Extremely quick thermalization in a macroscopic quantum system for a typical nonequilibrium subspace

Sheldon Goldstein[1], Takashi Hara[2], and Hal Tasaki[3]

Abstract

The fact that macroscopic systems approach thermal equilibrium may seem puzzling, for example, because it may seem to conflict with the time-reversibility of the microscopic dynamics. We here prove that in a macroscopic quantum system for a typical choice of "nonequilibrium subspace", any initial state indeed thermalizes, and in fact does so very quickly, on the order of the Boltzmann time $\tau_{\mathrm{B}} := h/(k_{\mathrm{B}}T)$. Therefore what needs to be explained is, not that macroscopic systems approach thermal equilibrium, but that they do so slowly.

1 Introduction

The recent renewed interest in the foundation of quantum statistical mechanics has led to a revival of the old approach by von Neumann to investigate the problem of thermalization only in terms of quantum dynamics in an isolated system [1, 2]. It has been demonstrated in some general or concrete settings that a pure initial state evolving under quantum dynamics indeed equilibrates or thermalizes[4] in a certain mathematical sense [3, 4, 5, 6, 7, 8, 9, 10]. See also [11, 12, 13, 14, 15] for closely related idea that a typical pure state of a macroscopic quantum system can fully describe thermal equilibrium.

Though the above mentioned works establish that a wide class of isolated quantum systems exhibits the approach to equilibrium, they provide basically no information about the necessary time scale. In our recent work [16], we have treated this problem of time scale in an abstract setting. We proved via an explicit (and artificial) construction that, depending on the system (or, more precisely, the choice of nonequilibrium subspace $\mathcal{H}_{\mathrm{neq}}$ introduced later), the relaxation time can be extremely large, exceeding the age of the universe, or ridiculously short.

This motivates us to study the time scale of thermalization, or, more precisely, that of the escape from the nonequilibrium subspace $\mathcal{H}_{\mathrm{neq}}$, for various choices of $\mathcal{H}_{\mathrm{neq}}$. It may be natural to first focus on a setting in which the "nonequilibrium subspace" $\mathcal{H}_{\mathrm{neq}}$ is not, in fact, a realistic nonequilibrium subspace, but rather is chosen in a completely random manner. One might hope that for such a subspace one generically has a realistic relaxation time. If this were so, it would seem reasonable to believe that the same thing would probably be true for a realistic nonequilibrium subspace (unless we have reasons

[1]Departments of Mathematics and Physics, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA

[2]Faculty of Mathematics, Kyushu University, Moto-oka, Nishi-ku, Fukuoka 819-0395, Japan

[3]Department of Physics, Gakushuin University, Mejiro, Toshima-ku, Tokyo 171-8588, Japan

[4]Usually equilibration means the approach to a certain macroscopic state while thermalization means the approach to thermal equilibrium.

to expect otherwise). This is basically von Neumann's philosophy in [1]. Unfortunately this hope turns out to be far too optimistic.

In the present paper, we report on our recent mathematically rigorous (and probably rather surprising) result implying that, for *any* initial state (and hence for those that are far from equilibrium), the relaxation time in the above setting is quite short, of order the Boltzmann time $\tau_{\rm B} := h/(k_{\rm B}T)$. Since $\tau_{\rm B} \approx 10^{-13}$ s at room temperature, we have basically proved that our coffee is no longer hot after, say, a micro-second. This physically absurd but mathematically rigorous conclusion has the important implication that the moderately slow decay observed in reality is *not* a typical property for a random choice of $\mathcal{H}_{\rm neq}$. It suggests that we have to take into account some essential features of realistic nonequilibrium subspaces in order to fully understand the problem of thermalization in isolated quantum systems. We shall discuss possible directions for future research.

In the present paper, which is mainly addressed to physicists, we discuss our main result, its physical implications, and rough ideas of the proof. Full mathematical details can be found[5] in the accompanying mathematical paper [17].

Recently there have been several related works on the time scales for equilibration in isolated quantum systems. See, for example, [18, 19, 20, 21, 22, 23].

2 Setting, background, and main result

Setting: Our setting is basically the same as in our previous work [16]. Following von Neumann [1, 2] and Goldstein, Lebowitz, Mastrodonato, Tumulka, and Zanghì [6], we consider an abstract model of an isolated macroscopic quantum system in a large volume. A typical example is a system of N particles confined in a box of volume V, where the density N/V is kept constant when V becomes large. In what follows we assume that the volume V is fixed, and do not discuss the V dependence of various quantities explicitly.

Let $\hat{H}$ be the Hamiltonian, and denote by E_α and $|\psi_\alpha\rangle$ the eigenvalue and the normalized eigenstate, respectively, of $\hat{H}$, i.e., $\hat{H}|\psi_\alpha\rangle = E_\alpha|\psi_\alpha\rangle$. We here take the microcanonical point of view, and focus on the energy interval $[U - \Delta U, U]$. Here U is an arbitrary energy that we shall fix, and ΔU denotes the range of energy, which is small from the macroscopic point of view, but is still large enough to contain many energy levels. We also assume $\Delta U \gg k_{\rm B}T$, where the temperature T is defined below the theorem. The choice of ΔU is somewhat arbitrary.

We relabel the index α so that the energy eigenvalues $E_\alpha \in [U - \Delta U, U]$ precisely correspond to the indices $\alpha = 1, \ldots, D$. We shall work with the Hilbert space $\mathcal{H}$ spanned by all $|\psi_\alpha\rangle$ with $\alpha = 1, \ldots, D$. The space $\mathcal{H}$ is often called the *microcanonical energy shell*.

Thermal equilibrium: Let us explain and formulate what we mean by thermal equilibrium. We shall here take an operational point of view, and characterize thermal

[5] The variational estimate (11) does not appear in the mathematical paper [17].

equilibrium in terms of observation of macroscopic observables. Let $\hat{M}_1, \ldots, \hat{M}_k$ be a family of macroscopic observables in which one is interested, and $\bar{M}_1, \ldots, \bar{M}_k$ be their values in the thermal equilibrium corresponding to the energy U.

We say that the system is in thermal equilibrium if, for any $j = 1 \ldots, k$, a measurement of $\hat{M}_j$ gives a value sufficiently close to the equilibrium value $\bar{M}_j$ with probability close to one[6]. Clearly the standard equilibrium ensembles satisfy this property provided that $\hat{M}_j$ exhibits small fluctuation (in the ensemble). But the system need not be described by an ensemble to be in equilibrium if we use our operational definition. There can be a pure state $|\varphi\rangle \in \mathcal{H}$ which does the complete job, i.e., for any j, a measurement of $\hat{M}_j$ in $|\varphi\rangle$ gives a value close to $\bar{M}_j$ with probability close to one. We say that a pure state $|\varphi\rangle$ with this property represents thermal equilibrium[7]. This is in contrast with other formulations of thermality that would demand that the system as a whole be described by a suitable density matrix, for example the canonical, or that the reduced density matrix for a small subsystem be canonical (while the system as a whole might be in a pure state).

Pure states representing thermal equilibrium are indeed abundant. It is believed and can be shown in many cases that an overwhelming majority of states in the microcanonical energy shell $\mathcal{H}$ represent thermal equilibrium [2, 6, 7, 11, 12, 13].

In the present approach, we further assume that pure states representing thermal equilibrium can be fully characterized by a subspace of $\mathcal{H}$. To be precise, we assume that the microcanonical energy shell $\mathcal{H}$ is decomposed into the equilibrium and the nonequilibrium subspaces as $\mathcal{H} = \mathcal{H}_{\rm eq} \oplus \mathcal{H}_{\rm neq}$. Any state $|\varphi\rangle$ which is close enough to $\mathcal{H}_{\rm eq}$ represents thermal equilibrium, and a state not close to $\mathcal{H}_{\rm eq}$ does not[8]. The fact that an overwhelming majority of states in $\mathcal{H}$ represent thermal equilibrium implies that the dimension d of $\mathcal{H}_{\rm neq}$ satisfies[9] $1 \leq d \ll D$.

Thermalization and the time scale: We take an arbitrary normalized initial state $|\varphi(0)\rangle \in \mathcal{H}$, which may be close to $\mathcal{H}_{\rm neq}$, and consider its time-evolution

$$|\varphi(t)\rangle = e^{-i\hat{H}t/\hbar}\,|\varphi(0)\rangle. \tag{1}$$

Let $\hat{P}_{\rm neq}$ be the projection onto $\mathcal{H}_{\rm neq}$. If it happens for some $\tau > 0$ that

$$\frac{1}{\tau}\int_0^\tau dt\,\langle\varphi(t)|\hat{P}_{\rm neq}|\varphi(t)\rangle \ll 1, \tag{2}$$

[6]In general one cannot simultaneously measure different $\hat{M}_j$'s. But it was argued by von Neumann [1] that by "rounding" the original observables one my convert them into a mutually commutable family of observables, which can be measured simultaneously. See also [2, 6].

[7]In [14, 15], the term "thermal pure quantum states" has been coined for pure states representing thermal equilibrium.

[8]When the (rounded) observables $\hat{M}_1, \ldots, \hat{M}_k$ commute with each other and also with the projection onto $\mathcal{H}$, then we choose $\mathcal{H}_{\rm eq}$ as the subspace spanned by simultaneous eigenstates of $\hat{M}_1, \ldots, \hat{M}_k$ with eigenvalues close to $\bar{M}_1, \ldots, \bar{M}_k$. See [1, 2, 6].

[9]More precisely, we assume $1 \leq d \leq \text{const}\, D\,(\log D)^{-4}$.

we can say that, within the interval $[0, \tau]$, the state $|\varphi(t)\rangle$ escapes from the vicinity of $\mathcal{H}_{\rm neq}$, or comes close to $\mathcal{H}_{\rm eq}$ and stays there for most time. Thus the system thermalizes within τ.

The subspace $\mathcal{H}_{\rm neq}$ (or $\mathcal{H}_{\rm eq}$) should be determined from the macroscopic observables $\hat{M}_1, \ldots, \hat{M}_k$ as we have discussed above. It is not an easy task, however, to characterize $\mathcal{H}_{\rm neq}$ in general or in a concrete setting. Nor is it easy to usefully estimate the relaxation time for any specific physical $\mathcal{H}_{\rm neq}$. We shall therefore take an abstract approach in which we regard $\mathcal{H}_{\rm neq}$ as a general d-dimensional subspace of $\mathcal{H}$, and try to elucidate the relation between $\mathcal{H}_{\rm neq}$ and the associated relaxation time.

Our finding in the previous work [16] is that, for certain choices of $\mathcal{H}_{\rm neq}$, the required time τ becomes as large as $d\,\hbar/\Delta U$ or as small as $\hbar/\Delta U$.

Main theorem: In the present paper we shall focus on the relaxation time associated with a generic subspace $\mathcal{H}_{\rm neq}$. For this purpose we choose the subspace in a completely random manner as follows. Let $\mathcal{U}(\mathcal{H})$ be the group of all unitary transformations on $\mathcal{H}$. For each $\hat{U} \in \mathcal{U}(\mathcal{H})$ and $j = 1, \ldots, d$, we define a normalized state $|\xi^{(j)}\rangle = \hat{U}|\psi_j\rangle$. We then consider the d-dimensional subspace $\mathcal{H}_{\rm rnd}$ spanned by $|\xi^{(1)}\rangle, \ldots, |\xi^{(d)}\rangle$, and the corresponding projection operator

$$\hat{P}_{\rm rnd} = \sum_{j=1}^{d} |\xi^{(j)}\rangle\langle\xi^{(j)}|. \tag{3}$$

By drawing $\hat{U} \in \mathcal{U}(\mathcal{H})$ according to the unique Haar measure on $\mathcal{U}(\mathcal{H})$, we can generate the d-dimensional subspace $\mathcal{H}_{\rm rnd}$ and the associated projection $\hat{P}_{\rm rnd}$ in a completely random manner. The random subspace $\mathcal{H}_{\rm rnd}$ is our probabilistic model for $\mathcal{H}_{\rm neq}$.

Then our main result, implying that relaxation typically can't take much longer than the Boltzmann time $\tau_{\rm B}$ (see below), is the following.

Theorem: Let $1 \le d \ll D$. With probability close to one, we have[10]

$$\frac{1}{\tau}\int_0^\tau dt\, \langle\varphi(t)|\hat{P}_{\rm rnd}|\varphi(t)\rangle \lesssim \frac{\tau_{\rm B}}{\tau}, \tag{4}$$

for *any* normalized $|\varphi(0)\rangle \in \mathcal{H}$ and any τ such that $0 < \tau \le \tau_{\rm max}$, where $\tau_{\rm max} := \tau_{\rm B}\, \min\{(D/d)^{1/4}, D^{1/6}\}$ is a large constant[11].

Note that the probability in the theorem is with respect to the random choice of the subspace $\mathcal{H}_{\rm rnd}$. Also note that we do not rely on the energy eigenstate thermalization hypothesis[12]. Mathematically precise assumptions and statements can be found in [17].

[10] Throughout the present paper, $A \sim B$ means that A/B is close to one. $A \approx B$ means the weaker relation that A/B is $O(1)$. Likewise, $A \lesssim B$ and $A \lessapprox B$ mean $A \le B'$ with $B' \sim B$ and $B' \approx B$, respectively.

[11] If $\tau/\tau_{\rm B}$ is much larger, the right-hand side of (4) should be replaced by d/D.

[12] One might say that the energy eigenstate thermalization hypothesis is automatically satisfied if we take the subspace randomly.

In (4), we have introduced the Boltzmann time $\tau_{\rm B} := h/(k_{\rm B}T)$, where the inverse temperature T is determined by the standard formula $(k_{\rm B}T)^{-1} = \partial \log \Omega(E)/\partial E|_{E=U}$ involving the number of states $\Omega(E)$. We assume that $\Omega(E)$ exhibits the normal behavior $\Omega(E) \approx \exp[V\,\sigma(E/V)]$, where the entropy density $\sigma(\epsilon)$ is an increasing function [24]. For our theorem to be valid, it is essential to take into account properly the behavior of $\Omega(E)$.

Let us compare our result with the recent work by Malabarba, García-Pintos, Linden, Farrelly, and Short [22], who studied the equilibration (not thermalization) in isolated quantum systems[13]. They proved, among other things, that most observables equilibrate quite rapidly. This result is quite similar to ours but there are some essential differences. The initial state needs to have a sufficiently broad energy distribution in [22], while it is arbitrary in our work. The observable is drawn randomly *after* fixing the initial state in [22], while the random subspace is fixed at the beginning here.

3 Discussions

Atypicality of realistic thermalization: We first note that the Boltzmann time is $\tau_{\rm B} \sim 1.6 \times 10^{-13}\,{\rm s}$ at $T \sim 300\,{\rm K}$. Then even for $\tau \approx 10^{-6}\,{\rm s}$, the right-hand side of (4) does not exceed 10^{-6}. This means that the system, starting from *any* initial state, equilibrates much before one micro-second.

Although the Boltzmann time $\tau_{\rm B}$ is an important time scale for many quantum phenomena, it is simply absurd that a macroscopic system always thermalizes so quickly. We also note that the bound (4) reflects no information about the size of the system. This is unphysical since, in general, the relaxation time should increase with the size of the system. Since the theory is rigorous, the only reasonable conclusion is that realistic nonequilibrium subspaces form exceptions to the bound of the theorem, or, in other words, the moderately slow decay observed in reality is not typical (if we assume random $\mathcal{H}_{\rm neq}$).

We note that one may also interpret our theorem as providing information about how quickly states vary in a macroscopic quantum system. We find that, with probability close to one, any state (including mixed states) in $\mathcal{H}_{\rm rnd}$ escapes from $\mathcal{H}_{\rm rnd}$ on the order of the Boltzmann time.

The atypicality of slow thermalization in reality may not be too surprising, especially after knowing about it. Given the energy shell $\mathcal{H}$, the nonequilibrium subspace $\mathcal{H}_{\rm neq}$, in reality, is determined not in a random manner, but through the values of macroscopic quantities that we use to characterize the system. Recall that many of the standard macroscopic quantities are expressed as a sum (or an integral) of locally conserved observables, and the Hamiltonian of a realistic system consists of more or less short-range interactions. This means that the corresponding subspace $\mathcal{H}_{\rm neq}$ and the projection

[13] In [22], equilibration means that the expectation values of selected projections are close to those in the equilibrium state. We note that our formulation is made to study only thermalization (in our sense).

operator $\hat{P}_{\rm neq}$ should be special. It is likely, for example, that the commutator $[\hat{H}, \hat{P}_{\rm neq}]$ is smaller for realistic subspaces compared with randomly chosen ones.

Towards a realistic nonequilibrium subspace: In order to fully understand the problem of the approach to thermal equilibrium in macroscopic quantum systems, it may be essential to characterize realistic nonequilibrium subspaces, and to investigate the accompanying time scale.

The escape from a single state may provide a hint for such considerations (see also [23]). Take an arbitrary state $|\xi\rangle \in \mathcal{H}$ such that $\langle\psi_\alpha|\xi\rangle$ is negligible (in a certain rough sense) unless α is such that $E_\alpha \in [\tilde{E}, \tilde{E} + \Delta E] \subset [U - \Delta U, U]$ for some energy $\tilde{E}$ and energy width ΔE. Then take an initial state $|\varphi(0)\rangle \in \mathcal{H}$ which is close to $|\xi\rangle$, and examine how quickly the state escapes from the vicinity of $|\xi\rangle$. As is well-known as one form of the "uncertainty relation between time and energy", the overlap $\left|\langle\xi|\varphi(t)\rangle\right|^2$ changes (and hence decays) in the time scale of order $\tau_{\rm esc} := h/\Delta E$.

It may be possible that a subspace $\mathcal{H}_{\rm neq}$ is also characterized by certain energy scale ΔE, which is similarly related to the associated relaxation time. This picture is true, at least for some examples, as we shall see now.

In the present setting of completely random $\mathcal{H}_{\rm neq}$, we see that each $|\xi^{(j)}\rangle$ is characterized by the width (in the above sense) $\Delta E \approx k_{\rm B}T$. This is because, in a macroscopic system, most of the energy eigenvalues E such that $U - \Delta U \le E \le U$ are found in the smaller range with $U - {\rm const}\, k_{\rm B}T \le E \le U$, where the constant is of $O(V^0)$. From this the conclusion that the escape time coincides with the Boltzmann time does indeed follow.

Recall that we made a reasonable assumption that $\Delta U \gg k_{\rm B}T$. In the unphysical case with $\Delta U \ll k_{\rm B}T$, our inequality (4) should be modified and the right-hand side becomes $h/(\Delta U\, \tau)$. This is again consistent with our picture since each $|\xi^{(j)}\rangle$ now has width ΔU, which corresponds to the escape time $\tau_{\rm esc} = h/\Delta U$.

Two examples of $\mathcal{H}_{\rm neq}$ in our previous work [16] are also consistent with the picture of escape time. In Theorem 1, where one finds extremely slow decay, we have $\Delta E = \Delta U/d$, which means $\tau_{\rm esc} = h\, d/\Delta U$. In Theorem 2, where one finds quick decay, we have $\Delta E \sim \Delta U$, which means $\tau_{\rm esc} \sim h/\Delta U$.

This observation suggests that a realistic nonequilibrium subspace $\mathcal{H}_{\rm neq}$, determined through macroscopic quantities that we use to characterize the system, is associated with a certain energy width ΔE. It is possible that the escape time $\tau_{\rm esc}$, which may take a reasonable value depending on the value of ΔE, essentially determines the relaxation time.

Of course it is also likely that the above picture of the escape from a single state is too naive or has only limited applicability. In some systems it may happen that $|\varphi(t)\rangle$ is "trapped" in the vicinity of $\mathcal{H}_{\rm neq}$ in a more intricate manner. In such a situation the relaxation time should also depend on the number of independent $|\xi^{(j)}\rangle$'s that the state $|\varphi(t)\rangle$ should go through.

It is certainly an interesting challenge to examine these pictures in interacting many-body quantum systems. For the moment we only have limited rigorous results in abstract

and artificial settings. In particular Theorem 1 of [16] treats examples with unphysically long relaxation time, and the theorem of the present paper shows that a typical $\mathcal{H}_{\rm neq}$ leads to unphysically short relaxation time. The reality should lie in between these two extreme theorems, remaining to be understood.

4 Sketch of proof

Since a mathematically complete proof can be found in [17], we here briefly describe the essential ideas.

Let $|\varphi(0)\rangle = \sum_{\alpha=1}^{D} c_\alpha |\psi_\alpha\rangle$ be an arbitrary normalized initial state. Then we write

$$\frac{1}{\tau}\int_0^\tau dt\, \langle\varphi(t)|\hat{P}_{\rm rnd}|\varphi(t)\rangle = \sum_{\alpha,\beta=1}^{D} (c_\alpha)^* Q_{\alpha\beta}\, c_\beta, \tag{5}$$

where the matrix $\mathsf{Q} = (Q_{\alpha\beta})$ is defined by

$$Q_{\alpha\beta} := S_{\alpha\beta}\, P_{\alpha\beta}, \tag{6}$$

with $P_{\alpha\beta} := \langle\psi_\alpha|\hat{P}_{\rm rnd}|\psi_\beta\rangle$, and

$$S_{\alpha\beta} := \frac{1}{\tau}\int_0^\tau dt\, e^{i(E_\alpha - E_\beta)t/\hbar}. \tag{7}$$

The matrix Q is the Hadamard product of the matrices $\mathsf{P} = (P_{\alpha\beta})$ and $\mathsf{S} = (S_{\alpha\beta})$. Note also that Q is hermitian since both P and S are hermitian.

Let $\mu_{\rm max}$ be the maximum eigenvalue of S. When τ is not too large, we have the following bound for $\mu_{\rm max}$.

Proposition 1: If $D \gg 1$ and $0 < \tau \le \tau_{\rm max}$, we have $\mu_{\rm max} \lesssim D\,\tau_{\rm B}/\tau$.

Let $\lambda_{\rm max}$ be the maximum eigenvalues of Q. The following proposition, which bounds $\lambda_{\rm max}$ by a bound on $\mu_{\rm max}$, is our main mathematical result.

Proposition 2 Let $1 \le d \ll D$. Suppose that there is a constant $\bar{\mu}_{\rm max}$ such that $\mu_{\rm max} \le \bar{\mu}_{\rm max}$ and $\bar{\mu}_{\rm max} \gg d$. Then we have $\lambda_{\rm max} \lesssim \bar{\mu}_{\rm max}/D$ with probability close to one.

When $0 < \tau \le \tau_{\rm max}$, Proposition 1 allows us to choose $\bar{\mu}_{\rm max} \simeq D\,\tau_{\rm B}/\tau$, which satisfies $\bar{\mu}_{\rm max} \gg d$. By noting that (5) implies

$$\frac{1}{\tau}\int_0^\tau dt\, \langle\varphi(t)|\hat{P}_{\rm rnd}|\varphi(t)\rangle \le \lambda_{\rm max}, \tag{8}$$

we get the desired theorem from Proposition 2.

On Proposition 1: Since the proof in [17] of the upper bound is rather technical, let us here give a simple variational argument (which indeed proves the corresponding *lower* bound) which sheds light on the nature of $\mu_{\max}$ and the corresponding eigenvector.

It is convenient to define the matrix $\tilde{\mathsf{S}} = (\tilde{S}_{\alpha\beta})$ by

$$\tilde{S}_{\alpha\beta} := \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} dt\, e^{i(E_\alpha - E_\beta)t/\hbar}. \tag{9}$$

The matrix $\tilde{\mathsf{S}}$ has exactly the same eigenvalues as S, and is real symmetric. Let us approximate the density of states as $\rho(E) \sim D\tilde{\beta}\, e^{\tilde{\beta}(E-U)}$, where $\tilde{\beta} = (k_{\rm B}T)^{-1}$ is the inverse temperature. We have extended the range of energy to $-\infty < E \le U$. The approximation is reasonable in a macroscopic system.

Since $\tilde{S}_{\alpha\beta}$ has oscillating signs, the eigenvector corresponding to $\mu_{\max}$ should have nonvanishing components concentrated in a small energy range with high density of states. This motivates us to choose a normalized trial state $\varphi_\alpha = f(E_\alpha)$ with $f(E) := A\, e^{\gamma(E-U)}$. The normalization constant is

$$A \sim \sqrt{(2\gamma + \tilde{\beta})/(D\tilde{\beta})}. \tag{10}$$

Then the expectation value is easily computed as

$$\begin{aligned}\sum_{\alpha,\beta=1}^{D} \varphi_\alpha \tilde{S}_{\alpha\beta}\varphi_\beta &\sim \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} dt \int_{-\infty}^{U} dE\, dE'\, \rho(E)\,\rho(E')\, f(E)\, f(E')\, e^{i(E-E')t/\hbar} \\ &= D\,\frac{\tilde{\beta}\hbar}{\tau}\,\frac{2\gamma+\tilde{\beta}}{\gamma+\tilde{\beta}}\, 2\arctan\Big[\frac{\tau}{2(\gamma+\tilde{\beta})\hbar}\Big],\end{aligned} \tag{11}$$

which becomes $D\,\tau_{\rm B}/\tau$ when $\tau/\hbar \gg \gamma \gg \tilde{\beta}$.

On Proposition 2: It can be shown that Q is positive semi-definite. Thus we have $\mathrm{Tr}[\mathsf{Q}^n] \ge (\lambda_{\max})^n$. Then note that

$$\begin{aligned}\overline{\mathrm{Tr}[\mathsf{Q}^n]} &= \sum_{\alpha_1,\ldots,\alpha_n=1}^{D} \overline{Q_{\alpha_1\alpha_2}\, Q_{\alpha_2\alpha_3}\, \cdots\, Q_{\alpha_n\alpha_1}} \\ &= \sum_{\alpha_1,\ldots,\alpha_n=1}^{D} S_{\alpha_1\alpha_2}\, \cdots\, S_{\alpha_n\alpha_1}\, \overline{P_{\alpha_1\alpha_2}\, \cdots\, P_{\alpha_n\alpha_1}},\end{aligned} \tag{12}$$

where the bar denotes the average over the random choice of $|\xi^{(j)}\rangle$'s. The definition (3) implies $P_{\alpha\beta} = \langle\psi_\alpha|\hat{P}_{\rm rnd}|\psi_\beta\rangle = \sum_{j=1}^{d}\langle\psi_\alpha|\xi^{(j)}\rangle\langle\xi^{(j)}|\psi_\beta\rangle = \sum_{j=1}^{d}\xi_\alpha^{(j)}(\xi_\beta^{(j)})^*$, where we wrote $\xi_\alpha^{(j)} := \langle\psi_\alpha|\xi^{(j)}\rangle$. Then the expectation in (12) is written as

$$\overline{P_{\alpha_1\alpha_2}\, P_{\alpha_2\alpha_3}\, \cdots\, P_{\alpha_n\alpha_1}} = \sum_{j_1,\ldots,j_n=1}^{d} \overline{\xi_{\alpha_1}^{(j_1)}\,(\xi_{\alpha_2}^{(j_1)})^*\, \cdots \xi_{\alpha_n}^{(j_n)}\,(\xi_{\alpha_1}^{(j_n)})^*}\,. \tag{13}$$

In a crude approximation we treat $|\xi^{(j_1)}\rangle, \ldots, |\xi^{(j_n)}\rangle$ as independent random vectors each satisfying the normalization condition $\overline{\xi_\alpha^{(j)}\,(\xi_{\alpha'}^{(j)})^*} = D^{-1}\,\delta_{\alpha,\alpha'}$. This seems reasonable provided that[14] $n \ll d \ll D$, and was partially justified by Weingarten [25].

Then the relevant expectation becomes

$$\overline{\xi_{\alpha_1}^{(j_1)}\,(\xi_{\alpha_2}^{(j_1)})^*\,\xi_{\alpha_2}^{(j_2)}\,\cdots\,\xi_{\alpha_n}^{(j_n)}\,(\xi_{\alpha_1}^{(j_n)})^*} \sim \frac{1}{D^n}\prod_{s=1}^{n}\delta_{\alpha_s,\alpha_{s+1}}, \tag{14}$$

where we identified α_{n+1} with α_1. Substituted into (13), this approximation yields

$$\overline{P_{\alpha_1\alpha_2}\,P_{\alpha_2\alpha_3}\,\cdots\,P_{\alpha_n\alpha_1}} \sim \Big(\frac{d}{D}\Big)^n\prod_{s=1}^{n}\delta_{\alpha_s,\alpha_{s+1}}. \tag{15}$$

Note that the factor $\prod_{s=1}^{n}\delta_{\alpha_s,\alpha_{s+1}}$ imposes the constraint that $\alpha_1,\ldots,\alpha_n$ must be all identical. Thus, recalling (12), the present approximation gives

$$\overline{\mathrm{Tr}[\mathsf{Q}^n]} \overset{?}{\approx} \Big(\frac{d}{D}\Big)^n\sum_{\alpha=1}^{D}(S_{\alpha\alpha})^n = D\Big(\frac{d}{D}\Big)^n. \tag{16}$$

where we noted that $S_{\alpha\alpha}=1$ by (7). This cannot be the main contribution as it is too small and, moreover, independent of τ.

We next set $j_1=\cdots=j_n=j$ in the expectation value $\overline{\xi_{\alpha_1}^{(j_1)}\,(\xi_{\alpha_2}^{(j_1)})^*\,\cdots\,\xi_{\alpha_n}^{(j_n)}\,(\xi_{\alpha_1}^{(j_n)})^*}$ to get

$$\overline{\xi_{\alpha_1}^{(j)}\,(\xi_{\alpha_2}^{(j)})^*\,\xi_{\alpha_2}^{(j)}\,(\xi_{\alpha_3}^{(j)})^*\,\cdots\,\xi_{\alpha_n}^{(j)}\,(\xi_{\alpha_1}^{(j)})^*} = \overline{|\xi_{\alpha_1}^{(j)}|^2\,|\xi_{\alpha_2}^{(j)}|^2\cdots|\xi_{\alpha_n}^{(j)}|^2} \sim \frac{1}{D^n}, \tag{17}$$

where we assumed for simplicity that all $\alpha_1,\ldots,\alpha_n$ are distinct, and used the fact that for any j the coefficients $\xi_\alpha^{(j)}$ (with $\alpha=1,\ldots,D$) of the random vector $|\xi^{(j)}\rangle$ can be treated as independent random variables. Assuming that (14) and (17) give the dominant contributions, we find from (13) that

$$\overline{P_{\alpha_1\alpha_2}\,P_{\alpha_2\alpha_3}\,\cdots\,P_{\alpha_n\alpha_1}} \approx \Big(\frac{d}{D}\Big)^n\prod_{s=1}^{n}\delta_{\alpha_s,\alpha_{s+1}} + \frac{d}{D^n}. \tag{18}$$

Note that the first term in the right-hand side, which is (15), is larger but has the constraint on the α's while the second term is smaller but is (almost) free from constraint. Going back to (12), the second term yields a new contribution

$$\sum_{\alpha_1,\ldots,\alpha_n=1}^{D} S_{\alpha_1\alpha_2}\,S_{\alpha_2\alpha_3}\,\cdots\,S_{\alpha_n\alpha_1}\,\frac{d}{D^n} = \mathrm{Tr}[\mathsf{S}^n]\,\frac{d}{D^n}. \tag{19}$$

[14] It is sufficient to prove Proposition 2 for large enough d, since $\lambda_{\max}$ is easily found to be non-decreasing in d [17].

Since S is easily found to be positive semi-definite, we have $\mathrm{Tr}[\mathsf{S}^n] \leq D\,(\mu_{\max})^n \leq D\,(\bar{\mu}_{\max})^n$. We thus arrive at the estimate

$$\frac{1}{D}\,\overline{(\lambda_{\max})^n} \lessapprox \left(\frac{d}{D}\right)^n + \left(\frac{\bar{\mu}_{\max}}{D}\right)^n d \sim \left(\frac{\bar{\mu}_{\max}}{D}\right)^n d. \tag{20}$$

By taking n sufficiently large, Proposition 2 then follows from the Markov inequality.

Since $|\xi^{(j)}\rangle$'s are in fact correlated with each other, it is a nontrivial task to make the above estimate into a rigorous one. Fortunately we can make use of the recent exact integration formula due to Collins [26, 27], which gives

$$\begin{aligned}&\overline{\xi^{(j_1)}_{\alpha_1}\,(\xi^{(j_1)}_{\alpha_2})^*\,\xi^{(j_2)}_{\alpha_2}\,(\xi^{(j_2)}_{\alpha_3})^*\,\cdots\,\xi^{(j_n)}_{\alpha_n}\,(\xi^{(j_n)}_{\alpha_1})^*}\\ &\qquad = \sum_{\sigma,\tau\in\mathfrak{S}_n} I[\forall k, \alpha_k = \alpha_{\sigma(k)+1},\, j_k = j_{\tau(k)}]\,\mathrm{Wg}(\tau\sigma^{-1}),\end{aligned} \tag{21}$$

where $\mathfrak{S}_n$ is the group of all permutations of $\{1, 2, \ldots, n\}$. The Weingarten-Collins function $\mathrm{Wg}(\sigma)$ behaves for large D as $|\mathrm{Wg}(\sigma)| \approx 1/D^{n+|\sigma|}$, where $|\sigma|$ denotes the minimum number of transpositions necessary to express the permutation σ as their products. By making use of the formula (21) and some properties of the group $\mathfrak{S}_n$, we can control all the terms which contribute to $\overline{\mathrm{Tr}[\mathsf{Q}^n]}$, and prove the theorem [17].

We wish to thank Marcus Cramer, Tetsuo Deguchi, Takaaki Monnai, Shin-ichi Sasa, Akira Shimizu, Ayumu Sugita, Sho Sugiura, and Yu Watanabe for valuable discussions and comments. We also thank Benoît Collins, Hiroyuki Ochiai and Tomoyuki Shirai for their help in mathematical issues.

The present work was supported in part by grant no. 37433 from the John Templeton Foundation (S.G.), JSPS Grants-in-Aid for Scientific Research nos. 25610021 (T.H.) and 25400407 (H.T.).

References

[1] J. von Neumann, *Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik*, Z. Phys. **57**, 30 (1929); English translation (by R. Tumulka), *Proof of the Ergodic Theorem and the H-Theorem in Quantum Mechanics*. `arXiv:1003.2133`

[2] S. Goldstein, J. L. Lebowitz, R. Tumulka, N. Zanghì, *Long-time behavior of macroscopic quantum systems*, European Phys. J. H **35**, 173–200 (2010). `arXiv:1003.2129`

[3] H. Tasaki, *From Quantum Dynamics to the Canonical Distribution: General Picture and a Rigorous Example*, Phys. Rev. Lett. **80**, 1373–1376 (1998). `arXiv:cond-mat/9707253`

[4] P. Reimann, *Foundation of Statistical Mechanics under Experimentally Realistic Conditions*, Phys. Rev. Lett. **101**, 190403 (2008).
`arXiv:0810.3092`

[5] N. Linden, S. Popescu, A. J. Short, A. Winter, *Quantum mechanical evolution towards thermal equilibrium*, Phys. Rev. E **79**, 061103 (2009).
`arXiv:0812.2385`

[6] S. Goldstein, J. L. Lebowitz, C. Mastrodonato, R. Tumulka, and N. Zanghì, *On the Approach to Thermal Equilibrium of Macroscopic Quantum Systems*, Phys. Rev. E **81**, 011109 (2010).
`arXiv:0911.1724`

[7] H. Tasaki, *The approach to thermal equilibrium and "thermodynamic normality" — An observation based on the works by Goldstein, Lebowitz, Mastrodonato, Tumulka, and Zanghìin 2009, and by von Neumann in 1929*, unpublished note (2013).
`arXiv:1003.5424`

[8] P. Reimann and M. Kastner, *Equilibration of isolated macroscopic quantum systems*, New J. Phys. **14**, 043020 (2012).
`http://iopscience.iop.org/1367-2630/14/4/043020`

[9] P. Reimann, *Equilibration of Isolated Macroscopic Quantum Systems under Experimentally Realistic Conditions*, Phys. Scr. **86**, 058512 (2012).
`arXiv:1210.5821`

[10] J. Sato, R. Kanamoto, E. Kaminishi, T. Deguchi, *Exact relaxation dynamics of a localized many-body state in the 1D bose gas*, Phys. Rev. Lett. **108**, 110401 (2012).
`arXiv:1112.4244`

[11] S. Popescu, A. J. Short, A. Winter, *Entanglement and the foundation of statistical mechanics*, Nature Phys. **2** (11), 754–758 (2006).

[12] S. Goldstein, J. L. Lebowitz, R. Tumulka, N. Zanghì, *Canonical Typicality*, Phys. Rev. Lett. **96**, 050403 (2006).
`arXiv:cond-mat/0511091`

[13] A. Sugita, *On the Basis of Quantum Statistical Mechanics*, Nonlinear Phenomena in Complex Systems **10**, 192–195 (2007).
`arXiv:cond-mat/0602625`

[14] S. Sugiura, A. Shimizu, *Thermal Pure Quantum States at Finite Temperature* , Phys. Rev. Lett **108**, 240401 (2012).
`arXiv:1112.0740`

[15] S. Sugiura, A. Shimizu, *Canonical Thermal Pure Quantum State*, Phys. Rev. Lett **111**, 010401 (2013).
`arXiv:1302.3138`

[16] S. Goldstein, T. Hara, and H. Tasaki, *Time Scales in the Approach to Equilibrium of Macroscopic Quantum Systems*, Phys. Rev. Lett. **111**, 140401 (2013).
`arXiv:1307.0572`

[17] S. Goldstein, T. Hara, and H. Tasaki, *The approach to equilibrium in a macroscopic quantum system for a typical nonequilibrium subspace*, preprint (2014).
`arXiv:1402.3380`

[18] Vinayak and M. Znidaric, *Subsystem dynamics under random Hamiltonian evolution*, J. Phys. A: Math. Theor. **45**, 125204 (2012).
`arXiv:1107.6035`

[19] L. Masanes, A. J. Roncaglia, and A. Acin, *The complexity of energy eigenstates as a mechanism for equilibration*, Phys. Rev. E **87**, 032137 (2013).
`arXiv:1108.0374`

[20] F.G.S.L. Brandão, P. Ćwikliński, M. Horodecki, P. Horodecki, J. Korbicz, and M. Mozrzymas, *Convergence to equilibrium under a random Hamiltonian*, Phys. Rev. E **86**, 031101 (2012).
`arXiv:1108.2985`

[21] M. Cramer, *Thermalization under randomized local Hamiltonians*, New J. Phys. **14**, 053051 (2012).
`http://iopscience.iop.org/1367-2630/14/5/053051`

[22] A.S.L. Malabarba, L.P. García-Pintos, N. Linden, T.C. Farrelly, and A.J. Short, *Quantum Systems Equilibrate Rapidly for Most Observables*, Phys. Rev. E **90**, 012121 (2014).
`arXiv:1402.1093`

[23] T. Monnai, *General relaxation time of the fidelity for isolated quantum thermodynamic systems*, J. Phys. Soc. Jpn., **83**, 064001 (2014).
`arXiv:1403.6578`

[24] D. Ruelle, *Statistical Mechanics: Rigorous Results*, (World Scientific, 1999).

[25] D. Weingarten, *Asymptotic behavior of group integrals in the limit of infinite rank*, J. Math. Phys. **19**, 999–1001 (1978).

[26] B. Collins, *Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability*, IMRN, International Mathematics Research Notices, **17**, 953–982 (2003).
`arXiv:math-ph/0205010`

[27] B. Collins and P. Śniady, *Integration with respect to the haar measure on unitary, orthogonal and symplectic group*, Commun. Math. Phys. **264**, 773–795 (2006).
`arXiv:math-ph/0402073`