# Convex Duality and Entropy-Based Moment Closures: Characterizing Degenerate Densities

Cory D. Hauck
Computational Physics and Methods (CCS-2) &
Center for Nonlinear Studies (CNLS)
Los Alamos National Laboratory
cdhauck@lanl.gov

C. David Levermore
Department of Mathematics &
Institute for Physical Sciences and Technology
University of Maryland, College Park
lvrmr@math.umd.edu

André L. Tits
Department of Electrical and Computer Engineering &
Institute for Systems Research
University of Maryland, College Park
andre@umd.edu

November 15, 2007

**Abstract**

A common method for constructing a function from a finite set of moments is to solve a constrained minimization problem. The idea is to find, among all functions with the given moments, that function which minimizes a physically motivated, strictly convex functional. In the kinetic theory of gases, this functional is the kinetic entropy; the given moments are macroscopic densities; and the solution to the constrained minimization problem is used to formally derive a closed system of partial differential equations which describe how the macroscopic densities evolve in time. Moment equations are useful because they simplify the kinetic, phase-space description of a gas, and with entropy-based closures, they retain many of the fundamental properties of kinetic transport.

Unfortunately, in many situations, macroscopic densities can take on values for which the constrained minimization problem does not have a solution. Essentially, this is because the moments are not continuous functionals with respect to the $L^1$ topology. In this paper, we give a geometric description of these so-called *degenerate densities* in the most general possible setting. Our key tool is the complementary slackness condition that is derived from a dual formulation of a minimization problem with relaxed constraints. We show that the set of degenerate densities is a union of convex cones and, under reasonable assumptions, that this set is small in both a topological and measure theoretic sense. This result is important for further assessment and implementation of entropy-based moment closures.

**Keywords**: *convex duality, convex optimization, optimization in function spaces, kinetic theory, entropy-based closures, moment equations, gas dynamics.*

**AMS classification**: 49N15, 90C25, 82C40, 35A35, 94A17, 76N15, 14P10

## 1 Introduction

In gas dynamics, the kinetic description of a gas is often simplified using moment equations. In this reduced setting, a gas is characterized by a finite-dimensional vector $\boldsymbol{\rho}$ of densities that are moments of the kinetic distribution function $F$ with respect to polynomials of the microscopic velocity. Evolution equations for $\boldsymbol{\rho}$ are derived by taking moments of the Boltzmann equation which governs the evolution of $F$. The derivation requires that an approximation for $F$ be reconstructed from the densities $\boldsymbol{\rho}$, giving what is called a *closure.*

One well-known method for prescribing a closure is to find a function that minimizes the kinetic entropy subject to the constraint that its moments agree with $\boldsymbol{\rho}$. Such closures are called *entropy-based closures.* In recent years, they have generated substantial interest due to important structural properties which they inherit from the Boltzmann equation. These properties were first brought to light in [Lev96].

In cases where the moments are continuous with respect to the relevant topology, there is always an entropy minimizer [BL91, Jun00]. Unfortunately, in classical gas dynamics, this is not usually the case. As a result, there are often physically relevant densities for which the constrained entropy minimization problem does not have a solution. In such cases, entropy-based closures are not well-defined, and these densities are called *degenerate.* In this paper, we provide a geometric description for the set of degenerate densities in the most general possible setting. We believe that this description is an important step in assessing the practical usefulness of entropy-based closures.

## 1.1 Moment Systems and Entropy-Based Closures

Consider a gas that is enclosed in a container, represented mathematically by the set $\Omega \subset \mathbb{R}^d$ (typically $d = 3$). The kinetic distribution function $F = F(v, x, t)$ which describes the kinetic state of the gas is a non-negative function that is defined for positions $x \in \Omega$, velocities $v \in \mathbb{R}^d$, and times $t \geq 0$ so that, for any measurable set $\Lambda \subset \Omega \times \mathbb{R}^d$,

$$\int_\Lambda F(v, x, t)\; dvdx \tag{1}$$

gives the number of particles at time $t$ with positions $x$ and velocities $v$ such that $(v, x) \in \Lambda$. The evolution of $F$ is governed by the Boltzmann transport equation

$$\partial_t F + v \cdot \nabla_x F = \mathcal{C}(F)\,, \tag{2}$$

where $\mathcal{C}$ is an integral operator that describes the collisions between particles which drive the system to local thermal equilibrium.

Solutions of (2) formally satisfy the local balance law [CIP94]

$$\partial_t \mathcal{H}(F) + \nabla_x \cdot \mathcal{J}(F) = \mathcal{S}(F)\,, \tag{3}$$

where the functionals

$$\mathcal{H}(g) \equiv \langle g \log(g) - g \rangle \quad \text{and} \quad \mathcal{J}(g) \equiv \langle v(g \log(g) - g) \rangle \tag{4}$$

are the *kinetic entropy* and *kinetic entropy flux*, respectively, and

$$\mathcal{S}(g) \equiv \langle \log(g) \mathcal{C}(g) \rangle \tag{5}$$

is the *kinetic entropy dissipation.* Here and throughout this paper, $\langle \cdot \rangle$ denotes Lebesgue integration over all $v \in \mathbb{R}^d$, and we assume for the moment that the integrals in (4) and (5) are well-defined. According to Boltzmann's "H-Theorem" [CIP94],

$$\mathcal{S}(g) \leq 0\,, \tag{6}$$

with equality if and only if $\mathcal{C}(g) = 0$. In such cases, $g$ is said to be in a state of local thermal equilibrium; and it takes the form of a Maxwellian distribution

$$\mathcal{M}_{\rho,u,\theta}(v) \equiv \frac{\rho}{(2\pi\theta)^{d/2}} \exp\left( -\frac{|v - u|^2}{2\theta} \right)\,, \tag{7}$$

where $\rho$ and $\theta$ are positive scalars and $u \in \mathbb{R}^d$. In this way, $\mathcal{H}$ acts as a Lyapunov functional for (2).

In order to reduce computational cost, the kinetic description of a gas provided by $F$ is often simplified by retaining only a finite number of its velocity averages, or *moments.* Equations which govern the evolution of these moments are derived by integrating (2) with respect to a vector

$$\mathbf{m} = (m_0, \ldots, m_{n-1})^T \tag{8}$$

whose components are polynomials in $v$. Since $v$ commutes with the spatial gradient, these equations take the form

$$\partial_t \boldsymbol{\rho} + \nabla_x \cdot \langle v\mathbf{m}F \rangle = \langle \mathbf{m}\mathcal{C}(F) \rangle \ , \tag{9}$$

where the moments

$$\boldsymbol{\rho} = \boldsymbol{\rho}(x,t) \equiv \langle \mathbf{m}F \rangle \tag{10}$$

are the *spatial densities* associated with $F$. Here again, we assume that the integrals in (9) and (10) are well-defined.

In general, (9) is not a closed system because there is no way to express the flux terms $\langle v\mathbf{m}F \rangle$ and collision terms $\langle \mathbf{m}\mathcal{C}(F) \rangle$ in terms of $\boldsymbol{\rho}$. Furthermore, in a moment description, an exact expression for $F$ is not available. An alternative is to approximate $F$ by an ansatz of the form

$$\mathcal{F}[\boldsymbol{\rho}] = \mathcal{F}(v, \boldsymbol{\rho}(x,t)) \, , \tag{11}$$

By substituting $\mathcal{F}$ for $F$ in (9), the evolution of $\boldsymbol{\rho}$ can be approximated by the closed system of balance laws

$$\partial_t \boldsymbol{\rho} + \nabla_x \cdot \mathbf{f}(\boldsymbol{\rho}) = \mathbf{c}(\boldsymbol{\rho}) \, , \tag{12}$$

where the flux term $\mathbf{f}$ and collision term $\mathbf{c}$ are given by

$$\mathbf{f}(\boldsymbol{\rho}) = \langle v\mathbf{m}\mathcal{F}[\boldsymbol{\rho}] \rangle \qquad \text{and} \qquad \mathbf{c}(\boldsymbol{\rho}) = \langle \mathbf{m}\mathcal{C}(\mathcal{F}[\boldsymbol{\rho}]) \rangle \ . \tag{13}$$

One way to specify $\mathcal{F}$ is to invoke the principle of entropy minimization (or maximization in the physics community, where the term "entropy" refers to $-\mathcal{H}$ and has been widely used for over a century in the physics community). The probabilistic interpretation of entropy dates back to Boltzmann [Bol68, Bol77], who argued that the entropy of a system of identical particles depends on the number of microstates (particle arrangements in phase space) that are consistent with the macroscopic state of the system. This dependence is expressed by the famous logarithmic relationship known as *Boltzmann's entropy formula* [Cal85] (and also as Boltzmann's equation, although distinct from (2)) and was first presented in its popular form by Planck [Pla00, Pla01]. The practical application of entropy as a tool for statistical inference was championed by Jaynes although, in [Jay57], Jaynes himself attributes the original mathematical concepts to Gibbs, who generalized Boltzmann's entropy formula [Gib02]. Jaynes also credits Shannon [Sha48] for illuminating the central role that entropy plays in the theory of information. The relationship between statistics and information theory was further pursued by Kullback [Kul59]. Many of the first rigorous results concerning entropy minimization can be found in [Csi75] and references therein.

Closures which are based on the entropy minimization principle use the ansatz

$$\mathcal{F}[\boldsymbol{\rho}] = \arg \min_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m}g \rangle = \boldsymbol{\rho}\} \tag{14}$$

at each $x$ and $t$ to formally close (9). Here

$$\mathbb{F}_{\mathbf{m}} \equiv \left\{ g \in L^1(\mathbb{R}^d) : g \gneqq 0 \text{ and } |\mathbf{m}g| \in L^1(\mathbb{R}^d) \right\} \ , \tag{15}$$

and $|\cdot|$ is the standard Euclidean norm.

It is readily checked that $\mathcal{H}$ is strictly convex over $\mathbb{F}_{\mathbf{m}}$. Thus if the minimizer in (14) exists, it is unique and the closure is well-defined. In such cases, (12) is a hyperbolic system of PDE whose solutions satisfy the local dissipation law

$$\partial_t h(\boldsymbol{\rho}) + \nabla_x \cdot j(\boldsymbol{\rho}) = s(\boldsymbol{\rho}), \tag{16}$$

where

$$h(\boldsymbol{\rho}) \equiv \mathcal{H}(\mathcal{F}[\boldsymbol{\rho}]) \tag{17}$$

is a strictly convex function of $\boldsymbol{\rho}$ and

$$j(\boldsymbol{\rho}) \equiv \mathcal{J}(\mathcal{F}[\boldsymbol{\rho}]) \, , \qquad s(\boldsymbol{\rho}) \equiv \mathcal{S}(\mathcal{F}[\boldsymbol{\rho}]) \leq 0 \, . \tag{18}$$

Although any choice for the ansatz $\mathcal{F}[\boldsymbol{\rho}]$ will yield a system of the form (12), the entropy ansatz ensures that $s(\boldsymbol{\rho}) \leq 0$ and that $h$ is strictly convex. These conditions are important for two reasons. First, the dissipation law for a strictly convex function of $\boldsymbol{\rho}$, as given by (16), implies the existence of a well-posed linear $L^2$ (Hilbert space) theory for (12) [Str04]. Second, $h$ acts as a Lyapunov function for (12). To see this, note that (16) is simply (3) evaluated at $F = \mathcal{F}[\boldsymbol{\rho}]$; and like in Boltzmann's H-Theorem, $s(\boldsymbol{\rho})$ vanishes if and only if $\mathcal{C}(\mathcal{F}[\boldsymbol{\rho}]) = 0$, in which case $\mathcal{F}[\boldsymbol{\rho}]$ takes the form of a Maxwellian distribution [Lev96].

The entropy minimization procedure yields an entire hierarchy of systems with the aforementioned properties whose members are generated by appending an initial choice of $\mathbf{m}$ with additional polynomial components. For this reason, entropy-based closures have been applied to other areas of kinetic theory such as radiation transport [DF99, DK02] and charge transport in semiconductors [AMR03, DR03, JR04]. (Additional references for charge transport can be found in [AMR03].) In the case of gas dynamics, the moment hierarchy begins with the canonical choice $\mathbf{m} = (1, v, \frac{1}{2}|v|^2)^T$. For this choice, $\mathcal{F}[\boldsymbol{\rho}]$ is always a Maxwellian and the entropy-based closure generates Euler's equations for a compressible gas.

## 1.2 Realizability and Degenerate Densities

A density $\boldsymbol{\rho}$ is said to be *realized* by a function $g \in \mathbb{F}_{\mathbf{m}}$ if $\boldsymbol{\rho} = \langle \mathbf{m} g \rangle$. The set of all such *realizable* densities will be denoted by $\mathcal{R}_{\mathbf{m}}$. An entropy-based closure is applicable only to those realizable densities for which the minimization problem (14) with equality constraints has a solution. *If* the moments in (14) were continuous with respect to the $L^1$ topology, then there would always be a minimizer. Indeed, for such cases, Borwein and Lewis have shown in [BL91] that a constrained minimizer exists for a large class of convex functionals that include the classical entropy $\mathcal{H}$. However, in gas dynamics, the moments are typically *not continuous* in the $L^1$ topology. As a result, there are realizable densities for which the minimizer in (14) does not exist. For such densities, which we term *degenerate*, modifications must be made to the entropy-based procedure. There are essentially two approaches:

1. Show that the set of non-degenerate densities is invariant under the dynamics of the balance law (12) with the entropy-based closure (as discussed in [Jun98]) or impose such a condition in a way that is physically reasonable and mathematically justifiable.

2. Develop a modified closure that (i) is well-posed for *all* physically realizable values of $\boldsymbol{\rho}$, (ii) recovers the minimum entropy-based closures whenever the minimizer in (14) exists, and (iii) generates systems of hyperbolic PDE that dissipate a physically meaningful, convex entropy. This is the approach taken in [Sch04].

We define $\mathcal{D}_{\mathbf{m}}$ to be the set of all degenerate densities. In general, the set $\mathcal{D}_{\mathbf{m}}$ depends on $\mathbf{m}$, and understanding its geometry is critical to determining whether entropy-based closures can be used in practice. In either of the modified approaches listed above, it is important—at the very least—to show that $\mathcal{D}_{\mathbf{m}}$ is small in some sense, thereby minimizing the number of physically realizable spatial densities which require special treatment. In the first approach, this means limiting the number of initial conditions which must be discarded; in the second, it means limiting the number of physically realizable densities which require a modified closure.

Another reason to study $\mathcal{D}_{\mathbf{m}}$ is that the equilibrium densities, i.e., those densities which are moments of a Maxwellian distribution (7), lie on its boundary [Jun98, Jun00, Sch04, Lev96]. Because the kinetic entropy drives solutions of (3) toward local thermal equilibrium, we expect that trajectories defined by solutions to (16) will, at times, come very close to $\mathcal{D}_{\mathbf{m}}$. Thus it is very important to have a detailed understanding of its geometry.

Previous studies of the set $\mathcal{D}_{\mathbf{m}}$ can be found in [Jun98, Jun00, Sch04]. In [Jun98], Junk provides a geometric description for $\mathcal{D}_{\mathbf{m}}$ in a one-dimensional setting ($d = 1$) with $\mathbf{m} = (1, v, v^2, v^3, v^4)^T$. In turns out in this case that $\mathcal{D}_{\mathbf{m}}$ is a co-dimension one manifold. This result was discovered, in part, by extending the definition of $h$ given by (18) to include cases where the minimizer in (14) does not exist. This is done simply by replacing the minimum in (14) with an infimum, viz.

$$h_{\mathrm{J}}(\boldsymbol{\rho}) \equiv \inf_{g \in \mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m} g \rangle = \boldsymbol{\rho}\} \ . \tag{19}$$

Later, in [Jun00], Junk considers a more general case in which $\mathbf{m}$ consists of a radial component $|v|^N$, for some even integer $N \geq 2$, plus polynomials components of lower degree. For such cases, he provides an integrability condition to determine whether $\mathcal{D}_{\mathbf{m}}$ is non-empty. In practice, this condition is easily checked and extensible to more general choices of $\mathbf{m}$. However, a description of the geometry of $\mathcal{D}_{\mathbf{m}}$, as given in [Jun98], is still lacking for the general setting.

In [Sch04], Schneider introduces a different extension for $h$ by relaxing the constraints in (14):

$$h_{\mathrm{S}}(\boldsymbol{\rho}) \equiv \min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho} \right\} . \tag{20}$$

Here the notation $\langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho}$ means—roughly speaking—that inequalities between certain components are allowed. (See Section 3.2 for a precise definition.) The key difference between (14) and (20) is that the constraint set of the latter is closed in the weak-$L^1(\mathbb{R}^d)$ topology, whereas the constraint set of the former is not. Schneider uses this fact to prove that the minimizer in (20) with relaxed constraints always exists and is equal to the minimizer with equality constraints (14) when that minimizer exists (see our Theorem 3 and Corollary 4 below). In doing so, he provides a necessary and sufficient condition to determine whether a given density $\boldsymbol{\rho}$ is an element of $\mathcal{D}_{\mathbf{m}}$. However, this condition gives little insight into the geometry of $\mathcal{D}_{\mathbf{m}}$.

The main contribution of the present paper is a geometrical description of the set $\mathcal{D}_{\mathbf{m}}$ in the most general possible setting. Our results are based on a dual formulation of (20) and are summarized in the following theorems.

- In Theorems 14 and 16, we prove strong duality for both the equality constraint problem (19) and the relaxed constraint problem (20). One consequence of these theorems is that $h_{\mathrm{S}} = h_{\mathrm{J}}$, even when the infimum in (20) is not attained. In Theorem 14, we also prove *complementary slackness conditions* which relate the density $\boldsymbol{\rho}$ in (20) to the dual variable and serve as the basis of our geometrical description.

- In Theorem 25, we show that the set $\mathcal{D}_{\mathbf{m}}$ is a union of convex cones. The vertices of these cones are *non*-degenerate densities that lie on the boundary between the degenerate and non-degenerate densities in $\mathcal{R}_{\mathbf{m}}$. This conical description is based on the complementary slackness condition from Theorem 14.

- In Theorem 28 we show that, under reasonable assumptions, the set $\mathcal{D}_{\mathbf{m}}$ is a nowhere dense subset of $\mathcal{R}_{\mathbf{m}}$ that has Lebesgue measure zero and is restricted to the boundary of the *non*-degenerate, realizable densities. The assumptions we employ hold in all known cases. Whether they hold in general is an interesting and (to our knowledge) open question in analysis and algebraic geometry.

In the process of investigating $\mathcal{D}_{\mathbf{m}}$, we also recover and extend many previous results from both [Jun98,Jun00] and [Sch04].

The organization of the paper is as follows. In Section 2 we introduce some notation and background information. In Section 3 we review the entropy minimization problem. In Section 4 we give a dual formulation of the minimization problem with relaxed constraints (20) and prove duality theorems for both (19) and (20). We use these theorems to show that $h_S = h_J$ (even when the infimum in (19) is not attained) and to establish a complementary slackness condition. In Section 5 we review the formal structure of entropy-based closures for non-degenerate densities and determine how that structure differs for degenerate cases. In Section 6 we use the complementary slackness condition to describe the geometry of $\mathcal{D}_{\mathbf{m}}$. We then introduce the assumptions that allow us to make further assertions about the 'smallness' of $\mathcal{D}_{\mathbf{m}}$. At the end of the section, we present two examples. In Section 7 we give conclusions and discuss future work. Finally, in the appendix we provide a diagram and tables to assist the reader with notation.

## 2 Preliminaries

In this section, we introduce notation and present preliminary results. We refer the reader to the appendix for help in recalling the notation and useful properties for sets and mappings given throughout the paper.

## 2.1 Admissible Spaces

For a given moment system, the choice of $\mathbf{m}$ must satisfy criteria based on physical considerations. We require that components of $\mathbf{m}$ form a basis for an $n$ dimensional linear space $\mathbb{M}$ of multivariate polynomials over the field of real numbers that satisfies the following conditions:

$$\begin{aligned} &\text{I. } \mathbb{M} \supset \operatorname{span}\{1, v_1, \ldots v_d, |v|^2\}\,; \\ &\text{II. } \mathbb{M} \text{ is invariant under translation and rotation;} \\ &\text{III. The set } \mathbb{M}_{\mathrm{c}} \equiv \{p \in \mathbb{M} : \langle |p| \exp{(p)} \rangle < \infty\} \text{ has non-empty interior.} \end{aligned} \tag{21}$$

Our definition of $\mathbb{M}_{\mathrm{c}}$ is slightly different than the original definition given in [Lev96]. However, its interior is the same under both definitions.

Spaces that satisfy Conditions I-III are called *admissible.* According to Condition I, any set of moment equations will incorporate the conservation laws for mass, momentum, and energy which are given by the moments of the kinetic distribution function with respect to 1, $v$, and $\frac{1}{2}|v|^2$, respectively. In Condition II, invariance under translation and rotation means that for every $u \in \mathbb{R}^d$ and every orthogonal matrix $O$, the mappings $v \mapsto v - u$ and $v \mapsto O^T v$ map $\mathbb{M}$ onto itself. These properties will ensure that the moment equations are Galilean invariant—that is, invariant under the transformations $x \mapsto x - ut$ and $x \mapsto O^T x$. Condition III applies specifically to entropy-based closures. It turns out that the minimizer (14), if it exists, has the form $e^p$, where $p \in \mathbb{M}_{\mathrm{c}}$. Hence a non-empty interior for $\mathbb{M}_{\mathrm{c}}$ is a necessary requirement for any practical applications.

Typically an admissible space $\mathbb{M}$ is generated by the span of polynomial functions whose moments are physical quantities of specific interest. (Here the canonical examples are the polynomials 1, $v$, and $\frac{1}{2}|v|^2$). It may be that additional polynomial components are added to $\mathbf{m}$ to ensure that $\mathbb{M}$ is admissible. It should be noted that the vector $\mathbf{m}$ that generates a given $\mathbb{M}$ is not unique.

For convenience, we will assume, without loss of generality, that the components of $\mathbf{m}$ are homogeneous. We decompose $\mathbf{m}$ into sub-vectors:

$$\mathbf{m} = (\mathbf{m}_0^T, \mathbf{m}_1^T, \mathbf{m}_2^T, \ldots, \mathbf{m}_N^T)^T\,, \tag{22}$$

where the $n_j$ components of $\mathbf{m}_j$ are the $j^{th}$ degree polynomial components of $\mathbf{m}$. Consistency requires that $\sum_{j=0}^{N} n_j = n$. Any polynomial $p \in \mathbb{M}$ can be expressed as the sum of its homogeneous components:

$$p = \boldsymbol{\alpha}^T \mathbf{m} = \sum_{j=1}^{N} \boldsymbol{\alpha}_j^T \mathbf{m}_j\,, \tag{23}$$

were $\boldsymbol{\alpha} \in \mathbb{R}^n$ is a vector of constant coefficients that decomposes into sub-vectors

$$\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T, \ldots, \boldsymbol{\alpha}_N^T\right)^T\,. \tag{24}$$

We briefly outline how one can generate a space $\mathbb{M}$. Given the even integer $N \geq 2$ and $j < N$, let $\mathbb{Q}_j$ be the space of all homogeneous polynomials from $\mathbb{R}^d$ to $\mathbb{R}$ of degree $j$. For each $j$, $\mathbb{Q}_j$ can be composed into rotationally invariant subspaces in the following way [Fol76, Corollary 2.60]

$$\mathbb{Q}_j = \begin{cases} \mathbb{H}_j \oplus |v|^2 \mathbb{H}_{j-2} \oplus |v|^4\, \mathbb{H}_{j-4} \oplus \ldots \oplus |v|^j \mathbb{H}_0\,, & j \text{ even,} \\ \mathbb{H}_j \oplus |v|^2 \mathbb{H}_{j-2} \oplus |v|^4\, \mathbb{H}_{j-4} \oplus \ldots \oplus |v|^{j-1} \mathbb{H}_1\,, & j \text{ odd.} \end{cases} \tag{25}$$

Here $\mathbb{H}_k$ is the space of harmonic polynomials of degree $k$ given by

$$\mathbb{H}_k = |v|^k \operatorname{span}\left\{Y^k\left(\frac{v}{|v|}\right)\right\}\,, \tag{26}$$

and $Y^k$ maps vectors on the unit sphere $\mathbb{S}^{d-1}$ to the $k$-fold spherical harmonic tensor, which is unique modulo constant multiples. (Here the term "span" refers to all real linear combinations of the scalar components of the tensor.)

The decomposition in (25) is unique in the sense that no proper subset of the subspaces in (25) is rotationally invariant [Fol76]. Thus, in order to be rotationally invariant, an admissible space $\mathbb{M}$ must be a direct sum of some combination of the subspaces in (25) taken from each $\mathbb{Q}_j$, $j \leq N$. In addition, the condition of translational invariance implies that choices for larger values of $j$ will directly affect choices for smaller values of $j$. For example, inclusion of the term $|v|^j$ requires inclusion of the lower degree terms in the expansion of $|v-u|^j$.

To satisfy Condition III, $\mathbb{M}$ must include polynomials from $\mathbb{Q}_N$ which dominate the behavior of odd degree polynomials of lower degree for large $|v|$. In particular, $\mathbb{M}$ must include multiples of $|v|^N$. This is because spherical harmonics (both odd and even) other than $Y^0 \equiv 1$ take on both positive and negative values on the unit sphere. Excluding $|v|^N$ would therefore lead to polynomials $p$, all of which satisfy $\lim_{r\to\infty} p(r\omega) = \infty$ for all $\omega$ contained in some subset of $\mathbb{S}^{d-1}$ with positive Lebesgue measure. In such cases $\exp(p)$ is not integrable for any $p \in \mathbb{M}$, and Condition III is violated.

In applications it is sometimes convenient to represent components of $\mathbf{m}$ in tensor format. There are two reasons for this. The first reason is the convenience with which one can express $\mathbf{m}$ given (25) and (26). The second reason is that the evolution of the moment of $\langle TF \rangle$ for any $j$-fold tensor $T = T(v)$ depends on the divergence of the $(j+1)$-fold tensor $\langle vTF \rangle$ (refer to (9)).

The tensors in which we are interested are often symmetric and sometimes traceless. For example, the Gaussian closure which will be described in Section 5.3 is based on the vector

$$\mathbf{m} = \begin{pmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ v \\ v \vee v \end{pmatrix} = \begin{pmatrix} 1 \\ v \\ \left(v \vee v - \frac{1}{3}|v|^2 I\right) + \frac{1}{3}|v|^2 \end{pmatrix}, \tag{27}$$

where $v \vee v$ is the symmetric tensor product of $v$ with itself. [1] In the strict vector representation, $\mathbf{m}_2$ is composed only of the $d(d+1)/2$ linearly independent components of the tensor $v \vee v$. The components have the form $v_i v_j$ where $1 \leq i \leq d$ and $i \leq j \leq d$.

Vectors $\boldsymbol{\alpha} \in \mathbb{R}^n$ can also be represented by tensors, in which case the product in (23) is interpreted as a sum of tensor inner products. [2] However, for a given a polynomial $p$, the tensor form of $\boldsymbol{\alpha}$ in (23) is unique only under the additional requirement that it have the same symmetry properties as $\mathbf{m}$.

## 2.2 Cones

Many of the sets that we will encounter in this paper are cones [BNO03,Roc70]. A subset $C$ of $\mathbb{R}^k$ is a *cone* if, for all real numbers $\lambda > 0$, $y \in C$ if and only if $\lambda y \in C$. A cone is *solid* if it has non-empty interior. A closed cone $C$ is *pointed* if $-C \cap C$ is the origin. A closed cone that is convex, pointed, and solid is called *proper*. For example, the set $\mathbb{F}_{\mathbf{m}}$ is a solid, convex cone, whose closure in $L^1(\mathbb{R}^d)$ is proper. Several other cones will be introduced in the subsections that follow, and eventually, we will see that the set $\mathcal{D}_{\mathbf{m}}$ is also a cone.

Associated with every cone $C$ is its polar cone [3]

$$C^\circ \equiv \left\{ z \in \mathbb{R}^k : z^T y \leq 0 \quad \forall y \in C \right\} . \tag{28}$$

[1] Given a symmetric $j$-fold tensor $S$ and a symmetric $k$-fold tensor $T$, the symmetric tensor product of $S$ and $T$ is

$$S \vee T = T \vee S \equiv \frac{1}{(j+k)!} \sum_{\pi \in \Pi} S_{i_{\pi(1)},\dots i_{\pi(j)}} T_{i_{\pi(j+1)},\dots i_{\pi(j+k)}} ,$$

where $\Pi$ is the set of all permutation of the integers $1, \dots, j+k$.

[2] For $k > j$, the symmetric inner product (or contraction) of a symmetric $j$-fold tensor $S$ and a symmetric $k$-fold tensor $T$ is

$$(S \cdot T)_{i_{j+1},\dots,i_{j+k}} \equiv \sum_{i_1,\dots,i_j} S_{i_1,\dots,i_j} T_{i_1,\dots,i_j,i_{j+1},\dots,i_{j+k}} .$$

[3] The polar cone is the negative of the dual cone

$$C^* \equiv \left\{ z \in \mathbb{R}^k : z^T y \geq 0 \quad \forall y \in C \right\} .$$

It is readily checked that the polar of a proper cone is proper.

A vector $z \in \mathbb{R}^k$ is *tangent* to a subset $\Omega \subset \mathbb{R}^k$ at a point $y \in \Omega$ if $z = 0$ or if

$$\lim_{j\to\infty} \frac{y_j - y}{|y_j - y|} = \frac{z}{|z|} \tag{29}$$

for some sequence $\{y_j\}_{j=1}^{\infty} \subset \Omega$ such that $y_j \to y$, but $y_j \neq y$ for all $j$. The *tangent cone of* $\Omega$ *at* $y$, which we denote $\mathcal{TC}(\Omega, y)$, is the set of all vectors that are tangent to $\Omega$ at $y$. A vector $w \in \mathbb{R}^k$ is *normal* to $\Omega$ at $y \in \Omega$ if there exist sequences $\{y_j\}_{j=1}^{\infty} \subset \Omega$ and $\{w_j\}_{j=1}^{\infty} \subset \mathbb{R}^k$ such that

$$y_j \to y\,, \qquad w_j \to w\,, \qquad w_j \in (\mathcal{TC}(\Omega, y_j))^{\circ} \quad \forall\, j\,. \tag{30}$$

The *normal cone of* $\Omega$ *at* $y$, which we denote $\mathcal{NC}(\Omega, y)$, is the set of all vectors that are normal to $\Omega$ at $y$. For the important case that $\Omega$ is convex,

$$\mathcal{NC}(\Omega, y) = \left\{ z \in \mathbb{R}^k : z^T(y' - y) \leq 0\,, \quad \forall\, y' \in \Omega \right\}\,. \tag{31}$$

In particular, $\mathcal{NC}(\Omega, y)$ is convex. If $\partial\Omega$ is a $C^1$ (continuously differentiable) manifold containing $y$, then $\mathcal{NC}(\Omega, y)$ is a ray with base point at the origin that points in the outward normal direction to $\partial\Omega$ at $y$. More generally, given any $C^1$ manifold $M \ni y$ of dimension $j$, $\mathcal{NC}(M, y)$ is a subspace of dimension $n - j$. If $M \subset \Omega$, then $\mathcal{NC}(\Omega, y) \subset \mathcal{NC}(M, y)$. In Sections 5.4 and 6.4, we will use the notation $\mathcal{NC}_0(\Omega, y)$ to denote the normal cone without the origin:

$$\mathcal{NC}_0(\Omega, y) \equiv \mathcal{NC}(\Omega, y) \backslash \{0\}\,. \tag{32}$$

A particularly useful application of cones is to provide a partial ordering of elements in $\mathbb{R}^k$ (or, more generally, in any vector space). Given a pointed, convex cone $C$ and $y_1$ and $y_2$ in $\mathbb{R}^k$, we say that $y_1 \leq_C y_2$, or $y_2 \geq_C y_1$, if and only if $y_2 - y_1 \in C$.

## 2.3 Realizable Densities

Our motivation for solving (14), (19), or (20) is to find a closure for the moment equations (9). Thus we are only interested in constraints based on densities which are realizable, i.e., elements of the set

$$\mathcal{R}_{\mathbf{m}} \equiv \{\boldsymbol{\rho} \in \mathbb{R}^n : \boldsymbol{\rho} = \langle \mathbf{m} g \rangle\,,\, g \in \mathbb{F}_{\mathbf{m}}\}\,. \tag{33}$$

With this notation we formally define the set $\mathcal{D}_{\mathbf{m}}$:

$$\mathcal{D}_{\mathbf{m}} \equiv \{\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}} : \text{the minimizer in (14) does not exist}\}\,. \tag{34}$$

A density $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ has a natural decomposition based on the decomposition of $\mathbf{m}$ in (22):

$$\boldsymbol{\rho} = \left(\boldsymbol{\rho}_0^T, \boldsymbol{\rho}_1^T, \boldsymbol{\rho}_2^T, \ldots, \boldsymbol{\rho}_N^T\right)^T\,, \tag{35}$$

where $\boldsymbol{\rho}_j = \langle \mathbf{m}_j g \rangle$ for some $g \in \mathbb{F}_{\mathbf{m}}$. The set $\mathcal{R}_{\mathbf{m}}$ has several important properties, one of which is its relation to the cone

$$A_{\mathbf{m}} \equiv \left\{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}^T \mathbf{m} \leq 0\right\}\,. \tag{36}$$

It is straight-forward to verify that $A_{\mathbf{m}}$ is a proper cone.

**Theorem 1 (Junk [Jun00])** *The set $\mathcal{R}_{\mathbf{m}}$ is an open, convex, solid cone; and its closure is proper. In fact, $\mathcal{R}_{\mathbf{m}} = \operatorname{int} A_{\mathbf{m}}^{\circ}$, and every vector in $\mathcal{R}_{\mathbf{m}}$ is realized by a bounded, non-negative function with compact support.*

**Proof.** We refer the reader to Theorem A.2 of [Jun00] for a proof (which applies to the case $\mathbf{m}_N = |v|^N$, but can be modified to the general case with little effort). However, to provide the reader with some intuition, we show here that $\mathcal{R}_{\mathbf{m}} \subset \operatorname{int} A_{\mathbf{m}}^{\circ}$. Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$. Then $\boldsymbol{\rho} = \langle \mathbf{m} g \rangle$ for some $g \in \mathbb{F}_{\mathbf{m}}$ and according to (36)

$$\boldsymbol{\alpha}^T \boldsymbol{\rho} = \left\langle \boldsymbol{\alpha}^T \mathbf{m} g \right\rangle \leq 0 \tag{37}$$

for all $\boldsymbol{\alpha} \in A_{\mathbf{m}}$. Further, since $\boldsymbol{\alpha}^T \mathbf{m}$ is a polynomial, it can be zero only on a set of zero Lebesgue measure. Hence $\boldsymbol{\alpha}^T \boldsymbol{\rho} < 0$, which proves $\boldsymbol{\rho} \in \operatorname{int} A_{\mathbf{m}}^{\circ}$. ■

### 2.4 Exponentially Realizable Densities

We will see below that the minimizer of (20) has the form

$$G_{\boldsymbol{\alpha}} \equiv \exp(\boldsymbol{\alpha}^T \mathbf{m})\,, \tag{38}$$

where $\boldsymbol{\alpha}$ solves the dual problem to (20). The integral of $G_{\boldsymbol{\alpha}}$ is the *density potential*

$$h^*(\boldsymbol{\alpha}) \equiv \langle G_{\boldsymbol{\alpha}} \rangle\ , \tag{39}$$

which was introduced in [Lev98] as a tool for elucidating the formal structure of entropy-based closures. As the notation suggests, $h^*$ is the Legendre dual of $h$. In Section 5, we will discuss this relationship in more detail. The name "density potential" is derived from the fact that its formal derivative $\mathbf{r}$ generates the moments of $G_{\boldsymbol{\alpha}}$. Given the set

$$\mathcal{A}_{\mathbf{m}} \equiv \{\boldsymbol{\alpha} \in \mathbb{R}^n : G_{\boldsymbol{\alpha}} \in \mathbb{F}_{\mathbf{m}}\}\ , \tag{40}$$

$\mathbf{r} : \mathcal{A}_{\mathbf{m}} \to \mathbb{R}^n$ is defined by

$$\mathbf{r}(\boldsymbol{\alpha}) \equiv \langle \mathbf{m} G_{\boldsymbol{\alpha}} \rangle\ . \tag{41}$$

It should be noted in (40) that the condition $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ is stronger than $G_{\boldsymbol{\alpha}} \in L^1\left(\mathbb{R}^d\right)$, since the latter can still yield moments that are infinite. The image of $\mathcal{A}_{\mathbf{m}}$ under $\mathbf{r}$ is the set of *exponentially realizable densities*:

$$\mathcal{R}_{\mathbf{m}}^{\exp} \equiv \mathbf{r}(\mathcal{A}_{\mathbf{m}})\,. \tag{42}$$

The set $\mathcal{R}_{\mathbf{m}}^{\exp}$ is a solid cone. It need not be convex; nor is it necessarily open.

Since $\mathbf{r}(\mathcal{A}_{\mathbf{m}}) = \mathcal{R}_{\mathbf{m}}^{\exp}$, it is important to understand the structure of $\mathcal{A}_{\mathbf{m}}$. Its interior has a rather simple expression:

$$\operatorname{int} \mathcal{A}_{\mathbf{m}} = \left\{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_N^T \mathbf{m}_N(v) < 0 \quad \forall\, v \neq 0\right\} = \{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_N \in \operatorname{int} A_{\mathbf{m}_N}\}\ , \tag{43}$$

where

$$A_{\mathbf{m}_j} \equiv \left\{\boldsymbol{\alpha}_j \in \mathbb{R}^{n_j} : \boldsymbol{\alpha}_j^T \mathbf{m}_j \leq 0\right\}\ , \quad 1 \leq j \leq N\,, \tag{44}$$

is a proper cone for $j$ even. (It can be checked that Condition III of Section 2.1 is equivalent to $\operatorname{int} \mathcal{A}_{\mathbf{m}}$ being non-empty.) If $\boldsymbol{\alpha} \in \operatorname{int} \mathcal{A}_{\mathbf{m}}$, then the behavior of $p = \boldsymbol{\alpha}^T \mathbf{m}$ is dominated for large $|v|$ by the homogeneous component $p_N = \boldsymbol{\alpha}_N^T \mathbf{m}_N$, and

$$\lim_{|v| \to \infty} p(v) = \lim_{|v| \to \infty} p_N(v) = \lim_{|v| \to \infty} |v|^N p_N\left(v/|v|\right) = -\infty\,. \tag{45}$$

For such $\boldsymbol{\alpha}$, $G_{\boldsymbol{\alpha}}$ decays exponentially and the moments $\mathbf{r}(\boldsymbol{\alpha})$ are finite.

From (43), one can easily show that

$$\operatorname{cl} \mathcal{A}_{\mathbf{m}} = \{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}\} \qquad \text{and} \qquad \partial \mathcal{A}_{\mathbf{m}} \subset \{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_N \in \partial A_{\mathbf{m}_N}\}\ . \tag{46}$$

Even so, the boundary component $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$ is, in general, very complicated. If $\boldsymbol{\alpha} \in \partial \mathcal{A}_{\mathbf{m}}$, then $\boldsymbol{\alpha}_N \in \partial A_{\mathbf{m}_N}$ and $p_N(\lambda v) = 0$ for some $v \neq 0$ and all $\lambda \in \mathbb{R}$, and it may be that there are unbounded sequences $\{v_i\}_{i=1}^{\infty}$ such that $\lim_{i \to \infty} p(v_i) > -\infty$. In such cases, it is not clear whether the moments $\mathbf{r}(\boldsymbol{\alpha})$ are finite, i.e., whether $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. We will revisit this issue in Section 6.3. For now, we turn our attention to the entropy minimization problem (20).

# 3 Entropy Minimization

Most of this section reproduces and discusses the main result from [Sch04]. In this setting, we then state Theorem 9, which is the basis for our new results.

### 3.1 The Entropy Functional

Recall that the strictly convex *entropy functional* $\mathcal{H} : \mathbb{F}_{\mathbf{m}} \mapsto \mathbb{R} \cup \{\infty\}$ is given by

$$\mathcal{H}(g) \equiv \langle g \log g - g \rangle \,. \tag{47}$$

By employing the convention $0 \log 0 = 0$—which is consistent with the fact that $\lim_{z \to 0} z \log z = 0$—one can make sense of the integrand for those values of $v$ where $g(v) = 0$. There are functions $g \in \mathbb{F}_{\mathbf{m}}$ such that $\mathcal{H}(g) = +\infty$; however, in order for $\mathcal{H}(g)$ to be well-defined, the negative contribution to the integral, $\mathcal{H}^-(g)$, must be finite. We show this is indeed the case.

**Lemma 2** *For each $g \in \mathbb{F}_{\mathbf{m}}$, let $K_g = \left\{ v \in \mathbb{R}^d : g\,(v) \log(g\,(v)) - g\,(v) < 0 \right\}$. Then*

$$\mathcal{H}^-\,(g) \equiv -\int_{K_g} \left(g(v)\log(g(v)) - g(v)\right)\,dv \leq \int_{\mathbb{R}^d} \left(|v|^2 g(v) + e^{-|v|^2}\right)\,dv\,. \tag{48}$$

*In particular, $\mathcal{H}^-\,(g)$ is finite.*

The proof of this lemma is based on Young's inequality: for all $z, y > 0$,

$$z \log z - z \geq y \log y - y + (\log y)\,(z - y) \tag{49}$$

or, equivalently,

$$z \log z - z \geq z \log y - y\,. \tag{50}$$

These two inequalities follow immediately from convexity of the mapping $z \mapsto z \log z - z$.
**Proof.** Letting $z = g\,(v)$ and $y = e^{-|v|^2}$ in (50) gives—after integration over $K_g$,

$$\mathcal{H}^-\,(g) \leq \int_{K_g} \left(|v|^2 g(v) + e^{-|v|^2}\right)\,dv \leq \int_{\mathbb{R}^d} \left(|v|^2 g(v) + e^{-|v|^2}\right)\,dv\,, \tag{51}$$

which is finite since $|v|^2 \in \mathbb{M}$. ■

### 3.2 Schneider's Problem

Given $\boldsymbol{\rho} = (\boldsymbol{\rho}_0, \ldots, \boldsymbol{\rho}_N) \in \mathcal{R}_{\mathbf{m}}$, we seek a solution of (20), where the relation $\langle \mathbf{m} g \rangle \preceq^\circ \boldsymbol{\rho}$ (or, equivalently, $\boldsymbol{\rho} \succeq^\circ \langle \mathbf{m} g \rangle$) is a shorthand for

$$\langle \mathbf{m}_j g \rangle = \boldsymbol{\rho}_j\,,\ 0 \leq j \leq N-1\,, \tag{52a}$$

$$\langle \mathbf{m}_N g \rangle \leq_{A^\circ_{\mathbf{m}_N}} \boldsymbol{\rho}_N\,, \tag{52b}$$

and $A^\circ_{\mathbf{m}_N} \equiv (A_{\mathbf{m}_N})^\circ$. Note that (52b) means that

$$\boldsymbol{\alpha}_N^T \langle \mathbf{m}_N g \rangle \leq \boldsymbol{\alpha}_N^T \boldsymbol{\rho}_N \quad \text{whenever} \quad \boldsymbol{\alpha}_N^T \mathbf{m}_N \geq 0\,. \tag{53}$$

The components of $\langle \mathbf{m}_j g \rangle$, $0 \leq j < N$, will be referred to as *lower-order moments*, and the components of $\langle \mathbf{m}_N g \rangle$ will be referred to as *higher-order moments.*

The main result from [Sch04] concerning the minimization problem with relaxed constraints (20) is the following theorem.

**Theorem 3 (Schneider [Sch04])** *For any $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, there is a unique minimizer for the minimization problem (20). This minimizer has the form $G_{\boldsymbol{\alpha}}$ given by (38), where $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. Conversely, for each $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$,*

$$\mathcal{H}(G_{\boldsymbol{\alpha}}) = \min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^\circ \mathbf{r}(\boldsymbol{\alpha}) \right\}\,, \tag{54}$$

*where $\mathbf{r}(\boldsymbol{\alpha})$ is given by (41). Moreover, $G_{\boldsymbol{\alpha}}$ also satisfies the equality constraint problem (14) with $\boldsymbol{\rho} = \mathbf{r}(\boldsymbol{\alpha})$.*

We define $\mathbf{a} : \mathcal{R}_{\mathbf{m}} \to \mathcal{A}_{\mathbf{m}}$ as the mapping which assigns to $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ the vector $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ such that $G_{\boldsymbol{\alpha}}$ solves (20)—that is,

$$G_{\mathbf{a}(\boldsymbol{\rho})} \equiv \arg\min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho} \right\} . \tag{55}$$

The converse statement of Theorem 3 implies the following.

**Corollary 4** *Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. Then $G_{\mathbf{a}(\boldsymbol{\rho})}$ is the unique minimizer of the entropy minimization problem with equality constraints (14).*

To help the reader's intuition, we provide a proof for Theorem 3 with the use of three lemmas. The first lemma is used to prove existence of a minimizer for the minimization problem with relaxed constraints (20), and the first item of this lemma is a direct consequence of Lemma 2.

**Lemma 5 (Schneider [Sch04])** *The entropy functional $\mathcal{H}$ satisfies the following:*

1. *$\mathcal{H}(g) > -\infty$ for all $g \in \mathbb{F}_{\mathbf{m}}$.*
2. *$\mathcal{H}$ is convex, lower semi-continuous with respect to the norm $||g||_{L^1_{\mathbf{m}}(\mathbb{R}^d)} \equiv \langle |\mathbf{m} g| \rangle$.*
3. *Subsets of $\mathbb{F}_{\mathbf{m}}$ which are bounded in the $L^1_{\mathbf{m}}(\mathbb{R}^d)$ topology and on which $\mathcal{H}$ is bounded are weakly relatively compact in $L^1(\mathbb{R}^d)$.*

The second lemma is a statement about the constraint set

$$C_{\mathbf{m}}(\boldsymbol{\rho}) \equiv \left\{ g \in \mathbb{F}_{\mathbf{m}} : \langle \mathbf{m} g \rangle \preceq^{\circ} \boldsymbol{\rho} \right\} . \tag{56}$$

**Lemma 6** *For each $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the set $C_{\mathbf{m}}(\boldsymbol{\rho})$ is closed in the weak-$L^1$ topology.*

**Proof.** Let $\{g_k\}_{k=1}^{\infty}$ be any sequence in $C_{\mathbf{m}}(\boldsymbol{\rho})$ that converges in weak-$L^1(\mathbb{R}^d)$ to a function $g_*$. For the highest order moments, Fatou's Lemma implies that if $\boldsymbol{\alpha}_N^T \mathbf{m}_N \geq 0$, then

$$\boldsymbol{\alpha}_N^T \langle \mathbf{m}_N g_* \rangle \leq \lim_{k \to \infty} \boldsymbol{\alpha}_N^T \langle \mathbf{m}_N g_i \rangle = \boldsymbol{\alpha}_N^T \boldsymbol{\rho}_N . \tag{57}$$

For $j < N$, more can be said. We break up the integral $\langle \mathbf{m}_j g_k \rangle$ into two pieces:

$$\boldsymbol{\rho}_j = \langle \mathbf{m}_j g_k \rangle = \int_{|v|<R} \mathbf{m}_j g_k \, dv + \int_{|v|>R} \mathbf{m}_j g_k \, dv , \tag{58}$$

where $R > 0$ is an arbitrary constant. For the first term in (58), weak-$L^1(\mathbb{R}^d)$ convergence implies that

$$\int_{|v|<R} \mathbf{m}_j g_k \, dv \overset{k\to\infty}{\rightarrow} \int_{|v|<R} \mathbf{m}_j g_* \, dv , \quad 0 \leq j \leq N . \tag{59}$$

Meanwhile, in the second term

$$\frac{|\mathbf{m}_j|}{|v|^N} < \frac{C_0}{R^{N-j}} , \quad |v| > R , \ 0 \leq j < N \tag{60}$$

for some constant $C_0$ that is independent of $R$. Hence

$$\int_{|v|>R} \mathbf{m}_j g_k \, dv \leq \int_{|v|>R} \frac{|\mathbf{m}_j|}{|v|^N} |v|^N g_k \, dv < \frac{C_0}{R^{N-j}} \sup_k \left| \left\langle |v|^N g_k \right\rangle \right| . \tag{61}$$

Since $\{g_k\}_{k=1}^{\infty} \subset C_{\mathbf{m}}$, the sequence $\left\{ \left\langle |v|^N g_k \right\rangle \right\}_{k=1}^{\infty}$ is uniformly bounded in $k$, and it follows from (59) and (61) that

$$\lim_{k\to\infty} \left| \boldsymbol{\rho}_j - \langle \mathbf{m}_j g_k \rangle \right| \leq \lim_{k\to\infty} \int_{|v|>R} |\mathbf{m}_j g_k - \mathbf{m}_j g_*| \, dv \leq \frac{C_0}{R^{N-j}} \left( \sup_k \left\langle |v|^N g_k \right\rangle + \left\langle |v|^N g_* \right\rangle \right) < \frac{C_1}{R^{N-j}} \tag{62}$$

for some constant $C_1 > 0$ that is independent of $R$. Since $R$ can be arbitrarily large, we conclude that $\langle \mathbf{m}_j g_* \rangle = \boldsymbol{\rho}_j$ for all $j < N$. Hence $g_* \in C_{\mathbf{m}}(\boldsymbol{\rho})$. ■

The third lemma is used to prove the form of the minimizer. For any bounded measurable set $K \subset \mathbb{R}^d$ and any locally integrable function $g$, let

$$\langle g \rangle_K \equiv \int_K g(v)\,dv \qquad \text{and} \qquad \mathbb{F}^K_{\mathbf{m}} \equiv \left\{ g \in L_1\left(\mathbb{R}^d\right) : g \gneqq 0 \text{ and } \langle |m_i g| \rangle_K < \infty,\ i = 0, \ldots, n-1 \right\} . \tag{63}$$

On $\mathbb{F}^K_{\mathbf{m}}$, we define

$$\mathcal{H}^K(g) \equiv \langle g \log g - g \rangle_K \ , \tag{64}$$

As with $\mathcal{H}$, the negative contribution to $\mathcal{H}^K$ must be finite (see Lemma 2) in order for it to be well-defined and restricting $\mathrm{Dom}(\mathcal{H}^K)$ to $\mathbb{F}^K_{\mathbf{m}}$ ensures that this will be the case.

**Lemma 7 (Junk [Jun00, Jun98], Borwein-Lewis [BL91])** *For any bounded set $K \subset \mathbb{R}^d$ and any function $f \in \mathbb{F}^K_{\mathbf{m}}$, the problem*

$$\min_{g \in \mathbb{F}^K_{\mathbf{m}}} \left\{ \mathcal{H}^K(g) : \langle \mathbf{m} g \rangle_K = \langle \mathbf{m} f \rangle_K \right\} \tag{65}$$

*has a unique minimizer, which takes the form $G_{\boldsymbol{\alpha}}$ for some $\boldsymbol{\alpha} \in \mathbb{R}^n$.*

**Proof of Theorem 3.** The proof has three parts.

1. **Existence and uniqueness**. Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$. By Theorem 1, the set

$$C_{\mathbf{m}}(\boldsymbol{\rho}) \equiv \{\, g \in \mathbb{F}_{\mathbf{m}} : \langle \mathbf{m} g \rangle \preceq^\circ \boldsymbol{\rho} \} \tag{66}$$

contains bounded functions with compact support. Because such functions have finite entropy, the subset of $C_{\mathbf{m}}(\boldsymbol{\rho})$ on which $\mathcal{H}$ is finite is non-empty. Moreover, by Lemma 2, $\mathcal{H}$ is bounded below on $C_{\mathbf{m}}(\boldsymbol{\rho})$. Hence $h_{\mathrm{S}}(\boldsymbol{\rho})$ is finite, and there exists $\{g_i\}_{i=1}^\infty \subset C_{\mathbf{m}}(\boldsymbol{\rho})$ such that $\mathcal{H}(g_i) \to h_{\mathrm{S}}(\boldsymbol{\rho})$. By Lemma 5, there is a subsequence $\{g_{i_k}\}_{k=1}^\infty$ that converges in weak-$L^1$ to a function $\hat{g}_{\boldsymbol{\rho}}$, and since $C_{\mathbf{m}}(\boldsymbol{\rho})$ is closed (Lemma 6), $\hat{g}_{\boldsymbol{\rho}} \in C_{\mathbf{m}}(\boldsymbol{\rho})$. Finally, since $\mathcal{H}$ is lower semi-continuous (Lemma 5),

$$\mathcal{H}(\hat{g}_{\boldsymbol{\rho}}) \le \lim_{k \to \infty} \mathcal{H}(g_{i_k}) = h_{\mathrm{S}}(\boldsymbol{\rho})\,. \tag{67}$$

Thus $\hat{g}_{\boldsymbol{\rho}}$ attains the minimum in (20), and strict convexity of $\mathcal{H}$ implies that the minimizer is unique.

2. **Form of the minimizer**. According to Lemma 7, for any bounded set $K \subset \mathbb{R}^d$

$$\min \left\{ \mathcal{H}^K(g) : \langle \mathbf{m} g \rangle_K = \langle \mathbf{m} \hat{g}_{\boldsymbol{\rho}} \rangle_K \right\} \tag{68}$$

has a solution of the form $G_{\boldsymbol{\alpha}}$. We conclude then that $\hat{g}_{\boldsymbol{\rho}} = G_{\boldsymbol{\alpha}}$ on $K$; otherwise, the function

$$g^*_{\boldsymbol{\rho}}(v) = \left\{ \begin{array}{cc} G_{\boldsymbol{\alpha}}(v) & v \in K \\ \hat{g}_{\boldsymbol{\rho}} & v \notin K \end{array} \right. \tag{69}$$

would satisfy $\mathcal{H}(g^*_{\boldsymbol{\rho}}) \le \mathcal{H}(\hat{g}_{\boldsymbol{\rho}})$, an obvious contradiction. Since $K$ is arbitrary, we conclude that $\hat{g}_{\boldsymbol{\rho}} = G_{\mathbf{a}}$ and, in order to satisfy to constraints in (20), that $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$.

3. **Converse statement**. Applying Young's Inequality (49) to $z = g$ and $y = G_{\boldsymbol{\alpha}}$ and integrating over all velocity space gives

$$\mathcal{H}(g) \ge \mathcal{H}(G_{\boldsymbol{\alpha}}) + \boldsymbol{\alpha}^T \langle \mathbf{m}(g - G_{\boldsymbol{\alpha}}) \rangle \ . \tag{70}$$

By hypothesis, $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ which implies $\boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}$. Thus if $g \in \mathbb{F}_{\mathbf{m}}$ satisfies $\langle \mathbf{m} g \rangle \preceq^\circ \langle \mathbf{m} G_{\boldsymbol{\alpha}} \rangle$, then according to (52) and (53),

$$\boldsymbol{\alpha}^T \langle \mathbf{m}(g - G_{\boldsymbol{\alpha}}) \rangle = \sum_{j=1}^N \boldsymbol{\alpha}_j^T \langle \mathbf{m}_j (g - G_{\boldsymbol{\alpha}}) \rangle = \boldsymbol{\alpha}_N^T \langle \mathbf{m}_N (g - G_{\boldsymbol{\alpha}}) \rangle \ge 0\,. \tag{71}$$

Thus, from (70), $\mathcal{H}(g) \ge \mathcal{H}(G_{\boldsymbol{\alpha}})$. This concludes the proof.

■

The existence part of this proof provides some intuition as to why the optimization problem with equality constraints (14) may not always have a minimizer. Suppose that the minimizing sequence $\{g_{i_k}\}_{k=1}^{\infty}$ were restricted to the set

$$C^0_{\mathbf{m}}(\boldsymbol{\rho}) \equiv \{g \in \mathbb{F}_{\mathbf{m}} : \langle \mathbf{m} g \rangle = \boldsymbol{\rho}\} \tag{72}$$

rather than merely lying in $C_{\mathbf{m}}(\boldsymbol{\rho})$. Then $\{g_{i_k}\}_{k=1}^{\infty}$ would still converge in the weak-$L^1(\mathbb{R}^d)$ topology to $\hat{g}_{\boldsymbol{\rho}}$, with $\langle \mathbf{m}_j \hat{g}_{\boldsymbol{\rho}} \rangle = \boldsymbol{\rho}_j$ for $j < N$. However, the bound in (61) does not help when $j = N$. Hence there is no way to ensure that $\langle \mathbf{m}_N \hat{g}_{\boldsymbol{\rho}} \rangle = \boldsymbol{\rho}_N$—only that $\langle \mathbf{m}_N \hat{g}_{\boldsymbol{\rho}} \rangle \leq_{\mathcal{A}^{\circ}_{\mathbf{m}}} \boldsymbol{\rho}_N$. This is precisely why Schneider introduces the inequality constraint: $C_{\mathbf{m}}(\boldsymbol{\rho})$ is closed in the weak-$L^1$ topology whereas $C^0_{\mathbf{m}}(\boldsymbol{\rho})$ is not.

Such behavior begs the following question: For what values of $\boldsymbol{\rho}$ does a minimizing sequence for (14) *not* converge inside $C^0_{\mathbf{m}}(\boldsymbol{\rho})$? These will be the densities which make up the set $\mathcal{D}_{\mathbf{m}}$. In [Sch04], Schneider attempts to address this question in the following corollary to Theorem 3.

**Corollary 8 ( [Sch04])** *Given $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the minimizer in (14) exists if and only if there is no function of the form $G_{\boldsymbol{\alpha}}$ in $C_{\mathbf{m}}(\boldsymbol{\rho}) \backslash C^0_{\mathbf{m}}(\boldsymbol{\rho})$.*

Unfortunately, this result provides little understanding of the geometry of $\mathcal{D}_{\mathbf{m}}$. A more insightful point of view is given by the following theorem.

**Theorem 9** *Given $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the minimization problem with equality constraints (14) has a minimizer if and only if $\boldsymbol{\rho} \in \mathcal{R}^{\exp}_{\mathbf{m}}$. In other words,*

$$\mathcal{D}_{\mathbf{m}} = \mathcal{R}_{\mathbf{m}} \backslash \mathcal{R}^{\exp}_{\mathbf{m}} \,. \tag{73}$$

**Proof.** The "if" part of this theorem is just Corollary 4. The "only if" part will be proved at the end of w 4.3. ■

An immediate consequence of Theorem 9 is that $\mathcal{D}_{\mathbf{m}}$ is a cone. However, the essential point of the theorem is that when $\mathcal{D}_{\mathbf{m}}$ is non-empty, there are realizable densities $\boldsymbol{\rho}$ that *cannot* be realized by a functions of the form $G_{\boldsymbol{\alpha}}$. In other words, $\boldsymbol{\rho} \notin \mathcal{R}^{\exp}_{\mathbf{m}}$ even though $\mathbf{a}(\boldsymbol{\rho}) \in \mathcal{A}_{\mathbf{m}}$. It is this idea which lays the foundation for the results in [Jun98, Jun00], where a description of $\mathcal{D}_{\mathbf{m}}$ is given for the case $\mathbf{m}_N = |v|^N$. Theorem 9 will also be the basis for the new results of this paper. However, for a general admissible space $\mathbb{M}$, we will need to formulate the dual for relaxed constraint problem (20) and derive complementary slackness conditions in order to find a useful geometric description for $\mathcal{D}_{\mathbf{m}}$. In the process, we will recover and extend many of the results from [Jun98, Jun00] and [Sch04].

# 4 Dual Formulation

Because $\mathcal{H}$ is convex on $\mathbb{F}_{\mathbf{m}}$ and the constraints in (20) are linear, it is reasonable to apply a dual treatment to the relaxed-constraint problem, e.g. [Lue69, BV04, BNO03]. In this section, we prove two important duality theorems and the complementary slackness conditions that accompany them. We also give an alternate proof of the form of the minimizer in Theorem 3 and a proof of the "only if" part of Theorem 9.

## 4.1 The Dual Function

We define the Lagrangian function $\mathcal{L} : \mathbb{F}_{\mathbf{m}} \times \mathbb{R}^n \times \mathcal{R}_{\mathbf{m}} \to \mathbb{R} \cup \{\infty\}$ associated to (20), by

$$\mathcal{L}\left(g, \boldsymbol{\alpha}, \boldsymbol{\rho}\right) \equiv \mathcal{H}(g) + \boldsymbol{\alpha}^T \left(\boldsymbol{\rho} - \langle \mathbf{m} g \rangle\right) \tag{74}$$

and the dual function $\psi : \mathbb{R}^n \times \mathcal{R}_{\mathbf{m}} \to \mathbb{R} \cup \{-\infty\}$ by

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) \equiv \inf_{g \in \mathbb{F}_{\mathbf{m}}} \mathcal{L}\left(g, \boldsymbol{\alpha}, \boldsymbol{\rho}\right). \tag{75}$$

The dual function is closely related to the density potential $h^*$. In fact, we have the following.

**Theorem 10** *For all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ and $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$,*

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) = \mathcal{L}\left(G_{\boldsymbol{\alpha}}, \boldsymbol{\alpha}, \boldsymbol{\rho}\right) = \boldsymbol{\alpha}^T \boldsymbol{\rho} - h^*(\boldsymbol{\alpha})\,. \tag{76}$$

**Proof.** We apply Young's Inequality (50) and make the identification $z = g$ and $y = G_{\boldsymbol{\alpha}}$ to derive the point-wise inequality

$$(g \log g - g) - \boldsymbol{\alpha}^T \mathbf{m} g \geq -G_{\boldsymbol{\alpha}}. \tag{77}$$

Integration of (77) over $\mathbb{R}^d$ and addition of $\boldsymbol{\alpha}^T \boldsymbol{\rho}$ to both sides gives a lower bound on $\mathcal{L}$ and hence $\psi$:

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) \geq \boldsymbol{\alpha}^T \boldsymbol{\rho} - h^*(\boldsymbol{\alpha})\,. \tag{78}$$

For $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$, the definitions of $\mathcal{H}$, $G_{\boldsymbol{\alpha}}$, and $h^*$ (given in (47), (38), and (39), respectively) imply that

$$\mathcal{H}\left(G_{\boldsymbol{\alpha}}\right) = \boldsymbol{\alpha}^T \left\langle \mathbf{m} G_{\boldsymbol{\alpha}} \right\rangle - \left\langle G_{\boldsymbol{\alpha}} \right\rangle = \boldsymbol{\alpha}^T \left\langle \mathbf{m} G_{\boldsymbol{\alpha}} \right\rangle - h^*(\boldsymbol{\alpha})\,. \tag{79}$$

Thus by (74),

$$\mathcal{L}\left(G_{\boldsymbol{\alpha}}, \boldsymbol{\alpha}, \boldsymbol{\rho}\right) = \boldsymbol{\alpha}^T \boldsymbol{\rho} - h^*(\boldsymbol{\alpha})\,, \tag{80}$$

so that, from (75),

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) \leq \boldsymbol{\alpha}^T \boldsymbol{\rho} - h^*(\boldsymbol{\alpha})\,. \tag{81}$$

Together (78), (80), and (81) imply (76). ■

## 4.2 Smoothness Properties of the Dual Function

The following smoothness properties of $\psi$ will be used throughout the remainder of the paper.

**Theorem 11** *Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$. Then*

1. *$\psi(\cdot, \boldsymbol{\rho})$ is strictly concave on $\mathcal{A}_{\mathbf{m}}$ and infinitely Fréchet differentiable on* int $\mathcal{A}_{\mathbf{m}}$*, with derivatives*

$$\frac{\partial \psi}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}, \boldsymbol{\rho}) = \boldsymbol{\rho} - \mathbf{r}(\boldsymbol{\alpha})\,, \tag{82a}$$

$$\frac{\partial^{(i)} \psi}{\partial \boldsymbol{\alpha}^{(i)}}(\boldsymbol{\alpha}, \boldsymbol{\rho}) = -\left\langle \mathbf{m}^{\vee(i)} G_{\boldsymbol{\alpha}} \right\rangle\,, \qquad i > 1\,, \tag{82b}$$

*where $\mathbf{m}^{\vee(i)}$ is the $i^{th}$ tensor power of $\mathbf{m}$.*[4]

2. *For any $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{A}_{\mathbf{m}}$, the function*

$$\phi(\tau) \equiv \psi(\tau\boldsymbol{\alpha} + (1-\tau)\boldsymbol{\beta}, \boldsymbol{\rho}) \tag{83}$$

*is twice differentiable at each $\tau \in [0,1]$ (one-sided at endpoints) with derivatives*

$$\phi'(\tau) = (\boldsymbol{\alpha} - \boldsymbol{\beta})^T \left[\boldsymbol{\rho} - \mathbf{r}(\tau\boldsymbol{\alpha} + (1-\tau)\boldsymbol{\beta})\right]\,, \tag{84a}$$

$$\phi''(\tau) = -\left\langle \left((\boldsymbol{\alpha} - \boldsymbol{\beta})^T \mathbf{m}\right)^2 G_{\tau\boldsymbol{\alpha}+(1-\tau)\boldsymbol{\beta}} \right\rangle\,. \tag{84b}$$

*In particular, the function $\phi'(\tau)$ is a decreasing function of $\tau$.*

3. *The function $\psi(\cdot, \boldsymbol{\rho})$ is upper semi-continuous on $\mathcal{A}_{\mathbf{m}}$.*

[4] The tensor power of a symmetric tensor $S$ is defined recursively. For $n > 1$, $S^{\vee(n)} \equiv S \vee S^{\vee(n-1)}$ while $S^{\vee(1)} \equiv S$.

**Proof.** For the proofs of the first two statements above, we refer the reader to Lemmas 5.1 and 5.2 in [Jun00] along with a few comments. First, the lemmas in [Jun00] refer to $h^*$ rather than $\psi(\cdot,\boldsymbol{\rho})$. This makes little difference since the two functions differ only by a linear factor (see Theorem 10). Also, the proofs in [Jun00] are constructed specifically for the special case when $m_N = |v|^N$; however, modifications to the general setting are straight-forward. To prove the third statement we simply invoke Fatou's Lemma. Given a sequence $\left\{\boldsymbol{\alpha}_{(i)}\right\}_{i=1}^{\infty} \subset \mathcal{A}_{\mathbf{m}}$ with limit $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$,

$$\langle G_{\boldsymbol{\alpha}} \rangle \leq \lim_{i\to\infty} \left\langle G_{\boldsymbol{\alpha}_{(i)}} \right\rangle . \tag{85}$$

Hence $\lim_{i\to\infty} \psi(\boldsymbol{\alpha}_{(i)},\boldsymbol{\rho}) \leq \psi(\boldsymbol{\alpha},\boldsymbol{\rho})$. ■

**Corollary 12** *For all $\boldsymbol{\alpha} \in \operatorname{int}\mathcal{A}_{\mathbf{m}}$, $h^*_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = \mathbf{r}(\boldsymbol{\alpha})$ and $h^*_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = \left\langle \mathbf{m}\mathbf{m}^T G_{\boldsymbol{\alpha}} \right\rangle$, which is positive-definite on $\boldsymbol{\alpha} \in \operatorname{int}\mathcal{A}_{\mathbf{m}}$.*

Several remarks should be made concerning Theorem 11. First, statement 1 implies statement 2, but only for $\boldsymbol{\alpha},\boldsymbol{\beta} \in \operatorname{int}\mathcal{A}_{\mathbf{m}}$. Second, for $\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, $\phi''$ need not be continuous and higher derivatives may not exist . Finally, in spite of the smoothness properties given by Theorem 11, the dual function need not even be continuous on $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$. Indeed, given a sequence $\left\{\boldsymbol{\alpha}_{(i)}\right\}_{i=1}^{\infty} \in \mathcal{A}_{\mathbf{m}}$ with limit $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, it is possible that $h^*(\boldsymbol{\alpha}) < \lim_{i\to\infty} h^*(\boldsymbol{\alpha}_{(i)})$. As an example, consider the one-dimensional case ($d=1$) when $\mathbf{m} = \left(1,v,v^2,v^3,v^4\right)^T$. This case has been studied in detail in [Jun98]. Given the following five points in the $(v,w)$ plane:

$$(v_0,w_0) = (0,0) \ , \qquad (v_1,w_1) = (1,0) \ , \qquad (v_2,w_2) = \left(i,-i^2\right) \ ,$$
$$(v_3,w_3) = (2i,i) \ , \qquad (v_4,w_4) = (2i+1,0) \ ,$$

the unique degree four polynomial interpolating these points is

$$p_i(v) = \boldsymbol{\alpha}_{(i)}^T \mathbf{m}(v) = \sum_{j=0}^{4} \alpha_{(i)j} v^j \ , \tag{86}$$

where

$$\alpha_{0(i)} = 0 \, , \qquad \alpha_{1(i)} = \frac{2i+1}{4i-2} + \frac{4i^2+2i}{i^2-1} \, , \qquad \alpha_{2(i)} = -\frac{4i^2+6i+1}{i^2-1} - \frac{2i^2+4i+1}{4i^2-2i} \, ,$$

$$\alpha_{3(i)} = \frac{4i+2}{i^2-1} + \frac{3i+2}{4i^2-2i} \, , \qquad \alpha_{4(i)} = -\frac{1}{i^2-1} - \frac{1}{4i^2-2i} \, .$$

(The notation $\boldsymbol{\alpha}_{(i)}$ denotes a sequence of vectors rather than the usual notation $\boldsymbol{\alpha}_i$, which denotes the components of a single vector $\boldsymbol{\alpha}$ corresponding to polynomials of degree $i$). As $i \to \infty$,

$$\boldsymbol{\alpha}_{(i)} \to \boldsymbol{\alpha}_* = \left(0, \frac{9}{2}, -\frac{9}{2}, 0, 0\right)^T \quad \text{and} \quad G_{\boldsymbol{\alpha}_*} = \exp\left(-\frac{9}{2}v^2 + \frac{9}{2}v\right) \, . \tag{87}$$

The density potential $h^*(\boldsymbol{\alpha}_*)$ moments $\mathbf{r}(\boldsymbol{\alpha}_*)$ are finite. Therefore $\boldsymbol{\alpha}_* \in \mathcal{A}_{\mathbf{m}}$, but clearly $\boldsymbol{\alpha}_* \notin \operatorname{int}\mathcal{A}_{\mathbf{m}}$. Moreover, one may readily check that $p_i$ is positive and concave on the interval $[2i,2i+1]$ and hence,

$$h^*(\boldsymbol{\alpha}_{(i)}) = \left\langle G_{\boldsymbol{\alpha}_{(i)}} \right\rangle > \int_{2i}^{2i+1} e^{p_i(v)}\, dv > \int_{2i}^{2i+1} (1+p_i(v))\, dv > 1 + \frac{i}{2} \to \infty \qquad \text{as } i \to \infty \, . \tag{88}$$

Note that the second inequality above follows from the fact that $e^x > 1+x$, while the concavity of $p_i$ on $[2i,2i+1]$ implies that the graph of $p_i$ lies above the line segment $\ell$ joining the points $(2i,i)$ and $(2i+1,0)$ in the $(v,w)$ plane. Therefore the integral of $p_i$ over $[2i,2i+1]$ is bounded below by the area of the triangle formed by $\ell$, the $v$-axis, and the line $\{v=2i\}$. The area of this triangle is $i/2$. A similar argument shows that, for any $j \geq 0$, $\left\langle |v|^j G_{\boldsymbol{\alpha}_{(i)}} \right\rangle \to \infty$ as $i \to \infty$ while $\left\langle v^j G_{\boldsymbol{\alpha}_*} \right\rangle$ is finite.

The reason that $\psi(\cdot,\boldsymbol{\rho})$ is discontinuous at the boundary of $\mathcal{A}_{\mathbf{m}}$ is the same reason that the minimization problem (14) with equality constraints fails: because mass at the tails of the functions escapes as $i \to \infty$. In the example above, this is precisely what happens to the mass of $G_{\boldsymbol{\alpha}_{(i)}}$ that is supported on the interval $[2i, 2i+1]$. The same thing occurs with the minimizing sequence $\{g_{i_k}\}_{k=1}^{\infty}$ in the proof of Theorem 3. The difference is that, for $\{g_{i_k}\}_{k=1}^{\infty}$, only the highest moments fail to converge in the minimizing sequence, whereas none of moments in this example converge. The reason for this difference is that the moments $\langle \mathbf{m} g_{i_k}\rangle$ are all bounded. The moments of $\{G_{\boldsymbol{\alpha}_{(i)}}\}_{i=1}^{\infty}$ *would* converge if higher order moments were controlled in some way. Controlling the moments is, in effect, the same as requiring $\boldsymbol{\alpha}_i \to \boldsymbol{\alpha}_*$ along a specified path. In fact, we will see at the very end of Section 5.4 that the the map $\boldsymbol{\rho} \longmapsto \psi\left(\mathbf{a}(\boldsymbol{\rho}),\boldsymbol{\rho}\right)$ is continuous on $\mathcal{R}_{\mathbf{m}}$.

## 4.3 Duality Theorems

The main results of this subsection are based on the following strong duality theorem where, for a given cone $C$, the notations "$\leq_C$" and "$\geq_C$" are defined in the last paragraph of Section 2.2.

**Theorem 13 ( [Lue69])** *Consider the problem*

$$\begin{array}{ll} \textit{minimize} & f_0(x) \\ \textit{subject to} & f_i(x) \leq_{K_i} 0,\ i=1,\dots,m\,; \quad Ax=b\,, \end{array} \tag{89}$$

*where the functions $f_0,\dots,f_m : \mathbb{X} \to \mathbb{R}\cup +\infty$ are convex over a vector space $\mathbb{X}$, $A:\mathbb{X}\to\mathbb{R}^k$ is a linear mapping, $b\in\mathbb{R}^k$, and each $K_i$ is a proper cone for $i=1,\dots,m$. Let $D$ be the intersection of the domains of $f_0,\dots,f_m$ (i.e., $D$ is a convex set over which each $f_i$ is finite). Suppose there exists $\tilde{x}\in D$ with $f_i(\tilde{x})<0$, $i=1,..,m$, and $A\tilde{x}=b$. Further suppose that the set $\{Ax-b : x\in D\}$ contains a neighborhood of the origin. Then strong duality holds, i.e.,*

$$\inf\{f_0(x) : f_i(x)\leq_{K_i} 0,\ i=1,\dots,m\,;\ Ax=b\} = \sup_{\substack{\lambda_i \geq_{K_i^\circ} 0 \\ \nu\in\mathbb{R}^k}} \inf_{x\in D} \left\{f_0(x) + \Sigma_{i=1}^m \lambda_i f_i(x) + \nu^T(Ax-b)\right\} \tag{90}$$

*and the dual optimal value is attained whenever it is not $-\infty$.*

Theorem 13 follows from [Lue69, Exercise 8.7] and can be proven using arguments found in [Lue69, Chapter 8]. It can also be proven along the lines of similar results found in [BV04, Sections 5.3.2 and 5.9.1]. However, whereas those results require the existence of some $\tilde{x}$ in the relative interior of $D$, Theorem 13 requires only that $\tilde{x}\in D$. A side benefit of this is that there is no need to specify a topology on $\mathbb{X}$. In return, our condition that $\{Ax-b : x\in D\}$ contains a neighborhood of the origin is not present in the statements in [BV04].

To prove Theorem 13, one may repeat the arguments found in [BV04, Section 5.3.2] with the notation "$\leq$" changed to curly "$\preceq$". The only difference from that proof is in the contradiction argument showing (in the notation of [BV04]) that $\mu=0$ is not possible. The proof in [BV04] first shows, with logic that remains valid under our weaker assumptions on $\tilde{x}$, that if $\mu=0$, then there must exist $\nu\neq 0$ such that $\nu^T(Ax-b)\geq 0$ for all $x\in D$. At that point, our assumption that $\{Ax-b : x\in D\}$ contains a neighborhood of the origin immediately implies that $\nu=0$, which yields the requisite contradiction.

The statement of Theorem 13 is of much interest in the present context for two reasons: (i) our primal decision variable $g$ lies in an infinite-dimensional vector space and (ii) it is not straight-forward to show that the relative interior condition on $\tilde{x}$ (or in our case $\tilde{g}$) actually applies. (However, see [BL91, Definition 2.1], where the authors introduce the notion of *pseudo relative interior.*) On the other hand, that our additional condition on $\{Ax-b : x\in D\}$ holds is a direct consequence of the openness of $\mathcal{R}_{\mathbf{m}}$.

Direct application of Theorem 13 leads to the following results.

**Theorem 14** *Let $\boldsymbol{\rho}\in\mathcal{R}_{\mathbf{m}}$, and let $h_S$ and $\psi$ be given by (20) and (75), respectively. Then*

$$h_S(\boldsymbol{\rho}) = \max_{\boldsymbol{\alpha}\in\mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\alpha},\boldsymbol{\rho})\,, \tag{91}$$

*where the maximum on the right is attained by a unique* $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$*. If* $\hat{g}_{\boldsymbol{\rho}}$ *solves (20), then* $\hat{g}_{\boldsymbol{\rho}} = G_{\hat{\boldsymbol{\alpha}}}$*. Furthermore,* $\hat{g}_{\boldsymbol{\rho}}$ *and* $\hat{\boldsymbol{\alpha}}$ *satisfy the complementary slackness condition*

$$\hat{\boldsymbol{\alpha}}^T \boldsymbol{\rho} = \hat{\boldsymbol{\alpha}}^T \langle \mathbf{m} \hat{g}_{\boldsymbol{\rho}} \rangle = \hat{\boldsymbol{\alpha}}^T \langle \mathbf{m} G_{\hat{\boldsymbol{\alpha}}} \rangle \tag{92}$$

*and* $\hat{g}_{\boldsymbol{\rho}}$ *minimizes* $\mathcal{L}(g, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho})$ *over* $\mathbb{F}_{\mathbf{m}}$*, i.e.,*

$$\psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}) = \mathcal{L}(\hat{g}_{\boldsymbol{\rho}}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}) \; . \tag{93}$$

**Proof.** Theorem 14 may be recast in the form of Theorem 13 by setting $m = 1$ and introducing the following mapping of notation:

$$\begin{aligned}
&\mathbb{X} \mapsto L^1_{\mathbf{m}}(\mathbb{R}^d)\,; && f_1(x) \mapsto \boldsymbol{\rho}_N - \langle \mathbf{m}_N g \rangle\;; && Ax \mapsto \langle \mathbf{m}_j g \rangle\;, \quad j = 1, \dots, N-1\,;\\
&x \mapsto g\,; && K_1 \mapsto A^{\circ}_{\mathbf{m}_N}\,; && b \mapsto \boldsymbol{\rho}_j\,, \quad j = 1, \dots, N-1\,;\\
&f_0 \mapsto \mathcal{H}\,; && \lambda \mapsto \boldsymbol{\alpha}_N\,; && \nu \mapsto \boldsymbol{\alpha}_j\,, \quad j = 1, \dots, N-1\,.
\end{aligned}$$

All the conditions of Theorem 13 hold. However, we must be careful to ensure that $\mathcal{H}$ is restricted to a domain on which it is finite. Thus we consider the minimization problem over the set

$$\tilde{\mathbb{F}}_{\mathbf{m}} = \{ g \in \mathbb{F}_{\mathbf{m}} : \mathcal{H}(g) < \infty \} \; . \tag{94}$$

This set is convex and includes all bounded function in $\mathbb{F}_{\mathbf{m}}$ with compact support. Thus by Theorem 1, the moment mapping $g \mapsto \langle \mathbf{m} g \rangle$ maps $\tilde{\mathbb{F}}_{\mathbf{m}}$ onto $\mathcal{R}_{\mathbf{m}}$, and since $\mathcal{R}_{\mathbf{m}}$ is open, the set

$$\Big\{ \langle \mathbf{m}_i g \rangle - \boldsymbol{\rho}_i : g \in \tilde{\mathbb{F}}_{\mathbf{m}}\,,\ i < N \Big\} \tag{95}$$

contains a neighborhood of the origin. By the Polar Cone Theorem [BNO03, p.162], $(A^{\circ}_{\mathbf{m}_N})^{\circ} = A_{\mathbf{m}_N}$ so that strong duality holds, i.e.,

$$h_{\mathrm{s}}(\boldsymbol{\rho}) = \max_{\boldsymbol{\alpha} \in \mathbb{R}^n} \{ \psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) : \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N} \} \; . \tag{96}$$

Moreover, because $\psi$ is strictly concave, the maximum in (96) is attained by a *unique* $\hat{\boldsymbol{\alpha}} \in \{ \boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N} \}$. According to the constraint conditions in (52)

$$\langle \mathbf{m}_j \hat{g}_{\boldsymbol{\rho}} \rangle = \boldsymbol{\rho} \text{ for } j < N \qquad \text{and} \qquad \hat{\boldsymbol{\alpha}}_N^T \langle \mathbf{m}_N \hat{g}_{\boldsymbol{\rho}} \rangle \geq \hat{\boldsymbol{\alpha}}_N^T \boldsymbol{\rho}_N \; . \tag{97}$$

Thus $\hat{\boldsymbol{\alpha}}^T(\boldsymbol{\rho} - \langle \mathbf{m} \hat{g}_{\boldsymbol{\rho}} \rangle) \leq 0$ and

$$h_{\mathrm{s}}(\boldsymbol{\rho}) = \psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}) = \inf_{g \in \mathbb{F}_{\mathbf{m}}} \Big\{ \mathcal{H}(g) + \hat{\boldsymbol{\alpha}}^T (\boldsymbol{\rho} - \langle \mathbf{m} g \rangle) \Big\} \leq \mathcal{H}(\hat{g}_{\boldsymbol{\rho}}) + \hat{\boldsymbol{\alpha}}^T (\boldsymbol{\rho} - \langle \mathbf{m} \hat{g}_{\boldsymbol{\rho}} \rangle) \leq \mathcal{H}(\hat{g}_{\boldsymbol{\rho}}) = h_{\mathrm{s}}(\boldsymbol{\rho}) \, . \tag{98}$$

Equations (92) and (93) follow immediately.

To finish the proof, we need only show that $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$ and $\hat{g}_{\boldsymbol{\rho}} = G_{\hat{\boldsymbol{\alpha}}}$. For any non-negative function $g$, straight-forward calculation verifies that

$$g \log g - g - \hat{\boldsymbol{\alpha}}^T \mathbf{m} g = \phi(g) - G_{\hat{\boldsymbol{\alpha}}} \, , \tag{99}$$

where

$$\phi(g) \equiv \left[ g \log \left( \frac{g}{G_{\hat{\boldsymbol{\alpha}}}} \right) + (G_{\hat{\boldsymbol{\alpha}}} - g) \right] . \tag{100}$$

Applying (49) with $z = g/G_{\hat{\boldsymbol{\alpha}}}$ and $y = 1$ shows that for each $v \in \mathbb{R}^d$, $\phi(g(v)) \geq 0$, with equality if and only if $G_{\hat{\boldsymbol{\alpha}}}(v) = g(v)$. Now, for each $R > 0$, define the set $B_R \equiv \big\{ v \in \mathbb{R}^d : |v| < R \big\}$. Setting $g = \hat{g}_{\boldsymbol{\rho}}$ in (99) and integrating over $B_R$ gives

$$\mathcal{H}^{B_R}(\hat{g}_{\boldsymbol{\rho}}) - \Big\langle \hat{\boldsymbol{\alpha}}^T \mathbf{m} \hat{g}_{\boldsymbol{\rho}} \Big\rangle_{B_R} = \langle \phi(\hat{g}_{\boldsymbol{\rho}}) \rangle_{B_R} - \langle G_{\hat{\boldsymbol{\alpha}}} \rangle_{B_R} \; . \tag{101}$$

(Note that, since $\hat{g}_{\boldsymbol{\rho}} \in \mathbb{F}_{\mathbf{m}}$, all the integrals above are well-defined.) From (74), (101), (76), and (39), it follows that

$$\begin{aligned}\mathcal{L}\left(\hat{g}_{\boldsymbol{\rho}}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}\right) &= \mathcal{H}(\hat{g}_{\boldsymbol{\rho}}) + \hat{\boldsymbol{\alpha}}^T(\boldsymbol{\rho} - \langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle) \\ &= \mathcal{H}^{B_R}(\hat{g}_{\boldsymbol{\rho}}) + \mathcal{H}^{\mathbb{R}^d\setminus B_R}(\hat{g}_{\boldsymbol{\rho}}) + \hat{\boldsymbol{\alpha}}^T\left(\boldsymbol{\rho} - \langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle_{B_R} - \langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle_{\mathbb{R}^d\setminus B_R}\right) \\ &= \langle \phi(\hat{g}_{\boldsymbol{\rho}})\rangle_{B_R} - \langle G_{\hat{\boldsymbol{\alpha}}}\rangle_{B_R} + \mathcal{H}^{\mathbb{R}^d\setminus B_R}(\hat{g}_{\boldsymbol{\rho}}) + \hat{\boldsymbol{\alpha}}^T\left(\boldsymbol{\rho} - \langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle_{\mathbb{R}^d\setminus B_R}\right) \\ &= \mathcal{L}\left(G^{B_R}_{\hat{\boldsymbol{\alpha}}}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}\right) + \langle \phi(\hat{g}_{\boldsymbol{\rho}})\rangle_{B_R} + \mathcal{H}^{\mathbb{R}^d\setminus B_R}(\hat{g}_{\boldsymbol{\rho}}) - \hat{\boldsymbol{\alpha}}^T\langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle_{\mathbb{R}^d\setminus B_R}\,, \end{aligned} \tag{102}$$

where

$$G^{B_R}_{\hat{\boldsymbol{\alpha}}}(v) = \begin{cases} G_{\hat{\boldsymbol{\alpha}}}(v)\,, & v \in B_R \\ 0\,, & v \notin B_R \end{cases}. \tag{103}$$

Now since $\phi(g(v)) \geq 0$, the function $R \mapsto \Phi(R) \equiv \langle \phi(\hat{g}_{\boldsymbol{\rho}})\rangle_{B_R}$ is a non-negative and non-decreasing, and $\Phi(R) = 0$ if and only if $G_{\hat{\boldsymbol{\alpha}}}$ and $g$ agree on $B_R$. On the other hand, since $\mathcal{H}(\hat{g}_{\boldsymbol{\rho}})$ and $\langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle$ are finite,

$$\mathcal{H}^{\mathbb{R}^d\setminus B_R}(\hat{g}_{\boldsymbol{\rho}}) - \hat{\boldsymbol{\alpha}}^T\langle \mathbf{m}\hat{g}_{\boldsymbol{\rho}}\rangle_{\mathbb{R}^d\setminus B_R} \to 0 \quad \text{as} \quad R \to \infty\,. \tag{104}$$

It follows then from (102) that for $R$ is sufficiently large,

$$\mathcal{L}\left(\hat{g}_{\boldsymbol{\rho}}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}\right) > \mathcal{L}\left(G^{B_R}_{\hat{\boldsymbol{\alpha}}}, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}\right) \tag{105}$$

unless $G^{B_R}_{\hat{\boldsymbol{\alpha}}}$ agrees with $\hat{g}_{\boldsymbol{\rho}}$ on $B_R$. Since $\hat{g}_{\boldsymbol{\rho}}$ minimizes $\mathcal{L}\left(\cdot, \hat{\boldsymbol{\alpha}}, \boldsymbol{\rho}\right)$, we conclude that this exception is indeed the case. Moreover, since $R$ is arbitrary, it follows that $\hat{g}_{\boldsymbol{\rho}} = G_{\hat{\boldsymbol{\alpha}}}$. Finally, the fact that $\hat{g}_{\boldsymbol{\rho}} \in \mathbb{F}_{\mathbf{m}}$ implies that $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$. ■

Several remarks are in order here.

1. If $\boldsymbol{\rho} \in \mathcal{R}^{\exp}_{\mathbf{m}}$, then Theorem 14 can be proven more directly using Theorem 3. Indeed, weak duality is easy to show: If $g \in \mathbb{F}_{\mathbf{m}}$ satisfies the constraint conditions from (52), then

$$\mathcal{L}\left(g, \boldsymbol{\alpha}, \boldsymbol{\rho}\right) = \mathcal{H}(g) + \boldsymbol{\alpha}^T\left(\boldsymbol{\rho} - \langle \mathbf{m}g\rangle\right) \leq \mathcal{H}(g) \tag{106}$$

   for all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. Invoking the definitions of $\psi$ (75) and $h_S$ (20), we find that

$$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) = \inf_{g\in\mathbb{F}_{\mathbf{m}}} \mathcal{L}(g, \boldsymbol{\alpha}, \boldsymbol{\rho}) \leq \inf_{g\in\mathbb{F}_{\mathbf{m}}} \{\mathcal{L}(g, \boldsymbol{\alpha}, \boldsymbol{\rho}) : \langle \mathbf{m}g\rangle \preceq^\circ \boldsymbol{\rho}\} \leq \inf_{g\in\mathbb{F}_{\mathbf{m}}} \{\mathcal{H}(g) : \langle \mathbf{m}g\rangle \preceq^\circ \boldsymbol{\rho}\} = h_{\mathrm{S}}(\boldsymbol{\rho})\,. \tag{107}$$

   On the other hand, if $\boldsymbol{\rho} = \mathbf{r}(\hat{\boldsymbol{\alpha}})$ for some $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$, then it follows from Theorem 3, (76) and the definition of $\mathcal{H}$ (47) that

$$h_{\mathrm{S}}(\boldsymbol{\rho}) = \mathcal{H}(G_{\hat{\alpha}}) = \psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{\rho})\,. \tag{108}$$

   From (107) and (108), one can easily deduce strong duality (91) and the complementary slackness condition (92).

2. If it is known a priori that the maximum in (96) is attained by $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$, then the form of the minimizer follows almost immediately. In this case, $G_{\hat{\boldsymbol{\alpha}}} \in \mathbb{F}_{\mathbf{m}}$ (which is needed for $\mathcal{L}$ to be well-defined) so that (76) and (93) imply (108). Because $\mathcal{L}$ is strictly convex in its first argument, its minimizer is unique and consequently, $\hat{g}_{\boldsymbol{\rho}} = G_{\hat{\boldsymbol{\alpha}}}$.

3. Since $\boldsymbol{\rho}_j = \langle \mathbf{m}_j G_{\hat{\boldsymbol{\alpha}}}\rangle$ for $j < N$, the only nontrivial part of the complementary slackness condition (92) is

$$\hat{\boldsymbol{\alpha}}_N^T\boldsymbol{\rho}_N = \hat{\boldsymbol{\alpha}}_N^T\langle \mathbf{m}_N\hat{g}_{\boldsymbol{\rho}}\rangle = \hat{\boldsymbol{\alpha}}_N^T\langle \mathbf{m}_N G_{\hat{\boldsymbol{\alpha}}}\rangle\,. \tag{109}$$

   This relationship between $\hat{\boldsymbol{\alpha}}_N$ and $\boldsymbol{\rho}_N$ will be the key to characterizing the set $\mathcal{D}_{\mathbf{m}}$.

The following corollary will be used in Section 6. It is an immediate consequence of the complementary slackness condition.

**Corollary 15** *Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ and let $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$ solve (91). Then*

$$h_S(\boldsymbol{\rho}) = \min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \hat{\boldsymbol{\alpha}}^T \langle \mathbf{m} g \rangle = \hat{\boldsymbol{\alpha}}^T \boldsymbol{\rho} \right\} , \tag{110}$$

*and $G_{\hat{\boldsymbol{\alpha}}}$ is the unique minimizer.*

**Proof.** Let $g \in \mathbb{F}_{\mathbf{m}}$ be given. Using Young's Inequality (49) with $z = g$ and $y = G_{\hat{\boldsymbol{\alpha}}}$ gives

$$\mathcal{H}(g) \geq \mathcal{H}(G_{\hat{\boldsymbol{\alpha}}}) + \hat{\boldsymbol{\alpha}}^T \langle \mathbf{m}(g - G_{\hat{\boldsymbol{\alpha}}}) \rangle \tag{111}$$

which, given the complementary slackness condition (92), implies that

$$\mathcal{H}(g) \geq \mathcal{H}(G_{\hat{\boldsymbol{\alpha}}}) + \hat{\boldsymbol{\alpha}}^T \left( \langle \mathbf{m} g \rangle - \boldsymbol{\rho} \right) . \tag{112}$$

Thus if $g$ satisfies the constraints in (110), then $\mathcal{H}(g) \geq \mathcal{H}(G_{\hat{\boldsymbol{\alpha}}}) = h_{\mathrm{S}}(\boldsymbol{\rho})$. ■

A duality theorem similar to Theorem 14 holds for the minimization problem in (19) that defines $h_J(\boldsymbol{\rho})$. Like Theorem 14, it is a consequence of Theorem 13, and its proof is essentially the same.

**Theorem 16** *Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, and let $h_J(\boldsymbol{\rho})$ and $\psi$ be given by (19) and (75), respectively. Then*

$$h_J(\boldsymbol{\rho}) = \max_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) , \tag{113}$$

*where the maximum on the right is attained by a unique $\tilde{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$. Furthermore, if the infimum in (19) is attained by some function $\tilde{g}_{\boldsymbol{\rho}} \in \mathbb{F}_{\mathbf{m}}$ which satisfies the equality constraints of (19), then $\tilde{g}_{\boldsymbol{\rho}} = G_{\tilde{\boldsymbol{\alpha}}}$ and $\tilde{g}_{\boldsymbol{\rho}}$ minimizes $\mathcal{L}(g, \tilde{\boldsymbol{\alpha}}, \boldsymbol{\rho})$, i.e., $\psi(\tilde{\boldsymbol{\alpha}}, \boldsymbol{\rho}) = \mathcal{L}(\tilde{g}_{\boldsymbol{\rho}}, \tilde{\boldsymbol{\alpha}}, \boldsymbol{\rho})$.*

The careful reader may note that application of Theorem 13 to proving Theorem 16 initially gives a statement similar to (96), but without any constraint on $\boldsymbol{\alpha}$. However, the arguments which follow (96) show that $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ independently of this initial restriction.

Theorems 14 and 16 prove that the infima in (19) and (20) are equal—that is,

$$h_{\mathrm{S}}(\boldsymbol{\rho}) = h_{\mathrm{J}}(\boldsymbol{\rho}) = \max_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) , \tag{114}$$

even if the infimum in (19) is not attained. In light of (114), the definition of $h$ given in (17), which applies only to $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$, can be extended to all of $\mathcal{R}_{\mathbf{m}}$ by setting

$$h(\boldsymbol{\rho}) \equiv \max_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) . \tag{115}$$

In addition, we can now complete the proof of Theorem 9.

**Proof of Theorem 9.** We have already proven the "if" statement in Theorem 9. We now prove the "only if" statement. To this end, let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ be such that (14) has a minimizer. According to (114) this minimizer also the minimizer of (20) and is therefore given by $G_{\mathbf{a}(\boldsymbol{\rho})}$. Hence, the equality constraint conditions in (14) imply $\boldsymbol{\rho} = \left\langle \mathbf{m} G_{\mathbf{a}(\boldsymbol{\rho})} \right\rangle$, which means $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. ■

# 5 The Relationship between $\boldsymbol{\alpha}$ and $\boldsymbol{\rho}$

The formal structure of entropy-based closures depends heavily on the Legendre dual relationship between the functions $h$ and $h^*$ and their derivatives. In this section, we review this relationship for non-degenerate densities and show how Legendre duality ensures that the resulting system of PDE is symmetric hyperbolic. We then discuss what aspects of the dual relationship hold in degenerate densities. A similar analysis can be found in [Jun00] for the case $\mathbf{m}_N = |v|^N$.

## 5.1 Properties for Non-Degenerate Cases

Recall that the function $\mathbf{a}$ maps each $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ to the unique vector $\hat{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}}$ that solves (91). In particular,

$$\hat{g}_{\boldsymbol{\rho}} = G_{\mathbf{a}(\boldsymbol{\rho})} \qquad \text{and} \qquad h(\boldsymbol{\rho}) = \psi(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho}) \,. \tag{116}$$

It turns out that $\mathbf{a}$, when restricted to $\mathcal{R}_{\mathbf{m}}^{\exp}$, is the inverse of the function $\mathbf{r}$ defined in (41).

**Theorem 17** *The function* $\mathbf{r}$ *is one-to-one from* $\mathcal{A}_{\mathbf{m}}$ *onto* $\mathcal{R}_{\mathbf{m}}^{\exp}$ *with inverse* $\mathbf{a}$. *It is a diffeomorphism between* $\operatorname{int}\mathcal{A}_{\mathbf{m}}$ *and* $\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$.

**Proof.** We first identify $\mathbf{a}$ as the inverse of $\mathbf{r}$. Since $\mathbf{r}$ is (by definition) onto $\mathcal{R}_{\mathbf{m}}^{\exp}$, we need only to show that $\mathbf{a}(\mathbf{r}(\boldsymbol{\alpha})) = \boldsymbol{\alpha}$ for each $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. By the definition of $\mathbf{a}$,

$$\mathcal{H}(G_{\mathbf{a}(\mathbf{r}(\boldsymbol{\alpha}))}) = \min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \mathbf{r}(\boldsymbol{\alpha}) \right\} \,. \tag{117}$$

However, Theorem 3 implies that

$$\mathcal{H}(G_{\boldsymbol{\alpha}}) = \min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \langle \mathbf{m} g \rangle \preceq^{\circ} \mathbf{r}(\boldsymbol{\alpha}) \right\} \,. \tag{118}$$

Since this minimizer is unique, it follows that $\mathbf{a}(\mathbf{r}(\boldsymbol{\alpha})) = \boldsymbol{\alpha}$. If $\boldsymbol{\alpha} \in \operatorname{int}\mathcal{A}_{\mathbf{m}}$, then according to Corollary 12, $\mathbf{r}$ is the derivative of the density potential $h^*$ on $\operatorname{int}\mathcal{A}_{\mathbf{m}}$ and its Jacobian

$$\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}) = \frac{\partial^2 h^*}{\partial \boldsymbol{\alpha}^2}(\alpha, \boldsymbol{\rho}) = \left\langle \mathbf{m}\mathbf{m}^T G_{\boldsymbol{\alpha}} \right\rangle \tag{119}$$

is a positive-definite matrix. The inverse function theorem implies then that $\mathbf{r}$ is a diffeomorphism from $\operatorname{int}\mathcal{A}_{\mathbf{m}}$ onto $\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$. ■

The following corollary implies that $\mathcal{D}_{\mathbf{m}}$ cannot divide $\mathcal{R}_{\mathbf{m}}^{\exp}$ into disjoint subsets.

**Corollary 18** *The set* $\mathcal{R}_{\mathbf{m}}^{\exp}$ *is pathwise-connected.*

**Proof.** Given $\boldsymbol{\rho}_{(0)}, \boldsymbol{\rho}_{(1)} \in \mathcal{R}_{\mathbf{m}}^{\exp}$, we seek a continuous function $\Gamma : [0,1] \to \mathcal{R}_{\mathbf{m}}^{\exp}$ such that

$$\Gamma(0) = \boldsymbol{\rho}_{(0)} \quad \text{and} \quad \Gamma(1) = \boldsymbol{\rho}_{(1)} \,. \tag{120}$$

Convexity of $\mathcal{A}_{\mathbf{m}}$ implies that

$$\boldsymbol{\alpha}_\lambda \equiv \lambda \mathbf{a}(\boldsymbol{\rho}_{(0)}) + (1-\lambda)\mathbf{a}(\boldsymbol{\rho}_{(1)}) \in \mathcal{A}_{\mathbf{m}} \quad \forall\ \lambda \in [0,1] \,. \tag{121}$$

Thus, in view of Theorem 11, the function $\Gamma(\lambda) = \mathbf{r}(\boldsymbol{\alpha}_\lambda)$ satisfies (120). ■

An immediate consequence of Theorem 14 is that $h$ (as the maximum of a family of linear functions in $\rho$) is convex on $\mathcal{R}_{\mathbf{m}}$. However, more can be said if we restrict $h$ to convex subsets of $\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$.

**Theorem 19** *When restricted to* $\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$ *and* $\operatorname{int}\mathcal{A}_{\mathbf{m}}$, *respectively, the functions* $h$ *and* $h^*$ *are locally strictly convex, Legendre duals of one another.*

**Proof.** We first show that $h$ is the Legendre transform of $h^*$. From (76),

$$h(\mathbf{r}(\boldsymbol{\alpha})) + h^*(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T \mathbf{r}(\boldsymbol{\alpha}) \,, \quad \boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \,, \tag{122}$$

where, according to Corollary 12,

$$\mathbf{r}(\boldsymbol{\alpha}) = h^*_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \,, \quad \boldsymbol{\alpha} \in \operatorname{int}\mathcal{A}_{\mathbf{m}} \,. \tag{123}$$

We next show that the Legendre transform of $h^*$ recovers $h$. The inverse relationship between $\mathbf{a}$ and $\mathbf{r}$ (Theorem 17) implies that (122) may be rewritten in terms of $\boldsymbol{\rho} = \mathbf{r}(\boldsymbol{\alpha})$:

$$h(\boldsymbol{\rho}) + h^*(\mathbf{a}(\boldsymbol{\rho})) = \mathbf{a}(\boldsymbol{\rho})^T \boldsymbol{\rho} \,, \quad \boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp} \,. \tag{124}$$

Differentiating (124) and using (123) again gives

$$\mathbf{a}(\boldsymbol{\rho}) = h_{\boldsymbol{\rho}}(\boldsymbol{\rho}), \quad \boldsymbol{\rho} \in \operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp} \,. \tag{125}$$

Finally, for $\boldsymbol{\rho} \in \operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$,

$$h_{\boldsymbol{\rho}\boldsymbol{\rho}}(\boldsymbol{\rho}) = \frac{\partial \mathbf{a}}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho}) = \left[ \frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}(\mathbf{a}(\boldsymbol{\rho})) \right]^{-1} = \left[ h^*_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\mathbf{a}(\boldsymbol{\rho})) \right]^{-1} \tag{126}$$

which, by Corollary 12, is positive-definite. Thus $h$ and $h^*$ are strictly convex. ■

### 5.2 Application to Kinetic Moment Closures.

The dual relationship between $h$ and $h^*$ is used in [Lev96] to show that entropy-based closures formally produce hyperbolic systems which dissipate a convex entropy and satisfy an H-Theorem. Indeed, if $\boldsymbol{\rho} \in \operatorname{int} \mathcal{R}_{\mathbf{m}}^{\exp}$ and $\hat{\boldsymbol{\alpha}} = \mathbf{a}(\boldsymbol{\rho})$, then according to (123), the moment system (12) can be expressed in terms of $\hat{\boldsymbol{\alpha}}$:

$$\partial_t h^*_{\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}}) + \nabla_x \cdot j^*_{\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}}) = \mathbf{c}(h^*_{\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}})) \,, \tag{127}$$

where $j^*(\boldsymbol{\alpha}) \equiv \langle v G_{\boldsymbol{\alpha}} \rangle$ is the flux potential and

$$j^*_{\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}}) = \mathbf{f}(\boldsymbol{\rho}) \,. \tag{128}$$

Differentiating the left-hand side of (127) with respect to $\boldsymbol{\alpha}$ gives

$$h^*_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}}) \partial_t \hat{\boldsymbol{\alpha}} + j^*_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}}) \cdot \nabla_x \hat{\boldsymbol{\alpha}} = \mathbf{c}(h^*_{\boldsymbol{\alpha}}(\hat{\boldsymbol{\alpha}})) \,, \tag{129}$$

which has the form of a symmetric hyperbolic system [Eva02]. Furthermore, by multiplying (12) by $h_{\boldsymbol{\rho}}$ and applying relations (125) and (128), we find that $h(\boldsymbol{\rho})$ satisfies:

$$\partial_t h(\boldsymbol{\rho}) + \nabla_x \cdot j(\boldsymbol{\rho}) = \mathbf{a}(\boldsymbol{\rho})^T \mathbf{c}(\boldsymbol{\rho}), \tag{130}$$

where $j(\boldsymbol{\rho}) \equiv \mathbf{a}(\boldsymbol{\rho})^T \mathbf{f}(\boldsymbol{\rho}) - j^*(\mathbf{a}(\boldsymbol{\rho}))$. Then by (5) and (6),

$$\mathbf{a}(\boldsymbol{\rho})^T \mathbf{c}(\boldsymbol{\rho}) = \mathcal{S}(G_{\mathbf{a}(\boldsymbol{\rho})}) \leq 0 \tag{131}$$

with equality if and only if $G_{\mathbf{a}(\boldsymbol{\rho})}$ is a local Maxwellian (7). This is a direct analog of Boltzmann's H-Theorem for (2). (See [Lev96] for details.)

### 5.3 Non-Degenerate Examples

For $N = 2$, there are two possible closures: Maxwellian and Gaussian. Both are well-known, and in both cases, $\mathcal{A}_{\mathbf{m}} = \operatorname{int} \mathcal{A}_{\mathbf{m}}$ and $\mathcal{R}_{\mathbf{m}} = \mathcal{R}_{\mathbf{m}}^{\exp} = \operatorname{int} \mathcal{R}_{\mathbf{m}}^{\exp}$. Traditionally, these closures are expressed using so-called fluid variables:

$$\begin{array}{rlrl} \text{density:} & \rho = \langle F \rangle \,, & \text{temperature matrix:} & \Theta = \frac{\langle (v-u)\vee(v-u)\, F \rangle}{\langle F \rangle} \,, \\ \text{bulk velocity:} & u = \frac{\langle vF \rangle}{\langle F \rangle} \,, & \text{temperature:} & \theta = \frac{1}{3}\operatorname{trace}(\Theta) = \frac{\langle |v-u|^2 F \rangle}{3\langle F \rangle} \,. \end{array} \tag{132}$$

1. **Maxwellian closure**. If $\mathbf{m} = (1, v, \frac{1}{2}|v|^2)^T$, the ansatz $\mathcal{F}[\boldsymbol{\rho}]$ in (14) is a Maxwellian distribution:

$$\mathcal{M}_{\rho,u,\theta}(v) \equiv \frac{\rho}{(2\pi\theta)^{d/2}} \exp\left( -\frac{|v-u|^2}{2\theta} \right) \,. \tag{133}$$

The fluid variables are related to the densities $\boldsymbol{\rho}_i$ by

$$\boldsymbol{\rho}_0 = \rho \,, \qquad \boldsymbol{\rho}_1 = \rho u \,, \qquad \boldsymbol{\rho}_2 = \frac{1}{2}\rho u^2 + \frac{3}{2}\rho\theta \tag{134}$$

and to the vectors $\hat{\boldsymbol{\alpha}}_i$ by

$$\hat{\boldsymbol{\alpha}}_0 = \log\left( \frac{\rho}{(2\pi\theta)^{d/2}} \right) - \frac{|u|^2}{2\theta} \,, \qquad \hat{\boldsymbol{\alpha}}_1 = \frac{u}{\theta} \,, \qquad \hat{\boldsymbol{\alpha}}_2 = -\frac{1}{\theta} \,. \tag{135}$$

The moment equations in this case are the compressible Euler equations for a gas of point particles:

$$\partial_t n + \nabla_x \cdot (nu) = 0 \,, \tag{136a}$$

$$\partial_t \, (nu) + \nabla_x \cdot (nu \vee u + n\theta I) = 0 \,, \tag{136b}$$

$$\partial_t \left( \frac{1}{2} n|u|^2 + \frac{d}{2} n\theta \right) + \nabla_x \cdot \left( \frac{1}{2} n|u|^2 u + \frac{d+2}{2} n\theta u \right) = 0 \,. \tag{136c}$$

The spatial entropy,

$$h(\boldsymbol{\rho}) = \langle \mathcal{M}_{\rho,u,\theta} \log \mathcal{M}_{\rho,u,\theta} - \mathcal{M}_{\rho,u,\theta} \rangle = \rho \left[ \log \left( \frac{\rho}{(2\pi\theta)^{d/2}} \right) - \frac{d+2}{2} \right] , \tag{137}$$

is locally conserved by smooth solutions for (136), but is dissipated along shocks.

2. **Gaussian closure**. If $\mathbf{m} = (1, v, v \vee v)^T$, the ansatz $\mathcal{F}[\boldsymbol{\rho}]$ in (14) is a Gaussian distribution:

$$\mathcal{G}_{\rho,u,\Theta}(v) = \frac{\rho}{\sqrt{\det(2\pi\Theta)}} \exp\left( -\frac{1}{2}(v-u) \cdot \Theta^{-1} \cdot (v-u) \right) . \tag{138}$$

The fluid variables are related to the densities $\boldsymbol{\rho}_i$ by

$$\boldsymbol{\rho}_0 = \rho \,, \qquad \boldsymbol{\rho}_1 = \rho u \,, \qquad \boldsymbol{\rho}_2 = \rho u \vee u + n\Theta \tag{139}$$

and to the vectors $\hat{\boldsymbol{\alpha}}_i$ by

$$\hat{\boldsymbol{\alpha}}_0 = \log \left( \frac{\rho}{\sqrt{\det(2\pi\Theta)}} \right) - \frac{1}{2} u \cdot \Theta^{-1} \cdot u \,, \qquad \hat{\boldsymbol{\alpha}}_1 = \Theta^{-1} \cdot u \,, \qquad \hat{\boldsymbol{\alpha}}_2 = -\frac{1}{2}\Theta^{-1} \,. \tag{140}$$

The moment equations in this case are

$$\partial_t n + \nabla_x \cdot (nu) = 0 \,, \tag{141a}$$

$$\partial_t (nu) + \nabla_x \cdot (nu \vee u + n\Theta) = 0 \,, \tag{141b}$$

$$\partial_t (nu \vee u + n\Theta) + \nabla_x \cdot (nu \vee u \vee u + 3n\Theta \vee u) = \langle v \vee v \, \mathcal{C}(\mathcal{G}_{\rho,u,\Theta}) \rangle \,, \tag{141c}$$

and solutions to this system satisfy a local dissipation law for the spatial entropy

$$h(\boldsymbol{\rho}) = \langle \mathcal{G}_{\rho,u,\Theta} \log \mathcal{G}_{\rho,u,\Theta} - \mathcal{G}_{\rho,u,\Theta} \rangle = \rho \left[ \log \left( \frac{\rho}{\sqrt{\det(2\pi\Theta)}} \right) - \frac{d+2}{2} \right] . \tag{142}$$

Note that in both of the examples above, the expressions for $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{\rho}$ can be used to determine $\mathbf{a}(\boldsymbol{\rho})$ explicitly. However, generally speaking, an analytical solution is not available and a numerical solution must be computed via (91).

## 5.4 Properties for Degenerate Cases

If $\boldsymbol{\rho} \in \mathcal{D}_{\mathbf{m}}$, then the minimizer with equality constraints (14) does not exist and the entropy-based closure is not well-defined. Although it is possible to recover a well-defined closure using the relaxed constraints in (20), much of the formal structure is lost. For example, if $\boldsymbol{\rho} \in \mathcal{D}_{\mathbf{m}}$, then (123) and (126) no longer hold because $\mathbf{r}\left(\mathbf{a}(\boldsymbol{\rho})\right) \neq \boldsymbol{\rho}$ and, as shown in Corollary 22 below, $h_{\mathrm{S}}$ fails to be strictly convex on $\mathcal{R}_{\mathbf{m}}$ whenever $\mathcal{D}_{\mathbf{m}}$ is non-empty. Since many of the properties of entropy-based closures require $h$ to be strictly convex, this fact is critical.

The situation for degenerate densities may be best understood via the projection operator $\boldsymbol{\pi} : \mathcal{R}_{\mathbf{m}} \to \mathcal{R}_{\mathbf{m}}^{\exp}$, which assigns to each vector $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ the density which is realized by the minimizer of (20):

$$\boldsymbol{\pi}(\boldsymbol{\rho}) \equiv \mathbf{r}(\mathbf{a}(\boldsymbol{\rho})) = \left\langle \mathbf{m} G_{\mathbf{a}(\boldsymbol{\rho})} \right\rangle . \tag{143}$$

Before discussing $\boldsymbol{\pi}$ further, we introduce some notation that will be useful for the remainder of the paper. First we have the natural decompositions for $\mathbf{r}$, $\mathbf{a}$, and $\boldsymbol{\pi}$ based on the decomposition of $\mathbf{m}$ in (22):

$$\mathbf{r} = \left(\mathbf{r}_0^T, \mathbf{r}_1^T, \ldots, \mathbf{r}_N^T\right)^T , \qquad \mathbf{a} = \left(\mathbf{a}_0^T, \mathbf{a}_1^T, \ldots, \mathbf{a}_N^T\right)^T , \qquad \boldsymbol{\pi} = \left(\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \ldots, \boldsymbol{\pi}_N^T\right)^T . \tag{144}$$

With this notation,

$$G_{\mathbf{a}(\boldsymbol{\rho})} = \exp\left( \textstyle\sum_{j=1}^N \mathbf{a}_j(\boldsymbol{\rho})^T \mathbf{m}_j \right) , \quad \mathbf{r}_j(\boldsymbol{\alpha}) = \langle \mathbf{m}_j G_{\boldsymbol{\alpha}} \rangle \,, \quad \boldsymbol{\pi}_j(\boldsymbol{\rho}) = \mathbf{r}_j(\mathbf{a}(\boldsymbol{\rho})) \,. \tag{145}$$

Next, for any $\boldsymbol{\rho} \in \mathbb{R}^n$ and any $\boldsymbol{\zeta} \in \mathbb{R}^{n_N}$, we define

$$\boldsymbol{\rho} +_N \boldsymbol{\zeta} \equiv (\boldsymbol{\rho}_0^T, \boldsymbol{\rho}_1^T, \ldots, \boldsymbol{\rho}_N^T + \boldsymbol{\zeta}^T)^T \,. \tag{146}$$

This notation will often be applied to subsets of $\mathbb{R}^n$ and $\mathbb{R}^{n_N}$ in the context of set addition.

**Proposition 20** *Let $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ and let $\bar{\boldsymbol{\alpha}} = \mathbf{a}(\bar{\boldsymbol{\rho}})$. Then for any $\boldsymbol{\rho} \in \mathbb{R}^n$, the following are equivalent:*

$$\textit{i.}\ \ \boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N \in \mathcal{NC}(A_{\mathbf{m}_N}, \bar{\boldsymbol{\alpha}}_N)\,; \tag{147a}$$

$$\textit{ii.}\ \ (\boldsymbol{\alpha}_N - \bar{\boldsymbol{\alpha}}_N)^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) \leq 0, \quad \forall\ \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}\,; \tag{147b}$$

$$\textit{iii.}\ \ \bar{\boldsymbol{\alpha}}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) = 0 \quad \textit{and} \quad \boldsymbol{\alpha}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) \leq 0, \quad \forall\ \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}\,. \tag{147c}$$

**Proof.** Here $(i) \Leftrightarrow (ii)$ is just the definition of a normal cone (31), and the implication $(iii) \Rightarrow (ii)$ is clear. To prove $(ii) \Rightarrow (iii)$, we use the freedom to choose any $\boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}$. Setting $\boldsymbol{\alpha}_N = 0$ and then $\boldsymbol{\alpha}_N = 2\bar{\boldsymbol{\alpha}}_N$ in (147b) gives

$$\bar{\boldsymbol{\alpha}}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) \geq 0 \qquad \text{and} \qquad \bar{\boldsymbol{\alpha}}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) \leq 0\,, \tag{148}$$

respectively. We conclude that $\bar{\boldsymbol{\alpha}}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) = 0$ which, when substituted back into (147b), gives the inequality in $(iii)$. ■

**Lemma 21** *The projection $\boldsymbol{\pi}$ satisfies the following relations:*

$$\textit{i.}\ \ \boldsymbol{\pi}_j(\boldsymbol{\rho}) = \boldsymbol{\rho}_j\,, \quad \forall\ \boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}\ \textit{and}\ j < N\,; \tag{149a}$$

$$\textit{ii.}\ \ \mathbf{a}_N(\boldsymbol{\rho})^T \boldsymbol{\pi}_N(\boldsymbol{\rho}) = \mathbf{a}_N(\boldsymbol{\rho})^T \boldsymbol{\rho}_N\,, \quad \forall\ \boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}\,; \tag{149b}$$

$$\textit{iii.}\ \ \boldsymbol{\pi}(\boldsymbol{\rho}) = \boldsymbol{\rho}\ \textit{if and only if}\ \boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}\,; \tag{149c}$$

$$\textit{iv.}\ \ \boldsymbol{\pi}(\{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \bar{\boldsymbol{\alpha}}_N)\} \cap O) = \bar{\boldsymbol{\rho}}\,, \quad \forall\ \bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}\ \textit{and any}\ O \subset \mathcal{R}_{\mathbf{m}}\ \textit{containing}\ \bar{\boldsymbol{\rho}}\,\boldsymbol{;} \tag{149d}$$

$$\textit{v.}\ \ \boldsymbol{\pi}(\mathcal{D}_{\mathbf{m}}) = \mathbf{r}\,(\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}) = \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}\,; \tag{149e}$$

$$\textit{vi.}\ \ \mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho})) = \mathbf{a}(\boldsymbol{\rho})\,, \quad \forall\ \boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}\,; \tag{149f}$$

$$\textit{vii.}\ \ h(\boldsymbol{\pi}(\boldsymbol{\rho})) = h(\boldsymbol{\rho})\,, \quad \forall\ \boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}\,. \tag{149g}$$

**Proof.** We prove each statement in order.

1. Equation (149a) follows from the constraint conditions in (52a).

2. Equation (149b) is just a restatement of the nontrivial component of the complementary slackness condition (109) with $\hat{\boldsymbol{\alpha}} = \mathbf{a}(\boldsymbol{\rho})$.

3. By Theorem 17, $\boldsymbol{\pi} = \mathbf{r} \circ \mathbf{a}$ is the identity map on $\mathcal{R}_{\mathbf{m}}^{\exp}$. Thus $\boldsymbol{\pi}(\boldsymbol{\rho}) = \boldsymbol{\rho}$ if $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. However, the range of $\boldsymbol{\pi}$ is $\boldsymbol{\pi}(\mathcal{R}_{\mathbf{m}}) = \mathbf{r}(\mathcal{A}_{\mathbf{m}}) = \mathcal{R}_{\mathbf{m}}^{\exp}$. Thus if $\boldsymbol{\rho} \notin \mathcal{R}_{\mathbf{m}}^{\exp}$, then $\boldsymbol{\pi}(\boldsymbol{\rho})$ cannot equal $\boldsymbol{\rho}$.

4. Let $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ and let $O \subset \mathcal{R}_{\mathbf{m}}$ be an open set containing $\bar{\boldsymbol{\rho}}$ and let $\bar{\boldsymbol{\alpha}} = \mathbf{a}(\bar{\boldsymbol{\rho}})$. Choose any $\boldsymbol{\rho} \in \{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \bar{\boldsymbol{\alpha}}_N)\}$. Then $\boldsymbol{\rho} = \bar{\boldsymbol{\rho}}$ for $j < N$, so by Proposition 20, $\bar{\boldsymbol{\alpha}}^T\boldsymbol{\rho} = \bar{\boldsymbol{\alpha}}^T\bar{\boldsymbol{\rho}}$ and $\boldsymbol{\alpha}^T\boldsymbol{\rho} \leq \boldsymbol{\alpha}^T\bar{\boldsymbol{\rho}}$ for all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. Therefore

$$\psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\rho}}) = \psi(\bar{\boldsymbol{\alpha}}, \boldsymbol{\rho}) \leq \psi(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho}) \leq \psi(\mathbf{a}(\boldsymbol{\rho}), \bar{\boldsymbol{\rho}}) \leq \psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\rho}})\,. \tag{150}$$

Here the equality in (150) follows immediately from the definition of $\psi$ (75) and the fact that $\bar{\boldsymbol{\alpha}}^T\boldsymbol{\rho} = \bar{\boldsymbol{\alpha}}^T\bar{\boldsymbol{\rho}}$. The first inequality in (150) uses the fact that $\psi(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho})$ maximizes $\psi(\cdot, \boldsymbol{\rho})$ over all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$; the second uses the fact that $\boldsymbol{\alpha}^T\boldsymbol{\rho} \leq \boldsymbol{\alpha}^T\bar{\boldsymbol{\rho}}$ for all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$; and the third uses the fact that $\psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\rho}})$ maximizes $\psi(\cdot, \bar{\boldsymbol{\rho}})$ over all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$. We conclude from (150) that $\psi(\mathbf{a}(\boldsymbol{\rho}), \bar{\boldsymbol{\rho}}) = \psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\rho}})$. Since $\bar{\boldsymbol{\alpha}}$ is the *unique* maximizer of $\psi(\cdot, \bar{\boldsymbol{\rho}})$ over all $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$, it follows that $\mathbf{a}(\boldsymbol{\rho}) = \bar{\boldsymbol{\alpha}}$. Therefore $\boldsymbol{\pi}(\boldsymbol{\rho}) = \mathbf{r}(\mathbf{a}(\boldsymbol{\rho})) = \mathbf{r}(\bar{\boldsymbol{\alpha}}) = \bar{\boldsymbol{\rho}}$.

5. We first argue by contraction to show that $\boldsymbol{\pi}(\mathcal{D}_{\mathbf{m}}) \subset \mathbf{r}\,(\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}})$. Thus, suppose there exists $\boldsymbol{\rho} \in \mathcal{D}_{\mathbf{m}}$ and $\boldsymbol{\alpha} \in \operatorname{int}\mathcal{A}_{\mathbf{m}}$ such that $\boldsymbol{\pi}(\boldsymbol{\rho}) = \mathbf{r}(\boldsymbol{\alpha})$. We know that
$$\psi(\boldsymbol{\pi}(\boldsymbol{\rho}), \mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho}))) = \max_{\boldsymbol{\alpha}\in\mathcal{A}_{\mathbf{m}}} \psi(\boldsymbol{\pi}(\boldsymbol{\rho}), \boldsymbol{\alpha})\,, \tag{151}$$
and since $\psi$ is differentiable on $\operatorname{int}\mathcal{A}_{\mathbf{m}}$, first order optimality conditions imply that
$$\frac{\partial\psi}{\partial\boldsymbol{\alpha}}(\boldsymbol{\pi}(\boldsymbol{\rho}), \mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho}))) = \boldsymbol{\rho} - \boldsymbol{\pi}(\boldsymbol{\pi}(\boldsymbol{\rho})) = 0\,. \tag{152}$$
However, $\boldsymbol{\pi}$ is a projection; therefore, (152) implies that $\boldsymbol{\rho} = \boldsymbol{\pi}(\boldsymbol{\rho})$. According to (149c), this contradicts the assumption that $\boldsymbol{\rho} \in \mathcal{D}_{\mathbf{m}}$.

   We next show that $\mathbf{r}\,(\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}) \subset \boldsymbol{\pi}(\mathcal{D}_{\mathbf{m}})$. Let $\bar{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ and let $O \subset \mathcal{R}_{\mathbf{m}}$ be an open set containing $\mathbf{r}(\bar{\boldsymbol{\alpha}}) \in \mathcal{R}_{\mathbf{m}}^{\exp}$. Then choose (see (32))
$$\boldsymbol{\rho} \in \{\mathbf{r}(\bar{\boldsymbol{\alpha}}) +_N \mathcal{NC}_0(A_{\mathbf{m}_N}, \bar{\boldsymbol{\alpha}}_N)\} \cap O\,. \tag{153}$$
Since $\bar{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, $\bar{\boldsymbol{\alpha}}_N \in \partial A_{\mathbf{m}_N}$ (see 46) and this set is non-empty. By (149d), $\boldsymbol{\pi}(\boldsymbol{\rho}) = \mathbf{r}(\bar{\boldsymbol{\alpha}})$. Thus we need only show that $\boldsymbol{\rho} \in \mathcal{D}_{\mathbf{m}}$. If it is not, then $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ and $\boldsymbol{\pi}(\boldsymbol{\rho}) = \boldsymbol{\rho} = \mathbf{r}(\bar{\boldsymbol{\alpha}})$, which contradicts (153). Thus $\boldsymbol{\rho} \in \mathcal{D}_{\mathbf{m}}$.

   Finally, we show that $\mathbf{r}\,(\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}) = \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$. Because $\mathbf{r}$ is one-to-one on $\mathcal{A}_{\mathbf{m}}$ (Theorem 17),
$$\mathbf{r}\,(\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}) = \mathbf{r}(\mathcal{A}_{\mathbf{m}})\backslash\mathbf{r}(\operatorname{int}\mathcal{A}_{\mathbf{m}}) = \mathcal{R}_{\mathbf{m}}^{\exp}\backslash\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp} = \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}\,. \tag{154}$$

6. Given that $\mathbf{a}\circ\mathbf{r}$ is the identity map on $\mathcal{A}_{\mathbf{m}}$ (Theorem 17), $\mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho})) = (\mathbf{a}\circ\mathbf{r})(\mathbf{a}(\boldsymbol{\rho})) = \mathbf{a}(\boldsymbol{\rho})$ .

7. The proof is a simple calculation. For any $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, (149a), (149b), and (149f) give
$$h(\boldsymbol{\pi}(\boldsymbol{\rho})) = \psi(\mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho})), \boldsymbol{\pi}(\boldsymbol{\rho})) = \psi(\mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho})), \boldsymbol{\rho}) = \psi(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho}) = h(\boldsymbol{\rho})\,. \tag{155}$$

■

**Corollary 22** *The set $\mathcal{D}_{\mathbf{m}}$ is empty if and only if $\mathcal{A}_{\mathbf{m}}$ is open. If $\mathcal{D}_{\mathbf{m}}$ is non-empty, then $h$ fails to be strictly convex.*

**Proof.** The first statement is an immediate consequence of (149e). The second statement is a consequence of (149d) and (149g), which together imply that $h$ is constant on the cone $\{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))\}$ for any $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. If $\mathcal{D}_{\mathbf{m}}$ is non-empty, then by (149e), $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$ is also non-empty; and if $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$, then $\mathbf{a}(\bar{\boldsymbol{\rho}}) \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ and, consequently, $\mathbf{a}_N(\bar{\boldsymbol{\rho}}) \in \partial A_{\mathbf{m}_N}$ (see 46). As a result, $\{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}(\bar{\boldsymbol{\rho}}))\}$ is nontrivial and $h$ cannot be strictly convex on all of $\mathcal{R}_{\mathbf{m}}$. ■

It turns out that $\mathcal{A}_{\mathbf{m}}$ is open only for $N = 2$. (To see this fact, one need only realize that for $N > 2$, the vector $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}}$ corresponding to any Maxwellian $\mathcal{M}_{\rho,u,\theta}$ lies on the boundary $\partial\mathcal{A}_{\mathbf{m}}$.) Thus Corollary 22 show that the Maxwellian and Gaussian closures are the exception rather than the rule. However, in spite of the difficulties encountered for $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, (124) and (125) extend to all of $\mathcal{R}_{\mathbf{m}}$.

**Theorem 23** *For all $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$,*
$$h(\boldsymbol{\rho}) + h^*(\mathbf{a}(\boldsymbol{\rho})) = \mathbf{a}(\boldsymbol{\rho})^T\boldsymbol{\rho}\,, \tag{156}$$
*and the function $\mathbf{a}$ is the continuous Fréchet derivative of $h$ everywhere on $\mathcal{R}_{\mathbf{m}}$, i.e.,*
$$\mathbf{a}(\boldsymbol{\rho}) = h_{\boldsymbol{\rho}}(\boldsymbol{\rho})\,. \tag{157}$$

**Proof.** Let $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ and set $\bar{\boldsymbol{\rho}} = \boldsymbol{\pi}(\boldsymbol{\rho}) \in \mathcal{R}_{\mathbf{m}}^{\exp}$. By (124),

$$h(\boldsymbol{\pi}(\boldsymbol{\rho})) + h^*(\mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho}))) = \mathbf{a}(\boldsymbol{\pi}(\boldsymbol{\rho}))^T \boldsymbol{\pi}(\boldsymbol{\rho})\,. \tag{158}$$

However, according to Lemma 21, $\mathbf{a}(\boldsymbol{\rho}) = \mathbf{a}(\bar{\boldsymbol{\rho}})$, $h(\boldsymbol{\rho}) = h(\bar{\boldsymbol{\rho}})$, and $\mathbf{a}(\boldsymbol{\rho})^T\boldsymbol{\rho} = \mathbf{a}(\boldsymbol{\rho})^T\bar{\boldsymbol{\rho}}$. Therefore (158) and (156) are equivalent.

We now move on to proving (157). Using (116), we find that

$$h(\boldsymbol{\rho}+\boldsymbol{\delta}) = \psi\left(\mathbf{a}(\boldsymbol{\rho}+\boldsymbol{\delta}), \boldsymbol{\rho}+\boldsymbol{\delta}\right) \geq \psi\left(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho}+\boldsymbol{\delta}\right) = \psi\left(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho}\right) + \mathbf{a}(\boldsymbol{\rho})^T\boldsymbol{\delta} = h(\boldsymbol{\rho}) + \mathbf{a}(\boldsymbol{\rho})^T\boldsymbol{\delta} \tag{159}$$

and, similarly, that

$$h(\boldsymbol{\rho}) = \psi\left(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho}\right) \geq \psi\left(\mathbf{a}(\boldsymbol{\rho}+\boldsymbol{\delta}), \boldsymbol{\rho}\right) = \psi\left(\mathbf{a}(\boldsymbol{\rho}+\boldsymbol{\delta}), \boldsymbol{\rho}+\boldsymbol{\delta}\right) - \left[\mathbf{a}(\boldsymbol{\rho}+\boldsymbol{\delta})\right]^T\boldsymbol{\delta} = h(\boldsymbol{\rho}+\boldsymbol{\delta}) - \left[\mathbf{a}(\boldsymbol{\rho}+\boldsymbol{\delta})\right]^T\boldsymbol{\delta}\,. \tag{160}$$

Together (159) and (160) imply that

$$0 \leq h(\boldsymbol{\rho}+\boldsymbol{\delta}) - h(\boldsymbol{\rho}) - \mathbf{a}(\boldsymbol{\rho})^T\boldsymbol{\delta} \leq |\boldsymbol{\delta}|\,|\mathbf{a}(\boldsymbol{\rho}+\boldsymbol{\delta}) - \mathbf{a}(\boldsymbol{\rho})|\,. \tag{161}$$

Hence, to complete the proof, we only need to show that $\mathbf{a}$ is continuous.

Equation (159) implies also that $\mathbf{a}(\boldsymbol{\rho})$ is a *subgradient* of $h$ at $\boldsymbol{\rho}$ [Roc70, Section 23, p.214]. The set of all subgradients is called the *subdifferential* of $h$ at $\boldsymbol{\rho}$ and is denoted by $\partial h(\boldsymbol{\rho})$. It is a general result from convex analysis [Roc70, Theorem 24.7] that, because $h$ is convex, the set $\partial h(S) \equiv \bigcup_{\boldsymbol{\rho}\in K}\partial h(\boldsymbol{\rho})$ is bounded whenever $K \subset \mathbb{R}^n$ is bounded. In particular, if $\{\boldsymbol{\rho}_{(i)}\}_{i=1}^{\infty} \subset \mathcal{R}_{\mathbf{m}}$ converges to $\boldsymbol{\rho}_* \in \mathcal{R}_{\mathbf{m}}$, then $\{\mathbf{a}(\boldsymbol{\rho}_{(i)})\}_{i=1}^{\infty}$ is a bounded sequence. Let $\boldsymbol{\alpha}_*$ be any subsequential limit for this sequence. Then

$$\psi(\mathbf{a}(\boldsymbol{\rho}_*), \boldsymbol{\rho}_*) = \lim_{i\to\infty}\psi\left(\mathbf{a}(\boldsymbol{\rho}_*), \boldsymbol{\rho}_{(i_k)}\right) \leq \lim_{i\to\infty}\psi\left(\mathbf{a}(\boldsymbol{\rho}_{(i_k)}), \boldsymbol{\rho}_{(i_k)}\right) \leq \psi(\boldsymbol{\alpha}_*, \boldsymbol{\rho}_*) \leq \psi(\mathbf{a}(\boldsymbol{\rho}_*), \boldsymbol{\rho}_*)\,, \tag{162}$$

where $\{i_k\}_{i=1}^{\infty}$ is any sequence of integers such that $\boldsymbol{\alpha}_* = \lim_{i\to\infty}\mathbf{a}(\boldsymbol{\rho}_{(i_k)})$. The first and last inequalities in (162) follow because $\psi(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho})$ maximizes $\psi(\cdot, \boldsymbol{\rho})$, whereas the middle inequality is a consequence of the fact that $\psi(\cdot, \boldsymbol{\rho})$ is upper semi-continuous (Theorem 11).

From (162), we deduce that $\psi(\boldsymbol{\alpha}_*, \boldsymbol{\rho}) = \psi(\mathbf{a}(\boldsymbol{\rho}_*), \boldsymbol{\rho})$ and since $\mathbf{a}(\boldsymbol{\rho}_*)$ is the *unique* minimizer of $\psi(\cdot, \boldsymbol{\rho})$, it follows that $\boldsymbol{\alpha}_* = \mathbf{a}(\boldsymbol{\rho}_*)$. Because $\{\mathbf{a}(\boldsymbol{\rho}_{(i)})\}$ is bounded and all of its converging subsequences converges to $\mathbf{a}(\boldsymbol{\rho}_*)$, it follows then that

$$\lim_{i\to\infty}\mathbf{a}(\boldsymbol{\rho}_{(i)}) = \mathbf{a}(\boldsymbol{\rho}_*) \tag{163}$$

Thus $\mathbf{a}$ is continuous and $h$ is continuously differentiable. ■

Note that, as a consequence of Theorem 23, $h(\boldsymbol{\rho}) = \psi(\mathbf{a}(\boldsymbol{\rho}), \boldsymbol{\rho})$ is a differentiable on all of $\mathcal{R}_{\mathbf{m}}$ even though $\psi(\cdot, \boldsymbol{\rho})$ may not be continuous for $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$. We alluded to this fact earlier in Section 4.2.

# 6 Geometry of $\mathcal{D}_{\mathbf{m}}$

In this section, we give a description of the geometry of the set $\mathcal{D}$. The main results are given in Theorem 25, which shows that $\mathcal{D}$ is a union of cones, and in Theorem 28, which concludes that, with additional assumptions, $\mathcal{D}$ is small in both a topological and a measure-theoretic sense. We begin with some motivation for why such results are important.

## 6.1 Motivation: Behavior of the Closure Near Degeneracy

Even though $\mathcal{D}_{\mathbf{m}}$ is usually non-empty, there is evidence to suggest that if $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ initially, then densities in $\mathcal{D}_{\mathbf{m}}$ might never be attained. To investigate this possibility, we introduce the function $\chi : \mathcal{R}_{\mathbf{m}} \to \mathbb{R}$, defined by

$$\chi(\boldsymbol{\rho}) \equiv \int_{\mathbb{R}^d} |v\mathbf{m}(v)|\, G_{\mathbf{a}(\boldsymbol{\rho})}(v)\, dv\,. \tag{164}$$

For the entropy-based closure, $\chi$ is closely related to the flux $\mathbf{f}$ in (13), and we show below that $\chi$ becomes unbounded as $\boldsymbol{\rho}$ approaches $\mathcal{D}_{\mathbf{m}}$. As pointed out in [Jun98], such divergent behavior raises the possibility that $\mathcal{R}_{\mathbf{m}}^{\exp}$ is invariant under the dynamics of the closure.

**Proposition 24** *Let $\{\boldsymbol{\rho}_{(j)}\}_{j=1}^{\infty}$ be a sequence in $\mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$ such that $\boldsymbol{\rho}_{(j)} \to \boldsymbol{\rho}_* \in \mathcal{D}_{\mathbf{m}}$, and for each $j$, let $\chi_j \equiv \chi(\boldsymbol{\rho}_{(j)})$. Then $\{\chi_j\}_{j=1}^{\infty}$ is unbounded.*

**Proof.** Since $\{\boldsymbol{\rho}_{(j)}\}_{j=1}^{\infty} \subset \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$,

$$\mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_{(j)})) = \boldsymbol{\rho}_{(j)}\,, \quad j = 1, 2, \dots\ , \tag{165}$$

and taking limits on both sides give

$$\lim_{j\to\infty} \mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_{(j)}) = \boldsymbol{\rho}_* \tag{166}$$

We proceed by showing that *if* $\{\chi_j\}_{j=1}^{\infty}$ is bounded, then

$$\lim_{j\to\infty} \mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_{(j)}) = \mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_*)). \tag{167}$$

Together (166-167) will then imply that $\boldsymbol{\rho}_* \in \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$ which, by contradicting our hypothesis, proves the claim. Hence, suppose that $\{\chi_j\}_{j=1}^{\infty}$ is bounded. To conclude (167), we calculate

$$\begin{aligned}
\left|\mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_*)) - \mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_{(j)})\right| &= \left|\left\langle \mathbf{m}G_{\mathbf{a}(\boldsymbol{\rho}_*)}\right\rangle - \left\langle \mathbf{m}G_{\mathbf{a}(\boldsymbol{\rho}_{(j)})}\right\rangle\right| \\
&\le \int_{\mathbb{R}^d} |\mathbf{m}(v)| \left|G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v) - G_{\mathbf{a}(\boldsymbol{\rho}_{(j)})}(v)\right| dv \\
&= \int_{|v|>R} |\mathbf{m}(v)| \left|G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v) - G_{\mathbf{a}(\boldsymbol{\rho}_{(j)})}(v)\right| dv && (168) \\
&\quad + \int_{|v|<R} |\mathbf{m}(v)| \left|G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v) - G_{\mathbf{a}(\boldsymbol{\rho}_{(j)})}(v)\right| dv && (169)
\end{aligned}$$

where $R > 0$ is an arbitrary constant. We handle the integrals for $|v| > R$ and $|v| < R$ in (168) separately. For $|v| > R$,

$$\begin{aligned}
\int_{|v|>R} |\mathbf{m}(v)| \left|G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v) - G_{\mathbf{a}(\boldsymbol{\rho}_j)}(v)\right| dv &\le \int_{|v|>R} \frac{|v\mathbf{m}(v)|}{R} \left|G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v) - G_{\mathbf{a}(\boldsymbol{\rho}_j)}(v)\right| dv \\
&\le \frac{1}{R}\int_{\mathbb{R}^d} |v\mathbf{m}(v)| \left|G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v) - G_{\mathbf{a}(\boldsymbol{\rho}_j)}(v)\right| dv \le \frac{C}{R}\,, && (170)
\end{aligned}$$

where

$$C \equiv 2\max\left\{\chi(\boldsymbol{\rho}_*), \sup_j\{\chi_j\}\right\}\,. \tag{171}$$

For $|v| < R$, continuity of $\mathbf{a}$ (see Theorem 23) implies $\mathbf{a}(\boldsymbol{\rho}_{(j)}) \to \mathbf{a}(\boldsymbol{\rho}_*)$. Hence the sequence $G_{\mathbf{a}(\boldsymbol{\rho}_{(j)})}$ is uniformly bounded on $\{v \in \mathbb{R}^d : |v| \le R\}$. By the Lebesgue Bounded Convergence Theorem,

$$\lim_{j\to\infty}\int_{|v|<R} |\mathbf{m}(v)| G_{\mathbf{a}(\boldsymbol{\rho}_j)}(v)\,dv = \int_{|v|<R} |\mathbf{m}(v)| G_{\mathbf{a}(\boldsymbol{\rho}_*)}(v)\,dv\,. \tag{172}$$

Together (168), (170), and (172) imply that

$$\lim_{j\to\infty}\left|\mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_*)) - \mathbf{r}(\mathbf{a}(\boldsymbol{\rho}_j)\right| \le \frac{C}{R}\,. \tag{173}$$

Because $R$ can be arbitrarily large, we conclude that (167) holds, which proves the claim. ■

Note that by uniformly bounding $\chi_j$ in the proof above, we are providing uniform control on the highest order moments in $\boldsymbol{\rho}_{(j)}$. In general, such control is not possible, which is why the minimizer in (14) with equality constraints does not always exist. (See the discussion following the proof of Theorem 3.)

The behavior of $\chi$ expressed in Proposition 24 was first observed by Junk for the one dimensional example in [Jun98]. In particular, for a sequence $\{\boldsymbol{\rho}_{(j)}\}_{j=1}^{\infty} \in \operatorname{int}\mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$, it was found that $\left\langle v\mathbf{m}_N G_{\mathbf{a}(\boldsymbol{\rho})}\right\rangle$ diverges to either positive or negative infinity as $\boldsymbol{\rho}_{(j)} \to \boldsymbol{\rho}_* \in \mathcal{D}_{\mathbf{m}}$, with the sign depending on the direction of approach.

Suppose now that it can be proven that $\mathcal{R}_{\mathbf{m}}^{\exp}$ is invariant under the dynamics of the the balance law (12) with the entropy-based closure. Then if $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{\exp}$ initially, the entropy minimization problem with equality constraints (14) will always have a solution and the formal properties of the closure based on the Legendre duality between $h$ and $h^*$ will be maintained. However, it must be shown—at a minimum—that $\mathcal{D}_{\mathbf{m}}$ is small in some sense, thereby limiting the number of initial conditions in $\mathcal{R}_{\mathbf{m}}$ which must be discarded in order to maintain a well-defined closure. In the following subsections, we use the complementary slackness conditions (92) to show that under reasonable hypotheses, $\mathcal{D}_{\mathbf{m}}$ is indeed a Lebesgue measure zero set.

## 6.2 The Complementary Slackness Condition and Normal Cones

From the complementary slackness condition (149b), we obtain the following result.

**Theorem 25** *The set $\mathcal{R}_{\mathbf{m}}$ can be expressed as the following union of cones.*

$$\mathcal{R}_{\mathbf{m}} = \bigcup_{\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}} \bar{\boldsymbol{\rho}} +_{N} \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}})) \,. \tag{174}$$

The proof of this theorem uses the following lemma.

**Lemma 26** *Let $\mathbf{m}$ be a vector whose polynomial components form the basis for an admissible space $\mathbb{M}$. Then $A_{\mathbf{m}} \subset \{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}\}$, where $A_{\mathbf{m}}$ and $A_{\mathbf{m}_N}$ are defined in (36) and (44), respectively.*

**Proof.** Let $\boldsymbol{\alpha} \in A_{\mathbf{m}}$ and let $v_* \in \mathbb{R}^d$ be fixed. Because the components of $\mathbf{m}_i$ are homogeneous polynomials of degree $i$, for any $\lambda > 0$,

$$0 \geq \frac{1}{\lambda^N} \boldsymbol{\alpha}^T \mathbf{m}(\lambda v_*) = \sum_{i=0}^{N} \frac{\lambda^i}{\lambda^N} \boldsymbol{\alpha}_i^T \mathbf{m}_i(v_*) \,. \tag{175}$$

Taking the limit $\lambda \to \infty$ in (175) gives $\boldsymbol{\alpha}_N^T \mathbf{m}_N(v_*) \leq 0$, and since $v_*$ is arbitrary, we conclude that $\boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}$. ■

**Proof of Theorem 25.** Suppose that $\boldsymbol{\rho} \in \mathbb{R}^n$ and that $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp}$. Before making any further assumptions about $\boldsymbol{\rho}$ or any relationship between $\boldsymbol{\rho}$ and $\bar{\boldsymbol{\rho}}$, we note that from Proposition 20, we obtain the following set of equivalent statements:

$$\text{i. } \boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N \in \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))\,; \tag{176a}$$

$$\text{ii. } (\boldsymbol{\alpha}_N - \mathbf{a}_N(\bar{\boldsymbol{\rho}}))^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) \leq 0, \quad \forall\, \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}\,; \tag{176b}$$

$$\text{iii. } \mathbf{a}_N(\bar{\boldsymbol{\rho}})^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) = 0 \quad \text{and} \quad \boldsymbol{\alpha}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) \leq 0, \quad \forall\, \boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}\,. \tag{176c}$$

We first show containment of the left-hand side of (174). Given $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, let $\bar{\boldsymbol{\rho}} = \boldsymbol{\pi}(\boldsymbol{\rho}) \in \mathcal{R}_{\mathbf{m}}^{\exp}$. Then $\boldsymbol{\rho}_j = \bar{\boldsymbol{\rho}}_j$ for $j < N$ and, by (149f), $\mathbf{a}(\bar{\boldsymbol{\rho}}) = \mathbf{a}(\boldsymbol{\rho})$. Thus, from (149b),

$$\mathbf{a}_N(\bar{\boldsymbol{\rho}})^T \boldsymbol{\rho}_N = \mathbf{a}_N(\bar{\boldsymbol{\rho}})^T \bar{\boldsymbol{\rho}}_N \,. \tag{177}$$

Meanwhile, the constraint conditions in (52) imply that

$$\boldsymbol{\alpha}_N^T \boldsymbol{\rho}_N \leq \boldsymbol{\alpha}_N^T \bar{\boldsymbol{\rho}}_N \quad \forall\, \boldsymbol{\alpha} \in A_{\mathbf{m}} \,. \tag{178}$$

We conclude from (176)-(178) that $\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N \in \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))$.

Next we show containment in the other direction. Suppose that $\boldsymbol{\rho}_j = \bar{\boldsymbol{\rho}}_j$ for $j < N$ and that $\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N \in \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))$. By Theorem 1, $\mathcal{R}_{\mathbf{m}} = \operatorname{int} A_{\mathbf{m}}^{\circ}$, so it is sufficient to prove that $\boldsymbol{\rho} \in \operatorname{int} A_{\mathbf{m}}^{\circ}$. Because $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \subset \mathcal{R}_{\mathbf{m}} = \operatorname{int} A_{\mathbf{m}}^{\circ}$, it follows that $\boldsymbol{\alpha}^T \bar{\boldsymbol{\rho}} < 0$ for all $\boldsymbol{\alpha} \in A_{\mathbf{m}}$. Furthermore, by Lemma 26, $\boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}$ for all such $\boldsymbol{\alpha}$. Hence from (178),

$$\boldsymbol{\alpha}^T \boldsymbol{\rho} = \boldsymbol{\alpha}^T \bar{\boldsymbol{\rho}} + \boldsymbol{\alpha}_N^T (\boldsymbol{\rho}_N - \bar{\boldsymbol{\rho}}_N) < 0 \quad \forall\, \boldsymbol{\alpha} \in A_{\mathbf{m}} \,. \tag{179}$$

This shows that $\boldsymbol{\rho} \in \operatorname{int} A_{\mathbf{m}}^{\circ}$ and concludes the proof. ■

For $\bar{\boldsymbol{\rho}} \in \operatorname{int} \mathcal{R}_{\mathbf{m}}^{\exp}$, $\mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))$ is just the origin in $\mathbb{R}^{n_N}$. In such cases, Theorem 25 is trivial and the construction $\bar{\boldsymbol{\rho}} +_N \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))$ does not generate any new densities. Therefore $\mathcal{D}_{\mathbf{m}}$ is constructed entirely by convex cones attached to $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$. Recall from (32), that $\mathcal{NC}_0(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}})) = \mathcal{NC}(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))\backslash\{0\}$. We have the following corollary.

**Corollary 27** *The degenerate densities are*

$$\mathcal{D}_{\mathbf{m}} = \bigcup_{\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}} \{\bar{\boldsymbol{\rho}} +_N \mathcal{NC}_0(A_{\mathbf{m}_N}, \mathbf{a}_N(\bar{\boldsymbol{\rho}}))\} = \bigcup_{\bar{\boldsymbol{\alpha}} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}} \{\mathbf{r}(\bar{\boldsymbol{\alpha}}) +_N \mathcal{NC}_0(A_{\mathbf{m}_N}, \bar{\boldsymbol{\alpha}}_N\} \,. \tag{180}$$

## 6.3 Smoothness Assumptions on $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$

Corollary 27 gives the degenerate densities associated with each $\bar{\boldsymbol{\rho}} \in \mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$. However, a clean description of $\mathcal{D}_{\mathbf{m}}$ requires also that $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$ itself have a nice structure. In particular, we would like to say that $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$ is a finite union of disjoint manifolds. At this point we are unable to prove such a result in general, in part due to the complicated structure of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ to which we alluded in Section 2.4. We therefore make two assumptions. The first assumption says that $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ is a union of disjoint manifolds with dimensional restrictions that are related to the dimensions of the normal cones in (180) in such a way as to ensure that $\mathcal{D}_{\mathbf{m}}$ is a lower dimensional subset of $\mathcal{R}_{\mathbf{m}}$. The second assumption says that the mapping $\mathbf{r}$ is diffeomorphic when restricted to each of these manifolds. Thus each dimension $k$ manifold in $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ will map to a dimension $k$ manifold in $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$. Before stating our assumptions, we define the orthogonal projections $\mathcal{P}_N : \mathbb{R}^n \mapsto \mathbb{R}^{n_N}$ and $\mathcal{P}_{\tilde{N}} : \mathbb{R}^n \mapsto \mathbb{R}^{n-n_N}$ by

$$\mathcal{P}_N(\boldsymbol{\alpha}) \equiv (0, \ldots, 0, 0, \boldsymbol{\alpha}_N^T)^T \quad \text{and} \quad \mathcal{P}_{\tilde{N}}(\boldsymbol{\alpha}) \equiv \boldsymbol{\alpha} - \mathcal{P}_N(\boldsymbol{\alpha}) = (\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, \ldots, \boldsymbol{\alpha}_{N-1}^T, 0)^T \,. \tag{181}$$

**Assumption I** The vector $\mathbf{m}$ is such that the set $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ can be decomposed into a finite collection $\mathcal{S}$ of disjoint, smooth ($C^\infty$) manifolds in $\mathbb{R}^n$. Furthermore, if $S$ is one such manifold, then $\mathcal{P}_N$ projects $S$ onto a manifold $S_N \subset \partial A_{\mathbf{m}_N}$ with co-dimension at least one in $\mathbb{R}^{n_N}$ and $\mathcal{P}_{\tilde{N}}$ projects $S$ onto a manifold $S_{\tilde{N}}$ of co-dimension at least one in $\mathbb{R}^{n-n_N}$.

We call $\mathcal{S}$ a *stratification* of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$; the manifolds $S$ that make up $\mathcal{S}$ are called *strata*. We fully expect that $\mathcal{S}$ can be chosen so that, for each $S \in \mathcal{S}$, the projection $S_N$ is indeed a manifold. If so, $S_N$ will certainly have co-dimension of one or more since, by (46), $S_N \subset \partial A_{\mathbf{m}_N}$. Furthermore, if $\boldsymbol{\alpha}_N \in \partial A_{\mathbf{m}_N}$, then $\boldsymbol{\alpha}_N^T \mathbf{m}_N(\lambda\omega) = 0$ for some $\omega \in \mathbb{S}^{d-1}$ and all $\lambda \in \mathbb{R}$, which means that $\mathbf{m}_N$ no longer provides uniform control over lower degree polynomials. Thus, in order to maintain the integrability condition (40) that defines $\mathcal{A}_{\mathbf{m}}$, we expect further restrictions on the components $\boldsymbol{\alpha}_j$ for $j < N$. This is the motivation for the co-dimension one restriction on the manifold $S_{\tilde{N}}$ in Assumption I. Since, in general, $\dim(S) \leq \dim(S_N) + \dim(S_{\tilde{N}})$, these restrictions together imply that $S$ itself has co-dimension of at least two in $\mathbb{R}^n$.

It should be noted that Assumption I is known to hold for at least two cases:

$$\text{i.} \;\; d = 1 \,\text{and}\; N \geq 2\,; \tag{182a}$$

$$\text{ii.} \;\; d > 1, \;\; N = 4, \;\text{ and }\; \mathbf{m}_4 = |v|^4\,. \tag{182b}$$

(Whether or not Assumption I holds in any other case is, to our knowledge, an open question.) For the first case above, $\boldsymbol{\alpha}_j = \alpha_j$ and $n = N + 1$. For $i = 1, \ldots, N/2$, we define the sets

$$\mathcal{A}_{\mathbf{m}}^{2i} = \{\boldsymbol{\alpha} \in \mathbb{R}^n : \alpha_j = 0 \text{ for } 2i < j \leq N \text{ and } \alpha_{2i} < 0\}\,. \tag{183}$$

Clearly each $\mathcal{A}_{\mathbf{m}}^{2i}$ is a manifold of dimension $2i + 1$ such that

$$\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}} = \bigcup_{i=1}^{N/2-1} \mathcal{A}_{\mathbf{m}}^{2i} \qquad \text{and} \qquad \mathcal{A}_{\mathbf{m}}^{N} = \operatorname{int} \mathcal{A}_{\mathbf{m}}. \tag{184}$$

For the second case, $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}} = \{\boldsymbol{\alpha} \in \mathbb{R}^n : \boldsymbol{\alpha}_2^T \mathbf{m}_2 < 0\}$. If $\mathbf{m}_2 = |v|^2$, then $G_{\boldsymbol{\alpha}}$ has the form of a Maxwellian distribution (133) on $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$; if $\mathbf{m}_2 = v \vee v$, then $G_{\boldsymbol{\alpha}}$ has the form of a Gaussian distribution (138) on $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$.

One possible way to prove that Assumption I always holds is to show that the integrability condition which defines $\mathcal{A}_{\mathbf{m}}$ can be expressed as a family of polynomial equalities and inequalities for $\boldsymbol{\alpha}$. Sets expressed in this way are called *semi-algebraic* and are known to have a stratification with special properties [BR90, Mil68]. One can show for example that the sets $A_{\mathbf{m}_j}$ ($j$ even) and $\operatorname{cl}\mathcal{A}_{\mathbf{m}}$ are semi-algebraic. One can also show that the interiors and boundaries of these sets are semi-algebraic. See [Hau06] for details.

**Assumption II** The vector $\mathbf{m}$ is such that if Assumption I holds and if $S$ is an element of the stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, then for each $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$, the restriction of $\psi(\cdot, \boldsymbol{\rho})$ to $S$ is infinitely Fréchet differentiable on $S$.

One may easily verify that Assumption II also holds for the cases in (182). When both Assumptions I and II hold, $\mathbf{r}$ is a smooth diffeomorphism with inverse $\mathbf{a}$ when restricted to any manifold in the stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$.

## 6.4 Fiber Bundles

The construction of $\mathcal{D}_{\mathbf{m}}$ by attaching cones to the densities $\mathcal{R}_{\mathbf{m}}^{\exp} \cap \partial\mathcal{R}_{\mathbf{m}}^{\exp}$ is very similar to the construction of a fiber bundle. A *(continuous) fiber bundle* $(\mathcal{B}, B, F, \mathcal{P})$ [Hir97] consists of topological spaces $\mathcal{B}$, $B$, and $F$ along with a projection $\mathcal{P} : \mathcal{B} \to B$ such that, for every $y \in B$, there is a neighborhood $O \subset B$ containing $y$ such that $\mathcal{P}^{-1}(O)$ is homeomorphic to $O \times F$. In addition, if $\phi$ is this homeomorphism and $\Pi$ is the natural projection of $O \times F$ onto $O$ (i.e., $\Pi(y \times F) = y$ for all $y \in O$), then $\Pi(\phi(\mathcal{P}^{-1}))$ is the identity on $O$. The space $B$ is called the *base space*; $F$ is called the *fiber space* and often $\mathcal{B}$ itself is called the bundle. Roughly speaking, $\mathcal{B}$ is constructed by attaching to each point in $B$ a (topologically equivalent) copy of $F$ that varies continuously from point to point in the base space. If Assumptions I and II hold, then for each manifold $S$ in a stratification $\mathcal{S}$ of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, the manifold $\mathbf{r}(S)$ acts like a base space; the cones $\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)$, $\boldsymbol{\alpha} \in S$, are like fibers; and $\boldsymbol{\pi}$ is the projection onto the base space. The entire structure is

$$\mathcal{B}(S) = \bigcup_{\boldsymbol{\alpha} \in S} \{\mathbf{r}(\boldsymbol{\alpha}) +_{N} \mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)\} \ , \tag{185}$$

and, in view of Corollary 27, $\mathcal{D}_{\mathbf{m}} = \bigcup_{S \in \mathcal{S}} \mathcal{B}_0(S)$, where

$$\mathcal{B}_0(S) = \mathcal{B}(S) \backslash \mathbf{r}(S) = \bigcup_{\boldsymbol{\alpha} \in S} \{\mathbf{r}(\boldsymbol{\alpha}) +_{N} \mathcal{NC}_0(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)\} \ . \tag{186}$$

Unfortunately, we cannot conclude that $\mathcal{B}(S)$ is a bundle even with Assumptions I and II. In short, we have been unable to show a local homeomorphism between the base-fiber product space and the inverse image $\boldsymbol{\pi}^{-1}(S)$. However the sets taken from the examples in Section 6.6 below are all fiber bundles. This is fairly easy to check because, in these examples, the convex cones $A_{\mathbf{m}_N}$ and $\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)$, $\boldsymbol{\alpha}_N \in \partial A_{\mathbf{m}_N}$ have explicit expressions that are (relatively) simple.

## 6.5 'Smallness' of $\mathcal{D}_{\mathbf{m}}$

If Assumptions I and II hold, we can show that $\mathcal{D}_{\mathbf{m}}$ is small in the following sense.

**Theorem 28** *Suppose Assumptions I and II hold. Then $\mathcal{D}_{\mathbf{m}}$ has zero Lebesgue measure, $\operatorname{int}\mathcal{R}_{\mathbf{m}}^{\exp}$ is a dense subset of $\mathcal{R}_{\mathbf{m}}$, and $\mathcal{D}_{\mathbf{m}} \subset \partial\mathcal{R}_{\mathbf{m}}^{\exp}$.*

**Proof .** The basic idea of the argument is that the image of a smooth map from a lower dimensional space to a higher dimensional space has zero Lebesgue measure. We will construct such a map $F$ whose image covers a portion of $\mathcal{D}_{\mathbf{m}}$. We can then cover $\mathcal{D}_{\mathbf{m}}$ with the images from a countable number of similar maps.

Let $\mathcal{S}$ be a stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ as provided by Assumption I and let $S \in \mathcal{S}$ have dimension $j$. According to Assumption I, $S_N \equiv \mathcal{P}_N S \subset \partial A_{\mathbf{m}_N}$ and $S_{\tilde{N}} \equiv \mathcal{P}_{\tilde{N}} S$ are smooth manifolds with dimensions which we denote by $j_N$ and $j_{\tilde{N}}$, respectively. In general, $\dim(S) \leq \dim(S_N) + \dim(S_{\tilde{N}})$, and in view of Assumption I, $\dim(S_{\tilde{N}}) < n - n_N$. Therefore

$$j \leq j_N + j_{\tilde{N}} < j_N + (n - n_N) \, . \tag{187}$$

This inequality is the key to our result. For any $\boldsymbol{\alpha} \in S$, the normal cone $\mathcal{NC}(S_N, \boldsymbol{\alpha}_N)$ is a subspace of dimension $n_N - j_N$ and one can readily show that

$$\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N) \subset \mathcal{NC}(S_N, \boldsymbol{\alpha}_N)\,, \quad \boldsymbol{\alpha} \in S\,. \tag{188}$$

We can therefore proceed with the proof by considering the set

$$\mathcal{K}_{\mathbf{m}} \equiv \bigcup_{S \in \mathcal{S}} \bigcup_{\boldsymbol{\alpha} \in S} K(\boldsymbol{\alpha}) \supset (\mathcal{R}^{\exp}_{\mathbf{m}} \cap \partial \mathcal{R}^{\exp}_{\mathbf{m}}) \cup \mathcal{D}_{\mathbf{m}}\,, \tag{189}$$

where the affine spaces

$$K(\boldsymbol{\alpha}) \equiv \mathbf{r}(\boldsymbol{\alpha}) +_N \mathcal{NC}(S_N, \boldsymbol{\alpha}_N)\,, \quad \boldsymbol{\alpha} \in S\,, \tag{190}$$

are constructed by attaching $\mathcal{NC}(S_N, \boldsymbol{\alpha}_N)$ to $\mathbf{r}(\boldsymbol{\alpha}) \in \mathbf{r}(S) \subset \mathcal{R}^{\exp}_{\mathbf{m}} \cap \partial \mathcal{R}^{\exp}_{\mathbf{m}}$.

Let $U \subset S$ be the non-empty intersection of $S$ with a bounded open ball in $\mathbb{R}^n$. Because $S$ is a manifold, there exists a smooth diffeomorphism $\tau : U \to \mathbb{R}^j$ such that $\tau(U)$ is the open unit disk $\mathbb{D}^j$. Define a second mapping $\mathbf{V} : U \to \mathbb{R}^{n_N \times (n_N - j_N)}$ such that $\mathbf{V}(\boldsymbol{\alpha})$ is a matrix whose $(n_N - j_N)$ columns are vectors in $\mathbb{R}^{n_N}$ that form a basis for $\mathcal{NC}(S, \boldsymbol{\alpha}_N)$. Since $S$ is smooth, this basis can be chosen to vary smoothly over $\boldsymbol{\alpha} \in U$. Then using $\tau$ and $\mathbf{V}$, define $F : \mathbb{R}^j \times \mathbb{R}^{n_N - j_N} \to \mathbb{R}^n$ by

$$F(\mathbf{y}, \mathbf{b}) \equiv \mathbf{r}(\tau^{-1}(\mathbf{y})) +_N \mathbf{V}(\tau^{-1}(\mathbf{y})) \cdot \mathbf{b}\,. \tag{191}$$

In view of Assumptions II, $F$ is smooth and by (187), $j + (n_N - j_N) < n$. Thus, by [Hir97, Proposition 1.2], the image $F(\mathbb{D}^j \times \mathbb{R}^{n_N - j_N}) = \bigcup_{\boldsymbol{\alpha} \in U} K(\boldsymbol{\alpha})$ has zero Lebesgue measure. Because measure is countably subadditive, repeating this argument for each $j$-ball $U$ in a countable cover of $S$ and then for each $S \in \mathcal{S}$ shows that $\mathcal{K}_{\mathbf{m}}$ has zero Lebesgue measure. Since $\mathcal{D}_{\mathbf{m}} \subset \mathcal{K}_{\mathbf{m}}$, $\mathcal{D}_{\mathbf{m}}$ also has zero Lebesgue measure, and since $\mathcal{R}_{\mathbf{m}} \backslash \mathcal{K}_{\mathbf{m}} \subset \operatorname{int} \mathcal{R}^{\exp}_{\mathbf{m}}$, $\operatorname{int} \mathcal{R}^{\exp}_{\mathbf{m}}$ and $\mathcal{R}_{\mathbf{m}}$ have the same closure. (Otherwise, there would exist an open set of positive measure contained in $\mathcal{K}_{\mathbf{m}}$.) Therefore $\mathcal{D}_{\mathbf{m}} \subset \partial \mathcal{R}^{\exp}_{\mathbf{m}}$. ■

## 6.6 Examples

We will assume that Assumptions I and II hold in the following examples.

1. **Junk's Example.** The case $m_N = |v|^N$ has been studied in [Jun98, Jun00, Sch04], particularly when $N = 4$. For general $N$,

$$A_{\mathbf{m}_N} = \{\boldsymbol{\alpha}_N \in \mathbb{R} : \boldsymbol{\alpha}_N \le 0\} \qquad \text{and} \qquad \partial A_{\mathbf{m}_N} = \{0\}\,. \tag{192}$$

If $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}$ and $\mathbf{a}_N(\boldsymbol{\rho}) = 0$, then $\mathbf{a}_{N-1}(\boldsymbol{\rho}) = 0$ as well; otherwise, $G_{\mathbf{a}(\boldsymbol{\rho})} \notin \mathbb{F}_{\mathbf{m}}$. With this fact in mind, we conclude from Corollary 15 that $G_{\mathbf{a}(\boldsymbol{\rho})}$ is actually the minimizer of $\mathcal{H}$ subject to fewer constraints:

$$\mathcal{H}(G_{\mathbf{a}(\boldsymbol{\rho})}) = \min_{g \in \mathbb{F}_{\mathbf{m}}} \left\{ \mathcal{H}(g) : \langle \mathbf{m}_j g \rangle = \boldsymbol{\rho}_j\,,\ j \le N - 2 \right\}. \tag{193}$$

Let $\bar{\mathbf{m}}$ contain the components of $\mathbf{m}$ of degree $\bar{N} \equiv N - 2$ and less:

$$\bar{\mathbf{m}} \equiv (\mathbf{m}_0, \mathbf{m}_1, \ldots, \mathbf{m}_{N-2})^T\,, \tag{194}$$

and let the variables $\bar{\boldsymbol{\rho}}$ and $\bar{\boldsymbol{\alpha}}$ and the functions $\bar{\mathbf{r}}$ and $\bar{\mathbf{a}}$ be defined similarly. For this example,

$$\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}} \subset \{\boldsymbol{\alpha} \in \mathbb{R}^n : \bar{\boldsymbol{\alpha}} \in \mathcal{A}_{\bar{\mathbf{m}}}\,,\ \boldsymbol{\alpha}_{N-1} = 0\,,\ \boldsymbol{\alpha}_N = 0\}\,, \tag{195}$$

but these two sets are not necessarily equal, since that latter may include $\boldsymbol{\alpha}$ for which $G_{\boldsymbol{\alpha}} \in \mathbb{F}_{\bar{\mathbf{m}}}$, but $G_{\boldsymbol{\alpha}} \notin \mathbb{F}_{\mathbf{m}}$. However, one may readily conclude that $G_{\boldsymbol{\alpha}} \in \mathbb{F}_{\mathbf{m}}$ for all $\bar{\boldsymbol{\alpha}} \in \operatorname{int} \mathcal{A}_{\bar{\mathbf{m}}}$. Hence,

$$\{\boldsymbol{\alpha} \in \mathbb{R}^n : \bar{\boldsymbol{\alpha}} \in \operatorname{int} \mathcal{A}_{\bar{\mathbf{m}}}\,,\ \boldsymbol{\alpha}_{N-1} = 0\,,\ \boldsymbol{\alpha}_N = 0\} \subset \mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}\,. \tag{196}$$

Let $\mathcal{S}$ be a stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial \mathcal{A}_{\mathbf{m}}$. The projection of any manifold $S \in \mathcal{S}$ onto $\partial A_{\mathbf{m}_N}$ is the origin in $\mathbb{R}^{n_N}$, so the normal cone attached to $\boldsymbol{\alpha} \in S$ is just the non-negative axis:

$$\mathcal{NC}(A_{\mathbf{m}_N}, \alpha_N) = \{\boldsymbol{\sigma}_N \in \mathbb{R} : \boldsymbol{\sigma}_N \ge 0\} = A^{\circ}_{\mathbf{m}_N}\,. \tag{197}$$

Therefore

$$\mathcal{D}_{\mathbf{m}} = \{\boldsymbol{\rho} : \boldsymbol{\rho}_N > \mathbf{r}_N(\boldsymbol{\alpha}),\ \boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}\}\ . \tag{198}$$

Because $A_{\mathbf{m}_N}$ is one-dimensional, the inequality in (198) is scalar.

If $N = 4$, the situation simplifies further, because $\operatorname{int} \mathcal{A}_{\bar{\mathbf{m}}} = \mathcal{A}_{\bar{\mathbf{m}}}$ and the inclusion in (195) becomes an equality. In addition, $\mathcal{R}_{\bar{\mathbf{m}}} = \mathcal{R}^{\exp}_{\bar{\mathbf{m}}}$ and $\bar{\mathbf{r}}$ is a diffeomorphism on all of $\mathcal{A}_{\bar{\mathbf{m}}}$. Therefore

$$\begin{aligned} \mathcal{D}_{\mathbf{m}} &= \{\boldsymbol{\rho} : \boldsymbol{\rho}_N > \mathbf{r}_N(\boldsymbol{\alpha}),\ \bar{\boldsymbol{\alpha}} \in \mathcal{A}_{\bar{\mathbf{m}}},\ \boldsymbol{\alpha}_N = \boldsymbol{\alpha}_{N-1} = 0\}\ . \\ &= \left\{\boldsymbol{\rho} : \boldsymbol{\rho}_N > \mathbf{r}_N(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})),\ \boldsymbol{\rho}_{N-1} = \mathbf{r}_{N-1}(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})),\ \bar{\boldsymbol{\rho}} \in \mathcal{R}_{\bar{\mathbf{m}}}\right\}\ .latex \end{aligned} \tag{199}$$

The components $\mathbf{r}_N(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}}))$ and $\mathbf{r}_{N-1}(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}}))$ are simple to compute since $\mathbf{r}(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})) = \langle \mathbf{m} G_{\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})} \rangle$ and $\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})$ has an explicit formula when $\bar{N} = 2$. (See the examples in Section 5.3.)

2. **A Non-Junkian Example.** The situation becomes more complicated when $\mathbf{m}_N$ includes polynomials other than $|v|^N$ because the inequality constraints from the relaxed minimization problem (20) are no longer scalar. The simplest example of this type occurs when

$$\mathbf{m}_N = (v \vee v)\,|v|^{N-2}\,. \tag{200}$$

We examine in detail the two-dimensional case ($d = 2$) and write $\boldsymbol{\alpha}_N \in A_{\mathbf{m}_N}$ in the form of a symmetric matrix:

$$\boldsymbol{\alpha}_N = \begin{pmatrix} (\boldsymbol{\alpha}_N)_{11} & (\boldsymbol{\alpha}_N)_{12} \\ (\boldsymbol{\alpha}_N)_{21} & (\boldsymbol{\alpha}_N)_{22} \end{pmatrix} = \begin{pmatrix} a+b & c \\ c & a-b \end{pmatrix}\ . \tag{201}$$

As a matrix, $\boldsymbol{\alpha}_N$ must be negative definite. Thus, with respect to the $(a, b, c)$ coordinates, the set $A_{\mathbf{m}_N}$ is a cone in $\mathbb{R}^3$ that can be found in a high school geometry text:

$$A_{\mathbf{m}_N} = \left\{(a,b,c) \in \mathbb{R}^3 : a \le -\sqrt{b^2+c^2}\right\} \qquad \text{and} \qquad \partial A_{\mathbf{m}_N} = \left\{(a,b,c) \in \mathbb{R}^3 : a = -\sqrt{b^2+c^2}\right\}\ . \tag{202}$$

Let $\mathcal{S}$ be the stratification of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$ and let $S \in \mathcal{S}$ so that $S_N \in \partial A_{\mathbf{m}_N}$. The set $\partial A_{\mathbf{m}_N}$ itself has a stratification $\mathcal{T}$ consisting of two manifolds: $T_1$ is the origin in $\mathbb{R}^3$ and $T_2$ is the remainder of the cone. We consider $S_N$ as a subset of each manifold separately.

(a) $\boldsymbol{\alpha}_N \in T_1$. In this case, $a = b = c = 0$ and

$$\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N) = A^{\circ}_{\mathbf{m}_N} = \{\boldsymbol{\alpha}_N : \boldsymbol{\alpha}_N \ge 0\}\ . \tag{203}$$

The situation essentially reduces to the Junkian case. The fiber bundle associated with $S \subset \{\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}} : \boldsymbol{\alpha}_N = 0\}$ is

$$\mathcal{B}(S) = \left\{\boldsymbol{\rho} : \boldsymbol{\rho}_N \ge_{A^{\circ}_{\mathbf{m}_N}} \mathbf{r}_N(\boldsymbol{\alpha}),\ \boldsymbol{\alpha} \in S\ \right\}\ . \tag{204}$$

and if $N = 4$,

$$\mathcal{B}(S) = \left\{\boldsymbol{\rho} : \boldsymbol{\rho}_N \ge_{A^{\circ}_{\mathbf{m}_N}} \mathbf{r}_N(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})),\ \boldsymbol{\rho}_{N-1} = \mathbf{r}_{N-1}(0,0,\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})),\ \bar{\boldsymbol{\rho}} \in \mathcal{R}_{\bar{\mathbf{m}}}\right\} \tag{205}$$

where $\bar{\mathbf{a}}(\bar{\boldsymbol{\rho}})$ has an explicit formula. (See the examples in Section 5.3.) However, unlike the Junkian case, the inequalities in (204) and (205) are no longer scalar. Rather, it must be understood in terms of the polar cone $A^{\circ}_{\mathbf{m}_N}$.

(b) $\boldsymbol{\alpha}_N \in T_2$. In this case $a \le -|b| < 0$. In the $(a, b, c)$ coordinates, $\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)$ is a ray

$$\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N) = \left\{\lambda\left(\sqrt{b^2+c^2}, b, c\right) : \lambda \ge 0\right\}\ , \tag{206}$$

which can then be re-expressed in terms of the components of $\boldsymbol{\alpha}_N$ by inverting (201). The bundle associated with any $S \subset \{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}} : \boldsymbol{\alpha}_N \ne 0\}$ is

$$\mathcal{B}(S) = \left\{\boldsymbol{\rho} : \boldsymbol{\rho}_N = \mathbf{r}_N(\boldsymbol{\alpha}) + \mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N), \boldsymbol{\rho}_j = \mathbf{r}_j(\boldsymbol{\alpha})\,,\ \ j < N\,, \boldsymbol{\alpha} \in S\right\}\ . \tag{207}$$

The set $\mathcal{D}_{\mathbf{m}}$ is the union of sets of the form $\mathcal{B}_0(S) = \mathcal{B}(S) \backslash \mathbf{r}(S)$, where $\mathcal{B}(S)$ is a bundle of the type given in (204) or (207).

One should note from these examples that our ability to identify degenerate densities is currently limited by our inability to explicitly identify the elements of $\mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$. However, because the set $A_{\mathbf{m}_N}$ is semialgebraic, one should presumably be able to compute $\mathcal{NC}(A_{\mathbf{m}_N}, \boldsymbol{\alpha}_N)$ for any given $\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{m}} \cap \partial\mathcal{A}_{\mathbf{m}}$, even though such computations will likely be much more tedious than in the examples given above.

# 7 Conclusions and Discussion

We have given in this paper a description of the set $\mathcal{D}_{\mathbf{m}}$ of degenerate densities based on a geometric interpretation of the complementary slackness conditions associated with the dual formulation of (20). Roughly speaking the set $\mathcal{D}_{\mathbf{m}}$ is constructed by attaching a convex cone to every point in the boundary component $\partial\mathcal{R}^{\exp}_{\mathbf{m}} \cap \mathcal{R}^{\exp}_{\mathbf{m}}$. This description recovers and extends previous results concerning the constrained entropy minimization problem.

Analytically, we see three important open questions that must be solved. First, one must determine if Assumptions I and II hold in a setting that is more general than the examples in (182). Concerning Assumption I, this means understanding the structure of the set of polynomials $p$ for which $v \mapsto p(v)e^{p(v)}$ is Lebesgue integrable. For example, do the coefficients of such polynomials form a semialgebraic set? Second, it must be determined whether the sets $\mathcal{R}^{\exp}_{\mathbf{m}}$ and $\mathcal{R}_{\mathbf{m}}$ are invariant under the dynamics of the balance law (12) with the entropy-based closure. (Although not discussed it in this paper, such a condition on $\mathcal{R}_{\mathbf{m}}$ is obviously necessary for entropy-based closures to have any practical application.) Finally, it must be determined whether the existence of degenerate densities and the dynamics of (12) near such densities are simply artifacts of the entropy-based closure or if they actually reflect some physically relevant properties of the original Boltzmann equation (2).

Numerically speaking, a full implementation of entropy-based closures for gas dynamics faces many challenges. (An implementation has been attempted in [TP97], although the issue of degenerate densities was not addressed.) Clearly a discretization of (12) must preserve any invariant properties of $\mathcal{R}_{\mathbf{m}}$ and $\mathcal{R}^{\exp}_{\mathbf{m}}$ with respect to the balance law (12). As pointed out in [Jun98], even if $\mathcal{R}^{\exp}_{\mathbf{m}}$ *is* invariant under (12), solving the dual optimization problem (74) becomes extremely difficult for $\boldsymbol{\rho}$ *near* $\mathcal{D}_{\mathbf{m}}$ because the function $h^*$ is very hard to evaluate. The reason for this is that as $\boldsymbol{\alpha}$ approaches $\partial\mathcal{A}_{\mathbf{m}}$, the function $G_{\boldsymbol{\alpha}}$ can develop isolated modes that are often overlooked in a numerical quadrature. The result is a regularization effect in which accuracy is lost. In addition, the matrix $\langle \mathbf{m}\mathbf{m}^T G_{\boldsymbol{\alpha}} \rangle$ becomes poorly conditioned near the boundary of $\mathcal{A}_{\mathbf{m}}$. Any minimization algorithm for (74) must be carefully formulated in order to overcome these challenges. Furthermore, as with the degenerate densities themselves, one must determine if these difficulties are by-products of the closure or related in some way to the dynamics of the Boltzmann equation.

# 8 Appendix

The purpose of this appendix is to provide the reader with a reference for important notation used in the main body of the paper. We includes tables of important sets and mappings (Tables 1 and 2) and also a diagram (Figure 1) emphasizing the relationships between different sets. Recall that a cone is proper when it is closed, pointed, convex, and has non-empty interior (see Section 2.2).

| *set* | *lies in ...* | *defining equation(s)* | *important properties* |
|---|---|---|---|
| $\mathbb{F}_{\mathbf{m}}$ | $L^1(\mathbb{R}^d)$ | 15 | convex cone; closure is proper |
| $A_{\mathbf{m}_j}$ | $\mathbb{R}^{n_j}$ | 44 | proper cone for $j$ even |
| $A_{\mathbf{m}}$ | $\mathbb{R}^n$ | 36 | proper cone |
| $\mathcal{A}_{\mathbf{m}}$ | $\mathbb{R}^n$ | 40 | $\mathrm{int}(\mathcal{A}_{\mathbf{m}}) = \mathbb{R}^{n-n_N} \times \mathrm{int}\, A_{\mathbf{m}_N}$;<br>$\mathrm{cl}(\mathcal{A}_{\mathbf{m}}) = \mathbb{R}^{n-n_N} \times A_{\mathbf{m}_N}$;<br>$\partial\mathcal{A}_{\mathbf{m}} \subset \mathbb{R}^{n-n_N} \times \partial A_{\mathbf{m}_N}$ |
| $\mathcal{R}_{\mathbf{m}}$ | $\mathbb{R}^n$ | 33 | open, solid, convex cone;<br>$\mathcal{R}_{\mathbf{m}} = \mathrm{int}\, A_{\mathbf{m}}^{\circ}$ |
| $\mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$ | $\mathbb{R}^n$ | 42 | solid cone; in general, not convex or open; $\mathcal{R}_{\mathbf{m}}^{\mathrm{exp}} \subset \mathcal{R}_{\mathbf{m}}$;<br>$[\mathcal{R}_{\mathbf{m}} \subset \mathrm{cl}(\mathrm{int}\, \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}})]$ |
| $\mathcal{D}_{\mathbf{m}}$ | $\mathbb{R}^n$ | 34, 73 | $\mathcal{D}_{\mathbf{m}} = \mathcal{R}_{\mathbf{m}} \backslash \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$; cone;<br>[zero Lebesgue measure] |

Table 1: A list of important sets and properties used in the paper. Properties in brackets are known to hold under Assumptions I and II

| *function* | *domain/range* | *defining equation(s)* | *important properties* |
|---|---|---|---|
| $\mathbf{m}$ | $\mathbb{R}^d \to \mathbb{R}^n$ | 8 | polynomial components; see (21) |
| $\mathcal{H}$ | $\mathbb{F}_{\mathbf{m}} \to \mathbb{R} \cup \{\infty\}$ | 4, 47 | strictly convex and bounded below on $\mathbb{F}_{\mathbf{m}}$ |
| $G_{\boldsymbol{\alpha}}$ | $\mathcal{A}_{\mathbf{m}} \to \mathbb{F}_{\mathbf{m}}$ | 38 | positive; convex on $\mathcal{A}_{\mathbf{m}}$ |
| $\mathbf{r}$ | $\mathcal{A}_{\mathbf{m}} \to \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$ | 41 | bijective on $\mathcal{A}_{\mathbf{m}}$;<br>diffeomorphic on $\mathrm{int}\, \mathcal{A}_{\mathbf{m}}$;<br>derivative of $h^*$ on $\mathrm{int}\, \mathcal{A}_{\mathbf{m}}$ |
| $\mathbf{a}$ | $\mathcal{R}_{\mathbf{m}} \to \mathcal{A}_{\mathbf{m}}$ | 55 | continuous on $\mathcal{R}_{\mathbf{m}}$;<br>diffeomorphic on $\mathrm{int}\, \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$;<br>$\mathbf{a} \circ \mathbf{r}$ is identity on $\mathcal{A}_{\mathbf{m}}$ |
| $h$ | $\mathcal{R}_{\mathbf{m}} \to \mathbb{R}$ | 14, 19, 20, 115 | convex, differentiable on $\mathcal{R}_{\mathbf{m}}$;<br>strictly convex on $\mathrm{int}\, \mathcal{R}_{\mathbf{m}}^{\mathrm{exp}}$;<br>Legendre dual of $h^*$ on $\mathrm{int}\, \mathcal{A}_{\mathbf{m}}$; |
| $h^*$ | $\mathcal{A}_{\mathbf{m}} \to \mathbb{R}$ | 39 | strictly convex; directionally differentiable on $\mathcal{A}_{\mathbf{m}}$;<br>differentiable on $\mathrm{int}\, \mathcal{A}_{\mathbf{m}}$;<br>Legendre dual of $h$ on $\mathrm{int}\, \mathcal{A}_{\mathbf{m}}$;<br>generally not continuous at $\partial\mathcal{A}_{\mathbf{m}} \cap \mathcal{A}_{\mathbf{m}}$ |
| $\mathcal{L}$ | $\mathbb{F}_{\mathbf{m}} \times \mathbb{R}^n \times \mathcal{R}_{\mathbf{m}} \to \mathbb{R} \cup \{\infty\}$ | 74 | strictly convex with respect to first argument |
| $\psi$ | $\mathbb{R}^n \times \mathcal{R}_{\mathbf{m}} \to \mathbb{R} \cup \{-\infty\}$ | 75 | strictly concave;<br>$\psi(\boldsymbol{\alpha}, \boldsymbol{\rho}) = \boldsymbol{\alpha}^T \boldsymbol{\rho} - h^*(\boldsymbol{\alpha})$ |

Table 2: A list of important functions and properties used in the paper.

**Acknowledgements** *The authors would like to thank Dr. Jack Calcut, who was a valuable reference for many technical items in Sections 6.3 - 6.6 related to algebraic geometry, including the proof of Theorem 28. We would also like to thank the referees who provided insightful comments and suggestions. The work of the first author was funded in part by the Department of Mathematics at the University of Maryland, College Park under a National Science Foundation VIGRE grant DMS-0240049 and in part*

$$\begin{tikzcd}
 & \mathbb{R} & & \\
\mathbb{F}_{\mathbf{m}} \arrow[ur, "g\longmapsto\mathcal{H}(g)"] \arrow[rr, shift left, "g\longmapsto\langle\mathbf{m}g\rangle"] & & \mathcal{R}_{\mathbf{m}} \arrow[ul, "h(\boldsymbol{\rho})\longleftarrow\!\shortmid\boldsymbol{\rho}"'] \arrow[ll, shift left, "G_{\mathbf{a}(\boldsymbol{\rho})}\longleftarrow\!\shortmid\boldsymbol{\rho}"] \arrow[r, "\boldsymbol{\rho}\longmapsto\boldsymbol{\pi}(\boldsymbol{\rho})"] \arrow[d, "\boldsymbol{\rho}\longmapsto\mathbf{a}(\boldsymbol{\rho})"'] & \mathcal{R}_{\mathbf{m}}^{\exp} \\
\mathbb{M}_c \arrow[u, "p\longmapsto\exp(p)"] & & \mathcal{A}_{\mathbf{m}} \arrow[ull, "\exp(\boldsymbol{\alpha}^T\mathbf{m})=G_{\boldsymbol{\alpha}}\longleftarrow\!\shortmid\boldsymbol{\alpha}"] \arrow[ll, "\boldsymbol{\alpha}^T\mathbf{m}\longleftarrow\!\shortmid\boldsymbol{\alpha}"'] \arrow[ur, "\boldsymbol{\alpha}\longmapsto\mathbf{r}(\boldsymbol{\alpha})=\big\langle\mathbf{m}\exp(\boldsymbol{\alpha}^T\mathbf{m})\big\rangle"'] &
\end{tikzcd}$$

Figure 1: A commutative diagram summarizing mappings and relationships between important sets.

*by the U.S. Department of Energy at Los Alamos Laboratory under contracts DE-AC52-06NA25396 and the DOE Office of Science Advanced Computing Research (ASCR) Program. The work of the third author was supported in part by the National Science Foundation under Grant DMI-0422931 and by the U.S. Department of Energy under Grant DEFG0204ER25655. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation or the US Department of Energy.*

# References

[AMR03] A. M. Anile, G. Mascali, and V. Romano, *Recent developments in hydrodynamical modeling of semiconductors*, Lecture Notes in Mathematics, vol. 1823, pp. 1–56, Springer-Verlag, Berlin, 2003, Lectures given at the C.I.M.E. Summer School held in Cetraro, Italy on July 15-22, 1998.

[BL91] J. M. Borwein and A. S. Lewis, *Duality relationships for entropy-like minimization problems*, SIAM J. Control Optim. **1** (1991), 191–205.

[BNO03] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, Massachusetts, 2003.

[Bol68] L. Boltzmann, *Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten*, Wien. Ber. **58** (1868), 517–560.

[Bol77] ______, *Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen*, Wien. Ber. **76** (1877), 373–435.

[BR90] R. Benedetti and Jean-Jacques Risler, *Real Algebraic and Semi-Algebraic Sets*, Actualités Mathématiques, Hermann Éditeurs des Sciences et des Arts, Paris, 1990.

[BV04] S. Boyd and L. Vandenderghe, *Convex Optimization*, Cambridge University Press, New York, 2004.

[Cal85] H. B. Callen, *Thermodynamics and an Introduction to Thermostatstics*, second ed., John Wiley and Sons, Inc., New York, 1985.

[CIP94] C. Cercignani, R Ilner, and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Applied Mathematical Sciences 106, Springer-Verlag, New York, 1994.

[Csi75] I. Csiszar, *I-divergence geometry of probability distributions and minimization problems*, Ann. Probab. **3** (1975), no. 1, 146–158.

[DF99] B. Dubroca and Jean-Luc Fuegas, *Étude théorique et numérique d'une hiérarchie de modèles aus moments pour le transfert radiatif*, C.R. Acad. Sci. Paris **I. 329** (1999), 915–920.

[DK02] B. Dubroca and A. Klar, *Half-moment closure for radiative transfer equations*, J. Comput. Phys. **180** (2002), 584–596.

[DR03] P. Degond and C. Ringhofer, *Quantum moment hydrodynamics and the entropy principle*, J. Stat. Phys. **112** (2003), 587–627.

[Eva02] L. C. Evans, *Partial Differential Equations*, Graduate Series in Mathematics 19, American Mathematical Society, Providence, RI, 2002.

[Fol76] G. B. Folland, *Introduction to Partial Differential Equations*, Princeton University Press, Princeton, New Jersey, 1976.

[Gib02] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, Charles Scribner's Son's, New York, 1902.

[Hau06] C. D. Hauck, *Entropy-based moment closures in semiconductor models*, Ph.D. thesis, University of Maryland, College Park, 2006.

[Hir97] M. W. Hirsch, *Differential Topology*, Graduate Texts in Mathematics, vol. 33, Springer-Verlag, New York, 1997.

[Jay57] E. T. Jaynes, *Information theory and statistical mechanics*, Phys. Rev. **106** (1957), no. 4, 620–630.

[JR04] M. Junk and V. Romano, *Maximum entropy systems of the semiconductor Boltzmann equation using Kane's dispersion relation.*, Continuum Mech. Therm. **17** (2004), 247–267.

[Jun98] M. Junk, *Domain of definition of Levermore's five moment system*, J. Stat. Phys. **93** (1998), no. 5-6, 1143–1167.

[Jun00] ———, *Maximum entropy for reduced moment problems*, Math. Mod. Meth. Appl. S. **10** (2000), no. 7, 1001–1025.

[Kul59] S. Kullback, *Information Theory and Statistics*, Wiley, New York, 1959.

[Lev96] C. D. Levermore, *Moment closure hierarchies for kinetic theory*, J. Stat. Phys. **83** (1996), 1021–1065.

[Lev98] ———, *Moment closure hierarchies for the Boltzmann-Poisson equation*, VLSI Design **6** (1998), 97–101.

[Lue69] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley and Sons, Inc., New York, 1969.

[Mil68] J. Milnor, *Singular Points of Complex Hypersurfaces*, Princeton University Press and the University Press of Tokyo, Princeton, New Jersey, 1968.

[Pla00] M. Planck, *Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum*, Verhandlungen der Deutschen Physikalischen Gesellschaft **2** (1900), 237–245.

[Pla01] ———, *Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum*, Annalen der Physik **4** (1901), 553–563.

[Roc70] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[Sch04] J. Schneider, *Entropic approximation in kinetic theory*, Math. Model. Numer. Anal. **38** (2004), 541–561.

[Sha48] C. E. Shannon, *A mathematical theory of communication*, Bell System Tech. J. **27** (1948), 379–423 and 623–656.

[Str04] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, 2nd ed., SIAM, Philadelphia, 2004.

[TP97] P. Le Tallec and J. P. Perlat, *Numerical analysis of Levermore's moment system*, INRIA preprint **3124** (1997), 1–36.