PHYSICAL REVIEW E **85**, 031151 (2012)

Shannon-entropy-based nonequilibrium "entropic" temperature of a general distribution

K. R. Narayanan[1] and A. R. Srinivasa[2,*]

[1]*Department of Electrical Engineering, Texas A&M University, College Station, Texas 77843, USA*

[2]*Department of Mechanical Engineering, Texas A&M University, College Station, Texas 77843, USA*

(Received 25 October 2010; revised manuscript received 31 May 2011; published 30 March 2012)

The concept of temperature is one of the key ideas in describing the thermodynamical properties of systems. In classical statistical mechanics of ideal gases, the notion of temperature can be described in at least two different ways: the kinetic temperature (related to the average kinetic energy of the particles) and the thermodynamic temperature (related to the ratio between infinitesimal changes in entropy and energy). For the Boltzmann distribution, the two notions lead to the same result. However, for nonequilibrium phenomena, while the kinetic temperature has been commonly used both for theoretical and simulation purposes, there appears to be no corresponding general definition of thermodynamic or entropic temperature. In this paper, we consider the statistical or Shannon entropy of a system and use the "de Bruijn identity" from information theory (see Appendix A 2 for a derivation of this identity) to show that it is possible to define a "Shannon temperature" or "entropic temperature" T for a *nonequilibrium* system as the ratio between the average curvature of the Hamiltonian function associated with the system and the trace of the *Fisher information matrix* of the nonequilibrium probability distribution (see Appendix A 1 for a definition of the Fisher information). *We show that this definition subsumes many other attempts at defining entropic temperatures for nonequilibrium systems and is not restricted to equilibrium or near equilibrium systems*. Intuitively, the gist of our approach is to use the Shannon or Gibbs entropy of a system and make use of the relation $dS = dQ_{\text{rev}}/T$ as a definition of *temperature*. We achieve this by positing a statistical notion of infinitesimal heating as the addition of uncorrelated random variables (in a special way). As an example of the utility of such a definition, we obtain the nonequilibrium entropic temperature for a system satisfying the Langevin equations. For such a system, we show that while the kinetic temperature is related to the changes in the energy of the system, the entropic or Shannon temperature is related to the changes in the entropy of the system. We show that this notion, together with the well known Cramer-Rao inequality in statistics demonstrates the validity of the second law of thermodynamics for such a nonequilibrium system.

DOI: 10.1103/PhysRevE.85.031151 PACS number(s): 05.70.Ln, 89.70.Cf

I. INTRODUCTION

Consider a system of N particles whose collective positions and momenta are given by the $3N$-dimensional vectors $\mathbf{q}$ and $\mathbf{p}$, respectively. Let the state of the system be described by a $6N$-dimensional vector-valued random variable $\mathbf{X}$ with associated probability density function (PDF) $f_{\mathbf{X}}(\mathbf{q},\mathbf{p})$. Let the system possess a Hamiltonian (or energy function) $H(\mathbf{q},\mathbf{p})$.

One of the interesting challenges in statistical mechanics is to be able to describe the interactions of "nonequilibrium" systems with their surroundings and to simulate their evolution toward equilibrium. There has been a tremendous interest in the literature (see, e.g., [1–8]), seeking to gain insight into the behavior of such systems from a statistical point of view and to study to what extent classical notions of heating, working, entropy (and the second law), free energy, temperature, etc., can be gainfully used to obtain further insight into such systems.

Any investigation into such systems immediately leads to a debate as to what we mean by the "entropy" of such a system. In this paper, we adopt the information theoretic notion of statistical entropy (Shannon or Gibbs entropy) as a measure of the uncertainty associated with the PDF of the system.

Jou and co-workers (see, e.g., [1–3,9]), in a series of papers have investigated which notions of temperature can be sensibly extended to a variety of nonequilibrium systems and shown that there are different notions of temperature that are useful for different purposes.

In this paper, we investigate the idea of *thermodynamic* or "entropic" or "Shannon" temperature of a nonequilibrium system using some tools from information theory and seek to answer the following questions:

(1) Which notions of temperature can be extended to classical nonequilibrium systems[1] in a manner that is consistent with classical thermodynamics and which simultaneously provides insight into the behavior of these systems?

(2) Is there a way to extend the thermodynamic relation (3), relating the derivative of the entropy with respect to the energy and the inverse temperature, to nonequilibrium systems?

(3) Is it possible to distinguish between working and heating for a nonequilibrium system (specifically one that satisfies Langevin dynamics) in a way that is intuitively appealing and physically meaningful in terms of requirements of the second law, etc.?

A simple canonical example of such endeavors has been the study of nonequilibrium systems satisfying Langevin dynamics ([9]). This is a system whose dynamics has been

---

[*] asrinivasa@tamu.edu

[1] There are systems that have the possibility of negative temperatures, for example, systems that have an upper bound to the energy. The approach presented in this paper is not applicable to such systems. Also we consider only those systems whose states can be modeled as being continuously varying.

©2012 American Physical Society

well characterized (see, e.g., [10]) and forms an ideal platform to see whether any of the classical notions can be extended to other nonequilibrium systems.

In Sec. II, we show that if we interpret the heating of a system statistically as adding an uncorrelated random variable to the state of the system, then we can show that both the entropy and the energy of the system increase *and that the de Bruijn identity of information theory (see [11-13]) shows that the ratio of the increase in entropy and that in the energy due to the addition of a statistically uncorrelated random variable is precisely the inverse temperature as defined here.*

In Sec. III, after a quick overview of standard results in classical equilibrium statistical mechanics, we discuss the different notions of temperature and then introduce a modification of the notion of equilibrium temperature due to Landau [9,14] and extend this notion to a general nonequilibrium system. In Sec. IV, we show that when one considers a Langevin system and considers the rate of change of the Shannon entropy of the system, this definition of a nonequilibrium temperature arises naturally and allows us to differentiate between heating and working in a completely intuitive way. We then show that the Cramer-Rao inequality of statistics ([15,16]), relating the Fisher information and the variance, can be reinterpreted in this case to mean that the internal entropy production is always non-negative.

II. PRELIMINARIES

Even though we are going to consider *general nonequilibrium systems whose PDFs are not necessarily related to the Boltzmann distribution* just in the way of background, we note that if the system were in equilibrium, its PDF would be the well-known Boltzmann distribution which is of the form [17]

$$f_{\mathrm{eq}}(\mathbf{q},\mathbf{p}) = \frac{e^{-\beta H(\mathbf{q},\mathbf{p})}}{Z}, \qquad Z := \int e^{-\beta H(\mathbf{q},\mathbf{p})}\, d\mathbf{q}\, d\mathbf{p}, \tag{1}$$

where $\beta = 1/kT$, k being the Boltzmann constant and T being the thermodynamic temperature of the system. In the above equation, Z is the partition function and the notation $\int(\cdot)d\mathbf{q}\,d\mathbf{p}$ stands for the integration over the $6N$-dimensional phase space. *For an equilibrium distribution*, it is common knowledge that the macroscopic properties can be obtained simply from the knowledge of the partition function (which happens to be the cumulant generating function); specifically, the entropy of the system and the average energy of the system are given, respectively, by

$$\begin{gathered}
\mathcal{E} := \langle H\rangle_{f_{\mathrm{eq}}} = -\frac{d\ln Z}{d\beta},\\
S := -\langle \ln f\rangle_{f_{\mathrm{eq}}} = \ln Z - \frac{\partial \ln Z}{\partial\beta},\\
S: = -k\langle \ln f\rangle_{f_{\mathrm{eq}}} = k\left(\ln Z - \frac{\partial \ln Z}{\partial\beta}\right)
\end{gathered} \tag{2}$$

where the notation $\langle\cdot\rangle_f$ denotes the expectation value with respect to the distribution f.

The temperature of the system (and hence the parameter β) has a variety of statistical interpretations; the three most common ones are listed below.

Thermodynamic temperature. It is also well known that for equilibrium distributions, the parameter β can be written in terms of the energy and entropy through the Gibbs relation [17]

$$k\beta = \frac{1}{\theta} = \frac{\partial S}{\partial\mathcal{E}}, \tag{3}$$

resulting in the name thermodynamic temperature for θ. This notion is the one that is most closely tied to thermodynamics and aspects of the second law, but thus far was confined to strictly equilibrium systems.

Kinetic temperature. For classical systems, the Hamiltonian is the sum of the kinetic energy and the potential energy v, i.e.,

$$H(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{N} \frac{\mathbf{p}^i\cdot\mathbf{p}^i}{2m} + v(\mathbf{q}). \tag{4}$$

The thermodynamic temperature can also be thought of as a measure of the variance of the momentum distribution and is proportional to the average kinetic energy of the particles, i.e., it is well known that

$$(3/2)NkT_{\mathrm{kin}} = \langle \mathbf{p}\cdot\mathbf{p}/2m\rangle_{f_{\mathrm{eq}}}. \tag{5}$$

This is by far the most popular notion of temperature and has been widely adopted for nonequilibrium systems also.

Curvature temperature: The curvature of the Hamiltonian and its relation to the temperature. For the equilibrium distribution of a thermodynamic system with a general Hamiltonian, we can show that temperature can also be interpreted in terms of the average curvature of the Hamiltonian (see Appendix A 4 for a simple derivation of this result), in the form

$$T_{\mathrm{curv}} = \frac{1}{k\beta} = \frac{\langle\nabla H\nabla H\rangle_f}{k\langle\nabla^2 H\rangle_f}, \tag{6}$$

where the notation ∇ stands for the gradient with respect to the phase space variables, ∇^2 stands for the Laplacian operator, and $\langle\cdot\rangle$ stands for the expectation. The above expression is based on the result that *for the Boltzmann distribution*, the divergence theorem implies that $\beta\langle\nabla H\nabla H\rangle = \langle\nabla^2 H\rangle$. As pointed out in [7], the above expression (with the momentum variables ignored) is derived in a very brief way by Landau and Lifshitz [14] (Eq. 33.14 on page 97). Rugh [18] and Baranyai [19] have approached this from a dynamical systems point of view and obtained a similar result based on ideas of ergodicity and dynamical systems theory (see also Sec. 3.2 of the review article by Casas-Vasquez and Jou [1]). This representation of temperature has been referred to as the "curvature temperature" and has been applied to nonequilibrium systems also (see, for example, [7,9]).

The results (5), (3), and (6) are equivalent for equilibrium systems defined by the Boltzmann distribution, and hence one may speak of "the temperature" for the system. For nonequilibrium systems, while (5) and (6) can be defined and used, they are not equal. Moreover, it is not evident how to generalize (3) to nonequilibrium systems. In the next section we show that *if we adopt the Shannon or information entropy of a distribution as the physical entropy of any system irrespective of whether it is in equilibrium or not, then we can use the de Bruijn identity of information theory (see Appendix) to define a thermodynamical temperature for any system.* We will further show that the difference between the kinetic temperature and the thermodynamic temperature defined here, can be used to gain insight into the evolution of the Langevin system toward equilibrium and the operation of the second law in this case.

The method used here is motivated by the observation that the result (3) can be given the following interpretation: For an equilibrium system the temperature is the ratio of the infinitesimal change in entropy due to an infinitesimal change in the internal energy *keeping all other macroscopic variables fixed*.

Viewed in this manner it is clear that changing the internal energy while keeping the macroscopic variables fixed is equivalent to infinitesimally heating the system. The de Bruijn identity (see [13]) can be adapted to provide a purely statistical representation of the above process (without the need for notions of equilibrium) and with the Shannon entropy in the numerator) and hence allows us to extend this notion to nonequilibrium systems. As shown in the Appendix, for the equilibrium Boltzmann distribution, the above definition directly leads to the curvature temperature defined in (6).

A. A statistical definition of nonequilibrium entropic or thermodynamic temperature

We are going to define the *entropic or thermodynamic temperature* for *any* system (irrespective of whether it is in equilibrium or not) by developing a statistical notion of "infinitesimally heating" a system as a particular form of perturbation of the microstate of the system in such a way that increases the Shannon entropy by an amount δS and the energy associated with the distribution by an amount δQ. We then show that if we calculate the ratio of the change in energy to that in the Shannon entropy, and by using the de Bruijn identity of information theory (see [13]), we obtain

$$\theta = \frac{\delta Q}{\delta S} = \frac{\langle \nabla^2 H\rangle_f}{k\,\mathrm{tr}\,\mathcal{I}}, \tag{7}$$

where the Fisher information matrix $\mathcal{I}$ associated with any distribution f is given by

$$\mathcal{I} = \langle \nabla(\ln f)\otimes\nabla(\ln f)\rangle_f \Rightarrow \mathrm{tr}\,\mathcal{I} = \langle \nabla(\ln f)\nabla(\ln f)\rangle_f \tag{8}$$

and is a measure of the "average curvature" of the probability density function and the notation ∇ stands for the gradient in phase space.

Our main contribution in this paper is to show that the definition (7) is natural (in a sense to be made precise), and deserves the name "thermodynamical temperature" for any distribution. Notice the close relationship of (7) with the Landau definition (6). At first sight, there appears to be a contradiction since, in the definition (6), the term $\langle\nabla^2 H\rangle_f$ appears in the *denominator* of the right-hand side while in (7), it appears in the *numerator*. This seeming contradiction is quickly resolved when we observe that *for the Boltzmann distribution the trace of the Fisher information is given by*

$$\mathrm{tr}\,\mathcal{I} = \langle \nabla(\ln f)\nabla(\ln f)\rangle_f = \beta^2\langle \nabla(H)\nabla(H)\rangle_f. \tag{9}$$

Upon observing that $\beta = 1/kT$ and substituting the above expression into the right-hand side of (7), we recover the form (6).

Whereas (6) was derived for an equilibrium distribution, the definition (7) is valid for a general nonequilibrium distribution and *makes no use of notions of equilibrium* as we shall see in Sec. III. In Appendix A 4 we show that for an equilibrium system whose PDF is given by the Boltzmann distribution (7) reduces to (6).

At this stage, it is fair to ask the questions: What is the use of any new definition of temperature? What insight does it provide regarding thermodynamical systems? To answer these questions, we show that this new definition allows us to separate the energy transfer in a nonequilibrium Langevin system into "heat" and "work" in an intuitively acceptable way and provides insight into the statistical aspects of the Langevin system. Furthermore, the proposed definition can also be used in computer simulations where we can explicitly compute the thermodynamical temperature (rather than simply using the kinetic temperature). The definition can serve as the basis for the exploration of small systems (composed of just a few thousand atoms or molecules) over extremely short times (femtoseconds) where the system is not likely to be in equilibrium.

III. THE DE BRUIJN IDENTITY AND THE TEMPERATURE FOR NONEQUILIBRIUM PROBABILITY DISTRIBUTIONS

In this section, we show that a definition of temperature based on infinitesimal changes in energy and entropy can be extended meaningfully to all kinds of nonequilibrium systems—provided one describes the notion of "heating" the system appropriately.

Recall from Sec. I that the state of the system under consideration is given by the positions $\mathbf{q}$ and the momenta $\mathbf{p}$ of the N particles. For the sake of convenience, we will treat $\mathbf{q}$,$\mathbf{p}$ as the components of a single $6N$-dimensional vector $\mathbf{x}$. Thus, we can describe the system statistically with a ($6N$-dimensional) vector-valued random variable $\mathbf{X}$ taking on values in the $6N$-dimensional phase space and whose PDF is $f_{\mathbf{X}}(\mathbf{x})$.

The total energy and Shannon entropy of such a system are both functionals of $f_{\mathbf{X}}(\mathbf{x})$ and are given, respectively, by

$$\begin{aligned}\mathcal{E} &:= \langle H\rangle_f = \int H(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x},\\ S &:= -\langle \ln f\rangle_f = -\int f_{\mathbf{X}}(\mathbf{x})\log\left[f_{\mathbf{X}}(\mathbf{x})\right]d\mathbf{x},\end{aligned} \tag{10}$$

where the notation $\langle\cdot\rangle_f$ denotes the expectation value with respect to the distribution f. Notice that these definitions are not restricted to equilibrium systems.

One of the oldest definitions of entropy of a system is the idea that it is $dS = dQ_{\mathrm{rev}}/\theta$ where dQ_{rev} is the energy supplied by infinitesimally heating the system. Now from a statistical point of view, adding a little bit of energy to a system will perturb the PDF in some way. This perturbation will then cause changes in entropy and energy and the ratio between these infinitesimal changes can be used as a measure of temperature.

In other words, since both S and $\mathcal{E}$ are functionals of $f_{\mathbf{X}}$, we can treat the distribution $f_{\mathbf{X}}$ itself as a parameter (i.e., we use the distribution $f_{\mathbf{X}}$ as a descriptor of the macrostate of the system) and intuitively define thermodynamic temperature as

$$\frac{1}{\theta} = \frac{dS(f_{\mathbf{X}})/df_{\mathbf{X}}}{d\mathcal{E}(f_{\mathbf{X}})/df_{\mathbf{X}}}. \tag{11}$$

In other words, given any probability distribution $f_{\mathbf{X}}$, we perturb it[2] to $f_{\mathbf{X}'}$ and calculate the corresponding perturbations in S and $\mathcal{E}$, namely, ΔS and $\Delta\mathcal{E}$. Then, we can define thermodynamic temperature as

$$\frac{1}{\theta} = \lim_{\Delta\mathcal{E}\to 0} \frac{\Delta S}{\Delta\mathcal{E}}. \tag{12}$$

At a minimum, the temperature obtained by perturbing $f_{\mathbf{X}}$ should satisfy the following criteria:

(a) θ must be non-negative.[3]

(b) θ should be equal to the thermodynamic temperature for the Boltzmann distribution.

(c) θ should represent "spread" of the distribution, i.e., the more spread out the distribution, the higher the temperature. This should be contrasted with Shannon entropy, which is a measure of the "uncertainty" associated with any distribution.

(d) θ should be a functional of the PDF $f_{\mathbf{X}}$.

Not all perturbations of the probability distribution would give rise to sensible definitions of temperature. In fact, it is quite possible to perturb the distribution of a classical system in a way which will produce results such as decreases in entropy. But as the classical definition (which introduces the notion of reversible heating) illustrates, only certain specific modes of perturbation will lead to sensible means for defining temperature. It is for this reason that we should consider only special classes of perturbations. The following example illustrates the difficulties of general perturbations.

Example. For the sake of this example, consider a system whose state is defined by a scalar random variable p which takes values y and whose Hamiltonian is simply $H(y) = y^2/2m$. Let $U_{[a,b]}(y)$ be the uniform distribution in the interval $[a,b]$ i.e.,

$$U_{[a,b]}(y) = \begin{cases} \frac{1}{|b-a|}, & \text{a} \leqslant \text{y} \leqslant \text{b}; \\ 0, & \text{otherwise}. \end{cases} \tag{13}$$

and let the probability distribution $f_P(y)$ be

$$f_P(y) = \tfrac{1}{2}U_{[-2,0]}(y) + \tfrac{1}{2}U_{[0,2]}(y).$$

Consider a perturbation of $f_P(y)$ to $f_{P'}(y)$ given by

$$f_{P'}(y) = \tfrac{1}{2}U_{[-(2/\gamma)-\Delta,-\Delta]}(y) + \tfrac{1}{2}U_{[(2/\gamma)+\Delta,+\Delta]}(y).$$

i.e., the perturbation shifts the distribution by a factor Δ and expands it by a factor γ.

It can be seen that

$$S(f_P) = k\ln 4, \tag{14}$$

$$S(f_{P'}) = k\ln\left(\frac{4}{\gamma}\right), \tag{15}$$

$$\mathcal{E}(f_P) = \frac{1}{4m}\frac{8}{3}, \tag{16}$$

$$\mathcal{E}(f_{P'}) = \frac{\gamma}{4m}\frac{\left(\frac{2}{\gamma}+\Delta\right)^3 - (\Delta)^3}{3}. \tag{17}$$

From (15), it can be seen that when $\gamma = 1$, $\Delta S = S(f_{\mathbf{X}'}) - S(f_{\mathbf{X}}) = 0$. However, for any $\Delta \geqslant 0$, $\Delta\mathcal{E} > 0$ and,–for this example,–$\frac{\Delta S}{\Delta\mathcal{E}} = 0$. Thus, simply shifting the probability distribution is an example of an *isentropic process*.

By choosing appropriate values for γ and Δ, it is possible to get negative values of $\frac{\Delta S}{\Delta\mathcal{E}}$, thus demonstrating the fact that not all perturbations are suitable for a meaningful definition of temperature. We will now introduce a specific form of perturbation for which the temperature defined in (7) will satisfy conditions 1–4 mentioned in Sec. III.

A. Additive perturbation as a statistical equivalent of infinitesimal heating

From the macroscopic perspective, one can view the definition of thermodynamic temperature in (3) as the mathematical embodiment of the following thought experiment: We increase the total energy of the particles by a small amount by heating the system. Then, we measure the change in entropy of the system. The ratio of the change in entropy to the change in energy is the inverse of the temperature. The key point to observe here is that this thought experiment depends upon the notion of heating the system, which guarantees both the entropy and energy increase (i.e., some sort of a diffusive process). We now propose a statistical realization of this notion.

We will describe the notion of heating a nonequilibrium system in a statistical way as follows: We heat the system by "perturbing" the random variable $\mathbf{X}$ by adding to it an *uncorrelated random variable* $\delta\mathbf{Y}$. When we do so, there will be a change in the total energy of the system as well as the entropy of the system. We can then define the temperature as the ratio between the change in the energy and that in the entropy [by generalizing the Gibbs relation (3)].

At this juncture, we note that Reguera *et al.* [20], have suggested an approach based on the consideration of diffusion processes in a generalized state space for the definition of thermodynamic forces. They suggest a definition of local thermodynamic temperature in a manner that is somewhat similar to ours [see Eq. (54) on page 21511 of [20]). In the current approach, however, it is the total entropy and not just the logarithm of the probability density function that is varied.

We will show that the temperature defined via the proposed generalization of the Gibbs relation (12) is related to the trace of the Fisher information matrix defined in (8), i.e., $\sum_i \mathcal{I}_{i,i}(f_{\mathbf{X}})$. To do this, we will first define an additive perturbation as follows.

Let $\mathbf{X}$ be the random variable of dimension $6N$, representing the state of the system. Let us assume that its distribution has a finite variance and density $f_{\mathbf{X}}$. Let $\mathbf{Y}$ be an *independently*

---

[2] Note that we use $f_{\mathbf{X}'}$ to denote the density function of the perturbed distribution instead of $f'_{\mathbf{X}}$. It must be understood that the random variable $\mathbf{X}$ is perturbed to obtain a new random variable $\mathbf{X}'$ whose PDF is $f_{\mathbf{X}'}$.

[3] We explicitly exclude systems with bounded state spaces, nuclear spin system lasers, vortex fluids, etc., which show negative absolute temperatures from consideration here. For these systems, due to the fact that they can have an upper bound to the energy per degree of freedom, adding energy by "heating" may decrease the entropy of the system.

distributed random variable whose distribution is symmetric with zero mean and whose covariance matrix is the identity matrix I. Now consider the random variable $\mathbf{X}' = \mathbf{X} + \sqrt{\delta}\mathbf{Y}$, where δ is a small number, representing a random perturbation of the state of the system. Since $\mathbf{X}$ and $\mathbf{Y}$ are independent, the PDF associated with $\mathbf{X}'$ is

$$f_{\mathbf{X}'}(\mathbf{x},\delta) = \int f(\mathbf{y}) N_{\delta\mathbf{Y}}(\mathbf{x}-\mathbf{y})\, d\mathbf{y}, \tag{18}$$

where $N_{\delta y}(\mathbf{x})$ is a symmetric PDF with covariance matrix δI_{ij}. Note that as δ tends to zero, the distribution approaches a Dirac measure centered at the origin and the effect of the "perturbation" of the state $\mathbf{X}$ disappears.

We now formally stipulate that the statistical equivalent of infinitesimally heating a system is the addition of a "small fraction" of a random variable with zero mean and *whose covariance matrix is the* $6N$*-dimensional identity matrix*, as defined in Eq. (18), in the limit as δ tends to zero. The idea is intimately related to the notion of a stochastic differential equation where we add infinitesimal Gaussian white noise to a variable.

The vector version of the de Bruijn identity (see Appendix A 2 for a definition and proof of the de Bruijn identity, as well as [13]) can be used to obtain a relationship between the derivative of entropy with respect to the perturbation parameter δ and the trace of the Fisher information matrix associated with $f_{\mathbf{X}}$. We present the main ideas behind the proof of this relationship since it is insightful. The key result regarding perturbations of this form is the following:

Lemma 3.1. If the covariance matrix of the random variable $\mathbf{Y}$ is the identity matrix, then the perturbed probability distribution $f_{\mathbf{X}'}$ satisfies a diffusion equation in the limit as δ vanishes, i.e.,

$$\left.\frac{\partial f_{\mathbf{X}'}}{\partial \delta}\right|_{\delta\to 0} = 1/2 \sum_{i=1}^{6N} \frac{\partial^2 f_{\mathbf{X}'}}{\partial x_i\, \partial x_i} = 1/2 \nabla^2 f_{\mathbf{X}'}, \tag{19}$$

where the notation ∇^2 stands for the Laplacian operator in the $6N$-dimensional phase space.

Proof (*see Appendix A 2*). There are a number of different ways to prove this result. If the PDF of $\mathbf{Y}$ is assumed to be Gaussian, then the result can be obtained by direct computation (see [12] Appendix C). The same result can also be established (if the perturbation is Gaussian), using techniques of the Ito calculus.The proof by Narayanan and Srinivasa [21] follows by noting that the moment generating function (MGF) of $\mathbf{X}'$ is the product of the MGFs of $\mathbf{X}$ and $\sqrt{\delta}\mathbf{Y}$. By expanding the latter as a Taylor series and allowing δ to go to zero, we obtain the result. Unlike the proof in [12], Narayanan and Srinivasa [21] make no specific assumptions regarding the form of the PDF other than it be symmetric and that its variance be the identity matrix. A different version of the proof is given in Appendix A 2.

B. Proposed definition of nonequilibrium thermodynamic temperature via the additive perturbation

Given the perturbation (18), we formally define the inverse temperature of the system to be

$$\frac{1}{\theta} = \lim_{\delta\to 0} \frac{S(f_{\mathbf{X}'}) - S(f_{\mathbf{X}})}{\mathcal{E}(f_{\mathbf{X}'}) - \mathcal{E}(f_{\mathbf{X}})} = \left.\frac{\partial S(f_{\mathbf{X}'})/\partial\delta}{\partial \mathcal{E}(f_{\mathbf{X}'})/\partial\delta}\right|_{\delta=0}. \tag{20}$$

Thus in order to obtain an explicit expression for the temperature, we need to find the derivatives of the entropy and the energy with respect to the parameter δ. This is a relatively straightforward operation: Note that the entropy and the total energy of the PDF $f_{\mathbf{X}'}$ are given by

$$\begin{aligned} S(f_{\mathbf{X}'}) &:= -k \int f_{\mathbf{X}'}(\mathbf{x},\delta) \ln\, f_{\mathbf{X}'}(\mathbf{x},\delta)\, d\mathbf{x}, \\ \mathcal{E}(f_{\mathbf{X}'}) &:= \int H(\mathbf{x}) f_{\mathbf{X}'}(\mathbf{x},\delta)\, d\mathbf{x}, \end{aligned} \tag{21}$$

where $H(\mathbf{x})$ is the Hamiltonian function.

We now differentiate these functions with respect to δ and get

$$\begin{aligned} \frac{\partial S(f_{\mathbf{X}'})}{\partial\delta} &= -k \int \frac{\partial f_{\mathbf{X}'}}{\partial\delta} \ln\, f_{\mathbf{X}'}\, d\mathbf{x}, \\ \frac{\partial \mathcal{E}(f_{\mathbf{X}'})}{\partial\delta} &= \int \frac{\partial f_{\mathbf{X}'}}{\partial\delta} H(\mathbf{x}) d\mathbf{x}, \end{aligned} \tag{22}$$

where we have simplified the rate of change of the entropy by using the result that $\frac{\partial}{\partial\delta}\int f_{\mathbf{X}'}(\mathbf{y},\delta) d\mathbf{y} = 0$.

Next, using the result (19), we obtain

$$\begin{aligned} \frac{\partial S(f_{\mathbf{X}'})}{\partial\delta} &= -k/2 \int (\nabla^2 f_{\mathbf{X}'}) \ln\, f_{\mathbf{X}'}(\mathbf{x},\delta) d\mathbf{x}, \\ \frac{\partial \mathcal{E}(f_{\mathbf{X}'})}{\partial\delta} &= 1/2 \int (\nabla^2 f_{\mathbf{X}'}) H(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{23}$$

Finally, upon using the divergence theorem and some manipulation assuming that the probability density vanishes sufficiently rapidly as the distance from the origin increases so that there are no contributions from the boundary terms at infinity, we obtain

$$\begin{aligned} \frac{\partial S(f_{\mathbf{X}'})}{\partial\delta} &= k/2 \int f_{\mathbf{X}'}(\mathbf{x},\delta)(\nabla\, \ln\, f_{\mathbf{X}'})(\nabla\, \ln\, f_{\mathbf{X}'}) d\mathbf{x}, \\ \frac{\partial \mathcal{E}(f_{\mathbf{X}'})}{\partial\delta} &= 1/2 \int f_{\mathbf{X}'}(x) \nabla^2 H(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{24}$$

Now comparing the first equation of (24) with the Fisher information matrix, and using the definition (20) for the thermodynamic temperature, we see that, as δ tends to zero,

$$\frac{1}{\theta} := \frac{\delta S}{\delta U} = \left.\frac{\partial S/\partial\delta}{\partial\mathcal{E}/\partial\delta}\right|_{\delta=0} = \frac{k\langle\|\nabla\, \ln\, f\|^2\rangle_f}{\langle\nabla^2 H\rangle_f} = \frac{k\, \mathrm{tr}\, \mathcal{I}}{\langle\nabla^2 H\rangle_f}, \tag{25}$$

thus establishing that *the thermodynamic temperature introduced in* (7) *also satisfies the Gibbs relation beyond equilibrium, provided the state of the system is perturbed in a certain way*. This provides yet another justification for the use of this temperature and helps clarify the notion of the statistical equivalent of heating a system.

IV. APPLICATION TO THE LANGEVIN SYSTEM

To show the practical significance of the definition (7), we first consider the evolution of the probability distribution of the velocity of a particle immersed in a fluid (heat bath) which satisfies the Langevin equations [10]. The analysis carried out here shows the remarkable interplay between the kinetic and thermodynamical temperatures of the system in the evolution of the entropy and the energy.

Since the heat bath is in equilibrium, it satisfies Boltzmann statistics and so its kinetic and thermodynamic temperatures are equal.

We will now show that

(1) the average rate of change of the kinetic energy of the particle is related to the difference in the *kinetic temperature* of the particle and that of the fluid;

(2) the rate of change of the entropy of the particle is related to the difference in the *nonequilibrium thermodynamic temperature* of the particle [as defined by (7)] and that of the heat bath; and

(3) the Cramer-Rao bound ([15,16]) (see Appendix A 3 for a definition) relating the Fisher information and the variance of a distribution is equivalent to the statement that the total entropy change be non-negative, i.e., it becomes a statement of the second law of thermodynamics for a Langevin system.

Consider a classical fluid which is in equilibrium at an absolute temperature T_∞ and consider a particle of mass m immersed in the fluid. For simplicity, we concentrate on any one component of the velocity of the particle and refer to it as v. We will refer to this particle as our subsystem and we will sometimes use the term heat bath to refer to the fluid. The velocity of this particle satisfies the well-known Langevin equation [10]. For brevity, we will refer to the subsystem and the surrounding fluid together as a Langevin system.

We show how the proposed notion of thermodynamical temperature provides valuable insight into the behavior of such Langevin systems. Specifically, we can use this definition to show that the energy transfer that occurs between the subsystem and the surrounding fluid can be meaningfully split into heat and work. We show that the "heat transfer" is driven by the difference in thermodynamic temperatures. We also show how, during the process of equilibration of the velocity distribution of the subsystem, entropy is produced by the process of internal dissipation which is "driven" by the difference in the thermodynamic temperatures of the body and the bath.

When demonstrating these results, in order to avoid needless complications involving integration by parts, which will unnecessarily complicate the ideas that we wish to present here, we will first restrict attention to just one component of the velocity.

A. The Langevin system and its associated Fokker-Planck equation

Following [10] (page 72, Eq. 6.2, and the lines following it), the dynamics of the motion of the subsystem is given by

$$\dot{v}(t) = -\gamma v(t) + \sqrt{\Gamma}\zeta(t), \tag{26}$$

where γ represents the damping due to the systemic component of the fluctuating force exerted on the particle by the heat bath, Γ is the strength of the stochastic force on the particle, and $\zeta(t)$ is a stationary, Gaussian δ-correlated Markov process with zero mean and an autocorrelation which is a Dirac measure of unit strength, i.e.,

$$E[\zeta(t)] = 0, \qquad E[\zeta(t)\zeta(t')] = \delta(t - t'), \tag{27}$$

where $E[\cdot]$ stands for the time average of the quantity.

Since the surrounding fluid is in equilibrium at a temperature T_∞, the fluctuation dissipation theorem demands that

$$\Gamma = 2m\gamma k T_\infty. \tag{28}$$

The probability density function of the velocity of the particle at time t is denoted as $f(v,t)$. Its evolution is governed by a Fokker-Planck equation (see [10]) which takes the form

$$\frac{\partial f}{\partial t} = \gamma \frac{\partial}{\partial v}(vf) + \frac{\gamma k T_\infty}{m}\frac{\partial f}{\partial v}, \tag{29}$$

where we have used the fluctuation dissipation theorem. Thus, the condition (28) ensures that the steady state velocity distribution of the particle will be the Boltzmann distribution corresponding to the fluid temperature T_∞.

In the sequel, rather than work with (26), we will use (29) to gain further insight into the process by which the velocity distribution reaches equilibrium.

B. The total energy and entropy of the subsystem

Since Eq. (26) is linear we can assume for simplicity that initially the particle is at rest, i.e., $v(t = 0) = 0$ so that by taking averages of the left and right hand sides of (26), and using the requirements (27) it is easy to see that *the average velocity of the particle at subsequent times is also zero.*[4] Thus, the average kinetic energy $\mathcal{E}(t)$ of the particle is

$$\mathcal{E}(t) = \frac{m}{2}\int_{-\infty}^{\infty} f(v,t)v^2\, dv := \frac{1}{2}kT_{\text{kin}}(t), \tag{30}$$

where $T_{\text{kin}}(t)$ is the *kinetic temperature of the particle* and is defined simply as $T_{\text{kin}}(t) = 2\mathcal{E}(t)/k$, where k is the Boltzmann constant.

Now the Shannon entropy of the subsystem is defined by

$$S(t) := -k\int_{-\infty}^{\infty} f(v,t)\ln\, f(v,t)dv. \tag{31}$$

With these two definitions in place, we can calculate the rate of change of energy and entropy of the subsystem.

C. Rate of change of energy: Role of the kinetic temperature

We can now calculate the rate of change of the kinetic energy of the particle $d\mathcal{E}/dt$ by using (29). In order to simplify the notation, we will set $\partial f/\partial v = f'$, $\partial f/\partial t = \dot{f}$, and $\partial\mathcal{E}/\partial t = \dot{\mathcal{E}}$. Thus, a straightforward calculation [using (29) and (30)] shows that

$$\dot{\mathcal{E}} = \frac{\gamma m}{2}\int_{-\infty}^{\infty}\left\{v^2(vf)' + \frac{kT_\infty}{m}v^2 f''\right\} dv. \tag{32}$$

Integrating the above equation by parts (the first term once and the second term twice), and using the conditions that[5]

---

[4] Here, by average, we mean ensemble average over many runs from the same starting velocity.

[5] These conditions follow from the fact that the partial differential equation (PDE) (29) is a linear PDE whose formal solution is given in, for example, [10], page 76, Eq. 6.14 and shows exponential decay of $f(v,t)$ with v, for all t. Thus $v^n p(\pm\infty) = v^n f'(\pm\infty) = f'\ \ln\ f(\pm\infty) = 0$.

$v^2 f'(\pm\infty) = v^n f(\pm\infty) = 0$, we obtain a remarkably simple expression for $\dot{\mathcal{E}}$, namely,

$$\dot{\mathcal{E}} = \gamma k\,[T_\infty - T_{\mathrm{kin}}(t)]\,. \tag{33}$$

In other words, the kinetic temperature (and hence the kinetic energy) shows exponential growth (from an initial value of 0) at a rate proportional to the difference between the fluid temperature and the kinetic temperature of the particle (subsystem). This is not a surprising result but has been detailed here for subsequent comparison with the rate of change of entropy and to show how the second law of thermodynamics is satisfied.

D. Rate of change of Shannon entropy and the nonequilibrium thermodynamic temperature

We now turn our attention to the rate of change of the entropy of the subsystem. We begin with the definition (31) for the entropy of the system, differentiating both sides with respect to time, and using (29) and after some manipulation using the fact that $\int_{-\infty}^{\infty} \dot{f}\,dv = 0$ we get

$$\dot{S} = -\gamma k \int_{-\infty}^{\infty} \left\{ (vf)' + \frac{kT_\infty}{m} f'' \right\} \ln f\, dv. \tag{34}$$

Again, integrating the above expression by parts [and using the fact that $vf(\pm\infty,t) = f' \ln f(\pm\infty,t) = 0$] and after some manipulation we get

$$\begin{aligned} \dot{S} &= \gamma k \left\{ T_\infty \frac{k}{m} \int_{-\infty}^{\infty} (f')^2/f\, dv - 1 \right\} \\ &= \gamma k \left\{ T_\infty \frac{k}{m} \int_{-\infty}^{\infty} f[(\ln f)']^2\, dv - 1 \right\}. \end{aligned} \tag{35}$$

Note that, unlike the case with the energy of the subsystem, the kinetic temperature is *not* involved. However, we note that the entropy rate involves a closely related statistical quantity—the Fisher information. We will elaborate on this point further.

The term

$$\mathcal{I} := \int_{-\infty}^{\infty} (f')^2/f\, dv = \int_{-\infty}^{\infty} f[(\ln f)']^2\, dv \tag{36}$$

is the Fisher information of the velocity distribution. It has several key properties of interest:

(i) It is non-negative.

(ii) It is in some sense a weighted measure of the curvature of the probability distribution function, so that extremely peaked or narrow distributions will correspond to large values of the integral.

(iii) For a Gaussian distribution (which corresponds to the steady state of the system, its value becomes the reciprocal of the variance, i.e., for Gaussian velocity distributions it is (to within a constant) the reciprocal of the absolute temperature.

Motivated by these remarks, and in view of (7) we will proceed to define the nonequilibrium thermodynamic temperature for the Langevin subsystem as

$$1/\theta := \frac{k}{m} \int_{-\infty}^{\infty} f[(\ln f)']^2\, dv. \tag{37}$$

It is an easy matter to verify that this is the same as the definition (7) since the Hamiltonian of the subsystem is simply the kinetic energy; in other words, we have

$$1/\theta := \frac{k\langle(\ln f')^2\rangle_f}{\langle \nabla_v^2 H\rangle}. \tag{38}$$

At this stage, this is no more than a convenient definition, just like the kinetic temperature. But we will see that it has far reaching consequences.

Substituting the definition (37) into the right-hand side of (35) we arrive at a very revealing form for the rate of change of the entropy of the particle, i.e.,

$$\dot{S}_{\mathrm{particle}} = \frac{\gamma k (T_\infty - \theta)}{\theta}. \tag{39}$$

Thus, the entropy increase of the subsystem (subsystem being the particle) depends upon the difference in the thermodynamic temperature between the surrounding fluid and the subsystem and the entropy flow direction is from the "hotter" to the "colder" region—further justifying our use of the word "thermodynamic temperature" for the quantity defined in (37). Furthermore, comparing Eqs. (39) and (33) we see the different and parallel roles played by the two kinds of temperature for the nonequilibrium system. Note that the kinetic temperature of the system has no role to play in the entropy changes but only the thermodynamic temperature as defined in this paper.

E. Heating, working, and the second law of thermodynamics for the Langevin system

We are now in a position to gain insight into the thermodynamics of the Langevin system, when we observe that the above equation (39) is of the form

$$dS = \frac{dQ}{\theta}. \tag{40}$$

Based on this observation, we can define *the heat transfer into the subsystem (or energy transfer by heating) is $\dot{Q} = \gamma k (T_\infty - \theta)$. In words, the rate of heat exchange is proportional to the difference in the thermodynamic temperatures.* Furthermore, *the rate of entropy increase is equal to the rate of heating divided by the thermodynamic temperature of the system.*

We can go a lot further if we now rewrite Eq. (33) in the form

$$\dot{\mathcal{E}} = \gamma k (T_\infty - \theta) + \gamma K (\theta - T_{\mathrm{kin}}) \tag{41}$$

and use our identification of $\gamma K (T_\infty - \theta)$ as the energy transfer by heating; then the above equation can be written as

$$\dot{\mathcal{E}} = \dot{Q} + \dot{W}, \tag{42}$$

where $\dot{Q} = \gamma K (T_\infty - \theta)$ and $\dot{W} = \gamma K (\theta - T_{\mathrm{kin}})$. In other words, it becomes possible to split the energy supplied to the subsystem as that due to heating [which is proportional to $(T_\infty - \theta)$ and that due to working which is proportional to $(\theta - T_{\mathrm{kin}})$]. At this stage this is just a formal split which does not provide any particular insight. However, we can now combine this split with the second law of thermodynamics to gain some further insight.

To do so, we need to find out what is the net rate of entropy change of the subsystem and the surrounding fluid—i.e., the whole Langevin system—as the velocity distribution of the particle evolves. We have already calculated the rate of entropy change of the subsystem (39). Now we will evaluate that of the surroundings: Since the particle is immersed in a surrounding fluid, the entropy change of the surrounding fluid is simply the energy gained or lost by the fluid divided by the fluid temperature, i.e.,

$$\dot{S}_{\text{fluid}} = \frac{-\dot{\mathcal{E}}}{T_\infty}, \tag{43}$$

where $\dot{\mathcal{E}}$, given by (33) is the energy gain by the subsystem (and hence lost by the bath). Thus the net rate of change in entropy of the whole system (particle + bath) is obtained by substituting (41) into (43) and adding the result to (39). After some manipulation, we arrive at

$$\begin{aligned}\dot{S}_{\text{total}} &= \dot{S}_{\text{particle}} + \dot{S}_{\text{fluid}} = \dot{Q}\left\{\frac{1}{\theta} - \frac{1}{T_\infty}\right\} + \frac{\dot{W}}{T_\infty}\\ &= \frac{\gamma K (T_\infty - \theta)^2}{T_\infty \theta} + \frac{\gamma K (T_{\text{kin}} - \theta)}{T_\infty}.\end{aligned} \tag{44}$$

The first term on the right-hand side of the third line of (44) is *the entropy produced due to heat transfer across a finite temperature difference* and is clearly positive since the numerator is a quadratic function and the denominator is a product of positive numbers.

The second term is *the entropy produced by mechanical dissipation*. This term requires careful consideration. Recall that if the system had an equilibrium velocity distribution, then $T_{\text{kin}} = \theta$. On the other hand, if T_{kin} is not equal to θ, the system is capable of doing work. To understand the significance of this term, consider a stream of particles all moving with the same speed with half along the positive x- axis and half along the negative x axis. The thermodynamic temperature θ of such a system [as defined by (7)] is zero since the PDF is composed of the sum of two Dirac measures, but its kinetic temperature is related to the square of the common speed. It is clear that we can extract work out of the system because the speed is deterministic. On the other hand, if we immerse such a system in a surrounding fluid the system will lose its capability to do work as it equilibrates with the fluid. It is this phenomenon that we usually refer to as internal dissipation and we can clearly see that the second term on the third line of (44) is associated with this.

F. The Cramer-Rao inequality and its relationship to entropy production in Langevin systems

But what guarantee is there that the second term in (44) is positive, i.e., what guarantee is there that the work is dissipated, i.e., lost irrecoverably? The answer lies in the Cramer-Rao bound (see, e.g., [15,16]). The Cramer-Rao bound states that, for any distribution $p(x)$ with variance σ^2 and Fisher information $\mathcal{I}$,

$$\sigma^2 \geqslant \frac{1}{\mathcal{I}}. \tag{45}$$

For the system considered here, since $T_{\text{kin}} = k\sigma^2$ and $\theta = k/\mathcal{I}$, *the Cramer-Rao bound states that the kinetic temperature is always greater than the thermodynamic temperature, i.e.,* $T_{\text{kin}} \geqslant \theta$, *thus guaranteeing that the second term on the third line of* (44) *is also positive*.

Thus, it is possible to extract useful work out of a nonequilibrium system by using a Carnot cycle and connecting it to a heat bath which has the same temperature as the kinetic temperature of the nonequilibrium system (see also the discussion in Sec. 4.2.2 in [1]).

V. OTHER DEFINITIONS OF NONEQUILIBRIUM TEMPERATURE AND THEIR RELATIONSHIP TO THE CURRENT DEFINITION

We have thus seen that use of the thermodynamic temperature defined by (7) provides considerable insight into the dynamics of the Langevin system. It is natural to ask, “How does this definition of temperature in Eq. (7) relate to other definitions of temperature found in the literature?” In this section we will seek to answer this question.

There are a number of alternative definitions of temperature for nonequilibrium systems that have been developed with different purposes in mind. The recent review article by Casas-Vasquez and Jou [1] provides an excellent summary of the various definitions (see, in particular, Table A on page 1942). We will itemize several of these definitions and show that current definition subsumes these various definitions for the Langevin system in an intuitive and pleasing way:

(a) *Kinetic temperature*. Perhaps the most universally accepted is the kinetic temperature, which is simply a measure of the kinetic energy of a system composed of classical particles. We have seen in the previous section on Langevin systems that this definition of the temperature, while it is important for the description of the change in *internal energy, is not the appropriate descriptor of the change in the entropy of the system*. However, the Cramer-Rao bound of statistics which implies that $\theta \leqslant T_{\text{kin}}$, is an important restriction since it guarantees that the internal dissipation is positive. Indeed, the difference between them can be considered as a measure of the “available information content” or “available order” in the momentum distribution which can be harnessed to do work. Moreover, the fact that for equilibrium (when the velocity distribution is Gaussian) these two definitions coincide is an intuitive and important result.

(b) *Temperature based on the Gibbs relation*. Another version of the temperature is that which is based on a generalization of the Gibbs relation $1/\theta := \partial S/\partial \mathcal{E}$ to nonequilibrium systems, in a manner that is different than the approach used here. Criado-Sancho, Jou, and Vazquez [8] use this definition for the case of sheared ideal gases subjected to a nonvanishing viscous stress. They obtain a probability distribution by maximizing the Shannon entropy subject to constraints of prescribed stress and kinetic energy and then obtain a nonequilibrium temperature based on this. They point out that this entropy-based nonequilibrium temperature is always lower than the kinetic temperature.

A routine calculation using the velocity distribution function given in [8] and using the definition (7) shows that the nonequilibrium temperature defined by them is *identical* to that defined here for the special MaxEnt distribution considered by them. Then the fact that it is lower than the kinetic temperature is simply a result of the Cramer-Rao inequality as discussed in the previous section.

(c) *Flux temperature*. Another alternative that has been suggested is that based on the ratio between the heat flux into a system divided by the entropy flux. For the Langevin system, this definition also leads to the same thermodynamic temperature as seen from Eq. (39) with the heat flux being identified as $k(T_\infty - \theta)$.

VI. CONCLUSION

In this paper, we have demonstrated that the notion of thermodynamical temperature can be extended to nonequilibrium distributions in a relatively straightforward way as the ratio between the Fisher information associated with the probability distribution and the average curvature of the Hamiltonian. This is done by using the Shannon or Gibbs entropy of a distribution and then by utilizing the relationship $dS = dQ_{\text{rev}}/T$ as a means for defining temperature. The resulting temperature is roughly speaking, the ratio of the mean curvature of the probability distribution and that of the Hamiltonian function and is closely related to an idea due to Landau ([14]).

Furthermore we have seen that in a Langevin system we can associate two temperatures with the particle: the kinetic temperature associated with the variance of the velocity distribution and the *thermodynamic temperature* associated with its Fisher information as defined here. Both play a role in the thermodynamics of the Langevin system. Of the two it is the thermodynamic temperature and not the kinetic temperature that is associated with heat transfer and the flow of entropy into the system. Finally, the Cramer-Rao bound can be interpreted as a statement of the second law of thermodynamics in the form of the non-negativity of work dissipated by the Langevin system.

We also demonstrated that the Gibbs relation (showing equality between the inverse temperature and the rate of change of entropy with energy) can be extended to the general nonequilibrium cases via a perturbation which is "diffusive" and the use of the well-known de Bruijn identity of information theory. The method outlined here can be used to define a variety of other temperatures by either using conditional probability distributions or marginalizing the PDF.

ACKNOWLEDGMENTS

This work was supported in part by grant CMMI-1028894 from the National Science Foundation.

APPENDIX: FISHER INFORMATION, CRAMER-RAO BOUND, AND THE DE BRUIJN IDENTITY

Fisher information is a quantity that is commonly used in parametric estimation [22]. In this Appendix, we will briefly summarize some important relationships between the Fisher information, entropy, and the variance associated with a distribution.

1. Fisher information

For a *scalar* random variable X with a probability density function $f_X(x)$, the Fisher information with respect to the location family, namely, $\mathcal{I}(f_X)$, is given by

$$\mathcal{I}(f_X) = \langle [(\ln f_X)']^2\rangle_{f_X} = \int_{-\infty}^{\infty} f_X(x)\left[\frac{\partial}{\partial x}\ln f_X(x)\right]^2 dx = \int_{-\infty}^{\infty} f_X(x)\left[\frac{\frac{\partial}{\partial x} f_X(x)}{f_X(x)}\right]^2 dx. \tag{A1}$$

For a vector-valued random variable $\mathbf{X}$ which takes on values in a vector space with $6N$ components, the i,jth entry of the Fisher information matrix is given by

$$\mathcal{I}_{i,j}(f_{\mathbf{X}}) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{y})\left[\frac{\partial}{\partial y_i}\ln f_{\mathbf{X}}(\mathbf{y})\frac{\partial}{\partial y_j}\ln f_{\mathbf{X}}(\mathbf{y})\right] d\mathbf{y},\quad i,j = 1,\ldots,6N. \tag{A2}$$

The *Fisher information* is just the trace of this matrix and is given by

$$\mathcal{I} = \left\langle \frac{\nabla f \nabla f}{f^2}\right\rangle_f. \tag{A3}$$

There is an interesting relationship between Fisher information and the relative entropy between a distribution and a shifted version of the distribution which is used in this paper. We will briefly review this result for the scalar case here. Let us consider two distributions $f_X(x+t)$ and $f_X(x-t)$ which are shifted versions of $f_X(x)$, shifted by t to the left and right, respectively. Then, Fisher information $\mathcal{I}(f_X)$ can be expressed as [13,22]

$$\mathcal{I}(f_X) = \lim_{t\to 0}\frac{1}{t^2}[D(f_X(x+t)||f_X) + D(f_X(x-t)||f_X)], \tag{A4}$$

where $D(f_X||g_X)$ is the relative entropy (or Kullback-Leibler distance) between the distributions f_X and g_X. This can be shown using the following result from [22]

$$\frac{1}{2}\mathcal{I}(\omega_0) = \lim_{t\to 0}\frac{1}{t^2}D(f_X(x;\omega_0+t)||f_X(x;\omega_0)).$$

Now, considering two parametric families $f_X(x;\omega) = f_X(x-\omega)$ and $f_X(x;\omega) = f_X(x+\omega)$ and applying this result and taking the average we get the desired result. Note that for both these families $\mathcal{I}(\omega_0) = \mathcal{I}(f_X)$ as defined in (A1), which gives us the left-hand side of (A4). We will now show that there is a close relationship between changes in entropy and the Fisher information in the sense that Fisher information is related to the second derivative of the relative entropy functional.

2. de Bruijn identity

The relationship between the entropy functional of a distribution f_P and the Fisher information associated with it is given by the de Bruijn identity [11,13]. This is precisely stated below.

Lemma A 1. Let X be a scalar random variable with finite variance and PDF given by f_X. Let Y be an independent symmetric random variable with unit variance and PDF f_Y. Let X' be the random variable given by

$$X' = X + \sqrt{\delta}Y. \tag{A5}$$

Then,

$$\left.\frac{\partial S(f_{Y'})}{\partial\delta}\right|_{\delta=0} = \frac{k}{2}\mathcal{I}(f_X). \tag{A6}$$

The proof of the above lemma follows from the observation that *as δ tends to zero,* (A5) *where f_Y is symmetric and has unit variance, reduces to a convolution with a Gaussian random variable*. This result can be proved in two different ways: by considering the moment generating function (or Laplace transform) for the PDF of X' in terms of those of X and Y. The proof using Laplace transforms can be found in [21]. Hence, we will provide a proof based on a direct Taylor series expansion. We begin by noting that the PDF for X' is given by

$$f'_X(x,\delta) = \int_{-\infty}^{\infty} f_X(u)\frac{1}{\sqrt{\delta}}\, f_Y\left(\frac{u-x}{\sqrt{\delta}}\right) du. \tag{A7}$$

We notice that for any fixed value of x, the integrand as a function of u is sharply peaked around $u = x$ as δ tends to zero [this is evident since $\frac{1}{\sqrt{\delta}} f_Y(\frac{u}{\sqrt{\delta}})$ will approach a Dirac measure as δ tends to zero]. In other words, the distribution is very steep around $u = x$ as δ tends to zero. A more precise statement can be made, based on the results of [23] and [24]. Thus we assume that for the random variable Y, its PDF $f_Y(y)$ has a unique maximum at $y = 0$, $f''_Y(0) < 0$ (“smoothness”), with unit variance. Then the variance of the random variable $\sqrt{\delta}Y$ clearly tends to zero as δ tends to zero (“steepness”). Finally, we assume that for any $\epsilon > 0$,

$$\int_{-\epsilon}^{\epsilon} \frac{1}{\sqrt{\delta}} f_Y\left(\frac{u}{\sqrt{\delta}}\right) du \to 1 \qquad \text{as} \quad \delta \to 0. \tag{A8}$$

We will also assume that all the higher moments of the random variable Y are finite. Given these conditions of “concentration,” smoothness, and steepness (Chen [24]), we can see that the main contribution to the convolution integral comes from the region around $u = x$. Hence, expanding $f_{X'}(u)$ as a Taylor series around $u = x$ and using this expansion in (A7), we get

$$\begin{aligned} f_{X'}(x,\delta) &= \int_{-\infty}^{\infty}\left(f_X(x) + f'_X(x)(u-x) + \frac{f''_X(x)}{2}(u-x)^2 + \cdots\right) \\ &\quad \times \frac{1}{\sqrt{\delta}} f_Y\left(\frac{u-x}{\sqrt{\delta}}\right) du. \end{aligned} \tag{A9}$$

In carrying out the integration with respect to u, we make use of the fact that $\int_{-\infty}^{\infty}(u-x)^n \frac{1}{\sqrt{\delta}} f_Y(\frac{u-x}{\sqrt{\delta}})du$ is the nth central moment of the random variable $\sqrt{\delta}Y$. Noting that $\sqrt{\delta}Y$ has zero mean and a variance of δ, we obtain

$$f_{X'}(x,\delta) = f_X(x) + \frac{\delta}{2} f''_X(x) + o(\delta^2). \tag{A10}$$

Therefore, in the limit $\delta \to 0$, we obtain

$$\left.\frac{\partial f_{X'}}{\partial\delta}\right|_{\delta=0} = \frac{1}{2}\frac{\partial^2 f_X}{\partial x^2}. \tag{A11}$$

Now that we have an explicit expression for $\partial f_{X'}/\partial\delta$, it is a straightforward computation to see that

$$\begin{aligned} \frac{\partial S(f_{X'})}{\partial\delta} &= -k\frac{\partial}{\partial\delta}\int_{-\infty}^{\infty} f_{X'}(x)\ln\, f_{X'}(x)dx \\ &= -k\int_{-\infty}^{\infty}\frac{\partial f_{X'}(x)}{\partial\delta}\ln\, f_{X'}(x)dx, \end{aligned} \tag{A12}$$

where we have used the fact that $\int_{-\infty}^{\infty}\frac{\partial f_{X'}(x)}{\partial\delta}dx = 0$. Taking the limit as $\delta \to 0$, and using (A11) and the definition of Fisher information from (A1), we get

$$\begin{aligned} \lim_{\delta\to 0}\frac{\partial S(f_{X'})}{\partial\delta} &= \lim_{\delta\to 0} -\frac{k}{2}\int_{-\infty}^{\infty} f''_X(x)\ln\, f_{X'}(x)dx \\ &= -\frac{k}{2}\int_{-\infty}^{\infty} f''_X(x)\ln\, f_X(x)dx = \frac{k}{2}\mathcal{I}(f_X). \end{aligned} \tag{A13}$$

In the same manner, the vector version of the de Bruijn identity can be stated as follows [11]:

Lemma A 2. Let $\mathbf{X}$ be a vector random variable and PDF given by $f_{\mathbf{X}}$. Let $\mathbf{Y}$ be an *independently distributed random variable with zero mean and unit covariance matrix* C and I *(where I is the unit matrix) and satisfies the steepness and smoothness conditions stated earlier*. Now consider the random variable $\mathbf{X}' = \mathbf{X} + \sqrt{\delta}\mathbf{Y}$. Then,

$$\left.\frac{\partial S(f_{\mathbf{X}'})}{\partial\delta}\right|_{\delta=0} = \frac{k}{2}\mathrm{tr}\,\mathcal{I}(f_{\mathbf{X}}). \tag{A14}$$

The proof is identical to the previous version and is based on the expansion of $p(\mathbf{y})$ as a vector Taylor series around $\mathbf{x}$.

3. Cramer-Rao bound

One of the important uses for Fisher information is in determining a lower bound on the estimation error of an unbiased estimator through the Cramer-Rao bound (see [15,16]). An important consequence of the bound is that for any random variable X with a variance σ_X^2 and Fisher information $\mathcal{I}(f_X)$,

$$\sigma_X^2 \geqslant \frac{1}{\mathcal{I}(f_X)}, \tag{A15}$$

i.e., the inverse of the Fisher information is a lower bound to the variance with equality only for the Gaussian distribution. Thus, for the case of the equilibrium distribution with a quadratic Hamiltonian, we can show that this is an equality. For the vector case, the Cramer-Rao bound can be used to show that

$$\mathrm{cov}(\mathbf{X}) - \mathcal{I}^{-1}(f_{\mathbf{X}})$$

is positive semidefinite.

4. Relationship between Fisher information, de Bruijn identity, and temperature for the Boltzmann distribution

From the point of view of this paper, colloquially, the Cramer-Rao bound relates the two different measures of temperature—kinetic temperature, which is a measure of the variance of the momentum distribution, and the Fisher information, which is connected to the "infinitesimal" heating process discussed here. For the particular case of the Boltzmann distribution, many of the results can be computed explicitly.

a. Fisher information for the Boltzmann distribution: Scalar case

For ease of exposition, we will begin with the scalar case. For the equilibrium Boltzmann distribution, i.e., when $f_X = \frac{e^{-\beta H(x)}}{Z}$, where $Z = \int_y e^{-\beta H(x)}\, dx$, there is a simple relationship between β and Fisher information. This can be derived as follows:

$$\mathcal{I}(f_X) = \int_{-\infty}^{\infty} -\beta H'(x) \frac{e^{-\beta H(x)}}{Z}[-\beta H'(x)]dx. \tag{A16}$$

Integrating by parts we get

$$\begin{aligned}\mathcal{I}(f_X) = {} & \left[-\beta H'(x) \frac{e^{-\beta H(x)}}{Z}\right]_{-\infty}^{\infty} \\ & + \int_{-\infty}^{\infty} \frac{e^{-\beta H(x)}}{Z}[\beta H''(x)]dx \end{aligned} \tag{A17}$$

$$= \int \frac{e^{-\beta H(x)}}{Z}\left[\beta H''(x)\right] dx = \beta \langle H'' \rangle_{f_X}. \tag{A18}$$

The first term on the right-hand side is zero because we will assume that $H(x)$ tends to infinity as $|x|$ tends to infinity. In doing so, we explicitly eliminate from consideration, systems that have maximum energy ceilings. Hence,

$$\beta = \frac{1}{kT} = \frac{\mathcal{I}(f_X)}{\langle H'' \rangle_{f_X}} = \frac{\langle [(\ln f_X)']^2 \rangle_{f_X}}{\langle H'' \rangle_{f_X}}. \tag{A19}$$

b. Vector case

Using a derivation similar to that for the scalar-valued random variable case, it can be shown that for the vector-valued case, for the equilibrium distribution $f_{\mathbf{X}}(\mathbf{x})$,

$$\beta = \frac{\langle (\nabla \ln f_{\mathbf{X}})^2 \rangle_{f_{\mathbf{X}}}}{\langle \nabla_v^2 H \rangle}. \tag{A20}$$

c. Uncorrelated perturbation and curvature temperature for the Boltzmann distribution

For the Boltzmann distribution, the results can be obtained in a particularly simple and striking manner, namely, by the direct use of the results (A10) and (A11). We will do this only for the scalar case to show the results, the vector case showing no difficulty whatsoever. Thus we begin by considering a PDF for the random variable X of the form $f_X(x) = e^{-\beta H(x)}/Z$, where $\beta = 1/kT$, $H(x)$ is the Hamiltonian, and Z, the normalizing parameter, is the partition function.

As with the previous case, we consider an uncorrelated perturbation of the system of the form $X' = X + \sqrt{\delta} Y$ where the PDF for Y satisfies the same conditions as before, then by following the identical derivation as before, we have

$$\begin{aligned}\left.\frac{\partial f_{X'}}{\partial t}\right|_{t=0} = \frac{1}{2}\frac{\partial^2 f_X}{\partial x^2} = {} & -\beta H'^2(x) e^{-\beta H(x)}/Z \\ & + \beta^2 H''(x) e^{-\beta H(x)}/Z, \end{aligned} \tag{A21}$$

where we have substituted the Boltzmann distribution for $f_X(x)$. Now integrating both sides from $-\infty$ to ∞ and noting that the left-hand side—being the derivative of a PDF—goes to zero, and the right-hand side are the expectations of the corresponding quantities, we again obtain the result that

$$\beta = \frac{\langle H''(x) \rangle}{\langle H'(x) \rangle}, \tag{A22}$$

which upon substitution of $\beta = 1/kT$, leads to the form (6).

[1] J. Casas-Vazquez and D. Jou, Rep. Prog. Phys. **66**, 1937 (2003).

[2] D. Jou, M. Criado-Sancho, and J. Casas-Vazquez, Physica A **358**, 49 (2005).

[3] D. Jou and M. S. Mongiovi, Phys. Lett. A **359**, 183 (2006).

[4] M. Criado-Sancho, D. Jou, and J. Casas-Vazquez, Phys. Lett. A **350**, 339 (2006).

[5] E. Bertin, K. Martens, O. Dauchot, and M. Droz, Phys. Rev. E **75**, 031120 (2007).

[6] J. L. Garden, J. Richard, and H. Guillou, J. Chem. Phys. **129**, 044508 (2008).

[7] W. G. Hoover and C. G. Hoover, Phys. Rev. E **77**, 041104 (2008).

[8] A. Criado-Sancho, D. Jou, and J. Casas-Vazquez, Phys. Lett. A **372**, 2172 (2008).

[9] T. Hatano and D. Jou, Phys. Rev. E **67**, 026121 (2003).

[10] V. Balakrishnan, *Elements of Nonequilibrium Statistical Mechanics* (Ane Books, New Delhi, 2008).

[11] A. J. Stam, Inf. Control **2**, 101 (1959).

[12] O. Johnson, *Information Theory and the Central Limit Theorem* (Imperial College Press, London, 2004).

[13] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, New York, 2006).

[14] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Butterworth-Heinemann, 1980).

[15] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1999).

[16] C. R. Rao, Bull. Calcutta Math. Soc. **77**, 81 (1945).

[17] T. Callen, *Thermodynamics*, 2nd ed. (Wiley, New York, 1985).

[18] H. H. Rugh, Phys. Rev. Lett. **78**, 772 (1997).

[19] A. Baranyai, J. Chem. Phys. **112**, 3964 (2000).

[20] D. Reguera, J. M. Rubi, and J. M. G. Vilar, J. Phys. Chem. B **109**, 21502 (2005).

[21] K. R. Narayanan and A. R. Srinivasa, e-print arXiv:0711.1460.

[22] S. Kullback, *Information Theory and Statistics* (Dover, New York, 1959).

[23] M. Evans and T. Swartz, Stat. Sci. **10**, 254 (1995).

[24] C. F. Chen, J. R. Stat. Soc. Ser. B (Stat. Methodol.) **47**, 540 (1985).