PRESSURE AND HELMHOLTZ FREE ENERGY IN A DYNAMIC MODEL OF A LATTICE GAS

RICHARD HOLLEY
PRINCETON UNIVERSITY

1. Introduction

In this paper, we will study a model of an infinite volume one dimensional lattice gas. Our model differs from the usual model of a lattice gas in that the configuration of particles is a stochastic process. That is, the particles in the system will move around, and we will be studying properties of the system which are related to the motion. The particular interaction which governs the behavior of each particle will be introduced in Section 2. This interaction was discovered by F. Spitzer [4].

In Section 2, we will define the Helmholtz free energy in the usual way and prove that at constant temperature the Helmholtz free energy does not increase with time. In thermodynamics, this is usually derived as a consequence of the second law of thermodynamics. In Section 3, we will use the results obtained in Section 2 to prove that all shift invariant equilibrium states are limiting Gibbs distributions. Finally, in Section 4, we use the intuitive description of the interaction of the particles to motivate a definition for the pressure of a state. The usual definition of pressure used in statistical mechanics is only given for limiting Gibbs distributions, and the two definitions do not agree there. However, we will show that when they are both defined, they are both strictly increasing functions of the particle density at constant temperature. In the case of the usual definition this is well known.

In order to keep the notation as simple as possible, we will only consider one model in this paper. This model can clearly be generalized in several ways. Many of these generalizations can be found in [4]. The techniques in this paper are adequate to handle some, though by no means all, of these generalizations.

2. Helmholtz free energy

2.1. *Intuitive description.* We begin by giving an intuitive description of the stochastic process. For a careful proof that this process really exists the reader is referred to [1].

This work was prepared while the author was a Miller Fellow in the Statistics Department, University of California, Berkeley.

Let Z represent the integers, and give $\{0, 1\}$ the discrete topology. We will take $E = \{0, 1\}^Z$ with the product topology for the state space. If $\eta \in E$, η will be interpreted as a configuration of particles on the integers with a particle at x if and only if $\eta(x) = 1$.

Let V be a real valued function on the nonnegative integers such that for some positive integer L, $V(n) = 0$ if $n > L$. We define a pair potential $U(x, y)$ by the formula $U(x, y) = V(|x - y|)$. We think of $U(x, y)$ as the potential energy due to particles at x and y, and attribute half of this energy to each of the particles. If the system is in configuration η and $\eta(x) = 1$, then the particle at x has energy equal to

$$\tfrac{1}{2} \sum_y U(x, y)\eta(y), \tag{2.1}$$

$\frac{1}{2}U(x, x)$ represents the chemical potential of the particle at x.

Now let β be a positive constant which will represent the reciprocal of the temperature. Then if the system is in configuration η at time t, the particle at x attempts to make a jump during the time interval $(t, t + \Delta t)$ with probability

$$\exp \{\beta \sum_y U(x, y)\eta(y)\}\Delta t + o(\Delta t). \tag{2.2}$$

When a jump is attempted, the particle tries to move to the right one or to the left one, each with probability $\frac{1}{2}$. The direction of the attempted jump is independent of the time of the jump and the position of the particle. If the site where it is trying to go is unoccupied, it goes there. Otherwise, it remains where it is and starts over. It turns out that although the motion of each individual particle is not Markovian, the stochastic process of the configuration η_t is Markovian; we will denote the probability that the system goes from configuration η to a configuration in a Borel set $A \subseteqq E$ in time t by $P^\eta(\eta_t \in A)$.

The probability measures on the Borel subsets of E will be called states. If μ_0 is any initial state, we will define μ_t to be the state which assigns measure

$$\mu_t(A) = \int P^\eta(\eta_t \in A)\mu_0(d\eta) \tag{2.3}$$

to the set A.

A measure μ_0 will be called an equilibrium state, if $\mu_0 = \mu_t$ for all $t \geqq 0$. In [1], a set C_0 of equilibrium states for this Markov process is given. We will describe below a set C which certainly contains C_0 and at first glance looks as though it may be strictly larger than the closed convex hull of C_0. It is probably true that C is equal to the closed convex hull of C_0, although we have been unable to prove this. A proof that all of the measures in C are equilibrium states can be accomplished by a slight modification of the proof in [1].

2.2. *The set C.* Let N be an integer greater than L and let Y be a subset of the integers contained in $[-N, N]\backslash[-N + L, N - L]$. Define

$$(2.4)\quad \begin{aligned} &S(N, n, Y) \\ &= \left\{ \eta \in E \,\middle|\, \sum_{-N+L}^{N-L} \eta(x) = n, \quad \text{if} \quad N - L < |y| \leqq N \quad \text{then} \quad \eta(y) = 1 \right. \\ &\qquad \left. \text{if and only if} \quad y \in Y, \quad \text{and if} \quad |z| > N \quad \text{then} \quad \eta(z) = 0 \right\}. \end{aligned}$$

Now let $\nu_{N,n,Y}$ be the probability measure on $S(N, n, Y)$ given by the formula

$$(2.5)\qquad \nu_{N,n,Y}(\{\eta\}) = \psi(N, n, Y) \exp\left\{ -\tfrac{1}{2}\beta \sum_{a=-N}^{N} \sum_{b=-N}^{N} U(a, b)\eta(a)\eta(b) \right\}.$$

Here $\psi(N, n, Y)$ is the normalizing constant.

We may think of $\nu_{N,n,Y}$ as a probability measure on E which gives zero measure to the complement of $S(N, n, Y)$. Now let ν_N be any convex combination of the $\nu_{N,n,Y}$, where n is allowed to vary between 0 and $2N - 2L + 1$ and Y varies over all subsets of $[-N, N]\backslash[-N + L, N - L]$.

If we do this for each N, we get a sequence of probability measures on a compact metric space. The set C consists of all the possible weak limit points of sequences obtained in this way.

We still need a little more notation. Let Λ be a finite subset of Z and let μ be a probability measure on E. Then μ^Λ will denote the probability measure on the subsets of Λ defined by

$$(2.6)\qquad \mu^\Lambda(X) = \mu(\{\eta \mid \text{if} \quad x \in \Lambda, \quad \text{then} \quad \eta(x) = 1 \quad \text{if and only if} \quad x \in X\}).$$

If $\Lambda = [-N, N]$, we will use μ^N instead of $\mu^{[-N,N]}$.

Let us recall the definition of Helmholtz free energy in thermodynamics. If U represents the internal energy of a system in a certain state, and S and T are, respectively, the entropy and temperature of that state, then the Helmholtz free energy of that state is defined to be $U - ST$. For an infinite volume system, such as the one with which we are dealing, both the internal energy and the entropy may be infinite. In that case, this definition does not make any sense; however, we can define the Helmholtz free energy per site.

Let μ be a state on E. We define its Helmholtz free energy per site, $A(\mu)$, as follows

$$(2.7)\quad \begin{aligned} &A(\mu) \\ &= \limsup_{N\to\infty} \frac{1}{2N + 1}\left(\sum_{X \subseteqq [-N,N]} U(X)\mu^N(X) + T \sum_{X \subseteqq [-N,N]} \mu^N(X) \log \mu^N(X) \right). \end{aligned}$$

Here $U(X) = \frac{1}{2} \Sigma_{x,y\in X}\, U(x, y)$. The first sum in the definition of $A(\mu)$ represents the internal energy of the state μ between $-N$ and N, while the second sum represents the negative of the entropy of μ between $-N$ and N.

Throughout this paper 0 log (0) is understood to be zero. Some of our proofs require special attention to the case when $0 \log (0)$ appears in an expression. The modifications necessary then will be left to the reader.

An essential tool which we will need is the infinitesimal generator of the Markov process η_t. It is proved in [1] that if f is a continuous function on E which depends on only finitely many coordinates, then the infinitesimal generator Ω operating on f is given by

$$\Omega f(\eta) = \sum_{x, y \in Z, |x-y|=1} \eta(x)[1 - \eta(y)]c(x, \eta)[f(\eta_{x,y}) - f(\eta)], \tag{2.8}$$

where $c(x, \eta) = \frac{1}{2} \exp \{\beta \Sigma_{w \in Z} \eta(w) U(x, w)\}$, and

$$\eta_{x,y}(w) = \begin{cases} \eta(w) & \text{if } w \neq x, w \neq y, \\ 0 & \text{if } w = x, \\ 1 & \text{if } w = y. \end{cases} \tag{2.9}$$

The use we will make of this is given in the following lemma.

LEMMA 2.1. *Let* $W(X \cup Y, a) = \frac{1}{2} \exp \{\beta \Sigma_{c \in X \cup Y} U(a, c\}$ *and*

$$D(N, X, Y) = \{(a, b) | a \in X \cup Y, b \notin X \cup Y, |a - b| = 1, \text{ and } |a| \leqq N \text{ or } |b| \leqq N\}. \tag{2.10}$$

Then

$$\frac{d}{dt} \mu_t^N(X) = \sum_Y \sum_{(a,b) \in D(N,X,Y)} W(X \cup Y \cup b \backslash a, b) \mu_t^{N+L}(X \cup Y \cup b \backslash a) - \sum_Y \sum_{(a,b) \in D(N,X,Y)} W(X \cup Y, a) \mu_t^{N+L}(X \cup Y). \tag{2.11}$$

In (2.10) *the summation over* Y *is over all subsets* Y *of* $[-N - L, -N - 1] \cup [N + 1, N + L]$.

Throughout this paper if X and Y are subsets of the integers, and a and b are integers, we will write $X \cup Y \cup b \backslash a$ instead of $(X \cup Y \cup \{b\}) \backslash \{a\}$.

The proof of Lemma 2.1 is simply an application of (2.8) and will be left to the reader. Recall that we are assuming that $U(a, b) = 0$ if $|a - b| > L$.

THEOREM 2.1. *With* $A(\cdot)$ *and* μ_t *as defined above,* $A(\mu_t)$ *is a nonincreasing function of* t.

PROOF. Recalling that $\beta = 1/T$, we may revise (2.7) slightly to obtain

$$\begin{aligned} A(\mu_t) &= \limsup_{N \to \infty} \frac{1}{\beta} \frac{1}{2N + 1} \sum_{X \subseteq [-N, N]} \mu_t^N(X)[\log \mu_t^N(X) + \beta U(X)] \\ &= \limsup_{N \to \infty} \frac{1}{\beta} \frac{1}{2N + 1} \sum_{X \subseteq [-N, N]} \mu_t^N(X) \log \frac{\mu_t^N(X)}{P(X)}, \end{aligned} \tag{2.12}$$

where $P(X) = \exp \{-\beta U(X)\}$. To finish the proof, it will be sufficient to show that

$$\frac{d}{dt} \sum_{X \subseteq [-N, N]} \mu_t^N(X) \log \frac{\mu_t^N(X)}{P(X)} \tag{2.13}$$

is bounded above by some constant which is independent of μ_t and N. Now

$$\sum_{X \subseteq [-N,N]} \frac{d}{dt} \mu_t^N(X) = 0, \tag{2.14}$$

and therefore, interchanging the summation and differentiation in (2.13) yields

$$\sum_{X \subseteq [-N,N]} \left\{\frac{d}{dt} \mu_t^N(X)\right\} \log \frac{\mu_t^N(X)}{P(X)}. \tag{2.15}$$

Note that we are sure that the expression in (2.13) exists for all t only if we permit the derivative to take the value minus infinity. If for some X, $\mu_t^N(X) = 0$ and $(d/dt)\, \mu_t^N(X) > 0$, then both (2.13) and (2.15) are minus infinity. If for some X, $\mu_t^N(X) = 0$ and $(d/dt)\, \mu_t^N(X) = 0$, then by first using (2.8) to prove that $\mu_t^N(X)$ has two continuous derivatives, it can be seen that $(d/dt)\big(\mu_t^N(X) \log \mu_t^N(X)\big) = 0$. Thus, using our convention about $0 \log (0)$, we can then write

$$\frac{d}{dt}\big(\mu_t^N(X) \log \mu_t^N(X)\big) = \left(\frac{d}{dt} \mu_t^N(X)\right) \log \mu_t^N(X), \tag{2.16}$$

and it is still true that (2.15) equals (2.13).

We will omit the subscript t from the notation during the rest of the proof. Substituting (2.10) into (2.15), we have

$$\begin{aligned} \sum_X \Bigg[&\sum_Y \sum_{(a,b)} W(X \cup Y \cup b\backslash a, b)\mu^{N+L}(X \cup Y \cup b\backslash a) \\ &- \sum_Y \sum_{(a,b)} W(X \cup Y, a)\mu^{N+L}(X \cup Y)\Bigg] \log \frac{\mu^N(X)}{P(X)}, \end{aligned} \tag{2.17}$$

where the summations are on $(a, b) \in D(N, X, Y,)$. Now set

$$D_1(N, X, Y) = D(N, X, Y) \cap \{(a, b) \,|\, |a| \leqq N - L \text{ and } |b| \leqq N - L\} \tag{2.18}$$

and $D_2(N, X, Y) = D(N, X, Y)\backslash D_1(N, X, Y)$. Note that $D_1(N, X, Y)$ does not depend on Y; hence, we will write it $D_1(N, X)$. Expression (2.17) can be broken into two terms, one with D_1 replacing D and the other with D_2 replacing D. We first consider the expression resulting when D is replaced by D_2:

$$\begin{aligned} &\sum_X \Bigg[\sum_Y \sum_{(a,b)} W(X \cup Y \cup b\backslash a, b)\mu^{N+L}(X \cup Y \cup b\backslash a) \\ &\qquad - \sum_Y \sum_{(a,b)} W(X \cup Y, a)\mu^{N+L}(X \cup Y)\Bigg] \log \frac{\mu^N(X)}{P(X)} \\ &= \sum_X \sum_Y \sum_{(a,b)} W(X \cup Y \cup b\backslash a, b)\mu^{N+L}(X \cup Y \cup b\backslash a) \\ &\qquad \left[\log \frac{\mu^N(X)}{P(X)} - \log \frac{\mu^N(X \cup b\backslash a)}{P(X \cup b\backslash a)}\right], \end{aligned} \tag{2.19}$$

where the summations are on $(a, b) \in D_2(N, X, Y)$. One of a or b may not be in

$[-N, N]$. In that case, we must understand $\mu^N(X \cup b \backslash a)$ to be $\mu^N(X \cup b)$ if $a \notin [-N, N]$ or to be $\mu^N(X \backslash a)$ if $b \notin [-N, N]$. Similarly, for $P(X \cup b \backslash a)$.

If there exists X, Y, and $(a, b) \in D_2(N, X, Y,)$ such that $\mu^N(X) = 0$ and $\mu^{N+L}(X \cup Y \cup b \backslash a) > 0$, then (2.19) is $-\infty$. Therefore, since we are trying to show that (2.19) is bounded above, we may assume that if $\mu^{N+L}(X \cup Y \cup b \backslash a) > 0$, then $\mu^N(X) > 0$.

Now

$$\begin{aligned}(2.20)\quad & \mu^{N+L}(X \cup Y \cup b \backslash a)\left[\log \frac{\mu^N(X)}{P(X)} - \log \frac{\mu^N(X \cup b \backslash a)}{P(X \cup b \backslash a)}\right] \\ &= \mu^{N+L}(X \cup Y \cup b \backslash a)\Bigg[\log \frac{\mu^N(X)}{\mu^N(X \cup b \backslash a)} \\ &\qquad + \tfrac{1}{2}\beta \sum_{m,n \in X} U(m, n) - \tfrac{1}{2}\beta \sum_{m,n \in X \cup b \backslash a} U(m, n)\Bigg] \\ &\leqq -\frac{\mu^{N+L}(X \cup Y \cup b \backslash a)}{\mu^N(X)} \log\left(\frac{\mu^N(X \cup b \backslash a)}{\mu^N(X)}\right)\mu^N(X) \\ &\qquad + \mu^{N+L}(X \cup Y \cup b \backslash a)K,\end{aligned}$$

where $K = 2\beta \sum_{k=-L}^{L} |U(0, k)|$.

Since $\mu^{N+L}(X \cup Y \cup b \backslash a) \leqq \mu^N(X \cup b \backslash a)$ and $-x \log x \leqq e^{-1}$, the right side of (2.20) is bounded above by $e^{-1}\mu^N(X) + K\mu^{N+L}(X \cup Y \cup b \backslash a)$.

From the definition of $W(X \cup Y, a)$, we see that $W(X \cup Y, a) \leqq \frac{1}{2} \exp \frac{1}{2}K$. Substituting this bound and the bound for (2.20) into (2.19), we see that (2.19) is bounded above by

$$\begin{aligned}(2.21)\quad & \tfrac{1}{2} \exp\{\tfrac{1}{2}K\} \sum_X \sum_Y \sum_{(a,b) \in D_2(N,X,Y)} [\exp\{-1\}\mu^N(X) + K\mu^{N+L}(X \cup Y \cup b \backslash a)] \\ &\leqq \tfrac{1}{2} \exp\{\tfrac{1}{2}K - 1\}2^{2L}4L + \tfrac{1}{2}K \exp\{\tfrac{1}{2}K\}4L.\end{aligned}$$

In the future, we will denote the right side of (2.21) by K_1.

We now return to (2.17) and consider the expression obtained if $D(N, X, Y)$ is replaced by $D_1(N, X)$. The first thing to note is that

$$\begin{aligned}(2.22)\quad & \sum_X \Big[\sum_Y \sum_{(a,b) \in D_1(N,X)} W(X \cup Y \cup b \backslash a, b)\mu^{N+L}(X \cup Y \cup b \backslash a) \\ &\qquad - \sum_Y \sum_{(a,b) \in D_1(N,X)} W(X \cup Y, a)\mu^{N+L}(X \cup Y)\Big] \log \frac{\mu^N(X)}{P(X)} \\ &= \sum_X \Big[\sum_{(a,b) \in D_1(N,X)} W(X \cup b \backslash a, b)\mu^N(X \cup b \backslash a) \\ &\qquad - \sum_{(a,b) \in D_1(N,X)} W(X, a)\mu^N(X)\Big] \log \frac{\mu^N(X)}{P(X)}.\end{aligned}$$

For if $(a, b) \in D_1(N, X)$, then neither $W(X \cup Y, a)$ nor $W(X \cup Y \cup b \backslash a, b)$ depend on Y. Hence, we may first perform the summation on Y.

LEMMA 2.2. *Let D_1, W, and P be as above. For $z \geqq 0$ set $F(z) = z - z\log(z) - 1$. Then*

$$\sum_X \Big[\sum_{(a,b)\in D_1(N,X)} W(X \cup b\backslash a, b)\mu^N(X \cup b\backslash a) - \sum_{(a,b)\in D_1(N,X)} W(X, a)\mu^N(X)\Big] \log \frac{\mu^N(X)}{P(X)} \tag{2.23}$$

$$= \sum_X \sum_{D_1(N,X)} F\left(\frac{P(X)}{\mu^N(X)} \frac{\mu^N(X\cup b\backslash a)}{P(X\cup b\backslash a)}\right) W(X\cup b\backslash a, b) \frac{P(X\cup b\backslash a)}{P(X)} \mu^N(X).$$

REMARK 2.1. For (2.23) to be correct, we must make the following convention. For $a, b, c > 0$ we understand $F(\frac{a}{0}\frac{b}{c})\cdot 0$ to be minus infinity and $F(\frac{a}{0}\frac{0}{c})\cdot 0$ to be zero.

Let us assume Lemma 2.2 for the moment. One easily checks that $F(z) \leqq 0$ for all $z \geqq 0$, and thus the expression appearing in (2.23) is nonpositive. Since (2.17) is the sum of (2.22) and (2.19), we see from the above results that (2.17) is bounded above by K_1, which is independent of N and μ.

The proof will be complete as soon as we prove Lemma 2.2.

PROOF OF LEMMA 2.2. Let each $X \subseteqq [-N, N]$ be represented as $X_0 \cup X_1$, where $X_0 \subseteqq [-N, N]\backslash[-N + L, N - L]$ and $X_1 \subseteqq [-N + L, N - L]$. Then we can rewrite the left side of (2.23) to get

$$\sum_{X_0}\Bigg[\sum_{X_1}\sum_{(a,b)\in D_1(X_1)} W(X_0\cup X_1\cup b\backslash a, b)\mu^N(X_0\cup X_1\cup b\backslash a)\log\frac{\mu^N(X_0\cup X_1)}{P(X_0\cup X_1)} - \sum_{X_1}\sum_{(a,b)\in D_1(X_1)} W(X_0\cup X_1, a)\mu^N(X_0\cup X_1)\log\frac{\mu^N(X_0\cup X_1)}{P(X_0\cup X_1)}\Bigg]. \tag{2.24}$$

We write $D_1(X_1)$ instead of $D_1(N, X_0\cup X_1)$, since $D_1(N, X_0\cup X_1)$ depends only on X_1 and N is fixed throughout the proof.

We next notice that for fixed X_0 there is a kernel $\mathfrak{U}_{X_0}(\cdot,\cdot)$ defined on the subsets of $[-N + L, N - L]$ by the formula

$$\mathfrak{U}_{X_0}(A, B) = \begin{cases} W(X_0\cup A, a) & \text{if } B = A\cup b\backslash a \text{ for some } a\in A,\ b\notin A \text{ with } |a - b| = 1, \\ -\sum_{(a,b)\in D_1(A)} W(X_0\cup A, a) & \text{if } B = A, \\ 0 & \text{otherwise,} \end{cases} \tag{2.25}$$

and that

$$\sum_A P(X_0\cup A)\mathfrak{U}_{X_0}(A, B) = 0 \tag{2.26}$$

for all X_0 and all B.

This last assertion is an easy computation which will be left to the reader.

Now (2.24) may be rewritten to yield

$$\sum_{X_0}\left[\sum_A\sum_B \mathfrak{U}_{X_0}(A, B)\mu^N(X_0\cup A)\log\frac{\mu^N(X_0\cup B)}{P(X_0\cup B)}\right]. \tag{2.27}$$

Since $\Sigma_B\, \mathfrak{U}_{X_0}(A, B) = 0$, we have $\Sigma_A\,\Sigma_B\, \mu^N(X_0\cup A)\mathfrak{U}_{X_0}(A, B) = 0$ and

$$\sum_A\sum_B \mu^N(X_0\cup A)\log\left(\frac{\mu^N(X_0\cup A)}{P(X_0\cup A)}\right)\mathfrak{U}_{X_0}(A, B) = 0. \tag{2.28}$$

Using these observations and (2.26), we see, after some simplification, that (2.27) is equal to

$$\begin{aligned}(2.29)\quad &\sum_{X_0}\left[\sum_B\sum_A F\left(\frac{P(X_0\cup B)}{\mu^N(X_0\cup B)}\frac{\mu^N(X_0\cup A)}{P(X_0\cup A)}\right)\mathfrak{U}_{X_0}(A, B)\frac{P(X_0\cup A)}{P(X_0\cup B)}\mu^N(X_0\cup B)\right]\\ &= \sum_{X_0}\sum_B\sum_{A\neq B} F\left(\frac{P(X_0\cup B)}{\mu^N(X_0\cup B)}\frac{\mu^N(X_0\cup A)}{P(X_0\cup A)}\right)\mathfrak{U}_{X_0}(A, B)\frac{P(X_0\cup A)}{P(X_0\cup B)}\mu^N(X_0\cup B).\end{aligned}$$

We may delete the terms where $A = B$ because $F(1) = 0$. Now by substituting the formula for $\mathfrak{U}_{X_0}(A, B)$ into the right side of (2.29), the lemma is proved.

REMARK 2.2. Since $F \leqq 0$, and for $z \neq 1$, $F(z) < 0$, it is clear from Lemma 2.2 that if the expression in (2.23) is equal to zero and if A and B are two subsets of $[-N + L, N - L]$ with the same number of elements, then for all X_0 contained in $[-N, -N + L - 1]\cup[N - L + 1, N]$,

$$\frac{\mu^N(X_0\cup A)}{P(X_0\cup A)} = \frac{\mu^N(X_0\cup B)}{P(X_0\cup B)}. \tag{2.30}$$

Indeed, there is a finite sequence $A = B_0, B_1, \cdots, B_n = B$ such that $B_{i+1} = B_i\cup b_i\backslash a_i$ for some $a_i\in B_i$, $b_i\notin B_i$ with $|a_i - b_i| = 1$; and it is immediate that if the expression in (2.23) is equal to zero, then

$$\frac{\mu^N(X_0\cup B_i)}{P(X_0\cup B_i)} = \frac{\mu^N(X_0\cup B_{i+1})}{P(X_0\cup B_{i+1})}. \tag{2.31}$$

3. Shift invariant states

If $X \subseteqq Z$, we will set $X + a = \{x + a\,|\,x\in X\}$. A state μ is shift invariant if

$$\mu^\Lambda(X) = \mu^{\Lambda+a}(X + a) \tag{3.1}$$

for all finite subsets $\Lambda\subset Z$, all $X\subseteqq\Lambda$, and all $a\in Z$.

Let $\mathscr{M}$ be the space of all states on E and give $\mathscr{M}$ the weak topology. Denote by $\mathscr{M}_1$ the closed subspace of all shift invariant states.

We will need the following facts about shift invariant states.

PROPOSITION 3.1. *If μ_0 is shift invariant, then μ_t is also shift invariant for all $t \geqq 0$.*

PROPOSITION 3.2. *In the definition of $A(\mu)$ (see (2.7)), the limit supremum is actually the limit. Thus, if μ is shift invariant,*

$$A(\mu) = \lim_{N\to\infty} \frac{1}{\beta} \frac{1}{2N+1} \left[\sum_{X \subseteqq [-N,N]} \mu^N(X) \log \frac{\mu^N(X)}{P(X)} \right]. \tag{3.2}$$

A proof of (1) based on the techniques in [1] is routine and is left to the reader. A proof of (2) can be found in [2], Section 7.2.

If m is large enough so that $2^m - 1 \geqq L$, let

$$H_m(\mu) = \sum_X \sum_{(a,b)} F\left(\frac{P(X)}{\mu^{2^m-1}(X)} \frac{\mu^{2^m-1}(X \cup b\backslash a)}{P(X \cup b\backslash a)} \right) \cdot W(X \cup b\backslash a, b) \frac{P(X \cup b\backslash a)}{P(X)} \mu^{2^m-1}(X), \tag{3.3}$$

where $X \subseteqq [-2^m + 1, 2^m - 1]$ and $(a, b) \in D_1(2^m - 1, X)$. Using the convention in the Remark 2.1, it is easily seen that $H_m(\cdot)$ is an upper semicontinuous function on $\mathscr{M}$.

We will also need the following fact, which can be easily proved from the results in [1].

PROPOSITION 3.3. *The map $(\mu_0, t) \to \mu_t$ is continuous in the product topology on $\mathscr{M} \times [0, \infty)$.*

LEMMA 3.1. *Let $\mu \in \mathscr{M}_1$. Then $H_m(\mu) \leqq 2H_{m-1}(\mu)$.*

PROOF. Let

$$A_1(X) = \{(a, b) \in D_1(2^m - 1, X) \mid a \leqq -L - 1 \quad \text{and} \quad b \leqq -L - 1\} \tag{3.4}$$

and

$$A_2(X) = \{(a, b) \in D_1(2^m - 1, X) \mid a \geqq L + 1 \quad \text{and} \quad b \geqq L + 1\}. \tag{3.5}$$

Then, since $F \leqq 0$, $H_m(\mu)$ is less than or equal to the sum of the two expressions obtained if D_1 in (3.3) is replaced by A_1 and subsequntly by A_2.

Now each $X \subseteqq [-2^m + 1, 2^m - 1]$ can be written as $X = V_1 \cup V_2$, where $V_1 \subseteqq [-2^m + 1, -1]$ and $V_2 \subseteqq [0, 2^m - 1]$. We then notice that $A_1(V_1 \cup V_2)$ depends only on V_1, and hence may be written $A_1(V_1)$. In fact,

$$A_1(V_1) = D_1(2^{m-1} - 1, V_1 + 2^{m-1}) - (2^{m-1}, 2^{m-1}). \tag{3.6}$$

Similarly, if $(a, b) \in A_1(V_1)$, then $W(V_1 \cup V_2, a)$ depends only on V_1, and in fact,

$$W(V_1 \cup V_2, a) = W(V_1 + 2^{m-1}, a + 2^{m-1}). \tag{3.7}$$

And finally, if $(a, b) \in A_1(V_1)$, then it is easy to check that

$$\frac{P(V_1 \cup V_2 \cup b\backslash a)}{P(V_1 \cup V_2)} = \frac{P(V_1 \cup b\backslash a)}{P(V_1)}. \tag{3.8}$$

To simplify the notation, we will write $D_1(V_1)$ instead of $D_1(2^{m-1}, V_1 + 2^{m-1}) - (2^{m-1}, 2^{m-1})$, $W(V_1, a)$ instead of $W(V_1 + 2^{m-1}, a + 2^{m-1})$, and $\mu'(X)$ instead of $\mu^{2^m-1}(X)$.

$$(3.9)\quad \sum_X \sum_{(a,b)\in A_1(X)} F\left(\frac{P(X)}{\mu'(X)}\frac{\mu'(X\cup b\backslash a)}{P(X\cup b\backslash a)}\right) W(X\cup b\backslash a, b)\frac{P(X\cup b\backslash a)}{P(X)}\mu'(X)$$

$$= \sum_{V_1}\sum_{V_2}\sum_{(a,b)\in D_1(V_1)} F\left(\frac{P(V_1)}{\mu'(V_1\cup V_2)}\frac{\mu'(V_1\cup V_2\cup b\backslash a)}{P(V_1\cup b\backslash a)}\right) W(V_1\cup b\backslash a, b)$$

$$\cdot \frac{P(V_1\cup b\backslash a)}{P(V_1)}\frac{\mu'(V_1\cup V_2)}{\mu^{[-2^m+1,-1]}(V_1)}\mu^{[-2^m+1,-1]}(V_1)$$

$$\leqq \sum_{V_1}\sum_{(a,b)\in D_1(V_1)} F\left(\frac{P(V_1)}{\mu^{[-2^m+1,-1]}(V_1)}\frac{\mu^{[-2^m+1,-1]}(V_1\cup b\backslash a)}{P(V_1\cup b\backslash a)}\right)$$

$$\cdot W(V_1\cup b\backslash a, b)\frac{P(V_1\cup b\backslash a)}{P(V_1)}\mu^{[-2^m+1,-1]}(V_1).$$

The last inequality follows from Jensen's inequality, since F is concave. In the above argument, we have assumed that all $\mu^{[-2^m+1,-1]}(V_1) > 0$. If this is not the case, then the inequality in (3.9) follows directly from the convention given in Remark 2.1.

Since μ and P are both shift invariant, the right side of (3.9) is equal to $H_{m-1}(\mu)$. Similarly, if D_1 in (3.3) is replaced by A_2, the result is less than or equal to $H_{m-1}(\mu)$, and the lemma is proved.

LEMMA 3.2. *The function*

$$(3.10)\qquad H(\mu) = \frac{1}{\beta}\lim_{m\to\infty}\frac{1}{2^{m+1}-1}H_m(\mu)$$

exists on $\mathcal{M}_1$ (it is possibly minus infinity) and is upper semicontinuous there. Moreover, if $\mu_0 \in \mathcal{M}_1$, then $A(\mu_t) - A(\mu_0) \leqq \int_0^t H(\mu_s)\,ds$.

PROOF. Let $G(m) = \Pi_{j=m}^{\infty}\left[(2^{j+2}-2)/(2^{j+2}-1)\right]$. Then by Lemma 3.1, if $\mu \in \mathcal{M}_1$,

$$(3.11)\quad G(m+1)\frac{1}{2^{m+2}-1}H_{m+1}(\mu) \leqq G(m+1)\frac{2^{m+2}-2}{2^{m+2}-1}\frac{1}{2^{m+1}-1}H_m(\mu)$$

$$= G(m)\frac{1}{2^{m+1}-1}H_m(\mu).$$

Therefore, $G(m)(1/2^{m+1}-1)\,H_m(\mu)$ is a decreasing sequence of upper semicontinuous functions on $\mathcal{M}_1$. Hence, the limit exists and is upper semicontinuous on $\mathcal{M}_1$. Since $G(m)$ goes to one as m goes to infinity, this limit when divided by β is equal to $H(\mu)$.

In the proof of Theorem 2.1, we showed that

$$(3.12)\qquad \sum_{X\subseteqq[-N,N]}\mu_t^N(X)\log\frac{\mu_t^N(X)}{P(X)} - K_1 t$$

is a nonincreasing function of t. Therefore an application of Lebesgue's theorem and Fatou's lemma yields

$$\sum_{X \subseteq [-N,N]} \mu_t^N(X) \log \frac{\mu_t^N(X)}{P(X)} - \sum_{X \subseteq [-N,N]} \mu_0^N(X) \log \frac{\mu_0^N(X)}{P(X)} - K_1 t \leqq \int_0^t \left[\frac{d}{ds} \left(\sum_{X \subseteq [-N,N]} \mu_s^N(X) \log \frac{\mu_s^N(X)}{P(X)} \right) - K_1 \right] ds. \tag{3.13}$$

Now if $\mu_0 \in \mathscr{M}_1$, we may use (3.1) and (3.2) together with this inequality to get

$$\begin{aligned} A(\mu_t) - A(\mu_0) &\leqq \lim_{m \to \infty} \frac{1}{\beta(2^{m+1} - 1)} \int_0^t \frac{d}{ds} \left[\sum_X \mu_s^{2^m - 1}(X) \log \frac{\mu_s^{2^m-1}(X)}{P(X)} \right] ds \\ &\leqq \lim_{m \to \infty} \frac{1}{\beta(2^{m+1} - 1)} \int_0^t [H_m(\mu_s) + K_1] \, ds. \end{aligned} \tag{3.14}$$

In the middle expression, the summation extends to all $X \subseteq [-2^m + 1, 2^m - 1]$. The last inequality follows exactly as in the proof of Theorem 2.1.

Now since $0 < G(m) < 1$ and $H_m(\mu) \leqq 0$, $H_m(\mu) \leqq G(m)H_m(\mu)$. Therefore, using this inequality and monotone convergence, we have

$$\begin{aligned} A(\mu_t) - A(\mu_0) &\leqq \lim_{m \to \infty} \int_0^t \frac{1}{\beta(2^{m+1} - 1)} [G(m)H_m(\mu_s) + K_1] \, ds \\ &= \int_0^t H(\mu_s) \, ds, \end{aligned} \tag{3.15}$$

and the proof is complete.

LEMMA 3.3. *Let $\mu \in \mathscr{M}_1$. Then if $\mu \notin C$, $H(\mu) < 0$.*

PROOF. We may think of $\mu^{2^m - 1}$ as a measure on $\cup_Y \cup_n S(2^m - 1, n, Y)$ (here $S(N, n, Y)$ is as in Section 2.2), and it is easily seen that μ is the weak limit of the μ^{2^m-1}. From Remark 2.2, it is clear that if $H_m(\mu) = 0$, then μ^{2^m-1} is equal to one of the ν_{2^m-1} used to describe the set C. Thus, if $H(\mu) = 0$, then all $H_m(\mu)$ are zero, and therefore $\mu \in C$.

THEOREM 3.1. *Let $\mu_0 \in \mathscr{M}_1$ and suppose that $t_n \to \infty$ and that μ_{t_n} converges weakly to μ. Then $\mu \in C$.*

PROOF. Since $\mathscr{M}_1$ is weakly closed and each $\mu_{t_n} \in \mathscr{M}_1$, $\mu \in \mathscr{M}_1$. Suppose that $\mu \notin C$. Then by Lemma 3.3, $H(\mu) < 0$. Therefore, there is a $\delta > 0$ such that if

$$G_\mu = \{\nu \in \mathscr{M}_1 \mid H(\nu) < -\delta\}, \tag{3.16}$$

then $\mu \in G_\mu$. Since H is upper semicontinuous, G_μ is open. Therefore, there is an open subset $\hat{G}_\mu$ of $\mathscr{M}_1 \times [0, \infty)$ containing $(\mu, 0)$ and such that if $(\nu_0, s) \in \hat{G}_\mu$ then $\nu_s \in G_\mu$ (see (3.4)). Since $(\mu, 0) \in \hat{G}_\mu$, there is an open set $\bar{G}_\mu \subset \mathscr{M}_1$ and an $\varepsilon > 0$ such that $\mu \in \bar{G}_\mu$ and $\bar{G}_\mu \times [0, \varepsilon) \subseteq \hat{G}_\mu$. Thus, if $\nu_0 \in \bar{G}_\mu$ and $0 \leqq s < \varepsilon$, then $H(\nu_s) < -\delta$.

Since μ_{t_n} converges weakly to μ, $\mu_{t_n} \in \bar{G}_\mu$ for all sufficiently large n. Thus, by Lemma 3.2, for all sufficiently large n, $A(\mu_{t_n+\varepsilon}) - A(\mu_{t_n}) \leqq -\delta\varepsilon$.

This together with Theorem 2.1 implies that $\lim_{t\to\infty} A(\mu_t) = -\infty$. But it is easily seen from the definition of $A(\nu)$ that $\inf_{\nu\in\mathscr{M}} A(\nu) > -\infty$. This is a contradiction and completes the proof.

COROLLARY 3.1. *Let G be a weakly open subset of $\mathscr{M}$ containing C, and let $\mu_0 \in \mathscr{M}_1$. Then for all sufficiently large t, $\mu_t \in G$.*

PROOF. This follows immediately from the compactness of $\mathscr{M}$ and Theorem 3.1.

COROLLARY 3.2. *All shift invariant equilibrium states are elements of C.*

REMARK 3.1. It is clear from the proof of Lemma 3.1 that if the state at time t is shift invariant, but not an equilibrium state, then the Helmholtz free energy at all future times is strictly less than it is at time t.

4. Pressure

We first modify the Markov process η_t introduced in Section 1 in order to motivate our definition of pressure. We need a "wall" from which the particles "rebound," so we let Z^- be the negative integers and take $D = \{0, 1\}^{Z^-}$ as the state space. The intuitive description is the same as before except that no particle is allowed to jump from minus one to zero. A proof that such a process exists can be given by imitating the one in [1]. The attempted jumps from minus one to zero are to be thought of as collisions with the wall, and we will take the pressure of a state to be twice the expected number of collisions with the wall per unit time. Thus, if μ is a state, its pressure is given by the formula

$$p(\mu) = \int \eta(-1) \exp\left\{\sum_{k=-\infty}^{-1} U(-1, k)\eta(k)\right\} \mu(d\eta). \tag{4.1}$$

Pressure is usually defined only for equilibrium states, so perhaps this should be thought of as instantaneous pressure.

The temperature will no longer play a role in this section; thus, we have absorbed the β into the potential U. However, it will be important to display the chemical potential explicitly and so we will require $U(x, x) = 0$ for all x.

Now let γ be any real number and set $S_N = \{\eta \in D \mid \eta(X) = 0 \text{ if } x < -N\}$. We define a probability measure $\nu_{N,\gamma}$ with support S_N by the formula

$$\nu_{N,\gamma}(A) = \sum_{\eta\in A\cap S_N} \frac{1}{\theta(N, \gamma)} \exp\left\{\gamma \sum_{x=-N}^{-1} \eta(x) - \tfrac{1}{2}\sum_x \sum_y \eta(x)\eta(y)U(x, y)\right\}. \tag{4.2}$$

Here $\theta(N, \gamma)$ is the normalizing constant. It can be proved (see [3], footnote 7) that there is a probability measure ν_γ on D which is the weak limit of the $\nu_{N,\gamma}$. Just as in the case where the state space is E instead of D, it can be shown that the ν_γ are equilibrium states for the Markov process η_t. Then ν_γ are moreover the only states for which the pressure is usually defined.

The usual definition of pressure is given in terms of the normalizing constants $\theta(N, \gamma)$, and we will make use of the following result (see [2], Section 5.6).

LEMMA 4.1. *There are constants $\lambda(\gamma) > 1$ and $c(\gamma) > 0$ such that*

$$\lim_{N\to\infty} \lambda^{-N}(\gamma)\theta(N, \gamma) = c(\gamma). \tag{4.3}$$

The usual definition of pressure associated with the state ν_γ is taken to be $P(\gamma) = \log \lambda(\gamma)$ (see [2], Section 3.4).

THEOREM 4.1. *The functions $P(\gamma)$ and $p(\nu_\gamma)$ are both strictly increasing in γ.*

PROOF. A proof that $P(\gamma)$ is a strictly increasing analytic function of γ can be found in [2], Section 5.6; therefore, we will restrict our attention to $p(\nu_\gamma)$.

Since $\eta(-1)\exp\{\Sigma_{x=-\infty}^{-1} U(-1, x)\eta(x)\}$ is a continuous function on D, we have

$$\begin{aligned}
&\int \eta(-1)\exp\left\{\sum_{x=-\infty}^{-1} U(-1, x)\eta(x)\right\}\nu_\gamma(d\eta) \\
&\quad = \lim_{N\to\infty}\int \eta(-1)\exp\left\{\sum_{x=-\infty}^{-1} U(-1, x)\eta(x)\right\}\nu_{N,\gamma}(d\eta) \\
&\quad = \lim_{N\to\infty}\frac{1}{\theta(N, \gamma)}\sum_{\eta\in S'_N}\left[\exp\left\{\sum_{x=-\infty}^{-1} U(-1, x)\eta(x)\right\}\right. \\
&\qquad\qquad \left.\exp\left\{\gamma\sum_{x=-N}^{-1}\eta(x) - \tfrac{1}{2}\sum_x\sum_y \eta(x)\eta(y)U(x, y)\right\}\right].
\end{aligned} \tag{4.4}$$

Here we have set $S'_N = \{\eta \in S_N \mid \eta(-1) = 1\}$. Some elementary manipulations reduce this to

$$\lim_{N\to\infty}\frac{1}{\theta(N, \gamma)}\exp\{\gamma\}\sum_{\eta\in S''_N}\exp\left\{\gamma\sum_{x=-N}^{-2}\eta(x) - \tfrac{1}{2}\sum_x\sum_y\eta(x)\eta(y)U(x, y)\right\}. \tag{4.5}$$

In our last expression, $S''_N = S_N \backslash S'_N$. Now

$$\sum_{\eta\in S''_N}\exp\left\{\gamma\sum_{x=-N}^{-2}\eta(x) - \tfrac{1}{2}\sum_x\sum_y\eta(x)\eta(y)U(x, y)\right\} = \theta(N - 1, \gamma), \tag{4.6}$$

and thus

$$\begin{aligned}
p(\nu_\gamma) &= \lim_{N\to\infty}\exp\{\gamma\}\frac{1}{\theta(N, \gamma)}\theta(N - 1, \gamma) \\
&= \exp\{\gamma\}\frac{1}{\lambda(\gamma)} = \exp\{\gamma - P(\gamma)\}.
\end{aligned} \tag{4.7}$$

We have now reduced the problem to proving that $\gamma - P(\gamma)$ is a strictly increasing function of γ. To do this we clearly need more information about P, and this can be found in [2], Section 3.4. The crucial fact is that there is a function $f(\rho)$ defined on the interval $[0, 1)$ such that

$$P(\gamma) = \sup_{0\le\rho<1}(\rho\gamma - f(\rho)). \tag{4.8}$$

Therefore, $\gamma - P(\gamma) = \inf_{0 \leqq \rho < 1} [(1 - \rho)\gamma + f(\rho)]$, which is clearly a nondecreasing function of γ. As mentioned in the first line of the proof, $P(\gamma)$ is analytic; and thus if $\gamma - P(\gamma)$ is not strictly increasing, then it is constant. But this cannot be since $P(\gamma) > 0$ for all γ. Hence, $p(\nu_\gamma)$ is also a strictly increasing function of γ.

We conclude this section by giving a physical interpretation of Theorem 4.1. The parameter γ determines what the density $\rho(\gamma)$ of the state ν_γ will be; and in the case which we are considering, it is known that $\rho(\gamma)$ is a continuous strictly increasing function of γ. Thus, Theorem 4.1 tells us that at constant temperature the pressure is an increasing function of the density. Moreover, the pressure is a strictly increasing function of the density, and this is interpreted to mean that there is no change of phase for the model with which we are dealing.

The pressure $p(\nu)$ is defined for states other than the ν_γ, and Theorem 4.1 says nothing about how the pressure varies with the density for the other states. However, if ν is not an equilibrium state, one would not expect any nice relationship between the pressure and the density. Because of the results in Section 3, we feel that it is highly unlikely that there are any equilibrium states besides the ν_γ and convex combinations of the ν_γ. We are unfortunately unable to prove this.

REFERENCES

[1] R. HOLLEY, "A class of interactions in an infinite particle system," *Advances in Math.*, Vol. 5 (1970), pp. 291–309.

[2] R. RUELLE, *Statistical Mechanics*, New York, Benjamin, 1969.

[3] ———, "Statistical mechanics of a one-dimensional lattice gas," *Comm. Math. Phys.*, Vol. 9 (1968), pp. 267–278.

[4] F. SPITZER, "Interaction of Markov processes," *Advances in Math.*, Vol. 5 (1970), pp. 246–290.