

## The second law of thermodynamics as a deterministic theorem for quantum spin systems

Walter F. Wreszinski

*Instituto de Física,  
 Universidade de São Paulo (USP), Brazil  
 wreszins@gmail.com*

Received 25 May 2022

Accepted 15 August 2022

Published 21 September 2022

We review our approach to the second law of thermodynamics as a theorem asserting the growth of the mean (Gibbs–von Neumann) entropy of quantum spin systems undergoing automorphic (unitary) adiabatic transformations. Non-automorphic interactions with the environment, although known to produce on the average a strict reduction of the entropy of systems with finite number of degrees of freedom, are proved to conserve the mean entropy on the average. The results depend crucially on two properties of the mean entropy, proved by Robinson and Ruelle for classical systems and Lanford and Robinson for quantum lattice systems: upper semicontinuity and affinity.

*Keywords:* Second law; unitary evolution; quantum spin systems; mean entropy; upper semicontinuity; affinity.

Mathematics Subject Classification 2020: 81-XX, 82-XX

### 1. Introduction and Summary

In this paper, we review our approach [47] to the second law of thermodynamics as a theorem asserting the growth of the mean (Gibbs–von Neumann) entropy of a class of quantum spin systems undergoing automorphic (unitary) adiabatic transformations. The structure of the proofs in the original reference is clarified at several points in which only a sketch is given, at the same time allowing for a larger class of interaction potentials (of polynomial decrease rather than only exponential decrease at infinity). We also compare this framework with our previous approach to the second law, together with Narnhofer [30] based on the (quantum) Boltzmann entropy. Our main new result is that non-automorphic interactions with the environment, although known to produce on the average a strict reduction of the Boltzmann entropy of systems with finite number of degrees of freedom [30, Lemma 3], are proved to conserve the mean entropy on the average, as a consequence of the latter's property of affinity. As a consequence, we do not need to assume that such interactions are rare on the thermodynamic time scale, in order to assure the validity of the second law (see the remarks on the last paragraph of [30]).

As described by Wehrl in a still very readable paper [43, p. 227 et seq. “A paradox”], the common assertion of the second law of thermodynamics, that the entropy of a closed system never decreases, is in striking contradiction to the following fact. Let  $S_{N,V}^G$  denote the Gibbs–von Neumann entropy of a system of  $N$  particles in a box of volume  $V$ , i.e.

$$S_{N,V}^G = -k_B \text{Tr} \rho_{N,V} \log(\rho_{N,V}), \quad (1)$$

where  $\rho_{N,V}$  is a positive trace-class operator (density matrix) describing the system and  $\text{Tr}$  is the trace over the corresponding Hilbert space  $\mathcal{H}_{N,V}$ . For a system described by a time-independent Hamiltonian  $H_{N,V}$ , however, the density matrix  $\rho_{N,V}(t)$  at time  $t$  is obtained from the density matrix at time zero  $\rho_{N,V}$  from the formula

$$\rho_{N,V}(t) = \exp(-iH_{N,V}t) \rho_{N,V} \exp(iH_{N,V}t) \quad (2)$$

and since  $\exp(iH_{N,V}t)$  is a unitary operator, the repeated eigenvalues of  $\rho_{N,V}(t)$  are the same as the eigenvalues of  $\rho_{N,V}$ , and thus

$$S_{N,V}^G(t) = S_{N,V}^G(0). \quad (3)$$

As Wehrl remarks, “there is one way out of this dilemma: the time-evolution of a system is not described by the Schrödinger equation, but by some other equation. In fact, in statistical mechanics one uses, with great success, equations like the Boltzmann equation, the master equation, and other equations.” We refer to [39] for a very comprehensive discussion of the master equation approach, as well as references. This approach concerns open systems. In the present paper we shall be, however, concerned with *closed* systems — systems completely isolated from all external influences. Traditional thermodynamics takes this idealization as starting point, which is also done by Lieb and Yngvason in their rigorous axiomatic approach [21]. Examples such as the evolution of the Universe or the adiabatic irreversible expansion of a gas, in which the isolation of the system may be achieved to any degree of accuracy, suggests that this idealization also has a fundamental character, from the physical point of view (see also [47]).

We wish to maintain the universal definition (1) of the Gibbs–von Neumann entropy, and attempt to find a dynamical proof of the second law, continuing to assume a deterministic evolution (2), which implies (3). An important hint in this direction was provided by Penrose and Gibbs [33, p. 1959; 12] in the framework of classical statistical mechanics, whereby (1) becomes

$$S_G^\Gamma(\rho) = -k_B \int_\Gamma dx \rho(x) \log(\rho(x)) \quad (4)$$

(defining  $x \log(x) = 0$ , for  $x = 0$ , as we do throughout), where  $\Gamma$  denotes phase space and  $\rho$  is a phase space density, i.e. a non-negative function, absolutely continuous with respect to Lebesgue measure  $dx$ , such that

$$\int_\Gamma dx \rho(x) = 1. \quad (5)$$

The number of particles  $N$  and the box of volume  $V$  are implicit in  $\Gamma$ ; if necessary to make them explicit, we shall write  $\Gamma_{N,V}$ . Define [33] an adiabatic process as a process in which no heat enters or leaves the system, starting at time zero, when the Hamiltonian is  $H_{t=0}$  and the phase space density  $\rho_0$  is invariant under  $H_{t=0}$ , and ends at time one, when the Hamiltonian has changed to  $H_{t=1}$ , supposed to yield a mixing (hence ergodic) motion thereafter. In spite of the time-dependence of the Hamiltonian, Liouville's theorem remains valid, which permits to calculate the phase space density  $\rho_t$  for  $t \geq 0$ , but the final equilibrium is assumed to be described by the coarse-grained phase-space density  $\bar{\rho}$ , defined by the averaging prescription

$$\bar{\rho}(x) = \lim_{T \rightarrow \infty} \frac{\int_1^{1+T} \rho_t dt}{T}, \quad (6)$$

which exists for a.e.  $x \in \Gamma$ , by Birkhoff's ergodic theorem and which, for ergodic systems as assumed, yields an invariant density (see, e.g., [42, Chap. 1, paragraph 5]), just as equilibrium is modeled in Gibbs-ensemble theory. We may now state

**Theorem 1.1 (Penrose–Gibbs Theorem).** *Under the above assumptions,*

$$S_G^\Gamma(\rho_0) \leq S_G^\Gamma(\bar{\rho}). \quad (7)$$

As remarked by Penrose [33, p. 1959], mathematically, the inequality (7) illustrates the non-interchangeability of the functional  $S_G^\Gamma$ , given by (4), and the limit on the right-hand side of (6), since, by Liouville's theorem (the analogue of (3)), the left-hand side of (7) equals  $\lim_{T \rightarrow \infty} \frac{\int_1^{1+T} S_G^\Gamma(\rho_t) dt}{T}$ , while the right-hand side equals  $S_G^\Gamma(\lim_{T \rightarrow \infty} \frac{\int_1^{1+T} \rho_t dt}{T})$ .

In [47], we observed that, for finite systems, i.e. associated to a phase space  $\Gamma$  of finite measure (with  $N, V$  finite),  $S_G^\Gamma$ , given by (4), is a continuous functional of  $\rho$  (in the natural, weak\* topology, see [39, p. 57]), and therefore, unfortunately, (3) continues to hold.

For classical systems, it is the *mixing property* which determines the approach to equilibrium, in the sense

$$\lim_{t \rightarrow \infty} \int_\Gamma \rho_t(x) G(x) dx = \langle G \rangle_{\text{eq}} \equiv \int_\Gamma \bar{\rho}(x) G(x) dx \quad (8)$$

for  $G$  an observable, i.e. a continuous function over phase space (see [33, (1.35), p. 1949], and references given there). Equation (8) suggests regarding  $\rho$  as a *state*, i.e. a normalized linear functional over the algebra of observables, which turns out to be the natural choice for both classical and quantum statistical mechanics, in the former case the algebra being abelian. In this context, the notion of *mean entropy*

$$s_G(\rho) = \lim_{V \rightarrow \infty} \frac{S_{G^{N,V}}^\Gamma(\rho)}{V} \quad (9)$$

as a functional of the state  $\rho$  is the natural quantity to be considered. In (9), the limit  $V \rightarrow \infty$ , assumed to exist, is the thermodynamic limit in the sense of van Hove, whereby  $N \rightarrow \infty$ ,  $V \rightarrow \infty$ , with  $\frac{N}{V} = d$ ,  $d$  denoting the particle density in classical statistical mechanics [36], or the van Hove limit in the case of quantum spin systems.

The reason for considering the mean entropy  $s_G(\rho)$  rather than the entropy  $S_G^\Gamma(\rho)$  of finite systems is two-fold, and appears in both contexts of automorphic evolutions (Sec. 3) and non-automorphic evolutions (Sec. 4), the latter also relevant to the measurement problem studied by Hepp in his seminal paper [16].

First, the limit  $t \rightarrow \infty$  will not, in general, commute with the limit of infinite volume (for an explicit example of this fact in a soluble model in cosmology describing the CMB (cosmic microwave background) radiation, see [46]). Indeed, the rate of approach to equilibrium may grow indefinitely when the volume tends to infinity, and it is therefore of essential importance to consider the *densities*, such as the entropy density in (9), or the space-average of the magnetization in the model of Sec. 4.

Second, quantum systems imperatively require the use of infinite systems in the study of dynamics, since, for finite systems, the Hamiltonian has a discrete spectrum and the observables are quasi-periodic functions of time (see [33, Sec. 1.5]).

It turns out that consideration of  $s_G(\rho)$  rather than  $S_G^{\Gamma_{N,V}}(\rho)$  entails two bonuses: in the first place, it is upper-semicontinuous rather than continuous, leading to a refined form of Theorem 1.1, in which strict growth (of  $s_G(\rho)$ ) for large times is possible (Theorem 3.1). But even more is true: the property of affinity allows a proof of stability of the second law, in the form of Theorem 3.1, under what we believe to be a paradigmatic non-automorphic interaction with the environment (Theorem 4.1) — in sharp contrast to the situation for finite systems, in which a *reduction* of the average entropy (von Neumann or Boltzmann) occurs [30, Lemma 3].

As observed, a state  $\omega$  will be defined as a normalized, positive linear functional over the algebra  $\mathcal{A}$  of observables of an infinite system. Readers unfamiliar with this notion may consult the classic book [39]. This seems to be the adequate moment to make a brief interlude to explain our restriction to quantum spin systems. We believe that the forthcoming Theorem 3.1 should have a wide domain of applicability, including classical statistical mechanics, but consideration of the momenta still poses an open problem in the classical case (see [37, p. 1666]). Concerning the time evolution of quantum systems, only for quantum spin systems is the time-evolution an automorphism of the natural  $C^*$  algebra of observables, a fact used in Theorem 3.1; for quantum continuous systems, it is not, in general, the case. In order to start our explanation, we must introduce the main concepts.

An *automorphism*  $\tau_t(A)$  of  $\mathcal{A}$  is basic concept: it is a one-to-one mapping of  $\mathcal{A}$  onto  $\mathcal{A}$  which preserves the algebraic structure. For quantum spin systems, it equals the limit in the operator norm of the time-evolutes  $\exp(iH_\Lambda t)A\exp(-iH_\Lambda t)$

as the region  $\Lambda \nearrow \mathbf{Z}^\nu$ , and one speaks of a  $C^*$ -dynamical system  $(\mathcal{A}, \tau_t)$ . The letter  $\nu$  will always denote the spatial dimension.

Comparison with the discussion preceding Theorem 1.1 suggests the following definition (which follows [33] closely, adapting his discussion to a framework which includes states of infinite systems).

**Definition 1.2.** Let a  $C^*$ -dynamical system  $(\mathcal{A}, \tau_t)$  be given, with  $t \in [-r, \infty)$ . An *adiabatic transformation* consists of two successive steps. The *first step*, called preparation of the state, starts at some  $t = -r$ , with  $r > 0$ , when the state  $\omega_{-r}$  is invariant under the automorphism  $\tau_{-r}$ , and ends at  $t = 0$ . We require the cyclicity condition

$$\tau_{-r} = \tau_0 = \tau. \quad (10)$$

The *second step* is a dynamical evolution of the state  $\omega$  in the form

$$t \in \mathbf{R} \rightarrow \omega_t \equiv \omega \circ \tau_t, \quad (11)$$

where the circle denotes composition, i.e.  $(\omega \circ \tau_t)(A) = \omega(\tau_t(A))$ . We call (11) an automorphic evolution. It is assumed to be non-trivial, i.e. the state after preparation is *not* invariant under the evolution:

$$(\omega_0 \circ \tau) \neq \omega_0. \quad (12)$$

Note that the automorphism  $\tau$  refers to a *time-independent* interaction. Between  $t = -r$  and  $t = 0$  in the above definition a “time-dependent Hamiltonian” is supposed to act, see the above discussion of the Penrose–Gibbs theorem in the classical case. Indeed, according to Definition 1.2, the system is closed from  $t = 0$  to any  $t > 0$ , but not from  $t = -r$  to  $t = 0$ , where it is subject to *external conditions*, but is still thermally isolated. One form of the second law of thermodynamics may be found in the seminal paper of Pusz and Woronowicz [34]: if  $\omega_{-r}$  is a *passive state* (defined in [34] or [5, p. 101]), it is unable to perform work in a cyclic process. This statement is analogous to the Kelvin–Planck statement, rigorously formulated and proved as Theorem 3.4 of [21]. Under certain additional conditions ([34], or [5, Theorem 5.3.22]), it may be proved that  $\omega_{-r}$  is either a ground state, a thermal (KMS) state or a ceiling (infinite temperature) state. We shall refer to any of the latter states as an *equilibrium state*, and chose our  $\omega_{-r}$  in Definition 1.2 as an equilibrium state. As remarked in [5, (5.4.4), p. 211], passivity “reflects a property of stability of equilibrium states which is basically kinematic”. It does not include the second step in Definition 1.2, which is basically dynamic and may be called a relaxation. For finite systems, this process conserves the entropy, but we shall see that, for systems with infinite number of degrees of freedom, it may lead to growth of the mean entropy. In the explicit examples we are able to provide, the mean entropy is conserved and equals zero in the first step, while the growth appears in second step; it is a characteristic of states of infinite systems that pure states may tend to mixed states for large times, a fact observed by Hepp [16].

Equation (12) is a condition of *dynamical instability*. We shall correspondingly refer to  $\omega_0$  as an *unstable state*.

**Remark 1.1.** The “time-arrow” problem.

In Definition 1.2 we see that primitive causality requires a time interval  $r \neq 0$  preceding the time of appearance of the state  $\omega_0 = \omega$  during which the state is prepared. This is obviously not time-reversal invariant, and this mere fact implies the existence of an “arrow of time”, see also [47, 32]. We shall implicitly assume it in the present paper. See also Sec. 4.4.

## 2. Quantum Spin Systems

### 2.1. Generalities

A prototype, which we shall use, is the generalized Heisenberg Hamiltonian (generalized Ising model (gIm)) if  $J_1 = 0$ ):

$$H_\Lambda = -2 \sum_{x,y \in \Lambda} [J_1(x-y)(S_x^1 S_y^1 + S_x^2 S_y^2) + J_2(x-y)S_x^3 S_y^3], \quad (13)$$

where

$$\sum_{x \in \mathbf{Z}^\nu} |J_i(x)| < \infty \quad \text{and} \quad J_i(0) = 0 \quad \text{for } i = 1, 2. \quad (14)$$

Above,  $\vec{S}_x \equiv (S_x^1, S_x^2, S_x^3)$ , where  $S_x^i = 1/2\sigma_x^i$ ,  $i = 1, 2, 3$  and  $\sigma_x$  are the Pauli matrices at the site  $x$ . Above,  $H_\Lambda$  acts on the Hilbert space  $\mathcal{H}_\Lambda = \otimes_{x \in \Lambda} \mathbf{C}_x^2$  and  $\vec{S}_x$  is short for  $\mathbf{1} \otimes \cdots \otimes \vec{S}_x \otimes \cdots \otimes \mathbf{1}$ . We define  $\mathcal{A}(\Lambda) = B(\mathcal{H}_\Lambda)$ . These local algebras satisfy: (1) Causality:  $[\mathcal{A}(B), \mathcal{A}(C)] = 0$  if  $B \cap C = \emptyset$  and (2) Isotony:  $B \subset C \Rightarrow \mathcal{A}(B) \subset \mathcal{A}(C)$ .  $\mathcal{A}_L = \cup_B \mathcal{A}(B)$  is termed the *local algebra*; its closure with respect to the norm, the *quasilocal algebra* (observables which are, to arbitrary accuracy, approximated by observables attached to a *finite region*). The norm is  $A \in B(\mathcal{H}_\Lambda) \rightarrow \|A\| = \sup_{\|\Psi\| \leq 1} \|A\Psi\|$ ,  $\Psi \in \mathcal{H}_\Lambda$ . An automorphism  $\tau_t(A)$  equal to the norm limit of  $\lim_{\Lambda \nearrow \infty} \exp(iH_\Lambda t) A \exp(-iH_\Lambda t)$  for  $A \in \mathcal{A}(\Lambda)$ . The limit  $\lim_{\Lambda \nearrow \infty}$  will denote the van Hove limit [5, p. 287].

For a finite quantum spin system the Gibbs–von Neumann entropy is ( $k_B = 1$ )

$$S_\Lambda = -\text{Tr}(\rho_\Lambda \log \rho_\Lambda). \quad (15)$$

We may view  $\rho_\Lambda$  as a *state*  $\omega_\Lambda$  on  $\mathcal{A}(\Lambda)$  — a positive, normed linear functional on  $\mathcal{A}(\Lambda)$ :  $\omega_\Lambda(A) = \text{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda A)$  for  $A \in \mathcal{A}(\Lambda)$  (positive means  $\omega_\Lambda(A^\dagger A) \geq 0$ , normed  $\omega_\Lambda(\mathbf{1}) = 1$ ).

### 2.2. Finite versus infinite systems

This notion of state generalizes to systems with infinite number of degrees of freedom  $\omega(A) = \lim_{\Lambda \nearrow \infty} \omega_\Lambda(A)$ , at first for  $A \in \mathcal{A}_L$  and then to  $\mathcal{A}$ . The state

$\omega_\Lambda^t(A) = \omega_\Lambda(\exp(iH_\Lambda t)A\exp(-iH_\Lambda t))$  does *not* have a limit as  $\Lambda \nearrow \infty$  because the spectrum is discrete and the state an almost periodic function of  $t$ .

As we shall see later, for states of infinite systems, however, the situation is *entirely different!*

The function  $S_\Lambda^t \equiv S_\Lambda(\omega_\Lambda^t) = -\text{Tr}(\rho_\Lambda^t \log \rho_\Lambda^t) = S_\Lambda^0$ . It is a *continuous* functional of the state  $\omega_\Lambda^t$ . For a large system the *mean entropy* is the natural quantity from the physical standpoint:

$$s(\omega) \equiv \lim_{\Lambda \nearrow \infty} \left( \frac{S_\Lambda}{|\Lambda|} \right) (\omega). \quad (16)$$

The mean entropy has the following two properties [17]:

- (a)  $0 \leq s(\omega) \leq \log D$ , where  $D = 2S + 1$ ;
- (b)  $s$  is *upper semicontinuous*, that is  $\limsup_{n \rightarrow \infty} s(\omega_n) \leq s(\omega)$ .

For (b) see [4, Appendix 2, Lemma I.1]:  $\omega_n$  is a sequence of states such that  $\omega_n \rightarrow \omega$  in the weak\*-topology, i.e.  $\omega_n(A) \rightarrow \omega(A) \forall A \in \mathcal{A}$ . One simple example of (b) is a characteristic function of a closed set. For the statistical thermodynamical significance of this property, see [39, p. 55].

Let  $\Gamma = \mathbf{Z}^\nu$  and  $P_0(\Gamma)$  denote the collection of all finite parts of  $\Gamma$ . We consider in this paper, quantum spin systems described by a Hamiltonian for any finite region  $\Lambda \subset \mathbf{Z}^\nu$

$$H_\Phi(\Lambda) = \sum_{X \subset \Lambda} \Phi(X), \quad (17)$$

where  $\Phi$  is a translation-invariant interaction function from  $P_0(\Gamma)$  into the Hermitian elements of the quasi-local algebra  $\mathcal{A}$ , such that  $\Phi(X) \in \mathcal{A}(X)$  and

$$\Phi(X + x) = \tau_x(\Phi(X)) \forall X \subset \mathbf{Z}^\nu, \quad \forall x \in \mathbf{Z}^\nu. \quad (18)$$

Although general quantum lattice systems could be treated, including Fermion lattice systems, we shall for simplicity restrict ourselves to quantum spin systems with two-body interactions. In this case, stability requires that

$$\sum_{0 \neq x \in \mathbf{Z}^\nu} \|\Phi(\{0, x\})\| < \infty. \quad (19)$$

If, in (13), we consider the Ising limit, i.e. set  $J_1 \equiv 0$ , and define  $J_2 = J$ , we obtain the generalized Ising model (gIm) studied by Emch [8] and Radin [35] (we changed the overall sign to agree with Radin's notation)

$$H_\Lambda = \frac{1}{2} \sum_{x, y \in \Lambda} J(|x - y|) \sigma_x^3 \sigma_y^3 \quad (20)$$

with

$$J(0) = 0 \quad \text{and} \quad \sum_{x \in \mathbf{Z}^\nu} |J(|x|)| < \infty. \quad (21)$$

W. F. Wreszinski

Let, now,  $\nu = 1$ , and define two subclasses of models of the gIm:

$$J(|x|) = \xi^{-|x|} \quad \text{with } \xi > 1 \quad \text{and} \quad x \neq 0 \quad (\text{exponential model } E_\xi) \quad (22)$$

and

$$J(|x|) = \frac{1}{|x|^\alpha} \quad \text{with } \alpha > 1 \quad \text{and} \quad x \neq 0 \quad (\text{Dyson model } D_\alpha). \quad (23)$$

The above conditions on  $\xi$  and  $\alpha$  are required for stability (21). Under certain conditions, the Dyson model displays a ferromagnetic phase transition [7]. With  $P_0(\mathbf{Z}^\nu)$  denoting the set of all finite subsets of  $\mathbf{Z}^\nu$  as before, for each triple  $A = (A_1, A_2, A_3)$ , where the  $A_i \in P_0(\mathbf{Z}^\nu, i = 1, 2, 3)$  are pairwise disjoint, define

$$\sigma^A \equiv \prod_{x_1 \in A_1} \sigma_{x_1}^1 \prod_{x_2 \in A_2} \sigma_{x_2}^2 \prod_{x_3 \in A_3} \sigma_{x_3}^3. \quad (24)$$

Let  $\omega$  be any state satisfying

$$\omega(\sigma^A) = 0 \quad \forall A \quad \text{such that } A_3 \neq \phi. \quad (25)$$

We have the following theorem.

**Theorem 2.1.** *For the gIm with interaction either of the exponential model  $E_\xi$  with  $\xi$  a transcendental number, or the Dyson model  $D_\alpha$ , and any initial state (at  $t = 0$ )  $\omega$  satisfying (25),*

$$\lim_{t \rightarrow \infty} (\omega \circ \tau_t) = \omega_{\text{eq}} \equiv \otimes_{x \in \mathbf{Z}} \text{Tr}_x, \quad (26)$$

where the limit on the left-hand side of (26) is taken in the weak\* topology.

For the proof, see [35, Corollary, p. 2953].

### 2.3. The Lieb–Robinson bound and a limit theorem of Nachtergaele, Sims and Young on local approximations to time-translation automorphisms

Quantum spin systems are similar to quantum fields because of the now famous Lieb–Robinson bound ([20], see also [5, p. 254]). Following the last reference, let us assume, instead of (19),

$$\|\Phi\|_\lambda \equiv \sum_{x \in \mathbf{Z}^\nu} \|\Phi(\{0, x\})\| \exp(\lambda|x|) < \infty \quad (27)$$

for some  $\lambda > 0$ . Then, for  $A, B \in \mathcal{A}_0$ ,

$$\|[(\tau_x \tau_t)(A), B]\| \leq 2\|A\| \|B\| \exp[-|t|(\lambda|x|/|t| - 2\|\Phi\|_\lambda)]. \quad (28)$$

In (28),  $(\tau_x \tau_t)$  may be replaced by  $(\tau_t \tau_x)$ . The commutator  $[(\tau_x \tau_t)(A), B]$  with  $B \in \mathcal{A}_0$  provides a measure of the dependence of the observation  $B$  at the point  $x$  at time  $t$  at the origin at time  $t = 0$ , showing that this effect decreases exponentially with time outside the cone

$$|x| < |t| \left( \frac{2\|\Phi\|_\lambda}{\lambda} \right). \quad (29)$$



Equation (29) means that physical disturbances propagate with “group velocity” bounded by

$$v_{\Phi} \equiv \inf_{\lambda} \left( \frac{2\|\Phi\|_{\lambda}}{\lambda} \right). \quad (30)$$

Such bounds have been considerably developed in the last thirty years, and the recent reference by Nachtergaele, Sims and Young [24] contains a comprehensive summary of some of these results, as well as an extensive list of references to earlier work on the subject.

Even more significant for us in Sec. 3 are new results in [24] on local approximations of time-translation automorphisms of local observables, which hold even under conditions weaker than (27), for which the interpretation of the Lieb–Robinson bounds in terms of a bounded “group velocity” of physical disturbances (28), (29) no longer applies. We now assume (only) (19) and follow [24] to introduce the notion of an  $F$  function on the metric space  $(\Gamma, d)$ , with  $d$  a distance function, in order to quantify the locality properties of an interaction. For definitiveness,  $\Gamma = \mathbf{Z}^{\nu}$ ,  $\mathcal{A}_{\Gamma}^{\text{loc}}$  denotes a local algebra, i.e.  $\mathcal{A}_X$  with  $X \in P_0(\Gamma)$ , with, as before,  $P_0(\Gamma)$  denoting the collection of all finite parts of  $\mathbf{Z}^{\nu}$  and  $d(x, y) \equiv \sum_{i=1}^{\nu} |x_i - y_i|$ , for  $x = (x)_i, y = (y)_i, i = 1, \dots, \nu$ .  $F$  is assumed to be a non-increasing function  $F : [0, \infty) \rightarrow (0, \infty)$  satisfying the properties:

(i)  $F$  is uniformly integrable over  $\Gamma$ , i.e.

$$\|F\| \equiv \sup_{x \in \Gamma} \sum_{y \in \Gamma} F(d(x, y)) < \infty. \quad (31)$$

(ii)  $F$  satisfies the convolution condition

$$C_F \equiv \sup_{x, y \in \Gamma} \sum_{z \in \Gamma} \frac{F(d(x, z))F(d(z, y))}{F(d(x, y))} < \infty. \quad (32)$$

An equivalent formulation of (32) is that there is a constant  $C_F < \infty$  such that

$$\sum_{z \in \Gamma} F(d(x, z))F(d(z, y)) \leq C_F F(d(x, y)) \quad \forall x, y \in \Gamma. \quad (33)$$

The  $F$  functions describe the decay of a given interaction. Let  $F$  be an  $F$ -function on  $(\Gamma, d)$  and  $\Phi : P_0(\Gamma) \rightarrow \mathcal{A}_{\Gamma}^{\text{loc}}$  be an interaction. The  $F$ -norm of  $\Phi$  is defined by

$$\|\Phi\|_F = \sup_{x, y \in \Gamma} \frac{1}{d(x, y)} \sum_{Z \in P_0, x, y \in Z} \|\Phi(Z)\| \quad (34)$$

from which

$$\sum_{Z \in \mathfrak{P}_0, x, y \in Z} \|\Phi(Z)\| \leq \|\Phi\|_F F(d(x, y)). \quad (35)$$

An important question will arise in Sec. 3 in connection with the following issue: automorphisms of strictly local observables delocalize along the quasi-local algebra  $\mathcal{A}$ , as seen explicitly in [35, p. 2951] in the limit  $\Lambda \nearrow \infty$ , but there are (again by

W. F. Wreszinski

the same formulae) local approximations to them. The following theorem will be of crucial importance for us in Sec. 3.

**Theorem 2.2.** *Let  $\Lambda_n \subset \Lambda_{n+1}, \forall n \geq 1$  denote an exhaustive sequence of increasing and absorbing finite regions  $\{\Lambda\}_n$ , denoted  $\Lambda_n \nearrow \Gamma$ ,  $A$  be an element of the strictly local algebra  $\mathcal{A}_\Gamma^{\text{loc}}$  and*

$$\tau_t^\Lambda = \exp(iH_\Lambda t)A \exp(-iH_\Lambda t) \quad (36)$$

with  $H_\Lambda$  defined by (17), under the condition

$$\|\Phi\|_F < \infty \quad (37)$$

for a given  $F$ -function  $F$ . Then,

$$\|\tau_t^{\Lambda_n}(A) - \tau_t^{\Lambda_m}(A)\| \leq I_t(\Phi) \sum_{x \in X} \sum_{y \in \Lambda_n - \Lambda_m} F(d(x, y)), \quad (38)$$

where

$$I_t(\Phi) \equiv \frac{2\|A\|}{C_F} (2C_F|t|\|\Phi\|_F) \exp(2C_F|t|\|\Phi\|_F). \quad (39)$$

Above, we may assume that  $A \in \mathcal{A}_X$  for some  $X \in P_0(\mathbf{Z}^\nu)$ , and since the sequence is exhaustive,  $\exists N \geq 1$  for which  $X \subset \Lambda_n \forall n \geq N$ . The integers  $n, m$  in (41) satisfy  $N \leq m \leq n$ . Further

$$\tau_t(A) = \lim_{\Lambda \nearrow \infty} \tau_t^\Lambda(A) \quad \forall A \in \mathcal{A}_\Gamma^{\text{loc}} \quad (40)$$

exists in norm, and the convergence is uniform for  $t$  in compact subsets of  $\mathbf{R}$ .

**Proof.** See [24, Theorem 3.4 and (3.78) of Theorem 3.5]. The right-hand side of (38) converges to zero, as  $n, m \rightarrow \infty$ , by (31), assumption (37) and definition (34).  $\square$

We have the following basic corollary.

**Corollary 2.3.** *Theorem of Nachtergaele, Sims and Young on local approximations to automorphisms of local observables*

*Under the assumptions of Theorem 2.2,*

$$\|\tau_t^{\Lambda_n}(A) - \tau_t^{\Lambda_m}(A)\| \leq I_t(\Phi) \sum_{x \in X} \sum_{y \in \Gamma - \Lambda_m} F(d(x, y)). \quad (41)$$

We shall assume that our two-body translation-invariant interactions satisfy the condition  $P$  below:

$P$  The interactions  $\Phi$  have polynomial decay at infinity, i.e. for some  $0 < \beta < \infty$ ,

$$|\Phi(\{0, x\})| \leq \beta(1 + |x|)^{-\alpha}, \quad (42)$$

where  $|x| = d(x, 0)$  and

$$\alpha > \nu. \quad (43)$$

Condition (43) is due to the stability requirement (19).

In our applications in Sec. 3 we shall use Corollary 2.3. The right-hand side of (41) converges to zero, as  $m \rightarrow \infty$ , under assumption (37), but the set  $X$  in (41) must be also finally made to grow indefinitely, due to the definition of the mean entropy. This will require condition  $P$  for suitable  $\alpha$ , as we shall see in connection with the forthcoming inequality of Fannes.

## 2.4. A theorem of Fannes and a main auxiliary theorem

The entropy satisfies the subadditivity property [17]

$$S_{\Lambda_1 \cup \Lambda_2}(\omega) \leq S_{\Lambda_1}(\omega) + S_{\Lambda_2}(\omega) \quad (44)$$

for  $\Lambda_1 \cap \Lambda_2 = \emptyset$ .

Of great importance in Sec. 3 is the following theorem, due to Fannes [10], let  $\omega_1$  and  $\omega_2$  be two states on  $\mathcal{A}$ , i.e.  $\omega_1, \omega_2 \in E_{\mathcal{A}}$ , and for  $\Lambda_0 \subset \mathbf{Z}^\nu$ , let  $\omega_{1, \Lambda_0}$  and  $\omega_{2, \Lambda_0}$  denote the corresponding density matrices. Denote by  $\lambda_k^1, \lambda_k^2, k = 1, \dots, D^{|\Lambda_0|}$  the repeated eigenvalues of  $\omega_{1, \Lambda_0}$  and  $\omega_{2, \Lambda_0}$ , respectively, in ascending order. Define

$$\epsilon_k = |\lambda_k^1 - \lambda_k^2| \quad (45)$$

and

$$a = \sum_k \epsilon_k. \quad (46)$$

### Theorem 2.4.

$$(i) \ a \leq \sup_{\|A\| \leq 1, A \in \mathcal{A}_{\Lambda_0}} |\text{Tr}[(\omega_{1, \Lambda_0} - \omega_{2, \Lambda_0})A]|, \quad (47)$$

$$(ii) \ |S_{\Lambda_0}(\omega_1) - S_{\Lambda_0}(\omega_2)| \leq |\Lambda_0| a \log D + e, \quad (48)$$

$$a \leq \|\omega_1 - \omega_2\|. \quad (49)$$

By (48) and (49),

$$|s(\omega_1) - s(\omega_2)| \leq \|\omega_1 - \omega_2\| \log D. \quad (50)$$

**Proof.** See [10]. Equation (47) is the last inequality of the second line of his (1), and (48) is the last inequality before the final statement in [10]. Due to (49), which is also shown in [10], it follows that  $0 \leq a \leq 2$ , and the constant in (48) is thus bounded by  $\sup_{0 \leq a \leq 2} (2a - a \log a) = e$ .  $\square$

Theorem 2.4 has some important consequences.

W. F. Wreszinski

**Proposition 2.5.** *Let  $\Lambda_0$  be a finite fixed subset of  $\mathbf{Z}^\nu$ ,  $\Lambda \supset \Lambda_0$  and*

$$\omega^{\Lambda,t} \equiv \exp(-iH_\Lambda t) \omega_\Lambda \exp(iH_\Lambda t). \quad (51)$$

*Then, the following inequality holds:*

$$|S((\omega_t)_\Lambda) - S(\omega^{\Lambda,t})| \leq |\Lambda_0| \log D \sup_{A \in \mathcal{A}_{\Lambda_0}, \|A\|=1} \|\tau_{t,\Lambda,A}\| + c_\Lambda, \quad (52)$$

where

$$c_\Lambda \equiv e + 2|\Lambda - \Lambda_0| \log D \quad (53)$$

and  $A - B$  denotes the complement of  $B$  in  $A$ . Above,

$$\omega_t \equiv \omega \circ \tau_t \quad (54)$$

and

$$\tau_{t,\Lambda,A} \equiv \tau_t(A) - \exp(iH_\Lambda t) A \exp(-iH_\Lambda t). \quad (55)$$

**Proof.** By the subadditivity property of the entropy (44) and property (1) of the mean entropy

$$|S((\omega_t)_\Lambda) - S(\omega^{\Lambda,t})| \leq |S((\omega_t)_{\Lambda_0}) - S(\omega_{\Lambda_0}^{\Lambda,t})| + 2|\Lambda - \Lambda_0| \log D, \quad (56)$$

where  $\omega_{\Lambda_0}^{\Lambda,t}$  denotes the restriction of  $\omega^{\Lambda,t}$  to  $\mathcal{A}_{\Lambda_0}$ . By (48),

$$|S((\omega_t)_{\Lambda_0}) - S(\omega_{\Lambda_0}^{\Lambda,t})| \leq |\Lambda_0| d_\Lambda \log(D) + e \quad (57)$$

with

$$d_\Lambda \equiv \sup_{A \in \mathcal{A}_{\Lambda_0}, \|A\| \leq 1} |\mathrm{Tr}_{\mathcal{H}_{\Lambda_0}} [((\omega_t)_{\Lambda_0} - \omega_{\Lambda_0}^{\Lambda,t}) A]|. \quad (58)$$

But

$$\begin{aligned} & \sup_{A \in \mathcal{A}_{\Lambda_0}, \|A\| \leq 1} |\mathrm{Tr}_{\mathcal{H}_{\Lambda_0}} [((\omega_t)_{\Lambda_0} - \omega_{\Lambda_0}^{\Lambda,t}) A]| \\ & \leq \sup_{A \in \mathcal{A}_{\Lambda_0}, \|A\| \leq 1} |\mathrm{Tr}_{\mathcal{H}_{\Lambda_0}} \left[ \omega_\Lambda \left( \lim_{n \rightarrow \infty} \sigma_{n,\Lambda,A} - \exp(itH_\Lambda) A \exp(-itH_\Lambda) \right) \right]|, \end{aligned}$$

where

$$\sigma_{n,\Lambda,A} \equiv \mathrm{Tr}_{\Lambda_n - \Lambda} [\omega_{\Lambda_n - \Lambda} (\exp(itH_\Lambda) A \exp(-itH_\Lambda))].$$

The above inequality yields, together with (56)–(58), (52) and (53).  $\square$

We now use Proposition 2.5, together with (41), to arrive at the main auxiliary theorem.

**Theorem 2.6.** *Let  $\omega$  be a translation-invariant state on  $\mathcal{A}$ , with a dynamics  $\tau_t$ ;  $t \in \mathbf{R}$  generated by a family of Hamiltonians (17), satisfying (18) and (19). If the interactions  $\Phi$  have polynomial decay at infinity (42), with*

$$\alpha > 2\nu \quad (59)$$

it follows that the mean entropy  $s(\omega_t)$  of  $\omega_t$ , given by (54) satisfies

$$s(\omega_t) = \lim_{m \rightarrow \infty} \frac{1}{|\Lambda_{0,m}|} S(\exp(-itH_{\Lambda_{0,m}})\omega_{\Lambda_{0,m}} \exp(itH_{\Lambda_{0,m}})), \quad (60)$$

where  $\{\Lambda_{0,m}\}$  is a sequence of boxes tending to  $\mathbf{Z}^\nu$  as  $m \rightarrow \infty$ .

**Proof.** By definition,

$$s(\omega_t) = \lim_{m \rightarrow \infty} \frac{1}{|\Lambda_{0,m}|} S((\omega_t)_{\Lambda_{0,m}}). \quad (61)$$

We now pick  $\Lambda_0 = \Lambda_{0,m} \subset \Lambda = \Lambda_m$  in (52) and (53), where, for each  $m \geq 1$ ,

$$\lim_{m \rightarrow \infty} \frac{|\Lambda_{0,m}|}{|\Lambda_m|} = 1 \quad (62)$$

and

$$\lim_{m \rightarrow \infty} \frac{|\Lambda_m - \Lambda_{0,m}|}{|\Lambda_{0,m}|} = 0. \quad (63)$$

By (52) and (53) of Proposition 2.5, we obtain the inequality

$$|S((\omega_t)_{\Lambda_m}) - S(\omega_{\Lambda_m,t})| \leq |\Lambda_{0,m}|(\log D)f_m, \quad (64)$$

where

$$f_m \equiv \sup_{A \in \mathcal{A}_{\Lambda_{0,m}}, \|A\| \leq 1} \|\tau_{t,\Lambda_m,A}\| + 2|\Lambda_m - \Lambda_{0,m}| \log D, \quad (65)$$

where, by definition (54) and (55),

$$\|\tau_{t,\Lambda_m,A}\| = \|\tau_t(A) - \tau_t^{\Lambda_m}(A)\|. \quad (66)$$

We now use (42) and (59). A suitable  $F$ -function in this case is [24, Appendix, (A.9)]

$$F(|x|) = (1 + |x|)^{-2\nu - \epsilon} \quad (67)$$

for some  $\epsilon > 0$ , with  $|x| = d(x, 0)$ . By (31), (67) and (62), it readily follows from (41) that

$$\lim_{m \rightarrow \infty} \sup_{A \in \mathcal{A}_{\Lambda_{0,m}}, \|A\| \leq 1} \|\tau_t(A) - \tau_t^{\Lambda_m}(A)\| = 0. \quad (68)$$

Putting, now, together (52), (53), (63) and (68), we obtain (60).  $\square$

It is to be noted that (59) is stronger than the stability condition (43). The existence of local approximations to automorphisms of fixed local observables holds, by (41), for any stable interaction, but (60), which will be crucially needed in Sec. 3, requires a sufficiently fast decay of the interactions.

### 3. The Second Law of Thermodynamics as a Theorem Asserting the Growth of the Mean Entropy Under Adiabatic Transformations

#### 3.1. A refined form of the Penrose–Gibbs Theorem 1.1

The Penrose–Gibbs Theorem 1.1 may now be stated in a refined form, which incorporates the assumption of an infinite number of degrees of freedom.

**Theorem 3.1.** *A statistical-thermodynamical version of the second law of thermodynamics Let a quantum spin system, described by a finite-region Hamiltonian (17), and the state  $\omega$  fulfill the assumptions of Theorem 2.6. Assume further that  $\omega_t \equiv \omega \circ \tau_t$  satisfies the condition*

$$\lim_{t \rightarrow \infty} \omega_t = \bar{\omega}, \quad (69)$$

where the limit is taken in the weak\*-topology, that the mean entropy  $s$  exists and is defined by (16), and denote the initial state by  $\omega_0 = \omega$ . Then

$$s(\omega) \leq s(\bar{\omega}). \quad (70)$$

In words: the mean entropy of the initial state may increase, but cannot decrease, towards that of the “coarse-grained” state  $\bar{\omega}$  under an adiabatic transformation (that is, with  $\omega_t$  in (69) defined as in Definition 1.2).

**Proof.** We consider the sequence  $\{\omega_n\}_{n=0,1,2,\dots}$ , which a fortiori satisfies

$$\omega_n \rightarrow \bar{\omega} \quad \text{as } n \rightarrow \infty \quad (71)$$

in the weak\*-topology. By property (b) of the mean entropy,  $s$  is upper semicontinuous in the weak\*-topology and therefore,

$$\limsup_{n \rightarrow \infty} s(\omega_n) \leq s\left(\lim_{n \rightarrow \infty} \omega_n\right) = s(\bar{\omega}), \quad (72)$$

where the limit on the right-hand side of (72) is taken in the weak\*-topology. By (60) and the invariance of the trace  $\text{Tr}_{\mathcal{H}_{\Lambda_{0,m}}}$ , for each  $m \geq 1$ , under unitary transformations on the space  $\mathcal{H}_{\Lambda_{0,m}}$ , it follows that  $s(\omega_n) = s(\omega_0) = s(\omega)$ , and (70) results from (72).  $\square$

**Remark 3.1.** Theorem 3.1 answers in the affirmative the questions raised by Ruelle in [37, (iv), p. 1666].

#### 3.2. Applications to the gIm

The gIm corresponds to the strong interaction limit in the language of the lattice gas (see, e.g., [5, p. 425]): we expect that *interactions* are crucial for the approach (or return) to equilibrium! Therefore, in spite of being special, the present case does have certain physically sound aspects. Another example of this feature is the fact that the special case of the gIm considered in [8] describes well certain experiments

of free induction decays in solids [22]. This dynamics also received attention more than forty years later [6], and even quite recently in [23].

We consider the two subclasses of models of the gIm for  $\nu = 1$ , the exponential model  $E_\xi$  (22) (which a fortiori satisfies (42) and (59)), and the Dyson model  $D_\alpha$  (23), with  $\alpha$  satisfying (59). Then (26) holds, which is (69), with

$$\bar{\omega} \equiv \otimes_{x \in \mathbf{Z}} \text{Tr}_x \quad (73)$$

as long as  $\omega = \omega_0$  satisfies (25). As a prototypical example, take

$$\omega_0 = \otimes_{x \in \mathbf{Z}} P_x^{+,1}, \quad (74)$$

where

$$\sigma_x^{1P_x^{+,1}} = P_x^{+,1}. \quad (75)$$

We have the following.

**Corollary 3.2.** *For the exponential model  $E_\xi$  (22) and the Dyson model  $D_\alpha$  (23) with  $\alpha > 2$  and  $\omega_0$  given by (74), Theorem 3.1 is fulfilled in a non-trivial way, with  $s(\omega_0) = 0$  and  $s(\bar{\omega}) = \log 2$  (with  $S = 1/2$ ).*

**Proof.** The finite-volume version satisfies

$$\frac{S_\Lambda(\omega_0)}{|\Lambda|} = (-\lambda \log \lambda)_{\lambda=0} = 0$$

for all  $\Lambda \subset \mathbf{Z}$ , hence  $s(\omega_0) = 0$ . The finite volume version

$$\frac{S_\Lambda(\bar{\omega})}{|\Lambda|} = \left( -\sum_{i=1}^2 \lambda_i \log \lambda_i \right)_{\lambda_1=\lambda_2=1/2} = \log 2$$

for all  $\Lambda \subset \mathbf{Z}$ , hence  $s(\bar{\omega}) = \log 2$ . □

**Remark 3.2.** Note that the two values  $s(\omega_0) = 0$  and  $s(\bar{\omega}) = \log 2$  correspond to the two extreme values in property (1) of the mean entropy. Thus,  $\bar{\omega}$  corresponds to a maximum of the mean entropy, which is identified with the “infinite temperature state”. The initial state  $\omega_0$  may be seen as a nonlocal perturbation of the equilibrium state  $\bar{\omega}$ , if (69) is viewed as a return to equilibrium, or as a nonlocal perturbation of the ferromagnetic ground state  $\omega_g \equiv \otimes_{x \in \mathbf{Z}} P_x^{\pm,3}$ . In the latter case,  $\bar{\omega}$  would be viewed as a non-equilibrium stationary state. A model of the preparation of the system, suitable as the first step in Definition 1.2, consists in the following.

Schematically, a  $\text{CaF}_2$  crystal is placed in a magnetic field, thus determining the  $z$ -direction. Subsequently, a rf-pulse is applied to the sample, turning the net nuclear magnetization in the  $x$ -direction. The latter is then measured as a function of time: an oscillatory function slowly damping to zero ([22, Fig. 1.1] in the original experiment). See also [8].

## 4. Stability of the Second Law Under Interactions with the Environment

### 4.1. *The many histories picture strictly within Schrödinger quantum mechanics, K systems and their physical relevance*

In this section, we introduce the many-histories picture strictly within Schrödinger quantum mechanics, which will be used in our main theorem in this section. In the context of infinite systems, their introduction is due to Narnhofer and Thirring [29, 28], whose motivation was the proof of the macroscopic purification of states observed in Nature as a consequence of interactions with the environment. For this purpose, they needed the concept of (quantum) K system, introduced by Emch [9] and themselves (see [27] and references given there, as well as [40]).

A K system is a standard, and central, concept in the theory of classical dynamical systems (see, e.g., [42, Definition 4.7, p. 101]). A prototype of a K system is the baker's map [18, p. 74], which is also an Anosov system [2], and displays exponential instabilities leading to the mixing property (8). Such systems are “memoryless”, i.e. they possess a stochastic character which justifies equilibrium statistical mechanics. The microscopic mechanism of this loss of memory is the sensitive (exponential) dependence on initial conditions, produced by the “defocalizing shocks” between the particles (e.g., gas molecules) occurring in caricatures of interacting classical systems, such as the hard sphere gas, see [33, p. 1950] for a pedagogic introduction and early references.

An algebraic K system is a family of subalgebras  $\mathcal{A}_n, n \in \mathbf{Z}$  of an algebra  $\mathcal{A}$ , with the properties: (i)  $\mathcal{A}_{n+1} \supseteq \mathcal{A}_n$ ; (ii)  $\bigcup_n \mathcal{A}_n = \mathcal{A}$ ; (iii)  $\bigcap_n \mathcal{A}_n = z\mathbf{1}$  and an automorphism  $\sigma$  of  $\mathcal{A}$  with  $\sigma(\mathcal{A}_n) = \mathcal{A}_{n+1}$ .  $\bigcap_n$  means the set-theoretic intersection.

As in [27], we take  $\mathcal{A}_n$  as von Neumann algebras, with  $\bigcup_n$  meaning algebraic union together with strong closure in the representation with a given invariant state: the von Neumann K-system  $(\mathcal{A}_n, \sigma, \omega)$ . The isomorphism  $\sigma : \mathcal{A}_n \rightarrow \mathcal{A}_{n+1}$  has, by (iii), no non-trivial  $\sigma$ -invariant subalgebra  $\mathcal{B} \subset \mathcal{A}_n$ . We identify  $\sigma^{-1}$  with the time evolution, and assume that  $\omega$  is a KMS state with modular automorphism  $\sigma^{-t}$ , i.e.

$$\omega(AB) = \omega(\sigma^i(B)A) \quad \forall A, B \in \mathcal{A}. \quad (76)$$

We have to refer to [5, p. 84], for the above terminology, but remark that it will not be used in the sequel.

In the next subsection, we shall be concerned with the above mentioned generalization in [29, 40] of the many-histories interpretation of quantum mechanics [13, 11, 31] to (infinite) von Neumann K systems  $(\mathcal{A}_n, \sigma, \omega)$  with a set of projections  $P_\alpha \in \mathcal{A}, \alpha \in \{1, \dots, r\}$ . The operator  $\sigma^{-1}$  assigns to each  $P_\alpha$  a “time evolution”  $\sigma^{-t}(P_\alpha) = P_\alpha(t)$  and in this way, a sequence of “events”  $\{P_{\alpha_1}(t_1), P_{\alpha_2}(t_2), \dots, P_{\alpha_n}(t_n)\}$  (a “history”, briefly written  $\underline{\alpha}$  for the index set or the corresponding vector) and a probability distribution over the set  $\{\underline{\alpha}\}$  of histories

$$W(\underline{\alpha}) \geq 0 \quad \text{with} \quad \sum_{\underline{\alpha}} W(\underline{\alpha}) = 1 \quad (77)$$



and such that

$$W(\underline{\alpha}) \equiv \omega(P_{\alpha_1}(t_1) \cdots P_{\alpha_n}(t_n) \cdots P_{\alpha_1}(t_1)). \quad (78)$$

The analogue of (78) for systems with finite number of degrees of freedom is

$$\begin{aligned} W(\underline{\alpha}) &= \text{Tr} P_{\alpha_n}(t_n) \cdots P_{\alpha_1}(t_1) \rho P_{\alpha_1}(t_1) \cdots P_{\alpha_n}(t_n) \\ &= \text{Tr} \sqrt{\rho} P_{\alpha_1}(t_1) \cdots P_{\alpha_n}(t_n) P_{\alpha_{n-1}}(t_{n-1}) \cdots P_{\alpha_1}(t_1) \sqrt{\rho}. \end{aligned}$$

As Wightman observes in his review of the quantum theory of measurement [44], “The above formulae were already written down by Aharonov *et al.* [1] and used by them in a discussion of time reversal invariance in quantum measurement theory. These authors regarded the formulae as standard quantum mechanics”.

For the infinite systems the density matrix does not exist, but the notion of a state as a positive linear functional  $\omega$  as generalization of  $\text{Tr} \sqrt{\rho} A \sqrt{\rho} = \omega(A)$  carries over, with (78) replacing the standard formulae.

What we need for the main theorem of this section, Theorem 4.1, is the above framework for infinite K systems and histories. The Narnhofer–Thirring model used there is a K system (trivially) and the framework applies, as long as we have to do only with finite times, which is the case, there, too. The following results, which we present here for completeness, have to do with the limit of infinite times. They are, however, crucial, for the purification result of [29], which is the most important physical application of the theory.

The probability distributions (78) have the properties: (a) *decoherence*

$$\begin{aligned} &\omega(P_{\alpha_1}(t_1) \cdots P_{\alpha_n}(t_n) \cdots P_{\alpha'_1}(t_1) \cdots P_{\alpha'_n}(t_n)) \\ &\rightarrow \delta_{\underline{\alpha}, \underline{\alpha}'} \quad \text{as } t_i - t_{i+1} \rightarrow \infty \forall i, \end{aligned}$$

(b) *symmetry*  $W(\alpha_1, \dots, \alpha_n)$  tends to a symmetric function in the indices  $\alpha_1, \alpha_n$  as  $t_i - t_{i+1} \rightarrow \infty \forall i$ .

Property (a) is a consequence of the fact that K systems are asymptotically abelian, i.e. classical for large times (see [40] and references given there) and property (b) means that for long times the system forgets all causal links, in analogy to the “memoryless” character of classical K systems.

We have the following fundamental theorem of Narnhofer and Thirring [27, 40].

*Narnhofer–Thirring Theorem*

Let  $(\mathcal{A}_n, \sigma, \omega)$  be a von Neumann K system, with  $\sigma^{-1}$  the time evolution. Given a set  $P_1, \dots, P_r \in \mathcal{A}$ ,  $\epsilon > 0$  and  $n \in \mathbb{N}$ ,  $n < \infty$ , there exists a  $T < \infty$  such that for each history  $W(\underline{\alpha})$  given by (78),

$$\left| W(\underline{\alpha}) - \prod_{i=1}^n \omega(P_{\alpha_i}) \right| < \epsilon \quad (79)$$

whenever  $t_{i+1} - t_i > T \forall i \in [1, n]$ .

It is to be emphasized that Narnhofer and Thirring [29] and Thirring [40] use the history interpretation as a simple description of the essential effect of a measurement, but strictly within the Copenhagen interpretation. This follows because only the so-called *consistent histories* are considered, i.e. such that

$$\omega(P_{\underline{\alpha}'} P_{\underline{\alpha}}) = \delta_{\underline{\alpha}', \underline{\alpha}}. \quad (80)$$

This is necessary for an interpretation in terms of classical probabilities and is satisfied as soon as (79) holds.

It may be important, to avoid any confusion, to remark that Griffiths [13], see also [14, 15] bases his interpretation of quantum mechanics on the *probability* theory of consistent families of histories. This approach has resolved some important paradoxes in the conventional theory, see in this connection [15]. We thank Professor Griffiths for these remarks.

#### 4.2. Non-automorphic interactions with the environment

No physical system is entirely isolated, although isolation may be brought about, in principle, to an arbitrary degree of accuracy. Even concerning the Universe — the prototype of an isolated system — there will be “events”, i.e. interactions with the environment, which may be modeled by the state collapse of a given state, that is, non-automorphic evolutions. One such model is the Narnhofer–Thirring model [29, 28].

As a physical motivation, consider the question of the macroscopic purification of states by interactions with the environment [29], such as a magnet below the transition temperature (for a rigorous model of this situation, see [5, Theorem 6.2.48, p. 320]). It is known that domains of such a magnet are, in Nature, found in a definite direction. To quote [29] (see also [30]): “even if nobody looks at them, there will be enough ‘events’ (i.e. interactions with the environment) to purify the state over the classical part”. Such are prototypical of the unavoidable interactions of an (even to a good approximation “closed”) system with the environment, and were first shown to drive the state into a factor state by Narnhofer [25] and Narnhofer and Robinson [26]. We refer to Wightman’s review paper for superselection sectors due to environmental interactions of mesoscopic systems [45].

We consider as the classical quantity of the model the mean magnetization

$$\vec{m} = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{x=-N}^N \vec{\sigma}_x. \quad (81)$$

Classically, pure states are those in which all spins except a finite number of them point in the same direction  $\vec{m}$ . Below the phase transition temperature the state  $\omega$  is a mixture of pure, or more generally, extremal invariant states  $\omega_1, \omega_2$  (see [5, Theorem 6.2.48, p. 320]):

$$\omega = \mu\omega_1 + (1 - \mu)\omega_2 \quad \text{with } 0 < \mu < 1. \quad (82)$$

By (78), the histories are also convex combinations

$$W_\omega(\underline{\alpha}) = \mu W_{\omega_1}(\underline{\alpha}) + (1 - \mu) W_{\omega_2}(\underline{\alpha}). \quad (83)$$

It was proved in [29] that, for long histories  $n \rightarrow \infty$  and  $t_{i+1} - t_i \rightarrow \infty$ , and a dynamics described by (von Neumann) K-system they purify in the sense that in (83) either  $W_{\omega_1}$  or  $W_{\omega_2}$  dominates, i.e.  $\frac{W_{\omega_i}}{W_{\omega_j}} < \epsilon$  for arbitrarily small  $\epsilon$  and  $i \neq j$ : which one dominates depends on the history, although one cannot dominate over the other for all histories since

$$\sum_{\underline{\alpha}} W_{\omega_1}(\underline{\alpha}) = \sum_{\underline{\alpha}} W_{\omega_2}(\underline{\alpha}) = 1. \quad (84)$$

As in [29], consider the simplest non-trivial case with two states  $\omega_{1,2}$  and two projectors  $P, 1 - P$ , and denote  $\omega_{1,2}(P) = p_{1,2}$ . The model of a measuring apparatus will be defined next, following [29]. The algebra is the spin algebra  $\mathcal{A} = \{\vec{\sigma}_x\}_{x \in \mathbf{Z}}$ . The shift  $\vec{\sigma}_x \rightarrow \vec{\sigma}_{x+1}$  is taken as a discrete evolution. We first describe the system  $S$ . Let  $\Omega_{1,2}$  denote two disjoint states

$$\begin{aligned} |\vec{n}\rangle : \Omega_1 &= \otimes_{x=-\infty}^{\infty} |\vec{n}\rangle_x \\ \text{with } \vec{\sigma}_x \cdot \vec{n}_x |\vec{n}\rangle_x &= |\vec{n}\rangle_x \quad \text{with } \vec{n}_x^2 = 1; \\ |\vec{m}\rangle : \Omega_2 &= \otimes_{x=-\infty}^{\infty} |\vec{m}\rangle_x \quad \text{with } \vec{n} \neq \vec{m}. \end{aligned}$$

The two corresponding representations are denoted  $\Pi_1, \Pi_2$  on (incomplete tensor product) Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively. The corresponding mean magnetizations are

$$\vec{M}_{\vec{n}} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{x=-N}^N \Pi_1(\vec{\sigma}_x) = \vec{n} \mathbf{1} \quad (85)$$

and

$$\vec{M}_{\vec{m}} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{x=-N}^N \Pi_2(\vec{\sigma}_x) = \vec{m} \mathbf{1}. \quad (86)$$

The mixed state

$$\omega \equiv \mu \omega_1 + (1 - \mu) \omega_2 \quad (87)$$

is obtained by a vector in the orthogonal sum of  $\Pi_1$  and  $\Pi_2$ :

$$|\Omega_S\rangle = \sqrt{\mu} |\Omega_1\rangle \oplus \sqrt{1-\mu} |\Omega_2\rangle \in \mathcal{H}_1 \oplus \mathcal{H}_2 \equiv \mathcal{H}_S$$

$$\text{with } \Pi = \Pi_1 \oplus \Pi_2 \quad \text{and} \quad (\Omega_S, \Pi(\vec{\sigma}_x) \Omega) = \mu \vec{n} + (1 - \mu) \vec{m}.$$

The corresponding representation  $\Pi$  is reducible: there are two “superselection sectors” [45], the mean magnetization

$$\vec{M} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{x=-N}^N \Pi(\vec{\sigma}_x) = \mu \vec{n} \mathbf{1}_1 + (1 - \mu) \vec{m} \mathbf{1}_2$$

lies in the center  $Z = \Pi(\mathcal{A}) \cap \Pi(\mathcal{A})'$ , where the prime denotes the commutant, i.e. the set of bounded operators in the representation space which commute with  $\Pi(\mathcal{A})$ , which is not a multiple of unity. For the measuring device (“apparatus”)  $A$ , Narnhofer and Thirring also use a Hilbert space description, representing a device measuring one of two spin directions  $\vec{s}$  and  $-\vec{s}$  (“up” or “down”): the state of the device measuring the direction of  $\vec{\sigma}_i$  may be represented by a two-dimensional vector

$$\begin{bmatrix} u_i \\ d_i \end{bmatrix}$$

(pointer up and down). The measuring array is again an infinite tensor product  $\otimes_i \begin{bmatrix} u_i \\ d_i \end{bmatrix}$  which belongs to  $\mathcal{H}_A$  (where  $A$  stands for “apparatus”), and we start with a state  $|\Omega\rangle_A$ , where all  $u_i$  are zero. For the time evolution, we take a shift  $U$ , and then consider an instantaneous measurement of the direction  $\vec{s}$  of a spin, the corresponding projector being

$$P_k = \frac{1}{2}(1 + \vec{\sigma}_k \cdot \vec{s}) = |s\rangle_k \times |s\rangle_k^\dagger.$$

If the answer is one, we have the pointer unchanged, if the answer is zero, we turn the pointer up. The turning of the pointer is effected by an operator  $\tau$ ,

$$\tau_k \begin{bmatrix} u_i \\ d_i \end{bmatrix} = \begin{bmatrix} d_i \\ u_i \end{bmatrix}.$$

Thus, the effect of measuring  $\vec{\sigma}_1$  is

$$V_1 = P_1 + (1 - P_1)\tau_1$$

or written in full detail, with operators on  $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_A$ ,

$$V_1 = (\Pi_1(P_1) \oplus \Pi_2(P_1)) \otimes \mathbf{1} + (1 - \Pi_1(P_1)) \oplus ((1 - \Pi_2(P_1))) \otimes \tau_1. \quad (88)$$

The time evolution  $U$  between the measurements shifts by one unit

$$U\Pi_{1,2}(P_k) = \Pi_{1,2}(P_{k+1})U \quad \text{together with } U\tau_k = \tau_{k+1}U, \quad (89)$$

so that the full time evolution of  $|\Omega\rangle = |\Omega_S \otimes |\Omega\rangle_A$  after  $n$  time units is

$$|\Omega(n)\rangle = V_1 U V_1 U \cdots V_1 U |\Omega\rangle = V_1 V_2 \cdots V_n |\Omega\rangle \quad (90)$$

since  $U^k|\Omega\rangle = |\Omega\rangle$ . The results of the measurements are encoded in the  $\mathcal{H}_A$ -part of  $|\Omega(n)\rangle$ , so we decompose  $|\Omega(n)\rangle$ , given by (90) in an orthogonal basis of  $\mathcal{H}_A$ ,

$$|\Omega(n)\rangle = \sum_{\alpha_i=0}^1 v(\underline{\alpha}) \tau_1^{\alpha_1} \cdots \tau_n^{\alpha_n} |\Omega_A\rangle \quad (91)$$

with  $v(\underline{\alpha}) \in \mathcal{H}_S$ . Wherever  $\alpha_i = 0$ , the corresponding spin is in direction  $\vec{s}$ , for  $\alpha_i = 1$  we have  $-\vec{s}$ :

$$v(\underline{\alpha}) = v_1(\underline{\alpha}) \oplus v_2(\underline{\alpha}) \quad (92)$$

together with

$$P_1|n\rangle_1 = |s\rangle_1(s|n) \quad (93)$$

and

$$(1 - P_1)|n\rangle = |-s\rangle(-s|n). \quad (94)$$

If we introduce  $|(s|n)|^2 = p_1$ , thus  $|(-s|n)|^2 = 1 - p_1$ , and similarly  $|(s|m)|^2 = p_2$ ,  $|(-s|m)|^2 = 1 - p_2$ , and if  $\alpha$  contains  $l$  zeros and  $(n - l)$  ones, we have

$$\|v_1(\underline{\alpha})\|^2 = \mu p_1^l (1 - p_1)^{n-l}; \|v_2(\underline{\alpha})\|^2 = (1 - \mu) p_2^l (1 - p_2)^{n-l} \quad (95)$$

together with

$$W(\underline{\alpha}) = \|v(\underline{\alpha})\|^2. \quad (96)$$

Note that there is no collapse of the wave-function after each measurement — the evolution is unitary and only at the end the configuration of classical pointers is read. We shall refer to the above model as the *Narnhofer–Thirring model*.

### 4.3. Invariance of the mean entropy on the average and comparison with finite systems

We are now ready to present our main results. First, let us examine finite systems or more precisely, systems with a finite number of degrees of freedom. We consider, as in [30], a finite composite system consisting of a finite system together with a measurement apparatus, for example a finite version of the Narnhofer–Thirring model, described by a collection of projectors  $\{P_\alpha\}$ ,  $\alpha \in [1, n]$ , such that

$$P_\alpha P_{\alpha'} = P_\alpha \delta_{\alpha, \alpha'}. \quad (97)$$

Inspired by von Neumann ([41, Chap. V], see also the discussion in Wightman [44, p. 431] concerning degenerate spectra), it is natural to take the  $\{P_\alpha\}$  ( $\alpha = 1, \dots, n$ ) as “macroscopic observables”, and define a density matrix  $\rho$  as *decoherent* with respect to the set  $\{P_\alpha\}$  if

$$P_{\alpha'} \rho P_\alpha = 0 \quad \text{if } \alpha \neq \alpha'. \quad (98)$$

When a measurement is performed and the position of the pointer was observed, the probability of its position can only be evaluated if we repeat the measurement. In the individual observation it has a definite value, and the measured system is in a pointer position: reduction of the wave-function turns into the collapse of the wave-function. We have gained knowledge, not necessarily in the individual case, but on the average, and indeed, by [30, Lemma 3], the average von Neumann entropy  $S_{\text{av}}$ , defined by

$$S_{\text{av}} = \sum_{\alpha=1}^n \text{Tr}(\rho P_\alpha) S(\rho_\alpha), \quad (99)$$

where

$$S(\rho) = -\text{Tr}(\rho \log(\rho)) \quad (100)$$

(with  $k_B = 1$ ) is the von Neumann entropy, is *reduced*; above,

$$\rho_\alpha \equiv [\text{Tr}(\rho P_\alpha)]^{-1} P_\alpha \rho P_\alpha \quad (101)$$

is one of the collapsed states, obtained with probability  $\text{Tr}(\rho P_\alpha)$  through unitary evolution of an initial decoherent state, i.e. which satisfies (98). This follows from the strict concavity of the entropy of a finite system, i.e.

$$S_\Lambda(\alpha \rho_\Lambda^1 + (1 - \alpha) \rho_\Lambda^2) > \alpha S_\Lambda(\rho_\Lambda^1) + (1 - \alpha) S_\Lambda(\rho_\Lambda^2). \quad (102)$$

The mean entropy is, however, *affine* [17]:

$$s(\alpha \omega_1 + (1 - \alpha) \omega_2) = \alpha s(\omega_1) + (1 - \alpha) s(\omega_2). \quad (103)$$

For the statistical-thermodynamical significance of the property of affinity, see [39, p. 60].

We now consider the Narnhofer–Thirring model. After  $n$  (time) steps, we use the notation

$$\underline{\alpha}_n = (\alpha_1, \dots, \alpha_n). \quad (104)$$

**Theorem 4.1.** *Let  $\omega$  be the state (87). Then, in the Narnhofer–Thirring model with initial state  $\omega$ , the mean entropy is conserved on the average by the collapse of  $\omega$ :*

$$s_{\text{av}} = s(\omega), \quad (105)$$

where

$$s_{\text{av}} \equiv \sum_{\underline{\alpha}_n} W_\omega(\underline{\alpha}_n) s(\omega_{\underline{\alpha}_n}) \quad (106)$$

and

$$\omega_{\underline{\alpha}_n} \equiv \omega(P_{\underline{\alpha}_n} \cdot) [W_\omega(\underline{\alpha}_n)]^{-1}. \quad (107)$$

**Proof.** Reading the pointer in a position  $\underline{\alpha}_n$ , i.e. after  $n$  time steps, the initial state collapses to (107). The property of affinity (103) now implies that

$$\begin{aligned} s_{\text{av}} &= \sum_{\underline{\alpha}_n} W_\omega(\underline{\alpha}_n) s(\omega_{\underline{\alpha}_n}) \\ &= s\left(\sum_{\underline{\alpha}_n} W_\omega(\underline{\alpha}_n) [W_\omega(\underline{\alpha}_n)]^{-1} \omega(P_{\underline{\alpha}_n})\right) \\ &= s(\omega) \end{aligned}$$

by linearity of  $\omega$  and the fact that  $\sum_{\underline{\alpha}_n} P_{\underline{\alpha}_n} = 1$ . □

As a consequence of Theorem 4.1, we have the following.

**Corollary 4.2.** *The statistical thermodynamical version of the second law of thermodynamics Theorem 3.1 is stable under non-automorphic interactions with the environment, exemplified by the Narnhofer–Thirring model, in sharp contrast to the second law of thermodynamics for the quantum Boltzmann entropy [30].*

The last sentence of Corollary 4.2 follows from [30, Lemma 3].

#### 4.4. Physical interpretation

In his paper “Against measurement”, John Bell [3] insisted on the necessity of physical precision regarding such words as reversible, irreversible, information (whose information? information about what?).

In the adiabatic transformation there is a first step, a finite preparation time during which external forces act, at the end of which the Hamiltonian associated to the initial equilibrium state is restored, and remains so “forever” during the second step. The dynamics of preparation of the state is therefore *not* time-reversal invariant, leading to the *time arrow* mentioned in Remark 1.1.

Given a time arrow, the process  $\omega_1(0) \rightarrow \omega_2(\infty)$  is reversible (irreversible) if and only if the inverse process  $\omega_2(0) \rightarrow \omega_1(\infty)$  is possible (impossible). The first alternative takes place if and only if  $s(\omega_1) = s(\omega_2)$ , the second one if and only if  $s(\omega_1) < s(\omega_2)$ . Infinite time  $t = \infty$  means that  $t$  is much larger than a quantity  $t_D$ , the relaxation time, or decoherence time in measurement theory. Of course, irreversibility is incompatible with time-reversal invariance, because the entropy or the mean entropy cannot both strictly increase and strictly decrease with time. This is the Schrödinger paradox [38], cited in Lebowitz’s inspiring review of the issue of time-assymetry [19].

Entropy  $S_\Lambda = |\Lambda| \log D - I_\Lambda$ , with  $I_\Lambda$  denoting the (quantum) information. For quantum spin systems  $0 \leq S_\Lambda/|\Lambda| \leq \log D$ , and therefore,  $0 \leq I_\Lambda/|\Lambda| \leq \log D$ . It attains its maximum value for pure states, which are characterized by  $S_\Lambda = 0$ . Under “collapse”, each collapsed state is pure and therefore, information is gained: this explains that the (Boltzmann and von Neumann) entropies are reduced, on the average, violating the second law (on the average). Equivalently, entropy is, in Boltzmann’s sense, a measure of a macrostate’s wealth of “microstates”, and therefore grows by mixing, but it turns out that this growth is *not* extensive and disappears upon division by  $|\Lambda|$ , i.e. taking the infinite volume limit (inequalities of Lanford and Robinson [17]), so that the affinity property (103) results and with it, the stability of the second law in the form of Theorem 3.1 under interactions with the environment (Corollary 4.2).

## 5. Conclusion

One central and dominating feature of the analysis over finite versus infinite dimensional spaces is that in the infinite dimensional case the solution may depend

*discontinuously* on the parameters of the problem. Indeed, infinite systems may exhibit *singularities*, not present in finite macroscopic systems, well-known in the theory of *phase transitions*: they are parametrized by *critical exponents*, which, moreover, display *universal* properties, in excellent agreement with experiment! Another example is furnished by the upper-semicontinuity of the mean entropy (versus continuity of the finite entropy).

In greater generality, the physical “N large but finite” differs qualitatively from “N infinite” because the latter exhibits *universal* properties not found in finite systems. Two examples of these universal properties, crucial in our approach, are the upper semicontinuity and affinity of the mean entropy, whose finite-volume counterparts (continuity of  $\frac{S_\Lambda}{|\Lambda|}$  and strict concavity of  $\frac{S_\Lambda}{|\Lambda|}$ ) are not universal because, not being uniform in  $|\Lambda|$ , they depend on the volume  $|\Lambda|$  of the system. The fact that (only) “N infinite” is in good agreement with experiment is explained by the fact that, with  $N \approx 10^{24}$ , macroscopic systems are extremely close to infinite systems (the success of the thermodynamic limit!)

It is well known that the thermodynamic entropy, which is relevant to the second law, does not, in general, coincide with the (microscopic) Gibbs–von Neumann entropy, except for a limited class of states which includes the standard equilibrium states. Our approach introduces two types of coarse-graining in the entropy; the first, in space, by building the mean entropy. This enables consideration of states of infinite systems which may (and do sometimes) have non-trivial ergodic properties: starting from an unstable state at an initial time  $t = 0$ , they may approach a different (“equilibrium”) state as  $t \rightarrow \infty$  — i.e. there is also a “coarse graining” in time. The important issue here is the fact that a sequence of pure states may tend to a mixed state, a fact first observed by Hepp [16], see also [29], and valid only for states of infinite systems. In the examples given, the first step of the adiabatic transformation — which is the thermodynamically non-trivial step — yields no change in the mean entropy (it remains equal to the initial value of zero), it is only the second step, which is thermodynamically trivial — a relaxation — which yields a change in the specific entropy. This feature is not realistic and is due to the extreme scarcity of models in which the time evolution of states is under control. Both Theorems 3.1 and 4.1 are, however, of general validity, because of the universal properties mentioned above, and of the natural type of coarse-graining associated to the use of the mean entropy (a density).

There is what seems to us to be a close analogy with the situation in classical dynamical systems as regards the intuitive reason for this possible growth of the mean entropy with time. If one looks at the entropy  $S(U_t, V_t)$  of a system with finite number of degrees of freedom,  $(U_t, V_t)$  are analogous to the individual “trajectories” of a classical map, which, for a (time-reversible) K system, such as the baker map [18, Fig. 1.2.4 and Sec. 1.3] are highly erratic and do not tend to any limit as  $t \rightarrow \infty$ . On the other hand, in  $s(u_t, v_t)$ , where  $s$  is the mean entropy and  $u, v$  the energy and volume densities, the  $u, v$  densities may be expected to behave qualitatively as



the density  $\rho_t$ , e.g., in the baker transformation [18, Example 4.1.3, p. 48],

$$\lim_{t \rightarrow \infty} \rho_t = \bar{\rho} \equiv 1, \quad (108)$$

Eq. (108) is analogous to (129), where, for  $t = n = 1, 2, \dots$ , i.e. discrete time units,

$$\rho_n = P^n \rho_0 \quad (109)$$

and  $P$  denotes the Ruelle–Perron–Frobenius operator [18] which maps densities to densities, and  $\rho_0$  denotes an arbitrary initial density

$$\rho_0 \in L^1(X), \quad \rho_0 \geq 0, \quad \int_X \rho_0(x) dx = 1 \quad (110)$$

with  $X$  denoting the unit square in  $\mathbf{R}^2$ . By (110),  $\rho_0$  cannot be a delta function, i.e. individual trajectories are excluded.  $P$  is a Markov operator [18], with a unique fixed point, which is the function

$$\bar{\rho} = 1 \forall x \in X \quad (111)$$

(analogous to the uniform “microcanonical” distribution on the square). The reduction of information which results from building a density provides a “restoring force” towards equilibrium (111), with maximal disorder or entropy. See also [46].

Our main new result in this paper is that non-automorphic interactions with the environment, although known to produce on the average a strict reduction of the Boltzmann entropy of systems with finite number of degrees of freedom [30, Lemma 3], are proved to conserve the mean entropy on the average, as a consequence of the latter’s property of affinity. As a consequence, we do not need to assume that such interactions are rare on the thermodynamic time scale, in order to assure the validity of the second law (see the remarks on the last paragraph of [30]) and *stability* with regard to interactions with the environment follows instead (Corollary 4.2). We view this as an additional important indication of the naturalness of the present approach, suggesting that it may turn out to be of general validity, including classical mechanics (see the introduction) and relativistic quantum field theory, but serious open problems of technical nature remain in both cases.

## Acknowledgments

I am greatly indebted to Bruno Nachtergaele, who provided me with two crucial elements of the proof of Theorem 4.1: his theorem (together with Sims and Young) on local approximations to automorphisms of local observables, and Fannes’ beautiful inequality.

I also thank very heartily Heide Narnhofer for the very enjoyable and instructive collaboration [30].

The present paper owes very much to several people: Elliott Lieb, for his encouraging remarks on the related [46], Derek Robinson for a correspondence in which he mentioned the property of affinity, Prof. D. Ruelle for his encouraging remarks in correspondence and for asking the “right questions” in [37], Professor Oliver

Penrose for correspondence on the Penrose–Gibbs theorem, and Geoffrey Sewell for helping me to formulate the time-arrow problem in a most concise way.

I should also like to thank an anonymous referee for helpful remarks concerning a previous version of Definition 1.2.

## References

- [1] Y. Aharonov, P. G. Bergmann and J. L. Lebowitz, Time symmetry in the quantum process of measurement, *Phys. Rev. B* **134** (1964) B1410–B1416.
- [2] D. V. Anosov, Ergodic properties of geodesic flows on closed Riemannian manifolds of negative curvature, *Proc. Steklov Inst. Math.* **90** (1967) 1–235.
- [3] J. S. Bell, Against measurement, *Phys. World* **3** (1990) 33–41.
- [4] P. Blanchard and E. Brüning, *Variational Methods in Mathematical Physics: A Unified Approach* (Springer Verlag, Berlin, 1992).
- [5] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol. 2, 2nd edn. (Springer Verlag, 1997).
- [6] H. Cho, T. D. Ladd, J. Baugh, D. G. Cory and C. Ramanathan, Multi-spin dynamics of the solid state NMR free induction decay, *Phys. Rev. B* **72** (2005) 054427.
- [7] F. J. Dyson, Existence of a phase transition in a one-dimensional Ising ferromagnet, *Comm. Math. Phys.* **12** (1969) 91–107.
- [8] G. G. Emch, Nonmarkovian model of the approach to equilibrium, *J. Math. Phys.* **7** (1966) 1198–1206.
- [9] G. G. Emch, Generalized  $K$  flows, *Comm. Math. Phys.* **49** (1975) 191–215.
- [10] M. Fannes, A continuity property of the entropy density for quantum lattice systems, *Comm. Math. Phys.* **31** (1973) 291–294.
- [11] M. Gell-Mann and J. B. Hartle, Alternative decohering histories in quantum mechanics, in *Proc. 25th Int. Conf. High Energy Physics*, eds. K. K. Phua and Y. Yamaguchi (World Scientific, Singapore, 1991), pp. 1303–1310.
- [12] J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Dover, New York, 1960).
- [13] R. B. Griffiths, Consistent histories and the interpretation of quantum mechanics, *J. Stat. Phys.* **36** (1984) 219–272.
- [14] R. B. Griffiths, The consistent histories approach to quantum mechanics, in *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, June 6, 2019).
- [15] R. B. Griffiths, Bell inequalities versus quantum physics: A reply to Lambare, *Phys. Rev. A* **104** (2021) 066202.
- [16] K. Hepp, Quantum theory of measurement and macroscopic observables, *Helv. Phys. Acta* **45** (1972) 237–248.
- [17] O. E. Lanford and D. W. Robinson, Mean entropy of states in quantum statistical mechanics, *J. Math. Phys.* **9** (1968) 1120–1125.
- [18] A. Lasota and M. C. Mackey, *Probabilistic Properties of Deterministic Systems* (Cambridge University Press, 1985).
- [19] J. L. Lebowitz, Time-asymmetric macroscopic behavior: An overview, in *Boltzmann's Legacy*, eds. G. Gallavotti *et al.* (European Mathematical Society, 2008).
- [20] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, *Comm. Math. Phys.* **28** (1972) 251–257.
- [21] E. H. Lieb and J. Yngvason, The physics and mathematics of the second law of thermodynamics, *Phys. Rep.* **310** (1999) 1–96.
- [22] I. J. Lowe and R. E. Norberg, Free induction decays in solids, *Phys. Rev.* **107** (1957) 46–61.

- [23] J. R. Maze, A. Dréau, V. Waselowski, H. Duarte, J. F. Roch and V. Jacques, Free induction decays of single spins in diamond, *New J. Phys.* **14** (2012) 103041.
- [24] B. Nachtergaele, R. Sims and A. Young, Quasi-locality bounds for quantum lattice systems I — Lieb–Robinson bounds, quasi-local maps and spectral flow automorphisms, *J. Math. Phys.* **60** (2019) 061101.
- [25] H. Narnhofer, Stability of pure thermodynamical phases in quantum statistics, in *Phase Transitions*, ed. R. Kotecky (World Scientific, Singapore, 1993), pp. 150–158.
- [26] H. Narnhofer and D. W. Robinson, Dynamical stability and pure thermodynamical phases, *Comm. Math. Phys.* **41** (1975) 89–97.
- [27] H. Narnhofer and W. Thirring, Clustering for algebraic  $K$  systems, *Lett. Math. Phys.* **30** (1994) 307–316.
- [28] H. Narnhofer and W. Thirring, Why Schrödinger’s cat is most likely to be alive or dead, ESI preprint (1996) 34.
- [29] H. Narnhofer and W. Thirring, Macroscopic purification of states by interactions, in *On Quanta, Mind and Matter*, eds. A. Amann, H. Armanspacher and U. Müller-Herold (Kluwer Academic Publishers, 1999), pp. 105–118.
- [30] H. Narnhofer and W. F. Wreszinski, On reduction of the wave packet, decoherence, irreversibility and the second law of thermodynamics, *Phys. Rep.* **541** (2014) 249–278.
- [31] R. Omnès, *The Interpretation of Quantum Mechanics* (Princeton University Press, 1994).
- [32] R. E. Peierls, *Surprises in Theoretical Physics* (Princeton University Press, 1979).
- [33] O. Penrose, Foundations of statistical mechanics, *Rep. Progr. Phys.* **42** (1979) 1937–2006.
- [34] W. Pusz and S. L. Woronowicz, Passive states and KMS states for general quantum systems, *Comm. Math. Phys.* **58** (1978) 273–290.
- [35] C. Radin, Approach to equilibrium in a simple model, *J. Math. Phys.* **11** (1970) 2945–2955.
- [36] D. W. Robinson and D. Ruelle, Mean entropy of states in classical statistical mechanics, *Comm. Math. Phys.* **5** (1967) 288–300.
- [37] D. Ruelle, States of classical statistical mechanics, *J. Math. Phys.* **8** (1967) 1657–1668.
- [38] E. Schrödinger, The spirit of science, in *What is Life? And Other Scientific Essays* (Doubleday Anchor Books, Garden City, New York, 1965), pp. 229–250.
- [39] G. L. Sewell, *Quantum Theory of Collective Phenomena* (Oxford University Press, 1986).
- [40] W. Thirring, The histories of chaotic quantum systems, *Helv. Phys. Acta* **69** (1996) 706–716.
- [41] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1996).
- [42] P. Walters, *Ergodic Theory - Introductory Lectures*, Lecture Notes in Mathematics, Vol. 458 (Springer, 1975).
- [43] A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50** (1978) 221–260.
- [44] A. S. Wightman, Some comments in the quantum theory of measurement, in *Probabilistic Methods in Mathematical Physics*, eds. F. Guerra, M. Ioffredo and C. Marchioro (World Scientific, 1992), pp. 411–438.
- [45] A. S. Wightman, Superselection sectors: Old and new, *Nuovo Cim. B* **110** (1995) 751–769.
- [46] W. F. Wreszinski, One or two small points in thermodynamics, preprint (2020), arXiv:2003.03800.
- [47] W. F. Wreszinski, Irreversibility, the time arrow and a dynamical proof of the second law of thermodynamics, *Quantum Stud. Math. Found.* **7** (2020) 125–136.