PHYSICAL REVIEW E **70**, 017102 (2004)

Rényi entropies and observables

Bernhard Lesche
Departamento de Física, Universidade Federal de Juiz de Fora, Juiz de Fora MG, Brazil
(Received 23 March 2004; published 26 July 2004)

Evidence is given that Rényi entropies of macroscopic thermodynamic systems defined on the bases of probabilities of microstates cannot be related to observables. The notion of observable is clarified.

DOI: 10.1103/PhysRevE.70.017102 PACS number(s): 65.40.Gr

Recently Abe [1] and Jizba and Arimitsu [2] investigated an issue that I raised in 1981 [3]. In that paper I showed that it is possible to find pairs of probability assignments that are so close to each other that they cannot be distinguished by any reasonable test and yet their Rényi entropies, with parameters $\alpha \neq 1$, may differ considerably. This result was used to argue that, for macroscopic thermodynamic systems, these entropies could not be related to observable quantities. Abe extended the idea to other types of entropies, and Jizba and Arimitsu criticized the arguments. On reading my old paper again I noticed that it was extremely concise and some points were formulated in an improper way. This may have led modern authors to some wrong interpretations. Therefore I shall begin the present investigation clarifying the corresponding points.

The central point in Ref. [3] was not the question of whether Rényi entropies were observable. In fact Rényi entropies are not observable. Not even Shannon entropy is observable. An observable is a collection of experimental yes-no questions about a system that forms a Boolean algebra, together with a corresponding collection of Borel sets on the real line. Observables can be represented by essentially self-adjoint operators on Hilbert space that are independent of the state of the physical system. Therefore clearly $I_1(\rho) = -\mathrm{Tr}(\rho \ln \rho)$ is not an observable. Shannon entropy and all other Rényi entropies are state functionals—i.e., mappings that map the set of states into the real numbers. It may seem that the distinction between observables and state functionals is a mere question of words. However, all understanding of nature is based on language (mathematical and ordinary language) and the improper use of words can quickly lead to wrong results. For instance, we find in Ref. [2] (second page, top of second column) "…. systems whose statistical fluctuations in $G(x)$ would change too dramatically with a small change in the state variable x." Here $G(x)$ stands for a Rényi entropy or Tsallis-Havrda-Charvat entropy or any other state functional and x is the probability assignment. This is wrong. $G(x)$ will not show any statistical fluctuations. A given ensemble—i.e., a class of independent experiments defined by some experimental procedure—is described by a probability assignment x. Therefore, all members of the ensemble correspond to the same value $G(x)$. On the other hand, if one measured an observable, different members of the ensemble may give different results. This variation of experimental outcomes among different members of an ensemble is called statistical fluctuation. The instability of state functionals has nothing to do with statistical fluctuations.

The central point of the investigation of Ref. [3] was the question whether Rényi entropies could be related to observables of a thermodynamic system. Let us see what is meant by "relate a state functional to an observable." Let $\hat{A}, \hat{B}, \ldots, \hat{H}$ be macroscopic commeasurable observables used for the thermodynamic description of a macroscopic system. The total number of these observables may be large, but it is supposed to be extremely small in comparison with the total number N of microscopic degrees of freedom. Let $|a,b,\ldots,h,j\rangle$ be a basis of common eigenstates, where $a,b,\ldots,h$ are the corresponding eigenvalues and j is an index of degeneracy. Let $W_{a,b,\ldots,h}$ be the dimension of the common eigenspace with eigenvalues $a,b,\ldots,h$. These degrees of degeneracy are typically of the order of M^N, where M stands for some large number. The Boltzmann entropy is then defined as

$$\hat{S} = \sum_{a,b,\ldots,h,j} |a,b,\ldots,h,j\rangle \ln W_{a,b,\ldots,h} \langle a,b,\ldots,h,j|. \quad (1)$$

This is an observable. The typical states discussed in thermodynamics have the form

$$\rho = \sum_{a,b,\ldots,h,j} |a,b,\ldots,h,j\rangle \frac{P_{a,b,\ldots,h}}{W_{a,b,\ldots,h}} \langle a,b,\ldots,h,j|. \quad (2)$$

That means the probabilities of microstates $|a,b,\ldots,h,j\rangle$ are constant within the common eigenspaces. The Shannon–von Neumann entropy of such a kind of state is

$$\begin{aligned} I_1(\rho) = -\mathrm{Tr}(\rho \ln \rho) &= -\sum_{a,b,\ldots,h,j} \frac{P_{a,b,\ldots,h}}{W_{a,b,\ldots,h}} \ln \frac{P_{a,b,\ldots,h}}{W_{a,b,\ldots,h}} \\ &= -\sum_{a,b,\ldots,h} P_{a,b,\ldots,h} \ln \frac{P_{a,b,\ldots,h}}{W_{a,b,\ldots,h}} \\ &= \sum_{a,b,\ldots,h} P_{a,b,\ldots,h} \ln W_{a,b,\ldots,h} \\ &\quad - \sum_{a,b,\ldots,h} P_{a,b,\ldots,h} \ln P_{a,b,\ldots,h}. \end{aligned} \quad (3)$$

The second term is usually negligible as compared to the first one. Therefore, the Shannon entropy is approximately equal to the expectation value of the Boltzmann entropy. In most cases the probability distributions of the macroscopic variables are sharply peaked around the expectation values. In these cases one has the following relation of Shannon entropy and observable Boltzmann entropy: *The* *probability of observing an eigenvalue of entropy far away from the Shannon value is extremely small*. Note that this kind of relation between state functional and observable necessarily involves probabilities or expectation values. The issue of Ref. [3] was the question of whether one could find observables $\hat{I}_\alpha$ that might be related to the Rényi entropies $I_\alpha=(1-\alpha)^{-1}\ln[\Sigma_i(p_i)^\alpha]$ in a similar way.

Before investigating the question of instabilities of Rényi entropies let us add one more remark on state functionals and observables: Thermodynamic systems are special because of their large number of microscopic degrees of freedom. But apart from that they are nothing special. So if Rényi entropies are state functionals and not observables, this should be true for different systems too. So how can it be that Rényi entropies are routinely measured in numerous situations such as cryptography, chaotic dynamical systems, earthquake analysis, etc. [2]? In fact, what is routinely being measured is

$$I_\alpha(r) = \frac{1}{1-\alpha}\ln\left(\sum_i (r_i)^\alpha\right), \tag{4}$$

where r_i are relative frequencies. Relative frequencies can be measured; probabilities cannot. A probability P of some event is again a state functional (a linear one) and it can be related to a relative frequency R in Z repeated experiments (with $Z \gg 1/P$), very much the same way as Shannon entropy can be related to Boltzmann entropy: *The probability of finding R far away from P is very small*. Again this statement is a probabilistic one. A probability is nothing but an opinion (which may be based on objective facts by information theoretic rules). We do not measure opinions in physics laboratories. Relative frequencies and mean values can be measured; probabilities and expectation values cannot be measured. The pseudo "measurement" of Rényi entropies by substituting probabilities by relative frequencies cannot be performed in the case of Rényi entropies that are defined on the bases of probabilities of microstates of thermodynamic systems. The probabilities of individual microstates p_i are far too small to be related to relative frequencies. This is also true for the case of Shannon entropy. Therefore it is remarkable that Shannon entropy can be related to an observable.

Now let us address the question of the unstable behavior of Rényi entropies. Jizba and Arimitsu [2] argue that it is not enough to give one example of a pair of probability assignments that are so close that no test can distinguish them and whose Rényi entropies differ considerably in order to show that Rényi entropies cannot be related to observables. Their argument is correct. In fact, if the problem was limited to a small number and a well known type of states and moreover to states that were not of special interest, one could remedy the problem easily by excluding the problematic cases. In fact, Jizba and Arimitsu show that for large numbers of microstates the problematic sector is confined to a set of Bhattacharyya measure zero. This fact is interesting in itself and adds one more item to phenomena related to very large numbers. However, the small Bhattacharyya measure of the problematic cases may not be a valid argument to save the usefulness of Rényi entropies of thermodynamic systems. After all we are not talking about a set of measure zero in phase space, but about a measure in the space of probability assignments. The exact significance of the Bhattacharyya measure in statistical descriptions of thermodynamic systems has to be given. We shall give evidence that in fact all interesting initial states describing thermodynamic systems are exactly of the problematic type.

It is not of much interest to investigate a system when it already reached its final equilibrium. So let us study an initial state with entropy less than $\ln n$, where n is the total number of microstates. In any normal macroscopic experiment starting from a nonequilibrium state means that the initial entropy is smaller than $\ln n$ by some macroscopic amount—i.e., by some number of the order N. Furthermore, the initial macrostate $[a_I, b_I, \ldots, h_I]$ is usually known. Therefore, the typical initial state is of the form

$$\rho = \sum_j |a_I, b_I, \ldots h_I, j\rangle p_j \langle a_I, b_I, \ldots h_I, j|. \tag{5}$$

The special characteristics of this kind of state are the following: (1) The number of occupied microstates is of the order M^N and therefore usually any individual probability p_j is extremely small (of the order of M^{-N}). (2) The number of empty microstates is larger than the number of occupied microstates by a huge factor, which is also of the order of $\tilde{M}^N$. Essentially these two characteristics make this sort of state a problem case. I shall not give the most general proof but limit myself to the simplest, but most important, example, which shows the general idea clearly. I shall assume the p_j to be constant:

$$\rho = \sum_j |a_I, b_I, \ldots h_I, j\rangle \frac{1}{W_{a_I, b_I, \ldots h_I}} \langle a_I, b_I, \ldots h_I, j|. \tag{6}$$

All Rényi entropies of this state have the same value, which is the Boltzmann entropy of the macro-state $[a_I, b_I, \ldots, h_I]$:

$$I_\alpha(\rho) = S(a_I, b_I, \ldots h_I) = \ln W_{a_I, b_I, \ldots h_I}. \tag{7}$$

Now, imagine that a friend of ours enters the laboratory and criticizes our experiment. He claims our preparation of state may in some cases result in the macro-state $[\bar{a}, \bar{b}, \ldots, \bar{h}]$. His density operator would be

$$\tilde{\rho} = \sum_j |a_I, b_I, \ldots, h_I, j\rangle \frac{1-\delta}{W_{a_I, b_I, \ldots, h_I}} \langle a_I, b_I, \ldots, h_I, j| + \sum_j |\bar{a}, \bar{b}, \ldots, \bar{h}, j\rangle \frac{\delta}{W_{\bar{a}, \bar{b}, \ldots, \bar{h}}} \langle \bar{a}, \bar{b}, \ldots, \bar{h}, j|. \tag{8}$$

Now, if the probability δ is, say, 10^{-100}, we will not be able to convince our friend that our probability assignment is better than his by showing experimental results. The number 10^{-100} is far too small to get sufficient statistics in 2×10^{10} years (age of the universe) even if we could perform 10^{20} experiments per second. However, 10^{-100} is astronomically huge in comparison with the probabilities of individual microstates. The Rényi entropy of the friend's probability assignment is

$$I_\alpha(\tilde{\rho}) = \frac{1}{1-\alpha}\ln\{(1-\delta)^\alpha (W_{a_I,\ldots,h_I})^{1-\alpha} + \delta^\alpha (W_{\bar{a},\bar{b},\ldots,\bar{h}})^{1-\alpha}\}$$
$$= \frac{1}{1-\alpha}\ln\left\{(W_{\bar{a},\bar{b},\ldots,\bar{h}})^{1-\alpha}\delta^\alpha \times\left(1+\frac{(1-\delta)^\alpha}{\delta^\alpha}\frac{(W_{a_I,\ldots,h_I})^{1-\alpha}}{(W_{\bar{a},\bar{b},\ldots,\bar{h}})^{1-\alpha}}\right)\right\}$$
$$= \ln W_{\bar{a},\bar{b},\ldots,\bar{h}} + \frac{\alpha}{1-\alpha}\ln\delta + \frac{1}{1-\alpha}\ln\left(1+\frac{(1-\delta)^\alpha}{\delta^\alpha}\frac{(W_{a_I,\ldots,h_I})^{1-\alpha}}{(W_{\bar{a},\bar{b},\ldots,\bar{h}})^{1-\alpha}}\right). \tag{9}$$

The first term is of order N. If $|1-\alpha| \gg |\ln\delta|/N$, the second term is negligible as compared to the first one (for instance, with $N\approx 10^{24}$ and $\delta = 10^{-100}$ our argument is restriction to α values with $|1-\alpha| \gg 10^{-22}$). To estimate the third term we now distinguish the following two cases: (1) If $\alpha > 1$, we shall assume that our friend thought of a state $[\bar{a},\bar{b},\ldots,\bar{h}]$ with smaller entropy than the main state $[a_I,b_I,\ldots,h_I]$. If $S(a_I,\ldots,h_I) - S(\bar{a},\ldots,\bar{h})$ is macroscopic (of the order N), the third term is clearly also negligible as compared to the first one. (2) If $\alpha < 1$, we assume that the friend thought of a state $[\bar{a},\bar{b},\ldots,\bar{h}]$ whose entropy is macroscopically larger than $S(a_I,\ldots,h_I)$. Again the third term will be negligible. So, in either case, the Rényi entropy of the friend's probability assignment would essentially be the entropy of the irrelevant state $[\bar{a},\bar{b},\ldots,\bar{h}]$, which is far away from our value [Eq. (7)].

One may easily extend the idea to states that do not have a large region of unoccupied microstates. For example, one may think of a sharply peaked probability distribution of Boltzmann entropies. Generally emptying the meaningless small tail of the distribution will give similar disasters.

The argument that the problematic sector of probability assignments is limited to a set of measure zero would mean that we have no chance to encounter such type of problem in our daily work. The present counterexample shows that the problematic behavior of Rényi entropies shows up for states that are really used in common descriptions of thermodynamic systems. Therefore, one may conclude that the Bhattacharyya measure is not a relevant measure to judge the unstable behavior of Rényi entropies. If this measure really had significance, the situation of Rényi entropies of thermodynamic systems would be even worse. Let us take a closer look at the demonstration of Ref. [2], which shows that the unstable sector is limited to a small measure. In order to define the Bhattacharyya measure the authors associate a vector ξ to a probability assignment $\mathcal{P}=\{p_1,\ldots,p_n\}$ putting $\xi_i=\sqrt{p_i}$. The Bhattacharyya measure is then defined by the ordinary (appropriately normalized) surface area on the unit l_2-sphere in the ξ-space. For the case $\alpha > 1$ Jizba and Arimitsu show that for $\varepsilon > 0$ and any p with $1 < p < \alpha/(\alpha-1)$ the inequalities

$$\|\xi\|_{2\alpha} \geq E(\|\xi'\|_{2\alpha})\exp\{-2\varepsilon[E(\|\xi'\|_{2\alpha})]^{p-1}\}, \tag{10}$$

$$\|\xi\|_{2\alpha} \leq E(\|\xi'\|_{2\alpha})\exp\{\varepsilon[E(\|\xi'\|_{2\alpha})]^{p-1}\} \tag{11}$$

hold for almost all ξ (their Bhattacharyya measure is arbitrarily close to 1 as n increases). In these relations $\|\xi\|_q$ designates the Hölder l_q norm of the vector ξ and $E[\|\xi'\|_{2\alpha}]$ is the mean value of $\|\xi'\|_{2\alpha}$ calculated with the Bhattacharyya measure. Now let us take two arbitrary probability assignments $\mathcal{P}$ and $\mathcal{Z}$ from this set of measure of almost 1 that satisfies inequalities (10) and (11). Their difference in Rényi entropy satisfies the relation

$$\frac{|I_\alpha(\mathcal{P})-I_\alpha(\mathcal{Z})|}{I_{\alpha MAX}} = \frac{2\alpha}{(\alpha-1)\ln n}\left|\ln\left(\frac{\|\xi^{(\mathcal{P})}\|_{2\alpha}}{\|\xi^{(\mathcal{Z})}\|_{2\alpha}}\right)\right| \leq \frac{2\alpha}{(\alpha-1)\ln n}\left|\ln\left(\frac{\exp\{\varepsilon[E(\|\xi\|_{2\alpha})]^{p-1}\}}{\exp\{-2\varepsilon[E(\|\xi\|_{2\alpha})]^{p-1}\}}\right)\right| = \frac{6\alpha\varepsilon}{(\alpha-1)\ln n}[E(\|\xi\|_{2\alpha})]^{p-1}. \tag{12}$$

According to Ref. [2], $E(\|\xi\|_{2\alpha})$ approaches zero when n goes to infinity. Furthermore, the last expression is proportional to ε, which may be arbitrarily small. So what Jizba and Arimitsu show is that the Rényi entropies are essentially constant on a set of measure almost 1 (the case $\alpha < 1$ is similar). Thus these functions would not be interesting at all. Jizba and Arimitsu also suggest to remove the instability problem by coarse graining the probability assignments. However, in the counterexample given in this paper the states are already coarse grained and this coarse graining is based on the macroscopic input information. Any further coarse graining would change the physics. At any rate, coarse graining only substitutes a problematic pair of probability assignments ($\mathcal{P}$, $\mathcal{P}'$) by a different one ($\tilde{\mathcal{P}},\tilde{\mathcal{P}}'$) with $I_\alpha(\tilde{\mathcal{P}})\approx I_\alpha(\tilde{\mathcal{P}}')$. Of course one can always find infinitely many $\tilde{\mathcal{P}}'$ in a δ vicinity of $\mathcal{P}$ that have practically the same Rényi entropy as the original assignment. That does not solve the problem. The problematic assignment $\mathcal{P}'$ still exists. The only way out seems to be to exclude the problematic $\mathcal{P}'$. However, this should not be done in a causistic manner but by some general rule. The counterexample given in the present paper seems to make it difficult to find such general rule. The physicist who still believes that Rényi entropies of thermodynamic systems can be related to observables should give the corresponding observable explicitly, indicate how it can be measured in laboratories, and describe the way it can be related to a Rényi entropy.

Summarizing we found more evidence that indicates that the usefulness of Rényi entropies, with $\alpha \neq 1$ is limited to systems with a small number n of states. Further we clarified the notions of observable and state functional. This distinction is essential to understand the whole problem. For example, the fact that some relationships between observables show discontinuities at first-order phase transitions has nothing to do with the present issue. The stability of a state functional is only a necessary condition for the existence of an associated observable. It is remarkable that Shannon entropy does have an associated observable. One may now, following the ideas of Ref. [1], investigate whether other stable entropies also have corresponding observables.

It is a pleasure to thank C. Tsallis and S. Abe for bringing my attention back to a subject that I abandoned many years ago.

[1] S. Abe, Phys. Rev. E **66**, 046134 (2002).
[2] P. Jizba and T. Arimitsu, Phys. Rev. E **69**, 026128 (2004).
[3] B. Lesche, J. Stat. Phys. **27**, 419 (1982).