Maximum Entropy Moment Systems and Galilean Invariance

M. Junk[*] A. Unterreiter[†]

December 7, 2001

Abstract

In this article, we investigate the maximum entropy moment closure in gas dynamics. We show that the usual choice of polynomial weight functions may lead to hyperbolic systems with an unpleasant state space: equilibrium states are boundary points with possibly singular fluxes. In order to avoid singularities, the necessity arises to find weight functions which growing sub-quadratically at infinity. Unfortunately, this requirement leads to a conflict with Galilean invariance of the moment systems because we can show that rotational and translational invariant, finite dimensional function spaces necessarily consist of polynomials.

1 Introduction

Generally speaking, maximum entropy moment systems are extensions of the Euler equations of gas dynamics where non-trivial heat flux and stresses are taken into account. The derivation is based on the Boltzmann equation which models the microscopic particle movement in rarefied gases. By taking suitable weighted averages of the particle distribution functions (velocity moments), the variables of the moment system are obtained. Their time evolution follows from the Boltzmann equation with the additional assumption that the particle distribution function has the particular form of an entropy maximizing function which just depends on the considered moments. For a detailed description of this approach, we refer to [13] where several structural features of the resulting moment systems are investigated. The systems are strictly hyperbolic because a convex entropy exists. Moreover, Galilean invariance is guaranteed if the weight functions which define the moment variables, span a translational and rotational invariant space. This condition is naturally satisfied by certain classes of polynomials which are generally used as weight functions [7, 2, 15, 13].

However, for certain model problems [8, 3, 10], polynomial weight functions lead to bad side effects: once polynomials weights are used which grow super-quadratically, the state space of the moment system is not convex and the equilibria are located on the boundary of the state space. Moreover, the flux

[*]Universität Kaiserslautern, D–67663 Kaiserslautern, (junk@mathematik.uni-kl.de).

[†]Universität Kaiserslautern, D–67663 Kaiserslautern, (unterreiter@mathematik.uni-kl.de).

function is possibly singular which has the surprising effect that arbitrarily close to equilibrium, arbitrarily large characteristic speeds can appear.

In this article, we extend these results to general maximum entropy moment systems. We show that equilibrium states are always on the boundary of the state space if polynomial weight functions of degree larger than two are used (see Section 3).

In view of this result, one may try to use other non-polynomial weight functions which grow sub-quadratically at infinity. Unfortunately, this is not possible if we want to keep the property that the weight functions span a rotational and translational invariant space which is a prerequisite for Galilean invariance of the moment system (see Section 4).

This demonstrates the delicate relation between Galilean invariance on the one hand and a reasonable structure of the state space on the other hand: unpleasant features of the maximum entropy systems in gas dynamics can not be avoided by choosing non-polynomial weight functions if one wants to keep Galilean invariance.

Our results are in a certain sense complementary to the related work [16] which shows that Galilean invariant systems with a *particular structure* (where most fluxes coincide with variables) are *necessarily* polynomial moment systems obtained from the Boltzmann equation. Complementing this, we show that Galilean invariant systems with an *arbitrary structure* can only be moment systems based on translational and rotational invariant families of weight functions if the weights are suitable polynomials.

2 Maximum entropy in gas dynamics

The mathematical description of slightly rarefied gas flows requires a system of evolution equations for macroscopic quantities like mass density, flow velocity, temperature, heat flux, and stresses. The system should obey the basic physical properties of *Galilean invariance* and, in view of the second law of thermodynamics, it should satisfy a suitable H-theorem. A standard approach to derive such a system is based on the Boltzmann equation for the distribution function $f(t, \boldsymbol{x}, \boldsymbol{v})$ of the gas particles [1]

$$\frac{\partial f}{\partial t} + v_j \frac{\partial f}{\partial x_j} = \frac{1}{\epsilon}\mathcal{C}(f), \qquad \boldsymbol{x}, \boldsymbol{v} \in \mathbb{R}^d, t \geq 0. \tag{1}$$

Note that in (1) and in the sequel, we make use of Einstein's summation convention. We will not specify the collision operator $\mathcal{C}$ in detail but only mention those properties which are important for our investigations. Before going into details, we remark that many physically relevant quantities are obtained as velocity moments of f. Using $\langle \cdot, \cdot \rangle$ to denote $\boldsymbol{v}$-integration, let us consider linearly independent weight functions $a_i : \mathbb{R}^d \mapsto \mathbb{R}$ and the corresponding moments

$$\rho_i(t, \boldsymbol{x}) = \langle f(t, \boldsymbol{x}), a_i \rangle = \int_{\mathbb{R}^d} a_i(\boldsymbol{v}) f(t, \boldsymbol{x}, \boldsymbol{v})\, d\boldsymbol{v}, \qquad i = 1, \ldots, n. \tag{2}$$

As most prominent examples, we mention $a_1, \ldots, a_{d+2} = 1, v_1, \ldots, v_d, \|\boldsymbol{v}\|^2$ with associated macroscopic quantities

$$\begin{aligned} \varrho &= \langle f, 1 \rangle && \text{mass density} \\ \varrho \boldsymbol{u} &= \langle f, \boldsymbol{v} \rangle && \text{momentum density} \\ E &= \left\langle f, |\boldsymbol{v}|^2/2 \right\rangle && \text{energy density} \end{aligned} \tag{3}$$

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

In the case of a slightly rarefied gas (i.e. if the Knudsen number ϵ in (1) is a small parameter), numerical simulations of (1) are very expensive because of the high dimensionality of the problem and the stiffness of the right hand side. Since, from a physical point of view, one is rather interested in functionals of f than in f itself, it is a natural idea to derive equations directly for the functionals. Multiplying (1) with $\boldsymbol{a} = (a_1, \ldots, a_n)^T$ and integrating over $\boldsymbol{v}$, we obtain

$$\frac{\partial \boldsymbol{\rho}}{\partial t} + \frac{\partial}{\partial x_j} \langle f, v_j \boldsymbol{a} \rangle = \langle \mathcal{C}(f), \boldsymbol{a} \rangle . \tag{4}$$

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

$$f(t, \boldsymbol{x}, \boldsymbol{v}) = F(\boldsymbol{\rho}(t, \boldsymbol{x}), \boldsymbol{v})$$

for some suitable F. A method to obtain F is the maximum entropy approach where $F(\boldsymbol{\rho}, \boldsymbol{v})$ is the solution of the problem

$$\begin{gathered} \text{maximize } -H(f) \\ \text{with } f \geq 0 \text{ and } \langle f, \boldsymbol{a} \rangle = \boldsymbol{\rho} \end{gathered} \tag{5}$$

where $H(f) = \langle f \ln f - f, 1 \rangle$ is a strictly convex functional (we call $H(f)$ the entropy, at variance with much of the physics literature which would denote it as neg-entropy). Variants and generalizations of this basic idea have been pursued by several authors (see, for example, [4, 2, 6, 15, 13]).

The formal solution of (5) is obtained with the method of Lagrange multipliers. We introduce the Lagrange functional

$$L(f, \boldsymbol{\lambda}) := H(f) - \boldsymbol{\lambda} \cdot (\boldsymbol{\rho} - \langle f, \boldsymbol{a} \rangle)$$

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

$$0 = \delta L(f_{\boldsymbol{\lambda}}, \boldsymbol{\lambda}) = -\ln f_{\boldsymbol{\lambda}} + \boldsymbol{\lambda} \cdot \boldsymbol{a}$$

so that

$$f_{\boldsymbol{\lambda}} = \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}). \tag{6}$$

In (6), the Lagrange multipliers $\boldsymbol{\lambda}$ have to satisfy the moment constraints $\boldsymbol{\rho} = \langle f_{\boldsymbol{\lambda}}, \boldsymbol{a} \rangle$. It can be shown [11] that the maximum entropy problem (5) has the unique solution (6) whenever such multipliers $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\rho})$ can be found. In this case, we set $F(\boldsymbol{\rho}, \boldsymbol{v}) = f_{\boldsymbol{\lambda}(\boldsymbol{\rho})}(\boldsymbol{v})$.

Using the maximum entropy distribution, we can now close the moment system (4) and obtain

$$\frac{\partial \boldsymbol{\rho}}{\partial t} + \frac{\partial}{\partial x_j} \boldsymbol{G}_j(\boldsymbol{\rho}) = \boldsymbol{P}(\boldsymbol{\rho}) \tag{7}$$

where

$$\boldsymbol{G}_j(\boldsymbol{\rho}) = \langle F(\boldsymbol{\rho}, \boldsymbol{v}), v_j \boldsymbol{a}(\boldsymbol{v}) \rangle, \quad \boldsymbol{P}(\boldsymbol{\rho}) = \langle \mathcal{C}(F(\boldsymbol{\rho}, \boldsymbol{v})), \boldsymbol{a}(\boldsymbol{v}) \rangle$$

Note that the domain $\mathcal{U}$ of $\boldsymbol{G}_j$ and $\boldsymbol{P}$ is given by those moment vectors for which the solution $F(\boldsymbol{\rho}, \boldsymbol{v})$ of (5) exists.

Now, (7) is a system of evolution equations for the relevant physical quantities and it remains to check additional properties like the existence of an entropy inequality or Galilean invariance. To discuss these topics in more detail, we first list certain properties of the collision operator $\mathcal{C}$. Following [13], we assume that $\mathcal{C}(f)$ is defined for $f \in \mathcal{D}(\mathcal{C})$, where $\mathcal{D}(\mathcal{C})$ is a set of non-negative distribution functions. Furthermore, $\mathcal{C}$ acts only on the $\boldsymbol{v}$ dependence of f locally at each $(t, \boldsymbol{x})$.

The first assumption on $\mathcal{C}$ is related to dynamical conservation and can be formulated as equivalence of

$$\begin{aligned} &i) \quad \langle g, \mathcal{C}(f) \rangle = 0 \quad \text{for every } f \in \mathcal{D}(\mathcal{C}) \\ &ii) \quad g \in \operatorname{span}\{1, v_1, \ldots, v_d, |\boldsymbol{v}|^2\}. \end{aligned}$$

Hence, if we include the moments (3) in the vector $\boldsymbol{\rho}$, then the corresponding components of the right hand side $\boldsymbol{P}$ in (7) vanish, which means that mass, momentum, and energy are locally conserved quantities in the evolution.

Our second assumption is that $\mathcal{C}$ satisfies the dissipation relation

$$\langle \ln f, \mathcal{C}(f) \rangle \leq 0, \qquad \forall f \in \mathcal{D}(\mathcal{C}) \tag{8}$$

where equality is required to hold if and only if $\mathcal{C}(f) = 0$ which happens exactly for the Maxwellian densities

$$F(\varrho, \varrho \boldsymbol{u}, E, \boldsymbol{v}) = \frac{\varrho}{(2\pi\theta)^{\frac{d}{2}}} \exp\left(-\frac{|\boldsymbol{v} - \boldsymbol{u}|^2}{2\theta} \right) \tag{9}$$

where $\theta = (E/\varrho - |\boldsymbol{u}|^2/2)$ is the rescaled temperature of the gas. Multiplying (1) with $h'(f)$ where $h(f) = f \ln f - f$ and using (8) and mass conservation, we find that the entropy dissipates, i.e.

$$\frac{\partial h(f)}{\partial t} + \boldsymbol{v} \cdot \nabla h(f) = \langle \ln f, \mathcal{C}(f) \rangle \leq 0.$$

Integrating this equation over $\boldsymbol{v}$ and introducing $H(f) = \langle h(f), 1 \rangle$, we recover an entropy dissipation law (H-theorem). In [13], it is shown that this entropy dissipation law is still satisfied if we replace f by the maximum entropy distribution $F(\boldsymbol{\rho})$. Hence, the physical entropy $\eta(\boldsymbol{\rho}) = H(F(\boldsymbol{\rho}))$ is a locally strictly convex mathematical entropy for the system (7) which implies that (7) is symmetric hyperbolic.

Our last assumption on $\mathcal{C}$ concerns the behavior under translational and rotational transformations. Introducing for any vector $\boldsymbol{u} \in \mathbb{R}^d$ and any rotation matrix $R \in SO(d)$

$$(\mathcal{T}_{\boldsymbol{u}} f)(\boldsymbol{v}) = f(\boldsymbol{v} + \boldsymbol{u}), \qquad (\mathcal{T}_R f)(\boldsymbol{v}) = f(R\boldsymbol{v}), \qquad \boldsymbol{v} \in \mathbb{R}^d$$

the assumption on $\mathcal{C}$ reads

$$\mathcal{T}_{\boldsymbol{u}} \mathcal{C}(f) = \mathcal{C}(\mathcal{T}_{\boldsymbol{u}} f), \qquad \mathcal{T}_R \mathcal{C}(f) = \mathcal{C}(\mathcal{T}_R f). \tag{10}$$

Due to (10), the collision process is Galilean invariant and the Galilean invariance of the Boltzmann equation (1) follows, as we show now. First, we introduce the Galilean transformation Γ which associates physical quantities in a reference space-time coordinate system $(t, \boldsymbol{x})$ to a moving system $(t, \boldsymbol{y})$. If $\boldsymbol{u} \in \mathbb{R}^d$ is a constant vector and $R \in SO(d)$ a fixed rotation matrix, we have the following relations for space coordinates and velocity variables

$$\boldsymbol{y} = R\boldsymbol{x} + t\boldsymbol{u}, \qquad \boldsymbol{w} = R\boldsymbol{v} + \boldsymbol{u}.$$

We thus obtain for the distribution function Γf in the moving system

$$(\Gamma f)(t, \boldsymbol{y}, \boldsymbol{w}) = f(t, R^T(\boldsymbol{y} - t\boldsymbol{u}), R^T(\boldsymbol{w} - \boldsymbol{u})).$$

Assumption (10) together with the locality of $\mathcal{C}$ in $(t, \boldsymbol{x})$ implies $\Gamma \mathcal{C}(f) = \mathcal{C}(\Gamma f)$. Furthermore, it is a simple exercise in differentiation to show that

$$\Gamma \frac{\partial}{\partial x_j} = R_{kj} \frac{\partial}{\partial y_k} \Gamma, \qquad \Gamma \frac{\partial}{\partial t} = \frac{\partial}{\partial t} \Gamma + u_k \frac{\partial}{\partial y_k} \Gamma,$$

and obviously, $\Gamma v_j = R_{kj}(w_k - u_k)$. Thus,

$$\Gamma \left(\frac{\partial}{\partial t} + v_j \frac{\partial}{\partial x_j} - \mathcal{C} \right) = \left(\frac{\partial}{\partial t} + w_k \frac{\partial}{\partial y_k} - \mathcal{C} \right) \Gamma \tag{11}$$

which shows Galilean invariance of (1).

The same property can be obtained for the corresponding moment system (7) under suitable assumptions on the weight functions a_i. Here, we adopt the sufficient condition introduced in [13] which requires that the finite dimensional function space spanned by the weight functions $a_1, \ldots, a_n$

$$\mathbb{M} = \left\{ \sum_{i=1}^{n} \beta_i a_i : \beta_i \in \mathbb{R} \right\}$$

is invariant under rotations and translations, i.e.

$$\mathcal{T}_{\boldsymbol{u}} \mathbb{M} \subset \mathbb{M}, \qquad \mathcal{T}_R \mathbb{M} \subset \mathbb{M}, \qquad \text{if } \boldsymbol{u} \in \mathbb{R}^d,\ R \in SO(d). \tag{12}$$

In other words, $\mathcal{T}_{\boldsymbol{u}} a_i$ and $\mathcal{T}_R a_i$ are linear combinations of the functions a_i so that there are matrices $\Lambda(\boldsymbol{u}) \in \mathbb{R}^{n\times n}$ and $\Lambda(R) \in \mathbb{R}^{n\times n}$ such that

$$\mathcal{T}_{\boldsymbol{u}}\boldsymbol{a} = \Lambda(\boldsymbol{u})\boldsymbol{a}, \qquad \mathcal{T}_R\boldsymbol{a} = \Lambda(R)\boldsymbol{a}, \qquad \text{if } \boldsymbol{u} \in \mathbb{R}^d,\ R \in SO(d). \tag{13}$$

Note that $\Lambda(\boldsymbol{u}), \Lambda(R)$ are invertible with inverse $\Lambda(-\boldsymbol{u}), \Lambda(R^T)$. We prove

Proposition 1. *Assume $F(\boldsymbol{\rho}, \boldsymbol{v}) = \exp(\boldsymbol{\lambda}(\boldsymbol{\rho})\cdot \boldsymbol{a}(\boldsymbol{v}))$ satisfies $\langle F(\boldsymbol{\rho}), \boldsymbol{a}\rangle = \boldsymbol{\rho}$. If the weight functions satisfy* (13)*, then the moments $\Gamma\boldsymbol{\rho} = \langle \Gamma F(\boldsymbol{\rho}), \boldsymbol{a}\rangle$ in the moving system can be written as*

$$(\Gamma\boldsymbol{\rho})(t, \boldsymbol{y}) = \Lambda(R)\Lambda(\boldsymbol{u})\boldsymbol{\rho}(t, R^T(\boldsymbol{y} - t\boldsymbol{u})). \tag{14}$$

Moreover, we have the commutation relation

$$\Gamma F(\boldsymbol{\rho}) = F(\Gamma\boldsymbol{\rho}) \tag{15}$$

and the moment system (7) *is Galilean invariant, i.e.*

$$\frac{\partial \Gamma\boldsymbol{\rho}}{\partial t} + \frac{\partial}{\partial y_k}\boldsymbol{G}_k(\Gamma\boldsymbol{\rho}) = \boldsymbol{P}(\Gamma\boldsymbol{\rho}). \tag{16}$$

Proof. Since both translations and rotations have Jacobian determinant one, we find

$$\Gamma\boldsymbol{\rho} = \langle \Gamma F(\boldsymbol{\rho}), \boldsymbol{a}\rangle = \langle \mathcal{T}_R\mathcal{T}_{\boldsymbol{u}}\Gamma F(\boldsymbol{\rho}), \mathcal{T}_R\mathcal{T}_{\boldsymbol{u}}\boldsymbol{a}\rangle = \Lambda(R)\Lambda(\boldsymbol{u})\left\langle \mathcal{T}_R\mathcal{T}_{\boldsymbol{u}}\Gamma F(\boldsymbol{\rho}), \boldsymbol{a}\right\rangle.$$

Since

$$(\mathcal{T}_R\mathcal{T}_{\boldsymbol{u}}\Gamma F(\boldsymbol{\rho}))(t, \boldsymbol{y}, \boldsymbol{w}) = F(\boldsymbol{\rho}(t, R^T(\boldsymbol{y} - t\boldsymbol{u})), \boldsymbol{w})$$

and $\boldsymbol{\rho} = \langle F(\boldsymbol{\rho}), \boldsymbol{a}\rangle$, we arrive at (14). Using the structure of $F(\boldsymbol{\rho}, \boldsymbol{v}) = \exp(\boldsymbol{\lambda}(\boldsymbol{\rho})\cdot \boldsymbol{a}(\boldsymbol{v}))$, we obtain with

$$\boldsymbol{a}(R^T(\boldsymbol{w} - \boldsymbol{u})) = (\mathcal{T}_{R^T}\mathcal{T}_{-\boldsymbol{u}}\boldsymbol{a})(\boldsymbol{w}) = \Lambda(R^T)\Lambda(-\boldsymbol{u})\boldsymbol{a}(\boldsymbol{w})$$

that

$$\Gamma F(\boldsymbol{\rho}) = \exp\left([\Lambda(-\boldsymbol{u})^T\Lambda(R^T)^T\boldsymbol{\lambda}(\boldsymbol{\rho})]\cdot \boldsymbol{a}\right).$$

Since $\langle \Gamma F(\boldsymbol{\rho}), \boldsymbol{a}\rangle = \Gamma\boldsymbol{\rho}$, we conclude (using uniqueness of the representation (6))

$$\boldsymbol{\lambda}(\Gamma\boldsymbol{\rho}) = \Lambda(-\boldsymbol{u})^T\Lambda(R^T)^T\boldsymbol{\lambda}(\boldsymbol{\rho}).$$

Hence, $\Gamma F(\boldsymbol{\rho}) = \exp(\boldsymbol{\lambda}(\Gamma\boldsymbol{\rho})\cdot \boldsymbol{a}) = F(\Gamma\boldsymbol{\rho})$, which shows (15). Finally, Galilean invariance of the moment system is easily shown by writing (7) as

$$\left\langle \frac{\partial F(\boldsymbol{\rho})}{\partial t} + v_j\frac{\partial}{\partial x_j}F(\boldsymbol{\rho}) - \mathcal{C}(F(\boldsymbol{\rho})), \boldsymbol{a}\right\rangle = \mathbf{0}.$$

Changing the integration variable $\boldsymbol{v}$ to $\boldsymbol{w} = R^T(\boldsymbol{v} - \boldsymbol{u})$ and evaluating the expression at $\boldsymbol{x} = R^T(\boldsymbol{y} - t\boldsymbol{u})$, we find

$$\left\langle \Gamma\left(\frac{\partial}{\partial t} + v_j\frac{\partial}{\partial x_j} - \mathcal{C}\right)F(\boldsymbol{\rho}), \mathcal{T}_{R^T}\mathcal{T}_{-\boldsymbol{u}}\boldsymbol{a}\right\rangle = \mathbf{0}.$$

Using (11) and (15), we conclude

$$\Lambda(R^T)\Lambda(-\boldsymbol{u})\left\langle \left(\frac{\partial}{\partial t} + w_k\frac{\partial}{\partial y_k} - \mathcal{C}\right)F(\Gamma\boldsymbol{\rho}), \boldsymbol{a}\right\rangle = \mathbf{0}$$

which is exactly (16) since $\Lambda(R^T)\Lambda(-\boldsymbol{u})$ is an invertible matrix. $\square$

3 Polynomial weight functions

Examples of weight functions which exhibit the invariance property are given by suitably chosen families of polynomials. Since the moment system should typically contain the conservation equations of mass, momentum, and energy, we include $\{1, v_1, \dots, v_d, |\boldsymbol{v}|^2\}$ in the set of weight functions. Note that the span of these polynomials is indeed invariant under rotations and translations.

By taking additional weight functions, one can 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. Following [13], we mention as examples for the three-dimensional case

$$
\begin{array}{llllllll}
\text{10 moments:} & 1 & v_i & |\boldsymbol{v}|^2/2 & v_iv_j - |\boldsymbol{v}|^2\delta_{ij} & & \\
\text{14 moments:} & 1 & v_i & |\boldsymbol{v}|^2/2 & v_iv_j - |\boldsymbol{v}|^2\delta_{ij} & |\boldsymbol{v}|^2 v_i & |\boldsymbol{v}|^4 \\
\text{21 moments:} & 1 & v_i & |\boldsymbol{v}|^2/2 & v_iv_j - |\boldsymbol{v}|^2\delta_{ij} & v_iv_jv_k & |\boldsymbol{v}|^4 \\
\text{26 moments:} & 1 & v_i & |\boldsymbol{v}|^2/2 & v_iv_j - |\boldsymbol{v}|^2\delta_{ij} & v_iv_jv_k & |\boldsymbol{v}|^2 v_iv_j \\
\text{35 moments:} & 1 & v_i & |\boldsymbol{v}|^2/2 & v_iv_j - |\boldsymbol{v}|^2\delta_{ij} & v_iv_jv_k & v_iv_jv_kv_l
\end{array}
\tag{17}
$$

Two aspects are important here: First of all, the generated function spaces are rotational and translational invariant. Furthermore, the highest degrees are always taken even to ensure that $\exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}(\boldsymbol{v}))$ is integrable for a reasonable range of parameters $\boldsymbol{\lambda}$.

At first glance, the associated moment systems (7) is very promising: it generalizes the Euler equations of gas dynamics, it is strictly hyperbolic, Galilean invariant, and possesses an entropy inequality. However, a closer look reveals a fundamental problem if weight functions of fourth (or higher) order are used. To explain this problem, let us consider, for example, the 14-moment case. As always, the maximum entropy distribution function has the structure $F(\boldsymbol{\rho}, \boldsymbol{v}) = \exp(\boldsymbol{\lambda}(\boldsymbol{\rho}) \cdot \boldsymbol{a}(\boldsymbol{v}))$, where in the 14-moment case (using Einstein's summation convention)

$$
\exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}(\boldsymbol{v})) = \exp(\lambda + \lambda_i v_i + \lambda_{ij} v_i v_j + \bar{\lambda}_i |\boldsymbol{v}|^2 v_i + \lambda_{14} |\boldsymbol{v}|^4) \tag{18}
$$

is the set of all admissible Lagrange multipliers. Obviously, this distribution function is only integrable if either $\lambda_{14} < 0$, or if $\lambda_{14} = 0, \bar{\lambda}_i = 0$ and (λ_{ij}) is strictly negative definite. In the following, we denote the set of $\boldsymbol{\lambda}$ for which $\exp(\boldsymbol{\lambda} \cdot \boldsymbol{a})$ is integrable by L, i.e.

$$
L = \{\boldsymbol{\lambda} \in \mathbb{R}^n : \langle \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}), |\boldsymbol{a}| \rangle < \infty\}.
$$

The physical equilibrium states of the gas are given by Maxwellian states (9) which are of the form (18) with

$$
\lambda^* = \ln \frac{\varrho}{(2\pi\theta)^{3/2}} - \frac{|\boldsymbol{u}|^2}{2\theta}, \qquad \lambda_i^* = \frac{u_i}{\theta}, \quad i = 1, 2, 3, \qquad \lambda_{ij}^* = -\frac{1}{2\theta}\delta_{ij}
$$

and $\bar{\lambda}_i^* = \lambda_{14}^* = 0$. These states belong to ∂L because any small deviation $\boldsymbol{\lambda}$ from $\boldsymbol{\lambda}^*$ with $\lambda_{14} > 0$ does not belong to L. As a consequence, the equilibrium moments $\boldsymbol{\rho}^*$ from the set

$$\mathcal{U}^* = \left\{ \left\langle \varrho(2\pi\theta)^{-d/2} \exp(-|\boldsymbol{v} - \boldsymbol{u}|^2/2), \boldsymbol{a} \right\rangle : \varrho, T > 0, \boldsymbol{u} \in \mathbb{R}^d \right\}$$

are also on the boundary of the state space $\mathcal{U}$

$$\mathcal{U} = \{ \langle \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}), \boldsymbol{a} \rangle : \boldsymbol{\lambda} \in L \}$$

which is the domain of definition of equation (7). In most practical applications, there are regions in the physical space where states are close to equilibrium. Hence, $\boldsymbol{\rho}(t, \boldsymbol{x})$ touches or at least comes very close to ∂U for certain t and $\boldsymbol{x}$. As example, let us consider a typical Riemann initial value, consisting of two separated equilibrium states. Since 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 [17], little can be said about solvability. Also, the local existence result for smooth solutions [14] requires that the range of the initial values is contained in a compact subset of the *interior* of the state space [14]. Hence, initial values obtained by adding smooth and compactly supported perturbances to an equilibrium state are *not* covered by the local existence result if equilibrium states are located on $\partial\mathcal{U}$.

It turns out that equilibrium points at ∂U is not a mere technicality. For a model problem ($d = 1$ and $\boldsymbol{a}(v) = (1, v, v^2, v^3, v^4)$), it can be shown that $\mathcal{U}$ is not convex and that the flux function $\boldsymbol{G}$ in (7) 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 [8, 10, 3, 11].

We now show that equilibrium states on the boundary $\partial\mathcal{U}$ always appear when the set of weight functions contains, apart from $1, v_i, |\boldsymbol{v}|^2$, other functions with super-quadratic growth at infinity. First, we need a general result on the solvability of moment problems for which we refer to [12, 9, 11]. Introducing the set of moments of non-negative distribution functions

$$\mathcal{M} = \{ \langle f, \boldsymbol{a} \rangle : f : \mathbb{R}^d \to \mathbb{R}^+ \quad \text{measurable} \}$$

we have:

Theorem 2. *For some $n \in \mathbb{N}$ let $a_1, \ldots, a_n$ be a set of linearly independent polynomials on $\mathbb{R}^d$ and assume $\boldsymbol{\rho} \in \mathbb{R}^n$. Then $\boldsymbol{\rho} \in \mathcal{M}$ if and only if $\boldsymbol{\beta} \cdot \boldsymbol{\rho} < 0$ for all non-zero vectors $\boldsymbol{\beta} \in \mathbb{R}^n$ which satisfy $\boldsymbol{\beta} \cdot \boldsymbol{a}(\boldsymbol{v}) \leq 0$ for all $\boldsymbol{v} \in \mathbb{R}^d$.*

Using this theorem, we can prove the following result.

Proposition 3. *Let $n \in \mathbb{N}$ and $\mathcal{P} = \{a_1, \ldots, a_n\}$ be a set of linearly independent polynomials including $\{1, v_1, \ldots, v_d, |\boldsymbol{v}|^2\}$. Assume that the polynomials $a_m, , \ldots, a_n$ are homogeneous and of maximal degree in $\mathcal{P}$ which is larger than two and even. Then every neighborhood of an equilibrium state $\boldsymbol{\rho}^*$ contains moment vectors $\boldsymbol{\rho} \in \mathcal{M}$ which are not admissible states, i.e. $\boldsymbol{\rho} \notin \mathcal{U}$.*

Proof. For some given $\boldsymbol{\rho}^* \in \mathcal{U}^*$, we first construct $\bar{\boldsymbol{\rho}} \in \mathcal{M}$ which is the moment vector of some non-negative density f. Then, we show $\bar{\boldsymbol{\rho}} \notin \mathcal{U}$ although $|\boldsymbol{\rho}^* - \bar{\boldsymbol{\rho}}|$ is as small as we want. For $\epsilon > 0$ and $\bar{\boldsymbol{v}} \in \mathbb{R}^d$ with $|\bar{\boldsymbol{v}}| = 1$, we set

$$\bar{\boldsymbol{\rho}} = \boldsymbol{\rho}^* + \epsilon(0, \dots, 0, a_m(\bar{\boldsymbol{v}}), \dots, a_n(\bar{\boldsymbol{v}})).$$

To see that $\bar{\boldsymbol{\rho}}$ is the moment vector of some non-negative density, we use Theorem 2. If $\mathbf{0} \neq \boldsymbol{\beta} \in \mathbb{R}^n$ with $\boldsymbol{\beta} \cdot \boldsymbol{a}(\bar{\boldsymbol{v}}) \leq 0$ for all $\boldsymbol{v} \in \mathbb{R}^d$ (such a vector exists because $\mathcal{P}$ contains the constant one function) and if $2l$ is the highest degree appearing in $\mathcal{P}$, we have

$$0 \geq \lim_{r \to \infty} \frac{\boldsymbol{\beta} \cdot \boldsymbol{a}(r\bar{\boldsymbol{v}})}{r^{2l}} = \sum_{i=m}^{n} \beta_i a_i(\bar{\boldsymbol{v}}),$$

where we have used the homogeneity of $a_m, \dots, a_n$ and the fact that $a_1, \dots, a_{m-1}$ are of degree $< 2l$. Hence

$$\boldsymbol{\beta} \cdot \bar{\boldsymbol{\rho}} = \boldsymbol{\beta} \cdot \boldsymbol{\rho}^* + \sum_{i=m}^{n} \beta_i a_i(\bar{\boldsymbol{v}}) \leq \boldsymbol{\beta} \cdot \boldsymbol{\rho}^* < 0$$

so that $\bar{\boldsymbol{\rho}} \in \mathcal{M}$ according to Theorem 2.

We prove $\bar{\boldsymbol{\rho}} \notin \mathcal{U}$ in an indirect way. Assume $\bar{\boldsymbol{\rho}} \in \mathcal{U}$. Then there exists $\bar{\boldsymbol{\lambda}} \in L$ such that $\langle \exp(\bar{\boldsymbol{\lambda}} \cdot \boldsymbol{a}), \boldsymbol{a} \rangle = \bar{\boldsymbol{\rho}}$. Similarly, $\boldsymbol{\rho}^* = \langle \exp(\boldsymbol{\lambda}^* \cdot \boldsymbol{a}), \boldsymbol{a} \rangle$ for some $\boldsymbol{\lambda}^* \in L$ with $\lambda_m^* = \dots = \lambda_n^* = 0$ (because $a_m, \dots, a_n$ are of degree $2l > 2$ and thus do not appear in equilibrium distributions). Defining

$$\boldsymbol{\lambda}(s) = \boldsymbol{\lambda}^* + s(\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^*), \qquad s \in [0, 1]$$

and using the convexity of the exponential function, we find an s-independent bound $\exp(\boldsymbol{\lambda}(s) \cdot \boldsymbol{a}) \leq \exp(\boldsymbol{\lambda}^* \cdot \boldsymbol{a}) + \exp(\bar{\boldsymbol{\lambda}} \cdot \boldsymbol{a})$ which suffices to show that

$$g(s) = \langle \exp(\boldsymbol{\lambda}(s) \cdot \boldsymbol{a}), 1 \rangle - \boldsymbol{\lambda}(s) \cdot \bar{\boldsymbol{\rho}}, \qquad s \in [0, 1]$$

is smooth and strictly convex. Since

$$g'(s) = \langle \exp(\boldsymbol{\lambda}(s) \cdot \boldsymbol{a}), \boldsymbol{a} \rangle \cdot (\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^*) - (\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^*) \cdot \bar{\boldsymbol{\rho}}$$

we conclude $g'(1) = 0$ so that $g'(0) < 0$ because $g''(s) > 0$ in $[0, 1]$. Hence,

$$\begin{aligned} 0 > g'(0) &= (\langle \exp(\boldsymbol{\lambda}^* \cdot \boldsymbol{a}), \boldsymbol{a} \rangle - \bar{\boldsymbol{\rho}}) \cdot (\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^*) \\ &= (\boldsymbol{\rho}^* - \bar{\boldsymbol{\rho}}) \cdot (\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^*) = -\epsilon \sum_{i=m}^{n} \bar{\lambda}_i a_i(\bar{\boldsymbol{v}}) \end{aligned}$$

where we have used the definition of $\bar{\boldsymbol{\rho}}$ and the fact that $\lambda_i^* = 0$ for $i = m, \dots, n$. Since a_i are continuous functions, we obtain for some suitable $\delta > 0$

$$\sum_{i=m}^{n} \bar{\lambda}_i a_i(\bar{\boldsymbol{v}}) \geq \mu > 0, \qquad \text{for } |\boldsymbol{v} - \bar{\boldsymbol{v}}| < \delta$$

so that

$$\lim_{r\to\infty}\frac{\bar{\boldsymbol{\lambda}}\cdot\boldsymbol{a}(r\boldsymbol{v})}{r^{2l}}\geq\mu>0,\qquad\text{for }|\boldsymbol{v}-\bar{\boldsymbol{v}}|<\delta.$$

In particular, the integral over $\exp(\bar{\boldsymbol{\lambda}}\cdot\boldsymbol{a})$ is infinite because the exponential exceeds one on a domain $\{r\boldsymbol{v} : r>R, |\boldsymbol{v}-\bar{\boldsymbol{v}}|<\delta\}$ which has infinite measure. Since we have assumed $\bar{\boldsymbol{\lambda}}\in L$, i.e. $\langle\exp(\bar{\boldsymbol{\lambda}}\cdot\boldsymbol{a}),1\rangle<\infty$, this is a contradiction. We conclude with the remark that one can generalize the definition of $\bar{\boldsymbol{\rho}}$ in to

$$\bar{\boldsymbol{\rho}}=\boldsymbol{\rho}^*+\epsilon\sum_{i=1}^{s}\sigma_i(0,\ldots,0,a_m,(\bar{\boldsymbol{v}}_i),\ldots,a_n(\bar{\boldsymbol{v}}_i))$$

where $|\boldsymbol{v}_i|=1$ and $\sigma_i\geq 0$, $\sigma_1>0$, without changing much of the proof. □

The proof clearly shows that equilibrium points are on the boundary because weight functions are present which dominate $|\boldsymbol{v}|^2$ for large $|\boldsymbol{v}|$. Hence, to avoid the disadvantage $\mathcal{U}^*\subset\partial\mathcal{U}$, we should choose weight functions which satisfy

$$\lim_{|\boldsymbol{v}|\to\infty}\frac{|a_i(\boldsymbol{v})|}{1+|\boldsymbol{v}|^2}=0,\qquad i=2,\ldots,n \tag{19}$$

where we assume $a_1(\boldsymbol{v})=|\boldsymbol{v}|^2$. Under this condition, the state space $\mathcal{U}$ coincides with the open convex set $\mathcal{M}$ of moments belonging to non-negative distribution functions [9]. In particular, equilibrium points are then always interior points which allows the application of existence results and rules out the case that the flux functions $\boldsymbol{G}_j$ become singular in equilibrium states.

In order to keep the advantage of the polynomial weight functions (that the associated moments have a physical interpretation as averaged fluxes), one could, for example, resort to suitable truncations: if $\{a_1,\ldots,a_{m-1}\}=\{1,v_1,\ldots,v_d,|\boldsymbol{v}|^2\}$ and $a_m,\ldots,a_n$ are polynomial weight functions of degree higher than two, we define

$$\tilde{a}_i(\boldsymbol{v})=a_i(\boldsymbol{v})\quad\text{for }i=1,\ldots,m-1,\qquad\tilde{a}_i(\boldsymbol{v})=a_i(\boldsymbol{V}(\boldsymbol{v}))\quad\text{for }i=m,\ldots,n$$

where

$$V_k(\boldsymbol{v})=\begin{cases}v_k & |v_k|\leq v_{max}\\ v_{max} & \text{else}\end{cases},\qquad k=1,\ldots,d.$$

If v_{max} is chosen large enough, the new weight functions $\tilde{\boldsymbol{a}}$ obviously coincide with the polynomials $\boldsymbol{a}$ for all "reasonable" velocities $\boldsymbol{v}\in\mathbb{R}^d$. However, the new system $\{\tilde{a}_1,\ldots,\tilde{a}_n\}$ is clearly not translation invariant so that, while gaining a better structure of the state space, we destroy Galilean invariance of the moment system. Intuitively, it is clear that any modification of the polynomial weight functions at high velocities (which is necessary to ensure (19)) includes a possible conflict with Galilean invariance because large speeds v_{max} in one coordinate system can be small in a frame of reference which itself moves with the speed v_{max}.

Therefore, we now pose the following question: Is it possible to find families of weight functions which include $a_1(\boldsymbol{v}) = |\boldsymbol{v}|^2$ as well as $\{1, v_1, \dots, v_d\}$ and give rise to rotation and translation invariant function spaces while satisfying condition (19)?

Unfortunately, the answer is *no*, as we prove in the next section: under suitable assumptions, a finite dimensional function space is rotation and translation invariant if and only if it is a rotation and translation invariant space of polynomials – which contradicts (19), once we want to take more weight functions than $1, v_i, v_i v_j$.

4 Invariant function spaces

In the following, $\mathcal{F}(A, B)$ denotes the set of functions mapping the set A into the set B. We recall that translations and rotations are introduced as linear operators on $\mathcal{F}(\mathbb{R}^d, \mathbb{R})$ according to

$$(\mathcal{T}_{\boldsymbol{x}}\varphi)(\boldsymbol{z}) = \varphi(\boldsymbol{z} + \boldsymbol{x}), \qquad (\mathcal{T}_R\varphi)(\boldsymbol{z}) = \varphi(R\boldsymbol{z}), \qquad \forall \boldsymbol{z} \in \mathbb{R}^d$$

where $\boldsymbol{x} \in \mathbb{R}^d$ and $R \in SO(d)$. We put

$$\mathcal{T}_* = \{\mathcal{T}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^d\} \cup \{\mathcal{T}_R : R \in SO(d)\}.$$

Using this notation, our aim is to characterize finite dimensional subspaces $\mathbb{M} \subset \mathcal{F}(\mathbb{R}^d, \mathbb{R})$ which are invariant under the family $\mathcal{T}_*$. We start with some preliminary considerations.

4.1 Generalities

If $\mathbb{M} \subset \mathcal{F}(\mathbb{R}^d, \mathbb{R})$ is finite dimensional, we can fix a basis

$$a_1, \dots, a_n \in \mathbb{M},$$

where $n \in \mathbb{N}$. In the sequel we set for $\boldsymbol{z} \in \mathbb{R}^d$,

$$\boldsymbol{a}(\boldsymbol{z}) := (a_1(\boldsymbol{z}), \dots, a_n(\boldsymbol{z}))^T.$$

Proposition 4. *There are $\boldsymbol{z}_1, \dots, \boldsymbol{z}_n \in \mathbb{R}^d$ such that $\boldsymbol{a}(\boldsymbol{z}_1), \dots, \boldsymbol{a}(\boldsymbol{z}_n)$ are linearly independent.*

Proof. It suffices to prove: $U := \mathsf{span}(\{\boldsymbol{a}(\boldsymbol{z}) : \boldsymbol{z} \in \mathbb{R}^d\}) = \mathbb{R}^n$. If $\mathsf{dim}(U) < n$, then there are $\eta_1, \dots, \eta_n \in \mathbb{R}$ such that $\eta_1 y_1 + \dots + \eta_n y_n = 0$ for all $(y_1, \dots, y_n) \in U$. As a consequence, $\eta_1 a_1(\boldsymbol{z}) + \dots + \eta_n a_n(\boldsymbol{z}) = 0$ for all $\boldsymbol{z} \in \mathbb{R}^d$, i.e. $a_1, \dots, a_n$ are not linearly independent in $\mathcal{F}(\mathbb{R}^d, \mathbb{R})$. $\square$

Proposition 5. *Let $\Omega, \Omega_\circ \in \mathbb{R}^{n \times n}$ and let $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ be a basis of $\mathbb{R}^n$. Assume*

$$\forall j \in \{1, \dots, n\} : \quad \Omega \cdot \mathbf{w}_j = \Omega_\circ \cdot \mathbf{w}_j.$$

Then $\Omega = \Omega_\circ$.

Proof. Evident. ☐

Corollary 1. *Let* $\Omega, \Omega_\circ \in \mathbb{R}^{n\times n}$*. Assume*

$$\forall \boldsymbol{z} \in \mathbb{R}^d: \quad \Omega \cdot \boldsymbol{a}(\boldsymbol{z}) = \Omega_\circ \cdot \boldsymbol{a}(\boldsymbol{z}).$$

Then $\Omega = \Omega_\circ$*.*

Proof. Take $\boldsymbol{z} = \boldsymbol{z}_j$, $j = 1, \dots, n$ as in Proposition 4 and apply Proposition 5. ☐

Proposition 6. *Let* $\Omega \in \mathcal{F}(\mathbb{R}, \mathbb{R}^{n\times n})$*. Let* $\boldsymbol{\psi} \in \mathcal{F}(\mathbb{R}^d, \mathbb{R}^n)$*. Assume*

$$\forall \boldsymbol{z} \in \mathbb{R}^d: \quad \lim_{t\to 0} \Omega(t) \cdot \boldsymbol{a}(\boldsymbol{z}) = \boldsymbol{\psi}(\boldsymbol{z}).$$

Then there is a matrix $\Omega_\circ \in \mathbb{R}^{n\times n}$ *such that* $\lim_{t\to 0} \Omega(t) = \Omega_\circ$*.*

Proof. We introduce for $t \in \mathbb{R}$ the matrix

$$\Omega_1(t) := \Omega(t) \cdot \begin{pmatrix} a_1(\boldsymbol{z}_1) & \dots & a_1(\boldsymbol{z}_n) \\ \vdots & \vdots & \vdots \\ a_n(\boldsymbol{z}_1) & \dots & a_n(\boldsymbol{z}_n) \end{pmatrix} =: \Omega(t) \cdot \mathbf{V}.$$

Due to assumption we have

$$\lim_{t\to 0} \Omega_1(t) = \begin{pmatrix} \psi_1(\boldsymbol{z}_1) & \dots & \psi_1(\boldsymbol{z}_n) \\ \vdots & \vdots & \vdots \\ \psi_n(\boldsymbol{z}_1) & \dots & \psi_n(\boldsymbol{z}_n) \end{pmatrix} =: \mathbf{U}.$$

Hence

$$\lim_{t\to 0} \Omega(t) = \lim_{t\to 0} \Omega_1(t) \cdot \mathbf{V}^{-1} = \mathbf{U} \cdot \mathbf{V}^{-1}.$$

☐

4.2 Translation Invariance

Now let $\mathbb{M}$ be an n-dimensional subspace of $\mathcal{F}(\mathbb{R}^d, \mathbb{R})$ which is invariant under translations $\mathcal{T}_{\boldsymbol{x}}$, $\boldsymbol{x} \in \mathbb{R}^d$. Then, we can find for each $j \in \{1, \dots, n\}$ functions $\lambda_{j,1}, \dots, \lambda_{j,n} : \mathbb{R}^d \to \mathbb{R}$ such that

$$\mathcal{T}_{\boldsymbol{x}} a_j = \lambda_{j,1}(\boldsymbol{x}) a_1 + \ldots + \lambda_{j,n}(\boldsymbol{x}) a_n.$$

We introduce the matrix-valued function

$$\Lambda : \mathbb{R}^d \to \mathbb{R}^{n\times n}, \quad \Lambda(\boldsymbol{x}) = \left(\lambda_{j,l}(\boldsymbol{x})\right)_{j,l=1,\dots,n}.$$

Then we have for all $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^d$,

$$\boldsymbol{a}(\boldsymbol{x}+\boldsymbol{z}) = \begin{pmatrix} a_1(\boldsymbol{x}+\boldsymbol{z}) \\ \vdots \\ a_n(\boldsymbol{x}+\boldsymbol{z}) \end{pmatrix} = \Lambda(\boldsymbol{x}) \cdot \begin{pmatrix} a_1(\boldsymbol{z}) \\ \vdots \\ a_n(\boldsymbol{z}) \end{pmatrix} = \Lambda(\boldsymbol{x}) \cdot \boldsymbol{a}(\boldsymbol{z}). \qquad (20)$$

In particular, if we set $\boldsymbol{x} = \mathbf{0}$, then

$$\boldsymbol{a}(\boldsymbol{z}) = \Lambda(\mathbf{0}) \cdot \boldsymbol{a}(\boldsymbol{z}),$$

hence, since $a_1, \dots, a_n$ is a basis of $\mathbb{M}$, $\Lambda(\mathbf{0}) = \mathsf{id}_{n\times n}$, which is the $n\times n$-identity matrix. Furthermore, if we set $\boldsymbol{z} = \mathbf{0}$, then

$$\boldsymbol{a}(\boldsymbol{x}) = \Lambda(\boldsymbol{x}) \cdot \boldsymbol{a}_\circ, \tag{21}$$

where $\boldsymbol{a}_\circ := \boldsymbol{a}(\mathbf{0})$. As a consequence of (20) we have for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^d$,

$$\begin{aligned}(\Lambda(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{y})) \cdot \boldsymbol{a}(\boldsymbol{z}) &= \Lambda(\boldsymbol{x}) \cdot (\Lambda(\boldsymbol{y}) \cdot \boldsymbol{a}(\boldsymbol{z})) = \Lambda(\boldsymbol{x}) \cdot \boldsymbol{a}(\boldsymbol{y}+\boldsymbol{z}) \\ &= \boldsymbol{a}(\boldsymbol{x} + (\boldsymbol{y}+\boldsymbol{z})) = \boldsymbol{a}((\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}) = \Lambda(\boldsymbol{x}+\boldsymbol{y}) \cdot \boldsymbol{a}(\boldsymbol{z}),\end{aligned}$$

hence by Corollary 1

$$\Lambda(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{y}) = \Lambda(\boldsymbol{x}+\boldsymbol{y}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d.$$

In a similar way we deduce

$$\Lambda(\boldsymbol{y}) \cdot \Lambda(\boldsymbol{x}) = \Lambda(\boldsymbol{y}+\boldsymbol{x}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d,$$

and therefore due to $\boldsymbol{x}+\boldsymbol{y} = \boldsymbol{y}+\boldsymbol{x}$,

$$\Lambda(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{y}) = \Lambda(\boldsymbol{y}) \cdot \Lambda(\boldsymbol{x}) = \Lambda(\boldsymbol{x}+\boldsymbol{y}). \tag{22}$$

Now we introduce for $k \in \{1, \dots, d\}$ the k-th unit vector $\boldsymbol{e}_k$ of $\mathbb{R}^d$ the one-parameter family

$$\Lambda_k : \mathbb{R} \to \mathbb{R}^{n\times n}, \quad \Lambda_k(t) = \Lambda(t\boldsymbol{e}_k),$$

of matrices. Due to (20) we have for all $k \in \{1, \dots, d\}$, for all $t \in \mathbb{R}$, and for all $\boldsymbol{z} \in \mathbb{R}^d$,

$$\Lambda_k(t) \cdot \boldsymbol{a}(\boldsymbol{z}) - \boldsymbol{a}(\boldsymbol{z}) = \boldsymbol{a}(\boldsymbol{z} + t\boldsymbol{e}_k) - \boldsymbol{a}(\boldsymbol{z}),$$

such that

$$\lim_{t\to 0} \frac{\Lambda_k(t) \cdot \boldsymbol{a}(\boldsymbol{z}) - \boldsymbol{a}(\boldsymbol{z})}{t} = (\partial_k \boldsymbol{a})(\boldsymbol{z}) \tag{23}$$

if $\boldsymbol{a}$ possesses first partial derivatives in $\boldsymbol{z}$. Note that, because of translation invariance, this differentiability property follows already if we assume it in a single point.

Lemma 7. *Let $\mathbb{M} = \mathrm{span}\{a_1, \dots, a_n\}$ be translation invariant and assume that there exists $\bar{\boldsymbol{x}} \in \mathbb{R}^d$ where the functions $a_1, \dots, a_n$ possess all partial derivatives $(\partial_k a_i)(\bar{\boldsymbol{x}})$, $i = 1, \dots, n$, $k = 1, \dots, d$. Then partial derivatives of a_i exist in every point $\boldsymbol{z} \in \mathbb{R}^d$.*

Proof. Writing

$$\begin{aligned}\frac{a_i(\boldsymbol{z}+t\boldsymbol{e}_k)-a_i(\boldsymbol{z})}{t} &= \frac{a_i(\bar{\boldsymbol{x}}+t\boldsymbol{e}_k+(\boldsymbol{z}-\bar{\boldsymbol{x}}))-a_i(\bar{\boldsymbol{x}}+(\boldsymbol{z}-\bar{\boldsymbol{x}}))}{t} \\ &= \Lambda(\boldsymbol{z}-\bar{\boldsymbol{x}})\cdot\frac{a_i(\bar{\boldsymbol{x}}+t\boldsymbol{e}_k)-a_i(\bar{\boldsymbol{x}})}{t}\end{aligned}$$

we conclude that $(\partial_k a_i)(\boldsymbol{z}) = \Lambda(\boldsymbol{z}-\bar{\boldsymbol{x}})(\partial_k a_i)(\bar{\boldsymbol{x}})$. ☐

Hence, if $\boldsymbol{a}$ has all partial derivatives in a single point, then (23) holds and with Proposition 6, we can deduce: There is for each $k \in \{1,\dots,d\}$ a matrix $\Omega_k \in \mathbb{R}^{n\times n}$ such that

$$\lim_{t\to 0}\frac{\Lambda_k(t)-\mathsf{id}_{n\times n}}{t} = \Omega_k. \tag{24}$$

As a consequence, Λ_k is differentiable at $t=0$ and $\Omega_k = \Lambda_k'(0)$, $k\in\{1,\dots,d\}$.

Due to (22) we have for all $s,t\in\mathbb{R}$ and for all $k\in\{1,\dots,d\}$ the semigroup property

$$\Lambda_k(s+t) = \Lambda_k(s)\cdot\Lambda_k(t) = \Lambda_k(t)\cdot\Lambda_k(s),$$

and - naturally - $\Lambda_k(0) = \mathsf{id}_{n\times n}$. We deduce for all $t\in\mathbb{R}$, for all $\boldsymbol{\psi}\in\mathbb{R}^n$ and for all $k\in\{1,\dots,d\}$,

$$\begin{aligned}\lim_{s\to 0}\frac{\Lambda_k(t+s)\cdot\boldsymbol{\psi}-\Lambda_k(t)\cdot\boldsymbol{\psi}}{s} &= \lim_{s\to 0}\frac{\Lambda_k(s)-\mathsf{id}_{n\times n}}{s}\cdot(\Lambda_k(t)\cdot\boldsymbol{\psi}) \\ &= \Omega_k\cdot(\Lambda(t)\cdot\boldsymbol{\psi}),\end{aligned}$$

i.e. if we consider the mapping $\boldsymbol{\phi}(.) : \mathbb{R}\to\mathbb{R}^n, \boldsymbol{\phi}(t) = \Lambda(t)\cdot\boldsymbol{\psi}$, then $\boldsymbol{\psi}(.)$ is differentiable (thus, continuous) and satisfies for each $t\in\mathbb{R}$ the ODE

$$\boldsymbol{\phi}'(t) = \Omega_k\cdot\boldsymbol{\phi}(t),$$

subject to the initial condition

$$\boldsymbol{\phi}(0) = \boldsymbol{\psi}.$$

As a consequence, we have for each $t\in\mathbb{R}$ and for each $k\in\{1,\dots,d\}$,

$$\boldsymbol{\phi}(t) = \exp(t\Omega_k)\cdot\boldsymbol{\psi}, \tag{25}$$

and therefore due to Proposition 5,

$$\Lambda_k(t) = \exp(t\Omega_k), \quad k\in\{1,\dots,d\}. \tag{26}$$

Due to (22) we have for all $k,l\in\{1,\dots,d\}$ the commutator relation

$$[\Lambda_k,\Lambda_l] = \Lambda_k\cdot\Lambda_l - \Lambda_l\cdot\Lambda_k = \mathbf{0}_{n\times n}.$$

Hence, with (24),

$$[\Omega_k,\Omega_l] = \mathbf{0}_{n\times n}, \quad k,l\in\{1,\dots,d\}, \tag{27}$$

and therefore for all $s, t \in \mathbb{R}$ and for all $k, l \in \{1, \ldots, d\}$,

$$\exp(t\Omega_k + s\Omega_l) = \exp(t\Omega_k) \cdot \exp(s\Omega_l) = \exp(s\Omega_l) \cdot \exp(t\Omega_k).$$

Thus, we have for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$\Lambda(\boldsymbol{x}) = \exp\left(\sum_{k=1}^{d} x_k \Omega_k\right), \quad [\Omega_k, \Omega_l] = \mathbf{0}_{n\times n}, \quad k, l \in \{1, \ldots, d\}, \tag{28}$$

and, due to (21),

$$\boldsymbol{a}(\boldsymbol{x}) = \Lambda(\boldsymbol{x}) \cdot \boldsymbol{a}_\circ = \exp\left(\sum_{k=1}^{d} x_k \Omega_k\right) \cdot \boldsymbol{a}_\circ, \tag{29}$$

respectively

$$\boldsymbol{a}(\boldsymbol{x}) = \exp(x_1\Omega_1) \cdot \ldots \cdot \exp(x_d\Omega_d) \cdot \boldsymbol{a}_\circ. \tag{30}$$

Note that relation (30) drastically restricts the structure of the basis $\{a_1, \ldots, a_n\}$ of the translation invariant subspace $\mathbb{M}$ because the entries of an exponential matrix $\exp(x_k\Omega_k)$ are combinations of polynomials, exponential functions and sine and cosine functions. More precisely, if the spectrum of Ω_k is given by

$$\sigma(\Omega_k) = \{\xi_1^k, \ldots, \xi_{r(k)}^k\}, \qquad 1 \leq r(k) \leq n$$

where ξ_w^k are the (possibly complex) eigenvalues of Ω_k with respective multiplicities $\mu_1^k, \ldots, \mu_{r(k)}^k$, then each entry of the matrix $\exp(x_k\Omega_k)$ is the sum of terms of the form $p_w(x_k)\exp(x_k\xi_w^k)$, $w = 1, \ldots, r(k)$, where $p_w(.)$ is a (complex) polynomial of degree less or equal $\mu_w^k - 1$. Carrying out the matrix multiplications of (30) we obtain for all $j = 1, \ldots, n$ and for all $\boldsymbol{x} \in \mathbb{R}^d$,

$$\begin{aligned} a_j(\boldsymbol{x}) \in \operatorname{span}_{\mathbb{C}} \Big\{ &\boldsymbol{x}^{\boldsymbol{\alpha}} \exp(\boldsymbol{x} \cdot \boldsymbol{\xi}) : \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d, \\ &\alpha_1 + \ldots + \alpha_d \leq nd, \boldsymbol{\xi} = (\xi_1, \ldots, \xi_d) \in \sigma(\Omega_1) \times \ldots \times \sigma(\Omega_d) \Big\}, \end{aligned} \tag{31}$$

where "$\operatorname{span}_{\mathbb{C}}\{\ldots\}$" is the set of all complex linear combinations of elements of $\{\ldots\}$. Obviously, real eigenvalues ξ_w^k give rise to exponential behavior in direction x_k (unless $\xi_w^k = 0$ which leads to a purely polynomial behavior). If an eigenvalue ξ_w^k is not real, the complex conjugate is also an eigenvalue (because Ω_k is real) so that complex ξ_w^k lead to combinations of sine and cosine functions.

We summarize our results in a final

Theorem 8. *Let $\mathbb{M}$ be a translation invariant, n-dimensional subspace of the space $\mathcal{F}(\mathbb{R}^d, \mathbb{R})$ with the property that the functions in $\mathbb{M}$ possess all partial derivatives in some point $\bar{\boldsymbol{x}} \in \mathbb{R}^d$. Then, all functions in $\mathbb{M}$ are analytic (combinations of polynomials, exponential functions and sine and cosine functions). More precisely, if $a_1, \ldots, a_n$ is any basis of $\mathbb{M}$, the vector $\boldsymbol{a} = (a_1, \ldots, a_n)^T$ can be written as*

$$\boldsymbol{a}(\boldsymbol{x}) = \exp\left(\sum_{k=1}^{d} x_k \Omega_k\right) \cdot \boldsymbol{a}(\mathbf{0}), \qquad \forall \boldsymbol{x} \in \mathbb{R}^d$$

for certain commuting matrices $\Omega_k \in \mathbb{R}^{n\times n}$, $k = 1, \ldots, d$.

4.3 Translation Invariance and Rotation Invariance

While translation invariance already puts a strong constraint on the subspace $\mathbb{M}$, we will now see that the additional assumption of rotation invariance further restricts the structure: the possibility that $\mathbb{M}$ contains exponential or sine and cosine functions is ruled out by rotation invariance.

We start with the observation that (using (31)): For all $R \in SO(d)$, for all $j = 1, \dots, n$ and for all $\boldsymbol{x} \in \mathbb{R}^d$,

$$a_j(R \cdot \boldsymbol{x}) \in \operatorname{span}_{\mathbb{C}} \left\{ \boldsymbol{x}^{\boldsymbol{\alpha}} \exp(\boldsymbol{x} \cdot (R^T \cdot \boldsymbol{\xi})) : \boldsymbol{\alpha} \in \mathbb{N}_0^d, |\boldsymbol{\alpha}| \leq nd, \boldsymbol{\xi} \in \Xi \right\} \tag{32}$$

where $|\boldsymbol{\alpha}| = \alpha_1 + \ldots + \alpha_d$ and $\Xi = \sigma(\Omega_1) \times \ldots \times \sigma(\Omega_d)$. On the other hand, due to the assumed rotation invariance, we also have for all $R \in SO(d)$, for all $j = 1, \dots, n$ and for all $\boldsymbol{x} \in \mathbb{R}^d$,

$$\begin{aligned} a_j(R \cdot \boldsymbol{x}) &\in \operatorname{span}_{\mathbb{C}}\{a_1(\boldsymbol{x}), \dots, a_n(\boldsymbol{x})\} \\ &\subseteq \operatorname{span}_{\mathbb{C}} \left\{ \boldsymbol{x}^{\boldsymbol{\alpha}} \exp(\boldsymbol{x} \cdot \boldsymbol{\xi}) : \boldsymbol{\alpha} \in \mathbb{N}_0^d, |\boldsymbol{\alpha}| \leq nd, \boldsymbol{\xi} \in \Xi \right\}. \end{aligned} \tag{33}$$

We set for $\boldsymbol{\omega} \in \mathbb{R}^d$ with $|\boldsymbol{\omega}| = 1$ and $t \in \mathbb{R}$,

$$\boldsymbol{a}_{\boldsymbol{\omega}}(t) := (a^1_{\boldsymbol{\omega}}(t), \dots, a^n_{\boldsymbol{\omega}}(t))^T := \boldsymbol{a}(t\boldsymbol{\omega}).$$

Then we have due to (31) for all $j \in \{1, \dots, n\}$ and for all $t \in \mathbb{R}$,

$$(a^j_{\boldsymbol{\omega}})(t) \in \operatorname{span}_{\mathbb{C}} \{ t^\nu \exp(t\boldsymbol{\omega} \cdot \boldsymbol{\xi}) : 0 \leq \nu \leq nd, \boldsymbol{\xi} \in \Xi \}$$

Let $\boldsymbol{\omega}' \in \mathbb{R}^d$ with $|\boldsymbol{\omega}'| = 1$. Then there is $R' \in SO(d)$ with $\boldsymbol{\omega} = R' \cdot \boldsymbol{\omega}'$. We deduce from (33) for all $j = 1, \dots, n$ and for all $t \in \mathbb{R}$,

$$(a^j_{\boldsymbol{\omega}})(t) \in \bigcap_{\boldsymbol{\omega}' \in S^2(d)} \operatorname{span}_{\mathbb{C}} \left\{ t^\nu \exp(t\boldsymbol{\omega}' \cdot \boldsymbol{\xi}) : 0 \leq \nu \leq nd, \boldsymbol{\xi} \in \Xi \right\}, \tag{34}$$

where $S^2(d)$ is the unit sphere in $\mathbb{R}^d$. Introducing

$$\Xi^* := (\sigma(\Omega_1) \setminus \{0\}) \times \ldots \times (\sigma(\Omega_d) \setminus \{0\}).$$

we have for all $j = 1, \dots, n$, for all $\boldsymbol{\omega} \in S^2(d)$ and for all $t \in \mathbb{R}$,

$$(a^j_{\boldsymbol{\omega}})(t) = p_{\boldsymbol{\omega}}(t) + \sum_{\boldsymbol{\xi} \in \Xi^*} p_{\boldsymbol{\omega}, \boldsymbol{\xi}}(t) \exp(t\boldsymbol{\omega} \cdot \boldsymbol{\xi}), \tag{35}$$

where $p_{\boldsymbol{\omega}}(.)$, $p_{\boldsymbol{\omega},\boldsymbol{\xi}}$ are (possibly vanishing) polynomials in t of order less or equal nd. We need two auxiliary results.

Proposition 9. *Let $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_\rho \in \mathbb{C}^d \setminus \{0\}$, $\rho \in \mathbb{N}$. Let $\boldsymbol{\omega} \in S^2(d)$. Then there is $\boldsymbol{\omega}' \in S^2(d)$ such that*

$$\{\boldsymbol{\omega} \cdot \boldsymbol{\xi}_l : 1 \leq l \leq \rho\} \cap \{\boldsymbol{\omega}' \cdot \boldsymbol{\xi}_l : 1 \leq l \leq \rho\} = \emptyset.$$

Proof. We set $\alpha_l := \boldsymbol{\omega} \cdot \boldsymbol{\xi}_l$, $1 \leq l \leq \rho$. Since $\boldsymbol{\xi}_{l'} \neq 0$, $1 \leq l' \leq \rho$, the set $A_{l,l'} := \{\boldsymbol{z} \in \mathbb{R}^d : \boldsymbol{z} \cdot \boldsymbol{\xi}_{l'} = \alpha_l\}$ is a hyperplane in $\mathbb{R}^d$. We certainly have

$$\{\alpha_l : 1 \leq l \leq \rho\} \cap \{\boldsymbol{z} \cdot \boldsymbol{\xi}_l : 1 \leq l \leq \rho\} = \emptyset$$

for all $\boldsymbol{z} \in \mathbb{R}^d \setminus \bigcup_{l,l'=1}^{\rho} A_{l,l'}$. Since $S^2(d)$ is not contained in the union of finitely many hyperplanes there is $\boldsymbol{\omega}' \in S^2(d) \cap (\mathbb{R}^d \setminus \bigcup_{l,l'=1}^{\rho} A_{l,l'})$. □

Proposition 10. *Let $\alpha_1, \ldots, \alpha_\rho \in \mathbb{C}$, $\rho \in \mathbb{N}$, such that $\alpha_l \neq \alpha_{l'}$ whenever $l \neq l'$, $l, l' \in \{1, \ldots, \rho\}$ and let $P_1, \ldots, P_\rho$ be non-vanishing complex polynomials. Then $\{P_l(t) \exp(\alpha_l t) : 1 \leq l \leq \rho\}$ is a linear independent subset of $\mathcal{F}(\mathbb{R}, \mathbb{C})$.*

Proof. The functions are contained in a fundamental system of an appropriately constructed ODE system of first order with constant coefficient. □

Choosing $\boldsymbol{\omega}'$ as in Proposition 9 and ignoring vanishing terms we deduce from (35)

$$(a^j_{\boldsymbol{\omega}})(t) = p_{\boldsymbol{\omega}}(t) + \sum_{l=1}^{\rho} p_{\boldsymbol{\omega},l}(t) \exp(t\alpha_l) = p_{\boldsymbol{\omega}'}(t) + \sum_{l=1}^{\rho'} p_{\boldsymbol{\omega}',l}(t) \exp(t\alpha_{\rho+l}), \quad (36)$$

where the polynomials $p_{\boldsymbol{\omega},1}, \ldots, p_{\boldsymbol{\omega},\rho}, p_{\boldsymbol{\omega}',1}, \ldots, p_{\boldsymbol{\omega}',\rho'}$ do not vanish (but empty sums with $\rho = 0$ or $\rho' = 0$ are possible), and $\alpha_1, \ldots, \alpha_{\rho+\rho'}$ are complex numbers with $\alpha_l \neq \alpha_{l'}$ for $l \neq l'$, $l, l' \in \{1, \ldots, \rho + \rho'\}$. Hence by Proposition 10: $\rho = \rho' = 0$. As a consequence, $\alpha^j_{\boldsymbol{\omega}}$ is for each $\boldsymbol{\omega} \in S^2(d)$ and for each $j \in \{1, \ldots, n\}$ a polynomial of degree less or equal nd.

Due to (31) the function a_j, $j \in \{1, \ldots, n\}$, has an expansion in a Taylor series at $\boldsymbol{x} = 0$. Since each $a^j_{\boldsymbol{\omega}}$, $\boldsymbol{\omega} \in S^2(d)$, $j = 1, \ldots, n$, is a polynomial of degree less or equal nd, the Taylor series of a_j, $j = 1, \ldots, n$, has only finitely many non-vanishing terms. Thus, a_j is for each $j = 1, \ldots, n$, a polynomial, and we have shown:

Theorem 11. *Let $\mathbb{M}$ be a translation and rotation invariant, finite dimensional subspace of $\mathcal{F}(\mathbb{R}^d, \mathbb{R})$ with the property that the functions in $\mathbb{M}$ possess all partial derivatives in some point $\bar{\boldsymbol{x}} \in \mathbb{R}^d$. Then, $\mathbb{M}$ is a space of polynomials.*

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*, Nonlinearity, **14**, 2001, 881–906.

[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] H. Grad, *On the Kinetic Theory of Rarefied Gases*, Comm. Pure & Appl. Math. **2**, 1949, 331–407.

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

[9] M. Junk, *Maximum entropy for reduced moment problems*, Math. Models Methods Appl. Sci. **10**, 2000, 1001–1025.

[10] M. Junk, *Maximum entropy moment problems and extended Euler equations*, proceedings of IMA Workshop "Simulation of Transport in Transition Regimes", to appear

[11] M. Junk, *Moment problems in kinetic theory*, Habilitationsschrift, Universität Kaiserslautern, 2001.

[12] A. S. Lewis, *Consistency of moment systems*, Can. J. Math. **47**, 1995, 995–1006.

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

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

[15] I. Müller, T. Ruggeri, *Rational Extended Thermodynamics.* 2nd Edition, Springer Tracts in Natural Philosophy, Springer New York, 1998.

[16] T. Ruggeri, *Galilean invariance and entropy principle for systems of balance laws*, Continuum Mech. Thermodyn. **1**, 1989, 3–20.

[17] J. Smoller, *Shock waves and reaction-diffusion equations*, Springer, 1983.