

1. Gibbs states and Markov random fields

Let Λ be a finite set and let $\mathcal{P}(\Lambda)$ denote the set of subsets of Λ . In this chapter we will look at various classes of probability measures on $\mathcal{P}(\Lambda)$ that might arise in simple physical and biological models. The points of Λ can be interpreted as sites, each of which can be either empty or occupied by a particle (or some other entity); the subset $A \in \mathcal{P}(\Lambda)$ will be regarded as describing the state of the model when the points of A are occupied and the points of $\Lambda - A$ are empty. The elements of $\mathcal{P}(\Lambda)$ will sometimes be called *configurations*. Most physical models (even simple ones) involving configurations would be dynamic in nature; the probability measures that we will look at will describe the distribution of the configurations when the model is in some state of dynamic equilibrium.

The set Λ , representing the sites in the model, can be expected to have some additional structure, for example we might know the distances between the sites, or we might know that certain sites are connected. We will consider structures on Λ of the latter kind, thus we suppose that the points of Λ are the vertices of some finite graph $\mathcal{G} = (\Lambda, e)$, where e is the set of edges of \mathcal{G} . We do not allow \mathcal{G} to have any multiple edges or loops. The following notation and definitions will be used: If $x, y \in \Lambda$ and there is an edge of the graph between x and y then we will say that x and y are *neighbours*. Let $c: \Lambda \times \Lambda \rightarrow \{0, 1\}$ be given by

$$c(x, y) = \begin{cases} 1 & \text{if } x \text{ and } y \text{ are neighbours,} \\ 0 & \text{otherwise.} \end{cases}$$

(Note that $c(x, x) = 0$ for all $x \in \Lambda$, since we did not allow \mathcal{G} to have any loops.) If $A \in \mathcal{P}(\Lambda)$ then we define $\partial A \in \mathcal{P}(\Lambda)$ by

$$\partial A = \{y \in \Lambda - A : c(x, y) = 1 \text{ for some } x \in A\}.$$

Cambridge University Press

978-0-521-09011-7 - Gibbs States on Countable Sets

Christopher J. Preston

Excerpt

[More information](#)

2

Gibbs states on countable sets

∂A is called the *boundary of A* ; if $x \in \Lambda$ then we will write ∂x instead of $\partial\{x\}$ (and in fact we will in general tend to write x instead of $\{x\}$, thus for example if $A \subset \Lambda$ then we will write $A \cup x$ instead of $A \cup \{x\}$). For any $A \in \mathcal{P}(\Lambda)$, $|A|$ will denote the number of points in A ; a subset $B \in \mathcal{P}(\Lambda)$ will be called a *simplex* of the graph if $B \neq \emptyset$ and if given any $x, y \in B$ with $x \neq y$ then $c(x, y) = 1$. We will say that $B \in \mathcal{P}(\Lambda)$ is an *n -simplex* (for $n \geq 0$) if B is a simplex and $|B| = n + 1$. Note that for all $x \in \Lambda$ $\{x\}$ is a simplex of the graph.

We will let $\mathcal{S}(\Lambda)$ denote the set of probability measures on $\mathcal{P}(\Lambda)$ (with $\mathcal{P}(\Lambda)$ given the σ -algebra $\mathcal{P}(\mathcal{P}(\Lambda))$, i.e. the σ -algebra of all subsets of $\mathcal{P}(\Lambda)$). An element $\mu \in \mathcal{S}(\Lambda)$ describes the distribution of configurations when some model is in dynamic equilibrium; because of this we will call $\mathcal{S}(\Lambda)$ the set of *states* on Λ . Since $\mathcal{P}(\Lambda)$ is a finite set we will identify $\mu \in \mathcal{S}(\Lambda)$ with its density; thus we consider μ as a function from $\mathcal{P}(\Lambda)$ to \mathbb{R} (where \mathbb{R} denotes the real numbers) with the properties $\mu(A) \geq 0$ for all $A \in \mathcal{P}(\Lambda)$, and

$$\sum_{A \subset \Lambda} \mu(A) = 1.$$

The first class of states that we consider are the Gibbs states, which arise in models in statistical physics. A function

$$V: \mathcal{P}(\Lambda) \rightarrow \mathbb{R}$$

will be called a *potential* on Λ if $V(\emptyset) = 0$. For any potential V on Λ we define the *Gibbs state with potential V* to be the state π on Λ given by

$$\pi(A) = Z^{-1} \exp V(A) \quad \text{for all } A \in \mathcal{P}(\Lambda),$$

where Z is the correct normalizing constant, i.e.

$$Z = \sum_{B \subset \Lambda} \exp V(B).$$

If V is a potential on Λ then we define $J_V: \mathcal{P}(\Lambda) \rightarrow \mathbb{R}$ by

$$J_V(A) = \sum_{X \subset A} (-1)^{|A-X|} V(X).$$

Cambridge University Press

978-0-521-09011-7 - Gibbs States on Countable Sets

Christopher J. Preston

Excerpt

[More information](#)*Gibbs states and Markov random fields*

3

Then for any $A \in \mathcal{P}(\Lambda)$ we have

$$V(A) = \sum_{B \subset A} J_V(B).$$

(This follows immediately from what was once called the inclusion–exclusion principle and now goes under the name of the Möbius inversion formula; in any case it is a simple matter to check directly that it holds.) We will call J_V the *interaction potential* corresponding to V . Conversely, given a potential Φ let us denote by U_Φ the potential given by

$$U_\Phi(A) = \sum_{B \subset A} \Phi(B).$$

Thus Φ is the interaction potential corresponding to U_Φ .

It should be noted that if μ is any state on Λ with positive density (i.e. $\mu(A) > 0$ for all $A \subset \Lambda$) then μ is the Gibbs state with potential V where V is the potential given by

$$V(A) = \log \left[\frac{\mu(A)}{\mu(\emptyset)} \right] \quad \text{for any } A \subset \Lambda.$$

Thus the class of Gibbs states consists of all states on Λ with positive density, and hence the introduction of potentials should be regarded as merely a convenient way of describing the states with positive density. The potentials that we will be most interested in, in this chapter, are those that have some connection with the graph structure: a potential V will be called a *nearest neighbour potential* if $J_V(A) \neq 0$ only when A is a simplex of the graph.

The second class of states on Λ that we look at are the Markov random fields. If $\mu \in \mathcal{S}(\Lambda)$ then we will say that μ is a *Markov random field* if:

- (i) μ has positive density, i.e. $\mu(A) > 0$ for all $A \subset \Lambda$.
- (ii) Given $x \notin A \subset \Lambda$ then the conditional probability (with respect to μ) that a configuration contains x , given that the configuration is A on $\Lambda - x$, is the same as the conditional probability that the configuration contains x given that the

configuration is $A \cap \partial x$ on ∂x ; i.e.

$$\frac{\mu(A \cup x)}{\mu(A \cup x) + \mu(A)} = \sum_{B \subset \Lambda - (\partial x \cup x)} \mu((A \cap \partial x) \cup x \cup B) / \sum_{B \subset \Lambda - (x \cup x)} [\mu((A \cap \partial x) \cup x \cup B) + \mu((A \cap \partial x) \cup B)].$$

This says that the probability of there being a particle at x , given a particular configuration of particles on $\Lambda - x$, only depends on what happens on the neighbours of x ; thus in some sense particles do not interact unless they occupy neighbouring sites.

From the above description of a Markov random field one might expect a connection between Markov random fields and Gibbs states whose potentials have some connection with the graph structure. In fact the main point of this chapter is to show that Markov random fields and Gibbs states with nearest neighbour potentials are the same.

The above definition of a Markov random field is clearly quite cumbersome to work with, so we will define another class of states, called nearest neighbour states, which have a simpler definition and then show that they are the same as Markov random fields. Let $\mu \in \mathcal{S}(\Lambda)$; we will call μ a *nearest neighbour state* if:

- (i) $\mu(A) > 0$ for all $A \subset \Lambda$.
- (ii) Given $x \notin A \subset \Lambda$ then

$$\frac{\mu(A \cup x)}{\mu(A)} = \frac{\mu((A \cap \partial x) \cup x)}{\mu(A \cap \partial x)}.$$

Note that (ii) is equivalent to

$$\frac{\mu(A \cup x)}{\mu(A \cup x) + \mu(A)} = \frac{\mu((A \cap \partial x) \cup x)}{\mu((A \cap \partial x) \cup x) + \mu(A \cap \partial x)} \quad \text{for all } x \notin A \subset \Lambda,$$

which says that the conditional probability that the configuration contains x , given that the configuration is A on $\Lambda - x$, is the same as the conditional probability that the configuration contains x given that the configuration is $A \cap \partial x$ on $\Lambda - x$.

Gibbs states and Markov random fields 5

PROPOSITION 1.1 If $\mu \in \mathcal{S}(\Lambda)$ then μ is a Markov random field if and only if it is a nearest neighbour state.

Proof Let μ be a Markov random field and let $x \notin A \subset \Lambda$. Then

$$\begin{aligned} \frac{\mu(A \cup x)}{\mu(A \cup x) + \mu(A)} &= \sum_{B \subset \Lambda - (\partial x \cup x)} \mu((A \cap \partial x) \cup x \cup B) \\ &\times \left[\sum_{B \subset \Lambda - (\partial x \cup x)} [\mu((A \cap \partial x) \cup x \cup B) + \mu((A \cap \partial x) \cup B)] \right]^{-1} \\ &= \frac{\mu((A \cap \partial x) \cup x)}{\mu((A \cap \partial x) \cup x) + \mu(A \cap \partial x)}, \end{aligned}$$

since the middle line is unchanged on replacing A by $A \cap \partial x$. Thus μ is a nearest neighbour state. Conversely, suppose that μ is a nearest neighbour state and again let $x \notin A \subset \Lambda$. If

$$B \subset \Lambda - (\partial x \cup x)$$

then

$$\begin{aligned} \frac{\mu((A \cap \partial x) \cup B \cup x)}{\mu((A \cap \partial x) \cup B)} &= \frac{\mu(((A \cap \partial x) \cup B) \cap \partial x) \cup x)}{\mu(((A \cap \partial x) \cup B) \cap \partial x)} \\ &= \frac{\mu((A \cap \partial x) \cup x)}{\mu(A \cap \partial x)} = \frac{\mu(A \cup x)}{\mu(A)}. \end{aligned}$$

Thus

$$\mu(A) \mu((A \cap \partial x) \cup B \cup x) = \mu(A \cup x) \mu((A \cap \partial x) \cup B)$$

and therefore

$$\begin{aligned} \mu(A) \sum_{B \subset \Lambda - (\partial x \cup x)} \mu((A \cap \partial x) \cup B \cup x) \\ = \mu(A \cup x) \sum_{B \subset \Lambda - (\partial x \cup x)} \mu((A \cap \partial x) \cup B). \end{aligned}$$

From this it is easy to see that μ is a Markov random field. \square

We will now show that Markov random fields and Gibbs states with nearest neighbour potentials are the same. We start by proving:

PROPOSITION 1.2 Let π be the Gibbs state with nearest neighbour potential V . Then π is a Markov random field.

Proof By Proposition 1.1 we need only show that π is a

6

Gibbs states on countable sets

nearest neighbour state, and it is clear that π has positive density. If $x \notin A \subset \Lambda$ then we have

$$\frac{\pi(A \cup x)}{\pi(A)} = \exp [V(A \cup x) - V(A)].$$

But

$$\begin{aligned} V(A \cup x) - V(A) &= \sum_{B \subset A \cup x} J_V(B) - \sum_{B \subset A} J_V(B) \\ &= \sum_{B \subset A} J_V(B \cup x) = \sum_{B \subset A} J_V((B \cap \partial x) \cup x) \\ &= V((A \cap \partial x) \cup x) - V(A \cap \partial x). \end{aligned}$$

Therefore
$$\frac{\pi(A \cap x)}{\pi(A)} = \frac{\pi((A \cap \partial x) \cup x)}{\pi(A \cap \partial x)},$$

and thus π is a nearest neighbour state. \square

The proof of the converse of this result is just as easy:

PROPOSITION 1.3 Let μ be a Markov random field. Then there exists a unique nearest neighbour potential V such that μ is the Gibbs state with potential V .

Proof It is clear that V is unique since we are forced to define V by

$$V(A) = \log \left[\frac{\mu(A)}{\mu(\emptyset)} \right] \quad \text{for all } A \in \mathcal{P}(\Lambda).$$

Then we have that μ is the Gibbs state with potential V , so it only remains to prove that V is a nearest neighbour potential. Let $A \in \mathcal{P}(\Lambda)$ with A not a simplex of the graph. Then there exists $x, y \in A$ with $x \neq y$ and $c(x, y) = 0$. Put $B = A - x - y$; then if $X \subset B$ we have

$$\begin{aligned} \frac{\mu(X \cup x \cup y)}{\mu(X \cup x)} &= \frac{\mu(((X \cup x) \cap \partial y) \cup y)}{\mu((X \cup x) \cap \partial y)} \\ &= \frac{\mu((X \cap \partial y) \cup y)}{\mu(X \cap \partial y)} = \frac{\mu(X \cup y)}{\mu(X)}; \end{aligned}$$

thus
$$V(X \cup x \cup y) - V(X \cup x) - V(X \cup y) + V(X) = 0.$$

Gibbs states and Markov random fields 7

Now we have

$$\begin{aligned}
 J_V(\mathcal{A}) &= \sum_{E \subset \mathcal{A}} (-1)^{|\mathcal{A}-E|} V(E) \\
 &= \sum_{X \subset B} [(-1)^{|\mathcal{A}-X \cup x \cup y|} V(X \cup x \cup y) + (-1)^{|\mathcal{A}-X \cup x|} V(X \cup x) \\
 &\quad + (-1)^{|\mathcal{A}-X \cup y|} V(X \cup y) + (-1)^{|\mathcal{A}-X|} V(X)] \\
 &= \sum_{X \subset B} (-1)^{|\mathcal{A}-X|} [V(X \cup x \cup y) - V(X \cup x) - V(X \cup y) + V(X)] \\
 &= 0,
 \end{aligned}$$

and thus V is a nearest neighbour potential. \square

For the sake of reference we now collect together our previous results and call them a theorem.

THEOREM 1.1 The following are equivalent for $\mu \in \mathcal{S}(\Lambda)$:

- (i) μ is a Markov random field;
- (ii) μ is a nearest neighbour state;
- (iii) μ is a Gibbs state with potential V for some nearest neighbour potential V .

A re-examination of the proof of Proposition 1.3 suggests the following characterization of nearest neighbour potentials:

PROPOSITION 1.4 Let V be a potential on Λ . Then V is a nearest neighbour potential if and only if given $x, y \in \Lambda$ with $x \neq y$ and $c(x, y) = 0$ and given $X \subset \Lambda - x - y$ then

$$V(X \cup x \cup y) - V(X \cup x) - V(x \cup y) + V(X) = 0.$$

Proof If the condition holds, then the proof of Proposition 1.3 shows that V is a nearest neighbour potential. Conversely suppose that V is a nearest neighbour potential and let X, x, y be as in the hypothesis. If $B \subset X$ then just as in Proposition 1.3 we have

$$\begin{aligned}
 J_V(B \cup x \cup y) &= \sum_{Y \subset B} (-1)^{|B-Y|} [V(Y \cup x \cup y) - V(Y \cup x) \\
 &\quad - V(Y \cup y) + V(Y)].
 \end{aligned}$$

But $B \cup x \cup y$ is not a simplex of the graph since $c(x, y) = 0$, and hence $J_V(B \cup x \cup y) = 0$. Thus for all $B \subset X$ we have

$$\sum_{Y \subset B} (-1)^{|B-Y|} [V(Y \cup x \cup y) - V(Y \cup y) - V(Y \cup x) + V(Y)] = 0.$$

Therefore

$$\begin{aligned} &V(X \cup x \cup y) - V(X \cup x) - V(X \cup y) + V(X) \\ &= \sum_{B \subset X} \sum_{Y \subset B} (-1)^{|B-Y|} [V(Y \cup x \cup y) - V(Y \cup x) - V(Y \cup y) + V(Y)] \\ &= 0. \square \end{aligned}$$

We will use the (confusing) terminology of the physicists and say that the graph \mathcal{G} is a *cubic lattice* if \mathcal{G} contains no 2-simplexes (and thus contains no n -simplexes for $n \geq 2$). The most common occurrences of cubic lattices are as finite subsets of \mathbb{Z}^ν (where for $\nu \geq 1$, \mathbb{Z}^ν denotes the points of \mathbb{R}^ν which have integer coordinates, considered as a graph by defining two points to be neighbours if the distance between them is exactly 1). Let $\mathcal{G} = (\Lambda, e)$ be a cubic lattice and let V be a nearest neighbour potential on Λ . Then we must have $J_V(A) = 0$ if $|A| \geq 3$, and thus defining $H: \Lambda \times \Lambda \rightarrow \mathbb{R}$ by

$$H(x, y) = \begin{cases} \frac{1}{2} V(\{x, y\}) & \text{if } x \neq y, \\ V(\{x\}) & \text{if } x = y, \end{cases}$$

then it is easy to check that we have

$$V(A) = \sum_{x \in A} \sum_{y \in A} H(x, y) \quad \text{for all } A \in \mathcal{P}(\Lambda).$$

It is also clear that $H(x, y) = 0$ if $x \neq y$ and $c(x, y) = 0$. This shows that for cubic lattices we can write nearest neighbour potentials in the form of nearest neighbour pair potentials as they are usually defined in elementary physics.

NOTES The definition of a Gibbs state (on a finite subset of \mathbb{Z}^ν) goes back to the classical work of Gibbs (1902). Markov random fields (on \mathbb{Z}^ν) were first introduced by Dobrushin (1968a). Spitzer (1971a) and Averintsev (1970) showed that for finite

Gibbs states and Markov random fields

9

subsets of \mathbb{Z}^{ν} the class of Markov random fields and Gibbs states with nearest neighbour potentials were the same (in Spitzer's paper translation invariance was assumed). The equivalence of Markov random fields and Gibbs states with nearest neighbour potentials on general finite graphs was first obtained by Hammersley and Clifford (1971). Their proof was long and involved and a considerably easier proof is given in Preston (1973), an even simpler proof is given in Grimmett (1973), and yet another proof is given in Sherman (1973). The proof given in this book is an adaptation of Grimmett's proof. For a different approach to this material see Suomela (1972).

Cambridge University Press

978-0-521-09011-7 - Gibbs States on Countable Sets

Christopher J. Preston

Excerpt

[More information](#)

2. Interacting particle systems

The aim of this chapter is to investigate some simple dynamic models that have the probability measures considered in the previous chapter as their equilibrium states. We will look at models in which the dynamics is random; in fact we will consider Markov chains. As before $\mathcal{G} = (\Lambda, e)$ will be an arbitrary finite graph, and for convenience we will denote $\mathcal{P}(\Lambda)$ by Γ .

Let $\{P_t\}_{t \geq 0}$ be a semi-group on Γ ; i.e. for each $t \geq 0$ we have $P_t: \Gamma \times \Gamma \rightarrow \mathbb{R}$ with the properties:

- (i) $0 \leq P_t(A, B)$ for all $A, B \in \Gamma, t \geq 0$;
- (ii) $\sum_{B \in \Gamma} P_t(A, B) = 1$ for all $A \in \Gamma, t \geq 0$;
- (iii) $\sum_{X \in \Gamma} P_t(A, X) P_s(X, B) = P_{t+s}(A, B)$ for all $A, B \in \Gamma, s, t \geq 0$;
- (iv) $\lim_{t \rightarrow 0} P_t(A, B) = I(A, B)$ for all $A, B \in \Gamma$, where

$$I(A, B) = \begin{cases} 1 & \text{if } A = B, \\ 0 & \text{otherwise.} \end{cases}$$

The semi-group $\{P_t\}_{t \geq 0}$ is interpreted as describing a model which has the property that if the model has configuration $A \in \Gamma$ at time s then the probability that it will have configuration $B \in \Gamma$ at time $s + t$ is $P_t(A, B)$.

It is a well-known result that there exists a unique function $G: \Gamma \times \Gamma \rightarrow \mathbb{R}$ satisfying:

- (i) $G(A, B) \geq 0$ if $A, B \in \Gamma$ and $A \neq B$;
- (ii) $\sum_{B \in \Gamma} G(A, B) = 0$ for all $A \in \Gamma$;
- (iii) $P_t = \exp(tG)$.

(Here we consider P_t and G as $|\Gamma| \times |\Gamma|$ matrices and define

$$\exp(tG) = \sum_{n=0}^{\infty} \frac{t^n}{n!} G^n,$$