

MICROSCOPIC FOUNDATIONS OF THERMODYNAMICS AND
GENERALIZED STATISTICAL ENSEMBLES

Michele Campisi

Dissertation Prepared for the Degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2008

APPROVED:

Donald. H. Kobe, Committee Chair
William. D. Deering, Committee Member
Duncan. L. Weathers, Committee Member
Jacek. M. Kowalski, Committee Member
Chris. L. Littler, Chair of the Department of
Physics
Sandra L. Terrell, Dean of the Robert B. Toulouse
School of Graduate Studies

Campisi, Michele, Microscopic Foundations of Thermodynamics and Generalized Statistical Ensembles. Doctor of Philosophy (Physics), May 2007, 115 pp., 13 illustrations, bibliography, 81 titles.

This dissertation aims at addressing two important theoretical questions which are still debated in the statistical mechanical community. The first question has to do with the outstanding problem of how to reconcile time-reversal asymmetric macroscopic laws with the time-reversal symmetric laws of microscopic dynamics. This problem is addressed by developing a novel mechanical approach inspired by the work of Helmholtz on monocyclic systems and the Heat Theorem, i.e., the Helmholtz Theorem. By following a line of investigation initiated by Boltzmann, a Generalized Helmholtz Theorem is stated and proved. This theorem provides us with a good microscopic analogue of thermodynamic entropy. This is the volume entropy, namely the logarithm of the volume of phase space enclosed by the constant energy hyper-surface. By using quantum mechanics only, it is shown that such entropy can only increase. This can be seen as a novel rigorous proof of the Second Law of Thermodynamics that sheds new light onto the arrow of time problem. The volume entropy behaves in a thermodynamic-like way independent of the number of degrees of freedom of the system, indicating that a whole thermodynamic-like world exists at the microscopic level. It is also shown that breaking of ergodicity leads to microcanonical phase transitions associated with nonanalyticities of volume entropy.

The second part of the dissertation deals with the problem of the foundations of generalized ensembles in statistical mechanics. The starting point is Boltzmann's work on

statistical ensembles and its relation with the Heat Theorem. We first focus on the nonextensive thermostatistics of Tsallis and the associated deformed exponential ensembles. These ensembles are analyzed in detail and proved (a) to comply with the requirements posed by the Heat Theorem, and (b) to interpolate between canonical and microcanonical ensembles. Further they are showed to describe finite systems in contact with finite heat baths. Their mechanical and information-theoretic foundation, are highlighted. Finally, a wide class of generalized ensembles is introduced, all of which reproduce the Heat Theorem. This class, named the class of dual orthodes, contains microcanonical, canonical, Tsallis and Gaussian ensembles as special cases.

Copyright 2007
by
Michele Campisi

ACKNOWLEDGMENTS

In compiling this dissertation I have benefited from many valuable comments and feedbacks coming from collaborators and other researchers in the field.

Many exchanges of ideas with Baris Bagci helped in refining my views on many arguments presented in the dissertation.

Most of the work presented here has been published in peer-reviewed journals. Feedback and comments from anonymous referees of *Studies in History and Philosophy of Modern Physics*, *Physics Letters A* and *Physica A* are thankfully acknowledged.

Most of all I must thank my supervisor, Dr. Donald H. Kobe, for the many interesting discussion we had on these topics. His constant enthusiasm and encouragement have been truly invaluable.

This dissertation would have never been possible without the love and the support of my wife Elisa.

CONTENTS

ACKNOWLEDGMENTS	iii
LIST OF FIGURES	vii
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. THE PROBLEM OF TIME REVERSAL ASYMMETRY OF THERMODYNAMICS	4
CHAPTER 3. HEAT THEOREM AND THE GENERALIZED HELMHOLTZ THEOREM	9
3.1. Surface vs. Volume Entropy	9
3.1.1. Properties of the Surface Entropy	10
3.1.2. Properties of the Volume Entropy	11
3.2. Helmholtz Theorem	14
3.3. Generalized Helmholtz Theorem	16
3.4. Concluding Remarks	23
CHAPTER 4. VOLUME ENTROPY AND THE LAW OF ENTROPY INCREASE	25
4.1. Quantum Volume Entropy	26
4.2. Proof of the Entropy Principle	27
4.3. Thomson's Formulation of the Second Law and the Minimal Work Principle	32
4.4. Comparison with other Quantum Entropies	34
4.5. Classical case	35

4.6. The role of the Initial Equilibrium	37
CHAPTER 5. APPLICATIONS OF VOLUME ENTROPY: MICROCANONICAL PHASE TRANSITIONS IN SMALL SYSTEMS	41
5.1. Overview of Microcanonical Phase Transitions	41
5.1.1. Why use Volume Entropy	42
5.2. Lennard-Jones Chain	43
5.3. Discussion	47
CHAPTER 6. OPEN ISSUES IN NONEXTENSIVE THERMODYNAMICS	51
CHAPTER 7. TSALLIS ENSEMBLES AND THE HEAT THEOREM	53
7.1. Definition of Orthode	53
7.2. Definition of Tsallis (Escort) Ensembles	54
7.3. Orthodicity of the Canonical Ensemble	55
7.4. Orthodicity of the Microcanonical Ensemble	56
7.5. Tsallis Escort Ensembles are Exact Orthodes	57
7.6. Remarks	59
7.6.1. Duality	59
7.6.2. Connection with Rényi Entropy	59
7.6.3. Mechanical Additivity	60
7.6.4. Exact Orthodicity	60
7.6.5. Free Energy	61
7.6.6. Connection with Tsallis Entropy	61
CHAPTER 8. MORE ON TSALLIS ESCORT ENSEMBLES	63
8.1. Interpolating between Canonical and Microcanonical Ensembles	63
8.2. Interpretation: Finite Heat Baths	66
8.3. Fluctuations in the Tsallis Escort Ensemble	69
8.4. Information-Theory and Tsallis Ensembles	71
8.4.1. Tsallis Escort Ensembles Maximize Rényi Entropy	71

8.4.2. Information Theoretic justification for Maximum Rényi Entropy Principle	72
8.5. New insights into the Microcanonical Entropy Principle	75
8.5.1. Canonical MEP	76
8.5.2. Microcanonical MEP	76
CHAPTER 9. GENERALIZED DUAL ENSEMBLES	78
9.1. Definition of Dual Orthodes	78
9.1.1. Recovery of known cases	81
9.2. Derivation of the Ideal Gas Thermodynamics	83
9.3. Examples	85
9.3.1. Tsallis Escort Ensembles	85
9.3.2. Gaussian Ensemble	86
9.3.3. Fermi-like Ensemble	89
9.4. Discussion	91
CHAPTER 10. CONCLUSIONS AND PERSPECTIVES	95
10.1. Conclusions	95
10.2. Perspectives	97
APPENDIX A. UNITARITY AND LINEARITY OF THE OPERATOR \mathcal{U}	100
APPENDIX B. CONSTRAINED MAXIMIZATION OF VOLUME ENTROPY	102
APPENDIX C. SOME CALCULATIONS REGARDING DUAL ENSEMBLES	105
C.1. Proof of the Generalized Equipartition Theorem of Eq. (180)	106
C.2. Proof of Orthodicity in the β, V parametrization	106
C.3. Large n behavior of the Coefficients $A_{n,\sigma}$	107
C.4. Large n behavior of the Coefficients A_n for the Fermi-like Case	108
REFERENCES	109

LIST OF FIGURES

- 4.1 Visual representation of the effect of a non adiabatic perturbation on a quantum system which is initially described by decreasing probabilities. After shaking, the initial accumulation towards the left would flatten out, and the average value of $\ln(n + 1/2)$ (i.e. the entropy) would increase. 38
- 5.1 Lennard-Jones box potential profile $\varphi(x, V)$ for $V = 4 > V_c$. 45
- 5.2 Phase space structure for a particle of mass $m = 1$ in a Lennard-Jones box of size $V = 4 > V_c$. The separatrix corresponds to the critical energy $E_c = -0.0308$. Below the critical energy there are two distinct trajectories (the dynamics is not ergodic). 45
- 5.3 Entropy versus energy for a particle of mass $m = 1$ in a Lennard-Jones box of size $V = 4 > V_c$. The discontinuity of entropy at the critical energy signals a discontinuous phase transition. 46
- 5.4 Temperature versus Energy for a particle of mass $m = 1$ in a Lennard-Jones box of size $V = 4 > V_c$. The curve displays a region of negative heat capacity. At the critical energy the temperature goes to zero. 47
- 5.5 Entropy versus Energy for $E < E_c$. Right below the critical energy the entropy function (thick line) is convex (the thin straight line is only a guide for the eye). 48
- 8.1 Plot of q-exponential (Eq. 130) for large negative values of q . The function tends to the Heaviside step function. 64
- 8.2 Plot of q-exponential derivative (Eq. 133) for large negative values of q . The function approaches a Dirac delta function. 65

8.3 Two-fold foundations of Tsallis escort ensembles: mechanical (based on Boltzmann's orthodes method) and information-theoretic (based on the maximum Rényi entropy principle)	73
9.1 Solution of Eq. (209), for different values of σ and n	88
9.2 Gaussian fit of Tsallis statistics for $\alpha_q = 50, n = 60, U = 10$. The dual Gaussian ensemble approximates Tsallis escort ensembles for large values of n and α_q	89
9.3 Solution of Eq. (214) for different values of n .	90
9.4 Generalization scheme. Tsallis escort ensembles, Gaussian ensembles, Fermi-like ensemble, and possibly other kinds of ensembles belong to the class of dual orthodes. Canonical and microcanonical are special instances of Tsallis and Gaussian ensembles. For large systems ($n \gg 1$) Gaussian and Tsallis escort ensemble coincide if $\sigma = \frac{1}{1-q} = C_V \gg 1$.	92

CHAPTER 1

INTRODUCTION

This dissertation collects calculations, results and ideas, that stemmed in the last few years from a continuous meditation about a few central issues that the theory of statistical mechanics naturally poses. In order to give the reasons for studying the “Microscopic Foundations of Thermodynamics and Generalized Statistical Ensembles” it is perhaps best to tell how and why I became involved in this subject. As a student of statistical physics I was never really satisfied with the textbook justification of “Boltzmann’s principle”:

$$(1) \quad S_W = k_B \ln W$$

As a matter of fact this is always presented as a postulate, namely something that is not derived from some other more fundamental principles, and that everybody accepts on the basis of experience. Although I did not have any reason to doubt its validity, I still could not accept “Boltzmann’s principle” as a fundamental principle. The questions that kept running through my mind were: why is Clausius entropy related to the number of available states, and why is that relation logarithmic? The questions are perfectly legitimate. Besides the undisputable phenomenological validity of Eq. (1) and its connection with Boltzmann’s H-theorem (which refers to rarefied gases only), textbooks often do not really provide any bold theoretical foundation for it¹. Reading about Boltzmann’s studies on the *heat theorem* from Gallavotti’s *Short Treatise of Statistical Mechanics* [2] made me realize that a deeper reason for Eq. (1) actually existed. Gallavotti’s book has the merit of emphasizing the role played by the heat

¹A.I. Khinchin had to comment about this that “All existing attempts to give a general proof of this postulate must be considered as an aggregate of logical and mathematical errors superimposed on a general confusion in the definition of the basic quantities” [1]

theorem in Boltzmann's construction of statistical mechanics. In particular, if one considers the continuum counterpart of S_W , that is the logarithm of the area of the hypersurface of constant energy, $S_\Omega = \ln \Omega$, Gallavotti shows that

$$(2) \quad dS_\Omega = \frac{dE + PdV}{T} + O\left(\frac{1}{N}\right),$$

where P is the pressure, V is the volume, T is the temperature and N is the number of particles composing the gas. This began to enlighten my understanding of why the Boltzmann entropy is fundamental: in the thermodynamic limit ($N \rightarrow \infty$) it generates the exact differential $\frac{\delta Q}{T}$, which is not at all a trivial result. This fact justifies the use of Boltzmann entropy in deriving the thermodynamic relations by differentiating it with respect to energy and volume. At this point another theorem discussed by Gallavotti attracted my attention, namely the Helmholtz theorem (see Sec. 3.2). This says that $\frac{dE+PdV}{T}$ is an exact differential even in small one-dimensional systems, thus suggesting that the thermodynamic laws already exist at the microscopic level. Noticing this motivated the subsequent search for a good microscopic or "mechanical" analogue of entropy which forms a bridge from the Helmholtz theorem, valid for one degree of freedom, to Eq. (2), valid for a large number of degrees of freedom. The result is the generalized Helmholtz theorem (see Sec. 3.3). Its associated entropy is the volume entropy, namely the logarithm of the portion of phase space *enclosed* by the energy hypersurface, not its surface area!

At this point the research took two different parallel paths. On one hand, since the volume entropy is an adiabatic invariant, one is naturally led to guess that for non adiabatic perturbations it can only increase. Thus I started wondering under what hypothesis and in what sense this could be proved. The result is a novel proof of the second law which is presented in Chapter 3. On the other hand, since Boltzmann used the heat theorem as a rationale for canonical and microcanonical ensembles, one is naturally led to check if other ensembles also satisfy it. I checked with the Tsallis ensembles associated with the so called nonextensive thermodynamics, and

surprisingly enough, they did so! The second part of the dissertation deals with this fact. I also define an even more generalized class of ensembles that comply with the requirements posed by the heat theorem.

The dissertation can ideally be divided in two parts. The first, composed of Chapters 2 to 5, covers The Helmholtz theorem, the generalized Helmholtz theorem, the the proof of the second law based on volume entropy and more applications of volume entropy. The second, composed of Chapters 6 to 9, covers the Tsallis Ensembles, their “mechanical” foundations and their generalizations. In Chapter 2 and Chapter 6 the reader can find more technical introductions to the first and second part of the dissertation, respectively.

Most of the material presented here has been published in refereed Journals. Part of it is currently under consideration for publication. Roughly, the content of Chapter 2 has been adapted from Ref. [3], Chapter 3 from Ref. [4], Chapter 4 from Ref. [5] (currently unpublished), Chapter 6 from Ref. [6], Chapter 7 from Ref. [7], and Chapter 8 and Appendix C from Ref. [8]. Section 8.5 and Appendix B contain material covered in Ref. [9] (currently unpublished).

These studies have largely contributed to sharpening my own understanding of statistical mechanics as well as thermodynamics, both ordinary and nonextensive. It is my hope that they will also be of help and interest to others as well.

CHAPTER 2

THE PROBLEM OF TIME REVERSAL ASYMMETRY OF THERMODYNAMICS

Explaining the time reversal asymmetric nature of macroscopic phenomena such as those entailed in the second law of thermodynamics from the time reversal symmetric laws of microscopic dynamics is one of the most challenging problems of modern theoretical physics. Since the birth of statistical mechanics with Boltzmann and Gibbs, scientists in this field have been continuously working on this puzzling problem. A wide range of approaches and views have been developed which have all contributed, in different ways, to the development of the multi-faceted, continuously-developing field of statistical physics. Boltzmann was the first to attack the problem with his celebrated H-Theorem. In his approach the time reversal asymmetry stems from the assumption of molecular chaos (*Stoßzahl Ansatz*). Along the years many others have tried to explain the second law using different approaches. Tolman for example first proposed the idea of *coarse graining* as a mechanism that breaks time reversal invariance. In his famous book [10], the increase of coarse grained entropy is proved for infinitesimal abrupt changes of an external parameter that does work on a macroscopic system. Yet more recently the Brussels-Austin school of Y. Prigogine proposed another solution based on Rigged Hilbert Spaces (see [11] for a recent review of the Brussels-Austin school). For some authors irreversibility stems from the very large number of degrees of freedom that occur in macroscopic systems. For others it is a consequence of non-integrability of the equations of motion. For still others it is our limited possibility to know the exact time evolution of macroscopic systems that

determines the time-reversal symmetry breaking. Chaos and sensitivity to initial conditions are believed to play an important role, too. One thing is sure: there is no universal consensus about this problem.

In this first part of the thesis we propose yet another approach toward the solution of this puzzling problem. In order to do so, we shall go back to the works of Boltzmann on monocyclic systems [12]. These works were inspired by previous works of Helmholtz [13] on one-dimensional Hamiltonian systems, whose main result is known (not as widely as one would expect though) as the Helmholtz theorem¹. These works of Boltzmann were very closely connected to his celebrated *ergodic hypothesis*. If we look back at the development of Boltzmann's work in statistical physics [14] we see that he did not develop further those ideas and turned his attention to other methods like the well-known counting method which led him to the famous formula $S_W = \ln W$ and the already mentioned H-Theorem. Here we shall revive that initial research line of Boltzmann which he did not investigate further. We will undertake the task of developing it to the point of appreciating its actual relevance for the still open problem of reconciling the second law with mechanics (classical or quantum).

Before embarking on such an adventure it is important to ask the question: What do we exactly mean here by *second law*? This may sound like a trivial question, but it is not indeed. As pointed out nicely in Ref. [15] there are many different formulations of the second law that are not all exactly equivalent. Further the second law lends itself to possible hasty conclusions which are very controversial. One such conclusion is Clausius' celebrated statement: "The entropy of the universe tends to a maximum" [15]. In order to make our study as clear as possible we shall refer to the second law as the principle of entropy increase as formulated by Clausius [15]:

THE ENTROPY PRINCIPLE: For every *nicht umkehrbar* process in a thermally isolated system which begins and ends in an equilibrium state, the entropy of the final state is greater than or

¹Not to be confused with the better known Helmholtz theorem of vector calculus

equal to that of the initial state. For every *umkehrbar* process in a thermally isolated system the entropy of the final state is equal to that of the initial state

The expressions *nicht umkehrbar* and *umkehrbar* could be translated into the current scientific English as *non-quasi-static* and *quasi-static* respectively.² It must be stressed that the entropy principle refers to transformations caused by the variation of some external parameter, and is not at all a statement about the *spontaneous tendencies* of physical systems. If the variation of the external parameter acts in such a way as to drive the system out of equilibrium (non-quasi-static process) the entropy will increase. If it acts in such a way that the system remains arbitrarily close to equilibrium (quasi-static process) then the entropy will not change.

A few commentators have expressed the idea that the second law might not entail any time-asymmetry [16]. According to them the time asymmetry stems from another phenomenological fact that is often confused with the second law itself, namely

THE EQUILIBRIUM PRINCIPLE: An isolated system in an arbitrary initial state within a finite fixed volume will spontaneously attain a unique state of equilibrium.

The authors of Ref. [16] refer to this basic principle as the *minus first law of thermodynamics* to stress the fact that it is more fundamental than all the other thermodynamic laws. In fact the Zeroth, First and Second laws all make use of the concept of equilibrium.

In this part of the thesis we will address the second law as expressed by the entropy principle enunciated above. For the sake of completeness and simplicity we shall re-express it as the following three statements:

²Often the expressions *irreversible* and *reversible* are used as synonymous of non-quasi-static and quasi-static

Statement 1 (Heat Theorem). There exists a state function S , called entropy, such that

$$(3) \quad dS = \frac{\delta Q}{T}$$

where dS is an exact differential, δQ is the heat exchanged (not an exact differential) and T is the absolute temperature.

Statement 2 (Adiabatic Invariance). For any quasi-static process occurring in a thermally isolated system that begins and ends in an equilibrium state we have

$$(4) \quad \Delta S = 0.$$

Statement 3 (Entropy increase). For any non quasi-static process occurring in a thermally isolated system that begins and ends in an equilibrium state we have

$$(5) \quad \Delta S \geq 0.$$

The first two statements evidently pertain to *equilibrium thermodynamics* whereas the third pertains to *non-equilibrium thermodynamics*. We shall address the first two statements in Chapter 3 and the third in Chapter 4. As we will see, our approach based on Boltzmann's early studies on *monocyclic* systems leads naturally to identify the statistical mechanical analog of thermodynamic entropy as the so-called *volume entropy*. Such an entropy will provide a quite satisfactory (although perhaps not completely exhaustive) explanation of all the three statements above. New light will be shed also on the previously mentioned idea that the second law is itself time-reversal symmetric. The arrow of time stems at the level of equilibrium attainment, which chronologically and logically comes before the processes occur that are mentioned in the entropy principle (which are responsible for the entropy increase).

In Chapter 5 we shall see more applications of such entropy which are not related to the fundamental laws of thermodynamics, but rather to some other very important macroscopic phenomena, i.e., phase transition. We will see how the volume entropy

can also be employed to implement a simple model of phase transitions without invoking the *thermodynamic limit* in which the number of degrees of freedom tends to infinity.

CHAPTER 3

HEAT THEOREM AND THE GENERALIZED HELMHOLTZ THEOREM

3.1. Surface vs. Volume Entropy

Since the pioneering works of Boltzmann and Gibbs two possibilities have been given for the microcanonical analysis of physical systems, which correspond to two different definitions of entropy (see for example the textbook of Gibbs [17] or the more recent textbook of Huang [18]). We shall refer to these two entropies as “surface entropy” and “volume entropy”. The *surface entropy* is defined as:

$$(6) \quad S_{\Omega}(E, V) = \ln \Omega(E, V)$$

where E is the energy, V the volume of the system, and

$$(7) \quad \Omega(E, V) = \int \frac{d\mathbf{z}}{h^{3N}} \delta [E - H(\mathbf{z}, V)].$$

The *volume*¹ entropy is:

$$(8) \quad S_{\Phi}(E, V) = \ln \Phi(E, V)$$

where

$$(9) \quad \Phi(E, V) = \int \frac{d\mathbf{z}}{h^{3N}} \theta [E - H(\mathbf{z}, V)].$$

Throughout this dissertation we adopt a system of units where k_B , the Boltzmann constant, is equal to 1. The symbol $\delta(x)$ represents the one-dimensional Dirac delta-function and $\theta(x)$ represents the Heaviside step-function. The symbol $\mathbf{z} = (\mathbf{p}, \mathbf{q}) = (p_1, \dots, p_{3N}, q_1, \dots, q_{3N})$ is a point in the $6N$ -dimensional phase space Γ . The Hamiltonian is $H(\mathbf{z}, V) = K(\mathbf{p}) + \varphi(\mathbf{q}, V)$, where $K(\mathbf{p}) = \frac{\mathbf{p}^2}{2m}$ and the potential φ includes both particle-particle and container-particle interactions. The container “coordinate”

¹Sometimes referred to as “bulk”.

(i.e., its volume V) is treated as an “external” parameter. h is a constant with the dimensions of an action introduced as a convention ². The quantity $\Omega(E, V)$ represents the volume of the infinitesimally thin shell of constant energy E in the phase space Γ (surface integral), while $\Phi(E, V)$ represents the volume of the region enclosed by the hyper-surface of constant energy E (volume integral). Following Ref. [1, see p.32], we assume the energy has a lower bound, which for convenience will be set equal to zero. With a further assumption of “smoothness” of the hyper-surfaces of constant energy, the following relation holds:

$$(10) \quad \Omega = \frac{\partial}{\partial E} \Phi.$$

For ideal systems (for which inter-particle interaction energy is negligible) with a very large number of constituents N , the volume integral tends to the exponential form $\Phi \propto e^E$ [18], hence one has

$$(11) \quad \Omega \rightarrow \Phi \text{ as } N \rightarrow \infty.$$

This relation expresses the well-known fact that the two possible microcanonical descriptions associated, respectively, with the volume and surface entropies are *equivalent*. Far from the thermodynamic limit and for systems with long-range interaction the two entropies can lead to very different results. One example of this *inequivalence* will be offered in Chap. 5. For the purposes of this dissertation, it is worth discussing some properties of these two entropies in more detail. This discussion helps to confront them, and to introduce the Helmholtz theorem, whose generalization is the main result that we present in this chapter.

3.1.1. *Properties of the Surface Entropy*

From a formal point of view, the surface entropy plays an important role in statistical mechanics for its close connection with the microcanonical distribution. Defining

²I do not include the celebrated Gibbs correction factor $\frac{1}{N!}$ for it is not essential in this context.

the microcanonical distribution as:

$$(12) \quad \rho_\mu(\mathbf{z}, E, V) \frac{d\mathbf{z}}{h^{3N}} = \frac{\delta[E - H(\mathbf{z}, V)]}{\text{normalization}} \frac{d\mathbf{z}}{h^{3N}},$$

the normalization factor would be given exactly by the surface integral Ω . This is why often Ω is referred to as the partition function of the microcanonical ensemble. Thanks to Eq. (12) the surface integral enters explicitly in the expression of the microcanonical average of a physical quantity on the surface of constant energy, i.e.,

$$(13) \quad \langle f \rangle_\mu \equiv \frac{\int \frac{d\mathbf{z}}{h^{3N}} f(\mathbf{z}) \delta[E - H(\mathbf{z}, V)]}{\Omega(E, V)}.$$

From a mathematical perspective Ω plays also the role of the Jacobian of the change of variables $\mathbf{z} \rightarrow E$, which simplifies the evaluation of average values of observables of the type $f(E(\mathbf{z}))$:

$$(14) \quad \frac{d\mathbf{z}}{h^{3N}} = \Omega(E, V) dE.$$

Because of this relation Ω is also important in the passage from classical to quantum statistical mechanics; in fact Ω counts the “density of states at a given energy” (degeneracy).

From a physical point of view, the surface entropy is interesting as, in the limit of very large N , it approaches Boltzmann’s counting entropy ($S_W = \ln W$) of the given equilibrium state. Nevertheless the counting entropy is at the same time more and less general than the surface entropy. It is more general in the fact that it applies both to equilibrium and out of equilibrium states, whereas the surface entropy is restricted to equilibrium; less general in the fact that, unlike the surface entropy, it applies only to ideal gases [14].

3.1.2. *Properties of the Volume Entropy*

After the seminal work of P. Hertz [19], it is a known fact that the volume integral Φ is an adiabatic invariant, and it has been recognized that this very fact plays a significant role in the mechanical foundations of thermodynamics. To show this, let

us first provide a definition of adiabatic invariant. Let $H(\mathbf{z}, V)$ be the Hamiltonian of a system which depends on an external parameter whose value changes in time according to some law: $V = V(t)$, and let τ be the characteristic time of variation of V . Through the dependence on V the Hamiltonian depends explicitly on time, hence the energy is not conserved (work is performed on the system by changing the parameter V from outside), and accordingly the energy will change in time with some temporal law $E(t)$.

DEFINITION 3.1. A function $I(E, V)$ is called an *adiabatic invariant* if, in the limit of very slow variation of $V(t)$, namely as $\tau \rightarrow \infty$, $I(E(t), V(t)) \rightarrow \text{const.}$

Note that the qualifier “adiabatic” in the context of Hamiltonian mechanics has quite a different meaning than in thermodynamics, where it is used as synonymous of “thermally isolated”. For example, with reference to Definition 3.1, in mechanics an adiabatic transformation is a very slow transformation during which the adiabatic invariant remains constant. This meaning of “adiabatic” matches with that of the thermodynamic expression “quasi-static”, rather than “thermally isolated”.³ In other words, in studying thermodynamic systems as many-body mechanical systems interacting with an externally driven field of forces, it is quite reasonable to model quasi-static transformations as adiabatic processes [21, see Sec. 2.3.2]. It is natural, then, to model the quantities which remain constant during quasi-static transformations as mechanical adiabatic invariants.

Now consider Statement 2 of the entropy principle (see Chap. 2), $\Delta S = 0$, and recall that it refers to quasi-static transformations. It is a basic fact of thermodynamics that along a quasi-static transformation between two equilibrium states of a thermally isolated system, the entropy (which, for the exactness of the heat differential,

³This is also reflected by the fact that a quasi static-process is understood as such a slow process that it can be thought of as a virtual sequence of equilibrium states, just as the adiabatic process is so slow that the motion can be thought of as driven by a sequence of “frozen”(i.e., time independent) Hamiltonians [20].

is a function of E and V) is a constant. In modelling this fact from a mechanical perspective, as Hertz has done, it is natural to require the mechanical expression of entropy to be an adiabatic invariant. Hertz had also shown that, for any number of degrees of freedom, the volume integral (Eq. 9) is an adiabatic invariant *provided that the system is ergodic*.⁴ This means that the candidate mechanical expression of entropy should be a function of Φ : $S = f(\Phi)$. If one takes a look at the equipartition theorem (which is valid independent of the dimensionality of the system as well), the reason why Hertz was led to choose the entropic function to be of the form of a logarithmic function becomes clear. Note in fact that the equipartition theorem [1, p. 104] reads⁵:

$$(15) \quad \langle K \rangle_\mu = \frac{1}{2} \left\langle \mathbf{p} \cdot \frac{\partial H}{\partial \mathbf{p}} \right\rangle_\mu = \frac{3N}{2} \frac{\Phi(E, V)}{\Omega(E, V)}$$

or equivalently (using Eq. (10)):

$$(16) \quad \left(\frac{2 \langle K \rangle_\mu}{3N} \right)^{-1} = \frac{\partial}{\partial E} \ln \Phi(E, V).$$

From this perspective, the equipartition theorem expresses a very general dynamical-geometrical property of Hamiltonian systems which, as soon as we agree to name the quantity $\frac{2 \langle K \rangle_\mu}{3N}$ as the “absolute temperature”, reads as the fundamental thermodynamic relation $T^{-1} = \frac{\partial S}{\partial E}$. As a natural consequence, we have to agree to name the quantity $\ln \Phi(E, V)$ the “entropy”. In sum, *for ergodic systems*, the volume entropy has two remarkably good properties: it is an adiabatic invariant, that is, it reproduces Statement 1 of the second law, and it is consistent with the equipartition theorem. The approach of Hertz has been recently re-expressed by Berdichevsky [23] adopting a modern viewpoint. In particular Berdichevsky showed that under the hypothesis of *metric indecomposability* (see Definition 2 below) the volume entropy is an adiabatic invariant and that any other adiabatic invariant is necessarily a function of Φ only.

⁴This is why Φ is sometimes referred to, in literature, as *the ergodic adiabatic invariant* [22, 20].

⁵This holds for free gases as well as for interacting systems

This latter statement leaves no room for alternative definitions of entropy within the approach of Hertz and it is in contrast with the claim by Toda *et.al.* [21], that the surface integral is also an adiabatic invariant. Indeed, that the surface integral is not in general an adiabatic invariant can be seen by considering as a counter example the 1D harmonic oscillator (which is trivially ergodic) with slowly varying frequency ω . The ergodic adiabatic invariant in this case would be the action variable $\Phi = \frac{E}{\omega}$ [24]. Therefore the surface integral $\Omega = \frac{\partial}{\partial E}\Phi = \frac{1}{\omega}$, would be trivially non-constant. However, on the basis of Eq. (11), in the thermodynamic limit and for ideal systems, it is expected that the surface entropy approaches adiabatic invariance.

3.2. Helmholtz Theorem

In the previous section we have pointed out that the volume entropy provides a good mechanical analogue of thermodynamic entropy at least for Statement 2 of the second law. As we will see next, the volume entropy also accounts quite well for Statement 1 too. First we are going to introduce an interesting, but unfortunately not very well known, theorem in 1D Hamiltonian mechanics due to Helmholtz. The aim of this theorem is to provide a mechanical model of Statement 1 of the second law of thermodynamics for one-dimensional systems. Boltzmann intended to generalize this theorem to multi-dimensional systems in order to unify micro- and macro- physics into one consistent picture using ergodicity as the key tool. In the next section we are going to revive Boltzmann's argument and provide a generalization of the Theorem, based on the modern ergodic theory.

The Helmholtz theorem applies to one-dimensional conservative systems in a confined potential where there is only one periodic trajectory per energy level, i.e., it applies to what Helmholtz called one-dimensional *monocyclic* systems [2, p. 38]. To cite some examples, the harmonic oscillator and a particle in a box are one-dimensional monocyclic systems, but a particle inside a double-well potential is not because under

a certain energy value there are two distinct trajectories per energy level⁶. Roughly speaking, the theorem applies to one-dimensional Hamiltonian systems with a U-shaped potential. Let us recall the theorem without proof:

Theorem 1 (Helmholtz). Let $H(p, q; V) = \frac{p^2}{2m} + \varphi(q; V)$ be the Hamiltonian of a one-dimensional *monocyclic* system. Let a *state* be characterized by the set of quantities:

$$\begin{aligned}
 (17) \quad & E = \text{total energy} = K + \varphi \\
 & T = \text{twice the time average of the kinetic energy} = 2 \langle K \rangle_t \\
 & V = \text{the external parameter} \\
 & P = \text{minus the time average of } \frac{\partial \varphi}{\partial V} = \langle -\frac{\partial \varphi}{\partial V} \rangle_t.
 \end{aligned}$$

Then the differential

$$(18) \quad \frac{dE + PdV}{T}$$

is exact⁷, and $S_H(E, V)$ ⁸, defined as:

$$(19) \quad S_H(E, V) = \ln 2 \int_{x_-(E, V)}^{x_+(E, V)} \frac{dx}{h} \sqrt{2m(E - \varphi(x, V))}$$

is the generating function, i.e.:

$$(20) \quad dS_H = \frac{dE + PdV}{T}$$

where the symbols $x_{\pm}(E, V)$ denote the turning points of the trajectory, i.e., the roots of the equation $E - \varphi(x, V) = 0$, and time averages are calculated over one period τ of the trajectory $(x(t), p(t))$:

$$(21) \quad \langle f(x, p) \rangle_t \doteq \frac{1}{\tau} \int_0^{\tau} dt f(x(t), p(t))$$

As the reader can easily note, this theorem says that there is a class of mechanical systems which provide a model of thermodynamic behavior as described in Statement

⁶A double-well potential system will be discussed in Chap. 5 to illustrate the mechanism of microcanonical phase transitions as captured by the volume entropy when ergodicity breaks down

⁷This is often referred to as the *heat theorem* [2].

⁸The subscript H stands for Helmholtz.

1 of the second law of thermodynamics. Once we attach the thermodynamic names of “temperature” and “entropy” to T and S_H , respectively⁹, the theorem reads exactly as Statement 1 of the second law of thermodynamics (see Chap.2). For such systems, the heat theorem holds as a consequence of Hamiltonian mechanics alone. Later we will see that the same can be said of multidimensional systems provided certain assumptions are made. In developing the generalization we are going to adopt a historical point of view. In particular we are going to consider Boltzmann’s ideas about the extension of Helmholtz’s result to many particle systems and we are going to develop those. Our starting point will be Gallavotti’s reconstruction of Boltzmann’s ergodic hypothesis [2, see Sec. 9.1 and Appendix 9.A.3].

3.3. Generalized Helmholtz Theorem

3.3.0.1. *Discrete Version.* According to Gallavotti [2, see Sec. 9.1 and App. 9.A.3], based on a discrete picture of phase space, Boltzmann saw how the Helmholtz theorem could be applied to multi-dimensional systems. The argument of Gallavotti goes as follows: imagine that we divide the phase space into cells of the same size and then consider the time evolution of the system as a map which transforms a cell into another, that is, a *permutation* of the phase space cells on the hyper-surface of constant energy. This permutation is in general *decomposable* into sub-permutations, or *cycles*, involving only a fraction of the total number of cells. This happens, for instance, when extra integrals of motion are present besides the energy: cells with different values of these integrals do not transform into each other, but only among themselves to define the aforementioned cycles. The discrete ergodic hypothesis of Boltzmann-Gallavotti (the reader may find a detailed account in [2, Sec. 1.3]) is that there is only *one* cycle per energy level, i.e., that the discrete trajectory visits *all* the phase space points (cells) with a given energy. A system satisfying such a hypothesis will be referred to as a discrete ergodic system in this paper. For such systems, the

⁹ P can be proved to be the average momentum transfer to the confining walls per unit time, therefore it has its own mechanical interpretation of pressure.

motion can be followed along a curvilinear abscissa running along the trajectory itself, reducing the dynamics of the multi-dimensional system to that of a one-dimensional *monocyclic* system to which the Helmholtz theorem can be applied. More explicitly, let us discretize the phase space into hyper-cubic cells of volume $h^{3N} = (\Delta q \Delta p)^{3N}$. Consider the trajectory in the discrete configuration space as an ordered sequence of cells labelled by a counter j , and let $\ell = j\Delta q$ be the distance travelled by the representative point (ℓ is, so to speak, a “discrete curvilinear abscissa”). Let us now define the sequence

$$(22) \quad \ell \rightarrow \mathbf{q}(\ell)$$

which assigns the cell centered at \mathbf{q} to the travelled distance ℓ . The representative point spends different times in different cells. Denoting the sojourn time in the j^{th} cell with the symbol Δt_j allows us to define the velocity $\dot{\ell} \doteq \frac{\Delta q}{\Delta t_j}$ associated with ℓ . For such a discrete system then, the conservation of energy statement would read

$$(23) \quad \frac{1}{2m} p_\ell^2 + \varphi(\mathbf{q}(\ell), V) = E.$$

where $p_\ell = m\dot{\ell}$. If the original system is a discrete ergodic system, the above equation describes the motion of a one-dimensional system living in the phase space (ℓ, p_ℓ) , where it draws *one* closed trajectory for each energy specification: namely it is a one-dimensional *monocyclic* system. Therefore, the Helmholtz theorem applies, suggesting that

“Thermodynamic relations are [...] very general and simple consequences of the structure of the equations of motion: they hold for small and large systems, from one degree of freedom [...] to 10^{23} degrees of freedom [...]” [2, p. 266]

which is a remarkable result indeed because it implies that an expression for the entropy in terms of the Hamiltonian must exist for systems of any size. Such a mechanical expression of entropy, namely the multi-dimensional counterpart of the

Helmholtz entropy (Eq. (19)) is missing in the Boltzmann-Gallavotti argument, but it can be found by continuing their argument. Let us express the Helmholtz entropy associated with a discrete ergodic system as¹⁰:

$$(24) \quad S_H(E, V) = \ln \int_0^{\mathcal{N}\Delta q} \frac{d\ell}{h} \sqrt{2m [E - \varphi(\mathbf{q}(\ell), V)]},$$

where $\mathcal{N}\Delta q$ is the total distance travelled in one cycle. Note that the right hand side of Eq. (24) depends explicitly on the sequence $\mathbf{q}(\ell)$ which afterwards can be computed only after integrating the complete set of Hamilton equations, which, as N gets large, becomes a practically impossible task. In order to evaluate S_H we will use the following “trick”. As noticed before, the dynamics of the system can be alternatively depicted either in the $6N$ dimensional phase space Γ or in the 2 dimensional phase space (ℓ, p_ℓ) . If we express the kinetic energy in both representations, we get:

$$(25) \quad \frac{1}{2m} p_\ell^2 = \frac{1}{2m} \sum_{i=1}^{3N} p_i^2.$$

By taking the time average and then applying the equipartition theorem (Eq. (16)) to both sides of Eq. (25) we get (remembering that on the left we are dealing with a 1-dimensional system, whose curve of constant energy “encloses” an area A , and on the right we have a $3N$ -dimensional system whose hyper-surface of constant energy encloses a volume Φ):

$$(26) \quad 3N \frac{\partial}{\partial E} \ln A(E, V) = \frac{\partial}{\partial E} \ln \Phi(E, V),$$

Then, by integration (with the condition $A^{3N}|_{E=0} = \Phi|_{E=0}$) follows the simple formula:

$$(27) \quad \ln \Phi = 3N \ln A.$$

¹⁰Here the symbol $\int_0^{\mathcal{N}\Delta q} d\ell f(\ell)$ is a convenient notation for $\Delta q \sum_{j=0}^{\mathcal{N}} f(j\Delta q)$.

Note that the area under the trajectory in the phase space of a one-dimensional system can be expressed as the action integral [24]:

$$(28) \quad A = \oint \frac{p_\ell d\ell}{h}.$$

On the other hand, $p_\ell = \sqrt{2m [E - \varphi(\mathbf{q}(\ell), V)]}$; hence $\ln A$ is the Helmholtz entropy and therefore:

$$(29) \quad S_\Phi(E, V) = 3NS_H(E, V).$$

This result is surprising as it reveals that (apart from the multiplicative factor) the Helmholtz entropy of a discrete ergodic system is nothing but the volume entropy. Therefore the volume entropy generates the heat differential of discrete ergodic systems. This result is important because it shows that the discrete version of volume entropy satisfies Statement 1 of the second law for discrete ergodic systems as a consequence of the microscopic equation of motion. This completes Boltzmann-Gallavotti's ideas (see the quotation above) by providing the missing entropy function.

3.3.0.2. *Continuum Version.* In order to pass to the continuum let us focus on Eq. (29) and note that, though S_H requires a discrete space to be well-defined (otherwise how could the one-dimensional phase trajectory fill the $6N - 1$ dimensional hypersurface?), the volume entropy appearing on the left hand side of Eq. (29) can be safely defined on the continuum (which is indeed the usual case). This suggests (a) that we consider the volume entropy as a good continuum counterpart of the Helmholtz entropy of a discrete ergodic system, and (b) that a general formulation of the Helmholtz theorem may be attempted which does not rely on a discrete structure. Proceeding in this direction would require in particular the employment of a notion of ergodicity based on the continuum instead of the Boltzmann-Gallavotti discrete one.

The notion of *metric indecomposability* seems to be very well suited for this purpose. Let us recall that in the modern ergodic theory [1],

DEFINITION 3.2. A portion Π of the phase space is called metrically indecomposable, if:

- Π is invariant under the Liouville evolution induced by H , and
- Π cannot be represented in the form $\Pi = \Pi_1 + \Pi_2$, where Π_1 and Π_2 are invariant disjoint subsets of positive measure.

Two remarks are needed at this point: (1) metric indecomposability extends the notion of one-dimensional monocyclicity to any dimension, and (2) metric indecomposability extends the Boltzmann-Gallavotti discrete notion of ergodicity to the continuum. The first remark follows from the fact that in the case of one-dimensional monocyclic systems, any hyper-surface of constant energy in the phase space (they are indeed curves) is metrically indecomposable. This is a consequence of the fact that each trajectory draws *one* complete constant energy curve. For example, for a non-monocyclic system like a particle inside a double well potential, all the curves with energy lower than a certain value are in fact decomposable into two sub-curves drawn by distinct trajectories with same energy (see Chap. 5 and Figure 5.2)). The second remark follows as one notes that the Boltzmann-Gallavotti ergodicity expresses the impossibility of decomposing the time-evolution permutation into sub-permutations as much as metric indecomposability expresses the impossibility of decomposing the Liouville flow into disjoint sub-flows: they express the same concept in a discrete and a continuum space, respectively.

Two analogous remarks apply to the volume entropy: (1') the volume entropy is the generalization of the Helmholtz entropy (Eq. (19)) to any dimension, (2') the volume entropy is the generalization of entropy for discrete ergodic systems in (Eq. (24) to continuum systems. Remark (1') follows by direct check that the volume entropy in one-dimension is equal to the Helmholtz entropy, viz. by integration over dp one gets:

$$(30) \quad \int \frac{dx dp}{h} \theta \left(E - \frac{p^2}{2m} - \varphi(x, V) \right) = 2 \int_{x_-(E,V)}^{x_+(E,V)} \frac{dx}{h} \sqrt{2m(E - \varphi(x, V))}.$$

Remark (2') has been already discussed (see point (a) above).

It is not a surprise, then, that volume entropy and metric indecomposability can be employed to formulate a generalized Helmholtz theorem (GHT), which on one hand extends the Helmholtz theorem to any dimension, and on the other extends the Boltzmann-Gallavotti ideas to the continuum:

Theorem 2 (Helmholtz, generalized). Let $H(\mathbf{p}, \mathbf{q}; V)$ be the Hamiltonian of a mechanical system with $3N$ degrees of freedom. Let any hyper-surfaces of constant energy in the $6N$ -dim phase space Γ be metrically indecomposable. Let a *state* be characterized by the set of quantities:

(31)

$$E = \text{total energy} = K + \varphi$$

$$T = \text{twice the time average of the kinetic energy per degree of freedom} = \frac{2\langle K \rangle_t}{3N}$$

$$V = \text{the external field}$$

$$P = \text{time average of } -\frac{\partial \varphi}{\partial V} = -\left\langle \frac{\partial \varphi}{\partial V} \right\rangle_t.$$

Then, the differential

$$(32) \quad \frac{dE + PdV}{T}$$

is exact, and the volume entropy,

$$(33) \quad S_{\Phi}(E, V) = \ln \Phi(E, V),$$

is the generating function, i.e.:

$$(34) \quad dS_{\Phi} = \frac{dE + PdV}{T}.$$

Proof. The differential of $S_{\Phi}(E, V)$ is:

$$(35) \quad dS_{\Phi} = \frac{\partial S_{\Phi}}{\partial E} dE + \frac{\partial S_{\Phi}}{\partial V} dV.$$

Using the definition of Eq.(8):

$$\frac{\partial S_{\Phi}}{\partial E} = \frac{1}{\Phi} \left(\frac{\partial}{\partial E} \int \frac{d\mathbf{z}}{h^{3N}} \theta [E - H(\mathbf{z}, V)] \right) =$$

$$\begin{aligned}
&= \frac{1}{\Phi} \left(\int \frac{d\mathbf{z}}{h^{3N}} \delta [E - H(\mathbf{z}, V)] \right) = \\
(36) \quad &= \frac{\Omega}{\Phi} \\
\frac{\partial S_{\Phi}}{\partial V} &= \frac{1}{\Phi} \left(\frac{\partial}{\partial V} \int \frac{d\mathbf{z}}{h^{3N}} \theta [E - H(\mathbf{z}, V)] \right) = \\
&= -\frac{1}{\Phi} \left(\int \frac{d\mathbf{z}}{h^{3N}} \delta [E - H(\mathbf{z}, V)] \frac{\partial}{\partial V} H(\mathbf{z}, V) \right) = \\
(37) \quad &= -\frac{\Omega}{\Phi} \left\langle \frac{\partial H}{\partial V} \right\rangle_{\mu}
\end{aligned}$$

where use is made of Eqs. (6) and (13) and of the relation $\frac{d}{dx}\theta(x) = \delta(x)$. For Birkhoff's ergodic theorem [1], metric indecomposability of the hyper-surface of constant energy implies that the time average of any summable phase function f over (almost) any trajectory belonging to the hyper-surface does not depend on the trajectory itself and is equal to its phase average, i.e.¹¹,

$$(38) \quad \langle f \rangle_t = \langle f \rangle_{\mu}$$

where now the time average is calculated over the *Poincaré recurrence time* τ . Thanks to Eq.(38), we can interchangeably use the subscript μ or t for the averages. This implies:

$$(39) \quad -\left\langle \frac{\partial H}{\partial V} \right\rangle_{\mu} = -\left\langle \frac{\partial H}{\partial V} \right\rangle_t = P.$$

Thanks to the equipartition theorem (Eq. (16)), we also get:

$$(40) \quad \frac{\Phi}{\Omega} = \frac{2\langle K \rangle_{\mu}}{3N} = \frac{2\langle K \rangle_t}{3N} = T.$$

Combining all these together the thesis follows straightforwardly:

$$(41) \quad dS_{\Phi} = \frac{1}{T}dE + \frac{P}{T}dV.$$

¹¹For clarity, systems with metrically indecomposable hyper-surfaces will be referred to, in the text, as Birkhoff ergodic systems.

3.4. Concluding Remarks

It is worth to notice that the generalized Helmholtz theorem can be easily extended to any number of external “coordinates” $\vec{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$. By defining the conjugate “forces” with:

$$(42) \quad P_j = - \left\langle \frac{\partial H(\mathbf{z}, \vec{\lambda})}{\partial \lambda_j} \right\rangle_t \quad j = 1 \dots r,$$

the theorem would read:

$$(43) \quad dS_\Phi = \frac{dE + \sum P_j d\lambda_j}{T}.$$

An important fact is that, adopting the Helmholtz-Boltzmann viewpoint, we can derive *equilibrium* thermodynamics from mechanics without incurring the “methodological paradox”, mentioned by Khinchin [1, p. 41], of neglecting inter-molecular interactions. Indeed, we are forced to include the interaction term in the Hamiltonian, because otherwise the system would fail to be ergodic and Boltzmann’s ideas (either in their original or modern form) would not apply. This is because ergodicity means a situation of complete *non-integrability*¹², which ultimately stems from the impossibility of reducing the system (through a canonical transformation) to a simpler one with no interactions. Surprisingly, the “disaster” of non-integrability turns out, indeed, to be the savior: in the very case when the possibility of a microscopic description is completely out of our reach, the thermodynamic one is made available.

Remarkably, the generalized Helmholtz theorem tells that the heat theorem holds mechanically for Birkhoff ergodic systems *of any size*, which, therefore, provide a good mechanical model of thermodynamics. Taking Gallavotti’s reconstruction as the starting point, the Helmholtz-Boltzmann theory has been re-expressed within the frame of the modern ergodic theory and has been found to agree with Berdichevsky’s [23] reconstruction of Hertz’ theory: both deal with Birkhoff ergodic systems¹³ and

¹²The presence of integrals of motion besides the energy, in fact, would decompose the surface of constant energy.

¹³Recall that the volume entropy is the *ergodic adiabatic invariant*.

through different paths (in one case the starting point is the requirement that the entropy generate the heat differential; in the other case the entropy is required to be an adiabatic invariant) reach the same mechanical expression of entropy, and the same thermodynamics. Adopting Hertz's approach, the heat theorem would follow as a corollary. Adopting that of Helmholtz-Boltzmann adiabatic invariance would follow. Both theories, remarkably, are based on microscopic dynamics and the ergodic hypothesis only: never are statistical notions (like probability distributions or the law of large numbers) invoked. In this sense the generalized Helmholtz theorem is a theorem of ergodic Hamiltonian mechanics that has a straightforward thermodynamic interpretation.

To summarize, Statements 1 and 2 of the second law hold *mechanically*, i.e., as a consequence of Hamilton's equations of motion, for ergodic systems, provided that we identify the thermodynamic entropy with the quantity $\ln \Phi$, i.e., the volume entropy in Eq. (8).

In the next Chapter we will see that Statement 3 can be explained in terms of volume entropy too. This time a bit of statistics though will be needed.

CHAPTER 4

VOLUME ENTROPY AND THE LAW OF ENTROPY INCREASE

In the previous chapter we have seen that the volume entropy (8) provides a good mechanical analogue of thermodynamic entropy in the sense that it accounts quite well for Statement 1 and Statement 2 of the second law of thermodynamics (see Chap. 1). These are the two statements that pertain to equilibrium, whereas Statement 3 pertains to out of equilibrium and expresses the law of entropy increase. Obviously the fact that the volume entropy accounts for Statement 1 and Statement 2 suggests that it might turn out to be very useful in addressing Statement 3 as well. It is important to stress that Statements 1 and 2 are accounted for by the volume entropy *mechanically*, i.e., no probabilistic assumptions were introduced and the equations $dS = \frac{dE+pdV}{T}$ (Statement 1) and $\Delta S = 0$ (Statement 2) hold as a consequence of *ergodic Hamiltonian mechanics* only. Simple considerations, however, suggest that Statement 3 could be proved only in some statistical or averaged sense. Consider, for example, a particle of mass m in a one-dimensional box of length L . The particle bounces back and forth inside the box. Let E be the energy of the particle. Imagine that we can change the length of the box by moving the right wall of the box. The volume entropy of this elementary ergodic system is simply $S_{\Phi}(E, L) = \ln(2L\sqrt{2mE}/h)$. Now imagine that we perform a very fast compression of the box, much faster than the particle period of motion $T = \sqrt{\frac{2m}{E}}L$. Let $L - \Delta L$ be the final length of the box. Imagine that during this transformation the particle is far from the moving wall and does not bounce against it. Its energy would not change but the change of its volume entropy would be $\Delta S_{\Phi} = \ln(1 - \Delta L/L)$. Thus ΔS_{Φ} would be negative. This simple example should convince us that any purely mechanical attempt to derive the entropy increase principle (i.e., Statement 3 of the

second law) on the basis of volume entropy would be in vain. We certainly need to add some statistical ingredient if we want to prove it.

Thus, we are going to assume that the initial energy of our insulated system is not known. All we know is that it is within some range $E, E + dE$ with some probability $p(E)\Omega(E)dE$. The symbol $\Omega(E)$ denotes the “density of states” at energy E (Eq. 9). For example, if we first place the system in thermal contact with a heat bath at temperature T , and then we remove the contact, we will not know for certain what the energy of the system will be, but we will know that $p(E) = \frac{e^{-E/T}}{Z(T)}$ where $Z(T)$ is the canonical partition function (i.e., the normalization).

In this chapter we will prove that, provided that $p(E)$ is a decreasing function of E , the *expectation value* of the volume entropy will be larger than its initial value. It turns out that such a proof is much easier in quantum mechanics than in classical mechanics. Therefore we shall first quantize the volume entropy and then study its behavior under the action of a varying parameter, that is, a time-dependent perturbation.

4.1. Quantum Volume Entropy

Let us first consider the one-dimensional (1D) case. In 1D the volume entropy is:

$$(44) \quad S_{\Phi} = \ln \int_{H \leq E} \frac{dx dp}{h}.$$

This expression can be conveniently reexpressed as the logarithm of the *reduced action* [24]

$$(45) \quad S_{\Phi} = \ln \oint \frac{p dx}{h}.$$

which is the Helmholtz entropy (19). Quantization of the Helmholtz entropy is almost immediate. Indeed, using a colorful expression, I would say that Eq. (45) invites the reader to quantize. Using the semiclassical approximation of Bohr-Sommerfeld [25] and setting h equal to Planck’s constant allows us to see that S_{Φ} is a quantized

quantity whose possible values are (within the range of validity of the approximation):

$$(46) \quad S_{\Phi} = \ln \left(n + \frac{1}{2} \right).$$

We can extend this line of reasoning to multidimensional systems whose dynamics is ergodic. In this general case the volume entropy is given by Eq. (8). Again using the quasi-classical viewpoint [26] the integral in Eq. (8) approximately counts the number of quantum states not above a certain energy $\varepsilon_n = E$. Since the levels are nondegenerate this number is $n + \frac{1}{2}$, where one considers that the vacuum state counts as a half state. The levels are nondegenerate because the corresponding classical dynamics is ergodic. This can be understood by noticing that ergodicity implies that the Hamiltonian is the only integral of motion. This, translated into the language of quantum mechanics, says that the Hamiltonian alone constitutes a complete set of commuting observables, so that the only quantum number is n .

At this point, it is quite easy to construct the quantum version of volume entropy. Consider a finite (i.e., not necessarily macroscopic) *nondegenerate* quantum system. Let \hat{N} be the *quantum number operator*, i.e.,:

$$(47) \quad \hat{N} \doteq \sum_{k=0}^K k |k\rangle \langle k|$$

where $\{|k\rangle\}$ is the complete orthonormal set of the Hamiltonian's eigenstates. K , the total number of energy levels, can be infinite. The eigenvectors of \hat{N} are the energy eigenvectors, and the eigenvalues are the corresponding quantum numbers. Then the *quantum volume entropy operator* can be defined as:

$$(48) \quad \hat{S} \doteq \ln \left(\hat{N} + \frac{1}{2} \right)$$

We adopt a system of units where k_B , the Boltzmann constant, is equal to 1.

4.2. Proof of the Entropy Principle

Armed with a quantum mechanical analogue of thermodynamic entropy (48), we can now study its evolution under a time-dependent perturbation. As prescribed by

the entropy principle (see Chap. 1) we shall assume that the system is thermally isolated from the environment. As discussed previously, the system energy is not known. This means that the system is assumed to be in a statistical mixture of states, described by a density matrix $\hat{\rho}_i$, rather than a pure state $|k\rangle$. As prescribed by Statement 3 we shall also assume that the system is initially at equilibrium. We will translate this thermodynamic notion into the quantal requirement that $\frac{\partial \hat{\rho}_i}{\partial t} = 0$. So the system is at equilibrium whenever $\frac{\partial \hat{\rho}}{\partial t} = 0$ and it is out of equilibrium whenever $\frac{\partial \hat{\rho}}{\partial t} \neq 0$. At $t = t_i$, we switch on a perturbation. This is implemented by changing the value of some external parameter λ during the course of time: $\lambda = \lambda(t)$. λ can be for example the volume V of a vessel containing the system, or the value of some external field like an electric or a magnetic field. At time $t = t_{off}$, the perturbation is switched off. We assume that at some time $t_f \geq t_{off}$ any transient effect has vanished and the system has attained a new equilibrium state described by some $\hat{\rho}_f$, such that $\frac{\partial \hat{\rho}_f}{\partial t} = 0$. Thus, before time t_i and after t_f the system is at equilibrium, and for $t_i < t < t_f$ it is out of equilibrium. Due to the perturbation the Hamiltonian changes from the initial value \hat{H}_i to the final value \hat{H}_f , and accordingly the quantum entropy operator will change in time and move from \hat{S}_i to \hat{S}_f . We introduce the following time-dependent orthonormal basis set $\{|k, t\rangle\}$. The vectors $|k, t\rangle$ are defined as the eigenvectors of the “frozen” Hamiltonian $\hat{H}(t)$. That is:

$$(49) \quad \hat{H}(t) = \sum_{k=0}^K \varepsilon_k(t) |k, t\rangle \langle k, t|.$$

Since at time t_i $\frac{\partial \hat{\rho}_i}{\partial t} = 0$, then $[\hat{\rho}_i, \hat{H}] = 0$. This means that $\hat{\rho}_i$ is diagonal over the initial basis $\{|k, t_i\rangle\}$:

$$(50) \quad \hat{\rho}(t_i) = \sum_{k=0}^K p_k |k, t_i\rangle \langle k, t_i|$$

As anticipated in the introduction we shall assume that the initial probability distribution p_i is decreasing:

$$(51) \quad p_0 \geq p_1 \geq \dots \geq p_i \geq \dots$$

Our definition of quantum entropy, (48), is essentially an *equilibrium* definition. Using the bases $\{|k, t\rangle\}$, we can extend the definition to the *out of equilibrium* case, as follows:

$$(52) \quad \hat{S}(t) \doteq \sum_{k=0}^K \ln \left(k + \frac{1}{2} \right) |k, t\rangle \langle k, t|.$$

We shall assume that nondegeneracy is kept at all times. This implies that there is *no level crossing*, and ensures that the quantum number operator gives the correct eigenvalues at all times. The same assumption is used in Ref. [27] to ensure the proper ordering of energy eigenvalues. Note that, unlike the Hamiltonian's spectrum (49), the spectrum of the quantum entropy (52) is time-independent. We define the transition probabilities:

$$(53) \quad |a_{kn}(t_f)|^2 = |\langle n, t_f | \hat{U}(t_i, t_f) | k, t_i \rangle|^2,$$

where

$$(54) \quad \hat{U}(t_i, t) = \mathcal{T} \exp \left(-\frac{i}{\hbar} \int_{t_i}^t \hat{H}(s) ds \right)$$

is the time evolution operator expressed in terms of the time-ordered exponential $\mathcal{T} \exp$. The $|a_{kn}(t_f)|^2$'s represent the probabilities that the system will be found in the state $|n, t_f\rangle$ at time t_f provided that it was in the state $|k, t_i\rangle$ at time t_i . They satisfy the relations [27]:

$$(55) \quad \sum_{k=0}^K |a_{kn}(t_f)|^2 = \sum_{n=0}^K |a_{kn}(t_f)|^2 = 1$$

and

$$(56) \quad |a_{kn}(t_f)|^2 \geq 0.$$

For the change in the expectation value of the quantum entropy \hat{S} of Eq. (48) we have:

$$(57) \quad S_f - S_i = Tr \left[\hat{\rho}_f \hat{S}_f \right] - Tr \left[\hat{\rho}_i \hat{S}_i \right] = \sum_{n=0}^K (p'_n - p_n) \ln \left(n + \frac{1}{2} \right)$$

where

$$(58) \quad p'_n = \sum_{k=0}^K p_k |a_{kn}(t_f)|^2$$

is the probability that the system is in state $|n, t_f\rangle$ provided that the initial probabilities were p_n .

Using the “summation by parts” rule [27]:

$$(59) \quad \sum_{n=0}^K a_n b_n = a_K \sum_{n=0}^K b_n - \sum_{m=0}^{K-1} (a_{m+1} - a_m) \sum_{n=0}^m b_n,$$

Eq. (57) becomes

$$(60) \quad S_f - S_i = \sum_{m=0}^{K-1} \ln \left(\frac{m + \frac{3}{2}}{m + \frac{1}{2}} \right) \sum_{n=0}^m (p_n - p'_n)$$

We have:

$$(61) \quad \begin{aligned} \sum_{n=0}^m (p_n - p'_n) &= \sum_{n=0}^m p_n - \sum_{n=0}^m \sum_{i=0}^K p_i |a_{in}(t_f)|^2 \\ &= \sum_{n=0}^m p_n \left(1 - \sum_{i=0}^m |a_{in}(t_f)|^2 \right) - \sum_{n=0}^m \sum_{i=m+1}^K p_i |a_{in}(t_f)|^2 \end{aligned}$$

From Eqs. (55) and (56) we have $(1 - \sum_{i=0}^m |a_{in}(t_f)|^2) \geq 0$ and $|a_{in}(t_f)|^2 \geq 0$, therefore using the ordering of probabilities (51) we get (see also [27]):

$$(62) \quad \sum_{n=0}^m (p_n - p'_n) \geq p_m \sum_{n=0}^m \left(1 - \sum_{i=0}^m |a_{in}(t_f)|^2 \right) - p_m \sum_{n=0}^m \sum_{i=m+1}^K |a_{in}(t_f)|^2$$

$$(63) \quad = mp_m - p_m \sum_{n=0}^m \sum_{i=0}^K |a_{in}(t_f)|^2 = 0$$

where we used Eq. (55) in the last line. Noting that $\ln \left(\frac{m + \frac{3}{2}}{m + \frac{1}{2}} \right) > 0$ in Eq. (60), we finally reach the conclusion that:

$$(64) \quad S_f \geq S_i.$$

This inequality holds for any transformation acting on a thermally insulated, nondegenerate quantum system which is initially at equilibrium with a decreasing ordering

of probabilities. It is not hard to see that the equal sign holds for adiabatic transformation (Statement 2). For this case the nondegeneracy assumption ensures that the *quantum adiabatic theorem* holds [28]. This ensures that the transition probability between states with different quantum number will be null during an adiabatic transformation:

$$(65) \quad |a_{in}(t_f)|^2 = \delta_{in}.$$

Therefore, for an adiabatic transformation we get $p'_i = p_i$ (see Eq. 58), so we obtain equality in Eq. 64

$$(66) \quad S_f = S_i.$$

This concludes our quantum mechanical proof of the entropy principle, that is Statements 2 and 3 of the second law. Note that the result in Eq. (66) is not surprising because the quantum entropy operator has been defined as the quantum counterpart of a classical adiabatic invariant (i.e., the volume integral Φ of Eq. (9))

We have established and used the following correspondences between thermodynamics and quantum mechanics:

- entropy $\rightleftharpoons \ln \left(\hat{N} + \frac{1}{2} \right)$
- equilibrium $\rightleftharpoons \frac{\partial \hat{\rho}}{\partial t} = 0$
- (non)quasi-static process \rightleftharpoons (non)adiabatic perturbation

Since the equilibrium condition for the final state has not been used in the proof, inequality (64) holds for any $t \geq t_i$, where at time t the system is possibly out of equilibrium. Note that this by no means implies that

$$(67) \quad S(t) \doteq Tr[\hat{S}(t)\hat{\rho}(t)]$$

is a monotonically increasing function of time. All we can say is that if at times $t_1 < t_2 < \dots < t_n < \dots$ the density matrix is diagonal and its spectrum is monotonically

decreasing, then:

$$(68) \quad S(t_1) \leq S(t_2) \leq \dots \leq S(t_n) \leq \dots$$

If at time t_A the system is out of equilibrium, or it is at equilibrium but not with decreasing ordering of probabilities, there can well be a time $t_B > t_A$ such that $S(t_A) > S(t_B)$. It is important to stress that, when the system is out of equilibrium the quantity $S(t)$ should not be regarded as the system's thermodynamic entropy, which is essentially an equilibrium property. Thus $S(t)$ is only one of the many possible out-of-equilibrium generalizations of entropy. What makes it special is that it proves effective in addressing the entropy principle.

4.3. Thomson's Formulation of the Second Law and the Minimal Work Principle

The present proof of the entropy principle is very close, in the approach and methods, to a result discussed recently in Ref. [29]. The authors there considered the following alternative formulation of the second law, which is attributed to Kelvin (W. Thomson):

THOMSON'S FORMULATION: No work can be extracted from a closed equilibrium system during a cyclic variation of a parameter by an external source.

If we denote the work done by the external source as W , the principle can be expressed simply as:

$$(69) \quad W \geq 0$$

The proof of [29] goes like the one we have proposed above for the entropy principle. In this case one wants to study the following quantity:

$$(70) \quad W \doteq Tr[\hat{H}_f \hat{\rho}_f] - Tr[\hat{H}_i \hat{\rho}_i]$$

for a cyclic process. This means that the final Hamiltonian is assumed to be equal to the initial one $\hat{H}_f = \hat{H}_i \doteq \hat{H}_0$. Thus:

$$(71) \quad W \doteq \sum_{n=0}^K \varepsilon_n (p'_n - p_n),$$

where ε_n are the eigenvalues of \hat{H}_0 . These are ordered according to $\varepsilon_1 < \varepsilon_2 < \dots < \varepsilon_i < \dots$. The eigenvalues ε_n play the same role here as the entropy eigenvalues $\ln(n + 1/2)$, in Eq. (57). Thus it is immediate to see that, under the assumption of decreasing probabilities (51), Eq. (69) holds quantum-mechanically.

This result has been extended to the case of a possibly noncyclic transformation in Ref. [27]. The authors have found that

$$(72) \quad W - \widetilde{W} = \sum_{n=0}^K \varepsilon'_n (p'_n - p_n) \geq 0,$$

where ε'_n are the eigenvalues of the final Hamiltonian, W is the work actually performed on the system and \widetilde{W} is the work that would have been performed if the same transformation had been carried adiabatically. The proof is formally equivalent to the one discussed here. Equation (72) expresses the minimal work principle: whenever we perform a nonadiabatic transformation, we spend more work than we would have if performing an adiabatic one.

The formal similarity of Eq. (72) and Eq. (57) proves that the formulation of the second law as a principle of minimal work or as a principle of entropy increase are *equivalent*. In particular it is easily seen that:

$$(73) \quad \text{sign}(W - \widetilde{W}) = \text{sign}(S_f - S_i),$$

so the two principles are equivalent. Further, whenever one is violated, the other will be too. Cases where the minimal work principle is violated because of level crossing are discussed by [27]. In those case The entropy principle would also be violated.

The history behind this kind of quantum mechanical proof of the second law is relatively recent, and can traced back at last to the works of Ref. [30] and [31].

Due to the lack of a suitable quantum mechanical analogue of entropy, though, the application of such quantal approaches has remained restricted to the analysis of statements that concern work, whose mechanical definition is quite straightforward. To the best of my knowledge, similar arguments and approaches have been previously proposed for addressing the entropy principle only in the relatively un-known work of Tasaki [32].

4.4. Comparison with other Quantum Entropies

The employment of the quantum volume entropy improves quite a lot over previous attempts at explaining the entropy principle based on quantum entropies. In fact, the employment of the entropy in Eq. (67) has many advantages over other quantum mechanical entropies present in literature. In contrast with von Neumann entropy:

$$(74) \quad S_{vN} = -Tr[\hat{\rho}(t) \ln \hat{\rho}(t)]$$

which is constant in time because of unitarity of the quantum mechanical time-evolution operator, the expectation value of the quantum operator \hat{S} in Eq. (67) does change in time, and it has been proved to increase under the assumption discussed. Tolman's coarse-grained entropy [10]:

$$(75) \quad S_{cg}(t) = - \sum_{\nu} P_{\nu}(t) \ln P_{\nu}(t)$$

does change in time and it is an adiabatic invariant [10]. Nonetheless it fails in accounting for the inequality in the case of nonadiabatic perturbations. All we know is that for an infinitesimal *abrupt* transformation that begins and ends in a *canonical equilibrium* state ¹ the corresponding change in S_{cg} is nonnegative [10]. But this does not ensure that for any *finite* nonadiabatic transformation the change would be nonnegative as required by Clausius formulation. Tolman's argument that any finite transformation could be reproduced by a sequence of many infinitesimal abrupt transformations each followed by attainment of a canonical equilibrium does not seem

¹That is $\hat{\rho} = Z^{-1}e^{-\beta\hat{H}}$

to be tenable. In fact, as a result of a finite nonadiabatic transformation, the system could well end up in a noncanonical distribution [27]. Further, Tolman’s definition of entropy of Eq. (75) applies only to macroscopic equilibrium systems. On the contrary the result proved in Sec. 4.2 holds *no matter the number of degrees of freedom* of the system. Thus the quantum volume entropy might turn out to be very useful in the novel and fast growing field of *quantum thermodynamics* of nanoscale systems. See for example Ref. [33] and Ref. [34].

4.5. Classical case

The result of Eq. (64) can be proved also classically. Let the system be initially distributed according to some probability distribution function $p_0(E)$. Let $p_1(E)$ be the final distribution. Let $\Phi_0(E)$ and $\Phi_1(E)$ denote the volumes enclosed by the hyper-surfaces $H_0(\mathbf{q}, \mathbf{p}) = E$ and $H_1(\mathbf{q}, \mathbf{p}) = E$ respectively, where H_0 and H_1 are the initial and final Hamiltonians. The volume entropy of a representative point that at time t_i lies on the hypersurface $H_0(\mathbf{q}, \mathbf{p}) = E$ is $\ln \Phi_0(E)$. We have a similar expression for time t_f . Then:

$$(76) \quad S_f - S_i = \int_0^\infty dE \Omega_1(E) p_1(E) \ln \Phi_1(E) - \int_0^\infty dE \Omega_0(E) p_0(E) \ln \Phi_0(E)$$

where Ω_r denotes the initial ($r = 0$) or final ($r = 1$), density of states. Note that according to Eq. (10), we have:

$$(77) \quad dE \Omega_r(E) = d\Phi_r(E).$$

Thus we can make the change of variable $E \leftrightarrow \Phi_r$ in the integrals. Let² $P_r(\Phi_r) \doteq p_r(E(\Phi_r))$, then we have (after dropping the subscript in Φ_r):

$$(78) \quad S_f - S_i = \int_0^\infty d\Phi (P_1(\Phi) - P_0(\Phi)) \ln \Phi$$

This is the classical analogue of Eq. (57). The role of $n + \frac{1}{2}$ is played by the “enclosed volume” Φ , and the discrete probabilities p_n, p'_n are now probability density functions

²Note that $\Phi(E)$ is a monotonically increasing function of E by definition, so it is always possible to define the inverse function $E(\Phi)$.

$P_0(\Phi)$ and $P_1(\Phi)$. Since the evolution is deterministic, it is possible to express the final probability in terms of the initial one as

$$(79) \quad P_1(\Phi) = \int_0^\infty d\Theta A(\Phi, \Theta) P_0(\Theta)$$

where $A(\Phi, \Theta)$ is the Green function associated to the evolution of probabilities in Φ space. That is, $A(\Phi, \Theta)$ represents the evolution at time t_f of a Dirac delta centered around Θ at time t_i . If we denote the time evolution operator that evolves probabilities in Φ space from time t_i to time t_f as $\mathcal{U}(t_i, t_f)$, A is defined as:

$$(80) \quad A(\Phi, \Theta) = \mathcal{U}(t_i, t_f) \delta(\Phi - \Theta)$$

The function $A(\Phi, \Theta)$ is the classical counterpart of the transition probability $|a_{kn}|^2$. Evidently, thanks to the classical adiabatic theorem, we have for an adiabatic switching:

$$(81) \quad A(\Phi, \Theta) = \delta(\Phi - \Theta)$$

and $S_f = S_i$ as is known from Hertz's theory (see Chap. 2). The equation above is the classical counterpart of Eq. (65). For nonadiabatic switching we expect $A(\Phi, \Theta)$, considered as a function of Φ , to be bell-shaped with some finite width. The problem of determining the shape of A has been studied by Jarzinsky [20], who proved that, within second order of adiabatic perturbation theory, A actually drifts and diffuses according to an effective Fokker-Planck equation. Since $A(\Phi, \Theta)$ represents a probability distribution function in Φ space, it satisfies:

$$(82) \quad A(\Phi, \Theta) \geq 0$$

and

$$(83) \quad \int_0^\infty d\Phi A(\Phi, \Theta) = 1.$$

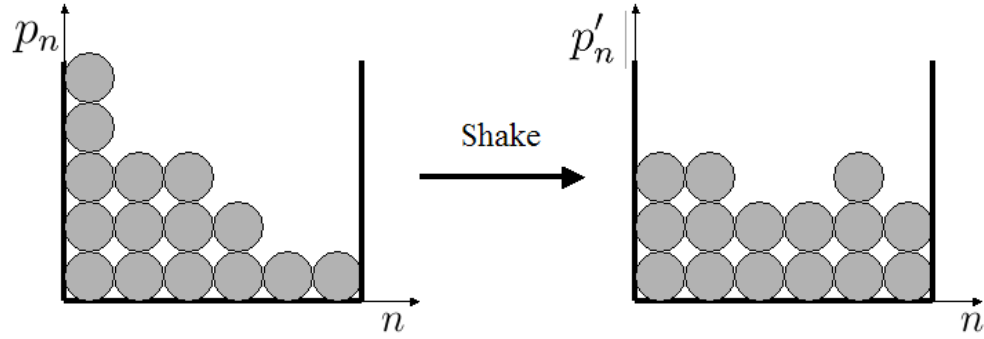
Using Liouville's Theorem it is also possible to prove that:

$$(84) \quad \int_0^\infty d\Theta A(\Phi, \Theta) = 1;$$

see Appendix A. The analogy with the quantum case has been completely established now, and the proof of the entropy principle follows by repeating the same steps. The requirement on the initial distribution $P_0(\Phi)$ is that it be a decreasing function of Φ . Since $\Phi_0(E)$ is increasing, this requirement translates into the requirement that $p_0(E)$ be a decreasing function of E .

4.6. The role of the Initial Equilibrium

Inequality (64) for entropy holds as a direct consequence of the time reversal symmetric microscopic laws of quantum or classical mechanics. As such, it does not entail any arrow of time. The reason for the emergence of the \geq sign in Eq. (64) should be looked for, rather, in the fact that we have considered only a certain restricted subset of all possible initial conditions. To explain this point it might be useful to see our ensemble of systems as a box containing many balls (see Fig. 4.1). Each ball represents an element of the ensemble. The box is divided into labelled cells that represent the quantum states. The cell closest to the left wall is the state with $n = 0$, its right neighbor cell is the state $n = 1$ and so forth. At time t_i , the balls are distributed in the box according to some probability p_n . We can consider the time-dependent perturbation acting on the system as the action of shaking the box. The effect of the shaking is that of flattening out the initial distribution. Thus if initially we had some accumulation of balls towards the left side of the box, we expect the final state to be flatter. If we look at the average value of n or any other increasing function of n , like for example $\ln(n + 1/2)$, we would record an increase of such values. This is a mere consequence of the fact that initially we had an accumulation towards the left. If initially we have had an accumulation towards the right, again the shaking would flatten out the distribution, but this time we would see a decrease of the average value of n and of $\ln(n + 1/2)$. If instead the initial distribution were flat, we would see no change in those quantities. Indeed it is quite easy to see that the sign of inequality (64) would be reversed if an increasing ordering of initial probability were assumed. Therefore, for such subset of the set of all possible initial distributions, we



$$S_i = \langle \ln(n + 1/2) \rangle_i \leq \langle \ln(n + 1/2) \rangle_f = S_f$$

FIGURE 4.1. Visual representation of the effect of a non adiabatic perturbation on a quantum system which is initially described by decreasing probabilities. After shaking, the initial accumulation towards the left would flatten out, and the average value of $\ln(n + 1/2)$ (i.e. the entropy) would increase.

would actually have a law of entropy decrease! This reflects the fact that there is no asymmetry in the time evolution of the volume entropy operator. Thus, in principle, it should be possible to observe a decrease of entropy if the initial equilibrium would be given by an *inverted population*. In other words, we should be able to observe an inverted second law of thermodynamics in *negative absolute temperature* systems. Indeed experimental evidence of this exists since the very pioneering works of Pound, Purcell and Ramsey on spin systems [35, 36, 37, 38]. They observed that

“when a negative temperature spin system was subjected to resonance radiation, more radiant energy was given off by the spin system than was absorbed [38].”

This means that it is possible to extract work from a negative temperature system by means of a cyclic transformation. In other words, for negative temperature systems

we already have experimental evidence that, in contrast to Eq. (69):

$$(85) \quad W \leq 0$$

Because of the *equivalence* of the minimal work principle and the entropy principle, in this case, in contrast to Eq. (64) we would also have:

$$(86) \quad S_f \leq S_i$$

The fact that the law of entropy increase is overwhelmingly more often observed than its mirror-image law of entropy decrease is a consequence of the fact that positive temperatures are overwhelmingly more common than negative ones. The former in fact is the *natural* state of matter, whereas the second can only be created *artificially* and only in few very special cases. Ramsey [38] already pointed out that very strict conditions must be met for a system to be capable of negative temperatures: (a) the system must be at equilibrium, (b) there must be an upper limit in the Hamiltonian's spectrum, and (c) the system must be thermally isolated from the environment. The second requirement is very restrictive as most systems have an unbounded kinetic energy term in the Hamiltonian³. Also the requirement (c) is restrictive in the sense that thermal insulation can be achieved only approximately and for a certain amount of time. On the contrary, the inevitable thermal contact of our system with its environment would eventually bring it to the monotonically decreasing Gibbs state

$$(87) \quad p_i = Z^{-1} e^{-\beta \epsilon_i}.$$

The latter describes the natural state of matter, and as such is the *inevitable* initial condition of any thermodynamic experiment⁴. Thus the time asymmetry of the laws

³See [39] for a recent and interesting example of negative kinetic temperature, though.

⁴If we consider that in experiments on negative temperature systems one has first to create an inverted population from a natural one, we will see that indeed the total entropy change would be positive. The entropy spent to create the inverted population is larger than that gained back when applying the resonant radiation.

of thermodynamics arises at the level of the initial thermal equilibrium, rather than in the second law itself. This seems to be in agreement with the view expressed by Brown and Uffink [16], according to which the second law does not entail any time asymmetry. The origin of the arrow of time should be found, instead, in the minus first law of thermodynamics, namely the *equilibrium principle* cited in Chap. 2.

CHAPTER 5

APPLICATIONS OF VOLUME ENTROPY: MICROCANONICAL PHASE TRANSITIONS IN SMALL SYSTEMS

5.1. Overview of Microcanonical Phase Transitions

The work of D.H.E. Gross [40] has recently focused attention on the fact that a microcanonical description of systems which may display phase transitions is in general more adequate than the traditional canonical one. This is because the canonical description may “smear out” important information contained in the richer microcanonical description [40]. For example negative specific heats, which have recently been observed experimentally in mesoscopic systems [41, 42, 43], can be accounted for in the microcanonical ensemble but not in the canonical one [44]. Indeed it is well known that the canonical ensemble and microcanonical ensemble are not in general equivalent, even when the thermodynamic limit is considered [2, 40, 45, 46].

The statistical mechanical analysis of physical systems based on the canonical ensemble is quite well established with universal consensus. Roughly speaking one has to compute the partition function $Z(\beta)$ and derive the thermodynamics of the system from the free energy $F = -\beta^{-1} \ln Z(\beta)$. For the microcanonical ensemble there is not such general consensus. Since the pioneering works of Boltzmann and Gibbs, as we have seen in Chap. 2, two possibilities have been given for the microcanonical analysis of physical systems, which correspond to surface and volume entropy in Eq. (6) and (8), respectively.

In the literature about microcanonical phase transitions the surface entropy is certainly the most popular. For example, the method of Barré *et. al.* [45], based on large deviation techniques, uses the surface entropy. The surface entropy is also used in Rugh’s microcanonical formalism adopted in Ref. [47]. The strongest advocate of

surface entropy is perhaps Gross [48]. Nonetheless, pioneers of microcanonical phase transitions, such as Thirring [49] and Lynden-Bell [44], used the volume entropy.

5.1.1. *Why use Volume Entropy*

As discussed previously (see Sec. 3.1) the two entropies are expected to coincide in the thermodynamic limit (see Eq. (11)). However when the number of degrees of freedom of the system under study is small relevant differences may appear. Therefore, it is necessary to choose the entropy properly. Some authors [50, 51] have already pointed out that the surface entropy is not adequate when dealing with small systems because it does not account properly for finite-size effects. On the other hand, since the volume entropy accounts for all three statements of the second law as a consequence of the properties of the Hamiltonian time evolution, and *no matter the number of degrees*, then the volume entropy must naturally be considered a *better mechanical analogue* of thermodynamic entropy than surface entropy.

The surface entropy does not generally reproduce statements 1 and 2. No general proof exists of Statement 3 from surface entropy that would not involve taking the thermodynamic limit. Concerning Statement 2, we have already discussed in Sec. 3.1.2 the fact that the surface entropy is not an adiabatic invariant. That Statement 1 is only satisfied approximately by surface entropy has been proved by Gallavotti [2]. He finds:

$$(88) \quad dS_{\Omega} = \frac{dE + PdV}{T} + O\left(\frac{1}{N}\right)$$

This result is not a surprise if we consider that in general $S_{\Phi} \neq S_{\Omega}$, $dS_{\Phi} = \frac{dE+PdV}{T}$ and the fact (Eq. (11)) that the surface entropy approaches the volume entropy for N tending to infinity.

Recently more authors are becoming aware of the theoretical value of volume entropy. For example, on the basis of a Laplace transform technique for the microcanonical ensemble, Pearson *et. al.* [52] reached the conclusion that the volume entropy “is the most correct definition for the entropy, even though it is unimportant

for any explicit numerical calculation”, meaning that in the thermodynamic limit the difference with surface entropy becomes negligible. On the other hand, for a small system, such intrinsic correctness of the volume entropy becomes very important. Adib [51] argues that the finite size corrections to surface entropy found in Ref. [53] would be unnecessary if the volume entropy were used instead.

It is worth mentioning that the volume entropy has another property that is particularly important for small systems which have negative heat capacity, namely it is a naturally *nonextensive* entropy. According to Lynden-Bell [44], systems with negative heat capacity are necessarily *nonextensive*. The property of nonextensivity of volume entropy follows directly from the composition rule of enclosed volumes Φ_i , $i = 1, 2$, of two systems with total energy $E = E_1 + E_2$ [1]:

$$(89) \quad \Phi(E) = \int dE_1 \Omega_1(E_1) \Phi_2(E - E_1)$$

which is not a simple multiplication but a form of convolution which accounts for all possible partitions of energies between the two systems. On the other hand, despite what is often stated in literature, the surface entropy is also nonextensive. The composition rule for surface integrals Ω is convolution, not multiplication [1]:

$$(90) \quad \Omega(E) = \int dE_1 \Omega_1(E_1) \Omega_2(E - E_1).$$

In summing, the volume entropy accounts for certain basic principles of thermodynamics, like the heat theorem and Clausius formulation of the second law equally well for large and small systems, whereas the surface entropy accounts for them only in the case of large systems. For this reason the volume entropy is the most appropriate *mechanical analogue of thermodynamic entropy* for both large and small systems.

5.2. Lennard-Jones Chain

According to the Helmholtz theorem (see Sec. 3.2) the mechanical analogue of thermodynamic entropy of a one dimensional system is the Helmholtz entropy in Eq.

19. However the theorem holds if there is only one trajectory per energy level (ergodicity). In this section we want to study the properties of the Helmholtz entropy when the ergodic conditions breaks down. In particular we are going to study the thermodynamic behavior of a 1D system that has two possible disconnected trajectories for energies below a certain critical value and only one above that value. Although the Helmholtz theorem does not hold for energy below the critical value, it is easy to adapt it and give the expression for the entropy if we know on what trajectory the system *actually* is. In this case P and T would be calculated as time averages over the *actual* trajectory and the entropy would be given by the area enclosed by *that* trajectory only. Let us illustrate this with a practical example. Let us consider a 1D chain composed of three particles which interact via a Lennard-Jones potential. Let us fix the position of two of them and let us place the third one in between, so that the first two particles act as walls of a 1D box. Let us now study the behavior of the particle inside the box. Let the interaction potential be:

$$(91) \quad u(x) = \frac{1}{x^{12}} - \frac{1}{x^6}$$

and let us place the “walls” at $x = \pm V/2$. Then the particle in the box is subject to the following potential:

$$(92) \quad \varphi(x, V) = u(x + V/2) + u(x - V/2)$$

For values of V larger than a certain critical value $V_c \simeq 2.5$, this system has a critical energy $E_c(V) = \phi(0, V)$ such that for energy below $E_c(V)$ ergodicity is broken and there are two trajectories per energy level. Above E_c the dynamics is ergodic and there is only one trajectory per energy level. Figure 5.1 shows the graph of $\varphi(x, 4)$ and Figure 5.2 shows a contour plot of various energy levels in phase space for a particle of mass $m = 1$ in the Lennard-Jones box of size $V = 4 > V_c$. For energy $E = E_c$ we have a separatrix. Below E_c the curve of constant energy splits into two disconnected curves, whereas for values of E larger than E_c we have only one curve. Below the critical energy the volume integral Φ is given by the area enclosed by *one*

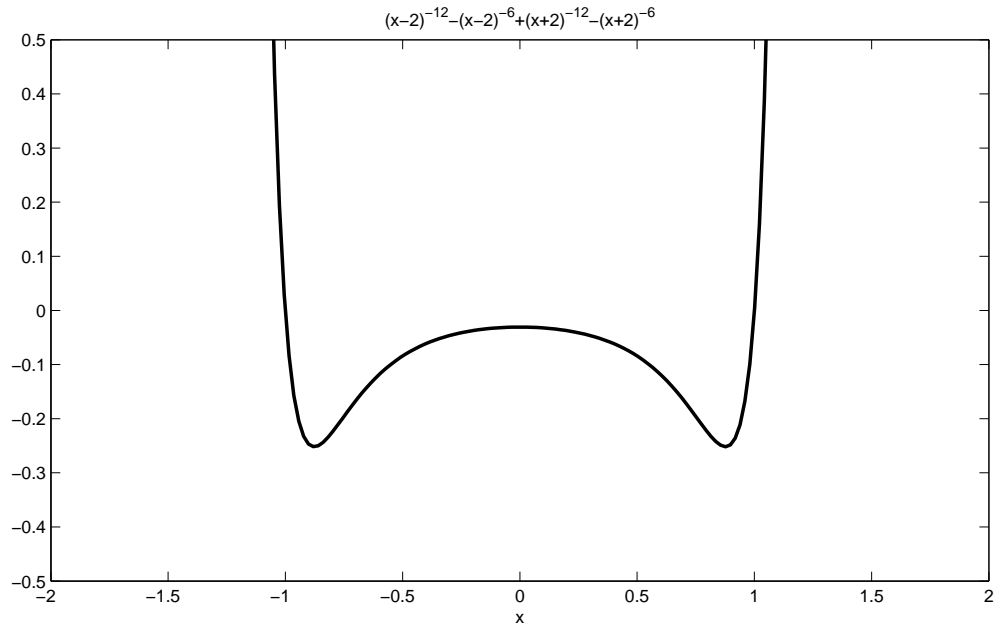


FIGURE 5.1. Lennard-Jones box potential profile $\varphi(x, V)$ for $V = 4 > V_c$.

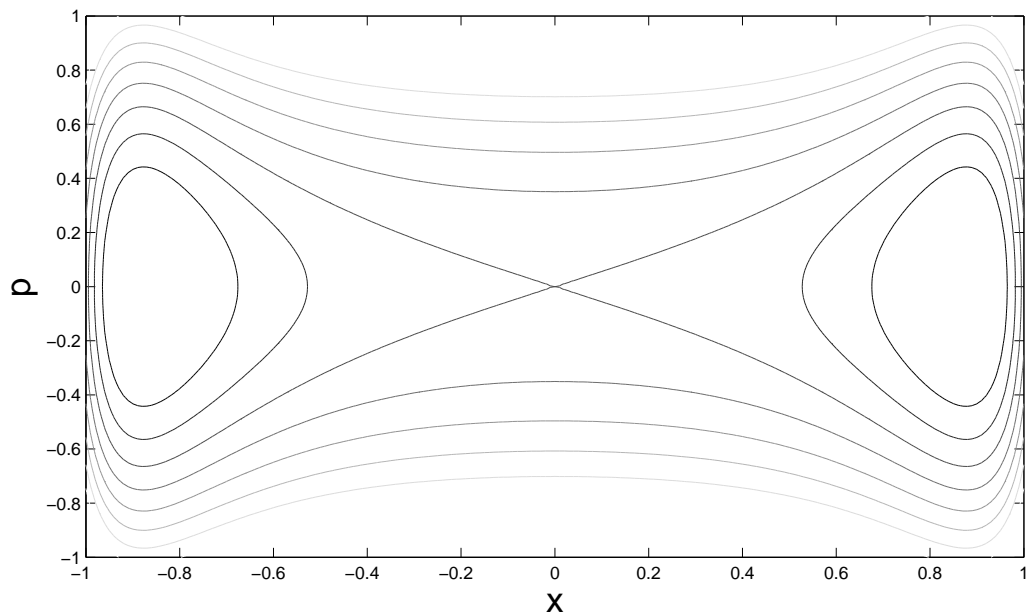


FIGURE 5.2. Phase space structure for a particle of mass $m = 1$ in a Lennard-Jones box of size $V = 4 > V_c$. The separatrix corresponds to the critical energy $E_c = -0.0308$. Below the critical energy there are two distinct trajectories (the dynamics is not ergodic).

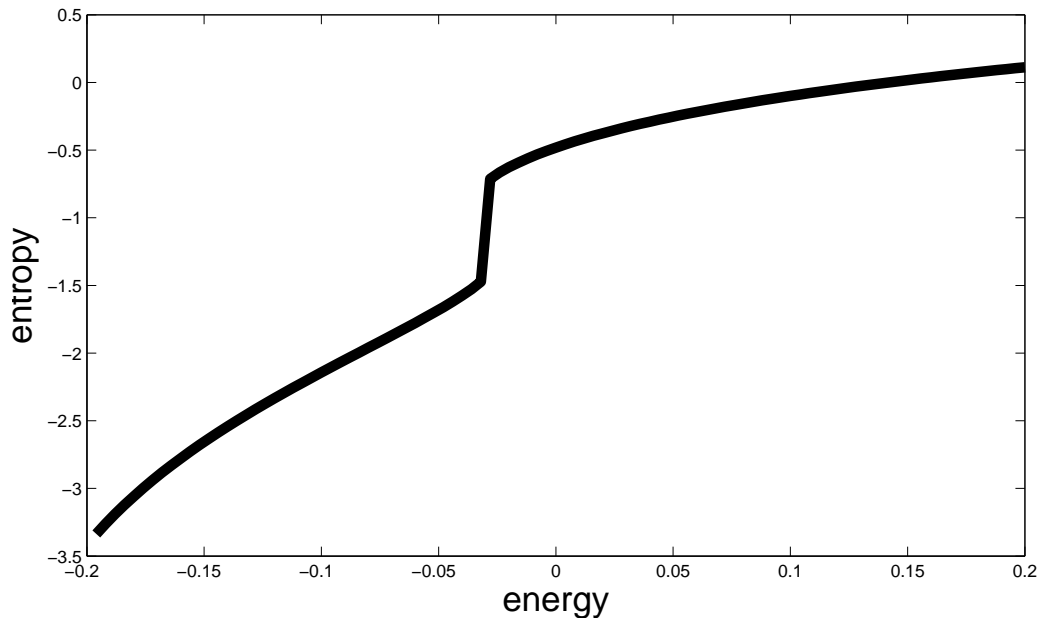


FIGURE 5.3. Entropy versus energy for a particle of mass $m = 1$ in a Lennard-Jones box of size $V = 4 > V_c$. The discontinuity of entropy at the critical energy signals a discontinuous phase transition.

of the two possible trajectories. As the energy crosses the critical value the integral Φ jumps discontinuously. In formulae we have:

$$(93) \quad \Phi(E) = \left[\frac{1}{2} \theta(E_c - E) + \theta(E - E_c) \right] \int [2m(E - \varphi(x, V))]_+^{1/2} dx.$$

The symbol $[y]_+^{1/2}$ denotes a function that is equal to \sqrt{y} for $y \geq 0$ and is null otherwise. The volume entropy, which is calculated by taking the logarithm of the expression above, then displays a jump at the critical energy as well. Figure 5.3 shows a plot of S as a function of E , for the values $m = 1$ and $V = 4 > V_c$. The critical energy is $E_c = -0.0308$. Figure 5.4 shows the temperature plotted against the energy calculated, according to the microcanonical equipartition theorem of Eq. (16):

$$(94) \quad T \doteq \langle 2K \rangle_t = \frac{\Phi(E)}{\Omega(E)} = \left(\frac{\partial S_\Phi}{\partial E} \right)^{-1}.$$

There is a region of negative slope in the graph which correspond to a *negative heat capacity*.

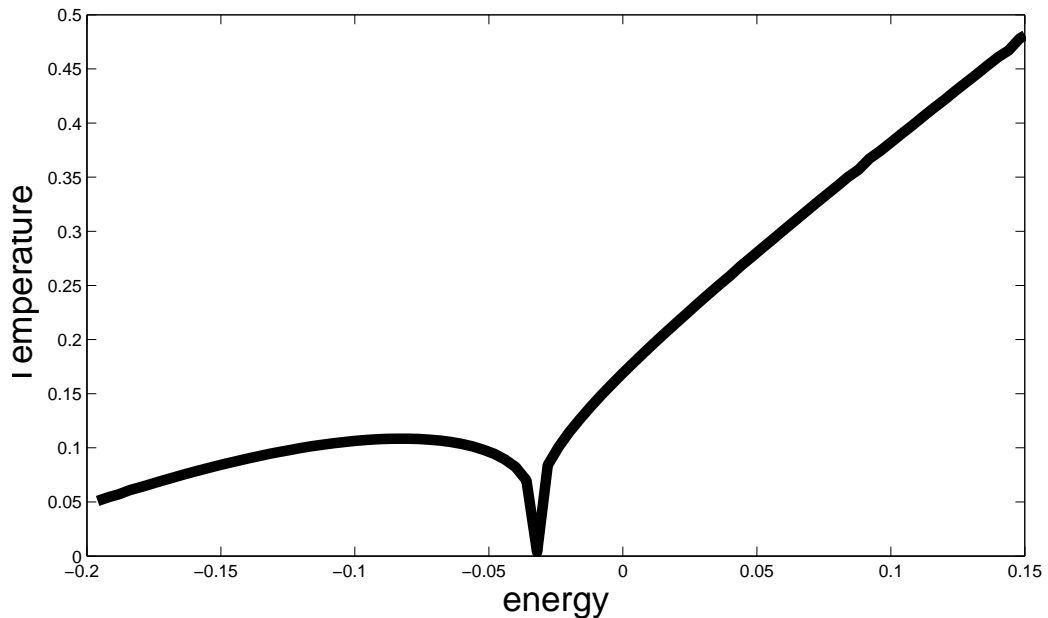


FIGURE 5.4. Temperature versus Energy for a particle of mass $m = 1$ in a Lennard-Jones box of size $V = 4 > V_c$. The curve displays a region of negative heat capacity. At the critical energy the temperature goes to zero.

5.3. Discussion

The example provided in the previous section is perhaps too simple to be of interest to any specific physical problem. Nonetheless it illustrates qualitatively the mechanism of microcanonical phase transition as captured by the volume entropy. Such phase transitions are associated with the crossing of separatrix trajectories, for which the dynamics of the system has no finite time scale. The figures show neatly that at the separatrix energy the entropy has a discontinuous jump, the temperature goes to zero, and for energies below the critical energy we have a region of negative heat capacity. These are not special features of the system studied.¹ Whenever a separatrix is crossed there is a sudden *open-up* of a larger portion of phase space

¹A similar behavior has been observed in a chain of particles interacting via a ϕ^4 potential studied with volume entropy [54] and has been predicted for the pendulum [50].

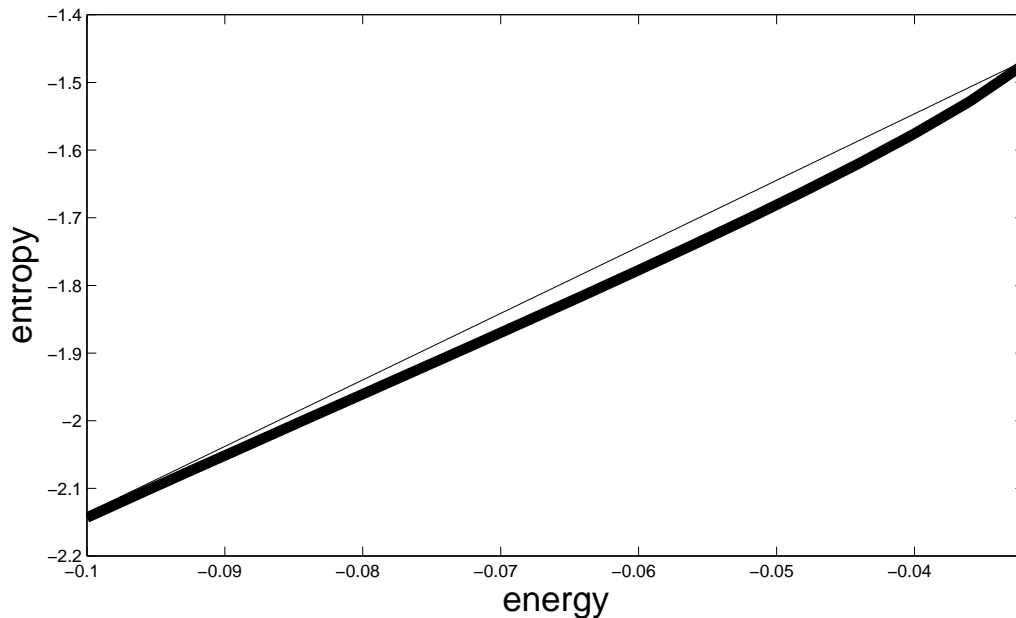


FIGURE 5.5. Entropy versus Energy for $E < E_c$. Right below the critical energy the entropy function (thick line) is convex (the thin straight line is only a guide for the eye).

for the trajectory to *enclose* which leads to a discontinuity in the entropy². Further, at the separatrix, the period of motion, which for a well known theorem of classical mechanics is given by $\Phi' = \Omega$ [24], becomes infinite. Therefore the temperature, *i.e.*, $T = \Phi/\Omega$ goes to zero. Since the temperature $T \doteq 2 \langle K \rangle_t = \langle p^2/m \rangle_t$ is positive definite, below the critical energy there *necessarily* is a region of negative slope, that is negative heat capacity. The appearance of a negative heat capacity is associated with a convex dip in the entropy (see Fig. 5.5) which signals the approach to the separatrix from below. Such convexity appears *necessarily* because $\lim_{E \rightarrow E_c^-} \frac{\partial S(E)}{\partial E} = \lim_{E \rightarrow E_c^-} \frac{1}{T(E)} = +\infty$.

It is important to notice that using the surface entropy would lead to a *drastically different* result. In this case the entropy, instead of having a discontinuous jump, would diverge! Furthermore the temperature would be calculated as $T_\Omega = \Omega/\Omega'$,

²The idea that microcanonical phase transitions are associated to sudden open-up of phase space has been expressed also in Ref. [48]

which might not tend to zero at the critical energy! Note also that T_Ω is not proportional to the average kinetic energy and can be negative. Therefore, in agreement with Ref. [50] we believe that surface entropy is not suited for low dimensional systems with broken ergodicity.

The volume entropy could be used to address microcanonical phase transitions in small dimensional systems with either long- or short- range interactions, like the ϕ^4 model, chains of particles interacting via Lennard-Jones potential [55] or the Hamiltonian Mean Field model [56]. All these models are expected to undergo a breaking of ergodicity [55, 57]; thus there are separatrix trajectories and possible phase transitions that the volume entropy can detect.

One of the most well-known no-go theorems of statistical mechanics says that it is impossible to have phase transitions in 1D chains of particles interacting via short ranged forces [58]. Our result is seemingly in contrast with that theorem. Indeed, that theorem says that the *canonical partition function* $Z(\beta)$ is analytical. This does not preclude the *microcanonical volume integral* $\Phi(E)$ from being non-analytical. In other words the theorem says that it is impossible to have *canonical* phase transitions in 1D chains, thus leaving open the possibility of *microcanonical* phase transitions. As a matter of fact the existence of a microcanonical phase transition in the 1D chain studied above confirms that the two ensembles are *nonequivalent*.

The advantage of using the volume entropy is that it provides a good mechanical analogue of thermodynamic entropy even for small system, thus accounting properly for the finite-size effects. As the development of technology is allowing experimentalists to probe the thermodynamic behavior of smaller and smaller systems, this is becoming an increasingly important task. The main limitation of the present approach is that it is restricted to classical statistical mechanics, and thus it does not account for quantum-mechanical phenomena. More work in this direction is certainly necessary.

Recently also Dunkel and Hilbert have pointed out the importance of adopting the volume entropy instead of surface entropy [59]. They studied the thermodynamic behavior of few simple models of interacting particles inside a box. They studied the square-well interaction potential (which is exactly solvable [60]), the diatomic Lennard-Jones molecule (exactly solvable), many particle Lennard-Jones chains as well as the Takahashi gas [59]. The more interesting feature that appears when dealing with many degrees of freedom is the appearance of oscillations in the $T(E)$ graph. Each oscillation corresponds to one particle jumping from the bound state to the free state. The graphs show a low-energy smooth cluster phase region, an oscillatory intermediate region, corresponding to the phase transition whose extension can be understood as the *latent heat*, and a high-energy smooth region corresponding to the gas phase.

The idea of non-analyticity of microcanonical entropy as a signature of phase transitions appears also in the work of M. Kastner *et.al.* [61]. These authors have put forward a very useful equation that relates the degree of non-analyticity of the microcanonical entropy with the order of the saddle points in the potential energy landscape. According to [61] the order of non-analyticity of the microcanonical entropy increases with the number of degrees of freedom so no phase transition should be observable in macroscopic isolated systems. The authors argue that these may result from the dense accumulation of many non-analytical points in the thermodynamic limit.

CHAPTER 6

OPEN ISSUES IN NONEXTENSIVE THERMODYNAMICS

Around 20 years ago a generalization of the standard Gibbs-Boltzmann statistics was proposed by C. Tsallis [62]. The generalization scheme proposed by Tsallis was based on the adoption of a generalized Shannon's informational entropy, which, when maximized under the constraints of normalization and average energy, leads to deformed exponential statistics. Since then, the corresponding generalized thermostatics, also named non-extensive thermodynamics, has proved to be a very powerful tool of investigation within the most diverse fields.

Many authors have addressed two problems connected with the employment of Tsallis nonextensive thermostatics. The first is that of providing a theoretical foundations other than the information-theoretical one. The second is that of defining an even larger class of ensembles which include the former ones. This second part of the dissertation will address both these two problems.

First we are going to study the *mechanical foundations of Tsallis ensembles*. In order to do so, we shall return again to Boltzmann's work on the *heat theorem* [12]. Thus as in First part of the dissertation we are taking again the fundamental equation of thermodynamics:

$$(95) \quad \frac{dE + PdV}{T} = \text{exact differential} = dS$$

as our starting point. Boltzmann used the heat theorem as a criterion to assess whether a certain statistical ensemble provides *a mechanical model of thermodynamics*, and called the ensembles that satisfy such a criterion *orthodes*. In Chap. 7 we shall provide a definition of orthode and we will see that Tsallis ensembles satisfy the orthodicity requirements, not only microcanonical and canonical ensembles.

Other interesting properties of Tsallis ensembles are investigated in Chap. 8. By slightly modifying an already present theorem due to Almeida [63], we show that they describe the behavior of (possibly small) classical systems in contact with *finite heat baths*. We also prove that in the limit of infinite bath heat capacity (thermalization) the Tsallis ensemble reduces to the canonical one, and that, in the limit of null heat capacity (thermal insulation), they reduce to the microcanonical one as well. This allows us to reconsider the standard derivation of the microcanonical ensemble from a maximum entropy principle.

Finally, in Chap. 9 we generalize further and show that the family of Tsallis ensemble is a subset of a more general class of orthodes that we shall call *dual orthodes*. Canonical, microcanonical, Tsallis and Gaussian ensembles all belong to the class of dual orthodes. We also address questions concerning their *equivalence*.

CHAPTER 7

TSALLIS ENSEMBLES AND THE HEAT THEOREM

7.1. Definition of Orthode

The idea on which Boltzmann's notion of orthode is based is quite simple. Consider a generic statistical ensemble of distributions, $\rho(\mathbf{z}; \lambda_0, \lambda_i)$ parameterized by one "internal" parameter λ_0 and a given number of "external" parameters λ_i $i = 1, 2, 3, \dots$. The symbol \mathbf{z} denotes the phase space point vector $\mathbf{z} = (p_1, p_2, \dots, p_n, q_1, q_2, \dots, q_n)$. We assume the system to be Hamiltonian with Hamilton function of the form $H = K(\mathbf{p}) + \varphi(\mathbf{q}; \lambda_i)$. The Hamiltonian depends explicitly on the external parameters, λ_i $i = 1, 2, 3, \dots$, which are also called "generalized displacements". For the sake of simplicity, we consider only one external parameter $\lambda_1 = V$. This may be, for example, the volume of a vessel containing the system. The kinetic energy is assumed to be of the form $K(\mathbf{p}) = \frac{\mathbf{p}^2}{2m}$. Then we define the *macroscopic state* of the system by the set of following quantities:

$$\begin{aligned}
 U_\rho &\doteq \langle H \rangle_\rho && \text{"energy"} \\
 T_\rho &\doteq \frac{2\langle K \rangle_\rho}{n} && \text{"doubled kinetic energy per degree of freedom"} \\
 V_\rho &\doteq \langle V \rangle_\rho && \text{"generalized displacement"} \\
 P_\rho &\doteq \left\langle -\frac{\partial H}{\partial V} \right\rangle_\rho && \text{"generalized conjugated force"},
 \end{aligned}
 \tag{96}$$

where $\langle \cdot \rangle_\rho$ denotes the average over the given ensemble ρ . Now let the parameters λ_0, V change by infinitesimal amounts, and calculate the corresponding change in the macroscopic state. If the state changes in such a way that the fundamental equation of thermodynamics, namely

$$\tag{97} \quad \frac{dU_\rho + P_\rho dV_\rho}{T_\rho} = \text{exact differential},$$

holds, then we say that the ensemble provides a good mechanical model of thermodynamics. Equation (97), that is the *heat theorem* (Statement 1 in Chap. 2), expresses the fact that the heat differential δQ admits an integrating factor (T^{-1}), whose physical interpretation is that of reciprocal absolute temperature (namely doubled average kinetic energy per degree of freedom). Correspondingly the function that generates this exact differential is interpreted as the physical entropy S of the system. The ensembles that satisfy Eq. (97) were called *orthodes* by Boltzmann [12] (see also [2] and [3] for modern expositions). Boltzmann proved that microcanonical and canonical ensembles are *orthodes*, and placed this fact at the very heart of statistical mechanics. After introducing Tsallis ensembles in the next section, we are going to review the proofs of orthodicity of canonical and microcanonical ensembles and prove that Tsallis (escort) ensembles are also orthodic.

7.2. Definition of Tsallis (Escort) Ensembles

The Tsallis escort ensembles of indices $q \leq 1$ can be parameterized either via U, V or β, V and are defined as

$$(98) \quad \rho_q(\mathbf{z}; \square, V) = \frac{[1 - \beta(1 - q)(H(\mathbf{z}; V) - U)]^{\frac{q}{1-q}}}{N_q},$$

where

$$(99) \quad N_q(\square, V) = \int d\mathbf{z} [1 - (1 - q)\beta(H(\mathbf{z}; V) - U)]^{\frac{q}{1-q}}.$$

The symbol “ \square ” has to be replaced by either U or β according to the parametrization adopted. In the (U, V) parametrization, one first fixes U and then adjusts β in such a way that:

$$(100) \quad U = \langle H \rangle_{\rho_q},$$

In this parametrization Eq. (100) defines the function $\beta(U, V)$. In the (β, V) parametrization, β is fixed instead and U is adjusted accordingly, in such a way that Eq. (100) defines the function $U(\beta, V)$.

We shall refer to the possibility of two distinct parameterizations as a property of *duality*. Therefore we shall say that Tsallis escort ensembles are *dual* ensembles.

Let us explain why we are interested in the distributions (98) with index $q \leq 1$. According to the established theory of nonextensive thermodynamics, the probability distribution function of a nonextensive system is obtained via maximization of Tsallis entropy ($S_{ts} \doteq \frac{\sum p_i^q - 1}{1-q}$) subject to constraints on normalization ($\sum p_i = 1$) and escort-averaged energy ($\frac{\sum p_i^q \varepsilon_i}{\sum p_i^q} = U$). The result of the maximization procedure is the Tsallis ensemble ($p_i = \mathcal{N}_q^{-1} [1 - \beta(1 - q)(\varepsilon_i - U)]^{\frac{1}{1-q}}$) [64]¹. The rule for the calculation of average quantities is that the escort version of the distribution p_i should be employed. This is defined as $P_i = \frac{p_i^q}{\sum p_i^q}$ [64]. In other words, according to the theory of nonextensive thermodynamics, averages should be calculated over the distribution $P_i = \mathcal{N}_q^{-1} [1 - \beta(1 - q)(\varepsilon_i - U)]^{\frac{q}{1-q}}$, or, in the continuum case, over the distribution in Eq. (98)². If we want to study the foundations of nonextensive theory adopting Boltzmann's method, then we must check the orthodicity of the escort ensemble (98) because this is the ensemble that we have to employ when defining the state of the system (96). As we will see below, we are interested in $q \leq 1$, because this is a necessary condition for orthodicity of the corresponding escort ensemble. But first let us recall the proofs of orthodicity of the canonical and microcanonical ensembles

7.3. Orthodicity of the Canonical Ensemble

The canonical ensemble is parameterized by one internal parameter, usually indicated by the Greek letter β , and the external parameter V :

$$(101) \quad \rho_c(\mathbf{z}; \beta, V) = \frac{e^{-\beta H(\mathbf{z}; V)}}{Z(\beta, V)}$$

¹Note that the quantities p_i , U , \mathcal{N}_q , β introduced here correspond respectively to $\tilde{p}_i^{(nor)}$, $\tilde{U}^{(nor)}$, $\tilde{Z}_q^{(nor)}$ and β^* of Ref. [64]. Note also that β does not correspond to the Lagrange multiplier used in the Tsallis maximum entropy principle.

²Note the exponent $\frac{1}{1-q}$ of p_i , as opposed to the exponent $\frac{q}{1-q}$ of P_i .

where

$$(102) \quad Z(\beta, V) \doteq \int d\mathbf{z} e^{-\beta H(\mathbf{z}; V)},$$

The following function:

$$(103) \quad S_c(\beta, V) = \beta U_c(\beta, V) + \ln Z(\beta, V)$$

generates the heat differential, and therefore is the entropy associated with the canonical orthode. To prove that, we have to calculate the partial derivatives of S_c :

$$(104) \quad \frac{\partial S_c}{\partial \beta} = U_c + \beta \frac{\partial U_c}{\partial \beta} - \langle H \rangle_c = \beta \frac{\partial U_c}{\partial \beta}$$

$$(105) \quad \frac{\partial S_c}{\partial V} = \beta \frac{\partial U_c}{\partial V} - \beta \left\langle \frac{\partial H}{\partial V} \right\rangle_c = \beta \frac{\partial U_c}{\partial V} + \beta P_c$$

where the symbol $\langle \cdot \rangle_c$ denotes average over the canonical distribution (101) and the state definition (96) has been used. Further according to the canonical equipartition theorem³

$$(106) \quad \left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_c = \frac{1}{\beta}$$

where i counts the degrees of freedom. Therefore, by comparison with (96), $T_c = \frac{1}{\beta}$.

Combining Eqs. (104)-(106) together we get:

$$(107) \quad dS = \beta \left(\frac{\partial U_c}{\partial \beta} d\beta + \frac{\partial U_c}{\partial V} dV + P_c dV \right) = \frac{dU_c + P_c dV_c}{T_c},$$

which proves that the canonical ensemble is an exact orthode.

7.4. Orthodicity of the Microcanonical Ensemble

Let us recall the definition of the microcanonical ensemble from Chap. 3. The microcanonical ensemble is parameterized by (U, V) :

$$(108) \quad \rho_{mc}(\mathbf{z}; U, V) = \frac{\delta(U - H(\mathbf{z}; V))}{\Omega(U, V)}$$

³In this work we do not adopt the rule that repeated indices are summed

where δ denotes Dirac's delta function and

$$(109) \quad \Omega(U, V) \doteq \int d\mathbf{z} \delta(U - H(\mathbf{z}; V))$$

denotes the density of states or structure function [1]. Then the generalized Helmholtz theorem of Chap. 3 tells us that the volume entropy

$$(110) \quad S_{mc}(U, V) = \ln \int d\mathbf{z} \theta(U - H(\mathbf{z}; V))$$

generates exactly the heat differential. Note that the microcanonical equipartition theorem of Eq. (15) is essential in proving the generalized Helmholtz theorem.

7.5. Tsallis Escort Ensembles are Exact Orthodes

In order to prove orthodicity of Tsallis escort ensembles of Eq. (98) let us first introduce the following quantity:

$$(111) \quad \mathcal{N}_q(\square, V) \doteq \int d\mathbf{z} [1 - (1 - q)\beta(H(\mathbf{z}; V) - U)]^{\frac{1}{1-q}}.$$

As we have seen above, the proof of orthodicity of an ensemble is based on the corresponding *equipartition theorem* (Eq. (106) for the canonical ensemble and Eq. (15) for the microcanonical one). Interestingly there exists a Tsallis equipartition theorem as well [65]:

$$(112) \quad T_q = \left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_{\rho_q} = \frac{1}{\beta}.$$

The proof of this formula is based on the well known fact that, as a consequence of Eq. (100) one has:

$$(113) \quad \mathcal{N}_q(\square, V) = N_q(\square, V).$$

Using the Tsallis equipartition of Eq. (112) and Eq. (113) we can prove that the following quantity

$$(114) \quad S_q(\square, V) = \ln \mathcal{N}_q(\square, V)$$

generates the heat differential both in the U, V and β, V parameterizations. Let us treat the U, V case first. Remember that within this parametrization β is a function of U, V . Therefore:

$$\begin{aligned}
\frac{\partial \mathcal{N}_q}{\partial U} &= \frac{\partial}{\partial U} \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{1}{1-q}} \\
&= \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{q}{1-q}} \left[-\frac{\partial \beta}{\partial U}(H - U) + \beta \right] \\
(115) \quad &= -N_q \frac{\partial \beta}{\partial U} \langle H - U \rangle_{\rho_q} + N_q \beta = N_q \beta = \frac{\mathcal{N}_q}{T_q}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \mathcal{N}_q}{\partial V} &= \frac{\partial}{\partial V} \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{1}{1-q}} \\
&= \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{q}{1-q}} \left[\frac{\partial \beta}{\partial V}(U - H) - \beta \frac{\partial H}{\partial V} \right] \\
(116) \quad &= N_q \frac{\partial \beta}{\partial V} \langle U - H \rangle - N_q \beta \left\langle \frac{\partial H}{\partial V} \right\rangle = N_q \beta P = \mathcal{N}_q \frac{P_q}{T_q}
\end{aligned}$$

where according to the general state definition of Eq. (96) $P_q = \langle \frac{\partial H}{\partial V} \rangle_{\rho_q}$ and $U = U_q$.

Then, combining Eqs. (112), (115) and (116) together, we obtain:

$$(117) \quad dS_q = \frac{d\mathcal{N}_q}{\mathcal{N}_q} = \frac{dU_q + P_q dV}{T_q},$$

which proves orthodicity in the U, V parametrization.

Similarly, within the β, V parametrization one has $U = U(\beta, V)$, and thus:

$$\begin{aligned}
\frac{\partial \mathcal{N}_q}{\partial \beta} &= \frac{\partial}{\partial \beta} \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{1}{1-q}} \\
&= \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{q}{1-q}} \left[-(H - U) + \beta \frac{\partial U}{\partial \beta} \right] \\
(118) \quad &= -N_q \langle H - U \rangle + N_q \beta \frac{\partial U}{\partial \beta} = \frac{\mathcal{N}_q}{T_q} \frac{\partial U}{\partial \beta},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \mathcal{N}_q}{\partial V} &= \frac{\partial}{\partial V} \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{1}{1-q}} \\
&= \int d\mathbf{z} [1 - (1 - q)\beta(H - U)]^{\frac{q}{1-q}} \left[\beta \frac{\partial U}{\partial V} - \beta \frac{\partial H}{\partial V} \right]
\end{aligned}$$

$$(119) \quad = N_q \beta \frac{\partial U}{\partial V} - N_q \beta \left\langle \frac{\partial H}{\partial V} \right\rangle = \frac{N_q}{T_q} \left(\frac{\partial U}{\partial V} + P_q \right).$$

Therefore, we obtain:

$$(120) \quad dS_q = \frac{dN_q}{N_q} = \frac{\frac{\partial U}{\partial \beta} d\beta + \frac{\partial U}{\partial V} dV + P_q dV}{T_q}.$$

However, dU is

$$(121) \quad dU = \frac{\partial U}{\partial \beta} d\beta + \frac{\partial U}{\partial V} dV,$$

so Eq. (120) becomes:

$$(122) \quad dS_q = \frac{dU_q + P_q dV}{T_q}$$

which proves the orthodicity of Tsallis ensembles within the β, V parametrization.

7.6. Remarks

7.6.1. *Duality*

The fact that two parameterizations are available for the Tsallis escort ensembles can be thought of as a *duality* property that these ensemble (98) do not share with the canonical and microcanonical ensembles. Furthermore it suggests that the Tsallis ensemble can be thought of as a *hybrid* ensemble, in between the canonical and the microcanonical ensemble. In fact, we see in the Chap. 8 that the Tsallis escort ensembles interpolate between canonical and microcanonical ensembles as q ranges from 1 to $-\infty$.

7.6.2. *Connection with Rényi Entropy*

It is important to point out that the mechanical treatment we have presented above is independent and alternative to the standard information theoretic approach. We have considered an ensemble and verified whether it is an orthode, adopting the same procedure used by Boltzmann, much before any information-theoretic approach

was developed. The entropy associated with the Tsallis escort orthodes, i.e., Eq. (114), *coincides* with the Rényi entropy [66]

$$(123) \quad \bar{S}_a[\rho] = \frac{1}{1-a} \ln \int d\mathbf{z} \rho(\mathbf{z})^a$$

of order $a = 1/q$ calculated over the power-law distribution in Eq. (98). This fact follows straightforwardly from Eq. (113). In the Chap. 8 we show that, on the other hand, the constrained maximization of the $1/q$ -Rényi entropy leads back to the q -Tsallis escort ensembles in Eq. (98).

7.6.3. Mechanical Additivity

It is known that the Rényi entropy is an extensive and additive quantity [67], where the qualifier “additive” refers to the factorization of probabilities. Adopting a mechanical viewpoint rather than information theoretic, we may ask whether the Rényi entropy of Eq. (114) is additive in a mechanical sense, referring rather to the decomposition of the system into two noninteracting sub-systems with Hamiltonian H_1 and H_2 with $H = H_1 + H_2$. Then it is easy to see that if the canonical entropy is additive (thanks to the factorization of the exponential ($e^{-\beta(H_1+H_2)} = e^{-\beta H_1} e^{-\beta H_2}$) [1]). The entropy in Eq. (114), is not, because the power-law does not factorize.

7.6.4. Exact Orthodicity

In our view, it is very important to stress that Eq. (117) holds independent of the number of degrees of freedom of the system, i.e., the ensemble (98) is an *exact orthode*. In summary, the main result found here states that the Tsallis escort ensemble provides a mechanical model of equilibrium thermodynamics even for low dimensional systems. Exact orthodicity is not an exclusive feature of the Tsallis ensemble, as microcanonical and canonical ensembles are exact orthodes too. Therefore, in principle, all three ensembles are potentially valid for describing low dimensional systems. The result proved above holds for interacting Hamiltonians, *without restriction to ideal systems*. In these cases one should first check that the function $\beta(U, V)$ exists and is differentiable. Whether the result also holds for systems with long-range interactions,

like in the Hamiltonian Mean Field model introduced by Antoni and Ruffo [56], is a very interesting problem which deserves further studies.

7.6.5. *Free Energy*

Due to orthodicity, namely the fact that standard thermodynamic relations exist among the quantities U, P, T, V, S , one can construct the thermodynamic potentials as usual by means of Legendre transformations. For example, the free energy is given by:

$$(124) \quad F = U - TS = U - T \ln \mathcal{N}_q.$$

This expression is in agreement with the results of Ref. [68, see Eq. (29)] where the free energy is derived from the requirement that it should be expressed as a function of the physical temperature β^{-1} . Here, Eq. (124) is obtained by adopting Boltzmann's general method, namely by proving that the heat theorem holds within the Tsallis escort ensemble. Our approach is a very fundamental and unifying one. Thanks to Boltzmann's concept of orthodicity, Tsallis escort ensembles have been framed within the general theory of statistical ensembles beside the microcanonical and the canonical ensembles.

7.6.6. *Connection with Tsallis Entropy*

Let us now focus on the connection of this work with the standard nonextensive approach where the leading role is played by the Tsallis entropy. It is quite simple to show that, for any monotonic C^1 function $g(x)$, the quantity $\mathcal{N}_q g'(\mathcal{N}_q) \beta$ is an integrating factor for $dE + PdV$ and the associated entropy is $S^{(g)} = g(\mathcal{N}_q)$. In fact, using Eqs. (115) and (116) would lead to

$$(125) \quad dS^{(g)} = \mathcal{N}_q g'(\mathcal{N}_q) \beta (dE + PdV).$$

Using the condition that the integrating factor is the average kinetic energy, i.e., β , leads us to $\mathcal{N}_q g'(\mathcal{N}_q) = 1$, which returns us back to the logarithmic entropy $S = \log \mathcal{N}_q + const$. On the other hand, using the condition that the integrating factor is

equal to the Lagrange multiplier used in the Tsallis maximization procedure, which in our notation would read $\beta_L = \beta \mathcal{N}_q^{1-q}$, leads to the Tsallis entropy form:

$$(126) \quad S_q^{ts}(\mathcal{N}_q) = \alpha(\mathcal{N}_q^{1-q} - 1) + const = \ln_q \mathcal{N}_q + const$$

where the symbol $\ln_q(x)$ indicates the q-deformed logarithm [69]. Therefore

$$(127) \quad dS_q^{ts} = \beta_L \delta Q$$

This fact has also been acknowledged very recently in [70]. The present approach makes evident that the quantity $\delta Q = dE + PdV$ admits infinitely many different integrating factors associated with as many different entropies. The Tsallis entropy is the one associated with the integrating factor β_L . Therefore, using the Tsallis entropy the thermodynamic relations would still hold, even though the integrating factor would not coincide with the average kinetic energy. Nonetheless, the fact that the Lagrange multiplier β_L does not have such a straightforward physical interpretation as the quantity β , also poses some questions regarding the physical interpretation of the associated entropy (126) too. On the contrary, the fact that the logarithmic entropy $S = \ln \mathcal{N}$ is associated with the *physical temperature* i.e.,

$$(128) \quad dS = \beta \delta Q,$$

makes it the ideal candidate to play the role of *physical entropy* or *Clausius entropy* within the nonextensive framework.

CHAPTER 8

MORE ON TSALLIS ESCORT ENSEMBLES

8.1. Interpolating between Canonical and Microcanonical Ensembles

As anticipated in Chap. 7, the Tsallis escort ensembles of Eq. (98) have an interesting duality property: they can either be parameterized through U, V or β, V . In addition, regardless of the parametrization adopted, these ensembles are exact orthodes. This duality suggested that they are a sort of *hybrid* of the microcanonical and canonical ensembles. This is true not only from a qualitative point of view but also from a quantitative one.

As is well known, the ensemble (98) is indeed a generalization of the canonical one to which it tends when q goes to 1. In fact, from the properties of the q -exponential [69], we have:

$$(129) \quad \lim_{q \rightarrow 1} \rho_q = \frac{e^{-\beta H}}{Z} = \rho_c, \quad \lim_{q \rightarrow 1} S^{[q]} = \beta U + \ln Z = S_c.$$

Note that, in this limit the explicit dependence of the distribution ρ_q on U disappears, namely the *duality* property is lost in the limit $q \rightarrow 1$, and the only possible parametrization is the (β, V) one.

In a similar manner, the microcanonical orthode is a special case of the Tsallis orthode recovered in the limit $q \rightarrow -\infty$, in which case the explicit dependence on β disappears. In this limit, again, the *duality* is lost and the only possible parametrization is the (U, V) one. Considering the q -exponential function:

$$(130) \quad e_q(x) = \begin{cases} 0 & 1 + (1 - q)x < 0 \\ [1 + (1 - q)x]^{1/(1-q)} & 1 + (1 - q)x \geq 0 \end{cases}$$

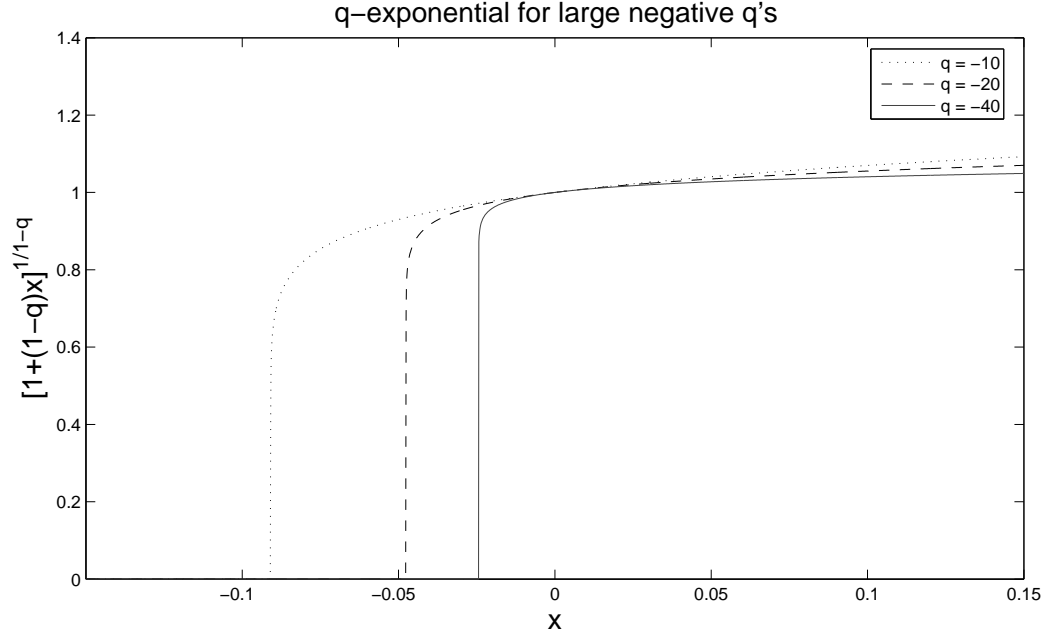


FIGURE 8.1. Plot of q -exponential (Eq. 130) for large negative values of q . The function tends to the Heaviside step function.

it is easily seen that:

$$(131) \quad \lim_{q \rightarrow -\infty} e_q(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} = \theta(x).$$

This fact is illustrated in Figure 8.1. Therefore the Rényi entropy of Eq. (114) tends to the microcanonical entropy, that is the volume entropy of Eq. (8), when q goes to $-\infty$

$$(132) \quad \lim_{q \rightarrow -\infty} S^{[q]} = S_{mc},$$

where we have used the fact that $\theta(\beta x) = \theta(x)$ for $\beta > 0$. Furthermore, the Tsallis ensemble (98) is expressed in terms of the “derivative” of the q -exponential

$$(133) \quad e'_q(x) = \begin{cases} 0 & 1 + (1 - q)x < 0 \\ [1 + (1 - q)x]^{q/(1-q)} & 1 + (1 - q)x \geq 0 \end{cases}$$

which tends to a Dirac delta function. This is graphically illustrated in Fig. 8.2. A

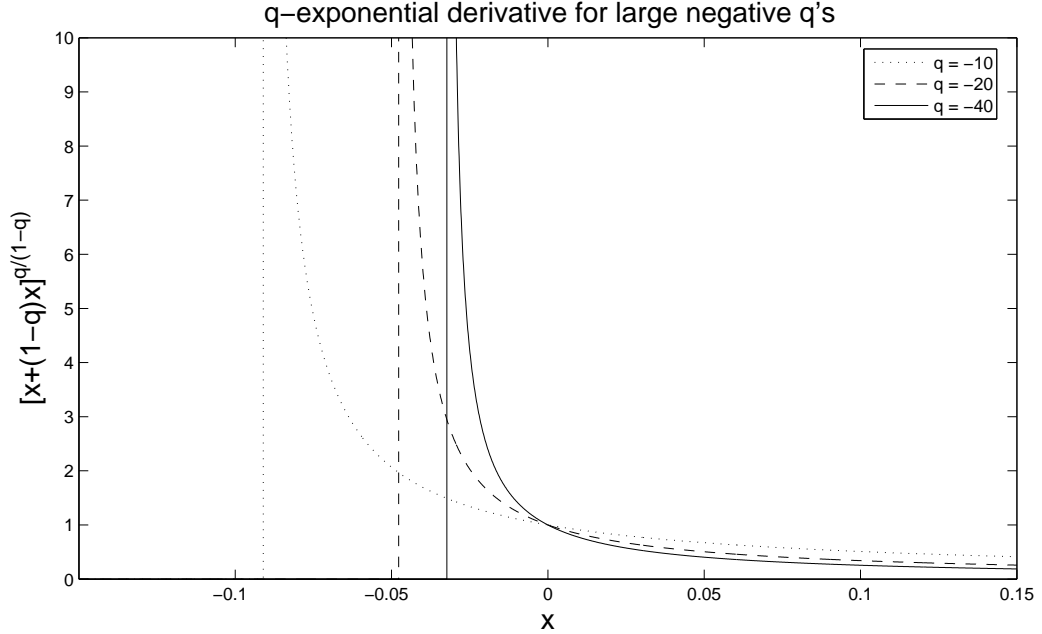


FIGURE 8.2. Plot of q -exponential derivative (Eq. 133) for large negative values of q . The function approaches a Dirac delta function.

simple way to prove this result is to consider the Fourier transform of the function e'_q :

$$(134) \quad \widehat{e}'_q(\omega) = e^{-i\omega\alpha_q} \frac{\Gamma(\alpha_q + 1)}{(-i\omega\alpha_q)^{\alpha_q}}$$

where for simplicity we have used the notation $\alpha_q = 1/(1 - q)$ and Γ is the Gamma function. Now it is easily seen that, in the limit $q \rightarrow -\infty$ ($\alpha_q \rightarrow 0^+$), $\widehat{e}'_q(\omega) \rightarrow 1$ which is the Fourier transform of the Dirac delta. Therefore the function $e'_q(x)$ tends to the delta function. From this it follows that the distribution ρ_q tends to the microcanonical distribution of Eq. (12):

$$(135) \quad \lim_{q \rightarrow -\infty} \rho_q = \frac{\delta(U - H(\mathbf{z}; V))}{\int \delta(U - H(\mathbf{z}; V))} = \rho_{mc}$$

It is evident that, due to the property $\delta(\beta x) = \beta^{-1}\delta(x)$, in the limit $q \rightarrow -\infty$, the explicit dependence on β disappears. This is understood also based on the fact that the distribution is extremely peaked, so the average of H always equals U , regardless of the value taken by β .

The microcanonical equipartition theorem also is recovered in the limit $q \rightarrow -\infty$. For finite q 's, thanks to Eq. (113), we have $N_q = \mathcal{N}_q$ hence, from Eq. 112

$$(136) \quad T_q = \frac{1}{\beta} \quad (\text{finite } q\text{'s}).$$

Therefore the canonical equipartition theorem (see Eq. 106) is trivially recovered for $q = 1$. Nonetheless, when q goes to infinity, the relation $N_q = \mathcal{N}_q$ stops holding. In fact we have

$$(137) \quad \lim_{q \rightarrow -\infty} \mathcal{N}_q = \Phi \quad , \quad \lim_{q \rightarrow -\infty} N_q = \frac{1}{\beta} \Omega$$

so that $\lim_{q \rightarrow -\infty} T^{[q]} = \frac{\Omega}{\Phi}$ -namely the microcanonical equipartition theorem of Eq. (15) is recovered as well.

8.2. Interpretation: Finite Heat Baths

We shall now investigate the physical meaning of Tsallis escort *hybrid* statistics. In other words, we shall ask ourselves in what physical situation we expect to observe them. It is well known that the microcanonical ensemble describes the statistical properties of isolated systems whereas the canonical one describes the properties of systems in contact with a heat bath. Both ensembles apply to two ideal (nonetheless very useful) cases: that of a system in contact with a heat bath with infinite capacity (the temperature is fixed) and that of an isolated system, namely a system in contact with a bath with null heat capacity (the energy is fixed). Between these two extremal cases lie the physically realistic cases of systems in contact with *finite heat baths, where both energy and temperature are allowed to fluctuate*. Therefore we find it reasonable to expect such systems to obey Tsallis statistics of some order q , where q accounts for the heat capacity of the bath. The idea of a finite heat bath is not new in the context of nonextensive thermodynamics. It was first proposed in [71], and then further developed in [63], although the limiting case of null heat capacity was never investigated before. In particular Almeida [63] considered an isolated system of total energy a composed of two non interacting subsystems: the system of interest

(labelled by the subscript 1) and its complement, the bath (labelled by 2). The total Hamiltonian splits into the sum of the two sub-system Hamiltonians, $H = H_1 + H_2$. Using the density of states of the total system (Ω), and that of the bath (Ω_2), one can express the distribution law for the component 1 in its phase space as [1]:

$$(138) \quad p_1(H_1) = \frac{\Omega_2(a - H_1)}{\Omega(a)}.$$

By defining the inverse temperature of the bath as $\beta_2 \doteq \frac{\Omega'_2}{\Omega_2}$, Almeida proved that the bath heat capacity ($C_V^{-1} \doteq \frac{\partial}{\partial E_2} \frac{1}{\beta_2}$) is given by the expression $C_V = \frac{1}{1-q}$, if and only if:

$$(139) \quad \frac{\Omega_2(a - H_1)}{\Omega_2(a)} = e_q(-\beta_2(a)H_1);$$

namely, if and only if $p_1(H_1) \propto e_q(-\beta_2(a)H_1)$. Equation (139) was first put forward by Abe and Rajagopal [72] as an asymptotic formula valid for $H_1 \ll H_2$. The result of Almeida constitutes an improvement and a clarification of the results of Ref. [72]: an improvement because it shows that Eq. (139) holds exactly and a clarification because it provides a physical interpretation of the power-law index $\frac{1}{1-q}$ in terms of the bath heat capacity C_V .

Although Almeida's theorem is in line with our interpretation based on finite heat baths, we believe it to be still subject to further improvement. As Eq. (139) shows, Almeida's theorem leads to a form of Tsallis distribution expressed in terms of e_q rather than the form investigated here expressed in terms of e'_q . The latter would correspond to the *escort* version of the former. As the reader can easily notice the microcanonical distribution is not a special case of the e_q -type distribution. In order to improve the theorem we notice that the definition $\beta_2 \doteq \frac{\Omega'_2}{\Omega_2}$ adopted by Almeida is not consistent with the microcanonical equipartition prescription of Eq. (15). If one adopts the correct definition

$$(140) \quad \beta_2 \doteq \frac{\Omega_2}{\Phi_2}$$

of the bath's temperature, namely if one replaces Ω_2 with Φ_2 , then Almeida's theorem would read: $C_V = \frac{1}{1-q} \Leftrightarrow \frac{\Phi_2(a-H_1)}{\Phi_2(a)} = e_q(-\beta_2(a)H_1)$. By taking the derivative of the latter with respect to H_1 , we obtain the following:

Theorem 3.

$$(141) \quad \frac{\Omega_2(a - H_1)}{\Omega_2(a)} = e'_q(-\beta_2(a)H_1)$$

if and only if

$$(142) \quad C_V \doteq \left(\frac{\partial}{\partial U_2} \frac{1}{\beta_2} \right)^{-1} = \frac{1}{1-q}.$$

Now we notice that $\beta_2(a)$ is the inverse physical temperature that the bath would have if it were isolated and its energy were a . Thanks to Eq. (142) such temperature is given by $\frac{1}{\beta_2(a)} = (1-q)a$. Instead, in the physical situation under study, the bath is in contact with system 1 and its average energy U_2 is smaller than a . The actual temperature in the composite system may be expressed by $\frac{1}{\beta} = (1-q)U_2$. Therefore $\beta(a) = \beta U_2/a$, hence Eq. (141) may be rewritten as $p_1(H_1) \propto \left[1 + \left(\frac{a}{U_2} - 1 \right) - (1-q)\beta H_1 \right]^{\frac{q}{1-q}}$. Using the relation $a = U_1 + U_2$ then p_1 would read exactly as the Tsallis escort distribution of Eq. (98).

In summary, the theorem says that if the heat capacity of the bath is $C_V = \frac{1}{1-q}$, then the component 1 obeys the Tsallis escort distribution law of Eq. (98) of index q . This theorem is consistent with our physical interpretation, according to which q should account for the finiteness of the bath heat capacity. In particular it reproduces well the two limiting cases: if $q \rightarrow 1$ then C_V goes to infinity, namely we are in the case of an infinite bath, and accordingly we get the canonical ensemble. If $q \rightarrow -\infty$ then $C_V \rightarrow 0$, namely we are in the isolated case, and accordingly we get the microcanonical ensemble. By substituting Eqs. (136) and (142) into Eqs. (98) and (114), we are finally in the position to rewrite the distribution law for system 1 and its entropy in

terms of physical quantities as:

$$(143) \quad \rho_{C_V}(\mathbf{z}; U, V) = \frac{\left[1 - \frac{(H(\mathbf{z}; V) - U)}{C_V T}\right]^{C_V - 1}}{N_{C_V}(U, V)}$$

$$(144) \quad S_{C_V}(U, V) = \ln \int d\mathbf{z} \left[1 - \frac{(H(\mathbf{z}; V) - U)}{C_V T}\right]^{C_V},$$

where for simplicity we dropped the subscript 1.

8.3. Fluctuations in the Tsallis Escort Ensemble

Calculation of fluctuations of energy for systems obeying Tsallis escort statistics is interesting. Let us first assume that the system of interest has a power law density of states:

$$(145) \quad \Omega(E) = \text{const} \times E^{\gamma - 1}$$

which is tantamount to assuming that the heat capacity at constant volume of the system is γ . Let us denote by c the constant heat capacity at constant volume of the bath. According to Eq. (143), the system is distributed according to:

$$(146) \quad \rho_c(E) E^{\gamma - 1} dE = N_c^{-1} \left[1 + \frac{\gamma}{c} \left(1 - \frac{E}{U}\right)\right]^{c-1} E^{\gamma - 1} dE$$

where $U = \gamma T$ and we have the cut-off condition:

$$(147) \quad E < U \left(1 + \frac{c}{\gamma}\right) \doteq \bar{U}.$$

It is straightforward to check that:

$$(148) \quad \langle E \rangle = U.$$

We are interested in the relative square fluctuation σ^2 , defined as $\frac{\langle (U-E)^2 \rangle}{U^2} = \frac{\langle E^2 \rangle - U^2}{U^2}$.

One has:

$$(149) \quad \langle E^2 \rangle = \frac{\int_0^{\bar{U}} \left[1 + \frac{\gamma}{c} \left(1 - \frac{E}{U}\right)\right]^{c-1} E^{\gamma+1} dE}{\int_0^{\bar{U}} \left[1 + \frac{\gamma}{c} \left(1 - \frac{E}{U}\right)\right]^{c-1} E^{\gamma-1} dE} = U^2 \frac{(\gamma + c)(1 + \gamma)}{\gamma(1 + \gamma + c)}.$$

So the relative squared fluctuation is:

$$(150) \quad \sigma^2 = \frac{\langle E^2 \rangle - U^2}{U^2} = \frac{c}{\gamma} \frac{1}{1 + \gamma + c}$$

It is worth examining some limiting cases:

$c \gg \gamma$: In this case $\sigma^2 \simeq \frac{1}{\gamma}$. This correspond to the canonical case (infinite heat bath), where, as is well known, the relative square fluctuation of the energy of the system goes as the inverse of the number of degrees of freedom in the system.

$c \ll \gamma$: In this case $\sigma^2 \simeq 0$. This correspond to the microcanonical case (absent heat bath)

$c \simeq \gamma \gg 1$: In this case $\sigma^2 \simeq \frac{1}{\gamma+c}$. This correspond to the gaussian case where both system and bath consist of a very large number of constituents (see Sec. 9.3.2). In this case the relative square fluctuation of the system goes as the inverse of the total number of degrees of freedom of system plus bath, i.e., $c + \gamma$.

It is evident that our distinction between system and bath is purely arbitrary. Indeed one could consider the system as the bath and the bath as the system of interest and all our reasonings would equally well apply. Thus by labelling our two subsystems as subsystem 1 and 2 one can express the fluctuations of both in terms of their respective heat capacities c_1 and c_2 . Evidently:

$$(151) \quad \sigma_1^2 = \frac{c_2}{c_1} \frac{1}{1 + c_1 + c_2}$$

$$(152) \quad \sigma_2^2 = \frac{c_1}{c_2} \frac{1}{1 + c_1 + c_2}.$$

More physical insight can be gained from the two equations above, if we re-express them as:

$$(153) \quad \frac{\sigma_1}{\sigma_2} = \frac{c_2}{c_1}$$

$$(154) \quad \sigma_1 \sigma_2 = \frac{1}{1 + c_1 + c_2}.$$

Using the notation $\Delta E_{1,2} = \sqrt{\langle E_{1,2}^2 \rangle - U_{1,2}^2}$, and noting that, because of thermal equilibrium, $U_1/c_1 = U_2/c_2 = T$, Eq. (153) says that:

$$(155) \quad \Delta E_1 = \Delta E_2$$

That is, the absolute fluctuations in the two systems are identical. In order to understand Eq. (155) let us first consider the limiting situation in which system 1 is much larger than system 2. This means that system 1 is almost in a microcanonical condition and system 2 is almost in a canonical one. System 1 absolute fluctuations are small because of the microcanonical condition, whereas system 2 absolute fluctuations are small because its share of the total energy is small. The same happens if we exchange 1 and 2. The other limiting situation is that in which the two systems have the same heat capacity, in which case they trivially must have the same absolute fluctuations. Eq. (153) says that for all the intermediate cases, the two fluctuations will keep adjusting so as to remain balanced. In the end, the physical reason for Eq. (155) is quite simple. Since the two systems form an *isolated* system, a temporary change in the energy of one is accompanied by the same and opposite change of the other, in such a way that the total energy remains constant.

Eq. (154), expresses an interesting *Complementarity Relation*, which resembles Heisenberg's relations between fluctuations of position and momentum in quantum mechanical *coherent states* of minimal indeterminacy. Eq. (154) says that the product of the relative fluctuation of the two subsystems is constant and determined by the total heat capacity $c_1 + c_2$. The larger the relative fluctuations in system 1, the smaller in system 2 and vice-versa.

8.4. Information-Theory and Tsallis Ensembles

8.4.1. Tsallis Escort Ensembles Maximize Rényi Entropy

Let us now investigate on the information-theoretic basis of Tsallis escort ensembles. As we have pointed out above, the physical entropy associated with them, Eq. (114), *coincides* with the $1/q$ -order Rényi entropy (see Eq. 123) calculated over the

Tsallis escort distribution, Eq. (98). It is quite striking to see that, on the other hand, the Tsallis escort ensembles (98), are obtained from the maximization of the $1/q$ -order Rényi entropy

$$(156) \quad \bar{S}_{\frac{1}{q}}[\rho] = \frac{q}{q-1} \ln \int d\mathbf{z} \rho(\mathbf{z})^{1/q}$$

under the two constraints on the normalization and average energy. The variation of the following functional, obtained with the method of Lagrange multipliers:

$$(157) \quad \mathcal{F}[\rho] = \frac{q}{q-1} \ln \int d\mathbf{z} \rho^{1/q} - \lambda \left(\int d\mathbf{z} \rho - 1 \right) - \beta \left(\int d\mathbf{z} H \rho - U \right)$$

leads to:

$$(158) \quad \frac{\rho^{\frac{1-q}{q}}}{(q-1) \int d\mathbf{z} \rho^{1/q}} - \lambda - \beta H = 0$$

from which it is easily inferred that

$$(159) \quad \rho \propto [\lambda + \beta H]^{\frac{q}{1-q}}.$$

Multiplying Eq. (158) by ρ and integrating over $d\mathbf{z}$ as described in [64, 73] leads to:

$$(160) \quad \lambda = \frac{1}{q-1} - \beta U$$

which combined with (159) leads exactly to the power-law ensemble (98). In this way the information-theoretic approach and the mechanical one are connected reciprocally through a closed loop. Therefore the Tsallis escort ensembles have the same two-fold foundation (i.e., information-theoretic and mechanical) that the canonical ensemble has. This point is illustrated in Fig. 8.3.

8.4.2. *Information Theoretic justification for Maximum Rényi Entropy Principle*

The maximum entropy principle (MEP) of Eq. (157) arises naturally if the Shore and Johnson axioms [74, 75] are considered. In the traditional approach of Shannon, adopted also by Rényi, one sets a few axioms that the information measure should satisfy and then finds the appropriate entropy form that satisfy the axioms in question. Shore and Johnson were the first who shifted the focus of information theory from the

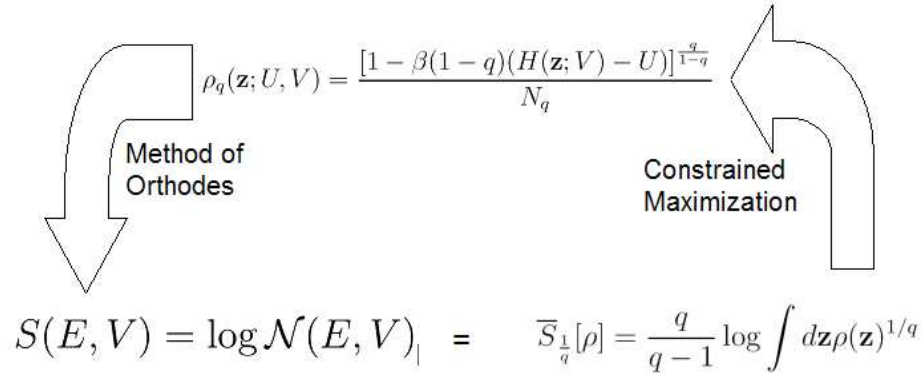


FIGURE 8.3. Two-fold foundations of Tsallis escort ensembles: mechanical (based on Boltzmann’s orthodes method) and information-theoretic (based on the maximum Rényi entropy principle)

information measure to the MEP itself. Thus they axiomatized the *inference rule*, according to which, a *prior probability* r , which represents the initial knowledge of the system is *updated* to a *posterior probability* p when new *information*, in the form of *linear constraints* I , is gained. Shore and Johnson expressed this using the notation:

$$(161) \quad p = I \circ r.$$

The updating procedure is what we apply when using the MEP. We start from a flat distribution, and we update it to a new distribution which accounts with the least possible bias for the information contained in the constraints on average energy and normalization. Shore and Johnson set the following five axioms:

- Axiom I (Uniqueness): If the same problem is solved twice, then the same answer is expected to result both times.
- Axiom II (Invariance): The same answer is expected when the same problem is solved in two different coordinate systems, in which the posteriors in the two systems should be related by the coordinate transformation.

- Axiom III (System Independence): It should not matter whether one accounts for independent information about independent systems separately in terms of their marginal distributions or in terms of the joint distribution.
- Axiom IV (Subset Independence): It should not matter whether one treats independent subsets of the states of the systems in terms of their separate conditional distributions or in terms of the joint distribution.
- Axiom V (Expansibility): In the absence of new information, the prior (i.e., the reference distribution) should not be changed.

The interested reader may find an exhaustive discussion of the meaning and reasonability of these axioms in Ref. [76]. These five axioms select an *equivalence class* of inference rules rather than a single rule. For example, if the rule “maximize H under the constraint I ” satisfies the axioms, then all the rules of the type “maximize H' under the constraints I ” will satisfy the axioms if H and H' have the same maxima. In this case we say that the two rules are equivalent. Thus the rule “maximize H under the constraints I ” is equivalent to any rule of the type “maximize $f(H)$ under the constraints I ”, if f is a strictly monotonic function. According to the original works of Shore and Johnson, the five axioms are satisfied by any rule *equivalent* to the rule “maximize $-\int p(x) \ln p(x)$ under the constraints I ”. Later on Uffink found a flaw in their proof and proved that the following theorem holds indeed [76]:

Theorem 4 (Shore-Johnson-Uffink). An inference rule satisfies all five axioms, if and only if it is equivalent to one of the rules

$$\text{“maximize } H_\alpha \text{ under the constraint } I\text{”}$$

where

$$(162) \quad H_\alpha[p, r] \doteq \left[\int dx p(x)^\alpha r(x)^{1-\alpha} \right]^{\frac{1}{1-\alpha}}.$$

If we assume the prior probability r to be flat, then the theorem evidently tells that any inference rule equivalent to the rule

“maximize $H_\alpha[p]$ under the constraint I ”,

where

$$(163) \quad H_\alpha[p] \doteq \left[\int dx p(x)^\alpha \right]^{\frac{1}{1-\alpha}}$$

satisfies the five axioms. If we take the logarithm of $H_\alpha[p]$ we immediately recognize the Rényi entropies of Eq. (123). Therefore the interpolating maximum entropy principle in Eq. (157) satisfies the five axioms (see also [77]). This can be seen as a justification of the interpolating MEP based on few very general and reasonable consistency requirements that any maximum entropy principle should satisfy.

It is equally evident that, by taking the deformed logarithm of order α (defined as $\ln_\alpha(x) = \frac{1-x^{1-\alpha}}{1-\alpha}$) of H_α , the maximization of Tsallis entropy

$$(164) \quad S_\alpha[p] = \frac{1 - \int dx p(x)^\alpha}{1 - \alpha}$$

satisfies the five axioms as well. This can be seen as an important justification of the maximum Tsallis entropy principle, which, since its very introduction [62], has found many applications. Nonetheless Tsallis entropy does not seem suitable in this specific context concerning systems in equilibrium with finite heat baths, because Eq. (164) does not interpolate correctly. In fact, as $\alpha \rightarrow 0$ it tends to $\ln_\alpha \Phi$. Instead $\ln \Phi$ is the correct equilibrium microcanonical entropy.

8.5. New insights into the Microcanonical Entropy Principle

So far we have seen that the Tsallis escort distributions (98) interpolate between canonical and microcanonical ensemble. They maximize the $1/q$ -order Rényi entropy, and they represent the physical situation of a system in contact with a finite heat bath. This allows us to take the limits $q \rightarrow 1$ and $q \rightarrow -\infty$ in the maximum Rényi entropy principle of Eq. (157) and interpolate between a *canonical maximum entropy principle* (cMEP) and a *microcanonical maximum entropy principle* (μ cMEP).

8.5.1. Canonical MEP

Let $q \rightarrow 1$ in Eq. (157). As is well known [66], Rényi entropy of order 1 is nothing but Shannon's entropy, thus Eq. (157) becomes:

$$(165) \quad \mathcal{F}[\rho] = - \int \rho \ln \rho - \lambda \left(\int d\mathbf{z} \rho - 1 \right) - \beta \left(\int d\mathbf{z} H \rho - U \right)$$

whose maximization (see for example [18]) is the canonical ensemble.

8.5.2. Microcanonical MEP

It is more interesting, though, to consider the limit $q \rightarrow -\infty$. In this limit the integrand $\rho^{1/q}$ tends to 1. We have to be careful with the integration domain at this point, because if the integration is carried over the whole space \mathbb{R}^{6n} , the integral would diverge. Instead, if one considers that the domain of integration is bounded by the condition that the system's energy does not exceed the total energy of system plus bath, which is fixed, we obtain the following cut-off condition:

$$(166) \quad H(\mathbf{z}) \leq U + \frac{T}{1-q}$$

where U is the system's average energy and $T/(1-q) = C_V T$ is the bath's average energy. Now if q goes to 1, the right hand side of Eq. (167) goes to ∞ and the integration domain becomes \mathbb{R}^{6n} for the canonical case. Instead, for q tending to $-\infty$ the cut-off becomes

$$(167) \quad H(\mathbf{z}) \leq U.$$

Thus the microcanonical distribution maximizes the following:

$$(168) \quad \mathcal{F}[\rho] = \ln \int_{H \leq U} d\mathbf{z} - \lambda \left(\int d\mathbf{z} \rho - 1 \right) - \beta \left(\int d\mathbf{z} H \rho - U \right)$$

(the explicit maximization of the functional in (168) is carried in appendix B). Thus we obtain the μ CMEP in the form of constrained maximization of the *volume entropy* of Eq. (8) with the normalization constraint *and* the average energy constraint. This is in contrast with the traditional derivation of the microcanonical ensemble from a maximum entropy principle. According to the commonly accepted derivation, in order

to obtain the microcanonical distribution one should maximize Shannon information under the only constraint on normalization¹

$$(169) \quad \delta \left[- \int_{H=U} d\mathbf{z} \rho \ln \rho - \left(\lambda \int_{H=U} d\mathbf{z} \rho - 1 \right) \right] = 0$$

which leads to the following constant distribution:

$$(170) \quad \rho(\mathbf{z}) = \frac{1}{\Omega(U)}.$$

The assumption underlying this procedure is that the point \mathbf{z} *already* belongs to the surface of energy U . Maximizing the Shannon information over such surface then leads to a *flat* distribution which says that all the states on the surface are equiprobable. The maximum entropy principle in (169) leads to the principle of *a priori equi-probability*. With some abuse of notation this is often referred to as the microcanonical ensemble. The notation is abused because strictly speaking the microcanonical ensemble is defined on the whole phase space and it constrains the system state to lie on a certain surface of constant energy as in Eq. (12). Instead in using Eq. (169) one assumes the states to be already belonging to some constant energy surface. The microcanonical distribution is assumed in Eq. (169) rather than derived from it. Our μ cMEP, instead derives the condition of constant energy from the maximization of the proper microcanonical entropy, that is, the volume entropy of Eq. (8).

¹Without loss of generality we now switch to the discrete case, which is more common in literature

CHAPTER 9

GENERALIZED DUAL ENSEMBLES

Let us now move to generalize Tsallis escort ensemble and define a new class of ensembles that satisfy the requirements of the heat theorem.

9.1. Definition of Dual Orthodes

So far we have seen how the Tsallis escort ensembles of index $q \leq 1$ are orthodes with a special *duality* property. We have also mentioned that the microcanonical and canonical ensembles are two special instances of Tsallis escort ensembles. Now we shall see that the property of orthodicity can be proved for a general class of ensembles that share with the Tsallis escort ensemble the fact that both U and β appear explicitly in their expression. They can be considered as parameterized by either U, V or β, V , depending on which parameter is kept fixed and which one is adjusted in such a way that $U = \langle H \rangle$. We shall call these ensembles *dual ensembles*.

The generalization is purely formal, in the sense that we shall assume that all the integrals and derivatives written exist. Further we shall assume that, for given U, V (or β, V), the equation $U = \langle H \rangle$ admits a solution $\beta(U, V)$ (or $U(\beta, V)$) which is of class C^1 . These are conditions that must be checked, a posteriori, on a case by case basis, depending on the explicit form of the Hamiltonian and of the distribution. Thus, let us consider a generic ensemble of the form:

$$(171) \quad \rho(\mathbf{z}; U, V) = \frac{f[\beta(U - H(\mathbf{z}; V))]}{G(U, V)},$$

where G and β are assumed to be differentiable functions of U, V . This means that we are adopting the (U, V) parametrization, alternatively we could have adopted the β, V parametrization if G and U are assumed to be differentiable functions of (β, V) .

In the (U, V) parametrization The values of G and β are fixed by the two constraints

$$(172) \quad G(U, V) = \int d\mathbf{z} f[\beta(U - H(\mathbf{z}; V))]$$

and

$$(173) \quad U = \int d\mathbf{z} H(\mathbf{z}; V) \rho(\mathbf{z}; U, V).$$

Throughout this section the symbol $\langle \cdot \rangle_\rho$ denotes average over the dual distribution (171). Consider now a differentiable function $F(x)$ such that $F(x) > 0$, $F'(x) = f(x)$, and define the following function:

$$(174) \quad S_\rho(U, V) \doteq \ln \mathcal{G}(U, V),$$

where

$$(175) \quad \mathcal{G}(U, V) \doteq \int d\mathbf{z}, F[\beta(U - H(\mathbf{z}; V))]$$

and all the integrals are extended to the definition domain \mathcal{D} of the distribution ρ . We shall also assume that a cut-off condition exists such that F is null on the boundary $\partial\mathcal{D}$ of the domain \mathcal{D} :

$$(176) \quad F[\beta(U - H(\mathbf{z}; V))]_{\mathbf{z} \in \partial\mathcal{D}} = 0,$$

which is of central importance for the theory that we are developing. Let us define the macroscopic state:

$$(177) \quad \begin{aligned} U_\rho &\doteq \langle H \rangle_\rho = U \\ T_\rho &\doteq \frac{2\langle K \rangle_\rho}{n} \\ V_\rho &\doteq V \\ P_\rho &\doteq \left\langle -\frac{\partial H}{\partial V} \right\rangle_\rho. \end{aligned}$$

Before proving that the ensembles of the form (171) are orthodes let us state the following *generalized equipartition theorem*:

Theorem 5. In the U, V parametrization the average of $p_i \frac{\partial H}{\partial p_i}$ is¹:

$$(178) \quad \left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_\rho = \frac{1}{\beta(U, V)} \frac{\mathcal{G}(U, V)}{G(U, V)}.$$

In the β, V parametrization the average of $p_i \frac{\partial H}{\partial p_i}$ is:

$$(179) \quad \left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_\rho = \frac{1}{\beta} \frac{\mathcal{G}(\beta, V)}{G(\beta, V)}$$

The proof is provided in Appendix C.1. It involves writing the integral expression of the quantity $\left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_\rho$, integrating by parts over p_i and using the cut-off condition (176). The proof structure is the same as the structure of the proof of Tsallis equipartition theorem of Ref. [65].

Comparing with the macroscopic state definition (177), the generalized equipartition theorem can be reexpressed as the following compact formula :

$$(180) \quad T_\rho = \frac{1}{\beta} \frac{\mathcal{G}}{G}$$

where it is intended that Eq. (178) is used in the U, V parametrization and Eq. (179) in the β, V one. Equation (180) tells that for generic dual statistics the quantity β *might not* coincide with the inverse physical temperature. Let us now evaluate the partial derivatives of the entropy function in Eq. (174):

$$(181) \quad \begin{aligned} \frac{\partial S_\rho}{\partial U_\rho} &= \frac{1}{\mathcal{G}} \frac{\partial}{\partial U} \int d\mathbf{z} F(\beta(U - H)) \\ &= \frac{1}{\mathcal{G}} \int d\mathbf{z} f(\beta(U - H)) \left[\frac{\partial \beta}{\partial U} (U - H) + \beta \right] \\ &= -\frac{G}{\mathcal{G}} \frac{\partial \beta}{\partial U} \langle H - U \rangle_\rho + \frac{G}{\mathcal{G}} \beta = \frac{1}{T_\rho}. \end{aligned}$$

In order to obtain the last equality we used the first definition in (177) and Eq. (180).

$$\begin{aligned} \frac{\partial S_\rho}{\partial V_\rho} &= \frac{1}{\mathcal{G}} \frac{\partial}{\partial V} \int d\mathbf{z} F(\beta(U - H)) \\ &= \frac{1}{\mathcal{G}} \int d\mathbf{z} f(\beta(U - H)) \left[\frac{\partial \beta}{\partial V} (U - H) - \beta \frac{\partial H}{\partial V} \right] \end{aligned}$$

¹The label i counts the degrees of freedom. Repeated indices are not summed.

$$\begin{aligned}
&= \frac{G}{\mathcal{G}} \frac{\partial \beta}{\partial V} \langle U - H \rangle_\rho - \frac{G}{\mathcal{G}} \beta \left\langle \frac{\partial H}{\partial V} \right\rangle_\rho \\
(182) \quad &= \frac{P_\rho}{T_\rho}.
\end{aligned}$$

In order to obtain the last equality we used the first and fourth definitions in (177) and Eq. (180). From Eqs. (181) and (182) we get:

$$(183) \quad dS_\rho = \frac{dU_\rho + P_\rho dV_\rho}{T_\rho}$$

Therefore the differential $\frac{dU_\rho + P_\rho dV_\rho}{T_\rho}$ is exact and the entropy is given by Eq. (174). This implies that the ensembles of the form (171) are orthodes, namely they provide good mechanical models of thermodynamics. In Appendix C.2 we provide a proof that the heat theorem is also satisfied if the alternative β, V parametrization is adopted. The proof is essentially the same as for the Tsallis case (see Sec. 7.5). Thus, we have found that the class of orthodes, whose known representatives have been for more than one century only a few (canonical, microcanonical, grand-canonical and pressure ensemble [2]) is indeed quite vast and can include other statistics.

9.1.1. Recovery of known cases

9.1.1.1. *Canonical.* The canonical ensemble is a very special case of dual orthode where the parameter U does not appear explicitly in the expression of the distribution. This case is obtained with the choice:

$$f(x) = F(x) = e^x.$$

In this case we get

$$(184) \quad \rho(\mathbf{z}; \beta, V) = \frac{e^{\beta(U-H)}}{\int d\mathbf{z} e^{\beta(U-H)}} = \frac{e^{-\beta H}}{\int d\mathbf{z} e^{-\beta H}}.$$

The average energy U cancels in the last term of (184). In this sense we refer to the canonical ensemble as a case of “hidden dual orthode”. The canonical entropy is recovered by taking the natural logarithm of $\mathcal{G} = \int d\mathbf{z} e^{\beta(U-H)}$:

$$S(\beta, V) = \beta U + \ln \int d\mathbf{z} e^{-\beta H}$$

Note also that, from Eqs. (172) and (175), $G = \mathcal{G}$ in this specific case, therefore the generalized equipartition theorem (180) gives $T_\rho = \frac{1}{\beta}$. In this way the canonical equipartition theorem (106) is recovered too.

9.1.1.2. *Microcanonical.* The microcanonical ensemble is recovered with the choice

$$\begin{aligned} f(x) &= \delta(x) \\ F(x) &= \theta(x) \end{aligned}$$

From the properties of the Dirac delta the distribution in Eq. (171) is:

$$(185) \quad \rho(\mathbf{z}; U, V) = \frac{\delta(\beta(U - H))}{\int d\mathbf{z} \delta(\beta(U - H))} = \frac{\delta(U - H)}{\int d\mathbf{z} \delta(U - H)}.$$

As with the canonical case (184), the last term in (185), does not depend explicitly on β , hence the microcanonical case is also a case of “hidden dual statistics”. The microcanonical equipartition theorem (15) is also recovered. From (180) one gets:

$$(186) \quad T_\rho = \frac{1}{\beta} \frac{\mathcal{G}}{G} = \frac{\int d\mathbf{z} \theta(U - H)}{\int d\mathbf{z} \delta(U - H)} = \frac{\Phi}{\Omega},$$

where we have used the relations $\theta(ax) = \theta(x)$ (for $a > 0$) and $\delta(ax) = a^{-1}\delta(x)$.

9.1.1.3. *Tsallis (escort).* The Tsallis escort case is recovered with the choice

$$(187) \quad \begin{aligned} f(x) &= [1 + (1 - q)x]^{\frac{q}{1-q}} \\ F(x) &= [1 + (1 - q)x]^{\frac{1}{1-q}} \end{aligned}$$

(188)

In this case one finds $\mathcal{G} = \mathcal{N}_q$ and $G = N_q$ in Eqs. (172) and (175), so from Eq. (180) the Tsallis equipartition theorem (112), is recovered:

$$(189) \quad T_\rho = \frac{1}{\beta} \frac{\mathcal{G}}{G} = \frac{1}{\beta} \frac{\mathcal{N}_q}{N_q}.$$

Canonical and microcanonical cases are both included in the family of Tsallis distributions as special cases corresponding to the values $q = 1$ and $q = -\infty$, as from Chap. 8.

For future reference, let us remember that we have $\mathcal{N}_q = N_q$ for finite q 's. Therefore Eq. (189) becomes:

$$(190) \quad T_q = \frac{1}{\beta} \quad |q| < \infty.$$

9.2. Derivation of the Ideal Gas Thermodynamics

In the ideal gas case the potential energy $\varphi(\mathbf{z}; V)$ is a box potential which constrains the coordinates to lie in an interval of measure $L = V^{1/3}$, where we assume for simplicity a cubic box of volume V . The box potential reduces the integration over the configuration space to a domain of measure $V^{n/3}$, where $n = 3N$ is the total number of degrees of freedom and N is the number of particles. The Hamiltonian is purely kinetic: $H = \sum_{i=1}^{3N} \frac{p_i^2}{2m}$. Assuming that the function $\beta(U, V)$ exists, the equation of state is obtained from Eq. (182), namely $\frac{\partial S_\rho}{\partial V} = \frac{P_\rho}{T_\rho}$:

$$(191) \quad \begin{aligned} \frac{\partial S_\rho}{\partial V} &= \frac{1}{\mathcal{G}} \frac{\partial}{\partial V} \int_0^{V^{1/3}} d^n \mathbf{q} \int d^n \mathbf{p} F[\beta(U - H(\mathbf{p}))] \\ &= \frac{1}{\mathcal{G}} \frac{\partial}{\partial V} V^{n/3} \int d^n \mathbf{p} F[\beta(U - H(\mathbf{p}))] \\ &= \frac{n}{3} \frac{1}{V} \end{aligned}$$

from which the standard ideal gas law is easily obtained:

$$(192) \quad P_\rho V_\rho = \frac{n}{3} T_\rho = N T_\rho$$

The fact that the standard ideal gas law is found to hold for any dual statistics should not be surprising since it generalizes a result already found within the nonextensive thermodynamics [68] (see also the recent work of [78]). Let us now focus on the form of the function $\beta(U, V)$ in the ideal gas case. Using the standard change of variable [1] $\frac{\mathbf{p}^2}{2m} = e$, $d^n \mathbf{p} = c_n e^{\frac{n}{2}-1} de$, followed by the change of variable $y = \beta e$, the condition $\langle H \rangle = U$ is expressed as:

$$(193) \quad I(A, n) \doteq \int_0^{\bar{y}(A)} dy y^{\frac{n}{2}-1} (A - y) f(A - y) = 0$$

where $A = \beta U$, and \bar{y} is a cut-off possibly infinite and possibly depending on A . We shall refer to Eq. (193) as the energy constraint equation. From such equation it is easily inferred that, if for a given number of degrees of freedom n , a solution A_n of (193) exists, then the function $\beta(U, V)$ exists and is given by:

$$(194) \quad \beta = \frac{A_n}{U}.$$

For example, within the canonical ensemble one has $\beta = \frac{n}{2U}$. In the ideal gas case $\langle K \rangle = U$, so, from the definition of macroscopic state (177) one has

$$(195) \quad T_\rho = \frac{2U}{n}.$$

Because of orthodicity we have $\frac{\partial S_\rho}{\partial U} = \frac{1}{T_\rho} = \frac{n}{2U}$ for any dual ensemble. This implies that the entropy is:

$$(196) \quad S_\rho(U, V) = \frac{n}{2} \ln U + \frac{n}{3} \ln V + \text{const}$$

Therefore the standard ideal gas thermodynamics is recovered for any dual orthode. This means that the canonical or microcanonical statistics are not the only statistics necessary to obtain the ideal gas thermodynamics. On the contrary, ordinary thermodynamics may follow from nonordinary ensembles that belong to the class of dual orthodes.

As the generalized equipartition theorem (180) suggests, in general the standard relation $T_\rho = \frac{1}{\beta}$ does not hold. For example, from Eqs. (194) and (195), one easily finds the following formula:

$$(197) \quad T_\rho = \frac{1}{\beta} \frac{\frac{n}{2}}{A_n}.$$

By comparison with the Eq. (180), one also deduces that

$$(198) \quad \frac{\mathcal{G}}{G} = \frac{\frac{n}{2}}{A_n}.$$

The previous formula can also be derived directly by considering the explicit expression of \mathcal{G} (we adopt the U, V representation),

$$(199) \quad \mathcal{G}(U, V) = c_n V^{\frac{n}{3}} \int_0^{\bar{U}} de e^{\frac{n}{2}-1} F[A_n(1+e/U)]$$

where \bar{U} is the cut-off energy value. Equation (198) follows after an integration by parts and the definition of U from Eq. (177).

9.3. Examples

9.3.1. Tsallis Escort Ensembles

As an illustration of the theory let us first apply it to the Tsallis escort orthodes of indices $q \leq 1$. As we will see, this is a quite special case that can be worked analytically. The ensembles are (we adopt the U, V representation):

$$(200) \quad \rho_q(\mathbf{z}; U, V) = \frac{\left[1 - \frac{\beta}{\alpha_q}(H(\mathbf{z}; V) - U)\right]^{\alpha_q-1}}{N(U, V)},$$

where for simplicity we have adopted the notation $\alpha_q = \frac{1}{1-q}$ ². The cut-off condition (see Eq. (167)), $H \leq U + \frac{\alpha_q}{\beta}$, is dictated by the fact that the expression within square parentheses has to be non-negative. In this case the energy constraint integrals (193) can be evaluated analytically:

$$(201) \quad I_q(A, n) = -\left(\frac{n}{2} - A\right) \frac{(A + \alpha_q)^{\alpha_q+n/2-1} \Gamma(\alpha_q + 1) \Gamma(\frac{n}{2})}{\alpha_q^{\alpha_q} \Gamma(\frac{n}{2} + \alpha_q + 1)}.$$

The solutions $A_{n,q}$ of the equations $I_q(A, n) = 0$ are $A_{n,q} = \frac{n}{2}$, no matter the value of q . Therefore one finds, from Eq. (194), the relation $T^{[q]} = \frac{1}{\beta}$, which is in agreement with the Tsallis equipartition theorem (see Eq. (190))

Using Eq. (194) one can express the Tsallis escort ensemble, in the ideal gas case, as

$$\rho_q(\mathbf{z}; U, V) = \frac{\left[1 + \frac{(1-q)n}{2} \left(1 - \frac{H(\mathbf{z}; V)}{U}\right)\right]^{\frac{q}{1-q}}}{\int d\mathbf{z} \left[1 - \frac{(1-q)n}{2} \left(1 - \frac{H(\mathbf{z}; V)}{U}\right)\right]^{\frac{q}{1-q}}}$$

²Note that, according to Eq. (142), α_q represents the heat capacity of the bath

where the U, V parametrization has been adopted. Alternatively, adopting the β, V parametrization the Tsallis escort ensemble would be:

$$\rho_q(\mathbf{z}; U, V) = \frac{[1 + (1 - q)(\frac{n}{2} - \beta H(\mathbf{z}; V))]^{\frac{q}{1-q}}}{\int d\mathbf{z} [1 + (1 - q)(\frac{n}{2} - \beta H(\mathbf{z}; V))]^{\frac{q}{1-q}}}.$$

Applying Eqs. (174) and (175) gives the entropy. In the U, V representation, it reads:

$$(202) \quad S^{[q]}(U, V) = \frac{n}{3} \ln V + \ln L_{n,q}(U) + \ln c_n,$$

where

$$(203) \quad L_{n,q}(U) = U^{\frac{n}{2}} \int_0^{1+\frac{2}{n(1-q)}} dx x^{\frac{n}{2}-1} \left[1 + \frac{(1-q)n}{2} (1-x) \right]^{\frac{1}{1-q}}.$$

In agreement with Eq. (196), the dependence of the entropy on U is of the type $\frac{n}{2} \ln U$. The integral $L_{n,q}$ has been obtained by using the cut-off condition $\bar{e} = U \left(1 + \frac{2}{n(1-q)} \right)$, and the change of variable $x = e/U$, where e denotes energy.

9.3.2. Gaussian Ensemble

Since the fundamental work of Khinchin [1] based on the application of the central limit theorem, it is known that the distribution law for a large component of a large Hamiltonian isolated systems of total energy a is well approximated by the following Gaussian distribution:

$$(204) \quad \rho = \frac{e^{\beta(a-H)} \exp \left[-\frac{(A_1-H)^2}{2B_2} \right]}{\text{normalization}},$$

where the quantities A_1 and B_2 being defined in terms of the Laplace transforms $Z_i(\beta)$ of the structure functions $\Omega_i(x)$ of the system ($i = 1$) and the heat bath ($i = 2$):

$$A_1 = -\frac{d \ln Z_1}{d\beta},$$

$$B_2 = \frac{d^2 \ln Z_2}{d\beta^2}.$$

According to Khinchin the quantity A_1 is a good approximation to the average energy U of the system ($\frac{U-A_1}{U} = O(\frac{1}{N_1})$), where N_1 is the number of particles in the system).

The width of the distribution, B_2 , can be expressed, in the case of an ideal gas bath (see Chapter 5, Section 22 of Ref. [1]) as:

$$(205) \quad B_2 = \frac{3N_2}{2\beta^2}.$$

Here N_2 denotes the number of particles in the bath. Hence the ensemble in Eq. (204) can be re-expressed in the form of a dual ensemble:

$$(206) \quad \rho_\sigma(\mathbf{z}; U, V) = e^{\beta(U-H)} \frac{\exp\left[-\frac{(\beta(U-H))^2}{2\sigma}\right]}{G},$$

where $\sigma = \frac{3N_2}{2}$ is indeed the heat capacity of the heat bath, namely σ plays the same role here as the parameter $\alpha_q = \frac{1}{1-q}$ in the Tsallis escort ensembles (see Eq. (142)). The distribution goes rapidly to zero, correspondingly the cut-off boundary is at infinity (in other words the domain of definition of (206) is \mathbb{R}^{6N}). The Gaussian ensemble is reproduced with the choice:

$$(207) \quad f(x) = e^x \exp\left[-\frac{x^2}{2\sigma}\right]$$

$$(208) \quad F(x) = \sqrt{\frac{\pi\sigma}{2}} e^{\frac{\sigma}{2}} \left(1 + \operatorname{erf}\left[\frac{x-\sigma}{\sqrt{2\sigma}}\right]\right).$$

The energy constraint condition (193) in this ensemble is

$$(209) \quad I_\sigma(A) = \int_0^\infty dy y^{\frac{n}{2}-1} (A-y) e^{A-y-\frac{(A-y)^2}{2\sigma}} = 0.$$

The solutions $A_{n,\sigma}$ of this equation have been evaluated numerically for $n = 10, 20, \dots, 100$, $\sigma = 2, 10, 20$, and shown in Figure 9.1. For all the values of σ investigated, $A_{n,\sigma} \simeq \frac{n}{2}$. The large n behavior of $A_{n,\sigma}$ is investigated in Appendix C.3. The entropy, in the U representation is:

$$(210) \quad S(U, V) = \frac{n}{3} \ln V + \ln Y_{n,\sigma}(U) + \ln c_n \sqrt{\frac{\pi\sigma}{2}} e^{\frac{\sigma}{2}}$$

where

$$(211) \quad Y_{n,\sigma}(U) = U^{\frac{n}{2}} \int_0^\infty dx x^{\frac{n}{2}-1} \left(1 + \operatorname{erf}\left[\frac{A_{n,\sigma}(1-x) - \sigma}{\sqrt{2\sigma}}\right]\right) -$$

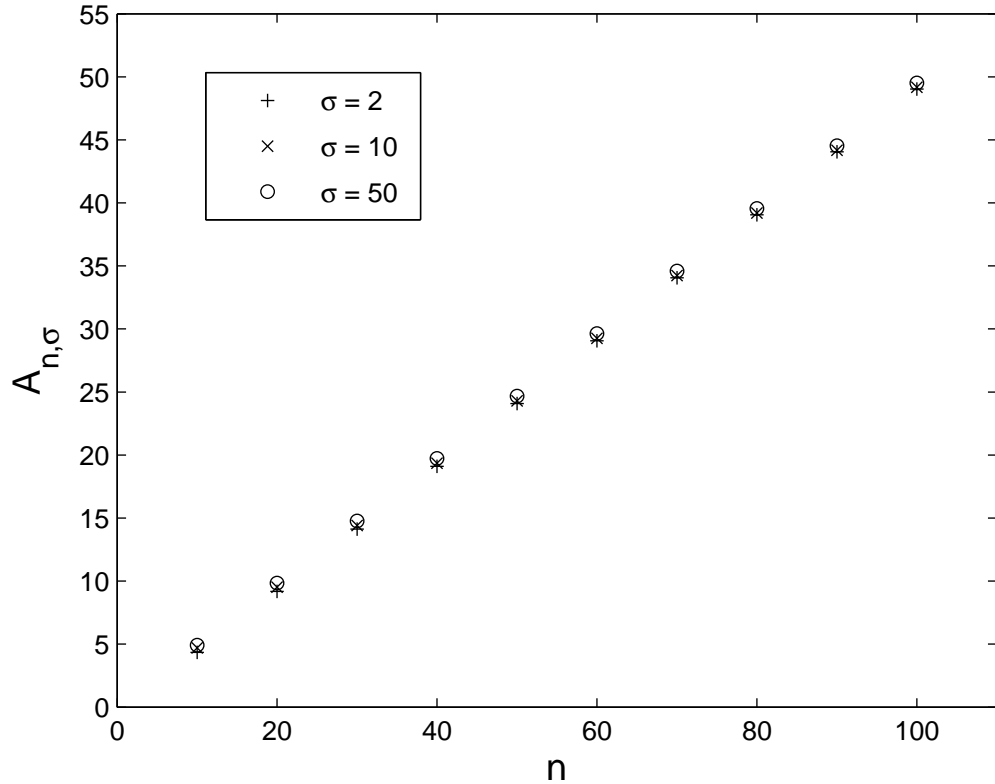


FIGURE 9.1. Solution of Eq. (209), for different values of σ and n

In agreement with Eq. (196), the dependence of the entropy on U is of the type $\frac{n}{2} \ln U$. The integral $Y_{n,\sigma}$ has been obtained by using the change of variable $x = e/U$, where e denotes energy.

The Gaussian ensemble of Eq. (206) interpolates between canonical and micro-canonical ensembles as does the Tsallis escort ensemble [79]. In fact, on one hand $f(x) \rightarrow e^x$ as $\sigma \rightarrow +\infty$, and, on the other $f(x) \rightarrow \sqrt{2\pi\sigma}\delta(x)$ as $\sigma \rightarrow 0$ (the vanishing term $\sqrt{2\pi\sigma}$ is not a problem because it will be cancelled with the normalization).

Furthermore, based on the fact that the Gaussian ensemble of index σ well describes the statistics of a large component of a large isolated system, we deduce that it must well approximate the Tsallis statistics of index $\alpha_q = \sigma$ in the case $n, \alpha_q \gg 1$. This fact is illustrated in Figure 9.2, where we have plotted the unnormalized Tsallis

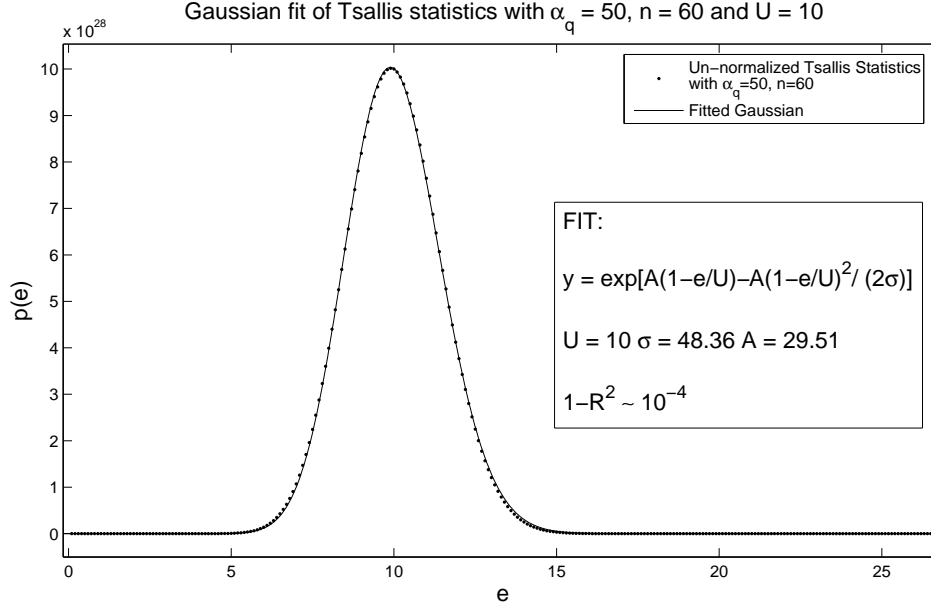


FIGURE 9.2. Gaussian fit of Tsallis statistics for $\alpha_q = 50$, $n = 60$, $U = 10$. The dual Gaussian ensemble approximates Tsallis escort ensembles for large values of n and α_q

escort distribution $\left(1 + \frac{n}{2\alpha_q}(1 - \frac{e}{U})\right)^{\alpha_q-1} e^{\frac{n}{2}-1}$, for $\alpha_q = 50$, $n = 60$, $U = 10$, and fitted it to the unnormalized gaussian statistics $\exp[A(1 - e/U) - \frac{(A(1-e/U))^2}{2\sigma}]e^{\frac{n}{2}-1}$ with U, σ, A as free parameters. The fit is very good ($1 - R^2 = O(10^{-4})$) and the fitting parameter matched quite well the expected values: $\sigma = 48.36 \simeq \alpha_q = 50$, $U = 10$, $A = 29.51 \simeq A_{n=60, \sigma=50} = 29.63$.

9.3.3. Fermi-like Ensemble

The theory developed so far allows us to construct mechanical models of thermodynamics with the most diverse types of distributions. For example one may ask whether it would be possible to have an ensemble with a Fermi-like distribution. This is possible for the ideal gas case. We will describe this ensemble as an illustration of the theory, without discussing whether it really applies to some many-particle physical system. The Fermi-like statistics uses the choice:

$$(212) \quad F(x) = \ln(e^x + 1)$$

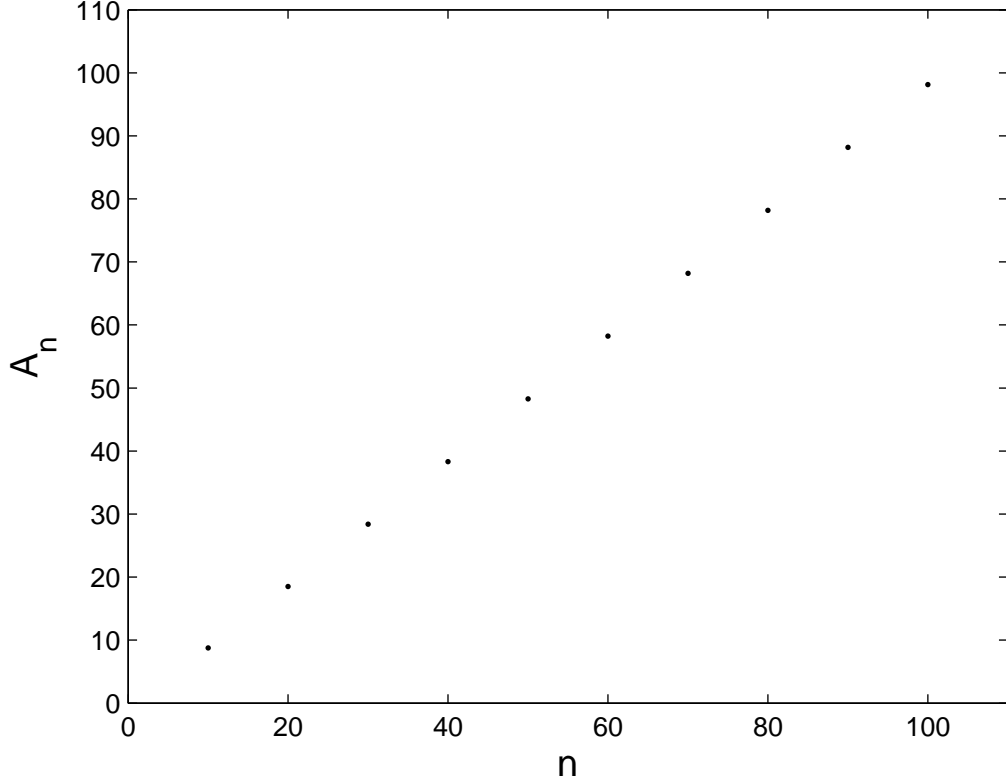


FIGURE 9.3. Solution of Eq. (214) for different values of n .

$$(213) \quad f(x) = \frac{1}{e^{-x} + 1}$$

with no finite cut-off. The solutions A_n of the energy constraint equation (193):

$$(214) \quad I(A, n) = \int_0^\infty dy \frac{y^{\frac{n}{2}-1}(A-y)}{e^{y-A} + 1} = 0$$

have been evaluated numerically and shown in Fig. 9.3 for the values $n = 10, 20, \dots, 100$.

From the figure we see that, $A_n \simeq n$. This fact is analyzed in more details in Appendix C.4

Adopting the (U, V) parametrization, the ensemble is:

$$(215) \quad \rho(\mathbf{z}; U, V) = \frac{\left[e^{-A_n(1-\frac{H(\mathbf{z}; V)}{U})} + 1 \right]^{-1}}{\int d\mathbf{z} \left[e^{-A_n(1-\frac{H(\mathbf{z}; V)}{U})} + 1 \right]^{-1}}$$

or, alternatively in the (β, V) parametrization:

$$(216) \quad \rho(\mathbf{z}; \beta, V) = \frac{[e^{-A_n + \beta H(\mathbf{z}; V)} + 1]^{-1}}{\int d\mathbf{z} [e^{-A_n + \beta H(\mathbf{z}; V)} + 1]^{-1}}.$$

Adopting the (U, V) parametrization, the entropy in Eq. (174) is given by

$$(217) \quad S(U, V) = \ln c_n + \frac{n}{3} \ln V + \ln J_n(U),$$

where

$$(218) \quad J_n(U) = U^{\frac{n}{2}} \int_0^\infty dx x^{\frac{n}{2}-1} \cdot \ln (e^{A_n(1-x)} + 1).$$

In agreement with Eq. (196), the dependence of the entropy on U is of the type $\frac{n}{2} \ln U$. The integral J_n has been obtained by using the change of variable $x = e/U$, where e denotes energy.

9.4. Discussion

In this Chapter we addressed some fundamental issues raised recently in statistical mechanics, namely whether a theoretical basis can be provided for nonstandard (i.e., neither canonical nor microcanonical) ensembles, which are often encountered in the most diverse fields of physics. Other authors have addressed the same problem and developed different approaches (see for example [80] and [81]). Here we used Boltzmann's original approach based on the "heat theorem", in order to examine the subject from another point of view. By generalizing the *duality* property observed in Tsallis escort ensembles, we have been able to define a class of *dual statistics*, which includes the Tsallis escort ensembles as particular cases. The generalization scheme is represented in Fig. 9.4. Thanks to the proposed generalization it is possible to provide a theoretical basis for many nonstandard statistics other than Tsallis', such as the gaussian and the Fermi-like statistics. For all such nonstandard orthodes the heat theorem holds, namely the usual thermodynamic relations are recovered. In this dissertation we have also provided a general formula for the entropy associated with any dual orthode. This allowed us to obtain the expression of entropy for the

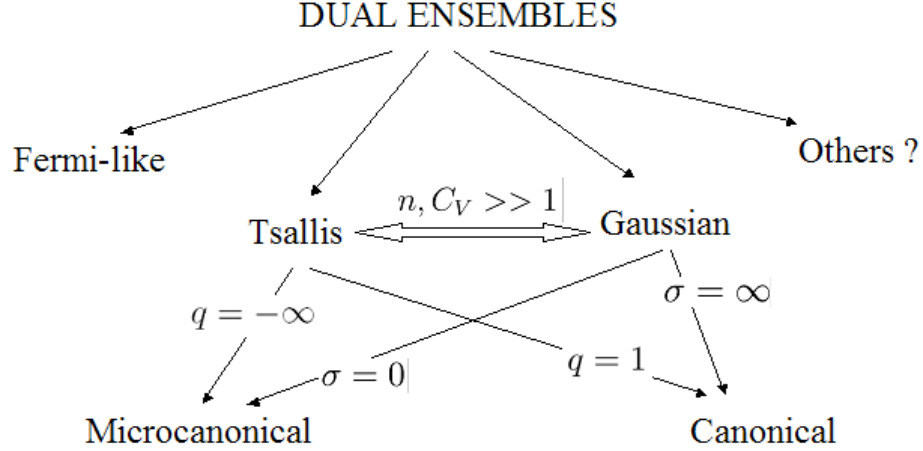


FIGURE 9.4. Generalization scheme. Tsallis escort ensembles, Gaussian ensembles, Fermi-like ensemble, and possibly other kinds of ensembles belong to the class of dual orthodes. Canonical and microcanonical are special instances of Tsallis and Gaussian ensembles. For large systems ($n \gg 1$) Gaussian and Tsallis escort ensemble coincide if $\sigma = \frac{1}{1-q} = C_V \gg 1$.

Gaussian ensemble, which has never been done before. To use the same expression as Gallavotti [2], all dual orthodes provide “mechanical models of thermodynamics”. This result is not trivial since previously the only known orthodes were the microcanonical ensemble, the canonical ensemble and its variants like the grand canonical ensemble and the pressure ensemble [2]. Despite the fact that the class of orthodes is quite vast, the microcanonical and the canonical ensembles still play a special role in statistical mechanics because they are cases of “hidden dual statistics”. They do not rely on the employment of the energy constraint, which constitutes the mechanism through which it is possible to construct nonstandard dual orthodes.

The theory of dual orthodes allows us to generalize the definition of escort distribution. From Eqs. (187)(188), we see that the function $f = F'$ generates the escort Tsallis distribution, whereas F generates the ordinary Tsallis distribution. Thus, within the framework of dual ensembles we can say that the distribution $G^{-1}f(\beta(U - H))$ is the escort version of the distribution $\mathcal{G}^{-1}F(\beta(U - H))$. This definition of generalized ordinary and escort ensembles in terms of a function F and its derivative f is similar to that introduced by Naudts in the definition of generalized exponential families [80]. We notice also a similarity with the work of Hanel and Thurner [81]. In the work of Ref. [81] the generalized ensemble is of the form $\mathcal{E}(\beta(U - H) - \gamma)$. Although the normalization is accounted for in a different way (i.e., through γ), in both cases the dependence upon the Hamiltonian H comes through the same expression $\beta(U - H)$.

One interesting result of the present work is that the classical ideal gas thermodynamics follows from any dual ensemble, thus revealing that this occurrence is not special for the standard statistics. This fact was already noticed for the Tsallis case a few years ago [68], although its orthodicity was not yet clearly recognized. It also suggests that there is a one-to-one correspondence of states obtained in different dual ensembles, i.e. there is an *equivalence* of all dual ensembles. Such equivalence holds no matter the number of degrees of freedom, and, of course, may break as one considers more realistic Hamiltonian models of systems with phase transitions. We also like to stress that all dual orthodes are equivalent also in another sense, namely the sense first investigated by Gibbs [17] for the canonical and microcanonical ensembles. In the thermodynamic limit ($n \rightarrow \infty$) and for the free gas model, the canonical ensemble is so peaked around the average energy value that it is practically undistinguishable from the microcanonical one [18]. The same kind of equivalence occurs for any dual statistics provided that $A_n \rightarrow \infty$ in the thermodynamic limit, which is indeed the case observed in the examples considered in this paper. This behavior is because the distribution is expressed in terms of the quantity $f\left(A_n \frac{U-e}{U}\right)$, which tends to the Dirac

delta centered around U , times an unimportant proportionality factor:

$$(219) \quad f\left(A_n \frac{U - H}{U}\right) \rightarrow \text{const} \times \delta(U - H)$$

Equation (219) follows from the asymptotic formula $h(kx) \rightarrow \frac{\int dy h(y)}{k} \delta(x)$, as $k \rightarrow \infty$.

CHAPTER 10

CONCLUSIONS AND PERSPECTIVES

10.1. Conclusions

Following the line of ideas of Boltzmann, in this dissertation we have taken the heat theorem

$$\frac{\delta Q}{T} = \text{exact differential}$$

and placed it at the very center of an investigation of the microscopic foundations of thermodynamics and of statistical ensembles.

Although such line of ideas of Boltzmann is more than a century old, adopting it proves to be extremely fruitful in addressing two topics which are nowadays still very actual and debated. Those are the problem of reconciling the non time-reversal symmetric second law of thermodynamics with the time-reversal symmetric microscopic law of motion, and the problem of providing a theoretical foundation for ensembles other than the standard microcanonical and canonical ones.

We started by considering the Helmholtz theorem. According to this theorem the thermodynamic relations hold mechanically (without probabilistic assumptions) in the case of one-dimensional *monocyclic* systems. Thanks to a discrete picture of the phase space, Boltzmann was able to apply Helmholtz theorem to multi-dimensional *ergodic* systems, suggesting that the thermodynamic relations we observe in macroscopic systems at equilibrium are a direct consequence of the microscopic laws of dynamics alone. Here we have reviewed Boltzmann's argument and showed that, using the language of the modern ergodic theory, it can be safely re-expressed on a continuum phase space as a generalized Helmholtz theorem (GHT), which has been readily proved. Along the way the agreement between the Helmholtz-Boltzmann theory and that of P. Hertz (based on *adiabatic invariance*) has been revealed. Both

theories, in fact, lead to define the entropy as the logarithm of the phase space volume enclosed by the constant energy hyper-surface (*volume entropy*).

Armed with such a good mechanical analogue of entropy we studied its behavior under non-quasi-static perturbations and found a very interesting result. Independent of the system's number of degrees of freedom, the expectation value of the volume entropy can only increase under the unitary (quantum or classical) time-evolution, provided the initial state is an equilibrium state at positive absolute temperature and the system is non-degenerate. This can be considered as a novel rigorous proof of the second law that is an improvement over previous attempts. In particular it illuminates on the fact that time-reversal asymmetry stems at the level of the equilibrium principle (that is the minus first law of thermodynamics) rather than at the level of the second law. The asymmetry appears because we are looking only at the evolution of a subset of all possible initial equilibrium states.

Based on the results presented in the first part of the dissertation we can say that the macroscopically observed second law actually can be understood on the basis of microscopic dynamics, without the need to invoke the thermodynamic limit. We have also noticed that even phase transition can exist in microscopic systems. Thus it seems that a whole new world of microscopic thermodynamics, completely similar and analogous to the well known macroscopic thermodynamics is amenable to be discovered, studied and observed.

Still following Boltzmann, in the second part of the dissertation we have applied the heat theorem to the study of the foundations of statistical ensembles. We know from Boltzmann that microcanonical and canonical ensembles satisfy the prescriptions imposed by the heat theorem. Here we have seen that a class of ensembles (that is Tsallis escort ensembles) that are at the basis of the recently developed field of non-extensive thermodynamics, also satisfy those prescriptions. This provides such ensembles with a *mechanical foundation* other than the usual information-theoretic one from which they were first derived. Indeed we have seen that the two approaches

(i.e., mechanical and information-theoretic) are mutually consistent with one another, with Rényi entropy acting as a link. An interesting result that we have found is that Tsallis escort ensembles interpolate between the canonical ensemble and the microcanonical ensemble. This holds also for the associated Rényi entropies which correctly interpolate between the canonical entropy and the microcanonical entropy (that is the volume entropy). The physical explanation of this is that Tsallis escort ensembles describe systems in contact with finite heat baths.

We concluded the dissertation by unfolding a general mathematical structure that underlies the microcanonical ensemble, the canonical ensemble and the interpolating Tsallis (escort) ensembles. That is, all these ensembles belong to the class of what we called the *dual orthodes*. All its members satisfy the heat theorem and produce the same ideal gas thermodynamic relations. The mathematical structure of the dual orthodes captures in a unifying picture the fundamental features of canonical and microcanonical ensembles and at the same time extends those to other types of statistics, such as the Tsallis (escort) ones, the gaussian ones and possibly others.

10.2. Perspectives

Ideas, methods and results presented in this work are certainly amenable to further investigation and development. For example it would be very interesting to apply the general theory developed in Chap. 4 to some specific, possibly exactly solvable, models. This could provide interesting information on the actual time dependence of the entropy one should expect from an experimental test of the theory. It is interesting to notice that the *damped harmonic oscillator* which is often considered a prototype model of *dissipative system*, fits quite well in the theory developed in Chap. 4. In fact the damped harmonic oscillator is actually a Hamiltonian system with the Hamiltonian depending explicitly on time:

$$H(x, p, t) = e^{-\gamma t} \frac{p^2}{2m} + e^{\gamma t} \frac{k}{2} x^2$$

Thus the expectation value of its volume entropy increases as it should be for a dissipative system. A specific study might be devoted to the explicit calculation of the function $S(t)$ in Eq. (67) for an initial canonical equilibrium (if the initial state has been realized by thermalization with an infinite bath) or an initial Tsallis escort equilibrium (if it has been realized by thermalization with a finite heat bath).

In Sec. 4.6 we have discussed the role played by the initial equilibrium in the second law. We have found that a law of entropy decrease exists for initial inverted populations. We also have discussed the fact that the latter can only be created artificially by spending a larger amount of entropy than that returned during the time dependent perturbation. Our discussion on the thermalization with finite heat baths of Sec. 8.2 suggests a possible way to create an inverted population at no entropic cost. If we thermalize our system with a finite heat bath with constant heat capacity $C_V < 1$ then the resulting equilibrium distribution in Eq. (143) of our system will be *increasing* !!! Such inverted population would have been created via thermalization, in a spontaneous way, rather than via a forceful entropically expensive pumping like in laser systems. This is just an hypothesis whose actual feasibility certainly needs further investigation, both on the theoretical and experimental level.

Our statement according to which the entropy of a negative entropy system decreases is in evident disagreement with the standard understanding of negative temperature thermodynamics according to which instead it increases [38]. Further studies aimed at resolving this disagreement might turn out to be fruitful in improving the general theory of thermodynamics of negative temperature systems as it is currently understood.

In Chap. 5 we have addressed microcanonical phase transitions in small isolated systems and observed that when the ergodicity breaks down the microcanonical entropy becomes singular. Considering that perfect thermal isolation is not practically possible, and that one might be in general interested in the thermodynamic behavior

of a *small system* in contact with a *small bath*, it is meaningful to study the thermodynamics of such system in the Tsallis (escort) ensemble.¹ It is expected that the degree of non-analyticity of the associated Rényi entropy decreases with increasing heat capacity of the bath and becomes zero for infinite bath, as required by Ruelle's Theorem [58]. This issue is certainly worth exploring.

In Chap. 9 we have introduced a quite wide class of generalized ensembles: the dual orthodoxes. Their theoretical justification is based on the fact that they comply with the requirements posed by the heat theorem. Some important questions are still to be addressed. Gaussian, Tsallis, canonical and microcanonical ensembles are known to describe certain physical situations of systems in thermal contact with heat baths with constant heat capacity. But we know that constant specific heats actually do not exist, and thus it is worth asking ourselves what would be the distribution function of a system in contact with a bath with a certain temperature dependent heat capacity $C_V(T)$. Certainly this is an interesting question in its own right. It would also be interesting to verify if such a distribution belongs to the class of dual ensembles.

¹This idea has been recently expressed also in Ref. [59].

APPENDIX A

UNITARITY AND LINEARITY OF THE OPERATOR \mathcal{U}

Let us prove Eq. (84). According to Liouville Theorem the Hamiltonian flow is incompressible, i.e., it is volume preserving. This means that the phase space measure

$$(220) \quad \frac{d\mathbf{q}d\mathbf{p}}{h^{3N}}$$

is preserved during the time evolution. Since $\frac{d\mathbf{q}d\mathbf{p}}{h^{3N}} = \Omega(E)dE = d\Phi$ (Eqs. (10) and (14)), then the measure $d\Phi$ is also preserved. This implies that the time evolution operator $\mathcal{U}(t_i, t_f)$ is *unitary*, that is norm preserving:

$$(221) \quad \int d\Phi \mathcal{U}(t_i, t_f) f(\Phi) = \int d\Phi f(\Phi).$$

Further it implies that any distribution that is constant throughout the whole phase space does not change in time, that is:

$$(222) \quad \mathcal{U}(t_i, t_f) 1 = 1.$$

The operator $\mathcal{U}(t_i, t_f)$ is evidently *linear*, thus:

$$\int d\Theta A(\Theta, \Phi) = \int d\Theta \mathcal{U}(t_i, t_f) \delta(\Theta - \Phi) = \mathcal{U}(t_i, t_f) \int d\Theta \delta(\Theta - \Phi) = \mathcal{U}(t_i, t_f) 1 = 1$$

where we have used, in order, Eq. (80), linearity, the relation $\int dx \delta(x) = 1$ and Eq. (222).

APPENDIX B

CONSTRAINED MAXIMIZATION OF VOLUME ENTROPY

Let us first consider the discrete case. Maximization of the discrete counterpart of Eq. (168) reads:

$$(223) \quad \frac{\partial}{\partial p_n} \left[\ln \sum_{\{n|E_n \leq U\}} 1 - \lambda \sum_n p_n - \beta \sum_n E_n p_n \right] = 0$$

hence we get:

$$(224) \quad \lambda + \beta E_n = 0.$$

Multiplying by p_n and summing over n , we get (using the constraints $\sum p_n = 1$, $\sum p_n E_n = U$):

$$(225) \quad \lambda + \beta U = 0.$$

This implies that $E_n = U$, therefore any state that does not satisfy such relation will have zero probability. Since there are g_U states on the level of energy U , using the principle of a priori equi-probability, the output of the maximization procedure will be:

$$(226) \quad p_n = \frac{\delta_{E_n, U}}{g_U}$$

where $\delta_{E_n, U} = 1$ if $E_n = U$ and $\delta_{E_n, U} = 0$ otherwise. That is, p_n is the discrete micro-canonical ensemble. The passage to the continuum is performed via the substitutions:

$$(227) \quad \sum \rightarrow \int d\mathbf{z}$$

$$(228) \quad p_n \rightarrow \rho(\mathbf{z})$$

$$(229) \quad g_n \rightarrow \Omega(e)$$

Using the substitutions (227), (228) and (229) the continuum μ cMEP becomes:

$$(230) \quad \delta \left[\ln \Phi(U) - \lambda \int d\mathbf{z} \rho(\mathbf{z}) - \beta \int d\mathbf{z} \rho(\mathbf{z}) H(\mathbf{z}) \right] = 0$$

where:

$$\Phi(U) \doteq \int_{H \leq U} d\mathbf{z} = \int d\mathbf{z} \theta(U - H(\mathbf{z})).$$

Eq. (230) then leads to the microcanonical ensemble

$$\rho(\mathbf{z}) = \frac{\delta(U - H(\mathbf{z}))}{\Omega(U)}.$$

APPENDIX C

SOME CALCULATIONS REGARDING DUAL ENSEMBLES

C.1. Proof of the Generalized Equipartition Theorem of Eq. (180)

Let us calculate explicitly the average value of $p_i \frac{\partial H}{\partial p_i}$.¹ Regardless of the parametrization one has:

$$\begin{aligned}
 \left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_\rho &= \frac{1}{G} \int d\mathbf{z} p_i \frac{\partial H}{\partial p_i} f[\beta(U - H)] \\
 &= -\frac{1}{\beta G} \int d\mathbf{z} p_i \frac{\partial}{\partial p_i} F[\beta(U - H)] \\
 &= -\frac{1}{\beta G} \left[[p_i F[\beta(U - H(\mathbf{z}; V))]]_{\mathbf{z} \in \partial \mathcal{D}} - \int d\mathbf{z} F[\beta(U - H)] \right].
 \end{aligned}
 \tag{231}$$

Where we integrated by parts over p_i to obtain the third line. From Eqs. (175) and (176) it follows:

$$\left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle_\rho = \frac{1}{\beta} \frac{\mathcal{G}}{G}.
 \tag{232}$$

According to the parametrization adopted this would be either Eq. (178) or (179).

C.2. Proof of Orthodicity in the β, V parametrization

In the β, V parametrization U is a function of (β, V) . Therefore F is a function of $(\mathbf{z}; U, V)$, and \mathcal{G} (Eq. 175) and S_ρ (Eq. 174) are functions of (β, V) . Let us calculate the partial derivative of S with respect to β :

$$\begin{aligned}
 \frac{\partial S_\rho}{\partial \beta} &= \frac{1}{\mathcal{G}} \frac{\partial}{\partial \beta} \int d\mathbf{z} F(\beta(U - H)) \\
 &= \frac{1}{\mathcal{G}} \int d\mathbf{z} f(\beta(U - H)) \left[(U - H) + \beta \frac{\partial U}{\partial \beta} \right] \\
 &= -\frac{G}{\mathcal{G}} \langle H - U \rangle_\rho + \frac{G}{\mathcal{G}} \beta \frac{\partial U}{\partial \beta} = \frac{1}{T_\rho} \frac{\partial U_\rho}{\partial \beta},
 \end{aligned}
 \tag{233}$$

where we used the equation $U_\rho = \langle H \rangle_\rho$ from the state definition (177) and the generalized equipartition theorem of Eqs. (180) and (232). Now, let us calculate the

¹The label i counts the degrees of freedom. Repeated indices are not summed.

partial derivative of S with respect to V :

$$\begin{aligned}
\frac{\partial S_\rho}{\partial V_\rho} &= \frac{1}{\mathcal{G}} \frac{\partial}{\partial V} \int d\mathbf{z} F(\beta(U - H)) \\
&= \frac{1}{\mathcal{G}} \int d\mathbf{z} f(\beta(U - H)) \left[\beta \frac{\partial U}{\partial V} - \beta \frac{\partial H}{\partial V} \right] \\
&= \beta \frac{G}{\mathcal{G}} \frac{\partial U}{\partial V} - \frac{G}{\mathcal{G}} \beta \left\langle \frac{\partial H}{\partial V} \right\rangle_\rho \\
(234) \qquad &= \frac{1}{T_\rho} \frac{\partial U_\rho}{\partial V_\rho} + \frac{P_\rho}{T_\rho},
\end{aligned}$$

where we used the expression for P_ρ from Eq. (177) and the generalized equipartition theorem of Eq. (180) and (232). Combining these together we obtain the heat theorem:

$$(235) \qquad dS = \frac{\frac{\partial U_\rho}{\partial \beta} d\beta + \frac{\partial U_\rho}{\partial V_\rho} dV_\rho + P_\rho dV_\rho}{T_\rho} = \frac{dU_\rho + P_\rho dV_\rho}{T_\rho}.$$

C.3. Large n behavior of the Coefficients $A_{n,\sigma}$

Equation (209) can be recast in the following form:

$$(236) \qquad \int_0^\infty dy g_n(y) e^{-\frac{(A-y)^2}{2\sigma}} = A \int_0^\infty dy g_n(y) y^{-1} e^{-\frac{(A-y)^2}{2\sigma}}$$

where

$$(237) \qquad g_n(y) \doteq \frac{y^{\frac{n}{2}} e^{-y}}{\int_0^\infty y^{\frac{n}{2}} e^{-y}} \theta(y)$$

and θ is Heaviside's step function. The Fourier transform of Eq. (237) is:

$$(238) \qquad \hat{g}_n(k) = (1 - ik)^{-\frac{n}{2}-1}$$

Using the formula $\lim_{N \rightarrow \infty} \left(1 + \frac{x}{N}\right)^N = e^x$, we have:

$$(239) \qquad \lim_{n \rightarrow \infty} \hat{g}_n(k) = e^{ik\left(\frac{n}{2}+1\right)}$$

therefore the inverse Fourier transform gives

$$(240) \qquad \lim_{n \rightarrow \infty} g_n(y) = \delta\left(y - \left(\frac{n}{2} + 1\right)\right).$$

Using this result, Eq. (236) becomes for very large n :

$$(241) \quad A \simeq \frac{n}{2},$$

which does not depend on σ .

C.4. Large n behavior of the Coefficients A_n for the Fermi-like Case

Figure 9.3 suggests that, in the limit $n \rightarrow \infty$, $A_n \simeq n$. In this appendix section we provide a simple consistency argument to support the claim that $A_n \simeq n$. Let us assume that for very large n , $A_n \simeq n$. Then we should have from Eq. (214):

$$(242) \quad \int_0^\infty dy \frac{y^{\frac{n}{2}}}{e^{y-n} + 1} \simeq n \int_0^\infty dy \frac{y^{\frac{n}{2}-1}}{e^{y-n} + 1}.$$

Equating the first derivative of the integrand on the left hand side to zero, gives:

$$y = \frac{n}{2}(1 + e^{n-y}),$$

which is satisfied for $y = n$. This means that an extremum (a maximum as we will see) is attained for $y = n$. The value taken by the integrand at the maximum is $\frac{n^{\frac{n}{2}}}{2}$ which increases very quickly. The second derivative, calculated at $y = n$, is:

$$(243) \quad \frac{1}{4} \left[\left(\frac{n}{2} - 1 \right) n^{-\frac{n}{2}+1} - n^{\frac{n}{2}} \right]$$

which tends to $-\infty$ very quickly. This indicates that the integrand becomes very sharply peaked around $y = n$ as n increases. Therefore, as an approximation, we can replace y with n , and see that $y^{\frac{n}{2}} \simeq ny^{\frac{n}{2}-1}$, thus getting Eq. (242).

REFERENCES

- [1] A.I. Khinchin. *Mathematical foundations of statistical mechanics*. Dover, New York, 1949.
- [2] G. Gallavotti. *Statistical mechanics. A short treatise*. Springer Verlag, Berlin, 1995.
- [3] M. Campisi. On the mechanical foundations of thermodynamics: The generalized Helmholtz theorem. *Studies in History and Philosophy of Modern Physics*, 36:275–290, 2005.
- [4] M. Campisi. Statistical mechanical proof of the second law of thermodynamics based on volume entropy. *Studies in History and Philosophy of Modern Physics*, in press.
- [5] M. Campisi. Microcanonical phase transitions in small systems. *arXiv:0709.1082v1*.
- [6] M. Campisi and Gokhan B. Bagci. Tsallis ensemble as an exact orthode. *Physics Letters A*, 362(1):11–15, 2007.
- [7] M. Campisi. On the limiting cases of nonextensive thermostatics. *Physics Letters A*, 366(4-5):335–338, 2007.
- [8] M. Campisi. Thermodynamics with generalized ensembles: The class of dual orthodes. *Physica A*, 385:501–517, 2007.
- [9] M. Campisi and D. H. Kobe. Maximum entropy principle for the microcanonical ensemble. *arXiv:0709.1071v1*.
- [10] R. T. Tolman. *The principles of statistical mechanics*. Oxford University Press, London, 1938.

- [11] R. C. Bishop. Nonequilibrium statistical mechanics Brussels-Austin style. *Studies In History and Philosophy of Modern Physics*, 35(1):1–30, 2004.
- [12] L. Boltzmann. Über die Eigenschaften monocyclischer und anderer damit verwandter Systeme. *Crelle's Journal*, 98:68–94, 1884. Reprinted in Hasenöhl (ed.), *Wissenschaftliche Abhandlungen*, vol. 3, pp. 122-152. New York: Chelsea.
- [13] H. Helmholtz. Principien der Statik monocyclischer Systeme. In *Wissenschaftliche Abhandlungen*, vol III, p.142-162 and p.179-202; Studien zur Statik monocyclischer Systeme, in *Wissenschaftliche Abhandlungen*, vol III, p.163-172 and p.173-178. Leipzig, 1895.
- [14] J. Uffink. Boltzmann's work in statistical physics. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Winter 2004 Edition)* The Metaphysics Research Lab, Stanford University, Stanford, 2004. <http://plato.stanford.edu/entries/statphys-Boltzmann/>.
- [15] J. Uffink. Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Modern Physics*, 32(3):305–394, 2001.
- [16] H. R. Brown and J. Uffink. The origins of time-asymmetry in thermodynamics: The minus first law. *Studies In History and Philosophy of Modern Physics*, 32(4):525–538, 2001.
- [17] J.W. Gibbs. *Elementary principles in statistical mechanics*. Yale University Press, Yale, 1902. Reprinted by Dover, New York, 1960.
- [18] K. Huang. *Statistical mechanics*. John Wiley & Sons, Singapore, 2nd edition, 1963.
- [19] P. Hertz. Über die mechanischen Grundlagen der Thermodynamik. *Annalen der Physik*, 33 (Leipzig):225–274 and 537–552, 1910.
- [20] C. Jarzynski. Diffusion equation for energy in ergodic adiabatic ensembles. *Physical Review A*, 46:7498–7509, 1992.
- [21] M. Toda, R. Kubo, and N. Saitô. *Statistical physics*, volume 1. Springer-Verlag, Berlin, 1983.

- [22] R. Brown, E. Ott, and C. Gregobi. Ergodic adiabatic invariants in chaotic systems. *Physical Review Letters*, 59:1173–1176, 1987.
- [23] V. L. Berdichevsky. *Thermodynamics of chaos and order*. Addison Wesley Longman, Essex, 1997.
- [24] L.D. Landau and E.M. Lifshitz. *Mechanics*. Pergamon, Oxford, 1960.
- [25] L.D. Landau and E.M. Lifshitz. *Quantum mechanics, non-relativistic theory*. Pergamon, London, 1958.
- [26] L.D. Landau and E.M. Lifshitz. *Statistical physics, I*. Pergamon, Oxford, 1978.
- [27] A. E. Allahverdyan and Th. M. Nieuwenhuizen. Minimal work principle: Proof and counterexamples. *Physical Review E*, 71(4):046107, 2005.
- [28] A. Messiah. *Quantum mechanics*, volume II. North Holland, Amsterdam, 1962.
- [29] A. E. Allahverdyan and Th. M. Nieuwenhuizen. A mathematical theorem as the basis for the second law: Thomson’s formulation applied to equilibrium. *Physica A*, 305(3-4):542–552, 2002.
- [30] A. Lenard. Thermodynamical proof of the Gibbs formula for elementary quantum systems. *Journal of Statistical Physics*, 19:575, 1978.
- [31] I. M. Bassett. Alternative derivation of the classical second law of thermodynamics. *Physical Review A*, 18(5):2356–2360, 1978.
- [32] H. Tasaki. Statistical mechanical derivation of the second law of thermodynamics. *arXiv:cond-mat/0009206*, 2000.
- [33] Th.M. Nieuwenhuizen, P.D. Keefe, and V. Špicka, editors. *Frontiers of Quantum*, volume 29 of *Physica E*, 2005.
- [34] T. D. Kieu. The second law, Maxwell’s demon, and work derivable from quantum heat engines. *Physical Review Letters*, 93(14):140403, 2004.
- [35] R. V. Pound. Nuclear spin relaxation times in single crystals of LiF. *Physical Review*, 81(1):156, 1951.
- [36] E. M. Purcell and R. V. Pound. A nuclear spin system at negative temperature. *Physical Review*, 81(2):279–280, 1951.

- [37] N. F. Ramsey and R. V. Pound. Nuclear audiofrequency spectroscopy by resonant heating of the nuclear spin system. *Physical Review*, 81(2):278–279, 1951.
- [38] N. F. Ramsey. Thermodynamics and statistical mechanics at negative absolute temperatures. *Physical Review*, 103(1):20–28, 1956.
- [39] A. P. Mosk. Atomic gases at negative kinetic temperature. *Physical Review Letters*, 95(4):040403, 2005.
- [40] D. H. E. Gross. Microcanonical thermodynamics and statistical fragmentation of dissipative systems. The topological structure of the N-body phase space. *Physics Reports*, 1997.
- [41] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmüller, B. von Issendorff, and H. Haberland. Negative heat capacity for a cluster of 147 sodium atoms. *Physical Review Letters*, 86(7):1191–1194, 2001.
- [42] F. Gobet, B. Farizon, M. Farizon, M. J. Gaillard, J. P. Buchet, M. Carré, and T. D. Märk. Probing the liquid-to-gas phase transition in a cluster via a caloric curve. *Physical Review Letters*, 87(20):203401, 2001.
- [43] F. Gobet, B. Farizon, M. Farizon, M. J. Gaillard, J. P. Buchet, M. Carré, P. Scheier, and T. D. Märk. Direct experimental evidence for a negative heat capacity in the liquid-to-gas phase transition in hydrogen cluster ions: Backbending of the caloric curve. *Physical Review Letters*, 89(18):183403, 2002.
- [44] D. Lynden-Bell. Negative specific heat in astronomy, physics and chemistry. *Physica A*, 263(1-4):293–304, 1999.
- [45] J. Barre, F. Bouchet, T. Dauxois, and S. Ruffo. Large deviation techniques applied to systems with long-range interactions. *Journal of Statistical Physics*, 119:677–713(37), May 2005.
- [46] H. Touchette, R. S. Ellis, and B. Turkington. An introduction to the thermodynamic and macrostate levels of nonequivalence of ensembles. *Physica A*, 340:138–146, 2004.

- [47] A. Campa and S. Ruffo. Microcanonical solution of the mean-field ϕ^4 model: Comparison with time averages at finite size. *Physica A*, 369(2):517–528, 2006.
- [48] D. H. E. Gross and J. F. Kenney. The microcanonical thermodynamics of finite systems: The microscopic origin of condensation and phase separations, and the conditions for heat flow from lower to higher temperatures. *The Journal of Chemical Physics*, 122(22):224111, 2005.
- [49] W. Thirring. Systems with negative specific heat. *Zeitschrift für Physik*, 235:339–352, 1970.
- [50] J. Naudts. Boltzmann entropy and the microcanonical ensemble. *Europhysics Letters*, 69(5):719–724, 2005.
- [51] A. Adib. Does the Boltzmann principle need a dynamical correction? *Journal of Statistical Physics*, 117:581–597, 2004.
- [52] E. M. Pearson, T. Halicioglu, and W. A. Tiller. Laplace-transform technique for deriving thermodynamic equations from the classical microcanonical ensemble. *Physical Review A*, 32(5):3030–3039, 1985.
- [53] M. Bianucci, R. Mannella, B. West, and P. Grigolini. From dynamics to thermodynamics: linear response and statistical mechanics. *Physical Review E*, 51:3002–3022, 1995.
- [54] F. Cardin and M. Favretti. On the Helmholtz-Boltzmann thermodynamics of mechanical systems. *Continuum Mechanics and Thermodynamics*, 16(1-2):15–29, 2004.
- [55] L. Galgani and A. Scott. Planck-like distributions in classical nonlinear mechanics. *Physical Review Letters*, 28(18):1173–1176, 1972.
- [56] M. Antoni and S. Ruffo. Clustering and relaxation in hamiltonian long-range dynamics. *Physical Review E*, 52(3):2361–2374, 1995.
- [57] D. Mukamel, S. Ruffo, and N. Schreiber. Breaking of ergodicity and long relaxation times in systems with long-range interactions. *Physical Review Letters*, 95:240604, 2005.

- [58] D. Ruelle. Statistical mechanics of a one-dimensional lattice gas. *Communications in Mathematical Physics*, 9(4):267–278, 1968.
- [59] J. Dunkel and S. Hilbert. Phase transitions in small systems: Microcanonical vs. canonical ensembles. *Physica A*, 370(2):390–406, 2006.
- [60] S. Hilbert and J. Dunkel. Nonanalytic microscopic phase transitions and temperature oscillations in the microcanonical ensemble: An exactly solvable one-dimensional model for evaporation. *Physical Review E*, 74(1):011120, 2006.
- [61] M. Kastner, S. Schreiber, and O. Schnetz. Phase transitions from saddles of the potential energy landscape. *Physical Review Letters*, 99(5):050601, 2007.
- [62] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, 52:479, 1988.
- [63] M. P. Almeida. Generalized entropy from first principles. *Physica A*, 300:424–432, 2001.
- [64] S. Abe and G. B. Bagci. Necessity of q-expectation value in nonextensive statistical mechanics. *Physical Review E*, 71, 2005.
- [65] S. Martínez, F. Pennini, A. Plastino, and C. Tessone. On the equipartition and virial theorems. *Physica A*, 305, 2002.
- [66] A. Rényi. On measures of information and entropy. In *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, pages 547–561, 1960.
- [67] E. Vives and A. Planes. Is Tsallis thermodynamics nonextensive? *Physical Review Letters*, 88(2):020601, 2002.
- [68] S. Abe, S. Martinez, F. Pennini, and A. Plastino. Nonextensive thermodynamic relations. *Physics Letters A*, 281:126, 2001.
- [69] C. Tsallis, R. S. Mendes, and A. R. Plastino. The role of constraints within generalized nonextensive statistics. *Physica A*, 261:534, 1998.
- [70] S. Abe. Temperature of nonextensive systems: Tsallis entropy as Clausius entropy. *Physica A*, 368(2):430–434, 2006.

- [71] A. R. Plastino and A. Plastino. From Gibbs microcanonical ensemble to Tsallis generalized canonical distribution. *Physics Letters A*, 193:140, 1994.
- [72] S. Abe and A. K. Rajagopal. Macroscopic thermodynamics of equilibrium characterized by power law canonical distributions. *Europhysics Letters*, 55:6–11, 2001.
- [73] A.G. Bashkurov. On maximum entropy principle, superstatistics, power-law distribution and renyi parameter. *Physica A*, 340:153–162, 2004.
- [74] J. Shore and R. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26(1):26–37, 1980.
- [75] R. Johnson and J. Shore. Comments on and correction to 'axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy'. *IEEE Transactions on Information Theory*, 29(6):942–943, 1983.
- [76] J. Uffink. Can the maximum entropy principle be explained as a consistency requirement? *Studies In History and Philosophy of Modern Physics*, 26(3):223–261, 1995.
- [77] A. G. Bashkurov. Rényi entropy as a statistical entropy for complex systems. *Journal Theoretical and Mathematical Physics*, 149(2):1559–1573, 2006.
- [78] M. Marino. A generalized thermodynamics for power-law statistics. *Physica A*, 2007. in press.
- [79] M. S. S. Challa and J. H. Hetherington. Gaussian ensemble as an interpolating ensemble. *Physical Review Letters*, 60(2):77–80, 1988.
- [80] J. Naudts. Parameter estimation in non-extensive thermostatics. *Physica A*, 365:42–49, 2006.
- [81] R. Hanel and S. Thurner. Generalized Boltzmann factors and the maximum entropy principle: Entropies for complex systems. *Physica A*, 380:109–114, 2007.