Axioms for the Boltzmann Distribution*

Adam Brandenburger† and Kai Steverson‡

January 1, 2017
Very Preliminary

Abstract

A fundamental postulate of statistical mechanics is that all microstates in an isolated system are equally probable. This postulate, which goes back to Boltzmann, has often been criticized for not having a clear physical foundation. In this note, we provide a derivation of the canonical (Boltzmann) distribution that avoids this postulate. In its place, we impose two axioms with physical interpretations. The first axiom (Thermal Equilibrium) requires that, as our system of interest comes into contact with different heat baths, the ranking of states of the system by probability is unchanged. Physically, this axiom is a statement that in thermal equilibrium, population inversions do not arise. The second axiom (Energy Exchange) requires that, for any heat bath and any probability distribution on states, there is a universe consisting of a system and heat bath that can achieve this distribution. Physically, this axiom is a statement that energy flows between system and heat bath are unrestricted. We show that our two axioms identify the Boltzmann distribution.

1 Introduction

The postulates of statistical mechanics have been examined and debated ever since the beginnings of the field in the nineteenth century. A central postulate, put in place by Boltzmann, is that there is equal a priori probability that an isolated system will be found

*Financial support from the NYU Stern School of Business, NYU Shanghai, and J.P. Valles is gratefully acknowledged. We are greatly indebted to Paul Glimcher, since it is joint work with him that provided the inspiration for this note. We thank Shane Manesfield for very helpful conversations.

†Stern School of Business, Tandon School of Engineering, NYU Shanghai, New York University, New York, NY 10012, U.S.A., adam.brandenburger@stern.nyu.edu, http://www.adambrandenburger.com

‡Institute for the Interdisciplinary Study of Decision Making, New York University, New York, NY 10012, U.S.A., ksteverson@nyu.edu, https://sites.google.com/site/ksteverson/

in any one of its microstates which are compatible with the overall constraints placed on the system. In Planck's words, "all microscopic states are equally probable in dynamics" (Planck, 1932, p.225). The equal-probability assumption has been rationalized in several ways. One can simply appeal to the Laplacian stance of insufficient reason. The observer's knowledge of the system does not yield a distinction among the microstates, so no distinction can legitimately be introduced via their probabilities of occurrence (Mandl, 1988, p.41). Ergodic arguments are a more sophisticated way of arriving at the same conclusion.

Investigations into the status of the equal-probability postulate have continued. Jaynes (1957) replaced this assumption with a maximum-entropy principle (a principle of "maximum noncommitment with respect to missing information" (op.cit., p.602)) in order to derive the canonical (Boltzmann) distribution in the microcanonical ensemble. Goldstein et al. (2006) proved that, for quantum systems, the canonical distribution arises for almost all wave functions of the universe (system plus heat bath). Popescu, Short, and Winter (2006) showed that, even without energy constraints, a "general canonical principle" can be established for quantum systems, under which a system will almost always behave as if the universe is in the equal-probability state.

In this note, we take a different route. (Our analysis is for classical systems, but should be extendable to quantum systems.) We replace the equal-probability postulate with two physically interpretable axioms, which we show characterize the canonical (Boltzmann) distribution.

2 Axioms

In the usual (textbook) derivation, one fixes a heat bath $\mathbb{B}$ at a temperature T and a system $\mathbb{S}$ with possible states s_i, for $i = 1, 2, \ldots, n$. The system $\mathbb{S}$ specifies an energy level E_i for each state s_i. (See Figure 1.) The probability assigned to state s_i depends on the system $\mathbb{S}$ and the heat bath $\mathbb{B}$ and can therefore be written as $p_{\mathbb{S}}(s_i, \mathbb{B})$. One then appeals to the equal-probability postulate to write the ratios of probabilities of states as

$$\frac{p_{\mathbb{S}}(s_i; \mathbb{B})}{p_{\mathbb{S}}(s_j; \mathbb{B})} = \frac{\Omega_{\mathbb{B}}(E_{\text{total}} - E_i)}{\Omega_{\mathbb{B}}(E_{\text{total}} - E_j)}, \tag{1}$$

where E_{total} is the total energy of the composite $\mathbb{S} + \mathbb{B}$, so that $\Omega_{\mathbb{B}}(E_{\text{total}} - E_i)$ is then the number of microstates of $\mathbb{B}$. A Taylor expansion of the entropy $S_{\mathbb{B}}(E_{\text{total}} - E_i) = k \ln \Omega_{\mathbb{B}}$ of $\mathbb{B}$ (where k is the Boltzmann constant), and use of the formula $\partial S_{\mathbb{B}} / \partial E_{\text{total}} = 1/T$, yields the Boltmann distribution

$$p_{\mathbb{S}}(s_i;\mathbb{B}) = \frac{1}{Z}e^{-\frac{E_i}{kT}}, \tag{2}$$

where $Z = \Sigma_j e^{-E_j/kT}$ is the partition function (e.g., Mandl, 1988, pp.52-56).

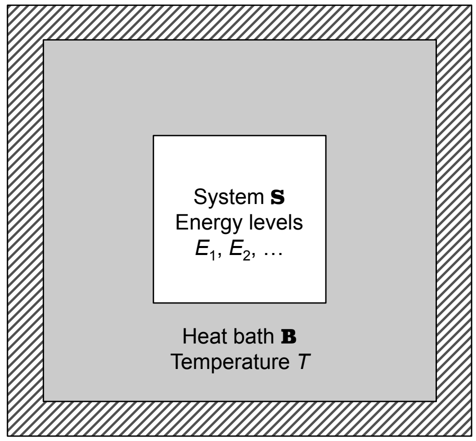

Figure 1

Our derivation will also begin with ratios of probabilities, as in Equation (1), but will not assume the equal-probability postulate. Our axioms are stated over a family $\{\mathbb{S}, \mathbb{S}', \mathbb{S}'', \ldots\}$ of systems and a family $\{\mathbb{B}, \mathbb{B}', \mathbb{B}'', \ldots\}$ of heat baths. All systems are defined on the same fixed underlying finite set of states $\{s_1, s_2, \ldots, s_n\}$.

Axiom 1. ***(Thermal Equilibrium)*** *Associated with each heat bath $\mathbb{B}$ there is a strictly increasing function $G_{\mathbb{B}} : (0, \infty) \to (0, \infty)$ such that for any system $\mathbb{S}$ and pair of states s_i and s_j, the ratio equation*

$$G_{\mathbb{B}}\left(\frac{p_{\mathbb{S}}(s_i;\mathbb{B})}{p_{\mathbb{S}}(s_j;\mathbb{B})}\right) = G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i;\mathbb{B}')}{p_{\mathbb{S}}(s_j;\mathbb{B}')}\right) \tag{3}$$

is satisfied.

To understand this axiom, it helps to see what it does not say. Suppose, instead, we require equality of the probability ratios without the transformations via the $G_{\mathbb{B}}$-functions. (Formally, all $G_{\mathbb{B}}$-functions are then the identity.) This axiom would say the heat baths play no role, which would be wrong at the physical level. The $G_{\mathbb{B}}$-functions allow for the idea that the probability of a state of the system will depend on the particular heat bath $\mathbb{B}$ to which the system is attached. But we require these functions to be strictly increasing. This condition ensures that the probabilistic ranking of states of the system does not change with changes in the heat bath. This is physically correct, since we are considering systems in equilibrium and, therefore, population inversions are not possible. If state s_i is more likely than another state s_j, this is because s_i is lower energy than s_j. In thermal equilibrium, the same probabilistic ranking of states will hold whether the heat bath is $\mathbb{B}$ or $\mathbb{B}'$. Here is the formal statement.

Lemma 1. *If* $p_{\mathbb{S}}(s_i; \mathbb{B}) \geq p_{\mathbb{S}}(s_j; \mathbb{B})$, *then* $p_{\mathbb{S}}(s_i; \mathbb{B}') \geq p_{\mathbb{S}}(s_j; \mathbb{B}')$.

Proof. We can write

$$\frac{p_{\mathbb{S}}(s_i; \mathbb{B})}{p_{\mathbb{S}}(s_j; \mathbb{B})} \geq \frac{p_{\mathbb{S}}(s_j; \mathbb{B})}{p_{\mathbb{S}}(s_i; \mathbb{B})},$$

so that, since $G_{\mathbb{B}}$ is increasing,

$$G_{\mathbb{B}}\left(\frac{p_{\mathbb{S}}(s_i; \mathbb{B})}{p_{\mathbb{S}}(s_j; \mathbb{B})}\right) \geq G_{\mathbb{B}}\left(\frac{p_{\mathbb{S}}(s_j; \mathbb{B})}{p_{\mathbb{S}}(s_i; \mathbb{B})}\right).$$

But, using Equation (3),

$$G_{\mathbb{B}}\left(\frac{p_{\mathbb{S}}(s_i; \mathbb{B})}{p_{\mathbb{S}}(s_j; \mathbb{B})}\right) = G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i; \mathbb{B}')}{p_{\mathbb{S}}(s_j; \mathbb{B}')}\right) \text{ and } G_{\mathbb{B}}\left(\frac{p_{\mathbb{S}}(s_j; \mathbb{B})}{p_{\mathbb{S}}(s_i; \mathbb{B})}\right) = G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_j; \mathbb{B}')}{p_{\mathbb{S}}(s_i; \mathbb{B}')}\right),$$

and, therefore,

$$G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i; \mathbb{B}')}{p_{\mathbb{S}}(s_j; \mathbb{B}')}\right) \geq G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_j; \mathbb{B}')}{p_{\mathbb{S}}(s_i; \mathbb{B}')}\right),$$

from which, since $G_{\mathbb{B}'}$ is increasing,

$$\frac{p_{\mathbb{S}}(s_i; \mathbb{B}')}{p_{\mathbb{S}}(s_j; \mathbb{B}')} \geq \frac{p_{\mathbb{S}}(s_j; \mathbb{B}')}{p_{\mathbb{S}}(s_i; \mathbb{B}')},$$

or $p_{\mathbb{S}}(s_i; \mathbb{B}') \geq p_{\mathbb{S}}(s_j; \mathbb{B}')$, as required. □

Our second axiom is designed to capture the fact that a heat bath $\mathbb{B}$ is very large compared with a system $\mathbb{S}$, so that any energy flows are possible between the two at the given temperature of the bath. We say this formally by fixing a heat bath $\mathbb{B}$ and some probability distribution on the states $\{s_1, s_2, \ldots, s_n\}$. We then say that we can attach a system $\mathbb{S}$ to $\mathbb{B}$ so that the desired probabilities are obtained. Physically, we know we can do this. Indeed, Equation (2) for the Boltzmann distribution tells us there are energy levels E_i, for $i = 1, 2, \ldots, n$, that yield the probabilities in question. (If λ_i is the probability of state i, then we set $E_i = -kT \ln \lambda_i$.) So, we attach a system $\mathbb{S}$ with these energy levels to the heat bath $\mathbb{B}$. Since $\mathbb{B}$ is very large compared with $\mathbb{S}$, we can always do this at the prevailing temperature T. Here is the formal statement. (We assume that λ has full support, i.e, that $\lambda_i > 0$ for all i. This guarantees that all ratios of probabilities are well-defined.)

Axiom 2. ***(Energy Exchange)*** *For any heat bath $\mathbb{B}$ and any full-support probability distribution $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ on $\{s_1, s_2, \ldots, s_n\}$, there is a system $\mathbb{S}$ such that $p_{\mathbb{S}}(\cdot; \mathbb{B}) = \lambda$.*

3 Result

We can now state our result, which is an axiomatic derivation of the Boltzmann distribution.

Theorem 1. *Suppose that Axioms 1 and 2 are satisfied. Then there exist functions $T : \{\mathbb{B}, \mathbb{B}', \mathbb{B}'', \ldots\} \to (0, \infty)$ and $E : \{s_1, s_2, \ldots, s_n\} \times \{\mathbb{S}, \mathbb{S}', \mathbb{S}'', \ldots\} \to (0, \infty)$ such that for each heat bath $\mathbb{B}$ and system $\mathbb{S}$, and for each $i = 1, 2, \ldots, n$,*

$$p_{\mathbb{S}}(s_i; \mathbb{B}) = \frac{1}{Z(\mathbb{B}, \mathbb{S})} e^{-\frac{E(s_i, \mathbb{S})}{T(\mathbb{B})}}, \tag{4}$$

where $Z(\mathbb{B}, \mathbb{S}) = \Sigma_j e^{-E(s_j, \mathbb{S})/T(\mathbb{B})}$.

Equation (4) is the Boltzmann distribution, with temperature $T(\cdot)$ (as a function of the heat bath) and energy levels $E(s_1, \cdot), E(s_2, \cdot), \ldots, E(s_n, \cdot)$ (as a function of the system). (We get $k = 1$ since temperature and energy are not measured in physical units here.) Notice that only positive temperatures are possible under our treatment. This makes sense, since we have assumed thermal equilibrium, and negative temperatures arise can only in systems which are (temporarily) out of equilibrium (e.g., Braun et al., 2013). Also, as expected in an abstract treatment, the fundamental quantity that emerges is $E(\cdot, \cdot)/T(\cdot)$, namely, entropy. We can be more precise about this last point by establishing the uniqueness properties of the functions T and E that represent a given heat bath and system.

Theorem 2. *Assume that, for each heat bath $\mathbb{B}$, it is not the case that all states have equal probability. Suppose a system $\mathbb{S}$ satisfies Equation (4) with functions E and T. Then $\mathbb{S}$ satisfies Equation (4) with functions $\widetilde{E}$ and $\widetilde{T}$ if and only if there are real numbers $\alpha > 0$ and β such that*

$$\begin{aligned} E(s_i, \mathbb{S}) &= \alpha \widetilde{E}(s_i, \mathbb{S}) + \beta \text{ for all states } s_i, \\ T(\mathbb{B}) &= \alpha \widetilde{T}(\mathbb{B}) \text{ for all heat baths } \mathbb{B}. \end{aligned}$$

(Physically speaking, the equal-probability case ruled out is that of infinite temperature.) Notice that the scaling factor for T is the same as the multiplicative factor in the affine transformation of E. It follows that, while the ratios $E(\cdot, \cdot)/T(\cdot)$ are not unique, the differences between these ratios, i.e., the entropy differences

$$\frac{E(s_i, \mathbb{S}) - E(s_j, \mathbb{S})}{T(\mathbb{B})}$$

between states, are unique. Again, we expect this on physical grounds.

References

1. Aczél, J., *Lectures on Functional Equations and Their Applications*, Academic Press, 1966.

2. Braun, S., J. Ronzheimer, M. Schrebier, S. Hodgman, T. Tom, I. Bloch, and U. Schneider, "Negative Absolute Temperature for Motional Degrees of Freedom," *Science*, 339, 2013, 52-55.

3. Goldstein, S., J. Lebowitz, R. Tumulka, and N. Zanghi, "Canonical Typicality," *Physical Review Letters*, 96, 2006, 050403.

4. Jaynes, E., "Information Theory and Statistical Mechanics," *Physical Review*, 106, 1957, 620-630.

5. Mandl, F., *Statistical Physics*, Wiley, 2nd ed., 1988.

6. Planck, M., *Theory of Heat*, trans. by L. Brose, Macmillan, 1932.

7. Popescu, S., A. Short, and A. Winter, "Entanglement and the Foundations of Statistical Mechanics," *Nature Physics*, 2, 2006, 754-758.

Proofs

We choose heat bath $\mathbb{B}$ as a reference point. Since $G_{\mathbb{B}}$ is strictly increasing, it is invertible. Therefore, for any (other) heat bath $\mathbb{B}'$, we can define a function $H_{\mathbb{B}'} : (0, \infty) \to (0, \infty)$ by

$$H_{\mathbb{B}'}(t) = G_{\mathbb{B}}^{-1}(G_{\mathbb{B}'}(t)).$$

By Axiom 1, we have that for any system $\mathbb{S}$ and pair of states r, s,

$$G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(r; \mathbb{B}')}{p_{\mathbb{S}}(s; \mathbb{B}')}\right) = G_{\mathbb{B}}\left(\frac{p_{\mathbb{S}}(r; \mathbb{B})}{p_{\mathbb{S}}(s; \mathbb{B})}\right),$$

so that

$$H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(r;\mathbb{B}')}{p_{\mathbb{S}}(s;\mathbb{B}')}\right)=\frac{p_{\mathbb{S}}(r;\mathbb{B})}{p_{\mathbb{S}}(s;\mathbb{B})}. \tag{5}$$

It follows that for any triplet of states s_i, s_j, s_k,

$$H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i;\mathbb{B}')}{p_{\mathbb{S}}(s_j;\mathbb{B}')}\right)\times H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_j;\mathbb{B}')}{p_{\mathbb{S}}(s_k;\mathbb{B}')}\right)=\frac{p_{\mathbb{S}}(s_i;\mathbb{B})}{p_{\mathbb{S}}(s_j;\mathbb{B})}\times\frac{p_{\mathbb{S}}(s_j;\mathbb{B})}{p_{\mathbb{S}}(s_k;\mathbb{B})}=\frac{p_{\mathbb{S}}(s_i;\mathbb{B})}{p_{\mathbb{S}}(s_k;\mathbb{B})}.$$

We can also write

$$H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i;\mathbb{B}')}{p_{\mathbb{S}}(s_k;\mathbb{B}')}\right)=\frac{p_{\mathbb{S}}(s_i;\mathbb{B})}{p_{\mathbb{S}}(s_k;\mathbb{B})}.$$

Putting these two equations together yields

$$H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i;\mathbb{B}')}{p_{\mathbb{S}}(s_j;\mathbb{B}')}\right)\times H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_j;\mathbb{B}')}{p_{\mathbb{S}}(s_k;\mathbb{B}')}\right)=H_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(s_i;\mathbb{B}')}{p_{\mathbb{S}}(s_k;\mathbb{B}')}\right). \tag{6}$$

We want to turn Equation (6) into the Cauchy functional equation. To do so, we need an intermediate result. (This result assumes that there are at least three states. The case of two states is treated later.)

Lemma 2. *For any $t, u \in (0,\infty)$, we can choose states s_i, s_j, s_k and a full-support probability distribution λ on $\{s_1, s_2, \ldots, s_n\}$ so that*

$$\frac{\lambda(s_i)}{\lambda(s_j)}=t \text{ and } \frac{\lambda(s_j)}{\lambda(s_k)}=u. \tag{7}$$

Proof. Choose three distinct states s_i, s_j, s_k, and set

$$\begin{aligned}\lambda(s_i) &= \frac{tuv}{1+u+tu},\\ \lambda(s_j) &= \frac{uv}{1+u+tu},\\ \lambda(s_k) &= \frac{v}{1+u+tu},\end{aligned}$$

where $v=1$ if $n=3$ (there are three states in total) and $v=\frac{1}{2}$ if $n>3$. Also, if $n>3$, set

$$\lambda(s)=\frac{1}{2(n-3)},$$

for all $s \neq s_i, s_j, s_k$. It is easy to check that λ has full support and that Equation (7) is satisfied. Also, if $n=3$, then

$$\sum_s \lambda(s)=\frac{tuv+uv+v}{1+u+tu}=v=1,$$

and if $n > 3$,

$$\sum_s \lambda(s) = \sum_{s \neq s_i, s_j, s_k} \frac{1}{2(n-3)} + \frac{tuv + uv + v}{1 + u + tu} = \frac{1}{2} + v = 1,$$

so that λ is a well-defined probability distribution on the states. $\square$

By Axioms 1 and 2, there is a system $\mathbb{S}$ so that Equation (3) is satisfied and $p_{\mathbb{S}}(\cdot; \mathbb{B}') = \lambda$. But then $p_{\mathbb{S}}(\cdot; \mathbb{B}')$ also satisfies Equation (6), and, therefore, using Equation (7), we obtain

$$H_{\mathbb{B}'}(t) \times H_{\mathbb{B}'}(u) = H_{\mathbb{B}'}(tu), \tag{8}$$

for any $t, u \in (0, \infty)$. Moreover, the functions $G_{\mathbb{B}}$ and $G_{\mathbb{B}'}$ are increasing and therefore have at most a countable number of discontinuities, from which it follows that $H_{\mathbb{B}'}$ can have at most a countable number of discontinuities. This allows us to apply a version of the Cauchy functional theorem (Axćel, 1966, p.41, Theorem 3) to Equation (8), to conclude that there exists a function $T : \{\mathbb{B}, \mathbb{B}', \mathbb{B}'', \ldots\} \to (0, \infty)$ such that

$$H_{\mathbb{B}'}(t) = t^{T(\mathbb{B}')}. \tag{9}$$

Note that $T(\mathbb{B}) = 1$ (this is why we called $\mathbb{B}$ a reference point). Also, from Equation (9) we get

$$G_{\mathbb{B}'}(t) = G_{\mathbb{B}}(t^{T(\mathbb{B}')}).$$

Since $G_{\mathbb{B}'}$ and $G_{\mathbb{B}}$ are both strictly increasing, it follows that $T(\mathbb{B}') > 0$ for all heat baths $\mathbb{B}'$. Next, using Equation (9) in Equation (5), we find

$$\frac{p_{\mathbb{S}}(r; \mathbb{B}')}{p_{\mathbb{S}}(s; \mathbb{B}')} = \left(\frac{p_{\mathbb{S}}(r; \mathbb{B})}{p_{\mathbb{S}}(s; \mathbb{B})} \right)^{\frac{1}{T(\mathbb{B}')}}.$$

Summing over all states r yields

$$\frac{1}{p_{\mathbb{S}}(s; \mathbb{B}')} = \sum_r \left(\frac{p_{\mathbb{S}}(r; \mathbb{B})}{p_{\mathbb{S}}(s; \mathbb{B})} \right)^{\frac{1}{T(\mathbb{B}')}},$$

which we can invert to get

$$p_{\mathbb{S}}(s; \mathbb{B}') = \frac{p_{\mathbb{S}}(s; \mathbb{B})^{\frac{1}{T(\mathbb{B}')}}}{\sum_r p_{\mathbb{S}}(r; \mathbb{B})^{\frac{1}{T(\mathbb{B}')}}}. \tag{10}$$

Finally, to make Equation (10) into the Boltzmann distribution, define $E : \{s_1, s_2, \ldots, s_n\} \times \{\mathbb{S}, \mathbb{S}', \mathbb{S}'', \ldots\} \to (0, \infty)$ by $E(r, \mathbb{S}) = -\ln p_{\mathbb{S}}(r; \mathbb{B})$ for each state r.

This completes the proof of Theorem 1, except for the case of two states. (The case of one state is trivial.) Here, we define the function T directly, by requiring it to give the solution to each equation

$$\left(\frac{p_{\mathbb{S}}(s_1;\mathbb{B}')}{p_{\mathbb{S}}(s_2;\mathbb{B}')}\right)^{T(\mathbb{B}')} = \frac{p_{\mathbb{S}}(s_1;\mathbb{B})}{p_{\mathbb{S}}(s_2;\mathbb{B})},$$

as we vary the heat bath $\mathbb{B}'$. We can argue similarly to Lemma 1 to see that $p_{\mathbb{S}}(s_1;\mathbb{B}') \geq p_{\mathbb{S}}(s_2;\mathbb{B}')$ if and only if $p_{\mathbb{S}}(s_1;\mathbb{B}) \geq p_{\mathbb{S}}(s_2;\mathbb{B})$. It follows that we will get $T(\mathbb{B}') > 0$ for all $\mathbb{B}'$, as required.

We should also establish that our axioms identify the Boltzmann distribution and not some subfamily of this distribution. To show this, start by supposing that Equation (4) holds. Define $G_{\mathbb{B}'} : (0,\infty) \to (0,\infty)$ by $G_{\mathbb{B}'}(t) = t^{T(\mathbb{B}')}$. Then for any system $\mathbb{S}$ and pair of states r, s, we can write

$$G_{\mathbb{B}'}\left(\frac{p_{\mathbb{S}}(r;\mathbb{B}')}{p_{\mathbb{S}}(s;\mathbb{B}')}\right) = G_{\mathbb{B}'}\left(e^{\frac{E(s,\mathbb{S})-E(r,\mathbb{S})}{T(\mathbb{B}')}}\right) = e^{E(s,\mathbb{S})-E(r,\mathbb{S})}.$$

Since the right-hand side is independent of $\mathbb{B}'$, we see that Equation (3) is satisfied, which establishes Axiom 1. For Axiom 2, fix a heat bath $\mathbb{B}'$ and a full-support probability distribution λ on the states. Let $T(\mathbb{B}')$ be arbitrary and set $E(s_i,\mathbb{S}) = -kT(\mathbb{B}')\ln\lambda_i$ for each i. Then $p_{\mathbb{S}}(\cdot;\mathbb{B}') = \lambda$, as required.

We now turn to the proof of Theorem 2. Suppose a system $\mathbb{S}$ satisfies Equation (4) for two pairs of functions E, T and $\widetilde{E}, \widetilde{T}$. Equation (4) implies that for any states s_1, s_2, s,

$$\frac{E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})}{E(s_2,\mathbb{S}) - E(s,\mathbb{S})} = \left(\ln\frac{p_{\mathbb{S}}(s_1;\mathbb{B})}{p_{\mathbb{S}}(s_2;\mathbb{B})}\right)\left(\ln\frac{p_{\mathbb{S}}(s;\mathbb{B})}{p_{\mathbb{S}}(s_2;\mathbb{B})}\right)^{-1} = \frac{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})}{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s,\mathbb{S})}.$$

Rearranging gives

$$(E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})) \times (\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s,\mathbb{S})) = (E(s_2,\mathbb{S}) - E(s,\mathbb{S})) \times (\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})),$$

from which,

$$E(s,\mathbb{S}) = E(s_2,\mathbb{S}) - \frac{E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})}{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})} \times (\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s,\mathbb{S})),$$

and, therefore,

$$E(s,\mathbb{S}) = \frac{E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})}{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})} \times \widetilde{E}(s,\mathbb{S}) + E(s_2,\mathbb{S}) - \frac{E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})}{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})} \times \widetilde{E}(s_2,\mathbb{S}).$$

Now set

$$\alpha = \frac{E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})}{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})} \text{ and } \beta = E(s_2,\mathbb{S}) - \frac{E(s_2,\mathbb{S}) - E(s_1,\mathbb{S})}{\widetilde{E}(s_2,\mathbb{S}) - \widetilde{E}(s_1,\mathbb{S})} \times \widetilde{E}(s_2,\mathbb{S}).$$

By assumption, there are states s_1, s_2 such that $p_{\mathbb{S}}(s_1;\mathbb{B}) > p_{\mathbb{S}}(s_2;\mathbb{B})$. (There is no loss of generality in labeling these two states this way.) It follows that $\alpha > 0$.

Next observe that, for any heat bath $\mathbb{B}'$,

$$\ln\left(\frac{p_{\mathbb{S}}(s_2;\mathbb{B}')}{p_{\mathbb{S}}(s_1;\mathbb{B}')}\right) = \frac{E(s_1,\mathbb{S}) - E(s_2,\mathbb{S})}{T(\mathbb{B}')} = \frac{\widetilde{E}(s_1,\mathbb{S}) - \widetilde{E}(s_2,\mathbb{S})}{\widetilde{T}(\mathbb{B}')},$$

from which, using the relationship between E and $\widetilde{E}$, we get

$$\frac{\alpha\widetilde{E}(s_1,\mathbb{S}) - \alpha\widetilde{E}(s_2,\mathbb{S})}{T(\mathbb{B}')} = \frac{\widetilde{E}(s_1,\mathbb{S}) - \widetilde{E}(s_2,\mathbb{S})}{\widetilde{T}(\mathbb{B}')}.$$

By assumption, $p_{\mathbb{S}}(s_1;\mathbb{B}) \neq p_{\mathbb{S}}(s_2;\mathbb{B})$. (Again, there is no loss of generality in using the state labels s_1 and s_2.) It follows that $\widetilde{E}(s_1,\mathbb{S}) \neq \widetilde{E}(s_2,\mathbb{S})$ and, therefore, $T(\mathbb{B}') = \alpha\widetilde{T}(\mathbb{B}')$, as claimed. This completes the proof of the forward direction of Theorem 2.

For the reverse direction, suppose that a system $\mathbb{S}$ satisfies Equation (4) for the functions E and T, and let $\alpha > 0$ and β be real numbers. Equation (4) then yields, for any heat bath $\mathbb{B}'$,

$$p_{\mathbb{S}}(s;\mathbb{B}') = \frac{e^{-\frac{E(s,\mathbb{S})}{T(\mathbb{B}')}}}{\sum_j e^{-\frac{E(s,\mathbb{S})}{T(\mathbb{B}')}}} = \frac{e^{-\frac{\alpha E(s,\mathbb{S})}{\alpha T(\mathbb{B}')}}}{\sum_j e^{-\frac{\alpha E(s,\mathbb{S})}{\alpha T(\mathbb{B}')}}} = \frac{e^{-\frac{\beta}{\alpha T(\mathbb{B}')}} e^{-\frac{\alpha E(s,\mathbb{S})}{\alpha T(\mathbb{B}')}}}{e^{-\frac{\beta}{\alpha T(\mathbb{B}')}} \sum_j e^{-\frac{\alpha E(s,\mathbb{S})}{\alpha T(\mathbb{B}')}}} = \frac{e^{-\frac{\alpha E(s,\mathbb{S})+\beta}{\alpha T(\mathbb{B}')}}}{\sum_j e^{-\frac{\alpha E(s,\mathbb{S})+\beta}{\alpha T(\mathbb{B}')}}},$$

from which we see that the system $\mathbb{S}$ satisfies Equation (4) for the functions $\alpha E + \beta$ and αT, as we needed to show.

To prove that entropy differences are unique, as asserted after the statement of Theorem 2, first suppose that a system $\mathbb{S}$ satisfies Equation (4) for the functions E, T and $\widetilde{E}, \widetilde{T}$. Theorem 2 tells us that there are real numbers $\alpha > 0$ and β such that $E(\cdot,\mathbb{S}) = \alpha\widetilde{E}(\cdot,\mathbb{S}) + \beta$

and $T(\cdot) = \alpha\widetilde{T}(\cdot)$. It follows that for any pair of states s_i, s_j, and any heat bath $\mathbb{B}'$,

$$\frac{E(s_i,\mathbb{S})}{T(\mathbb{B}')} - \frac{E(s_j,\mathbb{S})}{T(\mathbb{B}')} = \frac{\alpha\widetilde{E}(s_i,\mathbb{S}) + \beta}{\alpha\widetilde{T}(\mathbb{B}')} - \frac{\alpha\widetilde{E}(s_j,\mathbb{S}) + \beta}{\alpha\widetilde{T}(\mathbb{B}')} = \frac{\widetilde{E}(s_i,\mathbb{S})}{\widetilde{T}(\mathbb{B}')} - \frac{\widetilde{E}(s_j,\mathbb{S})}{\widetilde{T}(\mathbb{B}')},$$

as claimed. Conversely, suppose a system $\mathbb{S}$ satisfies Equation (4) for the functions E, T, and there exist functions $\widetilde{E}, \widetilde{T}$ such that

$$\frac{\widetilde{E}(s_i,\mathbb{S})}{\widetilde{T}(\mathbb{B}')} - \frac{\widetilde{E}(s_j,\mathbb{S})}{\widetilde{T}(\mathbb{B}')} = \frac{E(s_i,\mathbb{S})}{T(\mathbb{B}')} - \frac{E(s_j,\mathbb{S})}{T(\mathbb{B}')}.$$

This says that for each heat bath $\mathbb{B}'$, there is a number $\gamma_{\mathbb{B}'}$ such that

$$\frac{\widetilde{E}(s_i,\mathbb{S})}{\widetilde{T}(\mathbb{B}')} = \frac{E(s_i,\mathbb{S})}{T(\mathbb{B}')} + \gamma_{\mathbb{B}'}.$$

It follows that

$$\frac{e^{-\frac{\widetilde{E}(s_i,\mathbb{S})}{\widetilde{T}(\mathbb{B}')}}}{\sum_j e^{-\frac{\widetilde{E}(s_j,\mathbb{S})}{\widetilde{T}(\mathbb{B}')}}} = \frac{e^{-\frac{E(s_i,\mathbb{S})}{T(\mathbb{B}')} - \gamma_{\mathbb{B}'}}}{\sum_j e^{-\frac{E(s_j,\mathbb{S})}{T(\mathbb{B}')} - \gamma_{\mathbb{B}'}}} = \frac{e^{-\gamma_{\mathbb{B}'}} e^{-\frac{E(s_i,\mathbb{S})}{T(\mathbb{B}')}}}{e^{-\gamma_{\mathbb{B}'}} \sum_j e^{-\frac{E(s_j,\mathbb{S})}{T(\mathbb{B}')}}} = \frac{e^{-\frac{E(s_i,\mathbb{S})}{T(\mathbb{B}')}}}{\sum_j e^{-\frac{E(s_j,\mathbb{S})}{T(\mathbb{B}')}}} = p_{\mathbb{S}}(s_i;\mathbb{B}'),$$

from which we see that the system $\mathbb{S}$ satisfies Equation (4) for the functions $\widetilde{E}, \widetilde{T}$.