Equilibration of quantum systems and subsystems

Anthony J Short
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
E-mail: a.j.short@damtp.cam.ac.uk

New Journal of Physics 13 (2011) 053009 (10pp)
Received 14 February 2011
Published 4 May 2011
Online at http://www.njp.org/
doi:10.1088/1367-2630/13/5/053009

Abstract. We unify two recent results concerning equilibration in quantum theory. We first generalize a proof of Reimann (2008 Phys. Rev. Lett. 101 190403), that the expectation value of ‘realistic’ quantum observables will equilibrate under very general conditions, and discuss its implications for the equilibration of quantum systems. We then use this to re-derive an independent result of Linden et al (2009 Phys. Rev. E 79 061103), showing that small subsystems generically evolve to an approximately static equilibrium state. Finally, we consider subspaces in which all initial states effectively equilibrate to the same state.

Contents

1. Introduction 1
2. Equilibration of expectation values 2
3. Distinguishability 4
4. Effective equilibration of large systems 6
5. Equilibration of small subsystems 7
6. Universality of equilibrium states 8
7. Conclusions 9
Acknowledgments 9
References 9

1. Introduction

Recently there has been significant progress in understanding the foundations of statistical mechanics, based on fundamentally quantum arguments [1]–[12]. In particular, Reimann [1, 2]
has shown that the expectation value of any ‘realistic’ quantum observable will equilibrate to an approximately static value, given very weak assumptions about the Hamiltonian and initial state. Interestingly, the same assumptions were used independently by Linden et al [4, 5], to prove that any small quantum subsystem will evolve to an approximately static equilibrium state (such that even ‘unrealistic’ observables on the subsystem equilibrate). In this paper, we unify these two results, by deriving the central result of Linden et al [4] from a generalization of Reimann’s result. We also offer a further discussion and extension of Reimann’s results, showing that systems will appear to equilibrate with respect to all reasonable experimental capabilities. Finally, we identify subspaces of initial states which equilibrate to the same state.

2. Equilibration of expectation values

We prove below a generalization of Reimann’s result that the expectation value of any operator will almost always be close to that of the equilibrium state [1]. We extend his results to include non-Hermitian operators (which we will need in section 5 to prove equilibration of subsystems), correct a subtle mistake made in [2] when considering degenerate Hamiltonians, and improve the bound obtained by a factor of 4. As in [2, 5], we make one assumption in the proof, which is that the Hamiltonian has non-degenerate energy gaps. This means that given any four energy eigenvalues \( E_k, E_l, E_m, E_n \),

\[
E_k - E_l = E_m - E_n \Rightarrow \begin{cases} E_k = E_l \text{ and } E_m = E_n, \\ E_k = E_m \text{ and } E_l = E_n. \end{cases}
\]  

(1)

Note that this definition allows degenerate energy levels, which may arise due to symmetries. However, it ensures that all subsystems physically interact with each other. In particular, given any decomposition of the system into two subsystems \( H = H_A \otimes H_B \), equation (1) will not be satisfied by any Hamiltonian of the form \( H = H_A \otimes I_B + I_A \otimes H_B \) (unless either \( H_A \) or \( H_B \) is proportional to the identity). To see this, consider the four energy eigenstates \( |k\rangle = |0\rangle|0\rangle, |l\rangle = |0\rangle|1\rangle, |m\rangle = |1\rangle|0\rangle, |n\rangle = |1\rangle|1\rangle \), which are products of eigenstates of \( H_A \) and \( H_B \).

**Theorem 1** (Generalization of Reimann’s result [1]). Consider a d-dimensional quantum system evolving under a Hamiltonian \( H = \sum_n E_n P_n \), where \( P_n \) is the projector onto the eigenspace with energy \( E_n \). Denote the system’s density operator by \( \rho(t) \), and its time-averaged state by \( \omega \equiv \langle \rho(t) \rangle_t \). If \( H \) has non-degenerate energy gaps, then for any operator \( A \),

\[
\sigma_A^2 \equiv \langle |\text{tr}(A \rho(t)) - \text{tr}(A \omega)|^2 \rangle_t \leq \frac{\Delta(A)^2}{4d_{\text{eff}}} \leq \frac{\|A\|^2_{\text{eff}}}{d_{\text{eff}}}
\]  

(2)

where \( \|A\| \) is the standard operator norm\(^1\),

\[
\Delta(A) \equiv 2 \min_{c \in \mathbb{C}} \| A - cI \|,
\]  

(3)

and

\[
d_{\text{eff}} \equiv \frac{1}{\sum_n (\text{tr}(P_n \rho(0)))^2}.
\]  

(4)

\(^1\|A\| = \sup_{\langle A\rangle = 1} \sqrt{\langle \langle v | A | v \rangle : |v\rangle \in \mathcal{H} \text{ with } \langle v | v \rangle = 1} \), or equivalently \( \|A\| \) is the largest singular value of \( A \).
This bound will be most significant when the number of different energies incorporated in the state, characterized by the effective dimension $d_{\text{eff}}$, is very large. Note that $1 \leq d_{\text{eff}} \leq d$, and that $d_{\text{eff}} = N$ when a measurement of $H$ would yield $N$ different energies with equal probability. For pure states $d_{\text{eff}} = \text{tr}(\sigma^2)^{-1}$ as in [4, 5], but it may be smaller for mixed states when the Hamiltonian is degenerate.

The quantity $\Delta(A)$ gives the range of eigenvalues when $A$ is Hermitian, and gives a slightly tighter bound than the operator norm. Following [2], we could improve the bound further by replacing $\Delta(A)$ with a state- and Hamiltonian-dependent term\(^2\); however, we omit this step here for simplicity.

**Proof.** To avoid some difficulties, which arise when considering degenerate Hamiltonians, we initially consider a pure state $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$, then extend the results to mixed states via purification.

We can always choose an energy eigenbasis such that $|\psi(t)\rangle$ has non-zero overlap with only a single energy eigenstate $|n\rangle$ of each distinct energy, by including states $|n\rangle = P_n|\psi(0)\rangle/\sqrt{\langle\psi(0)|P_n|\psi(0)\rangle}$ whenever $\langle\psi(0)|P_n|\psi(0)\rangle > 0$. The state at time $t$ is then given by

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |n\rangle,$$

where $c_n = \langle n|\psi(0)\rangle$. This state will evolve in the subspace spanned by $\{|n\rangle\}$ as if it were acted on by the non-degenerate Hamiltonian $H' = \sum_n E_n |n\rangle\langle n|$.

For any operator $A$, it follows that $\sigma_A^2 = \left\langle |\text{tr}(A\rho(t) - \omega)|^2 \right\rangle_t$

$$= \left\langle \left| \sum_{n\neq m} c_n^* c_m e^{i(E_m - E_n) t/\hbar} \langle m|A|n\rangle \right|^2 \right\rangle_t$$

$$= \sum_{n\neq m} c_n^* c_m \left| e^{i(E_m - E_n) t/\hbar} \langle m|A|n\rangle \right|^2$$

$$= \sum_{n,m} |c_n|^2 |c_m|^2 \langle m|A|n\rangle\langle n|A^\dagger|m\rangle - \sum_n |c_n|^4 \langle n|A|n\rangle^2$$

$$\leq \text{tr}(AA^\dagger) \text{tr}(\omega^2)$$

$$\leq \|A\|^2 \text{tr}(\omega^2)$$

$$= \|A\|^2 \text{tr} \left( \left( \sum_n |c_n|^2 \langle n\rangle \right)^2 \right)$$

$$= \|A\|^2 \sum_n (\text{tr}(P_n\rho(0)))^2$$

$$= \frac{\|A\|^2}{d_{\text{eff}}}.$$

\(^2\) In particular, we could replace $\Delta(A)$ with $\Delta''(A) = \min_A 2\|\tilde{A}\|$, where the operators $\tilde{A}$ are obtained by subtracting any function of $H$ from $A$ and projecting onto the support of $\omega$.

*New Journal of Physics* **13** (2011) 053009 (http://www.njp.org/)
In the fourth line, we have used the assumption that the Hamiltonian has non-degenerate energy gaps, in the sixth line we have used the Cauchy–Schwartz inequality for operators with scalar product $\text{tr}(A^*B)$ and the cyclic symmetry of the trace, and in the seventh line we have used the fact that for positive operators $P$ and $Q$, $\text{tr}(PQ) \leq ||P||\text{tr}(Q)$. This gives the weaker bound in the theorem.

To obtain the tighter bound, we note that $\sigma_A$ is invariant if $A$ is replaced by $\tilde{A} = A - cI$ for any complex $c$. Performing this substitution with $c$ chosen so as to minimize $||\tilde{A}||$ we can replace $||A||$ with $||\tilde{A}|| = \Delta(A)/2$.

An extension to mixed states can be obtained via purification, following the approach discussed in [5]. Given any initial state $\rho(0)$ on $\mathcal{H}$, we can always define a pure state $|\phi(0)\rangle$ on $\mathcal{H} \otimes \mathcal{H}$ such that the reduced state of the first system is $\rho(0)$. By evolving $|\phi(t)\rangle$ under the joint Hamiltonian $H' = H \otimes I$, we will recover the correct evolution $\rho(t)$ of the first system, and $H'$ will have non-degenerate energy gaps whenever $H$ does. The expectation value of any operator $A$ for $\rho(t)$ will be the same as the expectation value of $A' = A \otimes I$ on the total system, and we also obtain $\Delta(A') = \Delta(A)$, $||A|| = ||A'||$, and $d'_{\text{eff}} = d_{\text{eff}}$. However, note that $\text{tr}(\omega^2)$ does not equal $\text{tr}(\omega^2)$. Using the result for pure states, we can obtain (2) in the mixed state case from

$$\sigma_A^2 = \sigma_A' \leq \frac{\Delta(A)^2}{4d'_{\text{eff}}} = \frac{\Delta(A)^2}{4d_{\text{eff}}}.$$  \hfill (7)

This completes the proof. \hfill $\blacksquare$

In [1], Reimann proves that $\sigma_A^2 \leq \Delta(A)^2 \text{tr}(\omega^2)$ when $A$ is Hermitian and the Hamiltonian has non-degenerate levels as well as non-degenerate gaps. However, it appears that there is a subtle mistake in [2] when extending this proof to degenerate Hamiltonians. Specifically, the step from equation (D.11) to (D.12) in [2] does not follow if the state has support on more than one eigenstate in a degenerate subspace. A counterexample is provided by the mixed state $\rho(0) = \frac{1}{2}|0\rangle\langle 0| \otimes I$, of a qubit and a $k$-dimensional system, with $H = (|0\rangle\langle 1| + |1\rangle\langle 0|) \otimes I$ and $A = (|0\rangle\langle 0| - |1\rangle\langle 1|) \otimes I$. In this case $\sigma_A^2 = \frac{1}{2}$, $\Delta(A) = 2$ and $\text{tr}(\omega^2) = \frac{1}{2k}$, giving $\sigma_A^2 > \Delta(A)^2 \text{tr}(\omega^2)$ when $k > 4$. However, subsequently in [2], $\text{tr}(\omega^2)$ is replaced by an upper bound of $\max_{\nu} \text{tr}(\rho(0)P_{\nu})$, and this also upper bounds $d'_{\text{eff}}$, so later results are unaffected. Note that the bound given by Theorem 1 for the same example is satisfied tightly for all $k$, as $d_{\text{eff}} = 2$ and thus $\sigma_A^2 = \frac{1}{2} = \frac{\Delta(A)^2}{4d_{\text{eff}}}$. 

3. Distinguishability

When $A$ represents a realistic physical observable and $\rho(0)$ a realistic initial state, it is argued in [1] that the difference between $\text{tr}(A\rho(t))$ and $\text{tr}(A\delta)$ will almost always be less than realistic experimental precision. The fact that this holds for all realistic observables suggests that $\rho(t)$ will be indistinguishable from $\omega$ for the overwhelming majority of times.

However, this issue is somewhat subtle, as the fact that two states yield the same expectation value for a measurement does not necessarily imply that they cannot be distinguished by it. For example, a measurement yielding an equal mixture of $+1$ and $-1$ outcomes for one state and always yielding $0$ for a second state clearly can distinguish the two states, despite the expectation values in the two cases being identical. We could rule out this example by considering the expectation value of $A^2$, which must also equilibrate,
but it is not clear how to derive a good bound on the distinguishability in general using this approach.

Furthermore, even though any particular realistic measurement cannot distinguish $\rho(t)$ from $\omega$ for almost all times, this does not imply that for almost all times, no realistic measurement can distinguish $\rho(t)$ from $\omega$. This is because the optimal measurement to distinguish the two states may change over time. This issue is discussed in a slightly different context in [3].

Finally, the measurement precision is not easy to define for measurements with discrete outcomes, which may not even have numerical labels.

To address these issues, we first note that the most general quantum measurement is not described by a Hermitian operator, but by a positive operator valued measure (POVM). For simplicity, we consider POVMs with a finite set of outcomes, which is reasonable for realistic measurements, as even continuous outputs such as pointer position cannot be determined or recorded with infinite precision. For example, the measurement of a continuous observable $A$ with precision $\delta A$ can be modeled by a discrete measurement with $\Delta(A)/\delta A$ outcomes. Alternatively, our results can be extended to continuous output sets using measure theory.

A general measurement $M$ is described by giving a positive operator $M_r$ for each possible measurement result $r$, satisfying $\sum_r M_r = I$. The probability of obtaining result $r$ when measuring $M$ on $\rho$ is given by $\text{tr}(M_r \rho)$.

Suppose you are given an unknown quantum state, which is either $\rho_1$ or $\rho_2$ with equal probability. Your maximum success probability in guessing which state you were given after performing the measurement $M$ is

$$p_{\text{succ}}^M = \frac{1}{2} (1 + D_M(\rho_1, \rho_2)), \quad (8)$$

where

$$D_M(\rho_1, \rho_2) \equiv \frac{1}{2} \sum_r |\text{tr}(M_r \rho_1) - \text{tr}(M_r \rho_2)|. \quad (9)$$

We refer to $D_M(\rho_1, \rho_2)$ as the distinguishability of $\rho_1$ and $\rho_2$ using the measurement $M$.

Similarly, the optimal probability of guessing which state you have using any single measurement from a set $\mathcal{M}$ is

$$p_{\text{succ}}^{\mathcal{M}} = \frac{1}{2} (1 + D_{\mathcal{M}}(\rho_1, \rho_2)), \quad (10)$$

where

$$D_{\mathcal{M}}(\rho_1, \rho_2) \equiv \max_{M \in \mathcal{M}} D_M(\rho_1, \rho_2). \quad (11)$$

$D_{\mathcal{M}}(\rho_1, \rho_2)$ expresses the distinguishability of the states given the set of measurements $\mathcal{M}$. Note that any sequence of measurements is itself a measurement, and can be included as an element of $\mathcal{M}$ if desired. Furthermore,

$$0 \leq D_{\mathcal{M}}(\rho_1, \rho_2) \leq D(\rho_1, \rho_2) \leq 1, \quad (12)$$

where $D(\rho_1, \rho_2) \equiv \frac{1}{2} \text{tr}|\rho_1 - \rho_2|$ is the trace-distance between $\rho_1$ and $\rho_2$. This is because the trace-distance is equal to the distinguishability of $\rho_1$ and $\rho_2$ using the optimal measurement.

---

3 Defining $p_r = \text{tr}(M_r \rho_1)$ and $q_r = \text{tr}(M_r \rho_2)$, the optimal strategy given result $r$ is to guess $\rho_1$ if $p_r \geq q_r$ and $\rho_2$ otherwise. This gives $p_{\text{succ}}^M = \frac{1}{2} \sum_r \max(p_r, q_r) = \frac{1}{4} \sum_r (p_r + q_r + |p_r - q_r|) = \frac{1}{2} (1 + D_M(\rho_1, \rho_2))$.
which is a two-outcome projective measurement onto the positive and negative eigenspaces of $(\rho_1 - \rho_2)$ \[13\]. Applying (9) to this measurement we obtain the definition of the trace distance given. Hence the trace-distance is equal to $D_M(\rho_1, \rho_2)$ when $M$ includes all measurements.

4. Effective equilibration of large systems

For typical macroscopic systems, the dimension of $\mathcal{H}$ will be incredibly large (e.g. for Avagardo’s number $N_A$ of spin-1/2 particles, we would have $d > 10^{10^{23}}$), and it is unrealistic to be able to perform any measurement with this many outcomes, let alone all such measurements. For practical purposes, we are therefore restricted to some set of realistic physical measurements $M$. In this case, we would expect $M$ to be a finite set, as all realistic experimental setups (including all settings of variable parameters) will be describable within a finite number of pages of text.

We say that a state effectively equilibrates if
\[
(D_M(\rho(t), \omega))_t \ll 1.
\]
This means that for almost all times, it is almost impossible to distinguish the true state $\rho(t)$ from the equilibrium state $\omega$ using any achievable measurement.

We can obtain an upper bound on the average distinguishability as a corollary of theorem 1.

**Corollary 1.** Consider a quantum system evolving under a Hamiltonian with non-degenerate energy gaps. The average distinguishability of the system’s state $\rho(t)$ from $\omega$, given a finite set of measurements $M$, satisfies
\[
(D_M(\rho(t), \omega))_t \leq \frac{\sum_{M \in M} \sum_r \Delta(M_r)}{4\sqrt{d_{\text{eff}}}} \leq \frac{N(M)}{4\sqrt{d_{\text{eff}}}},
\]
where $N(M)$ is the total number of outcomes for all measurements in $M$.

The first bound will be tighter when measurements are imprecise, as each outcome is weighted by $\Delta(M_r) \in [0, 1]$, reflecting its usefulness in distinguishing states (in particular, $\Delta(M_r)$ is the maximum difference in probability of result $r$ occurring for any two states). Note that for any measurement on a $d_S$-dimensional subsystem\(^4\) $\sum_r \Delta(M_r) \leq d_S$.

**Proof.**
\[
(D_M(\rho(t), \omega))_t = \left( \max_{M(\cdot) \in M} D_{M(\cdot)}(\rho(t), \omega) \right)_t \\
\leq \sum_{M \in M} (D_M(\rho(t), \omega))_t \\
= \frac{1}{2} \sum_{M \in M} \sum_r \langle |\text{tr}(M_r \rho(t)) - \text{tr}(M_r \omega)| \rangle_t \\
\leq \frac{1}{2} \sum_{M \in M} \sum_r \sqrt{\langle |\text{tr}(M_r \rho(t)) - \text{tr}(M_r \omega)|^2 \rangle_t},
\]
\(^4\) Write $M_r = M_r^S \otimes I$, where the first subsystem is the one on which the measurement acts. Then note that $\sum_r \Delta(M_r) = \sum_r \Delta(M_r^S) \leq \sum_r \text{tr}_S(M_r^S) = \text{tr}_S(I) = d_S$.
\[
\sum_{M \in \mathcal{M}} \sum_{r} \sqrt{\sigma_{M,r}^{2}} \leq \sum_{M \in \mathcal{M}} \sum_{r} \Delta(M_r) \quad \text{and} \quad \Delta(\mathcal{M}) \leq \frac{N(\mathcal{M})}{4\sqrt{d_{\text{eff}}}}. \tag{15}
\]

In realistic experiments, we would expect the bound on the right of (14) to be much smaller than 1, implying that the state of the system effectively equilibrates to \(\omega\). Consider again our system of \(N_A\) spins. If \(d_{\text{eff}} \geq d^{0.1}\), even if we take \(\mathcal{M}\) to include any experiment whose description could be written in \(10^{19}\) words, each of which generates up to \(10^{21}\) bytes of data, we would still obtain

\[
\langle D_{\mathcal{M}}(\rho(t), \omega) \rangle_t \leq 1/(10^{102}).
\]

5. Equilibration of small subsystems

Now consider that the system can be decomposed into two parts, a small subsystem of interest \(S\), and the remainder of the system which we refer to as the bath \(B\). Then \(\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B\), where \(\mathcal{H}_{S/B}\) has dimension \(d_{S/B}\). It is helpful to define the reduced states of the subsystem \(\rho_S(t) = \text{tr}_B(\rho(t))\) and \(\omega_S = \text{tr}_B(\omega)\).

In such cases, it was shown in [4, 5] that for sufficiently large \(d_{\text{eff}}\) the subsystem’s state fully equilibrates, such that for the vast majority of times, \(\rho(t)\) and \(\omega\) are almost impossible to distinguish using any measurement on the subsystem (even ‘unrealistic’ ones). In particular, when \(\rho(t)\) is pure and the Hamiltonian has non-degenerate energy levels as well as non-degenerate energy gaps, it is proven in [4] that

\[
\langle D(\rho_S(t), \omega_S) \rangle_t \leq \frac{1}{2} \frac{d_S^2}{d_{\text{eff}}}. \tag{16}
\]

Extending this result to degenerate Hamiltonians and initially mixed states is discussed in [5].

We cannot recover this bound directly from (14) by considering the set of all measurements on the subsystem, because this set contains an infinite number of measurements. However, we can derive (16) from theorem 1 by considering an orthonormal operator basis for the subsystem, given by the \(d_S^2\) operators [14]

\[
F_{(d_S k_0 + k_1)} = \frac{1}{\sqrt{d_S}} \sum_l \epsilon^{\frac{2\pi|l|}{d_S}} \langle l | (d_S k_0 + k_1) \mod d_S \rangle \langle l |,
\]

where \(k_0, k_1 \in \{0, 1, \ldots, d_S - 1\}\) and the states \(|l\rangle\) are an arbitrary orthonormal basis for the subsystem. Then writing \((\rho_S(t) - \omega_S) = \sum_k \lambda_k(t) F_k\) we have

\[
\langle D(\rho_S(t), \omega_S) \rangle_t \leq \frac{1}{2} \left( \text{tr} \left( \sum_k \lambda_k(t) F_k \right) \right)_t \leq \frac{1}{2} \left( \text{tr} \left( \sum_{kl} \lambda_k(t) \lambda^*_k(t) F_k^\dagger F_k \right) \right)_t.
\]

New Journal of Physics 13 (2011) 053009 (http://www.njp.org/)
\[ \leq \frac{1}{2} \sqrt{d_S \sum_{kl} \langle \lambda_k(t) \lambda_l^*(t) \rangle_t \text{tr}(F_k^\dagger F_k)} \]

\[ = \frac{1}{2} \sqrt{d_S \sum_k \langle |\lambda_k(t)|^2 \rangle_t} \]

\[ = \frac{1}{2} \sqrt{d_S \sum_k \langle |\text{tr}((\rho(t) - \omega) F_k^\dagger \otimes I)|^2 \rangle_t} \]

\[ \leq \frac{1}{2} \sqrt{d_S \sum_k \frac{\| F_k^\dagger \otimes I \|^2}{d_{\text{eff}}}} \]

\[ \leq \frac{1}{2} \sqrt{\frac{d_{\text{eff}}^2}{d_{\text{eff}}}}. \] (18)

In the second line, we have used a standard relation between the 1- and 2-norm, and in the sixth line we have used theorem 1 for the non-Hermitian operator \( F_k^\dagger \otimes I \). Note that \( \sqrt{d_S F_k} \) is unitary, and thus \( \| F_k^\dagger \otimes I \| = \frac{1}{\sqrt{d_S}} \).

6. Universality of equilibrium states

We have so far been concerned with when states equilibrate, rather than the nature of their equilibrium state. However, one of the notable properties of equilibration is that many initial states effectively equilibrate to the same state, determined only by macroscopic properties such as temperature. Given a particular Hamiltonian and a set of realistic measurements \( \mathcal{M} \), we can construct a partition of the Hilbert space into a direct sum of subspaces \( \mathcal{H} = \bigoplus k \mathcal{H}_k \), such that all states within \( \mathcal{H}_k \) with large enough \( d_{\text{eff}} \) effectively equilibrate to the same state \( \Omega_k \).

One way to achieve this is to choose the subspaces such that each projector \( \Pi_k \) onto \( \mathcal{H}_k \) commutes with the Hamiltonian, and such that any two energy eigenstates in \( \mathcal{H}_k \) are hard to distinguish, i.e. for some fixed \( \epsilon \) satisfying \( 0 < \epsilon \ll 1 \), and all normalized energy eigenstates \( |i\rangle, |j\rangle \in \mathcal{H}_k \)

\[ D_M(\langle i|i\rangle, \langle j|j\rangle) \leq \epsilon. \] (19)

When \( d_{\text{eff}} \) is sufficiently large, it follows that all states in \( \mathcal{H}_k \) effectively equilibrate to \( \Omega_k = \Pi_k/\text{tr}(\Pi_k) \), as

\[ \langle D_M(\rho(t), \Omega_k) \rangle_t \leq \langle D_M(\rho(t), \omega) \rangle_t + \langle D_M(\omega, \Omega_k) \rangle_t \]

\[ \leq \frac{N(\mathcal{M})}{4\sqrt{d_{\text{eff}}}} + \sum_{i,j} \frac{(i|\omega|i)}{\text{tr}(\Pi_k)} D_M(|i\rangle\langle i|, |j\rangle\langle j|) \]

\[ \leq \frac{N(\mathcal{M})}{4\sqrt{d_{\text{eff}}}} + \epsilon. \] (20)
where the sums in the second line are over an eigenbasis of $\omega$ (which is also a basis of $\mathcal{H}_k$), and we have used the fact that $D_M(\rho, \sigma)$ satisfies the triangle inequality ($D_M(\rho, \sigma) \leq D_M(\rho, \tau) + D_M(\tau, \sigma)$) and convexity,

$$D_M\left(\sum_i p_i \rho_i, \sigma\right) \leq \sum_i p_i D_M(\rho_i, \sigma),$$  \hspace{1cm} (21)

where $p_i \geq 0$ and $\sum_i p_i = 1$.

When $\mathcal{H}_k$ can be chosen to be a small band of energies, the equilibrium state $\Omega_k$ will be the usual microcanonical state.

7. Conclusions

To summarize, we have shown that two key results of [1, 4] about the equilibration of large systems can be derived from very weak assumptions (non-degenerate energy gaps, and sufficiently large $d_{	ext{eff}}$) and a single theorem (theorem 1). In particular, for the vast majority of times, the state of an isolated quantum system will be almost indistinguishable from its equilibrium state $\omega$ using any realistic experiment, and the state of a small subsystem will be almost indistinguishable from $\omega_S$ using any experiment.

Although the first result has a similar flavor to the classical equilibration of coarse-grained observables such as density and pressure, it is really much stronger, as it encompasses any measurement you could describe and record the data from in a reasonable length of text, including microscopic measurements. The second result has no classical analogue, as it yields an essentially static description of the true micro-state of a subsystem, rather than the rapidly fluctuating dynamical equilibrium of particles in classical statistical mechanics. Given the difficulty of proving similar results in the classical case, it seems that quantum theory offers a firmer foundation for statistical mechanics.

Acknowledgments

The author thanks Peter Reimann and Sabri Al-Safi for helpful discussions and the Royal Society for their support.

References

[1] Reimann P 2008 Phys. Rev. Lett. 101 190403
[2] Reimann P 2010 New J. Phys. 12 055027
[3] Reimann P 2007 Phys. Rev. Lett. 99 160404
[4] Linden N, Popescu S, Short A and Winter A 2009 Phys. Rev. E 79 061103
[5] Linden N, Popescu S, Short A and Winter A 2010 New J. Phys. 12 055021
[6] Gemmer J, Michel M and Mahler G 2004 Quantum Thermodynamics (Lecture Notes in Physics vol 657) (Berlin: Springer)
[7] Goldstein S, Lebowitz J L, Tumulka R and Zanghì N 2006 Phys. Rev. Lett. 96 050403
[8] Goldstein S, Lebowitz J L, Mastrodonato C, Tumulka R and Zanghì N 2010 Phys. Rev. E 81 011109
[9] Popescu S, Short A J and Winter A 2006 Nat. Phys. 2 754–8
[10] Tasaki H 1998 *Phys. Rev. Lett.* **80** 1373–6
[11] Gogolin C 2010 arXiv:1003.5058
[12] Gogolin C 2010 *Phys. Rev. E* **81** 051127
[13] Nielsen M and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press) p 405
[14] Schwinger J 1960 *Proc. Natl Acad. Sci. USA* **46** 570