A Space-Time Conservative Method for Hyperbolic Systems with Stiff and Non-Stiff Source Terms

Shamsul Qamar and Gerald Warnecke

1 Institute for Analysis and Numerics, Otto-von-Guericke University, PSF 4120, D-39106 Magdeburg, Germany.

Received 31 August 2005; Accepted (in revised version) 2 December 2005

Communicated by Chi-Wang Shu

Abstract. In this article we propose a higher-order space-time conservative method for hyperbolic systems with stiff and non-stiff source terms as well as relaxation systems. We call the scheme a slope propagation (SP) method. It is an extension of our scheme derived for homogeneous hyperbolic systems [1]. In the present inhomogeneous systems the relaxation time may vary from order of one to a very small value. These small values make the relaxation term stronger and highly stiff. In such situations underresolved numerical schemes may produce spurious numerical results. However, our present scheme has the capability to correctly capture the behavior of the physical phenomena with high order accuracy even if the initial layer and the small relaxation time are not numerically resolved. The scheme treats the space and time in a unified manner. The flow variables and their slopes are the basic unknowns in the scheme. The source term is treated by its volumetric integration over the space-time control volume and is a direct part of the overall space-time flux balance. We use two approaches for the slope calculations of the flow variables, the first one results directly from the flux balance over the control volumes, while in the second one we use a finite difference approach. The main features of the scheme are its simplicity, its Jacobian-free and Riemann solver-free recipe, as well as its efficiency and high of order accuracy. In particular we show that the scheme has a discrete analog of the continuous asymptotic limit. We have implemented our scheme for various test models available in the literature such as the Broadwell model, the extended thermodynamics equations, the shallow water equations, traffic flow and the Euler equations with heat transfer. The numerical results validate the accuracy, versatility and robustness of the present scheme.

Key words: Hyperbolic systems with relaxation; stiff systems; space-time conservative and Jacobian-free method; high order accuracy; discontinuous solutions.

1 Introduction

Hyperbolic conservation laws with source terms arise in a variety of manners. They describe non-equilibrium flows with relaxation models, reacting flows and flows with phase transition. There are a variety of physical phenomena where hyperbolic systems with relaxation arise namely, nonlinear waves [41, 46], gas flows with relaxation [16], viscoelasticity [38], multiphase and phase transitions [40], reacting Euler equations [21], discrete velocity models in kinetic theory [18], gas with vibrational degrees of freedom [45], hyperbolic models for semiconductors [2, 4], reactive flows and radiation hydrodynamics [29]. The development of efficient numerical schemes for these hyperbolic systems is quite challenging, since in many applications the relaxation time of the source term varies from 1 to very small values if one compares it with the characteristic time scale of the hyperbolic system.

The study of relaxation problems was initiated by Whitham [46] for linear problems. For nonlinear hyperbolic systems of two equations the stability of the equilibrium equation, under the subcharacteristic condition, and the zero relaxation limit were proved by Liu [28] and Chen, Levermore and Liu [13], respectively.

It is usually impossible to separate physical problems into non-stiff and stiff regimes, especially for multiple relaxation times, hence one needs a solution to the full relaxation system in all possible cases. The construction of the schemes that work for all ranges of the relaxation time, using coarse grids that do not resolve the small relaxation time, has been studied mainly in the context of upwind methods using a method of line approach combined with suitable splitting techniques [9, 23]. Alternatively, the approaches based on the method of characteristics were also considered [6, 34]. Most of these schemes are based on the solution to the Riemann problem and, except for the scheme proposed in [9], do not provide uniform second-order accuracy with respect to the relaxation parameter without resolving the small scales or initial layer. Numerical schemes for hyperbolic systems with more general stiff relaxation terms have also been studied in [21, 24, 31].

More recently Liotta, Romano and Russo [27] extended the central schemes [32] to hyperbolic systems with source terms. They derived and analyzed both the explicit and implicit formulations. Later on, Pareschi [34] extended the schemes in order to treat the stiff sources in more efficient manner. Apart from central schemes, Pareschi and Russo [35] developed implicit-explicit (IMEX) Runge-Kutta methods up to order 3 for hyperbolic systems with source terms. In both central and Runge-Kutta schemes they obtained excellent results especially for the stiff source terms where the behavior of the scheme is more important to be analyzed. We also consider the same numerical test cases with the same parameters which were presented in these articles. Hence one may compare our results with those presented in [23, 27, 34].

Apart from the above mentioned schemes there is another family of schemes called space-time conservation element and solution element (CE/SE) methods of Chang [10]. Like central schemes, these schemes also do not need Riemann solvers. However, unlike the upwind schemes and central schemes, the flow variable distribution inside the solution
element (SE) is not calculated through a reconstruction procedure using its neighbouring values at the same time level. Instead they are calculated as a part of local space-time flux conservation in the linear case. However in the nonlinear case Chang [10] reverted to reconstruction by finite differences. A similar procedure where nodal values and the slopes are the computational degrees of freedom, was used by Xu [48] for an upwind kinetic scheme, by van der Vegt et al. [44] in their space-time discontinuous Galerkin method and by Qiu and Shu [37] in their Hermite type scheme. We proposed in [1] a way to maintain the slope propagation idea even in the nonlinear case.

Recently Yu and Chang [47] have extended the CE/SE method to hyperbolic systems with source terms. They tested their scheme for both stiff and non-stiff source terms. Their results validate the applicability of the CE/SE method for hyperbolic systems with relaxation. The extension of our scheme also follows similar concepts. The level of difficulty of the present scheme is almost the same as in the original CE/SE method presented in [47]. The approximation of the source terms is similar to that in [47] but is different in the way we calculate the flux integrals in the convection part of the scheme. The different ways of approximating the flux integrals lead to a striking difference between the two schemes.

In this article we extend our second-order scheme [1] to hyperbolic systems with stiff and non-stiff source terms as well as relaxation systems. We derive the scheme in a simple and straightforward way by using the basic concept of finite volume schemes and conservation laws. The scheme utilizes the advantages of both the CE/SE method of Chang’s [10, 47] and the central schemes of Nessyahu and Tadmor [32]. However, unlike the CE/SE method the present scheme is Jacobian-free and hence like the central schemes can also be applied to any hyperbolic system. The scheme uses space-time control volumes in order to compute the conservative flow variables and their slopes. Note that the so called slope propagation technique has also been studied by several other authors, for example Bouchut et al. [8] as well as Tang and Wu [42]. Unlike the CE/SE method [10, 47], we do not assume the space-time linear variation of the fluxes in each element. We assume the linear variation of the conservative flow variables only. The fluxes are calculated from the flow variables at the midpoint of the faces of the space-time control volumes and the flux integrals are approximated by using midpoint rule. The source term is included as volumetric integral over space-time control volume such that it becomes an integral part of the overall space-time flux balance. In particular, the treatment of the source term is done in two ways. The first way is a straightforward extension of the original SP-method and is suitable for moderately stiff source terms with relaxation time up to the order of $10^{-5}$. This method will fail when it is used to solve very stiff relaxation systems. This failure is due to amplification effects by the source terms over the differences of the flow properties at adjacent nodes at the same time level. Hence it means that the source term effects become more important in the steady state situation. For highly stiff source terms the failure is removed in such a way that the source-term effects are only restricted to the mesh nodes at the updated time level. As a result we can write our scheme in the $\vartheta$—scheme form with $\vartheta$ ranging from 1/2 to 1. Hence for a non-stiff source terms one takes $\vartheta = 1/2$, while for highly stiff source terms $\vartheta = 1$. One may use other values of $\vartheta$ as well,
for example $\vartheta = 1/4$. In case of the extended central scheme, it has been shown in [27] that this choice of $\vartheta$ gives a scheme which is stable for both stiff and non-stiff source terms. Although we also define our scheme for $\vartheta = 1/4$, we only give the numerical results for $\vartheta = 1$ and $\vartheta = 1/2$. The present method is robust and stable for treating under-resolved stiff source terms. The above fact is also supported by one of the numerical examples presented in this article.

Note that the present scheme, which is only derived for the one-dimensional hyperbolic systems with source terms, can easily be extended to the multidimensional case by using a similar procedure which we have used in [1] for the same scheme in case of hyperbolic systems without source term.

The main features of our scheme are as follows:

(i) Space and time are unified.
(ii) The discrete space-time control volumes are the basic conservation regions.
(iii) In contrast to the usual MUSCL-type approach, the conservative flow variables and their derivatives are considered as independent variables. Even in the case of non-linear conservation laws our scheme follows a similar procedure.
(iv) The mesh is staggered in time with stencil similar to that of central schemes [32] and Chang’s method [10].
(v) The extension of the scheme to hyperbolic systems with relaxation is natural.
(vi) The scheme enjoys advantages of both central schemes and the CE/SE method.
(vii) The Courant-Friedrichs-Lewy (CFL) condition is independent of the small relaxation time and is solely determined by the non-stiff convection part.
(viii) The scheme correctly captures all the discontinuous features of the problem and gives oscillations-free correct shock positions.
(ix) The scheme gives the correct macroscopic behavior with high order accuracy by using coarse grid that does not resolve the small relaxation time.

The outline of the paper is the following. In Section 2, we briefly give the mathematical model of the problem under consideration. In Section 3, we derive our scheme for the one-dimensional hyperbolic systems with relaxation. In this section we also show that our scheme possesses the discrete analog of the continuous asymptotic limit outlined in Section 2. In Section 4, we present the numerical results of our scheme for several well known relaxation models which are available in the literature. Finally in Section 5 we give conclusions and remarks.

2 Mathematical model

In this article we are interested in hyperbolic systems with stiff and non-stiff source terms. A general hyperbolic system of this type can be written as

$$
\partial_t U + \partial_x F(U) = S(u), \quad x \in \mathbb{R}, \quad t \in \mathbb{R}^+,
$$

(2.1)
where $U \in \mathbb{R}^m$, $F(U) : \mathbb{R}^m \rightarrow \mathbb{R}^m$, $S(U) : \mathbb{R}^m \rightarrow \mathbb{R}^m$. An important class of such systems are the systems of relaxation type. They are given by

$$\partial_t U + \partial_x F(U) = \frac{1}{\varepsilon} R(U), \quad x \in \mathbb{R}, \quad t \in \mathbb{R}^+, \quad (2.2)$$

where $R(U) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $\varepsilon > 0$ is the stiffness parameter. This system according to Whitham [46] and Liu [28] is a relaxation system if there exists a constant $n \times m$ matrix $Q$ with rank($Q$) = $n < m$ such that

$$QR(U) = 0 \quad \forall U \in \mathbb{R}^m. \quad (2.3)$$

This gives $n$ independent conserved quantities $u = QU \in \mathbb{R}^n$. Moreover we assume that equation $R(U) = 0$ can be uniquely solved in terms of $u$, i.e.

$$U = E(u) \quad \text{such that} \quad R(E(u)) = 0. \quad (2.4)$$

The image of $E$ represents the manifold of local equilibria of the relaxation operator $R$. Using (2.3) in (2.2) we obtain a system of $n$ conservation laws which is satisfied by every solution of (2.2)

$$\partial_t (QU) + \partial_x (QF(U)) = 0. \quad (2.5)$$

For values of the relaxation parameter $\varepsilon$ from (2.2) close to zero, we get $R(U) = 0$ which by (2.4) implies $U = E(u)$. In this case system (2.2) is well approximated by the equilibrium system [13]

$$\partial_t (u) + \partial_x G(u) = 0, \quad (2.6)$$

where $G(u) = QF(E(u))$.

System (2.6) is the formal limit of system (2.2) as $\varepsilon \rightarrow 0$. The solution $u(x,t)$ of the system will be the limit of $QU$, with $U$ being a solution of system (2.2), provided a suitable condition on the characteristic velocities of systems (2.2) and (2.6) is satisfied, the so called subcharacteristic condition [13, 46].

## 3 The one-dimensional SP-method

In this section we derive the SP-method for one-dimensional hyperbolic systems of relaxation type. Let us consider the initial value problem

$$\partial_t U + \partial_x F(U) = \frac{1}{\varepsilon} R(U), \quad x \in \mathbb{R}, \quad t \in \mathbb{R}^+, \quad (3.1)$$

$$U(x,0) = \Phi(x), \quad -\infty < x < \infty, \quad (3.2)$$

where $U \in \mathbb{R}^m$, $f(U) : \mathbb{R}^m \rightarrow \mathbb{R}^m$, $R(U) : \mathbb{R}^m \rightarrow \mathbb{R}^m$, for $m \geq 1$. 
Here $U(x,t)$ is a vector of conserved variables, $F(U)$ is the corresponding vector of fluxes, and $R(U)$ is a vector of source terms. Let $x_1 = x$, and $x_2 = t$ be the coordinates of a two-dimensional Euclidean space $\mathbb{R}^2$. The integral form of the above equation is given by

$$\oint_{S(V)} h_k \cdot dS = \frac{1}{\varepsilon} \int_V R_k(U) dV, \quad \forall k = 1, 2, \cdots, m,$$

(3.3)

where $S$ is the boundary of an arbitrary space-time region $V$ in $\mathbb{R}^2$ and the space-time vector $h_k \overset{\text{def}}{=} (F_k, U_k)$, where $U_k, F_k, R_k$ are the components of vectors $U, F$ and $R$, respectively. Equation (3.3) is enforced over a space-time region $V$, called control volume (CV), in which discontinuities of the flow variables are allowed. Here $dS = \mathbf{n} d\sigma$, where $d\sigma$ is area of the face of the space-time control volume and $\mathbf{n}$ is an outward unit normal. We can see that $h_k \cdot dS$ is the space-time flux coming out of the region through the surface element $dS$.

In the present approach, like the original $\alpha$-scheme of Chang [10], the number of CVs associated with each grid point is identical to the number of unknowns. In one dimension the unknowns are $U$ and $U_x$, therefore we need two control volumes to determine them. To proceed we divide the entire computational spatial domain into non-overlapping cells, which are line elements of mesh length $\Delta x$ in this case, see Fig. 1. The dashed lines indicate the boundaries of the CVs. We introduce the grid points

$$x_i = i \cdot \Delta x \quad \text{for} \quad i \in \mathbb{Z} \quad \text{and} \quad x_{i+\frac{1}{2}} = x_i + \frac{\Delta x}{2}.$$ 

The grid points actually used are denoted by small dots, where the hollow dots represent the grid points at the previous time step, while the filled dot is the node at the updated time. The mesh is staggered in time with a time interval $\Delta t = t^{n+1} - t^n$, $\forall n \in \mathbb{N}_0$. 

Figure 1: Geometrical representation of the Control volumes.
Consider the case where \( n \in \mathbb{N}_0 \) is an even number. At each grid point, we construct two CVs as
\[
\text{CV-I} = [x_i, x_{i+1}] \times [t^n, t^{n+1}], \quad \text{CV-II} = [x_{i+1}, x_{i+1}] \times [t^n, t^{n+1}].
\]
For \( n \) odd everything is shifted by \( \frac{\Delta x}{2} \). We approximate the flow variables in each cell \( I^n_i \) by
\[
U(x, t) = \sum \left[ U^n_i + (U_x)^n_i (x - x_i) + (U_t)^n_i (t - t^n) \right] \chi_i(x).
\]
Here, \( \chi_i(x) \) is the characteristic function of the cell \( I^n_i := \{ \xi \mid |\xi - x_i| \leq \frac{\Delta x}{2} \} \), centered around \( x_i = i \cdot \Delta x \). By using (3.1) one can replace the \( (U_t)^n_i \) in (3.4) by space derivatives as
\[
U(x, t) = \sum \left[ U^n_i + (U_x)^n_i (x - x_i) + \frac{1}{\varepsilon} R^n_i - (F_x)^n_i \right] (t - t^n) \chi_i(x).
\]
Note that the third terms in (3.4) and (3.5) are redundant with respect to later integration. But, we have chosen to retain them in the presentation. For flows in one spatial dimension using a fixed spatial domain for CVs, we can rewrite (3.3) in a vector form as
\[
\oint_{c.c.} (F dt - U dx) = \frac{1}{\varepsilon} \int_V R(U) \, dV,
\]
where c.c. indicates that the line integration is carried out in the counterclockwise direction. Applying equation (3.3) to the control volume CV-I and using the midpoint rule on the time integrals of the fluxes we get after dividing by \( \Delta x \)
\[
- \left[ U^n_i + \frac{\Delta x}{4} (U_x)^n_i \right] + 2 \lambda F(U^n_{i+1} + \frac{\Delta x}{4} (U_x)^{n+1}_{i+\frac{1}{2}}) + U^{n+1}_{i+\frac{1}{2}} - \frac{\Delta x}{4} (U_x)^{n+1}_{i+\frac{1}{2}} - 2 \lambda F(U^n_{i+\frac{1}{2}})
\]
\[
= \frac{\Delta t}{2 \varepsilon} \left( R(U^{n+1}_{i+\frac{1}{2}}) + R(U^n_i) \right),
\]
where \( \lambda = \frac{\Delta t}{\Delta x} \). For the approximation of the source term space-time integrals we have used the quadrature formula which uses the midpoint rule for space and time, i.e.
\[
\frac{1}{\varepsilon} \frac{\Delta x}{\Delta t} \int_{t^n}^{t^{n+1}} \int_{x_i}^{x_{i+\frac{1}{2}}} R(U(x, t)) \, dx \, dt \approx \frac{1}{\varepsilon} \frac{\Delta x}{\Delta t} \frac{1}{4} R^n_{i+\frac{1}{4}} = \frac{\Delta t}{2 \varepsilon} \left( R(U^{n+1}_{i+\frac{1}{2}}) + R(U^n_i) \right).
\]
Due to the midpoint rule the above choice for approximating the source term satisfies the second-order accuracy condition in the rarefied regime where \( \Delta t / \varepsilon \) is of the order of 1, see also Gabetta et al. [17] as well as Liotta et al. [27]. The midpoint values \( U^n_{i+\frac{1}{2}} \) may be
predicted via (3.5) from the known data, while $\tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$ must still be determined. Similarly, the flux balance over the control volume CV-II gives

$$- \left[ U_{i+1}^n - \frac{\Delta x}{4} (U_x)_i^{n+1} \right] + 2\lambda F(U_{i+\frac{1}{2}}^{n+1}) + \left[ U_{i+\frac{1}{2}}^{n+1} + \frac{\Delta x}{4} (U_x)_i^{n+1} \right] - 2\lambda F(\tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}) = \frac{\Delta t}{2\varepsilon} \left( R(U_{i+1}^n) + R(U_{i+1}^n) \right).$$

(3.9)

Now we can find $U_{i+\frac{1}{2}}^{n+1}$ and $(U_x)_{i+\frac{1}{2}}^{n+1}$ from the above two equations. Adding (3.7) and (3.9) we get

$$U_{i+\frac{1}{2}}^{n+1} = \frac{1}{2} (U_i^n + U_{i+1}^n) + \frac{\Delta x}{8} \left[ (U_x)_i^n - (U_x)_{i+1}^n \right] + \lambda \left[ F(U_i^{n+\frac{1}{2}}) - F(U_{i+1}^{n+\frac{1}{2}}) \right]$$

$$+ \frac{\Delta t}{4\varepsilon} \left[ R(U_i^n) + 2R(U_{i+\frac{1}{2}}^{n+1}) + R(U_{i+1}^n) \right].$$

(3.10)

Subtracting (3.7) from (3.9), we obtain

$$\Delta x (U_x)_{i+\frac{1}{2}}^{n+1} = 2(U_{i+1}^n - U_i^n) - \frac{\Delta x}{2} \left[ (U_x)_i^n + (U_x)_{i+1}^n \right]$$

$$- 4\lambda \left[ F(U_i^{n+\frac{1}{2}}) - 2F(\tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}) + F(U_{i+1}^{n+\frac{1}{2}}) \right]$$

$$+ \frac{\Delta t}{\varepsilon} \left[ R(U_i^n) - R(U_{i+1}^n) \right].$$

(3.11)

Note that (3.10) is similar to the extended central schemes for source terms presented in [17, 27]. However, due to (3.11) the present scheme is still different from these central schemes as the reconstruction is not used here.

We still need in (3.10) and (3.11) the predicted values $U_i^{n+\frac{1}{2}}$ in each cell $I_i$ as well as the cells interface values $\tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$. They can be obtained by using Taylor expansion and by approximation of the flux derivative as follows. First we define

$$U_i^{n+\frac{1}{2}} = U_i^n + \frac{\Delta t}{2} (U_i)_i^n = U_i^n + \frac{\Delta t}{2} \left[ \frac{1}{\varepsilon} R(U_i^{n+\frac{1}{2}}) - \frac{\Delta F(U_i^n)}{\Delta x} \right], \quad \text{for any } i \in \mathbb{Z}.$$

(3.12)

Note that the use of $R(U_i^{n+\frac{1}{2}})$ is important in the case of stiff source terms, such as hyperbolic systems near fluid dynamic regime. This helps in the stability of the numerical scheme by reducing the effect of the stiff source terms. The computation of flux differences $\Delta F_i^n = \Delta F(U_i^n)$ can be done by using the min-mod nonlinear limiter, see [22, 32]. In this case for any grid function \{F_i^n\} and parameter $1 \leq \alpha \leq 2$ we set

$$\Delta F_i^n = MM \left( \alpha \Delta_- F_i^n, \frac{(\Delta_- F_i^n + \Delta_+ F_i^n)}{2}, \alpha \Delta_+ F_i^n \right), \quad \Delta_\pm F_i^n = \pm (F_{i+1}^n - F_i^n).$$

(3.13)
and $MM$ denotes the min-mod nonlinear limiter

$$MM\{x_1, x_2, \ldots\} = \begin{cases} 
\min_i \{x_i\} & \text{if } x_i > 0 \; \forall i, \\
\max_i \{x_i\} & \text{if } x_i < 0 \; \forall i, \\
0 & \text{otherwise}.
\end{cases}$$  \hspace{1cm} (3.14)

Note that (3.12) gives the predicted value at the forward half-time step with respect to the data at the initial time step. Next we set

$$\tilde{U}^{n+1}_{i+\frac{1}{2}} = U^{n+1}_{i+\frac{1}{2}} - \frac{\Delta t}{2} (U_i)_{i+\frac{1}{2}} = U^{n+1}_{i+\frac{1}{2}} - \frac{\Delta t}{2} \left[ \frac{1}{\epsilon} R(U^{n+1}_{i+\frac{1}{2}}) - \frac{\tilde{\Delta} F(U^{n+1}_{i+\frac{1}{2}})}{\Delta x} \right].$$  \hspace{1cm} (3.15)

The updated values of the flow variables $U^{n+1}_{i+\frac{1}{2}}$ have been calculated in (3.10) and our formula (3.15) is only needed in (3.11) afterwards. The approximation (3.15) is the predicted value at the backward half time step with respect to the data at the updated time step. Here,

$$\tilde{\Delta} F(U^{n+1}_{i+\frac{1}{2}}) = MM \left( \alpha \tilde{\Delta}_- F^{n+1}_{i+\frac{1}{2}}, \frac{\tilde{\Delta}_+ F^{n+1}_{i+\frac{1}{2}} + \tilde{\Delta}_- F^{n+1}_{i+\frac{1}{2}}}{2}, \alpha \tilde{\Delta}_+ F^{n+1}_{i+\frac{1}{2}} \right),$$  \hspace{1cm} (3.16)

where

$$\tilde{\Delta}_- F^{n+1}_{i+\frac{1}{2}} = F(U^{n+1}_{i+\frac{1}{2}}) - F((U')^n_{i+\frac{1}{2}}), \quad \tilde{\Delta}_+ F^{n+1}_{i+\frac{1}{2}} = F((U')^n_{i+\frac{1}{2}}) - F(U^{n+1}_{i+\frac{1}{2}}),$$  \hspace{1cm} (3.17)

and

$$(U')^n_{i+\frac{1}{2}} = U^n_i - \Delta t \left[ \frac{1}{\epsilon} R(U^n_{i+\frac{1}{2}}) - \frac{\Delta F(U^n_i)}{\Delta x} \right], \quad \forall i \in \mathbb{Z}. $$  \hspace{1cm} (3.18)

Again the use of $R(U^{n+1}_{i+\frac{1}{2}})$ is important in the case of stiff source terms, however one can also use $R(U^n_i)$ in the case of non-stiff source terms. The values $(U')^n_{i+\frac{1}{2}}$ are predicted at the updated time step in which the values $\Delta F(U^n_i)$ are calculated from the known initial data through equation (3.13) and $MM$ is the minmod limiter (3.14). The equations (3.10)-(3.18) form a complete scheme.

Note that the equations (3.10) and (3.12) are implicit, hence in the case of nonlinear source terms one needs to use some root finding routine, for example the Newton method will be the good choice.

The above scheme is a non-dissipative scheme. However, in the case of problems with discontinuities we need more dissipation for the smooth transition of shocks. Hence we have to limit the slopes given by (3.11). In this article we use two different procedures to limit our slopes. In the first procedure we limit directly the slopes calculated in (3.11),
while in the second procedure instead of (3.11) we use a central finite difference procedure to calculate the derivatives of the flow variables. Let us define for each \( k = 1, 2, \cdots, m \)

\[
(V_k)_{i+\frac{1}{2}}^{n+1} = \left( (U_k')_{i+\frac{1}{2}}^{n+1} - (U_k)_{i+\frac{1}{2}}^{n+1} \right), \quad (W_k)_{i+\frac{1}{2}}^{n+1} = \left( (U_k)_{i+\frac{1}{2}}^{n+1} - (U_k')_{i+\frac{1}{2}}^{n+1} \right),
\]

where \((U_k')_{i+\frac{1}{2}}^{n+1}\) are the componentwise predicted values given by (3.18) and \((U_k)_{i+\frac{1}{2}}^{n+1}\) are the components of the conserved variables at updated time \( t^{n+1} \) which are given by (3.10).

**Limiter 1:** Direct procedure for limiting derivatives in (3.11)

Using the equation (3.11) and (3.19) we can get for the parameter \( 1 < \alpha < 2, \) the following expression for the discrete slopes at updated time

\[
\Delta x (U_k x)_{i+\frac{1}{2}}^{n+1} = \Delta x (U_k x)_{i+\frac{1}{2}}^{n+1} \cdot \varphi \left( \frac{\alpha (W_k)_{i+\frac{1}{2}}^{n+1}}{\Delta x (U_k x)_{i+\frac{1}{2}}^{n+1}} \right) \cdot \varphi \left( \frac{\alpha (V_k)_{i+\frac{1}{2}}^{n+1}}{\Delta x (U_k x)_{i+\frac{1}{2}}^{n+1}} \right),
\]

where \((U_k x)_{i+\frac{1}{2}}^{n+1}\) represent the components of vector \((U_k x)_{i+\frac{1}{2}}^{n+1}\) and \( \varphi(r) \) is some limiting function, e.g. the minmod function [26] as given by

\[
\varphi(r) = \begin{cases} 
0, & \text{for } r < 0, \\
r, & \text{for } 0 < r < 1, \\
1, & \text{for } r > 1.
\end{cases}
\]

Note that the smaller the \( \alpha \), the smaller the oscillations, but at the same time, only taking reasonably large \( \alpha \) leads to a uniformly second-order accuracy. The optimal \( \alpha \) depends on the problem at hand. However, our experiments indicate that the value of \( \alpha \) between 1.4 and 1.7 is a good choice. This procedure may give some small oscillations in the results, but this result is expected since our scheme is still non-dissipative and uniformly second-order and thus more oscillatory than the (formally first order) minmod reconstruction. Hence for strong shocks problem one may need to take smaller values of the parameter \( \alpha \).

**Limiter 2:** Central finite difference procedure for the slopes

In this case instead of equation (3.11), we use a finite difference procedure to calculate the slopes \( U_k x \). Using (3.19) we get the following relation for the slopes at updated time

\[
\Delta x (U_k x)_{i+\frac{1}{2}}^{n+1} = \frac{|V_k|^\alpha W_k + |W_k|^\alpha V_k}{|V_k|^\alpha + |W_k|^\alpha} \quad \forall k = 1, 2, \cdots, m.
\]

Again \( \alpha \) is adjustable parameter and can be between 1 or 2. However, in our calculation we take \( \alpha = 1 \). Note that the weighted average technique applied in (3.22) is identical to the one used in [10–12, 49] as well as by Tang and Wu [42].

The scheme derived above works well for solving conservation laws with non-stiff and moderately stiff source terms, i.e. for relaxation time up to the order of \( 10^{-5} \). However,
when the source term becomes more stiff the method would fail. With stiff source terms, the dominant effects in (3.10) and (3.11) are from the source terms. In such a situation a small difference in the values of the conservative flow variables and their derivatives at two consecutive nodes, for example \( x_{i-1} \) and \( x_i \), will be amplified by the stiffness factor. This results in huge differences in the values of the variables in (3.10)-(3.15). As a result the numerical solution will converge to a wrong solution. Furthermore, in the case of nonlinear source terms any root finding method will fail to find the solutions of these implicit equations.

### 3.1 Modification of the method for very stiff source terms

Due to the above difficulties in the scheme given by (3.10)-(3.15), we treat the source term in a manner to avoid these problems. To this end, for highly stiff relaxation systems we consider that all the source-term effects are restricted to the mesh points \((x_i, t^{n+1})\) at the new time level. Hence in contrast to the above method we do not include the source term effect in the predicted values (3.12) and (3.15) of the flow variables. Furthermore, the integration of the source-term over space-time region is totally based on the information over the updated time \( t^{n+1} \). In summary we get the following \( \theta \)-scheme

\[
U_{i+\frac{1}{2}}^{n+1} = \frac{1}{2}(U^n_i + U^n_{i+1}) + \frac{\Delta x}{8} \left[(U_x)_i^n - (U_x)_{i+1}^n \right] + \lambda \left[F(U_i^{n+\frac{1}{2}}) - F(U_{i+1}^{n+\frac{1}{2}})\right] \\
+ \frac{\Delta t}{2\varepsilon} \left[(1 - \theta)R(U^n_i) + 2\theta R(U_{i+\frac{1}{2}}^{n+1}) + (1 - \theta)R(U_{i+1}^n)\right], \tag{3.23}
\]

and

\[
\Delta x(U_x)_{i+\frac{1}{2}}^{n+1} = 2(U^n_{i+1} - U^n_i) - \frac{\Delta x}{2} \left[(U_x)_i^n + (U_x)_{i+1}^n \right] \\
- 4\lambda \left[F(U_i^{n+\frac{1}{2}}) - 2F(\tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}) + F(U_{i+1}^{n+\frac{1}{2}})\right] \\
+ 2(1 - \theta) \frac{\Delta t}{\varepsilon} \left[R(U_{i+1}^n) - R(U_i^n)\right], \tag{3.24}
\]

where either \( \theta = 1 \) or \( \theta = \frac{1}{2} \). The predicted values \( U^{n+\frac{1}{2}}_i \) and \( \tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} \) are given by

\[
U^{n+\frac{1}{2}}_i = U^n_i + \frac{\Delta t}{2} (U^n_i)^n = U^n_i + \frac{\Delta t}{2} \left[\frac{1}{\varepsilon} R(U_i^n) - \frac{\Delta F(U_i^n)}{\Delta x}\right], \tag{3.25}
\]

\[
\tilde{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = U_{i+\frac{1}{2}}^{n+1} - \frac{\Delta t}{2} (U_{i+\frac{1}{2}}^{n+1}) = U_{i+\frac{1}{2}}^{n+1} - \frac{\Delta t}{2} \left[\frac{1}{\varepsilon} R(U_{i+\frac{1}{2}}^{n+1}) - \frac{\Delta F(U_{i+\frac{1}{2}}^{n+1})}{\Delta x}\right]. \tag{3.26}
\]

In the above scheme for \( \theta = \frac{1}{2} \) we get back the scheme (3.10)-(3.15) for the non-stiff source terms. In our numerical results we give one example where the above method for \( \theta = \frac{1}{2} \)
fails to give the equilibrium solutions for very stiff source term with \( \epsilon = 10^{-8} \), while the scheme works well for \( \vartheta = 1 \).

One can choose other values of \( \vartheta \) as well with slight modifications in the scheme, for example \( \theta = \frac{1}{4} \). In this case the above scheme becomes

\[
U_{i+\frac{1}{2}}^{n+1} = \frac{1}{2} (U_i^n + U_{i+1}^n) + \frac{\Delta x}{8} \left[ (U_x)_i^n - (U_x)_{i+1}^n \right] + \lambda \left[ F(U_i^{n+\frac{1}{2}}) - F(U_{i+1}^{n+\frac{1}{2}}) \right] + \frac{\Delta t}{2\epsilon} \left[ \frac{3}{4} R(U_i^{n+\frac{1}{2}}) + \frac{1}{2} R(U_{i+1}^{n+\frac{1}{2}}) + \frac{3}{4} R(U_{i+1}^{n+\frac{1}{2}}) \right],
\]

(3.27)

and

\[
\Delta x(U_x)_{i+\frac{1}{2}}^{n+1} = 2(U_{i+1}^n - U_i^n) - \frac{\Delta x}{2} \left[ (U_x)_i^n + (U_x)_{i+1}^n \right] - 4\lambda \left[ F(U_i^{n+\frac{1}{2}}) - 2F(U_{i+1}^{n+\frac{1}{2}}) + F(U_{i+1}^{n+\frac{1}{2}}) \right] + \frac{3\Delta t}{2\epsilon} \left[ R(U_i^{n+\frac{1}{2}}) - R(U_{i+1}^{n+\frac{1}{2}}) \right],
\]

(3.28)

where

\[
U_i^{n+\frac{1}{2}} = U_i^n + \frac{\Delta t}{3} (U_i)_i^n = U_i^n + \frac{\Delta t}{3} \left[ \frac{1}{\epsilon} R(U_i^{n+\frac{1}{2}}) - \frac{\Delta F(U_i^n)}{\Delta x} \right].
\]

(3.29)

The values \( U_i^{n+\frac{1}{2}} \) and \( \tilde{U}_i^{n+\frac{1}{2}} \) are given by (3.25) and (3.26). Note that (3.23) and (3.27) are similar to the extended central schemes for source terms presented in [17, 27]. These central schemes have also been tested for different values of \( \epsilon \). In particular, it was found in [27] that the scheme (3.27) is stable for both stiff and non stiff source terms. However, due to (3.24) and (3.28) the present schemes are still different from these central schemes as the spatial reconstruction is not used here.

### 3.2 Accuracy analysis of the scheme

In this section we show the second-order accuracy of the above scheme. A similar analysis has already been carried out by Liotta et al. [27] for the extended central schemes. As we have mentioned before the equation (3.27) is similar to that presented in [27]. Here we also follow the same lines for the proof. The analysis for the accuracy is exactly the same for both the \( \theta = 1/2 \) scheme and the \( \theta = 1/4 \) scheme.

We shall show that the scheme is second-order accurate both in the rarefied regime (\( \epsilon = 1 \)) and the fluid dynamic limit (\( \epsilon = 0 \)). We apply our scheme to the system

\[
\begin{align*}
    u_t + v_x &= 0, \\
v_t + u_x &= -\frac{1}{\epsilon}(v - au), \\
u(x, 0) &= u(x), \quad v(x, 0) = v(x),
\end{align*}
\]

(3.30)
where \( a \in \mathbb{R}, |a| < 1 \). The accuracy of the scheme is studied by comparing the Taylor expansion in time of the exact solution and the numerical solution after one time step. In the fluid dynamic limit \( v = au \), the system reduces to the single equation
\[
\begin{align*}
    u_t + u_x &= 0, \\
    u(x, 0) &= u^0(x).
\end{align*}
\] (3.31)

For small value of \( \epsilon \), the behavior of the system is described by the Navier-Stokes limit
\[
\begin{align*}
    u_t + au_x &= \epsilon(1-a^2)u_{xx}, \\
    v &= au - \epsilon(1-a^2)u_x, \\
    u(x, 0) &= u^0(x).
\end{align*}
\] (3.33)

We shall consider separately the different regimes.

**Rarefied regime**: \( \epsilon = 1 \). The Taylor expansion of the exact solution is given by
\[
\begin{align*}
    u(x, \Delta t) &= u^0 - \Delta tv^0_x + \frac{1}{2} \Delta t^2 (u^0_{xx} + v^0 - au^0_x) + O(\Delta t^3), \\
    v(x, \Delta t) &= v^0 - \Delta t(u^0_x + v^0 - au^0) + \frac{1}{2} \Delta t^2 (v^0_{xx} + u^0_x + v^0 - au^0 + v^0_x - au^0) + O(\Delta t^3),
\end{align*}
\] (3.36)

with an obvious notation.

**Fluid dynamic limit**: \( \epsilon = 0 \). The Taylor expansion of the exact solution is given by
\[
\begin{align*}
    u(x, \Delta t) &= u^0 - \Delta tau^0_x + \frac{1}{2} \Delta t^2 a^2 u^0_{xx} + O(\Delta t^3),
\end{align*}
\] (3.38)

**Thin regime**: \( \epsilon \ll 1 \). The Taylor expansion of the exact solution is given by
\[
\begin{align*}
    u(x, \Delta t) &= u^0 - \Delta tau^0_x + \frac{1}{2} \Delta t^2 a^2 u^0_{xx} + \epsilon \Delta t(1-a^2)(u^0_{xx} - a\Delta tu^0_{xxx}) + O(\Delta t^3, \epsilon^2).
\end{align*}
\] (3.39)

The above expressions are valid for the point-wise value of the exact solution. Similar expressions can be derived for the cell average \( \overline{u}(x, t) \) defined as
\[
\overline{u}(x, t) = \frac{1}{h} \int_{-h/2}^{h/2} u(x + \xi, t) d\xi.
\]

To the same order of accuracy the expressions for the Taylor expansion of the solution are the same, except that \( u^0(x) \) will be substituted by
\[
\overline{u}^0(x) = \frac{1}{h} \int_{-h/2}^{h/2} u(x + \xi) d\xi.
\]
where \( h \equiv \Delta x \). The relation between \( u^0(x) \) and \( \tilde{u}^0(x) \) for a smooth function is

\[
u^0(x) = \tilde{u}^0(x) - \frac{1}{24}u^0_{xx}(x)h^2 + O(h^4) .
\]

By applying scheme (3.27), one computes \( \tilde{u}^{1}_{i+\frac{1}{2}} \) and \( \tilde{v}^{1}_{i+\frac{1}{2}} \) as regular functions of \( \epsilon, h, \lambda \),

\[
\tilde{u}^{1}_{i+\frac{1}{2}} = U(U_i, U_{i+1}, U'_i, U'_{i+1}, \lambda, h, \epsilon),
\]

\[
\tilde{v}^{1}_{i+\frac{1}{2}} = V(U_i, U_{i+1}, U'_i, U'_{i+1}, \lambda, h, \epsilon),
\]

where \( U'_i := \Delta x (U_x)_i \) and \( U \) denote the field variables \((\bar{u}, \bar{v})\). Direct computation shows that

\[
V(\cdot, \epsilon = 0) = aU(\cdot, \epsilon = 0).
\]

The dependence on \( \epsilon \) is regular, and the function can be expanded in \( \epsilon \). To zeroth order in \( \epsilon \) one has

\[
\tilde{u}^{1}_{i+\frac{1}{2}} = \frac{1}{2} (\tilde{u}^0_i + \tilde{u}^0_{i+1}) - \frac{1}{8} (u'_{i+1} - u'_i) - \lambda a(\tilde{u}^0_{i+1} - \tilde{u}^0_i) + \frac{1}{2} a \lambda^2 (v'_{i+1} - v'_i) .
\]

The Taylor expansion of the exact solution is given by (3.38), which, for the cell average, becomes

\[
\bar{u}(x_{i+\frac{1}{2}}, \lambda h) = \bar{u}^0(x_{i+\frac{1}{2}}) - \lambda h a u^0_x(x_{i+\frac{1}{2}}) + \frac{1}{2} \lambda^2 h^2 a^2 u^0_{xx}(x_{i+\frac{1}{2}}) + O(h^3).
\]

For a regular initial condition with

\[
v^0(x) = a u^0(x)
\]

one has

\[
\tilde{u}^0(x_{i+\frac{1}{2}}) = \frac{1}{2} (\bar{u}^0_i + \bar{u}^0_{i+1}) - \frac{1}{8} (u'_{i+1} - u'_i) + O(h^3),
\]

\[
u^0_x(x_{i+\frac{1}{2}}) = \frac{1}{h} (\tilde{u}^0_{i+1} - \tilde{u}^0_i) + O(h^2),
\]

\[
v^0_{xx}(x_{i+\frac{1}{2}}) = \frac{1}{h^2} (v'_{i+1} - v'_i) + O(h).
\]

These expression can be easily proved under the following assumptions:

(i) the function \( u^0(x) \) is smooth,

(ii) the approximation of the derivatives is first-order accurate but depends smoothly on \( x \), i.e.,
\[ \frac{u_j'}{h} = \frac{\partial u_0}{\partial x}(x_i) + c(x_i)h + O(h^2), \]

\[ \frac{v_j'}{h} = \frac{\partial v_0}{\partial x}(x_i) + d(x_i)h + O(h^2), \]

with \( c(x) \) and \( d(x) \) smooth. Making use of these relations in (3.42), (3.43), (3.45), one obtains

\[ \pi_{i+\frac{1}{2}}^1 - \pi(x_i+\frac{1}{2}, \lambda h) = O(h^3) \]

and therefore the scheme is second-order accurate in the fluid dynamic limit.

**Analysis for \( \epsilon = 1 \).** In order to check the accuracy for \( \epsilon = 1 \), we expand the numerical solution \((\pi_{i+\frac{1}{2}}^1, \pi_{i+\frac{1}{2}}^1)\) making use of the following Taylor expansion around \( x_{i+\frac{1}{2}} \),

\[
\begin{align*}
\bar{u}_{i+1} &= \bar{u} + \frac{1}{2}u_x h + \frac{1}{8}u_{xx} h^2 + O(h^3), \\
\bar{v}_{j} &= \bar{v} - \frac{1}{2}u_x h + \frac{1}{8}u_{xx} h^2 + O(h^3), \\
\frac{u_j'}{h} &= u_x + \left( c + \frac{1}{2}u_{xx} \right) h + O(h^2), \\
\frac{v_j'}{h} &= u_x + \left( c - \frac{1}{2}u_{xx} \right) h + O(h^2),
\end{align*}
\]

where \( \bar{u} = \pi(x_{i+\frac{1}{2}}) \), \( u_x \equiv u_x(x_{i+\frac{1}{2}}) \), \( u_{xx} \equiv u_{xx}(x_{i+\frac{1}{2}}) \), \( c \equiv c(x_{i+\frac{1}{2}}) \). Analogous formulae hold for \( v \). These expressions are in agreement with the expressions (3.45) and (3.46).

Substituting the expansion into the expression of \( U \) and \( V \), and making use of (3.45), one obtains

\[
\begin{align*}
\pi_{i+\frac{1}{2}}^1 &= \pi^0 - h\lambda v^0 + \frac{1}{2}h^2\lambda^2(v^0_x - au^0_x + u^0_{xx}) + O(h^3), \\
\pi_{i+\frac{1}{2}}^1 &= v^0 + h\lambda(au^0_x - v^0 - u_x) + \frac{1}{2}h^2\lambda^2(v^0 - au^0_x + u^0_x - au^0_x + v^0_x) + O(h^3),
\end{align*}
\]

which is in agreement with the Taylor expansion of the exact solution.

**Analysis for \( \epsilon \ll 1 \).** The Taylor expansion in \( h \) and \( \epsilon \) of the numerical solution \( U \) is

\[ \pi_{j+\frac{1}{2}}^1 = \pi^0 - h\lambda u^0_x + \frac{1}{2}h^2\lambda^2 v^0_{xx} + \epsilon h\lambda(u^0_{xx} - av^0_{xx}) + 2\epsilon(au^0_x - v^0_x) + O(\epsilon^2). \]

Assuming that the initial condition is in agreement with the Chapman-Enskog expansion of the solution of system (3.30), i.e., that

\[ v^0 = au^0 + O(\epsilon), \]

then the term \( 2\epsilon(au^0_x - v^0_x) \) can be neglected since it is \( O(\epsilon^2) \), and the numerical solution is in agreement with the exact solution (3.39), if we neglect terms \( O(h^3, \epsilon h^2, \epsilon^2) \).
3.3 Zero relaxation limit

It is very important to check the consistent behavior of the method when $\varepsilon \to 0$. However, it is very straightforward to see that our scheme satisfies this property. Multiplying both sides of equations (3.23)-(3.26) with $\varepsilon$ and then letting $\varepsilon \to 0$ one gets

$$R(U^n) = 0, \quad R(U^{n+1}_i) = 0, \quad R(U^{n+1}_{i+\frac{1}{2}}) = 0. \quad (3.48)$$

Thus by (2.4) these formulations in the scheme are projections of the solution onto its local equilibrium

$$U^n = \mathcal{E}(u^n), \quad U^{n+1} = \mathcal{E}(u^{n+1}), \quad U^{n+\frac{1}{2}} = \mathcal{E}(u^{n+\frac{1}{2}}), \quad U^{n+\frac{1}{2}} = \mathcal{E}(u^{n+\frac{1}{2}}). \quad (3.49)$$

By multiplying (3.23)-(3.26) with $Q$ and using (3.49) we have the limiting scheme

$$u^{n+1}_{i+\frac{1}{2}} = \frac{1}{2}(u^n_i + u^n_{i+1}) + \frac{\Delta x}{8} \left[ (u_x)_i^n - (u_x)_{i+1}^n \right] + \lambda \left[ F(u^{n+\frac{1}{2}}_i) - F(u^{n+\frac{1}{2}}_{i+1}) \right], \quad (3.50)$$

and

$$\Delta x(u_x)_{i+\frac{1}{2}}^{n+1} = 2(u^n_{i+1} - u^n_i) - \frac{\Delta x}{2} \left[ (u_x)_i^n + (u_x)_{i+1}^n \right] - 4\lambda \left[ F(u^{n+\frac{1}{2}}_i) - 2F(u^{n+\frac{1}{2}}_{i+\frac{1}{2}}) + F(u^{n+\frac{1}{2}}_{i+1}) \right], \quad (3.51)$$

where

$$u_i^{n+\frac{1}{2}} = u_i^n + \frac{\Delta t}{2} (u_t)_i^n = u_i^n - \frac{\lambda}{2} \Delta F(u_i^n), \quad (3.52)$$

$$\tilde{u}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = u_{i+\frac{1}{2}}^{n+1} + \frac{\Delta t}{2} (u_i^{n+1})_{i+\frac{1}{2}} = u_{i+\frac{1}{2}}^{n+1} + \frac{\lambda}{2} \tilde{\Delta} F(u_{i+\frac{1}{2}}^{n+1}). \quad (3.53)$$

Hence, due to equation (3.48) the scheme (3.50)-(3.53) is equivalent to the second-order SP-method [1] applied to a homogeneous hyperbolic system with $R \equiv 0$.

4 Numerical simulations

In this section we apply our scheme to well-known hyperbolic systems with relaxation which were used in the literature for the validation and benchmarking of different numerical schemes. All the numerical results with symbols $\circ, \ast, \cdot$ are calculated on a mesh $N = 200$ points, while the reference solutions with thin continuous line are obtained by using the same scheme on a 2000 grid points. We will give our results for two different slope procedures as outlined in Section 3. Hence except for Fig. 3 the plots in the left column are the results obtained with limiter 1, while the right column results are with limiter 2.
Table 1: $L^1$-error for $u$ (in the units of $10^{-3}$).

<table>
<thead>
<tr>
<th>Nr.</th>
<th>$\varepsilon = 10^{-2}$</th>
<th>$\varepsilon = 1$</th>
<th>$\varepsilon = 10^{-4}$</th>
<th>$\varepsilon = 10^{-6}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-100</td>
<td>3.69984000</td>
<td>2.93240000</td>
<td>2.84048000</td>
<td>2.71360000</td>
</tr>
<tr>
<td>100-200</td>
<td>0.92623999</td>
<td>0.62992000</td>
<td>0.74748000</td>
<td>0.50240000</td>
</tr>
<tr>
<td>200-400</td>
<td>0.23107999</td>
<td>0.16100000</td>
<td>0.20719999</td>
<td>0.14480999</td>
</tr>
<tr>
<td>400-800</td>
<td>0.05747000</td>
<td>0.04097000</td>
<td>0.06418000</td>
<td>0.03039999</td>
</tr>
</tbody>
</table>

Table 2: Convergence rate for $u$ in $L^1$-norm.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>$\varepsilon = 10^{-2}$</th>
<th>$\varepsilon = 1$</th>
<th>$\varepsilon = 10^{-4}$</th>
<th>$\varepsilon = 10^{-6}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-100-200</td>
<td>2.00</td>
<td>2.21</td>
<td>1.93</td>
<td>2.06</td>
</tr>
<tr>
<td>100-200-400</td>
<td>2.00</td>
<td>1.96</td>
<td>1.85</td>
<td>2.13</td>
</tr>
<tr>
<td>200-400-800</td>
<td>2.01</td>
<td>1.97</td>
<td>1.69</td>
<td>2.28</td>
</tr>
</tbody>
</table>

4.1 Experimental order of convergence (EOC)

In order to check the experimental convergence rate of our scheme we consider a simple linear system with stiff relaxation source term [20]

\[
\begin{align*}
\partial_t u + \partial_x v &= 0, \\
\partial_t v + \partial_x u &= (au - v)/\varepsilon.
\end{align*}
\]

As $\varepsilon \to 0$, the source term describes the equilibrium equation $v = au$. Substituting this equilibrium value into above equations, we obtain the equilibrium differential equation of the form

\[
\partial_t u + a\partial_x u = 0.
\]

As a numerical example we approximate the above system in $x \in [0, 2]$ with initial data $u(x, 0) = \sin(2\pi x)$, $v(x, 0) = au(x, 0)$. We take the parameter $a = 0.5$, $\Delta t = 0.2\Delta x$ and the final time is $t = 0.2$. Tables 1 and 2 show the corresponding $L^1$-errors and EOC for different values of $\varepsilon$, respectively. These results verify the second-order accuracy of the scheme. Note that for $\varepsilon = 10^2$, $\varepsilon = 1$ and $\varepsilon = 10^{-2}$ we have used $\theta = 1/2$ in the equations (3.23)-(3.24), while for the rest of the $\varepsilon$ values we have used $\theta = 1$. For the calculation of $L^1$-errors and EOC, we have not used the exact solution of the above linear equation (4.1). Instead we have obtained the $L^1$-errors by comparing to two numerical solution. In [1] we have already studied the EOC of the present scheme for homogeneous hyperbolic systems where we have calculated the $L^1$-error by comparing the exact and numerical solutions of the linear advection equation. The aim of the present study was to check the accuracy of the scheme for different values of $\varepsilon$. 

4.2 Broadwell model

Here we consider the Broadwell model of rarefied gas dynamics [7, 9, 23, 34]. This kinetic model is characterized by a hyperbolic system with relaxation of the form (2.2) for $m = 3$ with

$$U = (\rho, m, z), \quad F(U) = (m, z, m), \quad R(U) = \left(0, 0, \frac{1}{2} [\rho^2 + m^2 - 2\rho z] \right).$$  \hspace{1cm} (4.3)

Here $\varepsilon$ represents the mean-free path of particles. The density $\rho$ and momentum $m$ are the only conserved quantities.

In the fluid dynamics limit $\varepsilon \to 0$ we have

$$z = z_E \equiv \frac{\rho^2 + m^2}{2\rho},$$

and the Broadwell system is well approximated by the reduced system (2.5) for $n = 2$ with

$$U = (\rho, \rho u), \quad G(U) = \left(\rho u, \frac{1}{2} [\rho + \rho u^2] \right) \quad u = \frac{m}{\rho}. \hspace{1cm} (4.4)$$

We consider the following two Riemann problems [27, 34, 35] for this model:

RIEM1: $\rho_l = 2$, $m_l = 1$, $z_l = 1$, $x < 0.2$,

$$\rho_r = 1, \quad m_r = 0.13962, \quad z_r = 1, \quad x > 0.2.$$

RIEM2: $\rho_l = 1$, $m_l = 0$, $z_l = 1$, $x < 0$,

$$\rho_r = 0.2, \quad m_r = 0, \quad z_r = 1, \quad x > 0.$$  \hspace{1cm} (4.5)

The results for different values of the relaxation time $\varepsilon$ are shown in Figs. 2, 3 and 4 with CFL condition $\Delta t = 0.3\Delta x$. In Fig. 2 the first two results for the initial data RIEM1 with $\varepsilon = 1$ and $\varepsilon = 0.02$ are obtained from the scheme (3.23)-(3.26) where we choose $\vartheta = 1/2$. The last result for $\varepsilon = 10^{-8}$ is obtained from the same scheme with $\vartheta = 1$. The last results in Fig. 2 demonstrate that the solution of our scheme is projected to the equilibrium solution for $\varepsilon = 10^{-8}$. The left-hand plot in Fig. 3 shows that our scheme (3.23)-(3.26) with $\vartheta = 1/2$ works well up to $\varepsilon = 10^{-5}$ and the solution is an equilibrium solution for this value of $\varepsilon$. However, the right-hand plot shows that the solution does not converge to the equilibrium solution when $\varepsilon = 10^{-8}$. This experiment supports our claim that it is better to use the scheme with $\vartheta = 1$ when the relaxation is smaller than $10^{-5}$. The results in Fig. 4 for RIEM2 problem are obtained from the same scheme with $\vartheta = 1$ and $\varepsilon = 10^{-8}$. All the results presented here show that our scheme gives highly resolved solution looking superior to those presented in [23, 27, 34, 35].
Figure 2: Solution of Broadwell model for RIEM1.
4.3 Extended thermodynamics for a monoatomic gas

Here we apply our scheme to a more general relaxation system for which an explicit analytical expression of the eigenvalues and eigenvectors is not known. We consider the one-dimensional model of the extended thermodynamics for a monoatomic gas \[27, 34, 35\]. In the conservative form the equations can be written as a relaxation system of type (2.2) for \( m = 5 \) with

\[
\begin{align*}
U &= (\rho, m, z, w, h), \\
R(U) &= \left( 0, 0, \frac{2}{3} m^2 - w \rho, \frac{2}{3} \left[ \frac{10}{3} mz - mw - h \rho \right] \right), \\
F(U) &= \left( m, \frac{2}{3} z + w, \frac{1}{2} h, \frac{1}{\rho^2} \left[ \frac{9}{5} mw \rho + \frac{4}{15} h \rho^2 - \frac{4}{5} m^3 \right], \right. \\
& \quad \left. \frac{2}{\rho^3} \left\{ \rho^2 \left[ \frac{7}{5} wz + \frac{10}{9} z^2 + \frac{8}{5} m h \right] + m^2 \left[ \frac{79}{45} m^2 - \frac{16}{3} \rho z - \frac{28}{15} w \rho \right] \right\} \right),
\end{align*}
\]

For monoatomic gases \( \varepsilon \) is a constant proportional to the relaxation time of the system. The conserved quantities are the density \( \rho \), the momentum \( m \), and the energy \( z \). Other quantities of physical interest are the pressure \( p \), the heat flux \( q \), and the stress \( \sigma \) given by

\[
p = \frac{2 \rho z - m^2}{3 \rho}, \quad q = \frac{3h \rho^2 + 6m^3 - 10mz \rho - 6mw \rho}{6\rho^2}, \quad \sigma = \frac{3w \rho - 2m^2}{3 \rho}.
\]

As \( \varepsilon \to 0 \) we obtain the equilibrium relations

\[
w = w_E \equiv \frac{2m^2}{3 \rho}, \quad h = h_E \equiv \frac{mw}{\rho} + \frac{10m}{3 \rho},
\]
that correspond to $\sigma = q = 0$ and provide the usual Euler equations for a monoatomic gas characterized by the reduced system (2.5) for $n = 3$ with

$$ u = \begin{pmatrix} \rho, \rho u, \frac{1}{2} \rho u^2 + \frac{3}{2} p \end{pmatrix}, $$

$$ G(u) = \begin{pmatrix} \rho u, \rho u^2 + p, \frac{1}{2} \rho u^2 + \frac{5}{2} u p \end{pmatrix}, $$

$$ u = \frac{m}{\rho}. $$

Here we test our scheme for the generalized form of classical Sod problem [27, 34, 39]

$$ U = U_l = (1, 0, 5, 0, 0), \quad x < 0.5, $$

$$ U = U_r = (0.125, 0, 0.5, 0, 0), \quad x > 0.5. $$

The numerical results are given in Figs. 5 and 6 for a CFL condition $\Delta t = 0.1 \Delta x$. From the results it is clear that both types of limiters give highly resolved shear stress and heat flux profiles. However, one can see some oscillations in the last plot of the Limiter 1. We have used $\vartheta = 1/2$ in (3.23)-(3.26) for both types of limiters. Our scheme gives comparable results to those presented in [27, 34].

4.4 Shallow water equations

Here we consider a simple model of shallow water flow [23, 35]

$$ \partial_t h + \partial_x (hu) = 0, $$

$$ \partial_t (hu) + \partial_x (h + \frac{1}{2} h^2) = \frac{h}{\epsilon} \left( \frac{h}{2} - u \right), $$

(4.9)
Figure 5: Solution of extended thermodynamics model.
where \( h \) is the water height with respect to the bottom and \( hu \) the flux. The zero relaxation limit of this model is the inviscid Burger equation. The initial data are

\[
h = 1 + 0.2 \sin(8\pi x), \quad hu = \frac{h^2}{2},
\]

with \( x \in [0, 1] \). The solution at \( t = 0.5 \) in the stiff region \( \varepsilon = 10^{-8} \) with periodic conditions is given by Fig. 7. Here we have used the scheme (3.23)-(3.26) with \( \vartheta = 1 \) for both types of limiters. The results show that our scheme gives a resolved solution at 200 mesh points. Our results seem to be far better than those presented in [23, 35]. One can see that our results are comparable to the third order IMEX-SSP3-WENO scheme results in [35].

4.5 Traffic flow model

This problem was also considered by [35] for testing their schemes for hyperbolic systems with relaxation. This new macroscopic model of vehicular traffic was initially presented in [6]. The model consists of a continuity equation for the density \( \rho \) of the vehicles together
with an additional velocity equation that describes the mass flux variation due to the road conditions in front of the driver. The model can be written in conservative form as follows

\[
\begin{align*}
\partial_t \rho + \partial_x (\rho u) &= 0, \\
\partial_t (\rho w) + \partial_x (u \rho w) &= A \frac{\rho}{T} (V(\rho) - u),
\end{align*}
\] (4.11)

where \( w = u + P(\rho) \) with \( P(\rho) \) a given function describing the anticipation of road conditions in front of the driver and \( V(\rho) \) describe the dependence of the velocity with respect to the density for an equilibrium situation. The parameter \( T \) is the relaxation time and \( A > 0 \) is a positive constant.

If the relaxation time goes to zero, under the subcharacteristic condition

\[-P'(\rho) \leq V'(\rho) \leq 0, \quad \rho > 0,
\]

we obtain the Whitham [46] model

\[
\partial_t \rho + \partial_x (\rho V(\rho)) = 0.
\]

A typical choice for the function \( P(\rho) \) is give by

\[
P(\rho) = \begin{cases} 
\frac{c_v}{\gamma} \left( \frac{\rho}{\rho_m} \right)^\gamma & \gamma > 0, \\
\frac{c_v}{\gamma} & \gamma = 0,
\end{cases}
\]

where \( \rho_m \) is a given maximal density and \( c_v \) a constant with dimensions of velocity. In our numerical results we assume \( A = 1 \) and an equilibrium velocity \( V(\rho) \) fitting experimental
data [5]

\[ V(\rho) = u_m \frac{\pi/2 + \arctan \left( \frac{\alpha \rho - \beta}{\rho_m - 1} \right)}{\pi/2 + \arctan(\alpha/\beta)}, \]

with \( \alpha = 11, \, \beta = 0.22 \) and \( u_m \) a maximal speed. We consider \( \gamma = 0 \) and, in order to fulfill the subcharacteristic condition, assume \( c_v = 2 \). All quantities are normalized so that \( u_m = 1 \) and \( \rho_m = 1 \).

We consider a Riemann problem centered at \( x = 0 \) with left and right states

\[ \rho_L = 0.05, \quad u_L = 0.05, \quad \rho_R = 0.05, \quad u_R = 0.5. \]

The solution at \( t = 1.0 \) for \( T = 0.2 \) is given in Fig. 8. The figure shows the development of the density of the vehicles. Both limiters of the scheme work very well. However, Limiter 1 gives better results as compared to Limiter 2. We have used the scheme (3.23)-(3.26) with \( \vartheta = 1/2 \) and the results again verify that the scheme works well for \( \epsilon \) of the order \( 10^{-4} \). Our results seem better than those in [35].

4.6 The Euler equations with heat transfer

In this model the Euler equations for gas dynamics are coupled with a simplified heat transfer equation [23, 36], where the one-dimensional flow of gas is assumed to be in contact with a constant temperature bath. The evolution equations are

\[
\begin{align*}
\partial_t \rho + \partial_x \rho u &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2 + p) &= 0, \\
\partial_t (\rho E) + \partial_x (\rho u E + u p) &= -K \rho (T - T_0),
\end{align*}
\]
where $\rho$ is the density, $u$ the velocity, $E = e + u^2/2$ the energy per unit mass, $e$ the internal energy, $T$ the temperature, and $p$ the pressure. The closure equation is the usual $\gamma$-gas law $P = (\gamma - 1)\rho e$. We choose units of temperature so that $T = e$. Here the positives constants $K$ and $T_0$ are the heat transfer coefficient and temperature of the constant temperature bath, respectively. The characteristic speeds of the system are $u - c$, $u$ and $u + c$, where $c = \sqrt{\gamma p/\rho}$. At equilibrium $T = T_0$, or $E = T_0 + \frac{1}{2}u^2$, and the flow is governed by the Euler equations for isothermal flow:

$$\begin{align*}
\partial_t \rho + \partial_x \rho u &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2 + p_*) &= 0.
\end{align*}$$

(4.13)

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 speed are $u - c_*$ and $u + c_*$, where $c_* = \sqrt{p_*/\rho} = \sqrt{(\gamma - 1)e_0}$.

For our numerical test we use the following initial data [23, 47]:

$$\begin{align*}
\rho_l &= 1, \quad \rho_r = 0.2, \\
m_l = m_r = 0, \quad E_l = E_r = 1.0,
\end{align*}$$

where the initial jump is located at $x = 0.5$. We take $\gamma = 1.4$, $K = T_0 = 1$ and 200 mesh points in the spatial domain $[0,1]$. The final time of computation is $t = 0.3$ and we use reflecting boundary conditions. We use our scheme (3.23)-(3.26) with $\vartheta = 1$ and the results for density $\rho$, pressure $p$, velocity $u$ and departure from equilibrium $(T - T_0) \times 10^7$ are shown in Fig. 9. The solutions contain a left-moving rarefaction and a right-moving shock wave. The results obtained from both types of slope procedures are highly resolved and oscillation free. The comparison of the results with those presented in [23, 47] clearly demonstrates the accuracy of the present method. There are no overshoots at the tails of the rarefaction and the shock capturing is quite sharp. Furthermore the departure from equilibrium $(T - T_0)$ is far smaller than the one presented in [23]. Note that in this problem we have only presented the numerical solution with 200 mesh points without any reference solution on the fine grid.

5 Conclusions and remarks

In this article we extend our second-order accurate scheme [1] to hyperbolic systems with stiff and non-stiff source terms as well as relaxation systems. We derive the scheme in a simple and straightforward way by using the basic concept of finite volume schemes and conservation laws. The scheme utilizes the advantages of both the CE/SE method of Chang et al. [10,47] and central schemes of Nessyahu and Tadmor [32]. However, unlike the CE/SE method the present scheme is Jacobian-free and hence like the central schemes can
Figure 9: Solution of Euler equations with heat transfer.
also be applied to any hyperbolic system without knowing the details of characteristics. The scheme uses space-time control volumes in order to compute the conservative flow variables and their slopes. Unlike the CE/SE method [10, 47], we do not assume the space-time linear variation of fluxes in each element. We assume the linear variation of the conservative flow variables only. The fluxes are calculated from the flow variables at the midpoint of the faces of the space-time control volumes and the flux integrals are approximated by using midpoint rule. The source terms are included as volumetric integrals over the space-time control volume such that it becomes an integral part of the overall space-time flux balance. The source terms is also included in the Taylor expansion for the flow variables, see (3.5). In particular, the treatment of source terms is done in two ways. The first way is a straightforward extension of the original SP-method and is suitable for moderately stiff source terms with relaxation time up to the order of $10^{-5}$. This method fails when it is used to solve more stiff relaxation systems. This failure is due to amplification effects by the source terms on the differences of the flow variables at adjacent nodes at the same time level. For highly stiff source terms the failure is removed in such a way that the source-terms effects are restricted to the mesh nodes at updated time level, i.e. by making the computations more implicit. As a result we can write our scheme as a $\vartheta$—scheme where for a non-stiff source term we take $\vartheta = 1/2$, while for highly stiff source terms $\vartheta = 1$. The above fact is also supported by the numerical example for Broadwell model presented in Fig. 3. Several numerical test cases are considered in this article which were tested by several other authors on their schemes. We found that our scheme gives comparable and some times even better results in comparison to those schemes which we have cited in the corresponding numerical test cases of the present article. All the numerical experiments clearly demonstrate the applicability of the present scheme to complicated problems of relaxation type with very stiff source terms.

In the present study we have taken $\varepsilon$ to be some constant. However, in practice $\varepsilon$ may be a function of space-time or even solution dependent. In such situations, depending on the value of $\varepsilon$, one has to find a switching strategy to use either $\vartheta = 1/2$ scheme or $\vartheta = 1$ scheme. In these situations it will be better to try the $\vartheta = 1/4$ scheme which has been found more stable in a wide range of $\varepsilon$.

References


