COMMUN. MATH. SCI.
Vol. 11, No. 2, pp. 547–571

GLOBALLY HYPERBOLIC REGULARIZATION OF GRAD'S MOMENT SYSTEM IN ONE DIMENSIONAL SPACE*

ZHENNING CAI†, YUWEI FAN‡, AND RUO LI§

Abstract. In this paper, we present a regularization to the 1D Grad's moment system to achieve global hyperbolicity. The regularization is based on the observation that the characteristic polynomial of the Jacobian of the flux in Grad's moment system is independent of the intermediate coefficients in the Hermite expansion. The method does not rely on the form of the collision at all, thus this regularization is applicable to the system without collision terms. Moreover, the proposed approach is proved to be the unique one if only the last moment equation is allowed to be altered to match the condition that the characteristic speeds coincide with the Gauss-Hermite interpolation points. The hyperbolic structure of the regularized system, including the signal speeds, Riemann invariants, and the properties of the characteristic waves including the rarefaction wave, contact discontinuity, and shock are provided in the perfect formations.

Key words. Grad's moment system, regularization, global hyperbolicity, characteristic wave.

AMS subject classifications. 35L60, 35Q20, 76P05.

1. Introduction

Nowadays, the kinetic gas theory is drawing increased attention in the high-tech fields. The kinetic theory is considered a mesoscopic description of fluids, which is based on the classical Boltzmann equation from statistical physics. However, a full accurate mesoscopic model is still too complex for many problems. People have been looking for a median model between the classical macroscopic equations and the Boltzmann equation for a long time. This can be tracked back to the work of Burnett [5]. As is well known, the Burnett equations were proved to be linearly unstable by Bobylev [3]. Another way leading to linearly stable intermediate models is the moment method proposed by Grad [9]. Since this method was discarded by Grad himself, very few works contributed to this area in the last century. However, this field has becoming active in recent years, since people find that some traditional difficulties in the moment equations can be ignored by some regularizations to these models; see e.g. [12, 10, 19, 23].

This paper focuses on a major critique of the moment method — the lack of global hyperbolicity for Grad's moment system. This deficiency directly causes blow-ups when the distribution is far away from the equilibrium state. It has been reported that increasing the number of moments shows no improvements in numerical experiments [8]. Levermore's work [12] gave a theoretical way to approach the general globally hyperbolic moment equations, although it is still far from practical use due to the lack of an analytical form of his model. Later, using the Pearson-Type-IV distribution, Torrilhon [23] also proposed a 13-moment system, which is globally hyperbolic when reduced to the one-dimensional case, but its generalization to large number moment systems seems to be difficult. In this work, we concentrate on the simple 1D case and achieve a globally hyperbolic regularization to Grad's moment system.

*Received: February 2, 2012; accepted (in revised form): July 14, 2012. Communicated by Shi Jin.

†School of Mathematical Sciences, Peking University, Beijing, China (caizn@pku.edu.cn).
‡School of Mathematical Sciences, Peking University, Beijing, China (ywfan@pku.edu.cn).
§CAPT, LMAM & School of Mathematical Sciences, Peking University, Beijing, China (rli@math.pku.edu.cn).

The first essential observation is that the characteristic polynomial of the Jacobian of the flux of a general Grad's moment system has a simple expression, which only depends on the macroscopic velocity, temperature, and two other coefficients in the Hermite expansion of highest orders. This amazing result directly leads to the possibility of a globally hyperbolic regularization. It is found that these two coefficients take the eigenvalues away from the real axis, resulting in the non-hyperbolicity. We discover an elegant modification of the last equation of the moment system to eliminate the terms involving these two terms in the characteristic polynomial and obtain a globally hyperbolic system. This new hyperbolic system has many fascinating properties. All characteristic fields are either genuinely nonlinear or linearly degenerate. The investigation into the three kinds of elementary characteristic waves (rarefaction waves, contact discontinuities, and shock waves) illustrates substantial similarities with Euler equations. The regularization proposed is very different from the classical way, which tries to give a reasonable recovery of the truncated moments, which is justified in the view of characteristic speeds and order of accuracy. The convergence in the number of moments is illustrated through the numerical study of a shock tube problem.

The rest of this paper is arranged as follows: in Section 2, the Boltzmann equation and the moment method are revised. In Section 3, a detailed investigation on the hyperbolicity of the 1D Grad's moment system is carried out. The regularization of the 1D Grad's moment system to achieve global hyperbolicity is derived in Section 4, with detailed discussion on its properties. A short discussion on the moment equations with collision terms is put forward in Section 5. Section 6 is devoted to the numerical study of a shock tube problem. Finally, some concluding remarks are given in Section 7.

2. The moment method in kinetic theory

In the kinetic gas theory, the state of a gas on the microscopic level is described by the velocity distribution function on each spatial point $\boldsymbol{x}\in\Omega\subset\mathbb{R}^D$. For a time-evolving problem, the distribution function can be described as

$$F:\mathbb{R}^+\times\Omega\times\mathbb{R}^D\to\mathbb{R}^+\cup\{0\},\quad (t,\boldsymbol{x},\boldsymbol{\xi})\mapsto F(t,\boldsymbol{x},\boldsymbol{\xi}), \tag{2.1}$$

where t is the time and $\boldsymbol{\xi}$ denotes the velocity of microscopic gas particles. As in [9], we introduce the mass density

$$f(t,\boldsymbol{x},\boldsymbol{\xi})=mF(t,\boldsymbol{x},\boldsymbol{\xi}), \tag{2.2}$$

where m is the mass of the molecule. The physical case is $D=3$, while in this paper we only consider a 1D model problem. Thus, $\boldsymbol{x}$ and $\boldsymbol{\xi}$ will be written in plain font as x and ξ for the rest of the paper.

2.1. The Boltzmann equation and conservation laws. The mass density f satisfies the Boltzmann equation, which reads

$$\frac{\partial f}{\partial t}+\xi\frac{\partial f}{\partial x}=Q(f,f), \tag{2.3}$$

where $Q(f,f)$ is the collision term with a complex expression, which models the interaction between particles. In most of this paper, we only consider the collisionless case, thus $Q(f,f)=0$ will be assumed if not specified. However, the readers may keep in mind that our final aim is to provide an improved description of the Boltzmann equation with collision term using the moment method, and we will return to this topic in Section 5.

The basic variables, including the density, the momentum density, and total energy density, are defined as

$$\begin{aligned}\rho(t,x) &= \int_{\mathbb{R}} f(t,x,\xi)\,\mathrm{d}\xi,\\ \rho(t,x)u(t,x) &= \int_{\mathbb{R}} \xi f(t,x,\xi)\,\mathrm{d}\xi,\\ \frac{1}{2}\rho(t,x)|u(t,x)|^2+\frac{1}{2}\rho(t,x)\theta(t,x) &= \int_{\mathbb{R}} \frac{1}{2}|\xi|^2 f(t,x,\xi)\,\mathrm{d}\xi.\end{aligned} \tag{2.4}$$

Here u is the macroscopic velocity, and θ is the multiplication of gas constant and temperature. Multiplying the Boltzmann equation (2.3) by $(1,\xi,\xi^2/2)^T$, integrating both sides over $\mathbb{R}$ with respect to ξ, and then making some simplifications, we get the non-conservative form of the conservation laws as

$$\frac{\partial \rho}{\partial t}+u\frac{\partial \rho}{\partial x}+\rho\frac{\partial u}{\partial x}=0, \tag{2.5a}$$

$$\rho\frac{\partial u}{\partial t}+\frac{\partial p}{\partial x}+\rho u\frac{\partial u}{\partial x}=0, \tag{2.5b}$$

$$\frac{1}{2}\rho\frac{\partial \theta}{\partial t}+\frac{1}{2}\rho u\frac{\partial \theta}{\partial x}+\frac{\partial q}{\partial x}+p\frac{\partial u}{\partial x}=0, \tag{2.5c}$$

where p is the pressure and q is the heat flux. They are defined as

$$p=\rho\theta, \quad q=\frac{1}{2}\int_{\mathbb{R}}(\xi-u)^3 f\,\mathrm{d}\xi. \tag{2.6}$$

2.2. The moment method. The moment method was raised by Grad in [9], where a thirteen moment system was introduced. However, systems with large moment numbers were not investigated until recently (e.g. [24, 1, 6, 8]). Here we use the notations in [6, 8], and expand $f(t,x,\xi)$ as

$$f(t,x,\xi)=\sum_{k\in\mathbb{N}} f_k(t,x)\mathcal{H}_{\theta(t,x),k}\left(\frac{\xi-u(t,x)}{\sqrt{\theta(t,x)}}\right), \tag{2.7}$$

where

$$\mathcal{H}_{\theta,k}(v)=\frac{1}{\sqrt{2\pi}}\theta^{-\frac{k+1}{2}}\,He_k(v)\exp\left(-\frac{v^2}{2}\right), \tag{2.8}$$

where He_k is the k-th Hermite polynomial, defined by

$$He_k(x)=(-1)^k\exp\left(\frac{x^2}{2}\right)\frac{\mathrm{d}^k}{\mathrm{d}x^k}\exp\left(-\frac{x^2}{2}\right). \tag{2.9}$$

Based on this expansion, some simple properties can be deduced:

$$f_0=\rho, \quad f_1=f_2=0, \quad q=3f_3. \tag{2.10}$$

If we put (2.7) into the Boltzmann equation (2.3), the equation for each coefficient can be deduced as

$$\begin{aligned}\frac{\partial f_k}{\partial t} - f_{k-1}\frac{\theta}{\rho}\frac{\partial \rho}{\partial x} + (k+1) f_k \frac{\partial u}{\partial x} + \left(\frac{1}{2}\theta f_{k-3} + \frac{k-1}{2} f_{k-1}\right)\frac{\partial \theta}{\partial x} \\ -\frac{3}{\rho} f_{k-2}\frac{\partial f_3}{\partial x} + \theta \frac{\partial f_{k-1}}{\partial x} + u\frac{\partial f_k}{\partial x} + (k+1)\frac{\partial f_{k+1}}{\partial x} = 0, \quad \text{for } k \geqslant 3.\end{aligned} \tag{2.11}$$

For details, we refer the reader to [8]. The conservation laws (2.5) together with (2.11) form a moment system with an infinite number of equations. In order to get a closed system with a finite number of equations, one can follow Grad's idea [9] and let $f_{M+1} = 0$ for some $M \geqslant 3$. Thus a closed system with $M+1$ moments is obtained.

3. Hyperbolicity of Grad's moment systems

A 1D quasilinear system

$$\frac{\partial \boldsymbol{q}}{\partial t} + \mathbf{A}(\boldsymbol{q})\frac{\partial \boldsymbol{q}}{\partial x} = 0 \tag{3.1}$$

is called hyperbolic for a particular $\boldsymbol{q}_0$ if the matrix $\mathbf{A}(\boldsymbol{q}_0)$ is diagonalizable with real eigenvalues. For Grad's systems, the hyperbolicity can only be obtained where the distribution function is near Maxwellian [15, 4, 23]. The loss of hyperbolicity makes Grad's systems overdetermined for strongly non-equilibrium gases, and severely restricts the application of moment methods. In this section, we are going to study the 1D model problem and find the way in which high order moments affect the hyperbolicity of the moment system.

Let $\boldsymbol{w}_M = (\rho, u, \theta, f_3, \cdots, f_M)^T \in \mathbb{R}^{M+1}$, $M \in \mathbb{N}$, and $M \geqslant 2$. The Grad's moment system (2.5) and (2.11) with $f_{M+1} = 0$ is then written as

$$\frac{\partial \boldsymbol{w}_M}{\partial t} + \mathbf{A}_M \frac{\partial \boldsymbol{w}_M}{\partial x} = 0, \tag{3.2}$$

where $\mathbf{A}_M$ is a lower Hessenberg matrix defined as

$$\mathbf{A}_M = \begin{pmatrix} u & \rho & 0 & \multicolumn{2}{r}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\; 0} \\ \theta/\rho & u & 1 & 0 & \cdots\cdots\cdots\; 0 \\ 0 & 2\theta & u & 6/\rho & 0 \cdots\cdots\cdots 0 \\ 0 & 4f_3 & \rho\theta/2 & u & 4\;\; 0\; \cdots\cdots\; 0 \\ -\theta f_3/\rho & 5f_4 & 3f_3/2 & \theta & u\;\; 5\;\; 0 \cdots 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\theta f_{M-2}/\rho & M f_{M-1} & \frac{1}{2}[(M-2)f_{M-2} + \theta f_{M-4}] & -3f_{M-3}/\rho & 0 \cdots 0\; \theta\; u\; M \\ -\theta f_{M-1}/\rho & (M+1)f_M & \frac{1}{2}[(M-1)f_{M-1} + \theta f_{M-3}] & -3f_{M-2}/\rho & 0 \cdots\cdots 0\; \theta\;\; u \end{pmatrix}. \tag{3.3}$$

We write the matrix in a simplified formation with a translation and similarity transformation. Let us define

$$\boldsymbol{\Lambda} = \operatorname{diag}\left\{1, \rho\theta^{-1/2}, \frac{1}{2}\rho\theta^{-1}, \theta^{-3/2}, \cdots, \theta^{-M/2}\right\}, \qquad g_j = \frac{f_j}{\rho\theta^{j/2}}, \quad j = 3, \cdots, M. \tag{3.4}$$

Then

$$\mathbf{A}_M = u\mathbf{I} + \sqrt{\theta}\boldsymbol{\Lambda}^{-1}\tilde{\mathbf{A}}_M\boldsymbol{\Lambda}, \tag{3.5}$$

where $\tilde{\mathbf{A}}_M$ is defined as

$$\tilde{\mathbf{A}}_M = \begin{pmatrix} 0 & 1 & 0 & \cdots\cdots\cdots\cdots & & & & 0 \\ 1 & 0 & 2 & 0 & \cdots\cdots & & & 0 \\ 0 & 1 & 0 & 3 & 0 \cdots\cdots & & & 0 \\ 0 & 4g_3 & 1 & 0 & 4 & 0 & \cdots & 0 \\ -g_3 & 5g_4 & 3g_3 & 1 & 0 & 5 & 0 \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -g_{M-2} & Mg_{M-1} & (M-2)g_{M-2}+g_{M-4} & -3g_{M-3} & 0 \cdots 0 & 1 & 0 & M \\ -g_{M-1} & (M+1)g_M & (M-1)g_{M-1}+g_{M-3} & -3g_{M-2} & 0 \cdots 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3.6}$$

Thus, if

$$\tilde{\lambda}_j, \quad j=1,\cdots,M+1$$

are all the eigenvalues of $\tilde{\mathbf{A}}_M$, then

$$u+\tilde{\lambda}_j\sqrt{\theta}, \quad j=1,\cdots,M+1$$

are all the eigenvalues of $\mathbf{A}_M$.

The matrix $\tilde{\mathbf{A}}_M$ can be considered as "simple" in a sense. It contains only dimensionless variables $g_3,\cdots,g_M$ with linear dependence. The diagonal elements of $\tilde{\mathbf{A}}_M$ are all vanished, and the subdiagonal entries are all 1. The superdiagonal elements are equal to their row numbers. Meanwhile, apart from the tridiagonal part, only the first four columns are nonzero. This formation gives us possibility to study its eigenvalues.

We first present the main result of this section in Theorem 3.1. In this paper, $|\cdot|$ is used to denote the determinant of a matrix.

THEOREM 3.1. *The characteristic polynomial of $\tilde{\mathbf{A}}_M$ is*

$$\left|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M\right| = He_{M+1}(\lambda)-\frac{1}{2}(M+1)!\cdot[(\lambda^2-1)g_{M-1}+2\lambda g_M]. \tag{3.7}$$

The result is incredibly simple, and therefore gives us a realistic possibility to make some kind of regularization to gain global hyperbolicity, which will be discussed in the next section. To proof this theorem, we need the following two lemmas.

LEMMA 3.2. *Suppose that a square matrix $\mathbf{A}=(a_{ij})$ depends on N variables $x_1,\cdots,x_N$. Then the partial derivatives of $|\mathbf{A}|$ can be calculated as*

$$\frac{\partial|\mathbf{A}|}{\partial x_k} = \sum_{i,j}(-1)^{i+j}\frac{\partial a_{ij}}{\partial x_k}A^{ij}, \quad k=1,\cdots,N. \tag{3.8}$$

where A^{ij} is the (i,j)-th minor of $\mathbf{A}$, which is defined to be the determinant of the submatrix obtained by removing from $\mathbf{A}$ its i-th row and j-th column.

This is a familiar result in linear algebra, and will not be proved here.

LEMMA 3.3. *Define tridiagonal matrices*

$$\mathbf{D}_j = \begin{pmatrix} \lambda & -(j+1) & 0 & \cdots\cdots & & & 0 \\ -1 & \lambda & -(j+2) & 0 & \cdots & & 0 \\ 0 & -1 & \lambda & -(j+3) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots\cdots & & 0 & -1 & \lambda & -M \\ 0 & \cdots\cdots & & & 0 & -1 & \lambda \end{pmatrix}, \quad 0\leqslant j\leqslant M. \tag{3.9}$$

The following relations for the determinants of $\mathbf{D}_j$ *hold:*

$$|\mathbf{D}_j| = \lambda|\mathbf{D}_{j+1}| - (j+1)|\mathbf{D}_{j+2}|, \quad 0 \leqslant j \leqslant M-2. \tag{3.10}$$

Proof. For $0 \leqslant j \leqslant M-1$, D_j can be written as

$$\mathbf{D}_j = \begin{pmatrix} \lambda & -(j+1)\boldsymbol{e}_1^T \\ -\boldsymbol{e}_1 & \mathbf{D}_{j+1} \end{pmatrix}, \tag{3.11}$$

where $\boldsymbol{e}_1$ is the unit vector $(1,0,\cdots,0)^T$. When $\lambda \neq 0$, since

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \lambda^{-1}\boldsymbol{e}_1 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \lambda & -(j+1)\boldsymbol{e}_1^T \\ -\boldsymbol{e}_1 & \mathbf{D}_{j+1} \end{pmatrix} = \begin{pmatrix} \lambda & -(j+1)\boldsymbol{e}_1^T \\ \mathbf{0} & \mathbf{D}_{j+1} - (j+1)\lambda^{-1}\boldsymbol{e}_1\boldsymbol{e}_1^T \end{pmatrix}, \tag{3.12}$$

the equality

$$|\mathbf{D}_j| = \lambda \left| \mathbf{D}_{j+1} - (j+1)\lambda^{-1}\boldsymbol{e}_1\boldsymbol{e}_1^T \right| \tag{3.13}$$

is obtained by taking determinants on both sides of (3.12). When $0 \leqslant j \leqslant M-2$, we use (3.11) again and get

$$\begin{aligned} |\mathbf{D}_j| &= \lambda \left| \mathbf{D}_{j+1} - (j+1)\lambda^{-1}\boldsymbol{e}_1\boldsymbol{e}_1^T \right| = \lambda \begin{vmatrix} \lambda - (j+1)\lambda^{-1} & -(j+2)\boldsymbol{e}_1^T \\ -\boldsymbol{e}_1 & \mathbf{D}_{j+2} \end{vmatrix} \\ &= \lambda \left(\begin{vmatrix} \lambda & -(j+2)\boldsymbol{e}_1^T \\ -\boldsymbol{e}_1 & \mathbf{D}_{j+2} \end{vmatrix} + \begin{vmatrix} -(j+1)\lambda^{-1} & \mathbf{0} \\ -\boldsymbol{e}_1 & \mathbf{D}_{j+2} \end{vmatrix} \right) = \lambda|\mathbf{D}_{j+1}| - (j+1)|\mathbf{D}_{j+2}|. \end{aligned} \tag{3.14}$$

If $\lambda = 0$, the continuity of $|\mathbf{D}_j|$ with respect to λ gives the same result. ∎

Now we prove Theorem 3.1.

Proof of Theorem 3.1. We start the proof by calculating $\partial|\lambda\mathbf{I} - \tilde{\mathbf{A}}_M|/\partial g_j$ for $3 \leqslant j \leqslant M-3$. From (3.6), one may find that g_j only appears in five entries of the matrix. Their positions are

$$(j+2,1), \quad (j+1,2), \quad (j+2,3), \quad (j+4,3), \quad (j+3,4),$$

which are illustrated in figure 3.1(a). Thus, according to Lemma 3.2, only five terms appear in the right hand side of (3.8). Now we will consider them one by one. Below we denote $\lambda\mathbf{I} - \tilde{\mathbf{A}}_M = (c_{ij})$, and use $C^{i,j}$ to denote the (i,j)-th minor of $\lambda\mathbf{I} - \tilde{\mathbf{A}}_M$.

1. As in figure 3.1(b), $C^{j+2,1}$ is presented as the product of the determinants of two matrices. One is a lower triangular matrix whose diagonal elements are $-1, \cdots, -(j+1)$, and the other is a lower right block of $\lambda\mathbf{I} - \tilde{\mathbf{A}}_M$, which is actually $\mathbf{D}_{j+2}$ defined in (3.9). Therefore, we obtain

$$C^{j+2,1} = (-1)^{j+1}(j+1)! \cdot |\mathbf{D}_{j+2}|. \tag{3.15}$$

Since $c_{j+2,1} = g_j$, one has

$$\begin{aligned} (-1)^{j+2+1} \frac{\partial c_{j+2,1}}{\partial g_j} C^{j+2,1} &= (-1)^{j+1} \cdot 1 \cdot (-1)^{j+1}(j+1)! \cdot |\mathbf{D}_{j+2}| \\ &= (j+1)! \cdot |\mathbf{D}_{j+2}|. \end{aligned} \tag{3.16}$$

2. Figure 3.1(c) shows that $C^{j+1,2}$ is factored into three parts: the first part is λ, the second is a lower triangular matrix with diagonal elements $-2,\cdots,-j$, and the third one is $\mathbf{D}_{j+1}$. Since $c_{j+1,2}=-(j+1)g_j$, we get

$$(-1)^{j+1+2}\frac{\partial c_{j+1,2}}{\partial g_j}C^{j+1,2}=(-1)^{j+1}\cdot[-(j+1)]\cdot(-1)^{j-1}j!\lambda\cdot|\mathbf{D}_{j+1}| \\ =-(j+1)!\cdot\lambda|\mathbf{D}_{j+1}|. \tag{3.17}$$

3. $C^{j+2,3}$ is illustrated in figure 3.1(d), from which one finds that $C^{j+2,3}$ is the product of the determinants of three matrices. The first matrix is a 2×2 upper left block of $\lambda\mathbf{I}-\tilde{\mathbf{A}}_M$, for which we have

$$\begin{vmatrix}\lambda & -1\\ -1 & \lambda\end{vmatrix}=\lambda^2-1. \tag{3.18}$$

And the other two blocks are similar to the last case. Using $c_{j+2,3}=-jg_j$, we have

$$(-1)^{j+2+3}\frac{\partial c_{j+2,3}}{\partial g_j}C^{j+2,3}=(-1)^{j+1}\cdot(-j)\cdot(-1)^{j-1}\frac{1}{2}(j+1)!(\lambda^2-1)\cdot|\mathbf{D}_{j+2}| \\ =-\frac{j}{2}(j+1)!\cdot(\lambda^2-1)|\mathbf{D}_{j+2}|. \tag{3.19}$$

4. The structure of $C^{j+4,3}$ is plotted in figure 3.1(e), which is very similar as $C^{j+2,3}$. Therefore we directly write the result,

$$(-1)^{j+4+3}\frac{\partial c_{j+4,3}}{\partial g_j}C^{j+4,3}=(-1)^{j+1}\cdot(-1)\cdot(-1)^{j+1}\frac{1}{2}(j+3)!(\lambda^2-1)\cdot|\mathbf{D}_{j+4}| \\ =-\frac{1}{2}(j+3)!\cdot(\lambda^2-1)|\mathbf{D}_{j+4}|, \tag{3.20}$$

where we have used $c_{j+4,3}=-(j+2)g_{j+2}-g_j$. Note that we define $|\mathbf{D}_{M+1}|=1$ so that (3.20) is correct for $j=M-3$.

5. Similar to $C^{j+2,3}$ and $C^{j+4,3}$, the minor $C^{j+3,4}$ is also factored into the determinants of three matrices as in figure 3.1(f), while the first matrix is the 3×3 upper left block of $\lambda\mathbf{I}-\tilde{\mathbf{A}}_M$, whose determinant is

$$\begin{vmatrix}\lambda & -1 & 0\\ -1 & \lambda & -2\\ 0 & -1 & \lambda\end{vmatrix}=\lambda^3-3\lambda. \tag{3.21}$$

Thus the last term becomes

$$(-1)^{j+3+4}\frac{\partial c_{j+3,4}}{\partial g_j}C^{j+3,4}=(-1)^{j+1}\cdot 3\cdot(-1)^{j-1}\frac{1}{6}(j+2)!(\lambda^3-3\lambda)\cdot|\mathbf{D}_{j+3}| \\ =\frac{1}{2}(j+2)!\cdot(\lambda^3-3\lambda)|\mathbf{D}_{j+3}|. \tag{3.22}$$

Collecting (3.16), (3.17), (3.19), (3.20), and (3.22), we finally get

$$\frac{\partial|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|}{\partial g_j}=(j+1)!\cdot\Bigg[|\mathbf{D}_{j+2}|-\lambda|\mathbf{D}_{j+1}|-\frac{j}{2}(\lambda^2-1)|\mathbf{D}_{j+2}| \\ -\frac{(j+3)(j+2)}{2}(\lambda^2-1)|\mathbf{D}_{j+4}|+\frac{j+2}{2}(\lambda^3-3\lambda)|\mathbf{D}_{j+3}|\Bigg]. \tag{3.23}$$

This expression will be further simplified using Lemma 3.3. Since (3.11) also holds for $j=M-1$, if we define $|\mathbf{D}_{M+1}|=1$ the following relation is deduced:

$$
\begin{aligned}
&|\mathbf{D}_{j+2}|-\lambda|\mathbf{D}_{j+1}|-\frac{j}{2}(\lambda^2-1)|\mathbf{D}_{j+2}| \\
&=|\mathbf{D}_{j+2}|-\lambda(\lambda|\mathbf{D}_{j+2}|-(j+2)|\mathbf{D}_{j+3}|)-\frac{j}{2}(\lambda^2-1)|\mathbf{D}_{j+2}| \\
&=-\frac{j+2}{2}(\lambda^2-1)|\mathbf{D}_{j+2}|+(j+2)\lambda|\mathbf{D}_{j+3}| \\
&=-\frac{j+2}{2}(\lambda^2-1)[\lambda|\mathbf{D}_{j+3}|-(j+3)|\mathbf{D}_{j+4}|]+(j+2)\lambda|\mathbf{D}_{j+3}| \\
&=-\frac{j+2}{2}(\lambda^3-3\lambda)|\mathbf{D}_{j+3}|+\frac{(j+2)(j+3)}{2}(\lambda^2-1)|\mathbf{D}_{j+4}|.
\end{aligned}
\tag{3.24}
$$

Substituting this equation into (3.23), we conclude

$$
\frac{\partial|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|}{\partial g_j}=0, \quad 3\leqslant j\leqslant M-3. \tag{3.25}
$$

It is clear that $g_3,\cdots,g_{M-3}$ do not appear in the characteristic polynomial of $\tilde{\mathbf{A}}_M$.

For $j=M-2,M-1,M$, the entries containing g_j still appear in the matrix as in figure 3.1(a), while some items are missing due to the cut-off. Therefore, if we define $|\mathbf{D}_j|=0$ for $j>M+1$, then (3.23) still applies for $j=M-2,M-1,M$. Note that this definition leads to

$$
|\mathbf{D}_M|=\lambda|\mathbf{D}_{M+1}|-(M+1)|\mathbf{D}_{M+2}|, \tag{3.26}
$$

and therefore g_{M-2} does not appear in $|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|$ either. Moreover, we have

$$
\begin{aligned}
\frac{\partial|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|}{\partial g_{M-1}}&=M!\cdot\left[|\mathbf{D}_{M+1}|-\lambda|\mathbf{D}_M|-\frac{M-1}{2}(\lambda^2-1)|\mathbf{D}_{j+2}|\right] \\
&=-\frac{(M+1)!}{2}(\lambda^2-1),
\end{aligned}
\tag{3.27}
$$

$$
\frac{\partial|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|}{\partial g_M}=(M+1)!\cdot(-\lambda|\mathbf{D}_{M+1}|)=-(M+1)!\cdot\lambda. \tag{3.28}
$$

Since (3.27) and (3.28) hold for any g_j, $3\leq j\leq M$, we write the characteristic polynomial of $\tilde{\mathbf{A}}_M$ as

$$
|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|=C(\lambda)-\frac{(M+1)!}{2}[(\lambda^2-1)g_{M-1}+2\lambda g_M], \tag{3.29}
$$

where $C(\lambda)$ is a function of λ.

Now it only remains to determine $C(\lambda)$, which is done by assigning $g_3,\cdots,g_M$ to be zero, and then calculating the characteristic polynomial of $\tilde{\mathbf{A}}_M$. In this case, it is easy to find

$$
|\lambda\mathbf{I}-\tilde{\mathbf{A}}_M|=C(\lambda)=|\mathbf{D}_0|, \qquad \text{if } g_3=\cdots=g_M=0. \tag{3.30}
$$

Meanwhile, the following relation between $|\mathbf{D}_j|$ and Hermite polynomials is discovered:

$$
\begin{gathered}
|\mathbf{D}_{M+1}|=He_0(\lambda)=1, \qquad |\mathbf{D}_M|=He_1(\lambda)=\lambda, \\
|\mathbf{D}_j|=\lambda|\mathbf{D}_{j+1}|-(j+1)|\mathbf{D}_{j+2}|, \\
He_j(\lambda)=\lambda He_{j-1}(\lambda)-(j-1)He_{j-2}(\lambda).
\end{gathered}
\tag{3.31}
$$

This reveals that

$$|\mathbf{D}_j| = He_{M+1-j}(\lambda), \quad 0 \leqslant j \leqslant M+1. \tag{3.32}$$

Hence $C(\lambda) = |\mathbf{D}_0| = He_{M+1}(\lambda)$. This completes the proof of Theorem 3.1. ∎

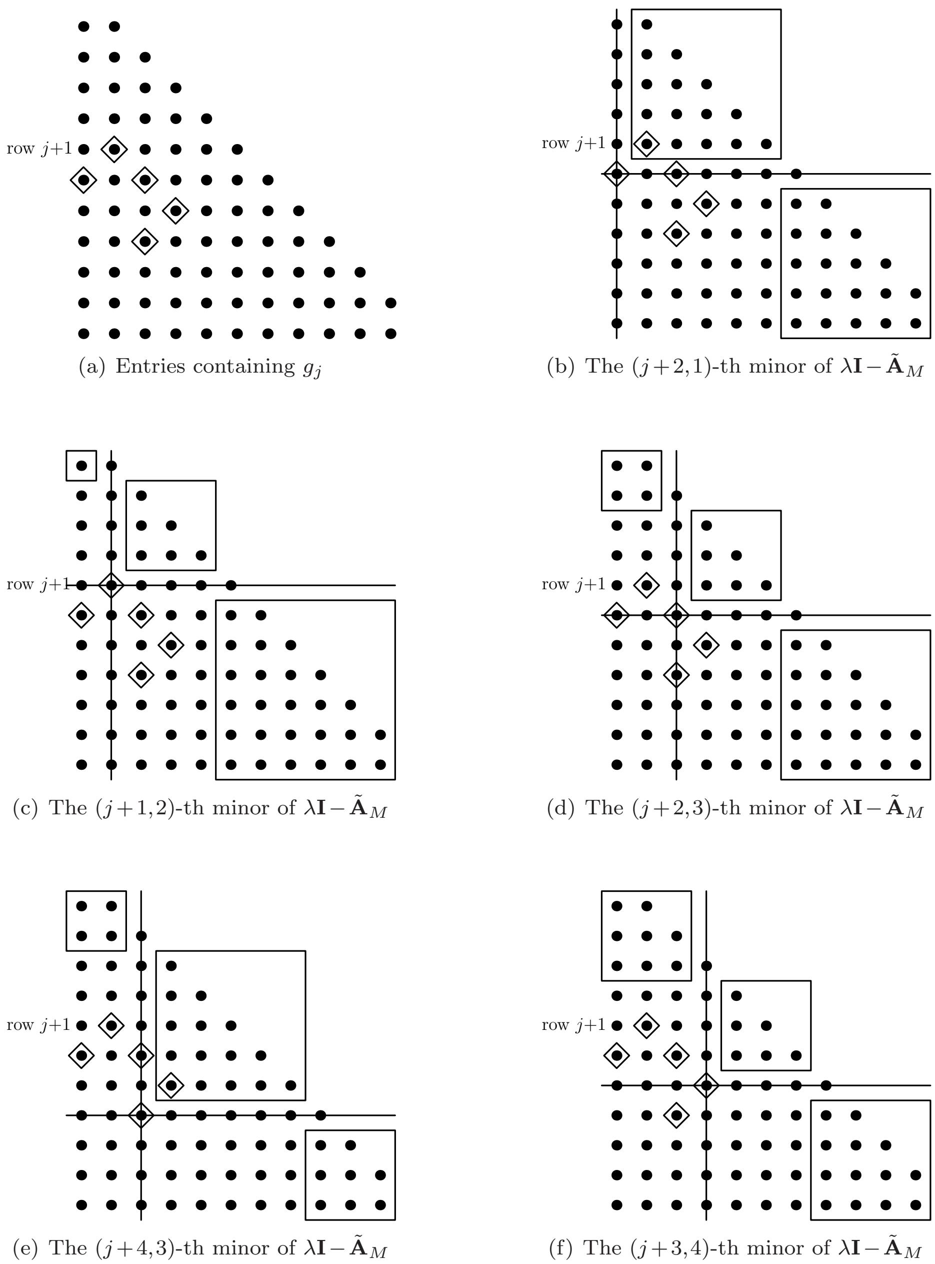
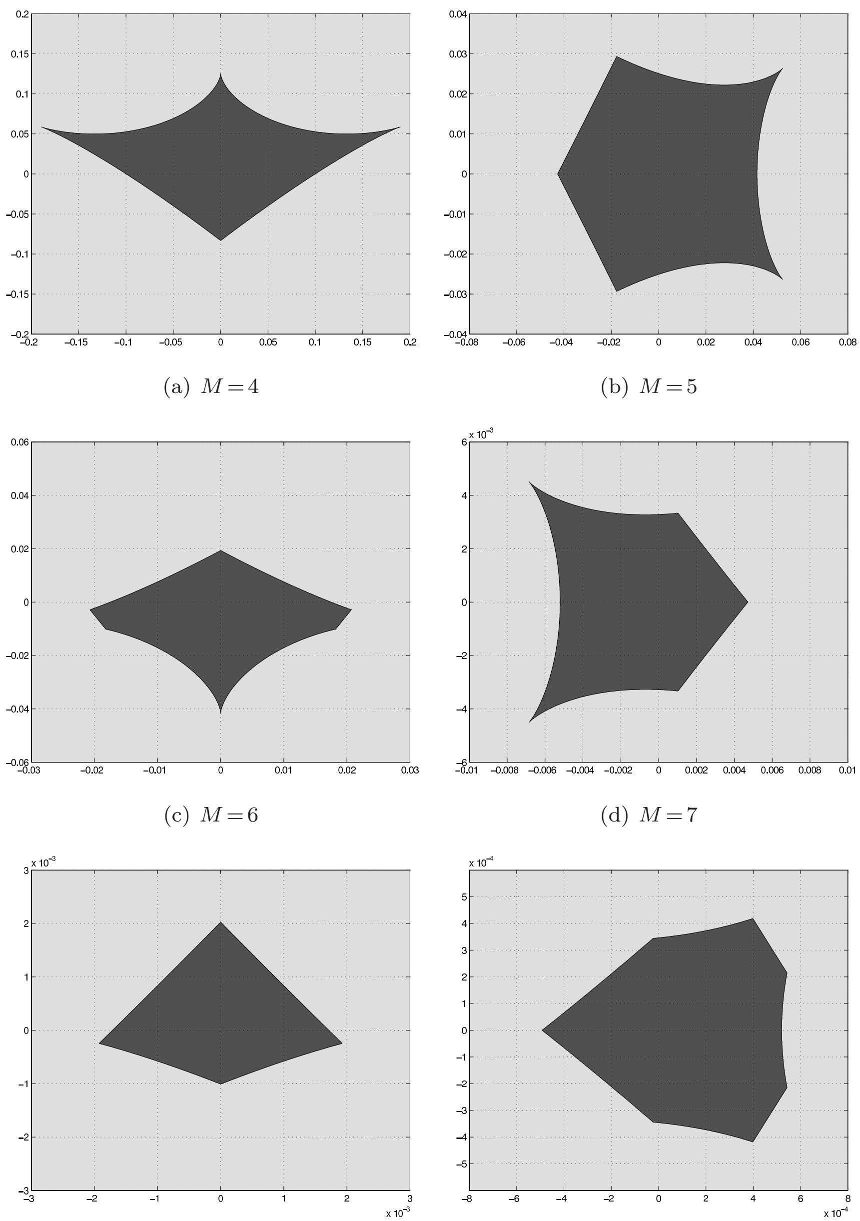


FIG. 3.1. *The Hessenberg matrix $\tilde{A}_M$ and its minors of the elements containing g_j.*

Theorem 3.1 reveals that hyperbolicity can only be obtained in a particular region $(g_{M-1}, g_M) \in \Omega_M$ for the $(M+1)$-moment system. Since the roots of Hermite polynomials are all real, the origin must lie in Ω_M. The region Ω_M for $M=4$ to 9 are plotted in figure 3.2, among which the result for $M=4$ has been obtained in [23], agreeing with ours with proper scaling and translation.



(a) $M=4$ (b) $M=5$

(c) $M=6$ (d) $M=7$

(e) $M=8$ (f) $M=9$

FIG. 3.2. *Hyperbolicity region of Grad's* $(M+1)$*-moment system. The* x*-axis is* g_{M-1} *and the* y*-axis is* g_M.

As a reference, the following corollary gives the characteristic polynomial of the original matrix $\mathbf{A}_M$.

COROLLARY 3.1. *The characteristic polynomial of* $\mathbf{A}_M$ *is*

$$\theta^{\frac{M+1}{2}} He_{M+1}\left(\frac{\lambda-u}{\sqrt{\theta}}\right) - \frac{(M+1)!}{2\rho}\left[\left((\lambda-u)^2-\theta\right)f_{M-1}+2(\lambda-u)f_M\right]. \tag{3.33}$$

Proof. This can be shown by direct calculation:

$$\begin{aligned}
&|\lambda\mathbf{I}-\mathbf{A}_M| = \left|\lambda\mathbf{I}-(u\mathbf{I}+\sqrt{\theta}\mathbf{\Lambda}^{-1}\tilde{\mathbf{A}}_M\mathbf{\Lambda})\right| \\
=&\theta^{\frac{M+1}{2}}\left|\frac{\lambda-u}{\sqrt{\theta}}\mathbf{I}-\mathbf{\Lambda}^{-1}\tilde{\mathbf{A}}_M\mathbf{\Lambda}\right| = \theta^{\frac{M+1}{2}}\left|\frac{\lambda-u}{\sqrt{\theta}}\mathbf{I}-\tilde{\mathbf{A}}_M\right| \\
=&\theta^{\frac{M+1}{2}}\Bigg\{He_{M+1}\left(\frac{\lambda-u}{\sqrt{\theta}}\right) \\
&\qquad -\frac{(M+1)!}{2}\left[\left(\frac{(\lambda-u)^2}{\theta}-1\right)\frac{f_{M-1}}{\rho\theta^{(M-1)/2}}+\frac{2(\lambda-u)}{\sqrt{\theta}}\frac{f_M}{\rho\theta^{M/2}}\right]\Bigg\} \\
=&\theta^{\frac{M+1}{2}} He_{M+1}\left(\frac{\lambda-u}{\sqrt{\theta}}\right) - \frac{(M+1)!}{2\rho}\left[\left((\lambda-u)^2-\theta\right)f_{M-1}+2(\lambda-u)f_M\right].
\end{aligned} \tag{3.34}$$

∎

4. Hyperbolic moment system

The loss of global hyperbolicity of Grad's moment system has long been considered as a failure of moment method. Recently, some encouraging progress has been made in this direction [12, 23]. However, in the case that the number of moments is greater than 10, Levermore's method leads to great difficulties for numerical implementation, since the moments cannot be analytically solved from the Lagrange multipliers[1]; and it has been demonstrated by Junk [11] that the domain of definition for a realizable distribution is not convex. Torrilhon's method mainly focuses on the 13-moment case in one space dimension, which seems not trivial to extend to the general case. To the best of our knowledge, no results for the general moment system have been published.

In this section, we provide the method to regularize the moment system based on the results in Section 3 to achieve global hyberbolicity. We discuss only the 1D case here and the multi-dimensional problems will be reported soon in later papers.

4.1. Construction of the hyperbolic moment system. For an $(M+1)$-moment system containing quantities $\{\rho,u,\theta,f_3,\cdots,f_M\}$, the Grad's moment system gives *accurate* evolution equations for most variables expect for f_M, since f_{M+1} appears in the accurate equation of f_M, and is forced to be zero in Grad's closure. Almost all the regularization methods in references are focused on the reconstruction of f_{M+1}, trying to express f_{M+1} as a function of the $M+1$ known variables in some possible ways such as Chapman-Enskog expansion or realizing a positive distribution [12, 19, 8, 23]. In this paper, we also limit our regularization to the modification of the equation of f_M. However, since f_{M+1} exists in this equation only in the form of its derivative, here we directly substitute $\partial f_{M+1}/\partial x$ with some other expression to gain global hyperbolicity.

Corollary 3.1 shows that the characteristic polynomial of $\mathbf{A}_M$ is independent of f_3, $\cdots$, f_{M-2}, and its dependence of f_{M-1} and f_M can be regarded as the result of

[1]A local system is required to be solved by Newton iteration on each grid for every time step. We refer the readers to [20] for details. There is no report indicating that such a system has a fast solver.

truncation. That is, if a Grad's system with $M+3$ or more variables is considered, then f_{M-1} and f_M do not affect the characteristic polynomial, either. Thus, it is reasonable to modify the matrix $\mathbf{A}_M$ such that its characteristic polynomial is a function only of u and θ. More precisely, the characteristic polynomial of the modified matrix should always be

$$\theta^{\frac{M+1}{2}} He_{M+1}\left(\frac{\lambda-u}{\sqrt{\theta}}\right), \tag{4.1}$$

which is obtained by substituting $f_{M-1}=f_M=0$ into (3.33). Recalling that only the equation of f_M is allowed to be changed, we summarize all the requirements and raise the following problem:

Find $M+1$ functions $a_j=a_j(\boldsymbol{w}_M)$, $j=1,\cdots,M+1$, such that for all $\rho,u,\theta,f_3,\cdots,f_M$,

$$\left|\lambda\mathbf{I}-\mathbf{A}_M-\sum_{j=1}^{M+1}a_j\mathbf{E}_{M+1,j}\right|=\theta^{\frac{M+1}{2}}He_{M+1}\left(\frac{\lambda-u}{\sqrt{\theta}}\right),$$

where $\mathbf{E}_{ij}$ denotes the matrix $\boldsymbol{e}_i\boldsymbol{e}_j^T$, and $\boldsymbol{e}_j$ is the unit vector whose j-th component is equal to 1.

If $a_j=a_j(\boldsymbol{w}_M)$, $j=1,\cdots,M+1$ is the solution of this problem, then a globally hyperbolic system can be obtained by substituting the matrix $\mathbf{A}_M$ in (3.2) with

$$\hat{\mathbf{A}}_M:=\mathbf{A}_M+\sum_{j=1}^{M+1}a_j\mathbf{E}_{M+1,j}. \tag{4.2}$$

The rest part of this section will be devoted to tackling this problem.

In order to simplify the notation, we use $S^{i,j}$ to denote the (i,j)-th minor of the matrix $\lambda\mathbf{I}-\mathbf{A}_M$, and define $S(k)$ as its *k-th order leading principal minor*, which is the determinant of the upper-left part of $\lambda\mathbf{I}-\mathbf{A}_M$ with k rows and k columns. According to the expression of $\mathbf{A}_M$ (3.3), it is not difficult to find

$$S^{M+1,1}=(-1)^M M!,\quad S^{M+1,2}=(-1)^{M-1}\frac{M!}{\rho}(\lambda-u), \tag{4.3a}$$

$$S^{M+1,3}=(-1)^{M-2}\frac{M!}{\rho}[(\lambda-u)^2-\theta], \tag{4.3b}$$

$$S^{M+1,j}=(-1)^{M+1-j}\frac{M!}{(j-1)!}S(j-1),\quad j=4,\cdots,M+1. \tag{4.3c}$$

Now we expand the characteristic polynomial of the matrix (4.2) as

$$\left|\lambda\mathbf{I}-\mathbf{A}_M-\sum_{j=1}^{M+1}a_j\mathbf{E}_{M+1,j}\right|=|\lambda\mathbf{I}-\mathbf{A}_M|-\sum_{j=1}^{M+1}(-1)^{M+1+j}a_jS^{M+1,j}. \tag{4.4}$$

In order that the above expression equals to (4.1), according to Corollary 3.1, we may choose a_j such that

$$\frac{(M+1)!}{2\rho}\left[\left((\lambda-u)^2-\theta\right)f_{M-1}+2(\lambda-u)f_M\right]+\sum_{j=1}^{M+1}(-1)^{M+1-j}a_jS^{M+1,j}\equiv 0. \tag{4.5}$$

The leading principal minor $S(k)$ is a polynomial in λ of degree k, since it is the characteristic polynomial of the $k\times k$ upper-left block of $\mathbf{A}_M$. Hence, $S^{M+1,j}$ is a polynomial in λ of degree $j-1$, which can be observed from (4.3). This observation directly leads to

$$a_j \equiv 0, \quad j=4,\cdots,M+1, \tag{4.6}$$

since the first term in (4.5) is a quadratic polynomial in λ. Then, we put (4.3a) and (4.3b) into (4.5), and some simplification gives

$$[(\lambda-u)^2-\theta]\left(\frac{M+1}{2}f_{M-1}+a_3\right)+(\lambda-u)[(M+1)f_M+a_2]+a_1\equiv 0. \tag{4.7}$$

Now, the choices of a_1, a_2, and a_3 are naturally given as

$$a_1=0, \quad a_2=-(M+1)f_M, \quad a_3=-\frac{M+1}{2}f_{M-1}. \tag{4.8}$$

For simplicity, the notation $\mathcal{R}_M$ is introduced as follows.

Definition 4.1. *The regularization term based on the characteristic speed correction is denoted as*

$$\mathcal{R}_M \triangleq \frac{M+1}{2}\left(2f_M\frac{\partial u}{\partial x}+f_{M-1}\frac{\partial\theta}{\partial x}\right). \tag{4.9}$$

Then we have the following theorem.

Theorem 4.1. *The moment system*

$$\frac{\partial \boldsymbol{w}_M}{\partial t}+\mathbf{A}_M\frac{\partial \boldsymbol{w}_M}{\partial x}-\mathcal{R}_M\boldsymbol{e}_{M+1}=0 \tag{4.10}$$

is strictly hyperbolic if $\theta>0$, and its characteristic speeds are

$$s_j=u+c_j\sqrt{\theta}, \quad j=1,\cdots,M+1, \tag{4.11}$$

where c_j is the j-th root of $He_{M+1}(x)$.

Proof. The equations (4.10) can be rewritten as

$$\frac{\partial \boldsymbol{w}_M}{\partial t}+\left(\mathbf{A}_M+\sum_{j=1}^{M+1}a_j\mathbf{E}_{M+1,j}\right)\frac{\partial \boldsymbol{w}_M}{\partial x}=0, \tag{4.12}$$

where a_j, $j=1,\cdots,M+1$ are defined in (4.6) and (4.8). As we have discussed above, (4.1) gives the characteristic polynomial of the matrix in the parentheses, which will be denoted by $\hat{\mathbf{A}}_M$ below as in (4.2). If $\theta>0$, one has

$$|\lambda\mathbf{I}-\hat{\mathbf{A}}_M|=\theta^{\frac{M+1}{2}}He_{M+1}\left(\frac{s_j-u}{\sqrt{\theta}}\right)=\theta^{\frac{M+1}{2}}He_{M+1}(c_j)=0. \tag{4.13}$$

Therefore, (4.11) gives all eigenvalues of $\hat{\mathbf{A}}_M$. Since the Hermite polynomial $He_{M+1}(x)$ has $M+1$ different zeros in $\mathbb{R}$ [17], all c_j's are distinct. Thus, the matrix $\hat{\mathbf{A}}_M$ has no duplicate eigenvalues, and hence is diagonalizable. This indicates that (4.10) is a strictly hyperbolic system. ∎

Comparing with the exact moment system (2.11), we see that the hyperbolic system (4.10) replaces $\partial f_{M+1}/\partial x$ by

$$-\frac{1}{M+1}\mathcal{R}_M = -f_M\frac{\partial u}{\partial x} - \frac{1}{2}f_{M-1}\frac{\partial \theta}{\partial x}. \tag{4.14}$$

This is a totally new way to regularize Grad's moment system.

REMARK 4.2. By modifying the last row of the matrix $\mathbf{A}_M$, the characteristic speeds can be appointed. Our regularization (4.10) selects a special set of characteristic speeds (4.11) such that they coincide with the Gauss-Hermite interpolation points. As discussed in [22], the characteristic speeds can be viewed as a sort of discretization of the distribution function. Therefore, the system (4.10) is similar to the "shifted and scaled discrete velocity model", with the expectation of spectral convergence when M goes to infinity. Meanwhile, unlike the ordinary discrete velocity model, the nonlinearity of Grad's moment systems introduced by shifting and scaling of the basis functions is preserved. Additionally, such regularization is only a slight modification based on the original Grad's moment system, and we will find in the next subsection that a number of interesting properties can be obtained.

4.2. Characteristic waves of the hyperbolic moment system. In this part, we will focus on the Riemann problem of (4.10). First, we claim that all characteristic fields of (4.10) are either genuinely nonlinear or linearly degenerate. To verify this, we write the right eigenvectors of $\hat{\mathbf{A}}_M$ in the following theorem.

THEOREM 4.3. *The right eigenvector of $\hat{\mathbf{A}}_M$ with eigenvalue $u+c_j\sqrt{\theta}$ is*

$$\boldsymbol{r}_j = (r_{j,1},\cdots,r_{j,M+1})^T, \quad j=1,\cdots,M+1, \tag{4.15}$$

where c_j is the j-th root of Hermite polynomial $He_{M+1}(x)$, and $r_{j,k}$ is defined as

$$\begin{gathered} r_{j,1}=\rho, \quad r_{j,2}=c_j\sqrt{\theta}, \quad r_{j,3}=(c_j^2-1)\theta, \\ r_{j,k}=\frac{He_{k-1}(c_j)}{(k-1)!}\rho\theta^{\frac{k-1}{2}} - \frac{c_j^2-1}{2}\theta f_{k-3} - c_j\sqrt{\theta}f_{k-2}, \quad k=4,\cdots,M+1. \end{gathered} \tag{4.16}$$

Proof. To prove this theorem, we need only to prove

$$\hat{\mathbf{A}}_M\boldsymbol{r}_j = (u+c_j\sqrt{\theta})\boldsymbol{r}_j. \tag{4.17}$$

Split $\hat{\mathbf{A}}_M$ by row as $\hat{\mathbf{A}}_M = (\boldsymbol{a}_1^T,\boldsymbol{a}_2^T,\cdots,\boldsymbol{a}_{M+1}^T)^T$, where $\boldsymbol{a}_k$ is the k-th row of $\hat{\mathbf{A}}_M$, $k=1,2,\cdots,M+1$. Thus (4.17) can be written as

$$(\boldsymbol{a}_1\boldsymbol{r}_j,\boldsymbol{a}_2\boldsymbol{r}_j,\cdots,\boldsymbol{a}_{M+1}\boldsymbol{r}_j)^T = (u+c_j\sqrt{\theta})\boldsymbol{r}_j, \quad \text{for } j=1,2,\cdots,M+1. \tag{4.18}$$

With the expression of $\hat{\mathbf{A}}_M$, the first four rows of (4.18) can be verified directly:

$$\begin{aligned} \boldsymbol{a}_1\boldsymbol{r}_j &= ur_{j,1}+\rho r_{j,2} = \rho(u+c_j\sqrt{\theta}) = r_{j,1}(u+c_j\sqrt{\theta}), \\ \boldsymbol{a}_2\boldsymbol{r}_j &= \theta/\rho\cdot r_{j,1}+ur_{j,2}+r_{j,3} = c_j\sqrt{\theta}(u+c_j\sqrt{\theta}) = r_{j,2}(u+c_j\sqrt{\theta}), \\ \boldsymbol{a}_3\boldsymbol{r}_j &= 2\theta r_{j,2}+ur_{j,3}+6/\rho\cdot r_{j,4} = (c_j^2-1)\theta(u+c_j\sqrt{\theta}) = r_{j,3}(u+c_j\sqrt{\theta}), \\ \boldsymbol{a}_4\boldsymbol{r}_j &= 4f_3r_{j,2}+\rho\theta/2\cdot r_{j,3}+ur_{j,4}+4r_{j,5} \\ &= He_3(c_j)\rho\theta^{3/2}/6\cdot(u+c_j\sqrt{\theta}) = r_{j,4}(u+c_j\sqrt{\theta}), \qquad \text{(only when } M\geqslant 4). \end{aligned} \tag{4.19}$$

For $5 \leq k \leq M$,

$$\begin{aligned}\boldsymbol{a}_k\boldsymbol{r}_j = &-\frac{\theta f_{k-2}}{\rho}r_{j,1} + kf_{k-1}r_{j,2} + \frac{1}{2}[(k-2)f_{k-2} + \theta f_{k-4}]r_{j,3} \\ &- \frac{3f_{k-3}}{\rho}r_{j,4} + \theta r_{j,j-1} + ur_{j,j} + kr_{j,j+1}.\end{aligned} \tag{4.20}$$

Then, we substitute (4.16) into (4.20) and get

$$\begin{aligned}\boldsymbol{a}_k\boldsymbol{r}_j &= \frac{He_{k-2}(c_j)}{(k-2)!}\theta^{k/2} + u\frac{He_{k-1}(c_j)}{(k-1)!}\theta^{(k-1)/2} + \frac{He_k(c_j)}{(k-1)!}\theta^{k/2} \\ &\quad + (-c_j\theta^{1/2})(c_j\sqrt{\theta}+u)f_{k-2} + [-\theta(c_j^2-1)(u+c_j\sqrt{\theta})/2]f_{k-3} \\ &= u\frac{He_{k-1}(c_j)}{(k-1)!}\theta^{(k-1)/2} + c_j\frac{He_{k-1}(c_j)}{(k-1)!}\theta^{k/2} \\ &\quad + (-c_j\theta^{1/2})(c_j\sqrt{\theta}+u)f_{k-2} + [-\theta(c_j^2-1)(u+c_j\sqrt{\theta})/2]f_{k-3} \\ &= (u+c_j\sqrt{\theta})r_{j,k}.\end{aligned} \tag{4.21}$$

For $k = M+1$, the situation is similar as $5 \leq k \leq M$. We expand $\boldsymbol{a}_{M+1}\boldsymbol{r}_j$ as

$$\begin{aligned}\boldsymbol{a}_{M+1}\boldsymbol{r}_j = &\frac{He_{M-1}(c_j)}{(M-1)!}\theta^{(M+1)/2} + u\frac{He_M(c_j)}{M!}\theta^{M/2} \\ &+ (-c_j\theta^{1/2})(c_j\sqrt{\theta}+u)f_{M-1} + [-\theta(c_j^2-1)(u+c_j\sqrt{\theta})/2]f_{M-2}.\end{aligned} \tag{4.22}$$

Here, c_j is the j-th root of Hermite polynomial $He_{M+1}(x)$. Hence, the recursion relation of Hermite polynomials gives

$$He_{M-1}(c_j) = \frac{c_j He_M(c_j)}{M}. \tag{4.23}$$

Substituting this equation into (4.22), we get

$$\boldsymbol{a}_{M+1}\boldsymbol{r}_j = (u+c_j\sqrt{\theta})r_{j,M+1}. \tag{4.24}$$

Collecting (4.19), (4.21), and (4.24), we finally arrive at (4.18). This completes the proof of the theorem. ∎

COROLLARY 4.2. *Each characteristic field of the hyperbolic system* (4.10) *is either genuinely nonlinear or linearly degenerate.*

Proof. Let $s_j = u + c_j\sqrt{\theta}$. We only need to verify that either $\nabla_{\boldsymbol{w}_M} s_j \cdot \boldsymbol{r}_j \equiv 0$ or $\nabla_{\boldsymbol{w}_M} s_j \cdot \boldsymbol{r}_j \not\equiv 0$ holds. Since

$$\nabla_{\boldsymbol{w}_M} s_j = \left(0, 1, \frac{1}{2}\frac{c_j}{\sqrt{\theta}}, 0, \cdots, 0\right)^T, \tag{4.25}$$

we have

$$\nabla_{\boldsymbol{w}_M} s_j \cdot \boldsymbol{r}_j = c_j\sqrt{\theta} + \frac{1}{2}c_j(c_j^2-1)\sqrt{\theta} = \frac{1}{2}c_j(c_j^2+1)\sqrt{\theta}. \tag{4.26}$$

If c_j is zero, the right hand side vanishes, while if c_j is nonzero, it is clear that $\nabla_{\boldsymbol{w}_M} s_j \cdot \boldsymbol{r}_j \not\equiv 0$. ∎

This corollary indicates the simplicity of characteristic waves in the solution of Riemann problems. Consider the following Riemann problem:

$$\begin{aligned}&\frac{\partial \boldsymbol{w}_M}{\partial t}+\hat{\mathbf{A}}_M\frac{\partial \boldsymbol{w}_M}{\partial x}=0,\\&\boldsymbol{w}_M(0,x)=\begin{cases}\boldsymbol{w}_M^L,\ x<0,\\ \boldsymbol{w}_M^R,\ x>0.\end{cases}\end{aligned}\tag{4.27}$$

A typical solution of this problem is the composition of at most $M+2$ intermediate states

$$\boldsymbol{w}_M^0=\boldsymbol{w}_M^L,\quad \boldsymbol{w}_M^1,\quad \cdots,\quad \boldsymbol{w}_M^N,\quad \boldsymbol{w}_M^{N+1}=\boldsymbol{w}_M^R,\qquad N\leqslant M,$$

which are connected by $N+1$ elementary waves: rarefaction waves, contact discontinuities, or shock waves. In order to get a full understanding of the hyperbolic moment system, these waves will be studied respectively below.

4.2.1. Rarefaction waves. As in all hyperbolic systems, the integral curves and the Riemann invariants are the major objects for the investigation of rarefaction waves. The parameterization of an integral curve of the vector field $\boldsymbol{r}_j$ satisfies

$$\tilde{\boldsymbol{w}}_M'(\zeta)=\boldsymbol{r}_j(\tilde{\boldsymbol{w}}_M(\zeta)),\tag{4.28}$$

where ζ is the parameter, and

$$\tilde{\boldsymbol{w}}_M(\zeta)=\left(\tilde{\rho}(\zeta),\tilde{u}(\zeta),\tilde{\theta}(\zeta),\tilde{f}_3(\zeta),\cdots,\tilde{f}_M(\zeta)\right)^T$$

denotes the integral curve in the $(M+1)$-dimensional phase space. For a given point $\boldsymbol{w}_M^0$ in the phase space, the integral curve through $\boldsymbol{w}_M^0$ can actually be analytically solved. Here we do not intend to write down the complete expressions, while the analytical solutions of $\rho(\zeta)$, $u(\zeta)$, and $\theta(\zeta)$ are given as

$$\tilde{\rho}(\zeta)=\rho^0\exp(\zeta),\tag{4.29a}$$

$$\tilde{u}(\zeta)=u^0+\frac{2c_j}{c_j^2-1}\sqrt{\theta^0}\left[\exp\left(\frac{c_j^2-1}{2}\zeta\right)-1\right],\tag{4.29b}$$

$$\tilde{\theta}(\zeta)=\theta^0\exp\left((c_j^2-1)\zeta\right).\tag{4.29c}$$

It is easy to verify that (4.29) satisfies the first three equations of (4.28). Note that in (4.28), only ρ, θ and f_{j-2}, f_{j-1} appear in the right hand side of f_j's equation, $j=3,\cdots,M$. Therefore, if the complete solution of $\tilde{\boldsymbol{w}}_M(\zeta)$ is needed, one can solve $f_j(\zeta)$ by explicit integration. Now we use (4.29) to give the j-th eigenvalue of $\mathbf{A}_M(\tilde{\boldsymbol{w}}_M(\zeta))$ as

$$s_j(\tilde{\boldsymbol{w}}_M(\zeta))=\tilde{u}(\zeta)+c_j\sqrt{\tilde{\theta}(\zeta)}=u^0+c_j\sqrt{\theta^0}+\frac{c_j^2+1}{c_j^2-1}c_j\sqrt{\theta^0}\left[\exp\left(\frac{c_j^2-1}{2}\zeta\right)-1\right].\tag{4.30}$$

It is not difficult to prove that $s_j(\tilde{\boldsymbol{w}}_M(\zeta))\gtrless s_j(\tilde{\boldsymbol{w}}_M^0)$ if and only if $c_j\zeta\gtrless 0$, which is helpful to predicate which part of the integral curve satisfies the entropy condition. And substitution of (4.29b) into (4.30) gives

$$s_j(\tilde{\boldsymbol{w}}_M(\zeta))-s_j(\tilde{\boldsymbol{w}}_M^0)=\frac{c_j^2+1}{2}(\tilde{u}(\zeta)-u^0).\tag{4.31}$$

Hence, $s_j(\tilde{\boldsymbol{w}}_M(\zeta)) \gtrless s_j(\tilde{\boldsymbol{w}}_M^0)$ holds if and only if $\tilde{u}(\zeta) \gtrless u^0$. Therefore, if the left state $\boldsymbol{w}_M^L$ and the right state $\boldsymbol{w}_M^R$ are connected by a single rarefaction wave, $u^L < u^R$ has to be satisfied, since the entropy condition requires $s_j(\boldsymbol{w}_M^L) < s_j(\boldsymbol{w}_M^R)$. Now let us turn to the pressure p. Equations (4.29a) and (4.29c) show that

$$\tilde{p}(\zeta) = \tilde{\rho}(\zeta)\tilde{\theta}(\zeta) = p^0 \exp(c_j^2 \zeta). \tag{4.32}$$

Therefore, the pressures on both sides of a rarefaction wave should satisfy

$$\begin{cases} p^L < p^R, & \text{if } c_j > 0, \\ p^L > p^R, & \text{if } c_j < 0. \end{cases}$$

Here we point out that the sign of c_j is as

$$c_j \begin{cases} > 0, & \text{if } j > (M+1)/2, \\ = 0, & \text{if } j = (M+1)/2, \\ < 0, & \text{if } j < (M+1)/2. \end{cases} \tag{4.33}$$

It is interesting that Riemann invariants exist for all genuinely nonlinear fields, and the following theorem gives its expressions.

THEOREM 4.4. *For the hyperbolic moment system* (4.10), *the Riemann invariants for the j-family are*

$$\begin{aligned} &R_1 = \rho\theta^{-1/(c_j^2-1)}, \quad R_2 = u - \frac{2c_j}{c_j^2-1}\sqrt{\theta}, \\ &R_k = C_{k,0}\rho\theta^{k/2} + \sum_{i=3}^{k} C_{k,i} f_i \theta^{(k-i)/2}, \quad k = 3, \cdots, M, \end{aligned} \tag{4.34}$$

where $C_{k,i}$ is defined recursively as

$$C_{k,k} = 1, \quad C_{k,k-1} = \frac{2c_j}{c_j^2-1}, \tag{4.35a}$$

$$C_{k,i} = \frac{1}{k-i}\left(C_{k,i+2} + C_{k,i+1}\frac{2c_j}{c_j^2-1}\right), \qquad i = 3, \cdots, k-2, \tag{4.35b}$$

$$C_{k,0} = \frac{2}{(1-c_j^2)k-2}\sum_{i=3}^{k}\frac{He_i(c_j)}{i!}C_{k,i}. \tag{4.35c}$$

Proof. We only need to prove

$$\nabla_{\boldsymbol{w}_M} R_k \cdot \boldsymbol{r}_j \equiv 0, \quad \forall k = 1, \cdots, M. \tag{4.36}$$

The verification in the cases $k=1$ and $k=2$ is straightforward:

$$\nabla_{\boldsymbol{w}_M} R_1 \cdot \boldsymbol{r}_j = \theta^{-1/(c_j^2-1)} \cdot \rho - \frac{1}{c_j^2-1}\rho\theta^{-1/(c_j^2-1)-1} \cdot (c_j^2-1)\theta = 0, \tag{4.37}$$

$$\nabla_{\boldsymbol{w}_M} R_2 \cdot \boldsymbol{r}_j = 1 \cdot c_j\sqrt{\theta} - \frac{c_j}{(c_j^2-1)\sqrt{\theta}} \cdot (c_j^2-1)\theta = 0. \tag{4.38}$$

If $k \geqslant 3$, the gradient of R_k is

$$\nabla_{\boldsymbol{w}_M} R_k = \Bigg(C_{k,0}\theta^{k/2}, 0, \frac{k}{2}C_{k,0}\rho\theta^{(k-1)/2} + \sum_{i=3}^{k-1}\frac{k-i}{2}C_{k,i}f_i\theta^{(k-i)/2-1}, \\ C_{k,3}\theta^{(k-3)/2}, C_{k,4}\theta^{(k-4)/2}, \cdots, C_{k,k}\theta^{(k-k)/2}\Bigg)^T. \tag{4.39}$$

With some rearrangement, $\nabla_{\boldsymbol{w}_M} R_k \cdot \boldsymbol{r}_j$ is simplified as

$$\begin{aligned} \nabla_{\boldsymbol{w}_M} R_k \cdot \boldsymbol{r}_j = & \left[\left(1 + \frac{k}{2}(c_j^2 - 1)\right)C_{k,0} + \sum_{i=3}^{k} C_{k,i}\frac{He_i(c_j)}{i!}\right]\rho\theta^{k/2} \\ & + \sum_{i=3}^{k-2}\left[\frac{c_j^2-1}{2}(k-i)C_{k,i} - c_jC_{k,i+1} - \frac{c_j^2-1}{2}C_{k,i+2}\right]f_i\theta^{(k-i)/2} \\ & + \left(\frac{c_j^2-1}{2}C_{k,k-1} - c_jC_{k,k}\right)f_{k-1}. \end{aligned} \tag{4.40}$$

We have that

- (4.35a) indicates that the last line of (4.40) is zero;
- (4.35b) indicates that the second line of (4.40) is zero;
- (4.35c) indicates that the first line of (4.40) is zero.

Thus (4.36) is proved. □

4.2.2. Contact discontinuities. According to the proof of Corollary 4.2, the contact discontinuities can only be found in the case of $c_j = 0$. Thus, if M is odd, no contact discontinuities exist in the characteristic waves. For contact discontinuities, the discussion on integral curves and Riemann invariants above is still valid. If we substitute $c_j = 0$ into (4.34), u, p, and f_3 can be found to be invariant across the contact discontinuity.

4.2.3. Shock waves. The discussion of shock waves requires additional care. As been well known, the jump condition on the shock wave is sensitive to the form of the hyperbolic equations. Therefore, before we give the Rankine-Hugoniot condition, it is necessary to rewrite (4.10) in an appropriate form. Though a conservative form is desired, the whole system can no longer be written as a conservation law since two terms are added to the last equation. Nevertheless, the conservative form of the first M equations remains. Thus (4.10) can actually be reformulated by M conservation laws and a single non-conservative equation. Precisely, if we let

$$\boldsymbol{q} = (q_0, \cdots, q_M)^T, \qquad q_j = \frac{1}{j!}\int_{\mathbb{R}} \xi^j f(\xi)\,\mathrm{d}\xi, \quad j = 0, \cdots, M, \tag{4.41}$$

(4.10) is reformulated as

$$\begin{aligned} & \frac{\partial q_j}{\partial t} + (j+1)\frac{\partial q_{j+1}}{\partial x} = 0, \quad j = 0, \cdots, M-1, \\ & \frac{\partial q_M}{\partial t} + \frac{\partial F(\boldsymbol{q})}{\partial x} - \mathcal{R}_M = 0. \end{aligned} \tag{4.42}$$

The relation between $\boldsymbol{q}$ and $\boldsymbol{w}_M$ is

$$f_j = \sum_{k=0}^{j} (-1)^{j-k} \frac{He_{j-k}(u/\sqrt{\theta})}{(j-k)!} \theta^{\frac{j-k}{2}} q_k, \quad u = q_1/q_0, \quad \theta = 2q_2/q_0 - (q_1/q_0)^2, \tag{4.43}$$

and $F(\boldsymbol{q})$ is defined as

$$F(\boldsymbol{q}) = (M+1) \sum_{k=0}^{M} (-1)^{M-k} \frac{He_{M+1-k}(u/\sqrt{\theta})}{(M+1-k)!} \theta^{\frac{M+1-k}{2}} q_k. \tag{4.44}$$

For convenience, we write (4.42) in the following form:

$$\frac{\partial \boldsymbol{q}}{\partial t} + \mathbf{B}(\boldsymbol{q}) \frac{\partial \boldsymbol{q}}{\partial x} = 0, \tag{4.45}$$

where $\mathbf{B}(\boldsymbol{q})$ is an $(M+1) \times (M+1)$ matrix.

Since (4.45) is still a non-conservative system, the DLM theory [13] is introduced when discussing the shock wave. A shock wave is a single jump discontinuity connecting two constant states $\boldsymbol{q}^L$ and $\boldsymbol{q}^R$ in a genuinely nonlinear field j, and $\boldsymbol{q}^L$, $\boldsymbol{q}^R$, and the propagation speed of the shock wave S_j should satisfy the following conditions:

- Generalized Rankine-Hugoniot condition:

$$\int_0^1 \left[S_j \mathbf{I} - \mathbf{B}\left(\boldsymbol{\Phi}(\nu; \boldsymbol{q}^L, \boldsymbol{q}^R)\right) \right] \frac{\partial \boldsymbol{\Phi}}{\partial \nu}(\nu; \boldsymbol{q}^L, \boldsymbol{q}^R) \, \mathrm{d}\nu = 0, \tag{4.46}$$

  where $\mathbf{I}$ is the identity matrix of order $M+1$, and $\boldsymbol{\Phi}(\nu; \boldsymbol{q}^L, \boldsymbol{q}^R)$ is a locally Lipschitz mapping satisfying

$$\boldsymbol{\Phi}(0; \boldsymbol{q}^L, \boldsymbol{q}^R) = \boldsymbol{q}^L, \quad \boldsymbol{\Phi}(1; \boldsymbol{q}^L, \boldsymbol{q}^R) = \boldsymbol{q}^R. \tag{4.47}$$

  We refer the readers to [13] for details. In Section 5, we will point out that the setup of $\boldsymbol{\Phi}$ is not crucial if the collision term presents.

- Entropy condition

$$s_j(\boldsymbol{q}^L) > S_j > s_j(\boldsymbol{q}^R). \tag{4.48}$$

It is obvious that the first M rows of (4.46) are independent of $\boldsymbol{\Phi}$; they are the same as the classical Rankine-Hugoniot conditions. This allows us to analyze the properties of the shock waves without regarding the form of $\boldsymbol{\Phi}$.

The first and second equations of (4.46) can be written as

$$\rho^L u^L - \rho^R u^R = S_j (\rho^L - \rho^R), \tag{4.49}$$

$$\rho^L (u^L)^2 + \rho^L \theta^L - \rho^R (u^R)^2 - \rho^R \theta^R = S_j (\rho^L u^L - \rho^R u^R). \tag{4.50}$$

Since $\rho^L \neq \rho^R$ and $\rho^L u^L \neq \rho^R u^R$ (otherwise $\boldsymbol{q}^L = \boldsymbol{q}^R$), one has

$$S_j = \frac{\rho^L u^L - \rho^R u^R}{\rho^L - \rho^R} \tag{4.51a}$$

$$= \frac{\rho^L (u^L)^2 + \rho^L \theta^L - \rho^R (u^R)^2 - \rho^R \theta^R}{\rho^L u^L - \rho^R u^R}. \tag{4.51b}$$

Putting (4.51a) into (4.48), and multiplying both sides with $(\rho^L-\rho^R)^2$, we get

$$\rho^L(u^L-u^R)(\rho^L-\rho^R)>c_j(\rho^L-\rho^R)^2\sqrt{\theta^R}, \tag{4.52a}$$

$$\rho^R(u^L-u^R)(\rho^L-\rho^R)<c_j(\rho^L-\rho^R)^2\sqrt{\theta^L}. \tag{4.52b}$$

If $c_j>0$, (4.52a) gives

$$(u^L-u^R)(\rho^L-\rho^R)>0. \tag{4.53}$$

Thus, we can divide both sides of (4.52) by $(u^L-u^R)(\rho^L-\rho^R)$ without changing the inequality sign, and the result is

$$\frac{\rho^L}{\sqrt{\theta^R}}>\frac{c_j(\rho^L-\rho^R)}{u^L-u^R}>\frac{\rho^R}{\sqrt{\theta^L}}, \tag{4.54}$$

from which one directly has

$$(\rho^L)^2\theta^L-(\rho^R)^2\theta^R>0. \tag{4.55}$$

Similarly, if $c_j<0$, we have

$$(u^L-u^R)(\rho^L-\rho^R)<0, \quad \text{and} \quad (\rho^L)^2\theta^L-(\rho^R)^2\theta^R<0. \tag{4.56}$$

Lemma 4.5. *For the hyperbolic moment system* (4.45)*, if* $\boldsymbol{q}^L$ *and* $\boldsymbol{q}^R$ *are connected by a* j*-shock wave, then the following inequalities hold:*

$$u^L>u^R, \qquad \text{and} \qquad \begin{cases} p^L>p^R, & \text{if } c_j>0, \\ p^L<p^R, & \text{if } c_j<0. \end{cases} \tag{4.57}$$

Proof. With some rearrangement, (4.51) can be reformulated as

$$(\rho^L-\rho^R)(\rho^L\theta^L-\rho^R\theta^R)=\rho^L\rho^R(u^L-u^R)^2. \tag{4.58}$$

Since the right hand side of (4.58) is positive, one and only one of the following two statements is true:

1. $\rho^L>\rho^R$ and $\rho^L\theta^L>\rho^R\theta^R$;
2. $\rho^L<\rho^R$ and $\rho^L\theta^L<\rho^R\theta^R$.

If $c_j>0$, equation (4.55) indicates that the first statement is true. Then, we can use (4.53) to conclude $u^L>u^R$. The conclusion for the case $c_j<0$ can be proved in the same way. ∎

Now, we summarize all our discussions on the entropy conditions of three types of waves in the following theorem:

Theorem 4.6. *For hyperbolic moment system* (4.10)*, if the wave of the* j*-th family is elementary, then its type can be determined by the value of* c_j *and the macroscopic velocities or pressures on both sides of the wave:*

	Velocity	Pressure
Contact discontinuity	$c_j=0,\ u^L=u^R$	$c_j=0,\ p^L=p^R$
Rarefaction wave	$c_j\neq 0,\ u^L<u^R$	$c_j>0,\ p^L<p^R$
		$c_j<0,\ p^L>p^R$
Shock wave	$c_j\neq 0,\ u^L>u^R$	$c_j>0,\ p^L>p^R$
		$c_j<0,\ p^L<p^R$

REMARK 4.7. It is not difficult to find that Euler equations are a special case of the proposed hyperbolic moment equations. In the case of $M=2$, we have $f_1 = f_2 = 0$, and thus the regularization vanishes. In other words, just like Grad's moment system, the hyperbolic system can be viewed as an extension of the Euler equations. Actually, all the discussions in this section, including the eigenvalues and eigenvectors, Riemann invariants, and the entropy condition, are valid for the 1D Euler equations with adiabatic index $\gamma=3$, while Grad's moment system is not able to preserve these criteria. In this respect, comparing with Grad's moment system, this regularized moment system is likely to be a more natural extension of Euler equations.

5. The case with collision terms

In this section, we will give a short discussion on the moment system with collision terms. For simplicity, the BGK collision operator [2] is considered. In this case, the Boltzmann equation (2.3) becomes

$$\frac{\partial f}{\partial t} + \xi \frac{\partial f}{\partial x} = \frac{1}{\tau}(f_M - f), \tag{5.1}$$

where τ is the relaxation time, and f_M is the Maxwellian:

$$f_M = \frac{\rho}{\sqrt{2\pi\theta}} \exp\left(-\frac{|\xi - u|^2}{2\theta}\right). \tag{5.2}$$

This equation leads to a very simple form of the collision term in the moment system as

$$\frac{\partial \boldsymbol{w}_M}{\partial t} + \mathbf{A}_M \frac{\partial \boldsymbol{w}_M}{\partial x} - \delta_H \mathcal{R}_M \boldsymbol{e}_{M+1} = -\frac{1}{\tau}\mathbf{P}\boldsymbol{w}_M, \tag{5.3}$$

where $\mathbf{P}$ is a diagonal matrix

$$\mathbf{P} = \mathrm{diag}\{0,0,0,1,\cdots,1\}, \tag{5.4}$$

and $\delta_H = 0$ corresponds to Grad's moment system, while $\delta_H = 1$ corresponds to the regularized moment system. Note that when considering the weak solution of (5.3), one still needs to rewrite (5.3) as equations of $\boldsymbol{q}$:

$$\begin{aligned} \frac{\partial q_j}{\partial t} + (j+1)\frac{\partial q_{j+1}}{\partial x} &= -\frac{1}{\tau}\mathcal{P}_j(q_0,\cdots,q_j), \quad j = 0,\cdots,M-1, \\ \frac{\partial q_M}{\partial t} + \frac{\partial F(\boldsymbol{q})}{\partial x} - \mathcal{R}_M &= -\frac{1}{\tau}\mathcal{P}_M(q_0,\cdots,q_M), \end{aligned} \tag{5.5}$$

where $\mathcal{P}_j$, $j=0,\cdots,M$ are the corresponding production terms. Then the first order derivative part of the last equation will still be treated using the DLM theory.

An important index that exhibits the quality of a collisional moment system is its order of accuracy in terms of τ. The conception of "order of accuracy" is based on the assumption that τ is a small quantity, and its precise definition can be found in [18, 7]. In [8], the order of magnitude for each moment has been deduced as

$$f_k \sim O(\tau^{\lceil k/3 \rceil}), \quad k \geqslant 3 \tag{5.6}$$

for the infinite moment system, which is obtained by the technique of Maxwellian iteration. It is easy to find that (5.6) remains correct for the regularized moment system (equation (5.3) with $\delta_H = 1$), since the order of $\mathcal{R}_M$ never exceeds the leading order term of f_M/τ. However, when $M=3m+1$, $m\geqslant 1$, the order of accuracy of the moment system is actually reduced by 2 with presentation of the regularization terms. This fact is not difficult to obtain and will be reported elsewhere. In general, the order of accuracy still goes to infinity as M increases.

Another issue is the choice of the path function $\mathbf{\Phi}$, which was introduced in (4.46). Let us restrict our discussion of its role in solving a Riemann problem of (5.3). First, we need to get some knowledge about the general behavior of the solution, referring to the careful study of the Riemann problem of 13-moment system in [21]. Roughly speaking, the solution shows a number of waves initially, then these waves are damping gradually, and eventually the solution tends to a smooth curve which is similar to the solution of Euler equations. The initial waves have no physical meaning due to the strong non-equilibrium which cannot be described by the moment system, while the solution gets close to the Boltzmann equation's solution only when the waves are fully dissipated. Later, this behavior is verified numerically for large number moment equations in [1], where the authors show that the speed of dissipation increases when the number of moments gets larger. It is expected that this also describes the evolution of the regularized moment system. Based on [21, 1], we have the following assertions for the regularized moment system:

1. If subshocks appear in the solution, the choice of $\mathbf{\Phi}$ indeed makes sense. In this situation, the system is inadequate for the description of the physical process, saying M needs to be increased.
2. $\mathbf{\Phi}$ affects the solution when the time t is very small. However, such solution has no physical significance, either. Only when the solution gets close enough to a smooth function, the moment system starts to show its ability to describe physics. Note that the smooth solution is independent of $\mathbf{\Phi}$; therefore, $\mathbf{\Phi}$ only affects the way in which the waves are damped, but does not affect the intrinsic constituent of the solution.

These two assertions indicate that the choice of $\mathbf{\Phi}$ is not crucial in solving a Riemann problem. We can simply use a linear function to connect any two states such that the numerical schemes can be constructed easily.

6. Numerical experimentation for a shock tube problem

In this section, a shock tube problem is studied numerically to show the behavior of the hyperbolic moment systems. We consider the following Riemann problem:

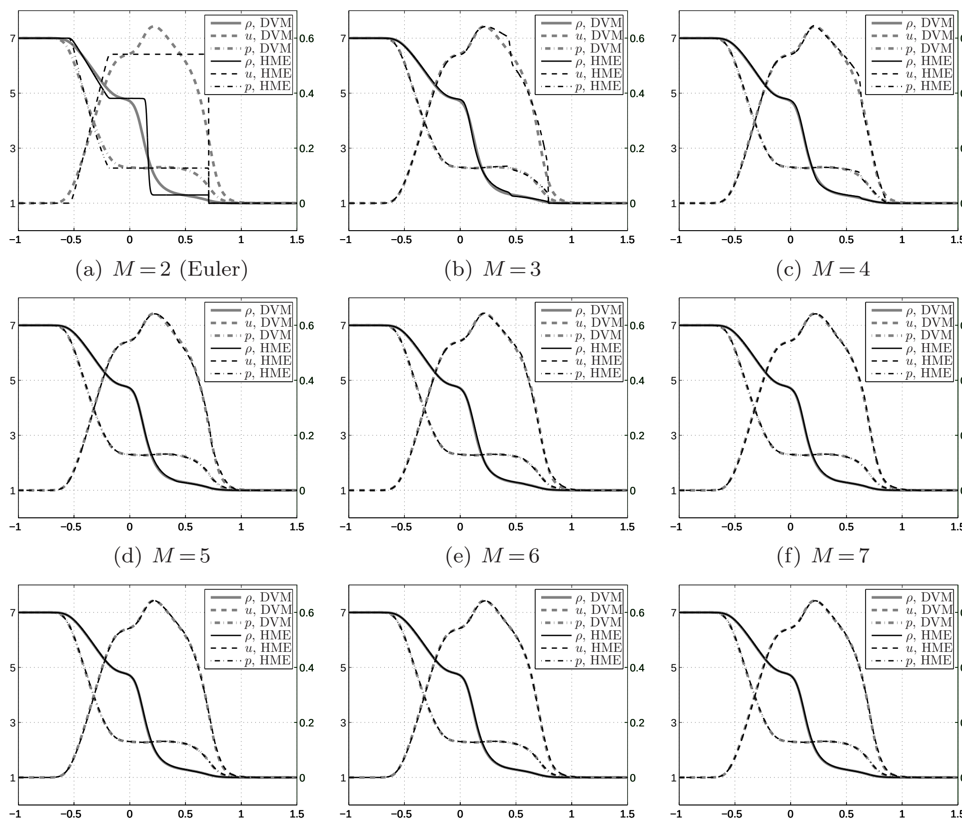
$$\begin{aligned}
&\frac{\partial \boldsymbol{w}_M}{\partial t}+\hat{\mathbf{A}}_M\frac{\partial \boldsymbol{w}_M}{\partial x}=-\frac{1}{\tau}\mathbf{P}\boldsymbol{w}_M,\\
&\boldsymbol{w}_M(0,x)=\begin{cases}\boldsymbol{w}_M^L,\ x<0,\\ \boldsymbol{w}_M^R,\ x>0,\end{cases}
\end{aligned}\tag{6.1}$$

where $\mathbf{P}$ is defined in (5.4) and the initial left and right states are

$$\boldsymbol{w}_M^L=(7,0,1,0,\cdots,0)^T,\quad \boldsymbol{w}_M^R=(1,0,1,0,\cdots,0)^T.\tag{6.2}$$

The relaxation time is chosen as $\tau=Kn/\rho$. Here two different cases $Kn=0.05$ and $Kn=0.5$ are considered. A nonconservative version of the HLL scheme [16] is employed to discretize the moment system.

The numerical results for $Kn=0.05$ with M ranging from 2 to 10 are listed in figure 6.1, in which the thin black lines are the numerical results of the hyperbolic moment equations (HME), and the thick gray lines are the results of Mieussens' discrete velocity model (DVM) [14], provided as reference solutions. The profiles of ρ, u, and p are drawn. It is clear that the solutions of hyperbolic moment systems converge to the solution of the Boltzmann equation when M increases. Note that when $M = 2$, the hyperbolic moment system is equivalent to the Euler equations, and the contact discontinuities and the shocks are obvious. When $M = 3$, a shock can still be found near $x = 0.75$. When M is greater than 5, the discontinuities are fully damped. This agrees with Torillhon's theory [1] that the discontinuities are damped faster when M is larger.



(a) $M = 2$ (Euler) (b) $M = 3$ (c) $M = 4$

(d) $M = 5$ (e) $M = 6$ (f) $M = 7$

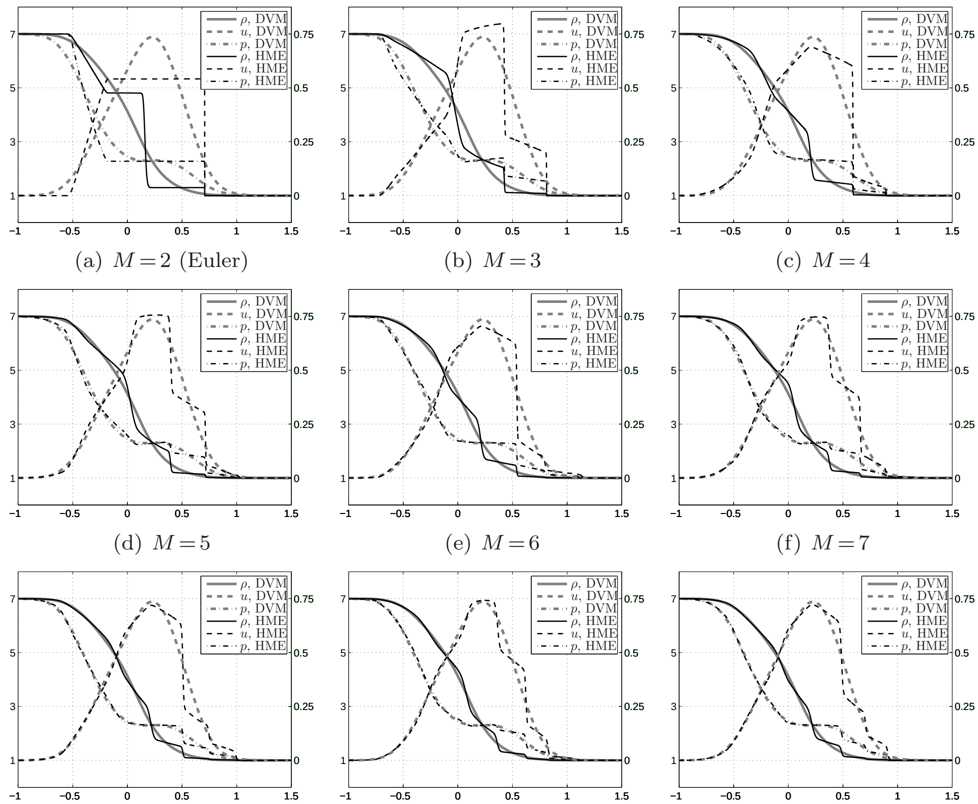
(g) $M = 8$ (h) $M = 9$ (i) $M = 10$

FIG. 6.1. *Numerical results of the shock tube problem for $Kn = 0.05$. The left y-axis is for ρ and p, and the right y-axis is for u.*

For a larger Knudsen number $Kn = 0.5$, the results are shown in figure 6.2. These results can also be considered as the solutions at $t = 0.03$ in the case of $Kn = 0.05$ (with proper scaling in the x direction). Thus these actually show the start-up phases of a shock tube by moment approximation. The discontinuities are clear for all choices of M, and the convergence can also be readily observed.

7. Concluding remarks

We regularize the 1D Grad's moment system to achieve global hyperbolicity for arbitrary order expansion. Fully investigations to the characteristic waves show that this set of equations may be a natural extension of Euler equations. Actually, the



(a) $M=2$ (Euler) (b) $M=3$ (c) $M=4$

(d) $M=5$ (e) $M=6$ (f) $M=7$

(g) $M=8$ (h) $M=9$ (i) $M=10$

FIG. 6.2. *Numerical results of the shock tube problem for $Kn=0.5$. The left y-axis is for ρ and p, and the right y-axis is for u.*

approach in this paper has been extended to two or three dimensional Grad's moment system, and the result is reported in a following paper.

Acknowledgment. Ruo Li is supported in part by the National Basic Research Program of China (2011CB309704), the National Science Foundation of China under grant 10731060 and NCET in China.

REFERENCES

[1] J.D. Au, M. Torrilhon, and W. Weiss, *The shock tube study in extended thermodynamics*, Phys. Fluids, 13, 2423–2432, 2001.

[2] P.L. Bhatnagar, E.P. Gross, and M. Krook, *A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems*, Phys. Rev., 94, 511–525, 1954.

[3] A.V. Bobylev, *The Chapman-Enskog and Grad methods for solving the Boltzmann equation*, Sov. Phys. Dokl., 27, 29–31, 1982.

[4] F. Brini, *Hyperbolicity region in extended thermodynamics with 14 moments*, Continuum Mech. Thermodyn., 13, 1–8, 2001.

[5] D. Burnett, *The distribution of velocities in a slightly non-uniform gas*, Proc. London Math. Soc., 39, 385–430, 1935.

[6] Z. Cai and R. Li, *Numerical regularized moment method of arbitrary order for Boltzmann-BGK equation*, SIAM J. Sci. Comput., 32, 2875–2907, 2010.

[7] Z. Cai, R. Li, and Z. Qiao, *NRxx simulation of microflows with Shakhov model*, SIAM J. Sci. Comput., 34, A339–A369, 2012.

[8] Z. Cai, R. Li, and Y. Wang, *Numerical regularized moment method for high Mach number flow*, Commun. Comput. Phys., 11, 1415–1438, 2012.

[9] H. Grad, *On the kinetic theory of rarefied gases*, Commun. Pure Appl. Math., 2, 331–407, 1949.

[10] S. Jin and M. Slemrod, *Regularization of the Burnett equations via relaxation*, J. Stat. Phys., 103, 1009–1033, 2001.

[11] M. Junk, *Domain of definition of Levermore's five-moment system*, J. Stat. Phys., 93, 1143–1167, 1998.

[12] C.D. Levermore, *Moment closure hierarchies for kinetic theories*, J. Stat. Phys., 83, 1021–1065, 1996.

[13] G.D. Maso, P.G. LeFloch, and F. Murat, *Definition and weak stability of nonconservative products*, J. Math. Pures Appl., 74, 483–548, 1995.

[14] L. Mieussens, *Discrete velocity model and implicit scheme for the BGK equation of rarefied gas dynamics*, Math. Models Methods Appl. Sci., 10, 1121–1149, 2000.

[15] I. Müller and T. Ruggeri, *Rational Extended Thermodynamics*, Second Edition, Springer Tracts in Natural Philosophy, Springer-Verlag, New York, 37, 1998.

[16] S. Rhebergen, O. Bokhove, and J.J.W. van der Vegt, *Discontinuous Galerkin finite element methods for hyperbolic nonconservative partial differential equations*, J. Comput. Phys., 227, 1887–1922, 2008.

[17] J. Shen and T. Tang, *Spectral and High-Order Methods with Applications*, Mathematics Monograph Series, Science Press, Beijing, P.R. China, 3, 2006.

[18] H. Struchtrup, *Macroscopic Transport Equations for Rarefied Gas Flows: Approximation Methods in Kinetic Theory*, Springer, 2005.

[19] H. Struchtrup and M. Torrilhon, *Regularization of Grad's 13 moment equations: Derivation and linear analysis*, Phys. Fluids, 15, 2668–2680, 2003.

[20] P.L. Tallec and J.P. Perlat, *Numerical analysis of Levermore's moment system*, Rapport de Recherche, INRIA Rocquencourt, Mar., 3124, 1997.

[21] M. Torrilhon, *Characteristic waves and dissipation in the 13-moment-case*, Continuum Mech. Thermodyn., 12, 289–301, 2000.

[22] M. Torrilhon, *Two dimensional bulk microflow simulations based on regularized Grad's 13-moment equations*, SIAM Multiscale Model. Simul., 5, 695–728, 2006.

[23] M. Torrilhon, *Hyperbolic moment equations in kinetic gas theory based on multi-variate Pearson-IV-distributions*, Commun. Comput. Phys., 7, 639–673, 2010.

[24] M. Torrilhon, J. Au, D. Reitebuch, and W. Weiss, *The Riemann-problem in extended thermodynamics*, in Hyperbolic Problems: Theory, Numerics, Applications, Vols I and II, H. Freistuühler and G. Warnecke, eds., International Series of Numerical Mathematics, Birkhäuser, 140, 79–88, 2001.