Eur. Phys. J. B **31**, 249–257 (2003)
DOI: 10.1140/epjb/e2003-00029-3

THE EUROPEAN PHYSICAL JOURNAL B

Distribution of local entropy in the Hilbert space of bi-partite quantum systems: origin of Jaynes' principle

J. Gemmer[a] and G. Mahler

Institut für Theoretische Physik, Universität Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany

Received 13 June 2002 / Received in final form 14 November 2002

Abstract. For a closed bi-partite quantum system partitioned into system proper and environment we interpret the microcanonical and the canonical condition as constraints for the interaction between those two subsystems. In both cases the possible pure-state trajectories are confined to certain regions in Hilbert space. We show that in a properly defined thermodynamical limit almost all states within those accessible regions represent states of some maximum local entropy. For the microcanonical condition this dominant state still depends on the initial state; for the canonical condition it coincides with that defined by Jaynes' principle. It is these states which thermodynamical systems should generically evolve into.

PACS. 03.65.-w Quantum mechanics – 05.70.-a Thermodynamics – 05.30.-d Quantum statistical mechanics

1 Introduction

Thermodynamics has been developed at a time, when the atomistic nature of matter was not well understood. Statistical mechanics tried to fill this gap, when, in turn, quantum mechanics as the micro-theory proper was unknown yet. Meanwhile quantum theory has invoked many changes, but the basic structure of modern theory is still close to that formulated by Gibbs some hundred years ago and therefore still plagued by old problems. Examples are the significance of ergodicity [1–4] the nature of irreversibility [4,5], the origin of the second law [4–7] and the relation between physics and information [8–10]. In spite of entropy being one of the most fundamental concepts in Gibbsian thermodynamics, there have been endless discussions about its nature and its precise definition [2,11]. As entropy (-change) is thermodynamically measurable (for any individual system), one would like to see its reduction to mechanics still to be based on observables. But entropy remains an alien concept within this frame: rather than to a point (state) it refers to a certain accessible volume in phase space, which depends on given macroscopic constraints such as energy, volume, etc. Nevertheless, this definition, eventually based on ergodicity, seems indispensable to calculate thermodynamical state functions.

On the other hand, one wants entropy to account also for the irreversibility, which is absent in the underlying microscopic equations of motion. Thus the concept of entropy appears to be related to the lack of knowledge that arises during an evolution too complicated to be traced in detail. Some sort of coarse-graining [2] is often invoked in this context, entropy would then be expected to increase up to the limits set by those macroscopic constraints. Such an information-theoretical approach also underlies Jaynes' principle [12,13]. This principle aims to guide our reasoning in the face of incomplete knowledge ("unbiased guess"). It is not at all restricted to physics. But from this point of view it seems as if entropy and the state functions would depend on the physicist's ability to minimize his lack of knowledge, which is, of course, unsatisfactory.

Within classical mechanics lack of knowledge may always be considered subjective: in principle any observable can be known with unlimited precision. This is different in quantum mechanics. From the uncertainty principle we know that there are always observables that are undetermined. Nevertheless, in single system scenarios, at least one observable can, in principle, be known exactly (pure states), while for compound systems there are even states where all observables referring to a specific subsystem are unknown, even if some compound observable of the full system was exactly predictable (pure but entangled and thus locally mixed states [9,14]). Compound systems might evolve from states that contain exact knowledge about some observable of each subsystem (pure product states) into the above mentioned states, featuring this fundamental lack of knowledge [15].

So, in the quantum domain we have two possible sources of ignorance: one being due to our inability to identify the initial state and calculate the evolution exactly, the other being intrinsic to the momentary state and thus present even for an"infinitely smart demon".

Here we want to show that in typical thermodynamical situations the fundamental lack of knowledge by far dominates over the subjective lack of knowledge in the following sense: Almost all the possible evolutions (of which we are typically unable to predict the actual one), will

---

[a] e-mail: `jochen@theo1.physik.uni-stuttgart.de`

eventually lead to states that are characterized by a maximum fundamental lack of knowledge about the considered subsystem; this lack is only limited by macroscopic constraints. Our considerations will entirely be based on quantum-mechanical system-environment-scenarios [5,8, 16].

2 State measures and their averages

2.1 Definitions

The entropy we are going to consider is the von Neumann entropy [14] given by

$$S(\hat{\rho}) := -k\mathrm{Tr}\left\{\hat{\rho}\ln\hat{\rho}\right\} \tag{2.1}$$

which is invariant with respect to any unitary transformation. In a diagonal representation of the density operator $\hat{\rho}$ this reads

$$S(\{W_n\}) := -k\sum_n W_n \ln W_n, \tag{2.2}$$

where the W_n's are the eigenvalues of $\hat{\rho}$, which are routinely interpreted as the probabilities to find the system in the eigenstate $|n\rangle$. Another measure for the broadness of the probability distribution is the purity P [17]:

$$P(\{W_n\}) := \sum_n W_n^2 \quad \text{or} \quad P(\hat{\rho}) := \mathrm{Tr}\left\{\hat{\rho}^2\right\} \tag{2.3}$$

which is also invariant with respect to unitary transformations. In general, these two measures do not uniquely map onto each other, but, as will be seen below, under some additional conditions maximum entropy states correspond to minimum purity states. Because of its less complicated structure, we mainly consider P and get back to S in the end.

Since we are going to deal with bi-partite systems partitioned into the considered system, labeled by g ("gas") and the surrounding labeled by c ("container") the pure state of the full system will be denoted as

$$|\psi\rangle = \sum_{A,B}\sum_{a,b}\psi_{ab}^{AB}|A,a\rangle\otimes|B,b\rangle \tag{2.4}$$

where $|A,a\rangle(|B,b\rangle)$ denotes the a-th(b-th) energy eigenstate with the energy eigenvalue $E_A(E_B)$ of the gas (container) system, with $a = 1, 2 ... N_A^g, b = 1, 2 ... N_B^c$. $N_A^g(N_B^c)$ are the respective degrees of degeneracy.

From any such wavefunction the reduced density operator of the subsystem, g, is found by tracing over the container system:

$$\hat{\rho}^g := \sum_{A,A',B}\sum_{a,a',b}\psi_{ab}^{AB}\left(\psi_{a'b}^{A'B}\right)^*|A,a\rangle\langle A',a'|. \tag{2.5}$$

For this state the purity is:

$$P^g = \sum_{ABCD}\sum_{abcd}\psi_{ab}^{AB}\left(\psi_{cb}^{CB}\right)^*\psi_{cd}^{CD}\left(\psi_{ad}^{AD}\right)^*. \tag{2.6}$$

Here and in the following $a, c, A, C,$ label the gas, $b, d, B, D,$ the container subsystem. (Note that $P^g = P^c$ as the total state is taken to be pure.)

2.2 Time and path averages

Now, let us consider the time average of some state measure M which could be the purity, the entropy, etc.

$$\overline{M} := \frac{1}{T}\int_0^T M(|\psi(t)\rangle)\mathrm{d}t. \tag{2.7}$$

Choosing a special parametrization for $|\psi\rangle$, we can convert the time integral into an integral over a trajectory generated by the total system's Hamiltonian for a given initial state. If this trajectory is represented by a "line" in a Cartesian space whose coordinates are the real (ψ_i) and imaginary (ψ_i') parts of the amplitudes of $|\psi\rangle$ in some basis $\{|i\rangle\}$, we can write instead of (2.7)

$$\overline{M} = \frac{\int_{|\psi(0)\rangle}^{|\psi(T)\rangle} M\left(\{\psi_i,\psi_i'\}\right)\frac{1}{v_{\mathit{eff}}}|\mathrm{d}|\psi\rangle|}{\int_{|\psi(0)\rangle}^{|\psi(T)\rangle}\frac{1}{v_{\mathit{eff}}}|\mathrm{d}|\psi\rangle|} \tag{2.8}$$

where $|\mathrm{d}|\psi\rangle|$ denotes the "length" of an infinitesimal step along this trajectory, in this representation of Hilbert space.

The advantage of this special parametrization derives from the fact that the effective velocity of a state vector $|\psi(t)\rangle$ along its trajectory through this Cartesian space

$$\begin{aligned} v_{\mathit{eff}}^2 &= \left|\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\right|^2 = \sum_i\left(\dot{\psi}_i^2 + \dot{\psi'}_i^2\right) \qquad (2.9)\\ &= \frac{1}{\hbar^2}\langle\psi(t)|\hat{H}^2|\psi(t)\rangle = \frac{1}{\hbar^2}\langle\psi(0)|\hat{H}^2|\psi(0)\rangle \end{aligned}$$

is constant on each trajectory and thus independent of the actual time t or the special point on the trajectory. Hence, the integral (2.8) simplifies to

$$\overline{M} = \frac{1}{L}\int_{|\psi(0)\rangle}^{|\psi(T)\rangle} M\left(\{\psi_i,\psi_i'\}\right)|\mathrm{d}|\psi\rangle| \tag{2.10}$$

where L is the length of the path. So, in this parametrization of Hilbert space, the time average of any state measure equals its path average along any given trajectory. We will exploit this relationship to infer the time-average (or even the "typical" momentary value) of M for cases, in which we know that M is constant over a large part of the Hilbert space region accessible by the respective trajectory, (see Sect. 3.2).

3 Microcanonical constraints

3.1 Accessible Hilbert space region

Any thermodynamical system has to allow for a partition of its full Hamiltonian in the following way:

$$\hat{H} =: \hat{H}^g + \hat{H}^c + \hat{I} \quad \text{with} \quad \left\langle\hat{I}\right\rangle \ll \left\langle\hat{H}^g\right\rangle, \left\langle\hat{H}^c\right\rangle \tag{3.1}$$

where $\hat{H}^g$ and $\hat{H}^c$ are the local Hamiltonians of the system and the environment, respectively, with $[\hat{H}^g, \hat{H}^c] = 0$,

$\hat{I}$ is some sort of interaction and the inequality for the expectation values has to hold for all states that the system can evolve into under given constraints. If a partition according to this weak coupling scheme was impossible, the idea of system proper and surrounding would be meaningless. Furthermore, weak coupling is required to justify the concepts of extensive/intensive variables, a basic feature of thermodynamics [18] (*cf.* also [19]).

Making such a partition possible, though, might involve a re-organisation of the Hamiltonian. Consider, for example, an ideal gas in a container. $\hat{H}^g$ in this case is a free particle Hamiltonian, $\hat{I}$ contains the interactions of all gas particles with all container particles. In this case, the expectation value of $\hat{I}$ for an energy eigenstate of the uncoupled problem $\hat{H}^g + \hat{H}^c$ would definitely not be small. But this deficiency can be overcome by defining an effective local Hamiltonian and an effective interaction:

$$\hat{H}^{g'} := \hat{H}^g + \hat{V}^g \quad \hat{I}' := \hat{I} - \hat{V}^g. \tag{3.2}$$

Here, $\hat{V}^g$ is some potential for the gas particles describing the mean effect of all container particles together. Substituting the real by the effective parts of the Hamiltonian obviously leaves the full Hamiltonian unchanged, but now the partition fits into the above scheme. In general, however, $\hat{I}'$ cannot be made zero. There are general theorems and concrete examples showing that such a "interaction reduction scheme" will only reduce the interaction, but the necessarily remaining effective interaction may have a significant effect [15,20].

It has often been claimed that a system whose inner energy is conserved could not interact with its environment. This, however, is typically not true (*cf.* [2,3]). A thermally isolated gas in a container, *e.g.*, definitely interacts with the container, otherwise the gas could not even have a well defined volume, thus there is a remaining effective interaction. But the mere existence of this interaction does not contradict the "microcanonicallity", the only constraint is, that it should not exchange energy. If the energies contained in the gas and the environment

$$E^{g'} := \left\langle \hat{H}^{g'} \right\rangle \quad E^c := \left\langle \hat{H}^c \right\rangle \tag{3.3}$$

are to be conserved, it follows that [6]

$$\left[\hat{H}^{g'}, \hat{I}'\right] = 0 \quad \left[\hat{H}^c, \hat{I}'\right] = 0. \tag{3.4}$$

Except for these constraints we need not specify $\hat{I}$ in more detail. All interactions that fulfill this relation will create perfectly microcanonical situations. Based on these commutator relations we find that for any energy eigenspace specified by A, B

$$\sum_{a,b} \left|\psi_{ab}^{AB}(t)\right|^2 = \sum_{a,b} \left|\psi_{ab}^{AB}(0)\right|^2 \tag{3.5}$$

is a conserved quantity, set by the initial state. Since we want to consider cases here that have zero local entropy in the beginning (product states), we get

$$\sum_{a,b} |\psi_{ab}^{AB}(0)|^2 = \sum_{a,b} |\psi_a^A(0)|^2 |\psi_b^B(0)|^2 = W_A^g W_B^c \tag{3.6}$$

where W_A^g (W_B^c) are the probabilities of finding the gas-(container-) system somewhere in the possibly highly degenerate subspace characterized by the energy eigenvalues E_A^g (E_B^c). This is the only constraint that microcanonical conditions impose on the accessible region of Hilbert space.

3.2 The "landscape" of P^g in the accessible region

We are not able to compute the time average of P^g according to (2.10) for we do not know $\hat{I}$ in detail, most likely we do not even know $\hat{H}^g$ and $\hat{H}^c$ in full detail, and even if we did, we could never hope to solve the Schrödinger equation for a system involving a macroscopic number of degrees of freedom. This is our subjective lack of knowledge. But, as will become clear shortly, this lack does not really matter.

To show this, we proceed as follows:
The minimum purity state consistent with the microcanonical conditions (3.5, 3.6) and its corresponding purity are:

$$\hat{\rho}_{min}^g = \sum_{A,a} \frac{W_A^g}{N_A^g} |A, a\rangle\langle A, a| \quad P_{min}^g = \sum_A \frac{(W_A^g)^2}{N_A^g} \cdot \tag{3.7}$$

We compute the average of P^g over the total accessible Hilbert space region. We will show that this average is very close to P_{min}^g for the large set of systems that fulfill the thermodynamic limit (see Sect. 5.2). Considering only these cases, we can conclude that $P^g \approx P_{min}^g$ for almost all states within this region. Since the only local state that has $P^g = P_{min}^g$ and is consistent with the microcanonical conditions is $\hat{\rho}_{min}^g$, all states within the accessible region that feature $P^g \approx P_{min}^g$ must yield reduced local states that are very close to $\hat{\rho}_{min}^g$.

Furthermore, as there is no other constraint for the paths entering (2.10) except to stay within the allowed region, any path will typically leave the very tiny region of maximum purity where it started if the initial state was a product state, and then venture through a region of almost minimum purity along almost all of its length. Thus in the long time limit any momentary state will be characterized by a maximum fundamental lack of knowledge, *i.e.*, a maximum entropy.

3.3 Hilbert space average of P^g

To calculate the Hilbert space average of P^g denoted as $\langle P^g \rangle$ we need a parametrization for $|\psi\rangle = \{\psi_{ab}^{AB}, \psi_{ab}^{AB'}\}$

confined to the allowed region (3.5, 3.6). The Hilbert space average can then be written as

$$\langle P^g \rangle = \frac{\int P^g\left(\{\phi_n\}\right) \det \mathcal{F} \prod_n \mathrm{d}\phi_n}{\int \det \mathcal{F} \prod_n \mathrm{d}\phi_n} \tag{3.8}$$

where ϕ_n is the respective set of parameters and $\mathcal{F}$ is the corresponding functional matrix.

According to (3.5) and (3.6) the real and imaginary parts of the amplitudes that correspond to a degeneracy subspace AB must be parametrized to lie on a hypersphere of radius $R = \sqrt{W_A^g W_B^c}$. Thus there is a corresponding set of parameters $\left\{\phi_n^{AB}\right\}$ for each degeneracy subspace on which the amplitudes of this and only this subspace depend. This means that the functional matrix $\mathcal{F}$ has block form and its determinant factorizes such that

$$\langle P^g \rangle = \frac{\int P^g \prod_{AB} \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}}{\prod_{AB} \int \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}} \cdot \tag{3.9}$$

As a consequence, the average over each term of P^g according to (2.6) factorizes and reduces to a product of averages over the degeneracy subspaces that the amplitudes in that very term correspond to.

The averages of the terms of P^g can be grouped to form "classes": It turns out that the averages of the terms of P^g belonging to the same class have the same functional dependence on the degree of degeneracy $N_{AB} := N_A^g N_B^c$ of the subspaces they correspond to and on the probability of those subspaces to be occupied $W_{AB} := W_A^g W_B^c$. Thus, we first have to evaluate the average of each term within a given class, add all terms and finally sum over all classes.

There are six classes to be considered:

I: All four amplitudes correspond to different states $((A \neq C) \vee (a \neq c)) \wedge ((B \neq D) \vee (b \neq d))$. Then each average is a product of averages of products of different Cartesian coordinates over full hyperspheres. As can be seen from the Appendix (7.8) those averages vanish.

II: The "gas indices" (indices referring to the gas subsystem) of all amplitudes correspond to the same gas state, but the container indices to two different container subspaces $((A = C) \wedge (a = c)) \wedge (B \neq D)$. They factorize into:

$$T^{II} = \left(\frac{\int |\psi_{ab}^{AB}|^2 \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}}{\int \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}}\right) \times \left(\frac{\int |\psi_{ad}^{AD}|^2 \det \mathcal{F}^{AD} \prod_n \mathrm{d}\phi_n^{AD}}{\int \det \mathcal{F}^{AD} \prod_n \mathrm{d}\phi_n^{AD}}\right) \cdot \tag{3.10}$$

The application of (7.8) yields

$$T^{II} = \frac{W_{AB}}{N_{AB}} \frac{W_{AD}}{N_{AD}} \cdot \tag{3.11}$$

There are $N_A^g N_B N_D^c$ terms belonging to that class and subspace-combination. Since subspaces that have $B = D$ are excluded we find, summing over subspaces:

$$\sum_{ABD} N_A^g (N_B^c N_D^c - \delta_{BD} (N_B^c)^2) T^{II} = \sum_A \frac{(W_A^g)^2}{N_A^g} \left(1 - \sum_B (W_B^c)^2\right) \tag{3.12}$$

III: The container indices of all amplitudes correspond to the same container state, but the gas indices to two different gas subspaces $(A \neq C) \wedge ((B = D) \wedge (b = d))$. By repeating the above procedure (II) we get:

$$\sum T^{III} = \sum_A \frac{(W_B^c)^2}{N_B} \left(1 - \sum_A (W_A^g)^2\right) \cdot \tag{3.13}$$

IV: The gas indices of all amplitudes correspond to the same gas state and the container indices to the same container subspace but to different container states within this subspace $((A = C) \wedge (a = c)) \wedge ((B = D) \wedge (b \neq d))$. Those terms are of the form

$$T^{IV} = \frac{\int |\psi_{ab}^{AB}|^2 |\psi_{ad}^{AB}|^2 \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}}{\int \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}} \tag{3.14}$$

To evaluate this equation one needs (7.9) and gets:

$$T^{IV} = \frac{(W_A^g)^2 (W_B^c)^2 \Gamma(N_{AB})}{\Gamma(N_{AB} + 2)} = \frac{(W_A^g)^2 (W_B^c)^2}{N_{AB}(N_{AB} + 1)} \cdot \tag{3.15}$$

For each subspace-combination there are $N_A^g N_B^c (N_B^c - 1)$ terms in that class. Thus, summing over subspaces we find:

$$\sum_{AB} N_{AB} N_B^c (N_B^c - 1) T^{IV} = \sum_{AB} \frac{(W_A^g)^2 (W_B^c)^2 (N_B^c - 1)}{N_{AB} + 1} \cdot \tag{3.16}$$

V: The container indices of all amplitudes correspond to the same container state and the gas indices to the same gas subspace but to different gas states within this subspace $((A = C) \wedge (a \neq c)) \wedge ((B = D) \wedge (b = d))$. For those terms the above calculation (IV) has to be repeated yielding:

$$\sum T^{V} = \sum_{AB} \frac{(W_A^g)^2 (W_B^c)^2 (N_A^g - 1)}{N_{AB} + 1} \cdot \tag{3.17}$$

VI: All four amplitudes correspond to the same state. These terms read:

$$T^{VI} = \frac{\int |\psi_{ab}^{AB}|^4 | \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}}{\int \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}} \cdot \tag{3.18}$$

Using (7.9) one gets:

$$T^{VI} = \frac{2(W_A^g)^2 (W_B^c)^2 \Gamma(N_{AB})}{\Gamma(N_{AB} + 2)} = \frac{2(W_A^g)^2 (W_B^c)^2}{N_{AB}(N_{AB} + 1)} \cdot \tag{3.19}$$

Since there are $N_A N_B$ terms in this class for each subspace, one gets:

$$\sum_{AB} T^{IV} N_{AB} = \sum_{AB} \frac{2(W_A^g)^2(W_B^c)^2}{N_{AB}+1} \,. \tag{3.20}$$

If we now finally sum up all the contributions from the six classes we get:

$$\begin{aligned}\langle P^g\rangle &= \sum_A \frac{(W_A^g)^2}{N_A^g}\left(1-\sum_B (W_B^c)^2\right)\\ &\quad+\sum_B \frac{(W_B^c)^2}{N_B^c}\left(1-\sum_A (W_A^g)^2\right)\\ &\quad+\sum_{A,B} \frac{(W_A^g)^2(W_B^c)^2(N_A^g+N_B^c)}{N_A^g N_B^c+1}\,.\end{aligned} \tag{3.21}$$

The Hilbert space average is thus a unique function of the invariants W_A^g, W_B^c, specified by the initial product state, and the degeneracies N_A^g, N_B^c.

If we, just to get in contact with previous results, ask for the average purity of an arbitrary state with no constraints at all we can enlarge the accessible region over all Hilbert space by formally taking both subsystems as completely degenerate. Doing so, we recover

$$\langle P^g\rangle = \frac{N^g+N^c}{N^g N^c+1} \tag{3.22}$$

as a special case [21].

If the degeneracy of the occupied energy levels is large enough that $N_A^g N_B^c + 1 \approx N_A^g N_B^c$, which should hold true for typical thermodynamical systems, (3.21) reduces to

$$\langle P^g\rangle \approx \sum_A \frac{(W_A^g)^2}{N_A^g} + \sum_B \frac{(W_B^c)^2}{N_B^c}\,. \tag{3.23}$$

The first sum in this expression is obviously exactly P^g_{min} (3.7), so that for systems and initial conditions, in which the second sum is very small, the allowed region almost entirely consists of states for which $P^g \approx P^g_{min}$. The second sum will be small if the container system occupies highly degenerate states, typical for thermodynamical systems, in which the surrounding is much larger than the considered system. This is the set of cases mentioned already in Section 3.2. Since the density operator that has $P^g = P^g_{min}$ and $S^g = S^g_{max}$ and is consistent with the microcanonical conditions is unique, the density operators with $P^g \approx P^g_{min}$ should not deviate much from this one and should therefore also have $S^g \approx S^g_{max}$. A more detailed lengthy but straightforward perturbative analysis shows that:

$$\langle S^g\rangle \approx S^g_{max}\left(\{W_A^g, N_A^g\}\right) - K\left(\sum_B \frac{(W_B^c)^2}{N_B^c}\right) \tag{3.24}$$

where K is a positive function that scales linearly with the system size of the gas system. (3.24) is valid for the thermodynamical regime, *i.e.*, if the second sum in (3.23) is small compared to the first one.

4 Canonical constraints

Our approach to the canonical conditions will be based on similar techniques as before. The possibility of a partition according to (3.1) is still assumed. But now there is no further constraint on the the interaction $\hat{I}$, since energy is allowed to flow from one subsystem to the other. The only constraint for the accessible region therefore derives from the initial state of the full system and the fact that the probability to find the total system at some energy E

$$W_E := \sum_{A,B/E} W_{AB} = \sum_{A,B/E}\sum_{a,b} |\psi_{ab}^{AB}|^2 \tag{4.1}$$

should be conserved. (Here $A, B/E$ stands for: all A, B such that $E_A + E_B = E$.) One could try to repeat the above calculation under this constraint, but now it turns out that the actual minimum purity is no longer near the average purity over the accessible region. Thus, one has to proceed in a slightly different way.

Contrary to the microcanonical case the probability to find the gas (container) subsystem at some given energy is no longer a constant of motion here. But we are going to prove that there is a predominant distribution, $\{W_{AB}^d\}$, which almost all states within the allowed region have in common. The subregion formed by these states will be called the "dominant region". Once the existence of such a dominant region has been established, we can use the results from the microcanonical conditions to argue that almost all states within this dominant region feature the maximum local entropy that is consistent with the predominant distribution.

Therefore, just like in the previous case, our subjective lack of knowledge about where to find the system within the accessible region should be irrelevant. The reduced local state $\hat{\rho}^g(t)$ as a function of the full state $|\psi(t)\rangle$ should always evolve into a state with a fixed probability distribution W_A^g and an almost time invariant entropy, which is the maximum entropy that is consistent with this (canonical) distribution. Nevertheless, the state of the full system continues to move with the constant velocity (2.10) in Hilbert space.

4.1 Identification of the dominant region

First, we calculate the size of a region in Hilbert space that is associated with a certain energy distribution $\{W_{AB}\}$. This size will then be maximized with respect to the W_{AB}'s under the condition of the energy probability distribution of the whole system $\{W_E\}$ being kept fixed, in order to find the predominant distribution $\{W_{AB}^d\}$. According to (3.9) the size of the region associated with the energy distribution $\{W_{AB}\}$ is:

$$A(\{W_{AB}\}) := \prod_{AB}\int \det \mathcal{F}^{AB} \prod_n \mathrm{d}\phi_n^{AB}. \tag{4.2}$$

Those integrals are just the surfaces of hyperspheres and can be done using the techniques described in the appendix:

$$A(\{W_{AB}\}) := \prod_{AB} W_{AB}^{N_{AB}-1/2} O(N_{AB}). \quad (4.3)$$

Here $O(N_{AB})$ is the surface area of a $2N_{AB}$-dimensional hypersphere of radius $R = 1$.

Instead of maximizing A directly, we choose to maximize $\ln A$; this is equivalent, since the logarithm is a monotonous function. Furthermore we drop all terms that do not depend on $\{W_{AB}\}$ since they are of no relevance for the maximum and set $N_{AB} - 1/2 \approx N_{AB}$, an approximation that is not necessary but simplifies the calculation and is definitely valid for large degrees of degeneracy. Introducing the Lagrange multipliers $\{\lambda_E\}$, the function we want to maximize with respect to the W_{AB} reads:

$$\ln \tilde{A} = \sum_{A,B} N_{AB} \ln W_{AB} - \sum_E \lambda_E \left(\sum_{A,B/E} W_{AB} - W_E \right). \quad (4.4)$$

This maximization is routinely done by solving the following set of equations:

$$\frac{\partial \ln \tilde{A}}{\partial W_{AB}} = 0 \quad (4.5)$$

and yields

$$W_{AB}^d = \frac{N_{AB}}{\lambda_{E=E_A+E_B}} \cdot \quad (4.6)$$

Finally, using (4.1) we find for the Lagrange multipliers

$$\lambda_E = \frac{N_E}{W_E} \quad N_E = \sum_{A,B/E} N_{AB}. \quad (4.7)$$

We have thus identified the energy probability distribution, which most of the states within the accessible region exhibit, *i.e.*, the energy probability distribution of the dominant region, $\{W_{AB}^d\}$.

4.2 Analysis of the size of the dominant region

So far we have only shown that among the regions with given energy probability distribution $\{W_{AB}\}$ there is a biggest (or smallest) one. But for our argument we need to show that this region $\tilde{A}^d$ is, indeed, extremely larger than all the others, that it really fills almost the entire accessible region. To examine the size of this region we need to know, how the size of a region depends on the corresponding distribution $\{W_{AB}\}$, if this distribution does not deviate much from the dominant distribution $\{W_{AB}^d\}$. Therefore we consider $W_{AB} =: W_{AB}^d + \epsilon_{AB}$, where the ϵ_{AB}'s are supposed to be small. For $\ln \tilde{A}$ we then find

$$\ln \tilde{A} = \sum_E \sum_{A,B/E} N_{AB} \ln(W_{AB}^d + \epsilon_{AB}) \quad (4.8)$$

with

$$\sum_{A,B/E} \epsilon_{AB} = 0. \quad (4.9)$$

The latter condition guarantees that $W_{AB} = W_{AB}^d + \epsilon_{AB}$ still belongs to the accessible region.

Expanding the logarithm to second order we get:

$$\ln \tilde{A} \approx \sum_E \sum_{A,B/E} N_{AB} \left(\ln(W_{AB}^d) + \frac{\epsilon_{AB}}{W_{AB}^d} - \frac{1}{2} \left(\frac{\epsilon_{AB}}{W_{AB}^d} \right)^2 \right) \cdot \quad (4.10)$$

Since the expansion is around an extremum the linear term should vanish. Indeed, using (4.9) the second summation over this term yields:

$$\sum_{A,B/E} N_{AB} \frac{\epsilon_{AB}}{W_{AB}^d} = \sum_{A,B/E} \lambda_E \epsilon_{AB} = 0 \quad (4.11)$$

so that, using (4.6) and (4.7), we finally find:

$$\tilde{A} \approx \tilde{A}^d \prod_{AB} \exp - \frac{\left(\sum_{A,B/E} N_{AB} \right)^2}{2 N_{AB} W_E} \epsilon_{AB}^2 \quad (4.12)$$

i.e. regions, $\tilde{A}$, that correspond to energy probability distributions that deviate from the dominant one are smaller than the dominant region, $\tilde{A}^d$. Since the smallest factor that can appear in the exponent of (4.12) for given N_E is $\frac{N_E}{2W_E}$, the regions $\tilde{A}$ will be extremely smaller already for very small deviations if the corresponding N_E's are large.

4.3 The canonical distribution

Finally, to find the marginal, dominant energy probability distribution W_A^d of the gas system one has to sum the compound probabilities W_{AB}^d over the irrelevant container system to obtain, using (4.7):

$$W_A^d = \sum_B W_{AB}^d = \sum_B \frac{N_{AB}}{\lambda_E} = N_A^g \sum_B \frac{N_B^c W_E}{N_E} \cdot \quad (4.13)$$

Going now from the degrees of degeneracy, N_A^g, N_B^c, to the state densities $n^g(E^g), n^c(E^c)$ and from sums to integrals, (4.13) reads:

$$W^d(E^g) = n^g(E^g) \int_{E_A}^{\infty} \frac{W(E)}{n(E)} n^c(E - E^g) \mathrm{d}E. \quad (4.14)$$

To better appreciate the meaning of this result we first consider the case of the full system energy being exactly specified in the beginning ($W(E) = \delta(E-U)$). Then (4.14) reduces to:

$$W^d(E^g) = \frac{n^c(U - E^g) n^g(E^g)}{n(U)} \cdot \quad (4.15)$$

This is precisely the energy probability distribution postulated in older approaches [22]. It is exactly what one gets for the considered system, if one assumes perfect ergodicity for the full system, an assumption that is not needed in our approach.

(4.14) is just a combination of such distributions weighted by the probability to find the total system at the corresponding energy and then adequately normalized.

For the remainder we follow those older approaches and their assumptions about the spectra of large systems. Thus we apply the standard expansion of $n^c(E-E_A)$ that is supposed to be valid even for large E^g, if the system is large,

$$n^c(E-E^g) \approx \mathrm{e}^{\ln n^c(E) - \frac{\mathrm{d}}{\mathrm{d}E}\ln n^c(E) E^g} \tag{4.16}$$

so that we can write

$$W^d(E_A) \approx n^g(E^g)\int_{E_A}^{\infty} \frac{W(E)n^c(E)}{n(E)}\mathrm{e}^{-\frac{\mathrm{d}}{\mathrm{d}E}\ln n^c(E)E^g}\mathrm{d}E. \tag{4.17}$$

Assuming the function $\frac{W(E)n^c(E)}{n(E)}$ to be still peaked sharply enough at some value $E = U$, such that $\frac{\mathrm{d}}{\mathrm{d}E}\ln n^c(E)$ does not vary much in the range where the former function takes on non-negligible values, (and again, this range can be fairly wide, if the system is large) we get

$$W^d(E_A) \propto n^g(E^g)\mathrm{e}^{-\frac{E^g}{kT}} \tag{4.18}$$

where we have defined $\frac{\mathrm{d}}{\mathrm{dU}}\ln n^c(U) := \frac{1}{kT}\cdot$

A set-up similar to the one that we discussed to derive the canonical distribution, but more specialized, has been analyzed by Tasaki [16]. His results for which entirely different techniques were used and some additional assumptions had to be made, are in good agreement with ours.

5 Discussion of the results

5.1 Attractor states

The accessible region of the pure state vector of the full, compound system can be decomposed into different zones, each of which yields local states with a common feature such as purity or energy probability distribution. Provided the thermodynamic limit (which is defined below) applies to the system, each accessible region is almost entirely filled with one dominant zone, corresponding to one special local state. This (equilibrium) state can thus be considered a local attractor. Hence even for trajectories starting in a tiny "non-equilibrium zone", corresponding to some strongly time dependent local state, after some time the gas system is most likely to be found in the time independent attractor state, no matter where exactly the total system keeps wandering around on its trajectory.

Those attractor states are as follows:

(i) Microcanonical conditions: Here we have found

$$\hat{\rho}^g_{mic.} \approx \sum_{A,a}\frac{W^g_A}{N^g_A}|A,a\rangle\langle A,a| \tag{5.1}$$

where A labels the respective energy eigenspace, a the different eigenstates within an eigenspace, W^g_A is the probability to find the system at the energy E_A and N^g_A is the degree of degeneracy of the corresponding energy level.

$\hat{\rho}^g_{mic.}$ is the state with the highest entropy that the initial state can possibly evolve into under strict energy conservation.

(ii) Canonical conditions: In this case we have shown that

$$\hat{\rho}^g_{can.} \approx \frac{\mathrm{e}^{\frac{-\hat{H}^g}{kT}}}{\mathrm{Tr}\left\{\mathrm{e}^{\frac{-\hat{H}^g}{kT}}\right\}} = \frac{\sum_{A,a}\mathrm{e}^{-\frac{E_A}{kT}}|A,a\rangle\langle A,a|}{\sum_A N_A\mathrm{e}^{-\frac{E_A}{kT}}}\cdot \tag{5.2}$$

This state is obviously the same one would have obtained, if one had applied Jaynes' principle, taking energy as the only relevant observable.

5.2 Quantum-thermodynamic limit

For the above to hold, the system has to fulfill certain requirements which then define the (quantum-) thermodynamic limit.

For both conditions the weak coupling limit should apply to the total composite system, *i.e.*:

$$\hat{H} =: \hat{H}^g + \hat{H}^c + \hat{I} \quad \text{with} \quad \left\langle \hat{I}\right\rangle \ll \left\langle \hat{H}^g\right\rangle, \left\langle \hat{H}^c\right\rangle\cdot \tag{5.3}$$

where $\hat{H}^g$ and $\hat{H}^c$ are the local Hamiltonians of the system and the environment, respectively, with $[\hat{H}^g, \hat{H}^c] = 0$, $\hat{I}$ is some sort of effective interaction. The expectation values are to be taken for the states that the system can possibly evolve into. If the coupling was too strong the notion of subsystems would become meaningless, if the coupling was too weak, thermalization times may become extremely long.

For a system to fulfill microcanonical conditions there are further requirements:

$$\left[\hat{H}^g, \hat{I}\right] = 0 \quad \left[\hat{H}^c, \hat{I}\right] = 0. \tag{5.4}$$

If (5.4) holds, no energy is exchanged between the two subsystems regardless of the strength of the interaction. Furthermore, the following condition has to be met:

$$\sum_B \frac{(W^c_B)^2}{N^c_B} \ll \sum_A \frac{(W^g_A)^2}{N^g_A}\cdot \tag{5.5}$$

This holds if the environment system, c, occupies energy levels of higher degeneracy than the considered system, g, and/or if its energy probability distribution is broader. This is likely to be the case, if the environment is much larger than the considered system.

For canonical conditions the full system state density $n(E)$ must be large at those energies that the full system occupies. In addition, the function $\frac{\mathrm{d}}{\mathrm{d}E^c}\ln n^c(E^c) =: \frac{1}{kT}$ has to be approximately constant over some energy range. Within this range the state density of the environment $n^c(E^c)$ has to be higher than the state density of the gas system $n^g(E^g)$ in the corresponding range. All this is, again, typical for "large systems" and for the container system being even larger than the gas system.

In this definition of the thermodynamical limit there is no necessity for the systems to consist of many particles (*cf.* [23]); nevertheless, the above mentioned criteria are most likely fulfilled by such systems.

5.3 Ergodicity

In classical statistics ergodicity (of isolated systems) is meant to imply hat the time average (of an individual system) equals the ensemble average. Experimentally accessible is only the former, but then it becomes questionable, under what conditions this equivalence should actually hold. And why should the (statistical) entropy based on the ensemble distribution have something to say about an individual system?

In the quantum treatment the reduced density operator $\hat{\rho}_g$ describes, indeed, the individual subsystem in interaction with its individual environment. Obviously the underlying quantum nature renders irrelevant questions like: How long do we have to wait until the single system has actually visited all those different pure states as required by the postulated ensemble-properties.

Only if one attempts to find a "quasi classical" interpretation of the reduced von Neumann entropy, the missing information about the actual local state may be ascribed to our "ignorance"

The fact that an individual system (embedded in a quantum environment) "simulates" a whole ensemble [24] may be considered a remarkable signature of "quantum parallelism" – a parallelism that has been implicitly anticipated since the beginnings of statistical mechanics.

This kind of "ergodicity" appears to be a natural ingredient of quantum mechanical system-environment scenarios. While the local entropy is a property of each individual total state vector, for the motion of this vector itself, ergodicity is not needed at all.

6 Concluding remarks

In this paper we have considered Hilbert space in quantum mechanics as the analog to phase space in classical mechanics. This analog is supported by the fact that in both cases "micro-states" correspond to state space points, and their dynamics to deterministic trajectories.

If the total system is partitioned into two subsystems (the considered or "gas" system, g, and the environment or "container", c), the reduced local state, $\hat{\rho}^g$, for the gas subsystem can be calculated from the pure total system state $|\psi\rangle$. Any such local state has typically non-zero entropy and different total system states may very well yield the same local system state. This fact, in turn, allows to define a probability distribution of local entropies on the space of total pure states. This distribution has been the main target of our investigation.

For composite systems we have specified classes of interactions, *i.e.*, microcanonical (mic.) and canonical (can.) Hamiltonians, respectively, which generate respective Hilbert space trajectories. These trajectories cannot cross the boundaries of different "accessible regions" which are set by the condition type ("mic." or "can."). This, again, is reminiscent of classical systems for which the trajectories may be confined to certain energy shells in phase space.

We have studied the fundamental mechanisms underlying Jaynes' principle with respect to energy exchange. Jaynes' principle, which may be taken as a rule for unbiased reasoning, thus coincides with emergent physical laws. Similar considerations should hold also for other interaction types like particle exchange (grand canonical condition). In any case, quantum mechanics should play a central role in the foundations of thermodynamics, not only, *e.g.*, in the low temperature limit. The predominance of local equilibrium states does not imply that non-equilibrium states were inaccessible or of minor interest. In fact any quantum machine, in particular the quantum computer, will require significant deviations from equilibrium. Machine design must specify means to prevent the system from running into the typical equilibrium behaviour.

We thank A. Otte, M. Stollsteimer, F. Tonner, M. Michel, H. Schmidt, T. Haury and P. Borowski for fruitful discussions. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

Appendix

The integrals that are to be evaluated in Section 3.3 are essentially of the form:

$$Z = \int z_l^{u_l}(\{\phi_n\}) z_m^{u_m}(\{\phi_n\}) \det \mathcal{F}(\{\phi_n\}) \prod_n^{d-1} \mathrm{d}\phi_n \quad (7.1)$$

where the u are integers and the z are parametrized by the ϕ to lie on a hypersphere of radius R and dimension d. (The z_l correspond to the real and imaginary parts of the amplitudes, respectively, d is the degree of degeneracy of a given subspace and R^2 the probability of this subspace to be occupied, *cf.* Sect. 3.2). The z_s are related to Cartesian coordinates x_s by:

$$\frac{z_s(\{\phi_n\})r}{R} = x_s \quad (7.2)$$

where r is a radial variable. The technique to solve this integral, Z, is basically the same as used to calculate the surface area of a hypersphere of arbitrary dimension. This surface area is also the special case of Z for $u_l = u_m = 0$.

Defining

$$Z_1 := \int_o^\infty \mathrm{e}^{-r^2} \left(\frac{r}{R}\right)^{u_l u_m + d - 1} \mathrm{d}r \quad (7.3)$$

the product $Z_2 = ZZ_1$ may be written as:

$$Z_2 := \int_o^\infty \mathrm{e}^{-r^2} \left(\frac{z_l r}{R}\right)^{u_l} \left(\frac{z_m r}{R}\right)^{u_m} \left(\frac{r}{R}\right)^{d-1} \times \det \mathcal{F} \prod_n^{d-1} \mathrm{d}\phi_n \mathrm{d}r. \quad (7.4)$$

Now, this is an integral over all space, written in angular and radial coordinates that can be converted to Cartesian coordinates yielding:

$$Z_2 = \int \mathrm{e}^{-\sum_s x_s^2} x_l^{u_l} x_m^{u_m} \prod_s^d \mathrm{d}x_s. \quad (7.5)$$

Since this integral factorizes completely, both Z_2 and Z_1 can be evaluated using standard tables of integrals, and depend only on R, d, u_l, u_m. (Note that Z_2 vanishes if any of the u is odd.) Z is then found to be

$$Z(R, d, u_l, u_m) = \frac{Z_2(R, d, u_l, u_m)}{Z_1(R, d, u_l, u_m)} \cdot \quad (7.6)$$

The averages that are to be computed in Section 3.3 are of the form:

$$A(R, d, u_l, u_m) := \frac{Z(R, d, u_l, u_m)}{Z(R, d, 0, 0)} \cdot \quad (7.7)$$

They are all invariant with respect to exchange of the u's. Here we need only:

$$A(R, d, 0, 1) = A(R, d, 1, 1) = 0, \quad A(R, d, 0, 2) = \frac{R^2}{d} \quad (7.8)$$

which could also have been found from simple symmetry considerations, and

$$A(R, d, 2, 2) = \frac{R^4 \Gamma\left(\frac{d}{2}\right)}{4\Gamma\left(\frac{d}{2}+2\right)} \quad A(R, d, 0, 4) = \frac{3R^4 \Gamma\left(\frac{d}{2}\right)}{4\Gamma\left(\frac{d}{2}+2\right)} \cdot \quad (7.9)$$

All A's are invariant with respect to exchange of the u's.

References

1. J.V. Neumann, Zeitschr. F. Physik **57**, 30 (1930)
2. J.M. Blatt, Progr. Theor. Phys. **22**, 745 (1959)
3. P.G. Bergmann, J.L. Lebowitz, Phys. Rev. **99**, 578 (1955)
4. O. Penrose, *Foundations of Statistical Mechanics* (Pergamon Press, New York, 1970)
5. Ya.P. Terlitskii, *Statistical Physics* (North Holland Publ., Amsterdam, 1991)
6. J. Gemmer, A. Otte, G. Mahler, Phys. Rev. Lett. **86**, 1927 (2001)
7. J. Berger, Physics Essays **7**, 281 (1994)
8. W. Zurek, J. Paz, Phys. Rev. Lett. **72**, 2508 (1994), Comment Phys. Rev. Lett. **75**, 351 (1995)
9. L.E. Ballantine, *Quantum Mechanics* (World Scientific, Singnapore, 1999)
10. R. Landauer, IBM J. Res. Develop. **5**, 183 (1961)
11. M. Li, P. Vitanyi, *An Introduction to Kolmogorov Complexity and its Applications*, Chap. 8 (Springer, New York, Berlin, 1993)
12. E.T. Jaynes, Phys. Rev. **106**, 620 (1957)
13. For a web-site dedicated to E.T. Jaynes, see `http://bayes.wustl.edu`
14. K. Blum, *Density matrix. Theory and applications* (Plenum Press, New York, 1981)
15. J. Gemmer, G. Mahler, Eur. Phys. J. D **17**, 385 (2001)
16. H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998)
17. C. Tsallis, J. Stat. Phys. **52**, 479 (1988)
18. B. Diu, C. Guthmann, D. Lederer, B. Roulet, *Elements de Physique Statistique* (Hermann Éditeurs des Sciences et des Arts, Paris, 1989)
19. A.E. Allahverdyan, Th.M. Nieuwenhuizen, Phys. Rev. Lett. **85**, 1799 (2000)
20. J. Gemmer, G. Mahler, Europhys. Lett. **59**, 159 (2002)
21. E. Lubkin, Math. Phys. **19**, 1028 (1978)
22. E. Schrödinger, *Statistical Thermodynamics* (Dover Publ., New York, 1989)
23. R.V. Jensen, R. Shankar, Phys. Rev. Lett. **54**, 1879 (1985)
24. In the older literature this situation has been characterized as an "improper mixture", *cf.*, *e.g.*, B. d'Espagnat *Conceptual Foundations of Quantum Mechanics* (Benjamin, London, 1976)