AN ENTROPY SATISFYING MUSCL SCHEME FOR SYSTEMS OF CONSERVATION LAWS

Frédéric Coquel[1] and Philippe G. LeFloch[2]

Abstract. For the high order numerical approximation of hyperbolic systems of conservation laws, we propose to use as a building principle an *entropy diminishing criterion* instead of the familiar total variation diminishing criterion introduced by Harten for scalar equations. Based on this new criterion, we derive *entropy diminishing projections* that ensure, both, the second order of accuracy and all of the classical discrete entropy inequalities. The resulting scheme is a nonlinear version of the classical Van Leer's MUSCL scheme. Strong convergence of this second order, entropy satisfying scheme is proved for systems of two equations. Numerical tests demonstrate the interest of our theory.

Contents

1. Introduction.

This paper treats the numerical approximation of the discontinuous solutions to an hyperbolic system of conservation laws

$$\partial_t u + \partial_x f(u) = 0, \qquad u(x,t) \in \mathcal{U}, \quad x \in \boldsymbol{R},\, t > 0, \tag{1.1}$$

First version: March 1995. Revised version: May 1995. Accepted for publication in Numerische Mathematik.

[1] Laboratoire d'Analyse Numérique, Université Paris VI, 75252 Paris, France.
[2] Centre de Mathématiques Appliquées (CMAP), CNRS UA756, Ecole Polytechnique, 91128 Palaiseau, France.

1991 *Mathematics Subject Classification*. Primary 65M12. Secondary 35L65.
Key words and phrases: conservation law, shock wave, entropy criterion, high order difference scheme.

supplemented with an initial condition at time $t = 0$

$$u(x, 0) = u_0(x), \qquad x \in \boldsymbol{R}, \tag{1.2}$$

and the entropy criterion

$$\partial_t U(u) + \partial_x F(u) \leq 0. \tag{1.3}$$

In (1.1), the subset $\mathcal{U} \subset \boldsymbol{R}^N$ is convex and open, the flux-function $f : \mathcal{U} \to \boldsymbol{R}^N$ is a given smooth function, and the initial data $u_0 : \boldsymbol{R} \to \mathcal{U}$ belongs to $L^\infty(\boldsymbol{R})^N$. It is assumed that (1.1) admits at least one strictly convex entropy pair, i.e. a pair $(U, F) : \mathcal{U} \times \mathcal{U} \to \boldsymbol{R} \times \boldsymbol{R}$ of class $\mathcal{C}^2$ such that $\nabla F = \nabla U \cdot \nabla f$, and $\nabla^2 U \geq C Id$ for some $C > 0$. For background on hyperbolic systems, we refer to Lax [25] and, for the theory of existence of entropy solutions, to Glimm [18] for initial data with small total variation, and DiPerna [16,17] for systems of two equations with L^∞ initial data. The main application in this paper will concern the system of gas dynamics.

Among the many efficient and robust numerical methods proposed those last fifteen years for systems of conservation laws, the second and higher order accurate versions of the (first order) Godunov scheme have been recognized as particularly attractive. The (second order) MUSCL scheme was introduced by Van Leer [26] for gas dynamic computations, and extended to the third order of accuracy (PPM method) by Colella-Woodward [11]. We recall that a Godunov-type scheme decomposes into two steps, as follows.

(1) *Time evolution step*: A piecewise polynomial approximation is provided at any time level. One has to solve a Cauchy problem for (1.1) with piecewise polynomial initial data. This sequence of non-interacting *generalized Riemann problems* can be solved up to any degree of approximation [1,4, 20, 28, 31]. This construction allows one to march in time with an approximation of the solution defined in a slab up to the next time level.
(2) *Projection step*: The solution above, generally, is not piecewise polynomial and a projection step is necessary. On the other hand, even if the solution was piecewise polynomial, it needs to be modified in order to prevent spurious numerical oscillations. The control of these oscillations requires the definition of a *nonlinear projection* operator.

Several strategies have been proposed in the literature that make the above method computationally attractive.The MUSCL scheme follows the above lines and uses

(1) classical Riemann solutions instead of generalized Riemann solutions;
(2) the now famous min-mod limitor for the projection step.

Specifically a piecewise constant approximation at a given time level is determined first by a standard L^2 projection; then an interpolation of the latter is built and nonlinearly corrected using the min-mod limitor. The min-mod function is primarily defined for scalar valued functions. Thus to apply the idea to systems of conservation laws, characteristic variables can be considered to formally reduce the system (1.1) to several scalar equations. MUSCL scheme is well recognized as a very successful tool for many applications especially in computational fluid dynamics.

From the theoretical standpoint, a complete analysis of the convergence of the MUSCL scheme was obtained in recent years. The results concern *scalar* conservation laws only and are based on the observation that one can enforce high order Godunov-type methods to mimic, at the discrete level, many of the properties of entropy solutions to (1.1). Such a scheme, by construction, is stable in the total variation norm (TVD), C.f. Harten [21], and therefore converges to a weak solution to (1.1). We observe that a high order scheme cannot satisfy all of the discrete entropy inequalities (1.3) satisfied by first order schemes [35]. (C.f. however [12, 13, 40] for schemes based on a non-homogeneous correction.) The consistency with the entropy criterion (1.3) was established under some technical assumption in a pioneering work by Osher [33], and recently proved in the general case by Lions-Sougadinis [30], Yang [41], and LeFloch-Liu [27]. The monotonicity property, (i.e. the solution operator for (1.1) preserves the region of monotonicity of the solution) was an essential ingredient in the proofs. Another approach was investigated by Brenier-Osher [5] and Goodman-LeVeque [19] based on a discrete version of Oleinik's entropy criterion. C.f. also, for related results, Nessyahu-Tadmor [32] and Osher-Tadmor [35]. More recently, Bouchut-Bourdarias-Perthame [3] have constructed a second-order TVD scheme for scalar conservation laws that satisfies a new discrete form of the entropy inequalities (1.3). The present work is motivated by this result and represents an extension to systems of [3].

Van Leer's approach seems difficult to analyze in the case of *systems*. We observe that it is based on the so-called Total Variation Diminishing (TVD) criterion satisfied by scalar equations only. The total variation of a solution to (1.1) in general is not a diminishing function of time. Van Leer's scheme is defined for systems based on the intuitive idea that the system (1.1) can be formally diagonalized and the nonlinear coupling between the equations can be neglected. Each characteristic variable is classically controlled via the min-mod limitor. (As we shall see, the theory of this paper establishes that the decoupling is in fact justified in a first approximation.)

The principal objective of this paper is deriving a new version of the MUSCL scheme for systems of conservation laws, which will have a rigorous mathematical basis. A rigorous extension of Van Leer's approach to systems requires a new concept of nonlinear projection that would not create spurious numerical oscillations. Since the TVD criterion does not make sense for a system, one has to look for another way to control the oscillations: a criterion that will ensure the strong convergence of the scheme. Recent work on the convergence of schemes has pointed out the relevance of discrete entropy inequalities. For multidimensional scalar conservation laws, entropy dissipation estimates have been deduced from the latters and have successfully replaced the (stronger) TVD estimate, C.f. [9, 10, 12, 13]. For first order schemes for one-dimensional systems, we refer to [6, 7, 17, 39]. We also refer to Johnson-Szepessy [23] for high-order finite element methods for systems.

Motivated by these results, we introduce in the present paper a notion of *entropy diminishing projection* that ensures, both, the second order of accuracy and all of the discrete entropy inequalities. This notion is based on a property satisfied by (1.1): the mathematical entropy associated with an entropy solution is a diminishing function of the time variable. So we require that the projection should diminish the entropy. The *existence* of high order entropy diminishing projections for systems is established in Section 2.

The resulting scheme satisfies all discrete entropy inequalities (1.3), preserves the (available) invariant regions, and can only converge to an entropy weak solution. Strong convergence of this second order, entropy satisfying scheme is proved for systems of two equations by using the compensated compactness method. As a matter of fact, the scheme is a nonlinear version of the classical Van Leer's MUSCL scheme. The projection operator turns out to be a *nonlinear* version of the min-mod limitor. Finally our results show that the derivation of discrete entropy inequalities given in [3] carries over to the case of systems.

Numerical experiments are performed at the end of this paper, that demonstrate the interest of the scheme proposed here. It is emphasized that non-homogeneous slope limitation is not necessary, neither for the theory nor for the pratical implementation.

The actual implementation uses the approximation of the generalized Riemann problem derived by Ben-Artzi-Falcovitz [1] and Bourgeade-LeFloch-Raviart [4]. At its current stage of development, the scheme in this paper gives satisfactory computational results but it should be considered as preliminary because of its relative complexity. We believe that further study of the method would lead to an alternative method for CFD. We refer to [15] for the Euler system of gas dynamics and further numerical experiments.

The content of this paper was announced in [14]. The derivation of slope limitations for the Euler system based on the specific entropy can be traced back to the paper by Khobalatte-Perthame [24].

A brief outline of the paper is as follows. In Section 2, we present the scheme and the notion of entropy diminishing projection and derive our new discrete entropy inequalities. Section 3 is devoted to the applications: linear hyperbolic systems, systems of two equations and the Lagrangian 3×3 gas dynamics system. Numerical tests are displayed in Section 4.

2. A General Framework.

The main goal of this section is to introduce a notion of entropy diminishing projection and establish the existence of such projections leading to a second order scheme.

2.1 A Class of High Order Schemes.

Let us specify some assumptions and notation that will be used throughout this paper. We assume that for each $u \in \mathcal{U}$ the Jacobian matrix $\nabla f(u)$ has N real and distinct eigenvalues denoted by $\lambda_i(u)$, $i = 1, \cdots, N$. Let $r_i(u)$, $i = 1, \cdots, N$, be a basis of right eigenvectors. Each characteristic field is assumed to be either genuinely nonlinear, i.e. $\nabla \lambda_i \cdot r_i = 1$, or linearly degenerate, i.e. $\nabla \lambda_i \cdot r_i = 0$.

Assuming that a projection operator is given, we recall how a high order Godunov-type scheme is classically defined. So let $\pi : L^\infty(0,1)^N \to \mathbf{P}(0,1)^N$ be a projection operator over a vector space $\mathbf{P}(0,1)^N$ of (smooth enough) vector-valued functions defined on the interval $(0,1)$. Denote by $\mathbf{P}_q(0,1)^N$ the space of all vector-valued polynomial functions of degree less or equal to a given integer q. The classical Godunov scheme will correspond to the simplest choice

$$\mathbf{P}(0,1)^N = \mathbf{P}_0(0,1)^N, \quad \pi v = \int_0^1 v(y)\, dy, \quad v \in L^\infty(0,1)^N.$$

Our main example will be based on $\mathbf{P}(0,1)^N = \mathbf{P}_1(0,1)^N$. One can also consider the space of all functions of the form $T(P)$ with $P \in \mathbf{P}_q(0,1)^N$, where the integer q and the smooth mapping $T : \mathcal{U} \to \boldsymbol{R}^N$ are fixed. A special case is given by the characteristic variables of (1.1).

We introduce a uniform mesh for time and space discretization: $t_n = n\tau$ $(n \geq 0)$ and $x_j = jh$ $(j \in \mathbf{Z})$. We also set $x_{j+1/2} = (j+1/2)h$ $(j \in \mathbf{Z})$. The increments τ and h will tend to zero with the ratio $\lambda = \tau/h$ kept constant and λ will have to satisfy a CFL stability restriction. For each $j \in \mathbf{Z}$, consider the operator $\sigma_{h,j} : L^\infty(x_{j-1/2}, x_{j+1/2})^N \to L^\infty(0,1)^N$ defined for $v \in L^\infty(x_{j-1/2}, x_{j+1/2})^N$ by

$$(\sigma_{h,j} v)(y) = v(x_{j-1/2} + hy), \quad y \in (0,1). \tag{2.1}$$

Abusing notation, we shall write $\sigma_{h,j} v$ for $\sigma_{h,j}(v_{|(x_{j-1/2},x_{j+1/2})})$. Let $\mathbf{P}_h(\boldsymbol{R})^N$ be the space of all vector-valued functions defined on $\boldsymbol{R}$ that are discontinuous at each "interface" $x_{j+1/2}$ and are polynomial functions in each cell $(x_{j-1/2}, x_{j+1/2})$:

$$\mathbf{P}_h(\boldsymbol{R})^N = \Big\{ w : \boldsymbol{R} \to \boldsymbol{R}^N \,/\, \forall j, \exists v_j \in \mathbf{P}(0,1)^N, w_{|(x_{j-1/2},x_{j+1/2})} = \sigma_{h,j} v_j \Big\}.$$

Using the operator π, we define a projection operator $\pi_h : L^\infty(\boldsymbol{R}) \to \mathbf{P}_h(\boldsymbol{R})^N$ by the relation

$$\pi_h = \sigma_{h,j}^{-1} \circ \pi \circ \sigma_{h,j} \qquad j = \cdots, -1, 0, 1, \cdots.$$

A second ingredient for the construction of the scheme is the generalized Riemann problem (G.R.P.) and its approximation. The latter is a Cauchy problem (1.1)–(1.3) with the initial data

$$u_0(x) = \begin{cases} u_{\mathrm{L}}(x) & \text{for } x < 0, \\ u_{\mathrm{R}}(x) & \text{for } x > 0, \end{cases} \tag{2.2}$$

where $u_{\mathrm{L}}(x)$ and $u_{\mathrm{R}}(x)$ are two sufficiently smooth functions. In recent years, this problem has received a lot of attention from the point of view of existence and approximation. It is known [20, 31] that the solution to the generalized Riemann problem has the same local structure as the one of the classical Riemann problem with inital data

$$u_0(x) = \begin{cases} u_{\mathrm{L}}(0-) & \text{for } x < 0, \\ u_{\mathrm{R}}(0+) & \text{for } x > 0. \end{cases} \tag{2.3}$$

In contrast with the classical problem, the solutions to the generalized Riemann problem are not given by a closed formula. The practical implementation of the schemes can be based on the general theory of approximation in [28]. For the gas dynamics systems, see [1,4].

Equipped with a projection operator and a (approximate) generalized Riemann solver, we now define the approximate solutions $u_h(x,t)$ to the problem (1.1)–(1.3) such that $u_h(t_n+)$ belongs to $\mathbf{P}_h(\boldsymbol{R})^N$ for all $n \geq 0$. Denote by u_j^n the cell average of $u_h(t_n+)$ in $(x_{j-1/2}, x_{j+1/2})$. The scheme is initiated in time by setting

$$u_h(0+) = \pi_h(u_0) \in \mathbf{P}_h(\boldsymbol{R})^N. \tag{2.4}$$

Suppose that the approximate solution $u_h(t)$ has been computed up to $t \leq t_n$.

Step 1: Time Evolution. Define $u_h(t)$ for $t \in (t_n, t_{n+1})$ as the entropy solution to (1.1)–(1.3) with the initial condition $u_h(t_n+)$ at time $t = t_n$. Under a suitable CFL condition, this Cauchy problem consists of a series of non-interacting generalized Riemann problems.

Step 2: Projection. Using the operator π_h, we compute the projection of $u_h(t_{n+1}-)$ and set

$$u_h(t_{n+1}+) = \pi_h(u_h(t_{n+1}-)) \in \mathbf{P}_h(\boldsymbol{R})^N. \tag{2.5}$$

In particular, the cell average values u_j^n clearly satisfy the discrete conservation law

$$u_j^{n+1} = u_j^n - \lambda\,(g_{j+1/2}^n - g_{j-1/2}^n), \tag{2.6}$$

where the numerical flux $g^n_{j+1/2}$ is given by

$$g^n_{j+1/2} = \frac{1}{\tau}\int_{t_n}^{t_{n+1}} f(u_h(x_{j+1/2}, t))\,dt.$$

In the following, we refer to the scheme above as the *high order Godunov-type scheme* based on the projection π. Observe that exact solutions to the G.R.P. will always be considered in the convergence analysis. The analysis of *approximate* solver is out of the scope of this paper. Note that the scheme is high order accurate in both space and time if π is a high order projection.

Remark 2.1. In Van Leer [26], a piecewise affine approximation is *reconstructed* from the cell average values u_j^{n+1}. Our standpoint is closer from the one in the discontinuous Galerkin finite element methods. As illustrated by (2.4), the values $u_h(t_{n+1}+)$ not only depend on u_j^{n+1} (i.e. the cell averages of $u_h(t_{n+1}-)$) but on the whole function $u_h(t_{n+1}-)$.

2.2 Entropy Diminishing Projections.

We now define the projection πv of any function $v \in L^\infty(0,1)^N$, which is the subject of this paper. The goal is to obtain a scheme that simultaneously

(1) is high-order accurate in the smooth regions of the solution,
(2) does not create spurious oscillations, that might prevent the strong convergence to a weak solution.

In order to achieve both properties, it is known that the projection should be a nonlinear operator.

The treatment of the scalar conservation laws is based on the property that the total variation (in space) of an entropy solution is a diminishing function of time. The concept of Total Variation Diminishing (TVD) scheme has been introduced with great success for the purpose of achieving (1) and (2) above. Recall that a uniform bound for the total variation is a sufficient condition for the scheme to converge in a strong topology. On the other hand, Harten [21] demonstrated that there exists *high order* accurate TVD schemes.

Considering next the numerical approximation of systems, we observe that the total variation of an entropy solution generally is not a diminishing function of time. The forthcoming analysis instead will be based on the entropy inequalities (1.3). As we shall see, the entropy inequalities provide natural stability properties for schemes of approximation for systems of conservation laws.

The distributional entropy inequality (1.3) implies that the entropy is a non-increasing function of time, i.e.

$$\int_{\mathbb{R}} U(u(x,t_2))\,dx \le \int_{\mathbb{R}} U(u(x,t_1))\,dx \qquad t_2 > t_1. \tag{2.7}$$

Letting $t_1 \to t_2 = t$ and at least when u is piecewise smooth, we arrice at the following inequality

$$\int_{\mathbb{R}} U(u(x,t+))\,dx \le \int_{\mathbb{R}} U(u(x,t-))\,dx, \qquad t > 0,$$

which can naturally be used a as building principle for numerical aproximation. More generally, one can deduce from (1.3), at least formally,

$$\int_{y_1}^{y_2} U(u(x,t+))\,dx \le \int_{y_1}^{y_2} U(u(x,t-))\,dx, \qquad y_1 < y_2,\ t > 0. \tag{2.8}$$

We suggest here to use (2.7)-(2.8) as a criterion for nonlinear stability: one should require (2.8) for suitable values of y_1 and y_2 (typically the mesh points) and *all* entropy functions of the system (1.1).

Let $\mathcal{E}$ be the set of all convex entropy functions of (1.1).

Definition 2.2. *The operator π is called an* entropy diminishing projection *if*

$$\int_0^1 U(\pi v)\,dx \ \le \int_0^1 U(v)\,dx \tag{2.9}$$

for any $U \in \mathcal{E}$ and $v \in L^\infty(0,1)^N$.

Roughly speaking the condition (2.9) means that π is a non-increasing operator for the "semi-norm" $||v||_U = \int_0^1 U(v)\,dx$. For second order projections πv of the form

$$\pi v(x) = \bar{v} + \beta(x - 1/2) \qquad \text{for some value } \beta \in \boldsymbol{R}^N, \tag{2.10}$$

where $\bar{v} = \int_0^1 v(x)\,dx$, we can prove the following result:

Proposition 2.3. *The set of all β in $\boldsymbol{R}^N$ satisfying (2.9) for all entropies $U \in \mathcal{E}$ is a non-empty, closed, and convex subset of $\boldsymbol{R}^N$.*

Proof of Proposition 2.3. For U in $\mathcal{E}$, we consider the real-valued function

$$H_U(\beta) = \int_0^1 U(\bar{v} + \beta(x - 1/2))\,dx. \tag{2.11}$$

Observe that H_U is a strictly convex function and $\nabla H_U(0) = 0$. Set $\bar{U} = \int_0^1 U(v)\,dx$. By Jensen's inequality, we have

$$H_U(0) = U(\bar{v}) \le \bar{U}.$$

Therefore the set $B_U = \{\beta \in \boldsymbol{R}^N \,/\, H_U(\beta) \le \bar{U}\}$ is a non-empty, closed, and convex subset of $\boldsymbol{R}^N$. The set of admissible β is the intersection of all B_U's, i.e.

$$\beta \in B \equiv \bigcap_{U \in \mathcal{E}} B_U.$$

Clearly B is a non-empty, closed, and convex subset of $\boldsymbol{R}^N$. □

Remark 2.4. 1. If v is a constant, then the Jensen's inequality in the proof above is a strict inequality, $B_U \setminus \{0\}$ is not empty, and the point $\beta = 0$ belongs to the interior of B_U. On the other hand, it is difficult to claim that the set B is non empty. It is conjectured that the second order of accuracy is obtained if β is chosen to be an extremal point of the set B and achieve $\max_{\beta \in B} |\beta|$.

2. We observe that (2.9) is equivalent to

$$C_U \cdot (\beta, \beta) \le \bar{U} - U(\bar{v}) \qquad \text{with} \qquad C_U = \int_0^1 \nabla^2 U(\pi v(x))(x - 1/2)^2\,dx. \tag{2.12}$$

3. For systems having few mathematical entropies, it might be necessary to require (2.9)-(2.10) for a broader class of functions U. To that purpose, $\mathcal{E}$ should be a cone of real-valued and convex functions defined on $\mathcal{U}$, containing the canonical projections on the axis of coordinates $g_i : \mathcal{U} \to \boldsymbol{R}, u = (u_1, u_2, \cdots, u_p) \to u_i$ $(i = 1, 2, \cdots, p)$. At least $\mathcal{E}$ should contain all of the convex mathematical entropies of (1.1).

Our next result (Theorem 2.7) provides a set of inequalities, that are sufficient to imply the condition (2.9) in Definition 2.2, but is explicit, and so can be used to construct projections for practical purposes. Consider the averaging operator $M : L^\infty(0,1)^N \to L^\infty(0,1)^N$ defined by

$$M(v)(x) = \frac{1}{x}\int_0^x v(y)\,dy, \qquad x \in (0,1),\, v \in L^\infty(0,1)^N. \tag{2.13}$$

The same formulation defines $M(\pi v)$. The quantity $M(\pi v) - M(v)$ will be used to estimate the interpolation error. Such averages can be found for instance in the ENO schemes by Harten-Engquist-Osher-Chakravarthy [22]. Note that $M(v) \in L^\infty(0,1)^N \cap W^{1,\infty}_{\mathrm{loc}}(0,1)^N$ and $\bar v \equiv \int_0^1 v(x)\,dx = M(v)(1)$. If v is a piecewise continuous function, then $M(v) \in W^{1,\infty}(0,1)^N$ and

$$\lim_{x\to 0} M(v)(x) = v(0+).$$

Since U is a convex function, one has

$$\begin{aligned}\int_0^1 U(\pi v)\,dx &\le \int_0^1 U(v)\,dx + \int_0^1 \nabla U(\pi v)\cdot\big(\pi v - v\big)\,dx \\ &\le \int_0^1 U(v)\,dx + \nabla U(\pi v(1))\cdot\big(M(\pi v)(1) - \bar v\big) \\ &\quad - \int_0^1 \nabla^2 U(\pi v)\cdot\left(\frac{d\,\pi v}{dx},\, M(\pi v) - M(v)\right)x\,dx,\end{aligned} \tag{2.14}$$

Proposition 2.5. *Assume that the operator π satisfies the condition*

$$M(\pi v)(1) = \bar v, \tag{2.15}$$

and for all $U \in \mathcal{E}$ and $v \in L^\infty(0,1)^N$

$$\nabla^2 U(\pi v)\cdot\left(\frac{d\,\pi v}{dx},\, M(\pi v) - M(v)\right) \ge 0 \quad \text{on } (0,1). \tag{2.16}$$

Then π is an entropy diminishing projection.

Remark 2.6. Using Kruzkov's entropy functions for scalar conservation laws, Bouchut-Bourdarias-Perthame [3] recently derived an inequality similar to (2.16). They proposed a second order TVD scheme consistent with all entropy inequalities (C.f. (2.20)).

Proposition 2.5 holds true for projections of any order of accuracy. We restrict now to second order and establish the existence of an entropy diminishing projection for arbitrary systems of conservation laws. Strict hyperbolicity of the system under consideration is assumed.

Theorem 2.7. *Consider continuous functions $v : (0,1) \to \boldsymbol{R}^N$ taking their values in a small neighborhood of a given state $v_* \in \mathcal{U}$. There exists an entropy diminishing projection π (consistent with all convex entropies of* (1.1)*) leading to a second-order Godunov-type scheme. For every function v, πv has the form* (2.10) *where the slope β is suitably chosen (C.f. proof below) in the closed, bounded, and convex set B of all vectors $\beta \in \boldsymbol{R}^N$ satisfying*

$$\beta_k(D_k - \beta_k) \geq 0, \quad k = 1, 2, \cdots, N \tag{2.17a}$$

with

$$\beta = \sum_{1 \leq i \leq N} \beta_i \, r_i(\pi v) \tag{2.17b}$$

and

$$D(v) = 2\, \frac{M(v) - \bar{v}}{x - 1} = \sum_{1 \leq i \leq N} D_i \, r_i(\pi v). \tag{2.18}$$

The set of inequalities (2.17) form a nonlinear system of coupled algebraic inequalities for the characteristic variables β_i. In the scalar case, the set B is an interval of $\boldsymbol{R}$ (possibly reduced to the point $\beta = 0$).

Remark 2.8. 1. Observe that no non-homogeneous slope limitation is *not* necessary in Theorem 2.7.

2. The slope β can be computed in practice by iterations. On the other hand, in a first approximation, it is possible to linearize (2.17) by replacing $r_i(\pi v)$ by $r_i(\bar{v})$: (2.17) then reduces to decoupled inequalities for the characteristic variables. Each such inequality coincides with the one given by another approach in [3] for scalar equations.

Proof of Theorem 2.7. First of all, we show how to derive (2.17) from (2.16). Dividing (2.16) by $1 - x$ and with the definition of $D(v)$ in (2.18) and straightforward calculations, we arrive at the inequality

$$\nabla^2 U(\pi v) \cdot \left(\frac{d\,\pi v}{dx}, \, D(v) - \frac{d\,\pi v}{dx} \right) \geq 0, \tag{2.19}$$

where we have used that $D(\pi v) = d(\pi v)/dx$. It remains to observe that, if U is an entropy function and since the system under consideration is strictly hyperbolic, the matrix $\left\{ \nabla^2 U \cdot \left(, r_i, r_j \right) \right\}_{1 \leq i,j \leq N}$ is a diagonal matrix with non negative diagonal coefficients.

Next we establish the existence of a solution β to the inequalities (2.17).

STEP 1: Consider a function v taking its values in a small ball around the constant state v_*. In particular the eigenvectors, eigenvalues, etc of the Jacobian matrix $\nabla f(u)$, for all u arising thereafter, will be always well-defined and in a distance of order $O(\epsilon)$ (ϵ small) from the eigenvectors of $\nabla f(v_*)$, i.e. for instance

$$|r_j(v) - r_j(v')| \; = \; O(\epsilon). \tag{2.20}$$

On the other hand, we observe that the inequalities (2.17) provide constraints on the projection operator for arbitrary functions v, but that $\beta = 0$ is always a trivial solution to that problem. A second constraint is that the projection should yield a second order accurate scheme. The later condition only concerns smooth functions v. And since a scaling has been used to transform the projection operator acting on functions defined on a shrinking interval, say $(0,h)$, to those defined on the (fixed) interval $(0,1)$, this constraint clearly concerns functions that are close to a constant state. The condition can be formulated as follows. We introduce an additional parameter $\delta << 1$ which plays the role of the mesh size h, We require that, for every smooth function v of the form $v(x) = \bar{v} + \delta\, w(x)$, the associated slope β should have the form $\beta = \delta\,\gamma$, and

$$\gamma_k = Dw_k(x_k) + O(\delta) \tag{2.21}$$

for every k and some $x_k \in (0,1)$. With this scaling, we also obtain

$$Mv(x) = \bar{v} + \delta M w(x), \qquad Dv(x) = \delta\, Dw(x).$$

The notation (2.18) is used here. When considering the projection π_h on $(0,h)$ associated with π. the condition (2.21) ensures that β is a first order approximation of dv/dx when v is smooth (so that Taylor expansion can be used).

Therefore we are led to the following problem. We are given a function $w : [0,1] \to \boldsymbol{R}^N$ smooth enough and such that

$$|Dw_k(x)| \le 1, \quad x \in [0,1], \tag{2.22}$$

where the Dw_k's are the characteristic coordinates on the constant basis $\big\{r_k(\bar{v})\big\}$, i.e.

$$Dw(x) = \sum_{k=1}^{N} Dw_k(x)\, r_k(\bar{v}).$$

Denote by Δ_1 the set of vectors of $\boldsymbol{R}^N$ characterized by the inequalities (2.22), and by Δ_2 the set defined by (2.22) with 1 replaced by 2 in the right hand side. We search for a slope $\gamma \in \Delta_2$, i.e.

$$\gamma = \sum_{k=1}^{N} \gamma_k\, r_k(\bar{v}), \qquad |\gamma_k| \le 2,$$

so that, for every $k = 1, 2, \cdots, N$, each function

$$\Phi_k(\gamma, x) = \gamma \cdot l_k\big(\bar{v} + (x - 1/2)\delta\gamma\big)\ \big(Dw(x) - \gamma\big) \cdot l_k\big(\bar{v} + (x-1/2)\delta\gamma\big)$$

would be non-negative for all $x \in [0,1]$. For convenience we are using here a decomposition over a constant set of eigenvectors $r_j(\bar{v})$ instead of the one in (2.17b).

STEP 2: First of all, we show that the problem under consideration has a very similar structure to the one in the scalar case ($N = 1$). To better describe these positivity conditions, let us observe that the function $\gamma_k \to \Phi_k(\gamma_1, \gamma_2, \cdots, \gamma_N, x)$ is strictly concave, at least for δ small enough. This is readily checked by observing that this function has the form

$$\begin{aligned} \Phi_k(\gamma, x) &= Q_k(\gamma_k, x) + \delta\, G(\gamma, x) \\ Q_k(\gamma_k, x) &= \gamma_k (Dw_k(x) - \gamma_k), \end{aligned}$$

which is concave w.r.t. γ_k for $\delta << 1$, $|w_i(x)| \le 1$, $|\gamma_i| \le 2$, and any bounded function G.

Since the minimum of a one-parameter family of concave functions remains concave, it follows that the function

$$\Phi_k(\gamma) = \min_{x\in[0,1]} \Phi_k(\gamma, x)$$

is a strictly concave function of the variable γ_k. The function $\gamma \to \Phi_k$ is Lipschitz continuous, at least, since the function "min" is Lipschitz continuous. Furthermore it is not hard to check that

$$\lim_{\gamma_k\to\pm 2} \Phi_k(\gamma) < 0.$$

Actually this inequality is true for the function $\Phi_k(\gamma, x)$: one has $\Phi_k(\gamma, x) \le \gamma_k\,(Dw_k(x) - \gamma_k) + O(\delta)$, therefore as $\gamma_k \to \pm 2$ the right handside of this inequality is negative since $|w_k(x)| \le 1$.

As a consequence, the function $\gamma_k \to \Phi_k(\gamma)$ admits at most two roots, but possibly only one or no root. We can choose γ to be such a root so that (2.21) is satisfied. Intuitively, one shall define γ to be the root with the largest modulus when they are two, and take $w_k = 0$ when no root exists.

STEP 3: We now give a rigorous definition for the value γ.

Let $\gamma^* \in \boldsymbol{R}^N$ be the slope constructed by keeping the quadratic part of Φ_k, i.e. Q_k introduced in Step 2:

$$\gamma_k^* = minmod_{x\in(0,1)} Dw_k(x)$$

for $k = 1, 2, \cdots, N$. Here *minmod* is the following extension of the classical *minmod* function (C.f. also Section 3.1 below):

$$minmod_{0\le x\le 1}\Big\{g(x)\Big\} = \begin{cases} \min g & \text{if the function } g > 0,\\ \max g & \text{if the function } g < 0,\\ 0 & \text{otherwise.}\end{cases}$$

Observe that γ^* is computed component after component. The (true) roots, those of Φ_k, can be checked to be at a distance of at most $O(\delta)$ from the γ_k^*'s. On the other hand, we have by construction (2.21) for γ^*, i.e.

$$\gamma_k^* = Dw_k(x_k) \tag{2.23}$$

for every k and some x_k.

The idea next is to correct the value γ_* and set

$$\gamma = \gamma^* + \delta\nu,$$

where ν is chosen so that to preserve the correct signs in the inequalities under consideration. In view of $\gamma = \gamma^* + O(\delta)$ and (2.23), the property (2.21) holds for γ, and the second order of accuracy is achieved.

It remains to choose ν. Consider first the case that $\gamma^* \cdot l_k(\bar v) > 0$ and let ν be such that

$$\nu \cdot l_k(\bar v) \; > \; 2 \sup_{1\le k\le N} |\nabla l_k(\bar v)|, \tag{2.24}$$

where the right handside above involves the Euclidian norm of the matrix-valued functions ∇l_k. Since $w \in \Delta_1$ and $\gamma \in \Delta_2$, we get for δ small enough

$$\gamma \cdot l_k\big(\bar v + (x-1/2)\delta\gamma\big) = \gamma^* \cdot l_k(\bar v) + \delta\Big(\nu \cdot l_k(\bar v) + (x-1/2)(w(x)-\gamma^*)\cdot \nabla l_k(\bar v)\cdot \gamma^*\Big) + O(\delta^2)$$

has the same sign as $\gamma^* \cdot l_k(\bar v) > 0$ since, in view of (2.24),

$$\nu \cdot l_k(\bar v) + (x-1/2)(w(x)-\gamma^*)\cdot \nabla l_k(\bar v)\cdot \gamma^* > 0.$$

Similarly

$$\begin{aligned}&\big(w(x)-\gamma\big)\cdot l_k\big(\bar v + (x-1/2)\delta\gamma\big)\\ &= \big(w(x)-\gamma^*\big)\cdot l_k(\bar v) + \delta\Big(\nu \cdot l_k(\bar v) + (x-1/2)(w(x)-\gamma^*)\cdot \nabla l_k(\bar v)\cdot \gamma^*\Big) + O(\delta^2)\end{aligned}$$

has the same sign as $\big(w(x)-\gamma^*\big)\cdot l_k(\bar v) > 0$. Since γ^* by construction satisfies

$$\gamma^* \cdot l_k(\bar v)\,\big(w(x)-\gamma^*\big)\cdot l_k(\bar v) \;\geq\; 0,$$

it follows that γ satisfies (2.17).

The case that $\gamma^* \cdot l_k(\bar v) < 0$ is analogous. Consider next the case that $\gamma^* \cdot l_k(\bar v) = 0$, which happens when the function $w(x)\cdot l_k(\bar v)$ changes its sign for $x \in (0,1)$. It is then necessary to choose ν such that for all $x \in (0,1)$

$$\big(\gamma^* + \delta\nu\big)\cdot l_k\big(\bar v + (x-1/2)\delta\,(\gamma^* + \delta\nu)\big) = 0.$$

(This condition implies that $\gamma = 0$ except when l_k are constants, say.) The proof is complete. □

2.3 Discrete Entropy Inequalities.

Finally we give the main consequences of Definition 2.2 and discuss the convergence of the high order Godunov-type scheme.

Theorem 2.9. *Assume that π is an entropy diminishing projection in the sense of* Definition 2.2 *and consider the approximations u_h constructed by the corresponding Godunov-type scheme. Then, for any entropy pair (U,F), the discrete entropy inequality*

$$\frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(x,t_{n+1}+))\,dx - \frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(x,t_n+))\,dx + \lambda\left(G^n_{j+1/2} - G^n_{j-1/2}\right) \leq 0 \tag{2.25}$$

holds true, where

$$G^n_{j+1/2} = \frac{1}{\tau}\int_{t_n}^{t_{n+1}} F(u_h(x_{j+1/2},t))\,dt \tag{2.26}$$

is a numerical entropy flux consistent with F. Suppose that the scheme is stable in the L^∞ norm.

(1) *Then any Young measure associated with the u_h's is an entropy measure-valued solution to* (1.1)–(1.3) *in the sense of DiPerna.*
(2) *When the compensated compactness method applies (e.g. $p = 2$), the high-order Godunov-type converges strongly.*

No strong compactness property is a priori available. The proof of Theorem 2.9 will be based on the following *entropy stability* property.

Proposition 2.10. *For every $U \in \mathcal{E}$, the scheme satisfies the estimate*

$$\int_{\mathbb{R}} U(u_h(x, t_n+))\, dx + E_h(t_n+) \le \int_{\mathbb{R}} U(u_0)\, dx \tag{2.27}$$

with

$$E_h(t_n+) = \frac{1}{2} \sum_{0 \le q \le n} \int_{\mathbb{R}} C_U^h \cdot \big(u_h(t_q+) - u_h(t_q-), u_h(t_q+) - u_h(t_q-)\big)\, dx \tag{2.28a}$$

and

$$C_U^h = \int_0^1 \nabla^2 U\big(u_h(t_n-) + s(u_h(t_n+) - u_h(t_n-))\big)\,(1-s)\, ds. \tag{2.28b}$$

Using a uniformly convex entropy function, one deduces from Proposition 2.10:

$$\sum_{0 \le n \le \infty} \int_{\mathbb{R}} \big|u_h(t_n+) - u_h(t_n-)\big|^2\, dx \le O(1). \tag{2.29}$$

Such an *entropy dissipation* estimate was first derived, for multidimensional scalar equations, in [10, 12, 13].

Proof of Proposition 2.10. The scheme consists of two steps. The first one is based on solving a generalized Riemann problem. Since entropy solutions are used, the distributional inequality (1.3) holds and after integration (C.f. (2.26) for the definition of the entropy-flux) one gets

$$\frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(t_{n+1}-, x))\, dx - \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(t_n+, x))\, dx + \lambda\big(G_{j+1/2}^n - G_{j-1/2}^n\big) \le 0. \tag{2.30}$$

The second step is based on an entropy diminishing projection operator π_h. Let us revisit the derivation of inequality (2.14). The following is indeed a sharper version of (2.14):

$$\begin{aligned}
&\int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(t_{n+1}+, x))\, dx - \int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(t_{n+1}-, x))\, dx \\
&= - \int_{x_{j-1/2}}^{x_{j+1/2}} \nabla^2 U(u_h(t_{n+1}+, x)) \cdot \big(u_h(t_{n+1}-, x) - u_h(t_{n+1}+, x)\big)\, dx \\
&\quad - \frac{1}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} C_U^h \cdot \big(u_h(t_{n+1}+, x) - u_h(t_{n+1}-, x), u_h(t_{n+1}+, x) - u_h(t_{n+1}-, x)\big)\, dx.
\end{aligned}$$

Our requirement (2.16) reads here

$$- \int_{x_{j-1/2}}^{x_{j+1/2}} \nabla^2 U(u_h(t_{n+1}+, x)) \cdot \big(u_h(t_{n+1}-, x) - u_h(t_{n+1}+, x)\big)\, dx \le 0.$$

Therefore the discrete entropy inequality holds:

$$\begin{aligned}
&\frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(t_{n+1}+, x)\, dx - \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} U(u_h(t_n+, x)\, dx + \lambda\big(G_{j+1/2}^n - G_{j-1/2}^n\big) \\
&\le -\frac{1}{2h} \int_{x_{j-1/2}}^{x_{j+1/2}} C_U^h \cdot \big(u_h(t_{n+1}+, x) - u_h(t_{n+1}-, x), u_h(t_{n+1}+, x) - u_h(t_{n+1}-, x)\big)\, dx.
\end{aligned}$$

Since U is a convex function, the right hand side of this inequality remains non-positive. The desired result follows by summation over space and time. □

Proof of Theorem 2.9. Since π is taken to be an entropy diminishing projection, the inequality (2.30) implies the discrete entropy inequality (2.25) for any convex entropy pair (U, F).

Assuming the scheme to be stable in the L^∞ norm, we consider a Young measure associated with u_h, that is a probability measure $\nu_{x,t}$ such that

$$\iint g(u_h(x,t))\,\varphi(x,t)\,dtdx \to \iint <\nu_{x,t)}, g> \ \varphi(x,t)\,dtdx \tag{2.31}$$

for all continuous function g and all test-function φ. Let φ be a non-negative test-function. Since u_h satisfies the inequality (1.3) in each strip $t_n < t < t_{n+1}$, one has

$$\begin{aligned}&\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}} U(u_h)\,\partial_t\varphi + F(u_h)\partial_x\varphi\,dtdx \\ &- \int_{\mathbb{R}} U(u_h(t_{n+1}-))\,\varphi(t_{n+1})\,dx + \int_{\mathbb{R}} U(u_h(t_n+))\,\varphi(t_n)\,dx \ \geq 0.\end{aligned}$$

After summation over time, we arrive at

$$\begin{aligned}0 \leq &\iint U(u_h)\,\partial_t\varphi + F(u_h)\,\partial_x\varphi\,dtdx \\ + &\sum_{n=0,1,2,\ldots} \int_{\mathbb{R}} \big(U(u_h(t_n+)) - U(u_h(t_n-))\big)\,\varphi(t_n)\,dx \ \equiv A_h^1 + A_h^2.\end{aligned} \tag{2.32}$$

In view of (2.31), one has

$$A_h^1 \to \iint <\nu, U> \ \partial_t\varphi + <\nu, F> \ \partial_x\varphi\,dtdx.$$

We write $A_h^2 = A_h^{2,1} + A_h^{2,2}$ with

$$\begin{aligned}A_h^{2,1} &= \sum_{n,j} \int_{x_{j-1/2}}^{x_{j+1/2}} \big(U(u_h(t_n+)) - U(u_h(t_n-))\big)\,\big(\varphi(t_n) - \varphi(t_n, x_j)\big)\,dx, \\ A_h^{2,2} &= \sum_{n,j} \varphi(t_n, x_j) \int_{x_{j-1/2}}^{x_{j+1/2}} \big(U(u_h(t_n+)) - U(u_h(t_n-))\big)\,dx.\end{aligned}$$

The diminishing entropy criterion (2.9) again yields

$$A_h^{2,2} \leq 0.$$

Using the estimate (2.29), one has

$$\begin{aligned}|A_h^{2,1}| &\leq \tau^{1/2}\,\|\nabla U\|_{L^\infty}\ \|\phi\|_{H_0^1}\,\big(\sum_{0\leq q\leq\infty} \int_{\mathbb{R}} |u_h(t_q+) - u_h(t_q-)|^2\,dx\big)^{1/2} \\ &\leq \ O(h^{1/2}).\end{aligned}$$

This proves the claim (1) in Theorem 2.9. In view of the full set of discrete entropy inequalities (2.25)-(2.26) and the entropy dissipation estimate (2.29), it is not hard to the compensated compactness method along the lines of DiPerna's work [17]. C.f. also Chen-Liu [8] and Coquel-LeFloch [12]. This completes the proof of Theorem 2.9. □

3. Applications.

This section focuses on the second order entropy diminishing projections. The convergence results below concern the second order Godunov-type scheme introduced in Section 2. Both conditions (2.16) and (2.17) will be dealt with. One of our main result is that the two conditions are equivalent when a large number of entropies are available: scalar equations, linear systems, and systems of two equations. For the full gas dynamics system however, the condition (2.16) is less restrictive than (2.17).

3.1 Linear Hyperbolic Systems.

In this section, we want to highlight some of the main features of our method and thus restrict attention to linear systems. Linear systems give rise to entropy diminishing projections that share most of the basic features encountered in the general case. Because of their very simplicity, linear systems lead to fully explicit formulas, which assess the role played by the characteristic variables in the entropy diminishing projection.

Let us begin with several notations. We consider an hyperbolic system with a flux-function of the form $f(u) = A\,u$, where A is a $N \times N$ matrix having real and distinct eigenvalues

$$\lambda_1 < ... < \lambda_{N_0} \leq 0 < \lambda_{N_0+1} < ... < \lambda_N$$

with N_0 fixed. Let $\{r_i\}_{1\leq ileN}$ and $\{l_i\}_{1\leq ileN}$ be a basis of right and left eigenvectors of A, respectively. We denote by $\alpha_i(u) \equiv l_i \cdot u$ the coordinates of a vector u on the basis of right eigenvectors. We suppose the standard CFL condition $\max_{1\leq i\leq N} |\lambda_i|\tau/h < 1$.

Solving generalized Riemann problems in the first step of the scheme is trivial. For $t_n < t < t_{n+1}$ and $x \in (x_{j-1/2}, x_{j+1/2})$, we obtain

$$\begin{aligned} u_h(x,t) = &\sum_{1\leq i\leq N_0} \alpha_i\big(u_h(x - x_{j+1/2} - \lambda_i t, t_n+)\big)\, r_i \\ &+ \sum_{N_0+1\leq i\leq N} \alpha_i\big(u_h(x - x_{j-1/2} - \lambda_i t, t_n+)\big)\, r_i, \end{aligned} \tag{3.1}$$

where the characteristic variables $\alpha_i(u) = l_i \cdot u$ play a central role in what follows.

The second step, i.e. projecting the solution $u_h(t_{n+1}-)$ is the subject of the following result. The following notation will be used. For each interval $(x_{j-1/2}, x_{j+1/2})$, the average value and the value at the two end points are decomposed as follows

$$u_j^{n+1} = \sum_{1\leq i\leq N} \alpha_{ji}^{n+1}\, r_i, \qquad u_h(x_{j+1/2}, t_{n+1}-) = \sum_{1\leq i\leq N} \alpha_{j+1/2,i}^{n+1}\, r_i,$$

and the slope β_j^{n+1} in the linear projection

$$u_h(x, t_{n+1}+) \equiv \pi_h u_h(x, t_{n+1}-) = u_j^{n+1} + \beta_j^{n+1}(x - x_j)$$

is decomposed in the form

$$\beta_j^{n+1} = \sum_{1\leq i\leq N} \beta_{ji}^{n+1}\, r_i.$$

Finally, the classical *minmod* function is

$$minmod\Big\{a,b,c\Big\} = \begin{cases} \min(a,b,c) & \text{if } a,b,c > 0, \\ \max(a,b,c) & \text{if } a,b,c < 0, \\ 0 \quad \text{otherwise.} \end{cases}$$

Theorem 3.1. *Consider a strictly hyperbolic, linear system.*

(1) *The conditions* (2.16) *and* (2.17) *in Proposition 2.5 and Theorem 2.7 respectively are* equivalent.

(2) *The following explicit formulas describe the (nonlinear and second order) entropy diminishing projection* (2.16), (2.17)*: for* $1 \le i \le N$

$$\beta_{ji}^{n+1} = \frac{2}{h}\, minmod\Big\{\alpha_{j+1/2,i}^{n+1} - \alpha_{ji}^{n+1}, \beta_{ji}^{n+1,\star}, \alpha_{ji}^{n+1} - \alpha_{j-1/2,i}^{n+1}\Big\}, \tag{3.2}$$

where $\beta_{ji}^{n+1,\star}$ *is given by*

$$\beta_{ji}^{n+1,\star} = \big(1 - \lambda_i \frac{\tau}{h}\big)\big(\alpha_{ji}^{n+1} - l_i \cdot u_h(x_{j-1/2} + \frac{\lambda_i \tau}{2}, t_{n+1}-)\big). \tag{3.3}$$

(3) *The corresponding second order Godunov-type scheme is stable in the* L^∞ *and* BV *norms, and converges strongly to the unique solution of* (1.1), (1.2).

To prove this statement, we need the following result which extends aresult in [3] to linear systems. Let us introduce an extension of the classical *minmod* function:

$$minmod_{0\le x\le 1}\Big\{g(x)\Big\} = \begin{cases} \min g & \text{if } g > 0, \\ \max g & \text{if } g < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

By $g > 0$, it is meant that $g(x) > 0$ for all $x \in (0,1)$. We use here the operator D defined by (2.19).

Lemma 3.2. *Let* v *be given in* $L^\infty(0,1)^N$*. Then the condition* (2.16) *for the projection* π *is equivalent to*

$$\begin{aligned} &\pi v(x) = \bar{v} + \beta(x - 1/2), \\ &\beta_i = \theta_i\, minmod_{0\le x\le 1}\Big\{l_i \cdot D(v)(x)\Big\}, \quad 1 \le i \le N, \end{aligned} \tag{3.5}$$

with $\theta_i \in (0,1)$ *arbitrary. Moreover, the second order of accuracy is achieved with the choice* $\theta_i = 1$, $1 \le i \le N$.

Proof of Lemma 3.2. The convex entropies associated with a strictly hyperbolic linear system are of the form

$$U(v) = \sum_{1\le i\le N} \psi_i\big(\alpha_i(v)\big) \quad \text{with} \quad v = \sum_{1\le i\le N} \alpha_i(v) r_i,$$

where the ψ_i's are arbitrary convex functions. The Hessian matrix of such a function writes

$$\nabla^2 U(v) = \sum_{1\le i\le N} \psi_i''\big(\alpha_i(v)\big) l_i \otimes l_i,$$

so that our requirement (2.16) is exactly equivalent to

$$\sum_{1\le i\le N} \psi_i''\big(\alpha_i(\pi v)\big)\beta_i\big(\frac{\beta_i - \alpha_i(D(v))(x)}{1-x}\big) \ge 0,$$

where we have used (2.19). To enforce the above inequality for all entropies, it is necessary (but also sufficient) to consider the functions $U_i(v) = 1/2\big(\alpha_i(v)\big)^2$, $1 \le i \le N$. This leads to a system of decoupled inequalities:

$$\beta_i\Big(\frac{\beta_i - \alpha_i(D(v))(x)}{1-x}\Big) \ge 0, \quad 1 \le i \le N. \tag{3.6}$$

As a consequence, entropy diminishing projections are defined in term of characteristic variables by

$$\beta_i = \theta_i \, minmod_{0\le x\le 1}\Big\{\alpha_i(D(v)(x))\Big\}, \quad \theta_i \in (0,1),\ 1 \le i \le N.$$

This completes the proof of Lemma 3.2. □

Proof of Theorem 3.1. Based on the explicit formula for u_h, C.f. (3.1), a straightforward calculation shows that, for $1 \le i \le N$, the function

$$x \to l_i \cdot D\big(\sigma_{h,j} u_h(t_{n+1}-)\big)(x)$$

is a *monotone* function on each of the intervals $(0, \lambda_i \frac{\tau}{h})$ and $(\lambda_i \frac{\tau}{h}, 1)$. Since it is continuous (C.f. its definition in (2.19)), it achieves it extremum values at either 0, 1, or $\lambda_i \frac{\tau}{h}$. Therefore we deduce (3.2), (3.3) from (3.4) and the additional observation

$$l_i \cdot D\big(\sigma_{h,j} u_h(t_{n+1}-)\big)(\lambda_i \frac{\tau}{h}) = 2 l_i \cdot \Big(\frac{u_j^{n+1} - \sigma_{h,j} u_h(t_{n+1}-)(\lambda_i \frac{\tau}{2h})}{1 - \lambda_i \frac{\tau}{h}}\Big).$$

The proof of the stability properties can be obtained by using on each characteristic variable the proof in [3] which treated scalar equations. It can be checked that

$$\min_{0\le x\le 1} \alpha_i\big(u_h(x, t_{n+1}-)\big) \le \alpha_i\big(u_h(x, t_{n+1}+)\big) \le \max_{0\le x\le 1} \alpha_i\big(u_h(x, t_{n+1}-)\big)$$

and

$$\int_{\boldsymbol{R}} |\frac{d}{dx}\alpha_i\big(u_h(x, t_{n+1}+)\big)|\, dx \le \int_{\boldsymbol{R}} |\frac{d}{dx}\alpha_i\big(u_h(x, t_{n+1}-)\big)|\, dx.$$

Obvious properties of the generalized Riemann problem for a linear hyperbolic system (C.f. in particular (3.1)) yield

$$\min_{0\le x\le 1} \alpha_i\big(u_0(x)\big) \le \alpha_i\big(u_h(x, t_{n+1}+)\big) \le \max_{0\le x\le 1} \alpha_i\big(u_0(x)\big) \tag{3.7}$$

and

$$\int_{\boldsymbol{R}} |\frac{d}{dx}\alpha_i\big(u_h(x, t_{n+1}+)\big)|\, dx \le \int_{\boldsymbol{R}} |\frac{d}{dx}\alpha_i\big(u_0(x)\big)|\, dx. \tag{3.8}$$

This provides the necessary estimates to conclude the strong convergence of the scheme, which completes the proof of Theorem 3.1. □

3.2 Systems of Two Equations.

Systems of two conservation laws admit infinitely many entropy functions. In the present subsection, we take advantage of this property to analyse more closely the scheme from Section 2 for systems of two equations, especially regarding the convex invariant domains.

Consider the following 2×2 system

$$\begin{aligned} &\partial_t u_1 + \partial_x f_1(u_1, u_2) = 0, \\ &\partial_t u_2 + \partial_x f_2(u_1, u_2) = 0, \end{aligned} \tag{3.9}$$

where $f = (f_1, f_2)$ satisfies $f_{1,2} f_{2,1} > 0$ (C.f. Smoller [37]). Here we have set $f_{i,j} = \partial f_i / \partial u_j$. This assumption implies that (3.9) is strictly hyperbolic. We consider a set of two Riemann invariants $\big(\Phi_1, \Phi_2\big)$, which by definition diagonalizes (3.9). Denote by $u(0)$ the initial condition at the time $t = 0$. Consider the following two assumptions

(1) For $i = 1, 2$ the i-Riemann invariant

$$\Phi_i \text{ is either convex or concave} \tag{3.10}$$

with respect to the conservative variables.

(2) For $i = 1, 2$, there exists one convex entropy function $U_{\star i}$ with the property

$$\begin{aligned} &U_{\star i}(u) \geq 0, \\ &U_{\star i}(u(0)) = 0 \\ &U_{\star i}(u) > 0 \quad \text{ for } |\Phi_i(u)| > C_i \end{aligned} \tag{3.11}$$

for some constants $C_i > 0$.

Under classical assumptions on the equation of state, the p-system of gas dynamics satisfies (3.10), while the isentropic Euler system satisfies (3.11) (C.f. [29] for a construction of the entropy and also [36] for the nonlinear elasticity system).

Theorem 3.3. *Consider a strictly hyperbolic system of two equations with two genuinely nonlinear characteristic fields,* (3.9).

(1) *The inequalities* (2.16) *and* (2.17) *are* equivalent*, and, for any function v in $L^\infty(0,1)^2$, read as*

$$\frac{d}{dx}\Phi_i(\pi v(x)) \Big\{ < \nabla_u \Phi_i(\pi v(x)), 2\frac{\bar{v} - M(v)(x)}{1 - x} > - \frac{d\Phi_i(\pi v)}{dx} \Big\} \geq 0. \tag{3.12}$$

for $x \in (0,1)$ and $i = 1, 2$.

(2) *Suppose that the system is endowed with bounded and convex invariant domains and one of the two assumptions* (3.10) *or* (3.11) *holds. Then the corresponding entropy diminishing Godunov-type scheme preserves the invariant domains of* (3.9).

This theorem provides the L^∞ stability of the scheme. From Theorem 2.9 and the results in [6, 8, 17], we immediately deduce that the second order Godunov-type scheme converges strongly. It is not hard to extend the convergence analysis to the isentropic Euler system of gas dynamics for a polytropic perfect gas.

Proof of Theorem 3.3. *Step 1: Proof of (1).* For any given entropy U of (3.9), the assumption of strict hyperbolicity implies that $\nabla^2 U(u) r_i(u),\ i = 1, 2$ is a left eigenvector of the matrix $\nabla f(u)$ associated with the eigenvalue $\lambda_i(u)$. Therefore we have

$$\nabla^2 U(u) \cdot \Big(r_1(u), r_2(u) \Big) = 0 \quad \text{for all } u \in \mathcal{U}.$$

With obvious notation, this identity reads

$$U_{,11}(u)r_{11}(u)r_{21}(u) + U_{,22}(u)r_{12}(u)r_{21}(u) + U_{,12}(u)\big(r_{11}(u)r_{22}(u) + r_{12}(u)r_{21}(u)\big) = 0.$$

The assumption $f_{1,2}(u)f_{2,1}(u) > 0$ guarantees that $r_{12}(u)r_{22}(u)$ never vanishes, so the latter can be rewritten as

$$U_{,11}\frac{r_{11}}{r_{12}}\frac{r_{21}}{r_{22}} + U_{,12}\big(\frac{r_{11}}{r_{12}} + \frac{r_{21}}{r_{22}}\big) = -U_{,22}. \tag{3.13}$$

Since U is convex, we have

$$U_{,12}^2 - U_{,11}U_{,22} \le 0,$$

so the following inequality holds:

$$U_{,12}^2 + \big(\frac{r_{11}}{r_{12}} + \frac{r_{21}}{r_{22}}\big)U_{,11}U_{,12} + \frac{r_{11}}{r_{12}}\frac{r_{21}}{r_{22}}U_{,11}^2 \le 0.$$

Any convex entropy U satisfies the property

$$U_{,12}(u) \in \Big(-\frac{r_{11}(u)}{r_{12}(u)}U_{,11}(u), -\frac{r_{21}(u)}{r_{22}(u)}U_{,11}(u)\Big) \quad \text{for all } u \in \mathcal{U}.$$

Therefore there exists a function $\alpha : \mathcal{U} \to [0,1]$ depending on U so that, in one hand,

$$U_{,12}(u) = \alpha(u)\big\{\kappa_1(u)U_{,11}(u)\big\} + (1-\alpha(u))\big\{\kappa_2(u)U_{,11}(u)\big\} \quad \text{for all } \ u \in \mathcal{U} \tag{3.14}$$

with $\kappa_i(u) = -r_{i1}(u)/r_{i2}(u), \ i = 1, 2$. On the other hand, (3.13) yields

$$U_{,22}(u) = \alpha(u)\big\{\kappa_1^2(u)U_{,11}(u)\big\} + (1-\alpha(u))\big\{\kappa_2^2(u)U_{,11}(u)\big\} \quad \text{for all } u \in \mathcal{U}. \tag{3.15}$$

Equipped with the two relations (3.14) and (3.15), we now rewrite the condition (2.16). For an arbitrary function v, one gets

$$\begin{aligned}\frac{d\pi v}{dx}\cdot\nabla^2 U(\pi v) = {}& \alpha(\pi v)\Big(\frac{d\pi w}{dx} + \kappa_1(\pi v)\frac{d\pi z}{dx}, \kappa_1(\pi v)\big(\frac{d\pi w}{dx} + \kappa_1(\pi v)\frac{d\pi z}{dx}\big)\Big)^T \\ & + (1-\alpha(\pi v))\Big(\frac{d\pi w}{dx} + \kappa_2(\pi v)\frac{d\pi z}{dx}, \kappa_2(\pi v)\big(\frac{d\pi w}{dx} + \kappa_2(\pi u)\frac{d\pi z}{dx}\big)\Big)^T.\end{aligned}$$

Since the two Riemann invariants satisfy $r_i(u)\cdot\nabla_u\Phi_i(u) = 0$, $i = 1, 2$, we obtain

$$\frac{\Phi_{i,2}(u)}{\Phi_{i,1}(u)} = -\frac{r_{i1}(u)}{r_{i2}(u)} = \kappa_i(u), \quad i = 1, 2$$

so

$$\begin{aligned}\frac{d\pi v}{dx}\cdot\nabla^2 U(\pi v) = {}&\alpha(\pi v)\Phi_{1,1}^{-2}(\pi v)\frac{d\Phi_1}{dx}(\pi v)\cdot\nabla\Phi_1(\widehat{\mathbf{w}}) \\ & + (1-\alpha(\pi v))\Phi_{2,1}^{-2}(\pi v)\frac{d\Phi_2}{dx}(\pi v)\cdot\nabla\Phi_2(\pi v).\end{aligned}$$

The condition (2.16) is equivalent to

$$\begin{aligned}&\alpha(\pi v)\Phi_{1,u}^{-2}(\pi v)\Big\{\frac{d\Phi_1}{dx}(\pi v) < \nabla\Phi_1(\pi v), \frac{M(\pi v)-M(v)}{1-x} > \Big\}\\ &+(1-\alpha(\pi v))\Phi_{2,u}^{-2}(\pi v)\Big\{\frac{d\Phi_2}{dx}(\pi v) < \nabla\Phi_2(\pi v), \frac{M(\pi v)-M(v)}{1-x} > \Big\} \geq 0.\end{aligned} \tag{3.16}$$

For any given $x \in (0,1)$, we consider two special entropy functions satisfying either

$$\alpha((\pi v)(x)) = 0 \quad \text{or} \quad \alpha((\pi v)(x)) = 1.$$

(Such entropies exists for general systems, at least locally in the phase space.) It follows that each term in the left hand side of (3.16) should be non-negative.

Step 2: Proof of (2). We first treat the case of the assumption (3.10). With no loss of generality, let us suppose that each $\Phi_i,\ i=1,2$ is a convex function of u. In what follows and depending on the sign of (the increasing function) $\frac{d\Phi_i(\pi v(x))}{dx}$, we shall need to consider one of the following two averages of v:

$$M_0(v)(x) = \frac{1}{x}\int_0^x v(y)dy, \quad M_1(v)(x) = \frac{1}{1-x}\int_x^1 v(y)dy. \tag{3.17}$$

Since $M_0(v)(1) = M_0(\pi v)(1) = \bar{v}$ and $M_0(v)(1) = M_0(\pi v)(1) = \bar{v}$ (see (2.15)), it can be easily checked that, for $0 \leq x \leq 1$,

$$\frac{M_0(v)(x) - M_0(\pi v)}{x-1} = \frac{M_1(v)(x) - M_1(\pi v)(x)}{x},$$

so that (3.12) (and so (2.16) and (2.17)) writes equivalently, for $i=1,2$,

$$\frac{d\Phi_i(\pi v)}{dx} < \nabla_u\Phi_i(\pi v), \frac{M_0(v)(x) - M_0(\pi v)(x)}{x-1} >\geq 0 \tag{3.18a}$$

or

$$\frac{d\Phi_i(\pi v)}{dx} < \nabla_u\Phi_i(\pi v), \frac{M_1(v)(x) - M_1(\pi v)(x)}{x} >\geq 0. \tag{3.18b}$$

Let $x \in (0,1)$ be given and, to begin with, let us suppose that $\frac{\Phi_i(d\pi v(x))}{dx} > 0$ for $i=1$ or $i=2$. Then choosing $M_1(v)(x)$ to express our requirement on the diminishing entropy projections, we have from (3.18b)

$$< \nabla_u\Phi_i(\pi v), \frac{M_1(v)(x) - M_1(\pi v)(x)}{x} >\geq 0,$$

so that

$$\begin{aligned}< \nabla_u\Phi_i(\pi v), M_1(v)(x) - \pi v(x) > &\geq < \nabla_u\Phi_i(\pi v), M_1(\pi v)(x) - \pi v(x) >\\ &= \frac{1}{2}\frac{d\Phi_i(\pi v)}{dx}(1-x),\end{aligned} \tag{3.19}$$

where we have used the identity $M_1(\pi v)(x) = \pi v(x) + ((1-x)/2)\frac{d\pi v}{dx}$. Equipped with this inequality, we now make use of our assumption of convexity on the Riemann invariants and we write for the real number x under consideration

$$\Phi_i(M_1(v)) - \Phi_i(\pi v) - < \nabla_u \Phi_i(\pi v)), (M_1(v) - \pi v) > \geq 0.$$

In view of (3.19), we therefore have

$$\Phi_i(\pi v) + \frac{1}{2}\frac{d\Phi_i(\pi v)}{dx}(1-x) \leq \Phi_i(M_1(v)),$$

and since, by assumption $\frac{d\Phi_i(\pi v)}{dx} > 0$, we deduce

$$\Phi_i(\pi v) \leq \Phi_i(M_1(v)) \leq \frac{1}{1-x}\int_x^1 \Phi_i(v(y))\,dy, \tag{3.20}$$

that is to say that $\pi v(x)$ belongs to the domain built on v.

Consider now the case that, for the given x under consideration, we have $\frac{d\Phi_i(\pi v(x))}{dx} < 0$. Then it is sufficient to consider $M_0(v)(x)$ in place of $M_1(v)(x)$. The proof follows the same steps: starting from (3.18a)

$$(x-1) < \nabla_u \Phi_i(\pi v), \frac{M_1(v)(x) - M_1(\pi v)(x)}{x-1} > \geq 0,$$

we conclude that

$$\Phi_i(\pi v) - \frac{1}{2}\frac{d\Phi_i(\pi v)}{dx}x \leq \Phi_i(M_0(v)) \leq \frac{1}{x}\int_0^x \Phi_i(v(y))\,dy. \tag{3.21}$$

At last, the case when $\frac{d\Phi_i(\pi v(x))}{dx}$ is zero on a nonempty set is obvious. We therefore have shown that an entropy diminishing projection preserves the invariant regions under the assumption (3.10). On the other hand, the exact solution operator preserves the invariant domains, and, as a consequence of (3.20)-(3.21), so does the scheme.

Turning now to the assumption (3.11), we recall that the entropy diminishing Godunov-type method satisfies a complete set of discrete entropy inequalities (see (2.25)). Uniform bound on the amplitude of the approximate solutions are deduced from entropy inequalities by suitable choices of entropy functions. The inequality (2.25) holds in particular for the entropies that are put forward in (3.11) and as a consequence we have at a given time $t = t_n$ and after summation over the space variable of (2.25),

$$\int_{\boldsymbol{R}} U_{\star i}\big((\pi v)(x,t_n)\big)\,dx \leq \int_{\boldsymbol{R}} U_{\star i}\big(u_h(x,0+)\big)\,dx \leq \int_{\boldsymbol{R}} U_{\star i}\big(u_0(x)\big)\,dx = 0.$$

Since, by assumption, $U_{\star i} \geq 0$, this completes the proof of Theorem 3.3. $\square$

Remark 3.5. The set $\mathcal{E}$ can be enlarged to include mathematical entropies for the *entropic* Euler system of gas dynamics, as it is natural from the genreal theory in [2]. The inequality (2.16) then produces a set of inequalities, which is distinct from (2.17).

3.3 Lagrangian Gas Dynamics.

We consider here the full system of gas dynamics in Lagrangian coordinates. Since $N \geq 3$, we do not have a large class of mathematical entropies. As a consequence, the equivalence between (2.16) and (2.17) is lost.

Consider the following system

$$\begin{aligned} &\partial_t \tau - \partial_x u = 0, \\ &\partial_t u + \partial_x p = 0, \\ &\partial_t e + \partial_x (pu) = 0, \end{aligned} \tag{3.22}$$

where $\tau > 0$, u, $e > 0$ denote the specific volume, velocity and internal energy per unit mass, respectively. The pressure $p > 0$ is given by

$$p = \frac{\gamma - 1}{\tau}\big(e - \frac{u^2}{2}\big), \qquad \gamma > 1.$$

This syetm has the form (1.1): it admits three real and distinct eigenvalues, i.e. $-g$, 0, and g, where $g = \sqrt{\frac{\gamma p}{\tau}}$ is the Lagrangian sound speed. The notation (2.13), (2.19) will be of use here. The following statement concerns first order projections $\pi v = \big(\pi\tau, \pi u, \pi e\big)$ of an arbitrary function $v = \big(\tau, u, e\big) : [0,1] \to \boldsymbol{R}_+ \times \boldsymbol{R} \times \boldsymbol{R}_+$.

Theorem 3.4. *Consider the gas dynamics system* (3.22) *and anarbitrary function* v. *The condition* (2.16) *is equivalent to the following* two *inequalities:*

$$\begin{aligned} &\frac{\pi p}{(\gamma-1)}\frac{d\pi\tau}{dx}\Big(D(\tau) - \frac{d\pi\tau}{dx}\Big) + \Big(\frac{d\pi e}{dx} - \pi u \frac{d\pi u}{dx}\Big)\Big(D(e) - \frac{d\pi e}{dx}\Big) \\ &\quad + \Big(\pi u\big(\pi u \frac{d\pi u}{dx} - \frac{d\pi e}{dx}\big) + \frac{(\pi\tau)^2 \pi p}{(\gamma-1)}\frac{d\pi u}{dx}\Big)\Big(D(u) - \frac{d\pi u}{dx}\Big) \geq 0, \end{aligned} \tag{3.23a}$$

$$\Big(\pi p \frac{d\pi\tau}{dx} - \pi u \frac{d\pi u}{dx} + \frac{d\pi e}{dx}\Big)\bigg(\pi p\big(D(\tau) - \frac{d\pi\tau}{dx}\big) - \pi u\big(D(u) - \frac{d\pi u}{dx}\big) + \big(D(e) - \frac{d\pi e}{dx}\big)\bigg) \geq 0. \tag{3.23b}$$

The condition (2.17) *is equivalent to the following* three *inequalities:*

$$\begin{aligned} &\bigg(\frac{\pi p}{(\gamma-1)}\frac{d\pi\tau}{dx} + \big(\pi u \pm \frac{\pi g}{(\gamma-1)}\big)\frac{d\pi u}{dx} - \frac{d\pi e}{dx}\bigg) \\ &\bigg(\frac{\pi p}{(\gamma-1)}\big(D(\tau) - \frac{d\pi\tau}{dx}\big) + \big(\pi u \pm \frac{\pi g}{(\gamma-1)}\big)\big(D(u) - \frac{d\pi u}{dx}\big) - \big(D(e) - \frac{d\pi e}{dx}\big)\bigg) \geq 0, \end{aligned} \tag{3.24a,b}$$

$$\Big(\pi p \frac{d\pi\tau}{dx} - \pi u \frac{d\pi u}{dx} + \frac{d\pi e}{dx}\Big)\bigg(\pi p\big(D(\tau) - \frac{d\pi\tau}{dx}\big) - \pi u\big(D(u) - \frac{d\pi u}{dx}\big) + \big(D(e) - \frac{d\pi e}{dx}\big)\bigg) \geq 0. \tag{3.24c}$$

In particular, the conditions (2.16) *and* (2.17) *are* not *equivalent.*

The result in Theorem 3.4 is a closely related to the lack of mathematical entropies. C.f. also Bonnetier-LeFloch [2].

Proof of Theorem 3.4. It is well known that all the entropies $U(\tau,u,e)$ of (3.22) are of the following form

$$U(\tau,u,e) = -G(S), \quad S = log\big((\gamma-1)\tau^{\gamma-1}(e-\frac{u^2}{2})\big),$$

where $G : \boldsymbol{R} \to \boldsymbol{R}$ is an arbitrary smooth function. Moreover, straigthforward calculations show that the Hessian matrix of $U(\tau,u,e)$ can be expressed as

$$\nabla^2 U(\tau,u,e) = \frac{\gamma^2}{a^4\tau^4} G'(S) \Big\{ (\gamma-1)R\ \nabla^2 U_\circ(\tau,u,e) + (1-\gamma R)\ \nabla^2 U^\circ(\tau,u,e) \Big\},$$

where $R = \frac{G''(S)}{G'(S)}$, $a^2 = \frac{\gamma p}{(\gamma-1)\tau}$, and and

$$\nabla^2 U_\circ(\tau,u,e) = \begin{pmatrix} \frac{a^4\tau^2}{\gamma^2} & \frac{a^2\tau u}{\gamma} & -\frac{a^2\tau}{\gamma} \\ \frac{a^2\tau u}{\gamma} & \frac{(a^2\tau^2-u^2+\gamma u^2)}{\gamma-1} & -u \\ -\frac{a^2\tau}{\gamma} & -u & 1 \end{pmatrix},$$

and

$$\nabla^2 U^\circ(\tau,u,e) = \begin{pmatrix} \frac{(\gamma-1)}{\gamma^2} a^4\tau^2 & 0 & 0 \\ 0 & \frac{a^2\tau^2}{\gamma} + u^2 & -u \\ 0 & -u & 1 \end{pmatrix}.$$

It can be checked that $U(\tau,u,e)$ is a convex function provided that

$$G'(S) \geq 0 \quad \text{and} \quad R \leq \frac{1}{\gamma}.$$

Choosing $R = 1/\gamma$ in the inequality (2.16) yields (3.23a). Choosing now R to be negative, we divide (2.16) by $-R$ and let R go to $-\infty$, which leads to the inequality (3.23b). Since the Hessian matrices of the entropies linearly depend on R, we conclude that the inequalities (3.23) are not only necessary but also sufficient to ensure (2.16) for all entropies.

On the other hand, in view of the expressions of the left eigenvectors to (3.22):

$$l_1 = \frac{1}{2a^2\tau} \begin{pmatrix} \frac{p}{(\gamma-1)} \\ \frac{g}{(\gamma-1)} + u \\ -1 \end{pmatrix},$$

$$l_2 = \frac{1}{a^2\tau} \begin{pmatrix} \frac{\gamma-1}{\gamma} a^2\tau \\ -u \\ 1 \end{pmatrix},$$

$$l_3 = \frac{1}{2a^2\tau} \begin{pmatrix} \frac{p}{(\gamma-1)} \\ -\frac{c}{(\gamma-1)} + u \\ -1 \end{pmatrix},$$

the inequalities (2.17) is seen to imply (3.24), which completes the proof of Theorem 3.4. □

4. Numerical Results.

This section presents numerical results obtained with the scheme introduced in this paper. For definitness we deal with the Lagrangian gas dynamic system (3.22). To begin with, we complete the description of the scheme with some formulas that are necessary for its implementation. Since no closed formula is available for the Generalized Riemann problem for (3.22), we use the approximation due to BenArtzi-Falcovitz [1]. We follow the general theory of approximation by LeFloch-Raviart [28]. We use the approach and notations in Bourgeade-LeFloch-Raviart [4].

We recall that (3.22) is a strictly hyperbolic system, which admits three distinct eigenvalues:

$$\lambda_1(U) = -g, \quad \lambda_2(U) = 0, \quad \lambda_3(U) = g.$$

Let us set $q = \gamma p \tau^{\gamma-1}$. We introduce the new variable $V = (\tau, u, S)$ where $S = p\tau^\gamma$. The components $\big(\alpha_i\big)_{1\le i\le 3}$ of V on the basis of right eigenvectors $\big\{r_i(V^0)\big\}_{1\le i\le 3}$ (V^0 is an arbitrary vector) are

$$\alpha_1 = \frac{1}{2}\big(\frac{u}{g^0} + \tau - \frac{S}{q^0}\big), \quad \alpha_2 = \frac{S}{q^0}, \quad \alpha_3 = \frac{1}{2}\big(-\frac{u}{g^0} + \tau - \frac{S}{q^0}\big) \tag{4.1}$$

or, equivalently,

$$\tau = \alpha_1 + \alpha_2 + \alpha_3, \quad u = g^0(\alpha_1 - \alpha_3), \quad S = q^0\alpha_2. \tag{4.2}$$

Setting $\xi = x/t$, we compute the first order approximation of the G.R.P. solution V:

$$V(x,t) = V^0\big(\xi\big) + t\,V^1\big(\xi\big) + ... \tag{4.3}$$

$$V^1(\xi) = \sum_{1\le j\le 3} \alpha_j(\xi)\, r_j(V^0(\xi)), \quad \xi \in R. \tag{4.4}$$

We denote by $\tau^0,\ u^0,\ p^0,\ g^0, ...$ (respectively $\tau^1,\ u^1,\ p^1,\ g^1, ...$) the quantities associated with the state V^0 (resp. V^1). In (4.3), $V^0\big(\xi\big)$ denotes the entropy weak solution of the classical Riemann problem for (3.22) with the initial data $\big(\tau_L(0-), u_L(0-), e_L(0-)\big)$ if $x < 0$ and $\big(\tau_R(0+), u_R(0+), e_R(0+)\big)$ if $x > 0$. It is a standard matter to determine the solution V^0: it is composed of four constant states separated by three elementary waves. The $1-$ and $3-$ waves are rarefaction waves or shock waves wheras the $2-$ wave is a contact discontinuity with zero speed. The index $*$ will be used to denote states nearby the contact discontinuity.

We only need list the formulas that are required to compute the function $V^1\big(\xi\big)$ in (4.4).

We first consider the intervals of smoothness of $V^0\big(\xi\big)$, i.e. each interval $(\bar\sigma_i^0, \underline{\sigma}_{i+1}^0)$ for $i = 0,1,2,3$ whoe endpoints are the wave speeds in the classical Riemann solution. Therein the function $V^0\big(\xi\big)$ is a constant and there exists three real constants $a_{i,j}$, $j = 1,2,3$, such that

$$\alpha_1(\xi) = (\xi + g_i^0)\, a_{i,1}, \quad \alpha_2(\xi) = \xi\, a_{i,2}, \quad \alpha_3(\xi) = (\xi - g_i^0)\, a_{i,3}. \tag{4.5}$$

On the other hand, consider the case that V^0 contains an i-rarefaction wave ($i = 1$ or $i = 3$). We have if $i = 1$ (for all $\xi \in (-g_L^0, -g_1^0)$):

$$\alpha_1(\xi) = \frac{1}{2(\gamma+1)}\alpha_2(\xi) - \frac{1}{2}\alpha_3(\xi), \tag{4.6a}$$

$$\alpha_2(\xi) = -\big(\frac{\xi}{-g_L^0}\big)^{(\gamma-1)/(\gamma+1)} \cdot g_L^0 \cdot a_{0,2}, \tag{4.6b}$$

$$\alpha_3(\xi) = \frac{g_L^0}{3\gamma-1}\big(\frac{-g_L^0}{\xi}\big)^{1/2}\Big\{-1 + \big(\frac{\xi}{-g_L^0}\big)^{\frac{3\gamma-1}{2(\gamma+1)}}\Big\}a_{0,2} - 2g_L^0\big(\frac{-g_L^0}{\xi}\big)^{1/2}a_{0,3}, \tag{4.6c}$$

and if $i=3$ (for all $\xi \in (g_2^0, g_R^0)$):

$$\alpha_1(\xi) = \frac{g_R^0}{3\gamma-1}\big(\frac{g_R^0}{\xi}\big)^{1/2}\Big\{-\big(\frac{\xi}{-g_R^0}+1\big)^{\frac{3\gamma-1}{2(\gamma+1)}}+1\Big\}a_{3,2} + 2g_R^0\big(\frac{g_R^0}{\xi}\big)^{1/2}a_{3,1}, \tag{4.7a}$$

$$\alpha_2(\xi) = -\big(\frac{\xi}{g_R^0}\big)^{(\gamma-1)/(\gamma+1)}\cdot g_R^0\cdot a_{3,2}, \tag{4.7b}$$

$$\alpha_3(\xi) = -\frac{1}{2}\alpha_1(\xi) + \frac{1}{2(\gamma+1)}\alpha_2(\xi). \tag{4.7c}$$

The constants $a_{0,j}$ and $a_{3,j}$ for $j=1,2,3$ are the same real numbers as the ones in the relations (4.5). We now turn to determining all of these constants. First of all, for $1\le j\le 3$, the constants are known in the two "extreme" regions:

$$a_{0,j} = a_{L,j}, \quad a_{3,j} = a_{R,j}. \tag{4.8}$$

The other constants $a_{i,j}$ for $1\le i\le 2$ and $1\le j\le 3$ are expressed in terms of the pressure $p_\star^1 = p^1(0)$ and the velocity $u_\star^1 = u^1(0)$ at the contact discontinuity. Furthermore these two unknowns are the solution of a 2×2 linear algebraic system that is to be defined later in this section. Assume first that $u_\star^1$ and $p_\star^1$ have been determined, then for $i=1,2$ we have

$$a_{i,1} = \frac{1}{2(g_i^0)^2}\Big\{u_\star^1 - \frac{p_\star^1}{g_i^0}\Big\}, \tag{4.9}$$

$$a_{i,3} = \frac{1}{2(g_i^0)^2}\Big\{u_\star^1 + \frac{p_\star^1}{g_i^0}\Big\}. \tag{4.10}$$

If V^0 contains a $k-$rarefaction wave, we have

$$a_{1,2} = \frac{q_L^0}{q_1^0}a_{L,2}, \quad \text{if } k=1 \tag{4.11}$$

$$a_{2,2} = \frac{q_R^0}{q_2^0}a_{R,2}, \quad \text{if } k=3. \tag{4.12}$$

If V^0 contains a $k-$shock wave, then

$$\begin{aligned} a_{1,2} = -\frac{u_\star^1}{(g_1^0)^2} + \frac{p_\star^1}{\sigma_1^0(g_1^0)^2} + \big(\frac{\partial G}{\partial p}\big)_1^0\Big\{-u_\star^1 + \frac{p_\star^1}{\sigma_1^0}\Big\} \\ + \big(\frac{\partial G}{\partial \tau_a}\big)_L^0\Big\{\tau_L^1 + \frac{u_L^1}{\sigma_1^0}\Big\} + \big(\frac{\partial G}{\partial p_a}\big)_L^0\Big\{p_L^1 - \frac{(g_L^0)^2}{\sigma_1^0}u_L^1\Big\}, \quad \text{if } k=1, \end{aligned} \tag{4.13}$$

and

$$\begin{aligned} a_{2,2} = -\frac{u_\star^1}{(g_2^0)^2} + \frac{p_\star^1}{\sigma_3^0(g_2^0)^2} + \big(\frac{\partial G}{\partial p}\big)_2^0\Big\{-u_\star^1 + \frac{p_\star^1}{\sigma_3^0}\Big\} \\ + \big(\frac{\partial G}{\partial \tau_a}\big)_R^0\Big\{\tau_R^1 + \frac{u_R^1}{\sigma_3^0}\Big\} + \big(\frac{\partial G}{\partial p_a}\big)_R^0\Big\{p_R^1 - \frac{(g_R^0)^2}{\sigma_3^0}u_R^1\Big\}, \quad \text{if } k=3. \end{aligned} \tag{4.14}$$

In (4.13) and (4.14), we have used

$$\frac{\partial G}{\partial \tau_a} = \frac{\tau}{\tau_a}, \quad \frac{\partial G}{\partial p_a} = \frac{\tau_a - \mu^2 \tau}{p + \mu^2 p_a}, \quad \frac{\partial G}{\partial p} = \frac{\mu^2 \tau_a - \tau}{p + \mu^2 p_a}, \tag{4.15}$$

and the notation

$$\begin{aligned} p &= p(\sigma_1^0 + 0), \quad p_a = p(\sigma_1^0 - 0), \cdots \quad \text{for a } 1-\text{wave}, \\ p &= p(\sigma_3^0 - 0), \quad p_a = p(\sigma_3^0 + 0), \cdots \quad \text{for a } 3-\text{wave}. \end{aligned}$$

We now complete the computation of $V^1(\xi)$ with the derivation of the 2×2 algebraic system in the two unknowns $p_\star^1$ and $u_\star^1$. If V^0 contains a 1−rarefaction wave, we have

$$u_\star^1 + \frac{1}{g_1^0} p_\star^1 = C_L \tag{4.16a}$$

with

$$C_L = \sqrt{g_1^0 g_L^0} \Big\{ - u_L^1 + \frac{1}{g_L^0} p_L^1 + \big(\frac{p_L^1}{g_L^0} + g_L^0 \tau_L^1 \big) \cdot \frac{1}{3\gamma - 1} \big[\big(\frac{g_1^0}{g_L^0} \big)^{(3\gamma-1)/2(\gamma+1)} - 1 \big] \Big\}. \tag{4.16b}$$

If V^0 contains a 3−rarefaction wave, then

$$u_\star^1 - \frac{1}{g_2^0} p_\star^1 = C_R \tag{4.17a}$$

with

$$C_R = \sqrt{g_2^0 g_R^0} \Big\{ - u_R^1 - \frac{1}{g_R^0} p_R^1 + \big(\frac{p_R^1}{g_R^0} + g_R^0 \tau_R^1 \big) \cdot \frac{1}{3\gamma - 1} \big[\big(\frac{g_2^0}{g_R^0} \big)^{(3\gamma-1)/2(\gamma+1)} - 1 \big] \Big\}. \tag{4.17b}$$

If V^0 contains a 1−shock wave, we have

$$\Big\{ 1 + \sigma_1^0 \big(\frac{\partial}{\partial p} \Phi_L \big)_1^0 \Big\} u_\star^1 + \Big\{ \frac{-\sigma_1^0}{(g_1^0)^2} - \big(\frac{\partial}{\partial p} \Phi_L \big)_1^0 \Big\} p_\star^1 = K_L \tag{4.18a}$$

with

$$K_L = \big(\frac{\partial}{\partial \tau_a} \Phi_L \big)_1^0 \cdot (\tau_L^1 \sigma_1^0 + u_L^1) + \big(\frac{\partial}{\partial p_a} \Phi_L \big)_1^0 \cdot (p_L^1 \sigma_1^0 - (g_L^0)^2 u_L^1) + u_L^1 \sigma_1^0 - p_L^1. \tag{4.18b}$$

If V^0 contains a 3−shock wave, then

$$\Big\{ 1 - \sigma_3^0 \big(\frac{\partial}{\partial p} \Phi_R \big)_2^0 \Big\} u_\star^1 + \Big\{ \frac{-\sigma_3^0}{(g_2^0)^2} + \big(\frac{\partial}{\partial p} \Phi_R \big)_2^0 \Big\} p_\star^1 = K_R \tag{4.19a}$$

with

$$K_R = -\big(\frac{\partial}{\partial \tau_a} \Phi_R \big)_2^0 \cdot (\tau_R^1 \sigma_3^0 + u_R^1) - \big(\frac{\partial}{\partial p_a} \Phi_R \big)_2^0 \cdot (p_R^1 \sigma_3^0 - (g_R^0)^2 u_R^1) + u_R^1 \sigma_3^0 - p_R^1. \tag{4.19b}$$

In (4.19), we have used the notation

$$\Phi(\tau_a, p_a, p) = (p - p_a)\sqrt{\frac{(1-\mu^2)\tau_a}{p+\mu^2 p_a}} \equiv \frac{(p-p_a)}{W_a(p)}, \quad \mu^2 = \frac{\gamma - 1}{\gamma + 1}. \tag{4.20}$$

Moreover, for a 1−shock wave, we have set

$$\begin{aligned}
\big(\frac{\partial}{\partial p}\Phi_L\big)_1^0 &= \frac{1}{W_L^0(p_\star^0)} - \frac{1}{2}\frac{u_\star^0 - u_L^0}{p_\star^0 + \mu^2 p_L^0},\\
\big(\frac{\partial}{\partial \tau_a}\Phi_L\big)_1^0 &= \frac{1}{2}\frac{u_\star^0 - u_L^0}{\tau_L^0},\\
\big(\frac{\partial}{\partial p_a}\Phi_L\big)_1^0 &= -\frac{1}{W_L^0(p_\star^0)} - \frac{\mu^2}{2}\frac{u_\star^0 - u_L^0}{p_L^0 + \mu^2 p_\star^0},
\end{aligned} \tag{4.21}$$

and, for a 3−shock wave,

$$\begin{aligned}
\big(\frac{\partial}{\partial p}\Phi_R\big)_2^0 &= \frac{1}{W_R^0(p_\star^0)} + \frac{1}{2}\frac{u_\star^0 - u_R^0}{p_\star^0 + \mu^2 p_R^0},\\
\big(\frac{\partial}{\partial \tau_a}\Phi_R\big)_2^0 &= -\frac{1}{2}\frac{u_\star^0 - u_R^0}{\tau_R^0},\\
\big(\frac{\partial}{\partial p_a}\Phi_R\big)_2^0 &= -\frac{1}{W_R^0(p_\star^0)} + \frac{\mu^2}{2}\frac{u_\star^0 - u_R^0}{p_R^0 + \mu^2 p_\star^0}.
\end{aligned} \tag{4.22}$$

We emphasize that the linear system defined by (4.16)–(4.22) is always numerically solvable. Finally, the values at the contact discontinuities of the specific volume relative to the function $V^1(\xi)$ are given by

$$\tau^1(0-) = \frac{u_\star^1}{g_1^0}, \quad \tau^1(0+) = \frac{u_\star^1}{g_2^0}. \tag{4.23}$$

The first term of the asymptotic expansion by LeFloch-Raviart, (4.3), provides a first order approximation for the G.R.P. It is used instead of the exact G.R.P. solution in the first step of the Godunov-type scheme in Section 2. The projection step is based on the second order entropy diminishing projection given in Theorem 2.7. In the present numerical experiments however, we only implement a simplified version of the inequalities (2.17) obtained by replacing $r_i(\pi v)$ by $r_i(\bar{v})$. Under this approximation, we are lead to considering one (decoupled) inequalities for each characteristic variable. The explicit computation then follows the formula given in Theorem 3.1 for the linear systems. The above numerical method has been implemented to solve one dimensional shock-tube problems.

Our figures also picture the solutions given by the classical Van Leer method: Specifically, we used the same first step as above, but the second step is based on the standard MUSCL recontruction in characteristic variables; the classical minmod function is used in the limitation procedure.

The numerical results computed by both schemes are displayed in Figures 4.1–4.3. Figure 4.1 corresponds to a Riemann solution made of two rarefaction waves, Figures 4.2 to a solution made of two shocks, and Figure 4.3 to a solution made of a shock and a rarefaction wave.

The exact solution is displayed as well. The scheme based on the new projectionmethod provides quite sharp discrete shock profiles that are quite comparable to the one provided by the Van Leer MUSCL scheme. The shock capturing capability of both schemes are quite similar. On the other hand, the new scheme happens to yield a somewhat better description of the rarefaction wave, especially regarding the treatment of the tails. The numerical results establish the practical interest of the concept of entropy diminishing method. For further analysis of this notion, C.f. [15].

References.

1. Ben-Artzi M. and Falcovitz J.C., A second-order Godunov-type scheme for compressible gaz dynamics, J. Comp. Phys. 55 (1984), 1–32.
2. Bonnetier E. and LeFloch P.G., Gas dynamics system, in preparation.
3. Bouchut F., Bourdarias C., and Perthame B., An example of MUSCL method satisfying all the entropy inequalities, C.R. Acad. Sc. Paris, Série I, 317 (1993), 619–624.
4. Bourgeade A., LeFloch P.G., and Raviart P.-A., An asymptotic expansion for the solution of the generalized Riemann problem. Part II: application to the gas dynamics equations, Ann. Inst. H. Poincaré, Nonlinear Analysis 6 (1989), 437–480.
5. Brenier Y. and Osher S., The one-sided Lipschitz condition for convex scalar conservation laws, SIAM J. Numer. Anal. 25 (1988), 8–23.
6. Chen G.-Q., The compensated compactness method and the system of isentropic gas dynamics, Preprint, Mathematical Sciences Research Institute, Berkeley (1991).
7. Chen G.-Q. and LeFloch P.G., Entropy flux-splittings for hyperbolic conservation laws. Part I: general framework, to appear in Comm. Pures Appl. Math. (1995).
8. Chen G.-Q. and Liu J.-G., Convergence of second-order schemes for isentropic gas dynamics, Math. of Comp. 61 (1993), 607–627.
9. Cockburn B., Coquel F. and LeFloch P.G., Error estimates for finite volume methods for multidimensional conservation laws, Math. of Comp. 63 (1994), 77–103.
10. Cockburn B., Coquel F. and LeFloch P.G., Convergence of finite volume methods for multidimensional conservation laws, SIAM J. Numer. Anal. (1995).
11. Colella P. and Woodward P. R., The piecewise parabolic method (PPM) for gas-dynamical simulations, J. Comp. Phys. 54 (1984), 174–201.
12. Coquel F. and LeFloch P.G., Convergence of finite difference schemes for conservation laws in several space variables: a general theory, SIAM J. Numer. Anal. 30 (1993), 675–700.
13. Coquel F. and LeFloch P.G., Convergence of finite difference schemes for scalar conservation laws in several space dimensions: the corrected antidiffusive flux approach, Math. of Comp. 57 (1991), 169–210.
14. Coquel F. and LeFloch P.G., A second order entropy satisfying scheme for systems of conservation laws, Note C.R. Acad. Sc., Paris, Série I, (1995).
15. Coquel F. and LeFloch P.G., in preparation (1995).
16. DiPerna R.J., Convergence of the viscosity method for isentropic gas dynamics, Comm. Math. Phys. 91 (1983), 1–30.
17. DiPerna R.J., Convergence of approximate solutions to conservation laws, Arch. Rat. Mech. Anal. 82 (1983), 27–70.
18. Glimm J., Solutions in the large for nonlinear hyperbolic systems of equations, Comm. Pure Appl. Math. 18 (1965), 697–715.
19. Goodman J. and LeVeque R., A geometric approach to high resolution TVD schemes, SIAM J. Numer. Anal. 25 (1988), 268–284.
20. Harabetian E., The generalized Riemann problem with analytic data, Trans. Amer. Math. Soc. (1988).
21. Harten A., On a class of high order resolution total-variation stable finite difference schemes, SIAM J. Numer. Anal. 21 (1974), 1–23.
22. Harten A., Engquist B., Osher S.J., and Chakravarthy S., Uniformly high order accurate essentially non-oscillatory schemes, J. Comp. Phys. 71 (1987), 231–303.

23. Johnson C., Szepessy A., and Hansbo P., On the convergence of shock capturing streamline diffusion finite element methods for hyperbolic conservation laws, Math. of comp. 54 (1990), 107–129.
24. Khobalatte B. and Perthame B., Maximum principle on the entropy and second-order kinetic schemes, Math. of Comp. 62 (1994), 119–135.
25. Lax P.D., Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves, Conf. Board Math. Sci. 11, SIAM, Philadelphia (1973).
26. Van Leer B., Towards the ultimate conservative difference schemes, V: A second order sequel to Godunov's method, J. Comp. Phys. 43 (1981), 357–372.
27. LeFloch P.G. and Liu J.-G., Discrete entropy and monotonicity criterion for hyperbolic conservation laws, C.R. Acad. Sc., Paris, 319 (1994), 881–886.
28. LeFloch P.G. and Raviart P.-A., An asymptotic expansion for the solution of the generalized Riemann problem. Part I: general theory, Ann. Inst. H. Poincaré, Nonlin. Anal. 5 (1988), 179–207.
29. Lions P.-L., Perthame B. and Tadmor E., Kinetic formulations for the p-system and Euler system, Comm. Math. Phys. 163 (1994), 415–431.
30. Lions P.-L. and P. Sougadinis, Convergence of MUSCL-type methods for scalar conservation laws, C.R. Acad. Sc. Paris, Série I, 311 (1990), 259–264.
31. Li T.-T. and W.-C. Yu, Boundary Value Problem for Quasilinear Hyperbolic Systems, Duke Univ. Math. Series (1985).
32. Nessyahu H. and E. Tadmor, Non-oscillatory central differencing for hyperbolic conservation laws, J. Comp. Phys. 87 (1990), 408–462.
33. Osher S.J., Convergence of generalized MUSCL schemes, SIAM J. Numer. Anal. 22 (1985), 947–961.
34. Osher S.J. and S. Chakravarthy, High resolution schemes and entropy condition, SIAM J. Numer. Anal. 22 (1985), 947–961.
35. Osher S.J. and E. Tadmor, On the convergence of difference approximations to scalar conservation laws, Math. of Comp. 50 (1988), 19–51.
36. Serre D., Domaines invariants pour les systèmes hyperboliques de lois de conservation, J. Diff. Eq. 69 (1987), 46–62.
37. Smoller J., Shock Waves and Reaction Diffusion Equations, Springer-Verlag, New York (1983).
38. Tadmor E., The numerical viscosity of entropy stable schemes for systems of conservation laws, Math. of Comp. 49 (1987), 91–103.
39. Tadmor E., Semi-discrete approximations to nonlinear systems of conservation laws; consistency and L^∞-stability imply convergence, ICASE Report 88-41 (1988).
40. Vila J.P., High-order schemes and the entropy condition for nonlinear hyperbolic systems of conservation laws, Math. of Comp. 50 (1988), 53–73.
41. Yang H., Nonlinear wave analysis and convergence of MUSCL schemes, IMA Preprint 697, Minneapolis (1991).