JOURNAL OF COMPUTATIONAL PHYSICS 122, 51–67 (1995)

Runge–Kutta Methods for Hyperbolic Conservation Laws with Stiff Relaxation Terms

SHI JIN

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332

Received February 18, 1994; revised February 22, 1995

Underresolved numerical schemes for hyperbolic conservation laws with stiff relaxation terms may generate unphysical spurious numerical results or reduce to lower order if the small relaxation time is not temporally well-resolved. We design a second-order Runge–Kutta type splitting method that possesses the discrete analogue of the continuous asymptotic limit, which thus is able to capture the correct physical behaviors with high order accuracy, even if the initial layer and the small relaxation time are not numerically resolved. © 1995 Academic Press, Inc.

1. INTRODUCTION

Hyperbolic systems with relaxations occur in the study of a variety of physical phenomena, for example, in linear and nonlinear waves [42, 36], in relaxing gas flow with thermal and chemical nonequilibrium [41, 9], in kinetic theory of rarefied gas dynamics [6], in viscoelasticity [33], and in multiphase and phase transitions [15, 38]. These problems can be described mathematically by the system of evolutional equations

$$\partial_t U + \nabla \cdot F(U) = \frac{1}{\varepsilon} Q(U), \quad U \in \mathbb{R}^N. \tag{1.1}$$

We will call this system the *relaxation system*. Here we use the term relaxation in the sense of Whitham [42] and Liu [29] to denote the relaxation term $Q(U)$ that determines uniquely the local equilibria $U = \mathscr{E}(u)$ for n ($n < N$) independent conserved quantities u. ε is called the *relaxation time*. As $\varepsilon \to 0$, u formally satisfies an $n \times n$ *equilibrium system*,

$$\partial_t u + \nabla \cdot f(u) = 0. \tag{1.2}$$

A system of conservation laws with relaxation is stiff when ε is small compared to the time scale determined by the characteristic speeds of the system and some appropriate length scales.

Theoretical study for these relaxation problems began by Whitham for linear problems [42]. For nonlinear hyperbolic systems of two equations the stability of the equilibrium equation and the zero relaxation limit were proved by Liu [29] and Chen, Levermore, and Liu [7] under an interlace condition between the characteristic speeds of the relaxation system and those of the equilibrium system. Such an interlace condition was referred to as the *subcharacteristic condition* by Liu [29].

We are interested in high order numerical methods for the *stiff* relaxation system (1.1). In particular we would like to investigate the possibility of obtaining the macroscopic behavior described by the equilibrium system (1.2) by solving the original relaxation system (1.1) with coarse grids (Δt, $\Delta x \gg \varepsilon$). Short of resolution of the small relaxation time ε, this approach is usually referred to as the *underresolved* numerical method. Of course one can just solve directly the equilibrium system, which may often be a simplification. However, in many circumstances, the relaxation time varies from order one to much smaller than unity. There it is usually impossible to split the problem into separated regimes and solve directly the equilibrium system in the stiff regime. The appearance of a wide range of relaxation time occurs in, for example, relaxing gases [9] and the hypersonic computations in reentry problems [12]. Thus one has to use *one* system, i.e., the original relaxation system, in the whole domain. Here, as a first, yet difficult and crucial, step toward developing a scheme that works for all ranges of the relaxation time, we will focus on the *stiff* regime. In this regime, reasonable schemes should allow the usage of time and spatial increments that are much bigger than the small relaxation time ε.

We call numerical schemes for the stiff relaxation system (1.1) *robust* in the following sense:

i. They should have a stability constraint independent of the small relaxation time. The Courant–Friedrichs–Lewy (CFL) number should be determined solely by the nonstiff convection part.

ii. They should be modern, high resolution shock-capturing and can properly handle the discontinuous features of the problem, yielding correct shock location and speed without numerical oscillations.

iii. They should give the correct macroscopic behavior with high order accuracy by using coarse grids that do not resolve the small relaxation time ε.

0021-9991/95 $12.00

Usually in a stiff source problem one can overcome the severe stability constraint by using implicit source terms during the time integration. In doing so one can expect a scheme with a CFL number independent of the small relaxation time ε; i.e., the CFL number will depend solely on the convection part. Since the only stiffness appears in the source term, it is very natural to use explicit convection terms [43]. Therefore, numerical stability is not the central issue here. Critical to hyperbolic systems with stiff source terms is that the underresolved numerical methods, although stable, may yield spurious numerical solutions that are totally unphysical. High order schemes may also reduce to lower order when the mesh fails to resolve the small relaxation time.

In this article we implement a second-order Godunov scheme (the MUSCL scheme by van Leer [40]) for the stiff relaxation system (1.1) under the subcharacteristic condition. The choice of MUSCL is not essential here, for other high resolution methods, such as the PPM method of Collela and Woodward [11] or the ENO scheme of Harten, Engquist, Osher, and Chakravarthy [19] may also serve our purpose. Here the Godunov scheme is meant for the convection part of the relaxation sytem only, and the Riemann problem does not take the effect of the source term into account. A method of line approach is considered here, combined with a Runge–Kutta method for time marching. The purpose of this paper is to show how a splitting second-order time discretization can be done to obtain a robust shock capturing method in the sense described above.

Earlier Pember studied similar problems in [30, 31]. Our understanding of this problem is that poor numerical results may be generated if the numerical scheme does not have the *correct asymptotic limit.* A scheme for the relaxation system (1.1) is said to have the correct asymptotic limit if for fixed Δx and Δt, as $\varepsilon \to 0$, the limiting scheme is a good (consistent, stable, and high order) discretization of the equilibrium system (1.2). In other words, the numerical scheme possesses a discrete analogy of the asymptotic limit that leads from the relaxation system (1.1) to the equilibrium system (1.2). We illustrate our idea through a model relaxation system to be specified below, and design a second-order time integration that works for the general relaxation systems defined in the beginning of this section.

The stiff source problem also arises in the computations of reacting flows. There the smeared numerical shock profile may trigger the reaction to the wrong equilibrium, thus causing incorrect shock speed [10, 28]. Various numerical methods are suggested in the literature, which require some sort of resolution in the reacting front [1, 2, 14, 17, 18, 39]. The stiff source terms in these problems have both stable and unstable local equilibria; thus they have essential differences from the relaxation systems we study here.

An important class of relaxation problems lies in the kinetic theory of rarefield gas dynamics. There the relaxation describes the interactions of particles and the relaxation time is the mean free path. When the mean free path is small, the kinetic equation approximates the compressible Euler or Navier–Stokes equations, known as the fluid dynamic limit. Numerical simulations of kinetic equations with small mean free path lead to the development of Boltzmann schemes or kinetic schemes for the compressible Euler equations [20, 32] that do not use the solution of the Riemann problem.

The correct asymptotic limit analysis was applied earlier in the literature. It was used to study and develop numerical schemes for the neutron transport equation in diffusive regimes [25, 26, 21, 22]. The diffusive behavior of spatially underresolved, semidiscrete high order Godunov schemes for hyperbolic systems with stiff relaxation terms was studied in [23], applying a combination of the correct asymptotic limit analysis and the modified equation analysis. The underresolved numerical method was also studied for hyperbolic systems in oscillatory fields; see, for example, [13].

We choose to analyze in detail the numerical discretizations of a prototypical relaxation model [7]

$$\partial_t h + \partial_x w = 0 \tag{1.3a}$$

$$\partial_t w + \partial_x p(h) = -\frac{1}{\varepsilon}(w - f(h)), \quad \varepsilon > 0, p'(h) > 0. \tag{1.3b}$$

This system is hyperbolic with two distinct real characteristics speeds $\pm\sqrt{p'(h)}$. The positive parameter ε is the relaxation time for the system. The relaxation term is stiff when $\varepsilon \ll 1$; that is, the relaxation time is much shorter than the time it takes for a hyperbolic wave (sound wave) to propagate over a gradient length. The leading term approximation to Eqs. (1.3) is

$$w = f(h), \tag{1.4a}$$

$$\partial_t h + \partial_x f(h) = 0. \tag{1.4b}$$

By looking for the $O(\varepsilon)$ correction to the approximation (1.4), one obtains a dissipative evolution equation [7],

$$w = f(h) - \varepsilon(p'(h) - f'(h)^2)\partial_x h, \tag{1.5a}$$

$$\partial_t h + \partial_x f(h) = \varepsilon\partial_x((p'(h) - f'(h)^2)\partial_x h), \tag{1.5b}$$

provided that the characteristic speed $f'(h)$ interlaces with those of system (1.3),

$$-\sqrt{p'(h)} \le f'(h) \le \sqrt{p'(h)}.$$

This is the subcharacteristic condition of Liu [29] for (1.3).

The asymptotic expansion here is analogous to the Chapman–Enskog expansion in rarefied gas dynamics modeled by the nonlinear Boltzman equation close to its fluid dynamic limit when the mean free path is small [5, .6]. Adopting the terminology of kinetic theory, the leading term approximation (1.4) is referred to as the Euler limit, while approximation (1.5) is usually called the Navier–Stokes limit. Equation (1.5a) can

be called the local Maxwellians or local equilibria. As the Chapman–Enskog expansion is formal in the sense that it may not be valid when the solution is near regions with large gradients, our numerical asymptotic analysis is only valid when the solution is smooth.

In Section 2 we perform a detailed initial layer analysis for (1.3). The result indicates that the initial layer projects the initial data to the local equilibrium. This information is needed since we want a scheme that does not resolve the initial layer. In Section 3 we begin our study with the first-order splitting method and Strang's splitting method. First we show that, by doing the first time step fully implicitly, the scheme automatically projects the initial data into the local equilibrium, thus the scheme does not need to resolve the initial layer nor to preprocess the initial data. We then show that the Strang splitting may fail to maintain its second-order accuracy as $\varepsilon \to 0$, thus it does not have a good limit when the mesh does not resolve ε. A second-order splitting scheme is developed which combines the high order Godunov schemes with an implicit ODE solver in a second-order total-variation-diminishing (TVD) Runge–Kutta formulation. This scheme is robust in the sense described above. In contrast to the conclusion of Pember in [31], where he conjectures that unsplit schemes must be used for the relaxation system, our analysis indicates that it is not the splitting that causes the spurious or poor solutions. Rather, any schemes, split or unsplit, violating the correct asymptotic limit lead to spurious or poor solutions. In Section 4 we show some numerical examples that seem to agree with our analysis.

Although the analysis and experiment are carried out on the model problem (1.3), the result extends far beyond this model. In Section 5 we apply the new splitting scheme developed in Section 3.5 to two more general relaxation systems, including the Broadwell model of the rarefied gas dynamics, and the Eulerian gas dynamics with heat transfer. Numerical results show that for these problems the new splitting scheme does give the correct equilibrium behavior without resolving the small relaxation time. Since our analysis concentrates on the time integrator, which is dimension independent, thus it also works for higher dimensional problems [24].

2. THE INITIAL LAYER ANALYSIS

Since the underresolved numerical schemes exhibit spurious behavior in the presence of the initial layer, such as the incorrect local equilibria and the wrong shock location, which do not appear if there is no initial layer, it is important to understand the initial layer behavior of the relaxation system. In this section we perform an initial layer analysis on the model system (1.3). The analysis here is in analogy to the similar analysis performed by Caflisch and Papanicolaou ([5]) on the Broadwell model of the Nonlinear Boltzmann Equation close to its fluid dynamic limit when the mean free path is small.

Introducing a stretched time variable

$$\tau = t/\varepsilon$$

and considering h and w as functions of τ and x, Eq. (1.3) then takes the form

$$\frac{1}{\varepsilon}\partial_\tau h + \partial_x w = 0, \tag{2.1a}$$

$$\frac{1}{\varepsilon}\partial_\tau w + \partial_x p(h) = -\frac{1}{\varepsilon}(w - f(h)), \tag{2.1b}$$

with initial conditions

$$h(0, x) = h^I(x), \quad w(0, x) = w^I(x).$$

We look for an expansion such that

$$h = h_0(t, x) + \varepsilon h_1(t, x) + \cdots + h_0^{IL}(\tau, x) + \varepsilon h_1^{IL}(\tau, x) + \cdots, \tag{2.2a}$$

$$w = w_0(t, x) + \varepsilon w_1(t, x) + \cdots + w_0^{IL}(\tau, x) + \varepsilon w_1^{IL}(\tau, x) + \cdots, \tag{2.2b}$$

where $h_0 + \varepsilon h_1$ and $w_0 + \varepsilon w_1$ are functions already determined by (1.3) up to the initial condition and an $O(\varepsilon^2)$ error.

We insert (2.2) into (2.1):

$$\frac{1}{\varepsilon}\partial_\tau[h_0(t, x) + \varepsilon h_1(t, x) + \cdots + h_0^{IL}(\tau, x) + \varepsilon h_1^{IL}(\tau, x) + \cdots] + \partial_x[w_0(t, x) + \varepsilon w_1(t, x) + \cdots + w_0^{IL}(\tau, x) + \varepsilon w_1^{IL}(\tau, x) + \cdots] = 0,$$

$$\frac{1}{\varepsilon}\partial_\tau[w_0(t, x) + \varepsilon w_1(t, x) + \cdots + w_0^{IL}(\tau, x) + \varepsilon w_1^{IL}(\tau, x) + \cdots] + \partial_x p(h_0(t, x) + \varepsilon h_1(t, x) + \cdots + h_0^{IL}(\tau, x) + \varepsilon h_1^{IL}(\tau, x) + \cdots) + \frac{1}{\varepsilon}[w_0(t, x) + \varepsilon w_1(t, x) + \cdots + w_0^{IL}(\tau, x) + \varepsilon w_1^{IL}(\tau, x) + \cdots - f(h_0(t, x) + \varepsilon h_1(t, x) + \cdots + h_0^{IL}(\tau, x) + \varepsilon h_1^{IL}(\tau, x) + \cdots)] = 0.$$

Equating to zero coefficients of equal powers of ε gives the following equations for the initial layer terms:

$$\partial_\tau h_0^{IL} = 0, \tag{2.3a}$$

$$\partial_\tau w_0^{IL} + w_0 + w_0^{IL} - f(h_0 + h_0^{IL}) = 0; \tag{2.3b}$$

$$\partial_\tau h_1^{IL} + \partial_t h_0 + \partial_x w_0 + \partial_x w_0^{IL} = 0, \tag{2.4a}$$

$$\partial_\tau w_1^{IL} + \partial_t w_0 + \partial_x p(h_0 + h_0^{IL}) + w_1 + w_1^{IL} - f'(h_0 + h_0^{IL})(h_1 + h_1^{IL}) = 0. \tag{2.4b}$$

In these equations all the terms of the interior expansion appearing on the right side are evaluated at $t = 0$ after the indicated operations have been performed.

We attempt now to solve (2.3) and (2.4) recursively so that $h_k^{IL}(\tau, x)$ and $w_k^{IL}(\tau, x)$ decay to zero as $\tau \to \infty$ uniformly along x derivatives for $k = 0, 1, \ldots$, and

$$h_0(0, x) + h_0^{IL} = h^I(x), \tag{2.5a}$$

$$w_0(0, x) + w_0^{IL} = w_I(x); \tag{2.5b}$$

$$h_k(0, x) + h_k^{IL} = 0, \quad k = 1, 2, \ldots, \tag{2.5c}$$

$$w_k(0, x) + w_k^{IL} = 0, \quad k = 1, 2, \ldots. \tag{2.5d}$$

LEMMA 2.1. *Let $w_0(0, x) = f(h_0(0, x))$ be the local equilibrium to the leading order and $h_0^{IL}(0, x) = 0$. Then the nonlinear system of ordinary differential equations* (2.3) *has a unique solution, exponentially decaying as $\tau \to \infty$, uniformly in x, with the initial condition* (2.5). *Moreover, x derivatives of the solution also decay exponentially as $\tau \to \infty$, uniformly in x.*

Proof. Since $h_0^{IL}(0, x) = 0$, one immediately obtains from Eq. (2.3) that

$$h_0^{IL}(\tau, x) = 0, \quad \text{for any } \tau \geq 0. \tag{2.6}$$

Thus (2.5a) gives

$$h_0(0, x) = h^I(x).$$

Applying (2.6) and the assumption $w_0(0, x) = f(h_0(0, x))$ in (2.3b) then gives

$$\partial_\tau w_0^{IL} = -w^{IL}. \tag{2.7}$$

The existence of a unique exponential decay solution w_0^{IL} is obvious from (2.7). The statement about x-derivative follows similarly after differentiating (2.7) with respect to x which is just a parameter here. The proof of the lemma is complete. ∎

Remark. Lemma 1 and (2.5) imply

$$h_0(0, x) = h^I(x), \quad h_0^{IL}(\tau, x) \equiv 0; \tag{2.8a}$$

$$w_0(0, x) = f(h^I(x)), \quad w_0^{IL} = w^I(x) - f(h^I(x)). \tag{2.8b}$$

With Lemma 1, the leading term interior approximation $\partial_t h_0 = -\partial_x f(h_0)$ from (1.2a), and (2.5), we can now reduce Eq. (2.4) to

$$\partial_\tau h_2^{IL} = -\partial_x w_0^{IL}, \tag{2.9a}$$

$$\partial_\tau w_1^{IL} = -w_1^{IL} + f'(h_0)^2 \partial_x h_0 - p'(h_0)\partial_x h_0 - w_1 + f'(h_0)(h_1 + h_1^{IL}). \tag{2.9b}$$

LEMMA 2.2. *Let $w_1(0, x) = f'(h_0(0, x))^2 \partial_x h_0(0, x) - p'(h_0(0, x))\partial_x h_0(0, x)$ and $h_1(0, x) = 0$. Then the nonlinear system of ordinary differential equations* (2.4) (*or equivalently* (2.9)) *has a unique solution exponentially decaying as $\tau \to \infty$, uniformly in x, with the initial condition* (2.5). *Moreover, x derivatives of the solution also decay exponentially as $\tau \to \infty$, uniformly in x.*

Proof. First, the exponential decay of h_1^{IL} and its x-derivatives can be easily seen from (2.9a) by the exponential decay of w_0^{IL} and its x-derivatives. Now, submitting $w_1(0, x) = f'(h_0(0, x))\partial_x w_0(0, x) - p'(h_0(0, x))\partial_x h_0(0, x)$ and $h_1(0, x) = 0$ in (2.9b) gives

$$\partial_\tau w_1^{IL} = f'(h_0) h_1^{IL} - w_1^{IL}. \tag{2.10}$$

This ordinary differential equation clearly has a unique exponential decay solution since h_1^{IL} has exponential decay. That x-derivatives of w_1^{IL} have exponential decay is also trivial by taking derivatives with respect to x on Eq. (2.10), using the fact that the x-derivatives of h_1^{IL} also have exponential decay. This completes the proof for the lemma. ∎

Remark. Lemma 2 and (2.8) give the initial layer solution,

$$w_1(0, x) = (f'(h^I(x))^2 - p'(h^I(x)))\partial_x h^I(x). \tag{2.11}$$

Higher-order terms and their initial conditions are determined similarly but are omitted here since the interior expansion (1.3) is only valid to $O(\varepsilon)$. By combining (2.8) with (2.11) we obtain the initial conditions for the relaxation equation (1.3) as

$$h(0, x) = h^I(x),$$

$$w(0, x) = f(h^I(x)) + \varepsilon[f'(h^I(x))^2 - p'(h^I(x))]\partial_x h^I(x).$$

By comparing with (1.3a) one sees that the initial layer projects the initial data to the local equilibria.

Remark. Similar initial layer analysis can be carried out for the more general relaxation system (1.1) and a similar conclusion may be drawn.

3. THE NUMERICAL DISCRETIZATIONS

We introduce the spatial grid points $x_{j+1/2}$, $j = \ldots, -1, 0, 1, \ldots$ with uniform mesh spacing $\Delta x = x_{j+1/2} - x_{j-1/2}$ for all j. The time level $t_0, t_1, \ldots$ is also spaced uniformly with space step $\Delta t = t^{n+1} - t^n$ for $n = 0, 1, 2, \ldots$. Here the assumption of a uniform grid is only for simplicity. We use U_j^n to denote the cell averages of U in the cell $[x_{j-1/2}, x_{j+1/2}]$ at time t^n:

$$U_j^n = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} U(t_n, x)\, dx.$$

Consider the one-dimensional relaxation system

$$\partial_t U + \partial_x F(U) = \frac{1}{\varepsilon} Q(U).$$

A spatial discretization in conservation form can be written as

$$\partial_t U_j + \frac{1}{\Delta x}(F_{j+1/2} - F_{j-1/2}) = \frac{1}{\varepsilon} Q_j,$$

where the numerical flux $F_{j+1/2}$ is to be defined in terms of the known cell-average numerical quantities, U_j's, and the averaged source term is defined by

$$Q_j = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} Q(U)\,dx = Q\left(\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} U\,dx\right) + O(\Delta x^2)$$
$$= Q(U_j) + O(\Delta x^2).$$

Thus, for sufficiently accurate spatial discretizations we have, with an accuracy of $O(\Delta x^2)$,

$$\partial_t U_j + \frac{1}{\Delta x}(F_{j+1/2} - F_{j-1/2}) = Q(U_j).$$

3.1. *The Shock Capturing Spatial Discretizations*

To define the convection flux $F_{j+1/2}$ we use the high order Godunov scheme of van Leer [40], that is based on Roe's approximate solution of the Riemann problem [34] for the homogeneous hyperbolic system,

$$\partial_t U + \partial_x F(U) = 0.$$

During the reconstruction step slope limiters [40] are applied in order to eliminate unphysical numerical oscillations. Note here that the reconstruction and the Riemann solver do not account for the presence of the source items.

3.2. *The Correct Asymptotic Limit Analysis*

For the relaxation system, it is natural to require that the numerical scheme possess a discrete analogy of the continuous asymptotic limit. Here the asymptotic analysis is defined in the following sense. First, the asymptotic expansion is carried out in terms of ε under the coarse scaling $\Delta x/\Delta t = O(1)$, $\varepsilon/\Delta t \ll 1$. Second, since the Chapman–Enskog expansion for the continuous relaxation system is valid only for a *smooth* solution, in our discrete system we have to impose the same assumption. Thus all the discrete spatial derivatives, including $f'(h)$, are assumed to be $O(1)$. Therefore, unless otherwise specified, we always have $\lambda\Delta_+ = O(\Delta t)$, where the operator Δ_+ is defined by $\Delta_+ V_{j-1/2} = V_{j+1/2} - V_{j-1/2}$ for any vector V. We *do not*, however, assume the time derivative to be $O(1)$. This allows us to determine the effect of the initial layer. Under these assumptions our asymptotic expansion illustrates the numerical behavior only in the smooth region.

A (high order) scheme is said to have the correct asymptotic limit if as $\varepsilon \to 0$, under the above assumptions, the limiting scheme becomes a good (high order) approximation of the equilibrium equations. The initial layer analysis in Section 3 suggests that the initial layer projects the initial data into a local equilibrium. This same projection also leads from the relaxation system to the equilibrium equation away from the initial layer. In order to have the correct asymptotic limit without resolving the initial layer, the numerical scheme should intrinsically have the same mechanism, that is to say, the scheme should project the numerical data, in or not in local equilibrium, into a local equilibrium at *every* time step. Such a projection in the first time step simulates the initial layer behavior without resolving the initial layer. At later times this same projection guarantees the correct numerical passage from the relaxation system to the equilibrium equation. Mathematically such a projection is realized through

$$Q(U^n) \approx 0 \quad \text{for all } n \geq 1,$$

up to some approximation error which is a function of ε and Δt. For the model problem (1.3), this implies

$$W^n - F(H^n) \approx 0 \quad \text{for all } n \geq 1.$$

3.3. *A First-Order Splitting Scheme*

By examining the asymptotics that leads from the relaxation system (1.1) to the equilibrium system (1.2), it is natural to design numerical schemes that simulate the same asymptotics. The simplest way is to use a first-order splitting scheme that combines a backward Euler method for the stiff source term with a forward Euler method for the convection term. It is given by

$$U^{(1)} = U^n + \frac{\Delta t}{\varepsilon} Q(U^{(1)}), \tag{3.1a}$$

$$U^{n+1} = U^{(1)} + \lambda\Delta_+ F^{(1)}_{j-1/2}. \tag{3.1b}$$

Roughly speaking, the first step being fully implicit always gives a projection into a local equilibrium $Q(U) \approx 0$ independently of the initial data. This local equilibrium, after applied to the second step, should give an equilibrium scheme that is consistent to the equilibrium system (1.2). The ODE solver being used here is the backward Euler method, which is both A-stable and L-stable. We show that this scheme has the correct asymptotic limit.

Applying (3.1) to the model problem (1.3), one has

$$H^{(1)} = H^n, \tag{3.2a}$$

$$W^{(1)} = W^n - \frac{\Delta t}{\varepsilon}(W^{(1)} - f(H^{(1)})); \tag{3.2b}$$

$$H^{n+1} = H^{(1)} + \lambda\Delta_+ W^{(1)}_{j-1/2}, \tag{3.2c}$$

$$W^{n+1} = W^{(1)} + \lambda\Delta_+ p^{(1)}_{j-1/2}. \tag{3.2d}$$

We have the following results:

• *If at* $t = t^n$,

$$W^n - f(H^n) = 0, \tag{3.3}$$

then at $t = t^{n+1}$,

$$W^{n+1} - f(H^{n+1}) = O(\Delta t). \tag{3.4}$$

In the intermediate step,

$$W^{(1)} - f(H^{(1)}) = 0. \tag{3.5}$$

If at $t = t^n$ *the solution is not in local equilibrium, then*

$$W^{n+1} - f(H^{n+1}) = O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right). \tag{3.6}$$

In the intermediate step,

$$W^{(1)} - f(H^{(1)}) = O\left(\frac{\varepsilon}{\Delta t}\right). \tag{3.7}$$

Case 1. *The initial data in a local equilibrium.* From (3.2b) and (3.3),

$$\begin{aligned} W^{(1)} - W^n &= -\frac{\Delta t}{\varepsilon}(W^{(1)} - W^n + W^n - f(H^{(1)})) \\ &= -\frac{\Delta t}{\varepsilon}(W^{(1)} - W^n + W^n - f(H^n)) \\ &= -\frac{\Delta t}{\varepsilon}(W^{(1)} - W^n). \end{aligned} \tag{3.8}$$

Thus

$$W^{(1)} - W^n = 0. \tag{3.9}$$

Applying this to (3.2b) implies

$$W^{(1)} - f(H^{(1)}) = 0.$$

Applying (3.9) in (3.2d) gives

$$\begin{aligned} &W^{n+1} - f(H^{n+1}) \\ &\quad = W^{(1)} - f(H^{n+1}) + O(\Delta t) \\ &\quad = W^n - f(H^n) + f(H^n) - f(H^{n+1}) + O(\Delta t) \\ &\quad = O(\Delta t). \end{aligned} \tag{3.10}$$

In the last equality of (3.10) we used $H^{n+1} - H^{(1)} = H^{n+1} - H^n = O(\Delta t)$. Thus (3.4) and (3.5) are true.

Case 2. *The initial data not in a local equilibrium.* Suppose $W^n - f(H^n) = O(1) \neq 0$. Then (3.9) does not hold. Instead, (3.8) only yields

$$W^{(1)} - W^n = -\frac{\Delta t}{\varepsilon}(W^{(1)} - W^n) + O\left(\frac{\Delta t}{\varepsilon}\right);$$

thus

$$W^{(1)} - W^n = \frac{1}{1 + \Delta t/\varepsilon} O\left(\frac{\Delta t}{\varepsilon}\right) = O(1).$$

Applying this in (3.2b) gives

$$W^{(1)} - f(H^{(1)}) = -\frac{\varepsilon}{\Delta t}(W^{(1)} - W^n) = O\left(\frac{\varepsilon}{\Delta t}\right).$$

Applying (3.2b) in (3.2d),

$$\begin{aligned} W^{n+1} - f(H^{n+1}) &= \frac{1}{1 + \Delta t/\varepsilon}\left(W^n + \frac{\Delta t}{\varepsilon} f(H^{(n)})\right) - f(H^{n+1}). \\ &= \frac{\varepsilon}{\Delta t}(W^n - f(H^n)) + f(H^n) - f(H^{n+1}) \\ &\quad + O\left(\left(\frac{\varepsilon}{\Delta t}\right)^2\right) \\ &= O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right). \end{aligned}$$

Hence (3.6) and (3.7) are true.

We have shown that, as long as the solution is bounded (by numerical stability), $W^n - f(H^n)$ will always be $O(\Delta t + \varepsilon/\Delta t)$ by the result (3.6) for general initial data, and at $t = t^{(1)} \in (t^{n-1}, t^{n+1})$ (3.7) is always valid. Thus we have the following conclusion.

• *For any initial data, the splitting scheme* (3.2) *always gives, for any* $n \geq 1$,

$$W^n - f(H^n) = O\left(\left(\Delta t + \frac{\varepsilon}{\Delta t}\right), \tag{3.11}$$

and

$$W^{(1)} - f(H^{(1)}) = O\left(\frac{\varepsilon}{\Delta t}\right), \tag{3.12}$$

for $t^{(1)} \in (t^{n-1}, t^n)$.

Moreover, as $\varepsilon \to 0$, by applying (3.12) in (3.2c), after ignoring the error terms, the splitting scheme (3.2) limits to the equilibrium scheme

$$H^{n+1} = H^n + \lambda \Delta_+ f^n_{j-1/2}, \tag{3.13a}$$

$$W^{n+1} = f(H^{n+1}), \tag{3.13b}$$

where

$$f^n_{j-1/2} = W^n_{j-1/2}\big|_{W^n = f(H^n)}. \tag{3.13c}$$

(3.13a) is clearly the forward Euler method for the equilibrium equation (1.4b). By (3.11) and (3.12), the scheme (3.2) approximates (3.13a) with an error $O(\varepsilon/\Delta t)$ and approximates (3.13b) with an error $O(\Delta t + \varepsilon/\Delta t)$. Thus, the splitting scheme always has the correct asymptotic limit independently of the initial data.

Remark. We use the first-order splitting method just to carry out the analysis and to illustrate the basic ideas. It does not mean that we advocate the use of a first-order scheme.

3.4. *The Strang Splitting*

A frequently used splitting method for an inhomogeneous hyperbolic system is Strang's splitting [37]. If we call the stiff ODE operators as $\mathcal{S}_1(t)$ and the homogeneous convection operator $\mathcal{S}_2(t)$ then the Strang splitting takes the form

$$U^{n+1} = \mathcal{S}_1\left(\frac{\Delta t}{2}\right) \mathcal{S}_2(\Delta t) \mathcal{S}_1\left(\frac{\Delta t}{2}\right) U^n. \tag{3.14}$$

This is a second-order splitting for $\varepsilon = O(1)$ and Δt, $\Delta x \ll \varepsilon$, as long as both $\mathcal{S}_1$ and $\mathcal{S}_2$ are of second-order discretizations. In this section we will show that as $\varepsilon \to 0$ while holding Δt and Δx fixed, the Strang splitting becomes only a *first*-order approximation to the equilibrium equation (1.4b).

For the model system (1.3), since the variable w is linear in the system, we can even assume that $\mathcal{S}_1$ is the exact solution operator. Thus in the stiff ODE step we do not introduce any numerical error. The exact ODE solver, of course, also projects the solution to a local equilibrium. For $\mathcal{S}_2$ we use the second-order explicit Runge–Kutta method. Applying the Strang splitting (3.14) to (1.3), one gets

$$H^* = H^n, \tag{3.15a}$$

$$W^* = \mathcal{S}_1\left(\frac{\Delta t}{2}\right)(H^n, W^n); \tag{3.15b}$$

$$H^{(1)} = H^* - \lambda \Delta_+ W^*_{j-1/2}, \tag{3.15c}$$

$$W^{(1)} = W^* - \lambda \Delta_+ p^*_{j-1/2}; \tag{3.15d}$$

$$H^{(2)} = H^{(1)} - \lambda \Delta_+ W^{(1)}_{j-1/2}, \tag{3.15e}$$

$$W^{(2)} = W^{(1)} - \lambda \Delta_+ p^{(1)}_{j-1/2}; \tag{3.15f}$$

$$H^{(3)} = \tfrac{1}{2}(H^* + H^{(2)}), \tag{3.15g}$$

$$W^{(3)} = \tfrac{1}{2}(W^* + W^{(2)}); \tag{3.15h}$$

$$H^{n+1} = H^{(3)}, \tag{3.15i}$$

$$W^{n+1} = \mathcal{S}_1\left(\frac{\Delta t}{2}\right)(H^{(3)}, W^{(3)}). \tag{3.15j}$$

As $\varepsilon \to 0$, (3.15a)–(3.15b) simply projects the solution into a local equilibrium

$$W^* = f(H^*) + O(\varepsilon). \tag{3.16}$$

Note that for $n \geq 2$, both (3.15a), (3.15b), and (3.15i), (3.15j) essentially make (3.16); thus we can disregard (3.15i)–(3.15j) in our analysis. Applying (3.16) to (3.15c), one can reduce the scheme (3.15) to

$$H^{(1)} = H^n - \lambda \Delta_+ f^n_{j-1/2} + O(\varepsilon), \tag{3.17a}$$

$$W^{(1)} = f(H^n) - \lambda \Delta_+ p^n_{j-1/2} + O(\varepsilon); \tag{3.17b}$$

$$H^{(2)} = H^{(1)} - \lambda \Delta_+ W^{(1)}_{j-1/2}, \tag{3.17c}$$

$$W^{(2)} = W^{(1)} - \lambda \Delta_+ p^{(1)}_{j-1/2}; \tag{3.17d}$$

$$H^{n+1} = \tfrac{1}{2}(H^n + H^{(2)}), \tag{3.17e}$$

$$W^{n+1} = \tfrac{1}{2}(W^n + W^{(2)}). \tag{3.17f}$$

Clearly (3.17a) is consistent to the equilibrium equation (1.4b), modulus an $O(\varepsilon)$ error. (3.17) overall seems to be a second-order Runge–Kutta method for (1.4b), except that one needs to justify that (3.17c) is consistent to (1.4b). This requires

$$W^{(1)} \approx f(H^{(1)}).$$

Using (3.17b) and (3.17a) implies

$$W^{(1)} = f(H^n) + O(\Delta t + \varepsilon) = f(H^{(1)}) + O(\Delta t + \varepsilon). \tag{3.18}$$

If one applies (3.18) in (3.17c) then one indeed gets a consistent discretization of the equilibrium equation (1.4b); *however*, the $O(\Delta t)$ error in (3.18) makes such an approximation only *first* order! Thus in the regime $\varepsilon \to 0$ and $\varepsilon/\Delta t \to 0$ the Strang splitting is only a first-order approximation to the equilibrium equation (1.4b).

Remark 1. A similar argument shows that one cannot improve the result by using higher order Runge–Kutta methods in the convection step.

Remark 2. If one uses a Godunov type integration in time in the convection step rather than a method of line approach, then such deterioration of the numerical results does not appear

[27]. The reason for this is that the Godunov type time marching scheme is a one step scheme that only uses the initial data obtained from the first step of the ODE solver (3.15a)–(3.15b), which is a good approximation of the local equilibrium. Thus the result of this paper applies only to the method of line approaches.

To fix the problem associated with the Runge–Kutta approach one just needs to add a good stiff ODE step between $t = t^{(1)}$ and $t = t^{(2)}$ in (3.17). This will reduce the error term in (3.18). This motivates the development of our second-order splitting scheme in the next section.

3.5. *A Second-Order Splitting Scheme*

In this section we introduce a second-order Runge–Kutta Godunov splitting scheme which combines two explicit steps for the convection terms and two implicit steps for the source terms. If one views (3.1) as a splitting method in the Euler setting, then this new splitting scheme is a natural second-order extension in the Runge–Kutta setting. It is a second-order method when ε is fixed, and it not only has the correct asymptotic limit but the limiting scheme, as $\varepsilon \to 0$ is again a second-order approximation to the equilibrium system. The scheme is

$$U^* = U^n + a\frac{\Delta t}{\varepsilon} Q(U^*), \tag{3.19a}$$

$$U^{(1)} = U^* - \lambda \Delta_+ F^*_{j-1/2}; \tag{3.19b}$$

$$U^{**} = U^{(1)} + b\frac{\Delta t}{\varepsilon} Q(U^{**}) + c\frac{\Delta t}{\varepsilon} Q(U^*), \tag{3.19c}$$

$$U^{(2)} = U^{**} - \lambda \Delta_+ F^{**}_{j-1/2}; \tag{3.19d}$$

$$U^{n+1} = \tfrac{1}{2}(U^n + U^{(2)}). \tag{3.19e}$$

The coefficients a, b, and c are to be determined. Roughly speaking, this scheme has projections into the local equilibrium at two intermediate time steps, t^* (which is the very first step!) and t^{**}, immediately followed by two convection steps. Due to the projection at t^* and t^{**}, these two convection steps will relax to a limiting equilibrium scheme for the equilibrium system.

Scheme (3.19) has the following general properties:

(a) If $Q = 0$ then (3.19) reduces to

$$U^{(1)} = U^n - \lambda \Delta_+ F^n_{j-1/2},$$

$$U^{(2)} = U^{(1)} - \lambda \Delta_+ F^{(1)}_{j-1/2},$$

$$U^{n+1} = \tfrac{1}{2}(U^n + U^{(2)}),$$

which is the second-order explicit TVD Runge–Kutta method [35].

(b) For fixed $\varepsilon = O(1)$ it is second order if

$$a = -1, \quad b = 1, \quad c = 2.$$

Proof. See Appendix.

Remark. Like the second-order time discretization of Engquist and Sjogreen [14], this scheme also contains a negative parameter a in the implicit term in (3.19). Since we only concentrate on the coarse ($\Delta t \gg \varepsilon$) regime, this drawback does not have any impact on the results presented in this article. In an upcoming article this obstacle will be removed [4].

(c) The L-stability analysis with $F = 0$ shows that this splitting method gives

$$U^{n+1} = \left(\frac{1 + q - q^2/2}{1 - q^2}\right) U^n \to \frac{1}{2} U^n \quad \text{as } q \to -\infty.$$

Although this method is not L-stable, it does damp any oscillation introduced by the transient behavior with a rate of $\frac{1}{2}$.

Applying scheme (3.19) to the model problem (1.3),

$$H^* = H^n, \tag{3.20a}$$

$$W^* = W^n - a\frac{\Delta t}{\varepsilon}(W^* - f(H^*)); \tag{3.20b}$$

$$H^{(1)} = H^* - \lambda \Delta_+ W^*_{j-1/2}, \tag{3.20c}$$

$$W^{(1)} = W^* - \lambda \Delta_+ p^*_{j-1/2}; \tag{3.20d}$$

$$H^{**} = H^{(1)}, \tag{3.20e}$$

$$W^{**} = W^{(1)} - b\frac{\Delta t}{\varepsilon}(W^{**} - f(H^{**})) - c\frac{\Delta t}{\varepsilon}(W^* - f(H^*)); \tag{3.20f}$$

$$H^{(2)} = H^{**} - \lambda \Delta_+ W^{**}_{j-1/2},$$

$$W^{(2)} = W^{**} - \lambda \Delta_+ p^{**}_{j-1/2}; \tag{3.20g}$$

$$H^{n+1} = \tfrac{1}{2}(H^n + H^{(2)}), \tag{3.20h}$$

$$W^{n+1} = \tfrac{1}{2}(W^n + W^{(2)}). \tag{3.20i}$$

Scheme (3.20) has the following properties:

(d) Although (3.20) contains implicit nonlinear terms, due to the special structure of the source term, one can avoid solving nonlinear algebraic equations. This is not true for more general source terms.

(e) Suppose $ab \neq 0$; i.e., both (3.20b) and (3.20f) are genuinely implicit, and $\Delta t \gg \varepsilon$. Then,

- *If at* $t = t^n$,

$$W^n - f(H^n) = 0, \tag{3.21}$$

then at $t = t^{n+1}$ *the scheme* (3.20) *gives*

$$W^{n+1} - f(H^{n+1}) = \tfrac{1}{2}(W^n - f(H^n)) + O(\Delta t). \tag{3.22}$$

In the intermediate steps,

$$W^* - f(H^*) = 0, \quad W^{**} - f(H^{**}) = O(\varepsilon). \tag{3.23}$$

If at $t = t^n$ *the solution is not a local equilibrium, then at* $t = t^{n+1}$,

$$W^{n+1} - f(H^{n+1}) = \tfrac{1}{2}(W^n - f(H^n)) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right), \tag{3.24}$$

and

$$W^* - f(H^*) = O\left(\frac{\varepsilon}{\Delta t}\right), \quad W^{**} - f(H^{**}) = O\left(\frac{\varepsilon}{\Delta t}\right). \tag{3.25}$$

We begin with Case 1. First, if (3.21) holds, then (3.20a), (3.20b) give

$$W^* - W^n = -a\frac{\Delta t}{\varepsilon}(W^* - W^n - f(H^n)) = -a\frac{\Delta t}{\varepsilon}(W^* - W^n);$$

thus

$$W^* - W^n = 0. \tag{3.26}$$

Applying (3.26) in (3.20b) gives

$$W^* - f(H^*) = -\frac{\varepsilon}{a\,\Delta t}(W^* - W^n) = 0. \tag{3.27}$$

By (3.20c), (3.20d)

$$H^{(1)} - H^n = H^{(1)} - H^* = O(\Delta t), \quad W^{(1)} - W^* = O(\Delta t). \tag{3.28}$$

Using (3.27) and (3.28) in (3.20f),

$$\begin{aligned} W^{**} - W^{(1)} &= -b\frac{\Delta t}{\varepsilon}(W^{**} - f(H^{**})) \\ &= -b\frac{\Delta t}{\varepsilon}(W^{**} - W^{(1)} + W^{(1)} - f(H^{(1)})) \\ &= -b\frac{\Delta t}{\varepsilon}(W^{**} - W^{(1)}) - b\frac{\Delta t}{\varepsilon}(W^* - f(H^*)) \\ &\quad + O\left(\frac{\Delta t^2}{\varepsilon}\right) \\ &= -b\frac{\Delta t}{\varepsilon}(W^{**} - W^{(1)}) + O\left(\frac{\Delta t^2}{\varepsilon}\right), \end{aligned}$$

so

$$\begin{aligned} W^{**} - W^{(1)} &= \frac{1}{1 + b(\Delta t/\varepsilon)} O\left(\frac{\Delta t^2}{\varepsilon}\right) \\ &= O\left(\frac{\varepsilon}{\Delta t}\right) O\left(\frac{\Delta t^2}{\varepsilon}\right) = O(\Delta t). \end{aligned} \tag{3.29}$$

Applying (3.29) in (3.20f), along with (3.27), gives

$$\begin{aligned} W^{**} - f(H^{**}) &= \frac{\varepsilon}{b\,\Delta t}\left[W^{**} - W^{(1)} + c\frac{\Delta t}{\varepsilon}(W^* - f(H^*))\right] \\ &= O\left(\frac{\varepsilon}{\Delta t}\right) O(\Delta t) = O(\varepsilon). \end{aligned}$$

(3.20g), (3.20h) give

$$H^{(2)} - H^{(1)} = H^{(2)} - H^{**} = O(\Delta t), \quad W^{(2)} - W^{**} = O(\Delta t). \tag{3.30}$$

Combining (3.30) with (3.28) gives

$$H^{(2)} - H^n = O(\Delta t).$$

Now from (3.20i), (3.20j),

$$\begin{aligned} W^{n+1} - f(H^{n+1}) &= \tfrac{1}{2}(W^n + W^{(2)}) - \tfrac{1}{2}f(H^n) - \tfrac{1}{2}f(H^{(2)}) + O(\Delta t^2) \\ &= \tfrac{1}{2}(W^n - f(H^n)) + \tfrac{1}{2}(W^{(2)} - f(H^{(2)})) + O(\Delta t^2) \\ &= \tfrac{1}{2}(W^n - f(H^n)) + \tfrac{1}{2}(W^{**} - f(H^{**})) + O(\Delta t) \\ &= \tfrac{1}{2}(W^n - f(H^n)) + O(\varepsilon + \Delta t) \\ &= \tfrac{1}{2}(W^n - f(H^n)) + O(\Delta t). \end{aligned}$$

Thus (3.22) and (3.23) are shown.

Next, assume that at $t = t^n$ the solution is not in local equilibrium, so $W^n - f(H^n) = O(1) \neq 0$. Then, due to the underresolution of the initial layer, (3.20a), (3.20b) imply

$$\begin{aligned} W^* - W^n &= -a\frac{\Delta t}{\varepsilon}(W^* - W^n) - a\frac{\Delta t}{\varepsilon}(W^n - f(H^n)) \\ &= -a\frac{\Delta t}{\varepsilon}(W^* - W^n) + O\left(\frac{\Delta t}{\varepsilon}\right), \end{aligned}$$

or

$$W^* - W^n = \frac{1}{1 + \Delta t/\varepsilon} O\left(\frac{\Delta t}{\varepsilon}\right) = O(1). \tag{3.31}$$

Applying (3.31) in (3.20b) gives

$$W^* - f(H^*) = O\left(\frac{\varepsilon}{\Delta t}\right).$$

Applying similar arguments to (3.20f) gives

$$W^{**} - f(H^{**}) = O\left(\frac{\varepsilon}{\Delta t}\right).$$

Furthermore, one has

$$W^{n+1} = \frac{1}{2}(W^n + W^{(2)}) = \frac{1}{2}(W^n + W^{**}) + O(\Delta t)$$

$$= \frac{1}{2}(W^n + f(H^{**})) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right).$$

Therefore,

$$W^{n+1} - f(H^{n+1}) = \frac{1}{2}W^n + \frac{1}{2}f(H^{**}) - \frac{1}{2}f(H^n) - \frac{1}{2}f(H^{(2)}) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right)$$

$$= \frac{1}{2}(W^n - f(H^n)) + \frac{1}{2}(f(H^{**}) - f(H^{(2)})) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right)$$

$$= \frac{1}{2}(W^n - f(H^n)) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right).$$

Thus (3.24) and (3.25) are true. ∎

Note that (3.24) and (3.25) are derived independently of $W^n - f(H^n)$ (as long as it is $O(1)$); thus they are true for all $n \geq 1$ independently of the initial data $W^0 - f(H^0)$. Therefore, we have the following:

• *For any $O(1)$ initial data, the splitting scheme* (3.20) *always gives*

$$W^n - f(H^n) = \frac{1}{2^n}(W^0 - f(H^0)) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right), \tag{3.32}$$

and

$$W^* - f(H^*) = O\left(\frac{\varepsilon}{\Delta t}\right), \quad W^{**} - f(H^{**}) = O\left(\frac{\varepsilon}{\Delta t}\right), \tag{3.33}$$

for all $n \geq 1$, where t^*, $t^{**} \in (t^{n-1}, t^n)$.

Proof. By (3.24) for all $n \geq 1$,

$$W^n - f(H^n) = \frac{1}{2}(W^{n-1} - f(H^{n-1})) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right)$$

$$= \frac{1}{2^n}(W^0 - f(H^0)) + \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-1}}\right) O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right) \tag{3.34}$$

$$= \frac{1}{2^n}(W^0 - f(H^0)) + O\left(\Delta t + \frac{\varepsilon}{\Delta t}\right).$$

(3.33) follows easily from (3.34) and the analysis that leads to (3.25).

By (3.32), $W^n - f(H^n)$ decays to zero with a decay rate of $\frac{1}{2}$, up to the error of $O(\Delta t + \varepsilon/\Delta t)$. Thus we have:

(f) Scheme (3.20) has the correct asymptotic limit as $t \to \infty$ for any $O(1)$ initial data. More specifically, as $\varepsilon \to 0$, it is limited to

$$H^{(1)} = H^n - \lambda \Delta_+ f^n_{j-1/2}, \tag{3.35a}$$

$$H^{(2)} = H^{(1)} - \lambda \Delta_+ f^{(1)}_{j-1/2}, \tag{3.35b}$$

$$H^{n+1} = \tfrac{1}{2}(H^n + H^{(2)}). \tag{3.35c}$$

$$W^{n+1} = f(H^{n+1}). \tag{3.35d}$$

Here $f^n_{j-1/2}$ and $f^{(1)}_{j-1/2}$ are defined the same way as in (3.13c). This is the *second*-order TVD Runge–Kutta method for the equilibrium equation (1.4), with the spatial discretizations $f_{j-1/2}$ being $W_{j+1/2}$-evaluated at the local equilibrium $W = f(H)$. Thus this new splitting method is limited to a second-order method of the equilibrium equation as $\varepsilon \to 0$, which is the major difference from the Strang splitting. By (3.33), the splitting scheme (3.20) approaches (3.35a)–(3.35c) with an error of $O(\varepsilon/\Delta t)$, and to (3.35d) with a decay rate of $\frac{1}{2}$ up to an error of $O(\Delta t + \varepsilon/\Delta t)$. In conclusion, the splitting scheme (3.20) always has the correct asymptotic limit in long time independently of the initial data, and the limiting scheme maintains its second-order accuracy and thus should perform better than the Strang splitting.

4. NUMERICAL EXAMPLES

We now test these methods on the following example:

$$\partial_t h + \partial_x w = 0, \tag{4.1a}$$

$$\partial_t w + \partial_x (h + \tfrac{1}{2}h^2) = -10^8 (w - \tfrac{1}{2}h^2). \tag{4.1b}$$

The limiting equations for this problem are

$$w = \tfrac{1}{2}h^2 - 10^{-8}(1 + h - h^2)\,\partial_x h, \tag{4.2a}$$

$$\partial_t h + \partial_x(\tfrac{1}{2}h^2) = 10^{-8}\,\partial_x[(1 + h - h^2)\,\partial_x h]. \tag{4.2b}$$

In the first example we choose the initial condition for h as

$$h^I(x) = \begin{cases} 1 & \text{for } 0 < x < 0.2, \\ 0.2 & \text{for } 0.2 < x < 1, \end{cases} \tag{4.3a}$$

while for w the non-local-equilibrium initial data are taken to be

$$w^I(x) = -\tfrac{1}{2}h^I(x)^2. \tag{4.3b}$$

In this example, $\varepsilon = 10^{-8}$. We use reflecting boundary conditions. In all the numerical examples presented in this section we always take $\Delta x = 10^{-2}$.

Given the initial condition (4.2a) the solution of the equilibrium equation (4.2b), to the leading order, forms a shock moving to the right with speed 0.6 determined by the Rankine–Hugoniot jump condition. Note that for this problem the CFL number

$$\mathrm{CFL} = \max_h \sqrt{h}\,\frac{\Delta t}{\Delta x} = \sqrt{2}\,\frac{\Delta t}{\Delta x} = \lambda.$$

We now test the three splitting schemes discussed in last section. In all the schemes we use CFL = 0.37 ($\Delta t = 0.0025$) and output the numerical solutions at $t = 0.5$ in Fig. 1. Figures 1a, b, and c show the results of h, w, and $w - f(h)$ given by the first-order splitting scheme (3.1), the Strang splitting (3.15), and the new splitting (3.20), respectively. All the schemes capture the correct equilibrium behavior, but Strang's splitting gives inferior results, compared with the other two splittings. In Fig. 1b we do not plot $w - f(H)$ for the Strang splitting since that is $O(\varepsilon)$ by the exact ODE solver that we use.

In the next example we still solve (4.1) but with the initial condition given by

$$h^I(x) = 1 + 0.2\sin(8\pi x) \tag{4.4a}$$

and the local equilibrium condition for w,

$$w^I(x) = \tfrac{1}{2}h^I(x)^2. \tag{4.4b}$$

The boundary condition is periodic. We choose $\Delta t = 0.005$ and output the solutions of the Strang splitting and the new splitting at $t = 0.3$ in Fig. 2. One can clearly see that the Strang splitting exhibits a typical first-order nature for a solution with complicated structures, while the new splitting gives a results of a typical second-order TVD (total-variation-diminishing) behavior.

5. SOME APPLICATIONS

In this section we apply the second-order splitting scheme (3.19) to two more general relaxation systems. These include the Broadwell model of the nonlinear Boltzmann equation of rarefied gas dynamics and the Eulerian gas dynamics with heat transfer. We believe that scheme (3.19) is also applicable to other discrete velocity kinetic equations and gas dynamics with thermo-nonequilibrium.

5.1. *The Broadwell Model*

A simple discrete velocity model for a gas was introduced by Broadwell [3]. It can be derived by looking for one-dimensional solutions of a four-velocity model. The gas is defined by a density function in phase space satisfying

$$\begin{aligned} \partial_t f_+ + \partial_x f_+ &= -\frac{1}{\varepsilon}(f_+ f_- - f_0^2), \\ \partial_t f_- - \partial_x f_- &= -\frac{1}{\varepsilon}(f_+ f_- - f_0^2), \\ \partial_t f_0 &= -\frac{1}{2\varepsilon}(f_+ f_- - f_0^2). \end{aligned} \tag{5.1}$$

Here f_+, f_-, and f_0 denote the particle density distribution at time t, position x, with velocity 1, -1, and 0, respectively; ε is the mean free path. The fluid dynamic variables for the Broadwell model are density ρ and momentum m defined by

$$\rho = f_+ + 2f_0 + f_-, \quad m = f_+ - f_-,$$

In addition, define

$$z = f_+ + f_-;$$

then the Broadwell equations can be rewritten as

$$\partial_t \rho + \partial_x m = 0, \tag{5.2a}$$

$$\partial_t m + \partial_x z = 0, \tag{5.2b}$$

$$\partial_t z + \partial_x m = \frac{1}{2\varepsilon}(\rho^2 + m^2 - 2\rho z). \tag{5.2c}$$

A local Maxwellian (or local equilibrium) in (5.1) is a density function that satisfies

$$f_0^2 = f_+ f_-,$$

or in fluid variables,

$$z = \frac{1}{2\rho}(\rho^2 + m^2). \tag{5.3}$$

As $\varepsilon \to 0$, the following model Euler equation can be derived by applying (5.3) in (5.2b):

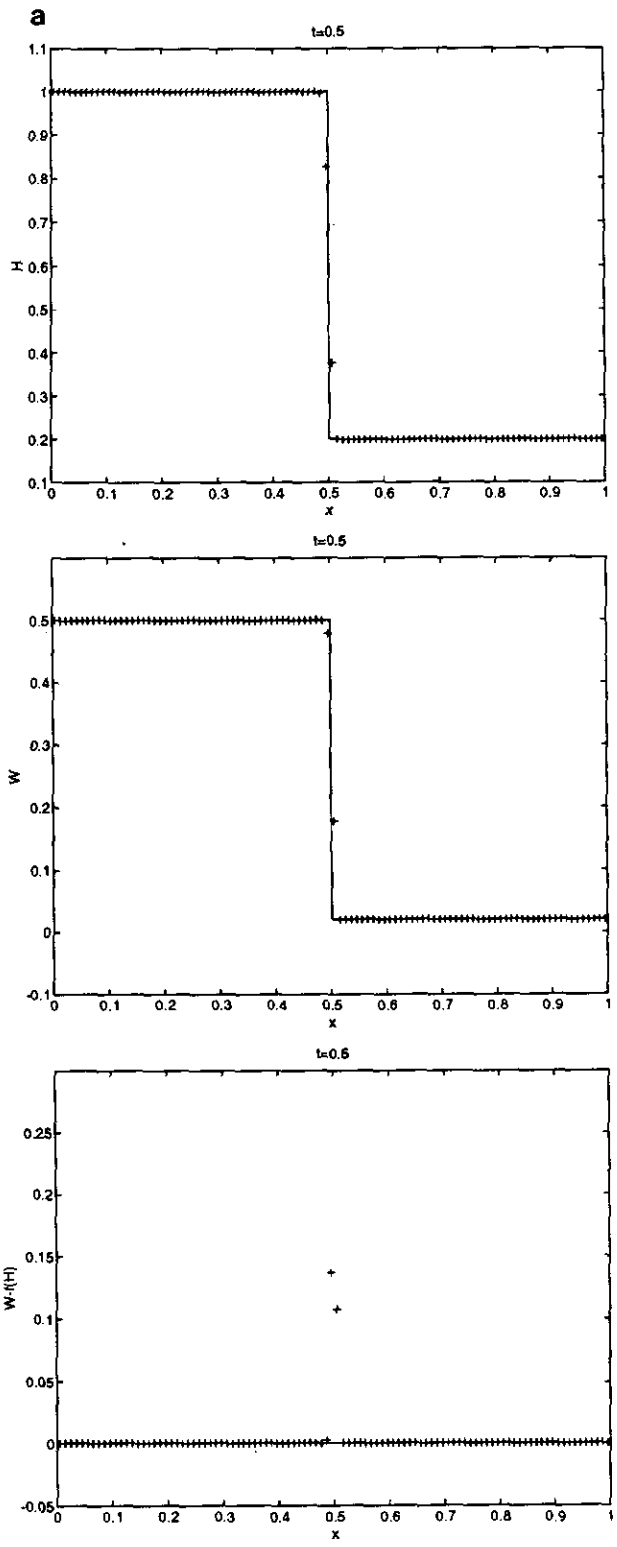

FIG. 1. Numerical solutions of Eqs. (4.1) and (4.3) at $t = 0.5$. The solid lines are the exact solutions; the "+" lines are the numerical solutions with $\Delta x = 0.01$, $\Delta t = 0.0025$ (CFL = 0.37). (a) The first-order splitting (3.1). (b) The Strang splitting (3.15). (No $w - f(h)$ is plotted since we used the exact ODE solver here; thus it is of $O(\varepsilon)$.) (c) The new splitting (3.20).

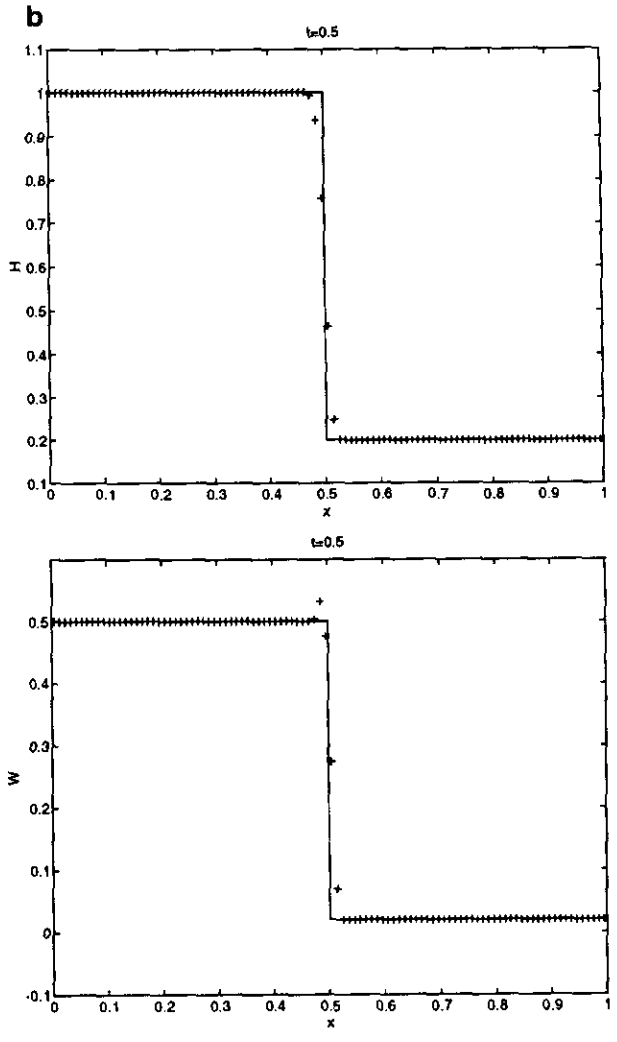

FIG. 1—*Continued*

$$\partial_t \rho + \partial_x m = 0,$$

$$\partial_t m + \partial_x \frac{1}{2\rho}(\rho^2 + m^2) = 0.$$

By including the $O(\varepsilon)$ correction, one obtains a model Navier–Stokes equation [3].

We now test the van Leer new splitting scheme for the Broadwell equation (5.2) with $\varepsilon = 10^{-8}$ by solving a Riemann problem with the initial data

$$\rho_l = 1, \quad \rho_r = 0.2, \quad m_l = m_r = 0, \quad z_l = z_r = 1. \tag{5.4}$$

Here the initial data are not a local equilibrium. The initial jump appears at $x = 0.5$. We integrate the Broadwell equation over [0, 1] with 200 spatial cells and $\Delta t = 0.0025$. The boundary condition is reflecting. The solution output at $t = 0.25$ and depicted in Fig. 3, contains a left-moving rarefaction and a

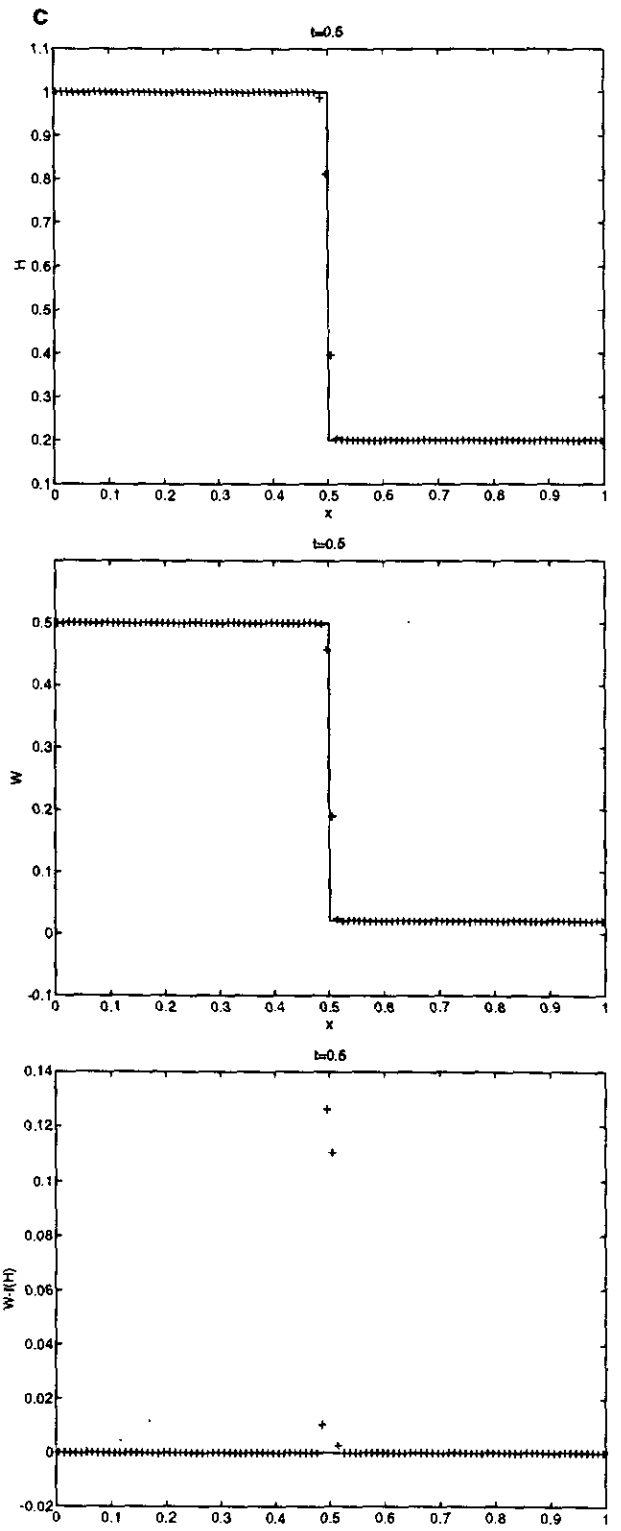

FIG. 1—*Continued*

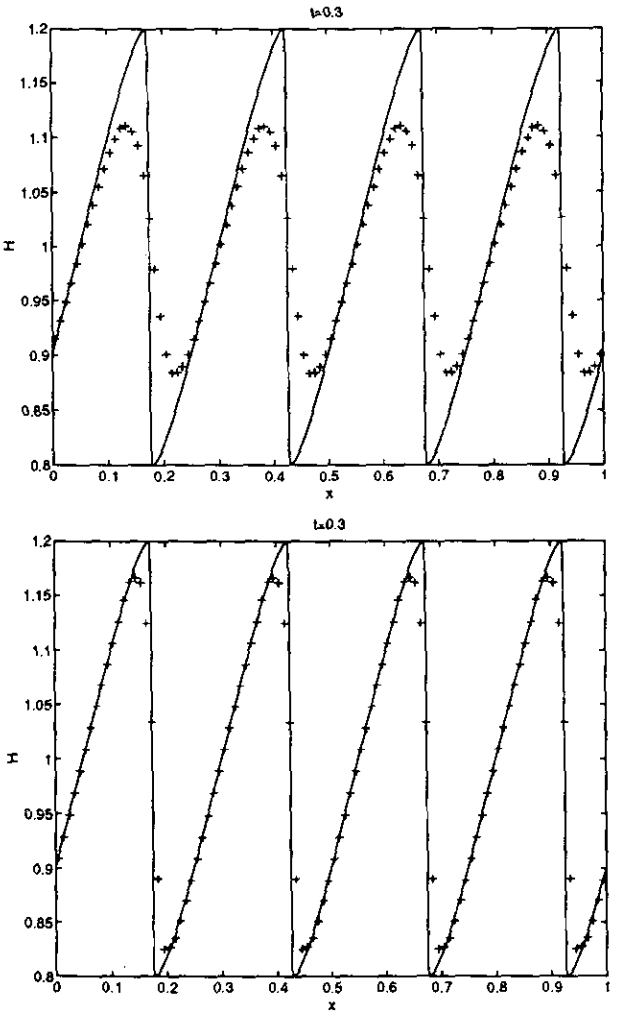

FIG. 2. Numerical solutions of Eqs. (4.1) and (4.4) at $t = 0.3$. The solid lines are the exact solutions; the "+" lines are the numerical solutions with $\Delta x = 0.01$, $\Delta t = 0.005$: above, the Strang splitting (3.15); below, the new splitting (3.20).

right-moving shock wave. Although the mean free path $\varepsilon = 10^{-8}$ is underresolved, the numerical scheme does capture the correct behavior given by the model Euler equations.

5.2. *Eulerian Gas Dynamics with Heat Transfer*

Consider the one-dimensional Euler equations for gas dynamics, coupled with a simplified heat transfer rate equation with a constant temperature bath [31]:

$$\partial_t \rho + \partial_x \rho u = 0, \tag{5.5a}$$

$$\partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0, \tag{5.5b}$$

$$\partial_t(\rho E) + \partial_x(\rho u E + up) = -K\rho(T - T_0). \tag{5.5c}$$

In this system, ρ is the density, u the velocity, $E = e + u^2/2$

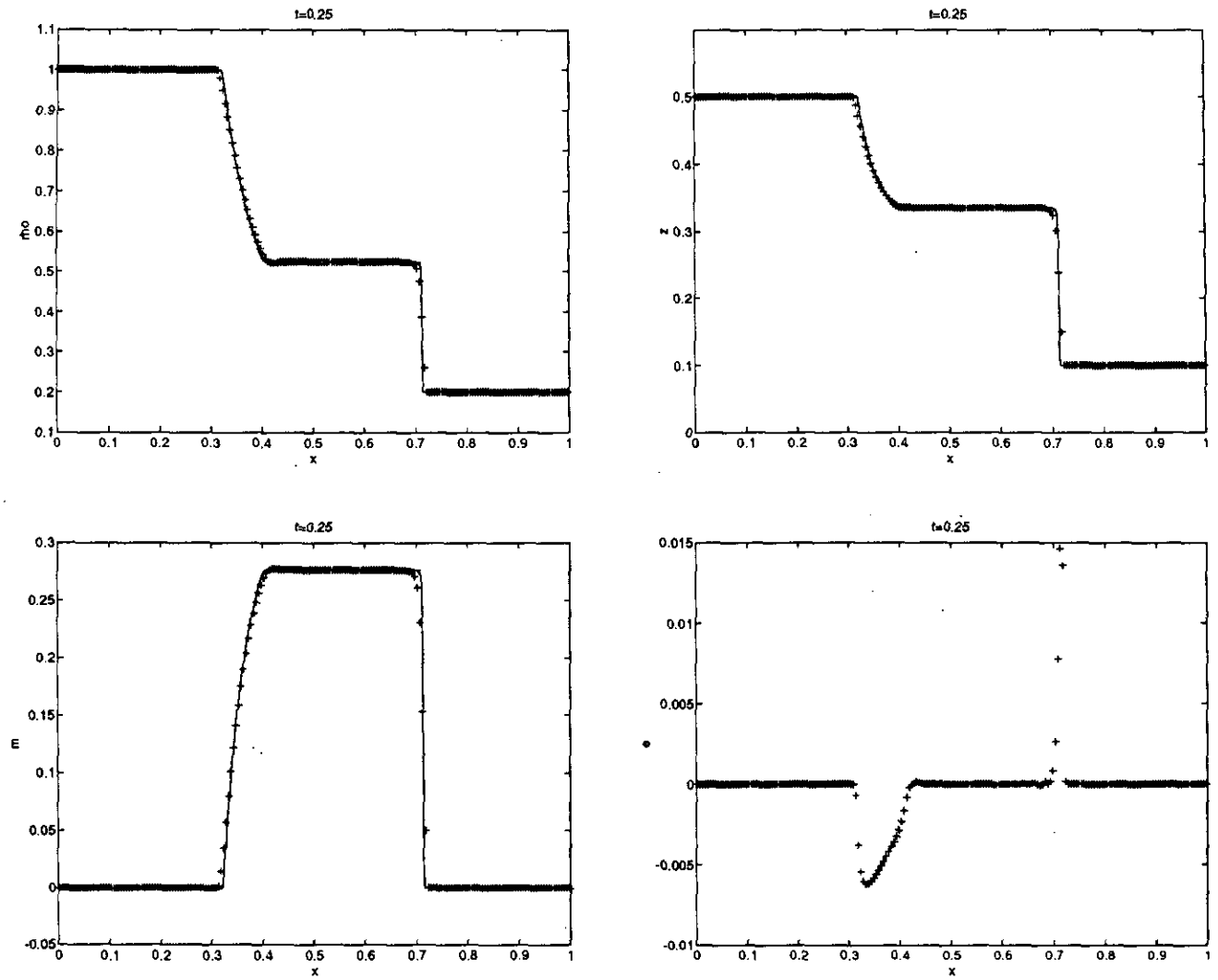

FIG. 3. Numerical solutions by the new splitting scheme (3.19) of the Broadwell equations (5.2) with initial data (5.4) not in the local equilibrium. The ρ, m, z and $e = z - (1/2\rho)(\rho^2 + m^2)$ at time $t = 0.25$ are depicted; $\varepsilon = 10^{-8}$, $\Delta x = 0.005$. The solid lines are the exact solutions; the "+" lines are the numerical solutions with $\Delta t = 0.0025$ (CFL = 0.5).

the energy per unit mass, e the internal energy, T the temperature, and p the pressure. Away from equilibrium we assume the gas is a γ-law gas, i.e., $p = (\gamma - 1)\rho e$. We choose units of temperature so that $T = e$. K and T_0 are positive constants. $K \gg 1$ is the heat transfer coefficient. T_0 is the temperature of the constant temperature bath. The characteristic speeds of the system are $u - c$, u, and $u + c$, where $c = \sqrt{\gamma p/\rho}$. At equilibrium,

$$T = T_0 \quad \text{or} \quad E = T_0 + \tfrac{1}{2}u^2,$$

the flow is governed by the Eulerian equations for isothermal flow:

$$\partial_t \rho + \partial_x \rho u = 0,$$

$$\partial_t(\rho u) + \partial_x(\rho u^2 + p_*) = 0.$$

The pressure p_* is governed by an isothermal gas law $p_*(\rho) = (\gamma - 1)\rho e_0$, where e_0 is the internal energy of the gas at $T = T_0$. The equilibrium characteristic speeds are $u - c_*$ and $u + c_*$, where

$$c_* = \sqrt{p/\rho} = \sqrt{(\gamma - 1)e_0}.$$

We now test the van Leer-new splitting scheme for Eqs. (5.5) with $K = 10^8$ by solving a Riemann problem with initial data

$$\rho_l = 1, \quad \rho_r = 0.2, \quad m_l = m_r = 0, \quad E_l = E_r = 1. \quad (5.6)$$

Here the initial data are not the local equilibrium. The initial

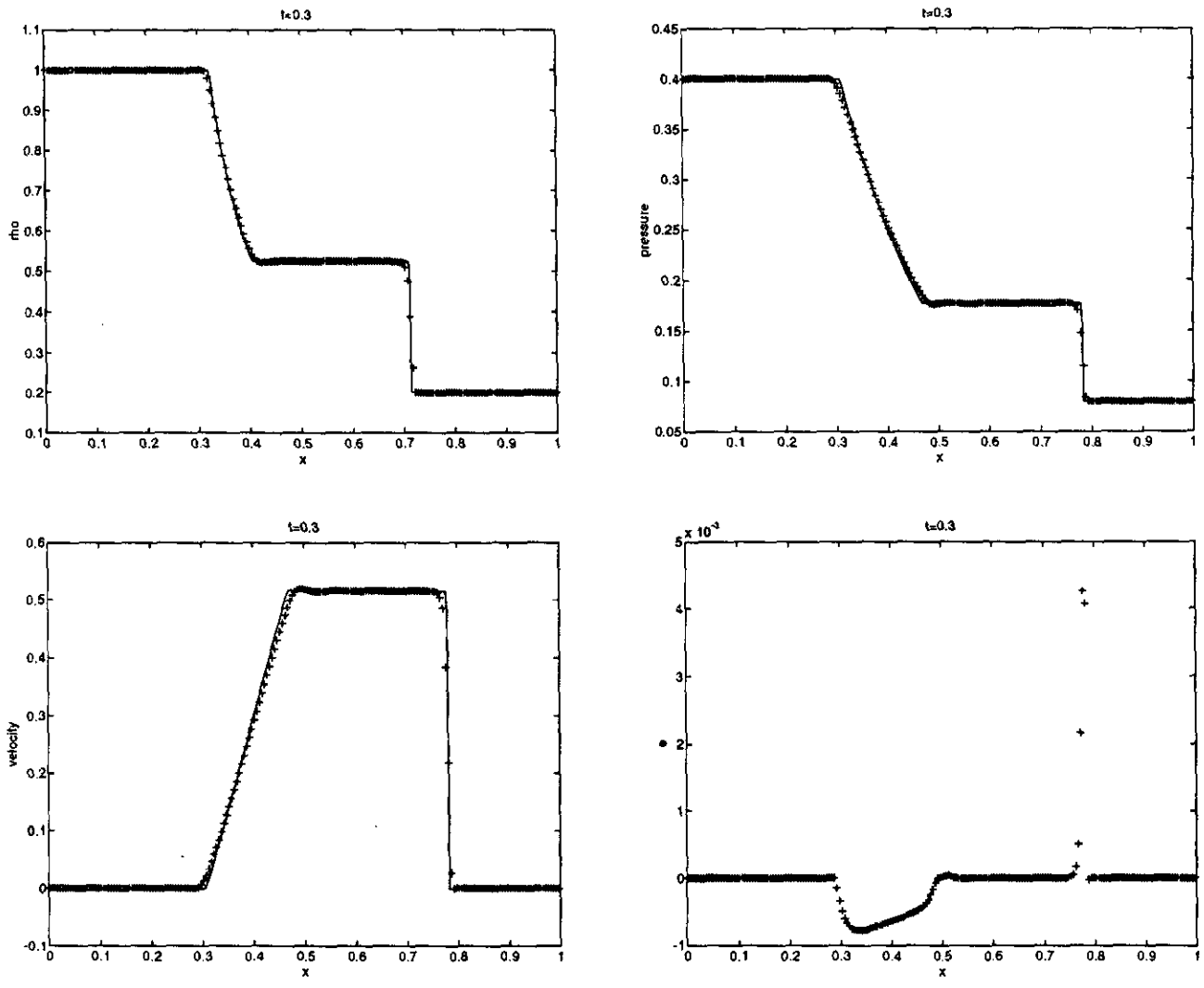

FIG. 4. Numerical solutions by the new splitting scheme (3.19) of the Euler equations with a heat transfer (5.5) with initial data (5.6) not in the local equilibrium. The density ρ, velocity u, pressure p, and $e = E - E_0$ at time $t = 0.3$ are depicted; $K = 10^8$, $\Delta x = 0.005$. The solid lines are the exact solutions; the "+" lines are the numerical solutions with $\Delta t = 0.002$.

jump appears at $x = 0.5$. We integrate over [0, 1] with 200 spatial cells and $\Delta t = 0.002$. The boundary condition is reflecting. The solution, output at $t = 0.3$ and depicted in Fig. 4, contains a left-moving rarefaction and a right-moving shock wave. Although the relaxation time $\varepsilon = 1/K = 10^{-8}$ is underresolved, the numerical scheme again captures the correct macroscopic behavior.

6. CONCLUSIONS

In this article we analyzed some underresolved, splitting schemes for hyperbolic systems with stiff relaxation terms. We indicate that such a stiff source problem is not merely a stability problem, and classical high order splitting method may fail to maintain the higher order accuracy when the relaxation time is not temporally resolved. To design a high order scheme that gives correct physical behavior, yet also maintains high order accuracy in the underresolved regime, the scheme should have a discrete analogue of the asymptotic limit of the continuous system. A new second-order splitting scheme is developed which has the correct asymptotic limit even if the initial layer and the small relaxation time is not resolved. The new scheme limits to a second-order scheme as the relaxation time shrinks to zero.

The asymptotic analysis carried out here is for a one-dimensional 2×2 p-systems. It is also applicable to general $N \times N$ relaxation systems in higher dimensions. Similarly, the second-order splitting scheme (3.19) can also be used for general hyperbolic systems with stiff relaxation terms (1.1) in any dimension. Besides the model problem (1.3), we have tested this splitting scheme on two more general one-dimensional relaxation sys-

tems in which the numerical results do have the correct asymptotic limit.

More importantly, the studies here also led to the development of the relaxation schemes for nonlinear systems of hyperbolic conservation laws [24]. This new class of TVD shock-capturing schemes do not use any Riemann solver and can be easily extended to higher dimensions.

APPENDIX: THE ORDER OF ACCURACY OF THE SPLITTING SCHEMES

Here we study the accuracy of the splitting scheme (3.19) for $\varepsilon = O(1)$. For simplicity consider the linear case with $f(U) = AU$ and $g(U) = BU$, where A and B are both constant matrices. Let $\varepsilon = 1$. Then (3.19) becomes

$$U^* = U^n - aB\,\Delta t\,U^*, \tag{A.1a}$$

$$U^{(1)} = U^* - A\,\Delta t\,U^*; \tag{A.1b}$$

$$U^{**} = U^{(1)} - bB\,\Delta t\,U^{**} - cB\,\Delta t\,U^*, \tag{A.1c}$$

$$U^{(2)} = U^{**} - A\,\Delta t\,U^{**}; \tag{A.1d}$$

$$U^{n+1} = \tfrac{1}{2}(U^n + U^{(2)}). \tag{A.1e}$$

Assume that $\|A\|\,\Delta t < 1$ and $\|B\|\,\Delta t < 1$ such that the invert of the matrices in the subsequent context is valid. From (A.1) we get

$$U^* = (I + aB\,\Delta t)^{-1}U^n, \tag{A.2a}$$

$$U^{(1)} = (I - A\,\Delta t)U^* = (I - A\,\Delta t)(I + aB\,\Delta t)^{-1}U^n, \tag{A.2b}$$

$$\begin{aligned} U^{**} &= (I + bB\,\Delta t)^{-1}(U^{(1)} - cB\,\Delta t\,U^*) \\ &= (I + bB\,\Delta t)^{-1}((I - A\,\Delta t - cB\,\Delta t) \\ &\quad (I + aB\,\Delta t)^{-1})U^n, \end{aligned} \tag{A.2c}$$

$$\begin{aligned} U^{(2)} &= (I - A\,\Delta t)U^{**} \\ &= (I - A\,\Delta t)(I + bB\,\Delta t)^{-1}((I - A\,\Delta t - cB\,\Delta t) \\ &\quad (I + aB\,\Delta t)^{-1})U^n. \end{aligned} \tag{A.2d}$$

Here (A.2b) uses (A.2a), (A.2c) uses (A.2b), and (A.2d) uses (A.2c). Note that in general $AB \neq BA$. After ignoring the $O(\Delta t^3)$ terms, we get from (A.2d) that

$$\begin{aligned} U^{(2)} = \{&I - [2A + (a + b + c)B]\,\Delta t \\ &+ [A^2 + (2a + b + c)AB + bBA \\ &+ ((a + b)(a + c) + b^2)B^2]\,\Delta t^2\}U^n. \end{aligned}$$

Therefore by (A.1e),

$$\begin{aligned} U^{n+1} = \{&I - [A + \tfrac{1}{2}(a + b + c)B]\,\Delta t \\ &+ \tfrac{1}{2}[A^2 + (2a + b + c)AB + bBA \\ &+ ((a + b)(a + c) + b^2)B^2]\,\Delta t^2\}U^n. \end{aligned} \tag{A.3}$$

From t_n to t_{n+1}, the exact solution is

$$\begin{aligned} U^{n+1} &= e^{-\Delta t(A+B)}U^n \\ &= [I - (A + B)\,\Delta t + \tfrac{1}{2}(A^2 + AB + BA + B^2)\,\Delta t^2 \\ &\quad + O(\Delta t^3)]U^n. \end{aligned} \tag{A.4}$$

By comparing (A.4) with (A.3) one should equate the coefficients of every corresponding terms. This gives the following system of linear equations:

$$\begin{aligned} a + b + c &= 2, \\ 2a + b + c &= 1, \\ b &= 1, \\ (a + b)(a + c) + b^2 &= 1. \end{aligned} \tag{A.5}$$

Equations (A.5) are consist of four equations but only three of them are independent. Solving (A.5) gives

$$a = -1, \quad b = 1, \quad c = 2.$$

With this choice of coefficients we get a second-order ODE solver.

ACKNOWLEDGMENTS

The author thanks Professors David Levermore and George Papanicolaou for their interest and support to this work. He also thanks Zhouping Xin and *the unknown referees for their critical comments on the earlier manuscript of* this paper. In particular, he is grateful for the tireless effort of Professor Randall LeVeque that helped to bring this work to its current shape. This research was supported by AFOSR Grant F49620-92-J0098 and NSF Grant DMS-9404157.

REFERENCES

1. J. Bell, P. Colella, J. Trangenstein, and M. Welcome, AIAA Paper 87-1168-CP; in *Proceedings, AIAA 8th Computational Fluid Dynamics Conference, Honolulu, Hawaii, June 9–11, 1987*, p. 717.
2. A. Bourlioux, A. Majda, and V. Roytburd, *SIAM J. Appl. Math.* **51,** 303 (1991).
3. J. E. Broadwell, *Phys. Fluids* **7,** 1243 (1964).
4. R. Caflisch, S. Jin, and G. Russo, *SIAM J. Numer. Anal.*, to appear.
5. R. Caflisch and G. Papanicolaou, *Commun. Pure Appl. Math.* **22,** 589 (1979).
6. C. Cercignani, *The Boltzmann Equation and Its Applications* (Springer-Verlag, New York, 1988).
7. G.-Q. Chen, C. D. Levermore, and T.-P. Liu, *Commun. Pure Appl. Math.* **47,** 787 (1994).
8. A. Chorin, *J. Comput. Phys.* **25,** 253 (1977).
9. J. F. Clarke, *Rep. Prog. Phys.* **41,** 807 (1978).
10. P. Colella, A. Majda, and V. Roytburd, *SIAM J. Sci. Stat. Comput.* **7,** 1059 (1986).
11. P. Collela and P. R. Woodward, *J. Comput. Phys.* **54,** 174 (1984).
12. F. Coron and B. Perthame, *SIAM J. Numer. Anal.* **28,** 26 (1991).
13. B. Engquist, *Lect. Notes in Math.*, Vol. 1270 (Springer-Verlag, New York/Berlin, 1987), p. 10.
14. B. Engquist and B. Sjogreen, UCLA CAM Report 91-03.

15. J. Glimm, *Lect. Notes in Phys.*, Vol. 344, (Springer-Verlag, New York/Berlin, 1986), p. 177.
16. S. K. Godunov, *Mat. Sb.* **47,** 271 (1959).
17. D. F. Griffiths, A. M. Stuart, and H. C. Yee, *SIAM J. Numer. Anal.* **29,** 1244 (1992).
18. E. Harabetian, *J. Comput. Phys.* **103,** 350 (1992).
19. A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, *J. Comput. Phys.* **71,** 231 (1987).
20. A. Harten, P. D. Lax, and B. van Leer, *SIAM Rev.* **25,** 35 (1983).
21. S. Jin and D. Levermore, *Transp. Theory Stat. Phys.* **20,** 413 (1991).
22. S. Jin and C. D. Levermore, *Transp. Theory Stat. Phys.* **22** 739 (1993).
23. S. Jin and D. Levermore, *J. Comput. Phys.*, submitted.
24. S. Jin and Z. P. Xin, *Commun. Pure Appl. Math.* **48,** 235 (1995).
25. E. W. Larsen, *Nucl. Sci. Eng.* **112,** 336 (1992).
26. E. W. Larsen, J. E. Morel, and W. F. Miller, Jr., *J. Comput. Phys.* **69,** 283 (1987).
27. R. J. LeVeque, private communication.
28. R. J. LeVeque and H. C. Yee, *J. Comput. Phys.* **86,** 187 (1990).
29. T.-P. Liu, *Commun. Math. Phys.* **108,** 153 (1987).
30. R. B. Pember, *SIAM J. Appl. Math.* **53,** 1293 (1993).
31. R. B. Pember, *SIAM J. Sci. Comput.* **14,** (1993).
32. B. Perthame, *SIAM J. Numer. Anal.* **27,** 1405 (1990).
33. M. Renardy, W. Hrusa and J. Nehel, "Mathematical Problems in Viscoelasticity," *Pitman Monographs and Surveys in Pure and Appl. Math.*, Vol. 35 (Longman Sci. Tech., Essex/New York, 1987).
34. P. L. Roe, *J. Comput. Phys.* **43,** 357 (1981).
35. C.-W. Shu and S. Osher, *J. Comput. Phys.* **77,** 439 (1988).
36. J. J. Stoker, *Water Waves* (Wiley, New York, 1958).
37. G. Strang, *Numer. Math.* **6,** 37 (1964).
38. I. Suliciu, *Int. J. Eng. Sci.* **28,** 829 (1990).
39. Z. Teng, A. Chorin, and T. Liu, *SIAM J. Appl. Math.* **42,** 964 (1982).
40. B. van Leer, *J. Comput. Phys.* **32,** 101 (1979).
41. W. Vincenti and C. Kruger, *Introduction to Physical Gas Dynamics* (Krieger, Melbourne, FL, 1982).
42. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).
43. H. C. Yee, *A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods*, Lecture Series, Vol. 4 (von Karman Institute for Fluid Dynamics, 1989).