

Canonical Probabilities by Directly Quantizing Thermodynamics

Blake C. Stacey¹

¹*Physics Department, University of Massachusetts Boston*

(Dated: December 21, 2020)

Boltzmann weighting, the canonical map from energies to probabilities, is shown to follow from the requirement that in quantum theory, state spaces compose by the tensor product.

We teach thermodynamics as a science of bulk matter, each body characterizable using a few parameters held constant or smoothly varying. Energy is useful work or waste heat, without much gradation of utility in between. Manipulations are the application of fire or of the sea [1, 2]. By contrast, quantum physics is a subject where bringing out its riches generally requires precision. The most intriguing and compelling kinds of quantum strangeness tend to wash out when we become sloppy in handling the world [3–5]. Bringing these subjects together is thus a bit of an adventure. My goal here is to provide a derivation of the canonical, or Boltzmann, class of distributions by starting with thermodynamics and just barely getting its toes wet in quantum theory. This originated with wondering whether a standard bit of pedagogy could be run backward, and it may hold some appeal for those who enjoy re-approaching the familiar from a different starting point. Our purpose will not be to study how a quantum state might “thermalize” in dynamical detail, but rather to understand why a particular class of states should be regarded as the “thermal” ones.

The Zeroth Law of thermodynamics states that thermal equilibrium is transitive, and so we can divide up the set of all possible preparations of thermodynamical systems into equivalence classes. In quantum theory, we encode the preparation of a system by ascribing a density matrix — a positive semidefinite operator of unit trace — to it. Thus, the meeting-of-the-minds for these two subjects, the quantization of thermodynamics, ought to center upon finding a method for generating density matrices given an equivalence class. The First Law tells us that energy is conserved. Quantum theory actively resists the idea that the values a physicist “measures” exist prior to the act of measurement [6]. How, then, can we match the call for constancy with the fact that quantum experiments are active interventions? We cannot with integrity live as though a quantum system has an energy value that an energy measurement simply *reveals*, but if we assume our laboratory abilities are limited, we can minimize the sinfulness of pretending otherwise. Mathematically, this means picking a scheme for building density matrices such that they all commute with one another, and so they can all be diagonalized in a common basis. If all the density matrices are diagonal in this “energy eigenbasis”, then we can come as close as quantum theory will allow to treating them as though they express ignorance about a value of the energy that pre-exists measurement [7]. Moreover, unitary time evolutions generated by a Hamiltonian diagonal in this basis will preserve all these density matrices, giving us the best match we can feasibly get with our classical notion of conservation.

We will at first assume that the density-matrix elements depend *only* upon the energy. Once we understand this case, we will generalize.

To summarize: What we want is a scheme for writing diagonal density matrices where each entry on the diagonal is a function of the eigenvalue of an observable. Requiring that our scheme plays nicely with the tensor product rule constrains what that function can

be. So, let's devise a scheme for generating density matrices with energy being the key observable, that is, matrices that are diagonal in the energy eigenbasis. We want to express the idea of equivalence classes, so we will say that when two preparations are equivalent, we will use the same function for mapping energy values to matrix elements.

The first part of the argument proceeds just as well in quantum theory and in classical probability, since by the above considerations the density matrices we want to construct are all simultaneously diagonalizable. We start by considering systems with *discrete* sets of possible energy levels. Our derivation will be based on the idea that if A and B are at the same temperature, a noninteracting composite system AB is also at that temperature. Suppose that E_j is an energy level of system A and E_k is an energy level of B . Then, if there is no interaction between the two systems, AB will have an energy level $E_j + E_k$. If we assume that for all systems prepared at temperature T ,

$$P(E_n) = \frac{1}{Z} f(E_n), \quad (1)$$

then we have

$$\frac{f(E_j)f(E_k)}{Z_A Z_B} = \frac{f(E_j + E_k)}{Z_{AB}}. \quad (2)$$

But we have the freedom to adjust f by an overall multiplicative constant, since the meaningful quantities are the probabilities and any prefactor will cancel when we divide by Z . So, we can declare

$$f(0) = 1, \quad (3)$$

which yields

$$Z_A Z_B = Z_{AB}, \quad (4)$$

and thus

$$f(E_j + E_k) = f(E_j)f(E_k). \quad (5)$$

which is the Cauchy functional equation, up to a logarithmic transform [8, §2.1.2]. Its solutions are $f(E) = 0$, which we can rule out by normalizability, and $f(E) = e^{g(E)}$, where g is an arbitrary function satisfying $g(E_1 + E_2) = g(E_1) + g(E_2)$. On physical grounds, we can assume that f is continuous, per the general smoothness we expect in thermodynamics problems. In fact, we only need to require continuity at a single point in order to fix the form of the function f completely:

$$f(E) = e^{-\beta E}, \quad (6)$$

where the “coolness” β labels the equivalence classes of thermal equilibrium. (The term “coolness” for inverse temperature is due to Baez and Pollard [9].) Showing that β ought to be inversely proportional to the temperature T that we define in phenomenological thermodynamics takes a little more work. For example, you can do a bit of elementary kinetic theory on an ideal gas [10] and convince yourself that the expectation value of a particle's kinetic energy is proportional to $k_B T$.

This argument is conceptually similar to one given by James Clerk Maxwell. To paraphrase his thoughts into our language, Maxwell argued that the components of a gas atom's velocity — call them v_x , v_y and v_z — are independent random variables.

$$p(v) = p(v_x)p(v_y)p(v_z). \quad (7)$$

But if no direction is preferred, then $p(v)$ can only be a function of the *magnitude* of v , or in other words, a function of v^2 . Likewise, we have no reason to prefer “up” over “down” motion along any axis, so the probability densities for each coordinate must be symmetrical about 0. We again obtain a functional equation where a sum becomes a product. The probability density is then an exponential function of v^2 , and to be well-behaved, it must be a decaying exponential. The rate at which the exponential decays is set by the temperature, since the temperature determines $\langle mv^2/2 \rangle$.

Calculating probabilities by taking e to the power of minus some energy divided by $k_B T$ is almost an instinctive move in statistical physics. It is the kind of result that is stronger than the argument used to derive it, and that is remembered when the rest of the course is forgotten.

To understand more fully the relation between β and T , we import more knowledge about classical thermodynamics. First, we perform the standard calculation for the expectation value of the energy:

$$\langle E \rangle := \sum_i E_i p(E_i) = \sum_i \frac{E_i e^{-\beta E_i}}{Z} = -\frac{1}{Z} \sum_i \frac{\partial}{\partial \beta} e^{-\beta E_i} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}, \quad (8)$$

which we can write compactly as

$$\langle E \rangle = -\frac{\partial \log Z}{\partial \beta}. \quad (9)$$

Moreover, since we have a probability distribution $f(E)/Z$, we can take the Shannon entropy of it. Importantly, the Shannon entropy can be motivated by information-theoretic arguments that make no reference to thermodynamics [11, 12]. From that angle, its relevance to heat flow is not so obvious! Working out the algebra,

$$\langle -\log p \rangle := -\sum_i p(E_i) \log p(E_i) = \beta \langle E \rangle + \log Z. \quad (10)$$

From thermodynamics, we have the phenomenological relation

$$E = F + TS, \quad (11)$$

where F is the Helmholtz free energy (meaning that we have finally gotten around to invoking the Second Law). These two expressions coincide if we make the identifications

$$E = \langle E \rangle, \quad S = k_B \langle -\log p \rangle, \quad \beta = \frac{1}{k_B T}, \quad F = -\frac{1}{\beta} \log Z. \quad (12)$$

The first identification is the obvious way to match the coarse-grained picture of thermodynamics with the finer-grained view of a statistical theory. To further motivate the remaining three, note that we want the free energy to be *extensive*: Given two independent, identically prepared systems, the free energy of the pair should be double that of either system alone. The partition functions multiply, so we turn to the logarithm to make a quantity that adds. Moreover, in thermodynamics, the entropy is the negative derivative of the Helmholtz free energy with respect to temperature. So, using the quotient and chain rules,

$$E = F - T \frac{\partial F}{\partial T} = -T^2 \frac{\partial (F/T)}{\partial T} = \frac{\partial (F/T)}{\partial (1/T)}. \quad (13)$$

Comparing this with Eq. (9), we see the parallel between (F/T) and $-\log Z$, and between $1/T$ and β . And if we make these identifications, we can see what the thermodynamic entropy must correspond to:

$$S = \frac{E - F}{T} = k_B \beta \langle E \rangle + k_B \log Z = -k_B \langle (-\beta E - \log Z) \rangle = -k_B \langle \log p \rangle. \quad (14)$$

Mathematically, it appears overall that we can say β is proportional to $1/T$, or we can match the Shannon entropy with the thermodynamic entropy, and either correspondence implies the other. It is not clear at this stage which identification ought to be taken as the more fundamental. However, if we go back and hit the books, we find that one introductory move is to develop a bit of “baby kinetic theory” for a classical ideal gas, and from this development show that if we have two gases in thermal equilibrium, the mean kinetic energies of their atomic center-of-mass motions are equal. The mean CM kinetic energy is therefore a suitable label for the equivalence classes under the Zeroth Law, and the simple choice is to *declare* that temperature is proportional to mean CM kinetic energy. Making the same declaration here would be no less respectable! From this perspective, identifying the characteristic energy scale β^{-1} with the thermodynamic temperature (up to a choice of units) is the more elementary move, and the correspondence between information-theoretic and thermodynamic entropies is the derived one.

To see how these choices of correspondence work in more detail, we consider an *infinitesimal change* of the information-theoretic entropy, which we imagine to be brought about by a small outside influence gently perturbing the energy levels [4, §9.1]. We can express this as

$$d(\langle -\log p \rangle) = - \sum_i d(p(E_i) \log p(E_i)), \quad (15)$$

which evaluates to

$$d(\langle -\log p \rangle) = \sum_i [d(E_i p(E_i)) - d(E_i) p(E_i)] \beta. \quad (16)$$

We can identify the first sum as our expected value for the change in total energy, and the second sum as the expected value of the *work* done on the system (provided the perturbation is sufficiently slow). The differential form of the first law says that

$$dE = dW + dQ, \quad (17)$$

where we have crossed the differentials on the right-hand side to indicate that they individually are path-dependent, while their sum is not. Having already asserted that the thermodynamic E maps onto the quantum-mechanical $\langle E \rangle$, we see that

$$d(\langle -\log p \rangle) = \beta d\langle Q \rangle. \quad (18)$$

The coolness β is the integrating factor necessary to turn the path-dependent infinitesimal heat flow dQ into an exact differential. Eq. (18) indicates what we said above, that assuming one correspondence implies the other.

Now, we move fully into quantum mechanics. Our path to the Boltzmann weighting formula $p(E_n) \propto e^{-\beta E_n}$ goes through just as well for diagonal density matrices being composed under the tensor product. Consequently, we have the functional form of the Gibbs states,

$$\rho = \frac{e^{-\beta H}}{\text{tr} e^{-\beta H}}, \quad (19)$$

and we can make use of them. For example, we can have a system to which we have ascribed a Gibbs state ρ_A interact with a second system to which we have assigned an arbitrary state ρ_B . The simplest such evolution is a joint unitary

$$\rho_A \otimes \rho_B \rightarrow \rho' := U(\rho_A \otimes \rho_B)U^\dagger. \quad (20)$$

This preserves the von Neumann entropy of the joint state:

$$S^{\text{vN}}(\rho') = S^{\text{vN}}(\rho_A) + S^{\text{vN}}(\rho_B). \quad (21)$$

The von Neumann entropy is subadditive, meaning that for the *marginals* of the time-evolved state, we have

$$S^{\text{vN}}(\rho') \leq S^{\text{vN}}(\rho'_A) + S^{\text{vN}}(\rho'_B). \quad (22)$$

Therefore, we can write

$$\Delta S_A^{\text{vN}} + \Delta S_B^{\text{vN}} \geq 0, \quad (23)$$

whereas the conservation of energy implies

$$\Delta \langle E_A \rangle + \Delta \langle E_B \rangle = 0. \quad (24)$$

The interaction will not in general preserve the form of the Gibbs state, but we can still compute the quantity $S^{\text{vN}} - \beta \langle E \rangle$ before and after. Prior to the interaction, this is just $\log Z$. Its change works out to be

$$\Delta(S_A^{\text{vN}} - \beta \langle E_A \rangle) = -\text{tr} \rho'_A (\log \rho'_A - \log \rho_A). \quad (25)$$

We recognize the expression for the *relative* von Neumann entropy:

$$S^{\text{vN}}(\sigma' | \sigma) := \text{tr} \sigma' (\log \sigma' - \log \sigma). \quad (26)$$

This is provably nonnegative, and so

$$\Delta(S_A^{\text{vN}} - \beta \langle E_A \rangle) = -S^{\text{vN}}(\rho'_A | \rho_A) \leq 0. \quad (27)$$

Combining this with the subadditivity and energy-conservation relations, we deduce

$$\Delta(S_B^{\text{vN}} - \beta \langle E_B \rangle) \geq 0. \quad (28)$$

Therefore, if a system for which our initial state ρ_B is arbitrary interacts with a system in thermal equilibrium, this quantity must generally increase. Repeatedly interacting with systems at the same temperature, our state for system B tends toward a limit where

$$S_B^{\text{vN}} - \beta \langle E_B \rangle = - \sum_i p(E_i) (\log p(E_i) + \beta E_i) \quad (29)$$

is maximized, assuming no selection rules that artificially constrain the accessibility of energy levels. Maximizing this quantity with the condition that $p(E_i)$ is properly normalized recovers the Gibbs state at coolness β [4, §9.2].

Now, we return to the question of dependencies on quantities other than the energy. What if there exist other “observables” that commute with each other and with the Hamiltonian, so that we can simultaneously diagonalize the whole lot of them? The above argument shows

that interacting with systems that have been ascribed Gibbs states will generally tend to wash out dependencies upon the other “quantum numbers”.

We conclude that the canonical mapping from energies to probabilities is just a prior that plays nicely with the tensor product.

Very often, what people do is to go the other way round, starting with the formula for the von Neumann entropy and then looking for states that maximize it given some constraint. This works! As mentioned above, that yields the Gibbs states. But it’s reasonable to ask why one would maximize the von Neumann entropy instead of any other function. The more time I put into convincing myself that, yes, probability theory does make sense in quantum physics after all [13, 14], the less convinced I was that the von Neumann entropy was the “obvious” choice to maximize. Indeed, extremizing other entropic quantities turns out to yield interesting geometrical results [15], so it is worth understanding a physical reason — clean behavior under composition — to focus on the Shannon and von Neumann formulae.

Brandenburger and Steverson also used the Cauchy functional equation to obtain the form of Gibbs states, from a different starting point [16]. They do not invoke quantum concepts, and they do not investigate in depth the question of how the β that arises from solving the Cauchy functional equation relates to the T of thermodynamics. Younger Halpern and Renes arrive at a special case of the Cauchy functional equation in Lemma 8 of [17], using more details about “resource theory” than we have developed here, and proceed from there to the Gibbs states. (The fact that Gibbs states are all diagonal in the preferred basis makes them poor resources, from a quantum perspective; nonzero off-diagonal elements, or “coherences”, are thermodynamically useful. This is consistent with our ethos of obtaining the Gibbs states by merely dipping our toes into the possibilities of quantum theory and resolutely holding out for quasiclassicality. In turn, there is the possibility of pushing still further into the quantum by seeking states that are *robustly* coherent under *changes* of basis [18].) Alterman and coauthors have also studied the origin of Boltzmann weighting and notions of quantum-classical correspondence, in a setting that uses much more detail about specific potential wells than we have here [19].

The fact that canonical partition functions for independent systems are multiplicative is a helpful and easy standard result. Here, we have seen that we can run that bit of pedagogy backwards, starting from a desideratum for how composition behaves and arriving at the Boltzmann weighting $e^{-\beta E}$. It is interesting to step back and consider just what portions of the quantum formalism we used in obtaining this result. To get the Gibbs states themselves, we did not invoke in much detail at all how quantum systems can be correlated, only using the tensor-product rule for compositing state spaces. Further knowledge about the space of possible correlated joint states did not enter until we started talking about thermalization. It is known that we can use a combination of Einstein and Gleason to argue that “locally quantum” systems should combine according to the tensor product rule [20, 21]. Whatever the space of joint states is, it at least contains the tensor-product states. Showing that the joint states are specifically positive semidefinite operators on that tensor-product Hilbert space requires further assumptions [22, 23]. However, we did not have to rely upon the positivity of correlated states. We can just say that each system is “locally quantum” and that when we consider them together but uncorrelated (on opposite sides of the lab, perhaps), we use the tensor product. This is enough to power the primary result we obtained above.

The line of inquiry followed here was motivated by QBism, a research program in quantum information that, among other things, takes a personalist Bayesian interpretation of probability [13, 14]. This version of Bayesianism de-emphasizes the setting of priors by

maximizing entropy, which then prompts the question of whether one can obtain the Gibbs states from an alternative way to express indifference to the fine details of a preparation.

-
- [1] M. Kardar, *Statistical Physics of Particles* (Cambridge University Press, 2007).
 - [2] J. Rau, *Statistical Physics and Thermodynamics: An Introduction to Key Concepts* (Oxford University Press, 2017).
 - [3] N. D. Mermin, “[Quantum mechanics vs local realism near the classical limit: A Bell inequality for spin \$s\$](#) ,” *Physical Review D* **22** (1980), 356–61.
 - [4] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, 1993).
 - [5] A. Garg, “Agnostic Detector Error, Wigner Functions, and the Classical Limit of the High-Spin Einstein-Podolsky-Rosen Experiment,” [arXiv:1811.12247](#) (2018).
 - [6] N. D. Mermin, “[Hidden variables and the two theorems of John Bell](#),” *Reviews of Modern Physics* **65** (1993), 803–15, [arXiv:1802.10119](#).
 - [7] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa and B. Schumacher, “[Noncommuting mixed states cannot be broadcast](#),” *Physical Review Letters* **76** (1996), 2818, [arXiv:quant-ph/9511010](#).
 - [8] J. Aczél, *Lectures on Functional Equations and Their Applications* (Academic Press, 1966).
 - [9] J. C. Baez and B. S. Pollard, “[Quantropy](#),” *Entropy* **17** (2015), 772–89, [arXiv:1311.0813](#).
 - [10] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, 1964), https://www.feynmanlectures.caltech.edu/I_39.html.
 - [11] J. C. Baez and T. Fritz, “[A Bayesian characterization of relative entropy](#),” *Theory and Application of Categories* **29** (2014), 421–56, [arXiv:1402.3067](#).
 - [12] T. Leinster, *Entropy and Diversity: The Axiomatic Approach* (Cambridge University Press, forthcoming 2021). [arXiv:2012.02113](#).
 - [13] C. A. Fuchs and R. Schack, “[Quantum-Bayesian coherence](#),” *Reviews of Modern Physics* **85** (2013), 1693–1715, [arXiv:1301.3274](#).
 - [14] C. A. Fuchs and B. C. Stacey, “QBism: Quantum Theory as a Hero’s Handbook.” In *Proceedings of the International School of Physics “Enrico Fermi,” Course 197 – Foundations of Quantum Physics*, edited by E. M. Rasel and W. P. Schleich (2019). [arXiv:1612.07308](#).
 - [15] B. C. Stacey, “Sporadic SICs and exceptional Lie algebras,” [arXiv:1911.05809](#) (2019).
 - [16] A. Brandenburger and K. Steverson, “[Axioms for the Boltzmann distribution](#),” *Foundations of Physics* **49** (2019), 444–56.
 - [17] N. Yunger Halpern and J. M. Renes, “[Beyond heat baths: Generalized resource theories for small-scale thermodynamics](#),” *Physical Review E* **93** (2016), 022126, [arXiv:1409.3998](#).
 - [18] B. C. Stacey, “Invariant off-diagonality: SICs as equicoherent quantum states,” [arXiv:1906.05637](#) (2019).
 - [19] S. Alterman, J. Choi, R. Durst, S. M. Fleming and W. K. Wootters, “[The Boltzmann distribution and the quantum-classical correspondence](#),” *Journal of Physics A* **51** (2018), 345301, [arXiv:1710.06051](#).
 - [20] C. A. Fuchs, “Quantum mechanics as quantum information (and only a little more),” [arXiv:quant-ph/0205039](#) (2002).
 - [21] H. Barnum, C. A. Fuchs, J. M. Renes and A. Wilce, “Influence-free states on compound quantum systems,” [arXiv:quant-ph/0507108](#) (2005).
 - [22] H. Barnum, S. Beigi, S. Boixo, M. B. Elliott and S. Wehner, “[Local quantum measurement](#)

- and no-signaling imply quantum correlations,” *Physical Review Letters* **104** (2010), 140401, [arXiv:0910.3952](#).
- [23] G. de la Torre, L. Masanes, A. J. Short and M. P. Müller, “[Deriving quantum theory from its local structure and reversibility](#),” *Physical Review Letters* **109** (2012), 90403, [arXiv:1110.5482](#).