Lecture Notes in Physics

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Editorial Policy

The series *Lecture Notes in Physics* (LNP), founded in 1969, reports new developments in physics research and teaching -- quickly, informally but with a high quality. Manuscripts to be considered for publication are topical volumes consisting of a limited number of contributions, carefully edited and closely related to each other. Each contribution should contain at least partly original and previously unpublished material, be written in a clear, pedagogical style and aimed at a broader readership, especially graduate students and nonspecialist researchers wishing to familiarize themselves with the topic concerned. For this reason, traditional proceedings cannot be considered for this series though volumes to appear in this series are often based on material presented at conferences, workshops and schools (in exceptional cases the original papers and/or those not included in the printed book may be added on an accompanying CD ROM, together with the abstracts of posters and other material suitable for publication, e.g. large tables, colour pictures, program codes, etc.).

Acceptance

A project can only be accepted tentatively for publication, by both the editorial board and the publisher, following thorough examination of the material submitted. The book proposal sent to the publisher should consist at least of a preliminary table of contents outlining the structure of the book together with abstracts of all contributions to be included.
Final acceptance is issued by the series editor in charge, in consultation with the publisher, only after receiving the complete manuscript. Final acceptance, possibly requiring minor corrections, usually follows the tentative acceptance unless the final manuscript differs significantly from expectations (project outline). In particular, the series editors are entitled to reject individual contributions if they do not meet the high quality standards of this series. The final manuscript must be camera-ready, and should include both an informative introduction and a sufficiently detailed subject index.

Contractual Aspects

Publication in LNP is free of charge. There is no formal contract, no royalties are paid, and no bulk orders are required, although special discounts are offered in this case. The volume editors receive jointly 30 free copies for their personal use and are entitled, as are the contributing authors, to purchase Springer books at a reduced rate. The publisher secures the copyright for each volume. As a rule, no reprints of individual contributions can be supplied.

Manuscript Submission

The manuscript in its final and approved version must be submitted in camera-ready form. The corresponding electronic source files are also required for the production process, in particular the online version. Technical assistance in compiling the final manuscript can be provided by the publisher's production editor(s), especially with regard to the publisher's own Latex macro package which has been specially designed for this series.

Online Version/ LNP Homepage

LNP homepage (list of available titles, aims and scope, editorial contacts etc.):
http://www.springer.de/phys/books/lnpp/

LNP online (abstracts, full-texts, subscriptions etc.):
http://link.springer.de/series/lnpp/

Sumiyoshi Abe Yuko Okamoto (Eds.)

Nonextensive Statistical Mechanics and Its Applications

Springer

Editors

Sumiyoshi Abe
College of Science and Technology
Nihon University
Funabashi
Chiba 274-8501, Japan

Yuko Okamoto
Department of Theoretical Studies
Institute for Molecular Science
Okazaki, Aichi 444-8585, Japan

Cover picture: see contribution by Tsallis in this volume.

Library of Congress Cataloging-in-Publication Data applied for.

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Nonextensive statistical mechanics and its applications / Sumiyoshi
Abe ; Yuko Okamoto (ed.). - Berlin ; Heidelberg ; New York ; Barcelona
; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer,
2001
(Lecture notes in physics ; Vol. 560)
(Physics and astronomy online library)
ISBN 3-540-41208-5

ISSN 0075-8450
ISBN 3-540-41208-5 Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera-ready by the authors/editors
Camera-data conversion by Steingraeber Satztechnik GmbH Heidelberg
Cover design: *design & production*, Heidelberg

Printed on acid-free paper
SPIN: 10786438 55/3141/du - 5 4 3 2 1 0

Preface

It is known that in spite of its great success Boltzmann–Gibbs statistical mechanics is actually not completely universal. A class of physical ensembles involving long-range interactions, long-time memories, or (multi-)fractal structures can hardly be treated within the traditional statistical-mechanical framework. A recent nonextensive generalization of Boltzmann–Gibbs theory, which is referred to as nonextensive statistical mechanics, enables one to analyze such systems. This new stream in the foundation of statistical mechanics was initiated by Tsallis' proposal of a nonextensive entropy in 1988. Subsequently it turned out that consistent generalized thermodynamics can be constructed on the basis of the Tsallis entropy, and since then we have witnessed an explosion in research works on this subject.

Nonextensive statistical mechanics is still a rapidly growing field, even at a fundamental level. However, some remarkable structures and a variety of interesting applications have already appeared. Therefore, it seems quite timely now to summarize these developments.

This volume is primarily based on The IMS Winter School on Statistical Mechanics: Nonextensive Generalization of Boltzmann–Gibbs Statistical Mechanics and Its Applications (February 15-18, 1999, Institute for Molecular Science, Okazaki, Japan), which was supported, in part, by IMS and the Japanese Society for the Promotion of Science. The volume consists of a set of four self-contained lectures, together with additional short contributions. The topics covered are quite broad, ranging from astrophysics to biophysics. Some of the latest developments since the School are also included herein.

We would like to thank Professors W. Beiglböck and H.A. Weidenmüller for their advice and encouragement.

Funabashi, *Sumiyoshi Abe*
Okazaki, *Yuko Okamoto*
November 2000

Contents

Part 1 Lectures on Nonextensive Statistical Mechanics

I. Nonextensive Statistical Mechanics and Thermodynamics: Historical Background and Present Status

IV. Computational Methods for the Simulation of Classical and Quantum Many Body Systems Sprung from Nonextensive Thermostatistics

---

Part 2 Further Topics

V. Correlation Induced by Nonextensivity and the Zeroth Law of Thermodynamics

VI. Dynamic and Thermodynamic Stability of Nonextensive Systems

VII. Generalized Simulated Annealing Algorithms Using Tsallis Statistics: Application to ±J Spin Glass Model

I. Nonextensive Statistical Mechanics and Thermodynamics: Historical Background and Present Status

C. Tsallis

Department of Physics, University of North Texas
P.O. Box 311427, Denton, Texas 76203-1427, USA
`tsallis@unt.edu`
and
Centro Brasileiro de Pesquisas Físicas
Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro-RJ, Brazil
`tsallis@cbpf.br`

Abstract. The domain of validity of standard thermodynamics and Boltzmann-Gibbs statistical mechanics is focused on along a historical perspective. It is then formally enlarged in order to hopefully cover a variety of anomalous systems. The generalization concerns *nonextensive* systems, where nonextensivity is understood in the thermodynamical sense. This generalization was first proposed in 1988 inspired by the probabilistic description of multifractal geometry, and has been intensively studied during this decade. In the present effort, we describe the formalism, discuss the main ideas, and then exhibit the present status in what concerns theoretical, experimental and computational evidences and connections, as well as some perspectives for the future. The whole review can be considered as an attempt to clarify our current understanding of the foundations of statistical mechanics and its thermodynamical implications.

1 Introduction

The present effort is an attempt to review, in a self-contained manner, a one-decade-old nonextensive generalization [1,2] of standard statistical mechanics and thermodynamics, as well as to update and discuss the recent associated developments [3]. Concomitantly, we shall address, on physical grounds, the domain of validity of the Boltzmann-Gibbs (BG) formalism, i.e., under what restrictions it is expected to be valid. Although only the degree of universality of BG thermal statistics will be focused on, let us first make some generic comments.

In some sense, every physical *phenomenon* occurs *somewhere* at *some time* [4]. Consistently, the ultimate (most probably unattainable!) goal of physical sciences is, in what theory concerns, to develop formalisms that approach as much as possible *universality* (i.e., valid for all phenomena), *ubiquity* (i.e., valid everywhere) and *eternity* (i.e., valid always). Since these words are very rich in meanings, let us illustrate what we specifically refer to through the best known physical formalism, namely Newtonian mechanics. After its plethoric verifications along the centuries, it seems fair to say that in some sense Newtonian mechanics is "eternal" and "ubiquitous". However, we do know that it is *not*

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 3–98, 2001.

universal. Indeed, we all know that, when the involved velocities approach that of light in the vacuum, Newtonian mechanics becomes only an approximation (an increasingly bad one for velocities increasingly closer to that of light) and reality appears to be better described by special relativity. Analogously, when the involved masses are as small as say the electron mass, once again Newtonian mechanics becomes but a (bad) approximation, and quantum mechanics becomes necessary to understand nature. Also, if the involved masses are very large, Newtonian mechanics has to be extended into general relativity. To say it in other words, we know nowadays that, whenever $1/c$ (inverse speed of light in vacuum) and/or h (Planck constant) and/or G (gravitational constant) are different from zero, Newtonian mechanics is, strictly speaking, false since it only conserves an asymptotic validity.

Along these lines, what can we say about BG statistical mechanics and standard thermodynamics? A diffuse belief exists, among not few physicists as well as other scientists, that these two interconnected formalisms are eternal, ubiquitous and universal. It is clear that, after more than one century highly successful applications of standard thermodynamics and the magnificent Boltzmann's connection of Clausius *macroscopic* entropy to the theory of probabilities applied to the *microscopic* world , BG thermal statistics can (and should!) easily be considered as one of the pillars of modern science. Consistently, it is certainly fair to say that BG thermostatistics and its associated thermodynamics are eternal and ubiquitous, in precisely the same sense that we have used above for Newtonian mechanics. But, again in complete analogy with Newtonian mechanics, we can by no means consider them as universal. It is *unavoidable* to think that, like all other constructs of human mind, these formalisms *must* have physical restrictions, i.e., domains of applicability, out of which they can at best be but approximations.

The *precise* mathematical definition of the domain of validity of the BG statistical mechanics is an extremely subtle and yet unsolved problem (for example, the associated canonical equilibrium distribution is considered a *dogma* by Takens [5]); such a rigorous mathematical approach is out of the scope of the present effort. Here we shall focus on this problem in three steps. The first one is deeply related to Krylov's pioneering insights [6] (see also [7–9]). Indeed, Krylov argued (half a century ago!) that the property which stands as the hard foundation of BG statistical mechanics, is *not ergodicity* but *mixing*, more precisely, quick enough, *exponential* mixing, i.e., *positive* largest Liapunov exponent. We shall refer to such situation as *strong chaos*. This condition would essentially guarantee physically short *relaxation times* and, we believe, thermodynamic *extensivity*. We argue here that whenever the largest Liapunov exponent vanishes, we can have slow, typically *power-law* mixing (see also [8,9]). Such situations will be referred as *weak chaos*. It is expected to be associated with algebraic, instead of exponential, relaxations, and to thermodynamic *nonextensivity*, hopefully for large classes of anomalous systems, of the type described in the present review.

The second step concerns the question of what geometrical structure can be responsible for the mixing being of the exponential or of the algebraic type. The picture which emerges (details will be seen later on) is that when the physically relevant phase space (or analogous quantum concept) is smooth, Euclidean-like (in the sense that it is continuous, differentiable, etc.), the mixing is of the exponential type. In contrast, if that space has a multifractal structure, then the mixing becomes kind of uneasy, and tends to be of the algebraic (or even slower) type.

The third and last step concerns the question of what kind of physical circumstances can produce a smooth (translationally invariant in some sense) or, instead, a multifractal (scaling invariant in some sense) structure for the relevant phase space. At first approximation the scenario seems to be as follows. If the effective microscopic interactions are *short*-ranged (i.e., *close* spatial connections) *and* the effective microscopic memory is *short*-ranged (i.e., *close* time connections, for instance Markovian-like processes) *and* the boundary conditions are *smooth, non(multi)fractal and* the initial conditions are standard ones *and* no peculiar mesoscopic dissipation occurs (e.g., like that occurring in various types of granular matter), etc, *then* the above mentioned space is smooth, and BG statistical mechanics appears to correctly describe nature. If one or more of these restrictions is violated (long-range interactions and/or irreducibly nonmarkovian microscopic memory [10] and/or multifractal boundary conditions and/or quite pathological initial conditions are imposed and/or some special types of dissipation are acting, etc., then the above mentioned space can be multifractally structured, and anomalous, nonextensive statistical mechanics seem to be necessary to describe nature (see also [11]).

To summarize the overall picture, we may say, roughly speaking, that a smooth relevant phase space tends to correspond to BG statistical mechanics, *exponential* mixing, energy-dependence of the canonical equilibrium distribution (i.e., the celebrated Boltzmann factor) and time-dependence of typical relaxation processes, and *extensive* thermodynamics (entropy, thermodynamic potentials and similar quantities proportional to the number of microscopic elements of the system). In contrast, a multifractally structured phase space tends to correspond to anomalous statistical mechanics (hopefully, for at least some of the typical situations herein described in some detail), *power-law* mixing, energy-dependence of the canonical equilibrium distribution and time-dependence of typical relaxation processes, and *nonextensive* thermodynamics (anomalous entropy, thermodynamic potentials and similar quantities). The basic group of symmetries would be continuous translations (or rotations) in the first case, and dilatations in the second one. (This opens, of course, the door to even more general scenarios, respectively associated to more complex groups of symmetries [12], but again this is out of our present scope). The actual situation is naturally expected to be more complex and cross-imbricated that the one just sketched here, but these would nevertheless be the essential guiding lines.

Before entering into the nonextensive thermostatistical formalism herein addressed, let us mention at least some of the thermodynamical anomalies that we have in mind as physical motivations. As argued above, it is nowadays quite well known that a variety of physical systems exists for which the powerful (and beautiful!) BG statistical mechanics and standard thermodynamics present serious difficulties, which can occasionally achieve the status of just plain failures. The list of such anomalies increases every day. Indeed, here and there, features are pointed out which defy (not to say, in some cases, that plainly violate!) the standard BG prescriptions. The violation is, in some examples, clear; in others, the situation is different. Either an explanation within the BG framework is, perhaps, possible but it has not yet been exhibited convincingly. Or the problem indeed lies out of the realm of standard equilibrium and nonequilibrium BG statistics. Our hope and belief is that the present nonextensive statistics might correctly cover at least some of the known anomalies. Within a long list that will be systematically focused on later on with more details, we may mention at this point systems involving long-range interactions [13–17] (e.g., $d = 3$ gravitation [18,19], ferromagnetism [20], spin-glasses [21]), long-range microscopic memory (e.g., nonmarkovian stochastic processes, on which much remains to be known, in fact) [10,22,23], pure-electron plasma two-dimensional turbulence [24], Lévy anomalous diffusion [25], granular systems [26], phonon-electron anomalous thermalization in ion-bombarded solids [27,28], solar neutrinos [29], peculiar velocities of galaxies [30], inverse bremsstrahlung in plasma [31], black holes [32], cosmology [33], high energy collisions of elementary (or more complex) particles [34–39], quantum entanglement [40], among others. Some of these examples clearly appear to be out of the domain of validity of the standard formalisms; others might be considered as more controversial. In any case, the present status of all of them, and even some others, will be discussed in Sections 3, 4 and 5.

2 Formalism

2.1 Entropy

As an attempt to overcome at least some of the difficulties mentioned in the previous Section, a proposal has been advanced, one decade ago [1], (see also [41,42]), which is based on a generalized entropic form, namely

$$S_q = k\frac{1-\sum_{i=1}^{W} p_i^q}{q-1} \qquad \left(\sum_{i=1}^{W} p_i = 1;\ q \in \mathcal{R}\right) \ , \tag{1}$$

where k is a positive constant and W is the total number of microscopic possibilities of the system. For the $q < 0$ case, care must be taken to exclude all those possibilities whose probability is not strictly positive, otherwise S_q would diverge. Such care is not necessary for $q > 0$; due to this property, the entropy is said to be *expansible* for $q > 0$; more explicitly, we can verify that $S_q(p_1, p_2, ..., p_W, 0) = S_q(p_1, p_2, ..., p_W)$ $(\forall\{p_i\}; \forall q)$. Expression (1) recovers (using $p_i^{q-1} = e^{(q-1)\ln p_i} \sim 1 + (q-1)\ln p_i$) the usual BG entropy $(-k_B \sum_{i=1}^{W} p_i \ \ln \ p_i)$ in the limit $q \to 1$, where k_B is the Boltzmann constant. The constant k presumably coincides with k_B for all values of q; however, nothing that we are presently aware of forbids it to be *proportional* to k_B, the proportionality factor being (for dimensional reasons) a pure number which might depend on q (clearly, this pure number must be unity for $q = 1$). In many of the applications along this text, we might without further notice (and without loss of generality) consider units such that $k = 1$.

The quantum version of expression (1) is given [43] by

$$S_q = k\frac{1 - \mathrm{Tr}\,\rho^q}{q-1} \qquad (\mathrm{Tr}\,\rho = 1)\ , \tag{2}$$

where ρ is the density operator. Of course, in the particular instance when ρ is diagonal in a W-dimensional Hilbert space, we recover Eq. (1).

The classical version of expression (1) is given by

$$S_q = k\frac{1 - \int d(\mathbf{x}/\sigma)\ [\sigma\ p(\mathbf{x})]^q}{q-1} \qquad (\int d\mathbf{x}\, p(\mathbf{x}) = 1)\ , \tag{3}$$

where $\mathbf{x} \equiv (x_1, x_2, ..., x_d)$ and $\sigma \equiv \Pi_{r=1}^{d}\sigma_r$, σ_r being, for all values of r, a characteristic constant whose dimension equals that of x_r, in such a way that all $\{x_r/\sigma_r\}$ are pure numbers. If all $\{x_r\}$ are already pure numbers, then $\sigma_r = 1$ $(\forall r)$, hence $\sigma = 1$. Of course, if $\sigma p(\mathbf{x}) = \sum_{i=1}^{W} \delta(\mathbf{x} - \mathbf{x}_i)$, $\delta(...)$ being the d-dimensional Dirac delta, we recover Eq. (1).

It is important that we point out right away that the Boltzmann entropy can be clearly differentiated (see for instance [44]) from the Gibbs entropy in what concerns the variables to which they apply. Moreover, besides Boltzmann and Gibbs, many other scientists, such as von Neumann, Ehrenfest, Szilard, Shannon, Jaynes, Kolmogorov, Sinai, Prigogine, Lebowitz, Zurek, have given invaluable contributions to the subject of the statistical entropies and their connections to the Clausius entropy. However, for simplicity, and because we are focusing on the *functional form* of the entropy, we shall here indistinctly refer to the $q = 1$ particular cases of Eqs. (1-3) as the *Boltzmann-Gibbs entropy.*

Another historical point which deserves to be mentioned at this stage is that, as we discovered along the years after 1988, the form (1) (occasionally with some different q-dependent factor) had already been introduced in the community of cybernetics and information long ago. More precisely, by Harvda and Charvat [45] in 1967, further discussed by Vajda [46] (who quotes [45]) in 1968, and again re-discovered in the initial form by Daroczy [47] (who apparently was unaware of his predecessors) in 1970. There are perhaps others, especially since in that community close to 25 (!) different entropic forms [48] have been advanced for a variety of specific purposes (e.g., image processing). Daroczy' s work became relatively known nowadays; ourselves, we mentioned it in 1991 [41], and some historical review was done in 1995 [49]; however, we are not aware of *any* exhaustive description of all these entropic forms and their interconnections. In fact, this would be a quite heavy task! Indeed, to the 20-25 entropic forms introduced in communities *other* than Physics, we must now add several more entropic forms that appeared (see, for instance, references [42,50–60] as well as the end of the present subsection 5.5) *within* the Physics community after paper [1]. In any case, at least as far as we know, it is allowed to believe that no proposal before this 1988 paper was advanced for generalizing, along the present nonextensive path, standard statistical mechanics and thermodynamics.

The entropic index q (intimately related to and determined by the microscopic dynamics, as we shall argue later on) characterizes the *degree of nonextensivity* reflected in the following *pseudo-extensivity* entropy rule

$$\frac{S_q(A+B)}{k} = \frac{S_q(A)}{k} + \frac{S_q(B)}{k} + (1-q)\frac{S_q(A)}{k}\frac{S_q(B)}{k} \; , \tag{4}$$

where A and B are two *independent* systems in the sense that the probabilities of $A+B$ *factorize* into those of A and of B (i.e., $p_{ij}(A+B) = p_i(A)p_j(B)$). We immediately see that, since in all cases $S_q \geq 0$ (*nonnegativity* property), $q<1$, $q=1$ and $q>1$ respectively correspond to *superadditivity (superextensivity), additivity (extensivity) and subadditivity (subextensivity).* The pseudo-extensivity property (4) can be equivalently written as follows:

$$\begin{aligned}&\frac{\ln[1+(1-q)S_q(A+B)/k]}{1-q}\\ &= \frac{\ln[1+(1-q)S_q(A)/k]}{1-q} + \frac{\ln[1+(1-q)S_q(B)/k]}{1-q} \; .\end{aligned} \tag{5}$$

We shall come back onto this form later on in connection with the widely known Renyi's entropy [61] (in fact first introduced, according to Csiszar [62], by Schutzenberger [63]).

Let us mention at this point that expression (1) exhibits a property which has apparently never been focused before, and which we shall from now on refer to as the *composability* property. It concerns the nontrivial fact that the entropy $S(A+B)$ of a system composed of two independent subsystems A and B can be calculated from the entropies $S(A)$ and $S(B)$ of the subsystems, *without any need of microscopic knowledge about A and B, other than the knowledge of some generic universality class, herein the nonextensive universality class, represented by the entropic index q*, i.e., without any knowledge about the microscopic possibilities of A and B nor their associated probabilities. This property is so obvious for the BG entropic form that the (false!) idea that *all* entropic forms automatically satisfy it could easily install itself in the mind of most physicists. To show counterexamples, it is enough to check that the recently introduced Anteneodo-Plastino's [50] and Curado's [55] entropic forms satisfy a variety of interesting properties, and nevertheless are *not* composable. See [64] for more details.

Another important (since it eloquently exhibits the surprising effects of nonextensivity) property is the following. Suppose that the set of W possibilities is arbitrarily separated into two subsets having respectively W_L and W_M possibilities ($W_L + W_M = W$). We define $p_L \equiv \sum_{i=1}^{W_L} p_i$ and $p_M \equiv \sum_{i=W_L+1}^{W} p_i$, hence $p_L + p_M = 1$. It can then be straightforwardly established that

$$S_q(\{p_i\}) = S_q(p_L, p_M) \; + p_L^q \; S_q(\{p_i/p_L\}) \; + \; p_M^q \; S_q(\{p_i/p_M\}) \; , \tag{6}$$

where the sets $\{p_i/p_L\}$ and $\{p_i/p_M\}$ are the *conditional* probabilities. This would precisely be the famous Shannon's property *were it not for the fact that, in front* *of the entropies associated with the conditional probabilities, appear* p_L^q *and* p_M^q *instead of* p_L *and* p_M . This fact will play, as we shall see later on, a central role in the whole generalization of thermostatistics. Indeed, since the probabilities $\{p_i\}$ are generically numbers between zero and unity, $p_i^q > p_i$ for $q<1$ and $p_i^q < p_i$ for $q>1$, hence $q<1$ and $q>1$ will respectively privilege the *rare* and the *frequent* events. This simple property lies at the heart of the whole proposal. Santos has recently shown [65], strictly following along the lines of Shannon himself, that, if we assume (i) continuity (in the $\{p_i\}$) of the entropy, (ii) increasing monotonicity of the entropy as a function of W in the case of equiprobability, (iii) property (4), and (iv) property (6), then *only one entropic form exists, namely that given in definition* (1).

The generalization of Eq. (6) to the case where, instead of two, we have R nonintersecting subsets $(W_1 + W_2 + ... + W_R = W)$ is straightforward [49]. To be more specific, if we define

$$\pi_j \equiv \sum_{W_j \; terms} p_i \quad (j = 1, 2, ..., R) \ , \tag{7}$$

(hence $\sum_{j=1}^{R} \pi_j = 1$), Eq. (4) is generalized into

$$S_q(\{p_i\}) = S_q(\{\pi_j\}) + \sum_{j=1}^{R} \pi_j^q S_q(\{p_i/\pi_j\}) \ , \tag{8}$$

where we notice, in the last term, the emergence of what we shall soon introduce generically as the unnormalized q-expectation value (of the *conditional* entropies $S_q(\{p_i/\pi_j\})$, in the present case).

Another interesting property is the following. The Boltzmann-Gibbs entropy S_1 satisfies the following relation [66]:

$$-k\left[\frac{d}{d\alpha}\sum_{i=1}^{W} p_i^{\alpha}\right]_{\alpha=1} = -k\sum_{i=1}^{W} p_i \, \ln \, p_i \equiv S_1 \ . \tag{9}$$

Moreover, Jackson introduced in 1909 [67] the following generalized differential operator (applied to an arbitrary function $f(x)$):

$$D_q \, f(x) \equiv \frac{f(qx) - f(x)}{qx - x}, \tag{10}$$

which satisfies $D_1 \equiv \lim_{q \to 1} D_q = \frac{d}{dx}$. Abe [66] recently remarked that

$$-k\left[D_q \, \sum_{i=1}^{W} p_i^{\alpha}\right]_{\alpha=1} = k\frac{1 - \sum_{i=1}^{W} p_i^q}{q-1} \equiv S_q \ . \tag{11}$$

This property provides an intuitive insight into the generalized entropic form S_q. Indeed, the inspiration for its use in order to generalize the usual thermal statistics came [1] from multifractals, and its applications concern, in one way or another, systems which exhibit scale invariance. Therefore, its connection with Jackson's differential operator appears to be rather natural. Indeed, this operator "tests" the function $f(x)$ under *dilatation* of x, in contrast to the usual derivative, which "tests" it under *translation* of x [68].

Another property which no doubt must be mentioned in the present introduction is that S_q is consistent with Laplace's maximum ignorance principle, i.e., it is extremum at *equiprobability* ($p_i = 1/W, \ \ \forall i$). This extremum is given by

$$S_q = k\frac{W^{1-q}-1}{1-q} \qquad (W \geq 1) \ , \tag{12}$$

which, in the limit $q \to 1$, reproduces Boltzmann's celebrated formula $S = k \ln\, W$ (carved on his marble grave in the Central Cemetery of Vienna). In the limit $W \to \infty$, S_q/k diverges if $q \leq 1$, and saturates at $1/(q-1)$ if $q > 1$. By using the q-logarithm function [69,70] (see Appendix), Eq. (12) can be rewritten in the following Boltzmann-like form:

$$S_q = k \ \ln_q \ W \ . \tag{13}$$

Finally, let us close the present set of properties by reminding that S_q has, with regard to $\{p_i\}$, a *definite concavity* for *all* values of q (S_q is always concave for $q > 0$ and always convex for $q < 0$). In this sense, it contrasts with Renyi's entropy (quite useful in the geometrical characterization of strange attractors and similar multifractal structures; see [71] and references therein) $S_q^R \equiv (\ln\ \sum_{i=1}^W p_i^q)/(1-q) = \{\ln\ [1+(1-q)S_q/k]\}/(1-q)$, which does not have this property for *all* values of q, but only for $q \leq 1$.

Let us now introduce, for an arbitrary physical quantity A, the following *unnormalized q-expectation value*

$$\langle A \rangle_q \equiv \sum_{i=1}^{W} p_i^q \ A_i \ , \tag{14}$$

as well as the *normalized q-expectation value*

$$\langle\langle A \rangle\rangle_q \equiv \frac{\sum_{i=1}^W p_i^q \ A_i}{\sum_{i=1}^W p_i^q} \ . \tag{15}$$

We verify that both $\langle A \rangle_1$ and $\langle\langle A \rangle\rangle_1$ coincide with the standard mean value $\langle A \rangle$ of a A. We also verify that

$$\langle\langle A \rangle\rangle_q = \frac{\langle A \rangle_q}{\langle 1 \rangle_q} \ , \tag{16}$$

and notice that, whereas $\langle\langle 1 \rangle\rangle_q = 1 \ (\forall q)$, in general $\langle 1 \rangle_q \neq 1$.

Let us now go back to the nonextensive entropy. We can easily verify that

$$S_q = k\langle -\ln_q p_i \rangle_q \tag{17}$$

and that

$$S_q = k\langle \ln_q(1/p_i)\rangle_1 \ . \tag{18}$$

For the $q=1$ case, the quantity $-\ln p_i = \ln(1/p_i)$ has been eloquently called *surprise* by Watanabe [72], and *unexpectedness* by Barlow [73]. The question which now arises is which quantity should we call *q-surprise* (or *q-unexpectedness*), $-\ln_q p_i$ or $\ln_q(1/p_i)$? The question is more than semantics since it will point the natural physical quantity whose appropriate average provides S_q. We can easily check that (i) $-\ln_0 p_i = 1-p_i$ plays the role of a separatrix, $-\ln_q p_i$ being convex for all $q>0$ and concave for all $q<0$; (ii) $\ln_2(1/p_i) = 1-p_i$ also plays the role of a separatrix, $\ln_q(1/p_i)$ being convex for all $q<2$ and concave for all $q>2$. Since concavity of S_q changes sign at $q=0$, there is a compelling reason for having a separatrix at that value, whereas no such reason exists for $q=2$. Consistently it is $-\ln_q p_i$ that we shall adopt as the q-quantity generalizing $-\ln p_i$. We notice also that it is the q-expectation values, and not the standard mean values, which naturally enter into the formalism. This is consistent with Eq. (8), for instance, and will prove to be of extreme mathematical utility in replacing *divergent* sums and integrals by *finite* analogous sums and integrals (see later on our discussion of Lévy-like anomalous superdiffusion).

If our system is a generic quantum one we must use, as already mentioned, the density operator ρ. The unnormalized and normalized q-expectation values of an observable A (which not necessarily commutes with ρ) are respectively given by

$$\langle A\rangle_q \equiv \mathrm{Tr}\ \rho^q A \tag{19}$$

and

$$\langle\langle A\rangle\rangle_q \equiv \frac{\mathrm{Tr}\ \rho^q A}{\mathrm{Tr}\ \rho^q} \ . \tag{20}$$

In particular, the entropy S_q is given by

$$S_q = \langle \hat{S}_q\rangle_q \ , \tag{21}$$

where the *entropy operator* is given by

$$\hat{S}_q \equiv -k\ln_q \rho \ . \tag{22}$$

It is worthy mentioning that the following pseudo-extensivity holds for the operators:

$$\begin{aligned}\frac{\hat{S}_q(A+B)}{k} &= \frac{\hat{S}_q(A)}{k} + \frac{\hat{S}_q(B)}{k} \\ &\quad + (q-1)\frac{\hat{S}_q(A)}{k}\frac{\hat{S}_q(B)}{k} \ ,\end{aligned} \tag{23}$$

where we must notice the appearance of a $(q-1)$ factor in front of the cross term, where there was a $(1-q)$ factor in Eq. (4)!

The same type of considerations hold, *mutatis mutandis*, in the case when our system is a generic classical one.

2.2 Canonical Ensemble

Once we have a generalized entropic form, say that given in Eq. (1) (or an even more general one, or a different one), we can use it in a variety of ways. For instance, if we are interested in cybernetics, information theory, some optimization algorithms, image processing, among others, we can take advantage of a particular form in a variety of manners. However, if our primary interest is Physics, this is to say the (qualitative and quantitative) description and possible understanding of phenomena occurring in nature, then we are naturally led to use the available generalized entropy in order to generalize statistical mechanics itself and, if unavoidable, even thermodynamics. It is along this line that we shall proceed from now on (see also [74]). To do so, the first nontrivial (and quite ubiquitous) physical situation is that in which a given system is in contact with a thermostat at temperature T. To study this, we shall follow along Gibbs' path and focus on the so called *canonical ensemble*. More precisely, to obtain the thermal equilibrium distribution associated with a conservative (Hamiltonian) physical system in contact with the thermostat we shall extremize S_q under appropriate constraints. These constraints are [42]

$$\sum_{i=1}^{W} p_i = 1 \qquad\qquad (\textit{norm constraint}) \tag{24}$$

and

$$\langle\langle \epsilon_i \rangle\rangle_q \equiv \frac{\sum_{i=1}^{W} p_i^q \epsilon_i}{\sum_{i=1}^{W} p_i^q} = U_q \quad (\textit{energy constraint}) \ , \tag{25}$$

where $\{\epsilon_i\}$ are the eigenvalues of the Hamiltonian of the system. We refer to $\langle\langle ... \rangle\rangle_q$ as the *normalized q-expectation value*, as previously mentioned, and to U_q as the *generalized internal energy* (assumed *finite* and fixed). It is clear that, in the $q \to 1$ limit, these quantities recover the standard mean value and internal energy respectively.

The outcome of this optimization procedure is given by

$$p_i = \frac{\left[1-(1-q)\beta(\epsilon_i - U_q)/\sum_{j=1}^{W}(p_j)^q\right]^{\frac{1}{1-q}}}{\bar{Z}_q} \tag{26}$$

with

$$\bar{Z}_q(\beta) \equiv \sum_{i=1}^{W} \left[1-(1-q)\beta(\epsilon_i - U_q)/\sum_{j=1}^{W}(p_j)^q\right]^{\frac{1}{1-q}} \ . \tag{27}$$

It can be shown that, for the case $q < 1$, the expression of the equilibrium distribution is complemented by the auxiliary condition that $p_i = 0$ whenever the argument of the function becomes negative (*cut-off* condition). Also, it can be shown [42] that

$$1/T = \partial S_q/\partial U_q, \ \ \forall q \quad (T \equiv 1/(k\beta)) \ . \tag{28}$$

Furthermore, it is important to notice that, if we add a constant ϵ_0 to all $\{\epsilon_i\}$, we have (as it can be self-consistently proved) that U_q becomes $U_q + \epsilon_0$, which leaves invariant the differences $\{\epsilon_i - U_q\}$, which, in turn, (self-consistently) leaves *invariant* the set of probabilities $\{p_i\}$, hence *all* the thermostatistical quantities. It is also trivial to show that, for the independent systems A and B mentioned previously, $U_q(A+B) = U_q(A) + U_q(B)$, thus recovering the same form of the standard ($q = 1$) thermodynamics.

It can be shown that the following relations hold:

$$\sum_{i=1}^{W} (p_i)^q = (\bar{Z}_q)^{1-q} \ , \tag{29}$$

hence

$$S_q = k \ln_q \ \bar{Z}_q \tag{30}$$

(which recovers Eq. (13) at the $T \to \infty$ limit), and also

$$F_q \equiv U_q - TS_q = -\frac{1}{\beta}\frac{(Z_q)^{1-q} - 1}{1-q} = -\frac{1}{\beta}\ln_q \ Z_q \tag{31}$$

and

$$U_q = -\frac{\partial}{\partial\beta}\frac{(Z_q)^{1-q} - 1}{1-q} = -\frac{\partial}{\partial\beta}\ln_q \ Z_q \ , \tag{32}$$

where Z_q is defined through

$$\frac{(Z_q)^{1-q} - 1}{1-q} = \frac{(\bar{Z}_q)^{1-q} - 1}{1-q} - \beta U_q \ , \tag{33}$$

or, more compactly,

$$\ln_q \ Z_q = \ln_q \ \bar{Z}_q - \beta U_q \ . \tag{34}$$

At this stage it is convenient to discuss *thermodynamic stability* for the present canonical ensemble. In other words, we desire to check that small fluctuations of the energy do not modify the macroscopic state of the system at equilibrium. For this to be so, S_q must be a concave function of U_q (typically $\partial^2 S_q/\partial U_q^2 < 0$) if $q > 0$, and a convex function (typically $\partial^2 S_q/\partial U_q^2 > 0$) if $q < 0$. But

$$\frac{\partial}{\partial U_q}\frac{\partial S_q}{\partial U_q} = \frac{\partial}{\partial U_q}\frac{1}{T} = -\frac{1}{T^2}\frac{\partial T}{\partial U_q} = -\frac{1}{T^2 C_q} \ , \tag{35}$$

where we have used the fact that $C_q \equiv T\partial S_q/\partial T = \partial U_q/\partial T$. Thermodynamic stability is therefore guaranteed if $C_q/q \geq 0$, for all q and all Hamiltonians (characterized by the spectra $\{\epsilon_i\}$). da Silva *et al.* [75] have recently proved the following relation (in quantum notation for brevity):

$$\frac{C_q}{qk} = \frac{\beta^2(\bar{Z}_q)^{3(q-1)}\mathrm{Tr}\{\rho[\rho^{q-1}(\mathcal{H} - U_q)]^2\}}{1 + 2q(q-1)\beta^2(\bar{Z}_q)^{4(q-1)}\mathrm{Tr}\{\rho[\rho^{q-1}(\mathcal{H} - U_q)]^2\}} \ . \tag{36}$$

Consequently , if $q \geq 1$ or $q < 0$, $C_q/q \geq 0$ as desired. The situation is more complex for $0 < q < 1$. A general proof is missing for this case. However, the analysis of some particular examples suggests a scenario which is quite satisfactory.

In [42,75,76] a one-body problem, namely when the energy spectrum is given by $\epsilon_n = an^r$ ($a > 0$; $r > 0$; $n = 0, 1, 2, ...$; $r = 1$ corresponds to a harmonic oscillator; $r = 2$ corresponds to a particle confined in a infinitely deep square well), has been discussed. In the *classical* limit when n can be considered as a continuous variable (and sums are to be replaced by integrals), the following result has been obtained:

$$\frac{C_q}{qk} \propto (\frac{kT}{a})^{\frac{1-q}{r-1+q}} \tag{37}$$

with a nonnegative proportionality coefficient. Consequently, in all cases, $C_q/q \geq 0$, as desired.

In the *quantum* case, it has been shown [75] that the interval $0 < q < 1$ can be separated in two cases, namely $0 < q < q^*$ (with $q^* < 1$), and $q^* \leq q < 1$. In the latter, $C_q \geq 0$ as desired. In the former, regions of kT/a exist for which formally C_q would be negative. However, fortunately enough, as conjectured in [42,56] and illustrated in [77], a Maxwell-like equal-area construction takes place in such a manner that $C_q \geq 0$ for *all* values of kT/a and *all* values of $q \in (0,\ q^*)$. Indeed, two branches can appear in F_q versus T, but the *lowest* one (which is therefore the physically relevant one) has the desired curvature! The general proof for the interval $0 < q < 1$ would of course be very welcome; in the meanwhile, everything we are aware of at the present moment points to a generic thermodynamical stability for the canonical ensemble. Moreover, it might well generically happen in the thermodynamic limit ($N \to \infty$) that discontinuities (in value or derivatives) in the thermal dependance of the specific heat become gradually washed out while N increases (see [78] for such examples).

Let us now make an important remark. If we take out as factors, in both numerator and denominator of Eq. (26), the quantity $\left[1 + (1-q)\beta U_q / \sum_{j=1}^{W} (p_j)^q\right]$, and then cancel them, we obtain

$$p_i(\beta) = \frac{[1-(1-q)\beta'\epsilon_i]^{\frac{1}{1-q}}}{Z'_q} \qquad \left(Z'_q \equiv \sum_{j=1}^{W} [1-(1-q)\beta'\epsilon_j]^{\frac{1}{1-q}}\right) \tag{38}$$

with

$$\beta' = \frac{\beta}{\sum_{j=1}^{W} (p_j)^q + \ (1-q)\beta U_q} \qquad (T' \equiv 1/(k\beta'))\ , \tag{39}$$

where β' is an increasing function of β [77].

Let us now comment on the all important question of the connection between experimental numbers (those provided by measurements), and the quantities that appear in the theory. The definition of the internal energy U_q, and consistently of $A_q \equiv \langle\langle A \rangle\rangle_q$ associated with an arbitrary observable A, suggests that it is A_q the mathematical object to be identified with the numerical value provided by the experimental measure. Later on, we come back onto this crucial point.

At this point let us make some observations about the set of *escort probabilities* [79] $\{P_i^{(q)}\}$ defined through

$$P_i^{(q)} \equiv \frac{p_i^q}{\sum_{j=1}^W p_j^q} \quad (\sum_{i=1}^W P_i^{(q)} = 1) \tag{40}$$

from which follows the inverse relation

$$p_i = \frac{[P_i^{(q)}]^{\frac{1}{q}}}{\sum_{j=1}^W [P_j^{(q)}]^{\frac{1}{q}}} \ , \tag{41}$$

hence

$$\sum_{i=1}^W p_i^q = \frac{1}{[\sum_{i=1}^W [P_i^{(q)}]^{\frac{1}{q}}]^q} \ . \tag{42}$$

The $W = 2$ illustration of $P_i^{(q)}$ is shown in Fig. 1.

As anticipated, $q < 1$ ($q > 1$) privileges the *rare* (*frequent*) events. Eqs. (40) and (41) have, within the present formalism, a role somehow analogous to the direct and inverse Lorentz transformations in Special Relativity (see [80] and references therein). These transformations have an interesting structure. Let us mention a few of their features. If we consider a set of nonvanishing probabilities $\{p_i\}$ ($p_i > 0, \ \forall i$) associated with a set of W possibilities (or an infinite countable set, i.e., $W \to \infty$) and a nonvanishing real number q, we can define the transformation $\mathcal{T}_q$ as follows:

$$\mathcal{T}_q(\{p_i\}) \equiv \left\{ \frac{p_i^q}{\sum_{j=1}^W p_j^q} \right\} \ . \tag{43}$$

$\mathcal{T}_q$ transforms the set $\{p_i\}$ into the set $\{P_i^{(q)}\}$. We can easily verify that (i) $\mathcal{T}_q^{-1} = \mathcal{T}_{1/q}$, (ii) $\mathcal{T}_q \mathcal{T}_{q'} = \mathcal{T}_{q'} \mathcal{T}_q = \mathcal{T}_{qq'}$, and (iii) $\mathcal{T}_1$ is the identity element; in other words, the set of transformations $\{\mathcal{T}_q\}$ exhibits the structure of a commutative Abelian group. Furthermore, if the $\{p_i\}$ are ordered in such a way that $p_1 \geq p_2 \geq p_3 \geq ... \geq p_W > 0$, $\mathcal{T}_q$ preserves (inverts) the ordering if $q > 0$ ($q < 0$). A trivial corollary is that $\mathcal{T}_q$ preserves equiprobability ($p_1 = p_2 = p_3 = ... = 1/W$) for any value of q. A full and rigorous study of the mathematical properties associated with these transformations is missing and would be very welcome.

Let us now focus on the expectation values. We notice that O_q becomes an usual mean value when expressed in terms of the probabilities $\{P_i^{(q)}\}$, i. e.,

$$O_q \equiv \frac{\sum_{i=1}^W p_i^q O_i}{\sum_{j=1}^W p_j^q} = \sum_{i=1}^W P_i^{(q)} O_i \ , \tag{44}$$

and

$$\sum_{i=1}^W P_i^{(q)} \epsilon_i = U_q \ . \tag{45}$$

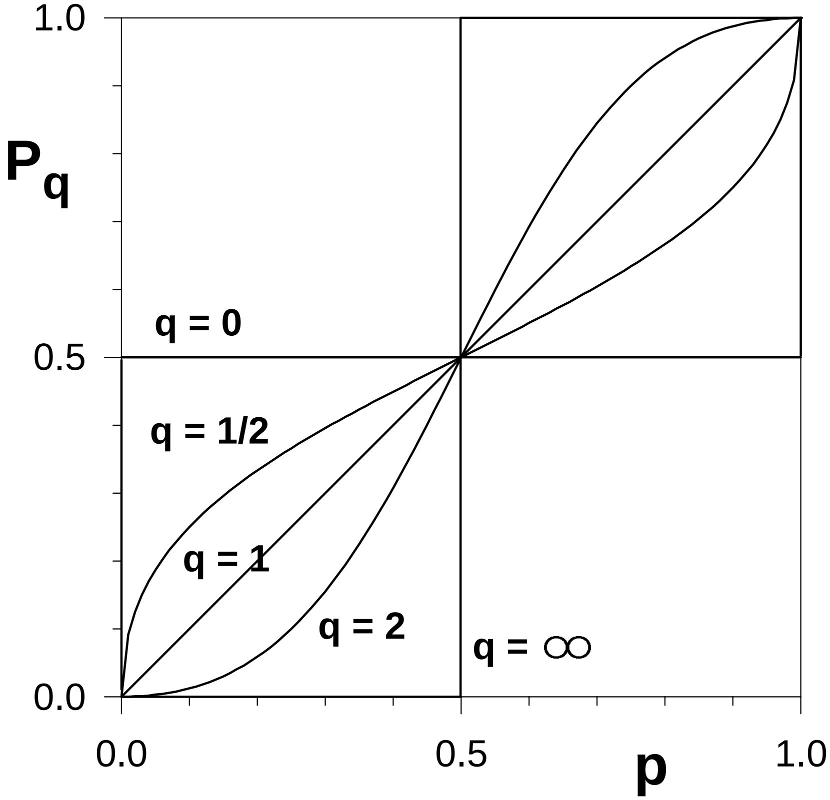

Fig. 1. $W = 2$ illustration of the *escort* probabilities: $P^{(q)} = \frac{p^q}{p^q+(1-p)^q}$.

The final equilibrium distribution reads

$$P_i^{(q)} = \frac{[1-(1-q)\beta'\epsilon_i]^{\frac{q}{1-q}}}{\sum_{k=1}^{W}[1-(1-q)\beta'\epsilon_k]^{\frac{q}{1-q}}} \ . \tag{46}$$

If the energy spectrum $\{\epsilon_i\}$ is associated with the set of *degeneracies* $\{g_i\}$, then the above probability leads to the following one (associated with the *level* ϵ_i and not the *state* i)

$$P(\epsilon_i) = \frac{g_i[1-(1-q)\beta'\epsilon_i]^{\frac{q}{1-q}}}{\sum_{all\ levels} g_k[1-(1-q)\beta'\epsilon_k]^{\frac{q}{1-q}}} \ . \tag{47}$$

If the energy spectrum $\{\epsilon_i\}$ is so dense that can practically be considered as a continuum, then the discrete degeneracies yield the function *density of states*

$g(\epsilon)$, hence

$$P(\epsilon) = \frac{g(\epsilon)[1-(1-q)\beta'\epsilon]^{\frac{q}{1-q}}}{\int d\epsilon' \; g(\epsilon')[1-(1-q)\beta'\epsilon']^{\frac{q}{1-q}}} \; . \tag{48}$$

The density of states is of course to be calculated for every specific Hamiltonian (given the boundary conditions). For instance, for a d-dimensional ideal gas of particles or quasiparticles, it is given [81] by $g(\epsilon) \propto \epsilon^{\frac{d}{r}-1}$, where r is the exponent characterizing the energy spectrum $\epsilon \propto K^r$ where K is the wavevector (e.g., $r = 1$ corresponds to the harmonic oscillator, $r = 2$ corresponds to a nonrelativistic particle in an infinitely high square well, etc.). In Figs. 2 and 3 we see typical energy distributions for the particular case of a constant density of states. Of course, the $q = 1$ case reproduces the celebrated Boltzmann factor. Notice the cut-off for $q < 1$ and the long algebraic tail for $q > 1$.

Since we have optimized the entropy, all the above considerations refer, strictly speaking, to *thermodynamic equilibrium.* The word *thermodynamic* makes allusion to "very large" ($N \to \infty$, where N is the number of microscopic particles of the physical system). The word *equilibrium* makes allusion to asymptotically large times ($t \to \infty$ limit) (assuming a stationary state is eventually achieved). The question arises: *which of them first?* Indeed, although both possibilities clearly deserve the denomination "thermodynamic equilibrium", nonuniform convergences might be involved in such a way that $\lim_{N\to\infty} \lim_{t\to\infty}$ could differ from $\lim_{t\to\infty} \lim_{N\to\infty}$. To illustrate this situation, let us imagine a classical Hamiltonian system including two-body interactions decaying at long distances as $1/r^\alpha$ in a d-dimensional space, with $\alpha \geq 0$ (we also assume that the potential presents no nonintegrable singularities, typically at the origin). If $\alpha > d$ the interactions are essentially *short-* ranged, the two limits just mentioned are basically interchangeable, and the prescriptions of standard statistical mechanics and thermodynamics are valid, thus yielding *finite* values for all the physically relevant quantities. In particular, the Boltzmann factor certainly describes reality, as very well known. But, if $0 \leq \alpha \leq d$, nonextensivity is expected to emerge, the order of the above limits becomes important because of nonuniform convergence, and the situation is certainly expected to be more subtle. More precisely, a crossover (between $q \neq 1$ and $q = 1$ behaviors) is expected to occur at $t = \tau(N)$. If $\lim_{N\to\infty} \tau(N) = \infty$, then we would indeed have *two (or even more) different and equally legitimate states of thermodynamic equilibrium*, instead of the familiar unique state. The conjecture is illustrated in Fig. 4. We could of course reserve the expression "thermodynamical equilibrium" for the *global* extrema of the appropriate thermodynamical energy. However, if the system is going to remain practically for ever in a *local* extremum, the distinction becomes physically artificial.

Since we are discussing thermodynamical equilibrium, it is relevant to say a few words on the present status of knowledge concerning the so-called 0^{th} principle of thermodynamics, or in other words, what happens with the *transitivity* of thermodynamical equilibrium between systems if $q \neq 1$? This important question is far from being transparent; it has already been addressed [82,83] though, in our opinion, only preliminarily. However, after the instructive illus-

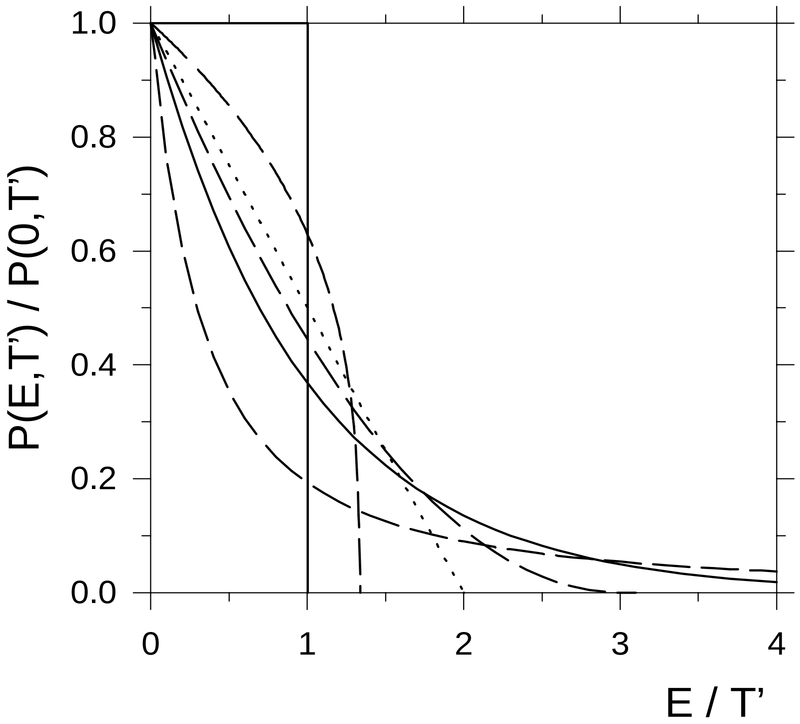

Fig. 2. Generalization (Eq. (48)) of the Boltzmann factor (recovered for $q = 1$) as function of the energy E at a given renormalized temperature T', assuming a constant density of states. From top to bottom at low energies: $q = 0,\ 1/4,\ 1/2,\ 2/3,\ 1,\ 3,\ \infty$ (the vertical line at $E/T' = 1$ belongs to the limiting $q = 0$ distribution; the $q \to \infty$ distribution collapses on the ordinate). All $q > 1$ curves have a $(T'/E)^{q/(q-1)}$ tail; all $q < 1$ curves have a cut-off at $E/T' = 1/(1 - q)$.

tration recently provided by Abe [84], a plausible scenario starts emerging. Let us assume a composed isolated Hamiltonian system $A + B$ (microcanonical ensemble) such that (in some sense to be further qualified) we consider (i) $\mathcal{H}(A+B) \sim \mathcal{H}(A) + \mathcal{H}(B)$, and (using quantum notation) (ii) $\rho(A+B) \sim \rho(A) \otimes \rho(B)$ (i.e., A and B are essentially independent in the sense of the theory of probabilities). We shall *also* assume that (iii) A and B are in thermal equilibrium, i.e., their energy distributions are essentially given by Eqs. (26) and (27). These three assumptions seem at first sight incompatible, since the power-law of a sum does not coincide with the product of the power-laws. In other words, the simultane-

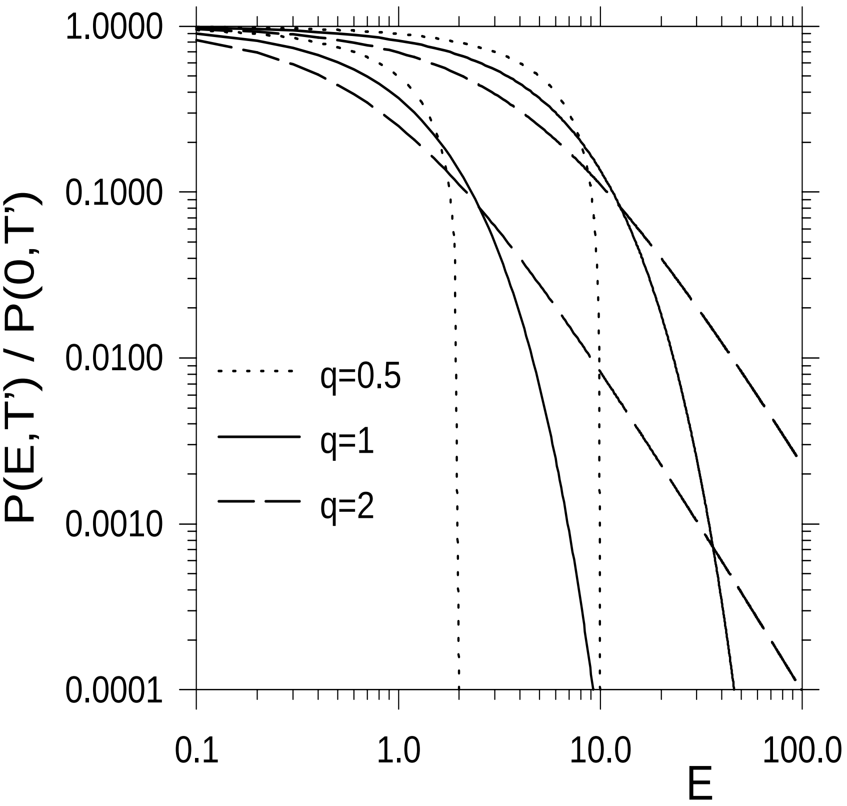

Fig. 3. Log-log plot of some cases like those of Fig. 2 ($T' = 1,\ 5$ for each value of q).

ous demand of the three hypothesis seems to lead to only one type of statistics, namely the $q = 1$ statistics. For this reason, it is stated in [82] that the present generalized statistics is incompatible with the 0^{th} principle, hence with thermodynamics. We believe this standpoint might be too narrow; it might well be that $q = 1$ is sufficient *but not necessary* for the compatibility of the above three hypothesis. Indeed, the thermodynamic limit ($N \to \infty$) might play a crucial role in the problem, as elegantly illustrated on a simple example by Abe [84]! This limit can be very subtle: for instance, several evidences are available [76,78,85] which show that the $N \to \infty$ and $q \to 1$ can be not commutative. (Since at $t \to \infty$ for fixed N we expect $q = 1$, this non commutativity might be directly related to the previously mentioned $(N, t) \to (\infty, \infty)$ noncommutativity). Let us now proceed with our argument. The three above assumptions imply that

$$U_q(A + B) \sim U_q(A) + U_q(B) \tag{49}$$

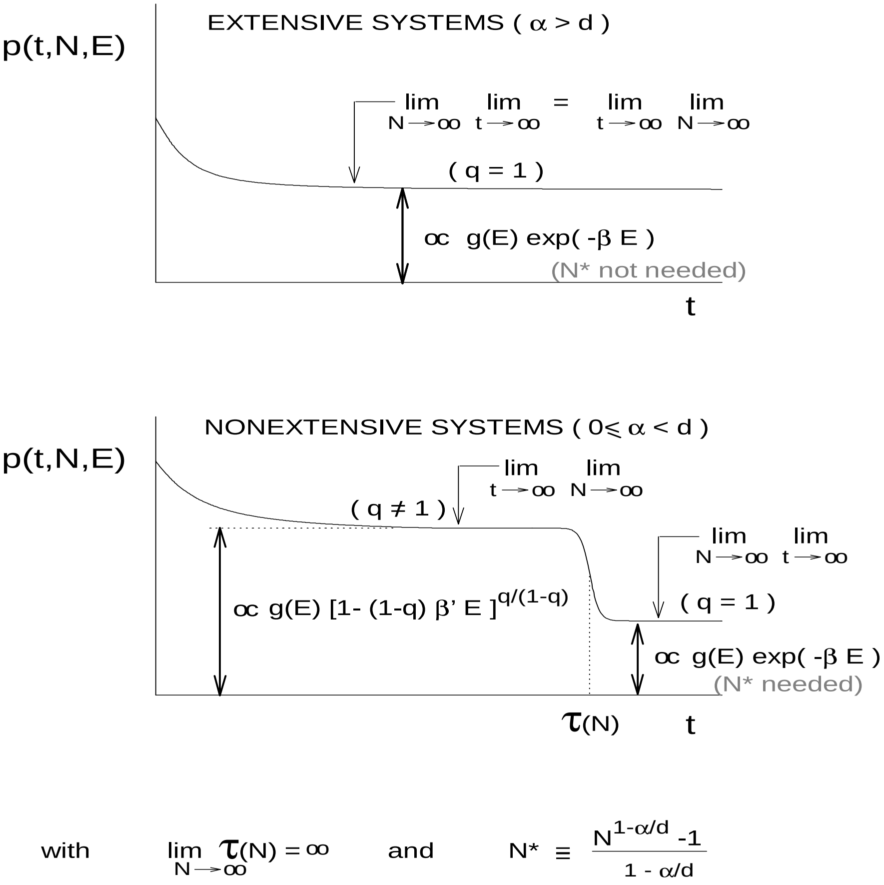

Fig. 4. One of the central conjectures of the present work, assuming a Hamiltonian system which includes two-body (attractive) interactions which, at long distances, decay as $r^{-\alpha}$. Only one long-standing macrostate is expected for $\alpha/d > 1$. More than one long-standing macrostates are expected for $0 \leq \alpha/d \leq 1$. The crossover at $t = \tau$ is expected to be slower than indicated in the figure (for space reasons). We are assuming that $\lim_{N\to\infty} \tau(N) = \infty$. The rescaling factor $\tilde{N} \equiv N^* + 1 = [N^{1-\alpha/d} - (\alpha/d)]/[1 - (\alpha/d)]$ is (is not) necessary if $0 \leq \alpha/d \leq 1$ ($\alpha/d > 1$).

and (taking $k = 1$ in Eq. (4))

$$S_q(A+B) \sim S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B) \tag{50}$$

which can be rewritten as

$$\begin{aligned} &\frac{\ln[1+(1-q)S_q(A+B)]}{1-q} \\ &\sim \frac{\ln[1+(1-q)S_q(A)]}{1-q} + \frac{\ln[1+(1-q)S_q(B)]}{1-q} \,. \end{aligned} \tag{51}$$

Since the system is isolated and at equilibrium, $U_q(A+B)$ and $S_q(A+B)$ are constants hence, by differentiating Eqs. (49) and (51), we obtain

$$\delta U_q(A) \sim -\delta U_q(B) \tag{52}$$

and

$$\frac{\delta S_q(A)}{\mathrm{Tr}[\rho(A)]^q} \sim -\frac{\delta S_q(B)}{\mathrm{Tr}[\rho(B)]^q} \ , \tag{53}$$

where we have used the definition of S_q. By dividing one by the other these two equations and using that $\partial S_q/\partial U_q = 1/T$ we straightforwardly obtain that

$$T_q(A) = T_q(B) \ , \tag{54}$$

where

$$T_q \equiv T \ \mathrm{Tr}\rho^q \ . \tag{55}$$

Also Rajagopal (see Section 6 of [86]) has obtained and commented Eq. (54) [87], as well as the analogous equalities for chemical potentials and pressures.

If our three primary hypothesis turn out to indeed be simultaneously compatible under some circumstances (presumably in the $t \to \infty$ limit of the $N \to \infty$ limit for long-range-interacting Hamiltonian systems), this generalized equality will play the role of 0^{th} principle. If we have in thermal contact systems with different entropic indexes, say q_A and q_B, it seems plausible that at equilibrium we have something like

$$T_{q_A}(A) = T_{q_B}(B) \quad (\forall (q_A, q_B)) \ . \tag{56}$$

See also [88]. The present result implies that, for positive temperatures, T_q is larger, equal or smaller than $T_1 \equiv T$ if q is smaller, equal or larger than unity (we remind that, say for $q > 1$, $\rho^q < \rho$ hence $\mathrm{Tr}\rho^q < \mathrm{Tr}\rho = 1$). Moreover, to measure temperatures T_q of all kinds of systems a standard thermometer should suffice. It is clear that the reader must be well aware of the speculative grounds on which we have discussed this important thermodynamic criterion. The subject is still unclear and can easily generate controversy. However, it is no doubt suggestive the fact that what appears in the equilibrium distribution (see Eq. (26)) precisely is T_q. Also, some suggestive evidences do exist for the above scenario. Indeed, at least three examples are available in the literature where values of q larger than unity seem to be accompanied by an effective temperature (presumably playing the role of T_q) which is *below* T. These three examples are: (i) Fig. 2(b) of [89] (remark that, while N increases, a finite, N-independent and time-independent temperature is emerging; for the infinitely-ranged model used in the paper we expect a long-tailed energy distribution, i.e., $q > 1$); (ii) in the fitting of the experimental data shown in [35], values of q *above* unity appear together with temperatures *below* those that would provide any Hagedorn-like ($q = 1$) fitting; (iii) in the diffusion of a quark in plasma, it has been recently obtained [39] $q > 1$ and a temperature *below* that of the thermal bath. In spite of these evidences, one must be extremely cautious in such a delicate matter. Just to play the devil's advocate, one could for instance wonder how correct the numerical value attributed in [89] to the temperature T is. Indeed, in that molecular dynamics work, to each degree of kinetic freedom the value $\frac{1}{2}kT$ has been associated by the authors. This association is obviously correct for any classical Hamiltonian system, *as long as the velocities distribution is Maxwellian*. However, for the metastable state of a long-range-interacting system the velocities distribution are more likely to be *non Maxwellian*! (possibly not even Gaussian). Under these circumstances, and even if the linearity with T was maintained, the corresponding factor would *not* be 1/2! (see [90]).

A wealth of works has shown that the above described nonextensive statistical mechanics retains much of the formal structure of the standard theory. Indeed, many important properties have been shown to be *q-invariant*. Among them, it is mandatory to mention
(i) the Legendre transformations structure of thermodynamics [41,42];
(ii) the H-theorem (macroscopic time irreversibility), more precisely, that, in the presence of some irreversible, master-equation-like physical evolution, $dS_q/dt \geq 0,\ = 0$ and ≤ 0 if $q > 0,\ = 0$ and < 0, respectively, the equalities holding for equilibrium [91,92];
(iii) the Ehrenfest theorem (correspondence principle between classical and quantum mechanics) [43];
(iv) the Onsager reciprocity theorem (microscopic time reversibility) [93,94];
(v) the Kramers and Wannier relations (causality) [94];
(vi) the factorization of the likelihood function (Einstein' 1910 reversal of Boltzmann's formula) [49]; more precisely,

$$W_q(A+B) = W_q(A)\ W_q(B) \qquad (\forall q) \tag{57}$$

if A and B are independent, the likelihood function being given by

$$W_q(\{p_i\}) \propto e_q^{S_q(\{p_i\})}; \tag{58}$$

(vii) the Bogolyubov inequality [95]; more precisely,

$$F_q^{(0)} + \langle\langle \mathcal{H} - \mathcal{H}_0 \rangle\rangle_q^{(0)} \geq F_q, \tag{59}$$

where F_q is the exact free energy associated with the Hamiltonian $\mathcal{H}$ we want to solve, and $F_q^{(0)}$ is the free energy associated with the variational Hamiltonian $\mathcal{H}_0$ we have adopted to discuss $\mathcal{H}$, otherwise unsolvable;
(viii) thermodynamic stability (i.e., a definite sign for the specific heat: $C_q/q \geq 0$) [96,75];
(ix) classical equipartition theorem (in particular, $\langle\langle$*total kinetic energy*$\rangle\rangle_q = NdkT_q/2$) and virial theorem [97,98];
(x) the Pesin equality [99].

In contrast with the above quantities and properties, which are q-invariant, some others *do* depend on q, such as
(i) the specific heat [100];
(ii) the magnetic susceptibility [101];

(iii) the fluctuation-dissipation theorem (of which the two previous properties can be considered as particular cases) [101];
(iv) the Chapman-Enskog expansion, the Navier-Stokes equations and related transport coefficients [102];
(v) the Vlasov equation [103,104];
(vi) the Langevin, Fokker-Planck and Lindblad equations [105–109];
(vii) stochastic resonance [110];
(viii) the mutual information or Kullback-Leibler entropy [92,111];
(ix) the Lie-Trotter formula [112].

A remark is necessary with regard to both sets just mentioned. Indeed, these properties have in fact been studied, whenever applicable, mostly within unnormalized q-expectation values for the constraints, rather than within the normalized ones that we are using herein. Nevertheless, in principle they still hold because they have been established for fixed β, which, through Eq. (39), implies fixed β'. However, the proofs using *normalized* q-expectation values should be checked case by case. Various of these checks can be found in Ref. [86], entirely written in terms of normalized q-expectation values.

Finally, let us mention various important theoretical tools which enable the thermostatistical discussion of complex nonextensive systems, and which are now available (within the unnormalized and/or normalized versions for the q-expectation values) for arbitrary q. We refer to
(i) Linear response theory [94];
(ii) Perturbation expansion [113];
(iii) Variational method (based on the Bogoliubov inequality) [113];
(iv) Many-body Green functions [114];
(v) Path integral and Bloch equation [115], as well as related properties [116];
(vi) Dynamical themostatting for the canonical ensemble [117];
(vii) Simulated annealing and related optimization, Monte Carlo and Molecular dynamics techniques [118–129];
(viii) Information theory and related issues (see [43,74,130,131] and references therein);
(ix) Entropic lower and upper bounds [132–134] (related to Heinberg uncertainty principle);
(x) Quantum statistics [135] and those associated with the Gentile and the Haldane exclusion statistics [136,137]. In particular, Fermi-Dirac and Bose-Einstein (escort) distributions could be generalizable as follows

$$n_{\mathbf{k}} = \frac{1}{[1+(q-1)\beta(\epsilon_{\mathbf{k}}-\mu]^{\frac{q}{q-1}} \pm 1} \,, \tag{60}$$

where $\mathbf{k}$ is the wave vector, β and μ are effective inverse temperature and chemical potential respectively, and ± respectively correspond to fermions and bosons. The degree of validity of this expression needs to be further clarified in at least three points: (a) It has been originally deduced [135] using a factorization which is in principle valid only for $q \approx 1$ (but which perhaps becomes valid for arbitrary q in the $N \to \infty$ limit; see also [138]); (b) It has been deduced for an *ideal* gas, for which the possible need for $q \neq 1$ is far from transparent; (c) Its rededuction in the framework of normalized q-expectation values is needed. In spite of these fragilities, and interestingly enough, it has in its favor an impressively good fitting of high temperature experimental data obtained in electron-positron collisions [35].

3 Theoretical Evidence and Connections

3.1 Lévy-Type Anomalous Diffusion

An enormous amount of phenomena in Nature follow the Gaussian distribution: measurement error distributions, height and weight distributions in biological individuals of given species, Brownian motion of particles in fluids, Maxwell-Boltzmann distribution of particle velocities in a variety of systems, noise distribution in uncountable electronic devices, energy fluctuations at thermal equilibrium of many systems, to only mention a few. Why is it so? Or, equivalently, what is their (thermo)statistical foundation? This fundamental problem has already been addressed, particularly by Montroll, and satisfactorily answered (see [25] and references therein). The answer basically relies onto two pillars, namely the BG entropy and the standard central limit theorem. However, the Gaussian is *not* the only ubiquitous distribution: we also similarly observe Lévy distributions (in micelles [139], supercooled laser [140], fluid motion [141], wandering albatrosses [142], heart beating [143], turbulence [144], DNA [145], financial data [146–148], among many others). So, once again, what is the (thermo)statistical foundation of their ubiquity? This relevant question has also been addressed, once again by Montroll and collaborators [25] among others. In this case however, a satisfactory answer has been missing for a long time. The first successful step toward (what we believe to be) the solution was performed in 1994 by Alemany and Zanette [149], who showed that the generalized entropic form S_q was able to provide a *power*-law (instead of the *exponential*-law associated with Gaussians) decrease at long distances. Many other works followed along the same lines [150,151]. In [151] it was exhibited how the Lévy-Gnedenko central limit theorem (see, for instance, [152] and references therein) also plays a crucial role by transforming, through successive iterations of the jumps, the power- law obtained from optimization of S_q into the specific power-law appearing in Lévy distributions. Summarizing, in complete analogy with the above mentioned Gaussian case (and which is recovered in the more powerful present formalism as the $q = 1$ particular case), the answer once again relies onto two pillars, which now are the generalized entropy S_q and the Lévy-Gnedenko central limit theorem.

The arguments have been very recently re-worked out [90] on the basis of the normalized q-expectation values introduced in [42]. These are the results that we briefly recall here.

Let us write S_q as follows:

$$S_q[p(x)] = k\frac{1 - \int_{-\infty}^{\infty} \frac{dx}{\sigma}\, [\sigma\, p(x)]^q}{q-1} \,, \tag{61}$$

where x is the distance of one jump, and $\sigma > 0$ is the characteristic length of the problem. We optimize (maximize if $q > 0$, and minimize if $q < 0$) S_q with the norm constraint $\int_{-\infty}^{\infty} dx\ p(x) = 1$, as well as with the constraint

$$\langle\langle x^2 \rangle\rangle_q \equiv \frac{\int_{-\infty}^{\infty} dx\ x^2\ [p(x)]^q}{\int_{-\infty}^{\infty} dx\ [p(x)]^q} = \sigma^2 \ . \tag{62}$$

We straightforwardly obtain the following one-jump distribution.
If $q > 1$:

$$p_q(x) = \frac{1}{\sigma} \Big[\frac{q-1}{\pi\ (3-q)}\Big]^{1/2} \frac{\Gamma(\frac{1}{q-1})}{\Gamma(\frac{3-q}{2(q-1)})} \frac{1}{\big[1 + \frac{q-1}{3-q}\ \frac{x^2}{\sigma^2}\big]^{1/(q-1)}} \ . \tag{63}$$

If $q = 1$:

$$p_q(x) = \frac{1}{\sigma} \Big[\frac{1}{2\pi}\Big]^{1/2}\ e^{-(x/\sigma)^2/2} \ . \tag{64}$$

If $q < 1$:

$$p_q(x) = \frac{1}{\sigma} \Big[\frac{1-q}{\pi\ (3-q)}\Big]^{1/2} \frac{\Gamma(\frac{5-3q}{2(1-q)})}{\Gamma(\frac{2-q}{1-q})} \Big[1 - \frac{1-q}{3-q}\ \frac{x^2}{\sigma^2}\Big]^{1/(1-q)} \tag{65}$$

if $|x| < \sigma[(3-q)/(1-q)]^{1/2}$ and zero otherwise.

We see that the support of $p_q(x)$ is compact if $q \in (-\infty, 1)$, an exponential behavior is obtained if $q = 1$, and a power-law tail is obtained if $q > 1$ (with $p_q(x) \propto (\sigma/x)^{2/(q-1)}$ in the limit $|x|/\sigma \to \infty$). Also, we can check that $\langle\langle x^2 \rangle\rangle_1 = \langle x^2 \rangle_1 = \int_{-\infty}^{\infty} dx\ x^2\ p_q(x)$ is finite if $q < 5/3$ and diverges if $5/3 \leq q \leq 3$ (the norm constraint cannot be satisfied if $q \geq 3$). Finally, let us mention that the Gaussian ($q = 1$) solution is recovered in both limits $q \to 1 + 0$ and $q \to 1 - 0$ by using the $q > 1$ and the $q < 1$ solutions respectively. This family of solutions is illustrated in Fig. 5.

We focus now the N-jump distribution $p_q(x, N) = p_q(x) * p_q(x) * ... * p_q(x)$ (N-folded convolution product). If $q < 5/3$, the standard central limit theorem applies, hence, in the limit $N \to \infty$, we have

$$p_q(x, N) \sim \frac{1}{\sigma} \Big[\frac{5-3q}{2\pi(3-q)N}\Big]^{1/2} \exp\Big(-\frac{5-3q}{2(3-q)N} \frac{x^2}{\sigma^2}\Big) \ , \tag{66}$$

i.e., the attractor in the distribution space is a Gaussian, consequently we have *normal* diffusion. If, however, $q > 5/3$, then what applies is the Lévy-Gnedenko central limit theorem, hence, in the limit $N \to \infty$, we have

$$p_q(N, x) \sim L_\gamma(x/N^{1/\gamma}) \ , \tag{67}$$

where L_γ is the Lévy distribution with index $\gamma < 2$ given by

$$\gamma = \frac{3-q}{q-1} \qquad (5/3 < q < 3) \tag{68}$$

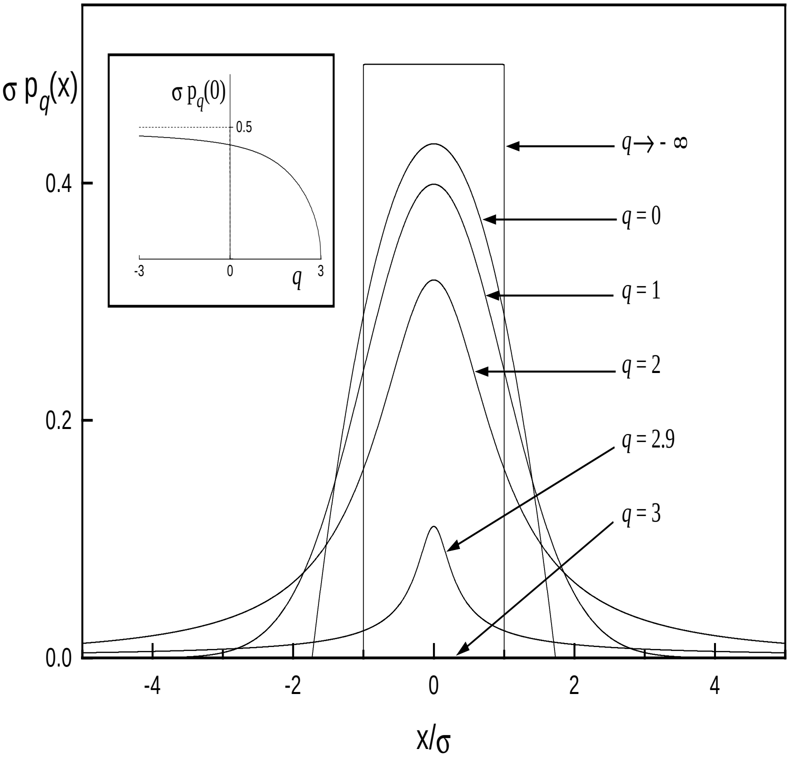

Fig. 5. The one-jump distributions $p_q(x)$ for typical values of q. The $q \to -\infty$ distribution is the uniform one in the interval $[-1, 1]$; $q = 1$ and $q = 2$ respectively correspond to Gaussian and Lorentzian distributions; the $q \to 3$ is completely flat. For $q < 1$ there is a cut-off at $|x|/\sigma = [(3 - q)/(1 - q)]^{1/2}$.

Through the Fourier transforms of both Eq. (66) and (67), we can characterize the *width* Δ_q (dimensionless diffusion coefficient) of $p_q(x, N)$. We obtain

$$\Delta_q \equiv \frac{3-q}{5-3q} \qquad (q < 5/3) \tag{69}$$

and

$$\Delta_q = \frac{2}{\pi^{1/2}} \left[\frac{q-1}{3-q}\right]^{\frac{3-q}{2(q-1)}} \Gamma\left[\frac{3q-5}{2(q-1)}\right] \qquad (5/3 < q < 3)\ . \tag{70}$$

These results are depicted in Fig. 6. This result should be measurable in specifically devised experiments. More details can be found in [90] and references therein. What we wish to retain in this short review is that the present formalism is capable of (thermo)statistically founding, in an unified and simple manner, both Gaussian and Lévy behaviors, very ubiquitous in Nature (respectively associated with normal diffusion and a certain type of anomalous superdiffusion).

The special values $q = 5/3$ and $q = 3$ correspond to the $d = 1$ case that we have considered here. In d dimensions, these values respectively become (see the article by Tsallis *et al.* in the book edited by Shlesinger *et al.* [25]) $q = (4+d)/(2+d)$ and $q = (2+d)/d$. These results together with some illustrative values obtained from experimental data are shown in Fig. 7 (from Zanette's article in the Brazilian Journal of Physics [150]).

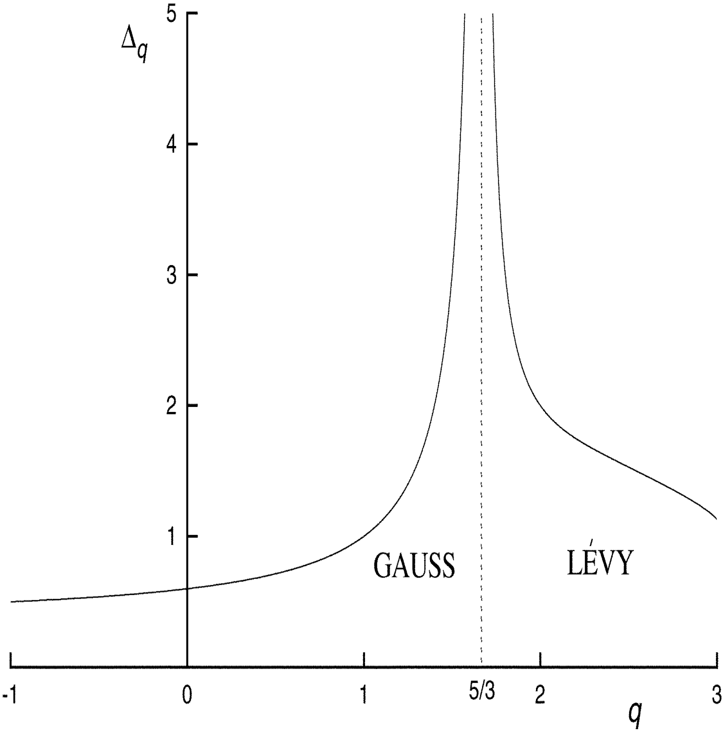

Fig. 6. The q-dependence of the dimensionless diffusion coefficient Δ_q (*width* of the properly scaled distribution $p_q(x,N)$ in the limit $N \to \infty$). In the limits $q \to 5/3-0$ and $q \to 5/3+0$ we respectively have $\Delta_q \sim [4/9]/[(5/3)-q]$ and $\Delta_q \sim [4/(9\pi^{1/2}]/[q-(5/3)]$; also, $\lim_{q\to 3} \Delta_q = 2/\pi^{1/2}$.

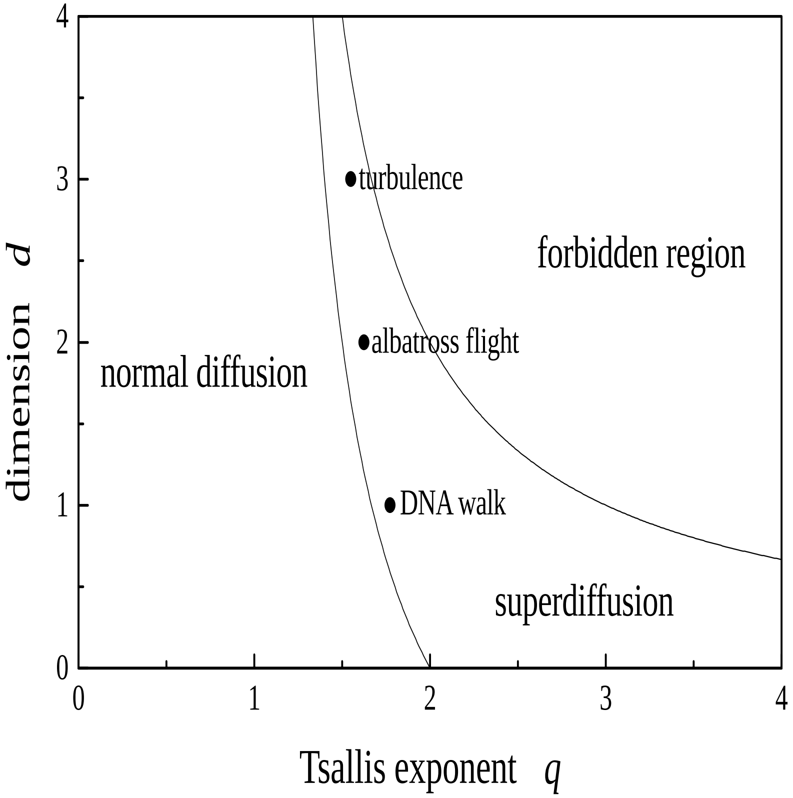

Fig. 7. From Zanette's article in Brazilian Journal of Physics [150]. The relevant regions in the (q, d) space for Lévy-like anomalous diffusion. The special points in the superdiffusive region correspond to experimental measurements: [144] for turbulence, [142] for the albatross flight and [145] for the DNA.

3.2 Correlated-Type Anomalous Diffusion

There are some phenomena exhibiting anomalous (super and sub) diffusion of a type which differs from the one discussed in the previous subsection. We refer to the so called *correlated*-type of diffusion. We consider here a quite large class of them, namely those associated with the following generalized, Fokker-Planck-like equation [107]:

$$\frac{\partial}{\partial t}[p(x,t)]^{\mu} = -\frac{\partial}{\partial x}\{F(x)[p(x,t)]^{\mu}\} + D\frac{\partial^2}{x^2}[p(x,t)]^{\nu} \ , \tag{71}$$

where $(\mu,\nu)\in\mathcal{R}^2$, D is a dimensionless diffusion-like constant, $F(x)\equiv -dV/dx$ is a dimensionless external force (drift) associated with a potential $V(x)$, and (x,t) is a dimensionless $1+1$ space-time. If $\mu=1$, we can interpret $p(x,t)$ as a probability distribution since $\int dx\; p(x,t)=1, \quad \forall t$ can be satisfied. If $\mu\neq 1$, then $p(x,t)$ must be seen as a density function. The word "correlated" is frequently used in this context due to the fact that $D(\partial^2/\partial x^2)[p(x,t)]^\nu = (\partial/\partial x)\{D\nu[p(x,t)]^{\nu-1}\,(\partial/\partial x)\,p(x,t)\}$, i.e., an effective diffusion emerges, for $\nu\neq 1$, which depends on $p(x,t)$ itself, a feature which is natural in the presence of correlations. The $\mu=1$ particular case of this nonlinear equation is commonly denominated "Porous medium equation", and corresponds to a variety of physical situations (see [107] and references therein for several examples; see also [153]).

The first connection of Eq. (71) with the present nonextensive statistical mechanics was established in 1995 by Plastino and Plastino [106]. They considered a particular case, namely $\mu=1$ and $F(x)=-k_2x$ with $k_2>0$ (so called Uhlenbeck-Ornstein processes), and found an exact solution which has the form of Eq. (63-65). Their work was generalized in [107] where arbitrary μ and $F(x)=k_1-k_2x$ were considered. The explicit exact solution of Eq. (71), for *all values* of (x,t), was once again found by proposing an Ansatz of the form of Eqs. (63-65), i.e., the form which optimizes S_q with the associated simple constraints. This form eventually turns out to be the Barenblatt one. By introducing this Ansatz into Eq. (71) we can verify, after some tedious but rather elementary algebra, that an exact solution is given by [107]

$$p_q(x,t)=\frac{\{1-(1-q)\beta(t)[x-x_M(t)]^2\}^{1/(1-q)}}{Z_q(t)}\;, \tag{72}$$

where

$$q=1+\mu-\nu \tag{73}$$

and

$$\frac{\beta(t)}{\beta(0)}=\left[\frac{Z_q(0)}{Z_q(t)}\right]^{2\mu} \tag{74}$$

with

$$Z_q(t)=Z_q(0)\left[\left(1-\frac{1}{K_2}\right)e^{-t/\tau}+\frac{1}{K_2}\right]^{1/(\mu+\nu)}, \tag{75}$$

$$K_2\equiv\frac{k_2}{2\nu D\beta(0)[Z_q(0)]^{\mu-\nu}}\;, \tag{76}$$

$$x_M(t)=\frac{k_1}{k_2}+\left[x_M(0)-\frac{k_1}{k_2}\right]\;e^{-k_2t}\;, \tag{77}$$

and

$$\tau\equiv\frac{\mu}{k_2(\mu+\nu)}\;. \tag{78}$$

An extreme case of this class of solutions would be when at $t=0$ we have a Dirac delta (i.e., $p_q(x,0)=\delta(x)$); this case corresponds to the limit $\beta(0)\to\infty$.

Summarizing, by using the form which optimizes S_q, it has been possible to find the physically relevant solution of a *nonlinear* equation in partial derivatives with *integer* derivatives. It can be shown [154] that the problem that was solved in the previous subsection corresponds to a *linear* equation in partial derivatives but with *fractional* derivatives. We believe that we are allowed to say that an unusual mathematical versatility has been observed, within the present nonextensive formalism, in this couple of nontrivial examples of anomalous diffusion. The discussion of an unifying equation which simultaneously is *nonlinear* and has *fractional* derivatives remains to be done. Let us finally mention that equations similar to Eq. (71) but also including either an absorbing term [155] or a nonlinear reaction term [156] have as well been exactly solved recently. The simultaneous inclusion of *both* terms also remains to be done.

3.3 Charm Quark Diffusion in Quark–Gluon Plasma

In a recent theoretical work, Walton and Rafelski [39] used a Fokker-Planck equation and perturbative Quantum Chromodynamics techniques to calculate the energy dependence of the energy loss per unit distance traveled by a quark inside a quark-gluon plasma. They applied their theory to a charm quark with mass $m_c = 1.5$ GeV interacting with thermal gluons at $T_b = 500$ MeV (b stands for *bath*). In their phenomenological approach, q and T_T were left as fitting parameters (of course, $T_T = T_b$ if $q = 1$). Their results are exhibited in Fig. 8, and their best fitting was obtained for $q = 1.114$ and $T_T = 135.2$ MeV. We notice that, as other analogous systems, the fact that $q > 1$ comes together with $T_T < T_b$. We also notice that a small discrepancy of q from unity can carry substantial modifications in measurable physical quantities (the case of the solar neutrino problem is even more remakable in this sense).

3.4 Self-Gravitating Systems

It is known since long [19] that self-gravitating systems exhibit anomalous thermodynamics. This comes from a two-folded cause: the short-distance singularity of the gravitational potential, as well as its long-distance tail. The first one is in some sense less severe since, on physical grounds, a cut-off is ultimately expected to exist (for instance, due to quantum effects). The second one is heavier in thermodynamical consequences. Indeed, if we have, say, a classical d-dimensional system with a two-body interacting potential which decays, at long distances, like $r^{-\alpha}$, it is long known [15] that standard, extensive thermodynamics are perfectly well defined if $\alpha/d > 1$ (from now on referred to as *short-range* interactions). But nonextensivity is expected to emerge if $0 \leq \alpha/d \leq 1$ (*long-range* interactions). Newtonian gravitation ($d = 3$ and $\alpha = 1$) clearly belongs to the anomalous class. (In fact, d-dimensional gravitation, i.e., $\alpha = d - 2$ belongs to the anomalous class for $d \geq 2$). One of the known thermodynamical anomalies associated with Newtonian gravitation is the fact that, within BG statistical mechanics, it is not possible to have the total mass, the total energy and the total entropy *simultaneously* finite, as physically desirable. This was, in fact, the

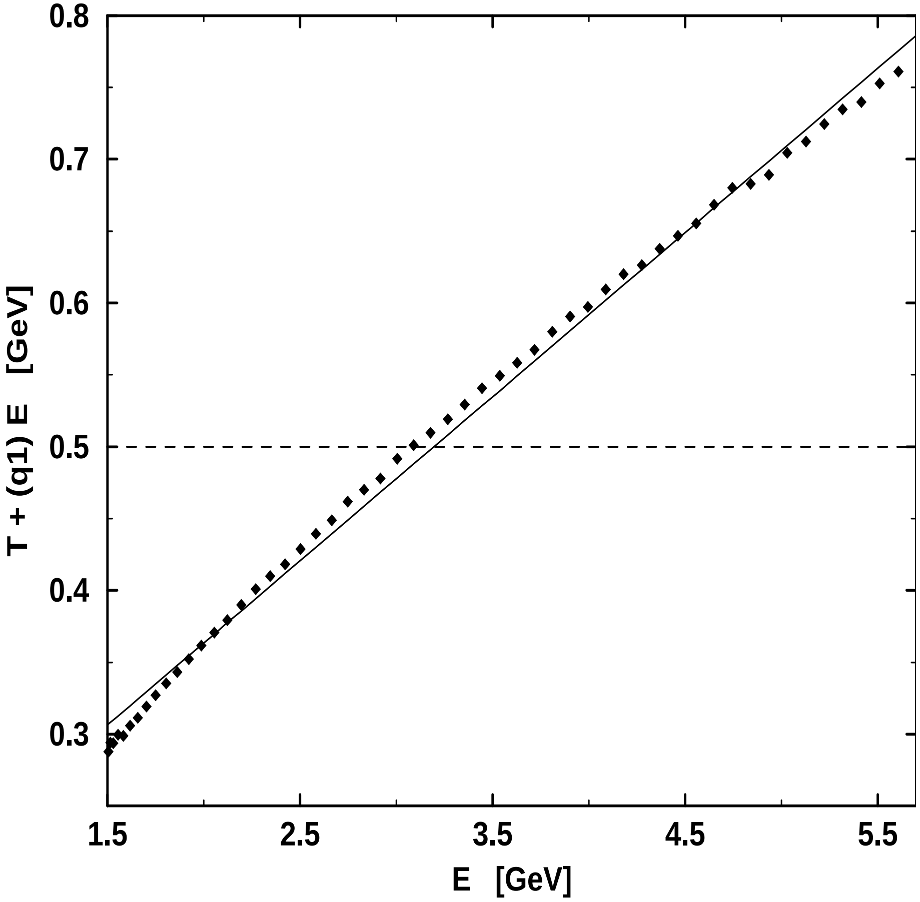

Fig. 8. Quark energy dependence of the loss of energy per unit distance traveled by the quark (with a 1.5 MeV mass) in the gluon plasma at a temperature of 500 MeV (diamonds: perturbative QCD calculation; dashed line: $q = 1$ scenario (Boltzmann) with a 500 MeV temperature; solid line: $q = 1.114$ scenario with a 135.2 MeV temperature).

first physical application of nonextensive statistics. Indeed, without entering in details now, Plastino and Plastino [103] were the first to show, in 1993, that this physically desirable situation *can* be achieved if we allow q to sufficiently differ from unity ! In fact, it can be shown (by considering the Vlasov equation in d-dimensional Schuster spheres) that the problem becomes a mathematically well posed one if $q < q^*$, where the critical value q^* is given [103] (see also [104]) by

$$q^* = \frac{8 - (d-2)^2}{8 - (d-2)^2 + 2(d-2)} \ . \tag{79}$$

For $d = 3$ we recover the (by now quite well known) 7/9 value. Also, we notice that $D = 2$ implies $q^* = 1$, which is very satisfactory since it is known that $d < 2$ gravitation is tractable within standard thermodynamics.

The present formalism has in fact been applied to a variety of astrophysical ([157] and references therein) and cosmological [158] self-gravitating systems. In fact, similar types of anomalies have already been encountered in long-range Ising ferromagnets [20], and are generically known since many decades [13,15,16,18].

3.5 Zipf–Mandelbrot Law

The problem we focus here first appeared in Linguistics. However, its relevance is quite broad, as it will soon become clear. Suppose we take a given text, say Cervante's *Don Quijote*, and order all of its words from the most to the less frequent; we refer to the ordered position of a given word as its *rank* R (low rank means high frequency ω of appearance in the text, and high rank means low frequency). Zipf [159] discovered that, in this as well as in a variety of similar problems, the following law is satisfied:

$$\omega = A\, R^{-\xi} \qquad (Zipf\ law)\ , \tag{80}$$

where $A > 0$ and $\xi > 0$ are constants (the value initially adopted by Zipf was $\xi = 1$). Later on, Mandelbrot [146] suggested that such behavior was reflecting a kind of fractality hidden in the problem; moreover, he suggested how the Zipf law could be numerically improved:

$$\omega = \frac{A}{(D+R)^{\xi}} \qquad (Zipf - Mandelbrot\ law;\ D > 0)\ . \tag{81}$$

This expression has been useful in a variety of analysis, and has provided satisfactory fittings with experimental data. The connection we wish to mention here is that in 1997 Denisov [160] showed that, by extending (to arbitrary q) the well known Sinai-Bowen- Ruelle thermodynamical formalism of symbolic dynamics (i.e., by considering S_q instead of S_1), the Zipf-Mandelbrot law can be *deduced.* He obtained

$$\xi = \frac{1}{q-1}, \tag{82}$$

hence

$$\omega \propto \frac{1}{[1+(q-1)R/d]^{1/(q-1)}} \qquad (q > 1)\ , \tag{83}$$

where $d \equiv (q-1)D > 0$. Clearly, to make the discussion complete, a model would be welcome, which would provide quantities such as q and d. Nevertheless, Denisov's arguments have the deep interest of explicitly exhibiting that the Zipf-Mandelbrot law can be seen as having a nonextensive foundation. Fittings with experimental data will be shown later on in connection with the citations of scientific papers.

3.6 Theory of Financial Decisions: Risk Aversion

An important problem in the theory of financial decisions is how to take into account extremely relevant phenomena such as the risk aversion human beings (hence financial operators) quite frequently feel. This kind of problem has, since long, been extensively studied by Tversky [161] and co-workers. The situation can be illustrated as follows. What do you prefer, to earn 85,000 dollars or to play a game in which you have 0.15 probability of earning nothing and 0.85 probability of earning 100,000 dollars ? *You are allowed to participate in the game only once!* In fact, most people prefer to take the money. The problem of course is the fact that the expectation value for the gain is one and the same (more precisely 85,000 dollars) for *both* choices, and therefore this mathematical tool does not reflect reality ! The same problem appears if one expects to loose 85,000 and the chance is given for playing a game in which, if you win, you pay nothing, but, if you loose, you pay 100,000 dollars. In this case, most people choose to play. So, the experimental facts are that most human beings are *risk-averse* when they expect to gain, and *risk-seeking* when they expect to loose ! The problem is how to put this into mathematics. One traditional manner is to make standard averages, not on the *gain*, but on the *utility*, defined as a *nonlinear* function of the gain. We wish to show here a different possibility, which places the nonlinearity on the probabilities themselves. (Nothing forbids of course to consider nonlinearities on both the utility and the probabilities).

Following along the lines of [162], let us introduce, for the above gain problem, normalized q-expectation values as follows:

$$\langle\langle gain \rangle\rangle_1^{take\ the\ money} = 85,000 \tag{84}$$

and

$$\begin{aligned}\langle\langle gain \rangle\rangle_q^{play\ the\ game} &= \frac{100,000 \times 0.85^q + 0 \times 0.15^q}{0.85^q + 0.15^q} \\ &= \frac{100,000 \times 0.85^q}{0.85^q + 0.15^q} \ .\end{aligned} \tag{85}$$

Since most people would prefer the money, this means that most people have $q < 1$ for this particular decision problem.

For the loss problem we have:

$$\langle\langle gain \rangle\rangle_1^{pay\ the\ money} = -85,000 \tag{86}$$

and

$$\begin{aligned}\langle\langle gain \rangle\rangle_q^{play\ the\ game} &= \frac{-100,000 \times 0.85^q + 0 \times 0.15^q}{0.85^q + 0.15^q} \\ &= \frac{-100,000 \times 0.85^q}{0.85^q + 0.15^q} \ .\end{aligned} \tag{87}$$

Since in this case most people would prefer to play, this means that, consistently with the previous result, most people have $q < 1$ for the particular decision problem we are considering now. In some sense, we have some epistemological progress! Indeed, the statement "most people have (for this type of amount of money) $q < 1$", *unifies* the previous two *separate* statements concerning expectation to gain and expectation to loose.

Let us address now the following question: how can we measure the value of q associated with a particular individual ? We illustrate this interesting point with the example of the gain. The person is asked to choose between having V dollars or playing a game in which, if the person wins, the prize will be 100,000 dollars and, if the person looses, he (she) will receive nothing. As before, the person is informed that his (her) probability of winning is 0.85 (hence, the probability of loosing is 0.15). Then we keep gradually changing the value V and asking what is the preference. At a certain critical value, noted V_c, the person will change his (her) mind. Then, the value of q to be associated with that person, for that problem, is given by the following equality

$$\frac{100,000 \times 0.85^q}{0.85^q + 0.15^q} = V_c \ . \tag{88}$$

(See Fig. 9.) The ideally rational operator corresponds to $q = 1$. For this gain problem, the risk-averse operators correspond to $q < 1$, and the risk-seeking ones to $q > 1$.

This particular manner of formulating the problem is no doubt appealing. However, is it the only one along nonextensive lines? The answer is *no*. Let us be more specific. We can use *unnormalized* q-expectation values instead of the *normalized* ones we have just used. In this case Eqs. (84) and (85) are to be replaced by

$$\langle gain \rangle_1^{take\ the\ money} = 85,000 \tag{89}$$

and

$$\langle gain \rangle_q^{play\ the\ game} = 100,000 \times 0.85^q + 0 \times 0.15^q = 100,000 \times 0.85^q \tag{90}$$

and Eq. (88) is to be replaced by

$$100,000 \times 0.85^q = V_c \ . \tag{91}$$

If so formulated, the conclusion is that most people have $q > 1$ (instead of $q < 1$ obtained in the previous formulation). The analysis of the loss problem also provides for most people $q > 1$. Consequently, also in this formulation we have the benefit of unification of the gain and loss problems.

At this point we have to face an ambiguity: both criteria, respectively using normalized and unnormalized q-expectation values, unify the gain and the loss problems *but* the former attributes to most people values of $q < 1$, whereas the latter attributes $q > 1$! Which one is the correct one for this particular problem?

As an attempt to solve the ambiguity, let us address a different, though similar, problem. We propose to the candidate to choose to play with one of two boxes. He (she) is informed that in box A there are (exactly) 100 balls, 50 of them being red, the other 50 being white. The person will have to declare a

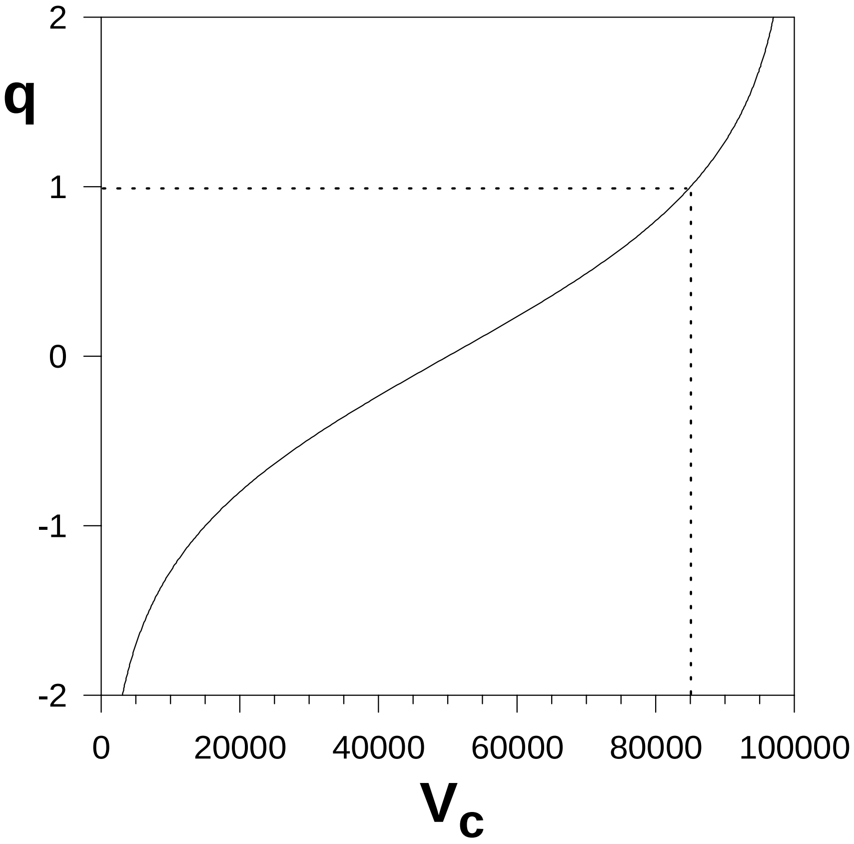

Fig. 9. The index q to be associated with a person whose critical value corresponding to Eq. (88) is V_c. People with $q < 1$ ($q > 1$) tend to avoid (seek) risks for that particular game. The case $q = 1$ corresponds to an ideally rational agent.

color, and then randomly take off a ball. If it has the chosen color, the person will earn 100 dollars. If it has the other color, the person will receive nothing. The person is also informed that in box B there are also (exactly) 100 balls, some are red, some are white, but nobody knows how many of each, though we do know that no other colors are in the box. As before, the person will have to declare a color, and then randomly take off a ball. As for the other box, if the ball has the chosen color, the person will receive 100 dollars; if it has the other color, he (she) will receive nothing. These are the two choices. *The person will be given only one opportunity for playing.* The person is now asked to choose *one* box for playing. The experimental outcome is that most people choose box A (possibly because their anxiety is smaller with regard to that particular box, because they have some supplementary information about it... though this information is completely useless!). Let us write down the associated expectation values. If we use the normalized ones we have

$$\langle\langle gain\rangle\rangle_1^{box\ A} = 100 \times 0.5 + 0 \times 0.5 = 50 \tag{92}$$

and

$$\langle\langle gain\rangle\rangle_q^{box\ B} = \frac{100 \times 0.5^q + 0 \times 0.5^q}{0.5^q + 0.5^q} = 50 \ \ (\forall q) \ . \tag{93}$$

Since both possibilities yield one and the same result for all values of q, this criterion is unable to reproduce reality, i.e., the fact that most people choose the box A. This is a consequence – regrettable in the present occasion! – of the fact that escort distributions preserve equiprobability.

If we use, instead, the unnormalized q-expectation values, we have

$$\langle gain\rangle_1^{box\ A} = 100 \times 0.5 + 0 \times 0.5 = 50 \tag{94}$$

and

$$\langle gain\rangle_q^{box\ B} = 100 \times 0.5^q + 0 \times 0.5^q = 50 \times 0.5^{q-1} \ . \tag{95}$$

Most people have, within the present model, the feeling that $\langle gain\rangle_1^{box\ A} > \langle gain\rangle_q^{box\ B}$, hence $q > 1$. So, this criterion not only is capable of differentiating the two boxes, but also unifies this game with the 85,000 game. Summarizing, the comparison of *unnormalized* q-expectation values yields the unified conclusion that "most people have $q > 1$" for *three* different games, namely the 85,000 gain, the 85,000 loss and the box games. For the box game, the value q of a given person can be "measured" by using

$$100 \times 0.5^q = V_c \times 0.5 \ , \tag{96}$$

where V_c is the value to be offered (instead of 100) if the person chooses to play with box A, whereas 100 is maintained if the person chooses to play with the other box.

We can formalize these games and the associated decisions as follows. The person is offered to choose, for playing a single time, between two games, one of them being perceived as "safe" (to be therefore calculated using $q = 1$), the other one perceived as "unsafe" (to be therefore calculated using $q \neq 1$). In the safe game the probability of wining is p_s and that of loosing $(1 - p_s)$, receiving respectively V^s_{gain} and V^s_{loss}. In the unsafe game, the probability of wining is p_u and that of loosing $(1-p_u)$, receiving respectively V^u_{gain} and V^u_{loss}. We then have

$$\langle gain\rangle_1^{(s)} = p_s \times V^s_{gain} + (1 - p_s) \times V^s_{loss} \tag{97}$$

and

$$\langle gain\rangle_q^{(u)} = p_u \times V^u_{gain} + (1 - p_u) \times V^u_{loss} \ , \tag{98}$$

and the critical point (which allows for the measure of q for a given person) satisfies

$$p_s \times V^s_{gain} + (1 - p_s) \times V^s_{loss} = p_u \times V^u_{gain} + (1 - p_u) \times V^u_{loss} \ . \tag{99}$$

All three games that we have discussed above are particular instances of these last equations. Summarizing, it appears that unnormalized q-expectation values can be used to modelize the fact that most people are risk-averse when expecting to gain something substancial (”A bird in the hand is worth two in the bush”, says the popular dictate), and risk seeking when expecting to loose. This situation corresponds to $q > 1$, hence $\sum_i p_i^q < 1$ (which we might call *subadditive probabilities*). A few people are the other way around, which corresponds to $q < 1$, hence $\sum_i p_i^q > 1$ (which we might call *superadditive probabilities*). It is clear that models for stock exchange can be formulated by using these remarks. Such an effort is presently in progress [163].

3.7 Physiology of Vision

Physiological perceptions such as the visual perception are since long known to focus upon *rare* events (e.g., a red spot on a white wall). Barlow [73], among others, has recurrently stressed our attention on the fact that, at the action decision level, the various possibilities should enter with a weight proportional to $-\ln p_i$, and *not* proportional to p_i, p_i being the a priori probability of occurrence of that particular event; indeed, $-\ln p_i$ diverges when $p_i \to 0$. He even argues that evolutionary arguments hold very well together with such hypothesis. To privilege rare events is precisely what happens, in the present formalism, whenever $q < 1$. Let us be more specific: if we consider the $0 < q << 1$ limit, we obtain [164]

$$S_q/k = \frac{1-\sum_{i=1}^W p_i^q}{q-1} \sim W - 1 + q[W - 1 - \sum_{i=1}^W (-\ln p_i)] \; , \tag{100}$$

$$\langle O \rangle_q \equiv \sum_{i=1}^W p_i^q \, O_i \sim \sum_{i=1}^W O_i - q \sum_{i=1}^W (-\ln p_i) \, O_i \; , \tag{101}$$

and

$$\begin{aligned} \langle\langle O \rangle\rangle_q &\equiv \frac{\sum_{i=1}^W p_i^q \, O_i}{\sum_{i=1}^W p_i^q} \\ &\sim \frac{\sum_{i=1}^W O_i}{W} \Big\{ 1 + q \Big[\frac{\sum_{i=1}^W (-\ln p_i)}{W} - \frac{\sum_{i=1}^W (-\ln p_i) \, O_i}{\sum_{i=1}^W O_i} \Big] \Big\} \; , \end{aligned} \tag{102}$$

where O is an arbitrary observable. Leaving aside several constant quantities that appear above, we immediately observe the prominent role which $-\ln p_i$ plays in these expressions. Consistently, the $q \to 0$ limit of the present formalism could well be of some utility in the theoretical analysis of a variety of physiological phenomena.

4 Experimental Evidence and Connections

4.1 $D = 2$ Turbulence in Pure-Electron Plasma

A few years ago, in 1994, Huang and Driscoll [24] exhibited some quite interesting nonneutral plasma experimental results obtained in pure-electron plasma (confined in a 20 cm long and 6 cm wide metallic cylindrical Penning trap with a 10^{-10} torr vacuum in its interior) in the presence of an external axial magnetic field (507 Gauss). In the interval 2-100 ms after every single electric shot (generating the electron plasma), it was observed a turbulent axisymmetric metaequilibrium state, the electronic density radial distribution of which was measured. Its average (over typically 100 shots) monotonically decreased with the radial distance, disappearing at some radius sensibly *smaller* than the radius of the container (i.e., a *cut-off* was observed). The experiment was recently redone [165] under slightly modified experimental conditions (a slow external rotation was imposed in such a way as to compensate the small energy dissipation existing in the plasma), and essentially the same metaequilibrium state was observed during lapses of time as long as 27 hours, or even longer! In addition to the 1994 experiment, the authors also proposed [24] a phenomenological theory trying to reproduce the experimentally observed profile. Their proposal consisted on the optimization, for a given model, of a functional of the electron density $\rho(r)$ under constraints, namely conservation of total mass, angular momentum and energy. They presented four different attempts. The first one (*Point Vortex Maximum Entropy*) consisted in optimizing, for a point vortex representation of the plasma, the BG entropy: it failed in reproducing the experimental data. The second attempt (*Fluid Maximum Entropy*) was essentially the same as the previous one, but using a fluid model for the plasma: the failure was even bigger. They assumed next that the problem possibly relied, not so much in the particular plasma model, but rather in the chosen functional to be optimized. In their third attempt (*Global Minimum Enstrophy*), they turned back to the point vortex model, but optimized the *enstrophy* instead of the BG entropy. The result was better than the two first attempts, but had the physically unacceptable feature of producing a *negative* electron density at sufficiently high radius. They then addressed their fourth attempt (*Restricted Minimum Enstrophy*), whose only difference with the third one was the fact of introducing an *out-of-the pocket* cut-off of the electron density at the proper value of the radius. This procedure was, finally, successful, and a very good first-approximation fitting was obtained! The effort done by Huang and Driscoll was, on top of the high merit of a remarkable experiment, extremely pedagogical and elucidating: the main theoretical problem was *not* the model, but rather the choice of the functional to be optimized, i.e., the *statistics*.

The next important step in this story was done by Boghosian. He realized in 1995 and published [104] in 1996 that the Huang and Driscoll fourth, successful attempt *precisely* corresponds to the optimization of S_q with $q = 1/2$! Indeed, by following the recipes of the present generalized thermostatistics, he re-obtained, for the electron density profile, the *same* differential equation produced within the Restricted Minimum Enstrophy phenomenological theory, with the supplementary bonus of *not* having to introduce, in an *ad hoc* manner, the necessary cut-off. Indeed, as already argued, all $q<1$ cases exhibit a cut-off intrinsic to the formalism, and the radial position of that cut-off nicely fits the experimental value.

The next step was performed in 1997 by Anteneodo and myself [166] (in fact, after related remarks by Boghosian himself). The Restricted Minimum Enstrophy theory is based on the enstrophy functional, which belongs to the general discussion of Casimir invariants; its form is in fact that of the order 2 Casimir invariant. Consequently, an epistemologically conservative theoretical viewpoint is to appreciate Boghosian's effort as just a formal interesting remark, with no real physical necessity. It happens, however, that, for $r \to r_c - 0$, ($r_c \equiv$ cut-off radius) the enstrophy theory yields $\rho(r) \propto (r_c - r)$ whereas the experimental data fit much better a *vanishing* derivative at r_c! We followed along Boghosian's lines and generalized his theory for arbitrary q [166]. We obtained the generalized differential equation for $\rho(r)$ and showed that $\rho(r) \propto (r_c - r)^{q/(1-q)}$. Consequently, the experimental data fit better for q *slightly above* $1/2$. This, together with the numerical solution of the differential equation, advanced $q \simeq 0.55$ as a better value for a satisfactory overall fitting. (Better fittings would probably demand for a model more sophisticated than the point vortex one used here). The conceptually important point of this discussion is that Casimir invariants are characterized by *integer* exponents (in $\rho(r)$), hence *none* of them can be related to a value of q close to 0.55.

The last step of this analysis was performed very recently by Anteneodo [167], and addressed the type of q-expectation values to be used. Indeed, the calculations above recalled [104,166] were done by using unnormalized q-expectation values. However, as already mentioned and used in the present review, it has been recently argued [42] that normalized q-expectation values should be used instead. It is therefore important to check that the present discussion and results for turbulence remain essentially invariant. This is now done [167], and it is this theory that we present in what follows.

The generalized entropy and associated constraints are given by

$$S_q[g] \equiv \frac{1}{q-1} \int (g - g^q) \mathrm{d}^2\mathbf{r} \ , \tag{103}$$

$$\begin{aligned}
&\int g \mathrm{d}^2\mathbf{r} = 1 \qquad\qquad (\textit{mass conservation}) \ , \\
&\frac{\int r^2 g^q \mathrm{d}^2\mathbf{r}}{\int g^q \mathrm{d}^2\mathbf{r}} = L_q \equiv L \quad (\textit{angular momentum conservation}) \ , \\
&\frac{-\frac{1}{2}\int \frac{\phi}{\phi^\star} g^q \mathrm{d}^2\mathbf{r}}{\int g^q \mathrm{d}^2\mathbf{r}} = U_q \equiv U \quad (\textit{energy conservation}) \ ,
\end{aligned} \tag{104}$$

where $g(r)$ is the probability distribution. Moreover, the scaled electrostatic potential

$$\frac{\phi(r)}{\phi^\star} \equiv \frac{\int g^q(r')G(\mathbf{r},\mathbf{r}')\mathrm{d}^2\mathbf{r}'}{\int g^q \mathrm{d}^2\mathbf{r}} \qquad \text{with } \nabla^2 G(\mathbf{r},\mathbf{r}') = 4\pi\delta(\mathbf{r}-\mathbf{r}'), \tag{105}$$

satisfies

$$\nabla^2 \frac{\phi}{\phi^\star} = 4\pi \frac{g^q}{\int g^q \mathrm{d}^2\mathbf{r}} \ . \tag{106}$$

The constrained optimization of $S_q[g]$ ($\delta(S_q - \alpha\int g\mathrm{d}^2\mathbf{r} - \lambda L_q - \beta U_q)$) now yields

$$\begin{aligned} &\frac{1-q\,g_q^{q-1}}{q-1} - \alpha - \frac{\lambda}{N}qr^2\,g_q^{q-1} + \frac{\beta}{N}q\frac{\phi_q}{\phi^\star}g_q^{q-1} \\ &+q(L\frac{\lambda}{N} - 2U\frac{\beta}{N})g_q^{q-1} = 0 \ , \end{aligned} \tag{107}$$

(where $\int g^q \mathrm{d}^2\mathbf{r} \equiv N$) or

$$\frac{g_q^{1-q} - q}{q-1} - \alpha q^{1-q} - \frac{\lambda}{N}qr^2 + \frac{\beta}{N}q\frac{\phi_q}{\phi^\star} + q(L\frac{\lambda}{N} - 2U\frac{\beta}{N}) = 0 \ , \tag{108}$$

or, taking the Laplacian of both sides,

$$[1+\alpha(1-q)]\frac{\nabla^2 g_q^{1-q}}{q-1} - 4\frac{\lambda}{N}q + 4\pi\frac{\beta}{N^2}qg_q^q = 0 \tag{109}$$

which can be rewritten as

$$g_q'' - q\frac{(g_q')^2}{g_q} + \frac{g_q'}{r} = g_q^q(B^\dagger g_q^q - A^\dagger) \ , \tag{110}$$

where $A^\dagger \equiv 4q\frac{\lambda}{N}/[1+\alpha(1-q)]$ and $B^\dagger \equiv 4\pi q\frac{\beta}{N^2}/[1+\alpha(1-q)]$.

Alternatively, identifying $\rho_q \equiv g_q^q/N$, we have

$$\rho_q'' - \frac{2q-1}{q}\frac{(\rho_q')^2}{\rho_q} + \frac{\rho_q'}{r} = q\rho_q^{\frac{2q-1}{q}}(B\rho_q - A) \ , \tag{111}$$

with $A \equiv A^\dagger N^{\frac{q-1}{q}}$ and $B \equiv B^\dagger N^{\frac{2q-1}{q}}$. This equation precisely is the one appearing in [166], which, for $q = 1/2$, recovers that of [104]. For any chosen q, the values of the parameters (A, B) are obtained by imposing the experimental values of total angular momentum and energy. This phenomenological theory has, therefore, only one fitting parameter (q). As said before, $q = 1/2$ exactly reproduces the Huang and Driscoll's Restricted Minimum Enstrophy profile. As already mentioned, the best overall fitting is, however, obtained for a value of q *slightly above* 1/2, like $q \approx 0.55$. From this point of view, the present formalism appears to us as the *only* satisfactory phenomenological theoretical approach available in the literature at the present time. It is but fair to mention here that, in a recent paper, Brands *et al.* [168] disagree with this standpoint. They claim to have proved that the nonextensive approach is unsatisfactory. However, they have incorrectly used it (more precisely, they have not used normalized q-expectation values, among other controversial statements), so their conclusions can be subject for doubt. A point which obviously remains to be dynamically explained is what is the microscopic reason for q having that particular value which fits the data.

4.2 Solar Neutrino Problem

As easily conceivable, the core of the Sun is a very complex and turbulent plasma, within which an enormous amount of nuclear reactions take place. Many of them constitute chains of nuclear reactions in which neutrinos are produced. For instance, the p-p chain is described in [169]. Through a quite complete analysis of the production of neutrinos within the so called *Solar Standard Model* (SSM), it is possible to predict the neutrino flux onto the Earth. However, the actual flux measured in a variety of underground laboratories (Gallex, Sage, Kamiokande, Super-Kamiokande, Homestake) roughly amounts to only half of the predicted value ! This problem is currently referred to as the "solar neutrino problem". Two nonexclusive sources of explanation of this enigmatic discrepancy are: (i) the possible neutrino oscillations, which would make that only part of the predicted value would be detectable on the Earth; (ii) the current use of the SSM might be incorrect because it uses BG thermal statistics, which could be inappropriate for the solar plasma. Clayton [29] was the first to address the second possibility, as far as 25 years ago! Indeed, he assumed an hypothetic distribution of energies essentially given by

$$p(E) \propto e^{-\beta E} \, e^{-\delta(\beta E)^2} \; . \tag{112}$$

The particular value $\delta = 0$ obviously recovers BG statistics. Clayton showed that a small value of δ ($\delta \simeq 0.01$) was enough to make the theory consistent with the experimental data that were available at that time. Quarati and co-workers remarked (preliminarily in 1996 [170], and in more refined calculations since then [157]) that, since the needed δ is very small, the Ansatz distribution could as well be the power-law one which appears in the present formalism. By identifying the first corrections (to BG) of both distributions, they obtained

$$\delta = \frac{1-q}{2} \; . \tag{113}$$

Consequently, values of q quite close to unity are enough to fit the solar neutrino discrepancy. Once again, we verify the extreme efficiency that modifications of the statistics can have. The reason for which it is so in this particular probem is eloquently shown in Fig. 10 (from Coraddu *et al.* in [157]).

4.3 Peculiar Velocities in Sc Galaxies

From the data obtained by the Cosmic Background Explorer (COBE), it has been possible to infer the distribution of peculiar velocities of certain groups

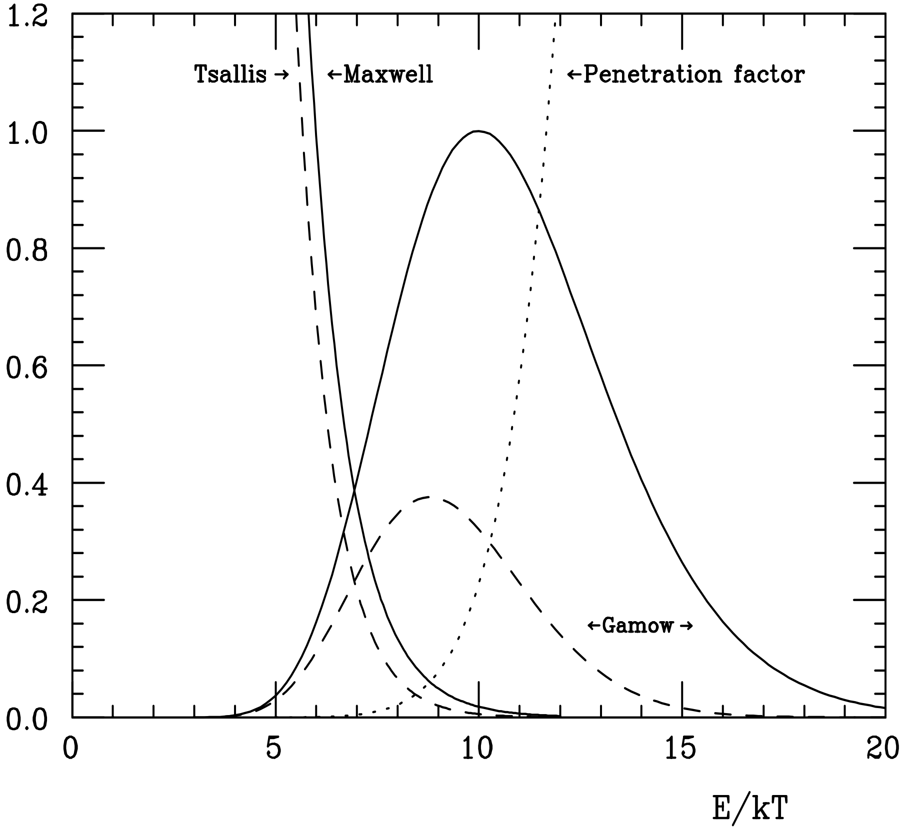

Fig. 10. From Coraddu *et al.* in [157]. The penetration factor independs from the statistics, hence from q. On the contrary, the thermal distribution of course depends on q, and the authors have illustrated the cases $q = 1$ (solid line) and $q = 0.98$ (dashed line); since only the tails of the distributions are relevant in this case, the distributions are here multiplied by a huge factor 400, to make the tiny effect visible. The Gamow peak, which effectively determines the flux of neutrinos, is given by the product of the penetration factor and the thermal distribution: We can see how different the Gamow peak is for $q = 1$ (solid line) and for $q = 0.98$ (dashed line)! This effect is therefore in principle capable of explaining sensible deffects in the number of neutrinos arriving on Earth.

of spiral (Sc) galaxies (we recall that by *peculiar* velocity we mean the residual velocity after the global universe expansion velocity has been subtracted). Bahcall and Oh [30] developed four theoretical attempts (namely Cold Dark Matter with $\Omega = 0.3$ and with $\Omega = 1.0$, Hot Dark Matter with $\Omega = 1.0$ and Primeval Barionic Isotropic with $\Omega = 0.3$). All the attempts were done within BG statistics. The less unsatisfactory fitting was obtained for the CDM model with $\Omega = 0.3$. In fact, all the attempts exhibit a long tail towards high velocities, whereas the experimental data show a pronounced cut-off at about 500 $Km\ s^{-1}$. It is relevant to mention that all the models that were used had several fitting parameters, and nevertheless could not get rid of the tail. A fitting was then advanced [171] using the present formalism with only two free parameters, one of them being q and the other one a characteristic velocity. The function that was used was the q-generalized Maxwell distribution, essentially corresponding to an ideal classical gas. The quality of the fitting is quite remarkable, far better than those corresponding to the already mentioned four attempts. Once again, one sees that modifications of the statistics can be sensibly more efficient than modifications of the model. A famous example along this line is provided by the completely different physics associated with a gas of free fermions or of free bosons, i.e., a Fermi-Dirac ideal gas or a Bose-Einstein ideal gas (same model but different statistics).

4.4 Nonlinear Inverse Bremsstrahlung Absorption in Low Pressure Argon Plasma

Liu *et al.* [31] provided in 1994 strong evidence of the existence of non-Maxwellian velocity distributions in a specific plasma experiment, where low pressure argon is exposed to pulsed discharges. During the afterglow, measurements of the inverse bremsstrahlung of intense microwaves is performed. The experimental setting is such that Coulombian collisions are dominant. The experimental data were fitted with the following flat-topped distribution:

$$f(v) \propto exp[-(v/v_m)^m] \tag{114}$$

with $m \geq 2$. Souza and myself [172] showed in 1997 that the same data can equally well be fitted with

$$f(v) \propto [1-(1-q)(v/v_q)^2]^{q/(1-q)} \tag{115}$$

with $q \geq 1$. In both fittings, the exponents m and q depend on the microwave power. In order to discriminate between the two fitting functions, quite precise and systematic experiments would be needed, in particular exploring the actual dependence of the results on the power.

4.5 Cosmic Microwave Background Radiation

The most accurate data concerning the cosmic microwave background radiation have been obtained with the FIRAS (Far-infrared absolute spectrophotometer) instrument in the COBE (Cosmic background explorer) satellite [173]. These data are known to follow, in the $2 - 20\ cm^{-1}$ region, Planck's black-body law. In 1995, Sa Barreto, Loh and myself [174], as well as Plastino, Plastino and Vucetich [175] (and several others since then), analyzed within what precision one is allowed to assume $q = 1$. The result that has systematically come out from these analyses is $|q-1| < 10^{-4}$. If new observations were performed in the future which would be say 10 times more precise than the available ones [173], this bound would be attained. Consequently, we would know better within what degree of confidence extensive thermostatistics can be used for this cosmological problem. If $q \neq 1$ turns out to be clearly confirmed, it is not excluded that we would have to revise our notions about the structure of space-time at the appropriate scales (possibly, Planck's length). It might come out that the physics at that level are better described by *finite-difference* equations than by *differential* equations! If so, the notion of a smooth and continuous space-time would probably have to be abandoned, as suspected by Einstein in his last years! [176].

4.6 Electron–Positron Collisions

The electron-positron annihilation into a virtual photon and the subsequent creation of a quark-antiquark pair provides the cleanest environment for the hadroproduction. Each of the two initial partons begins a complex cascade related to the strong-coupling long-distance regime of Quantum Chromodynamics. A partially successful global description of the hadroproduction has been provided through a thermodynamical equilibrium approach, mainly that of Hagedorn in 1965 [177]. This theory provides the following prediction:

$$\frac{1}{\sigma}\frac{d\sigma}{dp_T} \simeq c p_t^{3/2} \exp -p_t/T_0 \quad (p_t > T_0) \ , \tag{116}$$

where σ is the distribution of the transverse momenta p_t, T_0 is a characteristic temperature which Hagedorn predicts to be *independent* from the electron-positron collision energy W in the mass center referential, and c is a constant. The physical cause of this insensibility to W of T_0 would be the impossibility of heating the system above T_0; indeed, larger W would induce the creation of more mesons instead of further heating. This theory fits the data quite well for small W, say $W < 10$ *Gev*, but exhibits a pronounced failure for W increasing up to say 160 *Gev*. Very recently, Bediaga, Curado and Miranda [35] have used, along Hagedorn's lines, the present generalized statistics. To be more precise, Hagedorn arrived [177] at expression

$$\frac{1}{\sigma}\frac{d\sigma}{dp_t} = c p_t \int_0^\infty dp_l P(p_l) \tag{117}$$

and used $P(p_l) = \exp[-\frac{1}{T_0}(p_l^2 + p_t^2 + m^2)^{1/2}]$, where m is the rest mass of the particles. Solving this integral within the realistic assumption that $m, T_0 << |p_t|$, Hagedorn finally arrived to Eq. (116). What Bediaga *et al.* essentially did was to phenomenologically use $P(p_l) = [1 - \frac{1-q}{T_0}(p_l^2 + p_t^2 + m^2)^{1/2}]^{q/(1-q)}$, which recovers

Hagedorn's Boltzmann weight as the $q = 1$ particular case. The solution of this integral leads to an expression involving several hypergeometric functions, hence becoming a tractable problem. The results are indicated in Figs. 11 and 12. Remark that (i) q varies smoothly and monotonically with varying W (Hagedorn's theory is recovered in the $W \to 0$ limit), and (ii) $T_0 \simeq 0.12$ *Gev* (which is of the order of magnitude of the meson mass!) and practically independs from W as desirable from Hagedorn's arguments. Due to the considerable experimental span (almost 5 decades in the ordinate and 3 in the abcissa), these results can be considered as a strong evidence of the applicability of the nonextensive thermostatistics to specific anomalous systems. Further studies are welcome in order to enlighten the microscopic mechanism (e.g., increase of the effective range of nonmarkovian microscopic memory) which produces the phenomenological dependence of q on W.

The nonextensivity-like phenomenon observed in electron-positron annihilation also is observed in several others high-energy collisions, in particular involving heavy ions. For instance, Wilk and co-workers [37] have satisfactorily fitted an experimental [38] distribution of transverse momenta in $S - S$ collisions with $q = 1.015$. Also very recently, Alberico *et al.* [178] have proposed a possible nonextensive origin of the so-called *blue shift factor* in $Pb - Pb$ and similar collisions.

4.7 Emulsion Chamber Observation of Cosmic Rays

Cosmic rays can be observed by using a variety of detectors (such as Pb detectors; see [179] and references therein). Typically, showers of (clustered or individual) elementary particles appear which start at the so called vertex. These vertex are localized at various depths. The distribution of their depths can be measured (see [34] and references therein for the measurements done at the Mount Pamir lead chambers) . This distribution was recently fitted by Wilk and Wlodarcsyc [34] with the $q = 1.3$ function which emerges within the present formalism.

4.8 Reassociation of Heme–Ligands in Folded Proteins

In the folded conformational state, proteins might exhibit fractal effects. One such case might be the time evolution of the re-association of molecules that have been taken away from their natural positions. For instance, if O_2 molecules are dissociated, through light flashes, from their natural Fe positions in a heme protein and reach positions outside the heme pocket, they tend to start rebinding, and, for so doing, they might have to follow a fractal path, or be under the dynamical influence of fractal excitations (e.g., fractons). Anyhow, this re-association phenomenon has been lengthily studied by Frauenfelder *et al.* [180]. If we define $\xi \equiv N(t)/N(0)$ where $N(t)$ is the number of molecules that have not yet re-associated at time t, the $\xi(t)$ monotonically vanishes with t. The results obtained by photo-dissociating CO molecules from Sigma Type 2 sperm whale Myoglobin (Mb) dissolved in a glycerol-water solution are shown in Fig. 10. For

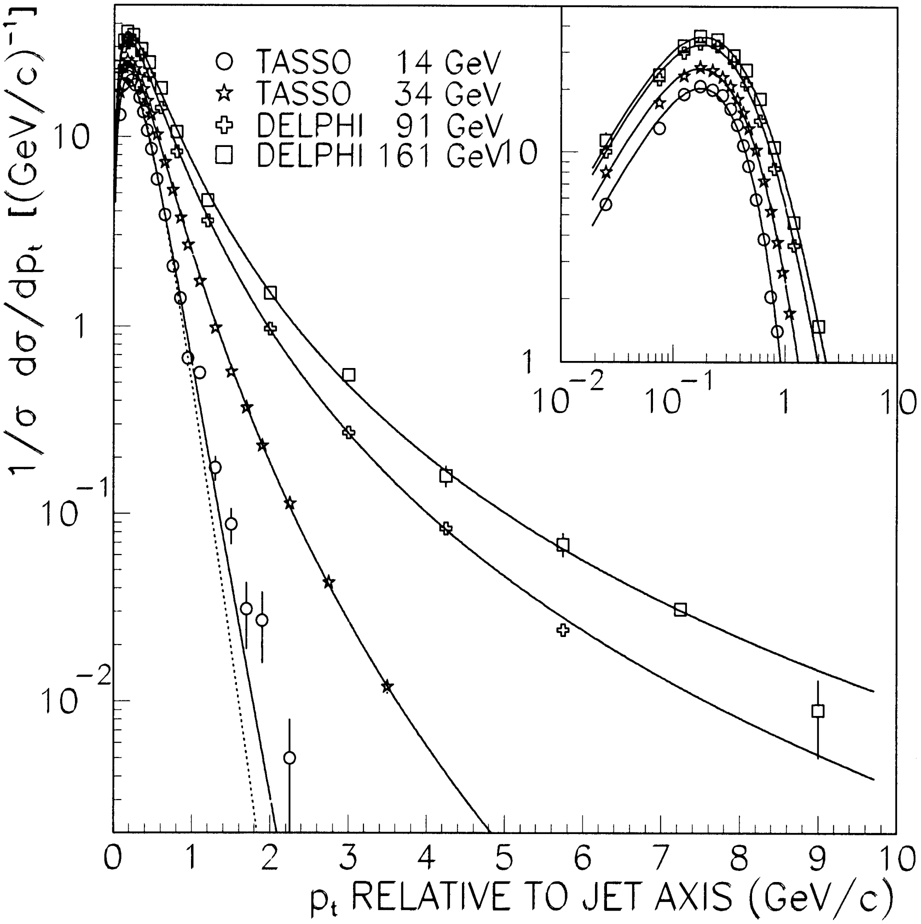

Fig. 11. Distribution of the transverse momenta p_T obtained in electron-positron frontal collisions of energy W varying from 14 to 161 *Gev*. The dotted line corresponds to $q = 1$ (i.e., a Hagedorn type of fitting as given by Eq. (116)) for all values of W. The solid lines correspond to $q \neq 1$ fittings.

times not too long, the experimental data have been fitted by Frauenfelder *et al.* [180] with

$$\xi = (1 + t/t_0)^{-n} \ , \tag{118}$$

where t_0 and n smoothly depend on the temperature T. Bemski, Mendes and myself [181] have argued that, within the generalized formalism, the following equation naturally appears:

$$\frac{d\xi}{dt} = -\lambda_q \, \xi^q \quad (\lambda_q \geq 0;\ q \geq 1) \ . \tag{119}$$

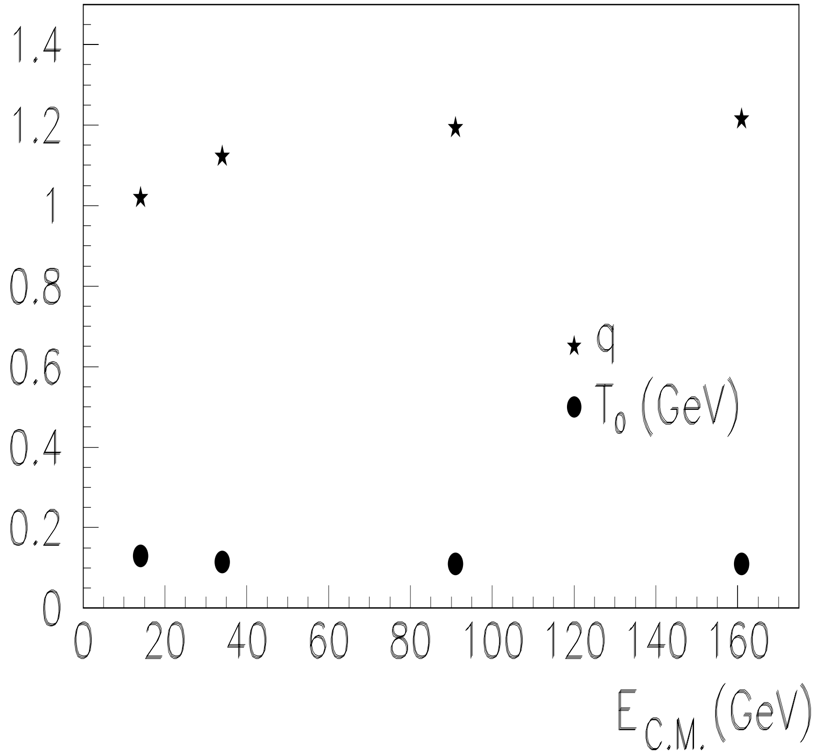

Fig. 12. The values of q and T_0 used in the fittings of Fig. 11. When W approaches zero, q approaches unity, i.e., Hagedorn's theory; T_0 is essentially insensitive to W, as physically desirable.

Its solution is given by

$$\xi = \frac{1}{[1 + (q-1)\lambda_q t]^{\frac{1}{q-1}}} \ . \tag{120}$$

This expression recovers, for $q = 1$, the usual exponential relaxation (i.e., $\xi = e^{-\lambda_1 t}$), and reproduces the Frauenfelder form through the identifications $1/(q-1) \equiv n$ and $1/[(q-1)\lambda_q] \equiv t_0$. Besides reobtaining the Frauenfelder empirical law, the present scheme allows for a better approximation if a crossover is admitted. More precisely, the above differential equation can be generalized as follows:

$$\frac{d\xi}{dt} = -\mu_r \, \xi^r - (\lambda_q - \mu_r) \, \xi^q \quad (r \leq q) \ . \tag{121}$$

The general solution involves [181] a hypergeometric function, hence it is tractable. The fitting is shown in Figs. 13 and 14. To better understand the physical regions involved in Eq. (121), we have exhibited in Fig. 15 two typical situations, namely corresponding to $r = 1$ and to $r > 1$.

A microscopic model which would justify the above phenomenological differential equation would be welcome. Such model would have the uneasy task of simultaneously recovering, in the experimental temperature range, the following behaviours:
(i) $1/(q-1) \propto T$;
(ii) $r(T) \approx 2$;
(iii) $\lambda_q \sim A_\lambda e^{B_\lambda T}$;
(iv) $\mu_r \sim A_\mu e^{B_\mu T}$ with $A_\mu << A_\lambda$ and $B_\mu \approx 2B_\lambda$.

4.9 Diffusion of Hydra Vulgaris

Upadhyaya *et al.* [182] are presently performing interesting experiments on *Hydra Vulgaris* (a cylindrical body column with inner and outer cells, respectively referred to as endodermal and ectodermal respectively) in physiological solution. The endodermal cells are more adhesive than the ectodermal ones. The authors have measured the velocity distribution $P(|V_y|)$ of the "vertical" component of the velocity during the diffusion of endodermal Hydra cells in an ectodermal aggregate. The results are presented in Fig. 16, where the velocity unit is $10^{-6} m/hour$ and the probability is represented by the histogram of the number of counts. These results were fitted with

$$P(|V_y|) = \frac{a}{(1+b|V_y|^2)^c} \tag{122}$$

with the values of (a, b, c) indicated in the figure. Through the identifications

$$a = P(0) \ , \quad b = (q-1)/V_0^2 \ , \quad c = \frac{q}{q-1} \ , \tag{123}$$

we precisely have the law which emerges within the present formalism, namely

$$P(|V_y|) = \frac{P(0)}{[1+(q-1)(V_y/V_0)^2]^{q/(q-1)}} \tag{124}$$

with $q = 1.53$. The next desirable step of course is to formulate a specific model for Hydra which would lead to this law, but this remains to be done.

4.10 Citations of Scientific Papers

An interesting study was recently done by Redner [183], in which the statistics of citations of scientific papers is focused. He exhibited the number $N(x)$ of papers which have been cited x times for two long series, namely one (6 716 198 citations of 783 339 papers) from the Institute of Scientific Information (ISI) and another one (351 872 citations of 24 296 papers) from the Physical Review

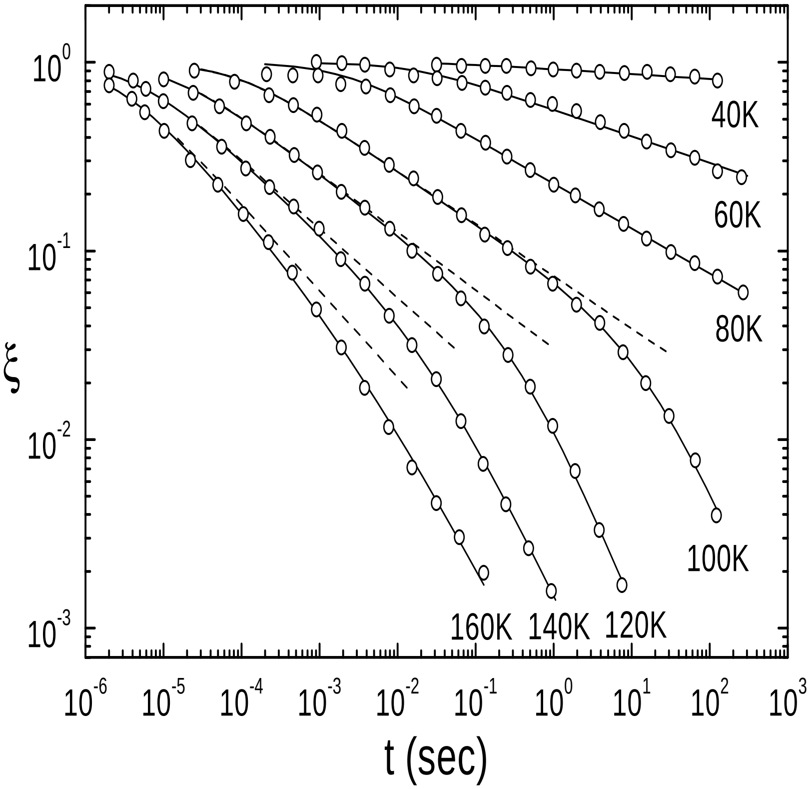

Fig. 13. Time evolution of $\xi \equiv N(t)/N(0)$ associated with *MbCO* in glycerol-water. Dots: experimental data. Dashed lines: fittings with Frauenfelder's empirical law (Eq. (118) or Eq. (120)). Solid lines: fittings with the solutions of Eq. (121) (see Fig. 14).

D (PRD). As expected, in both examples, $N(x)$ monotonically decreases with x. Redner fitted the (relatively) low-x data with a stretched exponential of the form

$$N(x) = N(0)\ e^{-(x/x_0)^{\beta}} \tag{125}$$

with $\beta = 0.44$ and 0.39 for the series ISI and PRD respectively. Also, he remarked that the large- x data exhibit a power law, namely close to $\propto 1/x^3$. He argues that this different functional behavior for low and large values of x must reflect different phenomenologies in these two regimes. In contrast with this viewpoint, Albuquerque and myself [184] argue that this is not necessarily so since the data

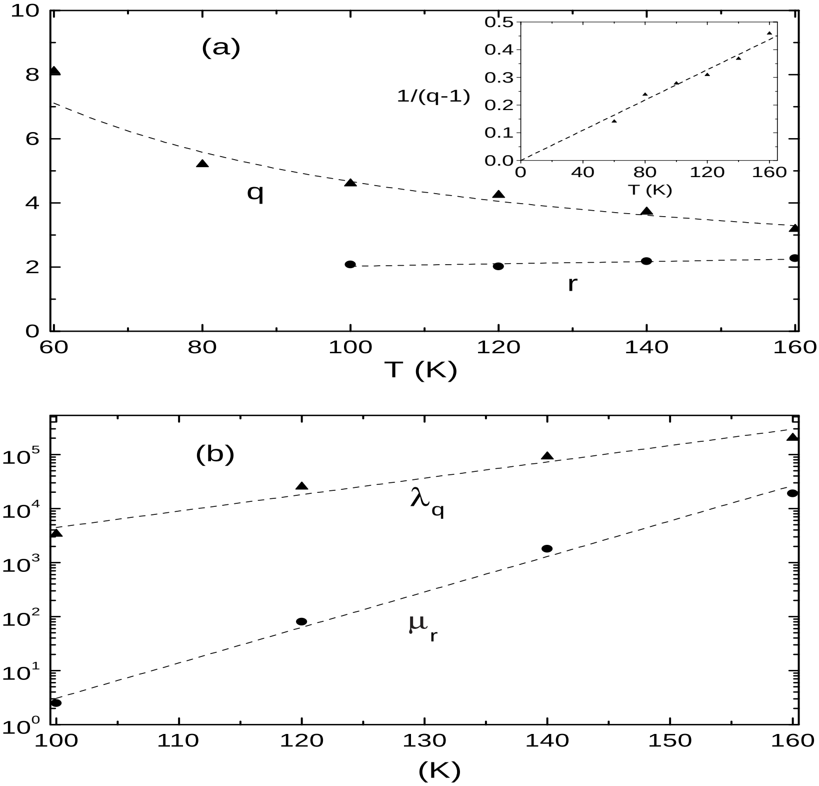

Fig. 14. Temperature dependence of the parameters used to fit the experimental data of Fig. 13: (a) the exponents, and (b) the coefficients.

can be quite satisfactorily fitted with a *single* function, namely

$$N(x) = \frac{N(0)}{[1 + (q-1)\lambda x]^{q/(q-1)}} \tag{126}$$

with $q = 1.53$ and 1.64 for the series ISI and PRD respectively: see Fig.17. The satisfactory quality of the fittings is, after all, not so surprising, since we have mentioned earlier in this paper the connection [160] of this formalism with the Zipf law.

4.11 Electroencephalographic Signals of Epilepsy

It is since long known that the analysis of signals can be done within formalisms which use entropic forms. One such application has been recently done on EEG

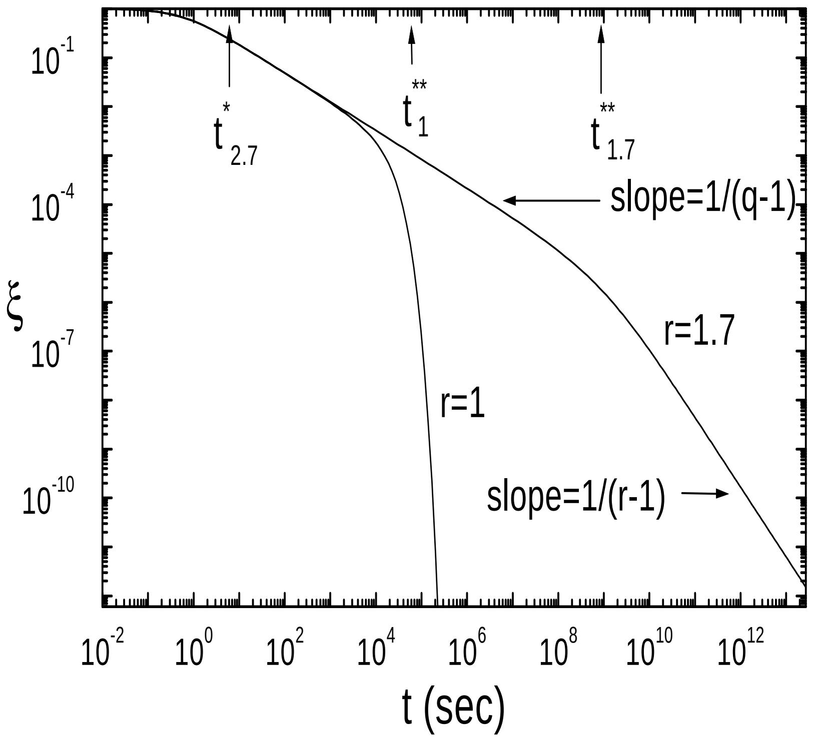

Fig. 15. Typical time dependence of ξ. There are two crossover times, namely t_q^* and t_r^{**} (with $1 \leq r < q$), typically separating three distinct physical regimes: for $0 \leq t << t_q^*$, ξ linearly decreases with t; for $t_q^* << t << t_r^{**}$, $\xi \propto 1/t^{\frac{1}{q-1}}$; for $t_r^{**} << t$, $\xi \propto 1/t^{\frac{1}{r-1}}$ if $r > 1$ (ξ exponentially decreases with t if $r = 1$). The two examples exhibited in the figure correspond to $q = 2.7$, $\lambda_q = 1$, $\mu_r = 10^{-5}$ and $r = 1,\ 1.7$. Erratum in [181]: where we see in this figure $t_{1.7}^{**}$ (correct), we see in [181] $t_{2.7}^{**}$ (wrong).

records of epileptic humans and turtles [185]. The simultaneous use of wavelet-based multiresolution analysis including the nonextensive entropy S_q leads to signals whose interpretation can be clinically neat and pharmacologically convenient. The authors of this novel processing suggest perspectives for building up automatic detection devices.

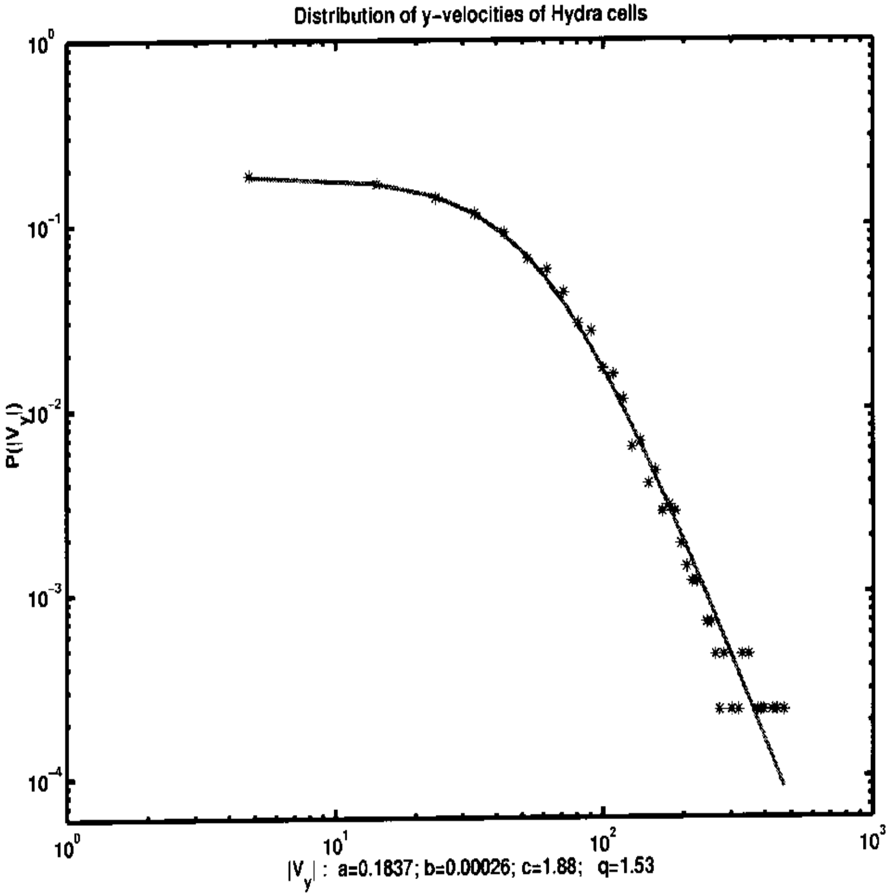

Fig. 16. Distribution of the "vertical" velocities during diffusion of endodermal Hydra cells in an ectodermal aggregate. The abcissa units are 10^{-6} $m/hour$. The fitting was obtained using $q = 1.53$ (see the text).

4.12 Cognitive Psychology

The development of artificial neural networks and their connections with statistical mechanics (e.g., the Hopfield model for associative memory) makes quite natural the approach of cognitive problems with the present nonextensive formalism. Within this philosophy, we performed [186] an experiment of learning/memorization (of 5 × 5 and 7 × 7 square matrix having circles and crosses randomly distributed once for ever) with students of the University level; 150 students were interviewed, the first 30 in order to optimize the experimental protocol, and the other 120 to make the measurements of the time-evolution of the total amounts of errors when the original matrix was successively shown and hidden. The average results were then fitted with those obtained, for the

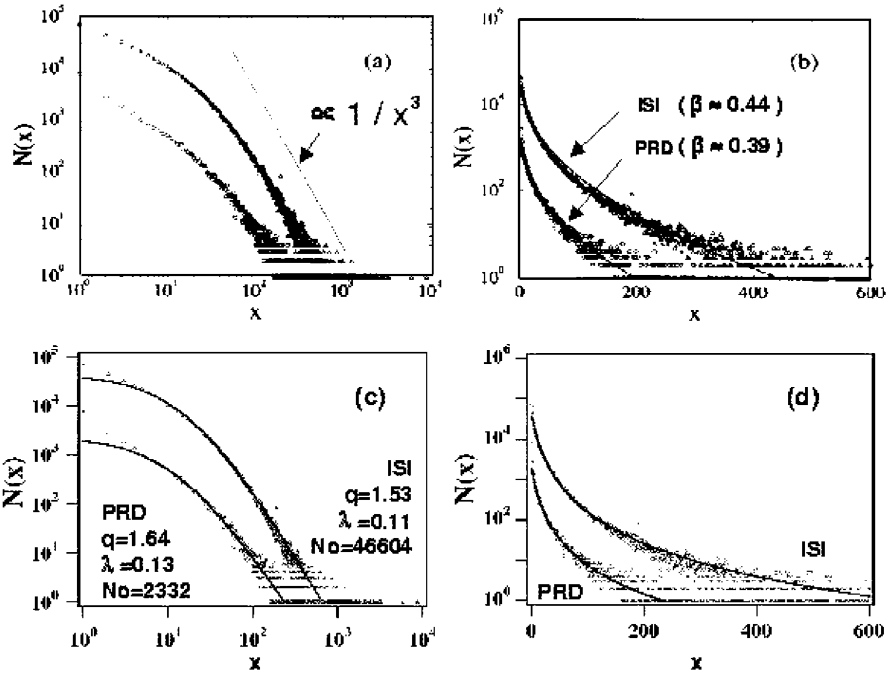

Fig. 17. Distribution of ISI and PRD papers having received x citations. (*a*) and (*b*) exhibit the fittings in [183]; (*c*) and (*d*) exhibit our present fittings (see the text).

same task, with a learning machine [187] having a perceptron architecture and an internal dynamics based on the Langevin equation [105] generalized by Stariolo to arbitrary q. The (average) learning time of the machine turned out to monotonically increase with q, exhibiting a practically divergent derivative at $q = 1$. The best human-machine fit occurred for q slightly above unity. More experiments and comparisons along these lines would be very welcome. Indeed, they would help better understanding some cognitive phenomena, on one hand, and could generate efficient machines for specific tasks, on the other.

4.13 Fully Developed Turbulence and Financial Markets

In 1996 Ghashghaie *et al.* [188] compared financial data with those obtained from turbulent measurements such as those exhibited in Fig. 18, and showed very similar behaviors when appropriate scalings are used. Ramos *et al.* [189] have recently shown that all these data can be satisfactorily fitted with the functional form which emerges from the present formalism. Olsen and Associates data containing bid-ask quotes for (US dollar)-(German mark) exchange rates (1,472,241 records) are presented in Fig. 19 (probability density $P_\tau(z_\tau)$ of price changes; $z_\tau \equiv z(t) - z(t+\tau)$ with $\tau = 640s$, $5120s$, $40960s$ and $163840s$

from top to bottom in the figure). The turbulent flow data presented in Fig. 18 are those in [190] (probability density $P_r(v_r)$ of velocity differences; $v_r \equiv v(x) - v(x+r)$ for spatial scale differences $r = 3.3\eta$, 18.5η, 138η and 325η from top to bottom in the figure, where η is the Kolmogorov scale, i.e., the critical limit for occurrence of viscous dissipation). All these curves exhibit a slight left-right asymmetry (skewness), which has been phenomenologically taken into account by Ramos *et al.* by using the same q for both sides but different widths (i.e., allowing for $\beta^- \neq \beta^+$).

Besides these fittings the authors developed a theory which essentially uses only one experimental value, namely that of the intermittency exponent μ, known to be close to 0.25 [190]. The results they obtained are shown in Fig. 20. Let us remind at this point the definition of μ. The energy spectrum (inertial range) is given by

$$E_k \propto \epsilon^{2/3}\, k^{-5/3}\, (k/K)^{-\mu/9} \;, \tag{127}$$

where ϵ is the energy input rate, and k and K are respectively the wave numbers associated with the eddies and that associated with L, the largest size of the grid. Within Kolmogorov theory $\mu = 0$.

The next step connecting the present nonextensive formalism with fully developed turbulence was accomplished by Arimitsu and Arimitsu [191]. By using a scenario similar to the binomial multiplicative one, they obtained the following relation between q and μ:

$$q = 1 - \frac{1 + \mu + \left\{\log_2\left[1 + (1-2^{-\mu})^{1/2}\right]\right\}\left\{\log_2\left[1 - (1-2^{-\mu})^{1/2}\right]\right\}}{\log_2\left[1 + (1+2^{-\mu})^{1/2}\right] - \log_2\left[1 + (1-2^{-\mu})^{1/2}\right]} \tag{128}$$

where $\log_2$ denotes the logarithm with basis 2. The authors used the experimental value [192] $\mu \approx 0.235$ and obtained $q \approx 0.237$. They then used this value of q and obtained a distribution of turbulent kinetic energy remarkably close to both the experimental one and to that obtained in the theory which assumes binomial multiplicative processes.

The next contribution to this fascinating phenomenon, also using nonextensive thermostatistics, is due to Beck [193]. Using a Langevin-like equation, he obtained the following distribution (normalized to unit variance) for the radial velocity difference u between two points separated by r:

$$p(u) = \frac{1}{Z_q}\left\{1 + \beta(q-1)\left[\frac{u^2}{2} - \frac{1}{R_\lambda^{1/2}}\left(u - \frac{u^3}{3}\right)\right]\right\}^{-\frac{1}{q-1}} \;, \tag{129}$$

where Z_q and β are determined (as functions of q) by $\int_{-\infty}^{\infty} du\; p(u) = 1$ and $\langle u^2\rangle - \langle u\rangle^2 = 1$ (hence, if $R_\lambda >> 1$, $\beta \approx 2/(5-3q)$). The entropic index q is in turn determined by

$$\frac{1}{1-q} = 1 + \log_2 \frac{r}{\eta} \;. \tag{130}$$

It should be noticed that, in contrast with the two, already quite satisfactory, preceding theoretical approaches [189,191] (which both used one free parameter, namely μ) this is a *no-free-parameter* theory. Once the experimental values for (R_λ, η, r) are chosen, the distribution $p(u)$ completely follows. The results are exhibited in Fig. 21. Let us also stress that, within this attempt, the necessary skewness emerges as well. In addition to this, the moment scaling exponents ζ_m (defined through $\langle |u|^m \rangle \sim r^{\zeta_m}$) for $m=$ 2, 3, ..., 8 were calculated. The values obtained are respectively 0.68 [0.705], 1 [1] (by construction), 1.30 [1.305], 1.58 [1.595], 1.82 [1.79], 2.02 [2.06] and 2.16 [2.225], where between brackets we have indicated the average of two sets of experimental data. The average of the absolute values of the discrepancies is 1.6%, which is quite remarkable.

Finally, the most recent contribution to this *saga* was provided by Arimitsu and Arimitsu [195]. Indeed, in an approach related to their first one (and which again uses the experimental value for μ as the unique parameter which is external to the theory) they focused on the m-dependence of ζ_m. We reproduce in Fig. 22 their result compared to those obtained within other available theories, as well as to experimental data. Their curve satisfactorily fits the data *up to* $m = 30$, which is quite remarkable. As a comparison, notice that the range presently illustrated by Beck only goes up to $m = 8$.

As a final comment, let us stress that, whereas both Ramos *et al.* and Beck theories yield $q > 1$, the Arimitsu's approach yields $q < 1$. There is nevertheless no contradiction. Although using the same notation, they are referring in fact to different indexes. Indeed, the Arimitsu's refer to the sensitivity to the initial conditions, whereas the other two theories refer to energy-like indexes. This delicate point is not yet well understood and is further analyzed in subsection 5.7.

5 Computational Evidence and Connections

5.1 Thermalization of a Hot Gas Penetrating in a Cold Gas

In 1991, Waldeer and Urbassek [27] made, assuming $d = 3$ Newtonian mechanics, a computational simulation in which a certain amount of high energy particles penetrate into a cold gas and are thermalized through the interactions between molecules. The cold gas is initially put at BG thermal equilibrium at temperature T_C. The high energy particles at time $t = 0$ are randomly distributed in energy at a quite high energy per particle. The interaction potential was assumed to be hard sphere at short distances and decreasing, at long distances, like $r^{-\alpha}$. They analyzed three typical situations, one with $\alpha \to \infty$, hence well above d (i.e., very short range interactions), the second one with $\alpha = 4$ (i.e., short range interactions), and the last one with $\alpha = 8/3$, which is below d (i.e., long range interactions). In their simulation, they follow the time evolution of the energy distribution of the hot particles. After a transient, this distribution evolves with a regular pattern. For $\alpha > d$, this pattern basically is the BG distribution with a temperature $T(t)$ which gradually approaches T_C from above (with $\lim_{t\to\infty} T(t) = T_C$), in other words, through curves which approximately are straight lines in a log-linear plot. For $\alpha < d$, this approximation occurs through curves which are close to straight lines... in a log-log plot! See Figs. 1, 2 and 3

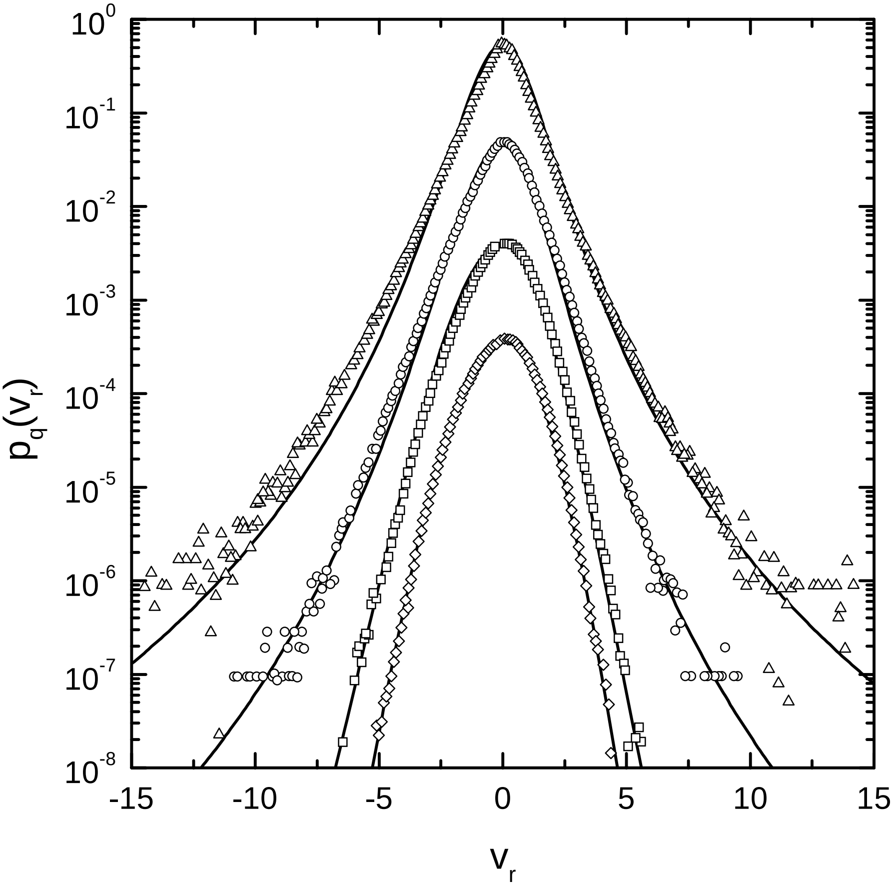

Fig. 18. Standardized distributions of velocity differences $v_r \equiv v(x) - v(x+r)$ where r is the spatial scale, with $L/\eta = 454$, L and η being respectively the integral and Kolmogorov scales. From top to bottom: $r/L=$ 0.0073, 0.0407, 0.3036 and 0.7150. The experimental data are from [190]. The solid curves have been obtained through fittings using asymmetric q-distributions (see the text). From top to bottom: $(q, \beta^-, \beta^+)=$ (1.26,0.69,0.88), (1.20,0.66,0.82), (1.11,0.55,0.76) and (1.08,0.62,0.70). For better visibility the curves have been vertically shifted with respect to each other.

of [27]. Notice that the curvature in these log-log plots tends to be even slightly *upwards* for $\alpha < d$, whereas it is *downwards* for $\alpha > d$. This power-law behavior is typical of $q > 1$. This peculiarity was invoked by Koponen [28] in 1997 as a justification for using the present generalized formalism to discuss electron-phonon relaxation in ion-bombarded solids if the interactions are long-ranged. A study like that of Waldeer and Urbassek [27] which would systematically address the details of that thermalization by gradually varying α across d is missing and would certainly be very welcome.

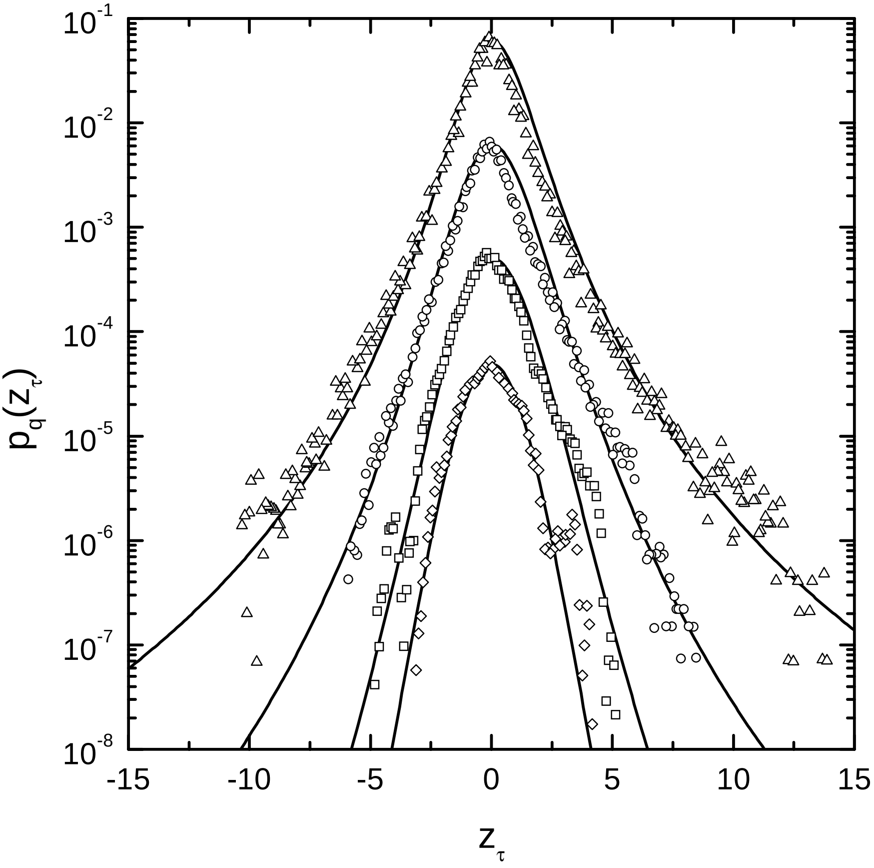

Fig. 19. Standardized distributions of price changes $z_\tau \equiv z(t) - z(t+\tau)$ for (US dollar)-(German mark) exchange rates where τ is the temporal scale, with the integral scale $\tau_L = 186,265s$. From top to bottom: τ/τ_L= 0.0035, 0.0276, 0.2210, 0.8838. The experimental data are from Olsen and Associates (see [188]). The solid curves have been obtained through fittings using asymmetric q-distributions (see the text). From top to bottom: (q, β^-, β^+)= (1.35,1.12,0.98), (1.26,0.83,0.72), (1.16,0.75,0.61) and (1.11,0.75,0.77). For better visibility the curves have been vertically shifted with respect to each other.

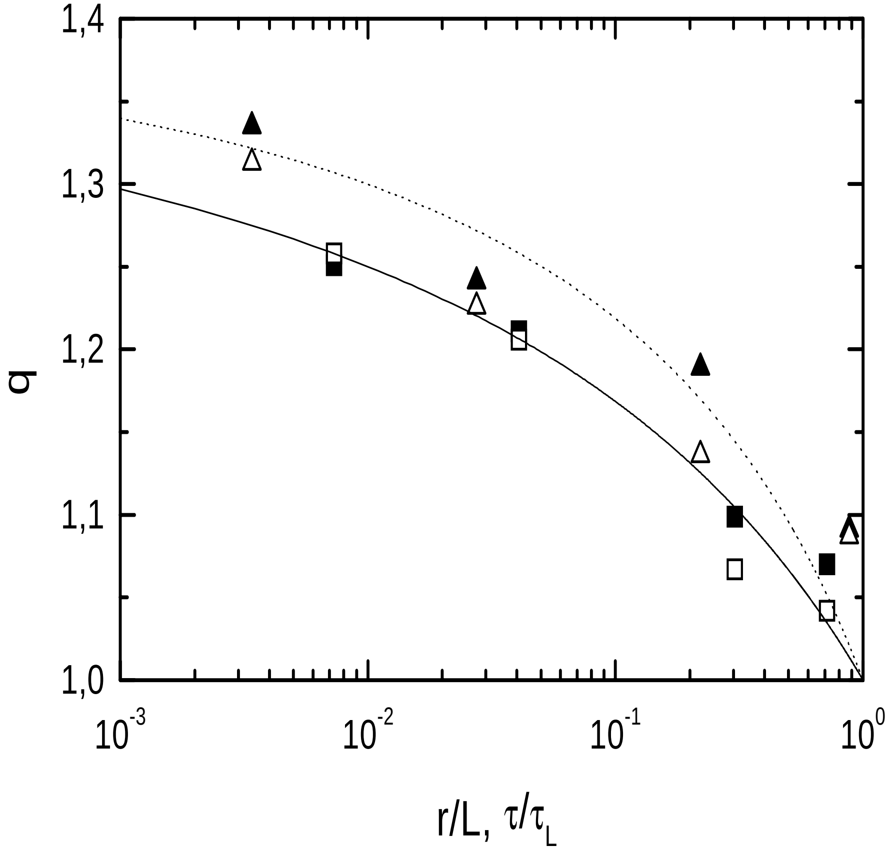

Fig. 20. Dependence of q on normalized spatial (r/L) and temporal (τ/τ_L) scales, respectively corresponding to turbulence and financial markets. Turbulence: solid line (theory by Ramos *et al.*), open squares (asymmetric fitting; see Fig. 18) and solid squares (symmetric fitting, i.e., imposing $\beta^+ = \beta^- \equiv \beta$). Financial markets: dotted line (theory by Ramos *et al.*), open triangles (asymmetric fitting; see Fig. 19) and solid triangles (symmetric fitting, i.e., imposing $\beta^+ = \beta^- \equiv \beta$).

5.2 Long-Range Classical Hamiltonian Systems: Static Properties

Let us focus here on what we refer to as *weak* violation of BG statistics. We use this expression to distinguish it from what we call *strong* violation of BG statistics. Both of them lead to nonextensive quantities, but, whereas the strong violation concerns $q \neq 1$, the weak one concerns $q = 1$ calculations. To make all this explicit we shall here focus on classical systems, i.e., all observables are

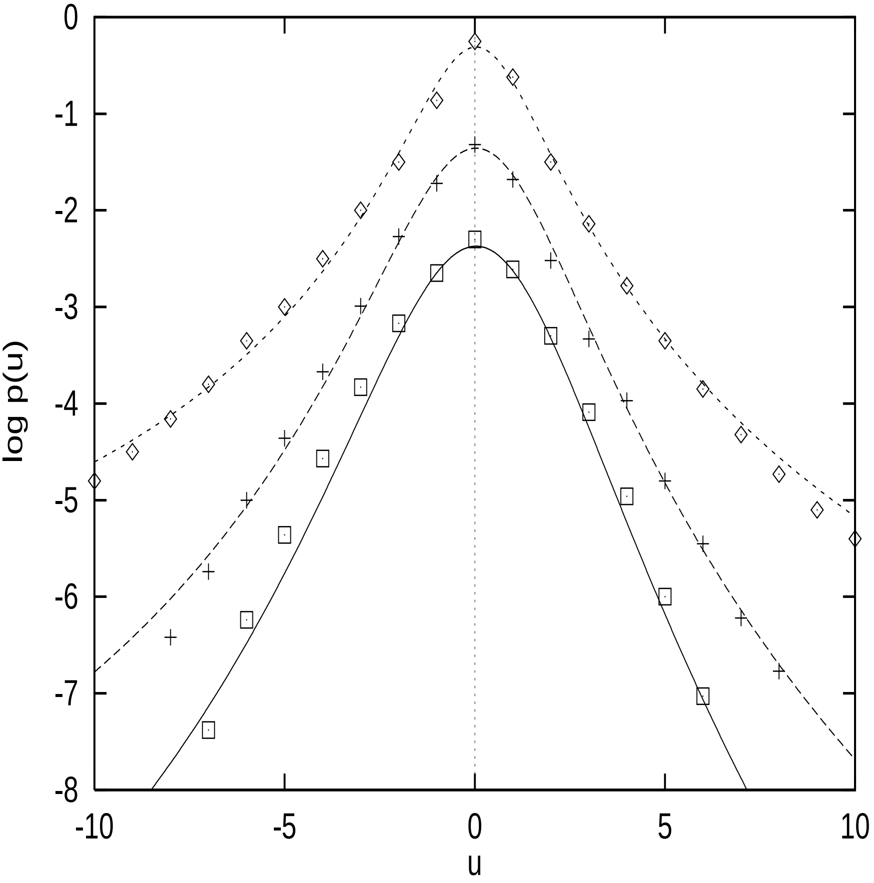

Fig. 21. Distribution of velocity differences (normalized to have unity variance) for Reynolds number $R_\lambda = 852$. From top to bottom: r/η= 3.3, 23.6 and 100. The experimental data are from [194]. The curves are from Beck's theory. For better visibility the curves have been vertically shifted with respect to each other.

assumed to commute. Let us consider the following paradigmatic Hamiltonian:

$$H = \frac{1}{2m} \sum_{i=1}^{N} p_i^2 + \sum_{i \neq j} V(r_{ij}) \ , \tag{131}$$

where m is a microscopic mass, $\{p_i, r_i\}$ are the d-dimensional linear momenta and positions associated with N particles, and $r_{ij} \equiv r_j - r_i$. A typical situation is that of a finite confined system but, if some care is taken, the system could as well be thought of as having periodic boundary conditions. To be specific, let

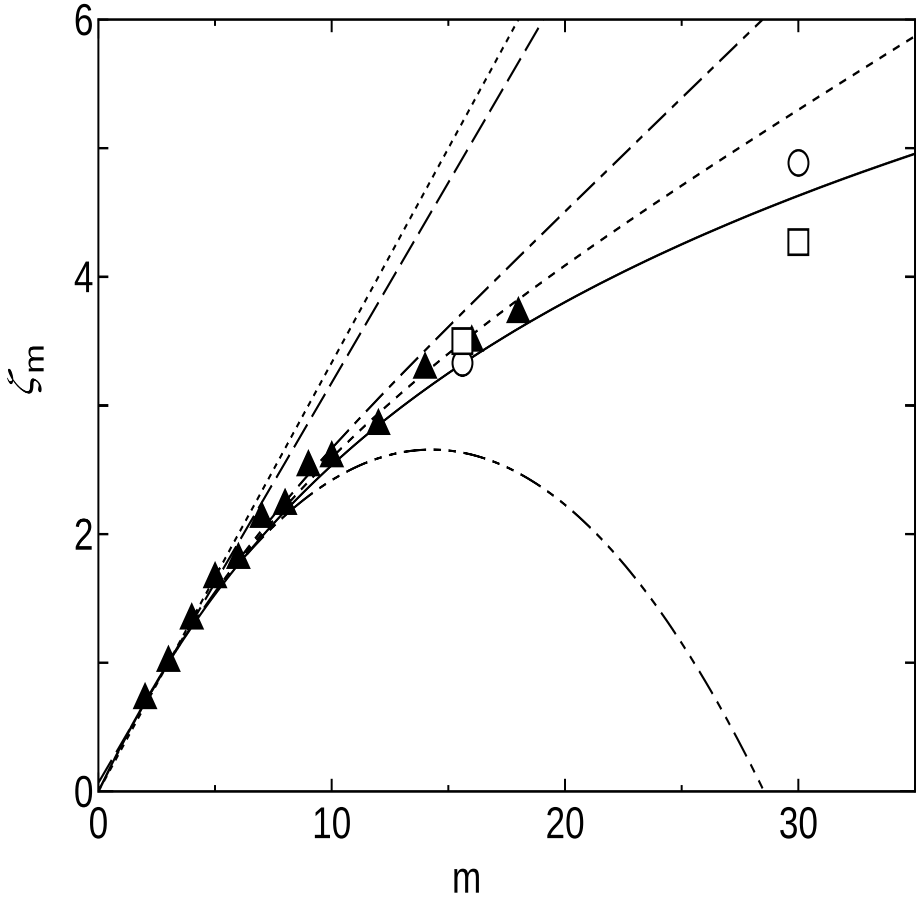

Fig. 22. The m-dependence of the scaling exponent ζ_m. Theoretical curves (from top to bottom): 1941 Kolmogorov theory, β-model (with $D_\beta = 2.8$), p-model (with $\mu = 0.235$), log-Poisson model (fitted), present Arimitsu and Arimitsu model (solid curve; with $\mu = 0.235$, which here corresponds to $q = 0.370$, instead of $q = 0.237$ obtained in their previous approach [191]) and log-normal mode (with $\mu = 0.235$). Experimental data: solid triangles from Anselmet *et al.* in [192]; squares and cercles from [196].

us assume

$$V(r_{ij}) = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{\alpha}} \quad (A > 0;\ B > 0;\ 0 \le \alpha < 12)\ , \tag{132}$$

where, in order to avoid any singularity at the origin (for any dimension d not exceedingly high), we have assumed, for the repulsive term, the Lennard-Jones exponent 12. What we desire to focus on in the present discussion is possible singularities associated with infinite distances, i.e., the effects of *long*-range (attractive) interactions. The case $(\alpha, d) = (6, 3)$ precisely recovers the standard

Lennard-Jones fluid; the case $(\alpha, d) = (1, 3)$ is asymptotically equivalent to Newtonian gravitation; the case $(\alpha, d) = (d-2, d)$ is asymptotically equivalent to d-dimensional gravitation (i.e., the one associated with the solutions of the d-dimensional Poisson equation); the case $(\alpha, d) = (3, 3)$ basically reproduces the distance dependence of permanent dipole-dipole interaction. The range of the (attractive) interaction increases when α decreases; $\alpha \to 12$ corresponds to very short-ranged interactions, whereas $\alpha = 0$ corresponds to the situation of the Mean Field Approximation, where every particle (attractively) interacts with every other with the *same* strength, in all occasions.

A typical quantity to be calculated within BG statistics is the following one (basically related to the $T = 0$ internal energy per particle):

$$\int_1^\infty dr\; r^{d-1}\, r^{-\alpha} \;, \tag{133}$$

where the distances r have been expressed in units of a characteristic length of the problem. We immediately verify that this integral *converges* if $\alpha > d$, and *diverges* if $0 \le \alpha \le d$. Consequently, thermodynamic calculations in the $0 \le \alpha \le d$ case have to be done with some care, and not blindly following the standard rules associated with BG statistics (i.e., $q = 1$). It is in this sense that we use the expression "weak" violation of BG statistics. The care to which we refer is the fact that we have to strictly consider the *finite* size of the physical system. Consistently, a relevant quantity that emerges naturally is

$$N^* \equiv d \int_1^{N^{1/d}} dr\; r^{d-1}\, r^{-\alpha} = \frac{N^{1-\alpha/d} - 1}{1 - \alpha/d} \;. \tag{134}$$

We can check that, in the $N \to \infty$ limit, we have

$$N^* \sim \begin{cases} \frac{1}{\alpha/d - 1} & \text{if} \quad \alpha/d > 1 \;, \\ \ln N & \text{if} \quad \alpha/d = 1 \;, \\ \frac{N^{1-\alpha/d}}{1-\alpha/d} & \text{if} \quad 0 \le \alpha/d < 1 \;. \end{cases} \tag{135}$$

As it will become transparent later on, what these regimes imply is that the system is *extensive* for $\alpha/d > 1$ (hence standard thermodynamics apply), whereas it is *nonextensive* for $0 \le \alpha/d \le 1$, and special scalings become necessary [197,198] in order to have both a mathematically well posed problem, and a physical unfolding (or qualification) of the nonextensive region. The $\alpha/d > 1$ regime has since long been analyzed [15], and it is well known that extensivity (or, *stability*, as also referred to) is lost for $\alpha/d \le 1$. However, to the best of our knowledge, the scalings associated with the quantity N^*, as well as its numerically efficient collapsing properties, were introduced for the first time by Jund *et al.* [198] in 1995.

A quantity related to N^*, namely $\tilde{N}$ turns out to be even more convenient. It is defined through (see Eq. (102) of [3])

$$\tilde{N} \equiv N^* + 1 = \frac{[N^{1-\alpha/d} - \alpha/d]}{1 - \alpha/d} \;. \tag{136}$$

In the $N \to \infty$ limit, we have

$$\tilde{N} \sim \begin{cases} \frac{\alpha/d}{\alpha/d-1} & \text{if} \quad \alpha/d > 1 \ , \\ \ln N & \text{if} \quad \alpha/d = 1 \ , \\ \frac{N^{1-\alpha/d}}{1-\alpha/d} & \text{if} \quad 0 \leq \alpha/d < 1 \ . \end{cases} \tag{137}$$

In the limit $\alpha/d \to \infty$, $\tilde{N} \to 1$; in the limit $\alpha/d \to 1+0$, $\tilde{N} \sim 1/(\alpha/d-1)$; in the limit $\alpha/d \to 0$, $\tilde{N} \sim N$. Roughly speaking, $\tilde{N}$ characterizes the *effective* number of neighbors that can be associated with a given particle. This is the convenience to which we referred above. See Fig. 23.

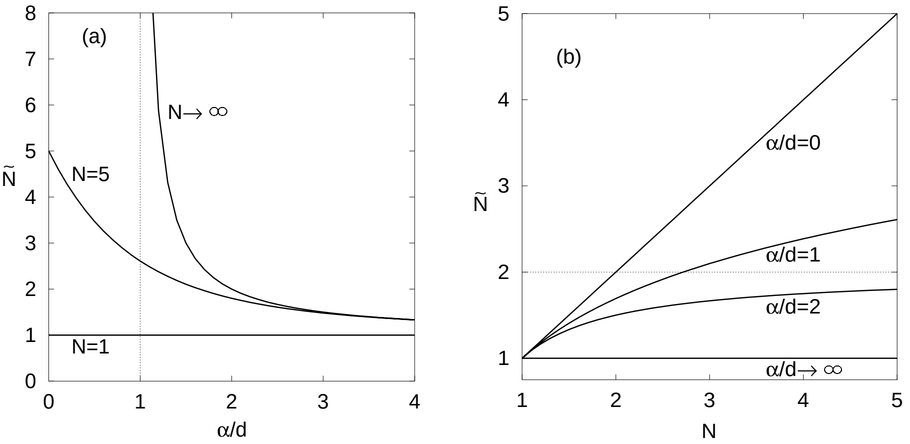

Fig. 23. The rescaling function $\tilde{N}(N, \alpha/d)$ versus α/d for typical values of N (a), and versus N for typical values of α/d (b). For fixed $\alpha/d \geq 0$, $\tilde{N}$ monotonically increases with N increasing form 1 to ∞; for fixed $N > 1$, $\tilde{N}$ monotonically decreases for α/d increasing from 0 to ∞. $\tilde{N}(N,0) = N$, thus recovering the Mean Field Approximation usual rescaling; $\lim_{N\to\infty} \tilde{N}$ diverges for $0 \leq \alpha/d \leq 1$, thus separating the extensive from the nonextensive region; $N(\infty, \alpha/d) = (\alpha/d)/[(\alpha/d)-1]$ if $\alpha/d > 1$; $\lim_{\alpha/d\to\infty} \tilde{N} = 1$, thus recovering *precisely* the traditional intensive and extensive thermodynamical quantities; $\tilde{N}(N,1) = \ln N$.

We are ready now to present the kind of size-scalings we expect to be necessary for thermodynamically describing a generic *classical* Hamiltonian system with the type of interactions above mentioned. Let us focus on a simple fluid, and start with the standard case, i.e., $\alpha > d$. Its Gibbs energy $G(T,p,N)$ is given by

$$\frac{G(T,p,N)}{N} \sim \frac{U(T,p,N)}{N} - T\,\frac{S(T,p,N)}{N} + p\,\frac{V(T,p,N)}{N} \ , \tag{138}$$

where U, S, V, N, T and p respectively are the total internal energy, total entropy, total volume, total number of particles, temperature and pressure. In the $N \to \infty$ limit, we obtain

$$g(T,p) = u(T,p) - T\, s(T,p) + p\, v(T,p) \ , \tag{139}$$

where the corresponding densitary variables have been introduced.

In contrast with the above, if we have $0 \le \alpha \le d$, the scalings are different, namely

$$\frac{G(T,p,N)}{N\,\tilde{N}} \sim \frac{U(T,p,N)}{N\,\tilde{N}} - \frac{T}{\tilde{N}}\ \frac{S(T,p,N)}{N} + \frac{p}{\tilde{N}}\ \frac{V(T,p,N)}{N} \ . \tag{140}$$

Consistently we have

$$g(\tilde{T},\tilde{p}) = u(\tilde{T},\tilde{p}) - \tilde{T}\, s(\tilde{T},\tilde{p}) + \tilde{p}\, v(\tilde{T},\tilde{p}) \ , \tag{141}$$

where

$$\tilde{T} \equiv \frac{T}{\tilde{N}} \ , \quad \tilde{p} \equiv \frac{p}{\tilde{N}} \ . \tag{142}$$

These equations recover the previous ones (i.e., Eqs. (138,139)), i.e., those associated with $\alpha > d$, as a particular case. Indeed, for $\alpha > d$, $\tilde{N}$ becomes a constant. This constant becomes unity in the $\alpha/d \to \infty$ limit, hence $\tilde{T} \to T$, $\tilde{p} \to p$, etc.; this is a nice illustration of what was already mentioned, namely that $\tilde{N}$ is more convenient that N^*.

So, we see that long range interactions have important thermodynamical consequences, namely

(i) the *energy* quantities (G, U, which, in the usual Legendre-transform structure, appear *alone*), which are *extensive* for $\alpha > d$, *loose* their extensivity;

(ii) the *non-energy* quantities (S, V, which normally appear, in the Legendre structure, in *canonical pairs* with intensive quantities), which are *extensive* for $\alpha > d$, *preserve* their extensivity, though with rescaled control parameters;

(iii) the control parameters (T, p), that are *intensive* for $\alpha > d$, *loose* their intensivity.

Consistently, to have mathematically well defined and physically useful equations of states and related quantities, everything must refer to *finite* quantities, hence, we must express all relations with the above rescaled variables. This does *not* imply that thermal equilibrium occurs through sharing equal values of $\tilde{T}$, $\tilde{p}$, etc. The zero-th principle of thermodynamics appears to hold in the usual way, *even if we have long range interactions in the system.* Although we have illustrated these features on a fluid, it is clear that the same considerations hold for all types of thermodynamical systems (magnets, dielectric substances, elastic solids, etc.).

Two particular remarks must be made at this point:

(i) When every element of the system equally interacts with each other (i.e., $\alpha = 0$), $\tilde{N} = N$, and consequently $\tilde{T} = T/N$. In what concerns the thermostatistical approach of a system, this is equivalent to dividing the microscopic coupling constants by N, a familiar feature that is artificially imposed in *all* Mean Field calculations. We have used the word "artificial" because, whenever

$\alpha \le d$, the Hamiltonian which includes the *microscopic* interactions *indeed is nonextensive!* (and so is U). To divide the two-body coupling constants by N when $\alpha = 0$ (or, by $\tilde{N}$, when $\alpha > 0$) certainly is an artificial manner of forcing to be extensive a Hamiltonian which physically is not. This practice is traditionally frequent among magneticiens (who divide J by N), but certainly not among astronomers, who normally do not even think about what would be a very strange way of renormalizing the gravitational constant G!

(ii) If a singularity (for example, a critical phenomenon) occurs under particular physical conditions, it must generically occur at *finite* values of $\tilde{T}$, $\tilde{p}$, etc., and *not* at finite values of T, p, etc. Let us illustrate this for the simple case of a critical temperature: $\tilde{T}_c$ must be finite, hence, if $\alpha/d > 1$, it must be $T_c(\alpha, d) \propto 1/(\alpha/d - 1)$ This implies that T_c must *generically* diverge for *all* classical systems at the extensive-nonextensive frontier. More precisely

$$T_c(\alpha, d) \sim \frac{A(d)}{\alpha/d - 1} \quad (\alpha/d \to 1 + 0) \ , \tag{143}$$

where $A(d)$ is a system-dependent finite constant. In fact, let us anticipate that this precise behavior has been observed in the systems available in the literature [198] ($d = 2$ and $d = 3$ Lennard-Jones-like fluids, $d = 1$ and $d = 2$ Ising and Potts ferro- and antiferromagnets, etc.), with no exception. To illustrate the connection between fluid models like the extended Lennard-Jones one above considered, and localized spin systems, let us briefly focus on the Ising ferromagnet. The simplest long-range N-spin cubic-lattice Hamiltonian of this kind is given by [20]

$$H = -J \sum_{i \neq j} \frac{S_i S_j}{r_{ij}^{\alpha}} \quad (J > 0;\ \alpha \ge 0;\ S_i = \pm 1\ \forall i) \ , \tag{144}$$

where, for $d = 1$, r_{ij} =1, 2, 3, ...; for $d = 2$, $r_{ij} = 1,\ 2^{1/2},\ 2, ...$; for $d = 3$, $r_{ij} = 1,\ 2^{1/2},\ 2^{1/3},\ 2, ...$; and so on. Clearly, the limit $\alpha \to \infty$ recovers the first- neighbor d-dimensional spin 1/2 ferromagnet, whereas $\alpha = 0$ corresponds to the Mean Field Approximation. For this model, $kTc(\alpha, d)/J$ diverges for $0 \le \alpha/d < 1$, decreases for α/d increasing above unity, and approaches the first-neighbor value (0 for $d = 1$, 2.269... for $d = 2$, etc.) for $\alpha/d \to \infty$. Also, $kT_c/J \sim A(d)/[(\alpha/d) - 1]$ in the $\alpha/d \to 1 + 0$ limit. The introduction of $\tilde{T}$ nicely enables the *unfolding* of the region where T_c diverges. Indeed, $k\tilde{T}_c/J \equiv kT_c/(J\tilde{N})$ is *finite* in *both* extensive ($\alpha/d > 1$) and nonextensive ($0 \le \alpha/d \le 1$) regions, thus providing an enlightening unification. Finally, let us mention that, it seems that all equations of states (e.g., $\lim_{N \to \infty} M(T, N)/N$ can be, for all $\alpha/d \le 1$, mapped into that associated with the Molecular Field Approximation. This simplifying feature appears to hold *only* for the static thermodynamic properties, and not for the dynamical ones, as will become clear later on (in subsection 5.7).

(iii) It is interesting to illustrate, say on the internal energy U, the role played by N. At fixed temperature $\tilde{T}$, in the $N \to \infty$ limit, we have for any fixed value $\lambda > 0$ that

$$U(\lambda N) \sim \frac{(N\lambda)^{1-\alpha/d} - \alpha/d}{N^{1-\alpha/d} - \alpha/d} \, \lambda \, U(N) \ . \tag{145}$$

This relation implies that, for $\alpha/d \geq 1$,

$$U(\lambda N) \sim \lambda\, U(N) \tag{146}$$

and that, for $0 \leq \alpha/d < 1$,

$$U(\lambda N) \sim \lambda^{2-\alpha/d}\, U(N)\ . \tag{147}$$

So, if we eat an ice-cream whose weight is the *double* of another one, we shall get into our body the *double* of calories. But, were it not for the gravitational singularity at the origin which introduces complications, if we "eat" a double-mass galaxy (i.e., $\alpha = 1$ and $d = 3$), instead of a single-mass one, we would get $2^{5/3} \approx 3.2$ as much calories! As we see, size scaling at the thermodynamical limit ignores the microscopic interactions if they are short-ranged, but it does "remember" the ratio α/d if the interactions are long-ranged.

To close this subsection, let us emphasize that what we have been focusing on here is what we refer to as the *weak* violation of BG statistical mechanics (see Fig. 4). These are analytical or Monte Carlo $q = 1$ calculations (i.e., the energy distribution obeys by imposition the Boltzmann factor), *but* the variables must be scaled with $\tilde{N}$, which is not at all necessary for the standard, short-ranged interacting systems.

5.3 Long-Range Tight-Binding Systems

In the previous subsection we addressed classical systems. It is clear, however, that similar nonextensivity is expected to emerge in quantum systems if long-range interactions are present. One such Hamiltonian is the tight-binding-like which follows [199,200]:

$$H = \sum_{i=1}^{N} \epsilon_i c_i^{+} c_i + \sum_{i,j \neq i} \frac{V}{r_{ij}^{\alpha}} c_i^{+} c_j \quad (\alpha \geq 0;\ r_{ij} = 1, 2, 3, ...)\ , \tag{148}$$

where c_i^{+} and c_i are the creation and annihilation operators associated with electrons on site i, the $\{\epsilon_i\}$ are the on-site energies, and V is the inter-site energy. The $T = 0$ electron diffusive properties corresponding to this Hamiltonian exhibit a variety of anomalies intimately related to $\tilde{N}$, as preliminary shown by Nazareno and Brito [199]. More details have been recently exhibited by Borland, Menchero and myself [200]: see Fig. 24.

5.4 Granular Systems

In 1995, Taguchi and Takayasu [26] simulated a vertically vibrated bed of powder with inelastic collisions and studied the distribution of horizontal velocities. In the lower layers (so called *solid phase*) they observed a standard Maxwellian (Gaussian) distribution of velocities. The situation was sensibly different in the upper layers (so called *fluidized phase*). Indeed, there the distribution was a

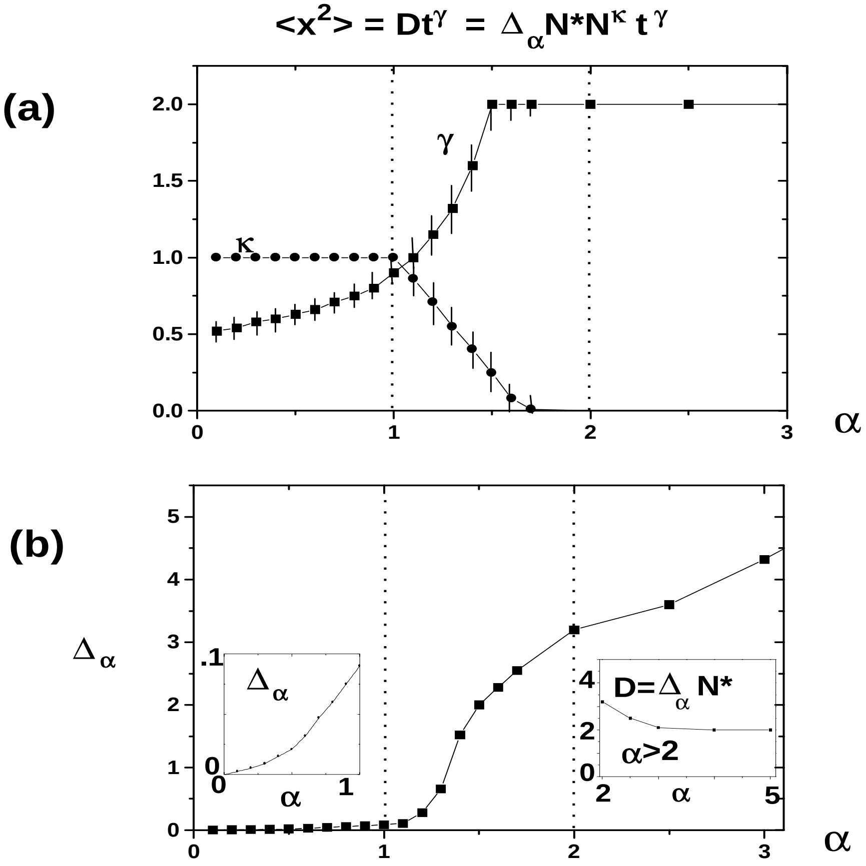

Fig. 24. Some of the various anomalies present in this quantum problem are shown for the time evolution of the mean square displacement, namely the α-dependences of the exponents κ and γ (a), and of the amplitude Δ_α (b).

Student's t-distribution, precisely the one appearing in Eq. (63) with $q = 5/3$ (precisely the value which, for $d = 1$, separates the Gaussian from the Lévy regime!), and assuming an energy proportional to the $(velocity)^2$ (together with $d = r = 2$, hence a constant density of states). This anomaly must be related to the fractal-like granular clusterization which occurs in real space [201] but a deep analysis would be welcome. Also would further simulations, for instance of the cooling type. Studies of such computational models, either externally forced or just left to their own isolated evolution, can provide important physical insights, especially if quantities like the energy distribution, the Liapunov spectrum (or at least its maximum value) or possible multifractality are focused on.

5.5 $d = 1$ Dissipative Systems

One-dimensional maps constitute the simplest systems which might present chaos. Basically they consist of the following recurrent equation:

$$x_{t+1} = h(x_t; a) \;\; (t = 0, 1, 2, ...) \; , \tag{149}$$

where $h(x; a)$ is a rather simple nonlinear function of x, and a is a control parameter. Typically, both x and a are real numbers (but higher-dimensional situations are of course possible, and frequently studied). The logistic map, for instance, exhibits this structure. Typically, for $a < a_c$, the system exhibits simple orbits, the attractor being a cycle whose number of elements is finite. For $a > a_c$, the system can exhibit attractors with an infinite number of elements. The value a_c is the critical one, usually referred to as the *chaos threshold*; the associated attractor typically constitutes a *multifractal* characterized by the so called *multifractal function* $f(\alpha_H)$, where α_H is the *Holder* (or *crowding*) *exponent*. The $f(\alpha_H)$ function is generically concave, attains its maximum at a value of α_H in the interval $[\alpha_H^{min}, \alpha_H^{max}]$ and this maximum equals the *fractal* or *Hausdorff dimension* d_f. An important feature of this type of maps is the sensitivity to the initial conditions (and, of course, to the rounding at any intermediate calculation). More precisely, if we note $\Delta x(0)$ a small variation in the initial condition x_0, and follow its time evolution $\Delta x(t)$, we can define the *sensitivity function* $\xi(t)$ as $\xi(t) \equiv \lim_{\Delta x(0) \to 0} \frac{\Delta x(t)}{\Delta x(0)}$. At most values of a, $\xi(t)$ satisfies $d\xi/dt = \lambda_1 \xi$, hence

$$\xi = \exp(\lambda_1 \, t) \; , \tag{150}$$

where λ_1 is the so called *Liapunov exponent.* If $\lambda_1 > 0$ ($\lambda_1 < 0$) the system is said *strongly sensitive (insensitive)* to the initial conditions. The $\lambda_1 = 0$ possibility can also occur and is referred to as the *marginal* case. In this situation $\xi(t)$ satisfies [99,202,203] $d\xi/dt = \lambda_q \, \xi^q$, hence

$$\xi = [1 + (1-q)\lambda_q \, t]^{\frac{1}{1-q}} \tag{151}$$

which recovers Eq. (150) as the $q = 1$ case. Two generic $\lambda_1 = 0$ possibilities exist for which ξ is finite and nonvanishing for *all* values of t, namely $q < 1$ with $\lambda_q > 0$ (*weakly sensitive* to the initial conditions), and $q > 1$ with $\lambda_q < 0$ (*weakly insensitive* to the initial conditions). For instance, the logistic map exhibits $q = 1$ for almost all values of a but exhibits $q < 1$ at the chaos threshold and $q > 1$ at every doubling-period as well as tangent bifurcations. More two generic $\lambda_1 = 0$ possibilities exist in principle, namely $q < 1$ with $\lambda_q < 0$ (which makes ξ to vanish at a *finite* time $t_0 = 1/[(q-1)\lambda_q]$), and $q > 1$ with $\lambda_q > 0$ (which makes ξ to diverge at a *finite* time $t_\infty = 1/[(q-1)\lambda_q]$). Finally, an extremely marginal case can in principle occur, namely that for which λ_q vanishes for *all* values of q. In that case, the sensitivity or insensitivity to the initial conditions must be extremely weak (for instance, of the logarithmic type).

Of all the above possibilities, by far the most interesting is in some sense the first of them, i.e., when $q < 1$ and $\lambda_q > 0$. Indeed, in this case we do have permanent mixing, but at a slow (power-law) rate. It is this situation that we primarily focus on in what follows. First of all, for this case, Eq. (151) must be understood only as the *upper bound* of $\xi(t)$. Indeed, $\xi(t)$ presents complex, large, and ever lasting fluctuations as t increases (see also [204], where this behavior is described; these references contain the first observations and discussions of the $\xi \propto t^{1/(1-q)}$ law, which can be now seen as the asymptotic region $t >> 1$ of Eq. (151)).

The sensitivity to the initial conditions of three different families of maps at their respective edges of chaos have been recently studied [99,202,203] in a systematic numerical way. These families of maps are natural generalizations (to arbitrary power at their extremum) of the logistic, periodic and circle maps. They are respectively defined as follows:

$$x_{t+1} = 1 - a|x_t|^{z_L}$$

$$(t = 0,1,2,...;\ x_t \in [-1,1];\ 0 < a \le 2;\ z_L \ge 1)\ , \tag{152}$$

$$x_{t+1} = d\cos(\pi |x_t - 1/2|^{z_P/2})$$

$$(t = 0,1,2,...;\ x_t \in [-d,d];\ d > 0;\ z_P \ge 1)\ , \tag{153}$$

$$\theta_{t+1} = \Omega + \left[\theta_t - \frac{1}{2\pi}\sin(2\pi\theta_t)\right]^{z_C/3} \qquad [\mathrm{mod}(1)]$$

$$(t = 0,1,2,...;\ \theta_t \in [0,1];\ 0 \le \Omega \le 1;\ z_C > 0)\ , \tag{154}$$

where

$$\begin{aligned} z_L &= 2: \ \text{the usual logistic map}\ , \\ z_P &= 2: \ \text{the usual periodic map}\ , \\ z_C &= 3: \ \text{the usual circle map}\ , \end{aligned} \tag{155}$$

respectively. The logistic-like and the periodic-like families belong to the same set of universality classes, namely characterized by $z_L = z_P$. The circle-like family belongs to a different set of universality classes, characterized by z_C. Some of the results are exhibited in Fig. 25, where we see (q^*, d_f) for typical values of the exponents $z_L = z_P$ and z_C (q^* denotes the value, appearing in Eq. (151), which characterizes the upper bound mentioned above). As we see, $d_f = 1$ does *not* guarantee $q = 1$, which would be the false idea we could have if we only knew the results associated with say logistic-like maps. Consequently, the idea is reinforced that what is relevant for nonextensivity is *not* the degree of occupancy of the phase space, but rather the speed at which *mixing* occurs within it.

It has been shown that a large class of such systems (for which $f(\alpha_H^{min}) = f(\alpha_H^{max}) = 0$) verify, at the chaos threshold, the following scaling law [203]:

$$\frac{1}{1-q^*} = \frac{1}{\alpha_H^{min}} - \frac{1}{\alpha_H^{max}} \tag{156}$$

This is a fascinating relation. Indeed, its left-hand member concerns the dynamics of the sensitivity to initial conditions of the map, whereas its right-hand member concerns pure, though nontrivial, geometry. Under what precise mathematical conditions does it hold? How should it be generalized in order to also cover

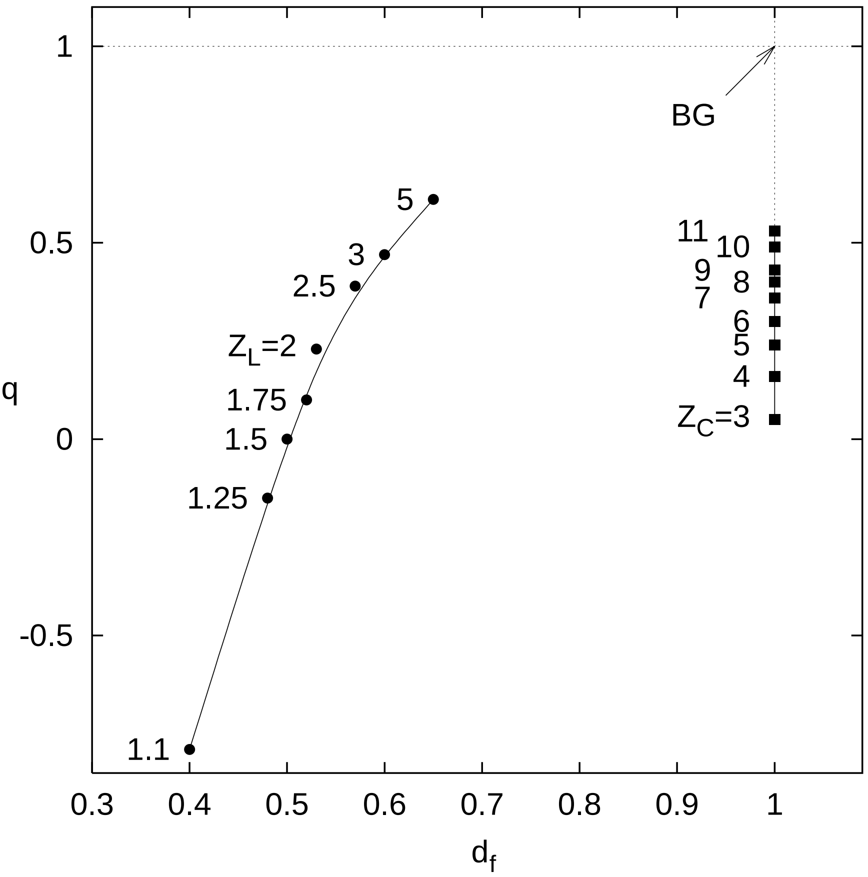

Fig. 25. The special value $q^* \equiv Q \equiv q_S$ of the entropic index, and the fractal dimension d_f of the attractor at the edge of chaos of the logistic-like and periodic-like families of maps ($z_L = z_P$) and cercle-like familiy of maps (z_C); the solid lines that run between the numerical data are guides to the eye. The corner $q = d_f = 1$ corresponds to a typical Boltzmann-Gibbs situation. For $z_L \to \infty$ and $z_C \to \infty$ the points approach the BG corner. However, it can be shown that $\lim_{z_L \to \infty} q^*(z_L) < 1$; probably the same occurs for z_C.

the standard case of Euclidean geometry ($\alpha_H^{min} = \alpha_H^{max} = d_f = 1$) for which one expects $q^* = 1$? (Should we also consider simultaneously $f(\alpha_H^{min}) = f(\alpha_H^{max}) = 1$?; Could Eq. (156) be generalized into say $1/(1 - q^*) = 1/[\alpha_H^{min} - f(\alpha_H^{min})] - 1/[\alpha_H^{max} - f(\alpha_H^{max})]$, the Euclidean case thus corresponding to a special limit of the type $\alpha_H^{min}/f(\alpha_H^{min}) = \alpha_H^{max}/f(\alpha_H^{max}) = q^* = 1$?). What happens for two- or more-dimensional maps? What happens if, instead of maps, we have ordinary (or even partial) differential equations? To answer all these questions, computational effort is invaluable in order to prepare the road to analytical proofs.

Let us make at this point an important clarification to avoid an easy confusion in the mind of more than one reader. In multifractal theory there exists an index frequently noted q, but which we shall here note q_M (M stands for multifractal). The $f(\alpha_H)$ function appears as a convenient Legendre transfom (see for instance [79]) which involves q_M. This transformation provides the relation $\alpha_H(q_M)$. The quantities α_H^{min} and α_H^{max} earlier introduced, and appearing in Eq. (147), satisfy

$$\alpha_H^{min} = \lim_{q_M \to \infty} \alpha_H(q_M) \tag{157}$$

and

$$\alpha_H^{max} = \lim_{q_M \to -\infty} \alpha_H(q_M) \; . \tag{158}$$

Consequently, although multifractals inspired the definition of S_q and although several mathematical properties exhibit naturally the same (or similar) form (see for instance [79]), the two formalisms are deeply different. In multifractal theory, as just mentioned, a Legendre transformation is operated on q_M, thus yielding α_H, whereas, in nonextensive statistical mechanics, q is maintained fixed all the way long (since it is determined by the microscopic, or mesoscopic, dynamics of the system). The Legendre transform of the latter is precisely that usually done in standard, extensive thermodynamics. The Legendre transform of the former is a convenient manner for describing the complex geometry of multifractals. These comments might seem absolutely dispensable for many readers, but experience has proved just the opposite!

Before proceeding, it is mandatory to clarify what the index q appearing in the differential equation yielding Eq. (151) has to do with the one appearing in the present generalized entropy. In fact, they are one and the same, and the connection is established through a generalization of the so called Pesin equality or identity. Let us illustrate the basic ideas on the logistic map herein considered, fixing $a = 2$ (which corresponds to its largest Liapunov exponent, namely $\ln 2$). Assume that we make a partition of the x interval into a large number W of equally small windows, chose arbitrarily *one* of those windows and randomly put a large number N of points inside. By identifying the probabilities with the ratios of points in each one of the W windows (the approximation becomes exact in the $N \to \infty$ limit), we can calculate the entropy at each value of time t. More precisely, we allow all the N points to evolve according to the map, which in turn determines the time evolution of the entropy. The $t = 0$ value of the BG entropy vanishes for all values of W, i.e., $S_1(0) = 0$. The N points will evolve until the attractor is achieved. The entropy $S_1(t)$ will evolve with t until arrival to a saturation value $S_1(\infty)$ which depends on (W, N); necessarily $\lim_{N\to\infty} S_1(\infty) < \ln\, W$ (this upper bound is attained if and only if we have uniform occupancy of the entire phase space, which is not the generic case, and certainly not the $a = 2$ logistic map case). In the $W \to \infty$ limit, the asymptotic growth of $S_1(t)$ is in fact linear (see, for instance, [205]), which enables the following characterization of the so called Kolmogorov-Sinai entropy:

$$K_1 = \lim_{t\to\infty} \lim_{W\to\infty} \lim_{N\to\infty} \frac{S_1(t)}{t} \; . \tag{159}$$

Quite generically, the Pesin *inequality* holds, i.e., (for one-dimensional nonlinear dynamical systems)

$$K_1 \le \lambda_1 \ \text{ if } \lambda_1 > 0 \tag{160}$$

and $K_1 = 0$ otherwise. The simplest situation occurs when there is no *escape* (see [9]). In this case the *equality* holds. Let us from now on address those particular systems for which this is true [9,206], i.e.,

$$K_1 = \lambda_1 \ \text{ if } \lambda_1 > 0 \tag{161}$$

and $K_1 = 0$ otherwise.

This type of analysis is convenient either if we have simple orbits (i.e., strong insensitivity to the initial conditions, i.e., for $\lambda_1 < 0$) or if we have strong chaos (i.e., strong sensitivity to the initial conditions, i.e., $\lambda_1 > 0$). But this analysis is a very poor one at say the edge of chaos, where $\lambda_1 = 0$ and we have *weak* sensitivity to the initial conditions. It is to unfold this type of situation that S_q becomes extremely useful. Let us show how. At the chaos threshold we have $K_1 = \lambda_1 = 0$. But if we follow the same procedure we just described for calculating K_1, but using instead $S_q(t)$, an interesting phenomenon can be revealed, which we describe now. Let us first define the following generalized Kolmogorov-Sinai entropy:

$$K_q = \lim_{t\to\infty} \lim_{W\to\infty} \lim_{N\to\infty} \frac{S_q(t)}{t} \ . \tag{162}$$

A value q^* is generically expected to exist [207] such that, for $q > q^*$ ($q < q^*$), $K_q = 0$ (K_q diverges), and K_{q^*} is *finite*! Furthermore, it can be argued [99] that the above Pesin equality can be generalized as follows:

$$K_{q^*} = \lambda_{q^*} \ \text{ if } \lambda_{q^*} > 0 \tag{163}$$

and $K_{q^*} = 0$ otherwise. It is through this important type of (in)equality that the connection emerges between S_q and the power-law time-dependence of the sensitivity to the initial conditions. The particular value q^* above described is what was numerically calculated in [99,202,203], and satisfies the scaling (156).

Some of the above statements can be trivially checked with the logistic map at its chaotic region (i.e., for $\lambda_1 > 0$). We know in that case that $S_1(t) \propto t$, hence (assuming that the scaling with time is essentially the same as for the simple case of equiprobability, for which $S_1 = \ln W$) the total number of possibilities $W(t)$ grows *exponentially* with t. For any $q > 1$, $S_q(t)$ is *always* bounded, then K_q necessarily vanishes. For any $q < 1$, $S_q(t)$ grows like the $1/(1-q)$ power of $W(t)$, which in turn, as said before, grows exponentially with t, hence necessarily $K_q \to \infty$. We conclude that $q^* = 1$! An analogous picture holds [208,209] for weak chaos, but with $q^* < 1$. See Fig. 26 and Fig. 27.

Summarizing, we have verified that, for maps like the logistic one at its edge of chaos, a special value of q, noted q^* can be found in at least *three different manners*. These are (i) through the sensitivity to the initial conditions (slope in the log-log plot of $\xi(t)$ versus t); (ii) through the multifractal function $f(\alpha_H)$ (measuring α_H^{min} and α_H^{max} and using Eq. (156)); (iii) through the time evolution of S_q (value of q for which, in the limit of very large W and even larger N, S_q increases linearly with t). This scenario is expected to be much more ubiquitous, appearing in a vast class of nonlinear dynamical systems, and not only in the logistic map.

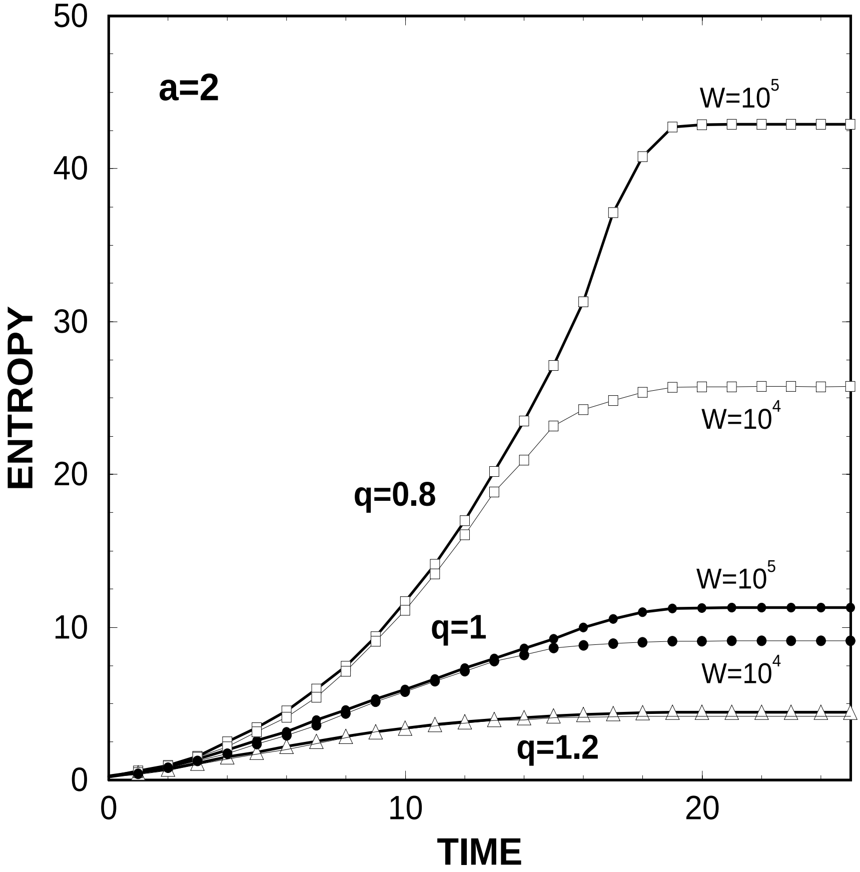

Fig. 26. Time dependence of S_q for the logistic map for the control parameter $a = 2$ (positive Liapunov exponent, i.e., *strong* chaos). The mathematical objects we are interested in emerge at the limit $W \to \infty$, where W is the number of intervals in our partition of the phase space. In that limit, the scenario which numerically emerges is that, for $q = 1$, $\lim_{t\to\infty}(S_q/t)$ is finite, whereas it vanishes for any $q > 1$ and diverges for any $q < 1$; we say that in this case $q^* = 1$. The fact that, in a situation like this one, $\lim_{t\to\infty}(S_1/t)$ is finite is well expected. This known result is now placed in a wider perspective.

At this point a clarification might be useful. It is quite common (and we have ourselves here inclined to this tendency) that the expression *Kolmogorov-*

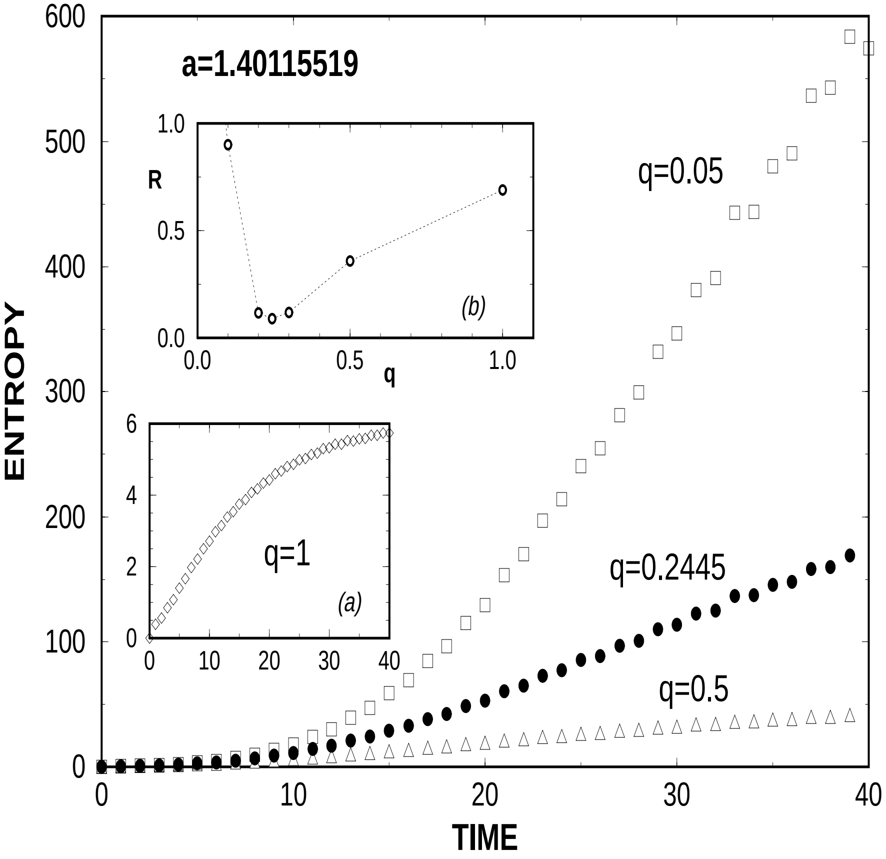

Fig. 27. Time dependence of S_q for the logistic map at the edge of chaos, i.e., for the control parameter $a = 1.40115519$ (zero Liapunov exponent, i.e., *weak* chaos). Once we have taken the $W \to \infty$ limit, the $\lim_{t\to\infty}(S_q/t)$ is now finite only for $q = q^* \equiv 0.2445...$; it vanishes for any $q > q^*$ and diverges for any $q < q^*$. For instance, see in one of the insets how S_1 is bending approaching saturation. In the other inset, a conveniently defined linearity coefficient R (strictly zero only when the curve is ideally linear) is depicted as a function of q: we verify that its minimum indeed occurs in the neighborhood of q^*. These results and those shown in Fig. 26 become unified within the present framework.

Sinai entropy (K) is indistinctly used for the case where the entropy is defined through a *single trajectory* visiting the elements of a partition of phase space (i.e., the time evolution associated with *one* initial condition, in a Boltzmann-type scenario), and for the case where the entropy is defined through a *distribution of probabilities* associated with the partition in phase space (i.e., the parallel time evolution of *many* initial conditions, in a Gibbs-type scenario). Although it is believed that almost always both constructions eventually lead to the same number for K, by all means this must be proved case by case, and the original definition of Kolmogorov-Sinai exclusively concerns the first definition, i.e., that using a single trajectory.

Let us briefly present a conjecture [210] concerning the more general situation where some escape does exist. Let us denote by γ_1 the *escape rate* whenever we are facing a *repeller*, instead of an attractor. This quantity is defined as follows. At time $t = 0$ we distribute (say homogeneously) $M(0) >> 1$ points in the phase space, and let them evolve according to a map which makes (for instance, through a *hole*) regularly disappear some of them. It is generically expected that the asymptotic time behavior of the number of points $M(t)$ still remaining in the system is given [9] by $M(t)/M(0) \sim e^{-\gamma_1 t}$. The repeller is defined as the set of all points in phase space that *never* disappear; it typically has zero measure. Then, Pesin inequality can be transformed into an equality, namely [9], $\gamma_1 = \lambda_1(repeller) - K_1(repeller)$. The Pesin equality is recovered for $\gamma_1 = 0$. Now, what happens whenever we are at a possible edge of chaos of the dynamical system (i.e., when $\lambda_1(repeller) = 0$)? It obviously becomes very natural to expect that a power-law replaces the exponential behavior we had for $\lambda_1 > 0$. Consistently, we expect now the escape to asymptotically follow something like $M(t)/M(0) \sim [1 - |(1-q)\gamma_q|t]^{-|1/(1-q)|}$, and we expect also something like $\gamma_q = \lambda_q(repeller) - K_q(repeller)$. The particular case $\gamma_q = 0$ recovers the structure of Eq. (163). The particular case $q = 1$ recovers the result known [9] for exponential escape. Analytical or numerical evidences along this conjecture would of course be extremely welcome.

Let us end this Section by a short remark which further clarifies the connections of the present formalism with notions currently used for chaotic systems. A possible generalization of the usual thermodynamical formalism for chaos (see [79]) has been recently proposed by Johal and Rai [59]. Following along their lines, we can define a quite general entropic form as follows:

$$S_{q,q'} \equiv \frac{\ln_{q'} \sum_i p_i^q}{1-q} \ . \tag{164}$$

We can straightforwardly verify that

$$S_{q,0} = \frac{1 - \sum_i p_i^q}{q-1} \ , \tag{165}$$

which is the entropy used here;

$$S_{q,1} = \frac{\ln \sum_i p_i^q}{1-q} \ , \tag{166}$$

which is Renyi's entropy;

$$S_{q,2} = \frac{1}{q-1}\left[\frac{1}{\sum_i p_i^q} - 1\right] \; , \tag{167}$$

which is the entropy discussed in [53,54,58]; and

$$S_{1/q,1+q} = \frac{1}{q-1}\left[1 - \frac{1}{(\sum_i p_i^{1/q})^q}\right] \; , \tag{168}$$

which is the (auxiliary) entropy introduced in [42] and discussed in detail in [56]. We also verify that $S_{1,q'}$ recovers the B-G entropy, $\forall q'$. However, let us make it clear that we do not attribute to definition (164) any special physical meaning. Were it not for the application (in order to set a nonextensive thermodynamical description of chaos) that, for this particular function of $\sum_i p_i^q$, is done in [59], we would consider it not much more than an unifying mathematical curiosity.

5.6 Self-Organized Criticality

In the previous example, fine tuning (e.g., $a = a_c$) is necessary to observe the anomalous ($q \neq 1$) behavior. Let us address dissipative systems with many degrees of freedom, very particularly those which do not need fine tuning. Would robust systems like those exhibiting self-organized criticality [211] (SOC) also present $q \neq 1$ behavior? The answer is *yes*, as it has been clearly exhibited [212] in at least three computational systems, namely the Bak-Sneppen model for biological evolution, the Suzuki-Kaneko model for imitation games and the Bak-Tang-Wiesenfeld model for sandpiles. In these systems, the Hamming distance plays the role played by ξ in the previous ones (i.e., the Hamming distance is proportional to $t^{1/(1-q)}$). Also, the relevance of the order of the $t \to \infty$ and $N \to \infty$ limits has been exhibited. Like in the conjectural Fig. 4, the $q \neq 1$ behavior is observed only in the $\lim_{t\to\infty} \lim_{N\to\infty}$ order. On what model ingredients does q depend? Is a taxonomy in universality classes analogous to that of standard critical phenomena possible? Is a relevant multifractal $f(\alpha_H)$ function hidden somewhere? Does a scaling law like that of Eq. (156) still hold? Additional computational effort is very welcome.

5.7 Long-Range Classical Hamiltonian Systems: Dynamic Properties

Let us finally focus on the "heart" of statistical mechanics, the dynamics of the systems on which Boltzmann himself was meditating, namely the Hamiltonian systems with many degrees of freedom. Although lots of interesting quantum nonextensive phenomena exist, here we shall restrict ourselves to the classical ones. We expect them to be able to provide nonextensive anomalies in a kind of pure, or simpler manner. Since a classical canonical Hamiltonian must satisfy the Liouville theorem, the Liapunov spectrum must be symmetric with respect to zero, the corresponding eigenvalues being necessarily coupled in pairs of positive and negative values with the same absolute value. Consequently, to study the sensitivity to the initial conditions it suffices to study the maximum Liapunov exponent. If it is positive, the system will generically be *strongly* chaotic, and will therefore easily satisfy the ergodic/mixing hypothesis (equality of time and ensemble averages). If, however, the maximum Liapunov exponent vanishes, the entire spectrum will necessarily vanish, hence the system will be, at most, *weakly* chaotic, and will therefore have difficulties in satisfying the ergodic/mixing hypothesis, at least before extremely large times (reflecting the macroscopic size of the system). The $d = 1$ coupled planar rotators N-body model with a two-body coupling constant proportional to $1/r^\alpha$ ($r \equiv$ distance between two given rotators) has been recently studied (for $\alpha = 0$ in [89,213], and, for $\alpha \geq 0$ in [214]) in the microcanonical ensemble. It has been established that, above a critical (conveniently normalized) total energy, the maximum (conveniently normalized) Liapunov exponent is, in the $N \to \infty$ limit, positive (zero) for $\alpha > 1$ ($0 \leq \alpha \leq 1$). More precisely, this maximum Liapunov exponent is proportional to $1/N^\kappa$ where $\kappa(\alpha)$ appears to be a monotonic function which decreases from $\kappa(0)$ to zero while α increases from 0 to 1, and remains zero for all $\alpha \geq 1$ (for $\alpha = 1$ it might well be that the maximum Liapunov exponent decreases like $1/\ln N$). It must be recalled that it is only for $\alpha > 1$ that the standard BG prescriptions provide *finite* integrals in the relevant calculations. Preliminary results are already available for the $d = 1$ classical Heisenberg-ferromagnet-like rotators (i.e., three-dimensional rotators instead of planar ones) [215] as well as for the $d = 1$ classical Lennard-Jones-like fluid [216]. In these three $d = 1$ models, the numerical results are well fitted by $\kappa = (1 - \alpha)/(3 - \alpha)$.

If we were to discuss the d-dimensional versions of the same models (Lennard-Jones-like fluids as well as coupled n-vector classical rotators, where $n = 2$ corresponds to the planar ones, $n = 3$ corresponds to the Heisenberg-like ones and $n \to \infty$ corresponds to the spherical ones) we would certainly have an α-dependence of κ. It is certainly possible that, within these classes of models, it is $\kappa(\alpha, d)$, i.e., the same for all values of n and for the fluid (and even for d-dimensional gravitation if the singularity at the origin is excluded by some relatively hard core). Heuristic arguments as well as some recent numerical results [217] for the $(d, n) = (2, 2)$ and $(3, 2)$ models suggest a fitting like the following one:

$$\kappa = \frac{1 + (\alpha/d)^2}{3 + (\alpha/d)^2} \quad (0 \leq \alpha/d \leq 1) \ . \tag{169}$$

This fitting clearly satisfies the trivial expectation that, for $\alpha = 0$, κ does not depend on d since all elements (rotators, particles) of the model are coupled to all other elements with the same strength. It also satisfies the recent exact result [218] for the $\alpha = 0$ case of the $(d, n) = (1, 2)$ model, namely that $\kappa = 1/3$.

In a recent paper, an essentially $\alpha = 0$ model was considered [219], and, under certain circumstances, a crossover was found between anomalous (at times smaller than $\tau(N)$) and normal (at times larger than $\tau(N)$) diffusion, with $\tau(N) \propto N$. What happens if $\alpha > 0$? Does τ scale like $\tilde{N} \equiv N^* + 1 = (N^{1-\alpha/d} -$

$\alpha/d)/(1-\alpha/d)$? What happens for other models? The behaviors observed for this particular model that was studied, as well as for the $\alpha = 0$ case of the $(d, n) = (1, 2)$ model [89], are consistent with Fig. 4. But is it exactly that conjecture that is going on? Only the study of the energy distributions (of single particles or of relatively large subsystems of the N-body system) themselves can provide the answer. What about the distributions of velocities? Are they Maxwellian (i.e., Gaussian) for $\alpha/d > 1$ and non-Maxwellian otherwise? Are they Lévy's or Student's t-distributions for $\alpha/d \leq 1$? If so, what is the dependence of $q(\alpha, d)$? Maybe $q(\alpha/d)$? Are the associated fluctuations anomalously time-correlated? Can nonmarkovian processes be present when the system is nonextensive (i.e., when $0 \leq \alpha/d < 1$)? Plenty of intriguing questions that, sooner or later, will have to be answered, partly through computational work (at least the first approaches). Better sooner than later!

Before ending this Section, let us mention that the scenario which apparently emerges for $0 \leq \alpha/d \leq 1$ is that the sensitivity to the initial conditions seems to be generically characterized by $q_S \leq 1$ (S stands for *sensitivity*; q_S is sometimes noted Q as well as q^* in the literature and in the present review), whereas the energy distribution seems to be generically characterized by $q_E \geq 1$ (E stands for *energy*). Naturally, we expect a universal relationship $q_E = g(q_S)$ such that $g(1) = 1$ (corresponding to $\alpha/d \geq 1$). This relationship probably is very simple; one such possibility is $q_E + q_S = 2$. Indeed, in this case, q_S varying from 1 to $-\infty$ would imply q_E varying from 1 to ∞, hence the exponent $q_E/(q_E - 1)$ associated with the escort distribution would be *larger* that unity (limit of integrability, up to infinity, of $1/\epsilon^{q_E/(q_E-1)}$). Also, $q_S = 0$ would correspond to the exponent $1/(q_E - 1) = 1$ (limit of integrability, up to infinity, of $1/\epsilon^{1/(q_E-1)}$).

5.8 Symbolic Sequences

An important progress has been recently done [220] by Grigolini and co-workers which, in addition to its intrinsic interest, provides an important insight on what is expected to occur for the long-range Hamiltonians just considered. The authors focused on a simple model which provides stochastic one-dimensional binary symbolic sequences. The symbols are sequentially chosen with a power-law correlation, more precisely the probability $p(l)$ of generating at the next step an l-sized string with the same symbol is assumed to be proportional to $1/l^{\alpha+2}$ ($\alpha \geq 0$). Then the q-generalized Kolmogorov-Sinai entropy K_q was numerically studied. This is to say the entropy S_q was studied as a function of "time" (subsequence fixed *width* to be run along the entire chain), and the special value q^* was determined for which S_q increases *linearly* with time (i.e., K_q is finite for $q = q^*$, vanishes for $q > q^*$ and diverges for $q < q^*$). Grigolini *et al.* exhibited the very eloquent fact that q^* monotonically increases from 0 to 1 when α increases from 0 to 1, and q^* remains equal to unity for all values of $\alpha \geq 1$, thus recovering well known results for short-range problems. Notice that the second moment $\langle l^2 \rangle \equiv \sum_{l=1}^{\infty} p(l)\, l^2$ is finite for $\alpha > 1$, whereas it diverges for $\alpha \leq 1$.

At this point it is convenient to make a few generic remarks. For the standard, extensive systems we have the following Kolmogorov-Sinai scenario. For strongly chaotic systems, the Liapunov spectrum contains *positive* exponents so that the number W of phase space visited regions grows exponentially with time t, i.e., $\ln W \propto t$, hence the entropy S_1, which scales as $\ln W$, increases *linearly* with time. Consequently, K_1 is finite and $q^* = 1$. In contrast with this picture, if the mixing is weak (power-law) because all Liapunov exponents are zero, we expect $W \propto t^{\delta}$ for long times, with $\delta > 0$. In this case we expect S_q to scale, for long times, like $[W^{1-q} - 1]/[1-q]$, i.e., like W^{1-q}, which is proportional to $t^{\delta(1-q)}$. Consequently, S_q increases *linearly* with time if and only if $q = q^*$ where q^* satisfies

$$1 - q^* = \frac{1}{\delta} \quad (\delta \geq 0) \; . \tag{170}$$

The corollary is that two typical behaviors are anticipated, namely (i) if δ varies, for a class of systems, between 0 and ∞, then q^* varies from $-\infty$ to 1 (as for the logistic-like maps, examined in subsection 5.5); (ii) if δ varies, for another class of systems, between 1 and ∞, then q^* varies from 0 to 1 (as for the model of symbolic sequences described in the present subsection).

5.9 Optimization Techniques; Simulated Annealing

The so called *Optimization problem* consists basically in determining the *global* minimum (or minima, if degeneracy is present) of a given cost function $E(x)$, where x is a discrete or continuous d-dimensional variable. This problem can become extremely complex depending on the cost function having a large number of *local* minima, and on the dimension d being high. For the ubiquitous cases (in physics, chemistry, neural networks, engineering, finances, etc.) for which analytic discussion is not tractable, a variety of computational algorithms have been developed. A special place among these is occupied, because of its efficiency and paradigmatic value, by the *Simulated Annealing* (SA) introduced in 1983 by Kirkpatrick *et al.* [221]. Its denomination comes from its total analogy with the well known annealing technique, frequently used in Metallurgy for making a molten metal to reach its crystalline state (global minimum of the relevant thermodynamic energy). In SA, one or more artificial temperatures are introduced and gradually cooled, acting as a source of stochasticity, extremely convenient for eventually detrapping from local minima. Near the end of the process, the system hopefully is in the attractive basin of one of the global minima, the temperature is practically zero, and the algorithm asymptotically becomes a steepest descent one. The challenge is to cool the temperature the quickest we can *but* still having the guarantee that no definitive trapping in any local minimum will occur. More precisely speaking, we search for the quickest annealing (i.e., in some sense approaching a *quenching*) which preserves the probability of ending in a global minimum being equal to one. SA strictly follows a BG scheme. Let us illustrate for continuous x. The system "tries" to visit, according to a *visiting distribution* assumed to be Gaussian in the neighborhood of its actual state. The jump is always accepted if it is "downhill", i.e., if the cost function decreases. If it is "uphill", the jump *might* be accepted with a probability given by the Boltzmann factor corresponding to that cost function. It has been shown that the probability of ending on the global minimum equals unity if $T(t)$ decreases logarithmically with time t. This algorithm is sometimes referred to as *Classical Simulated Annealing* (CSA) or Boltzmann machine. We easily recognize that, if instead of decreasing, the temperature was maintained fixed, this procedure precisely would be the well known Metropolis *et al.* one for simulating BG thermostatistical equilibrium.

This optimization machine has been generalized within the present statistics as follows [118]. The visiting distribution is generalized to be a q_V-Gaussian, and the acceptance Boltzmann factor is generalized to be a q_A-generalized factor, where q_V and q_A respectively are the *visiting* and *acceptance* entropic indexes. The cooling schedule is generalized as follows:

$$T(t) = T(1)\, \frac{\ln_q\,[1/2]}{\ln_q\,[1/(t+1)]} \qquad (t = 1,\ 2,\ 3,\ ...)\ . \tag{171}$$

This is the *Generalized Simulated Annealing.* This machine is characterized by (q_V, q_A). The choice $(1,1)$ corresponds to CSA, and the choice $(2,1)$ corresponds to the so called *Fast Simulated Annealing* (FSA). The CSA corresponds to a cooling given by

$$T(t) = T(1)\, \frac{\ln 2}{\ln(1+t)} \qquad (t = 1,\ 2,\ 3,\ ...)\ . \tag{172}$$

The FSA corresponds to a faster cooling given by

$$T(t) = T(1)\, \frac{1}{t} \qquad (t = 1,\ 2,\ 3,\ ...)\ . \tag{173}$$

The limiting case $q_V = 3$ corresponds to

$$T(t) = T(1)\, \frac{3}{(t+1)^2 - 1} \qquad (t = 1,\ 2,\ 3,\ ...)\ . \tag{174}$$

These particular cases illustrate the great computational advantage that can be obtained by speeding up the algorithm by conveniently choosing q_V (see also [128]). In practice, a convenient choice for q_V is somewhat below 3, the exact value depending on the phase space dimension d. The choice of q_A seems to be more model-dependent. Details can be seen in a by now vast literature [119–126], in which applications have been done and variations have been performed concerning a variety of classical and quantum physical problems, the Traveling Salesman Problem, and many others. The first application [127] in quantum chemistry concerned simple molecules of the series $CH_3 - R$ and some others, including the $H_2 0_3$ one, by using the MOPAC program package. Nowadays, Straub (in Boston), Okamoto (in Okazaki), Bisch (in Rio de Janeiro) and Ellis-Mundim (respectively in Chicago and Salvador), to name but a few, are currently improving and applying these techniques to complex molecules such as polypeptides, in particular with the aim of studying the important, though hard, *protein folding* problem. The quickly expanding bibliography in these optimization applications of nonextensivity is included in [2].

6 Final Remarks

Boltzmann-Gibbs statistical mechanics and standard thermodynamics do not seem to be universal. They have domains of applicability quite poorly known nowadays. The precise knowledge of the restrictions for their validity is conceptually and practically very important. A nonextensive generalization of these formalisms is now available [1,41,42]; see Table 1 (Appendix). It has been developed to cover at least some of these difficulties. Several important types of systems have been focused on in the present paper, which should substantially clarify the situation. These efforts span a wide epistemological variety, which goes from clear-cut theories to phenomenological ones, to quite well or less well understood fittings and connections. As exhibited at length here, the areas on which this formalism has been satisfactorily applied includes physics, astronomy, chemistry, mathematics, biology, economics, linguistics, cognitive psychology, etc. However, in spite of all this sensible progress, some inter-related crucial points are still to be understood and established on a neat and transparent basis. These include
(i) the zeroth principle of thermodynamics and its connections with the thermodynamic limit, properties which would in principle exhibit the mathematical connection of the *weak* violation of the BG statistics (i.e., introduction of $\tilde{N}$ within the $q = 1$ formalism) with the *strong* one ($q \neq 1$);
(ii) the functional dependence of q on (α, d) for long-range interacting Hamiltonians ($0 \leq \alpha/d \leq 1$) and its connection with anomalous diffusion. Why for these systems universality classes seem to emerge whose structure greatly resembles that of traditional critical phenomena (in the sense that q depends only on a few structural ingredients), whereas for other systems, such as the electron-positron annihilation experiments (and similar high energy reactions) and turbulence, q *continuously* varies with some energy-like variables (e.g., the collision energy W for the electron-positron experiments, r/η for turbulence)? The latter clearly remind what occurs in the $d = 2$ XY model where, below the Kosterlitz-Thouless temperature, the critical exponents are known to continuously depend on temperature;
(iii) the physical interpretation of the $\{p_i\}$ distribution and of the escort one $\{P_i\}$, as well as clear-cut prescriptions for using one or the other when fitting experimental data (in the meanwhile, it appears that $\{P_i\}$ is the one to be used for equilibrium distributions of kinetic, potential or total energies, whereas $\{p_i\}$ is the one to be used for real-space diffusive nonequilibrium phenomena, either Lévy or Student's [222] distributions according to whether correlations between jumps are suspected to be absent or present);
(iv) what are the generic physical conditions for using the microcanonical, canonical and grand-canonical ensembles, under what exact and fully specified conditions they are expected to be thermodynamically equivalent, and the possible relevance for the so called thermogravitational catastrophe;

(v) the clear connection with microscopic dynamic properties such as (partial) lack of ergodicity and mixing, the generalization (to weak chaos) of the Pesin inequality, the complete domain of validity of scaling relations connecting q to multifractality, and the possible relevance for SOC, spin-glasses and similar phenomena;
(vi) the clear physical connection with quantum groups and, in general, deformations of relevant Lie algebras, and through these, the possible relevance for quantum gravity and the deep (possibly discrete, multifractal- like) structure of space-time;
(vii) the clear physical meaning of the connections that are nowadays emerging between nonextensivity and quantum entanglement. Indeed, a *pure state* $|\psi\rangle$ of a quantum system composed by subsystems A and B is said to be entangled if it cannot be written in the tensor product form $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ (see, for instance, [223]). Consistently, we consider a *statistical mixture* characterized by the density operator ρ to be entangled if it cannot be written as $\rho = \rho_A \otimes \rho_B$. This is what opens the door to consider that there is entanglement whenever $S_q(A+B) \neq S_q(A)+S_q(B)+(1-q)S_q(A)S_q(B)$. What happens when the number of subsystems is very large?, and when each of them has a very large number of elements? We have seen in subsection 5.5 that a *special*, and unique, value of q exists such that, if the initial state of the nonlinear dynamical system is appropriately prepared, the entropy asymptotically increases *linearly* with time. Could we expect something similar to this with regard to the number of particles N in the system? Would these considerations be useful for understanding entanglement within the system, which could nevertheless be disentangled with the rest of the universe? Answering these questions constitutes to our eyes an interesting research project! Also, perhaps, very useful for quantum computation, teleportation, cryptographic applications and related matters.

A practical warning might be worthy at this point. Whenever the reader wants to compare values of q obtained for different specific systems by different authors, a careful inspection of the corresponding papers is recommended. Indeed, since along this decade three (slightly) different versions of the formalism were introduced (see [42]), one must be careful in checking the versions that those particular papers are using. If we call q_I, q_{II} and q_{III} the entropic indexes appearing respectively in versions I (standard mean values), II (unnormalized q-expectation values) and III (normalized q-expectation values), the following relations hold whenever velocity, energy (or similar) distributions (e.g., p_q) are being focused on:

$$q_I - 1 = 1 - q_{II} = 1 - q_{III} \ . \tag{175}$$

For the corresponding escort distributions (e.g., proportional to p_q^q, a concept which was not present in version I), $q_{II}/(1-q_{II}) = q_{III}/(1-q_{III})$ play the role of $1/(q_I-1)$, hence $q_{II} = q_{III} = 1/q_I$. The connections between the variables which play a temperature role, so to say T_I, T_{II} and T_{III}, are somewhat more subtle (see [42]). Of course, this warning does not apply to whatever *only* involves the entropy, without any reference to physical constraints introduced through mean values For instance, it does not apply to the sensitivity to the initial conditions and related concepts such as the multifractal spectrum $f(\alpha_H)$ and the present generalization of the Kolmogorov-Sinai entropy (see also discussion about q_S and q_E in subsection 5.7).

On general grounds, one might think of two conjectural contexts, to be clarified (i.e., rigorously formulated), confirmed or refuted. The first of these contexts can somehow (on intuitive grounds) be formulated as follows. *Strongly* mixing phenomena are ubiquitous in Nature; essentially, they are driven by microscopic interactions which are short-ranged in space-time (short-range forces, short-range memory, nonfractal boundary conditions); their basic geometry tends to be continuous, smooth, like the Euclidean one; their thermodynamics is extensive; their central laws (energy distribution at equilibrium, time-relaxation towards equilibrium) generically are exponentials; and their thermo statistical foundation lies within Boltzmann- Gibbs statistics (i.e., $q = 1$). But *weakly* mixing phenomena also are ubiquitous in Nature (e.g., biological, socio-economical, human cognitive phenomena, etc.); essentially, they are driven by microscopic interactions which are long-ranged in space time (long-range forces, nonmarkovian memory, fractal boundary conditions); their basic geometry tends to be discrete, multifractal-like; their thermodynamics is nonextensive; their central laws (energy distribution at equilibrium, time-relaxation towards equilibrium) generically are power-laws; and the thermo statistical foundation of (at least some of) them (hopefully!) lies within the $q \neq 1$ statistics. The allowance for nonextensivity, in general, and for a nonextensive entropy, in particular, appears to be the "price" to be paid in order that Boltzmann's "mechanical" (i.e., *one* system evolving along time) manner of thinking about macroscopic systems coincides, at the level of the concrete mathematical results to be compared with the experimental data, with Gibbs' "ensemble" (i.e., *many* systems at a fixed time) manner of thinking. This coincidence of results is, since one century, well known and understood for standard systems. Our aim here is to extend it to a large variety of anomalous systems.

In other words, this conjecture concerns the following logical chain. Whenever we have *quick* sensitivity to the initial conditions (i.e., positive Liapunov exponents), on one hand the relaxation towards equilibrium of a generic physical observable typically occurs exponentially in time (and the inverse relaxation time essentially scales with the largest Liapunov exponent), and on the other hand the entropy which is useful is the BG one (hence *extensive* with regard to independent systems). This entropy naturally yields, for Hamiltonian systems, to an energy distribution which is exponential (Boltzmann factor); this in turn implies, if the system is classical, that the distribution of velocities is Maxwellian (i.e., Gaussian). This is the kingdom of the exponential function! But, generically, whenever we have *slow* sensitivity to the initial conditions (i.e., zero Liapunov exponents), on one hand the relaxation towards equilibrium of a generic physical observable typically occurs algebraically in time (and the inverse relaxation time essentially scales with the largest *generalized* Liapunov exponent), and on the other hand the entropic form which appears to be useful is S_q (hence *nonextensive* with regard to independent systems). This entropy naturally yields, for Hamiltonian systems, to an energy distribution which is a power-law; this in turn implies, if the system is classical, that the distribution of velocities is not Maxwellian (i.e., non Gaussian; perhaps a Lévy distribution or a Student's t-distribution). This is the kingdom of the power function!

The second of these conjectural contexts is, at the present moment, so hard to rationalize that I dare to mention it here *only* because, after having been exposed to so many mathematical and physical arguments (that have been included in the present review), the reader might accept to honor me with his (her) indulgence, and have a look at the following few, intuitive lines. I believe that a deep analogy (maybe a kind of isomorphism, through the use of mathematical structures like the co-homology groups) exists between crystallographic structures such as crystal - quasicrystal - fluid, and nonlinear dynamics such as integrable - (weak) chaotic - (strong) chaotic. In some sense, they appear as space and time versions of the same mathematical structures. The first case concerns crystals (i.e., d-dimensional Bravais lattices) and integrable dynamics (i.e., motion on simple orbits), and its essential invariance is the *discrete translational* one. The third case concerns strongly disordered systems like fluids (liquid, gases) and strongly chaotic dynamics, and its essential invariance is the *continuous translational* one. Finally, the second, and intermediate, case is by far the most subtle one (and probably this is why it is the one that humanity took the longest time to discover), and concerns quasicrystals (e.g., Penrose tilings, amorphous substances like glasses, spin-glasses, and other structures known to have (multi)fractal scalings; probably most of the so called complex spatial phenomena belong to this group) and weakly ergodic dynamics (e.g., edge of chaos, strange attractors, self-organized criticality, probably most of the so called complex time phenomena); its essential invariance is the *dilatation* one. In the first case we have the (space or time) *highest predictability*, and statistical methods are out of place. In the third case we have the (space or time) *lowest predictability*, and statistical methods exhibit their full power. Finally, in the second case, we have an *intermediate predictability*, and the statistical methods have to be "intrinsically nonlinear" in some sense, in order to be applicable and useful. There will be no surprise for the reader if, at this point, I admit that I believe that the statistical mechanics to be associated with the third present case of course is the BG one, whereas it might be the $q \neq 1$ statistical mechanics the one to be associated with the present second case ! (see also [224], where some preliminary, but nevertheless concrete, calculations exhibit this kind of connections). Let us now remind that Wiles' 1995 celebrated proof [225] of Fermat's last theorem was deeply related to quasicrystals since it was based on the proof of the Taniyama-Shimura conjecture about modular elliptic curves and used certain Hecke algebras. Consequently, I hardly dare to explicitly state a simple and unavoidable corollary, namely that, if my present second conjecture turns out to be, in some nontrivial and precise sense, correct, then the $q \neq 1$ statistical mechanics must be related to Fermat's last theorem!

Through the complete analysis, in more detailed terms, of the various aspects tackled in the present review, we should be able to learn a lot and, very especially, *(precisely) when* the celebrated Boltzmann factor is the correct theoretical description of natural systems at thermal equilibrium! This famous, ubiquitous, and so useful factor could then cease to be a "dogma", as referred to by Takens [5], and could climb to the status of a theorem!

Acknowledgements

Along the years, I have had extremely fruitful discussions with, and received enlightening comments from S. Abe, C. E. Aguiar, F. C. Alcaraz, C. Anteneodo, P. Bak, M. Baranger, G. Bemski, B. M. Boghosian, E. P. Borges, L. Borland, M. O. Caceres, S. A. Cannas, A. Coniglio, A. Craievich, E. M. F. Curado, D. J. Driebe, P. M. Duxbury, J.-P. Eckmann, A. Erzan, R. Ferreira, M. E. Fisher, M. Gell-Mann, P. Grigolini, G. R. Guerberoff, H. J. Herrmann, T. A. Kaplan, D. H. Kobe, T. Kodama, P. T. Landsberg, V. Latora, E. K. Lenzi, K. Loerintz, M. L. Lyra, A. C. N. de Magalhaes, S. D. Mahanti, R. Maynard, R. S. Mendes, H. N. Nazareno, Y. Okamoto, I. Oppenheim, A. Overhauser, T. J. P. Penna, A. Plastino, A. R. Plastino, D. Prato, I. Procaccia, P. Quarati, A. K. Rajagopal, A. Rapisarda, N. Scafetta, D. A. Stariolo, J. E. Straub, F. A. Tamarit, U. Tirnakli, L. Tisza, B. Widom, H. S. Wio, D. H. Zanette, and so many others that it would be an almost impossible task to name them all. I heartily acknowledge them all. Finally, I am very grateful to A. Upadhyaya, J. P. Rieu, J. A. Glazier and Y. Sawada for making their results available to me prior to publication. This work was partially supported by CNPq, PRONEX/FINEP and FAPERJ (Brazilian Agencies).

Appendix: *q*-Exponential and *q*-Logarithm Functions

Let us define as follows the *q-exponential* function [69,70]:

$$e_q^x \equiv [1+(1-q)\,x]^{\frac{1}{1-q}} \qquad (x \in \mathcal{R};\ q \in \mathcal{R}) \ . \tag{176}$$

Its inverse, *q-logarithm* function is given by

$$\ln_q\, x \equiv \frac{x^{1-q}-1}{1-q} \qquad (x \in \mathcal{R}^+;\ q \in \mathcal{R}) \ . \tag{177}$$

We easily verify that

$$e_1^x \equiv \lim_{q\to 1+0} e_q^x = \lim_{q \to 1-0} e_q^x = e^x \ \ (\forall x) \ , \tag{178}$$

$$\ln_1\, x \equiv \lim_{q\to 1+0} \ln_q\, x = \lim_{q \to 1-0} \ln_q\, x = \ln\, x \ \ (\forall x) \ , \tag{179}$$

as well as that

$$e_q^{\ln_q\, x} = \ln_q\, e_q^x = x \ \ (\forall x;\ \forall q) \ . \tag{180}$$

Let us be more detailed about the above definitions. For $q<1$, the q-exponential function vanishes for $x \leq -1/(1-q)$ and continuous and monotonically increases from 0 to ∞ when x increases from $-1/(1-q)$ to ∞. For $q>1$, the q-exponential function continuous and monotonically increases from 0 to ∞ when x increases from $-\infty$ to $1/(q-1)$, remaining divergent for $x > 1/(q-1)$. See Fig. 28.

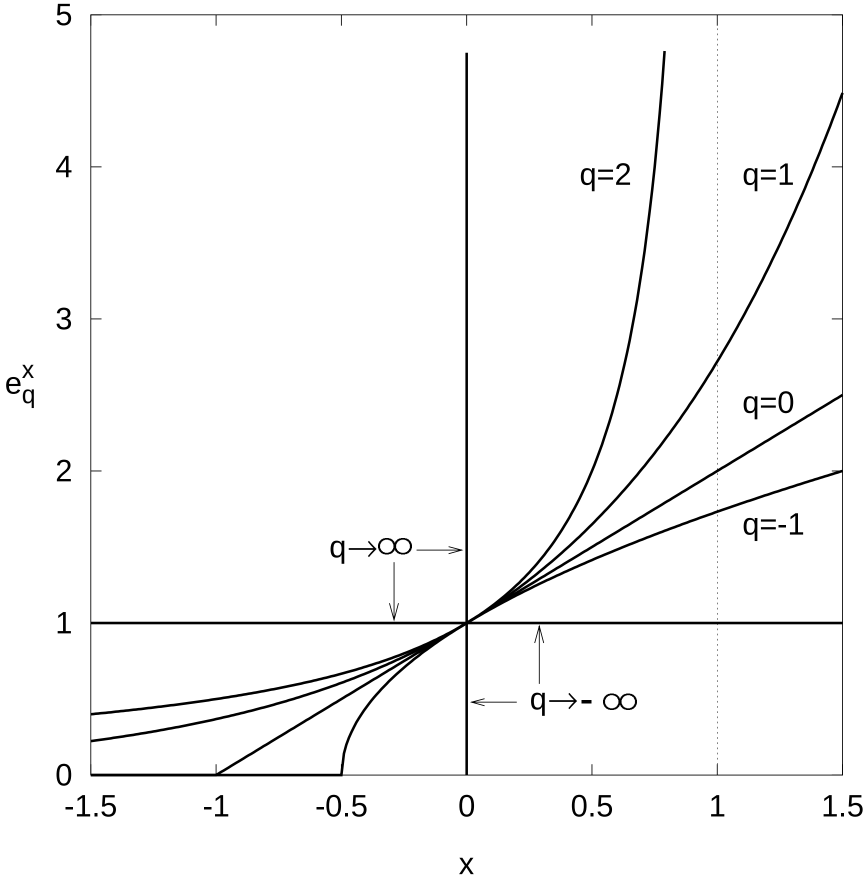

Fig. 28. The q-exponential function for typical values of q. For all $q < 1$ the function vanishes for $x \leq -1/(1-q)$; for all $q > 1$ the function diverges at $x = 1/(q-1)$ (the dotted line indicates the vertical asymptote for $q = 2$).

The following property is satisfied:

$$[e_{1/q}^{-qx}]^{1/q} = \frac{1}{e_q^x} \quad (\forall x;\ \forall q) \ , \tag{181}$$

or, equivalent and more symmetrically,

$$[e_q^{x/q^{1/2}}]^{q^{1/2}} \ e_{1/q}^{-q^{1/2}x]^{1/q^{1/2}}} = 1 \quad (\forall x;\ \forall q) \ , \tag{182}$$

as well as their corresponding inverse ones, namely

$$-\frac{1}{q}\ln_{1/q}(1/x^q) = \ln_q x \quad (\forall x;\ \forall q)\ , \tag{183}$$

and, equivalently and more symmetrically,

$$q^{1/2}\ln_q x^{\frac{1}{q^{1/2}}} + \frac{1}{q^{1/2}}\ln_{1/q}(1/x^{q^{1/2}}) = 0 \quad (\forall x;\ \forall q)\ . \tag{184}$$

Also,

$$e_{2-q}^{-x} = \frac{1}{e_q^x} \quad (\forall x;\ \forall q)\ , \tag{185}$$

or, equivalent and more symmetrically,

$$e_{1+q}^{x}\, e_{1-q}^{-x} = 1 \quad (\forall x;\ \forall q)\ , \tag{186}$$

as well as their corresponding inverse ones, namely

$$-\ln_{2-q}(1/x) = \ln_q x \quad (\forall x;\ \forall q)\ , \tag{187}$$

and, equivalently and more symmetrically,

$$\ln_{1+q} x + \ln_{1-q}(1/x) = 0 \quad (\forall x;\ \forall q)\ . \tag{188}$$

Other interesting properties are:

$$e_q^{x+y+(1-q)xy} = e_q^x\, e_q^y \quad (\forall (x,y);\ \forall q)\ , \tag{189}$$

and

$$\ln_q(xy) = \ln_q\ x + \ln_q\ y + (1-q)(\ln_q\ x)(\ln_q\ y) \quad (\forall (x,y);\ \forall q)\ , \tag{190}$$

as well as

$$\frac{d}{dx}e_q^x = (e_q^x)^q \quad (\forall x\ \forall q)\ , \tag{191}$$

$$\frac{d}{dx}\ln_q x = \frac{1}{x^q} \quad (\forall x\ , \forall q)\ , \tag{192}$$

$$\ln_q(\frac{x}{y}) = \frac{1}{y^{1-q}}[\ln_q\ x - \ln_q\ y] \quad (\forall (x,y);\ \forall q)\ , \tag{193}$$

$$\ln_q \frac{1}{x} = -\frac{1}{x^{1-q}}\ln_q x \quad (\forall x,\ \forall q)\ . \tag{194}$$

The following expansions are useful in a variety of applications:

$$e_q^x = 1 + x + \frac{q}{2}x^2 + \frac{q(2q-1)}{6}x^3 + \frac{q(2q-1)(3q-2)}{24}x^4 + ... \\ + \frac{1}{n!}\Big[\Pi_{i=1}^{n-1}(iq-i+1)\Big]x^n + ... \quad (\forall q;\ x \to 0)\ , \tag{195}$$

$$\ln_q(1+x) = x - \frac{q}{2}x^2 + \frac{q(q+1)}{6}x^3 - \frac{q(q+1)(q+2)}{24}x^4 + ...$$
$$+ \frac{(-1)^{n+1}}{n!}\Big[\Pi_{i=0}^{n-2}(q+i)\Big]x^n + ... \quad (\forall q;\ x \to 0)\ , \tag{196}$$

$$e_q^x = e^x\Big[1 + \frac{x^2}{2}(q-1) + \frac{8x^3+3x^4}{24}(q-1)^2$$
$$+ \frac{12x^4+8x^5+x^6}{48}(q-1)^3 + ...\Big] \quad (\forall x;\ q \to 1)\ . \tag{197}$$

Finally, the following integral formula can be applied in a variety of circumstances (it shows [226] in particular that *renormalizability* can occur without *factorizability*):

$$\int_{-\infty}^{\infty} dx_2\ \ e_q(-a_1|x_1|^{z_1} - a_2|x_2|^{z_2})$$
$$= \ A_2\, e_{q'}(-a_1'|x_1|^{z_1}) \quad (a_1, a_2, z_1, z_2 > 0)\ , \tag{198}$$

where

$$\frac{1}{1-q'} = \frac{1}{1-q} + \frac{1}{z_2}\ , \tag{199}$$

$$a_1'(1-q') = a_1(1-q)\ , \tag{200}$$

and

$$A_2 = \int_{-\infty}^{\infty} d\xi\ e_q(-a_2|\xi|^{z_2}) = \frac{2\,\Gamma\left(1+\frac{1}{z_2}\right)}{a_2^{1/z_2}}$$
$$\times \begin{cases} \frac{\Gamma\left(1+\frac{1}{1-q}\right)}{(1-q)^{1/z_2}\,\Gamma\left(1+\frac{1}{z_2}+\frac{1}{1-q}\right)} & \text{for } q<1 \\ 1 & \text{for } q=1 \\ \frac{\Gamma\left(\frac{1}{q-1}-\frac{1}{z_2}\right)}{(q-1)^{1/z_2}\,\Gamma\left(\frac{1}{q-1}\right)} & \text{for } q>1 \end{cases}\ . \tag{201}$$

Many other elegant properties concerning the generalization of the geometric and hyperbolic functions [70] as well as of Laplace transform [227] are available in the literature.

As we can see in Table 1, the present functions and their properties enable the expression of important relationships in a convenient Boltzmann-Gibbs-like form.

Table 1. Some useful formulae written in a Boltzmann-Gibbs-like form

Equiprobability entropy	$S_q = k \ln_q\ W$
Generic entropy	$S_q = -k \langle \ln_q \rho \rangle_q$
Canonical equilibrium distribution	$\rho_q = \frac{e_q^{-\beta(\mathcal{H}-U_q)/\mathrm{Tr}\rho_q^q}}{\mathrm{Tr}\ e_q^{-\beta(\mathcal{H}-U_q)/\mathrm{Tr}\rho_q^q}} = \frac{e_q^{-\beta' \mathcal{H}}}{\mathrm{Tr} e_q^{-\beta' \mathcal{H}}}$ $(\beta' \equiv \frac{\beta}{\mathrm{Tr}\rho_q^q + (1-q)\beta U_q})$
Partition functions	$\bar{Z}_q = \mathrm{Tr}\ e_q^{-\beta(\mathcal{H}-U_q)/\mathrm{Tr}\rho_q^q}$ $(\ln_q Z_q = \ln_q \bar{Z}_q - \beta U_q)$
Internal energy	$U_q = -\frac{\partial}{\partial \beta} \ln_q Z_q$
Free energy	$F_q = U_q - T S_q = -\frac{1}{\beta} \ln_q Z_q$
Anomalous diffusion probability distribution	$p_q(x) = \frac{e_q^{-\beta x^2}}{\int dy\ e_q^{-\beta y^2}}$
Sensitivity to the initial conditions ($d = 1$)	$\lim_{\Delta x(0) \to 0} \frac{\Delta x(t)}{\Delta x(0)} = e_q^{\lambda_q\ t}$
Rate of increase of entropy	$\lim_{t \to \infty} \frac{S_q(t)}{t} = \lambda_q$
Likelihood function	$W_q(\{p_i\}) \propto e_q^{S_q(\{p_i\})}$
Power-law interactions ($\propto R^{-\alpha}$)	$\frac{U(N,T)}{N \tilde{N}} \sim u(\frac{T}{\tilde{N}})$ $(\tilde{N} \equiv N^* + 1;\ N^* \equiv \ln_{[\alpha/d]} N)$
Simulated annealing (cooling rythm)	$\frac{T(t)}{T(1)} = \frac{\ln_q\ [1/2]}{\ln_q\ [1/(t+1)]}$

References

1. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
2. See http://tsallis.cat.cbpf.br/biblio.htm for a regularly updated bibliography on the subject.
3. C. Tsallis, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 1 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm]. The present review is an extended and updated version of the one just quoted.
4. With *somewhere*, we are by no means restricting to *local* phenomena; in particular, quantum or other nonlocalities are of course not excluded. Analogous considerations are applicable to the use of the expression *some time*.
5. F. Takens, *Structures in dynamics - Finite dimensional deterministic studies*, eds. H. W. Broer, F. Dumortier, S. J. van Strien, and F. Takens (North-Holland, Amsterdam, 1991), p. 253. [In his words: "The values of p_i are determined by the following dogma: if the energy of the system in the i^{th} state is E_i and if the temperature of the system is T then: $p_i = \exp\{-E_i/kT\}/Z(T)$, where $Z(T) = \sum_i \exp\{-E_i/kT\}$, (this last constant is taken so that $\sum_i p_i = 1$) . This choice of p_i is called *Gibbs distribution*. We shall give no justification for this dogma; even a physicist like Ruelle disposes of this question as " deep and incompletely clarified"."]
6. N. Krylov, Nature **153**, 709 (1944). [In his words: " In the present investigation, the notion of ergodicity is ignored. I reject the ergodical hypothesis completely: it is both insufficient and unnecessary for statistics. I use, as starting point, the notion of motions of the mixing type, and show that the essential mechanical condition for the applicability of statistics consists in the requirement that in the

phase space of the system all the regions with a sufficiently large size should vary in the course of time in such a way that while their volume remains constant – according to Liouville's theorem – their parts should be distributed over the whole phase space (more exactly over the layer, corresponding to given values of the single-valued integrals of the motion) with a steadily increasing degree of uniformity. (...) The main condition of mixing, which ensures the fulfillment of this condition, is a sufficiently rapid divergence of the geodetic lines of this Riemann space (that is, of the paths of the system in the n-dimensional configuration space), namely, an exponential divergence (cf. Nopf[1]). "]. For full details on this pioneering approach see N.S. Krylov, *Works on the Foundations of Statistical Physics*, translated by A. B. Migdal, Ya. G. Sinai, and Yu. L. Zeeman, Princeton Series in Physics (Princeton University Press, Princeton, 1979).

7. R. Balescu, *Equilibrium and Non-equilibrium Statistical Mechanics* (Wiley, New York, 1975), p. 727. [In his words: "It therefore appears from the present discussion that the *mixing property of a mechanical system is much more important for the understanding of statistical mechanics than the mere ergodicity.* (...) A detailed rigorous study of the way in which the concepts of mixing and the concept of large numbers of degrees of freedom influence the macroscopic laws of motion is still lacking."]
8. M. Courbage and D. Hamdan, Phys. Rev. Lett. **74**, 5166 (1995).
9. J.R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*, Cambridge Lecture Notes in Physics **14**, eds. P. Goddard and J. Yeomans (Cambridge University Press, Cambridge, 1999), footnote on p. 9. [In his words: "It is worth mentioning that there are examples of mixing systems with no non-zero Liapunov exponents. The concepts of ergodicity, mixing, and chaos can be quite subtle."]
10. N. G. van Kampen, Braz. J. Phys. **28**, 90 (1998) [accessible at http://www.sbf.if.usp.br/WWW_pages/Journals/BJP/Vol28/Num2/index.htm].
11. G. M. Zaslavsky, Physics Today **52**, 39 (August 1999).
12. Private discussion with M. Gell-Mann.
13. L. Tisza, Ann. Phys. **13**, 1 (1961). [Or, in *Generalized Thermodynamics* (MIT Press, Cambridge, 1966), p. 123]. [In his words: "The situation is different for the additivity postulate P a2, the validity of which cannot be inferred from general principles. We have to require that the interaction energy between thermodynamic systems be negligible. This assumption is closely related to the homogeneity postulate P d1. From the molecular point of view, additivity and homogeneity can be expected to be reasonable approximations for systems containing many particles, provided that the intramolecular forces have a short range character."]
14. H. Grad, in *Rarefied Gas Dynamics*, ed. J. A. Laurmann (Academic Press, New York, 1963), p. 26. [A variety of anomalous effects on the spectrum of the collision operator of the Boltzmann equation, due to the range of the forces, are discussed in this article.]
15. M. E. Fisher, Arch. Rat. Mech. Anal. **17**, 377 (1964); J. Chem. Phys. **42**, 3852 (1965); J. Math. Phys. **6**, 1643 (1965); M. E. Fisher and D. Ruelle, J. Math. Phys. **7**, 260 (1966); M. E. Fisher and J. L. Lebowitz, Commun. Math. Phys. **19**, 251 (1970).
16. P.T. Landsberg, *Thermodynamics and Statistical Mechanics* (Oxford University Press, Oxford, 1978; also Dover, New York, 1990), p. 102. [In his words: "The presence of long-range forces causes important amendments to thermodynamics, some of which are not fully investigated as yet."]

17. W. Thirring, Foundations of Physics **20**, 1103 (1990).
18. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981), footnote on p. 114. [In his words: "Actually an additional stability criterion is needed, see M. E. Fisher, Archives Rat. Mech. Anal. **17**, 377 (1964); D. Ruelle, *Statistical Mechanics, Rigorous Results* (Benjamin, New York 1969). A collection of point particles with mutual gravitation is an example where this criterion is not satisfied, and for which therefore no statistical mechanics exists."]
19. P. T. Landsberg, J. Stat. Phys. **35**, 159 (1984); L. G. Taff, *Celestial Mechanics* (John Wiley and Sons, New York, 1985), p. 437; W. C. Saslaw, *Gravitational Physics of Stellar and Galactic Systems* (Cambridge University Press, Cambridge, 1985), p. 217; J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, 1987), p. 267; D. Pavon, Gen. Rel. and Gravit. **19**, 375 (1987); H. E. Kandrup, Phys. Rev. A **40**, 7265 (1989); H. S. Robertson, *Statistical Thermophysics* (Prentice Hall, Englewood Cliffs, New Jersey, 1993), p. 96.
20. B. J. Hiley and G. S. Joice, Proc. Phys. Soc. **85**, 493 (1965).
21. G. Kotliar, P. W. Anderson, and D. L. Stein, Phys. Rev. B **27**, 602 (1983); see also K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986), p. 924. [The system which is analyzed is a $d = 1$ Ising spin-glass whose random coupling constants follow a centered Gaussian distribution and decay with distance as $1/r^{\alpha}$. It can be seen that the size scaling is still given by $\tilde{N} \equiv N^* + 1$, to be introduced later on, but with α replaced by 2α; in other words, the system is extensive if and only if $2\alpha > d$.]
22. H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1984), p. 9.
23. M. O. Caceres, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 124 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
24. X.-P. Huang and C. F. Driscoll, Phys. Rev. Lett. **72**, 2187 (1994).
25. E. M. Montroll and B. J. West, in *Fluctuation Phenomena*, eds. E. W. Montroll and J. L. Lebowitz (North-Holland, Amsterdam, 1979) [2nd edition: North-Holland Personal Library (1987)]; E. Montroll and M. F. Shlesinger, J. Stat. Phys. **32**, 209 (1983); *Lévy Flights and Related Topics in Physics*, eds. M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch (Springer-Verlag, Berlin, 1995); P. Allegrini, P. Grigolini, and B. J. West, Phys. Rev. E **54**, 4760 (1996); B. J. West and P. Grigolini, Phys. Rev. E **55**, 99 (1997); P. Grigolini, A. Rocco, and B. J. West, Phys. Rev. E **59**, 2603 (1999); B. J. West, *Physiology, Promiscuity and Prophecy at the Millenium: A Tale of Tails* (World Scientific, Singapore, 1999).
26. Y.-H. Taguchi and H. Takayasu, Europhys. Lett. **30**, 499 (1995).
27. K. T. Waldeer and H. M. Urbassek, Physica A **176**, 325 (1991).
28. I. Koponen, Phys. Rev. E **55**, 7759 (1997).
29. D. C. Clayton, Nature **249**, 131 (1974).
30. N. A. Bahcall and S. P. Oh, Astrophys. J. **462**, L49 (1996).
31. J. M. Liu, J. S. De Groot, J. P. Matte, T. W. Johnston, and R. P. Drake, Phys. Rev. Lett. **72**, 2717 (1994).
32. J. Maddox, Nature **365**, 103 (1993).
33. H. P. de Oliveira, S. L. Sautu, I. D. Soares, and E. V. Tonini, Phys. Rev. D **60**, 121301 (1999).
34. G. Wilk and Z. Wlodarcsyk, Nucl. Phys. B (Proc. Suppl.) **75A**, 191 (1999); G. Wilk and Z. Wlodarcsyk, Phys. Rev. D **50**, 2318 (1994); G. Wilk and Z. Wlodarcsyk, Phys. Rev. Lett. **84**, 2770 (2000); M.L.D. Ion and D.B. Ion, Phys. Lett. B **482**, 57 (2000).

35. I. Bediaga, E. M. F. Curado, and J. Miranda, Physica A **286**, 156 (2000); C. Beck, Physica A **286**, 164 (2000).
36. O. V. Utyuzh, G. Wilk, and Z. Wlodarcsyk, preprint (1999) [hep-ph/9906442].
37. O. V. Utyuzh, G. Wilk, and Z. Wlodarcsyk, preprint (1999) [hep-ph/9906500].
38. T. Alber *et al.*, Eur. Phys. J. C **2**, 643 (1998).
39. D. B. Walton and J. Rafelski, Phys. Rev. Lett. **84**, 31 (2000).
40. S. Abe and A. K. Rajagopal, Phys. Rev. A **60**, 3461 (1999); A. Vidiella-Barranco, Phys. Lett. A **260**, 335 (1999).
41. E. M. F. Curado and C. Tsallis, J. Phys. A **24**, L69 (1991); (Corrigenda) **24**, 3187 (1991) and **25**, 1019 (1992).
42. C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica A **261**, 534 (1998).
43. A. R. Plastino and A. Plastino, Phys. Lett. A **177**, 177 (1993).
44. J. L. Lebowitz, Physica A **194**, 1 (1993); Physics Today **46**, 32 (1993); Physica A **263**, 516 (1999), and references therein.
45. J. Harvda and F. Charvat, Kybernetica **3**, 30 (1967).
46. I. Vajda, Kybernetika **4**, 105 (1968) [in Czech].
47. Z. Daroczy, Inf. and Control **16**, 36 (1970).
48. A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978); I. J. Taneja, Advances in Electronics and Electron Physics **76**, 327 (1989); M. Behara, *Additive and Nonadditive Measures of Entropy* (Wiley Eastern, New Delhi, 1990); M. Basseville, Institut de Recherche en Informatique et Systemes Aleatoires -IRISA (France), Report 1020 (May 1996).
49. C. Tsallis, Chaos, Solitons and Fractals **6**, 539 (1995).
50. C. Anteneodo and A. R. Plastino, J. Phys. A **32**, 1089 (1999).
51. A. R. R. Papa, J. Phys. A **31**, 5271 (1998).
52. E. P. Borges and I. Roditi, Phys. Lett. A **246**, 399 (1998).
53. P. T. Landsberg and V. Vedral, Phys. Lett. A **247**, 211 (1998).
54. P. T. Landsberg, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 46 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
55. E. M. F. Curado, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 36 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
56. R. P. Di Sisto, S. Martinez, A. R. Plastino, and A. Plastino, Physica A **265**, 590 (1999).
57. R. S. Johal, Phys. Rev. E **58**, 4147 (1998).
58. A. K. Rajagopal and S. Abe, Phys. Rev. Lett. **83**, 1711 (1999).
59. R. S. Johal and R. Rai, Physica A **282**, 525 (2000).
60. R. Rossignoli and N. Canosa, Phys. Lett. A **264**, 148 (1999).
61. A. Renyi, in *Proc. Fourth Berkeley Symposium*, 1960, Vol. 1 (University of California Press, Berkeley, Los Angeles, 1961), 547; *Probability Theory* (North-Holland, 1970) and references therein.
62. I. Csiszar, *Information measures: A critical survey*, in *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and the European Meeting of Statisticians*, 1974 (Reidel, Dordrecht, 1978), p. 73.
63. M. P. Schutzenberger, *Contribution aux applications statistiques de la theorie de l'information*, Publ. Inst. Statist. Univ. Paris **3**, 3 (1954).
64. M. Hotta and I. Joichi, Phys. Lett. A **262**, 302 (1999).
65. R. J. V. dos Santos, J. Math. Phys. **38**, 4104 (1997); see also S. Abe, Phys. Lett. A **271**, 74 (2000).

66. S. Abe, Phys. Lett. A **224**, 326 (1997).
67. F. Jackson, Mess. Math. **38**, 57 (1909); Quart. J. Pure Appl. Math. **41**, 193 (1910).
68. It is not hard to propose Jackson-like generalized differential operators associated with operations *other* than dilatations, for instance those associated with the conformal group. The procedure can be illustrated on a transformation of a simple, one-dimensional variable x. We assume $x' = x + f(x,a)$ with $f(x,0) = 0$ $(\forall x)$. The generalized derivative would then be [private discussion with M. A. Rego-Monteiro] $D_a F(x) \equiv [F(x') - F(x)]/[x' - x]$. In the limit $a \to 0$ limit we have, if $f(x,a)$ is differentiable with respect to a, that $f(x,a) \sim f'(x,0)a$, hence $x' \sim x + f'(x,0)a$, hence $D_0 F(x) \equiv \lim_{a\to 0} D_a F(x) = F'(x)$, thus recovering the usual derivative as the $a = 0$ special case. Now comes the nontrivial step: if we knew for this generic situation on what, instead of $\sum_i^W p_i^\alpha$, we should apply D_a, we would be able to define classes of entropic forms, respectively associated with generic symmetry groups. The BG entropy would be the one essentially corresponding to translations; the present $q \neq 1$ would correspond to dilatations; many more could emerge! [12]. This speculative scenario would in fact be physically appealing: the entropy being a measure of the (lack of) information, it seems natural that it be adapted to and reflecting some essential symmetry of the system it is supposed to "inform" on. To be concrete, if a many-body nonlinear dynamical system had neither exponential nor power-law time-dependence of the sensitivity to the initial conditions, but rather say a *logarithmic* dependence, it would not be really surprising that the need appeared of an entropic form *different* from S_q.
69. C. Tsallis, Quimica Nova **17**, 468 (1994).
70. E. P. Borges, J. Phys. A **31**, 5281 (1998).
71. P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983) and Phys. Rev. A **28**, 2591 (1983); T. A. Halsey *et al.*, Phys. Rev. A **33**, 1141 (1986).
72. S. Watanabe, *Knowing and Guessing* (Wiley, New York, 1969).
73. H. Barlow, Vision. Res. **30**, 1561 (1990); see also G. Toulouse, J. Phys. I (France) **3**, 229 (1993).
74. A. Plastino and A. R. Plastino, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 50 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
75. L. R. da Silva, E. K. Lenzi, J. S. Andrade, and J. Mendes Filho, Physica A **275**, 396 (2000).
76. S. Abe, Phys. Lett. A **263**, 424 (1999); (Erratum) **267**, 456 (2000).
77. A. R. Lima and T. J. P. Penna, Phys. Lett. A **256**, 221 (1999).
78. F. D. Nobre and C. Tsallis, Physica A **213**, 337 (1995) [Erratum: **216**, 369 (1995)]; F. D. Nobre and C. Tsallis, Phil. Mag. B **73**, 545 (1996).
79. C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems* (Cambridge University Press, Cambridge, 1993).
80. C. Tsallis and S. Abe, Physics Today **51**, 114 (October, 1998).
81. M. C. S. Vieira and C. Tsallis, J. Stat. Phys. **48**, 97 (1987).
82. G. R. Guerberoff and G. A. Raggio, J. Math. Phys. **37**, 1776 (1996); G. R. Guerberoff, P. A. Pury, and G. A. Raggio, J. Math. Phys. **37**, 1790 (1996).
83. A. Chame, Physica A **255**, 423 (1998).
84. S. Abe, Physica A **269**, 403 (1999).
85. R. F. S. Andrade, Physica A **203**, 486 (1994).
86. A. K. Rajagopal, in the present volume.

87. Although Eq. (54) is easy to obtain (it is in fact a simple corollary of Eqs. (2) and (36) of [42]), its physical interpretation and its mathematical rigorous justification are highly nontrivial. Rajagopal has presented in Okazaki, in February 1999, his own viewpoint on the matter, which has been enlightening for all of us.
88. K. Sasaki and M. Hotta, preprint (1999) [cond-mat/9912454].
89. V. Latora, A. Rapisarda, and S. Ruffo, Phys. Rev. Lett. **83**, 2104 (1999).
90. D. Prato and C. Tsallis, Phys. Rev. E **60**, 2398 (1999).
91. A. M. Mariz, Phys. Lett. A **165**, 409 (1992); J. D. Ramshaw Phys. Lett. A **175**, 169 and 171 (1993).
92. L. Borland, A. R. Plastino, and C. Tsallis, J. Math. Phys. **39**, 6490 (1998); (Erratum) **40**, 2196 (1999).
93. M. O. Caceres, Phys. Lett. A **218**, 471 (1995); A. Chame and E. V. L. de Mello, Phys. Lett. A **228**, 159 (1997).
94. A. K. Rajagopal, Phys. Rev. Lett. **76**, 3469 (1996).
95. E. K. Lenzi, L. C. Malacarne, and R. S. Mendes, Phys. Rev. Lett. **80**, 218 (1998).
96. C. Tsallis, Phys. Lett. A **206**, 389 (1995).
97. A. R. Plastino, A. Plastino, and C. Tsallis, J. Phys. A **27**, 5707 (1994).
98. T. Yamano, Phys. Lett. A **264**, 276 (1999).
99. C. Tsallis, A. R. Plastino, and W.-M. Zheng, Chaos, Solitons and Fractals **8**, 885 (1997).
100. E. P. da Silva, C. Tsallis, and E. M. F. Curado, Physica A **199**, 137 (1993); (Erratum) **203**, 160 (1994).
101. C. Tsallis, in *New Trends in Magnetism, Magnetic Materials and Their Applications*, eds. J. L. Moran-Lopez and J. M. Sanchez (Plenum Press, New York, 1994); A. Chame and E. V. L. de Mello, J. Phys. A **27**, 3663 (1994).
102. B. M. Boghosian, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 90 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
103. A. R. Plastino and A. Plastino, Phys. Lett. A **174**, 384 (1993); J. J. Aly, *Minimum energy / maximum entropy states of self-gravitating systems*, in *N–body problems and gravitational dynamics*, Proceedings of the Meeting held at Aussois–France (21–25 March 1993), eds. F. Combes and E. Athanassoula (Publications de l'Observatoire de Paris, Paris, 1993), page 19; A. R. Plastino and A. Plastino, Phys. Lett. A **193**, 251 (1994); A. R. Plastino and A. Plastino, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 1 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
104. B. M. Boghosian, Phys. Rev. E **53**, 4754 (1996).
105. D. A. Stariolo, Phys. Lett. A **185**, 262 (1994); L. Borland, Phys. Lett. A **245**, 67 (1998).
106. A. R. Plastino and A. Plastino, Phys. Lett. A **222**, 347 (1995).
107. C. Tsallis and D. J. Bukman, Phys. Rev. E **54**, R2197 (1996).
108. A. Compte and D. Jou, J. Phys. A **29**, 4321 (1996); D. A. Stariolo, Phys. Rev. E **55**, 4806 (1997); L. Borland, Phys. Rev. E **57**, 6634 (1998); L. Borland, Phys. Lett. A **245**, 67 (1998).
109. A. K. Rajagopal, Physica A **253**, 271 (1998).
110. H. S. Wio and S. Bouzat, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 135 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].

111. C. Tsallis, Phys. Rev. E **58**, 1442 (1998).
112. A. K. Rajagopal and C. Tsallis, Phys. Lett. A **257**, 283 (1999).
113. E. K. Lenzi, L. C. Malacarne, and R. S. Mendes, Phys. Rev. Lett. **80**, 218 (1998).
114. A. K. Rajagopal, R. S. Mendes and E. K. Lenzi, Phys. Rev. Lett. **80**, 3907 (1998); E. K. Lenzi, R. S. Mendes and A. K. Rajagopal, Phys. Rev. E **59**, 1398 (1999); R. S. Mendes, Braz. J. Phys. **29**, 66 (1999).
115. R. A. Treumann, Phys. Rev. E **57**, 5150 (1998); E. K. Lenzi, L. C. Malacarne, and R. S. Mendes, preprint (1998); J. E. Straub and T. Whitfield, preprint (1998).
116. A. K. Rajagopal, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 61 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
117. A. R. Plastino and C. Anteneodo, Ann. Phys. **255**, 250 (1997).
118. C. Tsallis and D. A. Stariolo, Notas de Fisica/CBPF (Brazil) 026 (June 1994) and Physica A **233**, 395 (1996); D.A. Stariolo and C. Tsallis, Ann. Rev. Comput. Phys., Vol. II, ed. D. Stauffer (World Scientific, Singapore, 1995)), page 343.
119. T. J. P. Penna, Phys. Rev. E **51**, R1 (1995).
120. I. Andricioaei and J. E. Straub, Pys. Rev. E **53**, R3055 (1996); Physica A **247**, 553 (1997); J. E. Straub and I. Andricioaei, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 179 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
121. J. Schulte, Phys. Rev. E **53**, 1348 (1996).
122. U. H. E. Hansmann, Physica A **242**, 250 (1997); Chem. Phys. Lett. **281**, 140 (1997); U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **56**, 2228 (1997); U. H. E. Hansmann, M. Masuya, and Y. Okamoto, Proc. Natl. Acad. Sci. USA **94**, 10652 (1997); U. H. E. Hansmann, F. Eisenmenger, and Y. Okamoto, Chem. Phys. Lett. **297**, 374 (1998); U. H. E. Hansmann and Y. Okamoto, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 187 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
123. P. Serra, A. F. Stanton, S. Kais, and R. E. Bleil, J. Chem. Phys. **106**, 7170 (1997); P. Serra and S. Kais, Chem. Phys. Lett. **275**, 211 (1997).
124. Y. Xiang, D. Y. Sun, W. Fan, and X. G. Gong, Phys. Lett. A **233**, 216 (1997).
125. M. R. Lemes, C. R. Zacharias, and A. Dal Pino Jr., Phys. Rev. B **56**, 9279 (1997).
126. D. E. Ellis, K. Mundim, V. P. Dravid, and J. W. Rylander, *Hybrid classical and quantum modeling of defects, interfaces and surfaces*, in *Computer Aided–Design of High–Temperature Materials* (Oxford University Press, Oxford, 1999), page 350; M. A. Moret, P. M. Bisch, and F. M. C. Vieira, Phys. Rev. E **57**, R2535 (1998); K. C. Mundim, T. Lemaire, and A. Bassrei, Physica A **252**, 405 (1998); M. A. Moret, P. G. Pascutti, P. M. Bisch, and K. C. Mundim, J. Comput. Chem. **19**, 647 (1998); L. Guo, D. E. Ellis, and K. C. Mundim, J. Porphyrins and Phthalocyanines **3**, 196 (1999); K. C. Mundim and D. E. Ellis, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 197 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
127. K. C. Mundim and C. Tsallis, Int. J. Quantum Chem. **58**, 373 (1996).
128. H. Nishimori and J. Inoue, J. Phys. A **31**, 5661 (1998).
129. R. Salazar and R. Toral, Phys. Rev. Lett. **83**, 4233 (1999); Computer Physics Communications **121/122**, 40 (1999); *A Monte Carlo method for the numerical simulation of Tsallis statistics*, preprint (1999) [cond-mat/9911063]; A. R. Lima, J. S. Sa Martins, and T. J. P. Penna, Physica A **268**, 553 (1999).

130. A. R. Plastino, H. G. Miller, A. Plastino, and G. D. Yen, J. Math. Phys. **38**, 6675 (1997).
131. S. Abe and A. K. Rajagopal, Physica A (2001), in press.
132. A. K. Rajagopal, Phys. Lett. A **205**, 32 (1995).
133. M. Portesi and A. Plastino, Physica A **225**, 412 (1996).
134. D. B. Ion and M. L. D. Ion, Phys. Rev. Lett. **81**, 5714 (1998); Phys. Rev. Lett. **83**, 463 (1999); Phys. Rev. E **60**, 5261 (1999).
135. F. Buyukkilic and D. Demirhan, Phys. Lett. A **181**, 24 (1993); F. Buyukkilic, D. Demirhan, and A. Gulec, Phys. Lett. A **197**, 209 (1995); F. Pennini, A. Plastino, and A.R. Plastino, Phys. Lett. A **208**, 309 (1995); F. Pennini, A.R. Plastino, and A. Plastino, Physica A **235**, 388 (1997); U. Tirnakli, F. Buyukkilic, and D. Demirhan, Phys. Lett. A **245**, 62 (1998).
136. A. K. Rajagopal, Physica B **212**, 309 (1995).
137. A. K. Rajagopal, Phys. Lett. A **214**, 127 (1996).
138. E. K. Lenzi, R. S. Mendes, and A. K. Rajagopal, Physica A **286**, 503 (2000).
139. A. Ott, J. P. Bouchaud, D. Langevin, and W. Urbach, Phys. Rev. Lett. **65**, 2201 (1990); J. P. Bouchaud, A. Ott, D. Langevin, and W. Urbach, J. Phys. II (France) **1**, 1465 (1991).
140. F. Bardou, J. P. Bouchaud, O. Emile, A. Aspect, and C. Cohen-Tannoudji, Phys. Rev. Lett. **72**, 203 (1994).
141. T. H. Solomon, E. R. Weeks, and H. L. Swinney, Phys. Rev. Lett. **71**, 3975 (1993).
142. G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley, Nature **381**, 413 (1996).
143. C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger, Phys. Rev. Lett. **70**, 1343 (1993).
144. R. A. Antonia, N. Phan-Thien, and B. R. Satyoparakash, Phys. Fluids **24**, 554 (1981).
145. C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Scirotino, M. Simons, and H. E. Stanley, Nature **356**, 168 (1992); H. E. Stanley, S. V. Buldyrev, A. L. Goldberger, S. Havlin, R. N. Mantegna, C.-K. Peng, M. Simons, and M. H. R. Stanley, in *Lévy Flights and Related Topics*, eds. M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch (Springer, Berlin, 1995), p. 331.
146. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1983).
147. B. B. Mandelbrot, *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk* (Springer-Verlag, New York, 1997), and references therein.
148. R. N. Mantegna, and H. E. Stanley, Nature **376**, 46 (1995).
149. P. A. Alemany and D. H. Zanette, Phys. Rev E **49**, R956 (1994).
150. D. H. Zanette and P. A. Alemany, Phys. Rev. Lett. **75**, 366 (1995); **77**, 2590 (1996); M. O. Caceres and C. E. Budde, Phys. Rev. Lett. **77**, 2589 (1996); D. H. Zanette, Braz. J. Phys. **29**, 107 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm]; A. Robledo, Phys. Rev. Lett. **83**, 2289 (1999).
151. C. Tsallis, S. V. F. Levy, A. M. C. de Souza, and R. Maynard, Phys. Rev. Lett. **77**, 5422 (1996) [Erratum **77**, 5442 (1996)]; C. Tsallis, Physics World **10**, 42 (July 1997).
152. V. V. Uchaikin and V. M. Zolotarev, *Chance and Stability - Stable Distributions and Their Applications*, Series *Modern Probability and Statistics* (VSP, The Netherlands, 1999).
153. E. W. Montroll and B. J. West, in *Studies in Statistical Mechanics*, Vol. VII, eds. E. W. Montroll and J. L. Lebowitz (North-Holland, Amsterdam, 1979), p. 62.

154. A. S. Chaves, Phys. Lett. A **239**, 13 (1998); M. P. Almeida, Phys. Lett. A **249**, 560 (1998).
155. G. Drazer, H. S. Wio and C. Tsallis, Phys. Rev. E **61**, 1417 (2000).
156. A. R. Plastino, M. Casas, and A. Plastino, Physica A (2000), in press.
157. P. Quarati, A. Carbone, G. Gervino, G. Kaniadakis, A. Lavagno, and E. Miraldi, Nucl. Phys. A **621**, 345c (1997); G. Kaniadakis, A. Lavagno, M. Lissia, and P. Quarati, Physica A **261**, 359 (1998); M. Coraddu, G. Kaniadakis, A. Lavagno, M. Lissia, G. Mezzorani, and P. Quarati, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 153 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
158. V. H. Hamity and D. E. Barraco, Phys. Rev. Lett. **76**, 4664 (1996); D. F. Torres, H. Vucetich, and A. Plastino, Phys. Rev. Lett. **79**, 1588 (1997).
159. G. K. Zipf, *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Cambridge-MA, 1949); see also R. Gunther, L. Levitin, B. Schapiro, and P. Wagner, Intern. J. Theor. Phys. **35**, 395 (1996).
160. S. Denisov, Phys. Lett. A **235**, 447 (1997).
161. D. Kahneman and A. Tversky, Econometrica **47**, 263 (1979); A. Tverky and D. Kahneman, Journal of Risk and Uncertainty **5**, 297 (1992).
162. C. Tsallis, Chaos, Solitons and Fractals **6**, 539 (1995).
163. C. Tsallis, A. S. Martinez and R. Maynard, unpublished.
164. C. Tsallis, Physica A **221**, 277 (1995).
165. X.-P. Huang, F. Anderegg, E. M. Hollmann, C. F. Driscoll, and T. M. O'Neil, Phys. Rev. Lett. **78**, 875 (1997); F. Anderegg, X.-P. Huang, C. F. Driscoll, E. M. Hollmann, T. M. O'Neil, and D. H. E. Dubin, Phys. Rev. Lett. **78**, 2128 (1997).
166. C. Anteneodo and C. Tsallis, J. Mol. Liq. **71**, 255 (1997).
167. C. Anteneodo, private communication. This theory was first presented in [3].
168. H. Brands, P. H. Chavanis, R. Pasmanter, and J. Sommeria, Phys. Fluids **11**, 3465 (1999).
169. V. Berezinsky, *Solar Neutrino Problem*, 30eme Rencontre de Moriond (12-18 March, 1995) [Laboratori Nazionali del Gran Sasso, Report 14 (June 1995)].
170. G. Kaniadakis, A. Lavagno, and P. Quarati, Phys. Lett. B **369**, 308 (1996).
171. A. Lavagno, G. Kaniadakis, M. Rego-Monteiro, P. Quarati, and C. Tsallis, Astrophys. Lett. and Comm. **35**, 449 (1998).
172. C. Tsallis and A. M. C. de Souza, Phys. Lett. A **235**, 444 (1997).
173. J. C. Mather *et al.*, Astrophys. J. **420**, 439 (1994); D.J. Fixsen *et al.*, Astrophys. J. **420**, 457 (1994).
174. C. Tsallis, F. C. Sa Barreto, and E. D. Loh, Phys. Rev. E **52**, 1447 (1995).
175. A. R. Plastino, A. Plastino, and H. Vucetich, Phys. Lett. A **207**, 42 (1995).
176. A. Einstein, in *Physics and Reality.* [In his words: "To be sure, it has been pointed out that the introduction of a spacetime continuum may be considered as contrary to nature in view of the molecular structure of everything which happens on a small scale. It is maintained that perhaps the success of the Heisenberg method points to a purely algebraical method of description of nature, that is to the elimination of continuum functions from physics. Then, however, we must also give up, by principle, the spacetime continuum."]; in *The Meaning of Relativity.* [In his words: "One can give good reasons why reality cannot at all be represented by a continuous field. From the quantum phenomena it appears to follow with certainty that a finite system of finite energy can be completely described by a finite set of numbers (quantum numbers). This does not seem to be in accordance

with a continuum theory, and must lead to an attempt to find a purely algebraic theory for the description of reality. But nobody knows how to obtain the basis of such a theory."]

177. R. Hagedorn, Suppl. Nuovo Cimento **3**, 147 (1965).
178. W. M. Alberico, A. Lavagno, and P. Quarati, Eur. Phys. J. C **12**, 499 (1999).
179. C. M. G. Lattes, Y. Fujimoto, and S. Hasegawa, Phys. Rep. **65**, 151 (1980).
180. R. H. Austin, K. Beeson, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, and V. P. Marshall, Phys. Rev. Lett. **32**, 403 (1974); R. H. Austin, K. Beeson, L. Eisenstein, and H. Frauenfelder, Biochemistry **14**, 5355 (1975); F. Parak and H. Frauenfelder, Physica A **201**, 332 (1993); H. Frauenflder and P. G. Wolynes, Physics Today **58** (February 1994); P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **267**, 1169 (1995).
181. C. Tsallis, G. Bemski, and R. S. Mendes, Phys. Lett. A **257**, 93 (1999); Erratum: see caption of Fig. 15.
182. A. Upadhyaya, J. P. Rieu, J. A. Glazier, and Y. Sawada (1998), private communication.
183. S. Redner, Eur. Phys. J. B **4**, 131 (1998).
184. C. Tsallis and M. P. de Albuquerque, Eur. Phys. J. B **13**, 777 (2000).
185. L. G. Gamero, A. Plastino, and M. E. Torres, Physica A **246**, 487 (1998); A. Capurro, L. Diambra, D. Lorenzo, O. Macadar, M. T. Martin, C. Mostaccio, A. Plastino, J. Perez, J. Perez, E. Rofman, M. E. Torres, and J. Velluti, Physica A **257**, 149 (1998); A. Capurro, L. Diambra, D. Lorenzo, O. Macadar, M. T. Martin, C. Mostaccio, A. Plastino, J. Perez, E. Rofman, M. E. Torres, and J. Velluti, Physica A **265**, 235 (1999); M. T. Martin, A. R. Plastino, and A. Plastino, Physica A **275**, 262 (2000).
186. A. C. Tsallis, C. Tsallis, A. C. N. de Magalhaes, and F.A. Tamarit, to be published.
187. S. A. Cannas, D. A. Stariolo, and F. A. Tamarit, Network: Computation in neural sciences **7**, 141 (1996).
188. S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Nature **381**, 767 (1996).
189. F. M. Ramos, C. Rodrigues Neto, and R. R. Rosa, preprint (1999) [cond-mat/9907348].
190. B. Chabaud, A. Naert, J. Peinke, F. Chilla, B. Castaing, and B. Hebral, Phys. Rev. Lett. **73**, 3227 (1994).
191. T. Arimitsu and N. Arimitsu, Phys. Rev. E **61**, 3237 (2000).
192. F. Anselmet, Y. Cagne, E. J. Hopfinger, and R. A. Antonia, J. Fluid Mech. **140**, 63 (1984); C. Meneveau and K. R. Sreenivasan, Nucl. Phys. B (Proc. Suppl.) **2**, 49 (1987).
193. C. Beck, Physica A **277**, 115 (2000).
194. B. Castaing, Y. Gagne, and E. J. Hopfinger, Physica D **46**, 177 (1990).
195. T. Arimitsu and N. Arimitsu, J. Phys. A **33**, L 235 (2000).
196. C. Meneveau and K. R. Sreenivasan, J. Fluid Mech. **224**, 429 (1991).
197. C. Tsallis, Fractals **3**, 541 (1995).
198. P. Jund and S. G. Kim, Phys. Rev. B **52**, 50 (1995); J. R. Grigera, Phys. Lett. A **217**, 47 (1996); S. A. Cannas and F. A. Tamarit, Phys. Rev. B **54**, R12661 (1996); S. A. Cannas and A. C. N. Magalhaes, J. Phys. A **30**, 3345 (1997); L. C. Sampaio, M. P. de Albuquerque, and F. S. de Menezes, Phys. Rev. B **55**, 5611 (1997); S. E. Curilef, Ph.D. thesis (CBPF, Rio de Janeiro, 1997); S. Curilef and C. Tsallis, Phys. Lett. A **264**, 270 (1999); H. H. A. Rego, L. S. Lucena, L. R. da

Silva, and C. Tsallis, Physica A **266**, 30 (1999); F. Tamarit and C. Anteneodo, Phys. Rev. Lett. **84**, 208 (2000).

199. H. N. Nazareno and P. E. de Brito, Phys. Rev. B **60**, 4629 (1999).
200. L. Borland and J. G. Menchero, Braz. J. Phys. **29**, 169 (1999); L. Borland, J. G. Menchero, and C. Tsallis, Phys. Rev. B **61**, 1650 (2000).
201. I. Goldhirsch and G. Zanetti, Phys. Rev. Lett. **70**, 1619 (1993).
202. U. M. S. Costa, M. L. Lyra, A. R. Plastino, and C. Tsallis, Phys. Rev. E **56**, 245 (1997).
203. M. L. Lyra and C. Tsallis, Phys. Rev. Lett. **80**, 53 (1998); U. Tirnakli, C. Tsallis, and M. L. Lyra, Eur. Phys. J. B **10**, 309 (1999); C. R. da Silva, H. R. da Cruz, and M. L. Lyra, in *Nonextensive Statistical Mechanics and Thermodynamics*, eds. S. R. A. Salinas and C. Tsallis, Braz. J. Phys. **29**, 144 (1999) [accessible at http://sbf.if.usp.br/WWW_pages/Journals/BJP/Vol29/Num1/index.htm].
204. P. Grassberger and M. Scheunert, J. Stat. Phys. **26**, 697 (1981); T. Schneider, A. Politi, and D. Wurtz, Z. Phys. B **66**, 469 (1987); G. Anania and A. Politi, Europhys. Lett. **7**, 119 (1988) [It can be shown that from its Eq. (2) can be derived the present Eq. (156)]; H. Hata, T. Horita, and H. Mori, Progr. Theor. Phys. **82**, 897 (1989).
205. V. Latora and M. Baranger, Phys. Rev. Lett. **82**, 520 (1999).
206. R. C. Hilborn, *Chaos and Nonlinear Dynamics* (Oxford University Press, New York, 1994), page 390.
207. G. Guerberoff, private communication (1998).
208. V. Latora, M. Baranger, A. Rapisarda, and C. Tsallis, Phys. Lett. A **273**, 97 (2000).
209. S. Montangero, L. Fronzoni, and P. Grigolini, preprint (1999) [cond-mat/9911412].
210. Private discussion with P. Grigolini.
211. P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).
212. F. A. Tamarit, S. A. Cannas, and C. Tsallis, Eur. Phys. J. B **1**, 545 (1998); A. R. R. Papa and C. Tsallis, Phys. Rev. E **57**, 3923 (1998); see also A. Bhowal, Physica A **247**, 327 (1997).
213. V. Latora, A. Rapisarda, and S. Ruffo, Phys. Rev. Lett. **80**, 692 (1998).
214. C. Anteneodo and C. Tsallis, Phys. Rev. Lett. **80**, 5313 (1998).
215. L. Borland and C. Tsallis, in preparation (1999).
216. E. P. Borges and C. Tsallis, in preparation (1999).
217. A. Campa, A. Giansanti, D. Moroni, and C. Tsallis, cond-mat/0002168
218. M.-C. Firpo, Phys. Rev. E **57**, 6599 (1998).
219. M. Antoni and A. Torcini, Phys. Rev. E **57**, R6233 (1998).
220. M. Buiatti, P. Grigolini, and L. Palatella, Physica A **268**, 214 (1999).
221. S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi, Science **220**, 671 (1983).
222. A. M. C. de Souza and C. Tsallis, Physica A **236**, 52 (1997).
223. S. Lloyd, Phys. Rev. A **61**, R010301 (2000).
224. P. Gaspard and X.-J. Wang, Proc. Natl. Acad. Sci. USA **85**, 4591 (1988); P. A. Alemany, Phys. Lett. A **235**, 452 (1997); C. Tsallis, L. R. da Silva, R. S. Mendes, R. O. Vallejos, and A. M. Mariz, Phys. Rev. E **56**, R4922 (1997).
225. A. Wiles, Annals of Mathematics **142**, 443 (1995); R. Taylor and A. Wiles, Annals of Mathematics **142**, 553 (1995).
226. R. S. Mendes and C. Tsallis, *Renormalization group approach to nonextensive statistical mechanics*, preprint (2000) [cond-mat/0003365].
227. E. K. Lenzi, E. P. Borges, and R. S. Mendes, J. Phys. A **32**, 8551 (1999).

II. Quantum Density Matrix Description of Nonextensive Systems

A.K. Rajagopal

Naval Research Laboratory, Washington D.C. 20375-5320, USA

Abstract. A description of nonextensive systems based on the Tsallis generalization of the Boltzmann-Gibbs (BG) formalism is given in terms of quantum density matrix. After developing reasons for using the density matrix description as well as generalizations of the BG-formalism in the first section, we formulate the theory of quantum entangled states and information theory in the second section. In the third section, the maximum Tsallis entropy principle with given normalized q-mean value constraints is developed in detail, leading to quantum statistical mechanics of nonextensive systems. In the fourth section, time-dependent unitary dynamics is given. Here, the Green function theory as well as linear response theory are described in detail. A brief description of nonunitary (Lindblad) dynamics is outlined in the fifth section. In the final section, open problems and possible resolution are discussed as concluding remarks. In eight Tables, the summaries of the various sections are given with a view to give intercomparison of the present developments with the familiar ones found in the literature. In Appendix, several forms of the q-Kullback-Leibler entropy are given in seeking a guide to introduce invariance principles in the theory of nonextensive Tsallis entropy.

1 General Remarks

It is nearly 130 years since Boltzmann (1866) [1] and Gibbs (1876) [2] laid the foundation for the statistical description of the thermodynamic equilibrium of a collection of mutually interacting classical particles by introducing the well-known entropy functional expressed in terms of a stationary distribution function of particles in phase space. After this seminal work, many ideas such as the law of large numbers, ensemble theory, etc. were used to better understand the Boltzmann-Gibbs description. Several ensembles were introduced but they all led to one thermodynamics! The dynamical version of the classical distribution function follows the Liouville equation derived from the classical equations of motion of the particles composing the system. The stationary solution of this equation is used in setting up the equilibrium statistical mechanics through the maximum entropy (maxent) principle. It is in this way that the classical dynamical aspects of the system were incorporated into the description of macroscopic systems in thermodynamic equilibrium. In 1927, von Neumann [3] gave a general description of quantum systems by introducing the density matrix, which replaces the classical probability distribution by a traceclass (often taken as unit trace corresponding to the total probability being equal to unity), positive semi-definite, Hermitian operator, $\hat{\rho}$, corresponding to the properties of the classical probability and the concomitant generalization of entropy functional stated now

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 99–156, 2001.

in terms of the trace over an appropriate functional of the density matrix. The trace structure is important in that it makes the entropy functional independent of unitary transformations on the state space. Its dynamical version obeys the von Neumann equation derived from the quantum equations of motion for the system. This equation is linear so that the superposition principle of quantum theory is incorporated into the formulation. Again its stationary solution plays an important role in setting up quantum statistical mechanics of systems. Specifically, the density matrix with unit trace describes "mixed" states whenever $\hat{\rho}^2 < \hat{\rho}$ and so the system cannot be described by a wave function. In the case of "pure" states, wave function description holds. Here, the condition for a pure state is that the density matrix is idempotent: $\hat{\rho}^2 = \hat{\rho}$ implying that a state function description suffices where $\hat{\rho} = |\Psi\rangle\langle\Psi|$ with $|\Psi\rangle$ the quantum state. It is important to point out that Shannon [4] in 1948, in an entirely different context, introduced the idea of entropy associated with classical discrete probabilities, thus introducing information theory as a new area of science. This involved the introduction of the idea of additivity of entropy associated with composite systems along with their "conditional" and "mutual" entropies. This formulation, in contrast to the above, did not depend on the underlying mechanism for the statistical features, but employed only the basic structure of the multivariate probabilities. Recently, it has become possible to set up quantum communication system for which the Shannon description is generalized in terms of the density matrix, without invoking the underlying quantum dynamics, but only the quantum superposition principle, which leads to new features that are not in the classical probability framework (e.g., entanglement, the Bell inequalities, etc.). Unlike the classical probability description, the quantum version involving the superposition principle leads to the new idea of probability amplitude. In this respect, quantum information theory is different from its classical Shannon version. Only in 1957, Jaynes [5] gave a comprehensive maxent principle generalizing Boltzmann-Gibbs, von Neumann, and incorporating Shannon. He replaced the ensembles by linearly independent "constraints" in the maximization procedure with the introduction of corresponding Lagrange multipliers, which led to the same thermodynamics as before, in the form of general Legendre transformations. The Lagrange multipliers must be determined in terms of the given constraints, which may or may not be possible. Furthermore, the requirements of the positivity and finite trace of the resulting density matrix must be obeyed. Thus Jaynes simplified the statistical mechanical procedures enormously by avoiding the use of large-numbers approximation, ensembles, heat bath, etc.. Jaynes applied his procedure to examine many issues both in statistical mechanics and in other areas of sciences where probabilistic description is called for. In the latter cases, judicious choice of constraints placed on the maxent principle played the role of the Hamiltonian in equilibrium statistical mechanics. In this way, Jaynes introduced a new method for mathematical analysis of problems with probabilistic basis. This was considerably generalized beyond what Jaynes originally imagined to include quantum communication [6] by reformulating the Shannon theory in terms of the density matrix. (The dynamical description of the many-particle system has its own history and is a separate topic in itself.) With this brief history as introduction, we will now give an account of the new developments in the formalism that have taken place in the last decade. We will present this work in terms of the density matrix as it is the most general way of stating these ideas, provides unified, consistent picture of entangled systems besides encompassing the classical probability version of the theory of both statistical mechanics and information theory as special cases. There have been an enormous number of applications of this formulation over the last decade, as was the case with the original Jaynes maxent principle. Only a few of these will be touched upon in this sequel, as they will be discussed in much greater detail by the other colleagues, especially by Professor Constantino Tsallis in this volume. Our presentation here is thus hoped to be complementary to these other presentations.

In this first Section, we provide three examples as motivation for the use of density matrix approach. Two of these examples are from phenomenological considerations and one based on a simple Hamiltonian of a system. We also give arguments for the need for a generalization of the Boltzmann-Gibbs description of statistical mechanics. We follow this up with a quick derivation of the maxent principle of Jaynes [5] to establish a framework for subsequent development in succeeding sections in the non-Gibbsian context. In Section 2, the ideas of entanglement, decoherence, and quantum information theory based on the density matrix are developed. We do this by employing the exactly soluble model of interacting radiation and a two-level system (the Jaynes-Cummings model, JC for short) as a **guide** for the general development given here. This development is particularly relevant in the context of quantum devices of nanometric sizes that are being contemplated currently and perhaps will be fabricated in the next century. Even though this description is based on the underlying JC Hamiltonian, the conclusions about the various aspects of the information content in the density matrix of a composite system are general. In Section 3, the Tsallis formalism in its most recent version is presented along with the maxent principle and generalized statistical mechanics that go with it. Also in this section, a calculation of the isothermal response of a system property to a static perturbation is given. In Section 4, we address time-dependent phenomena by first using the unitary time evolution of the density matrix. A detailed discussion of the dynamical linear response theory is given. More generally, in this section, the Green function theory of nonextensive systems is also outlined. In Section 5, a brief description of nonunitary time evolution which is of current interest in understanding fast dynamical processes on time scales of pico- and femtoseconds in various contexts such as nanometric electronic devices, fast chemical reactions, etc., is given. Each of these sections contain respective summaries in the form of Tables. In the final Section 6, some concluding remarks in the form of some questions that are worrisome to me personally and prospects for future work are given in the light of two recent applications of the Tsallis form of the entropy.

A. Motivation for Generalization of Boltzmann–Gibbs Description

We first recall that among the statistical distributions, the **exponential-class** (e.g., Gaussian) played the central role in the analysis of many phenomena. (See [5] for a discussion of these aspects.) These can all be derived from a maxent principle subject to some constraints in which the entropy functional is chosen to be the Gibbs-von Neumann form, with $\mathrm{Tr}\hat{\rho} = 1$:

$$S_1 = -\mathrm{Tr}(\hat{\rho} \ln \hat{\rho}) \equiv -\langle \ln \hat{\rho} \rangle_1 \ . \tag{1}$$

Once the density matrix is determined, the mean values of any other physical quantity represented by a hermitian operator $\hat{A}$ of interest can be calculated by performing a calculation of the trace given by $\langle \hat{A} \rangle_1 = \mathrm{Tr}(\hat{A}\hat{\rho})$. This procedure covers a large class of extensive systems, for which the entropy is additive. There are many other probability distributions possessing long tails such as Pareto, Lévy, etc. which are of the monomial-class, not related to the exponential class. These are not derivable from the maxent principle with the Gibbs-von Neumann form for the entropy functional which were based on the assumed additive property of their entropy functional. These cover many other physical phenomena such as systems with fractally structured space-time, for example, which do not come under the rubric of "extensive" class of systems traditionally treated in physical and other sciences because the additive property does not hold for these systems. It is to encompass both of these situations that Tsallis [7] in 1988 introduced the following entropy functional:

$$S_q = \frac{1}{q-1}\mathrm{Tr}(\hat{\rho} - \hat{\rho}^q) = -\mathrm{Tr}(\hat{\rho}\mathrm{Ln}_q\hat{\rho}) \equiv -\langle \mathrm{Ln}_q\hat{\rho} \rangle_1 \ , \tag{2}$$

where $\mathrm{Tr}\hat{\rho} = 1$ and $\mathrm{Ln}_q\hat{\rho} \equiv \left(\hat{\rho}^{q-1} - 1\right)/(q-1)$. Here, q is a real parameter which characterizes the nonextensive feature (nonadditivity) of the system; for $q \to 1$, $\mathrm{Ln}_q\hat{\rho} \to \ln\hat{\rho}$ and Eq. (2) goes over to the von Neumann entropy, S_1, in Eq. (1). The 1988 version has since evolved [8] to make the formalism more complete and we use this form in this sequel. The density matrix so determined may then be used to compute normalized q-mean values of physical quantities defined by $\left\langle \hat{A} \right\rangle_q = \mathrm{Tr}(\hat{\rho}^q\hat{A})/\mathrm{Tr}\hat{\rho}^q$, for reasons of consistency of the formulation, which will be discussed later. It should be remarked that the criterion of using the limit $q = 1$ to give the extensive results and the preservation of Legendre transform properties of the functional used do not lead to a unique form of the q-entropy. By including a consistent formulation of information theory which now requires defining "conditional" and "mutual" entropies, along with statistical mechanics, a complete q-formulation as presented here may be one of the possible generalizations. (See however a discussion on this point in Section 6. In Table 1, we display the two formalisms for easy comparison. The difference between the two lies mainly in their additive properties (see Table 2).

As far as we are aware, there is no mathematical or physical argument to rule out the applicability of the Tsallis ensemble nor do we know any demonstration that the exponential-class covers every conceivable situation in physical and

Table 1. Definitions of entropy

Extensive – von Neumann	**Nonextensive – Tsallis**
$S_1 = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho})$ with $\mathrm{Tr}\hat{\rho} = 1$.	$S_q = \frac{1}{q-1}\mathrm{Tr}(\hat{\rho} - \hat{\rho}^q) \equiv -\mathrm{Tr}(\hat{\rho}\mathrm{Ln}_q\hat{\rho})$. Here, $\mathrm{Ln}_q\hat{\rho} \equiv \frac{1}{q-1}\left(\hat{\rho}^{q-1} - 1\right)$, $\mathrm{Tr}\hat{\rho} = 1$ (q: a real number).
$\langle\hat{A}\rangle_1 = \mathrm{Tr}(\hat{A}\hat{\rho})$: mean value of $\hat{A}$	$\langle\hat{A}\rangle_q = \mathrm{Tr}(\hat{A}\hat{\rho}^q)/\mathrm{Tr}\hat{\rho}^q$: normalized q-mean value of $\hat{A}$.

other sciences so that the universality of the Boltzmann-Gibbs ensemble may be considered as the only one paramount form. Essentially, then, the exponential weights that have appeared in the past few decades are found to be replaced by a monomial weights in the last decade by invoking the maxent principle with a different form for the entropy functional. In discussing sensitivity to initial conditions of nonlinear dynamical systems, similar "exponential" and "power" law sensitivities have been discussed recently [9,10], in the context of the use of the Kolmogorov-Sinai and classical Tsallis entropies in their quantification.

Table 2. Dissimilar properties of the entropies

Extensive – von Neumann	**Nonextensive – Tsallis**
Additivity property of S_1: If $\hat{\rho}(A \cup B) = \hat{\rho}(A) \otimes \hat{\rho}(B)$, then $S_1(A \cup B) = S_1(A) + S_1(B)$.	Additivity property of S_q: If $\hat{\rho}(A \cup B) = \hat{\rho}(A) \otimes \hat{\rho}(B)$, then $S_q(A \cup B) = S_q(A) + S_q(B) +$ $+ (1-q)S_q(A)S_q(B)$, which means that it is superadditive for $q < 1$ (entropy of whole is greater than the sum of its parts) and subadditive for $q > 1$.
Additivity property of mean values: $\langle\hat{O}\rangle_1(A \cup B) = \langle\hat{O}\rangle_1(A) + \langle\hat{O}\rangle_1(B)$	Additivity property of normalized q-mean values: $\langle\hat{O}\rangle_q(A \cup B) = \langle\hat{O}\rangle_q(A) + \langle\hat{O}\rangle_q(B)$
The Kubo dynamic linear response function obeys the fluctuation-dissipation theorem.	The q-Kubo dynamic linear response function obeys (a different of) form the fluctuation-disssipation theorem [11].

Also, the Tsallis ensemble with q different from 1 deals with Hamiltonians of systems with long-range interactions which may exhibit nontrivial anomalies in their ergodicity and mixing properties. For systems which are noninteracting or interacting systems with short-range forces, one has q less than or equal to 1 (Boltzmann-Gibbs class included). In Table 2, we display the dissimilar properties of the two entropies. Just as the work of Jaynes spawned a large body of work in many areas of science, so has the work of Tsallis, and one may access these by going to the website quoted in [7] where the burgeoning literature on the subject since 1988 is periodically updated.

B. Motivation for Using Density Matrix Description

There are several reasons for our choice of presenting this work in terms of the density matrix. Besides the mathematical reasons of generality and elegance, it covers the quantum probability ideas which include the superposition principle, not present in classical probability description. We will exhibit this distinction in several places in the course of this development. Also, density matrix description is eminently suited to discuss many commonly occurring physical situations. We consider here three such examples. Two of these examples describe phenomenological situations while the third one is based on an underlying system Hamiltonian.

(a) Photon Polarization (Polarized Light): Light being a transverse vector is described in terms of the two orthogonal components of its polarization vector. In nature, light occurs with arbitrary polarization not corresponding to these directions and by suitable optical means (e. g., Nicol prism), one may obtain light of given polarization state. We use this phenomenology to construct a density matrix description of light with arbitrary polarization as a 2×2 matrix:

$$\hat{\rho} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} , \quad \mathrm{Tr}\hat{\rho} = \rho_{11} + \rho_{22} = I \ , \tag{3}$$

where I is the intensity of light, and $\hat{\rho}$ satisfies the hermiticity condition $\rho_{12} = \rho_{21}^*$. The positive semi-definiteness of the density matrix implies that its eigenvalues are greater than or equal to zero, and whose sum is I, the intensity of light, as in Eq. (3). An immediate consequence of this is that

$$\begin{aligned} &\det\hat{\rho} = \rho_{11}\rho_{22} - |\rho_{12}|^2 \geq 0 \ , \\ &\mathrm{Tr}\hat{\rho}^2 - (\mathrm{Tr}\hat{\rho})^2 = -2\det\hat{\rho} \leq 0 \ . \end{aligned} \tag{4}$$

In the optics literature, one often expresses the polarization matrix in terms of the Stokes parameters, which is another way of parametrizing the density matrix in Eq. (3), which takes account of the various properties of the light polarization stated here. Thus,

$$\hat{\rho} = \frac{I}{2} \begin{pmatrix} 1 + \eta_3 & \eta_1 - i\eta_2 \\ \eta_1 + i\eta_2 & 1 - \eta_3 \end{pmatrix} \ . \tag{5}$$

From the condition derived in Eq. (4), it follows that the Stokes parameters obey the inequality

$$P^2 \equiv \eta_1^2 + \eta_2^2 + \eta_3^2 \leq 1 \ . \tag{6}$$

This leads to another important representation of the polarization, called the Poincare sphere representation. The eigenvalues associated with the density matrix in Eq. (5) are $\lambda_{1,2} = I(1 \pm P)/2$. Once the Stokes parameters are determined, the character of the light beam is known. The Tsallis information content in the beam is then represented by the Tsallis entropy, after appropriately redefining it in terms of the normalized density matrix, which depends only on the parameter, P, introduced in Eq. (6):

$$S_q = -\left[\left(\frac{1+P}{2}\right)\mathrm{Ln}_q\left(\frac{1+P}{2}\right) + \left(\frac{1-P}{2}\right)\mathrm{Ln}_q\left(\frac{1-P}{2}\right)\right] \ . \tag{7}$$

In the limit $q \to 1$, this goes to the traditional standard result of the information content in a beam of light. If $P = 0$, the beam is “unpolarized”, and for $P = 1$ it is completely polarized as in laser beam, and any other intermediate values of P, the beam is partially polarized as in Sun light.

(b) Particle Spin Polarization (Magnetism): In the phenomenological description of magnetic states of electron systems, one employs the spin-1/2 feature of the electron spin. The density matrix is again of the 2×2 form as above, separating the spin-dependence from the space-time dependence, so that the interpretation of the elements of the density matrix are now in terms of the space and time coordinate representation of the electron. In this case, we may write it in spin-space symbolically as

$$\hat{\rho}(\overrightarrow{r}, t; \overrightarrow{r}', t') = \begin{pmatrix} \rho_{\uparrow\uparrow}(\overrightarrow{r}, t; \overrightarrow{r}', t') & \rho_{\uparrow\downarrow}(\overrightarrow{r}, t; \overrightarrow{r}', t') \\ \rho_{\downarrow\uparrow}(\overrightarrow{r}, t; \overrightarrow{r}', t') & \rho_{\downarrow\downarrow}(\overrightarrow{r}, t; \overrightarrow{r}', t') \end{pmatrix} \ . \tag{8}$$

Now the diagonal parts in the space-time representation of this density matrix has the interpretation in terms of the particle density and spin density vector. To show this, we represent this matrix in terms of the unit and the Pauli spin matrices in the spin-space,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tau_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tau_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tau_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{9}$$

Then we express Eq. (8) in the form

$$\begin{aligned} \hat{\rho}(\overrightarrow{r}, t; \overrightarrow{r}', t') = {} & [I\hat{\rho}_0(\overrightarrow{r}, t; \overrightarrow{r}', t') + \tau_x\hat{\rho}_x(\overrightarrow{r}, t; \overrightarrow{r}', t') \\ & + \tau_y\hat{\rho}_y(\overrightarrow{r}, t; \bar{r}', t') + \tau_z\hat{\rho}_z(\overrightarrow{r}, t; \overrightarrow{r}', t')]/2 \ , \end{aligned} \tag{10}$$

where $\hat{\rho}_0(\overrightarrow{r}, t; \overrightarrow{r}', t') = \mathrm{tr}[I\hat{\rho}(\overrightarrow{r}, t; \overrightarrow{r}', t')]$, $\hat{\rho}_\alpha(\overrightarrow{r}, t; \overrightarrow{r}', t') = \mathrm{tr}[\tau_\alpha\hat{\rho}(\overrightarrow{r}, t; \overrightarrow{r}', t')]$, with $\alpha = x, y, z$, and tr means the trace over spin space only. In the diagonal space-time representation, we have then

$$n(\overrightarrow{r}, t) = \mathrm{tr}[I\hat{\rho}(\overrightarrow{r}, t; \overrightarrow{r}, t)] \ , \overrightarrow{s}(\overrightarrow{r}, t) = \mathrm{tr}[\overrightarrow{\tau}\hat{\rho}(\overrightarrow{r}, t; \overrightarrow{r}, t)] \ . \tag{11}$$

The magnetic properties of the system may then be represented in terms of this elegant form of the density matrix. Similar representations can be constructed for scattering problems involving spins.

(c) Systems at Finite Temperature: The most commonly occurring mixed state is that of a system at finite temperature. Here, the states of the system are weighted by suitable factors so that the system properties may be studied at any physical situation. Thus, if we label the energy eigenstates of the Hamiltonian of the system by a and the corresponding weights by $w(a)$, the density matrix has the form

$$\hat{\rho} = \sum_a |a\rangle w(a) \langle a| \ , \quad \mathrm{Tr}\hat{\rho} = 1 = \sum_a w(a) \ , \tag{12}$$

provided that, for simplicity of presentation, a is assumed to be discrete and nondegenerate. The weights are determined by the maxent principle. One can see at once that Eq. (12) implies the condition of "mixed" state, namely, $\hat{\rho}^2 < \hat{\rho}$. In the Gibbsian scheme, the weight factors are proportional to $\exp[-E(a)/kT]$, where $E(a)$ is the energy state associated with state a, T is the tempersature of the system, and k is the Boltzmann constant. In the next subsection, these ideas will be dealt with in some detail.

C. Maximum Entropy Principle of Jaynes Constraints Replace Ensembles

Let us consider the maxent principle of Jaynes [5] in the context of statistical mechanics. One could write this principle in a more general form but we focus here on problems of statistical mechanics. Suppose the system Hamiltonian and number operators be $\hat{H}$ and $\hat{N}$, respectively, whose mean values are given by $\mathrm{Tr}(\hat{\rho}\hat{H}) = \tilde{U}_1$ and $\mathrm{Tr}(\hat{\rho}\hat{N}) = \tilde{N}_1$. The maxent principle requires maximization of the von Neumann entropy functional of the density matrix subject to these two constraints along with the normalization of the density matrix. Introducing three Lagrange multipliers, λ_0, λ_1, and λ_2 corresponding to these three constraints, we have

$$\begin{aligned}
&\delta \left\{ S_1[\hat{\rho}] - \lambda_0 \mathrm{Tr}\hat{\rho} - \lambda_1 \mathrm{Tr}[\hat{\rho}(\hat{H} - \tilde{U}_1)] - \lambda_2 \mathrm{Tr}[\hat{\rho}(\hat{N} - \tilde{N}_1)] \right\} = 0 \\
&\implies \quad \ln \hat{\tilde{\rho}} = -\lambda_0 - \hat{X}_1(\lambda_1, \lambda_2) \ , \\
&\hat{\tilde{\rho}} = \left[\tilde{Z}_1(\lambda_1, \lambda_2)\right]^{-1} \exp\left[-\hat{X}_1(\lambda_1, \lambda_2)\right] , \\
&\hat{X}_1(\lambda_1, \lambda_2) = \lambda_1(\hat{H} - \tilde{U}_1) + \lambda_2(\hat{N} - \tilde{N}_1) \ , \\
&\tilde{Z}_1(\lambda_1, \lambda_2) = \mathrm{Tr}\exp\left[-\hat{X}_1(\lambda_1, \lambda_2)\right] \quad \text{(partition function).}
\end{aligned} \tag{13}$$

The traditional partition function found in the textbooks does not have the constant factor $\exp(\lambda_1\tilde{U}_1 + \lambda_2\tilde{N}_1)$, and we write these in the above form to be consistent with the formulation to be presented later when we consider the nonextensive theory. In fact, we have

$$\tilde{Z}_1(\lambda_1, \lambda_2) = Z_1(\lambda_1, \lambda_2) \exp(\lambda_1\tilde{U}_1 + \lambda_2\tilde{N}_1) \ . \tag{13a}$$

where Z_1 is the traditional partition function found in standard textbooks. This derivation shows that for $\tilde{U}_1$ and $\tilde{N}_1$ to be finite, the spectrum of the Hamiltonian and the number operators must be bounded from below. It is implied in this derivation that calculating $\tilde{U}_1$ and $\tilde{N}_1$, by using the density matrix found in Eq. (13), which are now found in terms of the Lagrange multipliers, one can solve for λ_1 and λ_2, for a given set of $\tilde{U}_1$ and $\tilde{N}_1$. In the case of classical probabilities, there are counterexamples to this requirement. One example of this may be mentioned in passing, namely, given the first two moments, μ_1 and μ_2, of a probability distribution defined on a semi-infinite domain, the maxent solution does not exist for all values of the two given moments if $\mu_2 > 2\mu_1^2$ [D. C. Dowson and A. Wragg, IEEE Trans. Info. Theory **IT-19**, 689 (1973)]. Another problem that often crops up is that the resulting density matrix must be positive in the space defined. One suggestion is to use an additional inequality constraint associated with the requirement of positivity of the density matrix. With these caveats in mind, then the entropy associated with this density matrix is a function of the Lagrange multipliers for given $\tilde{U}_1$ and $\tilde{N}_1$:

$$S_1\left(\lambda_1, \lambda_2 \middle| \tilde{U}_1, \tilde{N}_1\right) = -\operatorname{Tr}(\hat{\tilde{\rho}} \ln \hat{\tilde{\rho}}) = \ln \tilde{Z}_1(\lambda_1, \lambda_2) \; . \tag{14}$$

This entropy has zero first derivatives with respect to the Lagrange parameters showing that it is stationary and it must attain its maximum in this parameter space, for given fixed values of $\tilde{U}_1$ and $\tilde{N}_1$. Indeed we have

$$\begin{aligned}
&\frac{\partial S_1}{\partial \lambda_1} = \frac{\partial \ln\ \tilde{Z}_1(\lambda_1,\lambda_2)}{\partial \lambda_1} = 0 \; , \\
&\frac{\partial S_1}{\partial \lambda_2} = \frac{\partial \ln\ \tilde{Z}_1(\lambda_1,\lambda_2)}{\partial \lambda_2} = 0 \; , \\
&\frac{\partial^2 \ln\ \tilde{Z}_1(\lambda_1,\lambda_2)}{\partial \lambda_1^2} = \left\langle \left(\hat{H} - \tilde{U}_1\right)^2 \right\rangle_1 \equiv \sigma_1^2(\hat{H}) \\
&\qquad = -\frac{\partial \tilde{U}_1}{\partial \lambda_1} > 0 \; , \\
&\frac{\partial^2 \ln\ \tilde{Z}_1(\lambda_1,\lambda_2)}{\partial \lambda_2^2} = \left\langle \left(\hat{N} - \tilde{N}_1\right)^2 \right\rangle_1 \equiv \sigma_1^2(\hat{N}) \\
&\qquad = -\frac{\partial \tilde{N}_1}{\partial \lambda_2} > 0 \; , \\
&\frac{\partial^2 \ln\ \tilde{Z}_1(\lambda_1,\lambda_2)}{\partial \lambda_1 \partial \lambda_2} = \left\langle \left(\hat{H} - \tilde{U}_1\right) \left(\hat{N} - \tilde{N}_1\right) \right\rangle_1 \equiv \sigma_1^c(\hat{H},\hat{N}) \\
&\qquad = -\frac{\partial \tilde{U}_1}{\partial \lambda_2} = -\frac{\partial \tilde{N}_1}{\partial \lambda_1} \; , \\
&\sigma_1^2(\hat{H})\sigma_1^2(\hat{N}) - \left[\sigma_1^c(\hat{H},\hat{N})\right]^2 > 0 \; .
\end{aligned} \tag{15}$$

This nicely shows that the maxent principle requires that the fluctuations in energy and number, and the correlation coefficient between number and energy form a negative definite 2 × 2 covariance matrix, as spelled out by Eq. (15). To make contact with the traditional thermodynamics, one recognizes the Lagrange multiplier going with the constant energy constraint, λ_1, should be identified with the inverse temperature, β, and the one going with number constraint, λ_2, with $-\beta\mu$, where μ is the chemical potential of the system and are determined by the given values of mean energy and number of particles. Moreover, the Gibbs free energy is identified to be $F_1(\beta,\mu) = -\beta^{-1}\ln Z_1(\beta,\mu)$, and hence the thermodynamic result

$$F_1(\beta,\mu) = \tilde{U}_1 - \mu\tilde{N}_1 - \beta^{-1}S_1\left(\beta,\mu\middle|\tilde{U}_1,\tilde{N}_1\right). \tag{16}$$

This shows that maxent principle implies a minimum free energy principle! Furthermore, the thermodynamic stability conditions requiring the specific heat and compressibility to be positive are nothing but the conditions obtained in Eq. (15) on the second derivatives of the entropy functional. If we use $\tilde{U}_1$ and $\tilde{N}_1$ as variables with fixed β and μ, Eq. (16) can be used in the sense of a Legendre transformation so that $\partial S_1\big/\partial\tilde{U}_1 = \beta$, $\partial S_1\big/\partial\tilde{N}_1 = -\beta\mu$. The matrix of the second derivatives of S_1 with respect to $\tilde{U}_1$ and $\tilde{N}_1$ is minus the inverse of that of the the second derivatives of $S_1 = \ln\tilde{Z}_1$ with respect to the Lagrange multipliers, λ_1 and λ_2.

In the canonical ensemble description, one uses only the energy constraint in the maxent procedure, and similar results follow with the chemical potential dropped in the above analysis. This shows the power of the method of Jaynes which overcomes the concepts of heat bath, large numbers approximation, etc.. This derivation also shows that "thermodynamics" is nothing but an aspect of Legendre transformation and maximum entropy or minimum free energy principle! An important aspect of this development is that when λ_1 goes to ∞ (corresponding to zero temperature in the familiar notation) the entropy S_1 goes to zero and the density matrix collapses to the ground state of the system, which is the lowest eigenstate of the Hamiltonian. This idea is at the heart of the "annealing algorithm" which provides a way of exploring the lowest state of a many-body system, for example. It may be interesting to note the following Lemma which compares entropies when one has different number of constraints. The lemma states as follows.

LEMMA: The entropy of a system with one constraint is larger than the entropy of the same system with an additional constraint. The proof of this is based on the Kullback-Leibler inequality concerning two density matrices:

$$K_1(\hat{\rho}_1;\hat{\rho}_2) = \mathrm{Tr}[\hat{\rho}_2(\ln\hat{\rho}_2 - \ln\hat{\rho}_1)] \geq 0 \;, \tag{17}$$

with $\mathrm{Tr}\hat{\rho}_2 = \mathrm{Tr}\hat{\rho}_1 = 1$.

The proof of this inequality will be given in the next Section. We now consider a system with a given Hamiltonian. We construct two density matrices associated with it: one, $\hat{\rho}_2$, as in the grand canonical ensemble with two constraints $\tilde{U}_1$ and $\tilde{N}_1$, and another, $\hat{\rho}_1$, with only one constraint, $\tilde{U}_1$, the same value of the mean energy. Then we have

$$S_1(\hat{\rho}_1) = \ln\ Z_1(\beta_1) + \beta_1 \tilde{U}_1,\ \hat{\rho}_1 = (Z_1)^{-1} \exp(-\beta_1 \hat{H})\ . \tag{18}$$

From Eq. (17) and the results of the grand canonical ensemble, we have

$$-S_1(\hat{\rho}_2) + S_1(\hat{\rho}_1) \geq 0\ . \tag{19}$$

More generally, this shows that as we keep increasing the number of constraints in the maxent procedure, we get a nesting of the entropies as above with the entropy with one constraint being the largest. Another important consequence of this is that if $\hat{\rho}$ is an arbitrary density matrix, then the free energy associated with it obeys the inequality: $F_1(\hat{\rho}) = \mathrm{Tr}[\hat{\rho}(\hat{H} - \mu\hat{N} + \beta^{-1} \ln \hat{\rho})] \geq F_1(\hat{\hat{\rho}} = \hat{\rho}_{\max})$, where $\hat{\hat{\rho}}$ is given in Eq. (13).

In Table 3, we summarize the definitions of standard ensembles used in statistical mechanics as well as the corresponding constraints that go with each of them. It may be of interest to note that in the context of the experimental investigations on Bose condensation of small number of atoms, the question of ensembles and fluctuations therein has been discussed in [12].

Table 3. Standard ensembles and relation to maximum entropy principle

Microcanonical	**Canonical**	**Grand Canonical**
Total Number of particles N and the total energy E prescribed	System in contact with bath of given temperature T	System exchanges N and E with a bath
All configurations with N, E equally probable: $W(N, E)$	Prescribed parameters are T and N $\big(\beta = (kT)^{-1}\big)$: $Z(N, \beta)$, $\nu \equiv e^{-\varepsilon\beta}$	Chemical potential, μ, and T are control variables: $R(u, \nu) = R(\mu, T)$, $u \equiv e^{\beta\mu}, \nu \equiv e^{-\varepsilon\beta}$
	In the maxent terms, the constraints is the mean energy of the N-particle system. Correspondingly, the Lagrangean parameter is β.	In the maxent terms, the constraints are are the mean energy and the mean number. The Lagrangean parameters are β and μ.

2 Theory of Entangled States and Its Implications: Jaynes–Cummings Model

The motivation for the theory of entangled states is perhaps best illustrated by studying an exactly soluble model of an interacting system which brings out various aspects of the underlying concepts associated with any general formulation of several interacting systems, such as "marginal" density matrices of the subsystems and their entropies, etc.. We use the notations in [13] in setting up the density matrix for the Jaynes-Cummings (JC) model [14] of a generic two interacting systems, one being a single-mode radiation field and the other, a two-level atom. This model has become the paradigm for discussing many of the current problems in quantum optics [15]. This is also generic model for discussing interacting two-level system with a boson field. With this model as a *guide*, we discuss (a) entanglement and decoherence concepts (b) entropy and information content in the total system as well as in the subsystems and (c) comparison of two systems employing the various entropy concepts of the entangled system.

The Hamiltonian of the two-level atom interacting with a single given mode of quantized radiation field of a given frequency, ω, characterized by the usual creation, $a^\dagger$, and destruction, a, operators and in the rotating wave approximation is given by

$$H_{A+R} = \hbar\omega a^\dagger a + \frac{\hbar\omega_0}{2}\,\sigma_z + \hbar\kappa(a^\dagger\sigma_- + a\sigma_+)\ . \tag{20}$$

The Hamiltonian of the two level atom is represented by the z-component of the Pauli matrices, the energy separation of the two levels is ω_0, and κ is the interaction strength between the radiation and the atom. Exact solutions of this interacting system are [13,14]

$$\begin{aligned}
&H_{A+R}\,|\varphi(n,1)\,\rangle = \hbar\,[\omega(n+1/2)+\lambda_n]\,|\varphi(n,1)\,\rangle\ ,\\
&H_{A+R}\,|\varphi(n,2)\,\rangle = \hbar\,[\omega(n+1/2)-\lambda_n]\,|\varphi(n,2)\,\rangle\ ,\\
&H_{A+R}\,|0,-1\,\rangle = -\hbar\omega_0/2\,|0,-1\,\rangle\ ,\\
&|\varphi(n,1)\,\rangle = \cos\theta_n\,|n+1,-1\,\rangle + \sin\theta_n\,|n,+1\,\rangle\ ,\\
&|\varphi(n,2)\,\rangle = -\sin\theta_n\,|n+1,-1\,\rangle + \cos\theta_n\,|\ n,+1\,\rangle\ ,\\
&\tan\theta_n = \kappa\sqrt{(n+1)}\Big/\left(\varDelta\omega/2+\lambda_n\right)\ ,\\
&(\varDelta\omega) \equiv \omega-\omega_0,\ \lambda_n \equiv \left[(\varDelta\omega/2)^2+\kappa^2(n+1)\right]^{1/2}\ ,\\
&\sum_{n=0}^{\infty}\left[|\varphi(n,1)\,\rangle\langle\varphi(n,1)| + |\varphi(n,2)\,\rangle\langle\varphi(n,2)\,|\right]\\
&\qquad + |0,-1\,\rangle\langle 0,-1\,|\ =\ 1\ ,
\end{aligned} \tag{21}$$

where the last equation is the completeness relation, and $\{|\varphi(n,1)\,\rangle, |\varphi(n,2)\,\rangle, |0,-1\,\rangle\}$ form an orthonormal set. Here, n is a positive integer running from zero to infinity labeling the quantum numbers of the photon field, and $+1, -1$ represent the upper and lower quantum states of the two-level atom. The states $\{|\varphi(n,i)\rangle\}$ $(i=1,2)$ and $|0,-1\rangle$ represent the exact entangled states of the atom-radiation system. The angles, θ_n, are measures of the entanglement; it is zero in the noninteracting case when there is no entanglement, and is maximum at resonance, $\varDelta\omega = 0$ when $\theta_n = \pi/4$ for all n. In Eq. (21), we have also expressed these in terms of the states of the radiation and the atom. The density matrix of this system is now constructed in this entangled state representation by using the maxent principle with a given mean total energy of the system [13]. This construction provides us a way of realizing the various aspects of the entanglement – from classical to fully quantum – in terms of the system parameters, a novel aspect of this model. This density matrix is diagonal in this representation, given by

$$\begin{aligned}
&\hat{\rho}_{A+R} = |0,-1\,\rangle w(0)\,\langle 0,-1\,| \\
&\quad + \sum_{n=0}^{\infty} \left[|\varphi(n,1)\,\rangle w(n,1)\,\langle\varphi(n,1)\,| + |\varphi(n,2)\,\rangle w(n,2)\,\langle\varphi(n,2)\,| \right] \quad , \\
&w(0) = \exp\left(\beta\hbar\omega_0/2\right)/Z_{A+R} \ , \\
&w(n,1) = \exp\left\{-\beta\hbar\left[\omega(n+1/2)+\lambda_n\right]\right\}/Z_{A+R} \ , \\
&w(n,2) = \exp\left\{-\beta\hbar\left[\omega(n+1/2)-\lambda_n\right]\right\}/Z_{A+R} \ , \\
&Z_{A+R} = \sum_{n=0}^{\infty} \left\{ \exp\left(-\beta\hbar\left[\omega(n+1/2)+\lambda_n\right]\right) \right. \\
&\qquad \left. + \exp\left(-\beta\hbar\left[\omega(n+1/2)-\lambda_n\right]\right) \right\} + \exp(\beta\hbar\omega_0/2) \ . \qquad (22)
\end{aligned}$$

The weights $w(0)$, $w(n,1)$, and $w(n,2)$ are the occupation probabilities of the entangled eigenstates of the total system Hamiltonian in Eq. (20), that is, $|0,-1\rangle$, $|\varphi(n,1)\,\rangle$, and $|\varphi(n,2)\,\rangle$, respectively. Here, $\beta = T^{-1}$ $(k_B \equiv 1)$ is the usual inverse temperature. Clearly, using the orthonormality of the states, we have the relation

$$\sum_{n=0}^{\infty} \left[w(n,1) + w(n,2) \right] + w(0) = \mathrm{Tr}\hat{\rho}_{A+R} = 1 \ . \qquad (23)$$

We now express the density matrix in terms of the photon and atomic states using the relations in Eq. (21), which will display explicitly the entangled nature of the system between the atomic and the photon states:

$$\begin{aligned}
\hat{\rho}_{A+R} \ &= \sum_{n=0}^{\infty} |n+1,-1\,\rangle\langle\, n+1,-1| \\
&\qquad \times \left(w(n,1)\cos^2\theta_n + w(n,2)\sin^2\theta_n \right) \\
&+ w(0)\,|0,-1\,\rangle\langle\, 0,-1| \\
&+ \sum_{n=0}^{\infty} |n,+1\,\rangle\langle\, n,+1| \left(w(n,1)\sin^2\theta_n + w(n,2)\cos^2\theta_n \right) \\
&+ \sum_{n=0}^{\infty} \left[|n+1,-1\rangle\langle n,+1| + |n,+1\rangle\langle n+1,-1| \right] \\
&\qquad \times \left(w(n,1) - w(n,2) \right) \cos\theta_n \sin\theta_n \ . \qquad (24)
\end{aligned}$$

The fourth term on the right-hand side of the above equation is off-diagonal in both the photon number and atom states and thus represents the entanglement of the atomic and photonic states and hence also the "decoherence" features of the interacting system. From this expression, we obtain the "marginal" density matrix of the atomic system by taking trace over the photon number states, which is found to be diagonal in the atomic states but representing a mixed character as expected:

$$\begin{aligned}
&f_{(-)}\,|{-1}\,\rangle\langle\,-1| + f_{(+)}\,|{+1}\,\rangle\langle\,+1| = \hat{\rho}_A = \mathrm{Tr}_R\hat{\rho}_{A+R}\ ,\\
&f_{(-)} = w(0) + \sum_{n=0}^{\infty}(w(n,1)\cos^2\theta_n + w(n,2)\sin^2\theta_n)\ , \qquad (24a)\\
&f_{(+)} = \sum_{n=0}^{\infty}(w(n,1)\sin^2\theta_n + w(n,2)\cos^2\theta_n)\ .
\end{aligned}$$

Thus $f_{(-)}$ and $f_{(+)}$ are the occupation probabilities of the (-1) and $(+1)$ states of the atom with $f_{(-)} + f_{(+)} = 1$. Note that these retain the features of the interaction. Similarly, the "conditional" density matrix of the radiation is obtained by taking trace over the atomic states, which is diagonal in the photon number states:

$$\begin{aligned}
&\sum_{n=0}^{\infty} p_n|n\rangle\langle n| = \hat{\rho}_R = \mathrm{Tr}_A\hat{\rho}_{A+R}\ ,\\
&p_0 = w_0 + w(0,1)\sin^2\theta_0 + w(0,2)\cos^2\theta_0\ ,\\
&p_n = w(n,1)\sin^2\theta_n + w(n,2)\cos^2\theta_n + w(n-1,1)\cos^2\theta_{n-1}\\
&\quad + w(n-1,2)\sin^2\theta_{n-1} \quad (n = 1,2,3,\cdots)\ . \qquad (24b)
\end{aligned}$$

Here, p_n are the occupation probabilities of the n-th photon state with $\sum_{n=0}^{\infty} p_n = 1$, retaining features of the interaction. These reduce to the familiar noninteracting occupation probabilities when the interaction is absent as will be indicated subsequently.

It should be mentioned that, in Ref. [18], the exact solutions given in Eq. (21) can also be used in discussing time evolution of an initially specified system density matrix, from which one can learn about the time dependent features of the subsystems with a view to elucidate similar ideas about the marginal density matrices etc. in the quantum optical context.

Having obtained the relevant density matrices, one then constructs the various entropies in understanding the properties of the interacting system. A formal theoretical discussion of this problem was recently given by Cerf and Adami (CA) [16,17] who focused mainly on the quantum information theory aspects of the problem. CA proposed definitions of "conditional" and "mutual" amplitude operators which generalize their classical probability theory counterparts. The formalism utilized by Rajagopal *et al.* [18] differs from CA in that they could explicitly obtain by appealing to the definitions of "marginal" density matrices to elucidate the quantum information features in the JC model. An important result of the CA investigation is that quantum conditional entropy can be negative, in contrast to the classical counterpart, which must be positive. This is an aspect of quantum entanglement. We will now describe how one may construct such a theory of general entangled pair of systems using the lessons learned by the exact solutions of the JC model. It is interesting to note that the formal results of CA seem to be borne out by the work in Ref. [18] in the parameter space of the ensembles considered. Here, we present this version of the theory of entangled systems.

Consider now a general entangled pair of interacting systems A and B, whose composite density matrix is defined in its diagonal form in terms of the appropriate complete set of orthonormal entangled states, $\{|\varGamma\rangle\}$, with the properties:

$$\begin{aligned} &\langle \varGamma|\varGamma'\rangle = \delta(\varGamma,\varGamma') \ \ \text{(Kronecker delta)} \ , \\ &\sum_{\varGamma} |\varGamma\rangle\langle\varGamma| = \hat{I}(A,B) \ \ \text{(unit operator)} \ . \end{aligned} \tag{25}$$

We here use a discrete set for the sake of simplicity. Thus we have for the composite density matrix the representation

$$\begin{aligned} &\hat{\rho}(A,B) = \sum_{\varGamma} |\varGamma\rangle P(\varGamma)\langle\varGamma| \ , \\ &0 \le P(\varGamma) \le 1 \ , \ \sum_{\varGamma} P(\varGamma) = 1 \ . \end{aligned} \tag{26}$$

We interpret $P(\varGamma)$ as the probability of finding the AB-system in the entangled state $|\varGamma\rangle$.

The systems A and B, when they are not interacting, are described by their own private complete orthonormal sets of states, $\{|a\rangle\}$ and $\{|b\rangle\}$, respectively as follows:

$$\begin{aligned} &\langle a|a'\rangle = \delta(a,a'), \ \sum_{a} |a\rangle\langle a| = \hat{I}(A) \ , \\ &\langle b|b'\rangle = \delta(b,b'), \ \sum_{b} |b\rangle\langle b| = \hat{I}(B) \ . \end{aligned} \tag{27}$$

Here, $\hat{I}(A)$ and $\hat{I}(B)$ are the unit operators in A-space and B-space, respectively. In terms of these states, we may express the interacting AB-system as follows:

$$\begin{aligned} &|\varGamma\rangle = \sum_{a,b} |a,b\rangle U^{(\varGamma)}(a,b) \ , \ |a,b\rangle = |a\rangle|b\rangle \ , \\ &\sum_{a,b} U^{(\varGamma)}(a,b) U^{(\varGamma')^*}(a,b) = \delta(\varGamma,\varGamma') \ , \\ &\sum_{\varGamma} U^{(\varGamma)}(a,b) U^{(\varGamma)^*}(a',b') = \delta(a,a')\delta(b,b') \ . \end{aligned} \tag{28}$$

Then Eq. (26) takes the form

$$\hat{\rho}(A,B) = \sum_{\varGamma}\sum_{a,b}\sum_{a',b'} |\, a,b\,\rangle \ U^{(\varGamma)}(a,b) P(\varGamma) U^{(\varGamma)^*}(a',b') \langle a',b'\,| \ . \tag{29}$$

Taking the trace over the B-space leaves us with an operator in A-space, which we will now show is the "marginal" density matrix of A within the entangled system. Thus, we have

$$\begin{aligned}\hat{\rho}_1(A) &= \mathrm{Tr}_B \hat{\rho}(A,B) = \sum_b \langle b|\hat{\rho}(A,B)|b\rangle = \sum_{a,a'} |a\rangle \rho_1(a,a')\langle a'| \ , \\ \rho_1(a,a') &= \sum_{\Gamma}\sum_{b} \ U^{(\Gamma)}(a,b)P(\Gamma)U^{(\Gamma)^*}(a',b) \ . \end{aligned} \tag{30}$$

We will now show that these expressions obey the required properties of a density matrix in A-space by demonstrating hermiticity, positivity, and traceability of the marginal density matrix of A in Eq. (30).

Hermiticity: Upon taking the complex conjugate of $\rho_1(a,a')$ in Eq. (30), we have

$$\begin{aligned}\rho_1^*(a,a') &= \sum_{\Gamma}\sum_{b} U^{(\Gamma)^*}(a,b)P(\Gamma)U^{(\Gamma)}(a',b) \\ &= \rho_1(a',a) \ . \end{aligned} \tag{31}$$

Positivity: Consider a diagonal matrix element of $\hat{\rho}_1(A)$ in Eq. (30) in an arbitrary A-state, we have

$$\begin{aligned}\langle \psi(A) \mid \hat{\rho}_1(A) \mid \psi(A)\rangle &= \sum_{\Gamma}\sum_{b} \mid F^{(\Gamma)}(\psi,b) \mid^2 P(\Gamma) \geq 0 \ , \\ F^{(\Gamma)}(\psi,b) &= \sum_{a} \langle \psi \mid a\rangle \, U^{(\Gamma)}(a,b) \ . \end{aligned} \tag{32}$$

Traceability: By taking the trace over the A-states of Eq. (30), we have

$$\mathrm{Tr}_A \hat{\rho}_1(A) = \sum_{\Gamma}\sum_{a,b} \mid U^{(\Gamma)}(a,b) \mid^2 \, P(\Gamma) = 1 \ , \tag{33}$$

following from Eq. (28).

We now invoke a theorem which states that a positive semi-definite, hermitian, traceclass operator (finite trace) necessarily possesses a complete orthonormal set of eigenvectors and associated eigenvalues such that a diagonal representation for it exists. We therefore employ a complete orthonormal set $\{|\alpha\rangle\}$, in terms of which $\hat{\rho}_1(A)$ is diagonal:

$$\begin{aligned}\hat{\rho}_1(A) &= \sum_{\alpha} |\alpha\rangle p_1(\alpha)\, \langle \alpha| \ , \\ p_1(\alpha) &\geq 0 \ , \ \sum_{\alpha} p_1(\alpha) = 1 \ . \end{aligned} \tag{34}$$

In exactly similar manner, we may define the marginal density matrix of the system B and express it in diagonal form in terms of its own orthonormal complete set $\{| \, \beta \, \rangle\}$:

$$\hat{\rho}_2(B) = \sum_{\beta} | \, \beta \, \rangle q_2(\beta) \, \langle \beta \, | \ ,$$
$$q_2(\beta) \geq 0 \ , \ \sum_{\beta} q_2(\beta) = 1 \ . \tag{35}$$

One interprets $p_1(\alpha)$ as the probability of finding the system A in the state $|\alpha\rangle$ within the entangled AB-system, and similarly the other one. (Compare these with Eqs. (24a) and (24b) of the JC model.) Henceforth, we shall employ these bases to re-express equivalently the density matrix of the entangled AB-system. This will enable us to develop a theory of the entangled system in an economical and succinct fashion, **analogously** to the one formulated in [16,17], as follows:

$$| \, \Gamma \, \rangle = \sum_{\alpha,\beta} | \, \alpha, \beta \, \rangle \ V^{(\Gamma)}(\alpha,\beta), \ | \, \alpha, \beta \, \rangle = | \ \alpha \ \rangle \ | \ \beta \, \rangle \ \ ,$$
$$\sum_{\alpha,\beta} V^{(\Gamma)}(\alpha,\beta) V^{(\Gamma')^*}(\alpha,\beta) = \delta(\Gamma, \Gamma') \ ,$$
$$\sum_{\Gamma} V^{(\Gamma)}(\alpha,\beta) V^{(\Gamma)^*}(\alpha',\beta') = \delta(\alpha,\alpha')\delta(\beta,\beta') \ ,$$
$$\hat{\rho}(A,B) = \sum_{\Gamma} \sum_{\alpha,\beta} \sum_{\alpha',\beta'} | \, \alpha, \beta \, \rangle \ V^{(\Gamma)}(\alpha,\beta) P(\Gamma) V^{(\Gamma)^*}(\alpha',\beta') \, \langle \, \alpha', \beta' \, | \ \ ,$$
$$\hat{\rho}_1(A) = \sum_{\alpha} | \, \alpha \, \rangle \ p_1(\alpha) \, \langle \alpha \, | \ , \quad \hat{\rho}_2(B) = \sum_{\beta} | \, \beta \, \rangle q_2(\beta) \, \langle \beta \, | \ ,$$
$$p_1(\alpha) = \sum_{\Gamma} \sum_{\beta} | \ V^{(\Gamma)}(\alpha,\beta) \, |^2 \ P(\Gamma) \ ,$$
$$q_2(\beta) = \sum_{\Gamma} \sum_{\alpha} | \ V^{(\Gamma)}(\alpha,\beta) \, |^2 \ P(\Gamma) \ . \tag{36}$$

Note that the expressions for the quantum probabilities, $p_1(\alpha)$ and $q_2(\beta)$, in terms of the quantum probability, $p(\Gamma)$, of the composite system involve the absolute magnitudes of the overlap matrix elements of the entangled states with the product states of the individual A- and B-states defined above, and thus display the quantum entanglement features of these entities.

We may now express the marginal density matrices in the full Hilbert space of the entangled system as follows:

$$\hat{\rho}_1(A) \otimes \hat{I}(B) \ , \quad \hat{I}(A) \otimes \hat{\rho}_2(B) \ . \tag{37}$$

An explicit example of the above formulation is the exact solution of the JC model of the interacting two-level system (atom) with a single mode of the quantized radiation field presented at the beginning of this discussion. In this example, as in many others in physics, the systems A and B when not interacting are independent (i.e., their

density matrices commute) and they are entangled only when there is an interaction between the two systems.

We first consider the concepts of conditional and mutual entropies as was done in Refs. [16,17] but in terms of the density matrices of the subsystems constructed here. Define the von Neumann entropies associated with the three density matrices considered above as follows:

$$\begin{aligned} S_{(1)}(A) &= -\mathrm{Tr}_A[\hat{\rho}_1(A)\ln\hat{\rho}_1(A)] = -\sum_{\alpha} p_1(\alpha)\ln p_1(\alpha)\ , \\ S_{(2)}(B) &= -\mathrm{Tr}_B[\hat{\rho}_2(B)\ln\hat{\rho}_2(B)] = -\sum_{\beta} q_2(\beta)\ln q_2(\beta)\ , \\ S(A,B) &= -\mathrm{Tr}_{A,B}[\hat{\rho}(A,B)\ln\hat{\rho}(A,B)] = -\sum_{\Gamma} P(\Gamma)\ln P(\Gamma)\ . \end{aligned} \tag{38}$$

These are all nonnegative. The "conditional entropy" of A given B and vice versa are then **defined** as

$$S(A|B) = S(A,B) - S_{(2)}(B)\ , \quad S(B|A) = S(A,B) - S_{(1)}(A)\ . \tag{39}$$

The "mutual entropy" is defined as

$$S(A:B) = S_{(1)}(A) + S_{(2)}(B) - S(A,B)\ . \tag{40}$$

Using the expressions in Eq. (36), these may be rewritten in the form

$$\begin{aligned} S(A|B) &= -\sum_{\Gamma}\sum_{\alpha,\beta} |V^{(\Gamma)}(\alpha,\beta)|^2 P(\Gamma)\ln[P(\Gamma)/q_2(\beta)]\ , \\ S(B|A) &= -\sum_{\Gamma}\sum_{\alpha,\beta} |\ V^{(\Gamma)}(\alpha,\beta)\ |^2\ P(\Gamma)\ln[P(\Gamma)/p_1(\alpha)]\ , \\ S(A:B) &= \sum_{\Gamma}\sum_{\alpha,\beta} V^{(\Gamma)}(\alpha,\beta)\ |^2\ P(\Gamma)\ln[P(\Gamma)/(p_1(\alpha)q_2(\beta))]\ . \end{aligned} \tag{41}$$

When A and B are noninteracting, the composite system density matrix is just a product of those of the A and B systems. In such a case, the mutual entropy is zero and is thus a measure of interaction or entanglement!

Special Cases:

(a) **Uncorrelated System:** Here, $\hat{\rho}(A,B) = \hat{\rho}_1(A)\otimes\hat{\rho}_2(B)$, so that the eigenstates of the AB-system are just the product states of A and B, $|\Gamma\rangle = |\alpha\rangle|\beta\rangle$, and the probability of the AB-system is just the product of the probabilities, $P(\Gamma) \equiv p_1(\alpha)q_2(\beta)$.

(b) **Phase-Correlated System:** Here, the above factorization does not occur but the overlap matrix elements are mere phases: $|\ \Gamma = (\alpha,\beta)\rangle \equiv |\alpha,\beta\rangle e^{i\varphi(\alpha,\beta)}$, where (α,β) is a composite index and $P(\Gamma) = P(\alpha,\beta) \neq p_1(\alpha)q_2(\beta)$.

But we do have $p_1(\alpha)$ and $q_2(\beta)$ as the marginal probabilities associated with $P(\alpha,\beta)$:

$$p_1(\alpha) = \sum_\beta P(\alpha,\beta) \ , \quad q_2(\beta) = \sum_\alpha P(\alpha,\beta) \ . \tag{42}$$

An immediate consequence of this is

$$P(A|B) = \frac{P(\alpha,\beta)}{q_2(\beta)} \geq 0 \ , \tag{43}$$

which is the conditional probability of A given B. Since $\sum_\alpha P(\alpha,\beta)/q_2(\beta) = 1$, we have $P(\alpha,\beta)\,/q_2(\beta) \leq 1$. Similarly, the other conditional probability is defined. This in turn implies

$$S(A|B) \geq 0 \ , \quad S(B|A) \geq 0 \ . \tag{44}$$

Since $\ln x \geq (1 - x^{-1})$ for positive x, we have

$$S(A:B) = \sum_{\alpha,\beta} P(\alpha,\beta) \ln[P(\alpha,\beta)/(p_1(\alpha) q_2(\beta))] \geq 0 \ . \tag{45}$$

We also have the following inequality for the classical entropies:

$$S(A:B) \leq \min\left[S_{(1)}(A), S_{(2)}(B)\right] \ . \tag{45a}$$

(c) **Quantum Correlated System:** In this case, the interacting state is more complicated. The inequality Eq. (45) holds but the marginal probabilities are more complicated than in Eq. (42), and the inequality, $P(\alpha,\beta)\,/q_2(\beta) \leq 1$, does not hold. Therefore Eq. (44) does not hold. Also Eq. (45a) has to be replaced by the Araki-Lieb inequality:

$$0 \leq S(A:B) \leq 2 \ \min \ \left[S_{(1)}(A), S_{(2)}(B)\right] \ . \tag{46}$$

In fact, we may identify "quantum correlation" by the violation of the inequality in Eq. (44) or by the statement that the "negative values" of the conditional entropy. Another way of stating this is that the region between the inequalities in Eqs. (45a) and (46) depicts the quantum correlation, beyond the classical. This is the point made in Refs. [16,17] and found explicitly in the example of the JC model by Rajagopal *et al.* [18].

We have so far considered the traditional von Neumann entropies. We now turn our attention to the Tsallis-type entropies. There are two papers, one by Tsallis [19] and the other by Borland *et al.* [20], which deal with classical two-variable probabilities within the Tsallis entropy. The central issue here is the generalization of the conditional entropy and the mutual entropy defined above. These papers base their analysis on a generalization of the classical Kullback-Leibler (KL) entropy:

$$\begin{gathered} K(p;p') \equiv \sum_a p(a) \ln[p(a)/p'(a)] \geq 0 \ , \\ \sum_a p(a) = 1 = \sum_a p'(a) \ . \end{gathered} \tag{47}$$

Our generalization to the Tsallis-type entropy is

$$K_q(p;p') = \sum_a p(a)\left[\mathrm{Ln}_q p(a) - \mathrm{Ln}_q p'(a)\right] \ . \tag{48}$$

Note that this quantity goes to Eq. (47) when $q \to 1$. It is different from the definition in Refs. [19,20]. (However, it turns out that this quantity may have a problem with the nonnegativity condition, in general. See Appendix for a further discussion.) The definition of KL entropy for a pair of density matrices is given in Eq. (17). The correspondence to Eq. (48) is established as follows. Express these density matrices in their own diagonal representations:

$$\begin{aligned}&\hat\rho = \sum_a |\, a\,\rangle p(a)\,\langle\, a\,| \ , \quad \hat\rho' = \sum_b |\, b\,\rangle p'(b)\ \langle b\,| \ \ ,\\ &0 \le p(a), p'(b) \le 1 \ , \quad \sum_a p(a) = 1 = \sum_b p'(b) \ ,\end{aligned} \tag{49}$$

and $\langle\, a\,|\, b\,\rangle \neq 0$. $\{|\, a\,\rangle\}$ and $\{|\, b\,\rangle\}$ are complete, orthogonal sets. Then Eq. (17) takes the form

$$K_1(\hat\rho;\hat\rho') = \sum_{a,b} |\langle a|b\rangle|^2 p(a)\ \ln\left[p(a)/p'(b)\right] \ge 0 \ . \tag{17a}$$

The inequality follows by the same argument as before. This result goes over to the classical case when the overlap is unity for $a = b$ and zero otherwise. [This observation gives the hint as to the generalization of Eq. (48) for the density matrix description.] It is here that our approach differs from Refs. [16,17] as well as from Refs. [19,20]. The definitions of the Tsallis-type entropies associated with the various density matrices given in Eq. (36) are

$$\begin{aligned}S_q(A,B) &= \frac{1}{1-q}\mathrm{Tr}\left\{\hat\rho^q(A,B)\left[\hat I(A,B) - \hat\rho^{1-q}(A,B)\right]\right\}\\ &= -\mathrm{Tr}[\hat\rho(A,B)\mathrm{Ln}_q\hat\rho(A,B)]\\ &= -\sum_\Gamma P(\Gamma)\mathrm{Ln}_q P(\Gamma)\\ &= -\sum_\Gamma\sum_{\alpha,\beta}|V^{(\Gamma)}(\alpha,\beta)|^2 P(\Gamma)\mathrm{Ln}_q P(\Gamma) \ ,\\ \mathrm{Ln}_q\hat\rho(A,B) &\equiv \frac{1}{q-1}[\hat\rho^{q-1}(A,B) - \hat I(A,B)] \ ,\\ S_{q(1)}(A) &= \frac{1}{1-q}\mathrm{Tr}_A\left\{\hat\rho_1^q(A)[\hat I(A) - \hat\rho_1^{1-q}(A)]\right\}\\ &= -\mathrm{Tr}_A[\hat\rho_1(A)\mathrm{Ln}_q\hat\rho_1(A)]\\ &= -\sum_\alpha p_1(\alpha)\mathrm{Ln}_q p_1(\alpha)\\ &= -\sum_\Gamma\sum_{\alpha,\beta}|V^{(\Gamma)}(\alpha,\beta)|^2 P(\Gamma)\mathrm{Ln}_q p_1(\alpha) \ ,\end{aligned}$$

$$\begin{aligned}S_{q(2)}(B) &= \frac{1}{1-q}\mathrm{Tr}_B\left\{\hat{\rho}_2^q(B)[\hat{I}(B)-\hat{\rho}_2^{1-q}(B)]\right\}\\ &= -\mathrm{Tr}_B[\hat{\rho}_2(B)\mathrm{Ln}_q\hat{\rho}_2(B)]\\ &= -\sum_{\beta} q_2(\beta)\mathrm{Ln}_q q_2(\beta)\\ &= -\sum_{\Gamma}\sum_{\alpha,\beta}|V^{(\Gamma)}(\alpha,\beta)|^2 P(\Gamma)\mathrm{Ln}_q q_2(\beta)\ .\end{aligned} \tag{50}$$

We then **define** the conditional q-entropy of entangled systems as follows (see *Note Added*):

$$\begin{aligned}S_q(A|B) &= \frac{1}{1-q}\mathrm{Tr}\left\{\hat{\rho}(A,B)[\hat{\rho}^{q-1}(A,B)-(\hat{I}(A)\otimes\hat{\rho}_2(B))^{q-1}]\right\}\\ &= S_q(A,B)-S_{q(2)}(B)\ .\end{aligned} \tag{51}$$

We also **define** the q-mutual entropy by

$$S_q(A:B) = \frac{1}{q-1}\mathrm{Tr}\left\{\hat{\rho}(A,B)[\hat{\rho}^{q-1}(A,B)-(\hat{\rho}_1(A)\otimes\hat{\rho}_2(B))^{q-1}]\right\}\ . \tag{52}$$

Adding $\hat{I}(A,B)-\hat{I}(A,B)$ in $[\cdots]$ and noting the identity

$$\begin{aligned}&\left(\hat{I}(A)-\hat{\rho}_1^{q-1}(A)\right)\otimes\left(\hat{I}(B)-\hat{\rho}_2^{q-1}(B)\right)\\ &\quad= \hat{I}(A)\otimes\left(\hat{I}(B)-\hat{\rho}_2^{q-1}(B)\right)+\left(\hat{I}(A)-\hat{\rho}_1^{q-1}(A)\right)\otimes\hat{I}(B)\\ &\qquad-\left(\hat{I}(A)\otimes\hat{I}(B)-\hat{\rho}_1^{q-1}(A)\otimes\hat{\rho}_2^{q-1}(B)\right)\ ,\end{aligned} \tag{53}$$

we obtain

$$\begin{aligned}S_q(A:B) &= S_{q(1)}(A)+S_{q(2)}(B)-S_q(A,B)\\ &+\frac{1}{1-q}\mathrm{Tr}\left[\hat{\rho}(A,B)\left(\hat{I}(A)-\hat{\rho}_1^{q-1}(A)\right)\otimes\left(\hat{I}(B)-\hat{\rho}_2^{q-1}(B)\right)\right]\ .\end{aligned} \tag{54}$$

In the above, some of the definitions are different from those given by the authors of Ref. [20], and some of the results generalize theirs. In Eq. (54), we see a structure similar to Eq. (40) except for the last term. However, it is fair to say that establishing a connection between the conditional and mutual entropies in the nonextensive case is still an open problem.

These results on information theory are summarized below in Tables 4 and 5. The criteria used in defining the above quantities were that they should go to the known results when the probability of the composite system is the product of its marginals and when $q \to 1$.

Table 4. Comparison of classical and quantum information theories (extensive systems)

Classical Information Theory (Shannon)	**Quantum Information Theory (Cerf and Adami)**
If A takes a value a (discrete, for simplicity) with probability $p(a)$, then the Shannon entropy is defined by $S_1(A) = -\sum_a p(a) \ln p(a) \geq 0$.	If $\hat{\rho}(A)$ represents the density matrix of system A, its eigenvalue $p(\alpha)$ is the probability of finding A in the eigenstate of $\hat{\rho}(A)$. Then the von Neumann entropy is defined by $S_1(A) \equiv -\mathrm{Tr}[\hat{\rho}(A) \ln \hat{\rho}(A)]$ $= -\sum_\alpha p(\alpha) \ln p(\alpha)$.
A combined classical system AB is characterized by a joint probability $P(a,b)$ and therefore the joint entropy is $S_1(A,B) = -\sum_{a,b} P(a,b) \ln P(a,b) \geq 0$. One also defines "marginal" probabilities given by $p_1(a) = \sum_b P(a,b),\ q_2(b) = \sum_a P(a,b)$.	Here, a quantum system AB is described by its joint density matrix $\hat{\rho}(A,B)$ with its joint entropy given by $S_1(A,B) \equiv -\mathrm{Tr}[\hat{\rho}(A,B) \ln \hat{\rho}(A,B)]$ $= -\sum_\Gamma P(\Gamma) \ln P(\Gamma)$, where Γ is the quantum number of the AB-system. Tr here is trace over the eigenstates of the combined system. The "marginal" density matrices are defined by $\hat{\rho}_1(A)$ $= \mathrm{Tr}_B \hat{\rho}(A,B)$, $\hat{\rho}_2(B) = \mathrm{Tr}_A \hat{\rho}(A,B)$, where partial traces are taken.

The "conditional" entropy is defined by $S_1(A\|B) = -\sum_{a,b} P(a,b) \ln P(a\|b)$ $= S_1(A,B) - S_{1(2)}(B)$, where $P(a\|b) = P(a,b)/q_2(b) \leq 1$ is the conditional probability of A given B. The inequality follows from $\sum_a P(a\|b) = 1$. In the Shannon theory, the conditional entropy is always nonnegative and $\max[S_{1(1)}(A), S_{1(2)}(B)] \leq S_1(A,B)$.	The new proposal due to CA for defining the quantum conditional entropy is $S_1(A\|B) = -\mathrm{Tr}[\hat{\rho}(A,B) \ln \hat{\rho}(A\|B)]$, $\ln \hat{\rho}(A\|B) \equiv$ $\ln \hat{\rho}(A,B) - \hat{I}(A) \otimes \ln \hat{\rho}(B)$. The "conditional" matrix defined here is a positive Hermitian operator but is not a density matrix! This is due to "quantum entanglement" and leads to the difference from the Shannon result. Explicitly, in terms of the eigenstates, $S_1(A\|B) = S_1(A,B) - S_{1(2)}(B)$ $= \sum_{\Gamma} \sum_{\alpha,\beta} \left\| V^{(\Gamma)}(\alpha,\beta) \right\|^2 P(\Gamma)$ $\times \ln[P(\Gamma)/q_2(\beta)]$, $q_2(\beta) = \sum_{\Gamma} \sum_{\alpha} \left\| V^{(\Gamma)}(\alpha,\beta) \right\|^2 P(\Gamma)$. Unlike the classical quantity, $S_1(A\|B)$ can be negative.
The "mutual" or "correlation" Shannon entropy is a measure of the correlation that may exist in the joint probability and is defined by $S_1(A:B) = -\sum_{a,b} P(a,b) \ln P(a:b)$ $= S_{1(1)}(A) + S_{1(2)}(B) - S_1(A,B) \geq 0$, where $P(a:b) = p_1(a)q_2(b)/P(a,b)$. It also obeys the inequality $S_1(A:B) \leq \min\,[S_{1(1)}(A), S_{1(2)}(B)]$.	The quantum "mutual" entropy is a measure of entanglement of the two subsystems and is defined by $S_1(A:B) \equiv -\mathrm{Tr}[\hat{\rho}(A,B) \ln \hat{\rho}(A:B)]$ $= S_{1(1)}(A) + S_{1(2)}(B) - S_1(A,B)$ $= \sum_{\Gamma} \sum_{\alpha,\beta} \|V^{(\Gamma)}(\alpha,\beta)\|^2 P(\Gamma)$ $\times \ln[P(\Gamma)/(p_1(\alpha)q_2(\beta))]$. $\ln \hat{\rho}(A:B) \equiv -\ln \hat{\rho}(A,B)$ $+ \ln(\hat{\rho}_1(A) \otimes \hat{\rho}_2(B))$. This obeys the inequalities $0 \leq S_1(A:B) \leq 2\min[S_{1(1)}(A), S_{1(2)}(B)]$. The second inequality is due to Araki and Lieb.

Table 5. Comparison of classical and quantum information theories (nonextensive systems)

Classical Information Theory	**Quantum Information Theory**
If A takes a value a (discrete, for simplicity) with probability $p(a)$, then the classical Tsallis entropy is defined in the form $S_q(A) = -\sum_a p(a)\mathrm{Ln}_q p(a)$.	If $\hat{\rho}(A)$ represents the density matrix of a system A, its eigenvalue $p(\alpha)$ is the probability of finding it in the (for simplicity, discrete) eigenstate of $\hat{\rho}(A)$, then the quantum Tsallis entropy is defined in the form $S_q(A) \equiv -\mathrm{Tr}[\hat{\rho}(A)\mathrm{Ln}_q\hat{\rho}(A)]$ $= -\sum_\alpha p(\alpha)\mathrm{Ln}_q p(\alpha)$.
A combined classical system AB is characterized by a joint probability $P(a,b)$ and so the joint Tsallis entropy is $S_q(A,B) = -\sum_{a,b} P(a,b)\mathrm{Ln}_q P(a,b)$. One also defines the marginal probabilities as $p_1(a) = \sum_b P(a,b),\ q_2(b) = \sum_a P(a,b)$.	Here, a quantum system AB is described by its joint density matrix $\hat{\rho}(A,B)$, with which the joint Tsallis entropy is given by $S_q(A,B) \equiv -\mathrm{Tr}[\hat{\rho}(A,B)\mathrm{Ln}_q\hat{\rho}(A,B)]$ $= -\sum_\Gamma P(\Gamma)\mathrm{Ln}_q P(\Gamma)$, where Γ is the quantum number of AB-system, and $\mathrm{Ln}_q\hat{\rho}(A,B) \equiv \frac{\hat{\rho}^{q-1}(A,B) - \hat{I}(A,B)}{q-1}$. Here, Tr is trace over the eigenstates of the combined system. The marginal density matrices are defined by $\hat{\rho}_1(A) = \mathrm{Tr}_B\hat{\rho}(A,B)$, $\hat{\rho}_2(B) = \mathrm{Tr}_A\hat{\rho}(A,B)$, where partial traces are taken.

The "conditional" Tsallis entropy is defined by $S_q(A \mid B) \equiv S_q(A,B) - S_{q(2)}(B)$ $= \sum_{a,b} P(a,b)\frac{P^{q-1}(a,b) - q_2^{q-1}(b)}{1-q}.$ The conditional probability of A given B is $P(a \mid b) = P(a,b)/q_2(b) \le 1.$ The last inequality follows from $\sum_{a} P(a \mid b) = 1.$	The new proposal in [21] for defining the quantum conditional Tsallis entropy is to employ the eigenvalues of the marginal density matrices defined above. The "conditional" matrix defined from them is a positive hermitian operator but is not a density matrix! This is due to "quantum entanglement" and leads to the difference from the classical result. Explicitly, in terms of the eigenstates, $S_q(A \mid B) \equiv \mathrm{Tr}\hat{\rho}(A,B)$ $\times \left\{ \frac{\hat{\rho}^{q-1}(A,B) - (\hat{I}(A)\otimes\hat{\rho}_2(B))^{q-1}}{1-q} \right\}$ $= S_q(A,B) - S_{q(2)}(B)$ $= \sum_{\Gamma}\sum_{\alpha,\beta} \lvert V^{(\Gamma)}(\alpha,\beta)\rvert^2 P(\Gamma)$ $\times \frac{P^{q-1}(\Gamma) - q_2^{q-1}(\beta)}{1-q},$ $q_2(\beta) = \sum_{\Gamma}\sum_{\alpha} \lvert V^{(\Gamma)}(\alpha,\beta)\rvert^2 P(\Gamma).$
The "mutual" or "correlation" Tsallis entropy is a measure of correlation that may exist in the joint probability and is defined by $S_q(A:B) \equiv \sum_{a,b} P(a,b)$ $\times \left\{ \frac{P^{q-1}(a,b) - (p_1(a)q_2(b))^{q-1}}{q-1} \right\},$ $P(a:b) \equiv p_1(a)q_2(b)/P(a,b).$	The quantum "mutual" Tsallis entropy is a measure of quantum entanglement of the two subsystems and is defined by $S_q(A:B) \equiv \mathrm{Tr}\hat{\rho}(A,B)$ $\times \left\{ \frac{\hat{\rho}^{q-1}(A,B) - (\hat{\rho}_1(A)\otimes\hat{\rho}_2(B))^{q-1}}{q-1} \right\}$ $= \sum_{\Gamma}\sum_{\alpha,\beta} \lvert V^{(\Gamma)}(\alpha,\beta)\rvert^2 P(\Gamma)$ $\times \left\{ \frac{P^{q-1}(\Gamma) - (p_1(\alpha)q_2(\beta))^{q-1}}{q-1} \right\}.$

3 Variational Principle

We will now discuss statistical mechanics of nonextensive systems employing the Tsallis entropy. We will discuss (a) the generalized maxent principle using normalized q-mean values as constraints leading to the formulation of a (b) generalized statistical mechanics: free energy and Legendre transformation, and also consider the (c) isothermal response of the system to an external probe.

We maximize the Tsallis entropy given by Eq. (2) subject to the most common constraints employed in generalized statistical mechanics, namely given normalized "q-mean values" [8] of energy and number of particles in the system as in Section 1:

$$\begin{aligned} \left\langle \hat{H} \right\rangle_q &= \mathrm{Tr}(\hat{\rho}^q \hat{H})/\mathrm{Tr}\hat{\rho}^q \equiv \tilde{U}_q \ , \\ \left\langle \hat{N} \right\rangle_q &= \mathrm{Tr}(\hat{\rho}^q \hat{N})/\mathrm{Tr}\hat{\rho}^q \equiv \tilde{N}_q \ , \end{aligned} \tag{55}$$

where $\hat{H}$ and $\hat{N}$ are the system Hamiltonian operator and the system number operator, respectively. It is in this form that we get a consistent theory with all the important properties of the conventional statistical mechanics, as will be seen presently. Furthermore, the normalized q-mean values defined here lead to "connected" diagrams in any perturbation expansion of the mean value of any operator without divergences associated with the disconnected terms, in complete parallel to the usual formulation of quantum statistical mechanics of many-body systems. Using the standard method of Lagrange multipliers as in Section 1, we have the result

$$\begin{aligned} \hat{\tilde{\rho}} &= e_q\left(-\hat{X}_q(\lambda_1,\lambda_2)\right) \Big/ \tilde{Z}_q\left(\lambda_1,\lambda_2\right) \ , \\ \tilde{Z}_q\left(\lambda_1,\lambda_2\right) &= \mathrm{Tr} e_q\left(-\hat{X}_q(\lambda_1,\lambda_2)\right) \ , \\ \hat{X}_q\left(\lambda_1,\lambda_2\right) &= \left[\lambda_1(\hat{H}-\tilde{U}_q)+\lambda_2(\hat{N}-\tilde{N}_q)\right]/c_q \ , \\ e_q(\hat{A}) &= \begin{cases} \left[1+(1-q)\hat{A}\right]^{1/(1-q)} \ , \ \text{if Spec}\left[1+(1-q)\hat{A}\right] \geq 0 \ , \\ 0 \ , \quad \text{otherwise} \ , \end{cases} \end{aligned} \tag{56}$$

where $c_q \equiv \mathrm{Tr}\hat{\tilde{\rho}}^q$. Here, Spec $\left[1+(1-q)\hat{A}\right] \geq 0$ means that there is a natural cut off when the spectrum of the operator in the parentheses has negative values, for reasons of reality of the trace. In the limit $q \to 1$, there is the condition of the boundedness from below of the spectrum of the Hamiltonian. λ_1 and λ_2 are the Lagrange multipliers that go with the energy and number constraints defined above, with the interpretation of inverse "temperature" and "chemical potential" which will be introduced later. It is in this form that the Tsallis variational principle has the desirable properties; invariance with respect to shift of the origin of the energy, q-mean value of a c-number being itself, and additivity of the energy. These nice features come with the renormalization of the Lagrange multipliers with a self-consistency relation for c_q:

$$c_q = \mathrm{Tr}[e_q(-\hat{X}_q(\lambda_1, \lambda_2))]^q \, / \, (\tilde{Z}_q)^q \ . \tag{57}$$

From Eq. (56), we also have a relation between c_q and $\tilde{Z}_q$. Noting

$$\left(\hat{\rho}\tilde{Z}_q\right)^{1-q} = 1 - (1-q)\hat{X}_q(\lambda_1, \lambda_2) \ ,$$

we have

$$c_q = (\tilde{Z}_q)^{1-q} \ . \tag{58}$$

From Eqs. (57) and (58), we obtain another expression for $\tilde{Z}_q$:

$$\tilde{Z}_q = \mathrm{Tr}[e_q(-\hat{X}_q(\lambda_1, \lambda_2))]^q \ . \tag{59}$$

With these relations, one finds that the normalized q-mean value takes the form

$$\left\langle \hat{A} \right\rangle_q \equiv \tilde{A}_q = \mathrm{Tr}\left\{ \left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2) \right) \right]^q \hat{A} \right\} / \tilde{Z}_q \ . \tag{55a}$$

In this form, only the connected terms appear in any diagrammatic expression for the q-mean value in perturbation theory.

One observes, from Eq. (56) and the definitions in Eq. (55), that the following relations hold:

$$\frac{\partial \tilde{Z}_q}{\partial \lambda_1} = 0 = \frac{\partial \tilde{Z}_q}{\partial \lambda_2} \ . \tag{60}$$

The corresponding "fluctuations" calculated from the second derivatives are

$$\begin{aligned}
&\left(\tilde{Z}_q\right)^{-q} \frac{\partial^2 \tilde{Z}_q}{\partial \lambda_1^2} = q\left(\tilde{Z}_q\right)^{-q} \\
&\times \mathrm{Tr}\left\{ \left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2) \right) \right]^{2q-1} \frac{(\hat{H} - \tilde{U}_q)^2}{c_q^2} \right\} \ , \\
&\left(\tilde{Z}_q\right)^{-q} \frac{\partial^2 \tilde{Z}_q}{\partial \lambda_1^2} = q\left(\tilde{Z}_q\right)^{-q} \\
&\times \mathrm{Tr}\left\{ \left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2) \right) \right]^{2q-1} \frac{(\hat{N} - \tilde{N}_q)^2}{c_q^2} \right\} \ , \\
&\left(\tilde{Z}_q\right)^{-q} \frac{\partial^2 \tilde{Z}_q}{\partial \lambda_1 \partial \lambda_2} = q\left(\tilde{Z}_q\right)^{-q} \\
&\times \mathrm{Tr}\left\{ \left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2) \right) \right]^{2q-1} \frac{(\hat{H} - \tilde{U}_q)(\hat{N} - \tilde{N}_q)}{c_q^2} \right\} \ .
\end{aligned} \tag{61}$$

It is important to point out that these are not the true fluctuations of the type in Eq. (15) even though these reduce to those given there in the limit $q = 1$. In fact, in the notation of Eq. (15), we have explicitly the following expressions to exhibit this difference:

$$\begin{aligned}
&\left(\tilde{Z}_q\right)^{-q} \frac{\partial^2 \tilde{Z}_q}{\partial \lambda_1^2} = -\frac{\partial \tilde{U}_q}{\partial \lambda_1} \equiv \tilde{\sigma}_q^2(\hat{H}) \quad \left(\neq \left\langle (\hat{H} - \tilde{U}_q)^2 \right\rangle_q\right) , \\
&\left(\tilde{Z}_q\right)^{-q} \frac{\partial^2 \tilde{Z}_q}{\partial \lambda_2^2} = -\frac{\partial \tilde{N}_q}{\partial \lambda_2} \equiv \tilde{\sigma}_q^2(\hat{N}) \quad \left(\neq \left\langle (\hat{N} - \tilde{N}_q)^2 \right\rangle_q\right) , \\
&\left(\tilde{Z}_q\right)^{-q} \frac{\partial^2 \tilde{Z}_q}{\partial \lambda_1 \partial \lambda_2} = -\frac{\partial \tilde{U}_q}{\partial \lambda_2} = -\frac{\partial \tilde{N}_q}{\partial \lambda_1} \\
&\qquad \equiv \tilde{\sigma}_q^c(\hat{H}, \hat{N}) \quad \left(\neq \left\langle (\hat{H} - \tilde{U}_q)(\hat{N} - \tilde{N}_q) \right\rangle_q\right) . \qquad (61a)
\end{aligned}$$

In deriving these results, we used Eq. (60) with the equivalent expression for $\tilde{Z}_q$ given in Eq. (59), which shows

$$\begin{aligned}
&\mathrm{Tr}\left\{ \left[e_q(-\hat{X}_q(\lambda_1, \lambda_2)) \right]^{2q-1} (\hat{H} - \tilde{U}_q) \right\} = 0 , \\
&\mathrm{Tr}\left\{ \left[e_q(-\hat{X}_q(\lambda_1, \lambda_2)) \right]^{2q-1} (\hat{N} - \tilde{N}_q) \right\} = 0 . \qquad (61b)
\end{aligned}$$

Eq. (60) assures us of the extremum feature of the Tsallis entropy. Furthermore, that the Tsallis entropy takes its maximum at these places requires negative definiteness of the matrix given in Eq. (61a), $\partial^2 \tilde{Z}_q / \partial \lambda_i \partial \lambda_j$. This leads to the requirements that

$$\begin{aligned}
&\tilde{\sigma}_q^2(\hat{H}),\ \tilde{\sigma}_q^2(\hat{N}) > 0 , \\
&\tilde{\sigma}_q^2(\hat{H})\ \tilde{\sigma}_q^2(\hat{N}) - \left[\tilde{\sigma}_q^c(\hat{H}, \hat{N}) \right]^2 > 0 . \qquad (62)
\end{aligned}$$

These are the q-versions of Eq. (15). The connection to thermodynamic language is obtained when as before we formally identify the Lagrange multipliers λ_1 and λ_2 with the inverse temperature β and $-\beta\mu$ with the chemical potential μ, respectively. We can now define a "partition function" which gives us the usual results by appropriately shifting the origins of the energy and number. It leads to the Legendre tranformation when we change our reference from given constraints to given Lagrange multipliers as in the conventional theory discussed earlier. The maximum value of the entropy is now given by

$$S_q = -\frac{1 - c_q}{1 - q} = \frac{\left(\tilde{Z}_q\right)^{1-q} - 1}{1 - q} \equiv \ln_q \tilde{Z}_q , \qquad (63)$$

where $\ln_q x \equiv (x^{1-q} - 1)/(1 - q) = x^{1-q} \mathrm{Ln}_q x$.

We now define the "q-free energy" as follows:

$$\tilde{F}_q \equiv \tilde{U}_q - \mu \tilde{N}_q - \frac{1}{\beta}\, S_q . \qquad (64)$$

By using the definitions and the relations derived above, the right hand side of this expression is found to be

$$\tilde{F}_q = \frac{1}{\beta(1-q)}\left[1 - \left(\tilde{Z}_q\right)^{1-q} + \beta(1-q)\left(\tilde{U}_q - \mu\tilde{N}_q\right)\right] . \tag{65}$$

We redefine the "q-partition function" by

$$\begin{aligned}\left(Z_q\right)^{1-q} &= \left(\tilde{Z}_q\right)^{1-q} - \beta(1-q)\left(\tilde{U}_q - \mu\tilde{N}_q\right) , \\ \tilde{F}_q &= -\frac{1}{\beta}\ \ln_q Z_q .\end{aligned} \tag{66}$$

From these, we obtain the familiar looking expressions

$$\begin{aligned}\frac{\partial \ln_q Z_q}{\partial \beta} &= -\tilde{U}_q , \\ \frac{\partial \ln_q Z_q}{\partial\left(\mu\beta\right)} &= \tilde{N}_q .\end{aligned} \tag{67}$$

Using Eqs. (60), (66), and (67), we also obtain

$$\frac{\partial S_q}{\partial \tilde{U}_q} = \beta , \quad \frac{\partial S_q}{\partial \tilde{N}_q} = -\beta\mu . \tag{68}$$

The matrix of second derivatives of S_q with respect to $\tilde{U}_q$ and $\tilde{N}_q$ is then minus the inverse matrix of the second derivatives of $\ln_q \tilde{Z}_q$ w.r.t. λ_1 and λ_2.

As in the case of the limit $q \to 1$ discussed earlier, in the limit of zero temperature, λ_1 goes to ∞, and we have $S_q = 0$, thus reaching the lowest (ground) state of the Hamiltonian with eigenvalue, $\tilde{U}_q \to E_0$. In this limit, we also have $\tilde{N}_q \to 0$ and $c_q \to 1$. This shows the possible q-generalization of the annealing algorithm mentioned earlier.

This completes the discussion of the development of equilibrium statistical mechanics of nonextensive systems using the Tsallis ensemble.

Now, to obtain the expressions for the isothermal admittance, we consider a static applied force X which couples to the system via a hermitian operator, $\hat{A}$, and calculate the response of a normalized q-mean value of the required physical entity represented by a Hermitian operator, $\hat{B}$, to this external static probing field. Thus we need to calculate the expression

$$\begin{aligned}\left\langle \Delta\hat{B}\right\rangle_q^{(\mathrm{iso})} &= \mathrm{Tr}\left[\hat{\tilde{P}}\left(\hat{H} - X\hat{A}; q, \tilde{\beta}\right)\hat{B}\right] - \mathrm{Tr}\left[\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\beta}\right)\hat{B}\right] \\ &\equiv \tilde{\chi}_{BA}^{(q)}(T)X ,\end{aligned} \tag{69}$$

to first order in X, which we have expressed in the conventional linear response form, and which defines the isothermal admittance, $\tilde{\chi}_{BA}^{(q)}(T)$. Here, for simplicity, we have employed the canonical ensemble scheme with only the normalized q-mean value of the Hamiltonian as a constraint in the maxent principle and so we obtain

$$\begin{aligned}
\hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right) &= \hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right)/\tilde{Z}_q\left(\tilde{\beta}\right) \ , \\
\hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right) &= \left[e_q\left(-\tilde{\beta}\left(\hat{H}-\tilde{U}_q\right)\right)\right]^q \ , \\
\tilde{Z}_q\left(\tilde{\beta}\right) &= \mathrm{Tr}e_q\left(-\tilde{\beta}\left(\hat{H}-\tilde{U}_q\right)\right) \ , \qquad (70)
\end{aligned}$$

where $\tilde{\beta} = \beta/c_q$ and $c_q = \left[\tilde{Z}_q\left(\tilde{\beta}\right)\right]^{1-q}$. It should be stressed that the expressions given in Eq. (70) are the results following from the application of the maximum Tsallis entropy procedure subject to the given constraint on the normalized q-mean value of the Hamiltonian. The notation of tilde on the operators is used to indicate such maxent solutions. [This feature implies that we have to hold the constraint fixed while calculating the derivative with respect to the Lagrange multiplier. Thus, we have the relations of the forms given in Eqs. (60) and (61b). With these in mind, the following results in the Identity, Theorem A, and Theorem B can be derived. Also, we must point out that, in the context of linear response theory, we are considering small deviations from equilibrium even when an external probe is applied to the system. The interested reader may obtain the basic steps required in deriving the results below by requesting the author for a copy of proofs of the Identity, Theorem A, and Theorem B.] To calculate the difference in Eq. (69), we analyze the operator $\hat{\tilde{Q}}\left(\hat{H}-X\hat{A};q,\tilde{\beta}\right)$. We first give a general result in the form of an identity which is a generalization of that due to Karplus and Schwinger [22] for the exponential operators to which ours reduces when $q \to 1$.

Identity:

$$\begin{aligned}
\hat{\tilde{Q}}\left(\hat{H}-X\hat{A};q,\tilde{\beta}\right) &= \hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right) + qX\hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right) \\
&\quad \times \int_0^{\tilde{\beta}} d\tilde{\lambda}\left[\hat{\tilde{Q}}\left(\hat{H};q,\tilde{\lambda}\right)\right]^{-1}\hat{\tilde{A}}\left(\hat{H},\hat{A};q,\tilde{\lambda}\right)\hat{\tilde{Q}}\left(\hat{H}-X\hat{A};q,\tilde{\lambda}\right) \ , \\
\hat{\tilde{A}}\left(\hat{H},\hat{A};q,\tilde{\lambda}\right) &\equiv \left[1-\tilde{\lambda}(1-q)\left(\hat{H}-\tilde{U}_q-X\left(\hat{A}-\tilde{A}_q\right)\right)\right]^{-1} \\
&\quad \times \left(\hat{A}-\tilde{A}_q\right)\left[1-\tilde{\lambda}(1-q)\left(\hat{H}-\tilde{U}_q\right)\right]^{-1} \ . \qquad (71)
\end{aligned}$$

This identity is established by calculating the derivative of $\left\{\left[\hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right)\right]^{-1}\hat{\tilde{Q}}\left(\hat{H}-X\hat{A};q,\tilde{\beta}\right)\right\}$ with respect to $\tilde{\beta}$ explicitly with holding $\tilde{A}_q$ and $\tilde{U}_q$, rearranging the resulting terms, integrating the expression so obtained, and finally expressing it in the form given above. To first order in X, the above expression simplifies to the form

$$\hat{\tilde{Q}}\left(\hat{H}-X\hat{A};q,\tilde{\beta}\right) \cong \hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right) + qX\hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right)$$

$$
\begin{aligned}
&\times \int_0^{\tilde{\beta}} d\tilde{\lambda} \left[\hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\lambda}\right)\right]^{-1} \hat{\tilde{A}}\left(\hat{H}; q, \tilde{\lambda}\right) \hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\lambda}\right) \ , \\
\hat{\tilde{A}}\left(\hat{H}; q, \tilde{\lambda}\right) &= \left[1 - \tilde{\lambda}(1-q)\left(\hat{H} - \tilde{U}_q\right)\right]^{-1} \left(\hat{A} - \tilde{A}_q\right) \\
&\times \left[1 - \tilde{\lambda}(1-q)\left(\hat{H} - \tilde{U}_q\right)\right]^{-1} \ . \qquad (72)
\end{aligned}
$$

Considering a similar equation for $e_q\left(\hat{H} - X\hat{A}; q, \tilde{\beta}\right)$, which has the same appearance as Eq. (72) but with no factor of q and only the operator difference $\left(\hat{A} - \tilde{A}_q\right)$ in the second terms of Eq. (72), and taking its trace to obtain the change in the partition function expressed in the equivalent form in Eq. (70), due to the presence of the perturbation, we find that there is no change in the partition function in the presence of the external field to first order in X:

$$
\begin{aligned}
&\tilde{Z}_q\left(\hat{H} - X\hat{A}; \hat{\beta}\right) \cong \tilde{Z}_q\left(\tilde{\beta}\right) + \\
&X\int_0^{\tilde{\beta}} d\tilde{\lambda}\ \mathrm{Tr}\left\{\hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\beta}\right)\left[\hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\lambda}\right)\right]^{-1}\left(\hat{A} - \tilde{A}_q\right)\hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\lambda}\right)\right\} \\
&= \tilde{Z}_q\left(\tilde{\beta}\right) \ , \qquad (73)
\end{aligned}
$$

since the second trace term in Eq. (73) vanishes upon using the cyclic property of trace operation. Then, Eq. (72) with this result yields an equation for $\hat{\tilde{P}}\left(\hat{H} - X\hat{A}; q, \tilde{\beta}\right)$:

$$
\begin{aligned}
\hat{\tilde{P}}\left(\hat{H} - X\hat{A}; q, \hat{\beta}\right) &\cong \hat{\tilde{P}}\left(\hat{H}; q, \tilde{\beta}\right) + qX\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\beta}\right) \\
&\times \int_0^{\tilde{\beta}} d\tilde{\lambda}\left[\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\lambda}\right)\right]^{-1}\hat{\tilde{A}}\left(\hat{H}; q, \tilde{\lambda}\right)\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\lambda}\right) \ . \qquad (74)
\end{aligned}
$$

From Eqs. (69) and (74), we finally obtain the isothermal response function in the form

$$
\begin{aligned}
\tilde{\chi}_{BA}^{(q)} = q\int_0^{\tilde{\beta}} d\tilde{\lambda}\mathrm{Tr}&\left\{\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\beta}\right)\left[\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\lambda}\right)\right]^{-1}\right. \\
&\left.\times\hat{\tilde{A}}\left(\hat{H}; q, \tilde{\lambda}\right)\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\lambda}\right)\hat{B}\right\} \ . \qquad (75)
\end{aligned}
$$

It is important to see if this expression can be derived from the free energy expression in the form given by Eq. (65) but now calculated for a Hamiltonian of the form $\hat{H} - X\hat{A} - Y\hat{B}$. Thus

$$
\begin{aligned}
&\tilde{F}_q = -\frac{1}{\beta}\ \ln_q Z_q \ , \\
&\left(Z_q\right)^{1-q} = \left(\tilde{Z}_q\right)^{1-q} - \beta(1-q)\left(\tilde{U}_q - X\tilde{A}_q - Y\tilde{B}_q\right) \ , \qquad (76) \\
&\tilde{Z}_q = \mathrm{Tr}e_q\left(-\tilde{\beta}\left[\hat{H} - \tilde{U}_q - X\left(\hat{A} - \tilde{A}_q\right) - Y\left(\hat{B} - \tilde{B}_q\right)\right]\right) \ .
\end{aligned}
$$

Taking the first partial derivative of this expression with respect to Y yields (remembering that the mean values appearing in these expressions are held fixed as given constraints in deriving the above, as in the derivation of Eq. (60))

$$-\frac{\partial \tilde{F}_q}{\partial Y} = \tilde{B}_q = \mathrm{Tr}\left[\hat{\tilde{P}}\left(\hat{H} - X\hat{A} - Y\hat{B}; q, \tilde{\beta}\right)\hat{B}\right] . \tag{77}$$

From Eq. (74), we see at once that

$$\left.\frac{\partial^2 \tilde{F}_q}{\partial X \partial Y}\right|_{X=0=Y} = \tilde{\chi}^{(q)}_{BA}(T) . \tag{78}$$

This is a familiar form for the isothermal susceptibility, here derived for the normalized q-mean value formalism. Maxent implies minimum free energy and so the stability condition may be stated as the positivity of the isothermal response matrix, if we have more than one external probing fields. This means that only then the system is stable with respect to external perturbations.

We give below two theorems in the normalized q-mean value formalism which generalize similar theorems recently given in Ref. [23].

THEOREM A:

$$\begin{aligned}
&\frac{\partial \hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right)}{\partial \lambda} \\
&\qquad = -q \int_0^{\tilde{\beta}} d\tilde{u}\, \hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right) \left[\hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{u}\right)\right]^{-1} \\
&\qquad\quad \times \frac{\partial \hat{\tilde{A}}(\lambda, \tilde{u})}{\partial \lambda}\, \hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{u}\right) , \\
&\frac{\partial \hat{\tilde{A}}(\lambda, \tilde{u})}{\partial \lambda} = \left[1 - \tilde{u}(1-q)\left(\hat{H}(\lambda) - \tilde{U}_q(\lambda)\right)\right]^{-1} \frac{\partial\left(\hat{H}(\lambda) - \tilde{U}_q(\lambda)\right)}{\partial \lambda} \\
&\qquad\qquad\qquad \times \left[1 - \tilde{u}(1-q)\left(\hat{H}(\lambda) - \tilde{U}_q(\lambda)\right)\right]^{-1} .
\end{aligned} \tag{79}$$

Here, the Hamiltonian is assumed to depend on a parameter λ.
Proof: Consider

$$\begin{aligned}
&\hat{f}_q\left(\lambda, \tilde{\beta}\right) = \partial \hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right) \Big/ \partial \lambda , \\
&\hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right) = \left[e_q\left(-\tilde{\beta}\left(\hat{H}(\lambda) - \tilde{U}_q(\lambda)\right)\right)\right]^q , \\
&\tilde{U}_q(\lambda) = \mathrm{Tr}\left[\hat{H}(\lambda)\hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right)\right] \Big/ \tilde{Z}_q\left(\lambda, \tilde{\beta}\right) .
\end{aligned} \tag{80}$$

Differentiating the first expression in Eq. (80) w.r.t. $\tilde{\beta}$, and interchanging the orders of differentiation and carrying out the $\tilde{\beta}$-differen-tiation and remembering the constraint while differentiating the various terms in Eq. (80), we obtain

$$\begin{aligned}
\frac{\partial \hat{f}_q\left(\lambda,\tilde{\beta}\right)}{\partial\tilde{\beta}} &+ q\left[1-\tilde{\beta}(1-q)\left(\hat{H}(\lambda)-\tilde{U}_q(\lambda)\right)\right]^{-1} \\
&\times\left(\hat{H}(\lambda)-\tilde{U}_q(\lambda)\right)\hat{f}_q\left(\lambda,\tilde{\beta}\right) \\
&= -q\,\frac{\partial\hat{\tilde{A}}\left(\lambda,\tilde{\beta}\right)}{\partial\lambda}\,\hat{\tilde{Q}}\left(\hat{H}(\lambda);q,\tilde{\beta}\right), \\
\frac{\partial\hat{\tilde{A}}\left(\lambda,\tilde{\beta}\right)}{\partial\lambda} &= \left[1-\tilde{\beta}(1-q)\left(\hat{H}(\lambda)-\tilde{U}_q(\lambda)\right)\right]^{-1}\frac{\partial\left(\hat{H}(\lambda)-\tilde{U}_q(\lambda)\right)}{\partial\lambda} \\
&\times\left[1-\tilde{\beta}(1-q)\left(\hat{H}(\lambda)-\tilde{U}_q(\lambda)\right)\right]^{-1}\,. \qquad (81)
\end{aligned}$$

By a direct calculation, this equation may be rewritten in the form

$$\begin{aligned}
\hat{\tilde{Q}}\left(\hat{H}(\lambda);q,\tilde{\beta}\right)\,&\frac{\partial\left\{\left[\hat{\tilde{Q}}\left(\hat{H}(\lambda);q,\tilde{\beta}\right)\right]^{-1}\hat{f}_q\left(\lambda,\tilde{\beta}\right)\right\}}{\partial\tilde{\beta}} \\
&= -q\,\frac{\partial\hat{\tilde{A}}\left(\lambda,\tilde{\beta}\right)}{\partial\lambda}\,\hat{\tilde{Q}}\left(\hat{H}(\lambda);q,\tilde{\beta}\right)\,. \qquad (82)
\end{aligned}$$

Integrating this equation using the facts that $\hat{\tilde{Q}}\left(\hat{H}(\lambda);q,\tilde{\beta}=0\right)=0$ and $\hat{f}_q\left(\lambda,\tilde{\beta}=0\right)=0$, we obtain the theorem stated above. This theorem reduces to the Wilcox theorem for an exponential operator in the limit $q\to 1$. An immediate consequence of this theorem leads to a generalization of a result due to Kubo [24]:

THEOREM B: If $\hat{A}$ is an arbitrary operator, then the commutator $\left[\hat{A},\hat{\tilde{Q}}\left(\hat{H};q,\tilde{\beta}\right)\right]$ is given by

$$\begin{aligned} \left[\hat{A}, \hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\beta}\right)\right] &= q\hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\beta}\right) \int_0^{\tilde{\beta}} d\tilde{u} \left[\hat{\tilde{Q}}\left(\hat{H}; q, \tilde{u}\right)\right]^{-1} \\ &\qquad \times \left[\hat{H}, \hat{\tilde{A}}(\tilde{u})\right] \hat{\tilde{Q}}\left(\hat{H}; q, \tilde{u}\right) \ , \\ \hat{\tilde{A}}(\tilde{u}) &= \left[1 - \tilde{u}(1-q)\left(\hat{H} - \tilde{U}_q\right)\right]^{-1} \hat{A} \left[1 - \tilde{u}(1-q)\left(\hat{H} - \tilde{U}_q\right)\right]^{-1} \ . \end{aligned} \tag{83}$$

Proof: We deduce this theorem by making a similarity transformation of $\hat{H}$ as $\hat{H}(\lambda) = e^{\lambda\hat{A}} H e^{-\lambda\hat{A}}$ in Theorem A. Then, it follows that

$$\hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right) = e^{\lambda\hat{A}} \hat{\tilde{Q}}\left(\hat{H}; q, \tilde{\beta}\right) e^{-\lambda\hat{A}} \ , \tag{84}$$

and similarly the other operators appearing in Eqs. (80) and (81) upon using their definitions. Therefore, the following expressions that appear in Eq. (79) finally lead to Eq. (83):

$$\begin{aligned} \frac{\partial \hat{\tilde{Q}}\left(\hat{H}(\lambda); q, \tilde{\beta}\right)}{\partial\lambda} &= e^{\lambda\hat{A}} \left[\hat{A}, \hat{\tilde{Q}}\left(\hat{A}; q, \tilde{\beta}\right)\right] e^{-\lambda\hat{A}} \ , \\ \frac{\partial \hat{\tilde{A}}\left(\lambda, \tilde{u}\right)}{\partial\lambda} &= e^{\lambda\hat{\tilde{A}}} \left[\hat{\tilde{A}}(\tilde{u}), \hat{H}\right] e^{-\lambda\hat{\tilde{A}}} \ . \end{aligned}$$

4 Time-Dependence: Unitary Dynamics

In this section, we will discuss the traditional unitary time evolution of the density matrix, (a) the Liouville-von Neumann equation for the density matrix. After a few observations about this equation, we derive (b) the dynamical linear response of a nonextensive system to a time-dependent external probe [11], which is the counterpart of the isothermal response discussed in the last section. More generally, (c) unlike the published versions so far [25], the Green function theory of many-particle systems [26] is developed in terms of the normalized q-mean values [27].

The density matrix $\hat{\rho}(t)$ obeys the standard quantum Liouville-von Neumann equation:

$$i\hbar \, \frac{\partial\hat{\rho}(t)}{\partial t} = \left[\hat{H}(t), \hat{\rho}(t)\right] \ , \tag{85}$$

where $\hat{H}(t)$ is in general a time-dependent hermitian Hamiltonian, with a given initial condition $\hat{\rho}(t = 0)$ at, say, $t = 0$. This is a linear equation for $\hat{\rho}(t)$ and thus the superposition principle of quantum theory is incorporated. One immediate consequence of this equation is that it preserves the hermiticity of the density matrix as well as its trace, both of which are easily verified by direct calculation. Introduce the time evolution operator

$$\hat{U}(t) = \mathrm{T}\exp\left[-\frac{i}{\hbar}\int_0^t dt'\hat{H}(t')\right] \ , \tag{86}$$

where T stands for the time-ordering symbol. This operator is unitary, i.e., $\hat{U}^\dagger(t)\hat{U}(t) = \hat{I} = \hat{U}(t)\hat{U}^\dagger(t)$, where $\hat{U}^\dagger(t)$ is the hermitian conjugate of the time evolution operator. It is found that they obey the equations

$$i\hbar\frac{\partial \hat{U}(t)}{\partial t} = \hat{H}(t)\hat{U}(t) \ , \quad -i\hbar\ \frac{\partial \hat{U}^\dagger(t)}{\partial t} = \hat{U}^\dagger(t)\hat{H}(t) \ . \tag{87}$$

Often, the first and second equations in Eq. (87) are called the “forward” and the “backward” propagating equations, respectively. This terminology is useful later, when we discuss nonunitary time evolution. This shows that the time evolution of the density matrix is just a unitary transformation of the given initial density matrix:

$$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t) \ . \tag{88}$$

Eq. (88) is the solution to Eq. (85), and has the properties of a density matrix. This shows that the trace of the density matrix is independent of time. Accordingly, the entropy functional in both the forms of Eqs. (1) and (2) are time-independent and thus do not evolve in time. This is because the time dependence of any well behaved operator functional of the initial density matrix also evolves in time unitarily:

$$\hat{F}\left[\hat{\rho}(t)\right] = \hat{U}\hat{F}\left[\hat{\rho}(0)\right]\hat{U}^\dagger(t) \ . \tag{89}$$

The mean value of any operator $\hat{A}(t)$, which may have its own private time dependence (time-dependent electric field, for example), is defined by the normalized q-mean value

$$\begin{aligned}\left\langle \hat{A}(t) \right\rangle_q &= \mathrm{Tr}\left[\hat{A}(t)\hat{\rho}^q(t)\right] \Big/ \mathrm{Tr}\hat{\rho}^q(t) \\ &= \mathrm{Tr}\left[\hat{U}^\dagger(t)\hat{A}(t)\hat{U}(t)\hat{\rho}^q(0)\right] \Big/ \mathrm{Tr}\hat{\rho}^q(0) \ . \end{aligned} \tag{90}$$

This equation shows that there is time evolution of the q-mean value of an operator, and this fact is used in the next stage of our development of the theory. It may not be out of place here to point out that the Wigner transformation of the density matrix defines a quantum phase-space distribution function, f_W, which shares all the properties of the classical phase space distribution except that its quantum nature reveals itself by not being positive everywhere. The corresponding time-dependent equation arises from Eq. (85) which is the Fokker-Planck-like equation for f_W. These are standard textbook results collected here for easy reference. It should be pointed out that there is a “stationary action principle” that replaces the maxent principle in the time-independent case discussed earlier. This is weaker than the maxent principle in the sense that the action principle only gives Eq. (85) as the stationary point of the action when the initial condition on the density matrix is specified. Often, one uses the maxent principle to specify this initial density matrix. But in quantum optics, for example, one uses the density matrix of an initially prepared state, such as the coherent state of the radiation (“laser”) and the initially prepared state of the “atom” as in the JC model. See, for example, Ref. [18]. This completes the general discussion of unitary time evolution of the density matrix and the concomitant time evolution of entropy in general and mean value of any operator.

The first application of the above time-dependent equation is the calculation of the dynamical response function of a system to an external time-dependent probe. This is a perturbation theory result. We present here the formulation in terms of normalized q-mean values unlike in our earlier published work [11]. In developing the dynamical linear response of a physical quantity represented by a hermitian operator, $\hat{B}$, we consider its equilibrium normalized q-mean value $\left\langle \hat{B} \right\rangle_q^{(0)}$ considered as the given initial value, being driven away by means of a time-dependent external field $X(t)$. In Table 6, we give some commonly occurring probing fields and the corresponding linear response functions.

Table 6. Examples of linear response to $H_{\text{ext}} = -X(t)\hat{A}$

$\hat{A}$	$X(t)$	χ_{AB} **(Response)**
Density, $\hat{n}(\overrightarrow{r})$	Electric potential, $\phi(\overrightarrow{r}, t)$	Density-density longitudinal dielectric function, $\chi_{n,n}$
Spin density (magnetization), $\overrightarrow{S}(\overrightarrow{r})$ or $\overrightarrow{M}(\overrightarrow{r})$	Magnetic field, $\overrightarrow{B}(\overrightarrow{r}, t)$	Magnetic susceptibility tensor, $\overset{\leftrightarrow}{\chi}_{s_i, s_j}$
Current density, $\overrightarrow{J}(\overrightarrow{r}, t)$	Vector potential, $\overrightarrow{A}(\overrightarrow{r}, t)$	Current-current correlation function, dieletric tensor

In the usual way, the q-mean value of the deviation $\hat{B} - \left\langle \hat{B} \right\rangle_q^{(0)} \equiv \Delta\hat{B}$ is given by

$$\begin{aligned} \langle \Delta\hat{B}(t)\rangle_q &= \mathrm{Tr}\left[\hat{\rho}^q(t)\hat{B}\right] \Big/ \mathrm{Tr}\hat{\rho}^q(t) - \mathrm{Tr}\left[\hat{\rho}^q(0)\hat{B}\right] \Big/ \mathrm{Tr}\hat{\tilde{\rho}}^q(0) \\ &= \mathrm{Tr}\left[\hat{\rho}^q(t)\Delta\hat{B}\right] \Big/ \mathrm{Tr}\hat{\rho}^q(t) \end{aligned} \tag{91}$$

with $\mathrm{Tr}\hat{\rho}(t) = 1 = \mathrm{Tr}\hat{\tilde{\rho}}(0)$. Here, $\hat{\tilde{\rho}}(0)$ is the equilibrium density matrix determined from the maximum Tsallis entropy of the system for given constraints and $\hat{\rho}(t)$ is the time-dependent density matrix obeying the quantum Liouville-von Neumann equation determined by the Hamiltonian incorporating the effect of $X(t)$:

$$\hat{H} - X(t)\hat{A} \, . \tag{92}$$

$\hat{A}$ is the operator conjugate to the time-dependent field $X(t)$. For simplicity of presentation, we use here the equilibrium canonical ensemble prescription as before. Thus the equilibrium density matrix with temperature (Lagrange) parameter β is given as in Eq. (70). (See the development above.)
We rewrite Eq. (91) in terms of this notation:

$$\left\langle \Delta\hat{B}(t)\right\rangle_q \equiv \mathrm{Tr}\left[\hat{P}(t)\hat{B}\right] - \mathrm{Tr}\left[\hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right)\hat{B}\right] \;, \tag{93}$$

where $\hat{P} = \hat{\rho}^q(t)/\mathrm{Tr}\hat{\rho}^q(t)$. The density matrix $\hat{\rho}(t)$ now obeys the standard quantum Liouville-von Neumann equation

$$i\hbar\,\frac{\partial\hat{\rho}(t)}{\partial t} = \left[\hat{H} - X(t)\hat{A}, \hat{\rho}(t)\right] \tag{94}$$

with the initial condition as given above in Eq. (70). We take the initial condition specified at some remote past, by considering **the initial time $t = 0$ in Eq. (91) as now pushed back to that at** $t = -\infty$. This time evolution is unitary as shown above and so the equation obeyed by any power of $\hat{\rho}(t)$ has the same form as in Eq. (94):

$$i\hbar\,\frac{\partial\hat{\rho}^q(t)}{\partial t} = \left[\hat{H} - X(t)\hat{A}, \hat{\rho}^q(t)\right] \;. \tag{95}$$

Taking traces on both sides of this equation, we have the important result that $c_q = \mathrm{Tr}\hat{\rho}^q(t)$ is independent of time, t, and hence the equation for $\hat{P}(t)$ in Eq. (93) takes the same form as Eq. (95), which we rewrite in the form

$$i\hbar\,\frac{\partial\hat{P}(t)}{\partial t} = \left[\hat{H}, \hat{P}(t)\right] - \left[\hat{A}, \hat{P}(t)\right]X(t) \tag{96}$$

with $\hat{P}(t = -\infty) = \hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right)$ in Eq. (70). This equation is converted into an integral equation by a standard procedure so that a solution to linear order in $X(t)$ can be found as follows:

$$\begin{aligned}
\hat{P}(t) &= \hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right) - \frac{1}{i\hbar}\int_{-\infty}^{t} dt' e^{-i\hat{H}(t-t')/\hbar} \\
&\qquad\qquad \times \left[\hat{A}, \hat{P}(t')\right] X(t') e^{i\hat{H}(t-t')/\hbar} \\
&\cong \hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right) - \frac{1}{i\hbar}\int_{-\infty}^{t} dt' e^{-i\hat{H}(t-t')/\hbar} \\
&\qquad\qquad \times \left[\hat{A}, \hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right)\right] X(t') e^{i\hat{H}(t-t')/\hbar} \;.
\end{aligned} \tag{97}$$

From Eqs. (93) and (97), we find

$$\left\langle \Delta\hat{B}(t)\right\rangle_q = \int_{-\infty}^{t} dt' \tilde{\phi}_{BA}^{(q)}(t-t')X(t') \;,$$

$$\begin{aligned}\tilde{\phi}^{(q)}_{BA}(t) &= -\frac{1}{i\hbar}\,\mathrm{Tr}\left\{\left[\hat{A},\hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right)\right]\hat{B}(t)\right\}\\ &= \frac{1}{i\hbar}\,\mathrm{Tr}\,\left\{\left[\hat{A},\hat{B}(t)\right]\hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right)\right\}\,,\end{aligned} \tag{98}$$

where $\hat{B}(t) = e^{i\hat{H}t/\hbar}\hat{B}e^{-i\hat{H}t/\hbar}$. If $X(t) = X_0\cos\omega t$, which is the usual experimental harmonic applied field, Eq. (98) may be expressed in a more familiar form,

$$\begin{aligned}\left\langle \Delta\hat{B}(t)\right\rangle_q &= \int_{-\infty}^{t} dt'\tilde{\phi}^{(q)}_{BA}(t-t')X(t')\\ &= \int_{-\infty}^{t} dt'\tilde{\phi}^{(q)}_{BA}(t-t')X_0\left[e^{i\omega t'}+e^{-i\omega t'}\right]\Big/2\\ &= \int_{0}^{\infty} dx\tilde{\phi}^{(q)}_{BA}(x)X_0\left[e^{-i\omega x}e^{i\omega t}+e^{i\omega x}e^{-i\omega t}\right]\Big/2\\ &= \mathrm{Re}\left\{\tilde{\chi}^{(q)}_{BA}(\omega)X_0e^{i\omega t}\right\}\,,\end{aligned} \tag{99}$$

where $\tilde{\chi}^{(q)}_{BA}(\omega) = \lim_{\varepsilon\to+0}\int_0^\infty dx\tilde{\phi}^{(q)}_{BA}(x)e^{-i\omega x-\varepsilon x}$.

We will first deduce from this the counterparts of all the known relationships in the extensive ($q \to 1$) case. To this end, define the "relaxation function" as the relaxation of $\left\langle \Delta\hat{B}(t)\right\rangle_q$ after the removal of the external disturbance as in Kubo [24]:

$$\begin{aligned}\tilde{\Phi}^{(q)}_{BA}(t) &= \lim_{\varepsilon\to+0}\int_{t}^{\infty} dt'\tilde{\phi}^{(q)}_{BA}(t')e^{-\varepsilon t'}\\ &= \sum_{i,j}\left(\frac{\tilde{P}_q(i)-\tilde{P}_q(j)}{E_i-E_j}\right)\langle i|\hat{A}|j\rangle\langle j|\hat{B}|i\rangle e^{i(E_j-E_i)t/\hbar}\,,\end{aligned} \tag{100}$$

where $\tilde{P}_q(i) \equiv \left[1-\tilde{\beta}(1-q)\left(E_i-\tilde{U}_q\right)\right]^{q/(1-q)}\Big/\tilde{Z}_q$. We have used the complete set of eigenfunctions of the Hamiltonian operator $\hat{H}|i\rangle = E_i|i\rangle$ in deriving the second expression in Eq. (100). From this, we deduce three important relations:

1) $\tilde{\Phi}^{(q)}_{BA}(t)$ is real, which is proved by taking the complex conjugate of both sides of Eq. (100), interchanging the i, j summations and using the hermitian characters of the matrix elements.

2) $\tilde{\Phi}^{(q)}_{BA}(t) = \tilde{\Phi}^{(q)}_{AB}(-t)$ (time-reversal symmetry). This follows by interchanging the roles of the operators $\hat{A}$ and $\hat{B}$, interchanging the i and j summations, changing t to $-t$, and comparing the result with the above expression.

3) Defining

$$\tilde{\sigma}^{(q)}_{BA}(\omega) = \int_0^\infty dt\,\tilde{\Phi}^{(q)}_{BA}(t)e^{-i\omega t}\,, \tag{101}$$

we obtain the **Onsager relationships**

$$\mathrm{Re}\tilde{\sigma}_{AB}^{(q)}(\omega) = \mathrm{Re}\tilde{\sigma}_{BA}^{(q)}(-\omega),\ \mathrm{Im}\tilde{\sigma}_{AB}^{(q)}(\omega) = -\mathrm{Im}\tilde{\sigma}_{BA}^{(q)}(-\omega)\ . \tag{102}$$

This follows from the result 2) above.

4) Defining the normalized q-mean value over the anticommutator combination of the operators which represents the fluctuation

$$\begin{aligned}\tilde{\Psi}_{BA}^{(q)}(t) &= \frac{1}{2}\ \mathrm{Tr}\left[\hat{\tilde{P}}\left(\hat{H};q,\tilde{\beta}\right)\left\{\hat{A}\hat{B}(t)+\hat{B}(t)\hat{A}\right\}\right] \\ &= \sum_{i,j}\left(\frac{\tilde{P}_q(i)-\tilde{P}_q(j)}{2}\right)\left\langle i|\hat{A}|j\right\rangle\left\langle j|\hat{B}|i\right\rangle e^{i(E_j-E_i)t/\hbar}\ . \end{aligned}\tag{103}$$

We may now derive **the dynamical fluctuation-dissipation theorem** by formally writing $e^{i(E_j-E_i)t/\hbar} = e^{iE_jt/\hbar}e^{-iE_it'/\hbar}$ and setting $t = t'$ at the end of the calculation. We then obtain a relationship between the two functions defined in Eqs. (100) and (103)

$$\tilde{\Psi}_{BA}^{(q)}(t,t') = \frac{\tilde{a}_q(-\partial_{t'})+\tilde{a}_q(\partial_t)}{2\left[\tilde{a}_q\left(-\partial_{t'}\right)-\tilde{a}_q\left(\partial_t\right)\right]}\ i\hbar\left(\partial_{t'}+\partial_t\right)\tilde{\Phi}_{BA}^{(q)}(t,t')\ , \tag{104}$$

where $\tilde{a}_q(\partial_t) = \left[1-\tilde{\beta}(1-q)\left(-i\hbar\partial_t-\tilde{U}_q\right)\right]^{q/(1-q)}$. When we take the limit $q \to 1$, this reduces to the result of Kubo [24] for extensive systems with the same notations

$$\Psi_{BA}(t) = E_\beta(-i\partial_t)\Phi_{BA}(t)\ ,\ E_\beta(x) \equiv \frac{\hbar x}{2}\coth\,\frac{\beta\hbar x}{2}\ . \tag{105}$$

5) Finally, we find that the **Kramers-Kronig relation** holds for the real and imaginary parts of the susceptibility defined by Eq. (100):

$$\begin{aligned}\tilde{\chi}_{BA}^{(q)}(\omega) &= \mathrm{Re}\tilde{\chi}_{BA}^{(q)}(\omega)+i\mathrm{Im}\tilde{\chi}_{BA}^{(q)}(\omega)\ , \\ \mathrm{Re}\tilde{\chi}_{BA}^{(q)}(\omega) &= P.V.\sum_{i,j}\frac{\left[\tilde{P}_q(i)-\tilde{P}_q(j)\right]\left\langle i|\hat{A}|j\right\rangle\left\langle j|\hat{B}|i\right\rangle}{\hbar\omega+E_i-E_j}\ , \\ \mathrm{Im}\tilde{\chi}_{BA}^{(q)}(\omega) &= \pi\sum_{i,j}\left[\tilde{P}_q(i)-\tilde{P}_q(j)\right]\left\langle i|\hat{A}|j\right\rangle\left\langle j|\hat{B}|i\right\rangle\delta\left(\hbar\omega+E_i-E_j\right)\ . \end{aligned}\tag{106}$$

Here, *P.V.* stands for the principal value. We thus note that all the relations obtained in the statistical mechanics of extensive systems hold in the normalized q-mean value formalism of the Tsallis ensemble theory, except that the fluctation-dissipation theorem appears in a different form.

The static limit of the dynamical framework is subtle even for the extensive case and this continues to be so here also. From Eq. (98), we have $\tilde{\phi}_{BA}^{(q)}(t) =$

$-(1/i\hbar)$ Tr $\left\{[\hat{A}, \hat{\tilde{P}}(\hat{H}; q, \tilde{\beta})]\tilde{B}(t)\right\}$. In view of Eq. (83) in Theorem B, after some manipulation, this takes the form

$$\tilde{\phi}^{(q)}_{BA}(t) = -q\mathrm{Tr}\Big\{\hat{\tilde{P}}\left(\hat{H}; q, \tilde{\beta}\right)\int_0^{\tilde{\beta}} d\tilde{u}\left[\hat{\tilde{P}}\left(\hat{H}; q, \tilde{u}\right)\right]^{-1} \times \hat{\tilde{A}}(\tilde{u})\hat{\tilde{P}}\left(\hat{H}; q, \tilde{u}\right)\dot{\hat{B}}(t)\Big\}\ . \tag{107}$$

The overdot on the operator $\hat{B}$ denotes its time derivative. This is another version of the linear response function which reduces to the corresponding Kubo expression in the limit $q \to 1$.

We compare the properties of the extensive and nonextensive theories in Table 7.

Table 7. Similar properties of the entropies

Extensive – von Neumann	**Nonextensive – Tsallis**
Positivity: $S_1 \geq 0$, with the equality for pure states	Positivity: $S_q \geq 0$, with the equality for pure states
The microcanonical ensemble has equiprobbility.	The microcanonical esemble has equiprobbility.
S_1 is concave, **implying the maxent principle.**	S_q is concave for $q > 0$ and convex for $q < 0$, **implying the maxent principle for** $q > 0$ **and the minimum entropy principle for** $q < 0$.
For canonical ensemble, $\hat{\rho}_1 = \exp\left(-\beta\hat{H}\right)/Z_1$, where $Z_1 = \mathrm{Tr}\exp(-\beta\hat{H})$ is the partition function.	For canonical ensemble, $\hat{\tilde{\rho}} = e_q\left(-\tilde{\beta}\left(\hat{H} - \tilde{U}_q\right)\right)/\tilde{Z}_q$, where $\tilde{Z}_q = \mathrm{Tr}e_q\left(-\tilde{\beta}\left(\hat{H} - \tilde{U}_q\right)\right)$ with $e_q(\hat{A}) \equiv \left(1 + (1-q)\hat{A}\right)^{1/(1-q)}$ is the q-partition function. $\tilde{\beta} = \beta/c_q$ and $c_q = \mathrm{Tr}\hat{\tilde{\rho}}^q = \left(\tilde{Z}_q\right)^{1-q}$. Note that there is a natural cutoff in energy when $q < 1$.

The Legendre transform structure of the free energy defined by $F_1 = -\beta^{-1} \ln Z_1 = U_1 - TS_1$, $U_1 = \frac{\partial}{\partial \beta}\left(\beta F_1\right)$, $S_1 = -\frac{\partial F_1}{\partial T}$, and $\frac{\partial S_1}{\partial U_1} = \beta$.	The Legendre transform structure of the q-free energy defined by $\tilde{F}_q = -\beta^{-1} \ln_q Z_q = \tilde{U}_q - \beta^{-1} S_q$, where $Z_q^{1-q} \equiv \tilde{Z}_q^{1-q} - \beta(1-q)\tilde{U}_q$. $\tilde{U}_q = \frac{\partial}{\partial \beta}\left(\beta \tilde{F}_q\right)$, $S_q = -\frac{\partial \tilde{F}_q}{\partial T}$, and $\frac{\partial S_q}{\partial \tilde{U}_q} = \beta$.
Causality (Kramers-Kronig), time-reversal symmetry, and Onsager reciprocity are obeyed by the Kubo dynamic linear response function.	Causality (Kramers-Kronig), time-reversal symmetry, and Onsager reciprocity are obeyed by the q-Kubo dynamic linear response function [11].

We now turn our attention to the discussion of Green function theory and its implications in the nonextensive context. The Green function method is perhaps one of the most powerful field theoretical techniques used to study interactions and correlations in many particle systems with its constituent particles obeying either Bose or Fermi statistics. In the Gibbsian ensemble theory of extensive systems, this method acquired a new look because it could accommodate finite temperature-dependence equally well as its time-dependence. Given these, it is natural to inquire if one can develop a corresponding Green function theory when the Tsallis ensemble is employed for describing nonextensive systems. This would then complete the program of describing nonextensive systems in much the same manner as extensive ones. We adopt the second-quantized creation and annihilation operators of the particle in the Heisenberg representation as in the book of Kadanoff and Baym (KB) [26] to describe a many-particle system whose Hamiltonian and number operators are $\hat{H}$ and $\hat{N}$, respectively. In this way, we describe the nonextensive bosonic and fermionic systems in equilibrium at arbitrary temperatures by maximizing the Tsallis entropy subject to the constraints of the normalized q-mean values of $\hat{H}$ and $\hat{N}$ [27]. This presentation is different from the one given in [25]where the normalized q-mean values were not used. Using the standard method of Lagrange multipliers, we have the result given in Eq. (56) for the density matrix, which is expressed in a slightly different notation by introducing the following quantities:

$$\begin{aligned} \hat{X}_q\left(\lambda_1,\lambda_2\right) &= \hat{Y}_q\left(\lambda_1,\lambda_2\right) - a_q\left(\lambda_1,\lambda_2\right) \ , \\ \hat{Y}_q\left(\lambda_1,\lambda_2\right) &= \left(\lambda_1 \hat{H} + \lambda_2 \tilde{N}\right) / c_q\left(\lambda_1,\lambda_2\right), \\ a_q\left(\lambda_1,\lambda_2\right) &= \left(\lambda_1 \hat{U}_q + \lambda_2 \tilde{N}_q\right) / c_q\left(\lambda_1,\lambda_2\right) \ . \end{aligned} \tag{108}$$

We now define the tow-point q-Green function (we use units with $\hbar = 1$):

$$G^{(q)}\left(1, 1'; \beta_q, \mu_q\right) = i\left\langle \mathrm{T}\left(\Psi(1)\Psi^{\dagger}(1')\right)\right\rangle_q$$
$$\equiv -i\,\frac{\mathrm{Tr}\left\{\left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2)\right)\right]^q \mathrm{T}\left(\Psi(1)\Psi^{\dagger}(1')\right)\right\}}{\mathrm{Tr}\left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2)\right)\right]^q}\ . \tag{109}$$

Here, 1 refers to a space-time coordinate of a particle situated at position $\vec{r}_1$ at time t_1, and T is the Wick time-ordering symbol

$$\begin{aligned}\mathrm{T}\left(\Psi(1)\Psi^{\dagger}(1')\right) &= \Psi(1)\Psi^{\dagger}(1') \qquad \text{for} \quad t_1 > t_{1'}\ ,\\ &= \pm\Psi^{\dagger}(1')\Psi(1) \quad \text{for} \quad t_1 < t_{1'}\ .\end{aligned} \tag{110}$$

The creation $\Psi^{\dagger}(1')$ and annihilation $\Psi(1')$ operators obey the equal time canonical commutation relations (CCR):

$$\begin{aligned}&\Psi(\vec{r}, t)\Psi^{\dagger}(\vec{r}', t) \pm \Psi^{\dagger}(\vec{r}', t)\Psi(\vec{r}, t) = \delta^{(D)}(\vec{r} - \vec{r}')\ ,\\ &\Psi(\vec{r}, t)\Psi(\vec{r}', t) \pm \Psi(\vec{r}', t)\Psi(\vec{r}, t) = 0 \text{ and h.c.}\ .\end{aligned} \tag{111}$$

In the above and in subsequent analysis, the upper sign refers to bosons and the lower to fermions. The definitions of the n-point q-Green functions follow in the same fashion. It should be noted that the conventional grand canonical ensemble results given in KB are obtained when we take the limit $q \to 1$ in all the above expressions. We now use the following integral relationship, which is **valid only for** $q < 1$:

$$\begin{aligned}\left[e_q\left(-\hat{X}_q(\lambda_1, \lambda_2)\right)\right]^q &= \int_{-\infty}^{+\infty} du\tilde{K}_q(u; a_q)e^{-(1-q)(1+iu)\hat{Y}_q(\lambda_1, \lambda_2)}\ ,\\ \tilde{K}_q(u; a_q) &\equiv \frac{\Gamma\left(\frac{1}{1-q}\right)e^{(1+iu)[1+(1-q)a_q(\lambda_1, \lambda_2)]}}{2\pi\ (1+iu)^{1/(1-q)}}\ ,\end{aligned} \tag{112}$$

which is obtained from the formula

$$\frac{y^{z-1}}{\Gamma(z)} = \frac{1}{2\pi}\int_{-\infty}^{-\infty} du\ \frac{e^{(1+iu)y}}{(1+iu)^z}$$
$$(y > 0 \text{ and } \mathrm{Re}z > 0)\ .$$

Then the q-Green function is formally expressed as a parametric integral over a conventional grand canonical Green function, but with the **same** Lagrange parameters λ_1 and λ_2 as in Eq. (56). This point will be important in later discussions. The other quantities are also formally expressed similarly in terms of their $q \to 1$ counterparts. For $q > 1$, a contour deformation technique is employed to rewrite Eq. (112) in a more suitable form [25]. Thus, we have

$$\begin{aligned}
&\tilde{Z}_q\left(\lambda_1,\lambda_2\right)=\int_{-\infty}^{+\infty}du\tilde{K}_q(u;a_q)Z_1\left(\frac{(1-q)(1+iu)}{c_q\left(\lambda_1,\lambda_2\right)}\ (\lambda_1;\lambda_2)\right),\\
&Z_1\left(\lambda_1,\lambda_2\right)=\mathrm{Tr}e^{-\hat{Y}_q(\lambda_1,\lambda_2)}\ ,\qquad\qquad (113)\\
&a_q(\lambda_1,\lambda_2)=\frac{1}{\tilde{Z}_q}\int_{-\infty}^{+\infty}du\tilde{K}_q(u;a_q)\mathrm{Tr}\left\{e^{-(1-q)(1+iu)\hat{Y}_q(\lambda_1,\lambda_2)}\hat{Y}_q(\lambda_1,\lambda_2)\right\}.
\end{aligned}$$

Here, we have introduced the following abbreviation for the two arguments:

$$\begin{aligned}
&\frac{(1-q)(1+iu)}{c_q\left(\lambda_1,\lambda_2\right)}\ (\lambda_1;\lambda_2)\\
&\qquad\equiv\left(\frac{(1-q)(1+iu)}{c_q\left(\lambda_1,\lambda_2\right)}\ \lambda_1\ ,\frac{(1-q)(1+iu)}{c_q\left(\lambda_1,\lambda_2\right)}\ \lambda_2\right)\ . \qquad (113a)
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
G^{(q)}(1,1';\lambda_1,\lambda_2)&=\frac{1}{\tilde{Z}_q}\int_{-\infty}^{+\infty}du\tilde{K}_q(u;a_q)Z_1\left(\frac{(1-q)(1+iu)}{c_q(\lambda_1,\lambda_2)}(\lambda_1;\lambda_2)\right)\\
&\qquad\times G^{(1)}\left(1,1';\frac{(1-q)(1+iu)}{c_q(\lambda_1,\lambda_2)}(\lambda_1;\lambda_2)\right)\ ,\\
&\qquad\qquad\qquad\qquad\qquad\qquad (114)\\
G^{(1)}(1,1';\lambda_1,\lambda_2)&=-i\frac{\mathrm{Tr}\left[e^{-\hat{Y}_q(\lambda_1,\lambda_2)}\mathrm{T}\left(\Psi(1)\Psi^\dagger(1')\right)\right]}{\mathrm{Tr}e^{-\hat{Y}_q(\lambda_1,\lambda_2)}}\ .
\end{aligned}$$

For $q>1$, we use a general contour integral representation of the gamma function. One may consult [25] for details, as well as for general representations of the quantities given here. We now introduce the correlation functions which are related to the full Green functions:

$$\begin{aligned}
&G_>^{(q)}\left(1,1';\lambda_1,\lambda_2\right)=-i\left\langle\Psi(1)\Psi^\dagger(1')\right\rangle_q\ ,\\
&\qquad\qquad\qquad\qquad\qquad\qquad (115a)\\
&G_<^{(q)}\left(1,1';\lambda_1,\lambda_2\right)=\mp i\left\langle\Psi^\dagger(1')\Psi(1)\right\rangle_q\ ,\\
&G^{(q)}\left(1,1';\lambda_1,\lambda_2\right)=\begin{cases}G_>^{(q)}\left(1,1';\lambda_1,\lambda_2\right) & \text{for}\quad t_1>t_{1'}\ ,\\ G_<^{(q)}\left(1,1';\lambda_1,\lambda_2\right) & \text{for}\quad t_1<t_{1'}\ .\end{cases} \qquad (115b)
\end{aligned}$$

Using Eq. (114), each of these correlation functions can be expressed in terms of the usual ($q\to1$) Green functions given in KB. Formally, those usual Green functions have "periodicity"property in a complex time-domain and so can be expressed in terms of a spectral weight function, $A(1,1')$, which reflects only the properties of the system Hamiltonian. Furthermore, the basic permutation symmetry of the system is accounted for by the introduction of the occupation number of a mode of energy ω in the grand canonical ensemble: $f(\omega;\lambda_1,\lambda_2)\equiv$

$[\exp(\lambda_1\omega + \lambda_2) \mp 1]^{-1}$. We thus rewrite the expressions for Eqs. (115a) and (115b) in terms of the frequency variable which is the Fourier transform variable with respect to time differences, in the following way:

$$\begin{aligned} iG^{(q)}_{>}\left(\vec{r}_1, \vec{r}_{1'}; \omega; \lambda_1, \lambda_2\right) = & \frac{1}{\tilde{Z}_q}\int_{-\infty}^{+\infty} du\tilde{K}_q(u, a_q) \\ & \times Z_1\left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)}(\lambda_1; \lambda_2)\right) \\ & \times\left[1 \pm f\left(\omega;\ \frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)}(\lambda_1; \lambda_2)\right)\right] \\ & \times A(\vec{r}_1, \vec{r}_{1'}; \omega)\ , \end{aligned} \tag{116a}$$

$$\begin{aligned} iG^{(q)}_{<}\left(\vec{r}_1, \vec{r}_{1'}; \omega; \lambda_1, \lambda_2\right) = & \frac{1}{\tilde{Z}_q}\int_{-\infty}^{+\infty} du\tilde{K}_q(u, a_q) \\ & \times Z_1\left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)}(\lambda_1; \lambda_2)\right) \\ & \times\left[\pm f\left(\omega;\ \frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)}(\lambda_1; \lambda_2)\right)\right] \\ & \times A(\vec{r}_1, \vec{r}_{1'}; \omega)\ . \end{aligned} \tag{116b}$$

From this, we see that the spectral weight function is just

$$i\left\{G^{(q)}_{>}\left(1, 1'; \lambda_1, \lambda_2\right) - G^{(q)}_{<}\left(1, 1'; \lambda_1, \lambda_2\right)\right\} = A(1, 1') \tag{117}$$

in view of the CCR given in Eq. (111). An important sum rule regarding the spectral weight reflecting the equal-time commutation relation is then given as follows:

$$\begin{aligned} & i\int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\left\{G^{(q)}_{>}\left(\vec{r}_1, \vec{r}_{1'}; \omega; \lambda_1, \lambda_2\right) - G^{(q)}_{<}\left(\vec{r}_1, \vec{r}_{1'}; \omega; \lambda_1, \lambda_2\right)\right\} \\ & = \int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\ A\left(\vec{r}_1, \vec{r}_{1'}; \omega\right) = \delta^{(D)}\left(\vec{r}_1 - \vec{r}_{1'}\right)\ . \end{aligned} \tag{118}$$

It is worth pointing out that the physical interpretation of the correlation functions for any q as transition probabilities for adding and removing a particle holds as for the case $q \to 1$ described by KB. Also, many important mean values of physical quantities of an interacting system may be expressed in terms of the spectral weight function and the occupation probabilities. We give here a few of these which will be useful for later discussions. Following KB, the q-mean energy of an interacting uniform system of particles in D-dimensions contained in a volume Ω, having the usual two-body interactions, is found to be

$$\frac{\tilde{U}_q}{\Omega} = \frac{1}{\tilde{Z}_q}\int_{-\infty}^{+\infty} du\tilde{K}_q(u; a_q) Z_1\left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)}\ (\lambda_1; \lambda_2)\right)$$

$$\times \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \int \frac{d^D \overrightarrow{p}}{(2\pi)^D} \left(\frac{\omega + \overrightarrow{p}^2/2m}{2} \right)$$
$$\times f \left(\omega; \frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right) A(\overrightarrow{p}; \omega) \ , \tag{119}$$

where $A(\overrightarrow{p}; \omega)$ is the spatial Fourier transform of $A\left(\overrightarrow{r}_1, \overrightarrow{r}_{1'}; \omega\right)$. Similarly, the single-particle momentum distribution function is

$$\tilde{N}_q(\overrightarrow{p}) = \frac{1}{\tilde{Z}_q} \int_{-\infty}^{+\infty} du \tilde{K}_q(u; a_q) Z_1 \left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right)$$
$$\times \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \, f \left(\omega; \frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right) A(\overrightarrow{p}; \omega) \ , \tag{120}$$

and the single-particle frequency distribution function is

$$\tilde{N}_q(\omega) = \frac{\Omega}{\tilde{Z}_q} \int_{-\infty}^{+\infty} du \tilde{K}_q(u; a_q) Z_1 \left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right)$$
$$\times f \left(\omega; \frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right) \int \frac{d^D \overrightarrow{p}}{(2\pi)^D} \, A(\overrightarrow{p}; \omega) \ . \tag{121}$$

The chemical potential is determined by the expression for the q-mean value of the total number operator:

$$\tilde{N}_q = \frac{\Omega}{\tilde{Z}_q} \int_{-\infty}^{+\infty} du \tilde{K}_q(u; a_q) Z_1 \left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right)$$
$$\times \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \int \frac{d^D \overrightarrow{p}}{(2\pi)^D}$$
$$\times f \left(\omega; \frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \, (\lambda_1; \lambda_2) \right) A(\overrightarrow{p}; \omega) \ . \tag{122}$$

Experimentally, one can measure the momentum distribution function in Eq. (120) either by Compton scattering or by positron annihilation.

For the sake of completeness, it is also important to examine the fluctuations given in Eq. (61) in terms of the integral representation:

$$\frac{\partial^2 \tilde{Z}_q}{\partial \lambda_1^2} = q \mathrm{Tr} \left\{ \left[e_q \left(-\hat{X}_q(\lambda_1, \lambda_2) \right) \right]^{2q-1} \frac{\left(\hat{H} - \tilde{U}_q \right)^2}{c_q^2(\lambda_1, \lambda_2)} \right\} \ ,$$

$$\frac{\partial^2 \tilde{Z}_q}{\partial \lambda_2^2} = q \mathrm{Tr} \left\{ \left[e_q \left(-\hat{X}_q(\lambda_1, \lambda_2) \right) \right]^{2q-1} \frac{\left(\hat{N} - \tilde{N}_q \right)^2}{c_q^2(\lambda_1, \lambda_2)} \right\} \ ,$$

$$\frac{\partial^2 \tilde{Z}_1}{\partial \lambda_1 \partial \lambda_2} =$$

$$q\mathrm{Tr}\left\{\left[e_q\left(-\hat{X}_q(\lambda_1,\lambda_2)\right)\right]^{2q-1}\frac{\left(\hat{H}-\hat{U}_q\right)\left(\hat{N}-\tilde{N}_q\right)}{c_q^2(\lambda_1,\lambda_2)}\right\}\ ,$$

$$\left[e_q\left(-\hat{X}(\lambda_1,\lambda_2)\right)\right]^{2q-1}=\int_{-\infty}^{+\infty}\frac{du}{2\pi}\,\frac{\Gamma\left(\frac{q}{1-q}\right)e^{(1+iu)(1+(1-q)a_q)}}{(1+iu)^{q/(1-q)}}$$
$$\times e^{-\dfrac{(1+iu)(1-q)}{c_q(\lambda_1,\lambda_2)}\left(\lambda_1\hat{H}+\lambda_2\hat{N}\right)}\ . \tag{123}$$

We will now illustrate some consequences of this theory by examining the "free particle" systems for which $q=1$ results are known.

Ideal free gas:
Recall that the Gibbsian "fugacity" is defined by

$$z=e^{\lambda_2}\ . \tag{124}$$

For the ideal gas, the energy eigenvalues are given by $\varepsilon(\vec{p})=\vec{p}^{\,2}/2m$, m is the mass and $\vec{p}$ is the linear momentum of the particle. Therefore the spectral weight function is given by

$$A(\vec{p};\omega)=2\pi\ \delta\left(\omega-\vec{p}^{\,2}/2m\right)\ . \tag{125}$$

We consider three dimensional systems so that $D=3$ in the formulas derived above. Then the grand partition functions for the fermi (F), bose (B), and maxwell (M) systems are given in any standard text book (e.g., [28]), and we will recast them here in our notation:

$$\ln Z_1^F(\lambda_1,\lambda_2)\ =\Omega\left(\frac{m}{2\pi\lambda_1}\right)^{3/2}\sum_{l=1}^{\infty}\frac{(-1)^{l+1}e^{-\lambda_2 l}}{l^{5/2}}\ , \tag{126}$$

$$\ln Z_1^B(\lambda_1,\lambda_2)\ =\Omega\left(\frac{m}{2\pi\lambda_1}\right)^{3/2}\sum_{l=1}^{\infty}\frac{e^{-\lambda_2 l}}{l^{5/2}}-\ln\left(1-e^{-\lambda_2}\right)\ , \tag{127}$$

$$\ln Z_1^M(\lambda_1,\lambda_2)=\Omega\left(\frac{m}{2\pi\lambda_1}\right)^{3/2}e^{-\lambda_2}\ . \tag{128}$$

Note that we need to evaluate only the formal structure of the partition functions in the limit $q\to 1$ because the Lagrange multipliers of the original Tsallis maxent problem are the same as those in the above expressions. We now evaluate all the relevant quantities using these expressions and finally examine the results. For all these noninteracting cases, q can only be less than 1 as otherwise the expressions diverge.

Case (a): Maxwellons:
We first consider the case of classical ideal gas composed of N particles confined in the volume Ω, Using Eq. (128) in Eq. (113), with corresponding modifications,

we have

$$
\begin{aligned}
\tilde{Z}_q^M &= \int_{-\infty}^{+\infty} du \tilde{K}_q\left(u;a_q\right) \frac{1}{(1-q)(1+iu)} \\
&\quad \times \exp\left\{ \Omega \left(\frac{mc_q(\lambda_1,\lambda_2)}{2\pi(1-q)(1+iu)\lambda_1} \right)^{3/2} e^{-(1-q)(1+iu)\lambda_2/c_q(\lambda_1,\lambda_2)} \right\} \\
&= \sum_{N=0}^{\infty} \frac{\left[\Omega \left(\frac{mc_q(\lambda_1,\lambda_2)}{2\pi(1-q)\lambda_1} \right)^{3/2} \right]^N \Gamma\left(\frac{2-q}{1-q}\right)}{N! \Gamma\left(\frac{2-q}{1-q} + \frac{3}{2}\ N\right)} \\
&\quad \times \left[1 + (1-q)\left(a_q - N\lambda_2/c_q(\lambda_1,\lambda_2)\right)\right]^{\frac{1}{1-q}+\frac{3}{2}\ N} \\
&\equiv \sum_{N=0}^{\infty}\ \tilde{\tilde{Z}}_q^M\left(N;\lambda_1,\lambda_2\right)\ .
\end{aligned}
\tag{129a}
$$

From Eqs. (59) and (112), we also have an equivalent expression for the partition function, which reads

$$
\begin{aligned}
\tilde{Z}_q^M &= \int_{-\infty}^{+\infty} du \tilde{K}_q\left(u;a_q\right) \\
&\quad \times \exp\left\{ \Omega \left(\frac{mc_q(\lambda_1,\lambda_2)}{2\pi(1-q)(1+iu)\lambda_1} \right)^{3/2} e^{-(1-q)(1+iu)\lambda_2/c_q(\lambda_1,\lambda_2)} \right\} \\
&= \sum_{N=0}^{\infty} \frac{\left[\Omega \left(\frac{mc_q(\lambda_1,\lambda_2)}{2\pi(1-q)\lambda_1} \right)^{3/2} \right]^N \Gamma\left(\frac{1}{1-q}\right)}{N! \Gamma\left(\frac{1}{1-q} + \frac{3}{2}\ N\right)} \\
&\quad \times \left[1 + (1-q)\left(a_q - N\lambda_2/c_q(\lambda_1,\lambda_2)\right)\right]^{\frac{q}{1-q}+\frac{3}{2}\ N} \\
&\equiv \sum_{N=0}^{\infty}\ \tilde{\tilde{Z}}_q^M\left(N;\lambda_1,\lambda_2\right)\ .
\end{aligned}
\tag{129b}
$$

We now calculate the q-mean value of the energy in Eq. (119) as follows:

$$
\begin{aligned}
\frac{\tilde{U}_q^M}{\Omega} &= \frac{1}{\tilde{Z}_q^M}\ \int_{-\infty}^{+\infty} du \tilde{K}_q(u;a_q) Z_1^M \left(\frac{(1-q)(1+iu)}{c_q(\lambda_1,\lambda_2)}\ (\lambda_1;\lambda_2) \right) \\
&\quad \times \int\ \frac{d^3\overrightarrow{p}}{(2\pi)^3}\ e^{-\frac{(1-q)(1+iu)}{c_q(\lambda_1,\lambda_2)}(\lambda_1 \overrightarrow{p}^2/2m+\lambda_2)}\ \frac{\overrightarrow{p}^2}{2m} \\
&= \frac{1}{\Omega\tilde{Z}_q^M}\ \frac{3c_q(\lambda_1,\lambda_2)}{2\lambda_1} \sum_{N=0}^{\infty} N\ \frac{\left[\Omega \left(\frac{mc_q(\lambda_1,\lambda_2)}{2\pi(1-q)\lambda_1} \right)^{3/2} \right]^N \Gamma\left(\frac{2-q}{1-q}\right)}{N! \left(\frac{2-q}{1-q} + \frac{3}{2}\ N\right)} \\
&\quad \times \left[1 + (1-q)\left(a_q - N\lambda_2/c_q(\lambda_1,\lambda_2)\right)\right]^{\frac{1}{1-q}+\frac{3}{2}\ N}\ .
\end{aligned}
\tag{130}
$$

Similarly, the q-mean value of the total number operator from Eq. (122) is

$$\begin{aligned} \frac{\tilde{N}_q^M}{\Omega} &= \frac{1}{\tilde{Z}_q^M} \int_{-\infty}^{+\infty} du \tilde{K}_q(u; a_q) Z_1^M \left(\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} \; (\lambda_1; \lambda_2) \right) \\ &\quad \times \int \frac{d^3 \overrightarrow{p}}{(2\pi)^3} \; e^{-\frac{(1-q)(1+iu)}{c_q(\lambda_1, \lambda_2)} (\lambda_1 \overrightarrow{p}^2 / 2m + \lambda_2)} \\ &= \frac{1}{\Omega \tilde{Z}_q^M} \sum_{N=0}^{\infty} N \, \frac{\left[\Omega \left(\frac{m c_q(\lambda_1, \lambda_2)}{2\pi (1-q) \lambda_1} \right)^{3/2} \right]^N \Gamma\left(\frac{1}{1-q} \right)}{N! \left(\frac{1}{1-q} + \frac{3}{2} \; N \right)} \\ &\quad \times \left[1 + (1-q) \left(a_q - N\lambda_2 / c_q(\lambda_1, \lambda_2) \right) \right]^{\frac{q}{1-q} + \frac{3}{2} \; N} \; . \end{aligned} \tag{131}$$

Eqs. (130) and (131) have to be solved in terms of the Lagrange multipliers for given constraints of the q-mean values of the energy and number operators.

Corresponding expressions for the canonical ensemble were obtained by Abe [29], which are formally reproduced by picking the coefficient of N in the above series expressions and setting λ_2 equal to zero. His results were based on the normalized q-mean value constraints, and as such while he did not have infinities in the thermodynamic limit that plagued the earlier non-normalized q-mean value theory of the classical ideal gas [30], but he obtained the negative q-specific heat for $0 < q < 1$. In the grand canonical expressions given above, it should be noted that the infinite series in N has a natural cut off given by

$$1 + (1-q) \left(a_q - N_c \, \frac{\lambda_2}{c_q(\lambda_1, \lambda_2)} \right) = 0 \; . \tag{132}$$

so that the series is terminated after $N = N_c$.

We note that, from Eqs. (129a) and (130), we get an identity

$$\begin{aligned} 0 &= \sum_{N=0}^{\infty} \left(\tilde{U}_q^M - \frac{3 c_q(\lambda_1, \lambda_2)}{2\lambda_1} \; N \right) \frac{\left[\Omega \left(\frac{m c_q(\lambda_1, \lambda_2)}{2\pi (1-q) \lambda_1} \right)^{3/2} \right]^N \Gamma\left(\frac{2-q}{1-q} \right)}{N! \Gamma\left(\frac{2-q}{1-q} + \frac{3}{2} \; N \right)} \\ &\quad \times \left[1 + (1-q) \left(a_q - N\lambda_2 / c_q(\lambda_1, \lambda_2) \right) \right]^{\frac{1}{1-q} + \frac{3}{2} \; N} \; . \end{aligned} \tag{133}$$

Similarly, from Eqs. (129b) and (131), we get another identity

$$\begin{aligned} 0 &= \sum_{N=0}^{\infty} \left(\tilde{N}_q^M - N \right) \frac{\left[\Omega \left(\frac{m c_q(\lambda_1, \lambda_2)}{2\pi (1-q) \lambda_1} \right)^{3/2} \right]^N \Gamma\left(\frac{1}{1-q} \right)}{N! \Gamma\left(\frac{1}{1-q} + \frac{3}{2} \; N \right)} \\ &\quad \times \left[1 + (1-q) \left(a_q - N\lambda_2 / c_q(\lambda_1, \lambda_2) \right) \right]^{\frac{q}{1-q} + \frac{3}{2} \; N} \; . \end{aligned} \tag{134}$$

These identities are useful in calculating the derivatives of the partition function in Eq. (129a) with respect to the Lagrange multipliers as in Eqs. (60) and (61) to determine “fluctuations” in this formalism. The first derivatives vanish upon using Eqs. (133) and (134). The second derivatives are nonzero and are simplified further after using the expressions in Eq. (123). We will not belabor this point here, as it requires more numerical analysis.

Case (b) Bosons:
We focus on the region of Bose condensation where the chemical potential is set equal to zero, thereby allowing some simplification. Then, the calculations are almost parallel to the classical ideal gas case considered above. We have

$$\tilde{Z}_1^B(\beta_q^c) \cong \exp\left[\Omega\left(\frac{m}{2\pi\lambda_1^c}\right)^{3/2}\zeta(5/2)\right] , \tag{135}$$

where $\zeta(s) = \sum_{l=1}^{\infty} l^{-s}$ is the zeta function of order s. Details of this type of calculation are given in [25,27] and we will not repeat them here. Further consequences of these require numerical analysis to include self-consistency and is being done currently.

5 Time-Dependence: Nonunitary Dynamics

The Lindblad equation is for the density matrix describing the time evolution of dissipative quantum systems under only the most general requirements of linearity, locality in time, and conservation of probability (traceclass property of the density matrix). For a recent description of this and the corresponding literature, see [33]. By its very nature, this time evolution is nonunitary and hence the corresponding manipulations are different and more complex, and such details may be found in [33]. The forward and backward time evolutions (compare Eq. (87)) are different in this formalism as well as the mean value of an arbtrary operator for the same reason (compare Eq. (90)). It replaces the nonunitary stochastic Fokker-Planck dynamics for the classical distribution function. By taking a Wigner-transform of the Lindblad equation for the density matrix, a Fokker-Planck-like equation for the quantum phase-space distribution function is derived. See Table 8 for its comparison with its unitary counterpart. The time-dependent Tsallis entropy functional is now time-dependent unlike in the unitary evolution case. This equation has only recently been employed to discuss real physical phenomena such as time-scale of quantum operation of nanometric devices [34] etc.. This is a separate subject deserving its own full discussion. We will be content here with just giving the salient features of its consequences as presented in Table 8.

Table 8. Comparison of von Neumann and Lindblad time evolutions

Unitary Evolution – von Neuman	**Nonunitary Evolution – Lindblad**
$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}\ \left[\hat{H}, \hat{\rho}\right]$, $\hat{\rho}$ is Hermitian. The initial condition is usually specified to suit the physical situation under study. For example, one may choose to maximize entropy (von Neumann or Tsallis) given some physical constraints such as energy.	$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}\ \left[\hat{H}, \hat{\rho}\right]$ $-\frac{1}{2\hbar}\sum_i A^{(i)}\left(\hat{Q}_i^{\dagger}\hat{Q}_i\hat{\rho} + \hat{\rho}\hat{Q}_i^{\dagger}\hat{Q}_i - 2\hat{Q}_i\hat{\rho}\hat{Q}_i^{\dagger}\right)$. The operators $\hat{Q}_i$ are arbitrary and represent the decay processes. $\hat{\rho}$ is Hermitian iff $A^{(i)}$ are real. The initial condition is usually specified to suit the physical situation under study.
$\hat{\rho}$ is traceclass: $\mathrm{Tr}\hat{\rho}(t) = \mathrm{Tr}\hat{\rho}(0) = 1$. $\hat{\rho}$ is a positive operator; its eigenvalues are positive.	$\hat{\rho}$ is traceclass: $\mathrm{Tr}\hat{\rho}(t) = \mathrm{Tr}\hat{\rho}(0) = 1$. $\hat{\rho}$ is a positive operator; its eigenvalues are positive, iff $A^{(i)}$ are positive.
von Neumann entropy $S_1 = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho})$ is independent of time and so there is no possibility of transition from pure to mixed states and vice versa.	von Neumann entropy $S_1 = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho})$ depends on time; $\frac{dS_1(t)}{dt} \geq 0$ provided $\hat{Q}_i^{\dagger} = \hat{Q}_i$. Thus there is possibility of transitionfrom pure to mixed states and vice versa.
Tsallis entropy $S_q = -\mathrm{Tr}(\hat{\rho}\mathrm{Ln}_q\hat{\rho})$ is independent of time and so there is no possibility of transition from pure to mixed states and vice versa.	Tsallis entropy $S_q = -\mathrm{Tr}(\hat{\rho}\mathrm{Ln}_q\hat{\rho})$ depends on time; provided $\frac{dS_q(t)}{dt} \geq 0$ for $q > 0$, and so there is possibility of transition from pure to mixed states and vice versa.
$i\hbar\ \frac{d}{dt}\mathrm{Tr}\left(\hat{O}(t)\hat{\rho}(t)\right) = \mathrm{Tr}\left(i\hbar\ \frac{\partial\hat{O}(t)}{\partial t}\ \hat{\rho}(t)\right)$ $-\ \mathrm{Tr}\left(\left[\hat{H}, \hat{O}(t)\right]\hat{\rho}(t)\right)$.	$i\hbar\ \frac{d}{dt}\mathrm{Tr}\left(\hat{O}(t)\hat{\rho}(t)\right) = \mathrm{Tr}\left(i\hbar\ \frac{\partial\hat{O}(t)}{\partial t}\ \hat{\rho}(t)\right)$ $-\mathrm{Tr}\left(\left[\hat{H}, \hat{O}(t)\right]\hat{\rho}(t)\right) - \frac{i}{2}\sum_i A^{(i)}$ $\times\mathrm{Tr}\left(\left\{\begin{array}{c}\hat{O}(t)\hat{Q}_i^{\dagger}\hat{Q}_i + \hat{Q}_i^{\dagger}\hat{Q}_i\hat{O}(t)\\ -2\hat{Q}_i^{\dagger}\hat{O}(t)\hat{Q}_i\end{array}\right\}\hat{\rho}(t)\right)$.
Action principle–variational approach [31]	Action principle–variational approach [32,33]

6 Concluding Remarks

Instead of a summary, we end this description with some thoughts on several questions that need to be asked in the context of the development of the formalism. They fall into two classes. The first class of questions concerns some experimental consequences of the formalism while the second class concerns some theoretical points. There are two new papers applying the Tsallis entropy that have attracted my attention which I will point out. Finally, there is an interesting question concerning the Boltzmann-Einstein relationship in the Tsallis entropy context, which will be answered.

Experimental consequences in the q-formalism parallel to those in ordinary equilibrium statistical mechanics such as thermal, mechanical, and particle equilibrium as in the standard textbook [35] need to be explored. This depends on the additive properties of the entropy and the normalized q-mean values defined in Section 2. Consider a closed system in thermal equilibrium. The value of the Tsallis entropy will depend on the q-energy, q-number, and external variables such as volume. Consider now two subsystems, for the present with the same q-index, in contact with each other and in thermal equilibrium. Thermal equilibrium means that the entropy of the total system must be maximum with respect to small transfers of q-energy from one to the other. Thus we have

$$\begin{aligned}\delta S_q(A\cup B) = {} & \frac{\delta S_q(A)}{\delta \tilde{U}_q(A)}\left[1+(1-q)S_q(B)\right]\delta\tilde{U}_q(A) \\ & + \frac{\delta S_q(B)}{\delta \tilde{U}_q(B)}\left[1+(1-q)S_q(B)\right]\delta\tilde{U}_q(B) = 0 \ .\end{aligned}$$

We know also that the total q-energy is held constant, because the system is thermally closed (microcanonical ensemble or the fixed total energy constraint) $\delta\tilde{U}_q(A)+\delta\tilde{U}_q(B)=0$. Therefore, we obtain

$$\begin{aligned}\delta S_q(A\cup B) \propto {} & \left\{ \frac{\delta S_q(A)}{\delta \tilde{U}_q(A)}\left[1+(1-q)S_q(A)\right]^{-1} \right. \\ & \left. - \frac{\delta S_q(B)}{\delta \tilde{U}_q(B)}\left[1+(1-q)S_q(B)\right]^{-1} \right\} \delta\tilde{U}_q(A) = 0 \ .\end{aligned}$$

As $\delta\tilde{U}_q(A)$ is an arbitrary variation, we must have the following relation in thermal equilibrium:

$$\frac{\delta S_q(A)}{\delta \tilde{U}_q(A)}\left[1+(1-q)S_q(A)\right]^{-1} = \frac{\delta S_q(B)}{\delta \tilde{U}_q(B)}\left[1+(1-q)S_q(B)\right]^{-1} \ .$$

If we define a quantity β_q by

$$(\beta_q) = \frac{\delta S_q}{\delta \tilde{U}_q}\left[1+(1-q)S_q\right]^{-1} \equiv \frac{\delta S_q}{\delta \tilde{U}_q}\ (c_q)^{-1}$$

after using Eq. (63), then in thermal equilibrium, $\beta_q(A) = \beta_q(B)$. Should we interpret this as the inverse temperature of the q-system in analogy with the familiar $q = 1$ case? By following similar arguments as in [35], this time with the two systems occupying respective volumes,

$$\frac{\delta S_q(A)}{\delta \Omega(A)} \left[1 + (1-q) S_q(A)\right]^{-1} = \frac{\delta S_q(B)}{\delta \Omega(B)} \left[1 + (1-q) S_q(B)\right]^{-1}$$

is the condition for mechanical equilibrium. We may define q-pressure by

$$\beta_q \Pi_q = \frac{\delta S_q}{\delta \Omega} \left[1 + (1-q) S_q\right]^{-1} \equiv \frac{\delta S_q}{\delta \Omega} (c_q)^{-1} \, .$$

Similarly, for particle equlibrium, we obtain

$$\frac{\delta S_q(A)}{\delta \tilde{N}_q(A)} \left[1 + (1-q) S_q(A)\right]^{-1} = \frac{\delta S_q(B)}{\delta \tilde{N}_q(B)} \left[1 + (1-q) S_q(B)\right]^{-1}$$

and the definition of q-chemical potential

$$-\beta_q \mu_q = \frac{\delta S_q}{\delta \tilde{N}_q} \left[1 + (1-q) S_q\right]^{-1} \equiv \frac{\delta S_q}{\delta \tilde{N}_q} \, (c_q)^{-1} \, .$$

With these new identification of the Lagrange parameters, we may have different definitions of the q-specific heat and q-compressiblity, which are something that seemed to be called for in our discussion of fluctuations in the q-formalism in Section 3. Two other questions come to mind in this context; can we put two systems with two different q's in thermal contact? How about one with $q = 1$ and another with q not equal to unity? This latter question may give us a clue to measure q-temperature using conventional thermometer?!

On some theoretical points that are still worrisome to me, there is really only one basic cause for worry in all the three cases presented below:

(1) The non uniqueness in the definition of the Tsallis entropy of nonextensive systems: Several entropy functionals may be set up [36] such that the nonadditivity is characterized in a generic Tsallis form given in Table 2, with the only criterion that it goes to the von Neumann form for extensive systems in the limit $q \to 1$. For example, we may define another entropy functional, $S_q^{(a)} = S_q / c_q$, from the definition in Eq. (2). This leads to the pseudoadditivity relation for the entropy in Table 2 with the nonadditivity term appearing with $(q-1)$ instead of $(1-q)$ there. They all possess the Legendre transform structure as well. The question is whether one can have at least one other criterion to make a unique choice or, is it that the nature of the beast is such that this non uniqueness allows for many other forms for the probability distribution which are not either exponential or monomial form and are pertinent to some other hitherto undiscovered physical situation?

(2) There is a paper [37] claiming to prove the uniqueness of the Tsallis form similar to the well known one for the Shannon structure from general mathematical axioms. This proof is equally valid under certain conditions for other

definitions mentioned above [36]! This means that the uniqueness proof depends on the underlying axioms concerning the form of the non extensive character.

(3) Nonuniquensss in defining information theory based on the Tsallis entropy (Abe and A. K. R. [38]). In [36], there was a study reported where the authors suggest that perhaps information theory should be developed in a consistent fashion. The reason for the various definitions arises from the ambiguous interpretation of Eq. (2), either as $\langle \mathrm{Ln}_q \hat{\rho} \rangle_1$ which is based on the traditional definition of the mean value or as the un-normalized q-mean value, or as the normalized q-mean value, $\langle \ln_q \hat{\rho} \rangle_q$, provided that in the latter case we divide the unnormalized expression by c_q. The first one is equivalent to the second expression but the interpretation is different and crucial. All these go to Eq. (1) in the limit $q \to 1$! My point has been to use the quantum version of the Kullback-Leibler entropy a là Section 2 as a guide (see Appendix), in which case, as far as the entropy definition is concerned, we use the structure given in Eq. (2) which is the third form of the three mentioned above and employ the normalized q-mean value for calculating mean values which preserves the usual structure that the mean value of a constant is itself. The reason is that in the quantum version, the overlap of the wave functions appear, and the definition of the "conditional density matrix" does not seem to uniquely exist in contrast to the fact that "conditional probability" exists in the classical description. But we can construct "conditional entropy", which has the property that it should go to zero when the density matrix factorizes.

Much of these uncertain feelings may be traced to the sole guiding principle that we must get back to the von Neumann entropy in the limit $q \to 1$. This seems to be at the heart of the non uniqueness! There must be another principle which makes the definitions unambiguous! Our suggestion at present is the quantum density matrix approach! A discussion of these points is called for. At this juncture, part of a rubai'yi # 191 of the Ruba'iyat of Omar Khayyam (trans. Peter Avery and John Heath-Stubbs, Penguin Classics, 1979) comes to mind:

"I have meditated for seventy-two years night and day
To learn that nothing has been learned at all."

Among the recent applications, two studies have attracted my attention.

(1) Correlations in quantum many particle systems by Zeische *et al.* [39]. These authors suggest that Tsallis-type entropy defined in terms of the eigenvalues of one-particle density matrix will exhibit correlation effects in an interacting system. They call the Tsallis entropy as the q-order idempotency, which is zero when there is no correlation, and increases with increasing correlation. This therefore is a good test for approximation methods which claim to take account of correlations (basis sets, the Slater determinants, etc.) in computations of properties of many particle systems. The authors test their idea on the two-site Hubbard model and on the Bardeen-Cooper-Schrieffer model of superconductivity.

(2) Entropic lower bound for quantum scattering by Ion and Ion [40] based on Tsallis entropy. These authors define the angular and angular-momentum Tsallis entropies for any quantum scattering states. Since the two basis states are a Fourier series transform pair, they use the Riesz theorem instead of the Sobolev inequality (used in [41] for a Fourier integral transform pair of probability amplitudes) to derive an entropic lower bound! They then show that the experimental pion-nucleus entropies are well described by their optimal entropies.

Another point of interest is to display the nature of the famous Boltzmann formula for entropy as $S = k \ln W(N)$, where $W(N)$ is the degeneracy of the microstates. In our view for a many particle system, $W(N)$ is just the largest degeneracy allowed for occupying a given N-particle state with a given energy. Thus, the density matrix is dominated by the term with $n = N$ in the eigenstate expansion: $\hat{\rho} = \sum\limits_n W(n)p(n)|n\rangle\langle n|$. Since $\mathrm{Tr}\hat{\rho} = 1$, the dominating term saturates the trace and, therefore, we have $p(N)W(N) = 1$. Consequently, the Tsallis entropy given by Eq. (2) in this representation of the density matrix becomes $S_q = \sum_n W(n)p(n)\left[1 - p^{q-1}(n)\right]/(q-1) \cong \left[W^{1-q}(N) - 1\right]/(1-q) \equiv \ln_q W(N)\ \ (k = 1)$. Hence we get the Tsallis entropy to be $\ln_q W(N)$. Since $W(N)$ is an intrinsic property of the system under consideration, as it depends only on the symmetry properties of the system, it is the same one that is employed in examining the fermi or bose or maxwell system of particles. This expression when used in the von Neumann entropy with large number approximation yields the traditional statistical number distributions for the fermi, bose, and maxwell cases.

Note Added

We update here the manuscript originally composed a little more than a year ago with some new developments that we have been involved with in the intervening months by giving a brief account of these.

(1) A. K. Rajagopal and S. Abe, "Implications of Form Invariance to the Structure of Nonextensive Entropies", Phys. Rev. Lett. **83**, 1711 (1999). In this paper, the form invariance of the maxent principle and the metric structure in quantum density matrix suitably generalized to nonextensive situations is shown to determine a structure of possible nonextensive entropies.

(2) S. Abe and A. K. Rajagopal, "Quantum entanglement inferred by the principle of maximum nonadditive entropy", Phys. Rev. A **60**, 3461 (1999). By extending the work in A. K. Rajagopal, "Quantum entanglement and the maximum entropy states from the Jaynes principle", Phys. Rev. A **60**, 4338 (1999) to the nonadditive Tsallis entropy, the problem of quantum-state inference and the concept of quantum entanglement are studied. In particular, we examine in detail the 2-qubit system.

(3) S. Abe and A. K. Rajagopal, "Nonadditive conditional entropy and its significance for local realism", Physica A (2000), in press. The concept of conditional entropy is extended to the nonadditive composite systems by using the form invariance of the structures given by Khinchin's axiomatic foundations of information theory and the pseudoadditivity of Tsallis entropy. A criterion de-

duced from this for separability of density matrices for validity of local realism is examined in detail for the 2-qubit system. In another paper, S. Abe, “Axioms and uniqueness theorem for Tsallis entropy”, Phys. Lett. A **271**, 74 (2000), the Shannon-Khinchin axioms for ordinary information entropy is generalized to the nonextensive systems based on the concept of nonadditive conditional entropy, thus providing a complete proof of the uniqueness theorem for the Tsallis entropy.

(4) S. Abe and A. K. Rajagopal, “Nonuniqueness of canonical ensemble theory arising from microcanonical basis”, Phys. Lett. A (2000), in press, and S. Abe and A. K. Rajagopal, “Microcanonical foundation for systems with power-law distributions”, J. Phys. A (2000), in press. In these two papers, by an examination of the traditional microcanonical foundation of the Boltzmann-Gibbs canonical distribution, it is revealed that alternate paths exist to understand nonexponential distributions that also occur so commonly in nature.

(5) S. Abe and A. K. Rajagopal, “Justification of power-law canonical distributions based on generalized central limit theorem”, Europhys. Lett., in press. This paper employs the generalized law of large numbers (due to Lévy and Gnedenko) to justify the power law distributions appearing in fractal structures etc. by extending the discussion given by Khinchin for deriving Gibbsian canonical ensemble theory based on the well-known law of large numbers (ordinary central limit theorem).

Acknowledgments: It is a pleasure to thank Professors S. Abe and Y. Okamoto for providing me this opportunity to present my version of various aspects of the work on Tsallis ensemble by inviting me to present lectures in their IMS Winter School on Statistical Mechanics. My words are inadequate to express my indebtedness to Professors S. Abe, E. K. Lenzi, R. S Mendes, and C. Tsallis who endured my many e-mails, faxes, exchanges of preprints of their articles and notes, which has resulted in collaborations electronically during the preparation of this article. Many thanks are due to Ms. Socorro and Ms. Valeria for typing the manuscript and for patiently enduring the hardship of correcting it through faxes which were often unreadable etc.. It is my pleasant duty to place on record the enormous help of Professor S. Abe in preparing this lecture note. He has gone over it with extreme care and has made many important and crucial changes with my consultation, thus improving the notes considerably from its original version. This work is supported in part by the United States Office of Naval Research.

Appendix: q-Kullback–Leibler Entropy as a Guide?

In this, several forms of the definitions of the q-Kullback-Leibler entropy, all of which lead to the correct known expression for $q = 1$, are considered and only one of them is shown to be logically the appropriate choice. The usual expression for the Kullback-Leibler (KL) entropy is [Eq. (17)]

$$K_1(\hat{\rho};\hat{\rho}') = \mathrm{Tr}[\hat{\rho}(\ln\hat{\rho} - \ln\hat{\rho}')] \geq 0\ . \tag{A.1}$$

If one chooses the equilibrium density matrix in the canononical ensemble, for simplicity, $\hat{\rho}' = \hat{\rho}_e \exp\left[-\lambda_1(\hat{H} - \tilde{U}_1)\right]/\tilde{Z}_1$, (which is a slightly modified version of Eq. (18)), then we have the variational statement that any arbitrary density matrix one has the inequality

$$K_1(\hat{\rho}; \hat{\rho}_e) = S_1(\hat{\rho}_e) - S_1(\hat{\rho}) + \lambda_1 \mathrm{Tr}[\hat{\rho}(\hat{H} - \tilde{U}_1)] \geq 0 \ . \tag{A.2}$$

We want to use this as a guide to choose the correct form of the q-KL entropy which gives us this as the $q \to 1$ result.
A few possible choices of the q-KL entropy are

$$\begin{aligned} &(a) \;\; K_q^{(BPT)}(\hat{\rho}; \hat{\rho}') = \mathrm{Tr}[\hat{\rho}^q(\ln_q \hat{\rho} - \ln_q \hat{\rho}')] \geq 0 \ , \\ &(b) \;\; K_q^{(BPT)(a)}(\hat{\rho}; \hat{\rho}') = \mathrm{Tr}[\hat{\rho}^q\,(\ln_q \hat{\rho} - \ln_q \hat{\rho}')]/\mathrm{Tr}\hat{\rho}^q \geq 0 \ , \\ &(c) \;\; K_q(\hat{\rho}; \hat{\rho}') = \mathrm{Tr}[\hat{\rho}\,(\mathrm{Ln}_q\hat{\rho} - \mathrm{Ln}_q\hat{\rho}')] \ . \end{aligned} \tag{A.3}$$

In the above, $q \geq 0$ and $\ln_q x = (x^{1-q} - 1)/(1-q) = x^{1-q}\mathrm{Ln}_q x$. All these go over to Eq. (A.1) for $q \to 1$. (However, the third one turns out to have a problem with the nonnegativity condition.) If we put Eq. (2) as our second criterion, then we have two choices for the equilibrium q-density matrix, one arising from maxent using the normalized q-mean value constraints on the entropy defined in Eq. (2), and second arising from the alternate form, $S_q^{(a)}(\hat{\rho}) = S_q(\hat{\rho})/\mathrm{Tr}\hat{\rho}^q$, mentioned in Sec. 6. We find that the second expression in Eq. (A.3) with the maxent associated with $S_q^{(a)}$ gives the same form as Eq. (A.2) even though limit the $q \to 1$ is correctly given with either choice of the maxent density matrix in (a) and (b) whereas (c) does not give the same limit! We therefore give the answer for case (b) here:

$$\begin{aligned} \hat{\rho}_e^{(a)} &= \left[1 - c_q^{(a)}\lambda_1(1-q)\left(\hat{H} - \tilde{U}_q^{(a)}\right)\right]^{1/(1-q)} \Big/ \tilde{Z}_q^{(a)} \ , \\ \tilde{Z}_q^{(a)} &= \mathrm{Tr}\left[1 - c_q^{(a)}\lambda_1(1-q)\left(\hat{H} - \tilde{U}_q^{(a)}\right)\right]^{1/(1-q)} \ , \\ c_q^{(a)} &= \left[\tilde{Z}_q^{(a)}\right]^{1-q} = \left[1 - (1-q)S_q^{(a)}\right]^{-1} \ , \\ K^{(BPT)(a)}\left(\hat{\rho}; \hat{\rho}_e^{(a)}\right) &= S_q^{(a)}\left(\hat{\rho}_e^{(a)}\right) - S_q^{(a)}(\hat{\rho}) + \lambda_1 \frac{\mathrm{Tr}[\hat{\rho}^q\left(\hat{H} - \tilde{U}_q^{(a)}\right)]}{\mathrm{Tr}\hat{\rho}^q} \\ &\geq 0 \ . \end{aligned} \tag{A.4}$$

References

1. L. Boltzmann, Wien Ber. **53**, 195 (1866).
2. J. W. Gibbs, Tr. Conn. Acad. **3**, 108 (1876); see also his *Scientific Papers* Vols. 1 and 2 (Dover, New York, 1961).
3. J. von Neumann, Gott. Nach., **273** (1927); see also his book, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).

4. C. E. Shannon, Bell Syst. Tech. L., **27**, 379 (1946).
5. E. T. Jaynes, Phys. Rev. **106**, 620 (1957); see also his collection of papers, *E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, ed. R. D. Rosenkrantz (Kluwer, Dordrecht, 1989).
6. See for example, M. Ohya and D. Pertz, *Quantum Entropy and Its Use*, (Springer-Verlag, Berlin, 1993).
7. C. Tsallis, J. Stat. Phys. **52**, 479 (1988). Since this paper first appeared, a large number of papers based on this work have been and continues to be published in a wide variety of topics. A list of this vast literature which is continually being updated and enlarged may be obtained from http://tsallis.cat.cbpf.br/biblio.htm.
8. C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica A **261**, 534 (1998).
9. M. L. Lyra and C. Tsallis, Phys. Rev. Lett. **80**, 53 (1998).
10. C. Anteneodo and C. Tsallis, Phys. Rev. Lett. **80**, 5313 (1998).
11. A. K. Rajagopal, Phys. Rev. Lett. **76**, 3469 (1996).
12. N. L. Balazs and T. Bergeman, Phys. Rev. A **58**, 2359 (1998).
13. W. H. Louisell, *Quantum Statistical Properties of Radiation* (Wiley, New York, 1990), p. 323-328.
14. E. T. Jaynes and F. W. Cummings, Proc. I. R. E., **51**, 89 (1963). See also H. Paul, Ann. der Phys. **11**, 411 (1963), who also reported an exact solution to the JC model but did not explore in full detail the ramifications of the solution.
15. See for example various articles in *Physics and Probability – Essays in honor of Edwin T. Jaynes*, eds. W. T. Grandy, Jr. and P. W. Milonni (Cambridge University Press, Cambridge, 1993).
16. N. J. Cerf and C. Adami, Phys. Rev. Lett. **79**, 5194 (1997).
17. N. J. Cerf and C. Adami, Physica D **120**, 62 (1998).
18. A. K. Rajagopal, K. L. Jensen, and F. W. Cummings, Phys. Lett. A **259**, 285 (1999).
19. C. Tsallis, Phys. Rev. E **58**, 1442 (1998).
20. L. Borland, A. R. Plastino, and C. Tsallis, J. Math. Phys. **39**, 6490 (1998); (E) **40**, 2196 (1999).
21. Unpublished notes of S. Abe and A. K. Rajagopal. See *Note Added.*
22. R. Karplus and J. Schwinger, Phys. Rev. **73**, 1020 (1948).
23. A. K. Rajagopal, Braz. J. Phys. **29**, 61 (1998).
24. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
25. A. K. Rajagopal, R. S. Mendes, and E. K. Lenzi, Phys. Rev. Lett. **80**, 3907 (1998). See also an expanded version of this work by E. K. Lenzi, R. S. Mendes, and A. K. Rajagopal, Phys. Rev. E **59**, 1398 (1999).
26. L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).
27. E. K. Lenzi, R. S. Mendes, and A. K. Rajagopal, cond-mat/9904100. See also S. Abe, Eur. Phys. J. B **9**, 679 (1999).
28. K. Huang, *Statistical Mechanics* (Wiley, New York, 1963).
29. S. Abe, Phys. Lett. A **263**, 424 (1999); (E) **267**, 456 (2000).
30. A. R. Plastino, A. Plastino, and C. Tsallis, J. Phys. A: Math. Gen. **27**, 5707 (1994).
31. A. K. Rajagopal, Phys. Rev. A **54**, 3916 (1996).
32. A. K. Rajagopal, Phys. Lett. A **228**, 66 (1997).
33. A. K. Rajagopal, Physica A **253**, 271 (1998).
34. A. K. Rajagopal, Phys. Lett. A **246**, 237 (1998).
35. C. Kittel, *Elementary Statistical Physics* (Wiley, New York, 1958), Chapter 1.
36. P. T. Landsberg and V. Vedral, Phys. Lett. A **247**, 211 (1998).

37. R. J. V. dos Santos, J. Math. Phys. **38**, 4104 (1997).
38. S. Abe and A. K. Rajagopal (private ruminations, 1998).
39. P. Zeische, O. Gunnarsson, W. John, and H. Beck, Phys. Rev. B **55**, 10 270 (1997).
40. D. B. Ion and M. L. D. Ion, Phys. Rev. Lett. **81**, 5714 (1998).
41. A. K. Rajagopal, Phys. Lett. A **205**, 32 (1995).

III. Tsallis Theory, the Maximum Entropy Principle, and Evolution Equations

A.R. Plastino

Faculty of Astronomy and Geophysics, National University La Plata, C.C. 727,
(1900) La Plata, Argentina

Abstract. The generalized thermostatistics advanced by Tsallis in order to treat nonextensive systems has greatly increased the range of possible applications of statistical mechanics to the description of natural phenomena. Here we consider some aspects of the relationship between Tsallis' theory and Jaynes' maximum entropy (MaxEnt) principle. We review some universal properties of general thermostatistical formalisms based on an entropy extremalization principle. We explain how Tsallis formalism provides a useful tool in order to obtain MaxEnt solutions of nonlinear partial differential equations describing nonextensive physical systems. In particular, we consider the case of the Vlasov-Poisson equations, where Tsallis MaxEnt principle leads in a natural way to the stellar polytrope solutions. We pay special attention to the "P-picture" formulation of Tsallis generalized thermostatistics based on the entropic functional $\tilde{S}_q$ and standard linear constraints.

1 Introduction

In spite of its great success, the statistical mechanics paradigm based on the Boltzmann-Gibbs entropy seems to be unable to deal with many interesting physical scenarios [1–3]. Astronomical self-gravitat-ing systems constitute an important illustrative example of these difficulties [4–8]. A considerable effort has been devoted by astrophysicists to achieve a thermostatistical understanding of self-gravitating systems along the lines of standard statistical mechanics [6,7]. The nonextensivity effects associated with the long range of the gravitational interaction were the main reason for the failure of those attempts [5,7].

In 1988 Tsallis advanced a nonextensive generalization of the celebrated Boltzmann-Gibbs (BG) entropic measure [9]. The new entropy functional introduced by Tsallis [9] along with its associated generalized thermostatistics [10,11] is nowadays being hailed as the possible basis of a theoretical framework appropriate for the study of nonextensive systems [12–14]. This entropy has the form

$$S_q = \frac{1}{q-1}\left(1 - \int f(\mathbf{x})^q \, d\mathbf{x}\right). \tag{1}$$

where $\mathbf{x}$ is a dimensionless state-variable, f stands for the probability distribution and the Tsallis parameter q is any real number. The standard Boltzmann-Gibbs entropy

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 157–191, 2001.

$$S = -\int f \ln f d\mathbf{x} \tag{2}$$

is recovered in the limit case $q \to 1$. A simple illustration of the nonextensive character of the Tsallis measure S_q is provided by its behaviour in the case of a composite system constituted by two statistically independent subsystems A and B (that is, the joint probabilities verify $f_{A+B}(\mathbf{x}, \mathbf{x}') = f_A(\mathbf{x}) f_B(\mathbf{x}')$). The Tsallis entropy of the composite system $A + B$ verifies the so called q-additivity rule,

$$S_q(A + B) = S_q(A) + S_q(B) + (1 - q) S_q(A) S_q(B). \tag{3}$$

The standard additivity property of the entropy is recovered when $q = 1$. It is clear from equation (3) that q can be interpreted as a measure of the degree of nonextensivity.

Many relevant mathematical properties of the standard thermostatistics are preserved by Tsallis' formalism or admit suitable generalizations [12–20]. Tsallis' proposal was shown to be consistent both with Jaynes' information theory formulation of statistical mechanics [21], and with the dynamical thermostatting approach to statistical ensembles [22].

The recent development of many applications of Tsallis' theory is beginning to provide a picture of the kind of physical problems where the new approach may be useful. Self-gravitating systems were the first concrete physical systems discussed within the nonextensive thermostatistics [23,24]. That early application, in turn, inspired Boghosian's treatment of the two dimensional pure electron plasma, leading to the first experimental confirmation of Tsallis theory [25]. A possible solution of the solar neutrino puzzle based on Tsallis thermostatistics has been advanced [26]. Some cosmological implications of Tsallis proposal have also been worked out [27]. The behaviour of dissipative low dimensional chaotic systems [28,29], chaotic Hamiltonian systems with long range interactions [30], and self organized critical systems [31] have been discussed in connection with the new approach. Tsallis entropy has also been advanced as the basis of a thermostatistical foundation of Lévy flights and distributions [32]. Tsallis nonextensive statistical formalism proved to be a useful framework for the analysis of many interesting properties of nonlinear Fokker-Planck equations [33–39]. H-theorems based upon S_q have been analyzed both in the discrete case through a direct application of the master equation[40,41], and in the continuous one in connection with the Liouville and Fokker-Planck equations [42].

The q-generalization of the maximum entropy principle based upon the extremalization of Tsallis measure S_q under appropriate constraints plays a fundamental role within Tsallis theory and its applications. Consequently, it is to be expected that the interaction between Jaynes information theory approach to statistical mechanics, on the one hand, and Tsallis nonextensive formalism, on the other, will provide a fertile basis for new developments in statistical physics. We want to discuss here some aspects of the relationship between Jaynes' and Tsallis' ideas. As an illustration, we will consider their application to the study of two nonlinear partial differential equations: the nonlinear Fokker-Planck equations, and the Vlasov-Poisson equations.

2 Jaynes Maximum Entropy Principle

Although we are going to be concerned mainly with classical systems, many of the ideas and techniques discussed are also applicable to quantum systems. A classical Hamiltonian system in equilibrium at temperature T is described by the Gibbs canonical ensemble. The concomitant phase space probability distribution is

$$f_{Gibbs}(q_i, p_i) \;=\; \frac{1}{Z} e^{-\beta H(q_i, p_i)}, \tag{4}$$

where $\beta = 1/kT$ (k stands for the Boltzmann constant) and the partition function z is given by

$$z \;=\; \int e^{-\beta H(q_i, p_i)}\, d\Omega. \tag{5}$$

The canonical distribution (4) is the one that maximizes Boltzmann-Gibbs entropy,

$$S = -k \int f(q_i, p_i) \,\ln f(q_i, p_i)\, d\Omega \quad , \tag{6}$$

under the constraints imposed by normalization

$$\int f(q_i, p_i)\, d\Omega \;=\; 1, \tag{7}$$

and the mean value of the energy

$$\langle E \rangle \;=\; \int f(q_i, p_i)\, H(q_i, p_i)\, d\Omega. \tag{8}$$

In the above equations, the pair (q_i, p_i) denotes the complete set of phase space canonical coordinates, and $d\Omega$ stands for the corresponding phase space volume element. The fact that the canonical distribution constitutes an extremum of the entropy functional under appropriate constraints was already pointed out by Gibbs himself in his famous book on statistical mechanics [43]. However, Jaynes was the one to elevate this property to the status of a fundamental principle [44–48]. Inspired in Shannon's information theory, Jaynes reinterpreted and generalized the above mentioned extremum principle. According to Jaynes maximum entropy (MaxEnt) principle, when we have incomplete information about a probability distribution $f(\mathbf{x})$ ($\mathbf{x} \in R^D$ stands for a point in the relevant phase space. In the particular case of a Hamiltonian system's phase space we have $\mathbf{x} = (q_i, p_i)$.), we must adopt the probability distribution compatible with the available data that maximizes the Boltzmann-Gibbs-Shannon entropy

$$S[f] = -\int f(\mathbf{x}) \ln f(\mathbf{x}) d\mathbf{x}. \tag{9}$$

Following Shannon's ideas, Jaynes interpreted the above functional as a measure of the missing information about the system under consideration. Consequently, the extremalization of the entropy under the constraints imposed by the known (incomplete) information yields a probability distribution that incorporates the available data without assuming any further information that we don't actually have [46–49]. In this sense, Jaynes' MaxEnt prescription can be regarded as a mathematical formulation of the celebrated "Occam's razor" principle.

Usually, the input partial information consists of M relevant mean values,

$$\langle A_i \rangle = \int A_i(\mathbf{x}) \, f(\mathbf{x}) \, d\mathbf{x}, \qquad (i = 1, \ldots, M). \tag{10}$$

By introducing appropriate Lagrange multipliers (λ_i, $i = 1, \ldots, M$), Jaynes MaxEnt variational problem leads to the well known solution

$$f(\mathbf{x}) = \frac{1}{Z} \exp\left(-\sum_i \lambda_i \, A_i(\mathbf{x}) \right), \tag{11}$$

where the partition function Z is given by

$$Z = \int \exp\left(-\sum_i \lambda_i \, A_i(\mathbf{x}) \right) d\mathbf{x}. \tag{12}$$

The relevant mean values $\langle A_i \rangle$ and the associated Lagrange multipliers λ_i verify by the well known Jaynes relations

$$\frac{\partial S}{\partial \langle A_i \rangle} = \lambda_i, \tag{13}$$

and

$$\frac{\partial \ln Z}{\partial \lambda_i} = -\langle A_i \rangle. \tag{14}$$

Within Jaynes approach to Boltzmann-Gibbs thermostatistics, different choices for the relevant quantities $A_i(q_i, p_i)$ give rise to the various ensembles of standard statistical mechanics [46–48]. In particular, if we happen to know only the mean value of the energy (i.e., the Hamiltonian $H(q_i, p_i)$ is the only relevant quantity), the MaxEnt prescription yields the Gibbs canonical ensemble. If we have as prior information both the mean values of the energy and the number of particles, we get the macrocanonical ensemble. However, Jaynes approach allows to consider more general statistical ensembles [46–48]. When the relevant quantities involved are constants of motion, the associated MaxEnt distribution describes an equilibrium situation. But if the mean values of quantities that are not constant of motion are included as input information, time dependent nonequilibrium processes are also within the scope of the theory. In this way, Jaynes procedure greatly extends the range of phenomena that can be treated by the methods of statistical mechanics [47].

Although originally advanced as a new foundation for Boltzmann-Gibbs statistical mechanics, Jaynes MaxEnt principle is nowadays applied to a wide range of problems such as image reconstruction and other inverse problems with noisy and incomplete data [50,51] and time series analysis [52].

The standard MaxEnt principle based on the Boltzmann-Gibbs-Shannon entropy measure, in spite of its numerous and impressive successful applications, has serious limitations. There are important probability distributions in statistical physics and related fields that are not derivable from that principle in a natural way [3,12]. There is a growing body of evidence indicating that the generalized MaxEnt principle based on Tsallis nonextensive entropy, and the concomitant thermostatistics, may provide powerful tools for the analysis of problems outside the scope of the standard Boltzmann-Gibbs-Jaynes formalism [1–3].

3 General Thermostatistical Formalisms

Tsallis bold attempt to build up a complete thermostatistical formalism based upon a generalized nonlogarithmic entropy functional has raised many interesting issues related both to the mathematical structure and physical implications of general thermostatistical formalisms [53–56]. Tsallis pioneering work has stimulated the exploration of the properties of other generalized or alternative information measures [57–60]. On the other hand, it has been recently realized that some physically relevant features are shared by extended families of thermostatistical formalisms based on the extremalization of an entropy functional [53–56].

We now know that many important probability distribution appearing in statistical physics and related fields are derivable from an appropriate MaxEnt prescription [3,60]. In order to specify any particular implementation of the MaxEnt approach, one needs to provide the entropic functional that is to be extremalized, and the form of the constraints involved in the concomitant variational process. One may classify the MaxEnt formalisms in two main groups, according to the kind of constraints used.

3.1 MaxEnt Formalisms with Standard Linear Constraints

Here we consider generalizations of the canonical ensemble for Hamiltonian systems. However, all the discussion can easily be extended to the general scenario (not necessarily Hamiltonian) where more than one relevant mean values describing a system are known.

Let us now consider the general entropic functional

$$S_G = \int C(f)\, d\Omega. \tag{15}$$

The associated generalized canonical ensemble is obtained maximizing S_G under the constraints imposed by normalization

$$\int f(q_i, p_i)\, d\Omega \;=\; 1, \tag{16}$$

and the mean value of the energy,

$$\langle E \rangle \;=\; \int H(q_i, p_i)\, f(q_i, p_i)\, d\Omega. \tag{17}$$

By introducing the Lagrange multipliers α and β, associated respectively to the normalization and the energy constraints, the concomitant variational problem reads

$$\delta \left(S_G \,-\, \beta \int H\, f\, d\Omega \,-\, \alpha \int f\, d\Omega \right) \;=\; 0, \tag{18}$$

and its solution is given by

$$C'(f) \;=\; \alpha + \beta H. \tag{19}$$

If $g(x)$ denotes the inverse function of $C'(x)$,

$$g(C'(x)) \;=\; C'(g(x)) \;=\; x, \tag{20}$$

the solution of equation (19) may be written as

$$f(q_i, p_i) \;=\; g\big[\alpha \,+\, \beta H(q_i, p_i)\big]. \tag{21}$$

A similar result obtains when we have as prior information the mean values of more than one relevant quantities. Of course, in that case we have to introduce as many Lagrange multipliers as constraints.

3.2 MaxEnt Formalisms with Generalized Nonlinear Constraints

Tsallis and Curado introduced the idea of generalized mean values [10]. They reformulated the original [9] Tsallis entropy maximization principle using q-generalized constraints of the form [10]

$$\langle H \rangle_q \;=\; \int H(q_i, p_i)\, f^q(q_i, p_i)\, d\Omega. \tag{22}$$

The probability distribution obtained when Tsallis entropy is extremalized under the constraints imposed by (standard) normalization and the q-mean value of the energy reads [10]

$$f(q_i, p_i) \;=\; \frac{1}{Z_q}\, [1 - (1-q)\beta H(q_i, p_i)]^{1/(1-q)}, \tag{23}$$

where the partition function Z_q is defined by

$$Z_q = \int [1-(1-q)\beta H(q_i,p_i)]^{1/(1-q)}\, d\Omega. \tag{24}$$

It is important to stress that within this scheme, the canonical q-mean value of a quantity $A(q_i,p_i)$,

$$\langle A\rangle_q = \frac{1}{Z_q^q}\int A(q_i,p_i)\,[1-(1-q)\beta H(q_i,p_i)]^{q/(1-q)}\, d\Omega, \tag{25}$$

is still equal to the average of A computed with a weighting function (i.e., f^q) that depends on the canonical variables (q_i,p_i) only through the Hamiltonian function.

A new kind of normalized nonlinear q-mean values (denoted here by double angle brackets)

$$\langle\langle A\rangle\rangle_q = \frac{\int f^q(q_i,p_i)\,A(q_i,p_i)\,d\Omega}{\int f^q(q_i,p_i)\,d\Omega}, \tag{26}$$

have been recently introduced by Tsallis and coworkers [61]. However, it can be shown that a MaxEnt formalisms based upon these normalized q-mean values, after an appropriate redefinition of the entropy functional, is mathematically equivalent to a MaxEnt scheme based on standard linear constraints. Hence, these kind of generalized mean values can be incorporated within the MaxEnt formalisms considered in the previous subsection. This issue, which we think involves a relevant aspect of Tsallis theory, will be discussed in some detail in the next section.

3.3 Tsallis Entropy Plus Escort Mean Values

Tsallis and coworkers [61] have recently introduced a slight (but significant) modification of the original Tsallis nonextensive formalism. This new proposal is based on two main ingredients, which are i) an entropic (or information) measure, and ii) an appropriate form for the mean energy constraint. In this section we consider a system with discrete states because the matters under consideration are very general and hold both for classical and quantum scenarios. All the discussion can easily be recast in terms of classical systems with a continuous phase space.

In order to obtain the q-nonextensive canonical ensemble, Tsallis' functional (p_i denotes the probability of the microstate i with eigenenergy e_i)

$$S_q = \frac{1-\sum_i p_i^q}{q-1}, \tag{27}$$

is to be extremalized under the constraints imposed by normalization,

$$\sum_i p_i = 1, \tag{28}$$

and a normalized version of Tsallis generalized q-mean values,

$$\langle\langle U\rangle\rangle_q = \frac{\sum_i \epsilon_i p_i^q}{\sum_i p_i^q}. \tag{29}$$

The above MaxEnt procedure leads to the variational problem

$$\delta\left\{S_q - \beta\langle\langle U\rangle\rangle_q - \alpha\sum_i p_i\right\} = 0 \tag{30}$$

A detailed analysis of many properties of this formalism is given in [61]. This new MaxEnt scheme has the remarkable property of being reformulable entirely in terms of *standard linear constraints*. This reformulation of the q-nonextensive formalism is based on the concept of *escort distributions* [49]. Given the probability distribution $\{p_i\}$, the associated escort distribution $\{P_i\}$ is defined by the transformation

$$P_i = \frac{p_i^q}{\sum_j p_j^q}. \tag{31}$$

From the normalization requirement on the p_i's, it is easy to arrive at

$$p_i^q = \frac{P_i}{\left(\sum_j P_j^{1/q}\right)^q}, \tag{32}$$

which, in turn, allows us to write both the entropy S_q and the energy constraint $\langle U\rangle_q$ in terms of the new variables P_i,

$$\tilde{S}_q\left[P_i\right] = \frac{1-\left(\sum_i P_i^{1/q}\right)^{-q}}{q-1}, \tag{33}$$

and

$$\langle\langle U\rangle\rangle_q = \sum_i \epsilon_i P_i. \tag{34}$$

It is also evident that the new variables P_i are appropriately normalized

$$\sum_i P_i = 1. \tag{35}$$

Making use of the above expressions it is possible now to reformulate our MaxEnt variational principle under the guise

$$\delta\left\{\tilde{S}_q\left[P_i\right] - \beta\langle\langle U\rangle\rangle_q - \alpha\sum_i P_i\right\} = 0. \tag{36}$$

It is important to stress that the equations (30) and (36) represent one and the same MaxEnt principle, expressed in terms of two different sets of variables. In order to distinguish the two (mathematically equivalent) formulations, we are going to refer to the version given by equation (30) as the "p-picture". On the other hand, the formulation embodied in Equation (36) will be called the "P-picture". The new P-picture is just a mathematical reformulation of the MaxEnt scheme associated with the original p-picture. However, the entropic functional (33) is a legitimate measure in its own right, whose properties can be studied without any reference to the associated p-picture. Two interesting characteristics of the entropy functional (33) are that it complies with Tsallis q-generalized additivity law, and that its associated MaxEnt canonical ensemble is invariant under an homogeneous shift of the energy eigenvalues. Of course, both these features are already known from the p-picture analysis [61]. However, their reconsideration in terms of the P-picture provides further insight into the properties of more general thermostatistical formalisms.

3.4 Invariance Under Uniform Shifts of the Hamiltonian Eigenenergies

One of the most disturbing properties of the Tsallis-Curado MaxEnt variational principle, which is based upon the q-generalized constraints

$$\langle U \rangle_q^{(TC)} = \sum_i \epsilon_i p_i^q, \tag{37}$$

is that the concomitant generalization of the canonical distribution lacks invariance under uniform shifts of the Hamiltonian's energy spectrum

$$\{\epsilon_i\} \to \{\epsilon_i + \varDelta\}. \tag{38}$$

However, it was recently discovered [61] that this fundamental invariance principle of statistical mechanics is recovered if Tsallis entropy S_q is extremalized under the constraints imposed by the standard normalization prescription, along with the escort mean values

$$\langle\langle U \rangle\rangle_q = \frac{\sum_i \epsilon_i p_i^q}{\sum_i p_i^q}. \tag{39}$$

The P-picture formulation (36) of the Tsallis MaxEnt principle with normalized q-constraints (30) provides a deeper understanding of the invariance under uniform energy shifts [56]. The P-picture formulation is characterized by the fact that both the normalization and the mean value constraints adopt the standard *linear* dependence on the probabilities. This means that when the mean values are computed in the ordinary and intuitively appealing way, the energy shift invariance holds for a whole monoparametric family of entropies embracing the standard logarithmic measure as a particular case. This fact immediately suggests the question of how large is the complete set of entropy functionals sharing this property. Let us consider the general MaxEnt problem of extremalizing the measure [56]

$$S_G = \sum_i C(p_i) \tag{40}$$

under the constraints imposed by normalization

$$\sum_i p_i = 1 \tag{41}$$

and the *linear* energy mean value

$$\langle U \rangle = \sum_i \epsilon_i p_i . \tag{42}$$

The associated variational problem

$$\delta \left\{ S_G - \beta \langle U \rangle - \alpha \sum_i p_i \right\} = 0 \tag{43}$$

leads to

$$C'(p_i) = \beta \epsilon_i + \alpha \quad (i = 1 \ldots, N). \tag{44}$$

The probabilities $\{p_i\}$ complying with the above set of equations constitute the "generalized" canonical ensemble

$$p_i = g(\alpha + \beta \epsilon_i) \quad (i = 1 \ldots, N), \tag{45}$$

$g(x)$ being the inverse function of $C'(x)$. Notice that $g(x)$ is a universal function, in the sense that its form does not depend on the eigenenergies of the particular system under consideration. On the other hand, and because of the normalization condition, the Lagrange multiplier $\alpha = \alpha(\beta)$ can be regarded as a function of the inverse temperature β. The form of the function $\alpha(\beta)$ does depend on the eigenenergies.

Let us now consider two systems related by a uniform shift in the eigenenergies $\epsilon'_i = \epsilon_i + \Delta$, and characterized by the same value of β. That is, we compare two systems that have different energy spectrum but are at the same temperature. The new canonical distribution associated with the shifted eigenenergies is

$$p'_i = g(\alpha' + \beta \epsilon'_i) = g(\alpha' + \beta \Delta + \beta \epsilon_i) \quad (1 = 1, \ldots, N). \tag{46}$$

Comparing now the two canonical distributions (45) and (46), it is easy to see that the normalization requirement implies

$$\alpha \ = \ \alpha' + \beta \Delta, \tag{47}$$

and consequently

$$p_i \ = \ p'_i. \quad (1 = 1, \ldots, N). \tag{48}$$

We conclude that the invariance under uniform spectral translations is a UNIVERSAL property of the MaxEnt prescription. It is verified by any entropy of the form $S = \sum_i C(p_i)$, whenever standard linear constraints are adopted [56]. It is instructive to arrive at the above result by an alternative reasoning based entirely on the variational problem itself (i.e. without recourse to the explicit form of the canonical distribution). Let us assume that, in the original "unshifted" system, the concrete value adopted by the mean energy is U. That is, we maximize the entropy under the requirements that the mean energy be equal to U and that the probabilities add up to unity. We also consider the "shifted" MaxEnt problem, where the mean energy of the shifted Hamiltonian is $U + \Delta$. The energy constraints associated, respectively, to our original and shifted Hamiltonians are

$$U = \sum_i p_i \epsilon_i, \tag{49}$$

and

$$U + \Delta = \sum_i p_i \epsilon_i' = \sum_i p_i \epsilon_i + \sum_i p_i \Delta. \tag{50}$$

These two equations, considered simultaneously with the normalization requisite, clearly define the same set of constraints on the probabilities. Hence, the solution to both variational problems must be one and the same. *The specific functional form of the maximized entropy does not play any role in the above argument* [56].

The above discussion leads to some interesting conclusions:

- The factorizability of the exponential function is not the fundamental reason for the shift invariance of the standard Gibbs canonical ensemble.
- The first proposal by Tsallis to generalize the Gibbs-Jaynes approach to thermostatistics [9], based upon his nonextensive entropy, but keeping the usual linear mean energy constraint, *does preserve the shift invariance.* This important property was overlooked because of a misidentification of the Lagrange multiplier associated with the energy [9].
- The underlying reason for the shift invariance exhibited by the nonextensive thermostatistics when escort mean values are adopt-ed, is that the corresponding formalism can be reformulated entirely in terms of linear mean values. In fact, any form of generalized escort mean values

  $$\langle\langle U \rangle\rangle_\phi = \frac{\sum_i \phi(p_i)\epsilon_i}{\sum_i \phi(p_i)} \tag{51}$$

  based on an arbitrary function $\phi(x)$ would do. However, in the particular case of Tsallis' nonextensive formalism, there are plenty of other mathematical reasons, as well as physical evidences, that point towards a simple power law as the appropriate choice [61].

3.5 Other Universal Properties of General Thermostatistical Formalisms

Other two important universal properties of general thermostatistical formalism are the Legendre transform structure of the concomitant maximum entropy scheme, and the classical equipartition theorem. The thermostatistical formalism associated with the extremalization of the general entropy functional (40) under the constraints of normalization (41) and mean energy (42) complies with the relations associated with the Legendre transform structure of thermodynamics. Within general thermostatistical formalisms we have that [53–55]

$$\frac{dS_G}{d\langle U\rangle} = \beta. \tag{52}$$

If we define the Jaynes' parameter

$$\lambda_J(\beta) = S_G - \beta\langle U\rangle, \tag{53}$$

which is the Legendre transform of S_G, then,

$$\frac{d\lambda_J}{d\beta} = -\langle U\rangle. \tag{54}$$

Relations (52-54) constitute the core of the connection between statistical mechanics and thermodynamics [9,10,55]. They hold true also for generalized nonlinear mean values [53–55], and when more than one relevant mean value is given as input prior information [53].

There exists also an appropriate generalization of the classical equipartition theorem that is verified within a large family of thermostatistical formalisms [62]. The contribution to the total mean energy of a system in thermal equilibrium at temperature $T = 1/k\beta$ due to an homogeneous term J of the Hamiltonian is given by [62]

$$\langle J\rangle = \frac{L}{\gamma}\, kT\, I(T), \tag{55}$$

where L is the number of arguments (canonical coordinates and momenta) entering the function g and γ denotes the corresponding degree of homogeneity. $I(T)$ is a function of the temperature that does not depend upon the detailed form of function g, nor upon its particular arguments, nor on γ. In the particular case of Boltzmann-Gibbs thermostatistics we have, of course, $I = 1$.

4 Time Dependent MaxEnt

Within Jaynes approach to time dependent problems one does not try to follow the evolution of the system in all its full detail. Instead, a reduced description of its behaviour based on the maximum entropy principle is considered [63–66]. At each time, the system is described by a probability distribution function (or statistical operator, in the case of quantum mechanics) that maximizes the Shannon entropy ($S = -\int f(\mathbf{x}) \ln f(\mathbf{x}) d\mathbf{x}$ or $S = -Tr(\rho \ln \rho)$) under the constraints imposed by normalization and the instantaneous mean values of an appropriate set of relevant quantities. This procedure was introduced by Jaynes in order to provide a new formulation of statistical mechanics based on information theory [46]. Jaynes applied these ideas to the von Neuman equation and the Liouville equation for Hamiltonian systems. Within that context, all the ensembles appearing in equilibrium statistical mechanics can be obtained if we choose constants of motion as relevant constraints. However, Jaynes' maxent scheme can also be implemented including non-conserved quantities within the set of relevant observables. This indeed constitutes one of the most interesting aspects of Jaynes' approach, since it allows for the description of non-equilibrium situations. For instance, let us consider a quantum system with Hamiltonian $\hat{H}$ described by the density operator $\hat{\rho}$ whose evolution obeys the von Neuman equation

$$\frac{d\hat{\rho}}{dt} = -i\,[\hat{H}, \hat{\rho}], \tag{56}$$

where Planck's constant was set equal to 1. The time derivative of the expectation value of an observable $\hat{O}_i$ is then given by

$$\frac{d\langle \hat{O}_i \rangle}{dt} = i\,\langle [\hat{H}, \hat{O}_i] \rangle. \tag{57}$$

Now, suppose we have a set of M observables $\hat{O}_i$, $i = 1, \ldots, M$ that close a semialgebra under commutation with the Hamiltonian operator

$$[\hat{H}, \hat{O}_i] = \sum_{j=1}^{M} g_{ij}\, \hat{O}_j, \quad i = 1, \ldots, M, \tag{58}$$

where g_{ij} are the concomitant structure constants. It is easy to verify that the mean values of the M observables evolve according to the closed set of linear ordinary differential equations,

$$\frac{d\langle \hat{O}_i \rangle}{dt} = i \sum_{j=1}^{M} g_{ij}\, \langle \hat{O}_j \rangle, \quad i = 1, \ldots, M. \tag{59}$$

Under these circumstances, the M observables $\langle \hat{O}_i \rangle$ are as good as integrals of motions in order to implement Jaynes approach [63–66]. Solving the linear differential equations (59), the expectation values $\langle \hat{O}_i \rangle(t)$ at any time t can be computed from their values at an initial time t_0. Consequently, the associated MaxEnt statistical operator $\hat{\rho}_{ME}$ can be obtained at any time t, as happens in the equilibrium situation. Indeed, it can be shown that this MaxEnt density operator constitutes an exact solution to the von Neuman equation.

Of course, the closure condition (58) does not always hold. However, even if the relevant observables $\hat{O}_i$ do not close a semialgebra with the Hamiltonian, it is still possible to close the set of equations (57) in a nonlinear and approximate way by recourse to an appropriate maximum entropy scheme. The expectation values appearing on the right hand sides of equations (57) can be evaluated using a maximum entropy density matrix $\hat{\rho}_{Me}$ determined, at each time t, by the constraints imposed by the M instantaneous expectation values $\langle \hat{O}_i \rangle$. As an example of this scheme, let us consider a quantum many body system. If the set of relevant observables $\hat{O}_i$ consists only of one-particle operators, then the maximum entropy approximate closure scheme leads to the well known time dependent Hartree-Fock approximation.

These MaxEnt ideas have been applied to a variety of classical evolution equations. A maximum entropy scheme has been numerically implemented for the cosmic ray transport equation [67]. The MaxEnt formalism has been applied to the Liouville equation associated with dynamical systems endowed with divergenceless phase space flows [68], to the (linear) Fokker-Planck equation [69], and to a more general family of evolution equations with the form of linear continuity equations [70]. The MaxEnt-Minimum Norm approach has been extended to the time dependent case [71]. The inverse problem of reconstructing the underlying microscopic dynamics from time-series by recourse to maximum entropy ideas has also been addressed [52].

5 Time-Dependent Tsallis MaxEnt Solutions of the Nonlinear Fokker–Planck Equation

5.1 The Nonlinear Fokker-Planck Equation

The nonlinear Fokker-Planck equation is characterized by a diffusion term that depends on a power η of the probability density, the drift term being linear as in the ordinary case. When $\eta = 1$ the standard linear Fokker-Planck is recovered. The nonlinear Fokker-Planck equation provides useful mathematical models for a variety of interesting physical processes such as the percolation of gases through porous media [72], thin liquid films spreading under gravity [73], surface growth [74], and some self-organizing phenomena [75]. It has also been used to describe systems showing anomalous diffusion of the correlated type [38,39]. The nonlinear, generalized N-dimensional Fokker-Planck equation is given by

$$\frac{\partial F}{\partial t} = L_R \, F \; + \; L_D \; (F^\eta) \, , \tag{60}$$

where $F(\mathbf{x}, t)$ is the normalized distribution function, and L_R and L_D are linear differential operators.

Earlier work [33–35] on the application of Tsallis formalism to the nonlinear Fokker-Planck equation was based on the unnormalized Tsallis q-constraints. Within that formulation of Tsallis theory, the nonlinearity in the diffusion equation and the entropic index q parametrizing the Tsallis entropy are connected by the relation $\eta = 2 - q$. It was shown that there are particular exact solutions of the nonlinear Fokker-Planck equation characterized by η that maximize the Tsallis entropy of index $q = 2 - \eta$ [35]. As we shall see, if the q-thermostatistics with normalized q-values is used, this relation must be changed into $q = 1/\eta$.

The drift term

$$L_R \, F \; = \; -\sum_{i=1}^{N} \frac{\partial (K_i \, F)}{\partial x_i} \tag{61}$$

is due to the deterministic forces associated to the drift vector $\mathbf{K}(\mathbf{x}) \in R^N$ of components K_i, while the nonlinear diffusion term

$$L_D \;\; (F^\eta) \; = \; \sum_{i,j=1}^{N} \frac{\partial}{\partial x_i} \left(D_{ij}(\mathbf{x}) \; \frac{\partial F^\eta}{\partial x_j} \right) \tag{62}$$

describes the effect of stochastic forces characterized by the diffusion tensor $\mathbf{D}(\mathbf{x}) \in R^{N\times N}$, of components D_{ij}. Note that the diffusion coefficients may depend on the state variable $\mathbf{x}$. In the limit case $\eta \to 1$ we recover the standard N-dimensional *linear* Fokker-Planck equation,

$$\frac{\partial F}{\partial t} \; = \; L_{FP} F, \tag{63}$$

which can be written in terms of just one single linear differential operator [76] $L_{FP} = L_R + L_D$.

It will prove useful for our forthcoming discussions to introduce the adjoint operators $L_R^\dagger$ and $L_D^\dagger$, defined by

$$\int F_1 \;\; (L_R \; F_2) \;\; d\mathbf{x} \;\; = \;\; \int \left(L_R^\dagger \; F_1 \right) \;\; F_2 \; d\mathbf{x}, \tag{64}$$

and

$$\int F_1 \;\; (L_D \; F_2) \;\; d\mathbf{x} \;\; = \;\; \int \left(L_D^\dagger \; F_1 \right) \;\; F_2 \; d\mathbf{x}, \tag{65}$$

for any two probability distributions F_1 and F_2. These adjoint operators are given by

$$L_R^\dagger \; = \sum_{i=1}^{N} K_i \, \frac{\partial}{\partial x_i} \tag{66}$$

and

$$L_D^\dagger \; = \; \sum_{i,j=1}^{N} \frac{\partial}{\partial x_j} \;\; D_{ij}(\mathbf{x}) \;\; \frac{\partial}{\partial x_i}. \tag{67}$$

Defined in a similar way, the adjoint operator $L_{FP}^\dagger = L_R^\dagger + L_D^\dagger$ plays an important role in the study of the *linear* Fokker-Planck equation [76].

5.2 Generalized Maximum Entropy Approach

q-Maximum Entropy Approximation As mentioned in Section IV, the main idea of the MaxEnt approach to time dependent probability distribution functions is to study the behaviour of a relatively small number of relevant mean values instead of trying to compute the evolution of the complete distribution $f(\mathbf{x}, t)$. Here we shall reformulate these ideas within the q-nonextensive formalism based upon Tsallis entropy and the *normalized* q-constraints. The q-MaxEnt approach based on unnormalized q-values was developed in [35]. Let us consider a set of M normalized q-mean values

$$\begin{aligned} \langle\langle A_i \rangle\rangle_q &= \frac{\int f^q \, A_i(\mathbf{x}) \, d\mathbf{x}}{\int f^q \, d\mathbf{x}} \\ &= \int F \, A_i(\mathbf{x}) \, d\mathbf{x} \qquad (i = 1, ..., M), \end{aligned} \tag{68}$$

where

$$F(\mathbf{x}) = \frac{f^q(\mathbf{x})}{\int f^q \, d\mathbf{x}}. \tag{69}$$

The formalism which we are now going to develop can alternatively be obtained using the unnormalized mean values $\langle A_i \rangle_q \equiv \int f^q A_i d\mathbf{x}$ [10] as well as by using the standard mean values $\langle A_i \rangle = \int A_i f d\mathbf{x}$. It is important to stress here that, among these three types of mean values, the normalized q-mean values $\langle\langle A_i \rangle\rangle_q$ are the ones best suited in order to develop a well-behaved thermostatistical formalism [61]. However, due to the fact that some relevant properties are shared by the three kinds of mean values (namely, the power-law form of Tsallis MaxEnt distributions and the Legendre transform structure of the concomitant thermostatistical formalism) [61], the results of the present application can be formulated in terms of any of those three types of mean values. Here we choose to work with the mean values given by (68) in order to illustrate that it is possible to implement Jaynes approach to time dependent problems on the basis of the Tsallis formalism with escort constraints. Besides, it is important to consider a q-MaxEnt scheme using the escort mean values because these ideas may be indispensable in order to apply Tsallis formalism to other nonlinear equations, like the Vlasov-Poisson system that we are going to discuss in the next section. However, if the Tsallis MaxEnt distribution is used only as an ansatz to find *exact* solutions for the nonlinear Fokker-Planck equation, the formalism based upon the unnormalized q-mean values leads, by recourse to simpler mathematical manipulations, to the same results as the ones we obtain here [35].

We assume that the evolution of the probability F is governed by the nonlinear Fokker-Planck equation (60). Consequently, $f(\mathbf{x}, t)$ does not verify that evolution equation. But f will play no role within our considerations. This implies that we are going to study Tsallis MaxEnt (approximate) solutions within the P-picture discussed in section III. We are going to use the P-picture formulation as a MaxEnt formalism in its own right. The concomitant MaxEnt solutions are to be obtained maximizing the functional

$$\tilde{S}_q[F] = \frac{1 - \left(\int F^{1/q} d\mathbf{x}\right)^{-q}}{q-1} \tag{70}$$

under appropriate constraints. We want to stress that the functional (70) is just the standard (!) Tsallis entropy written in terms of the escort "P" variables (now F).

The time derivatives of the M relevant moments of $F(\mathbf{x}, t)$ are given by (in this and the next sections we drop both the subindex q and the $\langle\langle\rangle\rangle$ notation from the mean values, since they are ordinary linear mean values on F)

$$\frac{d\langle A_i\rangle}{dt} = \int \{L_R\, F \;+\; L_D\; (F^\eta)\}\; A_i\; d\mathbf{x} \qquad (i = 1, \ldots, M). \tag{71}$$

Unfortunately, these equations do not, in general, constitute a closed system of ordinary differential equations of motion for the mean values $\langle A_i\rangle$. The integrals appearing on the right hand sides of equations (71) are not, in general, equal to linear combinations of the original mean values $\langle A_i\rangle$. Here enters the Maximum Entropy Principle. We can "close" the set (71), in an approximate way, by evaluating the right hand sides using, at each instant of time, the MaxEnt distribution $F_{ME}(\mathbf{x}, t)$ that maximizes Tsallis' entropy (70) under the constraints imposed by normalization and the M instantaneous values of $\langle A_i\rangle$. The concomitant variational problem has the analytical solution

$$\begin{aligned} F_{ME}(\mathbf{x}, t) \;=\; & \left\{\int F_{ME}^{1/q} d\mathbf{x}\right\}^{\frac{q(1+q)}{1-q}} \\ & \times \left[(q-1)\alpha \;-\; (1-q)\sum_{i=1}^{M} \lambda_i(t)\; A_i(\mathbf{x})\right]^{\frac{q}{1-q}}, \end{aligned} \tag{72}$$

where the ($\lambda_i,\;\; i = 1, \ldots, M$) are appropriate Lagrange multipliers that guarantee compliance with the given constraints, and α is the Lagrange multiplier associated with the normalization requirement. The generalized entropy $\tilde{S}_q$, the relevant mean values, and the concomitant Lagrange multipliers, are related by Jaynes' thermodynamical relations,

$$\frac{\partial\; \tilde{S}_q}{\partial\; \langle A_i\rangle} = \lambda_i. \tag{73}$$

Behaviour of the Generalized Entropy $\tilde{S}_q$ The time derivative of the entropy constitutes one of the most important qualitative features characterizing the behaviour of probability density functions associated with irreversible processes. When that derivative has a definite sign, an H-theorem holds, and a useful mathematical realization of “the arrow of time” becomes available. Given an approximate scheme for solving the evolution equations describing the system under study, it is crucial to know how close the behaviour of the entropy evaluated on the approximate solutions follows the evolution of the entropy of the exact solutions. Here we shall compare the behaviour of the Tsallis entropy $\tilde{S}_q$ corresponding to exact solutions of the nonlinear Fokker-Planck equation, with the behaviour associated to our MaxEnt solutions.

If $F(\mathbf{x},t)$ is an exact solution of (1), we have

$$\frac{d\tilde{S}_q}{dt} = \frac{1}{q-1}\left\{\int F^{1/q} d\mathbf{x})\right\}^{-1-q} \int F^{\frac{1-q}{q}} \left\{L_R\, F \;+\; L_D\; (F^{\eta})\right\}\; d\mathbf{x}, \tag{74}$$

which, after introducing the functional

$$D_q\left[F\right] = \frac{1}{q-1}\left\{\int F^{1/q} d\mathbf{x})\right\}^{-1-q} \int F^{\frac{1-q}{q}} \left\{L_R\, F \;+\; L_D\; (F^{\eta})\right\}\; d\mathbf{x}, \tag{75}$$

can be cast under the guise

$$\frac{d\tilde{S}_q}{dt} = D_q\left[F\right]. \tag{76}$$

This last expression is preserved by the maximum entropy approach, as can be seen as follows. We have

$$\frac{d\tilde{S}_q\left[F_{ME}\right]}{dt} = \sum_{i=1}^{M} \frac{\partial\;\tilde{S}_q}{\partial\;\langle A_i\rangle}\frac{d\langle A_i\rangle}{dt} = \sum_{i=1}^{M}\lambda_i \frac{d\langle A_i\rangle}{dt} \tag{77}$$

which yields

$$\frac{d\tilde{S}_q\left[F_{ME}\right]}{dt} = \int \left\{L_R\; F_{ME} \;+\; L_D\; (F_{ME}^{\eta})\right\}\; \left[\sum_{i=1}^{M}\lambda_i\; A_i\right]\; d\mathbf{x}. \tag{78}$$

But

$$\sum_{i=1}^{M}\lambda_i\; A_i = \frac{1}{q-1}\; \left[\int F_{ME}^{1/q} d\mathbf{x}\right]^{-1-q} F_{ME}^{\frac{1-q}{q}} \;-\; \alpha \tag{79}$$

so that

$$\begin{aligned}\frac{d\tilde{S}_q\left[F_{ME}\right]}{dt} = \frac{1}{q-1}\int &\left\{L_R\; F_{ME} \;+\; L_D\; (F_{ME}^{\eta})\right\} \\ &\times\left\{\;\left[\int F_{ME}^{1/q} d\mathbf{x}\right]^{-1-q} F_{ME}^{\frac{1-q}{q}} \;-\; (q-1)\alpha\right\}\; d\mathbf{x}.\end{aligned} \tag{80}$$

We now assume that (remember that α is just a number and does not depend on $\mathbf{x}$)

$$\int (L_R F_{ME})\ \alpha\ d\mathbf{x}\ =\ \int \left(L_R^{\dagger}\alpha\right)\ F_{ME}\ d\mathbf{x} = 0, \tag{81}$$

and

$$\int (L_D F_{ME})\ \alpha\ d\mathbf{x}\ =\ \int \left(L_D^{\dagger}\alpha\right)\ F_{ME}\ d\mathbf{x} = 0. \tag{82}$$

These last two equations involve an integration by parts procedure. We assume that F_{ME} verifies appropriate boundary conditions (essentially, it goes to zero fast enough with $|\mathbf{x}| \to 0$) in order for the "integrated part" being zero. Then we obtain

$$\begin{aligned} \frac{d\tilde{S}_q\left[F_{ME}\right]}{dt} &= \frac{1}{q-1}\left[\int F_{ME}^{1/q} d\mathbf{x}\right]^{-1-q} \\ &\quad \times \int \left\{L_R\ F_{ME}\ +\ L_D\ \left(F_{ME}^{\eta}\right)\right\}\ F_{ME}^{\frac{1-q}{q}}\ d\mathbf{x} \\ &= D_q\left[F_{ME}\right] \end{aligned} \tag{83}$$

and have therewith proved that

$$\frac{d\tilde{S}_q[F_{ME}]}{dt} = D_q\left[F_{ME}\right]. \tag{84}$$

We can conclude that the functional relation giving the time derivative of the Tsallis entropy $\tilde{S}_q$ in terms of the approximate maxent ansatz F_{ME} is the same as the one verified in the case of the unknown *exact solutions*. This important property is verified in general, for any (exact) solution of the nonlinear Fokker-Planck equation and regardless of the particular set of relevant mean values $\langle A_i \rangle$ employed in order to build up the corresponding Tsallis maximum entropy approximation. The present derivation explicitly makes use of the q-MaxEnt form of the (approximate) q-MaxEnt solution. However, it is possible that a similar property holds within a more general context involving other kinds of evolution equations endowed with an associated "natural" entropic measure.

Hamiltonian Structure We shall now consider the equations of motion for the Lagrange multipliers λ_i, and study how are they related to the equations of motion of the associated mean values $\langle A_i \rangle$. Making use of Jaynes' thermodynamic relations, we have

$$\begin{aligned} \frac{d\lambda_i}{dt} &= \sum_{j=1}^{M} \frac{\partial \lambda_i}{\partial \langle A_j \rangle}\,\frac{d\langle A_j \rangle}{dt} \\ &= \sum_{j=1}^{M} \frac{\partial^2 \tilde{S}_q}{\partial \langle A_i \rangle_q \partial \langle A_j \rangle}\,\frac{d\langle A_j \rangle}{dt} \end{aligned}$$

$$= \sum_{j=1}^{M} \frac{\partial \lambda_j}{\partial \langle A_i \rangle} \frac{d\langle A_j \rangle}{dt}, \tag{85}$$

which can be recast as

$$\frac{d\lambda_i}{dt} = \frac{\partial}{\partial \langle A_i \rangle} \left[\sum_{j=1}^{M} \lambda_j \frac{d\langle A_j \rangle}{dt} \right] - \sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \left(\frac{d\langle A_j \rangle}{dt} \right). \tag{86}$$

Making use now of the equations of motion (71) for the relevant mean values, and of the expression (77) for the time derivative of the entropy, we finally obtain

$$\begin{aligned} \frac{d\lambda_i}{dt} &= \frac{\partial}{\partial \langle A_i \rangle} \left[\frac{d\tilde{S}q[F_{ME}]}{dt} \right] \\ &\quad - \sum_{j=1}^{M} \lambda_j \frac{\partial}{\partial \langle A_i \rangle} \int \{ L_R \; F_{ME} \; + \; L_D \; (F_{ME}^{\eta}) \} \; A_j \; d\mathbf{x}. \end{aligned} \tag{87}$$

Introducing now the Hamiltonian

$$\begin{aligned} &H(\langle A_1 \rangle, \ldots, \langle A_M \rangle, \lambda_1, \ldots \lambda_M) \\ &= D_q[F_{ME}] - \sum_{j=1}^{M} \lambda_j \int \{ L_R \; F_{ME} + L_D \; (F_{ME}^{\eta}) \} \, A_j d\mathbf{x}, \end{aligned} \tag{88}$$

the equations of motion for the relevant mean values and their associated Lagrange multipliers can be put in a Hamiltonian way,

$$\frac{d\langle A_i \rangle}{dt} = -\frac{\partial H}{\partial \lambda_i}, \tag{89}$$

and

$$\frac{d\lambda_i}{dt} = \frac{\partial H}{\partial \langle A_i \rangle}. \tag{90}$$

In the expression (88) defining our Hamiltonian function, the MaxEnt distribution F_{ME} should be regarded as parametrized by the set of M relevant mean values $\langle A_i \rangle$. That is to say, the specific values adopted by the M quantities $\langle A_i \rangle$ determine, via the MaxEnt recipe, a particular distribution F_{ME}. In this way, the functionals of F_{ME} appearing in (88) are functions of the M quantities $\langle A_i \rangle$. Consequently, the only dependence of the Hamiltonian on the Lagrange multipliers is the one explicitly shown in equation (88). It is important to realize that ours is not a time dependent Hamiltonian, since its functional dependence on the relevant mean values and their associated Lagrange multipliers is dictated by the MaxEnt procedure and does not depend on time.

Note also that although all the solutions of our time dependent MaxEnt scheme evolve according to the Hamiltonian equations (89,90), not all the orbits associated with the Hamiltonian (88) constitute realizations of our MaxEnt approach. The orbits that are relevant to our problem are those whose initial conditions verify

$$\begin{aligned}\langle A_j\rangle = &\left\{\int F_{ME}^{1/q} d\mathbf{x})\right\}^{\frac{q(1+q)}{1-q}} \\ &\times \int \left[(q-1)\alpha - (1-q)\sum_{i=1}^{M} \lambda_i(t) A_i(\mathbf{x})\right]^{\frac{q}{1-q}} A_j(\mathbf{x}) d\mathbf{x} \\ &(i = 1, \ldots, M). \qquad (91)\end{aligned}$$

These M equations determine an M-dimensional submanifold of the full $2M$ phase space $(\langle A_1\rangle, \ldots, \langle A_M\rangle_q, \lambda_1, \ldots, \lambda_M)$. This submanifold constitutes an invariant set of our Hamiltonian dynamical system. That is to say, any orbit with initial conditions belonging to the set stays within it forever.

Summing up, we conclude that within the present time dependent thermostatistical context the relevant mean values and their concomitant Lagrange multipliers are conjugate variables not only in the thermodynamical sense, but also in the Hamiltonian phase space sense.

Exact *q*-MaxEnt Solutions It is known from earlier work [33,34] that the one dimensional nonlinear Fokker-Planck equation with constant diffusion coefficient and linear drift has particular exact solutions exhibiting the Tsallis q-MaxEnt form given by

$$F_{ME} = \frac{1}{Z}\left[1 - (1-q)\beta(x-x_0)^2\right]^{1/(1-q)}, \qquad (92)$$

where Z, β, and x_0 are time dependent parameters obeying a system of coupled ordinary differential equations [33]. Moreover, there are more general nonlinear Fokker-Planck equations with state-dependent diffusion that are also endowed with q-MaxEnt exact time dependent solutions [35]. By recourse to a redefinition of q, as well as an appropriate identification of the other parameters involved, it is easy to see that the exact solutions (92) constitute a particular instance of the MaxEnt solutions (72) (which in general are approximate but in this particular case are exact solutions). An important point is that the solutions (92) were obtained using the formulation of Tsallis nonextensive thermostatistics based on the unnormalized q-constraints (22). Within that frame, the Tsallis parameter is related to the power η in the diffusion term by the relation $\eta = 2 - q$. Within the present reformulation that connection adopts a different form. Let us denote the power appearing in the exact solution (92) by ξ. Then, the power in the solution and the power in the diffusion term are related by

$$\xi = \frac{1}{1+\eta}. \tag{93}$$

The above relation characterizes the particular exact solution to the nonlinear Fokker-Planck equation and does not depend on the particular formulation of Tsallis formalism used to obtain it. Within the present "P-picture" formulation of Tsallis thermostatistics with escort constraints we have $\xi = q/(1+q)$ and consequently,

$$q = \frac{1}{\eta}, \tag{94}$$

which is now the relation between the nonlinearity in the Fokker-Planck equation and Tsallis parameter.

6 Tsallis Nonextensive Thermostatistics and the Vlasov–Poisson Equations

The first physical application of Tsallis entropy was concerned with stationary maxent solutions of the Vlasov-Poisson equations for self-gravitational systems [23]. It is well known that the standard Boltz-mann-Gibbs thermostatistics is unable to provide a useful description of this type of systems [4,6–8]. Besides, the connection between Tsallis entropy and gravitation constituted the first clue for the characterization of the kind of problems where the generalized formalism might be useful.

Although the standard Boltzmann-Gibbs formalism (associated with the value $q = 1$ of Tsallis parameter) has serious difficulties in dealing with gravitation in three dimensional space, it shows no problems in the case of one or two dimensions. On the other hand, it is possible to obtain (at least in the Vlasov approximation) physically acceptable answers in $D = 3$, if we employ an appropriate value of $q < 7/9$. These facts suggest that dimensionality plays an important role in this problem. The aim of the present section is to discuss some aspects of the application of the Tsallis formalism to the Valsov-Poisson equations. The Tsallis nonextensive approach to the Vlasov-Poisson equations will be considered for the general case of D spatial dimensions, in order to provide an illustrative example of a problem where Tsallis parameter q depends on the spatial dimentionality. Similar calculations using Tsallis formalism with unnormalized q-constraints can be found in [77].

6.1 Long-Range Interactions and Nonextensivity

The thermodynamics of self-gravitating systems is characterized by the fact that the total energy, as well as other variables usually regarded as additive quantities, lose their extensive character. This suggests that a non-additive entropic measure, like the one advanced by Tsallis, may be more approrpiate for dealing with gravitation than the standard Boltzamnn-Gibbs entropy. The non-extensive behaviour exhibited by self-gravitating systems is due to the long range character of the gravitational interaction. The early applications of Tsallis' Thermostatistics to gravitational systems stimulated the exploration, within Tsallis formalism, of other many body problems showing non-extensive effects due to the presence of long range interaction [78–81]. These studies considered systems with Lennard-Jones like interparticle potentials, exhibiting a repulsive behaviour at short distances together with an attractive long range interaction falling as $r^{-\alpha}$. This kind of potentials, while being useful in order to illuminate the thermodynamical implications of long range forces, do differ in an essential way from the gravitational interaction. The Gibbs canonical ensemble for a system of a N particles interacting via a Lennard-Jones like potential, enclosed within a box of volume V, is well defined and has a convergent partition function. On the other hand, and due to *the singularity at the origin* of the gravitational potential, the Gibbs canonical ensemble for a system of N gravitationally interacting particles (even if they are enclosed within a finite box) has a *divergent partition function* and thus is not well defined. The non-extensive properties of the Lennard-Jones like gases are more apparent when we try to define the $N \to \infty$ thermodynamic limit. The main results obtained so far in connection with this issue deal with the exotic scaling laws that are to be employed in orther to define the thermodynamic limit in a sensible way. Cogent evidence has already been obtained, showing that the correct scaling (with particle number) of the thermodynamic variables such as internal and free energy, is the one proposed by Tsallis, that goes with NN^*, where

$$N^* = \frac{N^{(1-(D/\alpha))}}{1-(D/\alpha)}, \tag{95}$$

D being the spatial dimension. Unfortunatelly, this scaling with the number of particles is not applicable to the strict gravitational case, since the canonical ensemble (for a given finite number of particles) is meaningless. There is another important feature of the thermodynamics of systems endowed with long range interactions that has not yet been fully addressed by the simplified models discussed in [78–81]. This property is *the breaking of the translational symmetry of space implicitly assumed in standard thermodynamics* [82]. This property of systems with long range interactions is illustrated by the polytropic model discussed in [23] and reviewed here.

6.2 The Vlasov–Poisson Equations

The Vlasov-Poisson equations constitute important an widely used mathematical tools for the study of the structure and evolution of astrophysical self-gravitating N-body systems like galaxies and galaxy clusters [4,7]. Let us consider a system of N identical stars of mass m. Vlasov equation reads [4]

$$\frac{\partial F}{\partial t} + \mathbf{v} \cdot \frac{\partial F}{\partial \mathbf{x}} - \frac{\partial \phi}{\partial \mathbf{x}} \cdot \frac{\partial F}{\partial \mathbf{v}} = 0, \tag{96}$$

where $F(\mathbf{x}, \mathbf{v}) d\mathbf{x} d\mathbf{v}$ denotes the number of stars in the $2D$-dimension-al volume element $d\mathbf{x} d\mathbf{v}$ in position-velocity space, and the gravitational potential $\phi(\mathbf{x})$ is given by the D-dimensional Poisson equation

$$\nabla^2 \phi = D V_1 G \rho. \tag{97}$$

In the last equation G is the universal gravitational constant, V_1 stands for the volume of a D-dimensional unitary sphere, and the mass density ρ is given by

$$\rho(\mathbf{x}) = m \int F(\mathbf{x}, \mathbf{v}) \, d\mathbf{v}. \tag{98}$$

The Vlasov equation provides only an approximate description of the dynamics of a gravitational N-body system. Within this approximation, each particle of the system is assumed to move under the influence of a smooth gravitational field generated by the average bulk distribution of mater of the N-body system. This means that the effect of close encounters between particles are neglected. This is regarded as a very good approximation for galaxies [4]. However, it is not always applicable in stellar dynamics. For example, it is not a good approximation for globular clusters [4].

In the particular case of a central potential $\phi(r)$ depending only on the radial coordinate r, the Laplacian operator adopts the form [83],

$$\nabla^2 \phi = \frac{1}{r^{D-1}} \frac{d}{dr} \left\{ r^{D-1} \frac{d\phi}{dr} \right\}. \tag{99}$$

It is sometimes useful to write Vlasov equation under the guise

$$\left(\frac{dF}{dt} \right)_{orbit} = 0, \tag{100}$$

where the total time derivative is evaluated along the orbit of any individual star moving in the potential ϕ. In the case of stationary solutions (i.e. $\partial F / \partial t = 0$), equation (100) implies that $F(\mathbf{x}, \mathbf{v})$ depends on the coordinates and the velocity components only through integration constants C_i of the motion in the potential ϕ. This result,

$$F_{est.} = F(C_1, \ldots C_L), \tag{101}$$

constitutes the well known Jeans' Theorem in Stellar Dynamics [4].

The solutions describing spherically symmetric systems with iso-tropic velocity distributions depend only on the energy, and it is convenient to write them in the form

$$F = F(\epsilon), \tag{102}$$

where

$$\epsilon = \Phi(\mathbf{x}) - \frac{1}{2} v^2, \tag{103}$$

and

$$\Phi = \phi_0 - \phi. \tag{104}$$

The quantities ϵ and Φ are usually called the relative energy and potential (both per unit mass), respectively [4]. The constant ϕ_0 is chosen in such a way that the relative potential vanishes at the boundary of the system.

6.3 MaxEnt Stationary Solutions to the Vlasov–Poisson Equations

Present day galaxies are described by stationary solutions to the Vlasov-Poisson equations. These equilibrium configurations are usually assumed to be the result of appropriate relaxation processes such as the ones denoted by the names "phase mixing" and "violent relaxation" [4]. These kind of processes are characterized by a loss of memory about the initial conditions of the system. It has been argued that the final equilibrium state should be determined from a MaxEnt variational principle [6]. The maximization of Boltzmann' logarithmic entropy under the constraints imposed by the conservation of the total mass and energy leads to the stellar isothermal sphere distribution, which in $D = 3$ looks like

$$F(\epsilon) = A\,(2\pi\sigma^2)^{-3/2}\,\exp(\epsilon/\sigma^2), \tag{105}$$

where σ stands for the velocity dispersion and A is an appropriate normalization constant. Unfortunately, the isothermal sphere is characterized by an *infinite* total mass [6]. The real meaning of this unphysical result is that the posed variational problem does not have a solution. For any given star distribution $F(\mathbf{x}, \mathbf{v})$, it is always possible to obtain a new distribution $F^*(\mathbf{x}, \mathbf{v})$ with the same mass and energy but showing a larger Boltzmann entropy [4]. For given values of total mass and energy, entropy is not bounded from above. An illustration of this feature of self-gravitating systems is provided by the so-called "red giant structure". This configuration is characterized by a high density inner core surrounded by a diluted extended "atmosphere". The core accounts for almost all the energy of the system. The outer envelope makes the main contribution to the total entropy. By increasing the concentration of the core while simultaneously expanding the "atmosphere", the entropy can be raised as much as wanted without changing the total energy. [4].

Summing up, a MaxEnt approach based on the Gibbs-Boltzmann entropic measure seems to be unable to characterize relaxed self-gravitating systems. This fact has motivated the exploration of alternative entropy functionals [8] of the form

$$S = \int C[F]\,d\mathbf{x}\,d\mathbf{v}, \tag{106}$$

$C[F]$ being a convex function that vanishes for $F = 0$. Here enters Tsallis generalized entropy S_q. It has been shown that adopting appropriate values for Tsallis parameter q, the extremization of Tsallis functional

$$S_q = \frac{1}{q-1}\left(1 - \int f^q d\mathbf{x} d\mathbf{v}\right) \qquad (107)$$

under the constraints imposed by the total mass and energy leads to sensible stellar distributions. Indeed, the so-called stellar polytropes are obtained [23]. These distributions are widely employed in the modelization of astrophysical objects like galaxies [4], although their relationship with Tsallis entropy was not known. The mass density of a stellar polytrope behaves in the same fashion as the density of a self-gravitating sphere constituted by a gas with a polytropic equation of state,

$$p = \text{const.} \times \rho^{\gamma}. \qquad (108)$$

The exponent γ in the equation of state is related to the polytropic index n by

$$\gamma = 1 + \frac{1}{n}. \qquad (109)$$

The gravitational behaviour of polytropic gasses is very important in theoretical astrophysics. These kind of systems was already introduced by Lord Kelvin in 1862 [84] in considerations relating to the convective equilibrium of fluids under the influence of gravity. A detailed account of the mathematical properties of polytropic and isothermal gas spheres has been given by Chandrasekhar [85]

Stellar polytropes constitute the most simple, but still physically acceptable, models for equilibrium stellar systems. In the earlier days of stellar dynamics, polytropes where considered as possible realistic models for galaxies and globular clusters. Nowadays it is known that stellar polytropes do not fit properly the corresponding observational data. However, they still play an important role in theoretical astrophysics, and are widely employed as a first theoretical approach, both in numerical simulations and analytical studies, for the description of stellar systems. From our point of view, they are relevant because they illustrate how a generalized thermostatistical formalism, based on a non-extensive entropy, is able to deal in a physically sensible way with self-gravitating systems. Moreover, the present formalism is relevant in connection with the empirical verifications of Tsallis theory that we have up to now. Bogoshian's application of Tsallis entropy to the pure electron plasma [25] *is based essentially on the same formalism as the one developed for the discussion of stellar polytropes.* Besides, it was recently shown [86] that a Tsallis distribution fits the observed velocity distribution of galaxy clusters in a much better way than other models based on the standard entropy. The success of Tsallis thermostatistics within this context is presumably due to the long-range gravitational forces involved. And, so far, the stellar polytropes are the only detailed theoretical model establishing a link between Tsallis distributions and N-body self-gravitational systems. For all these reasons we believe that a further study of the relationship between stellar polytropes and Tsallis statistics is worthwhile, and will contribute to the understanding of non-extensive systems.

6.4 Tsallis MaxEnt Solutions to the Vlasov–Poisson Equations

Here we are going to work within the "P-picture" formulation of Tsallis nonextensive formalism with escort constraints (see section III). We will extremize the Tsallis entropy of the star distribution function $F(\mathbf{x},\mathbf{v})$,

$$\tilde{S}_q[F] = \frac{1}{q-1}\left[1 - \left(\int F^{1/q}\, d\mathbf{x}\, d\mathbf{v}\right)^{-q}\right], \tag{110}$$

under the constraints imposed by the total mass

$$\frac{M}{m} = \int F(\mathbf{x},\mathbf{v})\, d\mathbf{x}\, d\mathbf{v}, \tag{111}$$

and the total energy E,

$$\frac{E}{m} = \frac{1}{2}\int F(\mathbf{x},\mathbf{v})\,\big(v^2 + \phi(\mathbf{x})\big)\; d\mathbf{x}\, d\mathbf{v}. \tag{112}$$

In order to avoid some misunderstandings that arose in the literature in connection with this last formula [25], we must stress that the above expression deals with the system's gravitational potential in a self consistent way. The potential ϕ appearing in (112) is generated by the mass distribution ρ associated with the distribution F itself. We are not considering a given external potential. In such a case, the appropriate expression would be

$$\frac{E^*}{m} = \int F(\mathbf{x},\mathbf{v})\left(\frac{1}{2}v^2 + \phi(\mathbf{x})\right) d\mathbf{x}\, d\mathbf{v}. \tag{113}$$

Introducing now appropriate Lagrange multipliers λ and β, our variational problem can be put in the form

$$\delta\left(\tilde{S}_q - \beta\frac{E}{m} - \lambda\frac{M}{m}\right) = 0, \tag{114}$$

yielding the Tsallis MaxEnt solution

$$\begin{aligned} F_{ME} &= \left[\int F_{ME}^{1/q} d\mathbf{x} d\mathbf{v}\right]^{\frac{(1+q)q}{1-q}} \left[(q-1)\lambda - (1-q)\beta e\right]^{\frac{q}{1-q}} \\ &= const \times (\phi_0 - e)^{\frac{q}{1-q}}\,, \end{aligned} \tag{115}$$

where

$$e = \frac{1}{2}v^2 + \phi \tag{116}$$

is the energy (per unit mass) of an individual star and

$$\phi_0 = -\lambda/\beta \tag{117}$$

Thus, the MaxEnt distribution adopts the form of a stellar polytrope,

$$F_{ME} = B\epsilon^{\alpha} \tag{118}$$

with

$$\alpha = \frac{q}{1-q}, \tag{119}$$

B being a constant.

The cut-off condition on the generalized MaxEnt distribution was originally introduced by Tsallis in a somewhat ad-hoc way. However, within the present application of Tsallis generalized thermostatistics, the cut-off prescription admits of a clear physical interpretation. Within the polytrope stellar distributions, the Tsallis cut-off corresponds to the escape velocity from the system [4].

The mass density distribution associated with the politropic distribution is

$$\rho(\mathbf{x}) = DV_1 B \int_0^{\sqrt{2\Phi}} \left(\Phi - \frac{1}{2}v^2\right)^{\alpha} v^{D-1}\, dv \tag{120}$$

This expression can be simplified making the change of variables

$$v^2 = 2\Phi \cos^2\theta, \tag{121}$$

wich yields

$$\rho = 2^{D/2} DV_1 B \Phi^{\alpha+D/2} \int_0^{\pi/2} \sin^{2\alpha+1}\theta\, \cos^{D-1}\theta d\theta. \tag{122}$$

Introducing now the polytropic index

$$n = \alpha + \frac{D}{2}, \tag{123}$$

we obtain

$$\rho = 2^{D/2} DV_1 B \Phi^{n} \int_0^{\pi/2} \sin^{2n-D+1}\theta\, \cos^{D-1}\theta d\theta. \tag{124}$$

This last equation can be recast in the simpler form

$$\rho = C\Phi^n, \tag{125}$$

where C is now a constant given by

$$C = 2^{D/2} DV_1 B \int_0^{\pi/2} \sin^{2n-D+1}\theta\, \cos^{D-1}\theta d\theta. \tag{126}$$

Inserting this expression in the D-dimensional Poisson equation (97) we obtain a non-linear ordinary differential equation for Φ (recall that $\Phi = \phi_0 - \phi$) ,

$$\frac{1}{r^{D-1}}\frac{d}{dr}\left[r^{D-1}\frac{d\Phi}{dr}\right] + DV_1GC\Phi^n = 0. \tag{127}$$

The expression for the mass density ρ also allows us to obtain the radial derivative of the total mass $M(r)$ contained within radious r,

$$\frac{dM}{dr} = DV_1r^{D-1}\rho = DV_1r^{D-1}C\Phi^n. \tag{128}$$

Inserting this last result in the differential equation verified by Φ, it is easy to see that

$$\left[r^{D-1}\frac{d\Phi}{dr}\right] + GM(r) = const, \tag{129}$$

which evaluated in $r = 0$ leads to

$$\left[r^{D-1}\frac{d\Phi}{dr}\right] + GM(r) = 0. \tag{130}$$

From this last equation it follows that $\Phi \sim r^{2-D}$ is the limiting asimptotic behaviour still yielding a *finite* total mass.

6.5 *D*-Dimensional Schuster Spheres

In terms of the dimensionless variables

$$h = \Phi/\Phi_0, \tag{131}$$

and

$$y = [DV_1GC\Phi_0^{n-1}]^{1/2}\, r, \tag{132}$$

the differential equation for Φ reads

$$\frac{1}{y^{D-1}}\frac{d}{dy}\left[y^{D-1}\frac{dh}{dy}\right] + h^n = 0. \tag{133}$$

If we adopt the polytropic index

$$n = \frac{D+2}{D-2}, \tag{134}$$

the above equation admits the solution

$$h(y) = \left[1 + \frac{y^2}{D(D-2)}\right]^{\frac{2-D}{2}}, \tag{135}$$

which, in term of the physical variables reads,

$$\Phi = \Phi_0 \left[1 + \frac{j^2 r^2}{D(D-2)} \right]^{\frac{2-D}{2}}, \tag{136}$$

where

$$j = [D V_1 G C \Phi_0^{n-1}]^{1/2}. \tag{137}$$

From equation (130) it follows that the D-dimensional Schuster Sphere shows the limiting asimptotic behaviour for a finite total mass. Although these solutions have infinite spacial extent, the density falls rapidly enough so as to have a finite total mass. Furthermore, they show, within the larger set of Tsallis MaxEnt solutions, the limiting behaviour between sensible solutions with finite mass and unphysical solutions with divergent mass.

The exponent α appearing in the stellar polytrope distribution is now

$$\alpha = n - \frac{D}{2} = \frac{8 - (D-2)^2}{2(D-2)}. \tag{138}$$

Notice that for a space dimension $D > D_l$, where

$$D_l = 2 + 2\sqrt{2}, \tag{139}$$

the D-dimensional Schuster polytrope is a monotonic *increasing* function of the energy.

Since the exponent α of a polytropic distribution is related to q by the equation $\alpha = q/(1-q)$, it is easy to verify that the Tsallis parameter characterizing the D-dimensional Schuster distributions is given by

$$q = \frac{8 - (D-2)^2}{8 - (D-2)^2 + 2(D-2)}. \tag{140}$$

It is interesting to notice that in the limit $D \to 2$, we obtain $q \to 1$. This is consistent with the known fact that the standard Boltzmann-Gibbs thermostatistics is able to deal in a physically acceptable way with self-gravitating systems with spatial dimension less than 2 [7,87]. In the case of $D = 3$, we recover the known value $q = 7/9$ [25] corresponding to the limit value of q providing a polytropic distribution with finit total mass.

6.6 Tsallis MaxEnt Time-Dependent Solutions

By recourse to a Tsallis MaxEnt scheme similar to the one discussed in section V for the nonlinear Fokker-Plack equation, it is also possible to obtain time dependent approximate solutions to the Vlasov-Poisson equations [24]. This approach, first developed in [24] is again based on the idea of following the time evolution of the mean values of a small set of M relevant dynamical quantities $A_i(\mathbf{x}, \mathbf{v})$,

$$\langle A_i \rangle = \frac{1}{h} \int F(\mathbf{x}, \mathbf{v})\, A_i(\mathbf{x}, \mathbf{v})\, d\mathbf{x}\, d\mathbf{v} \qquad (i = 1, \ldots, M), \tag{141}$$

where h is just a constant with the dimensions of the element of volume in the pertinent $(\mathbf{x}, \mathbf{v})$ space. The q-MaxEnt scheme developed in [24] was based on the unnormalized q-constraints. However, it can be reformulated in terms of normalized q-mean values [88]. If the distribution function $F(\mathbf{x}, \mathbf{v})$ evolves according to the Vlasov equation the time derivative of the relevant mean values is given by

$$\frac{d}{dt}\langle A_i\rangle = \frac{1}{h}\int F(\mathbf{x}, \mathbf{v})\,(\mathbf{v}\cdot\nabla_x A_i - \nabla\phi\cdot\nabla_v A_i)\,d\mathbf{x}\,d\mathbf{v} \qquad (i = 1, \ldots, M). \tag{142}$$

Unfortunatelly, this last set of equations does not constitute, in general, a closed set of M ordinary differential equations for the mean values $\langle A_i\rangle$. However, the MaxEnt approach allows us to build up, at each time t, a q-MaxEnt distribution [88]

$$\begin{aligned} F_{ME}(\mathbf{x}\mathbf{v}, t) &= \left[\int F_{ME}^{1/q}\,d\mathbf{x}\,d\mathbf{v}\right]^{\frac{(1+q)q}{1-q}} \\ &\times \left[(q-1)\lambda_0 - (1-q)\sum_{i=1}^{M}\lambda_i(t)A_i(\mathbf{x}, \mathbf{v})\right]^{q/(1-q)}, \end{aligned} \tag{143}$$

taking as constraints the instantaneous values adopted by the relevant moments $\langle A_i\rangle$ and introducing appropriate Lagrange multipliers λ_i $(i = 1, \ldots, M)$. The right hand sides of equations (142) can then be evaluated by recourse to this q-MaxEnt approximation. In such a manner a *closed* system of equations is formally obtained. The ensuing system, however, becomes highly non-linear. As a counterpart, the Vlasov equation, a partial differential equation, becomes now a system of ordinary differential equations for the evolution of the relevant mean values. In [24] it has been shown that such an approach is indeed useful because it does preserve some important properties of the *exact* evolution equation.

The exact equations of motion (142) corresponding to the M chosen relevant moments are verified by the Tsallis q-MaxEnt approximate solution $F_{ME}(\mathbf{x}, \mathbf{v}, t)$ by construction. The evolution of the M relevant mean values and their associated Lagrange multipliers is governed by a set of $2M$ ordinary differential equations that can be cast into a Hamiltonian form [24,88]. This means that the relevant mean values $\langle A_i\rangle$ and their concomitant Lagrange multipliers λ_i are canonicaly conjugate dynamical variables. This Hamiltonian structure implies that the associated flow in phase space does not have sinks, nor does it have sources. Moreover, the Tsallis entropy $\tilde{S}_q$ evaluated upon the MaxEnt approximate solution $F_{ME}(\mathbf{x}, \mathbf{v})$ is a constant of the motion. Consequently, the evolution of the system does not exhibit any relaxation process. This is an important characteristic of the exact solutions that is preserved by the q-MaxEnt approximations. All these general properties of the q-MaxEnt scheme do not depend on the dimensionality of the the (one-particle) configuration space. They hold true for any number of spatial dimensions. However, in the particular case of $q = -1$ and $D = 1$ or $D = 3$, the MaxEnt approximation generates, if the relevant mean values are properly chosen, *exact solutions to the Vlasov-Poisson equations* [24,88]. Indeed, various of the known exact solutions of the Valsov-Poisson equations constitute particular instances of our general q-MaxEnt approach [24,88].

It is worth remarking that these time dependent q-MaxEnt solutions of the Vlasov equation are not related in a direct fashion to the previously discussed stationary MaxEnt solutions. The total energy does not appear as a constraint within the time dependent scenario, as happens in the stationary case. Consequently, the $q = -1$ value of the Tsallis parameter yielding exact time dependent solutions should not be regarded as belonging to the "allowed" $q < 7/9$ range of q-values associated with the physially acceptable q-MaxEnt stationary solutions. These two special q-values, $7/9$ and -1, arise from quite different settings and seem to be completely unrelated.

7 Conclusions

We have discussed some aspects of the relationship between Tsallis' nonextensive thermostatistical formalism and Jaynes' information theory approach to statistical mechanics. We have seen that there are fundamental features of the maximum entropy principle that are universal and hold true for general thermostatistical formalism regardless of the particular form of the entropic functional used. Tsallis thermostatistics, however, is also endowed with important nonuniversal characteristics.

We have shown how Jaynes maxent approach to evolution equations can be generalized and implemented within Tsallis formalism. As an illustration we considered its application to the nonlinear Fokker-Planck equation. We have also considered MaxEnt Tsallis distributions in connection with stationary solutions of the D-dimensional Vlasov-Poisson equations. In particular, we provided an analitical relation between Tsallis parameter q and the dimension D of physical space for the special case of D-dimensional Schuster solutions.

Acknowledgements
Special words of gratitude are due to two of my coauthors in this field: Profs. C. Tsallis and A. Plastino. I wish also to acknowledge extremely fruitful discussions, along the years, with H. Miller, H. Vucetich, J. A. S. Lima, L. Borland, E. M. F. Curado, F. Pennini, C. Anteneodo, R. S. Mendes, W.-M. Zheng, A. K. Rajagopal, S. Abe, B. M. Boghosian, L. R. da Silva, F. Solms, D. Malaza, R. DiSisto, S. Martinez, A. R. R. Papa, E. K. Lenzi, R. Silva , and S. Curilef. This work was partially supported by the Argentine National Research Council (CONICET).

References

1. C. Tsallis, Physica A **221**, 277 (1995), and references therein.
2. C. Tsallis, Fractals **6**, 539 (1995), and references therein; an updated bibliography can be found in http://tsallis.cat.cbpf.br/biblio.htm
3. C. Tsallis, Physics World **10**, 42 (July 1997).
4. J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, 1987).
5. A. R. Plastino and A. Plastino, in *Condensed Matter Theories*, Volume **11**, ed. E. Ludeña (Nova Science Publishers, New York, 1996), p. 341.
6. D. Lynden-Bell, Mon. Not. R. Astron. Soc. **124**, 279 (1967).
7. W. C. Saslaw, *Gravitational Physics of Stellar and Galactic Systems*, (Cambrideg University Press, Cambridge, 1985).
8. S. Tremaine, M. Henon, and D. Lynden-Bell, Mon. Not. R. Astron. Soc. **219**, 285 (1986).
9. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
10. E. M. F. Curado and C. Tsallis, J. Phys. A **24**, L69 (1991); Corrigenda **24**, 3187 (1991) and **25**, 1019 (1992).
11. A. R. Plastino and A. Plastino, in *Condensed Matter Theories*, Volume **11**, ed. E. Ludeña (Nova Science Publishers, New York, 1996), p. 327.
12. D. H. Zanette and P. A. Alemany, Phys. Rev. Lett. **75**, 366 (1995); M. O. Caceres and C. E. Budde, ibid **77**, 2589 (1996); D. H. Zannette and P. A. Alemany, ibid **77**, 2590 (1996).
13. A. K. Rajagopal, Phys. Rev. Lett. **76**, 3469 (1996).
14. E. K. Lenzi, L. C. Malacarne, and R. S. Mendes, Phys. Rev. Lett. **80**, 218 (1998).
15. A. R. Plastino, A. Plastino, and C. Tsallis, J. Phys. A, **27**, 5707, (1994).
16. G. A. Raggio, J. Math. Phys. **36**, 4785 (1995).
17. A. M. Meson and F. Vericat, J. Math. Phys. **37**, 4480 (1996).
18. R. J. V. dos Santos, J. Math. Phys. **38**, 4104 (1997).
19. R. Silva Jr., A. R. Plastino, and J. A. S. Lima, Phys. Lett. A **249**, 401 (1998).
20. S. Abe, Eur. Phys. J. B **9**, 679 (1999).
21. A. R. Plastino and A. Plastino, Phys. Lett. **A177**, 177 (1993).
22. A. R. Plastino and C. Anteneodo, Ann. Phys. **255**, 250 (1997).
23. A. R. Plastino and A. Plastino, Phys. Lett. **A174**, 384 (1993); J. J. Aly, in *N-Body Problems and Gravitational Dynamics*, Proc. of the Meeting held at Aussois-France (21-25 March 1993), eds. F. Combes and E. Athanassoula (Publications de l'Observatoire de Paris, 1993).
24. A. R. Plastino and A. Plastino, Phys. Lett. A **193**, 251 (1994).
25. B. M. R. Boghosian, Phys. Rev. E **53**, 4754 (1995); C. Anteneodo and C. Tsallis, J. Mol. Liq. **71**, 255 (1997).
26. G. Kaniadakis, A. Lavagno, and P. Quarati, Phys. Lett. B **369**, 308 (1996); P. Quarati, A. Carbone, G. Gervino, G. Kaniadakis, A. Lavagno, and E. Miraldi, Nucl. Phys. A **621**, 345c (1996); G. Kaniadakis, A. Lavagno, M. Lissia, and P. Quarati, Physica A **261**, 359 (1998).
27. V. H. Hamity and D. E. Barraco, Phys. Rev. Lett. **76**, 4664 (1996); D. F. Torres, H. Vucetich, and A. Plastino, Phys. Rev. Lett. **79**, 1588 (1997).
28. C. Tsallis, A. R. Plastino and W.-M. Zheng, Chaos, Solitons and Fractals **8**, 885 (1997); M. L. Lyra and C. Tsallis, Phys. Rev. Lett. **80**, 53 (1998).
29. U. M. S. Costa, M. L. Lyra, A. R. Plastino, and C. Tsallis, Phys. Rev. E **56**, 245 (1997).

30. C. Anteneodo and C. Tsallis, Phys. Rev. Lett. **80**, 5313 (1998).
31. F. A. Tamarit, S. A. Cannas, and C. Tsallis, Eur. Phys. J. B **1**), 545 (1998); A. R. R. Papa and C. Tsallis, Phys. Rev. E **57**, 3923 (1998).
32. P. A. Alemany and D. H. Zanette, Phys. Rev. E **49**, 956 (1994); D. H. Zanette and P. A. Alemany, Phys. Rev. Lett. **75**, 366 (1995); M. O. Caceres and C. E. Budde, ibid **77**, 2589 (1996); D. H. Zannette and P. A. Alemany, ibid **77**, 2590 (1996); C. Tsallis, S. V. F. Levy, A. M. C. Souza and R. Maynard, Phys. Rev. Lett. **75**, 3589 (1995); Erratum **77**, 5442 (1996).
33. A. R. Plastino and A. Plastino, Physica A **222**, 347 (1995).
34. C. Tsallis and D. J. Bukman, Phys. Rev. E **54**, R2197 (1996).
35. L. Borland, F. Pennini, A. R. Plastino, and A. Plastino, Eur. Phys. J. B **12**, 285 (1999).
36. L. Borland, Phys. Rev E **57**, 6634 (1998).
37. D. A. Stariolo, Phys. Rev. **E 55**, 4806 (1997).
38. A. Compte and D. Jou, J. Phys. A **29**, 4321 (1996).
39. A. Compte, D. Jou, and Y. Katayama, J. Phys. A **30**, 1023 (1997).
40. A. M. Mariz, Phys. Lett. A **165**, 409 (1992).
41. J. D. Ramshaw, Phys. Lett. A **175**, 169 (1993).
42. L. Borland, A. R. Plastino, and C. Tsallis, J. Math. Phys. **39**, 6490 (1998); Erratum **40**, 2196 (1999).
43. J.W. Gibbs, *Elementary Principles in Statistical Mechanics* (Yale University Press, New Haven, 1902).
44. E.T. Jaynes, Phys. Rev. **106**, 620 (1957).
45. E.T. Jaynes, Phys. Rev. **108**, 171 (1957).
46. A. Katz, *Statistical Mechanics* (Freeman, San Francisco, 1967).
47. W. Grandy, *Foundations of Statistical Mechanics* (Reidel, Dordrecht, 1987).
48. R. Balian, *From Microphysics to Macrophysics* (Springer-Verlag, Berlin, 1992).
49. C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*, (Cambridge University Press, Cambridge, 1993).
50. J. Skilling and R. K. Bryan, Mon. Not. R. Astron. Soc. **211**, 111 (1984).
51. V. J. Cornwell and K. F. Evans, Astron. Astrophys. **143**, 77 (1985).
52. L. Borland, Physica D **99**, 175 (1996).
53. A. Plastino and A. R. Plastino, Phys. Lett. A **226**, 257 (1997).
54. R. S. Mendes, Physica A **242**, 299 (1997).
55. E. M. F. Curado, Braz. J. Phys. **29**, 36 (1999).
56. R. P. Di Sisto, S. Martinez, R. B. Orellana, A. R. Plastino, and A. Plastino, Physica A **265**, 590 (1999).
57. A. R. R. Papa, J. Phys. A: Math. Gen. **31**, 5271 (1998).
58. S. Abe, Phys. Lett. A **224**, 326 (1997).
59. E. P. Borges and I. Roditi, Phys. Lett. A **246**, 399 (1998).
60. C. Anteneodo and A.R. Plastino, J. Phys. A: Math. Gen. **32**, 1089 (1999).
61. C. Tsallis, R. S. Mendes and A. R. Plastino, Physica A **261**, 534 (1998).
62. A. R. Plastino and J. A. S. Lima, *Equipartition and Virial Theorems within General Thermostatistical Formalisms*, (1999) submitted for publication.
63. D. Otero, A. Plastino, A. N. Proto, and G. Zannoli, Phys. Rev. A **26**, 1209 (1982).
64. D. Otero, A. Plastino, A. N. Proto, and G. Zannoli, Z. Phys. A **316**, 323 (1984).
65. E. Duering, D. Otero, A. Plastino, and A. N. Proto, Phys. Rev. A **32**, 2455 (1985).
66. E. Duering, D. Otero, A. Plastino, and A. N. Proto, Phys. Rev. A **35**, 2314 (1985).
67. P. Hick and G. Stevens, Astron. Astrophys. **172**, 350 (1987).
68. A. R. Plastino and A. Plastino, Physica A **232**, 458 (1996).

69. A. R. Plastino, H. G. Miller, and A. Plastino, Phys. Rev. E **56**, 3927 (1997).
70. A. Plastino, A. R. Plastino, and H. G. Miller, Phys. Lett. A **232**, 349 (1997).
71. E. D. Malaza, H. G. Miller, A. R. Plastino, and F. Solms, Physica A **265**, 224 (1999).
72. M. Muskat, *The Flow of Homogeneous Fluids Through Porous Media* (McGraw-Hill, New York, 1937).
73. J. Buckmaster, J. Fluid Mech. **81**, 735 (1977).
74. H. Spohn, J. Phys. (France) I **3**, 69 (1993).
75. J. M. Carlson, E. R. Grannan, C. Singh, and G.H. Swindle, Phys. Rev. E **48** 688 (1993).
76. H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, New York, 1989).
77. A. R. Plastino and A. Plastino, Braz. J. Phys. **29**, 79 (1999).
78. L. S. Lucena, L. R. da Silva, and C. Tsallis, Phys. Rev. E **51**, 6247 (1995).
79. J. R. Grigera, Phys. Lett. A **217**, 47 (1996).
80. P. Jund, S.G. Kim, and C. Tsallis, Phys. Rev. B **52**, 50 (1995).
81. S. Curilef and C. Tsallis, *Critical behaviour in Lennard-Jones like fluids: Long-range interactions and nonextensivity*, preprint (1998).
82. J. D. Brown and J. W. York Jr., in *Physiscal Origins of Time Assymetry*, eds. J.J. Halliwell, J. Perez-Mercader, and W.H. Zurek (Cambrideg University Press, Cambridge, 1996), p. 465-474.
83. P. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
84. W. Thomson (Lord Kelvin), in: *Mathematical and Physical Papers*, Vol.3 (Cambridge University Press, Cambridge, 1911), p. 255-360.
85. S. Chandrasekhar, *An Introduction to the Theory of Stellar Structure* (University of Chicago Press, Chicago, 1958).
86. A. Lavagno, G. Kaniadakis, M. Rego-Monteiro, P. Quarati, and C. Tsallis, Astroph. Lett. Comm. **35**, 449 (1998).
87. G. Rybicki, Astrophys. and Sapce Sci. **14**, 56 (1971).
88. A. R. Plastino and A. Plastino, unpublished.

IV. Computational Methods for the Simulation of Classical and Quantum Many Body Systems Arising from Nonextensive Thermostatistics

I. Andricioaei and J.E. Straub

Department of Chemistry, Boston University, Boston, MA 02215, USA

Abstract. This chapter presents a variety of computational methods that make use of essential features of the non-extensive thermostatistics proposed by Tsallis. The basic properties of the non-extensive thermostatistics are derived and discussed in the context of the law of mass action and reaction kinetics. Those results are used to motivate and derive Monte Carlo and Molecular Dynamics algorithms which isolate low lying energy minima or effectively sample the extensive or non-extensive thermostatistical distributions. Algorithmic protocols are defined for cases of equilibrium sampling and simulated annealing for both classical and quantum simulations of many body systems. A practical method for the optimization of the resulting simulation algorithms is described. A method for global optimization based on the simulated annealing Cauchy-Lorentz "density packets" is also presented. Applications to a variety of problems related to simulated annealing and enhanced sampling in the context of model, spin, atomic and biomolecular systems are discussed.

1 Background and Focus

The computer simulation of complex systems such as biomolecules, liquids and glasses is often made difficult due to the ruggedness of the underlying "energy landscape." In a Monte Carlo or Molecular Dynamics simulation, one hopes to generate the equilibrium Boltzmann distribution of states [1–3]. However, for many systems of interest, high energetic barriers separate basins on the potential surface as depicted in Fig. 1. This makes it difficult for a Monte Carlo or Molecular Dynamics trajectory to move between basins and sample the equilibrium distribution of states. This problem is referred to as "broken ergodicity" [4].

To compute an equilibrium thermodynamic average of a given property one must average over the statistically relevant configurations of the system. For a system of N atoms in a d-dimensional space this amounts to computing the partition function by evaluating a $2dN$ dimensional phase space integral. Any point in configuration space (excluding maxima, saddle points and ridges) will be mapped by a steepest descent to a minimum on the potential energy surface. The set of all points that map to a given minimum form the basin of attraction. Therefore, the configuration space can be decomposed into a set of basins of attraction and the integral over phase space can be expressed as a sum of integrals each over a given basin of attraction. An accurate calculation of the phase

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 193–233, 2001.

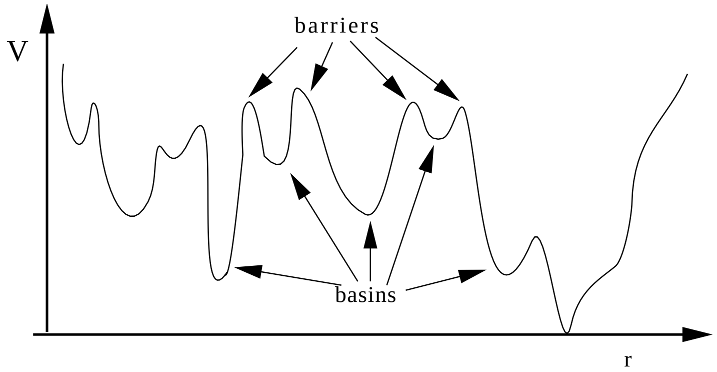

Fig. 1. Schematic showing a rugged energy landscape with multiple basins separated by barriers of various heights. The energy landscapes of liquid, glass and biomolecular systems are often characterized by significant numbers of thermodynamically important basins separated by barriers on many energy scales.

space integral will involve a sum over the local *intrabasin* averages using the correct relative *interbasin* statistical weights. An effective computational method for evaluating such integrals must couple *local* sampling of the basins together with a convenient means of *global* transportation between basins.

For complex systems these basins are numerous. Moreover, they are connected by energy barriers that can range in size from the minuscule to the enormous on the thermal energy scale. Depending on the details of the connectivity of such minima, one may find thermostatistically important regions of configuration space that are separated by large energy barriers such that transitions between those regions – using standard Monte Carlo or Molecular Dynamics simulation methods – occur on a time scale that is very long compared with the simulation time scale. In such cases, it is essential to employ simulation methods that can effectively sample all statistically relevant regions in the available simulation time.

Overcoming broken ergodicity is a non-trivial problem that can be approached in two ways. One way is to exploit specific features of a system to develop a *tailored* method. For example, it may be that the energy minima are regularly distributed in space and can easily be found by symmetry considerations. Such tailored methods may be effective for a given system but inapplicable to other problems. A second approach is to develop a *general* method designed to be effective for a class of systems. While the development of general methods is more challenging, the potential impact is also much greater.

The development of such a method is difficult for complex systems that are characterized by many minima separated by a broad distribution of length scales. For such systems, random long range moves will typically lead to high energy regions of the potential surface which are usually rejected. Small steps lead to slow, activated barrier crossing which is an inefficient means of moving between thermodynamically important basins. The key to an effective algorithm is to identify the optimal balance of the time spent moving between thermodynamically important basin regions of configuration space and the time spent moving between basins in thermodynamically unimportant (high energy barrier) regions of the potential.

In this chapter we focus on computational methods that employ the special features of the non-extensive thermostatistics proposed by Tsallis. As we will see, special features of the non-extensive statistical mechanics lead to accelerated sampling of configurational space that can effectively address the problem of broken ergodicity in complex systems. This is accomplished by accelerating the rate of energy barrier crossing and "smoothing over" phase transitions that lead to critical slowing down. Moreover, extra "parameters" are needed to define the relevant ensemble in the non-extensive thermostatistics. By optimizing these additional variables, a computational method based on the non-extensive thermostatistical distributions may be tailored to a specific problem of interest in a way that cannot be accomplished for general methods based on the standard extensive thermostatistics.

In this chapter we begin with a concise summary of essential features of the non-extensive thermostatistics. We then explore the nature of equilibria and kinetics in the non-extensive thermostatistics. Those results motivate the development of algorithms for enhanced simulated annealing, Monte Carlo averaging, and Molecular Dynamics that are discussed in detail. An effective technique for the optimization of the proposed computational methods is described. The closing sections explore the relationship between the non-extensive statistics and path integral representations of the density matrix and a novel method for the simulated annealing of the non-extensive classical density distribution.

2 Basic Properties of Tsallis Statistics

The background that we need to develop the simulation methods described in this chapter starts with the seminal work of Tsallis describing a possible generalization of the standard Gibbs-Boltzmann statistical mechanics [5]. While this material is developed in detail in other chapters, it is summarized here for reasons of completeness and the need to define the particular version of the theory that is employed.

Tsallis began with a reëxpression of the Gibbs-Shannon entropy formula

$$S = \lim_{q\to 1} S_q = \lim_{q\to 1} \frac{k}{q-1} \int p_q(\mathbf{\Gamma})(1 - [p_q(\mathbf{\Gamma})]^{q-1})d\mathbf{\Gamma} \tag{1}$$

$$= -k \int p(\mathbf{\Gamma}) \ln p(\mathbf{\Gamma}) d\mathbf{\Gamma} \tag{2}$$

where $d\mathbf{\Gamma} = d\mathbf{r}^N d\mathbf{p}^N / h^{dN}$ is a phase space increment. On the right hand side of this expression, the identity $\ln x = \lim_{n\to 0}(x^n - 1)/n$ has been used to transform the logarithm and q is a real number [5,6]. A similar result was previously presented in the context of generalized information entropies but had apparently not been applied to describe the physical world [7]. Tsallis's bold move was to do exactly that and the result has been a productive reëxamination of the central dogma of thermostatistics.

2.1 From the Thermal Density Distribution to Tsallis Statistics

For the purposes of this chapter, it is more useful to develop the non-extensive thermostatistics from a reëxpression of the classical density distribution $\exp(-\beta\mathcal{H})$. Rather than writing

$$e^{-\beta\mathcal{H}} = \left(e^{-\beta\mathcal{H}/P}\right)^P , \tag{3}$$

as is commonly done in discussions of path integral representations of the density distribution [8], suppose that we express the exponential as a limit

$$e^{-\beta\mathcal{H}} = \lim_{P\to\infty} \left(\frac{1}{1+\beta\mathcal{H}/P}\right)^P . \tag{4}$$

Now suppose that discard the limit

$$\left(\frac{1}{1+\beta\mathcal{H}/P}\right)^P \tag{5}$$

and consider the argument for arbitrary P. If we substitute

$$P = \frac{1}{1-q} \tag{6}$$

we find that $P = 1, 2, 3, 4 \ldots \infty$ becomes $q = 0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4} \ldots 1$ and

$$e^{-\beta\mathcal{H}} \to (1-(q-1)\beta\mathcal{H})^{\frac{1}{q-1}} . \tag{7}$$

The right hand side of this expression is the Tsallis thermostatistical density distribution that was originally derived by extremizing the "entropy" S_q subject to the constraints that the distribution is normalized and that the average energy is computed in the *standard way.*

Now suppose that we instead define

$$P = \frac{1}{q-1} \tag{8}$$

so that $P = 1, 2, 3, 4 \ldots \infty$ becomes $q = 2, \frac{3}{2}, \frac{4}{3}, \frac{5}{4} \ldots 1$. The resulting distribution is

$$e^{-\beta\mathcal{H}} \to (1-(1-q)\beta\mathcal{H})^{\frac{1}{1-q}} . \tag{9}$$

The right hand side of the expression is precisely the Tsallis thermostatistical density distribution $p_q(\mathbf{\Gamma})$ derived by extremizing S_q subject to the constraints that the distribution is normalized and the average energy is defined in terms of the *"q-expectation" value.*

How can we interpret these results? Tsallis showed how a generalized statistics can originate from the Gibbs entropy formula and the identity $p \ln p = \lim_{n \to \infty} \frac{1}{n} p(p^n - 1)$. He then stripped away the limit and interpreted the kernel where $n = q - 1$

$$p \ln p \to \frac{1}{q-1} p(p^{q-1} - 1). \tag{10}$$

However, it is possible to reach the same distribution from a very different starting point – the equilibrium density distribution – by rewriting the density distribution using the identity $\exp(-h) = \lim_{P \to \infty}(1 + h/P)^{-P}$ and removing the limit to interpret the kernel where $P = 1/(q-1)$

$$e^{-\beta \mathcal{H}} \to (1 - (1-q)\beta \mathcal{H})^{\frac{1}{1-q}} \tag{11}$$

in the spirit of Tsallis statistics.

For the later expression, derived from Eq. (10) using the constraint based on the q-expectation value, when $P = 1$ we have what will prove to be the interesting case of $q = 2$. In the limit that $P = \infty$, we recover the Gibbs-Boltzmann statistical distribution. Intermediate values of P provide cases in between these limits.

2.2 Specific Defining Properties of the Tsallis Statistical Distributions

Tsallis noted a number of properties of S_q, which he referred to as a "generalized entropy," and the associated statistical distributions. He found that much of the standard mathematical structure of Gibbs-Boltzmann statistical mechanics is preserved. This is interesting in itself. However, even more interesting was what is not preserved. Thermodynamic state functions such as the entropy and energy were no longer extensive functions of the system. This prompted the use of a generalized formalism based on the non-additive function S_q to again derive, for non-extensive systems, a variety of results of the standard statistical mechanics (see [9] and references therein).

As is discussed elsewhere in this book, at least three "conventional" versions of the theory of non-extensive Tsallis thermostatistics exist. The following discussion is presented to familiarize the reader with the particular incarnation of the non-extensive Tsallis statistical distributions we will employ. The properties described are important for the development of the computational methods summarized in this chapter. However, please note that many of the results relating to the equipartitioning of energy, the treatment of ideal systems, and the normalization of statistical averages in the non-extensive thermostatistics have been developed elsewhere in a slightly different manner in the different context of the description of physical systems using the non-extensive thermostatistics [9].

Defining the Temperature When $q \neq 1$ the relative probability of two points in phase space depends on the choice of the zero of energy [5]. By defining

$$\bar{\mathcal{H}}(\mathbf{\Gamma}) = \frac{1}{\beta(q-1)} \ln\left[1 - (1-q)\beta\mathcal{H}(\mathbf{\Gamma})\right], \tag{12}$$

the probability of being at a point $\mathbf{\Gamma}$ in phase space can be written in the familiar form

$$p(\mathbf{\Gamma}) = \frac{1}{Zh^{dN}} \exp(-\beta\bar{\mathcal{H}}(\mathbf{\Gamma})), \tag{13}$$

where Z is the partition function

$$Z = \frac{1}{h^{dN}} \int \exp(-\beta\bar{\mathcal{H}}(\mathbf{\Gamma}))d\mathbf{\Gamma}. \tag{14}$$

However, for a constant potential shift ϵ the relative probability

$$\frac{p_q(\mathbf{\Gamma}_{\text{new}})}{p_q(\mathbf{\Gamma}_{\text{old}})} = \left[\frac{1 - (1-q)\beta(\mathcal{H}(\mathbf{\Gamma}_{\text{new}}) + \epsilon)}{1 - (1-q)\beta(\mathcal{H}(\mathbf{\Gamma}_{\text{old}}) + \epsilon)}\right]^{\frac{1}{1-q}} \tag{15}$$

depends on ϵ. This ratio can be rewritten as

$$\frac{p_q(\mathbf{\Gamma}_{\text{new}})}{p_q(\mathbf{\Gamma}_{\text{old}})} = \left[\frac{1 - (1-q)\beta'\mathcal{H}(\mathbf{\Gamma}_{\text{new}})}{1 - (1-q)\beta'\mathcal{H}(\mathbf{\Gamma}_{\text{old}})}\right]^{\frac{1}{1-q}}, \tag{16}$$

where potential energy shift ϵ has been absorbed in an effective "temperature"

$$\frac{1}{\beta'} = \frac{1}{\beta} + (q-1)\epsilon. \tag{17}$$

In the $q = 1$ limit, the effective temperature equals the standard temperature. Otherwise, adding a constant shift to the potential energy is equivalent to rescaling the temperature at which the canonical probability distribution is computed.

Weight Dependence on the Zero of Energy In the $q \neq 1$ regime, for certain points in phase space the probability $p_q(\mathbf{\Gamma})$ may be negative or even imaginary. We then say that the probability of the system accessing that point in phase space is zero. This may be the case even when the energy is finite. Of course, the exact range of phase space over which the probability is set to zero will also depend on the choice of the zero of energy. In our computational applications, the zero of energy is typically chosen such that all of configuration space is accessible (represented by a real, non-negative probability).

Normalizing q-Expectation Values Equilibrium average properties are calculated using a statistical weighting of the probability $p_q(\mathbf{\Gamma})$ of Eq. (13) raised to the power of q as required by the generalized statistical mechanics. The so-called q-expectation value is written

$$\langle \ldots \rangle_q = \int [p_q(\mathbf{\Gamma})]^q \ldots d\mathbf{\Gamma}. \tag{18}$$

In general, $\langle 1 \rangle_q \neq 1$ for $q \neq 1$. Clearly, this is an odd "average!" It is also inconvenient as it requires evaluation of Z_q. To address this last problem we employ a simple generalization of the q-expectation average [10]

$$\langle \mathcal{A} \rangle_q = \frac{\int \mathcal{A}(\mathbf{\Gamma})[1-(1-q)\beta\mathcal{H}(\mathbf{\Gamma})]^{\frac{q}{1-q}} d\mathbf{\Gamma}}{\int [1-(1-q)\beta\mathcal{H}(\mathbf{\Gamma})]^{\frac{q}{1-q}} d\mathbf{\Gamma}}. \tag{19}$$

It is obviously normalized and convenient to apply. Also, unlike one current version of the non-extensive thermostatistics, the statistical weight is not dependent on the thermal average of the potential energy (which in turn depends on the statistical weight, which depends on the thermal energy, and so … and must be computed by iteration).

Equipartitioning the Kinetic Energy Is the "temperature" $1/\beta$ related to the variance of the momentum distribution as in the classical equipartition theorem? It happens that there is no simple generalization of the equipartition theorem of classical statistical mechanics. For a $2N$ dimensional phase space $\mathbf{\Gamma} = (x_1 \ldots x_N, p_1, \ldots p_N)$ the ensemble average for a *harmonic* system is

$$\frac{\langle p_k^2 \rangle_q}{\langle 1 \rangle_q} = \frac{\int [p_q(\mathbf{\Gamma})]^q p_k^2 d\mathbf{\Gamma}}{\int [p_q(\mathbf{\Gamma})]^q d\mathbf{\Gamma}} = \frac{1}{\beta} \frac{1}{1-(q-1)N} \tag{20}$$

where we assume unit mass. For the case of $q = 1$ we find the standard result that $\langle p_k^2 \rangle = 1/\beta$. In general we find that the average is proportional to $1/\beta$ but not equal to it. The situation is equally strange for the unnormalized "multifractal" average where

$$\langle p_k^2 \rangle_q = \text{stuff} \times \frac{1}{\beta^{1+(1-q)N/2}} \tag{21}$$

but the "stuff" is a q-dependent constant that may be negative or imaginary!

A related result is that the distribution of momenta cannot be written as a product of single particle distributions and we find that

$$\langle p_k^2 + p_k^2 \rangle_q \neq 2\langle p_k^2 \rangle_q. \tag{22}$$

The result is surprising – there is no simple linear scaling of the variance of the momentum with the number of degrees of freedom.

Non-extensivity of the Entropy and Long-Range Interactions When $q = 1$ the extensivity of the entropy can be used to derive the Boltzmann entropy equation $S = k \ln W$ in the microcanonical ensemble. When $q \neq 1$, it is the odd property that the generalization of the entropy S_q is not extensive that leads to the peculiar form of the probability distribution. The non-extensivity of S_q is an intriguing property that has led to speculation that Tsallis statistics may be applicable to a variety of systems where interaction length scales comparable to the system size violate the assumptions underlying Gibbs-Boltzmann statistics [11,9].

2.3 Generalized Partition Functions for Ideal Systems

The properties of ideal systems that are predicted by the non-extensive thermostatistics have been described elsewhere with an emphasis on the possible values of q for an ideal gas or perfect crystal. Our purpose is different. Here we simply evaluate the partition function, in a given version of the non-extensive thermostatistics defined above, for two ideal systems to note generic properties of the results. These results will lead us to recognize a general problem with the evaluation of non-extensive averages of this kind for systems with a large number of particles and suggest how we might "fix" it.

Translational Partition Function Let us focus on the $q > 1$ regime for systems with a Hamiltonian of the form

$$\mathcal{H}(\mathbf{\Gamma}) = \sum_{k}^{N} \frac{1}{2m} \mathbf{p}_k^2 + \mathcal{V}(\mathbf{r}^N), \tag{23}$$

in d-dimensional space. The partition function can be written as

$$\begin{aligned} Z_q = & \left(\frac{1}{\Lambda \sqrt{q-1}} \right)^{dN} \frac{\Gamma(\frac{1}{q-1} - \frac{dN}{2})}{\Gamma(\frac{1}{q-1})} \\ & \times \int \left[1 - (1-q)\beta \mathcal{V}(\mathbf{r}^N) \right]^{\frac{1}{1-q} + \frac{dN}{2}} d\mathbf{r}^N \end{aligned} \tag{24}$$

where $\Lambda = \sqrt{h^2 \beta / 2\pi m}$ is the thermal de Broglie wavelength of the kth oscillator. For an ideal gas of N particles where $\mathcal{V} = 0$ we find

$$Z_q^{trans} = \left(\frac{L}{\Lambda \sqrt{q-1}} \right)^{dN} \frac{\Gamma(\frac{1}{q-1} - \frac{dN}{2})}{\Gamma(\frac{1}{q-1})} \tag{25}$$

where L is the length of a side of a cubic box containing the ideal gas. In the limit that $q \to 1$ we can use the asymptotic approximation $\Gamma(x+a)/\Gamma(x+b) = x^{a-b}$, good for large x, to show that the standard partition function for an ideal gas is recovered

$$\lim_{q \to 1} Z_q = \left(\frac{L}{\Lambda} \right)^{dN} . \tag{26}$$

One Way to Define the Vibrational Partition Function It is clear that when $q \neq 1$ the thermostatistical weight is not separable into a translational and configurational component even for a separable Hamiltonian. To make this point explicit, consider a system of N one-dimensional harmonic oscillators with the Hamiltonian

$$\mathcal{H}(\mathbf{\Gamma}) = \sum_{k}^{N} \left[\frac{1}{2m} p_k^2 + \frac{1}{2} m \omega_k^2 x_k^2 \right] . \tag{27}$$

The canonical ensemble partition function is the phase space integral

$$Z_q = \frac{1}{h^N} \int dx^N \int dp^N \left[1 - (1-q)\beta \mathcal{H}(\mathbf{\Gamma}) \right]^{\frac{1}{1-q}} . \tag{28}$$

The configuration integral can be evaluated and the partition function is

$$Z_q = \left(\frac{2\pi}{h\beta(q-1)} \right)^N \frac{\Gamma(\frac{1}{q-1} - N)}{\Gamma(\frac{1}{q-1})} \prod_{k}^{N} \frac{1}{\omega_k} \tag{29}$$

Note that there is not a unique separation of the partition function as $Z_q = Z_q^{\mathrm{trans}} Z_q^{\mathrm{vib}}$. However, using the result for the ideal gas translational partition function

$$Z_q^{\mathrm{vib}} = \left(\frac{2\pi}{m\beta(q-1)} \right)^{N/2} \frac{1}{L^N} \frac{\Gamma(\frac{1}{q-1} - N)}{\Gamma(\frac{1}{q-1} - \frac{N}{2})} \prod_{k}^{N} \frac{1}{\omega_k} \tag{30}$$

In the limit that $q \to 1$ we see that

$$\lim_{q \to 1} Z_q = \left(\frac{2\pi}{\beta h} \right)^N \prod_{k}^{N} \frac{1}{\omega_k} \tag{31}$$

and the canonical ensemble partition function for N harmonic oscillators in classical Gibbs-Boltzmann statistics is recovered.

2.4 Problems Arise for Many-Body Systems

Consider a system of N particles in d dimensions. Using the standard procedure employed above of integrating over the momenta in Cartesian coordinates, we can write the average of a mechanical property $\mathcal{A}(\mathbf{r}^N)$ using Eq. (19) as [10]

$$\langle \mathcal{A} \rangle_q = \frac{\int \mathcal{A}(\mathbf{r}^N) \left[1 - (1-q)\beta \mathcal{V}(\mathbf{r}^N) \right]^{\frac{q}{1-q} + \frac{dN}{2}} d\mathbf{r}^N}{\int \left[1 - (1-q)\beta \mathcal{V}(\mathbf{r}^N) \right]^{\frac{q}{1-q} + \frac{dN}{2}} d\mathbf{r}^N} . \tag{32}$$

This definition of the normalized statistical average is based on and proportional to the q-expectation value. However, it is more useful since it is not necessary to evaluate the partition function to compute the average. Nevertheless, this result is unsatisfactory. It is poorly behaved in the $N \to \infty$ thermodynamic limit. We will address this problem in the next section in the context of Maxwell-Tsallis statistics.

It is interesting to compare our Tsallis sampling algorithm with the multicanonical sampling method [12]. In the latter, a random walk in energy is the essential feature used for enhanced sampling with great success in a number of applications. It has been shown [13], however, that in the thermodynamic limit the multicanonical algorithm is identical with the regular Metropolis scheme. The reasoning is that, in a system with a large number of particles N, the entropy $S(E)$ is a smooth function of the energy E, and can thus be expanded to first order in ΔE. In this case, the acceptance probability of the multicanonical sampling update becomes

$$\lim_{N\to\infty} \exp(-\Delta S) = \exp(-\frac{\partial S}{\partial E}\Delta E) = \exp(-\frac{\Delta E}{T}), \tag{33}$$

where we use the equality $\partial S/\partial E = 1/T$.

In the thermodynamic limit, the Tsallis updating scheme has the form

$$\lim_{N\to\infty} \exp(-\frac{\Delta \bar{E}}{T}) = \left(\frac{E}{E+\Delta E}\right)^{\frac{q}{1-q}} \tag{34}$$

which is basically unity since the energy change ΔE is local (i.e., small) and E is of order N.

It is ironic that in the $N \to \infty$ limit the Tsallis sampling algorithm has the main feature for which the multicanonical algorithm was designed – it performs a random walk in energy – while the multicanonical algorithm loses this feature.

2.5 Enforcing Separability Using Maxwell–Tsallis Statistics

Take the equilibrium distribution to be a hybrid product of (1) a Tsallis statistical distribution (for the configurations) and (2) a Maxwell distribution (for the momenta) as [14,10]

$$p_q(\mathbf{r}^N, \mathbf{p}^N) = \exp[-\beta(\mathcal{T}(\mathbf{p}^N) + \bar{\mathcal{V}}(\mathbf{r}^N))] \tag{35}$$

where

$$\bar{\mathcal{V}}(\mathbf{r}^N) = \frac{1}{\beta(q-1)} \ln\left[1 - (1-q)\beta\mathcal{V}(\mathbf{r}^N)\right] \tag{36}$$

is the transformed potential and

$$\mathcal{T}(\mathbf{p}^N) = \sum_k^N \frac{1}{2m_k} p_k^2 \tag{37}$$

is the kinetic energy of the N-body system defined in the usual way.

This distribution can also be derived in a more systematic way from the classical density distribution

$$e^{-\beta\mathcal{H}} = \left(e^{-\beta\mathcal{H}/P}\right)^P = \left(e^{-\beta\mathcal{T}/P}e^{-\beta\mathcal{V}/P}\right)^P \tag{38}$$

where we have separated the kinetic energy and potential energy contributions. Now suppose that we carry out the approximation described above for the Boltzmann factor alone. We find

$$e^{-\beta\mathcal{H}} \approx e^{-\beta\mathcal{T}}\left(\frac{1}{1+\beta\mathcal{V}/P}\right)^P . \tag{39}$$

This is precisely the expression for the Maxwell-Tsallis statistical distribution where $P = 1/(q-1)$.

2.6 Corrected Ensemble Averages

Consider a system of N particles in d dimensions. Using the standard procedure of integrating over the momenta in Cartesian coordinates, we can write the average of a mechanical property $\mathcal{A}(\mathbf{r}^N)$ using Eq. (19) as

$$\langle\mathcal{A}\rangle_q = \frac{\int \mathcal{A}(\mathbf{r}^N)\left[1-(1-q)\beta\mathcal{V}(\mathbf{r}^N)\right]^{\frac{q}{1-q}}\,d\mathbf{r}^N}{\int \left[1-(1-q)\beta\mathcal{V}(\mathbf{r}^N)\right]^{\frac{q}{1-q}}\,d\mathbf{r}^N}. \tag{40}$$

This definition is based on and proportional to the q-expectation value. This result lacks the odd N-dependence in the exponent of the configurational weighting function found for the case of pure Tsallis statistics in Eq. (32). While it is not clear which result is “right,” this expression is certainly more satisfying in the $N \to \infty$ thermodynamic limit.

This definition is not consistent with the current, self-consistent version of non-extensive Tsallis statistics where the thermostatistical weight depends on the ensemble averaged energy which depends on the thermostatistical weight and so on. The definition of the ensemble average employed here is justified in that it is convenient to use in developing computational algorithms for enhanced sampling and optimization, it behaves well in the large N limit, and it provides normalized statistical averages without depending implicitly on the averaged energy.

3 General Properties of Mass Action and Kinetics

It is useful to explore the generalized forms of the law of mass action and kinetic rate constants in the non-extensive thermostatistics. Specific features and scaling relations emerge that will be essential motivations for the computational methods presented in the sections that follow. Consider the equilibrium

$$\alpha \rightleftharpoons \gamma. \tag{41}$$

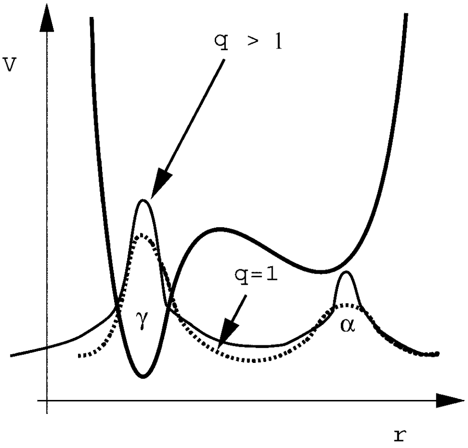

Fig. 2. Schematic of the equilibrium probability distributions for reactant (α) and product (γ) macrostates in a double well potential (thick line). Relative to the standard Boltzmann distribution (dashed line), the Tsallis statistical distribution (thin line) shows an enhanced probability for the system to be found in the barrier region.

The general form of the equilibrium constant is

$$K_{eq} = \frac{X_\gamma}{X_\alpha} = \frac{\int [p_q]^q \theta_\gamma(\mathbf{\Gamma}) d\mathbf{\Gamma}}{\int [p_q]^q \theta_\alpha(\mathbf{\Gamma}) d\mathbf{\Gamma}} \tag{42}$$

where X_α and X_γ are mole fractions of reactants and products, and the Heaviside functions $\theta_\alpha(\mathbf{\Gamma})$ and $\theta_\gamma(\mathbf{\Gamma})$ are unity for phase space points in well α and γ, respectively, and zero otherwise (see Fig. 2). zero otherwise. We will restrict our evaluation to one dimension; extension to many dimensions is straightforward.

In the limit that the barrier height is large compared with the thermal energy, it is standard practice to expand the potential near the reactant well minimum to quadratic order and approximate the integral by the contribution near the well minimum. This approximation is well justified when the thermal distribution is Gaussian and relatively short-ranged. However, for $q > 1$ the approximation is not so easily justified. For $q = 2$ the equilibrium distribution for a harmonic well is a Cauchy-Lorentz distribution so that it is likely that significant contributions from anharmonicity far from the well minimum will contribute to the integral. Aware of this limitation, we carry out the integral using a harmonic approximation to the α and γ potential wells. For the α state well

$$\mathcal{V}(x) \simeq \mathcal{V}(x_\alpha) + m\omega_\alpha^2 (x - x_\alpha)^2/2. \tag{43}$$

We follow this with a low temperature approximation to the integral over the well's phase space population. The resulting equilibrium constant is

$$K_{eq}(\beta; q) = \frac{X_\gamma}{X_\alpha} = \frac{\omega_\alpha}{\omega_\gamma} \left[\frac{1 - (1-q)\beta\mathcal{V}(x_\gamma)}{1 - (1-q)\beta\mathcal{V}(x_\alpha)} \right]^{\frac{1}{1-q}} . \tag{44}$$

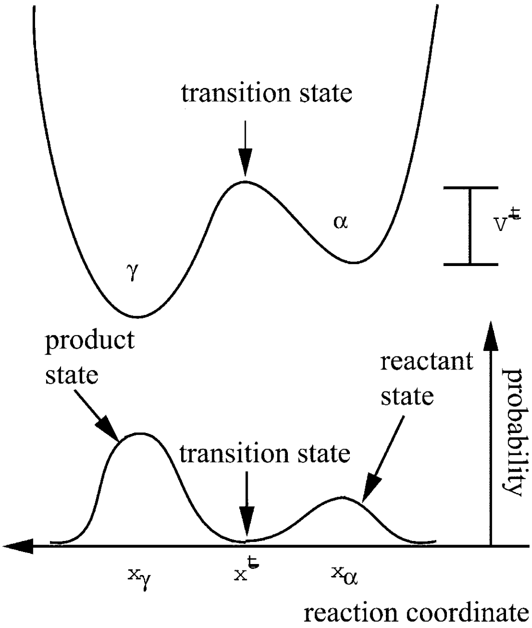

Fig. 3. Schematic showing the potential energy surface along the reaction coordinate indicating the transition state dividing surface separating the reactant and product macrostate distributions.

When $q \neq 1$, the equilibrium distributions are more delocalized and the low temperature approximation may not be well justified.

3.1 Transition State Theory for Rates of Barrier Crossing

The transition state theory estimate of the rate constant for barrier crossing provides an upper bound on the rate of transition between well-defined reactant and product states. An assumption of the theory is that once a reactant acquires enough energy to cross the barrier, it will cross the transition state and be deactivated as product. Dynamical recrossings of the transition state, associated with weak or strong damping, or nonadiabatic transitions, lead to reductions in the rate of barrier crossing from the transition state theory estimate. In this section, we examine the transition state theory rate constant for Tsallis statistics. This discussion is restricted to one-dimension, but the generalization to many dimensions is straightforward.

To evaluate the rate constant in the transition state theory approximation, we need only know the form of the equilibrium distribution. Consider a one-dimensional bistable potential with the transition state positioned along the reaction coordinate x at $x = x^{\ddagger}$ as in Fig. 3. The TST rate for forward reaction is defined as

$$k_{TST}(\beta;q) = \frac{\int \delta(x - x^{\ddagger}) v \Theta(v) [p_q(\mathbf{\Gamma})]^q d\mathbf{\Gamma}}{\int \Theta(x^{\ddagger} - x) [p_q(\mathbf{\Gamma})]^q d\mathbf{\Gamma}} \tag{45}$$

where $p_q(\mathbf{\Gamma}) = p_q(x, v)$ is the generalized statistical distribution

$$p_q(x,v) = \frac{1}{Z_q h} \left[1 - (1-q)\beta \mathcal{H}(x,v)\right]^{\frac{1}{1-q}} \tag{46}$$

and

$$Z_q = \frac{1}{h} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dv [1 - (1-q)\beta \mathcal{H}(x,v)]^{\frac{1}{1-q}} \tag{47}$$

is the generalized partition function and $\mathcal{H}(x,v) = mv^2/2 + \mathcal{V}(x)$ is the Hamiltonian.

For $q < 1$ there can be difficulties that arise from distributions which have zero probability in the barrier region and zero rate constant. In our analysis we assume that for any q the zero of energy is chosen such that the probability $p_q(\mathbf{\Gamma})$ is positive and real for all $\mathbf{\Gamma}$. The transition state theory rate constant as a function of the temperature and q is

$$k_{TST}(\beta;q) \\ = \frac{\Gamma(\frac{q}{q-1})}{\Gamma(\frac{q}{q-1} - \frac{1}{2})} \left(\frac{q-1}{2\pi m \beta}\right)^{1/2} \left[1 - (1-q)\beta \mathcal{V}(x^{\ddagger})\right]^{\frac{1}{1-q}} \frac{1}{\chi(\beta;q)} \tag{48}$$

where

$$\chi(\beta;q) = \int_{-\infty}^{x^{\ddagger}} [1 - \beta(1-q)\mathcal{V}(x)]^{\frac{q}{1-q} + \frac{1}{2}} \, dx. \tag{49}$$

We can approximate this fraction of states in the reactant well, by expanding the potential in a harmonic approximation and assuming that the temperature is low compared with the barrier height. This leads to an estimate for the rate constant

$$k_{TST}(\beta;q) = \frac{\omega_\alpha}{2\pi} \left[\frac{1 - (1-q)\beta \mathcal{V}(x^{\ddagger})}{1 - (1-q)\beta \mathcal{V}(x_\alpha)}\right]^{\frac{1}{1-q}}. \tag{50}$$

As we expect, in the limit that $q \to 1$ the standard transition state theory result

$$k_{TST}(\beta; q=1) = \frac{\omega_\alpha}{2\pi} \exp[-\beta(\mathcal{V}(x^{\ddagger}) - \mathcal{V}(x_\alpha))] \tag{51}$$

is recovered. For the special case of $q = 2$ we find

$$k_{TST}(\beta; q=2) = \frac{\omega_\alpha}{2\pi} \left[\frac{1 + \beta \mathcal{V}(x_\alpha)}{1 + \beta \mathcal{V}(x^{\ddagger})}\right]. \tag{52}$$

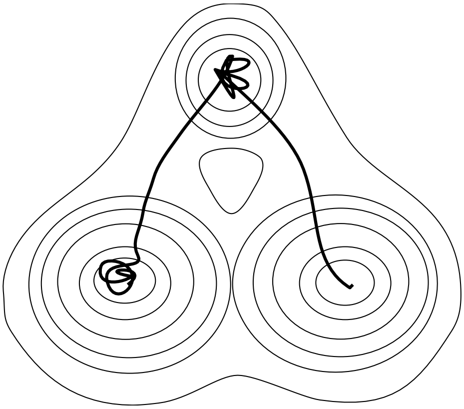

Fig. 4. The two-dimensional potential considered in the text has two deep holes, seen at the bottom of the plot, and a less deep hole centered above them.

This remarkable result, anticipated in earlier examinations of the rate of the optimal cooling schedule for generalized simulated annealing is that the temperature dependence of the rate constant is found to be a power law for the case $q > 1$ rather than the standard exponential Arrhenius temperature dependence found for $q = 1$. This result suggests that the rate of relaxation to the equilibrium distribution from a non-equilibrium distribution will be substantially enhanced when $q > 1$ relative to the $q = 1$ result. It also suggests that the rate of sampling phase space should be substantially enhanced through an increased rate of barrier crossing. These features lead to enhanced simulated annealing methods as well as algorithms for thermostatistical averaging.

3.2 Master Equations and Relaxation to Equilibrium

The relaxation of a system to equilibrium can be modeled using a master equation

$$\frac{dP_i}{dt} = \sum_{j \neq i} \left[L_{ij} P_j(t) - L_{ji} P_i(t) \right] . \tag{53}$$

Using the results of the previous section, the elements of the transition matrix from state j to state i can be estimated in the transition state theory approximation

$$L_{ij} = \frac{\omega_j}{2\pi} \left[\frac{1 + (q-1)\beta \mathcal{V}(x_{ij}^{\ddagger})}{1 + (q-1)\beta \mathcal{V}(x_j)} \right]^{\frac{1}{1-q}} , \tag{54}$$

where the total phase space probability of the jth state is proportional to

$$M_j = \frac{2\pi}{\beta \omega_j} \left[1 + (q-1)\beta \mathcal{V}(x_j) \right]^{\frac{1}{1-q}} . \tag{55}$$

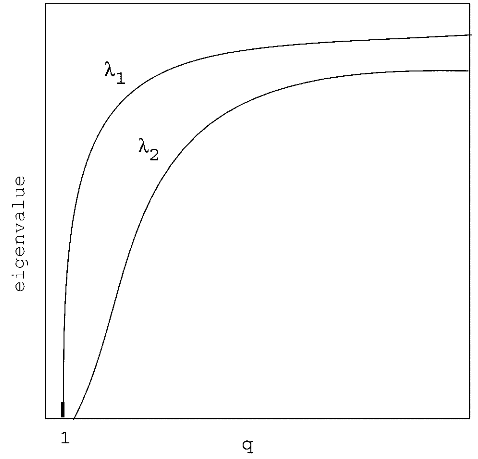

Fig. 5. The dependence on q of the non-zero eigenvalues of the two-dimensional, three-hole model described in the text shows the rapid onset of escape from wells as soon as q exceeds unity and the saturation at higher values of q.

The symmetric transmission coefficients are defined $B_{ij} = L_{ij}M_i$. The general solutions are of the form

$$P_i(t) = \sum_n a_n \phi_i^{(n)} \exp(-\lambda_n t) \tag{56}$$

in terms of the eigenfunctions $\phi_i^{(n)}$ and eigenvalues λ_n of the transmission matrix $\hat{L}$.

We have calculated the eigenvalues for a two-dimensional model system described by a potential function depicted in Fig. 4 and consisting of three holes, two deeper holes of equal depth, and a more shallow hole. The barrier between the deeper holes is higher than the other two barriers, which have equal height [10]. The eigenvalues are shown in Fig. 5 as a function of q. One of the three eigenvalues is zero, for any q, for reasons of conservation of probability. In the figure, we plot the other two as a function of q, together with their ratio. There is to be observed, for a thermal energy of a tenth of the well depth, the quick increase in the magnitude of the eigenvalues for $q > 1$. This implies fast relaxation to equilibrium even at low temperatures, due to the delocalized character of the Tsallisian distributions.

Do we expect this model to be accurate for a dynamics dictated by Tsallis statistics? A jump diffusion process that randomly samples the equilibrium canonical Tsallis distribution has been shown to lead to anomalous diffusion and Lévy flights in the $5/3 < q < 3$ regime [15]. Due to the delocalized nature of the equilibrium distributions, we might find that the macrostates of our master equation are not well defined. Even at low temperatures, it may be difficult to identify distinct macrostates of the system. The same delocalization can lead to large transition probabilities for states that are not adjacent in configuration space. This would be a violation of the assumptions of the transition state theory – that once the system crosses the transition state from the reactant macrostate it will be deactivated and equilibrated in the product state. Concerted transitions between spatially far-separated states may be common. This would lead to a highly connected master equation where each state is connected to a significant fraction of all other macrostates of the system [16,17].

4 Tsallis Statistics and Simulated Annealing

Versions of the simulated annealing, Monte Carlo and Molecular Dynamics methods that sample Tsallis thermostatistical distributions have been applied to a variety of model and molecular systems. These applications consist of two types. In the first, the Tsallis statistical distributions are generated to produce a trajectory with enhanced sampling of the conformational space of the system. However, the goal is to compute Gibbs-Boltzmann statistical averages or discover low-lying energy minima of the system. In such cases, the non-extensive Tsallis statistics are used as a novel means to a standard end. In the second, the point is precisely to explore the properties of systems that are well described by the non-extensive Tsallis statistics. Representative studies in this and the following sections demonstrate the accomplishments and promise of these methods.

Finding the ground state conformation of biologically important molecules has an obvious importance, both from the academic and pragmatic points of view [18]. The problem is hard for biomolecules, such as proteins, because of the ruggedness of the energy landscape which is characterized by an immense number of local minima separated by a broad distribution of barrier heights [19,20]. Algorithms to find the global minimum of an empirical potential energy function for molecules have been devised, among which a central role is played by the simulated annealing methods [21,22]. Once a cooling schedule is chosen, representative configurations of the allowed microstates are generated by methods either of the molecular dynamics (MD) or Monte Carlo (MC) types. For biomolecular simulations, simulated annealing is traditionally built on an MD approach [23] where the dynamics of the system is simulated by integrating the Newtonian equations of motion and the temperature is controlled through coupling to a heat bath. If the MC approach is used, after having drawn a new configuration, it is accepted or rejected according to a probability of, for example, the Metropolis type [24]

$$p = \min\left[1, \exp\left(-\beta \Delta \mathcal{V}\right)\right], \tag{57}$$

where $\beta = 1/kT$ and $\Delta\mathcal{V}$ is the change in potential energy. This acceptance probability has the desirable features that (i) it obeys detailed balance and (ii) it reduces to a steepest descent minimizer at low temperature (where only moves which decrease the potential energy are accepted). In addition to the standard Metropolis Monte Carlo protocol, several other smarter MC algorithms have been designed using atomic moves biased by the forces acting upon the atoms in the molecule [25,26] or by relaxing the restriction to Markov processes [27].

4.1 Essential Features and Algorithms

Inspired by the non-extensive thermostatistics, a new generalized simulated annealing algorithm was suggested based on the acceptance probability

$$p = \min\left[1, (1-(1-\mathrm{q})\beta\Delta\mathcal{V})^{\frac{1}{1-\mathrm{q}}}\right] \tag{58}$$

where $\Delta\mathcal{V}$ is the change in the potential energy. This probability has the desirable property that it reduces to a steepest descent in energy for $T \to 0$. This method was shown [28] to be faster than both the classical simulated annealing and the fast simulated annealing methods [29] and was employed to find close to optimal solutions to the traveling salesman problem [30].

However, it is easy to demonstrate that the acceptance criterion Eq. (58) does not obey detailed balance. It is true that detailed balance is a sufficient but not necessary condition for the convergence to the equilibrium distribution. Even so, the acceptance in Eq. (58) does not, in general, converge towards the generalized distribution of Eq. (13).

We proposed a generalized acceptance probability

$$p = \min\left[1, \left(\frac{1-(1-q(T))\beta\mathcal{V}_{\mathrm{new}}}{1-(1-q(T))\beta\mathcal{V}_{\mathrm{old}}}\right)^{\frac{q(T)}{1-q(T)}}\right], \quad \lim_{T\to 0} q(T) = 1, \tag{59}$$

that obeys detailed balance [31]. Moreover, this acceptance probability does tend towards the generalized equilibrium distribution in Eq. (13). The parameter q is varied as a monotonically decreasing function of temperature. Starting with a convenient value of q at the initial temperature, q tends towards 1 as the temperature decreases during annealing. Since $q \to 1$ as $T \to 0$, the desirable reduction to a steepest descent at low temperature is preserved.

Note that the probability is raised to the power of q as required by the generalized statistical mechanics Eq. (13). In this generalization, the average of an observable O is defined as $O = \sum {p_i}^q O_i$ and thus the detailed balance must be written as $W_{ij}p_i^q = W_{ji}p_j^q$, where W_{ij} are the elements of the transition matrix. The fact that the probability distribution is raised to the power of q allows one to use simulation binning to compute average properties of the system according to the standard definition of the statistical average [32].

This acceptance probability is, in the spirit of the generalized statistical mechanics of Tsallis, invariant under the transformation

$$\bar{\mathcal{V}}_i \to \bar{\mathcal{V}}_i + \bar{\mathcal{V}}_0, \tag{60}$$

where $\bar{\mathcal{V}}_0$ is a constant shift in energy and, by definition,

$$\bar{\mathcal{V}} = \frac{q}{\beta(q-1)} \ln\left[1-(1-q)\beta\mathcal{V}\right]. \tag{61}$$

For $q \to 1$ the property that the probability distribution does not depend on the choice of the zero of energy is thus recovered. Also, using the definition Eq. (61) the acceptance probability Eq. (59) can be written in the more familiar form

$$p = \min\left[1, \exp\left(-\beta\Delta\bar{\mathcal{V}}\right)\right]. \tag{62}$$

It was shown [33] that when the maximum entropy formalism is applied to the entropy postulated by Tsallis Eq. (2) one is able to recover the general Lévy probability distribution (corresponding to a fractal random walk the dimension of which is determined by q). A variational entropic formalism based on the Boltzmann entropy is unable to do this. Using the acceptance probability proposed here and an initial value of $q(T) > 1$, a Markov chain generated at constant temperature will converge to a Lévy distribution. For the particular case of $q = 2$, the Lévy distribution is a Cauchy-Lorentz distribution which is the same distribution used for the attempted moves in the fast simulated annealing method of Szu and Hartley [29].

The Metropolis acceptance criterion in Eq. (59) is not the only solution that guarantees convergence towards the probability distribution in Eq. (13). It is known that a solution of the type proposed by Barker

$$W_{ij} = \alpha_{ij} \frac{\rho_j}{\rho_j + \rho_i}, \tag{63}$$

where α_{ij} is the *a priori* symmetric transition matrix and ρ_i the probability distribution function for the ith state, also satisfies detailed balance [2,3]. Thus, by constructing the acceptance probability using Eqs. (13) and (63) we recover the familiar form of acceptance widely used in simulations of spin systems

$$p = \frac{1}{2}\left[1 - \tanh\left(\beta \Delta \bar{\mathcal{H}}/2\right)\right]. \tag{64}$$

A simulation using this acceptance probability will tend towards the generalized probability distribution of Tsallis, as does Eq. (59). It will obey detailed balance and, by making $q \to 1$ as the temperature decreases, it will behave like a steepest descent at low temperatures.

4.2 Temperature Scaling in Simulated Annealing

Using these results, we can derive a scaling relation for the optimal cooling schedule in a simulated annealing optimization protocol (see Fig. 6). We suppose that the relevant energy scales of $\mathcal{V}(\mathbf{r}^N)$ are bounded by $\Delta\mathcal{V}$, the difference in energy between the ground and first excited state minima, and $\mathcal{V}^\ddagger$, the highest barrier on the potential surface accessed from the global energy minimum. The final temperature (maximum β) reached in a simulated annealing run must be small enough so that at equilibrium the mole fraction in the global energy minimum basin is significant. In other words, based on Eq. (44) we demand that $K_{eq}^{\max} = K_{eq}(\beta_{\max}; q)$.

The time that the trajectory must spend at $\beta_{\max}$ to ensure that the equilibrium distribution is sampled is at least τ_{min} – the time required to surmount the largest barrier separating the global energy minimum from other thermodynamically important states. Using Eq. (50) we find

$$\tau_{\min} = \frac{2\pi}{\omega_\alpha}\left[1 - (1 - \frac{1}{\eta})\frac{\mathcal{V}^\ddagger}{\Delta\mathcal{V}}\right]^{\frac{1}{1-q}}. \tag{65}$$

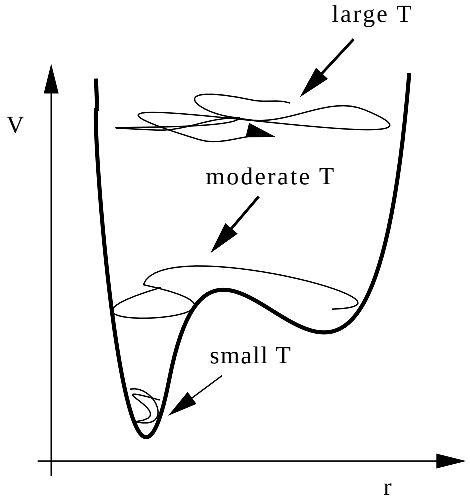

Fig. 6. Schematic showing the general protocol of simulated annealing.

where

$$\eta = \left(\frac{\omega_\alpha}{\omega_\gamma} K_{eq}^{\max} \right)^{q-1} . \tag{66}$$

$K_{eq}^{\max}$ is the maximum allowable equilibrium constant for the ground and first excited state populations at the final and lowest temperature reached in the annealing run, $\beta_{\max}$. For most cases of interest, we expect that $\eta << 1$.

In the limit $q \to 1$ of Gibbs-Boltzmann statistics, using the fact that $\lim_{x\to 0}[1 - a(1 - b^x)]^{1/x} = b^a$, we find that

$$\tau_{min} = \left(\frac{2\pi}{\omega_\alpha} \right) \eta^{-\frac{\mathcal{V}^\ddagger}{\Delta\mathcal{V}}} . \tag{67}$$

The time for classical simulated annealing increases exponentially as a function of the ratio of the energy scales $\mathcal{V}^\ddagger/\Delta\mathcal{V}$. However, for $q > 1$ the situation is qualitatively different. As a result of the weak temperature dependence in the barrier crossing times, the time for simulated annealing increases only weakly as a *power law*.

4.3 Further Applications of Enhanced Simulated Annealing

Having noted the effect of enhanced diffusion for systems sampling underlying Tsallis statistical distributions, Tsallis and Stariolo proposed the Generalized Simulated Annealing (GSA) method [28]. They assigned a visiting distribution that was Lévy-like (described by a parameter q_v) and the acceptance probability (controlled by a parameter q_a). In the GSA method the Monte Carlo walk does not converge to the Tsallis probability distribution. However, it does reduce to a steepest descent at low temperature. The effectiveness of the GSA method was demonstrated by isolating the global optimization of a model function of 4 variables with 16 minima.

In the work of Moret and coworkers [34], the GSA method of Tsallis and Stariolo was implemented with a modification – the Lévy form of the visiting visiting probability distribution was replaced by a distribution of jumps numerically calculated by inverting the power series representation of the cumulative probability distribution for jumps in GSA. As in the GSA, the Monte Carlo walk will not sample the Tsallis distribution. The method was applied to study the conformational optimization of $(H_2O)_3$ and polypeptides composed of 5 and 20 residues. Other applications have included linear-chain polymers, ceramics and ligand-protein interactions [35].

Salazar and Toral [36] have employed a standard Hybrid Monte Carlo method coupled with a simulated annealing protocol employing an exponential temperature decay. That work was then generalized to a Hybrid Monte Carlo method for sampling the Tsallis statistical distribution. They applied their annealing algorithm to find the global minima of a discrete random phase sine-Gordon Model. The results of their careful study support the conclusions drawn earlier [31,28] that the simulated annealing method is made more effective for values of $q > 1$ relative to the standard case of $q = 1$.

Another important application of the non-extensive thermostatistics has come in the context of the protein folding problem (specifically the conformational optimization of peptides) in our work [14] and that of Okamoto and Hansmann [37]. In those applications, Monte Carlo annealing methods that sample distributions derived from the non-extensive thermostatistics were found to lead to enhanced sampling and a more rapid convergence to low lying energy minima.

There have been additional applications of the GSA method and some formal results indicating that the method is weakly ergodic. Xiang and coworkers studied the Thomson model of N equal point charges constrained to a spherical surface [38]. They employed a schedule of linearly decreasing negative values of q_a. Nishimori and Inoue [39] also examined the GSA method, proving that the distribution sampled by the GSA in a sufficiently long Monte Carlo walk is independent of the starting probability distribution.

To date, all indications are that the Generalized Simulated Annealing methods based on a sampling of Tsallis statistic distributions show a faster convergence to the global minimum than the Classical Simulated Annealing approach [21] or the Fast Simulated Annealing method of Szu and Hartley [29].

4.4 The Importance of Detailed Balance

In earlier work on GSA methods, it was noted that two important properties are that (1) the method should satisfy detailed balance and sample the correct generalized distribution and (2) that at low temperatures the Monte Carlo method should reduce to a steepest descent in energy. In the earliest work, one or the other criterion was met, but not both [30,28]. It is important to note that both criteria – the condition of detailed balance and the reduction to a steepest descent method at low temperature – are met by the form of the generalized simulated annealing proposed by Andricioaei and Straub [31] but not for the standard GSA employed in most work to date (see above). While the detailed balance condition is sufficient but not necessary for sampling the correct underlying distribution, there appears to be no good reason why it should be violated. *When the asymptotic distribution is known, arguments for rates of convergence are easily made.* It is strongly advised that the method one employs *does* satisfy detailed balance. This is true for optimization methods and even more important for that case of *equilibrium* sampling discussed in the following section.

5 Tsallis Statistics and Monte Carlo Methods

The development of a Monte Carlo method that may be used to sample the Tsallis thermostatistical distributions is useful for two reasons. There is an interest in exploring thermodynamic averages over non-extensive thermostatistical distributions for complex systems; there is also an interest, which is the focus of this chapter, on developing optimization and enhanced sampling algorithms for computing Gibbs-Boltzmann thermal averages using a Monte Carlo walk that samples the Tsallis thermostatistical distributions.

In this section we describe a variety of Monte Carlo methods for the evaluation of non-extensive and extensive thermostatistical averages. All of the methods discussed will make use of features of the generalized statistics of Tsallis.

5.1 Monte Carlo Estimates of Tsallis Statistical Averages

A point in phase space or configuration space can be said to have the probability

$$p_q(\mathbf{r}^N) \sim \exp[-\beta \bar{\mathcal{V}}] \tag{68}$$

where $\bar{\mathcal{V}}$ is the effective potential energy [5]. The question arises, how might we simulate systems that are well described by the Tsallis statistical distributions when $q \neq 1$? One possibility is the direct implementation of the Metropolis Monte Carlo method where the acceptance probability is a function of the change in the effective potential energy $\Delta\bar{\mathcal{V}}$.

1. A new configuration is chosen within a region with uniform constant probability.
2. The point in configuration space is accepted or rejected according to the criterion
$$p = \min\left[1, \exp[-\beta \Delta\bar{\mathcal{V}}]\right] \tag{69}$$
where the change in the effective potential energy is $\Delta\bar{\mathcal{V}}$.
3. Repeat the previous steps.

This Monte Carlo algorithm will satisfy detailed balance and eventually sample the equilibrium Tsallis distribution.

We have implemented the generalized Monte Carlo algorithm using a hybrid MD/MC method composed of the following steps.

1. Velocities are randomly chosen from a Maxwell distribution at a given temperature.
2. The positions and velocities are updated for a time step Δt according to Newton's equation of motion using the force deriving from $\bar{\mathcal{V}}$.
3. The point (phase space point or configuration) is accepted or rejected according to the criterion
$$p = \min\left[1, \exp[-\beta(\Delta\mathcal{T} + \Delta\bar{\mathcal{V}})]\right] \tag{70}$$
where $\Delta\mathcal{T}$ is the change in standard classical kinetic energy and $\Delta\bar{\mathcal{V}}$ is the change in the effective potential energy
$$\bar{\mathcal{V}}(\mathbf{r}^N) = \frac{1}{\beta(q-1)} \ln\left[1 - (1-q)\beta\mathcal{V}(\mathbf{r}^N)\right]. \tag{71}$$
4. Return to 1.

When the integrator used is reversible and symplectic (preserves the phase space volume) the acceptance probability will exactly satisfy detailed balance and the walk will sample the equilibrium distribution $[p_q(\mathbf{r}^N)]^q$.

A similar algorithm has been used to sample the equilibrium distribution $[p_q(\mathbf{r}^N)]^q$ in the conformational optimization of a tetrapeptide [31] and atomic clusters at low temperature [14]. It was found that when $q > 1$ the search of conformational space was greatly enhanced over standard Metropolis Monte Carlo methods. In this form, the velocity distribution can be thought to be Maxwellian.

5.2 Monte Carlo Estimates of Gibbs–Boltzmann Statistical Averages

While the Monte Carlo algorithms described above will sample the generalized thermostatistical distributions, they can also be used to effectively sample the configuration space that is statistically important to the standard canonical ensemble probability density. The first method of this kind was the "q-jumping Monte Carlo" method of Andricioaei and Straub [14].

1. At random a choice is made to make a uniformly distributed "local" move, with probability $1 - P_J$, or a global "jump" move, with probability P_J.
2. Either (a) the *local* trial move is sampled from a uniform distribution, for example from a cube of side Δ *or* (b) the *jump* trial move is sampled from the generalized statistical distribution
$$T(\mathbf{r} \to \mathbf{r}') = p_q(\mathbf{r}') \tag{72}$$
at $q > 1$.

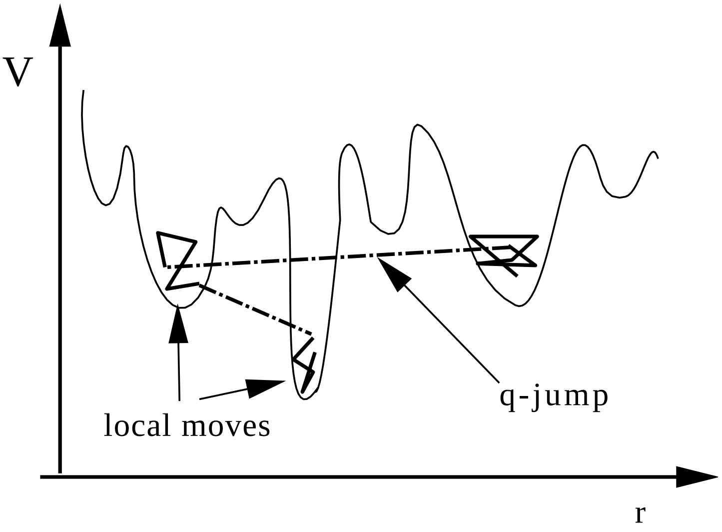

Fig. 7. Schematic representation of a q-jumping Monte Carlo trajectory. Initially, a Monte Carlo walk with $q > 1$ is run and far separated configurations are stored. A q-jumping Monte Carlo run is then generated where a local $q = 1$ Monte Carlo search is punctuated by random "jump" moves sampled from the the $q > 1$ distribution.

3. Either (a) the *local* move is accepted or rejected by the standard Metropolis acceptance criterion with probability

$$p = \min\left[1, \exp[-\beta \Delta \bar{\mathcal{V}}]\right] . \tag{73}$$

or (b) the *jump* trial move is accepted or rejected according to the probability

$$p = \min\left[1, \exp[-\beta \Delta \bar{\mathcal{V}} \left(\frac{p_q(\mathbf{r})}{p_q(\mathbf{r}')}\right)^q]\right] \tag{74}$$

where the bias due to the non-symmetric trial move has been removed. This algorithm satisfies detailed balance.

Fig. 7 shows a Monte Carlo walk that might be generated by the q-jumping Monte Carlo method. The delocalized Tsallis statistical distributions are sampled to provide *long range moves* between well-separated but thermodynamically significant basins of configuration space. Such long range moves are randomly shuffled with short range moves which provide good local sampling within the basins. Note that this Monte Carlo walk samples the equilibrium *Gibbs-Boltzmann* distribution.

When $q > 1$, it has been shown that the generalized q-jumping Monte Carlo trajectories will cross barriers more frequently and explore phase space more efficiently than standard Monte Carlo (without jump moves). (For a review of recent methods for enhanced phase-space sampling see [4].) We have shown how this property can be exploited to derive effective Monte Carlo methods which provide significantly enhanced sampling relative to standard methods.

5.3 Monte Carlo Algorithm for Spin Systems

We have explored a modification of the Tsallis statistical MC [40] based on the Glauber acceptance probability sometimes known as Barker sampling. In the computation of equilibrium thermodynamic averages for spin systems it was shown [41] that for two-state problems such as the Ising model the Glauber (Barker) acceptance is favorable. The so-called asymmetric acceptance in Eq. (68) is just one of the solutions that guarantees convergence towards the distribution probability in Eq. (13). It is known from the theory of Monte Carlo simulation that a symmetrical solution of the type

$$W(x \to x') = T(x \to x') \frac{\rho_{x'}}{\rho_x + \rho_{x'}}, \tag{75}$$

where $T(x \to x')$ is the *a priori* symmetric transition matrix and ρ_x the probability distribution function for the x state, also satisfies detailed balance. Thus, by constructing the acceptance probability

$$p = \frac{1}{2}\left[1 - \tanh\left(\beta \Delta \bar{\mathcal{H}}/2\right)\right] \tag{76}$$

where $\bar{\mathcal{H}} = \frac{q}{\beta(q-1)} \ln\left[1 - (1-q)\beta\mathcal{H}\right]$, we retrieve a form reminiscent of the Glauber function [42], widely used in simulations of spin systems. A simulation using this acceptance probability will also tend towards the generalized probability distribution of Tsallis, just as Eq. (68) does. It will obey detailed balance,

$$[p_q(x)]^q W(x \to x') = [p_q(x')]^q W(x' \to x) \tag{77}$$

where $W(x \to x')$ is an element of the transition matrix and, by making $q \to 1$ as the temperature decreases, it will behave like a steepest descent at low temperature.

We performed simulations on a two-dimensional Ising spin system, with a 16 site cell with periodic boundary conditions (wraped around a torus) [40]. Because of the form of the Tsallis probability, the energy cannot be negative for $q > 1$. For this reason we have chosen the following expression of the Ising Hamiltonian, with 0 and 1 for the possible values of the spin

$$\mathcal{H} = J \sum (1 - \delta_{s_i, s_j}) \tag{78}$$

where the sum extends over nearest neighbor spin pairs. The energy associated with the bond between nearest neighbors is 0 if the spins are parallel and 1 if anti-parallel. We have implemented the generalized Monte Carlo algorithm using a single spin flip procedure composed of the following steps.

1. A spin is picked at random and flipped.
2. The change of energy is calculated. Because of the short range interactions of the Ising systems, only nearest neighbors count.
3. The spin configuration is accepted or rejected according to the criterion Eq. (76).
4. Return to 1.

Since this acceptance will exactly satisfy detailed balance, the walk will converge towards sampling the equilibrium distribution $[p_q(x)]^q$.

Applications of this algorithm to explore the nature of simulated annealing and phase changes in a two dimensional Ising system are discussed in the following section.

5.4 Further Applications of Generalized Monte Carlo

In the work of Lima, Martins and Penna, the magnetic phase behavior of a two-dimensional nearest neighbor Ising system was studied using the non-extensive thermostatistics proposed by Tsallis [43]. Using a combination of a Monte Carlo method and the broad histogram technique, a nearest-neighbor Ising model was examined for a range of q values. Two important conclusions were drawn from this thorough study. The first conclusion is a technical one – while the generalized Monte Carlo method of Andricioaei and Straub was found to sample well the Tsallis distribution, the equivalent Monte Carlo method based on the GSA algorithm did not. This is a clear demonstration of the importance of satisfying detailed balance in a Monte Carlo method. The second conclusion is of more general importance – the finite temperature phase transition exhibited by the two-dimensional Ising system in the thermodynamic limit for $q = 1$ appears to be washed out for the non-extensive case of $q > 1$. This is a general feature that appears to be related to the success of simulated annealing method based on sampling the generalized thermostatistical distributions.

It has long been recognized that when an annealing run is interrupted by a phase change, the large energy fluctuations and relaxation times necessitate a long period of time being spent in the cooling schedule moving through the phase change. Therefore, simulated annealing methods employing cooling rates inversely proportional to the heat capacity of the system have proved to be more effective than monotonic cooling schedules [44,45]. Better yet, by avoiding first order phase transitions altogether, the generalized simulated annealing method circumvents this problem.

It is also possible to generalize the “parallel tempering” Monte Carlo method using Tsallis statistical distributions [46,47]. In the standard parallel tempering method, a set of parallel Monte Carlo walkers are set to run at a variety of temperatures. The manifold of temperatures is chosen such that there is a significant overlap in the distribution functions of adjacent temperatures. In addition to the standard Monte Carlo moves at a fixed temperature, moves that exchange walkers at two temperatures (while maintaining the positions of the walkers) are

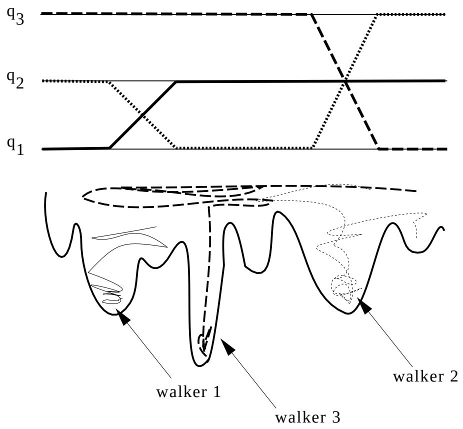

Fig. 8. Schematic representation of three Monte Carlo walkers moving over a rugged energy landscape in the generalized parallel tempering method.

occasionally attempted. Therefore, the path of a single walker will have an enhanced probablity of overcoming barriers on the potential energy surface through excursions at higher temperatures.

The "generalized parallel tempering" Monte Carlo method is depicted in Fig. 8. For the same reasons that the q-Jumping Monte Carlo method proved to be superior to the J-Walking Monte Carlo method in the simulation of atomic clusters [14], one can expect that the generalized parallel tempering method will provide enhanced sampling relative to the standard parallel tempering method using the same number of walkers.

6 Tsallis Statistics and Molecular Dynamics

Standard Molecular Dynamics for an ergodic system generates time averages which agree with averages taken over a microcanonical distribution. To compute averages such as Eq. (40) for the generalized canonical ensemble probability density using MD, we will employ a trick. We define a molecular dynamics algorithm such that the trajectory samples the distribution $p_q(\mathbf{r}^N)$ by having the trajectory move on a temperature dependent, but *static*, effective potential [14]. The equation of motion takes on a simple and suggestive form

$$m_k \frac{d^2\mathbf{r}_k}{dt^2} = -\nabla_{\mathbf{r}_k}\bar{\mathcal{V}} = \frac{-q}{[1-(1-q)\beta\mathcal{V}(\mathbf{r}^N)]}\ \nabla_{\mathbf{r}_k}\mathcal{V}(\mathbf{r}^N) \qquad (79)$$

for a particle of mass m_k, at position $\mathbf{r}_k$, and $\bar{\mathcal{V}}$ defined by Eq. (71). It is known that in the canonical ensemble a constant-temperature molecular dynamics algorithm generates samples from the configuration space according to the Boltzmann probability. As a result, this generalized molecular dynamics will sample the Tsallis statistical distribution $p_q(\mathbf{r}^N)$.

The effective force employed is the “exact” force for standard molecular dynamics, $-\nabla_{\mathbf{r}_k}\mathcal{V}$, scaled by

$$\alpha_q(\mathbf{r}^N;\beta) = \frac{q}{1-(1-q)\beta\mathcal{V}(\mathbf{r}^N)} \tag{80}$$

which is a function of the potential energy. This scaling function is unity when $q = 1$ but can otherwise have a strong influence on the dynamics. Assume that the potential is defined to be a positive function. In the regime $q > 1$, the scaling function $\alpha_q(\mathbf{r}^N,\beta)$ is largest near low lying minima of the potential. In barrier regions, where the potential energy is large, the scaling function $\alpha_q(\mathbf{r}^N,\beta)$ is small. This has the effect of reducing the magnitude of the force in the barrier regions. A particle attempting to pass over a potential energy barrier will meet with less resistance when $q > 1$ than when $q = 1$. At equilibrium, this leads to more delocalized probability distributions with an increased probability of sampling barrier regions and making transitions between thermodynamically relevant basins.

6.1 Molecular Dynamics Estimates of Gibbs–Boltzmann Statistical Averages

The generalized molecular dynamics described above generates trajectories which, averaged over time, sample the *Tsallis* statistical distribution. To compute averages over the *Gibbs-Boltzmann* distribution we may simply reweight each measurement as

$$\begin{aligned}\langle\mathcal{A}\rangle &= \left\langle \frac{\mathcal{A}\,\mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{\Gamma})}}{[1-(1-q)\beta\mathcal{H}(\boldsymbol{\Gamma})]^{\frac{q}{1-q}}}\right\rangle_q \\ &\quad\times\left\langle \frac{\mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{\Gamma})}}{[1-(1-q)\beta\mathcal{H}(\boldsymbol{\Gamma})]^{\frac{q}{1-q}}}\right\rangle_q^{-1} .\end{aligned} \tag{81}$$

Using this expression, the standard ($q = 1$) Gibbs-Boltzmann equilibrium average properties may be calculated over a trajectory which samples the generalized statistical distribution for $q \neq 1$ with the advantage of enhanced sampling for $q > 1$.

This method leads to an enhanced sampling of configuration space. However, it suffers a bit from the ease with which the trajectory moves through high energy regions in the $q > 1$ regimes [14]. Most of those regions of high potential energy are not thermodynamically important. It is good to visit them, as the q-jumping Monte Carlo does, but only on the way to another thermodynamically important region. The q-jumping Monte Carlo method has the advantage that the trajectory *directly* samples the Gibbs-Boltzmann distribution – no reweighting is necessary. Therefore, the walk is compelled to spend time in thermodynamically significant regions.

This effective MD method is consistent with Tsallis statistics in as much as a long time dynamical average for an ergodic system will provide results identical to the average over the what might be called a Maxwell-Tsallis statistical distribution. However, it cannot be said to tell us about the true dynamics of the system.

6.2 Rate Constants for Maxwell–Tsallis Statistics

The principles of microscopic reversibility and detailed balance dictate relationships between the rate constants for moving between macrostates and the equilibrium constants for the probability of being found in one macrostate or another. Earlier, we evaluated the rate and equilibrium constants in the "pure" Tsallis statistics. Here we derive those results using Maxwell-Tsallis statistics – the statistics appropriate for a system undergoing Molecular Dynamics using the effective potential $\bar{\mathcal{V}}$.

For $q = 1$, the normal transition state theory rate constant is independent of temperature at high temperatures and varies exponentially with temperature in the limit of low temperatures

$$k_{TST} \sim \exp(-\beta \mathcal{V}^{\ddagger}) \tag{82}$$

where kT is small compared with the barrier height $\mathcal{V}^{\ddagger}$.

For Maxwell-Tsallis statistics we expect to find a standard Newtonian dynamics on the Tsallis effective potential energy surface $\bar{\mathcal{V}}$. Therefore, the rate of activated barrier crossing should vary as

$$k_{TST} \sim \exp(-\beta \bar{\mathcal{V}}^{\ddagger}). \tag{83}$$

where $\bar{\mathcal{V}}^{\ddagger}$ is the activation energy on the Tsallis effective potential energy surface. As we have seen, the particular value of $\bar{\mathcal{V}}^{\ddagger}$ will depend on the choice of the zero of energy in the "true" potential energy function $\mathcal{V}$ when $q \neq 1$. In the nonextensive regime of $q \neq 1$, at low temperature the temperature scaling of the rate is a weak inverse power law and the rate of activated barrier crossing is found to be significantly larger than for $q = 1$. This is the origin of the enhanced sampling generated by Molecular Dynamics using a Tsallis statistical effective energy function.

6.3 Further Applications of Generalized Molecular Dynamics

The effectiveness of enhanced sampling generated by Molecular Dynamics methods that employ the Tsallis thermostatistical effective potential energy function $\bar{\mathcal{V}}$ has been beautifully demonstrated by Pak and Wang [48]. In their simulations of a 16-residue model peptide, they found that the peptide rapidly found the global free energy basin corresponding to an α-helical conformation. By increasing the probability of visiting regions of conformational space with high energy, the generalized Molecular Dynamics method increases the rate of barrier crossing in a way that allows the simulation trajectory to sample what appears to be reversible folding of the peptide.

Using the same method of molecular dynamics on the transformed potential surface, Pak and Wang have modeled the flexible docking of streptavidin with biotin and of the protein kinase C with phorbol-13-acetate [49]. The ligands are fully flexible and a partial flexibility of the receptor active site is allowed. The authors report, in both cases, accelerated docking and correct prediction of the docked structures indicating that the method may provide an effective means of optimization in the important problem of flexible docking and drug design.

7 Optimizing the Monte Carlo or Molecular Dynamics Algorithm Using the Ergodic Measure

How can a generalized Monte Carlo or Molecular Dynamics algorithm be tailored to a particular application? How can its performance be optimized? One useful criterion is to vary the algorithmic parameters (such as step size distribution and frequency) and additional variables (q and the zero of energy) to *maximize the rate of self-averaging* of the property under study. To do this, we employ the ergodic measure to estimate the rate of self-averaging [50,11,52].

Self-averaging is a necessary but not a sufficient condition for the ergodic hypothesis to be satisfied. The rate of self-averaging for a given property is expected to be proportional to the rate of phase space sampling. It is most convenient to use the potential energy metric defined for two independent trajectories α and β. We define the "move average" over the Monte Carlo trajectory of the potential energy $\mathcal{V}_j$ for the jth particle along the α trajectory after n moves as

$$\sqsubseteq_j^{\alpha}(n) = \frac{1}{n} \sum_k^n w(\mathbf{r}_k^N) \mathcal{V}_j(\mathbf{r}_k^N). \tag{84}$$

$w(\mathbf{r}_k^N)$ is a weighting factor for any property at a given position on the kth step $\mathbf{r}_k^N$. For a Metropolis Monte Carlo run the weighting factor is unity. For a generalized Monte Carlo run the weighting factor is given by Eq. (81). The ergodic measure is then defined as the sum over N particles

$$d_{\mathcal{V}}(n) = \frac{1}{N} \sum_j \left[\sqsubseteq_j^{\alpha}(n) - \sqsubseteq_j^{\beta}(n) \right]^2 \tag{85}$$

so that for an ergodic system if $n \to \infty$ then $d_{\mathcal{V}}(n) \to 0$. For either Monte Carlo or Molecular Dynamics, for large n we expect the form of the convergence to be

$$d_{\mathcal{V}}(n) = d_{\mathcal{V}}(0) \frac{1}{D_{\mathcal{V}} n} \tag{86}$$

where $D_{\mathcal{V}}$ is a rate for self-averaging over the two independent trajectories (see Fig. 9).

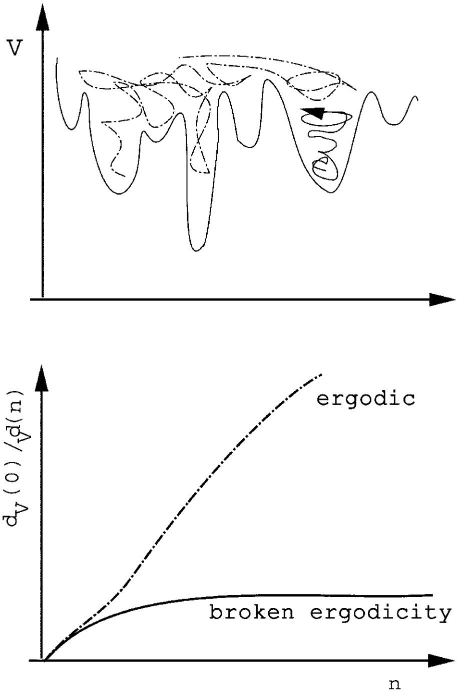

Fig. 9. Schematic representation of two trajectories on a rugged energy landscape. One trajectory is effectively trapped in a single basin while a second trajectory effectively samples the energy surface. Dynamics of the first kind will lead to "broken ergodicity" – two trajectories initially starting in distinct energy basins will not self-average on the simulation time scale. Dynamics of the second kind will lead to self-averaging and ergodic sampling.

We associate rapid and effective sampling of phase space with a large value of D_V. The choice of the potential energy metric is arbitrary. However, it has been shown to be a good measure of the extent of phase space sampling in a variety of systems. For the Monte Carlo and Molecular Dynamics algorithms described in this chapter, we vary the parameters of the algorithm to determine the optimal parameter set which maximizes D_V (see Fig. 10). This provides an effective means of optimizing the algorithm for the particular problem of interest by maximizing the rate of self-averaging [14].

8 Tsallis Statistics and Feynman Path Integral Quantum Mechanics

Our initial presentation of the origin of the Tsallis statistical distributions began with the expression for the classical density distribution. A similar starting

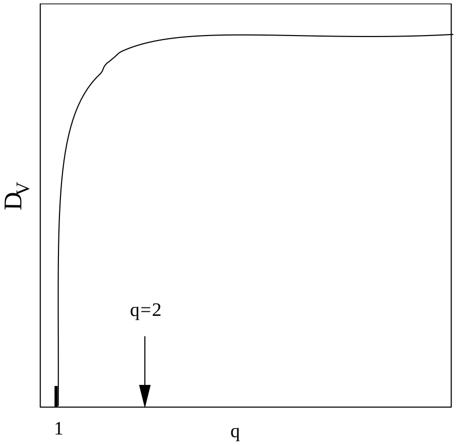

Fig. 10. Schematic representation of the generalized diffusion constant $D_{\mathcal{V}}$ as a function of the parameter q. For many applications, the general features of (1) a dramatic increase in the rate of self-averaging for $q > 1$ relative to $q = 1$ and (2) a saturation in the rate for higher values of q are noted.

point – that of the quantum mechanical thermal density matrix – is commonly used to derive the path integral formulation of quantum statistical mechanics. In this section we explore the possibility of a generalized quantum statistical mechanics that may prove useful as a computational tool for enhanced sampling and optimization in complex systems.

The starting point for the calculation of any equilibrium average of a many body quantum system is the density matrix. Written in the position representation

$$\rho(\mathbf{r}, \mathbf{r}'; \beta) = \langle \mathbf{r} | e^{-\beta \mathcal{H}} | \mathbf{r}' \rangle. \tag{87}$$

The equilibrium average of an operator $\mathcal{A}$ can be written

$$\langle \mathcal{A} \rangle = \frac{1}{Z} \int d\mathbf{r} \int d\mathbf{r}' \rho(\mathbf{r}, \mathbf{r}'; \beta) \langle \mathbf{r} | \mathcal{A} | \mathbf{r}' \rangle \tag{88}$$

where the partition function is

$$Z(\beta) = \int d\mathbf{r} \rho(\mathbf{r}, \mathbf{r}; \beta). \tag{89}$$

In the standard formulation of Feynman path integrals [8,53], the density matrix operator $\exp(-\beta\mathcal{H})$ is rewritten

$$e^{-\beta \mathcal{H}} = \left(e^{-\beta \mathcal{H}/P} \right)^P. \tag{90}$$

Interpreting the exponential as a propagator, the ratio β/P is often referred to as a step in imaginary time. In the position representation, this form of the density matrix becomes

$$\rho(\mathbf{r}_0, \mathbf{r}_P; \beta) = \int d\mathbf{r}_1 \dots d\mathbf{r}_{P-1} \rho(\mathbf{r}_0, \mathbf{r}_1; \beta) \dots \rho(\mathbf{r}_{P-1}, \mathbf{r}_P; \beta) \tag{91}$$

which is *exact* for any value of $P \geq 1$.

To make this expression useful for the numerical simulation of quantum systems, the operator expression [54]

$$e^{-\frac{\beta}{P}\mathcal{H} + \frac{1}{2}\left(\frac{\beta}{P}\right)^2 [\mathcal{T}, \mathcal{V}]} = e^{-\frac{\beta}{P}\mathcal{T}} e^{-\frac{\beta}{P}\mathcal{V}} \tag{92}$$

where $\mathcal{H} = \mathcal{T} + \mathcal{V}$ and $[\mathcal{T}, \mathcal{V}] = \mathcal{T}\mathcal{V} - \mathcal{V}\mathcal{T}$ is employed to prove that

$$e^{-\beta\mathcal{H}} = \lim_{P \to \infty} \left[e^{-\frac{\beta}{P}\mathcal{T}} e^{-\frac{\beta}{P}\mathcal{V}} \right]^P . \tag{93}$$

In what is known as the primitive approximation, the density matrix is written

$$\rho(\mathbf{r}_0, \mathbf{r}_1; \beta) \approx \int d\mathbf{r}_{\frac{1}{2}} \langle \mathbf{r}_0 | e^{-\frac{\beta}{P}\mathcal{T}} | \mathbf{r}_{\frac{1}{2}} \rangle \langle \mathbf{r}_{\frac{1}{2}} | e^{-\frac{\beta}{P}\mathcal{V}} | \mathbf{r}_1 \rangle \tag{94}$$

where the kinetic energy operator can be expressed in an integral approximation to an eigenfunction expansion as

$$\langle \mathbf{r}_0 | e^{-\frac{\beta}{P}\mathcal{T}} | \mathbf{r}_{\frac{1}{2}} \rangle \approx \left(\frac{mP}{2\hbar^2\beta} \right)^{-3N/2} e^{-\frac{mP}{2\hbar^2\beta}(\mathbf{r}_0 - \mathbf{r}_{\frac{1}{2}})^2} \tag{95}$$

and the potential energy operator is simply

$$\langle \mathbf{r}_{\frac{1}{2}} | e^{-\frac{\beta}{P}\mathcal{V}} | \mathbf{r}_1 \rangle = e^{-\frac{\beta}{P}\mathcal{V}(\mathbf{r}_{\frac{1}{2}})} \delta(\mathbf{r}_1 - \mathbf{r}_{\frac{1}{2}}). \tag{96}$$

In the limit of large enough P, we can combine these last two expressions and integrate over $\mathbf{r}_{\frac{1}{2}}$ to obtain a good approximation to $\rho(\mathbf{r}_0, \mathbf{r}_1; \beta)$. Combining this result with Eq. (91) we arrive at the path integral representation of the density matrix

$$\rho(\mathbf{r}_0, \mathbf{r}_P; \beta) \approx \int d\mathbf{r}_0 \dots d\mathbf{r}_{P-1} \left(\frac{mP}{2\hbar^2\beta} \right)^{-3NP/2} \times \exp\left(-\sum_{k=1}^{P} \left[\frac{mP}{2\hbar^2\beta} (\mathbf{r}_{k-1} - \mathbf{r}_k)^2 + \mathcal{V}(\mathbf{r}_k) \right] \right) . \tag{97}$$

This expression is only approximate for finite P and only appropriate when the size of the box in which the particle is contained is large compared to $\sqrt{\beta\hbar^2/mP}$.

In the standard interpretation of this result, the quantum density matrix of an N-body system is approximated as the classical equilibrium average of a

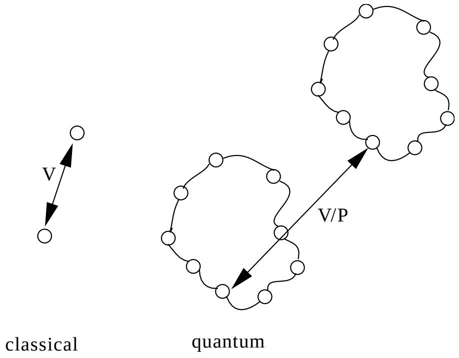

Fig. 11. Schematic showing the classical and path integral quantum mechanical representations of two interacting particles. In the classical case the particles interact with a potential $\mathcal{V}$ while in the path integral case the pseudoparticles interact with a potential $\mathcal{V}/P$.

system of N ring polymers. Each ring polymer is composed of P beads. The kth bead of polymer M is joined to its two neighboring beads by a Hooke's law potential with force constant $\kappa = mP/\hbar^2\beta$. For a pairwise additive potential energy function, the ring polymers interact in a pairwise manner. The kth bead of one ring polymer interacts with the kth bead of all other ring polymers through the scaled potential energy function $\mathcal{V}/P$.

8.1 Connection to Tsallis Statistics

A brief reflection on our derivation of the Tsallis thermostatistical probabilities and that of the primitive form of the path integral propagator suggests a hybrid formulation of the standard path integral expressions for the density matrix Eqs. (93) and (97). By defining the effective potential

$$\bar{\mathcal{V}} = \frac{P}{\beta} \ln\left(1 + \frac{\beta}{P}\mathcal{V}\right) \tag{98}$$

and

$$\lim_{P\to\infty} \bar{\mathcal{V}} = \mathcal{V} \tag{99}$$

it follows that

$$e^{-\beta\mathcal{H}} = e^{-\beta(\mathcal{T}+\mathcal{V})} = \lim_{P\to\infty} e^{-\beta(\mathcal{T}+\bar{\mathcal{V}})}. \tag{100}$$

Following the derivation above we find, up to the leading order terms in $1/P$, that

$$e^{-\frac{\beta}{P}(\mathcal{T}+\bar{\mathcal{V}})+\left(\frac{\beta}{P}\right)^2[\mathcal{T},\mathcal{V}]-\frac{1}{2}\left(\frac{\beta}{P}\right)^3[\mathcal{T},\mathcal{V}]} \approx e^{-\frac{\beta}{P}\mathcal{T}}\left(\frac{1}{1+\frac{\beta}{P}\mathcal{V}}\right) \tag{101}$$

where we have used the result that $[\mathcal{T},\bar{\mathcal{V}}]=[\mathcal{T},\mathcal{V}]-(\beta/2P)[\mathcal{T},\mathcal{V}^2]+\ldots$ and the fact that the limit of the product of two functions is the product of the limits [55,54]. Substituting this result in Eq. (100) leads to

$$e^{-\beta\mathcal{H}}=\lim_{P\to\infty}\left[e^{-\frac{\beta}{P}\mathcal{T}}\left(\frac{1}{1+\frac{\beta}{P}\mathcal{V}}\right)\right]^P . \tag{102}$$

This formula, expressed in the position representation, is the principal result of this section

$$\rho(\mathbf{r}_0,\mathbf{r}_P;\beta)\approx\int d\mathbf{r}_0\ldots d\mathbf{r}_{P-1}\left(\frac{mP}{2\hbar^2\beta}\right)^{-3NP/2} \tag{103}$$
$$\times\exp\left(-\sum_{k=1}^{P}\left[\frac{mP}{2\hbar^2\beta}(\mathbf{r}_{k-1}-\mathbf{r}_k)^2+\bar{\mathcal{V}}(\mathbf{r}_k)\right]\right).$$

This expression stands as an alternative to the standard path integral formulation of the density matrix in Eq. (97). In this case, for a pairwise additive potential energy function, the ring polymers interact in a pairwise manner. This is depicted in Fig. 11. The kth bead of one ring polymer interacting with the kth bead of all other ring polymers through the scaled, P-dependent, effective potential energy function $\bar{\mathcal{V}}/P$.

8.2 Feynman–Tsallis Path Integrals for the Simulation of Quantum Many-Body Systems

It is possible to apply Eq. (104) in path integral simulations of condensed phase systems. Ordinarily, in systems where quantum effects are most important, one must approach large values of P before there is convergence to the exact quantum mechanical average. As we have shown, if the necklace samples the Tsallis statistical distribution it should visit regions of higher potential energy more frequently and the distribution should be significantly more delocalized than the standard representation for the same number of beads in the necklace P. This implies that this representation might provide faster convergence to the quantum mechanical limit of $P \to \infty$ than the standard form.

It should be noted that while this formalism is intriguing, there is some cost for the potential gain in using the proposed propagator relative to the primitive form. The proposed form should lead to *enhanced sampling* of chain conformations relative to the primitive form. However, while the proposed form of the short time propagator is correct to *first order* in β/P, as it must be, the primitive form is correct to second order. Numerical tests will demonstrate the relative merits of each approach.

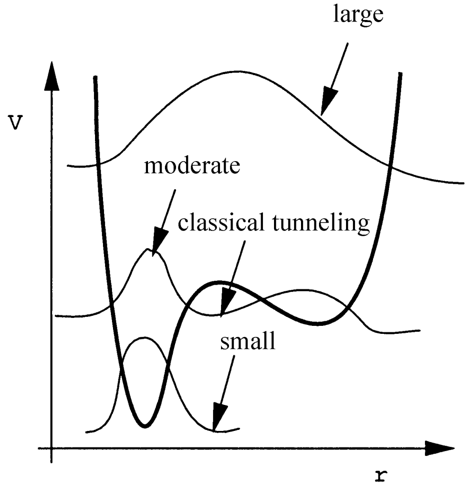

Fig. 12. Schematic showing the evolution of a density distribution under the equations of motion for wave packet dynamics.

9 Simulated Annealing Using Cauchy–Lorentz "Density Packet" Dynamics

In this section we propose a means of performing simulated annealing by propagation not a *point* in phase space but the continuous *density distribution* ρ represented by a mobile basis set. The motion of the basis set through phase space is optimally determined through a time-dependent variational principle. This technique has been explored earlier [56] in the context of the Gibbs-Boltzmann statistical mechanics. It is attractive for two reasons. (1) It provides enhanced sampling through the direct simulation of a non-local density distribution that explores *regions* of the potential surface at an instant. (2) It also provides the feature of *classical tunneling* as a density packet may move through a barrier whose height is greater than the average energy of the packet [22] (see Fig. 12). Here we generalize the method to the case of the non-extensive thermostatistics. We first treat the case of a single particle in d-dimensional space. The many-body problem extension is discussed at the end of the section.

The density distribution $\rho(\mathbf{r}, \mathbf{p}, t)$ gives the density of phase points at a point in phase space $(\mathbf{r}, \mathbf{p})$ at time t [57]. It represents an infinite number of identical systems, continuously distributed in phase space. The average of an observable property over this collection of systems is an integral over all phase space of that property weighted by $\rho(\mathbf{r}, \mathbf{p}, t)$; for example, the average position is $\int d^d\mathbf{r} \int d^d\mathbf{p}\ \rho(\mathbf{r}, \mathbf{p}, t)\ \mathbf{r}$. Calculating the time evolution of $\rho(\mathbf{r}, \mathbf{p}, t)$ is equivalent to calculating the trajectories of an infinite number of identical systems, differing only in their initial positions in phase space.

The time evolution of $\rho(\mathbf{r},\mathbf{p},t)$ is given by the Liouville equation

$$\frac{\partial}{\partial t}\rho(\mathbf{r},\mathbf{p},t) = -\mathcal{L}\rho(\mathbf{r},\mathbf{p},t) \tag{104}$$

where $\mathcal{L}$ is the Liouville operator

$$\mathcal{L} = \frac{\mathbf{p}}{M}\cdot\frac{\partial}{\partial \mathbf{r}} + F(\mathbf{r})\cdot\frac{\partial}{\partial \mathbf{p}} \tag{105}$$

and where $F(\mathbf{r})$ is the force and M the mass. $F(\mathbf{r})$, $\mathbf{r}$ and $\mathbf{p}$ are d-dimensional vectors. Below, we shall work with the positive square root of the density distribution, $\psi = \sqrt{\rho}$, which also satisfies the Liouville equation

$$\frac{\partial}{\partial t}\psi(\mathbf{r},\mathbf{p},t) = -\mathcal{L}\psi(\mathbf{r},\mathbf{p},t). \tag{106}$$

We work with $\psi(\mathbf{r},\mathbf{p},t)$ in part because the variational equations are more pleasing for $\psi(\mathbf{r},\mathbf{p},t)$, but more importantly because we can show that they conserve energy and normalization, regardless of the size of our basis set. In contrast, experience with most other mobile basis sets in quantum mechanics shows that most mobile basis sets do not rigorously conserve energy and normalization except in the limit that the basis set is complete, i.e., that the solution is exact. We may hope that many observable properties will not require an exact solution of $\psi(\mathbf{r},\mathbf{p},t)$, but only certain averaged properties of it. That is, we expect there to be a hierarchy of gross properties and more detailed, higher order properties [58]. Therefore it is extremely helpful to have energy and normalization conservation at every order of the basis set expansion. We shall demonstrate norm conservation later in this section.

Our variational principle, adapted from the Dirac-Frenkel variational principle for quantum dynamics [59], is a functional least-squares fit to the Liouville equation for ψ, with the error defined in terms of an error operator $\eta = \mathcal{L} + \partial/\partial t$. The error functional to be minimized is an integral of the square of the error $\eta\psi$ integrated over all time and all phase space

$$\begin{aligned} I &= \int dt \int d^d\mathbf{r} \int d^d\mathbf{p}\, \Big[\eta\psi(\mathbf{r},\mathbf{p},t)\Big]^2 \\ &= \int dt\, \langle \eta\psi(t)|\eta\psi(t)\rangle \end{aligned} \tag{107}$$

where we have adopted the Dirac ket notation

$$\langle f(t)|\mathcal{A}(\mathbf{r},\mathbf{p})|g(t)\rangle = \int d^d\mathbf{r} \int d^d\mathbf{p}\; f(\mathbf{r},\mathbf{p},t)\; \mathcal{A}(\mathbf{r},\mathbf{p})\; g(\mathbf{r},\mathbf{p},t). \tag{108}$$

Minimizing I by setting $\delta I/\delta\dot{\psi}(\mathbf{r}',\mathbf{p}',t')$ equal to zero yields the Liouville equation for $\psi(\mathbf{r}',\mathbf{p}',t')$. (The definition of the functional derivative is $\delta f(t)/\delta f(t') = \delta(t-t')$ [60]). Note that the derivative is with respect to $\dot{\psi}$ and not ψ itself because the Liouville equation is an initial value problem: given ψ, the equation gives $\dot{\psi}$.

In practice, we minimize the error functional I by first expanding ψ in terms of a basis set, and then setting the functional derivative of I with respect to all the first time derivatives of the parameters in the basis set equal to zero. In the limit that the basis set is complete (i.e., that the basis set is large enough), this procedure is equivalent to solving the Liouville equation exactly. In our past work we have used a basis set of Gaussian functions since the Gaussian functions represent the canonical ensemble probability for a locally harmonic potential when $q = 1$. Here our basis set is chosen to be a sum of Cauchy-Lorentz functions which appear in the non-extensive thermostatistics as the equilibrium statistical distribution for a locally harmonic potential when $q = 2$. Therefore, ψ can be written

$$\psi = \sum_{n=1}^{L} c_n \phi_n^2$$

where

$$\phi_n(\mathbf{r}) = \frac{1}{1 + a_n \mathbf{r}^2} \tag{109}$$

and the ϕ_n are Cauchy-Lorentz functions.

Setting $\delta I/\delta\dot{\chi}(t')$ to zero, where $\chi(t') = \{c_n, \ldots\}$ is a set of parameters defining the wave packet yields a set of coupled equations linear in the $\dot{\chi}(t')$. The equations are of the form [61]

$$\frac{\delta I}{\delta\dot{\chi}(t')} = 2\int dt \; \langle \frac{\delta\dot{\psi}(t)}{\delta\dot{\chi}(t')} | \eta\psi(t)\rangle = 0. \tag{110}$$

To show that the variational equations above preserve the normalization, we must show $\langle\psi|\dot{\psi}\rangle = 0$. From the variational condition $\langle\phi_n|\eta\psi\rangle = 0$ by Eq. (110), the norm preservation condition can be reëxpressed as $\langle\psi|\mathcal{L}\psi\rangle = 0$, which can be rewritten in terms of ρ (after dividing by 2) as

$$\int d^d\mathbf{r} \int d^d\mathbf{p} \; \mathcal{L}\rho = \int d^d\mathbf{r} \int d^d\mathbf{p} \left(\frac{\mathbf{p}}{M} \cdot \frac{\partial}{\partial \mathbf{r}} + F(\mathbf{r}) \cdot \frac{\partial}{\partial \mathbf{p}}\right) \rho. \tag{111}$$

since $\rho(\mathbf{r}, \mathbf{p})$ goes to zero at $\mathbf{r} = \pm\infty$ and $\mathbf{p} = \pm\infty$ integration of the above expression gives zero.

When many particles are introduced, the density distribution may in the first approximation be written as a product of the density distributions of the N individual particles

$$\rho(\mathbf{r}^N, \mathbf{p}^N, t) = \prod_{k=1}^{N} \rho_k(\mathbf{r}_k, \mathbf{p}_k, t). \tag{112}$$

This is in the spirit of the Hartree approximation of quantum mechanics. The force $F(\mathbf{r})$ is taken to be the integral of the force weighted by the total density distribution integrated over all the other particles; that is, $F(\mathbf{r})$ is averaged over the positions of all the other particles, a mean field force [62]. For a pairwise additive potential, the force on particle 1 due to particle k can be written in terms of the distribution function $\rho_k(\mathbf{r}_k, \mathbf{p}_k)$ for the kth particle

$$\langle F(\mathbf{r}_1)\rangle_k = \int d\mathbf{r}_k \int d\mathbf{p}_k \rho_k(\mathbf{r}_k, \mathbf{p}_k) F(|\mathbf{r}_1 - \mathbf{r}_k|). \tag{113}$$

The total force on a single particle will be the sum of the forces of interaction with all other atoms. Similarly, the potential energy can be written as a sum over the individual pair interaction energies

$$\langle \mathcal{V}\rangle_{ij} = \sum_{i>j=1}^{N} \int d\mathbf{r}_i \int d\mathbf{p}_i \int d\mathbf{r}_j \int d\mathbf{p}_j \; \rho_i(\mathbf{r}_i, \mathbf{p}_i) \; \rho_j(\mathbf{r}_j, \mathbf{p}_j) \mathcal{V}(|\mathbf{r}_i - \mathbf{r}_j|). \tag{114}$$

When the potential is not pairwise additive, the density distribution functions for two or more particles correlations must be used.

How might this method be applied? A set of Cauchy-Lorentz “density packets” can initially be parameterized to extend broadly over the potential energy surface. The force used in the Liouville operator may be the standard force or the scaled effective force of Eq. (80). The exact choice will depend on the nature of the system studied. The temperature or energy is initially taken to be high and the packets are propagated in time. The temperature is then then gradually lowered as the packet is annealed to a low lying basin on the potential energy surface. Similar annealing schemes using Gaussian packets have been successful in a variety of optimization problems [22,63]. However, Gaussian packets are relatively local functions. The use of Cauchy-Lorentz packets, which extend more broadly over the energy surface, should be even more effective in isolating low lying energy basins on the rugged energy landscapes of peptides and proteins.

We gratefully acknowledge the National Science Foundation for support and the Center for Scientific Computing and Visualization at Boston University for computational resources. JES thanks the Institute for Advanced Studies at The Hebrew University of Jerusalem for support and hospitality. We also thank Troy Whitfield for helpful comments.

References

1. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, New York, 1996).
2. J. P. Valleau and S. G. Whittington, in *Statistical Mechanics, Part A*, ed. B. J. Berne (Plenum Press, New York, 1976).
3. J. P. Valleau and G. M. Torrie, in *Statistical Mechanics, Part A*, ed. B. J. Berne (Plenum Press, New York, 1976).

4. B. J. Berne and J. E. Straub, Curr. Opin. Struc. Bio. **7**, 181 (1997).
5. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
6. E .M. F. Curado and C. Tsallis, J. Phys. A **24**, L69 (1991).
7. A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978); Z. Daróczy, Inf. Control **16**, 36 (1970).
8. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
9. C. Tsallis, Braz. J. Phys. **29**, 1 (1999).
10. J. E. Straub and I. Andricioaei, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, eds. P. Deuflhard, J. Hermans, B. Leimkuhler, A. Mark, S. Reich, and R. D. Skeel (Springer-Verlag, Berlin, 1999).
11. J. Maddox, Nature **365**,103 (1993).
12. B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
13. K. K. Bhattacharya and J. P. Sethna, Phys. Rev. E **57**, 2553 (1998).
14. I. Andricioaei and J. E. Straub, J. Chem. Phys. **107**, 9117 (1997).
15. C. Tsallis, S. V. F. Levy, A. M. C. de Souza, and R. Maynard, Phys. Rev. Lett. **75**, 3589 (1995).
16. R. Zwanzig, J. Chem. Phys. **103**, 9397 (1995).
17. Such “global” master equations have recently gained attention in the study of protein folding. See, for example, E. I. Shakhnovich and A. M. Gutin, Eurphys. Lett. bf 9, 569 (1989); J. G. Saven, J. Wang, and P. G. Wolynes, J. Chem. Phys. **101**, 11037 (1994).
18. E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing* (Wiley, New York, 1989).
19. H. Frauenfelder and P. G. Wolynes, Physics Today, p. 58 (February, 1994).
20. J. E. Straub and J.-K. Choi, J. Phys. Chem. **98**, 10978 (1994); J. E. Straub and D. Thirumalai, Proc. Natl. Acad. Sci. USA **90**, 809 (1993); J. E. Straub and D. Thirumalai, Proteins **15**, 360 (1993); J. E. Straub, A. Rashkin, and D. Thirumalai, J. Am. Chem. Soc. **98**, 10978 (1994).
21. S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Science **220**, 671 (1983).
22. J. E. Straub, in *Recent developments in theoretical studies of proteins*, ed. R. Elber (World Scientific, Singapore), in press.
23. S. H. Northrup and J. A. McCammon, Biopolymers **19**, 1001 (1980).
24. N. Metropolis, A. W. Rosenbluth, N. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
25. C. Pangali, M. Rao, and B. J. Berne, Chem. Phys. Lett. **55**, 413 (1978).
26. P. J. Rossky, J. D. Doll, and H. L. Friedman, J. Chem. Phys. **69**, 4628 (1978).
27. R .H. Swendsen, D. Bouzida, and S. Kumar, Technical Report 91-121-NAMS-27, Carnegie Mellon University (1991).
28. C. Tsallis and D. A. Stariolo, Physica A **233**, 395 (1996).
29. H. Szu and R. Hartley, Phys. Lett. A **122**, 157 (1987).
30. T. J .P. Penna, Phys. Rev. E **51**, R1 (1995).
31. I. Andricioaei and J. E. Straub, Phys. Rev. E **53**, R3055 (1996).
32. If the acceptance probability is chosen to be $p^{1/q(T)}$, where p is the acceptance probability in (59), then the limiting probability distribution is p_q. However, during an equilibrium simulation, average properties calculated as $O = \sum p_i O_i$ will diverge, whereas averages based on the acceptance (59), $O = \sum p_i^q O_i$, will converge. As a simple example, consider the average potential energy of a harmonic oscillator and let $q = 2$. Then $\int_0^{1nfty} p(x) x^2 dx \propto \int_0^\infty (1 + \beta x^2)^{-1} x^2 dx$ diverges, while $\int_0^\infty p^2(x) x^2 dx \propto \int_0^\infty (1 + \beta x^2)^{-2} x^2 dx$ converges.

33. P. A. Alemany and D. H. Zanette, Phys. Rev. E **49**, R956 (1994).
34. M. A. Moret, P. G. Pascutti, P. M. Bisch, and K. C. Mundim, J. Comput. Chem. **19**, 647 (1998).
35. K. C. Mundim and D. E. Ellis, Braz. J. Phys. **29**, 199 (1999).
36. R. Salazar and R. Toral, J. Stat. Phys. **89**, 1047 (1999).
37. U. H. E. Hansmann and Y. Okamoto, Braz. J. Phys. **29**, 187 (1999).
38. Y. Xiang, D. Y. Sun, W. Fan, and X. G. Gong, Phys. Lett. A **233**, 216 (1997).
39. H. Nishimori and J. Inoue, J. Phys. A **31**, 5661 (1998).
40. I. Andricioaei and J. E. Straub, Physica A **247**, 553 (1997).
41. G. W. Cunningham and P. H. E. Meijer, J. Comput. Phys. **20**, 50 (1976).
42. R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
43. A. R. Lima, J. S. S. Martins, and T. J. P. Penna, Physica A **268**, 533 (1999).
44. P. Salamon, J. D. Nulton, J. R. Harland, J. Pedersen, G. Ruppeiner, and L. Liao, Comput. Phys. Commun. **49**, 423 (1988).
45. J. Ma and J. E. Straub, J. Chem. Phys. **101**, 533 (1994); *ibid.* **103**, 9113 (1995).
46. E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
47. C. J. Geyer and E. A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
48. Y. Pak and S. M. Wang, J. Chem. Phys. **111**, 4359 (1999).
49. Y. Pak and S. Wang, J. Phys. Chem. B **104**, 354 (2000).
50. R. D. Mountain and D. Thirumalai, J. Phys. Chem. **93**, 6975 (1989).
51. J. E. Straub and D. Thirumalai, Proc. Natl. Acad. Sci. USA **90**, 809 (1993); J. E. Straub and D. Thirumalai, Proteins **15**, 360 (1993).
52. J. E. Straub, A. Rashkin, and D. Thirumalai, J. Am. Chem. Soc. **116**, 2049 (1994).
53. D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
54. A. K. Rajagopal and C. Tsallis, Phys. Lett. A **257**, 283 (1999).
55. J. E. Straub and I. Andricioaei, Braz. J. Phys. **29**, 179 (1999).
56. J. Ma, Ph.D. thesis (Boston University, 1996).
57. R. C. Tolman, *The principles of Statistical Mechanics* (Dover, New York, 1979).
58. D. Hsu and D. Coker, J. Chem. Phys. **96**, 4266 (1992).
59. J. Frenkel, *Wave Mechanics: Advanced General Theory* (Clarendon Press, Oxford, 1934). In the Dover edition (1950), see p. 253. Frenkel cites an appendix in the Russian edition of Dirac's *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1930).
60. See, for example, Chapter 2 in H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Massachusetts, 1980).
61. For readers unfamiliar with functional derivatives, simply drop the time integral in I and take partial derivatives of I with respect to the $\dot{\chi}$ as though they were normal variables, rather than functions. See, for example, Chapter 2 in Ref. [60].
62. A. D. McLachlan, Mol. Phys. **8**, 39 (1964).
63. P. Amara and J. E. Straub, J. Phys. Chem. **99**, 14840 (1995).

V. Correlation Induced by Nonextensivity and the Zeroth Law of Thermodynamics

S. Abe

College of Science and Technology, Nihon University, Funabashi, Chiba 274-8501, Japan

Abstract. To establish the zeroth law of thermodynamics in nonextensive statistical mechanics, it is essential to realize factorization of the Tsallis probability distribution. Here, the factorizability and the property of correlation induced by nonextensivity are discussed. Employing the classical gas model, it is shown how the correlation is suppressed in the thermodynamic limit.

A nonextensive generalization of Boltzmann-Gibbs statistical mechanics proposed by Tsallis is based on the following postulate for the entropy functional [1]:

$$S_q[p] = \frac{k}{1-q}\left(\sum_n p_n^q - 1\right), \tag{1}$$

where $\{p_n\}$ is the normalized probability distribution of the system under consideration and q is a positive parameter describing the degree of nonextensivity. k is the Boltzmann constant and is henceforth set equal to unity for simplicity. This generalized entropy is concave and satisfies the H-theorem. The discrete notation is employed here but the case of continuous configuration can be treated in an analogous way. The equilibrium distribution is obtained through maximum entropy principle under the constraints on normalization of $\{p_n\}$ and the normalized q-expectation value of the energy of the system. The latter is given as follows [2]:

$$\begin{aligned} U_q &= < H >_q \\ &= \frac{1}{c_q}\sum_n p_n^q\, \varepsilon_n, \end{aligned} \tag{2}$$

$$c_q \equiv \sum_n p_n^q, \tag{3}$$

where H and ε_n are the system Hamiltonian and its value in the nth configuration, respectively. Then, maximization of the Tsallis entropy in Eq. (1) yields the equilibrium probability distribution

$$p_n^{(e)} = \frac{1}{Z_q(\beta)}\Big[1 - (1-q)(\beta/c_q)(\varepsilon_n - U_q)\Big]^{1/(1-q)}, \tag{4}$$

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 237–242, 2001.

where

$$Z_q(\beta) = \sum_n \Big[1 - (1-q)(\beta/c_q)(\varepsilon_n - U_q)\Big]^{1/(1-q)}. \tag{5}$$

Here, β is the Lagrange multiplier associated with the constraint in Eq. (2). From eqs. (3)-(5), one finds the identical relation

$$c_q = \Big[Z_q(\beta)\Big]^{1-q}. \tag{6}$$

Actually, β is identified with the inverse temperature, since the equilibrium Tsallis entropy $S_q^{(e)} = S_q[p^{(e)}]$ and the generalized internal energy satisfy U_q the thermodynamic relation

$$\frac{\partial S_q^{(e)}}{\partial U_q} = \beta. \tag{7}$$

Now, assume the total system be divided into two subsystems, A and B, and the joint probability distribution be factorized, that is,

$$p_{mn}(A,B) = p_m(A) p_n(B). \tag{8}$$

Then the Tsallis entropy in Eq. (1) satisfies

$$S_q(A,B) = S_q(A) + S_q(B) + (1-q) S_q(A) S_q(B), \tag{9}$$

where $S_q(A) \equiv S_q[p(A)]$ and so on. As pointed out by Rajagopal in his contribution to this volume, the maximum entropy condition on the total entropy $S_q^{(e)}(A,B)$ at the thermal equilibrium state leads to

$$\begin{aligned} 0 &= \delta S_q^{(e)}(A,B) \\ &= \frac{\partial S_q^{(e)}(A)}{\partial U_q(A)} \Big[1 + (1-q) S_q^{(e)}(B)\Big] \delta U_q(A) \\ &\quad + \frac{\partial S_q^{(e)}(B)}{\partial U_q(B)} \Big[1 + (1-q) S_q^{(e)}(A)\Big] \delta U_q(B). \end{aligned} \tag{10}$$

Since the total internal energy is fixed, holds the identity

$$\delta U_q(A) + \delta U_q(B) = 0. \tag{11}$$

Therefore, from Eq. (10), it follows that

$$\begin{aligned} &\frac{\partial S_q^{(e)}(A)}{\partial U_q(A)} \Big[1 + (1-q) S_q^{(e)}(A)\Big]^{-1} \\ &\qquad = \frac{\partial S_q^{(e)}(B)}{\partial U_q(B)} \Big[1 + (1-q) S_q^{(e)}(B)\Big]^{-1}, \end{aligned} \tag{12}$$

which can also be expressed as

$$\tilde{\beta}_q(A) = \tilde{\beta}_q(B), \tag{13}$$

where $\tilde{\beta}_q$ is the renormalized inverse temperature: $\tilde{\beta}_q = \beta/c_q$.

Equation (13) defines the equivalence relation between temperatures of the subsystems in equilibrium, on which the zeroth law of thermodynamics in nonextensive statistical mechanics can be based. An essential point regarding this relation is the factorization in Eq. (8). In the extensive limit $q \to 1$, the Tsallis entropy converges to the Boltzmann-Shannon entropy, $S_q[p] \to S[p] = -\sum_n p_n \ln p_n$, which satisfies the additivity as seen from Eq. (9). The corresponding equilibrium probability distribution has the exponential form, $p_n^{(e)} \sim \exp(-\beta\varepsilon_n)$, which can be obtained from Eq. (4) in the above limit. Therefore, if the subsystems are dynamically independent, i.e., $H = H(A) + H(B)$, then the joint probability distribution is obviously factorized. In the nonextensive case, however, the probability distribution in Eq. (4) is not apparently factorized and accordingly correlation is always induced. Thus the validity of the equivalence relation in Eq. (13) (and therefore the zeroth law of thermodynamics) is highly nontrivial.

In what follows, we report a recent result [3] in the study of the property of correlation induced by nonextensivity. Specifically we consider the classical ideal gas model. Since the gas particles in this model are all dynamically independent, the property of correlation induced by nonextensivity will be revealed in a pure manner.

The Hamiltonian of the system is given by

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{P}_i^2}{2m}, \tag{14}$$

where m is the common particle mass, $\boldsymbol{P}_i$ the D-dimensional momentum of the ith particle and N the number of the particles. In this case, the probability distribution which maximizes the Tsallis entropy becomes

$$\begin{aligned} & f^{(e)}\,(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N) \\ & = \frac{1}{Z_q(\beta)} \left[1 - (1-q)(\beta/c_q) \left(\sum_{i=1}^{N} \frac{\boldsymbol{P}_i^2}{2m} - U_q \right) \right]^{1/(1-q)} . \end{aligned} \tag{15}$$

The generalized partition function in this equation is

$$\begin{aligned} Z_q(\beta) = {} & \frac{V^N}{N! h^{DN}} \int \prod_{i=1}^{N} d^D \boldsymbol{P}_i \\ & \times \left[1 - (1-q)(\beta/c_q) \left(\sum_{i=1}^{N} \frac{\boldsymbol{P}_i^2}{2m} - U_q \right) \right]^{1/(1-q)} , \end{aligned} \tag{16}$$

where V and h are the spatial volume of the system and the linear size of the elementary cell in phase space, respectively. The normalization condition on $f^{(e)}$

is

$$\frac{V^N}{N!h^{DN}} \int \prod_{i=1}^{N} d^D \boldsymbol{P}_i \ f^{(e)}(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N) = 1. \tag{17}$$

The normalized q-expectation value of a physical quantity $Q = Q(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ is defined by

$$\begin{aligned} < Q >_q = \frac{V^N}{c_q N!h^{DN}} \int \prod_{i=1}^{N} d^D \boldsymbol{P}_i \ Q(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N) \\ \times \Big[f^{(e)}(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N) \Big]^q, \end{aligned} \tag{18}$$

where

$$c_q = \frac{V^N}{N!h^{DN}} \int \prod_{i=1}^{N} d^D \boldsymbol{P}_i \Big[f^{(e)}(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N) \Big]^q. \tag{19}$$

It is possible to show that the present model is well defined in the two ranges of the nonextensivity parameter q: $0 < q < 1$ and $1 < q < 1 + 2/(DN)$. Since N is very large, the latter essentially collapses to the extensive limit $q \to 1 + 0$. Therefore, in what follows, we restrict ourselves to the range

$$0 < q < 1. \tag{20}$$

In this case, the probability distribution in Eq. (15) can be written in the form

$$f^{(e)}(\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N) = C \left(\varLambda^2 - \sum_{i=1}^{N} \boldsymbol{P}_i^2 \right)^{1/(1-q)}, \tag{21}$$

where C is the normalization factor and

$$\varLambda^2 = \frac{2mc_q}{(1-q)\beta} \left[1 + (1-q) \frac{\beta U_q}{c_q} \right]. \tag{22}$$

Thus, the support of the distribution function is the interior of the hypersphere with the radius $\varLambda$ in the DN-dimensional momentum space.

Using Eq. (15), the generalized partition function Z_q in Eq. (16), the generalized internal energy U_q in Eq. (18) with $Q = H$, and the factor c_q in Eq. (19) are respectively calculated to be

$$\begin{aligned} Z_q(\beta) = \frac{\varGamma\left(\frac{2-q}{1-q}\right)}{\varGamma\left(\frac{2-q}{1-q} + \frac{DN}{2}\right)} \frac{V^N}{N!h^{DN}} \left[\frac{2\pi m c_q}{(1-q)\beta} \right]^{DN/2} \\ \times \left[1 + (1-q) \frac{\beta U_q}{c_q} \right]^{1/(1-q)+DN/2}, \end{aligned} \tag{23}$$

$$U_q = \frac{\frac{DN}{2\beta}}{[Z_q(\beta)]^q} \frac{\Gamma\left(\frac{2-q}{1-q}\right)}{\Gamma\left(\frac{2-q}{1-q} + \frac{DN}{2}\right)} \frac{V^N}{N! h^{DN}} \left[\frac{2\pi m c_q}{(1-q)\beta}\right]^{DN/2} \times \left[1 + (1-q)\frac{\beta U_q}{c_q}\right]^{1/(1-q)+DN/2}, \tag{24}$$

$$c_q = \frac{1}{[Z_q(\beta)]^q} \frac{\Gamma\left(\frac{1}{1-q}\right)}{\Gamma\left(\frac{1}{1-q} + \frac{DN}{2}\right)} \frac{V^N}{N! h^{DN}} \left[\frac{2\pi m c_q}{(1-q)\beta}\right]^{DN/2} \times \left[1 + (1-q)\frac{\beta U_q}{c_q}\right]^{q/(1-q)+DN/2}. \tag{25}$$

A point is that equations (24) and (25) give rise to the relation

$$\frac{\beta U_q}{c_q} = \frac{DN}{2}. \tag{26}$$

This enables us to evaluate all physical quantities in the closed forms. For example, from eqs. (6), (25) and (26), the solution for c_q is obtained as follows:

$$c_q = \left\{ \frac{\Gamma\left(\frac{1}{1-q}\right)}{\Gamma\left(\frac{1}{1-q} + \frac{DN}{2}\right)} \frac{V^N}{N! h^{DN}} \left[\frac{2\pi m}{(1-q)\beta}\right]^{DN/2} \times \left[1 + (1-q)\frac{DN}{2}\right]^{q/(1-q)+DN/2} \right\}^{\frac{2(1-q)}{2-(1-q)DN}}. \tag{27}$$

Now, we wish to discuss the property of correlation between the Hamiltonians of the ith and jth particles, $H_i = \boldsymbol{P}_i^2/2m$ and $H_j = \boldsymbol{P}_j^2/2m$. For this purpose, we define the generalized variance, covariance, and correlation coefficient respectively as follows:

$$(\Delta_q H_i)^2 = < H_i^2 >_q - < H_i >_q^2, \tag{28}$$

$$C_q(H_i, H_j) = < H_i H_j >_q - < H_i >_q < H_j >_q, \tag{29}$$

$$\rho_q(H_i, H_j) = \frac{C_q(H_i, H_j)}{\sqrt{(\Delta_q H_i)^2 (\Delta_q H_j)^2}}. \tag{30}$$

Straightforward calculation shows

$$(\Delta_q H_i)^2 = \frac{c_q^2}{2\beta^2} \times \frac{2D + (1-q)D^2(N-1)}{4 - 2q + (1-q)DN}, \tag{31}$$

$$C_q(H_i, H_j) = -\frac{c_q^2}{2\beta^2} \times \frac{(1-q)D^2}{4-2q+(1-q)DN}, \tag{32}$$

$$\rho_q(H_i, H_j) = -\frac{(1-q)D}{2+(1-q)D(N-1)}. \tag{33}$$

ρ_q vanishes in the extensive limit $q \to 1-0$, as it should do. An important point is that the correlation is suppressed also in the thermodynamic limit, N, $V \to \infty$ with a fixed density N/V:

$$\rho_q(H_i, H_j) \sim -\frac{1}{N}, \tag{34}$$

showing factorization of the distribution function.

The above result suggests that the zeroth law of thermodynamics can be established in nonextensive statistical mechanics in the thermodynamic limit based on the equivalence relation prescribed by Eq. (13). The present discussion is however entirely dependent on the simple classical ideal gas model. It is extremely important to extend the idea to the case of a general Hamiltonian system.

Acknowledgements

Enlightening discussions with A. R. Plastino, A. K. Rajagopal, and C. Tsallis are gratefully acknowledged.

References

1. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
2. C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica A **261**, 534 (1998).
3. S. Abe, Physica A **269**, 403 (1999).

VI. Dynamic and Thermodynamic Stability of Nonextensive Systems

J. Naudts[1] and M. Czachor[1,2,3]

[1] Departement Natuurkunde, Universiteit Antwerpen, UIA,
Universiteitsplein 1, 2610 Antwerpen, Belgium
[2] Katedra Fizyki Teoretycznej i Metod Matematycznych, Politechnika Gdańska
ul. Narutowicza 11/12, 80—952 Gdańsk, Poland
[3] Arnold Sommerfeld Institute for Mathematical Physics
Technical University of Clausthal, 38678 Clausthal-Zellerfeld, Germany

Abstract. We prove uniqueness of the equilibrium states of q-thermodynamics for $1 < q \leq 2$, using both the unnormalized and the recently introduced normalized energy functional. The proof follows from thermodynamic stability of equilibrium states in case of the unnormalized energy. Dynamic stability is shortly discussed.

1 Introduction

The recent [1] introduction of a normalized energy functional in the formalism of non-extensive thermodynamics eliminates some anomalies which plagued the theory. The price one pays is that it is no longer straightforward to establish the existence and uniqueness of equilibrium states. Therefore the present paper presents a rather formal treatment of the problem, based on convexity arguments.

We consider quantum systems with a finite number of degrees of freedom. The state of the system is described by a density matrix ϱ on a Hilbert space $\mathcal{H}$. The Hamiltonian H is bounded from below, and has isolated eigenvalues with finite multiplicity. For such systems one expects thermodynamic stability of the equilibrium state, which is a stronger statement than just existence and uniqueness. Still stronger is the concept of dynamic stability. We can present only preliminary results concerning each of these concepts.

The next section introduces Tsallis entropy and the different choices of the energy functional. Sect. 3 introduces both linear and nonlinear dynamics. Next a short discussion is given of dynamic stability. Sect. 5 discusses thermodynamic stability in the case of unnormalized energy. It is followed by the proof of the first theorem. Finally, Sect. 7 discusses uniqueness of the equilibrium state in the case of normalized energy, followed by the proof of the second theorem.

2 Nonextensive Thermodynamics

The canonical ensemble of thermodynamics is based on a pair consisting of an entropy functional $\mathcal{S}(\varrho)$ and an energy functional $\mathcal{U}(\varrho)$. Together they determine

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 243–252, 2001.

the free energy $\mathcal{F}(\varrho)$ at temperature T by

$$\mathcal{F}(\varrho) = \mathcal{U}(\varrho) - T\mathcal{S}(\varrho)\ . \tag{1}$$

The equilibrium state is the density matrix which minimizes (1).

The Tsallis entropy [2] is defined by

$$\mathcal{S}_q(\varrho) = k_B \frac{1 - \operatorname{Tr} \varrho^q}{q-1}\ . \tag{2}$$

If $q > 0$, then $\mathcal{S}_q(\varrho) \geq 0$ for all ϱ. In the limit $q = 1$ the Shannon entropy is recovered. Originally [2] the standard expression for the energy $\mathcal{U}(\varrho)$ was used. Given a Hamiltonian H let

$$\mathcal{U}_q^{(1)}(\varrho) = \operatorname{Tr} \varrho H\ . \tag{3}$$

Quite soon [3] it was observed that it is more natural to replace ϱ by ϱ^q in the previous expression, i.e. one introduced

$$\mathcal{U}_q^{(2)}(\varrho) = \operatorname{Tr} \varrho^q H\ . \tag{4}$$

Recently [1], the following modification has been proposed

$$\mathcal{U}_q^{(3)}(\varrho) = \frac{\operatorname{Tr} \varrho^q H}{\operatorname{Tr} \varrho^q}\ . \tag{5}$$

The effect on thermodynamics of replacing $\mathcal{U}_q^{(2)}$ by $\mathcal{U}_q^{(3)}$ is not so big. In fact, the equilibrium states remain the same but the temperature scales are modified. Using $\mathcal{U}_q^{(3)}$ the temperature is always positive and the low temperature limit corresponds to the ground state of the system, as one is used to. Using $\mathcal{U}_q^{(2)}$ the temperature has a lower bound, which can even be negative, and the formalism is not invariant under shifts of the origin of the energy scale. Hence $\mathcal{U}_q^{(3)}$ is the preferred expression. However $\mathcal{U}_q^{(2)}$ is used below as an intermediate step in the study of $\mathcal{U}_q^{(3)}$.

3 Nonlinear von Neumann Equation

There are two ways to associate a quantum mechanical time evolution with the energy functional (5). Introduce a new density matrix σ by

$$\sigma = \frac{\varrho^q}{\operatorname{Tr} \varrho^q}\ . \tag{6}$$

In terms of σ the conventional expression for the average energy $\mathcal{U}_q^{(3)}(\varrho) = \operatorname{Tr} \sigma H$ is obtained. The corresponding choice of time evolution is that of the Heisenberg picture

$$\sigma(t) = e^{-i\hbar^{-1}Ht} \sigma e^{i\hbar^{-1}Ht}\ . \tag{7}$$

It satisfies the von Neumann equation

$$i\hbar\frac{\mathrm{d}\sigma}{\mathrm{d}t}=[H,\sigma]\,. \tag{8}$$

Alternatively, an equation of motion is derived following the procedure of [4] (note that in [4] the energy functional $\mathcal{U}_q^{(2)}$ was used instead of $\mathcal{U}_q^{(3)}$). Write (5) in a 1-homgeneous way

$$\mathcal{U}_q^{(3)}(\varrho)=(\,\mathrm{Tr}\,\varrho)\frac{\mathrm{Tr}\,\varrho^q H}{\mathrm{Tr}\,\varrho^q}\,. \tag{9}$$

Introduce an effective Hamiltonian $\hat{H}(\varrho)$ by

$$\hat{H}(\varrho)=\frac{\delta}{\delta\varrho}\mathcal{U}_q^{(3)}(\varrho)\,. \tag{10}$$

The precise meaning of this expression is that in a suitable orthonormal basis $\{\psi_n\}_n$ one has

$$\langle\psi_n|\hat{H}(\varrho)|\psi_m\rangle=\frac{\partial\mathcal{U}_q^{(3)}(\varrho)}{\partial\varrho_{n,m}} \tag{11}$$

with $\varrho_{n,m}=\langle\psi_m|\varrho|\psi_n\rangle$. The 1-homogeneity implies

$$\mathcal{U}_q^{(3)}(\varrho)=\,\mathrm{Tr}\,\varrho\hat{H}(\varrho)\,. \tag{12}$$

The corresponding choice of time evolution is described by the nonlinear von Neumann equation

$$i\hbar\frac{\mathrm{d}\varrho}{\mathrm{d}t}=[\hat{H}(\varrho),\varrho]\,. \tag{13}$$

An important property of (13) is that it conserves the spectrum of ϱ. Note that the effective Hamiltonian satisfies

$$[\hat{H}(\varrho),\varrho]=\frac{\mathrm{Tr}\,\varrho}{\mathrm{Tr}\,\varrho^q}[H,\varrho^q]\,. \tag{14}$$

If $[H,\varrho]=0$ then it is easy to calculate $\hat{H}(\varrho)$ from the definition (10). One obtains, putting $\mathrm{Tr}\,\varrho=1$,

$$\hat{H}(\varrho)=\mathcal{U}_q^{(3)}(\varrho)\mathbf{1}+q\varrho^{q-1}\frac{H-\mathcal{U}_q^{(3)}(\varrho)\mathbf{1}}{\mathrm{Tr}\,\varrho^q}\,. \tag{15}$$

Both equations of motion (8) and (13) are important in the context of the $q\leftrightarrow 1/q$ duality introduced in [1] (see [5]). Nonlinear equations of the type (13) occur often in the literature, although not with $\hat{H}(\varrho)$ of the form (15).

4 Dynamic Stability

Sufficient conditions for a density matrix ϱ_0 to be dynamically stable are that ϱ_0 is invariant under the time evolution and that there exists a norm $||\cdot||$ such that the distance $||\varrho(t)-\varrho_0||$ does not depend on time for any other density matrix ϱ. See [6] for a definition of dynamic stability in the context of classical mechanics, or [7] for an introductory text. Such an invariant norm exists, as is shown below. Hence all invariant density matrices satisfying $[H,\varrho_0]=0$ are dynamically stable.

Fix a Hamiltonian H with discrete spectrum, and a constant γ such that $\mathbf{1}+\gamma H$ is strictly positive (by which it is meant that $1+\gamma E_{\mathrm{m}}>0$ with E_{m} the ground state energy of H). Introduce an operator norm $||\cdot||_\gamma$ by

$$||A||_\gamma^2 = \operatorname{Tr}(\mathbf{1}+\gamma H)|A|^2 . \tag{16}$$

Then one has for any pair of density matrices ϱ,ϱ_0, with $[\varrho_0,H]=0$,

$$\begin{aligned} ||\varrho-\varrho_0||_\gamma^2 &= \operatorname{Tr}(\mathbf{1}+\gamma H)(\varrho-\varrho_0)^2 \\ &= \operatorname{Tr}\varrho^2 - 2\operatorname{Tr}\varrho\varrho_0 + \operatorname{Tr}\varrho_0^2 \\ &\qquad +\gamma\operatorname{Tr}H\varrho^2 - 2\gamma\operatorname{Tr}H\varrho\varrho_0 + \gamma\operatorname{Tr}H\varrho_0^2 . \end{aligned} \tag{17}$$

Each of the terms in the latter expression is time-invariant, for any of the dynamics discussed in the previous section. Hence $||\varrho(t)-\varrho_0||_\gamma$ does not depend on time t. Therefore any invariant density matrix ϱ_0 is dynamically stable w.r.t. any of the $||\cdot||_\gamma$ norms. Note that $||\cdot||_0$ is the Hilbert-Schmidt norm. The stability of invariant density matrices for the time evolution of the linear von Neumann equation w.r.t. the Hilbert-Schmidt norm belongs to the common knowledge.

The above result suggests that the dynamics of quantum systems is more regular than that of classical mechanics. In fact it is so regular that any invariant density matrix is dynamically stable w.r.t. an infinite set of norms. However, in classical mechanics it is implicitly understood that the energy functional $\mathcal{U}(\varrho)$ is continuous in the norm w.r.t. which ϱ_0 is dynamically stable. In the present context this extra condition is in general *not* satisfied because the Hamiltonian H can be unbounded. Hence dynamic stability in a more strict sense is still an open issue.

One way to obtain stronger results is by the energy-Casimir method (see [7]). The method starts by finding a conserved quantity $C(\varrho)$ (typically a Casimir of the Lie bracket which defines the dynamics) such that the function $\mathcal{U}_q(\varrho)+C(\varrho)$ is extremal at ϱ_0. In the present situation the obvious choice is $C(\varrho)=-T\mathcal{S}(\varrho)$ because the equilibrium density matrix ϱ_0 minimizes the free energy. The next step of the method is to find quadratic forms which bound the energy and the entropy from below and from above. In the quantum context this means that one should find quadratic forms Q_1 and Q_2 such that for all variations $\delta\varrho$ of ϱ_0

$$Q_1(\delta\varrho) \le \mathcal{U}_q(\varrho_0+\delta\varrho) - \mathcal{U}_q(\varrho_0) - \operatorname{Tr}\hat{H}(\varrho_0)\delta\varrho \tag{18}$$

and

$$Q_2(\delta\varrho) \le C(\varrho_0+\delta\varrho) - C(\varrho_0) - \operatorname{Tr}\frac{\delta C}{\delta\varrho}(\varrho_0)\delta\varrho . \tag{19}$$

Require that $Q_1(\delta\varrho) + Q_2(\delta\varrho) > 0$ for all $\delta\varrho \neq 0$. Then these quadratic forms determine a norm $||\cdot||$ by

$$||\delta\varrho||^2 = Q_1(\delta\varrho) + Q_2(\delta\varrho)\ . \tag{20}$$

Given these quadratic forms, and the associated norm, one has to prove continuity of $\mathcal{U}$ and C in the vicinity of ϱ_0, in the following sense:

$$|\mathcal{U}_q(\varrho_0 + \delta\varrho) - \mathcal{U}_q(\varrho_0)| \leq K_1 ||\delta\varrho||^\alpha \tag{21}$$

and

$$|C(\varrho_0 + \delta\varrho) - C(\varrho_0)| \leq K_2 ||\delta\varrho||^\alpha \tag{22}$$

for constants $K_1 > 0$, $K_2 > 0$, and $\alpha > 0$, and for sufficiently small $||\delta\varrho||$. We did not succeed to complete this step. Probably, progress can be made by first studying explicit examples.

5 Thermodynamic Stability

The invariant density matrix ϱ_0 is formally stable if the free energy $\mathcal{F}(\varrho)$ is minimal at ϱ_0 and the matrix of second order variations is positive definite.

A short calculation shows that the free energy $\mathcal{F}^{(2)}(\varrho) = \mathcal{U}_q^{(2)}(\varrho) - TS_q(\varrho)$ is minimal at $\varrho = \varrho_0$, with the equilibrium density matrix ϱ_0 given by

$$\varrho_0 = \frac{1}{\zeta(\beta, q)} \left(\mathbf{1} + \beta(q-1)H\right)^{1/(1-q)} \tag{23}$$

and with

$$\zeta(\beta, q) = \ \mathrm{Tr}\ \left(\mathbf{1} + \beta(q-1)H\right)^{1/(1-q)} \tag{24}$$

(assuming $\mathbf{1} + \beta(q-1)H > 0$ and $\beta = 1/k_B T$).

We can now formulate our first result.

Theorem 1. *Assume that* $1 < q \leq 2$. *Take a Hamiltonian* H *and an inverse temperature* $\beta = 1/k_B T$ *such that simultaneously* $1+\beta(q-1)H$ *is strictly positive and* $H^{-1/(q-1)}$ *has a finite trace. Let* ϱ_0 *be the equilibrium density matrix given by* (23) *and* $\gamma = \beta(q-1)$. *Then*

$$\mathcal{F}^{(2)}(\varrho) - \mathcal{F}^{(2)}(\varrho_0) \geq \frac{q}{2} k_B T ||\varrho - \varrho_0||_\gamma^2 \tag{25}$$

for all density matrices ϱ.

The theorem shows that, if a state ϱ has a free energy which differs only slightly from that of ϱ_0, then the distance between ϱ and ϱ_0 is small in the $||\cdot||_\gamma$ norm. The converse is not necessarily true. The theorem shows further that the density matrix ϱ_0 is formally stable, and is an absolute minimum of the free energy $\mathcal{F}^{(2)}$. For this reason it is called a a thermodynamic equilibrium state.

6 Proof of Theorem 1

Two lemmas are needed. The following result is a variant of the Klein inequality (see e.g. [8], 2.5.2, [9], 2.1.7).

Lemma 1. *Let A, B and C be self-adjoint operators with discrete spectrum. Let f be a convex function. Assume $C \geq 0$ and $CB = BC$. Then*

$$\mathrm{Tr}\, C\left(f(A) - f(B) - (A-B)f'(B)\right) \geq 0\,. \tag{26}$$

Proof. Let $\{\phi_i\}_i$ be a basis in which A is diagonal, and $\{\psi_i\}_i$ a basis in which B and C are diagonal. Let $A\phi_i = a_i\phi_i$, $B\psi_i = b_i\psi_i$ and $C\psi_i = c_i\psi_i$. Denote $\lambda_{i,j} = \langle\phi_i|\psi_j\rangle$. Then

$$\begin{aligned}&\langle\phi_i|C\left(f(A) - f(B) - (A-B)f'(B)\right)\phi_i\rangle\\ &\qquad = \sum_j c_j|\lambda_{i,j}|^2\left(f(a_i) - f(b_j) - (a_i - b_j)f'(b_j)\right) \geq 0.\end{aligned} \tag{27}$$

This proves the lemma.

Take

$$f = \frac{q}{2}f_2 - f_q \qquad \text{with} \qquad f_q(x) = \frac{x - x^q}{q-1}\,. \tag{28}$$

One has $S_q(\varrho) = k_B\,\mathrm{Tr}\, f_q(\varrho)$. It is easy to check that f is convex on the interval $[0,1]$, provided $0 < q \leq 2$. Hence Klein's inequality implies that

$$\begin{aligned}&\frac{q}{2}\,\mathrm{Tr}\, C\left(f_2(A) - f_2(B) - (A-B)f_2'(B)\right)\\ &\qquad \geq\ \mathrm{Tr}\, C\left(f_q(A) - f_q(B) - (A-B)f_q'(B)\right)\end{aligned} \tag{29}$$

for operators A and B with spectrum in $[0,1]$, and any positive operator C commuting with B.

Lemma 2.

$$\mathrm{Tr}\,(\varrho - \varrho_0)H = k_BT\,\mathrm{Tr}\,(\mathbf{1} + \beta(q-1)H)(\varrho - \varrho_0)f_q'(\varrho_0)\,. \tag{30}$$

Proof.

$$\begin{aligned}&\mathrm{Tr}\,(\mathbf{1} + \beta(q-1)H)(\varrho - \varrho_0)f_q'(\varrho_0) - \beta\,\mathrm{Tr}\, H(\varrho - \varrho_0)\\ &\quad = \mathrm{Tr}\,(\varrho - \varrho_0)f_q'(\varrho_0) + \beta(q-1)\,\mathrm{Tr}\, H(\varrho - \varrho_0)f_q'(\varrho_0)\\ &\qquad\quad - \beta\,\mathrm{Tr}\, H(\varrho - \varrho_0)\\ &\quad = \mathrm{Tr}\,(\varrho - \varrho_0)f_q'(\varrho_0) + \beta\,\mathrm{Tr}\, H(\varrho - \varrho_0)(\mathbf{1} - q\varrho_0^{q-1})\\ &\qquad\quad - \beta\,\mathrm{Tr}\, H(\varrho - \varrho_0)\\ &\quad = \mathrm{Tr}\,(\varrho - \varrho_0)f_q'(\varrho_0) - \beta q\,\mathrm{Tr}\, H(\varrho - \varrho_0)\varrho_0^{q-1}\\ &\quad = \frac{1}{q-1}\,\mathrm{Tr}\,(\varrho - \varrho_0)\Big(\mathbf{1} - q\varrho_0^{q-1}(\mathbf{1} + \beta(q-1)H)\Big) = 0\,.\end{aligned} \tag{31}$$

The latter follows because (23) implies that $\varrho_0^{q-1}(\mathbf{1} + \beta(q-1)H)$ is a multiple of $\mathbf{1}$. This proves the lemma.

Recall that

$$\mathcal{U}_q^{(2)}(\varrho) = \operatorname{Tr} \varrho^q H = \operatorname{Tr} \varrho H - (q-1) \operatorname{Tr} f_q(\varrho) H \ . \tag{32}$$

Hence, assuming $\operatorname{Tr} \varrho = 1$, one has

$$\mathcal{F}_q^{(2)}(\varrho) = \operatorname{Tr} \varrho H - k_B T \operatorname{Tr} C f_q(\varrho) \tag{33}$$

with $C = \mathbf{1} + \beta(q-1)H$. Therefore

$$\begin{aligned} \mathcal{F}_q^{(2)}(\varrho) - \mathcal{F}_q^{(2)}(\varrho_0) = {} & \operatorname{Tr} (\varrho - \varrho_0) H \\ & - k_B T \operatorname{Tr} C (f_q(\varrho) - f_q(\varrho_0)) \ . \end{aligned} \tag{34}$$

Using the result of Lemma 2 one finds

$$\begin{aligned} & \mathcal{F}_q^{(2)}(\varrho) - \mathcal{F}_q^{(2)}(\varrho_0) \\ & \quad = -k_B T \operatorname{Tr} C (f_q(\varrho) - f_q(\varrho_0) - (\varrho - \varrho_0) f_q'(\varrho_0)) \ . \end{aligned} \tag{35}$$

From (29) with $A = \varrho$, $B = \varrho_0$, then follows

$$\begin{aligned} \mathcal{F}_q^{(2)}(\varrho) - \mathcal{F}_q^{(2)}(\varrho_0) & \geq -k_B T \frac{q}{2} \operatorname{Tr} C \left(f_2(\varrho) - f_2(\varrho_0) - (\varrho - \varrho_0) f_2'(\varrho_0) \right) \\ & = k_B T \frac{q}{2} \operatorname{Tr} C (\varrho - \varrho_0)^2 \\ & = k_B T \frac{q}{2} ||\varrho - \varrho_0||_\gamma^2 \end{aligned} \tag{36}$$

with $\gamma = \beta(q-1)$. This ends the proof of the theorem.

7 Minima of $\mathcal{F}^{(3)}$

The variational principle states that the free energy (1) is minimal in equilibrium. Its origin is that entropy $\mathcal{S}_q(\varrho)$ should be maximal under the constraint that the energy $\mathcal{U}_q(\varrho)$ has a given value. To study $\mathcal{F}^{(3)}$ it is easier to consider the problem of minimizing $\mathcal{U}_q^{(3)}(\varrho)$ under the constraint that the entropy $\mathcal{S}_q^{(3)}(\varrho)$ has a given value. The reason for this is that at constant entropy the denominator of (5) is also constant. By the method of Lagrange multipliers, the minimum of $\mathcal{F}^{(3)}$ is the solution of the problem of minimizing $\mathcal{U}_q^{(3)}(\varrho)$ given $\mathcal{S}_q^{(3)}(\varrho)$. By variation of the value of $\mathcal{S}_q^{(3)}(\varrho)$ one then obtains a solution of the original variational principle. The study of $\mathcal{F}^{(3)}$ which follows below is based on this idea.

From now on assume that $1 < q \leq 2$ – the situation for $0 < q < 1$ is more complicated due to the possibility of a high energy cutoff, as discussed in [1]; theorem 1 is only valid for $q \leq 2$. The equilibrium state is of the form $\varrho_0 = \varrho_\alpha$ with

$$\varrho_\alpha = \frac{1}{\zeta_\alpha} \left(\frac{1}{\alpha \mathbf{1} + H} \right)^{1/(q-1)} \tag{37}$$

and

$$\zeta_\alpha = \operatorname{Tr} \left(\frac{1}{\alpha \mathbf{1} + H} \right)^{1/(q-1)} \ . \tag{38}$$

(See [1].) The parameter α should be such that $\alpha\mathbf{1}+H$ is strictly positive, in order to guarantee that the denominator is strictly positive.

Note that

$$\begin{aligned}\varrho_\alpha^q H &= \zeta_\alpha^{1-q}\frac{\varrho_\alpha H}{\alpha\mathbf{1}+H} = \zeta_\alpha^{1-q}\varrho_\alpha\left(\mathbf{1}-\frac{\alpha\mathbf{1}}{\alpha\mathbf{1}+H}\right)\\ &= \zeta_\alpha^{1-q}\varrho_\alpha - \alpha\varrho_\alpha^q\ . \end{aligned} \tag{39}$$

From (39) it follows that

$$\mathcal{U}_q^{(3)}(\varrho_\alpha) = \frac{1}{\zeta_\alpha^{q-1}}\frac{1}{\operatorname{Tr}\varrho_\alpha^q} - \alpha\ . \tag{40}$$

Hence the free energy equals

$$\mathcal{F}_q^{(3)}(\varrho_\alpha) = \frac{1}{\zeta_\alpha^{q-1}}\frac{1}{\operatorname{Tr}\varrho_\alpha^q} - \alpha - \frac{1}{\beta(q-1)}(1-\operatorname{Tr}\varrho_\alpha^q) \tag{41}$$

with $\beta = 1/k_BT$ as usual. Variation w.r.t. α gives

$$\begin{aligned}\frac{\partial}{\partial\alpha}\mathcal{F}_q^{(3)}(\varrho_\alpha) &= -(q-1)\frac{1}{\zeta_\alpha^q}\frac{1}{\operatorname{Tr}\varrho_\alpha^q}\frac{\partial\zeta_\alpha}{\partial\alpha} - 1\\ &\quad - \left(\frac{1}{\zeta_\alpha^{q-1}}\frac{1}{(\operatorname{Tr}\varrho_\alpha^q)^2} - \frac{1}{\beta(q-1)}\right)\frac{\partial}{\partial\alpha}\operatorname{Tr}\varrho_\alpha^q\ . \end{aligned} \tag{42}$$

Using

$$\frac{\partial}{\partial\alpha}\zeta_\alpha = -\frac{1}{q-1}\operatorname{Tr}\frac{1}{(\alpha\mathbf{1}+H)^{q/(q-1)}} = -\frac{1}{q-1}\zeta_\alpha^q\operatorname{Tr}\varrho_\alpha^q \tag{43}$$

there follows

$$\frac{\partial}{\partial\alpha}\mathcal{F}_q^{(3)}(\varrho_\alpha) = \left(\frac{1}{\beta(q-1)} - \frac{1}{\zeta_\alpha^{q-1}}\frac{1}{(\operatorname{Tr}\varrho_\alpha^q)^2}\right)\frac{\partial}{\partial\alpha}\operatorname{Tr}\varrho_\alpha^q\ . \tag{44}$$

One can show that $\operatorname{Tr}\varrho_\alpha^q$ is a strictly decreasing function of α (see the appendix). Therefore the extrema of $\mathcal{F}_q^{(3)}(\varrho_\alpha)$ are reached when α is solution of

$$\beta(q-1) = \zeta_\alpha^{q-1}(\operatorname{Tr}\varrho_\alpha^q)^2\ . \tag{45}$$

This equation can be written out as

$$\beta(q-1) = \frac{\left(\operatorname{Tr}(\alpha\mathbf{1}+H)^{-q/(q-1)}\right)^2}{\left(\operatorname{Tr}(\alpha\mathbf{1}+H)^{-1/(q-1)}\right)^{1+q}}\ . \tag{46}$$

One can convince oneself that this equation has a unique solution $\alpha(\beta)$ for each $\beta>0$ (see the appendix).

We can now formulate our main result.

Theorem 2. *Assume that* $1<q\le 2$ *and that* $H^{-1/(q-1)}$ *has a finite trace. Then the free energy* $\mathcal{F}_q^{(3)}(\varrho)$ *has a unique minimum. It occurs at* $\varrho=\varrho_\alpha$ *with* α *the unique solution of* (46).

8 Proof of Theorem 2

Let ϱ be an arbitrary density matrix. There exists γ such that $\mathcal{S}_q(\varrho) = \mathcal{S}_q(\varrho_\gamma)$. Now Theorem 1 implies

$$\mathcal{U}_q^{(2)}(\varrho) - T'\mathcal{S}_q(\varrho) \geq \mathcal{U}_q^{(2)}(\varrho_\gamma) - T'\mathcal{S}_q(\varrho_\gamma) \tag{47}$$

with $k_B T' = \gamma(q-1)$. But because the entropies are equal one has also $\mathcal{F}_q^{(3)}(\varrho) \geq \mathcal{F}_q^{(3)}(\varrho_\gamma)$. Because α is the solution of (46) one has $\mathcal{F}_q^{(3)}(\varrho_\gamma) \geq \mathcal{F}_q^{(3)}(\varrho_\alpha)$. The two inequalities together yield

$$\mathcal{F}_q^{(3)}(\varrho) \geq \mathcal{F}_q^{(3)}(\varrho_\alpha) \, . \tag{48}$$

Finally, equality in (48) implies $\gamma = \alpha$ and $\varrho = \varrho_\gamma$ because of the underbound of Theorem 1. Hence the theorem is proved.

9 Conclusions

Our results bring additional support for the recently introduced normalized energy functional. Theorem 2 proves the uniqueness of the equilibrium density matrix ϱ_0 at any positive temperature for $1 < q \leq 2$. The situation for $0 < q < 1$ is more complicated and has not been treated here. The stronger result of thermodynamic stability with an underbound as obtained in Theorem 1 could only be proved in the context of the unnormalized energy functional. It is quite likely that a similar result holds in the context of the normalized energy functional. However, it is not clear how to tackle this problem. Concerning dynamic stability, which is the strongest of the stability properties under consideration here, no result could be obtained. Probably progress can be made by studying dynamic stability in explicit examples.

Acknowledgements
Our collaboration is supported by the Flemish-Polish project No. 007 and is financed in part by KBN Grant No. 2 P03B 163 15 and the Alexander von Humboldt Foundation.

Appendix

We show that $\mathrm{Tr}\, \varrho_\alpha^q$ is a strictly decreasing function of α and that (46) has a unique solution for each $\beta > 0$.

First note that

$$\begin{aligned} \frac{\partial \zeta_\alpha}{\partial \alpha} &= \frac{1}{1-q} \mathrm{Tr}\, (\alpha \mathbf{1} - \beta(1-q)H)^{q/(1-q)} \\ &= \frac{1}{1-q} \zeta_\alpha^q \, \mathrm{Tr}\, \varrho_\alpha^q \, . \end{aligned} \tag{49}$$

A short calculation using (49) gives

$$\begin{aligned}(q-1)\frac{\partial}{\partial\alpha}\,\mathrm{Tr}\,\varrho^q &= (q-1)\frac{\partial}{\partial\alpha}\zeta_\alpha^{-q}\,\mathrm{Tr}\,(\alpha\mathbf{1}-\beta(1-q)H)^{q/(1-q)}\\ &= -q(q-1)\zeta_\alpha^{-q-1}\,\mathrm{Tr}\,(\alpha\mathbf{1}-\beta(1-q)H)^{q/(1-q)}\frac{\partial\zeta_\alpha}{\partial\alpha}\\ &\qquad -q\zeta_\alpha^{-q}\,\mathrm{Tr}\,(\alpha\mathbf{1}-\beta(1-q)H)^{(2q-1)/(1-q)}\\ &= q\zeta_\alpha^{-q-1}\left[f(q)^2-f(1)f(2q-1)\right]\end{aligned} \tag{50}$$

with

$$f(x)=\ \mathrm{Tr}\,(\alpha\mathbf{1}-\beta(1-q)H)^{x/(1-q)}\ . \tag{51}$$

Now, the function f is strictly log-convex (assume that H is not a multiple of the identity). Hence the r.h.s. of (50) is negative. This shows that $\mathrm{Tr}\,\varrho_\alpha^q$ is strictly decreasing.

Equation (46) has at least one solution, because its rhs tends to $+\infty$ as α tends to $-E_\mathrm{m}$ and to zero as α goes to $+\infty$. Hence it takes on all positive values. Therefore a solution exists for any $\beta>0$.

Take the logarithm of the r.h.s. of (46). Its derivative w.r.t. α equals

$$-\frac{2q}{q-1}\frac{f_\alpha(2q-1)}{f_\alpha(q)}+\frac{1+q}{q-1}\frac{f_\alpha(q)}{f_\alpha(1)} \tag{52}$$

with

$$f_\alpha(x)=\ \mathrm{Tr}\,(\alpha\mathbf{1}+H)^{-x/(q-1)}\ . \tag{53}$$

Since f_α is strictly log-convex, and $q+1<2q$, and $q=(1/2)1+(1/2)(2q-1)$, expression (52) is negative. This implies that the r.h.s. of (46) is decreasing in α. Hence (46) has a unique solution which is denoted $\alpha(\beta)$. Moreover, $\alpha(\beta)$ is a decreasing function of β.

References

1. C. Tsallis, R. S. Mendes, and A. R. Plastino, *The role of constraints within generalized nonextensive statistics.* Physica A **261**, 543-554 (1998).
2. C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics.* J. Stat. Phys. **52**, 479 (1988).
3. E. M. F. Curado and C. Tsallis, *Generalized statistical mechanics: connection with thermodynamics.* J. Phys. A **24**, L69-L72 (1991).
4. M. Czachor and J. Naudts, *Microscopic Foundation of Nonextensive Statistics.* Phys. Rev E **59**, R2497-R2500 (1999).
5. J. Naudts, *Dual description of nonextensive ensembles*, preprint (1999) [cond-mat/9904070 v2].
6. D. D. Holm, J. E. Marsden, T. S. Ratiu, and A. Weinstein, *Nonlinear stability of fluid and plasma equilibria.* Phys. Rep. **123**, 2-116 (1985).
7. J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry* (Springer-Verlag, New York, 1994).
8. D. Ruelle, *Statistical Mechanics: Rigorous Results* (Benjamin, New York, 1969).
9. W. Thirring, *Lehrbuch der Mathematischen Physik*, Vol. 4 (Springer-Verlag, Wien, 1980).

VII. Generalized Simulated Annealing Algorithms Using Tsallis Statistics: Application to ±*J* Spin Glass Model

J. Kłos[1,2] and S. Kobe[1]

[1] Institut für Theoretische Physik, Technische Universität Dresden,
D-01062 Dresden, Germany
[2] Instytut Fizyki, Uniwersytet im. A. Mickiewicza, Coll. Physicum,
ul. Umultowska 85, 61-614 Poznań, Poland

Abstract. Recently, the Tsallis generalized distribution of states has been adapted to various optimization algorithms. New Monte Carlo acceptance probabilities have been worked out and tested in a number of physical systems. In the paper we have applied the generalized simulated annealings to a ± J spin glass of 70 × 70 spins on a square lattice. Since the system ground-state energy was known exactly the ability of the generalized acceptance probabilities to find low-energy states has been checked out. Efficiency of the new annealing procedures has also been compared with that of the traditional method with respect to the number of annealing steps.

1 Generalized Acceptance Probabilities

Simulated annealing (SA) is a computational tool to deal with hard optimization problems [1]. The method proceeds in such a way that at each annealing step n the temperature $T(n)$ is given and the system is allowed to walk on its phase space long enough so that it will thermalize at $T(n)$. Then the temperature is decreased according to a given annealing schedule and the process is repeated till the final temperature is sufficiently low. Some of the recent applications of SA [2,3] have been based on the Tsallis statistics which unlike the usual Boltzmann-Gibbs theory can be used for nonextensive systems [4,5].

The paper is organized as follows. We first recall the generalized simulated annealings (SAs) based on the Tsallis formalism. Next, we describe the system considered and the organization of the programs. The exact knowledge of the system ground-state energy enables us not only to examine the time and energy characteristics of the generalized SAs but also to check their efficiency quantitatively.

The Tsallis generalized acceptance probability (AP) is defined as

$$p = \min[1, [1 - (1-q)\beta \Delta E]^{\frac{1}{(1-q)}}], \tag{1}$$

where $\Delta E = E_{new} - E_{old}$, $\beta = 1/kT$ (E_{new} and E_{old} are energies after and before a random reconfiguration; in the following $k = 1$ is used) and q is the Tsallis parameter [5]. For $q \to 1$ the standard AP is recovered. Since (1) does

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 253–258, 2001.

not obey detailed balance, Andricioaei and Straub have recently replaced it by another AP [3]

$$p = \min\left[1, \left(\frac{1-[1-q(T)]\beta E_{new}}{1-[1-q(T)]\beta E_{old}}\right)^{\frac{q(T)}{[1-q(T)]}}\right],$$

$$\lim_{T\to 0} q(T) = 1. \tag{2}$$

In (1) q is kept constant whereas in (2) it is a monotonically decreasing function of temperature T.

2 Model and Simulations

In our work we have tested the above SAs for the 2D Edward-Anderson spin glass system of 70×70 spins whose ground state energy is known to us [6]. Ising spins are located at each site of a square lattice with periodical boundary conditions. The ferromagnetic and antiferromagnetic bonds $J_{ij} = \pm 1$ between the nearest neighbouring sites are distributed randomly and uniformly so that the fraction of each of them is $c = 0.5$. The system energy is given by

$$E = -\sum_{i<j} J_{ij} S_i S_j + 2N, \tag{3}$$

where $S_i, S_j = \pm 1$ are Ising spins, N is their total number, and the sum on the right hand side runs over the nearest neighbouring sites. The last term in (3) is a shift in the energy scale so that it becomes nonnegative [7]. The SAs are tested by applying Monte Carlo (MC) algorithms within single-spin dynamics [8] which consist of the following steps:

1. Set the temperature T and q (the initial temperature T_i is chosen considerably higher than the largest ΔE).
2. Pick a spin at random.
3. Compute E_{new}, E_{old}, and the energy change associated with that flip.
4. Calculate the acceptance probability p for that flip.
5. Draw a random number z uniformly distributed between zero and unity.
6. If $z < p$ flip the spin, otherwise do not flip it.
7. Repeat step 2. T and q are kept constant for $100N$ trial reconfigurations or for $10N$ successful rearrangements, whichever comes first [2].
8. Record the energy.
9. Change T (and q) according to the annealing schedule. Start with step 2.

The initial temperature T_i is lowered according to $T = T_i \left(\frac{T_f}{T_i}\right)^{\frac{n}{N_a}}$, where $n = 0, 1, 2, ..., N_a$, and N_a is the number of annealing steps. T_f is the final temperature. In (2) the initial Tsallis parameter $q = q_i$ is also decreased exponentially to one, starting from values higher than one [3]. The AP (1) is set to zero whenever the argument of the power law acceptance function is negative.

3 Results

For each of the APs we have carried out 20 independent runs of $N_a = 1000$ steps for one special sample of the system size $N = 70 \times 70$. As measured in annealing steps (Fig. 1) low energies are found faster with (1) for $q < 1$. It is particularly the case for $q < 0$ which seems consistent with Penna's solution of the traveling salesman problem [2]. However, the minimizing process stops at relatively high temperatures and the system is trapped in a set of isoenergetic states. For q nearly one no major difference from the regular SA can be seen, and for $q > 1$ there is no energy decrease at all. The SA (2) gives different results. Only for q_i extremely close to one the annealing looks similar to the traditional one. In

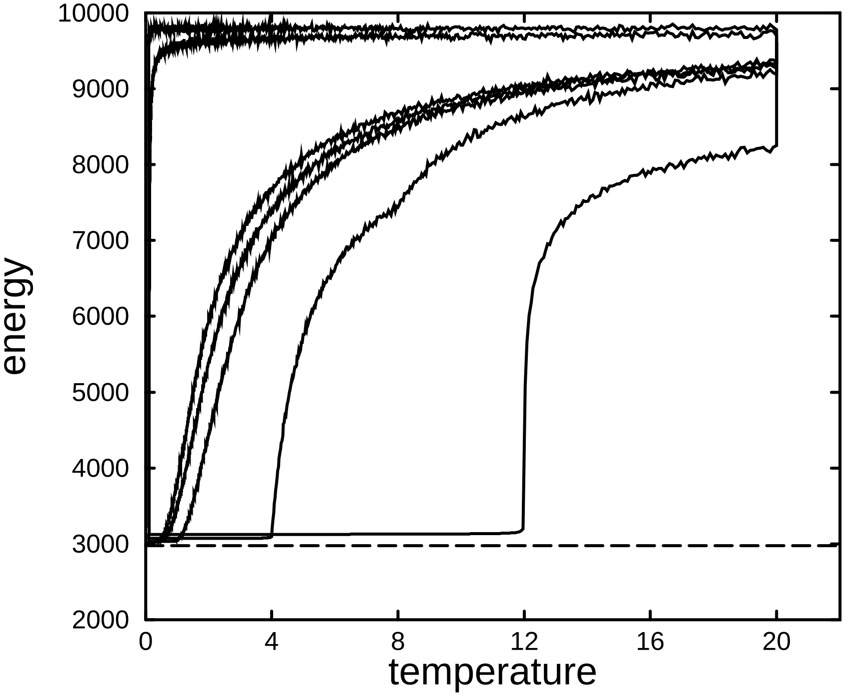

Fig. 1. The energy E versus temperature T for the average of 20 runs. A 70×70 spin system with the ground-state energy of 2976 (dashed line) is used. $T_i = 20$, $T_f = 0.1$. From the bottom right hand side corner to the top left hand side one: a) AP (1) with $q = -2, 0, 0.8$, respectively. In this case the energy decreases very rapidly even in the 'zeroth' annealing step. b). The standard SA ($q = 1$). c). AP (2) with $q_i = 1.0001, 1.01, 2$, respectively.

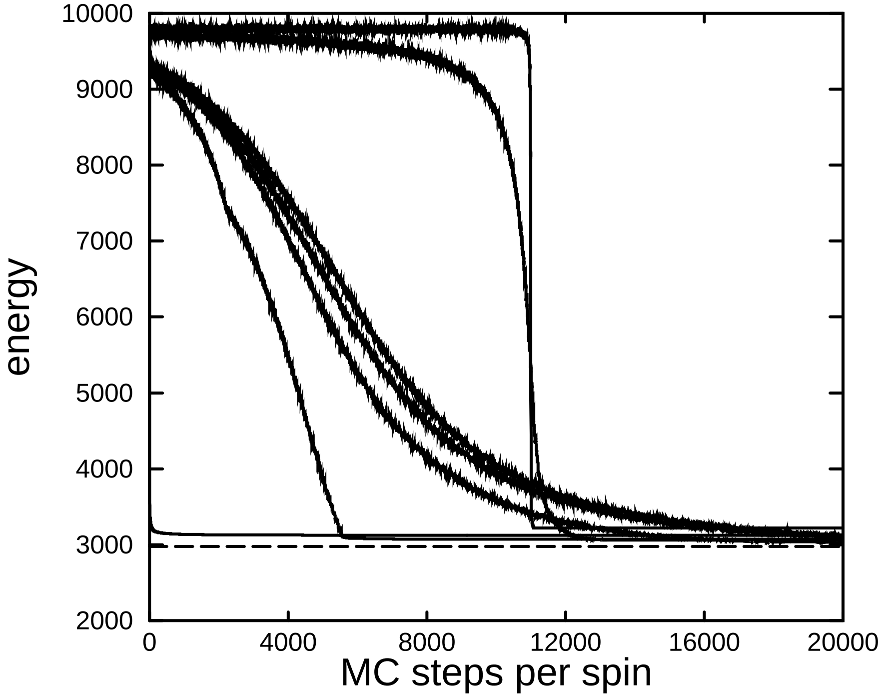

Fig. 2. The energy E versus MC steps per spin for the average of 20 runs for the same system as in Fig 1. The same order as in Fig. 1 but from the bottom left hand side corner to the top right hand side one.

other cases most of the annealing steps the energy fluctuates around a constant value and drops at the end. No special advantage over the standard SA can be observed in this approach. With respect to the number of MC steps per spin (Fig. 2), as previously, eq. (1) leads to low energies extremely fast for $q < 0$, and an immediate trapping in local energy minima is present. The algorithm (2) is slower and for q_i nearly one it can only be as fast as the regular SA.

We also have been interested in the precise values of the energies found at the end of each run (Tab. I). Using (1) the closer to one q is the lower are both the lowest final energy E_l and its average $< E_l >$ (a). However, the energies found are not qualitatively lower than those obtained using the standard scheme (b). For $q = -2$ the minimizing is very fast but the system is trapped in a subspace of states with relatively high energies. When applying (2), for q_i slightly larger than one we have arrived at relatively low final energies (c). Although some of the E_l values are even a little lower than those in (b) we doubt that the difference is sufficient to consider this AP as more efficient. The tendency is

Table 1. The same system as in Fig. 1. The lowest final energy E_l out of 20 independent runs found for each pair of T_i and T_f using: a) AP (1), b) standard SA, c) AP (2). $< E_l >$ is the average over the 20 runs of E_l obtained in each of them.

a	T_i	T_f	q	E_l	$< E_l >$	b	q	E_l	$< E_l >$	c	q_i	E_l	$< E_l >$
	20	0.1	0.8	3012	3032.00		1	3008	3026.20		2	3152	3217.80
	10	0.1	0.8	3012	3027.00		1	3012	3024.80		1.01	3044	3066.00
	20	0.1	0	3044	3069.89		1				1.01	3036	3060.60
	10	0.1	-2	3096	3122.67		1				1.0001	3004	3020.60
	10	0.001	-2	3096	3124.20		1	3020	3031.4		1.0001	3012	3025.80
	20	0.1	-2	3092	3125.33		1				1.0001	3008	3024.40
	20	0.001	-2	3076	3119.40		1	3008	3031.4		1.0001	3004	3030.00

such that for larger q_i both E_l and $< E_l >$ are higher as well. Although, none of the algorithms has arrived at the true ground state, the precise knowledge of its energy ($E_{ground} = 2976$, i. e. $\frac{E_{ground}}{N} - 2 = -1.393$) enables us to conclude that in terms of energy the final states found were relatively close to it. Actually, the ratio $|\frac{E_l - E_{ground}}{E_{ground} - 2N}|$ is of the order of only 0.4 percent. Otherwise, the lowest found state with the energy $E_l = 3004$ belongs to the 7th excited level above the exact ground one.

4 Summary

In the paper we have performed Monte Carlo simulations of $\pm J$ spin glass using Tsallis generalized mechanics. We have discussed two proposals for the use of Tsallis weights in simulated annealing simulated annealing algorithms and compared them to the standard one. In frames of the described cooling protocol none of the generalized schemes has proven qualitatively more efficient in finding states of low energies than that one based on the Metropolis acceptance probability. After a carefull choice of the free parameters it is only possible to arrive at states whose energies compare with those found with the standard Monte Carlo method. However, if one is only interested in very quick but rather imperfect minimization the Tsallis generalized algorithm seems the most useful. This approach could be used, for instance, as an introductory step for further deterministic optimization. Of course, our results are not fully general because they refer only to one cooling schedule which actually is very widespread in the literature but not the only. However, after comparing them with those obtained for other systems, they seem to confirm the general opinion that the presented methods are strongly model dependent. Therefore, we think that each of them should be exactly tested before a serious application to a certain system.

Acknowledgements

We acknowledge support from Graduiertenkolleg "Struktur- und Korrelationseffekte in Festkörpern". We thank J. Weißbarth and M. Nogala for helpful discussions and the authors of [6] for allowance to use their computer data.

References

1. S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Science **220**, 671 (1983).
2. T. J. P. Penna, Phys. Rev. E **51**, R1 (1995).
3. I. Andricioaei and J. E. Straub, Phys. Rev. E **53**, R3055 (1996).
4. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
5. D. A. Stariolo and C. Tsallis, in *Annual Reviews of Computational Physics II*, ed. D. Stauffer (World Scientific, Singapore, 1995), p. 343.
6. C. De Simone, M. Diehl, M. Jünger, P. Mutzel, G. Reinelt, and G. Rinaldi, J. Stat. Phys. **80**, 487 (1995).
7. I. Andricioaei and J. E. Straub, Physica A **247**, 553 (1997).
8. K. Binder and D. W. Herrmann, *Monte Carlo Simulation in Statistical Physics*, Springer Series in Solid-State Sciences, Vol. 80 (Springer-Verlag, Berlin, 1988).

VIII. Protein Folding Simulations by a Generalized-Ensemble Algorithm Based on Tsallis Statistics

Y. Okamoto[1] and U.H.E. Hansmann[2]

[1] Department of Theoretical Studies, Institute for Molecular Science
and Dept. of Funct. Mol. Sci., The Graduate University for Advanced Studies
Okazaki, Aichi 444-8585, Japan
[2] Department of Physics, Michigan Technological University
Houghton, MI 49931-1291, U.S.A.

Abstract. We review uses of Tsallis statistical mechanics as a generalized-ensemble simulation algorithm in the protein folding problem. A simulation based on this algorithm performs a random walk in energy space, and it allows one not only to find the global-minimum-energy conformation but also to obtain probability distributions in canonical ensemble for a wide temperature range from only one simulation run. The folding properties of a penta peptide, Met-enkephalin, are studied by this algorithm.

1 Introduction

The protein folding problem is one of the most challenging problems in theoretical molecular science and biophysics (for reviews, see, for instance, Refs. [1,2]). Proteins under their native physiological conditions spontaneously fold into unique three-dimensional structures (tertiary structures) in the time scale of miliseconds to minutes. Although the tertiary structures of proteins appear to be dependent on various environmental factors within the cell, it was shown by experiments *in vitro* that unfolded proteins can refold back into their native conformations once the denaturants are removed, implying that the three-dimensional structure of a protein is determined solely by its amino-acid sequence information [3]. This gave many theoreticians a hope that protein folding can be reproduced by computer simulations. However, this has not been accomplished to date even for small proteins.

The difficulty comes from the fact that the number of possible conformations for each protein is astronomically large [4]. Simulations by conventional methods such as Monte Carlo or molecular dynamics algorithms in canonical ensemble will necessarily get trapped in one of many local-minimum states in the energy function. In order to overcome this multiple-minima problem, many methods have been proposed (for reviews, see, for example, Refs. [5]-[8]). Generalized-ensemble algorithms, most well-known of which is the multicanonical approach [9], are powerful ones and were first introduced to the protein-folding problem in Ref. [10]. Simulations in the multicanonical ensemble perform 1D random walk in energy space, which allows the system to overcome any energy barrier. Besides multicanonical algorithms, simulated tempering [11] and *1/k*-sampling [12] have

S. Abe and Y. Okamoto (Eds.): LNP 560, pp. 259–274, 2001.

been shown to be equally effective generalized-ensemble methods in the protein folding problem [13]. The simulations are usually performed with Monte Carlo (MC) scheme, but recently molecular dynamics (MD) version of multicanonical algorithm was also developed [14]-[16].

The generalized-ensemble approach is based on non-Boltzmann probability weight factors, and in the above three methods the determination of the weight factors is non-trivial. We have shown that a particular choice of the Tsallis weight factor [17] can be used for generalized-ensemble simulations [18,19]. (For a review, see Ref. [20]. Similar ideas based on the Tsallis weight were also pursued in Refs. [21,22]. See the lecutre notes by Straub in this volume.) The advantage of this ensemble is that it greatly simplifies the determination of the weight factor. We remark that simulated annealing versions of the method have been applied to the protein folding problem in Refs. [21,23].

In this article, we review a generalized-ensemble algorithm based on Tsallis statistics. The performances of the algorithm are tested with the system of an oligopeptide, Met-enkephalin, and the free energy landscape of this peptide is studied in the context of the protein folding problem.

2 Methods

2.1 Energy Function of Protein Systems

The total potential energy function E_{tot} that we used is one of the standard ones. Namely, it is given by the sum of the electrostatic term E_C, 12-6 Lennard-Jones term E_{LJ}, and hydrogen-bond term E_{HB} for all pairs of atoms in the molecule together with the torsion term E_{tor} for all torsion angles:

$$\begin{aligned}
E_{tot} &= E_C + E_{LJ} + E_{HB} + E_{tor} \ , \\
E_C &= \sum_{(i,j)} \frac{332\ q_i q_j}{\epsilon\ r_{ij}} \ , \\
E_{LJ} &= \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \ , \\
E_{HB} &= \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \ , \\
E_{tor} &= \sum_{i} U_i \left(1 \pm \cos(n_i \chi^i) \right) \ .
\end{aligned} \tag{1}$$

Here, r_{ij} is the distance (in Å) between atoms i and j, ϵ is the dielectric constant, and χ^i is the torsion angle for the chemical bond i. Each atom is expressed by a point at its center of mass, and the partial charge q_i (in units of electronic charges) is assumed to be concentrated at that point. The factor 332 in E_C is a constant to express energy in units of kcal/mol. These parameters in the energy function as well as the molecular geometry were adopted from ECEPP/2 [24].

The computer code KONF90 [25] was used for the present simulations. We neglected the solvent contributions for simplicity and set the dielectric constant ϵ equal to 2. The peptide-bond dihedral angles ω were fixed at the value 180° for simplicity. The remaining dihedral angles ϕ and ψ in the main chain and χ in the side chains thus constitute the variables to be updated in the simulations (there are 19 such angles for Met-enkephalin). One MC sweep consists of updating all these angles once with Metropolis evaluation [26] for each update.

2.2 Generalized-Ensemble Algorithm with Tsallis Statistics

In the canonical ensemble at temperature T each state with potential energy E is weighted by the Boltzmann factor

$$w_B(E,T) = e^{-\beta E} \ , \tag{2}$$

where the inverse temperature is given by $\beta = 1/k_B T$ with Boltzmann constant k_B. This weight factor gives the usual bell-shaped canonical probability distribution of energy:

$$P_B(E,T) \propto n(E)\ w_B(E,T) \ , \tag{3}$$

where $n(E)$ is the density of states. For systems with many degrees of freedom, it is usually very difficult to generate a canonical distribution at low temperatures. This is because there are many local minima in the energy function, and simulations will get trapped in states of these local minima.

Generalized-ensemble algorithms are the methods that perform random walks in energy space, allowing simulations to escape from any state of energy local minimum. To name a few, multicanonical algorithms [9], simulated tempering [11], and *1/k*-sampling [12] are such algorithms. Here, we discuss one of the latest examples of simulation techniques in generalized ensemble [18,19], which is based on Tsallis statistics [17]. The probability weight factor of Tsallis statistics is given by

$$w(E) \propto [1 + (q-1)\beta E]^{-\frac{1}{q-1}} \ , \tag{4}$$

which tends to the Boltzmann factor of Eq. (2) for $q \to 1$, and therefore regular statistical mechanics is recovered in this limit. In our formalism of Tsallis statistics as a generalized-ensemble simulation algorithm, we use the following probability weight factor:

$$w(E) = \left(1 + \beta_0 \frac{E - E_{GS}}{m}\right)^{-m} \ , \tag{5}$$

where $T_0 = 1/k_B \beta_0$ is a low temperature, E_{GS} is the global-minimum potential energy, and $m (> 0)$ is a free parameter, the optimal value of which will be given below. This is the Tsallis weight of Eq. (4) at a fixed temperature T_0 with the following choice of q parameter:

$$q = 1 + \frac{1}{m} \ . \tag{6}$$

Note that through the substraction of E_{GS} it is ensured that the weight will always be positive definite, which is required for computer simulations (otherwise, the quantity can become complex).

The above choice of the weight was motivated by the following reasoning [18]. We are interested in an ensemble where not only the low-energy region can be sampled efficiently but also the high-energy states can be visited with finite probability. In this way the simulation can overcome energy barriers and escape from local minima. The probability distribution of energy should resemble that of an ideal low-temperature canonical distribution, but with a tail to higher energies. The Tsallis weight of Eq. (5) at low temperature T_0 has the required properties when the parameter m is carefully chosen. Namely, for suitable $m > 0$ it is a good approximation of the Boltzmann weight $w_B(E, T_0) = \exp(-\beta_0(E - E_{GS}))$ for $\beta_0(E - E_{GS})/m \ll 1$, while at high energies it is no longer exponentially suppressed but only according to a power law with the exponent m.

In this work we consider a system with continuous degrees of freedom. At low temperatures the harmonic approximation then holds, and the density of states is given by

$$n(E) \propto (E - E_{GS})^{\frac{n_F}{2}} , \tag{7}$$

where n_F is the number of degrees of freedom of the system under consideration. Hence, by Eqs. (5) and (7) the probability distribution of energy for the present ensemble is given by

$$P(E) \propto n(E)w(E) \propto (E - E_{GS})^{\frac{n_F}{2} - m} , \tag{8}$$

for $\beta_0 \frac{E - E_{GS}}{m} \gg 1$. This implies that we need $m > \frac{n_F}{2}$. For, otherwise, the sampling of high-energy configurations will be enhanced too much. On the other hand, in the limit $m \to \infty$ our weight tends for all energies to the Boltzmann weight and high-energy configurations will not be sampled.

In order for low-temperature simulations to be able to escape from energy local minima, the weight should start deviating from the (exponentially damped) Boltzmann weight at the energy near its mean value (because at low temperatures there are only small fluctuations of energy around its mean). In Eq. (5) we may thus set

$$\beta_0 \frac{< E >_T - E_{GS}}{m} = \frac{1}{2} . \tag{9}$$

The mean value at low temperatures is given by the harmonic approximation:

$$< E >_T \;\; = E_{GS} + \frac{n_F}{2} k_B T_0 = E_{GS} + \frac{n_F}{2\beta_0} . \tag{10}$$

Substituting this value into Eq. (9), we obtain the optimal value for the exponent m:

$$m_{opt} = n_F . \tag{11}$$

Hence, the optimal weight factor is given by [18]

$$w(E) = \left(1 + \beta_0 \frac{E - E_0}{n_F}\right)^{-n_F} , \tag{12}$$

where E_0 is the best estimate of the global-minimum energy E_{GS}.

We remark that the calculation of the weight factor is much easier than in other generalized-ensemble techniques, since it requires one to find only an estimator for the ground-state energy E_{GS}. Once the weight factor is given, we can implement the Metropolis MC algorithm [26] in a straightforward manner. We remark that the molecular dynamics algorithm in this generalized ensemble has also been developed in Ref. [19] (see also Ref. [22]).

As in the case of other generalized ensembles, we can use the reweighting techniques [27] to construct canonical distributions at various temperatures T. This is because the simulation by the present algorithm samples a large range of energies. The thermodynamic average of any physical quantity $\mathcal{A}$ can be calculated over a wide temperature range by

$$< \mathcal{A} >_T \ = \ \frac{\displaystyle\int dx \ \mathcal{A}(x) \ w^{-1}(E(x)) \ e^{-\beta E(x)}}{\displaystyle\int dx \ w^{-1}(E(x)) \ e^{-\beta E(x)}} \ , \tag{13}$$

where $w(E)$ is the weight in Eq. (12) and x stands for configurations.

3 Results

The effectiveness of the algorithm presented in the previous section is tested for the system of an oligopeptide, Met-enkephalin. This peptide has the amino-acid sequence Tyr-Gly-Gly-Phe-Met.

It is known from our previous work that the global-minimum value of KONF90 energy for Met-enkephalin is $E_{GS} = -12.2$ kcal/mol [28]. This ground-state structure, which is referred to as Structure A below, is shown in Fig. 1. (This structure is essentially the same as the ground-state conformation found previously in Refs. [29,30].) The figure was created with Molscript [31] and Raster3D [32]. It is a superposition of ball-and-stick and space-filling representations. The latter representation was added in order to give a rough idea of the volume of the peptide as discussed below.

The peptide has essentially a unique three-dimensional structure at temperatures $T \leq 50$ K, and the average energy is about -11 kcal/mol at $T = 50$ K [30,28]. Hence, in the present work we always set $T_0 = 50$ K (or, $\beta_0 = 10.1$ $[\frac{1}{\text{kcal/mol}}]$) in our new probability weight factor in Eq. (12). All simulations were started from completely random initial configurations (Hot Start).

To demonstrate that thermalization is greatly enhanced in our ensemble, we first compare the "time series" of energy as a function of MC sweep. In Fig. 2 we show the results from a regular canonical MC simulation at temperature $T = 50$ K (dotted curve) and those from a generalized-ensemble simulation of the new algorithm (solid curve). Here, the weight we used for the latter simulation is given by Eq. (12) with $n_F = 19$ and $E_0 = E_{GS} = -12.2$ kcal/mol. For the canonical run the curve stays around the value $E = -7$ kcal/mol with small thermal fluctuations, reflecting the low-temperature nature. The run has apparently

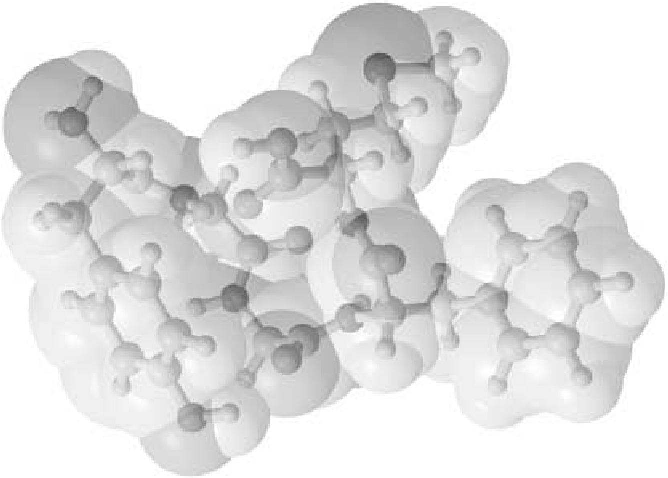

Fig. 1. Ground-state structure (Structure A) of Met-enkephalin for KONF90 energy function. The figures were created with Molscript [31] and Raster3D [32].

been trapped in a local minimum, since the mean energy at this temperature is $< E >_T = -11.1$ kcal/mol as found by a multicanonical simulation in Ref. [28]. On the other hand, the simulation based on the new weight covers a much wider energy range than the canonical run. It is a random walk in energy space, which keeps the simulation from getting trapped in a local minimum. It indeed visits the ground-state region several times in 1,000,000 MC sweeps. These properties are common features of generalized-ensemble methods.

We now examine the dependence of the simulations on the values of the exponent m in our weight (see Eqs. (5) and (12)) and demonstrate that $m = n_F$ is indeed the optimal choice. Setting $E_0 = E_{GS} = -12.2$ kcal/mol, we performed 10 independent simulation runs of 50,000 MC sweeps with various choices of m. In Table 1 we list the lowest energies obtained during each of the 10 runs for five choices of m values: 9.5 ($= \frac{n_F}{2}$), 14, 19 ($= n_F$), 50, and 100. The results from regular canonical simulations at $T = 50$ K with 50,000 MC sweeps are also listed in the Table for comparison. If m is chosen to be too small (e.g., $m = 9.5$), then the weight follows a power law in which the suppression for higher energy region is insufficient (see Eq. (8)). As a result, the simulations tend to stay at high energies and fail to sample low-energy configurations. On the other hand, for too large a value of m (e.g., $m = 100$), the weight is too close to the canonical weight, and therefore the simulations will get trapped in local minima. It is clear from the Table that $m = n_F$ is the optimal choice. In this case the simulations found the ground-state configurations 80 % of the time (8 runs out of 10 runs). This should be compared with 90 % and 40 % for multicanonical annealing

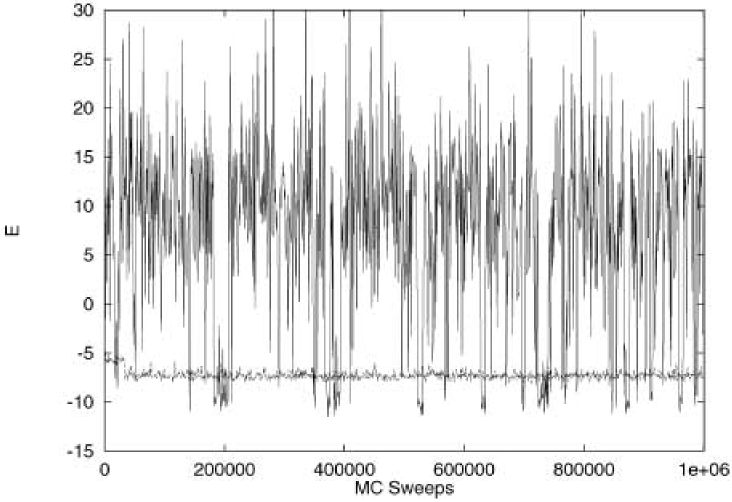

Fig. 2. Time series of the total energy E_{tot} (kcal/mol) from a regular canonical simulation at temperature $T = 50$ K (dotted curve) and that from a simulation of the present method with the parameters: $E_0 = -12.2$ kcal/mol, $n_F = 19$, and $T = 50$ K (solid curve).

and simulated annealing algorithms, respectively, in simulations with the same number of MC sweeps [28].

The weight factor of the present algorithm just depends on the knowledge of the global-minimum energy E_{GS} (see Eq. (12)). If its value is known, which is the case for some systems, the weight is completely determined. However, if E_{GS} is not known, we have to obtain its best estimate E_0. In Table 2 we list the lowest energies obtained during each of 10 independent simulation runs of 200,000 MC sweeps with $m = n_F = 19$. Four choices were considered for the E_0 value: $-12.2, \ -13.2, \ -14.2$, and -15.2 kcal/mol. We remark that E_0 has to underestimate E_{GS} to ensure that $E - E_0$ cannot become negative. Our data show that an accuracy of $1 \sim 2$ kcal/mol in the estimate of the global-minimum energy is required for Met-enkephalin.

Since the simulation by the present algorithm samples a large range of energies (see Fig. 2), we can use the reweighting techniques [27] of Eq. (13) to construct canonical distributions and calculate thermodynamic quantities as a function of temperature over a wide temperature range.

All thermodynamic quantities were then calculated from a single production run of 1,000,000 MC sweeps which followed 10,000 sweeps for thermalization. At the end of every fourth sweep we stored the energy of the conformation, the corresponding volume, and the overlaps of the conformations with the (known) reference states for further analyses. Here, we approximate the volume of the peptide by its solvent excluded volume (in Å^3) which is calculated by a variant [33] of the double cubic lattice method [34]. Our definition of the overlap, which

Table 1. Lowest energy (in kcal/mol) obtained by the present method with several different choices of the exponent m.

E_0	$E_{GS} = -12.2$	-12.2	-12.2	-12.2	-12.2	
m	$\frac{n_F}{2} = 9.5$	14	$n_F = 19$	50	100	∞^{a}
Run						
1	0.8	−5.2	−11.8	−6.9	−6.8	−4.2
2	−1.4	−2.6	−11.5	−7.1	−7.7	−5.2
3	0.1	−6.8	−11.5	−6.9	−4.9	−11.8
4	0.5	−5.5	−11.7	−8.2	−9.9	−7.1
5	−1.0	−3.4	−11.6	−7.4	−12.0	−3.3
6	1.1	−6.4	−11.6	−10.1	−8.8	0.9
7	−1.3	−5.1	−8.5	−8.7	−8.7	−5.3
8	0.4	−3.3	−9.7	−10.8	−9.5	−6.3
9	1.2	−8.1	−11.6	−12.0	−6.8	−6.4
10	1.2	−3.3	−11.9	−10.8	−9.5	−4.7
$< E_{min} >^{\mathrm{b}}$	0.2 (1.0)	−5.0 (1.8)	−11.1 (1.1)	−8.9 (1.9)	−8.5 (2.0)	−5.3 (3.2)
$n_{GS}{}^{\mathrm{c}}$	0/10	0/10	8/10	1/10	1/10	1/10

[a] The case for $m = \infty$ stands for a regular canonical run at $T = 50$ K.
[b] $< E_{min} >$ is the average of the lowest energy obtained by the 10 runs (with the standard deviations in parentheses).
[c] n_{GS} is the number of runs in which a conformation with $E \leq -11.0$ kcal/mol (the average energy at $T = 50$ K) was obtained.

measures how much a given conformation resembles a reference state, is given by

$$O(t) = 1 - \frac{1}{90\ n_F} \sum_{i=1}^{n_F} \left| \alpha_i^{(t)} - \alpha_i^{(RS)} \right| \ , \tag{14}$$

where $\alpha_i^{(t)}$ and $\alpha_i^{(RS)}$ (in degrees) stand for the n_F dihedral angles of the conformation at t-th Monte Carlo sweep and the reference state conformation, respectively. Symmetries for the side-chain angles were taken into account and the difference $\alpha_i^{(t)} - \alpha_i^{(RS)}$ was always projected into the interval $[-180°, 180°]$. Our definition guarantees that we have

$$0 \leq \ < O >_T \ \leq 1 \ . \tag{15}$$

We remark that the average overlap $< O >_T$ approaches its limiting value zero (for $T \rightarrow \infty$) only very slowly as the temperature increases. For instance, at

Table 2. Lowest energy (in kcal/mol) obtained by the present method with several different choices of the free parameter E_0.

E_0	$E_{GS} = -12.2$	-13.2	-14.2	-15.2
m	$n_F = 19$	19	19	19
Run				
1	-11.8	-11.1	-10.5	-9.0
2	-11.9	-10.8	-8.3	-10.3
3	-11.9	-11.3	-11.6	-9.7
4	-11.9	-10.2	-10.9	-10.8
5	-11.8	-11.2	-6.9	-9.2
6	-11.3	-11.5	-10.8	-9.6
7	-11.9	-11.3	-8.3	-10.3
8	-11.8	-11.4	-5.9	-6.8
9	-12.0	-11.5	-10.6	-8.6
10	-11.7	-10.0	-10.3	-8.9
$< E_{min} >$[a]	-11.8 (0.2)	-11.0 (0.5)	-9.4 (1.9)	-9.3 (1.1)
n_{GS}[a]	10/10	7/10	1/10	0/10

[a] The definitions for $< E_{min} >$ and n_{GS} are given in the caption of Table 1.

$T = 1000$ K we found that the overlap O_A with the ground-state conformation (Structure A) still has an average value of $< O_A > \approx 0.3$. This is because $< O >_T \ = 0$ corresponds to a completely random distribution of dihedral angles which is energetically highly unfavorable due to the steric hindrance of both main and side chains. Note the obvious limit: $O_A \to 1$, as $T \to 0$, indicating that only the ground-state conformation exists at $T = 0$ K.

We expect the folding of proteins and peptides to occur in a multi-stage process. A common scenario for folding may be that first the polypeptide chain collapses from a random coil to a compact state. This coil-to-globule transition is characterized by the collapse transition temperature T_θ. In the next stage, a set of compact structures are explored and there exists a transition from one of many local minima in the set of compact structures into the native (ground-state) conformation. This second transition is characterized by the folding temperature T_f $(\leq T_\theta)$.

The first process is connected with a collapse of an extended coil structure into an ensemble of compact structures. This transition should be connected with a pronounced change in the average volume of the peptide as a function of temperature. The transition temperature T_θ can then be located from the position of the maximum of

$$\frac{d < V >_T}{dT} \equiv \beta^2 \left(< V E_{tot} >_T - < V >_T < E_{tot} >_T\right) \ . \tag{16}$$

The average volume and its derivative as a function of temperature are shown in Fig. 3. The second quantity measures the steepness of the decrease in volume. To quantify its value we divided our time series in 4 bins corresponding to 250,000 sweeps each, determined the position of the maximum for both quantities in each bin and averaged over the bins. In this way we found a transition temperature $T_\theta = 310 \pm 20$ K [35].

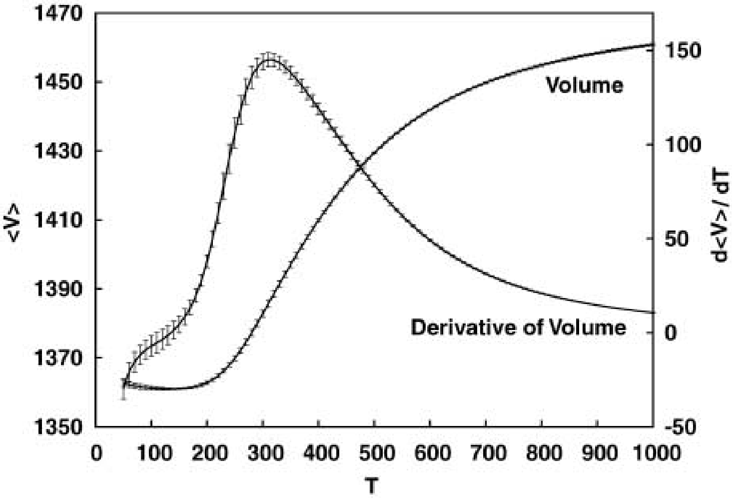

Fig. 3. Average volume $< V >_T$ and its derivative $d < V >_T /dT$ as a function of temperature.

The second transition which should occur at a lower temperature T_f is that from a set of compact structures to the "native conformation" that is considered to be the ground state of the peptide. This transition should be characterized by a rapid change in the average overlap $< O_A >_T$ with the ground-state conformation (see Eq. (14)) and a corresponding maximum in

$$\frac{d < O_A >_T}{dT} \equiv \beta^2 \left(< O_A E_{tot} >_T - < O_A >_T < E_{tot} >_T\right) \ . \tag{17}$$

Both quantities are displayed in Fig. 4, and we indeed find the expected behavior. The rapid change in the overlap is clearly visible and occurs at a temperature lower than the first transition temperature T_θ. We again try to determine its value by searching for the peak in $d < O_A >_T /dT$ in each of the 4 bins and averaging over the obtained values. In this way we find a transition temperature of $T_f = 230 \pm 30$ K [35].

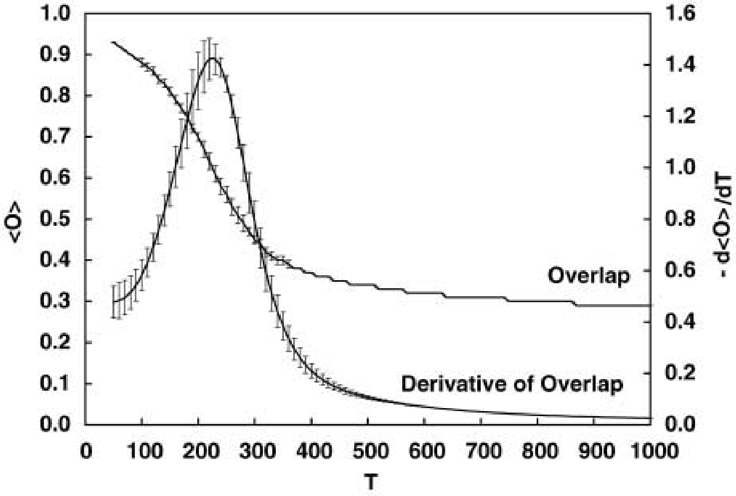

Fig. 4. Average overlap $< O_A >_T$ with the ground-state conformation and its derivative $-d < O_A >_T /dT$ as a function of temperature.

We now apply the above algorithm to a more direct evaluation of the free-energy landscape of Met-enkephalin [36]. The ensemble of low-temperature conformations of this peptide has been studied in detail [30,28,37,38]. These studies imply that there are two major groups of well-defined compact structures which are characterized (and stabilized) by specific hydrogen-bonding patterns. In Fig. 5 we show a sketch of the two structures. The figures were created with RasMol [39]. Structure A is the ground-state conformation in ECEEP/2 and is stabilized by the hydrogen bonds between Gly-2 and Met-5. Structure B, the second-lowest energy state, is characterized by the hydrogen bonds between Tyr-1 and Phe-4.

Since it is not feasible to plot the free energy G as a function of all order parameters, we only plot G as a function of a suitable combination of the relevant order parameters of the molecule. We chose to plot the free energy $G(V, O_A)$ as a function of volume V and overlap O_A with the known ground state (Structure A) and $G(O_A, O_B)$ as a function of the overlaps with the ground state (O_A) and with Structure B (O_B), where we have

$$G(V, O_A) = -k_B T \log P(V, O_A) \ , \tag{18}$$

and

$$G(O_A, O_B) = -k_B T \log P(O_A, O_B) \ . \tag{19}$$

Here, $P(V, O_A)$ and $P(O_A, O_B)$ are respectively the probabilities to find a peptide conformation with values V, O_A and O_A, O_B. We study these quantities for temperatures $T = 1000$ K, $T = T_\theta = 300$ K, $T = T_f = 230$ K, and $T = 150$ K. They are shown in Figs. 6 and 7 [36].

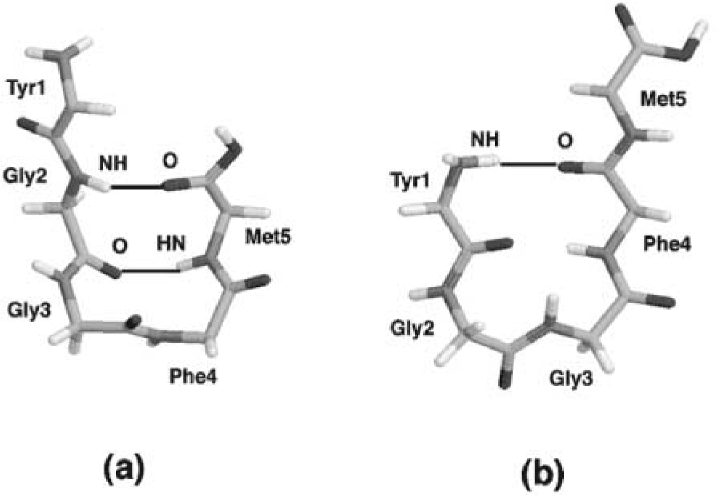

Fig. 5. Backbone structures of the two dominant low-energy structures with their characteristic hydrogen bonding: (a) Structure A, the ground state in ECEPP/2 (with a KONF90 energy of −12.2 kcal/mol) and (b) Structure B, the local minimum with the second-lowest potential energy (−11.0 kcal/mol in KONF90). The figures were created with RasMol [39].

In Fig. 6(a) we show the free energy landscape as a function of volume and overlap with the known ground state (Structure A) at the high-temperature region ($T = 1000$ K). In the contour plots, the contour lines mark multiples of k_BT (therefore different for different temperatures but appropriate to understand the folding mechanism). We see that the free energy has its minimum at large volumes (≈ 1470 Å^3) and values of the overlap $O_A \approx 0.3$. Small volumes and larger values of the overlap are suppressed by many orders of k_BT. Hence, extended random coil structures are favored at this temperature. The picture changes dramatically once we reach the collapse temperature T_θ, shown in Fig. 6(b). At this temperature a large part of the V-O_A space can be sampled in a simulation. The contour plot shows that regions with both small and large volumes and almost all values of O_A lie within the 2 k_BT contour. This indicates that at this temperature the cross over between extended and compact structures happens with a small thermodynamic barrier between them. By lowering the temperature to $T_f = 230$ K, we now observe strong evidence for a funnel-like landscape [40,41] (Fig. 6(c)). At this temperature the drive towards the native configuration is

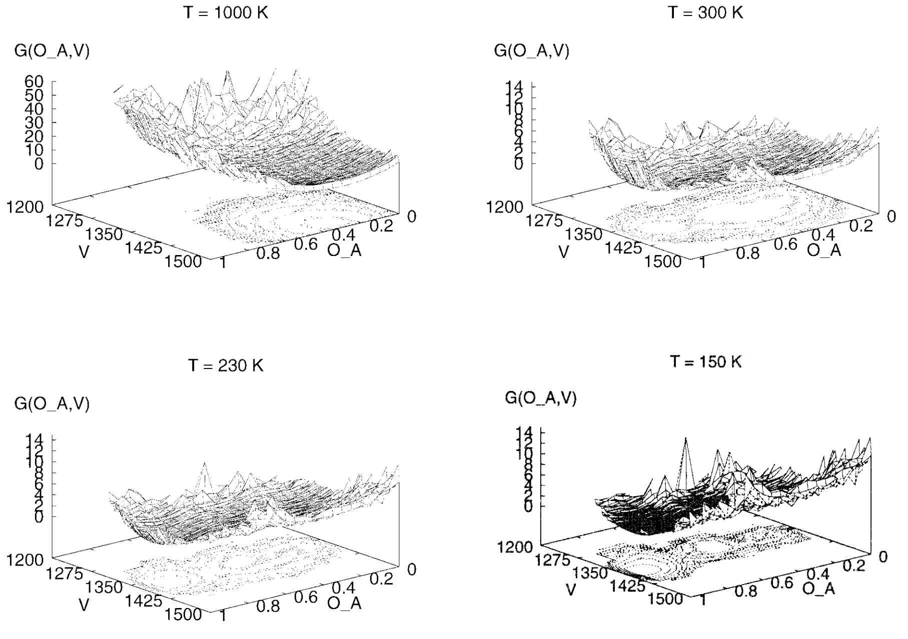

Fig. 6. Free energy $G(V, O_A)$ (in kcal/mol) as a function of both peptide volume V (in Å^3) and overlap O_A (as defined in the text) for (a) $T = 1000$ K, (b) $T = 300$ K, (c) $T = 230$ K, and (d) $T = 150$ K. Both the free energy surface and the contour plot are shown. The contour lines are multiples of k_BT. $G(V, O_A)$ was normalized such that $\min(G(V, O_A)) = 0$.

dominant and no long-lived traps exist. There is clearly a gradient towards the ground-state structure ($O_A \approx 1$), but other structures with similar volume (characterized by values of $O_A \approx 0.5$) are only separated by free energy barriers of order 1 k_BT. Below this temperature we expect that the ground state is clearly favored thermodynamically and separated from other low energy states by free energy barriers of many orders of k_BT. This can be seen in Fig. 6(d) where at $T = 150$ K where other low energy states have free energies of 3 k_BT higher than the ground state and are separated by an additional barrier of 2 k_BT.

The above picture is supported by the plots for the free energy as a function of both the overlap O_A with the ground state and the overlap O_B with Structure B. Fig. 7(a) shows again the high-temperature situation. The free energy has its minimum at small values of the overlap indicating that both conformers appear with only very small frequency at high temperature. At $T = 300$ K, the collapse temperature, again a large part of the space of possible configurations (characterized by values of O_A and O_B) lies within the $2k_BT$ contour as is clear from Fig. 7(b). At the folding temperature $T_f = 230$ K a funnel in the energy

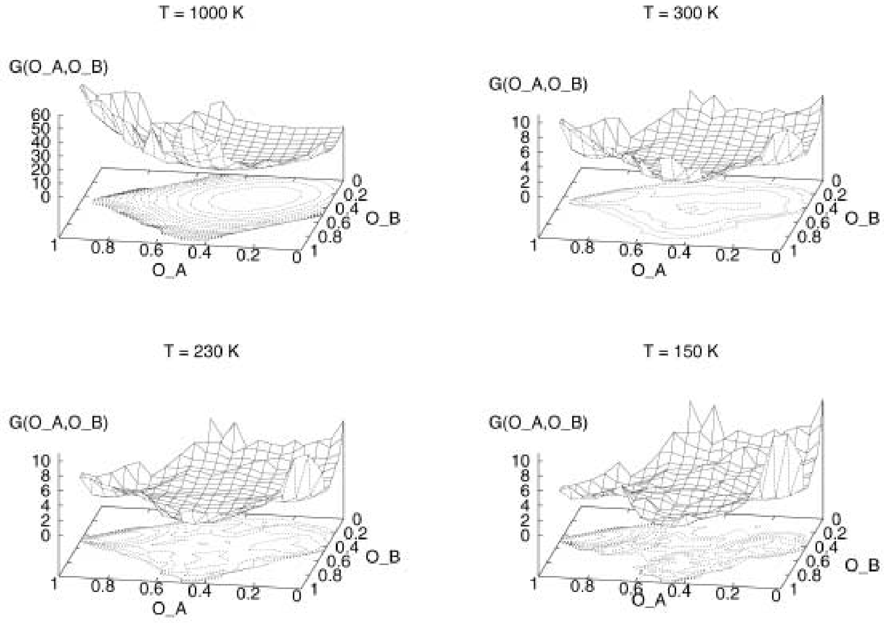

Fig. 7. Free energy $G(O_A, O_B)$ as a function of both overlaps O_A and O_B (as defined in the text) for (a) $T = 1000$ K, (b) $T = 300$ K, (c) $T = 230$ K, and (d) $T = 150$ K. Both the free energy surface and the contour plot are shown. The contour lines are multiples of k_BT. $G(O_A, O_B)$ was normalized such that $\min(G(O_A, O_B)) = 0$.

landscape appears with a gradient towards the ground state, but Fig. 7(c) shows that there are various other structures, the most notable of which is Structure B (where $O_B \approx 1$), with free energies 3 k_BT higher than the ground-state conformation but separated from each other and the ground state only by free energy barriers less than 1 k_BT. No other long-lived traps are populated. Hence, the funnel at T_f is reasonably smooth. Folding routes include direct conversion from random-coil conformations into Structure A or some short trapping in Structure B region before reaching Structure A region, but at the folding temperature it is possible to reach the ground state from any configuration without getting kinetically trapped. Canonical Monte Carlo runs at a fixed temperature ($T = 230$ K) were performed and confirmed this picture (data not shown). We observed that some of the runs went directly from the unfolded state to the ensemble of folded conformations in State A, while in other runs trapping in State B occurred first before folding into the ground-state structure. These kinetic runs therefore support our observation that Met-enkephalin is a good folder. Fig. 7(d) shows the situation for $T = 150$ K where we expect an onset of glassy behavior. Again one sees a funnel-like bias toward the ground state, however, the funnel is no longer smooth and the free energy landscape is rugged. Free energy barriers of many

$k_B T$ now separate different regions and would act as long-lived kinetic traps in a canonical simulation rendering folding at this temperature extremely difficult.

4 Conclusions

In this article we have reviewed the uses of Tsallis statistical mechanics in the protein folding problem. A generalized-ensemble algorithm based on Tsallis probability weight factor was introduced. While the determination of the optimal weight factor for other generzlized-ensemble algorithms is non-trivial, its determination in the Tsallis generalized-ensemble algorithm is greatly simplified. The effectiveness of the method was tested for a penta peptide, Met-enkephalin, and the folding properties of this peptide were elucidated.

Acknowledgements
The authors would like to thank Masato Masuya and José Onuchic for collaborations on the works presented in the present article. Our simulations were performed on the computers of the Computer Center at the Institute for Molecular Science, Okazaki, Japan. This work is supported, in part, by a grant from the Research for the Future Program of Japan Society for the Promotion of Science (JSPS-RFTF98P01101).

References

1. M. Vásquez, G. Némethy, and H. A. Scheraga, Chem. Rev. **94**, 2183 (1994).
2. M. Levitt, Curr. Opin. Struct. Biol. **6**, 193 (1996).
3. C. J. Epstain, R. F. Goldberger, and C. B. Anfinsen, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 439 (1963).
4. a) C. Levinthal, J. Chim. Phys. **65**, 44 (1968).
   b) D. B. Wetlaufer, Proc. Natl. Acad. Sci. USA **70**, 691 (1973).
5. B. J. Berne and J. E. Straub, Curr. Opin. Struct. Biol. **7**, 181 (1997).
6. Y. Okamoto, Recent Res. Devel. in Pure & Applied Chem. **2**, 1 (1998).
7. U. H. E. Hansmann and Y. Okamoto, Curr. Opin. Struct. Biol. **9**, 177 (1999).
8. U. H. E. Hansmann and Y. Okamoto, in *Ann. Rev. Comput. Phys. VI*, ed. by D. Stauffer (World Scientific, Singapore, 1999) pp. 129–157.
9. B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).
10. U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
11. a) A. P. Lyubartsev, A. A. Martinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
    b) E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
12. B. Hesselbo and R. B. Stinchcombe, Phys. Rev. Lett. **74**, 2151 (1995).
13. U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **18**, 920 (1997).
14. U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger, Chem. Phys. Lett. **259**, 321 (1996).
15. N. Nakajima, H. Nakamura, and A. Kidera, J. Phys. Chem. **101**, 817 (1997).
16. C. Bartels and M. Karplus, J. Phys. Chem. B **102**, 865 (1998).

17. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
18. U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **56**, 2228 (1997).
19. U. H. E. Hansmann, F. Eisenmenger, and Y. Okamoto, Chem. Phys. Lett. **297**, 374 (1998).
20. U. H. E. Hansmann and Y. Okamoto, Braz. J. Phys. **29**, 187 (1999).
21. I. Andricioaei and J. E. Straub, Phys. Rev. E **53**, R3055 (1996).
22. I. Andricioaei and J. E. Straub, J. Chem. Phys. **107**, 9117 (1997).
23. U. H. E. Hansmann, Physica A **242**, 250 (1997).
24. a) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **79**, 2361 (1975).
    b) G. Némethy, M. S. Pottle, and H. A. Scheraga, J. Phys. Chem. **87**, 1883 (1983).
    c) M.J. Sippl, G. Némethy, and H. A. Scheraga, J. Phys. Chem. **88**, 6231 (1984).
25. a) H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, Chem. Lett. **1991**, 213 (1991).
    b) Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, Protein Eng. **4**, 639 (1991).
26. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
27. A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
28. U. H. E. Hansmann and Y. Okamoto, J. Phys. Soc. Jpn. **63**, 3945 (1994); Physica A **212**, 415 (1994).
29. Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. USA **84**, 6611 (1984).
30. Y. Okamoto, T. Kikuchi, and H. Kawai, Chem. Lett. **1992**, 1275 (1992).
31. P. J. Kraulis, J. Appl. Cryst. **24**, 946 (1991).
32. a) D. Bacon and W. F. Anderson, J. Mol. Graphics **6**, 219 (1988).
    b) E. A. Merritt and M. E. P. Murphy, Acta Cryst. D **50**, 869 (1994).
33. M. Masuya, in preparation.
34. F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, and M. Scharf, J. Comput. Chem. **16**, 273 (1995).
35. U. H. E. Hansmann, M. Masuya, and Y. Okamoto, Proc. Natl. Acad. Sci. USA **94**, 10652 (1997).
36. U. H. E. Hansmann, Y. Okamoto, and J. N. Onuchic, PROTEINS: Struct. Funct. Genet. **34**, 472 (1999).
37. F. Eisenmenger and U. H. E. Hansmann, J. Phys. Chem. B **101**, 3304 (1997).
38. A. Mitsutake, U. H. E. Hansmann, and Y. Okamoto, J. Mol. Graphics Mod. **16**, 226 (1998).
39. R. A. Sayle and E. J. Milner-White, TIBS **20**, 374 (1995).
40. J. N. Onuchic, Z. Luthey-Schulten, and P.G. Wolynes, Ann. Rev. Phys. Chem. **48**, 545 (1997).
41. K. A. Dill and H. S. Chan, Nature Struct. Biol. **4**, 10 (1997).

Subject Index