MAXIMUM ENTROPY MOMENT PROBLEMS AND EXTENDED EULER EQUATIONS

MICHAEL JUNK*

Abstract. The reduction of kinetic equations to moment systems leads to a closure problem because material laws have to be expressed in terms of the moment variables. In the maximum entropy approach, the closure problem is solved by assuming that the kinetic distribution function maximizes the entropy under some constraints. For the case of Boltzmann equation, the resulting hyperbolic moment systems are investigated. It turns out that the systems generally have non-convex domains of definition. Moreover, the equilibrium state is typically located on the boundary of the domain of definition where the fluxes are singular. This leads to the strange property that arbitrarily close to equilibrium the characteristic velocities of the moment system can be arbitrarily large.

Key words. maximum entropy, moment methods, moment realizability

AMS(MOS) subject classifications. 82C70, 35L65, 82B40

1. Introduction. In this article, a maximum entropy moment system is studied which is based on the Boltzmann equation of gas dynamics with scalar kinetic velocities. However, the observations also apply to more general cases: whenever the velocity space in the underlying kinetic equation is unbounded, when the entropy functional is essentially given by $H(f) = -\int_{\mathbb{R}^d} f \log f \, d\boldsymbol{c}$, when equilibrium states are related to Maxwellian distributions and when velocity moments of order four and higher are used (these assumptions are satisfied for the Boltzmann equation of gas dynamics with 3D velocities or the semiconductor Boltzmann equation with parabolic bands). It turns out that in such cases the equilibrium states are located on the boundary of the domain of definition of the maximum entropy system. Moreover, the flux is singular in these states which leads to a very unexpected behavior of the systems. Note that there are kinetic equations and corresponding entropy functionals where these problems do not appear, as for example the radiative transfer equation or the semiconductor Boltzmann equation with bounded Brillouin zone.

2. The maximum entropy approach. A standard mathematical model describing the flow of rarefied gases is given by the Boltzmann equation for the distribution function $f_\epsilon(t, \boldsymbol{x}, \boldsymbol{v})$ of the gas particles [1]

$$\frac{\partial f_\epsilon}{\partial t} + v_j \frac{\partial f_\epsilon}{\partial x_j} = \frac{1}{\epsilon} J(f_\epsilon) \qquad (2.1)$$

(Einstein's summation convention is used for repeated indices). Physically relevant quantities are obtained as functionals of f_ϵ in the form of velocity moments. Using $\langle \cdot, \cdot \rangle_{\boldsymbol{v}}$ to denote $\boldsymbol{v}$-integration, we introduce moment

---
* FB Mathematik, Universität Kaiserslautern, Erwin-Schrödinger-Straße, 67663 Kaiserslautern, Germany,(junk@mathematik.uni-kl.de).

functions $\psi_i : \mathbb{R}^d \mapsto \mathbb{R}$ and the corresponding moments

$$\mu_i = \langle f, \psi_i \rangle, \qquad i = 1, \ldots, m. \tag{2.2}$$

As most prominent examples, we mention

$$\begin{aligned} \rho &= \langle f, 1 \rangle & \text{mass density} \\ \rho \boldsymbol{u} &= \langle f, \boldsymbol{v} \rangle & \text{momentum density} \\ E &= \langle f, |\boldsymbol{v}|^2/2 \rangle & \text{energy density} \end{aligned}$$

Other polynomial moments include the stress tensor and the energy flux.

In the case of slightly rarefied gas where the Knudsen number ϵ is a small parameter, numerical simulations of (2.1) are very expensive because of the high dimensionality of the problem and the stiffness of the right hand side. In view of the fact that one is rather interested in functionals of f_ϵ than in f_ϵ itself, it is a natural idea to derive equations directly for the functionals. Multiplying (2.1) with $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_m)^T$ and integrating over $\boldsymbol{v}$, we obtain

$$\frac{\partial \boldsymbol{\mu}}{\partial t} + \frac{\partial}{\partial x_j} \langle f, v_j \boldsymbol{\psi} \rangle = \langle J(f), \boldsymbol{\psi} \rangle. \tag{2.3}$$

The system would be closed if the particle distribution could be expressed in terms of the moment vector $\boldsymbol{\mu}$

$$f_\epsilon(t, \boldsymbol{x}, \boldsymbol{v}) = F(\boldsymbol{\mu}(t, \boldsymbol{x}), \boldsymbol{v})$$

A method to obtain such a relationship is the maximum entropy approach where $F(\boldsymbol{\mu}, \boldsymbol{v})$ is taken as solution of the problem

$$\begin{aligned} &\text{maximize } H(f) = - \langle f \ln f - f, 1 \rangle \\ &\quad \text{with } f \geq 0 \text{ and } \langle f, \boldsymbol{\psi} \rangle = \boldsymbol{\mu} \end{aligned} \tag{2.4}$$

Variants and generalizations of this basic idea have been pursued by several authors [4, 2, 6, 11, 9]. We remark that the maximum entropy distribution $F(\boldsymbol{\mu}, \boldsymbol{v})$ can be viewed as *the* solution of $\langle f, \boldsymbol{\psi} \rangle = \boldsymbol{\mu}$ which contains least additional information in the sense of Shannon [5]. Moreover, the entropy maximization is in accordance with the Boltzmann equation (2.1). In fact, in the space homogeneous case, the entropy $H(f_\epsilon)$ of the solution of (2.1) grows in time and tends to the maximum for $t \to \infty$ with a speed proportional to $1/\epsilon$. Natural constraints are given by the fact that mass, momentum, and energy are conserved

$$\langle f_\epsilon, (1, \boldsymbol{v}, |\boldsymbol{v}|^2/2) \rangle = (\rho, \rho \boldsymbol{u}, E) = const.$$

In this case, the limit distribution is given by the Maxwellian

$$F(\rho, \rho \boldsymbol{u}, E, \boldsymbol{v}) = \frac{\rho}{(2\pi\theta)^{\frac{3}{2}}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}|^2}{2\theta} \right)$$

where $\theta = (E/\rho - |\boldsymbol{u}|^2/2)$ is the temperature of the gas.

For general ψ_i, the formal solution of (2.4) is obtained with the method of Lagrange multipliers. We introduce the Lagrange functional

$$\mathcal{L}(f, \boldsymbol{\lambda}) := H(f) + \boldsymbol{\lambda} \cdot (\langle f, \boldsymbol{\psi} \rangle - \boldsymbol{\mu})$$

where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers. The necessary condition that all directional derivatives vanish in the maximum F leads to

$$0 = \delta\mathcal{L}(F, \boldsymbol{\lambda}) = -\ln F + \boldsymbol{\lambda} \cdot \boldsymbol{\psi}$$

so that

$$F = \exp(\boldsymbol{\lambda} \cdot \boldsymbol{\psi}). \tag{2.5}$$

Finally, the Lagrange multipliers $\boldsymbol{\lambda}$ have to be chosen in such a way that the moment constraints $\boldsymbol{\mu} = \langle \exp(\boldsymbol{\lambda} \cdot \boldsymbol{\psi}), \boldsymbol{\psi} \rangle$ are satisfied. In [5], a strict proof is given that the maximum entropy problem (2.4) has the unique solution (2.5) whenever such multipliers $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\mu})$ can be found.

Using the maximum entropy distribution, we can now close the moment system (2.3) and obtain

$$\frac{\partial \boldsymbol{\mu}}{\partial t} + \frac{\partial}{\partial x_j} \boldsymbol{G}_j(\boldsymbol{\mu}) = \boldsymbol{P}(\boldsymbol{\mu}) \tag{2.6}$$

where $\boldsymbol{G}_j$ and $\boldsymbol{P}$ are given by

$$\boldsymbol{G}_j(\boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}, \boldsymbol{v}), v_j \boldsymbol{\psi}(\boldsymbol{v}) \rangle, \quad \boldsymbol{P}(\boldsymbol{\mu}) = \langle J(F(\boldsymbol{\mu}, \boldsymbol{v})), \boldsymbol{\psi}(\boldsymbol{v}) \rangle$$

Note that the domain of definition $\mathcal{U}$ of $\boldsymbol{G}_j$ and $\boldsymbol{P}$ is given by those moment vectors for which the solution $F(\boldsymbol{\mu}, \boldsymbol{v})$ of (2.4) exists.

In [9] it is shown that

$$\eta(\boldsymbol{\mu}) = -H(F(\boldsymbol{\mu})) = \langle F(\boldsymbol{\mu}, \boldsymbol{v}) \ln F(\boldsymbol{\mu}, \boldsymbol{v}) - F(\boldsymbol{\mu}, \boldsymbol{v}), 1 \rangle_{\boldsymbol{v}} \qquad \boldsymbol{\mu} \in \mathcal{U}$$

is a (locally) strictly convex entropy for the system (2.6) which implies that (2.6) has the nice property of being symmetric hyperbolic. Another nice feature of (2.6) is that it contains the Euler equations of gas dynamics as special case. Setting $\psi_1(\boldsymbol{v}) = 1$, $\psi_{i+1}(\boldsymbol{v}) = v_i$, $\psi_5(\boldsymbol{v}) = |\boldsymbol{v}|^2/2$, equation (2.6) turns into the Euler system. By taking additional moment functions, one can thus generate hierarchies of symmetric hyperbolic systems which all contain the equations of mass, momentum and energy conservation but which allow, for example, a non-trivial heat flux and thus extend the classical Euler system. As examples, we mention

10 moments:	1	v_i	$|v|^2/2$	$v_i v_j - |\boldsymbol{v}|^2 \delta_{ij}$		
14 moments:	1	v_i	$|v|^2/2$	$v_i v_j - |\boldsymbol{v}|^2 \delta_{ij}$	$|\boldsymbol{v}|^2 v_i$	$|\boldsymbol{v}|^4$
21 moments:	1	v_i	$|v|^2/2$	$v_i v_j - |\boldsymbol{v}|^2 \delta_{ij}$	$v_i v_j v_k$	$|\boldsymbol{v}|^4$
26 moments:	1	v_i	$|v|^2/2$	$v_i v_j - |\boldsymbol{v}|^2 \delta_{ij}$	$v_i v_j v_k$	$|\boldsymbol{v}|^2 v_i v_j$
35 moments:	1	v_i	$|v|^2/2$	$v_i v_j - |\boldsymbol{v}|^2 \delta_{ij}$	$v_i v_j v_k$	$v_i v_j v_k v_l$

Despite the nice properties listed above, the models involving velocity moments of fourth order or higher suffer from severe drawbacks. For example, in the case of fourteen moments, the domain of definition $\mathcal{U}$ of $\eta, \boldsymbol{G}_j$, and $\boldsymbol{P}$ is not convex. Moreover, the equilibrium states are located on $\partial\mathcal{U}$ so that a typical Riemann initial value, consisting of two separated equilibrium states, leads to a problem where little can be said about solvability. Note that the existence theory for solutions to Riemann problems with small jumps requires that the initial states are in the interior of the domain of definition and *not* on the boundary [12]. Also, the local existence result for smooth solutions [10] requires that the range of the initial values is contained in a compact subset of the *interior* of the domain of definition. Hence, initial values obtained by adding smooth and compactly supported disturbances to an equilibrium state are *not* covered by the local existence result if equilibrium states are located on $\partial\mathcal{U}$.

As we will see in the next section, the fact that equilibrium points are located on ∂U is not a mere technicality. For a model problem, we show that the flux function of the maximum entropy system is actually singular in equilibrium points which leads to the surprising result that arbitrarily close to equilibrium, the maximal characteristic velocity of the system becomes arbitrarily large.

3. The model problem. To illustrate the problems that can occur in the maximum entropy approach, let us consider a model problem based on the Boltzmann equation with scalar velocities $v \in \mathbb{R}$ and moment functions $\psi(v) = (1, v, v^2, v^3, v^4)^T$. In this case, the maximum entropy system has the form

$$\frac{\partial \boldsymbol{\mu}}{\partial t} + \frac{\partial}{\partial x} \boldsymbol{G}(\boldsymbol{\mu}) = \boldsymbol{P}(\boldsymbol{\mu})$$

where $\boldsymbol{G} : \mathcal{U} \mapsto \mathbb{R}^5$ is given by

$$\boldsymbol{G}(\boldsymbol{\mu}) = (\mu_2, \mu_3, \mu_4, \mu_5, \left\langle F(\boldsymbol{\mu}, v), v^5 \right\rangle_v)^T \qquad (3.1)$$

In the following, we are investigating only this flux function so that the production terms $\boldsymbol{P}(\boldsymbol{\mu}) = \langle J(F(\boldsymbol{\mu})), \boldsymbol{\psi} \rangle$ need not be specified in detail. In fact, since $\boldsymbol{G}$ is independent of the collision model, the results automatically apply to a whole class of kinetic equations (see for example [3] for the case of Fokker-Planck equation).

We start with the investigation of the domain of definition $\mathcal{U}$ of $\boldsymbol{G}$. In order to be able to plot $\mathcal{U}$, we consider its intersection with the affine hyperplane $E = \{(1, 0, 1, q, s + 3)^T : q, s \in \mathbb{R}\}$ because $\mathcal{U} \cap E$ contains already full information about $\mathcal{U}$: to find out whether the maximum entropy problem (2.4) is solvable for some given $\boldsymbol{\mu}$ (i.e. whether $\boldsymbol{\mu} \in \mathcal{U}$), we first map $\boldsymbol{\mu}$ by some algebraic operation to its first three components and a normalized vector $\mu^* \in E$. If the maximum entropy problem can be solved for μ^* with result $\exp(\boldsymbol{\lambda}^* \cdot \boldsymbol{\psi})$ (i.e. if $\mu^* \in \mathcal{U} \cap E$), an application of the

algebraic operation in the reverse direction yields the solution $\exp(\boldsymbol{\lambda} \cdot \boldsymbol{\psi})$ of the original problem. On the other hand, if $\mu^* \notin \mathcal{U} \cap E$ it follows that $\mu \notin \mathcal{U}$.

The algebraic operation is based on a normalization of distribution functions. If $f \geq 0$ is given, we calculate ρ, u, θ which are equivalent to the first three moments of f

$$\rho = \langle f, 1 \rangle, \ \ u = \frac{1}{\rho} \langle f, v \rangle, \ \ \theta = \frac{1}{\rho} \langle f, v^2 - u^2 \rangle.$$

Then, f is mapped to (ρ, u, θ, f^*) where

$$f^*(w) = \frac{1}{\rho} f(\sqrt{\theta} w + u).$$

The moments of f^* have the form

$$\langle f^*, \psi \rangle = (1, 0, 1, q, 3 + s)^T, \qquad q, s \in \mathbb{R}$$

and thus are contained in E. Conversely, one can go back from (ρ, u, θ, f^*) to f using

$$f(v) = \frac{\rho}{\sqrt{\theta}} f^*((v - u)/\sqrt{\theta}).$$

Applying this procedure to the Maxwellian distribution

$$f(v) = \frac{\rho}{\sqrt{2\pi\theta}} \exp\left(-\frac{(v - u)^2}{2\theta} \right)$$
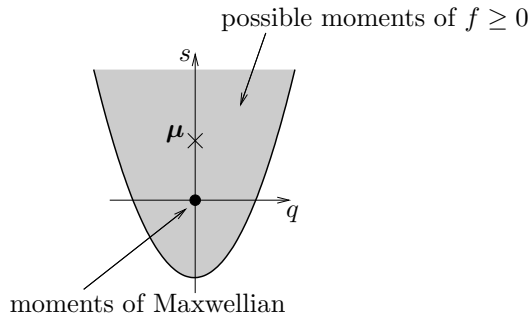
we recover the normalized Maxwellian

$$F^*(w) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2} \right)$$

with moments $(1, 0, 1, 0, 3)^T$ which are related to the origin $q = s = 0$ in the plane E.

By definition of the maximum entropy problem, all maximum entropy distributions are non-negative so that $\mathcal{U} \cap E$ is a subset of all moments of normalized non-negative distribution functions. In [7], it is shown that this set is given by $s \geq q^2 - 2$ (see fig. 1). While the maximum entropy problem is solvable for most moment vectors in this parabolic domain, there are some exceptions. To see this, we take $\boldsymbol{\mu}$ on the s-axis above the equilibrium point (see fig. 1), for example, $\boldsymbol{\mu} = (1, 0, 1, 0, 4)^T$. In a contradiction argument, we assume that the corresponding maximum entropy problem is solvable with solution f. Since f also solves the moment problem

$$\langle f, (1, v, v^2) \rangle = (1, 0, 1) \tag{3.2}$$

possible moments of $f \geq 0$

PSfrag replacements

$$\boldsymbol{\mu}$$

s

q

moments of Maxwellian

Fig. 1. *Reduced space of moments*

we conclude that $H(f) < H(F^*)$ because F^* is the maximum entropy solution for the constraint (3.2). However, we can show that

$$\sup\{H(f) : \langle f, \boldsymbol{\psi} \rangle = (1, 0, 1, 0, 4)^T\} = H(F^*)$$

by constructing a sequence of functions f_n which all have the correct moments and whose entropy converges to $H(F^*)$. The detailed construction can be found in [7] (see also [8] for a more general approach which includes higher dimensional cases). Here we just give the basic idea: we construct f_n by adding a small perturbation to the Maxwellian F^* as indicated in fig. 2. The perturbation should be symmetrically concentrated at $v = n$ and

PSfrag replacements

$$\frac{1}{n^4}$$

$-n$ 0 n v
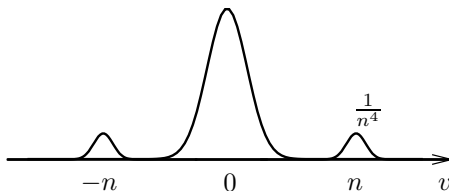
Fig. 2. *Small perturbation of the Maxwellian*

$v = -n$ so that odd moments of the distribution function vanish exactly. If the height of the little humps is taken of the order $1/n^4$, we can estimate the moments of the perturbation as $(\mathcal{O}(n^{-4}), 0, \mathcal{O}(n^{-2}), 0, \mathcal{O}(1))^T$. Hence, for large n, the perturbation practically contributes only to the highest moment. After a suitable rescaling, one can achieve that the combined distribution function consisting of Maxwellian and perturbation has the moment vector $(1, 0, 1, 0, 4)^T$ for every n larger than some n_0. Letting n tend to infinity, we find $f_n \to F^*$ in $\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^\infty(\mathbb{R})$ and also $H(f_n) \to H(F^*)$ which yields the contradiction.

In summary, we can say that very small packets of very fast particles are the reason for non-solvability of the maximum entropy problem. These packets are negligible for the value of the entropy but are important for the

highest moment. Since the argument can be repeated for every moment vector which differs from the one of the Maxwellian by a positive contribution in the highest moment, we find that the domain of definition $\mathcal{U} \cap E$ does not contain the half line $s > 0$. A detailed analysis shows that these are the only exceptions [7, 8]. It immediately follows from fig. 3 that the
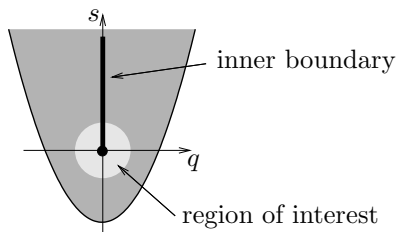

FIG. 3. *Non-convex domain of definition* $\mathcal{U}$

domain of definition is not convex and that the equilibrium point is located on ∂U so that the most interesting states for practical purposes are close to ∂U.

An estimate of the flux (3.1) close to the inner boundary $\{(0, s) : s > 0\}$ has been derived in [7]. It turns out that

$$\left| \left\langle F(\boldsymbol{\mu}, v), v^5 \right\rangle_v \right| \geq \frac{s}{|q|} \qquad q \neq 0$$

so that the flux is clearly singular at $\{(0, s) : s > 0\}$. If the equilibrium is approached on a curve $q \mapsto (q, q^\alpha)$ with $0 < \alpha < 1$, we even find that the flux is *singular in equilibrium*. The singularity of the flux is also reflected by the characteristic velocities of the system. To demonstrate this phenomenon, we calculate the maximal characteristic speed of the five-moment system along two curves given by constant $q = 0.1$ and $q = 0.05$ in the moment space $\mathcal{U} \cap E$ (see fig. 4). In equilibrium, the maximal char-
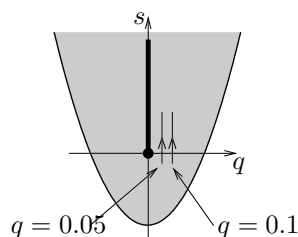

FIG. 4. *Curves along which maximal characteristic speeds are calculated*

acteristic speed is $\sigma_{max} \approx 3$. Essentially the same value is obtained on both curves for $s < 0$ (see fig. 5). However, if s enters the region $s > 0$, the maximal characteristic speed suddenly increases very strongly. For $q = 0.1$,
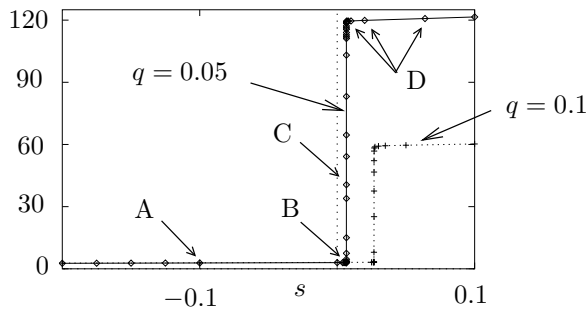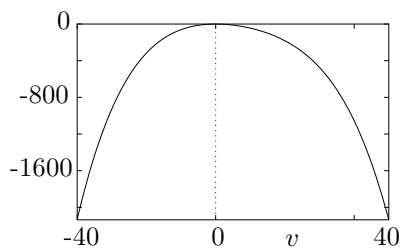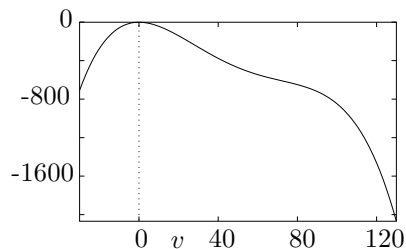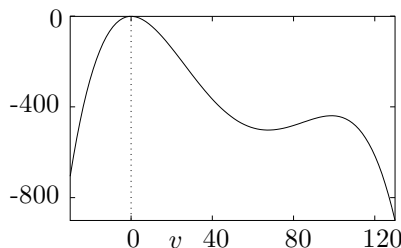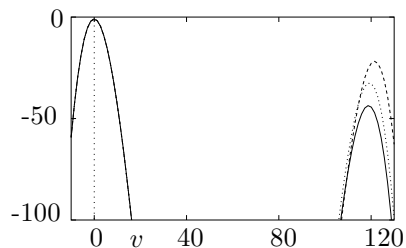
FIG. 5. *Maximal characteristic speeds of the 5-moment system*

FIG. 6. $\ln(F)$ *at point A in fig. 5*

FIG. 7. $\ln(F)$ *at point B in fig. 5*

FIG. 8. $\ln(F)$ *at point C in fig. 5*

FIG. 9. $\ln(F)$ *at point D in fig. 5*

it rises to $\sigma_{max} \approx 60$ and for $q = 0.05$ to $\sigma_{max} \approx 120$. Calculations for even smaller q show that the values of σ_{max} increase further in accordance with the singularity of the flux. To understand this behavior, we consider the maximum entropy distribution functions belonging to the moment vectors along the curve $q = 0.05$. For moment vectors with $s < 0$, the logarithm of $F(\boldsymbol{\mu}, v)$ appears like a negative parabola (see fig. 6). This is reasonable because $\ln F^* = -v^2/2 - \ln \sqrt{2\pi}$. Nevertheless, a slight bias to the set of positive velocities is visible in fig. 6. which reflects the fact that the third moment q is positive. If s is increased and approaches the point B in fig. 5, the corresponding distribution function develops a shoulder at $v \approx 80$

(see fig. 7). Increasing s further, the shoulder develops into a second peak (fig. 8) which rises for increasing s. In fig. 9, three consecutive distribution functions corresponding to D in fig. 5 are shown. Note that the position of the peaks is directly related to the characteristic velocity. Physically it means that the maximum entropy distribution contains a small packet of very fast particles (the value of the distribution function is only around $\exp(-30)$ at $v \approx 120$). Hence, it is not surprising that also in the moment system information is transported with this speed.

Recalling that the high velocity particles are the reason for the breakdown of the solvability of the maximum entropy problem, it is interesting to note that the same phenomenon leads to structural deficiencies of the moment system like the singular flux or the large characteristic velocities.

The problem of fast particles in the maximum entropy distribution, on the other hand, is related to the fact that the velocity domain is unbounded, that moment functions are used which grow faster than v^2 for $|v| \to \infty$, and that the entropy functional does not control small contributions at high velocities. Thus, the structural deficiencies can be avoided by removing any of these properties, i.e. to work on bounded velocity spaces, or to use moment functions with slow growth or to switch over to other entropy functionals. For the specific application of Boltzmann equation, however, none of these modifications is very natural.

4. Conclusion. Although the maximum entropy approach formally leads to moment systems with very nice properties, we have seen that the equations can suffer from severe drawbacks if the underlying maximum entropy problem is not solvable. For a model problem, we have demonstrated that the domain of definition is not convex, that equilibrium states are located on the boundary of the domain of definition and that the flux is singular in equilibrium points, leading to arbitrarily large characteristic speeds arbitrarily close to equilibrium. From a microscopic point of view, the problems occur because of small groups of very fast particles appearing in the maximum entropy distribution functions. We stress that such a situation can only occur if the underlying velocity space is unbounded.

REFERENCES

[1] C. CERCIGNANI, *The Boltzmann Equation And Its Applications*, Springer, 1988.
[2] W. DREYER, *Maximization of the entropy in non–equilibrium*, J. Phys. A: Math. Gen., **20**, 1987, 6505–6517.
[3] W. DREYER, M. JUNK, M. KUNIK, *On the approximation of kinetic equations by moment systems*, appears in Nonlinearity
[4] B.C. EU, *A modified moment method and irreversible thermodynamics*, J. Chem. Phys. **73**, 1980, 2958–2969.
[5] S. IHARA, *Information Theory for Continuous Systems*, World Scientific, 1993.
[6] A.N. GORBAN AND I.V. KARLIN, *Method of Invariant Manifolds and the Regularization of Acoustic Spectra*, Transp. Theory Stat. Phys. 23, 1994, 559–632.

[7] M. JUNK, *Domain of Definition of Levermore's Five-Moment System*, J. Stat. Phys., 93, 1998, 1143–1167.

[8] M. JUNK, *Maximum entropy for reduced moment problems*, Math. Models Methods Appl. Sci., to appear.

[9] C. D. LEVERMORE, *Moment Closure Hierarchies for Kinetic Theories*, J. Stat. Phys., 83, 1996, 1021–1065.

[10] A. MAJDA, *Compressible fluid flow and systems of conservation laws in several space variables*, Springer, 1984.

[11] I. MÜLLER, T. RUGGERI, *Rational Extended Thermodynamics*. 2nd Edition, Springer Tracts in Natural Philosophy, Springer New York, 1998.

[12] J. SMOLLER, *Shock waves and reaction-diffusion equations*, Springer, 1983.