A Relativistically Invariant Formulation of The Canonical Ensemble

Alexander Taskov
Imperial College London
alexander.taskov17@imperial.ac.uk

Abstract

An invariant Boltzmann distribution is derived by exploiting the symmetries between relativistic energy and momentum, ultimately leading to the realization of an inverse four-temperature β^μ. The form of the four temperature is constrained and the Boltzmann distribution is shown to agree with the Maxwell-Jütner distribution. It is further shown that taking the distribution off-shell leads to a natural PDE for the partition function similar in form to the Klein-Gordon equation. This ultimately leads to a generalization of the wick rotation correspondence between quantum and statistical mechanics, as well as an interesting interpretation of Lorentz group gauge theories.

1 Introduction

Relativistic thermodynamics is not a well understood subject. Much of its mainstream interest has died out since the beginning of the 20th century, having been overshadowed by the likes of quantum theory and general relativity. As such, there has hardly been any good consensus on matters as fundamental as the transformation law for temperatures. Considering the statistical mechanical origins of the discovery of quantum physics, as well as the relatively recent discovery of the thermodynamic nature of black holes, one might expect investigations into deeper thermodynamic theories to have taken more of a centre stage among physicists, although this has not been the case.

In a previous paper relativistic thermodynamics was explored and a covariant Boltzmann distribution derived, however the foundational concepts of the derivation were not considered with much depth. The purpose of this writing will then be to understand the formula better, both in its origin and its tangible consequences. Throughout the article we take $c = k = \hbar = 1$

2 Covariant Boltzmann Distribution

Consider an isolated, stationary ensemble of subsystems with conserved four momentum. Minimizing its entropy requires

$$\frac{\partial S}{\partial n_i} = \frac{\partial}{\partial n_i} \log \frac{N!}{\prod_{i=1}^{N} n_i!} = 0 \tag{1}$$

where N is the total number of subsystems and n_i is the occupancy of the ith system state. In the typical derivation of the canonical ensemble we now apply the energy conservation restriction as a Lagrange constraint, giving (after Stirling's approximation)

$$\log n_i + \beta\epsilon_i = 0 \tag{2}$$

where β is the inverse temperature (Lagrange multiplier) and ϵ_i is the energy of the ith subsystem state. It will be convenient to write this in the continuous limit, where the state occupancy becomes a density $n_i = n(\epsilon_i) \to \rho(\epsilon)$, and we vary the entropy

$$S = \int d\epsilon \rho \log \rho \tag{3}$$

with the constraint

$$\int d\epsilon \cdot \epsilon\rho = E \tag{4}$$

to get

$$\begin{aligned} \delta S = \delta \int d\epsilon \cdot \rho \log \rho + \beta\epsilon\rho = 0 \\ \int d\epsilon \cdot \delta\rho \log \rho + \delta\rho + \beta\epsilon\rho = 0 \\ \implies \log \rho + \beta\epsilon + 1 = 0 \end{aligned} \tag{5}$$

This, however, reasonlessly favours the 0th component of the four-momentum. If we take into account all other conserved components of the four momentum, we instead get

$$\log \rho + \beta^0 p^0 - \beta^1 p^1 - \beta^2 p^2 - \beta^3 p^3 + 1 = 0 \tag{6}$$

or, in Einstein notation

$$\log \rho + \beta^\mu p^\nu \eta_{\mu\nu} + 1 = 0 \tag{7}$$

where the components β^μ are just a series of Lagrange multipliers corresponding to the conserved components of the four-momentum, and p^μ are the components of the four-momentum of the relevant subsystem state. This equation then implies

$$\rho = \frac{1}{Z} e^{-\beta^\mu p^\nu \eta_{\mu\nu}} = \frac{1}{Z} e^{-\beta \cdot p} \tag{8}$$

were Z is the normalizing partition function. Thus far we have ignored the parity symmetry of the system. In the rest frame, this will amount to an invariance under $p^i \to -p^i$. In the $\bar{p}$ four momentum frame this becomes $p \to \Lambda_{2(\bar{p}-p)}p$, ie. $p = \bar{p} - \Delta p \to p = \bar{p} + \Delta p$. For this to be maintained in Eqn. (8) we require $\beta \propto \bar{p}$. That is

$$\rho = \frac{1}{Z} e^{-b\bar{p}\cdot p} \tag{9}$$

In the $\bar{p} = (\bar{m}, 0, 0, 0)$ rest frame, this would mean all momenta are equally likely, which obviously isn't the case. The equations break down because we haven't encoded the mass-momentum-energy relationship, which can be done with a simple state degeneracy function

$$\rho = \frac{1}{Z} \delta^{(4)}(p^2 - m^2) e^{-\beta\cdot p} \tag{10}$$

where now if we wish to know the probability of finding a subsystem with 3-momentum $\vec{p}$ we can integrate over the energy states, and use the rest frame where $\beta \propto (1, 0, 0, 0)$

$$\rho(\vec{p}) = \int dp^0 \rho(p) = \frac{1}{Z} e^{-\beta^0 \sqrt{\vec{p}^2 + m^2}} \tag{11}$$

which is just the Maxwell-Jütner distribution. It is clear now that relativistic inverse temperature is a four-vector, which is determined by the average system momentum up to a constant. Changing between frames is then just a matter of Lorentz transforming, which will leave the distribution invariant. The distribution being invariant is important, as it keeps the entropy invariant as well; entropy is determined by the microstates of the system, and so shouldn't change in any decent relativistic theory.

The general covariance of the distribution ensures a straightforward generalization to curved spacetime. Consider the same system falling along a geodesic with tangent $\propto \bar{p}^\mu$. The total momentum is

$$P^\mu \propto \bar{p}^\mu \propto \int d^4p \rho(p) p^\mu \tag{12}$$

and by definition obeys the geodesic equation $\frac{d\bar{p}}{d\tau} = (\bar{p} \cdot \nabla)\bar{p} = 0$. This conservation then gives the four-momentum constraint necessary to retrieve the given relativistic Boltzmann distribution, however requires that the subsystems are confined to the same neighbourhood such that we can meaningfully add their momenta. Tidal forces in general prevent this from always being true for any ensemble of subsystems moving independently along geodesics. For example a free gas of non-interacting particles falling towards a black-hole cannot obey these relationships for this reason, however a gas of non-interacting particles confined to a box can (to some consistent approximation). Again, local parity means $\beta \propto \bar{p}$, except now $\bar{p}$ has a more interesting interpretation as the geodesic tangent. We might then treat the free non-interacting gas case as an ensemble of systems with diverging inverse four-temperatures. Thus, the spacetimes varying pressure properties means the state of a gas spread throughout a spacetime is not necessarily one of uniform temperature when ignoring the gravitating properties of the gas (but trivially uniform otherwise). Interestingly, perhaps the same conclusions can be reached from the other direction: the equilibrium state is only non-uniform when ignoring the thermodynamic properties of the spacetime.

3 Symmetry and Entropy

Let us now ponder upon the nature of the Lagrange constraints which lead to these equations. Although perfectly physically motivated as constraints, their inclusion as modifications to the entropy is purely mathematical. Fundamentally these constraints are conservational, and so arise from symmetries. In the case of four-momentum, translation symmetry leads to a temperature. Consider the translation $\rho \to e^{\delta\tau v\cdot p}\rho$, where p is the momentum generating the translation in the direction v. With this the entropy becomes

$$\begin{aligned} S &= \int d^4p e^{\delta\tau v\cdot p}\rho \log e^{\delta\tau v\cdot p}\rho \\ &= \int d^4p e^{\delta\tau v\cdot p}\big(\rho\log\rho + \rho\delta\tau v\cdot p\big) \end{aligned} \tag{13}$$

and the variation

$$\begin{aligned} \delta S = \int d^4p e^{\delta\tau v\cdot p}\delta\rho\big(\log\rho + 1 + \delta\tau v\cdot p\big) &= 0 \\ \implies \rho \propto e^{-\delta\tau v\cdot p} \end{aligned} \tag{14}$$

A system falling along a geodesic will not be invariant if translated off that geodesic, and so $v \propto \bar{p}$, therefore giving us the previous expression. The constant b is then of the form $b \propto \frac{\bar{m}}{\delta\tau}$ for some equilibrium relaxation time $\delta\tau$.

This derivation of the constraint can be interpreted as follows: consider such a translationally invariant gas of particles in thermodynamic equilibrium. If the gas is dense enough then we can partition it off into a set of similar subsystems. Translation invariance implies each partition should have the same distribution of momenta, and so if a particle of momentum p crosses out of a partition, another particle of a similar momentum must cross into the partition. In a short time frame more particles of high momentum will cross between partitions, and so it is easier for a system containing many particles of lower momenta to satisfy this simultaneous crossing condition than one with higher momenta particles. On the other hand, satisfying this condition for lower momenta particles puts a higher constraint on their location in their partition, and so we expect a non-trivial distribution of momenta throughout the system.

Alternatively: translation symmetry means measuring the momentum of a particle in one partition will tell us there should exist a particle of a similar momentum in the other partitions, and so the uncertainty associated with measuring the distribution will in general decrease.

In either case, the symmetry means an observer measuring the system gains more information from each measurement than if the system didn't have that symmetry. We can expect from this that the Gibbs entropy formula is altered, and therefore so are the conditions on its maximization which fix the density function. It is easy to see from the above derivations that continuous symmetries will modify the entropy exactly as their conserved generators would normally constrain the Lagrangian, therefore always giving the appropriate Boltzmann distribution.

4 PDE of The Partition Function

Another point of interest is that Eqn. (10) is on-shell. As we saw, this is important for the statistical formulation, however the actual thermodynamics should only see coarse-grained, averaged out versions of the overall system. Let us then leave the distribution off-shell and encode mass-energy-momentum equivalence slightly differently. Consider the partition function in this case

$$Z = \int d^4p g(p) e^{-\beta\cdot p} \tag{15}$$

where $g(p)$ is some degeneracy. If we define $\partial_\mu = \frac{\partial}{\partial\beta^\mu}$ then

$$-\partial_\mu Z = \bar{p}_\mu Z \tag{16}$$

and so we can encode $E^2 - p^2 = m^2$ on the level of the overall system by demanding

$$\partial_\mu \partial^\mu Z = \bar{m}^2 Z \tag{17}$$

which is incredibly interesting in its similarity to the Klein-Gordon equation. In fact, the off-shell framework implies the microscopic statistical behaviour is unobservable, and replaced by this partition "wavefunction" probability density for the systems temperature. Furthermore, the general solution to this PDE can be written as a linear combination of $e^{\sigma\beta\cdot p}$ terms, where $\sigma^2 = 1$. This solution is interesting in that it forms a representation of the (non-compact) hyperbolic rotation (boost) group, and thus has no good orthogonality relationship (like eg. the fourier mode solutions of the KGE). It is this orthogonal mixing of phases that underlies Quantum decoherence, and so we expect there to be no collapse in this framework. In fact, thermodynamic systems brought into contact will typically reach a mutual equilibrium, in some sense becoming "more coherent". These solutions also imply that off-shell thermodynamics and Quantum mechanics can be related by $\sigma\beta\cdot p \to ip\cdot x$, or more generally $\sigma S \to i\mathcal{S}$, where S is the entropy and $\mathcal{S}$ the action. This transformation is in some sense a more general form of the Wick rotation, and fundamentally concerns a compactification of the boost group. If we then imagine forming a "hypercomplex" scalar field $\varphi = \varphi^\mu e_\mu = \varphi^0\sigma + \varphi^1 i + \varphi^2 j + \varphi^3 k$ where $i^2 = j^2 = k^2 = -1$, then gauging the symmetry $\varphi \to e^{e^\mu \theta_\mu}\varphi$, the gauge field produced would be some Lorentzian manifold connection $\Gamma^\rho_{\mu\nu} e_\rho e^\mu e^\nu$. The 0 component of such a field would then be classically statistical, while the rest would be quantum mechanical.

5 Conclusion

The purpose of this writing was to understand the nature of the Boltzmann distribution in a relativistic setting. In Section I. the relativistic energy-momentum equivalence was exploited in the derivation of the canonical ensemble, implying the existence of an inverse four-temperature, and returning an invariant Boltzmann distribution. It was then shown that parity symmetry demands the four-temperature be proportional to the systems expected momentum, such that a thermodynamic system localized around some point in a spacetime will vary in temperature as it falls along a geodesic. In Section II. the Lagrange constraints were reinterpreted as modifications to the Gibbs entropy arising from the symmetries generated by the relevant conserved quantities. Finally in Section III. it was shown that the mass-energy-momentum relationship can be encoded at the thermodynamic level, leading to a Klein-Gordon like PDE for the partition function. The hyperbolic phase solutions of this PDE then implied a generalization of the Wick rotation relating quantum and statistical mechanics from $t \to it$ to $\sigma S \to i\mathcal{S}$, as well as a statistical mechanical interpretation of the boost part of Lorentz group gauge theories.