ENTROPY PRODUCTION IN QUANTUM SPIN SYSTEMS.

(revised Nov. 2000)

by David Ruelle*.

Abstract. We consider a quantum spin system consisting of a finite subsystem connected to infinite reservoirs at different temperatures. In this setup we define nonequilibrium steady states and prove that the rate of entropy production in such states is nonnegative.

Keywords: statistical mechanics, nonequilibrium, entropy production, quantum spin systems, reservoirs.

* IHES. 91440 Bures sur Yvette, France. <ruelle@ihes.fr>

For several decades, Joel Lebowitz has been the soul of research in statistical mechanics. He now plays a central role in the development of new ideas which reshape our understanding of nonequilibrium. The present paper, dedicated to Joel on his 70-th birthday, extends some of the new ideas to quantum systems.

Introduction.

Consider a physical situation where a "small" system S is connected to different "large" heat reservoirs R_a ($a = 1, 2, \ldots$) at different inverse temperatures β_a. We want to define nonequilibrium steady states for the total system $L = S + R_1 + R_2 + \ldots$, and verify that the rate of entropy production in such states is ≥ 0. The model which we discuss in this paper is that of a fairly realistic quantum spin system. In what follows we first describe the model and state our assumptions (A1), (A2), (A3). In this setup we introduce nonequilibrium steady states ρ as states which, in the distant past, described noninteracting reservoirs at different temperatures. Under suitable conditions we check that our definition does not depend on where we place the boundary between the small system and the reservoirs. Our definition of the entropy production e_ρ also does not depend on where the boundary between the small system and the reservoirs is placed. With this definition we prove $e_\rho \geq 0$. By contrast with an earlier paper [4], we omit here assumptions of asymptotic abelianness in time which are difficult to verify, the definition of nonequilibrium steady states is more general, but we obtain less specific results.

Description of the model.*

Let L be a countably infinite set. For each $x \in L$, let $\mathcal{H}_x$ be a finite dimensional complex Hilbert space, and write $\mathcal{H}_X = \otimes_{x\in X}\mathcal{H}_x$ if X is a finite subset of L. We let $\mathcal{A}_X$ be the C*-algebra of bounded operators on $\mathcal{H}_X$, and if $Y \subset X$ we identify $\mathcal{A}_Y$ with a subalgebra of $\mathcal{A}_X$ by the map $\mathcal{A}_Y \mapsto \mathcal{A}_Y \otimes \mathbf{1}_{\mathcal{H}_{X\setminus Y}} \subset \mathcal{A}_X$. We write L as a finite union $L = \cup_{a\geq 0} R_a$, where $R_0 = S$ is finite (*small system*) and the R_a with $a > 0$ are infinite (*reservoirs*). We can then define the *quasilocal* C* algebras $\mathcal{A}_a$, $\mathcal{A}$ as the norm closures of

$$\bigcup_{X\subset R_a} \mathcal{A}_X \quad , \quad \bigcup_{X\subset L} \mathcal{A}_X$$

repectively. Note that all these algebras have a common unit element $\mathbf{1}$. In this setup we assume that an *interaction* $\Phi : X \mapsto \Phi(X)$ is given such that $\Phi(X)$ is a selfadjoint element of $\mathcal{A}_X$ for every finite $X \subset L$. Also, for each reservoir, we prescribe an inverse temperature $\beta_a > 0$ and a state σ_a on $\mathcal{A}_a$.

The assumptions (A1), (A2), (A3).

(A1) *The interaction* Φ *satisfies*

$$||\Phi||_\lambda = \sum_{n\geq 0} e^{n\lambda} \sup_{x\in L} \sum_{X\ni x:\mathrm{card}X=n+1} ||\Phi(X)|| < \infty$$

* See [3], [1].

for some $\lambda > 0$.

The importance of this assumption is that it allows us to equip $\mathcal{A}$ with a one-parameter group (α^t) of automorphisms* defining a *time evolution*. Introduce a linear operator $\delta : \cup_{X\subset L}\mathcal{A}_X \to \mathcal{A}$ such that

$$\delta A = i \sum_{Y:Y\cap X\neq\emptyset} [\Phi(Y), A] \qquad \text{if} \qquad A \in \mathcal{A}_X$$

If $A \in \mathcal{A}_X$, one checks that

$$||\delta^m A|| \leq ||A|| e^{\lambda \mathrm{card} X} m! (2\lambda^{-1}||\Phi||_\lambda)^m$$

The strongly continuous one-parameter group (α^t) of $*$-automorphisms of $\mathcal{A}$ is given by

$$\alpha^t A = \sum_{m=0}^{\infty} \frac{t^m}{m!} \delta^m A$$

if $A \in \cup_{X\subset L}\mathcal{A}_X$ and $|t| < \lambda/2||\Phi||_\lambda$. (More generally one could take $A \in \mathcal{A}_\lambda$, where $\mathcal{A}_\lambda$ is defined in the Appendix). Let

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

for finite $\Lambda \subset L$. Writing $\Lambda \to L$ if Λ eventually contains each finite $X \subset L$ we have, assuming $A \in \mathcal{A}$,

$$\lim_{\Lambda\to L} ||e^{itH_\Lambda} A e^{-itH_\Lambda} - \alpha^t A|| = 0$$

uniformly for t in compact intervals of $\mathbf{R}$.

(A2) $\Phi(X) = 0$ *if* $X \cap S = \emptyset$, $X \cap R_a \neq \emptyset$, $X \cap R_b \neq \emptyset$ *for different* $a, b > 0$.

Note that the description of the interaction Φ is somewhat ambiguous because anything ascribed to $\Phi(X)$ might also be ascribed to $\Phi(Y)$ for $Y \supset X$. Condition (A2) means that in our accounting, if a part of the interaction connects two different reservoirs, it must also involve the small system S.

(A3) *If* $a > 0$, *let* Φ_a *be the restriction of the interaction* Φ *to subsets of* R_a *and write*

$$H_{a\Lambda} = \sum_{X\subset R_a\cap\Lambda} \Phi_a(X) = H_{R_a\cap\Lambda}$$

Let also the interactions $\Psi_{(\Lambda)}$ *be given such that*

$$||\Psi_{(\Lambda)}||_\lambda \leq K < \infty \tag{1}$$

---

* See [1] Theorem 6.2.4 (or [3] Section 7.6).

and write

$$B_{a\Lambda} = \sum_{X \subset R_a \cap \Lambda} \Psi_{(\Lambda)}(X)$$

We assume that, for a suitable sequence $\Lambda \to L$,

$$\lim_{\Lambda \to L} \frac{\mathrm{Tr}_{\mathcal{H}_{R_a \cap \Lambda}} (e^{-\beta_a (H_{a\Lambda} + B_{a\Lambda})} A)}{\mathrm{Tr}_{\mathcal{H}_{R_a \cap \Lambda}} e^{-\beta_a (H_{a\Lambda} + B_{a\Lambda})}} = \sigma_a(A)$$

if $A \in \mathcal{A}_a$*: this defines a state* σ_a *on* $\mathcal{A}_a$*, depending on the choice of* $(\Psi_{(\Lambda)})$ *and the sequence* $\Lambda \to L$*. Furthermore we assume that for each finite* X *there is* Λ_X *such that* $\Psi_{(\Lambda)}(Y) = 0$ *if* $\Lambda \supset \Lambda_X$ *and* $Y \subset X$*; therefore*

$$||[B_{a\Lambda}, A]|| = 0 \tag{2}$$

if $\Lambda \supset \Lambda_X$ *and* $A \in \mathcal{A}_X$.

In particular we can take all $\Psi_{(\Lambda)} = 0$. Using (3) below, it is readily verified that σ_a is a β_a-KMS state (see [2]) for the one-parameter group $(\breve{\alpha}_a^t)$ of automorphisms of $\mathcal{A}_a$ corresponding to the interaction Φ_a. [I do not know which of the β_a-KMS states can be obtained in this manner].

Note that the assumptions (A1), (A2), (A3) can be explicitly verified in specific cases. From (A3) we obtain the following result.

Lemma.

$$\lim_{\Lambda \to L} ||e^{it(H_{a\Lambda} + B_{a\Lambda})} A e^{-it(H_{a\Lambda} + B_{a\Lambda})} - \breve{\alpha}_a^t A|| = 0 \tag{3}$$

for $a > 0$*, and*

$$\lim_{\Lambda \to L} ||e^{it(H_\Lambda + \sum_{a>0} B_{a\Lambda})} A e^{-it(H_\Lambda + \sum_{a>0} B_{a\Lambda})} - \alpha^t A|| = 0 \tag{4}$$

uniformly for t *in compact intervals of* $\mathbf{R}$.

We prove (4). Write $\alpha_\Lambda^t A = e^{it(H_\Lambda + \sum_{a>0} B_{a\Lambda})} A e^{-it(H_\Lambda + \sum_{a>0} B_{a\Lambda})}$ and $\delta_\Lambda A = i[H_\Lambda + \sum_{a>0} B_{a\Lambda}, A]$. If $A \in \cup_X \mathcal{A}_X$ we see using (1) that

$$\alpha_\Lambda^t A = \sum_{m=0}^{\infty} \frac{t^m}{m!} \delta_\Lambda^m A$$

converges uniformly in Λ for $|t| < \lambda / 2(||\Phi||_\lambda + K)$. Using also (2), it is shown in the Appendix that $\delta_\Lambda^m A \to \delta^m A$ in $\mathcal{A}$ when $\Lambda \to L$. Therefore

$$\lim_{\Lambda \to L} ||\alpha_\Lambda^t A - \alpha^t A|| = 0$$

when $A \in \cup_X \mathcal{A}_X$, uniformly for $|t| \leq T < \lambda/2(||\Phi||_\lambda + K)$. But the condition $A \in \cup_X \mathcal{A}_X$ is removed by density, and the condition $|t| \leq T < \lambda/2(||\Phi||_\lambda + K)$ by use of the group property. The proof of (3) is similar. $\square$

The KMS state σ.

The interaction $\sum_{a>0} \beta_a \Phi_a$, evaluated at X is $\beta_a \Phi_a(X)$ if $X \subset R_a$ and 0 if X is not contained in one of the R_a. The corresponding one-parameter group (β^t) of automorphisms of $\mathcal{A}$ has, according to (A3), the KMS state* $\sigma = \otimes_{a \geq 0} \sigma_a$ where σ_0 is the normalized trace on $\mathcal{A}_0 = \mathcal{A}_S$. In fact

$$\sigma(A) = \lim_{\Lambda \to L} \frac{\mathrm{Tr}_{\mathcal{H}_\Lambda}(\exp(-\sum_a \beta_a (H_{a\Lambda} + B_{a\Lambda}))A)}{\mathrm{Tr}_{\mathcal{H}_\Lambda} \exp(-\sum_a \beta_a (H_{a\Lambda} + B_{a\Lambda}))} \tag{5}$$

Nonequilibrium steady states.

We call *nonequilibrium steady states* (NESS) associated with σ the limits when $T \to \infty$ of

$$\frac{1}{T} \int_0^T dt\, (\alpha^t)^* \sigma$$

using the w^*-topology on the dual $\mathcal{A}^*$ of $\mathcal{A}$. With respect to this topology, the set Σ of NESS is compact, nonempty, and the elements of Σ are $(\alpha^t)^*$-invariant states on $\mathcal{A}$.

This definition generalizes that given in [4] where, under stringent asymptotic abelianness conditions, the existence of a single NESS was obtained.

Dependence on the decomposition $L = S + R_1 + R_2 + \ldots$**

Our definition of σ, and therefore of Σ depends on the choice of a decomposition of L into small system and reservoirs. If S is replaced by a finite set $S' \supset S$ and the R_a by correspondingly smaller sets $R'_a \subset R_a$ one checks that (A1), (A2), (A3) remain valid. If Φ'_a is the restriction of Φ to subsets of R'_a, the replacement of $\sum \beta_a \Phi_a$ by $\sum \beta_a \Phi'_a$ changes (β^t) to a one-parameter group (β'^t) and σ to a state σ'. These changes are in fact bounded perturbations covered by Theorem 5.4.4 and Corollary 5.4.5 of [1]. The map $\sigma \to \sigma'$ (of KMS states for (β^t) to KMS states for (β'^t)) is nonlinear (as can be guessed from (5)) and therefore we cannot expect that $\frac{1}{T} \int_0^T dt\, (\alpha^t)^* \sigma'$ has the same limit as $\frac{1}{T} \int_0^T dt\, (\alpha^t)^* \sigma$ in general, but the deviation is not really bad. The (central) decomposition of KMS states into extremal KMS states gives factor states. If σ is assumed to be a factor state, and (α^t) is asymptotically abelian, one finds that $\lim \frac{1}{T} \int_0^T dt\, (\alpha^t)^* \sigma$ does not depend on the decomposition $L = S + R_1 + R_2 + \ldots$, as the following result indicates.

Proposition.

* The state σ corresponds to the inverse temperature $+1$ rather than the inverse temperature -1 favored in the mathematical literature.

** This section and the following Proposition are in the nature of a technical digression, and may be omitted by the reader essentially interested in the positivity of the entropy production.

Using the above notation, assume that σ is a factor state, and that

$$\lim_{t\to\infty} ||[\alpha^t A, B]|| = 0$$

when $A, B \in \mathcal{A}$. Then, when $T \to \infty$,

$$\lim \frac{1}{T}\int_0^T dt\,(\alpha^t)^*\sigma' = \lim \frac{1}{T}\int_0^T dt\,(\alpha^t)^*\sigma$$

Let us introduce the GNS representation $(\mathcal{H}, \pi, \Omega)$ associated with σ so that if

$$\rho = \lim \frac{1}{T}\int_0^T dt\,(\alpha^t)^*\sigma$$

we have

$$\rho(A) = \lim \frac{1}{T}\int_0^T dt\,(\Omega, \pi(\alpha^t A)\Omega)$$

By restricting T to a subsequence we may assume that in the weak operator topology

$$\lim \frac{1}{T}\int_0^T dt\,\pi(\alpha^t A) = \bar{A} \in \pi(\mathcal{A})''$$

and by assumption we also have $\bar{A} \in \pi(\mathcal{A})'$, hence $\bar{A} \in \pi(\mathcal{A})' \cap \pi(\mathcal{A})'' = \{\lambda \mathbf{1}\}$ since σ is a factor state.

But we may write $\sigma'(\cdot) = (\Omega', \pi(\cdot)\Omega')$: this follows from the perturbation theory of [1] (see proof of Theorem 5.4.4). We have thus

$$\lim \frac{1}{T}\int_0^T dt\,\sigma'(\alpha^t A) = \lim \frac{1}{T}\int_0^T dt\,(\Omega', \pi(\alpha^t A)\Omega')$$

$$= \lim \frac{1}{T}\int_0^T dt\,(\Omega, \pi(\alpha^t A)\Omega) = \lim \frac{1}{T}\int_0^T dt\,\sigma(\alpha^t A)$$

as announced. □

Entropy production.

For finite $\Lambda \subset L$ we have defined

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

but H_L, H_{R_a} do not make sense. We can however define

$$[H_L, H_{R_a}] = \lim_{\Lambda\to L} [H_\Lambda, H_{R_a\cap\Lambda}] = \lim_{\Lambda\to L} [H_\Lambda, H_{a\Lambda}]$$

We have indeed

$$[H_\Lambda, H_{a\Lambda}] = [H_\Lambda - H_{a\Lambda}, H_{a\Lambda}] = [H_\Lambda - \sum_{b>0} H_{b\Lambda}, H_{a\Lambda}]$$

and (A2) gives

$$H_\Lambda - \sum_{b>0} H_{b\Lambda} = \sum_{x\in S} \sum_{X: x\in X\subset\Lambda} \frac{1}{\mathrm{card}(X\cap S)}\Phi(X)$$

[implying the existence of the limit $\lim_{\Lambda\to L}(H_\Lambda - \sum_{b>0} H_{b\Lambda}) = H_L - \sum_{b>0} H_{R_b} \in \mathcal{A}$]. Using (A1) we obtain

$$||[\Phi(X), H_{a\Lambda}]|| \leq 2\lambda^{-1}||\Phi||_\lambda ||\Phi(X)|| e^{\lambda \mathrm{card} X}$$

hence

$$\sum_{X\ni x} ||[\Phi(X), H_{a\Lambda}]|| \leq 2\lambda^{-1}||\Phi||_\lambda e^\lambda ||\Phi||_\lambda$$

and $[H_\Lambda, H_{a\Lambda}]$ has a limit $[H_L, H_{R_a}] \in \mathcal{A}$ when $\Lambda \to L$ with

$$||[H_L, H_{R_a}]|| \leq 2\mathrm{card} S \lambda^{-1} e^\lambda ||\Phi||_\lambda^2$$

The operator

$$i[H_L, H_{R_a}]$$

may be interpreted as the rate of increase of the energy of the reservoir R_a or (since this energy is infinite) rather the rate of transfer of energy to R_a from the rest of the system. According to conventional wisdom we define the rate of entropy production in an $(\alpha^t)^*$-invariant state ρ as

$$e_\rho = \sum_{a>0} \beta_a \rho(i[H_L, H_{R_a}])$$

(this definition does not require that $\rho \in \Sigma$).

Remark.

If we replace S by a finite set $S' \supset S$ and the R_a by the correspondingly smaller sets $R'_a \subset R_a$, we have noted earlier that (A1), (A2), (A3) remain satisfied. As a consequence of (A1) we have

$$i[H_L, H_{R_a} - H_{R'_a}] = \lim_{\Lambda\to L} i[H_\Lambda, H_{a\Lambda} - H'_{a\Lambda}] = \lim_{\Lambda\to L} \delta(H_{a\Lambda} - H'_{a\Lambda})$$

(where the operator δ has been defined just after (A3)), hence

$$\rho(i[H_L, H_{R_a} - H_{R'_a}]) = \lim_{\Lambda\to L} \rho(\delta(H_{a\Lambda} - H'_{a\Lambda})) = 0$$

i.e., the rate of entropy production is unchanged when S and the R_a are replaced by S' and the R'_a. The reason why we do not have $\rho(i[H_L, H_{R_a}]) = 0$ is mathematically because

H_{R_a} is "infinite" ($H_{R_a} \notin \mathcal{A}$), and physically because our definition of $\rho(i[H_L, H_{R_a}])$ takes into account the flux of energy into R_a from S, but not the flux at infinity.

Theorem.

The entropy production in a NESS *is nonnegative, i.e.,* $e_\rho \geq 0$ *if* $\rho \in \Sigma$.

We have seen that

$$[H_L, H_{R_a}] = \lim_{\Lambda \to L} [H_\Lambda, H_{a\Lambda}]$$

$$= \lim_{\Lambda \to L} [H_\Lambda - \sum_{b>0} H_{b\Lambda}, H_{a\Lambda}]$$

Therefore, using (A3) and $[H_{b\Lambda} + B_{b\Lambda}, \sum_{a>0} \beta_a (H_{a\Lambda} + B_{a\Lambda})] = 0$, we find

$$\sum_{a>0} \beta_a [H_L, H_{R_a}] = \lim_{\Lambda \to L} [H_\Lambda - \sum_{b>0} H_{b\Lambda}, \sum_{a>0} \beta_a H_{a\Lambda}]$$

$$= \lim_{\Lambda \to L} [H_\Lambda - \sum_{b>0} H_{b\Lambda}, \sum_{a>0} \beta_a (H_{a\Lambda} + B_{a\Lambda})]$$

$$= \lim_{\Lambda \to L} [H_\Lambda + \sum_{b>0} B_{b\Lambda}, \sum_{a>0} \beta_a (H_{a\Lambda} + B_{a\Lambda})]$$

in the sense of norm convergence.

We also have, for some sequence of values of T tending to infinity and all $A \in \mathcal{A}$,

$$\rho(A) = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt\, \sigma(\alpha^t A) = \lim_{T \to \infty} \lim_{\Lambda \to L} \frac{1}{T} \int_0^T dt\, \sigma(\alpha^t_\Lambda A)$$

where, by (4),

$$\alpha^t_\Lambda A = e^{it(H_\Lambda + \sum_{a>0} B_{a\Lambda})} A e^{-it(H_\Lambda + \sum_{a>0} B_{a\Lambda})} \to \alpha^t A \text{ in norm}$$

when $\Lambda \to L$, uniformly for $t \in [0, T]$.

Write

$$H_{B\Lambda} = H_\Lambda + \sum_{a>0} B_{a\Lambda}$$

$$G_\Lambda = \sum_{a>0} \beta_a (H_{a\Lambda} + B_{a\Lambda}) + \log \mathrm{Tr}_{\mathcal{H}_\Lambda} \exp(-\sum_{a>0} \beta_a (H_{a\Lambda} + B_{a\Lambda}))$$

Then the entropy production is

$$e_\rho = \rho(i \sum_{a>0} \beta_a [H_L, H_{R_a}]) = \lim_{T \to \infty} \lim_{\Lambda \to L} \frac{i}{T} \int_0^T dt\, \sigma(e^{itH_{B\Lambda}} [H_{B\Lambda}, G_\Lambda] e^{-itH_{B\Lambda}})$$

and the convergence when $\Lambda \to L$ of the operator $(e^{itH_{B\Lambda}}[H_{B\Lambda}, G_\Lambda]e^{-itH_{B\Lambda}})$ is uniform for $t \in [0, T]$. According to (A3) we may choose the Λ tending to L such that $\mathrm{Tr}_{\mathcal{H}_\Lambda} e^{-G_\Lambda}(\cdot)$ tends to $\sigma(\cdot)$ in the w^*-topology, hence

$$e_\rho = \lim_{T\to\infty} \lim_{\Lambda\to L} \frac{i}{T} \int_0^T dt\, \mathrm{Tr}_{\mathcal{H}_\Lambda}(e^{-G_\Lambda} e^{itH_{B\Lambda}}[H_{B\Lambda}, G_\Lambda]e^{-itH_{B\Lambda}})$$

$$= \lim_{T\to\infty} \lim_{\Lambda\to L} \frac{1}{T} \int_0^T dt\, \mathrm{Tr}_{\mathcal{H}_\Lambda}(e^{-G_\Lambda} \frac{d}{dt}(e^{itH_{B\Lambda}} G_\Lambda e^{-itH_{B\Lambda}}))$$

$$= \lim_{T\to\infty} \lim_{\Lambda\to L} \frac{1}{T} \big(\mathrm{Tr}_{\mathcal{H}_\Lambda}(e^{-G_\Lambda} e^{iTH_{B\Lambda}} G_\Lambda e^{-iTH_{B\Lambda}}) - \mathrm{Tr}_{\mathcal{H}_\Lambda}(e^{-G_\Lambda} G_\Lambda)\big)$$

and the Theorem follows from the Lemma below, applied with $A = G_\Lambda$, $U = e^{iTH_{B\Lambda}}$ and $\phi(s) = -e^{-s}$.

Lemma.

Let A, U be a hermitean and a unitary $n \times n$ matrix respectively, and $\phi : \mathbf{R} \to \mathbf{R}$ be an increasing function. Then

$$\mathrm{tr}(\phi(A) U A U^{-1}) \le \mathrm{tr}(\phi(A) A)$$

As R. Seiler kindly pointed out to me, this lemma can be obtained readily from O. Klein's inequality

$$\mathrm{tr}(f(B) - f(A) - (B - A) f'(A)) \ge 0$$

where A, B are hermitean and f convex: take $B = UAU^{-1}$ and $\phi = f'$. $\square$

Remark.

We have

$$\sum_{a>0} \rho(i[H_L, H_{R_a}]) = 0$$

because

$$-\sum_{a>0} \rho(i[H_L, H_{R_a}]) = \lim_{\Lambda\to L} \rho(i[H_\Lambda, H_\Lambda - \sum_{a>0} H_{a\Lambda}]) = \frac{d}{dt} \rho(\alpha^t \sum_{X: X\cap S \ne \emptyset} \Phi(X))|_{t=0} = 0$$

where we have used the fact that ρ is $(\alpha^t)^*$-invariant. In particular, in the case of two reservoirs

$$0 \le e_\rho = (\beta_1 - \beta_2) \rho(i[H_L, H_{R_1}])$$

so that if the temperature β_1^{-1} is less than β_2^{-1}, *i.e.*, $\beta_1 - \beta_2 > 0$, the flux of energy into R_1 is ≥ 0: heat flows from the hot reservoir to the cold reservoir.

Proving strict positivity of e_ρ.

It is an obvious challenge to prove that $e_\rho \ne 0$. A natural situation to discuss would correspond to $R_a = \mathbf{Z}^\nu$ and Φ_a translationally invariant. But we need then $\nu \ge 3$ as discussed in [4]. Indeed, for $\nu < 3$ one expects a nonequilibrium steady state to be in fact an equilibrium state at a temperature intermediate between the original temperatures of the reservoirs. Instead of a quantum spin system as described above, a gas of noninteracting fermions would probably be easier to treat first.

Complements and relation with recent work of Jakšić and Pillet.

After this paper was submitted for publication, two interesting contributions were posted to the mp_arc archive: one by Jakšić and Pillet* and one by Maes et al.** In this Section and the next two, I am complying with the editor's request to take into account remarks by the referees, and in particular to discuss the relations of my work with the two references mentionned above.

Note that the definition of entropy production used above is based on the thermodynamic relation $dQ = kT\,dS$ or, in the present case $dS = \sum_a (kT_a)^{-1} dQ_a$. It can be considered a drawback that this definition does not relate directly to a microscopically defined entropy-like quantity, as is done in the papers of Jakšić and Pillet, and Maes et al. We now discuss in detail the approach of Jakšić and Pillet, and its relation with the present paper.***

We are given a C*-algebra $\mathcal{A}$ with identity, an element $V = V^* \in \mathcal{A}$, time evolutions $(\breve{\alpha}^t)$, (α^t) (*i.e.*, strongly continuous one-parameter groups of $*$-automorphisms of $\mathcal{A}$) such that

$$\alpha^t(A) = \breve{\alpha}^t(A) + \sum_{n\geq 1} i^n \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{n-1}} dt_n [\breve{\alpha}^{t_n}(V), [\ldots[\breve{\alpha}^{t_1}(V), A]]]$$

and an $(\breve{\alpha}^t)$-invariant state σ on $\mathcal{A}$. Therefore (α^t) is a local perturbation by V of the "free" evolution given by $(\breve{\alpha}^t)$ and σ is an invariant state for the "free" evolution. We furthermore assume that

(C1) There exists a time evolution (β^t) for which σ is a KMS state at inverse temperature $+1$

(C2) V is in the domain of the infinitesimal generator δ_β of (β^t).

[In fact Jakšić and Pillet assume a temperature -1 in (C1); our choice of temperature $+1$ will bring a change of sign below in the definition of the entropy production. In the situation discussed earlier we have

$$V = \sum_{X \cap S \neq \emptyset} \Phi(X)$$

* V. Jakšić and C.-A. Pillet. "On entropy production in quantum statistical mechanics." mp_arc 00-309.

** Chr. Maes, F. Redig, and M. Verschuere. "Entropy production for interacting particle systems." mp_arc 00-357.

*** We have changed the notation of [2] to align it with the one used above.

hence $||V||_\lambda \leq ||\Phi||_\lambda \mathrm{card} S$, and $V \in \mathcal{A}_\lambda$. Note that $\mathcal{A}_\lambda$ is in the domain of the infinitesimal generator δ_β of (β^t) (see the Appendix), hence (C2) holds. The advantage of the approach of Jakšić and Pillet is that σ can be an arbitrary KMS state: the existence of "boundary terms" $B_{a\Lambda}$ as in (A3) is not required].

In this setup one introduces the observable

$$-\delta_\beta(V)$$

and the *entropy production* in the state ρ is defined as

$$\rho(-\delta_\beta(V))$$

[In our situation we have

$$-\delta_\beta(V) = -\sum_{a>0} \beta_a \sum_{X \subset R_a} \sum_{Y: Y \cap S \neq \emptyset} i[\Phi(X), \Phi(Y)]$$

$$= \sum_{a>0} \beta_a i[H_L, H_{R_a}]$$

so that $\rho(-\delta_\beta(V)) = e_\rho$ is indeed the rate of entropy production in the state ρ].

Finite dimensional digression.

For the purpose of motivation we discuss now the case where $\mathcal{A}$ would be the algebra of $n \times n$ matrices, and consider two states on $\mathcal{A}$ given by density matrices μ, ν. A relative entropy is then defined by

$$\mathrm{Ent}(\mu|\nu) = -\mathrm{tr}(\mu \log \mu - \mu \log \nu) \leq 0$$

If (α^t) is a one parameter group of $*$-automorphisms of $\mathcal{A}$ we have thus

$$\frac{d}{dt}\mathrm{Ent}(\mu \circ \alpha^t|\nu) = \mathrm{tr}(\mu \frac{d}{dt}\alpha^t(\log \nu))$$

Suppose now that ν is preserved by the "free" evolution $(\breve{\alpha}^t)$, and that (α^t) is a perturbation of $(\breve{\alpha}^t)$, so that

$$\alpha^t(A) = e^{i(H+V)t} A e^{-i(H+V)t} \qquad , \qquad \breve{\alpha}^t(A) = e^{iHt} A e^{-iHt}$$

then

$$\frac{d}{dt}\alpha^t(\log \nu) = \alpha^t(i[V, \log \nu])$$

Define now (β^t) by

$$\beta^t(A) = e^{-it \log \nu} A e^{it \log \nu}$$

so that ν is the corresponding KMS state (at inverse temperature $+1$). Then if δ_β is the infinitesimal generator of (β^t) we have

$$i[V, \log \nu] = \delta_\beta(V)$$

hence

$$\frac{d}{dt}\alpha^t(\log \nu) = \alpha^t(\delta_\beta(V))$$

$$\frac{d}{dt}\mathrm{Ent}(\mu \circ \alpha^t|\nu) = \mu(\alpha^t(\delta_\beta(V)))$$

We obtain thus

$$\mathrm{Ent}(\mu \circ \alpha^T|\nu) - \mathrm{Ent}(\mu|\nu) = \int_0^T (\mu \circ \alpha^t)(\delta_\beta(V))\, dt$$

or, taking $\mu = \nu = \sigma$,

$$0 \le -\mathrm{Ent}(\sigma \circ \alpha^T|\sigma) = \int_0^T (\sigma \circ \alpha^t)(-\delta_\beta(V))\, dt$$

The infinite dimensional situation.

If μ, ν are two faithful normal states on a von Neumann algebra $\mathcal{M}$ [in our case $\pi_\sigma(\mathcal{A})''$], Araki has introduced a relative entropy $\mathrm{Ent}(\mu|\nu)$ in terms of a relative modular operator associated with μ, ν. We must refer the reader to [1] Definition 6.2.29 for details. Using this definition, Jakšić and Pillet have worked out an infinite dimensional version of the finite dimensional calculation given above. They are able to prove the formula

$$\int_0^T (\sigma \circ \alpha^t)(-\delta_\beta(V))\, dt = -\mathrm{Ent}(\sigma \circ \alpha^T|\sigma) \ge 0$$

which can be interpreted as an entropy balance, and gives in the limit

$$\rho(-\delta_\beta(V)) \ge 0$$

if ρ is a NESS. The proof is fairly technical.

The approach of Jakšić and Pillet has the interest of great generality. In particular σ can be an arbitrary KMS state. Also, instead of a spin lattice system one can consider fermions on a lattice. For a noninetacting fermion model, Jakšić and Pillet have announced a proof of strict positivity of the entropy production, as had been suggested above.

Appendix: the algebras $\mathcal{A}_\lambda$.

The purpose of this Appendix is to complete the proof of (4) by establishing (10) below. On the way to this result we introduce "partial traces" π_Λ, and algebras $\mathcal{A}_\lambda$ which are of interest in their own right.

For finite $\Lambda \subset L$, a map $\pi_\Lambda : \cup_X \mathcal{A}_X \to \mathcal{A}_\Lambda$ is defined by

$$\pi_\Lambda A = \lim_{Y \to L \setminus \Lambda} \frac{\mathrm{tr}_{\mathcal{H}_Y} A}{\dim \mathcal{H}_Y}$$

If the ϕ_i form an orthonormal basis of $\mathcal{H}_Y$, and $\psi', \psi'' \in \mathcal{H}_\Lambda$ we have

$$(\psi', \frac{\mathrm{tr}_{\mathcal{H}_Y} A}{\dim \mathcal{H}_Y} \psi'') = \frac{1}{\dim \mathcal{H}_Y} \sum_i (\phi_i \otimes \psi', A\phi_i \otimes \psi'')$$

hence $||\pi_\Lambda A|| \le ||A||$. The properties of the following lemma are then readily checked.

Lemma

The map π_Λ extends to a unique linear norm-reducing map $\mathcal{A} \to \mathcal{A}_\Lambda$. Furthermore

$$\pi_\Lambda A = A \qquad \text{if} \qquad A \in \mathcal{A}_\Lambda$$

$$\pi_\Lambda A^* = (\pi_\Lambda A)^*$$

$$\pi_\Lambda \pi_{\Lambda'} = \pi_{\Lambda'} \pi_\Lambda$$

Choose now some $\lambda > 0$. For $A \in \mathcal{A}_\Lambda$, define

$$||A||_\lambda = \inf\{ \sum_{X \subset \Lambda} ||A_X|| e^{\lambda\, \mathrm{card} X} : \sum_X A_X = A\}$$

By compactness we may replace the inf by min. If Λ is replaced by a larger set Λ', and $\sum_Y A_Y = A$ with $Y \subset \Lambda'$, we have

$$\sum_{Y \subset \Lambda'} ||A_Y|| e^{\lambda\, \mathrm{card} Y} \ge \sum_Y ||\pi_\Lambda A_Y|| e^{\lambda\, \mathrm{card}(Y \cap \Lambda)}$$

with $\sum_Y \pi_\Lambda A_Y = \pi_\Lambda A = A$. Therefore $||A||_\lambda$ does not depend on the choice of Λ provided $A \in \mathcal{A}_\Lambda$. We have thus a norm $||.||_\lambda$ on $\cup_X \mathcal{A}_X$, and we may define the Banach space $\mathcal{A}_\lambda$ by completion.

Proposition.

The inclusion map $\cup_X \mathcal{A}_X \to \mathcal{A}$ extends to a norm-reducing map $\omega : \mathcal{A}_\lambda \to \mathcal{A}$ and ω is injective.

ω is norm-reducing because $||A|| \le ||A||_\lambda$ for $A \in \cup_X \mathcal{A}_X$.

Note now that $\pi_\Lambda : \cup_X \mathcal{A}_X \to \mathcal{A}_\Lambda$ reduces the $||.||_\lambda$-norm and extends thus to a linear norm-reducing map $\mathcal{A}_\lambda \to \mathcal{A}_{\Lambda\lambda}$ where $\mathcal{A}_{\Lambda\lambda}$ is $\mathcal{A}_\Lambda$ equipped with the $||.||_\lambda$-norm. Assume that $A \in \mathcal{A}_\lambda$ with $||A||_\lambda = a > 0$. We may choose Λ and $B \in \mathcal{A}_\Lambda$ such that $||A - B||_\lambda < a/3$, hence $||B||_\lambda > 2a/3$. Now $\omega A = 0$ would imply $\pi_\Lambda A = 0$ hence

$$\frac{2a}{3} < ||B||_\lambda = ||\pi_\Lambda(B - A)||_\lambda \le ||A - B||_\lambda < \frac{a}{3}$$

Therefore ω must be injective. $\square$

Corollary.

$\mathcal{A}_\lambda$ is identified by ω to a dense $*$-subalgebra of $\mathcal{A}$; $\mathcal{A}_\lambda$ is then a Banach algebra with respect to the norm $||.||_\lambda$. Taking $\lambda = 0$ we may define $\mathcal{A}_0 = \mathcal{A}$. With this definition, if $\lambda < \mu$ we have $\mathcal{A}_\lambda \supset \mathcal{A}_\mu$, and the map $\mathcal{A}_\mu \to \mathcal{A}_\lambda$ is norm-reducing.

If $A, B \in \mathcal{A}_\Lambda$ we may choose $A_X, B_X \in \mathcal{A}_X$ such that $A = \sum_{X\subset\Lambda} A_X$, $B = \sum_{X\subset\Lambda} B_X$, and

$$||A||_\lambda = \sum_{X\subset\Lambda} ||A_X||e^{\lambda\,\mathrm{card}X} \quad , \quad ||B||_\lambda = \sum_{X\subset\Lambda} ||B_X||e^{\lambda\,\mathrm{card}X}$$

Thus

$$||AB||_\lambda \leq \sum_X \sum_Y ||A_X A_Y||e^{\lambda\,\mathrm{card}(X\cup Y)}$$

$$\leq \sum_X \sum_Y ||A_X||.||A_Y||e^{\lambda(\mathrm{card}X+\mathrm{card}Y)} = ||A||_\lambda ||B||_\lambda$$

Therefore if A, B tend to limits A_∞, B_∞ in $\mathcal{A}_\lambda$, AB tends in $\mathcal{A}_\lambda$ to $A_\infty B_\infty$ and $||A_\infty B_\infty||_\lambda \leq ||A_\infty||_\lambda ||B_\infty||_\lambda$. The rest is clear. $\square$

If $||\Phi||_\lambda < \infty$ and $A_X \in \mathcal{A}_X$ the formula

$$\delta A_X = i \sum_{Y:Y\cap X\neq\emptyset} [\Phi(Y), A_X]$$

defines an element of $\mathcal{A}_\lambda$. If $\lambda > \mu \geq 0$, and $||\Phi||_\lambda < \infty$, one also checks that δ^m defines a map $\mathcal{A}_\lambda \to \mathcal{A}_\mu$ such that

$$||\delta A||_\mu \leq 2(\lambda-\mu)^{-1}||A||_\lambda||\Phi||_\lambda$$

$$||\delta^m A||_\mu \leq ||A||_\lambda m!(2(\lambda-\mu)^{-1}||\Phi||_\lambda)^m \tag{6}$$

[The proof of (6) is basically the same as that of the standard case $\mu = 0$].

We turn now to the proof of (10). We have $\delta_\Lambda = \delta'_\Lambda + \delta''_\Lambda$, where

$$\delta'_\Lambda A = i[H_\Lambda, A] \quad , \quad \delta''_\Lambda A = i[\sum_{a>0} B_{a\Lambda}, A]$$

and (1) and (6) (for $m = 1$) yield

$$||\delta A||_\mu \leq ||A||_\lambda.2(\lambda-\mu)^{-1}||\Phi||_\lambda$$

$$||\delta'_\Lambda A||_\mu \leq ||A||_\lambda.2(\lambda-\mu)^{-1}||\Phi||_\lambda$$

$$||\delta''_\Lambda A||_\mu \leq ||A||_\lambda.2(\lambda-\mu)^{-1}K$$

Given $\epsilon > 0$ and $A \in \mathcal{A}_\lambda$ we can find X such that $A = A_1 + A_2$ with $A_1 \in \mathcal{A}_X$ and $||A_2||_\lambda < \epsilon$. Therefore

$$||(\delta - \delta_\Lambda)A||_\mu \leq ||(\delta - \delta_\Lambda)A_1||_\mu + ||\delta A_2||_\mu + ||\delta'_\Lambda A_2||_\mu + ||\delta''_\Lambda A_2||_\mu$$
$$= ||(\delta - \delta_\Lambda)A_1||_\mu + \epsilon.2(\lambda - \mu)^{-1}(2||\Phi||_\lambda + K) \tag{7}$$

Taking $\Lambda \supset \Lambda_X$ we also have

$$\delta''_\Lambda A_1 = 0$$

by (2), and

$$(\delta - \delta'_\Lambda)A_1 = i \sum_{Y:Y\not\subset\Lambda, Y\cap X\neq\emptyset} [\Phi(Y), A_1]$$

so that

$$||(\delta - \delta'_\Lambda)A_1||_\mu \leq ||A_1||_\lambda.2(\lambda - \mu)^{-1}||\Phi||'_{X\lambda} \tag{8}$$

where $||\Phi||'_{X\lambda} = \sup_{x\in X} \sum_{Y\ni x, Y\not\subset X} e^{(\mathrm{card}Y-1)\lambda}||\Phi(Y)||$. When $\Lambda \to L$ we have $||\Phi||'_{X\lambda} \to 0$ and (7), (8) yield

$$\lim_{\Lambda\to L} ||(\delta - \delta_\Lambda)A||_\mu = 0 \tag{9}$$

We can now prove that, if $||\Phi||_\lambda < \infty$ and $A \in \mathcal{A}_\lambda$,

$$\lim_{\Lambda\to L} ||\delta^m A - \delta^m_\Lambda A|| = 0 \tag{10}$$

We have indeed

$$\delta^m A - \delta^m_\Lambda A = \sum_{k=0}^{m-1} \delta_\Lambda^{m-k-1}(\delta - \delta_\Lambda)\delta^k A$$

and, using (6),

$$||\delta^k A||_{2\lambda/3} \leq ||A||_\lambda.k!(\frac{6}{\lambda}||\Phi||_\lambda)^k$$

hence, by (9),

$$\lim_{\Lambda\to L} ||(\delta - \delta_\Lambda)\delta^k A||_{\lambda/3} = 0$$

so that, using (6),

$$||\delta_\Lambda^{m-k-1}(\delta - \delta_\Lambda)\delta^k A|| \leq ||\delta_\Lambda^{m-k-1}(\delta - \delta_\Lambda)\delta^k A||_0$$
$$\leq ||(\delta - \delta_\Lambda)\delta^k A||_{\lambda/3}(m-k-1)!(\frac{6}{\lambda}||\Phi||_\lambda)^{m-k-1}$$

which tends to zero when $\Lambda \to L$. This concludes the proof of (10).

References.

[1] O. Bratteli and D.W. Robinson. *Operator algebras and quantum statistical mechanics I, II.* Springer, New York, 1979-1981 (2-nd ed. 1987-1997).

[2] R. Haag, N.M. Hugenholtz, and M. Winnink. "On the equilibrium states in quantum statistical mechanics." Commun. Math. Phys. **5**,215-236(1967).

[3] D. Ruelle. *Statistical mechanics. Rigorous results.* Benjamin, New York, 1969.

[4] D. Ruelle. "Natural nonequilibrium states in quantum statistical mechanics." J. Statist. Phys. **98**,57-75(2000).