# Quantum versus classical foundation of statistical mechanics under experimentally realistic conditions

Peter Reimann and Mykhaylo Evstigneev
*Fakultät für Physik, Universität Bielefeld, 33615 Bielefeld, Germany*

(Received 18 December 2012; revised manuscript received 10 September 2013; published 11 November 2013)

Focusing on isolated macroscopic systems, described in terms of either a quantum mechanical or a classical model, our two key questions are how far does an initial ensemble (usually far from equilibrium and largely unknown in detail) evolve towards a stationary long-time behavior (equilibration) and how far is this steady state in agreement with the microcanonical ensemble as predicted by statistical mechanics (thermalization). A recently developed quantum mechanical treatment of the problem is briefly summarized, putting particular emphasis on the realistic modeling of experimental measurements and nonequilibrium initial conditions. Within this framework, equilibration can be proven under very weak assumptions about those measurements and initial conditions, while thermalization still requires quite strong additional hypotheses. An analogous approach within the framework of classical mechanics is developed and compared with the quantum case. In particular, the assumptions to guarantee classical equilibration are now rather strong, while thermalization then follows under relatively weak additional conditions.

DOI: 10.1103/PhysRevE.88.052114 PACS number(s): 05.20.−y, 05.30.−d, 05.45.−a

## I. INTRODUCTION

According to textbook statistical physics, all properties of a macroscopic system at thermal equilibrium are perfectly described by the canonical ensemble when weakly coupled to a thermal bath and by the microcanonical ensemble when isolated from the rest of the world. However, there still seems to be no truly satisfactory explanation of why this is so [1–3].

Here we adopt the most common viewpoint that investigations of this problem should be based on standard quantum or classical mechanics and should start with the treatment of isolated systems. Indeed, it is widely believed that by considering a system in contact with a thermal bath as a single isolated supersystem, the canonical formalism should be deducible from the microcanonical one. In turn, directly describing an open system alone (without the bath to which it is coupled) by means of standard quantum or classical mechanics does not seem possible.

While it is true that real systems are never perfectly isolated, it seems a quite reasonable working hypothesis (supported also by numerical simulations) that modeling them as perfectly isolated still provides a suitable description of the actual reality [1,4]. In the same vein, though classical models may possibly be inappropriate to describe *any* real macroscopic system [5], it remains interesting to understand why classical statistical mechanics agrees so well with many experiments and computer simulations of classical many-body systems.

Excellent reviews of the large number of pertinent classic (i.e., older) works and the ongoing debate about their physical implications are provided, e.g., by [1,2,6,7]. For more recent developments we refer to [8–10] and references therein. With our present work we further pursue a recently developed approach to the following two fundamental problems [11–17]: equilibration, i.e., how far a nonequilibrium initial ensemble evolves towards a stationary long-time behavior, and thermalization, i.e., to what extent this steady state is in agreement with the microcanonical ensemble predicted by equilibrium statistical mechanics.

The first key point of our approach is the modeling concept [2] that any given experimental system gives rise to a specific well-defined statistical ensemble (quantum mechanical density operator or classical probability density), the details of which are, however, unknown in practice. Typically, one only knows that in the initial state certain (usually macroscopic) observables are relatively sharply distributed about some approximately known mean values. A main challenge of the theory is to properly cope with this lack of knowledge. The second key point is not to modify or approximate in any way the exact time evolution according to standard quantum or classical mechanics. The third key point is to focus on experimentally realistic observables, exhibiting a finite range and a finite resolution.

Our present approach still covers essentially arbitrary (generic) systems and as such is complementary to many recent investigations of different specific quantum mechanical model classes, observables, and initial conditions, often with the main focus on the so-called eigenstate thermalization hypothesis and the role of quantum integrability (or nonintegrability) in this context (see, e.g., [8,9,18–39] and references therein). Indeed, since the basic laws of statistical physics are supposed to be of largely universal validity, also the foundation of such laws should not focus on any specific class of models, observables, or initial conditions.

In the first part of our paper (Secs. II–IV) we briefly summarize some previously obtained results in the quantum mechanical case from Refs. [11–17]. In the subsequent main part of the paper (Secs. V–VIII) the corresponding classical mechanical approach is outlined and compared with the quantum mechanical counterpart. Our present approach differs in the following important respects from some other well-known approaches, especially within the realm of ergodic theory [1,6,7].

First, it is not justified to consider only averages over sufficiently or even infinitely long time intervals in order to theoretically account for real measurements, as pointed out particularly clearly, e.g., by Sklar [1]: “If our macroscopic measurements could all be legitimately constructed as infinite time averages, then every macroscopic measurement would have to result in the equilibrium value for the quantity in

 

question. We could have no... ability to track the approach to equilibrium by following the variation of these quantities as they approached their final equilibrium values. But of course we can macroscopically determine the existence of [nonequilibrium] states." As a matter of fact, there can be little doubt that statistical mechanics is actually supposed to apply to arbitrarily short measurements, in agreement with all so far experimental and numerical experience. In our present approach, the main emphasis is on instantaneous rather than time-averaged measurements.

Second, it is not sufficient to focus on macroscopic observables, as done, e.g., in [4,40,41]. Quite the contrary, equilibrium statistical mechanics in fact also covers microscopic observables, e.g., the position and velocity, of one specific tracer particle within a fluid and its pertinent (probabilistic) predictions are in perfect agreement with numerical simulations and the rapidly increasing number of experiments on single molecules, nanoparticles, etc. In our present approach, arbitrary observables (with experimentally realistic range and resolution) are admitted.

Third, it is not sufficient to focus on an arbitrary but fixed energy hypersurface, as usually done in classical ergodic theory [6,7]. The reason is that the above-mentioned statistical ensemble that accounts for a given experimental system is largely unknown, but certainly cannot be realistically assumed to exhibit an infinitely sharp energy distribution. Our present approach admits a finite (and in detail unknown) spread of the energy distribution.

Fourth, the so-called ergodic hypothesis in its traditional form cannot be taken for granted, as done, e.g., in [6,7,42]. While this hypothesis has been commonly considered as not yet proven but probably true for a long time, the work by Markus and Meyer [43], Kolmogorov-Arnold-Moser (KAM) theory, and numerical evidence are making it increasingly clear that typical Hamiltonian systems exhibit a so-called divided phase space, consisting of a chaotic (ergodic) component of positive measure (chaotic sea) and a regular (nonergodic) component (union of the so-called KAM tori) of positive measure [1,44]. In our present approach, this fact is taken into account.

## II. QUANTUM MECHANICAL FRAMEWORK

We consider a large (macroscopic but finite) isolated system, modeled in terms of a Hilbert space $\mathcal{H}$ and a time-independent Hamiltonian $H : \mathcal{H} \to \mathcal{H}$ of the form

$$H = \sum_n E_n P_n, \tag{1}$$

where the $P_n$ are projectors onto the eigenspaces of $H$ with mutually different eigenvalues $E_n$ and multiplicities

$$\mu_n := \mathrm{Tr}\{P_n\}. \tag{2}$$

System states are described by density operators $\rho(t)$, evolving according to $\rho(t) = U_t \rho(0) U_t^\dagger$ with propagator $U_t := \exp\{-\mathrm{i}Ht\}$ and $\hbar = 1$. Exploiting (1) we thus can conclude that

$$\rho(t) = \sum_{mn} \rho_{mn}(0) \exp[-\mathrm{i}(E_m - E_n)t], \tag{3}$$

$$\rho_{mn}(t) := P_m \rho(t) P_n. \tag{4}$$

While we will be mainly interested in density operators describing some statistical ensemble (i.e., mixed states $\rho$ with $\mathrm{Tr}\{\rho^2\} < 1$), it is nevertheless worth pointing out that formally our considerations will also cover pure states ($\mathrm{Tr}\{\rho^2\} = 1$) as a special case.

Observables are represented by self-adjoint operators $A$ with expectation values $\mathrm{Tr}\{\rho(t)A\}$. In order to model real experimental measurements, however, it is not necessary to admit any arbitrary self-adjoint operator [40,45–51]. Rather, it is sufficient to focus on experimentally realistic observables in the following sense [11,14,52]: Any observable $A$ must represent an experimental device with a finite range of possible outcomes of a measurement

$$\Delta_A := \sup_{\|\psi\|=1} \langle\psi|A|\psi\rangle - \inf_{\|\psi\|=1} \langle\psi|A|\psi\rangle, \tag{5}$$

where the supremum and infimum are taken over all normalized vectors $|\psi\rangle \in \mathcal{H}$. Moreover, this working range $\Delta_A$ of the device must be limited to experimentally reasonable values compared to its resolution limit $\delta A$. All measurements known to the present authors yield less than 20 significant figures, i.e.,

$$\Delta_A / \delta A \leqslant 10^{20}. \tag{6}$$

Maybe some day 100 or 1000 significant figures will become feasible, but it seems reasonable that a theory that does not go very much beyond that will do. Note that similar restrictions also apply to numerical experiments by computer simulations.

The specific observable $A = P_n$ describes the population of the (possibly degenerate) energy level $E_n$ with expectation value (occupation probability)

$$p_n := \mathrm{Tr}\{P_n \rho(t)\}. \tag{7}$$

Note that $P_n$ commutes with $H$ from (1) and hence the level populations $p_n$ are conserved quantities.

For a system with $f$ degrees of freedom there are roughly $10^{O(f)}$ energy eigenstates with eigenvalues in every interval of 1 J beyond the ground state energy (see, e.g., [53] or Sec. 2.1 in [14]). The same estimate carries over to the number of energy eigenvalues under the assumption that their multiplicities (2) are much smaller than $10^{O(f)}$. For a macroscopic system with $f = O(10^{23})$, the energy levels are thus unimaginably dense on any decent energy scale and even the most careful experimentalist will not be able to populate only a few of them with significant probabilities $p_n$, in particular after averaging over the ensemble, i.e., over many repetitions of the experiment. In the generic case we thus expect [11,14,52] that, even if the system's energy is fixed up to an extremely small experimental uncertainty and even if the energy levels are populated extremely unequally, the largest ensemble-averaged level population $p_n$ will be extremely small (compared to $\sum_n p_n = 1$) and typically satisfy the rough estimate

$$\max_n p_n = 10^{-O(f)}. \tag{8}$$

## III. EQUILIBRATION IN THE QUANTUM CASE

Generically, the density operator $\rho(t)$ is not stationary right from the beginning, in particular for an initial condition $\rho(0)$ out of equilibrium. However, if the right-hand side of (3)

depends on $t$ initially, it cannot approach for large $t$ any time-independent equilibrium ensemble whatsoever. In fact, any mixed state $\rho(t)$ returns arbitrarily close (with respect to some suitable distance measure in Hilbert space) to its initial state $\rho(0)$ for certain, sufficiently large times $t$, as demonstrated, for instance, in Appendix D of Ref. [42]. We will therefore focus on the weaker notion of equilibration from [11–17], which merely requires the existence of a time-independent equilibrium state $\rho_{eq}$ (density operator) with the property that the difference

$$\sigma(t) := \mathrm{Tr}\{\rho(t)A\} - \mathrm{Tr}\{\rho_{eq}A\} \tag{9}$$

between the true expectation value $\mathrm{Tr}\{\rho(t)A\}$ and the equilibrium reference value $\mathrm{Tr}\{\rho_{eq}A\}$ is unresolvably small for the overwhelming majority of times $t$ contained in any sufficiently large (but finite) time interval $[0,T]$. In other words, the expectation values $\mathrm{Tr}\{\rho(t)A\}$ may still exhibit everlasting small fluctuations around their equilibrium value $\mathrm{Tr}\{\rho_{eq}A\}$, as well as very rare large excursions away from equilibrium (including the above-mentioned recurrences [2]), but quantitatively these fluctuations are either unobservably small compared to any reasonably achievable resolution limit or exceedingly rare on any realistic time scale after initial transients have died out. (Note that those initial transients become irrelevant if the time interval $[0,T]$ is chosen large enough.)

It seems quite plausible that if there is any such equilibrium ensemble $\rho_{eq}$ at all, then it should be given by the time-independent part of $\rho(t)$ from (3), i.e.,

$$\rho_{eq} := \sum_n \rho_{nn}. \tag{10}$$

Note that the time arguments of $\rho_{nn}$ have been omitted since these are conserved quantities according to (3) and (4).

Technically speaking (the reasons will become obvious below), the quantity of foremost interest is the time-averaged variance

$$\overline{\sigma^2(t)}^T := \frac{1}{T}\int_0^T dt\,\sigma^2(t) \tag{11}$$

following from (9) and (10). Considering and estimating such averages has a long tradition; see, e.g., [6,7,18,19,54–58]. Substantial progress along this line has been achieved quite recently in Refs. [11–17]. A particularly strong and general result has been obtained by Short and Farrelly [16] (for details see also [15,17]), showing that

$$\overline{\sigma^2(t)}^T \leqslant \frac{1}{2}\Delta_A^2 g \sum_n p_n^2 \tag{12}$$

for all sufficiently large $T$. Here $g$ denotes the maximal degeneracy of energy gaps

$$g := \max_{m\neq n} |\{(k,l) : E_k - E_l = E_m - E_n\}|, \tag{13}$$

where $|S|$ stands for the number of elements contained in the set $S$. In other words, $g$ is the maximal number of (exactly) coinciding energy differences among all possible pairs of distinct energy eigenvalues of the Hamiltonian $H$ from (1). In particular, Hamiltonians with degenerate energy eigenvalues and degenerate energy gaps are thus still admitted in (12). We further remark that the upper limit by Short and Farrelly [16] for the minimal $T$ value admitted in (12) depends on the spectrum of $H$ and on the level populations $p_n$, but not on the observable $A$. Finally, we note that the quite obvious relations

$$\sum_n p_n^2 \leqslant \max_n p_n \sum_n p_n = \max_n p_n, \tag{14}$$

$$\max_n p_n = \left(\max_n p_n^2\right)^{1/2} \leqslant \left(\sum_n p_n^2\right)^{1/2} \tag{15}$$

readily lead to the conclusion

$$\max_n p_n \text{ small} \Leftrightarrow \sum_n p_n^2 \text{ small}. \tag{16}$$

Hence, focusing on large effective dimensions [of the state $\rho(t)$] $d_{\mathrm{eff}} := 1/\sum_n p_n^2$ as done, e.g., in [12,13,15,16] is equivalent to our present assumption of small level populations [cf. (8)].

Next we define for any given $\delta A > 0$ and $T > 0$ the quantity

$$T_{\delta A} := |\{0 < t < T : |\mathrm{Tr}\{\rho(t)A\} - \mathrm{Tr}\{\rho_{eq}A\}| \geqslant \delta A\}|, \tag{17}$$

where $|M|$ denotes the size (Lebesgue measure) of the set $M$. According to (9), $T_{\delta A}$ thus represents the measure of all times $t \in [0,T]$ for which $|\sigma(t)| \geqslant \delta A$ holds true. It follows that $\sigma^2(t) \geqslant \delta A^2$ for a set of times $t$ of measure $T_{\delta A}$ and that $\sigma^2(t) \geqslant 0$ for all remaining times $t$ in $[0,T]$. Hence the temporal average (11) must be at least $\delta A^2 T_{\delta A}/T$ and we can conclude from (12) and (14) that

$$\frac{T_{\delta A}}{T} \leqslant \frac{1}{2}\left(\frac{\Delta_A}{\delta A}\right)^2 g \max_n p_n \tag{18}$$

for all sufficiently large $T$.

We remark that the original derivation of (12) from [16] and thus the result (18) is restricted to finite-dimensional Hilbert spaces $\mathcal{H}$. The generalization to infinite-dimensional systems has been worked out in [17], yielding instead of (18) the slightly weaker bound

$$\frac{T_{\delta A}}{T} \leqslant 2\left(\frac{\Delta_A}{\delta A}\right)^2 g \max_n p_n. \tag{19}$$

According to (17), the left-hand side of (19) represents the fraction of all times $t \in [0,T]$ for which there is an experimentally resolvable difference between the true expectation value $\mathrm{Tr}\{\rho(t)A\}$ and the time-independent equilibrium expectation value $\mathrm{Tr}\{\rho_{eq}A\}$. On the right-hand side of (19), $\Delta_A/\delta A$ is the range-to-resolution ratio, which can be considered as bounded according to (6) for all experimentally realistic measurements $A$. The next factor $g$ is the maximal degeneracy of energy gaps from (13). Finally, $\max_n p_n$ represents the largest ensemble-averaged occupation probability of the (possibly degenerate) energy eigenvalues $E_n$ [see (7)]. Typically, one expects that the rough upper bound (8) applies, except if certain energy eigenvalues are so extremely highly degenerate that the multiplicities defined in (2) severely reduce the pertinent energy level density compared to the nondegenerate case [see above (8)]. For a system with sufficiently many degrees of freedom $f$ and no exceedingly large degeneracies of energy eigenvalues and energy gaps [59] we thus can conclude from (19) with (6) and (8) that the system behaves in every possible experimental measurement exactly as if it were in

the equilibrium state $\rho_{eq}$ for the overwhelming majority of times within any sufficiently large (but finite) time interval $[0,T]$ [61], i.e., we recover equilibration in the sense proposed at the beginning of this section.

## IV. THERMALIZATION IN THE QUANTUM CASE

Next we address the somewhat related but still notably different issue of thermalization, i.e., the question of whether, and to what extent, the above-discussed equilibrium expectation value $\mathrm{Tr}\{\rho_{eq}A\}$ is in agreement with that predicted by the microcanonical ensemble. For the sake of simplicity, we confine ourselves to nondegenerate Hamiltonians throughout this section, i.e., all projectors $P_n$ in (1) are of the form $P_n = |n\rangle\langle n|$, where $|n\rangle$ denotes the energy eigenvector corresponding to the (nondegenerate) energy eigenvalue $E_n$. Furthermore, we assume without loss of generality that the eigenstates $|n\rangle$ are ordered according to their eigenvalues, i.e., $E_n \leqslant E_{n+1}$ for all $n$ [62]. According to (4), (7), and (10) it follows that

$$\mathrm{Tr}\{\rho_{eq}A\} = \sum_n p_n \langle n|A|n\rangle, \tag{20}$$

while the expectation value predicted by the microcanonical ensemble takes the form

$$\mathrm{Tr}\{\rho_{\mathrm{mic}}A\} = \sum_n p_n^{\mathrm{mic}} \langle n|A|n\rangle, \tag{21}$$

where the level populations $p_n^{\mathrm{mic}}$ are equal to a normalization constant if $E_n$ is contained within some small energy interval

$$I := [E_{\mathrm{mic}} - \Delta E, E_{\mathrm{mic}}] \tag{22}$$

and zero otherwise [53].

If (20) and (21) yield measurable differences for experimentally realistic $\rho(0)$ and $A$, the purely quantum mechanical prediction (20) is commonly considered as more fundamental [21,24,26]. In other words, provided equilibrium statistical mechanics itself is valid at all, its prediction (21) must agree with (20), at least within experimentally realistic resolution limits.

What are these validity conditions, beyond which the microcanonical formalism of equilibrium statistical mechanics may break down? The first well-known validity condition for the microcanonical formalism is, as said below (21), that only $E_n$ within some small energy interval (22) have a nonvanishing occupation probability. More generally, in equilibrium statistical mechanics it is taken for granted that the system energy is fixed up to unavoidable experimental uncertainties. In contrast, realistic initial conditions as discussed above Eq. (8) require that this energy uncertainty is much larger than $10^{-O(f)}$ J, which is obviously always fulfilled in practice, but we never introduced or exploited any type of upper limit for this uncertainty so far, i.e., the energy uncertainty may still be arbitrarily large in (19) and (20). In other words, for large energy uncertainties, our key relation (20) remains valid, while equilibrium statistical mechanics is likely to become invalid. This is clearly a not at all surprising case of disagreement between (20) and (21). To avoid such almost trivial cases, we henceforth take for granted that the system energy is known up to an uncertainty $\Delta E$ that is as small as possible, but still experimentally realistic.

The second (often tacit) validity condition of the microcanonical formalism is that the expectation values (21) are required (or assumed) to be (practically) independent of the exact choice of the interval $I$ in (22), i.e., of its upper limit $E_{\mathrm{mic}}$ and its width $\Delta E$. Essentially this means though that the value of the sum in Eq. (20) must be largely independent of the details of the weights $p_n$. The same conclusion also follows from the equivalence of the microcanonical and canonical ensembles (for all energies $E_{\mathrm{mic}}$), considered as a self-consistency condition for equilibrium statistical mechanics [49,50].

If the sum in Eq. (20) has the above-mentioned property of being largely independent of the weights $p_n$, we henceforth abbreviate this fact by saying that it has the property $P$. [A more precise definition of "largely independent" is difficult: Essentially, the value of the sum in (20) is supposed to exhibit negligible variations for all sets of weights $\{p_n\}$ arising under physically relevant circumstances.] Given property $P$ holds, the expectation values (20) and (21) are indeed practically indistinguishable.

Our first remark regarding property $P$ itself is that no experimentalist can control the populations $p_n$ of the unimaginably dense energy levels $E_n$, apart from the very gross fact that they are mainly concentrated within the interval $I$ from (22). Thus, if the details of the $p_n$ in (20) were to matter, then not only would equilibrium statistical mechanics break down, but also reproducing measurements, in particular in different laboratories, would be largely impossible.

The simplest way to guarantee property $P$ seems to require (or assume) that the expectation values $\langle n|A|n\rangle$ hardly vary within any small energy interval of the form (22). This is indeed part of a common conjecture about the semiclassical behavior of fully chaotic classical systems (see [19,58,63] and references therein). In particular, negligible variations of $\langle n|A|n\rangle$ for close by $n$ values imply Srednicki's eigenstate thermalization hypothesis [64] (anticipated in [65] and revisited in Ref. [26]), implying that each individual energy eigenstate $|n\rangle$ behaves like the equilibrium ensemble.

An alternative way to guarantee property $P$ follows from the argument by Peres [56] that even if the $\langle n|A|n\rangle$ may notably vary with $n$, the immense number of relevant summands in (20) may, for typical $A$ and $\rho(0)$, lead to a kind of statistical averaging effect and thus a largely $\rho(0)$-independent overall value of the sum. Numerically, the validity and possible failure of such conjectures and of property $P$ itself have been exemplified, e.g., in [21–23,26,28–34,36,37,39,65,66]. While the details, in particular the role of more basic system properties such as quantum ergodicity and integrability (or nonintegrability), are still not very well understood [6,7,19,25–27,56,58,63,67], equilibration in agreement with (20) was seen numerically in all known cases.

## V. CLASSICAL MECHANICAL FRAMEWORK

As in Sec. II, we consider an isolated system with $f \gg 1$ degrees of freedom, but now modeled classically in terms a $2f$-dimensional phase space $\Gamma$ and a time-independent Hamiltonian $H : \Gamma \to \mathbb{R}$. Given an initial condition $\phi_0 \in \Gamma$ at time $t_0 = 0$, the state at time $t$ can be written in terms of the propagator $\gamma_t$ induced by $H$ as $\phi(t) = \gamma_t(\phi_0)$. The fact that in reality the actual microstate $\phi(t)$ is not exactly

known is modeled by means of a probability density $\rho(\phi,t)$, being non-negative, Lebesgue integrable, and normalized on $\Gamma$ [2]. Further, $\rho(\phi,t)$ evolves from the initial ensemble $\rho_0(\phi) := \rho(\phi,0)$ according to

$$\rho(\phi,t) = \rho_0(\gamma_{-t}(\phi)). \tag{23}$$

In particular, pure states of the form $\rho(\phi,t) = \delta(\phi - \gamma_t(\phi_0))$ are thus excluded. The main place where this assumption turns out to be indispensable will be in Eq. (48) below.

The resulting expectation value for an observable $A(\phi)$ is given by

$$\langle A\rangle_t := \int d\phi\, \rho(\phi,t)A(\phi), \tag{24}$$

where the integral extends over the entire phase space $\Gamma$ and where we tacitly focus on Lebesgue integrable $A(\phi)$. From Liouville's theorem and $H(\gamma_t(\phi)) = H(\phi)$ (energy conservation) we can infer that

$$\int d\phi\, a(\phi)b(H(\phi)) = \int d\phi\, a(\gamma_t(\phi))b(H(\phi)) \tag{25}$$

for arbitrary functions $a : \Gamma \to \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$, and times $t \in \mathbb{R}$. With (23) we thus can rewrite (24) as

$$\langle A\rangle_t := \int d\phi\, \rho_0(\phi)A(\gamma_t(\phi)). \tag{26}$$

As in (5), we focus on observables $A(\phi)$ with a finite range

$$\Delta_A := \sup_\phi A(\phi) - \inf_\phi A(\phi), \tag{27}$$

where the supremum and infimum are taken over all $\phi \in \Gamma$. Moreover, the ratio between this range and the resolution limit $\delta A$ is again assumed to satisfy (6).

Before turning to the discussion of the admissible initial conditions $\rho_0(\phi)$, in particular the classical counterpart of (8), some very important but somewhat technical preliminaries are required. We start by recalling Birkhoff's theorem, asserting [6,7] that for any (Lebesgue integrable) $A(\phi)$ and almost all $\phi \in \Gamma$ the time average

$$\overline{A}(\phi) := \lim_{T\to\infty} \int_0^T \frac{dt}{T} A(\gamma_t(\phi)) \tag{28}$$

exists and is a constant of motion, i.e., $\overline{A}(\gamma_t(\phi)) = \overline{A}(\phi)$ for all $t$. Furthermore, based on numerical evidence, heuristic reasoning, and rigorous arguments in special cases, it is commonly conjectured [1,4,44,68,69] that a generic Hamiltonian system (with $f > 1$) exhibits a so-called divided phase space. More precisely, the total phase space $\Gamma$ can be decomposed into a set $\Gamma_{KAM} \subset \Gamma$, being the union of all the so-called Kolmogorov-Arnold-Moser tori, and a complement

$$\Gamma_{\rm sea} := \Gamma \setminus \Gamma_{KAM}, \tag{29}$$

essentially consisting of the so-called chaotic sea (or stochastic region), but for the sake of convenience will henceforth be understood to also contain a nonchaotic set of measure zero (e.g., so-called hyperbolic fixed points and periodic orbits). Their main properties are as follows [1,4,44,68,69].

(i) The dynamics leaves $\Gamma_{\rm sea}$ invariant, i.e., if $\phi \in \Gamma_{\rm sea}$ then $\gamma_t(\phi) \in \Gamma_{\rm sea}$ for all $t$. As a consequence, a similar invariance property applies to $\Gamma_{KAM}$.

(ii) The dynamics has the following so-called ergodicity property [70,71]: For almost all phase points $\phi \in \Gamma_{\rm sea}$, which furthermore belong to an arbitrary but fixed energy surface [i.e., $H(\phi) = E$ with an arbitrary but fixed $E$], the time average in (28) is given by one and the same value. This value, however, will in general be different for different energies $E$. In other words, for any $A(\phi)$ there exists a function $A^* : \mathbb{R} \to \mathbb{R}$ with the property that

$$\overline{A}(\phi) = A^*(H(\phi)) \quad \text{for almost all } \phi \in \Gamma_{\rm sea}. \tag{30}$$

(iii) The dynamics has the following so-called (strong) mixing property [70,71]: Almost all phase points $\phi,\phi' \in \Gamma_{\rm sea}$ with $H(\phi) = H(\phi')$ (i.e., $\phi$ and $\phi'$ belong to the same energy surface) exhibit an independent evolution over long times in the sense [72]

$$\lim_{T\to\infty} \int_0^T \frac{dt}{T} A(\gamma_t(\phi))A(\gamma_t(\phi')) = \overline{A}(\phi)\overline{A}(\phi')$$
$$\text{for almost all } \phi,\phi' \in \Gamma_{\rm sea} \quad \text{with} \quad H(\phi) = H(\phi'). \tag{31}$$

In particular, if $\phi'$ happens to coincide with $\phi$ then the long time evolution of $\phi'$ is obviously no longer independent from that of $\phi$ and thus the above relation generically breaks down.

(iv) The set $\Gamma_{KAM}$ of all KAM tori exhibits an extremely convoluted, nowhere dense, self-similar structure. For instance, let us focus on phase points up to some arbitrary but fixed energy $E$ and accordingly define the restricted sets $\Gamma(E) := \{\phi \in \Gamma : H(\phi) \leqslant E\}$, $\Gamma_{KAM}(E) := \Gamma(E) \cap \Gamma_{KAM}$, and $\Gamma_{\rm sea}(E) := \Gamma(E) \cap \Gamma_{\rm sea}$. Then, for all sufficiently large but finite $E$, all three sets $\Gamma(E)$, $\Gamma_{KAM}(E)$, and $\Gamma_{\rm sea}(E)$ are of positive but finite measure (see, e.g., Ref. [44] and p. 175 of Ref. [1]). Moreover, it is commonly expected [1,44,69,73,74] that for systems with sufficiently many degrees of freedom $f$ the ratios $\Gamma_{KAM}(E)/\Gamma(E)$ become arbitrarily small. Accordingly, it seems plausible that even the most careful experimentalist will not be able to populate the set $\Gamma_{KAM}$ of all KAM tori with significant probability on the average over many repetitions of the experiment. In other words, we expect that the initial ensemble $\rho_0(\phi)$ has the property that

$$\int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi) \quad \begin{array}{l}\text{becomes arbitrarily small}\\ \text{for sufficiently large } f,\end{array} \tag{32}$$

where $\int_{\Gamma_{KAM}}$ indicates an integration over the subset $\Gamma_{KAM}$ of the total phase space $\Gamma$ and the word "small" is meant in comparison with $\int d\phi\, \rho_0(\phi) = 1$.

This property (32) represents the announced classical counterpart of (8). In fact, according to the most common conjecture [69,73], an exponential decay with $f$ analogous to (8) may be expected also in (32). Explicit estimates are provided, e.g., at the end of Sec. 6.5a in Ref. [69] and in the paper by Falcioni *et al.*, see Ref. [73(b)], suggesting an exponential decay analogous to (8).

We remark that it is always possible to find sufficiently small hyperspheres (or hypercubes, etc.) $\Gamma' \subset \Gamma$ with the property that the measure of $\Gamma_{KAM} \cap \Gamma'$ approaches that of $\Gamma'$ arbitrarily closely [4]. Condition (32) expresses the expectation that the experimentalist cannot hit those tiny subsets with any notable probability, analogous to the unlikeliness of hitting any single energy eigenvalue expressed by (8), and in spite of

the fact that initial ensembles $\rho_0(\phi)$ far from equilibrium are usually concentrated themselves within very small phase space regions. In contrast, not hitting the subset $\Gamma_{KAM}$ at all, so that the integral (32) becomes strictly zero, also seems practically impossible.

## VI. EQUILIBRATION IN THE CLASSICAL CASE

Proceeding like in Sec. III, the quantity of foremost interest is the deviation

$$\sigma(t) := \langle A\rangle_t - \langle A\rangle_{eq} \tag{33}$$

of the true expectation value [cf. (24)] from the time-independent expectation value [75]

$$\langle A\rangle_{eq} := \int d\phi\, \rho_{eq}(\phi)A(\phi) \tag{34}$$

for a suitably defined equilibrium ensemble $\rho_{eq}(\phi)$. As one might have expected, the proper choice of the latter ensemble turns out to be the time-averaged true ensemble $\rho(\phi,t)$, i.e.,

$$\rho_{eq}(\phi) := \lim_{T\to\infty}\int_0^T \frac{dt}{T}\rho(\phi,t). \tag{35}$$

According to (23) and Birkhoff's theorem (28), the limit in (35) indeed exists [for almost all $\phi$; the remaining exceptional $\phi$ are irrelevant in (34)].

Exploiting (35), the equilibrium expectation value from (34) can be rewritten by means of (24), (26), and (28) as

$$\langle A\rangle_{eq} = \lim_{T\to\infty}\int_0^T \frac{dt}{T}\langle A\rangle_t = \int d\phi\, \rho_0(\phi)\overline{A}(\phi). \tag{36}$$

We further can conclude from (26), (33), and (36) that

$$\begin{aligned}\sigma(t) &= \int d\phi\, \rho_0(\phi)[A(\gamma_t(\phi)) - \overline{A}(\phi)]\\ &= \int dE\, p(E)D(E,t),\end{aligned} \tag{37}$$

$$D(E,t) := \int d\phi\, \rho(E,\phi)[A(\gamma_t(\phi)) - \overline{A}(\phi)], \tag{38}$$

$$p(E) := \int d\phi\, \delta(H(\phi) - E)\rho_0(\phi), \tag{39}$$

$$\rho(E,\phi) := \frac{\delta(H(\phi) - E)\rho_0(\phi)}{p(E)}, \tag{40}$$

where the integral in (37) and the definition (40) are tacitly restricted to $E$ values with $p(E) > 0$ [76]. Noting that $p(E)$ from (39) is non-negative we can infer from (37) that

$$\begin{aligned}\sigma^2(t) &\leqslant \left[\int dE\, p(E)|D(E,t)|\right]^2\\ &= \left[\int dE\sqrt{p(E)}\sqrt{q(E)}\right]^2,\end{aligned} \tag{41}$$

$$q(E) := p(E)D^2(E,t). \tag{42}$$

Applying the Cauchy-Schwarz inequality, it follows that

$$\sigma^2(t) \leqslant \int dE' p(E') \int dE\, q(E). \tag{43}$$

According to (39), the first integral is unity. With (42) and by indicating time averages by an overbar as in (11), it follows that

$$\overline{\sigma^2(t)}^T \leqslant \int dE\, p(E)\overline{D^2(E,t)}^T. \tag{44}$$

In view of (38), the last factor takes the form

$$\overline{D^2(E,t)}^T = \int d\phi \int d\phi' \rho(E,\phi)\rho(E,\phi')C(T,\phi,\phi'), \tag{45}$$

where we have introduced

$$\begin{aligned}C(T,\phi,\phi') &:= \overline{[A(\gamma_t(\phi)) - \overline{A}(\phi)][A(\gamma_t(\phi')) - \overline{A}(\phi')]}^T\\ &= \overline{A(\gamma_t(\phi))A(\gamma_t(\phi'))}^T - \overline{A(\gamma_t(\phi))}^T\overline{A}(\phi')\\ &\quad - \overline{A}(\phi)\overline{A(\gamma_t(\phi'))}^T + \overline{A}(\phi)\overline{A}(\phi').\end{aligned} \tag{46}$$

In the last two lines, the second and third terms have well-defined limits $T\to\infty$ for almost all $\phi$, $\phi'$ according to Birkhoff's theorem [see (28)]. To establish that the same applies for the first term, we consider an auxiliary dynamics on the product space $\Gamma^2 := \Gamma\times\Gamma$, governed by the Hamiltonian $H_{\rm tot}(\phi_{\rm tot}) := H(\phi) + H(\phi')$ with $\phi_{\rm tot} := (\phi,\phi') \in \Gamma^2$. Since this gives rise to a well-defined Hamiltonian dynamics, the existence of the limit in question follows by applying Birkhoff's theorem to this dynamics. Under the tacit assumption that the limit $T\to\infty$ commutes with the integrations in (44) and (45), the right-hand side in (44) thus exhibits a well-defined limit $T\to\infty$ [78]. Denoting this limit by $\overline{\sigma^2(t)}^\infty$, it follows for any given $\epsilon > 0$ that $|\overline{\sigma^2(t)}^T - \overline{\sigma^2(t)}^\infty| \leqslant \epsilon$ for all sufficiently large $T$. Hence

$$\overline{\sigma^2(t)}^T \leqslant \epsilon + \int dE\, p(E)\overline{D^2(E,t)}^\infty \tag{47}$$

for all sufficiently large $T$ [79].

Next, the integration domain $\Gamma\times\Gamma$ of the double integral in (45) is decomposed by exploiting that $\Gamma = \Gamma_{\rm sea}\cup\Gamma_{KAM}$ [see (29)] into the four subdomains $\Gamma_{\rm sea}\times\Gamma_{\rm sea}$, $\Gamma_{\rm sea}\times\Gamma_{KAM}$, $\Gamma_{KAM}\times\Gamma_{\rm sea}$, and $\Gamma_{KAM}\times\Gamma_{KAM}$. For the sake of later convenience, we reunify the second and the fourth of them into $\Gamma\times\Gamma_{KAM}$. Hence we are left with the three subdomains $\Gamma_{\rm sea}\times\Gamma_{\rm sea}$, $\Gamma\times\Gamma_{KAM}$, and $\Gamma_{KAM}\times\Gamma_{\rm sea}$. Noting that in (45) we have $H(\phi) = H(\phi') = E$ according to (40), it follows with (28), (31), and (46) that the contribution from the first subdomain $\Gamma_{\rm sea}\times\Gamma_{\rm sea}$ to (45) approaches zero in the limit $T\to\infty$,

$$\lim_{T\to\infty}\int_{\Gamma_{\rm sea}} d\phi \int_{\Gamma_{\rm sea}} d\phi' \rho(E,\phi)\rho(E,\phi')C(T,\phi,\phi') = 0. \tag{48}$$

We remark that in the case of a pure state of the form $\rho(\phi,t) = \delta(\phi - \gamma_t(\phi_0))$ the right-hand side of (45) becomes equal to $C(T,\phi_0,\phi_0)$, approaching $\overline{A^2(\gamma_t(\phi_0))}^\infty - [\overline{A}(\phi_0)]^2$ for $T\to\infty$ according to (46). As pointed out below (31), this difference is nonzero for any generic $\phi_0\in\Gamma_{\rm sea}$ and hence the same applies to the left-hand side of (48).

Next, denoting by $K$ the modulus of the contribution to (45) from the second subdomain $\Gamma\times\Gamma_{KAM}$ [see above (48)], it follows that

$$K \leqslant \int_\Gamma d\phi \int_{\Gamma_{KAM}} d\phi' \rho(E,\phi)\rho(E,\phi')|C(T,\phi,\phi')|. \tag{49}$$

Further, $|C(T,\phi,\phi')|$ is bounded from above by $\Delta_A^2$ according to (27) and (46), yielding

$$K \leqslant \Delta_A^2 \int_\Gamma d\phi\, \rho(E,\phi) \int_{\Gamma_{KAM}} d\phi' \rho(E,\phi'). \tag{50}$$

With (39) and (40) one readily sees that the first integral on the right-hand side of (50) is unity. Exactly the same upper bound as in (50) is readily recovered for the modulus of the contribution from the third subdomain $\Gamma_{KAM} \times \Gamma_{\rm sea}$ to (45). Altogether we thus obtain the upper bound

$$\int dE\, p(E) \overline{D^2(E,t)}^\infty \leqslant 2\Delta_A^2 \int dE\, p(E) \int_{\Gamma_{KAM}} d\phi\, \rho(E,\phi). \tag{51}$$

Taking into account (40) and (47), we finally can conclude that for any given $\epsilon > 0$

$$\overline{\sigma^2(t)}^T \leqslant \epsilon + 2\Delta_A^2 \int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi) \tag{52}$$

for all sufficiently large $T$.

Next we define for any given $\delta A > 0$ and $T > 0$ the classical counterpart of (17), namely,

$$T_{\delta A} := |\{0 < t < T : |\langle A\rangle_t - \langle A\rangle_{eq}| \geqslant \delta A\}|, \tag{53}$$

where $|M|$ denotes the size (Lebesgue measure) of the set $M$ and $\langle A\rangle_t$ and $\langle A\rangle_{eq}$ are defined by (26) and (34), respectively. Introducing $\epsilon' := \epsilon/\delta A^2$, it follows [as below (18)] that for any given $\epsilon' > 0$

$$\frac{T_{\delta A}}{T} \leqslant \epsilon' + 2\left(\frac{\Delta_A}{\delta A}\right)^2 \int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi) \tag{54}$$

for all sufficiently large $T$.

According to (53), the left-hand side of (54) represents the fraction of all times $t \in [0,T]$ for which there is an experimentally resolvable difference between the true expectation value from (26) and the time-independent equilibrium expectation value from (34). In (54) the range-to-resolution ratio of $\Delta_A/\delta A$ is bounded by (6), while $\int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi)$ is the initial population of the KAM tori, which is expected to become arbitrarily small for a system with sufficiently many degrees of freedom $f$ [see (32)]. In the generic case, this population $\int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi)$ is furthermore expected to be nonzero [see below (32)]. Hence we may choose $\epsilon'$ equal to the last term in (54) to arrive at the classical counterpart of (19), namely,

$$\frac{T_{\delta A}}{T} \leqslant 4\left(\frac{\Delta_A}{\delta A}\right)^2 \int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi) \tag{55}$$

for all sufficiently large $T$. In either case, it follows with (53) [analogously to (19)] that a system with sufficiently many degrees of freedom behaves as if it were in the equilibrium state $\rho_{eq}(\phi)$ for the overwhelming majority of times after initial transients have died out.

### A. Generalizations

In order to obtain (48) we have exploited the mixing condition (31) (see also [2]). However, the same relation (48) can readily be recovered under the alternative assumption that

$$\lim_{T\to\infty} \overline{\left[\int_{\Gamma_{\rm sea}} d\phi\, \rho(E,\phi)\{A(\gamma_t(\phi)) - \overline{A}(\phi)\}\right]^2}^T = 0 \tag{56}$$

for all $\rho(E,\phi)$ of the form (40) so that the phase space integral in (56) is effectively confined to an energy surface with an arbitrary but fixed $E$ (see also [76]). Property (56) follows [80] from the so-called weak mixing condition [71]

$$\lim_{T\to\infty} \overline{\left|\int_{\Gamma_{\rm sea}} d\phi\, \rho(E,\phi)\{A(\gamma_t(\phi)) - \overline{A}(\phi)\}\right|}^T = 0, \tag{57}$$

which in turn quite obviously follows from the mixing condition (31) (which therefore is sometimes also referred to as strong mixing [71]).

In summary, we thus found that either strong mixing (31) or weak mixing (57) [and taking for granted (6) and (32)] is sufficient for equilibration in the sense outlined at the beginning of Sec. III. However, it seems quite possible that a stronger form of equilibration, namely the asymptotic convergence of $\langle A\rangle_t$ for *all* sufficiently large $t$, could follow under much weaker conditions [40–42,74,81]. First, in view of (24) and (35) it is clear that if such a convergence does occur at all, then the limit will be given by (34). Second, one readily sees that for any (confining) system with a single degree of freedom ($f = 1$), such a convergence indeed occurs in the generic case, i.e., for nonsingular $\rho_0(\phi)$ [especially the energy distribution $p(E)$ from (39) must be nonsingular] and provided the frequency of the (analytically solvable) periodic motion depends in a nontrivial way on the energy [42]. The same applies for a system consisting of several noninteracting subsystems of this kind (see Chap. 3.4 in [42]). There seems no *a priori* reason that when the subsystems start to interact with each other, the long time limit of $\langle A\rangle_t$ should cease to exist. In other words, the convergence of (37) to zero for $T \to \infty$ may well be true under much weaker conditions than the corresponding convergence on every single energy surface, as required by (56) or by the weak mixing condition (57).

## VII. THERMALIZATION IN THE CLASSICAL CASE

As in Sec. IV, the problem of thermalization amounts, in view of the conclusion below (55), to showing that the equilibrium expectation value $\langle A\rangle_{eq}$ agrees, within the experimental resolution limit, with the microcanonical expectation value

$$\langle A\rangle_{\rm mic} = \int d\phi\, \rho_{\rm mic}(\phi) A(\phi), \tag{58}$$

provided the energy distribution $p(E)$ from (39) is peaked about its mean value $E_{\rm mean}$ with a small but still experimentally realistic dispersion $\delta E$. The microcanonical ensemble $\rho_{\rm mic}(\phi)$ in (58) is, as usual [53], assumed to be equal to some normalization constant if $H(\phi)$ is contained within the small energy interval $I$ from (22) and zero otherwise.

In particular, all the unknown details (see Sec. I) of the usually out of equilibrium initial condition $\rho_0(\phi)$ are postulated to be irrelevant in (39) and likewise for the exact values of $E_{\rm mic}$ and $\Delta E$ in (22) (see also the discussion of property $P$ in Sec. IV). For the sake of simplicity the latter two quantities

$E_{\rm mic}$ and $\Delta E$ are thus identified with their experimental counterparts $E_{\rm mean}$ and $\delta E$ from now on.

Intuitively, it is quite clear that the energy distribution from (39) remains unchanged as the system evolves in time. Formally, this is readily confirmed by means of (25), yielding, with (35),

$$p(E) = \int d\phi\, \delta(H(\phi) - E)\rho(\phi,t) = \int d\phi\, \delta(H(\phi) - E)\rho_{eq}(\phi) \tag{59}$$

for arbitrary times $t$.

Next, we introduce yet another time-independent probability density $\rho^*_{eq}(\phi)$, which attributes the same probability (59) as the true equilibrium ensemble $\rho_{eq}(\phi)$ to any given energy surface and is moreover constant on any such surface,

$$\rho^*_{eq}(\phi) := p(H(\phi))/\omega(H(\phi)), \tag{60}$$

where the normalization $\omega(E)$ denotes the area of the energy surface

$$\omega(E) := \int d\phi\, \delta(H(\phi) - E) \tag{61}$$

and (60) is replaced by $\rho^*_{eq}(\phi) := 0$ in the case $\omega(H(\phi)) = 0$. In other words, $\rho^*_{eq}(\phi)$ most closely resemble the true $\rho_0(t)$ among all time-independent densities that are constant on every energy surface. The corresponding expectation values are denoted by [cf. Eqs. (24) and (34)]

$$\langle A\rangle^*_{eq} := \int d\phi\, A(\phi)\rho^*_{eq}(\phi). \tag{62}$$

Our first main step in demonstrating thermalization [see above (58)] is based on the remarkable, completely general identity

$$\langle A\rangle_{eq} - \langle A\rangle^*_{eq} = \int d\phi[\overline{A}(\phi) - A^*(H(\phi))][\rho_0(\phi) - \rho^*_{eq}(\phi)], \tag{63}$$

where $A^*(H(\phi))$ is defined in (30) and whose derivation is provided in the Appendix. Due to (30), only $\phi \in \Gamma_{KAM}$ contribute to the integral in (63), i.e.,

$$\langle A\rangle_{eq} - \langle A\rangle^*_{eq} = \int_{\Gamma_{KAM}} d\phi[\overline{A}(\phi) - A^*(H(\phi))][\rho_0(\phi) - \rho^*_{eq}(\phi)]. \tag{64}$$

Taking into account $\rho_0(\phi) \geqslant 0$, $\rho^*_{eq}(\phi) \geqslant 0$, and Eq. (27), we thus obtain

$$|\langle A\rangle_{eq} - \langle A\rangle^*_{eq}| \leqslant \Delta_A\left[\int_{\Gamma_{KAM}} d\phi\, \rho_0(\phi) + \int_{\Gamma_{KAM}} d\phi\, \rho^*_{eq}(\phi)\right]. \tag{65}$$

The first integral becomes arbitrarily small for sufficiently large systems according to (32). The second integral is of the same type, except that the original density $\rho_0(\phi)$ is now uniformly redistributed within every energy surface [see below (61)]. Quite clearly, such a particularly tame density $\rho^*_{eq}(\phi)$ instead of the true (possibly far from equilibrium) initial density $\rho_0(\phi)$ is actually also covered by (32) as a special case.

In view of (6) we thus can conclude from (65) that the difference between $\langle A\rangle_{eq}$ and $\langle A\rangle^*_{eq}$ becomes unmeasurably small for sufficiently many degrees of freedom. Together with our previous result on equilibration [see below (54)] it follows that the system behaves as if it were in the equilibrium state $\rho^*_{eq}(\phi)$ for the overwhelming majority of times $t$ after initial transients have died out.

Finally, we introduce the average of an observable over an arbitrary but fixed energy surface [tacitly confining ourselves to $E$ values with $\omega(E) > 0$]

$$A_\omega(E) := \int d\phi\, A(\phi)\delta(H(\phi) - E)/\omega(E) \tag{66}$$

and, analogously to (39) and (59), the microcanonical energy distribution

$$p_{\rm mic}(E) := \int d\phi\, \delta(H(\phi) - E)\rho_{\rm mic}(\phi). \tag{67}$$

Combining these definitions with (58)–(60) and (62), it follows that

$$\langle A\rangle^*_{eq} = \int dE\, p(E)A_\omega(E), \tag{68}$$

$$\langle A\rangle_{\rm mic} = \int dE\, p_{\rm mic}(E)A_\omega(E). \tag{69}$$

These are the classical counterparts of Eqs. (20) and (21). Accordingly, the discussion below (21) can be readily carried over to the present case.

In particular, both $p(E)$ and $p_{\rm mic}(E)$ are assumed, as said below (58), to take only appreciable values within a small energy interval of width $\delta E$. Furthermore, the classical counterpart of the eigenstate thermalization hypothesis from Sec. IV now amounts to the condition that the variations of $A_\omega(E)$ from (66) must remain negligibly small (below the experimental resolution limit $\delta A$) when changing $E$ by less than the experimental energy resolution $\delta E$. Intuitively, the quite general validity of this condition seems considerably more plausible than its quantum mechanical counterpart, essentially amounting to some form of uniform continuity of $A(\phi)$ along directions perpendicular to the energy surfaces.

We also remark that $\delta E$ is meant here as the remnant uncertainty after the experimentalist has determined the energy of his system as accurately as possible. A violation of the above condition on $A_\omega(E)$ would imply that with the help of the observable $A(\phi)$ the energy uncertainty $\delta E$ could actually be reduced even further, in contradiction with the above meaning of $\delta E$. (This argument is not applicable in the quantum case since measuring $A$ would change the state of the system.)

We finally note that the often-considered limit $\delta E \to 0$ immediately implies $\langle A\rangle^*_{eq} = A_\omega(E) = \langle A\rangle_{\rm mic}$, but is no way out: If $A_\omega(E)$ would notably vary within the experimental energy uncertainty $\delta E$, this theoretical limit would be of little use to describe the corresponding real experiment. The fact that focusing on this limiting case is so often considered as sufficient actually confirms once again that the above-mentioned small variations of $A_\omega(E)$ upon small changes of $E$ have always been tacitly considered as a matter of course. In any case, if and only if the difference between (68) and (69) is negligibly small (below the experimental resolution limit) thermalization can be inferred in the sense that

a system with sufficiently many degrees of freedom behaves as if it were in the microcanonical state $\rho_{\mathrm{mic}}(\phi)$ for the overwhelming majority of times after initial transients have died out.

## VIII. SUMMARY AND DISCUSSION

Focusing on isolated macroscopic systems, described in terms of either a quantum mechanical or a classical model, our two key questions were in what sense and under which conditions does an initial ensemble (usually far from equilibrium and largely unknown in detail) evolve towards a stationary long-time behavior (equilibration) and to what extent this steady state is in agreement with the microcanonical ensemble predicted by equilibrium statistical mechanics (thermalization).

Much like in equilibrium statistical mechanics itself, the concomitant quantitative equilibration times are beyond the scope of our present approach. Indeed, none of our calculations admits any meaningful conclusion in this respect. The reason is that any further quantification of the relaxation process inevitably would require considerably more detailed specifications of system (Hamiltonian), initial condition (ensemble), and observables than in our present Secs. II and V. In particular, the question of why certain systems (glasses, Fermi-Pasta-Ulam model, etc.) equilibrate so slowly that they cannot be described in practice by equilibrium statistical mechanics goes beyond our present investigation.

In the quantum case, equilibration was established in the sense that deviations from a time-independent steady state become unmeasurably small for the overwhelming majority of times within any sufficiently large time interval. In doing so, the main assumptions were a finite range and resolution of observables, no exceedingly large degeneracies of energy levels and energy gaps, and small populations of single energy levels by the initial ensemble. Taking these prerequisites for granted appears to be quite convincing in modeling a real experiment whose details are unavoidably unknown in practice. In contrast, thermalization can so far be shown only under additional, not yet sufficiently well-established assumptions, e.g., the eigenstate thermalization hypothesis or Perez's argument (see Sec. IV), whose validity and limitations are presently at the focus of numerous numerical and analytical investigations [8,9,18–39].

In order to demonstrate equilibration in the same sense also in the classical case [42], we had to assume at least weak mixing within every energy surface [see (57)] and apart from a subset of phase space consisting of the union of all KAM tori. Further, we had to assume that the initial ensemble populates this subset only with very small probability. In support of both assumptions, quite suggestive theoretical arguments and numerical evidence can be provided for systems with sufficiently many degrees of freedom (see Sec. V). Regarding rigorous proofs, the problems seems to be extremely difficult in view of the quite limited progress during several decades concerning general (generic) high-dimensional Hamiltonian systems [1,4,44,68,69,73].

Once classical equilibration is established, thermalization can be shown by means of a relatively weak further assumption. On the one hand, this assumption may be viewed as the counterpart of the quantum mechanical eigenstate thermalization hypothesis; on the other hand, one expects that it will be satisfied whenever the observable satisfies some rather mild continuity conditions (see Sec. VII).

As mentioned in Sec. VI, it may well be that classical equilibration can be demonstrated under much weaker conditions than those we have employed here [40–42,74,81]. However, in order to subsequently demonstrate thermalization, Eqs. (63), (68), and (69) suggest that at least ergodicity (30) and condition (32) will still be required in the generic case, i.e., unless (63) happens to become very small due to accidental cancellations.

In spite of the common opinion that classical mechanics should follow from quantum mechanics, our present results in the classical case cannot be deduced from those in the quantum case. Generally speaking, the reason is that the classical limit $\hbar \to 0$ does not commute with the long time limit $T \to \infty$. More specifically, the derivations of relations of the type (12) in [11–17] break down when taking the limit $\hbar \to 0$ while keeping the time $T$ finite. Put differently, the energy spectrum must remain discrete and fixed while making $T$ sufficiently large. For similar reasons, the quantum recurrence theorem for mixed states (density operators) mentioned at the beginning of Sec. III does not survive in the classical limit (the Poincaré recurrence theorem concerns only pure states). The situation is comparable to the relation between quantum chaos and classical Hamiltonian chaos: In the quantum case the main tools are level statistics and random matrix theory and the obtained results do not allow one to recover the key features of chaos in the classical limit, namely, KAM theory and its implications regarding Lyapunov exponents and ergodicity. Accordingly, completely different approaches in the quantum and the classical case are also required and provided in the present work.

In particular, while our present quantum mechanical approach includes pure states as special cases [see below Eq. (4)], the classical counterpart breaks down for pure states [see below Eq. (23)]. In other words, while our present classical approach deals with instantaneous measurements on nontrivial statistical ensembles (mixed states), classical ergodic theory is mainly concerned with time-averaged properties of pure states (see also Sec. I).

A hierarchy of increasingly strong stochasticity properties are commonly used to characterize the degree of chaoticity of a classical Hamiltonian system, the weakest being ergodicity, followed by weak mixing, strong mixing, $K$ systems, Anosov systems, etc. [1,4]. Our present classical approach employs the first (weakest) three categories. The corresponding notions in the quantum case are not very clearly defined. Comparing our present findings in the quantum and the classical cases naturally suggests certain analogies or correspondences along these lines, but they do not seem to provide any additional physical insight.

## ACKNOWLEDGMENTS

Special thanks are due to Sheldon Goldstein for extremely insightful criticisms and hints. This work was supported by Deutsche Forschungsgemeinschaft under Grant No. RE1344/7-1.

## APPENDIX

This appendix provides the derivation of Eq. (63). By means of (39) and (60) we can rewrite (62) as

$$\langle A\rangle^*_{eq} = \int d\phi\, A(\phi) \int d\phi' \rho_0(\phi') \frac{\delta(H(\phi') - H(\phi))}{\omega(H(\phi))}. \tag{A1}$$

By exploiting the definition (66) we furthermore can conclude that

$$\langle A\rangle^*_{eq} = \int d\phi' \rho_0(\phi') A_\omega(H(\phi')). \tag{A2}$$

Equations (36) and (A2) finally yield

$$\langle A\rangle_{eq} - \langle A\rangle^*_{eq} = \int d\phi [\overline{A}(\phi) - A_\omega(H(\phi))]\rho_0(\phi). \tag{A3}$$

Next we can infer from (25) that $A(\phi)$ on the right-hand side of (66) can be replaced by $A(\gamma_t(\phi))$ for any $t$ and thus it can also be replaced by the time average from (28), i.e.,

$$A_\omega(E) := \int d\phi\, \overline{A}(\phi) \frac{\delta(H(\phi) - E)}{\omega(E)}. \tag{A4}$$

Further, recalling the definition of $\omega(E)$ from (61), one readily verifies that

$$A^*(E) = \int d\phi\, A^*(H(\phi)) \frac{\delta(H(\phi) - E)}{\omega(E)}. \tag{A5}$$

In combination with (A4) we thus obtain

$$A_\omega(E) = A^*(E) + \int d\phi\, D(\phi) \frac{\delta(H(\phi) - E)}{\omega(E)}, \tag{A6}$$

$$D(\phi) := \overline{A}(\phi) - A^*(H(\phi)). \tag{A7}$$

Introducing (A6) and (A7) into (A3) yields

$$\langle A\rangle_{eq} - \langle A\rangle^*_{eq} = \int d\phi\, D(\phi)\rho_0(\phi) - J, \tag{A8}$$

$$J := \int d\phi \int d\phi' D(\phi') \frac{\delta(H(\phi') - H(\phi))}{\omega(H(\phi))} \rho_0(\phi)$$

$$= \int d\phi' D(\phi') \frac{\int d\phi\, \delta(H(\phi') - H(\phi))\rho_0(\phi)}{\omega(H(\phi'))}. \tag{A9}$$

The last integral over $\phi$ equals $p(H(\phi'))$ according to (39) and after division by $\omega(H(\phi'))$ we recover $\rho^*_{eq}(\phi')$ from (60). Omitting the primes, (A9) thus takes the form

$$J = \int d\phi\, D(\phi)\rho^*_{eq}(\phi). \tag{A10}$$

With (A7) and (A8) we finally recover (63).


[1] L. Sklar, *Physics and Chance* (Cambridge University Press, Cambridge, 1993).
[2] O. Penrose, Rep. Prog. Phys. **42**, 1937 (1979).
[3] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics*, 2nd ed. (Springer, Berlin, 2009).
[4] A. S. Wightman, in *Statistical Mechanics at the Turn of the Decade*, edited by E. G. D. Cohen (Dekker, New York, 1971), pp. 1–32.
[5] D. Wallace, http://philsci-archive.pitt.edu/410.
[6] I. E. Farquhar, *Ergodic Theory in Statistical Mechanics* (Interscience, New York, 1961).
[7] R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics* (Pergamon, London, 1969).
[8] *Focus on Dynamics and Thermalization in Isolated Quantum Many-Body Systems*, edited by M. A. Cazalilla and M. Rigol, special issue of New J. Phys. **12** (5) (2010).
[9] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Rev. Mod. Phys. **83**, 863 (2011).
[10] V. I. Yukalnov, Laser Phys. Lett. **8**, 485 (2011).
[11] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008).
[12] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E **79**, 061103 (2009).
[13] N. Linden, S. Popescu, A. J. Short, and A. Winter, New J. Phys. **12**, 055021 (2010).
[14] P. Reimann, New J. Phys. **12**, 055027 (2010).
[15] A. J. Short, New J. Phys. **13**, 053009 (2011).
[16] A. J. Short and T. C. Farrelly, New J. Phys. **14**, 013063 (2012).
[17] P. Reimann and M. Kastner, New J. Phys. **14**, 043020 (2012).
[18] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998).
[19] M. Srednicki, J. Phys. A: Math. Gen. **32**, 1163 (1999).
[20] M. A. Cazalilla, Phys. Rev. Lett. **97**, 156403 (2006).
[21] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, Phys. Rev. Lett. **98**, 050405 (2007).
[22] C. Kollath, A. M. Läuchli, and E. Altman, Phys. Rev. Lett. **98**, 180601 (2007).
[23] S. R. Manmana, S. Wessel, R. M. Noack, and A. Muramatsu, Phys. Rev. Lett. **98**, 210405 (2007).
[24] M. Cramer, C. M. Dawson, J. Eisert, and T. J. Osborne, Phys. Rev. Lett. **100**, 030602 (2008).
[25] T. Barthel and U. Schollwöck, Phys. Rev. Lett. **100**, 100601 (2008).
[26] M. Rigol, V. Dunjko, and M. Olshanii, Nature (London) **452**, 854 (2008).
[27] G. Roux, Phys. Rev. A **81**, 053604 (2010).
[28] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).
[29] G. Biroli, C. Kollath, and A. M. Läuchli, Phys. Rev. Lett. **105**, 250401 (2010).
[30] C. Gogolin, M. P. Müller, and J. Eisert, Phys. Rev. Lett. **106**, 040401 (2011).
[31] T. N. Ikeda, Y. Watanabe, and M. Ueda, Phys. Rev. E **84**, 021130 (2011).
[32] M. C. Bañuls, J. I. Cirac, and M. B. Hastings, Phys. Rev. Lett. **106**, 050405 (2011).
[33] L. Campos Venuti, N. T. Jacobson, S. Santra, and P. Zanardi, Phys. Rev. Lett. **107**, 010403 (2011).
[34] N. T. Jacobson, L. C. Venuti, and P. Zanardi, Phys. Rev. A **84**, 022115 (2011).
[35] M. Kastner, Phys. Rev. Lett. **106**, 130601 (2011).
[36] K. Ji and B. V. Fine, Phys. Rev. Lett. **107**, 050401 (2011).

[37] E. Canovi, D. Rossini, R. Fazio, G. E. Santoro, and A. Silva, Phys. Rev. B **83**, 094431 (2011).

[38] M. Kastner, Cent. Eur. J. Phys. **10**, 637 (2012).

[39] M. Rigol and M. Srednicki, Phys. Rev. Lett. **108**, 110601 (2012).

[40] A. Y. Khinchin, *Mathematical Foundations of Quantum Statistics* (Graylock, New York, 1960).

[41] J. L. Lebowitz, arXiv:0709.0724v1; S. Goldstein, in *Chance in Physics: Foundations and Perspectives*, edited by J. Bricmont, D. Durr, M. C. Galavotti, G. Ghirardi, F. Petruccione, and N. Zanghi, Lecture Notes in Physics Vol. 574 (Springer, Berlin, in press); S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Phys. Rev. E **81**, 011109 (2010); S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Eur. J. Phys. H **35**, 173 (2010); S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Proc. R. Soc. London Ser. A **466**, 3203 (2010); H. Tasaki, arXiv:1003.5424v4.

[42] A. Hobson, *Concepts in Statistical Mechanics* (Gordon and Breach, New York, 1971).

[43] L. Markus and K. R. Mayer, Mem. Am. Math. Soc. **144**, 1 (1978).

[44] L. A. Bunimovich, Nonlinearity **21**, T13 (2008).

[45] E. P. Wigner, Am. J. Phys. **31**, 6 (1963).

[46] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992), Chap. XVII.7.

[47] A. Sugita, Nonlin. Phenom. Complex Syst. **10**, 192 (2007).

[48] O. Penrose, *Foundations of Statistical Mechanics* (Pergamon, Oxford, 1970), Chap. 1.

[49] A. Martin-Löf, *Statistical Mechanics and the Foundations of Thermodynamics*, Lecture Notes in Physics Vol. 101 (Springer, Berlin, 1979).

[50] H.-O. Georgii, J. Stat. Phys. **80**, 1341 (1995).

[51] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).

[52] P. Reimann, J. Stat. Phys. **132**, 921 (2008).

[53] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1980); B. Diu, C. Guthmann, D. Lederer, and B. Roulet, *Elements de Physique Statistique* (Hermann, Paris, 1996).

[54] G. Ludwig, Z. Phys. **150**, 346 (1958); **152**, 98 (1958).

[55] P. Bocchieri and A. Loinger, Phys. Rev. **114**, 948 (1959).

[56] A. Peres, Phys. Rev. A **30**, 504 (1984).

[57] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991).

[58] M. Srednicki, J. Phys. A: Math. Gen. **29**, L75 (1996).

[59] Note that no such degeneracies are actually encountered in the generic case [11,12,14,18,19,56,60], i.e., in the absence of special reasons such as additional conserved quantities [e.g., due to (perfect) symmetries or when the system consists of noninteracting subsystems] or fine-tuning of parameters ("accidental degeneracies").

[60] S. Goldstein, J. L. Lebowitz, J. L. Tumulka, and N. Zanghì, J. Stat. Phys. **125**, 1193 (2006).

[61] In particular, $T$ must obviously be much larger than the relaxation time in the case of a far-from-equilibrium initial condition $\rho(0)$.

[62] Note that gaps are a quite common feature of the energy spectra of single particles (or quasiparticles). In contrast, our present $E_n$ and $|n\rangle$ refer to the many-particle eigenenergies and eigenstates of the total (many-particle) Hamiltonian $H$ from Eq. (1). Even if the single-particle spectra exhibit gaps, the many-particle spectra are generically gapless.

[63] M. Feingold and A. Peres, Phys. Rev. A **34**, 591 (1986).

[64] M. Srednicki, Phys. Rev. E **50**, 888 (1994); arXiv:cond-mat/9410046.

[65] R. V. Jensen and R. Shankar, Phys. Rev. Lett. **54**, 1879 (1985).

[66] M. Feingold, N. Moiseyev, and A. Peres, Phys. Rev. A **30**, 509 (1984).

[67] S. Weigert, Physica D **56**, 107 (1992); B. Sutherland, *Beautiful Models* (World Scientific, Singapore, 2004), Chap. 2.1; A. Enciso and D. Peralta-Salas, Theor. Math. Phys. **148**, 1086 (2006) [Teoret. Mat. Fiz. **148**, 249 (2006)].

[68] J. Ford, Phys. Rep. **213**, 271 (1992).

[69] A. L. Lichtenberg and M. A. Liebermann, *Regular and Chaotic Motion* (Springer, Berlin, 1983), Sec. 6.5.

[70] P. Hartman and A. Wintner, Am. J. Math. **62**, 646 (1940).

[71] P. R. Halmos, Bull. Am. Math. Soc. **55**, 1015 (1949).

[72] The possibility of replacing (31) by substantially weaker conditions will be discussed in Sec. VI.

[73] C. Froeschlé and J.-P. Scheidecker, Phys. Rev. A **12**, 2137 (1975); M. Falcioni, U. M. Marconi, and A. Vulpiani, *ibid.* **44**, 2263 (1991); E. G. Altmann and H. Kantz, Europhys. Lett. **78**, 10008 (2007).

[74] R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1971).

[75] Our present notion of equilibration is thus very similar to the one by Hobson (see Chap. 5.2 in [42]), which is based on the concept of weak convergence. Its close relation to the notion of coarse graining is discussed, e.g., in [2].

[76] The $\delta$ function in (40) is a convenient formal way to effectively project the Euclidean Lebesgue measure of the phase space on the energy surfaces in the usual way; see, e.g., pp. 22–23 in [77] and pp. 1945–1946 in [2].

[77] T. Levi-Civita, J. Math. Phys. **13**, 18 (1934).

[78] Similarly, also the convergence of $\overline{\sigma^2(t)}^T$ for $T \to \infty$ can be deduced from (37). In both cases, a rigorous justification of commuting integrations with the $T \to \infty$ limit is, however, not obvious at all. It is closely related to the question of whether or not the characteristic convergence time of those quantities will be comparable to the notoriously large ergodicity time scales governing the convergence in (28) and (31). The analogous problem in the quantum case is treated, e.g., in Sec. 4.3 of [17].

[79] Since $\overline{\sigma^2(t)}^T$ does not necessarily approach $\overline{\sigma^2(t)}^\infty$ from below, $\epsilon$ is not superfluous in (47).

[80] Abbreviating the integral appearing in (56) as $h(t)$, it follows with (27) that $|h(t)|$ can be bounded from above by some finite constant $c$ for all $t$. Hence $\overline{[h(t)]^2}^T \leqslant c\overline{|h(t)|}^T$. As a consequence, (57) implies (56).

[81] D. Ruelle, *Chance and Chaos* (Princeton University Press, Princeton, 1991); L. Bunimovich, Chaos **13**, 903 (2003).