Equivalence and Nonequivalence of the Microcanonical and Canonical Ensembles: A Large Deviations Study

Hugo Touchette

Department of Physics and School of Computer Science
McGill University, Montréal, Québec, Canada
December 2003

A Thesis submitted to McGill University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy (Ph.D.) in Physics

htouchet@alum.mit.edu
Version of December 16, 2003

THESIS COMMITTEE

This thesis has been evaluated and approved on Tuesday November 25th, 2003 by a jury composed of the following supervisors and examiners:

- Claude Crépeau, Supervisor

  School of Computer Science, McGill University

- Michael C. Mackey, Co-Supervisor

  Department of Physiology and Department of Physics, McGill University

- Martin Grant, Internal Examiner

  Department of Physics, McGill University

- Stefano Ruffo, External Examiner

  Dipartimento de Energetica "Sergio Stecco", Universita de Firenze

Were also present at the Oral Defense:

- Hong Guo

  Department of Physics, McGill University

- Shaun Lovejoy

  Department of Physics, McGill University

- Nilima Nigam

  Department of Mathematics and Statistics, McGill University

SUMMARY

This thesis presents an in-depth study of statistical mechanical systems having *microcanonical* equilibrium properties, i.e., *energy*-dependent equilibrium properties, which cannot be put in correspondence with their *canonical* or *temperature*-dependent equilibrium properties. A general theory of these systems which focuses both on the thermodynamic and macrostate levels of description of systems is presented along the lines of a number of rigorous results derived recently by Ellis, Haven and Turkington (Journal of Statistical Physics, 2000). Several new results are also presented which relate the appearance of nonequivalent microcanonical and canonical properties with first-order (discontinuous) phase transitions and with nonequilibrium properties of systems.

Since the material presented in this thesis dwells on many elements of large deviations theory which are not familiar to physicists, a self-contained introduction to this theory has been included here. The presentation of the theory of nonequivalent microcanonical and canonical properties follows together with explicit computations carried out in the context of two simple spin models: a first original model involving a mixture of completely correlated and completely uncorrelated spins, and another model known as the mean-field Blume-Emery-Griffiths model.

Key words: Thermodynamics, equilibrium statistical mechanics, microcanonical and canonical ensembles, nonequivalence of ensembles, entropy functions, large deviations

RÉSUMÉ

Cette thèse présente une étude détaillée des systèmes statistiques possédant, à l'état d'équilibre, des propriétés physiques calculées en fonction de leur énergie (propriétés dites *microcanoniques*) n'ayant aucun équivalent au sein de leurs propriétés calculées en fonction de leur température (propriétés dites *canoniques*). Une théorie générale de ces systèmes faisant ressortir à la fois la description thermodynamique et la description en macro-états de la mécanique statistique d'équilibre est construite ici suivant un certain nombre de résultats formels obtenus récemment par Ellis, Haven et Turkington (Journal of Statistical Physics, 2000). Plusieurs nouveaux résultats reliant la non-équivalence des propriétés microcanoniques et canoniques aux transitions de phase de premier ordre (transitions discontinues) et aux propriétés hors d'équilibre des systèmes sont aussi présentés.

Comme la théorie présentée ici s'appuie sur un lot de résultats issus de la théorie mathématique des grandes déviations qui sont peu connus des physiciens, une introduction sommaire à cette dernière theorie est d'abord présentée. Suit la théorie portant sur la non-équivalence des propriétés microcanoniques et canoniques, puis plusieurs calculs explicites illustrant certains points de cette théorie dans le contexte de deux modèles physiques de spins bien précis: un premier modèle tout à fait original construit à partir d'un mélange de spins complètement corrélés et complètement non-corrélés, et un deuxième modèle connu sous le nom de modèle en champ moyen de Blume-Emery-Griffiths.

Mots clés: Thermodynamique, mécanique statistique d'équilibre, ensembles microcanonique et canonique, non-équivalence d'ensembles, fonctions d'entropie, grandes déviations

ACKNOWLEDGMENTS

This thesis rounds off three and a half years of research which can acutely be described as a random walk through a university system designed for the most part to encourage "deterministic" studies. The walk, though, has not been unpleasant, for I have been fortunate enough to meet along my way a number of people who have had the patience to allow me to advance at may own pace and choosing, and, in many diverse ways, have made my life easier. Let a few of these persons be thanked here: first, Claude Crépeau, my main supervisor and part initiator of my random walk, for his often expressed confidence in me; Michael C. Mackey, another active supporter of my case and valuable guide to me; and, finally, Richard S. Ellis, for his positive welcoming of a French-Canadian student who contacted him to work on some projects related to his. I would like to further express my gratitude to them as they have played an invaluable act of my life in the roles of motivators, mentors, examiners of this thesis, providers of comments, reference letters, and financial impetus.

Many thanks are also due to Seth Lloyd, André-Marie Tremblay, Shaun Lovejoy, Martin Grant and Bruce Turkington for having provided many pieces of information related to the content of the present thesis. Their guidance in various situations (internships, courses or simple discussions) has been a source of inspiration, as well as a source of interesting physics problems to solve in the years to come.

Hopefully, there is and always will be time for not solving physics problem, free time to gladly interact with the outer world and with friends. For the generous and unconditional offer of such good time in the last four years or so, I would like to thank my dear, long-time friends Jean-François Millette and Valérie Poulin. Special thanks are also due to Paul Dumais, Simon-Pierre Desrosier and Geneviève Arboit, my office mates, partners in computer glitches, and drinking buddies.

Finally, as no one can live without eating, drinking and sleeping, let me thank the granting agencies that funded my work: the Natural Sciences and Engineering Research Council of Canada (two years), the Fonds québecois de la recherche sur la nature et les technologies (one year), as well as the Cryptography and Quantum Information Laboratory of the School of Computer Science at McGill (2/3-year grant).

Évidemment, je ne pourrais clore cette section de remerciements sans ajouter un petit mot en français à l'intention de mes parents, frère et soeurs. Directement ou indirectement, leur existence m'amène à voir une utilité à cette vie qui va bien au-delà de ce que peut m'apporter ma «p'tite job» de physicien. Puissent-ils me poser encore plus de questions quant à l'utilité d'écrire un document de 150 pages que personne (ou presque) ne lira!

Contents

Nomenclature and Notations

RANDOM VARIABLES

Random variables are almost always represented by uppercase letters, while the specific outcomes of these random variables are denoted by lowercase letters. The only exception to this rule concerns the mean energy which is written as u, the total energy being written as U. The space of all events of a given random variable X, or **state space** of X, is denoted by the calligraphic letter $\mathscr{X}$. We write $X = x$, for example, to mean that the random variable X assumes the specific value $x \in \mathscr{X}$.

MEASURE-THEORETIC NOTATIONS

The formal notation $P(A \in da)$ stands as a shorthand for the probability assignment $P(A \in [a, a + da])$ which represents the probability that the random variable A takes on a value in the infinitesimal interval $[a, a + da]$. We write, in a similar manner,

$$P(A \in da) = P(da) = dP(a).$$

Note that although da is used in two different contexts, there can be no confusion: when da is preceded by the symbol "$\in$," da is to be interpreted as an infinitesimal interval of size da, whereas when we write $a + da$, what we mean of course is that da itself is the width of that interval.

SYMBOLS FOR PHYSICAL QUANTITIES AND FUNCTIONS

Following a common practice held in physics, different functions which refer to the same physical quantity are denoted in this thesis by the same letter. For example, the microcanonical entropy is written as $s(u)$ when it is a function of the mean energy u, and as $s(m)$ when it is a function of the macrostate value m. The use of this excessive and ambiguous notation is justified in part by the fact that physicists like to know what quantity they are dealing with even if it is written as a function of different quantities. For the case of the entropy, for example, they view s has being the quantity, and write $s = s(u)$ and $s = s(m)$

instead of $s = f(u)$ and $s = g(m)$ to express the quantity s as functions of the two other quantities u and m.

List of Symbols

$\asymp$	Exponential order approximation sign	1.2
inf	Infimum (generalized min)	
sup	Supremum (generalized max)	
A, X, Y, Z...	Random variables	
$\mathscr{A}, \mathscr{X}, \mathscr{Y}, \mathscr{Z}$...	Event (state) spaces	
a, x, y, z...	Events (values) of random variables	
$X^n = (X_1, X_2, \ldots, X_n)$	Sequence of random variables or microstate	1.1, 2.1
$x^n = (x_1, x_2, \ldots, x_n)$	Specific outcome of X^n	1.1, 2.1
$\mathscr{X} = \{x\}$	one-particle state space	2.1
$\mathscr{X}^n = \{x^n\}$	n-particle state space or microstate space	2.1
$P(\cdot)$	Probability measure	
$p(\cdot)$	Probability density function	
$E[X]$	Expected value of X	1.1
$\mathrm{var}(X)$, σ^2	Variance of X	1.1
$\lvert\mathscr{X}\rvert$	Cardinality or volume of $\mathscr{X}$	
A_n	Mean sum of n random variables	1.1
L_n	Empirical vector	1.1
$I(\cdot)$	Rate function	1.2
$W_n(k)$	Partition function	1.2
$\lambda(k)$	Free energy function	1.2
$D(P_1 \| P_2)$	Relative entropy between P_1 and P_2	1.1
I^*	Dual of I	3.2
I^{**}	Convex hull of I	3.2
$M_n(x^n)$	General macrostate	2.1
m	Macrostate value	2.1
$\mathscr{M} = \{m\}$	Macrostate space	2.1
$U_n(x^n)$	Energy function (Hamiltonian)	2.1
U	Energy value	
$u_n(x^n)$	Mean energy function (energy per particle)	2.1

$u(m)$	Energy representation function	2.2
u	Mean energy value	
T	Temperature	
β	Inverse temperature	2.3
β_c	Critical inverse temperature	3.3
$Z_n(\beta)$	Partition function	2.3
$\varphi(\beta)$	Free energy	2.3
$s(u)$	Microcanonical entropy function	2.2
$s^*(\beta) = \varphi(\beta)$	Dual of $s(u)$	3.3
$s^{**}(u)$	Concave hull of $s(u)$	3.3
$s(m)$	Macrostate entropy	2.2
$c_{\text{can}}(\beta)$	Canonical heat capacity	3.4
$c_{\text{micro}}(u)$	Microcanonical heat capacity	3.4
$P^u(\cdot)$	Microcanonical probability measure	2.2
$P_\beta(\cdot)$	Canonical probability measure	2.3
$I^u(\cdot)$	Microcanonical rate function	2.2
$I_\beta(\cdot)$	Canonical rate function	2.3
u_β	Canonical equilibrium mean energy	2.3
u_l	Lower nonequivalent mean energy	3.3
u_h	Upper nonequivalent mean energy	3.3
m^u	Microcanonical equilibrium value of M_n	2.2
m_β	Canonical equilibrium value of M_n	2.3
$\mathscr{E}^u = \{m^u\}$	Set of microcanonical equilibrium macrostates	2.2
$\mathscr{E}_\beta = \{m_\beta\}$	Set of canonical equilibrium macrostates	2.3

In the books of physics the formulation of the fundamental notions of probability theory as a rule is several decades behind the present scientific level, and the analytic apparatus of the theory of probability, mainly its limit theorems, which could be used to establish rigorously the formulas of statistical mechanics without any complicated special machinery, is completely ignored.

—A.I. Khinchin (1949), p. vii.

Introduction

Temperature is a correlate of energy; a measure of the energetic disorder that takes place at the microscopic level of all substance. This, in short, is what one learns about temperature by taking a course on equilibrium statistical mechanics or by reading textbooks on the subject. By learning how to visualize and model many-body physical systems as stochastic systems, one finally comes to realize what the nature of temperature is, how it is defined mathematically, and how it is measured. Rub your thumb on a desk, and you will immediately feel a rise of temperature on your skin. Why? Because what is measured by temperature is the "wildness" of the motion of the particles located under your thumb: the more wildly the particles move, the greater is their kinetic energy, and the greater, concomitantly, is their temperature. Temperature is thus a mechanical concept: it emerges as a quantity or as a sensation out of the random motion of a body's particles. But then why is temperature felt to be uniform if it emerges out of randomness? Because temperature is also statistical in nature: it is a measure of the energy per particle averaged over the whole of a body. From this viewpoint, the temperature of a single particle cannot be defined because its energy constantly fluctuates. But the temperature of macroscopic bodies can be defined because the *micro*scopic random movements of the particles constituting a large body tend to "average out" one another, so to speak, so as to let order and uniformity emerge at the *macro*scopic level, the level of human experience.

It has been one of the great successes of the theory of equilibrium statistical mechanics to explain what temperature is along the lines given above, and to overthrow, with these answers, the old and now futile belief which hold that heat was a fluid released by hot bodies [the enigmatic "caloric fluid"; see Truesdell (1980)]. Unfortunately, the same "energetic" or "kinetic" interpretation of temperature seems to have played a somewhat negative role in the development of the theory of equilibrium statistical mechanics by casting a shadow over one of its most fundamental problem: namely, to prove that the equilibrium properties of a body (*any* body) are the same whether they are calculated as a function of its internal energy or as a function of its temperature. This problem has evidently not much significance if one views temperature and energy as being one-to-one related. But what one has to realize is that this interpretation of temperature has gained a universal status only because it was verified in the past years to hold true for plenty of physical models, the perfect gas notably. This stands obviously as an insufficient proof of universality; a better, definite proof should proceed by proving in a rigorous manner that the body of equations used to predict the equilibrium properties of a system as a function of its internal energy are totally equivalent to the equations used to predict the same type of equilibrium properties as a function of the temperature of the system. Such a proof, as we shall see in this thesis, can be given for certain systems, but not all of them. In fact, we shall provide here a number of explicit examples of many-body systems which have, at equilibrium, two different sets of equilibrium properties: one set parameterized by the energy variable, and another set parameterized by the temperature variable.

The subject of the equivalence or nonequivalence of the energy and temperature dependent properties of systems is not new. Ever since Gibbs (1902) introduced his canonical probability distribution or "canonical ensemble" as a theoretical device to compute the equilibrium properties of systems as a function of their temperature, the question was posed as to whether this ensemble was equivalent to the other ensemble known to physicists: the so-called "Ergode" or microcanonical ensemble of Boltzmann (1877) which expresses, in probabilistic terms, the conservation of the energy of closed or isolated systems. To Gibbs, the solution of this problem seemed obvious, and nowadays it is his solution that can be found in almost every textbooks on statistical mechanics to clear the problem as solved. Gibbs' reasoning basically is that, although a system having a fixed temperature does not have, theoretically speaking, only one definite value of energy (the canonical distribution is "spread" over many energies), the fluctuations of the system's energy should become negligible in comparison with its total energy in the limit where the volume of the system tends to infinity. In this limit, the so-called thermodynamic limit, the system should thus appear to human observation as having a definite value of energy, which is exactly what the microcanonical ensemble assumes from the start.[1] Conclusion: *both the microcanonical and the canonical ensembles should predict the same equilibrium properties of systems in the thermodynamic limit.*

Gibbs's book is a milestone in the development of equilibrium statistical mechanics, and its publication, not surprisingly, did much to enforce the idea that it does not matter whether the equilibrium properties of a system are calculated from the point of view of the

[1] "For the average square of the anomalies of the energy, we find an expression which vanishes in comparison to the square of the average energy, when the number of degrees of freedom is indefinitely increased. An ensemble of systems in which the number of degrees of freedom is of the same order of magnitude as the number of molecules in the bodies with which we experiment, if distributed canonically, would therefore appear to human observation as an ensemble of systems in which all have the same energy." Gibbs (1902, p. xi)

microcanonical or the canonical ensemble, i.e., whether they are calculated as a function of the energy or the temperature of the system, respectively. Gibbs himself was able to find an explicit expression for the temperature of the perfect gas which shows that it is directly proportional to its internal energy per particle. Moreover, after the publication of Gibbs's book, it was discovered that many noninteracting systems had indeed the same microcanonical and canonical equilibrium properties. From this point on, it seemed then only a small step to take but to think that the equilibrium energy of *any* large-enough system could be related to its temperature. But the problem, unfortunately, is that this is not always the case.

In the past three and a half decades, numerous systems have been discovered to have microcanonical equilibrium properties which cannot be accounted for within the framework of the canonical ensemble. The nonequivalence of the two ensembles has been observed for these systems both at the thermodynamic and the macrostate levels of description of statistical mechanics, and, recently, a nearly complete theory of nonequivalent ensembles has appeared in an effort to shelve once and for all the idea that the energy variable could always be substituted for the temperature variable, and *vice versa,* when parameterizing the equilibrium properties of systems. Our goal in this thesis is to offer a unified presentation of this theory, emphasize its physical interpretation, and supplement it with original results of our own (see the next section for a list of the novel contributions).

A peculiarity of what will be discussed in these pages is worth noticing from the start. Among all the literature that has been published on the problem of nonequivalent microcanonical and canonical ensembles, we have decided in writing this thesis to focus our attention on one specific mathematical theory, due for the most part to Ellis, Haven and Turkington (2000) and to Eyink and Spohn (1993), which offers a "close-to-definitive" solution to this problem. The theory has, in the author's opinion, a definite virtue in that it emphasizes greatly the mathematical roots of the nonequivalence of ensembles problem, in addition to achieving an economy of thought rarely matched by the physics papers treating the same subject. However, we must mention that it has, in parallel, the unfortunate drawback of being not easily approached by physicists since it makes use of many mathematical concepts of the theory of large deviations and convex analysis which are not part of the average physicist's tool kit, so to speak. Great efforts, consequently, have been put into the writing of this thesis to produce a presentation of a complete theory of nonequivalent ensembles, based on the above authors' work, which may be easily understood by physicists. To assist us in the endeavour, we present herein two new simple physical models illustrating this theory, one of which originating directly from the theory of large deviations.

A last word of warning. Because our goal in this thesis is to emphasize the physical ideas, we have omitted in the text many mathematical details that mathematicians would surely consider essential, and have occasionally compromised perfect mathematical accuracy when it benefits the exposition. References treating these details, when they exist, are however always mentioned to assist the exigent reader. At the end, although physicists are likely to complain of seeing too much mathematics, and mathematicians not enough, the author believes that the balance between mathematics and physics reached in writing this work is a fair and healthy one.

CONTRIBUTIONS OF THIS THESIS

The contributions of this thesis can be divided *grosso modo* into two categories: the *pedagogical* contributions, which are the results of the author's desire to write a complete account of a theory which is easy to learn from, and the *scientific* contributions, which are, as the name suggests, the novel mathematical and physical results reported here for the first time. As part of the first category, we find

- Two self-contained introductory chapters (Chapters 1 and 2) which review all the basic elements of the theory of large deviations needed for properly understanding the content of the thesis. The first chapter is specifically devoted to the presentation of the mathematical theory of large deviations, while the second goes on to explain how the concepts of this theory enter in the formulation of most of the equations of equilibrium statistical mechanics.

- An integrated and unified presentation of the theory of the equivalence and nonequivalence of the microcanonical and canonical ensembles which stresses the mathematical roots of the problem (Chapters 3 and 4). Our discussion of this theory in these chapters always begins with the study of a mathematical problem which bears some relationship with the physical problems that we want to discuss. The purpose of these mathematical "preludes" is to introduce new mathematical concepts not previously seen, but, in a way, they are also there to provide a clear and bare light on the equivalence of ensembles problem which is devoid of all the preconceptions that physicists may have about this problem.

- An overview of the literature on the subject of nonequivalent ensembles as well as many notes which compare what has been studied in the past and what has emerged more recently as a complete theory of nonequivalent ensembles. In order not to hinder our presentation of this theory, we have put this material at the end of each chapter in a section called "Notes and Remarks."

- Simpler proofs of all the results pertaining to the equivalence or nonequivalence of the microcanonical and canonical ensembles. The order of presentation of these results also differ from what can be found in the scientific literature. Here, we emphasize the role played by the mean energy for establishing the equivalence or nonequivalence of the two ensembles, and then go on thereafter to prove their equivalence or nonequivalence in the most general way at the level of general macrostates (see Chapters 3 and 4).

In the second category, the category of scientific contributions, we find

- A thorough investigation of the relationship existing between nonequivalent ensembles and first-order (discontinuous) canonical phase transitions. Our discussion greatly unifies many results about this relationship that have appeared over the past years, in addition to supplement these results with others (see Chapters 3 and 4).

- A re-evaluation of the connection that exists between nonequivalent ensembles and the appearance of negative values of the heat capacity calculated from the point of view of the microcanonical ensemble (see Chapter 3).

- An important conjecture which provides a link between the "nonequivalent" microcanonical equilibrium properties of systems not realized in the canonical ensemble, on the one hand, and the nonequilibrium properties of these systems, on the other. (See Chapter 4.)

- Two new physical illustrations of nonequivalent microcanonical and canonical ensembles. Both illustrations involve simple spin systems. One of them is so simple, in fact, that it may well be presented to undergraduate students taking their first course in statistical mechanics (see Chapter 5).

- Finally, an extensive list of open problems, many of them directly suggested by the theory of large deviations (see Outlook and Open Problems).

These contributions, as well as many other results, will be summarized after they are presented in the text in a summary list located at the end of each chapter.

A NOTE TO THE EXAMINERS OF THIS THESIS

About half of what is contained in thesis is part of a paper entitled "*Thermodynamic versus statistical nonequivalence of ensembles for the mean-field Blume-Emery-Griffiths model*" (Los Alamos physics preprints archive: `cond-mat/0307007`) written recently by the author in collaboration with Richard S. Ellis and Bruce Turkington of the Department of Mathematics and Statistics, University of Massachusetts. Apart from being submitted to the Los Alamos preprints archive, this paper has been submitted for publication in the physics journal *Physica A* on June 30$^{\text{th}}$, 2003. It has been accepted for publication in this journal on November 5$^{\text{th}}$, 2003, and is expected to be published in its final form during the spring of 2004. The personal contributions of the author in the writing of this paper are emphasized at the end of the chapters in which these contributions appear. Further credits for the material presented in this thesis which is taken from other sources are also to be found at the end of each chapter.

1

Preliminaries on Large Deviations Theory

This first chapter contains an introduction to the theory of large deviations which emphasizes the elements of this theory that are required to understand the content of this thesis. We begin in the next section to present examples of large deviations which should help the reader to build some intuition and a "pre-knowledge" of what these deviations are. In the subsequent sections, we then spell out the fundamental properties of large deviations using an intuitive yet precise mathematical language familiar to theoretical physicists. The rather informal tone privileged for presenting the subject of large deviations reflects our belief that a complete understanding of all the subtleties of this subject is not necessary for our purposes. It would, in any case, require a whole book to develop the theory of large deviations using all the notations that mathematicians have come to devise only this theory. And that, we think, is best left to them.

1.1. EXAMPLES OF LARGE DEVIATIONS

Fraction of 1's in a Sequence of Random Bits

To begin, we shall consider a sequence $X^n = (X_1, X_2, \ldots, X_n)$ composed of n **binary random variables** taking values in the set $\{0, 1\}$. We suppose that these n random **bits** are mutually independent, and that they are generated from the uniform probability distribution $P(X = 0) = P(X = 1) = 1/2$. What we are interested to find is the probability that the fraction of 1's in the sequence X^n assumes a rational value $r \in \{0, 1/n, 2/n, \ldots, (n-1)/n, 1\}$. For a given **outcome** $x^n = (x_1, x_2, \ldots, x_n)$ of X^n, this ratio is mathematically defined as

$$R_n(x^n) = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{1.1}$$

so that what we are looking for is the following probability:

$$P(R_n = r) = \sum_{x^n : R_n(x^n) = r} P(x^n). \tag{1.2}$$

Here $P(x^n)$ is the probability of the sequence x^n; it is equal to 2^{-n} for all x^n (the random bits are unbiased).

To find the solution of this problem we simply need to count the number of binary n-tuples which are composed of rn 1's and $(1-r)n$ 0's. This is given of course by the binomial coefficient

$$\binom{n}{nr} = \frac{n!}{(rn)![(1-r)n]!}, \tag{1.3}$$

so that

$$P(R_n = r) = \frac{n!}{(rn)![(1-r)n]!} \frac{1}{2^n}. \tag{1.4}$$

This result, like many exact solutions, does not provide much insight as to how $P(R_n = r)$ behaves as a function of r and n. The following approximation, however:

$$\begin{aligned} P(R_n = r) &\simeq e^{-nI(r)} \\ I(r) &= \ln 2 + r \ln r + (1-r)\ln(1-r), \end{aligned} \tag{1.5}$$

which may be obtained using **Stirling's approximation** $n! \simeq n^n e^{-n}$, does tell us much about this behavior. It shows, among other things, that the sequences x^n whose ratio of 0's differs appreciably from the ratio of 1's are very unlikely to be generated in the process of producing bit strings with independent and unbiased bits. Indeed, we can see that $P(R_n = r)$ decays exponentially fast to 0 as $n \to \infty$ for all values of $r \neq 1/2$ (Figure 1.1), which means that only the sequences x^n which are such that $R_n(x^n) \to 1/2$ when $n \to \infty$ have a non-negligible probability to be observed in this limit. Because of this property, the latter sequences are called the **typical** sequences.

Sums of Gaussian Random Variables

We now seek the probability density $p(A_n = a)$ of the following mean sum:

$$A_n(X^n) = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1.6}$$

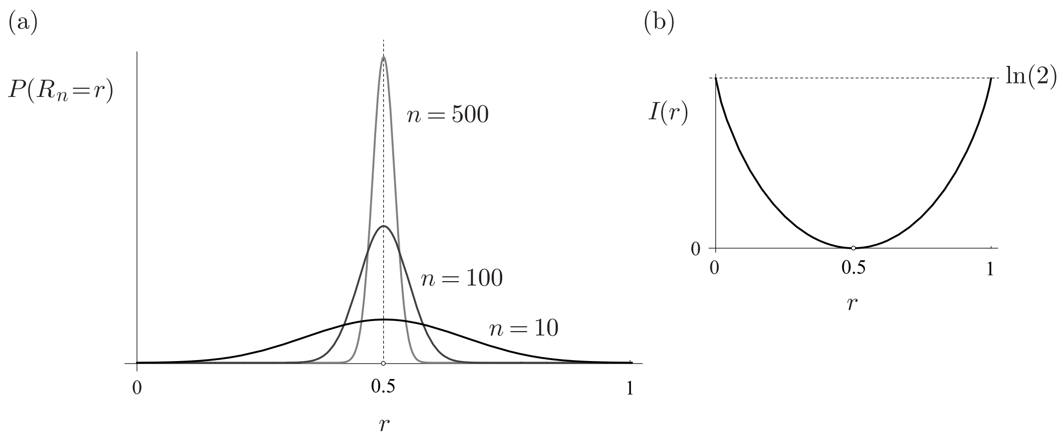

Figure 1.1: (a) Continuous interpolation of $P(R_n = r)$ with $r \in [0,1]$ for the binary string problem. (b) Rate of decay $I(r)$ of $P(R_n = r)$. The minimum of $I(r)$, which locates the maximum of $P(R_n = r)$ for all values of n is attained for $r = 1/2$.

under the assumption that the n random variables $X_1, X_2, \ldots, X_n$ are mutually independent and identically distributed (IID) according to the **Normal** or **Gaussian** probability density

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)/(2\sigma^2)}, \qquad x \in \mathbb{R}. \tag{1.7}$$

In the above function, μ represents the **expectation** of the random variable X, defined as

$$\mu = E[X] = \int_{-\infty}^{+\infty} x p(x) dx, \tag{1.8}$$

whereas σ^2 stands for the **variance**

$$\mathrm{var}(X) = E[(X-\mu)^2] = E[X^2] - E[X]^2. \tag{1.9}$$

Similarly as in (1.2), the density function $p(A_n = a)$ may be written formally as

$$p(A_n = a) = \int_{\{x^n : A_n(x^n) = a\}} p(x^n) dx^n, \tag{1.10}$$

where $p(x^n)$ is the probability density associated with the n outcomes of the random variables $X_1, X_2, \ldots, X_n$. We may write equivalently

$$p(A_n = a) = \int_{\mathbb{R}^n} \delta(A_n(x^n) - a) p(x^n) dx^n, \tag{1.11}$$

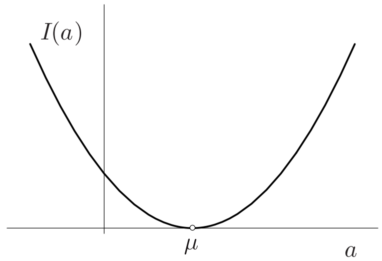

Figure 1.2: Generic parabolic form of the rate function $I(a)$ for the Gaussian sum problem ($a \in \mathbb{R}$). The minimum and zero of $I(a)$ is attained for the mean value $a = \mu$.

using **Dirac's delta distribution** $\delta(x)$. At this point it is common to use the Fourier integral representation of $\delta(x)$

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx} dk, \tag{1.12}$$

and the fact that $p(x^n)$ is the product measure of $p(x)$, to rewrite (1.11) as follows:

$$\begin{aligned} p(A_n = a) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk \int_{\mathbb{R}^n} e^{ik[A_n(x^n)-a]} p(x^n) dx^n \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{-ika} \prod_{j=1}^{n} \int_{-\infty}^{\infty} e^{ikx_j/n} p(x_j) dx_j \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{-ika} \left[\int_{-\infty}^{\infty} e^{ikx/n} p(x) dx\right]^n . \end{aligned} \tag{1.13}$$

Performing the two remaining integrals then leads to

$$p(A_n = a) = \sqrt{\frac{n}{2\pi\sigma^2}} e^{-n(a-\mu)^2/(2\sigma^2)}. \tag{1.14}$$

This results only confirms the well-known fact that a sum of IID Gaussian random variables is exactly Gaussian-distributed no matter how many random variables compose the sum.

Note that if we neglect the $\sqrt{n}$ term in front of the exponential entering in the above density, we obtain a result similar in form to the one found previously, namely,

$$p(A_n = a) \simeq e^{-nI(a)}, \qquad I(a) = \frac{(a-\mu)^2}{2\sigma^2}. \tag{1.15}$$

The functional exponent of decay $I(a)$ is positive and vanishes if $a = \mu$ (Figure 1.2). Thus, as before, any sequence x^n such that $A_n(x^n) \neq \mu$ for all n has an exponentially small probability to appear, whereas the sequences x^n such that $A_n(x^n) \to \mu$ in the limit $n \to \infty$ form the typical sequences which are the most likely to be observed in the long run.

Deviations of the Empirical Vector: Sanov's Theorem

For our last exercise we consider a generalization of the first problem. We are given a sequence $X^n = (X_1, X_2, \ldots, X_n)$ of n IID random variables which are assumed to be **discrete**. This only means that the outcomes of the X_i's are drawn from a finite countable set $\mathscr{X} = \{x\}$; the distribution used to generate the members of $\mathscr{X}$ is denoted by $P(x)$. We denote by $N_n(x)$ the number of times the symbol $x \in \mathscr{X}$ is seen in a given sequence $x^n = (x_1, x_2, \ldots, x_n)$ of outcomes, and we define the **frequencies of the symbols** found in x^n by

$$L_n(x) = \frac{N_n(x)}{n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i,x}. \tag{1.16}$$

The set of values $L_n(x)$ is obviously such that

$$\sum_{x\in\mathscr{X}} L_n(x) = 1. \tag{1.17}$$

What we want to find is $P(L_n = l)$, where L_n stands for the vector of components $L_n(x)$, and l is a particular value of L_n. As an example, let $\mathscr{X} = \{0,1\}$ and $x^n = (0,1,0,1,1,1)$. Then $N_{n=6}(0) = 2$ and $N_{n=6}(1) = 4$, so that

$$L_{n=6}(0) = \frac{2}{6} = \frac{1}{3}, \qquad L_{n=6}(1) = \frac{4}{6} = \frac{2}{3}, \tag{1.18}$$

and thus $L_{n=6} = (1/3, 2/3)$. In probability theory the vector L_n of symbol frequencies is called the **empirical vector**.

Building on our experience of the first problem, it should be clear that the probability distribution $P(L_n = l)$ is given by the multinomial distribution

$$P(L_n = l) = \frac{n!}{\prod_{x\in\mathscr{X}} [nl(x)]!} \prod_{x\in\mathscr{X}} P(x)^{nl(x)}. \tag{1.19}$$

Therefore, putting to use Stirling's approximation once more, we find

$$\begin{aligned} P(L_n = l) &\simeq e^{-nI(l)} \\ I(l) &= D(l||P) = \sum_{x\in\mathscr{X}} l(x) \ln \frac{l(x)}{P(x)}. \end{aligned} \tag{1.20}$$

This exponential estimate of $P(L_n = l)$ is known as **Sanov's Theorem** (Sanov, 1961); the quantity $D(l||P)$ is called the **relative entropy** or **Kullback-Leibler distance** between l and P. Notice that $D(l||P) > 0$ if $l \neq P$, i.e., if $l(x) \neq P(x)$ for at least one value $x \in \mathscr{X}$. Accordingly, if $P \neq l$, then we have that $P(L_n = l)$ goes to 0 exponentially fast as $n \to \infty$. If, on the other hand, we have $l(x) = P(x)$ for all $x \in \mathscr{X}$, then it is easily verified that $D(l||P) = 0$, so that $P(L_n = P) \to 1$ in the limit where $n \to \infty$ [see Cover and Thomas (1991)].

1.2. THE LARGE DEVIATION PRINCIPLE

Our goal in going through the previous examples was to demonstrate the ubiquity of the exponential decaying form e^{-nI} for approximating probability distributions or probability densities of sums of random variables. These approximations and the theory which studies these approximations, namely the theory of large deviations, play a central role in the asymptotic analysis of random sums because they embody basically all there is to know about the large-n behavior of their associated probability distributions or densities. They constitute, in some sense, a zeroth level of approximation, for they suppress any other conceivable polynomial factors in n in the limit where $n \to \infty$, and thus stand as dominant contributions of probability measures whenever they appear. Our aim in this section is to make this reasoning more rigorous by reviewing a few basic definitions and notations relating to large deviation approximations.

Definition of the Large Deviation Principle

In the jargon of large deviations theory, an approximation of the form $P(\text{event}) \simeq e^{-nI}$ is referred to as a **large deviation principle**. To make this a mathematically precise term, the following definition is usually given. Let $P(A_n \in B)$ be the probability that the random variable A_n takes on a value in the event set B. We say that $P(A_n \in B)$ satisfies a large deviation principle with **rate exponent** I_B if the limit

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(A_n \in B) = I_B$$

exists and is independent of n. In other words, $P(A_n \in B)$ satisfies a large deviation principle with rate exponent I_B when the former quantity, plotted as a function of n on a $1/n$-log scale, is seen to converge to a constant as $n \to \infty$. Those who are familiar with the small-o notation may also convince themselves that $P(A_n \in B)$ satisfies a large deviation principle with rate exponent I_B if $P(A_n \in B)$ is approximately exponential in n up to a

$o(n)$ correction term in the exponent or up to a $o(1)$ correction to I_B, i.e.,

$$P(A_n \in B) = e^{nI_B + o(n)} = e^{n[I_B + o(1)]}. \tag{1.21}$$

All the large deviations that we have treated in the previous section fall in this definition in an obvious manner, but they are more specific in a way because they refer to single events of the form $A_n = a$. For the binary string problem, for example, what we found was that

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(A_n = a) = I(a), \tag{1.22}$$

where $I(a)$ is a continuous function of a which, in this context, is called the **rate function** rather than the rate exponent. A similar result was also obtained for the case of the Gaussian mean sum, although, for this case, we have worked with the probability density of A_n not the probability that $A_n = a$ since the latter probability makes evidently no sense for continuous random variables. Our large deviation density result can nevertheless be translated into the language of probabilities: simply exploit the interpretation of probability densities to write

$$p(A_n = a)da = P(A_n \in [a, a + da]) \tag{1.23}$$

and

$$P(A_n \in [a, a + da]) \simeq e^{-nI(a)} da, \tag{1.24}$$

using the large deviation principle of (1.15). In this way, the rate function $I(a)$ is recovered just as for discrete random variables by taking the logarithmic limit in n:

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(A_n \in [a, a + da]) = I(a) - \lim_{n\to\infty} \frac{\ln da}{n} = I(a). \tag{1.25}$$

The Continuum Limit

We shall encounter in this thesis many cases of discrete random variables parameterized by n whose event space may be viewed as "converging" to a continuous space in the limit where $n \to \infty$. A situation of the sort was in fact already encountered when we considered the fraction R_n of 1's in binary strings made out of n random bits. In this example, we noted that R_n takes values in the set of rational values of the form i/n, $i = 0, 1, \ldots, n$. Now, what we want to point out about this example is that, as these rational values populate the real interval $[0, 1]$ more and more densely as $n \to \infty$, R_n can conveniently be thought of as a continuous random variable taking values anywhere in the continuous interval $[0, 1]$.

This way of thinking is obviously just an expedient for simplifying the analysis of R_n: the set of rational numbers $\mathbb{Q}$ cannot be taken to converge to the set $\mathbb{R}$ of real numbers! However, it makes life easier to assume that such a convergence takes place. To make sure that discrete random variables of this kind, which, in effect, "look" continuous as n gets large, are rigorously distinguished from real continuous random variables, which are continuous for all values of n, the former type of random variables will be referred to as being **quasi-continuous**.

In this thesis, we shall often appeal to this discrete-to-continuous limit, or so-called **continuum limit** as it is known in physics, to replace a "densely-defined" discrete probability distribution $P(R_n = r)$ ruling the outcomes of a quasi-continuous random variable R_n, such as the one found in Figure 1.1, by a continuous probability measure of the form $P(R_n \in [r, r+dr])$. The rationale for such a substitution is to be found in the fact that sums involving the probability distribution $P(R_n = r)$ can be approximated, in the continuum limit, by definite integrals involving probability densities, i.e.,

$$E[f(R_n)] = \sum_{r \in \mathscr{R}} f(r)\, P(R_n = r) \overset{n\to\infty}{\simeq} \int_{\mathscr{R}} f(r)\, p(R_n = r) dr, \tag{1.26}$$

where $f(R_n)$ is an arbitrary continuous function of R_n. This naturally suggests the following replacement rule:

$$\text{“}P(R_n = r)\text{”} \longrightarrow \text{“}P(R_n \in [r, r+dr])\text{”} \tag{1.27}$$

as a formal device for taking the continuum limit of R_n.

Equipped with this rule, it is the more reasonable to write a large deviation principle for R_n in the limit of very large n as

$$P(R_n \in [r, r+dr]) \asymp e^{-nI(r)} dr \tag{1.28}$$

instead of just

$$P(R_n = r) \asymp e^{-nI(r)}. \tag{1.29}$$

In this way, we shall be able to handle large deviations of quasi-continuous random variables using the exact same notations as continuous random variables. To spare us from the trouble of constantly writing the interval $[r, r+dr]$, we shall also use the following compact notation:

$$P(R_n \in dr) = P(R_n \in [r, r+dr]) \tag{1.30}$$

(see Nomenclature). Note that although the differential element dr above is not exponential in n, it has to be included in the expression of large deviation principles in order for expectation values to be correctly written as in

$$E[f(R_n)] = \int_{\mathscr{R}} f(r)\, P(R_n \in dr) = \int_{\mathscr{R}} f(r)\, e^{-nI(r)} dr. \tag{1.31}$$

Asymptotic Notation

We shall find it convenient, as a final expedient for cutting in the notations, to re-express the logarithmic limit involved in the definition of the large deviation principle using the formula

$$P(A_n \in da) \asymp e^{-nI(a)} da \tag{1.32}$$

to emphasize the fact that the probability $P(A_n \in da)$ has, to a first degree of approximation, the form of a decaying exponential as a function of n with $I(a)$ as the exponent of decay. The special sign '$\asymp$' is used here instead of the approximation sign '$\simeq$' to stress that, as $n \to \infty$, the *dominant* part of the exact expression of $P(A_n \in da)$ is the exponential function $e^{-nI(a)}$. In this sense, the sign '$\asymp$' may be interpreted as expressing a sort of equality relationship with respect to the logarithmic scale. That is, we may interpret the notation $y_n \asymp z_n$ as meaning that y_n "equals" z_n up to first order in their exponents, so that

$$\lim_{n\to\infty} \frac{1}{n} \ln y_n = \lim_{n\to\infty} \frac{1}{n} \ln z_n \tag{1.33}$$

This interpretation of '$\asymp$' is in perfect agreement with (1.21) as well as with our definition of large deviation approximations in general.

1.3. CALCULATING THE RATE FUNCTION

Many large deviations results can be derived, as was done in the previous sections, by having recourse to asymptotic formulae such as Stirling's approximation. In the general, however, it is more practical to derive the existence of large deviation principles and to calculate their associated rate functions by having recourse to a general-purpose result known as the **Gärtner-Ellis Theorem**. This theorem is, in essence, a generating-functional technique which transforms the problem of calculating $P(A_n \in da)$ into the often more tractable problem of inverting the Laplace transform of this probability measure.

Statement of the Gärtner-Ellis Theorem

Let us consider a random variable A_n, continuous or quasi-continuous, parameterized by n. The exact nature of A_n need not be specified at this point, but to put things on a concrete level, the reader may imagine that A_n is a mean sum of n random variables. Now, let us define the quantity

$$W_n(k) = E[e^{nkA_n}] \tag{1.34}$$

as the **generating function** or **partition function** of A_n and

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln W_n(k) \tag{1.35}$$

as the **cumulant generating function** or **free energy function** of A_n. In terms of the latter quantity, the **Gärtner-Ellis Theorem** states that if $\lambda(k)$ is differentiable over its entire domain of existence, including at the value $k = 0$, then $P(A_n \in da)$ satisfies a large deviation principle with a rate function $I(a)$ given by the Legendre-Fenchel transform of $\lambda(k)$; in symbols,

$$P(A_n \in da) \asymp e^{-nI(a)} da, \tag{1.36}$$

with

$$I(a) = \sup_k \{ka - \lambda(k)\}. \tag{1.37}$$

The supremum transform is what is called a **Legendre-Fenchel** transform.

The theorem of Gärtner and Ellis is very useful in practical calculations, as will be seen in the next section. But it is worth noting that not all large deviation results can be proved using this theorem. In the next chapters we shall be particularly interested to study rate functions which cannot be calculated as the Legendre-Fenchel transform of their associated free energy. The "plausibility proof" of the Gärtner-Ellis Theorem given in the next paragraphs should give some preliminary insights as to how these "anomalous" rate functions are constructed.

Plausibility Proof of the Gärtner-Ellis Theorem

Let us suppose that $P(A_n \in da)$ satisfies a large deviation principle with rate function $I(a)$; that is, suppose that $P(A_n \in da) \asymp e^{-nI(a)} da$. What we aim at is to find the explicit expression of $I(a)$. To arrive at this expression, we proceed to evaluate the partition function

$$W_n(k) = E[e^{nkA_n}] = \int_{\mathscr{A}} e^{nka} P(A_n \in da) \tag{1.38}$$

of A_n by exploiting the exponential order of $P(A_n \in da)$ and by working in the limit of very large values of n. Using the large deviation approximation for $P(A_n \in da)$, $W_n(k)$ can be expressed as

$$W_n(k) \asymp \int_{\mathscr{A}} e^{nka} e^{-nI(a)} da = \int_{\mathscr{A}} e^{n[ka - I(a)]} da. \tag{1.39}$$

At this point we approximate the integral as being given by the largest value of the integrand that lies in the range $\mathscr{A}$ of A_n. This approximation, known as **Laplace's Method** or simply as **Laplace's integral approximation** [see Bender and Orszag (1978)], is a natural approximation to consider here because the error incurred from using it is of the same order as the large deviation approximation itself. Hence, assuming that a largest integrand value exists and is unique, we obtain

$$W_n(k) \asymp \exp\left(n \sup_a \{ka - I(a)\} \right), \tag{1.40}$$

neglecting again any sub-exponential correction factors in the exact expression of $W_n(k)$.

The above equation is interesting for two reasons. First, it shows that if $P(A_n \in da)$ satisfies a large deviation principle with rate function $I(a)$, then $W_n(k)$ must also satisfy a form of large deviation principle, namely

$$W_n(k) \asymp e^{n\lambda(k)}, \tag{1.41}$$

with a "rate function" $\lambda(k)$ corresponding to the Legendre-Fenchel transform of $I(a)$:

$$\lambda(k) = \sup_a \{ka - I(a)\}. \tag{1.42}$$

Second, it shows that if the Legendre-Fenchel transform shown above could somehow be inverted, then we would be in a position to obtain $I(a)$ as a function of $\lambda(k)$. In this case, the obvious question that we have to face is, can this be achieved? namely, can Legendre-Fenchel transforms be inverted? The answer is yes: *If the rate function $I(a)$ is convex, then* *the Legendre-Fenchel transform is inverted by repeating the transform itself, so that*

$$I(a) = \sup_k \{ka - \lambda(k)\}. \tag{1.43}$$

This is as close as we can get to the statement of the Gärtner-Ellis Theorem without actually proving this theorem! To be more complete, we should go on to explain why Legendre-Fenchel transforms are invertible for the class of convex functions, and how the convexity property of $I(a)$ relates to the differentiability property of $\lambda(k)$. This we shall postpone to Chapter 3.

1.4. APPLICATIONS FOR SUMS OF IID RANDOM VARIABLES

To acquire a minimal working knowledge of the Gärtner-Ellis Theorem, and to indulge ourselves in the joy of doing simple calculations, we shall consider in this section a number of different sums of independent and identically distributed (IID) random variables, and calculate the rate functions associated with these sums. The case of IID random variables is easy to handle because $\lambda(k)$ reduces to

$$\begin{aligned} \lambda(k) &= \lim_{n\to\infty} \frac{1}{n} \ln E\left[\exp\left(k\sum_{i=1}^{n} X_i\right)\right] \\ &= \lim_{n\to\infty} \frac{1}{n} \ln \prod_{i=1}^{n} E[e^{kX_i}] \end{aligned} \tag{1.44}$$

as a result of the independency property, and to

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln E[e^{kX}]^n = \ln E[e^{kX}] \tag{1.45}$$

by taking into account the fact that the random variables $X_1, X_2, \ldots, X_n$ are identically distributed. Thus, all we have to do in order to find $I(a)$ is to evaluate the simple expectation value $E[e^{kX}]$ which involves no limits in n, take the logarithm of the result, and calculate finally its Legendre-Fenchel transform.

Gaussian Random Variables Revisited

Consider again the mean sum A_n of Gaussian IID random variables studied in Section 1.1. The cumulant generating function or free energy function associated with the Gaussian density of equation (1.7) is readily evaluated to be

$$\lambda(k) = \ln \int_{-\infty}^{\infty} p(x) e^{kx} dx = \mu k + \frac{1}{2}\sigma^2 k^2, \tag{1.46}$$

with $k \in \mathbb{R}$. Since $\lambda(k)$ is everywhere differentiable, the rate function $I(a)$ governing the probability density of A_n can be calculated as the Legendre-Fenchel transform of $\lambda(k)$:

$$I(a) = \sup_{k\in\mathbb{R}}\{ka - \lambda(k)\}. \tag{1.47}$$

To find the supremum value, we look for the value of k which achieves the supremum by taking the derivative with respect to k of the right-hand side of the above equation and make the result equal to 0. This yields

$$\lambda'(k) = \mu + \sigma^2 k = a, \tag{1.48}$$

so that

$$k(a) = \frac{a-\mu}{\sigma^2}. \tag{1.49}$$

The rate function is thus given by

$$I(a) = k(a)a - \lambda(k(a)) = \frac{(a-\mu)^2}{2\sigma^2}. \tag{1.50}$$

where $a \in \mathbb{R}$. This result matches exactly, as expected, the exponential estimate that we obtained before using a far more complicated method.

Binary ± 1 Random Variables

Suppose now that the mean sum A_n is composed of n IID **binary** random variables taking values in the set $\mathscr{X} = \{-1, +1\}$ according to the uniform probability distribution

$$P(X = -1) = P(X = +1) = \frac{1}{2}. \tag{1.51}$$

The free energy function associated with this probability distribution is

$$\lambda(k) = \ln\cosh k, \qquad k \in \mathbb{R}. \tag{1.52}$$

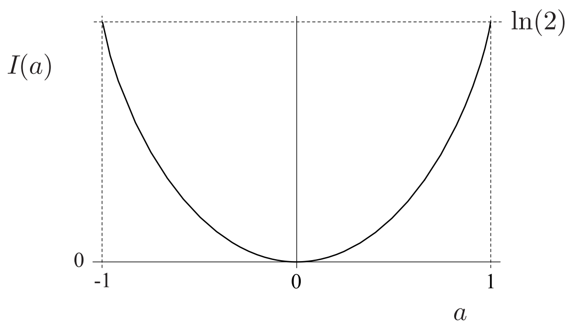

Figure 1.3: Rate function $I(a)$ for a mean sum of IID discrete random variables taking values in $\{-1,+1\}$.

It can easily be checked again that $\lambda(k)$ satisfies all the requirements of the Gärtner-Ellis Theorem which implies that $I(a)$ can be obtained exactly as before by solving the differential equation $\lambda'(k) = a$ for k and inserting the result in

$$I(a) = k(a)a - \lambda(k(a)). \tag{1.53}$$

What is found from these steps is

$$I(a) = a \operatorname{arctanh} a - \ln \cosh \operatorname{arctanh} a, \tag{1.54}$$

where $a \in [-1, 1]$. Using the two identities

$$\begin{aligned} \cosh a &= \frac{1}{\sqrt{1-\tanh^2 a}} \\ \operatorname{arctanh} a &= \frac{1}{2} \ln\left(\frac{1+a}{1-a}\right), \end{aligned} \tag{1.55}$$

we can also rewrite $I(a)$ in a more illuminating fashion as

$$I(a) = \frac{1+a}{2} \ln(1+a) + \frac{(1-a)}{2} \ln(1-a), \qquad a \in [-1, 1]. \tag{1.56}$$

The graph of this rate function is shown in Figure 1.3. Its shape is very similar to the shape of the rate function found previously in the context of the bit string problem (Figure 1.1), but the minimum and zero of $I(a)$ is now located at $a = 0$.

Sanov's Theorem Revisited

Although we stated the theorem of Gärtner and Ellis for the case of scalar random variables, its result may be applied to other kinds of random variables, even random vectors! In the case of vectors, the only modification that has to be made at the level of the free energy is to replace the ordinary product kA_n by the scalar product $k \cdot A_n$, k and A_n being now two vectors of the same dimension.

To illustrate this generalization, consider again the problem of determining the probability distribution $P(L_n = l)$ associated with the empirical vector L_n for sequences of IID random variables (see Section 1.1 for the definitions of these quantities). Given the definition of the empirical vector displayed in (1.16), the free energy is calculated as follows:

$$\begin{aligned}\lambda(k) &= \lim_{n\to\infty} \frac{1}{n} \ln E[e^{nk\cdot L_n}] \\ &= \lim_{n\to\infty} \frac{1}{n} \ln E\left[\exp\left(\sum_{i=1}^{n} \sum_{x\in\mathscr{X}} k(x)\delta_{X_i,x}\right)\right] \\ &= \lim_{n\to\infty} \frac{1}{n} \ln E\left[\exp\left(\sum_{i=1}^{n} k(X_i)\right)\right]. \end{aligned} \tag{1.57}$$

But since the random variables $X_1, X_2, \ldots, X_n$ are IID, we may forget about the limit in n so that

$$\begin{aligned}\lambda(k) &= \lim_{n\to\infty} \frac{1}{n} \ln E[e^{k(X)}]^n \\ &= \ln E[e^{k(X)}] \\ &= \ln\left(\sum_{x\in\mathscr{X}} P(x)e^{k(x)}\right). \end{aligned} \tag{1.58}$$

At this point, we find the rate function $I(l)$ associated with the large deviation principle

$$P(L_n \in dl) \asymp e^{-nI(l)}dl \tag{1.59}$$

similarly as in equation (1.48) by solving the set of equations

$$\frac{\partial}{\partial k(x)}\lambda(k) = l(x), \qquad x \in \mathscr{X} \tag{1.60}$$

for a given vector k. Despite the monstrous look of this vectorial differential equation involving a gradient with respect to k, its solution is easily found to satisfy

$$k_l(x) = \ln \frac{l(x)}{P(x)} + \lambda(k_l). \tag{1.61}$$

This last equation is an implicit equation satisfied by k_l; it can directly be substituted in

$$I(l) = k_l \cdot l - \lambda(k_l) \tag{1.62}$$

to obtain finally

$$I(l) = D(l||P) = \sum_{x\in\mathscr{X}} l(x) \ln \frac{l(x)}{P(x)}. \tag{1.63}$$

One may be interested to know that the above result which, we recall, goes by the same of Sanov's Theorem, holds for any type of event sets $\mathscr{X}$, even continuous ones. For this latter type of sets, the empirical vector $L_n(x)$ represents an empirical density of occurrence of the symbols $x \in \mathscr{X}$, and the derivatives involved in (1.60) must be interpreted as functional derivatives.

1.5. PROPERTIES OF $\lambda(k)$ AND $I(a)$

We now state and prove a number of properties satisfied by the free energy function $\lambda(k)$ and the rate function $I(a)$. The properties listed hold for *any* random variable A_n under the conditions stated, not just mean sums of IID random variables.

Properties of $\lambda(k)$ at $k = 0$

Since probability distributions are by definition normalized to 1, we must have

$$\lambda(k=0) = \lim_{n\to\infty} \frac{1}{n} \ln E[e^0] = 0. \tag{1.64}$$

Moreover,

$$\begin{aligned} \lambda'(k=0) &= \lim_{n\to\infty} \frac{1}{E[e^{knA_n}]} E[A_n e^{knA_n}]\bigg|_{k=0} \\ &= \lim_{n\to\infty} E[A_n], \end{aligned} \tag{1.65}$$

provided that $\lambda(k)$ exists. For the case of IID sums, this result reduces to

$$\lambda'(k=0) = E[X] = \mu. \tag{1.66}$$

In the same vein,

$$\begin{aligned}\lambda''(k=0) &= \lim_{n\to\infty} \frac{E[nA_n^2 e^{knA_n}]}{E[e^{knA_n}]} - \frac{E[nA_n e^{knA_n}]E[A_n e^{knA_n}]}{E[e^{knA_n}]^2}\Bigg|_{k=0} \\ &= \lim_{n\to\infty} nE[A_n^2] - nE[A_n]^2 \\ &= \lim_{n\to\infty} n\,\mathrm{var}(A_n). \end{aligned} \tag{1.67}$$

For sums of IID random variables, we thus have

$$\lambda''(k=0) = E[X^2] - E[X]^2 = \mathrm{var}(X) = \sigma^2. \tag{1.68}$$

Finally, from Jensen's inequality (see Cover and Thomas, 1991), we find

$$\begin{aligned}\lambda(k) &= \lim_{n\to\infty} \frac{1}{n} \ln E[e^{knA_n}] \\ &\geq \lim_{n\to\infty} \frac{1}{n} E[\ln e^{knA_n}] \\ &= \lim_{n\to\infty} kE[A_n] \end{aligned} \tag{1.69}$$

i.e.,

$$\lambda(k) \geq kE[X] = k\mu \tag{1.70}$$

for IID random variables, with equality in both results if and only if $k = 0$. This last property as well as the two first are illustrated in Figure 1.4 for the case of IID sums.

Convexity of $\lambda(k)$

The function $\lambda(k)$, if it exists, is always convex in k. This comes as a general consequence of Hölder's inequality

$$\sum_i |y_i z_i| \leq \left(\sum_i |y_i|^{1/p}\right)^p \left(\sum_i |z_i|^{1/q}\right)^q, \tag{1.71}$$

where $0 \leq p, q \leq 1$, $p + q = 1$ [see Zwillinger (1996) and Dembo and Zeitouni (1998)]. Indeed, using this inequality, we can write

$$\begin{aligned}\alpha \ln E[e^{nk_1A_n}] + (1-\alpha)\ln E[e^{nk_2A_n}] &= \ln E[e^{nk_1A_n}]^\alpha E[e^{nk_2A_n}]^{(1-\alpha)} \\ &\geq \ln E[(e^{nk_1A_n})^\alpha (e^{nk_2A_n})^{(1-\alpha)}] \\ &= \ln E[e^{n[\alpha k_1 + (1-\alpha)k_2]A_n}]. \end{aligned} \tag{1.72}$$

Whence

$$\lambda(\alpha k_1 + (1-\alpha)k_2) \leq \alpha\lambda(k_1) + (1-\alpha)\lambda(k_2). \tag{1.73}$$

Legendre Transform Equations

We have seen in treating all the previous examples of mean sums that the Legendre-Fenchel transform involved in the statement of the Gärtner-Ellis Theorem actually reduces to a derivative-transform given by

$$I(a) = k(a)a - \lambda(k(a)), \tag{1.74}$$

where $k(a)$ represents the unique root of the differential equation $\lambda'(k) = a$. Such a transform is known as a **Legendre transform**; it arises here because $\lambda(k)$ is everywhere differentiable, as required by the Gärtner-Ellis Theorem itself, and because $\lambda(k)$ is always convex, as mentioned above.

The Legendre transform enjoys many interesting properties. First, since $\lambda(k)$ is convex, $\lambda'(k)$ has to be monotonically increasing, and so the function $k(a)$ can be uniquely inverted so as to obtain the function $a(k)$. The latter function obviously satisfies the equation

$$\lambda'(k) = a(k) \tag{1.75}$$

The form of the Legendre transform also leads to

$$I'(a(k)) = k \qquad \text{or} \qquad I'(a) = k(a). \tag{1.76}$$

Indeed, differentiating expression (1.74) with respect to a yields

$$\begin{aligned} \frac{dI(a)}{da} &= \frac{dk(a)}{da}a + k(a) - \frac{d\lambda(k)}{dk}\frac{dk(a)}{da} \\ &= \frac{dk(a)}{da}a + k(a) - a\frac{dk(a)}{da} \\ &= k(a). \end{aligned} \tag{1.77}$$

The graphical interpretation of all these equations relating the slope values of $\lambda(k)$ and $I(a)$ to $a(k)$ and $k(a)$, respectively, is shown in Figure 1.4.

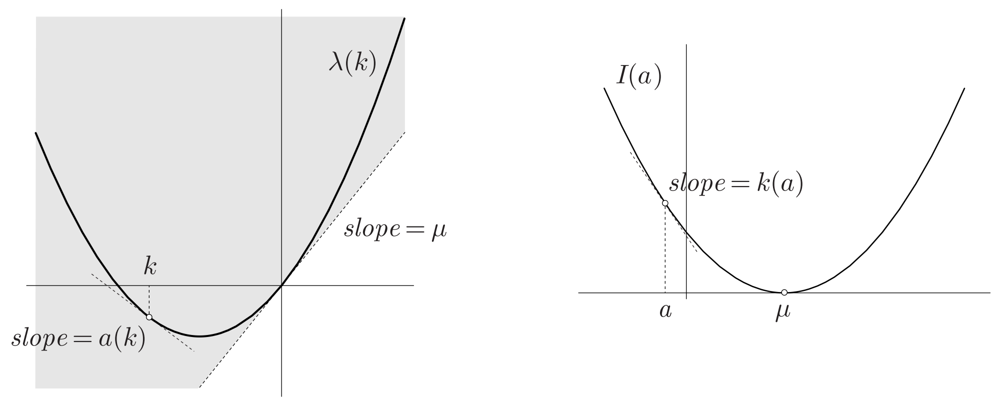

Figure 1.4: Generic properties of the free energy function $\lambda(k)$ (left) and rate function $I(a)$ (right).

Concavity of $I(a)$

Rate functions obtained from the Gärtner-Ellis Theorem are necessarily strictly convex. This comes as a consequence of the fact that Legendre transforms preserve the convexity or concavity properties of the functions on which they are applied. To be sure, differentiate another time the result of (1.77) so as to obtain the second derivative of $I(a)$:

$$\frac{d^2 I(a)}{da^2} = \frac{dk(a)}{da} = \left(\frac{da(k)}{dk}\right)^{-1} = \left(\frac{d^2\lambda(k)}{dk^2}\right)^{-1}. \tag{1.78}$$

This demonstrates that the curvature of $I(a)$ is the inverse curvature of $\lambda(k)$; for the case of IID sums, in particular,

$$I''(a=\mu) = [\lambda''(k=0)]^{-1} = \frac{1}{\sigma^2}. \tag{1.79}$$

Accordingly, since $\lambda(k)$ is always convex (i.e., $\lambda''(k) \geq 0$), $I(a)$ must always be convex as well (i.e., $I''(a) \geq 0$).

Zero of $I(a)$: The Law of Large Numbers

In the case where $I(a)$ is strictly convex, $I(a)$ must have a single minimum and zero located at the value $a = E[A_\infty] = a_{\text{typ}}$ (most probable or **typical** value). To prove this property, we use two of the properties proved above, namely $\lambda'(k=0) = a_{\text{typ}}$ and $\lambda(k=0) = 0$, to write

$$I(a_{\text{typ}}) = k(a_{\text{typ}})a_{\text{typ}} - \lambda(k(a_{\text{typ}})) = 0 \cdot a_{\text{typ}} - 0 = 0. \tag{1.80}$$

The existence of this unique zero of the rate function is guaranteed by the **Law of Large Numbers** which may loosely be stated as follows:

$$P(A_n \in [a_{\text{typ}}, a_{\text{typ}} + da]) \asymp e^{-nI(a_{\text{typ}})}da = 1. \tag{1.81}$$

In other terms

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} E[A_n] = a_{\text{typ}} \tag{1.82}$$

in probability. This shows that running averages converge to expectation values almost surely (see Figure 1.1).

It must be kept in mind that large deviations approximations are exponential-order approximations of probability measures; hence the fact that we directly get the value 1 for the probability of the event $A_n = a_{\text{typ}}$. Notice also that $I(a)$ may in general have more than one global minimum, and may have local minima in addition to global ones. The positions of the global minima give the most probable values of A_n just as in the case of a single minimum.

Positivity of $I(a)$

Rate functions are always such that $I(a) \geq 0$; see den Hollander (2000) or Dembo and Zeitouni (1998) for a proof of this result. A negative rate function would imply that large deviation probabilities diverge to infinity as $n \to \infty$.

Convex Minimum of $I(a)$: The Central Limit Theorem

The appearance of a single global minimum of strictly convex rate functions can be related to the **Central Limit Theorem**. Indeed, for "small" deviations of A_n close to the expected value $E[A_\infty]$, the rate function $I(a)$ can be expanded in a Taylor series to second order to obtain

$$I(a) \simeq \frac{1}{2} I''(a_{\text{typ}})(a - a_{\text{typ}})^2. \tag{1.83}$$

This quadratic approximation of $I(a)$ implies a Gaussian approximation for $P(A_n \in da)$ around $a = a_{\text{typ}}$:

$$P(A_n \in da) \asymp e^{-nc(a - a_{\text{typ}})^2} da, \tag{1.84}$$

with $c = I''(a_{\text{typ}})/2$. (The constant c is simply equal to σ^2 for IID sums.)

The Gaussian approximation can be shown to be valid so long as $A_n = O(1/\sqrt{n})$ or, equivalently, $nA_n = O(\sqrt{n})$ [see Dembo and Zeitouni (1998)]. This may begin to put some light on the meaning of the term "large deviations." On the one hand, a **small deviations** of the mean sum A_n is a value a for which the quadratic approximation to $I(a)$ can be considered to be a "good" approximation. A **large deviation**, on the other hand, is a mean sum value a for which the true rate function $I(a)$ departs in a non-negligeable way from the Gaussian quadratic approximation, and for which, therefore, the Central Limit Theorem ceases to be useful.

1.6. MORE CONSEQUENCES OF LAPLACE'S METHOD

Concentration Property

Imagine that we are interested to know the expected value of a given random variable A_n. If $P(A_n \in da)$ satisfies a large deviation function with convex rate function $I(a)$, then $E[A_n]$ can easily be calculated: simply apply the rule of Laplace's Method to obtain

$$\begin{aligned} E[A_n] &= \int_{\mathscr{A}} aP(A_n \in da) \\ &\asymp \int_{\mathscr{A}} ae^{-nI(a)} da \\ &\asymp \arg \inf_{a \in \mathscr{A}} I(a), \end{aligned} \tag{1.85}$$

provided that $I(a)$ has a unique global minimizer. If this is indeed the case, then what these equations say is that the outcomes of A_n are **exponentially concentrated** (in terms of their probabilities) around the most probable or **dominant** value of A_n. This, in a way, is just a re-statement of the Law of Large Numbers which states, informally again, that

$$P(A_n \in da) \to \delta(a - a_{\text{typ}}) da, \tag{1.86}$$

as $n \to \infty$, a_{typ} being the dominant or typical value of A_n (see Figure 1.1).

The Contraction Principle

Suppose now that we have worked out the rate function of a given random variable A_n, and that we want to derive the rate function of another random variable B_n which is a **contraction** of A_n, that is, which is such that $B_n = h(A_n)$ for some continuous mapping $h : \mathscr{A} \to \mathscr{B}$. Is there a way to compute $I_B(b)$ from $I_A(a)$? The answer is yes: simply apply

Laplace's Method in the course of calculating $P(B_n \in db)$ from $P(A_n \in da)$ to obtain

$$\begin{aligned} P(B_n \in db) &= \int_{\{a:h(a)=b\}} P(A_n \in da) \\ &\asymp \int_{\{a:h(a)=b\}} e^{-nI_A(a)} da \\ &\asymp \exp\left(-n \inf_{a:h(a)=b} I_A(a)\right) da \end{aligned} \tag{1.87}$$

and

$$P(B_n \in db) \asymp \exp\left(-n \inf_{a:h(a)=b} I_A(a)\right) \frac{db}{|h'(a)|}, \tag{1.88}$$

where the term $|h'(a)|$ arises as the Jacobian of the transformation $h(a) = b$. This proves that

$$P(B_n \in db) \asymp e^{-nI_B(b)} db \tag{1.89}$$

with

$$I_B(b) = \inf_{a:h(a)=b} I_A(a). \tag{1.90}$$

This general reduction of one rate function to another is called a **contraction principle**. It can be used, among many applications, to derive the rate functions of mean sums of IID random variables from Sanov's Theorem. For this precise case, the function $h(L_n)$ realizing the contraction between the empirical vector L_n and the scalar values a of the mean sum A_n is given by

$$h(L_n) = \sum_{x \in \mathscr{X}} x L_n(x) = \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{1.91}$$

SUMMARY OF CHAPTER

- **Large deviation principle:** The probability measure

  $$P(A_n \in da) = P(A_n \in [a, a + da]) \tag{1.92}$$

  is said to obey a large deviation principle if the limit

  $$\lim_{n\to\infty} -\frac{1}{n} \ln P(A_n \in da) = I(a) \tag{1.93}$$

exists, and is independent of n for all members a of the event space $\mathscr{A}$ of the random variable A_n. The formal notation

$$P(A_n \in da) \asymp e^{-nI(a)} da \tag{1.94}$$

is used as a shorthand notation for the logarithmic limit shown above.

- **Generating function or partition function:** $W_n(k) = E[e^{knA_n}]$.

- **Cumulant generating function or free energy function:**

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln W_n(k). \tag{1.95}$$

- **Legendre-Fenchel transform of I:** If $P(A_n \in da)$ satisfies a large deviation principle with rate function $I(a)$, then

$$\lambda(k) = \sup_a \{ka - I(a)\}. \tag{1.96}$$

The result holds independently of the form of $I(a)$, and yields an always convex function of k. If A_n is a vectorial random variable, then the product ka in the above formula should be understood as the scalar product $k \cdot a$.

- **Gärtner-Ellis Theorem:** If $\lambda(k)$ is differentiable everywhere over its range of convergence, including in a neighborhood of $k = 0$, then $P(A_n \in a) \asymp e^{-nI(a)} da$ with

$$I(a) = \sup_k \{ka - \lambda(k)\}. \tag{1.97}$$

The rate function thus calculated is always strictly convex.

- **Contraction principle:** Let A_n and B_n be two random variables having rate functions $I_A(a)$ and $I_B(b)$, respectively, and let $h : \mathscr{A} \to \mathscr{B}$ be a contraction of B_n to A_n, i.e., a continuous function mapping the event space $\mathscr{A}$ of A_n onto the event space $\mathscr{B}$ of B_n. Then,

$$I_B(b) = \inf_{a:h(a)=b} I_A(a). \tag{1.98}$$

NOTES AND REMARKS

References on Large Deviations Theory

Introductions to the theory of large deviations similar to the one given in this chapter can be found in the review papers of Oono (1989), Amann and Atmanspacher (1999), Ellis (1995, 1999) and Lewis and Russell (1996). The reader who wishes, on the other hand, to learn more about the mathematical subtleties of this theory is invited to consult the books of Dembo and Zeitouni (1998), Ellis (1985) and den Hollander (2000). However, let he or she be warned that much mathematical maturity is required to follow the content of these books: mathematicians take a great care of proving the existence of limits and in developing high-level notations to rigorously handle complicated limits such as those involved in what we called the continuum limit.

Our approach here has been to purposely avoid such notations, and much simplification is achieved by doing so. The fact is that physicists rarely care about rigor when they play with the continuum limit because, for them, it is a natural limit. They just think of it as they see the world that surrounds them: this apparently continuous world which is composed, at the very bottom, of discrete objects called particles.

The Genesis of Large Deviations in Brief

Much of what is referred to nowadays as the theory of large deviations emerged during the 1970s from the independent works of Donsker and Varadhan, and Freidlin and Wentzel (see Dembo and Zeitouni (1998) for an introduction to the original literature). However, the real roots of the subject seems to be much older. The Gärtner-Ellis Theorem, in particular, which is the result of the independent efforts of Gärtner (1977) and Ellis (1984), was already known to Cramér (1938) who proved it for the particular case of IID sums. Daniels (1954) obtained the same result by independent means apparently, and termed it the saddlepoint approximation.

From the point of view of a physicist, the theorem of Gärtner and Ellis is also very reminiscent of the techniques developed by Boltzmann (1877) and Gibbs (1902), the two physicists who laid down the foundations of equilibrium statistical mechanics. Boltzmann, in fact, seems to have anticipated Sanov's Theorem [Sanov (1961)], around 1877, as he was working on the equilibrium distribution of states of the perfect gas [see Ellis (1999)].

These historical facts are given here not in an attempt to revise the history of large deviations theory by excluding or marginalizing the role of mathematicians in the development of this theory. The fact is that neither Boltzmann nor Gibbs conceived their newly developed techniques as being part of a systematic mathematical theory which could encompass not only the probabilistic description of many-body systems, but any stochastic processes at large. If only for this reason, the two physicists cannot be considered as the true "founding fathers" of large deviations theory. However, the strong similarity between the techniques and results of large deviations theory and those of statistical mechanics cannot be thought of only as a coincidence, as will become clear in the next chapter. Boltzmann and Gibbs anticipated something, but we must thank the mathematicians for their work at formalizing completely and independently this "something."

2

Statistical Mechanics as a Large Deviations Theory

The previous chapter was a prelude aimed at introducing the mathematical tools that will be put to use from this point on to study the physical behavior of systems which are composed of many microscopic "particles" (molecules, spins, etc.). The probabilistic study of these **many-body systems** which appear to behave more or less randomly when probed at the *micro*scopic level, although they seem totally regular at the *macro*scopic level, has been initiated by physicists like Boltzmann, Gibbs and Maxwell more than a century ago. From their works has emerged a theory now referred to as **statistical physics** or **statistical mechanics**. We shall see in this chapter how such mathematical concepts as a large deviation principle and a rate function intervene in the very foundations of this theory, and especially in the construction of the so-called microcanonical and canonical ensembles, the two main "theoretical devices" of statistical mechanics with which the equilibrium properties of many-body systems are commonly calculated. An explicit calculation of these properties for a simple spin system is given in the last section to illustrate the usefulness of these concepts. By the end of this chapter, we hold that the reader should be convinced that statistical mechanics is, in retrospect, just an application of large deviations theory (see Table 2.1).

2.1. MODELING MANY-BODY SYSTEMS WITH RANDOM VARIABLES

Fundamental Terms and Definitions

To apply large deviations concepts or any other probabilistic concepts for studying physical many-body systems, we need first to state a number of postulates and definitions whose purpose is (i) to define what the term "many-body systems" means; (ii) establish an interpretation of the physical variables describing the states of such systems in terms of probabilistic concepts; and (iii) make precise the conditions (mathematical or physical) under which we expect a probabilistic analysis of these systems to be meaningful. In order not to unduly lengthen the list of such postulates and definitions, we take for granted that the reader has already some knowledge of the basic principles of statistical mechanics and an idea of the kind of systems for which statistical mechanics can be applied to [see, e.g., Reif (1965), Landau and Lifshitz (1991) or Balian (1991)]. Taking this knowledge as a *fait accompli*, the following few terms then need only be added:

- A many-body system is an assembly of n identical **subsystems** (e.g., particles, spins, proteins, etc.) which may be physically interacting or not.

- The **joint** or **collective state** of the n subsystems is represented abstractly by a sequence of n random variables $X^n = (X_1, X_2, \ldots, X_n)$, where X_i stands for the state of the i^{th} subsystem.

- In the jargon of statistical physics, a specific outcome x^n of X^n, which completely describes the state of a n-body system at the "microscopic" level, is called a **microstate**. The state space $\mathscr{X}^n = \{x^n\}$ of all microstates is the n-fold product space of the one-particle space $\mathscr{X} = \{x\}$.

- Since the state of a many-body system is modeled as a random variable, an *a priori* **joint probability measure** $P(dx^n)$ has to be specified on the set of all joint outcomes $x^n = (x_1, x_2, \ldots, x_n)$. The form of P is not arbitrary; in fact, it will be seen in the next section that P has to be chosen, for physical reasons, to be the **uniform** measure
$$P(dx^n) = \frac{dx^n}{|\mathscr{X}^n|}, \tag{2.1}$$
where $|\mathscr{X}^n| = |\mathscr{X}|^n$ is the **volume** of $\mathscr{X}^n$. Evidently, in order for this measure to be well defined, we have to assume that $\mathscr{X}$ is a **bounded** state space having a finite volume $|\mathscr{X}|$.

- The "physics" of an n-body system is described as a whole by its **Hamiltonian** or **energy function** $U_n(x^n)$. This function models all the possible interactions or dependencies that may exist between the n subsystems. Given $U_n(x^n)$, we define the **mean energy function** $u_n(x^n)$ or **mean energy per particle** by the ratio $u_n(x^n) = U_n(x^n)/n$.

- The interactions between the subsystems are such as to create a **thermodynamic behavior** of the whole system, i.e., a macroscopic behavior of the whole system which can be described by having recourse only to a few variables called **macrostates**. A macrostate, as the name suggests, is a macroscopic variable which, contrary to a microstate, does not provide a complete microscopic description of an n-body system, but only a "coarse-grained" description of it which uses a much smaller number of coordinates than the dimension of the microstate space $\mathscr{X}^n$. Mathematically, a macrostate is just a function $M_n(x^n)$ mapping the space $\mathscr{X}^n$ of microstates onto some space $\mathscr{M}$ which may be the real line, a vector space, etc.

- The thermodynamic or macroscopic behavior of an n-body system is characterized by an **equilibrium state**. The latter is defined operationally as corresponding to the *most probable* macrostate(s) value(s) of the system defined with respect to some probability measure which is yet to be specified.

- In trying to find the most probable values of the macrostates, the limit $n \to \infty$ with $u_n = U_n/n$ kept fixed will be assumed. Technically, we call this limit the **thermodynamic limit**; it entails, in many cases, the continuum limit.

- The equilibrium values of the macrostates enter in the parameterization of a certain number of **thermodynamic functions** which are useful for determining many physical properties of the system at equilibrium.

Physical Interpretation of the Formalism

The physical content of the formalism defined through the above points may be understood or "visualized" more clearly perhaps if we put it in the context of the most common many-body system of all: a gas. Typically the number of particles contained in a gas, e.g., the air filling the room you are in, is very large ($\sim 10^{20}$-10^{30}), whereas the scale at which these particles can be "seen" to evolve and interact is very small ($\sim 10^{-8}$-10^{-10} m). Because of the too large number of variables at play, we need obviously to abandon the hope of describing the physics of the gas by probing its microscopic deterministic evolution. What we can do, however, is to study the gas at the macroscopic level by abstracting the evolution of its constituting particles using some stochastic model, and then proceed to do a statistical analysis of this model to see if any physically realistic properties come out of it.

That such a probabilistic and statistical analysis of the gas system (and, by extension, any random many-body systems) should work is directly suggested by the following observations which follow from our analysis of large deviations:

- Random quantities, when combined together, tend to "average out" one another so as to let order and uniformity emerge from randomness (Law of Large Numbers;

macroscopic determinism).

- The repeated "measurements" of a random variable, say a macrostate $M_n(x^n)$, involving n sub-random variables $x_1, x_2, \ldots, x_n$ should tend in the same way to concentrate around some limiting "equilibrium" values (macroscopic determinism), despite the fact that the measured sequences of states x^n are randomly distributed (microscopic randomness).

- If the outcomes of the measurements of $M_n(x^n)$ are ruled by a large deviation principle, then the concentration effect should be "exponentially effective," just as a mean sum of random variables concentrates itself exponentially around the limiting value predicted by the Law of Large numbers. From this point of view, all that should matter if want to macroscopically describe a many-body system as a whole is to know the most probable values of the macrostates which are of interest to us.

- The most probable value(s) of a macrostate M_n can presumably be obtained mathematically by finding the minimum (or minima) of a certain rate function.

- Given the exponential concentration property, the most probable value(s) of a macrostate M_n should match, with a comfortable degree of confidence, the "laboratory" value(s) of the physical observable which is abstractly modelled by M_n.

These are *grosso modo* the basic precepts of equilibrium statistical physics; what remains to be done now, if we want to apply them to specific physical models, is to show how the probabilities for the microstates and the macrostates are to be constructed depending on the nature of the system studied.

Here we shall consider only two classes of systems: either the system studied is **closed** or **isolated** from its environment, in which case its energy is constant, or either the system is **open** and exchanges energy with another system which is large enough to be characterized by a fixed temperature. The first class of systems is modeled at the level of probabilities by the so-called **microcanonical ensemble**, whereas the second class is modeled by the **canonical ensemble**.

2.2. THE MICROCANONICAL ENSEMBLE

Definition

The fundamental probability distribution known as the microcanonical ensemble has been introduced by Boltzmann (1877) as a model of closed systems whose energy is kept constant.

It is anchored physically on the assumption that all microstates x^n such that $U_n(x^n) = U$, where U is the constant energy of the system, should be thought of as being equally probable (**equiprobability postulate**). Thus, what we call the **microcanonical ensemble** at the level of the microstates is the probability measure defined by conditioning the uniform prior

$$P(dx^n) = \frac{dx^n}{|\mathscr{X}^n|} \tag{2.2}$$

on the restricted set of all microstates $x^n \in \mathscr{X}^n$ having a mean energy $u_n(x^n)$ lying in the infinitesimal interval $[u, u + du]$; in symbols,

$$\begin{aligned} P^u(dx^n) &= P(dx^n | u_n \in du) \\ &= \frac{P(dx^n \cap \{x^n : u_n(x^n) \in du\})}{P(u_n \in du)} \\ &= \begin{cases} \dfrac{P(dx^n)}{P(u_n \in du)} & \text{if } u(x^n) \in du \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \tag{2.3}$$

where

$$P(u_n \in du) = \int_{\{x^n : u_n(x^n) \in du\}} P(dx^n) \tag{2.4}$$

is the probability that mean energy u_n lies in the infinitesimal interval du positioned at u. [Recall that we use the formal notation $P(u_n \in du)$ as a shorthand for $P(u_n \in [u, u + du])$.] The quantity $P(u_n \in du)$ entering in the above equations is there to make $P^u(x^n)$ a normalized measure:

$$\int_{\mathscr{X}^n} P^u(dx^n) = \frac{1}{P(u_n \in du)} \int_{\{x^n : u_n(x^n) \in du\}} P(dx^n) = 1. \tag{2.5}$$

To extend the microcanonical measure $P^u(dx^n)$ to the level of macrostates, we simply need at this stage to follow the standard rules of probability theory. For a given macrostate $M_n(x^n)$, for example, we define $P^u(M_n \in dm)$ to be the conditional probability given by

$$\begin{aligned} P^u(M_n \in dm) &= P(M_n \in dm | u_n \in du) \\ &= \frac{P(\{x^n : M_n(x^n) \in dm\} \cap \{x^n : u_n(x^n) \in du\})}{P(u_n \in du)}. \end{aligned} \tag{2.6}$$

It is this probability that one has to consider if one wants to find the most probable value(s) of M_n given that the system for which this macrostate is "measured" has a mean energy held fixed at a value u.

Microcanonical Large Deviations

The theory of large deviations enters in the description of the microcanonical ensemble as the basic tool for finding the values of M_n which maximize the microcanonical probability measure $P^u(M_n \in dm)$. To parallel the case of random mean sums, we should expect at this point to be able to find the most probable values of M_n by locating the minima of a given rate function. What we aim to show here is that if a large deviation principle holds for $P(M_n \in dm)$ as well as for $P(u_n \in du)$, and if there exists a contraction of the macrostate values into the mean energy values, then a large deviation principle also holds for $P^u(M_n \in dm)$. Moreover, in this case, the values of M_n which realize the global maximum of the microcanonical probability $P^u(M_n \in dm)$ also globally minimize the rate function of $P(M_n \in dm)$ under the constraint that $u_n(x^n) = u$.

To prove this result, consider the macrostate $M_n(x^n)$ taking values in the space $\mathscr{M}$, and suppose that a large deviation principle holds for this macrostate with respect to the *a priori* measure $P(dx^n)$, namely,

$$P(M_n \in dm) = \int_{\{x^n : M_n(x^n) \in dm\}} P(dx^n) \asymp e^{ns(m)} dm. \tag{2.7}$$

Note that to conform with the physicists' usage, we have written the "rate function" $s(m)$ in the expression of the large deviation probability without a minus sign. The negative rate function $s(m)$ thus defined is called an **entropy function**.

Suppose now that the energy per particle $u_n(x^n)$ can be rewritten as a function of the macrostate $M_n(x^n)$; that is, suppose that there exists a bounded, continuous function $u(m)$, which we shall call the **energy representation function**, that has the property that $u_n(x^n) = u(M_n(x^n))$ for any microstates x^n or, more generally, which is such that

$$|u_n(x^n) - u(M_n(x^n))| \to 0 \tag{2.8}$$

uniformly over all microstates as $n \to \infty$. Given that this function exists, it is readily seen that the most probable macrostate values m of $M_n(x^n)$ defined on the microcanonical set of microstates x^n such that $u_n(x^n) \in du$ are those which maximize the entropy function $s(m)$ subject to the constraint $u(m) = u$. To be sure, let us construct the explicit expression of the microcanonical measure for M_n. Assuming that $P(u_n \in du)$ satisfies a large deviation principle of the form

$$P(u_n \in du) \asymp e^{ns(u)}du, \tag{2.9}$$

we may write $P^u(M_n \in dm)$ as follows:

$$P^u(M_n \in dm) \asymp e^{-nI^u(m)}dm, \tag{2.10}$$

where

$$I^u(m) = \begin{cases} s(u) - s(m) & \text{if } u(m) = u \\ \infty & \text{otherwise.} \end{cases} \tag{2.11}$$

The function $I^u(m)$ has to be nonnegative for any $m \in \mathscr{M}$, since it is a rate function. Thus, we see that the microcanonical probability that $M_n(x^n)$ is near m must go to 0 at the exponential decaying rate $I^u(m)$ as $n \to \infty$ unless $I^u(m) = 0$, in which case $P^u(M_n \in dm) \to O(1)$ in the limit where $n \to \infty$. Since the minimizers of $I^u(m)$ must correspond to the constrained maximizers of $s(m)$, by virtue of (2.11), we are naturally led to define the set $\mathscr{E}^u$ of **microcanonical equilibrium macrostates** as the set containing all these constrained maximizers; in symbols,

$$\begin{aligned} \mathscr{E}^u &= \{m \in \mathscr{M} : I^u(m) = 0\} \\ &= \{m \in \mathscr{M} : m \text{ maximizes } s(m) \text{ with } u(m) = u\}. \end{aligned} \tag{2.12}$$

For later use, the members of $\mathscr{E}^u$ will be denoted by m^u.

Microcanonical Contraction Principle

The definition of the set $\mathscr{E}^u$ implies a useful variational formula for the function $s(u)$, the **microcanonical entropy function**. Indeed, since m^u is a global minimizer of the rate function $I^u(m)$ for a given value u of the mean energy, we must have $I^u(m^u) = 0$, and therefore

$$s(u) = \sup_{m:u(m)=u} s(m) = s(m^u) \tag{2.13}$$

by equation (2.11).

The same variational formula for $s(u)$ can also be seen as a consequence of the contraction principle presented in Section 1.6. Simply use the representation energy function as a contraction function to express the integral of (2.4) as an integral over the restricted macrostate space $\{m \in \mathscr{M} : u(m) = u\}$ instead of $\{x^n \in \mathscr{X}^n : u_n(x^n) = u\}$. This yields

$$\begin{aligned} P(u_n \in du) &= \int_{\{m:u(m)\in du\}} P(M_n \in dm) \\ &\asymp \int_{\{m:u(m)\in du\}} e^{ns(m)} dm. \end{aligned} \tag{2.14}$$

Then use Laplace's Method to locate the dominating contribution to the integral over m:

$$P(u_n \in du) \asymp \exp\left[n \sup_{m:u(m)=u} s(m)\right] dm, \tag{2.15}$$

and take, finally, the thermodynamic limit to obtain

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \ln P(u_n \in du) &= \sup_{m:u(m)=u} s(m) + \lim_{n\to\infty} \frac{1}{n} \ln dm \\ &= \sup_{m:u(m)=u} s(m). \end{aligned} \tag{2.16}$$

2.3. THE CANONICAL ENSEMBLE

Definition

While the microcanonical ensemble is defined in terms of a fixed value of the mean energy u, the canonical ensemble is defined in terms of a fixed value of the inverse temperature β. In the canonical ensemble, the relevant probability measure on $\mathscr{X}^n$ to consider is the **canonical probability measure** or **Gibbs measure** defined by

$$P_\beta(dx^n) = \frac{e^{-\beta U_n(x^n)}}{Z_n(\beta)} P(dx^n) = \frac{e^{-\beta n u_n(x^n)}}{Z_n(\beta)} P(dx^n), \tag{2.17}$$

where

$$Z_n(\beta) = \int_{\mathscr{X}^n} e^{-\beta n u_n(x^n)} P(dx^n). \tag{2.18}$$

(See Reif (1965), Landau and Lifshitz (1991) or Balian (1991) for a derivation of this probability measure as well as for a discussion of its interpretation.)

Following the terminology introduced in the previous chapter, we call $Z_n(\beta)$ the **partition function** of $u_n(x^n)$. From this quantity, we also define the **free energy** of $u_n(x^n)$ by the limit

$$\varphi(\beta) = \lim_{n\to\infty} -\frac{1}{n} \ln Z_n(\beta). \tag{2.19}$$

	Statistical mechanics	Large deviations theory
Large deviations	$P(u_n \in du) \asymp e^{ns(u)}$	$P(A_n \in da) \asymp e^{-nI(a)}$
Partition function	$Z_n(\beta) = E[e^{-\beta n u_n}]$	$W_n(k) = E[e^{knA_n}]$
Free energy	$\varphi(\beta) = \lim_{n\to\infty} -\frac{1}{n} \ln Z_n(\beta)$ $= \inf_u \{\beta u - s(u)\}$	$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln W_n(k)$ $= \sup_a \{ka - I(a)\}$

Table 2.1: Comparison of the quantities and fundamental mathematical equations of equilibrium statistical physics and large deviations theory.

Notice that the expressions of $Z_n(\beta)$ and $\varphi(\beta)$ differ slightly from the expressions of the partition and free energy functions given in the previous chapter. The difference, as can be seen, is only a matter of sign convention: physicists prefer to use the $-\beta$ term in the exponential function entering in the definition of the partition function, whereas mathematicians use the single k term. This only means of course that $Z_n(\beta) = W_n(k = -\beta)$. In a similar way, the "physical" free energy $\varphi(\beta)$ is defined with an extra minus sign so that $\varphi(\beta) = -\lambda(k = -\beta)$ (see Table 2.1).

Canonical Large Deviations

As we did in the case of the microcanonical ensemble, we now state the large deviation principle for the macrostate $M_n(x^n)$ with respect to the canonical ensemble, and then use this principle to define the set of canonical equilibrium macrostates.

For any macrostate value $m \in \mathscr{M}$ and any microstate $x^n \in \mathscr{X}^n$ satisfying $M_n(x^n) \in dm$, the continuity of the energy representation function implies that $u_n(x^n)$ is equal or is close to $u(M_n(x^n))$. Hence, we expect that

$$\begin{aligned} P_\beta(M_n \in dm) &= \int_{\{x^n : M_n(x^n) \in dm\}} P_\beta(dx^n) \\ &= \int_{\{x^n : M_n(x^n) \in dm\}} \frac{e^{-\beta n u_n(x^n)}}{Z_n(\beta)} P(dx^n) \end{aligned} \tag{2.20}$$

is equal or is close to

$$\frac{e^{-\beta n u(m)}}{Z_n(\beta)} \int_{\{x^n : M_n(x^n) \in dm\}} P(dx^n) = \frac{e^{-\beta n u(m)}}{Z_n(\beta)} P(M_n \in dm). \tag{2.21}$$

As a consequence, we can use the large deviation estimate (2.7) for $P(M_n \in dm)$ to write

$$P_\beta(M_n \in dm) \asymp e^{-nI_\beta(m)}dm, \tag{2.22}$$

where

$$I_\beta(m) = \beta u(m) - s(m) - \varphi(\beta). \tag{2.23}$$

This canonical large deviation principle for M_n shows that the most probable macrostate values with respect to Gibbs' measure P_β are those which minimize the quantity $\beta u(m) - s(m)$ for a fixed value of β. (The free energy function $\varphi(\beta)$ is only a constant of normalization.) Accordingly, we define the set $\mathscr{E}_\beta$ of **canonical equilibrium macrostates** associated with a given value β as follows:

$$\begin{aligned} \mathscr{E}_\beta &= \{m \in \mathscr{M} : I_\beta(m) = 0\} \\ &= \{m \in \mathscr{M} : \beta u(m) - s(m) \text{ is minimized}\}. \end{aligned} \tag{2.24}$$

As in the case of the microcanonical ensemble, the members of $\mathscr{E}_\beta$, which will be denoted by m_β, have exponentially more probability to be "observed" in the limit where $n \to \infty$ compared to any other macrostate values not contained in this set. This, we recall, is the precise explanation of the term "equilibrium macrostates."

Canonical Contraction Principles

One useful property of the canonical minimizers m_β is that they can be used to express the free energy function $\varphi(\beta)$ as a function of $u(m)$ and $s(m)$:

$$\varphi(\beta) = \inf_m \{\beta u(m) - s(m)\} = \beta u(m_\beta) - s(m_\beta). \tag{2.25}$$

This variational formula for $\varphi(\beta)$ is called the **macrostate representation** of the free energy. As in the case of the microcanonical entropy, we can derive this formula from the contraction principle: just re-express the integral entering in the definition of $Z_n(\beta)$ not as an integral over the set $\mathscr{X}^n$ of microstates, but as an integral over $\mathscr{M}$:

$$Z_n(\beta) = \int_{\mathscr{M}} e^{-\beta n u(m)} P(M_n \in dm). \tag{2.26}$$

Then use the large deviation principle for $P(M_n \in dm)$ to obtain

$$\begin{aligned} Z_n(\beta) &\asymp \int_{\mathscr{M}} e^{-n[\beta u(m)-s(m)]}dm. \\ &\asymp \exp\left(-n \inf_m \{\beta u(m) - s(m)\}\right), \end{aligned} \tag{2.27}$$

so that

$$\lim_{n\to\infty} -\frac{1}{n}\ln Z_n(\beta) = \inf_m \{\beta u(m) - s(m)\}. \tag{2.28}$$

Using a similar chain of arguments, we can also write

$$\begin{aligned} Z_n(\beta) &= \int_{\mathscr{U}} e^{-\beta n u} P(u_n(x^n) \in du) \\ &\asymp \int_{\mathscr{U}} e^{-n[\beta u - s(u)]}du \\ &\asymp \exp\left(-n \inf_u \{\beta u - s(u)\}\right). \end{aligned} \tag{2.29}$$

where $\mathscr{U}$ stands for the range of admissible values of the mean energy. The above equations clearly motivate the fundamental relationship

$$\varphi(\beta) = \inf_u \{\beta u - s(u)\} \tag{2.30}$$

which expresses the free energy $\varphi(\beta)$ as the Legendre-Fenchel transform of the microcanonical entropy $s(u)$. We call this formula the **thermodynamic representation** of $\varphi(\beta)$. The difference between this Legendre-Fenchel transform relating $\varphi(\beta)$ and $s(u)$ and the Legendre-Fenchel transform obtained in the previous chapter relating $\lambda(k)$ and $I(a)$ is, again, only a matter of sign convention (see Table 2.1).

Canonical Equilibrium Mean Energy

The mean energy has the role of a constant in the microcanonical ensemble: it is the parameter which defines P^u. But in the canonical ensemble, the mean energy is a random variable. The canonical probability measure which governs the outcomes of u_n is given by

$$P_\beta(u_n \in du) = \frac{e^{-\beta n u}}{Z_n(\beta)} P(u_n \in du). \tag{2.31}$$

Given the large deviation principle ruling the outcomes of $P(u_n \in du)$ and the entropy function $s(u)$ associated with this large deviation principle, we thus find

$$P_\beta(u_n \in du) \asymp e^{-nI_\beta(u)} du, \tag{2.32}$$

where

$$I_\beta(u) = \beta u - s(u) - \varphi(\beta). \tag{2.33}$$

This formula for the canonical rate function implies, among other things, that the equilibrium values of the mean energy in the canonical ensemble, which will be denoted by u_β, satisfies the two following differential equations:

$$\left.\frac{\partial I_\beta(u)}{\partial u}\right|_{u_\beta} = 0, \qquad \left.\frac{\partial^2 I_\beta(u)}{\partial u^2}\right|_{u_\beta} > 0, \tag{2.34}$$

if $I_\beta(u)$ is twice differentiable at u. In terms of the microcanonical entropy, this is equivalent to

$$s'(u)\big|_{u_\beta} = \beta, \qquad s''(u)\big|_{u_\beta} < 0 \tag{2.35}$$

if $s(u)$ is twice differentiable at u. Moreover, since $I_\beta(u_\beta) = 0$ by definition of the canonical equilibria, we must have

$$\varphi(\beta) = \beta u_\beta - s(u_\beta). \tag{2.36}$$

From this result we can proceed as we have done in Section 1.5 to differentiate the free energy with respect to β to find that u_β also satisfies

$$\varphi'(\beta) = u_\beta + \beta \frac{\partial u_\beta}{\partial \beta} - s'(u)\big|_{u_\beta} \frac{\partial u_\beta}{\partial \beta} = u_\beta. \tag{2.37}$$

That is, the slope value of φ at β, if it exists, gives the equilibrium value of the mean energy realized in the canonical ensemble with inverse temperature β.

2.4. ILLUSTRATIVE EXAMPLE: NONINTERACTING ± 1 SPINS

To illustrate the many equations derived so far, we consider in this section a simple spin system which is commonly presented to undergraduate physics students as the simplest microscopic model of magnetism [see, e.g., Reif (1965) or Balian (1991)]. The model consists of n noninteracting **spins** $X_1, X_2, \ldots, X_n$ taking values in the state space $\mathscr{X} = \{-1, +1\}$. The "free" nature of the spins simply means that we can write the total energy of the system as the sum of the individual energies of the spins which, for simplicity, will be assumed to be directly proportional to the spin value. Thus,

$$U_n(x^n) = \sum_{i=1}^{n} x_i \tag{2.38}$$

and

$$u_n(x^n) = \frac{U_n(x^n)}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2.39}$$

To complete the specification of the model, we set the prior probability on $\mathscr{X}^n$ to be equal to $P(x^n) = 2^{-n}$ for all microstates $x^n \in \mathscr{X}^n$.

We proceed to study the equilibrium properties of this spin model by choosing the empirical vector as the macrostate of interest. The latter is defined once again as

$$L_n(x) = \frac{\#\text{ spins } x \text{ in } x^n}{n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i,x}, \tag{2.40}$$

where x_i is the state of the i^{th} spin in the given microstate $x^n = (x_1, x_2, \ldots, x_n)$. The choice of the empirical vector is dictated by the fact that this quantity admits a simple energy representation given by

$$u(L_n = l) = \sum_{x\in\mathscr{X}} x l(x) = l(+1) - l(-1). \tag{2.41}$$

Moreover, we already know the entropy function associated with the large deviations of L_n. By Sanov's Theorem, we indeed have $P(L_n \in dl) \asymp e^{ns(l)}dl$ with

$$\begin{aligned} s(l) &= -\sum_{x\in\mathscr{X}} l(x)\ln l(x) - \ln 2 \\ &= -l(+1)\ln l(+1) - l(-1)\ln l(-1) - \ln 2. \end{aligned} \tag{2.42}$$

These properties of L_n allow us to construct the set $\mathscr{E}^u = \{l^u\}$ of equilibrium empirical vectors associated with the mean energy value u by maximizing the entropy measure $s(l)$ subject to the constraint $u(l) = u \in [-1, 1]$. Fortunately for us, solving this problem does not actually necessitate the maximization of $s(l)$ since the normalization constraint

$$l(-1) + l(+1) = 1 \tag{2.43}$$

on the components of the empirical vector reduces the number of independent components of l to one, while the microcanonical energy constraint

$$l(+1) - l(-1) = u \tag{2.44}$$

reduces this number by one more. This leaves us with no indeterminate, and so we find

$$l^u(-1) = \frac{1-u}{2}, \qquad l^u(+1) = \frac{1+u}{2} \tag{2.45}$$

as the two components of the microcanonical equilibrium empirical vector l^u associated with a given value u of the mean energy (Figure 2.1). From this solution, we calculate the microcanonical entropy function $s(u)$ by using the contraction formula $s(u) = s(l^u)$ which yields

$$s(u) = -\left(\frac{1-u}{2}\right)\ln\left(\frac{1-u}{2}\right) - \left(\frac{1+u}{2}\right)\ln\left(\frac{1+u}{2}\right) - \ln 2 \tag{2.46}$$

with $u \in [-1, +1]$. This entropy function is plotted in Figure 2.2.

In the case of the canonical ensemble, the set $\mathscr{E}_\beta = \{l_\beta\}$ of equilibrium empirical vectors parameterized by the inverse temperature β is found by maximizing the quantity $\beta u(l) - s(l)$. Here the maximization of $s(l)$ is necessary, but it is simple enough to be carried out analytically. The solution explicitly is

$$l_\beta(-1) = \frac{e^\beta}{e^\beta + e^{-\beta}}, \qquad l_\beta(+1) = \frac{e^{-\beta}}{e^\beta + e^{-\beta}}, \tag{2.47}$$

where $\beta \in \mathbb{R}$ (Figure 2.1). Using this result, we find the free energy function $\varphi(\beta)$ to be given by

$$\varphi(\beta) = \beta u(l_\beta) - s(l_\beta) = -\ln\cosh\beta \tag{2.48}$$

(Figure 2.2). The same result could have been obtained directly had we started with the definition of this quantity:

$$\begin{aligned} \varphi(\beta) &= \lim_{n\to\infty} -\frac{1}{n}\ln Z_n(\beta) \\ &= -\ln\left(\frac{1}{2}\sum_{x\in\{-1,1\}} e^{-\beta x}\right) \\ &= -\ln\cosh\beta. \end{aligned} \tag{2.49}$$

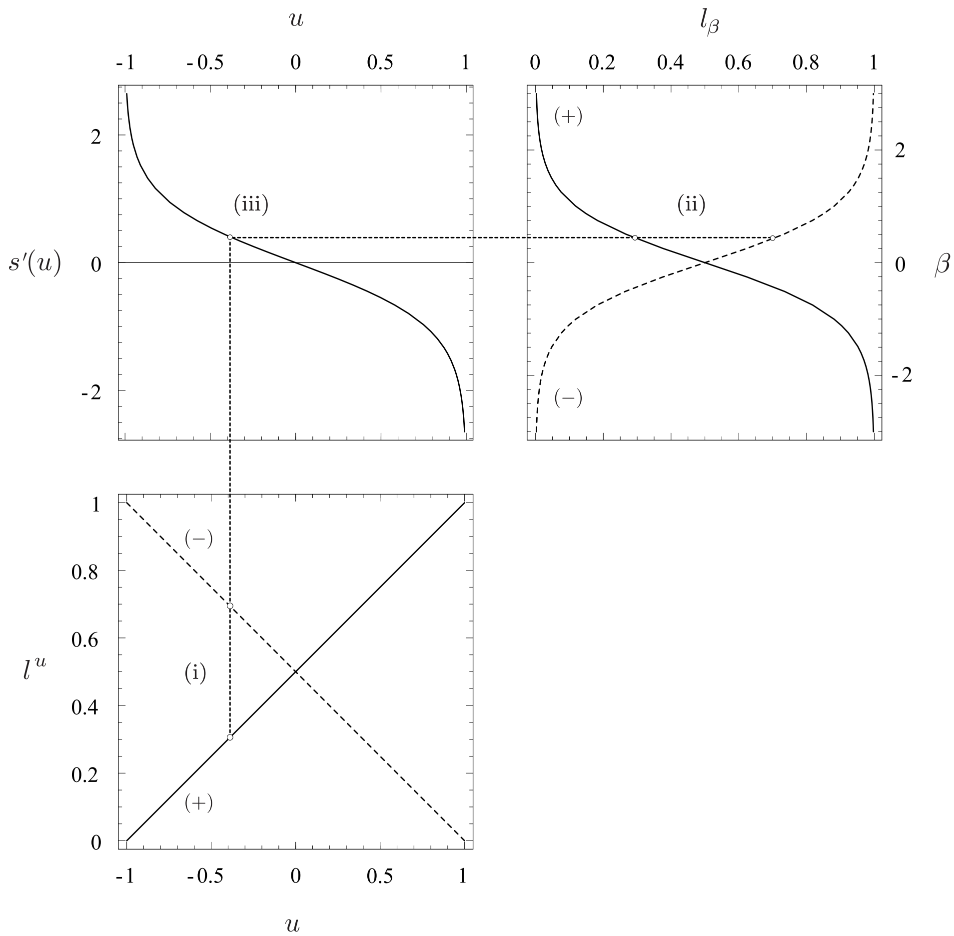

Figure 2.1: (Top left) Derivative of the microcanonical entropy. (Top right) -1 and $+1$ components of the canonical equilibrium empirical vector l_β. (Bottom left) -1 and $+1$ components of the microcanonical equilibrium empirical vector l^u. The dotted line indicates how the one-to-one correspondence between the microcanonical and canonical equilibrium values of the empirical vector is to be constructed: pick a microcanonical point l^u (i), and find a corresponding canonical point l_β such that $l^u = l_\beta$ (ii). The value β for which the equality between the two equilibrium macrostates holds should be such that $\beta = s'(u)$ (iii).

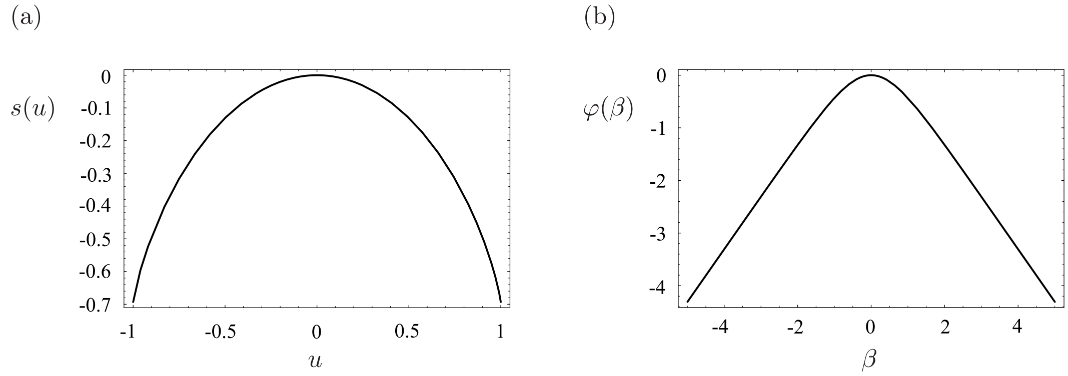

Figure 2.2: (a) Microcanonical entropy $s(u)$ for the noninteracting spin problem. (b) Associated free energy function $\varphi(\beta)$.

From this expression, we finally find

$$u_\beta = \varphi'(\beta) = -\tanh\beta \tag{2.50}$$

as the unique value of the equilibrium mean energy associated with a given value β of the inverse temperature.

We shall not reflect much on the physical implications of these results; they are, to a certain extent, "trivial" and can be found in virtually every undergraduate textbooks on statistical mechanics (see, e.g., the references mentioned in the Notes and Remarks section found at the end of the chapter). However, there are two subtle issues related to these calculations which are worth addressing in more detail; two issues which lie at the heart of this thesis, and which are but rarely addressed in textbooks.

The first of them has to do with the calculation of the microcanonical entropy function, and, more globally, with the applicability of the Gärtner-Ellis Theorem. In this section we have used the contraction principle applied to $s(l)$ to obtain the expression of $s(u)$, but we could have just as well calculated the microcanonical entropy by taking the Legendre-Fenchel transform of the corresponding free energy function $\varphi(\beta)$. In fact, given the isomorphism between the noninteracting spin model and the sum of IID binary random variables studied in Section 1.4, it is easily verified that the outcome of this transform, with $\varphi(\beta)$ given as in equation (2.48), coincides with what was found in (2.46). The reason explaining this coincidence is also already known to us: the free energy $\varphi(\beta)$ of the noninteracting spin model is everywhere differentiable, as required by the theorem of Gärtner and Ellis. But there is a missing piece in this explanation: we do not yet overly understand why we obtain the right answer under the conditions stated. More precisely, we have not explained up to now why the differentiability of $\varphi(\beta)$ is required to ensure that $s(u)$ equals the Legendre-Fenchel transform of $\varphi(\beta)$. Can we in fact imagine physical models which are characterized microcanonically by entropy functions which do not comply with the Gärtner-Ellis Theorem? This is a very important question that will find its answer in the next chapters.

The second problem raised by the above calculations is, in a sense, more fundamental than the first as it relates to the compatibility or incompatibility of the equilibrium macrostates predicted by the microcanonical and the canonical ensembles. Using the solutions for l^u and l_β found above, we may easily verify that the set $\mathscr{E}^u$ of microcanonical equilibrium macrostates can be put into a one-to-one correspondence with the set $\mathscr{E}_\beta$ of canonical equilibrium macrostates. That is, we may verify, using equation (2.50), that $l_\beta = l^{u_\beta}$ for all $\beta \in \mathbb{R}$. Conversely, we may invert the very same equation which relates u_β and β to obtain

$$\beta(u) = s'(u) = \tanh^{-1}(-u) = \frac{1}{2}\ln\left(\frac{1-u}{1+u}\right), \tag{2.51}$$

and check that $l^u = l_{\beta(u)}$ for all $u \in [-1,+1]$ (Figure 2.1). We thus see that the microcanonical and canonical ensembles actually give equivalent predictions for what regards the equilibrium value of the empirical vector. Faced with this surprising result, we should evidently ask ourselves if this is just an accident, a coincidence resulting from the choice of the model. Could it be that $\mathscr{E}^u$ is always equal to $\mathscr{E}_\beta$? or are there any models for which we may observe that $\mathscr{E}^u \neq \mathscr{E}_\beta$ for some values of u or β? This other question will also find its answer in the next chapters.

SUMMARY OF CHAPTER

- **Uniform prior measure:**

$$P(dx^n) = \frac{dx^n}{|\mathscr{X}^n|}. \tag{2.52}$$

  Here $|\mathscr{X}^n|$ represents the volume of the n-particle state space $\mathscr{X}^n$ which is assumed to be bounded.

- **Energy representation function:** Function $u(m)$ such that $u_n(x^n) = u(M_n(x^n))$ or

$$\lim_{n\to\infty} |u_n(x^n) - u(M_n(x^n))| = 0. \tag{2.53}$$

- **Microcanonical entropy:** The microcanonical entropy $s(u)$ is the negative rate function associated with the probability measure $P(u_n \in du)$, i.e.,

$$P(u_n \in du) \asymp e^{ns(u)} du. \tag{2.54}$$

- **Macrostate rate function:** $P(M_n \in dm) \asymp e^{ns(m)} dm$. The function $s(m)$ is the entropy function (negative rate function) associated with the large deviations of the macrostate M_n.

- **Microcanonical ensemble probability measure:** The microcanonical ensemble properties of a system are defined relatively to the set of all microstates x^n having a fixed values of the mean energy $u_n(x^n) = u$. At the macrostate level, a microcanonical probability measure is thus constructed as a conditional probability measure of the form

$$\begin{aligned} P^u(M_n \in dm) &= P(M_n \in dm | u_n \in du) \\ &= \frac{P(\{x^n : M_n \in dm\} \cap \{x^n : u_n \in du\})}{P(u_n \in du)}. \end{aligned} \tag{2.55}$$

- **Microcanonical rate function:** If the macrostate M_n and the mean energy function $u_n(x^n)$ each satisfy a large deviation principle with respect to the uniform prior probability measure P with entropy functions $s(m)$ and $s(u)$, respectively, then

$$\begin{aligned} P^u(M_n \in dm) &\asymp e^{-nI^u(m)} dm \\ I^u(m) &= \begin{cases} s(u) - s(m) & \text{if } u(m) = u \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{2.56}$$

  The rate function $I^u(m)$ is called the microcanonical rate function.

- **Microcanonical equilibria:** The equilibrium, i.e., most probable macrostate values m^u realized in the microcanonical ensemble for a fixed value u of the mean energy are determined as follows:

$$\begin{aligned} \mathscr{E}^u &= \{m \in \mathscr{M} : I^u(m) = 0\} \\ &= \{m \in \mathscr{M} : m \text{ maximizes } s(m) \text{ with } u(m) = u\} \\ &= \{m^u\}. \end{aligned} \tag{2.57}$$

- **Macrostate representation of the microcanonical entropy:**

$$s(u) = \sup_{m:u(m)=u} s(m) = s(m^u). \tag{2.58}$$

- **Canonical ensemble probability measure:** The canonical ensemble is defined at the level of the microstates x^n by the following probability measure:

$$P_\beta(dx^n) = \frac{e^{-\beta n u_n(x^n)}}{Z_n(\beta)} P(dx^n). \tag{2.59}$$

  The parameter β is called the inverse temperature.

- **Partition function:**

$$Z_n(\beta) = \int_{\mathscr{X}^n} e^{-\beta n u_n(x^n)} P(dx^n). \tag{2.60}$$

- **Free energy function:**

$$\varphi(\beta) = \lim_{n\to\infty} -\frac{1}{n} \ln Z_n(\beta). \tag{2.61}$$

- **Equilibrium value of the mean energy:** If $\varphi(\beta)$ is differentiable at β, then the most probable value of $u_n(x^n)$ in the canonical ensemble satisfies the equation $u = \varphi'(\beta)$. The unique value of mean energy determined by this equation is denoted by u_β.

- **Canonical rate function:**

$$\begin{aligned} P_\beta(M_n \in dm) &\asymp e^{-nI_\beta(m)} dm \\ I_\beta(m) &= \beta u(m) - s(m) - \varphi(\beta). \end{aligned} \tag{2.62}$$

- **Canonical equilibria:**

$$\begin{aligned} \mathscr{E}_\beta &= \{m \in \mathscr{M} : I_\beta(m) = 0\} \\ &= \{m \in \mathscr{M} : \beta u(m) - s(m) \text{ is minimized}\} \\ &= \{m_\beta\}. \end{aligned} \tag{2.63}$$

- **Thermodynamic representation of the free energy:**

$$\varphi(\beta) = \inf_u \{\beta u - s(u)\} = \beta u_\beta - s(u_\beta). \tag{2.64}$$

- **Macrostate representation of the free energy:**

$$\varphi(\beta) = \inf_m \{\beta u(m) - s(m)\} = \beta u(m_\beta) - s(m_\beta). \tag{2.65}$$

NOTES AND REMARKS

Further References on Statistical Mechanics

There is a plethora of textbooks on statistical mechanics that can be consulted to complete the material presented in this chapter. In addition to all the books already mentioned [Reif (1965), Landau and Lifshitz (1991), Balian (1991)], we may add the following ones: Wannier (1966), Huang (1987), Pathria (1996) and Salinas (2001). In this list, Balian (1991) stands as a useful source of information covering the historic development of statistical mechanics starting from the works of Boltzmann (1877), Gibbs (1902) and Einstein [see Einstein (1987)].

Unfortunately, none of these references treats the microcanonical and canonical ensembles from the point of view of large deviations theory. The use of this theory in connection with statistical mechanics is, in fact, relatively "contemporary" in comparison with statistical mechanics itself. The definitive and now classic reference on large deviations and statistical mechanics is the book by Ellis (1985). Since the content of this book is fairly mathematical, we suggest that the reader take a look at the review papers by Ellis (1995, 1999) before embarking on the reading of his book. The more physically-oriented review paper by Oono (1989) may also be a good entry point in the more mathematically-involved literature on large deviations. We also suggest the long paper by Lanford (1973), often cited as the first (mathematical physics) work on large deviations and equilibrium statistical mechanics as well as the papers by Lehtonen and Nummelin (1990) and LaCour and Schieve (2000) which all present standard applications of the large deviations formalism in the context of noninteracting systems.

We mention finally the works of Ellis, Haven and Turkington (2000), and Eyink and Spohn (1993) for those who would like to learn how to handle the thermodynamic limit in the most rigorous way (something that we have not done and will not do in this thesis).

A Translation Guide

The following few remarks are meant to complement our presentation of statistical mechanics and its affiliation with large deviations theory.

- What physicists call the free energy is the quantity $\beta\varphi(\beta)$, not $\varphi(\beta)$. In this thesis, we work mainly with $\varphi(\beta)$ because this quantity has the convenient property that it is always concave independently of the sign of β. The quantity $\beta\varphi(\beta)$, by contrast, is concave or convex depending on the sign of β (positive or negative, respectively).

- Physicists rarely care about specifying the *a priori* uniform probability $P(dx^n) = dx^n/|\mathscr{X}^n|$ in the definition of the microcanonical and canonical ensembles because the term $|\mathscr{X}^n|$ is only a constant. Their definition of the partition function, which usually takes the form
$$Z_n(\beta) = \int_{\mathscr{X}^n} e^{-\beta U_n(x^n)} dx^n \tag{2.66}$$
instead of
$$Z_n(\beta) = \int_{\mathscr{X}^n} e^{-\beta U_n(x^n)} P(dx^n) \tag{2.67}$$
as used here, reflects this choice.

  The reader can verify by him- or herself that the replacement of the normalized measure $dx^n/|\mathscr{X}^n|$ by the unnormalized (Lebesgue) measure dx^n changes absolutely nothing in the equations defining $\mathscr{E}^u$ and $\mathscr{E}_\beta$. What is "measured" by dx^n is the element of volume occupied by the microstate x^n rather than its probability. But probabilities need not be normalized to 1 in order to be physically meaningful: they can sum up to any value which means that any choice for $P(dx^n)$, as long as it is proportional to dx^n, will lead to the same results for m^u and m_β.

- The variational principle behind the definition of the microcanonical ensemble, which says precisely that the elements of $\mathscr{E}^u$ maximize the macrostate entropy function $s(m)$ subject to the energy constraint that $u(m) = u$, is what physicists call the **maximum entropy principle** [Jaynes (1957*a*, 1957*b*)]. The version of this principle that we have given here is *the* definitive version based on probability theory. In the same vein, the minimization of $\beta u(m) - s(m)$ which defines the elements of $\mathscr{E}_\beta$ is referred to by physicists as the **minimum free energy principle**, a misnomer perhaps given that $\beta u(m) - s(m)$ does not exactly represent the free energy.

3

Equivalence and Nonequivalence of Ensembles: Thermodynamic Level

We shall try to expound in this chapter the deep mathematical reasons explaining why the free energy function $\lambda(k)$ is required to be essentially smooth in the statement of the Gärtner-Ellis Theorem, and why, concomitantly, $\varphi(\beta)$ needs to be a smooth function of the inverse temperature β in order for the microcanonical entropy function $s(u)$ to be expressible as the Legendre-Fenchel transform of $\varphi(\beta)$. Following the approach of the first chapter, we begin by exploring the problem in bare, abstract mathematical terms to then study the physical consequences of what was learned from the mathematics. By passing to the physical world, we shall learn, among other things, that the two thermodynamic descriptions of statistical mechanical models obtained from the microcanonical and canonical ensembles differ from one another when the microcanonical entropy cannot be expressed as the Legendre-Fenchel transform of the canonical free energy. In such a case, we say that the two ensembles are **thermodynamically nonequivalent**.

3.1. THE MIXED SUM PROBLEM

The mathematical example that we intend to study in this section is due to Ioffe (1993). It has the interesting property that it involves a mean sum of random variables having a nonconvex rate function. Let $X^n = (X_1, X_2, \ldots, X_n)$ be a sequence of n independent and identically distributed (IID) Gaussian random variables having zero mean and unit variance, and let Y be a binary random variable such that

$$P(Y=-1) = P(Y=+1) = \frac{1}{2}. \tag{3.1}$$

The mean sum considered by Ioffe is then

$$A_n(X^n, Y) = Y + \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{3.2}$$

To find the form of the large deviations ruling the outcomes of the random variable A_n, we consider the following decomposition of $P(A_n \in da)$:

$$P(A_n \in da) = \sum_{y=\pm 1} P(A_n \in da | Y = y) P(Y = y). \tag{3.3}$$

Given what we know about Gaussian mean sums, it is plainly seen that if we fix the value of Y to -1, then A_n becomes a Gaussian mean sum centered at the expected value

$$E[A_n | Y = -1] = \int_{\mathscr{A}} a P(A_n \in da | Y = -1) = -1 \tag{3.4}$$

so that $P(A_n \in da | Y = -1)$ satisfies a large deviation principle with rate function

$$I_-(a) = \frac{(a+1)^2}{2}. \tag{3.5}$$

Similarly, for $Y = 1$, A_n becomes a Gaussian mean sum centered at the value $E[A_n | Y = 1] = 1$, which means that $P(A_n \in da | Y = 1)$ must satisfy a large deviation principle having a rate function given by

$$I_+(a) = \frac{(a-1)^2}{2}. \tag{3.6}$$

Combining these results in (3.3), we obtain

$$P(A_n \in da) \asymp \frac{1}{2} e^{-nI_-(a)} da + \frac{1}{2} e^{-nI_+(a)} da \asymp e^{-nI(a)} da \tag{3.7}$$

with

$$I(a) = \min\{I_-(a), I_+(a)\} = \begin{cases} I_-(a) & \text{if } a < 0 \\ I_+(a) & \text{if } a \geq 0. \end{cases} \tag{3.8}$$

It can be seen from Figure 3.1(a) that the above rate function has, contrary to all the simple mean sums that we have encountered so far, not one but *two* minima, and that it is furthermore not convex. Based on these observations, we should naturally expect that something wrong will happen if we try to apply the Gärtner-Ellis Theorem to find the rate

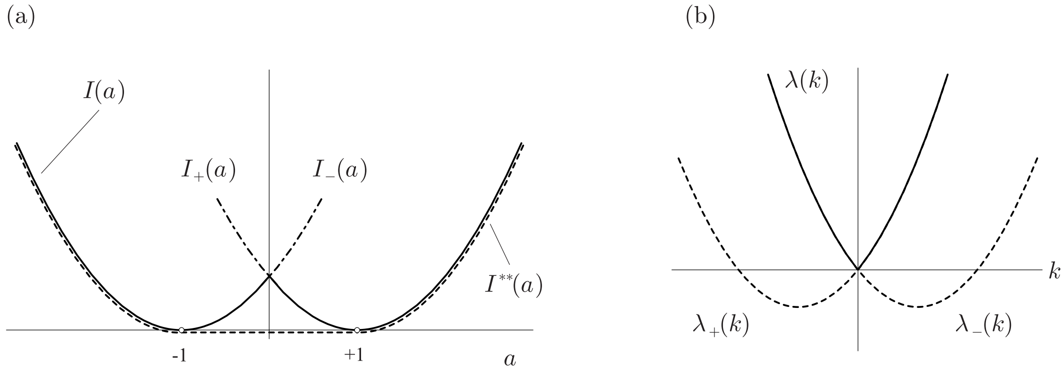

Figure 3.1: (a) (Full line) Nonconvex rate function $I(a)$ for the mixed Gaussian sum problem. (Dotted line) Function $I^{**}(a)$ corresponding to the convex envelope of $I(a)$. (b) (Full line) Free energy $\lambda(k)$ of A_n having a non-differentiable point at $k = 0$.

function of $P(A_n \in da)$ starting from the free energy of A_n, for we know that this theorem only yields strictly convex rate function (see Section 1.5). Let us verify that this is indeed the case. Using the expression of the free energy function for the Gaussian mean sum calculated in Sections 1.1 and 1.4, we proceed to the calculation of $\lambda(k)$ for the mixed sum (3.2) through the following steps:

$$\begin{aligned}
\lambda(k) &= \lim_{n\to\infty} \frac{1}{n} \ln E[e^{nkA_n}] \\
&= \lambda_{\text{Gauss}}(k) + \lim_{n\to\infty} \frac{1}{n} \ln E[e^{nkY}] \\
&= \frac{1}{2}k^2 + \lim_{n\to\infty} \frac{1}{n} \ln \frac{e^{nk} + e^{-nk}}{2} \\
&= \frac{1}{2}k^2 + |k|, \qquad k \in \mathbb{R}.
\end{aligned} \tag{3.9}$$

This result is consistent with the fact that $\lambda(k)$ must be a convex function of k, as well as with the fact that $\lambda(k)$ has to be equal to the Legendre-Fenchel transform of $I(a)$, as proved in Section 1.3. To verify the latter property, simply note that

$$\begin{aligned}
\lambda(k) &= \sup_a \{ka - I(a)\} \\
&= \sup_a \{ka - \min\{I_-(a), I_+(a)\}\} \\
&= \max\{\lambda_-(k), \lambda_+(k)\} \\
&= \begin{cases} \lambda_-(k) & \text{if } k < 0 \\ \lambda_+(k) & \text{if } k \geq 0, \end{cases}
\end{aligned} \tag{3.10}$$

where

$$\lambda_-(k) = \sup_a\{ka - I_-(a)\} = -k + \frac{1}{2}k^2$$
$$\lambda_+(k) = \sup_a\{ka - I_+(a)\} = k + \frac{1}{2}k^2 \tag{3.11}$$

are the two free energy functions associated with $I_-(a)$ and $I_+(a)$, respectively. These equations are illustrated in Figure 3.1(b).

Let us see now what we obtain if we invert the Legendre-Fenchel transform shown in the display (3.10) in the manner of the Gärtner-Ellis Theorem. Focusing our attention on the values k for which $\lambda(k)$ can be differentiated (i.e., $k \neq 0$), we find the supremum of $ka - \lambda(k)$ as we did in Section 1.4 by applying the standard rules of calculus to write

$$\sup_k\{ka - \lambda(k)\} = k(a)a - \lambda(k(a)), \tag{3.12}$$

where $k(a)$ is given by solving the differential equation $\lambda'(k) = a$. From (3.10), we directly find $k(a) - 1 = a$ for $k(a) < 0$ and $k(a) + 1 = a$ for $k(a) > 0$. We obtain, as as result,

$$\sup_k\{ka - \lambda(k)\} = \begin{cases} I_-(a) & \text{if } a < -1 \\ I_+(a) & \text{if } a > 1. \end{cases} \tag{3.13}$$

For $|a| \leq 1$, we may note that $ka \leq \lambda(k)$, and so

$$\sup_k\{ka - \lambda(k)\} = 0. \tag{3.14}$$

The combination of these two results thus yields

$$I^{**}(a) = \sup_{k\in\mathbb{R}}\{ka - \lambda(k)\} = \begin{cases} I_-(a) & \text{if } a < -1 \\ 0 & \text{if } a \in [-1, 1] \\ I_+(a) & \text{if } a > 1. \end{cases} \tag{3.15}$$

And there comes the surprise. This last rate function does not reproduce the true rate function (3.8) of A_n that was calculated above. In fact, it can be seen from Figure 3.1(a) that $I^{**}(a)$ coincides with $I(a)$ only for $a \in \mathbb{R}\backslash(-1, 1)$; for $a \in (-1, 1)$, $I^{**}(a)$ takes on the constant value 0 while $I(a)$ varies continuously between 0 and $\frac{1}{2}$.

3.2. PROPERTIES OF LEGENDRE-FENCHEL TRANSFORMS

To understand the origin of the properties of the function $I^{**}(a)$ that was just calculated, and understand furthermore how these properties compare with those of $I(a)$, we review in this section a few results pertaining to the theory of Legendre-Fenchel transforms. For a general and rigorous introduction to these transforms, including all the proofs which will be omitted here, see Rockafellar (1970).

Definition of Convex Conjugates

The double-star notation used to denote the Legendre-Fenchel transform of $\lambda(k)$ finds its meaning in the fact that the Legendre-Fenchel transform is usually denoted by a single star in the convex analysis literature [see, e.g., Rockafellar (1970)], and the fact that $\lambda(k)$ is itself the Legendre-Fenchel transform of the rate function I:

$$I^*(k) = \sup_a \{ka - I(a)\} = \lambda(k). \tag{3.16}$$

The function $I^*(k) = \lambda(k)$ is sometimes called the **convex conjugate** or **dual** of $I(a)$. By re-applying the Legendre-Fenchel transform, we thus obtain

$$\begin{aligned} I^{**}(a) &= (I^*)^*(a) \\ &= \sup_k \{ka - I^*(k)\} \\ &= \sup_k \{ka - \lambda(k)\}. \end{aligned} \tag{3.17}$$

From the expression of $I^*(k)$, we see that $I^{**}(a)$ is also equal to

$$\begin{aligned} I^{**}(a) &= \sup_k \left\{ ka - \sup_b \{kb - I(b)\} \right\} \\ &= \sup_k \inf_b \{k(a-b) + I(b)\}. \end{aligned} \tag{3.18}$$

Convexification of $I(a)$

The function $I^{**}(a)$ is always convex on its domain of definition, and equals the minimal convex function minorizing $I(a)$ for all a, which implies that $I(a) \geq I^{**}(a)$. It is called, because of these properties, the **convex hull** of $I(a)$. When $I(a) = I^{**}(a)$, we say that $I(a)$ is **convex** at the value a, whereas when $I(a) \neq I^{**}(a)$, we say that $I(a)$ is **nonconvex** at a. We also say that $I(a)$ is strictly convex at a if $I(a) = I^{**}(a)$ and $I^{**}(a)$ is itself strictly convex at a in the sense that it is not affine around a. For the mixed sum problem of the previous section, for example, we see that $I(a)$ is convex for $a \in \mathbb{R}\backslash(-1,1)$ and nonconvex for $a \in (-1,1)$. The function $I^{**}(a)$, for the same problem, is convex everywhere as expected, but has the particularity of being non-strictly convex over $(-1,1)$. In this case, we also say that $I^{**}(a)$ is **affine** or has an **affine part** over $(-1,1)$.

Inversion of the Legendre-Fenchel Transform

The above definitions together with the set of equations (3.17) provide us with a fundamental criterion for verifying that the rate function associated with a large deviation principle can be written as the Legendre-Fenchel transform of a free energy function. From (3.17), we see that the rate function $I(a)$ governing the large deviations of a given random variable A_n can be expressed as the Legendre-Fenchel of the free energy function $\lambda(k)$ of A_n if and only if $I(a)$ is globally convex on its domain of definition. More precisely, we have that

$$I(a) = \sup_k \{ka - \lambda(k)\} \tag{3.19}$$

is a valid equation at $A_n = a$ if and only if $I(a) = I^{**}(a)$. As a converse, we have then that

$$I(a) > \sup_k \{ka - \lambda(k)\} = I^{**}(a) \tag{3.20}$$

if and only if $I(a) \neq I^{**}(a)$. That is to say, $I(a)$ cannot be expressed as the Legendre-Fenchel transform of λ at a when $I(a) \neq I^{**}(a)$.

Convexity of $\lambda(k)$

The fact that the graph of $I(a)$ may contain nonconvex branches has no effect on the calculation of the free energy function $\lambda(k)$. The latter quantity is an always convex function of k, as was shown in Section 1.3, which means that the formula

$$\lambda(k) = \sup_a \{ka - I(a)\} = I^*(k) \tag{3.21}$$

holds no matter what form $I(a)$ has, be it convex or nonconvex. In fact, it can be verified that $\lambda(k)$ is not only the Legendre-Fenchel transform of $I(a)$, but also the Legendre-Fenchel transform of $I^{**}(a)$, as well as any other function having $I^{**}(a)$ as its convex hull. This motivates further our claim that the Legendre-Fenchel transform is in general a many-to-one mapping which can be inverted only for a restricted class of functions, namely the class

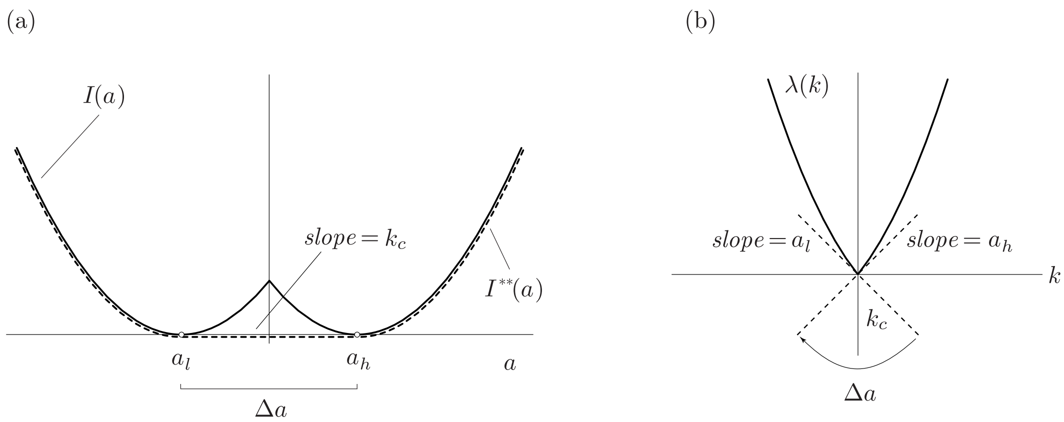

Figure 3.2: (a) Generic properties of a nonconvex rate function I (solid line), its convex hull I^{**} (dotted line), and their associated free energy function λ (b).

of convex functions.

Non-Differentiable Points of $\lambda(k)$

The regions of nonconvexity of $I(a)$, if there are any, are indicated at the level of the free energy function by the existence of points of $\lambda(k)$ where this function is non-differentiable. To be more precise, suppose that $I(a)$ is nonconvex over a single contiguous open interval (a_l, a_h) as in Figure 3.2. (The case of multiple non-overlapping ranges of nonconvexity is treated similarly.) Then it can be proved that $\lambda(k)$ must be such that $\lambda'(k)$ does not exist at the point $k_c = I'(a_l) = I'(a_h)$ [Figure 3.2(a)]. The left- and right-derivatives of $\lambda(k)$ must exist however, and must be equal to

$$\lim_{k \nearrow k_c} \lambda'(k) = \lambda'(k_c - 0) = a_l \tag{3.22}$$

and

$$\lim_{k \searrow k_c} \lambda'(k) = \lambda'(k_c + 0) = a_h, \tag{3.23}$$

respectively [Figure 3.2(b)]. The jump in the derivative of $\lambda(k)$ at k_c is thus found to be equal to

$$\Delta a = \lambda'(k_c + 0) - \lambda'(k_c - 0) = a_h - a_l. \tag{3.24}$$

To understand these equations, let us recall an important result of Chapter 1 (Section 1.5) which stated that the slopes values of $\lambda(k)$ are one-to-one related to the abscissa values of $I(a)$, while the slopes values of $I(a)$ are one-to-one related to the abscissa values of

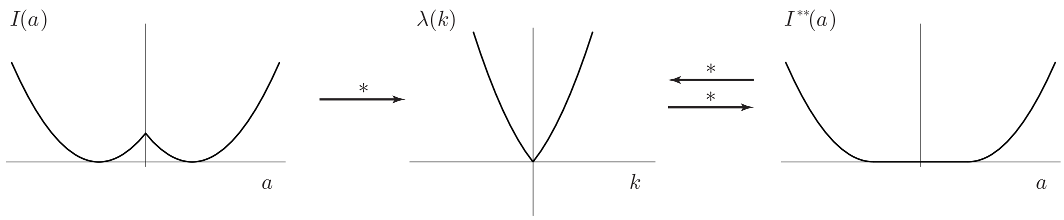

Figure 3.3: Non-involutive character of the Legendre-Fenchel transform ($*$-transform). The Legendre-Fenchel transform of a nonconvex rate function $I(a)$ (left) is the same as the Legendre-Fenchel transform of its convex hull $I^{**}(a)$ (right). The Legendre-Fenchel transform of the corresponding free energy $\lambda(k)$ (center) yields the convex hull of $I(a)$.

$\lambda(k)$. What we have to realize now is that this holds true provided that $I(a)$ is strictly convex, for if $I(a)$ is not convex over some interval, say again the open interval (a_l, a_h), then more than one value a can solve the equation $I'(a) = k$. In such a case, we see from the above equations that the Legendre-Fenchel transform acts on $I(a)$ in such a way that it assigns to $\lambda(k_c + 0)$ and $\lambda(k_c - 0)$ the two slope values corresponding to the two endpoints of the nonconvexity interval (a_l, a_h). The point k_c where $\lambda(k)$ thus becomes non-differentiable corresponds, as was claimed above, to the slope of the affine part of $I^{**}(a)$ which stretches across the two values a_l and a_h. In Figure 3.2, for example, we have $\lambda'(0^-) = a_l$ and $\lambda'(0^+) = a_h$ due to the fact that $I^{**\prime}(a) = 0$ for all $a \in [a_l, a_h]$.

The same applies obviously if $I(a)$ is convex on the same interval (a_l, a_h), but in a non-strict way as is the case for $I^{**}(a)$. This must be so because the free energy $\lambda(k)$ is at the same time the Legendre-Fenchel transform of $I(a)$ and of $I^{**}(a)$, i.e., $(I)^* = (I^{**})^* = \lambda$ (Figure 3.3). Thus, although a convex rate function $I(a)$ having one or more affine parts may in theory be expressed as the Legendre-Fenchel transform of a free energy function, one cannot distinguish, using the sole knowledge of the associated free energy, such a rate function from any other rate function which would be nonconvex over the parts where $I(a)$ is affine. This explains why $\lambda(k)$ is required to be everywhere differentiable in the statement of the Gärtner-Ellis Theorem: the *passage* of nonconvex rate functions to non-differentiable free energies is multi-valued, and cannot, as a result, be inverted.

3.3. EQUIVALENCE AND NONEQUIVALENCE RESULTS

Calculation of $s(u)$

All the results and comments of the preceding section carry over, with only minor modifications, to the microcanonical entropy function. The analog of $I^{**}(a)$ for $s(u)$ is the **concave**

hull defined by the modified Legendre-Fenchel transform

$$s^{**}(u) = \inf_{\beta}\{\beta u - \varphi(\beta)\}. \tag{3.25}$$

Since the entropy function $s(u)$ is a negative rate function, we now define $s(u)$ to be **concave** at u if $s(u) = s^{**}(u)$, and **nonconcave** at u otherwise. Furthermore, $s(u)$ is said to be **strictly concave** at u if $s(u) = s^{**}(u)$ and if $s^{**}(u)$ is itself strictly concave at u in the sense that it is locally not affine around u (Figure 3.4).

It follows from these definitions that if $s(u)$ is concave on its domain of definition, then $s(u)$ and $s^{**}(u)$ must coincide, so that

$$s(u) = \inf_{\beta}\{\beta u - \varphi(\beta)\}. \tag{3.26}$$

In this situation, we say that there is **equivalence** of the microcanonical and canonical ensembles at the thermodynamic level. We can be more precise in this definition, and say that the two ensembles are **thermodynamically equivalent** at the mean energy value u whenever $s(u) = s^{**}(u)$ (concave entropy), i.e., whenever the value of $s(u)$ at u is given by the Legendre-Fenchel transform (3.26). In the case where $s(u) = s^{**}(u)$ and $s^{**}(u)$ is strictly concave at u, then (3.26) actually reduces to the usual differential form of the Legendre transform. This is given by

$$s(u) = \beta(u)u - \varphi(\beta(u)), \tag{3.27}$$

where $\beta(u)$ is at the same time the unique root of the equation $\varphi'(\beta) = u$ and the slope value of $s(u)$, i.e., $\beta(u) = s'(u)$. Finally, when $s(u) \neq s^{**}(u)$ (nonconcave entropy), we say that there is **thermodynamic nonequivalence** of ensembles at u.

Concavity of $\varphi(\beta)$

As is the case for $\lambda(k)$, the non-invertibility of the Legendre-Fenchel transform for nonconcave functions has no effect on the calculation of $\varphi(\beta)$. The latter quantity is an always concave function of the inverse temperature, which means that the basic thermodynamic representation of $\varphi(\beta)$, given by

$$\varphi(\beta) = \inf_{\beta}\{\beta u - s(u)\}, \tag{3.28}$$

holds regardless of the form of $s(u)$. As we have just seen, however, $s(u)$ can be expressed as the Legendre-Fenchel transform of $\varphi(\beta)$ if and only if $s(u)$ is concave on its domain of

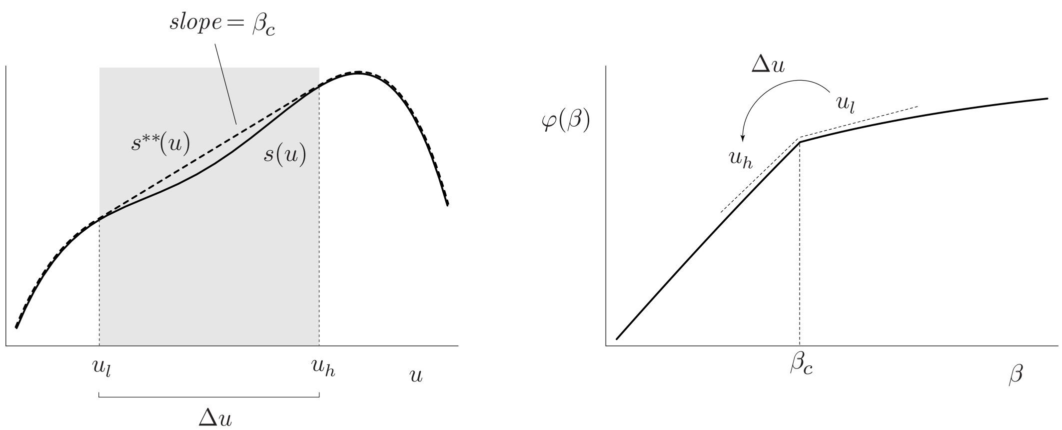

Figure 3.4: (a) Plot of a typical nonconcave entropy function $s(u)$ together with its concave envelope $s^{**}(u)$. The shaded area corresponds to the nonconcavity region of $s(u)$. This region determines the range of mean energy values for which we have thermodynamic nonequivalence of the microcanonical and canonical ensembles. (b) Corresponding free energy function $\varphi(\beta)$ obtained by calculating the Legendre-Fenchel transform of $s(u)$.

definition. In this sense, the microcanonical ensemble can be thought of as being more fundamental than the canonical ensembles.

Relationship with Canonical First-Order Phase Transitions

The apparent superiority of the microcanonical ensemble over the canonical ensemble does not prevent us from deriving a criterion based entirely on the canonical ensemble for verifying that the two ensembles are thermodynamically equivalent. Indeed, suppose that $\varphi(\beta)$ is differentiable for all β. Then the Gärtner-Ellis Theorem guarantees that, with respect to the *a priori* measure P, the energy per particle $u(x^n)$ satisfies a large deviation principle with entropy function $s(u)$ given by the Legendre-Fenchel transform of $\varphi(\beta)$. Because $\varphi(\beta)$ is assumed to be everywhere differentiable, the general theory of these transforms guarantees that $s(u)$ is strictly concave on its domain of definition. We conclude that *if $\varphi(\beta)$ is everywhere differentiable, then thermodynamic equivalence of ensembles holds for all admissible values of the mean energy.*

This result can be expressed in more physical terms by saying that the absence of a first-order phase transition in the canonical ensemble implies that the ensembles are equivalent at the thermodynamic level. By taking the contrapositive of this statement, we may also say that concomitant to a range of thermodynamic nonequivalence of ensemble is the presence of a first-order phase transition in the canonical ensemble. The precise critical value β_c

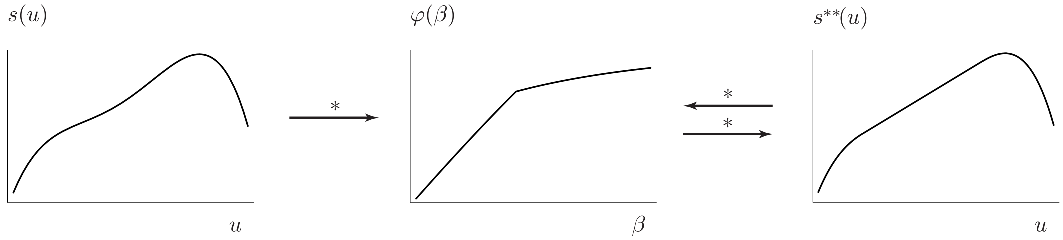

Figure 3.5: Non-involutive character of the Legendre-Fenchel transform ($*$-transform) now illustrated at the level of the microcanonical entropy (see Figure 3.3).

of the inverse temperature at which the phase transition appears is determined similarly as for $\lambda(k)$ by the properties of the Legendre-Fenchel transform. If $s(u) \neq s^{**}(u)$ for all $u \in (u_l, u_h)$, for example, then β_c must correspond to the slope of the affine part of $s^{**}(u)$ which is such that

$$\beta_c = s'(u_l) = s'(u_h) = s^{**\prime}(u_l) = s^{**\prime}(u_h) \tag{3.29}$$

(see Figure 3.4). The phase transition is also characterized by a jump of entropy

$$\Delta s = s(u_h) - s(u_l), \tag{3.30}$$

and a so-called **latent heat** which is given here by the length of the interval of thermodynamic nonequivalence:

$$\Delta u = u_h - u_l = \varphi'(\beta_c - 0) - \varphi'(\beta_c + 0). \tag{3.31}$$

(See Binney et al. (1992) for more information about first-order phase transitions and phase transitions in general.)

Unfortunately, the occurrence of a first-order phase transition in the canonical ensemble, as defined by the existence of a non-differentiable point of $\varphi(\beta)$, cannot be taken as a sufficient condition for concluding that there is thermodynamic nonequivalence of ensembles over some range of mean energy. To be sure, notice that the non-differentiability of $\varphi(\beta)$ at some critical inverse temperature β_c is related to one of the following situations: either $s(u)$ is not concave over some interval (u_l, u_h) or else $s(u)$ is affine with slope β_c over the same range of mean energy values (Figure 3.5). The latter situation, clearly, is not a situation of ensembles nonequivalence.

We shall revisit the subject of first-order canonical phase transitions and their relationship with nonequivalent ensembles in the next chapter to discuss a few subtle points about these transitions which cannot be addressed at this point. These subtleties are related, in short, to the fact that canonical first-order phase transitions arising from affine parts of $s(u)$ are, to some extent, of a different nature than those arising from nonconcave parts of $s(u)$. The difference is not perceptible from the point of view of the free energy because both "types" of first-order transitions are indistinctly associated with non-differentiable points of the free energy. But a thorough study of the canonical rate function $I_\beta(u)$ will reveal to us that the equilibrium mean energy u_β does actually behave differently around critical points depending on whether $s(u)$ has flat or nonconcave parts.

3.4. NEGATIVE HEAT CAPACITIES

It is common in the physics literature (see the many references mentioned at the end of this chapter) to characterize the microcanonical and canonical ensembles as being nonequivalent whenever the **heat capacity**, calculated microcanonically as a function of the mean energy, is found to be negative. At first, the idea of a microcanonical heat capacity is a little puzzling because this quantity, being defined mathematically as the rate of change of the equilibrium mean energy for given a change of temperature

$$c_{\text{can}}(T) = \frac{du}{dT} \tag{3.32}$$

or as

$$c_{\text{can}}(\beta) = \frac{du_\beta}{d(\beta^{-1})} = -\beta^2 \frac{du_\beta}{d\beta} \tag{3.33}$$

in terms of the inverse temperature $\beta = T^{-1}$ ($k_B = 1$), involves the very notion of temperature. However, some physicists have had the ingenuity to define a microcanonical analog of the heat capacity by formally substituting $s'(u)$ for β into equation (3.33), and take the derivative with respect to u (now interpreted as the mean energy of the microcanonical ensemble) rather than β so as to obtain

$$\begin{aligned} c_{\text{micro}}(u) &= -s'(u)^2 \left(\frac{du}{ds'(u)} \right) \\ &= -s'(u)^2 \left(\frac{ds'(u)}{du} \right)^{-1} \\ &= -s'(u)^2 s''(u)^{-1}. \end{aligned} \tag{3.34}$$

The assumption that $\beta = s'(u)$ violates, as will be seen in the next chapter, the fact that β has no "microcanonical" meaning when $s(u) \neq s^{**}(u)$. But if we take the definitions

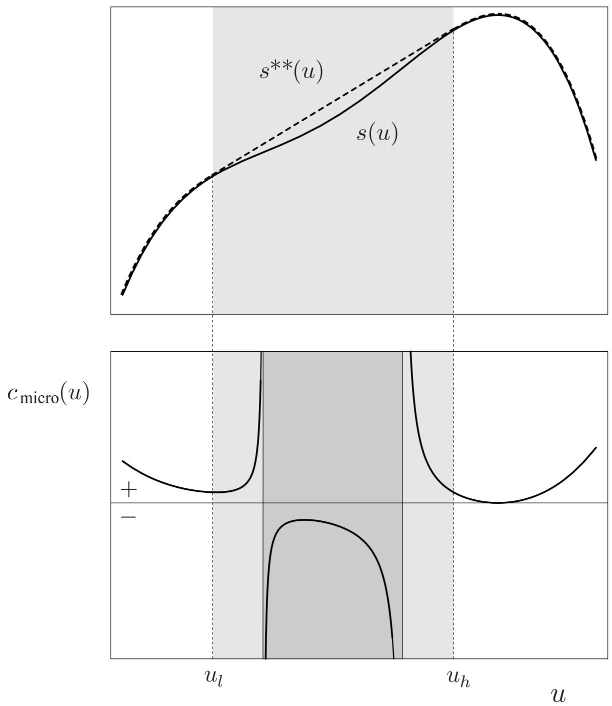

Figure 3.6: Illustration of the fact that negative values of the microcanonical heat capacity (dark shaded region) do not necessarily imply thermodynamic nonequivalence of ensembles (light shaded region).

of the two heat capacities above as they are given by equations (3.33) and (3.34), then the following results can be proved without difficulty:

- We have thermodynamic equivalence of ensembles at u if there exists β such that $c_{\text{can}}(\beta) = c_{\text{micro}}(u) < \infty$.

- Conversely, we have thermodynamic nonequivalence of ensembles at u if there exists no β for which $c_{\text{can}}(\beta) = c_{\text{micro}}(u) < \infty$.

In terms of the microcanonical specific heat only, we also have the following:

- If $c_{\text{micro}}(u) < 0$, then we have thermodynamic nonequivalence of ensembles at u.

The proofs of all these results follow simply from the fact that $\varphi(\beta)$ is always concave, which means that we must have $\varphi''(\beta) < 0$ for all β and thus $c_{\text{can}}(\beta) > 0$ for all β. The microcanonical entropy needs not be always concave, however, and the presence of a "kink" in the graph of $s'(u)$ must imply that $s''(u) > 0$ for some values of the mean energy, and thus that $c_{\text{micro}}(u) < 0$ for these values. In short, the canonical ensemble has no room for negative heat capacities, so to speak, but the microcanonical ensemble does.

In closing this section, let us note that the above thermodynamic definitions of nonequivalent ensembles based on the comparison of the microcanonical and canonical heat capacities cannot be applied at points of first-order phase transitions, either microcanonical or canonical, because such transitions lead the second derivatives of $s(u)$ and $\varphi(\beta)$, respectively, to diverge. Notice also that the occurrence of a negative heat capacity in the microcanonical ensemble provides only a *sufficient* condition for characterizing the microcanonical and canonical ensembles as being thermodynamically nonequivalent, not a *necessary* condition. It is not too difficult to find an example of a nonconcave entropy function $s(u)$ for which we have $c_{\text{micro}}(u) > 0$ at u, but $s^{**}(u) \neq s(u)$ (see Figure 3.6). Hence, although we may observe a positive value of the microcanonical heat capacity for some value u of the mean energy, we may be confronted with a case of thermodynamic nonequivalence of ensembles at u in the sense that the entropy at u may not be the Legendre-Fenchel transform of the free energy.

SUMMARY OF CHAPTER

- **Legendre-Fenchel inversion result:** The application of a Legendre-Fenchel transform to a free energy function yields, in general, the minimal convex envelope or convex hull of the associate rate function, not the rate function itself. For the rate function $I(a)$, as for example, the minimal convex envelope of $I(a)$ is defined as

$$\begin{aligned} I^{**}(a) &= \sup_k \{ka - \lambda(k)\} \\ &= \sup_k \{ka - \sup_b \{kb - I(b)\}\} \\ &= \sup_k \inf_b \{k(a-b) + I(b)\}. \end{aligned} \tag{3.35}$$

Thus $I^{**}(a) \leq I(a)$ with equality if and only if $I(a)$ is convex at a. In the context of the microcanonical entropy function, the same holds modulo some changes in sign; namely, the minimal concave envelope or concave hull of $s(u)$ is defined as

$$\begin{aligned} s^{**}(u) &= \inf_\beta \{\beta u - \varphi(\beta)\} \\ &= \inf_\beta \{\beta u - \inf_v \{\beta v - s(v)\}\} \\ &= \inf_\beta \sup_v \{\beta(u-v) + s(v)\}. \end{aligned} \tag{3.36}$$

Here we have $s^{**}(u) \geq s(u)$ with equality if and only if $s(u)$ is concave at u.

- **Thermodynamic equivalence of ensembles at u:** The microcanonical and canonical ensembles are thermodynamically equivalent at the mean energy value u if $s^{**}(u) = s(u)$, i.e., if $s(u)$ is concave at u. In case of thermodynamic equivalence, the value of $s(u)$ at u can be computed as the Legendre-Fenchel transform of the free energy $\varphi(\beta)$, i.e.,
$$s(u) = \inf_{\beta}\{\beta u - \varphi(\beta)\}. \tag{3.37}$$
In this case where $s(u)$ is strictly concave at u, then the Legendre-Fenchel transform actually reduces to the Legendre transform of $\varphi(\beta)$ given by
$$s(u) = \beta(u)u - \varphi(\beta(u)), \tag{3.38}$$
where $\beta(u) = s'(u)$.

- **Thermodynamic nonequivalence of ensembles at u:** The microcanonical and canonical ensembles are said to be thermodynamically nonequivalent at the mean energy value u if $s^{**}(u) > s(u)$, i.e., if $s(u)$ is nonconcave at u. In case of thermodynamic nonequivalence, $s(u)$ cannot be computed as the Legendre-Fenchel transform of $\varphi(\beta)$.

- **Global thermodynamic equivalence of ensembles and complete differentiability of the free energy:** If the free energy $\varphi(\beta)$ is everywhere differentiable (absence of first-order phase transitions), then the function $s(u)$ is everywhere concave on its domain of definition, which implies that we have equivalence of ensembles for all u. The converse statement does not hold since entropy functions which are concave but in a non-strict way are also characterized by non-differentiable free energies.

- **Local thermodynamic equivalence of ensembles and local differentiability of the free energy:** If $\varphi(\beta)$ is differentiable at β, then $s(u) = s^{**}(u)$ for $u = u_\beta$, where $u_\beta = \varphi'(\beta)$. (This result was not stated in the text, but it directly follows from the properties of Legendre transforms; see equation (3.38) above.)

- **Negativity of heat capacity:** A sufficient but non-necessary condition for having thermodynamic nonequivalence of ensembles at the mean energy value u is that $s''(u) > 0$. In terms of the microcanonical heat capacity
$$c_{\text{micro}}(u) = -s'(u)^2 s''(u)^{-1} \tag{3.39}$$

and the canonical heat capacity

$$c_{\text{can}}(\beta) = -\beta^2 \varphi''(\beta), \tag{3.40}$$

this condition translates into the following: the microcanonical and canonical ensembles are thermodynamically nonequivalent at u if $c_{\text{micro}}(u) \neq c_{\text{can}}(\beta)$ for all β.

SUPPLEMENTARY MATERIAL: MAXWELL'S CONSTRUCTION

A further characterization of the critical inverse temperature β_c, which signals the onset of a first-order phase transition in the canonical ensemble, can be given in terms of the three solutions of the equation $s'(u) = \beta_c$. We assume again that $s(u)$ has a single interval of nonconcavity. In Figure 3.7, the right, u_l is the smallest of these solutions and u_h the largest. We denote by u_m the intermediate solution of $s'(u) = \beta_c$. Because

$$s(u_l) = s^{**}(u_l), \qquad s(u_h) = s^{**}(u_h), \tag{3.41}$$

and

$$s^{**}(u_h) - s^{**}(u_l) = \beta_c(u_h - u_l), \tag{3.42}$$

it follows that

$$\int_{u_l}^{u_h} [\beta_c - s'(u)]\, du = \beta_c(u_h - u_l) - [s(u_h) - s(u_l)] = 0. \tag{3.43}$$

Rewriting this integral in terms of u_m, we see that

$$\int_{u_l}^{u_m} [\beta_c - s'(u)]\, du = \int_{u_m}^{u_h} [s'(u) - \beta_c]\, du = A. \tag{3.44}$$

This equation expresses the equal-area property of β_c, first observed by Maxwell (1875) [see also Huang (1987)].

NOTES AND REMARKS

References for Thermodynamic Nonequivalence of Ensembles

The idea of thermodynamic nonequivalence of ensembles originally made its way in physics not as a local concept defined for given values of the mean energy, as presented here, but as a global concept which applied whenever the graph of the microcanonical entropy $s(u)$

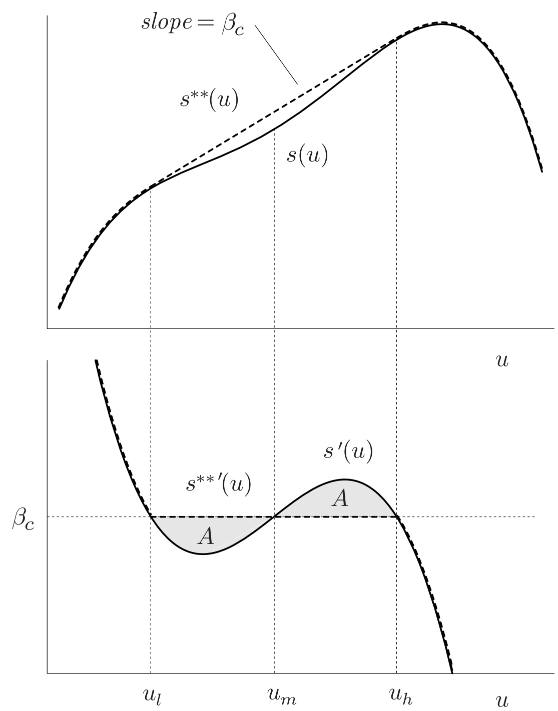

Figure 3.7: Maxwell's equal-area construction used to determine the value of β_c. See text for explanation.

showed a "kink" or, equivalently, whenever the graph of $s'(u)$ showed a "back-bending" shape as in Figure 3.7. This approach to the problem of nonequivalent ensembles has been propounded by a number of people, including Lynden-Bell and Wood (1968), who were among the first to observe such kinks in the entropy of certain gravitational many-body systems, and to relate them to negative values of the microcanonical heat capacity (see Lynden-Bell (1999) for a historical account). Similar theoretical observations have also been reported in similar contexts by Thirring (1970), Hertel and Thirring (1971), Gross (1997, 2001), and Chavanis and Ispolatov (2002) among others. For examples of lattice-spin systems displaying negative heat capacities, see Kiessling and Lebowitz (1997), Dauxois et al. (2000), Ispolatov and Cohen (2000), Antoni et al. (2002), Borges and Tsallis (2002), Barré et al. (2001), and Barré (2002).

For a recent survey of the subject of thermodynamic nonequivalence of ensembles and negative heat capacities, the reader is invited to consult Gross (1997, 2001), as well as the comprehensive collection of papers edited by Dauxois et al. (2002); both sources cover a wide range of physical models for which nonconcave anomalies of the microcanonical entropy have been observed, and contain much information about the "physics" of these models which will not be discussed in this thesis.

The reader will probably find it interesting to learn finally that various research groups have reported experimental measurements of negative heat capacities in "small" systems; see Schmidt et al. (2001), and Gobet et al. (2002) for experiments related to the melting of atom clusters, and D'Agostino and collaborators (2000, 2002) for results related to nuclear multifragmentation experiments.

Sources for Local Thermodynamic Nonequivalence of Ensembles

The rigorous approach to nonequivalent ensembles presented here, which explicitly focuses on the properties of Legendre-Fenchel transforms and on the local properties of $s(u)$, is due for the most part to Ellis et al. (2000) and Eyink and Spohn (1993) [see also Thirring (2002, Part II: §2.3, 2.4)]. Not all the results of this chapter are taken from these authors, however: the translation of the condition $s(u) \neq s^{**}(u)$ into an equivalent condition involving the microcanonical and canonical heat capacities is original to this thesis, as is the realization that the occurrence of a negative values for $c_{\text{micro}}(u)$ does not provide a necessary criterion for saying that the microcanonical and canonical ensembles are thermodynamically nonequivalent.

Allusions to Nonequivalent Ensembles and Phase Transitions

The idea that thermodynamic nonequivalent ensembles might be related to first-order phase transitions in the canonical ensemble seems to have been floating in the minds of physicists for some time now [see, e.g., Lynden-Bell (1968, 1999), Thirring (1970), and Gross (1997, 2001)]. The equal-area property mentioned above was in fact discovered by Maxwell in the course of his studies on van der Waal's equation and the equilibrium coexistence of different phases of matter (e.g., liquid-gas phases) which are usually associated with first-order transitions. It is important to note, however, that in Maxwell's case, the phase transition phenomenon is related not to a "back-bending" behavior of the graph of $s(u)$, but to a "back-bending" behavior of pressure-volume curves; see Huang (1987) and Griffiths (1967).

Concerning the microcanonical and canonical ensembles, most of the mentions found in the literature which try to relate the nonequivalence of these two ensembles to first-order phase transitions are vague, when they are explicit, and none of them express the actual rigorous relationship between the two phenomena which can only emerge by looking at the properties of Legendre-Fenchel transforms. The idea most recurrently found to explain first-order phase transitions, basically, is that, since a negative heat capacity system is inherently unstable (its gets colder by acquiring energy and gets hotter by giving energy), all values of the mean energy for which $c_{\text{micro}}(u) < 0$ must be avoided or "jumped over" by the system when its mean energy is not strictly kept constant as in canonical ensemble; see Thirring (1970), Lynden-Bell (1999). From the physical point of view, this is a very appealing explanation, but we have seen in this chapter that it cannot be entirely correct because the range of mean energy over which the microcanonical and canonical ensembles can be seen to be thermodynamically nonequivalent is, in general, larger than the interval of mean energy for which $c_{\text{micro}}(u) < 0$. We shall see in the next chapter that a more satisfying reason explaining why there is a phase transition is that all mean energy values u such that $s(u) \neq s^{**}(u)$ are not realized in the canonical ensemble as equilibrium mean energies.

To close this subsection, we would like to mention the works of Varchenko (1990) and Aicardi (2001) which take to a higher mathematical level the idea that thermodynamic singularities (e.g., phase transitions) can be understood by properly studying thermodynamic functions and their convex or concave hulls.

4

Equivalence and Nonequivalence of Ensembles: Macrostate Level

Our aim in this chapter is to show how the thermodynamic nonequivalence of the microcanonical and canonical ensembles reflects a deeper level of nonequivalence of these two ensembles that takes place at the level of the equilibrium values of macrostates. At this level, the natural questions to consider for comparing the two ensembles are the following. For every β and every m_β in the set $\mathscr{E}_\beta$ of canonical equilibrium macrostates, does there exist a value of u such that m_β lies in the set $\mathscr{E}^u$ of microcanonical equilibrium macrostates? Conversely, for every u and every $m^u \in \mathscr{E}^u$, does there exist a value of β such that $m^u \in \mathscr{E}_\beta$? In trying to relate the macrostate level of equivalence and nonequivalence of ensembles with the thermodynamic level of equivalence and nonequivalence, we may also ask whether there are thermodynamic conditions expressed in terms of properties of $s(u)$ or $\varphi(\beta)$ which ensure that there is a correspondence or a lack of correspondence between the members of $\mathscr{E}^u$ and those of $\mathscr{E}_\beta$. In particular, does equivalence of ensembles at the thermodynamic level implies equivalence of ensembles at the level of equilibrium macrostates? We shall see that the answers to all these questions depend, lo and behold, on the concavity properties of $s(u)$.

4.1. PRELIMINARY RESULTS FROM CONVEX ANALYSIS

We begin this chapter with yet another short incursion into the world of mathematics to acquaint ourselves, this time, with two basic results of convex analysis which will come to play a crucial role in the many proofs to come. The first of these results states that

$$s(v) \leq s(u) + \beta(v - u) \tag{4.1}$$

for all v if and only if $s(u) = s^{**}(u)$ and $\beta = s'(u)$. From a geometric point of view, this result only states that $s(u)$ is concave at u if and only if there exists a supporting line to the graph of $s(u)$ at $(u, s(u))$ which does not go under the graph of $s(u)$; see Figure 4.1(a). The second result sharpens the first one to cover cases of strict concavity; it states that $s(u)$ is strictly concave at u if and only if

$$s(v) < s(u) + \beta(v - u) \tag{4.2}$$

for all $v \neq u$ and $\beta = s'(u)$. That is, $s(u)$ is strictly concave at u if and only if the supporting line to the graph of $s(u)$ touches this graph only at $(u, s(u))$; see Figure 4.1(b).

We shall not go through the full, rigorous proofs of these results as they can be found in many textbooks on convex analysis [see, e.g., Rockafellar (1970)]. The geometrical interpretation of these results depicted in Figure 4.1 conveys, at any rate, a strong enough sense of their veracity. For instance, it is plainly seen from the plot of Figure 4.1(a) that, in the case where $s(u)$ is strictly concave at u, the only line passing through the point $(u, s(u))$ which does not pass through any other points of the graph of $s(u)$ is the tangent of $s(u)$ which satisfies the equation

$$f(v) = s'(u)(v - u) + s(u). \tag{4.3}$$

As a result, we must have $f(v) > s(v)$ for all $v \neq u$, as stated in (4.2). If $s(u)$ is non-strictly concave at u, however, then more than one point will touch the supporting line of $s(u)$ at u [see the endpoints of the interval delimited by the grey region in Figure 4.1(a)]. Finally, the plot of Figure 4.1(b) shows that if $s(u)$ is nonconcave at u, i.e., if $s(u) \neq s^{**}(u)$, then no line can be traced over the graph of $s(u)$ which passes only through the point $(u, s(u))$.

4.2. REALIZABILITY OF THE CANONICAL EQUILIBRIUM MEAN ENERGY

The first set of equivalence and nonequivalence of ensembles results that we obtain in this section express the necessary and sufficient conditions which guarantee that a given mean energy value u is realized (we also say “stabilized”) in the canonical ensemble as an equilibrium value of the mean energy macrostate $u_n(x^n)$. We shall see from these results that the canonical ensemble may realize at equilibrium less than the entire range of definition of $u_n(x^n)$, which is the range of definition of the microcanonical ensemble, and that multiple values of $u_n(x^n)$ may be realized at once at equilibrium for a given value β of the inverse temperature. These peculiarities of the canonical ensemble are, not surprisingly, related to the appearance of discontinuous (first-order) phase transitions in that ensemble.

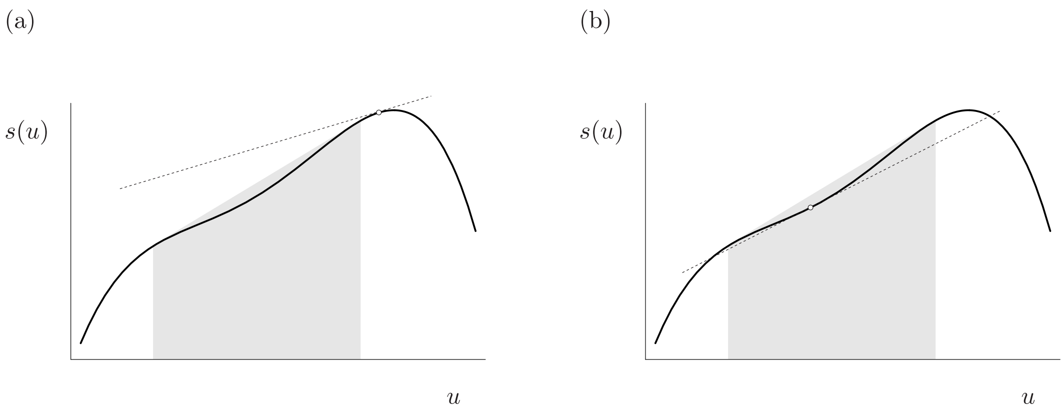

Figure 4.1: (a) Concavity point of the microcanonical entropy function $s(u)$ where $s(u) = s^{**}(u)$. (b) Nonconcavity point where $s(u) \neq s^{**}(u)$.

Stable Mean Energy Values

The mean energy value u is realized as an equilibrium value of the mean energy macrostate $u_n(x^n)$ in the canonical ensemble at $\beta = s'(u)$ if and only if $s(u) = s^{**}(u)$ at u. In terms of the canonical mean energy rate function, this can be expressed equivalently by saying that $I_\beta(u) = 0$ for $\beta = s'(u)$ if and only if $s(u) = s^{**}(u)$. The mean energy values realizing the global minimum of $I_\beta(u)$ are denoted, as before, by u_β, and are referred to as **stable** mean energy values.

To prove this result, we first prove that $s(u) = s^{**}(u)$ implies $I_\beta(u) = 0$ for $\beta = s'(u)$. The assumption that $s(u) = s^{**}(u)$ can be translated into the concavity result of the previous section as follows:

$$\beta u - s(u) \leq \beta v - s(v) \tag{4.4}$$

for all v and for $\beta = s'(u)$. The above inequality shows that u is a global minimizer of the quantity $\beta v - s(v)$ which is just another way to say, as was shown in Section 2.3, that $I_\beta(u) = 0$. To complete the proof, we proceed to prove the converse result. Assuming that $s(u) \neq s^{**}(u)$, we obtain

$$s(u) < s^{**}(u) = \inf_\gamma \{\gamma u - \varphi(\gamma)\} \tag{4.5}$$

where we have used the definition of the concave hull of $s(u)$. Going a step further, we can write

$$s(u) < \inf_\gamma \{\gamma u - \varphi(\gamma)\} \leq \beta u - \varphi(\beta) \tag{4.6}$$

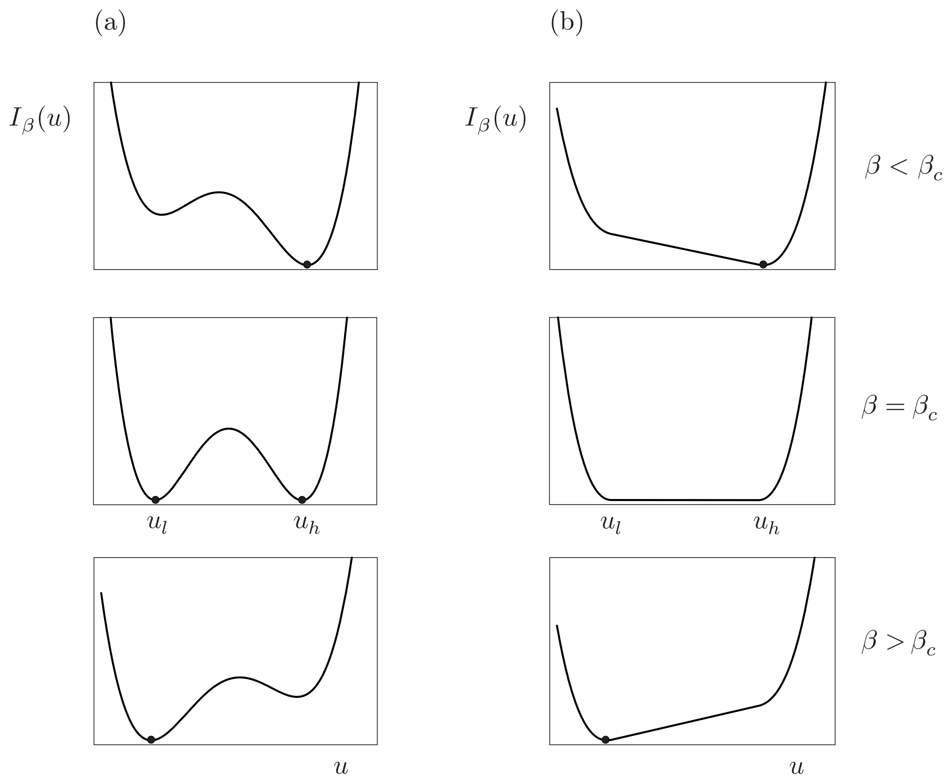

Figure 4.2: Form of the canonical rate function $I_\beta(u)$ and canonical equilibrium values of the mean energy for various values of β. (a) Case where $s(u)$ is strictly concave over (u_l, u_h). (b) Case where $s(u)$ is non-strictly concave over (u_l, u_h).

for all β or, equivalently,

$$\beta u - s(u) > \varphi(\beta) \tag{4.7}$$

for all β. This last inequality, together with the macrostate representation of the free energy function $\varphi(\beta)$, show that u cannot be a minimizer of $I_\beta(u)$ for any β. We conclude therefore that the value u is nowhere realized in the canonical ensemble as an equilibrium value of the mean energy.

To illustrate the connection between the minima and zeros of $I_\beta(u)$ and the concavity points of the microcanonical entropy $s(u)$, we present in Figure 4.2 different plots of $I_\beta(u)$ corresponding to various values of β and two different forms of $s(u)$. The plots on the left of this figure were obtained using the now well-studied nonconcave entropy function $s(u)$ displayed in the many figures of the previous chapter (also displayed in Figure 4.1), whereas those of on the right were obtained by using a non-strictly concave entropy function which coincides with the concave hull of entropy function used for the plots in (a). Among the

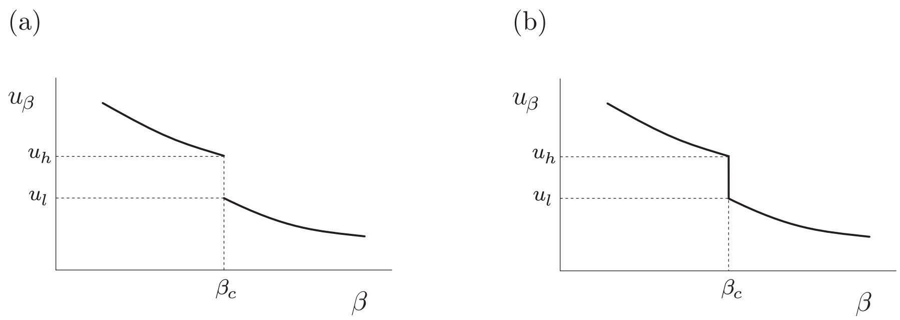

Figure 4.3: Two "flavors" of first-order or discontinuous canonical phase transitions illustrated at the level of u_β: (a) normal; (b) degenerate.

many properties of $I_\beta(u)$ illustrated by these plots, we see that

- The canonical equilibrium value of the mean energy never enters the region of non-concavity of $s(u)$, in agreement with the result proved above.
- The mean energy value u is the unique minima of $I_\beta(u)$ for $\beta = s'(u)$ when $s(u)$ is strictly concave at u.
- In the limit where $\beta \nearrow \beta_c$, i.e., where β moves to β_c from above, $u_\beta = u_h$ while for $\beta \searrow \beta_c$, we have $u_\beta = u_l$.
- At $\beta_c = s'(u_l) = s'(u_h)$, the two non-strictly concave endpoints u_l and u_h of the nonconcavity interval of $s(u)$ are realized canonically as the equilibrium values of $u_n(x^n)$.
- For an entropy function having a affine part over the open interval (u_l, u_h), all mean energy values $u \in (u_l, u_h)$ are realized canonically as equilibrium values of $u_n(x^n)$ for $\beta_c = s'(u_l) = s'(u_h)$.

The last point is important: it shows that when $s(u)$ is affine over some interval with slope β_c, the canonical ensemble at inverse temperature β_c does not "skip over" this interval like it does when $s(u)$ is nonconcave. This does not mean all the same that the β-behavior of $I_\beta(u)$ when $s(u)$ is affine should not be considered as leading to a true first-order phase transition. The point is that even though the equilibrium value of the mean energy u_β taken as a function of the inverse temperature does not jump discontinuously at β_c when $s(u)$ is affine, u_β still has an infinite slope at β_c (see Figure 4.3). In this sense, we may say that the phase transition associated with the affine part of $s(u)$ is a **degenerate** first-order transition. Note in fact that from a practical and less mathematical point of view, it is to be doubted that physicists doing real experiments in laboratories can actually measure a whole *continuum* of mean energy value for a perfectly *fixed* and controlled value of the inverse temperature. If only for this reason, a definition of first-order phase transitions based on the presence of an infinite value of $\partial u_\beta/\partial\beta$ appears to be more sensible than a definition based only on the presence of a discontinuous point in the graph of u_β versus β. The former definition includes the latter, in addition to be totally equivalent to the definition based on the non-differentiable points of $\varphi(\beta)$.

Metastable and Unstable Nonequivalent Mean Energy Values

We have seen in Section 2.3 that the equilibrium values of the mean energy, u_β, in the canonical ensemble with inverse temperature β must satisfy the equation

$$\frac{\partial I_\beta(u)}{\partial u} = 0 \qquad \text{or} \qquad s'(u) = \beta, \tag{4.8}$$

assuming that $I_\beta(u)$ or $s(u)$ is once differentiable at u. What we would like to stress now is that not all the smooth critical points of $I_\beta(u)$ determined by the above equations need to be realized canonically as equilibrium values of the mean energy: some of these critical points may actually correspond to local minima or local maxima of $I_\beta(u)$, in which case $P_\beta(u_n \in du) \to 0$ exponentially fast as $n \to \infty$. To determine the precise nature of these canonical "non-equilibrium" critical points, we look at the sign of the second u-derivative of $I_\beta(u)$: if $\partial^2 I_\beta(u)/\partial u^2 > 0$, u is a minimum of $I_\beta(u)$; if $\partial^2 I_\beta(u)/\partial u^2 < 0$ is a maximum of $I_\beta(u)$; if $\partial^2 I_\beta(u)/\partial u^2 = 0$ or if $\partial^2 I_\beta(u)/\partial u^2$ does not exist, the test fails. In terms of the microcanonical heat capacity, we are thus led to state the following:

- If $s(u) \neq s^{**}(u)$ and $c_{\text{micro}}(u) > 0$, then u is a **metastable critical point** of $I_\beta(u)$ for $\beta = s'(u)$ in the sense that it is a local minima of $I_\beta(u)$ for that particular value of β. Such a point is also called a **metastable critical mean energy** of the canonical ensemble.

- If $s(u) \neq s^{**}(u)$ and $c_{\text{micro}}(u) < 0$, then u is a local maximum of $I_\beta(u)$ for $\beta = s'(u)$. We call such a point an **unstable critical point** of $I_\beta(u)$ or an **unstable critical mean energy** of the canonical ensemble.

To better understand and "visualize" these results, we provide in Figure 4.4 a number of plots which illustrate in the most complete way how the critical mean energy values satisfying the equation $s'(u) = \beta$ relate to the critical points of $I_\beta(u)$ depending on the choice of β. Similar plots found in Figure 4.5 also show how the whole picture of critical points of $I_\beta(u)$ changes when we consider entropy functions and canonical rate functions having **corners** or **cusps**, i.e., non-smooth, non-differentiable points which are associated with *microcanonical* first-order phase transitions. We leave the reader with the captions of these figures for the detail of their interpretations.

4.3. EQUIVALENCE AND NONEQUIVALENCE RESULTS FOR GENERAL MACROSTATES

We now proceed to extend the equivalence and nonequivalence results derived for the mean energy to general macrostates. As in Chapter 2, we consider macrostates $M_n(x^n)$ which conform to the following hypotheses:

- $P(M_n \in dm)$ satisfies a large deviation principle with respect to the uniform prior probability P. The entropy function quantifying the rate of decay in the large deviation principle is denoted by $s(m)$.
- There exists an energy representation (total) function $u(m)$ mapping the macrostate space $\mathscr{M}$ onto the energy space $\mathscr{U}$ either exactly or asymptotically as $n \to \infty$.

We recall that, under these assumptions, the equilibrium values of $M_n(x^n)$ in the microcanonical ensemble correspond to those $m \in \mathscr{M}$ which maximize the entropy function $s(m)$ subject to the constraint that $u(m) = u$; in symbols,

$$\mathscr{E}^u = \{m \in \mathscr{M} : s(m) \text{ is maximized with } u(m) = u\} = \{m^u\}. \tag{4.9}$$

In the canonical ensemble, the equilibria of $M_n(x^n)$ are given by minimizing the quantity $\beta u(m) - s(m)$ for a given value of β with no constraint at all, so that

$$\mathscr{E}_\beta = \{m \in \mathscr{M} : \beta u(m) - s(m) \text{ is minimized}\} = \{m_\beta\}. \tag{4.10}$$

Realizability of Canonical Equilibria

The first result that we want to prove relating the equilibria of the microcanonical ensemble to those of the canonical ensemble reads in plain words as follows: *The canonical equilibrium values of the macrostate M_n obtained for a given value β of the inverse temperature are always realized somewhere in the microcanonical ensemble either for a fixed value u of the mean energy or for multiple values of the mean energy*. In more mathematical terms, what we want to prove

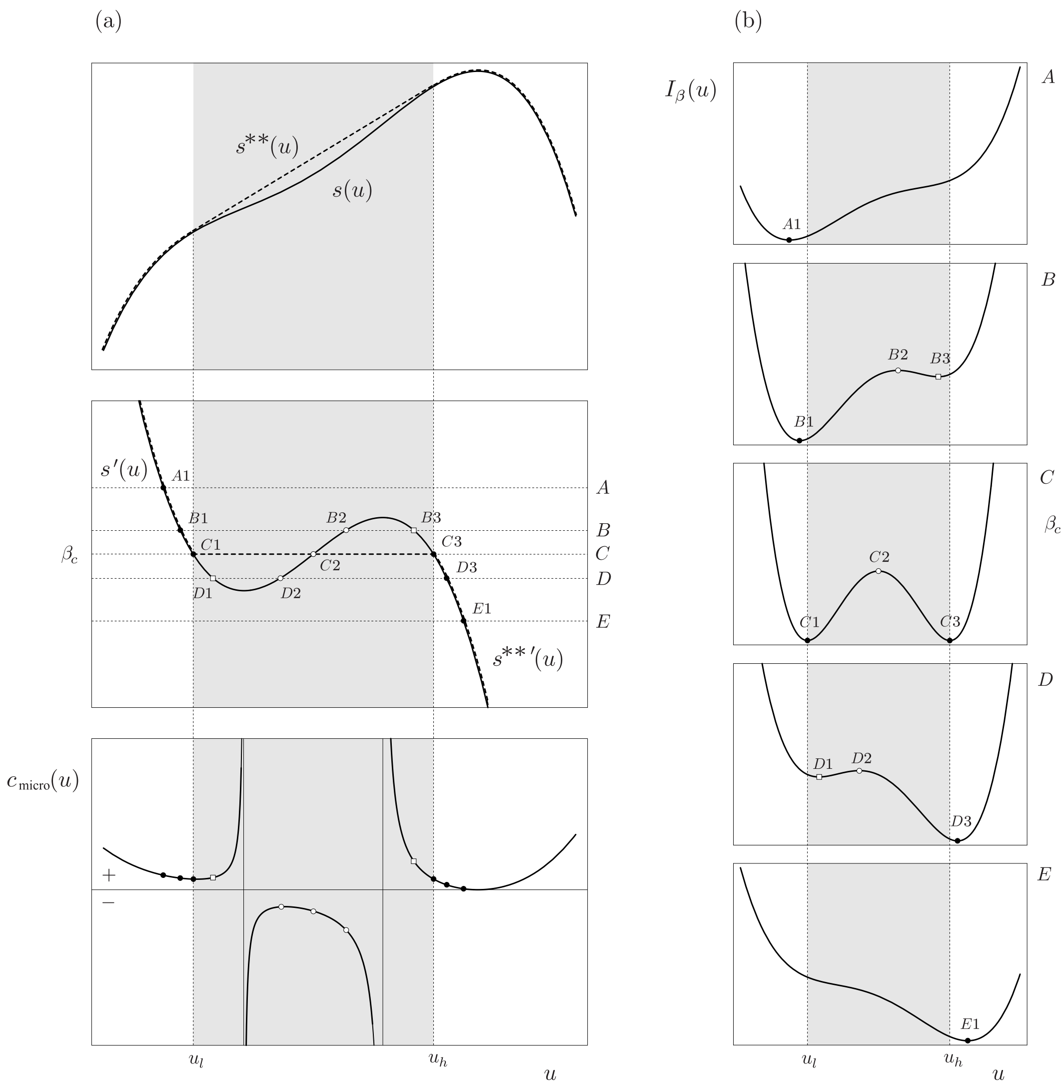

Figure 4.4: (a) (Top) Microcanonical entropy function $s(u)$ together with its concave envelope $s^{**}(u)$. (Middle) First derivative of $s(u)$ and $s^{**}(u)$. (Bottom) Plot of the microcanonical heat capacity as a function of the mean energy. The positivity or negativity of this last quantity determines whether the mean energy value u is a minimum (global: •; local: □) or a maximum (◦) point of the canonical rate function $I_\beta(u)$, respectively. (b) Form of $I_\beta(u)$ for different values of β displayed in plot of $s'(u)$. The shaded areas in all of these plots highlight the region of ensembles nonequivalence.

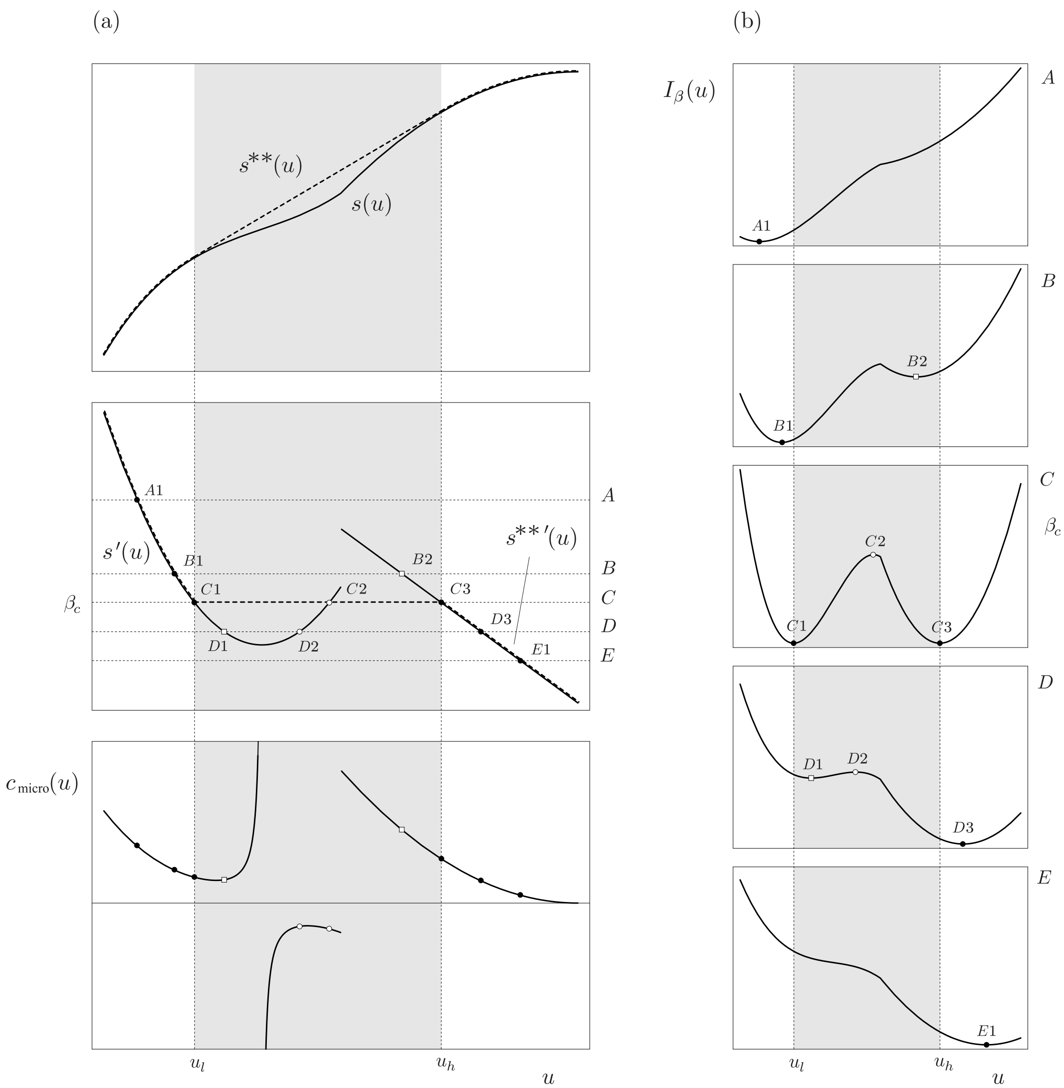

Figure 4.5: Same as in Figure 4.4, but now $s(u)$ has a "corner," i.e., a point u where $s'(u)$ jumps discontinuously. The corner of $s(u)$ translates into a corner of $I_\beta(u)$ as illustrated in the plots in (b).

is thus that

$$\mathscr{E}_\beta = \bigcup_{u \in u(\mathscr{E}_\beta)} \mathscr{E}^u, \tag{4.11}$$

where $u(\mathscr{E}_\beta)$ denotes the set of mean energy values u that can be written as $u(m)$ over all $m \in \mathscr{E}_\beta$, i.e., all mean energy values realized canonically.

To prove this result, we first show that

$$\mathscr{E}_\beta \subseteq \bigcup_{u \in u(\mathscr{E}_\beta)} \mathscr{E}^u. \tag{4.12}$$

Choose $m_\beta \in \mathscr{E}_\beta$ and denote the equilibrium value of the mean energy associated with this canonical equilibrium macrostate value by $u_\beta = u(m_\beta)$. Since m_β is by definition a global minimizer of $\beta u(m) - s(m)$, we have

$$\beta u(m_\beta) - s(m_\beta) \leq \beta u(m) - s(m) \tag{4.13}$$

for all $m \in \mathscr{M}$ with equality if and only if $m \in \mathscr{E}_\beta$. If we restrict the values of m in the above inequality to be such that their associated mean energy is equal to u_β, then the inequality above reduces to $s(m_\beta) \geq s(m)$ for all m such that $u(m) = u_\beta$. This shows that, with respect to the manifold of macrostate values $\{m \in \mathscr{M} : u(m) = u_\beta\}$ having a fixed mean energy value u_β, m_β is a global maximizer of $s(m)$; in other words, $m_\beta \in \mathscr{E}^{u_\beta}$. From this result, and the fact that there may be many mean energy values realized in the set $\mathscr{E}_\beta$, we arrive at (4.12).

To prove finally that the set containment (4.12) is actually an equality, let us prove the opposite containment relationship

$$\bigcup_{u \in u(\mathscr{E}_\beta)} \mathscr{E}^u \subseteq \mathscr{E}_\beta. \tag{4.14}$$

Choose $m^{u_\beta} \in \mathscr{E}^{u_\beta}$. By definition of the microcanonical equilibria, m^{u_β} maximizes $s(m)$ subject to the constraint $u(m) = u_\beta$, which means in symbols that $s(m^{u_\beta}) \geq s(m)$ for all m such that $u(m) = u_\beta$. But we just saw in the previous paragraph that $s(m_\beta) \geq s(m)$ for all macrostate value m such that $u(m) = u_\beta$. In order to avoid a contradiction, we must then have $s(m_\beta) = s(m^{u_\beta})$, and, as a result,

$$\begin{aligned} \inf_m \{\beta u(m) - s(m)\} &= \beta u(m_\beta) - s(m_\beta) \\ &= \beta u(m^{u_\beta}) - s(m^{u_\beta}), \end{aligned} \tag{4.15}$$

thereby proving that $m^{u_\beta} \in \mathscr{E}_\beta$. Repeating the argument for all $u_\beta \in u(\mathscr{E}_\beta)$, we arrive at the claimed result.

It may help in trying to understand this proof to recall that there can be one or more canonical equilibrium values of the mean energy associated with any given value β of the inverse temperature (see previous section). In such a case, it is naturally to be expected that the canonical ensemble may be decomposed into many microcanonical ensembles, each one realizing a single equilibrium mean energy value of the canonical ensemble. That such an expectation holds true, and holds furthermore at the level of the equilibrium values of M_n, is what we just proved.

Full Equivalence of Ensembles

If $s(u)$ *is strictly concave at* u*, then* $\mathscr{E}^u = \mathscr{E}_\beta$ *for* $\beta = s'(u)$. This is again intuitively expected considering that we have seen in the previous section that if $s(u)$ is strictly concave at u, then u is the unique equilibrium value of the mean energy realized in the canonical ensemble for $\beta = s'(u)$. Yet, since we have provided no detailed proof of this latter result, let us now take the time to concoct a satisfying proof of the result now in hand.

To begin, let us suppose that $s(u) = s^{**}(u)$. Following the concavity result of Section 4.1, this implies that

$$\beta u - s(u) \le \beta v - s(v) \tag{4.16}$$

for all v and $\beta = s'(u)$. Using the thermodynamic and macrostate representations of $\varphi(\beta)$ derived in Chapter 2, we also see that

$$\begin{aligned} \beta u - s(u) &\le \inf_v \{\beta v - s(v)\} \\ &= \varphi(\beta) \\ &= \inf_m \{\beta u(m) - s(m)\}. \end{aligned} \tag{4.17}$$

Now choose $m^u \in \mathscr{E}^u$. By definition of the microcanonical set $\mathscr{E}^u$, we must have $u(m^u) = u$ and $s(m^u) = s(u)$, and so

$$\beta u(m^u) - s(m^u) \le \inf_m \{\beta u(m) - s(m)\}. \tag{4.18}$$

We deduce from this inequality that m^u minimizes $\beta u(m) - s(m)$, i.e., that $m^u \in \mathscr{E}_\beta$. Since m^u is an arbitrary element of $\mathscr{E}^u$, it follows that $\mathscr{E}^u \subseteq \mathscr{E}_\beta$ with $\beta = s'(u)$. Thus, we have shown that if $s(u)$ is concave, then $\mathscr{E}^u \subseteq \mathscr{E}_\beta$. The converse of this statement is also true, for if we choose $m_\beta^u \in \mathscr{E}^u \cap \mathscr{E}_\beta$, then

$$\beta u(m_\beta^u) - s(m_\beta^u) = \inf_m \{\beta u(m) - s(m)\} = \varphi(\beta) \tag{4.19}$$

as a result of the fact that $m_\beta^u \in \mathscr{E}_\beta$ and

$$\beta u(m_\beta^u) - s(m_\beta^u) = \beta u - s(u) \tag{4.20}$$

as a result of the fact that $m_\beta^u \in \mathscr{E}^u$. Combining these two results, we obtain

$$\begin{aligned} \beta u - s(u) &= \varphi(\beta) \\ &= \inf_v \{\beta v - s(v)\} \\ &\leq \beta w - s(w) \end{aligned} \tag{4.21}$$

for all w, as claimed.

We now use this last result to complete the proof that $\mathscr{E}^u = \mathscr{E}_\beta$ by proving by contradiction that $\mathscr{E}^u$ cannot be a proper subset of $\mathscr{E}_\beta$ if $s(u)$ is strictly concave at u. We suppose as a working hypothesis that $\mathscr{E}^u \subset \mathscr{E}_\beta$. This hypothesis together with the result of equation (4.11), which expresses the realizability of the canonical equilibria, imply that the remaining part of $\mathscr{E}_\beta$ not covered by $\mathscr{E}^u$ must be covered by one or more other microcanonical sets. That is to say, there must exists $u' \neq u$ such that $\mathscr{E}^{u'} \subseteq \mathscr{E}_\beta$. This implies, as we have just seen, that $s(u)$ must be concave at u' so that

$$s(w) \leq s(u') + \beta(w - u') \tag{4.22}$$

for all w. As a specific case of this inequality, let $w = u$:

$$s(u) \leq s(u') + \beta(u - u'). \tag{4.23}$$

And now comes the contradiction: if $s(u)$ is strictly at u, then

$$s(v) < s(u) + \beta(v - u). \tag{4.24}$$

for all v, but if we put $v = u'$ in the above inequality and insert the result in (4.23), we obtain

$$s(u) \leq s(u') + \beta(u - u')$$

$$
\begin{aligned}
&< s(u) + \beta(u' - u) + \beta(u - u') \\
&= s(u),
\end{aligned} \tag{4.25}
$$

i.e., $s(u) < s(u)$ (!). To avoid this contradicting result, we have no other choice of course but to conclude that $\mathscr{E}^u = \mathscr{E}_\beta$ for $\beta = s'(u)$, thereby completing the proof. Note that this proves at the same time that $\mathscr{E}_{\beta=s'(u)}$ cannot realize more than one equilibrium value of the mean energy when $s(u)$ is strictly concave at u. In other words, the set $u(\mathscr{E}_\beta)$ must be a singleton set for $\beta = s'(u)$ when $s(u)$ is strictly concave at u.

Partial Equivalence of Ensembles

The next result is a variation on the theme of concavity: it states that *if $s(u)$ is concave at u but not strictly concave, then $\mathscr{E}^u \subsetneq \mathscr{E}_\beta$ for $\beta = s'(u)$.* Combining this result with the previous, we thus have that thermodynamic equivalence of ensembles at u implies either full or partial equivalence of ensembles at the level of the equilibrium macrostates for that particular value of the mean energy.

To prove the strict containment of $\mathscr{E}^u$ in $\mathscr{E}_{\beta=s'(u)}$, we simply need to use an intermediate result derived in the previous proof which stated that $s(u)$ is concave at u if and only if $\mathscr{E}^u \subseteq \mathscr{E}_\beta$ where $\beta = s'(u)$. This result applies here because $s(u)$ is assumed to be concave, though in a non-strict way. As a result, we must have

$$
s(v) \leq s(u) + \beta(v - u) \tag{4.26}
$$

for all v and $\beta = s'(u)$ as well as $\mathscr{E}^u \subseteq \mathscr{E}_{\beta=s'(u)}$. Now, since $s(u)$ is *non-strictly* concave at u, there must be another point of s different from $(u, s(u))$ which touches the supporting line of $s(u)$; that is, there must exists $u' \neq u$ such that

$$
s(u') = s(u) + \beta(u' - u). \tag{4.27}
$$

The combination of (4.26) and (4.27) thus yields

$$
s(v) \leq s(u) + \beta(v - u) = s(u') + \beta(v - u') \tag{4.28}
$$

for all v. This inequality only demonstrates that $s(u')$ is concave at u', which means that we must have $\mathscr{E}^{u'} \subseteq \mathscr{E}_\beta$. Since $u' \neq u$, we finally conclude that $\mathscr{E}^u$ must be but a proper subset of $\mathscr{E}_{\beta=s'(u)}$ since it is not the only set contained in $\mathscr{E}_{\beta=s'(u)}$. This is nothing unexpected considering that we have seen before that many values of the mean energy minimize the canonical rate function $I_\beta(u)$ for $\beta = s'(u)$ when $s(u)$ is non-strictly concave at u.

Nonequivalence of Ensembles

The final case that we have to consider to complete our comparison of the microcanonical and canonical sets of equilibrium macrostates is the case of a nonconcave entropy. For this case, the relationship between $\mathscr{E}^u$ and $\mathscr{E}_\beta$ may directly be guessed from our previous observation that u is nowhere realized in the canonical ensemble as an equilibrium value of the mean energy when $s(u)$ is nonconcave. Consequently, *if $s(u)$ is nonconcave at u*, we should have $\mathscr{E}^u \cap \mathscr{E}_\beta = \emptyset$ *for all* β.

Let us prove that is indeed a correct result. Assuming that $s(u)$ is nonconcave at u, we must have

$$s(u) < s^{**}(u) = \inf_\gamma \{\gamma u - \varphi(\gamma)\} \leq \beta u - \varphi(\beta) \tag{4.29}$$

for all β, as already stated in (4.5) and (4.6). Now, choose $m^u \in \mathscr{E}^u$ and any β. Since $u(m^u) = u$ and $s(m^u) = s(u)$, it follows from the above inequality that

$$\beta u(m^u) - s(m^u) > \varphi(\beta) = \inf_m \{\beta u(m) - s(m)\}. \tag{4.30}$$

This shows that m^u is not a minimizer of $\beta u(m) - s(m)$, and thus that $m^u \notin \mathscr{E}_\beta$. Since m^u is an arbitrary element of $\mathscr{E}^u$ and β is arbitrary, we conclude that $\mathscr{E}^u \cap \mathscr{E}_\beta = \emptyset$ for all β. Thus, *if there is thermodynamic nonequivalence of ensembles for some value of u, then the microcanonical equilibrium macrostates corresponding to that u are nowhere realized within the canonical ensemble.*

4.4. TWO REPRESENTATIONS OF THE MICROCANONICAL ENTROPY

Leyvraz and Ruffo (2002) have proposed two different macrostate representations of the microcanonical entropy function which illustrate a subtlety involved in the calculation of this function when the microcanonical and canonical ensembles are nonequivalent. Although their work does not appeal directly to the formalism of large deviations [it basically exploits the properties of a steepest descent approximation of an integral leading to $s(u)$], it is easy at this point of our study to trace a path leading to their results which uses nothing more than the concepts and quantities that we have defined in this thesis.

The basic result at play, once again, is the basic relationship $\varphi(\beta) = s^*(\beta)$ which expresses the free energy function $\varphi(\beta)$ as the Legendre-Fenchel transform of the microcanonical entropy function $s(u)$. We have seen in the previous chapter that this relationship cannot be inverted to express $s(u)$ as the Legendre-Fenchel transform of $\varphi(\beta)$ when $s(u)$ is nonconcave, and that what the Legendre-Fenchel transform of $\varphi(\beta)$ yields in this case is the concave hull of $s(u)$, not $s(u)$ itself. Thus, if we attempt to apply the Legendre-Fenchel transform to the basic macrostate representation of $\varphi(\beta)$, as in

$$\inf_\beta\{\beta u - \varphi(\beta)\} = \inf_\beta\{\beta u - \inf_m\{\beta u(m) - s(m)\}, \tag{4.31}$$

then what is obtained, in general, is

$$s^{**}(u) = \inf_\beta \sup_m\{\beta[u - u(m)] + s(m)\}. \tag{4.32}$$

Only in the case where $s(u)$ is concave at u is this formula a valid macrostate representation formula for calculating $s(u)$.

Let us see now what happens if we interchange the order of the infimum and supremum operators above. Taking first the inf over β before the sup over m, we get

$$\inf_\beta\{\beta[u - u(m)] + s(m)\} = \begin{cases} -\infty & \text{if } u(m) \neq u \\ 0 + s(m) & \text{if } u(m) = u. \end{cases} \tag{4.33}$$

Next, we evaluate the sup over the macrostate values m to obtain

$$\begin{aligned} \sup_m \inf_\beta\{\beta[u - u(m)] + s(m)\} &= \sup_m \left\{ \begin{array}{ll} -\infty & \text{if } u(m) \neq u \\ 0 + s(m) & \text{if } u(m) = u \end{array} \right\} \\ &= \sup_{m:u(m)=u} s(m) \\ &= s(u) \end{aligned} \tag{4.34}$$

using the thermodynamic representation formula of $s(u)$ for the last line. What results from these few lines of calculations is a new representation formula for $s(u)$:

$$s(u) = \sup_m \inf_\beta\{\beta[u - u(m)] + s(m)\} \tag{4.35}$$

which holds true regardless of the form of $s(u)$. The asymmetry between equations (4.32) and (4.35) is on its own quite interesting, and definitively offers a new perspective on the phenomenon of nonequivalent ensembles. However, it must be noted that the above equation is not very useful for practical calculations because the infimum over β accounts for the energy constraint $u(m) = u$ in a trivial manner.

4.5. CONJECTURE ABOUT NONEQUIVALENT MACROSTATES

We end this chapter on an open note by formulating a conjecture concerning equilibrium macrostates which are realized at equilibrium in the microcanonical ensemble but not in the canonical ensemble.

We have seen in Section 4.2 that some of the mean energy values u satisfying the differential equation $\beta = s'(u)$ are not realized in the canonical equilibrium because they correspond to local minima or maxima of the canonical rate function $I_\beta(u)$ rather than to global minima of this quantity. We have also seen, in this context, that the precise nature of these metastable or unstable mean energies, as we called them, can be assessed by looking at the sign of the microcanonical heat capacity $c_{\text{micro}}(u)$. The question that we would like to address in this last section is whether something similar holds for general macrostates, namely: could it be that the nonequivalent microcanonical equilibrium macrostates, defined mathematically as those $m^u \in \mathscr{E}^u$ such that $m^u \notin \mathscr{E}_\beta$ for all β, are not realized canonically because such macrostates correspond to local maxima of $I_\beta(m)$ or to local minima (or saddle-points?) of this function depending on the value of $c_{\text{micro}}(u)$? As an answer to this question, we conjecture the following two points:

- If $m^u \notin \mathscr{E}_\beta$ for all β and $c_{\text{micro}}(u) > 0$, then m^u is a **metastable** macrostate of the canonical ensemble, i.e., it is a local but not global minimum of $I_\beta(m)$. Without resorting to the microcanonical heat capacity, we may say more generally that m^u is metastable if u itself is metastable with respect to $I_\beta(u)$.

- If $m^u \notin \mathscr{E}_\beta$ for all β and $c_{\text{micro}}(u) < 0$, then m^u is a **saddle point** of the canonical rate function $I_\beta(m)$. More generally, m^u is a saddle point of the canonical ensemble whenever u is a local maximum of $I_\beta(u)$.

We do not have at this stage a satisfying proof of the above results, but we are tempted to think that they are valid for a wide range of statistical mechanical models which basically satisfy a large deviation principle with rate function $s(m)$ and which have an energy representation function. (These are the two working hypotheses of this chapter.) To support our claim, we mention the many hints which have suggested to us the very answer presented above. First, we know from the theory of Lagrange multipliers that the microcanonical critical points of $s(m)$ constrained by the condition $u(m) = u$ must necessarily correspond to extremal points of the quantity $\beta u(m) - s(m)$ and, by extension, of $I_\beta(m)$. Trying to relate the nonequivalent macrostate values m^u to the critical points of the "canonical" quantity $\beta u(m) - s(m)$ is thus a sensible idea to begin with. Second, we know that the microcanonical macrostates m^u globally maximize $s(m)$ on the manifold of macrostate values m such that $u(m) = u$ (see Section 4.3). On this manifold, the quantity $\beta u(m) - s(m)$ is thus globally minimized by m^u and equals

$$\beta u(m^u) - s(m^u) = \beta u - s(u). \tag{4.36}$$

At this point, it seems that our conjecture follows if only we can ascertain the stability of m^u against variations of the mean energy. This, we anticipate, should be determined by looking at the stability of u with respect to the quantity $\beta u - s(u)$ (see Section 4.2). Work aimed at consolidating these pieces of information into a rigorous proof is ongoing; see also the notes of this chapter for further hints.

SUMMARY OF CHAPTER

- **Concave function:** The function $s(u)$ is concave at u if and only if

$$s(v) \leq s(u) + \beta(v - u) \tag{4.37}$$

  for all v with $\beta = s'(u)$. The same function $s(u)$ is strictly concave at u if and only if

$$s(v) < s(u) + \beta(v - u) \tag{4.38}$$

  for all $v \neq u$ with $\beta = s'(u)$.

- **Realizability of the canonical equilibrium mean energy:** If $s(u)$ is concave at u, then u is realized in the canonical ensemble as an equilibrium value of the mean energy for $\beta = s'(u)$, i.e., $u_{\beta=s'(u)} = u$. In the case where $s(u)$ is strictly concave at u, $u_\beta = u$ is the unique value of the equilibrium value of the mean energy at $\beta = s'(u)$.

- **Non-realizability of the canonical equilibrium mean energy:** If $s(u)$ is nonconcave at u, then $u_\beta \neq u$ for all β. That is to say that u is nowhere realized in the canonical ensemble as an equilibrium value of the mean energy.

- **Realizability of canonical equilibrium macrostates:** Let β be given. Then

$$\mathscr{E}_\beta = \bigcup_{u \in u(\mathscr{E}_\beta)} \mathscr{E}^u, \tag{4.39}$$

  where $u(\mathscr{E}_\beta)$ stands for the set of all mean energy values u that can be written as $u(m)$

for some $m \in \mathscr{E}_\beta$.

- **Full macrostate equivalence of ensembles:** If $s(u)$ is strictly concave at u, then there exists a value of β equal to $s'(u)$ such $\mathscr{E}^u = \mathscr{E}_\beta$.

- **Partial macrostate equivalence of ensembles:** If $s(u)$ is concave at u but in a nonstrict fashion, then $\mathscr{E}^u \subsetneq \mathscr{E}_\beta$ for $\beta = s'(u)$ and $\mathscr{E}^u \neq \mathscr{E}_\beta$ for all other values of β.

- **Macrostate nonequivalence of ensemble:** If $s(u)$ is nonconcave at u, then $\mathscr{E}^u \cap \mathscr{E}_\beta = \emptyset$ for all β.

- **First-order canonical phase transitions:** There is macrostate nonequivalence of ensembles over the interval (u_l, u_h) of mean energy if and only if the canonical equilibrium mean energy value u_β, taken as a function of β, jumps discontinuously at $\beta_c = s'(u_l) = s'(u_h)$. There is macrostate nonequivalence of ensembles *or* partial equivalence of ensembles if u_β has an infinite slope at the same critical inverse temperature β_c. (These results follow by combining the results of Sections 4.2 and 4.3.)

- **Alternative macrostate representation of the entropy:** The concave hull $s^{**}(u)$ of the microcanonical entropy function $s(u)$ satisfies the following macrostate representation formula:

$$s^{**}(u) = \inf_\beta \sup_m \{\beta[u - u(m)] + s(m)\}. \tag{4.40}$$

  Interchanging the inf and sup in this formula yields a formula for $s(u)$:

$$s(u) = \sup_m \inf_\beta \{\beta[u - u(m)] + s(m)\}. \tag{4.41}$$

NOTES AND REMARKS

Misconceptions about Microcanonical and Canonical Ensembles Being Equivalent

The prevalent opinion that was forged over the years about the nonequivalence of ensembles problem is, as was mentioned in the introductory chapter, that the microcanonical and canonical ensembles always give equivalent descriptions of statistical mechanical models in the thermodynamic limit, no matter what model is considered. The following argument, found in the classic text of Landau and Lifshitz (1991, §29), is representative of this opinion:

> The possibility of applying the Gibbs distribution ... to closed systems is also seen from the fact that this distribution hardly differs from the microcanonical distribution, while being very much more convenient for practical calculations. For the microcanonical distribution is, roughly speaking, equivalent to regarding as equally probable all microstates of the body which correspond to a given value of its energy. The canonical distribution is "spread" over a certain range of energy values, but the width of this range (of the order of the mean fluctuation of the energy) is negligible for a macroscopic body.

Other similar arguments, which also concentrate on the negligibility of energy fluctuations in thermodynamic limit of the canonical ensemble, can be found in several popular textbooks on statistical mechanics; see, e.g., Huang (1987), Reif (1965, §6.7) and Balian (1991). One in particular which appears to us as being quite convincing (at first) can be found in Reif (1965, §6.7); it reads:

> If a macroscopic system A is in contact with a heat reservoir, the relative fluctuations in the energy of A are exceedingly small. Suppose now that A is removed from contact with the heat reservoir and is thermally insulated; then its total energy cannot change at all. But the distinction between this situation and the previous one is so small that it is really utterly irrelevant for most purposes; in particular, the mean values of all physical quantities (e.g., of the mean pressure or the mean magnetic moment of A) remain quite unaffected. Hence it makes no difference whether these mean values are calculated by considering the system to be isolated so that it has equal probability of being in any one of its states of accurately specified fixed energy, or by considering it to be in contact with a heat reservoir so that it is distributed over all its states in accordance with a canonical distribution.

We see now why this sort of argument must fall short in general: the range where the mean energy macrostate u_n is probabilistically seen to converge in the thermodynamic limit of the canonical ensemble does not necessarily coincide with the range of definition of u_n. The two ranges of mean energy do not coincide precisely when the microcanonical and canonical are nonequivalent (either thermodynamically or at the level of macrostates).

Sources on Macrostate Nonequivalence of Ensembles

The prehistory of macrostate nonequivalence of ensembles starts with Lax (1955) [see also Yan and Wannier (1965), and Wannier (1966, p. 67)] who noted that expected values calculated in the microcanonical ensemble were not always identical to similar expected values calculated in the canonical ensemble. Other works on the subject were subsequently published only much later in time by Eyink and Spohn (1993), Lewis et al. (1994*a*), Kiessling and Lebowitz (1997), Antoni et al. (2002) as well as by Ellis, Haven and Turkington (2000, 2002).

Our primary source for the material presented in this chapter was Ellis et al. (2000), although our presentation of this material differs greatly from the one found in this reference. We have tried here to put more in evidence the conditions needed for the mean energy value u to be realized at equilibrium in the canonical ensemble, in addition to stress the physical role played by the mean energy for determining the many relationships between the microcanonical and canonical sets of equilibrium macrostates. (In short, we have tried to put "physical flesh" on the many abstract mathematical results devised by Ellis and collaborators.) The proofs of these relationships given here are also much simplified compared to those found in Ellis et al. (2000). They are so basically because we considered in this chapter only mean energy values u lying in the interior of the domain of definition of $s(u)$, and assumed that $s(u)$ is differentiable at all such u. Furthermore, we did not find necessary to prove the existence of the thermodynamic limit. The reader is referred to Ellis et al. (2000) for complete proofs of more general results which hold under weaker assumptions, and for a rigorous mathematical treatment of the thermodynamic limit.

Here now are some remarks about the material presented in this chapter which is not to be found in the work of Ellis and collaborators.

The sup-inf representation formula for $s(u)$ is due to Leyvraz and Ruffo (2002), as was already mentioned. It can also be found in print in Oono (1989, p. 173) who did not consider, however, the interchangeability of the sup and inf operators and its consequences for the equivalence or nonequivalence of the microcanonical and canonical ensembles.

The study of the behavior of the canonical rate function $I_\beta(u)$ in relation to the nonconcavity regions and flat parts of microcanonical entropy $s(u)$ is original to this thesis, as is the connection of this behavior with first-order phase transition (both normal and degenerate); see next subsection.

Finally, to satisfy the examiners of this thesis, the author would like to add that the conjecture about the nonequivalent microcanonical macrostates, stated in Section 4.5, is his own personal invention which came to him while working on the numerical calculations reported in the next chapter. It can be found in the paper written by the author in collaboration with Richard S. Ellis and Bruce Turkington which report on these calculations [see Ellis et al. (2003)]. Nonequivalent macrostates having the property of being saddlepoints or local minima of $I_\beta(m)$ were also reported recently by Antoni et al. (2002) for a long-range Hamiltonian model different than the BEG model.

More Remarks on Nonequivalent Ensembles and Phase Transitions

The results relating the nonequivalence of the microcanonical and canonical at the level of the mean energy with first-order canonical phase transitions are an important original contribution of this thesis. To the best of the author's knowledge, the only references mentioning something related to these results are works by Chomaz et al. (2001), Gulminelli and Chomaz (2002), and Schmidt et al. (2001) which discuss the bimodality (i.e., double peak shape) of the canonical energy probability measure $P_\beta(u)$ in connection with first-order phase transitions and the nonconcave regions of $s(u)$. It should be noted, however, that the treatment of the phenomenon of nonequivalent ensembles offered by these authors is rather incomplete as it does not appeal to the properties of the Legendre-Fenchel transforms. They seem also not to have anticipated the role of entropy functions having affine parts in *degenerate* first-order canonical phase transitions. The appearance of this type of phase transitions constitutes an important subtlety of the theory of nonequivalent ensembles which is reported here for the first time.

Let us remark also that there exist many works in the physics literature which discuss various criteria for establishing the appearance of first-order phase transitions in finite-size systems based on the bimodal shape of $P_\beta(u)$ and its n-dependent behavior [see, e.g., Binder and Landau (1984), Challa et al. (1986), Borgs and Kotecký (1990), Lee and Kosterlitz (1991), Borgs and Janke (1992), and Borgs and Kotecký (1992)]. However, none of these papers tie the subject of the bimodality of $P_\beta(u)$ with the nonconcave points of $s(u)$, and, consequently, with nonequivalent microcanonical and canonical ensembles.

Another Conception of Equivalent Ensembles

Some authors use the term "equivalence of ensembles" in a different sense than the one studied in these pages. What they have in mind is to prove the equivalence of the microcanonical and canonical ensembles at the level of the probability distributions of *single* particles, that is mathematically to prove conditional limit results of the form

$$P^u(dx_i) = P(dx_i|u_n(x^n) \in du) \stackrel{n\to\infty}{\longrightarrow} P_\beta(dx_i), \tag{4.42}$$

where x_i refers to the state of the i^{th} "particle" in the joint state x^n. (The convergence in the limit is "in probability".) For examples of such results, see van Campenhout and Cover (1981), Cover and Thomas (1991, §12.6), Stroock and Zeitouni (1991), Lewis et al. (1994*a*), Lewis et al. (1994*b*) and Lewis et al. (1995).

A Canonical Entropy?

Leyvraz and Ruffo (2002), in relation to the inf-sup and sup-inf macrostate representations of $s(u)$, refer to $s^{**}(u)$ as a **canonical entropy** [see also Gross (1997)]. We would like to point out that this terminology is somewhat misleading because the mean energy values such $s(u) \neq s^{**}(u)$ are not represented in the canonical ensemble. Thus, $s^{**}(u)$ actually contains more information than what the canonical ensemble is able to provide. A more acceptable definition of a canonical entropy, in the author's opinion, is $s(\beta) = s(m_\beta)$ [see Ellis (1985, p. 77)]. From this definition, the following result is easily proved: if $m_\beta \in \mathscr{E}^{u_\beta}$ for some u_β, then $s(\beta) = s(u_\beta)$.

5

Illustrative Physical Examples

We study in this chapter two spin models which provide a clear illustration of the theory of nonequivalent ensembles that we have developed in the two last chapters. The two models are purposely simple; what we aim at is not to discuss their relevance as physical models, but to illustrate the equivalence or nonequivalence of the microcanonical and canonical ensemble in the most direct and understandable way possible.

5.1. THE HALF-BLOCKED SPIN MODEL

After having studied the mixed sum problem of Section 3.1, it is but a small task to devise a simple physical model which has a nonconcave entropy: just reconsider the mixed mean sum in question, and imagine that each of the random variables entering in that sum represents a spin variable or a group of spin variables. On the first hand, think of the n IID random variables involved in the mean sum (3.2) as representing n completely uncorrelated (independent) spins $X_1, X_2, \ldots, X_n$ (Figure 5.1) whose mean energy is given, as in the free spin problem (Section 2.4), by

$$u_{n,X}(X^n) = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{5.1}$$

On the other hand, think of the Y random variable entering in (3.2) as a **block-spin random variable** which models the combined effect of n completely correlated spins $Y_1, Y_2, \ldots, Y_n$ (Figure 5.1) taking the same value in the set $\mathscr{X} = \{-1, +1\}$, so that

$$u_{n,Y}(Y^n) = \frac{1}{n}\sum_{i=1}^{n} Y_i = Y. \tag{5.2}$$

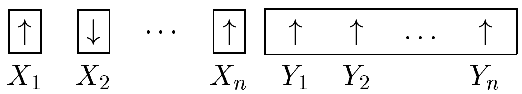

Figure 5.1: Half-block spin model: n spins out of the $2n$ spins are non-interacting (uncorrelated spins), while the remaining n spins take on the same value (completely correlated spins).

The combination of these two functions yields the **total mean energy** of the $2n$-spin system

$$u_n(Y^n, X^n) = u_n(Y, X^n) = Y + \frac{1}{n}\sum_{i=1}^{n} X_i \tag{5.3}$$

which has, as wanted, the form of the mean sum involved in the mixed sum problem. [*Note*: We have divided the total energy of the system by the factor n rather than by $2n$ in order not to carry a useless 2 in the mean energy function and in the calculations to come.]

At this point we straightforwardly apply all the techniques used before to study the mixed sum problem and the free spin model to find the microcanonical and canonical properties of the present "half-block" spin model. The microcanonical entropy, to begin with, may be calculated using a formula similar to equation (3.3) for $P(u_n \in du)$. In the present case, it yields:

$$s(u) = \begin{cases} s_0(u+1) & \text{if } u \in [-2, 0] \\ s_0(u-1) & \text{if } u \in (0, 2], \end{cases} \tag{5.4}$$

where

$$s_0(u) = -\left(\frac{1-u}{2}\right)\ln\left(\frac{1-u}{2}\right) - \left(\frac{1+u}{2}\right)\ln\left(\frac{1+u}{2}\right) - \ln 2 \tag{5.5}$$

is the entropy function associated with n independent spin (see Section 2.4). We next calculate the canonical free energy $\varphi(\beta)$ as was done before by separating the effect of the Y and X^n terms in the mean sum $u_n(Y, X^n)$. The Y term was already studied in Section 3.1, and leads to

$$\varphi_Y(\beta) = -|\beta|, \tag{5.6}$$

while the X^n term, treated in Section 2.4, was found to lead to

$$\varphi_{X^n}(\beta) = -\ln\cosh\beta. \tag{5.7}$$

The free energy being additive for independent random variables, we thus find

$$\varphi(\beta) = \varphi_Y(\beta) + \varphi_{X^n}(\beta) = -\ln\cosh\beta - |\beta|. \tag{5.8}$$

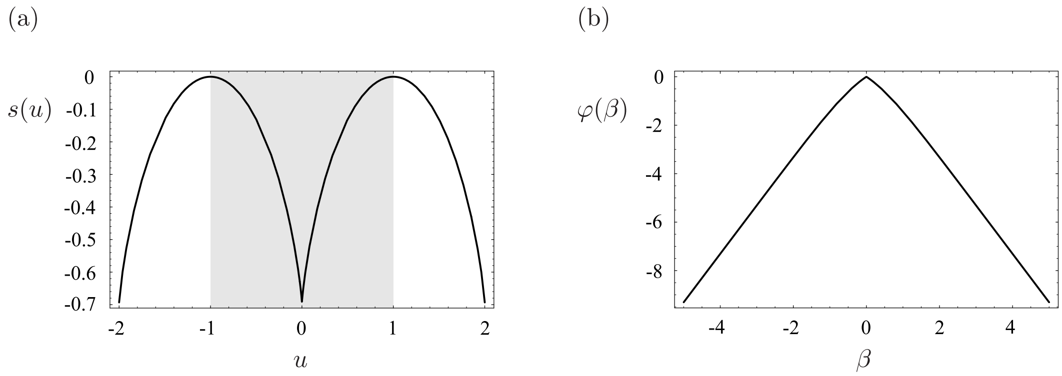

Figure 5.2: Nonconcave microcanonical entropy $s(u)$ (a) and free energy function $\varphi(\beta)$ (b) for the half-block spin problem. The shaded zone in the plot of $s(u)$ corresponds to the region of thermodynamic nonequivalence of ensembles where $s(u) \neq s^{**}(u)$.

To be able to compare the microcanonical and canonical ensembles, we finally calculate the concave hull of $s(u)$. No calculation is required to find this function as the graph of $s(u)$ shown in Figure 5.2 directly yields the answer:

$$s^{**}(u) = \begin{cases} s_0(u+1) & \text{if } u \in [-2,-1) \\ 0 & \text{if } u \in [-1,1] \\ s_0(u-1) & \text{if } u \in (1,2]. \end{cases} \tag{5.9}$$

From these results, we see that the microcanonical and canonical ensembles are thermodynamically nonequivalent for all $u \in (u_l, u_h) = (-1, 1)$ since over this interval of mean energy $s(u) \neq s^{**}(u)$ and thus

$$s(u) \neq \inf_{\beta}\{\beta u - \varphi(\beta)\}. \tag{5.10}$$

Following the general results derived in the last chapter, we expect accordingly to have nonequivalent microcanonical and canonical ensembles at the statistical level of this model, i.e., at the level of the vector $M_n = (Y, L_n)$ containing the value of the block spin variable and the empirical vector or statistical distribution $L_n = l$ of the n uncorrelated spins. To verify explicitly the nonequivalence of the two ensembles, we calculate the equilibrium values of (Y, L_n) using

$$u(Y = y, L_n = l) = y + \sum_{x=\pm 1} x l(x) \tag{5.11}$$

for the energy representation and

$$s(y,l) = s(l) = -\sum_{x=\pm 1} l(x)\ln l(x) + \ln 2 \tag{5.12}$$

for the macrostate entropy function. The form of this last function may be understood by noting that

$$P(Y=y, L_n=l) = P(Y=y)P(L_n=l), \tag{5.13}$$

and that $P(Y = y)$ does not scale exponentially with n while $P(L_n = l)$ does (recall Sanov's Theorem), so that

$$P(Y=y, L_n=l) \asymp e^{ns(l)}. \tag{5.14}$$

In the microcanonical ensemble, the mean energy constraint $u(y,l) = u$ together with the normalization constraint imposed on the empirical vector directly yield, as was the case in Section 2.4, the equilibrium values (y^u, l^u) without any need to maximize the entropy. The solution is

$$\begin{aligned}
l^u(-1) &= \begin{cases} -u/2 & \text{if } u \in [-2,0) \\ (2-u)/2 & \text{if } u \in (0,2] \end{cases} \\
l^u(+1) &= \begin{cases} (2-u)/2 & \text{if } u \in [-2,0) \\ u/2 & \text{if } u \in (0,2] \end{cases} \\
y^u &= \begin{cases} -1 & \text{if } u \in [-2,0) \\ +1 & \text{if } u \in (0,2] \end{cases}
\end{aligned} \tag{5.15}$$

for $u \in [-2,2]\backslash\{0\}$. For $u = 0$, we find the following special degenerate solution:

$$\begin{aligned}
(y^u, l^u(-1), l^u(+1)) &= (1,1,0) \\
(y^u, l^u(-1), l^u(+1)) &= (-1,0,1).
\end{aligned} \tag{5.16}$$

In the canonical ensemble, the equilibrium solutions (y_β, l_β) are found similarly as in Section 2.4 by minimizing once again the quantity $\beta u(y,l) - s(l)$. In the present case, we need to take proper care of the fact that $s(l)$ does not involve the value of the block spin variable. But this does not make altogether the problem less tractable analytically. In fact, except for the added variable Y, the solution is the same as for the simple non-interacting spin model:

$$l_\beta(-1) = \frac{e^\beta}{e^\beta + e^{-\beta}}, \qquad l_\beta(+1) = \frac{e^{-\beta}}{e^\beta + e^{-\beta}},$$
$$y_\beta = \begin{cases} -1 & \text{if } \beta > 0 \\ +1 & \text{if } \beta < 0. \end{cases} \tag{5.17}$$

At $\beta = 0$, we also have a degenerate equilibrium point corresponding to the two solutions:

$$(y_\beta, l_\beta(-1), l_\beta(+1)) = \left(-1, \frac{1}{2}, \frac{1}{2}\right)$$
$$(y_\beta, l_\beta(-1), l_\beta(+1)) = \left(+1, \frac{1}{2}, \frac{1}{2}\right). \tag{5.18}$$

All these quantities are illustrated graphically and compared together in Figure 5.3. As we have done in Section 2.4, we proceed to read this figure by choosing first a point in the microcanonical plot of l^u versus u, and then try to find an equivalent point in the canonical graph of l_β versus β. Following the theory developed in the previous chapter, if two such points exist for which $(y^u, l^u) = (y_\beta, l_\beta)$, then u must be such that $s(u) = s^{**}(u)$ and $\beta = s'(u)$. This is indeed observed for all u outside the shaded regions of mean energies. Inside those regions, however, we have macrostate nonequivalence of ensembles, since the points (y^u, l^u) cannot be mapped onto corresponding points (y_β, l_β) for any β, in agreement with the nonequivalence result of the last chapter.

One may be tempted to object that the empirical vectors l^u located within the nonequivalence region (shaded region) *can* be mapped onto "equivalent" empirical vectors l_β of the canonical ensemble. But the relationship between l^u and l_β is only an *incomplete* equivalence of ensembles, not a true *complete* equivalence of ensembles taking place at the level of the whole macrostate M_n which includes both L_n *and* Y. The fact is, in any case, that the canonical ensemble cannot entirely account for the microcanonical equilibria observed in the interval (u_l, u_h) because the former ensemble jumps over this interval in the manner of a first-order phase transition: y_β jumps from the value $y_\beta = -1$ to $y_\beta = +1$ (block spin reversal) as we positively move through the point $\beta = 0$ which makes u_β discontinuously jumps from the value $u_\beta = -1$ to the value $u_\beta = 1$. (l_β is continuous at $\beta = 0$.) In the microcanonical ensemble, by contrast, y^u undergoes a similar change only when the mean energy value $u = 0$ is crossed. It can also be noted that the canonical ensemble shows at $\beta = 0$ not one but *two* equilibrium values of (y, l) corresponding to the two coexisting equilibrium mean energies $u_{\beta=0} = -1$ and $u_{\beta=0} = 1$. In the jargon of the previous chap-

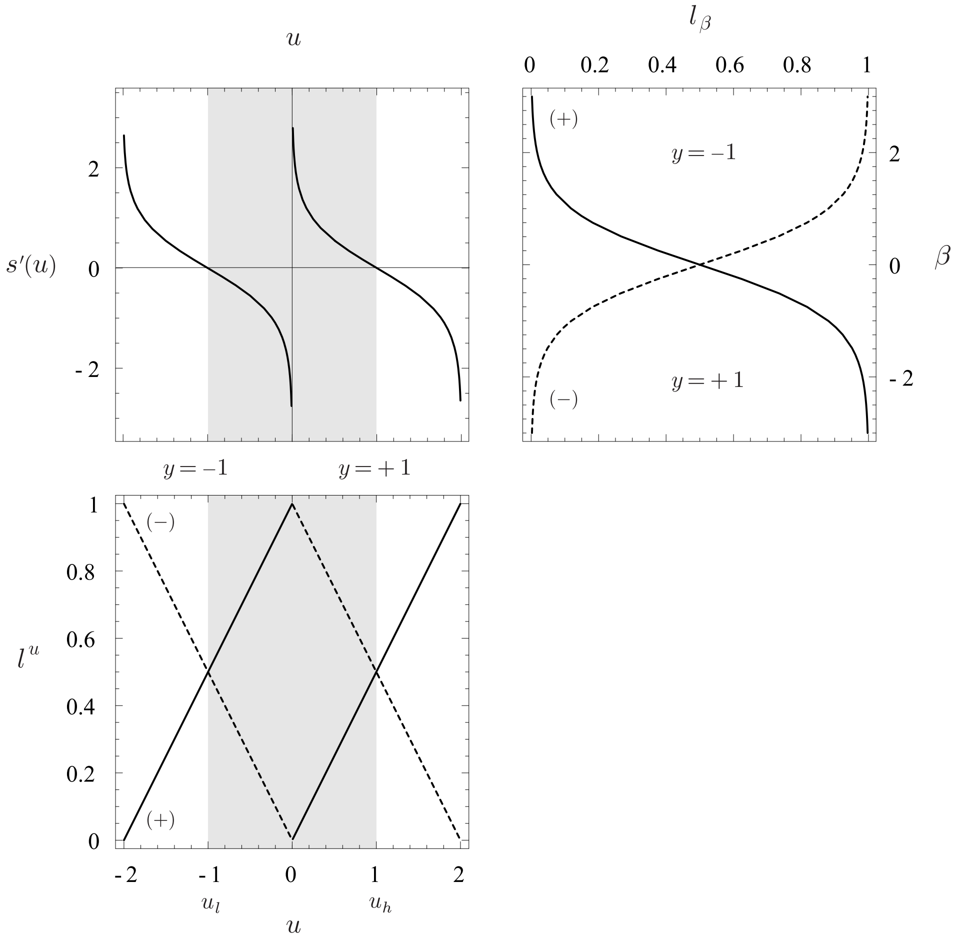

Figure 5.3: Equilibrium properties of the half-block spin model. (Top left) Derivative of the microcanonical entropy. (Middle right) -1 and $+1$ components of the canonical equilibrium empirical vector l^β. The equilibrium value of the Y component of this model is also shown on this graph. (Bottom left) -1 and $+1$ components of the microcanonical equilibrium empirical vector l^u together with the equilibrium value of Y. (Shaded region) Region of thermodynamic and macrostate nonequivalence of ensembles.

ter, we thus see that the points $u = -1$ and $u = 1$ correspond to cases of partial equivalence of ensembles for which we have $\mathscr{E}^u \subsetneq \mathscr{E}_\beta$ as a result of the fact that $\mathscr{E}_{\beta=0} = \mathscr{E}^{u=-1} \cup \mathscr{E}^{u=1}$.

5.2. THE MEAN-FIELD BLUME-EMERY-GRIFFITHS MODEL

The second model that we study for the purpose of illustrating the equivalence or nonequivalence of the microcanonical and canonical ensembles is a spin model originally devised by Blume, Emery and Griffiths (BEG) as a phenomenological mean-field model of the superfluid phases of liquid helium [Blume et al. (1971)]. The thermodynamic nonequivalence of the microcanonical and canonical ensembles for this model was proved by Barré, Mukamel and Ruffo (2001) who showed, via Landau expansion techniques, that the BEG microcanonical entropy has nonconcave anomalies for certain values of parameters defining this model. Our aim here, as it was in the previous section, is to show how the nonequivalence of the microcanonical and canonical ensembles for this model entails a more fundamental *statistical* nonequivalence of these two ensembles.

Definition of the Model

The model that we consider is not the actual original BEG model, but a variant of this model defined by the following Hamiltonian:

$$U(x^n) = \sum_{i=1}^{n} x_i^2 - K\left(\sum_{i=1}^{n} x_i\right)^2 \tag{5.19}$$

[see Barré et al. (2001) and Dauxois et al. (2002)]. In this formula, x_i represents a spin variable at site i taking values in the set $\mathscr{X} = \{-1, 0, +1\}$, and K is a positive real constant. As in the case of the noninteracting spin model, the macroscopic variable that we use to investigate the equivalence and nonequivalence of the microcanonical and canonical ensembles is the empirical vector

$$L_n = (L_n(-1), L_n(0), L_n(+1)), \tag{5.20}$$

whose three components $L_n(-1)$, $L_n(0)$, and $L_n(-1)$ give the proportion of spins in the microstate x^n that take the respective values -1, 0, and $+1$. The energy representation function associated with the mean energy

$$u(x^n) = \frac{U(x^n)}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - K\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \tag{5.21}$$

is trivially found to be

$$\begin{aligned} u(L_n = l) &= \sum_{x\in\mathscr{X}} x^2 l(x) - K\left(\sum_{x\in\mathscr{X}} x l(x)\right)^2 \\ &= l(+1) + l(-1) - K[l(+1) - l(-1)]^2 \end{aligned} \tag{5.22}$$

while the large deviation entropy function $s(l)$ associated with the probability measure $P(L_n \in dl)$ is given by

$$s(l) = -\sum_{x \in \mathscr{X}} l(x) \ln l(x) - \ln 3 \tag{5.23}$$

under the assumption that $P(x^n) = 3^{-n}$ for every $x^n \in \mathscr{X}^n$ (equiprobability hypothesis).

As before, these properties of L_n allow us to characterize the equilibrium macrostates with respect to the microcanonical and canonical ensembles as solutions of an appropriate optimization problem. In the case of the microcanonical ensemble, the set $\mathscr{E}^u = \{l^u\}$ of equilibrium empirical vectors l^u associated with the mean energy u are calculated in the thermodynamic limit by maximizing $s(l)$ over the macrostate space $\mathscr{L}$ subject to the constraint $u(l^u) = u$. Solving this problem necessitates only the maximization of a function of one variable, since the normalization constraint on the components of the empirical vector reduces the number of independent components of l^u to two, while the microcanonical energy constraint reduces this number by one more. On the other hand, the set $\mathscr{E}_\beta = \{l_\beta\}$ of canonical equilibrium empirical vectors l_β parameterized by the inverse temperature β is found by maximizing the quantity $\beta u(l) - s(l)$ over $\mathscr{L}$. In this case, we are faced with an unconstrained two-dimensional maximization problem involving the two components $l(-1)$ and $l(+1)$.

Another method for constructing $\mathscr{E}_\beta$ can be based on the determination of the canonical equilibrium value of the total spin per particle or magnetization macrostate. It is presented in the Supplementary Material section found at the end of this chapter. The advantage of this alternate method is that the associated minimization problem is one-dimensional rather than two-dimensional.

Equivalence of Ensembles for $K = 1.1111$

In Figure 5.4 we present a first set of solutions for $\mathscr{E}^u$ and $\mathscr{E}_\beta$ corresponding to the value $K = 1.1111$, together with a plot of the derivative of the microcanonical entropy function $s(u)$. Because neither of the two optimization problems involved in the definitions of $\mathscr{E}^u$ and $\mathscr{E}_\beta$ could be solved analytically, we provide from this point on numerical results obtained using various routines available in the scientific software Mathematica (see Notes and Remarks Section for the details of the numerical calculations). The top left plot of Figure 5.4 showing $s'(u)$ was obtained by calculating an empirical vector $l^u \in \mathscr{E}^u$, which satisfies $u(l^u) = u$ and $s(l^u) = s(u)$. The top right and the bottom left plots display, respectively, the canonical and microcanonical equilibrium components of the empirical

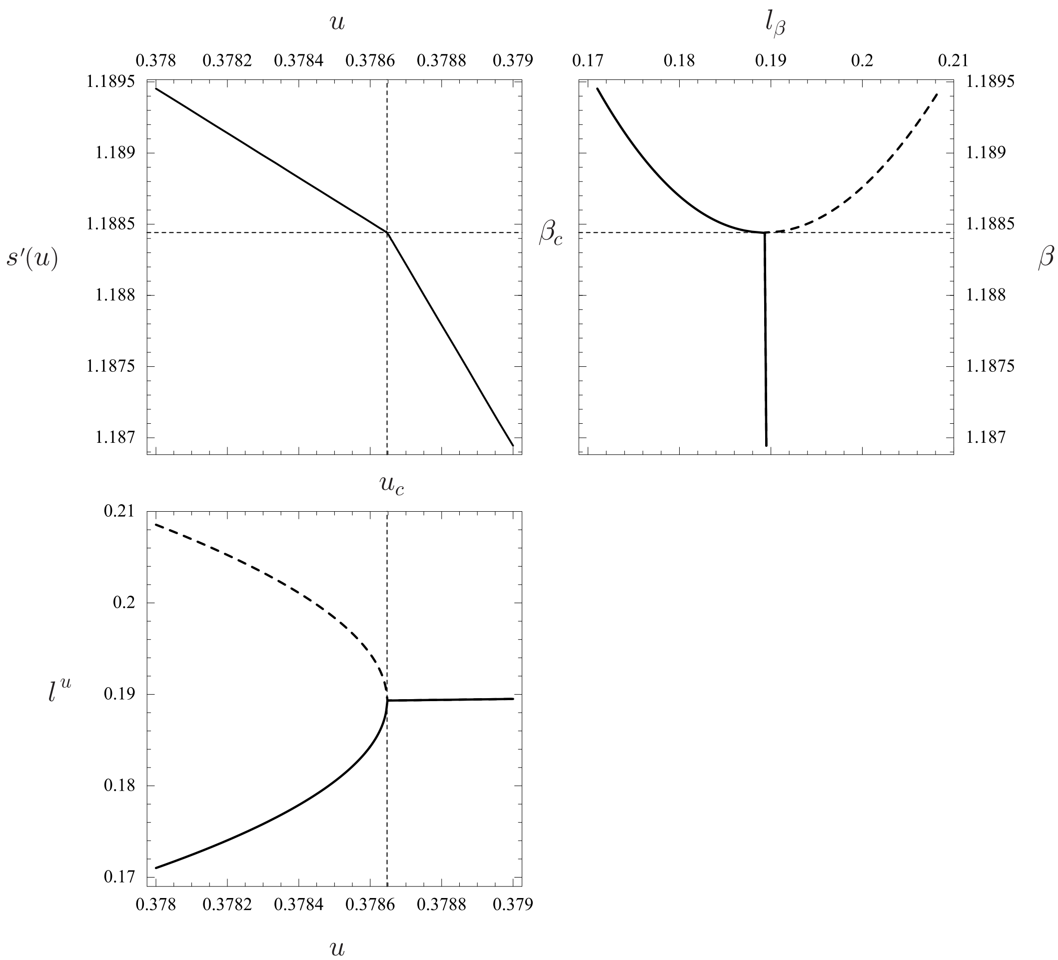

Figure 5.4: Full equivalence of ensembles for the BEG model with $K = 1.1111$. (Top left) Derivative of the microcanonical entropy $s(u)$. (Top right) The components $l(+1)$ and $l(-1)$ of the equilibrium empirical measure l_β in the canonical ensemble as functions of β. For $\beta > \beta_c$ the solid and dashed curves can be taken to represent $l(+1)$ and $l(-1)$, respectively, and vice versa. (Bottom left) The components $l(+1)$ and $l(-1)$ of the equilibrium empirical measure l^u in the microcanonical ensemble as functions of u. For $u < u_c$ the solid and dashed curves can be taken to represent $l(+1)$ and $l(-1)$, respectively, and vice versa.

vector as a function of the parameters β and u defining each of the two ensembles. In the top right plot, the solid curve can be taken to represent the spin $+1$ component of the equilibrium empirical vector l_β, while the dashed curve can be taken to represent the spin -1 component of the same equilibrium empirical vector. Since the BEG Hamiltonian satisfies the exchange symmetry $l(+1) \leftrightarrow l(-1)$, the roles of the solid and dashed curves can also be reversed. For $\beta \leq \beta_c$, the solid curve represents the common value of $l(+1) = l(-1)$. In all cases, the component $l(0)$ of l_β is determined by the normalization condition $l(0) = 1 - l(+1) - l(-1)$. The same explanation applies to the bottom left plot of l^u.

The first series of plots displayed in Figure 5.4 were designed to illustrate a case where $s(u)$ is concave and where, accordingly, we expect equivalence of ensembles. That the equivalence of ensembles holds in this case at the level of the empirical vector can be seen by noting that the solid and dashed curves representing the $l(+1)$ and $l(-1)$ components of l_β in the top right plot can be put in one-to-one correspondence with the solid and dashed curves representing the same two components of l^u in the bottom left plot. The one-to-one correspondence, as we now know, is defined by the derivative of the microcanonical entropy $s(u)$: for a given u we have $l^u = l_{\beta(u)}$ with $\beta(u) = s'(u)$. Moreover, since the monotonic function $s'(u)$ can be inverted to yield a function $u(\beta)$ satisfying $s'(u(\beta)) = \beta$, we have $l_\beta = l^{u(\beta)}$ for all β. Thus, the equilibrium statistics of the BEG model in the microcanonical ensemble can be translated unambiguously into equivalent equilibrium statistics in the canonical ensemble and vice versa. In this case, the critical mean energy u_c at which the BEG model goes from a high-energy phase of zero **magnetization**

$$m(l) = l(+1) - l(-1) \tag{5.24}$$

to a low-energy phase of nonzero magnetization in the microcanonical ensemble can be calculated from the viewpoint of the canonical ensemble by finding the critical inverse temperature β_c that determines the onset of the same phase transition in the canonical ensemble. Since the two ensembles are equivalent, both the microcanonical and canonical phase transitions must be of the same order, which in this case is second-order.

Nonequivalence of Ensembles for $K = 1.0817$

In the second series of plots in Figure 5.5, a case of ensemble nonequivalence corresponding to the value $K = 1.0817$ is shown. Since in the top left plot $s'(u)$ is not monotonic, $s(u)$ is not concave. As in many of the figures presented in the previous chapters, the open interval

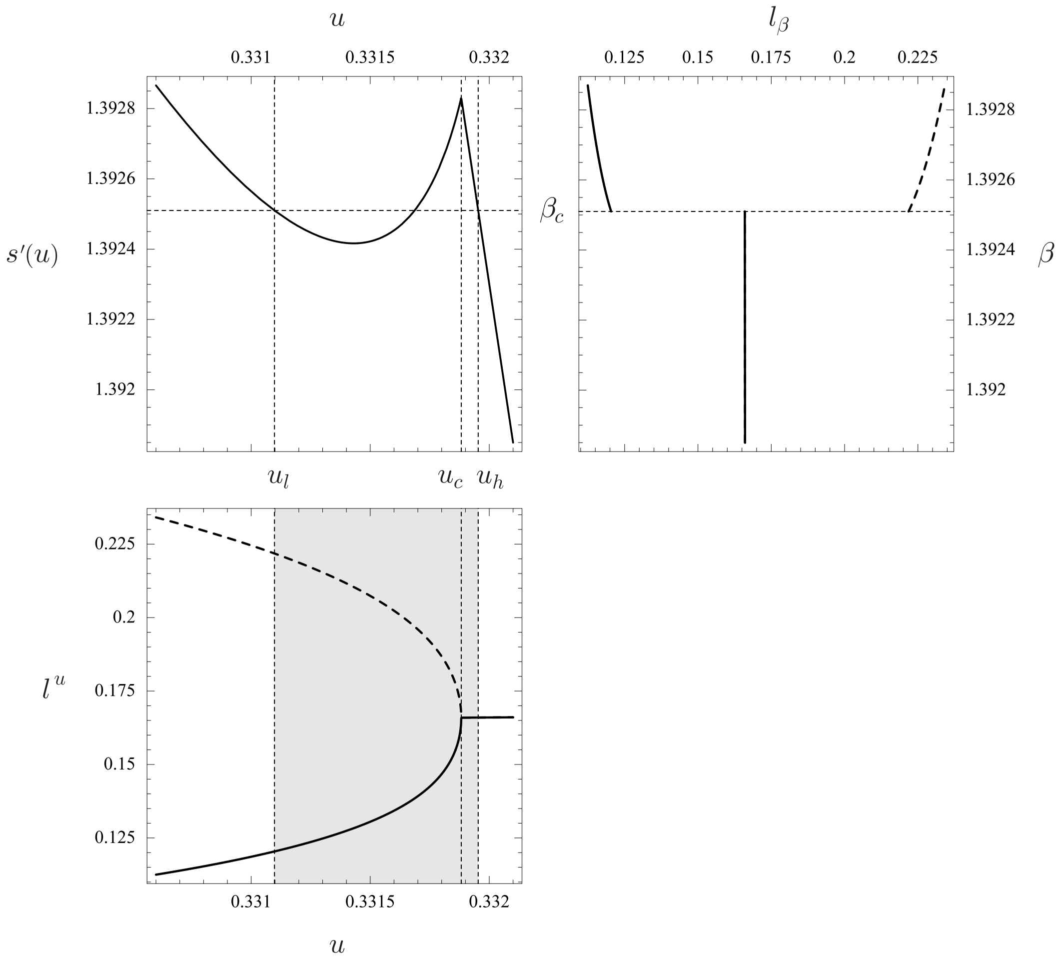

Figure 5.5: Equivalence and nonequivalence regions for the BEG model with $K = 1.0817$. The solid and dashed curves are interpreted as in Figure 5.4. The shaded area in the bottom left plot corresponds to the region of nonequivalence of ensembles delimited by the open interval (u_l, u_h). The ranges of the inverse temperature and the mean energy used to draw the plots were chosen so as to obtain a good view of the phase transitions.

(u_l, u_h) of mean energy values is the interval on which $s(u) \neq s^{**}(u)$; on this interval $s(u)$ is nonconcave and $s^{**}(u)$ is affine with slope β_c. By comparing the top right plot of l_β and the bottom left plot of l^u, we see that the elements of $\mathscr{E}^u$ cease to be related to elements of $\mathscr{E}_\beta$ for all mean energy values u in the interval (u_l, u_h). In fact, for any u in this interval of thermodynamic nonequivalence of ensembles (shaded region) no l_β exists that can be put in correspondence with an equivalent equilibrium empirical vector contained in $\mathscr{E}^u$. This lack of correspondence agrees with the rigorous results reviewed in Section 4.3. Thus, although the equilibrium macrostates l^u corresponding to $u \in (u_l, u_h)$ are characterized by a well defined value of the mean energy, it is impossible to assign a temperature to those macrostates from the viewpoint of the canonical ensemble. In other words, the canonical ensemble is blind to all mean energy values u contained in the domain of nonconcavity of $s(u)$. By decreasing β continuously through the critical value β_c, the equilibrium value of the energy per particle associated with the empirical vectors in $\mathscr{E}_\beta$ jumps discontinuously from u_l to u_h (canonical first-order phase transition). However, outside the range (u_l, u_h) we have equivalence of ensembles, and a continuous variation of β induces a continuous variation of u.

We can go further in our analysis of the plots of Figure 5.5 by noting that the phase transition exhibited in the microcanonical ensemble is second-order (continuous) whereas it is first-order (discontinuous) in the canonical ensemble. This provides another clear evidence of the nonequivalence of the two ensembles. Again, because the canonical ensemble is blind to all mean energy values located in the nonequivalence region, only a microcanonical analysis of the model can yield the critical mean energy u_c. As for the critical inverse temperature β_c, which signals the onset of the first-order transition in the canonical ensemble, its precise value can be found by calculating the slope of the affine part of $s^{**}(u)$ or, equivalently, by identifying the point of non-differentiability of $\varphi(\beta)$. It may also be found using Maxwell's equal-area construction.

Further Results for $K = 1.0805$

To conclude this section, we present in Figure 5.6 a final series of plots of $s'(u)$, l_β, and l^u corresponding to $K = 1.0805$, a slightly smaller value than the one considered in Figure 5.5. As in Figure 5.5, there also exists in Figure 5.6 an open interval (u_l, u_h) over which $s(u)$ is nonconcave. For $u \in (u_l, u_h)$ we consequently have nonequivalence of ensembles, illustrated by the shaded region in the bottom left plot. As in Figure 5.5, the nonequivalence of ensembles is associated with a first-order phase transition in the canonical ensemble determined by β_c. The microcanonical phase transition seen in Figure 5.6 is also first-order due to the jump in $s'(u)$ as u increases through the critical value u_c. By contrast, the microcanonical transition is second-order in Figure 5.5.

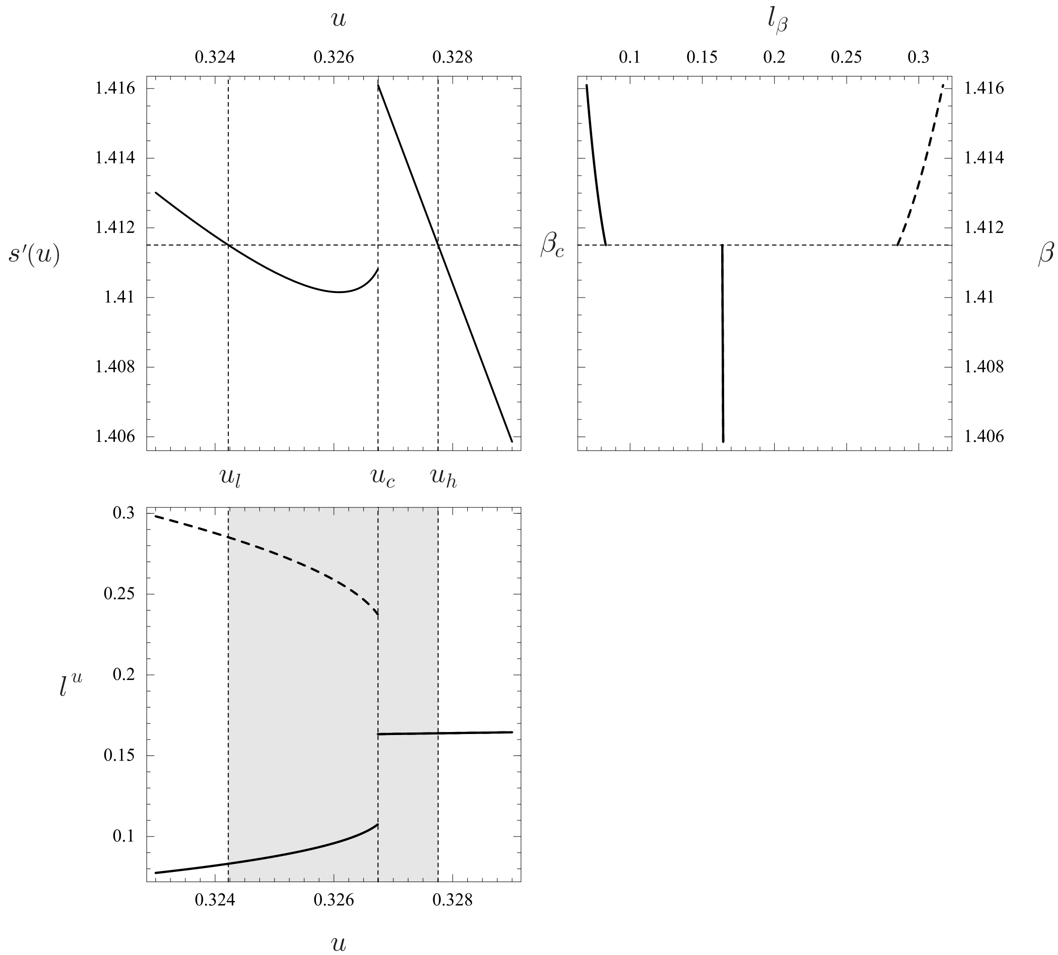

Figure 5.6: Equivalence and nonequivalence regions for the BEG model with $K = 1.0805$. The solid and dashed curves are interpreted as in Figure 5.4. The shaded area in the bottom left plot corresponds to the region of nonequivalence of ensembles delimited by the open interval (u_l, u_h).

SUPPLEMENTARY MATERIAL: EQUILIBRIUM EMPIRICAL VECTORS OF THE BEG MODEL

We show in this section that, given a value of β, the canonical equilibrium magnetization value m_β of the BEG model can be used to infer the canonical equilibrium empirical vector L_β. To prove this result, we start with Gibbs' canonical probability measure for the BEG model:

$$P_\beta(x^n) = \frac{1}{Z_n(\beta)} \exp\left\{-\beta\left[\sum_{i=1}^n x_i^2 - \frac{K}{n}\left(\sum_{i=1}^n x_i\right)^2\right]\right\} P(x^n). \tag{5.25}$$

In this expression, $Z_n(\beta)$ is the n-particle partition function, and $P(x^n) = 3^{-n}$ is the uniform probability measure on $\mathscr{X}^n = \{-1,0,1\}^n$. By absorbing the sum of the x_i^2 terms in the a priori measure P, we can re-write the Gibbs measure as

$$P_\beta(x^n) = \frac{1}{C}\exp\left[\frac{\beta K}{n}\left(\sum_{i=1}^n x_i\right)^2\right]\bar{P}_\beta(x^n), \tag{5.26}$$

having defined

$$\bar{P}_\beta(x^n) = \prod_{i=1}^n p_\beta(x_i) = \prod_{i=1}^n \frac{e^{-\beta x_i^2}}{\bar{Z}(\beta)}, \tag{5.27}$$

and

$$\bar{Z}(\beta) = \sum_{x_i\in\mathscr{X}} e^{-\beta x_i^2} = 2e^{-\beta} + 1. \tag{5.28}$$

The letter C above stands for the constant normalizing P_β.

At this point, we extend the new expression (5.26) for P_β to the space of the empirical vector so as to obtain

$$P_\beta(L(x^n)\in dl) = \frac{1}{C}\exp\left[n\beta K\left(\sum_{x\in\mathscr{X}} x\,l(x)\right)^2\right]\bar{P}_\beta(L(x^n)\in dl). \tag{5.29}$$

Also, since $\bar{P}_\beta(x^n)$ is given by the product measure (5.27) on $\mathscr{X}^n$, we can appeal to Sanov's Theorem to write

$$P_\beta(L(x^n)\in dl) \asymp e^{-nD(l||p_\beta)}, \tag{5.30}$$

where

$$D(l||p_\beta) = \sum_{x\in\mathscr{X}} l(x)\ln\frac{l(x)}{p_\beta(x)} \tag{5.31}$$

is the relative entropy (Kullback-Leibler distance) between $l(x)$ and $p_\beta(x)$. Sanov's Theorem thus implies the following large deviation estimate

$$P_\beta(L(x^n)\in dl) \asymp \exp\left\{-n\left[D(l||p_\beta) - \beta K\left(\sum_{x\in\mathscr{X}} x\,l(x)\right)^2 + c\right]\right\} \tag{5.32}$$

for $P_\beta(L(x^n)\in dl)$. Here, c is another constant (the free energy basically).

This last equation constitutes half of our proof: it shows that the value canonical equilibrium empirical vector L_β can be computed as

$$l_\beta = \arg\inf_l \left[D(l||p_\beta) - \beta K \left(\sum_{x\in\mathscr{X}} x l(x) \right)^2 \right]. \tag{5.33}$$

The other half proceeds by defining the quantity

$$m(l) = \sum_{x\in\mathscr{X}} x l(x) = l(+1) - l(-1)$$

as the magnetization of the system associated with a given value $L = l$ of the empirical vector, and in re-expressing the optimization problem above as follows:

$$\begin{aligned} l_\beta &= \arg \inf_{m\in[-1,+1]} \inf_{l:m(l)=m} \left[D(l||p_\beta) - \beta K m^2 \right] \\ &= \arg \inf_{m\in[-1,+1]} \left[\inf_{l:m(l)=m} D(l||p_\beta) - \beta K m^2 \right]. \end{aligned} \tag{5.34}$$

The constrained infimum in the square bracket is easy to solve; it has for solution

$$l_{\beta,m}(x) = \frac{p_\beta(x) e^{t(m)}}{W_\beta(t(m))}, \tag{5.35}$$

where

$$W_\beta(t(m)) = \sum_{x\in\mathscr{X}} p_\beta(x) e^{t(m)x}. \tag{5.36}$$

In this equation, $t(m)$ is the Lagrange multiplier associated with the constraint $m(l_{\beta,m}) = m$; its actual value as a function of m is determined by solving the differential equation

$$\frac{\partial}{\partial t} \ln W_\beta(t) = m. \tag{5.37}$$

The second infimum over the values of the magnetization can be solved by expanding $D(l_{\beta,m}||p_\beta)$ in (5.34), and by taking the derivative of the resulting expression. The net result obtained by following these steps is that the value of m where the infimum if achieved is given, for the specific case of the BEG model, by solving the equation

$$t(m) = 2\beta K m. \tag{5.38}$$

It can be shown without too much difficulty that the very same equation also determines the equilibrium value of the magnetization with respect to the canonical probability distribution P_β. Thus, we have shown that if one is able to compute the canonical equilibrium value of the magnetization (a one-dimensional optimization problem), then one is able to obtain an expression for l_β using equations (5.35) and (5.37).

NOTES AND REMARKS

Sources and Credits

The simple half-block spin model was imagined by the author at a physics conference held at the Santa Fe Institute (Santa Fe, New Mexico) during the week of April 11, 2002. The properties of this model are reported here for the first time.

The numerical results pertaining to the BEG model were also obtained by the author who used for the values of K the same values used by Barré et al. (2001). A paper written by the author in collaboration with Richard S. Ellis and Bruce Turkington which reports on these results was submitted recently for publication, as was already noted in the Introduction [see Ellis et al. (2003)]. The division of the work which led to the writing of this paper is as follows: the author imagined the problem and obtain all the numerics, while Ellis and Turkington provided much comments and useful hints as to how these results could be obtained, and how they relate to some of their rigorous results published in Ellis et al. (2000). Ellis suggested, in particular, the method for reducing the 2-D optimization problem needed to find the canonical equilibrium empirical vectors of the BEG model down to a 1-D optimization problem involving the magnetization variable (see Supplementary Material).

Numerical Analysis of the BEG model: Technical Notes

The mathematical software Mathematica 4.2 (trademark of Wolfram Research) was used in our study of the BEG model to numerically solve the optimization problems defining the microcanonical and canonical sets of equilibrium empirical vectors. All the calculations leading to the graphics presented in Figures 5.4-5.6 were performed in single precision numerical format (16 digits-precision), and were based on two numerical routines of Mathematica: `FindMinimum` and `NMinimize`. The details of the use of these routines is presented in the next points.

- Figure 5.4; $K = 1.1111$.
  - Microcanonical ensemble: Recall from Section 5.2 that the microcanonical equi-

librium values $l^u(+1)$ and $l^u(-1)$ are found by maximizing a simple function of one variable which results from inserting the microcanonical energy constraint $u(l) = u$ and the normalization constraint on l in the expression of the entropy function $s(l)$. To numerically locate the maximum (or maxima) of this 1-D function as a function of u, we simply used the symbolic routine `FindMinimum`. The syntax of this routine is

$$\texttt{FindMinimum}[-f(x), \{x, x_0, x_{\min}, x_{\max}\}],$$

where $f(x)$ is the function to maximize, x is the variable of maximization, and $x_0 \in [x_{\min}, x_{\max}]$ is the starting point used by numerical routine to locate the maximum of $f(x)$. The points $x_{\min}$ and $x_{\max}$ determine of course the range of values within which the maximum is to be located. In our case, $x_{\min} = 0$ and $x_{\max} = 1$.

– Stability and verification of the results: The microcanonical maximization problem for this value of K is numerically well-conditioned. The 1-D function to be maximized is smooth and presents only one maxima which splits continuously into two maxima at u_c (second-order phase transition). The numerical value of u_c that we have determined from our calculations matches, with the numerical precision stated (16 digits-precision), the exact critical value found from the Landau analysis of the BEG model [see Barré, Mukamel and Ruffo (2001)]. Furthermore, the author checked graphically that the positions of the maxima returned by the routine `FindMinimum` were indeed global maxima.

– Canonical ensemble: In this ensemble, we determined the values of $l_\beta(+1)$ and $l_\beta(-1)$ by minimizing the two-dimensional function $\beta u(l) - s(l)$ (the normalization constraint on l is assumed). The numerical routine that we used to solve this problem is the numerical equivalent of `FindMinimum` called `NMinimize` and whose syntax is

$$\texttt{NMinimize}[\{g(x, y), x > 0 \wedge y > 0 \wedge x + y < 1\}, \{x, y\}].$$

The advantage of using `NMinimize` over `FindMinimum` is only the speed at which the solutions are found: `NMinimize` is usually faster than `FindMinimum` because the former routine handles only numbers, whereas the latter tries to keep its input variables in symbolic form.

  - Stability of the results: The minimization of $\beta u(l) - s(l)$ is straightforward. This function is smooth, and the numerical estimate of β_c resulting from our calculations could be checked against the exact value of the critical inverse temperature determined by a Landau analysis of the model.

- Figure 5.5; $K = 1.0817$.

  - Microcanonical ensemble: Same as for $K = 1.1111$.
  - Canonical ensemble: For this value of K, we were careful to locate all the local minima of the quantity $\beta u(l) - s(l)$ for a given β using the method described for the previous value of K. To verify our results, we also used the routine `FindMinimum` with the syntax

$$\texttt{FindMinimum}[f(x,y),\{x,x_0,x_{\min},x_{\max}\},\{y,y_0,y_{\max},y_{\min}\}].$$

    Then we compared the various minima to determine which of them were global minima and which of them were local minima. The global minima were used as the values of l_β. Repeating this process for many values of β, we could numerically evaluate β_c.
  - Stability and verification of the results: We confirmed the validity of the results obtained with `FindMinimum` and `NMinimize` by calculating some points l_β using the 1-D technique described in the previous section. The value of the critical inverse temperature for the canonical first-order phase transition was also confirmed with another numerical method based on the determination of the value β_c for which the two symmetric minima $a \neq 0$ of the quantity

$$\varphi(\beta, a) = \beta K a^2 - \ln[1 + e^{-\beta}(e^{2K\beta a} + e^{-2K\beta a})] \tag{5.39}$$

    becomes global minima [see Barré, Mukamel and Ruffo (2001)].

- Figure 5.6; $K = 1.0805$.

  - Microcanonical ensemble: The 1-D function to maximize for this value of K presents many local maxima whose heights cross at some value u_c (first-order phase transitions). In order to locate the true global maximum (or maxima) of this function, we have used the numerical routine `NMinimize` with the option `RandomSearch` to force the localization of all local maxima of the function received as input. With this option, the local maxima found are compared together to determine the global one(s). The syntax of NMinimize with this option is

$$\texttt{NMinimize}[\{f(x), x > 0 \wedge x < 1\}, \{x\}, \texttt{Method} \to \texttt{``RandomSearch''}]$$

– Stability and verification of the results: The validity of the microcanonical results obtained for this value of K was directly verified by visualizing the shape (and behavior as a function of u) of the 1-D macrostate entropy function that has to be maximized in the microcanonical ensemble.

– Canonical ensemble: Same as for $K = 1.0817$.

Outlook and Open Problems

MAIN CONCLUSIONS

A complete technical summary of this thesis can be obtained by pasting together the summaries presented at the end of each chapter. Here we only recapitulate, in non-technical words, the most important results and conclusions in the order in which they were presented.

- The equilibrium properties of a many-body system calculated for fixed values of its internal energy may differ from its properties calculated in the situation where the system is in contact with a heat bath having a fixed temperature.

- At the theoretical level, this incompatibility between the energy-dependent and the temperature-dependent properties of a system is signalled by an incompatibility between the predictions of the microcanonical ensemble (constant energy ensemble) and the canonical ensemble (constant temperature ensemble).

- A first thermodynamic level of nonequivalence of the microcanonical and canonical ensembles takes place whenever the microcanonical entropy function, the basic thermodynamic function of the microcanonical ensemble, has nonconcave parts. In such a case, the Legendre-Fenchel structure relating the thermodynamic functions of the microcanonical ensemble to those of the canonical ensemble, and *vice versa*, is broken because the Legendre-Fenchel transform of nonconcave functions is non-invertible.

- More precisely, in case of thermodynamic nonequivalence of ensembles, the microcanonical entropy cannot be expressed as the Legendre-Fenchel transform of the free energy. What is obtained by taking this transform is the concave hull of the microcanonical entropy, i.e., the minimal concave envelope of the entropy function.

- The free energy can always be expressed as the Legendre-Fenchel transform of the microcanonical entropy function even if the latter function is nonconcave, i.e., even if there is thermodynamic nonequivalence of ensembles. This point and the previous one imply that the canonical ensemble properties of a system can always be derived from the microcanonical, although the contrary is not always true.

- The microcanonical and canonical ensembles are thermodynamically nonequivalent whenever the heat capacity, calculated microcanonically as a function of the mean energy, is negative. (This constitutes only a sufficient criterion, not a necessary one.)

- The thermodynamic level of nonequivalence of ensembles is but the zeroth level of nonequivalence; a more fundamental way to test whether the microcanonical and canonical ensemble are equivalent or nonequivalent is to compare the equilibrium values of macrostates calculated from the point of view of each ensemble (macrostate equivalence or nonequivalence of ensembles).

- The concavity properties of the microcanonical entropy (thermodynamic level) determines the relationships between the microcanonical and the canonical sets of equilibrium macrostates (macrostate level).

- Namely, if the entropy is strictly concave (case of thermodynamic equivalence), the two ensembles are also equivalent at the macrostate level (macrostate equivalence), in the sense that the microcanonical set of equilibrium macrostates can be put into a one-to-one correspondence with the canonical set of equilibrium macrostates.

- If the entropy is concave but not strictly, then the microcanonical set of equilibrium macrostates is but a strict subset of the set of canonical equilibrium macrostates.

- If the entropy is nonconcave (thermodynamic nonequivalence), then there exists microcanonical equilibrium macrostates which are nowhere realized in the canonical ensemble for all values of the temperature (macrostate nonequivalence).

- The fact that the microcanonical and canonical ensembles may be nonequivalent (either at the thermodynamic or at the macrostate level) has for consequence that not all mean energy values are realized in the canonical ensemble as equilibrium values of the mean energy. Thus, in case of nonequivalence of ensembles, the canonical ensemble skips over the microcanonical ensemble in the manner of a discontinuous (first-order) phase transition.

- The macrostate values realized at equilibrium in the microcanonical ensemble but not in the canonical ensemble (nonequivalent macrostates) correspond to critical nonequilibrium macrostate values of the canonical ensemble.

OPEN PROBLEMS

The present thesis hardly empties the subject of nonequivalent statistical mechanical ensembles. Many more problems could have been discussed in these pages, but due to space and time limitations, a number of them had to be put aside. By way of final conclusion, we provide below a partial list of such problems together with some clues, hints and references which may prove useful to solve them.

Conjecture About Metastable Nonequivalent Macrostates

Prove the conjecture stated in Section 4.5 about the equilibrium microcanonical nonequivalent macrostates realized as critical nonequilibrium points of the canonical ensemble. Try also to put in evidence, either experimentally or computationally, the existence of these states; see, e.g., Latora and Rapisarda (2001) and Latora et al (2001, 2002). See also the last problem of this section.

Other Physical Models with Nonequivalent Microcanonical and Canonical Ensembles

Revisit the numerous models which have been discovered in the past years to have nonconcave entropies to provide further illustrations of the phenomenon of macrostate nonequivalence of ensembles and its relation with the thermodynamic level of nonequivalent of ensembles; see, e.g., Thirring (1970), Kiessling and Lebowitz (1997) and Dauxois et al. (2000).

The design of new physical models having nonconcave anomalies in their microcanonical entropy function is also a quest worth pursuing. Models with entropy functions having affine parts, for instance, are not known to the author. Perhaps a good start for finding new physical models is to look at examples of sums of random variables which are already known to be characterized by nonconvex rate functions [see, e.g., Dinwoodie and Zabell (1992), Dinwoodie (1993) and Ellis (1995)].

Types of Interactions Leading to Nonequivalent Ensembles

The presence of long-range and mean-field (infinite range) interactions in statistical mechanical models seems to be a *causa sin qua non* for the appearance of nonequivalent microcanonical and canonical equilibrium properties; yet, no rigorous proof of this observation is known to exist at this time. (This issue is not without relationship with the other issue of determining which types of interactions are responsible for the appearance of first-order phase transitions in the canonical ensemble; see Section 4.2.) What is known is that the equivalence of the microcanonical and canonical ensemble holds rigorously for noninteracting systems because, for these systems, the optimization problems involved in the determination of the microcanonical and canonical equilibrium values of the empirical vector L (the most basic quantity to look at for noninteracting systems) is linear in L and involves an everywhere concave macrostate entropy function $s(L)$. According to the theory of convex functions, these two properties of L are necessary to guarantee that the microcanonical entropy, obtained by contracting the macrostate entropy $s(L)$, is everywhere concave. In view of this result, a question that seems to be of interest is: can this kind of argument based on convex analysis be generalized to other types of interactions?

A Limitation of the Theory of Macrostate Nonequivalence of Ensembles

The existence of an energy representation function $u(m)$ expressing the mean energy of a system as a function of some macrostate m of that system is a strong prerequisite, and, by consequence, a strong limitation of the theory presented in Chapter 4 about macrostate nonequivalence of ensembles. The existence of $u(m)$ poses no real problem for mean-field systems, but for systems having short-range interactions, such as the nearest-neighbor Ising model, it is not at all clear how $u(m)$ can be constructed, if it can be constructed at all. Perhaps there is a way to weaken the existence assumption of $u(m)$, or even get rid of it, and still be able to obtain useful results about macrostate equivalence or nonequivalence of ensembles. Work is ongoing on this question.

A Yang-Lee Theory of Nonequivalent Ensembles

Connect the Yang-Lee Theory of first-order phase transitions [Yang and Lee (1952), Lee and Yang (1952)] with the theory of nonequivalent microcanonical and canonical ensembles presented in this thesis. Further references on the Yang-Lee Theory can be found in Fisher (1965), Thompson (1972) and Salinas (2001).

Non-Thermodynamic Analogs of Nonequivalent Ensembles

The existence of nonconcave entropies is not *a priori* a problem of physics, but a problem of mathematics which emerge, as we have seen in this thesis, in connection with the theory of large deviations. From this extended perspective, it is to be expected that nonconvex entropies and non-differentiable free energies should arise in fields of physical investigation other than equilibrium statistical mechanics which make use of large deviation concepts.

Examples of such fields include the theory of dynamical systems and the theory of multifractal measures [see, e.g., Lopes (1990), Beck and Schlögl (1993), Veneziano (2002) and Zohar (1999)]. Finding nonconcave entropies in the context of these theories would mean that a non-thermodynamic analog of nonequivalent ensembles has been found.

Nonequilibrium Behavior of Nonequivalent Macrostates

The conjectured relationship between nonequivalent microcanonical macrostates and unstable or metastable macrostates of the canonical ensemble brings about many interesting questions about the nonequilibrium or dynamical behavior of these macrostates. Being unstable or metastable from the point of view of the canonical ensemble, such nonequivalent macrostates must, in effect, evolve towards more stable (canonical-equilibrium) macrostates as soon as the microcanonical energy constraint ceases to be enforced, i.e., as soon as the system initially assuming one of the nonequivalent macrostate values is put in contact with a heat reservoir. But how exactly is the system evolving towards canonical equilibrium? How much time does it take for the system to reach the new canonical-equilibrium state starting from its unstable or metastable state? At which "speed" does it evolve to equilibrium? What is the rate of decay? Also, how is the system going from a metastable state to a stable state? That is, what is the physical process taking the system from one state to another? Is the process instantaneous? Can the process be stopped or "frozen" in any way?

To answer these questions, the reader is referred to the abundant literature that already exists on the subject of metastable states and their appearance in first-order phase transitions; see, e.g., Griffiths et al. (1966), Langer (1967, 1969), Penrose and Lebowitz (1971), Binder (1973), Gilmore (1979), Agarwal and Shenoy (1981), and Gunton and Droz (1983). Before attempting to read these papers, the textbooks of Balescu (1975, 1997) and Zwanzig (2001) may also be read with profit to gain a basic knowledge of nonequilibrium statistical mechanics.

Bibliography

Agarwal, G. S. and Shenoy, S. R. (1981), 'Observability of hysteresis in first-order equilibrium and nonequilibrium phase transitions', *Phys. Rev. A* **23**, 2719–2723.

Aicardi, F. (2001), 'On the classification of singularities in thermodynamics', *Physica D* **158**, 175–196.

Amann, A. and Atmanspacher, H. (1999), 'Introductory remarks on large deviation statistics', *Journal of Scientific Exploration* **13**, 639–664.

Antoni, M., Ruffo, S. and Torcini, A. (2002), 'First and second order clustering transitions for a system with infinite-range attractive interaction', *Phys. Rev. E* **66**, 025103.

Balescu, R. (1975), *Equilibrium and Nonequilibrium Statistical Mechanics*, Wiley, New York.

Balescu, R. (1997), *Statistical Dynamics : Matter Out of Equilibrium*, Imperial College Press, London.

Balian, R. (1991), *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics*, Vol. I, Springer-Verlag, Berlin.

Barré, J. (2002), 'Microcanonical solution of lattice models with long range interactions', *Physica A* **305**, 172–177.

Barré, J., Mukamel, D. and Ruffo, S. (2001), 'Inequivalence of ensembles in a system with long-range interactions', *Phys. Rev. Lett.* **87**, 030601.

Beck, C. and Schlögl, F. (1993), *Thermodynamics of Chaotic Systems: An Introduction*, Cambridge University Press, Cambridge.

Bender, C. M. and Orszag, S. A. (1978), *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York.

Binder, K. (1973), 'Time-dependent Ginzburg-Landau theory of nonequilibrium relaxation', *Phys. Rev. B* **8**, 3423–3438.

Binder, K. and Landau, D. P. (1984), 'Finite-size scaling at first-order phase transitions', *Phys. Rev. B* **30**, 1477–1485.

Binney, J. J., Dowrick, N. J., Fisher, A. J. and Newman, M. E. J. (1992), *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*, Oxford University Press, Oxford.

Blume, M., Emery, V. J. and Griffiths, R. B. (1971), 'Ising model for the λ transition and phase separation in He^3-He^4 mixtures', *Phys. Rev. A* **4**, 1071–1077.

Boltzmann, L. (1877), 'Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respecktive den Sätzen über das Wärmegleichgewicht (On the relationship between the second law of the mechanical theory of heat and the probability calculus)', *Wiener Berichte* **2(76)**, 373–435.

Borges, E. P. and Tsallis, C. (2002), 'Negative specific heat in a lennard-jones-like gas with long-range interactions', *Physica A* **305**, 148–151.

Borgs, C. and Janke, W. (1992), 'New method to determine first-order transition points from finite-size data', *Phys. Rev. Lett.* **68**, 1738–1741.

Borgs, C. and Kotecký, R. (1990), 'A rigorous theory of finite-size scaling at first-order phase transitions', *J. Stat. Phys.* **61**, 79–118.

Borgs, C. and Kotecký, R. (1992), 'Finite-size effects at asymmetric first-order transitions', *Phys. Rev. Lett.* **68**, 1734–1737.

Challa, M. S. S., Landau, D. P. and Binder, K. (1986), 'Finite-size effects at temperature-driven first-order transitions', *Phys. Rev. B* **34**, 1841–1852.

Chavanis, P. H. and Ispolatov, I. (2002), 'Phase diagram of self-attracting systems', *Phys. Rev. E* **66**, 036109.

Chomaz, P., Gulminelli, F. and Duflot, V. (2001), 'Topology of event distributions as a generalized definition of phase transitions in finite systems', *Phys. Rev. E* **64**, 046114.

Cover, T. M. and Thomas, J. A. (1991), *Elements of Information Theory*, John Wiley, New York.

Cramér, H. (1938), Sur un nouveau théorème-limite de la théorie des probabilités, *in* 'Actualités scientifiques et industrielles', Vol. 736, Hermann, Paris, pp. 2–23.

D'Agostino, M., Bougault, R., Gulminelli, F., Bruno, M., Cannata, F., Chomaz, P., Gramegna, F., Iori, I., Neindre, N. L., Margagliotti, G. V., Moroni, A. and Vannini, G. (2002), 'On the reliability of negative heat capacity measurements', *Nucl. Phys. A* **699**, 795–818.

D'Agostino, M., Gulminelli, F., Chomaz, P., Bruno, M., Cannata, F., Bougault, R., Colonna, N., Gramegna, F., Iori, I., Neindre, N. L., Margagliotti, G., Mastinu, P., Milazzo, P., Moroni, A. and Vannini, G. (2000), 'Negative heat capacity in the critical region of nuclear fragmentation: an experimental evidence of the liquid-gas phase transition', *Phys. Lett. B* **473**, 219–225.

Daniels, H. E. (1954), 'Saddlepoint approximations in statistics', *Ann. Math. Stat.* **25**, 631–650.

Dauxois, T., Holdsworth, P. and Ruffo, S. (2000), 'Violation of ensemble equivalence in the antiferromagnetic mean-field XY model', *Eur. Phys. J. B* **16**, 659.

Dauxois, T., Ruffo, S., Arimondo, E. and Wilkens, M., eds (2002), *Dynamics and Thermodynamics of Systems with Long Range Interactions*, Vol. 602 of *Lecture Notes in Physics*, Springer, New York.

Dembo, A. and Zeitouni, O. (1998), *Large Deviations Techniques and Applications*, 2nd edn, Springer, New York.

den Hollander, F. (2000), *Large Deviations*, Fields Institute Monograph, Amer. Math. Soc., Providence, R.I.

Dinwoodie, I. H. (1993), 'Identifying a large deviation rate function', *Ann. Prob.* **21**, 216–231.

Dinwoodie, I. H. and Zabell, S. L. (1992), 'Large deviations for exchangeable random vectors', *Ann. Prob.* **20**, 1147–1166.

Einstein, A. (1987), *in* J. Stachel, ed., 'The Collected Papers of Albert Einstein', Princeton University Press, Princeton.

Ellis, R. S. (1984), 'Large deviations for a class of random vectors', *Ann. Prob.* **12**, 1–12.

Ellis, R. S. (1985), *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York.

Ellis, R. S. (1995), 'An overview of the theory of large deviations and applications to statistical mechanics', *Scand. Actuarial J.* **1**, 97–142.

Ellis, R. S. (1999), 'The theory of large deviations: From Boltzmann's 1877 calculation to equilibrium macrostates in 2D turbulence', *Physica D* **133**, 106–136.

Ellis, R. S., Haven, K. and Turkington, B. (2000), 'Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles', *J. Stat. Phys.* **101**, 999–1064.

Ellis, R. S., Haven, K. and Turkington, B. (2002), 'Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows', *Nonlinearity* **15**, 239–255.

Ellis, R. S., Touchette, H. and Turkington, B. (2003), 'Thermodynamic versus statistical nonequivalence of ensembles for the mean-field Blume-Emery-Griffiths model'. Accepted for publication in *Physica A*; LANL archive: `cond-mat/0307007`.

Eyink, G. L. and Spohn, H. (1993), 'Negative-temperature states and large-scale, long-lived vortices in two-dimensional turbulence', *J. Stat. Phys.* **70**, 833–886.

Fisher, M. E. (1965), The nature of critical points, *in* W. E. Brittin, ed., 'Statistical Physics, Weak Interactions, Field Theory', Vol. VIII C of *Lectures in Theoretical Physics*, The University of Colorado Press, Boulder, pp. 1–159.

Gärtner, J. (1977), 'On large deviations from the invariant measure', *Th. Prob. Appl.* **22**, 24–39.

Gibbs, J. W. (1902), *Elementary Principles in Statistical Mechanics with Especial Reference to the Rational Foundation of Thermodynamics*, Yale University Press, Yale, C.T. Reprinted by Dover, New York, 1960.

Gilmore, R. (1979), 'Catastrophe time scale and conventions', *Phys. Rev. A* **20**, 2510–2515.

Gobet, F., Farizon, B., Farizon, M., Gaillard, M. J., Buchet, J. P., Carré, M., Scheier, P. and Märk, T. D. (2002), 'Direct experimental evidence for a negative heat capacity in the liquid-to-gas phase transition in hydrogen cluster ions: backbending of the caloric curve', *Phys. Rev. Lett.* **89**, 183403.

Griffiths, R. B. (1967), 'Thermodynamic functions for fluids and ferromagnets near the critical point', *Phys. Rev.* **158**, 176–187.

Griffiths, R. B., Weng, C.-Y. and Langer, J. S. (1966), 'Relaxation times for metastable states in the mean-field model of a ferromagnet', *Phys. Rev.* **149**, 301–305.

Gross, D. H. E. (1997), 'Microcanonical thermodynamics and statistical fragmentation of dissipative systems: the topological structure of the n-body phase space', *Phys. Rep.* **279**, 119–202.

Gross, D. H. E. (2001), *Microcanonical Thermodynamics: Phase Transitions in "Small" Systems*, Vol. 66 of *Lecture Notes in Physics*, World Scientific, Singapore.

Gulminelli, F. and Chomaz, P. (2002), 'Failure of thermodynamics near a phase transition', *Phys. Rev. E* **66**, 046108.

Gunton, J. D. and Droz, M. (1983), *Introduction to the Theory of Metastable and Unstable*

States, Vol. 183 of *Lecture Notes in Physics*, Springer-Verlag, Berlin.

Hertel, P. and Thirring, W. (1971), 'A soluble model for a system with negative specific heat', *Ann. Phys. (NY)* **63**, 520.

Huang, K. (1987), *Statistical Mechanics*, Wiley, New York.

Ioffe, D. (1993), 'Two examples in the theory of large deviations', *Stat. Prob. Lett.* **18**, 297–300.

Ispolatov, I. and Cohen, E. G. D. (2000), 'On first-order phase transitions in microcanonical and canonical non-extensive systems', *Physica A* **295**, 475.

Jaynes, E. T. (1957*a*), 'Information theory and statistical mechanics', *Phys. Rev.* **106**, 620–630.

Jaynes, E. T. (1957*b*), 'Information theory and statistical mechanics. II', *Phys. Rev.* **108**, 171–190.

Khinchin, A. I. (1949), *Mathematical Foundations of Statistical Mechanics*, Dover, New York.

Kiessling, M. K.-H. and Lebowitz, J. (1997), 'The micro-canonical point vortex ensemble: beyond equivalence', *Lett. Math. Phys.* **42**, 43–56.

LaCour, B. R. and Schieve, W. C. (2000), 'Macroscopic determinism in noninteracting systems using large deviation theory', *J. Stat. Phys.* **99**, 1225–1249.

Landau, L. D. and Lifshitz, E. M. (1991), *Statistical Physics*, Vol. 5 of *Landau and Lifshitz Course of Theoretical Physics*, 3rd edn, Butterworth Heinemann, Oxford.

Lanford, O. E. (1973), Entropy and equilibrium states in classical statistical mechanics, *in* J. Ehlers, K. Hepp and H. A. Weidenmüller, eds, 'Statistical Mechanics and Mathematical Problems: Battelle Seattle 1971 Rencontres', Vol. 20 of *Lecture Notes in Physics*, Springer-Verlag, Berlin, pp. 1–113.

Langer, J. S. (1967), 'Theory of the condensation point', *Ann. Phys. (N.Y.)* **41**, 108–157.

Langer, J. S. (1969), 'Statistical theory of the decay of metastable states', *Ann. Phys. (N.Y.)* **54**, 258–275.

Latora, V. and Rapisarda, A. (2001), 'Dynamical quasi-stationary states in a system with long-range forces', *Chaos, Solitons and Fractals* **13**, 401–406.

Latora, V., Rapisarda, A. and Tsallis, C. (2001), 'Non-Gaussian equilibrium in a long-range Hamiltonian system', *Phys. Rev. E* **64**, 056134.

Latora, V., Rapisarda, A. and Tsallis, C. (2002), 'Fingerprints of nonextensive thermodynamics in a long-range Hamiltonian system', *Physica A* **305**, 129–136.

Lax, M. (1955), 'Relation between canonical and microcanonical ensembles', *Phys. Rev.* **97**, 1419.

Lee, J. and Kosterlitz, J. M. (1991), 'Finite-size scaling and Monte Carlo simulations of first-order phase transitions', *Phys. Rev. B* **43**, 3265–3277.

Lee, T. D. and Yang, C. N. (1952), 'Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model', *Phys. Rev.* **87**, 410–419.

Lehtonen, T. and Nummelin, E. (1990), 'Level I theory of large deviations in the ideal gas', *Int. J. Theoret. Phys.* **29**, 621–635.

Lewis, J. T., Pfister, C.-E. and Sullivan, G. W. (1994*a*), 'The equivalence of ensembles for lattice systems: some examples and a counterexample', *J. Stat. Phys.* **77**, 397–419.

Lewis, J. T., Pfister, C.-E. and Sullivan, W. G. (1994*b*), Large deviations and the thermodynamic formalism: a new proof of the equivalence of ensembles, *in* M. Fannes, C. Maes and A. Verbeure, eds, 'On Three Levels', Plenum Press, New York.

Lewis, J. T., Pfister, C.-E. and Sullivan, W. G. (1995), 'Entropy, concentration of probability and conditional limit theorem', *Markov Proc. Rel. Fields* **1**, 319–386.

Lewis, J. T. and Russell, R. (1996), 'An introduction to large deviations for teletraffic engineers', *Available at* `http://www.stp.dias.ie/APG/`.

Leyvraz, F. and Ruffo, S. (2002), 'Ensemble inequivalence in systems with long-range interactions', *J. Math. Phys. A: Math. Gen.* **35**, 285–294.

Lopes, A. O. (1990), 'Entropy and large deviation', *Nonlinearity* **3**, 527–546.

Lynden-Bell, D. (1999), 'Negative specific heat in astronomy, physics and chemistry', *Physica A* **263**, 293–204.

Lynden-Bell, D. and Wood, R. (1968), 'The gravo-thermal catastrophe in isothermal spheres and the onset of red-giant structure for stellar systems', *Mon. Notic. Roy. Astron. Soc.* **138**, 495.

Maxwell, J. C. (1875), 'On the dynamical evidence of the molecular constitution of bodies', *Nature* **II**, 357. Reprinted in *The Scientific Papers of James Clerk Maxwell*, vol. II, W. D. Niven (ed.), Cambridge University Press, Cambridge, 1890.

Oono, Y. (1989), 'Large deviation and statistical physics', *Prog. Theoret. Phys. Suppl.* **99**, 165–205.

Pathria, R. K. (1996), *Statistical Mechanics*, 2nd edn, Butterworth Heinemann, Oxford.

Penrose, O. and Lebowitz, J. L. (1971), 'Rigorous treatment of metastable states in the van der Waals-Maxwell theory', *J. Stat. Phys.* **3**, 211–236.

Reif, F. (1965), *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York.

Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press, Princeton.

Salinas, S. R. A. (2001), *Introduction to Statistical Physics*, Springer, New York.

Sanov, I. N. (1961), On the probability of large deviations of random variables, *in* 'Select. Transl. Math. Stat. and Prob., Vol.', Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., pp. 213–244.

Schmidt, M., Kusche, R., Hippler, T., Donges, J., Kronmüller, W., von Issendorff, B. and Haberland, H. (2001), 'Negative heat capacity for a cluster of 147 sodium atoms', *Phys. Rev. Lett.* **86**, 1191–1194.

Stroock, D. W. and Zeitouni, O. (1991), Microcanonical distributions, Gibbs' states, and the equivalence of ensembles, *in* R. Durrett and H. Kesten, eds, 'Random walks, Brownian motion and interacting particle systems: a festschrift in honor of Frank Spitzer', Birkhäuser, Boston, pp. 399–424.

Thirring, W. (1970), 'Systems with negative specific heat', *Z. Physik* **235**, 339–352.

Thirring, W. (2002), *Quantum Mathematical Physics: Atoms, Molecules and Large Systems*, 2nd edn, Springer, New York.

Thompson, C. J. (1972), *Mathematical Statistical Mechanics*, MacMillan, New York.

Truesdell, C. (1980), *The Tragicomical History of Thermodynamics, 1822-1854*, Springer Verlag, New York.

van Campenhout, J. M. and Cover, T. M. (1981), 'Maximum entropy and conditional probability', *IEEE Trans. Info. Th.* **27**, 483–489.

Varchenko, A. N. (1990), 'Evolutions of convex hulls and phase transitions in thermodynamics', *J. Sov. Math.* **52**, 3305–3325.

Veneziano, D. (2002), 'Large deviations of multifractal measures', *Fractals* **10**, 117–129.

Wannier, G. H. (1966), *Statistical Physics*, Wiley, New York.

Yan, C. C. and Wannier, G. H. (1965), 'Observations on the sperical model of a ferromagnet', *J. Math. Phys.* **6**, 1833–1838.

Yang, C. N. and Lee, T. D. (1952), 'Statistical theory of equations of state and phase transitions. I. Theory of condensation', *Phys. Rev.* **87**, 404–409.

Zohar, G. (1999), 'Large deviations formalism for multifractals', *Stoc. Proc. Appl.* **79**, 229–242.

Zwanzig, R. (2001), *Nonequilibrium Statistical Mechanics*, Oxford University Press, Oxford.

Zwillinger, D., ed. (1996), *Standard Mathematical Tables and Formulae*, 30 edn, CRC Press, New York.

Index