Journal of Statistical Physics, Vol. 111, Nos. 5/6, June 2003 (© 2003)

Heat Conduction Networks

Christian Maes,[1] Karel Netočný,[2] and Michel Verschuere[3]

Received April 16, 2002; accepted November 21, 2002

We study networks of interacting oscillators, driven at the boundary by heat baths at possibly different temperatures. A set of first elementary questions are discussed concerning the uniqueness of a stationary possibly Gibbsian density and the nature of the entropy production and the local heat currents. We also derive a Carnot efficiency relation for the work that can be extracted from the heat engine.

KEY WORDS: Heat current; entropy production; nonequilibrium state.

1. THE MODEL AND RESULTS

Consider a finite connected graph $G=(V,\sim)$ with vertex set V. Two vertices (= sites) $i \neq j \in V$ are called nearest neighbors if there is an edge between them: $i \sim j$. Every site $i \in V$ carries a momentum and position coordinate $(p_i, q_i) \in \mathbb{R}^2$. Generalizations to higher dimensional coordinate vectors are straightforward. We select a non-empty subset $\partial V \subset V$, called boundary sites, that, below, will be imagined connected to thermal baths at possibly different temperatures. States (p, q) are elements $((p_i, q_i), i \in V) \in \mathbb{R}^{2|V|}$ and ρ will denote a probability density (with respect to $dp\, dq = \prod_i dp_i\, dq_i$) on it.

The coupling between the degrees of freedom is modeled by the Hamiltonian

$$H(p,q)=\sum_{i\in V}\frac{p_i^2}{2}+U(q) \tag{1.1}$$

[1] Instituut voor Theoretische Fysica K.U. Leuven, B-3001 Leuven, Belgium; e-mail: christian.maes@fys.kuleuven.ac.be

[2] E-mail: karel.netocny@fys.kuleuven.ac.be

[3] Aspirant FWO Vlaanderen; e-mail: michel.verschuere@fys.kuleuven.ac.be

with a symmetric nearest neighbor potential

$$U(q) = \sum_i U_i(q_i) + \sum_{i \sim j} \lambda_{ij} \Phi(q_i - q_j) \tag{1.2}$$

where $\lambda_{ij} = \lambda_{ji} \neq 0$ whenever $i \sim j$ and Φ is even.

1.1. Hamiltonian Dynamics

The above mechanical system of coupled oscillators has a time-evolution given by Newton's equations of motion,

$$\begin{aligned} dq_i &= p_i \, dt, \\ dp_i &= -\frac{\partial U}{\partial q_i}(q) \, dt, \qquad i \in V \end{aligned} \tag{1.3}$$

which is well-defined under standard conditions on the U_i and Φ. We assume that these are infinitely differentiable and such that (1.3) yields existence of uniquely defined global solutions with Hamiltonian flow generated by the Liouville operator

$$\mathscr{L}_H \equiv p \cdot \nabla_q - \nabla_q U \cdot \nabla_p \tag{1.4}$$

where the dot product is a sum over $i \in V$.

The Gibbs measures with densities

$$\rho^\beta(p, q) \equiv \frac{1}{Z} e^{-\beta H(p, q)} \tag{1.5}$$

at inverse temperatures $\beta > 0$ are stationary under (1.3) assuming normalizability. They are however much more than that: they satisfy the Kubo–Martin–Schwinger (KMS) conditions

$$\begin{aligned} \int \frac{\partial f}{\partial q_i} \rho(p, q) \, dp \, dq &= \beta \int f \frac{\partial H}{\partial q_i} \rho(p, q) \, dp \, dq, \\ \int \frac{\partial f}{\partial p_i} \rho(p, q) \, dp \, dq &= \beta \int f \frac{\partial H}{\partial p_i} \rho(p, q) \, dp \, dq \end{aligned} \tag{1.6}$$

for smooth functions f, for all $i \in V$. Obviously, stationarity of ρ does not imply (1.6) and it has been a widely discussed question what kind of further regularity conditions, beyond stationarity under (1.3), must be imposed on densities ρ so that $\rho = \rho^\beta$ is a Gibbs measure (1.5). Often, and this is sometimes more efficient for infinite Hamiltonian systems, a noise or stochastic dynamics is added to (1.3) so that the randomized evolution makes all sufficiently regular stationary measures Gibbsian, see e.g., ref. 4 for an introduction. This randomization is usually of a bulk nature in the sense that it affects the whole volume of the system. Our first question takes a somewhat different approach. We ask whether stationarity together with the KMS-conditions (1.6) but only for $i \in \partial V$, suffice to make $\rho = \rho^\beta$. More generally, we impose

$$\int \frac{\partial f}{\partial p_i}\, \rho(p, q)\, dp\, dq = \beta_i \int f\, p_i\, \rho(p, q)\, dp\, dq, \qquad i \in \partial V \tag{1.7}$$

for all smooth f, where the $\beta_i > 0$ are thought of as the inverse temperatures of reservoirs connected to the $i \in \partial V$. A further motivation for this question comes from the heat bath dynamics that will be introduced below; there, the equations (1.7) at the boundary sites are equivalent with having zero entropy production (as will be explained in Section 3.3). Related to this, we observe that (1.7) implies

$$\int p_i^2\, \rho(p, q)\, dp\, dq = \frac{1}{\beta_i}, \qquad i \in \partial V$$

which says that the kinetic temperature at the boundary sites must be equal to the imagined corresponding reservoir temperatures. No matter how at the boundary the coupling with the reservoirs is actually performed, (1.7) thus implies that there is no net energy current into the reservoirs, hence a vanishing mean entropy production.

Question 1. Is it possible to find a density ρ that is stationary under (1.3) and that satisfies (1.7)?

Answer 1. In general: no, when $\beta_i \neq \beta_j$ for some $i, j \in \partial V$. If all $\beta_i = \beta$, further graph-dependent conditions on the potentials U_i and Φ can ensure that there is such a unique $\rho = \rho^\beta$ but it is very well possible to have still another solution ρ.

This question and answer will be put more precisely in Section 2.

1.2. Heat Bath Dynamics

In this case the dynamics is Hamiltonian except at the boundary ∂V where the interaction with the reservoirs has the form of Langevin forces as expressed by the Itô stochastic differential equations

$$\begin{aligned} dq_i &= p_i\, dt, \qquad i \in V \\ dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt, \qquad i \in V \setminus \partial V \\ dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt - \gamma p_i + \sqrt{\frac{2\gamma}{\beta_i}}\, dW_i(t), \qquad i \in \partial V \end{aligned} \tag{1.8}$$

The β_i are the inverse temperatures of the heat baths coupled to the boundary sites $i \in \partial V$; $W_i(t)$ are mutually independent, one-dimensional Wiener processes.

Again, standard conditions on the potentials allow the existence of the corresponding Markov diffusion process with a strongly continuous semigroup generated by

$$\mathscr{L} \equiv \mathscr{L}_H - \gamma \sum_{i \in \partial V} p_i \frac{\partial}{\partial p_i} + \sum_{i \in \partial V} \frac{\gamma}{\beta_i} \frac{\partial^2}{\partial p_i^2} \tag{1.9}$$

The Gibbs measure (1.5) is (generalized) reversible for the process whenever $\beta_i = \beta$, $i \in \partial V$ (also called, reversible in the Yaglom sense), in which case $\mathscr{L}^* = \pi \mathscr{L} \pi$ on $L^2(\rho^\beta)$. Here, the kinematical time-reversal π is given by the involution defined as: $\pi f(p, q) = f(-p, q)$. The question of uniqueness and mixing properties of a stationary density ρ has been discussed in refs. 1, 2, and 12 at least for G a one-dimensional lattice interval. Here we ask

Question 2. Is it possible that more than one smooth stationary density ρ exists?

Answer 2. Yes, depending on the graph and even for all $\beta_i = \beta$ and for nicely behaving interaction potentials, more than one such ρ may exist.

This will be taken up at the end of Section 3.

The rest of the questions have to do with the nonequilibrium problem where the temperatures of the reservoirs are different. Then, we expect to have heat currents in the system and positive entropy production. The first elementary question is therefore

Question 3. Suppose that ρ is a smooth and stationary density with mean entropy production rate $\dot{S}(\rho)$. Is $\dot{S}(\rho) > 0$ when some $\beta_i \neq \beta_j$?

Answer 3. Yes, quite generally. Nevertheless, depending on the complexity of the graph, we will need further conditions on the potential to prove the strict positivity.

We will start the discussion around entropy production in Section 3. Related to that is the following

Question 4. Does $\dot{S}(\rho) = 0$ imply that all local heat currents in the system are zero, and if some $\beta_i \neq \beta_j$, what is then the direction of the heat currents?

Answer 4. We prove that if there is a unique smooth stationary density, then vanishing entropy production implies that all local heat currents are zero. Only for the simplest graphs, we can prove that the direction of the heat currents is as expected from thermodynamics.

Question 4 becomes easier in the linear regime (where the differences between the β_i are assumed to be very small). This raises the question in what sense the minimum entropy production principle can characterize the stationary density up to linear order. We will see, also in Section 3, that this principle cannot be applied here.

Finally, in Section 4, we compute the maximal efficiency of our system. We couple the system to an external world on which it can exert forces. The efficiency is the ratio of the net work output to the energy input and we give an upper bound in terms of temperature differences as appears in the treatment of the Carnot engine.

1.3. Background

The above heat bath dynamics is not a microscopic dynamics but it results from a Markov approximation for the evolution of a small subsystem of oscillators coupled to large reservoirs. Derivations of such dynamics from first principles have been discussed in refs. 3 and 6 and more recently in refs. 1 and 2. We would like to draw the attention to the even more recent ref. 9 where the general philosophy is discussed on a mathematically elementary level. The nonequilibrium harmonic crystal was first solved in refs. 10 and 13.

2. KMS-CONDITIONS AT THE BOUNDARY

Here we study Question 1 and we are in the framework of Hamiltonian dynamics.

2.1. Propagation of KMS Conditions

We first introduce the notion of *non-degeneracy* of real functions. It depends on an integer n and later on it will be applied to the second derivative of the pair potential Φ.

Definition 2.1. We say that a function $f: \mathbb{R} \mapsto \mathbb{R}$ is n-non-degenerate whenever the set

$$\mathcal{U}^n = \{(q_1,..., q_n) \in \mathbb{R}^n : \exists (q'_1,..., q'_n) \in \mathbb{R}^n : \det f(q'_i - q_j) \neq 0\} \tag{2.1}$$

is dense in $\mathbb{R}^n$. Here, $\det f(q'_i - q_j)$ is the determinant of the $n \times n$ matrix with elements $f(q'_i - q_j)$.

Every real function $f \not\equiv 0$ is obviously 1-non-degenerate and all even real continuous functions are 2-non-degenerate. One can also verify that all polynomials of degree r are n-non-degenerate if and only if $r \geqslant n-1$. Here is a further characterization:

Proposition 2.2. Suppose we can write

$$f(q' - q) = \sum_{m=0}^{\infty} b_m(q)\, \phi_m(q')$$

absolutely convergent, where the $\{\phi_m\}$ are such that there is a non-empty set $B \subset \mathbb{R}$ for which,

$$\text{if } \sum_m \lambda_m \phi_m(q') = 0 \quad \text{for all } q' \in B, \quad \text{then } \lambda_m = 0 \quad \text{for all } m$$

If there are n different functions $b_{m_1},..., b_{m_n}$ for which the matrix $(b_{m_i}(q_j))$ has a non-vanishing determinant, $\det b_{m_i}(q_j) \neq 0$, for all $(q_1,..., q_n)$ in a dense set of $\mathbb{R}^n$, then f is n-non-degenerate.

Proof. By writing out the determinant as a sum over permutations, one can easily see that

$$\det f(q'_i - q_j) = \sum_{m_1,..., m_n} \det b_{m_i}(q_j) \prod_{k=1}^{n} \phi_{m_k}(q'_k)$$

Hence, from the hypothesis on the functions ϕ_m, if for all $(q'_1,..., q'_n) \in B^n$, $\det f(q'_i - q_j) = 0$, then the $(q_1,..., q_n)$ are such that $\det b_{m_i}(q_j) = 0$ for all possible choices $m_1,..., m_n$. Alternatively, if there exists a choice $b_{m_1},..., b_{m_n}$ so that $\det b_{m_i}(q_j) \neq 0$, then there must be $(q'_1,..., q'_n) \in B^n$ with $\det f(q'_i - q_j) \neq 0$. ▮

The following example helps to make the above proposition more explicit: for $f(q) = aq^2 + b$, we have $f(q' - q) = b_0(q) + b_1(q)\, q' + b_2(q)\, q'^2$ for $b_0(q) \equiv aq^2 + b$, $b_1(q) \equiv -2aq$, $b_2(q) \equiv a$. To check the $n = 2$-non-degeneracy, it suffices to observe that for the choice $m_1 = 1, m_2 = 2$ the determinant $b_2(q_2)\, b_1(q_1) - b_2(q_1)\, b_1(q_2) = 2a^2(q_1 - q_2)$ is non-zero (when $a \neq 0$) on the dense set $q_1 \neq q_2$.

Our first result of this section demonstrates how the boundary KMS conditions propagate to other sites provided that the pair potential is n-non-degenerate for n high enough. We start with some notation. The number of sites in a subset $A \subset V$ is denoted by $|A|$. Given $j \in A$, we introduce the set

$$N(j \mid A) = \{v \in V \setminus A;\, v \sim j\}\,. \tag{2.2}$$

Given a set of sites A and a site $i \in V \setminus A$, we use the notation $i \overset{n}{\sim} A$ whenever there is a $j \in A$ such that $i \sim j$ and $|N(j \mid A)| = n$.

Proposition 2.3. Fix an integer n and assume that Φ'' is n-non-degenerate. Let a set $A \subset V$ be given. If there is a smooth density $\rho > 0$ which is invariant under the Hamiltonian flow and satisfies the KMS conditions

$$\frac{\partial \ln \rho}{\partial p_j} = -\beta p_j \tag{2.3}$$

for all sites $j \in A$, then also

$$\frac{\partial \ln \rho}{\partial p_i} = -\beta p_i \tag{2.4}$$

for every site $i \overset{n}{\sim} A$.

Proof. Writing the density ρ in the form $\rho = \exp(-\beta H - W)$, the invariance under the Hamiltonian flow, $\{H, \rho\} = 0$, is equivalent to

$$p \cdot \nabla_q W - \nabla_q U \cdot \nabla_p W = 0\,. \tag{2.5}$$

Similarly, the KMS conditions (2.3) are

$$\frac{\partial W}{\partial p_k} = 0, \qquad k \in A \tag{2.6}$$

and, differentiating (2.5) with respect to p_k, one also gets

$$\frac{\partial W}{\partial q_k} = 0, \qquad k \in A \tag{2.7}$$

Since $i \overset{n}{\sim} A$, there is a site $j \in A$ such that $j \sim i$ and $|N(j \mid A)| = n$. Fixing such a j and taking the derivative of (2.5) with respect to q_j, we get, for all (p, q),

$$\sum_{v \in N(j \mid A)} \lambda_{vj} \Phi''(q_v - q_j) \frac{\partial W}{\partial p_v} = 0 . \tag{2.8}$$

To solve this equation we use that there are exactly n terms in the above sum and that $\partial W / \partial p_v$ does not depend on q_j. If $\{q_v\}_{v \in N(j \mid A)} \in \mathscr{U}^n$ with $\mathscr{U}^n$ being the same as in Definition 2.1, then there exist n values $q_j^{(v)}$, $v \in N(j \mid A)$ for q_j which, by substituting them in (2.8), give n linearly independent homogenous equations. Hence, as $i \in N(j \mid A)$ and W is a smooth function,

$$\frac{\partial W}{\partial p_i} = 0 \tag{2.9}$$

whenever $\{q_v\}_{v \in N(j \mid A)} \in \mathscr{U}^n$. Since $\mathscr{U}^n$ is dense in $\mathbb{R}^n$, the above equality is actually true for all $q_v \in \mathbb{R}$, $v \in N(j \mid A)$. ∎

2.2. Non-Existence of Stationary Density for Unequal Boundary Temperatures

To state the main result of this section, we first introduce the notion of the degree of a graph with boundary. It characterizes the complexity of the graph that will appear in a condition on the pair potential, see Theorem 2.4.

Let a graph $G = (V, \sim)$ with a boundary $\partial V \subset V$ be fixed and take a sequence $v_1,..., v_r$ of sites from $V \setminus \partial V$, $r \leqslant |V \setminus \partial V|$. Given an integer n, we say that this sequence is n-admissible if the following hold true:

(1) $v_1 \overset{m_1}{\sim} \partial V = \partial V_0$, $m_1 \leqslant n$,

(2) $v_2 \overset{m_2}{\sim} \partial V \cup \{v_1\} = \partial V_1$, $m_2 \leqslant n$,

⋮

(r) $v_r \overset{m_r}{\sim} \partial V \cup \{v_1,..., v_{r-1}\} = \partial V_{r-1}$, $m_r \leqslant n$, and $\partial V_r = \{v_r\} \cup \partial V_{r-1}$ is a connected set.

Moreover, if $\partial V_r = V$, we call this sequence complete. The smallest n such that there exists a n-admissible sequence of the non-boundary sites is called the degree of the graph G with boundary ∂V and we use the notation $n(G, \partial V)$ for it. Similarly, the smallest n such that there is a complete n-admissible sequence is the complete degree of the graph G with boundary

∂V and we use the notation $\bar{n}(G, \partial V)$. In case $\partial V = V$, we define $n(G, \partial V) = \bar{n}(G, \partial V) = 1$. In general, an easy upper bound on the (complete) degree of any graph with boundary is the maximal multiplicity of vertices.

We are now ready to formulate our statement on the existence and uniqueness of a stationary measure under boundary KMS conditions.

Theorem 2.4. Let the pair potential Φ be such that Φ'' is m-non-degenerate for all $m \leqslant n(G, \partial V)$. Then a smooth density $\rho > 0$ invariant under the Hamiltonian flow and satisfying the boundary KMS conditions

$$\frac{\partial \ln \rho}{\partial p_i} = -\beta_i p_i, \qquad \text{for all} \quad i \in \partial V \tag{2.10}$$

exists if and only if all the temperatures are equal: $\beta_i = \beta$, $i \in \partial V$. In this case and on the extra condition that Φ'' is m-non-degenerate for all $m \leqslant \bar{n}(G, \partial V)$, the density is unique and given by $\rho = \rho^\beta$.

The proof relies on the following simple lemma:

Lemma 2.5. Assume $\Phi'' \not\equiv 0$ and let $\rho > 0$ be a smooth density invariant under the Hamiltonian flow and satisfying (2.10) at two sites $i \sim j$ with inverse temperatures β_i and β_j. Then $\beta_i = \beta_j$.

Proof. Let $i \sim j$. By the same argument as in the proof of Proposition 2.2, the invariance under the Hamiltonian flow together with conditions (2.10) at $i \sim j$ imply

$$\frac{\partial \ln \rho}{\partial q_i} = -\beta_i \frac{\partial U}{\partial q_i}, \qquad \frac{\partial \ln \rho}{\partial q_j} = -\beta_j \frac{\partial U}{\partial q_j} \tag{2.11}$$

Differentiating further these equations with respect to q_j and q_i, respectively, we get

$$(\beta_i - \beta_j)\, \Phi''(q_i - q_j) = 0 \tag{2.12}$$

for all q_i, q_j, which proves the statement. ∎

Proof of Theorem 2.4. According to the assumptions, there is a sequence $v_1,..., v_r$ of sites from $V \setminus \partial V$ such that $v_k \overset{m_k}{\sim} \partial V_{k-1}$, $m_k \leqslant n(G, \partial V)$ for all $1 \leqslant k \leqslant r$ and ∂V_r is a connected set; recall the notation $\partial V_k = \partial V \cup \{v_1,..., v_k\}$ for $k = 0,..., r$. We will prove by induction that for any $0 \leqslant k \leqslant r$ the following is true: If A is any maximal connected component of ∂V_k, then all temperatures β_i, $i \in \partial V \cap A$, are equal, $\beta_i = \beta$, and the KMS condition

$$\frac{\partial \ln \rho}{\partial p_i} = -\beta_i p_i \tag{2.13}$$

is satisfied for all $i \in A$.

Let $k = 0$, first. As $\partial V_0 = \partial V$, one has for any $i, j \in \partial V$, $i \sim j$ that $\beta_i = \beta_j$ due to Lemma 2.5.

Fix an integer K, $1 \leqslant K \leqslant r-1$. Let the hypothesis be true for all $k < K$ and take $k = K$, now. Since $v_K \overset{m_K}{\sim} \partial V_{k-1}$ with an $m_K \leqslant n(G, \partial V)$, there is a site $v \in \partial V_{K-1}$ such that $|N(v \mid \partial V_{K-1})| = m_k$. Denote by A_v the connected component of ∂V_{K-1} for which $v \in A_v$. According to the induction hypothesis, the KMS conditions with an equal β are satisfied for all $i \in A$. As Φ'' is non-degenerate of order m_k, Proposition 2.3 implies the KMS condition with β to be also true at v_K. The proof of the hypothesis is finished by applying Lemma 2.5. Since ∂V_r is connected, the first part of the theorem is proven.

To prove the second part it suffices to realize that the sequence $v_1, \ldots, v_r$ may be chosen in such a way that $\partial V_r = V$ and to use that $\rho = \rho^\beta$ is the only density satisfying the KMS conditions at all sites with the same temperature β. ∎

2.3. Examples

To illustrate how Theorem 2.4 can be applied in case of various graphs, we discuss some examples.

Let G be a linear chain, i.e., $V = (v_1, \ldots, v_k)$ and $v_1 \sim v_2 \sim \cdots \sim v_k$. If $v_1 \in \partial V$ or $v_k \in \partial V$, then $n(G, \partial V) = \bar{n}(G, \partial V) = 1$. In this case, the condition of non-degeneracy in Theorem 2.4 only requires that $\Phi \not\equiv \text{const}$. On the other hand, if, for instance, $\partial V = \{v_j, v_{j'}\}$, $1 < j < j' < k$, then $n(G, \partial V) = \bar{n}(G, \partial V) = 2$ and the pair potential Φ is required to be non-quadratic: $\Phi'' \not\equiv \text{const}$. Another instance where both the degree and the complete degree are equal to 2 is discussed in the next section. We show that the uniqueness statement fails in this case. Another example with only the trivial condition on the pair potential $\Phi \not\equiv \text{const}$ is provided by the class of tree graphs, i.e., the graphs where any two vertices are connected via exactly one path. Assume G to be such a graph and denote by $V_1 \subset V$ the set of vertices which have exactly one neighbor (= vertices with multiplicity 1). If $V_1 \subset \partial V$, then indeed $n(G, \partial V) = \bar{n}(G, \partial V) = 1$. An example of such a graph is given in Fig. 1, with the boundary defined by $\partial V = \{1, 2, 5, 6\}$.

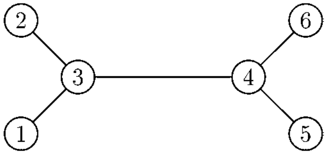

Fig. 1. Example of a tree graph if the boundary is formed by the sites 1, 2, 5, and 6.

An interesting class of graphs is obtained by taking subgraphs of the regular latices $\mathbb{Z}^d$. An obvious upper bound on both degrees is then $n(G, \partial V)$, $\bar{n}(G, \partial V) \leqslant 2d$. Depending on the boundary ∂V, these degrees can get smaller. Recall that when ∂V is a connected set, then $n(G, \partial V) = 1$. For the two-dimensional sublattice of Fig. 2 with the boundary $\partial V = \{1, 2,..., 8\}$ one can easily check that $n(G, \partial V) = \bar{n}(G, \partial V) = 1$, while in the case $\partial V = \{2, 7\}$, for instance, one finds $n(G, \partial V) = \bar{n}(G, \partial V) = 3$. We remark that a more physical model on a d-dimensional sublatice is obtained by replacing the scalar position-momentum coordinates $(p_i, q_i) \in \mathbb{R}^2$ assigned to each vertex $i \in V$ with the vector coordinates $(p_i, q_i) \in \mathbb{R}^{2d}$. Such a generalization is quite straightforward and we omit the details.

2.4. Example with Non-Unique Stationary Measure

We continue with Question 1 and look for non-uniqueness of stationary densities under the Hamiltonian flow when all temperatures at the boundary sites in (1.7) are equal. We give one specific counterexample to show that the condition of non-degeneracy in Theorem 2.4 plays an important role.

Consider the graph $G = (V, \sim)$ defined as follows: $V = \{1, 2, 3, 4\}$, $\partial V = \{1, 2\}$, and $1 \sim 3$, $3 \sim 2$, $2 \sim 4$, $4 \sim 1$, see Fig. 3.

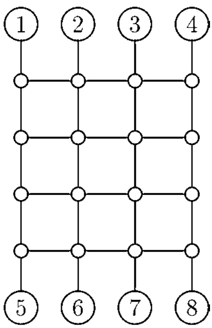

Fig. 2. Portion of square lattice; the non-degeneracy condition is trivial if the boundary is made by the sites 1,..., 8.

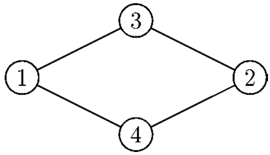

Fig. 3. Example with a non-unique stationary measure.

The self-potential and the pair interaction be purely quadratic: $U_i(q_i) = \alpha^2 q_i^2/2$ and $\Phi(q) = q^2/2$; take $\lambda_{ij} \equiv \lambda > 0$. Let the temperatures be equal, $\beta_1 = \beta_2 = \beta$, and consider the family of probability densities $\rho_c = \exp(-W_c)/Z_c$ parameterized by $c > -1/2$ (to allow normalizability) such that

$$W_c(p, q) \equiv \beta H(p, q) + \frac{c\beta}{2}\,[(p_3 - p_4)^2 + (\alpha^2 + 2\lambda)(q_3 - q_4)^2] \tag{2.14}$$

The remarkable thing is that $W_c - \beta H$ is a conserved quantity of the Hamiltonian dynamics even though it does not depend on all the coordinates of the system, $\{W_c, H\} = 0$. Since, trivially, the boundary KMS conditions (1.7) are verified, we thus have here a family of densities all different from the Gibbsian ρ^β, which are also invariant under the Hamiltonian flow and that satisfy the KMS conditions at the boundaries of the graph. This non-uniqueness of a stationary measure is not surprising, however, as one checks by a suitable canonical transformation that the Hamiltonian of the model decouples into the sum of two independent parts. Notice also that these ρ_c are not invariant under exchanging the momenta but they are invariant under a global sign reversal of the momenta, $\rho_c \pi(p, q) = \rho_c(p, q)$. This is not possible in infinite volume, see refs. 4 and 5. It is known that harmonic oscillators can have besides Gibbsian, additional regular stationary measures, see e.g., ref. 14. On the other hand, we do not believe that the quadratic nature of the potential is crucial; for other graphs, examples of non-uniqueness can probably also be obtained for coupled anharmonic oscillators.

3. ENTROPY PRODUCTION AND HEAT CURRENTS

We now move to the case of heat bath dynamics and we must first introduce what is needed for Questions 3 and 4.

At each boundary site $i \in \partial V$, we introduce the heat current into the corresponding heat bath as minus the work performed on the system by Langevin forces acting at the site. If $\omega = ((q(t), p(t)), t \in [-\tau, \tau])$ denotes the evolution of the system in the lapse of time $[-\tau, \tau]$, then the time-integrated boundary current at i is defined as

$$J_i^\tau(\omega) \equiv \int_{-\tau}^{\tau} \left[\gamma p_i^2(t)\, dt - \sqrt{\frac{2\gamma}{\beta_i}}\, p_i(t) \circ dW_i(t) \right] \tag{3.1}$$

where the "$\circ$" stresses the Stratonovich integration. The above definition relates the boundary currents J_i^τ, $i \in \partial V$ to the random trajectory ω. The realization $W_i(t)$ of the Wiener process appearing in (3.1) is determined uniquely by the equations of motion (1.8) (up to a redundant constant), once $\omega = ((q(t), p(t)), t \in [-\tau, \tau])$ is given. A global energy conservation statement follows from (1.8):

$$H(\omega_\tau) - H(\omega_{-\tau}) = -\sum_{i \in \partial V} J_i^\tau(\omega) \tag{3.2}$$

Consider the decomposition $H = \sum_i H_i$ of the Hamiltonian (1.1) into local versions

$$H_i(q, p) \equiv \tfrac{1}{2} p_i^2 + U_i(q_i) + \tfrac{1}{2} \sum_{j:\, j \sim i} \lambda_{ij} \Phi(q_i - q_j) \tag{3.3}$$

which should be interpreted as the energy stored at site $i \in V$. After defining local time-integrated currents $J_{ij}^\tau = -J_{ji}^\tau$ by $J_{ij}^\tau(\omega) \equiv \int_{-\tau}^{\tau} J_{ij}(q(t), p(t))\, dt$ where

$$J_{ij}(q, p) \equiv \tfrac{1}{2} \lambda_{ij} (p_i + p_j)\, \Phi'(q_i - q_j) \tag{3.4}$$

we obtain the next local version of the energy conservation:

$$H_i(\omega_\tau) - H_i(\omega_{-\tau}) = -J_i^\tau(\omega) - \sum_{j:\, j \sim i} J_{ij}^\tau(\omega) \tag{3.5}$$

where the first term on the right-hand side is to be omitted whenever $i \notin \partial V$. The conservation laws (3.2) and (3.5) have a number of implications for the steady state. Steady state expectations will be denoted by brackets and we will take for granted that it has a smooth stationary probability density $\rho = \exp[-W]$: if f depends on the state (p_t, q_t) at a single time, then

$$\langle f \rangle = \int dp\, dq\, f(p, q)\, \rho(p, q)$$

This is a non-trivial mathematical assumption but it is supported by the analysis of refs. 1, 2, 11, and 12 to which we refer for details. In the steady state regime, one has $\langle J_i^\tau \rangle = \sum_{j:\, j \sim i} \langle J_{ji}^\tau \rangle$ and $\sum_i \langle J_i^\tau \rangle = 0$.

Since $\langle \mathscr{L} H_i \rangle = 0$, we have

$$\sum_{j:\, j \sim i} \langle J_{ji} \rangle = \gamma \left(\langle p_i^2 \rangle - \frac{1}{\beta_i} \right), \qquad i \in \partial V \tag{3.6}$$

or

$$\frac{1}{2\tau} \langle J_i^\tau \rangle = \gamma \left(\langle p_i^2 \rangle - \frac{1}{\beta_i} \right), \qquad i \in \partial V \tag{3.7}$$

The total entropy production of the system plus reservoirs is the sum of the change of entropy in the system and the change of entropy in the reservoirs. From the point of view of the system, the change of entropy in the reservoirs corresponds to the entropy current J_S, that is the energy dissipated in the environment per unit time divided by the temperature of the reservoirs: in the steady state

$$J_S = \frac{1}{2\tau} \sum_{i \in \partial V} \beta_i \langle J_i^\tau \rangle \tag{3.8}$$

and the entropy of the system does not change, Hence, the total steady state entropy production equals (3.8). We can see from (3.7) that it is zero when the kinetic temperatures $\langle p_i^2 \rangle = 1/\beta_i$, $i \in \partial V$ as announced following (1.7).

3.1. Statistical Mechanical Entropy Production

It is important to go beyond averages to study further properties of the entropy production. We derive the formula for the variable entropy production in the line of thoughts of refs. 8 and 9. There, the variable entropy production is identified with that part in the action functional for the space-time distribution of the trajectories that breaks the time-reversal invariance. In ref. 11 a fluctuation symmetry is derived for this entropy production.

We first consider a reversible reference process $P_{\rho\beta}^{\kappa,\tau}$ corresponding to the dynamics

$$\begin{aligned} dq_i &= p_i \, dt, \qquad i \in V \\ dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt, \qquad i \in V \setminus \partial V \\ dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt - \gamma \kappa_i p_i \, dt + \sqrt{\frac{2\gamma}{\beta_i}}\, dW_i(t), \qquad i \in \partial V \end{aligned} \tag{3.9}$$

We take $\kappa_i \beta_i = \beta$, $\forall i \in \partial V$, so that the process (3.9) is reversible, as may be easily checked. As a consequence, we have $P^{\kappa,\tau}_{\rho^\beta} = P^{\kappa,\tau}_{\rho^\beta}\Theta$ where $(\Theta\omega)_t = \pi\omega_{-t}$ with $\omega = ((p_t, q_t), t \in [-\tau, \tau])$ a trajectory and π reverses the sign of the momenta. The stationary density is ρ^β just like in (1.5).

Let P^τ_ρ denote the pathspace measure obtained from the dynamics (1.8), started from initial states $\omega_{-\tau} = (q(-\tau), p(-\tau))$ that are sampled from the stationary probability density ρ. We compute the density of the process P^τ_ρ with respect to $P^{\kappa,\tau}_{\rho^\beta}$. Writing the Radon–Nikodym derivative in the form

$$dP^\tau_\rho(\omega) = e^{-A_\rho(\omega)}\, dP^{\kappa,\tau}_{\rho^\beta}(\omega)$$

the action functional A_ρ is simply found by application of a Girsanov formula, see ref. 7:

$$\begin{aligned}-A_\rho(\omega) = \sum_{i\in\partial V} \frac{1}{2}\bigg[\int_{-\tau}^{\tau} (\beta-\beta_i)\, p_i(t)\, dp_i(t) \\ + \int_{-\tau}^{\tau} (\beta-\beta_i)\frac{\partial U}{\partial q_i}(q(t))\, p_i(t)\, dt + \int_{-\tau}^{\tau} \gamma(\beta\kappa_i - \beta_i)\, p_i^2(t)\, dt\bigg] \\ + \ln \rho(\omega_{-\tau}) - \ln \rho^\beta(\omega_{-\tau})\end{aligned}$$

The first integral appearing in $A_\rho(\omega)$ is a stochastic Itô integral. The definition of this integral is itself not time-reversal invariant and the second integral is clearly antisymmetric under time-reversal (because the momenta change sign). The source of time-reversal breaking is $R^\tau_\rho(\omega) = A_{\rho\pi}(\Theta\omega) - A_\rho(\omega)$ and thus equals

$$\begin{aligned}R^\tau_\rho(\omega) = \sum_{i\in\partial V} (\beta-\beta_i)\int_{-\tau}^{\tau}\bigg[p_i(t)\circ dp_i(t) + \frac{\partial U}{\partial q_i}(q(t))\, p_i(t)\, dt\bigg] \\ + \ln \rho(\omega_{-\tau}) - \ln \rho(\omega_\tau) - \ln \rho^\beta(\omega_{-\tau}) + \ln \rho^\beta(\omega_\tau) \qquad (3.10)\end{aligned}$$

where the first stochastic integral is meant in Stratonovich sense; it gives $p_i^2(\tau)/2 - p_i^2(-\tau)/2$. After a number of straightforward manipulations (3.10) becomes

$$\begin{aligned}R^\tau_\rho(\omega) = \sum_{i\in\partial V} (\beta-\beta_i)\bigg[\left(\frac{p_i^2(\tau)}{2} - \frac{p_i^2(-\tau)}{2}\right) + U_i(q_i(\tau)) - U_i(q_i(-\tau)) \\ + \int_{-\tau}^{\tau} \sum_{j:\, i\sim j} \lambda_{ij}\Phi'(q_i(t) - q_j(t))\, p_i(t)\, dt\bigg] \\ - \ln \rho^\beta(\omega_{-\tau}) + \ln \rho^\beta(\omega_\tau) + \ln \rho(\omega_{-\tau}) - \ln \rho(\omega_\tau)\end{aligned}$$

In the above, all terms in β are seen to cancel when (1.5) is inserted. By adding and subtracting the terms

$$\int_{-\tau}^{\tau} \sum_{j:\, j \sim i} \lambda_{ij} \Phi'(q_i(t) - q_j(t))\, p_j(t)\, dt, \qquad i \in \partial V$$

and using (3.3)–(3.5), we get the final result

$$\begin{aligned} R_\rho^\tau(\omega) &= - \sum_{i \in \partial V} \beta_i \left[H_i(\omega_\tau) - H_i(\omega_{-\tau}) + \sum_{j:\, i \sim j} J_{ij}^\tau(\omega) \right] + \ln \rho(\omega_{-\tau}) - \ln \rho(\omega_\tau) \\ &= \sum_{i \in \partial V} \beta_i J_i^\tau(\omega) + \ln \rho(\omega_{-\tau}) - \ln \rho(\omega_\tau) \qquad (3.11) \end{aligned}$$

We call this quantity the total variable entropy production. The first sum is the variable change of entropy of the reservoirs. In case $\rho = \rho_{-\tau}$ would not be stationary but some initial density, the analysis above would be essentially unchanged:

$$R_{\rho_{-\tau}}^\tau(\omega) = \sum_{i \in \partial V} \beta_i J_i^\tau(\omega) + \ln \rho_{-\tau}(\omega_{-\tau}) - \ln \rho_\tau(\omega_\tau) \qquad (3.12)$$

but the two last terms now give the difference

$$[-\ln \rho_\tau(\omega_\tau)] - [-\ln \rho_{-\tau}(\omega_{-\tau})]$$

obtaining its interpretation from what we mean by the statistics of the states of the system. We have not assumed that the graph or V is large (in the thermodynamic sense) and we have dealt with a Markov evolution on the level of the microscopic states of the system without reference to additional macroscopic quantities for the system. That makes it difficult to introduce the Boltzmann entropy. If however we assume that the initial preparation of the system has been done in some macrostate M and that the states (p, q) were uniformly distributed in the phase space region $M - \delta \leqslant M(p, q) \leqslant M + \delta$ (with δ some tolerance), then $-\ln \rho_{-\tau}(p, q) \simeq S_B(M)$ where $S_B(M)$ is the Boltzmann entropy of macrostate M. Similarly, if at time τ, the system is found in macrostate M' and if the Markov approximation was appropriate in the sense that ρ_τ corresponds to the microcanonical ensemble with constraint $M' - \delta \leqslant M(p, q) \leqslant M' + \delta$, then also $-\ln \rho_\tau(p, q) \simeq S_B(M')$. In that case, $R_{\rho_{-\tau}}^\tau = \Delta S_{\text{env}} + \Delta S_{\text{sys}}$ is the sum of the change of entropy in the environment plus the change of entropy in the system. In the case that the system is small or no standard entropy considerations for it can be made, one may want to think of $\ln \rho(p, q)$ as an information potential.

3.2. Mean Entropy Production

Denote by $\mathbb{E}_{\rho_{-\tau}}$ the expectation in the process $P^{\tau}_{\rho_{-\tau}}$ started from $\rho_{-\tau}$. We assume that at time τ the evolved measure is described by a density ρ_{τ}.

Proposition 3.1. For every initial density $\rho_{-\tau}$ the mean entropy production over the time interval $[-\tau, \tau]$ is non-negative:

$$\mathbb{E}_{\rho_{-\tau}}[R^{\tau}_{\rho_{-\tau}}] = \sum_{i\in\partial V} \beta_i \mathbb{E}_{\rho_{-\tau}}[J^{\tau}_{i}] + S(\rho_{\tau}) - S(\rho_{-\tau}) \geqslant 0 \tag{3.13}$$

where $S(\rho) = -\int dp\, dq\, \rho(p, q) \ln \rho(p, q)$ is the Shannon entropy of the density ρ. In particular, the steady state currents satisfy the inequality

$$\sum_{i\in\partial V} \beta_i \sum_{j:\, j\sim i} \langle J_{ji} \rangle \geqslant 0 \tag{3.14}$$

The inequality (3.14) implies that the steady state entropy production (3.8) is always non-negative. For simpler graphs (basically, one-dimensional) this was already achieved in refs. 2 and 6. We will discuss later when the steady state entropy production is strictly positive, see Section 3.3 answering Question 3.

Proof of Proposition 3.1. The proof is just repeating the arguments of ref. 8. By definition,

$$\mathbb{E}_{\rho_{-\tau}}[e^{-R^{\tau}_{\rho_{-\tau}}}] = 1$$

and by convexity of the exponential, we get

$$\mathbb{E}_{\rho_{-\tau}}[R^{\tau}_{\rho_{-\tau}}] \geqslant 0$$

This expectation can be done from (3.12). The rest follows from inspecting (3.6)–(3.8). ▮

Suppose that the density at time $t \geqslant -\tau$ is ρ_t when started from $\rho_{-\tau}$. We can write

$$\mathbb{E}_{\rho_{-\tau}}[R^{\tau}_{\rho_{-\tau}}] = \int_{-\tau}^{\tau} \dot{R}^{\tau}_{\rho_{-\tau}}(t)\, dt$$

with, similar to (3.7)–(3.11),

$$\dot{R}^{\tau}_{\rho_{-\tau}}(t) \equiv \gamma \sum_{i\in\partial V} \beta_i \left[\int dp\, dq\, p_i^2 \rho_t(p, q) - \frac{1}{\beta_i} \right] + \frac{d}{dt} S(\rho_t) \tag{3.15}$$

The previous considerations thus identify the mean entropy production rate at time t (in the transient regime) with $\dot{R}^{\tau}_{\rho_{-\tau}}(t)$.

In the following proposition we derive another formula for it, which is explicitly positive. We define the functional

$$\dot{S}(\rho) \equiv \sum_{i \in \partial V} \frac{\gamma}{\beta_i} \int dp\, dq \left[\frac{e^{-\beta_i p_i^2/2}}{\sqrt{\rho}} \frac{\partial}{\partial p_i} (e^{\beta_i p_i^2/2} \rho) \right]^2 \tag{3.16}$$

on smooth densities ρ.

Proposition 3.2.

$$\dot{R}^{\tau}_{\rho_{-\tau}}(t) = \dot{S}(\rho_t) \tag{3.17}$$

Proof. We start by evaluating the time derivative of the Shannon entropy:

$$\frac{dS}{dt}(\rho) = -\int dp\, dq \frac{d\rho}{dt} \ln \rho = -\int dp\, dq\, (\mathscr{L}^{+}\rho) \ln \rho \tag{3.18}$$

where $\mathscr{L}^{+}$ is the forward generator (the adjoint of $\mathscr{L}$ with respect to $dp\, dq$). We split it into the Hamiltonian and the reservoir parts, $\mathscr{L}^{+} = \mathscr{L}^{+}_{H} + \mathscr{L}^{+}_{R}$, with

$$\mathscr{L}^{+}_{H}\rho = -p \cdot \nabla_q \rho + \nabla_q U \cdot \nabla_p \rho \tag{3.19}$$

and

$$\mathscr{L}^{+}_{R}\rho = \gamma \sum_{i \in \partial V} \left[\frac{\partial}{\partial p_i} (p_i \rho) + \frac{1}{\beta_i} \frac{\partial^2 \rho}{\partial p_i^2} \right] = \sum_{i \in \partial V} \frac{\gamma}{\beta_i} \frac{\partial X_i}{\partial p_i} \tag{3.20}$$

where we made use of the shorthand $X_i \equiv e^{-\beta_i p_i^2/2} \frac{\partial}{\partial p_i} (e^{\beta_i p_i^2/2} \rho)$. Using the invariance of the Shannon entropy under Hamiltonian flows, we get

$$\begin{aligned} \frac{dS}{dt}(\rho) = -\int dp\, dq\, (\mathscr{L}^{+}_{R}\rho) \ln \rho &= \sum_{i \in \partial V} \frac{\gamma}{\beta_i} \int dp\, dq\, X_i \frac{\partial}{\partial p_i} \ln \rho \\ &= \sum_{i \in \partial V} \frac{\gamma}{\beta_i} \int dp\, dq\, X_i \left(\frac{X_i}{\rho} - \beta_i p_i \right) \end{aligned} \tag{3.21}$$

Minus the second term reads

$$\gamma \sum_{i \in \partial V} \int dp\, dq\, p_i X_i = \gamma \sum_{i \in \partial V} \int dp\, dq\, p_i \left(\frac{\partial \rho}{\partial p_i} + \beta_i p_i \rho \right)$$

$$= \gamma \sum_{i \in \partial V} \beta_i \int dp\, dq\, \rho \left(p_i^2 - \frac{1}{\beta_i} \right) \tag{3.22}$$

Substituting (3.22) into (3.21), we immediately obtain the desired identity. ∎

We end here with some second look at the functional $\dot{S}(\rho)$ of (3.16). Clearly, it offers no advance to try to compute the steady state entropy production from it if we do not know the stationary density. Moreover, it gives as such no hint about the variable entropy production which is a function on the trajectories in phase space and it is hopelessly restricted to the Markov case. This contrasts with the approach of Section 3.1. Nevertheless, as a functional on densities, it invites checking the so called minimum entropy production principle. However, a prime feature of the functional $\dot{S}$ is that it does not depend on the interaction potential. So we can minimize it easily and there are many solutions ($\dot{S}$ is not strictly convex). This makes that the minimum entropy production is not applicable here to characterize, even approximately, the stationary density. Note that, if both the stationary distribution and the minimizer of the entropy production functional are unique, they indeed coincide up to the linear order when expanded around a reversible dynamics, see ref. 9, for instance.

3.3. Strict Positivity of Mean Entropy Production

In this section we formulate sufficient conditions on the model under which the steady state produces entropy with non-zero rate. The conditions are not necessary and they can often be relaxed as we demonstrate on a specific example in the next section. This provides the answer to Question 3.

Theorem 3.3. If the interaction potential Φ has a second derivative that is m-non-degenerate for all $m \leqslant n(G, \partial V)$, and if some reservoir temperatures are not equal, then the steady state entropy production is strictly positive.

Proof. We see from (3.2) that if ρ is a smooth stationary density and if its steady state entropy production $\dot{S}(\rho) = 0$, then

$$\frac{\partial}{\partial p_i} (e^{\beta_i p_i^2/2} \rho) = 0, \qquad i \in \partial V$$

This is equivalent with (1.7). Moreover, it also implies that $\mathscr{L}_R^+\rho=0$, see (3.19). Hence, ρ is invariant under the Hamiltonian flow: $\mathscr{L}_H^+\rho=\mathscr{L}^+\rho-\mathscr{L}_R^+\rho=0$. The statement now immediately follows from Theorem 2.4. ▮

Theorem 3.3 is not entirely satisfactory. It is true, that there exist interesting classes of graphs with $n(G,\partial V)=1$, see Section 2.3, for which the condition of the theorem just boils down to the natural condition $\Phi\not\equiv$ const. Yet, the theorem cannot be applied when the number of neighbors is large, when there are many loops and the potential Φ is not sufficiently non-degenerate. For example, if G is a part of 2-dimensional lattice, we have no general (= valid independently on the choice of the boundary) result in the case that Φ is a polynomial of order lower than 4. We are however convinced that for Theorem 3.3, the assumption of non-degeneracy is most often stronger than needed. This will be illustrated in Section 3.4.

3.4. Non-Unique Stationary Measure

We still need to answer Question 2 and this takes us back to the example discussed in Section 2.4; we refer again to Fig. 1. We keep the same notation as there.

We first observe that

$$\frac{\partial}{\partial p_1}[e^{\beta p_1^2/2}\rho_c]=0, \qquad \frac{\partial}{\partial p_2}[e^{\beta p_2^2/2}\rho_c]=0 \tag{3.23}$$

This is equivalent to zero entropy production, $\dot{S}(\rho_c)=0$, and it also implies $\mathscr{L}_R^+\rho_c=0$. Combined with $\{W_c,H\}=0$, we conclude that the measures are stationary under the heat bath dynamics: $\mathscr{L}^+\rho_c=0$. It is easy to see that not only the entropy production but all local currents are zero, for instance,

$$\int dp\,dq\,J_{13}(p,q)\,\rho_c(p,q)=\frac{\lambda}{2}\int dp\,dq\,\rho_c(p,q)\,(p_1+p_3)(q_1-q_3)=0 \tag{3.24}$$

as well as $\langle p_i q_j\rangle_{\rho_c}=0$ for all $i,j=1,2,3,4$. Another observation is that the kinetic temperatures at different sites are not equal, in general. Indeed, a simple calculation yields

$$\langle p_3^2\rangle_{\rho_c}=\langle p_4^2\rangle_{\rho_c}=\frac{1+c}{1+2c}\,\beta^{-1} \tag{3.25}$$

while $\langle p_1^2\rangle_{\rho_c}=\langle p_2^2\rangle_{\rho_c}=\beta^{-1}$.

Assume now that the temperatures of both reservoirs are different. We sketch how to prove in this case that $\dot{S}(\rho)=0$ and $\mathscr{L}^{+}\rho=0$ imply that $\beta_1=\beta_2$. We write again $\rho=\exp(-W)/Z$ and introduce, for convenience, $W_1=W-\beta_1 H$ and $W_2=W-\beta_2 H$. Then the condition $\dot{S}(\rho)=0$ is equivalent to

$$\frac{\partial W_1}{\partial p_1}=0, \qquad \frac{\partial W_2}{\partial p_2}=0 \tag{3.26}$$

Similarly, the stationarity condition $\mathscr{L}^{+}\rho=0$ is equivalent to $\{W, H\}=0$ which implies the equations

$$\{W_1, H\}=0, \qquad \{W_2, H\}=0 \tag{3.27}$$

Differentiating (3.27) with respect to p_1, respectively p_2, and using (3.26), we get

$$\frac{\partial W_1}{\partial q_1}=0, \qquad \frac{\partial W_2}{\partial q_2}=0 \tag{3.28}$$

Taking now the derivatives of (3.27) with respect to q_1 and q_2, we obtain the equations

$$\frac{\partial W_1}{\partial p_3}+\frac{\partial W_1}{\partial p_4}=0, \qquad \frac{\partial W_2}{\partial p_3}+\frac{\partial W_2}{\partial p_4}=0 \tag{3.29}$$

which are mutually in contradiction unless $\beta_1=\beta_2$. This shows that the assumption of non-degeneracy as stated in Theorem 3.3 is sometimes too strong. To apply it, we would need here $n=2$ but Φ'' is constant.

3.5. Local Heat Currents

We address here Question 4. A first answer concerning the direction of the heat currents is obtained from the positivity of the entropy production. Suppose for example that the boundary $\partial V=D_1\cup D_2$ splits into two non-empty disjoint regions with $\beta_i=\beta_1$ for $i\in D_1$ and $\beta_i=\beta_2$ for $i\in D_2$. Then, one has

$$\Psi_1=\sum_{i\in D_1}\left[\langle p_i^2\rangle-\frac{1}{\beta_1}\right]=-\Psi_2=-\sum_{i\in D_2}\left[\langle p_i^2\rangle-\frac{1}{\beta_2}\right]$$

which, for the steady state entropy production give

$$(\beta_1-\beta_2)\,\Psi_1\geqslant 0 \tag{3.30}$$

If therefore $\beta_1 \geqslant \beta_2$ (D_1 is colder than D_2), then $\Psi_1 \geqslant 0 \geqslant \Psi_2$ or the heat flows effectively from D_2 to D_1. All these inequalities become strict when the entropy production is strictly positive as discussed in Section 3.3.

Suppose now that the steady entropy production equals zero. We know that under certain conditions on the potential, as discussed in Section 3.3, we must have that all reservoir temperatures are equal to some β and that the density ρ^β is the unique stationary density. Here we show, without further explicit conditions on the potential but assuming uniqueness of the stationary density, that at any event, zero entropy production implies that all local heat currents are zero. We actually show something stronger: that the stationary density is invariant under sign reversals of all the momenta, $\rho(p, q) = \rho\pi(p, q)$. This will follow from the time-reversal invariance of the steady state process, $P^\tau_\rho = P^\tau_{\rho\pi}\Theta$ which is itself a consequence of

Theorem 3.4. Suppose that ρ is a stationary density for which $\dot{S}(\rho) = 0$. Then the generator satisfies $\mathscr{L} = \pi \mathscr{L}^* \pi$ in $L^2(\rho)$.

Corollary 3.5. Suppose that ρ is the unique stationary density and that $\dot{S}(\rho) = 0$. Then, $\rho = \rho\pi$ and as a result, for all nearest neighbors $i, j \in V$,

$$\int dp\, dq\, \rho(p, q)\, J_{ij}(q, p) = \int dp\, dq\, \rho(p, q)\, \tfrac{1}{2}\, \lambda_{ij}(p_i + p_j)\, \Phi'(q_i - q_j) = 0 \tag{3.31}$$

Proof of Theorem 3.4. We start by rewriting

$$\mathscr{L}_R f(p, q) = \gamma \sum_{i \in \partial V} \frac{1}{\beta_i}\, e^{\beta_i p_i^2/2} \left(\frac{\partial}{\partial p_i}\, e^{-\beta_i p_i^2/2}\, \frac{\partial f}{\partial p_i} \right)$$

so that it follows from $\dot{S}(\rho) = 0$ that

$$\int g(p, q)\, \mathscr{L}_R f(p, q)\, \rho(p, q)\, dp\, dq = -\gamma \sum_{i \in \partial V} \frac{1}{\beta_i} \int \frac{\partial g}{\partial p_i} \frac{\partial f}{\partial p_i}\, \rho(p, q)\, dp\, dq$$

which is symmetric under exchanging f with g. Moreover, $\mathscr{L}_R \pi = \pi \mathscr{L}_R$.

For the Hamiltonian flow, we have, by dynamical reversibility, $\pi \mathscr{L}_H \pi = -\mathscr{L}_H$ and again, since $\dot{S}(\rho) = 0$ combined with $\mathscr{L}^+ \rho = 0$ imply that $\mathscr{L}^+_H \rho = 0$, we also have $\mathscr{L}^*_H = \mathscr{L}^+_H = -\mathscr{L}_H$. ∎

Proof of Corollary 3.5. The symmetry $\mathscr{L} = \pi \mathscr{L}^* \pi$ in $L^2(\rho)$ of the generator implies that also $\rho\pi$ is stationary for the heat bath dynamics. Since it is assumed unique, we conclude that $\rho = \rho\pi$. ∎

Note that $\mathscr{L} = \pi\mathscr{L}^*\pi$ in $L^2(\rho)$ implies that $\mathscr{L}\rho\pi/\rho = 0$. An additional ergodicity requirement (that all ergodic components are π-invariant) thus makes $\rho = \rho\pi$ and can thus replace the uniqueness assumption of Corollary 3.5.

The corollary further shows that zero steady entropy production implies that the stationary density is invariant under reflection of the momenta. For our finite graphs, this property is certainly not enough to conclude that ρ is a convex combination of ρ^{β}'s, Gibbs measures at inverse temperature β, see Section 2.4 but, paradoxically, it may be sufficient for infinite graphs, see ref. 4.

4. GETTING WORK DONE

One of the prime applications of maintaining heat gradients is to get work done. This can be modeled after the scheme of Section 1.2 by inserting external forces $F_i = F_i(q)$ at the $i \in V \setminus \partial V$. We then have the forced heat bath dynamics

$$\begin{aligned}
dq_i &= p_i\, dt, \qquad i \in V \\
dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt + F_i\, dt, \qquad i \in V \setminus \partial V \\
dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt - \gamma p_i + \sqrt{\frac{2\gamma}{\beta_i}}\, dW_i(t), \qquad i \in \partial V
\end{aligned} \tag{4.1}$$

We will again assume the existence and smoothness properties of the process with stationary density ρ.

The work done by the system over the time-interval $[-\tau, \tau]$ is

$$W_o^\tau \equiv - \sum_{i \in V \setminus \partial V} \int_{-\tau}^{\tau} F_i\, p_i\, dt \tag{4.2}$$

We will show that for an arbitrary $\beta > 0$,

$$\langle W_o^\tau \rangle \leqslant \sum_{i \in \partial V} \left(\frac{\beta_i}{\beta} - 1 \right) \langle J_i^\tau \rangle \tag{4.3}$$

Fixing an arbitrary boundary site $v \in \partial V$ and taking $\beta = \beta_v$ in (4.3) yields

$$\langle W_o^\tau \rangle \leqslant \sum_{i \in \partial V} \frac{\beta_i - \beta_v}{\beta_v} \langle J_i^\tau \rangle \tag{4.4}$$

In fact, (4.3) and (4.4) are equivalent (as can be seen by eliminating $\langle J_v^\tau \rangle$ in (4.3) via (4.6)).

In the case where the boundary $\partial V = \{v, w\}$ contains just two sites (system coupled to two heat reservoirs) and we write $\beta_v \equiv 1/T_v$, $\beta_w \equiv 1/T_w$, then (4.4) reads as

$$\langle W_o^\tau \rangle \leqslant \left(\frac{T_v}{T_w} - 1\right) \langle J_w^\tau \rangle \tag{4.5}$$

which gives an upper bound on the work $\mathscr{W} \equiv \langle W_o^\tau \rangle$ that can be extracted from the engine: If $T_{\text{hot}} \equiv T_w > T_v \equiv T_{\text{cold}}$ and $Q_{\text{hot}} \equiv -\langle J_w^\tau \rangle > 0$ (energy input at the hot reservoir), we get as maximal efficiency that of the Carnot cycle:

$$\frac{\mathscr{W}}{Q_{\text{hot}}} \leqslant 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}}$$

For a heat pump, when now $T_{\text{hot}} = T_v > T_w = T_{\text{cold}}$ and $Q_{\text{cold}} = -\langle J_w^\tau \rangle > 0$ (energy input at the cold reservoir), we get from (4.5) a lower bound on the work $\mathscr{W}' = -\langle W_o^\tau \rangle$ that needs to be delivered on the system:

$$\frac{\mathscr{W}'}{Q_{\text{cold}}} \geqslant \frac{T_{\text{hot}}}{T_{\text{cold}}} - 1$$

Proof of (4.3). The first remark is that the global energy conservation relation (3.2) in the steady state must be changed to

$$\sum_{i \in \partial V} \langle J_i^\tau \rangle + \langle W_o^\tau \rangle = 0 \tag{4.6}$$

The inequality (4.3) is then obtained once we show that

$$\sum_{i \in \partial V} \beta_i \langle J_i^\tau \rangle \geqslant 0 \tag{4.7}$$

We only need to replace the right-hand side in (4.7) by β times the left-hand side of (4.6) for arbitrary $\beta > 0$.

It remains to show that (4.7) is true but that follows from the same analysis as in Section 3.1 that finally lead to (3.14). In other words, the expression for the total entropy production is not affected by the presence of the forces F_i. One way to see this via some regularization is by considering the following modification of (4.1):

$$\begin{aligned} dq_i &= p_i\, dt, \qquad i \in V \\ dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt + aF_i(q)\, dt - \kappa_i \varepsilon p_i\, dt + (4\varepsilon)^{1/4}\, dW_t, \qquad i \in V\setminus\partial V \\ dp_i &= -\frac{\partial U}{\partial q_i}(q)\, dt - \gamma\kappa_i\, p_i + \sqrt{\frac{2\gamma}{\beta_i}}\, dW_i(t), \qquad i \in \partial V \end{aligned} \tag{4.8}$$

for $\varepsilon > 0$. A first choice is taking $a = 0$, for $i \in \partial V$: $\kappa_i \beta_i = 1$ and for $i \in V\setminus\partial V$: $\kappa_i \sqrt{\varepsilon} = 1$. For this choice, the process is reversible exactly like in (3.9) with reversible measure $\rho^{\beta=1}$. A second choice is taking $a = \kappa_i = 1$ for all $i \in V$ which coincides with our original dynamics (4.1) when we let $\varepsilon \downarrow 0$. Just like in Section 3.1 we can compute the density of the process $P_\rho^{\tau,2}$, corresponding to our second choice, with respect to the process $P_{\rho^1}^{\tau,1}$ of our first choice:

$$dP_\rho^{\tau,2}(\omega) = e^{-A_\rho^\varepsilon(\omega)}\, dP_{\rho^1}^{\tau,1}(\omega)$$

where A_ρ^ε is given by a Girsanov formula. The source of time-reversal breaking is now $R_\rho^{\tau,\varepsilon}(\omega) \equiv A_{\rho\pi}^\varepsilon(\Theta\omega) - A_\rho^\varepsilon(\omega)$ and one can easily check that it coincides with $R_\rho^\tau(\omega)$ as given in (3.11) up to terms of order $\sqrt{\varepsilon}$ for each history ω. We now let $\varepsilon \downarrow 0$ and we get $R_\rho^{0,\tau} = R_\rho^\tau$. The rest is again an application of Proposition 3.1, inequality (3.14). ∎

ACKNOWLEDGMENTS

We thank F. Redig for many useful discussions.

REFERENCES

1. J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures, *Commun. Math. Phys.* **201**:657 (1999).
2. J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, Entropy production in nonlinear, thermally driven Hamiltonian systems, *J. Stat. Phys.* **95**:305 (1999).
3. G. W. Ford, M. Kac, and P. Mazur, Statistical mechanics of assemblies of coupled oscillators, *J. Math. Phys.* **6**:504–515 (1965).
4. J. Fritz, Stationary states of hamiltonian systems with noise, in *On Three Levels*, M. Fannes, C. Maes, and A. Verbeure, eds. (Plenum Press, New York, 1994), pp. 203–214.
5. J. Fritz, T. Funaki, and J. L. Lebowitz, Stationary states of random hamiltonian systems, *Probab. Theory Related Fields* **99**:211–236 (1994).

6. J. L. Lebowitz, Stationary nonequilibrium Gibbsian ensemble, *Phys. Rev.* **114**:1192–1202 (1959).
7. R. S. Lipster and A. N. Shiryayev, *Statistics of Random Processes I, II* (Springer-Verlag, New York/Heidelberg/Berlin, 1978).
8. C. Maes, Fluctuation theorem as a Gibbs property, *J. Stat. Phys.* **95**:367–392 (1999).
9. C. Maes and K. Netočný, Time-reversal and entropy, *J. Stat. Phys.* **110**:269–310 (2003).
10. H. Nakazawa, On the lattice thermal conduction, *Suppl. Prog. Theor. Phys.* **45**:231–262 (1970).
11. L. Rey-Bellet and L. E. Thomas, Fluctuations of the entropy production in anharmonic chains, *Ann. Henri Poincaré* **3**:483–502 (2002).
12. L. Rey-Bellet and L. E. Thomas, Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators, *Commun. Math. Phys.* **215**:1–24 (2000).
13. Z. Rieder, J. L. Lebowitz, and E. Lieb, Properties of a harmonic crystal in a stationary nonequilibrium state, *J. Math. Phys.* **8**:1073–1078 (1967).
14. H. Spohn and J. L. Lebowitz, Stationary non-equilibrium states of infinite harmonic systems, *Commun. Math. Phys.* **54**:97–120 (1977).