University of Massachusetts - Amherst

ScholarWorks@UMass Amherst

Mathematics and Statistics Department Faculty
Publication Series

Mathematics and Statistics

2006

Nonconcave entropies from generalized canonical ensembles

M Costeniuc

RS Ellis
University of Massachusetts - Amherst, rsellis@math.umass.edu

H Touchette

Follow this and additional works at: http://scholarworks.umass.edu/math_faculty_pubs

Part of the Physical Sciences and Mathematics Commons

Costeniuc, M; Ellis, RS; and Touchette, H, "Nonconcave entropies from generalized canonical ensembles" (2006). *Mathematics and Statistics Department Faculty Publication Series.* Paper 322.
http://scholarworks.umass.edu/math_faculty_pubs/322

This Article is brought to you for free and open access by the Mathematics and Statistics at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Mathematics and Statistics Department Faculty Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Nonconcave entropies from generalized canonical ensembles

Marius Costeniuc,[1, *] Richard S. Ellis,[2, †] and Hugo Touchette[3, ‡]

[1] *Max Planck Institute for Mathematics in the Sciences, Inselstr. 22-26, 04103 Leipzig, Germany*
[2] *Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA*
[3] *School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, UK*
(Dated: February 6, 2008)

It is well-known that the entropy of the microcanonical ensemble cannot be calculated as the Legendre transform of the canonical free energy when the entropy is nonconcave. To circumvent this problem, a generalization of the canonical ensemble which allows for the calculation of nonconcave entropies was recently proposed. Here, we study the mean-field Curie-Weiss-Potts spin model and show, by direct calculations, that the nonconcave entropy of this model can be obtained by using a specific instance of the generalized canonical ensemble known as the Gaussian ensemble.

Systems involving long-range interactions, such as gravitational forces or wave-particle interactions, have many unusual properties at equilibrium which set these systems apart from those involving short-range interactions [1, 2]. The root of most, if not all, of these unusual properties can be traced back to the fact that the microcanonical entropy function of long-range systems can be nonconcave as a function of the energy, even in the thermodynamic limit. By constrast, short-range systems can have nonconcave entropies at finite volume, but any nonconcave parts of the entropy must and will disappear in the thermodynamic limit, leaving a thermodynamic entropy which is always concave [3]. This has profound physical consequences because the concavity of the entropy determines ultimately whether the microcanonical ensemble (ME) is equivalent with the canonical ensemble (CE). If the entropy is concave, then the two ensembles are equivalent; otherwise they are nonequivalent at both the thermodynamic and macrostate levels [4, 5]. Thus short-range systems always have equivalent ensembles, but long-range systems need not; they can have equilibrium properties within the ME that have no counterparts in the CE. Many illustrations of this possibility have been given recently for a variety of systems, including those arising in the study of gravitation [6, 7], geostrophic turbulence [5, 8], plasmas [9], and magnetism [10, 11, 12, 13].

Systems with nonconcave entropies are special not only in terms of their equilibrium properties, but also in the way in which their entropy can be calculated. In cases of ensemble equivalence, it is well known that the microcanonical entropy function can be calculated as the Legendre transform of the canonical free energy function. This way of calculating the entropy goes back to Gibbs, and is nowadays the way of choice for calculating entropies because of the many practical advantages it offers. One well known advantage is that calculations carried out at the level of the free energy are generally more tractable than those carried out at the level of the entropy directly. Another advantage is that the definition of the free energy allows for many approximation schemes (e.g., perturbative expansions or variational principles), which are not available in the microcanonical ensemble.

The problem for systems having a nonconcave entropy is that the entropy does not correspond to the Legendre transform of the free energy anymore. This is obvious if one recalls that Legendre transforms yield only concave functions. Therefore, if one knows or suspects that a given system has a nonconcave entropy, then one also knows that this entropy cannot be calculated in the CE as the Legendre transform of the free energy. In this case, one must rely on the ME to perform that calculation (see, e.g., [12, 14]).

Our goal in this paper is to illustrate an alternative method for calculating entropy functions which goes beyond the CE in that it can be used to obtain nonconcave entropies. The method is not based on the ME. Rather, it is based on a *generalized canonical ensemble* (GCE) put forward recently by us, and works by modifying the structure of the Legendre transform through the use of a generalization of the free energy function. We have already presented the theory of the GCE and of its equivalence with the ME in [15, 16]; our goal here is to illustrate this theory with the help of a simple spin model known as the mean-field Curie-Weiss-Potts model [10, 13]. For this model, we shall calculate a nonconcave entropy function using the GCE, and show that the result agrees, in some appropriate limit, with the one obtained in the ME. In doing so, we shall explain a number of practical aspects of the GCE which have not been fully discussed before.

To begin, we review the definition of the GCE and the results establishing equivalence of this ensemble with the ME. Following our previous work [15, 16], we define the GCE by the following probability density over the microstates ω:

$$P_{n,g,\alpha}(\omega) = \frac{e^{-\alpha n h(\omega) - n g(h(\omega))}}{Z_{n,g}(\alpha)}. \tag{1}$$

In this expression, $h(\omega) = H(\omega)/n$ is the energy per particle or mean energy, α is a real parameter, $g(h)$ is a continuous but otherwise arbitrary function of h, and

$Z_{n,g}(\alpha)$ is a normalization constant given by

$$Z_{n,g}(\alpha) = \int e^{-\alpha n h(\omega) - n g(h(\omega))} d\omega. \qquad (2)$$

This function is naturally interpreted as a generalized partition function; from it, we define a generalized free energy via the usual limit

$$\varphi_g(\alpha) = -\lim_{n\to\infty} \frac{1}{n} \ln Z_{n,g}(\alpha). \qquad (3)$$

It is obvious that $P_{n,g,\alpha}(\omega)$ and, consequently, $Z_{n,g}(\alpha)$ and $\varphi_g(\alpha)$ reduce to their standard canonical-ensemble expressions for $g = 0$ [17]. This explains why we use the term "generalized" CE.

To complete this short review of the GCE, we state the generalized Legendre transform that will be used hereafter to obtain the entropy of the ME from the free energy of the GCE. At this point, we recall the thermodynamic-limit definition of the microcanonical entropy:

$$s(u) = \lim_{n\to\infty} \frac{1}{n} \ln \int \delta(h(\omega) - u)\, d\omega. \qquad (4)$$

In terms of $\varphi_g(\alpha)$ we then have the following: *If, for a given choice of g, $\varphi_g(\alpha)$ is differentiable at α, then*

$$s(u_{g,\alpha}) = \alpha u_{g,\alpha} + g(u_{g,\alpha}) - \varphi_g(\alpha), \qquad (5)$$

where $u_{g,\alpha} = \varphi_g'(\alpha)$ (see [16], Theorem 4). The choice $g = 0$ yields, as expected, the standard Legendre transform

$$s(u_\beta) = \beta u_\beta - \varphi(\beta), \qquad u_\beta = \varphi'(\beta), \qquad (6)$$

where $\varphi(\beta) = \varphi_{g=0}(\alpha = \beta)$ stands for the free energy of the CE. For the remaining, it is useful to note that $u_{g,\alpha}$ represents the equilibrium mean energy of the GCE with function g and parameter α [18]. This parallels the case of the CE for which $u_\beta = \varphi'(\beta)$ represents the equilibrium mean energy of the CE with inverse temperature β.

We now come to the main point of this paper, which is to consider a system known to have a nonconcave entropy, and show that the system's entropy can be derived from the sole point of view of the GCE by first calculating $\varphi_g(\alpha)$ for that system and then by applying to $\varphi_g(\alpha)$ the generalized Legendre transform shown in (5). The model that we consider for this purpose is the q-state mean-field Curie-Weiss-Potts (CWP) model [10, 13, 19] defined by the Hamiltonian

$$H(\omega) = -\frac{1}{2n} \sum_{i,k=1}^{n} \delta(\omega_i, \omega_k). \qquad (7)$$

In this expression, $\delta(x,y)$ is the Kronecker symbol, ω_i is a spin variable taking values in the set $\Lambda = \{\theta^1, \theta^2, \dots, \theta^q\}$, and ω represents the complete configuration of n spins, i.e., $\omega = (\omega_1, \omega_2, \dots, \omega_n)$. For simplicity, we consider

the case $q = 3$, so that spins can only take one of three possible values: θ^1, θ^2 and θ^3.

Despite the rather simple nature of the CWP model, a complete, analytical calculation of its generalized free energy seems out of reach. For a start, the integral defining the generalized partition $Z_{n,g}(\alpha)$ does not seem to be explicitly solvable, so that the calculation of $\varphi_g(\alpha)$ cannot proceed from a direct evaluation of Eqs. (2) and (3). Fortunately, there is an alternative way to calculate $\varphi_g(\alpha)$ suggested by large deviation techniques. It involves three steps [5, 15]:

(i) Rewrite the energy per spin $h(\omega) = H(\omega)/n$ as a function of some macrostate or order parameter $\nu(\omega)$. In our case, we choose the macrostate to be the empirical vector $\nu = (\nu_1, \nu_2, \nu_3)$, the jth component of which is defined by

$$\nu_j = \frac{1}{n} \sum_{k=1}^{n} \delta(\omega_k, \theta^j). \qquad (8)$$

A short calculation shows indeed that $h(\omega) = \tilde{h}(\nu(\omega))$, where $\tilde{h}(\nu) = -\frac{1}{2}\langle \nu, \nu \rangle$. The component ν_j represents the relative number of spins in ω equal to θ^j; $0 \le \nu_j \le 1$, $\sum_j \nu_j = 1$. For this reason, $\nu(\omega)$ is often called the statistical distribution of spin state or *spin distribution* for short. The function $\tilde{h}$ is called the energy representation function.

(ii) Derive the expression of an entropy function $\tilde{s}(\nu)$ for the macrostate ν. In the case of the empirical vector, that entropy function is well known to be given by the statistical (Boltzmann-Shannon) entropy

$$\tilde{s}(\nu) = -\sum_{j=1}^{3} \nu_j \ln \nu_j. \qquad (9)$$

(iii) Calculate $\varphi_g(\alpha)$, finally, using the following representation formula:

$$\varphi_g(\alpha) = \inf_\nu \left\{ \alpha \tilde{h}(\nu) + g(\tilde{h}(\nu)) - \tilde{s}(\nu) \right\}. \qquad (10)$$

The infimum is evaluated over all allowed values of ν, i.e., all triplets $\nu = (\nu_1, \nu_2, \nu_3)$ such that $0 \le \nu_j \le 1$ and $\sum_j \nu_j = 1$.

Equation (10) was solved numerically using a quadratic function g of the form $g(u) = \gamma u^2/2$, where $\gamma \ge 0$ [20]. This choice of g defines a GCE known as the *Gaussian ensemble* (GE) [21]. Fig. 1 shows the result of the numerical computation for three increasing values of the parameter γ. For each value, the free energy of the GE, which we denote by $\varphi_\gamma(\alpha)$, has the particularity that it possesses one nondifferentiable point, located at α_γ, separating two differentiable branches. The left branch is a linear function of α with a slope independent of γ. The value of the slope is equal to $-1/6$, which is the maximum admissible value of h [22]. To obtain the value of $s(u)$ at this specific value of h, which we denote by $u_{\max}$,
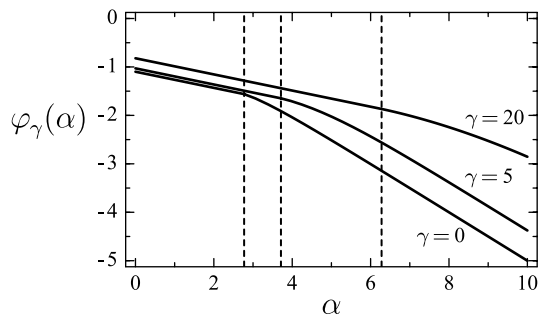
3

FIG. 1: Gaussian free energy of the CWP model for different values of γ.
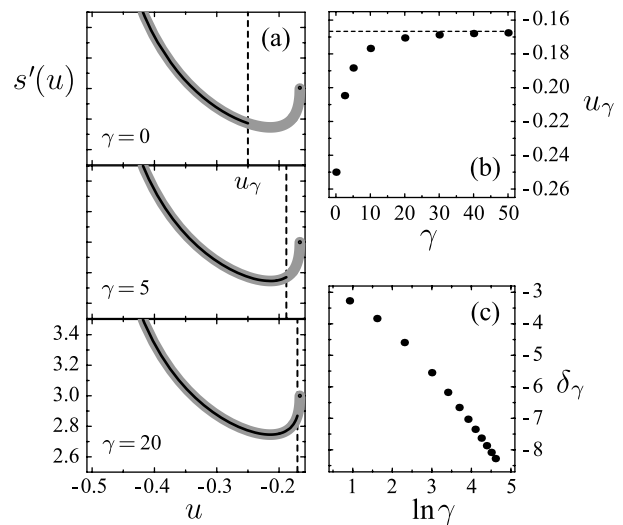

FIG. 2: (a) Derivative of the entropy function found for the CWP model using the GE (black line) and ME (gray line). The dashed lines show the mean energy value u_γ up to which there is equivalence between the GE and ME. (b) Convergence of u_γ towards $u_{\max}$. (c) $\delta_\gamma = \ln|u_{\max} - u_\gamma|$ versus $\ln\gamma$.

we simply need at this point to apply Eq. (5) to the left branch of $\varphi_\gamma(\alpha)$. The result obtained is $s(u_{\max}) = \ln 3$, which is the correct value found in the ME corresponding physically to the entropy of the excited state of the CWP model. The right, differentiable branch of $\varphi_\gamma(\alpha)$ can be treated similarly with the difference that $\varphi'_\gamma(\alpha)$ is not a constant but takes values in a range of the form $[u_{\min}, u_\gamma]$, where $u_{\min} = \varphi'_\gamma(\infty) = -1/2$ and $u_\gamma = \varphi'_\gamma(\alpha_\gamma + 0)$. The calculation of Eq. (5) for the right branch of φ_γ therefore yields $s(u)$ in that range. The result is displayed in Fig. 2(a) together with the exact entropy function $s(u)$ derived directly from the ME (see Eq. (4.3) in [13]). Note that, in order to accentuate the relatively shallow nonconcave region of $s(u)$ near $u_{\max}$, we have plotted the derivative of the entropy in Fig. 2(a) rather than the entropy itself.

The comparison in Fig. 2(a) of the entropy obtained in the GE with the "true" entropy of the ME shows that the use of the GE enables us to obtain $s(u)$ in a gradual manner by increasing the value of the parameter γ. For $\gamma = 0$, $s(u)$ is recovered from the ground-state mean energy $u_{\min} = -1/2$ up to the value $u_{\gamma=0} = -1/4$, while for $\gamma = 5$, $s(u)$ is recovered from $u_{\min}$ to $u_5 \approx -0.1883 > u_0$. Increasing the parameter γ further to $\gamma = 20$, we obtain $s(u)$ from $u_{\min}$ up to $u_{20} \approx -0.1706 > u_5$. The case $\gamma = 0$ corresponds to the CE, so that the part of $s(u)$ obtained in this ensemble corresponds to the concave part of $s(u)$ determined by Maxwell's construction or, equivalently, by the set of supporting lines of that function; see [16] for details. Thus, we see that a virtue of the GE with $\gamma > 0$ over the CE is that the former ensemble is able to recover nonconcave points of the entropy function while the latter is not.

By continuing to increase γ, we can calculate the value u_γ up to which the GE is able in to recover $s(u)$. The results are shown in Fig. 2(b), as well as in Fig. 2(c). The conclusion that we reach from these two figures is that the GE recovers the complete entropy function of the CWP model in the limit $\gamma \to \infty$, since $u_\gamma \to u_{\max} = -1/6$ in that limit. Hence, for a finite, positive value of γ, the GE recovers only a part of $s(u)$, but the part of $s(u)$ "missed"

by the GE can be made arbitrary small by choosing large enough values of γ. In this sense, we say that the GE and the ME are *asymptotically equivalent*. In general, we know that the GE should be equivalent with the ME at all values of h whenever γ is greater than the maximum value of $s''(u)$ (see [15], Theorem 5.2). In the case of the CWP, $s''(u)$ diverges at $u_{\max}$, which explains why the two ensembles become equivalent for all $u \in [u_{\min}, u_{\max}]$ only in the limit $\gamma \to \infty$. An analytical study of this limit is presented in [23] (see also [15, 16]).

We can go further in our study of the CWP model by showing that the GE can be used not only to calculate $s(u)$ but also to calculate the microcanonical equilibrium values of ν of that model. This aspect of the GE, which we refer to as the *macrostate level* of ensemble equivalence, is illustrated in Fig. 3. This figure shows two sets of plots. The one on the left shows the equilibrium values of ν calculated in the GE as a function of α [24], while the one on the right displays the same points but now as a function of their mean energy $u = \tilde{h}(\nu)$ (black line). Note that only the first component of ν is plotted because the equilibrium value of ν in the GE has the form $\nu = (a, b, b)$ [20]. The comparison with the ME is established by plotting the equilibrium value of a in the ME (see Eq. (4.1) in [13]) for all $u \in [u_{\min}, u_{\max}]$ (gray line).

The results obtained from these calculations show that the GE recovers the equilibrium macrostates of the ME for all $u \in [u_{\min}, u_\gamma]$, with u_γ approaching $u_{\max}$ for increasing γ. The value u_γ here is the same as the one found previously when calculating $s(u)$, so that the range of mean-energy values for which we have equivalence of ensembles at the level of ν is exactly the range over which
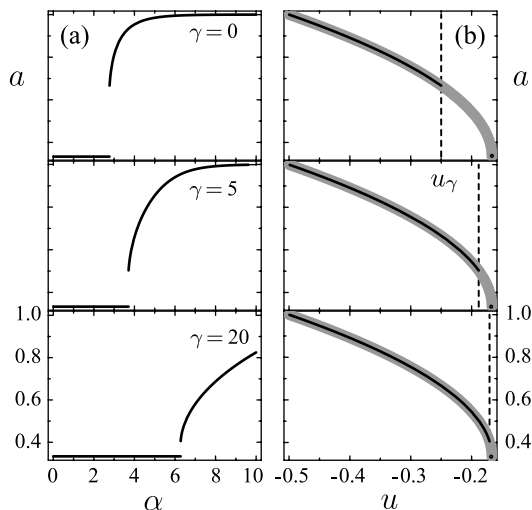
FIG. 3: (a) First component a of the equilibrium value of ν in the GE as a function of α. (b) Equilibrium value of a in the GE as a function of the equilibrium mean energy of that ensemble (black line); equilibrium value of a in the ME (gray line). The dashed lines show the mean energy value u_γ up to which there is equivalence between the GE and ME.

we have equivalence of ensembles at the *thermodynamic level*, i.e., the level of $s(u)$. This is not a coincidence but a direct result of the fact that equivalence of the GE and ME at the thermodynamic level entails, essentially, the equivalence of these ensembles at the macrostate level [15, 16]. As a result, we can conclude that the GE recovers, in the limit $\gamma \to \infty$, the microcanonical equilibrium values of ν for all $u \in [u_{\min}, u_{\max}]$, since it completely recovers $s(u)$ in the same limit.

In the end, it is interesting to note that the asymptotic equivalence of the GE and ME is reflected, at the macrostate level, by the fact that the jump of a, seen in the GE as a function of α, disappears as $\gamma \to \infty$ (Fig. 3). It can be proved in general that the disappearance of this jump, which is responsible for the nondifferentiable point of $\varphi_\gamma(\alpha)$ (Fig. 1), is a sufficient condition for the complete equivalence of the GE and ME (see [16], Theorem 4). From a physical point of view, this means that the absence of a first-order phase transition in the GE is a sufficient condition for the equivalence of the GE and ME.

To summarize, we have considered a mean-field version of the Potts model to show that the nonconcave entropy function of this spin model can be calculated using a generalization of the canonical ensemble known as the Gaussian ensemble. The large deviation formalism that we have used to study this model is totally general in that it can be applied to other models known to have nonconcave entropies. Our future work will aim at studying a number of these models (see, e.g., [25]), in addition to studying other types of generalized canonical ensembles, including the one defined by the function

$g(u) = \gamma|u|$. Although the Gaussian ensemble is thought to be universal, in the sense that it should be able to recover any form of entropy function [15, 16], other generalized ensembles could be useful, in that they may lead to more tractable calculations than those carried out in the Gaussian ensemble (see, e.g., [26]).

* Electronic address: marius.costeniuc@gmail.com
† Electronic address: rsellis@math.umass.edu
‡ Electronic address: htouchet@alum.mit.edu

[1] T. Dauxois, S. Ruffo, E. Arimondo, and M. Wilkens, eds., *Dynamics and Thermodynamics of Systems with Long Range Interactions*, vol. 602 of Lecture Notes in Physics (Springer, New York, 2002).
[2] D.H.E. Gross, Phys. Rep. **279**, 119 (1997).
[3] D. Ruelle, *Statistical Mechanics: Rigorous Results* (W.A. Benjamin, Amsterdam, 1969).
[4] G.L. Eyink and H. Spohn, J. Stat. Phys. **70**, 833 (1993).
[5] R.S. Ellis, K. Haven, and B. Turkington, J. Stat. Phys. **101**, 999 (2000).
[6] D. Lynden-Bell and R. Wood, Mon. Notic. Roy. Astron. Soc. **138**, 495 (1968); D. Lynden-Bell, Physica A **263**, 293 (1999).
[7] P.-H. Chavanis, Phys. Rev. E **65**, 1056123 (2002); *ibid*, Astron. & Astrophys. **401**, 15 (2003); P.-H. Chavanis and I. Ispolatov, Phys. Rev. E **66**, 036109 (2002).
[8] R.S. Ellis, K. Haven, and B. Turkington, Nonlinearity **15**, 239 (2002).
[9] R.A. Smith and T.M. O'Neil, Phys. Fluids B **2**, 2961 (1990).
[10] I. Ispolatov and E.G.D. Cohen, Physica A **295**, 475 (2000).
[11] J. Barré, D. Mukamel, and S. Ruffo, Phys. Rev. Lett. **87**, 030601 (2001).
[12] R.S. Ellis, H. Touchette, and B. Turkington, Physica A **335**, 518 (2004).
[13] M. Costeniuc, R.S. Ellis, and H. Touchette, J. Math. Phys. **46**, 063301 (2005).
[14] J. Barré, F. Bouchet, T. Dauxois, and S. Ruffo, J. Stat. Phys. **119**, 677 (2005).
[15] M. Costeniuc, R.S. Ellis, H. Touchette, and B. Turkington, J. Stat. Phys. **119**, 1283 (2005);
[16] M. Costeniuc, R.S. Ellis, H. Touchette, and B. Turkington, Phys. Rev. E **73**, 026105 (2006).
[17] We use α rather than β to emphasize the fact that α is not the physical inverse temperature when $g \neq 0$.
[18] H. Touchette, M. Costeniuc, R.S. Ellis, and B. Turkington, Physica A **365**, 132 (2006).
[19] F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).
[20] For a quadratic g, the minimizers in Eq. (10) have the form $\nu = (a, b, b)$ [13]. With this property, the calculation of $\varphi_\gamma(\alpha)$ is reduced to a simple 1D minimization problem.
[21] M.S.S. Challa and J.H. Hetherington, Phys. Rev. Lett. **60**, 77 (1988); *ibid*, Phys. Rev. A **38**, 6324 (1988).
[22] The range of admissible values of h for the general q-state CWP model is $[-1/2, -1/(2q)]$ [13].

[23] M. Costeniuc, R.S. Ellis, H. Touchette, and B. Turking-ton, Proc. MSRI Conf. (2007).

[24] The equilibrium values of ν in the GE correspond to the global minimizers obtained by solving Eq. (10).

[25] H. Touchette and C. Beck, cond-mat/0507379.

[26] R. Toral, Physica A **365**, 85 (2006).