PHYSICAL REVIEW LETTERS

Generalization of von Neumann's Approach to Thermalization

Peter Reimann

Fakultät für Physik, Universität Bielefeld, 33615 Bielefeld, Germany

Thermalization of isolated many-body systems is demonstrated by generalizing an approach originally due to von Neumann: For arbitrary initial states with a macroscopically well-defined energy, quantum mechanical expectation values become indistinguishable from the corresponding microcanonical expectation values for the overwhelming majority of all sufficiently late times. As in von Neumann's work, the eigenvectors of the Hamiltonian and of the considered observable are required to not exhibit any specially tailored (untypical) orientation relative to each other. But all of von Neumann's further assumptions about the admitted observables are abandoned.

The universal and irreversible tendency of nonequilibrium states towards thermal equilibrium is an everyday experience in the macroscopic world, but, in spite of more than a century of theoretical efforts, it has still not been satisfactorily reconciled with the basic laws of physics, which govern the microscopic world and which are fundamentally reversible [1]. The first quantum mechanical exploration of this problem is due to von Neumann [2], was unfortunately misunderstood for decades, but has recently been rehabilitated in a very enlightening commentary by Goldstein, Lebowitz, Tumulka, and Zanghì [3]. A major remaining bottleneck of von Neumann's approach is his notion of "macro-observer" or "macroscopic measurement," stipulating that all relevant observables can be approximated by commuting Hermitian operators with very high-dimensional common eigenspaces [2]. As an alternative, Goldstein *et al.* [3,4] suggested to consider "macroscopic observables" with the additional property (excluded in von Neumann's original treatment) that one of those eigenspaces is overwhelmingly large compared to all the others. In our present work, all such restrictions with respect to the considered observables are abandoned.

As in Refs. [2–4], we consider an isolated many-body system, whose energy E is known up to an uncertainty δE, which is small on the macroscopic but large on the microscopic scale. The system is modeled by a Hamiltonian H with eigenvalues E_n and eigenvectors $|n\rangle$, $n \in \mathbb{N}$. System states (pure or mixed) are described by density operators ρ, evolving in time according to the usual Liouville–von Neumann equation $\dot{\rho}(t) = i[\rho(t), H]/\hbar$. Observables are modeled by Hermitian operators A with expectation values $\langle A \rangle_\rho := \text{Tr}\{\rho A\}$. The preset energy interval $[E, E + \delta E]$ defines an energy shell, namely, the Hilbert space $\mathcal{H}$ spanned by all $|n\rangle$ with $E_n \in [E, E + \delta E]$. Without loss of generality, we assume that the corresponding labels are $n = 1, 2, \ldots, D$. For a macroscopic system with, say, $f \approx 10^{23}$ degrees of freedom, the dimensionality D of $\mathcal{H}$ is exponentially large in f [3], symbolically indicated as

$$D \approx 10^{\mathcal{O}(f)}. \qquad (1)$$

By definition, the probability to encounter a system energy outside $[E, E + \delta E]$ is negligibly small and is henceforth idealized as being strictly zero. As a consequence, the diagonal matrix elements ("level populations") $\rho_{nn} := \langle n|\rho|n \rangle$ vanish for all $n > D$, implying with Cauchy-Schwarz's inequality that $\rho_{mn} = 0$ if $m > D$ or $n > D$. Denoting by P the projector onto $\mathcal{H}$, the projection (or restriction) of A onto $\mathcal{H}$ takes the form $\tilde{A} := PAP$ and analogously $\tilde{H} := PHP$, etc. (note that $\tilde{\rho} = \rho$). It readily follows that $\text{Tr}\{\rho A\} = \text{Tr}\{\rho \tilde{A}\}$ and that $\tilde{H}$ yields the same time evolution of $\rho(t)$ as H. Hence we can and will restrict ourselves to the energy shell $\mathcal{H}$ from now on but, for convenience, omit the tilde symbols. Accordingly, P becomes the identity operator on $\mathcal{H}$, and the microcanonical density operator follows as $\rho_{\text{mc}} := P/D$ with expectation values $\langle A \rangle_{\text{mc}} := \text{Tr}\{\rho_{\text{mc}}A\}$.

The problem of thermalization is to show that $\langle A \rangle_{\rho(t)}$ evolves towards $\langle A \rangle_{\text{mc}}$ for arbitrary (possibly far from equilibrium) initial conditions $\rho(0): \mathcal{H} \to \mathcal{H}$. It is well known that this is impossible without additional assumptions on H and A. With respect to H, we adopt von Neumann's assumption [2] that the energy differences $E_m - E_n$ are finite and mutually different for all pairs $m \neq n$. Excluding nongeneric cases with additional conserved quantities (besides H), e.g., due to (perfect) symmetries or noninteracting subsystems, the validity of this assumption is by now commonly accepted [4–11]. Moreover, one expects that even considerably weaker assumptions will do [12–15].

Denoting by a_{max} and a_{min} the largest and smallest, respectively, among the D eigenvalues of A, the range of A is defined as $\Delta_A := a_{\text{max}} - a_{\text{min}}$. Furthermore, in any real (or numerical) experiment, $\langle A \rangle_\rho$ can be determined only with some finite accuracy δA. In practice, we thus can focus on measurements which yield at most, say, 20 relevant digits, i.e.,

$$\delta A \geq \Delta_A 10^{-20}. \qquad (2)$$

The eigenvectors of H and of A are related by some unitary basis transformation U. A key point of von Neumann's approach is the assumption that these two eigenbases do not exhibit any "special orientation" relative to each other [2,16]; i.e., the actual U is "typical" [3] among all possible unitary transformations $U: \mathcal{H} \to \mathcal{H}$ in the following sense: If a certain property can be shown to hold for the vast majority of U's (uniformly distributed according to the Haar measure [2–4,16]), then this property is supposed to hold for the actual U as well, unless there are special reasons to the contrary. Denoting by $\mu_U(X)$ the fraction (normalized measure) of all U's exhibiting a certain property X, a $\mu_U(X)$ value close to unity (zero) is thus assumed to generically imply (exclude) property X for the actual system. While a more rigorous justification is clearly very difficult, intuitively such a "typicality" argument is very convincing: If we imagine A as fixed and H as arising by randomly sampling its eigenvectors via U [4,16,17], the argument is essentially tantamount to the common lore of random matrix theory [3], which is well known to be extremely successful in practice [18]. In particular, $\mu_U(X)$ may be *formally* viewed as the probability of observing property X for a randomly sampled H (or U), however keeping in mind—exactly as in random matrix theory—that there is *no random sampling procedure in the real physical problem under consideration* [3,4]. In passing, we note that von Neumann actually adopted the complementary viewpoint of considering H as fixed while varying the eigenvectors of A [2].

By exploiting the above mentioned common assumptions about the energy eigenvalues E_n [2,4–11], one can infer [6–12] (see also [19]) that the quantity

$$\sigma^2(t) := [\langle A \rangle_{\rho(t)} - \langle A \rangle_{\bar{\rho}}]^2 \qquad (3)$$

satisfies the relations

$$\overline{\sigma^2(t)} = \sum_{m \neq n}^{D} |\rho_{mn}(0)|^2 |A_{mn}|^2 \leq \max_{m \neq n} |A_{mn}|^2, \qquad (4)$$

where $A_{mn} := \langle m|A|n \rangle$, $\rho_{mn}(0) := \langle m|\rho(0)|n \rangle$, and the overbar indicates an average over all times $t \geq 0$. In particular, $\bar{\rho} := \overline{\rho(t)}$ is an auxiliary density operator with matrix elements $\bar{\rho}_{mn} = \delta_{mn}\rho_{nn}(0)$, sometimes named the diagonal or generalized Gibbs ensemble [21]. The so-called eigenstate thermalization hypothesis (ETH) *conjectures* that, for a many-body system with $f \gg 1$ degrees of freedom, typical off-diagonal elements A_{mn} in (4) are exponentially small in f [8,21–25]. Within our present generalization of von Neumann's approach, we can actually *prove* that even their maximum on the right-hand side in (4) is typically so small that

$$\mu_U\left[\overline{\sigma^2(t)} \geq \epsilon\right] \leq 4 \exp\left\{-\frac{\epsilon D}{18\pi^3 \Delta_A^2} + 2 \ln D\right\} \qquad (5)$$

for any $\epsilon > 0$. Besides (4), the key ingredient in deriving this result is Levy's lemma (see [26,27], and further references therein), stating that

$$\text{Prob}[|g(\phi) - \langle g \rangle| \geq \epsilon] \leq 2 \exp\left\{-\frac{\epsilon^2(d+1)}{9\pi^3 \eta^2}\right\} \qquad (6)$$

for randomly and uniformly distributed points ϕ on the d-dimensional unit sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ and any Lipschitz continuous function $g: \mathbb{S}^d \to \mathbb{R}$ with Lipschitz constant η and mean value $\langle g \rangle$. Furthermore, any normalized $|\phi\rangle \in \mathcal{H}$ of the form $\sum_{n=1}^{D} c_n |n\rangle$ can be represented (via the real and imaginary parts of the c_n's) as a point ϕ on the $(2D-1)$-dimensional unit sphere. Finally, one can show [27] that $g(\phi) := \langle \phi|A|\phi \rangle$ is Lipschitz continuous with $\eta = \Delta_A$ and $\langle g \rangle = \langle A \rangle_{\text{mc}}$. Observing that randomizing ϕ is equivalent to randomizing U, we thus obtain

$$\mu_U[|\langle \phi|A|\phi \rangle - \langle A \rangle_{\text{mc}}| \geq \epsilon] \leq 2 \exp\left\{-\frac{2\epsilon^2 D}{9\pi^3 \Delta_A^2}\right\}. \qquad (7)$$

The remaining task is to connect this result for $\langle \phi|A|\phi \rangle$ with the maximal $|A_{mn}|$ in (4). The details are rather straightforward but tedious and thus provided as Supplemental Material in [19]. As an aside, it follows that von Neumann's main technical achievement (Appendix of Ref. [2]), as well as its further improvement by Pauli and Fierz [28], is in fact quite closely related to Levy's lemma (see also [19]).

Equation (5) represents the first main result of our Letter. By choosing, e.g., $\epsilon = D^{-1/2}\delta A^2$ in (5), it follows with (1) and (2) that the time-averaged variance from (3) remains extremely much smaller than δA^2 for "almost all" U: The fraction of the exceptional U's is an unimaginably small number of the order of 10^{-x} with $x \approx 10^{\mathcal{O}(f)}$, $f \approx 10^{23}$. Furthermore, the mere existence of the infinite time average in (4) implies that a similar estimate must also apply to averages of $\sigma^2(t)$ over finite time intervals $[0, T]$ with sufficiently large T [3,7,13,14]. Finally, the smallness of the latter time average implies (obviously or by Markov's inequality) that the averaged quantity (3) itself must be exceedingly small for most times $t \in [0, T]$ [3,7,10,14]. For example, for our above choice $\epsilon = D^{-1/2}\delta A^2$ and sufficiently large T, all the "bad times" $t \in [0, T]$ with $|\langle A \rangle_{\rho(t)} - \langle A \rangle_{\bar{\rho}}| \geq \delta A$ add up to a set, whose Lebesgue measure is smaller by (at least) a factor of the order $D^{1/4} \approx 10^{\mathcal{O}(f)}$ [cf. (1)] than the measure of all $t \in [0, T]$. Altogether, we thus can conclude that, for the overwhelming majority of U's, the difference $\langle A \rangle_{\rho(t)} - \langle A \rangle_{\bar{\rho}}$ remains below the resolution limit δA for the vast majority of times t contained in any sufficiently large time interval

$[0, T]$. The same conclusion carries over to our actual Hamiltonian H and observable A, given their eigenbases are related by a typical transformation U as discussed above. To establish quantitative bounds for T is a subject of considerable current interest [8,13,29–32] but goes beyond our present scope.

The salient point is that (5) holds independently of the initial condition $\rho(0)$. Once a pair H, A with a typical U is given, the above implications of (5) thus apply to *any* $\rho(0)$: No matter how far from equilibrium the system starts out, for almost all sufficiently late times it behaves practically as if it were in the state $\bar{\rho}$. Such an apparent convergence towards a steady state has been denoted as *equilibration*, e.g., in Refs. [9–15].

To demonstrate *thermalization*, we still have to show that the difference between $\langle A \rangle_{\bar{\rho}}$ and $\langle A \rangle_{\mathrm{mc}}$ is negligibly small. Recalling the definitions of these two expectation values, one readily sees that

$$B := \langle A \rangle_{\bar{\rho}} - \langle A \rangle_{\mathrm{mc}} = \sum_{n=1}^{D} \rho_{nn}(0)[A_{nn} - \langle A \rangle_{\mathrm{mc}}] \quad (8)$$

and hence that

$$|B| \leq \max_{n} |A_{nn} - \langle A \rangle_{\mathrm{mc}}|. \quad (9)$$

Similarly as above Eq. (5), one part (actually the better known part) of ETH consists in the *conjecture* that typical differences $A_{nn} - \langle A \rangle_{\mathrm{mc}}$ are exponentially small in f [8,21–25]. Within our present framework, we can prove that even their maximum in (9) is typically so small that

$$\mu_U(|B| \geq \epsilon) \leq 2 \exp\left\{-\frac{2}{9\pi^3}\frac{\epsilon^2 D}{\Delta_A^2} + \ln D\right\} \quad (10)$$

for any $\epsilon > 0$. This represents our second main result, whose derivation from (7) is quite obvious and is provided in full detail as Supplemental Material in [19]. Once again, it is crucial to note that (10) is independent of $\bar{\rho}$ [and thus of $\rho(0)$]: Given a pair H, A with a typical U, it follows from (8) and (10) that the difference $\langle A \rangle_{\bar{\rho}} - \langle A \rangle_{\mathrm{mc}}$ remains way below the resolution limit δA for *any* $\bar{\rho}$ [or $\rho(0)$]. Finally, upon considering A as fixed and H as a random matrix (see above), we can conclude that von Neumann's approach [2] in fact anticipates the verification of ETH from Ref. [22] within a random matrix theoretical framework; see also [3,17].

One readily sees that the measure of all U's which give simultaneously rise to both equilibration, as discussed in the paragraph below (7), *and* negligibly small B values according to (10) is still extremely close to unity. Hence, thermalization follows for any given pair H, A with a generic relative orientation of the eigenbases, no matter how the initial condition $\rho(0)$ is chosen. Along the same lines, one can infer the simultaneous thermalization of several (not necessarily commuting) observables [12,33], as long as their number remains "reasonable" (e.g., smaller than D).

Similarly as for U, let us now denote by V the unitary basis transformation between the eigenvectors of the density operator $\rho(0)$ and those of H. Likewise, $\mu_V(X)$ now represents the fraction (normalized measure) of all unitary transformations $V: \mathcal{H} \to \mathcal{H}$ which exhibit a certain property X. Furthermore, the usual von Neumann entropy is defined as $S[\rho] := -k_B \mathrm{Tr}\{\rho \ln \rho\}$ and satisfies $0 \leq S[\rho] \leq S[\rho_{\mathrm{mc}}] = k_B \ln D$. Hence, the entropy range is $\Delta_S = k_B \ln D$, and, similarly as in (2), experimentally resolvable entropy differences δS can be assumed to satisfy $\delta S \geq q\Delta_S = qk_B \ln D$ for some small but still reasonable q value. It follows that $\rho(0)$ entails a $\bar{\rho}$ with the properties that $S[\rho_{\mathrm{mc}}] - S[\bar{\rho}] \geq 0$ and, as demonstrated in detail in Supplemental Material [19],

$$\mu_V(S[\rho_{\mathrm{mc}}] - S[\bar{\rho}] \geq s) \leq k_B/s \quad (11)$$

for any $s > 0$. This is our third main result. By choosing $s = \delta S$ and recalling that $\delta S \geq qk_B \ln D$ (see above), it implies with (1) that the entropy of the diagonal ensemble $\bar{\rho}$, towards which the true $\rho(t)$ seems to equilibrate, differs from the microcanonical entropy only by an unmeasurably small amount for a generic $\rho(0)$, i.e., one without a specially tailored orientation of its eigenbasis relative to that of H. We remark that already von Neumann demonstrated a somewhat similar, so-called H theorem [2], however, for a differently defined entropy, whose physical relevance has been questioned, e.g., in Ref. [3]. Further related but different results about entropies of diagonal ensembles are also due to Ref. [34].

As shown in Refs. [9–11,19], an alternative upper bound for the left-hand side of (4) is given by $(\Delta_A^2/4)\mathrm{Tr}\{\bar{\rho}^2\}$. For the latter factor $\mathrm{Tr}\{\bar{\rho}^2\}$, a similar relation as in (11) is derived in Supplemental Material [19], yielding

$$\mu_V\left[\overline{\sigma^2(t)} \geq \epsilon\right] \leq \Delta_A^2/(2\epsilon D) \quad (12)$$

for any $\epsilon > 0$. By analogous arguments as in the discussion of (5) in the paragraph below (7), this amounts to an alternative demonstration of equilibration [7,9–11]. But in contrast to (5), which applies to arbitrary $\rho(0)$, provided the relative eigenbasis orientation of H and A is generic, the present findings now apply to arbitrary observables A, provided the eigenbases of H and $\rho(0)$ are in a generic constellation.

Finally, let us denote by W the unitary basis transformation between the eigenvectors of the density operator $\rho(0)$ and those of A and consider

$$\sigma_{\mathrm{mc}}^2(t) := [\langle A \rangle_{\rho(t)} - \langle A \rangle_{\mathrm{mc}}]^2. \quad (13)$$

Similarly as before, we now can show (see Supplemental Material [19]) that

$$\mu_W[\sigma_{\mathrm{mc}}^2(t) \geq \epsilon] \leq \Delta_A^2/(\epsilon D) \qquad (14)$$

for arbitrary t and $\epsilon > 0$ and that

$$\mu_W\left(\frac{1}{t_2 - t_1}\int_{t_1}^{t_2}\sigma_{\mathrm{mc}}^2(t)dt \geq \epsilon\right) \leq \Delta_A^2/(\epsilon D) \qquad (15)$$

for arbitrary $t_1 < t_2$. Note that, while (14) is a t-independent upper bound for the *measure* of all W's with $\sigma_{\mathrm{mc}}^2(t) \geq \epsilon$, this does not imply that the *set* of those W's is t independent. An analogous caveat applies to (15). Yet another crucial point is that (14) and (15) are valid for completely arbitrary (even time-dependent) Hamiltonians $H: \mathcal{H} \to \mathcal{H}$ [19,35,36].

A first remarkable implication of (14) and (15) follows by considering A as "given" (arbitrary but fixed): Namely, "most" [37] $\rho(0)$ then yield practically the same expectation value $\langle A \rangle_{\rho(t)}$ as ρ_{mc} for any arbitrary but fixed time point t but also for practically all times t within an arbitrary but fixed time interval $[t_1, t_2]$ [see the paragraph below (7)]. Put differently, nonequilibrium expectation values are "untypical" (even for $t = 0$); they require very special orientations W of the eigenbasis of $\rho(0)$ relative to that of A. In particular, for pure states $\rho(0) = |\psi\rangle\langle\psi|$ we recover the quintessence of so-called canonical typicality and related phenomena [26,33,38–40] (see also [36,41]).

Conversely, when considering $\rho(0)$ as given, it follows from (14) and (15) that most [37] measurement devices A cannot distinguish $\rho(t)$ from ρ_{mc} at any arbitrary but fixed time point t or for practically all times t within an arbitrary but fixed time interval $[t_1, t_2]$. This is the viewpoint adopted, e.g., in Refs. [35,42] but now formulated within our present generalization of von Neumann's original approach (see also [31]).

The fact that a nonequilibrium value of $\langle A \rangle_{\rho(0)}$ requires an untypical pair $\rho(0)$, A implies [3] that conclusions regarding thermalization can be drawn only from results concerning *all* orientations of $\rho(0)$ relative to A, as it is the case in von Neumann's approach [see below (7) and (10)], but not from results concerning *most* orientations, as in the above generalization (14) and (15) of the approach from Refs. [35,36,41,42]. In other words, it is not right to say that von Neumann's approach is inadequate to investigate thermalization, since the same applies to the approach from Refs. [35,36,41,42]. Rather, the two approaches are fundamentally different: One requires a generic eigenbasis constellation of H and A but admits any $\rho(0)$, and the other requires a generic eigenbasis constellation of $\rho(0)$ and A but admits any H.

In conclusion, von Neumann's demonstration of thermalization for isolated many-body systems has been generalized to arbitrary observables. The remaining prerequisites for thermalization are thus rather weak, namely, a Hamiltonian with generic eigenvalues E_n and a generic orientation of its eigenvectors relative to those of A, while the initial state $\rho(0)$ may still be chosen arbitrarily (mixed or pure, far from equilibrium or not). The first requirement (regarding E_n) is by now well established [4–11], and further generalizations like in Refs. [12–15] seem possible. With the second requirement, von Neumann essentially anticipated the foundation of random matrix theory [3], which is very difficult to justify rigorously but is extremely successful in practice, and can be corroborated by various intuitively convincing arguments [18]. For instance, since our mind is used to thinking about the physical world in terms of individual "particles," we mostly come up with single-particle observables A or sums thereof [kinetic energy (temperature), density, pressure, magnetization, etc.], whose eigenvectors are thus single-particle product states. In contrast, a generic Hamiltonian H includes particle-particle interactions, giving rise to a "completely different" eigenbasis without any "special relation" to that of A [6].

While von Neumann had in mind a preset H and a varying (or typical) eigenbasis of A [2], the mathematically equivalent but physically opposite viewpoint (fixed A, varying H) was emphasized, e.g., in Refs. [3,4,16,17]. Here, both views have been merged and significantly generalized by treating all three operators A, H, and $\rho(0)$ on an equal footing: After selecting two of them and assuming they exhibit a typical eigenbasis constellation, we were able to draw conclusions which are then entirely independent of the third one. Along these lines, we established the general new results (11) and (12) concerning the generic long-time behavior (equilibration) for expectation values of arbitrary (even untypical) observables and the entropy of the concomitant equilibrium states. Furthermore, our findings (14) and (15) significantly generalize previously known typicality results for arbitrary (even time-dependent) Hamiltonians. As a by-product, we thus obtained a unifying framework for several key aspects of thermalization, such as the validation of the eigenstate thermalization hypothesis [8,21,23,25] by means of the random matrix theory [22], recent explorations of "equilibration" [9–15] and "canonical typicality" [26,33,38–40], the long-lasting misjudgment of von Neumann's work [35,36,41,42], and its rehabilitation in Ref. [3].

[1] L. Sklar, *Physics and Chance* (Cambridge University Press, Cambridge, England, 1993).

[2] J. von Neumann, Z. Phys. **57**, 30 (1929) [English translation by R. Tumulka, Eur. Phys. J. H **35**, 201 (2010)].

[3] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Eur. Phys. J. H **35**, 173 (2010).

[4] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Phys. Rev. E **81**, 011109 (2010).

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

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

[7] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998).

[8] M. Srednicki, J. Phys. A **32**, 1163 (1999).

[9] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008).

[10] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E **79**, 061103 (2009).

[11] P. Reimann, New J. Phys. **12**, 055027 (2010).

[12] A. J. Short, New J. Phys. **13**, 053009 (2011).

[13] A. J. Short and T. C. Farrelly, New J. Phys. **14**, 013063 (2012).

[14] P. Reimann and M. Kastner, New J. Phys. **14**, 043020 (2012).

[15] P. Reimann, Phys. Scr. **86**, 058512 (2012).

[16] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, Proc. R. Soc. A **466**, 3203 (2010).

[17] M. Fierz, Helv. Phys. Acta **28**, 705 (1955).

[18] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Rev. Mod. Phys. **53**, 385 (1981).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.115.010403 for mathematical details, which includes Ref. [20].

[20] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics*, 2nd ed. (Springer, Berlin, 2009).

[21] M. Rigol, V. Dunjko, and M. Olshanii, Nature (London) **452**, 854 (2008).

[22] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991); New J. Phys. **12**, 075021 (2010); P. Reimann, *ibid.* **17**, 055025 (2015).

[23] M. Srednicki, Phys. Rev. E **50**, 888 (1994).

[24] S. Goldstein and R. Tumulka, in *Non-Equilibrium Statistical Physics Today: Proceedings of the 11th Granada Seminar on Computational and Statistical Physics*, edited by P. L. Garrido, F. de los Santos, and J. Marro (American Institute of Physics, New York, 2011), p. 155.

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

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

[27] S. Popescu, A. J. Short, and A. Winter, arXiv:quant-ph/0511225v3.

[28] W. Pauli and M. Fierz, Z. Phys. **106**, 572 (1937).

[29] S. Goldstein, T. Hara, and H. Tasaki, Phys. Rev. Lett. **111**, 140401 (2013).

[30] T. Monnai, J. Phys. Soc. Jpn. **82**, 044006 (2013).

[31] A. S. L. Malabarba, L. P. Garcia-Pintos, N. Linden, T. C. Farrelly, and A. J. Short, Phys. Rev. E **90**, 012121 (2014).

[32] S. Goldstein, T. Hara, and H. Tasaki, New J. Phys. **17**, 045002 (2015).

[33] P. Reimann, Phys. Rev. Lett. **99**, 160404 (2007).

[34] A. Polkovnikov, Ann. Phys. (Amsterdam) **326**, 486 (2011); L. F. Santos, A. Polkovnikov, and M. Rigol, Phys. Rev. Lett. **107**, 040601 (2011).

[35] P. Bocchieri and A. Loinger, Phys. Rev. **111**, 668 (1958).

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

[37] In principle, most is meant with respect to eigenbasis variations via *W*, but different spectra can also be readily included.

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

[39] C. Bartsch and J. Gemmer, Phys. Rev. Lett. **102**, 110403 (2009).

[40] S. Sugiura and A. Shimizu, Phys. Rev. Lett. **108**, 240401 (2012).

[41] P. Pechukas, J. Math. Phys. (N.Y.) **25**, 532 (1984).

[42] I. E. Farquhar and P. T. Landsberg, Proc. R. Soc. A **239**, 134 (1957).