

Relative entropy, Haar measures and relativistic canonical velocity distributions

Jörn Dunkel¹, Peter Talkner and Peter Hänggi

Institut für Physik, Universität Augsburg, Theoretische Physik I,
Universitätsstraße 1, D-86135 Augsburg, Germany
E-mail: joern.dunkel@physik.uni-augsburg.de

New Journal of Physics **9** (2007) 144

Received 8 March 2007

Published 22 May 2007

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/9/5/144

Abstract. The thermodynamic maximum principle for the Boltzmann–Gibbs–Shannon (BGS) entropy is reconsidered by combining elements from group and measure theory. Our analysis starts by noting that the BGS entropy is a special case of relative entropy. The latter characterizes probability distributions with respect to a pre-specified reference measure. To identify the canonical BGS entropy with a relative entropy is appealing for two reasons: (i) the maximum entropy principle assumes a coordinate invariant form and (ii) thermodynamic equilibrium distributions, which are obtained as solutions of the maximum entropy problem, may be characterized in terms of the transformation properties of the underlying reference measure (e.g. invariance under group transformations). As examples, we analyse two frequently considered candidates for the one-particle equilibrium velocity distribution of an ideal gas of relativistic particles. It becomes evident that the standard Jüttner distribution is related to the (additive) translation group on momentum space. Alternatively, imposing Lorentz invariance of the reference measure leads to a so-called modified Jüttner function, which differs from the standard Jüttner distribution by a prefactor, proportional to the inverse particle energy.

¹ Author to whom any correspondence should be addressed.

Contents

1. Introduction	2
2. Thermodynamic entropy, relative entropy and Haar measures	2
2.1. Standard formulation of the MEP	3
2.2. Relative entropy	4
2.3. Choice of the reference measure: group invariance and Haar measures	7
3. Relativistic velocity distributions	9
3.1. Standard Jüttner distribution: momentum translation symmetry	10
3.2. Modified Jüttner distribution: Lorentz symmetry	11
4. Summary	12
Acknowledgment	13
References	13

1. Introduction

The combination of variational principles and group symmetries has proven extremely fruitful in various fields of theoretical physics over the past century, with applications ranging from classical mechanics [1, 2] to quantum field theory [3]–[5]. In this paper, we would like to discuss how group and measure theoretical concepts may be incorporated into the maximum entropy principle (MEP) of canonical equilibrium thermostatics [6].² To this end, we follow up an idea by Ochs [7, 8] who demonstrated that the canonical Boltzmann–Shannon–Gibbs (BGS) entropy is a special case of *relative entropy* (section 2). The relative entropy [9]–[11] characterizes a probability distribution with respect to a pre-specified reference measure and allows a manifestly coordinate invariant formulation of the MEP. In particular, we will focus on how the choice of the reference measure affects the solution of the entropy maximization problem (i.e. the equilibrium distribution). Thereby, it will be clarified that an acceptable MEP must include a postulate that determines which specific reference measure has to be used for a given class of physical systems. To obtain a mathematically meaningful characterization of potential reference measures, one can study their symmetry properties by means of their transformation behaviour under group actions. The idea of combining measure and group theory goes back to the Hungarian mathematician Alfred Haar [12]. In the second part of the paper, this approach will be pursued in order to analyse the MEPs for two of the most frequently discussed candidates for the relativistic one-particle equilibrium velocity distribution (section 3).

2. Thermodynamic entropy, relative entropy and Haar measures

We start out by summarizing the standard formulation of the canonical MEP in section 2.1. We shall focus on the simplest paradigm, corresponding to a spatially homogeneous, ideal gas of classical particles, as this is sufficient for illustrating the main ideas. The concept of relative

² The MEP states that the thermodynamic equilibrium distribution can be obtained by maximizing a suitably chosen entropy functional under a given set of constraints.

entropy is reviewed in section 2.2. The choice of the reference measures and their characterization in terms of symmetry groups is discussed in section 2.3.

2.1. Standard formulation of the MEP

The canonical one-particle equilibrium velocity distribution for a non-relativistic gas of weakly interacting particles (e.g. atoms or molecules) is the Maxwell distribution, corresponding to the normalized probability density function (PDF)

$$f_M(\mathbf{v}) = \left(\frac{m\beta}{2\pi}\right)^{d/2} e^{-\beta m \mathbf{v}^2/2}, \quad \mathbf{v} = (v_1, \dots, v_d) \in \mathcal{V}_d := \mathbb{R}^d. \quad (1)$$

Here, $\beta = 1/T$ is the inverse temperature, m the mass of the particle, and \mathcal{V}_d denotes the space of the d -dimensional Cartesian velocity coordinates (throughout, we use units such that the speed of light $c = 1$, and Boltzmann constant $k_B = 1$). In principle, one can find several different arguments to justify equation (1) [13]; e.g. it can be shown [14] that the marginal one-particle PDF of an isolated, weakly interacting N -particle gas converges to $f_M(\mathbf{v})$ in the thermodynamic limit. An alternative derivation that will be focused on in the remainder of this paper is based on the canonical MEP.

The MEP approach starts from postulating a canonical BGS entropy functional of the form

$$S_B[f] := - \int_{\mathcal{V}_d} d^d \mathbf{v} f(\mathbf{v}) \log [f(\mathbf{v}) \cdot c_0^d], \quad (2a)$$

where $d^d \mathbf{v} = dv_1, \dots, dv_d$ denotes an infinitesimal volume element of the non-relativistic Cartesian velocity space \mathcal{V}_d . The velocity constant c_0 is formally required in equation (2a) to make the argument of the logarithm dimensionless; however, its actual value is irrelevant for the results to be obtained, so that without loss of generality we can fix $c_0 = c = 1$ throughout. The MEP associated with equation (2a) states that the thermodynamic equilibrium distribution f^* is obtained by maximizing $S_B[f]$ with respect to $f \geq 0$ under the normalization and mean energy constraints

$$1 = \int_{\mathcal{V}_d} d^d \mathbf{v} f(\mathbf{v}), \quad \epsilon = \int_{\mathcal{V}_d} d^d \mathbf{v} f(\mathbf{v}) E(\mathbf{v}). \quad (2b)$$

Here, $E(\mathbf{v}) = m\mathbf{v}^2/2$ is the non-relativistic kinetic energy of a single particle (measured in the laboratory-frame), and ϵ the mean energy per particle which is assumed to be known. By means of two Lagrangian multipliers (α, β) , the MEP results in the condition

$$\begin{aligned} 0 &\equiv \frac{\delta}{\delta f} \left\{ S_B[f] + \alpha \left[1 - \int_{\mathcal{V}_d} d^d \mathbf{v} f(\mathbf{v}) \right] + \beta \left[\epsilon - \int_{\mathcal{V}_d} d^d \mathbf{v} f(\mathbf{v}) E(\mathbf{v}) \right] \right\} \Big|_{f=f^*} \\ &= -(1 + \log f^*) - \alpha - \beta E. \end{aligned} \quad (3)$$

Solving this equation for f^* and determining (α, β) from the constraints (2b), one recovers the Maxwellian (1) with parameter $\beta = d/(2\epsilon)$. Hence, the MEP based on equation (2a) appears to be satisfactory at first sight, but a more careful analysis reveals the following drawback: in order to give the empirically established result (1), the BGS entropy (2a) must be written in terms of the ‘correct’ physical variable, and one has to use the ‘correct’ coordinate representation

(in the above case, \mathbf{v} , or some linear transformation as momentum $\mathbf{p} = m\mathbf{v}$, expressed in Cartesian coordinates). Otherwise, one does not obtain the correct one-particle equilibrium distribution (1).

To briefly illustrate this, consider the physically most relevant three-dimensional case ($d = 3$) and suppose that, instead of Cartesian coordinates (v_1, v_2, v_3) , we had started from polar coordinates $(v, \phi, \theta) \in [0, \infty) \times [0, 2\pi) \times [0, \pi] =: \mathbb{P}^3$, i.e., by naively writing

$$S_B[\bar{f}] := - \int_{\mathbb{P}^3} dv d\phi d\theta \bar{f}(v, \phi, \theta) \log \bar{f}(v, \phi, \theta), \quad (4a)$$

where $\bar{f}(v, \phi, \theta)$ is subject to the constraints

$$1 = \int_{\mathbb{P}^3} dv d\phi d\theta \bar{f}(v, \phi, \theta), \quad \epsilon = \int_{\mathbb{P}^3} dv d\phi d\theta \bar{f}(v, \phi, \theta) \bar{E}(v, \phi, \theta), \quad (4b)$$

and $\bar{E}(v, \phi, \theta) = mv^2/2$ is the energy expressed in polar coordinates. Maximizing $S_B[\bar{f}]$ under the constraints (4b) yields

$$\bar{f}^*(v, \phi, \theta) = \left(\frac{m\beta}{2\pi}\right)^{d/2} e^{-\beta mv^2/2}. \quad (5)$$

For comparison, by transforming the Maxwell PDF (1) to polar coordinates we find

$$\bar{f}_M(v, \phi, \theta) = \bar{J} \left(\frac{m\beta}{2\pi}\right)^{d/2} e^{-\beta mv^2/2}, \quad \bar{J} = v^2 \sin \theta, \quad (6)$$

where \bar{J} is the Jacobian of the coordinate transformation $(v_1, v_2, v_3) \mapsto (v, \phi, \theta)$. Upon comparing equations (6) and (5), we observe that $\bar{f}^* \neq \bar{f}_M$, due to the missing Jacobian prefactor in equation (5).

This simple example illustrates that the above entropy definition is implicitly coordinate dependent. This fact is somewhat unsatisfactory. If viewed as fundamental, then the MEP should be formulated in a form that works independently of the underlying coordinate representation. As we shall discuss next, this can be achieved by recognizing that the thermodynamic entropy (2a) is a special case of the so-called *relative entropy* [7]–[9].

2.2. Relative entropy

First, we summarize the definition of the relative entropy [7]–[9] and demonstrate its invariance under coordinate transformations. Subsequently, it will be shown how the BGS entropy (2a) is embedded into this concept.

Consider some set $\mathcal{X} \subseteq \mathbb{R}^d$ and two measures μ and ν on \mathcal{X} that are absolutely continuous with respect to each other (i.e. μ and ν have the same null sets in \mathcal{X} [15]). The *relative entropy* of μ with respect to ν is defined by³

$$S[\mu|\nu] := - \int_{\mathcal{X}} d\mu \log \frac{d\mu}{d\nu}(\mathbf{x}) = - \int_{\mathcal{X}} d\nu f_{\mu|\nu}(\mathbf{x}) \log f_{\mu|\nu}(\mathbf{x}), \quad (7a)$$

³ This definition of *relative entropy*, sometimes also referred to as generalized BGS entropy [7]–[9], does not only apply to continuous but also to discrete spaces \mathcal{X} . Depending on the problem under consideration, \mathcal{X} may be the one-particle or many-particle phase space, the velocity space, the configuration space, etc.

where the function

$$f_{\mu|\nu}(\mathbf{x}) := \frac{d\mu}{d\nu}(\mathbf{x}) > 0 \quad (7b)$$

is the so-called Radon–Nikodym density [15] of μ with respect to ν . The measure ν plays the role of a *reference measure*. We briefly illustrate the meaning of the Radon–Nikodym density by two simple examples.

The most prominent measure on \mathbb{R}^d is the Lebesgue measure, denoted by λ [15]. The measure λ assigns to any d -dimensional rectangular parallel-epiped $\mathcal{J}_d := [a_1, b_1] \times \cdots \times [a_d, b_d] \subset \mathbb{R}^d$ the intuitive measure number $\lambda(\mathcal{J}_d) = \prod_{i=1}^d (b_i - a_i)$, where it is assumed that $b_i > a_i$ holds $\forall i = 1, \dots, d$. If, for example, μ is a probability measure on $\mathcal{X} \subseteq \mathbb{R}^d$, then the Radon–Nikodym density $f_{\mu|\lambda}(\mathbf{x})$ of μ with respect to λ is the ‘ordinary’ PDF of μ .

As the second example, consider two measures μ, ν on $\mathcal{X} \subseteq \mathbb{R}^d$ with non-vanishing densities $f_{\mu|\lambda} > 0$ and $f_{\nu|\lambda} > 0$ on \mathcal{X} . In this case, the Radon–Nikodym density of μ with respect to ν is given by the quotient of their densities, i.e.,

$$f_{\mu|\nu}(\mathbf{x}) := \frac{d\mu}{d\nu}(\mathbf{x}) = \frac{f_{\mu|\lambda}(\mathbf{x})}{f_{\nu|\lambda}(\mathbf{x})}. \quad (8)$$

Accordingly, we may rewrite the relative entropy (7a) in terms of the two densities $f_{\mu|\lambda}$ and $f_{\nu|\lambda}$ as

$$S[\mu|\nu] = -\int_{\mathcal{X}} d\lambda f_{\mu|\lambda}(\mathbf{x}) \log \frac{f_{\mu|\lambda}(\mathbf{x})}{f_{\nu|\lambda}(\mathbf{x})} = -\int_{\mathcal{X}} d^d \mathbf{x} f_{\mu|\lambda}(\mathbf{x}) \log \frac{f_{\mu|\lambda}(\mathbf{x})}{f_{\nu|\lambda}(\mathbf{x})} =: s[f_{\mu|\lambda}|f_{\nu|\lambda}]. \quad (9)$$

In the second line, we have inserted the equivalent notation $d^d \mathbf{x}$ for the Lebesgue measure $d\lambda$ of an infinitesimal volume element in \mathbb{R}^d . Equation (9) will provide the basis for all subsequent considerations. We note that, in order to define relative entropy, it is *a priori* not required that the measures μ and ν are normalizable on $\mathcal{X} \subseteq \mathbb{R}^d$; it suffices to assume that they have the same null sets, i.e., $f_{\mu|\lambda}(\mathbf{x}) = 0$ implies $f_{\nu|\lambda}(\mathbf{x}) = 0$ and vice versa, so that the argument of the logarithm is well-defined.⁴

Before discussing how the BGS entropy (2a) arises as a special case of equation (9), it is useful to give the general, coordinate invariant form of the MEP with equation (9) serving as the starting point. For this purpose, we impose the constraints

$$1 = \int_{\mathcal{X}} d\mu = \int_{\mathcal{X}} d^d \mathbf{x} f_{\mu|\lambda}(\mathbf{x}), \quad (10a)$$

$$\epsilon = \int_{\mathcal{X}} d\mu E(\mathbf{x}) = \int_{\mathcal{X}} d^d \mathbf{x} f_{\mu|\lambda}(\mathbf{x}) E(\mathbf{x}), \quad (10b)$$

where $E > 0$ is a non-negative ‘energy’ function.⁵ Maximizing $S[\mu|\nu] = s[f_{\mu|\lambda}|f_{\nu|\lambda}]$ with respect to μ or, equivalently, with respect to $f_{\mu|\lambda}$, and taking into account the constraints (10a) and (10b),

⁴ Conventionally, we set $\frac{d\mu}{d\nu}(\mathbf{x}) = 1$, if $f_{\mu|\lambda}(\mathbf{x}) = f_{\nu|\lambda}(\mathbf{x}) = 0$.

⁵ In principle, one could also include more than two constraints.

leads to the condition

$$0 \equiv 1 + \log \frac{f_{\mu|\lambda}^*}{f_{\nu|\lambda}} + \alpha + \beta E. \quad (11)$$

Similar to equation (3), α and β have entered here as Lagrangian multipliers for the normalization and ‘energy’ constraints, respectively. From equation (11) the solution of the variational problem is obtained as

$$f_{\mu|\lambda}^*(\mathbf{x}) = f_{\nu|\lambda}(\mathbf{x}) e^{-(\alpha+1)-\beta E(\mathbf{x})}. \quad (12)$$

The parameters (α , β) are determined by means of the conditions (10a) and (10b). As it is evident from equation (12), the ‘equilibrium’ PDF $f_{\mu|\lambda}^*$ depends on the choice of the reference density $f_{\nu|\lambda}(\mathbf{x})$.

We next show that the relative entropy definition (9) is manifestly coordinate invariant. For this purpose, consider a change of coordinates $\mathbf{x} \mapsto \bar{\mathbf{x}}$, and denote by $\bar{\mathcal{X}}$ the range of the new coordinates. Using the following standard formulae for the transformation of volume elements and densities f :

$$d^d \mathbf{x} = \left(\frac{\partial \mathbf{x}}{\partial \bar{\mathbf{x}}} \right) d^d \bar{\mathbf{x}}, \quad \bar{f}(\bar{\mathbf{x}}) = \left(\frac{\partial \mathbf{x}}{\partial \bar{\mathbf{x}}} \right) f(\mathbf{x}(\bar{\mathbf{x}})), \quad (13)$$

where $\bar{J} = (\partial \mathbf{x} / \partial \bar{\mathbf{x}})$ is the Jacobian of the coordinate transformation, we find that

$$s[f_{\mu|\lambda} | f_{\nu|\lambda}] = - \int_{\mathcal{X}} d^d \mathbf{x} f_{\mu|\lambda}(\mathbf{x}) \log \frac{f_{\mu|\lambda}(\mathbf{x})}{f_{\nu|\lambda}(\mathbf{x})} = - \int_{\bar{\mathcal{X}}} d^d \bar{\mathbf{x}} \bar{f}_{\mu|\lambda}(\bar{\mathbf{x}}) \log \frac{\bar{f}_{\mu|\lambda}(\bar{\mathbf{x}})}{\bar{f}_{\nu|\lambda}(\bar{\mathbf{x}})} = s[\bar{f}_{\mu|\lambda} | \bar{f}_{\nu|\lambda}]. \quad (14)$$

Hence, the relative entropy is indeed independent of the choice of the coordinates, due the fact that the Jacobians in the argument of the logarithm cancel. As a consequence, the solution of the associated MEP becomes coordinate independent as well. To demonstrate this more explicitly, we first rewrite the constraint function E in terms of the new coordinates by defining $\bar{E}(\bar{\mathbf{x}}) := E(\mathbf{x}(\bar{\mathbf{x}}))$. Then, the constraints (10a) and (10b) may be expressed equivalently in the new coordinates as

$$1 = \int_{\mathcal{X}} d^d \mathbf{x} f_{\mu|\lambda}(\mathbf{x}) = \int_{\bar{\mathcal{X}}} d^d \bar{\mathbf{x}} \bar{f}_{\mu|\lambda}(\bar{\mathbf{x}}), \quad (15a)$$

$$\epsilon = \int_{\mathcal{X}} d^d \mathbf{x} f_{\mu|\lambda}(\mathbf{x}) E(\mathbf{x}) = \int_{\bar{\mathcal{X}}} d^d \bar{\mathbf{x}} \bar{f}_{\mu|\lambda}(\bar{\mathbf{x}}) \bar{E}(\bar{\mathbf{x}}). \quad (15b)$$

Hence, the solution of the associated variational problem reads

$$\bar{f}_{\mu|\lambda}^*(\bar{\mathbf{x}}) = \bar{f}_{\nu|\lambda}(\bar{\mathbf{x}}) e^{-(\alpha+1)-\beta \bar{E}(\bar{\mathbf{x}})} = \left(\frac{\partial \mathbf{x}}{\partial \bar{\mathbf{x}}} \right) f_{\nu|\lambda}(\mathbf{x}(\bar{\mathbf{x}})) e^{-(\alpha+1)-\beta E(\mathbf{x}(\bar{\mathbf{x}}))} = \left(\frac{\partial \mathbf{x}}{\partial \bar{\mathbf{x}}} \right) f_{\mu|\lambda}^*(\mathbf{x}(\bar{\mathbf{x}})). \quad (16)$$

This is indeed the correct transformation law for the equilibrium PDF $f_{\mu|\lambda}^*$ from equation (12); i.e., once the reference measure ν and its density are properly specified, the MEP and its solution become independent of the choice of the coordinates.

Finally, it is straightforward to see that the BGS entropy (2a) is a special case of equation (9): we identify $\mathcal{X} = \mathcal{V}_d = \mathbb{R}^d$ and fix the reference measure as the Lebesgue measure in velocity

space $\nu = \lambda$. Then, taking into account that $f_{\lambda|\lambda}(\mathbf{v}) \equiv 1$, equation (9) reduces to the BGS entropy (2a) i.e., explicitly,

$$S[\mu|\lambda] = - \int_{\mathcal{V}_d} d^d \mathbf{v} f_{\mu|\lambda}(\mathbf{v}) \log f_{\mu|\lambda}(\mathbf{v}) = S_B[f]. \quad (17)$$

We thus note that the canonical BGS entropy corresponds to a specific choice of the reference measure, namely, the Lebesgue measure in velocity space. Put differently, whenever one writes an entropy in the ‘standard’ form (17), one has implicitly fixed an underlying reference measure (defined with respect to some set of primary variables).

With regard to the subsequent discussion it will be important to keep in mind that the solution (12) of the coordinate invariant MEP is determined by two ingredients: (i) the ‘energy’ function E that specifies the mean value constraint and (ii) the underlying reference measure ν . While usually the energy function E is known, it is a not-so-trivial problem to identify the appropriate reference measure ν for a given class of physical systems. In the next section, we are going to discuss how one can classify reference measures according to their transformation properties under symmetry groups.

2.3. Choice of the reference measure: group invariance and Haar measures

The above discussion shows that the MEP is incomplete unless one is able to specify the reference measure ν on the state space \mathcal{X} . Put differently, before accepting the MEP as a truly fundamental principle, one has to find a general method that allows to determine ν for a given class of dynamical systems. A promising step towards solving this problem is to analyse potential reference measures with respect to their invariance properties under fundamental symmetry transformations. Conceptually, this idea is closely related to the theory of Haar measures [12, 13]. In a seminal paper [12] published in 1933, the Hungarian mathematician Alfred Haar studied the possibility to introduce a measure μ_\circ on a continuous group (\mathcal{G}, \circ) such that μ_\circ is invariant under the group multiplication ‘ \circ ’. To briefly sketch this idea, consider a subset \mathcal{A} of the group \mathcal{G} and some arbitrary, fixed group element $g \in \mathcal{G}$. By multiplying each element $a \in \mathcal{A}$ with g , the subset \mathcal{A} is mapped on to another subset of \mathcal{G} , denoted by

$$g \circ \mathcal{A} := \{g \circ a \mid a \in \mathcal{A}\}. \quad (18)$$

Now consider a measure μ_\circ on \mathcal{G} that assigns to $\mathcal{A} \subseteq \mathcal{G}$ some non-negative real number $\mu_\circ(\mathcal{A})$. The measure μ_\circ is said to be group invariant, if⁶

$$\mu_\circ(g \circ \mathcal{A}) = \mu_\circ(\mathcal{A}) \quad (19)$$

holds for any $g \in \mathcal{G}$ and $\mathcal{A} \subseteq \mathcal{G}$. Haar was able to prove the existence of an invariant measure μ_\circ , and its uniqueness apart from an irrelevant multiplicative constant for locally compact, topological groups. Such group invariant measures μ_\circ are referred to as Haar measures nowadays [15]. They give a mathematically precise meaning to the notion ‘uniform distribution’ by combining measure and group theoretical concepts. However, in physics one often encounters the slightly different situation, where a certain symmetry group acts on the domain \mathcal{X} of a vector space, e.g. by means of a matrix representation. In this case, it is a natural to extend

⁶ In the case of non-commutative (i.e. non-Abelian) groups, one may distinguish invariance under multiplications from the right or left.

the original ideas of Haar by considering measures on \mathcal{X} that are invariant under the group action.⁷

In order to link these concepts to thermodynamics, we return to the BGS entropy (17). This ‘canonical’ entropy was identified above as the relative entropy with respect to the Lebesgue measure λ on the non-relativistic velocity space $\mathcal{V}_d = \mathbb{R}^d$. Adopting the group-theoretical point of view, the *defining* property of the Lebesgue measure is given by the fact that λ is the only⁸ translation invariant measure on \mathcal{V}_d . To capture this fact more precisely, we define \mathbf{w} -parametrized translations $G_{\mathbf{w}}$ on \mathcal{V}_d by means of

$$G_{\mathbf{w}} : \mathbf{v} \mapsto \mathbf{w} + \mathbf{v}, \quad \forall \mathbf{w}, \mathbf{v} \in \mathcal{V}_d. \quad (20)$$

The velocity translations $G_{\mathbf{w}}$ form a group by means of the composition rule

$$G_{\mathbf{w}_1} \circ G_{\mathbf{w}_2} = G_{\mathbf{w}_1 + \mathbf{w}_2}. \quad (21)$$

Now consider some subdomain $\mathcal{A} \subset \mathcal{V}_d$ and define the translation $G_{\mathbf{w}}[\mathcal{A}]$ of \mathcal{A} by

$$G_{\mathbf{w}}[\mathcal{A}] := \{\mathbf{w} + \mathbf{v} \mid \mathbf{v} \in \mathcal{A}\}. \quad (22)$$

Then the Lebesgue measure λ is the only measure satisfying [15]

$$\lambda(G_{\mathbf{w}}[\mathcal{A}]) = \lambda(\mathcal{A}), \quad \forall \mathcal{A} \subset \mathcal{V}_d, \mathbf{w} \in \mathcal{V}_d, \quad (23a)$$

or, equivalently, in differential notation

$$d^d \mathbf{v} = d^d \mathbf{v}' \quad (23b)$$

where $\mathbf{v}' = G_{\mathbf{w}}(\mathbf{v}) = \mathbf{w} + \mathbf{v}$. This very translation invariance distinguishes the Lebesgue measure from all other measures that can be introduced on \mathcal{V}_d . Adopting the physicist’s point of view, the translation map (20) may be interpreted in two different ways.

1. Geometric interpretation: equation (20) describes a Galilei transformation, corresponding to a change of inertial coordinate systems.
2. Kinetic interpretation: equation (20) describes a momentum conservation law, with $\Delta \mathbf{p} = m \mathbf{w}$ corresponding to the particle’s momentum gain in a collision.

Both interpretations are equally plausible here, because non-relativistic momentum and velocity differ by a mass constant m only; in particular, the Lebesgue measure in velocity space transforms to a Lebesgue measure in momentum space, when changing from velocity to momentum coordinates in the non-relativistic case. However, regardless of this ambiguity in the interpretation of equation (20), it is evident that the Lebesgue measure in velocity space (or, equivalently, in momentum space) plays a distinguished role in non-relativistic physics: *it is the Haar measure of the Galilei group (or, equivalently, of the momentum translation group)*. This might explain why only the relative entropy with respect to this particular measure, $S[\mu|\lambda]$, yields the correct non-relativistic equilibrium distribution (1).

⁷ For example, in the one-dimensional case $d = 1$ the proper-orthochronous Lorentz group \mathcal{L}_+^\uparrow consists of boosts only and, therefore, it can be identified with the relativistic velocity space $\mathcal{RV}_1 = (-1, 1)$; hence, the action of \mathcal{L}_+^\uparrow on \mathcal{RV}_1 is just the action of \mathcal{L}_+^\uparrow on itself. This corresponds to the framework originally considered by Haar [12]. By contrast, in higher space dimensions $d > 1$ it is not possible anymore to identify the relativistic velocity space $\mathcal{RV}_d = \{\mathbf{v} \in \mathbb{R}^d \mid |\mathbf{v}| < 1\}$ directly with a subgroup of the Lorentz group, since then the number of group parameters is larger than d (cf chapter 6 in [16]). Nevertheless, also in this case one can find a Lorentz invariant measure on \mathcal{RV}_d , which is unique apart from an irrelevant multiplicative constant; cf discussion in section 3.

⁸ We omit the phrase ‘apart from an irrelevant multiplicative constant’ from now on.

In the remainder of this paper, we are going to study generalizations of the Maxwell distribution (1) in the framework of special relativity. In particular, we shall identify the reference measures underlying two of the most commonly considered relativistic one-particle equilibrium distributions.

3. Relativistic velocity distributions

Six years after Einstein [17, 18] had formulated his theory of special relativity, Ferencz Jüttner [19] presented in 1911 the first detailed study on the canonical thermostatics of a relativistic (quasi-)ideal gas of classical particles.⁹ As the main result of his paper, he proposed the following three-dimensional relativistic generalization of Maxwell's non-relativistic momentum distribution [19]–[21]:

$$\phi_J(\mathbf{p}) = \frac{\exp(-\beta E)}{\mathcal{Z}_0}, \quad \mathbf{p} \in \mathcal{RP}_3 := \mathbb{R}^3 \quad (24)$$

with $\beta = 1/T$ being the inverse temperature parameter, and

$$E = (m^2 + \mathbf{p}^2)^{1/2} = m\gamma(\mathbf{v}), \quad \mathbf{p} = m\mathbf{v}\gamma(\mathbf{v}), \quad (25)$$

the relativistic energy and the relativistic momentum with Lorentz factor $\gamma(\mathbf{v}) = (1 - \mathbf{v}^2)^{-1/2}$ (we continue to use units $k_B = c = 1$). The d -dimensional relativistic momentum space is denoted by \mathcal{RP}_d . The constant \mathcal{Z}_0 is determined by the normalization condition

$$1 = \int_{\mathbb{R}^3} d^3\mathbf{p} \phi_J(\mathbf{p}), \quad (26)$$

and, in the three-dimensional case, one finds [19]

$$\mathcal{Z}_0 = 4\pi m^3 \frac{K_2(\beta m)}{\beta m}, \quad (27)$$

where K_ν denotes the modified Bessel function of the second kind. The one-particle momentum distribution (24) refers to a laboratory rest frame, where the container enclosing the gas is at rest. As usual, it is assumed that for an ordinary hard box potential the spatial part of the one-particle *phase space* PDF is trivial (i.e. constant), corresponding to a spatially homogeneous particle distribution in the box.

The Jüttner function ϕ_J has been widely used in high energy and astrophysics over past decades [22]–[24]. However, in recent years several authors [25]–[29] argued that equation (24) might not represent the correct relativistic equilibrium distribution, and several alternatives were suggested. Generalizing to an arbitrary number of space dimensions d , the proposed candidates can be summarized in terms of the following η -parameterized momentum PDF:

$$\phi_\eta(\mathbf{p}) = \frac{\exp(-\beta E)}{\mathcal{Z} E^\eta}, \quad \mathbf{p} \in \mathcal{RP}_d = \mathbb{R}^d, \quad \eta \geq 0. \quad (28)$$

The normalization constant \mathcal{Z} depends on both η and d . For $\eta = 0$ the PDF (28) reduces to the standard Jüttner function (24), $\phi_0 \equiv \phi_J$. The most frequently considered modification corresponds to $\eta = 1$ [25]–[30], while one author [28] has also included the case $\eta = 2$. Compared with the Jüttner value $\eta = 0$, larger values $\eta > 0$ diminish the probability of particles having high

⁹ Jüttner's paper [19] appeared about 20 years prior to Haar's work on group invariant measures [12].

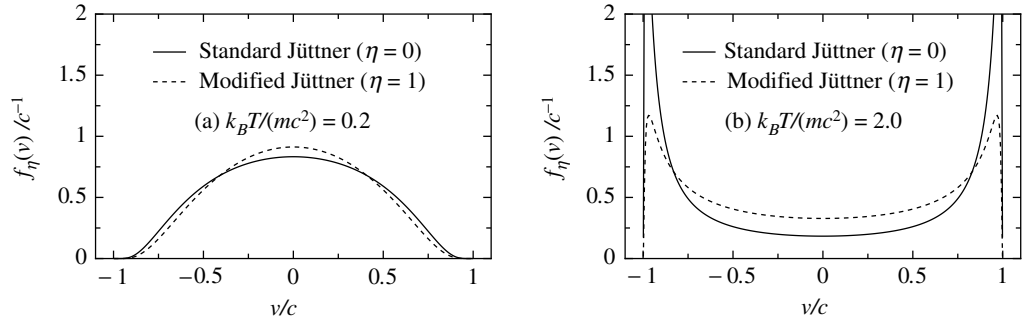


Figure 1. Relativistic velocity PDFs from equation (29) for the one-dimensional case $d = 1$. (a) At low temperatures $T \ll m$ both distributions approach a quasi-Gaussian (Maxwellian) shape. (b) At high temperatures $T \gg m$ the distributions become bimodal. For the same temperature value T , the standard Jüttner distribution with $\eta = 0$ exhibits a higher probability of large absolute velocities than the modified distribution with $\eta = 1$.

absolute momentum at same temperature $T = 1/\beta$. The one-particle velocity PDF corresponding to equation (28) is given by

$$f_\eta(\mathbf{v}) = \frac{m^{d-\eta} \gamma^{2+d}(\mathbf{v})}{\mathcal{Z}} \exp[-\beta m \gamma(\mathbf{v})] \quad (29)$$

with \mathbf{v} taking values in the relativistic velocity space $\mathcal{R}\mathcal{V}_d := \{\mathbf{v} \in \mathbb{R}^d \mid |\mathbf{v}| < 1\}$. Below we focus on the two most frequently considered values $\eta = 0$ (standard Jüttner distribution) and $\eta = 1$ (modified Jüttner distribution). Figure 1 shows the corresponding velocity PDFs f_0 and f_1 at two different temperature values for the one-dimensional case $d = 1$.

In the remainder, we will analyse the MEPs that give rise to the standard and modified Jüttner distributions, respectively. In particular, the different underlying reference measures shall be characterized by means of their invariance under group actions.

3.1. Standard Jüttner distribution: momentum translation symmetry

We first consider the MEP for the standard Jüttner distribution with $\eta = 0$. As discussed in section 2.2, the MEP becomes coordinate independent if expressed in terms of relative entropy. In the relativistic case, it is most convenient to use the momentum coordinate $\mathbf{p} \in \mathcal{R}\mathcal{P}_d := \mathbb{R}^d$. The Lebesgue measure on relativistic momentum space $\mathcal{R}\mathcal{P}_d$ has, by definition, a constant density denoted by ℓ . Without loss of generality, we choose the normalization $\ell(\mathbf{p}) = (mc)^{-d} = m^{-d}$ so that the integral of ℓ over some finite subset of $\mathcal{R}\mathcal{P}_d$ is a dimensionless number. With these preliminaries, we can state the MEP for the standard Jüttner function: maximization of the relative entropy

$$s_0[\phi|\ell] = - \int_{\mathcal{R}\mathcal{P}_d} d^d \mathbf{p} \phi(\mathbf{p}) \log \left[m^d \phi(\mathbf{p}) \right] \quad (30a)$$

under the constraints

$$1 = \int_{\mathcal{R}\mathcal{P}_d} d^d \mathbf{p} \phi(\mathbf{p}), \quad \epsilon = \int_{\mathcal{R}\mathcal{P}_d} d^d \mathbf{p} \phi(\mathbf{p}) E(\mathbf{p}), \quad (30b)$$

where now $E = (m^2 + \mathbf{p}^2)^{1/2}$ is the relativistic energy, yields the standard Jüttner distribution ϕ_J , corresponding to $\eta = 0$ in equation (28). It may be worth noting that, in the relativistic case, the Lebesgue measure on \mathcal{RP}_d does *not* transform into a Lebesgue measure on the relativistic velocity space \mathcal{RV}_d due to the nonlinear momentum–velocity relation $\mathbf{p} = m\mathbf{v}\gamma(\mathbf{v})$. Hence, if one rewrites the relative entropy s_0 in terms of the velocity \mathbf{v} , an additional determinant factor enters in the argument of the logarithm.

We now turn to the invariance properties of the specific reference measure, required to obtain the standard Jüttner distribution with $\eta = 0$. Analogous to the discussion in section 2.3, the Lebesgue measure in relativistic momentum space is singled out by the fact that it is the only translation invariant measure in momentum space; i.e., it is the Haar measure of the momentum translation group. Hence, the standard Jüttner function is consistent with the *kinetic interpretation* in section 2.3. Put differently, if the Jüttner function turns out to be the correct relativistic one-particle equilibrium distribution, then the maximum principle for the relative entropy should be completed by the postulate that the reference measure must be translation invariant in momentum space.

3.2. Modified Jüttner distribution: Lorentz symmetry

As the second example, we consider the modified Jüttner distribution with $\eta = 1$ in equation (28). It is straightforward to verify that this distribution is obtained by maximizing the relative entropy

$$s_1[\phi|\rho] = - \int_{\mathcal{RP}_d} d^d \mathbf{p} \phi(\mathbf{p}) \log \left[\frac{\phi(\mathbf{p})}{\rho(\mathbf{p})} \right], \quad \rho(\mathbf{p}) = 1/E \quad (31)$$

under the constraints (30b). In contrast to equation (30a), the reference density $\rho = 1/E$ is momentum dependent. The measure χ associated with ρ assigns to any subset $\mathcal{A} \subset \mathcal{RP}_d$ the measure number

$$\chi(\mathcal{A}) = \int_{\mathcal{A}} d^d \mathbf{p} \rho(\mathbf{p}) = \int_{\mathcal{A}} \frac{d^d \mathbf{p}}{E(\mathbf{p})}. \quad (32)$$

It is interesting to explore the invariance properties of this measure. For this purpose, we consider an arbitrary proper-orthochronous Lorentz transformation. Such transformations are either spatial rotations, or boosts, or a combination of both [16]. They act as linear transformations on the energy-momentum vector (E, \mathbf{p}) . Due to the fixed relation $E(\mathbf{p}) = (m^2 + \mathbf{p}^2)^{1/2}$ between energy and momentum, a Lorentz transformation can also be viewed as transformation that operates on the momentum coordinates \mathbf{p} alone, denoted by $L : \mathcal{RP}_d \rightarrow \mathcal{RP}_d$. The functions L are linear only in the case of pure rotations, but nonlinear otherwise [31]. However, analogous to equation (22), we may define the Lorentz transformation $L[\mathcal{A}]$ of a set $\mathcal{A} \subset \mathcal{RP}_d$ by

$$L[\mathcal{A}] := \{L(\mathbf{p}) \mid \mathbf{p} \in \mathcal{A}\}. \quad (33)$$

By taking into account the well-known fact that [23, 31, 32]

$$\frac{d^d \mathbf{p}}{E(\mathbf{p})} = \frac{d^d \mathbf{p}'}{E(\mathbf{p}')} \quad (34)$$

holds under Lorentz transformations $\mathbf{p} \mapsto \mathbf{p}'$, one then finds that

$$\chi(L[\mathcal{A}]) = \chi(\mathcal{A}), \quad \forall \mathcal{A} \in \mathcal{RP}_d. \quad (35)$$

Hence, the specific reference measure underlying the modified Jüttner distributions with $\eta = 1$ is distinguished by the property that it is Lorentz invariant. In view of the fact that the Lorentz

group is the relativistic counterpart of the Galilei group, one can say that the modified Jüttner distribution is obtained when adopting the *geometric interpretation* in section 2.3. Put differently, if the modified Jüttner function were the correct relativistic one-particle equilibrium distribution, then the maximum principle for the relative entropy should be completed by the postulate that the reference measure in momentum space must be invariant under the action of the fundamental symmetry group of the physical model (e.g., Galilei, Lorentz, etc).

3.2.1. Explicit example: one-dimensional case $d = 1$.

As remarked earlier, the one-dimensional case $d = 1$ is somewhat special, because (only) in this case the Lorentz boosts form a group that may be directly identified with the one-dimensional velocity space $\mathcal{RV}_1 := (-1, 1)$. The composition of two Lorentz boosts induces a group multiplication \oplus_ψ on \mathcal{RV}_1 , given by

$$v_3 := v_1 \oplus_\psi v_2 := \frac{v_1 + v_2}{1 + v_1 v_2}. \quad (36)$$

This group operation is well known as the Einstein addition of velocities. The task of introducing an invariant measure on the group $(\mathcal{RV}_1, \oplus_\psi)$ falls exactly into the class of problems originally considered by Haar [12]. The subscript ψ symbolizes that the Einstein addition \oplus_ψ is equivalent to an ordinary addition ‘+’ in the space $\Psi := (-\infty, \infty)$ of the rapidity variables $\psi := \operatorname{arvtanh} v$. Put differently, the maps ‘ $\operatorname{arctanh}$ ’ and ‘ tanh ’ induce a group isomorphism between $(\mathcal{RV}_1, \oplus_\psi)$ and $(\Psi, +)$. The latter fact makes it particularly simple to identify the Haar measure on $(\mathcal{RV}_1, \oplus_\psi)$: one merely needs to rewrite the Lebesgue measure λ_ψ on Ψ , which is invariant under the addition of rapidities, in terms of the velocity coordinate; in differential notation, one then finds

$$d\lambda_\psi = d\psi = \left(\frac{d\psi}{dv} \right) dv = \gamma^2(v) dv \propto \frac{dp}{E(p)}, \quad (37)$$

corresponding to the Lorentz invariant measure on \mathcal{RV}_1 and \mathcal{RP}_1 , respectively (cf equation (34)). As discussed above, using this measure as the reference measure in the MEP, one obtains the one-dimensional modified Jüttner distribution with $\eta = 1$. For comparison, the ordinary addition $p_3 := p_1 + p_2$ in momentum space $\mathcal{RP}_1 = (-\infty, \infty)$ induces another group operation \oplus_p on $\mathcal{RV}_1 = (-1, 1)$ by means of the map $v(p) = p/(m^2 + p^2)^{1/2}$. The corresponding velocity addition law reads explicitly

$$v_3 := v_1 \oplus_p v_2 := \frac{v_1 \sqrt{1 - v_2^2} + v_2 \sqrt{1 - v_1^2}}{\sqrt{1 - v_1^2 v_2^2 + 2v_1 v_2 \sqrt{1 - v_1^2} \sqrt{1 - v_2^2}}}. \quad (38)$$

Analogous to equation (37), the invariant Haar measure on $(\mathcal{RV}_1, \oplus_p)$ is obtained by expressing the Lebesgue measure λ_p on \mathcal{RP}_1 , which is invariant under the momentum addition, in terms of the velocity variable, yielding

$$d\lambda_p = dp \propto \gamma^3(v) dv. \quad (39)$$

As discussed in section 3.1, by using this measure in the MEP one is led to the standard Jüttner function.

4. Summary

We have studied the canonical MEP for thermodynamic equilibrium distributions by combining basic ideas from group and measure theory [10, 15]. It has been demonstrated that the concept of

relative entropy [7]–[9] provides a suitable basis for stating the MEP in a coordinate invariant way. Moreover, this approach clarifies that thermodynamic equilibrium distributions, if obtained from a MEP [6], are determined not only by their constraint functionals but also by the underlying reference measures. The latter may be characterized in terms of their symmetry properties, i.e., by their invariance under group actions.

As examples, we analysed the two most frequently considered candidates [19]–[30] for the relativistic generalization of the Maxwell distribution. It could be shown that the two candidate distributions are based on different underlying reference measures. The reference measure leading to a standard Jüttner distribution [19]–[21] is uniquely characterized by the fact that it is invariant under momentum translations, whereas the modified Jüttner distribution [25]–[30] is related to a Lorentz invariant reference measure in momentum space. Even though the above approach clarifies the underlying mathematical differences on a fundamental level, it does not permit a decision to be made as to which distribution is actually the better candidate, as either reference measure has its own merits. In our opinion, this ambiguity deserves further consideration in the future.

We conclude this paper by mentioning two applications. The correct relativistic equilibrium distribution is required in order to calculate the friction coefficients and noise correlation functions of relativistic Langevin equations (RLEs) self-consistently [33]. An accurate determination of these quantities is essential, e.g., if RLEs are employed to estimate the outcome of high energy collision experiments, as recently done by van Hees *et al* [34]. Another potential, astrophysical application concerns the Sunyaev–Zeldovich (SZ) effect [35, 36], i.e., the distortion of the cosmic microwave background (CMB) spectrum due to the interaction of CMB photons with the hot electron gas in clusters of galaxies. The size of relativistic corrections to the SZ effect depends essentially on the shape of the assumed relativistic electron velocity distribution.

Acknowledgment

We thank J Casado-Pascual, J Chluba, D Cubero, S Hilbert, K Sakmann, M Schindler, S Weber, and M Wubs for helpful discussions.

References

- [1] Goldstein H 1980 *Classical Mechanics (Addison-Wesley Series in Physics)* 2nd edn (Reading, MA: Addison-Wesley)
- [2] Debbasch F 1993 Classical mechanics and gauge theories *Am. J. Phys.* **61** 1131–3
- [3] Weinberg S 2002 *The Quantum Theory of Fields* vol 1 (Cambridge: Cambridge University Press)
- [4] Coleman S 1995 *Aspects of Symmetry (Selected Erice Lectures)* (Cambridge: Cambridge University Press)
- [5] Quigg C 1997 *Gauge Theories of the Strong, Weak and Electromagnetic Interactions* (Boulder: Westview Press)
- [6] Jaynes E T 1957 Information theory and statistical mechanics *Phys. Rev.* **106** 620–30
- [7] Ochs W 1976 Basic properties of the generalized Boltzmann–Gibbs–Shannon entropy *Rep. Math. Phys.* **9** 135–55
- [8] Ochs W 1976 A unique characterization of the generalized Boltzmann–Gibbs–Shannon entropy *Rep. Math. Phys.* **9** 331–54
- [9] Wehrl A 1978 General properties of entropy *Rev. Mod. Phys.* **50** 221–60
- [10] Sergi A and Giaquinta P V 2007 On the geometry and entropy of non-Hamiltonian phase space *J. Stat. Mech.: Theor. Exp.* P02013

- [11] Banavar J and Maritan A 2007 The maximum relative entropy principle *Preprint cond-mat/0703622*
- [12] Haar A 1933 Der Massbegriff in der Theorie der kontinuierlichen Gruppen *Ann. Math.* **34** 147–69
- [13] Becker R 1967 *Theory of Heat* (New York: Springer)
- [14] Matolcsi T, Kristóf J and Székely M 1996 On the momentum distribution of molecules of an ideal gas *Publ. Appl. Analysis* **7** 1–14
- [15] Halmos P R 1951 *Measure Theory (The University Series in Higher Mathematics)* (New York: van Nostrand)
- [16] Sexl R U and Urbantke H K 2001 *Relativity, Groups, Particles (Springer Physics)* (Wien: Springer)
- [17] Einstein A 1905 Zur Elektrodynamik bewegter Körper *Ann. Phys., Lpz.* **17** 891–921
- [18] Einstein A 1905 Ist die Trägheit eines Körpers von seinem Energiegehalt abhängig? *Ann. Phys., Lpz.* **18** 639–41
- [19] Jüttner F 1911 Das Maxwellsche Gesetz der Geschwindigkeitsverteilung in der Relativtheorie *Ann. Phys., Lpz.* **34** 856–82
- [20] Pauli W 1921 Relativitätstheorie *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen* vol V/2 (Leipzig: Teubner) pp 539–775
- [21] Synge J L 1957 *The Relativistic Gas* (Amsterdam: North-Holland)
- [22] Ter Haar D and Wegeland H 1971 Thermodynamics and statistical mechanics in the special theory of relativity *Phys. Rep.* **1** 31–54
- [23] Cercignani C and Kremer G M 2002 *The Relativistic Boltzmann Equation: Theory and Applications (Progress in Mathematical Physics)* vol 22 (Basel: Birkhäuser Verlag)
- [24] Liboff R 1990 *Kinetic Theory* (Englewood Cliffs, NJ: Prentice Hall)
- [25] Horwitz L P, Shashoua S and Schieve W C 1989 A manifestly covariant relativistic Boltzmann equation for the evolution of a system of events *Physica A* **161** 300–38
- [26] Horwitz L P, Schieve W C and Piron C 1981 Gibbs ensembles in relativistic classical and quantum mechanics *Ann. Phys.* **137** 306–40
- [27] Schieve W C 2005 Covariant relativistic statistical mechanics of many particles *Found. Phys.* **35** 1359–81
- [28] Lehmann E 2006 Covariant equilibrium statistical mechanics *J. Math. Phys.* **47** 023303
- [29] Dunkel J and Hänggi P 2007 One-dimensional nonrelativistic and relativistic Brownian motions: a microscopic collision model *Physica A* **374** 559–72
- [30] Dunkel J, Talkner P and Hänggi P 2007 Relativistic diffusion processes and random walk models *Phys. Rev. D* **75** 043001
- [31] van Kampen N G 1969 Lorentz-invariance of the distribution in phase space *Physica* **43** 244–62
- [32] Peskin M E and Schroeder D V 1995 *An Introduction to Quantum Field Theory* (Reading, MA: Perseus)
- [33] Dunkel J and Hänggi P 2006 Relativistic Brownian motion: from a microscopic binary collision model to the Langevin equation *Phys. Rev. E* **74** 051106
Dunkel J and Hänggi P 2006 *Phys. Rev. E* **74** 069902 (erratum)
- [34] van Hees H, Greco V and Rapp R 2006 Heavy-quark probes of the quark-gluon plasma and interpretation of recent data taken at the BNL relativistic heavy ion collider *Phys. Rev. C* **73** 034913
- [35] Sunyaev R A and Zeldovich Y B 1972 The observations of relic radiation as a test of the nature of x-ray radiation from the clusters of galaxies *Commun. Astrophys. Space Phys.* **4** 173
- [36] Itoh N, Kohyama Y and Nozawa S 1998 Relativistic corrections to the Sunyaev-Zeldovich effect for clusters of galaxies *Astrophys. J.* **502** 7–15