all entropy pairs (U,F) we have

$$\begin{aligned} U[v_v(t+k)] &\leq U[v_v(t)] - \lambda(F_{v+1/2} - F_{v-1/2}) \\ &+ \lambda \int_{v_{v-1/2}^G}^{v_v} U''(v) \left[ g^{(v-1/2)} \left( v; s^+ = s^- = \frac{1 + |s_v|}{2\lambda} \right) - f(v) \right] dv \\ &+ \lambda \int_{v_v}^{v_{v+1/2}^G} U''(v) \left[ g^{(v+1/2)} \left( v; s^+ = s^- = \frac{1 + |s_v|}{2\lambda} \right) - f(v) \right] dv, \end{aligned} \quad (6.20)$$

with a consistent numerical entropy flux

$$F_{v+1/2} = F(v_{v+1/2}^G) - U'(v_{v+1/2}^G) [h_{v+1/2} - f(v_{v+1/2}^G)]. \quad (6.20b)$$

For later purposes, we shall prefer a slightly different version of inequality (6.20), where the last two integrals on the right are shifted over the  $I_{v+1/2}$  interval. We state

THEOREM 6.11.

Consider the difference scheme (1.7a) satisfying the TVD condition

$$\lambda \left| \frac{\Delta g_{v \pm 1/2}}{\Delta v_{v \pm 1/2}} \right| \leq Q_{v \pm 1/2} \leq \frac{1 + |s_v|}{2}. \quad (6.21)$$

Then for all entropy pairs (U,F) we have

$$\begin{aligned} U[v_v(t+k)] &\leq U[v_v(t)] - \lambda (F_{v+1/2} - F_{v-1/2}) \\ &+ \lambda \int_{v_v}^{v_{v+1}} U''(v) \left[ g^{(v+1/2)} \left( v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv \end{aligned} \quad (6.22a)$$

with a consistent numerical entropy flux

$$\begin{aligned} F_{v+1/2} &= F(v_{v+1/2}^G) - U'(v_{v+1/2}^G) [h_{v+1/2} - f(v_{v+1/2}^G)] \\ &+ \lambda \int_{v_{v+1/2}^G}^{v_{v+1}} U''(v) \left[ g^{(v+1/2)} \left( v; s^+ = s^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv. \end{aligned} \quad (6.22b)$$

Remark. Theorem 6.11 finally leads us to the modified flux function<sup>(1)</sup> we were looking for; its restriction to the  $I_{v+1/2}$ -interval is given, according to (6.22a), by

$$\begin{aligned} g(v)|_{I_{v+1/2}} &= g^{(v+1/2)}(v) \\ &= g^{(v+1/2)}\left(v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda}\right). \end{aligned} \quad (6.23)$$

Proof. We first note the identity

$$\begin{aligned} \int_{v_v}^{v_{v+1/2}^G} U''(v) g^{(v+1/2)}(v; s_1^+, s_1^-) dv + \int_{v_{v+1/2}^G}^{v_{v+1}} U''(v) g^{(v+1/2)}(v; s_2^+, s_2^-) dv \\ = \int_{v_v}^{v_{v+1}} U''(v) g^{(v+1/2)}(v; s_1^+, s_2^-) dv. \end{aligned} \quad (6.24)$$

Indeed, of the two integrals on the left, the first depends only on  $s_1^+$ , the second only on  $s_2^-$ . Adding and subtracting

$$\lambda \int_{v_{v+1/2}^G}^{v_{v+1}} U''(v) \left[ g^{(v+1/2)}\left(v; s^+ = s^- = \frac{1 + |s_{v+1}|}{2\lambda}\right) - f(v) \right] dv$$

to the right-hand side of (6.20a) and using (6.24), the theorem then follows.

---

(1) To be precise,  $g(v)$  is double-valued in the critical neighborhood.

Remark 6.12. The two integrals appearing on the right of (6.22) are responsible for the unnaturally shifted estimates w.r.t. the  $v_v$ -grid value. If, in the previous proof, we instead add and subtract

$$\lambda \int_{v_{v-1}}^{v_v^G - 1/2} U''(v) \left[ g^{(v-1/2)} \left( v; s^+ = s^- = \frac{1 + |s_{v-1}|}{2} \right) - f(v) \right] dv$$

to the right-hand side of (6.20a) and use (6.24), we end up with a reversed shifted form

$$\begin{aligned} U[v_v(t+k)] &\leq U[v_v(t)] - \lambda (F_{v+1/2} - F_{v-1/2}) \\ &+ \lambda \int_{v_{v-1}}^{v_v} U''(v) \left[ g^{(v-1/2)} \left( v; s_{v-1}^+ = \frac{1 + |s_{v-1}|}{2\lambda}, s_v^- = \frac{1 + |s_v|}{2\lambda} \right) - f(v) \right] dv, \end{aligned} \quad (6.25a)$$

where

$$\begin{aligned} F_{v+1/2} &= F(v_{v+1/2}^G) - U'(v_{v+1/2}^G) [h_{v+1/2} - f(v_{v+1/2}^G)] \\ &- \lambda \int_{v_v}^{v_{v+1/2}^G} U''(v) \left[ g^{(v+1/2)} \left( v; s^+ = s^- = \frac{1 + |s_v|}{2\lambda} \right) - f(v) \right] dv. \end{aligned} \quad (6.25b)$$

Averaging of (6.22) and (6.25) results in the symmetric entropy estimate of this type,

$$\begin{aligned}
 U[v_v(t+k)] &\leq U[v_v(t)] - \lambda(F_{v+1/2} - F_{v-1/2}) \\
 &+ \lambda \int_{v_{v-1}}^{v_v} U''(v) \left[ g(v) \Big|_{I_{v-1/2}} - f(v) \right] dv + \lambda \int_{v_v}^{v_{v+1}} U''(v) \left[ g(v) \Big|_{I_{v+1/2}} - f(v) \right] dv.
 \end{aligned}
 \tag{6.26}$$

## 7. Necessity and Sufficiency of E-Fluxes

Theorem 6.11 shows that a desired cell entropy inequality, (5.2a), follows provided the last integral on the right of (6.22a) is negative

$$\int_{v_v}^{v_{v+1}} U''(v) \left[ g^{(v+1/2)} \left( v; s^+ = \frac{1 + |s_v|}{2\lambda}, s^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv \leq 0.
 \tag{7.1}$$

The requirement of negativity for all convex entropy functions  $U$ , in this case, is equivalent to the requirement that  $g(v) \Big|_{I_{v+1/2}}$  will lie below or above  $f(v)$ , depending on whether  $v_v \leq v_{v+1}$  or  $v_{v+1} \leq v_v$ ; taking into account the piecewise linear form of  $g(v) \Big|_{I_{v+1/2}}$  in (6.23) and (6.4), this requirement amounts to the following inequalities:

$$(i) \quad s_{v+1/2} \cdot [h_{v+1/2} - f(v)] \leq 0, \quad v \in I_{v+1/2} \tag{1}$$

---

(1) It is sufficient to consider  $v$  lying in between

$$v_{v+1/2}^+ (s^+ = \frac{1 + |s_v|}{2\lambda}) \quad \text{and} \quad v_{v+1/2}^- (s^- = \frac{1 + |s_{v+1}|}{2\lambda}).$$



$$(ii) \quad (a) \quad \lambda \quad \text{Max}_{|v-v_v| \leq \varepsilon_v^+} \left| \frac{f(v) - g_v}{v-v_v} \right| \leq \frac{1 + |s_v|}{2},$$

$$\varepsilon_v^+ = \frac{2}{1 + |s_v|} C_{v+1/2}^+ |\Delta v_{v+1/2}|$$

$$(b) \quad \lambda \quad \text{Max}_{|v-v_{v+1}| \leq \varepsilon_{v+1}^+} \left| \frac{f(v) - g_{v+1}}{v-v_{v+1}} \right| \leq \frac{1 + |s_{v+1}|}{2},$$

$$\varepsilon_{v+1}^- = \frac{2}{1 + |s_{v+1}|} C_{v+1/2}^- |\Delta v_{v+1/2}|.$$

Numerical fluxes satisfying condition (i) are called E fluxes after Osher [18]. Equivalently, the corresponding schemes are characterized as exactly those having more numerical viscosity than that of Godunov's scheme [26]. In terms of the modified viscosity coefficient  $\underline{Q}_{v+1/2}$  in (3.3), the first inequality, (i) reads, consult (2.5c),

$$Q_{v+1/2}^G \leq \underline{Q}_{v+1/2}. \quad (E)$$

Observe that two viscosity coefficients are comparable in the sense advocated in [26], provided that the same modified flux is used for the corresponding schemes; in the above E-condition, the original flux,  $f_v$ , was used in both cases, see Remark 3.1. Regarding the CFL condition in the second inequality (ii), it can be shifted to the more symmetric form in the spirit of Remark 6.12,

$$\lambda \max_{|v-v_v| \leq \epsilon_v^\pm} \left| \frac{f(v) - g_v}{v-v_v} \right| \leq \frac{1 + |s_v|}{2}, \quad (CFL)$$

$$\epsilon_v^\pm = \frac{2C_{v \pm 1/2}^\pm |\Delta v_{v \pm 1/2}|}{1 + |s_v|}.$$

With a similar shifting of the TVD condition (6.21), we arrive at

THEOREM 7.1 (Tadmor [26, Theorem 5.1]).

Consider an E-type TVD scheme (1.7a) such that

$$Q_{v+1/2} \leq \min\left(\frac{1 + |s_v|}{2}, \frac{1 + |s_{v+1}|}{2}\right) \quad (7.2)$$

and assume the following CFL condition holds

$$\lambda \max_{|v-v_v| \leq \epsilon_v^\pm} \left| \frac{f(v) - g_v}{v-v_v} \right| \leq \frac{1 + |s_v|}{2}, \quad (7.3)$$

$$\epsilon_v^\pm = \frac{2C_{v \pm 1/2}^\pm |\Delta v_{v \pm 1/2}|}{1 + |s_v|}.$$

Then, for all entropy pairs (U, F) the following cell entropy inequality is satisfied

$$U[v_v(t+k)] \leq U[v_v(t)] - \lambda(F_{v+1/2} - F_{v-1/2}). \quad (7.4)$$

Remark. We have preferred to center the TVD condition (7.2) and CFL condition (7.3) around the  $v_v$ -grid value. Therefore, no specific reference was given as to the consistent numerical entropy flux,  $F_{v+1/2}$  in (7.4): it may be chosen, throughout the computational grid, by (6.22b), (6.25b) or any convex

combination of the two.

Proof. The only part which requires clarification is the validity of the TVD constraint, see (6.21),

$$\lambda \left| \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2}. \quad (7.5)$$

Indeed, (3.1) and (2.5) yield

$$\lambda \left| \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq \lambda |a_{v+1/2}| + \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2}^G + \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right|; \quad (7.6)$$

since  $\tilde{g}_v / \Delta v_{v \pm 1/2}$  is positive by (3.7c), we find, as in (4.9b)

$$\left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq \frac{\text{Max}(|\tilde{g}_v|, |\tilde{g}_{v+1}|)}{|\Delta v_{v+1/2}|} \leq \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}}. \quad (7.7)$$

Augmented with the E condition, (7.6) and (7.7) yield (7.5)

$$\lambda \left| \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq Q_{v+1/2}^G + \frac{\tilde{g}_v + \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \leq Q_{v+1/2} + \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} \equiv Q_{v+1/2}.$$

Remark 7.2. It is the CFL-like condition (7.3) which throws some further light on the first-order accuracy limitation encountered with difference schemes satisfying all cell-wise entropy inequalities (5.2a). Indeed, in case  $\Delta v_{v+1/2} > 0$  (similarly  $\Delta v_{v-1/2} < 0$ ), then necessarily  $g_v = f_v + \frac{1}{\lambda} \tilde{g}_v$  will lie above (similarly below)  $f_v$  in contrast to the requirement induced by (7.1); specifically, as  $v$  approaches  $v_v$ , the CFL-like ratio on the left of (7.3)

$$\lambda \max_{|v-v_v| \leq \varepsilon_v^\pm} \left| \frac{f(v) - f(v_v)}{v-v_v} + \frac{1}{\lambda} \frac{\tilde{g}_v}{v-v_v} \right| \quad (7.8)$$

will diverge unless the correction  $\tilde{g}_v$  vanishes. In this case, the modified flux  $g_v$  is reduced to the differential one  $f_v$ , and we are back in the standard first-order accurate E-schemes, subject to the familiar

$$\lambda \max_{|v-v_v| \leq \varepsilon_v^\pm} \left| \frac{f(v) - f(v_v)}{v - v_v} \right| \leq \frac{1}{2}. \quad (7.9)$$

In particular, according to (7.9), the usual half CFL number is sufficient in this case, in agreement with [26]

$$\lambda \max_{|v-v_v| \leq |\Delta v_{v \pm 1/2}|} |f'(v)| \leq \frac{1}{2}. \quad (7.10)$$

Of course, (7.9) offers a more delicate alternative. Note that unlike the TVD constraint (1.8), some further information regarding the flux behavior in between the grid values, is always required for a cell entropy inequality to hold.

Next we show that the E condition is, in a sense described below, necessary for a cell entropy inequality, thus complementing the sufficiency of Theorem 7.1. To this end, we subtract  $v_v(t)$  from both sides of (1.1) and (1.3), divide by  $k = \Delta t$  and take the limit  $\Delta t \rightarrow 0$ , obtaining the semidiscrete approximation

$$\frac{dv_v}{dt} + \frac{1}{\Delta x} (h_{v+1/2} - h_{v-1/2}) = 0; \quad (7.11a)$$

here, the time derivative is understood in the distribution sense, and the

numerical flux is still allowed to depend on the mesh size,  $\Delta x$ ,

$$h_{v+1/2} = h(v_{v-p+1}, \dots, v_{v+p}; f, \lambda + 0, \Delta x). \quad (7.11b)$$

The corresponding cell entropy inequality will read

$$\frac{dU}{dt} [v_v(t)] + \frac{1}{\Delta x} (F_{v+1/2} - F_{v-1/2}) \leq 0; \quad (7.12a)$$

here,  $(U, F)$  is any consistent numerical entropy pair

$$F(w, w, \dots, w; f, \lambda + 0, \Delta x) = F(w) \equiv \int^w U^- f^-. \quad (7.12b)$$

In stating our next theorem, the terminology of separated data will be used: we shall say that the data are separated if all relevant grid values involved in the scheme's stencil, are separated w.r.t., say, the separating interval  $I_{j+1/2}$ ; that is, we either have

$$v_{j-k} \leq v_j < v_{j+1} \leq v_{j+k}, \quad k = 1, \dots, p. \quad (7.13a)$$

or

$$v_{j-k} \geq v_j > v_{j+1} \geq v_{j+k}, \quad k = 1, \dots, p. \quad (7.13b)$$

A monotone profile is, of course, the canonical example of separated data: any two consecutive grid values may serve as a separating interval in this case.

Equipped with this terminology, we may now turn to

THEOREM 7.3. (E-flux is necessary for cell entropy inequality).

Consider the semi-discrete approximation (7.11) and assume the cell entropy inequality (7.12) holds for all consistent numerical entropy pairs (U, F).

Then the numerical flux, h, is an E-flux in any separating interval.

Proof. Introducing the consistent numerical entropy flux

$$\underline{F}_{v+1/2} = F(v_{v+1}) - U^-(v_{v+1})[h_{v+1/2} - f(v_{v+1})], \quad (7.14a)$$

then the following equality [18, Section 3] holds

$$\frac{\partial U}{\partial t} [v_v(t)] + \frac{1}{\Delta x} [\underline{F}_{v+1/2} - \underline{F}_{v-1/2}] = \frac{1}{\Delta x} \int_{v_v}^{v_{v+1}} U''(w)[h_{v+1/2} - f(w)]dw. \quad (7.14b)$$

Subtracting (7.12a) from (7.14b) we find, after multiplication by  $\Delta x$

$$\int_{v_v}^{v_{v+1}} U''(w)[h_{v+1/2} - f(w)]dw + (H_{v+1/2} - H_{v-1/2}) \leq 0; \quad (7.15)_v$$

here,  $H_{v+1/2}$  stands for the difference  $F_{v+1/2} - \underline{F}_{v+1/2}$ , which is consistent with zero

$$H(w, w, \dots, w; f, \lambda \rightarrow 0, \Delta x) = F(w) - \underline{F}(w) = 0. \quad (7.16a)$$

Consider a  $(2p+1)$ -grid-values stencil separated by, say, the  $I_{j+1/2}$  interval; we shall extend it to all grid points, defining

$$v_{j+k} \equiv v_{j+p}, \quad v_{j-k} \equiv v_{j-p}, \quad k = p+1, p+2, \dots. \quad (7.16b)$$

Let  $v$  be any value in between  $v_j$  and  $v_{j+1}$  and apply  $(7.15)_v$  with Kruzkov's choice for entropy function:  $U(w) = |w - v|$ : for all but the  $v = j$  index, the first term in the left of  $(7.15)_v$  vanishes and we end up with

$$H_{j+1/2} \geq H_{j+3/2} \geq \dots \geq H_\infty = 0, \quad v = j+1, j+2, \dots, \quad (7.17a)$$

$$H_{j-1/2} \leq H_{j-3/2} \leq \dots \leq H_{-\infty} = 0, \quad v = j-1, j-2, \dots. \quad (7.17b)$$

Here, the equalities  $H_{\pm\infty} = 0$  on the right of (7.17) follow from the consistency (7.16a) and the definition of our extension in (7.16b). Complemented with  $(7.15)_j$ , which in this case reads

$$\begin{aligned} 0 &\geq \int_{v_j}^{v_{j+1}} \delta(w - v) [h_{j+1/2} - f(w)] dw + (H_{j+1/2} - H_{j-1/2}) \\ &= s_{j+1/2} [h_{j+1/2} - f(v)] + (H_{j+1/2} - H_{j-1/2}); \end{aligned} \quad (7.18)$$

we find, for all  $v$  in  $I_{j+1/2}$ ,

$$s_{j+1/2} [h_{j+1/2} - f(v)] \leq H_{j-1/2} - H_{j+1/2} \leq 0.$$

That is,  $h$ , is an E-flux over the  $I_{j+1/2}$  interval as asserted.

Next we have to confront the limitation of E schemes being at most first-order accurate [18]: indeed, according to the E-condition, their modified numerical viscosity is bounded away from that of the Lax-Wendroff scheme, hence (3.4) fails. Moreover, being an E-flux in separating intervals alone, still meets the limitation of first-order accuracy. The proof follows along the lines of [18].

## 8. Second-Order Resolution and a Cell Entropy Inequality

In light of our discussion in the last section, we are seeking cell entropy inequalities for SOR-TVD schemes, only for special entropy pairs: specifically, we shall consider the pair  $U(v) = \frac{v^2}{2}$ ,  $F(v) = \int^v w f'(w) dw$ . In the genuinely nonlinear case where  $f$  is, say, strictly convex, all (convex) entropy inequalities (and therefore convergence to the unique solution) then follows in the limit; see [14,5].

We start with 3-point E-type upwind schemes: upwinding simply means that we difference in the "streamwise" direction. In particular, considering the characteristic direction, we require either  $C_{v+1/2}^+$  or  $C_{v+1/2}^-$  to vanish depending on whether  $f'(v)$  is positive or negative throughout the  $I_{v+1/2}$  interval; in either of these two nonsonic cases, the sum of the incremental coefficients  $C_{v+1/2}^\pm$  equals their difference in absolute value, i.e.,

$$Q_{v+1/2}(f) = \lambda \left| \frac{\Delta f_{v+1/2}}{\Delta v_{v+1/2}} \right|, \quad f'(v)|_{I_{v+1/2}} \neq 0. \quad (8.1a)$$

This is, of course, nothing but a restatement of the well-known fact that away from sonic values, all upwind schemes coincide with Roe-Murman's (2.5d).

The SOR-TVD schemes to be considered are then constructed according to the piecewise constant viscosity modification of Theorem 4.7, where the viscosity coefficient used is given by

$$Q_{v+1/2} \equiv Q_{v+1/2}^{SOR} = Q_{v+1/2}(f) + \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right|. \quad (8.1b)$$

Here  $\tilde{g}$  is the flux correction, see (4.21b),



$$\tilde{g}_v = \frac{s_v}{2} \text{Min}([Q_{v\pm 1/2}(f) - \lambda^2(a^2)_{v\pm 1/2}] |\Delta v_{v\pm 1/2}|) \quad (8.1c)$$

with second-order resolution condition requiring

$$\lambda^2(a^2)_{v+1/2} \equiv \lambda^2(a_{v+1/2})^2 + O(|\Delta v|)_{v+1/2} \leq Q_{v+1/2}(f), \quad (8.1d)$$

with a first-order perturbation  $O(|\Delta v|)_{v+1/2}$ , which is yet to be precisely determined. Unless otherwise stated, the following CFL-like condition will be assumed

$$Q_{v+1/2}(f) \leq 1/3. \quad (8.1e)$$

In particular, the above use of the piecewise viscosity modification from Theorem 4.7 is justified, since (4.15) holds. Furthermore, the essential estimate (4.9b)

$$\left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq \frac{1}{2} [Q_{v+1/2}(f) - \lambda^2(a^2)_{v+1/2}],$$

yields the TVD requirement (6.21)

$$\begin{aligned} \lambda \left| \frac{\Delta g_{v\pm 1/2}}{\Delta v_{v\pm 1/2}} \right| &\leq \lambda \left| a_{v\pm 1/2} \right| + \left| \frac{\Delta \tilde{g}_{v\pm 1/2}}{\Delta v_{v\pm 1/2}} \right| \\ &\leq Q_{v\pm 1/2}(f) + \left| \frac{\Delta \tilde{g}_{v\pm 1/2}}{\Delta v_{v\pm 1/2}} \right| \equiv Q_{v\pm 1/2}^{\text{SOR}} \leq \frac{3}{2} Q_{v\pm 1/2}(f) \leq \frac{1}{2}, \end{aligned}$$

so that Theorem 6.11 applies.<sup>(1)</sup> According to this result, the desired cell entropy inequality holds for the above quadratic entropy provided

$$\int_{v_v}^{v_{v+1}} \left[ g^{(v+1/2)} \left( v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv \quad (8.2)$$

is nonpositive.

We first study the critical case, stating

Lemma 8.1. Consider the critical case  $|s_v| = |s_{v+1}| = 0$  (that is, both  $v_v$  and  $v_{v+1}$  are critical grid values), and assume the following CFL condition holds

$$\lambda \max_{|v-v_v| \leq |\Delta v_{v+1/2}|} |f'(v)| \leq 1/2. \quad (8.3)$$

We then have

$$\int_{v_v}^{v_{v+1}} \left[ g^{(v+1/2)} \left( v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv \leq 0. \quad (8.4)$$

Proof. We appeal to the conservative form of the modified scheme, quoted in

(4.22): by assumption  $s_v$  and therefore  $\tilde{g}_v$  vanish, and hence

$\tilde{h}_{v+1/2} = \frac{1}{2}[\tilde{g}_{v+1} - s_{v+1/2}|\tilde{g}_{v+1}|]$ ; since  $s_{v+1/2}$  agrees with the sign of  $\tilde{g}_{v+1}$ , we conclude that the numerical flux correction  $\tilde{h}_{v+1/2}$  vanishes in this

---

(1) In fact, twice the CFL limit (8.1e) can be used away from critical neighborhoods where  $|s_v| + |s_{v+1}| \neq 0$ .

case. Thus, we are left with the original E-flux we have started with. By Theorem 7.1 therefore, all entropy weighted integrals in (7.1) and in particular (8.4) are non-positive, provided the CFL condition (7.3) holds: the latter follows in view of (8.3) noting that, the modified flux corrections  $\tilde{g}_v$  and  $\tilde{g}_{v+1}$  vanish - consult Remark 7.2.

Next we turn to the noncritical case  $|s_v| + |s_{v+1}| \geq 1$ : calculating the area below the piecewise linear "double-wing" form of

$$g^{(v+1/2)}\left(v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda}\right),$$

see figure (6.4), we obtain

$$\begin{aligned} & \int_{v_v}^{v_{v+1}} \left[ g^{(v+1/2)}\left(v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda}\right) - f(v) \right] dv \\ &= \Delta v_{v+1/2} h_{v+1/2} + \frac{1}{2\lambda} \left[ \frac{2}{1 + |s_v|} [C_{v+1/2}^+]^2 \right. \\ & \quad \left. + \frac{2}{1 + |s_{v+1}|} (C_{v+1/2}^-)^2 \right] (\Delta v_{v+1/2})^2 - \int_{v_v}^{v_{v+1}} f(v) dv; \quad (8.5a) \end{aligned}$$

inserting the value of  $h_{v+1/2}$  from (1.7b) and (3.1),

$$h_{v+1/2} = \frac{f_v + f_{v+1}}{2} + \frac{1}{2\lambda} (\tilde{g}_v + \tilde{g}_{v+1}) - \frac{1}{2\lambda} Q_{v+1/2} \Delta v_{v+1/2},$$

we find

$$\begin{aligned}
 & \int_{v_v}^{v_{v+1/2}} \left[ g^{(v+1/2)} \left( v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv \\
 &= \Delta v_{v+1/2} \frac{f_v + f_{v+1}}{2} - \int_{v_v}^{v_{v+1}} f(v) dv \\
 &+ \frac{(\Delta v_{v+1/2})^2}{2\lambda} \left[ \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} - Q_{v+1/2} + \frac{2}{1 + |s_v|} (C_{v+1/2}^+)^2 + \frac{2}{1 + |s_{v+1}|} (C_{v+1/2}^-)^2 \right] \\
 & \hspace{15em} (8.5b)
 \end{aligned}$$

In the next proposition, we estimate from above the last brackets on the right of (8.5b). The somewhat technical proof is postponed to the end of this section.

Proposition 8.2. The following estimate holds

$$\begin{aligned}
 & \frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} - Q_{v+1/2} + \frac{2}{1 + |s_v|} (C_{v+1/2}^+)^2 + \frac{2}{1 + |s_{v+1}|} (C_{v+1/2}^-)^2 \\
 & \leq \frac{|s_v| + |s_{v+1}|}{2} [Q_{v+1/2}^2(f) - \lambda^2(a^2)_{v+1/2}]. \hspace{10em} (8.6)
 \end{aligned}$$

In view of (8.5b) and Proposition 8.2, we can now complement Lemma 8.1 with

Lemma 8.3. Consider the noncritical case  $|s_v| + |s_{v+1}| \geq 1$  (that is,  
either  $v_v$  or  $v_{v+1}$  is a noncritical value). We then have

$$\begin{aligned} & \int_{v_v}^{v_{v+1}} \left[ g^{(v+1/2)} \left( v; s_v^+ = \frac{1 + |s_v|}{2\lambda}, s_{v+1}^- = \frac{1 + |s_{v+1}|}{2\lambda} \right) - f(v) \right] dv \\ & \leq \frac{(\Delta v_{v+1/2})^2}{2\lambda} \left[ \frac{2\lambda}{(\Delta v_{v+1/2})^2} ET_{v+1/2} + \frac{|s_v| + |s_{v+1}|}{2} (Q_{v+1/2}^2(f) - \lambda^2 (a^2)_{v+1/2}) \right]. \end{aligned} \quad (8.7a)$$

Here ,  $ET_{v+1/2}$  stands for the error in the trapezoidal rule, applied over the  
 $I_{v+1/2}$  -interval

$$ET_{v+1/2} = \Delta v_{v+1/2} \left[ \frac{f(v_v) + f(v_{v+1})}{2} \right] - \int_{v_v}^{v_{v+1}} f(v) dv. \quad (8.7b)$$

To ensure the integral in the left is nonpositive, hence to guarantee a cell entropy inequality in this noncritical case, we would therefore like to choose  $\lambda^2 (a^2)_{v+1/2}$  which dominates the first two terms inside the right brackets of (8.7a).

We claim that such a choice admissible by the second-order resolution  
condition (8.1d)--is indeed available in all nonsonic intervals: according to the upwinding property (8.1a),

$$Q_{v+1/2}^2(f) = \lambda^2 |a_{v+1/2}|^2 \quad (8.8a)$$

and by a standard error estimate [4, Section 2 ]

$$ET_{v+1/2} = \frac{(\Delta v_{v+1/2})^3}{12} f''(\zeta), \quad \zeta \in I_{v+1/2}. \quad (8.8b)$$

It is here that we use the first-order flexibility in defining  $\lambda^2(a^2)_{v+1/2}$  previously noticed in Theorem 4.2. We choose

$$\lambda^2(a^2)_{v+1/2} = Q_{v+1/2}^2(f) + \lambda \frac{|\Delta v_{v+1/2}|}{3} \max_{I_{v+1/2}} |f''(v)|.$$

Unifying the critical and noncritical cases, we arrive at

$$\lambda^2(a^2)_{v+1/2} = Q_{v+1/2}^2(f) + \frac{\lambda}{3} |\Delta v_{v+1/2}| \max_{I_{v+1/2}} |f''(v)|. \quad (8.9)$$

In order to comply with the inequality on the right of (8.1d), we set in (8.1c)

$$\tilde{g}_v = \frac{s_v}{2} \min [Q_{v\pm 1/2}^2(f) - \lambda^2(a^2)_{v\pm 1/2}]^+ \cdot |\Delta v_{v\pm 1/2}|. \quad (8.10a)$$

In terms of the numerical flux modification

$$\tilde{h}_{v+1/2} = \frac{1}{2} [\tilde{g}_v + \tilde{g}_{v+1} - s_{v+1/2} |\tilde{g}_{v+1} - \tilde{g}_v|], \quad (8.10b)$$

the resulting scheme then reads:

$$v_v(t+k) = v_v(t) - \lambda(f_{v+1} - f_{v-1}) + \frac{1}{2} [\Delta(Q_{v+1/2}(f) \Delta v_{v+1/2}(t))] - (\tilde{h}_{v+1/2} - \tilde{h}_{v-1/2}) \quad (8.10c)$$

and we summarize what we have shown in

**THEOREM 8.4.** Let  $Q_{v+1/2}(f)$  be an upwind E-type viscosity coefficient, and assume the following CFL condition holds

$$\lambda \cdot \text{Max}|f'(v)| \leq 1/3. \quad (8.11)$$

Then the difference scheme (8.9), (8.10) satisfies

- (i) Total variation diminishing.
- (ii) Second Order resolution.
- (iii) The cell entropy inequality:<sup>(1)</sup>

$$\frac{1}{2} v_v^2(t+k) \leq \frac{1}{2} v_v^2(t) - \lambda(F_{v+1/2} - F_{v-1/2}).$$

**Remark 8.5.** As argued in Theorem 4.2, the scheme has formal second order accuracy in all intervals such that  $\tilde{g}_v$  and  $\tilde{g}_{v+1}$  do not vanish. According to (8.10a) therefore, all nonsonic noncritical intervals such that

---

(1) The numerical entropy flux is given in (6.22b).

$$Q_{v+1/2}(f) - \lambda^2(a^2)_{v+1/2} = \lambda |a_{v+1/2}| - \lambda^2 |a_{v+1/2}|^2 \\ - \lambda \frac{|\Delta v_{v+1/2}|}{3} \max_{I_{v+1/2}} |f''(v)| \geq 0,$$

are included; in view of the CFL condition (8.11), the first difference on the right exceeds  $2\lambda |a_{v+1/2}|/3$ , and the last inequality is valid provided that

$$\frac{1}{2} |\Delta v_{v+1/2}| \cdot \max_{I_{v+1/2}} |f''(v)| \leq |a_{v+1/2}|. \quad (8.12)$$

In other words, scheme (8.9), (8.10) is second order accurate at all noncritical neighborhoods satisfying (8.12). Observe that increasing  $\lambda^2(a^2)_{v+1/2}$  results in a similar increase of the modified viscosity coefficient,  $Q_{v+1/2}$  in (3.4a).

Remark 8.6. The additional viscosity added is directly related to the amount of  $\frac{u^2}{2}$  entropy loss across shock discontinuities which is precisely  $|ET_{v+1/2}|$ . This should indicate the possible generalizations to other entropy pairs; moreover, it seems to imply that second order resolution is the maximum possible for cell entropy satisfying schemes, the entropy loss being cubic [14].

As a special case, let us consider now the genuinely nonlinear model where  $f$  is, say, strictly convex. We distinguish the possibilities

(i) The shock case.  $\Delta v_{v+1/2} \leq 0$ . Since the trapezoidal error in (8.8b) as negative in this case, the admissible choice  $\lambda^2(a^2)_{v+1/2} = \lambda^2 |a_{v+1/2}|^2 = Q_{v+1/2}^2(f)$ , will make the integral (8.7a) non-positive, as required.



(ii) The rarefaction case.  $\Delta v_{v+1/2} > 0$ . As expected, this is the more delicate case because the sign of the trapezoidal error is now reversed to be positive, which will be compensated for by additional dissipation as before. To summarize, we choose

$$\lambda^2 (a^2)_{v+1/2} = Q_{v+1/2}^2(f) + \frac{\lambda}{3} (\Delta v_{v+1/2})^+ \max_{I_{v+1/2}} [f''(v)] \quad (8.13)$$

in which case we have

THEOREM 8.5. (The convergence of SOR-TVD schemes). Assume  $f$  is strictly convex,  $Q_{v+1/2}(f)$  is an upwind E-type viscosity and the CFL condition (8.11) holds. Then the difference scheme (8.13) (8.10) satisfies

(i) Total variation diminishing;

(ii) Second Order resolution where:

$$\frac{1}{2} (\Delta v_{v+1/2})^+ \max_{I_{v+1/2}} [f''(v)] \leq |a_{v+1/2}|;$$

(iii) A consistent quadratic cell entropy inequality;

and, as a consequence of (i) and (iii),

(iv) Convergence to the unique physically relevant solution.

Remark 8.6. The above mentioned quadratic cell entropy inequality is due to the related inequality of Theorem 6.11. If (the nonshifted form of) Theorem 6.10 is used instead, the same cell entropy inequality follows, this time with added dissipation compensating the rectangular rule error rather than the

trapezoidal one. Details are omitted. Recently, Osher [19] has considered SOR-TVD semidiscrete approximations to the genuinely nonlinear model (1.2). Convergence was then guaranteed with more dissipation added at shocks, rather than at the expected rarefactions. Indeed, the estimates used in [19] correspond to that quoted in our Theorem 6.10, and the dissipation unnaturally yet necessarily was added at shocks can be therefore attributed to the reversed sign of the rectangular rule error. We also note that in the cases studied in [19], [20] and in the last two theorems, formal second order accuracy had to be given up at strong jumps, which were measured w.r.t. the amount of convexity,  $\max_{I_{v+1/2}} [f''(v)]$ .

Remark 8.7. One can obtain similar results by starting with 3-point TVD schemes which are not necessarily of upwind type. For example, choosing the modified Lax-Friedrichs schemes where  $Q_{v+1/2}(f) \equiv 1/2$ , see [26], one can repeat the above arguments and conclude similar convergence results. Details are omitted.

We close this section with

Proof of proposition 8.2. The flux correction  $\tilde{g}_v$  in (8.1c) satisfies

$$\frac{\tilde{g}_v}{\Delta v_{v \pm 1/2}} \leq \frac{|s_v|}{2} [Q_{v \pm 1/2}(f) - \lambda^2 (a^2)_{v \pm 1/2}]; \quad (8.14)_v$$

summing  $(8.14)_v$  and  $(8.14)_{v+1}$  we find

$$\frac{\tilde{g}_v + \tilde{g}_{v+1}}{\Delta v_{v+1/2}} \leq \frac{|s_v| + |s_{v+1}|}{2} [Q_{v+1/2}(f) - \lambda^2 (a^2)_{v+1/2}]. \quad (8.15)$$

We recall the definition of the viscosity coefficient in (8.16)

$$-Q_{v+1/2}^{\text{SOR}} = -Q_{v+1/2}(f) - \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right|. \quad (8.16)$$

Finally, we examine the incremental coefficients  $C_{v+1/2}^{\pm}$  in (1.9): utilizing (3.1) we have

$$C_{v+1/2}^{\pm} = \frac{1}{2} [Q_{v+1/2}(f) + \left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \mp \lambda a_{v+1/2} \right| \mp \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}}]. \quad (8.17a)$$

Consequently, the following equality holds

$$\frac{1 - |s_v|}{1 + |s_v|} (C_{v+1/2}^+)^2 = \frac{1 - |s_v|}{4} [Q_{v+1/2}(f) - \lambda a_{v+1/2}]^2. \quad (8.17b)$$

Indeed, if  $|s_v| = 1$ , then both sides vanish; if otherwise  $s_v$  and therefore  $\tilde{g}_v$  vanish, then

$$\left| \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right| - \left( \frac{\Delta \tilde{g}_{v+1/2}}{\Delta v_{v+1/2}} \right) = \left| \frac{\tilde{g}_{v+1}}{\Delta v_{v+1/2}} \right| - \frac{\tilde{g}_{v+1}}{\Delta v_{v+1/2}},$$

vanishes as well, so that in view of (8.17a) equality holds in (8.17b).

Similar arguments yield

$$\frac{1 - |s_{v+1}|}{1 + |s_{v+1}|} (C_{v+1/2}^-)^2 = \frac{1 - |s_{v+1}|}{4} [Q_{v+1/2}(f) - \lambda a_{v+1/2}]^2. \quad (8.17c)$$

In view of the TVD constraint (1.8) we also have

$$\begin{aligned}
 (C_{v+1/2}^+)^2 + (C_{v+1/2}^-)^2 &= \frac{1}{2} \left[ (Q_{v+1/2}^{\text{SOR}})^2 + \lambda^2 \left( \frac{\Delta g_{v+1/2}}{\Delta v_{v+1/2}} \right)^2 \right] \\
 &\leq \left[ Q_{v+1/2}(f) + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \right]^2. \quad (8.17d)
 \end{aligned}$$

Adding (8.17b), (8.17c), and (8.17d) we arrive at

$$\begin{aligned}
 \frac{2}{1 + |s_v|} (C_{v+1/2}^+)^2 + \frac{2}{1 + |s_{v+1}|} (C_{v+1/2}^-)^2 \\
 \leq \left[ Q_{v+1/2}(f) + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \right]^2 + (2 - |s_v| - |s_{v+1}|) Q_{v+1/2}^2(f). \quad (8.18)
 \end{aligned}$$

The inequalities (8.15), (8.16) together with (8.18) amount to the desired inequality (8.6): if  $|s_v| = |s_{v+1}| = 1$ , the derived upper bound reads

$$Q_{v+1/2}^2(f) - \lambda^2 (a^2)_{v+1/2} + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \left[ -1 + 2Q_{v+1/2}(f) + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \right],$$

where in view of (8.14)<sub>v</sub> together with the CFL condition (8.1e), the last brackets on the right are indeed nonpositive

$$-1 + 2Q_{v+1/2}(f) + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \leq 1 + \frac{5}{2} Q_{v+1/2}(f) < 0;$$

if  $|s_v| + |s_{v+1}| = 1$ , we find

$$\begin{aligned}
 -\frac{1}{2} Q_{v+1/2}(f) - \frac{1}{2} \lambda^2 (a^2)_{v+1/2} + 2Q_{v+1/2}^2(f) \\
 + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \left[ -1 + 2Q_{v+1/2}(f) + \left| \frac{\tilde{\Delta g}_{v+1/2}}{\Delta v_{v+1/2}} \right| \right].
 \end{aligned}$$

As before, the last brackets on the right are nonpositive, while the sum of the preceeding three terms does not exceed the asserted value

$$\begin{aligned}
 & -\frac{1}{2} Q_{v+1/2}(f) + \frac{3}{2} Q_{v+1/2}^2(f) + \frac{1}{2} [Q_{v+1/2}^2(f) - \lambda^2(a^2)_{v+1/2}] \\
 & \leq \frac{1}{2} [Q_{v+1/2}^2(f) - \lambda^2(a^2)_{v+1/2}];
 \end{aligned}$$

finally, although not specifically referred to,  $\tilde{g}_v = \tilde{g}_{v+1} = 0$  in the critical case,  $|s_v| = |s_{v+1}| = 0$ , yielding

$$-Q_{v+1/2}(f) + 3Q_{v+1/2}^2(f) \leq 0.$$

### References

- [1] J. P. Boris and D. L. Book, "Flux corrected transport: I. SHASTA, a fluid transport algorithm that works," J. Comp. Phys., Vol. 11, 1973, pp. 38-69.
  
- [2] S. R. Chakravarthy and S. Osher, "Computing with high resolution upwind schemes for hyperbolic equations," to appear in Proceedings of AMS/SIAM, 1983 Summer Seminar, La Jolla, CA, (B. Engquist, S. Osher, and R. C. J. Somerville, eds.).
  
- [3] M. Crandall and A. Majda, "Monotone difference approximations for scalar conservative laws," Math. Comp., Vol. 34, 1980, pp. 1-21.
  
- [4] P. J. Davis and P. Rabinowitz, Methods of Numerical Integration, Academic Press, New York, 1975.
  
- [5] R. J. DiPerna, "Convergence of approximate solutions to conservation laws," Arch. Rat. Mech. and Analysis, Vol. 82, 1983, pp. 27-70.
  
- [6] B. Engquist and S. Osher, "One-sided difference approximations for nonlinear conservation laws," Math. Comp., Vol. 36, 1981, pp. 321-351.
  
- [7] S.K. Godunov, "A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics," Mat. Sb. Vol. 47, 1959, pp. 271-290.

- [8] A. Harten, "The method of artificial compression: I. Shocks and contact discontinuities," AEC Research & Development Rep. C00-3077-50, Courant Institute, New York University, June 1984.
- [9] A. Harten, "The artificial compression method for computation of shocks and contact discontinuities: III. Self-adjusting hybrid schemes," Math. Comp., Vol. 32, 1983, pp. 363-389.
- [10] A. Harten, "High resolution schemes for hyperbolic conservation laws," J. Comp. Phys., Vol. 49, 1983, pp. 357-393.
- [11] A. Harten, "On a class of high resolution total variation stable finite difference schemes," SINUM, Vol. 21, 1984, pp. 1-23.
- [12] A. Harten, J. M. Hyman and P. D. Lax, "On finite difference approximations and entropy conditions for shocks," Comm. Pure Appl. Math., Vol. 29, 1976, pp. 297-322.
- [13] S. N. Kruzhkov, "First order quasilinear equations in several independent variables," Math. USSR Sb., Vol. 10, 1970, pp. 217-243.
- [14] P. D. Lax, Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves, SIAM Regional Conference Lecturers in Applied Mathematics, No. 11, 1972.

- [15] P. D. Lax, and B. Wendroff, "Systems of conservation laws," Comm. Pure Appl. Math., Vol. 13, 1960, pp. 217-237.
- [16] E. M. Murman, "Analysis of embedded shock waves calculated by relaxation methods," AIAA J., Vol. 12, 1974, pp. 626-633.
- [17] A. Oleinik, "Discontinuous solutions of nonlinear differential equations," AMS Transl. Ser. 2, Vol. 26, 1963, pp. 95-172.
- [18] S. Osher, "Riemann solvers, the entropy condition and difference approximating," SINUM, Vol. 21, 1984, pp. 217-235.
- [19] S. Osher, "Convergence of generalized MUSCL schemes," NASA Langley Contractor Report 172306, 1984, SINUM, to appear.
- [20] S. Osher and S. R. Chakravarthy, "High resolution schemes and the entropy condition," SINUM, Vol. 21, 1984, pp. 955-984.
- [21] S. Osher and S. R. Chakravarthy, "Very high order accurate TVD schemes," ICASE Report 84-44, 1984, submitted to SINUM.
- [22] P. L. Roe, "Approximate Riemann solvers, parameter vectors and difference schemes," J. Comp. Phys., Vol. 43, No. 198, pp. 357-372.
- [23] R. Sanders, "On convergence of monotone finite difference schemes with variable spatial differencing," Math. Comp., Vol. 40, 1983, pp. 91-106.



- [24] P. K. Sweby, "High resolution schemes using flux limiters for hyperbolic conservation laws," SINUM, Vol. 21, 1984, pp. 995-1011.
  
- [25] E. Tadmor, "The large time behavior of the scalar, genuinely nonlinear Lax-Friedrichs scheme," Math. Comp. Vol. 43, 1984, pp. 353-368.
  
- [26] E. Tadmor, "Numerical viscosity and the entropy condition for conservative difference schemes," Math. Comp., Vol. 43, 1984, pp. 369-382.
  
- [27] B. Van Leer, "Towards the ultimate conservative difference scheme V.A. second-order sequel to Godunov's method," J. Comp. Phys., Vol. 32 1979, pp. 101-136.
  
- [28] B. Van Leer, "Towards the ultimate conservative difference scheme, II. Monotonicity and conservation combined in a second-order scheme," J. Comp. Phys, Vol. 14, 1974, pp. 361-376.

1 Report No NASA CR-172614 ICASE Report No. 85-28		2 Government Accession No		3. Recipient's Catalog No	
4 Title and Subtitle  ON THE CONVERGENCE OF DIFFERENCE APPROXIMATIONS TO SCALAR CONSERVATION LAWS				5 Report Date May 1985	
				6 Performing Organization Code	
7 Author(s) Stanley Osher and Eitan Tadmor				8 Performing Organization Report No 85-28	
9 Performing Organization Name and Address Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23665				10 Work Unit No	
				11 Contract or Grant No NAS1-17070	
12 Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546				13 Type of Report and Period Covered Contractor Report	
				14 Sponsoring Agency Code 505-31-83-01	
15 Supplementary Notes Langley Technical Monitor: J. C. South, Jr. Final Report				Submitted to Math. Comp.	
16 Abstract  We present a unified treatment of explicit in time, two level, second order resolution, total variation diminishing, approximations to scalar conservation laws. The schemes are assumed only to have conservation form and incremental form. We introduce a modified flux and a viscosity coefficient and obtain results in terms of the latter. The existence of a cell entropy inequality is discussed and such an equality for all entropies is shown to imply that the scheme is an E scheme on monotone (actually more general) data, hence at most only first order accurate in general. Convergence for TVD-SOR schemes approximating convex or concave conservation laws is shown by enforcing a single discrete entropy inequality.					
17 Key Words (Suggested by Author(s))  conservation law, finite-difference scheme, variation diminishing, entropy inequality			18 Distribution Statement  64 - Numerical Analysis  Unclassified - Unlimited		
19 Security Classif (of this report) Unclassified	20 Security Classif (of this page) Unclassified	21 No of Pages 74	22 Price A04		

**End of Document**