PHYSICAL REVIEW A VOLUME 38, NUMBER 9 NOVEMBER 1, 1988

Microscopic and macroscopic entropy

Kristian Lindgren
Physical Resource Theory Group, Chalmers University of Technology, S-412 96 Göteborg, Sweden

(Received 4 April 1988)

We consider physical systems in which microstates can be represented as symbol sequences. The internal correlations in a microstate give restrictions to its "randomness," quantified by the measure entropy of that microstate. It is shown that, in the thermodynamic limit, the measure entropy for the microstates has an ensemble average equal to the thermodynamic entropy. A simple explanation is that the correlations necessary to generate the whole ensemble can be found in almost any one of its microstates. If microscopic phases are present, there are subensembles, each given by the statistics in almost any one of its microstates. The typical microstate in an equilibrium ensemble can be found by maximization of the measure entropy under energy constraints and general constraints on probability distributions for symbol sequences. The typical microstate does not contain information in correlations of lengths greater than interaction distance, a fact which simplifies the calculation of the entropy for systems with finite interaction length. We exemplify this by deriving the entropy for a monatomic ideal gas and the one-dimensional Ising model with an external field.

I. INTRODUCTION

The entropy of a physical system is associated with the uncertainty of the actual microstate of the system, given the macrostate. The macrostate is a probability distribution over the possible microstates, and the entropy is defined as the ensemble average of $-\ln p_i$, where p_i is the probability for microstate i. Expressed in this way, the entropy is not an average of microscopic properties. However, for a system in the thermodynamic limit, all correlations necessary to generate the macrostate are present in each of the microstates, except in microstates of a subset of measure zero. Then there must exist some microscopic property that has an average equal to the statistical mechanics entropy. This property turns out to be the measure entropy that can be interpreted as a quantitative measure of randomness. For states which can be coded into symbol sequences, in the limit of infinite length, the measure entropy is the average Shannon information[1] per symbol, i.e., the information which remains when the information in correlations of all lengths has been subtracted from the total information per cell.

The idea of looking for this microscopic counterpart to entropy comes from Bennett,[2] who has shown that there is a microscopic property, the algorithmic complexity, which has an ensemble average equal to the thermodynamic entropy for an equilibrium system in the thermodynamic limit. The algorithmic complexity for a sequence of binary symbols is defined as the length of the shortest program, or algorithms, for a general-purpose computer, to generate the sequence.[3,4] The algorithmic complexity is generally applicable, since it does not presuppose any probability distributions on substrings of the binary sequence. The drawback is that there does exist a general procedure for finding the smallest program, and thus the algorithmic complexity is not generally computable. However, if the binary sequence is the outcome of a stationary stochastic process, the algorithmic complexity is often well approximated by the measure entropy.[5–7]

We consider physical systems in which the microstates can be expressed as binary sequences, and the extension to lattices of higher dimension is indicated. Then it follows that the thermodynamic entropy of a system, for which almost all microstates in the thermodynamic limit have a correlation information which decreases sufficiently fast with distance, is equal to the ensemble average of its microscopic measure entropies. We will discuss this relation in information-theoretical terms using the measure entropy and its complement, the information in correlations of different lengths.

It is well known that the equilibrium macrostate is the distribution with maximal entropy given some constraints, e.g., a specified expectation value on the energy.[8] If the entropy of the equilibrium macrostate in the thermodynamic limit is given by the measure entropy of a typical microstate, this microstate can be found by maximizing its measure entropy under the given constraints.

Note that a typical microstate is such that its probability distributions on finite subsequences describe the whole ensemble. Then it follows that the thermodynamics based on these microstates is formally identical to the thermodynamics of classical lattice gas theory (see, e.g., Ref. 9). Our aim is, however, to show that, in the thermodynamic limit, the measure entropy is a microscopic quantity which has an ensemble average equal to the thermodynamic entropy. The idea of such a quantity originates from Boltzmann,[10] who defined a microscopic property, the H function, which under certain conditions was proven to be nondecreasing in time—the H theorem. The aim was to establish a close relation between the H function and the thermodynamic entropy. Discussions of difficulties with Boltzmann's approach have been given in numerous works, e.g., Refs. 11 and 12. We expect the measure entropy for a microstate to be nondecreasing in time, implying that the second law of thermodynamics holds for single microstates. This will be discussed in terms of states in infinite cellular automata.

If there are states with different correlation functions, each having a maximum in measure entropy, we say that there are different microscopic phases. For a one-dimensional system, it is shown that if the number of microscopic phases increases less than exponentially with the size of the system, then the relation between measure entropy and thermodynamic entropy still holds. It is easily seen that, for a one-dimensional equilibrium system, there is no information in correlations longer than the interaction distance for the typical microstate. This corresponds to a result in lattice gas theory where such a statement holds for the macrostate.[9]

The typical microstate is derived for two simple systems. The measure entropy for the typical microstate in an ideal gas gives the Sackur-Tetrode equation, and the measure entropy for the microstate in the one-dimensional Ising model with an external field also agrees with the exact expression of the thermodynamic entropy.

II. MEASURE ENTROPY AND THERMODYNAMIC ENTROPY

We review some information-theoretical properties of lattice systems which are useful in the analysis of internal randomness and correlations of microstates. The microstates are assumed to be written in the form of binary symbol sequences.

Consider an infinite sequence of zeros and ones for which the probabilities $p_m(i_1 i_2 \cdots i_m)$ for length-m subsequences $i_1 i_2 \cdots i_m$, where $i_k \in \{0,1\}$ with $k=1,2,\ldots,m$, are well defined. The m-point Shannon entropy S_m is then defined as[1]

$$S_m = \sum_{i_1,i_2,\ldots,i_m} p_m(i_1 i_2 \cdots i_m) \ln \frac{1}{p_m(i_1 i_2 \cdots i_m)} . \tag{1}$$

It has the properties $\Delta S_m = S_m - S_{m-1} \geq 0$ ($S_0 = 0$) and $\Delta^2 S_m = \Delta S_m - \Delta S_{m-1} \leq 0$, $m \geq 2$. The second-order difference gives the information in correlations k_m of length m,

$$k_m = -\Delta^2 S_m = \sum_{i_1,i_2,\ldots,i_m} p_m(i_1 i_2 \cdots i_m) \ln \frac{p_m(i_1 i_2 \cdots i_m) p_{m-2}(i_2 i_3 \cdots i_{m-1})}{p_{m-1}(i_1 i_2 \cdots i_{m-1}) p_{m-1}(i_2 i_3 \cdots i_m)} , \tag{2}$$

which is a relative information or Kullback information between p_m and a maximum-entropy estimate $\tilde{p}_m(i_1 i_2 \cdots i_m) = p_{m-1}(i_1 i_2 \cdots i_{m-1}) p_{m-1}(i_2 i_3 \cdots i_m) / p_{m-2}(i_2 i_3 \cdots i_{m-1})$ of p_m given correlations $m' < m$.[13] The (spatial) measure entropy s_μ is defined as (see, e.g., Ref. 14)

$$s_\mu = \lim_{m \to \infty} \left[\frac{1}{m} S_m \right] = \lim_{m \to \infty} \Delta S_m , \tag{3}$$

which is the average entropy per cell. The entropy difference converges faster, and for Markov chains of memory m, i.e., sequences in which each symbol depends only on the m preceding ones, one has $s_\mu = \Delta S_m$. (The measure entropy for a lattice of dimension n is given by the average entropy per cell for an n-dimensional block when the side lengths increase to infinity.)

The total information of 1 bit, or equivalently ln 2, per lattice site can be decomposed into three non-negative terms,[13] $\ln 2 = (\ln 2 - S_1) + (S_1 - s_\mu) + s_\mu$. The first term can be interpreted as an information quantity due to different probabilities for 0's and 1's. The second term can be written as the sum of correlation information from all lengths, $\sum_m k_m$. Thus the third term s_μ can be interpreted as the random information per lattice site. The average m-point Shannon entropy approaches the measure entropy as $m \to \infty$, and the rate of convergence η is given by

$$\eta = \lim_{m \to \infty} (S_m - m s_\mu) . \tag{4}$$

This can be written as a product of mean correlation length (in information-theoretical terms) and correlation information,[7] and it was introduced by Grassberger as a measure of complexity for binary sequences.[15]

For physical systems we usually need the possibility of choosing among an infinite number of symbols, each corresponding to a certain physical state. In the limit of a continuum of states, represented by the set of real numbers, we will need a function for the density of states. Let the microstate be an infinite sequence of real numbers, and let $v_m(r_1 r_2 \cdots r_m)$ be the density of states for a subsequence $(r_1 r_2 \cdots r_m)$. The probability density for this sequence is denoted $p_m(r_1 r_2 \cdots r_m)$ and satisfies the normalization condition

$$\int dr_1 \int dr_2 \cdots \int dr_m \, v_m(r_1 r_2 \cdots r_m) p_m(r_1 r_2 \cdots r_m) = 1 . \tag{5}$$

Then the measure entropy for the infinite sequence is defined, cf. Ref. 16,

$$s_\mu = \lim_{m \to \infty} \left[\frac{1}{m} \int dr_1 \int dr_2 \cdots \int dr_m \, v_m(r_1 r_2 \cdots r_m) \times p_m(r_1 r_2 \cdots r_m) \times \ln \frac{1}{p_m(r_1 r_2 \cdots r_m)} \right] . \tag{6}$$

We consider a physical system with N available microstates, and we assume it to be the macrostate

$$P=\{p_i\}_{i=1}^{i=N} . \tag{7}$$

The thermodynamic entropy (in units of the Boltzmann constant k_B) is then

$$S(P)=\sum_{i=1}^{N} p_i \ln\frac{1}{p_i}=\left\langle \ln\frac{1}{p_i}\right\rangle . \tag{8}$$

Thus it is an ensemble average over $-\ln p_i$, a property which depends on the probability distribution P. Bennett has shown that if the macrostate is concisely describable, then the entropy S can be approximated with the ensemble average over the algorithmic information $H(i)$ of the microstates i. The following inequalities hold for the probability distribution P:[2]

$$S(P)\le \sum_{i=1}^{N} p_i H(i) \le S(P)+H(P)+O(1) . \tag{9}$$

Here $H(P)$ denotes the minimum information needed for describing the macrostate P. The $O(1)$ term comes from the ambiguity due to the choice of general-purpose computer. For a thermodynamic system, the entropy P is of the same order of magnitude as the number of particles, say, 10^{23} binary digits (bits), while $H(P)$ corresponds to some thousand bits in a nondeterministic algorithm (Monte Carlo program) which samples the ensemble. We also expect $H(P)$ to increase slower than linearly in the size of the system. Thus, the algorithmic information is a property of the microstate, and, in the thermodynamic limit, its ensemble average is equal to the thermodynamic entropy.

The relation (9) is conceptually important, but, since $H(i)$ is not a computable function, the relation cannot be directly used as a tool for calculating the entropy. However, since the algorithmic information has been shown to be equal to the measure entropy for almost all sequences which are outcomes of a stationary stochastic process,[5–7] there is a direct relation between the entropy $S(P)$ and the ensemble average of the measure entropy $s_\mu(i)$ of the microstates $i=1,2,\ldots,N$.

In the following we restrict ourselves to one-dimensional systems, and sketch a proof for the relation between the measure entropy of the microstates and the entropy of the macrostate. Further, we assume that the microstates can be described as length-L sequences of zeros and ones, $\alpha=(j_1 j_2 \cdots j_L)$. (Systems of higher dimension can be analyzed in a similar way.) We denote the set of microstates by Ω. Then the macrostate P_L is a probability distribution over Ω,

$$P_L=\{p(\alpha)\}_{\alpha\in\Omega} . \tag{10}$$

We divide Ω into nonoverlapping subsets Ω_k such that each microstate in Ω_k has approximately the same correlations for lengths $m\le M$. Here M is chosen to be large compared to typical correlation lengths that appear in the microstates, but small compared to L. (However there may occur low probability microstates with long correlation lengths for which such an inequality does not exist.)

Note that if a microstate α is generated by a nondeterministic algorithm which uses relatively short correlation lengths $l_c \ll M \ll L$, then the concepts defined in Eqs. (1)–(4) can be applied to that sequence. The microstates in Ω_k can then be regarded as generated by the same algorithm, but with different random numbers.

Let $Q_m(\alpha)$ denote the probability distribution over m-length sequences β in the microstate α,

$$Q_m(\alpha)=\{q_m(\alpha,\beta)\}_{\beta\in B_m} , \tag{11}$$

where B_m denote the set of all length-m binary sequences. Then the closeness between the elements in Ω_k can be expressed as

$$\frac{1}{m}\,|S(Q_m(\alpha_1))-S(Q_m(\alpha_2))| < \epsilon \ll 1 ,\qquad m\le M,\ \text{all } \alpha_1,\alpha_2\in\Omega_k . \tag{12}$$

Let $p_\Omega(\Omega_k)$ be the probability in P_L for a randomly chosen microstate to belong to Ω_k. In the limits $L\to\infty$, $M\to\infty$, and $L/M\to\infty$, we expect for simple physical systems that there are only a finite number of subensembles Ω_k with $p_\Omega(\Omega_k)>0$, and we say that they correspond to different microscopic phases. The microstates in such a subensemble have all the same internal correlations, and they can be regarded as being generated by a specific stochastic process σ, defined by the probability distribution $Q_M(\alpha)$ ($M\to\infty$ and $\alpha\in\Omega_k$). It is assumed that, for almost all microstates generated by σ, the information in correlations of length m, k_m, decreases sufficiently fast for η to be finite. (This may not be valid at phase-transition points.)

We use these facts in the following decomposition of specific thermodynamic entropy $s(P_L)=S(P_L)/L$, which holds in the limit $L\to\infty$:

$$s(P_L)=\frac{1}{L}\sum_\sigma p(\sigma)\sum_\alpha p(\alpha\,|\,\sigma)\ln\frac{1}{p(\alpha\,|\,\sigma)} +\frac{1}{L}\sum_\alpha p(\alpha)\sum_\sigma p(\sigma\,|\,\alpha)\ln\frac{p(\sigma\,|\,\alpha)}{p(\sigma)} . \tag{13}$$

Here $p(\alpha\,|\,\sigma)$ denotes the probability for the process σ to generate the sequence α, and $p(\sigma\,|\,\alpha)$ denotes the probability that microstate α has been generated by σ. Since almost all microstates generated by σ have same measure entropy, the first term is the ensemble average of the measure entropies. The last term is the relative information between $p(\sigma)$ and $p(\sigma\,|\,\alpha)$, averaged over α and divided by L. The relative information is less than or equal to the logarithm of the number of microscopic phases. Thus, if the number of microscopic phases increases slower than exponentially in L, this term vanishes. Then we get, in the thermodynamic limit, that the specific thermodynamic entropy is equal to the ensemble average of the measure entropies of the microstates,

$$s(P_L) \to \sum_{\sigma} p(\sigma) s_\mu(\alpha_\sigma) = \sum_{\alpha \in \Omega} p(\alpha) s_\mu(\alpha) \quad \text{as } L \to \infty \ , \tag{14}$$

where α_σ denotes a typical microstate in the subensemble corresponding to σ.

If the entropy of the distribution over microscopic phases is zero, i.e., all microstates can be viewed as being generated by the same algorithm, then Eq. (14) expresses the fact that, in the thermodynamic limit, the ensemble is given by the statistics in one microstate. To find the thermodynamic entropy is then equivalent to finding a typical microstate and calculating its measure entropy.

The thermodynamic entropy for an equilibrium system under energy constraints can be found by maximizing the macroscopic entropy with the constraint of a specified expectation value of the internal energy.[8] This problem is usually solved by introducing Lagrangian multipliers μ and β for the normalization and the energy constraints, respectively, and the solution is the Gibbs distribution. Then the multiplier β is equal to $1/(k_B T)$, and μ/β is the Helmholtz free energy.

The measure entropy of a typical microstate in an equilibrium ensemble is also given by the maximum under energy constraints, e.g., a specified energy for the microstate, and the constraints given by general properties of probability distributions on symbol sequences. These are the normalization constraint, the requirement of positive probabilities, and the condition that a summation over last index in an length-m distribution shall give the same length-$(m-1)$ distribution as is given by a summation over first index. If the interaction distance is finite, m, the typical microstate does not contain any information in correlations of lengths greater than m, and the measure entropy is equal to ΔS_m. This follows immediately from the fact that, without violating the constraints, the correlation information from lengths $m' > m$ can be decreased to zero by defining $p_{m'}(i_1 i_2 \cdots i_{m'}) = \tilde{p}_{m'}(i_1 i_2 \cdots i_{m'})$, step by step starting with $m' = m+1$, see Eq. (2). The maximization problem can be formulated as follows:

$$\max \left\{ \sum_{i_1, i_2, \ldots, i_m} p_m(i_1 i_2 \cdots i_m) \ln \frac{p_{m-1}(i_1 i_2 \cdots i_{m-1})}{p_m(i_1 i_2 \cdots i_m)} \right\} \tag{15}$$

under constraints on normalization and energy u,

$$\sum_{i_1, i_2, \ldots, i_m} p_m(i_1 i_2 \cdots i_m) = 1 \ , \tag{16}$$

$$\sum_{i_1, i_2, \ldots, i_m} p_m(i_1 i_2 \cdots i_m) h(i_1 i_2 \cdots i_m) = u \ , \tag{17}$$

where h is the Hamiltonian, and the general conditions for probability distributions on symbol sequences. If the problem is solved by introducing Lagrangian multipliers, the multiplier β associated with the energy constraint (17) has the same meaning here as in the maximum of the macrostate entropy, $\beta = 1/(k_B T)$. In Sec. III the maximization problem (15)–(17) is solved for an ideal gas and for the one-dimensional Ising model with an external field.

The maximization problem may give more than one solution, implying that there are different microscopic phases, each having the same value on the measure entropy of their typical microstates.

The second law of thermodynamics tells us that the entropy $S(P)$ for a closed thermodynamic system increases in time. Equivalently, the ensemble average of the measure entropy for the microstates must increase in time. But we expect a stronger statement than that to be valid: The measure entropy of a single microstate almost never decreases in time. We discuss this briefly in terms of infinite cellular automata, which define an interesting class of systems for studying microscopic dynamics. An infinite cellular automaton maps an infinite sequence onto the set of infinite sequences by a rule which depends on the local symbols only. In a previous article[7] we studied a class of deterministic rules which were sufficiently reversible to conserve the measure entropy of the sequence. But then the entropy is constant, and we do not get any dissipation. What happens in such a dynamics is that information moves into correlations of larger and larger distances, so that, locally, the entropy of the microstate appears to increase. It was also shown that if the rule is modified by an arbitrarily small amount of noise, then the measure entropy increases to a maximum value. In the time evolution of the microstate of a physical system, the information found in correlations must decrease, for example by an increased correlation length which makes the correlation information more sensitive to stochastic influence.

III. EXAMPLES: THE IDEAL GAS AND THE ONE-DIMENSIONAL ISING MODEL

We solve the maximization problem, Eqs. (15)–(17), for two simple examples: the ideal gas and the one-dimensional Ising model with external field.

Consider first a gas of N noninteracting point particles with mass m in a volume V, with a constraint on the expected value on the kinetic energy of the particles. Divide the volume V into boxes of volume $l^3 \ll V/N$. To each box two state variables are associated. The first variable is 1 if there is a particle in the box and 0 otherwise, and the second variable is the momentum $\mathbf{p}$ of the particle, if there is one. Then we get a cubic lattice of these pairs of variables, with a density of ones $p_1 = Nl^3/V$ and a density of zeros $p_0 = 1 - p_1$. The thermodynamic limit is given by $N \to \infty$ with N/V constant. Since the absence of interaction implies that there is no information in correlations, we are to maximize the measure entropy, which is given by the entropy of one cell (the density of states is l^3/h^3), cf. Eq. (6),

$$S_1 = p_0 \ln \frac{1}{p_0} + \frac{l^3}{h^3} \int d^3\mathbf{p}\, p_1 f(\mathbf{p}) \ln \frac{1}{p_1 f(\mathbf{p})} \ , \tag{18}$$

under normalization and energy constraints. Here, $f(\mathbf{p})$ is the probability density of $\mathbf{p}$, satisfying the normalization condition, Eq. (5). The maximum value is obtained for $f(\mathbf{p})$ being the Maxwell-Boltzmann distribution, and it is

$$S_1 = \frac{Nl^3}{V}\left[\ln\frac{V}{N}+\frac{5}{2}+\frac{3}{2}\ln\frac{2\pi m k_B T}{h^2}+O\left[\frac{Nl^3}{V}\right]\right] . \tag{19}$$

This is the measure entropy per lattice site, and if it is multiplied by $k_B V/l^3$, we get, in the limit $l \to 0$, the Sackur-Tetrode equation for thermodynamic entropy for N particles,

$$S = k_B N\left[\ln\frac{V}{N}+\frac{5}{2}+\frac{3}{2}\ln\frac{2\pi m k_B T}{h^2}\right] . \tag{20}$$

In the second example, we consider an infinite one-dimensional sequence of spins, and we denote spin up by 1 and spin down by 0, so that a microstate is an infinite sequence of zeros and ones. Suppose that the interaction energy between nearest neighbors of antiparallel and parallel spin is $+J$ and $-J$, respectively, and that an external field contributes with energy $+H$ and $-H$ for spin down and spin up, respectively. If the probabilities for configurations of pairs of spins are $p_0 = P(00)$, $p_1 = P(01) = P(10)$, and $p_2 = P(11)$, with $p_0 + 2p_1 + p_2 = 1$, then the energy e per site can be written

$$e = J(2p_1 - p_0 - p_2) + H(p_0 - p_2) . \tag{21}$$

Since the interaction is over pairs only, the measure entropy is equal to the entropy difference $S_2 - S_1$. Thus by introducing Lagrangian multipliers μ for the normalization constraint and β for the energy constraint we have to maximize

$$\begin{aligned} S(p_0,p_1,p_2,\beta,\mu) = {} & p_0 \ln\frac{p_0+p_1}{p_0} + p_1 \ln\frac{(p_0+p_1)(p_1+p_2)}{p_1^2} + p_2 \ln\frac{p_1+p_2}{p_2} \\ & + \beta[e - J(2p_1 - p_0 - p_2) - H(p_0 - p_2)] + \mu(1 - p_0 - 2p_1 - p_2) . \end{aligned} \tag{22}$$

The solution is

$$\mu = \ln\{\cosh(\beta H) + [\sinh^2(\beta H) + e^{-4\beta J}]^{1/2}\} , \tag{23}$$

$$p_1 = \frac{(e^{\beta H} - e^{-\mu})(e^{-\beta H} - e^{-\mu})}{2[\sinh^2(\beta H) + e^{-4\beta J}]^{1/2}} , \tag{24}$$

$$p_{1\pm1} = \frac{e^{\pm\beta H} - e^{-\mu}}{2[\sinh^2(\beta H) + e^{-4\beta J}]^{1/2}} , \tag{25}$$

which gives the thermodynamic entropy per spin (in units of k_B)

$$S = \mu + \frac{2\beta J e^{-4\beta J - \mu} - \beta H \sinh(\beta H)}{[\sinh^2(\beta H) + e^{-4\beta J}]^{1/2}} , \tag{26}$$

in agreement with the expression that one gets in traditional approaches.[17]

ACKNOWLEDGMENTS

I wish to thank Karl-Erik Eriksson and Mats Nordahl for many stimulating discussions. This work was supported by the Swedish Natural Science Research Council.

---

[1] C. E. Shannon, Bell System Tech. J. **27**, 379 (1948).

[2] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).

[3] A. N. Kolmogorov, Prob. Peredachi. Inf. **1**, 3 (1965).

[4] G. J. Chaitin, J. Assoc. Comput. Mach. **13**, 547 (1966).

[5] A. K. Zhvonkin and L. A. Levin, Russ. Math. Surveys **25**, 83 (1970).

[6] A. A. Brudno, Usp. Mat. Nauk. **33**, 207 (1977).

[7] K. Lindgren, Complex Systems **1**, 529 (1987).

[8] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).

[9] M. Fannes and A. Verbeure, Commun. Math. Phys. **96**, 115 (1984).

[10] L. Boltzmann, Wien. Ber. **66**, 275 (1872).

[11] P. Ehrenfest and T. Ehrenfest, *The Conceptual Foundations of Statistical Mechanics* (Cornell University Press, Ithaca, 1959).

[12] I. Prigogine, *From Being to Becoming* (Freeman, San Francisco, 1980).

[13] K.-E. Eriksson and K. Lindgren, Phys. Scr. **35**, 388 (1987).

[14] S. Wolfram, Physica D **10**, 1 (1984).

[15] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).

[16] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[17] C. J. Thompson, *Mathematical Statistical Physics* (Princeton University, Princeton, N.J., 1979).