Methods for calculating nonconcave entropies

Hugo Touchette
School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK
E-mail: H.Touchette@qmul.ac.uk

Received 2 March 2010
Accepted 12 April 2010
Published 14 May 2010

Online at stacks.iop.org/JSTAT/2010/P05008
doi:10.1088/1742-5468/2010/05/P05008

Abstract. Five different methods which can be used to analytically calculate entropies that are nonconcave as functions of the energy in the thermodynamic limit are discussed and compared. The five methods are based on the following ideas and techniques: (i) microcanonical contraction, (ii) metastable branches of the free energy, (iii) generalized canonical ensembles with specific illustrations involving the so-called Gaussian and Betrag ensembles, (iv) the restricted canonical ensemble, and (v) the inverse Laplace transform. A simple long-range spin model having a nonconcave entropy is used to illustrate each method.

Keywords: classical phase transitions (theory), metastable states

ArXiv ePrint: cond-mat/1003.0382
The recent study of many-body systems interacting via long-range potentials, such as gravitating particles or unscreened plasmas (see this issue and [1]–[4] for other examples), has revealed an interesting property of the entropy that went unnoticed for a surprisingly long time, namely, that it can be nonconcave as a function of the energy in the thermodynamic limit. It was known before this discovery that the entropy of finite-size systems could be nonconcave because of boundary or surface contributions (see, e.g., [5]–[7] and [8]–[10] for implicit references to this idea in the context of finite-size scaling). But, in almost all cases, it was assumed that this nonconcavity disappears on taking the thermodynamic limit because the ‘bulk’ entropy, which is supposedly always concave as a function of energy, dominates over the ‘surface’ entropy.

We now know that the situation is more complicated and, at the same time, more interesting. If the interaction in a homogeneous many-particle system is short-range (see [3] for a definition), then the thermodynamic entropy of this system is essentially always concave as a function of its energy, as was proved years ago by Ruelle [11] (see also Lanford [12]) using a separation or so-called ‘subadditivity’ argument. However, if the interaction is long-range, then the subadditivity argument does not work, and the entropy can be either concave or nonconcave [1]–[3]. The latter possibility has for some time been known to arise in mean-field systems, but the crucial and relatively recent input that the study of systems such as gravitating particles has provided is that nonconcave entropies also arise in systems involving physical interactions that are genuinely ‘long-range’. In this sense, nonconcave entropies cannot be dismissed as an artifact of

doi:10.1088/1742-5468/2010/05/P05008
mean-field approximations—they are ‘physical’. In fact, it is by now established that the nonconcavity of the entropy is related to many interesting physical phenomena, including:

- the existence of energy regions where the heat capacity, defined microcanonically, is negative (something forbidden in the canonical ensemble) [1]–[3], [13]–[15];
- the appearance of first-order phase transitions as well as metastable states in the canonical ensemble [3, 16, 17];
- the nonequivalence of the microcanonical and canonical ensembles at the thermodynamic and equilibrium macrostate levels [18, 19];
- a possible ergodicity breaking in microcanonical dynamics [20].

We will not be directly concerned here with any of these phenomena; instead, we will consider a problem of a more technical nature having to do with how entropies that are nonconcave can be calculated in practice. Our starting point is the age-old thermodynamic result stating that the entropy of a thermodynamic system is the Legendre transform of its free energy, and vice versa. This duality property can only be true, obviously, if the entropy is concave, since Legendre transforms only yield concave functions. Hence, if one knows or suspects that the entropy of a system is nonconcave, then this entropy cannot be obtained from the canonical ensemble by calculating the Legendre transform of the free energy. How can the entropy be calculated then?

The goal of this paper is to describe and illustrate in the simplest way possible the various methods that have been proposed in the last few years to answer this question. Five methods, which cover the latest studies on the topic of nonconcave entropies, will be covered:

(i) Microcanonical contraction (section 3).
(ii) Metastable branches of the canonical free energy (section 4).
(iii) Generalized canonical ensembles, with special emphasis on the Gaussian ensemble and a new ensemble called the Betrag ensemble (section 5).
(iv) The restricted canonical ensemble (section 6).
(v) The inverse Laplace transform (section 7).

Each of these methods will be illustrated with a simple spin system introduced in [21, 22] as a pedagogical model of equilibrium statistical mechanics having a nonconcave entropy. The model, as will become clear, is not meant to represent any real physical system, but has the advantage of being exactly solvable, which makes it useful for demonstrating how the methods listed above work in practice, and for illustrating along the way many general results about nonconcave entropies.

In theory, all of the methods that will be discussed can be used to calculate any nonconcave entropy, but we will see that some may be more effective or more ‘tractable’ than others in practice, depending on the system considered. The question of how to select the ‘right’ method for a given system will be discussed at the end of the paper, along with some open problems related to generalized canonical ensembles.
2. The canonical ensemble

Before we start discussing methods that can be used to calculate nonconcave entropies, let us convince ourselves that the Legendre transform of the canonical free energy does not yield the microcanonical entropy when the latter is nonconcave. This will give us the opportunity to introduce the basic notation used in this paper.

Let $H_N(\omega)$ be the Hamiltonian of a classical $N$-particle system, and let $\omega$ denote a configuration or microstate of this system, and $\Lambda_N$ its configuration space. We define the thermodynamic free energy or free energy density of the canonical ensemble via the limit

$$\varphi(\beta) = \lim_{N \to \infty} -\frac{1}{N} \ln Z_N(\beta),$$

where

$$Z_N(\beta) = \int_{\Lambda_N} e^{-\beta H_N(\omega)} d\omega$$

is the $N$-particle partition function. The problem that concerns us here is to determine whether $\varphi(\beta)$ can be used to obtain the thermodynamic entropy or entropy density of the microcanonical ensemble, defined by

$$s(u) = \lim_{N \to \infty} \frac{1}{N} \ln \Omega_N(u),$$

where

$$\Omega_N(u) = \int_{\omega \in \Lambda_N; h_N(\omega) = u} d\omega = \int_{\Lambda_N} \delta(h_N(\omega) - u) d\omega$$

is the density of states, which gives the volume (or, more pictorially, the number) of microstates $\omega$ that have a mean energy $h_N(\omega) = H_N(\omega)/N$ equal to $u$.

As mentioned in the introduction, the common answer to this problem given by most thermodynamics textbooks (see [23] for an exception) is that $s(u)$ can always be obtained as the Legendre transform of $\varphi(\beta)$, since $\varphi(\beta)$ and $s(u)$ are Legendre transforms of one another. This implies, incidentally, that one can also always calculate $\varphi(\beta)$ as the Legendre transform of $s(u)$. But is this really the complete answer? What if $\varphi(\beta)$ is nondifferentiable, as is the case when there is a first-order phase transition in the canonical ensemble? How does one define the Legendre transform for this case? Also, what if $s(u)$ is nonconcave?

The latter question naturally arises when studying long-range systems. If $s(u)$ is nonconcave, then the Legendre transform of $\varphi(\beta)$ cannot yield $s(u)$ simply because Legendre transforms only yield concave functions [24,25]. By reading a bit about convex analysis, one learns in fact that the Legendre transform of $\varphi(\beta)$, defined as

$$\varphi^*(u) = \inf_{\beta} \{\beta u - \varphi(\beta)\},$$

yields a concave function corresponding in general to the concave envelope of $s(u)$.1 Thus, if $s(u)$ is concave, then $s(u) = \varphi^*(u)$, which is to say that $s(u)$ is the Legendre transform

1 The transformation defined by equation (5) is actually a generalization of the Legendre transform known as the Legendre–Fenchel transform, which can be applied to nondifferentiable as well as nonconcave functions; see [24,25] for more details. In this paper, we refer to this transform as the Legendre transform for simplicity.
of \( \varphi(\beta) \). However, if \( s(u) \) is nonconcave, then \( s(u) \neq \varphi^*(u) \). In this case, only the concave part of \( s(u) \) will be recovered from the Legendre transform of \( \varphi(\beta) \).

These mathematical results are illustrated in the next example using a simple spin model which will stay with us for the rest of the paper. For proofs of these results, the reader should consult the classical book of Rockafellar [24] or the more readable treatise of van Tiel [25].

**Example 2.1 (The block-spin model [22]).** Consider the following Hamiltonian:

\[
H_N = \frac{N}{2} y + \sum_{i=1}^{N/2} \sigma_i, \tag{6}
\]

where \( y \) and \( \sigma_i, \ i = 1, 2, \ldots, N/2, \) are spin variables taking values in the set \( \{-1, +1\} \). The first term in the Hamiltonian represents the energy of a block of \( N/2 \) ‘frozen’ spins constrained to take the same spin value \( y \) (\( N \) is assumed to be even). The second term represents the energy of a second block of \( N/2 \) ‘free’ spins which do not interact with each other or with the first block of spins.

The entropy density \( s(u) \) of this spin system can easily be calculated from the definition of this quantity, i.e., from equations (3) and (4). This calculation is presented in [22] with the result

\[
s(u) = \frac{1}{2} \begin{cases} s_\sigma(2u + 1) & u \in [-1, 0) \\ s_\sigma(2u - 1) & u \in [0, 1], \end{cases} \tag{7}
\]

where

\[
s_\sigma(v) = - \left( \frac{1-v}{2} \right) \ln \left( \frac{1-v}{2} \right) - \left( \frac{1+v}{2} \right) \ln \left( \frac{1+v}{2} \right) \tag{8}
\]

is the entropy of the ‘free’ spins \( \sigma_i \). As is clear from figure 1, \( s(u) \) is a nonconcave function of \( u \), as it has more than one maximum\(^2\) and its derivative is non-monotonic.

To verify that this nonconcave entropy cannot be obtained from \( \varphi(\beta) \), we proceed to calculate \( \varphi^*(u) \). The direct evaluation of \( Z_N(\beta) \) yields for this model

\[
Z_N(\beta) = (e^{\beta N/2} + e^{-\beta N/2})(e^\beta + e^{-\beta})^{N/2}, \tag{9}
\]

\(^2\) It is easy to see that a concave function can only have one maximum.
and so
\[ \varphi(\beta) = -\frac{1}{2}|\beta| - \frac{1}{2} \ln(2 \cosh \beta). \]  

(10)

From this expression, plotted in figure 1, we then compute the Legendre transform defined in equation (5). This calculation can again be found in [22]; the result is
\[ \varphi^*(u) = \begin{cases} 
  s_\sigma(2u + 1) & u \in [-1, -\frac{1}{2}] \\
  \ln 2 & u \in [-\frac{1}{2}, \frac{1}{2}] \\
  s_\sigma(2u - 1) & u \in (\frac{1}{2}, 1]. 
\end{cases} \]  

(11)

This function is plotted in figure 1. We see, as announced, that \( \varphi^*(u) \) is a concave function corresponding to the concave envelope of \( s(u) \). The part of \( s(u) \) that coincides with \( \varphi^*(u) \), seen for \( u \in (-\frac{1}{2}, \frac{1}{2}) \), is called the nonconcave part of \( s(u) \).

We can get more insight into the results presented above and illustrated in figure 1 by noting the following extra results from convex analysis:

- The entropy \( s \) at the point \( u \) is equal to the Legendre transform of \( \varphi \) if and only if one can place a line above the graph of \( s(u) \) that touches \( s(u) \) without intersecting it. When this is possible, we say that \( s \) admits a supporting line at \( u \). Mathematically, this property is expressed as follows: \( s = \varphi^* \) at \( u \) if, and only if, there exists \( \beta \in \mathbb{R} \) such that
  \[ s(v) \leq s(u) + \beta(v - u) \]  
  for all \( v \). See [19, 26] for more details on the concept of supporting lines.

- The concave envelope or concave hull of \( s(u) \) is obtained by constructing the set of all the supporting lines of \( s(u) \); see figure 2.\(^3\)

- If \( \varphi \) is differentiable at \( \beta \), then
  \[ s(u_{\beta}) = \varphi^*(u_{\beta}) = \beta u_{\beta} - \varphi(\beta), \]  
  \[ u_{\beta} = \varphi'(\beta). \]  

(13)

The first result provides a useful geometric understanding of the points of the entropy that can or cannot be obtained from the Legendre transform of the free energy. This is illustrated in figure 2. As for the third result, it shows that \( s(u) \) is correctly given by the Legendre transform of \( \varphi(\beta) \) for all points \( u \) lying in the image of the derivative of \( \varphi(\beta) \), i.e., all points \( u \) such that \( u = \varphi'(\beta) \) for some \( \beta \in \mathbb{R} \). This implies, in particular, that if \( \varphi(\beta) \) is everywhere differentiable and the image of \( \varphi' \) coincides with the domain of \( s \) (i.e., the set of allowed or ‘realizable’ values for \( H_N/N \)), then \( s = \varphi^* \) holds globally. We will often come back to this result in the rest of the paper when treating generalized canonical ensembles.

\(^3\) Mathematically, the concave envelope of \( s(u) \) is also given by the smallest concave function that majorizes \( s(u) \) or by a geometrical construction known as Maxwell’s construction; see section 4 of [27] or chapter 3 of [21].

\(^4\) This result is essentially a consequence of the so-called Gärtner–Ellis theorem of large deviation theory; see section 5.2 of [26].
3. Microcanonical contraction

The block-spin model that we have studied in the previous example is simple enough that we can obtain its nonconcave entropy $s(u)$ directly from the definition of this quantity. But, of course, for more realistic and hence more complex models, one should not hope to be able to obtain $s(u)$ in this way. How can $s(u)$ be calculated then?

One answer finds its inspiration from large deviation theory, and attempts to derive $s(u)$ by maximizing another entropy subject to the energy constraint. The basis and hypotheses behind this method are the following [18]. Given the Hamiltonian $H_N(\omega)$, one must be able to find a macrostate $M_N(\omega)$ such that the following two properties are satisfied:

- The mean energy $h_N(\omega) = H_N(\omega)/N$ can rewritten as a function of $M_N(\omega)$ either exactly or asymptotically in the thermodynamic limit $N \to \infty$. Mathematically, this implies that there exists a function $\tilde{h}(m)$ such that
  \[
  |h_N(\omega) - \tilde{h}(M_N(\omega))| \to 0
  \]
  uniformly for all $\omega \in \Lambda_N$ as $N \to \infty$. The function $\tilde{h}(m)$ is called the energy representation function.

- There exists an entropy function $\tilde{s}(m)$ for $M_N(\omega)$, which is to say that the following limit exists:
  \[
  \tilde{s}(m) = \lim_{N \to \infty} \frac{1}{N} \ln \Omega_N(M_N = m),
  \]
  where
  \[
  \Omega_N(m) = \int_{\omega \in \Lambda_N : M_N(\omega) = m} d\omega = \int_{\Lambda_N} \delta(M_N(\omega) - m) d\omega
  \]
  counts the number of microstates such that $M_N(\omega) = m$. The function $\tilde{s}(m)$ is called the macrostate entropy.
When both of these conditions are satisfied, it is relatively easy to show (see [21,27]) that
\[ s(u) = \sup_{m: \tilde{h}(m) = u} \tilde{s}(m). \] (17)

This formula is what we refer to as a \textit{microcanonical contraction}. The word ‘contraction’ comes from the fact that this formula can be derived from a result known in large deviation theory as the \textit{contraction principle} [26].

The same result can also be seen as a form of maximum entropy principle expressing \( s(u) \) as the constrained maximization of the macrostate entropy \( \tilde{s}(m) \). It can be shown that the constrained maximizers of \( \tilde{s}(m) \) such that \( \tilde{h}(m) = u \) correspond physically to the equilibrium values of \( M_N \) in the microcanonical ensemble with mean energy \( u \). By denoting the set of such maximizers by \( \mathcal{E}^u \), we can therefore re-express equation (17) as \( s(u) = \tilde{s}(\mathcal{E}^u) \).

\textbf{Example 3.1.} The calculation of the entropy \( s(u) \) of the block-spin model via the contraction formula of equation (17) is presented in [22]. It is easy to see that a natural choice of macrostate for this model is \( m = (y, p) \), where \( y \) is the spin value of the ‘frozen’ block of spins, and \( p \) is the proportion of +1 spins in the block of ‘free’ spins. In terms of this macrostate, we obviously have
\[ \tilde{h}(y, p) = \frac{y}{2} + p - \frac{1}{2}. \] (18)
The macrostate entropy for this choice of macrostate is, up to a factor of 1/2, the Boltzmann–Shannon entropy:
\[ \tilde{s}(p) = -\frac{1}{2}p \ln p - \frac{1}{2}(1 - p) \ln(1 - p). \] (19)

This entropy does not depend on \( y \) because the block of ‘frozen’ spins does not contribute to the entropy of the model. We refer the reader to [22] again for the calculation of \( s(u) \) based on this energy representation function and macrostate entropy.

Other examples of calculations of entropies based on the microcanonical contraction formula include the mean-field Blume–Emery–Griffiths model [27,28], the mean-field Potts model [29,30] (see also Example 5.4 of [26]), the mean-field Hamiltonian model [30,31], the so-called mean-field \( \phi^4 \) model [32,33], as well as a variant of this model having a nonconcave entropy \( s(u, m) \) as a function of the energy \( u \) and magnetization \( m \) [34].

From this list and the form of equation (17), one might conclude that this equation is only good for mean-field models, as these are presumably the only models whose Hamiltonian can be re-expressed as a function of some specially chosen macrostates or ‘mean fields’. However, this is not the case. In theory, at least, it is always possible to express the Hamiltonian of any model, including short-range models, as a function of an infinite-dimensional macrostate known as the \textit{empirical process} (see [35] and section 5.3.4 of [26]). But given the infinite-dimensional nature of this macrostate, the calculation of \( s(u) \) from it is typically impractical if not impossible. For this reason, the microcanonical contraction formula has mostly, if not only, been used in the context of mean-field and long-range models, which are in any case the models for which nonconcave entropies are expected to arise.

doi:10.1088/1742-5468/2010/05/P05008
4. Metastable branches of the free energy

The microcanonical contraction formula discussed in the previous section involves a constrained maximization problem which can be transformed, following the theory of Lagrange multipliers, into an unconstrained maximization by considering the function

\[ G_\beta(m) = \tilde{s}(m) - \beta \tilde{h}(m), \]

which involves a Lagrange multiplier \( \beta \) associated with the constraint \( \tilde{h}(m) = u \). The question that we ask in this section is: can we obtain the set of constrained (global) maximizers of \( \tilde{s}(m) \) with \( \tilde{h}(m) = u \), which was denoted in the previous section by \( E_u \), from the set \( E_\beta \) of (global) unconstrained maximizers of the new function \( G_\beta(m) \)?

The answer is, no, at least if \( s(u) \) is nonconcave. Using techniques similar to those leading to the microcanonical contraction formula, it can indeed be proved that the canonical free energy \( \varphi(\beta) \) is given by the set \( E_\beta \) through the formula

\[ \varphi(\beta) = \inf_m \{ \beta \tilde{h}(m) - \tilde{s}(m) \} = -\sup_m G_\beta(m) = -G_\beta(E_\beta). \]

Therefore, if we were able to obtain \( E_u \) from \( E_\beta \), we would be in a position to obtain \( s(u) \) from \( \varphi(\beta) \). But we know that this is not possible when \( s(u) \) is nonconcave. Hence, \( E_\beta \) cannot be the same as \( E_u \) in general.

This result may be surprising but does not contradict the theory of Lagrange multipliers. What this theory actually says is that the global maximizers of \( \tilde{s}(m) \) subject to the constraint \( \tilde{h}(m) = u \) are contained in the set of critical points of \( G_\beta(m) \), which include the global maximizers of \( G_\beta(m) \), but also any local maximizers, minimizers and saddle points that this function may have. The theory simply does not say what the constrained maximizers of \( \tilde{s}(m) \) correspond to at the level of \( G_\beta(m) \). To obtain this information, one must go deeper into the structure of the microcanonical contraction formula, equation (17), and its canonical counterpart, equation (21), to find the following (see [17]):

(i) If \( s \) is nonconcave at \( u \), then the elements of \( E_u \) correspond either to local minima of \( G_\beta(m) \) or saddle points of this function depending on the local curvature of \( s(u) \).

(ii) If \( s \) is concave at \( u \), then the elements of \( E_u \) are also elements of \( E_\beta \) for some \( \beta \in \mathbb{R} \), which is consistent with the fact that \( s = \varphi^* \) in this case.

We reach two conclusions from these results. The first is that the microcanonical and canonical ensembles are equivalent at the level of thermodynamic properties and equilibrium values of macrostates when \( s(u) \) is concave [18,19,21,26]. The second is more pragmatic: it is possible to obtain \( s(u) \) from the knowledge of the critical points of \( G_\beta(m) \), but we must consider all the critical points of this function, not just its global maximizers [16]. We must locate, in particular, the local maxima of \( G_\beta(m) \), which corresponds physically to metastable values of \( M_N \) in the canonical ensemble, as well as saddle points of \( G_\beta(m) \), which correspond to unstable values of \( M_N \) in the same ensemble. This conclusion is put to use in the next example.

**Example 4.1.** For the block-spin model, the function \( G_\beta(m) \) has the simple form

\[ G_\beta(y, p) = \tilde{s}(p) - \beta \tilde{h}(y, p), \quad y \in \{-1, 1\}, \quad p \in [0, 1], \]

\[ \text{doi:10.1088/1742-5468/2010/05/P05008} \]
Methods for calculating nonconcave entropies

Figure 3. Left: plots of $\varphi_1(\beta)$ (blue), $\varphi_2(\beta)$ (purple), and $\varphi(\beta)$ (dashed). The branches of $\varphi_1(\beta)$ and $\varphi_2(\beta)$ that lie above $\varphi(\beta)$ are metastable branches of the free energy. Right: the complete nonconcave entropy is recovered by taking the Legendre transform of $\varphi_1(\beta)$ and $\varphi_2(\beta)$.

where $\tilde{s}(p)$ and $\tilde{h}(y,p)$ are the macrostate entropy and energy representation function, respectively, introduced in example 3.1. The calculation of the critical points of this form of $G_\beta(y,p)$ can be found in [22] as well as in section 5.1 of [21]. For the purpose of this section, there are two points to note about this solution:

(i) $G_\beta(y,p)$ has no saddle points, but has local maxima, i.e., metastable states, for all $\beta \in \mathbb{R}$;
(ii) if we denote the set of metastable points of $G_\beta(y,p)$ for a given $\beta$ by $\mathcal{M}_\beta$, then $s(u) = \tilde{s}(\mathcal{M}_\beta)$ for $u = \tilde{h}(\mathcal{M}_\beta) \in [-\frac{1}{2}, \frac{1}{2}]$.

The last point demonstrates that the metastable states of $G_\beta(y,p)$ can be used, as claimed, to recover nonconcave points of $s(u)$. In fact, for this particular model, $\mathcal{M}_\beta$ recovers the whole nonconcave region of $s(u)$. The set $\mathcal{E}_\beta$ of stable or equilibrium macrostates recovers only the concave parts of $s(u)$.

The nonconcave points of $s(u)$ can also be related, at the thermodynamic level, to metastable branches of the canonical free energy function $\varphi(\beta)$ rather than metastable states of the canonical ensemble, as was done above. This is illustrated next.

Example 4.2. The exact partition function shown in equation (9) can be put in the form

$$Z_N(\beta) = Z_N^{(1)}(\beta) + Z_N^{(2)}(\beta),$$

where

$$Z_N^{(1)}(\beta) = e^{\beta N/2}(e^\beta + e^{-\beta})^{N/2}, \quad Z_N^{(2)}(\beta) = e^{-\beta N/2}(e^\beta + e^{-\beta})^{N/2}.$$  \hfill (24)

From these two partition functions, it is natural to define two free energy functions, $\varphi_1(\beta)$ and $\varphi_2(\beta)$, using the definition of the free energy shown in equation (1):

$$\varphi_1(\beta) = -\frac{1}{2}\beta - \frac{1}{2}\ln(2 \cosh \beta), \quad \varphi_2(\beta) = \frac{1}{2}\beta - \frac{1}{2}\ln(2 \cosh \beta).$$  \hfill (25)

The relation between these two free energies and $\varphi(\beta)$ is easily found by using the expression of equation (23) in the limit defining $\varphi(\beta)$ to find

$$\varphi(\beta) = \inf\{\varphi_1(\beta),\varphi_2(\beta)\} = \begin{cases} \varphi_2(\beta) & \beta \leq 0 \\ \varphi_1(\beta) & \beta > 0. \end{cases}$$  \hfill (26)

doi:10.1088/1742-5468/2010/05/P05008
This result is illustrated in figure 3. The branches of $\varphi_1(\beta)$ and $\varphi_2(\beta)$ that do not contribute to $\varphi(\beta)$ can be interpreted as metastable branches of $\varphi(\beta)$, since they continue, in the sense of analytical continuation, the two ‘stable’ branches of $\varphi(\beta)$ while remaining above the ‘true’ minimal equilibrium free energy $\varphi(\beta)$.\(^5\) As was the case for the metastable states of $G_\beta(y,p)$ studied in the previous example, these metastable branches of $\varphi(\beta)$ completely determine $s(u)$ through the Legendre transform:

$$s(u) = \sup\{\varphi_1^*(u), \varphi_2^*(u)\} = \begin{cases} \varphi_1^*(u) & u \in [-1, 0) \\ \varphi_2^*(u) & u \in [0, 1] \end{cases}.$$  \(27\)

This result is also illustrated in figure 3, and is proved directly by calculating the Legendre transforms of $\varphi_1(\beta)$ and $\varphi_2(\beta)$.

The idea of analytically continuing the free energy around a phase transition point to characterize metastable states was studied for some time in the context of short-range models [36]–[39]. However, it was somewhat abandoned after it was realized that continued free energies do not provide correct estimates for the lifetime of metastable states. With hindsight, one could argue that these estimates were wrong because metastable states of short-range systems do not persist in the thermodynamic limit; they arise because of surface effects or, more precisely, because of a ‘sub-bulk’ nonconcavity of the entropy, which, as mentioned, disappears in the thermodynamic or ‘bulk’ limit. For this reason, metastable states of short-range systems cannot be associated with metastable branches of the free energy because, if they were, then the entropy would have to be nonconcave.

For long-range systems, the situation is different, since these can have states that are truly metastable in the thermodynamic sense, and are associated with metastable branches of $\varphi(\beta)$, as illustrated by the previous examples. The same phenomenon has also been studied in the context of gravitating systems; see, e.g., [40] for a recent review. Still, one must be careful: it is known that different entropies having the same concave envelope lead, through the Legendre transform, to the same $\varphi(\beta)$, so it is not possible to uniquely determine $s(u)$ by analytically continuing $\varphi(\beta)$. The existence of metastable branches of $\varphi(\beta)$ must be determined, ultimately, by calculating $Z_N(\beta)$, as was done in the previous example.

5. Generalized canonical ensembles

The use of generalized ensembles to obtain nonconcave entropies was extensively discussed in previous publications (see [17], [41]–[44]), so we will be brief here. The idea of this method is to obtain $s(u)$ from the Legendre transform of a modified or generalized free energy function having the form

$$\varphi_g(\beta) = \lim_{N \to \infty} -\frac{1}{N} \ln Z_{N,g}(\beta),$$  \(28\)

where

$$Z_{N,g}(\beta) = \int_{\Lambda_N} e^{-\beta H_N(\omega) - Ng(H_N(\omega)/N)} d\omega.$$  \(29\)

\(^5\) Recall that equilibrium states of the canonical ensemble correspond, according to Gibbs, to those states minimizing the free energy.

doi:10.1088/1742-5468/2010/05/P05008
Methods for calculating nonconcave entropies

is the generalized partition function. In these expressions, \( g \) is a function of the mean energy \( H_N/N \), assumed to be continuous. Different choices for this function determine different generalized canonical ensembles that can be used, under some conditions on \( g \) (see below), to obtain \( s(u) \) even when this function is nonconcave.

In the following, we will consider two generalized ensembles corresponding to two choices of \( g \), and will show that both of these ensembles recover the nonconcave entropy of the block-spin model, but that one is more effective than the other for this purpose. The general result at play behind these two ensembles was proved in [41,42], and can be stated in a simple form as follows: if, for a given choice of function \( g \), \( \varphi_g(\beta) \) is differentiable at \( \beta \), then

\[
s(u) = \beta u - \varphi_g(\beta) + g(u)
\]

for \( u = \varphi_g'(\beta) \). This modified Legendre transform, which is written in short form as \( s = \varphi_g^* + g \), generalizes the standard Legendre transform shown in equation (13) in an obvious way. In particular, if we are able to find a function \( g \) such that \( \varphi_g(\beta) \) is everywhere differentiable and the range of \( H_N/N \) coincides with the image of \( \varphi_g'(\beta) \), then \( s = \varphi_g^* + g \) for all \( u \) in the domain of \( s(u) \). We will see next that such a function can be constructed in some appropriate limit.

5.1. The Gaussian ensemble

The choice \( g(u) = \gamma u^2/2 \) with \( \gamma \in \mathbb{R} \) in equation (29) leads to a generalized ensemble known as the Gaussian ensemble. This ensemble was first introduced in the context of Monte Carlo simulations by Hetherington [45,46], who also discussed its physical interpretation in terms of finite-size heat baths (see [47]–[49]), and was later re-investigated in the context of nonconcave entropies in [41,42]. It has been applied so far to obtain the nonconcave entropies of two spin models, namely, the mean-field Potts model [44] and the mean-field Blume–Emery–Griffiths model [50]. We apply it next to the block-spin model.

**Example 5.1.** The first natural step to take in trying to calculate the generalized partition (29) with \( g(u) = \gamma u^2/2 \) is to use the Gaussian integral

\[
e^{-\gamma u^2/2} = \sqrt{\frac{\gamma}{2\pi}} \int_{-\infty}^{\infty} e^{-\gamma t^2/2 - i\gamma ut} \, dt,
\]

valid for \( \gamma > 0 \), to obtain

\[
Z_{N,\gamma}^G(\beta) = \sqrt{\frac{\gamma N}{2\pi}} \int_{-\infty}^{\infty} dt \, e^{-\gamma N t^2/2} Z_N(\beta + i\gamma t)
\]

for the Gaussian partition function. Unfortunately, the resulting integral over \( t \) cannot be evaluated in general, and in particular not for the block-spin model, despite the simplicity of this model. This point will be discussed in more detail in the concluding section. A similar integral can be obtained for \( \gamma < 0 \), and this one can actually be evaluated using a saddle point approximation, but, as will be discussed below, the case \( \gamma < 0 \) is not useful for obtaining nonconcave entropies.

It is possible in the end to calculate the Gaussian free energy \( \varphi_\gamma(\beta) \) of the block-spin model by generalizing the macrostate representation of \( \varphi(\beta) \) found in equation (21) to

\[
doi:10.1088/1742-5468/2010/05/P05008 12
Figure 4. Gaussian ensemble. Upper left: Gaussian free energy $\varphi^G_\gamma(\beta)$ of the block-spin model for different values of $\gamma$ (see right). Lower left: $\beta$-derivative of $\varphi^G_\gamma(\beta)$. Right: quadratic Legendre transform of the Gaussian free energy, which recovers the nonconcave entropy $s(u)$ (dashed line) as $\gamma \to \infty$. The entropy is recovered precisely where $\varphi^G_\gamma$ is differentiable.

The Gaussian ensemble:

$$\varphi^G_\gamma(\beta) = \inf_{y,p} \left\{ \beta \tilde{h}(y,p) + \frac{\gamma}{2} \tilde{h}(y,p)^2 - \tilde{s}(p) \right\}. \quad (33)$$

The minimization problem involved in this expression is easily solved. For $\gamma > 0$ and $\beta < 0$, the expression between the curly brackets above is globally minimized for $y = 1$ and $p$ solving

$$\beta + \gamma p - \tilde{s}'(p) = 0. \quad (34)$$

For $\gamma > 0$ and $\beta > 0$, on the other hand, the same expression is globally minimized for $y = -1$ and $p$ solving

$$\beta + \gamma(p - 1) - \tilde{s}'(p) = 0. \quad (35)$$

Both equations for $p$ are transcendental equations that can be solved numerically to obtain $\varphi^G_\gamma(\beta)$. The result of this numerical calculation is shown in figure 4. As can be seen, the Gaussian free energy $\varphi^G_\gamma(\beta)$ obtained for $\gamma > 0$ retains the nondifferentiable point of $\varphi(\beta) = \varphi^G_0(\beta)$ at $\beta = 0$, but tends to become ‘less nondifferentiable’ when $\gamma$ increases, as its left and right derivatives approach 0 for increasing $\gamma$. This behavior of $\varphi^G_\gamma(\beta)$ is illustrated in the lower left panel of figure 4, which shows the plot of $u_{\beta\gamma} = \partial_\beta \varphi^G_\gamma(\beta)$ for increasing values of $\gamma$. From this plot, we immediately see that the Legendre transform of the Gaussian ensemble, which takes the form

$$s(u_{\beta\gamma}) = \beta u_{\beta\gamma} + \frac{\gamma}{2} u_{\beta\gamma}^2 - \varphi^G_\gamma(\beta), \quad (36)$$

doi:10.1088/1742-5468/2010/05/P05008
recovered more and more nonconcave points of $s(u)$ as $\gamma$ increases. This is confirmed by the right-hand side plot of figure 4, which shows the part of the entropy $s(u)$ recovered by equation (36) for $\gamma = 0$ (canonical ensemble), $\gamma = 5$, $\gamma = 10$, and $\gamma = 15$.

The Gaussian ensemble is an interesting statistical ensemble not only because it can be used to recover nonconcave entropies, as illustrated above, but also because it allows for a natural ‘parabolic’ generalization of the concept of supporting lines discussed in section 2. Because of the quadratic nature of the function $g$ defining this ensemble, it can indeed be proved (see [41, 42]) that $s(u)$ is given by the ‘quadratic’ Legendre transform of $\varphi_\gamma(\beta)$ shown in equation (36) if

$$s(v) \leq s(u) + \beta(v-u) + \frac{\gamma}{2}(v-u)^2$$

for all $v$. We say in this case that $s$ admits a supporting parabola with curvature $\gamma$ at the point $u$; see figure 5. Therefore, the points of $s(u)$ that are recovered by the Gaussian ensemble with parameter $\gamma$ are all (and only) those points that admit a supporting parabola with curvature $\gamma$ or, equivalently, all the points of $s(u)$ coinciding with the parabolic hull of this function; see figure 5.

We see from this result that the entropy of the block-spin model is obtained only in the limit $\gamma \to \infty$ because the entropy of this model has a cusp or ‘corner’ at $u$, which can only be ‘supported’ by a degenerate parabola of infinite curvature, as shown in the center plot of figure 5. The same result also explains why the Gaussian ensemble is able to recover nonconcave points of $s(u)$: by modifying the Legendre transform with an added quadratic term, we are able to ‘reach’ with a supporting parabola points of $s(u)$ that cannot be ‘reached’ with a supporting line; see figure 5. This implies, naturally, that the Gaussian ensemble with $\gamma > 0$ can only recover more points of $s(u)$ as compared with the canonical ensemble, at least if $s(u)$ has a nonconcave region (and is nondegenerate). (Of course, if $s(u)$ is concave, then the Gaussian ensemble with $\gamma > 0$ necessarily recovers the whole of $s(u)$ just as the canonical ensemble does.) Conversely, the Gaussian ensemble with $\gamma < 0$ must recover fewer points of $s(u)$ than the canonical ensemble because points supported by a supporting line may not be supported by a parabola with inverted curvature. This explains our previous observation that the Gaussian ensemble with $\gamma < 0$ is not useful for obtaining nonconcave entropies.
Figure 6. Betrag ensemble. Upper left: Betrag free energy $\varphi_{B,\gamma}(\beta)$ of the block-spin model for different values of $\gamma$ (see right). Lower left: derivative of $\varphi_{B,\gamma}(\beta)$. Right: Deformed Legendre transform of the Betrag free energy, which recovers the nonconcave entropy $s(u)$ (dashed line) as $\gamma \to \infty$. The entropy is recovered precisely where $\varphi_{B,\gamma}$ is differentiable.

5.2. The Betrag ensemble

We now consider a different ensemble defined by the choice $g(u) = \gamma |u|$ with $\gamma \in \mathbb{R}$, which will be referred to as the Betrag ensemble\(^6\). This ensemble was mentioned in [44], and is somewhat related to piecewise linear Legendre transforms [51], but was never applied before to any equilibrium models.

Example 5.2. The partition function

$$Z^B_N(\beta) = \int_{\Lambda_N} e^{-\beta H_N(\omega) - \gamma |H_N(\omega)|} \, d\omega$$

associated with the Betrag ensemble can easily be calculated for the block-spin model because the energy of this model is positive when $y = 1$ and negative when $y = -1$. The term $|H_N|$ is easily separable, as a result, and we obtain

$$Z^B_N,\gamma(\beta) = Z^{(1)}_N(\beta - \gamma) + Z^{(2)}_N(\beta + \gamma),$$

where $Z^{(1)}_N(\beta)$ and $Z^{(2)}_N(\beta)$ are the two canonical partition functions defined in equation (24). Given the free energies $\varphi_1(\beta)$ and $\varphi_2(\beta)$ shown in equation (25), we therefore obtain

$$\varphi_{B,\gamma}(\beta) = \inf\{\varphi_1(\beta - \gamma), \varphi_2(\beta + \gamma)\} = \begin{cases} \varphi_1(\beta - \gamma) & \beta > 0 \\ \varphi_2(\beta + \gamma) & \beta \leq 0 \end{cases}$$

for the free energy of the Betrag ensemble.

\(^6\) This ensemble could also be called the ‘absolute value ensemble’, but German seems to provide a better name.
This free energy function is shown in figure 6. As for the Gaussian free energy, we see that $\varphi^B(\beta)$ has a nondifferentiable point at $\beta = 0$ for all the values of $\gamma$ considered, but that the image of the derivative of $\varphi^B(\beta)$, which we denote by $u_{\beta,\gamma}$, in the lower left plot of figure 6, fills more and more points of the interval $[-1, 1]$ as $\gamma$ is increased. These properties of $\varphi^B(\beta)$ were also observed for the Gaussian free energy $\varphi^G(\beta)$, and imply that the modified Legendre transform of the Betrag ensemble, given by 7
\begin{equation}
\begin{aligned}
s(u_{\beta,\gamma}) &= \beta u_{\beta,\gamma} + \gamma |u_{\beta,\gamma}| - \varphi^B(\beta),
\end{aligned}
\end{equation}
recovers more and more points of $s(u)$ as $\gamma$ is increased. This is illustrated in the right-hand side plot of figure 6.

As for the Gaussian ensemble, one can show that the Betrag ensemble recovers the full entropy only in the limit $\gamma \to \infty$. The comparison of figures 4 and 6 shows, however, that the Betrag ensemble is more efficient at obtaining the full entropy than the Gaussian ensemble. Indeed, both ensembles recover $s(u)$ over an interval of the form $I_\gamma = [-1, -u_\gamma] \cup [u_\gamma, 1]$, but $u_\gamma$ converges to 0 as $\gamma \to \infty$ faster for the Betrag ensemble than for the Gaussian ensemble; see figure 7. This difference in convergence is related to the way in which the two ensembles achieve equivalence [41]: for the Gaussian ensemble, the limit $\gamma \to \infty$ is required to recover the whole of $s(u)$ because, as already noted, $s(u)$ has a cusp at $u = 0$, whereas for the Betrag ensemble, the limit is needed because $s'(u)$ diverges around its cusp 8.

The possibility of applying the Betrag ensemble to other models rests on being able to separate the partition function of this ensemble into two sums: one involving microstates having a positive energy, and one involving the complementary set of microstates having a negative energy. Such a separation is easily achieved for the block-spin model because of the structure of its Hamiltonian, but one cannot be so optimistic, of course, as to assume that this sort of partitioning trick can be achieved for more realistic models; it all depends on the form of the Hamiltonian $H_N$ that one considers.

In addition to this consideration, it should be clear that, if the nonconcave region of $s(u)$ is located in a region of positive energy, then the Betrag ensemble will not be able

---

7 See the general result shown in equation (30).
8 This implies, in particular, that if $s(u)$ has a cusp with finite left and right derivatives, then the Betrag ensemble achieves equivalence for a finite value $\gamma > 0$, whereas the Gaussian ensemble still requires the limit $\gamma \to \infty$. 

doi:10.1088/1742-5468/2010/05/P05008
to recover or ‘express’ any nonconcave points of \( s(u) \). In this case, one must replace the function \( g(u) = \gamma |u| \) by \( g(u) = \gamma |u - u_0| \), where \( u_0 \) is some fixed value of the mean energy located inside the nonconcave region of \( s(u) \). In practice, this means that in order to use the Betrag ensemble in any useful way, one must have some prior information about where \( s(u) \) is nonconcave in order to choose the right \( u_0 \).

6. The restricted canonical ensemble

The concept of a restricted canonical ensemble or restricted partition function was developed by Penrose and Lebowitz [52] not for calculating nonconcave entropies, but as a way to study metastable states in the canonical ensemble. However, from the discussion of these two topics found in section 4, one should expect that restricted partition functions may also be useful for obtaining nonconcave entropies.

The idea behind restricted partition functions is, as the name suggests, to restrict the sum over all microstates \( \omega \in \Lambda_N \) defining \( Z_N(\beta) \) to a subset of \( \Lambda_N \) which will be denoted by \( R_N \). Thus instead of calculating \( Z(\beta) \), one attempts to calculate

\[
Z^R_N(\beta) = \int_{R_N} e^{-\beta H_N(\omega)} d\omega. \tag{42}
\]

The choice of \( R_N \) is determined by the fact that, when \( s(u) \) is concave, the sum over \( \Lambda_N \) in the standard partition function \( Z(\beta) \) is dominated in the thermodynamic limit \((N \to \infty)\) by a subset of microstates of \( \Lambda_N \) corresponding to the equilibrium states of the canonical ensemble having a fixed energy. However, when \( s(u) \) is nonconcave, there is a further important—yet subdominant—contribution to the sum of \( Z(\beta) \) coming from metastable or unstable states of the canonical ensemble. If one chooses \( R_N = \Lambda_N \), then only the dominant equilibrium states will contribute in the partition function. But if one selects \( R_N \) so as to exclude the dominant states, then \( Z^R_N(\beta) \) will be dominated by the metastable or unstable states. In this case one should be able to recover the nonconcave points of \( s(u) \), or at least some of them, by taking the Legendre transform of the free energy function \( \varphi^R(\beta) \) associated with \( Z^R_N(\beta) \). The next example shows how this works in practice.

Example 6.1. We have seen with the Betrag ensemble that the microstate space \( \Lambda_N \) of the block-spin model can be partitioned, with respect to \( H_N \), into microstates of positive and negative energy. Let us use this partition to define a restricted partition function \( Z^{\pm}_N(\beta) \) by summing only over the microstates having a positive energy:

\[
Z^{\pm}_N(\beta) = \sum_{y=1,\sigma_1,\ldots,\sigma_N} e^{-\beta H_N}. \tag{43}
\]

Going back to the example 4.2, it is easy to see that \( Z^{\pm}_N(\beta) \) is nothing but the ‘metastable’ partition function \( Z^{(2)}_N(\beta) \) defined in equation (24). Therefore,

\[
\varphi^*_+(u) = \varphi^*_2(u) = \frac{1}{2} s_\sigma (2u - 1) \tag{44}
\]

for \( u \in [0, 1] \). A similar result can be derived for \( u \in [-1, 0] \) by calculating a restricted partition function \( Z^{-}_N(\beta) \) for the microstates having a negative energy. In this case,
Methods for calculating nonconcave entropies

\[ Z_N^-(\beta) = Z_N^{(1)}(\beta) \] and so
\[ \varphi^*_- (u) = \varphi^*_1 (u) = \frac{1}{2} s_\sigma (2u + 1) \]
for \( u \in [-1, 0] \). Hence, although neither of the restricted partition functions recovers the whole of \( s(u) \), their combination does.

The difficulty of working with restricted partition functions is similar to that of working with the Betrag ensemble: in both cases, one must be able to calculate a partition function over some restricted set of microstates. Whether this can be done in practice depends on the Hamiltonian \( H_N \) considered and, more precisely, on the possibility of using symmetries of this Hamiltonian to partition \( \Lambda_N \) in easily definable sets of microstates. The block-spin model has such a symmetry, as we have seen, which allows for a straightforward calculation of \( Z_N^+(\beta) \) and \( Z_N^-(\beta) \), as well as \( Z_B^N(\beta) \), the Betrag partition function. In fact, for this model, the functions \( Z_N^+(\beta) \) and \( Z_N^-(\beta) \) merely recreate \( Z_B^N(\beta) \) but in two separate partition functions instead of one.

Of course, if \( H_N \) admits an energy representation function and macrostate entropy function, then restricted partition functions can be calculated for many conceivable restrictions of \( \Lambda_N \) simply by restricting the integrals over the macrostate \( M_N \) that result from the macrostate representation. In this context, the restriction method can be seen as a way to locate the critical points of the function \( G_\beta (m) \), considered in section 4, by restricting the range of values allowed for \( M_N \).

Finally, note that in the extreme case where the sum over states of the standard partition function \( Z_N(\beta) \) is restricted only to microstates of constant energy \( u \), the resulting restricted free energy function is necessarily equal to the entropy, up to some constant. This follows, of course, because the microcanonical ensemble is a special restricted ensemble that considers only microstates with a constant energy.

7. The inverse Laplace transform

The last method that we discuss has been introduced recently in [53]. Its basis is the inverse Laplace transform that expresses the density of states \( \Omega_N(u) \) in terms of the partition function \( Z_N(\beta) \):

\[ \Omega_N(u) = \frac{1}{2\pi i} \int_{r-i\infty}^{r+i\infty} Z_N(\beta) e^{\beta N u} d\beta. \]  

This integral is a complex integral along the path or contour \( \text{Re}(\beta) = r \), often referred to as the Bromwich contour. The value of \( r \) used to position this contour must be chosen in the region of convergence of \( Z_N(\beta) \), but is otherwise arbitrary.

Since the inverse Laplace transform expresses the density of states exactly in terms of the partition function, it can be used, obviously, to obtain \( s(u) \) from \( Z_N(\beta) \) even if the former is nonconcave. In fact, it is known that, since the entropy is a thermodynamic limit function, one needs to know in general only the asymptotic form of \( Z_N(\beta) \) as \( N \to \infty \) to obtain \( s(u) \) via the inverse Laplace transform. One has to be careful, however, to retain the dominant and subdominant terms of \( Z_N(\beta) \) when performing any approximations of the Bromwich integral. If one retains only the dominant term, then only the concave envelope of \( s(u) \) is recovered, in accordance with our discussion of metastable branches of \( \varphi(\beta) \) (see sections 4 and 6). This point is illustrated next.

doi:10.1088/1742-5468/2010/05/P05008 18
Example 7.1. Given \( \varphi(\beta) \), we can approximate the partition function as \( Z_N(\beta) \approx e^{-N\varphi(\beta)} \), plug this approximation into the integral of the inverse Laplace transform of equation (46), and then naively approximate the resulting integral by its saddle point (see, e.g., Appendix C.1 of [26]) to obtain

\[
\Omega_N(u) \approx e^{N \text{inf}_\beta \{ \beta u - \varphi(\beta) \}} = e^{N \varphi^*(u)} \tag{47}
\]

and so \( s(u) = \varphi^*(u) \). This result is correct, as we know, if \( s(u) \) is concave, but not if \( s(u) \) is nonconcave; see section 2 and especially equation (5).

A better approximation for \( Z_N(\beta) \) is suggested by equations (24) and (25):

\[
Z_N(\beta) \approx e^{-N\varphi_1(\beta)} + e^{-N\varphi_2(\beta)}. \tag{48}
\]

Using this expression in equation (46), we obtain

\[
\Omega_N(u) \approx e^{N \varphi_1^*(u)} + e^{N \varphi_2^*(u)} \tag{49}
\]

instead of equation (47), so

\[
s(u) = \text{sup}\{ \varphi_1^*(u), \varphi_2^*(u) \}. \tag{50}
\]

We know from equation (27) that this last formula recovers the correct entropy. Therefore, in the case of the block-spin model, the approximation shown in equation (49) is sufficient for obtaining \( s(u) \).

The previous example can be generalized to any model whose partition function can be put in the form

\[
Z_N(\beta) = \sum_j c_{N,j}(\beta) e^{-N\varphi_j(\beta)}, \tag{51}
\]

where \( \varphi_j(\beta) \) are concave and smooth functions of \( \beta \) that do not depend on \( N \), and \( c_{N,j}(\beta) \) are functions of \( \beta \) that are sub-exponential in \( N \). In [53] it is shown that if these assumptions are satisfied and the coefficients \( c_{j,N}(\beta) \) have no poles in the \( \beta \)-complex plane, then \( s(u) \) is given by a direct generalization of equation (50) involving the ‘metastable’ free energies \( \varphi_j(\beta) \). However, if any of these coefficients have poles in \( \beta \), then \( s(u) \) is given by a more complicated formula involving the \( \varphi_j(\beta) \) as well as the poles of \( c_{N,j}(\beta) \). The surprising effect of these poles is that they determine the presence of linear branches in the graph of \( s(u) \), which arise in short-range systems having first-order phase transitions. For more details on these results, the reader is referred again to [53].

8. Comments and open problems

The examples given in the previous sections are very simple, but provide nevertheless a useful guide as to how the different methods that we have covered in this paper can be applied in practice to obtain the nonconcave entropy of more realistic models. They provide, in particular, a good illustration of the properties that one should look for when selecting the right method to use. On the one hand, if the Hamiltonian considered has any symmetries that can be used to partition the microstate space \( \Lambda_N \) into easily definable regions having different energies, then it may be possible to obtain \( s(u) \) using the Betrag ensemble or the restricted canonical ensemble. On the other hand, if the Hamiltonian admits a macrostate representation, discussed in section 3, then the microcanonical
A Gaussian integral for the Gaussian ensemble: study the integral of equation (32), which expresses the Gaussian partition function in terms of a complex transform of the standard partition function, in order to see whether this integral can be approximated in any useful way. We have already commented on the fact that this integral cannot be solved for the block-spin model, and is unlikely to be computable in general. The reason for this is that the integrand $e^{-\gamma N t^2/2} Z_N(\beta + i\gamma t)$ is highly oscillatory when $\gamma > 0$, which prevents one from performing any form of saddle point approximation. Other types of approximation may be possible, however.

Generalized canonical ensembles and multicanonical simulation methods: there is a strong suggestion that the generalized canonical ensembles discussed in section 5 are related to a set of numerical methods known collectively as multicanonical methods or umbrella sampling methods (see, e.g., [55]–[61]). The exact connection, however, has yet to be made explicit.

Physical interpretation of generalized ensembles: we mentioned in section 5 that the Gaussian ensemble can be interpreted physically as a statistical-mechanical ensemble describing a sample system coupled to a finite-size heat bath (as opposed to the canonical ensemble which describes a sample system coupled to an infinite-size heat bath). Is there a similar physical interpretation for the Betrag ensemble? Are there other generalized ensembles for which a physical interpretation or ‘realization’ can be found or constructed?

Supporting functions for generalized canonical ensembles: we have seen that the concept of supporting lines, which provides a geometrical interpretation of the Legendre transform, is generalized in the Gaussian ensemble to the concept of supporting parabolas. It is not known whether other generalized ensembles admit a similar notion of supporting function. One may ask, for example, whether the supporting function of the Betrag ensemble is the ‘absolute value’ function. More generally, are there supporting functions of other types for other choices of $g(u)$?
Methods for calculating nonconcave entropies

- Moreau transforms: the quadratic Legendre transform of the Gaussian ensemble, defined in equation (36), appears to be related to a functional transform known in convex analysis as the Moreau transform [62]. Are there any known properties of the latter transform that could be used to simplify calculations in the Gaussian ensemble? Moreover, are there generalizations of the Legendre or the Moreau transform that could be used to define generalized canonical ensembles of other types?

Acknowledgments

I would like to thank Oscar Bandtlow for bits of German offered during tea time, Rosemary J. Harris for reading the manuscript, and an anonymous referee for bringing reference [23] to my attention. The hospitality of The Rockefeller University, where most of this paper was written, is also gratefully acknowledged. This work was supported by an RCUK Interdisciplinary Academic Fellowship.

References

[1] Dauxois T, Ruffo S, Arimondo E and Wilkens M (ed), 2002 Dynamics and Thermodynamics of Systems with Long Range Interactions vol 602 (New York: Springer)
[2] Campa A, Giansanti A, Morigi G and Labini F S (ed), 2008 Dynamics and Thermodynamics of Systems with Long-Range Interactions: Theory and Experiments (Melville, NY: AIP)
[3] Campa A, Dauxois T and Ruffo S, 2009 Phys. Rep. 480 57
[4] Dauxois T, Ruffo S and Cugliandolo L F (ed), 2010 Long-Range Interacting Systems ( Lecture Notes of the Les Houches Summer School 2008 vol 90) (Oxford: Oxford University Press)
[5] H"uller A, 1994 Z. Phys. B 95 63
[6] Ispolatov I and Cohen E G D, 2001 Physica A 295 475
[7] Behringer H and Pleimling M, 2006 Phys. Rev. E 74 011108
[8] Binder K and Landau D P, 1984 Phys. Rev. B 30 1477
[9] Challa M S S, Landau D P and Binder K, 1986 Phys. Rev. B 34 1841
[10] Lee J and Kosterlitz J M, 1991 Phys. Rev. B 43 3265
[11] Ruelle D, 1969 Statistical Mechanics: Rigorous Results (Amsterdam: Benjamin)
[12] Landford O E, Entropy and equilibrium states in classical statistical mechanics, 1973 Statistical Mechanics and Mathematical Problems ( Lecture Notes in Physics vol 20) ed A Lenard (Berlin: Springer) p 1
[13] Lynden-Bell D and Lynden-Bell R M, 1977 Mon. Not. R. Astron. Soc. 181 405
[14] Lynden-Bell D, 1999 Physica A 263 293
[15] Gross D H E, 2001 Microcanonical Thermodynamics: Phase Transitions in ‘Small’ Systems ( Lecture Notes in Physics vol 66) (Singapore: World Scientific)
[16] Bouchet F and Barré J, 2005 J. Stat. Phys. 118 1073
[17] Touchette H and Ellis R S, Nonequivalent ensembles and metastability, 2005 Complexity, Metastability and Nonextensivity ed C Beck, G Benedek, A Rapisarda and C Tsallis (Singapore: World Scientific) p 81
[18] Ellis R S, Haven K and Turkington B, 2000 J. Stat. Phys. 101 999
[19] Touchette H, Ellis R S and Turkington B, 2004 Physica A 340 138 [arXiv:cond-mat/0404655]
[20] Mukamel D, Ruffo S and Schreiber N, 2005 Phys. Rev. Lett. 95 240601
[21] Touchette H, Equivalence and nonequivalence of the microcanonical and canonical ensembles: a large deviations study, 2003 PhD Thesis Department of Physics, McGill University
[22] Touchette H, 2008 Am. J. Phys. 76 26 [arXiv:cond-mat/0504020]
[23] Šilhavý M, 1997 The Mechanics and Thermodynamics of Continuous Media (Berlin: Springer)
[24] Rockafellar R T, 1970 Convex Analysis (Princeton, NJ: Princeton University Press)
[25] van Tiel J, 1984 Convex Analysis: An Introductory Text (New York: Wiley)
[26] Touchette H, 2009 Phys. Rep. 478 1
[27] Ellis R S, Touchette H and Turkington B, 2004 Physica A 335 518 [arXiv:cond-mat/0307007]
[28] Barré J, Mukamel D and Ruffo S, 2001 Phys. Rev. Lett. 87 030601
[29] Costemine M, Ellis R S and Touchette H, 2005 J. Math. Phys. 46 063301 [arXiv:cond-mat/0410744]
[30] Barré J, Bouchet F, Dauxois T and Ruffo S, 2005 J. Stat. Phys. 119 677
[31] Campa A, Giansanti A, Mukamel D and Ruffo S, 2006 Physica A 365 120
Methods for calculating nonconcave entropies

[32] Hahn I and Kastner M, 2005 Phys. Rev. E 72 056134
[33] Hahn I and Kastner M, 2006 Eur. Phys. J. B 50 311
[34] Campa A, Ruffo S and Touchette H, 2007 Physica A 385 233
[35] Ellis R S, 1985 Entropy, Large Deviations, and Statistical Mechanics (New York: Springer)
[36] Langer J S, 1967 Ann. Phys., NY 41 108
[37] Langer J S, 1969 Ann. Phys., NY 54 258
[38] Newman C M and Schulman L S, 1977 J. Math. Phys. 18 23
[39] Newman C M and Schulman L S, 1980 J. Stat. Phys. 23 131
[40] Chavanis P H, 2006 Int. J. Mod. Phys. B 20 3113
[41] Costeniuc M, Ellis R, Touchette H and Turkington B, 2005 J. Stat. Phys. 119 1283 [arXiv:cond-mat/040868]
[42] Costeniuc M, Ellis R S, Touchette H and Turkington B, 2006 Phys. Rev. E 73 026105 [arXiv:cond-mat/0505218]
[43] Touchette H, Costeniuc M, Ellis R S and Turkington B, 2006 Physica A 365 132 [arXiv:cond-mat/0509802]
[44] Costeniuc M, Ellis R S and Touchette H, 2006 Phys. Rev. E 74 010105 [arXiv:cond-mat/0605213]
[45] Hetherington J H, 1987 J. Low Temp. Phys. 66 145
[46] Hetherington J H and Stump D R, 1987 Phys. Rev. D 35 1972
[47] Stump D R and Hetherington J H, 1987 Phys. Lett. B 188 359
[48] Challa M S S and Hetherington J H, 1988 Phys. Rev. A 38 6324
[49] Challa M S S and Hetherington J H, 1988 Phys. Rev. Lett. 60 77
[50] Frigori R B, Rizzi L G and Alves N A, 2009 arXiv:0910.0500v1
[51] Ellis R S, 1995 Scand. Actuarial J. 1 97
[52] Penrose O and Lebowitz J L, Towards a rigorous molecular theory of metastability, 1979 Fluctuation Phenomena ed E W Montroll and J L Lebowitz (Amsterdam: North-Holland) chapter 5, pp 293–340
[53] Touchette H, Harris R J and Tailleur J, 2010 Phys. Rev. E 81 030101 [arXiv:0912.3679v1]
[54] Neuhaus T and Hager J S, 2006 Phys. Rev. E 74 036702
[55] Torrie G M and Valeau J P, 1974 Chem. Phys. Lett. 28 578
[56] Berg B A and Neuhaus T, 1991 Phys. Lett. B 267 249
[57] Berg B B and Neuhaus T, 1992 Phys. Rev. Lett. 68 9
[58] Lee J, 1993 Phys. Rev. Lett. 71 211
[59] Janke W, 1998 Physica A 254 164
[60] Berg B A, 2003 Comput. Phys. Commun. 153 397
[61] Lundow P and Markström K, 2009 Cent. Eur. J. Phys. 7 490
[62] Rockafellar R T and Wets R J B, 1988 Variational Analysis vol 317 (New York: Springer)