Potentials for almost Markovian Random Fields

W. G. Sullivan

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia

Abstract. A positive almost Markovian random field is a probability measure on a lattice gas whose finite set conditional probabilities are continuous and positive. We show that each such random field has a potential and in the translation invariant case an absolutely convergent potential. We give a criterion for determining which random fields correspond to pair potentials, or in general n-body potentials. We show that two translation invariant positive almost Markovian random fields have the same finite set conditional probabilities if and only if one minimizes the specific free energy of the other.

1. Introduction

There are two natural approaches to the statistical properties of a lattice gas. The interaction can be described either in terms of conditional probabilities or in terms of potentials. For nearest neighbor pair interactions Averintsev [1] and Spitzer [11] showed these two approaches to be equivalent. Then Averintsev [2], Sherman [10], and Sullivan [12] extended this analysis to finite range interactions.

The type of interaction potential considered by Gallavotti *et al.* [5, 7] suggests an extension of the finite range results. Such potentials give rise to measures whose conditional probabilities are limits of conditional probabilities on finite subsets of the lattice. In the compact totally disconnected configuration space this amounts to the requirement of continuity for certain conditional probabilities. Interestingly, the resulting class of probability measures is slightly larger than the class of probability measures strictly associated with the interaction potentials mentioned above.

Nevertheless, we can find potentials with weaker convergence properties and the method of finding these potentials provides a criterion for determining whether a given probability measure corresponds to a potential involving n or fewer particles at a time.

With the additional assumption of translation invariance we can get better convergence for the potential and can prove the existence of specific free energy. We give a proof based on conditional probability of a specific free energy variational theorem.

2. Conditional Probabilities

We consider the lattice gas based on the v-dimensional integer lattice Z^v, each point of which to be occupied by one of the w particles of the set $W = \{0, 1, \ldots, w - 1\}$. The fundamental space $\Omega = W^{Z^v}$ is considered a measurable space with respect to the σ-field of sets generated as the product σ-field of the power set σ-fields of the factor spaces W. More generally, for $\Lambda \subset Z^v$ we say a subset of Ω is Λ measurable provided that it is contained in the σ-field generated as the product σ-field of the power set σ-fields of the factor spaces W in W^Λ and the indiscrete σ-fields $\{\emptyset, W\}$ of the factor spaces in W^{Λ^c}.

The symbol $\mathscr{L}$ denotes the set of finite subsets of Z^v. For $\Lambda \in \mathscr{L}$ we denote the number of elements in Λ by $N(\Lambda)$. A finite cylinder set is a finite union of sets of the form

$$\{\omega \in \Omega : \omega = x \text{ on } \Lambda\} \tag{2.1}$$

for $x \in \Omega$, $\Lambda \in \mathscr{L}$. Otherwise stated, a finite cylinder set is a subset of Ω which is Λ measurable for some $\Lambda \in \mathscr{L}$. A probability measure on Ω is determined by its values on the finite cylinder sets.

We assign the discrete topology to W and the Tychonoff topology to Ω, so that Ω is a totally disconnected compact Hausdorff space. The set of continuous real valued functions on Ω, $C(\Omega)$ is a Banach space in the sup norm, whose dual is the set of bounded signed measures on Ω. A finite cylinder function is a function on Ω which is Λ measurable for some $\Lambda \in \mathscr{L}$. Each finite cylinder function is continuous, and each continuous function is the uniform limit of a sequence of finite cylinder functions. If $f(x)$ is Λ measurable, then $f(x)$ depends only on the values of x on Λ. We employ without comment the natural correspondence between functions on W^Λ and Λ measurable functions on Ω.

Let μ be a probability measure on Ω. The simplest conditional probabilities are those of the form

$$\mu(\omega = x \text{ on } \Lambda \,|\, \omega = y \text{ on } \Lambda') = \mu(\omega = x \text{ on } \Lambda, y \text{ on } \Lambda') / \mu(\omega = y \text{ on } \Lambda') \tag{2.2}$$

for $x, y \in \Omega$, $\Lambda, \Lambda' \in \mathscr{L}$, $\Lambda \cap \Lambda' = \emptyset$, the expression being defined when the denominator is positive. More generally, one has the *finite set conditional probability*

$$\mu(\omega = x \text{ on } \Lambda \,|\, \omega = y \text{ on } \Lambda^c) \tag{2.3}$$

for $x, y \in \Omega$, $\Lambda \in \mathscr{L}$. For fixed Λ and x this is a Λ^c measurable function defined a.e. in y which satisfies the following:

$$\int_A \mu(\omega = x \text{ on } \Lambda \,|\, \omega = y \text{ on } \Lambda^c)\, \mu(dy) = \mu(\omega = x \text{ on } \Lambda, \omega \in A). \tag{2.4}$$

Here $x \in \Omega$ and A is any A^c measurable set. The Radon-Nikodym theorem assures the existence of finite set conditional probabilities (see Loève [8]).

Definition. An almost Markovian random field is a probability measure on Ω all of whose finite set conditional probabilities of the form (2.3) are continuous functions of y for x and A fixed. When these conditional probabilities are all positive, the random field is called *positive*.

When μ is a positive almost Markovian random field, we deduce from (2.4) and continuity that whenever $A, A' \in \mathscr{L}$, $A \cap A' = \emptyset$, $x, y \in \Omega$

$$\min_{z \in \Omega, z = y \text{ on } A'} \mu(\omega = x \text{ on } A \,|\, \omega = z \text{ on } A^c)$$

$$\leqq \mu(\omega = x \text{ on } A \,|\, \omega = y \text{ on } A') \tag{2.5}$$

$$\leqq \max_{z \in \Omega, z = y \text{ on } A'} \mu(\omega = x \text{ on } A \,|\, \omega = z \text{ on } A^c).$$

When we choose $A' = \emptyset$, expression (2.5) gives bounds for the values of μ on finite cylinder sets. In particular, μ is positive on nonempty finite cylinder sets.

For $\Psi \subset Z^\nu$ we use the notation $\mathscr{L}$-lim $A \to \Psi$ to indicate the limit taken on the net of increasing finite subsets of Ψ ordered by inclusion. From (2.5) and continuity we have

Lemma 1. *Let μ be a positive almost Markovian random field. Then $\mathscr{L}$-lim $A' \to A^c \mu(\omega = x$ on $A \,|\, \omega = y$ on $A') = \mu(\omega = x$ on $A \,|\, \omega = y$ on $A^c)$ for all $x, y \in \Omega$, $A \in \mathscr{L}$.*

It is convenient to introduce some special notation. For $x \in \Omega, j \in Z^\nu$, $A \in \mathscr{L}$ we use the left superscript

$$^j x, \,^A x$$

to denote the elements of Ω which are identically zero on the superscript set and equal to x otherwise, i.e. $[^A x]_k = 0$ if $k \in A$, $= x_k$ otherwise. We omit brackets from single point sets.

We shall consider only positive almost Markovian random fields. In this case it is more convenient to work with logarithms of ratios of conditional probabilities than with conditional probabilities themselves.

Definition. The *conditional energy field* $E_y^\mu(A, x)$ of the positive almost Markovian random field μ is defined with $x, y \in \Omega$, $A \in \mathscr{L}$

$$E_y^\mu(A, x) = -\log[\mu(\omega = x \text{ on } A \,|\, \omega = y \text{ on } A^c)/$$
$$\mu(\omega = 0 \text{ on } A \,|\, \omega = y \text{ on } A^c)]. \tag{2.6}$$

Proposition 1. *Let E^μ be the conditional energy field of the positive almost Markovian random field μ. Let $\Lambda', \Lambda'' \in \mathscr{L}$, $\Lambda' \cap \Lambda'' = \emptyset$ and set $\Lambda = \Lambda' \cup \Lambda''$. Then with $x, y \in \Omega$*

$$E_y^\mu(\Lambda, x) = E_z^\mu(\Lambda', x) + E_z^\mu(\Lambda'', x) \text{ with } z = x \text{ on } \Lambda', \ y \text{ on } \Lambda^c, \ 0 \text{ on } \Lambda'', \quad (2.7)$$

$$E_y^\mu(\Lambda, {}^{\Lambda'}x) = E_z^\mu(\Lambda'', x) \quad \text{with} \quad z = {}^{\Lambda'}y, \quad (2.8)$$

$$E_y^\mu(\Lambda, x) = \sum_{j \in \Lambda} E_{z(j)}^\mu(j, x) \quad \text{with} \quad [z(j)]_k = y_k \quad \text{for} \quad k \in \Lambda^c$$
$$= 0 \quad \text{for} \quad k \in \Lambda, k > j \quad (2.9)$$
$$= x_k \quad \text{for} \quad k \in \Lambda, k \leq j.$$

$<$ *is any linear order on Λ.*

Proof. For (2.7) and (2.8) apply Lemma 1 to the corresponding relations for conditional probabilities of the form (2.2). Note that (2.9) follows from (2.7) and (2.8).

We remark that (2.9) indicates that it is sufficient to know the one point conditional probabilities to obtain the finite set conditional probabilities.

3. Potentials

Definition. A *potential* is a real valued function $U(\Lambda, x)$ on $\mathscr{L} \times \Omega$ such that for each fixed $\Lambda \in \mathscr{L}$, $U(\Lambda, x)$ depends only on the values of x on Λ. A potential U such that

$$\Lambda' \in \mathscr{L}, \sum_{\Lambda' \cap \Lambda \neq \emptyset} |U(\Lambda', x) - U(\Lambda', {}^\Lambda x)| \quad (3.1)$$

converges uniformly in $x \in \Omega$ for each $\Lambda \in \mathscr{L}$ is called an *absolutely convergent potential*. A potential U such that

$$\mathscr{L} - \lim_{\Psi \to Z^\nu} \sum_{\Psi \supset \Lambda', \Lambda' \cap \Lambda \neq \emptyset} [U(\Lambda', x) - U(\Lambda', {}^\Lambda x)] \quad (3.2)$$

converges uniformly in $x \in \Omega$ for each $\Lambda \in \mathscr{L}$ is called an *$\mathscr{L}$-convergent potential*, and the limit is defined to be the *conditional energy field* $E^U(\Lambda, x)$ of U.

Clearly an absolutely convergent potential is $\mathscr{L}$-convergent. By elementary manipulations we have the following

Proposition 2. *A potential U is absolutely convergent ($\mathscr{L}$-convergent) if and only if*

$$\sum_{\Lambda \in \mathscr{L}, \Lambda \ni j} [U(\Lambda, x) - U(\Lambda, {}^j x)] \quad (3.3)$$

is absolutely convergent ($\mathcal{L}$-convergent) uniformly in $x \in \Omega$ for each $j \in Z^v$, in which case with $\Lambda \in \mathcal{L}$, $\Lambda = \Lambda' \cup \Lambda''$, $\Lambda' \cap \Lambda'' = \emptyset$

$$E^U(\Lambda, x) = E^U(\Lambda', {}^{\Lambda''}x) + E^U(\Lambda'', x), \tag{3.4}$$

$$E^U(\Lambda, {}^{\Lambda'}x) = E^U(\Lambda'', {}^{\Lambda'}x), \tag{3.5}$$

$$E^U(\Lambda, x) = \sum_{j \in \Lambda} E^U(j, x(j)), \quad [x(j)]_k = x_k \quad \text{for} \quad k \in \Lambda^c$$

$$= 0 \quad \text{for} \quad k \in \Lambda, \ k > j \tag{3.6}$$

$$= x_k \quad \text{for} \quad k \in \Lambda, \ k \leq j.$$

$<$ is any linear order on Λ.

Proposition 3. Let U be an $\mathcal{L}$-convergent potential. Then there exists a positive almost Markovian random field μ such that for each $\Lambda \in \mathcal{L}$ and $x \in \Omega$, $E^U(\Lambda, x) = E^\mu_x(\Lambda, x)$.

Proof. For $x, y \in \Omega$, $\Lambda \in \mathcal{L}$ we define a family of conditional probabilities

$$q_{\Lambda y}(x) = \exp(-E^U(\Lambda, z))/Z_{\Lambda y}, \quad z = x \quad \text{on} \quad \Lambda, y \quad \text{on} \quad \Lambda^c$$

$$Z_{\Lambda y} = \sum_{z \in \Omega, z = y \text{ on } \Lambda^c} \exp - E^U(\Lambda, z). \tag{3.7}$$

From $\mathcal{L}$-convergence we deduce that $q_{\Lambda y}(x)$ depends continuously on y for x and Λ fixed. Proposition 2 implies that the family of conditional probabilities $q_{\Lambda y}(x)$ is consistent in the sense of formula (1.6) of Dobrushin [3]. The proposition follows from Dobrushin's Theorem 1.

We shall now do the converse and get a potential for a given positive almost Markovian random field. There are many potentials which give rise to the same conditional energy field. The procedure of the proof of Theorem 1 below selects a certain one of these which has some rather nice properties.

Theorem 1. Every positive almost Markovian random field μ has an $\mathcal{L}$-convergent potential U such that $E^\mu_x(\Lambda, x) = E^U(\Lambda, x)$ for all $x \in \Omega$, $\Lambda \in \mathcal{L}$.

Before the proof we introduce some more notation.

$$\Delta^j E_y(\Lambda, x) = E_y(\Lambda, x) - E_y(\Lambda, {}^jx), \quad j \in Z^v,$$

$$\Delta^\Lambda = \prod_{j \in \Lambda} \Delta^j. \tag{3.8}$$

For the theorem only we define the operator ∇^j which can operate only on terms of the form $E_y(\Lambda, {}^jx)$ with

$$\nabla^j E_y(\Lambda, {}^jx) = E_y(\Lambda, x) - E_y(\Lambda, {}^jx). \tag{3.9}$$

Proof of Theorem 1. Define $U(\Lambda, x) = \Delta^{\Lambda} E_0^{\mu}(\Lambda, x)$. Now for any $\Psi \in \mathscr{L}, \Psi \supset \Lambda$ we have

$$\sum_{\Lambda' \subset \Psi, \Lambda' \cap \Lambda \neq \emptyset} [U(\Lambda', x) - U(\Lambda', {}^{\Lambda}x)] = E_0^{\mu}(\Psi, x) - E_0^{\mu}(\Psi, {}^{\Lambda}x). \quad (3.10)$$

To see this consider

$$\left[\prod_{j \in \Psi} (1 + \nabla^j) \right] E_0^{\mu}(\Psi, {}^{\Psi}x) = E_0^{\mu}(\Psi, x) \quad (3.11)$$

and

$$\left[\prod_{j \in \Psi \setminus \Lambda} (1 + \nabla^j) \right] E_0^{\mu}(\Psi, {}^{\Psi}x) = E_0^{\mu}(\Psi, {}^{\Lambda}x). \quad (3.12)$$

The terms in the left hand side of (3.11) which are not in the left hand side of (3.12) are equal to corresponding terms in the left hand side of (3.10). The right hand side of (3.10) equals $E_y^{\mu}(\Lambda, x)$ with $y = {}^{\Psi^c}x$. Hence the $\mathscr{L}$-limit as $\Psi \rightarrow Z^{\nu}$ is $E_x^{\mu}(\Lambda, x)$.

The potential produced by the proof of Theorem 1 has some very interesting properties. If one considers "interactions" in the sense of Lanford-Ruelle [7], the terms produced by the proof are exactly the terms of the "interaction". One might conjecture that this expansion of the potential should converge absolutely for all positive almost Markovian random fields. This conjecture is false, even in the translation invariant case. In Appendix 1 we give an example of a translation invariant $\mathscr{L}$-convergent potential which is not absolutely convergent. We draw the following conclusions from the proof of Theorem 1.

Corollary 1. *The conditional energy field $E_y^{\mu}(\Lambda, x)$ of the positive almost Markovian random field μ can be expressed in terms of $\mathscr{L}$-convergent potentials involving n or fewer bodies (points) if and only if*

$$\Delta^{\Lambda} E_y^{\mu}(\Lambda, x) = 0 \quad (3.13)$$

for all $\Lambda \in \mathscr{L}$ with $N(\Lambda) > n$ and all $x, y \in \Omega$.

Corollary 2. *Let $L \in \mathscr{L}$ and let μ be a positive almost Markovian random field such that $E_y^{\mu}(j, x)$ depends only on the values of x and y on $j + L$. Then $E_y^{\mu}(\Lambda, x)$ can be expressed by a potential U such that $U(\Lambda, x) = 0$ whenever $\{j - k : j, k \in \Lambda\} \not\subset L$.*

Proof. Let $\Lambda \in \mathscr{L}$, $j, k \in \Lambda$, $j - k \notin L$. It is straightforward to verify $\Delta^{jk} E_0^{\mu}(\Lambda, x) = 0$ so that $\Delta^{\Lambda} E_0^{\mu}(\Lambda, x) = 0$ for all $x \in \Omega$.

Corollary 2 gives an L-Markovian random field theorem similar to that of Averintsev [2]. Very elementary geometry yields the Markov random field theorem of Averintsev [1] and Spitzer [11].

4. Translation Invariance and Absolute Convergence

Most of the considerations thus far have not employed the additive structure of Z^ν and are applicable to a lattice gas over an arbitrary countable set, for which the concept of $\mathcal{L}$-convergence is particularly appropriate.

We show below that a translation invariant positive almost Markovian random field has an absolutely convergent potential. Theorem 1 gives an $\mathcal{L}$-convergent potential, but this need not always converge absolutely (see Appendix 1). We obtain approximating potentials by projecting the conditional energy field from the lattice Z^ν to spaces of the form $(Z_n)^\nu$, with Z_n the cyclic group of order n.

The translation operator is denoted by T_j for $j \in Z^\nu$. The action of T_j on $k \in Z^\nu$ and $x \in \Omega$ is

$$T_j k = k + j, \qquad [T_j x]_k = x_{k-j}. \tag{4.1}$$

This action induces natural actions on measures, potentials, etc.

Definition. A homogeneous conditional energy field is the conditional energy field of some translation invariant positive almost Markovian random field.

Let $E_y(\Lambda, x)$ be a homogeneous conditional energy field. It is convenient to consider the function $E_x(0, x)$ which we denote by $D(x)$. $D(x)$ represents the energy required to change the configuration of the origin from 0 to x_0, given the configuration x on the remainder of the lattice. From (2.9) and translation invariance

$$E_y(\Lambda, x) = \sum_{j \in \Lambda} D(T_{-j} x(j)), \qquad [x(j)]_k = y_k \quad \text{for} \quad k \in \Lambda^c$$

$$= 0 \quad \text{for} \quad k \in \Lambda, \ k > j \tag{4.2}$$

$$= x_k \quad \text{for} \quad k \in \Lambda, \ k \leqq j.$$

$<$ is any linear order on Λ.

The function $D(x)$ is continuous on Ω. The norm on $C(\Omega)$ gives a norm to the homogeneous conditional energy field E,

$$\|E\| = \sup_{x \in \Omega} |E_x(0, x)|. \tag{4.3}$$

Functions of the form $E_x(0, x)$ constitute a closed linear subspace of $C(\Omega)$ which can be characterized by the following

Proposition 4. *Let $D \in C(\Omega)$ be such that $D(^0 x) = 0$ for all $x \in \Omega$. Then D can be expressed in the form $D(x) = E_x^\mu(0, x)$ for some translation invariant positive almost Markovian random field μ if and only if for all $j \in Z^\nu$*

$$D(^j x) + D(T_{-j} x) = D(x) + D(T_{-j}{}^0 x). \tag{4.4}$$

Proof. Property (4.4) for $E_x(0, x)$ follows from (2.9) and translation invariance, which proves necessity. To prove sufficiency select a linear order for Z^v and use (4.2) to define $E_y(\Lambda, x)$. Property (4.4) implies that this $E_y(\Lambda, x)$ does not depend on the particular order chosen. We deduce that (2.7) and (2.8) are satisfied by $E_y(\Lambda, x)$, which is translation invariant. The reasoning of Proposition 3 is then applicable, and Dobrushin's theorem gives the required μ.

We now consider the process of projecting the homogeneous conditional energy field $E_y(\Lambda, x)$ onto a v-dimensional discrete "torus" embedded in Z^v. We define the n-cube C_n in Z^v and its state space Ω_n by

$$C_n = \{0, 1, \ldots, n-1\}^v, \qquad \Omega_n = W^{C_n}. \tag{4.5}$$

The action of translation is considered modulo n in each dimension on C_n, and we write $k \bmod C_n$ for $k \in Z^v$ to denote the projection of Z^v onto C_n. Let $D(x) = E_x(0, x)$ and define for $u \in \Omega_n$

$$
\begin{aligned}
D_n(u) = D(x), \quad x_j &= u_k \quad \text{for} \quad k = j \bmod C_n, \; k \neq 0 \\
&= 0 \quad \text{for} \quad 0 = j \bmod C_n, \; j > 0 \\
&= u_0 \quad \text{for} \quad 0 = j \bmod C_n, \; j \leq 0
\end{aligned}
\tag{4.6}
$$

with $<$ a fixed translation invariant linear order on Z^v.

Lemma 2. *Let $E_y(\Lambda, x)$ be a homogeneous conditional energy field and let D_n be defined by (4.6) with $D(x) = E_x(0, x)$. Then $D_n({}^0u) = 0$ for all $u \in \Omega_n$ and D_n satisfies for $u \in \Omega_n$, $k \in C_n$*

$$D_n({}^ku) + D_n(T_{-k}u) = D_n(u) + D_n(T_{-k}{}^0u). \tag{4.7}$$

Proof. We consider the larger cube C_{mn} and define for $i \in C_n$, $u \in \Omega_n$

$$
\begin{aligned}
D_n^m(i, u) &= [E_0(C_{mn}, x) - E_0(C_{mn}, y)]/m^v \\
x_j &= u_k \qquad \text{for} \quad k = j \bmod C_n \\
y_j &= [{}^iu]_k \quad \text{for} \quad k = j \bmod C_n.
\end{aligned}
\tag{4.8}
$$

By Proposition 1 for $i, k \in C_n$

$$D_n^m(i, {}^ku) + D_n^m(k, u) = D_n^m(i, u) + D_n^m(k, {}^iu). \tag{4.9}$$

Also by Proposition 1

$$
\begin{aligned}
m^v D_n^m(i, u) &= \sum_{j \in C_{mn}, \, i = j \bmod C_n} D(T_{-j}x(j)) \\
[x(j)]_k &= 0 \quad \text{for} \quad k \notin C_{mn} \\
&= u_k \quad \text{for} \quad k \in C_{mn}, \, i \neq k \bmod C_n \quad \text{or} \quad k \leq j \\
&= 0 \quad \text{for} \quad k \in C_{mn}, \, i = k \bmod C_n \quad \text{and} \quad k > j
\end{aligned}
\tag{4.10}
$$

with the linear order as in (4.6). When the integer m is large, the predominant fraction of the m^v terms in the right hand side of (4.10) are close to $D_n(T_{-i}u)$ and the remaining terms are each bounded. The conclusion follows from taking limit $m \to \infty$.

Definition. The *norm* $\|U\|$ of the translation invariant potential U is the finite or infinite value

$$\sup_{x \in \Omega} \sum_{\Lambda \in \mathscr{L}, \Lambda \ni 0} \left| U(\Lambda, x) - U(\Lambda, {}^0x) \right|. \tag{4.11}$$

From Proposition 2 we see that the absolutely convergent translation invariant potentials are just the ones with finite norms.

Theorem 2. *Let μ be a translation invariant positive almost Markovian random field with conditional energy field $E_y^\mu(\Lambda, x)$. Then for any $\varepsilon > 0$ there is a translation invariant absolutely convergent potential U such that for all $x \in \Omega$*

$$\left| E^U(0, x) - E_x^\mu(0, x) \right| < \varepsilon \tag{4.12}$$

and

$$\|U\| \leqq \|E^\mu\|. \tag{4.13}$$

Proof. We give in fact a finite range potential with these properties. Select an integer d such that $\left| E_x^\mu(0, x) - E_y^\mu(0, y) \right| < \varepsilon/2$ whenever x and y agree at all points of Z^v within Euclidean distance d of the origin. Select an integer n so that $\|E^\mu\| \left[1 - (n - 2d)^v / n^v \right] < \varepsilon/4$. On Ω_n we have the function D_n defined by (4.6) with $D(x) = E_x^\mu(0, x)$. By Lemma 2 and the finite lattice version of Proposition 4 there is a positive translation invariant probability measure ϱ on Ω_n such that for $u \in \Omega_n$ and the function f on Ω_n defined by

$$f(u) = - \log \left[\varrho(\omega = u) / \varrho(\omega = 0) \right] \tag{4.14}$$

we have

$$D_n(u) = f(u) - f({}^0u). \tag{4.15}$$

Consider f as a C_n measurable function on Ω and define

$$U(\Lambda, x) = f(T_j x)/n^v \quad \text{when} \quad T_j \Lambda = C_n$$
$$= 0 \quad \text{otherwise}. \tag{4.16}$$

From (4.6), (4.15) and (4.16) we have (4.13), since there are n^v terms in (4.11) for this U. At least $(n - 2d)^v$ of the terms of the sum for $n^v E^U(0, x)$ contain all points within distance d of the origin and so are within $\varepsilon/2$ of $E_x^\mu(0, x)$. The remaining terms are each bounded by $\|E^\mu\|$ which gives (4.12).

By considering the difference energy field $E^\mu - E^U$ and applying the usual 2^{-k} argument we get the following.

Corollary. *Every translation invariant positive almost Markovian random field has an absolutely convergent potential.*

5. Specific Free Energy

In this section we derive some properties of specific energies and specific free energies. Our definitions are based on probability theory. At constant temperature they differ by a constant from their thermodynamic counterparts. We shall take limits in the sense of increasing cubes, though, with some extra effort, the limits could be taken in the sense of Van Hove (see Ruelle [9]).

Lemma 3. *Let μ be a translation invariant positive almost Markovian random field. Then the following limit exists and is finite.*

$$\lim_{n \to \infty} -\log \mu(\omega = 0 \text{ on } C_n)/N(C_n). \tag{5.1}$$

The proof of this lemma for the case of finite range random fields is given in Sullivan [12]. With obvious refinements that proof carries over to this case.

Definition. The *μ-specific energy of the zero state $e_\mu(0)$ is the limit (5.1).*

It is of interest that the function $P(\Phi)$ considered in Gallavotti and Miracle-Sole [4] is exactly $e_\mu(0)$ for the case they consider. This gives another approach to the proof of Lemma 3.

For Lemma 4 we consider Z^ν to be ordered by a fixed translation invariant linear order $<$, e.g. lexicographical order.

Lemma 4. *Let ϱ be a translation invariant probability measure on Ω and let μ be a translation invariant positive almost Markovian random field. Then for $y \in \Omega$*

$$\lim_{n \to \infty} \int E_y^\mu(C_n, x) \varrho(dx)/N(C_n) \tag{5.2}$$

exists and has the value independent of y

$$\int E_{z(x)}^\mu(0, x) \varrho(dx), \quad [z(x)]_k = x_k \quad \text{for} \quad k \leq 0 \\ = 0 \quad \text{for} \quad k > 0. \tag{5.3}$$

Proof. We remark that for fixed n the integral in (5.2) can be written as a finite sum. Express $E_y^\mu(C_n, x)$ by (2.9). For large n the predominant fraction of the terms for the ϱ integral of $E_y^\mu(C_n, x)$ are close to (5.3) and the remaining terms are bounded. The lemma follows by taking the limit as $n \to \infty$.

Definition. The *μ-specific energy $e_\mu(\varrho)$* of the translation invariant probability measure ϱ is defined so that $e_\mu(\varrho) - e_\mu(0)$ equals (5.3).

By similar methods one can show that for all $y, x \in \Omega$

$$\lim_{n \to \infty} [E_y^\mu(C_n, x) - \log\mu(\omega = 0 \text{ on } C_n) + \log\mu(\omega = x \text{ on } C_n)]/N(C_n) = 0 \quad (5.4)$$

so we have the following

Proposition 5. *Let ϱ and μ be as in Lemma 4. Then*

$$e_\mu(\varrho) = \lim_{n \to \infty} \int - \log\mu(\omega = x \text{ on } C_n) \varrho(dx)/N(C_n). \quad (5.5)$$

Definition. Let μ and ϱ be translation invariant positive almost Markovian random fields. The *specific entropy* $s(\varrho)$ of ϱ is $e_\varrho(\varrho)$. The *μ-specific free energy* $f_\mu(\varrho)$ of ϱ is $e_\mu(\varrho) - s(\varrho)$. One can in fact define specific entropy for any translation invariant probability measure on Ω (see Ruelle [9]), and the definition agrees with the above in the almost Markovian case.

We now introduce the concept of *blocking*. In the simplest case we consider Z^ν partitioned into *blocks* isomorphic to C_n. We consider a new state space $W' = \Omega_n$ and principal space $\Omega' = W'^{Z^\nu}$. There is a natural one to one correspondence between elements of Ω and Ω'. A translation invariant probability measure on Ω yields a translation invariant probability measure on Ω'; a translation invariant probability measure on Ω' corresponds to a periodic probability measure on Ω. Let μ and ϱ be translation invariant positive almost Markovian random fields on Ω and let μ' and ϱ' be the corresponding measures on Ω'. Then we have

$$f_\mu(\varrho) = f'_\mu(\varrho')/N(C_n) \quad (5.6)$$

with f' denoting the specific free energy in Ω'. More generally blocking could be carried out with rectangular rather than cubic sets. Many results for translation invariant random fields can be extended to periodic random fields by blocking.

Below we give a variational characterization of translation invariant positive almost Markovian random fields. It is directed toward a larger class of random fields than that of Lanford-Ruelle [7] or Holley [6], but it is weaker than these theorems because it requires that both measures be almost Markovian. The proof employs relatively elementary probability theory, whereas the authors mentioned above use more elaborate considerations.

Theorem 3. *Let μ and ϱ be translation invariant positive almost Markovian random fields. Then $f_\mu(\varrho) \geqq 0$ and $f_\mu(\varrho) = 0$ if and only if $E_y^\mu(\Lambda, x)$ $= E_y^\varrho(\Lambda, x)$ for all $x, y \in \Omega, \Lambda \in \mathscr{L}$.*

Proof. From expression (5.5) we have that

$$f_\mu(\varrho) = \lim_{n \to \infty} \int \log[\varrho(\omega = x \text{ on } C_n)/\mu(\omega = x \text{ on } C_n)] \, \varrho(dx)/N(C_n). \quad (5.7)$$

For fixed n the integral is in fact a finite sum in the form of a free energy expression for a finite probability space, so $f_\mu(\varrho) \geq 0$. Next we number points of Z^ν so that $C_1, C_2, \ldots$ are filled in succession. Rather than write $j_0, j_1, \ldots$, we simply write $0, 1, \ldots$ to indicate those points of Z^ν. Let $x_0, x_1, \ldots$ be the values of $x \in \Omega$ on this enumeration with $x_0, \ldots, x_j$ filling C_n. We have

$$\mu(x_1 \ldots x_j) = \mu(x_1) \mu(x_2 | x_1) \ldots \mu(x_j | x_1 \ldots x_{j-1}) \quad (5.8)$$

so

$$\varrho(x_1 \ldots x_j) \log[\varrho(x_1 \ldots x_j)/\mu(x_1 \ldots x_j)]$$

$$= \varrho(x_1 \ldots x_j) \log[\varrho(x_1)/\mu(x_1)] + \cdots + \varrho(x_1 \ldots x_j) \log[\varrho(x_j | x_1 \ldots x_{j-1})/ \quad (5.9)$$

$$\mu(x_j | x_1 \ldots x_{j-1})].$$

The C_n approximation of the free energy (5.7) is just the sum over Ω_n of the terms of (5.9) divided by $N(C_n)$. These averages of the form

$$\sum_{x \in \Omega_n} \varrho(x_1 \ldots x_{j-1}) \varrho(x_j | x_1 \ldots x_{j-1}) \log[\varrho(x_j | x_1 \ldots x_{j-1})/\mu(x_j | x_1 \ldots x_{j-1})]$$

$$(5.10)$$

are nonnegative and appear in the approximate free energy of each larger cube. Since (5.7) is assumed to be zero, we must have for every $\varepsilon > 0$ an integer m and values $x_1, \ldots, x_{m-1}$ so that

$$\sum_{x_m \in W} \varrho(x_m | x_1 \ldots x_{m-1}) \log[\varrho(x_m | x_1 \ldots x_{m-1})/\mu(x_m | x_1 \ldots x_{m-1})] < \varepsilon. \quad (5.11)$$

Since W is finite and $\mu(x_m | x_1 \ldots x_{m-1})$ is bounded away from zero, this implies that

$$\max_{x_m \in W} |\log \varrho(x_m | x_1 \ldots x_{m-1}) - \log \mu(x_m | x_1 \ldots x_{m-1})| \quad (5.12)$$

can be made arbitrarily small by appropriate choice of m and $x_1, \ldots, x_{m-1}$. However, these are the wrong "boundary points" to make a direct conclusion about the conditional energy fields. Instead we conclude by the above reasoning and appropriate blocking, for any $\varepsilon > 0$ and any $\Lambda \in \mathscr{L}$ there exists a $\Lambda' \in \mathscr{L}$ with $\Lambda' \cap \Lambda = \emptyset$ and a $y \in \Omega$ such that

$$\max_{x \in \Omega} |\log \varrho(\omega = x \text{ on } \Lambda | \omega = y \text{ on } \Lambda') - \log \mu(\omega = x \text{ on } \Lambda | \omega = y \text{ on } \Lambda')| < \varepsilon.$$

$$(5.13)$$

From continuity it follows that $E_x^\mu(0, x) = E_x^\varrho(0, x)$ for all $x \in \Omega$, and the theorem follows from (2.9) and translation invariance.

Appendix 1

We give an example of a translation invariant $\mathscr{L}$-convergent potential which is not absolutely convergent. Let $W = \{0, 1\}$ and $\Omega = W^Z$. Define

$$U(\Lambda, x) = (-1)^n/(n \log(n+1)) \quad \text{if} \quad x \equiv 1 \text{ on } \Lambda$$

and Λ is a nonempty set of n adjacent points.

$$U(\Lambda, x) = 0 \text{ otherwise.}$$

We have

$$E^U(0, 1) = \sum_{n=1}^{\infty} (-1)^n n/(n \log(n+1))$$

since there are n intervals of length n containing 0. This series does not converge absolutely, so U does not converge absolutely.

$E^U(0, x)$ depends on the interval of ones of x about 0 and the position of 0 in this interval. It is straightforward to show that the series for $E^U(0, x)$ converges conditionally uniformly in $x \in \Omega$ in the sense of $\mathscr{L}$-convergence.

We remark that this potential does not converge in the sense of the norm of Lanford-Ruelle [7] nor in the weaker norm of Gallavotti and Miracle-Sole [4].

Appendix 2

To illustrate the first corollary to Theorem 1 we give a simple example of a potential with phase transition which is not an $\mathscr{L}$-convergent pair potential. Let $W = \{0, 1\}$, $\Omega = W^{Z^2}$. Let S be a unit square in Z^2. Define

$$U(\Lambda, x) = -\beta \text{ if } \Lambda \text{ is a translate of } S \text{ and } x$$
is constant on Λ.

$$U(\Lambda, x) = 0 \text{ otherwise.}$$

Since $\Delta^S E^U(S, 1) = -2\beta$, and $\mathscr{L}$-convergent potential giving the same conditional energy field as U must contain terms involving more than three points at a time.

One can show by a modification of Peierl's method of borders that the above potential exhibits a phase transition for sufficiently large positive β.

For the sake of comparison we note that the conditional energy field norm (4.3) and the potential norm (4.11) of this example both equal $4|\beta|$. The proof of Theorem 1 expresses the potential as an "interaction" in the sense of Lanford-Ruelle [7] with terms involving 1, 2, 3 and 4 points, whose norm in their sense is $36|\beta|$.

References

1. Averintsev, M. B.: On a method of describing descrete parameter random fields. Problemy Peradici Informacii **6**, 100—108 (1970)
2. Averintsev, M. B.: The description of Markov random field by Gibbs conditional distributions. Teor. Verojatnost i Primenen **17**, 21—35 (1972)
3. Dobrushin, R. L.: The description of a random field by means of conditional probabilities and conditions of its regularity. Theor. Probability Appl. **13**, 197—224 (1968)
4. Gallavotti, G., Miracle-Sole, S.: Statistical mechanics of lattice systems. Commun. math. Phys. **5**, 317—323 (1967)
5. Gallavotti, G., Miracle-Sole, S.: Correlation functions of a lattice gas. Commun. math. Phys. **7**, 274—288 (1968)
6. Holley, R.: Free energy in a Markovian model of a lattice spin system. Commun. math. Phys. **23**, 87—99 (1971)
7. Lanford, O. E., Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. Commun. math. Phys. **13**, 194—215 (1969)
8. Loève, M.: Probability theory. New York: D. Van Nostrand 1963
9. Ruelle, D.: Statistical mechanics, rigorous results. New York: Benjamin 1969
10. Sherman, S.: Markov random fields and Gibbs random fields. To appear in Israel J. Math
11. Spitzer, F.: Markov random fields and Gibbs ensembles. Amer. Math. Monthly **78**, 142—154 (1971)
12. Sullivan, W. G.: Finite range random fields and energy fields. To appear

W. G. Sullivan
School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332, USA

W. G. Sullivan's current address:
Dublin Institute for Advanced Studies
10 Burlington Road
Dublin 4, Ireland