NASA Technical Memorandum 101088

A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods

H. C. Yee

(NASA-TM-101088) A CLASS OF HIGH RESOLUTION EXPLICIT AND IMPLICIT SHOCK-CAPTURING METHODS (NASA. Ames Research Center) 222 p CSCL 12A
N89-25652
Unclas
G3/64 0217654

February 1989

NASA
National Aeronautics and
Space Administration

NASA Technical Memorandum 101088

A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods

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

February 1989

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

VON KARMAN INSTITUTE FOR FLUID DYNAMICS
Lecture Series 1989-04

Computational Fluid Dynamics
March 6-10, 1989
Rhode-St-Genèse, Belgium

A CLASS OF HIGH-RESOLUTION EXPLICIT AND IMPLICIT SHOCK-CAPTURING METHODS

H.C. YEE
NASA Ames Research Center
Computational Fluid Dynamics Branch
Moffett Field, California, 94035 USA

A CLASS OF HIGH-RESOLUTION EXPLICIT AND IMPLICIT SHOCK-CAPTURING METHODS[1]

H.C. Yee[2]
NASA Ames Research Center, Moffett Field, California, 94035 USA

Abstract

The development of shock-capturing finite difference methods for hyperbolic conservation laws has been a rapidly growing area for the last decade. Many of the fundamental concepts, state-of-the-art developments and applications to fluid dynamics problems can only be found in meeting proceedings, scientific journals and internal reports. This paper attempts to give a unified and generalized formulation of a class of high-resolution, explicit and implicit shock-capturing methods, and to illustrate their versatility in various steady and unsteady complex shock waves, perfect gases, equilibrium real gases and nonequilibrium flow computations. These numerical methods are formulated for the purpose of ease and efficient implementation into a practical computer code. The various constructions of high-resolution shock-capturing methods fall nicely into the present framework and a computer code can be implemented with the various methods as separate modules.

Included is a systematic overview of the basic design principle of the various related numerical methods. Special emphasis will be on the construction of the basic nonlinear, spatially second- and third-order schemes for nonlinear scalar hyperbolic conservation laws and the methods of extending these nonlinear scalar schemes to nonlinear systems via the approximate Riemann solvers and flux-vector splitting approaches. Generalization of these methods to efficiently include equilibrium real gases and large systems of nonequilibrium flows will be discussed. Some issues concerning the applicability of these methods that were designed for homogeneous hyperbolic conservation laws to problems containing stiff source terms and shock waves are also included. The performance of some of these schemes is illustrated by numerical examples for one-, two- and three-dimensional gas-dynamics problems. The use of the Lax-Friedrichs numerical flux to obtain high-resolution shock-capturing schemes is generalized. This method can be extended to nonlinear systems of equations without the use of Riemann solvers or flux-vector splitting approaches and thus provides a large savings for multidimensional, equilibrium real gases and nonequilibrium flow computations.

---

[1] Shorter version of this paper was published as a NASA TM-89464, May, 1987 entitled "Upwind and Symmetric Shock-Capturing Schemes".

[2] Research Scientist, Computational Fluid Dynamics Branch.

TABLE OF CONTENTS

I. INTRODUCTION

This paper is an outgrowth of a manuscript entitled "Upwind and Symmetric Shock-Capturing Schemes" in the proceedings of the Seminar on Computational Aerodynamics, University of California, Davis, spring, 1986, and in a NASA TM-89464, May, 1987. This expanded version includes approximately 35% new material, more updated information and an enlarged reference list. Over all, there is more than 50% increase in content. Many sections have been rewritten to make the paper more self-contained. The new material includes five new subsections in section III, three new subsections in section IV, one new subsection in section V, a complete new section VI, two new subsections in section VII, and approximately 90% increase in the reference list. Not all of the contents that appeared in NASA TM-89464 was review material. Consequently, material inside TM-89464 that was new and not published anywhere else will be referenced from time to time as TM-89464 or reference [1].

The objective of this paper is to give a unified and generalized formulation of a class of explicit and implicit, high-resolution, shock-capturing methods for hyperbolic conservation laws, and to illustrate their versatility in various steady and unsteady complex shock waves, perfect gases, equilibrium real gases and nonequilibrium flow computations. These methods are formulated for the purpose of ease and efficient implementation into new or existing computer codes. All of these methods can be implemented into the same computer code, sharing many of the common operations. The basic design principle and background information will be systematically reviewed.

Some issues concerning the applicability of these methods that were designed for homogeneous hyperbolic conservation laws to problems containing stiff source terms and shock waves are also included. The use of the Lax-Friedrichs numerical flux to obtain high-resolution shock-capturing schemes is generalized. This method can be extended to nonlinear systems of equations without the use of Riemann solvers or flux-vector splitting approaches and thus provides a large savings for multidimensional, equilibrium real gases and nonequilibrium flow computations.

The performance of some of these schemes is illustrated by numerical examples for one-, two- and three-dimensional gas-dynamics problems. The discussion on the development of conservative shock-capturing methods for hyperbolic conservation laws is the author's personal interpretation. The illustrations of the types of schemes and the areas of applications reflect the author's experiences and preferences for certain schemes. No attempt has been made to present a unified comparison except for the one-dimensional shock-tube problem. The number of theoretical and application papers on this topics is at least three time the number that the author has been able to mention, but those included should serve as a good starting point.

Much of the mathematical theory is omitted. However, when appropriate, sufficient references will be provided. It is assumed that the reader is familiar with the fundamentals of numerical analysis, fluid dynamics and computational fluid dynamics (CFD). Excellent lecture notes and review papers covering these subject areas can be found in [2-8]. Before getting into the main discussion, some pertinent terminology will be covered. The specifics to be addressed, and some aspects, limitation, and assumptions of shock-capturing schemes for hyperbolic conservation laws will be summarized. It is recommended that the reader refer to this introductory chapter for the limitations and assumptions stated when reading later sections.

Terminology: In this paper, the terms *explicit* and *implicit* schemes refer to *time* discretization, whereas the terms *symmetric* and *upwind* schemes refer to *spatial* discretization. The order of accuracy for time-accurate calculations refers to both the time and spatial discretization. On the other hand, the order of accuracy for steady-state calculations most often refers to the spatial discretization only. Spatial order of accuracy for high-resolution shock-capturing schemes usually refers to the order of accuracy away from discontinuities. Throughout the paper, the term "high-resolution shock-capturing schemes" is used in this context.

Upwind-differencing schemes attempt to discretize hyperbolic partial differential equations (PDEs) by using differences biased in the direction determined by the sign of the characteristic speed. Symmetric or centered schemes, on the other hand, try to discretize hyperbolic PDEs without any knowledge of the sign of the characteristic speed.

Shock-capturing schemes tend to treat shocks as a continuum, as opposed to shock-fitting, where almost always *a priori* knowledge of the shock location is needed.

Specifics to be Addressed: In this paper, only conservative finite-difference methods for hyperbolic conservation laws (i.e., for inviscid compressible flows) containing shock waves are addressed. The formulations are Eulerian, and the main emphasis is state-of-the-art of a class of high-resolution shock-capturing methods of the last decade. Theoretical discussions are for initial-value problems (IVPs). That is, the numerical schemes to be discussed are interior schemes (schemes for the interior points of the computational domain). When numerical boundary conditions (boundary schemes) are used, additional conditions have to be satisfied for the combined interior and boundary schemes (i.e. for the initial-boundary-value problems (IBVPs)) in order to maintain the same high-resolution and stability properties as the IVPs [9-11]. Also, the design principle of the schemes to be discussed are for homogeneous PDEs. For problems containing source terms, and especially stiff source terms and shock waves, the desired property of the scheme in question might be lost or no longer valid. See sections III for a discussion and a numerical study.

For compressible Navier-Stokes calculations, the physical problems considered here are assumed to be inviscid-dominated in the sense that moderate or strong viscous shock waves are present in the flow field such that high-resolution shock-capturing techniques are required. Thus, the numerical procedure described here for Navier-Stokes calculations is that a second-order, central-difference approximation is used for the diffusion terms and a high-resolution shock-capturing method is used for the inviscid part of the Navier-Stokes equations (Euler equations).

Hierarchy of Conservative Schemes for Hyperbolic Conservation Laws: The Hierarchy of conservative schemes for hyperbolic conservation laws can best be illustrated by figure 1.1. Let S_T be the set of all existing conservative schemes of any order for hyperbolic conservation laws which is the entire region shown in figure 1.1. We can break S_T into two parts, S_{UP} and S_C, where S_{UP} is the set of all existing upwind schemes of any order. Furthermore, let S_{ENO} be the set of all essentially nonoscillatory (ENO) [12-14] schemes of any order, let S_{TVD} be the set of all total variation diminishing (TVD) schemes [15-23] of any order and let S_M be the set of all monotone schemes [24-25]. Then $S_M \subset S_{TVD} \subset S_{ENO} \subset S_T$. In other words, the set of monotone schemes is the smallest set and is a subset of the set of TVD schemes. The set of all TVD schemes in turn is a subset of the ENO schemes. Definition and properties of these schemes will be described inside the text. This paper covers only a small subset of the shock-capturing schemes, namely, the TVD schemes.

Classical vs. Modern Shock-Capturing Methods: From an historical point of view, shock-capturing methods can be classified into two general categories: namely, classical and modern shock-capturing methods. In the case of classical shock-capturing methods, numerical dissipation terms are either linear such that the same amount is applied at all grid points or contain of adjustable parameters [26-28]. Classical shock-capturing methods only exhibit accurate results for smooth or weak shock solutions and are not robust enough for strong shock wave calculations. For strong shock waves, classical shock-capturing methods result in oscillatory solutions across the discontinuities and/or nonlinear instabilities [29-32]. For modern shock-capturing methods, however, the numerical dissipations are nonlinear such that the amount varies from one grid point to another and usually consists of automatic feedback mechanisms with hardly any adjustable parameters. These schemes [12-25] are stable for nonlinear scalar hyperbolic conservation laws. They exhibit high-resolution numerical results even for problems containing strong shock waves. Numerical dissipation terms similar to those of Jameson et al. [33] seem to fall in between the classical and modern shock-capturing methods.

Applicability of Scalar Schemes to Systems of Hyperbolic Conservation Laws: Basic modern shock-capturing methods for hyperbolic conservation laws are developed for linear and/or nonlinear scalar hyperbolic conservation laws. Extension of scalar methods to nonlinear systems is accomplished by assuming certain physical models or by local linearization. The mathematical foundation relies mainly on the scalar case. There is no identical theory for nonlinear systems or the multidimensional counterpart [34]. These schemes are formally extended to one- or higher-dimensional nonlinear systems of hyperbolic conservation laws via the so-called Riemann solvers (to be defined in section IV) and are evaluated by numerical experiments. Based on these facts, a major part of the discussion will be on modern shock-capturing schemes for the nonlinear scalar case and the methods of extending these nonlinear scalar schemes to nonlinear systems via the different types of Riemann solvers. However, numerical examples for one- and higher-dimensional nonlinear systems in gas dynamics will be stressed. Also, it is understood that even though the mathematical properties of modern shock-capturing methods are only valid for nonlinear scalar hyperbolic conservation laws or constant coefficient hyperbolic systems, in a loose and convenient way, the same properties are used and referred to when nonlinear systems or multidimensional problems are involved.

Guidelines and Usage of Modern Shock-Capturing Methods: Besides the fact that basic modern shock-capturing methods are developed for nonlinear scalar hyperbolic conservation laws, there are guidelines and limitations of these schemes that have to be considered. Some of the guidelines and limitations when appropriate are discussed at the beginning of each section. Specifically, a more detailed discussion on this subject for practical CFD applications appears in section 6.3.

In general, the best uses of modern shock-capturing schemes and in particular TVD-type of schemes are for problems containing moderate to strong shock waves and contact discontinuities. It is recommended not to apply any shock-capturing method (modern or otherwise) for ill-conditioned hyperbolic conservation laws, e.g., nearly incompressible hyperbolic conservation laws or very low Mach number flows [35,36]. For ill-conditioned hyperbolic conservation laws, if possible, make a proper transformation and transform the equation to a well-conditioned one before applying the numerical scheme [35,36]. When applying TVD-type of schemes for problems containing no discontinuities, the over all resolution compared with a higher-order classical shock-capturing scheme might be degraded due to the reduction of the scheme to first-order near extrema (points of extrema).

Outline of Paper: The outline of this paper is as follows. First, the basic properties of hyperbolic conservation laws and several schemes for linear hyperbolic equations will be reviewed. Then the various aspects of shock-capturing schemes for nonlinear scalar hyperbolic conservation laws will be discussed. These include monotone and first-order upwind schemes, deficiency of classical shock-capturing schemes, and methods of extending first-order TVD schemes to higher-order. Some issues concerning the applicability of these methods that were designed for homogeneous hyperbolic conservation laws to problems containing stiff source terms and shock waves also are included. It is important to investigate finite difference methods for nonhomogeneous hyperbolic PDEs since nonequilibrium fluid flows contain coupled stiff source terms and this is an area of vital interest for hypersonic flows and combustion. The use of the Lax-Friedrichs numerical flux to obtain high-resolution shock-capturing schemes is generalized. This method can be extended to a nonlinear system of equations without the use of Riemann solvers or flux-vector splitting approaches and thus provides a large savings for multidimensional, equilibrium real gas and nonequilibrium flow computations. A discussion on the importance of "asymptotic analysis of finite difference methods by the nonlinear dynamic approach" is briefly summarized. This subject is especially important for the analysis of numerical methods for nonhomogeneous hyperbolic PDEs with nonlinear source terms. Lastly formal extensions of nonlinear scalar TVD schemes to one- and higher-dimensional nonlinear systems of hyperbolic conservation laws will be reviewed. A method of extension, which is widely used and practical in terms of complex fluid dynamics problems, will be stressed. Generalization of these schemes to include steady and unsteady hypersonic equilibrium real gases and nonequilibrium flows will be described. Time-accurate as well as steady-state calculations for one-, two- and three-dimensional practical applications will be illustrated when appropriate.

II. PRELIMINARIES

The main difficulty in solving nonlinear hyperbolic PDEs is the need to allow for discontinuous solutions even when the initial data are smooth. For constant-coefficient hyperbolic PDEs, well-known stable, finite-difference methods are available in standard textbooks; see for example: Richtmyer and Morton [2], Mitchell [37], and Garabedian [38]. The theory is more complex for nonlinear hyperbolic PDEs. In order to motivate the ideas and set up the basic notations for nonlinear hyperbolic PDEs, some of the schemes originally designed for constant-coefficient cases are reviewed in this section.

2.1. An Upwind Scheme for Linear Hyperbolic PDEs

Consider the constant-coefficient scalar hyperbolic PDE

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, \tag{2.1}$$

where a is a real constant. Let u_j^n be the numerical approximation of the 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. Let $\lambda = \frac{\Delta t}{\Delta x}$. According to the characteristic theory of hyperbolic PDEs (direction of wave propagation), a simple first-order accurate explicit upwind scheme for an IVP of (2.1) can be written as

$$u_j^{n+1} = u_j^n - \lambda a \begin{cases} u_{j+1}^n - u_j^n & a < 0 \\ u_j^n - u_{j-1}^n & a > 0 \end{cases}. \tag{2.2}$$

Introducing the notation

$$a^+ = \frac{1}{2}(a + |a|); \quad a^- = \frac{1}{2}(a - |a|), \tag{2.3}$$

the scheme for positive or negative a can be rewritten as one equation

$$u_j^{n+1} = u_j^n - \lambda[a^+(u_j^n - u_{j-1}^n) + a^-(u_{j+1}^n - u_j^n)]. \tag{2.4}$$

Using the relationship between $a^{\pm}$, a and $|a|$, the scheme again can be written as

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n) + \frac{\lambda}{2}|a|(u_{j+1}^n - 2u_j^n + u_{j-1}^n). \tag{2.5}$$

Most often one recognizes the first form (2.4) as an upwind scheme but the second form (2.5) is less obvious. In this paper, the second form is preferred because when one extends the scheme to nonlinear equations and systems of equations, the second form is more compact to discuss and more efficient in terms of operations count [39-41]. Introducing a new notation "$h_{j+\frac{1}{2}}$", consider a scheme of the form

$$u_j^{n+1} = u_j^n - \lambda[h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n], \tag{2.6}$$

where $h_{j+\frac{1}{2}}$ is sometimes referred to as a "**numerical flux function**". Note that the notation $h_{j+\frac{1}{2}}$ will be used throughout this paper. For the previous example (2.5)

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[a(u_{j+1} + u_j) - |a|(u_{j+1} - u_j)\right]. \tag{2.7}$$

Higher-order upwind schemes can be obtained by increasing the stencils of the first-order scheme in the appropriate upwind directions. See references [2,37] for details.

Consider a hyperbolic system

$$\frac{\partial U}{\partial t} + A\frac{\partial U}{\partial x} = 0, \tag{2.8}$$

where U is a vector with m elements and A is an $m \times m$ constant matrix with real eigenvalues. Let $W = R^{-1}U$ and $R^{-1}AR = \Lambda$. One can transform the above system to a diagonal form

$$\frac{\partial W}{\partial t} + \Lambda\frac{\partial W}{\partial x} = 0, \qquad \Lambda = \mathrm{diag}(a^l),\ l = 1,\ldots,m. \tag{2.9}$$

Here $\mathrm{diag}(a^l)$ denotes a diagonal matrix with diagonal elements a^l. Applying the scalar scheme to the new variables, one obtains a scheme for the system case:

$$W_j^{n+1} = W_j^n - \frac{\lambda}{2}\Lambda(W_{j+1}^n - W_{j-1}^n) + \frac{\lambda}{2}|\Lambda|(W_{j+1}^n - 2W_j^n + W_{j-1}^n), \tag{2.10a}$$

with

$$|\Lambda| = \mathrm{diag}(|a^l|). \tag{2.10b}$$

This form looks exactly like the scalar case except it consists of m equations. Transforming back to the original variables, the scheme takes the form

$$U_j^{n+1} = U_j^n - \frac{\lambda}{2}A(U_{j+1}^n - U_{j-1}^n) + \frac{\lambda}{2}|A|(U_{j+1}^n - 2U_j^n + U_{j-1}^n), \tag{2.11a}$$

with

$$|A| = R|\Lambda|R^{-1}, \tag{2.11b}$$

or

$$U_j^{n+1} = U_j^n - \lambda\left[A^+(U_j^n - U_{j-1}^n) + A^-(U_{j+1}^n - U_j^n)\right], \tag{2.12a}$$

where

$$A^+ = \frac{1}{2}(A + |A|), \tag{2.12b}$$

$$A^- = \frac{1}{2}(A - |A|). \tag{2.12c}$$

Again, introducing a new notation "$\widetilde{F}_{j+\frac{1}{2}}$", consider a scheme of the form

$$U_j^{n+1} = U_j^n - \lambda[\widetilde{F}_{j+\frac{1}{2}}^n - \widetilde{F}_{j-\frac{1}{2}}^n], \tag{2.13}$$

where $\widetilde{F}_{j+\frac{1}{2}}$ is a numerical flux vector function. For the previous example (2.12)

$$\tilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[A(U_{j+1} + U_j) - |A|(U_{j+1} - U_j)\right]. \tag{2.14}$$

2.2. Centered (Symmetric) Schemes for Linear Hyperbolic PDEs

Several popular, spatially centered, second-order accurate schemes for the constant-coefficient scalar hyperbolic PDEs are as follows:

(i) Crank-Nicholson method:

$$u_j^{n+1} + \frac{\lambda a}{2}(u_{j+1}^{n+1} - u_{j-1}^{n+1}) = u_j^n - \frac{\lambda a}{2}(u_{j+1}^n - u_{j-1}^n). \tag{2.15}$$

(ii) Lax-Wendroff method:

$$u_j^{n+1} = u_j^n - \frac{\lambda a}{2}(u_{j+1}^n - u_{j-1}^n) + \frac{1}{2}\lambda^2 a^2(u_{j+1}^n - 2u_j^n + u_{j-1}^n). \tag{2.16}$$

(iii) MacCormack method:

$$u_j^{(1)} = u_j^n - \lambda a(u_{j+1}^n - u_j^n), \tag{2.17a}$$

$$u_j^{n+1} = \frac{1}{2}(u_j^n + u_j^{(1)}) - \frac{\lambda a}{2}(u_j^{(1)} - u_{j-1}^{(1)}). \tag{2.17b}$$

(iv) Lax-Friedrichs method:

$$u_j^{n+1} = u_j^n - \frac{\lambda a}{2}(u_{j+1}^n - u_{j-1}^n) + \frac{1}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n). \tag{2.18}$$

(v) Leap-frog method:

$$u_j^{n+1} = u_j^{n-1} - \lambda a(u_{j+1}^n - u_{j-1}^n). \tag{2.19}$$

Note that the Lax-Wendroff method can be rewritten as

$$u_j^{n+1} = u_j^n - \lambda(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n), \tag{2.20a}$$

with the numerical flux function $h_{j+\frac{1}{2}}$

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[a(u_{j+1} + u_j) - \lambda a^2(u_{j+1} - u_j)\right]. \tag{2.20b}$$

The numerical flux function $h_{j+\frac{1}{2}}$ for the Lax-Friedrichs method can be written as

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[a(u_{j+1} + u_j) - \frac{1}{\lambda}(u_{j+1} - u_j)\right]. \tag{2.21}$$

The Crank-Nicholson method, and other centered or upwind in space together with a class of linear multistep time-discretization methods can be expressed similarly. Note that the leap-frog method cannot be written in the form of equation (2.20a). Extension of these centered schemes to constant coefficient systems can be accomplished in a manner similar to the upwind methods.

III. SCHEMES FOR NONLINEAR SCALAR HYPERBOLIC CONSERVATION LAWS

As mentioned in the introduction, all of the discussion on shock-capturing methods for nonlinear hyperbolic conservation laws are for initial-value problems(IVPs). That is, the numerical schemes to be discussed are interior schemes (schemes for the interior points of the computational domain). When numerical boundary conditions (boundary schemes) are used, additional conditions have to be satisfied for the combined interior and boundary schemes (i.e. for the initial-boundary-value problems (IBVPs)) in order to maintain the same high-resolution and stability properties as the IVPs [9-11,42]. See section 3.12 or the cited references for a discussion.

It is also important to point out that the design principle of the schemes to be discussed is for homogeneous hyperbolic conservation laws (without source terms). For problems containing source terms, and especially stiff source terms and shock waves, the desired property of the scheme in question for unsteady computations might be lost or no longer valid. This point will be stressed in section 3.9.

The main theory for modern shock-capturing methods for nonlinear hyperbolic conservation laws considered in this paper relies heavily on the basic first-order upwind, the Lax-Friedrichs, the Lax-Wendroff, and the MacCormack methods. However, the resulting higher-order high-resolution modern shock-capturing methods, which are designed for the nonlinear hyperbolic conservation laws, when applied to constant-coefficient cases, are no longer linear finite-difference schemes (i.e., they are truly nonlinear finite-difference methods). This fact will be stressed in the current section.

3.1. Conservative Schemes and a Shock-Capturing Theory

The stability analysis of difference schemes for linear hyperbolic PDEs is very well established. However, the stability analysis of difference schemes for nonlinear hyperbolic PDEs is less developed. In general, the stability theory for linear difference equations is of use in checking the "local" stability of linearized equations obtained from truly nonlinear equations. However, in many instances when strong discontinuities are present, local stability is neither necessary nor sufficient for the nonlinear problems. One traditional remedy for removing instabilities is to introduce a "linear" numerical dissipation (or "artificial viscosity" or "smoothing term") into the difference schemes. One can do so by designing the scheme to be dissipative [2].

The majority of practical applications in fluid dynamics during the late seventies and early eighties is based on the traditional approach of adding an additional dissipation term to the numerical scheme to improve nonlinear stability. However, this approach alone will not guarantee convergence to a physically correct solution in the nonlinear case.

Lax and Wendroff [43] showed that the limit solution of any finite-difference scheme in a conservation form which is consistent with the conservation laws satisfies the jump conditions across a discontinuity *automatically*. This was a conceptual breakthrough which enabled the direct discretization of the conservation laws by introducing the notion of numerical dissipation. However, weak solutions (solutions with shocks and contact discontinuities) of hyperbolic conservation laws are not uniquely determined by their initial values; an entropy condition is needed to pick out the physically relevant solutions. The question arises whether finite-difference approximations converge to this particular solution. It is shown in references [24,25] that in the case of a single conservation law, monotone schemes (to be defined later) always converge to the physically relevant solution. If the scheme is not monotone, then it must be consistent with an entropy inequality for the assurance of convergence to a physically relevant solution [44,45]. Thus monotone schemes possess many desirable properties for the calculation of discontinuous solutions. Moreover, first-order upwind schemes share most of the properties of monotone schemes. The following is an introduction to monotone and first-order upwind schemes.

Before discussing monotone and first-order upwind schemes, it is important to mention a new subject area which in the author's opinion will have a dramatic impact on better understanding of numerical analysis for nonlinear ODEs and PDEs (e.g. nonlinear stability), and will provide insight on how well a numerical solution can mimic the true physics of the problem. This new subject area is hereafter referred to as "asymptotic analysis of finite difference methods by the nonlinear dynamic approach" or, for short, "chaotic dynamics". The sensitivity of numerical solutions to initial data and dependence of solutions on the discretized parameters (i.e., time step and numerical dissipation coefficients) for a fixed spatial mesh are absent from linear analysis and yet present quite often in nonlinear analysis. These phenomena are often unknown or ignored by practioners in CFD. Although the understanding of chaotic dynamical theory for PDEs and their discretized form is still in its infancy and theoretical development in this area is extremely difficult, the study of numerical analysis would not be complete without the utilization of the "nonlinear dynamic approach". For background material see references [29-32]. For an overview and potential application to CFD, see Yee [46] a paper in preparation. Since much of chaotic phenomena has a direct relation for problems containing nonlinear source terms, a short introduction to this subject will be discussed after the sections on numerical methods for nonhomogeneous hyperbolic PDEs.

3.2. Monotone and First-Order Explicit TVD Schemes

Consider the scalar hyperbolic conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \tag{3.1}$$

where $a(u) = \partial f / \partial u$ is the characteristic speed. A general three-point explicit difference scheme in conservation form can be written as

$$u_j^{n+1} = u_j^n - \lambda(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n), \tag{3.2}$$

where $h_{j+\frac{1}{2}}^n = h(u_j^n, u_{j+1}^n)$. The numerical flux function $h_{j+\frac{1}{2}}$ is required to be consistent with the conservation law in the following sense

$$h(u_j, u_j) = f(u_j). \tag{3.3}$$

Three familiar three-point conservative schemes with the numerical fluxes of the form (3.2) are the generalization of the Lax-Wendroff scheme with $h_{j+\frac{1}{2}} = \frac{1}{2}\left[f_{j+1} + f_j - \lambda(a_{j+\frac{1}{2}})^2(u_{j+1} - u_j)\right]$, the Lax-Friedrichs scheme with $h_{j+\frac{1}{2}} = \frac{1}{2}\left[f_{j+1} + f_j - \frac{1}{\lambda}(u_{j+1} - u_j)\right]$, and a generalization of the Courant-Isaacson-Rees (CIR) [47], (or the Roe's first-order upwind [48]) scheme with $h_{j+\frac{1}{2}} = \frac{1}{2}\left[f_{j+1} + f_j - |a_{j+\frac{1}{2}}|(u_{j+1} - u_j)\right]$.

Rewrite equation (3.2) as

$$u_j^{n+1} = G(u_{j-1}^n, u_j^n, u_{j+1}^n). \tag{3.4}$$

The numerical scheme (3.4) is said to be monotone if G is a monotonic increasing function of each of its arguments. Monotone schemes produce smooth transitions near discontinuities, but they are only first-order accurate [24,25]. Examples of monotone schemes are the Lax-Friedrichs scheme [25], the Godunov scheme [49], and the Engquist-Osher scheme [50].

In general, the class of first-order upwind schemes is larger than the class of monotone schemes. Not all first-order upwind schemes are monotone schemes. For example, the Godunov method is a first-order monotone upwind scheme and is of the form

$$h_{j+\frac{1}{2}} = \begin{cases} \min_{u_j \le u \le u_{j+1}} f(u) & u_j < u_{j+1} \\ \max_{u_j \ge u \ge u_{j+1}} f(u) & u_j > u_{j+1} \end{cases}. \tag{3.5}$$

The Engquist-Osher method is again a first-order monotone upwind scheme. However, the first-order upwind schemes of Huang [51] and Roe [48,52] are not monotone schemes [53]. It is instructive to see that all of these popular first-order upwind schemes (with explicit Euler time discretization) can be cast in the following form:

$$u_j^{n+1} = u_j^n - \lambda D_1^n - \lambda D_2^n. \tag{3.6}$$

Here, D_1 represents some forward difference of the f or u and D_2 represents some backward difference of the f or u. For scalar nonlinear hyperbolic conservation laws, the different representations for D_i, $i = 1, 2$, are not very crucial in terms of CPU operations count. However, in the implementation of these types of nonlinear schemes for systems of hyperbolic conservation laws via certain approximations or local linearizations (e.g., Riemann solvers; c.f. section IV), the operations count varies widely and depends on their original scalar representations. More importantly, equivalent representations of the same scheme for the scalar case turns out to be very different in operations count for their nonlinear system counterpart (employing the same Riemann solver). This fact, especially in multidimensional cases, will become apparent in section IV, where a pertinent discussion can be found.

For example, D_1 can be

$$D_1 = \mathcal{A}_1(f_j, f_{j+1}, u_j, u_{j+1})(f_{j+1} - f_j) \tag{3.7a}$$

or

$$D_1 = \mathcal{A}_2(f_j, f_{j+1}, u_j, u_{j+1})(u_{j+1} - u_j), \tag{3.7b}$$

and D_2 represents some backward difference of the f or u. For example, D_2 can be

$$D_2 = \mathcal{B}_1(f_{j-1}, f_j, u_{j-1}, u_j)(f_j - f_{j-1}) \tag{3.8a}$$

or

$$D_2 = \mathcal{B}_2(f_{j-1}, f_j, u_{j-1}, u_j)(u_j - u_{j-1}). \tag{3.8b}$$

Here $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}_1$, and $\mathcal{B}_2$ are some known functions of the arguments indicated above. As an example, consider the Engquist-Osher scheme, where the D_1 and D_2 for any convex flux function f are

$$D_1 = f^-_{j+1} - f^-_j, \qquad D_2 = f^+_j - f^+_{j-1}, \tag{3.9a}$$

with

$$f^+_j = f(\max(u_j, \overline{u})), \tag{3.9b}$$

$$f^-_j = f(\min(u_j, \overline{u})), \tag{3.9c}$$

and $\overline{u}$ is the sonic point of $f(u)$; i.e., $f'(\overline{u}) = 0$.

In [51], Huang introduced a first-order accurate upwind scheme

$$u^{n+1}_j = u^n_j - \frac{\lambda}{2}\left[1 - \operatorname{sgn}(a^n_{j+\frac{1}{2}})\right](f^n_{j+1} - f^n_j) - \frac{\lambda}{2}\left[1 + \operatorname{sgn}(a^n_{j-\frac{1}{2}})\right](f^n_j - f^n_{j-1}). \tag{3.10}$$

She was vague in defining $a_{j+\frac{1}{2}}$ for a general flux function f, but for Burgers' equation, she explicitly defined

$$a_{j+\frac{1}{2}} = (a_{j+1} + a_j)/2. \tag{3.11}$$

Here

$$D_1 = \frac{1}{2}\left[1 - \operatorname{sgn}(a_{j+\frac{1}{2}})\right](f_{j+1} - f_j), \tag{3.12}$$

which has the same form as (3.7a), and

$$D_2 = \frac{1}{2}\left[1 + \operatorname{sgn}(a_{j-\frac{1}{2}})\right](f_j - f_{j-1}), \tag{3.13}$$

which has the same form as (3.8a).

In [48], Roe defined

$$a_{j+\frac{1}{2}} = \begin{cases} (f_{j+1} - f_j)/\Delta_{j+\frac{1}{2}}u & \Delta_{j+\frac{1}{2}}u \neq 0 \\ a(u_j) & \Delta_{j+\frac{1}{2}}u = 0, \end{cases} \tag{3.14}$$

where $\Delta_{j+\frac{1}{2}}u = u_{j+1} - u_j$. This is equivalent to Huang's method for Burgers' equation.

With the definition (3.14), scheme (3.10) can be rewritten as a three-point central difference method plus a numerical viscosity term

$$u^{n+1}_j = u^n_j - \frac{\lambda}{2}\left[f^n_{j+1} - f^n_{j-1} - |a^n_{j+\frac{1}{2}}|\Delta_{j+\frac{1}{2}}u^n + |a^n_{j-\frac{1}{2}}|\Delta_{j-\frac{1}{2}}u^n\right]. \tag{3.15}$$

Now, using definition (3.14), $D_1 = \frac{1}{2}\left[a_{j+\frac{1}{2}} - |a_{j+\frac{1}{2}}|\right](u_{j+1} - u_j)$, which has the same form as (3.7b), and $D_2 = \frac{1}{2}\left[a_{j-\frac{1}{2}} + |a_{j-\frac{1}{2}}|\right](u_j - u_{j-1})$, which has the same form as (3.8b).

The numerical flux function as a function of D_1 can be written as

$$h_{j+\frac{1}{2}} = f_j + D_1. \tag{3.16}$$

Or, one can express (3.15) as (3.2) with

$$h_{j+\frac{1}{2}} = \frac{1}{2}[f_j + f_{j+1} - \psi(a_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u], \tag{3.17a}$$

and

$$\psi(a_{j+\frac{1}{2}}) = |a_{j+\frac{1}{2}}|. \tag{3.17b}$$

ψ is sometimes known as the coefficient of the numerical viscosity term. In this paper, the author prefers to use equation (3.17a) as the form of the first-order upwind numerical flux function. This form of the numerical flux function (3.17a) is not a common notation. But we can see later that representation (3.17a) is quite useful for the development of second-order TVD schemes, especially via the modified-flux approach [17]. Moreover, (3.17a) provides a more compact form for extension to nonlinear systems [39]. The two previously mentioned familiar schemes with the numerical flux function (3.17a) are the Lax-Wendroff scheme with $\psi(a_{j+\frac{1}{2}}) = \lambda(a_{j+\frac{1}{2}})^2$, and the Lax-Friedrichs scheme with $\psi(a_{j+\frac{1}{2}}) = \frac{1}{\lambda}$.

It is well known that (3.10) or (3.15) is not consistent with an entropy inequality, and the scheme might converge to a nonphysical solution. A slight modification of the coefficient of the numerical viscosity term [17],

$$\psi(z) = \begin{cases} |z| & |z| \geq \delta_1 \\ (z^2 + \delta_1^2)/2\delta_1 & |z| < \delta_1, \end{cases} \tag{3.18}$$

can remedy the entropy-violating problem. Here $\psi(z)$ is an entropy correction to $|z|$, where δ_1 is a small positive parameter or a function of z (see reference [53] for a formula for δ_1). Other ways of modifying (3.17b) to satisfy an entropy inequality can be found in [7,54,55]. One can view the size of δ_1 as a measure of the amount of numerical dissipation for the first-order upwind numerical flux (3.17a). $\delta_1 = 0$ is the least dissipative, and the larger the δ_1 the more dissipative the scheme becomes. Section VI discusses the use of δ_1 for steady-state, blunt-body hypersonic flows.

If we define

$$C^{\pm}(z) = \frac{1}{2}[\psi(z) \pm z], \tag{3.19}$$

then, this upwind scheme can be written as

$$u_j^{n+1} = u_j^n + \lambda C^-(a_{j+\frac{1}{2}}^n)\Delta_{j+\frac{1}{2}}u^n - \lambda C^+(a_{j-\frac{1}{2}}^n)\Delta_{j-\frac{1}{2}}u^n. \tag{3.20}$$

In other words, this conservative scheme can be viewed as a generalization of the Courant et al. (CIR) nonconservative upwind scheme [47].

3.3. Deficiency of Classical Shock-Capturing Methods

Although monotone schemes possess many desirable properties for the calculation of discontinuous solutions, they are only first-order accurate [24,25]. For complex flow-field structures, monotone and first-order upwind schemes are too diffusive. They cannot produce accurate solutions for complicated flow fields with a reasonable grid spacing. One needs higher-order shock-capturing methods. In the last ten years the emphasis has been on the development of better methods for problems with shocks. As discussed in the introduction, one can loosely divide higher-order shock-capturing methods into two classes. The classical one uses linear numerical dissipation; i.e., it uses the same amount everywhere or consists of adjustable parameters for each problem. The modern one uses nonlinear numerical dissipation. The amount varies from grid point to grid point and is built into the scheme with hardly any adjustable parameters.

Higher-order accurate classical upwind and symmetric (centered) shock-capturing schemes suffer from the following deficiencies: (1) they produce spurious oscillations whenever the solution is not smooth, (2) they may develop nonlinear instabilities when discontinuities are encountered, and (3) the scheme may select a nonphysical solution. Figure (3.1) shows an example of spurious oscillations associated with the classical shock-capturing method where the Burgers' equation is solved by the Lax-Wendroff method. Here the flux function $f(u) = u^2/2$. The initial condition is taken to be a sine wave and the boundary condition is taken to be periodic. The solid lines are the exact solutions at two different times and the circles are the solutions computed by the Lax-Wendroff method. At the time when the solution is still smooth the computed solution matches with the exact solution very well. However, at the time when the solution has developed into a shock, the scheme produces oscillations across the shock. The oscillations near the shock remain the same as the mesh is refined.

3.4. TVD Schemes: Background

There are basically two classes of modern shock-capturing schemes which are more appropriate for the calculation of weak solutions, namely the TVD [15-23] and ENO [12-14] schemes. In addition, these schemes should be consistent with an entropy inequality, second- or higher-order in smooth regions, and should produce high-resolution solutions near shock and contact discontinuities. Most of the available higher-order TVD and ENO schemes possess these properties. The main distinction between ENO and TVD methods is that certain types of ENO schemes can retain the same spatial order of accuracy even at points of extrema, whereas TVD schemes reduce to spatially first-order at these locations. The stencil of the globally higher-order ENO schemes is wider than TVD schemes. For example, to achieve second-order accuracy away from shocks, a five-point stencil is needed for the TVD schemes, whereas a seven-point stencil is needed for the ENO schemes. However, ENO schemes are not necessarily absolutely oscillation-free. They permit oscillation across discontinuities up to the order of the truncation error. The recently introduced total variation bounded (TVB) schemes such as Shu [56] can be globally high-order accurate in space. TVD and TVB schemes are a subset of ENO schemes and are more efficient in terms of operations count. ENO schemes are still in the development stage, whereas TVD schemes are more established in the sense of application to a wide range of multidimensional gas-dynamics problems. Only TVD schemes will be discussed here. Before a detailed discussion of the theory and the method of construction, the performance of a second-order TVD scheme will be illustrated. A second-order TVD scheme developed by Harten was applied to the same

Burgers' equation at the same two time instances as the Lax-Wendroff schemes as shown in figure (3.2). The solutions are very smooth near the shock.

This section is devoted to the introduction of TVD schemes. The definition and sufficient conditions for a scheme to be TVD will first be covered and then some first-order TVD scheme examples will be given. It turns out that all the monotone and first-order upwind schemes are first-order TVD schemes, and all first-order TVD schemes generate monotonic shock profiles. Unlike monotone schemes, not all TVD schemes are automatically consistent with an entropy inequality. Consequently, some mechanism may have to be explicitly added to TVD schemes to enforce the selection of the physical solution. An example is the entropy correction $\psi(z)$ ((3.18)) to $|z|$ for the first-order Roe scheme. It is emphasized here that the TVD property is only valid for homogeneous scalar hyperbolic conservation laws. For nonhomogeneous hyperbolic conservation laws, in order for the source term to not influence the TVD property, it is restricted to a special class of functions and fluid flows. For example, if the source term is contractive in the sense of stiff ordinary differential equations and if the governing equation contains no shock waves, it is expected that the source term will not influence the TVD property, especially if steady-state solutions are desired. This as well as TVD schemes for IBVPs will be subjects of discussions in later subsections.

Since the introduction of the notion of TVD schemes by Harten [17], many sufficient, necessary and sufficient and necessary conditions for nonlinear difference schemes to be TVD have been generalized. See for example [56-62]. Also, less restrictive conditions than TVD have been introduced. This leads to the more recent development of ENO schemes [12-14], TVB schemes [56], positive schemes [63,64] and nonlinear filter method [65]. Also, many of the popular schemes in use today are only either TVD in the semi-discrete (method-of-lines) sense [23,57,66] or only TVD for the constant coefficient case [19-21,73]. Also, a vigorous approach for a class of TVD multi-level type high-order time discretization was recently introduced by Shu [61].

Among all these variations, the basic fundamental idea can be best illustrated by Harten's original TVD notion for a one-parameter family of linear multistep time discretizations. The definition of TVB schemes and some examples will also be included in section 3.5.4.

In the following discussion, wherever there is no confusion, the term TVD scheme is loosely used for schemes that are TVD for (a) the fully discretized form, (b) the semidiscretized form, or (c) the frozen constant coefficient case.

Consider a one-parameter family of five-point difference schemes in conservation form,

$$u_j^{n+1} + \lambda\theta(h_{j+\frac{1}{2}}^{n+1} - h_{j-\frac{1}{2}}^{n+1}) = u_j^n - \lambda(1-\theta)(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n), \tag{3.21}$$

where $0 \le \theta \le 1$, $h_{j+\frac{1}{2}}^n = h(u_{j-1}^n, u_j^n, u_{j+1}^n, u_{j+2}^n)$, and $h_{j+\frac{1}{2}}^{n+1} = h(u_{j-1}^{n+1}, u_j^{n+1}, u_{j+1}^{n+1}, u_{j+2}^{n+1})$. The same numerical flux function $h_{j+\frac{1}{2}}$ with a different time-index appears on both sides of the equation. Let

$$\bar{h}_{j+\frac{1}{2}} = (1-\theta)h_{j+\frac{1}{2}}^n + \theta h_{j+\frac{1}{2}}^{n+1} \tag{3.22}$$

be another numerical flux function. Then (3.21) can be rewritten as

$$u_j^{n+1} = u_j^n - \lambda(\bar{h}_{j+\frac{1}{2}} - \bar{h}_{j-\frac{1}{2}}). \tag{3.23}$$

Here $\bar{h}_{j+\frac{1}{2}} = \bar{h}(u_{j-1}^n, u_j^n, u_{j+1}^n, u_{j+2}^n, u_{j-1}^{n+1}, u_j^{n+1}, u_{j+1}^{n+1}, u_{j+2}^{n+1})$ and is consistent with the conservation law (3.1) in the following sense

$$\bar{h}(u, u, u, u, u, u, u, u) = f(u). \tag{3.24}$$

This one-parameter family of schemes contains implicit as well as explicit schemes. When $\theta = 0$, (3.21) is an explicit method. When $\theta \neq 0$, (3.21) 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. It is not necessary to have the same numerical flux function appear in both the implicit and explicit operators. Other forms of difference schemes can also be analyzed. However, for implicit methods it will be more difficult to analyze the TVD property. To simplify the notation, rewrite equation (3.21) as

$$L \cdot u^{n+1} = R \cdot u^n, \tag{3.25}$$

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

$$(L \cdot u)_j = u_j + \lambda\theta(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}) \tag{3.26a}$$

$$(R \cdot u)_j = u_j - \lambda(1 - \theta)(h_{j+\frac{1}{2}} - h_{j-\frac{1}{2}}). \tag{3.26b}$$

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

$$TV(u^n) = \sum_{j=-\infty}^{\infty} |u_{j+1}^n - u_j^n| = \sum_{j=-\infty}^{\infty} |\Delta_{j+\frac{1}{2}} u^n|. \tag{3.27}$$

Here the general notation convention

$$\Delta_{j+\frac{1}{2}} z = z_{j+1} - z_j \tag{3.28}$$

for any mesh function z is used. The numerical scheme (3.21) for the initial-value problem (3.1) is said to be TVD if

$$TV(u^{n+1}) \leq TV(u^n). \tag{3.29}$$

The following sufficient conditions for (3.21) to be a TVD scheme are due to Harten [17]:

$$TV(R \cdot u^n) \leq TV(u^n) \tag{3.30a}$$

and

$$TV(L \cdot u^{n+1}) \geq TV(u^{n+1}). \tag{3.30b}$$

Assuming that the numerical flux $h_{j+\frac{1}{2}}$ in (3.21) is Lipschitz continuous, (3.21) can be written as

$$u_j^{n+1} - \lambda\theta\left(\widetilde{C}_{j+\frac{1}{2}}^{-} \Delta_{j+\frac{1}{2}} u - \widetilde{C}_{j-\frac{1}{2}}^{+} \Delta_{j-\frac{1}{2}} u\right)^{n+1} = u_j^n + \lambda(1 - \theta)\left(\widetilde{C}_{j+\frac{1}{2}}^{-} \Delta_{j+\frac{1}{2}} u - \widetilde{C}_{j-\frac{1}{2}}^{+} \Delta_{j-\frac{1}{2}} u\right)^n, \tag{3.31}$$

where $\tilde{C}^{\mp}_{j\pm\frac{1}{2}} = \tilde{C}^{\mp}(u_{j\mp1}, u_j, u_{j\pm1}, u_{j\pm2})$ are some bounded functions. Then Harten further showed that sufficient conditions for (3.30) are

(a) for all j

$$C^{\pm}_{j+\frac{1}{2}} = \lambda(1-\theta)\tilde{C}^{\pm}_{j+\frac{1}{2}} \geq 0 \tag{3.32a}$$

$$C^{+}_{j+\frac{1}{2}} + C^{-}_{j+\frac{1}{2}} = \lambda(1-\theta)(\tilde{C}^{+}_{j+\frac{1}{2}} + \tilde{C}^{-}_{j+\frac{1}{2}}) \leq 1, \tag{3.32b}$$

and

(b) for all j

$$-\infty < C \leq -\lambda\theta\tilde{C}^{\pm}_{j+\frac{1}{2}} \leq 0 \tag{3.33}$$

for some finite C. For example, when $\theta = 0$ and $\tilde{C}^{\pm} = C^{\pm} = \frac{1}{2}[\psi(z) \pm z]$ as defined in (3.19), the resulting scheme (3.21) is a first-order explicit upwind scheme, whereas when $\theta = 1$ with the same $\tilde{C}^{\pm}$, the scheme is a first-order implicit upwind method. Both of these methods satisfy conditions (3.32) and (3.33). By examining all the first-order upwind schemes and the Lax-Friedrichs method in section 3.3, it can be shown that they are first-order TVD schemes. However, the Lax-Wendroff method does not satisfy the TVD condition (3.30a). Conditions (3.32) and (3.33) are very useful in guiding the construction of higher-order-accurate TVD schemes which do not exhibit the spurious oscillations associated with the more classical second-order schemes. Other necessary and/or sufficient conditions for semi-discrete and fully-discrete difference methods for nonlinear hyperbolic PDEs can be found in references [23,57,61].

3.5. Higher-Order Explicit and Implicit TVD Schemes

The author is aware of primarily four different (and yet not totally distinct) design principles for the construction of high-resolution TVD-type schemes. They are as follows: (1) hybrid schemes such as the flux-corrected transport (FCT) of Boris and Book [67], Harten [68], and van Leer [69]; (2) second-order extension of Godunov's scheme by van Leer [15], and Colella and Woodward [16]; (3) the modified-flux approach of Harten [17]; and (4) the numerical fluctuation approach of Roe [18] and Sweby [70]. Also, Osher [59] has subsequently extended the first-order scheme of Engquist-Osher to second-order accuracy by using the above ideas. More recently, Jameson and Lax [23] extended the TVD idea for multi-point schemes. The following is a subjective interpretation of these design principles.

(1) The flux-corrected transport scheme is a two-step hybrid scheme consisting of a combined first- and second-order scheme. It computes a provisional update from a first-order scheme, and then filters the second-order corrections by the use of flux limiters to prevent occurrence of new extrema.

The idea of the hybrid scheme of Harten or van Leer is to take a high-order-accurate scheme and to switch it explicitly into a monotone first-order-accurate scheme when extreme points and discontinuities are encountered.

(2) Van Leer observed that one can obtain second-order accuracy in Godunov's scheme by replacing the piecewise-constant initial data of the Riemann problem with piecewise-linear initial data. The slope of the piecewise-linear initial data is chosen so that no spurious oscillations can occur. Colella and Woodward further refined van Leer's idea by using piecewise-parabolic initial data.

(3) The modified-flux TVD scheme is a technique to design a second-order accurate TVD scheme by starting with a first-order TVD scheme and applying it to a modified flux. The modified flux is chosen so that the scheme is second-order in regions of smoothness and first-order at points of extrema. Details of the construction of this scheme can be found in reference [17]. A discussion will be presented in a later section.

(4) The numerical fluctuation approach of Roe is a variation of the Lax-Wendroff scheme. Roe's variation depends on an average function. The average function is constructed (in such a way that spurious oscillations will not occur) by the use of flux limiters. As a matter of fact, under certain assumptions, a form of Roe's second-order scheme is equivalent to the modified-flux approach. Roe's second-order scheme was put into the TVD framework by Sweby [70].

Most of the above methods can also be viewed as spatially three-point central-difference schemes with an "appropriate" numerical dissipation or smoothing mechanism. "Appropriate" here means automatic feedback mechanism to control the amount of numerical dissipation, unlike the numerical dissipation used in linear theory. Design principles (2)-(4) are more closely related to the mathematical notion of TVD schemes devised by Harten. Hybrid types of schemes similar to design principle (1) do not fit in the same mathematical notion and will not be discussed in this paper. An excellent comparison between the performance of the various methods and ENO schemes on a representative scalar advection problem can be found in Zalesak [71].

In general, the basic idea of the above design principles is to construct a higher-order scheme which has properties similar to a first-order TVD scheme so that spurious oscillations cannot be generated. The main mechanisms for satisfying higher-order TVD sufficient conditions are some kind of limiting procedures called limiters (or flux limiters). They impose constraints on the gradients of the dependent variable u (slope limiters) or the flux function f (flux limiters). For constant coefficients, the two types of limiters are equivalent. One can obtain a second-order TVD scheme by modifying an upwind scheme or a centered scheme with proper limiters; i.e., if the scheme so constructed satisfies the TVD sufficient conditions. For nomenclature purposes, the term "upwind" or "symmetric" TVD schemes will be used to denote the original higher-order scheme before the application of limiters. Another way of distinguishing a higher-order upwind from a symmetric TVD scheme is that the numerical dissipation term corresponding to a higher-order upwind TVD scheme is upwind-weighted, whereas the numerical dissipation term corresponding to a symmetric TVD scheme is non-upwind-weighted (centered). This generic use of the notion upwind and symmetric TVD schemes no longer has its traditional upwinding and centering meanings. Since any high-resolution upwind (except first-order upwind schemes) or symmetric TVD schemes has the automatic feedback mechanism (slope or flux limiters) built into the scheme, a symmetric TVD scheme might become upwind biased or a high-resolution upwind scheme might become centered. Note that limiters are only used for schemes that are higher than first-order. In general, symmetric TVD schemes are simpler than the upwind TVD schemes. This point will become apparent later.

For the purposes of this paper, spatial order of accuracy of TVD schemes refers to the numerical solution that is away from extreme points (shocks and contact discontinuities).

3.5.1. Higher-Order Upwind TVD Schemes

Now we discuss the derivation of higher-order schemes in conjunction with TVD properties (i.e., the method of obtaining higher-order schemes using limiters). There are many variations but basically they can be divided into two general approaches: namely, (1) the so-called "MUSCL" (monotonic upstream schemes for conservation laws) approach due to van Leer, and (2) the others, hereafter grouped under the non-MUSCL approach. The methods of van Leer [15], Harten [17], Roe [18], Sweby [70], and Osher-Chakravarthy [22] for upwind TVD schemes, and the methods of Davis [19], Roe [20] and Yee [21] for symmetric TVD schemes are typical examples of the MUSCL and non-MUSCL approaches respectively. The "MUSCL" and "non-MUSCL" nomenclatures are used for convenience and ease of reference for later discussion. Note that one can obtain higher-order upwind TVD schemes via either the MUSCL or non-MUSCL approach.

For simplicity, take the forward Euler time differencing so the scheme has the form

$$u_j^{n+1} = u_j^n - \lambda\left[\tilde{h}_{j+\frac{1}{2}}^n - \tilde{h}_{j-\frac{1}{2}}^n\right], \tag{3.34a}$$

with

$$\tilde{h}_{j+\frac{1}{2}} = \frac{1}{2}\left[f_{j+1} + f_j + \phi_{j+\frac{1}{2}}\right]. \tag{3.34b}$$

Here the notation for the higher-order numerical flux $\tilde{h}_{j+\frac{1}{2}}$ is used to distinguish it from a first-order numerical flux $h_{j+\frac{1}{2}}$. Also the numerical fluxes described below will be a first-order time discretization for the MUSCL and Osher-Chakravarthy schemes. One would not recommend the use of the explicit Euler time discretization for these two methods, since if the limiters are not present, linear stability analysis shows that these two methods are unconditionally unstable. Forms that are second-order in time will be discussed at the appropriate places. Alternatively, one can discuss the various higher-order constructions via the semi-discrete approach. However, additional analyses are needed for the method-of-lines or multistage time differencing before the final fully discretized form is TVD. See references [57,61] for a discussion.

MUSCL (Monotonic Upstream Schemes for Conservation Laws) Approach: Consider a three-point explicit difference scheme (2.2) in conservation form,

$$\begin{aligned} u_j^{n+1} &= u_j^n - \lambda\left(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n\right) \\ &= u_j^n - \lambda\left[h(u_j^n, u_{j+1}^n) - h(u_{j-1}^n, u_j^n)\right] \end{aligned} \tag{3.35}$$

where the numerical flux $h_{j+\frac{1}{2}}$ is a function of u_j and u_{j+1}. Use the short-hand notation $\Delta_{j+\frac{1}{2}} = u_{j+1} - u_j$ (i.e., delete the u from $\Delta_{j+\frac{1}{2}}u$; $\Delta_{j+\frac{1}{2}}$ and $\Delta_{j+\frac{1}{2}}u$ will be used interchangebly throughout the text) and consider a first-order upwind numerical flux function of Roe,

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[f_j + f_{j+1} - |a_{j+\frac{1}{2}}|\Delta_{j+\frac{1}{2}}\right]. \tag{3.36}$$

Although the Roe's first-order numerical flux is used, the following discussion also applies to a variant of Roe's numerical flux and the Lax-Friedrichs numerical flux. See section 3.5.4 for a discussion.

In reference [15], van Leer observed that one can obtain spatially high-order accuracy in Godunov's scheme by replacing the piecewise constant initial data of the Riemann problem with piecewise linear initial data. The MUSCL approach applied to Roe's first-order numerical flux (instead of Godunov's scheme) to obtain a spatially higher-order differencing involves replacing the arguments u_{j+1} and u_j by $u^R_{j+\frac{1}{2}}$ and $u^L_{j+\frac{1}{2}}$, where u^R and u^L are defined for second- and third-order spatial differencings as follow:

$$u^R_{j+\frac{1}{2}} = u_{j+1} - \frac{1}{4}\left[(1-\bar{\eta})\Delta_{j+\frac{3}{2}} + (1+\bar{\eta})\Delta_{j+\frac{1}{2}}\right], \tag{3.37a}$$

$$u^L_{j+\frac{1}{2}} = u_j + \frac{1}{4}\left[(1-\bar{\eta})\Delta_{j-\frac{1}{2}} + (1+\bar{\eta})\Delta_{j+\frac{1}{2}}\right]. \tag{3.37b}$$

Here the spatial order of accuracy is determined by the value of $\bar{\eta}$:

$\bar{\eta} = -1$, fully upwind scheme
$\bar{\eta} = 0$, Fromm scheme
$\bar{\eta} = 1/3$, third-order upwind-biased scheme
$\bar{\eta} = 1$, three-point central-difference scheme

Various "slope" limiters are used to eliminate unwanted oscillations. A popular one is the "minmod" limiter; it modifies the upwind-biased interpolation as follows:

$$u^R_{j+\frac{1}{2}} = u_{j+1} - \frac{1}{4}\left[(1-\bar{\eta})\widetilde{\Delta}_{j+\frac{3}{2}} + (1+\bar{\eta})\widetilde{\widetilde{\Delta}}_{j+\frac{1}{2}}\right], \tag{3.38a}$$

$$u^L_{j+\frac{1}{2}} = u_j + \frac{1}{4}\left[(1-\bar{\eta})\widetilde{\widetilde{\Delta}}_{j-\frac{1}{2}} + (1+\bar{\eta})\widetilde{\Delta}_{j+\frac{1}{2}}\right], \tag{3.38b}$$

$$\widetilde{\Delta}_{j+\frac{1}{2}} = \text{minmod}\left(\Delta_{j+\frac{1}{2}}, \omega\Delta_{j-\frac{1}{2}}\right), \tag{3.38c}$$

$$\widetilde{\widetilde{\Delta}}_{j+\frac{1}{2}} = \text{minmod}\left(\Delta_{j+\frac{1}{2}}, \omega\Delta_{j+\frac{3}{2}}\right), \tag{3.38d}$$

where

$$\text{minmod}(x, \omega y) = \text{sgn}(x)\cdot\max\left\{0, \min\left[|x|, \omega y\,\text{sgn}(x)\right]\right\}, \tag{3.38e}$$

and $1 \le \omega \le \frac{3-\bar{\eta}}{1-\bar{\eta}}$ with $\bar{\eta} \neq 1$, and $\text{sgn}(x)$ denotes the sign of the variable x. Therefore, a spatially higher-order scheme can be obtained by simply redefining the arguments of $h_{j\pm\frac{1}{2}}$; i.e.

$$u^{n+1}_j = u^n_j - \lambda\left[h(u^R_{j+\frac{1}{2}}, u^L_{j+\frac{1}{2}}) - h(u^R_{j-\frac{1}{2}}, u^L_{j-\frac{1}{2}})\right]^n. \tag{3.39}$$

Applying the above to the first-order Roe scheme, the second-order numerical flux by the MUSCL approach denoted by $\tilde{h}^{VL}_{j+\frac{1}{2}}$ is

$$\tilde{h}^{VL}_{j+\frac{1}{2}} = h\left(u^R_{j+\frac{1}{2}}, u^L_{j+\frac{1}{2}}\right)$$

$$= \frac{1}{2}\left[f(u^R_{j+\frac{1}{2}}) + f(u^L_{j+\frac{1}{2}}) - \left|\frac{f(u^R_{j+\frac{1}{2}}) - f(u^L_{j+\frac{1}{2}})}{(u^R_{j+\frac{1}{2}} - u^L_{j+\frac{1}{2}})}\right|(u^R_{j+\frac{1}{2}} - u^L_{j+\frac{1}{2}})\right]. \tag{3.40}$$

For $\overline{\eta} = -1$, $\widetilde{\Delta_{j+\frac{1}{2}}} = \widetilde{\widetilde{\Delta_{j-\frac{1}{2}}}}$ can be the limiter function as follows:

$$\widetilde{\Delta_{j+\frac{1}{2}}} = \left\{\Delta_{j-\frac{1}{2}}\left[(\Delta_{j+\frac{1}{2}})^2 + \delta_2\right] + \Delta_{j+\frac{1}{2}}\left[(\Delta_{j-\frac{1}{2}})^2 + \delta_2\right]\right\} \Big/ \left[(\Delta_{j+\frac{1}{2}})^2 + (\Delta_{j-\frac{1}{2}})^2 + 2\delta_2\right]. \tag{3.41}$$

The parameter $10^{-7} \leq \delta_2 \leq 10^{-5}$ is a commonly used value in practical calculations. This limiter function is due to van Albada et al. [72].

One way to obtain a second-order time discretization (in addition to higher-order spatial discretization) is to replace the forward Euler time discretization by some linear multistep method, by a Runge-Kutta-type time discretization or a predictor-corrector-type time differencing. Another way (due to S. Hancock of Physics International, San Leandro, California, USA, unpublished) is to redefine $u^R_{j+\frac{1}{2}}$ and $u^L_{j+\frac{1}{2}}$ by the following:

$$u^L_{j+\frac{1}{2}} = (u_j^{n+\frac{1}{2}} + \frac{1}{2}\tilde{g}_j) \tag{3.42a}$$

$$u^R_{j+\frac{1}{2}} = (u_{j+1}^{n+\frac{1}{2}} - \frac{1}{2}\tilde{g}_{j+1}). \tag{3.42b}$$

Here $\tilde{g}_j$ is the same as $\widetilde{\Delta_{j+\frac{1}{2}}}$ in equations (3.41) or the minmod function (3.38e) (with $\omega = 1$, x replaced by $\Delta_{j+\frac{1}{2}}$ and y replaced by $\Delta_{j-\frac{1}{2}}$). The quantity $u_j^{n+\frac{1}{2}}$ is

$$u_j^{n+\frac{1}{2}} = u_j^n - \frac{\lambda}{2}\left\{f\left[(u_j^n + \frac{1}{2}\tilde{g}_j)\right] - f\left[(u_j^n - \frac{1}{2}\tilde{g}_j)\right]\right\}. \tag{3.43}$$

One can also obtain second-order time discretization by the predictor-corrector type of time differencing used by Colella and Woodward [16] and Colella [73].

Modified-Flux Approach (Harten): Now the second-order (space and time) TVD scheme of Harten [17] is considered. His method is sometimes referred to as the modified-flux approach. Apply the first-order TVD scheme to an appropriately modified flux

$$f \longleftarrow (f + g). \tag{3.44}$$

The second-order numerical flux looks exactly like the first-order scheme, except it is a function of $\tilde{f} = (f + g)$ instead of f. Thus, the second-order numerical flux denoted by $\tilde{h}^H_{j+\frac{1}{2}}$ is

$$\tilde{h}^H_{j+\frac{1}{2}} = \frac{1}{2}[\tilde{f}_j + \tilde{f}_{j+1} - \psi(\tilde{a}_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}], \tag{3.45a}$$

with

$$g_j = \text{minmod}(\sigma_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}, \sigma_{j-\frac{1}{2}}\Delta_{j-\frac{1}{2}}), \tag{3.45b}$$

$$\tilde{a}_{j+\frac{1}{2}} = a_{j+\frac{1}{2}} + \gamma_{j+\frac{1}{2}}, \tag{3.45c}$$

$$\gamma_{j+\frac{1}{2}} = \begin{cases} (g_{j+1} - g_j)/\Delta_{j+\frac{1}{2}} & \Delta_{j+\frac{1}{2}} \neq 0 \\ 0 & \Delta_{j+\frac{1}{2}} = 0, \end{cases} \tag{3.45d}$$

where the function $\psi(z)$ is an entropy correction to $|z|$ (e.g., equation (3.18)). For time-accurate calculations, $\sigma_{j+\frac{1}{2}} = \sigma(a_{j+\frac{1}{2}})$ and can be expressed as

$$\sigma(z) = \frac{1}{2}[\psi(z) - \lambda z^2] \geq 0. \tag{3.45e}$$

For steady-state calculation and/or implicit methods ($\theta \neq 0$ in (3.21)),

$$\sigma(z) = \frac{1}{2}\psi(z) \geq 0. \tag{3.45f}$$

In other words, (3.45) is a first-order numerical flux with f replaced by $\tilde{f}$ and the mean value characteristic speed $a_{j+\frac{1}{2}}$ replaced by the modified characteristic speed $\tilde{a}_{j+\frac{1}{2}} = a_{j+\frac{1}{2}} + \gamma_{j+\frac{1}{2}}$. Other more complicated forms of the g_j function which include artificial compression can be found in Harten's original paper. The current form (3.45) is quite diffusive, and a slight modification of this form without the use of *additional* artificial compression [17] by the author [74,75] will be discussed in section 3.5.2. The TVB scheme of Shu [56] and the generalization to higher-order TVD and TVB schemes using the first-order Lax-Friedrichs numerical flux [56,76] will be discussed in section 3.5.4. To illustrate the difference in shock resolution between equations (3.45) and the form suggested by the author, numerical examples for one-dimensional shock-tube problems [1] will be given in section IV and two-dimensional problems can be found in [41,74].

Roe-Sweby Second-order TVD Scheme: The scheme of Roe-Sweby [18,70] starts with a first-order upwind scheme whose numerical flux is

$$h_{j+\frac{1}{2}} = \frac{1}{2}[f_{j+1} + f_j - \text{sgn}(a_{j+\frac{1}{2}})(f_{j+1} - f_j)], \tag{3.46}$$

and adds a second-order correction term to $h_{j+\frac{1}{2}}$. The second-order (space and time) numerical flux is of the form

$$\tilde{h}^{RS}_{j+\frac{1}{2}} = h_{j+\frac{1}{2}} + \frac{\delta(r)}{2}[\text{sgn}(a_{j+\frac{1}{2}}) - \lambda a_{j+\frac{1}{2}}](f_{j+1} - f_j)], \tag{3.47a}$$

where

$$r = \frac{u_{j+1+\sigma} - u_{j+\sigma}}{\Delta_{j+\frac{1}{2}}}, \quad \sigma = \text{sgn}(a_{j+\frac{1}{2}}). \tag{3.47b}$$

Here $\delta(r)$ is a limiter and it can be

$$\delta(r) = \text{minmod}(1, r), \tag{3.47c}$$

$$\delta(r) = \frac{r + |r|}{1 + r}, \tag{3.47d}$$

$$\delta(r) = \max[0, \min(2r, 1), \min(2, r)]. \tag{3.47e}$$

Limiter (3.47d) is due to van Leer and a variant of (3.47d) with $\delta(r) = \frac{r+r^2}{1+r^2}$ is due to van Albada et al. [72]. The last limiter designed by Roe, nicknamed "superbee" [18], is the most compressive limiter among the three. It is especially designed for the computation of contact discontinuities. A wider class of limiters recently generalized by Roe is discussed in reference [60].

Osher-Chakravarthy TVD Scheme: Instead of using a MUSCL approach, Osher-Chakravarthy started with a one-parameter family of semi-discrete schemes with numerical flux

$$\begin{aligned}\tilde{h}_{j+\frac{1}{2}} = h_{j+\frac{1}{2}} &- \frac{(1-\bar{\eta})}{4}(\Delta_{j+\frac{3}{2}} f^{-}) - \frac{(1+\bar{\eta})}{4}(\Delta_{j+\frac{1}{2}} f^{-}) \\ &+ \frac{(1+\bar{\eta})}{4}(\Delta_{j+\frac{1}{2}} f^{+}) + \frac{(1-\bar{\eta})}{4}(\Delta_{j-\frac{1}{2}} f^{+}),\end{aligned} \tag{3.48}$$

where $h_{j+\frac{1}{2}} = h(u_j, u_{j+1})$ is some first-order upwind numerical flux. It can be the Engquist-Osher or Roe's first-order upwind numerical flux. Here $\bar{\eta}$ has the same meaning as before. They also derived formulas for more than five points and higher than third-order cases. Only the second- and third-order are discussed here since for practical computations, schemes larger than five-point stencil using limiters would be extremely CPU intensive.

The superscript + or − in (3.48) denotes the flux difference across the wave with positive or negative wave speed. To obtain a higher-order TVD scheme, they modified the last four terms on the right-hand side by utilizing "flux" limiters as

$$\begin{aligned}\tilde{h}^{OC}_{j+\frac{1}{2}} = h_{j+\frac{1}{2}} &- \frac{(1-\bar{\eta})}{4}(\widetilde{\Delta_{j+\frac{3}{2}}} f^{-}) - \frac{(1+\bar{\eta})}{4}(\widetilde{\widetilde{\Delta_{j+\frac{1}{2}}}} f^{-}) \\ &+ \frac{(1+\bar{\eta})}{4}(\widetilde{\Delta_{j+\frac{1}{2}}} f^{+}) + \frac{(1-\bar{\eta})}{4}(\widetilde{\widetilde{\Delta_{j-\frac{1}{2}}}} f^{+}),\end{aligned} \tag{3.49a}$$

with

$$\widetilde{\Delta_{j+\frac{3}{2}}} f^{-} = \text{minmod}[\Delta_{j+\frac{3}{2}} f^{-}, \omega\Delta_{j+\frac{1}{2}} f^{-}], \tag{3.49b}$$

$$\widetilde{\widetilde{\Delta_{j+\frac{1}{2}}}} f^{-} = \text{minmod}[\Delta_{j+\frac{1}{2}} f^{-}, \omega\Delta_{j+\frac{3}{2}} f^{-}], \tag{3.49c}$$

$$\widetilde{\Delta_{j+\frac{1}{2}}} f^{+} = \text{minmod}[\Delta_{j+\frac{1}{2}} f^{+}, \omega\Delta_{j-\frac{1}{2}} f^{+}], \tag{3.49d}$$

$$\widetilde{\widetilde{\Delta_{j-\frac{1}{2}}}} f^{+} = \text{minmod}[\Delta_{j-\frac{1}{2}} f^{+}, \omega\Delta_{j+\frac{1}{2}} f^{+}], \tag{3.49e}$$

and $1 \leq \omega \leq \frac{3-\bar{\eta}}{1-\bar{\eta}}$.

One way to obtain a second-order time discretization is to replace the forward Euler time discretization by some linear multistep method or by the Runge-Kutta type of time discretization. Note that the MUSCL way of obtaining higher-order time discretization is no longer valid in the Osher-Chakravarthy formulation.

Numerical experiments with the above four schemes on one- and two-dimensional steady-state calculations (section IV and [18,39,41,66,77-79]) show that they produce similar shock resolution.

In certain instances, the third-order scheme gives a slightly better result than the second-order method. In general, the improvement from second-order to third-order is far less pronounced than from first-order to second-order schemes. This is due partly to the fact that all TVD schemes of higher than first-order reduce to first-order at points of extrema. Knowing their performance, it is relevant to compare their computational effort.

Due to the design principle of this scheme, for $\bar{\eta} = -1$, the numerical flux $h^{OC}_{j+\frac{1}{2}}$ requires one more limiter than Harten's method. For $\bar{\eta} \neq -1$, two more limiters than van Leer's method and three more limiters than Harten's method are required in addition to an extra step of getting higher-order time discretization. The extra computations will become even more apparent as we extend these schemes to system cases. See reference [39] or section 4.6 for details.

3.5.2. Higher-Order Symmetric TVD Schemes, and Harten-Yee-Roe-Davis Generalization

The previous four methods are upwind methods. Next, the basic idea of second-order symmetric TVD schemes of Davis [19], Roe [20] and Yee [21] will be briefly described. Interested readers should consult the references cited for the actual construction of these methods.

In 1984, Davis [19] expressed a particular form of Roe-Sweby's second-order TVD scheme [18,70] as a sum of two terms. One term was the Lax-Wendroff method and the other term was an additional conservative dissipation term. He then simplified the scheme by eliminating the upwind weighting of the dissipation term and at the same time ensured that the simplified scheme still had the TVD property. Shortly after that, Roe [20] reformulated Davis's approach in a way that was easier to analyze and included a class of TVD schemes not observed by Davis. Subsequently, the author [21,80,81] generalized the Roe-Davis schemes to a one-parameter family of second-order explicit and implicit TVD schemes. The formulations of Roe-Davis can be considered as members of the explicit schemes. The main advantages of the author's formulation are that stiff problems can be handled by using implicit methods and that steady-state solutions are independent of the time step.

A general discussion and derivation with extensive numerical examples on symmetric TVD schemes can be found in a reference by Yee [21]. A careful examination of the modified-flux approach of Harten (later modified by Yee), Roe-Sweby, the symmetric TVD schemes of Davis, Roe, and Yee, and the higher-order TVD Lax-Friedrichs schemes [76] reveal that these schemes have a very similar structure and can be expressed in the same general form. They are simpler to implement than the MUSCL or the Osher-Chakravarthy schemes. Therefore, most of the numerical examples given later mainly employ these methods. Consider the general one-parameter family of explicit and implicit schemes of the form (3.21)

$$u_j^{n+1} + \lambda\theta(h^{n+1}_{j+\frac{1}{2}} - h^{n+1}_{j-\frac{1}{2}}) = u_j^n - \lambda(1-\theta)(h^n_{j+\frac{1}{2}} - h^n_{j-\frac{1}{2}}). \tag{3.50a}$$

The numerical flux for the second-order TVD schemes of Harten-Yee-Roe-Davis can be written as

$$\tilde{h}_{j+\frac{1}{2}} = \frac{1}{2}[f_j + f_{j+1} + \phi_{j+\frac{1}{2}}]. \tag{3.50b}$$

The schemes only differ in the forms of the ϕ function which are very similar to each other.

Harten-Yee Upwind Schemes : The ϕ function of Harten's original modified-flux scheme was discussed earlier. That form of the numerical flux is quite diffusive. The author's modification to equation (3.45a) is less diffusive and can be written as

$$\phi_{j+\frac{1}{2}} = \sigma(a_{j+\frac{1}{2}})(g_j + g_{j+1}) - \psi(a_{j+\frac{1}{2}} + \gamma_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}, \tag{3.51a}$$

$$\sigma(z) = \frac{1}{2}\psi(z) + \lambda\beta(1-\theta)z^2, \tag{3.51b}$$

$$g_j = \text{minmod}(\Delta_{j+\frac{1}{2}}, \Delta_{j-\frac{1}{2}}), \tag{3.51c}$$

$$\gamma_{j+\frac{1}{2}} = \sigma(a_{j+\frac{1}{2}})\begin{cases} \frac{g_{j+1}-g_j}{\Delta_{j+\frac{1}{2}}} & \Delta_{j+\frac{1}{2}} \neq 0 \\ 0 & \Delta_{j+\frac{1}{2}} = 0 \end{cases}. \tag{3.51d}$$

The coefficient β of the second term on the right-hand side of (3.51b) is the second-order time discretization. Here for an explicit method one sets $\beta = 1$, whereas for an implicit method or steady-state calculations, one sets $\beta = 0$ and $\theta \neq 0$. For implicit methods, it is not recommended that $\beta \neq 0$ be used since the resolution of the solution, in particular for steady-state computations, depends on the time step Δt. This modified form (3.51a) is just a change in the definition of the original g_j function of Harten (equation (3.45b)) by removing the $\sigma_{j\pm\frac{1}{2}}$ from (3.45b). In equation (3.51a), the $\sigma_{j+\frac{1}{2}}$ is then incorporated as a factor of $(g_j + g_{j+1})$. With the definition of (3.51), the scheme is only TVD for constant coefficient hyperbolic equations. However, it belongs to the class of positive schemes studied by Einfeldt [63].

One can generalize this method even further by including other limiters such as

$$g_j = \text{minmod}(\Delta_{j-\frac{1}{2}}, \Delta_{j+\frac{1}{2}}) \tag{3.51e}$$

$$g_j = \left(\Delta_{j+\frac{1}{2}}\Delta_{j-\frac{1}{2}} + |\Delta_{j+\frac{1}{2}}\Delta_{j-\frac{1}{2}}|\right)\Big/\left(\Delta_{j+\frac{1}{2}} + \Delta_{j-\frac{1}{2}}\right) \tag{3.51f}$$

$$g_j = \left\{\Delta_{j-\frac{1}{2}}\left[(\Delta_{j+\frac{1}{2}})^2 + \delta_2\right] + \Delta_{j+\frac{1}{2}}\left[(\Delta_{j-\frac{1}{2}})^2 + \delta_2\right]\right\}\Big/\left[(\Delta_{j+\frac{1}{2}})^2 + (\Delta_{j-\frac{1}{2}})^2 + 2\delta_2\right] \tag{3.51g}$$

$$g_j = \text{minmod}\left(2\Delta_{j-\frac{1}{2}}, 2\Delta_{j+\frac{1}{2}}, \frac{1}{2}(\Delta_{j+\frac{1}{2}} + \Delta_{j-\frac{1}{2}})\right) \tag{3.51h}$$

$$g_j = S \cdot \max\left[0, \min(2|\Delta_{j+\frac{1}{2}}|, S \cdot \Delta_{j-\frac{1}{2}}), \min(|\Delta_{j+\frac{1}{2}}|, 2S \cdot \Delta_{j-\frac{1}{2}})\right];\; S = \text{sgn}(\Delta_{j+\frac{1}{2}}). \tag{3.51i}$$

Here δ_2 is a small parameter to prevent division by zero. In practical calculations $10^{-7} \leq \delta_2 \leq 10^{-5}$ is a commonly used range. 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 arguments are of opposite sign. The limiter (3.51h) is due to Woodward and Colella [82].

Roe-Sweby Scheme: The Roe-Sweby numerical flux can be rewritten in the form (3.50) with (see section 4.6 for a derivation)

$$\phi_{j+\frac{1}{2}} = -\left[|a_{j+\frac{1}{2}}| - \frac{\delta(r)}{2}\left(|a_{j+\frac{1}{2}}| + \lambda\beta a^2_{j+\frac{1}{2}}\right)\right]\Delta_{j+\frac{1}{2}} \tag{3.52}$$

where r is the same as (3.47b).

Yee-Roe-Davis Symmetric Scheme: For the Yee-Roe-Davis symmetric TVD schemes, the $\phi_{j+\frac{1}{2}}$ can be expressed in the form

$$\phi_{j+\frac{1}{2}} = -\left[\lambda\beta(a_{j+\frac{1}{2}})^2\hat{Q}_{j+\frac{1}{2}} + \psi(a_{j+\frac{1}{2}})(\Delta_{j+\frac{1}{2}} - \hat{Q}_{j+\frac{1}{2}})\right], \tag{3.53a}$$

with $\hat{Q}_{j+\frac{1}{2}}$ chosen from

$$\hat{Q}_{j+\frac{1}{2}} = \text{minmod}(\Delta_{j+\frac{1}{2}}, \Delta_{j-\frac{1}{2}}) + \text{minmod}(\Delta_{j+\frac{1}{2}}, \Delta_{j+\frac{3}{2}}) - \Delta_{j+\frac{1}{2}}, \tag{3.53b}$$

$$\hat{Q}_{j+\frac{1}{2}} = \text{minmod}(\Delta_{j-\frac{1}{2}}, \Delta_{j+\frac{1}{2}}, \Delta_{j+\frac{3}{2}}), \tag{3.53c}$$

$$\hat{Q}_{j+\frac{1}{2}} = \text{minmod}\left[2\Delta_{j-\frac{1}{2}}, 2\Delta_{j+\frac{1}{2}}, 2\Delta_{j+\frac{3}{2}}, \frac{1}{2}(\Delta_{j-\frac{1}{2}} + \Delta_{j+\frac{3}{2}})\right]. \tag{3.53d}$$

Other forms can be found in Roe's paper [20] or one can design one's own limiter as long as a TVD sufficient condition (or a TVD necessary and sufficient condition) is satisfied. The coefficient β of the first term on the right-hand side of $\phi_{j+\frac{1}{2}}$ in (3.53a) is the second-order time discretization, and the last term can be viewed as a numerical dissipation term. The parameter β has the same meaning as in (3.51). For the explicit method ($\theta = 0$), if one sets $\beta = 1$ and $\hat{Q}$ to be the first limiter, it is the original explicit symmetric scheme of Davis. If one sets $\beta = 1$ and takes any of the three limiters, it is Roe's TVD Lax-Wendroff scheme [20]. Taking the implict scheme ($\theta \neq 0$) with $\beta = 0$ and any of the limiters, it is the form that the author proposed. It is suitable for time-accurate as well as steady-state computations [81,21]. See Roe [20] for more examples of the $\hat{Q}_{j+\frac{1}{2}}$ functions.

For analysis purposes it is sometimes convenient to let $\hat{Q}_{j+\frac{1}{2}} = Q_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}$ and to express the $\phi_{j+\frac{1}{2}}$ function for the explicit second-order symmetric TVD schemes [20] as

$$\phi_{j+\frac{1}{2}} = -\left[\lambda(a_{j+\frac{1}{2}})^2 Q_{j+\frac{1}{2}} + \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\right]\Delta_{j+\frac{1}{2}}. \tag{3.54a}$$

Let

$$r^- = \frac{\Delta_{j-\frac{1}{2}}}{\Delta_{j+\frac{1}{2}}}; \quad r^+ = \frac{\Delta_{j+\frac{3}{2}}}{\Delta_{j+\frac{1}{2}}}. \tag{3.54b}$$

The Q function can then be written as

$$Q(r^-, r^+) = \text{minmod}(1, r^-) + \text{minmod}(1, r^+) - 1, \tag{3.54c}$$

$$Q(r^-, r^+) = \text{minmod}(1, r^-, r^+), \tag{3.54d}$$

$$Q(r^-, r^+) = \text{minmod}\left(2, 2r^-, 2r^+, \frac{1}{2}(r^- + r^+)\right). \tag{3.54e}$$

The graphical representations of these three limiters for symmetric TVD schemes are shown in figure (3.3). Theoretically, one can design other limiters graphically by the aid of the sufficient condition. It is recommended that (3.53) be used instead of (3.54) for practical implementation since one does not have to make an extra logical test (or add a small number) for avoiding divison by zero if $\Delta_{j+\frac{1}{2}} = 0$.

3.5.3. Global Order of Accuracy of a Second-Order TVD Scheme [83]

One of the drawbacks of higher-order TVD schemes is that they reduce to first-order at points of extrema. In the modified-flux approach, for example, the form of g_j devised by Harten has the property of switching the second-order scheme into first-order at points of extrema (i.e., $g_j = 0$ at points of extrema). To see this, the behavior of the modified-flux approach around points of extrema is examined by considering its application to data where

$$u_{j-1} \le u_j = u_{j+1} \ge u_{j+2}. \tag{3.55}$$

In this case $g_j = g_{j+1} = 0$ in (3.45c), and thus the numerical flux (3.45a) becomes identical to that of the original first-order-accurate scheme. Consequently, the truncation error of the second-order scheme (3.34) together with (3.45) deteriorates to $O((\Delta x)^2)$ at j and $j+1$. This behavior is common to all TVD schemes, since this is one of the vehicles used to prevent spurious oscillations near a shock. Thus, a second-order TVD scheme must have a mechanism that switches itself into a first-order-accurate TVD scheme at points of extrema. Because of the above property, second-order-accurate TVD schemes are genuinely nonlinear; i.e., they are nonlinear even in the constant-coefficient case. Due to the uncertainty of the effect of the above property on the global order of accuracy, some numerical experiments were performed on the Harten scheme with an artifical compression [17] for Burgers' equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(u^2/2) = 0. \tag{3.56a}$$

Here the flux function $f(u) = u^2/2$. Since the theory of TVD schemes is only developed for initial-value problems at this point, a periodic problem was considered to avoid extra complication. The initial condition is the same as shown in figures (3.1) and (3.2), namely

$$u(x,0) = \sin \pi x, \qquad 0 \le x \le 2. \tag{3.56b}$$

The local error of the computation at each grid point $(j\Delta x, n\Delta t)$ is defined as

$$e_j = u_j^n - u(j\Delta x, n\Delta t), \tag{3.57}$$

where $u(j\Delta x, n\Delta t)$ is the exact solution of the differential equation (3.56). Here we assume that there is a fixed relation between Δt and Δx. The global order of accuracy m is determined by

$$||e|| = O(\Delta x^m) \tag{3.58}$$

as the mesh is refined for some norm.

To obtain the global order of accuracy numerically, the error at a fixed time was computed for a given mesh and repeated with increasingly finer meshes. Figure (3.4) shows the global order of accuracy of the second-order TVD scheme compared with the Lax-Wendroff method at time $t = 0.2$, when the solution is still smooth. The order of accuracy for the TVD scheme is 2 for the L_1 norm, around 3/2 for the L_2 norm, and 1 for the L_∞ norm. On the other hand, the order of accuracy for the Lax-Wendroff is 2 for all three norms. The main reason for the difference in the order of accuracy on the three norms for the TVD scheme is that the scheme automatically switches itself into first-order whenever extreme points are encountered. In this case there are two extreme points.

Next, the global order of accuracy of the two methods was examined at time $t = 1.0$ when a shock has developed. Figure (3.5a) shows the order of accuracy of the TVD scheme at $t = 1.0$, which is identical to the one at time $t = 0.2$. But the order of accuracy for the Lax-Wendroff is drastically degraded. It is 1 for the L_1 norm, around 1/2 for the L_2 norm, and 0 for the L_∞ norm. This is due to the inherent characteristic of the Lax-Wendroff method that causes this scheme to generate spurious oscillations near the shock. Figures (3.1) and (3.2) show the numerical solution of the Lax-Wendroff method compared with the second-order explicit TVD method at $t = 0.2$ and $t = 1.0$

3.5.4. Other High-Resolution Schemes with Related Properties

In sections (3.4) and (3.5.1)-(3.5.3), emphasis was on first-order upwind TVD schemes and their higher-order extension. Here, besides the symmetric TVD schemes (3.53), the use of the Lax-Friedrichs flux to obtain high-resolution shock-capturing schemes is generalized [76]. This method can be extended to a nonlinear system of equations without the use of Riemann solvers or flux-vector splitting approaches, and thus provides a large saving for multidimensional, equilibrium real gases and nonequilibrium flow computations.

The construction of second-order TVD schemes using the first-order Lax-Friedrichs numerical flux on a staggered grid was first suggested by Nessyahu and Tadmor [84]. Shu [56] suggested modifying many existing TVD schemes such as the modified flux, and the Osher-Chakravarthy methods to be TVB and at the same time use the Lax-Friedrichs numerical flux as a base. Here, the author suggests the use of MUSCL and first-order Lax-Friedrichs numerical flux to obtain high-resolution TVD or TVB schemes [76]. The advantage of using the MUSCL approach over the non-MUSCL approach together with the Lax-Friedrichs numerical flux is that when generalized to nonlinear system cases, there is no need to use any type of Riemann solver. For gas dynamics applications the limiter can be applied on the slope of the conservative variables or the primitive variables themselves, thus making the extension to multidimensional problems straightforward, since one does not have to use the local one-dimensional characteristic variables assumptions or the complicated truly multidimensional Riemann solvers. A more detailed discussion and a generalization including both the MUSCL and non-MUSCL formulation with one- and two-dimensional gas dynamics applications can be found in Yee [76]. This type of scheme, hereafter, referred to as "high-resolution TVD and TVB Lax-Friedrichs schemes" is briefly described here. The definition of TVB schemes and an example are also included.

TVB Schemes: The numerical method (3.21) for an initial-value problem of (3.1) is said to be total variation bounded (TVB) in the time interval $(0 \leq t \leq T)$ if

$$TV(u^n) \leq B, \tag{3.59}$$

for some fixed $B > 0$ depending only on u^0, and all possible n and time step Δt such that $n\Delta t \leq T$. TVB schemes are less restrictive than TVD schemes. Clearly TVD implies TVB. There are two advantages of TVB schemes over TVD schemes: (a) TVB schemes can be uniformly higher-order accurate in space including extrema points; (b) it is easier to devise boundary schemes that are TVB for the combined interior and boundary scheme.

In references [9,56], Shu showed procedures of modifying some existing TVD schemes such that the resulting schemes can be proven to be TVB and of globally higher-order accuracy in space, including extrema points. For example, by replacing the g_j function in equations (3.45b) and (3.51) by g_j^M below, the modified flux schemes can be made uniformly second-order accurate even at points of extrema. One can define

$$\begin{aligned} g_j^M =& \frac{1}{2}\text{minmod}(d_{j+\frac{1}{2}}, \omega d_{j-\frac{1}{2}} + M\Delta x^2 \text{sgn}(d_{j+\frac{1}{2}})) \\ &+ \frac{1}{2}\text{minmod}(d_{j-\frac{1}{2}}, \omega d_{j+\frac{1}{2}} + M\Delta x^2 \text{sgn}(d_{j-\frac{1}{2}})). \end{aligned} \tag{3.60}$$

with $d_{j+\frac{1}{2}} = \sigma_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}$ to replace (3.45b) for the original Harten scheme to be TVB, and $d_{j+\frac{1}{2}} = \Delta_{j+\frac{1}{2}}$ to replace (3.51) for the author's modification to Harten scheme to be TVB. Here $1 \leq \omega \leq 3$ and $M \geq 0$. Shu suggests setting $M = 50$ for the Burgers' equation computations. One can improve the global order of accuracy of the MUSCL scheme (3.39) by modifying $u^R_{j+\frac{1}{2}}$ and $u^L_{j+\frac{1}{2}}$ in equations (3.38) by

$$u^R_{j+\frac{1}{2}} = u_{j+1} - \frac{1}{4}\left[(1-\bar{\eta})\widetilde{\Delta_{j+\frac{3}{2}}}^{M} + (1+\bar{\eta})\widetilde{\widetilde{\Delta_{j+\frac{1}{2}}}}^{M}\right], \tag{3.61a}$$

$$u^L_{j+\frac{1}{2}} = u_j + \frac{1}{4}\left[(1-\bar{\eta})\widetilde{\widetilde{\Delta_{j-\frac{1}{2}}}}^{M} + (1+\bar{\eta})\widetilde{\Delta_{j+\frac{1}{2}}}^{M}\right], \tag{3.61b}$$

$$\widetilde{\Delta_{j+\frac{1}{2}}}^{M} = \text{minmod}(\Delta_{j+\frac{1}{2}}, \omega\Delta_{j-\frac{1}{2}} + M\Delta x^2 \text{sign}(\Delta_{j+\frac{1}{2}})), \tag{3.61c}$$

$$\widetilde{\widetilde{\Delta_{j+\frac{1}{2}}}}^{M} = \text{minmod}(\Delta_{j+\frac{1}{2}}, \omega\Delta_{j+\frac{3}{2}} + M\Delta x^2 \text{sign}(\Delta_{j+\frac{1}{2}})), \tag{3.61d}$$

where

$$\text{minmod}(x,y) = \text{sgn}(x)\cdot\max\left\{0, \min\left[|x|, y\,\text{sgn}(x)\right]\right\}, \tag{3.61e}$$

with $x = \Delta_{j+\frac{1}{2}}$ and $y = \omega\Delta_{j-\frac{1}{2}} + M\Delta x^2 \text{sgn}(\Delta_{j+\frac{1}{2}})$ in equation (3.61c).

One can also modify the Osher-Chakravarthy method by changing the appropriate flux limited functions as above. Modification of other methods can be found in Shu [56]. Although TVB schemes require slightly more operations per time step, the improvement in accuracy at extrema points often justifies the extra computation. Practical numerical testing is required before a conclusion can be drawn.

High-Resolution TVD and TVB Lax-Friedrichs Schemes: Recall that the first-order Lax-Friedrichs method

$$u_j^{n+1} = u_j^n - \lambda(h^n_{j+\frac{1}{2}} - h^n_{j-\frac{1}{2}}), \tag{3.62}$$

with

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[f_{j+1} + f_j - \frac{1}{\lambda}(u_{j+1} - u_j)\right] \tag{3.63}$$

satisfies the TVD condition (3.29). With a simple modification to the schemes discussed in sections 3.5.1 - 3.5.2, one can obtain a set of simpler higher-order TVD schemes than what was discussed in sections 3.5.1 - 3.5.2. Basically, if one replaces $\left|\frac{f(u^R_{j+\frac{1}{2}})-f(u^L_{j+\frac{1}{2}})}{(u^R_{j+\frac{1}{2}}-u^L_{j+\frac{1}{2}})}\right|$ in (3.40), $\psi(\tilde{a}_{j+\frac{1}{2}})$ in (3.45a), (3.51a), and (3.53a) by $1/\lambda$, high-resolution TVD schemes based on the Lax-Friedrichs numerical flux will result. Thus the corresponding MUSCL scheme using the Lax-Friedrichs numerical flux in (3.40) would be

$$\begin{aligned}\tilde{h}^{LF}_{j+\frac{1}{2}} &= h\left(u^R_{j+\frac{1}{2}}, u^L_{j+\frac{1}{2}}\right)\\ &= \frac{1}{2}\left[f(u^R_{j+\frac{1}{2}}) + f(u^L_{j+\frac{1}{2}}) - \frac{1}{\lambda}(u^R_{j+\frac{1}{2}} - u^L_{j+\frac{1}{2}})\right].\end{aligned} \tag{3.64}$$

The corresponding modified flux approach and symmetric TVD schemes using the Lax-Friedrichs numerical flux in (3.45a), (3.51a) and (3.53a) would be

$$\tilde{h}^{LF}_{j+\frac{1}{2}} = \frac{1}{2}(f_{j+1} + f_j + \phi_{j+\frac{1}{2}}), \tag{3.65}$$

$$\phi_{j+\frac{1}{2}} = \sigma(a_{j+\frac{1}{2}})(g_j + g_{j+1}) - \frac{1}{\lambda}\Delta_{j+\frac{1}{2}}, \tag{3.66}$$

for (3.45a) and (3.51a), and

$$\phi_{j+\frac{1}{2}} = -\left[\lambda\beta(a_{j+\frac{1}{2}})^2\hat{Q}_{j+\frac{1}{2}} + \frac{1}{\lambda}(\Delta_{j+\frac{1}{2}} - \hat{Q}_{j+\frac{1}{2}})\right], \tag{3.67}$$

for (3.53a). The situation is similar for the Roe-Sweby numerical flux (3.47a), and the Osher-Chakravarthy numerical flux (3.48). The proof of the TVD property and the order of accuracy of the above generalizations is left as an exercise for the reader. Some of the proof and related theory can be found in references [56,58,84,85].

The high-resolution TVD Lax-Friedrichs numerical flux construction is slightly more dissipative than the use of a first-order upwind numerical flux construction. However, for the MUSCL formulation, the Lax-Friedrichs numerical flux has the advantage of extending to a nonlinear system of equations without the use of Riemann solvers or flux-vector splitting approaches as mentioned earlier. This is due to the replacement of the dependence of a mean characteristic speed (e.g. $\psi(a_{j+\frac{1}{2}})$ of (3.51a)) by $\frac{1}{\lambda}$. This issue will become more apparent when nonlinear systems (sections IV-VI) and in particular multidimensional nonlinear systems of nonequilibrium flows (section VII) are involved.

By replacing all the limiter functions in the above high-resolution TVD Lax-Friedrichs schemes with the appropriate ones such as equations (3.60) and (3.61), one can obtain the high-resolution TVB Lax-Friedrichs schemes.

3.6. Predictor-Corrector TVD Schemes with Source Terms

All of the second-order explicit TVD schemes discussed so far are for homogeneous hyperbolic PDEs. Consider a nonlinear nonhomogeneous hyperbolic conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = s(u). \tag{3.68}$$

As noted at the end of section 3.3, the TVD property is only valid for the homogeneous part of equation (3.68). Certain types of source terms $s(u)$ might preserve the original TVD property of the homogeneous part of (3.68), and others might not. However, disregarding the type of bounded source terms, one is not precluded from the use of TVD schemes when source terms are present, but precaution has to be taken in the procedure of including the source term [86,87].

Other more sophisticated numerical methods of including source terms can be found in references [87-92]. Among these methods, Sweby [87] suggests a systematic way of approaching the problem. He utilizes a change of dependent variable to reduce the inhomogeneous nonlinear scalar hyperbolic conservation law to a homogeneous form which does possess the TVD property. He then suggests how this transformation can be used to apply TVD schemes to inhomogeneous equations. Although Sweby's systematic way of applying TVD schemes for nonhomogeneous PDEs seems to be the best way to obtain nonoscillatory solutions for problems containing source terms, Sweby's numerical study [87] indicated that there still remains a fundamental question of — what is the best treatment of the source term itself. Another drawback is that the method of tranforming to a homogeneous equation is of limited application for nonlinear systems of hyperbolic conservation laws since it is very unlikely that one can make this type of transformation for fully coupled nonlinear system cases, in particular with fully coupled source terms. Other issues related to this subject can be found in reference [86]. Here a formal extension of TVD schemes to include source terms is discussed. This method seems to work well for problems containing no shock waves, for non-stiff source terms, or for steady-state computations.

To formally include the source term efficiently, one can use (1) a two-step Lax-Wendroff-type method (e.g., explicit predictor-corrector MacCormack type method), (2) an operator-splitting procedure (similar to the time-splitting procedure in multidimensional problems except the operator-splitting procedure is on the homogeneous part and the source terms). or (3) the method-of-lines with a linear multistep approach. In this section, the predictor-corrector type method is discussed. The next two subsections will describe the other two alternatives.

A discussion and derivation of the two-step Lax-Wendroff method (without the TVD property) can be found in reference [93]. For steady-state application, to avoid additional treatment of intermediate boundary conditions and to save storage, a straightforward way of extending the second-order explict TVD scheme to include source terms is to first rewrite the numerical flux without the source term in two parts: namely, a predictor-corrector Lax-Wendroff part and a conservative numerical dissipation part. One then includes the source term in the predictor corrector step and considers the conservative numerical dissipation part as a second corrector step. Take for example the second-order explict symmetric TVD schemes ((3.50) together with (3.53)). The predictor-corrector scheme can be written as

$$u_j^{(1)} = u_j^n - \lambda\left(f_j^n - f_{j-1}^n\right) + \Delta t s_j^n \tag{3.69a}$$

$$u_j^{(2)} = \frac{1}{2}\left\{u_j^{(1)} + u_j^n - \lambda\left[f_{j+1}^{(1)} - f_j^{(1)}\right] + \Delta t s_j^{(1)}\right\} \tag{3.69b}$$

$$u_j^{n+1} = u_j^{(2)} + \frac{1}{2}\left[\bar{\phi}_{j+\frac{1}{2}} - \bar{\phi}_{j-\frac{1}{2}}\right] \tag{3.69c}$$

Here the superscript "(1)" and the overbar on $\phi_{j+\frac{1}{2}}$ and $\phi_{j-\frac{1}{2}}$ designate the values of the function evaluated at the intermediate solutions $u^{(1)}$ and $\bar{u}$. The value of $\bar{u}$ can be $\bar{u} = u_j^n$ or $\bar{u} = u_j^{(2)}$. Also $\phi_{j+\frac{1}{2}}$ has a slightly different form than (3.53a),

$$\phi_{j+\frac{1}{2}} = [\psi(\nu_{j+\frac{1}{2}}) - \nu_{j+\frac{1}{2}}^2](\Delta_{j+\frac{1}{2}} - \widehat{Q}_{j+\frac{1}{2}}) \tag{3.69d}$$

$$\nu_{j+\frac{1}{2}} = \lambda a_{j+\frac{1}{2}}. \tag{3.69e}$$

where $\psi(z)$ is (3.18).

The value $\widehat{Q}_{j+\frac{1}{2}}$ can be any of the forms defined in equations (3.53). The main difference between (3.69d,e) and (3.53a) is an extra term $\nu_{j+\frac{1}{2}}^2 \widehat{Q}_{j+\frac{1}{2}}$. However, (3.69) is the same as (3.50) together with (3.53) for the constant coefficient case. By defining a more complex $\phi_{j+\frac{1}{2}}$, scheme (3.69) can be made upwind-weighted and would belong to the class of upwind schemes. The derivation is straightforward and will not be given here.

For a high-resolution TVD Lax-Friedrichs scheme, $\psi(z) = \frac{1}{\lambda}$. In (3.69c), when $\bar{\phi}_{j+\frac{1}{2}}$ is based on $u_j^{(2)}$ rather than u_j^n, the method is no longer strictly TVD without the source term. When the source terms are present, it is not clear which approach is superior. Numerical experiments in reference [86] indicate that for the model problem (3.68) with $f(x) = u$ and $s(u) = -\bar{\mu}u(u-1)(u-\frac{1}{2})$, limiting based on u^n is preferable for small values of $\Delta t\bar{\mu}$ but that limiting based on $u^{(2)}$ may be more robust for larger values of $\Delta t\bar{\mu}$. See section 3.9 for a numerical example.

One can see that the formulation of this scheme is broken into two parts, namely, the predictor-corrector step of the MacCormack explicit scheme, and an appropriate conservative dissipation term. Here the predictor-corrector scheme is TVD in the sense of a constant-coefficient homogeneous case ($s = 0$) and with $\bar{\phi}_{j\pm\frac{1}{2}}$ evaluated at u^n instead of $u^{(2)}$. Although for the general nonlinear case, it appears to be difficult to prove that this predictor-corrector scheme is TVD, numerical experiments for one and higher-dimensional nonlinear homogeneous hyperbolic conservation laws show that (3.69) has TVD-type properties. Other equivalent predictor-corrector forms can also be used. This predictor-corrector TVD method is sometimes referred as the "TVD MacCormack" scheme. It is a slight modification of Roe's one-step TVD Lax-Wendroff scheme. If one sets $\widehat{Q}$ to be equation (3.53b), $\psi(z) = |z|$ and $\bar{u} = u^n$, the scheme is the same as described in Davis [19] and Kwong [94]. The reason for choosing the predictor-corrector step instead of the one-step Lax-Wendroff formulation is that the predictor-corrector form provides a natural and efficient inclusion of the source terms, especially for multidimensional problems [93].

3.7. Semi-Implicit Schemes for Problems Containing Stiff Source Terms [1,95,86]

The explicit TVD scheme (3.69) can be used for either time-accurate or steady-state calculations. It is second-order accurate in time and space. For stiff source terms and problems containing shock waves, a wrong speed of propagation will result. See section 3.9 for details. Moreover, if the source term is stiff, the restriction in the time step due to stability requirements is prohibitively small and (3.69) is not practical, especially for steady-state applications. In this section a semi-implicit method is described for steady-state computations (or for unsteady computations containing no shock waves). Another alternative is a fully implicit method. The basic implicit scheme and the related difficulty in extending the implicit method to higher dimensions with stiff source terms will be discussed in later sections.

The idea of treating the stiff term implicitly and the non-stiff term explicitly is a common procedure in numerical methods for stiff ordinary differential equations. The semi-implicit treatment for PDEs with stiff source terms in conjunction with classical shock-capturing methods is also a common procedure; see for example, reference [96]. What is proposed here is to replace the classical shock-capturing methods with a modern shock-capturing method. If one follows the idea of Bussing and Murman [96] in treating the source term implicitly, a semi-implicit predictor-corrector TVD scheme can easily be obtained. The basic idea is to treat the source term implicitly and the homogeneous part of the PDE with a predictor-corrector TVD scheme. For "extremely" stiff source terms, it is advisable to solve the resulting nonlinear system iteratively. However, for a "moderately" stiff source term, in order to avoid solving nonlinear equations iteratively, the Taylor expansion of the source term at time-level $n+1$ is truncated to first-order as

$$s_j^{n+1} \approx s_j^n + \left(\frac{\partial s}{\partial u}\right)_j^n (u_j^{n+1} - u_j^n). \tag{3.70}$$

The scheme can be written as a one-parameter family of time differencing schemes for the source term; i.e., the following formulation includes scheme (3.69). The semi-implicit scheme is

$$d_j^n \Delta u_j^{(1)} = -\frac{\Delta t}{\Delta x}\left(f_j^n - f_{j-1}^n\right) + \Delta t s_j^n, \tag{3.71a}$$

$$d = \left(1 - \Delta t \bar{\theta} \frac{\partial s}{\partial u}\right), \tag{3.71b}$$

$$d_j^{(1)} \Delta u_j^{(2)} = \frac{1}{2}\left\{-\Delta u_j^{(1)} - \lambda\left[f_{j+1}^{(1)} - f_j^{(1)}\right] + \Delta t s_j^{(1)}\right\}, \tag{3.71c}$$

$$u_j^{n+1} = u_j^{(2)} + \frac{1}{2}\left[\bar{\phi}_{j+\frac{1}{2}} - \bar{\phi}_{j-\frac{1}{2}}\right], \tag{3.71d}$$

with $u_j^{(1)} = \Delta u_j^{(1)} + u_j^n$, $u_j^{(2)} = \Delta u_j^{(2)} + u_j^{(1)}$ and $\bar{\phi}_{j+\frac{1}{2}}$ has the same meaning as in the previous section. Here, d is assumed to be nonzero; i.e., only the type of source terms such that d is invertible at each grid point is permissible. The parameter $\bar{\theta}$ is in the range $0 \le \bar{\theta} \le 1$. For $\bar{\theta} \neq 0$, the source term is treated implicitly. If $\bar{\theta} = 1$, the time differencing for the source term is first-order, and (3.71) is best suited for steady-state calculations. Note that the order of time-accuracy, which is determined by the parameter $\bar{\theta}$, has a different meaning than for the θ

appearing in the implicit method (3.21). To obtain a second-order time discretization, one can set $\bar{\theta} = 1/2$, and (3.71c) is replaced by

$$d_j^n \Delta u_j^{(2)} = -\lambda\left[f_{j+1}^{(1)} - f_j^{(1)}\right] + \Delta t s_j^n, \tag{3.71e}$$

and

$$u_j^{(2)} = u_j^n + \frac{1}{2}(\Delta u_j^{(1)} + \Delta u_j^{(2)}). \tag{3.71f}$$

Equation (3.71e) is very similar to (3.71c) except d and s are evaluated at u^n instead of $u^{(1)}$, and $u_j^{(2)}$ is (3.71f). By doing this, scheme (3.71,a,b,e,f) is second-order in time and space [86].

3.8. Splitting Methods and Implicit Methods for Problems Containing Source Terms [86]

In this section, two other alternatives in applying existing methods for homogeneous PDEs to include source terms are discussed. Instead of implementing the source terms directly into the predictor-corrector steps as in the previous sections, a splitting method in which the homogeneous part of the PDEs and the source term are handled in separate steps is used. The other is the method-of-lines approach (e.g., linear multistep methods).

Splitting methods: The semi-implicit predictor-corrector method attempts to handle the homogeneous terms and the source term simultaneously. An alternative approach is to employ a time-splitting in which one alternates between solving a system of conservation laws, with no source terms, and a system of ordinary differential equations modeling the source term. In the simplest case this splitting takes the form

$$u^{n+1} = \mathcal{L}_s^h \mathcal{L}_f^h u^n. \tag{3.72}$$

Here $\mathcal{L}_f^h$ represents the numerical solution operator for the conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \tag{3.73}$$

over a time step of length $h = \Delta t$, and $\mathcal{L}_s^h$ is the numerical solution operator for the ODE

$$\frac{\partial u}{\partial t} = s(u). \tag{3.74}$$

To maintain second-order accuracy, the Strang splitting [91] can be used, in which the solution u^{n+1} is computed from u^n by

$$u^{n+1} = \mathcal{L}_s^{h/2} \mathcal{L}_f^h \mathcal{L}_s^{h/2} u^n. \tag{3.75}$$

Naturally, when several time steps are taken the adjacent operators $\mathcal{L}_s^{h/2}$ can be combined to give

$$u^n = \mathcal{L}_s^{h/2} \mathcal{L}_f^h [\mathcal{L}_s^h \mathcal{L}_f^h]^{n-1} \mathcal{L}_s^{h/2} u^0. \tag{3.76}$$

In this form the method is nearly as efficient as (3.72).

The splitting approach has also frequently been used to solve reacting flow problems [98-100]. At first glance it may appear to be less satisfactory than an unsplit method such as the predictor-corrector method, since in reality the fluid dynamics and chemistry are strongly coupled and cannot be separated. However, the fact that the splitting (3.75) is second-order accurate suggests that the interaction of different effects is adequately modeled by a split method, at least for smooth solutions. Moreover, there are distinct advantages to the splitting from the standpoint of algorithm design. High quality numerical methods have been developed both for systems of conservation laws and for stiff systems of ordinary differential equations. By decomposing the problem into subproblems of these types, it is possible to take advantage of these methods directly. To some extent the mathematical theory that supports them can also be carried over. By alternating between using a high-resolution method for the conservation law and a stable stiff solver for the system of ODEs, one can easily derive a method with excellent prospects of stability on the full problem. By contrast, attempting to devise a good hybrid method handling both effects simultaneously with good accuracy and stability properties can be difficult, as has been seen in the previous section. (But in the stiff case, we will still see the problem of incorrect wave speeds with the splitting method.)

A split version of the method studied in Sections 3.6 and 3.7 might take the form

$$
\begin{aligned}
\mathcal{L}_s^{h/2}: \quad & \left[I - \frac{1}{4}\Delta t s'(u_j^n)\right]\Delta u_j^* = \frac{1}{2}\Delta t s(u_j^n) \\
& u_j^* = u_j^n + \Delta u_j^* \\[1em]
\mathcal{L}_f^h: \quad & \Delta u_j^{(1)} = -\frac{\Delta t}{\Delta x}\left[f(u_j^*) - f(u_{j-1}^*)\right] \\
& u_j^{(1)} = u_j^* + \Delta u_j^{(1)} \\
& \Delta u_j^{(2)} = -\frac{\Delta t}{\Delta x}\left[f(u_{j+1}^{(1)}) - f(u_j^{(1)})\right] \\
& u_j^{(2)} = u_j^* + \frac{1}{2}\left(\Delta u_j^{(1)} + \Delta u_j^{(2)}\right) \\
& u_j^{**} = u_j^{(2)} + (\phi_{j+1/2}^* - \phi_{j-1/2}^*) \\[1em]
\mathcal{L}_s^{h/2}: \quad & \left[I - \frac{1}{4}\Delta t s'(u_j^{**})\right]\Delta u_j^{**} = \frac{1}{2}\Delta t s(u_j^{**}) \\
& u_j^{n+1} = u_j^{**} + \Delta u_j^{**}.
\end{aligned}
\tag{3.77}
$$

Here ϕ^* involves limited fluxes as before, based on u^*. Alternatively, we can compute the limited value u_j^{**} based on $u^{(2)}$ and replace ϕ^* by $\phi^{(2)}$.

Each of these methods could be replaced by other well-known methods for the respective problems. For example, any implicit stiff solver, such as the trapezoidal method, could be used for $\mathcal{L}_s^{h/2}$ and any of a wide variety of high-resolution methods as discussed in sections 3.5 or 3.6 used for $\mathcal{L}_f^h$. Note that one does not have to use the predictor-corrector type of methods for the splitting methods for one dimension. Numerical examples will be presented in section 3.9.

Linear Multistep Implicit Method: One can include the source terms easily by using the method of lines together with semidiscrete TVD schemes. For example, one can use the same one-parameter family of time discretizations as (3.21) as follows:

$$
\begin{aligned}
u_j^{n+1} &+ \theta\left[\frac{\Delta t}{\Delta x}\left(\tilde{h}_{j+\frac{1}{2}}^{n+1} - \tilde{h}_{j-\frac{1}{2}}^{n+1}\right) - \Delta t s_j^{n+1}\right] \\
&= u_j^n - (1-\theta)\left[\frac{\Delta t}{\Delta x}\left(\tilde{h}_{j+\frac{1}{2}}^{n} - \tilde{h}_{j-\frac{1}{2}}^{n}\right) - \Delta t s_j^{n}\right].
\end{aligned}
\tag{3.78}
$$

Here θ has the same meaning as before. The numerical fluxes $\tilde{h}_{j\pm\frac{1}{2}}$ have the same meaning and form as the scheme for the homogeneous PDEs (3.1). Again, the MUSCL or non-MUSCL approaches as discussed in the previous section for the implicit method are applicable here.

3.9. A Study of Numerical Methods for Problems with Stiff Source Terms [86]

For unsteady applications, formally, one can extend the scheme to include stiff source terms as in earlier sections for steady-state calculations. However, the TVD criterion used in the extension is inappropriate for inhomogeneous problems. There are more sophisticated techniques which have been used successfully in certain problems. See for example [87,88,91,92]. The numerical study in reference [86] summarized below identified some of the essential numerical difficulties. The main difficulty seems to relate to obtaining the correct jumps in the correct locations which are absent from problems containing no shock waves.

The proper modeling of nonequilibrium gas dynamics is required in certain regimes of hypersonic flow. Often a wide range of time scales is present in the problem, leading to numerical difficulties as in stiff systems of ordinary differential equations. Stability can be achieved by using implicit methods, but other numerical difficulties are observed. The behavior of typical numerical methods on a model convection equation with a parameter-dependent source term is studied. The use of this simple equation is an attempt to bring out some of difficulties sure to be encountered in solving more realistic equations, although there are fundamental differences such that a single equation will probably be inadequate to model the hypersonic flow phenomena of interest.

Two approaches to incorporate the source terms were utilized: predictor-corrector methods with or without flux limiters, and splitting methods in which the fluid dynamics and the chemistry are handled in separate steps. Note that for the system counterpart, the homogeneous part of the conservation law and chemistry are handled in separate steps. Comparisons over a wide range of parameter values are made. On the whole, the splitting methods perform somewhat better.

In the study [86], the following questions were investigated: *i)* Can one develop stable methods? *ii)* Can one obtain "high resolution" results, with sharp discontinuities and second-order accuracy in smooth regions, and *iii)* Does one obtain the correct jumps in the correct locations?

As stated earlier, numerical stability is typically not a problem. A variety of excellent implicit methods have been developed for solving stiff systems of ODEs, and many of the same techniques can be applied to the stiff source terms in (3.68) to obtain stable methods for solving this system.

The second question is also not a problem and details can be found in [86]. In reference [86], it was shown that with some care, second-order accuracy and reasonably sharp discontinuities can be obtained.

The third question is the most interesting. For stiff reactions containing shock waves, it is possible to obtain stable solutions that look reasonable and yet are completely wrong, because the discontinuities are in the wrong locations. Stiff reaction waves move at nonphysical wave speeds, often at the rate of one grid cell per time step regardless of their proper speed.

This phenomenon has also been observed by Colella, Majda, and Roytburd [100] who made a similar study of the limiting behavior with increasing stiffness for various model systems. In particular they looked at the Euler equations coupled with a single chemistry variable representing the mass fraction of unburnt gas in a detonation wave.

Colella, Majda, and Roytburd applied Godunov's method and a high-resolution extension of Godunov's method [16] to this problem. The source terms are handled by splitting and solving the resulting ODEs exactly, so that stability is not a problem. However, they observe that on coarse grids the numerical solution is qualitatively incorrect.

The essential numerical difficulty can be identified and studied most easily by looking at even simpler equations. This same numerical behavior of discontinuities traveling at incorrect speeds can be observed in scalar problems. We have found it illuminating to study the model problem (3.68) with

$$s(u) = -\bar{\mu}u(u-1)(u-\frac{1}{2}). \tag{3.79}$$

This is the linear advection equation with a source term that is stiff for large $\bar{\mu}$. Along the characteristic $x = x_0 + t$, the solution to (3.79) evolves according to the ODE

$$\frac{d}{dt}u(x_0+t,t) = s(u(x_0+t,t)), \tag{3.80}$$

with initial data $u(x_0, 0)$. This equation has stable equilibria at $u = 0$ and $u = 1$ and an unstable equilibrium at $u = 1/2$. For large $\bar{\mu}$ and arbitrary initial data the ODE solution consists of a rapid transient with u approaching 0 (if $u(x_0,0) < 1/2$) or 1 (if $u(x_0,0) > 1/2$).

Consequently, the solution $u(x,t)$ to (3.68) together with (3.79) with initial data $u(x,0)$ rapidly approaches a piecewise constant traveling wave solution $w(x-t)$, where

$$w(x) = \begin{cases} 0 & \text{if } u(x,0) < \frac{1}{2} \\ \frac{1}{2} & \text{if } u(x,0) = \frac{1}{2} \\ 1 & \text{if } u(x,0) > \frac{1}{2}. \end{cases} \tag{3.81}$$

In particular, the solution with piecewise constant initial data

$$u(x,0) = \begin{cases} 1 & \text{if } x < x_0 \\ 0 & \text{if } x > x_0 \end{cases} \tag{3.82}$$

is simply $u(x,t) = u(x-t,0)$. In this case the ODE solution is in equilibrium on each side of the discontinuity, which theoretically behaves as it would if the source term were not present and we simply solved the linear convection equation $\partial u/\partial t + \partial u/\partial x = 0$. Numerically, however, the solution is smeared out due to discretization errors. This introduces nonequilibrium states into the wave front which turn on the source terms.

This linear discontinuity could easily be converted to a shock by replacing $\partial u/\partial x$ in (3.68) by $\partial f(u)/\partial x$ for some nonlinear flux function f.

Three numerical semi-implicit methods are used to compare their numerical behavior on the model problem for various values of $\bar{\mu}$. The first semi-implicit method that was considered is based on MacCormack's predictor-corrector method for conservation laws [27]. This second-order accurate method can be modified to include the source terms, which appear in each step of the method as in equation (3.71a,b,e,f).

In order to avoid oscillations near discontinuities, the semi-implicit MacCormack's method can be modified by adding a flux-correction step motivated by the theory of TVD methods (see section 3.7 and references [1,95]). Two different forms of this correction as in equation (3.71c,d) and (3.71e,f,d) were compared.

The third method (3.77) is the splitting method in which one alternates between solving the conservation laws (with no source terms) in one step, and the stiff systems of ODEs modeling the chemistry (with no fluid motion) in the second step. This approach has certain advantages in that high quality numerical methods exist for each of the subproblems. Combining these via splitting can yield stable, second-order accurate methods for the full problem.

Numerical tests on the model problem (3.68) and (3.79) reveal that methods can be devised by either of these approaches that will be stable and second-order accurate as the mesh is refined. However, for realistic choices of grid and time step, stiff reaction waves will have the nonphysical behavior described below.

Numerical Results on Discontinuous Data: To investigate the ability of this method to deal with propagating discontinuities, the following initial data for equation (3.68) and (3.79) were considered:

$$u(x,0) = \begin{cases} 1 & \text{if } x \leq 0.3 \\ 0 & \text{if } x > 0.3 \end{cases} \tag{3.83}$$

If the numerical computations were carried out to larger t, the solution would become nearly discontinuous. For larger values of $\bar{\mu}$ this sharpening happens more quickly. The values $u^{(1)} = u^{**} = u_j^n$ for $(s')_j^{(1)}$ and $s_j^{(1)}$ of (3.71c), and $s(u_j^{**})$ and $s((u')_j^{**})$ for (3.77) were used exclusively and the effects of the different limiters (no limiter, limiting based on u^n, and limiting based on $u^{(2)}$) were compared. Also $\Delta x = 0.02$, $\Delta t/\Delta x = 0.75$ and various values of $\bar{\mu}$ were taken. Note that due to scaling properties of the equation and method, results at time T with a particular value of $\bar{\mu}$ can equally well be regarded as results with $\bar{\mu}$ replaced by $\bar{\mu}/\widehat{\beta}$ for arbitrary $\widehat{\beta}$ at time $\widehat{\beta}T$ with time step $\widehat{\beta}\Delta t$ and grid spacing $\widehat{\beta}\Delta x$ (with x rescaled so that $[0,1]$ becomes $[0,\widehat{\beta}]$). Indeed, the critical dimensionless parameters that determine the performance of the method are the mesh ratios $\lambda = \Delta t/\Delta x$ and the product $\Delta t\bar{\mu}$ of the time step and reaction rate. The value $\Delta t\bar{\mu}$ determines the stiffness of the system. When $\Delta t\bar{\mu}$ is large, relaxation to equilibrium occurs on a time scale that cannot be temporally resolved on the grid.

All of the methods studied in this section give propagation of the step function (3.83) at incorrect speeds when the source term is sufficiently stiff; i.e., when $\bar{\mu}$ is sufficiently large. The quantity $\Delta t\bar{\mu}$, where Δt is the time step, is identified as the critical parameter affecting the propagation speed. Unless $\Delta t\bar{\mu}$ is much smaller than 1, numerical difficulties are observed.

Note that $\tau \equiv 1/\bar{\mu}$ is the relaxation time scale for the source term. Typically $\Delta t = O(\Delta x)$, where Δx is the spatial mesh width, and so Δt is the appropriate time scale for advection on the grid. Consequently, one can view $\Delta t\bar{\mu} = \Delta t/\tau$, the ratio of the advection time scale to the relaxation scale, as a sort of "cell Damköhler number".

The basic explanation is that numerical advection of the discontinuity gives a smeared representation, which includes intermediate states $0 < u < 1$ that are not in equilibrium. When $\Delta t\bar{\mu}$ is large, the source term restores near equilibrium in each time step, shifting the value in each cell towards 0 or 1 and consequently shifting the discontinuity to a cell boundary. It is thus not surprising that nonphysical propagation speeds of one cell per time step can be observed for large $\Delta t\bar{\mu}$.

Clearly this scalar model is inadequate as a full test of any numerical method. However, it does model one essential difficulty encountered in reacting flow problems, and is sufficient to point out difficulties that may arise also on more complicated systems of equations. Moreover, due to the simplicity of this equation, numerical problems that do arise can be easily understood and their source identified, yielding insight that may be valuable in developing better methods.

Figure 3.6 shows computed results at $t = 0.3$ for $\bar{\mu} = 1, 10, 100$, and 1000 ($\Delta t\bar{\mu} = 0.015, 0.15, 1.5$, and 15). Each row of figures illustrates a different value of $\Delta t\bar{\mu}$. The three figures in each row correspond to different choices of limiter. Several interesting things were observed from these graphs:

(i) For small $\Delta t\bar{\mu}$ (0.015) oscillations are visible if no limiter is used and to a lesser extent if the limiter is based on $u^{(2)}$, while limiting on u^n gives monotone results. This agrees with what is expected for the pure convection case ($\Delta t\bar{\mu} = 0$).

(ii) For larger $\Delta t\bar{\mu}$ (0.15 - 1.5) there is a slight overshoot in all cases of similar magnitude regardless of the limiter. Note that for the case of no limiter there is less oscillation here than with smaller $\Delta t\bar{\mu}$ due to the stabilizing effect of the source terms that tend to restore u towards 1.

(iii) For large $\Delta t\bar{\mu}$ (15), limiting on u^n appears to be unstable (there are large scale oscillations near $x = 0.3$ not visible in the figure) whereas limiting on $u^{(2)}$ or using no limiter gives stable results. In each case, however, the solution is completely wrong! The discontinuity has remained at its initial location $x = 0.3$ rather than propagating.

Note that for $\Delta t\bar{\mu} = 1.5$ there is also some discrepancy in the location of the discontinuity. The speed of propagation is slightly too small. For intermediate values of $\Delta t\bar{\mu}$ it is possible to obtain results with the discontinuity anywhere between 0.3 and 0.6. This phenomenon of wrong propagation speeds for large $\Delta t\bar{\mu}$ was discussed in more detail in reference [86].

The splitting method was used to compute the same model problem using the predictor-corrector form to solve the homogeneous PDE. This form was considered rather than other possibilities since it is the logical choice for comparison with the previous results. The ODE method used in (3.77) for $\mathcal{L}_s^{h/2}$ will be referred to as the "linearized implicit method". Figure 3.7 shows the same set of experiments as in Figure 3.6 but with the splitting method. The following was observed:

(i) For small $\Delta t\bar{\mu}$ either choice of limiter (based on $u^{(2)}$ or u^*) works well, and good results are obtained.

(ii) For $\Delta t\bar{\mu} = 15$ the discontinuity again moves at the wrong speed, now too fast. In fact it has moved to $x = 0.7$ and so is moving at speed $4/3$ rather than 1. Since $\Delta t/\Delta x = 3/4$, this indicates that the wave is moving at the speed of one mesh cell per time step.

(iii) A large overshoot occurs in one mesh cell behind the discontinuity for $\Delta t\bar{\mu} = 15$, regardless of the limiter used.

With regard to this last observation, it appears that the overshoot must originate within the ODE-solving step. The flux-limiter method is applied only to the homogeneous conservation law and should give no overshoots, at least in the case where one limits based on u^*. In other words, $\mathcal{L}_f^h$ keeps monotone data monotone and therefore the lack of monotoniciy must be generated by $\mathcal{L}_s^{h/2}$. Note that this solution operator works pointwise (for example u_j^* is a function only of u_j^n, independent of u_i^n for $i \neq j$), and so is oblivious to the gradient in u. What it does see, however, is a nonequilibrium value of u near the discontinuity. The linearized ODE method used in (3.77) is stable but converges in an oscillatory manner to the steady state of a stiff equation and one is seeing this here. In ODE terminology the $\mathcal{L}_s$ is not a L-stable method (see e.g. [101,102]).

It seems that these overshoots can be avoided by switching to a different ODE method. For example, if one leaves $\mathcal{L}_f^h$ unchanged but changes $\mathcal{L}_s^{h/2}$ to the trapezoidal method, then these overshoots disappear for this value of $\Delta t\bar{\mu}$ (Figure 3.8), but note that the propagation speed is still wrong. With the trapezoidal method used for the computations, the value u_j^* is obtained by solving the nonlinear equation

$$u_j^* = u_j^n + \frac{1}{4}\Delta t(s(u_j^n) + s(u_j^*)). \tag{3.84}$$

Although monotone profiles were obtained in Figure 3.8, the trapezoidal method also experiences overshoots if a still larger value of $\Delta t\bar{\mu}$ is used. The use of an L-stable method such as the backward Euler method might eliminate this problem more generally, but backward Euler is only first order accurate. One might consider the use of higher order BDF methods (the "stiffly stable methods" of [101]), but the second-order BDF method is already a two-step method and in the present context it appears to require a one-step method because of the nature of the splitting method. Implicit Runge-Kutta methods are a possibility. For reaction equations a special asymptotic method has been developed by Young and Boris [103] (see also [98]) which may avoid this problem. Another possibility is to use several steps of an ODE solver for $\mathcal{L}_s^{h/2}$; i.e., subdivide the time interval to a point where one can more adequately resolve the transient approach to equilibrium. It appears that this fails to achieve the goal of using time steps large relative to the fast time scales, but note that one would need to do this refinement in time only in regions where nonequlibrium conditions hold. At grid points where u starts out close to equilibrium (e.g. those for which $\Delta ts(u)$ is small, presumably most grid points), a single step of the linearized implicit method used in (3.77) is adequate to maintain stability.

This is another advantage of the splitting method – since the ODE solver is decoupled from the fluid solver and is applied at each grid point independently, it is easy to change the ODE solver or even to use different solvers at different points depending on the character of the flow. This approach is also advocated by Young and Boris [103], who suggest using their asymptotic integration method at stiff points and explicit Euler elsewhere.

It is emphasized here however, that improvements to the ODE solver cannot cure the problem of incorrect propagation speeds. A discussion on the source of nonphysical wave speeds and suggestion of ways (yet not too practical for physical applications) to overcome this difficulty can be found in reference [86].

3.10. Linearized Form of the Implicit TVD Schemes

All of the TVD and TVB schemes discussed above are *nonlinear schemes* in the sense that the final algorithm is *nonlinear even for the constant-coefficient case*. For implicit TVD schemes ($\theta \neq 0$ in equation (3.21)), the value of u^{n+1} is obtained as the solution of a system of nonlinear algebraic equations. To approximately solve this set of nonlinear equations noniteratively, a linearized version of these nonlinear equations is considered. For the non-MUSCL formulation (or for a spatially first-order implicit operator and MUSCL formulation for the explicit operator), linearized forms can easily be obtained. For illustration purposes, only the linearized form of implicit symmetric TVD schemes will be discussed. The same idea can be used for the implicit upwind TVD scheme (3.51) or the implicit schemes discussed in section 3.5.4. A detailed derivation can be found in Yee [74]. Also, unlike the Lax-Wendroff-type scheme, it is more straightforward to include the source terms for the implicit scheme (3.21). See sections 3.6 - 3.9 and 6.4 for a discussion.

3.10.1. Linearized Version for Constant-Coefficient Equations

For the linear scalar hyperbolic PDE (2.1), the numerical flux for the symmetric TVD scheme together with (3.54) can be written as

$$h_{j+\frac{1}{2}} = \frac{1}{2}\left[a(u_{j+1} + u_j) - |a|(1 - Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}\right]. \tag{3.85}$$

Substituting (3.85) in (3.21), one obtains

$$\begin{aligned} u_j^{n+1} &+ \frac{\lambda\theta}{2}\left[a u_{j+1} - |a|(1 - Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]^{n+1} \\ &- \frac{\lambda\theta}{2}\left[a u_{j-1} - |a|(1 - Q_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u\right]^{n+1} = \text{RHS of } (3.21). \end{aligned} \tag{3.86}$$

Here "RHS of (3.21)" means the right-hand side of equation (3.21) with $h^n_{j+\frac{1}{2}}$ defined in (3.85). Locally linearizing the coefficients of $(\Delta_{j\pm\frac{1}{2}}u)^{n+1}$ in (3.86) by dropping the time-index from $(n+1)$ to n, one obtains

$$\begin{aligned} u_j^{n+1} + \frac{\lambda\theta}{2}\Big[a u^{n+1}_{j+1} - a u^{n+1}_{j-1} &- |a|(1 - Q^n_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u^{n+1} \\ &+ |a|(1 - Q^n_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u^{n+1}\Big] = \text{RHS of } (3.21). \end{aligned} \tag{3.87}$$

Letting $d_j = u_j^{n+1} - u_j^n$ (the "delta" notation), equation (3.87) can be written as

$$e_1 d_{j-1} + e_2 d_j + e_3 d_{j+1} = -\lambda(h^n_{j+\frac{1}{2}} - h^n_{j-\frac{1}{2}}), \tag{3.88a}$$

where

$$e_1 = \frac{\lambda\theta}{2}\left[-a - |a|(1 - Q_{j-\frac{1}{2}})\right]^n, \tag{3.88b}$$

$$e_2 = 1 + \frac{\lambda\theta}{2}\left[|a|(1 - Q_{j-\frac{1}{2}}) + |a|(1 - Q_{j+\frac{1}{2}})\right]^n, \tag{3.88c}$$

$$e_3 = \frac{\lambda\theta}{2}\left[a - |a|(1 - Q_{j+\frac{1}{2}})\right]^n. \tag{3.88d}$$

The linearized form (3.88) 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, and u^{n+1}_{j+1}. Although the coefficients e_i involve five points, they are at the nth time-level.

The form of (3.88) is the same as (3.87) 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. One would expect that the linearized form (3.88) is still TVD. Numerical studies on one- and two-dimensional gas-dynamics problems support this hypothesis. It was found in reference [104] 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 in the implicit operator. Therefore, for steady-state applications, one might want to use the linearized form obtained by setting $Q_{j\pm\frac{1}{2}} = 0$ in (3.88); i.e., by redefining the coefficients in (3.88) as

$$e_1 = \frac{\lambda\theta}{2}(-a - |a|), \tag{3.89a}$$

$$e_2 = 1 + \lambda\theta(|a|), \tag{3.89b}$$

$$e_3 = \frac{\lambda\theta}{2}(a - |a|). \tag{3.89c}$$

Scheme (3.88a) together with (3.89) is spatially first-order accurate for the implicit operator and spatially second-order accurate for the explicit operator.

3.10.2. Linearized Version 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. Again, only the linearized form of the one-parameter implicit symmetric TVD schemes will be discussed. The same idea can be used for the implicit upwind TVD scheme and the implicit schemes discussed in section (3.5.4). For the symmetric TVD scheme, after substituting (3.54) in (3.21), one obtains

$$u^{n+1}_j + \frac{\lambda\theta}{2}\left[f_{j+1} - \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\Delta_{j+\frac{1}{2}}u\right]^{n+1} - \frac{\lambda\theta}{2}\left[f_{j-1} - \psi(a_{j-\frac{1}{2}})(1 - Q_{j-\frac{1}{2}})\Delta_{j-\frac{1}{2}}u\right]^{n+1} = \text{RHS of (3.21)}. \tag{3.90}$$

Unlike the constant-coefficient case, one also has to linearize $f_{j\pm1}^{n+1}$, $\psi(a_{j\pm\frac{1}{2}}^{n+1})$, and $Q_{j\pm\frac{1}{2}}^{n+1}$. Following the same procedure as in [21,74], two linearized versions of (3.90) are considered.

Linearized Nonconservative Implicit Form: Adding and subtracting f_j^{n+1} on the left-hand-side of (3.90) and using the relation (3.14), one can express (3.90) as

$$\begin{aligned} u_j^{n+1} &+ \frac{\lambda\theta}{2}\left[a_{j+\frac{1}{2}}^{n+1} - \psi(a_{j+\frac{1}{2}}^{n+1})(1 - Q_{j+\frac{1}{2}}^{n+1})\right]\Delta_{j+\frac{1}{2}}u^{n+1} \\ &- \frac{\lambda\theta}{2}\left[-a_{j-\frac{1}{2}}^{n+1} - \psi(a_{j-\frac{1}{2}}^{n+1})(1 - Q_{j-\frac{1}{2}}^{n+1})\right]\Delta_{j-\frac{1}{2}}u^{n+1} = \text{RHS of (3.21)}. \end{aligned} \tag{3.91}$$

Rewriting (3.91) in the same form as (3.31) and dropping the time-index of the coefficients of $\Delta_{j\pm\frac{1}{2}}u^{n+1}$ from $(n+1)$ to n , one obtains

$$\bar{e}_1 d_{j-1} + \bar{e}_2 d_j + \bar{e}_3 d_{j+1} = -\lambda(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n), \tag{3.92a}$$

where

$$\bar{e}_1 = \lambda\theta B^-, \tag{3.92b}$$

$$\bar{e}_2 = 1 - \lambda\theta\left(B^- + B^+\right), \tag{3.92c}$$

$$\bar{e}_3 = \lambda\theta B^+, \tag{3.92d}$$

and

$$B^{\pm} = \frac{1}{2}\left[\pm a_{j\pm\frac{1}{2}} - \psi(a_{j\pm\frac{1}{2}})(1 - Q_{j\pm\frac{1}{2}})\right]^n. \tag{3.92e}$$

Equation (3.92) is again 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, (3.92) is more applicable to steady-state calculations. A spatially first-order-accurate implicit operator similar to (3.92e) can be obtained for (3.92) 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 (3.89) might be possible [41,74]. 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 form

$$f_j^{n+1} - f_j^n = a_j^n(u_j^{n+1} - u_j^n) + O(\Delta t^2), \tag{3.93}$$

where $a_j^n = (\partial f/\partial u)_j^n$. Applying the first-order approximation of (3.93) and locally linearizing the coefficients of $(\Delta_{j\pm\frac{1}{2}}u)^{n+1}$ in (3.90) by dropping the time-index from $(n+1)$ to n, one obtains

$$u_j^{n+1} + \frac{\lambda\theta}{2}\Big[a_{j+1}^n u_{j+1}^{n+1} - a_{j-1}^n u_{j-1}^{n+1} - \psi(a_{j+\frac{1}{2}}^n)(1 - Q_{j+\frac{1}{2}}^n)\Delta_{j+\frac{1}{2}} u^{n+1} + \psi(a_{j-\frac{1}{2}}^n)(1 - Q_{j-\frac{1}{2}}^n)\Delta_{j-\frac{1}{2}} u^{n+1}\Big] = \text{RHS of (3.21)}. \tag{3.94}$$

Letting $d_j = u_j^{n+1} - u_j^n$, equation (3.94) can be written as

$$e_1 d_{j-1} + e_2 d_j + e_3 d_{j+1} = -\lambda(h_{j+\frac{1}{2}}^n - h_{j-\frac{1}{2}}^n), \tag{3.95a}$$

where

$$e_1 = \frac{\lambda\theta}{2}\left[-a_{j-1} - \psi(a_{j-\frac{1}{2}})(1 - Q_{j-\frac{1}{2}})\right]^n, \tag{3.95b}$$

$$e_2 = 1 + \frac{\lambda\theta}{2}\left[\psi(a_{j-\frac{1}{2}})(1 - Q_{j-\frac{1}{2}}) + \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\right]^n, \tag{3.95c}$$

$$e_3 = \frac{\lambda\theta}{2}\left[a_{j+1} - \psi(a_{j+\frac{1}{2}})(1 - Q_{j+\frac{1}{2}})\right]^n. \tag{3.95d}$$

The linearized form (3.95) is conservative and is a spatially five-point scheme with a tridiagonal system of linear equations. Scheme (3.95) is applicable to transient as well as steady-state calculations. As of this writing, the conservative linearized form (3.95) has not been proven to be TVD. Yet numerical study shows that for moderate CFL number, equation (3.95) produces high-resolution shocks and nonoscillatory solutions.

For steady-state applications, one can use a spatially first-order implicit operator for (3.95) by simply setting all the $Q_{j\pm\frac{1}{2}} = 0$; i.e., redefining (3.95b)-(3.95d) as

$$e_1 = \frac{\lambda\theta}{2}\left[-a_{j-1} - \psi(a_{j-\frac{1}{2}})\right]^n, \tag{3.96a}$$

$$e_2 = 1 + \frac{\lambda\theta}{2}\left[\psi(a_{j-\frac{1}{2}}) + \psi(a_{j+\frac{1}{2}})\right]^n, \tag{3.96b}$$

$$e_3 = \frac{\lambda\theta}{2}\left[a_{j+1} - \psi(a_{j+\frac{1}{2}})\right]^n. \tag{3.96c}$$

Numerical experiments with two-dimensional steady-state airfoil calculations show that this form (alternating direction implicit (ADI) version) is the most efficient (in terms of CPU time) among the various linearized methods for the case of $\theta = 1$. For time-accurate calculations, see section VI for a discussion.

3.11. Two-Parameter Family of Method-of-Lines TVD Schemes

The two-parameter family of explicit and implicit high-resolution schemes [66] presented here is based on a semi-discrete methodology and on the one-parameter family of TVD-type algorithms discussed in the previous sections. The idea is to use the same spatial discretization as before for the spatial derivatives (with $\beta = 0$) and to use the two-parameter family of linear multistep methods for the time derivatives. The original one-parameter family of TVD-type schemes is a subset of the two-parameter family of algorithms. Mathematical analysis similar to that in the previous section for the fully discretized form for the current larger family of schemes requires further investigation. Numerical experiments with steady and unsteady viscous calculations for a particular set of the two-parameter family [66] indicate that the TVD-type property is preserved. This two-parameter family of schemes takes the form

$$u_j^{n+1} + \frac{\lambda\theta}{1+\omega}\left[\widetilde{h}_{j+\frac{1}{2}}^{n+1} - \widetilde{h}_{j-\frac{1}{2}}^{n+1}\right] = u_j^n - \frac{\lambda(1-\theta)}{1+\omega}\left[\widetilde{h}_{j+\frac{1}{2}}^{n} - \widetilde{h}_{j-\frac{1}{2}}^{n}\right] + \frac{\omega}{1+\omega}(u_j^n - u_j^{n-1}). \tag{3.97}$$

Here the function $\widetilde{h}_{j+\frac{1}{2}}$ is the same higher-order numerical flux defined in the one-parameter family of schemes in section 3.5.2. This two-parameter family of algorithms contains first- and second-order implicit as well as explicit schemes. The scheme is temporally second-order if $\theta = \omega + \frac{1}{2}$ and first-order otherwise. When $\theta \neq 0$, algorithm (5) is an implicit scheme. In this paper, only the temporally first-order backward Euler ($\theta = 1$, $\omega = 0$) and the temporally second-order three-point backward differentiation ($\theta = 1$, $\omega = 1/2$) time differencing are investigated. One can use the same procedure as in section 3.10 to obtain a linearized form of implicit methods for (3.97). Note that ω in (3.97) is different from the ω in (3.38).

3.12. IBVPs and TVD/TVB Schemes

All of the previous discussion on TVD or TVB schemes was for IVPs. That is, these numerical schemes are interior schemes (schemes for the interior points of the computational domain). For IBVPs, one hopes that the boundary treatment still retains the TVD or TVB (or at least high-resolution) property of the scheme. However, analysis shows that when numerical boundary conditions (boundary schemes) are used additional conditions have to be satisfied for the combined interior and boundary schemes (i.e. for the initial-boundary-value problems) in order to maintain the same high-resolution and stability properties as the IVPs [9-11,42].

It is important to distinguish the difference between linearized stability for IBVPs and maintaining TVD or TVB properties for IBVPs. A combined interior and boundary scheme (of which the individual scheme is TVD for the IVP) might be stable in the linearized sense and yet might not satisfy the TVD or TVB property for the IBVP. For example, (3.21) with $\theta = 1/2$ and (3.51) is unconditionally stable in the linearized sense but TVD if $\lambda|a_{j+\frac{1}{2}}| < 2$. A common boundary treatment for use with high-order interior schemes is to use a lower-order scheme, such as a lower order upwind scheme and/or to use extrapolation near the boundary. These boundary schemes seem to be standard practice in CFD circles and most of the computations presented in this manuscript use these approaches. While these boundary treatments are widely used, there are, however, a few theoretical drawbacks concerning their usage; namely, (a) the local order of the scheme near the boundary is reduced, (b) these boundary treatments only proved to be linearly stable and (c) the proof of TVD and TVB properties for the combined interior and boundary schemes for the IBVPs appears to be extremely difficult. In reference [9], Shu introduces an approach to the treatment of boundaries which uses the same high-order scheme up to the boundary, plus extrapolation and an upwind treatment at the boundary. The resulting scheme is proved to be TVB for the scalar nonlinear case and for linear systems. His boundary treatment is based on the globally high-order TVB schemes discussed in [9,56,61]. His schemes have natural upwind-downwind decompositions which help to implement and prove the TVB boundary treatments. Interested readers should consult Shu's papers for details. Theoretical results for a combined interior and boundary scheme to be TVD remain an open question.

3.13. Asymptotic Analysis of Finite Difference Methods by the Nonlinear Dynamic Approach

This section presents a brief discussion on the uniqueness, stability and accuracy of difference methods for nonlinear PDEs by the "nonlinear dynamic approach". Many existing results for nonlinear dynamical systems such as chaos, bifurcations, and limit cycles have a direct application to problems containing nonlinear source terms such as the reaction-diffusion or the reaction-convection-diffusion equations. With the advent of increasing demand for numerical accuracy, stability, efficiency, and uniqueness of numerical solutions in modeling such equations, an interdisciplinary approach for the analysis of these systems is needed.

Dynamical systems is a broad area of inquiry focusing on the development in time — continuous or discrete — of a wide variety of systems. The models that simulate these systems describe generic phenomena of interest to biologists, economists, physicists, chemists and engineers. This subject area, which in the author's opinion would have a dramatic impact in the better understanding of numerical analysis for nonlinear ODEs and PDEs, would provide a better judgement on how well a numerical solution can mimic the true physics of the problems. This proposed subject area hereafter is referred to as "asymptotic analysis of finite difference methods by the nonlinear dynamic approach" or for short, "chaotic dynamics". The sensitivity of numerical solutions to initial data and the dependence of solutions on the temporal time step and numerical dissipation coefficients as parameters for a fixed spatial mesh are absent from linear analysis and yet present quite often in nonlinear analysis. These phenomena are often not known or ignored by practioners in CFD. Although the theory of chaotic dynamics for continuous systems and for discrete maps have flourished rapidly for the last decade, there are very few investigators addressing the issue of the correlation between the nonlinear dynamical behavior of the continuous systems and the corresponding discrete map resulting from finite difference discretizations. In addition, although the understanding of chaotic dynamical theory for PDEs and their discretized form is still in its infancy and theoretical development in this area is extremely difficult, the study of numerical analysis would not be complete without the utilization of the nonlinear dynamic approach and the correlation with the dynamical behavior of the corresponding PDEs. For background material see references [29-32,105-109]. For an overview and potential application to CFD, see Yee [46] a paper in preparation. The following is an attempt to give a flavor of the subject and at the same time provide a possible application in aerodynamics.

In particular, one of the recent emphases in computational fluid dynamics has been the development of appropriate finite difference methods for nonequilibrium gas dynamics in the hypersonic range. A nonlinear scalar model equation would be of the form

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \varepsilon \frac{\partial^2 u}{\partial x^2} + s(u), \qquad \varepsilon \text{ a constant} \tag{3.98}$$

for the reaction-convection-diffusion model, or

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=s(u), \tag{3.99}$$

without the diffusion term. The nonlinear source term (or the reaction term) $s(u)$ can be very stiff and the simplest nonlinear function $f(u)$ can be taken as $u^2/2$. In order to understand the fundamentals, (3.99) can be simplified as

$$\frac{\partial u}{\partial t}+\frac{\partial (u^2/2)}{\partial x}=-\bar{\mu}u(u-1)(u-\frac{1}{2}), \tag{3.100a}$$

or

$$\frac{\partial u}{\partial t}+c\frac{\partial u}{\partial x}=-\bar{\mu}u(u-1)(u-\frac{1}{2}), \qquad c \text{ a constant} \tag{3.100b}$$

Note that phenomena such as chaos, bifurcations and limit cycles seem to only relate to source terms $s(u)$ which are nonlinear in u. Equation (3.100b) leads to a linear advection equation with a source term that is stiff for large $\bar{\mu}$. Equation (3.98) can be viewed as a model equation in combustion or as one of the species continuity equations in nonequilibrium flows (except in this case, the source term is coupled with other species mass fractions). The above model equations are a good starting point in the investigation of correlation between theory of chaotic dynamical systems [110,111] and uniqueness, stability, accuracy and convergence rate of finite difference methods for computational fluid dynamics.

The main interest is to investigate what types of new phenomena arise from the numerical methods (e.g., spatio-temporal chaos for nonlinear partial difference equations) but not from the original PDE as a function of the stiff coefficient $\bar{\mu}$, the diffusion coefficient ε, and the time step Δt with a fixed (or variable) grid spacing Δx. The time step can vary a lot depending on whether the time discretization is explicit or implicit. More precisely, one wants to weed out all undesirable phenomena due to the numerical method (e.g., additional equilibrium points introduced by the numerical methods, domain of attraction, etc.) and to identify whether the numerical method really describes the true solution of the PDE under a prescribed initial and boundary condition with $\bar{\mu}$, ε, and the time step Δt being parameters. The study can be divided into steady and unsteady behavior with or without *shock waves*. But most of all, before anything else, one wants to fully understand the topographical behavior (e.g. solution trajectories in phase planes) of these PDEs as the parameters $\bar{\mu}$ and ε are varied. At this moment, understanding the topographical behavior of nonlinear dynamical systems relies heavily on numerical methods.

There is an added stumbling block as mentioned in the previous section, namely: the difficulties concerning the selection of proper numerical methods or devising suitable yet feasible schemes for practical computations for hyperbolic conservation laws with nonlinear stiff source terms. This stems partly from the fact that schemes were designed for homogeneous hyperbolic conservation laws. Most often for low speed flows, if source terms exist, they are not stiff and are well-behaved. However, with the increased interest in high speed and reentry vehicle type aerodynamics design, suitable and yet efficient numerical methods are a vital tool for the advancement in this area and such numerical methods have not yet been discovered. To gain some insight, conventional finite difference methods such as upwind or central differences plus numerical dissipation methods for the convection term and central differences for the diffusion term can be used for analysis. Time differencing can be the class of a two-parameter family of explicit and implicit linear multistep methods or multistage methods. Most of them are similar to the ones that Iserles [105], Mitchell [106] and Sleeman et al. [107] have studied for the nonlinear ODE dynamical systems.

Another difficulty is that combustion-related and high speed hypersonic flow problems usually contain multiple equilibrium states and shock waves that are inherent in the governing equations. Furthermore, additional equilibrium states might be introduced due to the time differencing. In many instances the stable and unstable states, whether due to the physics or spurious in nature (i.e., introduced by the numerical methods), are interwoven over the domain of interest and are usually very sensitive to initial conditions and the time steps (even when the chosen time step is within the linearized stability limit) as well as variation of parameters such as angle of attack, Reynolds number and coefficients of physical and numerical dissipations.

On the subject of sensitivity to initial data, the basin of attraction or domain of attraction (i.e. the domain for which the set of initial conditions that are time-asymptotically approach to a specified stable equilibrium state) might be very different between the governing differential equations and the discretized counterpart. The basin of attraction might contract in one direction, expand in the other direction or be very different from the original PDEs depending on the numerical methods. In many instances, even with the same spatial discretization but different time discretizations, the basins of attraction can also be extremely different. However, mapping out the basin of attraction for any nonlinear continuum dynamical system other than the very simple scalar equations relies on numerical methods. The type of nonlinear behavior and the sensitivity to initial conditions for both the PDEs and their discretized counterparts make the understanding of the true physics extremely difficult when numerical methods are the sole source. Under this situation, how can one delineate the numerical solutions that approximate the true physics from the numerical solutions that are spurious in nature? Numerical studies show that for relatively simple equations and well-known time discretizations with modest step-lengths, some schemes can converge to a false asymptote or wrong solution in a deceptively "smooth" manner [105].

The above subject area is especially important for employing a "time-dependent" approach to the steady state with a given initial data. In many CFD computations, the initial data are not known. A freestream condition or an intelligent guess for the initial condition is used. In particular, the controversy of the "existence of multiple steady-state solutions" will not be exactly resolved until there is a better understanding of basins of attraction for both the PDEs and the discretized equations.

I'd like to end this section with a direct quote from Sanz-Serna and Vadillo's paper [31]. This quote indicates the danger of relying on linearized stability and convergence theory in analyzing nonlinear dynamical problems. Reference [31] is one of the few papers trying to convey to numerical analysts the flavor of the powerful "nonlinear dynamic approach". Hopefully with section 3.13, I can convey to computational fluid dynamicists the flavor of the importance of the "nonlinear dynamic approach" in CFD analysis.

"Assume that the convergence of a numerical method has been established; it is still possible that for a given choice of Δt, or even for any such a choice, the qualitative behaviour of the numerical sequence u^0, u^1, ..., u^n, ... be competely different from that of the theoretical sequence $u(t_0)$, $u(t_1)$, ... $u(t_n)$, ... This discrepancy which refers to n tending to ∞, Δt fixed cannot be ruled out by the convergence requirement, as this involves a different limit process (namely Δt tending to 0)." ———————————————————————————————— "The fact that analyses based on linearization cannot accurately predict the qualitative behaviour of u^n for fixed Δt should not be surprising: there is a host of nonlinear phenomena (chaos, bifurcations, limit cycles ...) which cannot possibly be mimicked by a linear model."

The unresolved problem of fixed step stability for calculations involving nonlinear ODEs and its link with dynamical systems was disscussed by Iserles [105].

IV. EXTENSION OF NONLINEAR SCALAR TVD SCHEMES TO 1-D NONLINEAR SYSTEMS

Before getting into a detailed discussion, it is important to emphasize that first-order upwind schemes for one-dimensional nonlinear systems encountered in the literature differ mainly by their so-called "Riemann solver". For constant-coefficient systems, they all reduce to the CIR (Courant-Isaacson-Rees) method [47]. There exist three popular ways of extending scalar schemes to nonlinear systems (hereafter, referred to as Riemann solvers): the exact Riemann solvers [49,112], the approximate Riemann solvers [48,113], and the flux-vector splitting techniques [114,115]. In this paper, the discussion and numerical comparison only stress the use of high-resolution TVD schemes in conjunction with an approximate Riemann solver of Roe, and the flux-vector splittings of Steger-Warming and of van Leer for perfect gases. Comparison of these methods in conjunction with the generalized Roe's approximate Riemann solver of Vinokur [116], and the generalized flux-vector splittings of Vinokur and Montagné [117] for equilibrium real gases are also considered. Generalization of the high-resolution schemes to nonequilibrium flows was discussed in references [95,99,100,118-120] and will be discussed in section VII.

Recent developments of higher-order modern shock-capturing methods stress the nonlinear scalar case. The extension of higher-order modern shock-capturing methods to systems relies heavily on these Riemann solvers. Furthermore, since the extension is not unique, it depends on the form of the nonlinear schemes that one started with [39]. One form might be simpler than another for its system counterpart. The situation arises even when one starts with a scheme with two different representations for the scalar case. A discussion on this subject will be presented in section 4.6. For example, the Osher-Chakravarthy scheme is more complicated and more expensive in the system cases than other schemes under discussion. See reference [39] or section 4.6 for a discussion. Moreover, a comparison among the numerical results for schemes (3.39), (3.47), (3.49), (3.51), and (3.53) does not indicate any advantage of the more complicated schemes over the simpler ones. Based on this fact, numerical results presented here reflect the author's personal experiences and preferences for certain schemes. No attempt has been made to present a unified comparison. Also, no effort has been made to collect numerical results from investigators in related fields to illustrate the performance of similar schemes. Readers are encouraged to study the related theory [15-23,77,114,115] and numerical results of references [17,77,82,121-128].

Since the current discussion is on conservative shock-capturing finite-difference methods, nonconservative schemes such as the lambda scheme of Moretti [129] will not be discussed, although recent results of Moretti [130,131] show that the lambda scheme together with a shock-fitting procedure suggests an efficient alternative to shock-capturing methods. An excellent detailed discussion on the lambda scheme can be found in Pandolfi [91]. More detailed study and numerical tests are needed in this direction in the near future. Also, the FCT scheme of Boris and Book [67] for nonlinear systems does not make use of any type of Riemann solver and the flux limiters were applied directly to the flux functions even for a nonlinear coupled system of equations. Thus this scheme will not fall into the category of the current discussion. Historically, Boris and Book were two of the pioneers in introducing the concept of flux limiters for nonlinear scalar hyperbolic conservation laws. However, for nonlinear systems they applied flux limiters to the individual flux functions. In the author's opinion this is less effective than the use of Riemann solvers. A discussion is included below. Also the popular and efficient shock-capturing method of Jameson et al. will not be discussed here. The scheme of Jameson et al. [33] seems to fall in between the classical and modern shock-capturing schemes. Their scheme, originally designed for the Euler equations, uses two adjustable parameters to control the amount of numerical dissipation, and also does not use any Riemann solver. The relative advantages and disadvantages of the schemes of Moretti, Boris and Book, Jameson et al., and TVD schemes for weak and moderate shock-wave calculations are not clear at this point. However, for strong shock waves, especially in the hypersonic regime, TVD or TVB schemes in conjunction with the appropriate Riemann solvers are conjectured to perform better than the other aforementioned approaches. One possible efficient type of numerical method that can avoid the use of Riemann solvers and yet have the TVD or TVB properties are the high-resolution TVD and TVB Lax-Friedrichs methods using the MUSCL as suggested in section 3.5.4 and in reference [76]. See sections 4.4 and 4.5 for a discussion.

4.1. Methods of Extension (Riemann Solvers)

The following is a brief review of the methods of extending nonlinear scalar difference schemes to nonlinear systems of hyperbolic conservation laws. The objective is to give a flavor of the developments and therefore many of the details are left out. First the various methods of extension to systems and the *original* use of these methods in conjunction with the first-order upwind finite-difference methods will be discussed. All of the original methods of extension to systems were developed for perfect or thermally perfect gases and the resulting algorithms for the gas-dynamics equations were first-order accurate (except in flux-vector splitting approaches). In the subsequent section, generalization of these methods to equilibrium real gases will be described in conjunction with numerical schemes that are higher than first-order.

The conservation laws for the one-dimensional Euler equations for equilibrium real gases can be written in the form

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

where the column vectors U and $F(U)$ take the form

$$U = \begin{bmatrix} \rho \\ m \\ e \end{bmatrix}, \qquad F = \begin{bmatrix} \rho u \\ mu + p \\ eu + pu \end{bmatrix}. \tag{4.1b}$$

Here ρ is the density, $m = \rho u$ is the momentum per unit volume, p is the pressure, $e = \rho(\epsilon + \frac{1}{2}u^2)$ is the total internal energy per unit volume, and ϵ is the internal energy per unit mass. Note that there is an arbitrary additive constant in the definition of ϵ.

Many approximate Riemann solvers make use of the eigenvalues and eigenvectors of the Jacobian matrix $A = \partial F/\partial U$. For a general gas, one therefore requires the thermodynamic derivatives of p. In terms of the internal energy per unit volume $\tilde{\epsilon} = \rho\epsilon$, the differential of p can be written as

$$dp = \chi\, d\rho + \kappa\, d\tilde{\epsilon}, \tag{4.2a}$$

where the thermodynamic derivatives are defined as

$$\chi = \left(\frac{\partial p}{\partial \rho}\right)_{\tilde{\epsilon}} ; \qquad \kappa = \left(\frac{\partial p}{\partial \tilde{\epsilon}}\right)_{\rho} . \tag{4.2b}$$

Here the subscript $\tilde{\epsilon}$ (ρ) means the partial derivative of p with respect to ρ ($\tilde{\epsilon}$) by holding $\tilde{\epsilon}$ (ρ) constant. If $h = \epsilon + p/\rho$ is the specific enthalpy, one can obtain for the speed of sound c the relation

$$c^2 = \chi + \kappa h. \tag{4.3}$$

For a perfect gas with $p = \rho\epsilon(\gamma - 1)$, where γ is the ratio of specific heat, $\chi = 0$, and $\kappa = (\gamma - 1)$. For a thermally perfect gas $p = \rho w(\epsilon)$, the temperature $T = w(\epsilon)$ where w is a function of ϵ only, and $\chi = w(\epsilon)$.

The Jacobian matrix A takes the form

$$A = \begin{bmatrix} 0 & 1 & 0 \\ \chi - (2-\kappa)u^2/2 & (2-\kappa)u & \kappa \\ (\chi + \kappa u^2/2 - H)u & H - \kappa u^2 & (1+\kappa)u \end{bmatrix}, \tag{4.4}$$

where $H = h + u^2/2$ is the total enthalpy. The three eigenvalues of A are

$$a^1 = u - c, \quad a^2 = u, \quad \text{and} \quad a^3 = u + c. \tag{4.5}$$

The corresponding right-eigenvector matrix is

$$R = \begin{bmatrix} 1 & 1 & 1 \\ u - c & u & u + c \\ H - uc & \frac{u^2}{2} - \frac{\chi}{\kappa} & H + uc \end{bmatrix}, \tag{4.6}$$

while its inverse can be written as

$$R^{-1} = \begin{bmatrix} \frac{1}{2}(b_1 + \frac{u}{c}) & -\frac{1}{2}(ub_2 + \frac{1}{c}) & \frac{b_2}{2} \\ 1 - b_1 & b_2 u & -b_2 \\ \frac{1}{2}(b_1 - \frac{u}{c}) & -\frac{1}{2}(ub_2 - \frac{1}{c}) & \frac{b_2}{2} \end{bmatrix}, \tag{4.7}$$

where $b_1 = \frac{1}{c^2}[\kappa u^2/2 + \chi]$ and $b_2 = \kappa/c^2$.

In order to relate the variables p and c to the independent variables ρ and ϵ, it will also be convenient to introduce the nondimensional thermodynamic variables

$$\tilde{\gamma} = 1 + \frac{p}{\rho\epsilon}, \qquad \gamma = \frac{\rho c^2}{p}. \tag{4.8}$$

For a perfect gas, these two parameters are constant and equal to each other; for an equilibrium real gas, they are both arbitrary functions of ρ and ϵ. Note that while κ, c, and γ have well-defined values, the values of χ and $\tilde{\gamma}$ depend on the choice of arbitrary constant in the definition of ϵ.

Many existing conservative schemes for the system (4.1) use forward Euler time discretization. For simplicity of presentation, this time discretization is used for discussion in this section and the scheme can be written as

$$U_j^{n+1} = U_j^n - \lambda\left[\tilde{F}_{j+\frac{1}{2}}^n - \tilde{F}_{j-\frac{1}{2}}^n\right]. \tag{4.9}$$

The vectors $\tilde{F}_{j\pm\frac{1}{2}}$ are numerical flux vectors corresponding to $h_{j\pm\frac{1}{2}}$ for the scalar case. Note that the numerical flux function $\tilde{F}_{j+\frac{1}{2}}$ should not be confused with the flux function $F_j = F(U_j)$.

An excellent brief description of the various Riemann solvers for a perfect gas can be found in Roe [7,60]. Here a short narrative discussion of the exact Riemann solver and a more detailed description of the approximate Riemann solvers and flux-vector splitting approaches are included. For a systematic discussion of the topic see LeVeque [8].

CIR Method: The earliest method for gas-dynamic equations in characteristic form was proposed by Courant et al. [47]. Their procedure, sometimes called the CIR method, is to trace back from $(j\Delta x, (n+1)\Delta t)$ all three characteristic paths. Since the problem is nonlinear, the directions of these paths are not known exactly, but to a first approximation they can be taken equal to their known directions at $(j\Delta x, n\Delta t)$. Then each characteristic equation is solved using interpolated data at time $n\Delta t$, in the interval to the left of j for characteristics with positive speed, and in the interval to the right of j for characteristics with negative speed. The resulting method is only first-order accurate. It has a principal drawback in that the scheme cannot convey a shock wave with the proper speed because it is not a conservative scheme. This method was later rediscovered by Chakravarthy et al. [132] and was renamed the split-coefficient matrix method.

Exact Riemann Solver: Godunov [49] was the first to develop the idea of advancing the solution to the next time-level by solving a set of Riemann problems. Recall that the Riemann problem for any system of conservation laws arises if initial data are prescribed as two constant states ($U = U^L$ for $x < 0$, $U = U^R$ for $x > 0$). The solution then consists of centered waves. For the one-dimensional Euler equations, the solution consists of three waves; the inner one is a contact discontinuity, and the outer ones may be shock waves or rarefaction fans. The exact solution of this problem involves only algebraic equations. See references [2,49] for details.

Let $U_{j+\frac{1}{2}}^n$ be the average state over $((j \pm \frac{1}{2})\Delta x)$ at $(n\Delta t)$. The way Godunov used the Riemann solvers was to replace the data by an approximate distribution in which the state inside each interval is uniform and equal to U_j^n. For each interface $(j + \frac{1}{2})\Delta x$, one can solve the Riemann problem with $U^L = U_j$ and $U^R = U_{j+1}$. This gives an exact solution "$U_{j+\frac{1}{2}}^{n+\frac{1}{2}}$" to the approximate problem, assuming Δt is small enough that the waves from neighboring interfaces do not intersect. The solution at $(n+1)\Delta t$ can again be approximated by a piecewise uniform distribution, and then the process can be repeated. For the Godunov method, the numerical flux is

$$\tilde{F}_{j+\frac{1}{2}}^n = F(U_{j+\frac{1}{2}}^{n+\frac{1}{2}}). \tag{4.10}$$

More recently Ben-Artzi and Falcovitz [133] generalized the exact Riemann solver of Godunov to be second-order accurate. Their theory is too complicated to be summarized here. The versatility of their method remains to be shown.

Another closely related method, devised by Glimm [134], modified by Chorin [135], and further improved by Colella [112], is the random-choice method. It represents the staggered grid solution by randomly sampling the Riemann solutions generated at the previous time-level. Recently, Toro [136], Roe and Toro [137], and Toro and Clarke [138] extended the random-choice method to be second-order accurate. Their preliminary result is very encouraging for utilizing this type of scheme in combustion-related flows.

Approximate Riemann Solvers: Since the Riemann problems arising in Godunov's method relate only to an approximation of the data, one might reasonably be satisfied with approximate solutions of the Riemann problem if these solutions still describe the important nonlinear behavior. Roe [48], Osher-Solomon [139], Harten et al. [113], Davis [140], and Einfeldt [141] proposed methods for finding such solutions. The method of Osher-Solomon is quite complicated to explain, and uses a similar number of arithmetic operations (and is not as exact) as the Godunov method. The method of Harten et al., Davis and Einfeldt does not retain all the information about every wave. Numerical experiments [140,141] showed high-resolution solutions for one-dimensional shock tube problems but more extensive numerical testing on practical application is needed before a decisive conclusion can be made. Roe's idea is closely tied with the characteristic field decomposition method which is an effective approximate local decoupling of the nonlinear system and is widely used in CFD practical applications. Therefore, only Roe's approximate Riemann solver is described here. Interested readers should refer to the original references for details. Roe's approximate Riemann solver is a linear wave decomposition in which he required that there exists an average state $\overline{U}$ which is a nonlinear function of the left and right states $U^{L,R}$ (or $U_{j+\frac{1}{2}}$ for U_j and U_{j+1}) satisfying

(1) $F(U^R) - F(U^L) = A(U^R, U^L)(U^R - U^L) = A(\overline{U})(U^R - U^L)$
(2) $A(U^R, U^L)$ has real eigenvalues and a complete set of eigenvectors
(3) $A(U, U) = A(U)$
(4) $U^R - U^L = R(U^R, U^L)\alpha = R(\overline{U})\alpha$

where the kth element of α is the strength of the kth characteristic wave (or the jump in the characteristic variables). The main feature of the method that makes it valuable for nonlinear systems is that it returns the exact solution whenever U^L and U^R lie on opposite sides of a shock wave or a contact discontinuity. For the one-dimensional Euler equations for a perfect gas, the Roe averaged state can readily be obtained as

$$\overline{u} = \frac{u^L + Du^R}{1 + D}, \tag{4.11a}$$

$$\overline{H} = \frac{H^L + DH^R}{1 + D}, \tag{4.11b}$$

$$\overline{c}^2 = (\gamma - 1)\left[\overline{H} - \frac{1}{2}\overline{u}^2\right], \tag{4.11c}$$

with

$$D = \sqrt{\frac{\rho^R}{\rho^L}}. \tag{4.11d}$$

To make use of his approximate Riemann solver, Roe applied his average state (4.11) in conjunction with his first-order upwind scheme for the nonlinear scalar case to arrive at a first-order scheme for the nonlinear system case. The numerical flux is of the form

$$\tilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F_{j+1} + F_j - R_{j+\frac{1}{2}}|\Lambda_{j+\frac{1}{2}}|\alpha_{j+\frac{1}{2}}\right], \tag{4.12a}$$

where $R_{j+\frac{1}{2}} = R(U_{j+\frac{1}{2}})$ and

$$\Lambda_{j+\frac{1}{2}} = \begin{bmatrix} a^1_{j+\frac{1}{2}} & 0 & 0 \\ 0 & a^2_{j+\frac{1}{2}} & 0 \\ 0 & 0 & a^3_{j+\frac{1}{2}} \end{bmatrix}, \tag{4.12b}$$

$$\alpha_{j+\frac{1}{2}} = R^{-1}_{j+\frac{1}{2}}(U_{j+1} - U_j). \tag{4.12c}$$

The subscript "$j+\frac{1}{2}$" denotes the average state between j and $j+1$ using relation (4.11). It can shown that the Roe-averaged Mach number $\overline{M}$ can lie outside of the range (M^L, M^R). Roberts [142] investigated the phenomenon of noise radiated from slowly moving shock waves associated with (4.12). However, this type of deficiency can easily be corrected by replacing $|\Lambda_{j+\frac{1}{2}}|$ in (4.12a) by $\mathrm{diag}(\psi(a^l_{j+\frac{1}{2}}))$ or by using the Lax-Friedrichs numerical flux as suggested in section III. See also section VI for a discussion on high speed flows.

Flux-Vector Splittings: The simplest way of introducing upwinding into systems of hyperbolic conservation laws seems to be based on the representation of the flux vector $F(U)$ as the sum of $F^-(U) + F^+(U)$ such that one can apply forward- and backward-differencing to the Jacobian matrices $\frac{\partial F^-}{\partial U}$ and $\frac{\partial F^+}{\partial U}$, respectively. This would be equivalent to using an approximate Riemann solution in which the numerical flux $\tilde{F}_{j+\frac{1}{2}}$ is of the form

$$\tilde{F}_{j+\frac{1}{2}} = F^+(U^L_{j+\frac{1}{2}}) + F^-(U^R_{j+\frac{1}{2}}). \tag{4.13}$$

This numerical flux amounts to requiring that the Jacobian matrices $\frac{\partial F^-}{\partial U}$ and $\frac{\partial F^+}{\partial U}$ have no positive and no negative eigenvalues, respectively. A popular way is to split the flux according to the characteristic speeds $(u, u \pm c)$.

This idea, assuming that the flux is homogeneous of degree one (for thermally perfect gases), seems to have been first used in the context of astrophysical gas dynamics [143], and to have been rediscovered with a fuller mathematical development by Steger-Warming [114]. Note that Steger-Warming make use of the flux-vector splitting for a second-order upwind scheme in a non-MUSCL way and without the use of limiters (non-TVD method). Their final second-order upwind scheme cannot be represented in the form (4.13). Mulder and van Leer [144] and Anderson et al. [77] introduced the MUSCL approach with limiters into the Steger-Warming flux-vector splitting (TVD method, hereafter referred to as TVD flux vector-splitting method) and its numerical fluxes have the form (4.13). The MUSCL approach with limiters (TVD) formulation can produce a better shock resolution than the non-MUSCL and non-TVD flux-vector splitting approaches.

In 1982, van Leer [115] devised an alternative splitting for a perfect gas such that there are noticeably better results around sonic points and sharper shock transitions than can be obtained with the Steger-Warming splitting. For the derivation, see the original paper for details. In the next few sections, two generalizations of these flux-vector splittings and a generalization of Roe's average to equilibrium real gases will be described. In these formulations, the perfect-gas version is included as a particular case.

4.2. Description of the Riemann Solvers for Equilibrium Real Gases

Extensions of the exact Riemann solver of Godunov to certain types of equilibrium real gases have been obtained by Colella and Glaz [145], Dukowicz [146], Ben-Artzi and Falcovitz [147] and in a more general setting by Liu [148]. The derivations are quite involved and interested readers should refer to their papers for details. Efficient extensions of the Osher-Solomon Riemann solver for equilibrium real gases are not known at this point. The recent generalization of Roe's average by Vinokur [116], and generalization of Steger-Warming and van Leer flux-vector splittings by Vinokur and Montagné [117] will be discussed in this section. Comparison among the various generalization will also be included. These generalizations are also simpler to use than the exact Riemann solvers of references [145-147].

The following three subsections (4.2.1 - 4.2.3) were written by M. Vinokur of Sterling Software, Palo Alto, California. They are summaries of his papers [116,117,149].

4.2.1. An Approximate Riemann Solver (Generalized Roe Average [116])

Among the various approximate Riemann solvers [48,113-115] for a perfect gas, the most common one uses the Roe average [48] because of its simplicity and its ability to satisfy the jump conditions. Surprisingly, an exact definition of a Roe average for equilibrium real gases not only exists but is actually not unique. A number of different ways of obtaining a unique solution have been proposed. Due to the arbitrary nature of the equation of state, it is difficult to evaluate the various solutions. The amount of computation involved in the various solutions is also a function of the equation of state subroutine. Thus different formulas could prove to be more efficient when used with different equation of state routines.

While an exact definition of a Roe average for equilibrium real gases will be given below, its implementation is generally not practical. Various approximations to the Roe average for equilibrium real gases have been proposed [51,99,113,116,150-154]. The exact and approximate formulae of Vinokur [116] will be given here, whereas the others can be found in the appropriate references. The reason for choosing Vinokur's formulation is that his derivation is more systematic than those of other investigators. Comparisons with the formulae of other investigators will also be indicated. Generalizations of the Roe average for multidimensional flows are found in Vinokur [116]. The use of the approximate Riemann solver in conjunction with the numerical schemes will be discussed in sections 4.4 and 4.5.

The flux at a point separating two states U^L and U^R is based on the eigenvalues and eigenvectors of some average $\bar{A}$. The optimum choice for $\bar{A}$ is one satisfying

$$\Delta F = \bar{A} \Delta U, \tag{4.14}$$

where $\Delta(\cdot) = (\cdot)^R - (\cdot)^L$. This choice of $\bar{A}$ will capture discontinuities exactly. One way of obtaining $\bar{A}$ is to seek an average state $\bar{U}$, such that

$$\bar{A} = A(\bar{U}). \tag{4.15}$$

The notation $\bar{U}$ implies only those variables that appear explicitly in equation (4.15). Such a state is known as a Roe-averaged state. Expressions for a perfect gas were first devised by Roe [48] and are given by equation (4.11).

The entries in A depend explicitly on the thermodynamic variables h, χ, and κ, as well as the velocity u. Since the density is not explicitly required, one would expect the Roe-averaged state to depend on ρ^L and ρ^R through their ratio only. It is therefore convenient to define the parameter $D = \sqrt{\rho^R/\rho^L}$. We first examine the second component of equation (4.14). The average velocity $\bar{u}$ must be a linear combination of u^L and u^R. Recalling that u^L and u^R can be independently prescribed, we can readily establish the same $\bar{u}$ as in equation (4.11a) for a perfect gas. This definition will satisfy all the terms involving the velocity. Note that $\bar{u}$ always lies between u^L and u^R. The remaining terms in the equation result in the new condition

$$\bar{\chi}\Delta\rho + \bar{\kappa}\Delta\tilde{\epsilon} = \Delta p. \tag{4.16}$$

This is just the discrete form of equation (4.2a), averaged between the two states. This last condition is automatically satisfied for a perfect gas.

In order to satisfy the third component of equation (4.14) we also require $\bar{H}$ to have the same form as the perfect-gas version (equation (4.11b)). Using the definition of H, equations (4.11a) and (4.11b) can be combined to define the Roe-averaged specific enthalpy as

$$\bar{h} = \frac{h^L + Dh^R}{1 + D} + \frac{D}{2(1 + D)^2}(\Delta u)^2. \tag{4.17}$$

Note that $\bar{h}$ could lie outside the range given by h^L and h^R if Δu is sufficiently large. The Roe-averaged sound speed is given by (4.3) as

$$\bar{c}^2 = \bar{\chi} + \bar{\kappa}\bar{h}. \tag{4.18}$$

For a perfect gas, equations (4.11a), (4.11b), (4.17) and (4.18) are sufficient to define uniquely $\bar{a}^l$, $\bar{R}$, and $\bar{R}^{-1}$, since $\bar{\chi} = 0$ and $\bar{\kappa}$ is a given constant.

For an equilibrium real gas, equation (4.16) provides only one relation for the variables $\bar{\chi}$ and $\bar{\kappa}$. We thus have the paradoxical situation that not only does a Roe-averaged state exist for an equilibrium real gas, but also its precise value is not uniquely defined. For the special case in which states L and R are precisely those that satisfy the jump conditions across a discontinuity, equations (4.11a) through (4.18) are consistent with the exact Riemann solver, even though $\bar{\chi}$ and $\bar{\kappa}$ are not uniquely defined. For a shock wave one obtains

$$\bar{h} = \frac{\Delta\tilde{\epsilon}}{\Delta\rho} = \frac{h^L + D^2 h^R}{1 + D^2} \tag{4.19a}$$

and

$$\bar{c}^2 = \frac{\Delta p}{\Delta\rho}. \tag{4.19b}$$

For the special case of a thermally perfect gas, c^2 is a function of h only, and one can readily show that the values of $\bar{h}$ and $\bar{c}^2$ given by equations (4.19a) and (4.19b) can only satisfy this gas law if the function is linear. But this is precisely the definition of a perfect gas.

It is clear that unique values of $\bar{\chi}$ and $\bar{\kappa}$ must be defined in terms of the thermodynamic states L and R. Accurate numerical calculations for air [155] show that $\bar{\chi}$ and $\bar{\kappa}$ can have a nonmonotonic behavior if the states L and R are far apart. One way to obtain unique values of $\bar{\chi}$ and $\bar{\kappa}$ is to integrate equation (4.2a) along the straight-line path between states L and R in the ρ-$\tilde{\epsilon}$ plane. This results in the expressions

$$\bar{\chi} = \int_0^1 \chi[\rho(\tau), \tilde{\epsilon}(\tau)]\, d\tau, \tag{4.20a}$$

$$\bar{\kappa} = \int_0^1 \kappa[\rho(\tau), \tilde{\epsilon}(\tau)]\, d\tau, \tag{4.20b}$$

where

$$\rho(\tau) = \rho^L + \tau\, \Delta\rho, \tag{4.20c}$$

$$\tilde{\epsilon}(\tau) = \tilde{\epsilon}^L + \tau\, \Delta\tilde{\epsilon}. \tag{4.20d}$$

Equations (4.20a) to (4.20d) give unique definitions of $\bar{\chi}$ and $\bar{\kappa}$ satisfying equation (4.16) for arbitrary values of $\Delta\rho$ and $\Delta\tilde{\epsilon}$, including the limiting case $\Delta\rho = \Delta\tilde{\epsilon} = 0$. From equations (4.17) and (4.18) one sees that even if $\Delta u = 0$, $\bar{c}^2$ is not equal to the integrated average of c^2. Since h is a smoothly varying function (compared to χ and κ), it is reasonable to expect that $\bar{c}^2$ will always be positive.

Given an equation of state, or some interpolation representation (such as in reference [155]), the integrals in equations (4.20a) and (4.20b) can be evaluated for any two end states L and R. Since the exact evaluation may not be practical, some approximate quadratures may be required. Let $\hat{\chi}$ and $\hat{\kappa}$ be approximations to equations (4.20a) and (4.20b). They will not satisfy equation (4.16) exactly. One therefore requires values of $\bar{\chi}$ and $\bar{\kappa}$ satisfying equation (4.16) which are closest to $\hat{\chi}$ and $\hat{\kappa}$. This can be formulated geometrically by projecting the point $\hat{\chi}$, $\hat{\kappa}$ onto the straight line defined by equation (4.16). But in order for the Roe-averaged state to be independent of the arbitrary constant in the definition of ϵ, one must first recast the problem so that geometric relationships will not be affected by the choice of this constant. This can be accomplished if one first divides equation (4.16) by $\bar{\kappa}$. The slope of the straight line for the variables $1/\bar{\kappa}$ and $\bar{\chi}/\bar{\kappa}$ is now given by Δp and $\Delta\rho$, both of which are uniquely defined by states L and R. A further scale factor $\hat{s}$ with the dimension of $\bar{\chi}$ must be introduced, since $\bar{\chi}$ is not dimensionless. A natural choice for the scale factor $\hat{s}$ is an average value of c^2. In terms of arithmetic averages, which is equivalent to a trapezoidal quadrature approximation to the integrals, one obtains the expressions

$$\bar{\chi} = \frac{b\hat{\chi} + \hat{s}^2\, \Delta\rho\, \delta p}{b - \Delta p\, \delta p}, \tag{4.21a}$$

$$\bar{\kappa} = \frac{b\hat{\kappa}}{b - \Delta p\, \delta p}, \tag{4.21b}$$

where

$$\delta p = \Delta p - \widehat{\chi}\,\Delta\rho - \widehat{\kappa}\,\Delta\widetilde{\epsilon} \tag{4.21c}$$

$$b = (\widehat{s}\,\Delta\rho)^2 + (\Delta p)^2, \tag{4.21d}$$

$$\widehat{\chi} = (\chi^L + \chi^R)/2, \tag{4.21e}$$

$$\widehat{\kappa} = (\kappa^L + \kappa^R)/2, \tag{4.21f}$$

and

$$\widehat{s} = \widehat{c^2} = [(c^2)^L + (c^2)^R]/2. \tag{4.21g}$$

Note that if $\Delta p = \Delta\rho = 0$, equations (4.21a) and (4.21b) are replaced by $\bar{\chi} = \chi^L$ and $\bar{\kappa} = \kappa^L$. Other approximate formulas for $\widehat{\chi}$, $\widehat{\kappa}$, and $\widehat{s}$ may be found in reference [116]. Note that the earlier expressions for $\bar{\chi}$ and $\bar{\kappa}$ given by equations (6.1a) and (6.1b) of reference [150] are based on a projection onto the straight line defined by equation (4.16), and the choice h for the scale factor $\widehat{s}$. The resultant values are not independent of the arbitrary constant in the definition of ϵ, and also depend on Δu. These expressions therefore do not have as general a validity as those given by equations (4.21a) to (4.21g). The numerical experiments reported in reference [150] show no significant differences between the two expressions. The particular cases discussed in section 4.4 actually used modifications of the average of Carofano [99] given by equations (6.2a) to (6.2c) of reference [150]. Even though this average, as well as the arithmetic average suggested by Huang [51], does not satisfy the Roe condition, no significant differences among the various averages were observed. Also note that the expressions for $\bar{\chi}$ and $\bar{\kappa}$ given by equations (3.17a) and (3.17b) of reference [150] and also quoted in references [151] and [1] differ from the expressions given by equations (4.21a) to (4.21g), since the projection onto the straight line was defined in terms of $\widehat{1/\kappa}$ and $\widehat{\chi/\kappa}$, instead of $1/\widehat{\kappa}$ and $\widehat{\chi}/\widehat{\kappa}$.

The column vector $\bar{\alpha}$ can be expressed simply in term of Δp, $\Delta\rho$ and Δu as

$$\bar{\alpha} = \begin{bmatrix} \Delta\rho - \Delta p/\bar{c}^2 \\ \frac{1}{2}(\Delta p/\bar{c}^2 + \bar{\rho}\,\Delta u/\bar{c}) \\ \frac{1}{2}(\Delta p/\bar{c}^2 - \bar{\rho}\,\Delta u/\bar{c}) \end{bmatrix}, \tag{4.22a}$$

where one *formally* defines

$$\bar{\rho} = \sqrt{\rho^L \rho^R}. \tag{4.22b}$$

These expressions have the same form as for the perfect gas case. The quantities $\bar{\chi}$ and $\bar{\kappa}$ do not appear explicitly. One can show that unlike the perfect gas, $\bar{c}^2$ in (4.18) (depending on the formula) could be negative if the two states are far apart and $\widehat{\chi}$ and $\widehat{\kappa}$ are poorly approximated. The same situation holds true for the nonequilibrium flows to be described in section VII.

Comparison Among the Various Generalizations: The expressions for $\bar{A}$ given by references [51,99,113] do not satisfy the Roe condition (4.14). In the formulas for the Roe average derived by Grossman and Walters [152], γ and $\bar{\gamma}$ are introduced to relate c^2 to h using equation

(4.8). Their final expressions, which do not satisfy equation (4.14) exactly, are equivalent to setting $\chi = 0$, with $\bar{\gamma}$ and γ replaced by their arithmetic averages. The calculations of reference [155] show that setting $\chi = 0$ may not be justified. The simple arithmetic average may also not be justified if the two states are far apart.

Glaister [153] employs ϵ instead of $\tilde{\epsilon}$ in equations (4.2), and introduces $\bar{\rho}$ from equation (4.22b) as well as $\bar{\epsilon}$ given by a relation analogous to equation (4.11a) in order to obtain equation (4.16) in terms of $\Delta\epsilon$. He obtains his $\bar{\chi}$ and $\bar{\kappa}$ by averaging the results of integrating his equation (4.2a) along two piecewise straight-line paths, one passing through the state ρ^L, ϵ^R and the other passing through the state ρ^R, ϵ^L. The results can be simply expressed in terms of the values of p at the four states. Actually, Glaister's results do not reduce to Roe's average state solution for a perfect gas. This can be corrected by employing a weighted average of the results along the two paths in a manner similar to equation (4.11a). If the states L and R are far apart, and the equation of state is non-convex, the introduction of the two fictitious states could give poor results.

Liou et al. [154] also employ ϵ instead of $\tilde{\epsilon}$, and use the state $\bar{\rho}, \bar{\epsilon}$ defined by Glaister to calculate their approximate $\hat{\chi}$ and $\hat{\kappa}$. One can show that equation (4.16) in terms of $\Delta\epsilon$ can be satisfied by integrating the analogous equation (4.2a) along a piecewise straight-line path passing through the state $\hat{\rho}$, $\hat{\epsilon}$. Their approximate $\hat{\chi}$, $\hat{\kappa}$ thus corresponds to a one term quadrature approximation. Their method for obtaining the values of their $\bar{\chi}$ and $\bar{\kappa}$ satisfying equation (4.16) which are closest to their $\hat{\chi}$ and $\hat{\kappa}$ differs from that used by Vinokur [116]. Their formulae break down when either $\Delta\rho$ or $\Delta\epsilon$ approaches zero.

4.2.2. Generalized Steger-Warming Flux-Vector Splitting [117]

For subsonic flow, the eigenvalues of A are of mixed sign. In flux vector-splitting methods, the flux F is divided into several parts, each of which has a Jacobian matrix whose eigenvalues are all of one sign. The flux-vector splitting approach of Steger-Warming [114] made use of the assumption that the Euler equations are homogeneous of degree one. This homogeneity property is satisfied for a thermally perfect gas. While they only gave results for a perfect gas, the results for a thermally perfect gas are found in reference [117]. Various generalizations of the Steger-Warming flux-vector splitting for an arbitrary equilibrium real gas have been proposed [117,150-152,154,156,157]. The formulae of Vinokur and Montagné [117] will be given here, whereas the others can be found in the appropriate references. Comparisons with the formulae of other investigators will also be indicated.

The generalization to an arbitrary equilibrium real gas presented here makes direct use of the eigenvalues of A, and is an extension of the work of Sanders and Prendergast [143] and Montagné [157]. The basic idea is the observation that the eigenvalues are actually three velocities. One can associate with each eigenvalue a^l, $l = 1,2,3$, a stream with velocity a^l, and some unknown density ρ^l and specific internal energy ϵ^l. Each stream can then be characterized by the column vector

$$U^l = \begin{bmatrix} \rho^l \\ m^l \\ e^l \end{bmatrix}, \tag{4.23}$$

where $m^l = \rho^l a^l$ and $e^l = \rho^l[\epsilon^l + \frac{1}{2}(a^l)^2]$. The fictitious flux associated with each stream is assumed to be convective only, namely,

$$F^l = a^l U^l. \tag{4.24}$$

The six unknowns are determined from the conditions

$$U = \sum_{l=1}^{3} U^l \qquad \text{and} \qquad F = \sum_{l=1}^{3} F^l. \tag{4.25}$$

Since the second component of U and the first component of F give the same equation, only one degree of freedom remains. From the first two components of U and F one readily obtains

$$\rho^1 = \rho^3 = \frac{\rho}{2\gamma} \qquad \text{and} \qquad \rho^2 = \rho(1 - \frac{1}{\gamma}). \tag{4.26a}$$

The third components of U and F result in the relations

$$\epsilon^1 = \epsilon^3 \qquad \text{and} \qquad \sum_{l=1}^{3} \rho^l \epsilon^l = \rho\left(\epsilon - \frac{c^2}{2\gamma}\right). \tag{4.26b}$$

The relations in equations (4.26a) and (4.26b) have the same form as those for the Steger-Warming solution for a thermally perfect gas. In choosing the manner in which the remaining degree of freedom is expressed, we are guided by the fact that ϵ is not an absolute quantity. Consequently, we express the ϵ^i as

$$\epsilon^2 = \epsilon - \frac{(1-\alpha)c^2}{\gamma(\gamma - 1)}, \qquad \text{and} \qquad \epsilon^1 = \epsilon^3 = \epsilon + \left(\frac{1-\alpha}{\gamma} - \frac{1}{2}\right)c^2. \tag{4.26c}$$

Here the non-dimensional thermodynamic parameter α (not to be confused with $\alpha = R^{-1}(\overline{U})(U^R - U^L)$) has been chosen so that $\alpha = 0$ corresponds to the Steger-Warming solution for a thermally perfect gas. The final expressions for the F^l can be written in the form

$$F^2 = \frac{\rho(\gamma - 1)}{\gamma} \begin{bmatrix} a^2 \\ (a^2)^2 \\ \frac{(a^2)^3}{2} + \left[\epsilon - \frac{(1-\alpha)c^2}{\gamma(\gamma-1)}\right] a^2 \end{bmatrix}, \tag{4.27a}$$

and

$$F^l = \frac{\rho}{2\gamma} \begin{bmatrix} a^l \\ (a^l)^2 \\ \frac{(a^l)^3}{2} + \left[\epsilon + \left(\frac{1-\alpha}{\gamma} - \frac{1}{2}\right)c^2\right] a^l \end{bmatrix} \tag{4.27b}$$

for $l = 1$ or 3. We thus have a one-parameter family of flux-vector splittings, where α can be an arbitrary parameter. The total flux can be split according to the sign of the eigenvalues. For $-c < u < 0$ we, therefore, have

$$F^+ = F^3 \qquad \text{and} \qquad F^- = F^1 + F^2 \tag{4.28a}$$

and for $0 < u < c$

$$F^+ = F^2 + F^3 \qquad \text{and} \qquad F^- = F^1. \tag{4.28b}$$

For a perfect gas one can show that when $-c < u < 0$, the determinant of the split-flux Jacobian $A^+ = \partial F^+/\partial U$ is

$$\det(A^+) = \frac{(c+u)^3}{16\gamma^2}\left[3 - \gamma - 2\alpha\frac{(\gamma-1)}{\gamma}\right]. \tag{4.29a}$$

The determinant is the product of the three eigenvalues, and a necessary condition for A^+ to have eigenvalues that are all positive is that $\det(A^+) > 0$. It follows from equation (4.29a) that we must take $\alpha < \gamma(3-\gamma)/\left[2(\gamma-1)\right]$. For the region $0 < u < c$, one can show that the minimum of $\det(A^+)$ occurs when u approaches zero. This minimum value is

$$\det(A^+) = \frac{c^3}{16\gamma^3}\left[5 - 3\gamma - 2\alpha\frac{(3-\gamma)(\gamma-1)}{\gamma}\right]. \tag{4.29b}$$

The condition $\det(A^+) > 0$ requires that $\alpha < \gamma(5-3\gamma)/\left[2(3-\gamma)(\gamma-1)\right]$. For $1 < \gamma < 3$, this second condition on α automatically satisfies the previous inequality derived for $-c < u < 0$. In order to determine an optimum value of α, we note that in general the three eigenvalues of A^+ are discontinuous at $-c$, 0, and c. For $\alpha = 0$, two of the eigenvalues are continuous at $u = c$, and one can further show that all the eigenvalues are real and non-negative in the subsonic region when $\gamma < 5/3$. For a thermally perfect gas with vibrational excitation, one can show that all the eigenvalues remain real and non-negative when $\alpha = 0$. Because of the general nature of the variation of γ and ϵ, it is simplest to set $\alpha = 0$ for an arbitrary equilibrium real gas. The numerical results presented in section 4.4 were actually based on an earlier parameterization which was chosen to reduce to the perfect-gas solution. It corresponds to the present $\alpha = (\bar{\gamma} - \gamma)/(\bar{\gamma} - 1)$. The cases were also rerun with $\alpha = 0$, and no noticeable difference was observed.

Comparison Among the Various Generalizations: The generalization of Steger-Warming flux-vector splitting for an equilibrium real gas given by Palmer [156] and Grossman and Walters [152] is obtained by replacing γ with $\bar{\gamma}$ in Eqs. (4.8), (4.27a), and (4.27b), and setting $\chi = \alpha = 0$. Since accurate numerical calculations for air [155] show that γ and $\bar{\gamma}$ can differ by more than 20%, the effective sound speed can differ from the true sound speed by more than 10%. This could give rise to errors in transonic regions. The formula of Montagné [157] is based on the analysis of this paper, except that the arbitrary parameter is chosen by assuming that the internal energy per unit mass of each stream is the same. This is equivalent to setting $\alpha = (3-\gamma)/2$, and does not reduce to the Steger-Warming flux-vector splitting for a perfect gas. In the pseudo-splitting formula of Liou et al. [154], the homogeneous term AU is treated in the manner of Steger-Warming, and the inhomogeneous term $F - AU$ is treated in a central difference manner. The solution explicitly involves the additional parameter κ and is, therefore, not as simple a form of flux-vector splitting. It also appears that their split fluxes are discontinuous at $u = c$ and $u = -c$ for an arbitrary gas law.

4.2.3. Generalized van Leer Flux-Vector Splitting [117]

In a different approach, van Leer [115] constructed a flux-vector splitting for a perfect gas in terms of low-order polynomials of u, which gives continuous eigenvalues at $u = 0$ and $u = \pm c$. The splitting also has the desirable property that one of the eigenvalues of the split-flux Jacobians is identically zero. This results in a sharper capture of transonic shocks. Various generalizations of van Leer's flux-vector splitting for an arbitrary equilibrium real gas have been proposed [117,150-152,154,158]. The formulae of Vinokur and Montagné [117] will be given here, whereas the others can be found in the appropriate references. Due to the arbitrary nature of a real-gas law, the condition of one eigenvalue being identically zero cannot be satisfied exactly. Comparisons with the formulae of other investigators will also be indicated.

For $|u| < c$, the continuity requirements necessitate a factor $(u \pm c)^2$ in the formulas for $F^{\pm}$. The expressions for the first two components of $F^{\pm}$ that are given by the lowest-order polynomials in u are readily found to be

$$F^{\pm}_{mass} = \pm \frac{\rho}{4c}(u \pm c)^2 \tag{4.30a}$$

and

$$F^{\pm}_{mom} = \frac{\rho}{\gamma 4c}(u \pm c)^2 \left[2c \pm (\gamma - 1)u\right]. \tag{4.30b}$$

For a perfect gas they are identical to those derived by van Leer. The expression for the third component, which satisfies continuity and symmetry conditions, can be written most generally in the form

$$\begin{aligned} F^{\pm}_{ene} = &\pm \frac{1}{2(\gamma^2 - 1)} \frac{\rho}{4c}(u \pm c)^2 \left[2c \pm (\gamma - 1)u\right]^2 \pm \frac{\rho}{4c}(u \pm c)^2 \left[\epsilon - \frac{c^2}{\gamma(\gamma - 1)}\right] \\ &\pm \beta \frac{\rho}{4c}(u \pm c)^2 (u \mp c)^2, \end{aligned} \tag{4.30c}$$

where the arbitrary parameter β has been chosen so that $\beta = 0$ corresponds to van Leer's solution for a perfect gas. We again have a one-parameter family of flux-vector splittings. Note that when $\beta = 1/(\gamma + 1)$, $F^{\pm}_{ene} = F^{\pm}_{mass}H$. This choice of flux-vector splitting, which preserves the total enthalpy, was recommended for high Mach number flows by Hänel et al. [159].

For a perfect gas, the second term in $F^{\pm}_{ene}$ vanishes. Van Leer's condition of a vanishing eigenvalue for $A^{\pm}$ requires β to vanish also, so that $F^{\pm}_{ene}$ reduces to one term. One can readily show that the remaining eigenvalues are both of the proper sign for $1 \leq \gamma \leq 3$. For an equilibrium real gas, one requires at least two terms for $F^{\pm}_{ene}$. Because of the general nature of the variations of γ and ϵ, it is impossible to obtain the vanishing eigenvalue condition identically throughout the velocity range for any choice of β.

One can demonstrate readily that for a thermally perfect gas, the two-term solution has one eigenvalue that is of the wrong sign for the whole subsonic velocity range. Fortunately, the magnitude of the offending eigenvalue is extremely small, so that its effect on a numerical scheme is not noticeable. In view of this fact and because of the general nature of the variation of γ and ϵ, it is simplest to set $\beta = 0$ for an equilibrium real gas. The results presented in section 4.4 are obtained with this value of β and seem to show that this approximation is valid for a wide range of flow conditions.

Comparison Among the Various Generalizations: The generalization of van Leer flux-vector splitting given by Grossman and Walters [152] is again based on replacing γ by $\tilde{\gamma}$, and does not use the true sound speed. Since $\tilde{\gamma}$ is a variable, van Leer's condition of one zero eigenvalue can also not be obtained exactly in their formulation. For a thermally perfect gas, they also obtain one eigenvalue of the wrong sign. The formula of Montagné [158] differs from Eq. (4.30c), involving both γ and $\bar{\gamma}$ in the first term, with the second term absent. His solution is internally inconsistent, since it does not satisfy the condition $F = F^+ + F^-$. The formula of Liou et al. [154] is obtained from Eq. (4.30c) by setting $\beta = (\bar{\gamma} - \gamma)/(\gamma + 1)[\gamma(\bar{\gamma} - 1) + 2\bar{\gamma}]$. This does not satisfy the principle that the solution should be independent of the arbitrary constant in the definition of ϵ.

4.3. Extension to Systems via the Local-Characteristic Approach

In this section, the method sometimes referred to as the "local characteristic approach" (a generalization of Roe's approximate Riemann solver) in conjunction with TVD schemes is discussed. Consider a system of hyperbolic conservation laws of the form (4.1a) where U and F are vectors of m components. The idea of the local-characteristic approach is to extend the scalar TVD method to systems so that the resulting scheme is TVD for the "locally" frozen constant-coefficient system. The procedure is to define at each point a local system of characteristic variables W and to obtain a system of uncoupled scalar equations

$$W_t + \Lambda W_x = 0, \qquad W = R^{-1}U \tag{4.31a}$$

$$\Lambda = \mathrm{diag}(a^l). \tag{4.31b}$$

The matrix R can be (4.6) for the Euler equations, and is a transformation matrix such that Λ is diagonal. One then applies the nonlinear scalar scheme to each of the scalar characteristic equations. The final form after transforming back to the original variables looks like the scalar case except there is coupling between the characteristic variables through the eigenvectors R. The numerical flux for the corresponding explicit scheme (3.34) or implicit scheme (3.50) can be written in the form

$$\widetilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F_j + F_{j+1} + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\right], \tag{4.32a}$$

for the non-MUSCL formulation and

$$\widetilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F(U^R_{j+\frac{1}{2}}) + F(U^L_{j+\frac{1}{2}}) + \widehat{R}_{j+\frac{1}{2}}\widehat{\Phi}_{j+\frac{1}{2}}\right] \tag{4.32b}$$

for the MUSCL formulation. The matrix $R_{j+\frac{1}{2}}$ is R evaluated at some symmetric average of U_j and U_{j+1}. For example, $R_{j+\frac{1}{2}} = R(\frac{U_{j+1}+U_j}{2})$. Other approximate ways of obtaining symmetric averages are the Roe average for a perfect gas and its various generalizations to equilibrium real gases as discussed in sections 4.1 and 4.2. The lth element of $\Phi_{j+\frac{1}{2}}$, denoted by $\phi^l_{j+\frac{1}{2}}$ for the lth characteristic, has the same form as the scalar case except that $a_{j+\frac{1}{2}}$ is replaced by $a^l_{j+\frac{1}{2}}$ and $\Delta_{j+\frac{1}{2}}$ is replaced by $\alpha^l_{j+\frac{1}{2}}, l = 1, \ldots, m$. Here $\alpha^l_{j+\frac{1}{2}}$ are elements of $\alpha_{j+\frac{1}{2}} = R^{-1}_{j+\frac{1}{2}}(U_{j+1} - U_j)$. Specific forms of the $\phi^l_{j+\frac{1}{2}}$ for explicit and implicit methods will be discussed in sections 4.4 and 4.5. Here $\widehat{R}_{j+\frac{1}{2}}$ and $\widehat{\Phi}_{j+\frac{1}{2}}$ have the same meaning as (4.32a) except that the variables are in terms of $(U^R_{j+\frac{1}{2}}, U^L_{j+\frac{1}{2}})$ instead of (U_{j+1}, U_j)

The local-characteristic approach is more efficient than the exact [49] or approximate Riemann solvers of Osher-Solomon [139], and it provides a natural way to linearize the implicit TVD schemes [21,74]. The advantages of this approach as opposed to Davis' simplified approach [8] or the Boris and Book approach [67] to systems are that (a) the current approach in effect uses scalar schemes on each characteristic field so that the limiter used need not be the same for each field (e.g., one can use a more compressive limiter for the linear fields and use a less compressive limiter for the nonlinear fields as in the numerical examples of references [21,75,160-162]) and (b) one can even use different schemes for different fields. For the one-dimensional Euler equations, the characteristic fields consist of two nonlinear fields $u \pm c$ and a linear field u. The contact discontinuities are associated with the linear fields. It has been shown [21,75,160-162] that the two different fields required different amounts of numerical dissipation (i.e., different limiters). Often the limiters that are designed for the linear field might give spurious oscillation or nonphysical solutions for the nonlinear field. Numerical examples concerning this aspect will be discussed in subsections 4.4.3 and 4.4.4, and in section V.

4.4. Description of the Explicit Numerical Algorithms and Examples

The second-order in time and second- or third-order in space explicit-difference schemes considered here for both the MUSCL and the non-MUSCL approaches for the system case can be written in the same form as equation (4.9); i.e. $U_j^{n+1} = U_j^n - \lambda[\tilde{F}^n_{j+\frac{1}{2}} - \tilde{F}^n_{j-\frac{1}{2}}]$. These nonlinear scalar schemes as discussed in section III were purposely written in a more convenient form for the generalization to systems via the local characteristic approach. The various methods can be implemented into a new or existing computer code with minimum effort. Also the various methods fall nicely into the present framework and a computer code can be implemented with the various methods as separate modules. These schemes are written in a manner such that there is a smaller operations count than other equivalent forms for their scalar counterparts. For scalar problems, the difference in operations count between any two algorithms may be very small and yet the operations counts for their system counterparts might be vastly different. The situation occurs even though one starts with a different yet equivalent representation for the scalar case. More detail on this matter can be found in Yee [39] or section 4.6.

4.4.1. The Non-MUSCL Approach

The numerical flux functions $\tilde{F}_{j\pm\frac{1}{2}}$ for a non-MUSCL-type approach for both the upwind and symmetric TVD schemes [21,75] using the local-characteristic approach are given by equation (4.32); i.e., $\tilde{F}_{j+\frac{1}{2}} = \frac{1}{2}[F_j + F_{j+1} + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}]$.

Second-Order Symmetric TVD Scheme: The elements of the vector $\Phi_{j+\frac{1}{2}}$, denoted by $(\phi^l_{j+\frac{1}{2}})^S$, for a general second-order symmetric TVD scheme are

$$(\phi^l_{j+\frac{1}{2}})^S = -\lambda(a^l_{j+\frac{1}{2}})^2\hat{Q}^l_{j+\frac{1}{2}} - \psi(a^l_{j+\frac{1}{2}})[\alpha^l_{j+\frac{1}{2}} - \hat{Q}^l_{j+\frac{1}{2}}]. \tag{4.33a}$$

The value $a^l_{j+\frac{1}{2}}$ is the characteristic speed a^l evaluated at some symmetric average of U_j and U_{j+1}. The function ψ is an entropy correction to $|z|$. It can be the same as (3.18), which is repeated here for completeness:

$$\psi(z) = \begin{cases} |z| & |z| \geq \delta_1 \\ (z^2 + \delta_1^2)/2\delta_1 & |z| < \delta_1 \end{cases}. \tag{4.33b}$$

For the test problems containing only unsteady shocks to be shown, δ_1 is set to zero in most of the computations. Note that entropy-violating phenomena occur only for steady or nearly steady shocks. For steady-state problems containing strong shock waves, a proper control of the size of δ_1 is very important, especially for hypersonic blunt-body flows. A discussion is given in section 5.7. The 'limiter' function $\widehat{Q}^l_{j+\frac{1}{2}}$, expressed in terms of characteristic variables, can be of the form

$$\widehat{Q}^l_{j+\frac{1}{2}} = \text{minmod}(\alpha^l_{j-\frac{1}{2}}, \alpha^l_{j+\frac{1}{2}}) + \text{minmod}(\alpha^l_{j+\frac{1}{2}}, \alpha^l_{j+\frac{3}{2}}) - \alpha^l_{j+\frac{1}{2}}, \tag{4.33c}$$

$$\widehat{Q}^l_{j+\frac{1}{2}} = \text{minmod}(\alpha^l_{j-\frac{1}{2}}, \alpha^l_{j+\frac{1}{2}}, \alpha^l_{j+\frac{3}{2}}), \tag{4.33d}$$

$$\widehat{Q}^l_{j+\frac{1}{2}} = \text{minmod}\left[2\alpha^l_{j-\frac{1}{2}}, 2\alpha^l_{j+\frac{1}{2}}, 2\alpha^l_{j+\frac{3}{2}}, \frac{1}{2}(\alpha^l_{j-\frac{1}{2}} + \alpha^l_{j+\frac{3}{2}})\right]. \tag{4.33e}$$

The minmod function has the same meaning as in the scalar case.

Second-Order Upwind TVD Scheme: The elements of the vector $\Phi_{j+\frac{1}{2}}$ for a second-order upwind TVD scheme, denoted by $(\phi^l_{j+\frac{1}{2}})^U$, originally developed by Harten and later modified and generalized by Yee [74,75,163], are

$$(\phi^l_{j+\frac{1}{2}})^U = \sigma(a^l_{j+\frac{1}{2}})(g^l_{j+1} + g^l_j) - \psi(a^l_{j+\frac{1}{2}} + \gamma^l_{j+\frac{1}{2}})\alpha^l_{j+\frac{1}{2}}. \tag{4.34a}$$

The function $\sigma(z) = \frac{1}{2}[\psi(z) - \lambda z^2]$ and

$$\gamma^l_{j+\frac{1}{2}} = \sigma(a^l_{j+\frac{1}{2}}) \begin{cases} (g^l_{j+1} - g^l_j)/\alpha^l_{j+\frac{1}{2}} & \alpha^l_{j+\frac{1}{2}} \neq 0 \\ 0 & \alpha^l_{j+\frac{1}{2}} = 0 \end{cases}. \tag{4.34b}$$

Examples of the 'limiter' function g^l_j can be expressed as

$$g^l_j = \text{minmod}(\alpha^l_{j-\frac{1}{2}}, \alpha^l_{j+\frac{1}{2}}) \tag{4.34c}$$

$$g^l_j = \left(\alpha^l_{j+\frac{1}{2}}\alpha^l_{j-\frac{1}{2}} + |\alpha^l_{j+\frac{1}{2}}\alpha^l_{j-\frac{1}{2}}|\right) \Big/ \left(\alpha^l_{j+\frac{1}{2}} + \alpha^l_{j-\frac{1}{2}}\right) \tag{4.34d}$$

$$g^l_j = \left\{\alpha^l_{j-\frac{1}{2}}\left[(\alpha^l_{j+\frac{1}{2}})^2 + \delta_2\right] + \alpha^l_{j+\frac{1}{2}}\left[(\alpha^l_{j-\frac{1}{2}})^2 + \delta_2\right]\right\} \Big/ \left[(\alpha^l_{j+\frac{1}{2}})^2 + (\alpha^l_{j-\frac{1}{2}})^2 + 2\delta_2\right] \tag{4.34e}$$

$$g^l_j = \text{minmod}\left(2\alpha^l_{j-\frac{1}{2}}, 2\alpha^l_{j+\frac{1}{2}}, \frac{1}{2}(\alpha^l_{j+\frac{1}{2}} + \alpha^l_{j-\frac{1}{2}})\right) \tag{4.34f}$$

$$g^l_j = S \cdot \max\left[0, \min(2|\alpha^l_{j+\frac{1}{2}}|, S \cdot \alpha^l_{j-\frac{1}{2}}), \min(|\alpha^l_{j+\frac{1}{2}}|, 2S \cdot \alpha^l_{j-\frac{1}{2}})\right]; \; S = \text{sgn}(\alpha^l_{j+\frac{1}{2}}). \tag{4.34g}$$

Here δ_2 is a small parameter to prevent division by zero. In practical calculations $10^{-7} \leq \delta_2 \leq 10^{-5}$ is a commonly used range. The parameter $\delta_2 = 10^{-7}$in all of the calculations presented in sections 4.4.3 and 4.4.4. For $\alpha^l_{j+\frac{1}{2}} + \alpha^l_{j-\frac{1}{2}} = 0$, g^l_j is set to zero in (4.34d). The limiter (4.34f) is due to Woodward and Colella [82].

Roe-Sweby Second-order TVD Schemes: The elements of the vector $\Phi_{j+\frac{1}{2}}$ for the second-order Roe-Sweby upwind TVD scheme , denoted by $(\phi^l_{j+\frac{1}{2}})^{RS}$ are

$$(\phi^l_{j+\frac{1}{2}})^{RS} = -\left[|a^l_{j+\frac{1}{2}}| - \frac{\delta(r^l)}{2}\left(|a^l_{j+\frac{1}{2}}| - \lambda(a^l_{j+\frac{1}{2}})^2\right)\right]\alpha^l_{j+\frac{1}{2}} \tag{4.35}$$

where r^l is (3.47b) with the u's replaced by the lth local characteristic variable, and the $\Delta_{j+\frac{1}{2}}$ by $\alpha^l_{j+\frac{1}{2}}$.

An Explicit Predictor-Corrector TVD Scheme: With the above notation, a formal extension of the scalar explicit second-order TVD method (section 3.6) in a predictor-corrector form via the local-characteristic approach for the nonlinear hyperbolic system (4.1) can be written as

$$\Delta U^{(1)}_j = -\frac{\Delta t}{\Delta x}\left(F^n_j - F^n_{j-1}\right) \tag{4.36a}$$

$$U^{(1)}_j = \Delta U^{(1)}_j + U^n_j, \tag{4.36b}$$

$$\Delta U^{(2)}_j = \frac{1}{2}\left\{-\Delta U^{(1)}_j - \frac{\Delta t}{\Delta x}\left[F^{(1)}_{j+1} - F^{(1)}_j\right]\right\} \tag{4.36c}$$

$$U^{(2)}_j = \Delta U^{(2)}_j + U^{(1)}_j, \tag{4.36d}$$

$$U^{n+1}_j = U^{(2)}_j + \left[\overline{R}_{j+\frac{1}{2}}\overline{\Phi}_{j+\frac{1}{2}} - \overline{R}_{j-\frac{1}{2}}\overline{\Phi}_{j-\frac{1}{2}}\right] \tag{4.36e}$$

The elements of the vector $\Phi_{j+\frac{1}{2}}$ denoted by $(\phi^l_{j+\frac{1}{2}})^{PC}$ are the same as equations (3.69d) and (3.69e) with $a_{j+\frac{1}{2}}$ replaced by $a^l_{j+\frac{1}{2}}$, $\Delta_{j+\frac{1}{2}}$ replaced by $\alpha^l_{j+\frac{1}{2}}$, and $\widehat{Q}_{j+\frac{1}{2}}$ replaced by $\widehat{Q}^l_{j+\frac{1}{2}}$. $\overline{R}_{j+\frac{1}{2}}$ and $\overline{\Phi}_{j+\frac{1}{2}}$ can be evaluated at U^n or $U^{(2)}$ as discussed in sections 3.6 - 3.9.

High-Resolution TVD & TVB Lax-Friedrichs Schemes: The corresponding high-resolution TVD Lax-Friedrichs schemes for system cases can be obtained by changing the $\psi(z)$ in (4.33a) and (4.34a) with $\psi(z) = \frac{1}{\lambda}$ for any of the $(\phi^l_{j+\frac{1}{2}})^S$, $(\phi^l_{j+\frac{1}{2}})^U$ or $(\phi^l_{j+\frac{1}{2}})^{PC}$. In addition in changing the $\psi(z)$ to $\frac{1}{\lambda}$ one can obtain a high-resolution TVB Lax-Friedrichs method by changing the limiter function to be the same form as equation (3.60). Although using the Lax-Friedrichs numerical flux would introduce more numerical dissipation into the scheme, with this numerical flux, an entropy inequality is automatically satisfied. Thus one does not have to deal with an arbitrary parameter δ_1. In addition, at each grid point a savings of a (3×3) matrix multiplication with a vector is realized.

4.4.2. The MUSCL Approach

<u>MUSCL Approach Using an Approximate Riemann Solver</u>: The numerical flux function $\widetilde{F}_{j+\frac{1}{2}}$ for a MUSCL type approach of an upwind scheme as described in Yee [39] using the local-characteristic approach can be expressed as

$$\widetilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F(U^R_{j+\frac{1}{2}}) + F(U^L_{j+\frac{1}{2}}) + \widehat{R}_{j+\frac{1}{2}}\widehat{\Phi}_{j+\frac{1}{2}}\right], \tag{4.37a}$$

where the elements of $\widehat{\Phi}_{j+\frac{1}{2}}$ are

$$\widehat{\phi}^l_{j+\frac{1}{2}} = -\psi(\widehat{a}^l_{j+\frac{1}{2}})\widehat{\alpha}^l_{j+\frac{1}{2}}, \tag{4.37b}$$

$$\widehat{\alpha}_{j+\frac{1}{2}} = \widehat{R}^{-1}_{j+\frac{1}{2}}(U^R_{j+\frac{1}{2}} - U^L_{j+\frac{1}{2}}), \tag{4.37c}$$

where $\psi(z)$ can be $|z|$ or (4.33b). Here $\widehat{a}^l_{j+\frac{1}{2}}$ are the eigenvalues and $\widehat{R}_{j+\frac{1}{2}}$ is the eigenvector matrix of A, evaluated using a symmetric average between $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$; i.e.,

$$\widehat{a}^l_{j+\frac{1}{2}} = a^l(U^R_{j+\frac{1}{2}}, U^L_{j+\frac{1}{2}}), \tag{4.37d}$$

$$\widehat{R}_{j+\frac{1}{2}} = R(U^R_{j+\frac{1}{2}}, U^L_{j+\frac{1}{2}}). \tag{4.37e}$$

However, there are options in applying the limiters for system cases. Namely, one can impose limiters on the conservative, primitive, or characteristic variables. For a first-order time discretization, the simplest case is to define each of the elements of $U^{R,L}_{j+\frac{1}{2}}$ as in the scalar case (equation (3.38)). For a second-order time discretization, in addition to the various options in imposing the limiter, another step is needed. For the moment, let us assume that the variables for imposing the limiter are $\widetilde{W} = (\rho, u, p\epsilon)$. Denote P and P^{-1} as the matrices such that $U = P\widetilde{W}$ and $\widetilde{W} = P^{-1}U$. The vectors $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ for a second-order in time (due to S. Hancock of Physics International, unpublished work), second-order in space MUSCL approach can be

$$U^L_{j+\frac{1}{2}} = P(\widetilde{W}^{n+\frac{1}{2}}_j + \frac{1}{2}\widetilde{g}_j), \tag{4.38a}$$

$$U^R_{j+\frac{1}{2}} = P(\widetilde{W}^{n+\frac{1}{2}}_{j+1} - \frac{1}{2}\widetilde{g}_{j+1}). \tag{4.38b}$$

Here $\widetilde{g}_j$ is defined as in equations (4.34c)-(4.34g), except the arguments will be elements of $(\widetilde{W}^n_{j+1} - \widetilde{W}^n_j)$ and $(\widetilde{W}^n_j - \widetilde{W}^n_{j-1})$, with $\widetilde{W}^{n+\frac{1}{2}}_j = P^{-1}(U^{n+\frac{1}{2}}_j)$ where

$$U^{n+\frac{1}{2}}_j = U^n_j - \frac{\lambda}{2}\left\{F[P(\widetilde{W}^n_j + \frac{1}{2}\widetilde{g}_j)] - F[P(\widetilde{W}^n_j - \frac{1}{2}\widetilde{g}_j)]\right\}. \tag{4.38c}$$

One can define $\widetilde{W} = U$ the conservative variables or $\widetilde{W} = W$ the characteristic variables. A second-order in time but third-order in space scheme can be obtained by defining a different $\widetilde{g}_j$

function (see equation (3.38) for a formula). In all the shock tube computations in section 4.4.3, $\widetilde{W} = (\rho, u, \rho\epsilon)$.

MUSCL Approach Using the Lax-Friedrichs Numerical Flux [76]: The numerical flux function $\widetilde{F}_{j+\frac{1}{2}}$ for a MUSCL type approach using the Lax-Friedrichs numerical flux can be expressed as

$$\widetilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F(U^R_{j+\frac{1}{2}}) + F(U^L_{j+\frac{1}{2}}) + \widehat{\Phi}_{j+\frac{1}{2}}\right], \tag{4.39a}$$

where the lth element of $\widehat{\Phi}_{j+\frac{1}{2}}$ is

$$\widehat{\phi}^l_{j+\frac{1}{2}} = -\frac{1}{\lambda}(U^R_{j+\frac{1}{2}} - U^L_{j-\frac{1}{2}}) \tag{4.39b}$$

In addition to the fact that the Lax-Friedrichs numerical flux automatically satisfies an entropy inequality (thus one does not have to deal with an arbitrary parameter δ_1), there is also a tremendous savings in operations count (especially for multidimensional problems and/or nonequilibrium flows) in using this type of numerical flux instead of the Roe-type first-order upwind numerical flux for the MUSCL formulation if the limiter function is applied to the conservative or primitive variables instead of the characteristic variables. In problems containing contact discontinuities as well as shocks, one can use a more compressive limiter for the density and a less compressive limiter for the other variables. Note that one does not have a similar savings using the Lax-Friedrichs numerical flux for the non-MUSCL formulations.

MUSCL Approach Using Flux-Vector Splittings: The numerical flux $\widetilde{F}_{j+\frac{1}{2}}$ for either flux-vector splittings, can be expressed as

$$\widetilde{F}_{j+\frac{1}{2}} = F^+(U^L_{j+\frac{1}{2}}) + F^-(U^R_{j+\frac{1}{2}}), \tag{4.40}$$

where $F^{\pm}(U^{L,R}_{j+\frac{1}{2}})$ are evaluated using either the generalized Steger-Warming splitting or the generalized van Leer splitting. The vectors $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ are the same as in equation (4.38).

4.4.3. Comparative Study of TVD Schemes for Equilibrium Real Gases [150]

A detailed description of this study can be found in Montagné et al. [150]. Six one-dimensional shock-tube problems were considered.

Case	State	Density	Pressure	Temp.	Energy	Velocity	Mach
Case A	Left	0.0660	$9.84\ 10^4$	4390	$7.22\ 10^6$	0	0.0
	Right	0.0300	$1.50\ 10^4$	1378	$1.44\ 10^6$	0	0.0
Case B	Left	1.4000	$9.88\ 10^5$	2438	$2.22\ 10^6$	0	0.0
	Right	0.1400	$9.93\ 10^3$	2452	$2.24\ 10^6$	0	0.0
Case C	Left	1.2900	$1.00\ 10^5$	272	$1.95\ 10^5$	0	0.0
	Right	0.0129	$1.00\ 10^4$	2627	$2.75\ 10^6$	0	0.0
Case D	Left	1.0000	$6.50\ 10^5$	2242	$2.00\ 10^6$	0	0.0
	Right	0.0100	$1.00\ 10^3$	346	$2.50\ 10^5$	0	0.0
Case E	Left	0.0100	$5.73\ 10^2$	199	$1.44\ 10^5$	2200	7.8
	Right	0.1400	$2.23\ 10^4$	546	$4.00\ 10^5$	0	0.0
Case F	Left	0.0100	$5.73\ 10^2$	199	$1.44\ 10^5$	4000	14.6
	Right	0.0100	$5.73\ 10^2$	199	$1.44\ 10^5$	-4100	-14.5

Table 1. Initial conditions for the test cases.

The left and right states of the initial conditions for all six cases are tabulated in table 1. The cases have been ordered in the direction of increasing maximum Mach numbers encountered in the flow. The densities are given in kg/m^3, the pressures in $Newtons/m^2$, the temperatures in Kelvin, the internal energies in $kg(m/sec)^2$ and the velocities in m/sec.

The thermodynamic properties of equilibrium air were obtained from the curve fits of Srinivasan et al. [164]. These curve fits give analytic expressions for $\tilde{\gamma}$ in several ranges of density and internal energy. The values of γ, χ, and κ are then calculated from the derivatives of these analytical expressions. The numerical solutions were compared with an "exact solution" computed by solving the Rankine-Hugoniot jump conditions and integrating numerically the characteristic equations in the expansion fan. Figure 4.1 shows the distribution of Mach number and of the two quantities γ and $\tilde{\gamma}$ defined in (4.8). Not only the changes in their values but also the differences between them are indications of departure from the perfect-gas case. These differences do not necessarily occur at very high temperatures, but at intermediate temperatures when the vibration is excited and when the dissociation reactions start. Results for cases B, C and E are discussed here.

The combination of Riemann solvers and of differencing algorithms considered above yields five different schemes: a symmetric non-MUSCL scheme, an upwind non-MUSCL scheme, and three MUSCL-type schemes, depending on the Riemann solvers. The Roe-Sweby and the high-resolution Lax-Friedrichs schemes were not considered in this study. The study provided a check on the validity of the extended formulas, since theoretical prediction of their properties appears to be difficult due to the non-analytic form of the state equation. Comparisons were done on the accuracy and the robustness of the methods. The six test cases chosen were intended to highlight the effect of the high ratios in pressure or density related to shocks, and the effect of departure from a perfect gas in the state equation.

The five second-order explicit schemes tested were (a) the symmetric TVD scheme (4.33), (b) the upwind TVD scheme (4.34), (c) the upwind TVD scheme (4.37, 4.38), (d) the generalized van

Leer flux splitting (4.40) and (e) the generalized Steger-Warming flux splitting (4.40). Schemes (a) and (b) follow the non-MUSCL approach, while schemes (c)-(e) follow the MUSCL approach. The same approximate Riemann solver (local-characteristic approach, section 4.3) is used in the three schemes (a), (b) and (c). The limiter function used for each scheme remains the same for all the cases. Limiter (4.33e) is used for the symmetric TVD scheme (4.33), and limiter (4.34e) is used for the rest. For the MUSCL computations, the function $\psi(z) = |z|$. The time-step limit is expressed in terms of a CFL number related to the eigenvalues of the numerical fluxes. The CFL number is fixed at 0.9 in cases B and E. In case C, it has been fixed at 0.5 for the upwind non-MUSCL and MUSCL scheme and at 0.9 for the symmetric scheme. The actual CFL used for the flux splitting approaches is approximately 80% of the fixed CFL (see reference [150] for details). The number of discretization points is 141 in cases B and C, and only 81 for case E because the expansion fan is replaced by a shock. The time for stopping the computation has been chosen for each case in order to use the full computational domain. For a Δx normalized to 0.1 meter, these times, given in 10^{-3} seconds, are $t = 3.0$ for case B, $t = 5.0$ for case C and $t = 3.2$ for case E. The comparative study can be divided into two aspects, one on the differencing algorithms, and the other on the Riemann solvers. Figures (4.2)-(4.4) show the perfect-gas computations and figures (4.5)-(4.7) show the equilibrium real-gas computations. In all of the computations δ_1 (in equation (3.18)) was set to zero for the non-flux-split approach. The vector $\widetilde{W}$ in (4.38) is set to $(\rho, u, \rho\epsilon)$. Since the resolutions of shock and contact discontinuities do not always behave the same for the different variables, the computed solutions for $\gamma, \bar{\gamma}$, Mach number, energy, velocity and density are shown here for completeness.

Comparison of the Differencing Techniques: Parts (a)-(c) of figures (4.2)-(4.7) provide a comparison of the symmetric TVD and the upwind TVD non-MUSCL scheme, and the MUSCL scheme with the same approximate Riemann solver for both perfect and equilibrium real gases.

The three techniques give almost the same results in general and the differences are similar to those found for a perfect gas. The greatest difference occurs in test case C. But this case happens to be already a difficult one when the same initial conditions on density and energy are applied to a perfect gas. The major differences are between the symmetric scheme and the two upwind schemes. Although the symmetric scheme is generally more diffusive at the contact discontinuities, the situation is reversed in case E where the main shock is almost stationary and the flow behind it has a very low velocity. Furthermore, this symmetric scheme yields more stable results in that case. The influence of the limiters is the same as for a perfect gas as summarized in references [21,75] and in sections 5 and 6. A comparative study of flux limiters for case B will be discussed in the next subsection (4.4.4). This point, like some aspects related to the computational efficiency, is discussed more fully in [150].

The main difference in computational effort lies in the MUSCL and non-MUSCL approaches. The operations count between the non-MUSCL and MUSCL is within 30% for a perfect gas. However, due to extra evaluation in the curve fitting between the left and right states in an equilibrium real gas for the MUSCL formulation, additional computation is required for the MUSCL approach. The slight advantage of MUSCL over non-MUSCL is that MUSCL can be spatially third-order accurate. One word of caution is that experiences with the third-order case ($\eta = 1/3$ for equation (3.38)) do not show a very visible improvement over the second-order case. Part of the reason is that all TVD schemes reduce to first-order at points of extrema regardless of the order of accuracy in smooth regions.

Comparison of the Riemann Solvers: Parts (c)-(e) of figures (4.2)-(4.7) compare the three Riemann solvers for the MUSCL scheme for both perfect and equilibrium real gases.

In the test conditions used, a comparison between the two classes of flux splittings showed little difference when equilibrium air was used. The generalized van Leer splitting yields a sharper capture of the shocks than the generalized Steger-Warming splitting. The resolution and operations count of the different approximate Riemann solvers [51,99,116] are very similar. The results obtained with the approximate Riemann solver are very similar to the ones obtained with the generalized van Leer splitting. Actually, the generalized van Leer splitting seems to be less sensitive to the state equation for the shock resolution while the approximate Riemann solver is more accurate at the contact discontinuities.

It is important to note that flux-vector splittings make use of the sound speed only, whereas approximate Riemann solvers of the Roe-type make use of the thermodynamic dervatives χ and κ of equation (4.2). These thermodynamic derivatives put more stringent requirements on the curve fit that represents the thermodynamic properties of the gas. In this regard, the curve fits of Srinivasan et al. may be deficient for the approximate Riemann solver as can be seen from figure (4.1), case D. One probably needs more improved curve fits such as those of reference [155] before a definite conclusion can be drawn about the accuracy of the different Riemann solvers and schemes.

In conclusion, for the purpose of calculations in gas dynamics with equilibrium real gases, these numerical tests show that the simple extensions to an equilibrium real gas of the flux-vector splitting or the approximate Riemann solver presented in this section are valid. The main effect of using a real-gas equation of state is to exacerbate the problems of the methods for large discontinuities. Test case C is an example of such a situation. Similarily, it seems difficult to give a ranking of the methods. Depending on the case, each one presents some drawbacks or some advantages. The present results also indicate that the state equation does not have a very large effect on the general behavior of these methods for a wide range of flow conditions.

None of the differences observed for the explicit versions seems to be decisive for the one-dimensional tests, but factors such as stability and computational efficiency need further investigation in multidimensional tests. The main differences between the methods lie in their versatility in extending to implicit methods with efficient solution procedures, especially for multidimensional steady-state computations. Preliminary study shows certain advantages of the approximate Riemann solver over the flux-vector splitting approaches (see section 4.5 for a discussion).

4.4.4. Comparative Study of Flux Limiters for a 1-D Shock-Tube Problem [1]

The same five different explicit schemes discussed in the last subsection (4.4.3) were used to study the effect of different limiters on the accuracy of the schemes for both perfect and equilibrium real gases. Because of the anticipated decrease in accuracy for equilibrium real gases of the curve fits of Srinivasan et al. in providing the thermodynamic derivatives, only the result of case B is summarized and shown here.

Figure (4.8) provides a perfect-gas comparison of the five schemes in conjunction with the different limiters (4.33) and (4.34) corresponding to the appropriate schemes. Figure (4.9) provides the same comparison for an equilibrium real gas. A CFL = 0.9 was used for all computations. Parts (a) - (c) of figures (4.8) and (4.9) show the accuracy of Roe's first-order scheme, Harten's original modified-flux scheme (3.45) without the added artificial compression [17], and the author's modification to (3.45) (i.e. (4.34a)-(4.34c)) respectively. One can see the dramatic improvement of scheme (4.34a)-(4.34c) over (3.45). Harten's scheme (3.45) produced accuracy similar to that of Roe's first-order method.

Parts (d) - (m) of figures (4.8) and (4.9) summarize the non-MUSCL and MUSCL approaches for both perfect and equilibrium real gases for all the second-order methods in section 4.4. In the case of a non-MUSCL approach, limiter (4.33e) is the most accurate among (4.33c) - (4.33e) for the symmetric TVD scheme (4.33). As for the upwind schemes, limiters (4.34d) and (4.34e) are very similar, whereas limiter (4.34g) gives very accurate results for contact discontinuities but is sometimes too compressive, thus causing slight oscillations in smooth regions for high Mach number cases (result not shown). A combination of limiters such as (4.34d) or 4.34e) for the nonlinear fields and (4.34g) for the linear field seems to be a good compromise. In the case of the MUSCL approach, only limiters (4.34c) and (4.34e) were studied. Between the two limiters, (4.34e) produces higher shock resolution than (4.34c) (comparison not shown). Not shown is the Colella and Woodward limiter (4.34f). The resolution is somewhere between the superbee and (4.34e).

4.5. Description of an Implicit Numerical Algorithm and an Example

The corresponding implicit scheme of equation (3.21) for the system case via the local-characteristic approach can be written as

$$U_j^{n+1} + \theta \frac{\Delta t}{\Delta x}\left(\widetilde{F}_{j+\frac{1}{2}}^{n+1} - \widetilde{F}_{j-\frac{1}{2}}^{n+1}\right) = U_j^n - (1-\theta)\frac{\Delta t}{\Delta x}\left(\widetilde{F}_{j+\frac{1}{2}}^{n} - \widetilde{F}_{j-\frac{1}{2}}^{n}\right). \tag{4.41}$$

Here θ has the same meaning as in equation (3.21). The spatial accuracy of the scheme depends on the form of the numerical flux functions. Note that the same numerical flux (to be discussed in section VI) is also applicable for the two-parameter family of time differencings as discussed in section 3.11. Implicit methods via the flux-vector splitting approaches in conjunction with a MUSCL formulation can be found in references [77,122,123] and will not discussed here.

Non-MUSCL Approach: The numerical flux function $\widetilde{F}_{j+\frac{1}{2}}$ for a non-MUSCL type approach for both the upwind and symmetric TVD schemes has the same form as equation (4.32); i.e., $\widetilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F_j + F_{j+1} + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\right]$. The elements of the vector $\Phi_{j+\frac{1}{2}}$ denoted by $(\phi_{j+\frac{1}{2}}^l)^S$ for a general second-order symmetric TVD scheme are

$$(\phi_{j+\frac{1}{2}}^l)^S = -\psi(a_{j+\frac{1}{2}}^l)(\alpha_{j+\frac{1}{2}}^l - \widehat{Q}_{j+\frac{1}{2}}^l), \tag{4.42}$$

where $\alpha_{j+\frac{1}{2}}^l$ are elements of $R_{j+\frac{1}{2}}^{-1}(U_{j+1} - U_j)$. The function ψ is the same as in the corresponding scalar case. The limiter function $\widehat{Q}_{j+\frac{1}{2}}^l$ can be the same as (4.33).

The elements of the vector $\Phi_{j+\frac{1}{2}}$ denoted by $(\phi_{j+\frac{1}{2}}^l)^U$ for a second-order implicit upwind TVD scheme, originally developed by Harten [17] and later modified by Yee [74,75,163], are

$$(\phi^l_{j+\frac{1}{2}})^U = \frac{1}{2}\psi(a^l_{j+\frac{1}{2}})(g^l_{j+1} + g^l_j) - \psi(a^l_{j+\frac{1}{2}} + \gamma^l_{j+\frac{1}{2}})\alpha^l_{j+\frac{1}{2}}. \tag{4.43}$$

The $\gamma^l_{j+\frac{1}{2}}$ and the limiter function g^l_j can be the same as (4.34). The situation is similar for the Roe-Sweby numerical flux except the $(\phi^l_{j+\frac{1}{2}})^{RS}$ function in this case is (4.35) with the term $\lambda(a^l_{j+\frac{1}{2}})^2$ set to zero.

MUSCL Approach: The numerical flux function $\tilde{F}_{j+\frac{1}{2}}$ for a MUSCL-type approach of the implicit TVD upwind scheme has the same form as the explicit case, i.e., $\tilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F(U^R_{j+\frac{1}{2}}) + F(U^L_{j+\frac{1}{2}}) + \hat{R}_{j+\frac{1}{2}}\hat{\Phi}_{j+\frac{1}{2}}\right]$, except that the form of the $U^{R,L}_{j+\frac{1}{2}}$ is slightly different. The values $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ are the same as (4.38) without the additional step (4.38c); i.e., setting $U^{n+\frac{1}{2}}_j = U_j$.

High-Resolution Lax-Friedrichs Method: The TVD or TVB numerical flux function $\tilde{F}_{j+\frac{1}{2}}$ for a MUSCL or non-MUSCL approach using the Lax-Friedrichs numerical flux for (4.41) has the same form and properties as discussed in sections 4.4.1 and 4.4.2 for the explicit methods with the exception of dropping the $\lambda(a^l_{j+\frac{1}{2}})^2$ term in (4.33a), (4.34a), (4.35) and (4.36), and the appropriate limit for TVD or TVB methods.

A Conservative Linearized Form for Steady-State Applications: A conservative linearized form of (4.41) (see section 3.10) can be written as

$$\left[I + \theta\frac{\Delta t}{\Delta x}\left(H^x_{j+\frac{1}{2}} - H^x_{j-\frac{1}{2}}\right)\right]^n E^n = -\frac{\Delta t}{\Delta x}\left(\tilde{F}^n_{j+\frac{1}{2}} - \tilde{F}^n_{j-\frac{1}{2}}\right), \tag{4.44a}$$

with

$$E^n = U^{n+1} - U^n, \tag{4.44b}$$

where

$$H^x_{j+\frac{1}{2}} = \frac{1}{2}\left(A_{j+1} - \Omega^x_{j+\frac{1}{2}}\right). \tag{4.44c}$$

Note that the matrix $H^x_{j+\frac{1}{2}}$ in (4.44c) is different from the total enthalpy H in section 4.1. The nonstandard notation

$$H^x_{j+\frac{1}{2}}E = \frac{1}{2}\left(A_{j+1}E_{j+1} - \Omega^x_{j+\frac{1}{2}}E\right) \tag{4.44d}$$

is used, and for a first-order implicit operator, $\Omega^x_{j+\frac{1}{2}}E$ can be

$$\Omega^x_{j+\frac{1}{2}}E = R_{j+\frac{1}{2}}\mathrm{diag}\left[\psi(a^l_{j+\frac{1}{2}})\right]R^{-1}_{j+\frac{1}{2}}(E_{j+1} - E_j). \tag{4.44e}$$

Here A_{j+1}, is the Jacobian of F evaluated at $(j+1)$ and $E^n_j = U^{n+1}_j - U^n_j$. For steady-state applications, one can simplify (4.44e) as

$$\Omega^x_{j+\frac{1}{2}}E = \mathcal{M}_x I(E_{j+1} - E_j). \tag{4.44f}$$

The scalar value $\mathcal{M}_x$ is

$$\mathcal{M}_x = \max_l \psi(a^l_{j+\frac{1}{2}}), \tag{4.44g}$$

and I is the identity matrix. Note that (4.44f) involves scalar multiplications only. The solution using (4.44) is still spatially second-order (or third-order) accurate after it reaches steady state. A second-order form of (4.44e) for the non-MUSCL formulation can be found in references [21,74]. The nonconservative linearized form as discussed in the scalar case can be obtained similarly. See references [21,74] for details.

The author would like to point out an important distinction between the flux-vector splittings and the local-characteristic approach for implicit methods. Unlike flux-vector splitting approaches, implicit methods employing the local-characteristic approach (non-MUSCL or MUSCL) with first-order implicit operators such as (4.44f)) do not require the Jacobian of the $F^{\pm}$ fluxes. In many instances, the Jacobian of $F^{\pm}$ is relatively difficult to obtain. A similar difficulty applies to the MUSCL formulations via the local-characteristic approach if a second-order implicit operator is desired.

To show the stability and accuracy of an implicit upwind TVD-type method for steady-state application, results for a quasi-one-dimensional divergent nozzle problem [11] are presented. Figure (4.10b) shows the converged density distribution after 25 steps at a CFL number of 10^6 using 20 equal grid spacings and using (4.43) with the minmod function (4.34c) and $\theta = 1$. The value δ_1 in (3.18) is set to 0.125. The solid line is the exact solution and the diamonds are the computed solution. Only 14 points are plotted. The six points not shown on both ends of the x-axis are equal to the exact solution. The solution looks very much like the explicit upwind TVD scheme using the corresponding explicit numerical flux and the same limiter (figure 4.10a) except the implicit method has a tremendous gain in efficiency. Figures (4.10c) and (4.10d) show the same computation with a classical shock-capturing method and the first-order Steger-Warming flux-vector splitting methods.

4.6. On the Implementation of High-Resolution Schemes for Nonlinear Systems

This section discusses a practical point on implementing schemes that are designed for nonlinear scalar equations into nonlinear systems by Riemann solvers. This is largely related to the way second-order accuracy is achieved and to the types of limiters (flux or slope limiters). For the scalar case the difference in operations count between any two algorithms may be very small and yet the operations count for their system counterparts might be vastly different, especially for problems in multidimensions and in curvilinear coordinates. The situation occurs even though one starts with two different yet equivalent representations for the scalar case. In the discussion, only the MUSCL scheme (3.39), Harten scheme (3.45) with $\psi(z) = |z|$ and the Osher-Chakravarthy scheme (3.49) are considered. The results and discussion carry over to other schemes with similar structures as well.

To discuss the relative computational effort among the spatially five-point TVD schemes of Harten, MUSCL and Osher-Chakravarthy for system cases, the local characteristic method of generalization for systems is chosen. It is emphasized here that the scalar schemes of Harten, MUCSL, and Osher-Chakravarthy are used. Then extensions to systems are employed by the same method (i.e. Huang or Roe), not by their original Riemann solvers [69,139] for system cases.

For simplicity of presentation, assume $\psi(z)$ in (3.45) has the form $\psi(z) = |z|$. The use of an entropy-satisfying $\psi(z)$ is only a minor modification of what is about to be presented. Furthermore, since the main concern of this discussion is the spatial discretization, only the forward Euler time differencing is considered. Other explicit, implicit, or predictor-corrector types of time differencing do not alter the results.

Using the identity $|\widetilde{a}_{j+\frac{1}{2}}| = \mathrm{sgn}(\widetilde{a}_{j+\frac{1}{2}})\widetilde{a}_{j+\frac{1}{2}}$ and $\widetilde{a}_{j+\frac{1}{2}} = (\widetilde{f}_{j+1} - \widetilde{f}_{j+1})/(u_{j+1} - u_j)$, Harten's scheme can be rewritten as

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}\left[1 - \mathrm{sgn}(\widetilde{a}_{j+\frac{1}{2}}^n)\right](\widetilde{f}_{j+1}^n - \widetilde{f}_j^n) - \frac{\lambda}{2}\left[1 + \mathrm{sgn}(\widetilde{a}_{j-\frac{1}{2}}^n)\right](\widetilde{f}_j^n - \widetilde{f}_{j-1}^n). \tag{4.45}$$

The g_j function of (3.45b) can be defined in a slightly different form.

$$g_j = \mathrm{minmod}\left(\bar{\sigma}_{j+\frac{1}{2}}\Delta_{j+\frac{1}{2}}f, \bar{\sigma}_{j-\frac{1}{2}}\Delta_{j-\frac{1}{2}}f\right) \tag{4.46a}$$

with $\bar{\sigma}_{j+\frac{1}{2}} = \bar{\sigma}(a_{j+\frac{1}{2}})$ and

$$\bar{\sigma}(z) = \frac{1}{2}\left[\mathrm{sgn}(z) - \lambda z\right]. \tag{4.46b}$$

Here the identities $\mathrm{sgn}(\Delta_{j+\frac{1}{2}}f)\mathrm{sgn}(a_{j+\frac{1}{2}}) = \mathrm{sgn}(u_{j+1} - u_j)$ and $|\Delta_{j+\frac{1}{2}}f| = |a_{j+\frac{1}{2}}||u_{j+1} - u_j|$ are used. The above limiter g_j of (4.46) can be considered as a flux limiter since the flux f is limited. For scalar problems, equations (3.45a), and (4.45) are equivalent. The difference in operations count is very minor. Equation (3.45a) is preferred over (4.45) for its straightforward extension to system cases because u appears rather than f. In the extension to nonlinear systems via Huang's and Roe's generalizations, (3.45a) results in a smaller operations count than (4.45).

Similarly, the Roe-Sweby scheme (3.47b) can be rewritten in a similar form as (3.45a) (i.e. in terms of $u_{j+1} - u_j$ instead of $f_{j+1} - f_j$ using the above identities). The Roe-Sweby scheme written in terms of $u_{j+1} - u_j$ is expressed in equation (3.52). Again, with the same reasoning, equation (3.52) is preferred over (3.47b).

The numerical flux (3.45) for the Harten scheme for the system case can be found in section 4.4. His numerical flux $\widetilde{F}_{j+\frac{1}{2}} = \frac{1}{2}\left[F_j + F_{j+1} + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\right]$. Extension of the scalar TVD scheme to nonlinear system cases is not unique. Take, for example, the case where the numerical flux $\widetilde{h}_{j+\frac{1}{2}}$ in (3.45a) is equivalent to $\widetilde{h}_{j+\frac{1}{2}}$ in (4.45). The corresponding numerical fluxes for the system case have a very different operations count and are no longer equivalent to each other. If one started with (4.45), the system can be expressed as

$$U_j^{n+1} = U_j^n - \frac{\lambda}{2}\left[I - \mathcal{A}_{j+\frac{1}{2}}^n\right]\left[\mathcal{F}_{j+1}^n - \mathcal{F}_j^n\right] - \frac{\lambda}{2}\left[I + \mathcal{A}_{j-\frac{1}{2}}^n\right]\left[\mathcal{F}_j^n - \mathcal{F}_{j-1}^n\right] \tag{4.47a}$$

where

$$\mathcal{A}_{j+\frac{1}{2}} = R^{-1}_{j+\frac{1}{2}}\Lambda_{j+\frac{1}{2}}R_{j+\frac{1}{2}} \tag{4.47b}$$

$$\Lambda_{j+\frac{1}{2}} = \mathrm{diag}(\mathrm{sgn}(a^l_{j+\frac{1}{2}})) \tag{4.47c}$$

and

$$\mathcal{F}_j = F_j + R_jG_j \tag{4.47d}$$

with $G_j = (g^1_j, \ldots, g^m_j)^T$ where g^l_j can be (4.34c) -(4.34g), and diag(z^l) denoting a diagonal matrix with diagonal elements z^l. For (4.47), the eigenvector R is evaluated at $U_{j+\frac{1}{2}}$; i.e., $R_{j+\frac{1}{2}}$ is needed in both (4.32) and (4.47). Therefore scheme (4.32) is preferred over (4.47) since R_j is not required in (4.32). In other words, for scalar cases (3.45a) and (4.45) are identical, but in system cases (4.47) requires more work than (4.32). The same situation holds true for the Roe-Sweby scheme in the system cases. The MUSCL numerical flux vector $\tilde{F}^{VL}_{j+\frac{1}{2}}$ can be found in (4.37a).

The numerical flux for the Osher-Chakravarthy scheme for a system can be written as

$$\begin{aligned} \tilde{F}^{OC}_{j+\frac{1}{2}} = H_{j+\frac{1}{2}} &- \frac{(1-\bar{\eta})}{4}R_{j+\frac{3}{2}}\widetilde{\mathcal{A}^-_{j+\frac{3}{2}}} - \frac{(1+\bar{\eta})}{4}R_{j+\frac{1}{2}}\widetilde{\widetilde{\mathcal{A}^-_{j+\frac{1}{2}}}} \\ &+ \frac{(1+\bar{\eta})}{4}R_{j+\frac{1}{2}}\widetilde{\mathcal{A}^+_{j+\frac{1}{2}}} + \frac{(1-\bar{\eta})}{4}R_{j-\frac{1}{2}}\widetilde{\widetilde{\mathcal{A}^+_{j-\frac{1}{2}}}}. \end{aligned} \tag{4.48a}$$

Here the $H_{j+\frac{1}{2}}$ is the first-order numerical flux

$$H_{j+\frac{1}{2}} = \frac{1}{2}\left[F_j + F_{j+1} - R_{j+\frac{1}{2}}\hat{\hat{\Phi}}_{j+\frac{1}{2}}\right], \tag{4.48b}$$

where the lth element of $\hat{\hat{\Phi}}_{j+\frac{1}{2}}$ is $\phi^l_{j+\frac{1}{2}} = |a^l_{j+\frac{1}{2}}|\alpha^l_{j+\frac{1}{2}}$.

The lth elements of $\mathcal{A}^{\pm}$'s are

$$\mathcal{A}^{l-}_{j+\frac{1}{2}} = a^{l-}_{j+\frac{1}{2}}\alpha^l_{j+\frac{1}{2}} \tag{4.48c}$$

$$\mathcal{A}^{l+}_{j+\frac{1}{2}} = a^{l+}_{j+\frac{1}{2}}\alpha^l_{j+\frac{1}{2}} \tag{4.48d}$$

with $a^{l\pm}_{j+\frac{1}{2}} = 0.5(a^l_{j+\frac{1}{2}} \pm |a^l_{j+\frac{1}{2}}|)$, and $a^l_{j+\frac{1}{2}}$ is (4.5), and

$$\widetilde{\mathcal{A}^{l-}_{j+\frac{3}{2}}} = \mathrm{minmod}\left[\mathcal{A}^{l-}_{j+\frac{3}{2}}, \omega\mathcal{A}^{l-}_{j+\frac{1}{2}}\right] \tag{4.48e}$$

$$\widetilde{\widetilde{\mathcal{A}^{l-}_{j+\frac{1}{2}}}} = \mathrm{minmod}\left[\mathcal{A}^{l-}_{j+\frac{1}{2}}, \omega\mathcal{A}^{l-}_{j+\frac{3}{2}}\right] \tag{4.48f}$$

$$\widetilde{\mathcal{A}^{l+}_{j+\frac{1}{2}}} = \text{minmod}\left[\mathcal{A}^{l+}_{j+\frac{1}{2}}, \omega \mathcal{A}^{l+}_{j-\frac{1}{2}}\right] \tag{4.48g}$$

$$\widetilde{\widetilde{\mathcal{A}^{l+}_{j-\frac{1}{2}}}} = \text{minmod}\left[\mathcal{A}^{l+}_{j-\frac{1}{2}}, \omega \mathcal{A}^{l+}_{j+\frac{1}{2}}\right]. \tag{4.48h}$$

Collecting terms of $R_{j+\frac{1}{2}}$, (4.48a) can be rewritten as

$$\widetilde{F}^{OC}_{j+\frac{1}{2}} = \frac{1}{2}\Big\{F_j + F_{j+1} - R_{j+\frac{1}{2}}\Big[\widehat{\Phi}_{j+\frac{1}{2}} + \frac{1+\bar{\eta}}{4}\widetilde{\widetilde{\mathcal{A}^{-}_{j+\frac{1}{2}}}} - \frac{1+\bar{\eta}}{4}\widetilde{\mathcal{A}^{+}_{j+\frac{1}{2}}}\Big]$$

$$-\frac{1-\bar{\eta}}{4}R_{j+\frac{3}{2}}\widetilde{\mathcal{A}^{-}_{j+\frac{3}{2}}} + \frac{1-\bar{\eta}}{4}R_{j-\frac{1}{2}}\widetilde{\widetilde{\mathcal{A}^{-}_{j-\frac{1}{2}}}}\Big\}. \tag{4.49}$$

The terms that are common to the numerical fluxes of Harten's $\widetilde{F}^{H}_{j+\frac{1}{2}}$, Roe-Sweby $\widetilde{F}^{RS}_{j+\frac{1}{2}}$, MUSCL $\widetilde{F}^{VL}_{j+\frac{1}{2}}$ and Osher-Chakravarthy's $\widetilde{F}^{OC}_{j+\frac{1}{2}}$ are F_{j+1}, F_j, $R_{j+\frac{1}{2}}$ and $\alpha_{j+\frac{1}{2}}$. Both $\widetilde{F}^{H}$ and $\widetilde{F}^{VL}$ require one matrix-vector multiplication $R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}$ or $\widehat{R}_{j+\frac{1}{2}}\widehat{\Phi}_{j+\frac{1}{2}}$. Comparing all four numerical fluxes, the operations count between $\widetilde{F}^{H}$, $\widetilde{F}^{RS}_{j+\frac{1}{2}}$ and $\widetilde{F}^{VL}$ is very competitive. The first three terms on the right hand side of $\widetilde{F}^{OC}_{j+\frac{1}{2}}$ in (4.49) require similar operations count as $\widetilde{F}^{H}$, $\widetilde{F}^{RS}_{j+\frac{1}{2}}$ and $\widetilde{F}^{VL}$. The last two terms of (4.49) require approximately 2/5 of the computation of the numerical flux $\widetilde{F}^{OC}_{j+\frac{1}{2}}$. They contain the extra computations that are not present in $\widetilde{F}^{H}$, $\widetilde{F}^{RS}_{j+\frac{1}{2}}$ and $\widetilde{F}^{VL}$. Therefore, for $\bar{\eta} \neq -1$, $\widetilde{F}^{OC}$ is approximately 40% more expensive to use than $\widetilde{F}^{H}$ or $\widetilde{F}^{VL}$. For $\bar{\eta} = -1$, $\widetilde{F}^{OC}$ is approximately 25% more expensive than the other two schemes.

4.7. Splitting Methods for Problems Containing Source Terms

The various methods discussed in section III for problems with source terms can be formally extended to system cases. The extension is straightforward and only a splitting method is described here. The corresponding form of (3.75) for the system case can be written as

$$U^{n+1} = \mathcal{L}_S^{h/2}\mathcal{L}_F^{h}\mathcal{L}_S^{h/2}U^n. \tag{4.50}$$

A split version of the method studied in sections 3.6 and 3.7 for the system case can take the form

$$
\begin{aligned}
\mathcal{L}_S^{h/2}: \qquad & [I - \tfrac{1}{4}\Delta t S'(U_j^n)]\Delta U_j^* = \tfrac{1}{2}\Delta t S(U_j^n) \\
& U_j^* = U_j^n + \Delta U_j^* \\[1em]
\mathcal{L}_F^h: \qquad & \Delta U_j^{(1)} = -\frac{\Delta t}{\Delta x}\left[F(U_j^*) - F(U_{j-1}^*)\right] \\
& U_j^{(1)} = U_j^* + \Delta U_j^{(1)} \\
& \Delta U_j^{(2)} = -\frac{\Delta t}{\Delta x}\left[F(U_{j+1}^{(1)}) - F(U_j^{(1)})\right] \\
& U_j^{(2)} = U_j^* + \frac{1}{2}\left(\Delta U_j^{(1)} + \Delta U_j^{(2)}\right) \\
& U_j^{**} = U_j^{(2)} + (R_{j+\frac{1}{2}}^* \Phi_{j+1/2}^* - R_{j-\frac{1}{2}}^* \Phi_{j-1/2}^*) \\[1em]
\mathcal{L}_S^{h/2}: \qquad & [I - \tfrac{1}{4}\Delta t S'(U_j^{**})]\Delta U_j^{**} = \tfrac{1}{2}\Delta t S(u_j^{**}) \\
& U_j^{n+1} = U_j^{**} + \Delta U_j^{**}
\end{aligned}
\tag{4.51}
$$

Here S' is the Jacobian of the source term S and Φ^* involves limited fluxes as before, based on $(\alpha^*)_{j+\frac{1}{2}}^l$ and $(\alpha^*)_{j-\frac{1}{2}}^l$. Alternatively, one can compute the limited value based on $U^{(2)}$ and replace Φ^* by $\Phi^{(2)}$. Note that one does not have to use a predictor-corrector type of method for the homogeneous part of the PDEs.

Again, each of these methods could be replaced by other well-known methods for the respective problems. For example, any explicit or implicit time differencing such as the explicit Euler or the trapezoidal method could be used for $\mathcal{L}_S^{h/2}$, and any of a wide variety of high-resolution methods as discussed in sections 3.5 or 3.6 could be used for $\mathcal{L}_F^h$.

V. EXTENSIONS OF NONLINEAR SCALAR TVD SCHEMES TO HIGHER-DIMENSIONAL NONLINEAR SYSTEMS

This section describes the formal extension of higher than first-order TVD scalar schemes to multidimensional nonlinear systems of hyperbolic conservation laws. Time-accurate as well as steady-state calculations for two- and three-dimensional practical applications will be illustrated when appropriate.

At the present stage of development, truly multidimensional schemes are still in the research stage [60]. The available theory is too complicated for practical applications. Here the usual approach of applying the one-dimensional scalar scheme and the Riemann solver for each direction in multidimensional nonlinear problems is adopted. Therefore, highly skewed nonorthogonal grids should be avoided. As will be illustrated in later subsections, this method of extending one-dimensional schemes to multidimensional nonlinear systems is quite satisfactory.

In order to preserve the original second-order time-accuracy, the extension of the nonlinear scalar second-order explicit schemes to multidimensional problems can be accomplished by a predictor-corrector [1,27,95], Runge-Kutta multistage [33] or Strang-type of fractional-step [97] (time-splitting) method. All the time-accurate numerical examples illustrated in this section utilize the time-splitting and predictor-corrector methods. Therefore, only these two methods will be described here. For the implicit methods, only a simplified form will be briefly discussed, since extensions to other forms follow the same idea. Noniterative relaxation implicit methods other than alternating direction implicit (ADI) methods will not be discussed. Extensive work in the area of relaxation methods in conjunction with van Leer flux-vector splitting for perfect gases has been performed by Thomas, Walters and van Leer; see references [122,123,165] for details. All of the examples in higher dimensions employ coordinate transformations, and the schemes will be written in transformed coordinates. The formulation described here is valid for both two- and three-dimensional systems of conservation laws. Only the two-dimensional case will be described. For three-dimensional formulation, one only has to add an extra dimension and the corresponding numerical flux.

It is emphasized here that the time and spatial order of accuracy pertains to the scheme for nonlinear scalar hyperbolic conservation laws with uniform grid spacing, and the spatial accuracy is for numerical solutions that are away from discontinuities. The algorithms are formulated in a finite volume form which, for certain physical problems and grid distributions, can enhance stability and convergence rates for highly clustered or skewed grids and require only a slight modification from the form originally presented in Yee and Kutler [166] and Yee and Harten [41] for generalized geometries. The various related pseudo-finite volume formulations will also be discussed. However all of the two-dimensional calculations presented in this section employ fairly uniform grid spacing and use the original pseudo-finite volume form presented in [41]. A detailed comparison between the finite volume and the pseudo-finite volume formulations [41,66] by T. Aki of the National Aerospace Laboratory in Japan is included in section 5.2.4.

A numerical study in reference [66] for hypersonic flows shows that for fairly uniform or mildly clustered grids, the present finite volume and pseudo-finite volume forms behave the same as in reference [41] for inviscid flows. This is in contrast to the study of Takakura et al. [167] which claimed that their modified form is the correct finite difference formulation for generalized geometries. A comparison between the pseudo-finite volume formulation of Takakura et al. [167] and reference [41] on the same fairly uniform curvilinear grid for a blunt-body computation shows no noticeable difference in resolution.

Consider a two-dimensional system of hyperbolic conservation laws

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = 0. \tag{5.1}$$

Here U, $F(U)$ and $G(U)$ are column vectors of m components.

A generalized coordinate transformation of the form $\xi = \xi(x,y)$ and $\eta = \eta(x,y)$ which maintains the strong conservation-law form of equation (5.1) is given by

$$\frac{\partial \widehat{U}}{\partial t} + \frac{\partial \widehat{F}(\widehat{U})}{\partial \xi} + \frac{\partial \widehat{G}(\widehat{U})}{\partial \eta} = 0, \tag{5.2}$$

where $\widehat{U} = U/J$, $\widehat{F} = (\xi_x F + \xi_y G)/J$, $\widehat{G} = (\eta_x F + \eta_y G)/J$, and $J = \xi_x \eta_y - \xi_y \eta_x$, the Jacobian transformation. See Anderson [3] for a basic discussion. Let $A = \partial F/\partial U$ and $B = \partial G/\partial U$. Then the Jacobians $\widehat{A}$ and $\widehat{B}$ of $\widehat{F}$ and $\widehat{G}$ can be written as

$$\widehat{A} = (\xi_x A + \xi_y B) \tag{5.3a}$$

$$\widehat{B} = (\eta_x A + \eta_y B). \tag{5.3b}$$

Let the eigenvalues of $\widehat{A}$ be $(a_\xi^1, a_\xi^2, ..., a_\xi^m)$ and the eigenvalues of $\widehat{B}$ be $(a_\eta^1, a_\eta^2, ..., a_\eta^m)$. Denote R_ξ and R_η as the matrices whose columns are eigenvectors of $\widehat{A}$ and $\widehat{B}$, and denote R_ξ^{-1} and R_η^{-1} as the inverses of R_ξ and R_η.

Let the grid spacing be denoted by $\Delta\xi$ and $\Delta\eta$ such that $\xi = j\Delta\xi$ and $\eta = k\Delta\eta$. Let $U_{j+\frac{1}{2},k}$ denote some symmetric average of $U_{j,k}$ and $U_{j+1,k}$ (for example, $U_{j+\frac{1}{2},k} = 0.5 * (U_{j+1,k} + U_{j,k})$, or the Roe average [48] for a perfect gas and generalized Roe average of Vinokur [116] for equilibrium real gases). Let $a_{j+\frac{1}{2}}^l$, $R_{j+\frac{1}{2}}$, $R_{j+\frac{1}{2}}^{-1}$ denote the quantities a_ξ^l, R_ξ, R_ξ^{-1} related to $\widehat{A}$ evaluated at $U_{j+\frac{1}{2},k}$. Similarly, let $a_{k+\frac{1}{2}}^l$, $R_{k+\frac{1}{2}}$, $R_{k+\frac{1}{2}}^{-1}$ denote the quantities a_η^l, R_η, R_η^{-1} related to $\widehat{B}$ evaluated at $U_{j,k+\frac{1}{2}}$.

Let

$$\alpha_{j+\frac{1}{2}} = R_{j+\frac{1}{2}}^{-1}(U_{j+1,k} - U_{j,k}) \tag{5.4a}$$

be the difference of the characteristic variables in the local ξ-direction, and

$$\alpha_{k+\frac{1}{2}} = R_{k+\frac{1}{2}}^{-1}(U_{j,k+1} - U_{j,k}) \tag{5.4b}$$

be the difference of the characteristic variables in the local η-direction.

5.1. Description of the Explicit Numerical Algorithms

Fractional-Step Method: By using the Strang-type of fractional-step (time-splitting [97]) method, the one-dimensional second-order TVD scheme in section 4.4.1 can be implemented for the two-dimensional system (5.2) via the local-characteristic approach as follows:

$$\widehat{U}_{j,k}^{n+2} = \mathcal{L}_{\xi}^{h/2}\mathcal{L}_{\eta}^{h}\mathcal{L}_{\xi}^{h}\mathcal{L}_{\eta}^{h}\mathcal{L}_{\xi}^{h/2}\widehat{U}_{j,k}^{n}, \tag{5.5a}$$

where

$$\mathcal{L}_{\xi}^{h}\widehat{U}_{j,k}^{n} = \widehat{U}_{j,k}^{*} = \widehat{U}_{j,k}^{n} - \frac{\Delta t}{\Delta\xi}\left(\widetilde{F}_{j+\frac{1}{2},k}^{n} - \widetilde{F}_{j-\frac{1}{2},k}^{n}\right), \tag{5.5b}$$

$$\mathcal{L}_{\eta}^{h}\widehat{U}_{j,k}^{*} = \widehat{U}_{j,k}^{*} - \frac{\Delta t}{\Delta\eta}\left(\widetilde{G}_{j,k+\frac{1}{2}}^{*} - \widetilde{G}_{j,k-\frac{1}{2}}^{*}\right) \tag{5.5c}$$

with $h = \Delta t$. The functions $\widetilde{F}_{j+\frac{1}{2},k}$ and $\widetilde{G}_{j,k+\frac{1}{2}}$ (omitting the time superscript) are the numerical fluxes in the ξ- and η-directions evaluated at $(j+\frac{1}{2},k)$ and $(j,k+\frac{1}{2})$, respectively. Typically, $\widetilde{F}_{j+\frac{1}{2},k}$ for a non-MUSCL approach in a finite volume formulation can be expressed as

$$\widetilde{F}_{j+\frac{1}{2},k} = \frac{1}{2}\left[\left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}}(F_{j,k}+F_{j+1,k}) + \left(\frac{\xi_y}{J}\right)_{j+\frac{1}{2}}(G_{j,k}+G_{j+1,k}) + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\Big/J_{j+\frac{1}{2}}\right]. \tag{5.6a}$$

The corresponding pseudo-finite volume formulation will be discussed shortly. The explicit form of $R_{j+\frac{1}{2}}$ can be found in references [48,104,116,166]. To simplify the notation, the subscript $j+\frac{1}{2}$ is used to mean $j+\frac{1}{2},k$. The quantities $(\frac{\xi_x}{J})_{j+\frac{1}{2}}$ and $1/J_{j+\frac{1}{2}}$ are given by

$$\left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}} = \frac{1}{2}\left[\left(\frac{\xi_x}{J}\right)_{j,k} + \left(\frac{\xi_x}{J}\right)_{j+1,k}\right]; \qquad \frac{1}{J_{j+\frac{1}{2}}} = \frac{1}{2}\left(\frac{1}{J_{j+1,k}} + \frac{1}{J_{j,k}}\right). \tag{5.6b}$$

Also $(k_1)_{j+\frac{1}{2}} = \left[\frac{\xi_x}{J}\Big/\sqrt{(\frac{\xi_x}{J})^2 + (\frac{\xi_y}{J})^2}\right]_{j+\frac{1}{2}}$ and $(k_2)_{j+\frac{1}{2}} = \left[\frac{\xi_y}{J}\Big/\sqrt{(\frac{\xi_x}{J})^2 + (\frac{\xi_y}{J})^2}\right]_{j+\frac{1}{2}}$ used in $R_{j+\frac{1}{2}}$ are defined, for example, as

$$(k_1)_{j+\frac{1}{2}} = \frac{(\frac{\xi_x}{J})_{j+\frac{1}{2}}}{\sqrt{(\frac{\xi_x}{J})_{j+\frac{1}{2}}^2 + (\frac{\xi_y}{J})_{j+\frac{1}{2}}^2}}. \tag{5.6c}$$

The values ξ_x, ξ_y, η_x and η_y are evaluated by three-point central differences. Here the symbol $J_{j,k}$ is the Jacobian transformation evaluated at $(j\Delta\xi, k\Delta\eta)$. The averaged Jacobians $J_{j+\frac{1}{2}}$ are used here in order to preserve the free stream. Similarly, one can define the numerical flux $\widetilde{G}_{j,k+\frac{1}{2}}$ in this manner.

The $\Phi_{j+\frac{1}{2}}$ function in the ξ-direction for the symmetric or upwind TVD schemes can be expressed in the same way as equations (4.33) and (4.34) with the appropriate eigenvalues a_{ξ}^{l}

and $\alpha_{j+\frac{1}{2}}$ defined above. The situation is similar for the TVB or high-resolution Lax-Friedrichs schemes as discussed in section III. Note that for high-resolution Lax-Friedrichs schemes, there is little savings for each spatial direction via the non-MUSCL approach. However, the savings by the MUSCL approach (to be described next) in each spatial direction is greater. For gas dynamics applications the limiter can be applied to the slope of the conservative variables or to the primitive variables themselves, thus making the extension to multidimensional problems straightforward since one does not have to use the local one-dimensional characteristic variables assumption and in many instances, one does not have to use the complicated truly multidimensional Riemann solvers. Moreover, there is an additional savings from not having to compute α in equation (5.4a) and (5.4b), and not having to compute $R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}$ and $R_{k+\frac{1}{2}}\Phi_{k+\frac{1}{2}}$.

The numerical flux function $\widetilde{F}_{j+\frac{1}{2},k}$ for a MUSCL type approach, together with a Roe-type approximate Riemann solver in a finite volume formulation, for an upwind scheme as described in [39,66,95] can be expressed as

$$\begin{aligned}\widetilde{F}_{j+\frac{1}{2},k} = \frac{1}{2}\Bigg\{&\left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}}\left[F(U^R_{j+\frac{1}{2}}) + F(U^L_{j+\frac{1}{2}})\right] + \left(\frac{\xi_y}{J}\right)_{j+\frac{1}{2}}\left[G(U^R_{j+\frac{1}{2}}) + G(U^L_{j+\frac{1}{2}})\right]\\ &+\widehat{R}_{j+\frac{1}{2}}\widehat{\Phi}_{j+\frac{1}{2}}\Big/J_{j+\frac{1}{2}}\Bigg\}.\end{aligned} \tag{5.7}$$

The values $F(U^R_{j+\frac{1}{2}})$ and $F(U^L_{j+\frac{1}{2}})$ are the flux function F evaluated at $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ respectively, with $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ the same as in equations (3.38), (3.42) and (3.43) and section 4.4.2 for the ξ-direction. The matrix $\widehat{R}_{j+\frac{1}{2}}$ is the matrix whose columns are eigenvectors of $\widehat{A}$ evaluated at the symmetric average of $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$. The vector $\widehat{\Phi}_{j+\frac{1}{2}}$ is a function of $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$, and the corresponding $\phi^l_{j+\frac{1}{2}}$ can be (4.37) for the TVD method (or TVB methods using limiter (3.61)).

The numerical flux function $\widetilde{F}_{j+\frac{1}{2},k}$ for a MUSCL type approach, together with the high-resolution Lax-Friedrichs method in a finite volume formulation can be expressed as

$$\widetilde{F}_{j+\frac{1}{2},k} = \frac{1}{2}\left\{\left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}}\left[F(U^R_{j+\frac{1}{2}}) + F(U^L_{j+\frac{1}{2}})\right] + \left(\frac{\xi_y}{J}\right)_{j+\frac{1}{2}}\left[G(U^R_{j+\frac{1}{2}}) + G(U^L_{j+\frac{1}{2}})\right] + \widehat{\Phi}_{j+\frac{1}{2}}\Big/J_{j+\frac{1}{2}}\right\}. \tag{5.8a}$$

The values $F(U^R_{j+\frac{1}{2}})$ and $F(U^L_{j+\frac{1}{2}})$ are the flux function F evaluated at $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ respectively, with $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ the same as in equations (3.38), (3.42) and (3.43) and section 4.4.2 for the ξ-direction. The limiter function can be applied to the slope of the conservative variables or the primitive variables. The lth element of $\widehat{\Phi}_{j+\frac{1}{2}}$ is

$$\widehat{\phi}^l_{j+\frac{1}{2}} = -\frac{1}{\lambda}(U^R_{j+\frac{1}{2}} - U^L_{j+\frac{1}{2}}). \tag{5.8b}$$

Comparing (5.8) with (5.7) or (5.6), one can see (as discussed above) the dramatic savings by using the Lax-Friedrichs numerical flux instead of the other two approaches.

The numerical flux function $\widetilde{F}_{j+\frac{1}{2},k}$ for a MUSCL-type approach, together with a flux-vector splitting [77] referred to as the TVD flux-vector splitting method in this paper can be expressed as

$$\widetilde{F}_{j+\frac{1}{2},k} = \left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}}\left[F^-(U^R_{j+\frac{1}{2}}) + F^+(U^L_{j+\frac{1}{2}})\right] + \left(\frac{\xi_y}{J}\right)_{j+\frac{1}{2}}\left[G^-(U^R_{j+\frac{1}{2}}) + G^+(U^L_{j+\frac{1}{2}})\right], \qquad (5.9)$$

where $F^{\pm}(U^{R,L}_{j+\frac{1}{2}})$ are evaluated using either the Steger-Warming type [114,117] or van Leer type [115,117] flux-vector splittings. The vectors $U^R_{j+\frac{1}{2}}$ and $U^L_{j+\frac{1}{2}}$ are the same as in equations (3.38) and section 4.4.2. The quantity $F^-(U^R_{j+\frac{1}{2}})$ is the portion of the flux F with negative eigenvalues evaluated at $U^R_{j+\frac{1}{2}}$.

The difference in operations count between (5.6a), and (5.7) and (5.9) is less than 30% for a perfect gas. However, due to an extra evaluation per dimension in the curve fitting between the left and right states in an equilibrium real gas for (5.7) and (5.9), without further approximations, additional computation is required for the MUSCL approach (with the exception of the high-resolution Lax-Friedrichs method (5.8)). The slight advantage of (5.7) and (5.9) over (5.6d) is that (5.7) and (5.9) can be spatially third-order accurate. However, experiences with the third-order case do not show a very visible improvement over the second-order case for problems with discontinuities. Part of the reason is that all TVD-type schemes reduce to first-order at points of extrema regardless of the order of accuracy in smooth regions. Also, because of the similarity in shock resolution and yet higher operations count for equilibrium real gases and nonequilibrium flows of the MUSCL over the non-MUSCL approach using a Roe-type approximate Riemann solver, efforts are concentrated only on the non-MUSCL formulation both for explicit and implicit computations. At present no outstanding advantages or disadvantages between these formulations for a perfect gas have been observed. Further investigation is required along this line before a clearer comparison can be drawn. The performance of (5.8) is yet to be investigated.

Although the finite volume formulation (5.6a) - (5.6c) is recommended, the results obtained in this report used a slightly different formulation. In particular, three formulations (hereafter referred to as the pseudo-finite volume formulations) for the non-MUSCL schemes were investigated and are as follows

$$\widetilde{F}_{j+\frac{1}{2},k} = \frac{1}{2}\left[(\xi_x)_{j+\frac{1}{2}}(F_{j,k} + F_{j+1,k}) + (\xi_y)_{j+\frac{1}{2}}(G_{j,k} + G_{j+1,k}) + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\right] \Big/ J_{j+\frac{1}{2}} \qquad (5.10a)$$

with the corresponding quantities $(\xi_x)_{j+\frac{1}{2}}$, $J_{j+\frac{1}{2}}$ and $(k_1)_{j+\frac{1}{2}}$ of equations (5.6b,c) expressed as

$$(\xi_x)_{j+\frac{1}{2}} = \frac{1}{2}[(\xi_x)_{j+1,k} + (\xi_x)_{j,k}]; \quad J_{j+\frac{1}{2}} = \frac{1}{2}[J_{j+1,k} + J_{j,k}], \qquad (5.10b)$$

$$(k_1)_{j+\frac{1}{2}} = \frac{(\xi_x)_{j+\frac{1}{2}}}{\sqrt{(\xi_x)^2_{j+\frac{1}{2}} + (\xi_y)^2_{j+\frac{1}{2}}}}, \qquad (5.10c)$$

and

$$\widetilde{F}_{j+\frac{1}{2},k} = \frac{1}{2}\left[(\xi_x F + \xi_y G)_{j,k} + (\xi_x F + \xi_y G)_{j+1,k} + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}}\right] \Big/ J_{j+\frac{1}{2}} \tag{5.11}$$

and

$$\widetilde{F}_{j+\frac{1}{2},k} = \frac{1}{2}\left[\widehat{F}_{j+1,k} + \widehat{F}_{j,k} + R_{j+\frac{1}{2}}\Phi_{j+\frac{1}{2}} \Big/ J_{j+\frac{1}{2}}\right]. \tag{5.12}$$

Here $J_{j+\frac{1}{2}}$ and k_1 in equations (5.11) and (5.12) are the same as (5.10b,c). These forms are used by some practitioners in CFD. For highly skewed grids and nonuniform flows, equations (5.6) and (5.10) are preferred over (5.11) and (5.12). However, (5.10) and (5.11) do not preserve freestream whereas equations (5.6) and (5.12) do. All of the results presented in this section use (5.11). One of the blunt-body cases was rerun with equation (5.6) and (5.10)-(5.12) and no noticeable difference was observed. We expect all of the conclusions on the behavior of (5.10)-(5.12) to be carried over to equation (5.6), since all of the examples use mildly clustered yet quite regular and nearly orthogonal grids.

In two dimensions the present pseudo-finite volume formulations can be made 'truly' finite volume by a slight modification of equations (5.10)-(5.12); i.e., on the treatment of k_1 and $J_{j+\frac{1}{2}}$. However, the situation is different in three dimensions where finite volume formulations depart from finite difference formulations. See reference [168] for a discussion.

Predictor-Corrector Method: To preserve the original second-order time-accuracy, besides using the time-splitting approach, one can employ a predictor-corrector method similar to the one discussed in the nonlinear scalar case. With the same notation, a formal extension of the scalar explicit second-order TVD method in predictor-corrector form (via the local-characteristic approach for the nonlinear hyperbolic system (5.2)) in a pseudo-finite volume formulation can be written as

$$U^{(1)}_{j,k} = \widehat{U}^n_{j,k} - \frac{\Delta t}{\Delta\xi}\left(\widehat{F}^n_{j,k} - \widehat{F}^n_{j-1,k}\right) - \frac{\Delta t}{\Delta\eta}\left(\widehat{G}^n_{j,k+1} - \widehat{G}^n_{j,k}\right), \tag{5.13a}$$

$$U^{(2)}_{j,k} = \frac{1}{2}\left\{U^{(1)}_{j,k} + \widehat{U}^n_{j,k} - \frac{\Delta t}{\Delta\xi}\left[\widehat{F}^{(1)}_{j+1,k} - \widehat{F}^{(1)}_{j,k}\right] - \frac{\Delta t}{\Delta\eta}\left[\widehat{G}^{(1)}_{j,k} - \widehat{G}^{(1)}_{j,k-1}\right]\right\}, \tag{5.13b}$$

$$\widehat{U}^{n+1}_{j,k} = U^{(2)}_{j,k} + \frac{1}{2}\left[\frac{\overline{R}_{j+\frac{1}{2}}\overline{\Phi}_{j+\frac{1}{2}}}{J_{j+\frac{1}{2}}} - \frac{\overline{R}_{j-\frac{1}{2}}\overline{\Phi}_{j-\frac{1}{2}}}{J_{j-\frac{1}{2}}}\right] + \frac{1}{2}\left[\frac{\overline{R}_{k+\frac{1}{2}}\overline{\Phi}_{k+\frac{1}{2}}}{J_{k+\frac{1}{2}}} - \frac{\overline{R}_{k-\frac{1}{2}}\overline{\Phi}_{k-\frac{1}{2}}}{J_{k-\frac{1}{2}}}\right]. \tag{5.13c}$$

Here the superscript (1) and the overbar designate the values of the function evaluated at the intermediate solutions $U^{(1)}$ and $\overline{U}$. The value of $\overline{U}$ can be $\overline{U} = \widehat{U}^n_j$ or $\overline{U} = U^{(2)}_j$. The lth element $\phi^l_{j+\frac{1}{2}}$ of the vector $\Phi_{j+\frac{1}{2}}$ has the same form as the scalar case (3.69d) except the $a_{j+\frac{1}{2}}$ is replaced by $a^l_{j+\frac{1}{2}}$ and the $\Delta_{j+\frac{1}{2}}$ is replaced by $\alpha^l_{j+\frac{1}{2}}$ of equation (5.4a). See sections 3.3, 3.6, 3.7 for a discussion of equation (5.13c) and the usage of the scheme (5.13) in general. A finite volume formulation can be achieved by redefining (5.13a,b) and the definition of $J_{j\pm\frac{1}{2}}$ and $J_{k\pm\frac{1}{2}}$. The discussion on the high-resolution Lax-Friedrichs methods earlier in this section is also valid here. However, only the non-MUSCL formulation can be used in this case.

5.2. Time-Accurate Computations by Explicit Methods

Four time-accurate inviscid computations in the transonic and supersonic range for a perfect gas are described here. Hypersonic computations will be presented in the next section. All of the descriptions are summaries of the published works of references indicated in the subsection headings. The first three computations are external flows and the last one is an internal flow problem. Experimental data were available for three of the problems. In each of these three problems, good agreement between the computations and the experimental data were observed. All of the computations use the pseudo-finite volume form described in [41] and use a fixed grid without local grid adoptation or grid refinement.

5.2.1. Shock Wave Diffraction from an Obstacle [161]

There is a continuing interest in determining the diffraction loadings imposed on a stationary object during a blast-wave encounter, since this knowledge is important in designing the structure to survive such an event. The problem is especially suitable for numerical simulation, since experimental setups for such studies are usually very expensive and sometimes impractical.

A generic configuration of a class of objects of interest is shown in figure (5.1). The configuration has a broad base to maximize stability. To reduce the drag force acting on the body, the top is rounded off to minimize vortex shedding and flow separation. The objective is to ensure that the downward force is much larger than the lateral force generated on the body during a blast-wave encounter so that the object would suffer only minimal lateral motion and would not tip over.

Computational Domain and Grid Generation: For illustration purposes, a wedge angle of 40° and a rounded top with a radius of curvature of 0.17 times the base width was chosen for the current study. Different wedge angles and top curvatures were also computed but are not reported here. The flow features depend strongly on the wedge angle and top shape. The grids used were generated by a generalized Schwarz-Christoffel transformation for curved surfaces, and are orthogonal except at both ends of the body. Figure (5.2) shows the computational domain and grid distribution. A CFL number of 0.99 and a normal incident shock of Mach 2.0 in a perfect gas with $\gamma = 1.4$ and $\delta_1 = 0$ ($\psi(z)$ of equation (3.18)) were used in the computations. Roe's average was used to evaluate $R_{j+\frac{1}{2}}$ and $\alpha_{j+\frac{1}{2}}$.

Figure (5.3) shows a selected sequence of the diffraction process using the time-split second-order upwind scheme (4.34a) together with limiter (4.34d) for the nonlinear fields and (4.34g) for the linear fields. Limiter (4.34c) is the most dissipative limiter of the four limiters (4.34c)-(4.34g) and the results are not shown here. Figure (5.3a) shows the distinct formation of the triple point, with the Mach stem and the contact surface emanating from the triple point.

Figures (5.3b)-(5.3d) show the diffraction process as the shock wave traverses the rounded top. The triple point and the contact surface remain very distinct and the Mach stem has evolved into a curved diffracted shock. It remains almost perpendicular to the body surface at the impingement point. The diffracted shock is actually travelling slightly ahead of the incident shock. A careful examination of figures (5.3d)-(5.3f) reveals that the contact surface is beginning to roll up to form vortices. Figure (5.3g) shows the instant when the diffracted shock leaves the trailing edge of the body. Figure (5.3h) shows the fine resolution of the captured reflected shock at the trailing edge of the body and the emergence of a Mach stem located near $x = 0.8$.

Figure (5.4) shows the same sequence of the diffraction process computed using the predictor-corrector symmetric TVD scheme (5.13) together with limiter (4.33e). The shock resolution is almost identical to that of figure (5.3) except with slightly more diffusive slipstreams. Figure (5.5) shows the density contours comparing the upwind scheme (4.34a) using limiters (4.34d,f) with the predictor-corrector symmetric scheme (5.13) using limiters (4.33c)-(4.33e). The cut off appearance of the incident shocks indicates that the incident shock has travelled slightly beyond the right edge of the computational grid. Judging from the density contour plots of figures (5.4) and (5.5), it would appear that the result using (4.33e) is comparable to that of the upwind TVD scheme with limiter (4.34d) for the nonlinear fields and (4.34g) for the linear fields, although the flow field is slightly smoother than the upwind method away from discontinuities. Results from additional numerical experiments not shown here also indicate that the shock resolution of limiter (4.33e) is slightly better than the upwind scheme with limiter (4.34c). Considering that the predictor-corrector variant requires a smaller operations count and allows a larger time step, it offers a very attractive alternative to the upwind TVD scheme.

A comparison was also made between the time-splitting symmetric TVD scheme and the predictor-corrector scheme (5.13) using the same limiters ((4.33c) - (4.33e)). It was found that the predictor-corrector formulation produced slightly sharper shock resolution than the time-split symmetric form.

Numerical Results by the Explicit MacCormack Method: Using a Courant number of 0.99 and the same fine grid of figure (5.2) for the explicit MacCormack method with a smoothing coefficient of 0.2 resulted in negative pressure even before the incident shock reached the top of the body. The Courant number was reduced to 0.5 and the negative pressure did not show up until the incident shock reached $x = 0.182$. A sequence of the results by the explicit MacCormack method is shown in figure (5.6). Figure (5.6d) shows the instant when the negative pressure first appeared. It is observed from the thickness of the shocks and oscillations associated with them that the shock-capturing resolution of the explicit MacCormack method is inferior to that of the TVD schemes.

5.2.2. Shock Wave Diffraction from a Cylinder

In Yee and Kutler [166], a time-split form of the modified-flux approach with added artificial compression was used to simulate a planar moving shock wave impinging on a circular cylinder. The results of Yee and Kutler revealed the need of a better flux limiter for capturing contact discontinuities, and the need of fine enough grid resolution. In 1987, V.Y.C. Young of Martin Marietta recomputed the same problem with an incident shock of Mach 2.81 on a finer grid using the predictor-corrector symmetric TVD scheme with limiter (4.33e) (unpublished work).

Roe's average was used to evaluate $R_{j+\frac{1}{2}}$ and $\alpha_{j+\frac{1}{2}}$, and δ_1 was set to zero in $\psi(z)$ (equation (3.18)). This particular Mach number was chosen because experimental Schlieren photographs were available from A. Bryson of Stanford University. The detailed description of the shock-tube experiment can be found in Bryson and Gross [169].

Due to the symmetry of the problem, only the top half plane consisting of the semi-annular region between the radii of 0.5 and 3.0 was considered. In the angular direction, 362 rays with a uniform angular spacing of half a degree were used. The first and the second rays straddled the symmetry plane, and similarly for the last and the next to last rays. In the radial direction, 202 grid points were distributed with an exponential stretching. The first radial grid point was imbedded into the body as the image point, with the second radial grid point serving as the body grid point.

A Mach 2.81 normal shock wave was initially located at $x = -0.75$. The computation was stopped at approximately the same position as the Schlieren photograph ($x \approx 2$). The results in figure (5.7) show that a fairly detailed flow structure was obtained, especially in the wake region. The locations of the contact discontinuities emanating from the triple points were clearly captured. The contact discontinuities close to the centerline inside the wake region were also captured. The formation of vortices could also be spotted inside the wake. Overall, the result closely duplicated the experimental Schlieren photograph. Because the computation was inviscid, the locations of the discontinuities appeared to differ slightly from those of the experiment.

5.2.3. Complex Shock Reflections from Airfoils at High Angle of Attack [160]

An interesting shock-tube experiment was conducted by Mandella [170] and Mandella and Bershader [171]. A schematic sketch of the experimental apparatus and the flow field are shown in figure (5.9). After the diaphram ruptures, a planar-shock wave of Mach number $M_s = 2$ forms and travels down the shock tube. The shock tube has a 5cm × 5cm inner cross section and its end is open to the ambient atmosphere. In order to keep the shock wave and its induced flow two-dimensional, two parallel plexi-glass plates are attached to the end of shock tube. An NACA 0018 airfoil is mounted between the plates at an angle of attack of 30°. The airfoil is located at a distance of 3 1/3 times the shock tube height away from the shock tube end. The planar shock starts to diffract as soon as it leaves the shock tube, and forms a curved shock wave which finally impinges on the airfoil. The curved shock wave loses its energy in time as it diffracts and the Mach number decreases to approximately 1.5 at the instant of impingement.

Due to the positive angle of attack of the airfoil, an interesting feature of shock reflection is observed. There is a transition from a short moment of compression to expansion on the upper surface, and compression on the lower surface. Along the upper surface, the shock reflection retains the regular reflection up to the compressive region and then makes a rapid transition from regular to Mach reflection. A single Mach reflection forms a triple point from which the Mach stem, contact surface, and reflected shock emanate. Also a vortex starts to form and grows very slowly at the upper nose due to the sudden but mild expansion. Meanwhile, the transition process develops gradually but rather strongly from a regular to a Mach reflection on the compressive lower surface. Eventually, the Mach stem developed on the lower side wraps around the trailing edge of the airfoil and a vortex is generated due to the sudden strong expansion and the sharp trailing edge.

At the exit of the shock tube, another interesting and equally complex flow pattern is observed in the early stages of shock diffraction. The structure of flow fields behind the diffracting shock behaves in a pseudo-stationary way on a 90° convex corner until the Mach wave hits the plane of symmetry. A slipstream emanating from the convex corner rolls into a vortex-spiral. The reflected Mach wave, the incident planar shock, and its diffracted part form a triple point from which the contact surface emanates and ends at the slipstream encircling the vortex. The second shock appears and forms a triangular configuration with the tail of the Prandtl-Meyer fan (the terminator) and the slipstream. After the Mach wave hits the plane of symmetry, most of the features in the flow field retain their structures for some moments. The second shock grows with time and forms a Mach disk after hitting the plane of symmetry.

In order to properly simulate the experiment, the curved shock solution before it reaches the airfoil is required as an initial condition. Also time-dependent boundary conditions along the outer boundary are required. These requirements are met by simulating the whole experimental region from the shock tube to the airfoil. The computational domain would have to cover a large area with a reasonably fine grid throughout the entire domain or with a time-dependent local adaptive grid refinement in order to capture the complex shock structure as time evolved. To avoid unnecessary computations, to confine the physical domain to be smaller with simple fixed grid computations, and to make use of body-fitted coordinates around the airfoil, the numerical simulation is broken into two parts: (a) simulation of the shock tube that contains a 90° convex corner (with a computational domain extended to cover the region where the airfoil is supposed to be located but without the airfoil), and (b) simulation of a time-dependent curved shock interacting with an airfoil. Figures (5.10) and (5.11) show the computational domain for the two parts. The flow field solutions obtained in part (a) are stored and used with bilinear interpolation to obtain the necessary initial and time-dependent boundary conditions for the airfoil simulation. The time-dependent boundary conditions were obtained assuming that the presence of the airfoil had no effect on the flow at the outer boundaries.

The numerical simulation of the experiment is also conducted in a simpler way. The curved moving shock impinging on the airfoil is modelled as a constant-velocity, planar incident shock with approximately the same incident Mach number. Since the shock is planar, the implementation of the boundary condition along the outer boundary is easily handled with the known solutions. The numerical simulations of the curved and planar shock waves interacting with the airfoil will be discussed shortly. A better understanding of the difference in shock structure for the planar and curved shock simulations is useful for the vortex-blade interaction parametric study where a planar-shock wave is preferred. The second-order explicit upwind and symmetric TVD schemes with the Strang-type of time-splitting were applied to these complex physical problems solving the Euler equations of gas dynamics. Roe's average was used to evaluate $R_{j+\frac{1}{2}}$ and $\alpha_{j+\frac{1}{2}}$, and δ_1 was set equal to 0 in all computations.

<u>*Shock Diffraction on a 90° Convex Corner*</u>: Figure (5.12) shows three sequential instants of a diffracting shock wave of Mach 2 on a 90° convex corner. The time intervals are 50 μsec and 100 μsec. Computed density contours from the solution obtained with the second-order explicit time-split upwind TVD scheme (4.34a) together with limiter (4.34g) for the linear fields and (4.34c) for the nonlinear fields at a CFL number of 0.5 are shown in the right-hand column. The computed results agree quite well with the interferograms at each corresponding instant. The interferograms shown in the left column are missing in each corner of the photographs because

of the configuration of the test-section window. The diffracting shock and the second shock are captured within three to four grid points. The thickness of the shocks in the region for $x < 2.2$ is twice that of those in the rest of the region because the grid spacing is 2 times coarser in that area. The contact surface, slipstream, and the Prandtl-Meyer fan are also well captured. Their interaction with the vortex is shown in the first two instants. In the current case of a Mach 2 shock wave, the vortex core is not clearly defined. In fact, as the Mach number of the shock wave increases, the vortex core becomes more diffuse [172]. This appears in the numerical computations as well as in the interferograms. Numerical experiments carried out for this case show that a CFL number greater than 0.5 resulted in instability at the vortex core. Further study on grid resolution and the proper use of limiters for vortex-type flow fields is needed. The third instant (in which the test-section window is shifted downstream) shows a Mach disk formed after the second shock hits the plane of symmetry. The computed result also shows the well-captured Mach disk within 3-4 grid points following the contact surface.

Curved Shock Interaction with an Airfoil: Figure (5.13) shows the numerical results at six sequential instants while the traveling curved shock is progressing over a NACA 0018 airfoil at an angle of attack $\alpha = 30°$. The time intervals are 10 μsec except for the second interval which is 20 μsec. Computed density contours are shown for the solution obtained using the second-order explicit upwind TVD scheme (4.34a) together with limiter (4.34g) for the linear fields and (4.34d) for the nonlinear fields with a CFL number of 0.98. It should be noted that if the limiter given in equation (4.34g) is used for all of the fields, the solution diverges at a very early stage of the interaction. The authors conjectured that limiter (4.34g) is too compressive for the nonlinear fields. It was found that in general, the use of (4.34d) for the nonlinear fields produced sharper shock resolution than (4.34c). The use of (4.34g) for the linear fields produced sharper contact discontinuities than (4.34d). However, the use of (4.34d) for the nonlinear fields for the computation of shock diffraction on a 90° convex corner will give divergent solutions near the corner and the vortex core regions.

The Mach number of the time-dependent curved shock wave is approximately 1.5 at the moment of impingement. A strong vortex and the dense contours of density around it appear in figure (5.13e), since the Mach stem turns over an approximately 180° convex corner. The most inclusive features of the shock interaction with the airfoil are contained in figure (5.13f). At this instant, the foot of the Mach stem proceeds upstream and collides with the incident moving shock on the upper surface. The experimental interferogram and computed density and pressure contour plots at this time are shown in figure (5.14). It should be pointed out that the two little bumps on both the upper and lower surfaces two thirds of the way along the airfoil and the very thin layer wrapping all around the airfoil observed in the interferogram result from vibration of the experimental setup. Also, because the airfoil used in the experiment is a hand-made NACA 0018 airfoil, a slight difference exists between the experimental and computed configurations. The incident and reflected shocks, Mach stems, and contact surface on the lower surface are captured within three grid points. Also the vortices at the upper nose and the trailing edge of the airfoil are well captured. On the upper surface, the experimental results show a weak contact surface, which is not captured in the simulation. By increasing the grid resolution around that region, this slip surface is also expected to be captured. As expected, the contact surfaces cannot be seen in the pressure contour plot because the pressure is continuous across the contact surfaces.

Planar Shock Interaction with an Airfoil: For comparison, a planar incid~~ent~~ shock of Mach 1.5 at the moment of impingement was used to approximately model the curved shock to study the difference in the computed flow field structures. Figure (5.15) shows a schematic representation of the physical plane with its boundaries and initial conditions. The same C-grid as in the previous case is used. The initial condition is implemented in the same manner as in the shock-tube section. Along the outer boundary, the analytic boundary condition is used to trace the location of the moving planar shock as a function of time. The other boundary conditions (along the body and the wake region) are also implemented in the same manner as was used for numerical simulation of the curved incident shock. Figure (5.15) shows four sequential instants of the interferograms compared with the density contour plots of the numerical results obtained for both the curved and planar shocks. The time interval between the first and second interferograms is 20 μsec and the rest are 10 μsec apart. For both cases the numerical results are computed using the time-split second-order explicit upwind TVD scheme at a CFL number of 0.98. The curved shock solutions agree quite well with the interferograms. The planar-shock solutions also compare favorably with the experiment in terms of the shape and location of the discontinuities. There is, however, a slight difference in density level between the two simulations. The same format of comparison is shown in figure (5.16) using the second-order explicit time-split symmetric TVD scheme (4.33a,e) for both the curved and planar shocks. The shock resolution is similar to the upwind TVD scheme except that it is slightly more diffusive. The shock resolution for the symmetric TVD scheme (4.33a,e) is, however, sharper than (4.33a,c) and (4.33a,d), and the upwind TVD scheme (4.33a,c).

The numerical experiments show that both upwind and symmetric TVD schemes are quite stable and accurate even for higher Mach number shocks. The study also shows that for higher-Mach-number cases the symmetric TVD scheme is less sensitive to the numerical boundary condition treatment than the upwind scheme. For $M_s \geq 10$, one has to use a characteristic type of numerical boundary condition for the upwind TVD scheme. Aside from the difference in numerical boundary treatments, one advantage of symmetric TVD schemes over upwind TVD schemes is that the symmetric TVD schemes require less CPU time than the upwind schemes. Computations with the symmetric TVD scheme (4.33a,e) at a CFL number of 0.98 for a planar incident shock of $M_s = 20$ and $\alpha = 30°$ are shown in figure (5.17). The grid used is the same as for the previous cases.

5.2.4. Shock Propagation in a Channel with 90° Bends (T. Aki [162,173,174])

This computation was performed by T. Aki of the National Aerospace Laboratory, Tokyo, Japan. Detailed descriptions of the physical problem and numerical computations are reported in his original papers [162,173,174].

A shock wave transmitted through a bend in a channel modifies its shape as it travels around the bend, although its initial shape is planar and the bend curves smoothly. The modification of the shape results in an irregular force acting on the curved wall and nonuniform flow behind the transmitted shock.

The physical processes taking place during the transmission were little known and only qualitative arguments based on investigations of the shock processes for isolated concave and convex corners had been given. An experimental investigation of the shock process through 90° bends was performed by Takayama et al. in 1977. One can find the details of the experimental setup in reference [175].

A Mach reflection established on the outer wall communicates to a diffracted wave around the inner wall. The triple point moves successively toward the inner wall. The reflected wave emanating from the triple point interacts with a contact surface appearing behind the diffracted wave and eventually hits the inner wall and then proceeds along it.

The photograph (figure (5.18)) is an infinite fringe interferogram at this stage. The incident shock Mach number (at the entrance of the bend) is 2.2 and the test gas is air. The inner radius of curvature of the testing bend is 80 mm and its width is 40 mm. Therefore, the nondimensional inner curvature of the bend based on the width is 2. Takayama et al. concluded that the curvature of the bend must be equal to or greater than 2 in order to obtain a recovery of the planar shock after transmission under their experimental conditions. Generally speaking, the higher the incident shock Mach number is, the larger the curvature of the bend needed for a stable shock transmission; i.e. a recovery of planar shock front. After the stage shown in the photograph (figure (5.18)), the Mach reflected and diffracted shocks merge into one wave and one can in this case observe recovery of a planar shape within the bend itself.

The numerical boundary condition treatments are as follows: Let $j = 1$ be the inflow boundary. If the flow behind the incident shock is supersonic, then the variables on $j = 1$ are fixed at those obtained from the moving shock relations. If it is subsonic, then a procedure similar to that on the wall is employed. The variables on $j = 1$ are updated by using the Riemann invariants for the incoming and outgoing characteristics. The Riemann invariants for the incoming characteristics are postulated as those located far upstream; i.e. those behind the incident shock. At the outflow, the computation terminates when the transmitted shock front or part of it arrives at the outmost grid plane.

At the walls, let $k = 1$ be the grid plane on the inner wall. Initially, the first-order upwind scheme is used to evaluate U_1^* at the inner wall. Since this scheme contains only information for the left-running characteristics, U_1^* must be updated to U_1^{n+1} taking into account the effect of the right-running or reflected characteristic. The Riemann invariant for this wave, the entropy equation, and the boundary condition (vanishing normal velocity) are sufficient for updating U_1^*. Treatment on the outer wall (at $k = 121$ or 131 depending on the grid) is similar to reversing the characteristics to be considered.

In the computation using the pesudo-finite volume formulation, a total grid size of 581×121 is used with 450×121 of the grid located in the bent section. In the case of the time-split upwind TVD schemes using limiter (4.34g), a 400×121 grid in the bent section was used. A CFL number of 0.98, $\delta_1 = 0.1$ and Roe's average were used in all computations. Figures (5.18) and (5.19) show the pressure and density contours compared with experimental data for both the time-split upwind and symmetric TVD schemes using three different limiters. For the upwind TVD schemes, limiters (4.34d), (4.34d,f) and (4.34g) were used, corresponding to figure (5.18a,b,c). For the symmetric TVD schemes, limiters (4.33c), (4.33d) and (4.33e) were used corresponding to figure (5.19a,b,c). The overall performance compares very favorably with experimental data. The symmetric TVD scheme compares closely with the upwind TVD scheme, and with better accuracy than in the airfoil problem discussed earlier [160]. If one sets $\delta_1 = 0$ (equation (4.33b)), a slightly better shock resolution is expected (e.g. see reference [95,118]).

~ 2

Comparison Between the Finite Volume Formulation and the Pseudo-Finite Volume Formulations: Recently, Aki [174] did an extensive comparison between the different pseudo-finite volume formulations (equations (5.10a), (5.11), (5.12) and equation 5 of reference [41]) and the finite volume formulation (5.6a,b,c) using the same problem with two different fairly uniform grid sizes: (a) a 721 × 131 grid with 360 × 131 of the grid located in the bent section and (b) a 701 × 151 grid with 540 × 151 of the grid located in the bent section. In all of the computations, the same TVD method was employed; i.e., the second-order upwind TVD method (4.34) using the time-splitting approach with the superbee limiter (4.34g) for the linear field and the minmod limiter (4.34c) for the nonlinear field.

Figure (5.20) shows the comparison between the finite volume formulation (5.6a) and the various pseudo-finite volume formulations (5.10a), (5.11), (5.12) and equation (6.4) of reference [41] at one instance of the diffraction process using the (701 × 151) grid. No noticeable difference in resolution was observed among the five different formulations, independent of the two different grid sizes.

The main difference in accuracy at the walls for the four time-accurate calculations here is that each of these computations used slightly different outflow and wall numerical boundary conditions. Moreover, a different computer, different computer implementation, and a different density in grid spacing were used in these examples.

5.3. Description of an Implicit Numerical Algorithm for Steady-State Applications

A one-parameter family of explicit and implicit TVD schemes for the two-dimensional system (5.1) can be written as

$$\begin{aligned}&\hat{U}^{n+1}_{j,k} + \frac{\Delta t}{\Delta \xi}\theta\left[\tilde{F}^{n+1}_{j+\frac{1}{2},k} - \tilde{F}^{n+1}_{j-\frac{1}{2},k}\right] + \frac{\Delta t}{\Delta \eta}\theta\left[\tilde{G}^{n+1}_{j,k+\frac{1}{2}} - \tilde{G}^{n+1}_{j,k-\frac{1}{2}}\right]\\ &= \hat{U}^{n}_{j,k} - \frac{\Delta t}{\Delta \xi}(1-\theta)\left[\tilde{F}^{n}_{j+\frac{1}{2},k} - \tilde{F}^{n}_{j-\frac{1}{2},k}\right] - \frac{\Delta t}{\Delta \eta}(1-\theta)\left[\tilde{G}^{n}_{j,k+\frac{1}{2}} - \tilde{G}^{n}_{j,k-\frac{1}{2}}\right]. \qquad (5.14)\end{aligned}$$

Here $\tilde{F}^n_{j+\frac{1}{2},k}$ has the same form as (5.6a) and with the coressponding $\Phi_{j+\frac{1}{2}}$ in $\tilde{F}^n_{j+\frac{1}{2},k}$ has the same form as in the implicit method defined in section 4.5. Here the high-resolution Lax-Friedrichs method and Shu's TVB method as discussed in 3.5.4 can also be used here with similar CPU savings per time step as for the two-dimensional explicit method discussed earlier. The same numerical flux is also applicable for the two-parameter family of time differencings as discussed in section (3.11).

The solution procedures for steady-state calculations in two-dimensional and three-dimensional Euler and Navier-Stokes equations are as follows. For the explicit operator, the convection terms are discretized by a TVD scheme, and the diffusion terms are approximated by a central-difference method. For the implicit operator, a linearized conservative delta form can be constructed. For efficiency, a spatially first-order implicit operator, as in equation (4.44), was employed in most of the experiments. Algorithms for unsteady as well as viscous steady and unsteady blunt body flows will be discussed in the next section. A second-order time-accurate conservative noniterative ADI form will be discussed in section VI.

For steady-state applications, the resulting linearized delta form can be solved by some appropriate relaxation method other than ADI. This is the direction of current research. However, only a conservative noniterative ADI form [74] for the homogeneous PDE's will be described below. For steady-state applications, the numerical solution is independent of the time step. The implicit operator has a regular block-tridiagonal structure and the resulting block tridiagonal matrix is diagonally dominant. One can modify a standard central difference classical shock-capturing code by simply changing the conventional numerical dissipation term into the one designed for the TVD scheme. The only difference in computation is that the current scheme requires a more elaborate dissipation term for the explicit operator. No extra computation is required for the implicit operator. A second-order time-accurate, noniterative, conservative ADI form will be discussed in section VI.

A Conservative Linearized ADI Form for Steady-State Applications: A conservative linearized ADI form of equation (5.14) used mainly for steady-state applications as described in detail in reference [21,74], can be written as

$$\left[I + \frac{\Delta t}{\Delta\xi}\theta H^{\xi}_{j+\frac{1}{2},k} - \frac{\Delta t}{\Delta\xi}\theta H^{\xi}_{j-\frac{1}{2},k}\right]^n E^* = -\frac{\Delta t}{\Delta\xi}\left[\widetilde{F}^n_{j+\frac{1}{2},k} - \widetilde{F}^n_{j-\frac{1}{2},k}\right] - \frac{\Delta t}{\Delta\eta}\left[\widetilde{G}^n_{j,k+\frac{1}{2}} - \widetilde{G}^n_{j,k-\frac{1}{2}}\right], \tag{5.15a}$$

$$\left[I + \frac{\Delta t}{\Delta\eta}\theta H^{\eta}_{j,k+\frac{1}{2}} - \frac{\Delta t}{\Delta\eta}\theta H^{\eta}_{j,k-\frac{1}{2}}\right]^n E^n = E^*, \tag{5.15b}$$

$$\widehat{U}^{n+1} = \widehat{U}^n + E^n, \tag{5.15c}$$

where for the finite volume formulation (5.6a) or the pseudo-finite volume form (5.10a)

$$H^{\xi}_{j+\frac{1}{2},k}E^* = \frac{1}{2}\left[\widehat{A}_{j+1,k}E^*_{j+1,k} + \widehat{A}_{j,k}E^*_{j,k} - \Omega^{\xi}_{j+\frac{1}{2},k}E^*\right] \tag{5.15d}$$

$$\widehat{A}_{j,k} = \left[\left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}} A_{j,k} + \left(\frac{\xi_y}{J}\right)_{j+\frac{1}{2}} B_{j,k}\right] J_{j,k} \tag{5.15e}$$

and $\widehat{A}_{j+1,k} = \left[\left(\frac{\xi_x}{J}\right)_{j+\frac{1}{2}} A_{j+1,k} + \left(\frac{\xi_y}{J}\right)_{j+\frac{1}{2}} B_{j+1,k}\right] J_{j+1,k}$. Whereas for the pseudo-finite volume formulation (5.12) or reference [41]

$$H^{\xi}_{j+\frac{1}{2},k}E^* = \frac{1}{2}\left[\widehat{A}_{j+1,k}E^*_{j+1,k} - \Omega^{\xi}_{j+\frac{1}{2},k}E^*\right] \tag{5.15f}$$

is used. A similar form for $H^{\xi}_{j+\frac{1}{2},k}E^*$ for the pseudo-finite volume formulation (5.11) can also be obtained. Here $\Omega^{\xi}_{j+\frac{1}{2},k}E^*$, $\Omega^{\eta}_{j,k+\frac{1}{2}}E$ can be taken as (hereafter referred to as the full matrix form)

$$\Omega^{\xi}_{j+\frac{1}{2},k}E^* = R_{j+\frac{1}{2},k}\mathrm{diag}\left[\psi(a^l_{j+\frac{1}{2}})\right]R^{-1}_{j+\frac{1}{2},k}(E^*_{j+1,k} - E^*_{j,k}) \tag{5.15g}$$

$$\Omega^{\eta}_{j,k+\frac{1}{2}}E = R_{j,k+\frac{1}{2}}\mathrm{diag}\left[\psi(a^l_{k+\frac{1}{2}})\right]R^{-1}_{j,k+\frac{1}{2}}(E_{j,k+1} - E_{j,k}). \tag{5.15h}$$

Here $A_{j,k}$ and $B_{j,k}$ are Jacobians of F and G evaluated at (j,k) and $\widehat{A}_{j+1,k}$, $\widehat{B}_{j,k+1}$ are Jacobians of $\widehat{F}$ and $\widehat{G}$ evaluated at $(j+1,k)$ and $(j,k+1)$. The value $E^n_{j,k} = \widehat{U}^{n+1}_{j,k} - \widehat{U}^n_{j,k}$. The expression diag($z^l$) denotes a diagonal matrix with diagonal elements z^l.

The nonconservative linearized implicit form suitable for steady-state calculations [74] is also considered. Numerical study indicated that the latter form appears to be slightly less efficient in terms of convergence rate than the linearized conservative form (5.15). This conservative linearized implicit operator as well as the nonconservative linearized implicit operator (both suggested in reference [74]) for $\theta = 1$ was rediscovered two year later by Barth [176] as the approximate Jacobian linearization. To compute (5.15g,h), a triple matrix multiplication of dimension (4×4) has to be performed at every grid point. For steady-state applications, one can simplify (5.15g,h) as (hereafter referred to as the diagonal form)

$$\Omega^{\xi}_{j+\frac{1}{2},k} E = \mathcal{M}_{\xi} I (E_{j+1,k} - E_{j,k}) \tag{5.15i}$$

$$\Omega^{\eta}_{j,k+\frac{1}{2}} E = \mathcal{M}_{\eta} I (E_{j,k+1} - E_{j,k}), \tag{5.15j}$$

The scalar values $\mathcal{M}_{\xi}$ and $\mathcal{M}_{\eta}$ are the spectral radii of $\Omega^{\xi}_{j+\frac{1}{2},k}$ and $\Omega^{\eta}_{j,k+\frac{1}{2}}$:

$$\mathcal{M}_{\xi} = \max_{l} \psi(a^{l}_{j+\frac{1}{2}}) \tag{5.15k}$$

$$\mathcal{M}_{\eta} = \max_{l} \psi(a^{l}_{k+\frac{1}{2}}) \tag{5.15l}$$

and I is the identity matrix. Note that (5.15i,j) involve scalar multiplication only. The solution using (5.15i,j) is still second-order (or third-order) accurate after it reaches steady state. Other linearized implicit forms can be found in references [21,41,74].

All of the inviscid calculations shown in this section use the diagonal form for the subsonic, transonic and supersonic steady-state applications. A two-parameter family of time differencings with a similar type of linearized forms suitable for time-accurate calculations is reported in references [21,66,74] and is summarized in section VI. In section VI, both the full matrix form and the diagonal form are used for hypersonic steady viscous blunt-body computations. Some unsteady computations will also be described.

5.4. Subsonic, Transonic and Supersonic Steady-State Computations by an Implicit Method

The numerical experiments presented in this section were mainly performed on a NACA 0012 airfoil using the local-characteristic approach with $\delta_1 = 0.125$. For subsonic to low supersonic perfect gas flows, the resolution of the shock waves was found to be quite insensitive to δ_1 for $0.1 \le \delta_1 \le 0.125$ and a constant value seems to be sufficient. However, for hypersonic flows, especially for blunt-body flows, a constant δ_1 was found to be insufficient and a variable δ_1 depending on the spectral radii of the Jacobian matrices of the fluxes is needed. Moreover, a proper choice of the entropy parameter δ_1 for higher Mach number flows not only helps in preventing nonphysical solutions but can act, in some sense, as a control in the convergence rate and in the sharpness of shocks and slip surfaces (or shear layer in viscous flows). The smaller the nonzero δ_1 being used, the slower is the convergence rate. The larger the δ_1 being used, the larger is the numerical dissipation being added. See reference [66] or section VI for numerical examples. All the computations discussed here use the pseudo-finite volume form described in reference [41].

Generally, for inviscid time-accurate calculations, the non-MUSCL upwind TVD schemes produce sharper shocks than symmetric TVD schemes [8]. For the current implicit symmetric TVD scheme with limiter (4.33c) or (4.33d), this seems to be not the case for airfoil problems containing shock waves only. The symmetric method appeared to produce results almost identical to those from an upwind TVD scheme (4.34a)-(4.34c).

Numerical studies also show that there is no difference in resolution in using limiter (4.33c) or (4.33d) for the symmetric TVD scheme. Limiter (4.33e) produces slightly sharper shocks than (4.33c) and (4.33d). This conclusion was based on a numerical study for flow-field conditions ranging from subcritical to transonic and supersonic for the NACA 0012 airfoil. Also, since these test cases consist of shock waves only, the same limiter was used for all characteristic fields. Figures (5.21) and (5.22) show a comparison of the current method using limiter (4.33c) with the upwind scheme (5.15) for two inviscid steady-state airfoil calculations. The two solutions are almost indistinguishable. For the current calculations, the upwind TVD scheme requires approximately 35% more CPU time than the symmetric TVD scheme on a CrayX-MP computer.

Figures (5.23) and (5.24) show an inviscid comparison of the symmetric TVD scheme with the widely distributed computer code ARC2D, version 150 [177]. The free stream Mach numbers are $M_\infty = 1.2$ and 1.8, and the angle of attack is $\alpha = 7^\circ$. The pressure coefficient distributions (not shown) are identical between the two methods and yet the flow field appears very different. The symmetric TVD scheme gives a very well-ordered flow structure and can still capture the shocks with a coarse grid, especially near the trailing edge of the airfoil. On the other hand, the ARC2D code did rather poorly. The ARC2D, version 150 computer code is based on the Beam-Warming ADI algorithm [28], but uses a mixture of second- and fourth-order numerical dissipation terms. These numerical dissipation terms contain adjustable parameters. The values of the parameters on figures (5.23) and (5.24) are the same values as suggested in reference [177]. Other values of the parameters besides the one used in reference [177] were also studied. What is shown here is representative of the performance of ARC2D for this range of Mach numbers and angles of attack. For subsonic and transonic flow regimes the main advantage of TVD schemes over ARC2D for steady-state calculations is that one can capture the shock in one to two grid points without the associated spurious oscillations near shock waves as oppose to three to four. The flow away from the shock looks very much like that calculated by ARC2D. Note that in general, for two-dimensional blast wave calculations, the symmetric schemes usually can capture the shock in two to three grid points, but the slip surfaces are slightly more diffusive than the upwind TVD schemes as discussed in section 5.3.

The same problem was studied for the upwind TVD scheme, and the results and convergence rates were found to be almost identical to those for the symmetric TVD scheme. For figures (5.23) and (5.24), a residual of 10^{-12} can be reached at about 400-600 steps. ARC2D, however, required only 200-300 steps to converge to the same residual. The main reason for the faster convergence rate of ARC2D is by the absence of flux limiters. However, the gain in accuracy is more than compensated for the slower convergence rate.

Steady-state computations for higher Mach number flows in the hypersonic regime were reported in reference [66] and a summary can be found in the next section. Study shows that the

ADI form of the implicit scheme is very sensitive to initial conditions and numerical boundary condition treatments for blunt-body flows. The convergence rate is in general lower than for subsonic and supersonic flows.

5.5. A Thin-Layer Navier-Stokes Calculation

For steady-state application, a simple algorithm utilizing the TVD scheme for the Navier-Stokes equations is to difference the hyperbolic terms the same way as before, and then central difference the viscous terms. The final algorithm is the same as equation (5.15) except that the spatial central differencing of the viscous term is added to the right-hand side of equation (5.15). To illustrate the performance of this numerical procedure, the scheme was tested on a thin-layer Navier-Stokes airfoil calculation. The numerical solution shown below illustrates that this algorithm produces a fairly accurate solution for the case of an RAE 2822 airfoil calculation. Figure (5.25) is an example of the viscous case for the RAE 2822 airfoil using the implicit upwind TVD scheme using limiter (4.34c). The thin-layer Navier-Stokes equations with the algebraic turbulence model of Baldwin and Lomax [178] are used, and the transition is fixed at 3% of chord. The overall agreement with experiments is quite good. An L_2-norm residual of 10^{-7} can be reached in about 900 steps.

5.6. A 3-D Steady-State Computation by a Point-Relaxation Implicit Method [79]

This section shows an application of the symmetric TVD schemes for a three-dimensional steady-state computation. No evaluation or comparison with other techniques is given. Figure (5.26) shows pressure contours and energy contours in the plane of symmetry of the Aeroassist Flight Experiment (AFE) configuration at $V_\infty = 1429$ m/s, $p_\infty = 60.136\ N/m^2$, $T_\infty = 52.22°$K, $T_w = 300°$K, and $M_\infty = 9.86$. The configuration is a raked, elliptic cone with a circular shoulder. The body has a circular cross section when viewed perpendicular to the raked plane. The vehicle is designed as a test platform for a comprehensive series of experiments to define the flow field of an Aeroassisted Orbital Transfer Vehicle (AOTV) at high altitudes (above 75 km) returning from a Geosynchronous Earth Orbit (GEO).

The numerical method employs a finite-volume, point-relaxation implicit procedure of the symmetric TVD formulation (limiter (4.33d)) of the governing Navier-Stokes equations. Gauss-Seidel iteration is employed across data planes in the sweep direction (from the body, across the captured shock to the inflow boundary and back). Jacobi relaxation is used with respect to discretization within a single data plane. At each cell interface, Roe's averaging is used to define eigenvectors and eigenvalues, and δ_1 is set to a constant value of 0.2. Courant numbers up to 40 can be used to accelerate convergence. But Courant numbers of 1 to 2 must finally be used to damp high-frequency errors. In this computation, a coarse to fine grid sequence was used. The relatively coarse grid size (at close to the converged solution) was 64 cells between the body and the inflow boundary, 39 cells from the nose to shoulder, and 19 cells around the axis from 0° to 180°. The bow shock in this case is almost completely captured in two cells.

VI. NUMERICAL ISSUES FOR HIGH-RESOLUTION SHOCK-CAPTURING SCHEMES FOR INVISCID AND VISCOUS HYPERSONIC FLOWS [66]

Most shock-capturing methods are either inefficient for practical computations or only valid for transonic or supersonic perfect gas calculations. For hypersonic perfect gases, equilibrium real gases or nonequilibrium flows, improvement and modification to existing methods are necessary especially when a wide range of Mach numbers and a complex shock structure exist in the entire flow field. In addition, viscous hypersonic and nonequilibrium flow problems are generally stiff and implicit methods are often preferred over explicit methods. Some of the numerical issues for steady inviscid hypersonic blunt-body flow computations were addressed in [66,179]. A semi-implicit method and a fully implicit method for steady-state nonequilibrium flows were discussed in Yee and Shinn [95] and will be addressed in section VII. A basic study on numerical methods for unsteady inviscid nonequilibrium flows was presented in LeVeque and Yee [86] and in section 3.9. The objective of this section is to summarize our study in reference [66]. The main purpose is to discuss some of the numerical issues and show how one can extend and improve the existing implicit high-resolution shock-capturing schemes [21,41,74,95] for multidimensional compressible Euler and Navier-Stokes equations in the hypersonic, perfect and equilibrium real gas flow regimes.

In particular the following numerical issues are addressed:

1. Some numerical aspects of TVD-type schemes that affect the convergence rate for hypersonic Mach numbers and real gas flows but have negligible effect on low Mach number or perfect gas flows are identified.

2. The performance of the various linearized implicit forms of the schemes similar to the transonic flow study [21,41,74] is reexamined for hypersonic flows.

3. The behavior of these schemes with various temporal differencing but similar spatial discretization for inviscid and viscous flows is investigated. Studies indicated that their behavior in terms of stability and convergence rate is quite different between viscous and inviscid flows. However, with the proper choice of temporal discretization and suitable implicit linearization, these schemes are fairly efficient and accurate for very complex two-dimensional hypersonic inviscid and viscous shock interactions.

The behavior and performance of the schemes included in this section are based on numerical experiments from four of our conference papers [150,179-181] and previous transonic and supersonic studies. Readers are encouraged to refer to these references for more details. Most of the numerical solutions are compared with exact solutions or experimental data, or compared with existing schemes.

6.1. Description of the Numerical Algorithm

The conservation laws for the two-dimensional Navier-Stokes equations can be written in the form

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = \frac{1}{Re}\left[\frac{\partial F_v}{\partial x} + \frac{\partial G_v}{\partial y}\right], \tag{6.1a}$$

where $U = [\ \rho,\ m,\ n,\ e\ ]^T$, $F = [\ \rho u,\ mu+p,\ nu,\ eu+pu\ ]^T$, $G = [\ \rho v,\ mv,\ nv+p,\ ev+pv\ ]^T$, $F_v = [\ 0,\ \tau_{xx},\ \tau_{xy},\ f\ ]^T$, and $G_v = [\ 0,\ \tau_{xy},\ \tau_{yy},\ g\ ]^T$. Here ρ is the density, u and v are the velocity components, $m = \rho u$ and $n = \rho v$ are the x- and y-components of the momentum per unit volume, p is the pressure, $e = \rho[\epsilon + \frac{1}{2}(u^2 + v^2)]$ is the total energy per unit volume, and ϵ is the specific internal energy. For a perfect gas, we also have

$$\tau_{xx} = \mu(4u_x - 2v_y)/3, \tag{6.1b}$$

$$\tau_{xy} = \mu(u_y + v_x), \tag{6.1c}$$

$$\tau_{yy} = \mu(-2u_x + 4v_y)/3, \tag{6.1d}$$

$$f = u\tau_{xx} + v\tau_{xy} + \mu Pr^{-1}(\gamma - 1)^{-1}\frac{\partial a^2}{\partial x}, \tag{6.1e}$$

$$g = u\tau_{xy} + v\tau_{yy} + \mu Pr^{-1}(\gamma - 1)^{-1}\frac{\partial a^2}{\partial y}, \tag{6.1f}$$

where, for example, u_x is defined as $\partial u/\partial x$. The dynamic viscosity μ is given by Sutherland's formula. The Reynolds number is Re, the Prandtl number is Pr, the sound speed is a, and the ratio of specific heats is γ.

Under a generalized coordinate transformation, $\xi = \xi(x, y)$ and $\eta = \eta(x, y)$, equation (6.1) can be written in a form which maintains the strong conservation-law form as

$$\frac{\partial \widehat{U}}{\partial t} + \frac{\partial \widehat{F}}{\partial \xi} + \frac{\partial \widehat{G}}{\partial \eta} = \frac{1}{Re}\left[\frac{\partial \widehat{F}_v}{\partial \xi} + \frac{\partial \widehat{G}_v}{\partial \eta}\right], \tag{6.2}$$

where $\widehat{U} = U/J$, $\widehat{F} = (\xi_x F + \xi_y G)/J$, $\widehat{G} = (\eta_x F + \eta_y G)/J$, $\widehat{F}_v = (\xi_x F_v + \xi_y G_v)/J$, $\widehat{G}_v = (\eta_x F_v + \eta_y G_v)/J$, and $J = \xi_x \eta_y - \xi_y \eta_x$, the Jacobian transformation. Let $A = \partial F/\partial U$ and $B = \partial G/\partial U$. Then the Jacobians $\widehat{A} = \partial \widehat{F}/\partial \widehat{U}$ and $\widehat{B} = \partial \widehat{G}/\partial \widehat{U}$ can be written as

$$\widehat{A} = (\xi_x A + \xi_y B) \tag{6.3a}$$

$$\widehat{B} = (\eta_x A + \eta_y B). \tag{6.3b}$$

In this study the thin-layer Navier-Stokes approximation is assumed by dropping all the $\partial(\cdot)/\partial\xi$ derivatives in the viscous terms. Also, stability and convergence rates for the viscous results are for a perfect gas and laminar flows with adiabatic wall conditions.

In the application of TVD-type schemes for viscous flows, the physical problems considered here are assumed to be inviscid dominated in the sense that moderate or strong shock waves are present in the flow field such that high-resolution shock-capturing techniques are required. Thus the numerical procedures used here for the compressible Navier-Stokes calculations are a second-order central difference approximation for the diffusion terms and TVD-type schemes for the inviscid part of the Navier-Stokes equations. The question of whether the present numerical dissipation term (due to the TVD-type terms) has an adverse effect on the true viscosity terms in the boundary layer region is not known at this point. What we can conclude from the current study is that the portions of the solution slightly or far away from the boundary layer are quite accurately simulated.

To make this section more self-contained, some of the discussion on a two-parameter family of explicit and implicit high-resolution schemes is repeated here. This algorithm is based on a semi-discrete methodology and on the one-parameter family of TVD-type algorithms developed in references [17,41,41,150]. The idea is to use the same spatial discretization as in references [17,21,41,74] for the spatial derivatives and to use a two-parameter family of linear multistep methods for the time derivatives. The original one-parameter family of TVD-type schemes is a subset of the two-parameter family of algorithms. Mathematical analysis similar to that in [17,21,41,74] for the current larger family of schemes requires further investigation. For the one-parameter family of time differencings, these schemes are TVD for the one-dimensional constant coefficient hyperbolic equations. Also the MUSCL approach in conjunction with the Roe-type approximate Riemann solver [95] and TVD flux-vector splitting methods [77] falls nicely into the present framework.

Without loss of generality, the two-parameter family of implicit schemes for the Euler equations ($F_v = G_v = 0$) is presented here. For general Navier-Stokes equations, the appropriate three-point central differences of the viscous Jacobian terms should be added to the implicit operator and a central difference approximation for the diffusion terms should be added to the explicit operator. For time-accurate Navier-Stokes calculations or for faster convergence to steady state, the viscous Jacobian terms were found to be necessary.

Let $\lambda^{\xi} = \frac{\Delta t}{\Delta \xi}$ and $\lambda^{\eta} = \frac{\Delta t}{\Delta \eta}$; then a two-parameter family of explicit and implicit TVD-type algorithms in generalized coordinates for two-dimensional systems (6.1) with $F_v = G_v = 0$ can be written as

$$\Delta \widehat{U}^n_{j,k} + \frac{\lambda^{\xi}\theta}{1+\omega}\left[\widetilde{F}^{n+1}_{j+\frac{1}{2},k} - \widetilde{F}^{n+1}_{j-\frac{1}{2},k}\right] + \frac{\lambda^{\eta}\theta}{1+\omega}\left[\widetilde{G}^{n+1}_{j,k+\frac{1}{2}} - \widetilde{G}^{n+1}_{j,k-\frac{1}{2}}\right]$$

$$= -\frac{(1-\theta)\lambda^{\xi}}{1+\omega}\left[\widetilde{F}^{n}_{j+\frac{1}{2},k} - \widetilde{F}^{n}_{j-\frac{1}{2},k}\right] - \frac{(1-\theta)\lambda^{\eta}}{1+\omega}\left[\widetilde{G}^{n}_{j,k+\frac{1}{2}} - \widetilde{G}^{n}_{j,k-\frac{1}{2}}\right] + \frac{\omega}{1+\omega}\Delta \widehat{U}^{n-1}_{j,k}. \qquad (6.4)$$

Here $\Delta \widehat{U}^n_{j,k} = \widehat{U}^{n+1}_{j,k} - \widehat{U}^n_{j,k}$. The functions $\widetilde{F}_{j+\frac{1}{2},k}$ and $\widetilde{G}_{j,k+\frac{1}{2}}$ are the numerical fluxes in the ξ- and η-directions evaluated at $(j+\frac{1}{2},k)$ and $(j,k+\frac{1}{2})$, respectively. The numerical fluxes $\widetilde{F}_{j+\frac{1}{2},k}$ and $\widetilde{G}_{j,k+\frac{1}{2}}$ have the same meaning and form as in section V. The MUSCL and non-MUSCL approaches as discussed in section 4.5 for the implicit methods can be applied to each of the local ξ or η characteristic directions. This two-parameter family of algorithms contains first- and second-order implicit as well as explicit schemes. The scheme is temporally second-order if $\theta = \omega + \frac{1}{2}$ and first-order otherwise. When $\theta \neq 0$, algorithm (6.4) is an implicit scheme. In this paper, only the temporally first-order backward Euler ($\theta = 1$, $\omega = 0$) and the temporally second-order three-point backward differentiation ($\theta = 1$, $\omega = 1/2$) time differencing are investigated. The trapezoidal formulation ($\theta = 1/2$, $\omega = 0$) was tested on the same problem with an undesirable convergence rate and a smaller time step than the other two time differencings for steady-state computations and it is less stable for time-accurate computations. For the purpose of the current study, the trapezoidal formulation was ruled out for these reasons.

Detailed formulation and numerical studies for algorithm (6.4) with $\omega = 0$ for transonic, supersonic and hypersonic flows can be found in references [1,21,41,74,150,179-181]. The current study using the symmetric and upwind TVD-type scheme of (4.42) and (4.43) shows that, for viscous steady and unsteady flows, the temporally second-order implicit algorithm ($\theta = 1$, $\omega = 1/2$) appears to be slightly more stable and efficient than the temporally first-order implicit algorithm ($\theta = 1$, $\omega = 0$). Note that ω in (6.4) is different from the ω in (3.38).

6.2. A Conservative Linearized Implicit Form for Unsteady and Steady-State Calculations

To solve for U^{n+1} in (6.4) one normally needs to solve a set of nonlinear algebraic equations iteratively. One way to avoid this is to linearize the implicit operator and solve the linearized form by other means. Following the same procedure as in references [21,41,74], a conservative linearized (ADI) form of (6.4) for the numerical fluxes (5.6a) and (5.7) can be written as

$$\left[I + \frac{\lambda^{\xi}\theta}{1+\omega}H^{\xi}_{j+\frac{1}{2},k} - \frac{\lambda^{\xi}\theta}{1+\omega}H^{\xi}_{j-\frac{1}{2},k}\right]^{n} E^{*} =$$

$$-\frac{\lambda^{\xi}}{1+\omega}\left[\widetilde{F}^{n}_{j+\frac{1}{2},k} - \widetilde{F}^{n}_{j-\frac{1}{2},k}\right] - \frac{\lambda^{\eta}}{1+\omega}\left[\widetilde{G}^{n}_{j,k+\frac{1}{2}} - \widetilde{G}^{n}_{j,k-\frac{1}{2}}\right] + \frac{\omega}{1+\omega}\Delta\widehat{U}^{n-1}_{j,k}, \tag{6.5a}$$

$$\left[I + \frac{\lambda^{\eta}\theta}{1+\omega}H^{\eta}_{j,k+\frac{1}{2}} - \frac{\lambda^{\eta}\theta}{1+\omega}H^{\eta}_{j,k-\frac{1}{2}}\right]^{n} E^{n} = E^{*}, \tag{6.5b}$$

$$E^{n} = \Delta\widehat{U}^{n}; \qquad \widehat{U}^{n+1} = \widehat{U}^{n} + E^{n}, \tag{6.5c}$$

where $H^{\xi}_{j+\frac{1}{2},k}$ and $H^{\eta}_{j,k+\frac{1}{2}}$ have the same form as in section V, equations (5.15d) - (5.15j).

6.3. General Assumptions and Limitations on the Numerical Studies

The present study is by no means an exhaustive investigation. There are additional elements and parameters (other than the ones considered here) in the algorithm itself as well as in the physical problem, such as flow type and geometric complexity, that can affect or interfere with the performance of the numerical scheme. Even within the numerical issues listed in the introduction of this section, the study is limited to a sampling of the parameter range for the time-step size or CFL number and the form of the entropy parameter δ_1 in (4.33b). In particular, various strategies to speed up and stabilize the start-up solution from freestream conditions for steady computations have not been investigated. What is discussed here is intended to give interested readers some guidelines for the use of the algorithm. All of the numerical studies discussed in the subsequent sections rely on the following assumptions and considerations:

1. The numerical results and conclusions are for the non-MUSCL approach and for the particular flow type and geometry specified with a sampling of a narrow range of Mach numbers and time steps. The study was concentrated on the implicit scheme (6.5) together with (4.42) and (4.43). Results for viscous flow calculations are based on the shock wave dominated thin-layer Navier-Stokes equations for laminar flows.

2. The numerical boundary condition treatments are the same as in references [41,75,177] for the inviscid flow and as in reference [180] for the viscous flows. Normally a lower-order

upwind scheme or extrapolation near the boundary was used. Studies [10,42] showed that proper treatment of numerical boundary conditions has a major impact on the stability and convergence rate of the scheme. Therefore the boundary condition treatment used here affects the performance of the stability, accuracy and convergence rate of the present algorithm.

3. For steady-state computations, the convergence rate depends not only on all of the elements and parameters (to be discussed shortly), but more importantly also on the type of grid associated with the computation. Studies show that, in general, the numerical solution for a coarse nearly uniform orthogonal grid converges 1-3 times faster than similar finer grids, and possibly an order of magnitude or more faster than highly clustered or skewed grids. What will be presented in section 6.7 represents fairly uniform to mildly clustered grids. Most of the grids used for the numerical study were not very coarse; thus the number of iterations quoted is naturally higher than for the coarse grid counterpart.

4. For the non-interfering blunt-body flows, the convergence rate and behavior of the symmetric and upwind TVD-type schemes are very similar. However, for the interfering blunt-body flows containing slip or shear surfaces, the upwind scheme produces sharper weak solutions. Consequently, all of the illustrations and conclusions discussed in this section are for the upwind scheme (4.43) using limiter (4.34c). Other limiters can produce sharper discontinuities but are not as robust as limiter (4.34c). See section 6.7 or reference [1] for a discussion.

5. Research on equilibrium real gas effects on the performance of these schemes is only in the preliminary stage. All of the illustrations and conclusions for equilibrium real gases are for inviscid non-interfering blunt-body flows. Studies of viscous equilibrium real gas flows are in progress.

6. For steady-state computations using the backward Euler time differencing ($\theta = 1$, $\omega = 0$), a local time-stepping procedure similar to [33,177] was used. However, in comparing the convergence rate with the three-point backward differentiation time differencing ($\theta = 1$, $\omega = 1/2$) for the viscous flows, a constant time step was used.

Other issues such as reducing ADI factoring error, using multigrid, relaxation or conjugate gradient methods as an alternative to ADI, using local grid refinement to enhance resolution, etc., which are also sources of consideration to algorithm (6.5), are not considered here. These issues and the development of better algorithms are the subject of on-going research.

6.4. Enhancement of Stability and Convergence Rate for Hypersonic Flows

In reference [179] some elements and parameters which can affect the stability and convergence rate in high Mach number cases but which have negligible effect in low Mach number cases for steady-state inviscid blunt-body flows were identified. Further study [66] using scheme (6.5) together with the Harten and Yee upwind scheme (4.43) and the Yee-Roe-Davis symmetric scheme (4.42) indicated that the same elements and parameters can affect the stability and convergence rate at hypersonic speeds for viscous computations as well. They are: (1) the choice of the entropy-correction parameter δ_1, (2) the choice of the dependent variables on which the limiters are applied (related to proper scaling of the eigenvectors for high speed flows), and (3) the prevention of unphysical solutions during the initial transient stage. Our study indicates that these elements can also improve the stability of unsteady as well as steady hypersonic flows.

1. For Mach number ranging from 1.2 to 15, numerical experiments for one- and higher-dimensional unsteady flows containing unsteady shocks show that the second-order explicit TVD schemes (5.5) and (5.13) [1,75,150] are insensitive to the entropy correction for $0 \leq \delta_1 \leq 0.1$. In most cases $\delta_1 = 0$ was used. For $0.1 \leq \delta_1 \leq 0.25$, there is a possibility of improving stability in the sense of allowing a higher CFL number at the expense of a slight smearing of the discontinuities. However, for unsteady complex shock wave interactions, a small positive δ_1 or a variable δ_1 (to be discussed) can help stabilize the time-accurate implicit algorithm (6.5).

For subsonic to low supersonic steady-state NACA 0012 airfoil computations [21,41], the resolution of the shock waves was found to be quite insensitive to $0.1 \leq \delta_1 \leq 0.125$ and a constant value seems to be sufficient. However, for hypersonic flows, especially for blunt-body flows, a constant δ_1 or a variable δ_1 suggested by Harten and Hyman [53] was found to be insufficient, but a variable δ_1 depending on the spectral radius of the Jacobian matrices of the fluxes is very helpful in terms of stability and convergence rate. In fact, a proper choice of the entropy parameter δ_1 for higher Mach number flows not only helps in preventing nonphysical solutions but can act, as a control of the convergence rate and of the sharpness of shocks and slip surfaces (or shear layer in viscous flows). The smaller the δ_1 that is used, the slower is the convergence rate. The bigger the δ_1 that is being used, the larger is the numerical dissipation being added. However, δ_1 cannot be arbitrarily large.

For the present blunt-body steady-state calculations with Mach numbers $M > 4$, the initial flow conditions at the wall are obtained using the known wall temperature in conjunction with pressures computed from a modified Newtonian expression [182]. Also, for implicit methods a slow startup procedure from initial conditions [177] is necessary. Most importantly, experience indicates that if one sets δ_1 in equation (4.33b) as a function of the velocity and sound speed, i.e.,

$$(\delta_1)_{j+\frac{1}{2}} = \tilde{\delta}(|u_{j+\frac{1}{2}}| + |v_{j+\frac{1}{2}}| + c_{j+\frac{1}{2}}) \tag{6.6a}$$

$$(\delta_1)_{k+\frac{1}{2}} = \tilde{\delta}(|u_{k+\frac{1}{2}}| + |v_{k+\frac{1}{2}}| + c_{k+\frac{1}{2}}) \tag{6.6b}$$

with $0.05 \leq \tilde{\delta} \leq 0.25$, then blunt-body flows for $4 \leq M \leq 25$ appear to be stabilized and nonphysical solutions are less likely to occur. Equation (6.6) is written in Cartesian coordinates. In the case of generalized coordinates, the u and v should be replaced by the contravariant velocity components, and one half of the sound speed would be from the ξ-direction and the other half would be from the η-direction. For implicit methods, it is very important to use (6.6) in $\psi(z)$ on both the implicit and explicit operators since in a two-dimensional hypersonic flow field consisting of a mixture of subsonic and supersonic regions with Mach number ranging from 0 to hypersonic speeds, an entropy parameter like (6.6) is nonzero in all of the regions. The entropy parameter (6.6) seems to work well for blunt-body flows but whether this is also the right choice for configurations other than a blunt-body shape is an open question.

Recently, Müller [183] suggested a slightly different approach for viscous computation. At the present stage of development, the entropy parameter seems still highly geometric and problem dependent. A universal method is still yet to be discovered. Maybe the use of a high-resolution Lax-Friedrichs method (see sections III and IV) might be more robust in this respect at the expense of a slight degradation in the shock resolution. Again, extensive numerical experiments on a variety of geometries and flow conditions are needed before a conclusion can be drawn.

For unsteady hypersonic blunt body complex shock wave interactions, the entropy parameter (6.6) can help stabilize the time-accurate implicit algorithm. For most of the viscous and inviscid calculations shown, unless otherwise indicated, $\tilde{\delta}$ is set to 0.125.

2. Higher-order TVD schemes in general involve limiter functions. However, there are options in choosing the types of dependent variables when applying limiters for systems of hyperbolic conservation laws, in particular for systems in generalized coordinates. The choice of the dependent variables on which limiters are applied can affect the stability and convergence process. In particular, due to the nonuniqueness of the eigenvectors $R_{j+\frac{1}{2}}$, the choice of the characteristic variables on which the limiters are applied plays an important role in the stability and convergence rate as the Mach number increases. This is directly related to proper scaling of the eigenvectors for high speed flows. For moderate Mach numbers, the different choices of the eigenvectors have a negligible effect on the stability and convergence rate. However, for large Mach number cases, the magnitudes of all the variables at the jump of the bow shock are not the same. In general, the jumps are much larger for the pressures than for the densities or total energy. Studies indicated that employing the form $R_{j+\frac{1}{2}}$ such that the variations of the α are of the same order of magnitude as the pressure would be a good choice for hypersonic flows. The form similar to the one used by Gnoffo [184] or Roe and Pike [185] can improve the convergence rate over the ones used in references [104,166]. In all of the computations shown the form $R_{j+\frac{1}{2}}$ used is the same as in references [104,166] except for an extra factor of $1/c^2_{j+\frac{1}{2}}$. With this extra factor the variations of the α are in fact proportional to the pressure. Other forms of $R_{j+\frac{1}{2}}$ have not been investigated. With this type of scaling, the stability criterion is less restricted even for unsteady hypersonic speeds.

3. Due to the large gradients and to the fact that the initial conditions are far from the steady-state physical solution, the path used by the implicit method can go through states with negative pressures if a large time step is employed. A convenient way to overcome this difficulty is to fix a minimum non-negative allowed value for the density and the pressure. With this safety check, the scheme allows a much larger time step and converges several times faster.

In addition, since the Roe's average state allows the square of the average sound speed $c^2_{j+\frac{1}{2}}$ to lie outside the interval between c^2_j and c^2_{j+1} for equilibrium real gases, the average state $c^2_{j+\frac{1}{2}}$ might be negative even though c^2_j and c^2_{j+1} are positive during the transient stage when the initial conditions are far from the steady-state physical solution. In this case, we replace $c^2_{j+\frac{1}{2}}$ by $\max(c^2_{j+\frac{1}{2}}, \min(c^2_j, c^2_{j+1}))$. This latter safety check is in particular helpful for the symmetric TVD algorithm (4.42) but not necessary for the upwind algorithm (4.43).

6.5. Behavior of the Algorithm with Different Temporal Differencing

It is emphasized here that since the method (6.5) is written in the 'delta' formulation, both the backward Euler (first-order) and the three-point backward differentiation (second-order) time discretizations require the same amount of storage and a similar operations count. Therefore, the main consideration between the two time-differencing methods is their relative stability and convergence rate. The general behavior of the algorithm concluded here pertains to the explicit and implicit upwind and symmetric scheme (6.5) of Harten and Yee, and Yee-Roe-Davis using (4.33), (4.34), (4.42) and (4.43).

Inviscid Unsteady Flows: For inviscid unsteady flows, the explicit TVD-type methods [1,75,150] (section 5.1) are more efficient than the second-order implicit method (6.5). Unless the inviscid problem is stiff, there is no advantage of employing an implicit method for inviscid unsteady flows.

Inviscid Steady Flows: The backward Euler implicit method has a better stability and convergence rate than the three-point backward differentiation implicit method. Also a local time-stepping procedure [18,177] can speed up the convergence rate for the former time-differencing method whereas the same procedure has little effect on the convergence rate when compared with a fixed time-step procedure for the latter time-differencing method.

Viscous Unsteady Flows: Computations of unsteady viscous flows mainly use the second-order time differencing since a larger time step can be used (but still remain time-accurate at least slightly away from the wall) compared with the temporally first-order implicit method. Due to the highly clustered viscous grid used in contrast to their inviscid counterpart, solving a viscous unsteady complex shock interaction using an explicit TVD-type method is not practical due to its inherent time-step restriction. In certain cases, the time step might be an order of magnitude smaller than the implicit counterpart. A more detailed illustration of unsteady viscous hypersonic blunt-body computations with an impinging shock is reported in reference [181].

Viscous Steady Flows: At present there is no detailed viscous steady flow study comparing the first-order time differencing using a local time-stepping approach with the second-order time differencing using a constant time-step approach. But the general trend is that the second-order time differencing has slightly better stability and convergence rate than the first-order one. In particular, a summary using a sequence of fixed time-step approaches comparing the two time-differencing algorithms is discussed in section 6.6.

6.6. Numerical Results

The various numerical aspects discussed in the previous sections are evaluated in part by experience gained from one-dimensional shock tube results [150], from two-dimensional subsonic, transonic, supersonic and hypersonic speed flows results [1,21,41,74,179] and in part by a variety of two-dimensional steady and unsteady, viscous and inviscid hypersonic blunt-body flow computations in [66]. Some of the numerical studies are compared with shock-fitting computations and some are compared with experiments when available. More detailed computation of the physics of these problems can be found in references [180,181]. Six types of blunt-body test cases are illustrated in figures 6.1 - 6.14. Test cases 1 and 2 are inviscid, perfect gases and equilibrium real gases, non-interfering blunt-body flows. Test case 3 is a steady inviscid, perfect gas blunt-body flow with an impinging shock. Test cases 4-6 are viscous, steady and unsteady perfect gas blunt-body flows with and without impinging shocks. Unless indicated, all of the results and discussions for the viscous computations are based on the conservative linearized full matrix form (6.5) together with (5.15g,h) and the Harten and Yee numerical flux (4.43). From here on equation (6.5) means (6.5) together with (5.15g,h) and (4.43). If (5.15i,j) is used, it is referred to as the diagonal form.

Comparison Among the Various Linearized Implicit Methods: For the implicit operator, numerical experiments show that the linearized conservative form (6.5) converges slightly faster (while requiring almost identical CPU time per time step) than the linearized nonconservative form [186] for both viscous and inviscid flows. It appears also that when the freestream Mach number increases, the convergence rate of the linerarized conservative form (6.5) is better than a simplified version which replaces $\Omega^{\xi}_{j+\frac{1}{2},k}$ and $\Omega^{\eta}_{j,k+\frac{1}{2}}$ of (5.15g,h) by $\max_l\ \psi(a^l_{j+\frac{1}{2}})$ and $\max_l\ \psi(a^l_{k+\frac{1}{2}})$ times the identity matrix (equation (5.15i,j)). This is especially true for viscous flow computations. Due to the experience gained from the transonic and the inviscid hypersonic study, no detailed computations using the linearized nonconservative form were performed for viscous steady flow. All of the results and discussions for the viscous computations are based on the conservative linearized form.

Another area of investigation is that for viscous computations, the Jacobians of the viscous terms in the implicit operators are rather expensive to compute. For unsteady flows, these terms are needed to maintain the spatial order of accuracy. Whether the omission of these terms has a major impact on the stability and convergence rate of the algorithm for steady-state calculations is not known. Therefore, an investigation has been made on the difference in the convergence rate for the algorithm with or without the viscous terms in the implicit operator. A brief summary is included in one of the following subsections.

Choice of Limiters: Unlike flows with transonic and low supersonic shock waves, problems containing strong hypersonic shock waves are more sensitive to the treatment of limiters. Using the more diffusive limiter (4.33c) or (4.34c) turns out to be more stable than other more compressive limiters. In terms of shock resolution for both the symmetric and upwind TVD-type of schemes, the sequences written in equations (4.33c)-(4.33e) and (4.34c)-(4.34g) are in order of increasing accuracy. On the other hand, these sequences are in order of decreasing stability and convergence rate. The more compressive limiters like (4.34f) and (4.34g) have a very low stability and slow convergence rate for steady flows. The same conclusion applies for unsteady flows where the more compressive limiters have a very restricted time-step limit. From our experiences, it is not advisable to use (4.34f) and (4.34g) for complex steady shock wave interactions. In particular, limiter (4.34g) should be used only for the linear fields (i.e., for the u and v characteristic fields in the x- and y-direction respectively). This study was conducted in reference [1].

Convergence Rate of Explicit and Implicit TVD-type Schemes for Equilibrium Real Gas Flows: The five different second-order TVD methods previously studied [150] in one dimension yield very similar shock resolution for the blunt-body (non-interfering case) problem. In particular, for an inviscid blunt-body flow in the hypersonic equilibrium real gas range, the explicit second-order Harten and Yee, and Yee-Roe-Davis type TVD schemes [8,9,17] using a generalized approximate Riemann solver of Roe produce similar shock resolution but converge slightly faster (with comparable CPU-time per time step) than an explicit second-order van Leer type scheme using the generalized van Leer flux-vector splitting [150].

The freestream conditions for the current study are $M_\infty = 15$ and 25, $p_\infty = 1.22 \times 10^3\ N/m^2$, $\rho_\infty = 1.88 \times 10^{-2}\ kg/m^3$, and $T_\infty = 226$K. Figure 6.1 shows the upper half of the 61 × 33 grid used for the blunt-body problem. For the $M_\infty = 25$ case, the shock standoff distance is approximately fourteen points from the wall on the symmetry axis. The relaxation procedure for the explicit methods employs a second-order Runge-Kutta time discretization [179] with a CFL of 0.5 (solution not shown). The parameter $\tilde{\delta}$ is set to a constant value of 0.15. Pressure and Mach number contours (compared with shock-fitting solutions) converge and stabilize after 3000-4000 steps but the convergence rate is much slower for the density (with a 2-3 order of magnitude drop in the L_2-norm of the residual). The bow shock is captured in two to three grid points. The curve fits of Srinivasan et al. [164] are used to generate the thermodynamic properties of the gas.

The same flow condition was tested on the implicit scheme (6.5) and the convergence rate was found to be many times faster. Figures 6.2 and 6.3 show the Mach number, density, pressure and κ contours computed by the linearized conservative ADI form of the upwind scheme (6.5) with the first-order backward Euler ($\theta = 1$ and $\omega = 0$) for Mach numbers 15 and 25. Figure 6.4 shows the slight advantage of the convergence rate of the linearized conservative implicit scheme (6.5) over the linearized nonconservative implicit scheme (with $\theta = 1$, $\omega = 0$ and a fixed CFL of 15) suggested in reference [186]. The convergence rate and shock resolution for the symmetric TVD-type scheme (6.5) behave similarly. For $M_\infty = 15$ case, the L_2-norm residual stagnated after a drop of four orders of magnitude.

In general, for a perfect gas with $10 \leq M_\infty \leq 25$ and a not highly clustered grid, steady-state solutions can be reached in 600-800 steps with 12 orders of magnitude drop in the L_2-norm residual. However, the convergence rate is many times slower for the equilibrium real gas counterpart. Figure 6.5 shows the convergence rate for a perfect gas compared with an equilibrium real gas computation with a fixed CFL of 50. Note that the scale of the ordinates used in figure 6.5 for the perfect gas and an equilibrium real gas are not the same. The freestream conditions for an equilibrium real gas study are the same as figure 6.3. An important observation for the behavior of the convergence rate for the Mach 15 equilibrium real gas case is that the discontinuities of the thermodynamic derivatives which exist in the curve fits of Srinivasan et al. [164] might be the major contributing factor. This is evident from figures 6.2d and 6.3d and from a comparison with the convergence rate for the perfect gas. Another contributing factor is that the curve fits are accurate only for temperatures up to 6000°K. Since the temperature in this case is slightly above 6000°K, there is an uncertainty in the accuracy of the computed results. Further improvement of the existing curve fitting procedure is needed in order to assess the 'true' difference in convergence rate between perfect and equilibrium real gases.

Inviscid Impinging Shock Computations: Figures 6.6 and 6.7 show the schematic of the computational domain, the Mach contours and L_2-norm of the residual computed by the implicit upwind scheme (6.5) (with $\theta = 1$, $\omega = 0$) of an inviscid shock-on-shock interaction on a blunt body with radius R_l and thickness $D = 2R_l$ in the low hypersonic range $4 \leq M_\infty < 8$. Higher inviscid hypersonic Mach number $8 \leq M_\infty \leq 30$ computations using these scheme are also possible but are not shown here. Some viscous and inviscid studies on flow fields of this type were reported in references [182,187-189]. This flow field is typical of what may be experienced by the inlet cowl of a hypersonic aerodynamic vehicle. The freestream conditions for this flow field are the freestream Mach number $M_\infty = 4.6$, the freestream temperature $T_\infty = 167$K, and $\gamma = 1.4$ for a perfect gas. An oblique shock with an angle of 20.9° relative to the free stream impinges on the bow shock. Various types of interactions occur depending on where the impingement point is located on the bow shock. As shown by the Mach contours ranging from 0 to 4.55 in increments of 0.05, the impinging shock has caused the stagnation point to move away from its undisturbed location at the symmetry line. The surface pressures at the new stagnation point can be several times larger than those at the undisturbed location of the stagnation point. In addition, a slip surface emanates from the bow shock and impinging shock intersection point and is intercepted by a shock wave which starts at the upper kink of the bow shock. The interacting shock waves and slip surfaces are confined to a very small region and must be captured accurately by the numerical scheme if the proper surface pressures are to be predicted correctly. The 77×77 grid used and the convergence rate computed by the implicit scheme (6.5) are shown in figure 6.7. Though the flow pattern is significantly more complicated than for the previous cases, the convergence rate remains quite satisfactory. The convergence rate is for a variable Δt with a fixed CFL of 100. The scalloping of the L_2 residual norm is the characteristic of that specific CFL number. The shape of the scalloping seems to be a function of the CFL number. As shown in figure 6.6 at the inflow, all of the inviscid and viscous interfering blunt-body computations start with the appropriate freestream and oblique shock wave conditions as boundary conditions.

Viscous Steady Computations With or Without Impinging Shock: To keep the study tractable only two types of physical flow fields were chosen. The first is a viscous hypersonic ideal gas blunt-body flow at $M_\infty = 8.03$ and $T_\infty = 122.1$K with a laminar Reynolds number of 387,750 based on the body diameter . The second problem (with the same flow conditions) is similar to the inviscid shock-on-shock interaction where an oblique shock impinges on the bow shock of a blunt body at an angle of 19° relative to the free stream. A more detailed flow field computation of the six types of shock patterns categorized by Edney [186] is presented in reference [180]. For the convergence study only one type of interaction, namely the so-called Type III interaction, is considered. Also the study is restricted to only one type of time-stepping sequencing and only one value of the entropy-correction parameter. The computational meshes (not shown) consist of 181 points in the circumferential direction and 91 points in the direction normal to the body and are highly clustered in the wall region to resolve the viscous layer.

At this point, it is important to point out that the time-step sequence used for the viscous steady flows is very different from the inviscid study. Most of the inviscid computations use the same initial time-step input or a fixed CFL together with a local time-stepping procedure throughout the entire iteration process. The time-step sequence chosen for the viscous steady calculations is based on experience with a wide range of hypersonic flow simulations and consists of a sequence of fixed time steps with the procedure of doubling the time step every 100-400 time steps until the specified time step is reached. The initial time step is $\Delta t = 0.001$ which corresponds to a maximum Courant (CFL) number of 10 to 20 for the current problem and grid size. Larger values of the initial time step usually prevent convergence. The four specified time steps considered range from 0.001 to 0.008 with corresponding CFL numbers ranging from 20 to 200. Much larger maximum CFL (or specified time step) numbers are possible but do not improve the convergence rates. The value of the entropy-correction parameter was fixed at $\tilde{\delta} = 0.15$, again based on experience with a wide range of hypersonic flow field simulations.

The results of the blunt-body steady viscous flow obtained with the temporally second-order accurate algorithm (6.5) together with (5.15g,h) (hereafter referred to as the full matrix form) are shown in figure 6.8. Here algorithm (6.5) for the viscous computations means the appropriate three-point central differences of the viscous terms are added to the explicit and implicit operators of (6.5). Depicted are Mach contours ranging from 0 to 8 in increments of 0.1 and entropy contours normalized by the freestream value and ranging from 0 to 6.4 in increments of 0.1. The final view in figure 6.8 is the history of the L_2 norm residual. The residual drops to machine accuracy (10^{-14}) in less than 3200 steps. The corresponding results using the same algorithm (6.5) together with (5.15i,j) (hereafter referred as the diagonal form) are illustrated in figure 6.9. No noticeable difference in the numerical results is observed in the Mach number or entropy contours. However, the residual curves are very different. The residual for the diagonal scheme reached a plateau of 5×10^{-6} at 1500 steps and stayed at that level.

A more complex flow field which is far more difficult to compute is depicted in figures 6.10 and 6.11. The results using the same second-order time accurate full matrix algorithm are shown in figure 6.10. The convergence rate is slower than for the blunt body non-interfering case but is still satisfactory. The residual dropped seven orders of magnitude in 3000 steps. In both of the blunt-body flow cases with or without impinging shocks, steady state can be reached within 1000-1500 iterations. The extra iterations are needed only to bring the residual to a lower level but no change in the coutour plots or surface pressures at least to within 3-4 digits of accuracy is observed. However, the results shown in figure 6.11 using the diagonal scheme are not satisfactory. The residuals dropped less than two orders of magnitude in 3000 steps. The noise in the Mach number and entropy contours in the upper portion of the bow shock using the diagonal form of the scheme indicates that the algorithm has a problem reaching the converged steady-state solution.

All of the computations for figures 6.8 - 6.11 have the viscous terms included in the implicit operator. If the viscous terms are not included in the implicit operator, then the full matrix scheme becomes unstable for $\Delta t \geq 0.004$, whereas the diagonal scheme exhibits no change in convergence rate.

In summary, from the point of view based on the L_2-norm of the residuals, the best convergence rates were achieved by the full matrix form with the viscous terms included since it allowed the residual to drop to machine accuracy (10^{-14}). The diagonal form (5.15i,j) did not fare too well. Although there is a substantial savings in operations count per iteration ($\sim$ 20% less), the L_2-norm of the residual never dropped below 10^{-6} for all the time steps considered, and for the impinging shock case, the algorithm has a slight problem reaching the converged steady state. Moreover, the inclusion of the implicit viscous terms had little effect on the diagonal form of the scheme but is important for the full matrix form of the scheme. One way of taking advantage of the low operations count ($\sim$ 30% less) of the diagonal form (without the implicit viscous terms) is to use the scheme as an efficient way of obtaining a rough solution (from freestream) for the initial input to the full matrix algorithm. The temporally second-order time-differencing scheme had a marginal but beneficial effect on the convergence rates when compared with the temporally first-order scheme.

Viscous Steady and Unsteady Mach 15 Computations with Impinging Shock: Figure 6.12 illustrates the shock resolution of unsteady and steady thin-layer Navier-Stokes computations by the second-order time-accurate, full matrix algorithm (6.5). This steady test case is similar to the previous impinging shock study except the freestream Mach number is 15, the impingement shock angle is 22.75°, the freestream temperature is $T_\infty = 255.6$K, and the Reynolds number based on the diameter is 186,000. Shown are the Mach contours from 0 to 15 in increments of 0.1. For the unsteady computation, the impingement shock at an angle of 22.75° relative to the freestream moves downward across the bow-shock of the blunt body. The impingement shock velocity is 10% of the freestream velocity ($M_\infty = 15$). Although the impingement shock locations for the unsteady and steady computations are similar, the shock patterns are very different. A 241 × 141 non-adaptive grid is used for both computations. A time step of 0.002 (equivalent to a maximum CFL of 48) is used for steady-state computations whereas a time step of 0.0005 (equivalent to a maximum CFL of 10-12 at the vicinity of the boundary layer and a CFL of 1 at the rest of the flow field) is used for the unsteady computations. The steady-state solution can be reached in 1200 steps with a three order of magnitude drop in the L_2-norm residual. However, extra iterations are needed to bring the residual to a lower level but no change in the coutour plots or surface pressures at least to within 3-4 digits of accuracy is observed. More detailed study of the surface pressure and heat transfer rate of these types of shock-on-shock steady and unsteady numerical simulations were reported in references [180,181]. Figures 6.13 and 6.14 show the comparison between the steady and unsteady Mach 15 flows for six types of shock patterns where the flow conditions are the same as figure 6.12. Although no experimental data are available for the Mach 15 case, these figures serve as an example on the capability of these implicit schemes for steady and unsteady hypersonic viscous flows containing complex shock structures.

6.7. Concluding Remarks on the Numerical Study

A two-parameter family of implicit time-accurate shock-capturing algorithms has been shown to be applicable for hypersonic viscous flow computations. These algorithms are formulated in finite volume and pseudo-finite volume form and have been shown to be quite efficient and accurate for steady-state as well as unsteady viscous and inviscid hypersonic complex shock interactions. Some numerical aspects of these TVD-type algorithms that affect the convergence rate for hypersonic Mach numbers or equilibrium real gas flows but have negligible effect on low Mach numbers or perfect gas flows are identified. Improvements have been made to the algorithms to speed up the convergence rate in the hypersonic flow regime. Even with the improvements, though, the convergence is in general slower for equilibrium real gases than for a perfect gas. The nonsmoothness in the curve fits of Srinivasan et al. may be a major contributing factor in slowing down the convergence rate for equilibrium real gas flows. Also, the convergence rate is, in general, slower for viscous flows than for inviscid steady flows.

VII. EFFICIENT SOLUTION PROCEDURES FOR LARGE SYSTEMS WITH STIFF SOURCE TERMS [76,86,95]

In the application of modern shock-capturing methods like TVD-type of schemes to chemically reacting flows, Carofano [99] was the first to introduce the formalism that enabled full coupling in Harten's explicit TVD scheme for a two-species, two-dimensional unsteady flow in Cartesian coordinates. However, due to the system size and the varying multiple time-scale nature of the problem, the operations count increases nonlinearly as the number of species and/or the number of reactions increases. To avoid solving a large system, Gnoffo and McCandless [184] and Gnoffo et al. [79] uncoupled the species equations from the fluid dynamics equations and solved these two sets of systems of nonlinear partial differential equations in a time-lag fashion (loosely coupled method) by using a point-relaxation technique with a second-order symmetric TVD scheme of Yee [21,163] and an upwind TVD scheme of Osher-Chakravarthy [22]. Eberhardt and Brown [190] attempted to use the eigenvalues and eigenvectors of the fluid dynamics equations alone to obtain a "fully coupled" first-order explicit TVD scheme for a one-dimensional flow. The results of Eberhardt and Brown showed excessive smearing at the shock when compared with the true, fully coupled explict TVD result. Their motivation for designing such a coupling procedure was to optimize the operations count by avoiding multiplication of large matrices. However, as was demonstrated in [95] and summarized in this section, if one makes use of the unique structure of eigenvectors and eigenvalues for fluid flow of this type, the fully coupled formulation can be simplified even for a large number of species, thus providing a more efficient solution procedure than one might have anticipated. Moreover, using the eigenvalues and eigenvectors for the fully coupled equation set allows one to have the freedom of controlling the proper amount of numerical dissipation for the individual waves [21]. In particular, for the type of two-dimensional chemically reacting flows that are considered here, the number of linear waves is $ns + 1$ in each spatial direction where ns is the number of linearly independent species. Note that in order to capture contact discontinuities accurately, it is very important to apply the proper amount of numerical dissipation to the linear waves.

As discussed in sections 3.3, 3.6, and 3.7, the TVD property is only valid for homogeneous scalar hyperbolic conservation laws. Certain types of source terms might preserve the original TVD property of the homogeneous PDE and others might not. However, disregarding the type of bounded source terms, one is not precluded from the application of TVD schemes when source terms are present. But extreme precaution has to be taken in the procedure of including the source terms. This applies particularly to stiff source terms of the types in thermally and chemically nonequilibrium flows. Numerical difficulties associated with time-accurate computations for problems with stiff source terms were discussed in section III. However, in the use of time-accurate methods as relaxation procedures for steady-state computations, the semi-implicit algorithm discussed in section 3.7 appears to work quite well.

Three types of schemes are described. If the stiffness is entirely dominated by the source term, two semi-implicit TVD-type shock-capturing methods (as discussed in section III) are proposed for steady-state calculations provided that the Jacobian of the source terms possesses certain properties. One of the semi-implicit schemes can be viewed as a variant of the Bussing and Murman point-implicit predictor-corrector scheme [96] with a more appropriate numerical dissipation which provides for the computation of strong shock waves in the hypersonic regime and a speedup in the convergence rate for steady-state applications. The predictor-corrector scheme of Bussing and Murman in turn is the explicit MacCormack scheme with the source term treated implicitly. Another semi-implicit scheme is a splitting method as discussed in section 3.9 in which the homogeneous part of the PDEs and the source term are handled in separate steps. For multidimensional problems, the homogeneous part of the PDEs can be discretized in a variety of ways (predictor-corrector, time-splitting, or linear multistep methods). In the case of the time-splitting approach, the MUSCL formulation in conjunction with the Lax-Friedrichs numerical flux will result in an efficient high-resolution algorithm without the use of Riemann solvers.

However, if the stiffness is not solely dominated by the source terms (e.g., stiffness due to highly irregular grid and/or viscous flows), a fully implicit method would be a more efficient procedure. The situation is complicated by problems in more than one space dimension, and the presence of stiff source terms further complicates the solution procedures for alternating direction implicit (ADI) methods. In fact, there seems to be no straightforward way of efficiently treating general stiff source terms implicitly with ADI procedures. Several alternatives will be discussed. The fully implicit relaxation algorithm can be viewed as a variant of a fully coupled form of the algorithm proposed by Gnoffo and McCandless [184].

In the simplest case — the use of the MUSCL formulation and the Lax-Friedrichs numerical flux — one has the advantage of allowing the use of a larger time step and at the same time obtaining high resolution without the use of any Riemann solver [76]. The operations count per time step would be less than the loosely coupled method using a Riemann solver. The operations count per time step would be comparable with the Beam-Warming [28] type ADI method using a mixture of second- and fourth-order numerical dissipation terms [177].

An implicit algorithm with explicit coupling between fluid and species equations proposed by the author [1] will also be stressed here. Many existing perfect-gas or equilibrium real-gas computer codes can easily be modified to include this algorithm, which is a compromise between the loosely coupled implicit method of [184] and the fully coupled, fully implicit TVD method described here. To make this section more self-contained, some of the variables that were defined earlier will be repeated.

7.1. An Explicit Predictor-Corrector Algorithm for Systems with Source Terms

In this section, all the terms that are required for the basic TVD scheme for the compressible inviscid nonequilibrium reacting flow equations are derived. For simplicity, only chemically reacting flows are discussed. More general equation sets such as thermally nonequilibrium flows have similar properties and can be extended in a straightforward manner.

The Governing Equations: Consider a two-dimensional system of nonhomogeneous hyperbolic conservation laws,

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = S(U). \tag{7.1}$$

Here U, $F(U)$, $G(U)$, and $S(U)$ are column vectors of k components. Let $A = \partial F/\partial U$ and $B = \partial G/\partial U$, with $(a_x^1, a_x^2, ..., a_x^k)$ and $(a_y^1, a_y^2, ..., a_y^k)$ being the eigenvalues of A and B. Denote R_x and R_y as the matrices whose columns are eigenvectors of A and B, and denote R_x^{-1} and

R_y^{-1} as the inverses of R_x and R_y. In the case of the compressible inviscid flow equations with chemical reactions, the global continuity equation is replaced by the individual species continuity equations,

$$U = \begin{bmatrix} \rho^1 \\ \rho^2 \\ \cdot \\ \cdot \\ \cdot \\ \rho^{ns} \\ m \\ n \\ \epsilon \end{bmatrix}; \quad F(U) = \begin{bmatrix} c^1 m \\ c^2 m \\ \cdot \\ \cdot \\ \cdot \\ c^{ns} m \\ \frac{m^2}{\rho} + p \\ \frac{mn}{\rho} \\ (\epsilon + p)\frac{m}{\rho} \end{bmatrix}; \quad G(U) = \begin{bmatrix} c^1 n \\ c^2 n \\ \cdot \\ \cdot \\ \cdot \\ c^{ns} n \\ \frac{mn}{\rho} \\ \frac{n^2}{\rho} + p \\ (\epsilon + p)\frac{n}{\rho} \end{bmatrix}; \quad S(U) = \begin{bmatrix} s^1 \\ s^2 \\ \cdot \\ \cdot \\ \cdot \\ s^{ns} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{7.2}$$

Here $m = \rho u, n = \rho v$, and s^i represents the production of species from chemical reactions. The variables are the velocity components u and v, the pressure p, the total energy per unit volume ϵ, and the density of the ith species ρ^i. Also, $\rho = \sum_{i=1}^{ns} \rho^i$ and $c^i \rho = \rho^i$, where ns is the number of species in the model and c^i is the species mass fraction. Equation (7.2) assumes the ρ^i are linearly independent.

The eigenvalues of A and B are

$$(a_x^1, ..., a_x^{ns+3}) = (u, ..., u, u + a, u, u - a), \tag{7.3a}$$

$$(a_y^1, ..., a_y^{ns+3}) = (v, ..., v, v + a, v, v - a). \tag{7.3b}$$

Here the so-called "frozen sound speed" a is

$$a^2 = p_\rho + p_\epsilon (H - u^2 - v^2), \tag{7.4}$$

with

$$p_\rho = \sum_{i=1}^{ns} c^i p_{\rho^i}, \tag{7.5}$$

$$\sum_{i=1}^{ns} c^i = 1, \tag{7.6}$$

$$p_{\rho^i} = \frac{\partial p}{\partial \rho^i}\Big|_{m,n,\epsilon,\dots}, \tag{7.7}$$

$$p_\epsilon = \frac{\partial p}{\partial \epsilon}\Big|_{m,n,\rho^1,\dots}, \tag{7.8}$$

$$H = \frac{\epsilon + p}{\rho}. \tag{7.9}$$

The frozen sound speed a for the PDEs (7.1) (a without subscript or superscript and should not be confused with the a_x^i or a_y^i) defined in (7.4) has no physical meaning. It is defined here for the convenience of notation for the basic scheme.

The superscript n is used for the time-index and should not be confused with $n = \rho u$ in equation (7.2). Let $a^l_{j+\frac{1}{2}}$, $R_{j+\frac{1}{2}}$, $R^{-1}_{j+\frac{1}{2}}$ denote the quantities a^l_x, R_x, R^{-1}_x evaluated at some symmetric average of $U_{j,k}$ and $U_{j+1,k}$. Similarly, let $a^l_{k+\frac{1}{2}}$, $R_{k+\frac{1}{2}}$, $R^{-1}_{k+\frac{1}{2}}$ denote the quantities a^l_y, R_y, R^{-1}_y evaluated at some symmetric average of $U_{j,k}$ and $U_{j,k+1}$. In the case of chemically reacting flows, $a^l_{j+\frac{1}{2}}$, $R_{j+\frac{1}{2}}$, $R^{-1}_{j+\frac{1}{2}}$, $a^l_{k+\frac{1}{2}}$, $R_{k+\frac{1}{2}}$, and $R^{-1}_{k+\frac{1}{2}}$ are defined in ways similar to the ones described in [192] or used by Huang [51], Carofano [99], Shinn & Yee [118], or Gnoffo [79]. All of the calculations in this section use a variant of the symmetric average of Huang and Carofano. Whichever symmetric average one decides to use, it is best to choose the averaged state such that the quantities in (7.3) remains real.

In case one keeps the global continuity equation plus $ns - 1$ species continuity equations, the form and structure of the eigenvalues and eigenvectors set are very similar. Form (7.2) has the advantage of avoiding machine round off errors in the event the magnitude of one of the species densities is very small compared with the rest.

For thermally and chemically nonequilibrium flows, the eigenvalues and eigenvectors have a similar structure. For the two-dimensional system (7.1), if nt is the number of thermal energy variables, then the eigenvalues in the x-direction will have $(ns + nt + 1)$ "u" characteristics plus $u + a$ and $u - a$ characteristics. Here the values "a" will reflect the added thermal energy variables .

Riemann Solver for Nonequilibrium Flows: Extensions of the exact Riemann solver of Godunov to certain types of nonequilibrium flows have been obtained by Glaz et al. [191]. Extensions of the Osher-Solomon Riemann solver are not known at this point.

Basically, if one assume each species of the gas mixture behaves as a thermally perfect gas, the flux function again possesses the homogeneous property. The Steger-Warming flux-vector splitting carries over to nonequilibrium flows. However, the van Leer flux-vector splitting and Roe's approximate Riemann solver are no longer valid for nonequilibrium flows. Various approximations to van Leer and Roe's Riemann solvers have been proposed [192-195]. Basically their generalization for nonequilibrium flows are almost identical to the equilibrium real gas cases. They all still involve arbitrary constants. In the case of the generalized Roe average, there exist $ns - 1$ arbitrary constants instead of only one for the equilibrium real gases. See [192] for details. The basic difference in the formulations between references [192], [193] and [195] remains the same as for the equilibrium real gases discussion in section IV. Reference [194] basically discusses even more restricted flows than references [192] and [193].

Explicit Predictor-Corrector TVD Scheme: With the above notation, a formal extension of the scalar explicit second-order TVD method (section 3.6) in predictor-corrector form via the local-characteristic approach for the nonlinear hyperbolic system (7.1) with nonzero source terms can be written as

$$\Delta U^{(1)}_{j,k} = -\frac{\Delta t}{\Delta x}\left(F^n_{j,k} - F^n_{j-1,k}\right) - \frac{\Delta t}{\Delta y}\left(G^n_{j,k+1} - G^n_{j,k}\right) + \Delta t S^n_{j,k}, \tag{7.10a}$$

$$U^{(1)}_{j,k} = \Delta U^{(1)}_{j,k} + U^n_{j,k}, \tag{7.10b}$$

$$\Delta U^{(2)}_{j,k} = \frac{1}{2}\left\{-\Delta U^{(1)}_{j,k} - \frac{\Delta t}{\Delta x}\left[F^{(1)}_{j+1,k} - F^{(1)}_{j,k}\right] - \frac{\Delta t}{\Delta y}\left[G^{(1)}_{j,k} - G^{(1)}_{j,k-1}\right] + \Delta t S^{(1)}_{j,k}\right\}. \tag{7.10c}$$

$$U^{(2)}_{j,k} = \Delta U^{(2)}_{j,k} + U^{(1)}_{j,k}, \tag{7.10d}$$

$$U^{n+1}_{j,k} = U^{(2)}_{j,k} + \frac{1}{2}\left[\overline{R}_{j+\frac{1}{2}}\overline{\Phi}_{j+\frac{1}{2}} - \overline{R}_{j-\frac{1}{2}}\overline{\Phi}_{j-\frac{1}{2}}\right] + \frac{1}{2}\left[\overline{R}_{k+\frac{1}{2}}\overline{\Phi}_{k+\frac{1}{2}} - \overline{R}_{k-\frac{1}{2}}\overline{\Phi}_{k-\frac{1}{2}}\right]. \tag{7.10e}$$

The elements of the vector $\Phi_{j+\frac{1}{2}}$ are the same as equations (3.69d) and (3.69e), with $a_{j+\frac{1}{2}}$ replaced by $a^l_{j+\frac{1}{2}}$, $\Delta_{j+\frac{1}{2}}$ replaced by $\alpha^l_{j+\frac{1}{2}}$, and $\widehat{Q}_{j+\frac{1}{2}}$ replaced by $\widehat{Q}^l_{j+\frac{1}{2}}$. $\overline{R}_{j+\frac{1}{2}}$ and $\overline{\Phi}_{j+\frac{1}{2}}$ can be evaluated at U^n or $U^{(2)}$ as discussed in sections 3.6 - 3.9. The situation is similar for the the TVB method or the high-resolution Lax-Friedrichs method as discussed in section 3.5.4. Here Δy is the grid spacing with $y_k = k\Delta y$. Numerical experiments in reference [86] indicated that limiting based on U^n (for the second and third terms of (7.10e)) is preferable for nonstiff source terms, but that limiting based on $U^{(2)}$ may be more robust for stiff source terms. For steady-state calculations and for a contractive type of source term $S(U)$, it appears that (7.10e) might help improve the convergence rate by using the most updated information in the relaxation procedure. Other procedures are currently under investigation.

7.2. More Efficient Solution Procedures for Large Systems

The extra computation in (7.10) compared with a classical central-difference shock-capturing scheme such as the Lax-Wendroff method is due to the vectors $(R\Phi)_{j\pm\frac{1}{2}}$. At first glance, the vectors $\alpha_{j\pm\frac{1}{2}}$ and $(R\Phi)_{j\pm\frac{1}{2}}$ involve matrix and vector multiplication of dimension $ns+3$ for every grid point, and thus tend to discourage their adoption in problems other than ideal gas flows. Researchers such as Gnoffo and McCandless [184], and Eberhardt and Brown [190] were motivated to pursue other avenues to solve the complicated chemically reacting flow problems. However, as was illustrated in [95], if one makes use of the unique structure of the eigenvectors and eigenvalues for fluid flow of this type, the fully coupled formulation can be simplified even for a large number of species, and thus becomes a viable approach.

With straightforward manipulations, the computation for scheme (7.10) can be simplified tremendously. The corresponding vector α in equation (4.12c) for system (7.1)-(7.2), for example, can be expressed as

$$\alpha = a^2 \begin{bmatrix} \Delta\rho^1 - c^1 aa \\ \Delta\rho^2 - c^2 aa \\ \cdot \\ \cdot \\ \cdot \\ \Delta\rho^{ns} - c^{ns} aa \\ \frac{1}{2}\left(aa - \frac{u}{a}bb + \frac{\Delta m}{a}\right) \\ \frac{1}{a}(-vbb + \Delta n) \\ \frac{1}{2}\left(aa + \frac{u}{a}bb - \frac{\Delta m}{a}\right) \end{bmatrix}, \tag{7.11}$$

with

$$aa = \frac{1}{a^2}\left[\sum\nolimits_{i=1}^{ns} p_{\rho^i}\Delta\rho^i - p_e(u\Delta m + v\Delta n - \Delta e)\right], \tag{7.12}$$

$$bb = \sum\nolimits_{i=1}^{ns}\Delta\rho^i. \tag{7.13}$$

Here, for example, $(\Delta\rho^i)_{j+\frac{1}{2},k} = \rho^i_{j+1,k} - \rho^i_{j,k}$, and it is understood that the spatial indices in (7.11)-(7.13) are at $(j+\frac{1}{2},k)$. Similarily, $R\Phi$ also has a very simple form. For instance, the $R\Phi$ associated with the x-direction flux can be expressed as

$$R\Phi = \frac{1}{a^2}\begin{bmatrix} \phi^1 + c^1 k_1 \\ \phi^2 + c^2 k_1 \\ \cdot \\ \cdot \\ \cdot \\ \phi^{ns} + c^{ns} k_1 \\ uk_2 + k_3 \\ vk_2 + a\phi^{ns+2} \\ Hk_2 - \frac{1}{p_e}\sum_{i=1}^{ns}(a^i)^2\phi^i + av\phi^{ns+2} + uk_3 \end{bmatrix}, \tag{7.14a}$$

with

$$k_1 = \phi^{ns+1} + \phi^{ns+3}, \tag{7.14b}$$

$$k_2 = \sum\nolimits_{i=1, i\neq ns+2}^{ns+3}\phi^i, \tag{7.14c}$$

$$k_3 = a(\phi^{ns+1} - \phi^{ns+3}). \tag{7.14d}$$

where $(a^i)^2$ is the square of the frozen sound speed of the ith species. Here the spatial indices on (7.14) are at $(j+\frac{1}{2},k)$. As one can see, the terms in equations (7.11) and (7.14) due to the species equations are simple and do not require many operations. To be precise, for each additional species, only two operations contribute to the α or the $R\Phi$ operations. Therefore, the increase in the number of species equations is not as CPU-intensive as one might have anticipated. The inclusion of the thermally nonequilibrium flow analysis was later extended by Liu and Vinokur [192]. The property of the corresponding α and $R\Phi$ remains unchanged except more equations than (7.2) are involved.

7.3. A Semi-implicit Predictor-Corrector Algorithm and a 3-D Example

As mentioned in section 3.6 for the scalar schemes, the explicit TVD scheme (7.10) can be used for either time-accurate or steady-state calculation. It can be second-order accurate in time and space. However, if the source term is stiff, the restriction on the time step due to stability requirements is prohibitively small and (7.10) is not practical, especially for steady-state applications where a large system such as (7.2) is involved. In this section, a semi-implicit method for steady-state applications is described. Another alternative is a fully implicit method. The basic implicit scheme and the related difficulty in efficiently extending the implicit method to two dimensions with stiff source terms will be discussed in the next section.

If one follows the idea of Bussing and Murman [96] in treating the source term implicitly, a semi-implicit predictor-corrector TVD algorithm for steady-state computations can easily be obtained. It can be written as a one-parameter family of time-differencing schemes for the source term; i.e., the following formulation includes scheme (7.10). The scheme for system (7.1) can be written as

$$D^n_{j,k}\Delta U^{(1)}_{j,k} = -\frac{\Delta t}{\Delta x}\left(F^n_{j,k} - F^n_{j-1,k}\right) - \frac{\Delta t}{\Delta y}\left(G^n_{j,k+1} - G^n_{j,k}\right) + \Delta t S^n_{j,k}, \tag{7.15a}$$

$$D = \left(I - \Delta t\theta\frac{\partial S}{\partial U}\right), \tag{7.15b}$$

$$\widehat{D}_{j,k}\Delta U^{(2)}_{j,k} = \frac{1}{2}\left\{-\Delta U^{(1)}_{j,k} - \frac{\Delta t}{\Delta x}\left[F^{(1)}_{j+1,k} - F^{(1)}_{j,k}\right] - \frac{\Delta t}{\Delta y}\left[G^{(1)}_{j,k} - G^{(1)}_{j,k-1}\right] + \Delta t\widehat{\widehat{S}}_{j,k}\right\}, \tag{7.15c}$$

$$U^{n+1}_{j,k} = U^{(2)}_{j,k} + \frac{1}{2}\left[\overline{R}_{j+\frac{1}{2}}\overline{\Phi}_{j+\frac{1}{2}} - \overline{R}_{j-\frac{1}{2}}\overline{\Phi}_{j-\frac{1}{2}}\right] + \frac{1}{2}\left[\overline{R}_{k+\frac{1}{2}}\overline{\Phi}_{k+\frac{1}{2}} - \overline{R}_{k-\frac{1}{2}}\overline{\Phi}_{k-\frac{1}{2}}\right]. \tag{7.15d}$$

Here, D is assumed to be invertible; i.e., only source terms with Jacobians such that D is invertible at each grid point are permissible. Again, θ has the same meaning as in the scalar case. For $\theta \neq 0$, the source terms are treated implicitly. If $\theta = 1$, the time differencing for the source term is first-order and (7.15) is best suited for steady-state calculations. See sections 3.3, 3.6 - 3.9 and 7.1 for the discussion of equation (7.15d) and the usage of the scheme (7.15) in general. When $\theta = 1/2$ and $\widehat{S}_{j,k}$ in $\widehat{D}$ and $\widehat{\widehat{S}}_{j,k}$ are evaluated at U^n_j, the method is second-order in time and space while $\widehat{S}_{j,k}$ and $\widehat{\widehat{S}}_{j,k}$ evaluated at $U^{(1)}_{j,k}$, the scheme would only be second-order for steady-state calculations (i.e., first-order in time and second-order in space for unsteady problems). See reference [86] for a proof.

One can simplify equation (7.15) by partitioning the vectors U, F, G, S, and D in equation (7.2) as follows:

$$U = \begin{bmatrix} U^I \\ U^{II} \end{bmatrix}; \quad U^I = \begin{bmatrix} u^1 \\ \cdot \\ \cdot \\ \cdot \\ u^{ns} \end{bmatrix}, \quad U^{II} = \begin{bmatrix} u^{ns+1} \\ u^{ns+2} \\ u^{ns+3} \end{bmatrix}, \tag{7.16a}$$

$$F = \begin{bmatrix} F^I \\ F^{II} \end{bmatrix}, \quad G = \begin{bmatrix} G^I \\ G^{II} \end{bmatrix}, \tag{7.16b}$$

$$S = \begin{bmatrix} S^I \\ S^{II} \end{bmatrix}; \quad S^I = \begin{bmatrix} s^1 \\ \cdot \\ \cdot \\ \cdot \\ s^{ns} \end{bmatrix}, \quad S^{II} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{7.16c}$$

$$D = \begin{bmatrix} D^{11} & D^{12} \\ D^{21} & D^{22} \end{bmatrix}. \tag{7.16d}$$

Here D^{21} is a null matrix and D^{22} is an identity matrix. With the above definitions, the scheme can be greatly simplified. The procedures are as follows: taking the predictor step, for example, one first solves for $(\Delta U^{II})^{(1)}$ by

$$(\Delta U^{II})^{(1)}_{j,k} = -\frac{\Delta t}{\Delta x}\left[(F^{II}_{j,k})^n - (F^{II}_{j-1,k})^n\right] - \frac{\Delta t}{\Delta y}\left[(G^{II}_{j,k+1})^n - (G^{II}_{j,k})^n\right], \tag{7.17a}$$

then for $(\Delta U^{I})^{(1)}$ by

$$(D^{11}_{j,k})^n(\Delta U^{I})^{(1)}_{j,k} = (\text{ r.h.s })^I - D^{12}_{j,k}(\Delta U^{II})^{(1)}_{j,k}, \tag{7.17b}$$

where (r.h.s$)^I$ is the right-hand side of (7.17a) with all the indices "II" replaced by "I", and with the term $\Delta t(S^I_{j,k})^n$ added. In other words, one only has to invert the D^{11} matrix of dimension (ns, ns) instead of $(ns + 3, ns + 3)$. Similarly, one can simplify the corrector step in the same way. The solution obtained from the above procedure is used in (7.15d). Or, to explain it in another way, one solves the predictor step of the fluid equations

$$\frac{\partial U^{II}}{\partial t} + \frac{\partial F^{II}(U)}{\partial x} + \frac{\partial G^{II}(U)}{\partial y} = 0 \tag{7.18a}$$

explicitly, then uses the result to solve the predictor step of the species equations

$$\frac{\partial U^{I}}{\partial t} + \frac{\partial F^{I}(U)}{\partial x} + \frac{\partial G^{I}(U)}{\partial y} = S^I \tag{7.18b}$$

explicitly, with the exception that the chemical reaction terms are treated implicitly in (7.18b). One then repeats the same procedure for the first corrector step. The solution obtained from the first corrector step is then used to solve for α and $R\Phi$ in equations (7.11)-(7.14) for the complete system (7.1) so that one can solve for the second corrector step. Here, the first corrector step means the step to obtain $U^{(2)}$ from $U^{(1)}$, and the second corrector step means the step to obtain U^{n+1} from $U^{(2)}$ in (7.15). Note that the second corrector step is the important part of the algorithm that deviates from the Bussing and Murman method. This step, which ensures that the method will have the TVD-type properties, is designed to capture shock waves without the associated spurious oscillations.

In the case where S^{II} is not a null vector and it is not stiff, equation (7.17) is still applicable, except one has to add the term $\Delta t(S^{II}_{j,k})^n$ to the right-hand side of (7.17a). In steady-state calculations where body-fitted coordinates are used, one can further speed up the convergence rate by using a local time-stepping approach [33,177].

To verify the current approach and to make a fair comparison with a known method, the semi-implicit scheme (7.15) together with (4.33) was implemented into an existing three-dimensional code [196]. This existing full Navier-Stokes code, originally developed by A. Kumar at NASA Langley Research Center, contains the explicit MacCormack scheme with source terms treated implicitly. A detailed description of a numerical experiment and the extension of (7.15) to three-dimensional, chemically reacting flows in generalized coordinates is given in Shinn et al. [118]. Only partial results from reference [118] are presented here to illustrate the performance of the semi-implicit scheme. In reference [118], three symmetric averages for the eigenvalues and eigenvectors were studied. Aside from a slight difference in convergence history, no visible differences in resolution were observed among the different averages. The computations shown employed $\delta_1 = 0.1$ for the entropy-correction function $\psi(z)$ in equation (3.18).

Figures (7.1)-(7.3) show a preliminary test result for a three-dimensional, five-species viscous reacting flow using the semi-implicit TVD scheme (7.15) together with (4.33), as compared with an existing classical shock-capturing method which supplies numerical dissipation linearly [196]. The numerical result for the semi-implicit TVD method is shown to be oscillation-free around the shock, while the time spent per iteration is approximately double when compared with the method used in [196]. The configuration of the numerical experiment with one of the computation station $z = 0.29$cm is shown in figure (7.1). Although this is a two-dimensional flow, to check out the three-dimensional code it is computed as a three-dimensional flow with the appropriate boundary conditions in the y-direction. A uniform grid consisting of 31 points in the x-direction, 6 points in the y-direction and a viscous grid consisting of 51 points in the z-direction was used. The inflow conditions are: pressure $p = 1$ atm, temperature $T = 1200$ K, and Mach number $M = 4$ for a premixed air and hydrogen fuel. The species considered are H_2, O_2, OH, H_2O, and N_2, with the two reactions (N_2 being inert)

$$H_2 + O_2 \rightleftharpoons 2OH,$$

$$2OH + H_2 \rightleftharpoons 2H_2O.$$

The reaction rates for the above are given in reference [196].

7.4. An Operator Splitting Algorithm

An alternative to the previous semi-implicit algorithm is a splitting method as discussed in section 3.9 in which the homogeneous part of the PDEs and the source term are handled in separate steps. For multidimensional problems, the homogeneous part of the equations can be discretized in a variety of ways (predictor-corrector, time-splitting, or linear multistep methods). Take for example the time-splitting method as discussed in section 5.1 using the MUSCL approach based on the Lax-Friedrichs numerical flux. In this case, for problems containing large number of species ns, a very efficient algorithm will result since a Riemann solver is not necessary. The resulting algorithm would only be slightly more expensive than the Lax-Wendroff type method. The main reason of using the MUSCL approach with time-splitting over the non-MUSCL approach (both employing the Lax-Friedrichs numerical flux) is that when generalized to nonlinear system cases, there is no need for the MUSCL approach to use any type of Riemann solver. For gas dynamics applications the limiter can be applied to the slope of the conservative variables or to the primitive variables themselves, thus making the extension to multidimensional problems straightforward, since one does not have to use the local one-dimensional characteristic variables assumptions or the complicated truly multidimensional Riemann solvers. A splitting (dimensional splitting and splitting of the source term) algorithm can be written as

$$U^{n+2}_{j,k} = \mathcal{L}_S^{h/2} \mathcal{L}_\xi^h \mathcal{L}_\eta^h \mathcal{L}_S^h \mathcal{L}_\eta^h \mathcal{L}_\xi^h \mathcal{L}_S^{h/2} U^n_{j,k}, \tag{7.19}$$

where $\mathcal{L}_\xi^h$, $\mathcal{L}_\eta^h$ and $\mathcal{L}_S^h$ have the same meaning as in section 5.1. A more expensive approach compared with the use of the Lax-Friedrichs numerical flux is to use a predictor-corrector form

(5.13) for the homogeneous part of the PDEs. This will result in the following

$$U^{n+2}_{j,k} = \mathcal{L}_S^{h/2}\mathcal{L}_f^h\mathcal{L}_S^h\mathcal{L}_f^h\mathcal{L}_S^{h/2}U^n_{j,k}, \tag{7.20}$$

where, for example, $\mathcal{L}_f^h\widehat{U}^{n+1}_{j,k}$ can be (5.13). However, in this case one cannot make use of the MUSCL approach in conjunction with the Lax-Friedrichs numerical flux. The stability, efficiency and accuracy of using the Lax-Friedrichs numerical flux in conjunction with the MUSCL or non-MUSCL approaches for steady-state nonequilibrium flows remains an open question until extensive numerical experiments are performed on some practical nonequilibrium flow applications.

7.5. A Fully Implicit TVD Method and 3-D Examples

Another type of shock-capturing scheme that might be applicable to (7.1) for steady-state applications is a one-parameter family or a two-parameter family (6.4) of explicit and implicit TVD-type of schemes (section III). For the nonequilibrium equation (7.1), the one-parameter family of schemes can be written as

$$\begin{aligned} U^{n+1}_{j,k} &+ \theta\left[\frac{\Delta t}{\Delta x}\left(\tilde{F}^{n+1}_{j+\frac{1}{2},k} - \tilde{F}^{n+1}_{j-\frac{1}{2},k}\right) + \frac{\Delta t}{\Delta y}\left(\tilde{G}^{n+1}_{j,k+\frac{1}{2}} - \tilde{G}^{n+1}_{j,k-\frac{1}{2}}\right) - \Delta t S^{n+1}_{j,k}\right] \\ &= U^n_{j,k} - (1-\theta)\left[\frac{\Delta t}{\Delta x}\left(\tilde{F}^{n}_{j+\frac{1}{2},k} - \tilde{F}^{n}_{j-\frac{1}{2},k}\right) + \frac{\Delta t}{\Delta y}\left(\tilde{G}^{n}_{j,k+\frac{1}{2}} - \tilde{G}^{n}_{j,k-\frac{1}{2}}\right) - \Delta t S^{n}_{j,k}\right]. \end{aligned} \tag{7.21}$$

Here θ has the same meaning as before. The numerical fluxes $\tilde{F}_{j\pm\frac{1}{2},k}$ and $\tilde{G}_{j,k\pm\frac{1}{2}}$ have the same meaning and form as for the scheme for the homogeneous PDEs (5.1). Again, the MUSCL or non-MUSCL approaches as discussed in sections 4.5 and 5.3 for the implicit method are applicable here. The savings on operations count by using the MUSCL approach and the Lax-Friedrichs numerical flux are even greater for the implicit method than for (7.19), since for time-accurate calculations, an $((ns+3)\times(ns+3))$ triple matrix multiplication per time step at each grid point and in each direction is involved for the implicit operator using the full matrix conservative linearized form such as (6.5) together with (5.15g,h).

7.5.1 A Conservative Linearized Form for Steady-State Applications

Following the same procedure as in section (3.10) and in Yee [21,74], a conservative linearized form of (7.21) can be written as

$$\begin{aligned} \left\{I + \theta\left[\frac{\Delta t}{\Delta x}\left(H^x_{j+\frac{1}{2},k} - H^x_{j-\frac{1}{2},k}\right)^n + \frac{\Delta t}{\Delta y}\left(H^y_{j,k+\frac{1}{2}} - H^y_{j,k-\frac{1}{2}}\right)^n - \Delta t(A^S)^n_{j,k}\right]\right\}E^n \\ = -\frac{\Delta t}{\Delta x}\left(\tilde{F}^n_{j+\frac{1}{2},k} - \tilde{F}^n_{j-\frac{1}{2},k}\right) - \frac{\Delta t}{\Delta y}\left(\tilde{G}^n_{j,k+\frac{1}{2}} - \tilde{G}^n_{j,k-\frac{1}{2}}\right) + \Delta t S^n_{j,k}, \end{aligned} \tag{7.22a}$$

$$E^n = U^{n+1} - U^n, \tag{7.22b}$$

where the $H^x_{j+\frac{1}{2},k}$, $H^y_{j,k+\frac{1}{2}}$, $F_{j+\frac{1}{2},k}$ and $G_{j,k+\frac{1}{2}}$ have the same form as in equation (5.15) and section (4.5) except all the vectors are in Cartesian coordinates. The value $A^S_{j,k}$ is the Jacobian of the source terms S evaluated at (j,k)

7.5.2. Stiff Source Terms, ADI Approaches and Relaxation Methods

The stiff terms $(A^S)^n_{j,k}$ in the implicit operator (7.22) complicate the solution procedures for the commonly used ADI approaches. Normally, if $(A^S)^n_{j,k}$ are not stiff, one can reformulate (7.22) by an ADI procedure like the Beam-Warming [28] algorithm for an efficient solution process. Unfortunately, the $(A^S)^n_{j,k}$ considered here are stiff; consequently the additional higher order terms due to the ADI formulation can no longer be ignored. In a different context, Van Dalsen and Steger [197] suggested a remedy if $(A^S)^n_{j,k}$ is a diagonal matrix with identical diagonal elements. However, for chemically reacting flows, the matrix $(A^S)^n_{j,k}$ is full for the upper $(ns, ns+3)$ entities and no straightforward efficient way of utilizing ADI approaches for nonlinear system cases with general stiff source terms can be found.

The straightforward way of iteratively solving (7.21) as a set of nonlinear system of equations, or solving the linearized form (7.22) or similar non-ADI forms, appears to be quite expensive for large systems. Recently, Gnoffo et al. [79] successfully demonstrated the usefulness of a point-relaxation method on the implicit symmetric TVD scheme similar to (7.22) for a loosely coupled chemically nonequilibrium flow. Here a similar point-relaxation or line-relaxation method is recommended for the fully coupled system (7.21) or (7.22). Despite the fact that a larger equation set is involved than in [79], the extra operations are minimized by making use of the simplification procedure of section 7.2. For a point-relaxation method, the size of the matrix inversion for (7.22) is $(ns+3, ns+3)$ as opposed to the loosely coupled method of [79], where the size is (ns, ns). The gain in the freedom of controlling the appropriate amount of numerical dissipation for each individual wave more than compensates for the extra expense. More importantly, solving the fully coupled system is believed to lead to a better convergence rate than solving the loosely coupled approach. This is evident from the study by Gnoffo [119]. All of the necessary terms required for the implicit scheme (7.22) are derived in section 7.1. The implicit operator of (7.22a) is diagonally dominant for $(A^S)^n_{j,k} = 0$. Therefore, one has to make sure that the source term does not destroy the diagonally dominant property which is required for some relaxation methods. In the following subsection, an even simpler form than (7.22) is described [1,95].

7.5.3. An Implicit Algorithm with Explicit Coupling between Fluid and Species Equations [1,95]

To avoid the inversion of large matrices, one can further simplify (7.22) by requiring that the coupling between the fluid and species equations be explicit. With this relaxed requirement, one effectively solves the fluid and species equations separately. Unlike the loosely coupled method used in [79] or the chemistry-split technique used in [198], the eigenvalues and eigenvectors of the fully-coupled equations are coupled explicitly between the fluid and species equations. This can be accomplished by partitioning U, F, G, S, A^S in the same way as in (7.16) and partitioning the Jacobians A and B, and the numerical fluxes $\widetilde{F}_{j\pm\frac{1}{2},k}$ and $\widetilde{G}_{j,k\pm\frac{1}{2}}$ similarly. For example, the Jacobian A and the numerical fluxes $\widetilde{F}_{j\pm\frac{1}{2},k}$ can be partitioned as

$$A = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix}, \tag{7.23}$$

$$\widetilde{F}_{j\pm\frac{1}{2},k} = \begin{bmatrix} \widetilde{F}^{I}_{j\pm\frac{1}{2},k} \\ \widetilde{F}^{II}_{j\pm\frac{1}{2},k} \end{bmatrix}. \tag{7.24}$$

The dimensions of the subvectors and submatrices are the same as (7.16) if equation set (7.2) is chosen, and the A^{11}, A^{12}, A^{21}, and A^{22} in this case are

$$A^{11} = \begin{bmatrix} u(1-c^1) & -uc^1 & . \; . \; . & -uc^1 \\ -uc^2 & u(1-c^2) & . \; . \; . & -uc^2 \\ . & . & . \; . \; . & . \\ . & . & . \; . \; . & . \\ . & . & . \; . \; . & . \\ -uc^{ns} & -uc^{ns} & . \; . \; . & u(1-c^{ns}) \end{bmatrix}, \tag{7.25a}$$

$$A^{12} = \begin{bmatrix} c^1 & 0 & 0 \\ c^2 & 0 & 0 \\ . & . & . \\ . & . & . \\ . & . & . \\ c^{ns} & 0 & 0 \end{bmatrix}, \tag{7.25b}$$

$$A^{21} = \begin{bmatrix} p_{\rho^1} - u^2 & p_{\rho^2} - u^2 & . \; . \; . & p_{\rho^{ns}} - u^2 \\ -uv & -uv & . \; . \; . & -uv \\ u(p_{\rho^1} - H) & u(p_{\rho^2} - H) & . \; . \; . & u(p_{\rho^{ns}} - H) \end{bmatrix}, \tag{7.25c}$$

$$A^{22} = \begin{bmatrix} 2u - up_e & -vp_e & p_\epsilon \\ v & u & 0 \\ H - u^2 p_\epsilon & -uvp_e & u(1+p_e) \end{bmatrix}. \tag{7.25d}$$

Similarly, $E_{j,k}$ and Ω^x can be partitioned as

$$E_{j,k} = \begin{bmatrix} E^{I}_{j,k} \\ E^{II}_{j,k} \end{bmatrix}, \tag{7.26}$$

$$\Omega^x = \begin{bmatrix} (\Omega^x)^{11} & (\Omega^x)^{12} \\ (\Omega^x)^{21} & (\Omega^x)^{22} \end{bmatrix}. \tag{7.27}$$

Recall that two choices of equation sets are available: namely, equation set (7.2), or the full fluid equations (i.e., keeping the global continuity equation) plus the $ns - 1$ species equations. The procedure suggested here is best suited for the latter equation set, since modification of many existing implicit solver computer codes (perfect gases or equilibrium real gases) would be minimal. The dimensions of subvectors and submatrices for the latter equation set would be slightly different. For example, the dimensions for A^{11}, A^{12}, A^{21}, and A^{22} would be $(ns-1) \times (ns-1)$, $(ns-1) \times 4$, $4 \times (ns-1)$, and 4×4, respectively.

The procedures for either equation set are as follows. One solves the fluid equations (e.g., (7.18a)) implicitly, and then uses the result $(U^{II})^{n+1}$ to solve the species equation (e.g., (7.18b)) implicitly. In other words, one solves $(U^{II})^{n+1}$ from the following:

$$\left\{I+\theta\left[\frac{\Delta t}{\Delta x}\left((H^x_{j+\frac{1}{2},k})^{II} - (H^x_{j-\frac{1}{2},k})^{II}\right) + \frac{\Delta t}{\Delta y}\left((H^y_{j,k+\frac{1}{2}})^{II} - (H^y_{j,k-\frac{1}{2}})^{II}\right)\right]^n\right\}(E^{II})^n = (rhs)^{II}, \tag{7.28a}$$

$$(rhs)^{II} = -\frac{\Delta t}{\Delta x}\left[(\widetilde{F}^n_{j+\frac{1}{2},k})^{II} - (\widetilde{F}^n_{j-\frac{1}{2},k})^{II}\right] - \frac{\Delta t}{\Delta y}\left[(\widetilde{G}^n_{j,k+\frac{1}{2}})^{II} - (\widetilde{G}^n_{j,k-\frac{1}{2}})^{II}\right], \tag{7.28b}$$

$$(H^x_{j+\frac{1}{2},k})^{II} = \frac{1}{2}\left[(A_{j+1,k})^{22} - (\Omega^x_{j+\frac{1}{2},k})^{22}\right]. \tag{7.28c}$$

The explicit coupling between the fluid and species equations is imbedded in (7.28b). After obtaining $(U^{II})^{n+1}_{j,k}$ from (7.28), the solution is used to solve the species equation as follows:

$$\left\{I+\theta\left[\frac{\Delta t}{\Delta x}\left((H^x_{j+\frac{1}{2},k})^{I} - (H^x_{j-\frac{1}{2},k})^{I}\right)^n + \frac{\Delta t}{\Delta y}\left((H^y_{j,k+\frac{1}{2}})^{I} - (H^y_{j,k-\frac{1}{2}})^{I}\right)^n - \Delta t\left[(A^S)^n_{j,k}\right]^{11}\right]\right\}(E^I)^n$$

$$= (rhs)^I - (lhs)^{12}, \tag{7.29a}$$

with

$$(lhs)^{12} = \left\{\frac{\theta}{2}\left[\frac{\Delta t}{\Delta x}\left((A_{j+1,k})^{12} - (\Omega^x_{j+\frac{1}{2},k})^{12} - (A_{j-1,k})^{12} + (\Omega^x_{j-\frac{1}{2},k})^{12}\right)\right.\right.$$
$$\left.\left. + \frac{\Delta t}{\Delta y}\left((B_{j,k+1})^{12} - (\Omega^y_{j,k+\frac{1}{2}})^{12} - (B_{j,k-1})^{12} + (\Omega^y_{j,k-\frac{1}{2}})^{12}\right)\right]^n - \theta\Delta t\left[(A^S)^n_{j,k}\right]^{12}\right\}E^{II} \tag{7.29b}$$

and

$$(H^x_{j+\frac{1}{2},k})^{I} = \frac{1}{2}\left[(A_{j+1,k})^{11} - (\Omega^x_{j+\frac{1}{2},k})^{11}\right]. \tag{7.29c}$$

The term $(rhs)^I$ is the right-hand side of (7.28b) with all the indices "II" replaced by "I" and the source term $\Delta t(S^n_{j,k})^I$ added. Here the quantities in equation (7.29) are evaluated with updated information from equation (7.28). To simplify the procedure even further, one can drop the second term on the right-hand side of (7.29a) entirely; i.e., one can set the second term in equation (7.29b) to zero. In other words one can solve for $(U^I)^{n+1}$ without using the most updated information from (7.28a); thus E^{II} in (7.29b) in this case is zero. This method is especially attractive for the MUSCL formulation in conjunction with the Lax-Friedrichs numerical flux [76].

With this simplified procedure, one only has to solve two reduced systems of dimension 3 and ns (or 4 and $ns - 1$). For line-relaxation methods, this procedure can provide a large savings in operations count. Although numerical experiments on this simplified procedure have not yet been done, one would expect that the current method will give a faster convergence rate than the method of Gnoffo et al. [79] which was demonstrated to be applicable to many three-dimensional blunt-body problems. If point relaxation were used, the only difference between the two methods is that reference [79] uses the eigenvalues and eigenvectors of the individual subsystems (fluid and species) alone, whereas the current method couples the eigenvalues and eigenvectors of the full system explicitly. The use of the full eigenvalue and eigenvector set for the current method is believed to enhance the coupling between the two systems without imposing additional conditions as employed in reference [79]. A viscous hypersonic numerical computation by Gnoffo et al. [79] using a loosely-coupled point relaxation method and the symmetric TVD scheme (4.42) will be presented in section 7.5.4. A different viscous hypersonic numerical computation by Gnoffo [119] using a fully coupled point relaxation method and the symmetric TVD scheme (4.42) compared with a loosely coupled method will be presented in section 7.5.5.

For steady-state application, an algorithm utilizing the TVD scheme for viscous flows is to difference the hyperbolic terms the same way as before, and then central difference the viscous terms. The final algorithm is the same as equation (7.22) (or (7.28)-(7.27)), except that the spatial central differencing of the viscous term is added to the right-hand side of equation (7.22) (or (7.28)-(7.29)). Numerical tests, comparison with other approaches, and recommendations will be reported in a future publication. Some viscous results for a perfect gas with hypersonic speed can be found in reference [66] and of the previous section.

7.5.4. A Numerical Example for a Loosely Coupled Point-Relaxation Implicit Method [79]

To illustrate the performance of a point-relaxation implicit algorithm using the symmetric TVD scheme (4.33) for three-dimensional chemically nonequilibrium flows, Gnoffo et al. applied the scheme to solve a configuration similar to the ones used in section 5.6. However, a free stream Mach number of 32 was used in this computation. The free stream conditions for this case were $V_\infty = 8917$ m/sec, $p_\infty = 1.54\ N/m^2$, and $T_\infty = 197°K$. The maximum body diameter was 14 ft., corresponding to the full-scale Aeroassist Flight Experiment (AFE) vehicle. The kinetic model of Dunn and Kang [199] was used which involves 11 species, $(N, O, N_2, O_2, NO, N^+, O^+, N_2^+, O_2^+, NO^+, e)$ and 26 reactions.

Figure (7.4) shows contour plots of electron number density. The global contour plot in figure (7.4a) serves only to define the shock layer. The blowup of the shock layer near the far shoulder in figure (7.4b) shows that contour lines run nearly parallel with the body and shock. Number densities vary from $2 \times 10^{13}/cm^3$ at the body to $10^{15}/cm^3$ at the shock. Initial conditions for this nonequilibrium chemistry test case were taken from a converged fine-grid perfect-gas solution.

The same relaxation procedure and limiter as described in section 5.6 were used, except now the species equations are solved separately from the fluid equations. Roe's averaging is used to define eigenvectors and eigenvalues for the fluid and species subset of equations separately (i.e. the full eigenvectors and eigenvalues of the fully coupled system were not used). Other thermodynamic derivatives' average values were defined in reference [79]. Other details and procedures which allow a good control of this loosely coupled procedure were discussed in reference [79]. The next section illustrates a fully coupled point-relaxation method computation.

7.5.5. A Numerical Example for a Fully Coupled Point-Relaxation Implicit Method [119]

Future plans for space transportation and exploration call for mission trajectories with sustained and/or maneuvering hypersonic flight in the Earth's atmosphere at altitudes greater than 70 km and velocities greater than 9 km/s. Aeroassisted orbital transfer vehicles will use this domain in returning from geosynchronous Earth orbit to low Earth orbit for rendezvous with the space station. Lunar, planetary, and comet sample return missions will utilize the Earth's upper atmosphere for aerobraking as well. Hypersonic, air-breathing cruise vehicles may ultimately be called on to fly through this domain. Substantial portions of these mission trajectories, in the transitional regime between free molecular and continuum, will carry the vehicle through conditions resulting in chemical and thermal nonequilibrium within the surrounding shock layer.

Nonequilibrium processes in the shock layer and near wake of hypersonic vehicles alter the flowfield in three important ways. Radiative energy transfer rates are sensitive to the electron temperature. Shock standoff distances and potential shock-body interactions are sensitive to the degree of dissociation. Local sound speeds, which influence pressure levels over aerodynamic expansion and compression surfaces, are sensitive to the partition of energy among the translational, rotational and vibrational modes.

The modeled system includes 11 species continuity equations, 3 momentum equations, and 3 energy equations describing the conservation of vibrational, electronic, and total energies. Species 1 through 5 are the neutral components of air consisting of N, O, N_2, O_2, and NO. Species 6 through 10 are the ions corresponding to species 1 through 5, in which one electron has been removed. Species 11 is the free electrons. The thermal nonequilibrium is modeled using a three-temperature approximation (i.e., three energy equations) and can be found in reference [120]

The simulation employed a point-implicit relaxation algorithm for fully coupled three-dimensional, viscous, hypersonic flows in chemical and thermal nonequilibrium. The algorithm is derived using a finite volume formulation in which the inviscid components of the numerical flux across cell walls are based on the symmetric TVD-type scheme (4.42) together with limiter (4.33d). The relaxation strategy is well suited for computers employing either vector or parallel architectures. The exact relaxation procedure and discussion of efficient implementation for either vector or parallel supercomputer architectures can be found in the full paper [119].

Hypersonic Flows Over the Aeroassist Flight Experiment (AFE) Configuration: Pressure contours over an AFE model, including the sting, for Mach 10 flow at 0^o angle of attack are shown in figure 7.5. The four-domain grid used for this case is presented in figure 7.6. Comparisons with experimental data for pressure and heat transfer are shown in figure 7.7 on the forebody and for heat transfer on the sting in figure 7.8. (Experimental data are unpublished results of investigations by John Micol and William Wells at NASA Langley Research Center.) Differences between numerical predictions and experiments are generally very small and within experimental accuracy except for the heat transfer in the stagnation region. These differences are believed to be caused by increased numerical dissipation associated with the required entropy fix in an extensive region of near zero eigenvalues (i.e., very small velocites).

A four-domain grid (not shown) is also used for the investigation of the hypersonic flow in both chemical and thermal nonequilibrium at 8917 m/s and 78 km altitude. The multi-domain grid allows for better resolution of the free shear layer coming off the shoulder as compared with the resolution achievable using a single-domain structured grid. Note that a hexagonal carrier vehicle replaces the sting in the base for the flight case. Contours for translational temperature, vibrational temperature, Mach number, and electron number density are shown in figures 7.9 - 7.12. The free shear layer structure is well defined in these figures. Note that contour lines continue smoothly over domain boundaries. The gaps between domains is an artifact of the cell centered structure of the data sets. The boundaries of the domains coincide, but there is a gap between cell centers. The CFL number used over most of the relaxation process was varied between 1 and 20. Estimated CPU time for the forebody domain (39 × 24 × 64) starting from uniform initial conditions is approximately 10 to 15 Cray-2 hours. Additional cases involving changes in free stream conditions or in the chemical model require 5 to 10 Cray-2 hours for the forebody. Another 10 hours is required for the 3 base domains (31 × 24 × 32, 31 × 24 × 19, 26 × 24 × 10).

All of the results presented here for flow in chemical and thermal nonequilibrium use a fully coupled relaxation algorithm. That is to say that, relaxation procedure for all conservation laws are simultaneously and implicitly coupled. This procedure is in contrast to a loosely coupled relaxation algorithm used in earlier work [200] in which the solution of the equations for mixture continuity, momentum, and energy conservation was alternated with a solution for species continuity. Coupling of the solutions utilized a variable effective gamma (ratio of enthalpy to internal energy) and renormalization of mixture density as described in [200]. The loosely coupled scheme exhibited satisfactory convergence rates on coarse grid, inviscid problems but failed to achieve acceptable convergence on fine grid, viscous problems. Direct comparisons between the two algorithms on exactly the same problem were never made. The loosely coupled algorithm was not maintained once the superior convergence characteristics of the fully coupled algorithm became apparent. A very rough estimate of computational savings based on Gnoffo's personal experience with the two algorithms is that the fully coupled solution algorithm can achieve the same drop in error norm as achievable using the loosely coupled algorithm in about half the CPU time. More importantly, the fully coupled algorithm can be driven to acceptable convergence levels whereas the loosely coupled algorithm "thrashed" at unacceptably high values of the global error norm.

VIII. Concluding Remarks

A unified and generalized formulation of a class of high-resolution explicit and implicit shock-capturing methods for perfect gases, equilibrium real gases and nonequilibrium flows containing complex shock waves has been presented. These methods were formulated for the purpose of ease and efficient implementation into new or existing computer codes. Some of the schemes were presented in a more convenient form for the generalization to systems via approximate Riemann solvers. They were written in a manner such that there is a smaller CPU operations count than other equivalent forms for their scalar counterparts. The various high-resolution constructions fall nicely into the present framwork. All of these methods can be implemented into the same computer code sharing many of the common operations.

This unified formulation was accompanied by practical fluid dynamics computations for one- and higher-dimensional problems containing complex shock waves, including hypersonic speed flows. In time-accurate computations where experimental data were available, good agreement between the numerical results and the experimental data were observed. Steady-state computations performed well for subsonic, transonic, supersonic and hypersonic flows. However, the convergence rate for hypersonic, and in particular viscous flows is in general slower than the inviscid counterpart, and improvement in this area is the current pacing item.

Three interrelated special topics were emphasized:

(1) Some issues concerning the applicability of these methods, which were designed for homogeneous hyperbolic conservation laws, to problems containing stiff source terms and shock waves were discussed and were illustrated with steady and unsteady examples. It is important to investigate finite difference methods for nonhomogeneous hyperbolic PDEs since nonequilibrium fluid flows contain coupled stiff source terms and this is an area of vital interest for hypersonic flows and combustion.

(b) The use of the Lax-Friedrichs numerical flux to obtain high-resolution shock-capturing schemes was generalized. This method can be extended to nonlinear systems of equations without the use of Riemann solvers or flux-vector splitting approaches and thus provides a large savings for multidimensional, equilibrium real gases and nonequilibrium flows computations.

(c) The importance of "asymptotic analysis of finite difference methods by the nonlinear dynamic approach" was briefly discussed. Many existing results on nonlinear dynamical systems such as chaos, bifurcations, and limit cycles have a direct application to problems containing nonlinear source terms such as the reaction-diffusion or the reaction-convection-diffusion equations. With the advent of increasing demands on numerical accuracy, stability, efficiency, and uniqueness of numerical solutions in modeling such equations, an interdisciplinary fertilization for the analysis of these types of systems is needed.

The applicability of these high-resolution shock-capturing methods for equilibrium real gases was studied. Preliminary results for one-dimensional shock-tube and two-dimensional steady-state blunt body problems show that the shock and contact discontinuity resolution were not affected by the state equation for a wide range of flow conditions.

Three numerical algorithms for hyperbolic conservation laws that are suitable for large systems of fully coupled thermally and chemically nonequilibrium steady-state flows in the hypersonic regime were proposed. The specific properties of the governing equations for fluid flow of this type were taken into consideration for more efficient solution procedures. The main areas of consideration were to minimize the operations count, increase the allowable time-step constraint imposed by the stiff source terms, and expand the shock-capturing capability beyond classical approaches. Details of all these considerations were described. Preliminary test problems showed certain advantages of the high-resolution shock-capturing scheme over the classical ways of supplying numerical dissipation. In general, the fully coupled method converges faster than the loosely coupled method and at the same time no additional constraints are involved (section 7.5.5). More numerical testing and study will be pursued in the immediate future.

In summary, the performance of the schemes presented for one- and two-dimensional gas-dynamics problems in conjunction with the various Riemann solvers can be divided into the following aspects.

Looking at the Riemann solvers, in general the advantages of the generalized van Leer flux-vector splitting over the generalized Steger-Warming formulation remain for an equilibrium real gas, with slightly less dissipation at the discontinuities. The local-characteristic approach (approximate Riemann solver) gives results very similar to the generalized van Leer flux-vector splitting formulation.

For one-dimensional problems, the difference in computational effort required by the three Riemann solvers is small. The main difference lies in the MUSCL and non-MUSCL approaches. The operations count between the non-MUSCL and MUSCL is within 30% for a perfect gas. However, due to extra evaluation in the curve fitting between the left and right states in an equilibrium real gas for the MUSCL formulation, without further approximations, additional computation is required for the MUSCL approach. The amount of additional computation increases nonlinearly as the spatial dimension increases.

Looking at the numerical schemes, the main difference seems to occur between the upwind and the symmetric approaches. The upwind schemes give better results for contact discontinuities. On the other hand, symmetric schemes have better stability and produce shock resolution similar to that of the upwind schemes; yet they require less CPU time per time-step and are less sensitive to numerical boundary condition treatments.

In the case of a non-MUSCL approach, limiter (4.33e) is the most accurate among (4.33c) - (4.33e) for the symmetric TVD scheme. As for the upwind schemes, limiters (4.34d) and (4.34e) are very similar, whereas limiters (4.34f) and (4.34g) give very accurate results for contact discontinuities but are too compressive, causing slight oscillations in smooth regions for high Mach number cases. A combination of limiters such as (4.34d) or (4.34e) for the nonlinear fields and (4.34g) for the linear field seems to be a good compromise. In the case of the MUSCL approach, only limiters (4.34c) and (4.34e) were studied. Between the two limiters, (4.34e) produces better shock resolution than (4.34c).

None of the differences for the various approaches observed for the explicit schemes seems to be decisive in one-dimensional tests, but factors such as stability or computational efficiency need further investigation in multidimensional tests. The main differences between the methods lie in their versatility in extending to implicit methods with efficient solution procedures, especially for multidimensional steady-state computations. This study shows certain advantages of the local-characteristic approach over the flux-vector splitting approaches.

There is an important distinction between the flux-vector splittings and the local-characteristic approach for implicit methods. Unlike flux-vector splitting approaches, implicit methods employing the local-characteristic approach (non-MUSCL or MUSCL with first-order implicit operator such as (4.44f)) do not require the Jacobian of the $F^{\pm}$ fluxes. In many instances, the Jacobian of $F^{\pm}$ is relatively difficult to obtain. A similar difficulty applies to the MUSCL formulation via the exact Riemann solver or local-characteristic approach (if a second-order implicit operator is desired).

Another important fact for equilibrium real gases computations is that flux-vector splittings make use of the sound speed only, whereas the local-characteristic approach of the Roe-type makes use of the thermodynamic dervatives χ and κ of equation (4.2). These thermodynamic derivatives put more stringent requirements on the curve fit that represents the thermodynamic properties of the gas. In this regard, the curve fits of Srinivasan et al. may be deficient for the approximate Riemann solver as can be seen from figure (4.1), case D. One probably needs more improved curve fits than those of reference [164] before a definite conclusion can be drawn about the accuracy of the different Riemann solvers and schemes for equilibrium real gases.

ACKNOWLEDGMENT

The author wishes to thank all of her collaborators (T. Aki, P. Gnoffo, A. Harten, G. Klopfer, R. LeVeque, R.S. McCandless, J.-L. Montagné, Y. Moon, J.L. Shinn, K. Uenishi, M. Vinokur, R.F. Warming, and V.Y.C. Young) for making the contents of this paper more complete. Special thanks to M. Vinokur for writing sections 4.2.1 - 4.2.3, to T. Aki for contributing to section 5.2.4, and to P. Gnoffo for contributing to sections 7.4.4 and 7.4.5.

REFERENCES

[1] H.C. Yee, "Upwind and Symmetric Shock-Capturing Schemes," NASA TM-89464, May, 1987.

[2] R.D. Richtmyer and K.W. Morton, *Difference Methods for Initial-Value Problems*, Interscience-Wiley, New York, 1967.

[3] J.D. Anderson, Jr, "Introduction to Computational Fluid Dynamics," von Karman Institute for Fluid Dynamics, 1985 lecture series, Rhode-Saint-Genèse, Belgium.

[4] D.A Anderson, J.C. Tannehil and R.H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, McGraw-Hill, New York, 1984.

[5] G.A. Sod, *Numerical Methods in Fluid Dynamics*, Cambridge, Cambridge, 1985.

[6] R.M. Beam and R.F. Warming, "Implicit Numerical Methods for the Compressible Navier-Stokes and Euler Equations," Lecture Notes for Computational Fluid Dynamics, von Karman Institute for Fluid Dynamics, March 29 - April 2, 1982, Rohde-Saint-Genèse, Belgium.

[7] P.L. Roe, "Characteristic-Based Schemes for the Euler Equations," Ann. Rev. Fluid Mech., Vol. 18, pp. 337-365, 1986.

[8] R.J. LeVeque, "Numerical Solution of Conservation Laws," Lecture Notes for a special course in numerical analysis, Dept. of Math., University of Washington, Fall, 1988.

[9] S.-W Shu, "TVB Boundary Treatment for Numerical Solutions of Conservation Laws," Math. Comp. Vol. 49, No. 179, 1987, pp. 123-134.

[10] H.C. Yee, "Numerical Approximation of Boundary Conditions with Applications to Inviscid Equations of Gas Dynamics," NASA TM-81265, 1981.

[11] H.C. Yee, R.M. Beam and R.F. Warming, "Boundary Approximations for Implicit Schemes for One-Dimensional Inviscid Equations of Gasdynamics," AIAA J., Vol. 20, No. 9, pp. 1203-1211, 1982.

[12] A. Harten and S. Osher, "Uniformly High-Order Accurate Nonoscillatory Schemes I," SIAM J. Num. Analy. Vol. 24, No. 2, pp. 279-309, 1987.

[13] A. Harten, S. Osher, B. Engquist and S. Chakravarthy, "Some Results on Uniformly High Order Accurate Essentially Non-oscillatory Schemes," ICASE Report 86-18, Mar. 1986.

[14] A. Harten, "On High-Order Accurate Interpolation for Non-Oscillatory Shock-Capturing Schemes," The IMA Volumes in Mathematics and its Applications, Vol. 2, Springer-Verlag, pp. 71-106, 1986.

[15] B. van Leer, "Towards the Ultimate Conservation Difference Scheme V, A Second-Order Sequel to Godunov's Method," J. Comp. Phys. Vol. 32, pp. 101-136, 1979.

[16] P. Colella and P.R. Woodward, "The Piecewise Parabolic Method (PPM) for Gas Dynamical Simulations," J. Comp. Phys. Vol. 54, pp. 174-201, 1984.

[17] A. Harten, "On a Class of High Resolution Total-Variation-Stable Finite-Difference Schemes," SIAM J. Num. Anal., Vol. 21, pp. 1-23, 1984.

[18] P.L. Roe, "Some Contributions to the Modelling of Discontinuous Flows," Lectures in Applied Mathematics, Vol. 22 (Amer. Math. Soc., Providence, R.I., 1985).

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

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

[21] H.C. Yee, "Construction of Explicit and Implicit Symmetric TVD Schemes and Their Applications," J. Comput. Phys., Vol. 68, pp. 151-179, 1987; also NASA TM-86775 July 1985.

[22] S. Osher and S. Chakravarthy, "Very High Order Accurate TVD Schemes," The IMA Volumes in Mathematics and its Applications, Vol. 2, Springer-Verlag, pp. 229-274, 1986.

[23] A. Jameson and P.D. Lax, "Conditions for the Construction of Multi-Point Total Variation Diminishing Difference Schemes," ICASE Report No. 86-18, pp. 361-382, Mar. 1986.

[24] A. Harten, J.M. Hyman and P.D. Lax, "On Finite-Difference Approximations and Entropy Conditions for Shocks," Comm. Pure Appl. Math., Vol. 29, pp. 297-322, 1976.

[25] M.G. Crandall and A. Majda, "Monotone Difference Approximations for Scalar Conservation Laws," Math. Comp. Vol 34, No. 149, pp. 1-21, 1980.

[26] P.D. Lax and B. Wendroff, "Difference Schemes for Hyperbolic Equations with High Order of Accuracy," Commun. Pure Appl. Math., Vol. 17, pp. 381-398, 1964.

[27] R.W. MacCormack, "The Effect of Viscosity in Hypervelocity Impact Cratering," AIAA Paper No. 69-354, Cincinnati, Ohio, 1969.

[28] R.M. Beam and R. F. Warming, "An Implicit Finite-Difference Algorithm for Hyperbolic Systems in Conservation Law Form," J. Comp. Phys., Vol. 22, pp. 87-110, 1976.

[29] N.A. Phillips, "An Example of Nonlinear Computational Instability," *The Atmosphere and the Sea in Motion*, B. Bolin ed., Rockefeller Institute, New York, 1959.

[30] J.M. Sanz-Serna, "Studies in Numerical Nonlinear Instability I: Why do leapfrog schemes go unstable?," SIAM J. Sci. Stat. Comput. Vol. 6, 1985.

[31] J.M. Sanz-Serna and F. Vadillo, "Nonlinear Instability, the Dynamic Approach," in Proceedings Dundee, 1985, G.A. Watson and D.F. Griffiths, eds., Pitman, London.

[32] F. Vadillo and J.M. Sanz-Serna, "Studies in Numerical Nonlinear Instability II: A New Look at $u_t + uu_x = 0$," J. Comp. Phys., Vol. 66, pp225-238, 1986.

[33] 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.

[34] J.B. Goodman and R.J. LeVeque, "On the Accuracy of Stable Schemes for 2D Scalar Conservation Laws," Math. Comp., Vol. 45, 1985, pp. 15-21.

[35] B. Gustafsson, "A Semi-Implicit Method for Almost Incompressible Flow Problems on Parallel Computers," Proceedings of the 1st International Symposium on Computational

Fluid Dynamics, Tokyo, 1985.

[36] B. Gustafsson, “Unsymmetric Hyperbolic Systems and the Euler Equations at Low Mach-Number, “Uppsala University, Dept. of Computer Sciences, Report No. 104, May 1986.

[37] A.R. Mitchell, *Computational Methods in Partial Differential Equations*, Wiley and Son, New York, 1976.

[38] P.R. Garabedian, *Partial Differential Equations*, Wiley and Son, New York, 1964.

[39] H.C. Yee, “On the Implementation of a Class of Upwind Schemes for System of Hyperbolic Conservation Laws,” NASA TM-86839, Sept. 1985.

[40] 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, Vol. 22, 1985.

[41] H.C. Yee and A. Harten, “Implicit TVD Schemes for Hyperbolic Conservation Laws in Curvilinear Coordinates,” AIAA Paper No. 85-1513, Proceedings of the AIAA 7th CFD conference, Cincinnati, Ohio, July 1985; AIAA J., Feb. 1987.

[42] R.F. Warming and R.M. Beam, “Lecture Notes on the Stability and Accuracy of Difference Approximations for Initial-Boundary-Value Problems,” Manuscript in preparation.

[43] P.D. Lax and B. Wendroff, “Systems of Conservation Laws,” Commun. Pure Appl. Math., Vol. 13, pp. 217-237, 1960.

[44] A.Y. LeRoux, “A Numerical Conception of Entropy for Quasi-Linear Equations,” Math. Comput., Vol. 31, No. 140, pp. 848-872, 1977.

[45] A. Majda and S. Osher, “Numerical Viscosity and the Entropy Condition,” Commun. Pure Appl. Math., Vol. 32, pp. 797-838, 1979.

[46] H.C. Yee, “Asymptotic Analysis of Finite Difference Methods by the Nonlinear Dynamic Approach,” in preparation.

[47] R. Courant, E. Isaacson, and M. Rees, “On the Solution of Nonlinear Hyperbolic Differential Equations by Finite Differences,” Comm. Pure Appl. Math., Vol. 5, pp 243-255, 1952.

[48] P.L. Roe, “Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes,” J. Comp. Phys., Vol. 43, pp. 357-372, 1981.

[49] S.K. Godunov, “A Finite Difference Method for the Numerical Computation of Discontinuous Solutions of the Equations of Fluid Dynamics,” Mat. Sb. 47, pp. 357-393, 1959.

[50] Engquist and S. Osher, “One-Sided Difference Approximations for Non-linear Conservation Laws,” Math. Comput., Vol. 36, pp. 321-351, 1981.

[51] L.C. Huang, “Pseudo-Unsteady Difference Schemes for Discontinuous Solutions of Steady-State, One-Dimensional Fluid Dynamics Problems,” J. Comp. Phys., Vol. 42, pp. 195-211, 1981.

[52] P.L. Roe, “The use of the Riemann Problem in Finite-Difference Schemes,” Lect. Notes Phys. Vol. 141, pp. 354-359, 1980.

[53] A. Harten and J.M. Hyman, "A Self-Adjusting Grid for the Computation of Weak Solutions of Hyperbolic Conservation Laws," J. Comp. Phys., Vol. 50, pp. 235-269, 1983.

[54] P.K. Sweby and M.J. Baines, "Convergence of Roe's Scheme for the General Non-linear Scalar Wave Equation," Reading University Numerical Analysis Report, Aug. 1981.

[55] S. Osher, "Riemann Solvers, The Entropy Condition and Difference Approximations," SIAM J. Numer. Analy. Vol. 21, pp. 217-235, 1984.

[56] S.-W. Shu, "TVB Uniformly High-Order Schemes for Conservation Laws," Math. Comp., Vol. 49, No. 179, 1987, pp. 105-121.

[57] E. Tadmor, "Convenient Total Variation Diminishing Conditions for Nonlinear Difference Schemes," SIAM J. Numer. Anal., Vol. 25, No. 5 1988, pp. 1002-1014.

[58] S. Osher, "Convergence of Generalized MUSCL Schemes," SIAM J. Numer. Anal. Vol. 22, No. 5, 1985, pp. 947-961.

[59] S. Osher, "Shock Modeling in Transonic and Supersonic Flow," Recent Advances in Numerical Methods in Fluids, Vol. 4, Advances in Computational Transonics, W.G. Habashi Ed., Pineridge Press, 1984.

[60] P.L. Roe, "A Survey of Upwind Differencing Techniques," Proceedings of the 11th International Conference on Numerical Methods in Fluid Dynamics, June 1988, Williamsburg, VA., USA.

[61] S.-W. Shu, "Total-Variation-Diminishing Time Discretizations," SIAM J. Sci. Stat. Comput., Vol. 9, No. 6, 1988, pp. 1073-1084.

[62] J.B. Goodman and R.J. LeVeque, "A Geometric Approach to High-Resolution TVD Schemes," SIAM J. Numer. Anal., Vol. 25, 1988, pp. 268-284.

[63] B. Einfeldt, "Zur Numerik der Stroauflösenden Verfahren," Ph. D. dissertation, RWTH, Aachen, Feb. 1988.

[64] B. Einfeldt, "On Positive Shock-Capturing Schemes," College of Aeronautics, COA8810, August 1988, Cranfiled Institute Technology, Bedford, England.

[65] B. Engquist, P. Löstedt and B. Sjögreen, "Nonlinear Filters for Efficient Shock Computation," Dept. of Math., Math Report, 1988, UCLA, California, USA.

[66] H.C. Yee, G. Klopfer and J.-L. Montagné, "High-Resolution Shock-Capturing Schemes for Inviscid and Viscous Hypersonic Flows," NASA TM-100097, April, 1988.

[67] J.P. Boris and D.L. Book, "Flux-Corrected Transport. I. SHASTA, A Fluid Transport Algorithm That Works," J. Comp. Phys., Vol. 11, pp. 38-69, 1973.

[68] A. Harten, "The Artificial Compression Method for Computation of Shocks and Contact Discontinuities: III. Self-Adjusting Hybrid Schemes," Math. Comp., Vol. 32, No. 142, pp. 363-389, 1978.

[69] B. van Leer, "Towards the Ultimate Conservative Difference Scheme. II. Monotonicity and Conservation Combined in a Second-Order Scheme," J. Comp. Phys., Vol. 14, pp. 361-370, 1974.

[70] P.K. Sweby, "High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws," SIAM J. Num. Analy., Vol. 21, pp. 995-1011, 1984.

[71] S. Zalesak, "A Preliminary Comparison of Modern Shock-Capturing Scheme: Linear Advection," *Advances in Computer Methods for Partial Differential Equations, VI,* Vichnevetsky and Stepleman Eds., International Association for Mathematics and Computers in Simulation (IMACS), Rutgers University, New Brunswick, NJ, 1987.

[72] van Albada et al., Astron. Astrophysics, Vol. 108, 1982, pp. 76-84.

[73] P. Colella, "A Direct Eulerian MUSCL Scheme for Gas Dynamics," SIAM J. Scientific and Statistical Computing, Vol. 6 No. 1, 1985, pp. 104-117.

[74] H.C. Yee, "Linearized Form of Implicit TVD Schemes for Multidimensional Euler and Navier-Stokes Equations," Computers and Mathematics with Applications, Vol. 12A, pp. 413-432, 1986.

[75] H.C. Yee, "Numerical Experiments with a Symmetric High-Resolution Shock-Capturing Scheme," Proceedings of the 10th International Conference on Numerical Methods in Fluid Dynamics, June 1986, Beijing, China; also NASA TM-88325, June 1986.

[76] H.C. Yee, "High-Resolution Methods Using the Lax-Friedrichs Numerical Flux," in preparation.

[77] W.K. Anderson, J.L. Thomas and B. van Leer, "A Comparison of Finite Volume Flux Vector Splittings for Euler Equations," AIAA Paper No. 85-0122, 1985.

[78] S. Chakravarthy and K.Y. Szema, "An Euler Solver for Three-Dimensional Supersonic Flows with Subsonic Pockets," AIAA Paper No. 85-1703, June 1985.

[79] P.A. Gnoffo, R.S. McCandless and H.C. Yee "Enhancements to Program LAURA for Efficient Computation of Three-Dimensional Hypersonic Flow," AIAA Paper No. 87-0280, Jan. 1987.

[80] H.C. Yee, "Construction of a Class of Symmetric TVD Schemes" The IMA Volumes in Mathematics and its Applications, Vol. 2, Springer-Verlag, pp. 381-396, 1986.

[81] H.C. Yee, "Generalized Formation of a Class of Explicit and Implicit TVD Schemes," NASA TM-86775, July, 1985.

[82] P. Woodward and P. Colella, "The Numerical Simulation of Two-Dimensional Fluid Flow with Strong Shocks," J. Comput. Phys., Vol. 54, pp. 115-173, 1984.

[83] R.F. Warming and H.C. Yee, "Convergence Rate of Total Variation Diminishing (TVD) Schemes for Hyperbolic Conservation Laws," SIAM National Meeting, Denver, Colo., June 6-8, 1983.

[84] H. Nessyahu and E. Tadmor, "Non-Oscillatory Central Differencing for Hyperbolic Conservation Laws," ICASE Report No. 88-51, Sept. 1988.

[85] S. Osher, "Riemann Solvers, the Entropy Condtion and Difference Approximations," SIAM J. Numer. Anal., Vol. 21, 1984, pp. 217-238.

[86] R.J. LeVeque and H.C. Yee, "A Study of Numerical Methods for Hyperbolic Conservation Laws with Stiff Source Terms," NASA TM-100075, March 1988, also to appear J. Comput. Phys., 1989.

[87] P.K. Sweby, "TVD Schemes for Inhomogeneous Conservation Laws," Proceedings of the ICFD Conference, March 1988, Oxford, England.

[88] P.L. Roe, "Upwind Differencing Schemes for Hyperbolic Conservation Laws with Source Terms," Proc. Nonlinear Hyperbolic Problems, Eds. C. Carasso, P.A. Raviart, D. Serre, Springer Lecture Notes in Mathematics, 1986.

[89] M. Ben-Artzi and J. Falcovitz, "An Upwind Second-Order Scheme for Compressible Duct Flows," Dept. of Math., Technion Preprint Series No. MT625, Dec., 1983, Technion-Israel Institute of Technology, Haifa, Israel.

[90] P. Glaister, "Flux Difference Splitting Techniques for the Euler Equations in Non-Cartesian Geometry," Numerical Analysis Report 8/85, Dept. of Math., University of Reading, England.

[91] M. Pandolfi, "Upwind Formulations for the Euler Equations," von Karman Institute for Fluid Dynamics Lecture Series on Computational Fluid Dynamics, March 2-6, 1987, Rhode-St-Genèse, Belgium.

[92] M. Pandolfi, M. Germano and N. Botta, "Non-Equilibrium Reacting Hypersonic Flow About Blunt Bodies: Numerical Prediction, AIAA Paper No. 88-0514, Jan. 1988.

[93] R.F. Warming, P. Kutler and H. Lomax, "Second- and Third-Order Noncentered Difference Schemes for Nonlinear Hyperbolic Equations," AIAA J., pp. 189-196, Feb. 1973.

[94] C.M. Kwong, "Numerical Experiments with a Total Variation Diminishing (TVD) MacCormack Scheme," Proceedings of the 6th GAMM conference on Numerical Methods in Fluid Mechanics, Gottingen, West Germany, Sept. 1985.

[95] H.C. Yee and J.L. Shinn, "Semi-Implicit and Fully Implicit Shock-Capturing Methods for Hyperbolic Conservation Laws with Stiff Source Terms," AIAA Paper No. 87-1116, June, 1987, also to appear AIAA J., 1989.

[96] T.R.A. Bussing and E.M. Murman, "Finite Volume Method for the Calculation of Compressible Chemically Reacting Flows," AIAA Paper No. 85-0331, Jan. 1985.

[97] G. Strang, "On the Construction and Comparison of Difference Schemes," SIAM J. Numer. Anal. Vol. 5, pp. 506-517, 1968.

[98] J.P. Boris and E.S. Oran, *Numerical Simulation of Reactive Flow* (Elsevier, 1987).

[99] G.C. Carofano, "Blast Computation Using Harten's Total Variation Diminishing Scheme," Technical Report ARLCB-TR-84029, Oct. 1984.

[100] P. Colella, A. Majda, and V. Roytburd, "Theoretical and Numerical Structure for Reacting Shock Waves, SIAM J. Sci. Stat. Comput. Vol. 7, 1986, pp. 1059-1080.

[101] C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations* (Prentice-Hall, 1971).

[102] J.D. Lambert, *Computational Methods in Ordinary Differential Equations* (Wiley, 1973).

[103] T.R. Young and J.P. Boris, "A Numerical Technique for Solving Stiff Ordinary Differential Equations Associated with the Chemical Kinetics of Reactive-Flow Problems," J. Phys. Chem. Vol. 81, 1977, pp. 2424-2427.

[104] H.C. Yee, R.F. Warming and A. Harten, "Implicit Total Variation Diminishing Schemes (TVD) for Steady-State Calculations," J. Comput. Phys., Vol. 57, No. 3, pp. 327-360, 1985.

[105] A. Iserles, "Nonlinear Stability and Asymptotics of O.D.E. Solvers," International Conference on Numerical Mathematics, Singapore, (R.P. Agarwal, ed.), Birkhauser, to appear.

[106] A.R. Mitchell and D.F. Griffiths, "Beyond the Linearized Stability Limit in Non Linear Problems," Report NA/88 July 1985, Department of Mathematical Sciences, University of Dundee, Scotland U.K.

[107] B.D. Sleeman, D.F. Griffiths, A.R. Mitchell and P.D. Smith, "Stable Periodic Solutions in Nonlinear Difference Equations," SIAM J. Sci. Stat. Comput., Vol. 9, N0. 3, May 1988, pp. 543-557.

[108] A.R. Mitchell and J.C. Bruch, Jr., "A Numerical Study of Chaos in a Reaction-Diffusion Equation," Numerical Methods for PDEs, Vol. 1, 1985, pp. 13-23.

[109] D.F. Griffiths and A.R. Mitchell, "Stable Periodic Solutions of a Nonlinear Partial Difference Equation in Reaction Diffusion," Report NA/113, Jan. 1988, Dept. Math. and Compt. Science, University of Dundee, Scotland U.K.

[110] R.L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Addison Wesley, New York, 1987.

[111] R. Seydel, *From Equilibrium to Chaos*, Elsevier, New York, 1988.

[112] P. Colella, "Glimm's Method for Gas Dynamics," SIAM J. Sci. Stat. Comput., Vol. 30, pp. 76-110, 1982.

[113] A. Harten, P. D. Lax and B. van Leer, "On upstream Difference and Godunov-Type Schemes for Hyperbolic Conservative Laws," SIAM Rev., Vol. 25, pp. 35-61, 1983.

[114] J.L. Steger and R.F. Warming, "Flux Vector Splitting of the Inviscid Gasdynamic Equations with Application to Finite Difference Methods," J. Comput. Phys., Vol. 40, pp. 263-293, 1981.

[115] B. van Leer, "Flux-vector splitting for the Euler Equations," ICASE Report 82-30; Sept., 1982.

[116] M. Vinokur, "Flux Jacobian Matrices and Generalized Roe Average for an Equilibrium Real Gas," NASA Contractor Report 177512, Dec. 1988.

[117] M. Vinokur and J.-L. Montagné, "Generalized Flux-Vector Splitting for an Equilibrium Real Gas," NASA contractor report 177513, Dec. 1988.

[118] J.L. Shinn, H.C. Yee and K. Uenishi, "Extension of a Semi-Implicit Shock-Capturing Algorithm for 3-D Fully Coupled Chemically Reacting Flows in Generalized Coorinates," AIAA Paper No. 87-1577, June, 1987.

[119] P.A. Gnoffo, "Upwind-Biased Point-Implicit Relaxation Strategies for Hypersonic Flowfield Simulations on Supercomputers," in preparation.

[120] P.A. Gnoffo, R.N. Gupta and J.L. Shinn, "Conservation Equations and Physical Models for Hypersonic Air Flows in Thermal and Chemical Nonequilibrium," NASA TP-2867, Dec. 1988.

[121] C.C. Lytton, "Solution of the Euler Equations for Transonic Flow Past a Lifting Aerofoil - the Bernoulli Formulation," RAE Tech. Rep. 84080, 1984.

[122] J.L. Thomas and R.W. Walters, "Upwind Relaxation Algorithms for the Navier-Stokes

Equations," AIAA Paper No. 85-1501, July, 1985.

[123] J.L. Thomas, B. van Leer and R.W. Walters, "Implicit Flux-Split Schemes for the Euler Equations," AIAA Paper No. 85-1680, July, 1985.

[124] S.R. Chakravarthy and S. Osher, "High Resolution Applications of the Osher Upwind Scheme for the Euler Equations," AIAA Paper No. 83-1943, July 1983.

[125] S. R. Chakravarthy and S. Osher, "A New Class of High Accuracy TVD Schemes for Hyperbolic Conservation Laws," AIAA Paper No. 85-0363, AIAA 23rd Aerospace Sciences Meeting, Reno, Nevada, Jan. 14-17, 1985.

[126] A. Jameson, "Nonoscillatory Shock Capturing Scheme Using Flux Limited Dissipation," Lectures in Applied Mathematics, Vol. 22 (Amer. Math. Soc., Providence, R.I., 1985).

[127] G. Patnaik, J.P. Boris, R.H. Guirguis and E. Oran, "A Barely Implicit Correction for Flux-Corrected Transport," NRL Memo Rep. 5855, 1987.

[128] R. Lohner and G. Patnaik, "BIC-FEM-FCT: An Algorithm for Low-Mach-Number Flows," AIAA Paper No. 87-1146, June 1987.

[129] G. Moretti, "The Lambda Scheme," Computers and Fluids, Vol. 7, 1979.

[130] G. Moretti, "An Efficient Euler Solver with Many Applications," AIAA Paper No. 87-0352, Jan. 1987.

[131] G. Moretti, "An Optimistic Reappraisal of Computational Techniques in the Supercomputer Era," Proceedings of the ICFD Conference, March 21-24, 1988, Oxford, England.

[132] S. Chakravarthy, D. A. Anderson and M.D. Salas, "The Split Coefficient Matrix Method for Hyperbolic System of Gasdynamic Equations," AIAA Paper No. 80-0268, 1980.

[133] M. Ben-Artzi and J. Falcovitz, "A Second-Order Godunov-Type Scheme for Compressible Fluid Dynamics," J. Comput. Phys., Vol. 55, pp. 1-32, 1984.

[134] J. Glimm, "Solutions in the Large for Nonlinear Hyperbolic Systems of Equations," Commun. Pure Appl. Math., Vol. 18, pp. 697-715, 1965.

[135] A.J. Chorin, "Random Choice Solution of Hyperbolic Systems," J. Comput. Phys. Vol. 22, pp. 517-533, 1976.

[136] E.F. Toro, "A New Numerical Technique for Quasi-Linear Hyperbolic Systems of Conservation Laws," CoA Report No. 86/26, Dec. 1986, College of Aeronautics, Cranfield Institute of Technology, U.K.

[137] P.L. Roe and E.F. Toro, "A Second-Order Random-Choice Algorithm for Hyperbolic Systems of Equations," SIAM Conference on Numerical Combustion, March 9-11, 1987, San Francisco, Calif.

[138] E.F. Toro and J.F. Clarke, "Application of the Random Choice Method to Computing Problems of Solid-Propellant Combustion in a Closed Vessel," CoA Report NFP85/16, 1985, College of Aeronautics, Cranfield Institute of Technology, U.K.

[139] S. Osher and F. Solomon, "Upwind Schemes for Hyperbolic Systems of Conservation Laws," Math. Comp., Vol. 38, pp. 339-377, 1981.

[140] S.F. Davis, "Simplified Second-Order Godunov-Type Methods," Naval Surface Weapons Center Report, 1986, R44, Silver Spring, Maryland.

[141] B. Einfeldt, "On Godunov-Type Methods for Gas Dynamics," SIAM J. Numer. Anal., Vol. 25, No. 2, April 1988.

[142] T.W. Roberts, "The Behavior of Flux Difference Splitting Schemes Near Slowly Moving Shock Waves," Proceedings of the ICFD Conference, March 1988, Oxford, England.

[143] R. H. Sanders and K. H. Prendergast, "The Possible Relation of 3-Kiloparsec Arm to Explosions in the Galactic Nucleus," Astrophys. J., Vol. 188, pp. 489-500, 1974.

[144] W.A. Mulder and B. van Leer, "Implicit Upwind Methods for the Euler Equations," AIAA Paper No. 83-1930, July 1983.

[145] P. Colella and H.R. Glaz, "Efficient Solution Algorithm for the Riemann Problem for Real Gases," J. Comp. Phys., Vol. 59, pp. 254-289, 1985.

[146] J.K. Dukowicz, "A General, Non-Iterative Riemann Solver for Godunov's Method," J. Comput. Phys., Vol. 61, pp. 119-137, 1985.

[147] M. Ben-Artzi, "The Generalized Riemann Problem for Reactive Flows," LBL-21218, April 1987.

[148] T.P. Liu, "The Riemann Problem for General Systems of Conservation Laws," J. Diff. Equ., Vol. 18, 1975, pp. 218-234.

[149] M. Vinokur and Y. Liu, "Equilibrium Gas Flow Computations II. An Analysis of Numerical Formulations of Conservation Laws," AIAA Paper No. 88-0127, January 1988, Reno, Nevada.

[150] J.-L. Montagné, H.C. Yee and M. Vinokur, "Comparative Study of High-Resolution Shock-Capturing Schemes for Real Gas," Proceeding of the 7th GAMM conference on Numerical Methods in Fluid Mechanics, Sept. 9-11, 1987, Belgium, also NASA TM-86839, July 1987.

[151] M. Vinokur and Y. Liu, "Equilibrium Gas Flow Computations. II. An Analysis of Numerical Formulations of Conservation Laws", AIAA paper No. 88-0127, Jan. 1988.

[152] B. Grossman and R. W. Walters, "An Analysis of Flux-Split Algorithms for Euler's Equations with Real Gases", AIAA paper No. 87-1117, June 1987.

[153] P. Glaister, "An Approximate Linearised Riemann Solver fot the Euler Equations for Real Gases", J. Comp. Phys., Vol. 74, pp. 382-408, 1988.

[154] M.-S. Liou, B. van Leer and J.-S. Shuen, "Splitting of Inviscid Fluxes for Real Gases", NASA TM 100856, April 1988.

[155] Y. Liu and M. Vinokur, "Equilibrium Gas Flow Computations. I. Accurate and Efficient Calculation of Equilibrium Gas Properties", NASA Contractor Report, in preparation.

[156] G.Palmer, "An Implicit Flux-Split Algorithm to Calculate Hypersonic Flow Field in Chemical Equilibrium," AIAA paper No. 87-1580, June 1987.

[157] J.-L. Montagné, "Use of an Upwinded Scheme for Simulating NonViscous Flows of Real Gas at Equilibrium," La recherche Aerospatiale, Dec. 1986.

[158] J.-L. Montagné, "Rapport de Synthese Finale," ONERA, 23/1285 AY, Dec. 1985.

[159] D. Hänel, R. Schwane and G. Seider, "On the Accuracy of Upwind Schemes for the Solution of the Navier-Stokes Equations", AIAA paper No. 87-1105, June 1987.

[160] Y.J. Moon and H.C. Yee, "Numerical Simulation by TVD Schemes of Complex Shock Reflections from Airfoils at High Angle of Attack," AIAA Paper No. 87-0350, Jan. 1987.

[161] V.Y.C. Young and H.C. Yee, "Numerical Simulation of Shock Wave Diffraction by TVD Schemes," AIAA Paper No. 87-0112, Jan. 1987.

[162] T. Aki, " A Numerical Study of Shock Propagation in Channels with 90° Bends," a National Aerospace Laboratory technical report, Tokyo, Japan, 1987.

[163] H.C. Yee, "On Symmetric and Upwind TVD Schemes," Proceedings of the 6th GAMM conference on Numerical Methods in Fluid Mechanics, Gottingen, West Germany, Sept. 1985.

[164] S. Srinivasan, J.C. Tannehill, K.J. Weilmunster, "Simplified Curve Fit for the Thermodynamic Properties of Equilibrium Air," ISU-ERI-Ames 86401; ERI Project 1626; CFD15, 1986.

[165] R.W. Walter, J.L. Thomas and B. van Leer, "An Implicit Flux-Split Algorithm for the Compressible Euler and Navier-Stokes Equations," Proceedings of the 10th International Conference on Numerical Methods in Fluid Dynamics, June 1986, Beijing, China.

[166] H.C. Yee, and P. Kutler, "Application of Second-Order-Accurate Total Variation Diminishing (TVD) Schemes to the Euler Equations in General Geometries," NASA TM-85845, Aug., 1983

[167] Y. Takakura, T. Ishiguro, and S. Ogawa, "On the Recent Difference Schemes for Three-Dimensional Euler Equations," AIAA Paper No. 87-1151-CP, Proceedings of the AIAA 8th Computational Fluid Dynamics Conference, June 9-11, 1987, Honolulu, Hawaii, pp. 537-545.

[168] M. Vinokur, "An Analysis of Finite-Difference and Finite-Volume Formulations of Conservation Laws," NASA Contractor Report 177416, June 1986. Also, J. Comput. Phys. (in press)

[169] A.E. Bryson and R.W.F. Gross, "Diffraction of Strong Shocks by Cone, Cylinder, and Sphere," J. Fluid Mech., Vol. 10, pp. 1-16, 1961.

[170] M. Mandella, "Experimental and Analytical Studies of Compressible Vortices," Ph.D. thesis, Dept. of Applied Physics, Stanford University, Stanford, Calif., 1987.

[171] M. Mandella and D. Bershader, "Quantitative Study of Compressible Vortices: Generation, Structure and Interaction with Airfoils," AIAA Paper No. 87-0328, Jan. 1987.

[172] B.W. Skews, "The Perturbed Region Behind a Diffracting Shock Wave," J. Fluid Mechanic, Vol. 29, Part 4, pp. 705-719, 1967.

[173] T. Aki, "A TVD Scheme Study on Nonstationary Shock Reflection," Proceedings of the first SUJSCFD, Khadarovsh, Sept. 9-15, 1988.

[174] T. Aki, private communciation and manuscript in preparation 1989.

[175] K. Takayama et al., "Shock Propagation along 90 Degree Bends,", Report Inst. High Speed Mech., Tohoku Univ., Japan, Vol. 35, pp. 1-30, 1977.

[176] T.J. Barth, "Analysis of Implicit Local Linearization Techniques for Upwind and TVD Algorithms," AIAA Paper No. 87-0595, Reno, Nevada, 1987.

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

[178] B. Baldwin and H. Lomax, "Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows," AIAA Paper No. 78-257, 1978.

[179] J.-L. Montagné, H.C. Yee, G.H. Klopfer and M. Vinokur, "Hypersonic Blunt Body Computations Including Real Gas Effects," Proceedings of the 2nd International Conference in Hyperbolic Problems, March 14-18, 1988, Aachen, Germany, also NASA TM-100074, March 1988.

[180] G. Klopfer and H.C. Yee, "Viscous Hypersonic Shock on Shock Interaction on Blunt Cowl Lips," AIAA Paper No. 88-0233, AIAA 26th Aerospace Sciences Meeting, Jan. 11-14, 1988, Reno, Nevada.

[181] G. Klopfer, H.C. Yee and P. Kutler, "Numerical Study of Unsteady Viscous Hypersonic Blunt Body Flows With An Impinging Shock," Proceedings of the 11th International Conference on Numerical Methods in Fluid Dynamics, June 27 - July 1, 1988, Williamsburg, Virginia, also NASA TM-100096, April 1988.

[182] J.C. Tannehill, T.L. Holst and J.V. Rakich, "Numerical Computation of Two-Dimensional Viscous Blunt Body Flows with an Impinging Shock," AIAA J., Vol. 14, No. 2, 1976, pp. 204-211.

[183] B. Müller, "Simple Improvements of an Upwind TVD Scheme for Hypersonic Flow," Manuscript in preparation.

[184] P.A. Gnoffo and R.S. McCandless, "Three-Dimensional AOTV Flowfields in Chemical Nonequilibrium," AIAA Paper No. 86-0230, Jan. 1986.

[185] P.L. Roe and J. Pike, "Efficient Construction and Utilisation of Approximate Riemann Solutions, Computing Methods in Applied Sciences and Engineering, ed. R. Glowinski, North-Holland, Amsterdam, J.-L. Lions, 1984, pp. 499-518.

[186] B.E. Edney, "Anomalous Heat Transfer and Pressure Distributions on Blunt Bodies at Hypersonic Speeds in the Presence of an Impinging Shock," FFA Report 115, February 1968, The Aeronautical Research Institute of Sweden, Stockholm, Sweden.

[187] J.C. Tannehill, T.L. Holst and J.V. Rakich, "Comparison of a Two-Dimensional Shock Impingement Computation with Experiment," AIAA J., Vol. 14, No. 4, 1976, pp. 539-541.

[188] J.A. White and C.M. Rhie, "Numerical Analysis of Peak Heat Transfer Rates for Hypersonic Flow over a Cowl Leading Edge," AIAA Paper No. 87-1895, June 1987.

[189] K. Morgan, J. Peraire, R.R. Thareja and J.R. Stewart, "An Adaptive Finite Element Scheme for the Euler and Navier-Stokes Equations", AIAA Paper No. 87-1172-CP, Proceedings of the AIAA 8th Computational Fluid Dynamics Conference, June 9-11, 1987, Honolulu, Hawaii, pp. 749-757.

[190] S. Eberhardt and K. Brown, "A Shock Capturing Technique for Hypersonic, Chemically Relaxing Flows," AIAA Paper No. 86-0231, Jan. 1986.

[191] H. Glaz and P. Colella, "High-Resolution Calculations of Unsteady Two-Dimensional Nonequilibrium Gas Dynamics with Experimental Comparisons," AIAA Paper No. 87-

1293, June 8-10, 1987, Honolulu, Hawaii.

[192] Y. Liu and M. Vinokur, "Nonequilibrium Flow Computations: I. An Analysis of Numerical Formulation of Conservation Laws," NASA Contractor Report 177489, June 1988. Also, J. Comput. Phys. (to appear).

[193] B. Grossman and P. Cinnella, "Flux-Split Algorithms for Flows with Non-Equilibrium Chemistry and Vibrational Relaxation," ICAM Report 88-08-03, Dept. Aerospace & Ocean Enginerring, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, August 1988.

[194] B. Larrouturou and L. Fezoui, "On the Equations of Multi-Component Perfect or Real Gas Inviscid Flow," to appear in "Nonlinear Hyperbolic Problems" Carasso, Charrier, Hanouzet and Joly Eds., Lectrue Notes in Mathematics, Springer-Verlag, Heidelberg, 1989.

[195] R. Abgrall, "Preliminary Results on an Extension of Roe's Approximate Riemann Solver to Nonequilibrium Flows," INRIA/Sophia Antipolis Center Report, Feb., 1989, Valbonne, France.

[196] K. Uenishi and R.C. Rogers, "Three-Dimensional Computation of Transverse Hydrogen Jet Combustion in a Supersonic Air Stream," AIAA Paper No. 87-0089, Jan. 1987.

[197] W.R. Van Dalsen and J.L. Steger, "Using the Boundary-Layer Equations in Three-Dimensional Viscous Flow Simulation," NASA TM-88241, Mar., 1986.

[198] C.P. Li, "Chemistry-Split Techniques for Viscous Reactive Blunt Body Flow Computations," AIAA Paper No. 87-0282, Jan. 1987.

[199] M.C. Dunn and S.W. Kang, "Theoretical and Experimental Studies of Reentry Plasmas," NASA CR-2232, Apr. 1973.

[200] P.A. Gnoffo and F.A. Greene, "A Computational Study of the Flowfield Surrounding the Aeroassist Flight Experiment Vehicle," AIAA Paper No. 87-1575, 1987.

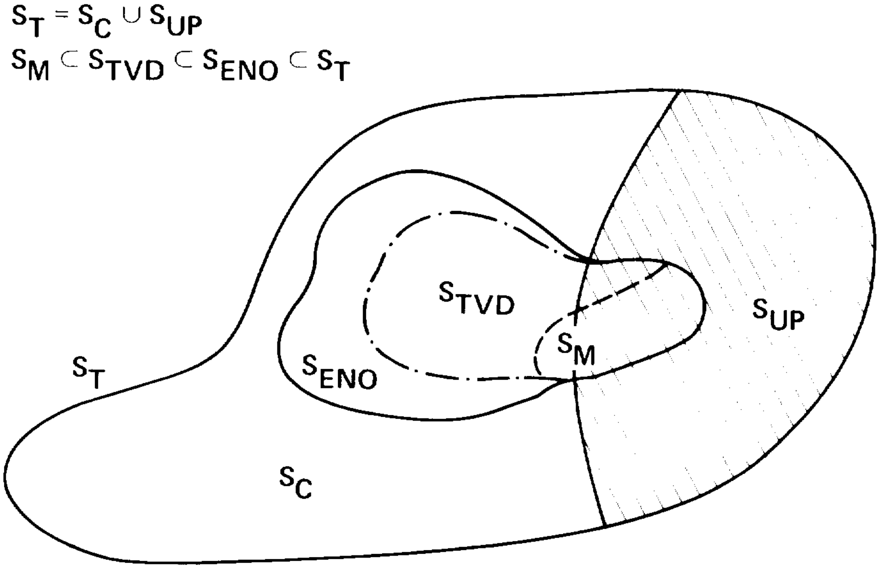


Fig. 1.1 Hierarchy of conservative schemes for hyperbolic conservation laws.

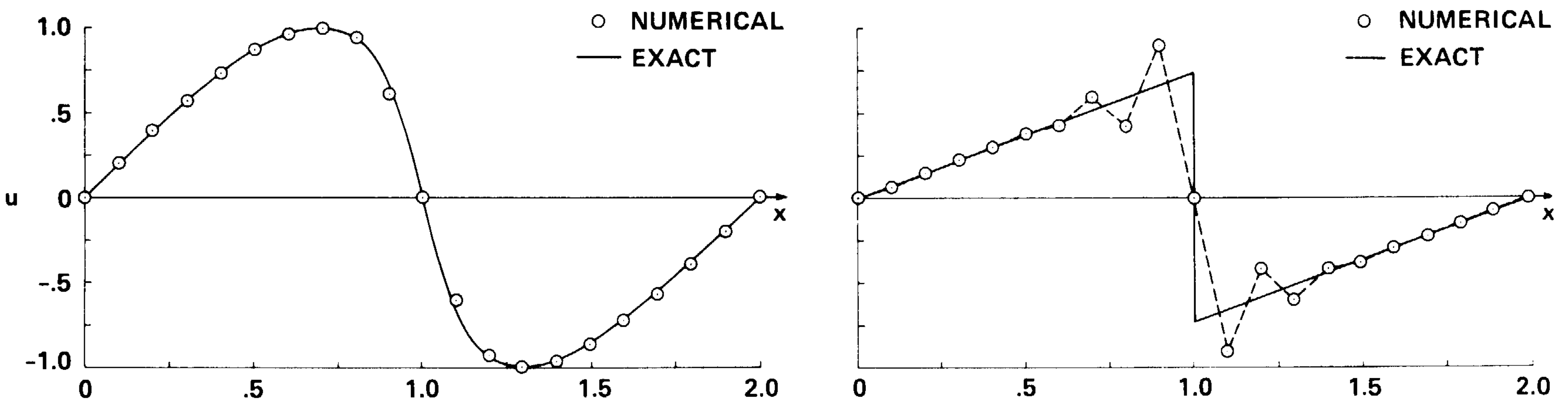


Fig. 3.1 Lax-Wendroff numerical solutions of Burgers' equation at t=0.2 (left) and t=1.0 (right).

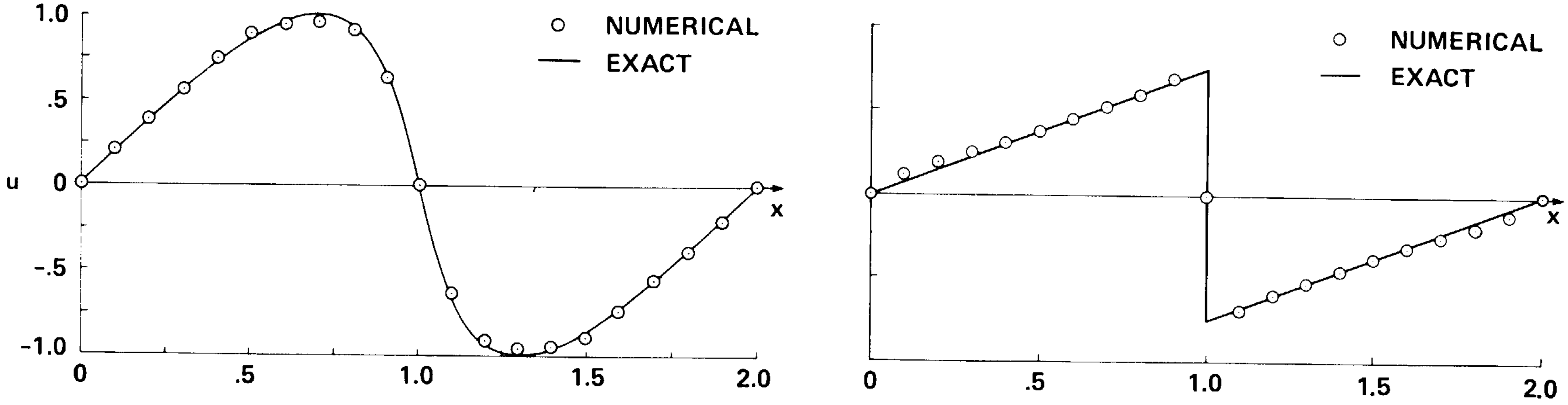
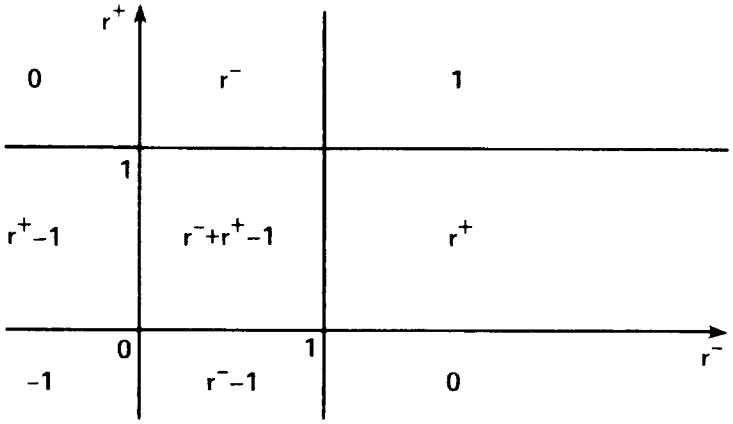
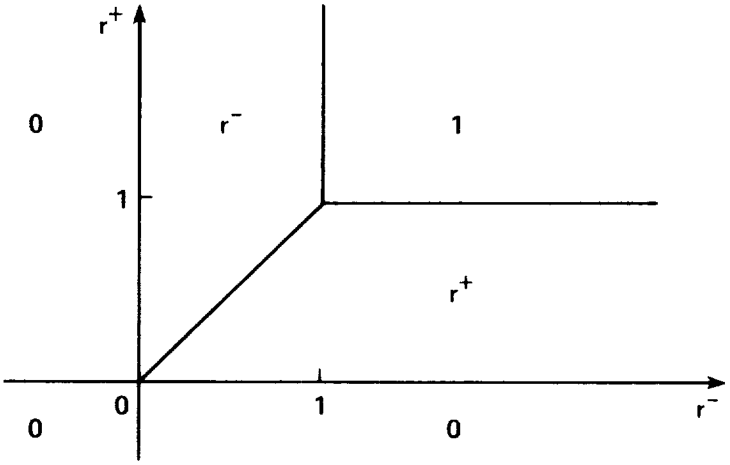


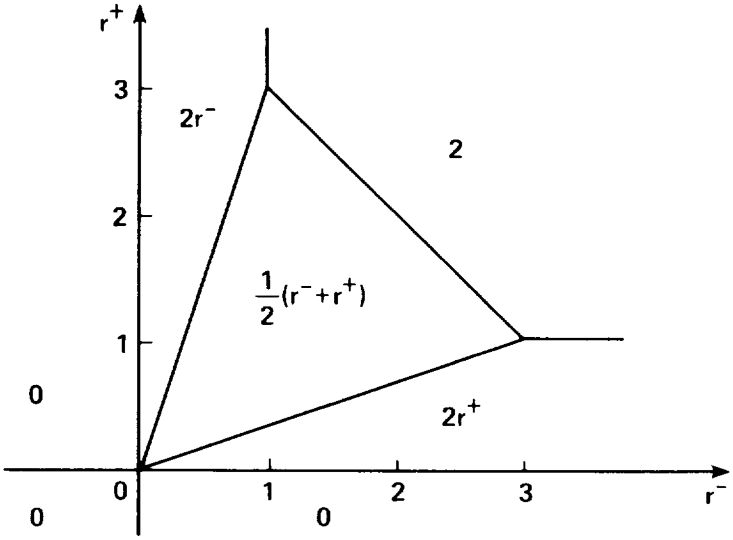
Fig. 3.2 Second-order TVD numerical solutions of Burgers' equation at t=0.2 (left) and t=1.0 (right).



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



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



(c) $Q(r^-, r^+) = \text{minmod}(2, 2r^-, 2r^+, \frac{1}{2}(r^- + r^+))$

Fig. 3.3. Graphical representation of three limiters (3.54,c,d,e).

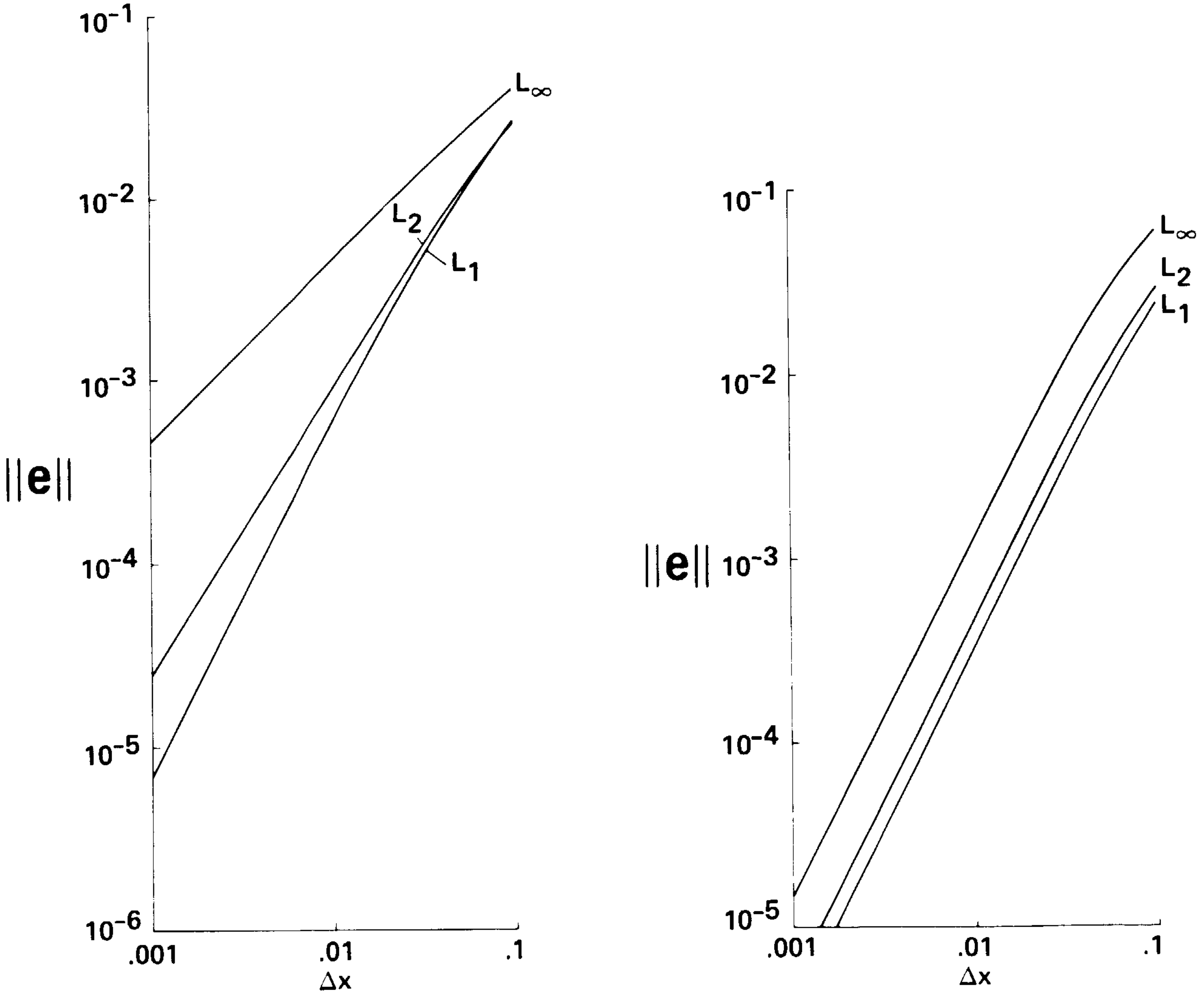


Fig. 3.4 Global order of accuracy at t=0.2 when the solution is still smooth.

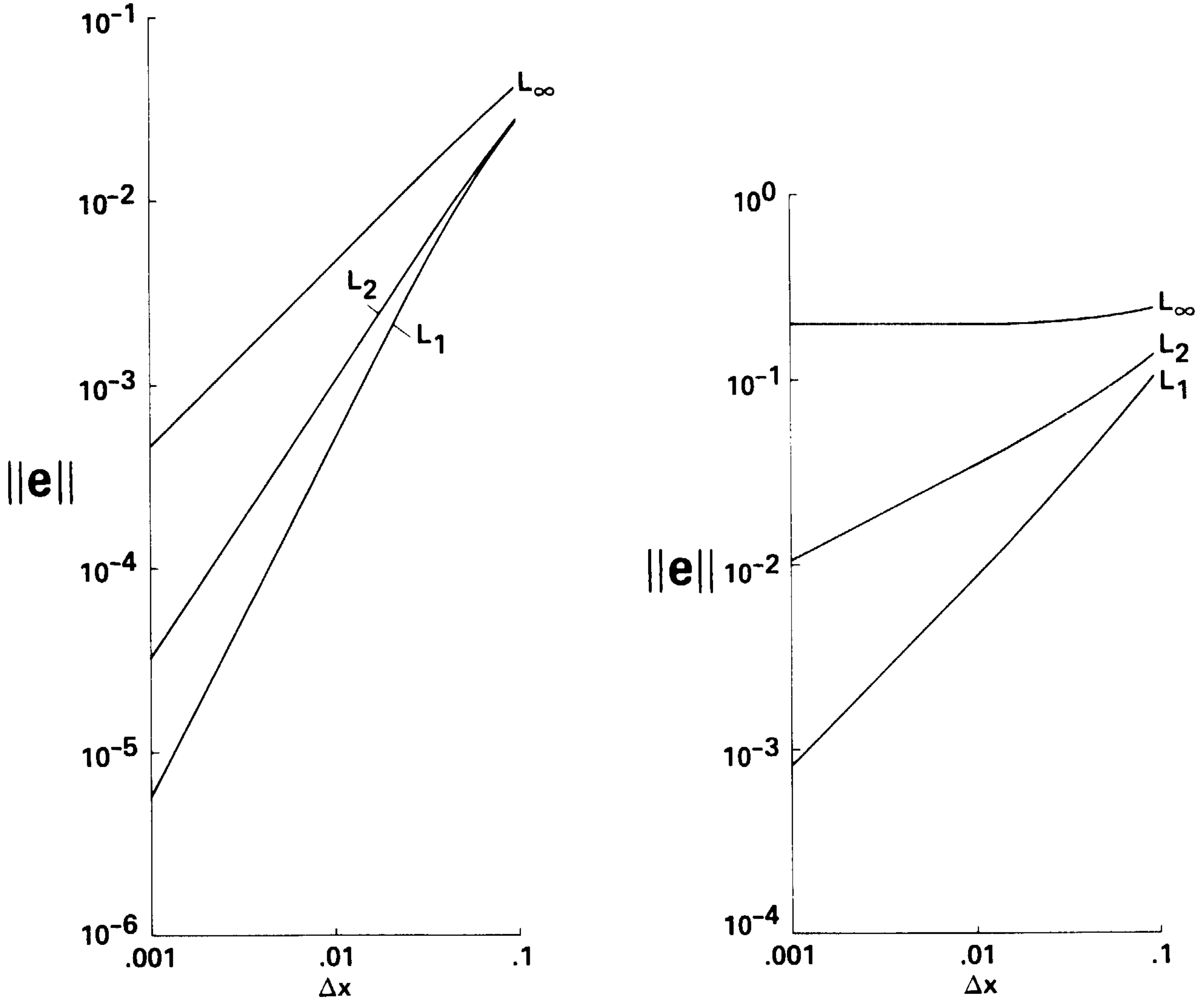


Fig. 3.5 Global order of accuracy at t = 1.0 when a shock solution has developed.

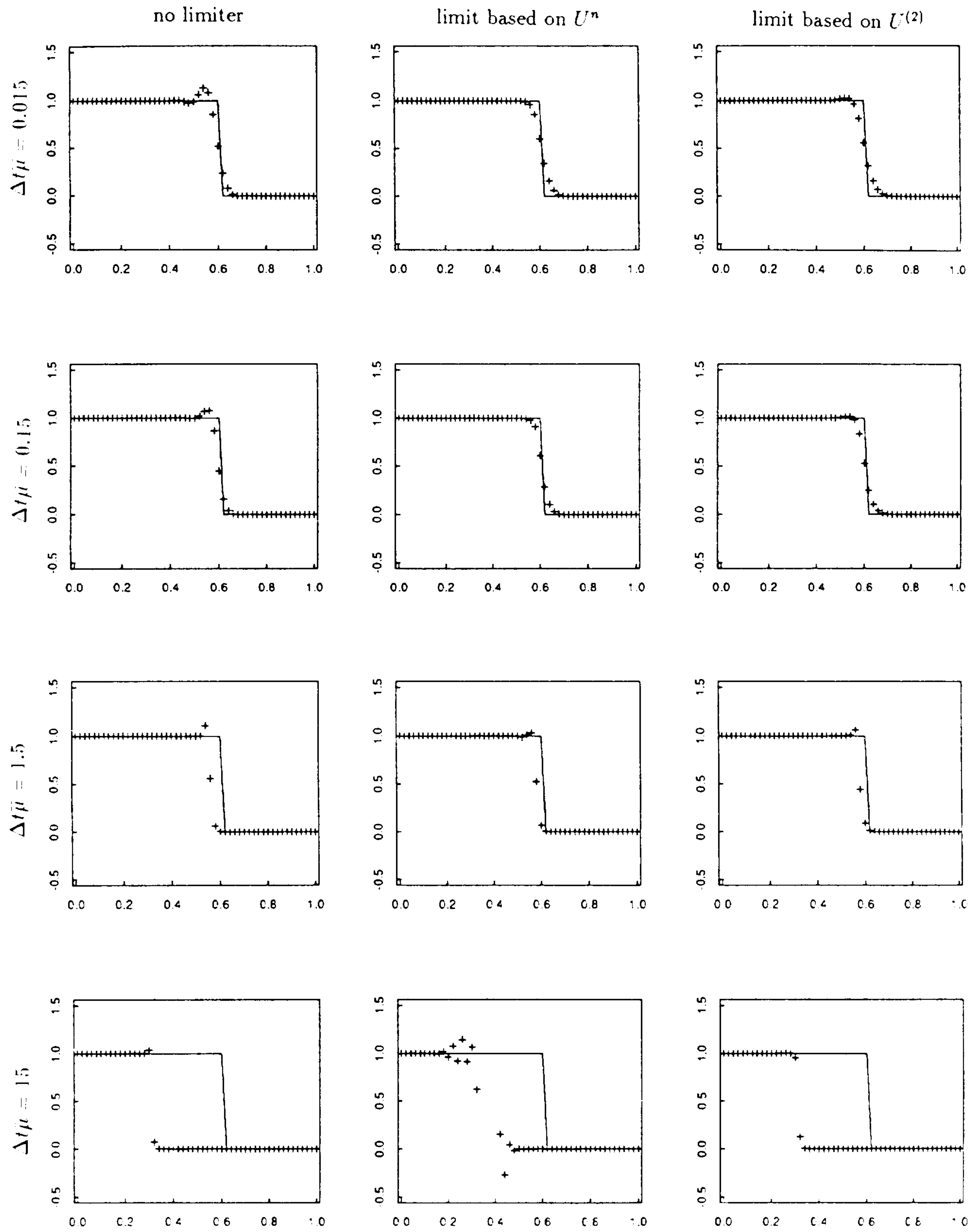


Fig. 3.6 Numerical results using extended MacCormack method with discontinuous initial data. — : true solution. + : computed solution.

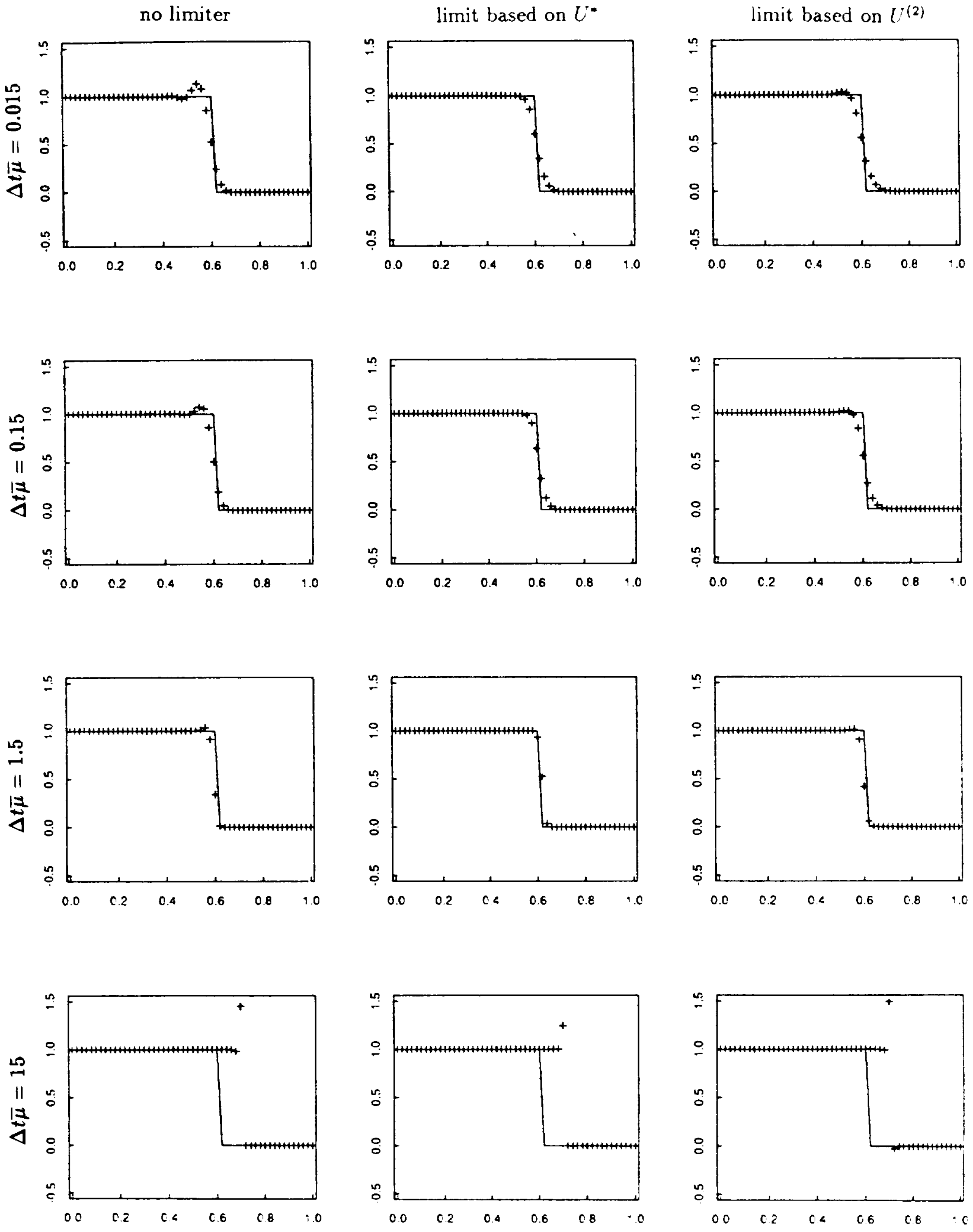
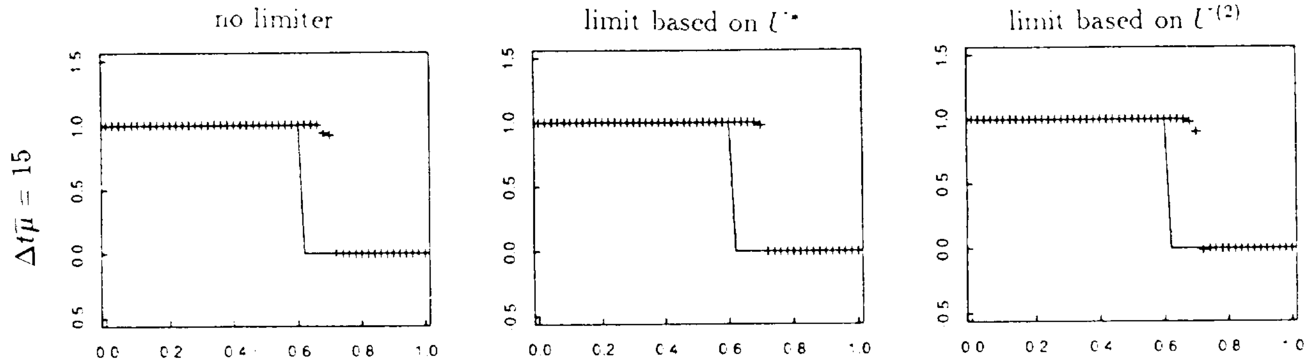


Fig 3.7 Numerical results using splitting method with discontinuous initial data. — : true solution, + : computed solution.



no limiter

limit based on l^{*}

limit based on $l^{*(2)}$

Fig. 3.8 Results for $\Delta t\bar{\mu} = 15$ when the trapezoidal method is used. — : true solution, + : computed solution.

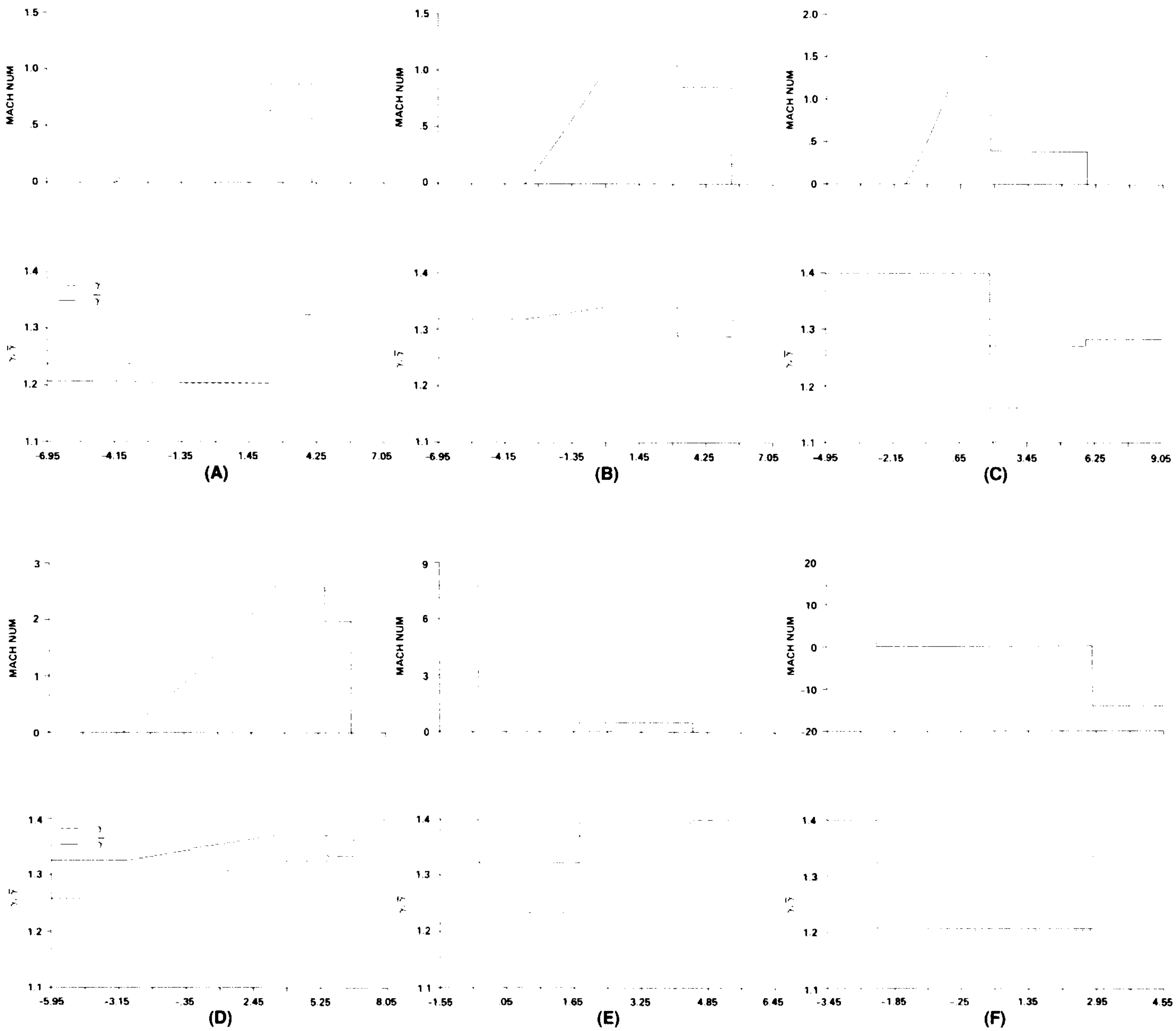


Fig. 4.1 "Exact solution" of the Mach number, γ, and $\bar{\gamma}$ for the six test cases at the same time instant as the numerical computation for equilibrium air.

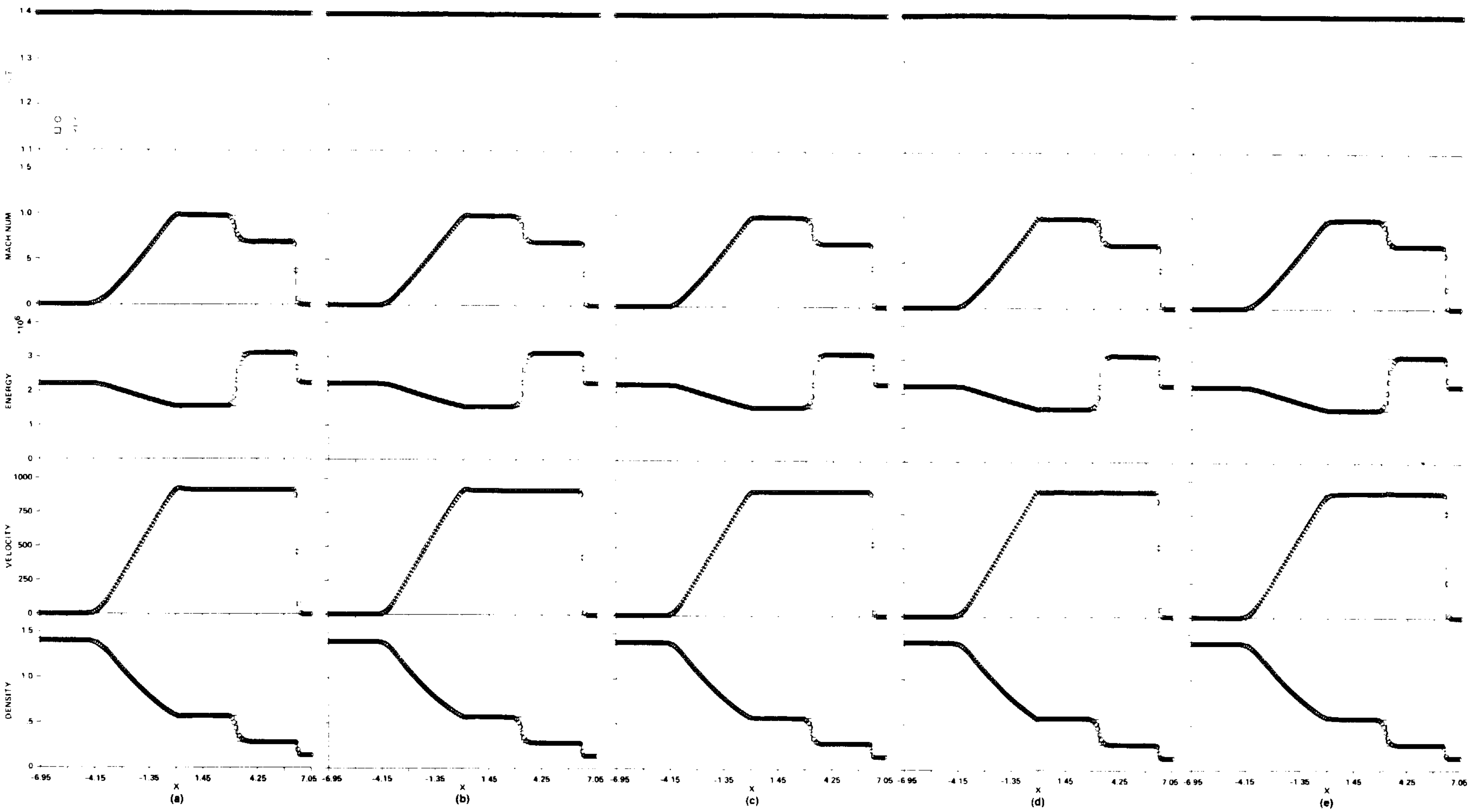


Fig. 4.2 Comparison of the schemes for test case B, for a **perfect gas.**
(a) Symmetric TVD, non-MUSCL; (b) upwind TVD, non-MUSCL, (c) upwind TVD, MUSCL,
(d) van Leer splitting, MUSCL; (e) Steger and Warming splitting, MUSCL.
Note: (a)-(c) use the identical approximate Riemann solver.

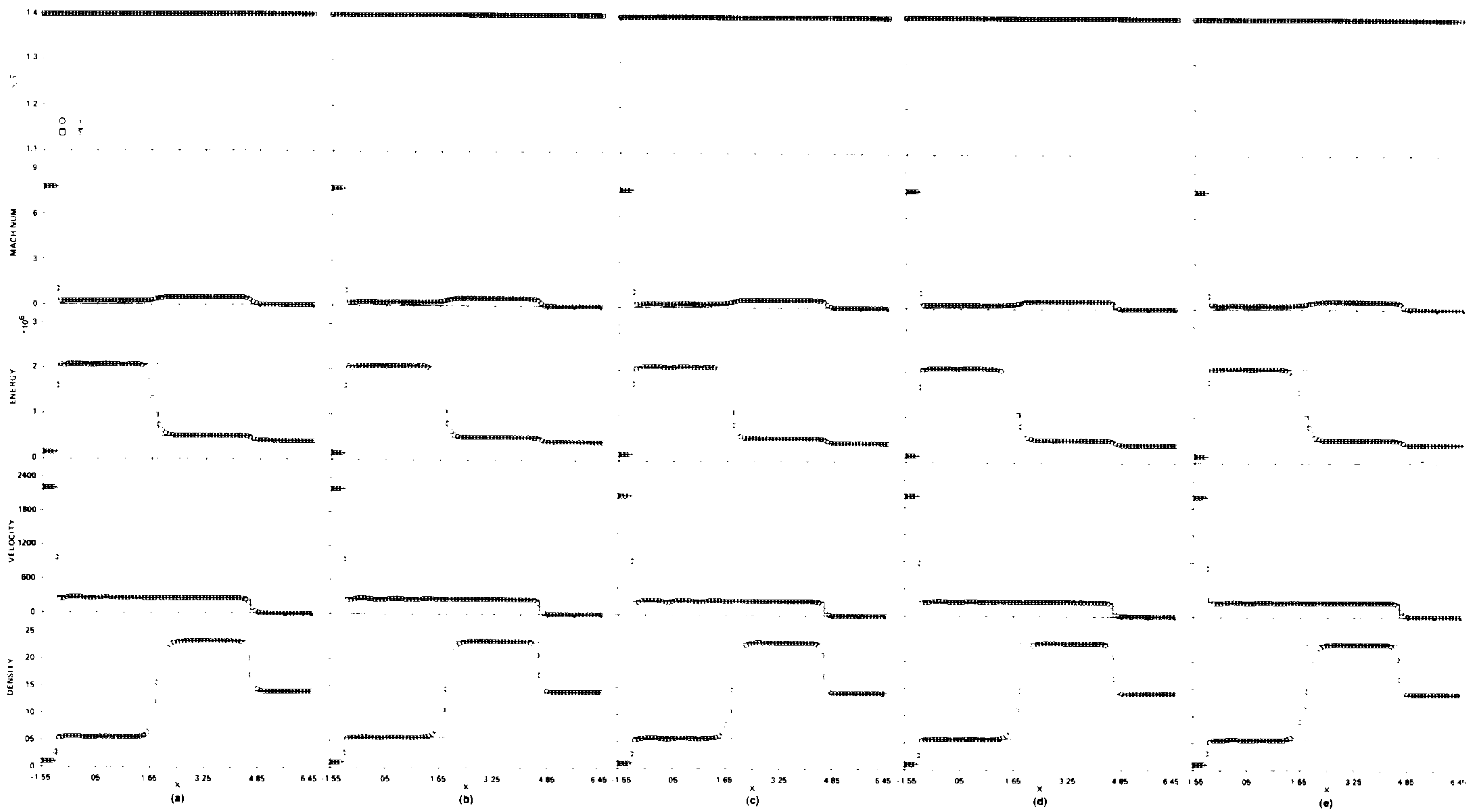


Fig. 4.3 Comparison of the schemes for test case E, for a **perfect gas.**

(a) Symmetric TVD, non-MUSCL; (b) upwind TVD, non-MUSCL; (c) upwind TVD, MUSCL;

(d) van Leer splitting, MUSCL; (e) Steger and Warming splitting, MUSCL.

Note: (a)-(c) use the identical approximate Riemann solver.

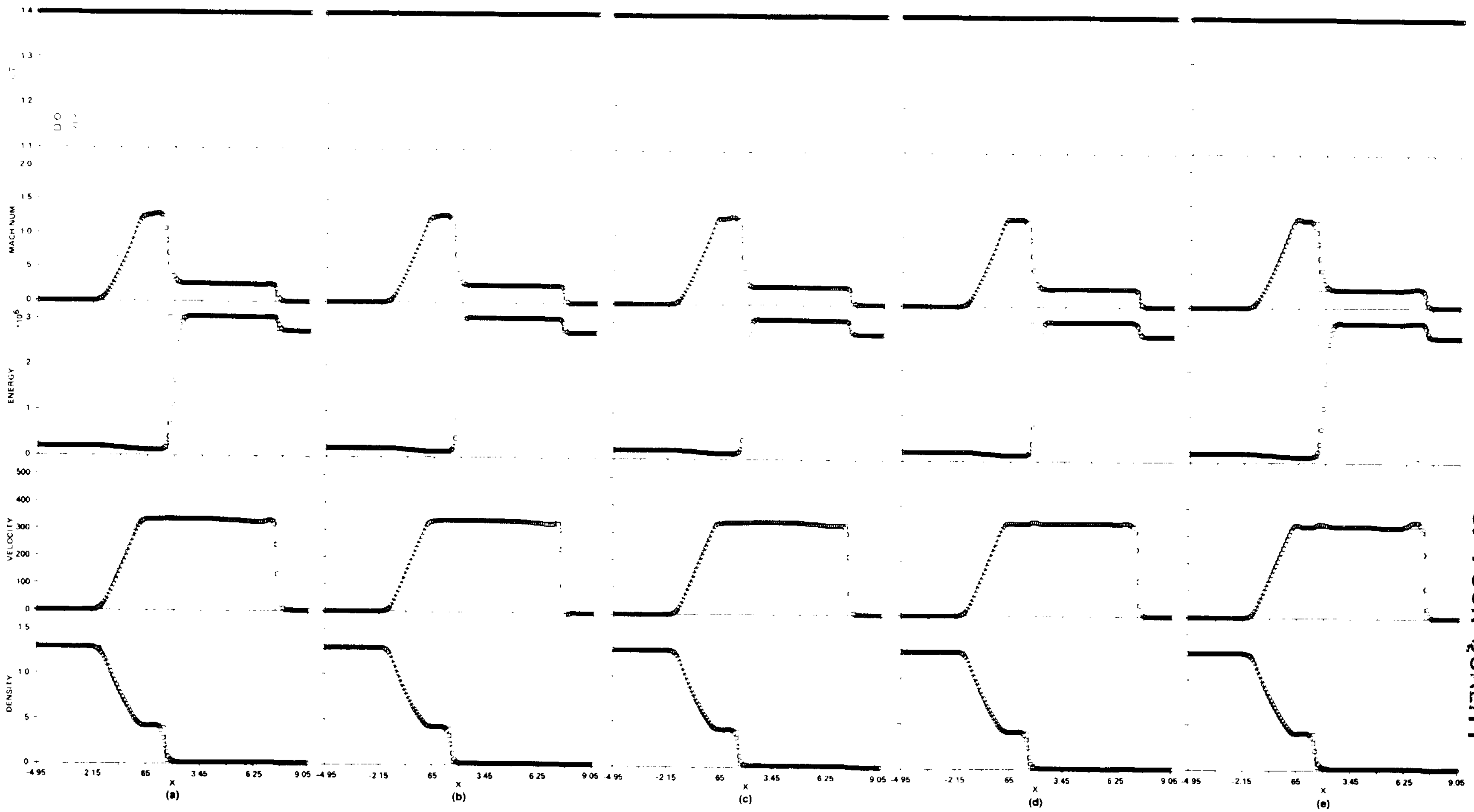


Fig. 4.4 Comparison of the schemes for test case C, for a **perfect gas.**
(a) Symmetric TVD, non-MUSCL; (b) upwind TVD, non-MUSCL; (c) upwind TVD, MUSCL;
(d) van Leer splitting, MUSCL; (e) Steger and Warming splitting, MUSCL.
Note: (a)-(c) use the identical approximate Riemann solver.

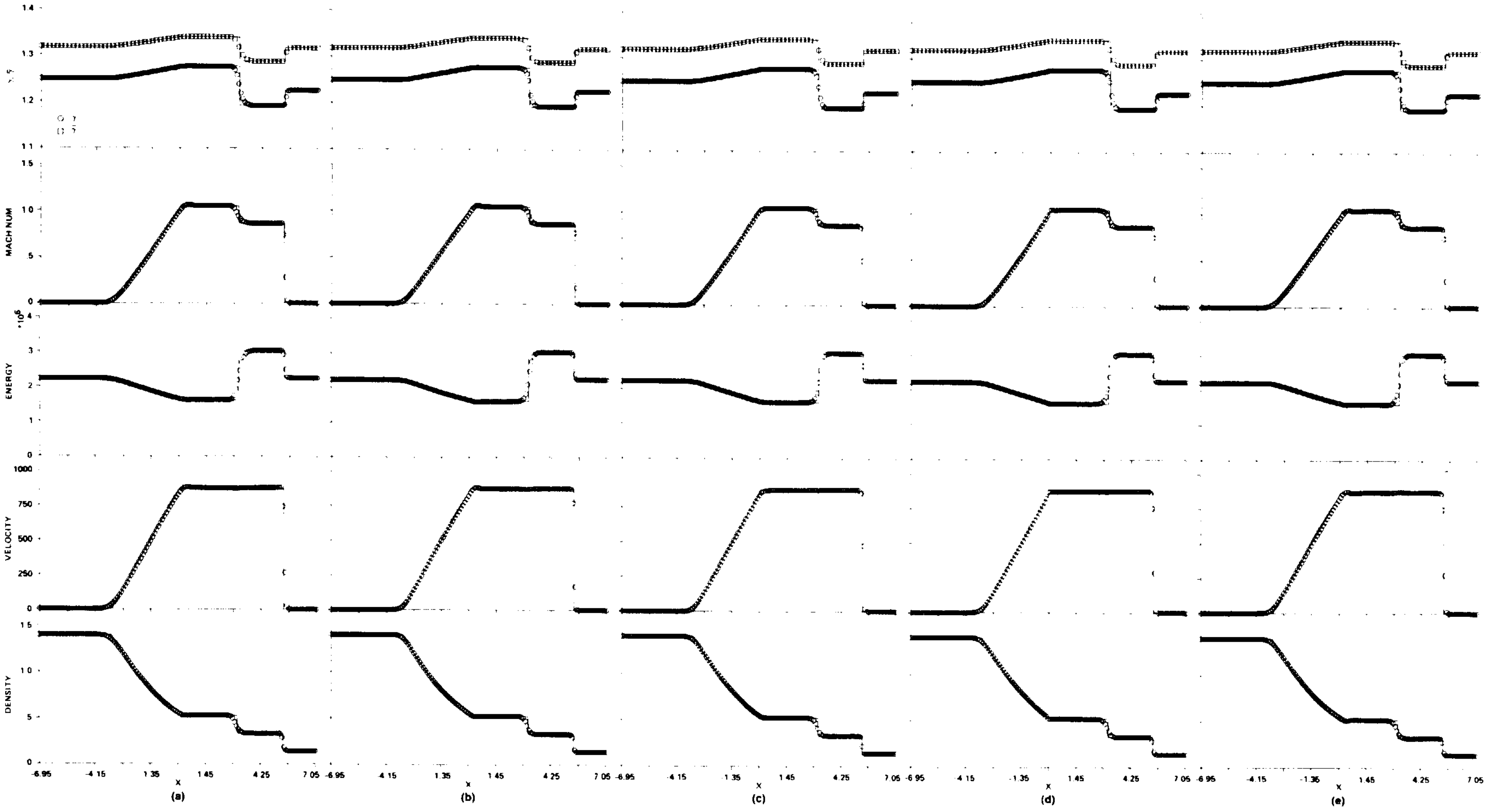


Fig. 4.5 Comparison of the schemes for test case B, for **equilibrium air.**
(a) Symmetric TVD, non-MUSCL; (b) upwind TVD, non-MUSCL; (c) upwind TVD, MUSCL;
(d) generalized van Leer splitting, MUSCL; (e) generalized Steger and Warming splitting, MUSCL.
Note: (a)-(c) use the identical approximate Riemann solver.

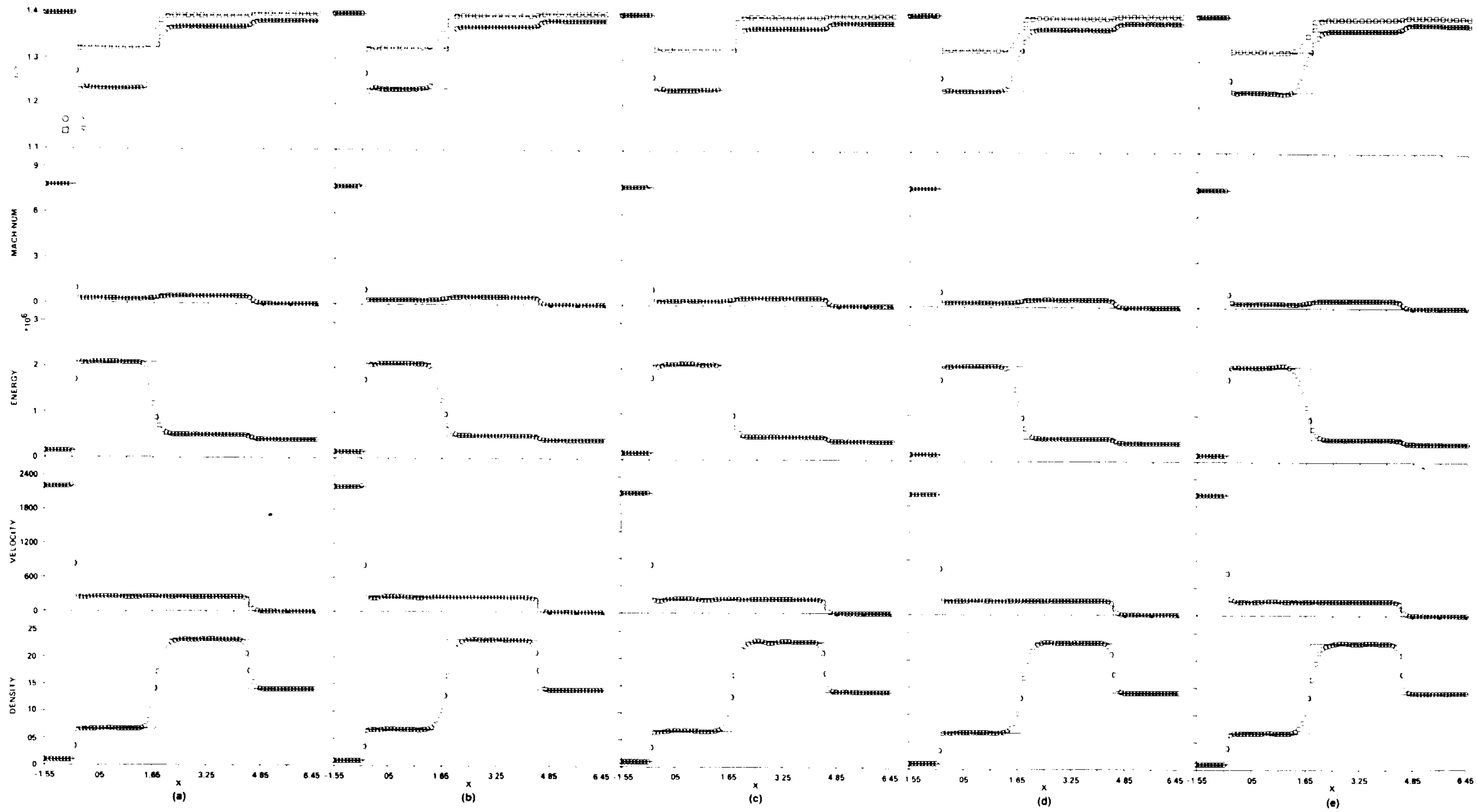


Fig. 4.6 Comparison of the schemes for test case E, for **equilibrium air.**
(a) Symmetric TVD, non-MUSCL; (b) upwind TVD, non-MUSCL; (c) upwind TVD, MUSCL;
(d) generalized van Leer splitting, MUSCL; (e) generalized Steger and Warming splitting, MUSCL.
Note: (a)-(c) use the identical approximate Riemann solver.

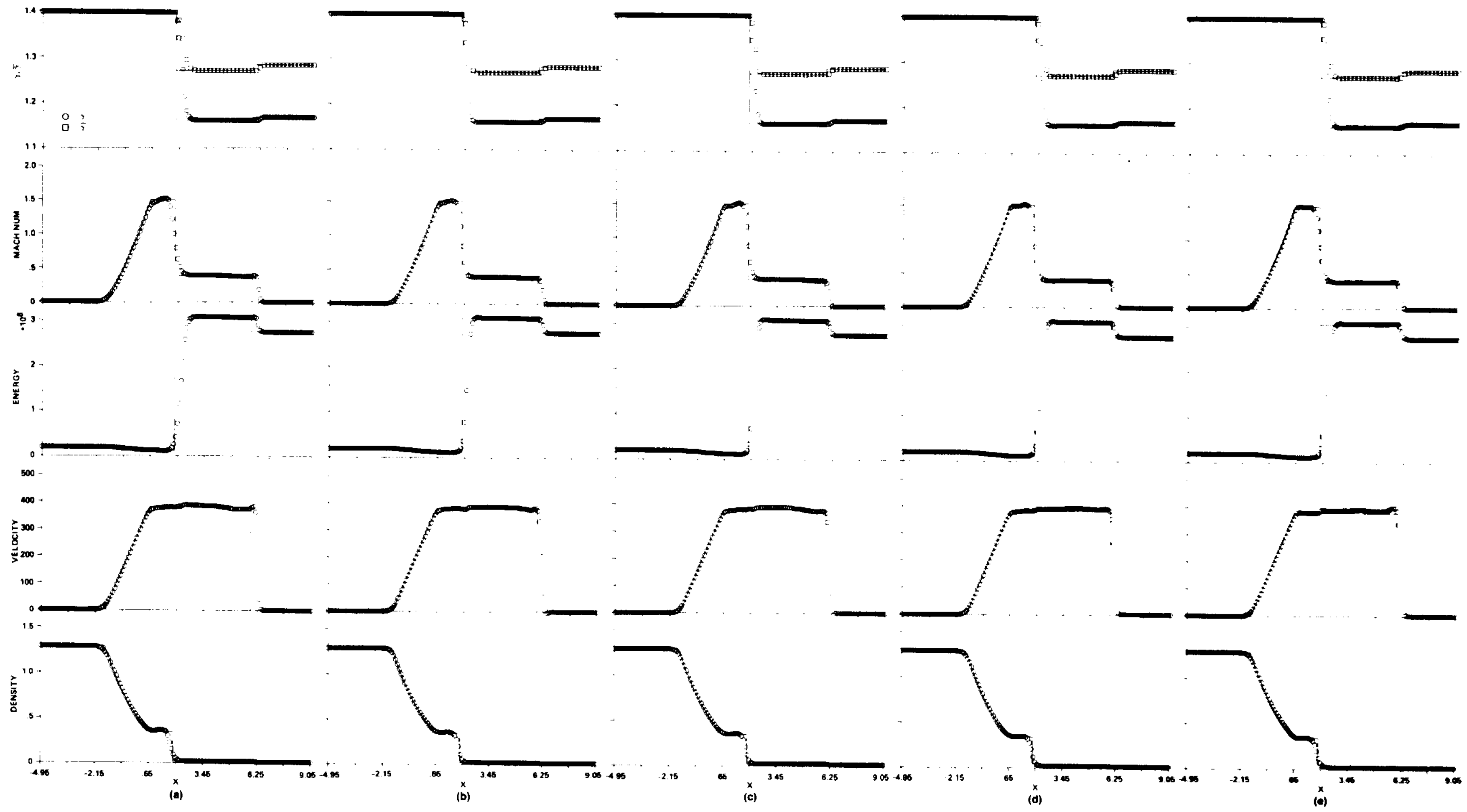


Fig. 4.7 Comparison of the schemes for test case C, for **equilibrium air.**
(a) Symmetric TVD, non-MUSCL; (b) upwind TVD, non-MUSCL; (c) upwind TVD, MUSCL;
(d) generalized van Leer splitting, MUSCL; (e) generalized Steger and Warming splitting, MUSCL.
Note: (a)-(c) use the identical approximate Riemann solver.

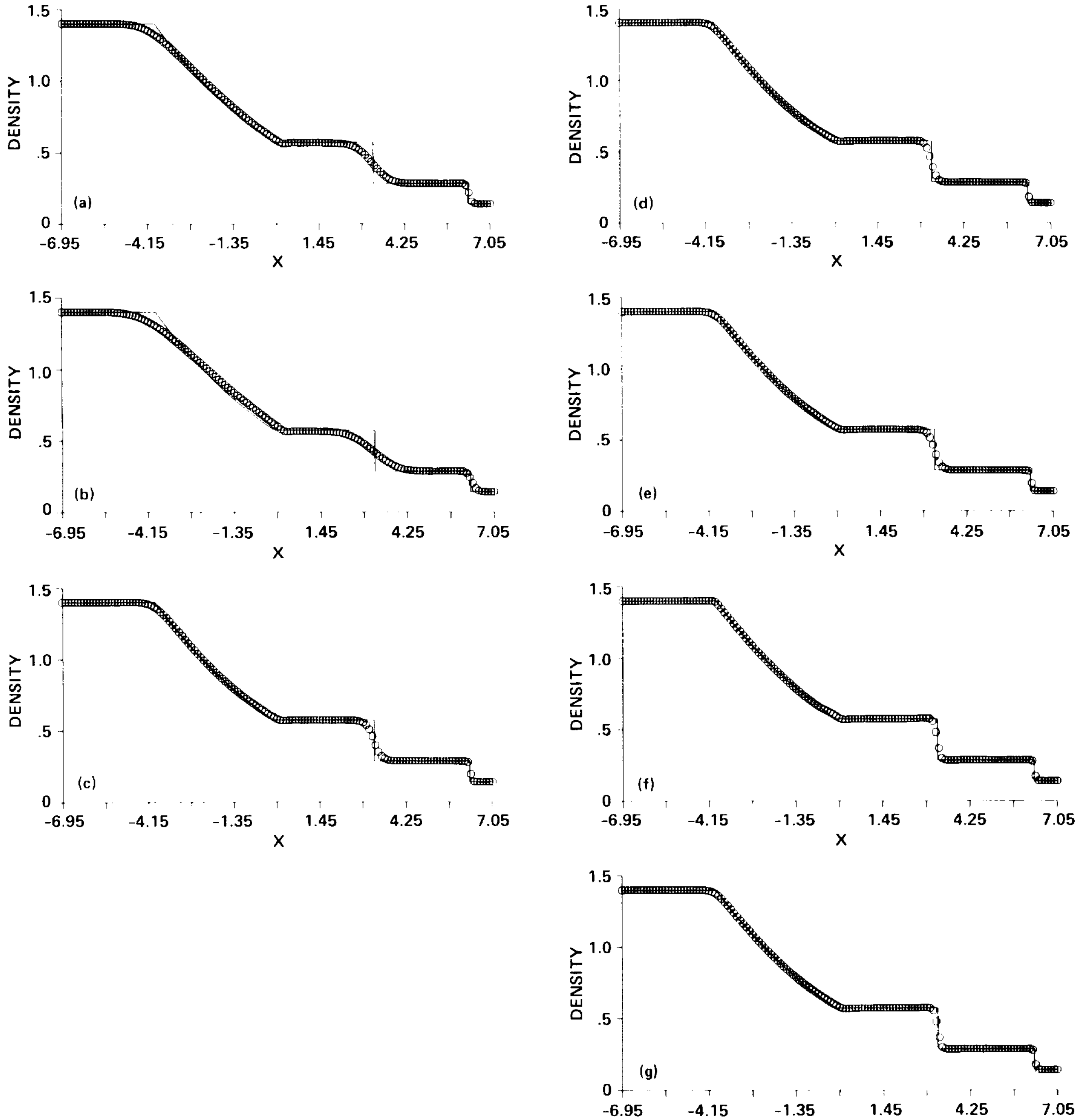


Fig. 4.8 Comparison of the schemes and limiters for test case B for a **perfect gas.**
(a) First-order Roe; (b) second-order Harten (3.45), non-MUSCL;
(c)-(g) upwind TVD (4.34), non-MUSCL, limiters (4.34c,d,e,g);
(h)-(j) symmetric TVD (4.33), non-MUSCL, limiters (4.33c,d,e);
(k) upwind TVD, MUSCL; (l) van Leer splitting, MUSCL;
(m) Steger and Warming splitting, MUSCL.
Note: (a)-(k) use Roe's average, and (k)-(m) use the same limiter (4.34e).

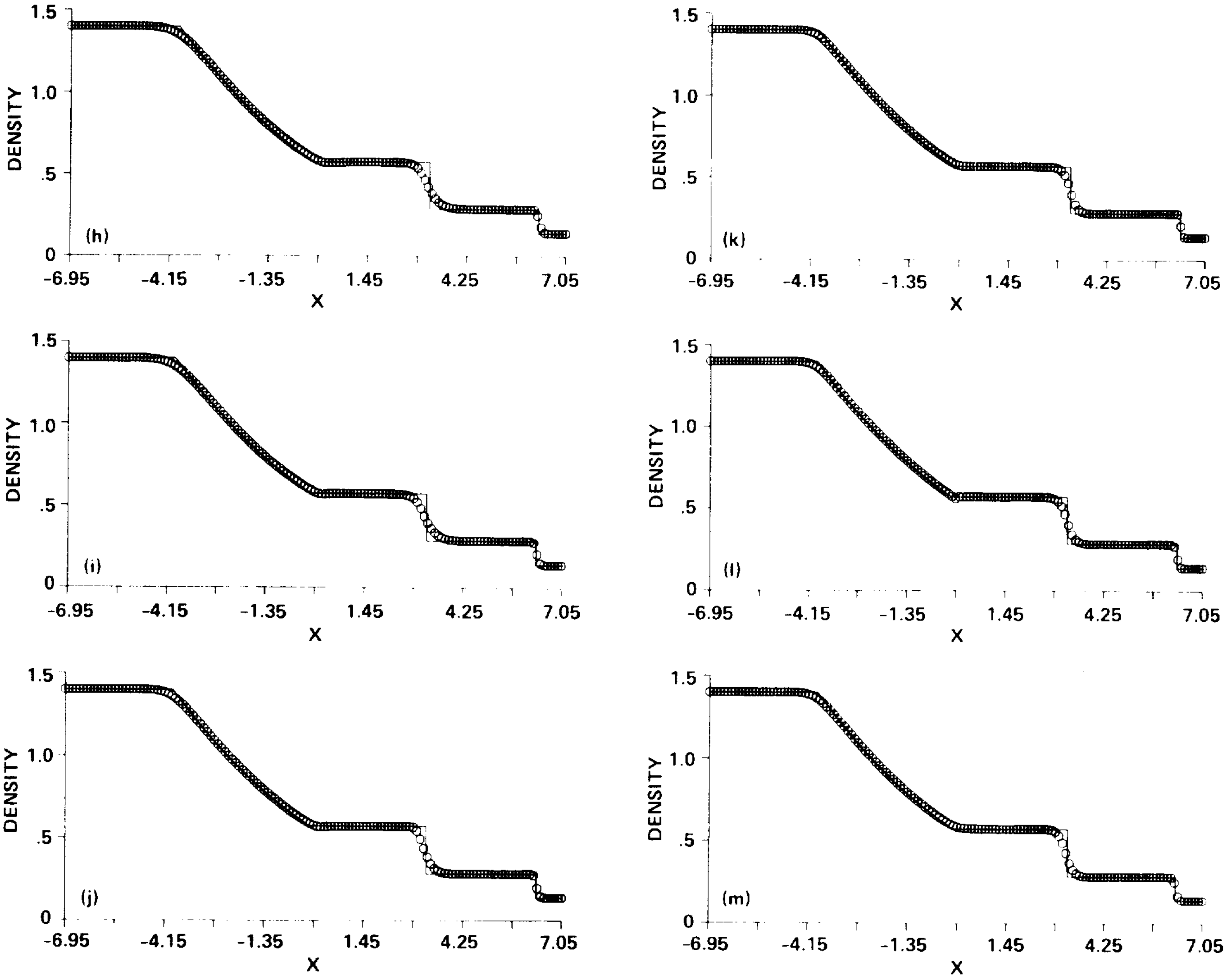


Fig. 4.8 (Concluded)

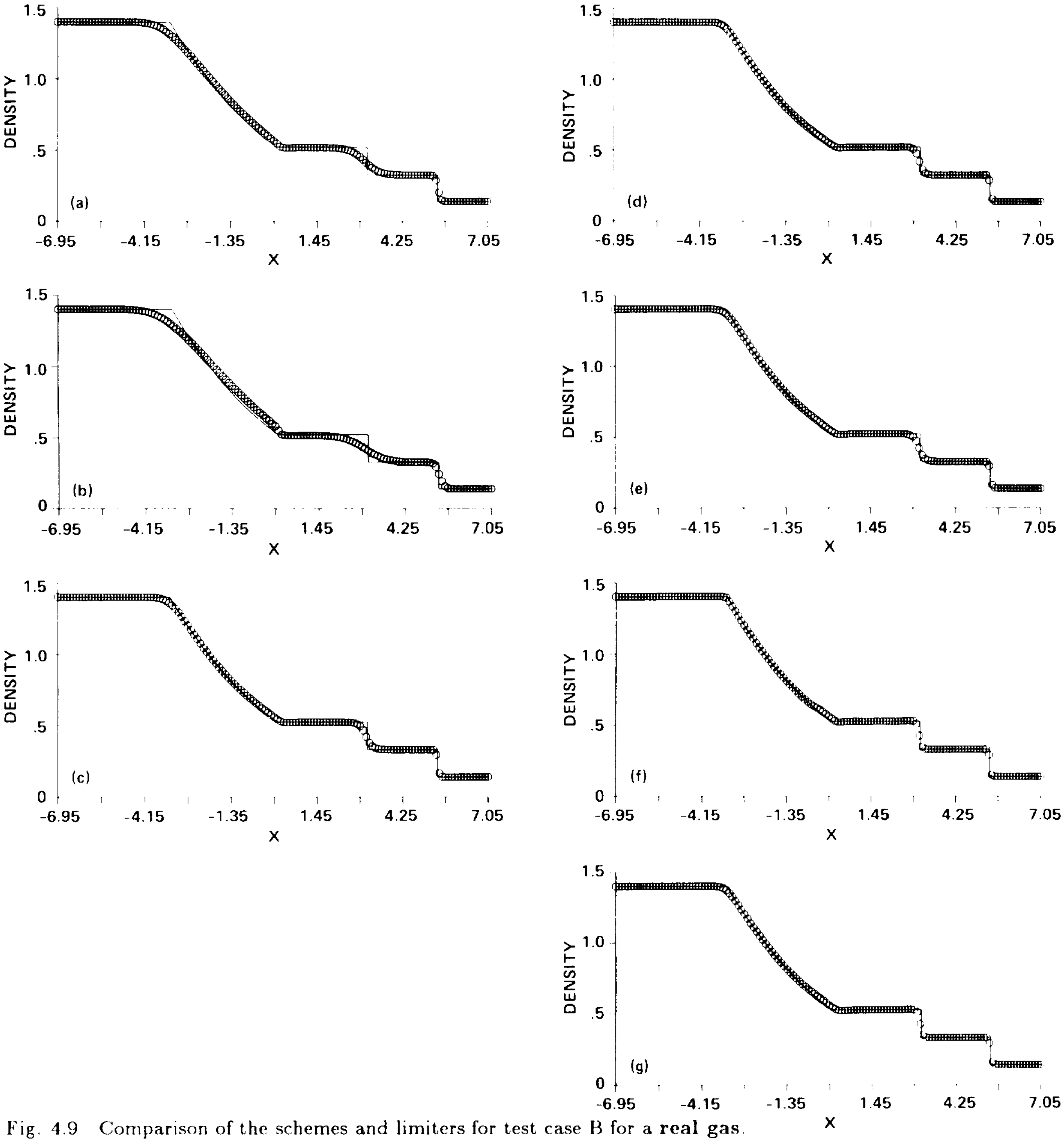


Fig. 4.9 Comparison of the schemes and limiters for test case B for a **real gas**.
(a) First-order Roe; (b) second-order Harten (3.45), non-MUSCL;
(c)-(g) upwind TVD (4.34), non-MUSCL, limiters (4.34c,d,e,g);
(h)-(j) symmetric TVD (4.33), non-MUSCL, limiters (4.33c,d,e);
(k) upwind TVD, MUSCL; (l) generalized van Leer splitting, MUSCL;
(m) generalized Steger and Warming splitting, MUSCL.
Note: (a)-(k) use the identical approximate Riemann solver, and (k)-(m) use the same limiter (4.34e).

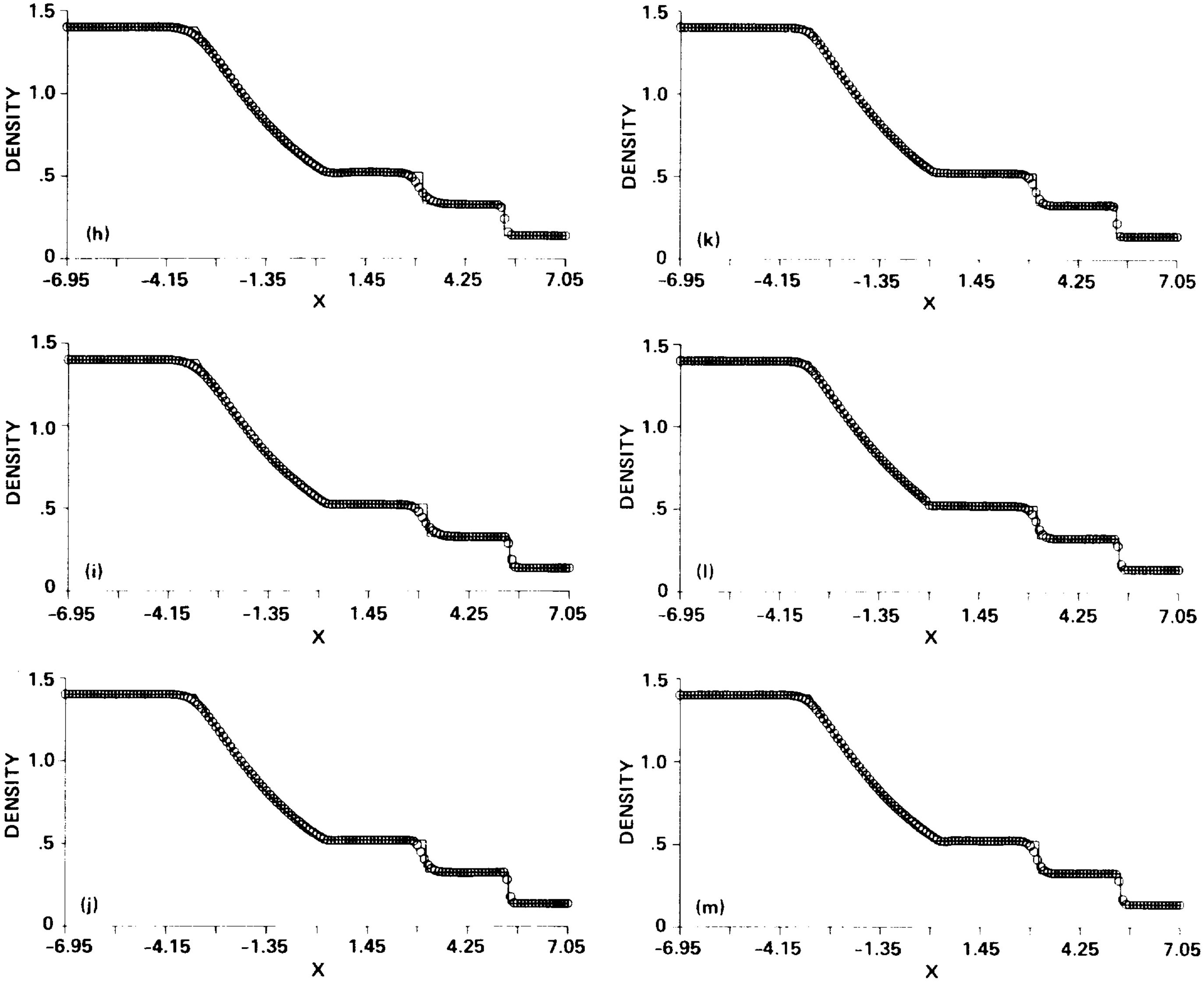
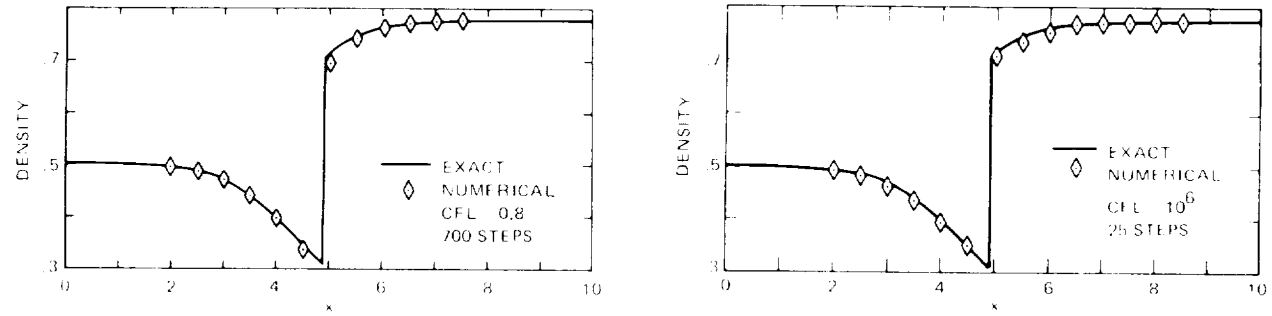
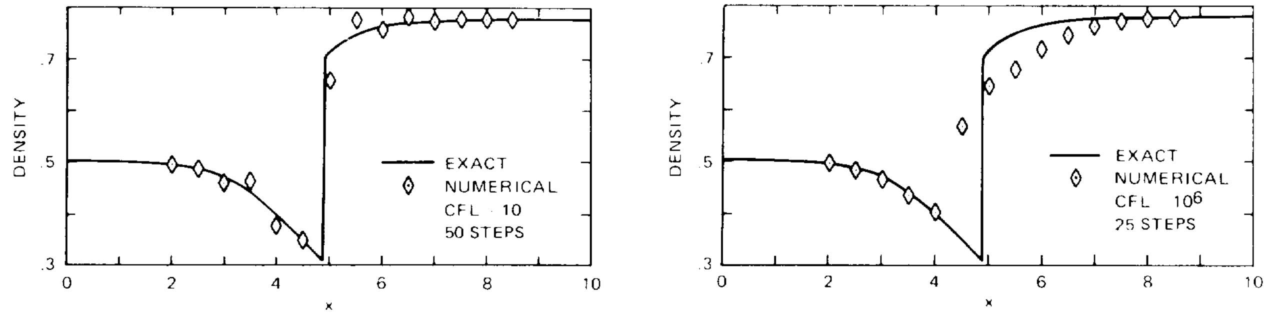


Fig. 4.9 (Concluded)



(a) Explicit *TVD* method.

(b) Implicit *TVD* method.



(c) Conventional implicit method

(d) First-order implicit flux-vector splitting method.

Fig. 4.10 Density distribution of a quasi-one-dimensional divergent nozzle for four different methods.

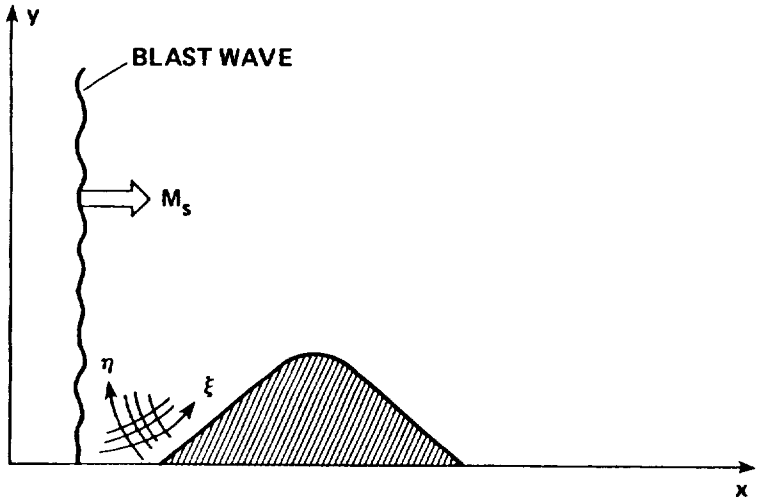


Fig. 5.1 Schematic of a shock wave diffraction problem.

ORIGINAL PAGE IS
OF POOR QUALITY

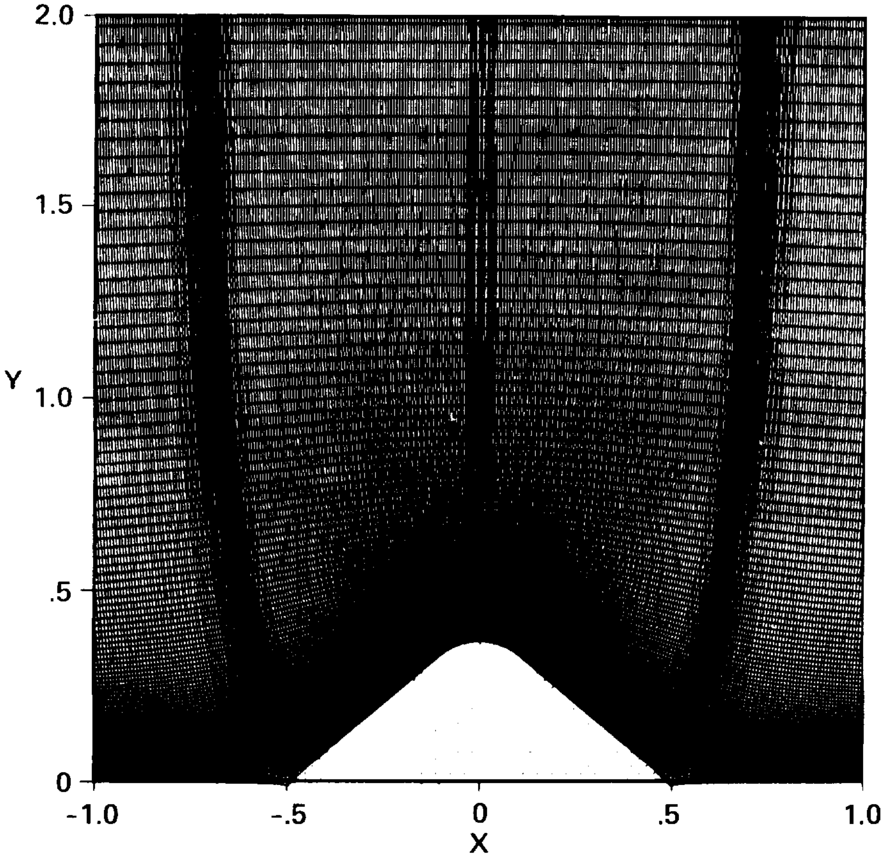


Fig. 5.2 The 317 × 100 orthogonal grid.

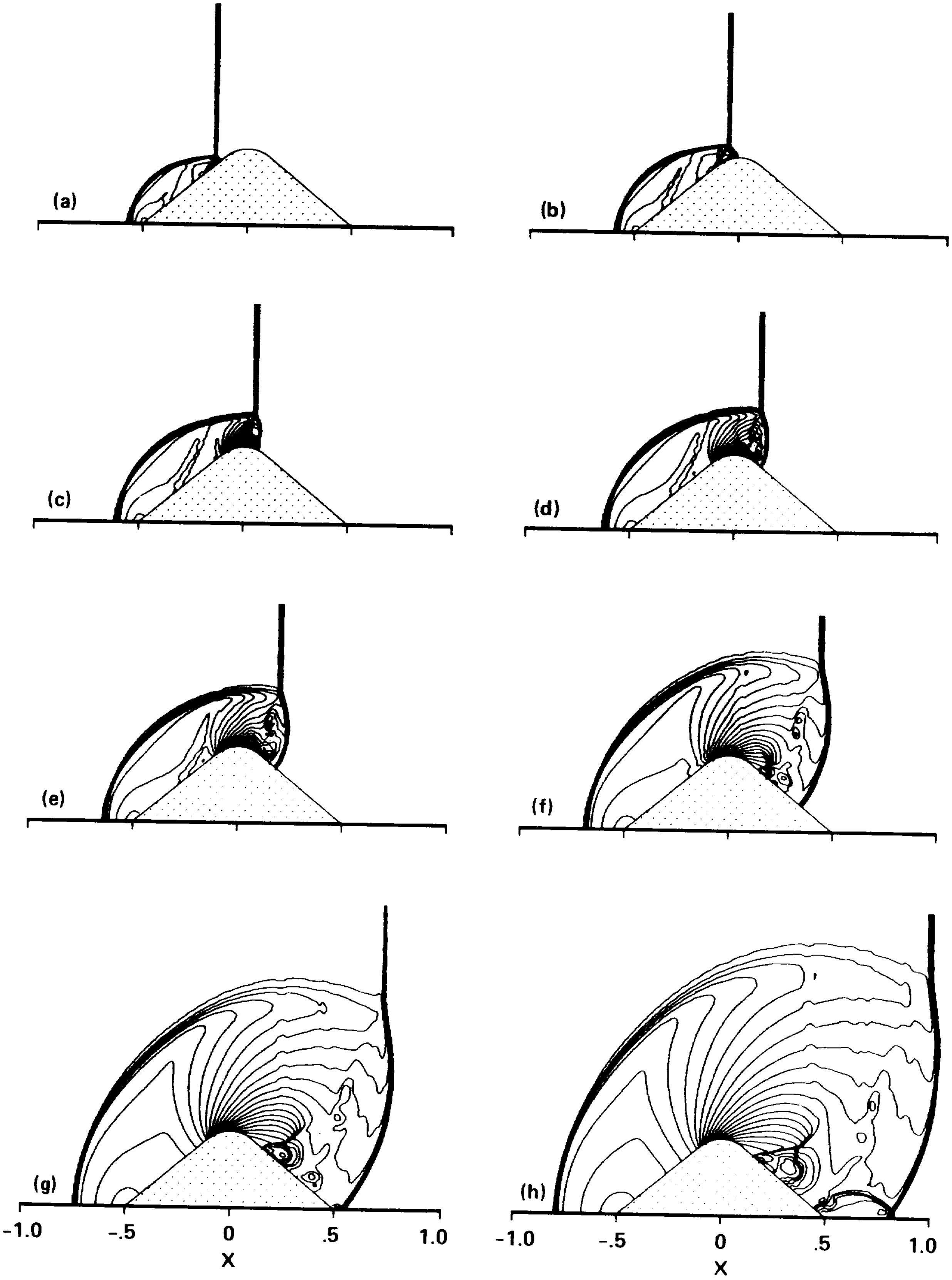


Fig. 5.3 Density contours computed by a time-split, non-MUSCL second-order upwind TVD scheme.

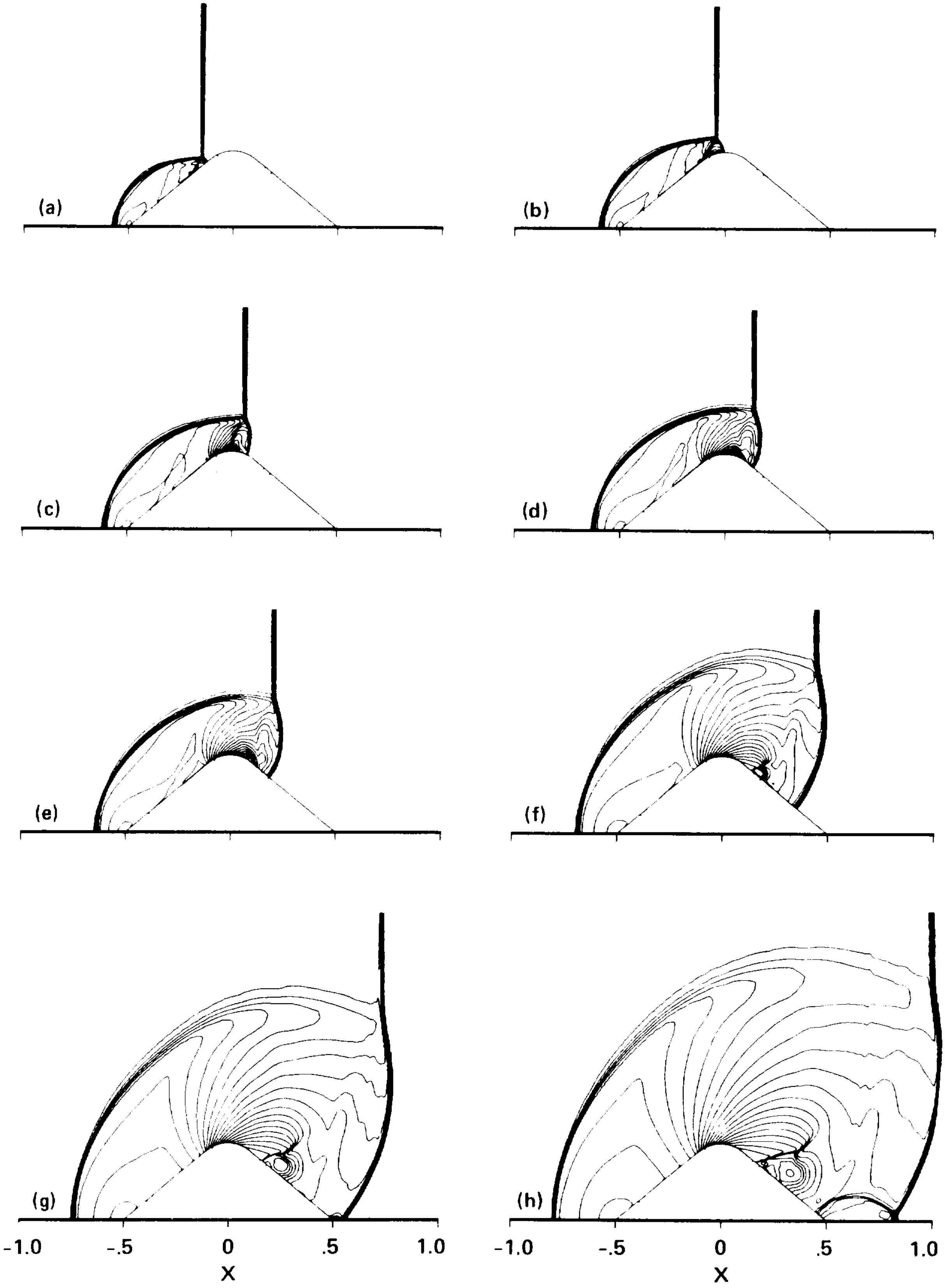


Fig. 5.4 Density contours computed by a predictor-corrector symmetric TVD scheme.

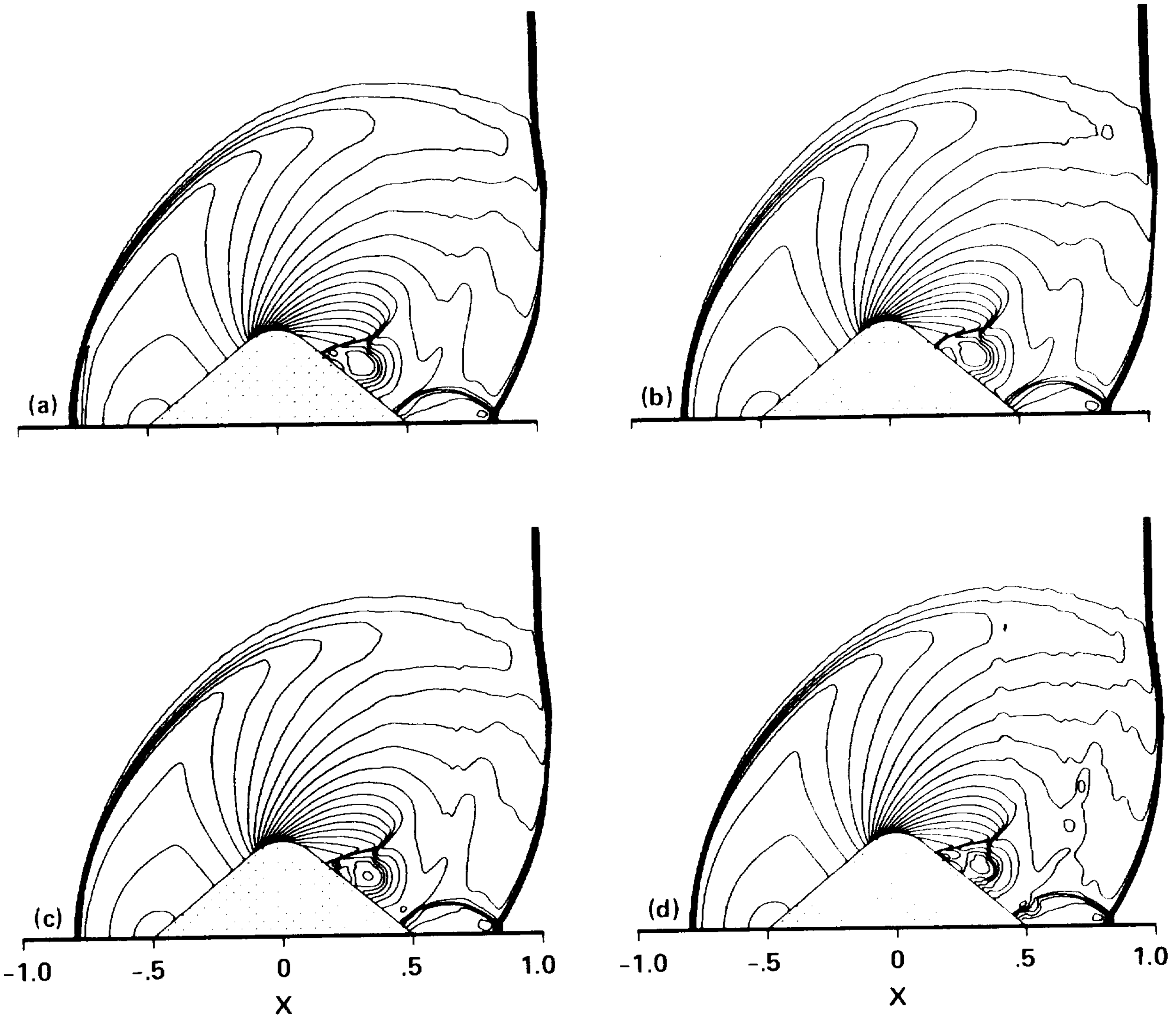


Fig. 5.5 Comparison of density contour plots from different TVD schemes.
(a) Predictor-corrector symmetric TVD with limiter (4.33c);
(b) predictor-corrector symmetric TVD with limiter (4.33d);
(c) predictor-corrector symmetric TVD with limiter (4.33e);
(d) time-split second-order upwind TVD, non-MUSCL with limiter (4.34d,g)

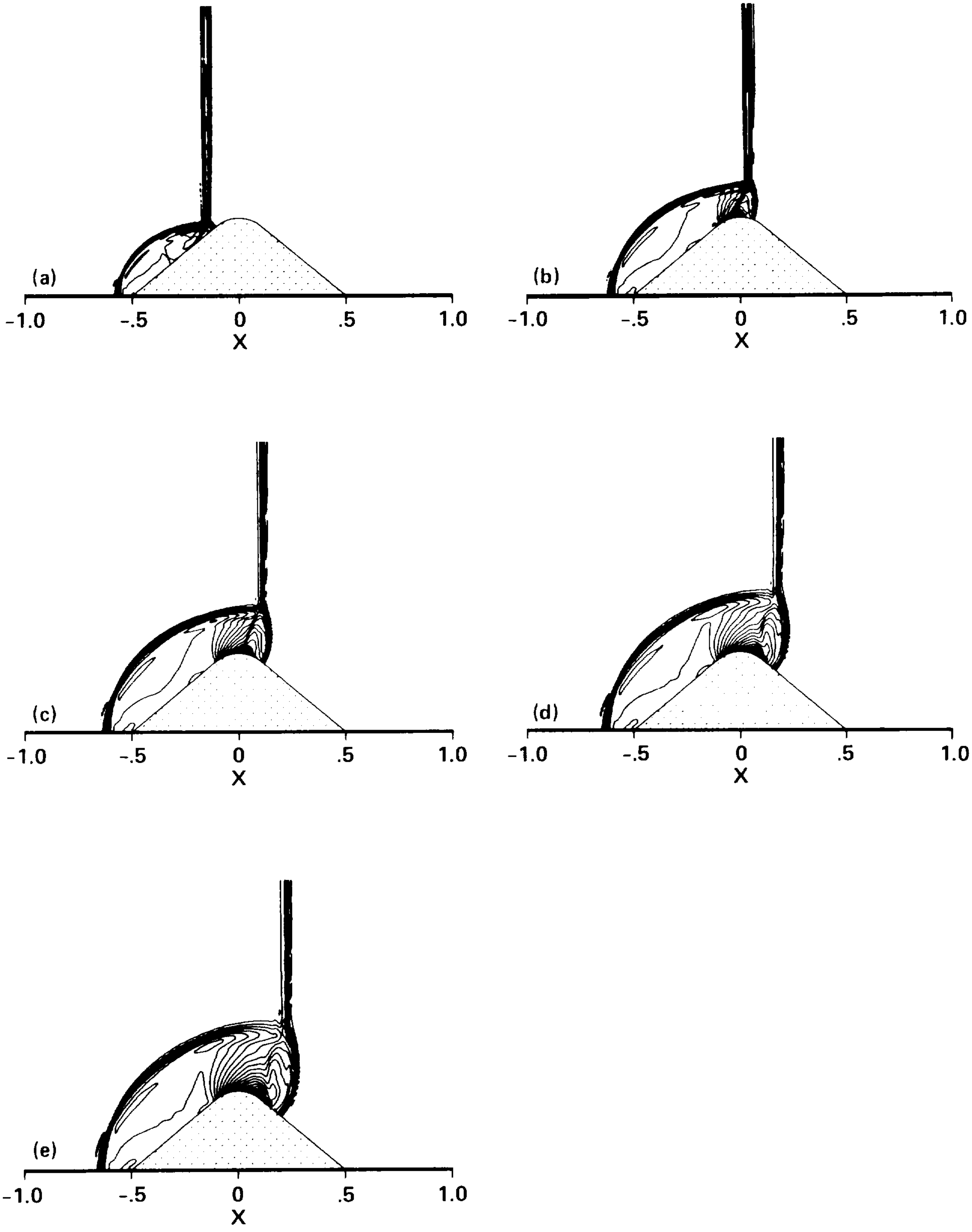
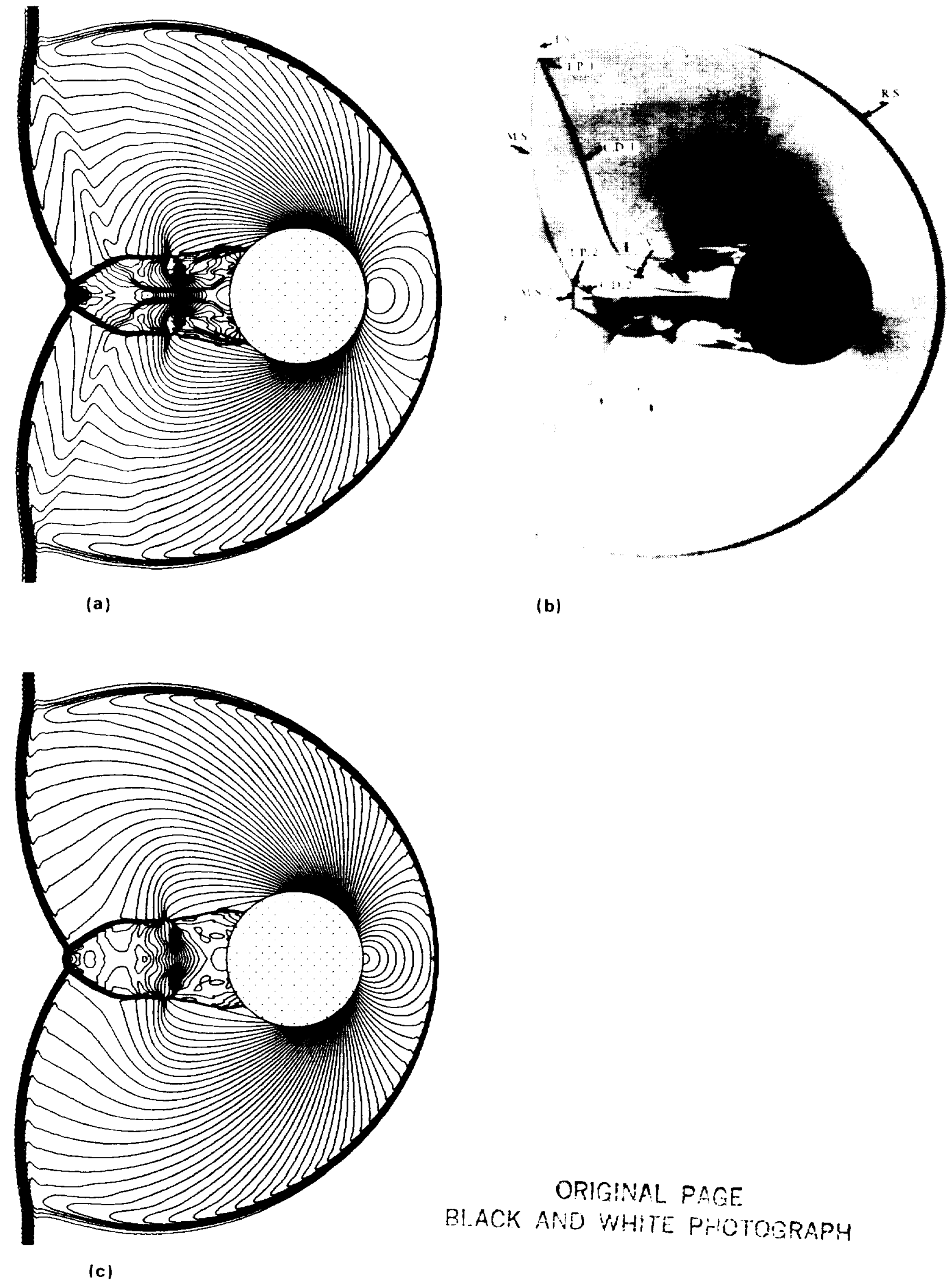


Fig. 5.6 Density contours computed by the explicit MacCormack method using the grid in Fig. 5.2.



ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Fig. 5.7 Density contours (a) and pressure contours (c) computed by a predictor-corrector symmetric TVD scheme compared with the Schlieren photograph (b).
Notes: T.P. = triple point; M.S. = Mach stem, V = vortex, C.D. = contact discontinuity; R.S. = reflected shock, I.S. = incident shock.

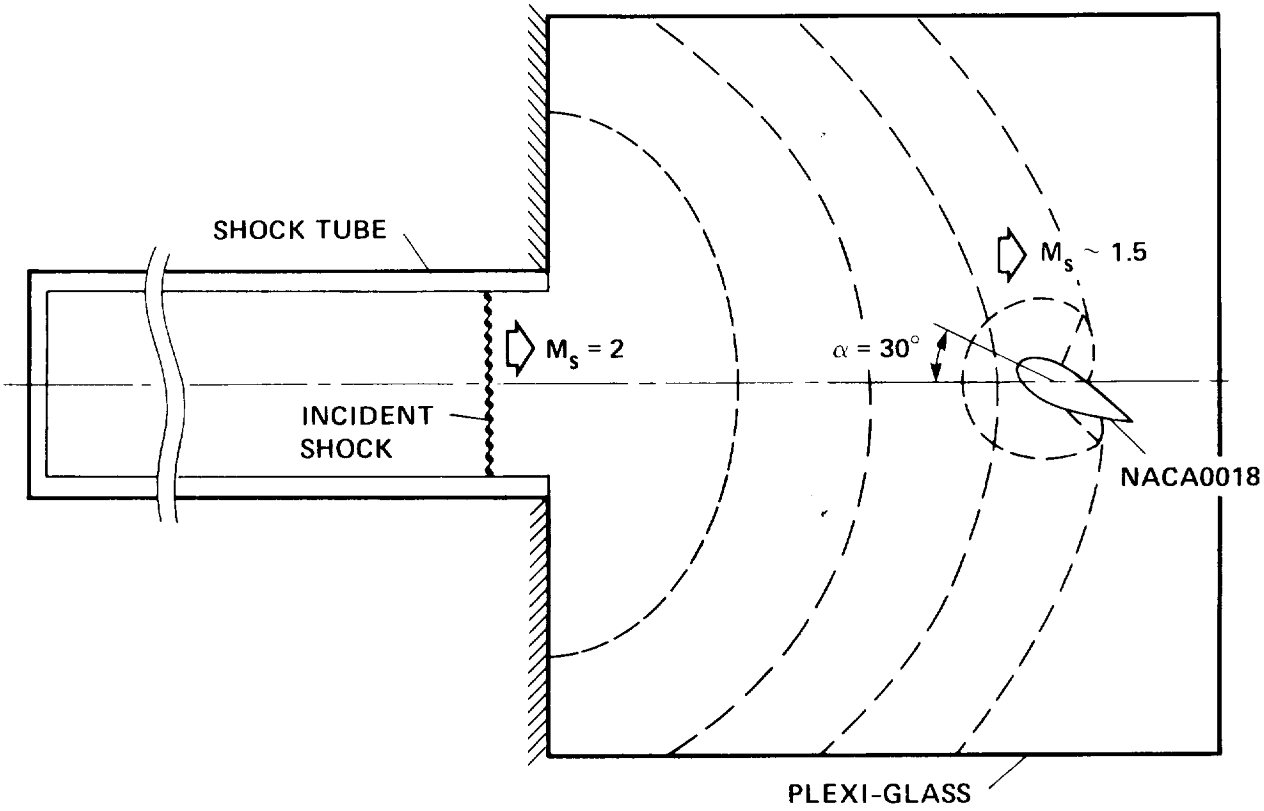


Fig. 5.8 Schematic of the experiment.

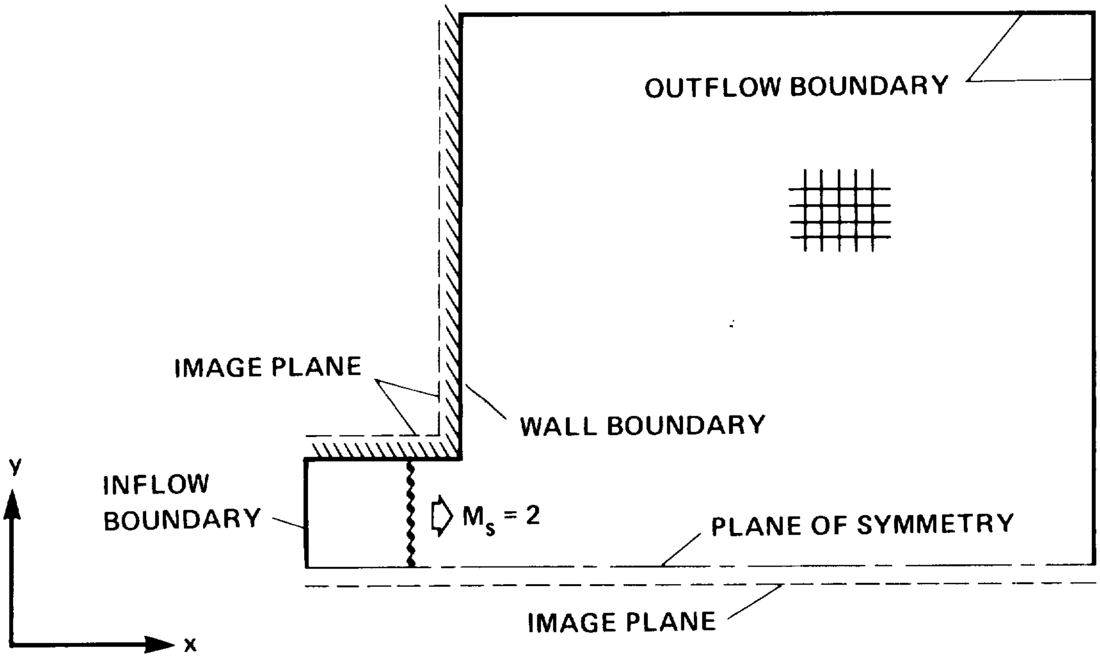


Fig. 5.9 Schematic of the shock tube computational domain (Cartesian grid).

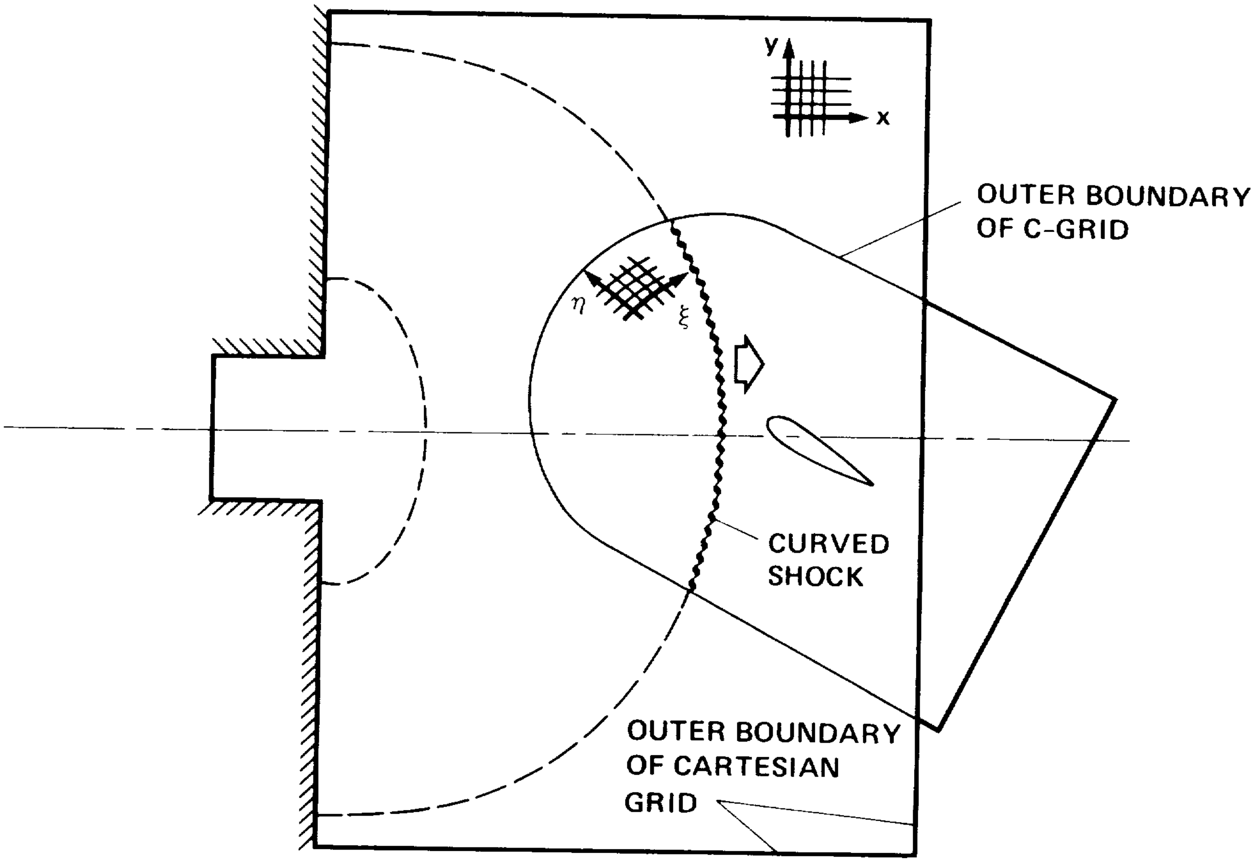


Fig. 5.10 Schematic of the computational domain (curved incident shock).

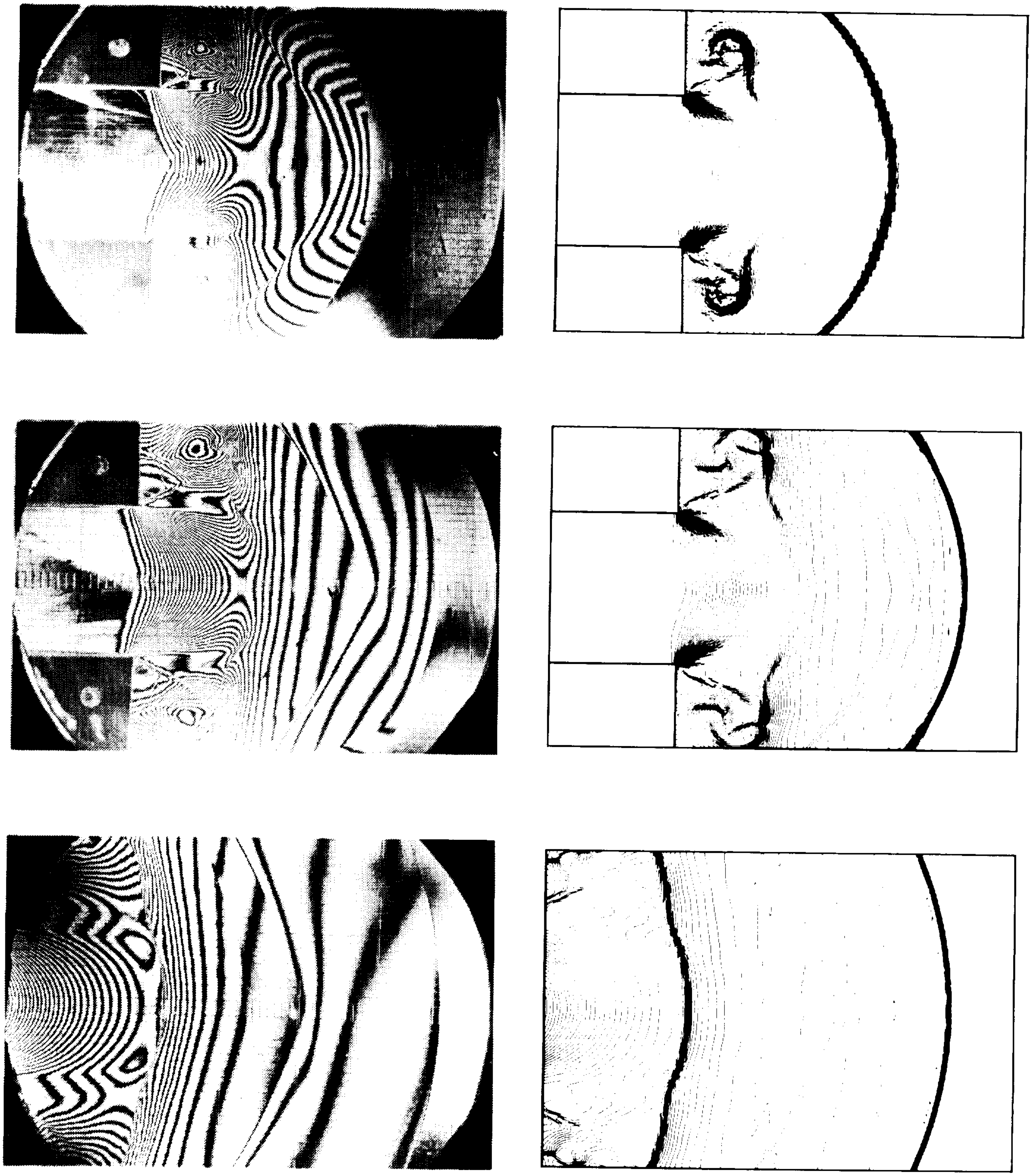


Fig. 5.11 Density contours computed by the time-split upwind TVD scheme (4.34a,c,g) compared with interferograms at approximately the same time instants.

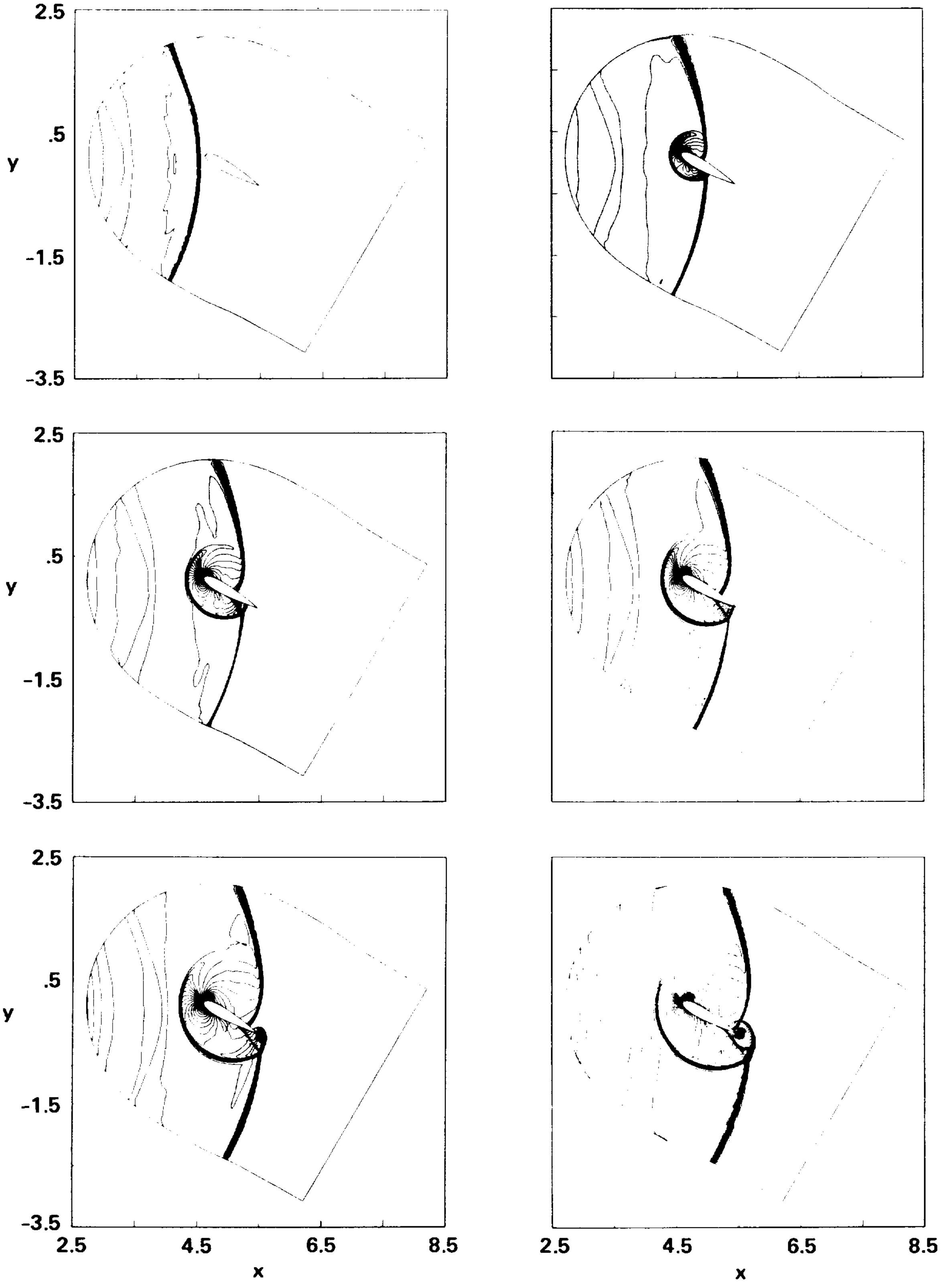


Fig. 5.12 Density contours computed by the time-split upwind TVD scheme (4.34a,d,g) at six sequential stages of the diffraction process.

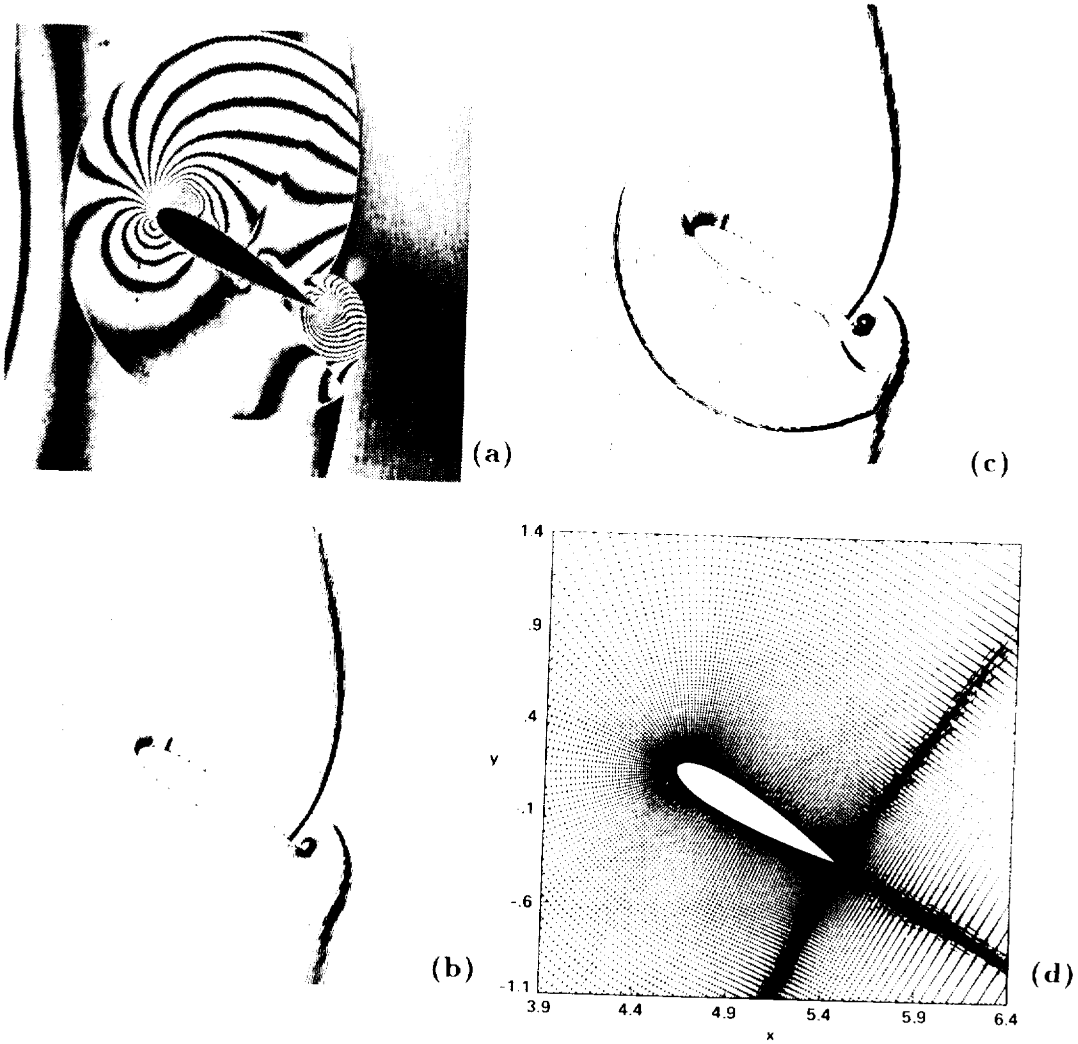


Fig. 5.13 Interferogram (a), Density contours (b), pressure contours (c), and grid (d) at an instant where a vortex is being formed near the trailing edge of the airfoil.

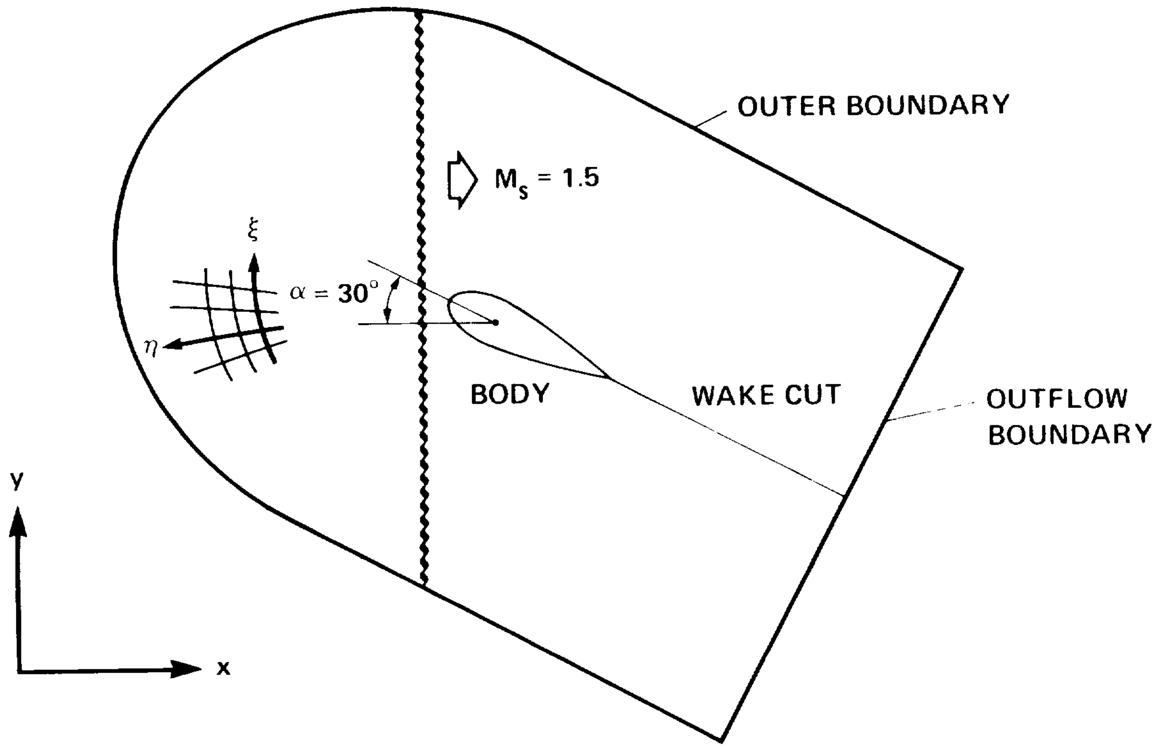


Fig. 5.14 Schematic of the computational domain (planar incident shock).

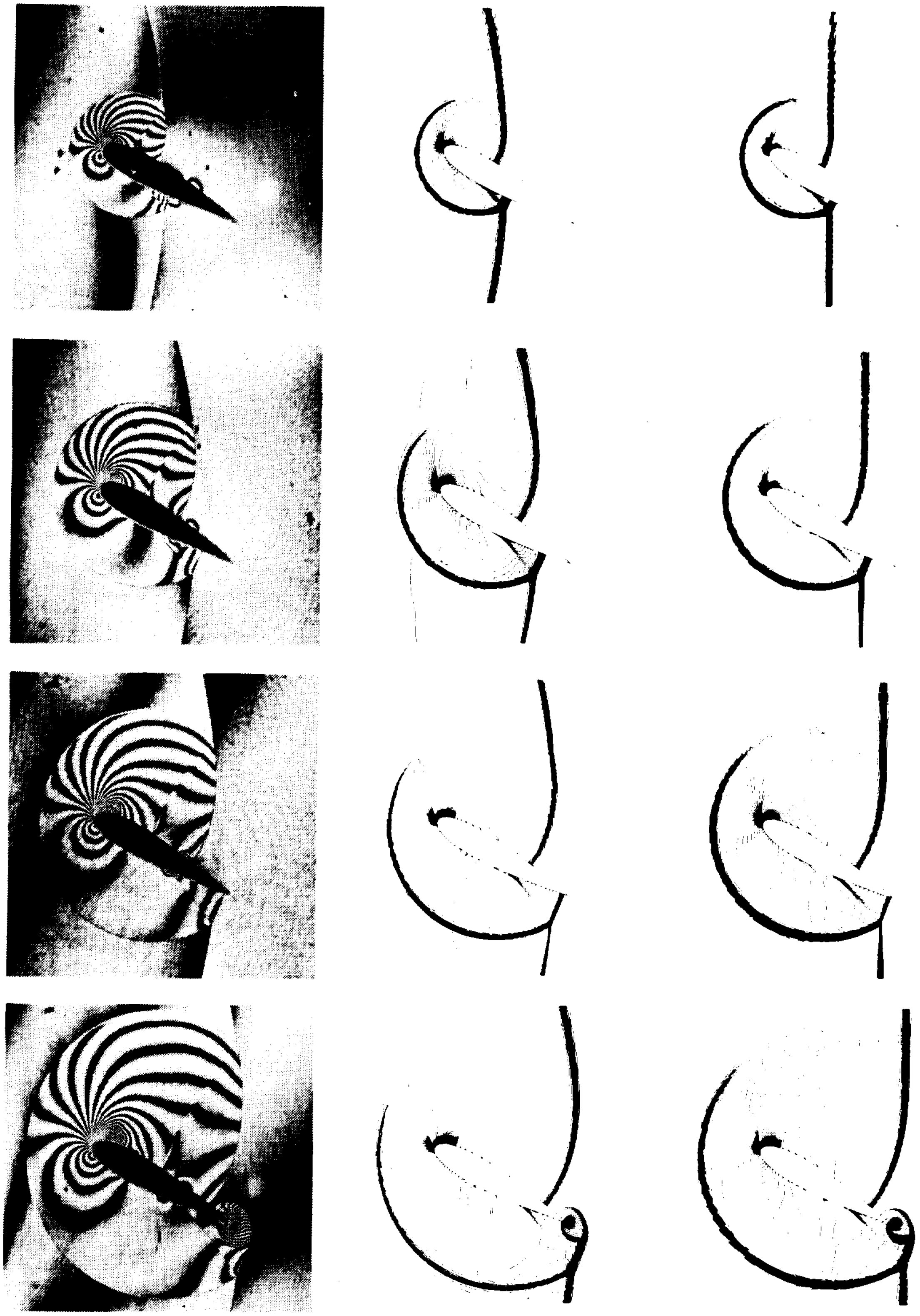


Fig. 5.15 Density contours computed by the time-split upwind TVD scheme (4.34a,d,g) using a curved shock (middle) and a planar shock (right) compared with the interferograms at approximately the same time instants.

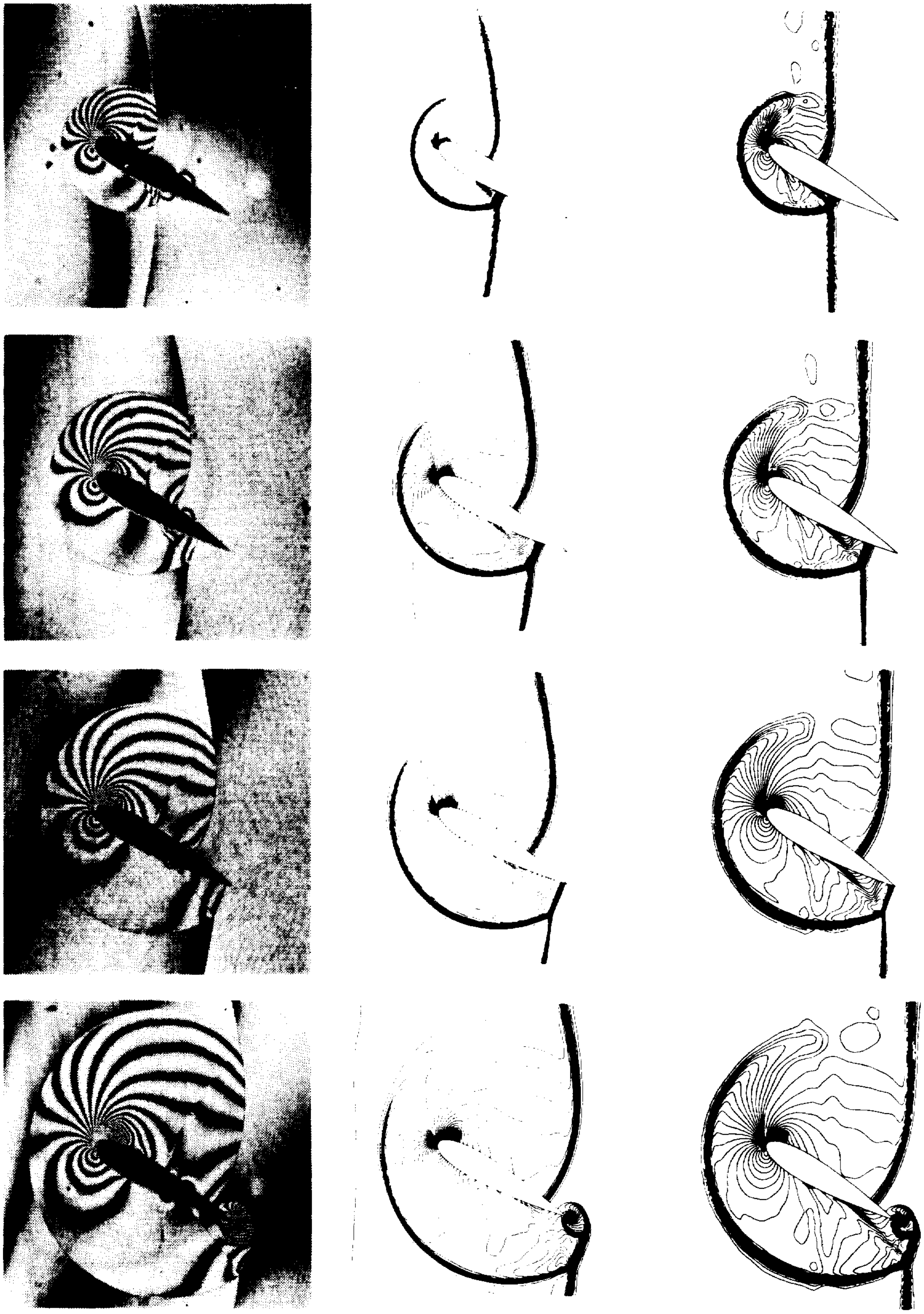


Fig. 5.16 Density contours computed by the time-split symmetric TVD scheme (4.33a,e) using a curved shock (middle) and a planar shock (right) compared with the interferograms at approximately the same time instants.

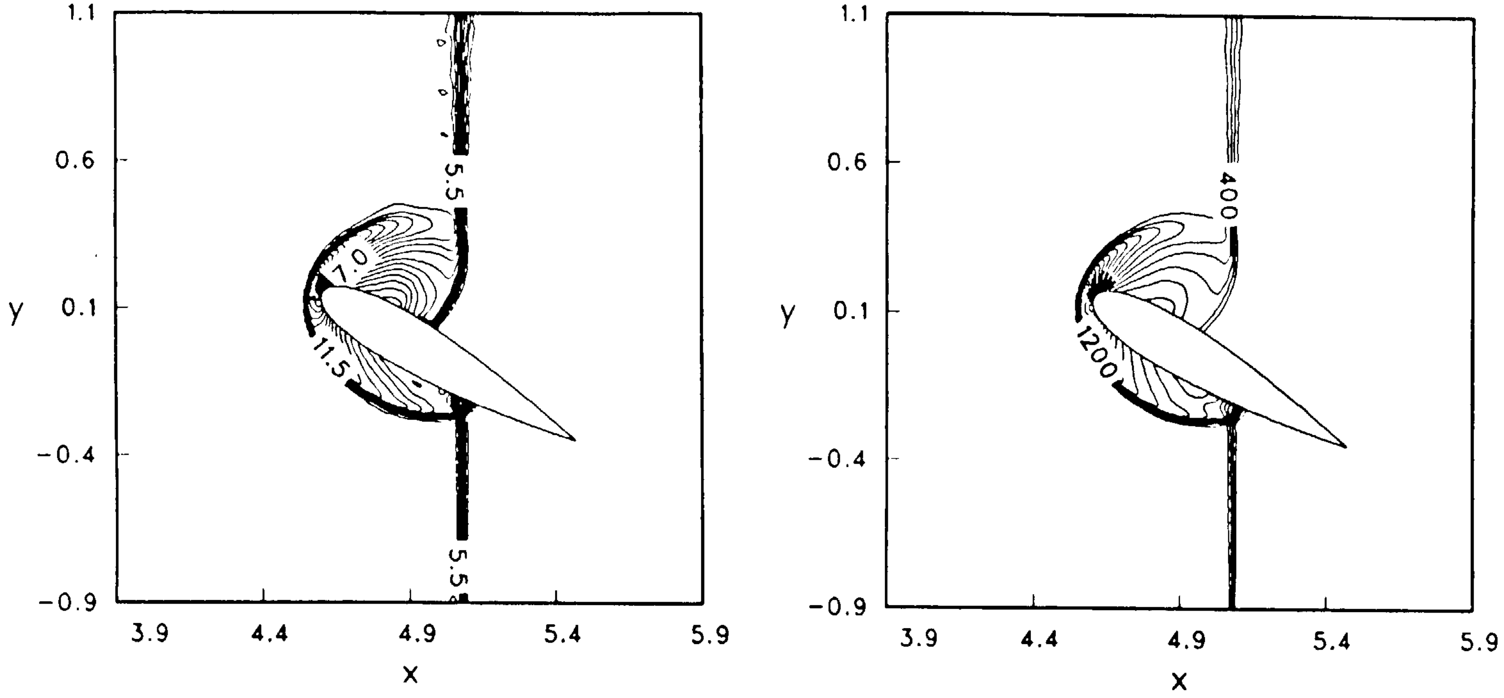


Fig. 5.17 Density contours (left) and pressure contours (right) computed by the time-split symmetric TVD scheme (4.33a,e) for the NACA 0018 airfoil with $M_s = 20, \alpha = 30^\circ$.

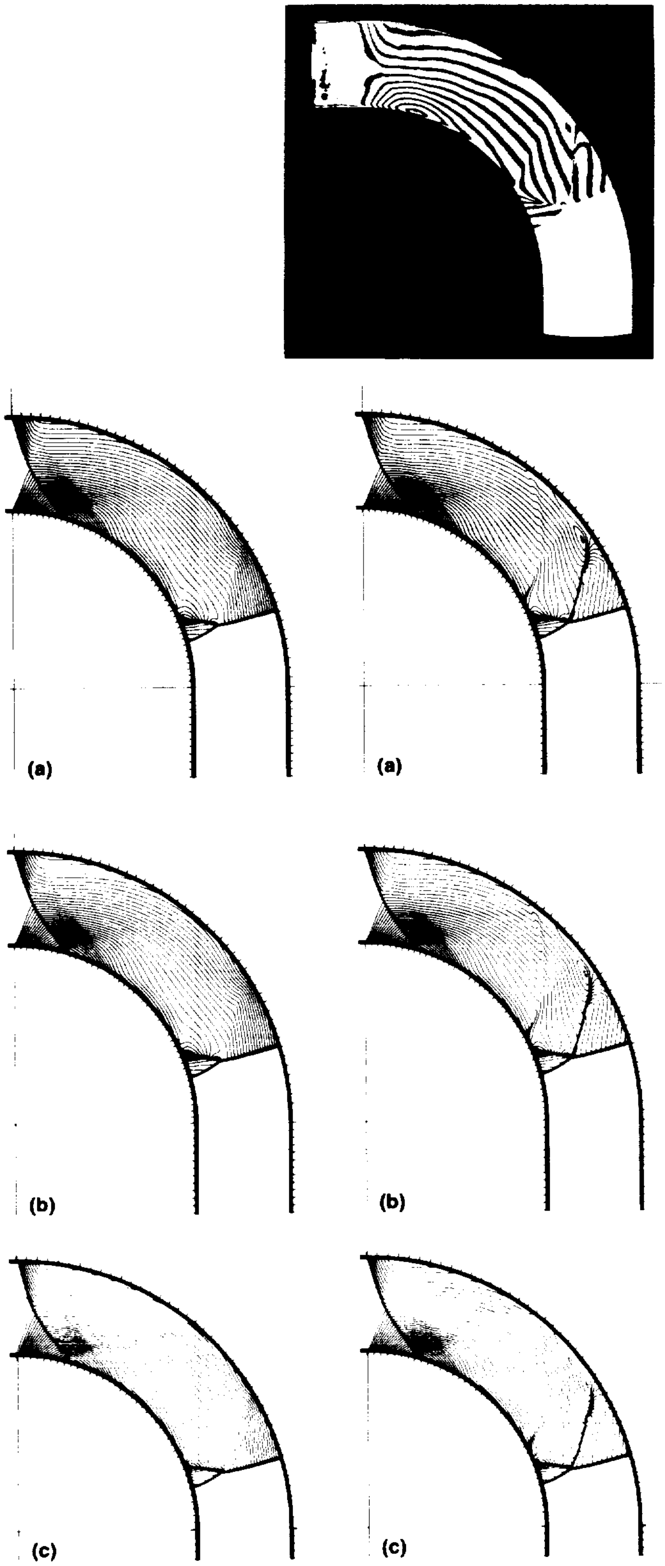
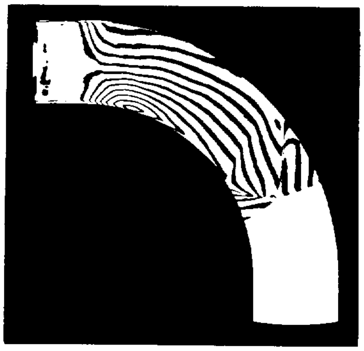


Fig. 5.18 Pressure contours (left) and density contours (right) computed by the time-split upwind TVD scheme (4.34) compared with experiment.
(a) Limiter (4.34d); (b) limiters (4.34d,g); (c) limiter (4.34g)

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH



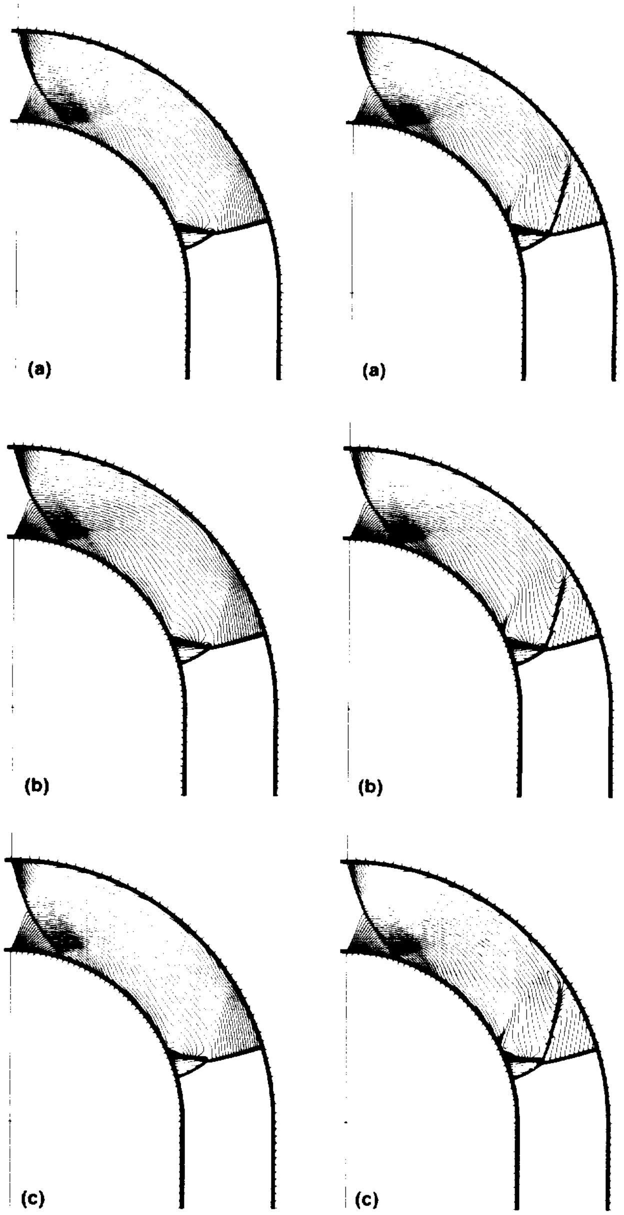
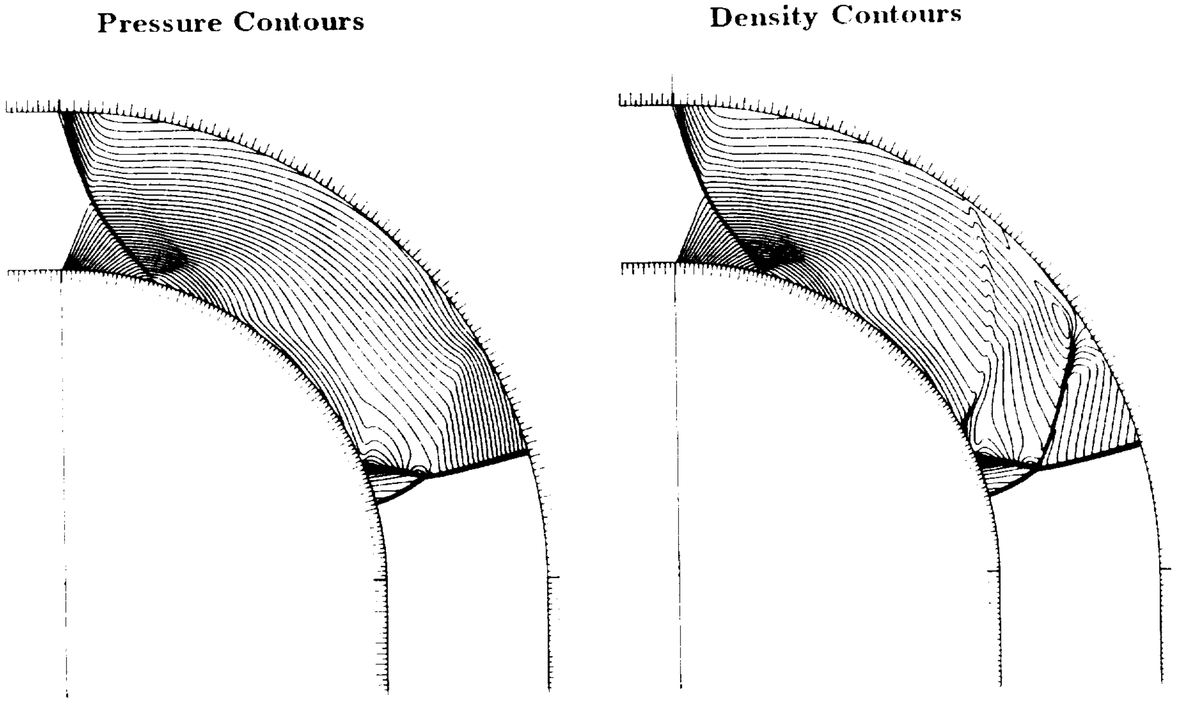


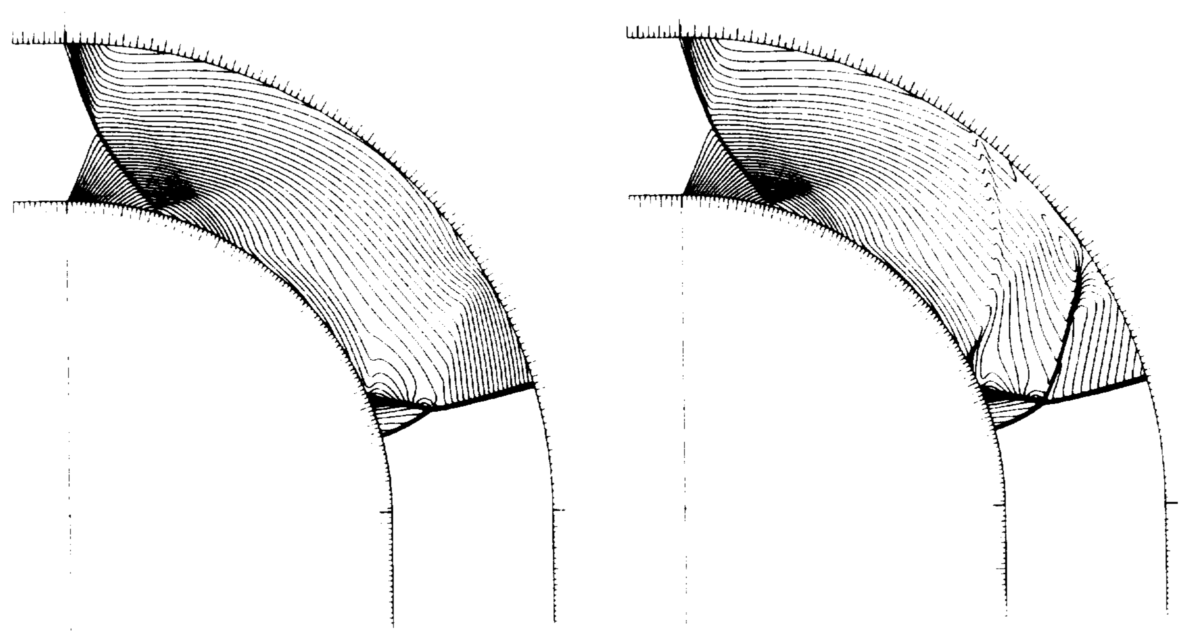
Fig. 5.19 Pressure contours (left) and density contours (right) computed by the time-split symmetric TVD scheme (4.33) compared with experiment.
(a) Limiter (4.33c); (b) limiters (4.33d); (c) limiter (4.33e).



Pressure Contours

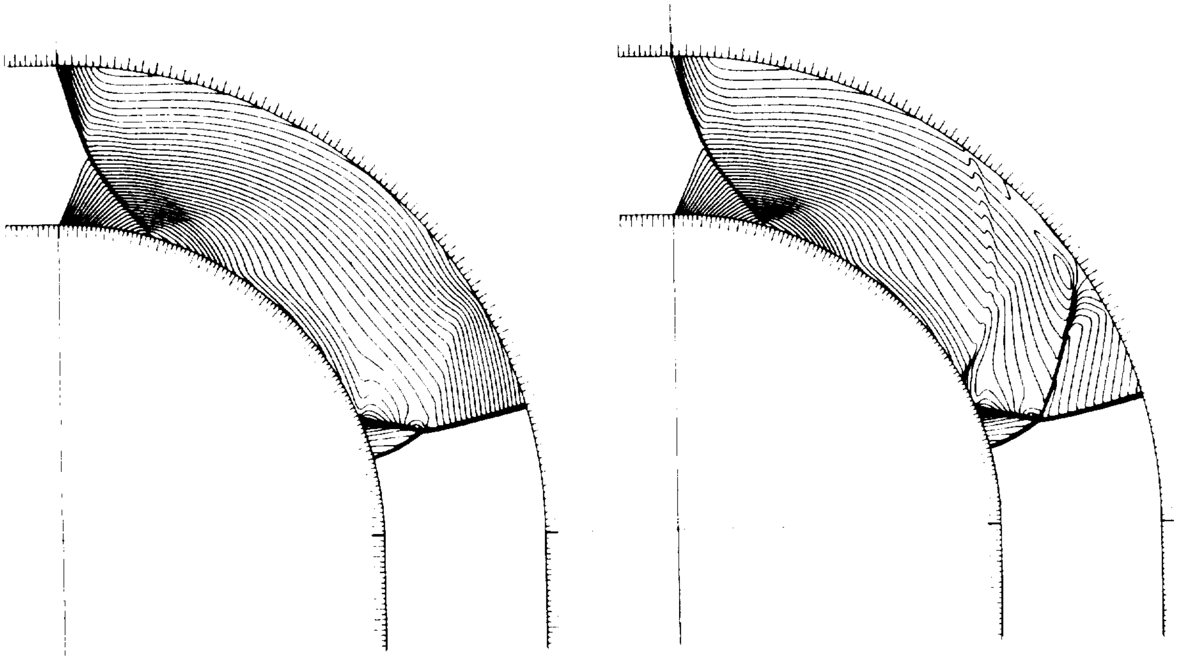
Density Contours

Finite Volume Formulation (5.6)



Pseudo Finite Volume Formulation (5.10)

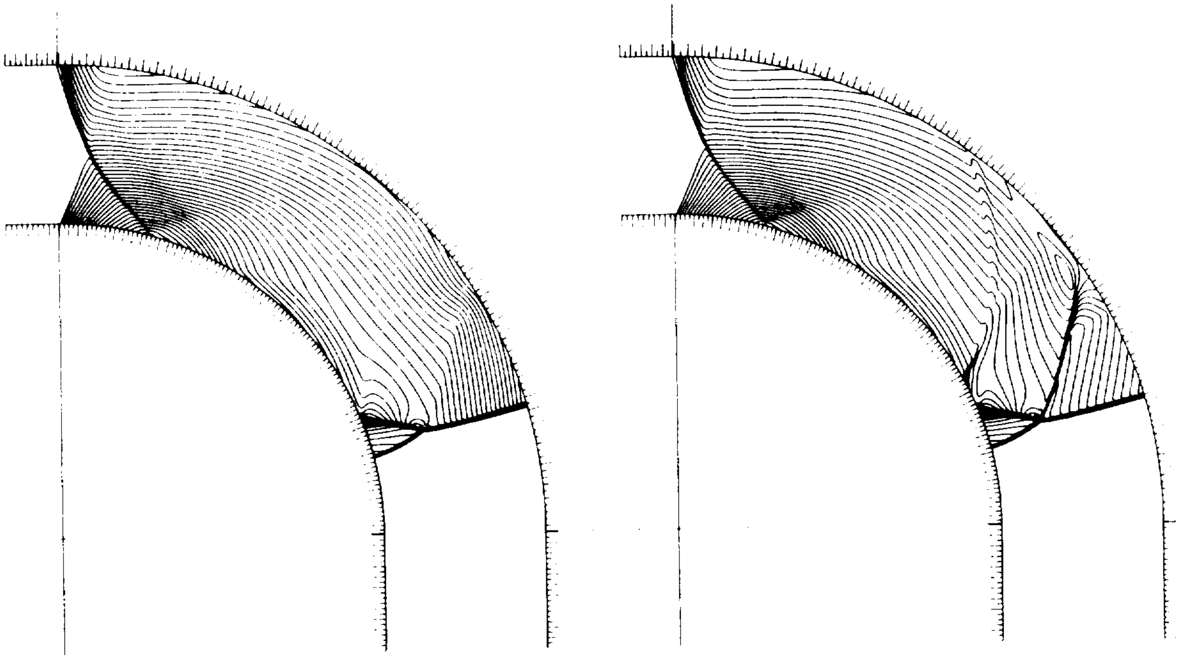
Fig. 5.20 Comparison between the finite volume formulation (5.6) and the pseudo finite volume formulations (5.10) - (5.12) and equation (5) of reference [24] respectively on a (701 × 151) grid.



Pressure Contours

Density Contours

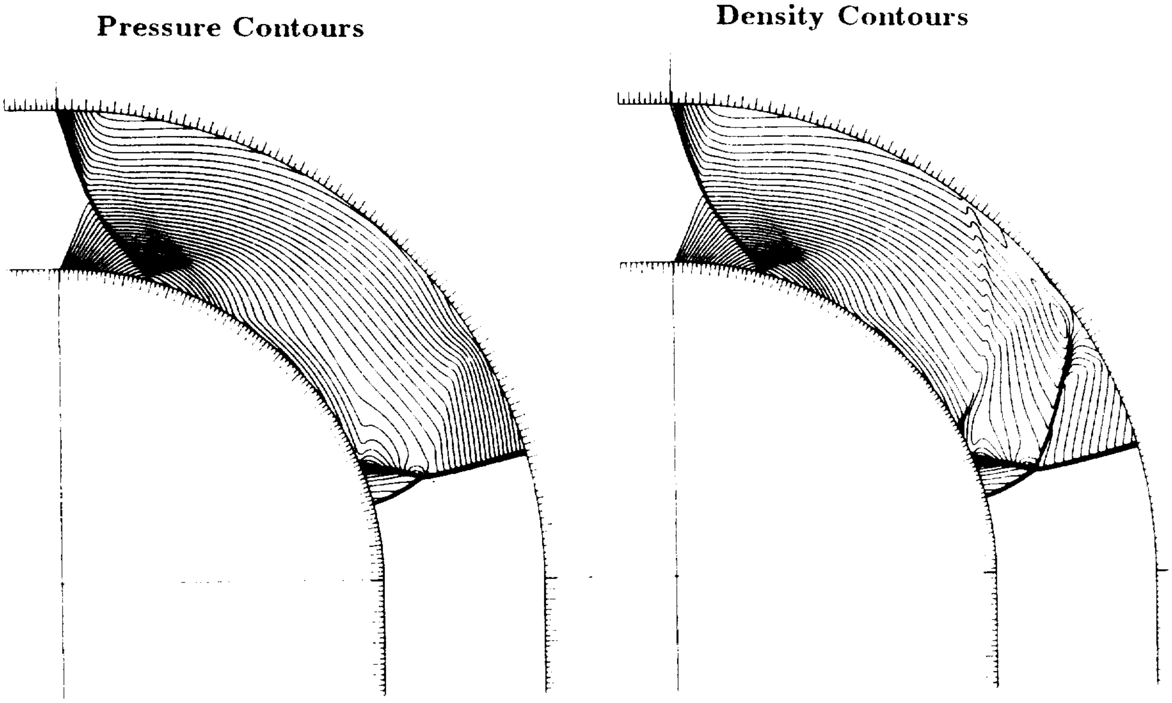
Pseudo Finite Volume Formulation (5.11)



Pseudo Finite Volume Formulation (5.12)

Fig. 5.20. Continued

Pressure Contours **Density Contours**



Pseudo Finite Volume Formulation (5) of Reference [24]

Fig. 5.20. Concluded

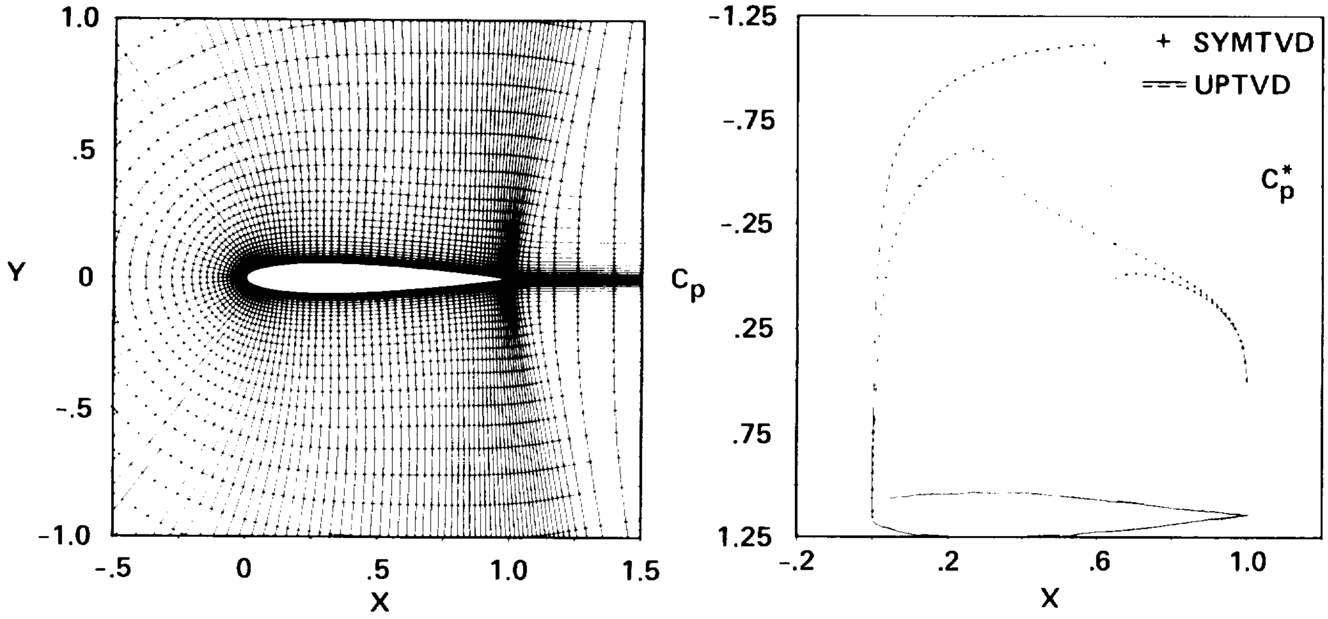


Fig. 5.21. Comparison of a symmetric TVD (SYMTVD) scheme with an upwind TVD (UPTVD) scheme for the NACA0012 airfoil with $M_\infty = 0.8$, $\alpha = 1.25$ using a 163 × 49 C grid.

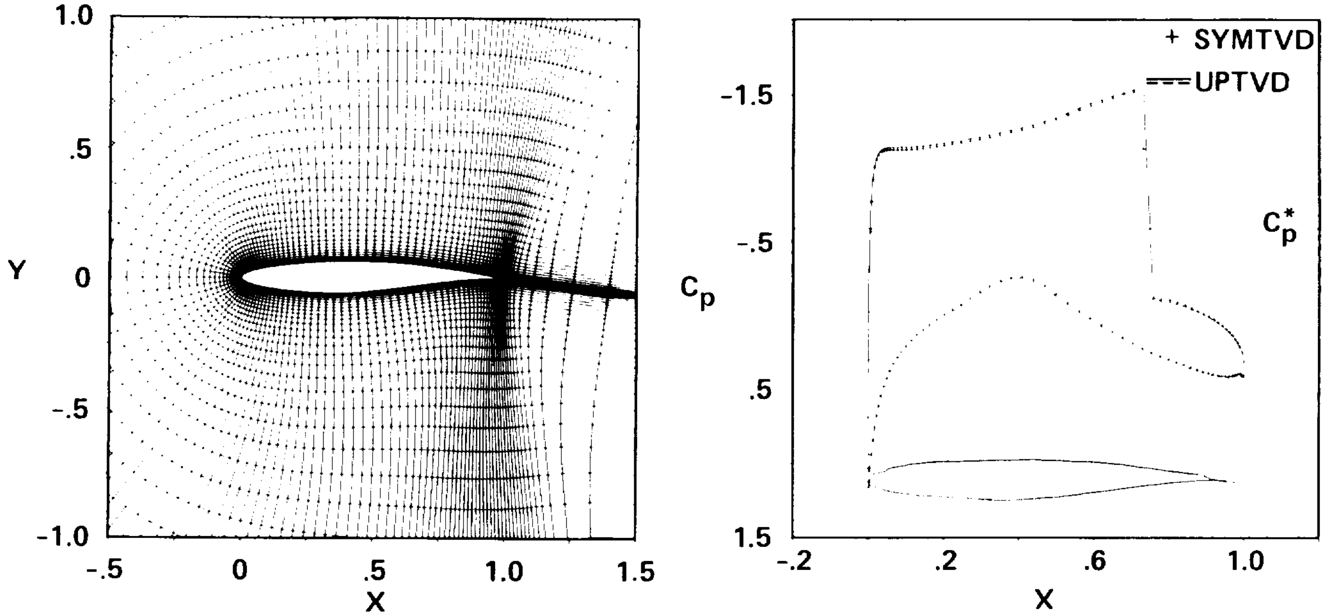


Fig. 5.22. Comparison of a symmetric TVD (SYMTVD) scheme with an upwind TVD (UPTVD) scheme for the RAE2822 airfoil with $M_\infty = 0.75$, $\alpha = 3$ using a 163 × 49 C grid.

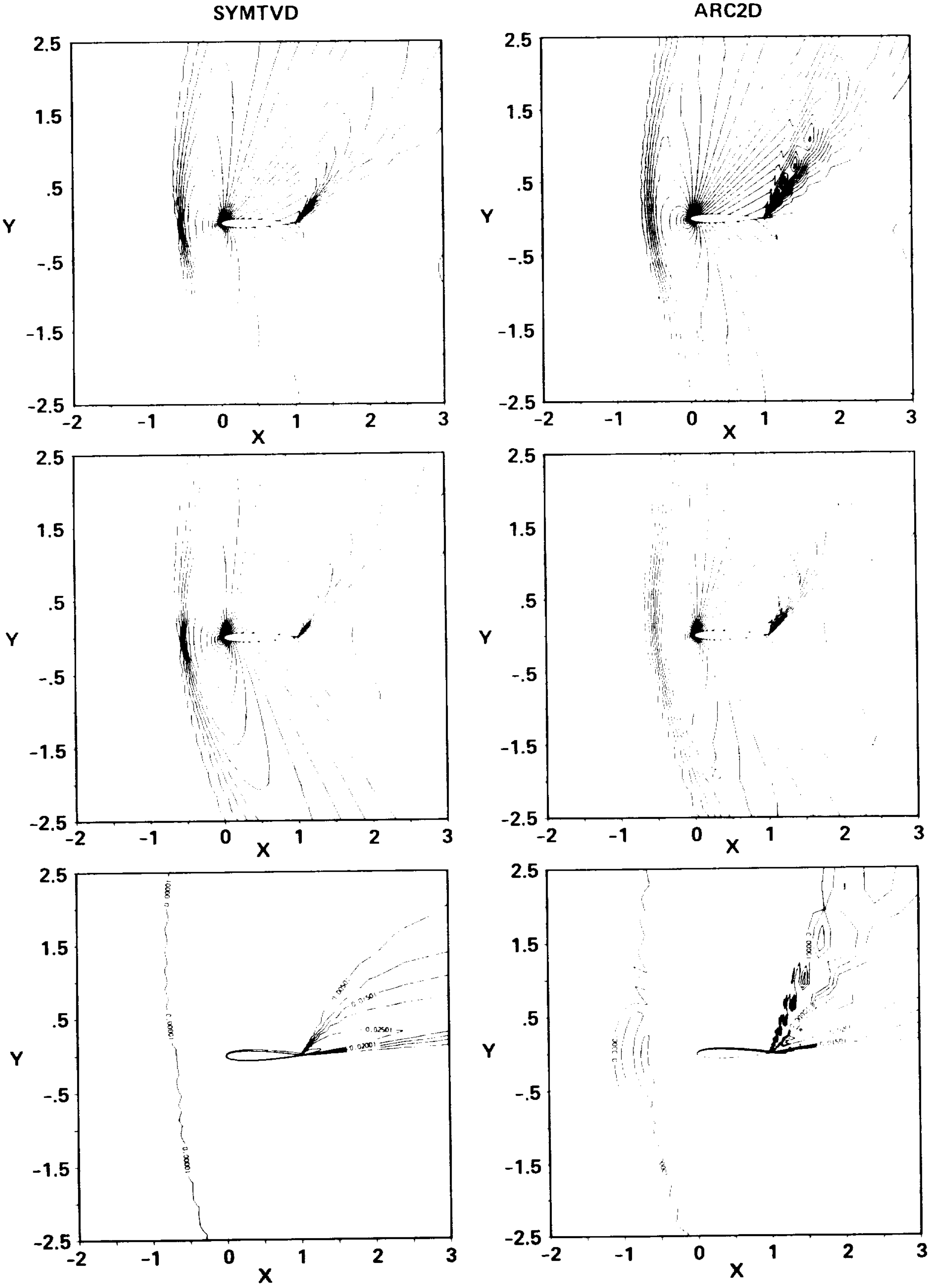


Fig. 5.23. Comparison of a symmetric TVD (SYMTVD) scheme with ARC2D (version 150) for the Mach contours, pressure contours and entropy contours of the NACA 0012 airfoil with $M_\infty = 1.2, \alpha = 7$ using a 163×49 C grid as shown in figure 5.21.

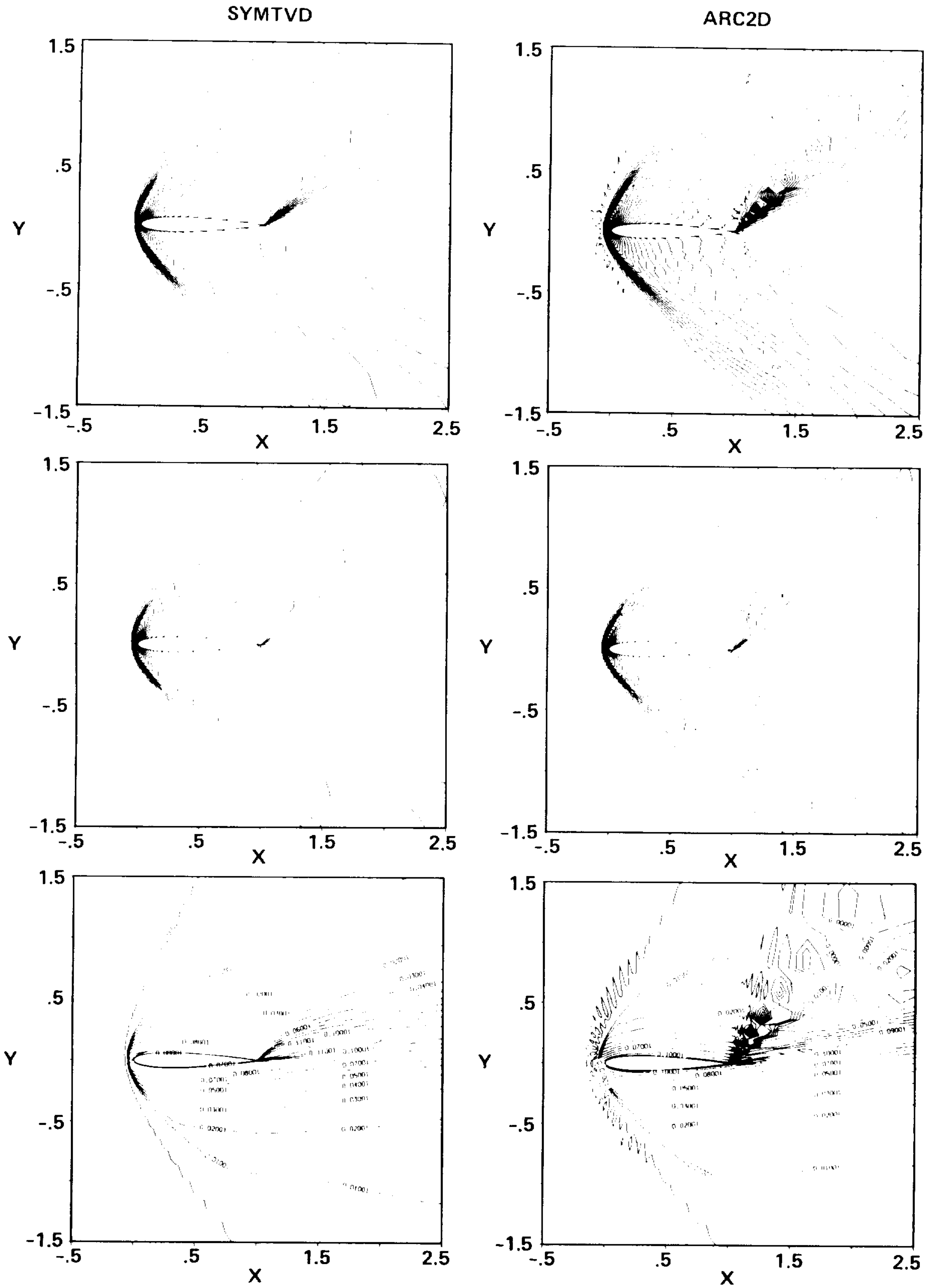


Fig. 5.24. Comparison of a symmetric TVD (SYMTVD) scheme with ARC2D (version 150) for the Mach contours, pressure contours and entropy contours of the NACA 0012 airfoil with $M_\infty = 1.8, \alpha = 7$ using a 163×49 C grid as shown in figure 5.21.

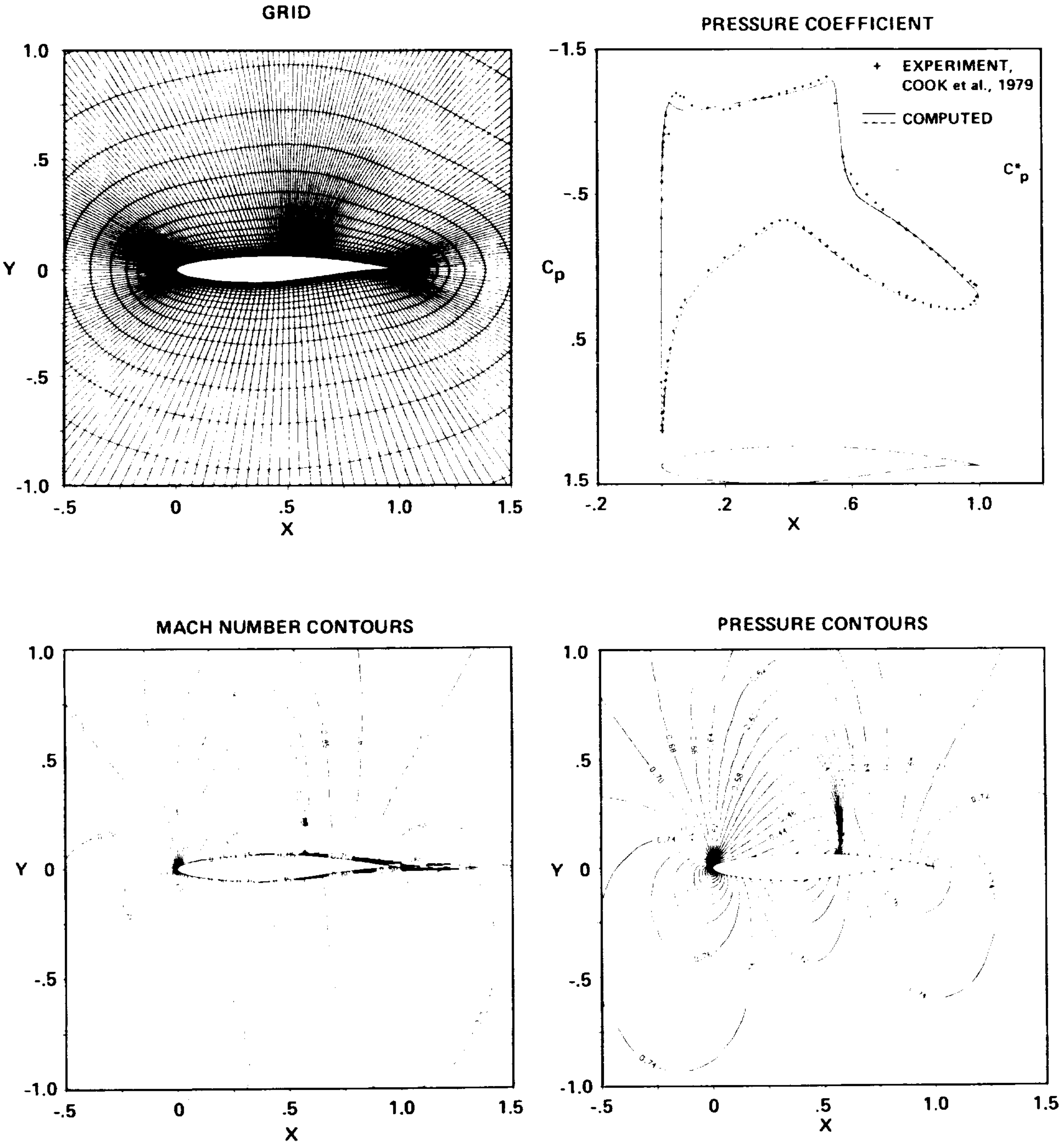


Fig. 5.25. A thin layer Navier-Stokes calculation for the RAE2822 airfoil using an upwind TVD scheme with a 249 × 51 O grid, $M_\infty = 0.73, \alpha = 2.79, Re = 6.5 \times 10^6$.

(a) Pressure.

(b) Energy.

Fig. 5.26. Contour plots in plane of symmetry of aerobrake flow field for Mach 10 flow at 0° angle of attack.

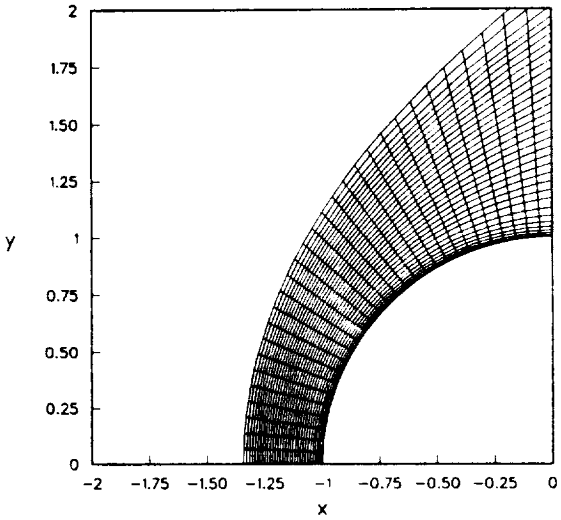


Fig. 6.1 **The 31 × 33 grid.**

C-3

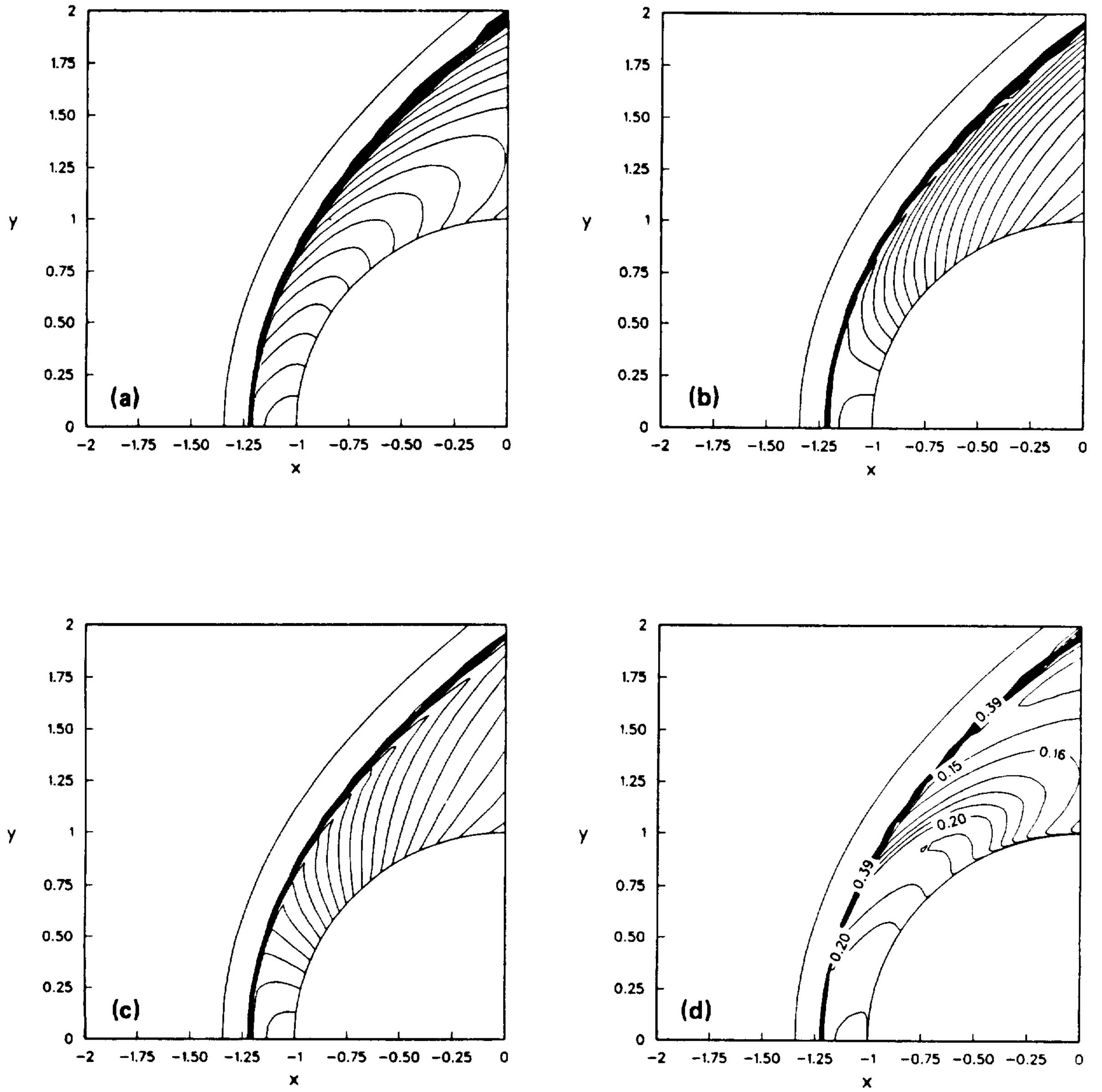


Fig. 6.2 The Mach contours (a), density contours (b), pressure contours (c) and κ (d) computed by the implicit scheme (6.5) ($\theta = 1$, $\omega = 0$) for an equilibrium **real gas** with $M_\infty = 15$.

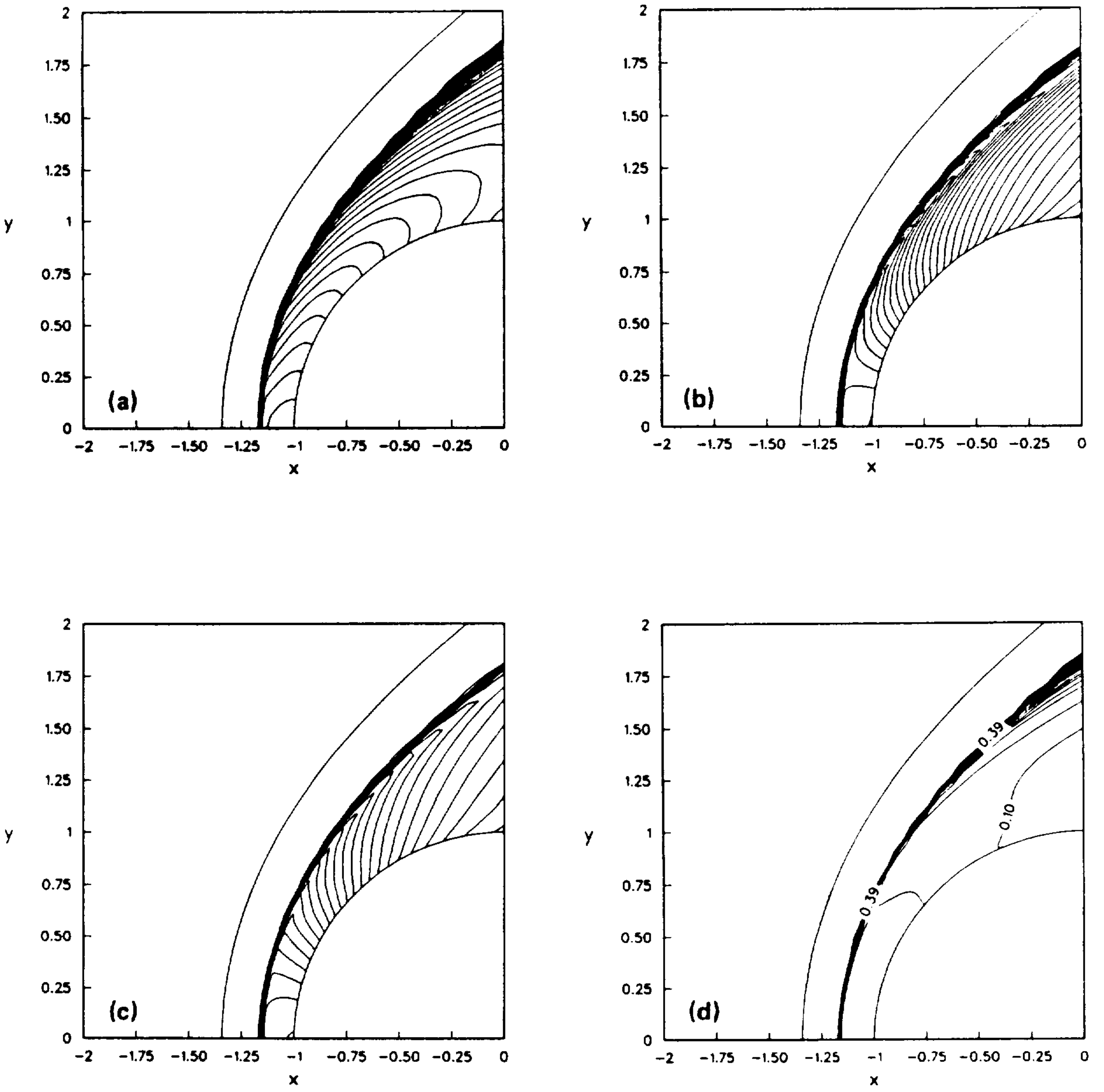


Fig. 6.3 The Mach contours (a), density contours (b), pressure contours (c) and κ (d) computed by the implicit scheme (6.5) ($\theta = 1$, $\omega = 0$) for an equilibrium real gas with $M_\infty = 25$.

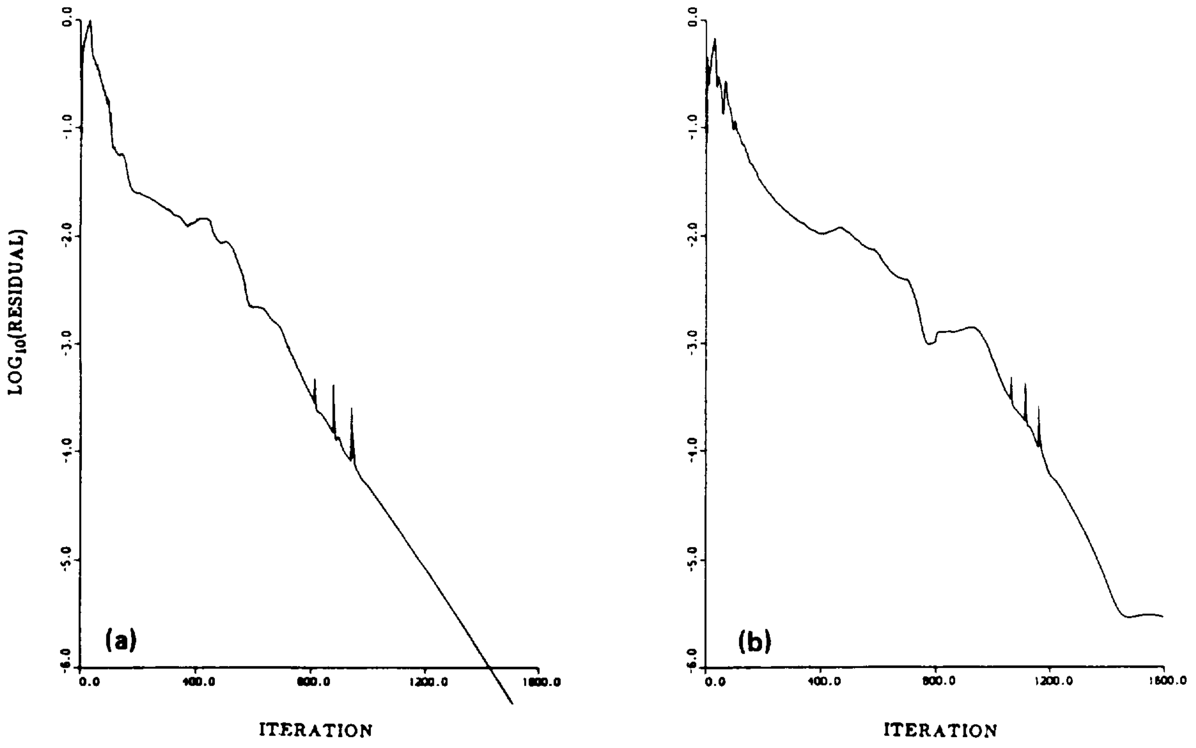


Fig. 6.4 Comparison of the L_2-norm residual of a linearized conservative implicit operator (a) and a linearized nonconservative implicit operator (b) for an equilibrium real gas with $M_\infty = 25$.

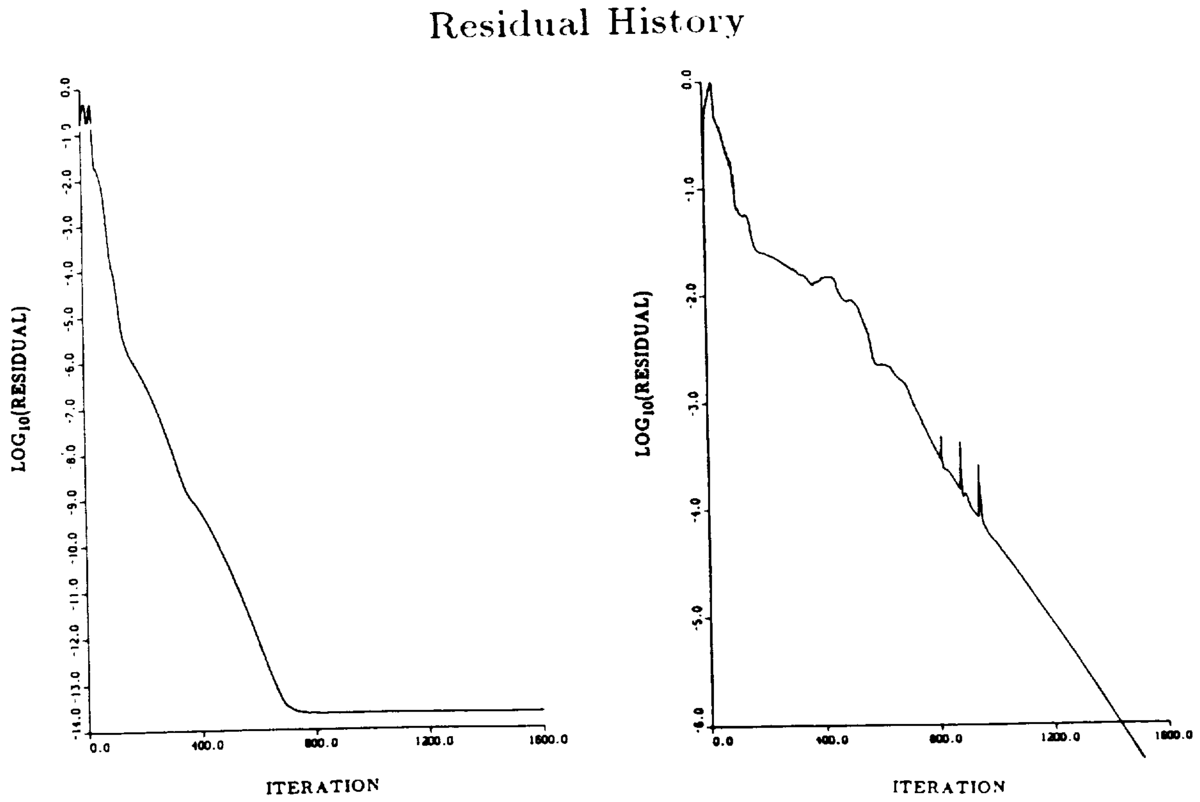


Fig. 6.5 Comparison of the L_2-norm residual of a **perfect gas** and **real gas** computed by the scheme (6.5) ($\theta = 1$, $\omega = 0$) with $M_\infty = 25$. Note that the scale of the ordinate for the perfect gas and the real gas is not the same.

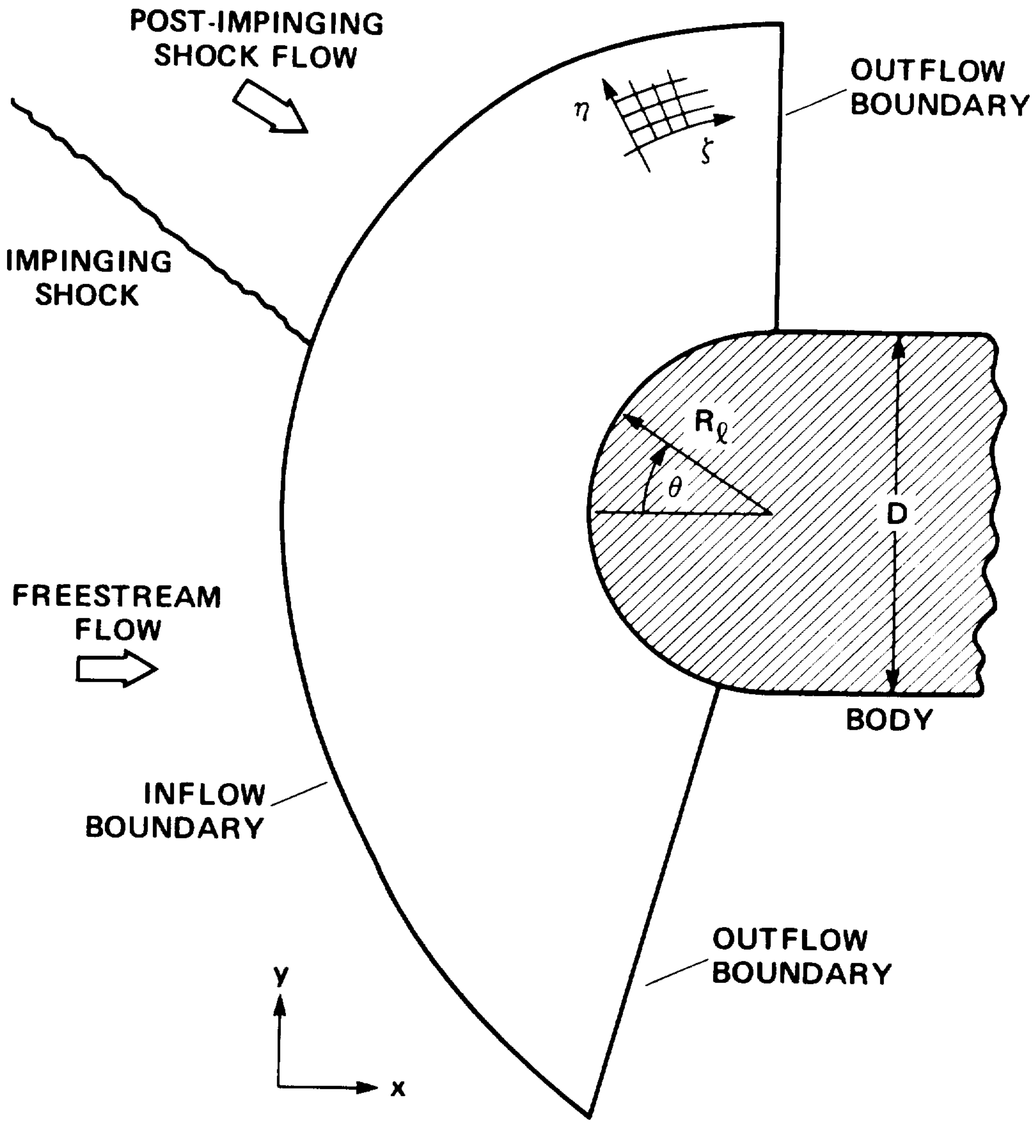


Fig. 6.6 Schematic of the computational domain for a blunt body flow with an impinging shock.

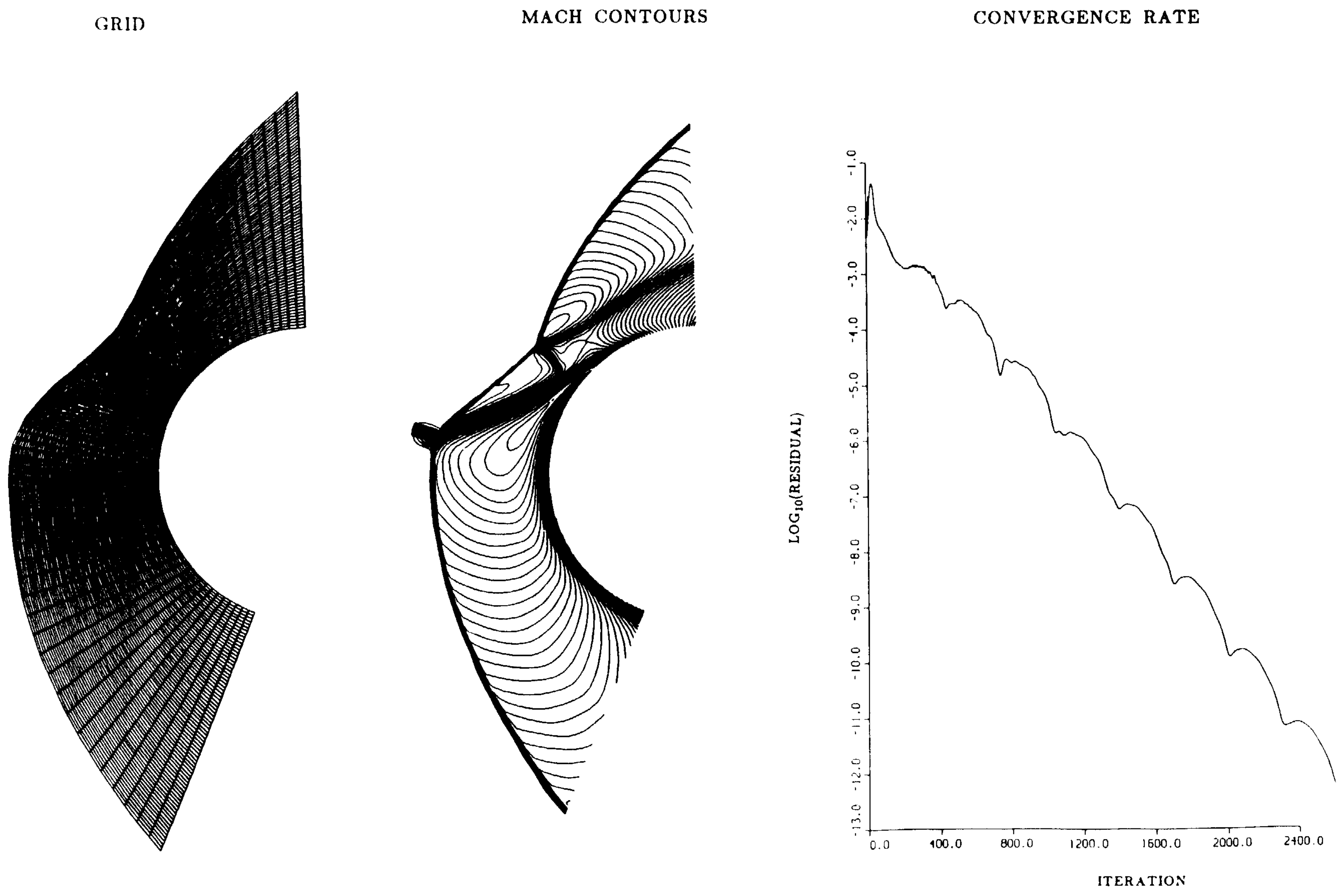


Fig. 6.7 Two-dimensional steady inviscid blunt-body flow with an impinging shock computed by the implicit scheme (6.5) ($\theta = 1$, $\omega = 0$) for a **perfect gas** with $M_\infty = 4.6$.

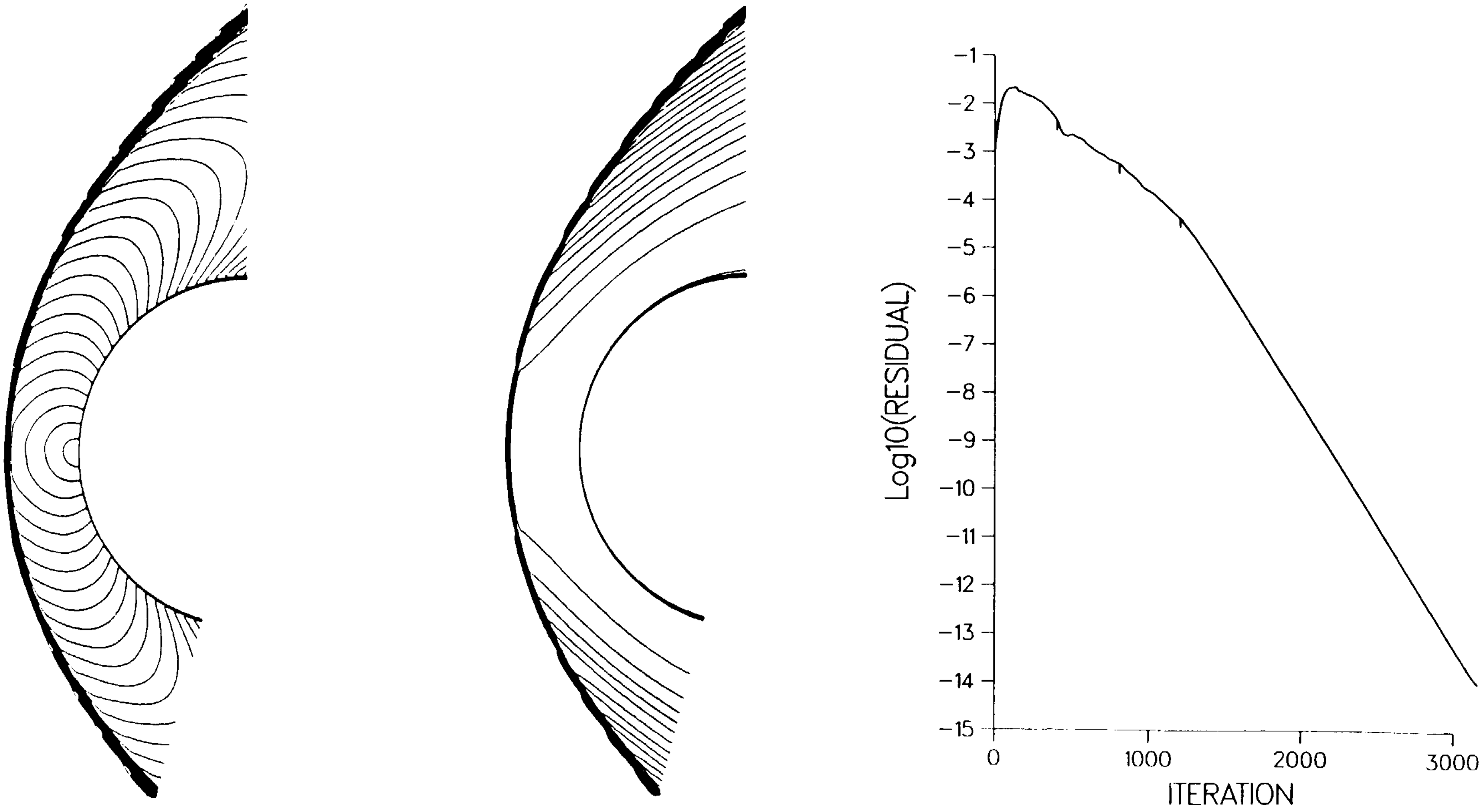


Fig. 6.8 Mach contours, entropy contours, and residual history for the steady viscous blunt body flow computed by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, full matrix form) with $M_\infty = 8.03$, $Re_D = 387{,}750$, $\gamma = 1.4$, and laminar flow.

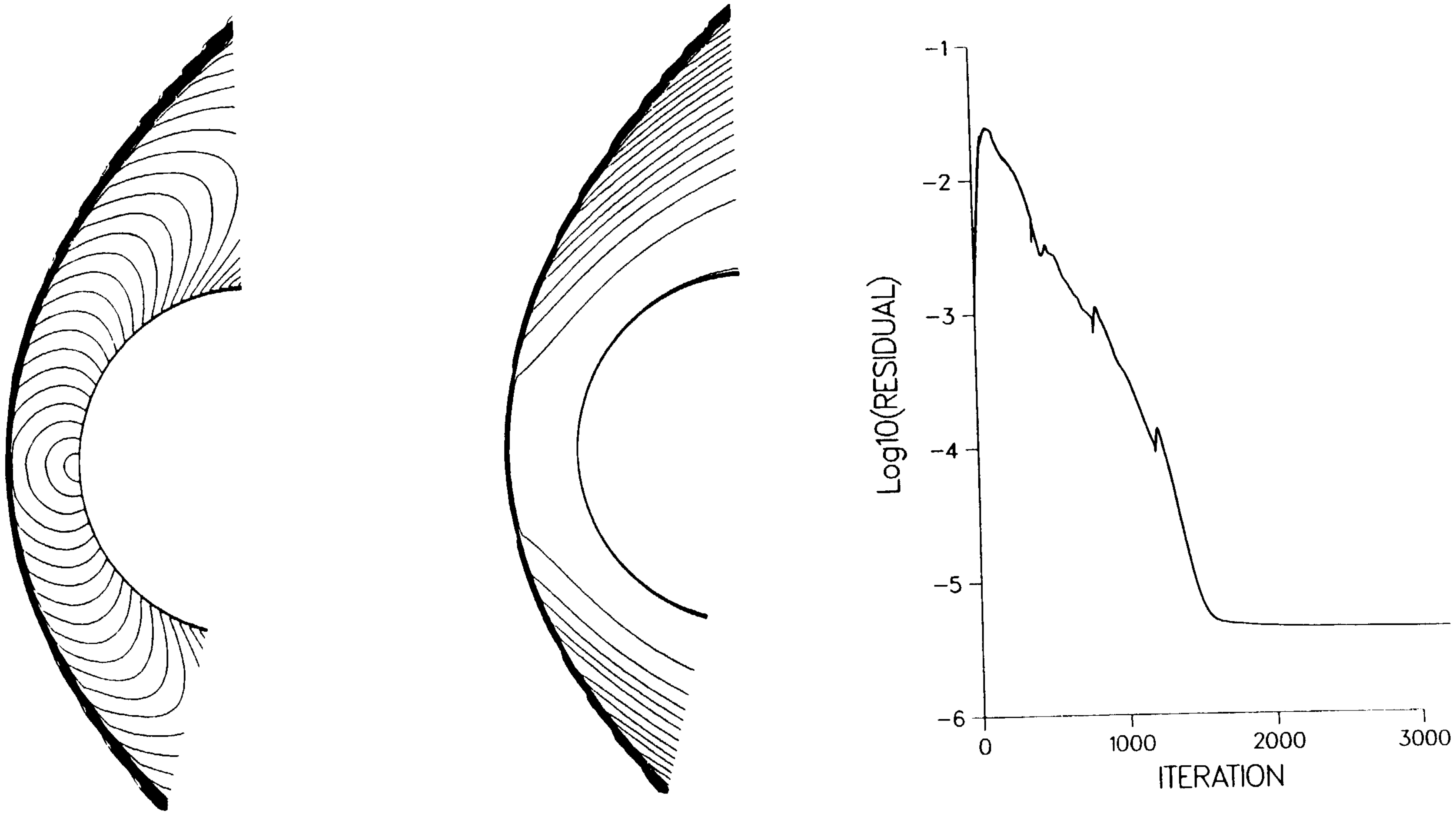


Fig. 6.9 Mach contours, entropy contours, and residual history for the steady viscous blunt body flow computed by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, diagonal form) with $M_\infty = 8.03$, $Re_D = 387{,}750$, $\gamma = 1.4$, and laminar flow.

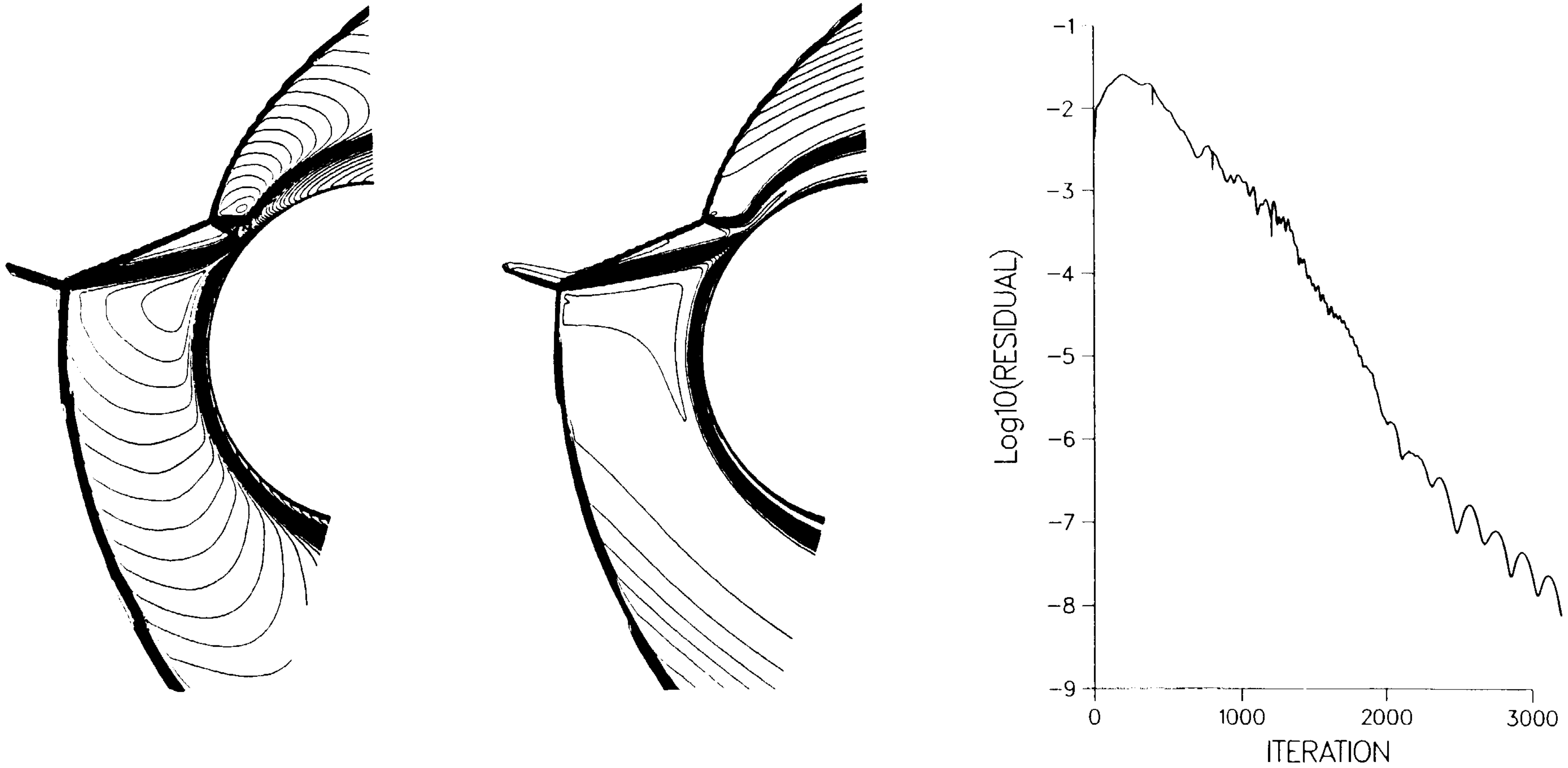


Fig. 6.10 Mach contours, entropy contours, and residual history for the steady viscous Type III shock interference flow computed by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, full matrix form) with $M_\infty = 8.03$, $Re_D = 387,750$, $\gamma = 1.4$, and laminar flow.

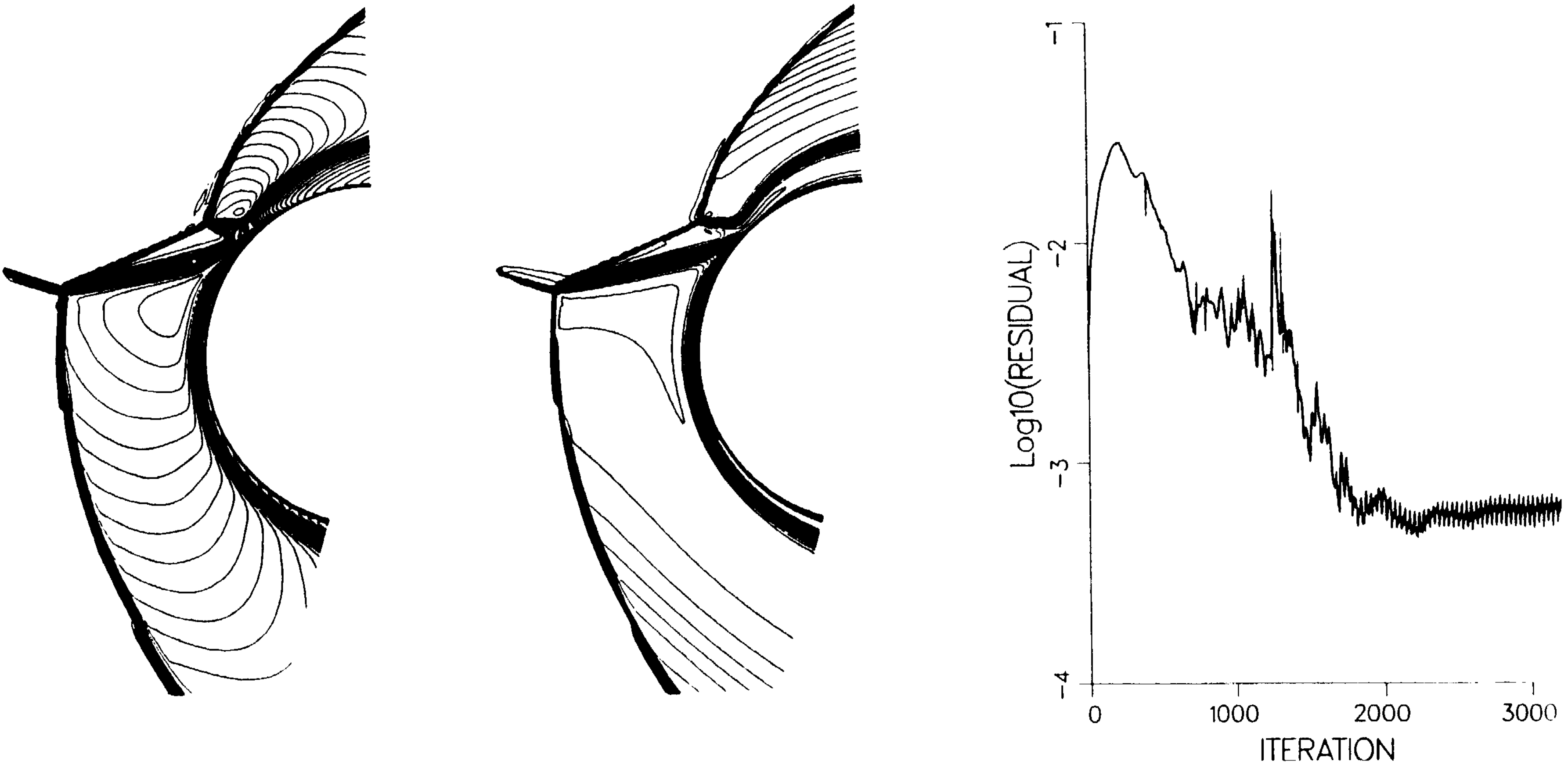


Fig. 6.11 Mach contours, entropy contours, and residual history for the steady viscous Type III shock interference flow computed by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, diagonal form) with $M_\infty = 8.03$, $Re_D = 387,750$, $\gamma = 1.4$, and laminar flow.

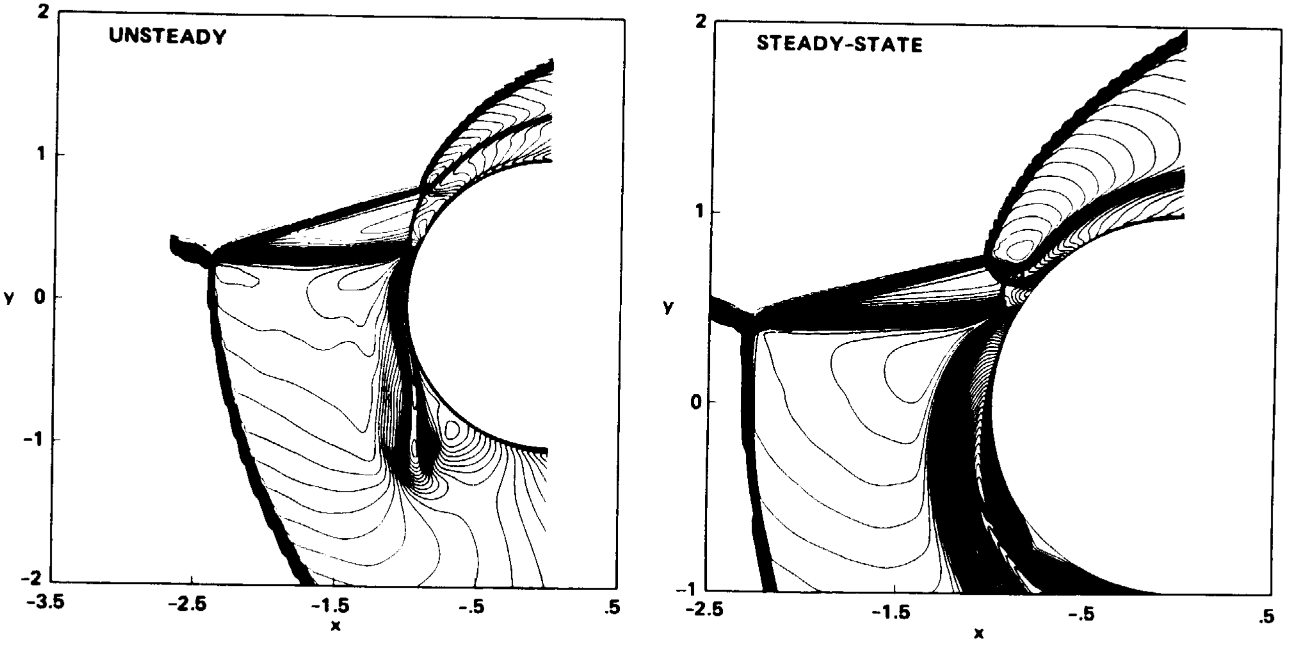
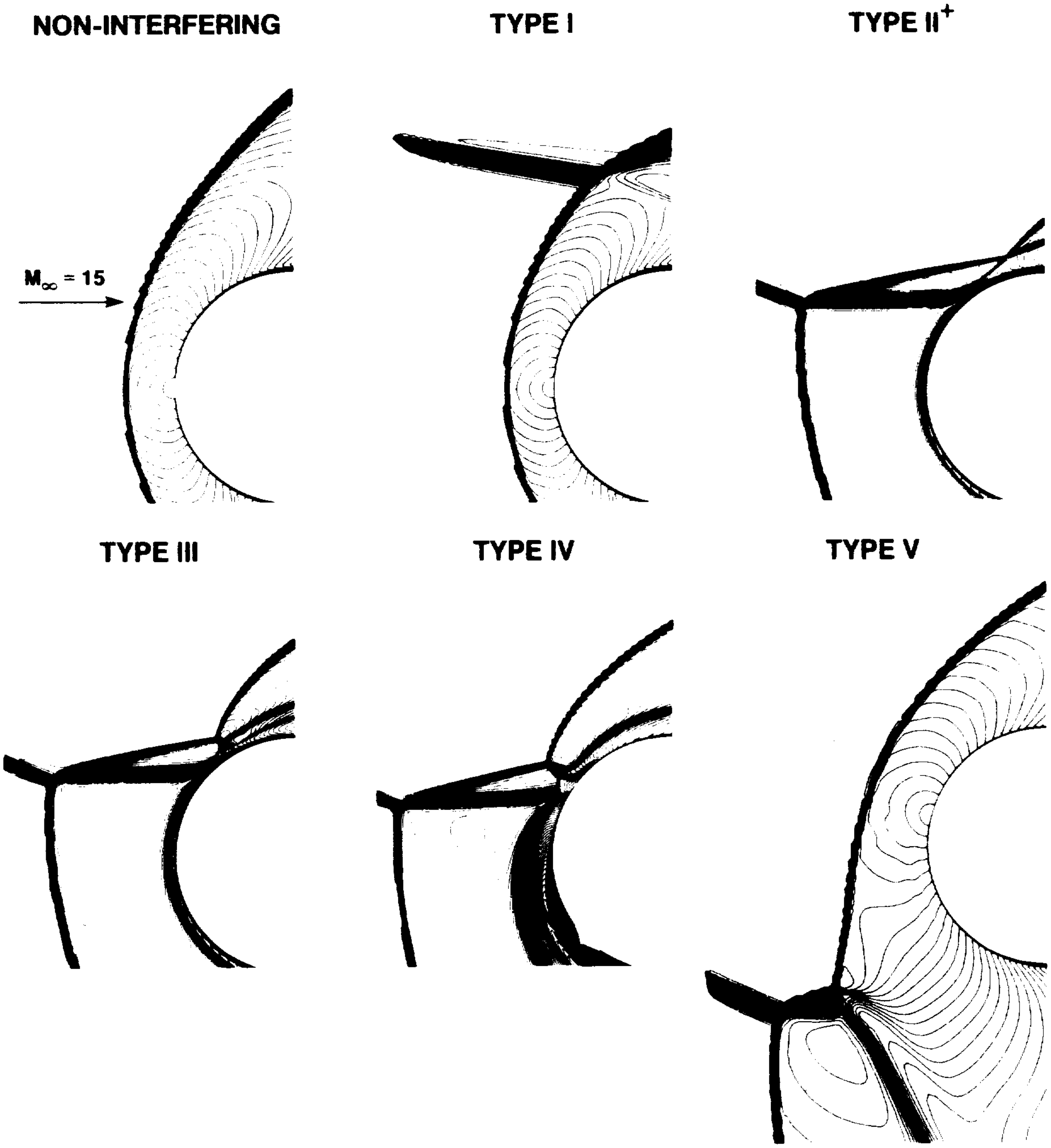


Fig. 6.12 The Mach contours of a two-dimensional steady viscous and unsteady hypersonic **perfect gas** computation by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, full matrix form) with $M_\infty = 15$ and $Re_D = 186{,}000$.


NON-INTERFERING

TYPE I

TYPE II$^+$

TYPE III

TYPE IV

TYPE V

Fig. 6.13 The Mach contours of a two-dimensional steady viscous hypersonic **perfect gas** computation by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, full matrix form) with $M_\infty = 15$ and $Re_D = 186,000$.

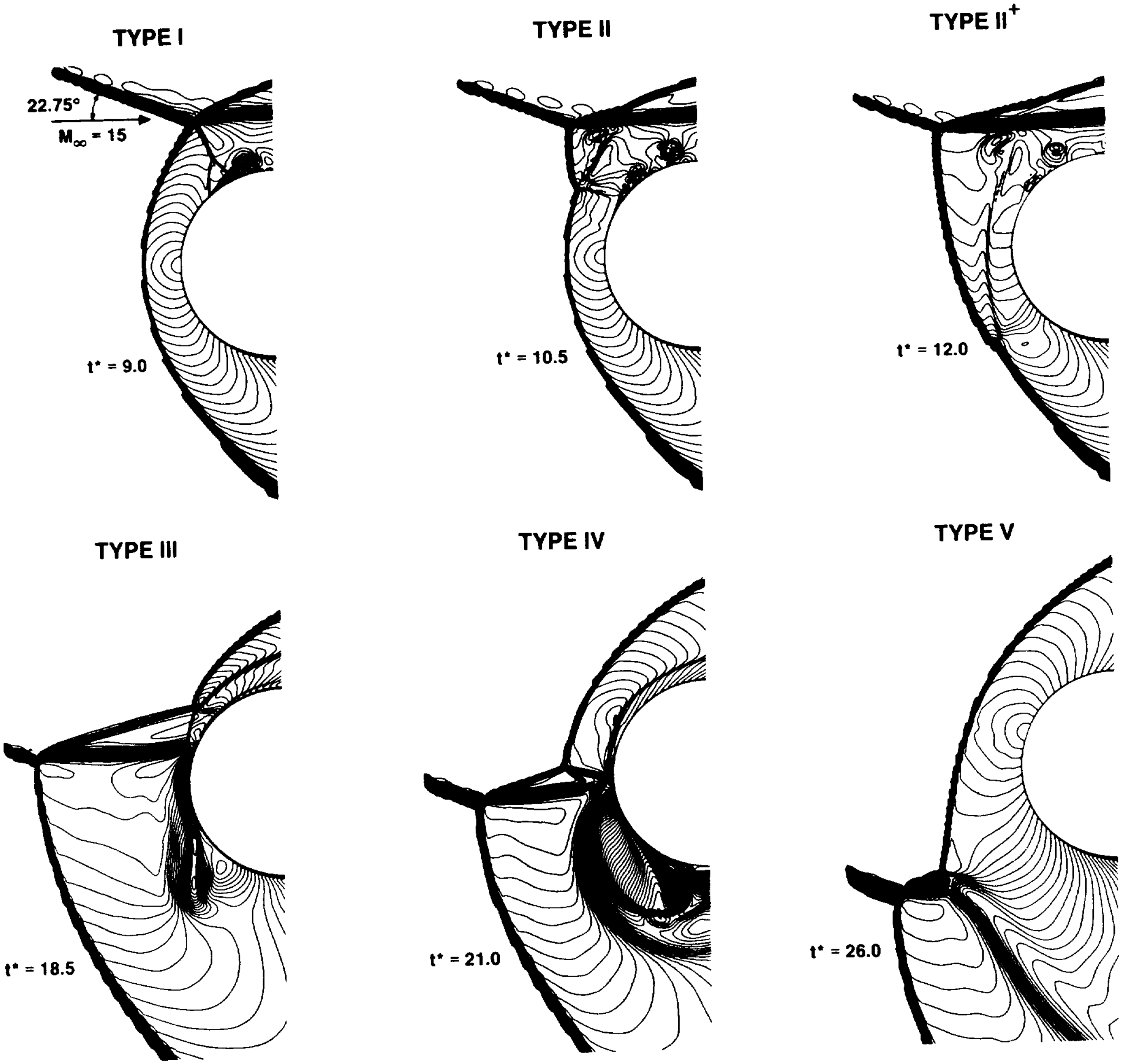


Fig. 6.14 The Mach contours of a two-dimensional unsteady viscous hypersonic **perfect** gas computation by algorithm (6.5) ($\theta = 1$, $\omega = 1/2$, full matrix form) with $M_\infty = 15$ and $Re_D = 186,000$. The indicated time (t^*) of each frame are normalized with the freestream velocity and the cowl lip thickness.

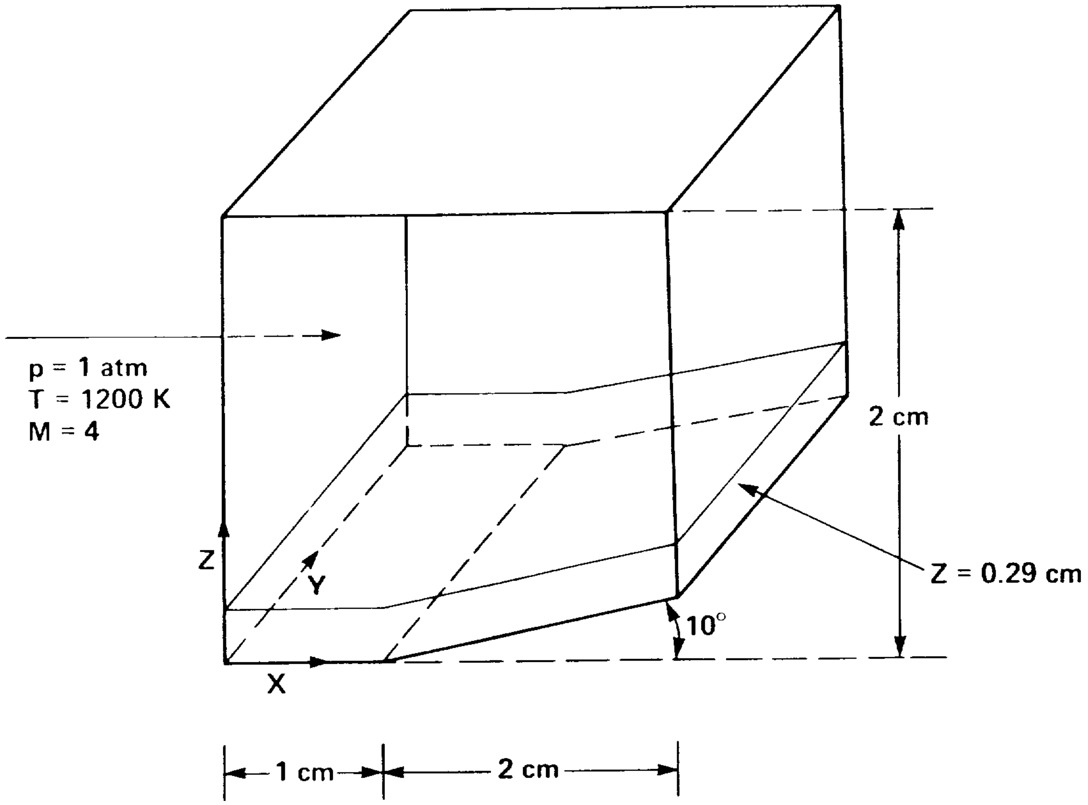


Fig. 7.1 Geometry and inflow condition.

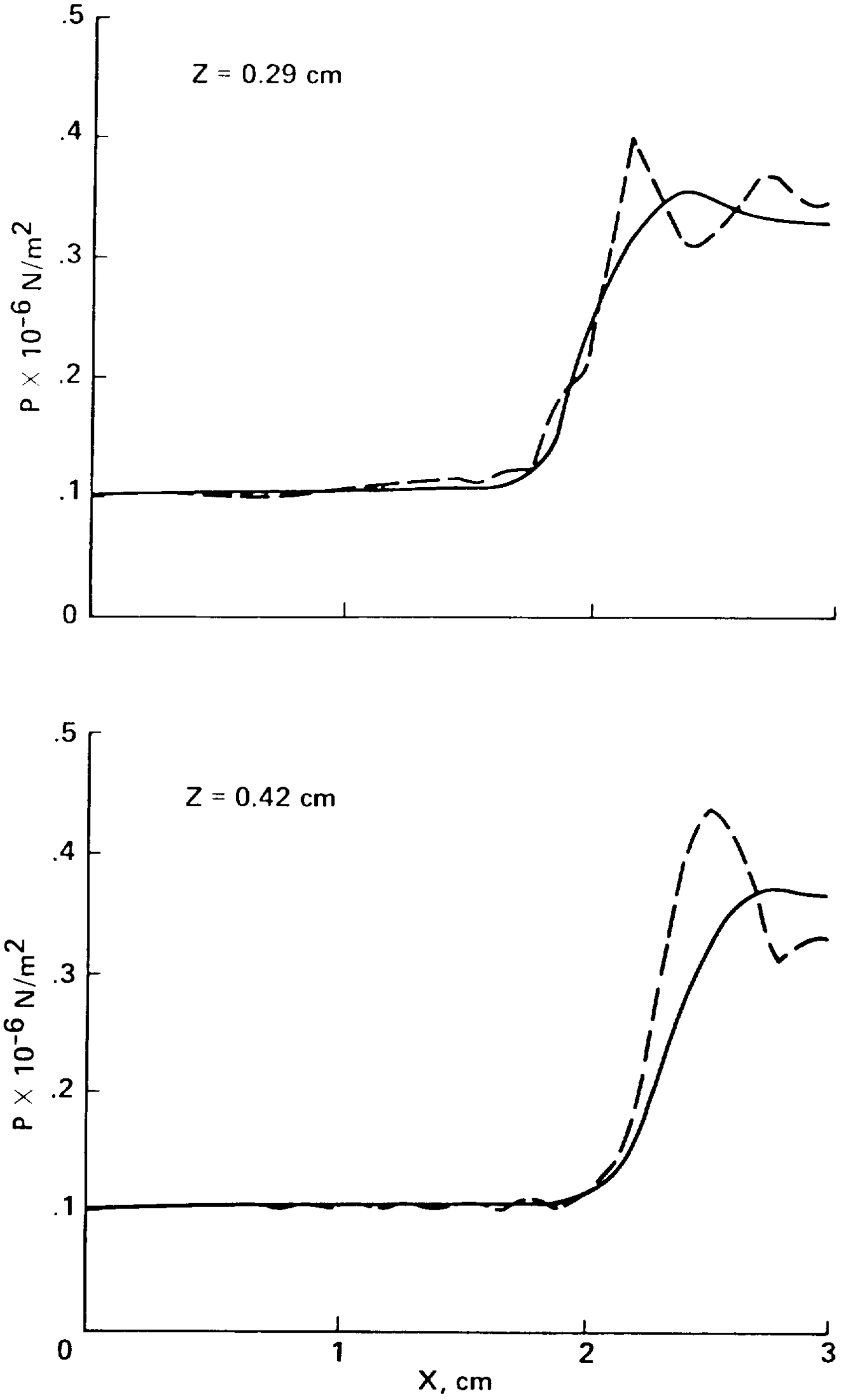


Fig. 7.2. Pressure profiles along $z = 0.29$ and $z = 0.42$ cm.
— Semi-implicit TVD method
- - Classical shock-capturing method

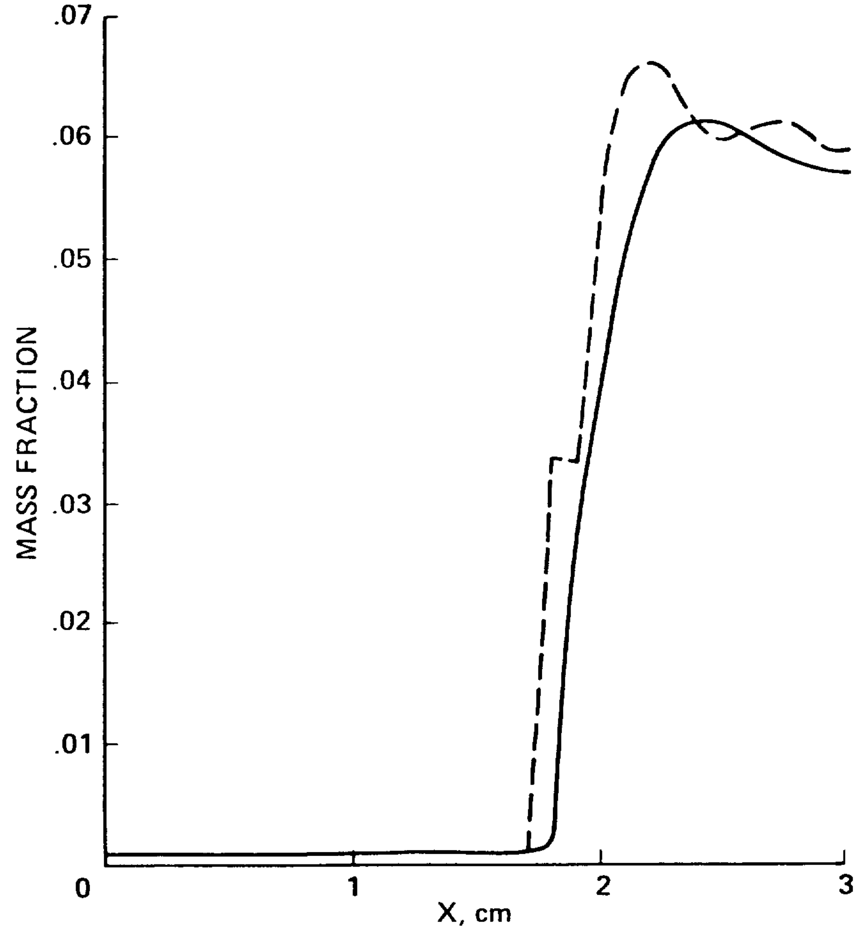
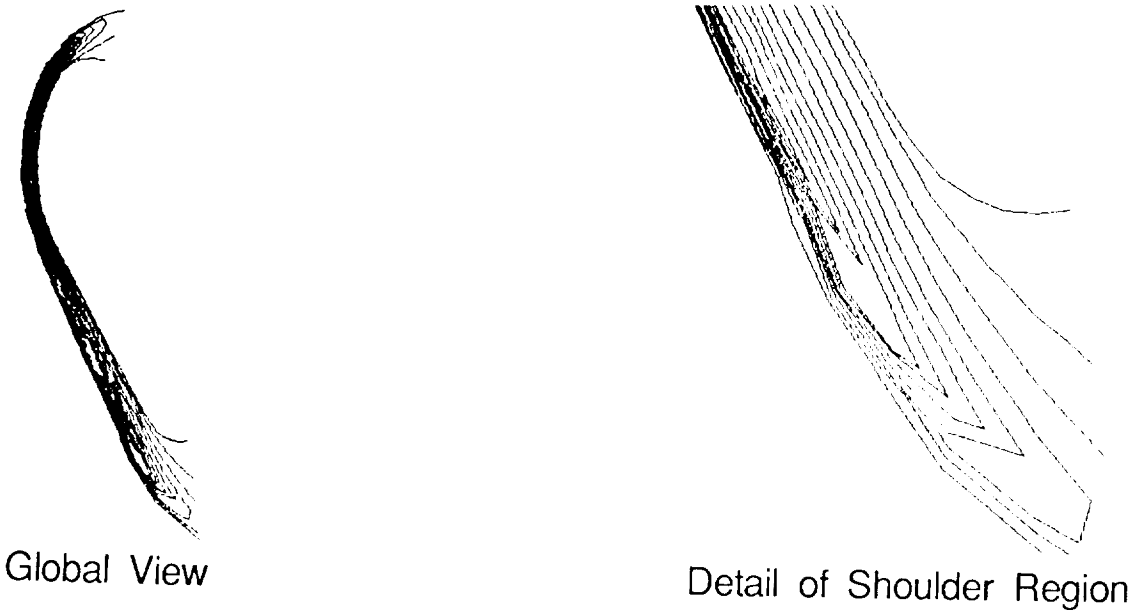


Fig. 7.3. Species distribution for OH along $z = 0.29$ cm.
— Semi-implicit TVD method
- - Classical shock-capturing method



Global View

Detail of Shoulder Region

Fig. 7.4. Contour plot of electron number density over aerobrake at maximum dynamic pressure ($M_\infty = 32$).

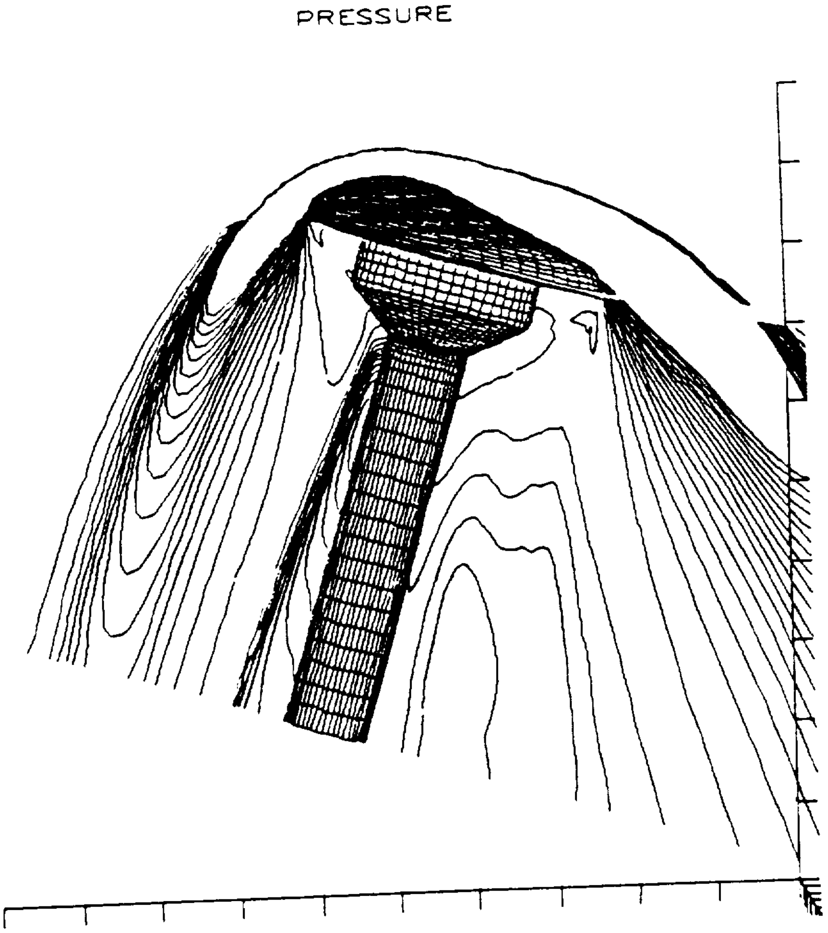


Fig. 7.5 Pressure contours in the near wake region of Aeroassist Flight Experiment (AFE) model, including sting at Mach 10.

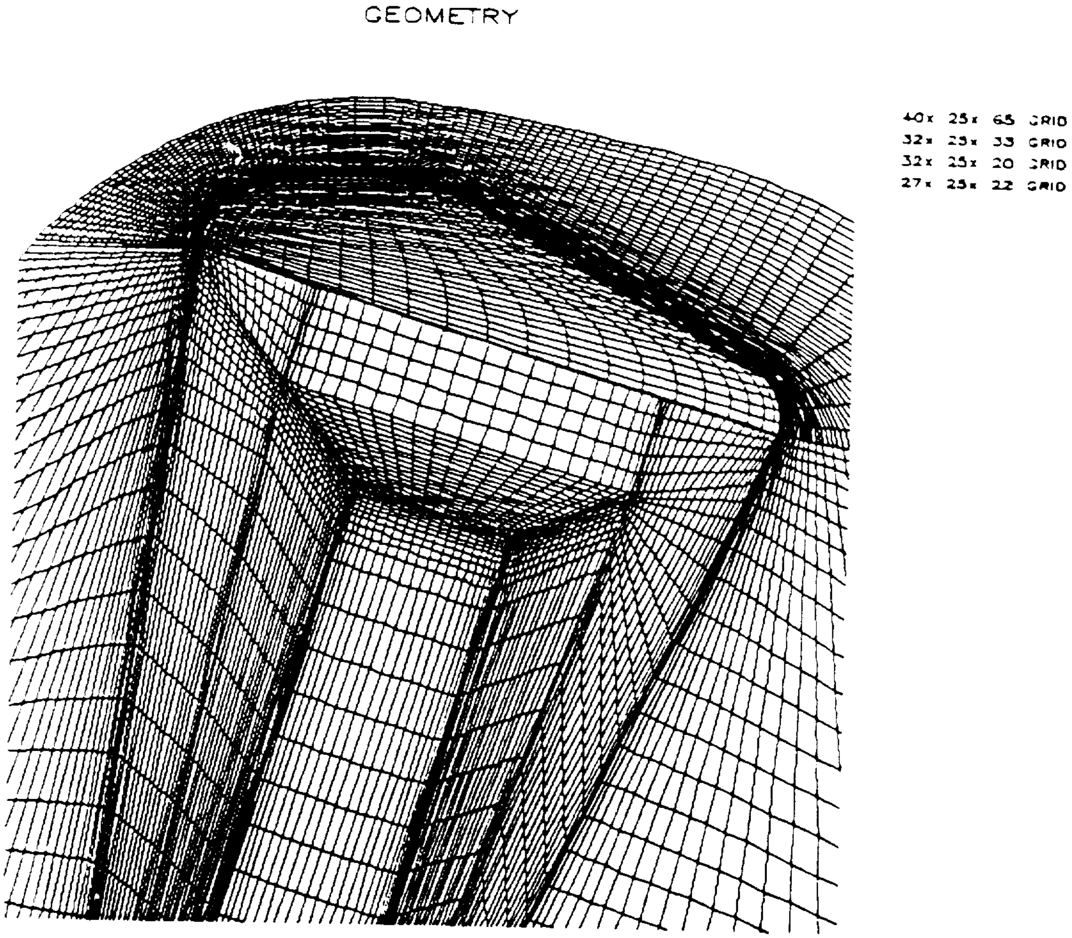
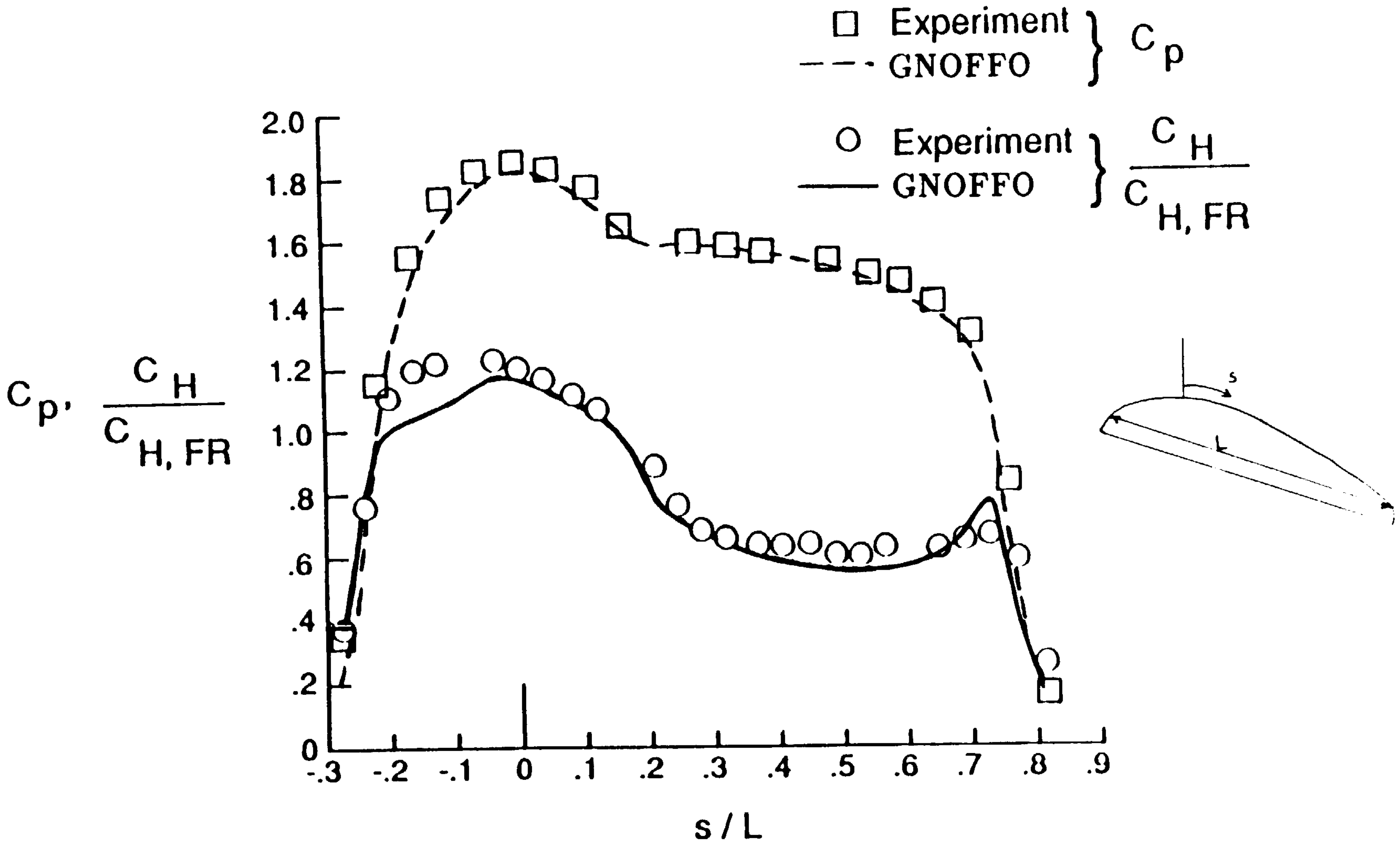
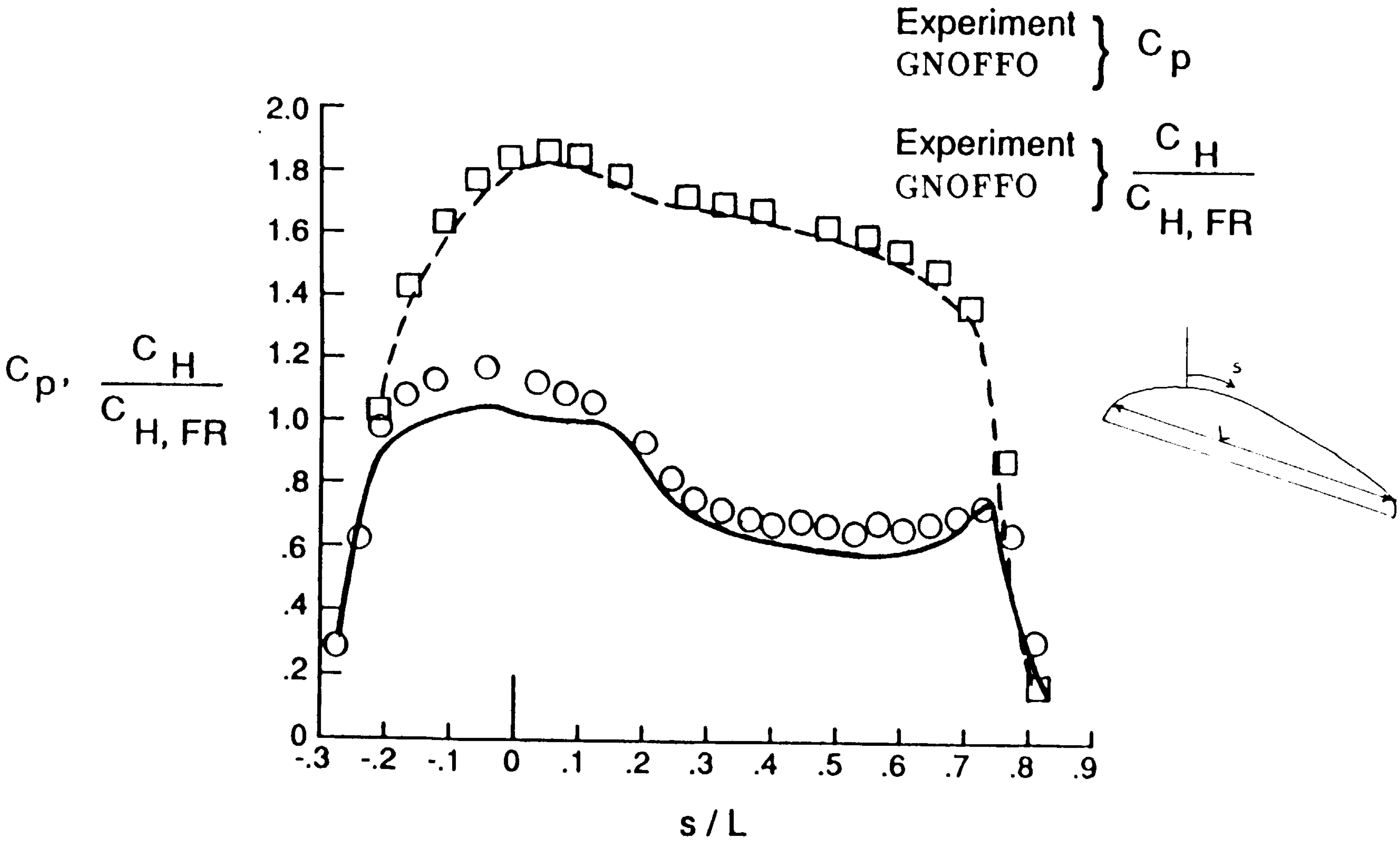


Fig. 7.6 Multidomain grid defining the surface and plane of symmetry for AFE test. The domains are approximately divided into forebody, outer wake, shear layer behind the shoulder, and inner wake core surrounding the sting.



(a) α = 0 degrees

Fig. 7.7 Comparisons between prediction and experiment for pressure and heating over the AFE at Mach 10.



(b) α = 5 degrees

Fig. 7.7 Concluded

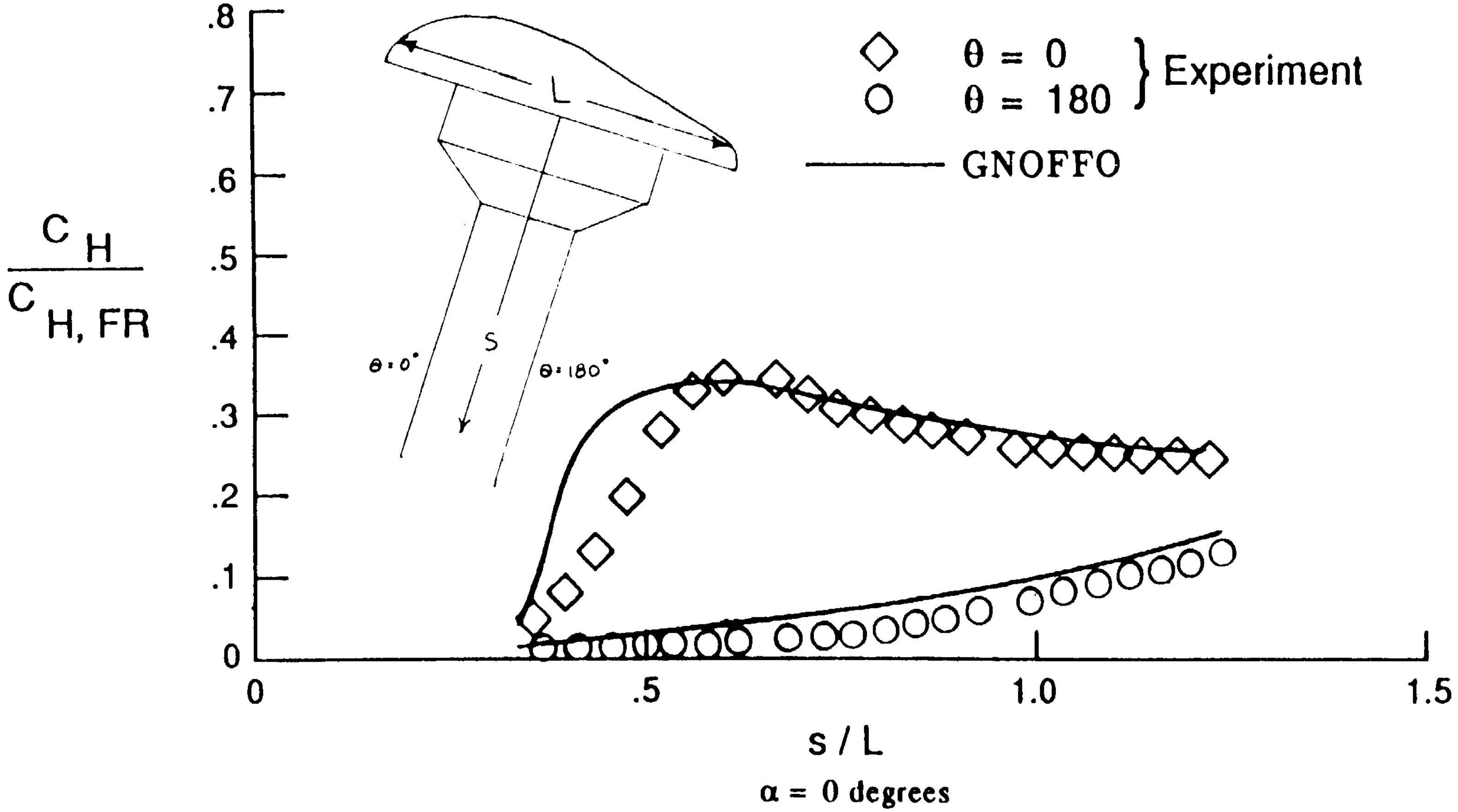


Fig. 7.8 Comparisons between prediction (laminar) and experiment for heating on sting supporting the AFE model at Mach 10.

TRANSLATIONAL TEMPERATURE

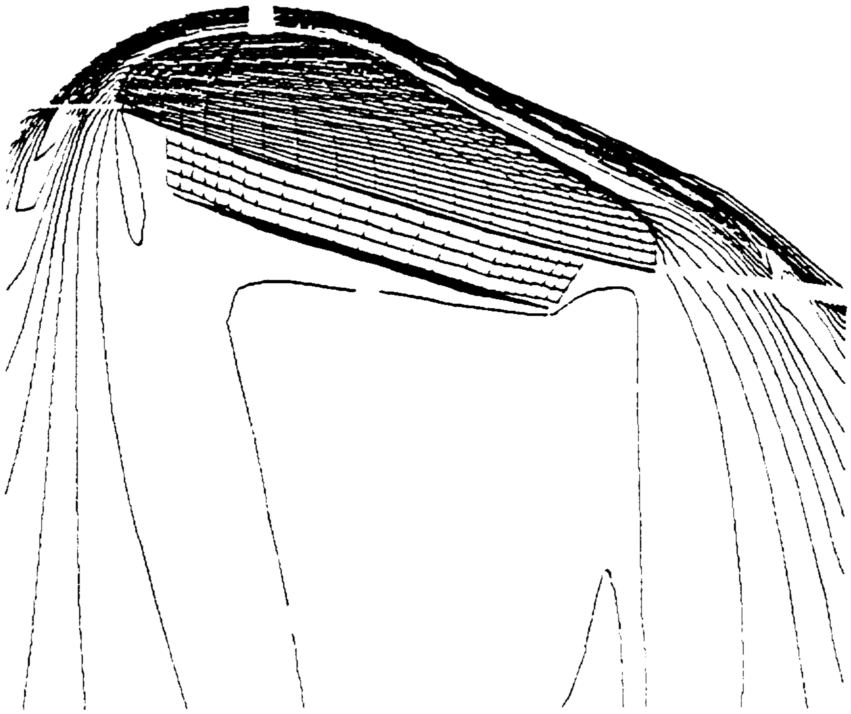


Fig. 7.9 Translational temperature contours for flow over AFE at 8917 m/s and 78 km altitude.

VIBRATIONAL TEMPERATURE

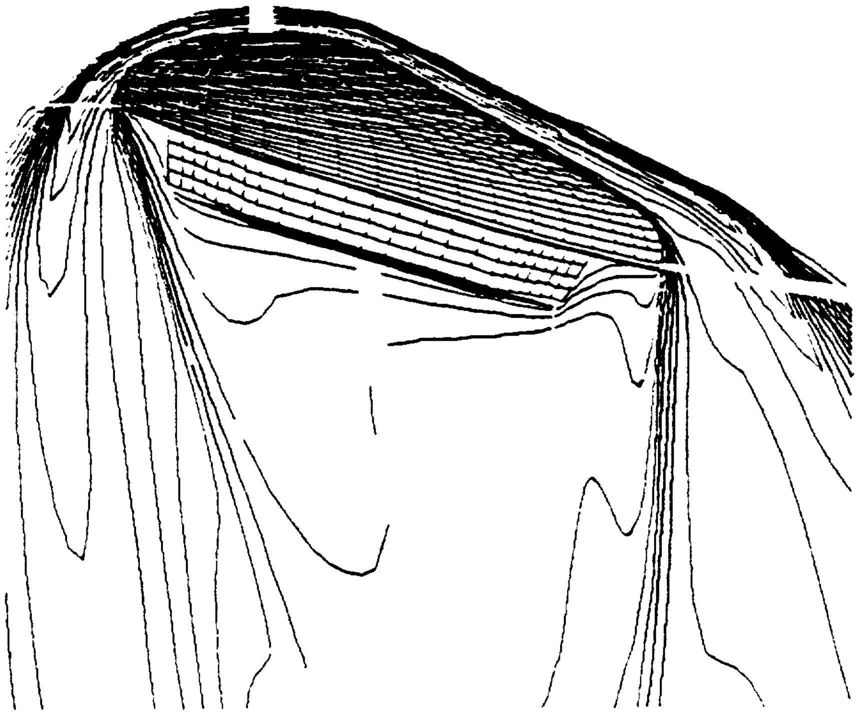


Fig. 7.10 Vibrational temperature contours for flow over AFE at 8917 m/s and 78 km altitude.

MACH NUMBER

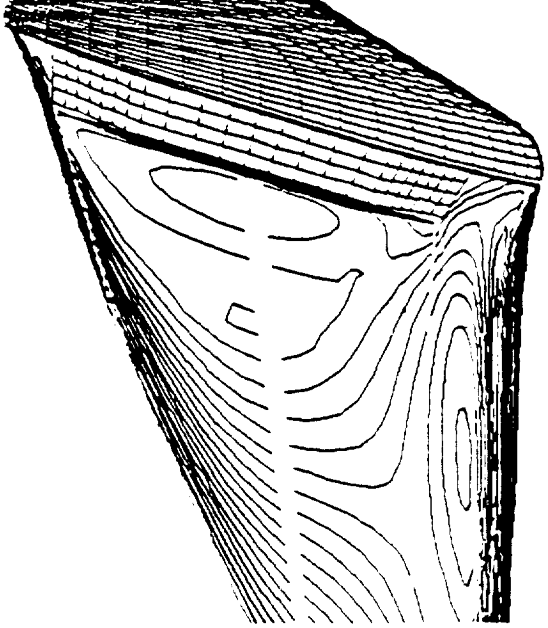


Fig. 7.11 Mach number contours showing the subsonic region for flow over AFE at 8917 m/s and 78 km altitude.

* U.S.GPO:1989-686-006/13901

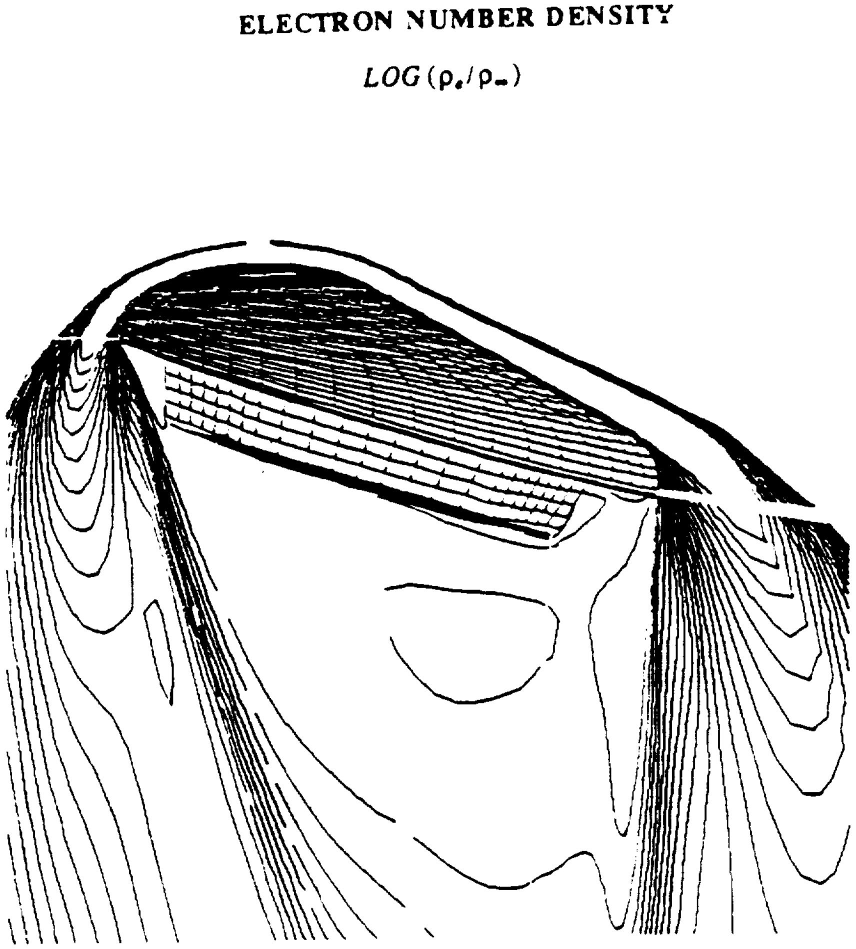


Fig. 7.12 Electron number density contours for flow over AFE at 8917 m/s and 78 km altitude.

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.
NASA TM-101088

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
A Class of High Resolution Explicit and Implicit Shock-Capturing Methods

5. Report Date
February 1989

6. Performing Organization Code

7. Author(s)
H. C. Yee

8. Performing Organization Report No.
A-89091

10. Work Unit No.
505-60

9. Performing Organization Name and Address
Ames Research Center
Moffett Field, CA 94035

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
National Aeronautics and Space Administration
Washington, DC 20546-0001

13. Type of Report and Period Covered
Technical Memorandum

14. Sponsoring Agency Code

15. Supplementary Notes

Point of Contact: H. C. Yee, Ames Research Center, MS 202A-1, Moffett Field, CA 94035
(415) 694-4769 or FTS 464-4769

16. Abstract

The development of shock-capturing finite difference methods for hyperbolic conservation laws has been a rapidly growing area for the last decade. Many of the fundamental concepts, state-of-the-art developments and applications to fluid dynamics problems can only be found in meeting proceedings, scientific journals and internal reports. This paper attempts to give a unified and generalized formulation of a class of high-resolution, explicit and implicit shock-capturing methods, and to illustrate their versatility in various steady and unsteady complex shock wave computations. Included is a systematic overview of the basic design principle of the various related numerical methods. Special emphasis will be on the construction of the basic nonlinear, spatially second- and third-order schemes for nonlinear scalar hyperbolic conservation laws and the methods of extending these nonlinear scalar schemes to nonlinear systems via the approximate Riemann solvers and flux-vector splitting approaches. Generalization of these methods to efficiently include equilibrium real gases and large systems of nonequilibrium flows will be discussed. Some issues concerning the applicability of these methods that were designed for homogeneous hyperbolic conservation laws to problems containing stiff source terms and shock waves are also included. The performance of some of these schemes is illustrated by numerical examples for one-, two- and three-dimensional gas-dynamics problems. The use of the Lax-Friedrichs numerical flux to obtain high-resolution shock-capturing schemes is generalized. This method can be extended to nonlinear systems of equations without the use of Riemann solvers or flux-vector splitting approaches and thus provides a large savings for multidimensional, equilibrium real gases and nonequilibrium flow computations.

17. Key Words (Suggested by Author(s))
Numerical method, Finite difference method, Computational fluid dynamics, System of hyperbolic conservation laws, Stiff problems, Shock capturing, Conservative differencing TVD schemes, Implicit methods

18. Distribution Statement
Unclassified-Unlimited

Subject Category - 64

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of pages
222

22. Price
A10

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