Typicality for Generalized Microcanonical Ensembles

Peter Reimann

Universität Bielefeld, Fakultät für Physik, 33615 Bielefeld, Germany
(Received 17 July 2007; published 19 October 2007; publisher error corrected 17 March 2008)

For a macroscopic, isolated quantum system in an unknown pure state, the expectation value of any given observable is shown to hardly deviate from the ensemble average with extremely high probability under generic equilibrium and nonequilibrium conditions. Special care is devoted to the uncontrollable microscopic details of the system state. For a subsystem weakly coupled to a large heat bath, the canonical ensemble is recovered under much more general and realistic assumptions than those implicit in the usual microcanonical description of the composite system at equilibrium.

Basic questions of statistical physics have gained renewed interest with the discovery of various work and fluctuation theorems [1]. A further topic which has attracted much attention concerns the foundation of the canonical formalism [2–5]. One of its key ingredients consists of shifting the focus from the traditional statistical equilibrium ensembles back to the role and predictability of one single experimental realization of a system (and its environment), described theoretically by a quantum mechanical pure state. In essence, the main message of the seminal works [2,3], for which the name "canonical typicality" has been coined in [2], is as follows. Consider the usual canonical setup, i.e., an isolated "supersystem," compound of a "large" thermal bath B, a comparatively "small" subsystem of actual interest S, and a negligibly weak coupling between them. Those energy eigenstates of the compound $S + B$ with energy-eigenvalues in the interval $[E − \Delta E, E]$ span a Hilbert space, from which we pick at random a pure state $|\psi\rangle$. The corresponding projector $|\psi\rangle\langle\psi|$ gives rise to a mixed state ρ_S of the subsystem S by tracing out the bath degrees of freedom. Now, the remarkable finding of Refs. [2,3] is that ρ_S will be extremely close to the standard canonical density operator ρ_{can} of the subsystem S for the overwhelming majority of random pure states $|\psi\rangle$, hence the name "canonical typicality." In other words, whatever is the (unknown) pure state of the compound $S + B$, the outcome of any experiment on the subsystem S is practically the same as if it were in the canonically mixed state ρ_{can}. For a more detailed, precise, and also more general exposition, we refer to the original Refs. [2,3]. Further, it should be pointed out that essentially the same conclusion could be inferred from formula (C.17) of the formidable prior work [4]. For less general system classes and/or after performing an additional time average, closely related results have been obtained even earlier in Ref. [5].

Here, we will show that a quite similar "typicality" property already holds for any isolated system, even when it cannot be decomposed into subsystem S and bath B. In the special case that such a decomposition is possible, the original "canonical typicality" is recovered by "trac-

ing out the bath," thereby shedding also new light on the role of entanglement. A further main point is to abandon the quite unrealistic assumption of the previous works [2,3] that all energy eigenstates belonging to the preset energy interval $[E − \Delta E, E]$ contribute to the pure state $|\psi\rangle$ with equal probabilities and all the other energy eigenstates are excluded. Rather, in reality, one only knows the occupation probabilities of the energy levels very roughly, and hence the unknown details should not matter in the final results. This problem (and our solution) is clearly not restricted to the issue of typicality but concerns the standard microcanonical formalism in general.

Setup.—We consider a quantum mechanical system, whose Hilbert space $\mathcal{H}$ is spanned by the orthonormal basis $\{|n\rangle\}_{n=1}^{N}$, $N \leq \infty$. Hence, any pure state $|\psi\rangle$ is of the form

$$|\psi\rangle = \sum \frac{c_n}{||c||} |n\rangle, \qquad (1)$$

where $c_n \in \mathbb{C}$, $c := (c_1, \ldots, c_N)$, $||c|| := \sqrt{\sum |c_n|^2}$, and the sum runs from $n = 1$ to N. The division by $||c||$ will be particularly convenient for our purposes.

As in Refs. [2–5], we assume that the system is in some pure state $|\psi\rangle \in \mathcal{H}$, but we do not know which one. In other words, the c_n in (1) are randomly drawn from some probability density $p(c)$. Denoting the corresponding ensemble average of any function $g(c)$ by

$$\overline{g(c)} := \int g(c) p(c) \prod_{n=1}^{N} d(\mathrm{Re}\,c_n) d(\mathrm{Im}\,c_n), \qquad (2)$$

the expectation value of an arbitrary observable $A = A^\dagger$: $\mathcal{H} \to \mathcal{H}$ takes the form

$$\langle A \rangle := \overline{\langle\psi|A|\psi\rangle} = \mathrm{Tr}(\rho A) \qquad (3)$$

$$\rho := \overline{|\psi\rangle\langle\psi|} = \sum_{n,m} \overline{(c_n c_m^*/||c||^2)} |n\rangle\langle m|. \qquad (4)$$

For infinite dimensional systems, well defined limits $N \to \infty$ in (1)–(4) and later on are tacitly taken for granted.

Next, we turn to our two key assumptions regarding $p(c)$: (i) The c_n in (1) are statistically independent and, moreover, c_n and $e^{i\varphi_n}c_n$ are equally likely for arbitrary phases φ_n, or equivalently, $p(c)$ is of the form

$$p(c) = \prod_{n=1}^{N} p_n(|c_n|). \tag{5}$$

As a consequence, the density operator (4) takes the form

$$\rho = \sum \rho_n |n\rangle\langle n|, \qquad \rho_n := \overline{|c_n|^2/||c||^2} \tag{6}$$

with $\rho_n \geq 0$ and $\sum \rho_n = 1$. (ii) The mixed state ρ has a low purity, i.e.,

$$\mathrm{Tr}\,\rho^2 = \sum \rho_n^2 \ll 1, \tag{7}$$

or equivalently, $\max \rho_n \ll 1$, or equivalently, there are not just a few dominating ρ_n (summing up to almost unity). Before justifying these two assumptions, we show what can be concluded from them.

Typicality.—Our first objective is to show that $\langle \psi | A | \psi \rangle$ is typically very close to the average (3), i.e.,

$$\sigma_A^2 := \overline{[\langle \psi | A | \psi \rangle - \langle A \rangle]^2} \tag{8}$$

is small in an appropriate sense. Postponing the formal proof to a later paper, we adopt here a more heuristic line of reasoning, similar in spirit to Ref. [2]. We first consider the deviation of $s(c) := ||c||^2$ from its average $\bar{s} = \overline{||c||^2} = \sum \overline{|c_n^2|}$. Equation (5) implies that the $|c_n|^2$ are independent random variables and Eq. (7) implies that a large number of them significantly contributes to the sum s. Taking for granted finite variances

$$q_n := \overline{(|c_n^2| - \overline{|c_n^2|})^2}/\overline{|c_n^2|}^2 = \overline{|c_n^4|}/\overline{|c_n^2|}^2 - 1, \tag{9}$$

the law of large numbers implies that $s(c) - \bar{s}$ is an unbiased random variable with a very small standard deviation compared to $\bar{s}$. Hence, for our present purpose of estimating (8), we can replace $||c||$ in (1) in very good approximation by $\bar{s}^{1/2}$. As a consequence, one finds from (6) that $\overline{|c_n^2|} = \rho_n \bar{s}$. Similarly, introducing (1) with $||c|| \simeq \bar{s}^{1/2}$ into (8) and exploiting (5) yields

$$\sigma_A^2 = \sum_{n,m} \overline{|c_n^2 c_m^2|} \frac{(1 - \delta_{nm})|\tilde{A}_{mn}|^2 + \tilde{A}_{nn}\tilde{A}_{mm}}{\bar{s}^2}$$

$$\tilde{A} := A - \langle A \rangle, \quad \tilde{A}_{nm} := \langle n | \tilde{A} | m \rangle. \tag{10}$$

Observing that $\overline{|c_n^2 c_m^2|} = \rho_n \rho_m \bar{s}^2$ for $n \neq m$ according to (5) and $\overline{|c_n^2|} = \rho_n \bar{s}$, that $\overline{|c_n^4|} = (q_n + 1)\rho_n^2 \bar{s}^2$ according to (9), and that $\sum \rho_n \rho_m \tilde{A}_{nn} \tilde{A}_{mm} = (\sum \rho_n \tilde{A}_{nn})^2 = 0$ according to (3), (6), and (10), yields

$$\sigma_A^2 = \sum_{n,m} \rho_n \rho_m \tilde{A}_{nm} \tilde{A}_{mn} + \sum (q_n - 1)\rho_n^2 \tilde{A}_{nn}^2. \tag{11}$$

Exploiting again (3), (6), and (10), the first sum can be identified with $\mathrm{Tr}(\rho\tilde{A})^2$. Performing this trace with the help of the eigenvectors $|\nu\rangle$ and eigenvalues $\tilde{a}_\nu$ of $\tilde{A}$ yields $\sum \tilde{a}_\nu \tilde{a}_\mu |\langle \nu | \rho | \mu \rangle|^2$. From (10), it follows that $\tilde{a}_{\min} := \min \tilde{a}_\nu \leq 0$, $\tilde{a}_{\max} := \max \tilde{a}_\nu \geq 0$, and hence $\Delta_A := \tilde{a}_{\max} - \tilde{a}_{\min} \geq |\tilde{a}_\nu|$ for any ν. As a consequence, the first sum in (11) is bounded by $\Delta_A^2 \mathrm{Tr}\rho^2$. Similarly, in the second sum, we exploit that $\tilde{A}_{nn}^2 \leq \Delta_A^2$, yielding

$$\sigma_A^2 \leq \Delta_A^2 (\max_n q_n)(\mathrm{Tr}\rho^2). \tag{12}$$

In the special case that ρ is of the standard microcanonical form (see below), the same result also follows from Eq. (C.17) in [4].

Chebyshev's inequality implies for any given $\epsilon > 0$ that σ_A^2/ϵ^2 is an upper bound for the probability that $|\langle \psi | A | \psi \rangle - \langle A \rangle|$ exceeds ϵ. Exploiting (12), one finally infers for K observables $\{A_k\}_{k=1}^K$ and any $\epsilon > 0$ that

$$\mathrm{Prob}(\max_{k \leq K} |\langle \psi | A_k | \psi \rangle - \langle A_k \rangle|/\Delta_{A_k} \geq \epsilon) \leq K(\max_n q_n)(\mathrm{Tr}\rho^2)/\epsilon. \tag{13}$$

This is the first main result of our Letter. With (10), one sees that Δ_A equals the difference between the maximal and minimal eigenvalues A and hence quantifies the full range of all *a priori* possible values of $\langle \psi | A | \psi \rangle$. In (13), we tacitly excluded trivial observables A_k with $\Delta_{A_k} = 0$. The q_n in (9) and hence $\max q_n$ are dimensionless, non-negative numbers, typically of the order of unity, e.g., any Gaussian factor p_n in (5) yields $q_n = 1$. Hence, (13) with (7) imply typicality: a randomly sampled pure state $|\psi\rangle \in \mathcal{H}$ is very likely to yield expectation values $\langle \psi | A_k | \psi \rangle$ very close to the ensemble averages $\langle A_k \rangle = \mathrm{Tr}(\rho A_k)$ simultaneously for a quite large number K of arbitrary but fixed observables $A_1, \ldots, A_K$.

Generalized microcanonical formalism.—So far, typicality (13) applies to any quantum mechanical system satisfying (5) and (7). Next, we specifically justify these assumptions (5) and (7) for an isolated system at thermal equilibrium with $f = \mathcal{O}(10^{23})$ degrees of freedom and with $|n\rangle$ being the eigenvalues of the Hamiltonian $H = \sum E_n |n\rangle\langle n|$, $E_n \geq E_{n-1}$, $E_1 > -\infty$.

The assumption that coefficients c_n and $e^{i\varphi_n}c_n$ occur with equal probability in (1) is quite suggestive. Indeed, upon time evolution, the eigenvectors $|n\rangle$ acquire factors of the form $e^{-iE_n t/\hbar}$. Taking for granted that $p(c)$ does not change with time at thermal equilibrium, the invariance under $c_n \mapsto e^{i\varphi_n}c_n$ follows under rather mild and generic incommensurability conditions for the E_n. Exploiting this property, one readily concludes that $\bar{c}_n = 0$ for all n and that $\overline{c_n c_m} = 0$ for all $n \neq m$. In other words, the c_n are uncorrelated. This does not yet imply independence in principle, but in practice, it almost always does, and hence assumption (5) is reasonable.

The starting point of the seminal works [2,3] is the assumption that in (1), all coefficients c_n corresponding

to energies E_n within some preset energy interval $[E - \Delta E, E]$ are "equally likely," while all other c_n are zero. Denoting by c' the vector of all nonzero c_n and by c'' those which must be zero, this means that $p(c)$ can be so chosen that all c' of equal length $||c'||$ must be equally probable and thus $p(c)$ must be of the form $g(||c'||)\delta(c'')$ for some (properly normalized, non-negative) function g. Further, the division by $||c||$ in (1) implies that any such g actually yields the same distribution of vectors $|\psi\rangle$. Choosing a Gaussian g, the c_n can thus without loss of generality be considered [2] as independent, Gaussian, and satisfying (5). Moreover, ρ from (6) becomes the standard microcanonical density operator ρ_{mic} with equal weights $\rho_n > 0$ if $E_n \in [E - \Delta E, E]$ and $\rho_n = 0$ otherwise [6].

On one hand, our present approach thus includes the standard microcanonical formalism [6] and the starting point of Refs. [2,3] as special cases. On the other hand, the above observation that an entire class of different $p(c)$ actually yields—due to the division by $||c||$ in (1)—the same distribution of vectors $|\psi\rangle$ still remains true, and hence the assumption that one of those equivalent $p(c)$ satisfies (5) is very weak indeed. On top of that, our above proof of typicality can even be significantly generalized beyond the independence assumption (5) itself. In particular, the Gaussian adjusted projected measures (GAP) from [7] are still admissible.

Before turning to condition (7), we recall some standard notions and general properties regarding the energy spectrum $\{E_n\}$ [6]. Denoting the number of states within $[E - \Delta E, E]$ by $\Omega(E)$ and Boltzmann's constant by k_B, entropy and temperature follow as

$$S(E) := k_B \ln\Omega(E), \qquad T(E) := 1/S'(E). \qquad (14)$$

One finds [6] that $\Omega(E)$ is a very rapidly increasing function of E with typical values in range of $10^{\mathcal{O}(f)}$, $f = \mathcal{O}(10^{23})$. Hence, $\Omega(E)$ is largely independent of ΔE (provided $\Delta E \gg k_B T$), and its derivative can be identified with the density of states,

$$\Omega'(E) = \sum \delta(E_n - E), \qquad (15)$$

where the delta functions are slightly washed out to yield smooth functions in (14).

In view of this tremendous density of energy levels E_n, it is indeed quite convincing that no real system can be prepared such that just a few of them are populated with appreciable probability ρ_n, implying that (7) is indeed satisfied. In fact, an even stronger statement is quite plausible and will be derived dynamically elsewhere, namely, that these populations ρ_n can be written in the form

$$\rho_n = h(E_n) \qquad (16)$$

with a smooth function h, exhibiting appreciable variations only on scales much larger that $E_n - E_{n-1}$.

Because of (16), the energy distribution $p(E) := \langle\delta(H - E)\rangle$ can be rewritten with (3) and (6) as

$\sum \rho_n \delta(E_n - E) = h(E) \sum \delta(E_n - E)$ and, after washing out the delta functions according to (15), as

$$p(E) = h(E)\Omega'(E). \qquad (17)$$

In reality, after the experimentalist has prepared the system as carefully as possible, the only available knowledge about h and hence ρ is that the probability density (17) exhibits a relatively sharp peak (but still very wide compared to $E_n - E_{n-1}$). All further details of $p(E)$ are completely fixed by the given experimental setup, but it is impossible to know them. The only way out is to verify that these details "do not matter." Experimentally, this seems indeed to be the case, but theoretically, it has apparently not been demonstrated so far. On the contrary, for the usually considered ρ_{mic}, the concomitant details of $p(E)$ are in fact quite unrealistic.

As a first example, we show that the celebrated relation

$$-k_B \mathrm{Tr}(\rho \ln\rho) = S(E^*) \qquad (18)$$

indeed holds for any sharply peaked $p(E)$ with E^* located in the peak region: Exploiting (6) and (16) yields

$$\mathrm{Tr}(\rho \ln\rho) = \sum h(E_n) \ln h(E_n)$$

$$= \int dE\, h(E) \ln h(E) \sum \delta(E_n - E).$$

Because of (15) and (17), we can conclude that

$$\mathrm{Tr}(\rho \ln\rho) = \int dE\, p(E) \ln h(E). \qquad (19)$$

The integrand is dominated by the sharp peak of $p(E)$ since the possibly comparable variations of $h(E)$ [cf. (17)] are tamed by the logarithm. Hence, there exists an energy E^* within the peak region with the property

$$\int dE\, p(E) \ln h(E) = \ln h(E^*) \int dE\, p(E).$$

Likewise, Eq. (17) and the normalization of $p(E)$ imply

$$1 = \int dE\, p(E) = p(E^*)\epsilon = h(E^*)\Omega'(E^*)\epsilon, \qquad (20)$$

where ϵ essentially represents the peak width of $p(E)$. Equations (14) imply $\Omega' = \Omega/k_B T$, yielding with (19) and (20)

$$\mathrm{Tr}(\rho \ln\rho) = -\ln\Omega(E^*) - \ln[\epsilon/k_B T(E^*)]. \qquad (21)$$

Since $\ln\Omega = \mathcal{O}(10^{23})$ [see below (14)], the last term in (21) is negligible for any realistic ϵ, and with (14), our proof of (18) is completed. A somewhat similar calculation has been performed in chapter 12.3.2 of [4], but its purpose and physical content is quite different.

An analogous line of reasoning yields $\mathrm{Tr}\rho^2 \approx 1/\Omega(E^*) = 10^{-\mathcal{O}(10^{23})}$; i.e., typicality (13) is extremely well satisfied for a very large number K of observables.

Canonical formalism.—As in the introduction, we consider a subsystem S in weak contact with a much larger bath B, resulting in a compound $S + B$ with a product Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B$ and a Hamiltonian $H = H_S \otimes 1_B + 1_S \otimes H_B$, where 1_S is the identity on $\mathcal{H}_S$ and similarly for 1_B. Given $H_S = \sum E_j^S |j\rangle_{SS}\langle j|$ and $H_B = \sum E_k^B |k\rangle_{BB}\langle k|$, we thus have $H = \sum_{jk} E_{jk} |jk\rangle\langle jk|$ with $E_{jk} := E_j^S + E_k^B$ and $|jk\rangle := |j\rangle_S |k\rangle_B$; i.e., previous labels n now become jk. Since only subsystem observables $A = A_S \otimes 1_B$ are of interest, (3) can be rewritten as

$$\langle A \rangle = \mathrm{Tr}_S(\rho_{\mathrm{can}} A_S), \qquad \rho_{\mathrm{can}} := \mathrm{Tr}_B(\rho), \qquad (22)$$

with Tr_S the partial trace over $\mathcal{H}_S$, and similarly for Tr_B. Likewise, $\langle \psi | A | \psi \rangle = \mathrm{tr}(|\psi\rangle\langle\psi| A)$ can be rewritten as

$$\langle \psi | A | \psi \rangle = \mathrm{tr}_S(\rho_S A_S), \qquad \rho_S := \mathrm{Tr}_B(|\psi\rangle\langle\psi|). \qquad (23)$$

Equation (13) implies "canonical typicality" in the sense that for the vast majority of pure states $|\psi\rangle$ of the compound $S + B$, the corresponding mixed state of the subsystem ρ_S yields practically the same result for K subsystem observables $A_{S,1}, \ldots, A_{S,K}$ as the ensemble averaged mixed state ρ_{can}. If the bath B is sufficiently much larger than the subsystem S, then the extremely low purity of the compound $S + B$ implies typicality even for *all* possible subsystem observables A_S, giving rise to a natural metric [3] according to which the reduced density operator ρ_S itself is close to ρ_{can}.

Finally, one finds that any $p(E)$ in (17) with a sharp peak near E^* results in the same canonical density matrix

$$\rho_{\mathrm{can}} = Z^{-1} \exp\{-H_S / k_B T(E^*)\}. \qquad (24)$$

The main line of reasoning to prove (24) is analogous to (18)–(21), while the somewhat more tedious details will be presented elsewhere. In particular, (24) implies that the expectation value (22) of arbitrary subsystem observables A_S are indeed independent of any further details of $p(E)$.

Conclusions.—We have shown that the overwhelming majority of pure states yields practically identical expectation values for any given (not extremely large) set of observables under conditions which are generically satisfied for isolated macroscopic systems at thermal equilibrium. A second main result is that the experimentally uncontrollable and hence unknown microscopic details of the system state are indeed irrelevant. In particular, for the practically most important system-plus-bath setup, the canonical ensemble (24) is recovered under much more general and realistic assumptions than those implicit in the usual microcanonical description of the composite system.

The finding that (18) does not depend on the unknown details of the equilibrium ensemble ρ also sheds new light on the usual information theoretical "derivation" of the microcanonical ensemble [6]: While ρ_{mic} indeed mini-mizes the information functional $\mathrm{Tr}\rho \ln\rho$, the information content of many other ρ's is almost equally low, and one cannot conclude that practically only the exact minimum ρ_{mic} occurs in reality.

The present approach generalizes the seminal works [2,3] on canonical typicality in two respects: The system need not be a subsystem-plus-bath compound, and the equilibrium ensemble need not be of the quite particular microcanonical form.

Given a compound $S + B$ in a pure state $|\psi\rangle$, a well known consequence of entanglement between subsystem S and bath B is a mixed state ρ_S after tracing out the bath according to (23). While entanglement has been proposed as the main origin of canonical typicality in Ref. [3], our present findings suggest that the main root is the typicality property of the entire compound, which is in turn not entangled with any further system.

We close with a simple but quite interesting observation regarding systems *out of equilibrium*: Specifically, assume that the system is isolated and has reached equilibrium for times $t \leq 0$, while for $t > 0$, an external perturbation sets in [1], giving rise to an explicitly time dependent Hamiltonian $H(t)$ and a corresponding propagator $U(t_2, t_1)$. Instead of propagating the states $|\psi\rangle$ beyond $t = 0$, we can switch to the Heisenberg picture and instead propagate the observables. In this way, by replacing for any given $t > 0$ the original observable A by $U^\dagger(t, 0) A U(t, 0)$, all the equilibrium typicality properties at $t = 0$ immediately carry over to the out of equilibrium situation for $t > 0$. For not explicitly time dependent A, the spectrum remains invariant under time propagation, and hence, (13) remains valid for any given $t > 0$ with ρ being the equilibrium density operator at $t = 0$.

---

[1] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998); S. Mukamel, Phys. Rev. Lett. **90**, 170604 (2003); U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005); R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).

[2] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, Phys. Rev. Lett. **96**, 050403 (2006).

[3] S. Popescu, A. J. Short, and A. Winter, Nature Phys. **2**, 758 (2006).

[4] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics* (Springer, Berlin, 2004).

[5] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998); R. V. Jensen and R. Shankar, *ibid.* **54**, 1879 (1985); P. Bocchieri and A. Loinger, Phys. Rev. **114**, 948 (1959).

[6] B. Diu, C. Guthmann, D. Ledere, and B. Roulet, *Elements de Physique Statistique* (Hermann, Paris, 1996).

[7] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, J. Stat. Phys. **125**, 1197 (2006).