Markov Random Fields and Measures with Nearest Neighbour Gibbs Potential

by

Nishant Chandgotia

B. Math., Indian Statistical Institute, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE STUDIES
(Mathematics)

The University Of British Columbia
(Vancouver)

August 2011

Abstract

This thesis will discuss the relationship between stationary Markov random fields and probability measures with a nearest neighbour Gibbs potential. While the relationship has been well explored when the measures are fully supported, we shall discuss what happens when we weaken this assumption.

Preface

This thesis came out of discussions and collaborations with Brian Marcus, Tom Meyerowitch, Guangyue Han, Ronnie Pavlov and Ori Gurel-Gurevich.

Table of Contents

List of Figures

Acknowledgments

I will like to thank my parents and my family for bringing me up and teaching me how to stand on my feet. I am grateful to Prof. Brian Marcus for helping me understand my abilities and to reach beyond my reach. I would not have learnt so much, if Tom Meyerovitch would not have been so kind with his time and so patient with my faltering steps. It was joy to be a student in Prof. David Brydges' probability classes and to have him examine my thesis. His humour and knowledge has taught me more than what books could ever have. Last but not the least I would like to thank my friends Marc Carnovale, Vasu Tewari and Gourab Ray. If it were not for their nonsensensical banter, I would not have survived so far away from home.

Chapter 1

Introduction

Stationary Markov random fields are probability measures arising naturally in ergodic theory, image processing, statistical physics among many other fields. In general, there is no compact way of representing them. However, if they are fully supported, it known by the Hammersley-Clifford theorem [2] that these measures have a nearest neighbour Gibbs potential. This is one of the avenues of understanding the structure of stationary Markov random fields. There are further generalisations of this result in [7].

We prove that the assumption on the support can be dropped in the Hammersley-Clifford Theorem for stationary Markov random fields with a finite state space on the $\mathbb{Z}$ lattice. However this generalisation fails to extend in higher dimensions.

In Chapter 2, we will discuss the classic Hammersley-Clifford theorem. This chapter follows the treatment in [11]. In Chapter 3, we will introduce some symbolic dynamics. This chapter can be skipped by any one familiar with the field. In Chapter 4, we prove that every stationary Markov random field on the $\mathbb{Z}$ lattice and finite state space is a Markov chain and consequently a measure with a nearest neighbour Gibbs potential. We begin by proving that the support of stationary Markov random fields are nearest neighbour shifts of finite type. Using the specific structure of shifts of finite type and the results in Chapter 2 we prove that the measures are Markov chains. In Chapter 5 we shall discuss how this result fails to extend in higher dimensions. Here we will also discuss the pivot property which gives another "compact" representation of the conditional probabilities of Markov random fields. Chapter 6 shall discuss some conjectures and ideas for further work.

Chapter 2

Hammersley-Clifford Theorem

An *undirected graph* is a tuple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$, the set of *vertices* is a countable set and $\mathcal{E}$, the set of *edges* is a subset of the set of unordered pairs of distinct elements from $\mathcal{V}$. An undirected graph is called *locally finite* if for all $x \in \mathcal{V}$ the set $\{y \mid (x,y) \in \mathcal{E}\}$ is finite. Vertices will be often referred to as sites. We will always assume that the graphs are locally finite.

$\mathcal{R}$ is some set which we will call the set of symbols or the alphabet. We shall focus on the case where $\mathcal{R}$ is finite but most of results in Chapter 2 will hold for a countable alphabet. We will sometimes assume that $\mathcal{R}$ has a special symbol '0'. We will define probability measures on the space $(\mathcal{R}^{\mathcal{V}}, \mathcal{F})$ where $\mathcal{F}$ is the sigma-algebra generated by cylinder sets which are given by

$$[a,A] = \{x \in \mathcal{R}^{\mathcal{V}} \mid x|_A = a\} \text{ where } A \subset \mathcal{V} \text{ is finite and } a \in \mathcal{R}^A$$

Given a set $A \subset \mathcal{V}$, elements of $\mathcal{R}^A$ will be called *configurations* on A. So the probability space is all ways of putting symbols on sites and the events which generate the sigma algebra are all ways in which symbols at certain sites have been previously fixed. The topology on the space is also generated by the cylinder sets. This is the same as the product topology of the discrete topology over $\mathcal{R}$. The closure of sets under this topology is denoted by $\overline{K}$

The *support* of a measure μ is defined as

$$supp(\mu) = \mathcal{R}^{\mathcal{V}} - \bigcup [a,A]$$

where the union is over all cylinder sets with 0 measure. Note that, since every cylinder set is open, the support of any measure is always a closed set.
The *boundary* of a set $A \subset \mathcal{V}$ is given by

$$\partial A = \{x \in \mathcal{V} - A \mid (x, y) \in \mathcal{E} \text{ for some } y \in \mathcal{V}\}$$

In this chapter we will assume that the support of any given measure has a *safe symbol* ‘0’. To be precise, given a measure μ on $\mathcal{R}^{\mathcal{V}}$, $A \subset \mathcal{V}$ and $x \in supp(\mu)$, y defined by

$$y(t) = \begin{cases} x(t) \text{ for } t \text{ in } A \\ 0 \text{ for t in } A^c \end{cases}$$

belongs to $supp(\mu)$.

This means that if at some sites some symbols can be placed with positive probabilities, those symbols can be replaced with ‘0’ and still the probability will be positive.

Definition. *A set $S \subset \mathcal{R}^{\mathcal{V}}$ is called a* topological Markov field *with respect to a graph $\mathcal{G}$ if for all $x, y \in S$ such that $x = y$ on ∂C for some C finite in $\mathcal{V}$, $z \in \mathcal{R}^{\mathcal{V}}$ defined by*

$$z = \begin{cases} x \text{ on } C \cup \partial C \\ y \text{ on} (C \cup \partial C)^c \end{cases}$$

is an element of S

This means that if we have configurations which agree on the boundary of some finite set then we can paste one on the other i.e. outside the boundary we see y and inside it we see x.

Definition. *A* Markov random field *on a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ with alphabet $\mathcal{R}$ is a probability measure μ on $(\mathcal{R}^{\mathcal{V}}, \mathcal{F})$ such that for all $A, B \subset \mathcal{V}$ finite such that $\partial A \subset B \subset A^c$ and $a \in \mathcal{R}^A, b \in \mathcal{R}^B$ satisfying $\mu([b, B]) > 0$*

$$\mu([a, A] \Big| [b, B]) = \mu([a, A] \Big| [b|_{\partial A}, \partial A]).$$

All this is saying is that probability of having something on a finite set conditioned on its complement is same as the probability of having it given its boundary.

Lemma 2.0.1. *Suppose μ is a Markov random field on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with an alphabet $\mathcal{R}$. Then $supp(\mu)$ is a topological Markov field.*

Proof. Let $C \subset \mathcal{V}$ be finite. Consider $x, y \in supp(\mu)$ such that $x = y$ on ∂C. Let $z \in \mathcal{R}^{\mathcal{V}}$ be such that

$$z = \begin{cases} x \text{ on } C \cup \partial C \\ y \text{ on} (C \cup \partial C)^c. \end{cases}$$

Consider a finite set $D \subset \mathcal{V}$ such that $C \cup \partial C \subset D$. We will prove that for any such D, $\mu([z|_D, D]) > 0$. This will prove that $z \in supp(\mu)$. Since the choice of x, y and D is arbitrary, this is sufficient to prove that $supp(\mu)$ is a topological Markov field.

$$\begin{aligned} \mu([z|_D, D]) &= \mu([z|_C, C] \,\big|\, [z|_{D-C}, D-C])\mu([z|_{D-C}, D-C]) \\ &= \mu([z|_C, C] \,\big|\, [z|_{\partial C}, \partial C])\mu([z|_{D-C}, D-C]) \text{ since } \mu \text{ is a Markov} \\ &\quad \text{random field} \\ &= \mu([x|_C, C] \,\big|\, [x|_{\partial C}, \partial C])\mu([y|_{D-C}, D-C]) > 0 \end{aligned}$$

□

To simplify notation we will sometimes use the following:

$$\mu(a \quad b \quad c) = \mu([a, A] \cap [b, B] \cap [c, C])$$

and

$$\mu(a \mid b) = \mu([a, A] \,\Big|\, [b, B])$$

The following lemma gives an equivalent way of defining Markov random fields when $\mathcal{V}$ is finite. Two sets $A, B \subset \mathcal{V}$ are independent if $A \cap B = \partial A \cap B = \partial B \cap A = \phi$ i.e. $A \cap B = \phi$ and no edges connect vertices of A to vertices of B.

Lemma 2.0.2. *Let μ be a probability measure on $\mathcal{R}^{\mathcal{V}}$ where $\mathcal{V}$ is finite. Then μ is a Markov random field on the graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ if and only if for all A, B independent* *and finite,* $a, a' \in \mathcal{R}^A, b, b' \in \mathcal{R}^B, c \in \mathcal{R}^C$ *where* $C = (A \cup B)^c$

$$\mu(a \quad b \quad c)\mu(a' \quad b' \quad c) = \mu(a' \quad b \quad c)\mu(a \quad b' \quad c) \qquad (2.0.1)$$

The proof of this lemma can be found in [11].

Now we shall define measures with nearest neighbour Gibbs potential. Motivated by topological dynamics, we will call an ordered pair $(x, y) \in \mathcal{R}^{\mathcal{V}}$ *homoclinic* if $x = y$ on all but finitely many points in $\mathcal{V}$. Suppose μ is a probability measure on $\mathcal{R}^{\mathcal{V}}$. For a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *an induced subgraph* is defined to be a tuple $H = (\mathcal{V}_1, \mathcal{E}_1)$ where $\mathcal{V}_1 \subset V$ and $\mathcal{E}_1 = \{(a, b) \in \mathcal{E} \mid a, b \in \mathcal{V}_1\}$. *Cliques* of a graph $\mathcal{G}$ are complete graphs which occur as induced subgraphs of the graph $\mathcal{G}$. Define $\mathcal{T}$ to be the set of all configuration on cliques with positive measure i.e.

$$\mathcal{T} = \{c \in \mathcal{R}^C \mid \text{C is a clique and } \mu([c, C]) > 0\}.$$

A *nearest neighbour Gibbs potential* is a function $V : \mathcal{T} \longrightarrow \mathbb{R}$. A measure μ has a nearest neighbour Gibbs potential V if μ is a probability measure such that for any homoclinic pair $(x, y) \in supp(\mu)$ and $T, \Lambda \subset \mathbb{Z}^2$ finite such that $x = y$ on Λ^c and $\Lambda \cup \partial\Lambda \subset T$

$$\frac{\mu([x|_T, T])}{\mu([y|_T, T])} = \prod_{C-cliques} e^{V(x|_C) - V(y|_C)}$$

Lemma 2.0.3. *Let* μ *be a probability measure on* $\mathcal{R}^{\mathcal{V}}$*. The following are equivalent :*

1. μ *has a nearest neighbour Gibbs potential* V

2. *For every finite* $A, B \subset \mathcal{V}$ *satisfying* $\partial A \subset B \subset A^c$ *and* $x \in supp(\mu)$

$$\mu([x|_A, A] \Big| [x|_B, B]) = \frac{\prod_{C-cliques \subset A \cup \partial A} e^{V(x|_C)}}{Z_{A, x|_B}}$$

where $Z_{A, x|_B}$ *is a normalising factor depending on A and* $x|_B$*.*

The proof is left to the reader. We will now state the Hammersley-Clifford Theorem.

Theorem 2.0.4. *Suppose* μ *is a probability measure on* $\mathcal{R}^{\mathcal{V}}$ *such that* $supp(\mu)$ *has a safe symbol '0'. Then,* μ *has a nearest neighbour Gibbs potential if and only if* μ *is a Markov random field.*

Proof. We will first prove that if μ has a nearest neighbour Gibbs potential then it is a Markov random field. This part does not require the assumption on the support. Let $A, B \subset \mathcal{V}$ finite such that $\partial A \subset B \subset A^c$. Consider $b \in \mathcal{R}^B$ such that $\mu([b,B]) > 0$.

Let $a \in \mathcal{R}^A$. We will first prove that

$$\mu([a,A] \,\Big|\, [b,B]) \leq \mu([a,A] \,\Big|\, [b|_{\partial B}, \partial B]) \text{ for all } a \in \mathcal{R}^A.$$

Since $[a,A] \cap [b,B] \subset [a,A] \cap [b|_{\partial B}, \partial B]$

$$\mu([a,A] \,\Big|\, [b|_{\partial B}, \partial B]) = 0 \Longrightarrow \mu([a,A] \,\Big|\, [b,B]) = 0. \tag{2.0.2}$$

Let us assume that $\mu([a,A] \,\Big|\, [b,B]) > 0$. Consider

$$x \in supp(\mu) \bigcap ([a,A] \cap [b,B]).$$

Then,

$$\begin{aligned}
\mu([a,A] \,\Big|\, [b,B]) &= \mu([x|_A, A] \,\Big|\, [x|_B, B]) \\
&= \frac{\prod_{C-cliques \subset A \cup \partial A} e^{V(x|_C)}}{Z_{A, x|_B}} \\
&\leq \frac{\prod_{C-cliques \subset A \cup \partial A} e^{V(x|_C)}}{Z_{A, x|_{\partial B}}} \text{ by equation 2.0.2} \\
&= \mu([x|_A, A] \,\Big|\, [x|_{\partial B}, \partial B]) \\
&= \mu([a,A] \,\Big|\, [b|_{\partial B}, \partial B]).
\end{aligned}$$

Therefore,

$$\mu([a,A] \,\Big|\, [b,B]) \leq \mu([a,A] \,\Big|\, [b|_{\partial B}, \partial B]) \text{ for all } a \in \mathcal{R}^A.$$

But,

$$1 = \sum_{a \in \mathcal{R}^A} \mu([a,A] \,\Big|\, [b,B]) \leq \sum_{a \in \mathcal{R}^A} \mu([a,A] \,\Big|\, [b|_{\partial B}, \partial B]) = 1.$$

Hence,

$$\mu([a,A] \,\Big|\, [b,B]) = \mu([a,A] \,\Big|\, [b|_{\partial B}, \partial B]) \text{ for all } a \in \mathcal{R}^A.$$

Thus μ is a Markov random field.

Now assume that μ is a Markov random field such that $supp(\mu)$ has a safe symbol '0'. We want to prove that μ has a nearest neighbour Gibbs potential. We will first prove this for the case when $\mathcal{V}$ is finite and then generalise this for countably infinite $\mathcal{V}$.

Suppose μ is a Markov random field. We will translate the question of whether or not there exists a Gibbs potential into a linear algebra problem. Consider a matrix $\mathbf{A}$ whose rows are indexed by elements of $supp(\mu)$ and the columns are indexed by elements of $\mathcal{T}$ where the entries are

$$\mathbf{A}_{a,c} = \begin{cases} 1 \text{ if } a|_C = c \\ 0 \text{ otherwise} \end{cases}$$

where $a \in supp(\mu)$ and $c \in \mathcal{R}^C$ for some clique C. Take the column vector $\mathbf{b}$ indexed by $supp(\mu)$ such that

$$\mathbf{b}_a = log(\mu(a))$$

Then finding a nearest neighbour Gibbs potential with normalising constant $Z = 1$ is equivalent to solving the equation $\mathbf{A}x = \mathbf{b}$. Note that the matrix might have countably infinite rows and columns. However since the graph is finite, every row has finitely many non-zero entries. It is proved in Chapter 7 that

$$\begin{aligned}
\mathbf{A}x = \mathbf{b} \text{ has a solution} \iff & \left(v\mathbf{A} = 0 \Longrightarrow v\mathbf{b} = 0\right) \text{ for all } v \text{ with finitely many} \\
& \text{non-zero entries} \\
\iff & \left(\sum_{i=1}^{k} n_{a_i}\mathbf{A}_{a_i,c} = 0 \text{ for all } c \in \mathcal{T} \Longrightarrow \sum_{i=1}^{k} n_{a_i}\mathbf{b}_{a_i} = 0\right)
\end{aligned} \tag{2.0.3}$$

Consider A, B independent and finite, $a, a' \in \mathcal{R}^A, b, b' \in \mathcal{R}^B, d \in \mathcal{R}^D$ where $D = (A \cup B)^c$. Let,

$$\begin{aligned}
\{x\} &= [a,A] \cap [b,B] \cap [d,D] \\
\{y\} &= [a',A] \cap [b',B] \cap [d,D] \\
\{z\} &= [a',A] \cap [b,B] \cap [d,D] \\
\{w\} &= [a,A] \cap [b',B] \cap [d,D]
\end{aligned}$$

Then by Equation (2.0.1)

$$\mu(\{x\})\mu(\{y\}) = \mu(\{z\})\mu(\{w\}).$$

it follows that

$$\begin{aligned}
\mathbf{A}_{x,c} &= \mathbf{A}_{z,c} + \mathbf{A}_{w,c} - \mathbf{A}_{y,c} \text{ for all } c \in \mathcal{T} \\
\mathbf{b}_x &= \mathbf{b}_z + \mathbf{b}_w - \mathbf{b}_y
\end{aligned} \tag{2.0.4}$$

Thus the Markov random field conditions correspond to Equations (2.0.3) where $k = 4$. Thus the heart of the question whether or not a Markov random field has a nearest neighbour Gibbs potential is whether the left null space of $\mathbf{A}$ is generated by vectors corresponding to the Markov random field conditions.

Now we will prove the following 'break-up' lemma. It says that any configuration in the support can be broken into configurations with non-zero symbols exactly on a clique. This is exactly where we need the support to have a safe symbol.

Lemma 2.0.5. *Suppose μ is a Markov random field such that $supp(\mu)$ has a safe* *symbol '0'. Then for any* $b \in supp(\mu)$ *there exists* $b_1, b_2 \dots b_k \in supp(\mu)$ *such that for each i,* $b_i = 0$ *exactly on* C_i^c *for some clique* C_i *and*

$$\begin{aligned} \mathbf{A}_{b,c} &= \sum_{i=1}^{k} r_i \mathbf{A}_{b_i,c} \textit{ for all } c \in \mathcal{T} \\ \mathbf{b}_b &= \sum_{i=1}^{k} r_i \mathbf{b}_{b_i} \end{aligned}$$

Proof. We will induct on the number of non-zero sites. If b has only 1 non-zero site, then we are done and there is nothing to prove. Suppose the result is true for the number of non-zero sites $\leq k$. Now assume that b has $k+1$ non zero sites. If the $k+1$ sites form a clique we are done. Otherwise, there exist independent sets $A, B \subset \mathcal{V}$ such that b is non zero on $A \cup B$. Let $C = (A \cup B)^c$. Then,

$$\{b\} = [b|_A, A] \cap [b|_B, B] \cap [b|_C, C]$$

Let x, y and $z \in supp(\mu)$ be given by

$$\begin{aligned} \{x\} &= [b|_A, A] \cap [0, B] \cap [b|_C, C] \\ \{y\} &= [0, A] \cap [b|_B, B] \cap [b|_C, C] \\ \{z\} &= [0, A] \cap [0, B] \cap [b|_C, C] \end{aligned}$$

where $[0, D] = \{x \in \mathcal{R}^{\mathcal{V}}$ such that $x(i) = 0$ for all $i \in D\}$ for all $D \subset \mathcal{V}$. By equations (2.0.4)

$$\begin{aligned} \mathbf{A}_{b,c} &= \mathbf{A}_{x,c} + \mathbf{A}_{y,c} - \mathbf{A}_{z,c} \text{ for all } c \in \mathcal{T} \\ \mathbf{b}_b &= \mathbf{b}_x + \mathbf{b}_y - \mathbf{b}_z \end{aligned}$$

and x, y and z have smaller number of non-zero sites than b. By induction we are done. □

Continuing the proof of Theorem (2.0.4), we can now assume that all the configurations are non-zero exactly on cliques. If not we can break up our configurations by using Lemma (2.0.5). The final component of the proof is the following observation.

Suppose a_i's are distinct configurations non-zero exactly on cliques C_i. Then

$$\sum_{i=1}^{k} n_{a_i}\mathbf{A}_{a_i,c} = 0 \text{ for all } c \in \mathcal{T} \Longrightarrow n_{a_i} = 0 \text{ for all i} \tag{2.0.5}$$

This can be proved by induction on k. It is certainly true when $k = 1$. Suppose it is true for $k < r$ and

$$\sum_{i=1}^{r} n_{a_i}\mathbf{A}_{a_i,c} = 0 \text{ for all } c \in \mathcal{T}$$

By renumbering we can assume that among C_i's, C_1 is a maximal subgraph. Let $c_1 = a_1|_{C_1}$.Then $\mathbf{A}_{a_i,c_1} = 1$ only for $i = 1$. Hence $n_{a_1} = 0$. Thus by induction we are done.

Therefore if $\sum_{i=1}^{r} n_{a_i}\mathbf{A}_{a_i,c} = 0$ for all $c \in \mathcal{T}$, then by Lemma (2.0.5) we can assume that the configurations are non-zero exactly on a clique. Then the observation above proves that n_{a_i}' s are 0 for all i. Hence the implication as given in Equation (2.0.3) holds. Therefore, the Markov random field has a nearest neighbour Gibbs potential.

Now we will prove the general case where the graph is infinite. The idea is to reduce it to the case where the graph is finite.
Let $0^{\mathcal{V}} \in supp(\mu)$ be the point such that $0^{\mathcal{V}}(v) = 0$ for all $v \in \mathcal{V}$. Let

$$H = \{a \in supp(\mu) \mid (a, 0^{\mathcal{V}}) \text{ is a homoclinic pair } \}$$

Let $\mathbf{A}$ be a matrix such that rows are indexed by pairs in H and the columns by elements of $\mathcal{T}$. H and $\mathcal{T}$ are countable and any pair in H is homoclinic. Suppose $c \in \mathcal{T}$ is a configuration on C and $a, b \in H$. The matrix entries are given by

$$\mathbf{A}_{(a,b),c} = \begin{cases} 1 & \text{if } a|_C = c, b|_C \neq c \\ -1 & \text{if } a|_C \neq c, b|_C = c \\ 0 & \text{otherwise} \end{cases}$$

Note that since a and b are homoclinic and the graph is locally finite, every row has at most finitely many non-zero entries. Let $\mathbf{b}$ be a column vector indexed by pairs in H and the entries are given by

$$\mathbf{b}_{(a,b)} = \log\left(\frac{\mu([a,\Lambda])}{\mu([b,\Lambda])}\right)$$

where $a \neq b$ exactly on T and $T \cup \partial T \subset \Lambda$. The existence of a solution to the equation $\mathbf{A}x = \mathbf{b}$ is equivalent to the measure having a nearest neighbour Gibbs potential. This is true because of the following. Take any homoclinic pair (a,b) in $supp(\mu)$ such that they differ exactly on the set T. Then $a', b' \in supp(\mu)$ defined by

$$a' = \begin{cases} a \text{ on } T \cup \partial T \\ 0 \text{ on } (T \cup \partial T)^c \end{cases}$$

$$b' = \begin{cases} b \text{ on } T \cup \partial T \\ 0 \text{ on } (T \cup \partial T)^c \end{cases}$$

satisfy

$$\log\left(\frac{\mu([a,\Lambda])}{\mu([b,\Lambda])}\right) = \log\left(\frac{\mu([a',\Lambda])}{\mu([b\prime,\Lambda])}\right)$$

Therefore the existence of a nearest neighbour Gibbs potential for pairs in H gives us existence of potentials for homoclinic pairs in general. A solution to $\mathbf{A}x = \mathbf{b}$ exists if and only if

$$v\mathbf{A} = 0 \Longrightarrow v\mathbf{b} = 0$$

where v has finitely many non-zero entries. The proof is given in the Chapter 7 at the end. Note that $\mathbf{A}$ will now have countably many rows and countably many columns. Restating it, it becomes

$$\sum_{i=1}^{k} n_{b_i,c_i}\mathbf{A}_{(b_i,c_i),c} = 0 \text{ for all } c \in \mathcal{T} \Longrightarrow \sum_{i=1}^{k} n_{b_i,c_i}\mathbf{b}_{(b_i,c_i)} = 0$$

Assume $\{(b_i,c_i)\}_{i=1}^{k}$ is a set of homoclinic pairs satisfying

$$\sum_{i=1}^{k} n_{b_i,c_i}\mathbf{A}_{(b_i,c_i),c} = 0 \text{ for all } c \in \mathcal{T}$$

Since the pairs are homoclinic, each pair agrees outside some finite $T \subset \mathcal{V}$. Let $T \cup \partial T = \Lambda$. Then, $\mathbf{A}_{(b_i,c_i),c} = 0$ for $c \in \mathcal{R}^C$ where $C \subset \mathcal{V} - \Lambda$ is a clique. Define a probability measure $\tilde{\mu}$ on $\mathcal{R}^{\Lambda \cup \partial \Lambda}$ where the graph structure is induced by $\mathcal{G}$ and the sigma-field is the power set. The measure is given by

$$\tilde{\mu}(a) = \mu([a|_\Lambda, \Lambda] \mid 0|_{\partial\Lambda}, \partial\Lambda)$$

Then $\tilde{\mu}$ is a Markov random field on the finite graph $\Lambda \cup \partial\Lambda$ satisfying the safe symbol requirement about its support.

Let $\tilde{\mathbf{A}}$ be its corresponding matrix whose rows are indexed by elements of $supp(\tilde{\mu})$ and columns are indexed by the set of configurations on cliques $\tilde{\mathcal{T}}$. The entries are given by

$$\begin{aligned} \tilde{\mathbf{A}}_{a,c} &= 1 \text{ if } a|_C = c \\ &= 0 \text{ otherwise} \end{aligned}$$

for $a \in supp(\tilde{\mu})$, $c \in \mathcal{R}^C$ and C is a clique in $\Lambda \cup \partial\Lambda$. Take a column vector $\tilde{\mathbf{b}}$ indexed by $supp(\mu)$ such that

$$\tilde{\mathbf{b}}_a = \log \tilde{\mu}(a)$$

Let, $\tilde{b}_i, \tilde{c}_i \in \mathcal{R}^{\Lambda \cup \partial\Lambda}$ such that

$$\begin{aligned} \tilde{b}_i &= b_i|_\Lambda \\ \tilde{c}_i &= c_i|_\Lambda \end{aligned}$$

Now,

$$\sum_{i=1}^{k} n_{b_i,c_i} \mathbf{A}_{(b_i,c_i),c} = 0 \text{ for all } c \in \mathcal{T}$$

Therefore,

$$\sum_{i=1}^{k} n_{b_i,c_i} \tilde{\mathbf{A}}_{\tilde{b}_i,c} - \sum_{i=1}^{k} n_{b_i,c_i} \tilde{\mathbf{A}}_{\tilde{c}_i,c} = 0 \text{ for all } c \in \tilde{\mathcal{T}}$$

By Lemma (2.0.5), we can assume that the configurations are non- zero exactly on cliques. By Equation (2.0.5) we have

$$\sum_{i=1}^{k} n_{b_i,c_i}\tilde{\mathbf{b}}_{\tilde{b}_i} - \sum_{i=1}^{k} n_{b_i,c_i}\tilde{\mathbf{b}}_{\tilde{c}_i} = 0$$

Now

$$\mathbf{b}_{b_i,c_i} = \tilde{\mathbf{b}}_{\tilde{b}_i} - \tilde{\mathbf{b}}_{\tilde{c}_i}$$

Hence,

$$\sum_{i=1}^{k} n_{b_i,c_i}\mathbf{b}_{(b_i,c_i)} = 0$$

Therefore, the proof is complete. □

To summarise, we began by looking at the case where the graph was finite. We broke our configurations up till the configurations were non- zero exactly on cliques. In the case where the graph was infinite, we replaced configurations by configurations which are '0' outside a finite set. This reduced the problem into the case where the graph was finite. We used the fact that the support of the measure had a safe symbol. In Chapter 5 we will see that the theorem fails otherwise. One deficiency in the proof is that it does not give a clear direction as to how to find the potential. For this one can refer to [11] where the proof uses Möbius inversion. The reason that we did not go with the seemingly easier proof is that we believe that this setting is the correct one for generalising the result beyond the safe symbol case. Related work on finite graphs can be found in [1].

Now we will consider stationary Markov random fields.

Let G be a subgroup of the automorphism group of the graph $\mathcal{G}$. Then G acts naturally on the configurations on the graph by

$$(ga)_v = a_{gv} \text{ for all } g \in G \text{ and } v \in \mathcal{V}$$

Thus it also acts on the borel measurable sets

$$gA = \{ga \mid a \in A\}$$

where A is a measurable set and on borel measures on $\mathcal{R}^{\mathcal{V}}$.

$$(g\mu)(A) = \mu(gA)$$

A measure μ is called stationary under the action of the group G if $g\mu = \mu$ for $g \in G$. A nearest neighbour Gibbs potential is called stationary under the action of a group G if $V(c) = V(gc)$. We will require the following result in the next chapter.

Theorem 2.0.6. *Let μ be a Markov random field on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with alphabet $\mathcal{R}$ that is stationary with respect to a group G and $supp(\mu)$ has a safe symbol '0'. Then the measure has a nearest neighbour Gibbs potential stationary with respect to G.*

Proof. We will consider a special class of potentials for the measure and then prove that it is unique and invariant under the action of G.

Lemma 2.0.7. *Suppose μ is a Markov random field on a graph $\mathcal{G}$ such that $supp(\mu)$ has a safe symbol '0'. Then there exists a unique nearest neighbour Gibbs potential $V : \mathcal{T} \longrightarrow \mathbb{R}$ such that $V(c) = 0$ whenever c is a configuration on a clique C such that $c(i) = 0$ for some $i \in C$*

All this is saying is that the potential is 0 when any of the symbols on the concerned clique is '0'. We will prove this lemma and the theorem for the case when $\mathcal{V}$ is finite. The case where $\mathcal{V}$ is infinite requires similar adjustments to the proof as in Theorem 2.0.4

We will begin by making some changes to the matrices **A** and **b**.

Let $\mathcal{T}_1 \subset \mathcal{T}$ be defined as

$$\mathcal{T}_1 = \{c \in \mathcal{R}^C \mid C \text{ is a clique and } c(i) = 0 \text{ for some } i \in C\}$$

Consider a matrix **A** whose rows are indexed by elements of $supp(\mu) \cup \mathcal{T}_1$ and columns are indexed by elements of $\mathcal{T}$. If $a \in supp(\mu)$ and $c \in \mathcal{R}^C$ for some clique C then

$$\mathbf{A}_{a,c} = \begin{cases} 1 \text{ if } a|_C = c, a \in supp(\mu) \\ 0 \text{ if } a|_C \neq c, a \in supp(\mu) \end{cases}$$

If $c \in \mathcal{T}_1$ and $c' \in \mathcal{R}^C$ for some clique C then

$$\begin{aligned}\mathbf{A}_{c,c'} &= 1 \text{ if } c = c' \\ &= 0 \text{ otherwise}\end{aligned}$$

Take the column vector $\mathbf{b}$ indexed by $supp(\mu) \cup \mathcal{T}_1$ defined by

$$\begin{aligned}\mathbf{b}_a &= log(\mu(a)) - log(\mu(0^{\mathcal{V}})) \text{ for } a \in supp(\mu) \\ &= 0 \text{ for } a \in \mathcal{T}_1\end{aligned}$$

Solving the equation $\mathbf{A}x = \mathbf{b}$ is equivalent to finding a nearest neighbour Gibbs potential satisfying the requirements of Lemma 2.0.7 with a normalising constant $\frac{1}{(\mu(0^{\mathcal{V}}))}$. As before,

$$\begin{aligned}\mathbf{A}x = \mathbf{b} \text{ has a solution} \quad &\Longleftrightarrow \quad \left(v\mathbf{A} = 0 \Longrightarrow v\mathbf{b} = 0\right) \text{ for all } v \text{ with finitely many} \\ &\qquad \text{non-zero entries} \\ &\Longleftrightarrow \quad \left(\sum_{i=1}^{k} n_{a_i}\mathbf{A}_{a_i,c} = 0 \text{ for all } c \in \mathcal{T} \Longrightarrow \sum_{i=1}^{k} n_{a_i}\mathbf{b}_{a_i} = 0\right)\end{aligned}$$

Suppose there exist distinct configurations $\{a_i\}_{i=1}^{l} \subset supp(\mu)$, $\{c_i\}_{i=1}^{k} \subset \mathcal{T}_1$ and $\{n_{a_i}\}_{i=1}^{l}$, $\{n_{c_i}\}_{i=1}^{k} \subset \mathbb{R}$ satisfying

$$\sum_{i=1}^{l} n_{a_i}\mathbf{A}_{a_i,c} + \sum_{i=1}^{k} n_{c_i}\mathbf{A}_{c_i,c} = 0 \text{ for all } c \in \mathcal{T}$$

We want to prove that

$$\sum_{i=1}^{l} n_{a_i}\mathbf{b}_{a_i} + \sum_{i=1}^{k} n_{c_i}\mathbf{b}_{c_i} = \sum_{i=1}^{l} n_{a_i}\mathbf{b}_{a_i} = 0 \tag{2.0.6}$$

By the 'break-up' lemma (Lemma(2.0.5)), we can assume that the configurations $\{a_i\}_{i=1}^{l}$ are non-zero exactly on cliques C_i for every i. We will prove by induction on l that

$$\sum_{i=1}^{l} n_{a_i}\mathbf{b}_{a_i} = 0.$$

Suppose $l = 1$. For all $c \in \mathcal{T} - \mathcal{T}_1$

$$n_{a_1}\mathbf{A}_{a_1,c} + \sum_{i=1}^{k} n_{c_i}\mathbf{A}_{c_i,c} = 0$$

This implies

$$n_{a_1}\mathbf{A}_{a_1,c} = 0 \text{ for all } c \in \mathcal{T} - \mathcal{T}_1$$

This implies that either $n_{a_1} = 0$ or $a_1 = 0^{\mathcal{V}}$. In either case, $n_{a_1}\mathbf{b}_{a_1} = 0$. Now assume the result for $l \leq t-1$ and

$$\sum_{i=1}^{t} n_{a_i}\mathbf{A}_{a_i,c} + \sum_{i=1}^{k} n_{c_i}\mathbf{A}_{c_i,c} = 0 \text{ for all } c \in \mathcal{T}$$

By renumbering, we can assume that among C_i's, C_1 is maximal. If C_1 is empty then $C_1, C_2 \dots C_t = \phi$. This implies $a_1, a_2 \dots, a_t = 0^{\mathcal{V}}$. This implies $\mathbf{b}_{a_i} = 0$ for all i. If C_1 is not empty assume that $a_1|_{C_1} = c_0 \in \mathcal{T} - \mathcal{T}_1$. Then $\mathbf{A}_{a_i,c_0}, \mathbf{A}_{c_j,c_0} = 0$ for all $i > 1$ and all j and $\mathbf{A}_{a_1,c_0} = 1$. Therefore $n_{a_1} = 0$. By the induction hypothesis the proof of Equation (2.0.6) is complete.

We will now prove by induction on the number of elements of a clique that the measure completely determines the potential as described in Lemma (2.0.7). Consider a clique C of size 1 and a non-zero configuration c on it. Take a configuration $x \in \mathcal{R}^{\mathcal{V}}$ given by

$$x = \begin{cases} c \text{ on C} \\ 0 \text{ otherwise} \end{cases}$$

Then, $V(c)$ is uniquely determined as $\log(\mu(x)) - \log(\mu(0^{\mathcal{V}}))$. Suppose the potential values of configurations on cliques of size less than r are known. Now take a clique C of size r and a configuration $c \in \mathcal{R}^{C}$ with every symbol being non-zero. Take a configuration $x \in \mathcal{R}^{\mathcal{V}}$ such that

$$x = \begin{cases} c \text{ on C} \\ 0 \text{ otherwise} \end{cases}$$

Then,

$$V(c) = \log(\mu(x)) - \log(\mu(0^{\mathcal{V}})) - \sum_{D \subsetneq C} V(x|_D).$$

So $V(c)$ is completely determined by potential on configurations of size less than r. (Since any induced subgraph of a clique is also a clique.) By induction we have proved the measure completely determines the potential as given in Lemma (2.0.7). Suppose a potential V as in Lemma (2.0.7) is determined by the Markov random field μ. Then consider the potential gV for some $g \in G$. It is a potential for $g\mu$. But since the measure is stationary $g\mu = \mu$. Therefore gV is a potential for μ. Also it satisfies the conditions as stated in Lemma (2.0.7). Since the potential satisfying the conditions is unique, therefore $gV = V$. Hence the potential is unique. $\square$

Chapter 3

Symbolic Dynamics

This chapter shall follow the treatment as given in [10]. Let $\mathcal{R}$ be a finite set called the alphabet. The *full shift* on alphabet $\mathcal{R}$ is the dynamical system $(\mathcal{R}^{\mathbb{Z}}, \sigma)$. The topology on $\mathcal{R}^{\mathbb{Z}}$ is given by the product of the discrete topology on $\mathcal{R}$ and $\sigma : \mathcal{R}^{\mathbb{Z}} \longrightarrow \mathcal{R}^{\mathbb{Z}}$ is given by the map

$$\sigma(x)_i = x_{i+1}$$

where $x_i = x(i)$. In fact $\mathcal{R}^{\mathbb{Z}}$ is a compact metrizable space. The map σ, called the *shift map* is a homeomorphism on the space $\mathcal{R}^{\mathbb{Z}}$.

Let G be the group $\{\sigma^i | i \in \mathbb{Z}\}$. A *shift space* is a dynamical system (X, σ) where $X \subset \mathcal{R}^{\mathbb{Z}}$ is closed and invariant under G. Since X is invariant under σ, it is a homeomorphism on the space X. An equivalent definition is given by *forbidden patterns* which follows.

Let $\mathcal{F} \subset \mathcal{R}^{\star}$ where $\mathcal{R}^{\star}$ is the set of all finite words in $\mathcal{R}$. Consider the space,

$$X_{\mathcal{F}} = \{x \in \mathcal{R}^{\mathbb{Z}} \mid \text{ no subword of x belongs to } \mathcal{F}\}$$

It can be checked that a subset $X \subset \mathcal{R}^{\mathbb{Z}}$ is a shift space if and only if there exists $\mathcal{F} \subset \mathcal{R}^{\star}$ such that $X = X_{\mathcal{F}}$. A shift space X is called a *shift of finite type* if $X = X_{\mathcal{F}}$ for some finite set $\mathcal{F}$. A good example to keep in mind is *the golden mean shift* which is defined as the shift space over the alphabet $\{0,1\}$ given by $X_{\{11\}}$, that is all bi-infinite sequences in $\{0,1\}$ such that no two $1's$ are adjacent.

A *sliding block code* is a continuous map ϕ from one shift space X to another shift space Y which commutes with the shift map, that is $\phi \circ \sigma = \sigma \circ \phi$. A sliding block code is called a *conjugacy* if it is a bijection. This also guarantees that the inverse is also a sliding block code since a bijective continuous map from one compact metric space to another is a homeomorphism. Conjugacy is the notion of "sameness" for shift spaces.

To relate all this to the framework of chapter 2, $\mathbb{Z}$ can be made into a graph $(\mathbb{Z}, \mathcal{E})$ where $\mathcal{E} = \{(n, n+1) | n \in \mathbb{Z}\}$. Every point in a shift space can thus be considered as a configuration on the graph $\mathbb{Z}$ and the alphabet $\mathcal{R}$. We shall cheat a little by calling $\mathbb{Z}$ both the graph and the vertex set.

The *language* of a shift space X denoted by $\mathcal{B}(X)$ is defined as all finite words which occur in elements of X. Define

$$\mathcal{B}_n(X) = \{w \in \mathcal{B}(X) \mid \text{ length of } w \text{ is } n\}$$

A shift space is called *irreducible* if for all $a, b \in \mathcal{B}(X)$ there exists a $w \in \mathcal{B}(X)$ such that $awb \in \mathcal{B}(X)$. (w might possibly be the empty word.) A shift space X is called a *nearest neighbour shift of finite type* if $ab, bc \in \mathcal{B}(X)$ and $b \in \mathcal{R}$ implies $abc \in \mathcal{B}(X)$; that is if the last symbol and the first symbol of two finite words agree we can paste them together. This is a shift of finite type since $X = X_{\mathcal{F}}$ where $\mathcal{F} = \mathcal{R}^2 - \mathcal{B}_2(X)$. The golden mean shift is an irreducible nearest neighbour shift of finite type. It is irreducible because if $w_1, w_2 \in \mathcal{B}(X_{\{11\}})$ then $w_1 0 w_2 \in \mathcal{B}(X_{\{11\}})$. The study of nearest neighbour shifts of finite type was motivated by the fact that the support of stationary Markov chains are nearest neighbour shifts of finite type. In fact they were initially called intrinsic Markov chains.[12]

Definition. *A shift space X is called* non-wandering *if for any open set $U \subset X$ there exists $n \in \mathbb{N}$ such that $\sigma^n(U) \cap U \neq \phi$*

We need the following result from ergodic theory in the following chapters.

Lemma 3.0.8. *Let μ be a stationary probability measure such that $supp(\mu) = X$ is a nearest neighbour shift of finite type. Then X is a finite union of disjoint irreducible nearest neighbour shifts of finite type.*

Proof. Firstly, we will consider the non-wandering property of the support which is the consequence of *Poincaré Recurrence Theorem.*[8] Let U be an open set in X. Then $\mu(U) > 0$. Then there exists a $n \in \mathbb{N}$ such that $\sigma^n(U) \cap U \neq \phi$. Take $a_1 \in \mathcal{R}$. If not the set $\{\sigma^i(U) \mid i \in \mathbb{N}\}$ consists of disjoint sets all of the same measure. Then by stationarity of the measure $\mu(X) > \sum_i \mu(\sigma^i(U)) > n\mu(U)$ for all n. Therefore there exists a natural number $n \in \mathbb{N}$ such that $\sigma^n(U) \cap U \neq \phi$. Consider ~ a relation on $\mathbb{R}$ defined by

$$a \sim b \text{ if there exists } x \in X \text{ satisfying } x_i = a, x_j = b \text{ for some } i < j$$

We claim that ~ is an equivalence relation. By the non-wandering property $a \sim a$ for all $a \in \mathcal{R}$. Let $a \sim b$. Then there exists w such that $awb \in \mathcal{B}(X)$. By the non-wandering property there exists u such that $awbuawb \in \mathcal{B}(X)$. Hence $b \sim a$ Let $a \sim b$ and $b \sim c$. Then since X is a nearest neighbour shift of finite type $a \sim c$. Hence ~ is an equivalence relation. ~ partitions the alphabet into $\mathcal{R}_1, \mathcal{R}_2 \dots \mathcal{R}_n$. Define

$$X_i = \{x \in X \mid x_0 \in \mathcal{R}_i\} = X \cap \mathcal{R}_i^{\mathbb{Z}}$$

Then X_i's are shift spaces which partition X. X_i's are irreducible nearest neighbour shifts of finite type. To prove that X_i's are irreducible, let $a, b \in \mathcal{B}(X_i)$ for some i where a ends with alphabet u and b begins with w. $u \sim w$. So, there exists d such that $udw \in \mathcal{B}(X \cap \mathcal{R}_i^{\mathbb{Z}})$. Since X is a nearest neighbour shift of finite type $adb \in \mathcal{B}(X \cap \mathcal{R}_{\mathbb{Z}}^{i})$. Therefore X_i is irreducible. □

The period of an element $x \in X$ is given by

$$period(x) = \min\{i \in \mathbb{N} \mid \sigma^i(x) = x\}$$

Suppose an irreducible shift of finite type X is given. The *period* of X is defined as

$$period(X) = \text{the greatest common divisor of the periods of elements of X.}$$

It can be proved that periodic points are dense in every irreducible shift of finite type [10]. So the definition makes sense. For example, the period of $X_{\{11\}}$ is 1 because it has an element of period 1, namely 0^∞. An alternative and equivalent description is the following. Suppose $\mathcal{R}_1, \mathcal{R}_2 \dots \mathcal{R}_p$ partition $\mathcal{R}$ such that if $x \in X, x_1 \in A_1$ then $x_i \in A_{\bar{i}}$ where $\bar{i} = i - kp \in [1,p]$ for some $k \in \mathbb{Z}$. The smallest such p is the period of X.

In the full shift, any symbol can appear after a given symbol. Irreducible shifts of finite type do not have the same but a related property. The *offset* of an irreducible nearest shift of finite type X with period p is the smallest natural number t such that for any $a \in \mathcal{R}_i, b \in \mathcal{R}_j$ and $T > t$, there exists $x^T \in X$ such that $x^T_1 = a, x^T_{Tp+j-i+1} = b$. It can be proved that every irreducible nearest neighbour shift of finite has a finite offset.[10] For example, the offset of the golden mean shift is 1. This is because take $a, b \in \{0,1\}$ then $a0b \in \mathcal{B}(X_{\{11\}})$ however if $a, b = 1$ then $ab \notin \mathcal{B}(X_{\{11\}})$.

The entire framework can be generalised from $\mathbb{Z}$ to $\mathbb{Z}^d$. The full d-dimensional shift on alphabet $\mathcal{R}$ is the dynamical system $(\mathcal{R}^{\mathbb{Z}^d}, \sigma_1, \sigma_2 \dots \sigma_d)$ where $\mathcal{R}^{\mathbb{Z}^d}$ has the product topology over the discrete topology on $\mathcal{R}$ and the shift maps are defined by

$$(\sigma_i(x))_v = x_{v+e_i} \text{where } v \in \mathbb{Z}^d$$

and e_i is the vector with the ith coordinate 1 and all other entries 0.

As before, σ_i's are homeomorphisms of the space $\mathcal{R}^{\mathbb{Z}^d}$. Let G be the group generated by $\sigma_1, \sigma_2 \dots \sigma_d$. A d-dimensional shift space is a dynamical system $(X, \sigma_1, \sigma_2 \dots \sigma_d)$ where $X \subset \mathcal{R}^{\mathbb{Z}^d}$ is closed and invariant under G. Let

$$\mathcal{R}^\star = \{\mathcal{R}^W \mid W \text{ is a finite subset of } \mathbb{Z}^d\}$$

For any $\mathcal{F} \subset \mathcal{R}^\star$ let

$$X_\mathcal{F} = \{x \in \mathcal{R}^{\mathbb{Z}^d} | \text{ no translate of a subword of } x \text{ belongs to } \mathcal{F}\}$$

It can be checked that X is a shift space if and only if there exists a $\mathcal{F} \subset \mathcal{R}^\star$ such that $X = X_\mathcal{F}$. A shift space is called a shift of finite type if and only if $X = X_\mathcal{F}$ for some finite $\mathcal{F} \subset \mathcal{R}^\star$. A shift space is called a nearest neighbour shift of finite type if $X = X_\mathcal{F}$ for some $\mathcal{F} \subset \mathcal{R}^\star$ which are patterns on edges of $\mathbb{Z}^d$. A sliding block code is a map between d- dimensional shift space $\phi : X \longrightarrow Y$ such that $\phi \circ \sigma_i = \sigma_i \circ \phi$. The map is called a conjugacy if it is bijective. Two shift spaces are called conjugate if there exists a conjugacy between them. It can be shown that any shift of finite type is conjugate to a nearest neighbour shift of finite type. As before we can look at $\mathcal{R}^{\mathbb{Z}^d}$ as the space of configuration on graph $\mathbb{Z}^d$ where there is an edge from x to y if $x = y \pm e_i$ for some i. Again, the support of any probability measure stationary with respect to G on the graph $\mathbb{Z}^d$ is a shift space. Define the *language* of a shift space X on a set $A \subset \mathbb{Z}^d$ as

$$\mathcal{B}_A(X) = \{x|_A \mid x \in X\}$$

Suppose a stationary probability measure μ on $\mathcal{R}^{\mathbb{Z}^d}$ is given. Then the following can be proved. For any finite $A \subset \mathbb{Z}^d$

$$\mathcal{B}_A(supp(X)) = \{w \in \mathcal{R}^A \mid \mu([w,A]) > 0\}$$

Chapter 4

Markov Random Fields in 1 Dimension

In the following μ will always be a stationary Borel probability measure on the space $\mathcal{R}^{\mathbb{Z}}$ for this chapter and the alphabet $\mathcal{R}$ will always be finite. A probability measure μ is called *a Markov chain* if

$$\mu([x,0] \,\Big|\, [x_{-1}x_{-2}\ldots x_{-r},\{-1,-2\ldots -r\}]) = \mu([x,0] \,\Big|\, [x_{-1},-1])$$

We prove the following result in this chapter:

Theorem 4.0.9. *The following are equivalent :*

1. *μ is stationary Markov random field.*
2. *μ is stationary Markov chain*
3. *μ is a stationary measure which has a stationary nearest neighbour Gibbs potential*

The following short hand notation will often be used

$$x_{i_1}x_{i_2}\ldots x_{i_r} = [x_{i_1},x_{i_2}\ldots x_{i_r},\{i_1,i_2\ldots i_r\}]$$

The result is not true if $\mathcal{R}$ is not finite.[10]

Proof. By Lemma (2.0.3), (3) implies (1). It can be easily checked that (2) implies (3). We need to prove that (1) implies (2).

We first consider the case where $supp(\mu) = \mathcal{R}^{\mathbb{Z}}$. By Theorem (2.0.6) the measure has a stationary nearest neighbour Gibbs potential V.

If the measure has the following 'mixing' condition : for all $x_0, x_{-1}, x_{-2} \ldots x_{-r} \in \mathcal{R}$

$$\lim_{n\to\infty} \mu(x_0 | x_{-1}x_{-2}\ldots x_{-r}, (x_n = a)) \text{ exists and is independent of a } \in \mathcal{R}$$

then

$$\begin{aligned}
\mu(x_0 \mid x_{-1}x_{-2}\ldots x_{-r}) &= \lim_{n\to\infty} \frac{\sum_{a\in\mathcal{R}} \mu(x_0 x_{-1}\ldots x_{-r}, x_n = a)}{\sum_{a\in\mathcal{R}} \mu(x_{-1}\ldots x_{-r} x_n = a)} \\
&= \lim_{n\to\infty} \frac{\mu(x_0 x_{-1}\ldots x_{-r}, x_n = a)}{\mu(x_{-1}x_{-2}\ldots x_{-r}, x_n = a)} \text{ choosing any } a \in \mathcal{R} \\
&= \lim_{n\to\infty} \mu(x_0 \mid x_{-1}, x_n = a) \text{ since } \mu \text{ is a Markov random field} \\
&= \mu(x_0 \mid x_{-1}) \text{ by the assumption above}
\end{aligned}$$

Therefore we need to prove that

$$\lim_{n\to\infty} \mu(x_0 \mid x_{-1}x_{-2}\ldots x_{-r}(x_n = a)) \text{ exists and is independent of a } \in \mathcal{R}$$

We require the following version of the Perron-Frobenius theorem. [6]

Theorem. *Suppose T is a matrix of size $n \times n$ with all entries positive. Then there exists $\lambda > 0$ and vectors v, w with all components positive such that*

$$\lim_{n\to\infty} \frac{(T^n)_{i,j}}{\lambda^n v_i w_j} = 1$$

Let T be a matrix whos rows and columns are indexed by elements of the alphabet $\mathcal{R}$ and $T_{a,b} = e^{V([ab,\{0,1\}])}$ and $\lambda > 0, v, w \in \mathbb{R}^{|\mathcal{R}|}$ be such that

$$\lim_{n\to\infty} \frac{(T^n)_{i,j}}{\lambda^n v_i w_j} = 1$$

Then,

$$\begin{aligned}
\lim_{n\to\infty} \mu(x_0 \mid x_{-1}x_{-2}\ldots x_{-r}, x_n = a) &= \lim_{n\to\infty} \mu(x_0 \mid x_{-1}, x_n = a) \\
&= \lim_{n\to\infty} \frac{\mu(x_{-1}x_0, x_n = a)}{\mu(x_{-1}x_n = a)} \\
&= \lim_{n\to\infty} \frac{\sum\limits_{x_1x_2\ldots x_{n-1}\in\mathcal{R}^{n-1}} \prod\limits_{i=-1}^{l-1} e^{V(x_i x_{i+1})}}{\sum\limits_{x_0'x_1'x_2'\ldots x_{n-1}'\in\mathcal{R}^{n}} \prod\limits_{i=-1}^{l-1} e^{V(x_i' x_{i+1}')}} \\
&= \lim_{n\to\infty} \frac{e^{V(x_{-1}x_0)}(T^n)_{x_0,a}}{(T^{n+1})_{x_{-1},a}} \\
&= \frac{e^{V(x_{-1}x_0)} v_{x_0} w_a}{\lambda v_{x_{-1}} w_a} \\
&= \frac{e^{V(x_{-1}x_0)} v_{x_0}}{\lambda v_{x_{-1}}}
\end{aligned}$$

where $x_{-1}' = x_{-1}, x_n' = x_n$. Thus if μ is a stationary Markov random field such that $supp(\mu) = \mathcal{R}^{\mathbb{Z}}$ then μ is a stationary Markov chain.

Now, we will consider the general case. Suppose μ is a stationary Markov random field. We will first prove that $supp(\mu)$ is a nearest neighbour shift of finite type.

Let $\mu(a_1a_2\ldots a_n), \mu(a_na_{n+1}\ldots a_{n+m}) > 0$. We want to prove $\mu(a_1a_2\ldots a_{n+m}) > 0$. Let $\tilde{\mu}$ be the independent product $\mu\times\mu$ i.e. a probability measure on $(\mathcal{R}\times\mathcal{R})^{\mathbb{Z}}$ such that for all $c_i, d_i \in \mathcal{R}$

$$\tilde{\mu}\left(\substack{c_1c_2\ldots c_k \\ d_1d_2\ldots d_k}\right) = \mu(c_1c_2\ldots c_k)\mu(d_1d_2\ldots d_k)$$

It can be checked that $\tilde{\mu}$ is a stationary probability measure. We can choose $e_1e_2\ldots e_m$ and $f_1, f_2\ldots f_{n-1}$ in $\mathcal{R}$ such that

$$\mu(a_1a_2\ldots a_ne_1e_2\ldots e_m)\mu(f_1f_2\ldots f_{n-1}a_n\ldots a_{m+n}) > 0$$

Then by stationarity of $\tilde{\mu}$ the support has the non-wandering property. So we can choose $g_1, g_2 \dots g_r$ and $h_1, h_2 \dots h_r$ in $\mathcal{R}$ such that

$$\mu(a_1 a_2 \dots a_n e_1 e_2 \dots e_m g_1 g_2 \dots g_r a_1 a_2 \dots a_n e_1 e_2 \dots e_m) > 0$$
$$\mu(f_1 f_2 \dots f_{n-1} a_n \dots a_{m+n} h_1 h_2 \dots h_r f_1 f_2 \dots f_{n-1} a_n \dots a_{m+n}) > 0$$

Then by the first equation

$$\mu([a_1 a_2 \dots a_n, \{1, 2 \dots n\}] \cap [a_n, 2n+m+r]) > 0$$

and by the second one

$$\mu([a_n a_{n+1} \dots a_{m+n} h_1 h_2 \dots h_r f_1 f_2 \dots f_{n-1} a_n, \{n, n+1 \dots 2n+m+r\}]) > 0$$

Since μ is a stationary Markov random field, by Lemma (2.0.1) we get

$$\mu([a_1 a_2 \dots a_n a_{n+1} \dots a_{m+n} h_1 h_2 \dots h_r f_1 f_2 \dots f_{n-1} a_n, \{1, 2 \dots 2n+m+r]) > 0$$

Therefore $\mu(a_1 a_2 \dots a_{m+n}) > 0$. Thus $supp(\mu)$ is a nearest neighbour shift of finite type and consequently by Lemma 3.0.8 it is a union of disjoint irreducible nearest neighbour shifts of finite type.

Thus we can now assume that $supp(\mu)$ is an irreducible nearest neighbour shift of finite type X with period p and offset t. Let the $A_1, A_2 \dots A_p$ be a partition of $\mathcal{R}$ such that if $x \in X$ and $x_1 \in A_1$ then $x_i \in A_{\bar{i}}$ as described in Chapter 3. Define for every $r \in \mathbb{N}$ and i

$$\mathcal{B}^i_r(X) = \{a_1 a_2 \dots a_r \in \mathcal{B}_r(X) \mid a_1 \in A_i\}$$

Let $x_0 \in \mathcal{R}$, $r \in \mathbb{N}$, ($i \in \mathbb{N}$ large) and $x_{-r} \dots x_{-2} x_{-1} \in \mathcal{B}^1_r(X)$

$$\begin{aligned}
\mu(x_0 | x_{-1} x_{-2} \dots x_{-r}) &= \sum_{a_1 a_2 \dots a_r \in \mathcal{B}^1_r(X)} \mu((x_{itp-r} = a_1) \dots (x_{itp-1} = a_r), x_0 \mid x_{-1} x_{-2} \dots x_{-r}) \\
&= \sum \mu(x_0 \mid x_{-1} x_{-2} \dots x_{-r} x_{itp-r} \dots x_{itp-1}) \\
&\quad \mu(x_{itp-r} \dots x_{itp-1} \mid x_{-1} x_{-2} \dots x_{-r}) \text{ by Bayes' Theorem} \\
&= \sum \mu(x_0 \mid x_{-1} x_{itp-r}) \mu(x_{itp-r} \dots x_{itp-1} \mid x_{-1} x_{-2} \dots x_{-r}) \\
&\quad \text{by the Markov random field property}
\end{aligned}$$

Now to understand the expression, we need to consider the map

$$\phi : X \longrightarrow \{\mathcal{B}_r^1(X)\}^{\mathbb{Z}}$$

given by

$$(\phi(x))_i = (x_{itp-r-j+1} \ldots x_{itp-1-j+1}) \text{ if } x_{-r} \in A_j$$

Then ϕ is surjective and continuous. Let μ' be a probability measure on $\{\mathcal{B}_r^1(X)\}^{\mathbb{Z}}$ given by the push forward of the map μ by ϕ i.e. $\mu'(U) = \mu(\phi^{-1}(U))$. Then it can be checked that μ' is a stationary markov random field such that $supp(\mu') = \{\mathcal{B}_r^1(X)\}^{\mathbb{Z}}$. Therefore μ' is a stationary Markov chain. Let it have a stationary distribution π. Then the following is true. [4]

$$\begin{aligned}
\lim_{i\to\infty} \mu(x_{itp-r} \ldots x_{itp-1} \mid x_{-1}x_{-2} \ldots x_{-r}) &= \lim_{i\to\infty} \mu'(x_{itp-r} \ldots x_{itp-1} \mid x_{-1}x_{-2} \ldots x_{-r}) \\
&= \pi(x_{itp-r} \ldots x_{itp-1})
\end{aligned}$$

Take a sequence $\{i_k\}_{k=1}^{\infty} \subset \mathbb{N}$ such that

$$\lim_{k\to\infty} \mu(x_0 \mid x_{-1}(x_{i_ktp-r} = a_1))$$

exists for all $a_1 \in A_1$. Then,

$$\begin{aligned}
\mu(x_0 \mid x_{-1}x_{-2} \ldots x_{-r}) &= \lim_{k\to\infty} \sum \mu(x_0 \mid x_{-1}x_{i_ktp-r})\mu(x_{i_ktp-r} \ldots x_{i_ktp-1} \mid x_{-1}x_{-2} \ldots x_{-r}) \\
&= \sum \pi(a_1a_2 \ldots a_r) \lim_{k\to\infty} \mu(x_0 \mid x_{-1}x_{i_ktp-r})
\end{aligned}$$

which is independent of $x_{-2}, x_{-3} \ldots x_{-r}$. Therefore,

$$\mu(x_0 \mid x_{-1}x_{-2} \ldots x_{-r}) = \mu(x_0 \mid x_{-1})$$

which proves that μ is a stationary Markov chain. $\square$

The theorem does not hold if the measure is not assumed to be stationary. Consider any probability measure μ such that $supp(\mu)$ is the shift space

$$\{0^\infty, 0^\infty 1^\infty, 1^\infty, 1^\infty 0 2^\infty, 2^\infty\}$$

To clarify the notation, $1^\infty 0 2^\infty$ refers to the point x such that

$$x_i = \begin{cases} 1 \text{ if } i < 0 \\ 0 \text{ if } i = 0 \\ 2 \text{ if } i > 0 \end{cases}$$

and all its shifts. The shift space is countable. It can be checked that any such measure is a Markov random field. But the measure cannot be a Markov chain of any order because the support is a shift space which is not a shift of finite type. This is an elaboration of an example as given in [3].

Chapter 5

Markov Random Fields in 2 Dimensions

In the last chapter, we saw that stationary Markov random fields in 1 dimension are measures with nearest neighbour Gibbs potentials. This chapter intends to show how this result fails in 2 dimensions. One of the key observations in the proof of Theorem 4.1 was that the support of any stationary Markov random field in 1 dimension is a nearest neighbour shift of finite type. Our first example will be that of a stationary Markov random field in 2 dimensions such that its support is not a nearest neighbour shift of finite type. In fact it is not a shift of finite type at all.

Let ν be some stationary probability measure on $\{0,1\}^{\mathbb{Z}}$. Consider the probability measure μ on $\{0,1\}^{\mathbb{Z}^2}$ such that

$$\mu([a,A]) = \begin{cases} 0 \text{ if } a(m,n) \neq a(m,n+1) \text{for some } m \text{ and } n \\ \nu([b,B]) \text{ where } B = \{x \mid \text{ there exists } y \text{ such that } (x,y) \in A\} \\ \text{and } b(t) = a(t,m) \text{ if } (t,m) \in B \text{ for some } m \in \mathbb{Z} \end{cases}$$

for all A finite, $a \in \{0,1\}^A$. The measure constrains the symbols to remain constant vertically and behave with accordance to ν horizontally. Then

$$\mu([a,A] \,\big|\, [b,B]) = \begin{cases} 1 \text{ if } a(i,j) = b(i,k) \text{ for } (i,j) \in A, (i,k) \in B \\ 0 \text{ otherwise} \end{cases}$$

provided $\partial A \subset B \subset A^c$ and $\mu([b,B]) > 0$. That is, given a boundary configuration with positive measure there is a single way to fill it in to be in the support of μ. It follows that

$$\mu([a,A] \mid [b,B]) = \mu([a,A] \mid [b|_{\partial A,}, \partial A])$$

This proves that μ is a stationary Markov random field. If we choose ν such that $supp(\nu)$ is not a nearest neighbour shift of finite type then μ is a stationary Markov random field whos support is not a nearest neighbour shift of finite type.There are many such examples.

We will now give an example showing that the equivalence of (1) and (3) in Theorem (4.0.9) fails in 2 dimensions. We will take an example on a finite graph as given in [11] and the extend it to the $\mathbb{Z}^2$ lattice. Recall the characterisation of Markov random fields on finite graphs from Equation (2.0.1): μ is a Markov random field on a finite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with alphabet $\mathcal{R}$ if and only if for all A, B independent $a, a' \in \mathcal{R}^A, b, b' \in \mathcal{R}^B, c \in \mathcal{R}^C$ where $C = (A \cup B)^c$

$$\mu(a \quad b \quad c)\mu(a' \quad b' \quad c) = \mu(a \quad b' \quad c)\mu(a' \quad b \quad c)$$

Consider $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{0, 1, 2 \ldots 5\}$ and $\mathcal{E}$ given by Figure 1.1.

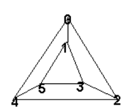


Figure 5.1: $\mathcal{G}$

The alphabet is $\mathcal{R} = \{a, b\}$. Consider the configurations given in Figure 1.2.
It can be checked that any pair of the configurations H, I, K, L and M differ on a connected subset of the vertices. Consider any probability measure μ on $\mathbb{R}^{\mathcal{V}}$ such that

$$supp(\mu) = \{H, I, J, K, L, M\} \tag{5.0.1}$$

and

$$\mu(\{H\})\mu(\{J\})\mu(\{L\}) \neq \mu(\{I\})\mu(\{K\})\mu(\{M\}) \tag{5.0.2}$$

We will use equation (5.0.1) to prove that μ is a Markov random field and equation

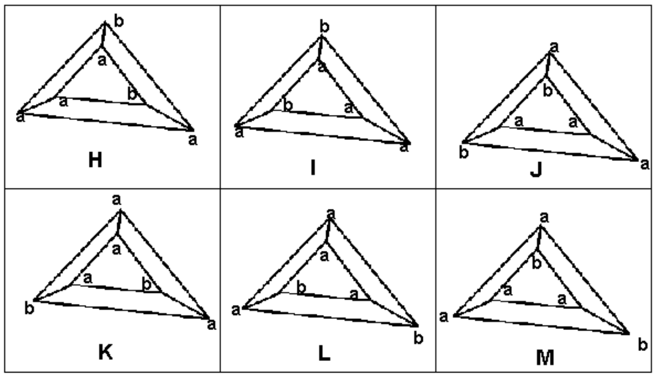


Figure 5.2: The Configurations

(5.0.2) to prove that it does not have a nearest neighbour Gibbs potential. Since any pair of the configurations H, I, J, K, L and M differ on a connected subset of the vertices the measure μ is a Markov random field. To see this, take two sets $A, B \subset \mathcal{V}$ which are independent. Let $C = (A \cup B)^c$. Consider configurations $x, x' \in \mathcal{R}^A, y, y' \in \mathcal{R}^B$ and $z \in \mathcal{R}^C$. Then $(x \quad y \quad z), (x' \quad y' \quad z) \in supp(\mu)$ implies $x = x'$ or $y = y'$. Therefore,

$$\mu(x \quad y \quad z)\mu(x' \quad y' \quad z) = \mu(x' \quad y \quad z)\mu(x' \quad y \quad z)$$

So the measure is a Markov random field.

Suppose the measure did have a nearest neighbour Gibbs potential V. Then,

$$\begin{aligned} \mu(\{H\})\mu(\{J\})\mu(\{L\}) &= \prod_{C \text{ cliques in } \mathcal{G}} e^{V(H|_C)} e^{V(J|_C)} e^{V(L|_C)} \\ &= \prod_{C \text{ cliques in } \mathcal{G}} e^{V(I|_C)} e^{V(K|_C)} e^{V(M|_C)} \\ &= \mu(\{I\})\mu(\{K\})\mu(\{M\}) \end{aligned}$$

since H, J and L together have the same collection of configurations on every clique as I, K and L. But since we have chosen a measure such that

$$\mu(\{H\})\mu(\{J\})\mu(\{L\}) \neq \mu(\{I\})\mu(\{K\})\mu(\{M\})$$

μ has no nearest neighbour Gibbs potential.

We will transport this example to one in the $\mathbb{Z}^2$ lattice by stretching out edges and introducing some new symbols. Consider the graph $\mathcal{G}'$ with vertices $\mathcal{V}' = \{1,2,3\ldots 25\}$ and edges as shown in Figure 3.

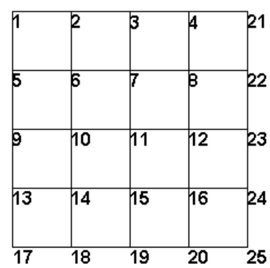
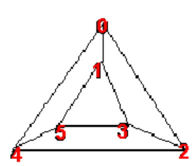


Figure 5.3: $\mathcal{G}'$

The new alphabet shall be $\mathcal{R}' = (\mathcal{R} \cup (\mathcal{R} \times \mathcal{R}) \cup \{\phi\}) \times \{1,2,3\ldots 25\}$. We will construct configurations H', I', J', K', L' and M' corresponding to the configurations H, I, J, K, L and M given before. The graph $\mathcal{G}'$ is identified with the the graph $\mathcal{G}$ as in Figure 1.4.



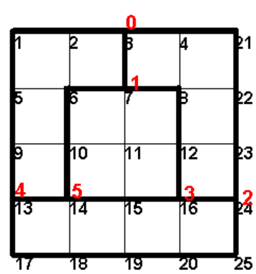


Figure 5.4: Correspondence between $\mathcal{G}$ and $\mathcal{G}'$

The darkened edges act like wires carrying information between the marked vertices. For the configurations in the support, the symbol on vertex 11 is always ϕ, The vertices 3, 7, 16, 24, 13 and 14 take symbols as on 0, 1, 3, 2, 4 and 5 respectively. The edges on the darkened edges carry information about the symbols from one end to another. For example look at Figure 1.5 to see how H' corresponds to H.

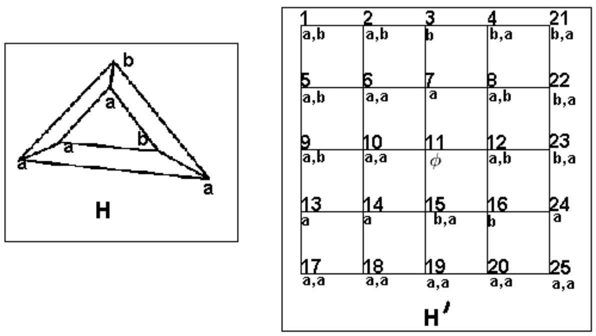


Figure 5.5: Correspondence between H and H'

To illustrate how the darkened edges act like wires note that in H vertex 0 has symbol b and vertex 4 has symbol a. Consequently, in H' vertex 3 has symbol b and vertex 13 has symbol a. Going clockwise vertex 13 comes before vertex 3. So the vertices in between have symbols (a,b). As in the last example, we consider a probability measure μ' on $\mathcal{R}'^{\mathcal{V}'}$ such that

$$\mu'(\{H'\})\mu'(\{J'\})\mu'(\{L'\}) \neq \mu'(\{I'\})\mu'(\{K'\})\mu'(\{M'\})$$

and

$$supp(\mu') = \{H', J', L', I', K', M'\}$$

For reasons as before the measure is a Markov random field such that it does not have a nearest neighbour Gibbs potential.

Now, divide the $\mathbb{Z}^2$ lattice into disjoint 5×5 grids. There are 25 ways of doing this. Choose the one in which the origin is left-upper corner of a certain grid. Now place configurations H', I', J', K', L', M' independently on the grids with probability

μ'. This gives a measure $\tilde{\mu}$ on ${\mathcal{R}'}^{\mathbb{Z}^2}$.Consider the measure ν on ${\mathcal{R}'}^{\mathbb{Z}^2}$ given by

$$\nu(U) = \frac{1}{25} \sum_{0 \le i,j \le 4} \tilde{\mu}(\sigma_1{}^i \sigma_2{}^j(U)) \text{ for any measurable set } U$$

Proposition 5.0.10. *ν is a stationary Markov random field that does not have a nearest neighbour Gibbs potential.*

Note, that in this example $supp(\nu)$ is a shift of finite type but not a nearest neighbour shift of finite type.

Proof. We will first prove that the measure is stationary. Take any cylinder set $[a,A]$ such that $\tilde{\mu}([a,A]) > 0$ where $a \in (\mathcal{R}')^A$ and $A \subset \mathbb{Z}^2$ finite. While setting up the configurations, we began by partitioning $\mathbb{Z}^2$ into 5×5 blocks. Let the partitioning sets be denoted by $\{T_i\}_{i=1}^{\infty}$. This also gives a partition of A, say $\{A_i\}_{i=1}^{n}$. Since the configurations are put independently on the grids with probability μ' on the 5×5 boxes

$$\begin{aligned}
\tilde{\mu}([a,A]) &= \tilde{\mu}(\bigcap_{i=1}^{n}[a|_{A_i}, A_i]) \\
&= \prod_{i=1}^{n} \tilde{\mu}([a|_{A_i}, A_i]) \\
&= \prod_{i=1}^{n} \tilde{\mu}(\sigma_k^5([a|_{A_i}, A_i])) \\
&= \tilde{\mu}(\sigma_k^5([a,A])) \\
&= \tilde{\mu}(\sigma_k^{-5}([a,A])) \text{ for } k = 1,2
\end{aligned}$$

Therefore, for any cylinder set $[a,A] \subset (\mathcal{R}')^{\mathbb{Z}^2}$

$$\tilde{\mu}([a,A]) = \tilde{\mu}(\sigma_1^5([a,A])) = \tilde{\mu}(\sigma_2^5([a,A]))$$

Since cylinder sets generate the borel sigma-algebra of ${\mathcal{R}'}^{\mathbb{Z}^2}$, for any measurable set U

$$\tilde{\mu}(U) = \tilde{\mu}(\sigma_1^5(U)) = \tilde{\mu}(\sigma_2^5(U))$$

Therefore, for any measurable set $U \subset \mathcal{R}'^{\mathbb{Z}^2}$

$$\nu(\sigma_1(U)) = \frac{1}{25}\sum_{0\le i,j\le 4}\tilde{\mu}(\sigma_1{}^i\sigma_2{}^j(\sigma_1(U))) = \frac{1}{25}\sum_{0\le i,j\le 4}\tilde{\mu}(\sigma_1{}^i\sigma_2{}^j(U)) = \nu(U)$$

and

$$\nu(\sigma_2(U)) = \frac{1}{25}\sum_{0\le i,j\le 4}\tilde{\mu}(\sigma_1{}^i\sigma_2{}^j(\sigma_2(U))) = \frac{1}{25}\sum_{0\le i,j\le 4}\tilde{\mu}(\sigma_1{}^i\sigma_2{}^j(U)) = \nu(U)$$

Therefore ν is stationary.

Now, we will prove that ν is a Markov random field. Suppose A is some non-empty finite subset of $\mathbb{Z}^2$. Let $x \in supp(\nu)$ and $a \in \mathcal{R}'^{\mathbb{Z}^2}$ be some configuration. Let B be another finite set such that $\partial A \subset B \subset A^C$. We want to prove

$$\nu([a,A]\Big|[x|_B,B]) = \nu([a,A]\Big|[x|_{\partial A},\partial A])$$

To construct configurations on $\mathbb{Z}^2$ we began by dividing $\mathbb{Z}^2$ into 5×5 blocks and then placing the configurations on these. Each symbol used has a number in between 1 and 25 which determines how this division has been done. By looking at $x(j,k)$ for any $(j,k) \in \partial A$ one can determine how the division has been made. Take some $(j,k) \in \partial A$. $\mathbb{Z}^2$ can be partitioned into 5×5 blocks T_i's and numbered according to Figure 1.3 such that the number on (j,k) is the same as the number designated to $a(j,k)$. This will give a partition of A say $\{A_i\}_{i=1}^n$. By renumbering the T_i's we can assume that each $A_i \subset T_i$. Since the configurations are put independently on the grids with probability μ' on the 5×5 boxes

$$\begin{aligned}
\nu([a,A]\Big|[x|_B,B]) &= \prod_{i=1}^n \nu([a|_{A_i},A_i]\Big|[x|_B,B])\text{by independence}\\
&= \prod_{i=1}^n \nu([a|_{A_i},A_i]\Big|[x|_{B\cap T_i},B\cap T_i])\\
&= \prod_{i=1}^n \nu([a|_{A_i},A_i]\Big|[x|_{\partial A_i\cap T_i},\partial A_i\cap T_i])\\
&\quad\text{by the Markov random field property of }\mu'\\
&= \nu([a,A]\Big|[x|_{\partial A},\partial A])
\end{aligned}$$

Now we will prove that the measure does not have a nearest neighbour Gibbs potential. Let $A \subset \mathbb{Z}^2$ be a 5×5 block. Let $H^o, I^o, J^o, K^o, L^o, M^o$ be configurations on A corresponding to H', I', J', K', L', M'. Now consider a configuration a on ∂A such that

$$\nu([H^o, A] \cap [a, \partial A]) > 0$$

This assumption implies that

$$\nu([H^o,A] \cap [a,\partial A])\nu([J^o,A] \cap [a,\partial A])\nu([L^o,A] \cap [a,\partial A]) > 0$$
$$\nu([I^o,A] \cap [a,\partial A])\nu([K^o,A] \cap [a,\partial A])\nu([M^o,A] \cap [a,\partial A]) > 0$$

Then,

$$\begin{aligned}
& \nu([H^o,A] \mid [a,\partial A])\nu([J^o,A] \mid [a,\partial A])\nu([L^o,A] \mid [a,\partial A]) \\
= \; & \mu'(\{H'\})\mu'(\{J'\})\mu'(\{L'\}) \\
\neq \; & \mu'(\{I'\})\mu'(\{K'\})\mu'(\{M'\}) \\
= \; & \nu([I^o,A] \mid [a,\partial A])\nu([K^o,A] \mid [a,\partial A])\nu([M^o,A] \mid [a,\partial A])
\end{aligned}$$

Therefore, the measure does not have a nearest neighbour Gibbs potential. □

We were prompted by this example to study Markov random fields from another direction. It turns out that if the support of a Markov random field has the pivot property(defined below) then there is another compact way of representing its conditional probabilities.

Let D be a set of finite subsets of $\mathbb{Z}^2$. Let $A \in D, X \subset \mathcal{R}^{\mathbb{Z}^2}$ a shift space and x, y be distinct elements of $\mathcal{B}_{A \cup \partial A}(X)$ such that $x = y$ on ∂A. Then a *pivot* from x to y in the shift space X is a sequence of points $x = x_1, x_2, x_3 \ldots x_n = y \in \mathcal{B}_{A \cup \partial A}(X)$ such that $x_i|_{\partial A} = x_j|_{\partial A}$ for all i and j and $|\{r \in A | x_i(r) \neq x_{i+1}(r)\}| = 1$

Definition. *A shift space $X \subset \mathcal{R}^{\mathbb{Z}^2}$ is said to have the $D-$pivot property if for all $A \in D$ and $x, y \in \mathcal{B}_{A \cup \partial A}(X)$ such that $x = y$ on ∂A there is a pivot from x to y in X.*

This means that any configuration can be changed site by site to obtain any other configuration if they agree on the boundary of an element of D. Note that a shift space with safe symbol or more specifically full support has the $D-pivot$ property where D is the set of all finite subsets of $\mathbb{Z}^2$.

This property was initially explored by us to give another proof of the Hammersley-Clifford Theorem. However the property has some other consequences which will be discussed in this chapter and the next.

Suppose a stationary Markov random field μ is given on $\mathcal{R}^{\mathbb{Z}^2}$ and D be the set of all finite subsets of $\mathbb{Z}^2$. Let $supp(\mu) = X$. The *specification of* μ is the function $\Delta : \{A \times \mathcal{B}_{A\cup\partial A}(X) | A \in D\} \longrightarrow \mathbb{R}$ given by

$$\Delta(A,x) = \mu([x|_A, A] \mid [x|_{\partial A}, \partial A])$$

The *infralocal specification of* μ is the function $\delta : \mathcal{B}_{\{(0,0)\}\cup\partial\{(0,0)\}}(X) \longrightarrow \mathbb{R}$ given by

$$\delta(x) = \mu([x|_{(0,0)}, \{(0,0)\}] \mid [x|_{\partial\{(0,0)\}}, \partial\{(0,0)\}])$$

Proposition 5.0.11. *Suppose D is a set of finite subsets of $\mathbb{Z}^2$ such that any finite subset of $\mathbb{Z}^2$ is contained in an element of D. Let X be a shift space with the $D-pivot$ property and μ be a stationary Markov random field such that $supp(\mu) = X$. Then there is an expression of the specification of μ in terms of the infralocal specification of μ.*

The proposition can be restated with $\mathbb{Z}^2$ replaced by any graph.

Proof. Suppose D, X and μ be as stated in the lemma. Suppose the infralocal specification of μ is given. We will determine the specification of the measure. Let $A \subset \mathbb{Z}^2$ be finite and $x, y \in \mathcal{B}_{A\cup\partial A}(X)$ be distinct such that $x = y$ on ∂A.Let $B \in D$ such that $A \cup \partial A \subset B$. Let $z \in \mathcal{B}_{B\cup\partial B}(X)$ such that $z = x$ on ∂A. Since μ is a Markov random field, $supp(\mu)$ is a topological Markov field. So there exist $x_1, y_1 \in \mathcal{B}_{B\cup\partial B}(X)$ such that $x_1 = x$ and $y_1 = y$ on $A \cup \partial A$ and $x_1 = y_1 = z$ on $B \cup \partial B - A \cup \partial A$.

Let $x = (x_1 \ldots x_n = y_1)$ be a $B - pivot$. Let $x_i \neq x_{i+1}$ at $v_i = (a_i, b_i)$ for all i. Then,

$$\begin{aligned}
\frac{\Delta(A,x)}{\Delta(A,y)} &= \frac{\mu([x|_A, A] \,\big|\, [x|_{\partial A}, \partial A])}{\mu([y|_A, A] \,\big|\, [y|_{\partial A}, \partial A])} \\
&= \frac{\mu([x, A \cup \partial A])}{\mu([y, A \cup \partial A])} \\
&= \frac{\mu([x_1, B \cup \partial B])}{\mu([x_n, B \cup \partial B])} \text{ since } \mu \text{ is a Markov random field} \\
&= \prod_{i=1}^{n-1} \frac{\mu([x_i, B \cup \partial B])}{\mu([x_{i+1}, B \cup \partial B])} \\
&= \prod_{i=1}^{n-1} \frac{\mu([x_i|_{\{v_i\} \cup \partial(\{v_i\})}, \{v_i\} \cup \partial(\{v_i\})])}{\mu([x_{i+1}|_{\{v_i\} \cup \partial(\{v_i\})}, \{v_i\} \cup \partial(\{v_i\})])} \\
&\quad \text{since } \mu \text{ is a Markov random field} \\
&= \prod_{i=1}^{n-1} \frac{\delta(\sigma_1^{a_i} \sigma_2^{b_i}(x_i))}{\delta(\sigma_1^{a_i} \sigma_2^{b_i}(x_{i+1}))}
\end{aligned}$$

□

Therefore the specification is completely determined by the infralocal specification. The proof gives an expression for the specification in terms of the infralocal specification. However the expression depends on the pivot chosen between the two configurations. This indicates that the infralocal specification is not arbitrary and depends on the shift space X. To understand the constraints satisfied by the infralocal specification, we have to gain a deeper insight into pivoting process. This is pivoting at a single site. One might ask whether nearest neighbour shifts of finite type always have this pivot property of some finite size. However there do exist shifts of finite type such that they do not have pivots of any fixed finite size. We will continue this discussion at the end of next chapter.

Chapter 6

Further Work

In this chapter, we will discuss some questions which arose in the study and, to the author, are still open. This chapter lacks proofs and the emphasis is on a quick introduction to various questions.

6.1 Generalising Theorem 2.0.4

In Theorem (2.0.4) we proved the equivalence of Markov random fields and measures with nearest neighbour Gibbs potential under the assumption that the support of the measure has a safe symbol. We were given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a Markov random field on $\mathcal{R}^{\mathcal{V}}$ such that its support had a safe symbol '0'

Let

$$H = \{a \in supp(\mu) \mid (a, 0^{\mathcal{V}}) \text{ is a homoclinic pair } \}$$

and $\mathbf{A}$ be a matrix whose rows are indexed by pairs in H and the columns by elements of $\mathcal{T}$ i.e. configurations on cliques with positive probability. Suppose $c \in \mathcal{T}$ is a configuration on C and $a, b \in H$. The matrix entries were given by

$$\mathbf{A}_{(a,b),c} = \begin{cases} 1 & \text{if } a|_C = c, b|_C \neq c \\ -1 & \text{if } a|_C \neq c, b|_C = c \\ 0 & \text{otherwise} \end{cases}$$

An essential part of the proof was to show that

$$V' = \{v \mid v\mathbf{A} = 0 \text{ and } v \text{ has finitely many non-zero entries}\}$$

is generated by

$$\tilde{V} = \{v \mid v\mathbf{A} = 0 \text{ and } v \text{ has at most 2 non-zero entries}\}$$

Now, let a set $X \subset \mathcal{R}^{\mathcal{V}}$ be given and H be the set of homoclinic pairs in X. As before $\mathbf{A}$ is a matrix whose rows are indexed by pairs in H and the columns by elements of $\mathcal{T}$ i.e. configurations on cliques in X. Suppose $c \in \mathcal{T}$ is a configuration on C and $a, b \in H$. The matrix entries are given by

$$\mathbf{A}_{(a,b),c} = \begin{cases} 1 & \text{if } a|_C = c, b|_C \neq c \\ -1 & \text{if } a|_C \neq c, b|_C = c \\ 0 & \text{otherwise} \end{cases}$$

Then X is called *Gibbs-feasible* if

$$V' = \{v \mid v\mathbf{A} = 0 \text{ and } v \text{ has finitely many non-zero entries}\}$$

is generated by

$$\tilde{V} = \{v \mid v\mathbf{A} = 0 \text{ and } v \text{ has at most 2 non-zero entries}\}$$

Then the proof of Theorem (2.0.4) shows that

Theorem 6.1.1. *Let μ be a Markov random field such that $supp(\mu)$ is Gibbs-feasible. Then μ has a nearest neighbour Gibbs potential.*

Question 1 What are nice conditions on a shift space X such that it becomes Gibbs-feasible?

One such nice condition is that X has a safe symbol.

6.2 Topological Markov Fields

Topological Markov fields are central to this thesis. They played an important role in the proof of Theorem 4.0.9 which states that stationary Markov random fields are Markov chains in 1 dimension. In the proof we showed that topological Markov fields on $\mathbb{Z}$ which support a stationary probability measure are nearest neighbour shifts of finite type. A natural generalisation of the class of shifts of finite type is the class of *sofic shifts* defined as shift spaces which are factors of shifts of finite type. The following is true.

Lemma 6.2.1. *Suppose $X \subset \mathcal{R}^{\mathbb{Z}}$ is a shift space such that it is a topological Markov field. Then X is a sofic shift.*

With the help of this lemma, the following is a consequence of the proof of Theorem (4.0.9).

Theorem 6.2.2. *Suppose $X \subset \mathcal{R}^{\mathbb{Z}}$ is a shift space. The following are equivalent:*

1. *X is the support of a stationary Markov chain.*
2. *X is the support of a stationary Markov random field.*
3. *X is a topological Markov field and $X \times X$ is non-wandering.*
4. *X is a topological Markov field and X is non-wandering.*
5. *X is a disjoint union of finitely many irreducible nearest neighbor shifts of finite type.*

However there does not seem to be any such theorem in dimensions greater than 1. The following aspect has not yet been explored but is of interest to the author. $X \subset \mathcal{R}^{\mathbb{Z}^2}$ is called a *global topological Markov field* if for all $x, y \in X$ such that $x = y$ on ∂C for some $C \subset \mathcal{V}$ which is not necessarily finite. Then $z \in \mathcal{R}^{\mathcal{V}}$ defined by

$$z = \begin{cases} x \text{ on } C \cup \partial C \\ y \text{ on } (C \cup \partial C)^c \end{cases}$$

is an element of X.

Question 2 Suppose X is a global topological Markov field and a shift space. Does this imply that it is sofic?

Another point of interest is the characterisation of the supports of a stationary Markov random field which are necessarily shift spaces. Suppose $X \subset \mathcal{R}^{\mathbb{Z}}$ is a shift space. Then X is the support of a stationary Markov random field if and only if it is support of a stationary probability measure and is a topological Markov field. This lead us to the following question.

Question 3 Let $X \subset \mathcal{R}^{\mathbb{Z}^2}$. If X is the support of a stationary Markov random field, then X is the support of a stationary probability measure and X is a topological Markov field. Is the converse true?

6.3 Pivot Property and 3-Checkerboard

The *n-checkerboard* is a shift space given by

$$X_n = \{x \in \{0,1,2\dots n-1\}^{\mathbb{Z}^2} \,|\, x(i,j) \neq x(i,j+1) \text{ and } x(i,j) \neq x(i+1,j) \text{ for all } i \text{ and } j\}$$

Then, as in [5], X_3 and X_n for $n \geq 6$ has the $D-pivot$ property where D is the set of diamonds D_n's.

$$D_n = \{(x,y) \in \mathbb{Z}^2 \,\Big|\, |x| + |y| \leq n\}$$

D can be replaced by the set of finite simply-connected subsets of $\mathbb{Z}^2$

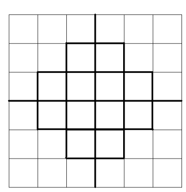


Figure 6.1: The thicker lines represent the subgraph induced by D_3

We will work with the 3-checkerboard, however the methods employed are more general.

Suppose μ is a stationary Markov random field such that $supp(\mu) = X_3$. It is to be noted that we do not know if such a measure exists. In the remainder of this chapter we shall derive some features of such a measure. Then by proof of Lemma 5.0.11 there is an expression for the specification Δ in terms of the infralocal specification δ. In X_3, the symbol at a site can change if and only if the neighbourhood is monochromatic i.e. a single symbol occupies the sites in the neighbourhood. Let p, x, q, r, y and $z \in \mathcal{B}_{D_2}(X_3)$ be as in Figure 1.7.

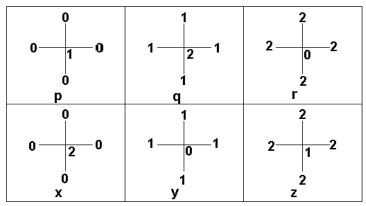


Figure 6.2: p,x,q,r,y and z

Let

$$
\begin{aligned}
v_1 &= \log(\delta(p)) - \log(\delta(x)) \\
v_2 &= \log(\delta(q) - \log(\delta(y)) \\
v_3 &= \log(\delta(r)) - \log(\delta(z)).
\end{aligned}
$$

Then v_1, v_2 and v_3 completely determine the specification Δ in the following way. Let $d = (d_1, d_2 \ldots d_m)$ be a D_n-pivot. Define

$$
\begin{aligned}
a_d = \; & |\{i | d_i(j) = 1, d_{i+1}(j) = 2 \text{ for some j} \in D_n\}| \\
& -|\{i | d_i(j) = 2, d_{i+1}(j) = 1 \text{ for some j} \in D_n\}| \\
b_d = \; & |\{i | d_i(j) = 2, d_{i+1}(j) = 0 \text{ for some j} \in D_n\}| \\
& -|\{i | d_i(j) = 0, d_{i+1}(j) = 2 \text{ for some j} \in D_n\}| \\
c_d = \; & |\{i | d_i(j) = 0, d_{i+1}(j) = 1 \text{ for some j} \in D_n\}| \\
& -|\{i | d_i(j) = 1, d_{i+1}(j) = 0 \text{ for some j} \in D_n\}|
\end{aligned}
$$

These record the number of switches of each kind. Suppose $A \subset \mathbb{Z}^2$ is finite. Let $x, y \in \mathcal{B}_{A \cup \partial A}(X_3)$ be distinct such that $x = y$ on ∂A. Let $z \in X_3$ such that $z = x$ on ∂A and k such that $A \cup \partial A \subset D_k$. Then as in the proof of Proposition 5.0.11, we consider points $x_1, y_1 \in \mathcal{B}_{D_k \cup \partial D_k}(X_3)$ such that

$$x_1 = x \text{ and } y_1 = y \text{ on } A \cup \partial A$$
$$\text{and } x_1 = y_1 = z \text{ on } D_k \cup \partial D_k - (A \cup \partial A)$$

Now consider the pivot $d = (d_1, d_2 \dots d_n)$ such that $d_1 = x_1$ and $d_n = y_1$ Then,

$$\frac{\Delta(A,x)}{\Delta(A,y)} = e^{v_1 a_d + v_2 b_d + v_3 c_d} \tag{6.3.1}$$

Suppose μ has a stationary nearest neighbour Gibbs potential V. We can assume that the potential is zero on configurations on single sites. Then,

$$v_1 = V(01) + V(10) + V\binom{0}{1} + V\binom{0}{1} - V(02) - V(20) - V\binom{0}{2} - V\binom{2}{0}$$
$$v_2 = V(12) + V(21) + V\binom{2}{1} + V\binom{1}{2} - V(01) - V(10) - V\binom{0}{1} - V\binom{1}{0}$$
$$v_3 = V(02) + V(20) + V\binom{2}{0} + V\binom{0}{2} - V(21) - V(12) - V\binom{2}{1} - V\binom{1}{2}$$

Therefore, $v_1 + v_2 + v_3 = 0$. The converse is also true.

Proposition 6.3.1. *Suppose μ is a stationary Markov random field such that $supp(\mu) = X_3$ Let v_1, v_2 and v_3 be as described above. Then μ has a stationary nearest neighbour Gibbs potential if and only if $v_1 + v_2 + v_3 = 0$*

Given this lemma, we would like to know if the infralocal specifications satisfy any relations. For this we refer to the proof of Proposition 5.0.11. There we derived an expression for the specification in terms of the infralocal specification. However this expression depended upon the pivot chosen between the two configurations. Different pivots might lead to different expressions. This will lead to an equation which the infralocal specification must satisfy. To be more formal, let X, μ and D be as in Proposition 5.0.11. Relationships arise in the following two ways:

1. Suppose $A \subset \mathbb{Z}^2$ finite such that $A \notin D$. Let $x, y \in \mathcal{B}_{A \cup \partial A}(X)$ be distinct such that $x = y$ on ∂A. Let $z_1, z_2 \in X$ such that $z_1 = z_2 = x$ on ∂A and $B \in D$ such that $A \cup \partial A \subset B$. Then as in the proof of the lemma, we consider points

$x_1, y_1 \in \mathcal{B}_{B\cup\partial B}(X)$ such that

$$x_1 = x \text{ and } y_1 = y \text{ on } A \cup \partial A$$
$$\text{and } x_1 = y_1 = z_1 \text{ on } B \cup \partial B - A \cup \partial A$$

Similarly $x_2, y_2 \in \mathcal{B}_{B\cup\partial B}(X)$ are chosen such that

$$x_2 = x \text{ and } y_2 = y \text{ on } A \cup \partial A$$
$$\text{and } x_2 = y_2 = z_2 \text{ on } B \cup \partial B - A \cup \partial A$$

Now consider pivots $a = (a_1, a_2 \dots a_n), b = (b_1, b_2 \dots b_n)$ such that $a_1 = x_1$, $a_n = y_1$, $b_1 = x_2$ and $b_n = y_2$ Let $a_i \neq a_{i+1}$ at $v_i = (l_i, m_i)$ and $b_i \neq b_{i+1}$ at $w_i = (n_i, k_i)$for all i. Then,

$$\frac{\Delta(A,x)}{\Delta(A,y)} = \prod_{i=1}^{m-1} \frac{\delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_i)|_{D_2})}{\delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_{i+1})|_{D_2})} = \prod_{i=1}^{n-1} \frac{\delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_i)|_{D_2})}{\delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_{i+1})|_{D_2})}$$

taking logarithm on both sides we get,

$$\begin{aligned} & \textstyle\sum_{i=1}^{m-1} (\log \delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_i)|_{D_2}) - \log \delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_{i+1})|_{D_2})) \\ = \quad & \textstyle\sum_{i=1}^{n-1} (\log \delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_i)|_{D_2}) - \log \delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_{i+1})|_{D_2})) \end{aligned}$$

2. Suppose $B \in D$. Let $x, y \in \mathcal{B}_{B\cup\partial B}(X)$ be distinct such that $x = y$ on ∂B. Consider pivots $a = (a_1, a_2 \dots a_m)$ and $b = (b_1, b_2 \dots b_n)$ such that $a_1 = b_1 = x$ and $a_2 = b_2 = y$. Let $a_i \neq a_{i+1}$ at $v_i = (l_i, m_i)$ and $b_i \neq b_{i+1}$ at $w_i = (n_i, k_i)$ for all i. Then,

$$\frac{\Delta(B,x)}{\Delta(B,y)} = \prod_{i=1}^{m-1} \frac{\delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_i)|_{D_2})}{\delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_{i+1})|_{D_2})} = \prod_{i=1}^{n-1} \frac{\delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_i)|_{D_2})}{\delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_{i+1})|_{D_2})}$$

Taking logarithm on both sides we get

$$\begin{aligned} & \textstyle\sum_{i=1}^{m-1} (\log \delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_i)|_{D_2}) - \log \delta(\sigma_1^{l_i}\sigma_2^{m_i}(a_{i+1})|_{D_2})) \\ = \quad & \textstyle\sum_{i=1}^{n-1} (\log \delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_i)|_{D_2}) - \log \delta(\sigma_1^{n_i}\sigma_2^{k_i}(b_{i+1})|_{D_2})) \end{aligned}$$

Note that these relationships are dependent on the shift space X rather than on the measure μ. Also the constraints of type (2) can be looked at as a special case of constraints of type (1). However checking constraints of type (1) presents an extra level of difficulty.

Suppose a shift space X is fixed. Let

$$U = \{\mu \mid \mu \text{ is a stationary Markov random field and } supp(\mu) = X\}$$

Given a probability measure $\mu \in U$, let δ be its infralocal specification. Consider the function

$$\delta_\mu : \{(x,y) \mid x,y \in \mathcal{B}_{D_2}(X) \text{ and } x = y \text{ on } \partial\{(0,0)\}\} \longrightarrow \mathbb{R}$$

given by

$$\delta_\mu(x,y) = \log(\delta(x)) - \log(\delta(y))$$

Then the set $\{\delta_\mu \mid \mu \in U\}$ is contained in a finite dimensional vector space satisfying equations of type (1) and (2).

Question 4 Given a shift space X, is there a finite procedure to determine the linear space generated by equations of type (1) and (2)?

These constraints depend on how much can configurations in X can pivot. For example, if X is the full shift then pivoting is very easy. Hence the infralocal specification of a stationary Markov random field supported on it is very constrained. However in X_3, the pivoting is very constrained and hence there are no relations arising of type (1) and (2).

To be more formal, suppose $x = (x_1, x_2, x_3 \ldots x_m)$ is an D_n- pivot in X_3 for some m. The following proposition will establish a consistency result between various pivots between two given configurations.

Proposition 6.3.2. *Let $x = (x_1, x_2 \ldots x_n)$ and $y = (y_1, x_2 \ldots y_r)$ be a D_m pivot for some m and $x_1 = y_1$ and $x_n = y_r$. Then,*

$$a_x = a_y, b_x = b_y \text{ and } c_x = c_y$$

This rules out any constraints of type (2). The following is a stronger version of the proposition given above. It rules out any constraints of type (1).

Proposition 6.3.3. *Let $x = (x_1, x_2 \dots x_n)$ and $y = (y_1, x_2 \dots y_r)$ be a D_m pivot for some m. Suppose there exists a set $A \subset D_m$ such that $x_1 = y_1$ and $x_n = y_r$ on $A \cup \partial A$. Also $x_1 = x_n$ and $y_1 = y_r$ on $D_m - A$. Then,*

$$a_x = a_y, b_x = b_y \textit{ and } c_x = c_y$$

Thus there are no constraints on v_1, v_2 and v_3 of type (1) and (2).

Question 5 Let D' be the set of all finite subsets of $\mathbb{Z}^2$. Let v_1, v_2 and $v_3 \in \mathbb{R}$. Then, using the expression in Equation 6.3.1, a function

$$\Delta : \{A \times \mathcal{B}_{A \cup \partial A}(X_3) \mid A \in D'\} \longrightarrow [0,1]$$

can be constructed such that for any finite $A \in \mathbb{Z}^2$ and $y \in \mathcal{B}_{\partial A}(X_3)$

$$\sum_{x \in \mathcal{B}_{A \cup \partial A}(X_3) \mid x = y \text{ on } \partial A} \Delta(A, x) = 1$$

Does there exist a stationary Markov random field μ such that $supp(\mu) = X_3$ and Δ is its specification? We are close to proving that there exist a stationary Markov random field μ such that $supp(\mu) = X_3$ for v_1, v_2 and $v_3 = 0$.

Chapter 7

A Linear Algebra Theorem

Typically linear algebra deals with finite dimensional vector spaces. However, in this thesis we have encountered the use of matrices with countably many rows and columns. This chapter is devoted to clarify the notions. It might be possible that functional analysis is a much better basis for these things. However, since the concepts are so simple, we shall keep it to this setting.

Given arbitrary non-empty sets M and N, a *matrix* of the size $M \times N$ is a function $A : M \times N \longrightarrow \mathbb{R}$ such that $\{y|A(x,y) \neq 0\}$ is finite for all $x \in M$. This is what is meant by having a matrix such that the rows are indexed by elements of M and the columns are indexed by elements of N. Suppose we are given A and B matrices of size $M \times N$ and $N \times K$ respectively. Then AB is a matrix of size $M \times K$ such that

$$AB(m,k) = \sum_{n \in N} A(m,n)B(n,k)$$

for all $m \in M$ and $k \in K$ whenever the sums are well defined.

Theorem. *Suppose M and N are countable sets and K is singleton. Let A be a matrix of size $M \times N$ and b a matrix of size $M \times K$. Then there exists a matrix x of size $N \times K$ such that $Ax = b$ if and only if*

$$vA = 0 \implies vb = 0$$

where v is a matrix of size $K \times M$ with finitely many non- zero entries and 0 is a *matrix of size $K \times N$ such that $0(a,b) = 0$ for all $a \in K$ and $b \in N$*

If M is finite, this is a well known linear algebra fact. One can for instance look in [9]. We shall assume this and prove the theorem. The idea is to take the solutions of finite parts and then stitch them together to get the entire solution. The theorem holds even if restrictions on the cardinality of M, N and K are removed.

Proof. Suppose A, b as stated in the theorem. Consider an increasing sequence of finite subsets $\{M_i\}$ of M such that

$$\bigcup_i M_i = M$$

Let A_i be a submatrix of A of size $M_i \times N$ and b_i be a submatrix of b of size $M_i \times K$ given by

$$A_i = A|_{M_i \times N}$$
$$b_i = b|_{M_i \times N}$$

Since M_i is finite and

$$vA_i = 0 \implies vb_i = 0,$$

for every i there exists a solution to the equation $A_i x = b_i$. Take an increasing sequence of finite subsets $\{N_j\}$ of N such that

$$\bigcup_j N_j = N$$

Let

$$S_j^i = \{x \in \mathbb{R}^{N_j \times K} | \text{ there exists } y \in \mathbb{R}^{N \times K} \text{ such that } A_i y = b_i \text{ and } y|_{N_j \times K} = x\}$$

be the projection of the solution space onto the N_jth coordinates.

Note that, each S_j^i is non-empty, $S_j^{i+1} \subset S_j^i$ and the projection of S_{j+1}^i onto the N_jth coordinates gives us S_j^i.

$$S_j^i = \{x \in \mathbb{R}^{N_j \times K} \mid \text{ there exists } y \in S_{j+1}^i \text{ such that } y|_{N_j \times K} = x\} \tag{7.0.1}$$

Therefore $\{S_j^i\}_{i=1}^{\infty}$ is a nested sequence of affine subspaces of a finite dimensional vector space. Hence the dimension of S_j^i is eventually a constant. Therefore for any j, $S_j^i = S_j^{i+1}$ for a large enough i. In particular, $\bigcap\limits_{i\in\mathbb{N}} S_j^i$ is non-empty for all j. We will now prove some consistency result of these solution sets by induction.

Lemma 7.0.4. *There exists a sequence $\{x_t\}_{t\in\mathbb{N}}$ such that $x_t \in \bigcap\limits_{i\in\mathbb{N}} S_t^i$ and $x_t|_{N_{t-1}\times K} = x_{t-1}$ for all t.*

Proof. Let $x_1 \in \bigcap\limits_{i\in\mathbb{N}} S_1^i$ be chosen. By equation (7.0.1), for every i there exists $x_2^i \in S_2^i$ such that $x_2^i|_{N_1\times K} = x_1$. Since the sequence $\{S_2^i\}_{i\in\mathbb{N}}$ is eventually constant there exists $x_2 \in \bigcap\limits_{i\in\mathbb{N}} S_2^i$ such that $x_2|_{N_1\times K} = x_1$. Let $x_1, x_2 \ldots x_r$ be chosen such that $x_t \in \bigcap\limits_{i\in\mathbb{N}} S_t^i$ for $1 \le t \le r$ and $x_t|_{N_{t-1}\times K} = x_{t-1}$ for $2 \le t \le r$. For every i, there exists $x_{r+1}^i \in S_{r+1}^i$ such that $x_{r+1}^i|_{N_r\times K} = x_r$ Since the sequence $\{S_{r+1}^i\}_{i\in\mathbb{N}}$ is eventually a constant there exists $x_{r+1} \in \bigcap\limits_{i\in\mathbb{N}} S_i^{r+1}$ such that $x_{r+1}|_{N_r\times K} = x_r$. Thus the induction is complete and the lemma is proved. □

Let a sequence $\{x_t\}_{t\in\mathbb{N}}$ be as in the lemma before. Let $x \in \mathbb{R}^{N\times K}$ such that $x|_{N_r\times K} = x_r$ for all $r \in \mathbb{N}$. Take an arbitrary $i \in \mathbb{N}$. Since every row has finitely many non-zero entries for a given $z \in \mathbb{R}^{N\times K}$, $A_i z$ is independent of $z|_{(N_r)^c\times K}$ for large enough r. Choose such a $r = r_0$. We have $x_{r_0} \in \bigcap\limits_{l\in\mathbb{N}} S_l^{r_0} \subset S_i^{r_0}$. By the definition of $S_{r_0}^i$, there exists $x_{r_0}^i$ such that $A_i x_{r_0}^i = b_i$ and $x_{r_0}^i|_{N_{r_0}} = x_{r_0}$. Therefore, $A_i x_{r_0}^i = A_i x = b_i$. Since i was arbitrary, $Ax = b$. □

Bibliography

[1] Drton, Mathias; Sturmfels, Bernd; Sullivant, Seth. Lectures on algebraic statistics *Oberwolfach Seminars, Birkhuser Verlag, Basel*, 39, 2009.

[2] J. M. Hammersley; P. Clifford. Markov fields on finite graphs and lattices. 1971.

[3] R.L. Dobruschin. Description of a random field by means of conditional probabilities and conditions for its regularity. *Teor. Verojatnost. i Primenen*, 13:201–229, 1968.

[4] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge,, 4 edition, 2010.

[5] Kari Eloranta. Diamond ice. *J. Statist. Phys.*, 96(5-6):1091–1109, 1999.

[6] F.R.Gantmacher. *The Theory of Matrices*, volume 2. AMS Chelsea Publishing, Providence, RI, 1998.

[7] Hans-Otto Georgii. *Gibbs measures and phase transitions*. de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1988.

[8] Paul R. Halmos. *Lectures in Ergodic Theory*. Chelsea Publishing Co., New York, 1960.

[9] Kenneth Hoffman; Ray Kunze. *Linear Algebra*. Prentice-Hall Mathematics Series Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961.

[10] Douglas Lind; Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge., 1995.

[11] John Moussouris. Gibbs and markov random systems with constraints. *Journal of Statistical Physics*, 10(1), 1974.

[12] William Parry. Intrinsic markov chains. *Trans. Amer. Math. Soc*, 112:55–66, 1964.