AN ABSTRACT OF THE THESIS OF

Clemens H. Glaffig for the degree of Master of Science in Mathematics presented on July 27, 1984.

Title: Gibbs States and Correlation

Redacted for Privacy

Abstract approved:

Edward C. Waymire

Certain important concepts from the theory of Gibbs states are first described in the simple setting of the finite volume case. With the extension to the infinite volume case, Gibbs states are defined, exhibiting two different approaches to the subject. The general structure of the set of Gibbs states is investigated, emphasizing the significant role of pure Gibbs states.

In a further restriction in physical assumptions, the relation of Gibbs states to Markov random fields is explored and this relation is used to detect infinite divisibility within the class of Gibbs states. For the case of infinitely divisible states on the Bethe lattice, the Levy-Khintchin representation of the correlation is used to prove these states to be extreme. Extreme states are characterized by their correlation function.

As a closure of this treatise, another approach to the theory, aimed at obtaining a different characterization of pure states, is introduced.

Gibbs States and Correlation

by

Clemens H. Glaffig

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed July 27, 1984

Commencement June 1985

APPROVED:

Redacted for Privacy

Assistant Professor of Mathematics in charge of major

Redacted for Privacy

Chairman of department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented July 27, 1984

Typed by Clover Redfern for Clemens H. Glaffig

ACKNOWLEDGEMENT

I wish to express my gratitude to Dr. E. Waymire for his assistance, guidance and encouragement during my stay in Corvallis.

TABLE OF CONTENTS

GIBBS STATES AND CORRELATION

Chapter 1. INTRODUCTION

During the last centuries, physics has been the most inspiring nonmathematical field for mathematics. One of the recent fields of interest of mathematicians in physics is statistical mechanics. The theory of random fields is a subject which arose from this, starting with the description of the complicated subject of ferromagnetism.

It is a known fact from quantum mechanics that every electron possesses an angular momentum, called spin, and associated with it a magnetic moment or magnetic spin. Furthermore additional magnetic moments occur, resulting from the rotation of electrons around protons. By this, an atom can be thought of as a little magnet, called an elementary magnet, and thus matter can be viewed as a collection of elementary magnets. These elementary magnets influence each other by speeding neighboring elements up or down in their rotation, changing their spins. This so called interaction results in most cases in a zero overall magnetic moment. Ferromagnetic matter has the property that the elementary magnets tend to line up, i.e. tend to have the same spin direction over certain local domains.

An interesting phenomenon occurs at a matter dependent temperature, called critical temperature. Above this temperature, the tendency to line up vanishes. Another form of realizing this critical temperature is obtained by the fact that ferromagnets may have a non zero magnetic moment, even though the external field is zero, which is not observed above the critical temperature. Much of the aim of modern mathematical statistical physics is to give precise theoretical descriptions of this phenomena. The physicist E. Ising presented in the early twenties of this century a model to describe this subject on a lattice, the set of sites of the lattice representing elementary magnets, by associating to each spin configuration on the lattice a certain probability of occurrence, governed by interaction, if the system is in equilibrium. The corresponding probability measures on the set of all spin configurations are called Gibbs states. General properties of Gibbs states are known for countable sets of sites without any structure. The predominant aim of investigating Gibbs states is to detect the different possibilities of magnetization for zero external field below some critical temperature, referred to as phase transition.

In this connection correlations between sets of spin variables play a basic role. So the major focus of this thesis is on correlation formulae for Gibbs states. The results are based on a method of Waymire, which is previously known to work for one-dimensional Ising models (see [18]). The method is explored for higher dimensional lattics configurations here.

Chapter 2. GIBBS STATES ON COUNTABLE SETS (THE FINITE VOLUME CASE)

Let S be a finite set of sites and let W be the set $\{1,-1\}$, representing the spin directions up and down. With the discrete topology and the usual multiplication W becomes a compact topological abelian group.

Let

$$\Omega = W^S = \{1,-1\}^S$$

be the finite topological product. Ω is called configuration space, the elements of Ω are called configurations. In the case of finite S all topological and measure theoretical considerations are entirely trivial. However these notions are essential to the case when S is infinite and so they will be brought out here for pedagogical reasons.

In this case Ω is trivially compact since the product topology on Ω is the discrete topology, generated by singleton sets of the form

$$\{\varepsilon_n : n \in S\} ,$$

where each $\varepsilon_n \in \{1,-1\}$.

Defining a group operation on Ω by coordinatewise multiplication, Ω becomes also a compact abelian topological group.

The topology on Ω can also be characterized to be the smallest topology on Ω for which all coordinate projections are continuous, i.e. if (X_n), $n \in S$ denotes the family of coordinatewise projections on Ω, $X_n \in C(\Omega)$ $\forall$ n.

Since

$$\underline{\bar{X}}:\Omega \to \Omega, \quad \underline{\bar{X}} := (X_n), \quad n \in S$$

is the identity map, $\underline{\bar{X}}$ is used to denote a specific configuration in Ω or a random variable as seen later.

The sites represent elementary magnets, which influence other sites, called interaction. A configuration then is a certain state of a ferromagnetic matter specifying the spin direction for every elementary magnet.

From now on only pairwise interaction will be considered. The interaction energy, I, between two sites n,m may then be represented by

$$I_{n,m}(\underline{\bar{X}}) = J_{n,m}X_nX_m \quad ,$$

where $J_{n,m}$ is the coupling constant between these sites.

The influence of a homogeneous external field H on the individual sites can be represented by

$$HX_n, \quad n \in S.$$

The total energy of a configuration in Ω becomes then

$$U(\underline{\bar{X}}) = -\beta \sum_{n,m \in S} J_{n,m}X_nX_m - \beta H \sum_{n \in S} X_n \quad , \tag{2.1}$$

where the first sum counts each pair only once, $\beta = 1/kT$, k = Boltzman's constant, T = absolute temperature.

Let A be the Borel-algebra of Ω. Since Ω is finite A coincides with the power set of Ω.

Definition 2-1: A finite dimensional Gibbs state is a probability measure on Ω of the form

$$P(\underline{\bar{X}}) = Z^{-1} \exp\{-U(\underline{\bar{X}})\} \quad , \tag{2.2}$$

where

$$Z = \sum_{\underline{\bar{X}} \in \Omega} \exp\{-U(\underline{\bar{X}})\}$$

is called partition function and normalizes (2.2) to a probability measure.

Since Ω is finite, P is specified in terms of its density with respect to counting measure, i.e.

$$\frac{dP}{dm}(\underline{\bar{X}}) = Z^{-1} \exp\{-U(\underline{\bar{X}})\} \quad ,$$

where $\frac{dP}{dm}$ is the Radon-Nikodym derivative.

(2.2) factors into

$$P(\underline{\bar{X}}) = Z^{-1} \exp\{-\beta \sum_{n,m \in S} J_{n,m} X_n X_m\} \prod_{n \in S} \exp\{-\beta H X_n\}$$

$$= Z^{-1} \exp\{-\beta I(\underline{\bar{X}})\} \prod_{n \in S} \exp\{-\beta H X_n\} \quad , \tag{2.3}$$

where $I(\underline{\bar{X}}) = \sum_{n,m \in S} I_{n,m}(\underline{\bar{X}})$ represents the total interaction potential.

Let a probability measure μ on Ω be defined by its projected measures:

$$\mu_n := X_n \mu .$$

μ_n is the probability measure on W with

$$\mu_n(\{-1\}) = \frac{\exp \beta H}{2 \cosh \beta H}, \quad \mu_n(\{1\}) = \frac{\exp\{-\beta H\}}{2 \cosh \beta H} .$$

μ is then the product measure of the μ_n's, $n \in S$.

(2.3) then becomes

$$P(\bar{\underline{X}}) = Z^{-1} \exp\{-\beta I(\bar{\underline{X}})\} \prod_{n \in S} \{\mu_n(X_n) 2 \cosh \beta H\}$$

$$= Z^{-1} \exp\{-\beta I(\bar{\underline{X}})\} \mu(\bar{\underline{X}}) 2^{|S|} (\cosh \beta H)^{|S|}$$

$$= Z_0^{-1} \exp\{-\beta I(\bar{\underline{X}})\} \mu(\bar{\underline{X}}) , \tag{2.4}$$

where

$$Z_0^{-1} = Z^{-1} 2^{|S|} (\cosh \beta H)^{|S|} \tag{2.5}$$

Note: $P \ll \mu \ll m$, where m is the counting measure on Ω and

$$\frac{dP}{d\mu}(\bar{\underline{X}}) = Z_0^{-1} \exp\{-\beta I(\bar{\underline{X}})\}$$

$$\frac{d\mu}{dm}(\bar{\underline{X}}) = \mu(\bar{\underline{X}}) .$$

Substituting this into (2.4) yields

$$P(\bar{\underline{X}}) = \frac{dP}{dm}(\bar{\underline{X}}) = \frac{dP}{d\mu}(\bar{\underline{X}}) \frac{d\mu}{dm}(\bar{\underline{X}}) .$$

By (2.2) and (2.4) follows

$$\mu(\underline{\bar{X}}) = \frac{Z_0}{Z} \cdot \exp\{-\beta H \sum_{n \in S} X_n\}$$

and thus

$$\frac{dP}{dm}(\underline{\bar{X}}) = \frac{dP}{d\mu}(\underline{\bar{X}}) \cdot \frac{Z_0}{Z} \exp\{-\beta H \sum_{n \in S} X_n\} ,$$

which is equivalent to (2.5).

Summing now (2.4) over all $\underline{\bar{X}} \in \Omega$ yields:

$$\sum_{\underline{\bar{X}} \in \Omega} P(\underline{\bar{X}}) = 1 - Z_0^{-1} \sum_{\underline{\bar{X}} \in \Omega} \exp\{-\beta I(\underline{\bar{X}})\}\mu(\underline{\bar{X}})$$

$$\Longleftrightarrow \quad Z_0 = \sum_{\underline{\bar{X}} \in \Omega} \exp\{-\beta I(\underline{\bar{X}})\}\mu(\underline{\bar{X}}) = E_\mu \exp\{-\beta I(\underline{\bar{X}})\} . \tag{2.6}$$

By (2.6) Z_0, as a function of β, can be interpreted as the moment generating function of the total interaction energy I.

2-1. The Use of the Partition Function

In (2.6) the partition function was probabilistically interpreted in terms of the total interaction energy. On the other hand the partition function is useful to express magnetization and susceptibility, indicating it to be more than just a normalization constant. In particular $Z(\beta,(J_{n,m}),H)$ is a physically meaningful quantity.

The magnetization of a configuration $\underline{\bar{X}} \in \Omega$ is defined by

$$M_S(\beta,H) := \sum_{n \in S} X_n = \{\text{\# of 1's}\} - \{\text{\# of -1's}\}$$

From (2.2), noting that Z is analytic as a function of β and H:

$$\frac{\partial}{\partial H} \ln Z(\beta,H) = Z^{-1} \frac{\partial}{\partial H} \sum_{\underline{\bar{X}}\in\Omega} \exp\{-\beta \sum_{n,m\in S} J_{n,m}X_nX_m - \beta H \sum_{n\in S} X_n\}$$

$$= \sum_{\underline{\bar{X}}\in\Omega} (-\sum_{n\in S} \beta X_n) Z^{-1} \exp\{-\beta \sum_{n,m\in S} J_{n,m}X_nX_m - \beta H \sum_{n\in S} X_n\}$$

$$= -\beta \sum_{\underline{\bar{X}}\in\Omega} \sum_{n\in S} X_n P(\underline{\bar{X}})$$

$$= -\beta E(\sum_{n\in S} X_n) = -\beta EM_S$$

or

$$EM_S(\beta,H) = -\frac{1}{\beta}\frac{\partial}{\partial H} \ln Z(\beta,H) \ , \tag{2.7}$$

where $-\frac{1}{\beta} \ln Z(\beta,H)$ is called the free energy and $M_S(\beta,H)$ is considered as a random variable.

Magnetic susceptibility $\chi_S(\beta,H)$ is defined as the rate of change of the expected magnetization resulting from an external field H

$$\chi_S(\beta,H) = \frac{\partial}{\partial H} EM_S(\beta,H) \ .$$

By (2.7):

$$\chi_S(\beta,H) = -\frac{1}{\beta}\frac{\partial^2}{\partial H^2} \ln Z(\beta,H) \ . \tag{2.8}$$

Also from (2.2):

$$\frac{\partial}{\partial \beta} \ln Z = Z^{-1} \frac{\partial}{\partial \beta} \sum_{\underline{\bar{X}}\in\Omega} \exp\{-\beta I(\underline{\bar{X}}) - \beta HM_S\}$$

$$= -\sum_{\underline{\bar{X}}\in\Omega} \{I(\underline{\bar{X}}+HM_S\} \cdot P(\underline{\bar{X}})$$

$$= -EI(\underline{\bar{X}}) - HEM_S \tag{2.9}$$

where now also $I(\bar{\underline{X}})$ is considered as a random variable.

2-2. Spontaneous Magnetization and Boundary Specifications

It can be seen from (2.7) that $EM_S(\beta,H)$ is continuous in H. For $H = 0$, $EX_n = 0$, $\forall\, n \in S$ since $P(\bar{\underline{X}}) = P(-\bar{\underline{X}})$. Therefore $EM_S(\beta,0) = 0$, again by (2.7). By continuity now follows:

$$\lim_{H\to 0} EM_S(\beta,H) = 0 \tag{2.10}$$

It is a known experimental fact that for sufficiently low temperatures, i.e. high β, spontaneous magnetization occurs for $H = 0$. Furthermore this spontaneous magnetization is known to obtain two distinct values corresponding to limits $H \to 0^+$, $H \to 0^-$, i.e. $\exists\, \mu^+(\beta)$, $\mu^-(\beta)$ such that

$$\lim_{H\to 0^+} EM_S(\beta,H) = \mu^+(\beta)$$

$$\lim_{H\to 0^-} EM_S(\beta,H) = \mu^-(\beta)$$

By (2.10) spontaneous magnetization cannot be observed in the finite dimensional model.

It seems to be natural then to extend the finite model by assuming $S \subset Z^n$ for some n and letting $|S| \to \infty$ in the sense of van Hove (see [7] p. 12). This can be achieved in several ways by specifying the boundary of S such that in the limit spontaneous magnetization occurs. In this case phase transition is said to occur. Now let S be a finite cube in Z^n. The two major cases to specify the boundary, i.e. the sites at the edge of the cube, are

(a) Fixed Boundary: All sites outside of S are fixed to a certain spin direction.

(b) Free Boundary: The boundary sites are treated as any other site.

2-3. The Roles of the LLN and the CLT

(a) Assume the case of infinite temperature, i.e. $\beta = 0$. From (2.2) follows for specified $\bar{\underline{X}} \in \Omega$

$$P(\bar{\underline{X}}) = \frac{1}{|\Omega|} = (\frac{1}{2})^{|S|} = \prod_{n \in S} \frac{1}{2}$$

and therefore P is the product measure of Bernoulli distributions $B(\frac{1}{2})$ on $X_n(\Omega) = W$. Since P represents the joint distribution of the (X_n)'s, viewed as random variables, the family of projections is an i.i.d. family of random variables. In this case $EX_n = 0$, Var $X_n = 1$.

Both the LLN and the CLT apply:

LLN: $$\lim_{|S| \to \infty} \frac{1}{|S|} \sum_{n \in S} X_n = \lim_{|S| \to \infty} \frac{1}{|S|} M_S = 0$$

with probability 1, where $|S| \to \infty$ again in the sense of van Hove.

This means that the average overall magnetization of the model is 0.

CLT: $\lim_{|S|\to\infty} \sqrt{|S|}M_S \to \text{Normal } (0,1)$

in distribution, i.e. M_S is for large $|S|$, after scaling, approximately normal distributed.

Empirical observations are quoted below to show that the effective absence of interaction at infinite temperature makes these results inapplicable at low temperatures.

(b) Assume now $\beta > 0$, $H = 0$ and S to be a cube in Z^n. Fixing the boundary to 1 and letting $|S| \to \infty$ yields a probability measure P^+ on W^{Z^n} and a probability measure P^- for -1 boundary. In the case of $n = 2$, i.e. $S \to Z^2$, it comes out that for a specific β_c, for $0 < \beta < \beta_c$, P^+ and P^- coincide, for $\beta > \beta_c$ they are different, reflecting experimental results. Therefore it is expected that the LLN and the CLT fail at low temperatures, moreover the distribution of the fluctuations of $M_S(\beta,H)$ in the limit should be a linear combination of bell shaped distributions about μ^+, μ^-.

(c) Assume now $H \neq 0$. Again by experimental results one expects that for the LLN

$$\lim_{|S|\to\infty} M_S(\beta,H) = \mu(\beta,H)$$

where

$$\lim_{H\to 0^+} \mu(\beta,H) \neq \lim_{H\to 0^-} \mu(\beta,H) \quad \text{if} \quad \beta > \beta_c$$

for a specified β_c and that the LLN holds for $0 < \beta < \beta_c$.

Chapter 3. GIBBS STATES ON COUNTABLE SETS (THE INFINITE VOLUME CASE)

In order to get a reasonable model to describe the physical subject of ferromagnetism it is necessary to extend the theory to the case of an infinite set of sites. For reasons discussed in (2) only then phenomena like phase transition can be observed. However, as will be seen, the properties of Gibbs states depend on the structure of the underlying set of sites.

3-1. Introduction and General Notations

Let the set of sites S now be a countable infinite set and let $W = \{1,-1\}$ again be the compact abelian group under multiplication and with discrete topology.

Define

$$\Omega := \{1,-1\}^S$$

as the set of all $\{1,-1\}$ valued functions on S.

Let (X_n), $n \in S$ again be the coordinate projections

$$X_n : \Omega \to W_n, \quad W_n = W \quad \forall \quad n \in S$$

$\bar{\underline{X}} \in \Omega$ will then again be denoted as

$$\bar{\underline{X}} = (X_n), \quad n \in S$$

Ω may also be viewed as the power set of S, where each $\bar{\underline{X}} \in \Omega$ is associated with a subset A of S by

$$\bar{\underline{X}} \sim A \quad \text{if} \quad A = \{n \in S,\ X_n = -1\}$$

Let $\mathcal{B}(S)$ be the set of finite subsets of S. Since Ω is a product of topological spaces, a topology on Ω is defined by the product topology, i.e. the smallest topology such that all X_n are continuous.

Then Ω has the following properties:

(i) Ω is compact, by Tychonoff.

(ii) Ω is metrizable: Since W is metrizable with the trivial metric and countable products of metric spaces are metrizable, Ω is metrizable.

A metric on Ω can be defined in the usual way:

$$d(\bar{\underline{X}},\bar{\underline{Y}}) := \sum_{n\in S} \frac{d_n(X_n,Y_n)}{2^n} \tag{3.1}$$

where $\bar{\underline{X}}$, $\bar{\underline{Y}} \in \Omega$, d_n the trivial metric on W and S is labelled (i.e. ordered and numbered). This metric induces the product topology.

Note: Since Ω is compact and satisfies the second axiom of countability it also follows by a lemma of Urysohn that Ω is metrizable. This follows also by a theorem of Bing-Nagata and Smirnow, who construct a metric in their proof which comes out to be equivalent to (3.1).

(iii) $\bar{\mathcal{C}} = C(\Omega)$ where $\mathcal{C}$ is the set of all real valued functions depending on at most finitely many coordinates and $\bar{\mathcal{C}}$ denotes closure in the sup-norm.

To see this let $f \in \mathcal{C}$. Then $\exists\, A \in \mathcal{B}(S)$ such that f depends only on coordinates in A and hence f can be viewed as a function on W^A.

Note: The induced topology on W^A is discrete since $|W^A| < \infty$.

If B is open in $\mathbb{R}$, $f^{-1}(B) \subset W^A$ and therefore open. So

$$\mathcal{C} \subset C(\Omega)$$

$\mathcal{C}$ now forms an algebra, containing the constants and separating points (take e.g. coordinate projections). With these properties it follows from the Stone-Weierstrass theorem that $\mathcal{C}$ is dense in $C(\Omega)$.

(iv) The finite dimensional cylinders form a basis for the product topology on Ω, where a finite dimensional cylinder $[A,F]$ is defined by

$$[A,F] := \{\underline{\bar{X}}\in\Omega,\ X_n=-1,\ n\in A, X_n=1,\ n\in F\backslash A\} \subset \Omega,$$

$$F \in \mathcal{B}(S),\ A \subset F$$

The set of finite intersections of the form

$$\bigcap_{n\in F} X_n^{-1}(\varepsilon_n) \quad \text{for some} \quad F \in \mathcal{B}(S)$$

where each $\varepsilon_n \in \{1,-1\}$, forms a basis for the product topology. Let B be an element of this basis

$$B = \bigcap_{n\in F} X_n^{-1}(\varepsilon_n) \quad \text{for fixed} \quad (\varepsilon_n),\ n \in F \in \mathcal{B}(S)$$

$$A := \{n\in F,\ \varepsilon_n=-1\}$$

then

$$B = \{\underline{\bar{X}}\in\Omega,\ X_n=\varepsilon_n,\ n\in F\} = \{\underline{\bar{X}}\in\Omega,\ X_n=-1,\ n\in A,\ X_n=1,\ n\in F\backslash A\}$$

$$= [A,F]$$

Now let $\mathcal{A}$ denote the Borel algebra on Ω, i.e. the σ-algebra generated by open sets. Each X_n is $\mathcal{A}$-measurable, so the site variables X_n, $n \in S$ may be regarded as random variables on $(\Omega,\mathcal{A})$.

<u>Definition 3-1</u>: A probability measure on $(\Omega,\mathcal{A})$ is called a random field on $(\Omega,\mathcal{A})$.

<u>Example 3-1</u>: $S = Z$, $\Omega = \{1,-1\}^Z$. A random field on Ω is a discrete time stochastic process with state space $\{1,-1\}$.

<u>Remark 3-1</u>: Ω is also a Polish space and a random field P is, as the distribution of the coordinate projections, a projective family of probability measures, i.e. $X_{F_1} P_{F_2} = P_{F_1}$, where $F_1 \subset F_2 \in \mathcal{B}(S)$ and P_{F_i} is the distribution of the coordinate projections onto sites in F_i, indexed by $\mathcal{B}(S)$. Therefore, by a theorem of Kolmogoroff (see [1] p. 347), it is enough to specify $(P_F), F \in \mathcal{B}(S)$ to construct P.

<u>Example 3-2</u>: Ferromagnet with infinite temperature;

(a) Without external field

$$P([A,F]) = (\tfrac{1}{2})^{|F|} \qquad F \in \mathcal{B}(S)$$

(b) With external field

$$P(X_n=1) = p \quad \forall\, n \in S, \quad 0 < p < 1$$

$$P([A,F]) = p^{|F\setminus A|}(1-p)^{|A|}$$

3-2. Definition of Gibbs States

For the finite dimensional case it was possible to define a Gibbs state by its probability for every configuration, i.e. probability mass function. This is not possible for the infinite volume case. The number of configurations is the same as the number of binary codes of numbers in $[0,1]$; every configuration represents a code. Therefore Ω has the same cardinality as the continuum. However it is possible to derive properties of Gibbs states in the finite case such that these properties are preserved by taking the thermodynamic limit, i.e. letting S tend to a countable infinite set, in a sense to be made precise. The following two different approaches define an infinite volume Gibbs state in an equivalent way. The equivalence is established in Theorem 3-1 below.

3-2.1. General Notations

A system of functions $\Phi := \{\Phi_F : \Omega_F \to \mathbb{R},\ F \in \mathcal{B}(S)$ where $\Omega_F = W^F$. is called an interaction potential on Ω if:

(i) $\Phi(\emptyset) = 0$

(ii) $\|\Phi\| = \sup\limits_{i\in S} \sum\limits_{A:i\in A} \sup\{|\Phi_A(\underline{\bar{X}})| : \underline{\bar{X}} \in \Omega_A\} < \infty$

Two sites $n,m \in S$ are said to interact if $\exists F \in \mathcal{B}(S)$ such that $n,m \in F$ and $\Phi_F \not\equiv 0$.

Definitions 3-2:

(a) Φ is said to have finite range if every site $n \in S$ interacts only with finitely many $m \in S$.

(b) Φ is called H-invariant if it is invariant under a family H of bijections on S.

(c) Φ is called a pair potential if $\Phi_F \not\equiv 0 \iff F = \{n,m\}$ or $F = \{n\}$; $n,m \in S$ where $\Phi_{\{n\}}$ will usually result from an external field. Φ is called a symmetric pair potential if in addition: Φ is invariant under $\phi: S \to S$,

$$\phi(n) = m, \quad \phi(m) = n, \quad \phi(k) = k, \quad k \neq m,n.$$

From now on consider Φ to be a symmetric, invariant pair potential with

$$\Phi_F(\underline{\bar{X}}) = \begin{cases} -JX_n X_m & \text{if} \quad F = \{n,m\} \\ -HX_n & \text{if} \quad F = \{n\} \\ 0 & \text{otherwise,} \end{cases}$$

where H is due to an external field.

Given a potential Φ, an energy function is defined for each $F \in \mathcal{B}(S)$ by

$$U_F^{\Phi}(\underline{\bar{X}}) = \beta \sum_{A \cap F \neq \emptyset} \Phi_A(\underline{\bar{X}}). \tag{3.1}$$

3-2.2. Gibbs States in the Sense of Dobrushin

Consider a sequence of finite sets (S_i), $i \in \mathbb{N}$ such that $S_i \subsetneq S_{i+1}$ $\forall$ i and $S = \bigcup_{i=1}^{\infty} S_i$.

Let for fixed i, $F_i \subset S_i$ and P_i be a Gibbs state on W^{S_i}.

$$\frac{P_i(X_n=\varepsilon_n, n\in S_i)}{P_i(X_n=\varepsilon_n, n\in S_i\backslash F_i)} = \frac{\exp\left\{-\beta J \sum\limits_{n,m\in S_i} \varepsilon_n\varepsilon_m - \beta H \sum\limits_{n\in S_i} \varepsilon_n\right\}}{\sum\limits_{\substack{\varepsilon_n,\varepsilon_m \\ n,m\in F_i}} \exp\left\{-\beta J \sum\limits_{n,m\in S_i} \varepsilon_n\varepsilon_m - \beta H \sum\limits_{n\in S_i} \varepsilon_n\right\}}$$

$$= \frac{\exp\left\{-\beta J \sum\limits_{n\cup m\in F_i} \varepsilon_n\varepsilon_m - \beta H \sum\limits_{n\in F_i} \varepsilon_n\right\}}{\sum\limits_{\substack{\varepsilon_n,\varepsilon_m \\ n,m\in F_i}} \exp\left\{-\beta J \sum\limits_{n\cup m\in F_i} \varepsilon_n\varepsilon_m - \beta H \sum\limits_{n\in F_i} \varepsilon_n\right\}}$$

$$=: P^{\Phi}_{F_i}(\underline{\bar{X}}), \tag{3.2}$$

where $\underline{\bar{X}} = (X_n)$, $n \in S_i$, $= (\varepsilon_n)$, $n \in S_i$.

Note that for every i

$$P^{\Phi}_{F_i}(\underline{\bar{X}}) = P_i(X_n=\varepsilon_n,\ n\in F_i \mid X_m=\varepsilon_m,\ m\in S_i\backslash F),$$

that is $P^{\Phi}_{F_i}(\underline{\bar{X}})$ is the conditional probability that (ε_n), $n \in F_i$, is the configuration inside, given the configuration outside is (ε_m), $m \in S_i\backslash F_i$.

Dobrushin used this specification to characterize Gibbs states on infinite S by letting $S_i \to S$ in the following sense.

Let A_F be the Borel algebra on Ω_F, $F \subset S$.

<u>Definition 3-3</u>: A random field on Ω is a Gibbs state with respect to Φ iff $\forall\, F \in \mathcal{B}(S)$ the conditional distribution $P(\cdot|A_{F^c})(\bar{X})$ is given by (3.2) where $F = F_i$.

Note: $P(\cdot|A_{F^c})$ is a probability mass function with

$$\frac{dP(\cdot|A_{F^c})}{dm} = p_F^{\Phi}\ , \quad m \text{ counting measure.}$$

Since

$$\int f dP = \int\int f dP(\cdot|A_{F^c}) dP_{F^c} \quad \forall\, f \in C(\Omega),$$

Definition 3-3 is equivalent to the condition

$$\int f dP = \int\int f p_F^{\Phi} dm dP_{F^c} \quad \forall\, f \in C(\Omega), \tag{3.3}$$

for random fields to be a Gibbs state.

In order to see that this definition is reasonable, consider the process (Y_i), $i \in \mathbb{N}$ defined by

$$\begin{aligned} Y_i &:= P(X_{F_i} \in B,\ B \in A_{F_i} | A_{S_i \setminus F_i}) \\ &= p_{F_i}^{\Phi}(\bar{X})\ , \quad \text{if } B \text{ is a singleton.} \end{aligned}$$

Note that $S_i \setminus F_i \to S \setminus F$ if $S_i \to S$, $F_i \to F$.

$$A_i := \sigma(X_m,\ m \in S_i \setminus F_i)$$

$(\mathcal{A}_i)$, $i \in \mathbb{N}$ is an increasing sequence of σ-algebras. Then

$$Y_i = P(X_{F_i} \in B | \mathcal{A}_i) \in \mathcal{A}_i \ ,$$

$$E(Y_i | \mathcal{A}_{i-1}) = E(P(X_{F_i} \in B | \mathcal{A}_i) | \mathcal{A}_{i-1})$$

$$= P(X_{F_i} \in B | \mathcal{A}_{i-1})$$

since $\mathcal{A}_{i-1} \subset \mathcal{A}_i$ (see [4] p. 35). Therefore (Y_i) is a martingale, $Y_i \geq 0 \ \forall \ i$ and hence by a martingale limit theorem

$$Y_i \to Y_\infty [P] \quad \forall \ B \in \mathcal{A}_{F_i}$$

([P] means with probability 1) where $Y_\infty \in \mathcal{A}_\infty = \mathcal{A}_{F^c}$, $Y_\infty = p_F^\Phi$ (see [1], p. 332).

Remark 3-2: (p_F^Φ), $F \in \mathcal{B}(S)$ is called local specification.

3-2.3. Gibbs States in the Sense of Lanford, Ruelle

Consider the function

$$T_A : \Omega \to \Omega, \qquad A \in \mathcal{B}(S)$$

$$T_A(\underline{\bar{X}}) = \begin{cases} X_n & n \in S \backslash A \\ -X_n & n \in A \end{cases}$$

i.e T_A changes spin direction inside of A. Let S be finite again and $A \subset F \subset S$. Let $C \in \mathcal{A}_{S \backslash F}$, $C = \{\underline{\bar{X}} \in \Omega,\ X_n = \varepsilon_n,\ n \in S \backslash F\}$.

Then

$$\frac{T_A P([\emptyset,F]\cap C)}{P([\emptyset,F]\cap C)} = \frac{P([A,F)\cap C)}{P([\emptyset,F]\cap C)}$$

$$= \exp\left\{2\beta \sum_{\substack{n\in A\\ m\in F\setminus A}} J + 2\beta \sum_{\substack{n\notin F\\ m\in A}} JX_n + 2\beta H|A|\right\}$$

$$=: h_F([A,F]\cap C)$$

where

$$h_F(\bar{\underline{X}}) := \exp\left\{\beta \sum_{n,m\in F} J + \beta \sum_{\substack{m\notin F\\ n\in F}} JX_m + \beta H|F| - \beta \sum_{n\cup m\in F} JX_n X_m - \beta H \sum_{n\in F} X_n\right\} \tag{3.4}$$

Condition (3.4) is the same as

$$\frac{dT_A P}{dP}(\bar{\underline{X}}) = h_F \circ T_A(\bar{\underline{X}}) \quad \text{for} \quad \bar{\underline{X}} \in [\emptyset,F] \tag{3.5}$$

<u>Definition 3-4</u>: A random field P is called a Gibbs state iff it satisfies (3.5) $\forall\ A \subset F \subset \mathcal{B}(S)$.

<u>Remark 3-3</u>: (3.5) is equivalent to:

$$\forall\ f \in C(\Omega): \quad \int_{[A,F]} f \circ T_A dP = \int_{[\emptyset,F]} f(h_F \circ T_A) dP\ . \tag{3.5'}$$

Note: $\quad Z_F(\bar{\underline{X}}) := \sum_{X_n, n\in F} h_F(\bar{\underline{X}})$.

then

$$\frac{h_F(\bar{\underline{X}})}{Z_F(\bar{\underline{X}})} = p_F^{\Phi}(\bar{\underline{X}}) \quad \forall\ F \in \mathcal{B}(S). \tag{3.6}$$

<u>Theorem 3-1</u>: The definitions 3-3 and 3-4 are equivalent.

Proof: (3.3) => (3.5): Let $F \in \mathcal{B}(S)$; since $P_{F^c} << P$, $\exists\, g \in \mathcal{A}$ such that

$$P_{F^c}(B) = \int 1_B g\,dP \quad \forall\, B \in \mathcal{A}_{F^c}$$

Let $f(\bar{\underline{X}}) = 1_{[A,F]}(\bar{\underline{X}})$ for $A \subset F$; $f \in \mathcal{C} \Rightarrow f \in C(\Omega)$

$$\int 1_{[A,F]}dP = \int\int 1_{[A,F]} p_F^{\Phi} dm dP_{F^c} = \int 1_{[A,F]} p_F^{\Phi} dP_{F^c}$$

$$= \int 1_{[A,F]} \frac{h_F}{Z_F} dP_{F^c} = \int 1_{[A,F]} \frac{h_F}{Z_F} g dP$$

$$\Rightarrow g(\bar{\underline{X}}) = \frac{Z_F(\bar{\underline{X}})}{h_F(\bar{\underline{X}})}, \quad \bar{\underline{X}} \in [A,F] \tag{3.7}$$

Note that h_F is constant in (3.7)

$$\int_{[A,F]} f \circ T_A dP = \int\int 1_{[A,F]}(f \circ T_A) p_F^{\Phi} dm dP_{F^c}$$

$$= \int 1_{[A,F]}(f \circ T_A) p_F^{\Phi} dP_{F^c}$$

$$= \int 1_{[\emptyset,F]} f \cdot (p_F^{\Phi} \circ T_A) dP_{F^c}$$

$$= \int 1_{[\emptyset,F]} f \cdot (h_F \circ T_A) dP$$

which is (3.5')

(3.5) => (3.3): For $f \in C(\Omega)$, $Ef = \int f dP = E(E(f|\mathcal{A}_{F^c}))$, $F \in \mathcal{B}(S)$. If the conditional expectation can be written as the integral with respect to a conditional probability, the only part to show would be that p_F^{Φ} is this conditional probability.

Since Ω is a complete, separable metric space, there exists a version of $P(\cdot|\mathcal{A}_{F^c})$ (see [2]).

Claim:

$$p_F^{\Phi} = P(\cdot\,|\,\mathcal{A}_{F^c})$$

To show:

(i) $p_F^{\Phi}(\underline{\bar{X}}) \in \mathcal{A}_{F^c}$

(ii) $P([A,F]\cap B) = \int 1_B p_F^{\Phi}([A,F])dP, \quad B \in \mathcal{A}_{F^c}$

(i) follows from the construction of p_F^{Φ} in (3-2.2).

(ii)

$$P_{F^c}(B) = \sum_{A\subset F} P([A,F]\cap B), \quad B \in \mathcal{A}_{F^c}, \; A \subset F \in \mathcal{B}(S)$$

$$= \sum_{A\subset F} T_A P([\emptyset,F]\cap B) = \sum_{A\subset F} \int_{[\emptyset,F]} 1_B dT_A P$$

$$= \sum_{A\subset F} \int_{[\emptyset,F]} 1_B (h_F \circ T_A) dP = \int_{[\emptyset,F]} 1_B (Z_F \circ T_A) dP$$

i.e.

$$\frac{dP_{F^c}}{dP} = Z_F \circ T_A \quad \text{on} \quad [\emptyset,F]$$

$$\int 1_B 1_{[A,F]} p_F^{\Phi} dP_{F^c} = \int 1_B (p_F^{\Phi}\circ T_A) 1_{[\emptyset,F]} dP_{F^c}$$

$$= \int_{[\emptyset,F]} 1_B (p_F^{\Phi}\circ T_A)(Z_F \circ T_A) dP$$

$$= \int_{[\emptyset,F]} 1_B (h_F \circ T_A) dP$$

$$= \int_{[A,F]} 1_B dP = P([A,F]\cap B)$$

3-3. General Notions to the Structure of the Set of Gibbs States

For a locally compact, countable at infinity space E, let

$$K(E) = \{f,\ f \in C(\Omega),\ T_f \text{ compact}\}$$

where $T_f = \overline{\{f \neq 0\}}$ and E is countable at infinity if there exists a sequence of compact sets with E as its union.

Ω is compact and therefore satisfies the conditions on E above. Since T_f is closed in Ω, T_f is compact $\forall\ f \in C(\Omega)$ so

$$K(\Omega) = C(\Omega)$$

Denote now by $M(\Omega)$ the set of Borel measures on Ω. By the Riesz representation theorem, there exists to every positive linear form I on $K(\Omega)$ exactly one Borel measure μ such that

$$I(f) = \int f d\mu \quad \forall\ f \in K(\Omega)$$

Note: Since Ω is metrizable, the Baire algebra coincides with the Borel algebra (see [1] p. 216).

Definition 3-4: A sequence (μ_n), n $\mathbb{N}$, $\mu_n \in M(\Omega)$ is said to converge weakly to a measure $\mu \in M(\Omega)$ if

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu \ , \quad \forall\ f \in K(\Omega)$$

Note: Weak convergence means convergence of the sequence of numbers $\int f d\mu_n$, this limit is unique. Then $f \to \lim_{n\to\infty} \int f d\mu_n$ is a positive linear form on $K(\Omega)$ and hence there exists an $\mu \in M(\Omega)$ such that

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu$$

Lemma 3-1: Let $\mu_n \in \mathcal{M}(\Omega)$ and $\mu_n \to \mu$ weakly, $\mu \in \mathcal{M}(\Omega)$. Then

$$(\mu_A)_n \to \mu_A \quad \text{weakly} \quad \forall\, A \subset S$$

Proof: $X_A : \Omega \to \Omega_A$, $X_A \in C(\Omega) \quad \forall\, A$

$$\int f dX_A \mu_n = \int f \circ X_A d\mu_n \to \int f \circ X_A d\mu = \int f dX_a \mu$$

$\forall\, f \in C(\Omega)$ since $f \circ X_A \in C(\Omega)$.

Let $\mathcal{G}(\Phi)$ be the set of Gibbs states on Ω with the same local specification for a given potential Φ.

Note: $\mathcal{G}(\Phi) \subset \mathcal{M}(\Omega)$

3-3.1. The Set of Gibbs States

In 3-2.2 the Gibbs states were characterized by their local specifications, namely a family of conditional probabilities indexed by $\mathcal{B}(S)$. However, it is not always true that two Gibbs states with the same local specification are the same. This phenomenon represents a point of central interest to the theory.

Definition 3-5: Phase transition is said to occur if

$$|\mathcal{G}(\Phi)| > 1$$

In the following, general properties of $G(\Phi)$ are established leading to distinguishing properties of certain Gibbs states in $G(\Phi)$ in case of phase transition.

Proposition 3-1: If for a pair potential Φ

$$\sum_{n \cup m \in F} J_{n,m} < \infty \quad \forall\ F \in \mathcal{B}(S)$$

then $G(\Phi)$ is weakly compact.

Proof: $C(\Omega)$ is a Banach space and therefore the unit ball in $C^*(\Omega)$ is weakly compact. $G(\Phi)$ is contained in the unit ball of $C^*(\Omega)$ $\forall\ \Phi$ since $M(\Omega) \subset C^*(\Omega)$ by the Riesz representation. Therefore $G(\Phi)$ is weakly compact if it is weakly closed and $G(\Phi)$ is weakly closed if every limit of a weakly converging sequence in $G(\Phi)$ is also in $G(\Phi)$, i.e. if for $\mu_n \in G(\Phi)$

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu \quad \forall\ f \in C(\Omega)$$

$$\Rightarrow \mu \in G(\Phi)$$

Since

$$\mu_n \in G(\Phi): \quad \int f d\mu_n = \int\int f p_F^{\Phi} dm d(\mu_{F^c})_n$$

If p_F^{Φ} is continuous on Ω_{F^c}, $\int f p_F^{\Phi} dm$ is continuous on Ω_{F^c}. Set

$$F = \int f p_F^{\Phi} dm$$

so

$$\int F d(\mu_{F^c})_n \to \int F d\bar{\mu} ,$$

where $\bar{\mu}$ is a random field on Ω_{F^c} and

$$\int F d\bar{\mu} = \int\int f p_F^{\Phi} dm d\bar{\mu} = \int f d\mu$$

By Lemma 3-1 $(\mu_{F^c}) \to \mu_{F^c}$ so $\bar{\mu} = \mu_{F^c}$ and

$$\int f d\mu = \int\int f p_F^{\Phi} dm d\mu_{F^c} \quad \text{i.e. } \mu \in \mathcal{G}(\Phi)$$

p_F^{Φ} is continuous on Ω_{F^c}:

$$p_F^{\Phi} \in C(\Omega_{F^c}) \iff \tilde{h}_F \in C(\Omega_{F^c})$$

where

$$\tilde{h}_F(\bar{\underline{X}}) = \exp\left\{-\beta J \sum_{n \cup m \in F} X_n X_m - \beta H \sum_{n \in F} X_n\right\}$$

To show: For $\bar{\underline{X}} \in \Omega$, $\varepsilon > 0$ $\exists\, \delta(\varepsilon) > 0$ such that for $\bar{\underline{Y}} \in \Omega$ with $d(\bar{\underline{X}},\bar{\underline{Y}}) < \delta$, $|\tilde{h}(\bar{\underline{X}})-\tilde{h}(\bar{\underline{Y}})| < \varepsilon$ where d is the metric defined in (3.1). Given $\bar{\underline{X}} \in \Omega$, $\varepsilon > 0$ choose $\gamma > 0$ such that

$$\tilde{h}_F(\bar{\underline{X}})(1-e^{-2\gamma}) < \varepsilon \quad \text{and} \quad \tilde{h}_F(\bar{\underline{X}})(e^{2\gamma}-1) < \varepsilon$$

By the assumption

$$\sum_{n \cup m \in F} J_{n,m} < \infty \quad \forall\, F \in \mathcal{B}(S)$$

there exists a finite set $M \in \mathcal{B}(S)$, $M \supset F$ with

$$\sum_{\substack{n \in F \\ m \notin M}} J_{n,m} < \gamma \tag{3.8}$$

Choose now $\bar{\underline{Y}} \in \Omega$ such that $Y_m = X_m \;\; \forall\, m \in M$ so

$$d(\bar{\underline{X}},\bar{\underline{Y}}) \leq \sum_{m \notin M} \frac{1}{2^m} =: \delta(\varepsilon)$$

By (3.8) and the definition of $\tilde{h}_F$:

$$\tilde{h}_F(\bar{\underline{X}})e^{-2\gamma} < \tilde{h}_F(\bar{\underline{Y}}) < \tilde{h}_F(\bar{\underline{X}})e^{2\gamma}$$

so

$$\tilde{h}_F(\bar{\underline{Y}}) - \tilde{h}_F(\bar{\underline{X}}) < (e^{2\gamma}-1)\tilde{h}_F(\bar{\underline{X}}) < \varepsilon$$

and

$$\tilde{h}_F(\bar{\underline{Y}}) - \tilde{h}_F(\bar{\underline{X}}) > (e^{-2\gamma}-1)\tilde{h}_F(\bar{\underline{X}}) > -\varepsilon$$

hence

$$|\tilde{h}(\bar{\underline{Y}})-\tilde{h}(\bar{\underline{X}})| < \varepsilon$$

and thus p_F^Φ is continuous on Ω_{F^c}.

Proposition 3-2: $G(\Phi)$ is convex.

Proof: Let

$$P = \alpha P_1 + (1-\alpha)P_2 \;, \qquad 0 < \alpha < 1,\; P_1,P_2 \in G(\Phi)$$

then

$$P_{F^c} = \alpha P_{1F^c} + (1-\alpha)P_{2F^c} \;, \qquad F \in B(\,S)$$

and

$$P(\cdot|A_{F^c}) = \alpha P_1(\cdot|A_{F^c}) + (1-\alpha)P_2(\cdot|A_{F^c})$$

$$= \alpha p_F^\Phi + (1-\alpha)p_F^\Phi$$

$$= p_F^\Phi$$

i.e. $\quad P \in G(\Phi)$

Proposition 3-3: $\mathcal{G}(\Phi)$ is non empty.

Proof: Let $F_\gamma \in \mathcal{B}(S)$ with $F_\gamma \uparrow S$. Define random fields on Ω by

(i) $P_\gamma([A,F]\cap B) = p_F^\Phi([A,F]\cap B\cap C)$ where $A \subset F \subset F_\gamma$, $B \in \mathcal{A}_{F_\gamma \setminus F}$,

B singleton and $C = \{\underline{\bar{X}}\in\Omega,\ X_n=-1\ \forall\ n\in F_\gamma^c\}$

(ii) $P_\gamma(X_n=-1) = 1 \quad \forall\ n \in F_\gamma^c$

Note: $P_\gamma(C) = 1$.

Since the set of random fields on Ω is contained in the unit ball of $C^*(\Omega)$, which is weakly compact, every sequence of random fields has a weakly converging subsequence. Assume therefore P_γ converges weakly to some random field P.

$$\begin{aligned}
P_\gamma([A,F]\cap B) &= \int_B P_\gamma([A,F]\,|\,\mathcal{A}_{F^c})dP_\gamma \\
&= \int_{B\cap C} P_\gamma([A,F]\,|\,\mathcal{A}_{F^c})dP_\gamma \\
&= P_\gamma([A,F]\,|\,\mathcal{A}_{F^c})(\{B\cap C\}) \\
&= P_\gamma([A,F]\,|\,\mathcal{A}_{F_\gamma \setminus F})(\{B\})
\end{aligned}$$

so

$$P_\gamma([A,F]) = P_\gamma([A,F]\,|\,\mathcal{A}_{F_\gamma\setminus F})$$

or

$$p_F^\Phi([A,F]\cap C) = P_\gamma([A,F]\,|\,\mathcal{A}_{F_\gamma\setminus F})$$

In the limit $\gamma \to \infty$: $p_F^\Phi([A,F])(\underline{\bar{X}}) = P([A,F]\,|\,\mathcal{A}_{F^c})(\underline{\bar{X}})$, $X \in \Omega_{F^c}$.

Therefore

$$P \in \mathcal{G}(\Phi)$$

An extremal point of $G(\Phi)$ is an extremal set in $G(\Phi)$ consisting of only one point, whereas an extremal set $E \subset G(\Phi)$ is such that every convex combination of two elements is in E only if both points are in E. These points play an important role in the theory of Gibbs states.

<u>Definition 3-6</u>: An extremal point of $G(\Phi)$ is called a pure state or pure phase, the set of pure states in $G(\Phi)$ is denoted by ext $G(\Phi)$.

<u>Remark 3-4</u>: Phase transition is equivalent to

$$|\text{ext } G(\Phi)| > 1$$

<u>Theorem 3-2</u>: [Krein Millman] (see [20] for proof). Let E be a locally convex topological vector space, satisfying the Hausdorff axiom. Let K be a non void, convex, compact subset. Then K has at least one extremal point and K is the closure of the convex hull of the set of all extreme points.

<u>Corollary 3-1</u>: $G(\Phi)$ has at least one pure phase and every Gibbs state which represents a non pure phase can be obtained as a limit of convex combinations of pure phases.

<u>Proof</u>: $C^*(\Omega)$ is locally convex, i.e. any of its open sets containing 0, contains a convex, balanced and absorbing open set, and is Hausdorff: The local convexity is easily seen and the separation axiom follows from the fact that the weak topology is Hausdorff, which can be proved by using the Hahn Banach theorem. $G(\Phi)$ is a

non void, convex, compact subset of $C^*(\Omega)$ and the result follows by Theorem 3-2.

<u>Definition 3-7</u>: A subset of a topological space B is called a G_σ-set if it is the intersection of a sequence of open sets.

<u>Definition 3-8</u>: Let K be convex, compact and μ be a probability measure on the Borel algebra of K. A point $x_0 \in K$ is called barycenter of μ if for each $f \in C(K)$ of the form $f(x) = \ell(x) + \alpha$, $\ell \in K^*$, $\alpha \in \mathbb{R}$

$$f(x_0) = \int f(x)\mu(dx) \tag{3.9}$$

The following theorem provides an extension of Corollary 3-1.

<u>Theorem 3-3</u>: [Choquet] (see [8] for proof). Let E be a locally convex topological vector space, which is Hausdorff and K a non void, compact, convex set such that the induced topology on K is metrizable. Then $\text{ext}\, K$ is a G_σ-set and hence belongs to the Borel algebra of K.
For every $x \in K$ $\exists$ at least one probability measure μ on K such that $\mu(\text{ext}\, K) = 1$ and x is the barycenter of μ.

A further result we need is

<u>Lemma 3-2</u>: If E is a compact metrizable space then $M(E)$ is metrizable in the weak topology (for a proof see [17], p. 148).

In Corollary 3-1 it was proved that every non pure Gibbs state is the limit of convex combinations of pure phases. The limit of convex combinations is an integral with respect to some probability measure.

Corollary 3-2: For every Gibbs state $\mu \in \mathcal{G}(\Phi)$ $\exists$ at least one probability measure P on $\mathcal{G}(\Phi)$ with its Borel algebra, such that

$$\mu(A) = \int \gamma(A)\,dP \qquad \forall\ A \in \mathcal{A}$$

where $\gamma \in \text{ext}\ (\Phi)$

Proof: Ω is compact and metrizable. By Lemma 3-2 $\mathcal{M}(\Omega)$ is metrizable and so is the induced topology on $\mathcal{G}(\Phi) \subset \mathcal{M}(\Phi)$. $\mathcal{G}(\Phi)$ is a non void, convex and compact subset of a locally convex topological vector space which is Hausdorff, namely $C^*(\Omega)$. Therefore by Theorem 3-3 $\exists$ to $\mu \in \mathcal{G}(\Phi)$ at least one such P on $\mathcal{G}(\Phi)$ such that $P(\text{ext}\ \mathcal{G}(\Phi)) = 1$ and μ is the barycenter of P:

$$f(\mu) = \int f(\gamma)P(d\gamma) \qquad \gamma \in \mathcal{G}(\Phi)$$

where f is as in Theorem 3-3. Set $f_A(\mu) = \mu(A)$ for $A \in \mathcal{A}$ then

$$\mu(A) = \int \gamma(A)P(d\gamma) = \int_{\text{ext}\ \mathcal{G}(\Phi)} \gamma(A)P(d\gamma)$$

From the above follows that $\mathcal{G}(\Phi)$ is completely determined if the pure phases are known. In the following various properties and characterizations of pure phases are presented.

Definition 3-9: $\mathcal{A}_\infty := \bigcap_{F \in \mathcal{B}(S)} \mathcal{A}_{F^c}$ is called tail σ-algebra. A probability measure P on Ω is said to have trivial tail if $P(E) = 0$ or 1 $\forall\ E \in \mathcal{A}_\infty$.

Note: If g is measurable with respect to $\mathcal{A}_\infty$, g is constant since if $g \in \mathcal{A}_{F^c}$, g is constant on Ω_F; for $g \in \mathcal{A}_\infty$, g is constant on Ω_F $\forall\, F \in \mathcal{B}(S)$.

Proposition 3-4: $P \in \mathcal{G}(\Phi)$ is extreme <=> P has trivial tail.

Proof: (a) Assume P is extreme and $E \in \mathcal{A}_\infty$ such that $0 < P(E) < 1$. For

$$A \in \mathcal{A}: \quad P(A) = P(A|E)P(E) + P(A|E^c)P(E^c)$$

The measure $P(\cdot|E)$ on Ω is in $\mathcal{G}(\Phi)$. To see this note that

$$P(\cdot|E|\mathcal{A}_{F^c}) = P(\cdot|\mathcal{A}_{F^c}) = p_F^\Phi \quad \text{for any} \quad F \in \mathcal{B}(S)$$

since

$$E \in \mathcal{A}_{F^c} \quad \forall\, F \in \mathcal{B}(S)$$

One also has $P(\cdot|E^c)$ is in $\mathcal{G}(\Phi)$ by the same reason. Set $\alpha = P(E)$, then $P(A) = \alpha P_1(A) + (1-\alpha)P_2(A)$ where $P_1 = P(\cdot|E)$, $P_2 = P(\cdot|E^c)$, contradicting that P is extreme.

(b) Assume P has trivial tail, $P \in \mathcal{G}(\Phi)$ and $P = \alpha P_1 + (1-\alpha)P_2$, $0 \le \alpha \le 1$, $P_1, P_2 \in \mathcal{G}(\Phi)$, $P_1 \ll P$ and therefore $\exists$ $g \in \mathcal{A}$ with

$$\frac{dP_1}{dP} = g \qquad g \ge 0$$

$$E_P g = \int g\,dP = P_1(\Omega) = 1$$

Let now $A \in \mathcal{A}_F$, $B \in \mathcal{A}_{F^c}$

$$P_1(A\cap B) = \int 1_A 1_B dP_1 = \iint 1_A 1_B dP_1(\cdot|\mathcal{A}_{F^c})dP_1$$

$$= \int 1_B P_1(A|\mathcal{A}_{F^c})dP_1 = \int 1_B P_1(A|\mathcal{A}_{F^c})gdP$$

On the other hand

$$P_1(A\cap B) = \int 1_A 1_B gdP = \int 1_B E(1_A g|\mathcal{A}_{F^c})dP$$

but since $P_1, P_2 \in \mathcal{G}(\Phi)$

$$P_1(A|\mathcal{A}_{F^c}) = P(A|\mathcal{A}_{F^c}) = E(1_A|\mathcal{A}_{F^c})$$

So

$$\int 1_B E(1_A g|\mathcal{A}_{F^c})dP = \int 1_B g E(1_A|\mathcal{A}_{F^c})dP$$

$$\Rightarrow g \in \mathcal{A}_{F^c} \quad \forall\, F \in \mathcal{B}(S)$$

$$\Rightarrow g \in \mathcal{A}_\infty \quad \Rightarrow g \text{ constant } [P]$$

Since $E_P g = 1 \Rightarrow g = 1$ $[P]$ and hence $P = P_1[P]$ i.e. P is extreme.

<u>Proposition 3-5</u>: If $P_1, P_2 \in \text{ext } \mathcal{G}(\Phi)$ and $P_1|_{\mathcal{A}_\infty} = P_2|_{\mathcal{A}_\infty}$ then $P_1 = P_2$.

<u>Proof</u>: Label all $F \in \mathcal{B}(S)$. Define $\mathcal{C}_1 := \mathcal{C}_{F_1^c}$, $\mathcal{C}_n := \bigcap_{k=1}^{n} \mathcal{A}_{F_k^c}$.
Note that

$$\mathcal{A}_{F_1^c} \cap \mathcal{A}_{F_2^c} = \mathcal{A}_{F_1^c\cap F_2^c}$$

so

$$\mathcal{C}_n = \mathcal{A}_{\bigcap_1^n F_k^c}$$

$\mathcal{C}_n$ is a decreasing sequence of σ-algebras with $\mathcal{C}_n \downarrow \mathcal{A}_\infty$. Therefore

$$P(\cdot|\mathcal{C}_n) \to P(\cdot|\mathcal{A}_\infty)[P] \quad \forall\, P \in \mathcal{G}(\Phi)$$

Since $P_1, P_2 \in \mathcal{G}(\Phi)$, $P_1(\cdot|\mathcal{C}_n) = P_2(\cdot|\mathcal{C}_n)$ and since they are extreme

$$P_1(\cdot|\mathcal{C}_n) \to P_1(\cdot|\mathcal{A}_\infty) = \text{const.} = c_1[P_1]$$

$$P_2(\cdot|\mathcal{C}_n) \to P_2(\cdot|\mathcal{A}_\infty) = \text{const.} = c_2[P_2]$$

If $c_1 \neq c_2$ let A_1 be the set of convergence to c_1, then $P_1(A_1) = 1$, $P_2(A_1) = 0$ and therefore, by the assumption, $c_1 = c_2$. For $B \in \mathcal{A}$:

$$P_1(B) = P_1(B|\mathcal{A}_\infty) = P_2(B|\mathcal{A}_\infty) = P_2(B)$$

<u>Proposition 3-6</u>: If $P_1 \neq P_2$, $P_1, P_2 \in \text{ext}\, \mathcal{G}(\Phi)$ then $P_1 \perp P_2$, i.e. $\min(P_1, P_2) = 0$.

<u>Proof</u>: By Proposition 3-5 $P_1 \neq P_2$ on $\mathcal{A}_\infty$, but both are tail trivial, i.e. $P_i(A) = 0$ or 1 for $A \in \mathcal{A}_\infty$, $i = 1,2$. $\exists\, A \in \mathcal{A}_\infty$ s. th.

$$P_1(A) = 0, \quad P_2(A) = 1$$

If $B \in \mathcal{A}$ with $P_2(B) > 0 \Rightarrow B \subset A[P_2]$ since if $B \cap A^c \neq \emptyset$ $P_2(B \cap A^c) = 0$. Also if $C \in \mathcal{A}$ with $P_1(C) > 0 \Rightarrow C \subset A^c[P_1] \Rightarrow P_1 \perp P_2$.

By Proposition 3-4 it is now possible to obtain conditions for a Gibbs state to be pure by conditions for tail triviality. The following result characterizes pure Gibbs states by conditions for tail triviality.

Theorem 3-4: (Lanford-Ruelle). Let P be a probability measure on Ω. Then P has trivial tail iff $\forall\, A \in \mathcal{A}$, $\varepsilon < 0$ $\exists\, F \in \mathcal{B}(S)$ such that

$$|P(A \cap B) - P(A)P(B)| < \varepsilon \qquad \forall\, B \in \mathcal{A}_{F^c} \tag{3.10}$$

Proof: Assume (3.10) holds. Let $E \in \mathcal{A}_\infty$, i.e. $E \in \mathcal{A}_{F^c}$ $\forall\, F \in \mathcal{B}(S)$. Set $A = E$ to get

$$P(E)^2 = P(E) \Rightarrow P(E) = 0 \text{ or } 1$$

Conversely define linear functionals on $L_1(\Omega)$ by

$$\ell_B(f) = \int 1_B f dP, \qquad B \in \mathcal{A}$$

so

$$\ell_B(1_A) = P(A \cap B)$$

$$|\ell_B f| \leq \int |f| dP = \| f \|_{L_1} \Rightarrow \ell_B \in L_1^*(\Omega) = L_\infty(\Omega)$$

$L_1(\Omega)$ is separable and complete and so $L_1^*(\Omega)$ has a sequentially compact unit ball. ℓ_B is contained in this unit ball $\forall\, B \in \mathcal{A}$. Consider now $F_n \in \mathcal{B}(S)$, $F_n \uparrow S$. Then $\exists$ a subsequence such that

$$\lim \ell_{B_{n_k}}(f) = \int fg dP, \qquad f \in L_1(\Omega) \tag{3.11}$$

$$g \in L_\infty(\Omega), \qquad B_{n_k} \subset F^c_{n_k}$$

Since no $\ell_{B_n} f$ depends on coordinates inside of F_n $\exists\, g \in L_\infty$ satisfying (3.11) and $g \in \mathcal{A}_\infty \Rightarrow g$ is constant. From (3.11) follows that for $\varepsilon > 0$ $\exists$ $F_{n_0} \in \mathcal{B}(S)$ such that

$$|\ell_{B_{n_0}}(f) - \int fgdP| < \varepsilon, \quad f \in L_1(\Omega) \tag{3.12}$$

$$B_{n_0} \subset F^c_{n_0}$$

Let $f = 1_A$, $A \in \mathcal{A}$ then (3.12) becomes

$$|\int 1_A 1_{B_{n_0}} dP - g \int 1_A dP| < \varepsilon$$

For $A = \Omega$ follows that $g = \lim P(B_{n_k})$. Therefore

$$|\int 1_A 1_{B_{n_1}} dP - \int 1_{B_{n_1}} dP \int 1_A dP| < \varepsilon, \quad \text{for some } n_1$$

$$\Longleftrightarrow \qquad |P(A \cap B_{n_1}) - P(A)P(B_{n_1})| < \varepsilon$$

where B_{n_1} arbitrary in $F^c_{n_1}$.

Remark 3-5: The following are equivalent to Theorem 3-4.

(i) Given any $f \in L_1(\Omega)$ there exists $F \in \mathcal{B}(S)$ such that

$$|\int fgdP - \int fdP \int gdP| \leq \| g \|_\infty \tag{3.13}$$

whenever $g \in L_\infty(\Omega_{F^c})$.

(ii) For each $f \in C(\Omega)$ there exists $F \in \mathcal{B}(S)$ such that

$$\left|\int fgdP - \int fdP \int gdP\right| \leq \| g \|_{\infty} \tag{3.14}$$

whenever $g \in \mathcal{A}_{F^c}$, $g \in C(\Omega)$ (see [19]).

Chapter 4. MARKOV RANDOM FIELDS AND NEAREST NEIGHBORHOOD GIBBS STATES

Up to now no special structure on the set of sites S was used. Let S now have a graph structure, i.e. S is the set of vertices of some graph, where a graph (Γ, E) is a set of sites Γ and a set of edges E such that:

(i) every edge connects two sites

(ii) two sites are connected by at most one edge

(iii) there are no edges connecting a site with itself (i.e. no loops).

Two sites are called neighbors if there exists an edge connecting them. For $A \in S$ the boundary of A is defined by

$$\partial A := \{n \in S \backslash A, \exists\, m \in A, \text{ n and m are neighbors}\}$$

i.e. ∂A is the set of neighbors of A outside of A.

A set $F \in \mathcal{B}(S)$ will be called connected if for $n, m \in F$ there exists a sequence of sites $n_0 = n$, $n_1, n_2 \ldots n_k = m$ in F such that n_i is a neighbor of n_{i+1}. A sequence of sites (n_i) where n_i is a neighbor of n_{i+1} is called a path.

A potential Φ is defined as in Chapter 3 and again only symmetric, invariant pair potentials are considered where

$$\Phi(\{X_n, X_m\}) = \Phi_{\{n,m\}}(\underline{\bar{X}})$$

$$= -J_{n,m} X_n X_m \;, \quad \underline{\bar{X}} \in \Omega,$$

$$J_{n,m} = J_{m,n} = J$$

and invariant now means invariant under graph-isomorphisms.

Definition 4-1: A pair potential is said to have the nearest neighborhood property if $J_{n,m} = 0$, whenever n and m are not neighbors.

As before Gibbs states are characterized by their local specifications, i.e. their conditional probabilities.

Definition 4-2: A random field on Ω is a Markov random field (MRF) if

(i) $P([A,F]|A_{F^c}) > 0 \quad \forall\, F \in \mathcal{B}(S),\ A \subset F$

(ii) $P([A,F]|A_{F^c}) = (P[A,F]|X_k,\ k \in \partial F) \quad \forall\, F \in \mathcal{B}(S)$ [Markov property].

It is easy to see that every nearest neighborhood Gibbs state satisfies (i) and (ii) since p_F^{Φ} has exponential form and $J_{n,m} = 0$ whenever $n \in \partial\{m\}$. On the other hand for a given MRFμ, it is possible to define a potential Φ such that the Markov property of the conditional distributions yields the nearest neighborhood property of the potential and $\mu \in \mathcal{G}(\Phi)$ (see [15]). Therefore one has the result:

Lemma 4-1: Every MRF on Ω is a nearest neighborhood Gibbs state and vice versa.

The conditional probabilities of a MRF, P, in Definition 4-2 are completely determined by a given nearest neighborhood potential Φ, if $P \in \mathcal{G}(\Phi)$, through

$$P([A,F]|A_{F^c}) = Z_F^{-1} \exp\left\{-\sum_{W\cap F\neq\emptyset} \Phi_W(\bar{\underline{X}})\right\} \tag{4.1}$$

for $A \subset F \in \mathcal{B}(S)$

The conditional probabilities in (4.1) again are determined by specifying them for connected $F \in \mathcal{B}(S)$. Since arbitrary $F \in \mathcal{B}(S)$ can be decomposed into

$$F = \sum_{i=1}^{n} F_i$$

where $F_i \in \mathcal{B}(S)$ connected and

$$[A,F] = \bigcap_{i=1}^{n} [A_i, F_i]$$

where $A_i = A \cap F_i$

These conditional probabilities of connected cylinder sets again are determined by probabilities of the form

$$\alpha_k^{(n)} := P(X_n = 1 | k \text{ neighbors of } n \text{ are } -1), \quad n \in S \tag{4.2}$$

Since

$$P([A,F]|A_{F^c}) = P([A,F]|X_k,\ k\in\partial F)$$

can be expressed in terms of $\alpha_k^{(n)}$, $k = 0,1\ldots|\partial\{n\}|$. So $\mathcal{G}(\Phi)$ denotes the set of MRF's such that probabilities of form (4.2) satisfy

$$\alpha_k^{(n)} = \frac{\exp\{-(N-2k)\beta J-\beta H\}}{2\cosh\{(N-2k)\beta J+\beta H\}}, \tag{4.3}$$

if N is the number of neighbors of $n \in S$ and Φ a nearest neighborhood (n.n.) potential.

A set of graphs for which calculations turn out to be fairly simple is the set of trees or Bethe Lattices.

Definition 4-2: A graph S is called a tree or a Bethe lattice if

(i) S is connected

(ii) S contains no circuits

(iii) Every site has the same number of neighbors.

T_N denotes the tree with N branches, i.e. each site $n \in T_N$ has $N+1$ neighbors. From the definition follows that to $n,m \in T_N$ there exists a unique path from n to m.

Note: For $N = 1$, $T_N = Z$ a Gibbs state on T_1 is referred to as the one dimensional Ising model.

For a connected finite subset of T_N there exists a simple but useful labeling: Let $M \in \mathcal{B}(T_N)$ be connected, $M = (1,\ldots,k)$ where each i has exactly one neighbor j with $i > j$ for $1 < i \leq k$; $(i) := j$.

Note: This labeling is not unique.

A state μ on $\Omega = \{1,-1\}^{T_N}$ is a nearest neighborhood Gibbs state if the conditional probabilities of the form (4.2) satisfy (4.3) and is then called Ising model on the tree.

For the repulsive case $J < 0$ it is possible to define a MRF such that for fixed site n transitions from sites an even number of branches away from n to sites an odd number of branches away from n have different probability than from sites an odd number away to sites an even number away. If this occurs the MRF is said to exhibit symmetry breaking.

Chapter 5. INFINITELY DIVISIBLE GIBBS STATES

Infinitely divisible distributions have the property that their characteristic function, i.e. Fourier-Stieltjes transform admits an exponential representation. The characteristic function of a Gibbs state is called a correlation function and is of significant importance in statistical mechanics, since correlation can be measured experimentally and represents therefore also a measure for the reference to the physical reality of the mathematical model.

The mathematical operation of convolving Gibbs states may be interpreted as to randomly change the spins.

5-1. General Notations

Let G be a locally compact abelian (LCA) group and $M^1(G)$ the set of finite, normalized Borel measures. The mapping

$$\phi : G^n \to G$$

defined by

$$\phi(g_1 \dots g_n) = \sum_{i=1}^{n} g_i \,, \quad g_i \in G$$

is continuous and therefore measurable. Hence ϕ induces a mapping $\bar{\phi} : M^1(G)^n \to M^1(G)$ by

$$\bar{\phi}(\mu_1 \otimes \dots \otimes \mu_n)(B) = \mu_1 \otimes \dots \otimes \mu_n(\phi^{-1}(B))$$

for any Borel set B.

Definition 5-1: $\mu_1 * \ldots * \mu_n := \bar{\phi}(\mu_1 \otimes \ldots \otimes \mu_n)$ is called the convolution of $\mu_1,\ldots,\mu_n$.

Definition 5-2: A group homomorphism $\alpha: G \to C$ is called an algebraic character if $|\alpha(x)| = 1 \ \forall\, x \in G$. The group of continuous characters of G is called the dual group $\hat{G}$ of G.

Remark 5-1: If G_i, $i \in I$ countable, are LCA groups then

$$\widehat{\otimes G_i} = \otimes \hat{G}_i$$

i.e. a continuous character on the direct product is a product of characters of each group.

Remark 5-2: The function

$$\hat{\mu}(\gamma) := \int \overline{\gamma(x)}\mu(dx) \qquad \gamma \in \hat{G}$$

on $\hat{G}$ is called Fourier Stieltjes transform of μ. Note that $\hat{\mu}$ is a complex valued function.

Properties of $\hat{\mu}$:

(i) $\hat{\mu}$ is uniformly continuous on G

(ii) $\widehat{\mu * \gamma} = \hat{\mu} \cdot \hat{\gamma} \qquad \forall\, \mu,\gamma \in M^1(G)$ since

$$\int \bar{\gamma} d\mu * \gamma = \iint \overline{\gamma(x+y)} d\mu \times \gamma = \iint \bar{\gamma}(x)\bar{\gamma}(y) d\mu d\gamma = \int \bar{\gamma} d\mu \int \bar{\gamma} d\gamma$$

(iii) The mapping $\mu \to \hat{\mu}$ is one to one.

Definition 5-3: A measure μ on G with $\mu(A \cdot x) = \mu(A)$ for every Borel set A and every $x \in G$ is called Haar measure on G.

Note: On every compact topological group exists a unique Haar measure (see [5]), e.g. on $\bar{\mathbb{R}}$ the Lebesgue-measure is the unique Haar measure.

With this it is now possible to define the Fourier transform also for functions:

<u>Definition 5-4</u>: For $f \in L_1(G)$, L_1 with respect to Haar measure λ on G,

$$\hat{f}(\gamma) := \int \bar{\gamma}(x) f(x) \lambda(dx), \quad \gamma \in G$$

is called the Fourier transform of f.

A useful fact for estimations is the Riemann Lebesgue Lemma: $\hat{f}$ vanishes at infinity $\forall f \in L_1(G)$, i.e. $\hat{f} \in K(\hat{G})$, where $K(\hat{G})$ is the set of functions f with compact T_f.

With this preparation, infinite divisibility can now be defined.

<u>Definition 5-5</u>: A measure $\mu \in M^1(G)$ is called infinitely divisible if for each n there exists $\mu_n \in M^1(G)$ such that

$$\mu = \mu_n * \dots * \mu_n \quad (n \text{ fold})$$

or equivalently by property (ii)

$$\hat{\mu} = (\hat{\mu}_n)^n .$$

Example 5-1: Define $\exp \mu$ as

$$\exp \mu = \exp\{-\mu(G)\}\{1 + \mu + \frac{\mu^2}{2!} + \dots + \frac{\mu^n}{n!} + \dots\}$$

for $\mu \in M^1(G)$. Then

$$(\widehat{\exp \mu})(\gamma) = \int\bar{\gamma} d \exp\mu - \int\bar{\gamma} \exp\{-\mu(G)\} \sum_{n=0}^{\infty} \frac{d\mu^n}{n!}$$

$$= \int\bar{\gamma} \exp\{-\mu(G)\} \sum_{n=0}^{\infty} \frac{(d\mu)^n}{n!}$$

$$= \sum_{n=0}^{\infty} \exp\{-\mu(G)\} \frac{(\int\bar{\gamma} d\mu)^n}{n!}$$

$$= \exp\{-\mu(G)\}\exp\{\int\bar{\gamma} d\mu\}$$

$$= \exp\{\int\bar{\gamma}-1 d\mu\}. \tag{5.1}$$

If $\mu_1, \mu_2 \in M^1(G)$, $\exp(\mu_1+\mu_2) = \exp \mu_1 * \exp \mu_2$ and from (5.1) follows that $\exp \mu$ is infinitely divisible.

Distributions as constructed in Example 5-1 are called elementary infinitely divisible distributions and every infinitely divisible distribution contains an elementary distribution as a factor (with respect to convolution). The advantage of infinitely divisible distributions is, that its characteristic function admits a canonical representation as stated below in a special case.

Theorem 5-1: On a totally disconnected LCA group G an infinitely divisible distribution has the form

$$\hat{\mu}(\gamma) = \gamma(x_0) \cdot \hat{\lambda}(\gamma) \exp\{\int(\gamma-1)dF\} \tag{5.2}$$

where $x_0 \in G$, λ the Haar measure on a compact subgroup of G, F a σ-finite measure, finite outside every neighborhood of the identity and

$$\int (1-\mathrm{Re}\gamma)\,dF < \infty, \qquad \forall\, \gamma \in \hat{G}.$$

F is called the Levy-Khintchin measure of μ. (See [12] for a proof and more general cases.)

5-2. The Case of the Ising Model

If a Gibbs state would be infinitely divisible, its correlation function would have the form (5.2). For a countable set of sites S,

$$\Omega = \bigotimes_{i \in S} W_i = W_i^S = \{1,-1\}^S$$

is the direct product of the multiplicative groups W_i. So by Remark 5-1

$$\hat{\Omega} = \bigotimes_{i \in S} \hat{W}_i$$

since $W_i = \{1,-1\}$ $\forall\, i \in S$ is a LCA group.

(a) Characters on W_i: Since the characters are linear $\gamma(-1) = \gamma(1)\gamma(-1)$ and hence $\gamma(1) = 1$ $\forall$ characters. From $\gamma(-1) \cdot \gamma(-1) = \gamma(1) = 1$ follows that $\gamma(-1)$ has to be real and $\gamma(-1) = 1$ or -1. Therefore $\hat{W}_i = \{\gamma_1, \gamma_2\}$ $\forall\, i \in S$, where

$$\gamma_1(1) = 1 \qquad \gamma_1(-1) = 1$$

$$\gamma_2(1) = 1 \qquad \gamma_2(-1) = -1$$

and both characters are continuous.

(b) Characters on Ω: To each character γ on Ω exist characters γ_{ε_i} on W_i, $\varepsilon_i \in \{1,2\}$, $i \in S$ such that

$$\gamma(\underline{\bar{X}}) = \prod_{i \in S} \gamma_{\varepsilon_i}(X_i) = \prod_{\substack{i \in S \\ \varepsilon_i = 2}} \gamma_{\varepsilon_i}(X_i) \ . \tag{5.3}$$

In order to have γ be well defined, the last product in (5.3) has to be finite, i.e.

$$\gamma(\underline{\bar{X}}) = \prod_{i \in A} \gamma_2(X_i), \qquad A \in \mathcal{B}(S). \tag{5.4}$$

This means that each character on Ω can be associated with some set in $\mathcal{B}(S)$. (5.4) is equivalent to

$$\gamma(\underline{\bar{X}}) = \gamma_A(\underline{\bar{X}}) = \prod_{i \in A} \gamma_2(X_i) = (-1)^{|A \cap \chi|} \tag{5.5}$$

where

$$A = \{i \in S, \ \varepsilon_i = 2\}$$

$$\chi = \{i \in S, \ X_i = -1\}.$$

On the otherhand, define for given $A \in \mathcal{B}(S)$

$$\gamma_A(\underline{\bar{X}}) = (-1)^{|A \cap \chi|}$$

γ_A is continuous ($\gamma_A \in C$), linear and has absolute value 1 and so $\gamma_A \in \hat{\Omega}$. Therefore:

Lemma 5-1: For the configuration space Ω over a countable set S

$$\hat{\Omega} = \mathcal{B}(S).$$

$\gamma \in \hat{\Omega}$ will be denoted by γ_D or D, $D \in \mathcal{B}(S)$, $D \sim \gamma$. Note: The group operaton on $\mathcal{B}(S)$ is symmetric difference.

(c) The Correlation Function: Let now P be a Gibbs state on Ω. The Fourier-Stieltjes transform $\hat{P}$ of P is called correlation function.

$$\hat{P}(\gamma_D) = \int \gamma_D dP = \int (-1)^{|D \cap \chi|} dP = \int \prod_D X_n dP = E \prod_D X_n , \qquad D \in \mathcal{B}(S)$$

$\hat{P}(\gamma_D)$ is called $|D|$-point correlation function.

Remark 5-3: A Gibbs state P is infinitely divisible if there exists for each n a probability distribution P_n, such that

$$\hat{P} = (\hat{P}_n)^n, \text{ where } P_n \text{ need not be a Gibbs state.}$$

The set $\{\gamma_D : \hat{\mu}(\gamma_D) = E\gamma_D \neq 0 =: B_0$ for μ a Gibbs state is an open subgroup of $\mathcal{B}(S)$, if $\mathcal{B}(S)$ is viewed as a group under symmetric difference, since B_0^c is closed. Let H be the annihilator of B_0 in Ω, i.e.

$$H = \{\underline{\bar{X}} \in \Omega, \ \gamma_D(\underline{\bar{X}}) = 1, \ \forall \ \gamma_D \in B_0\}$$

$$= \{\underline{\bar{X}} \in \Omega, \ \gamma_D(\underline{\bar{X}}) = 1, \ \forall \ \gamma_D \in \mathcal{B}(S) \text{ with } E\gamma_D \neq 0$$

then λ in Theorem 5-1 can be taken to be the normalized Haar measure on H. Note that H is a compact subgroup of Ω. With these notations:

Corollary 5-1: $H = \{\underline{\bar{X}} \in \Omega, \ X_n = 1 \text{ or } X_n = -1 \ \forall \ n \in S\}$

$$= \{\underline{\bar{X}}_+, \underline{\bar{X}}_-\} \quad \text{for 0-external field}$$

λ is Bernoulli $(\frac{1}{2})$ measure on H and

$$\hat{\lambda}(D) = \begin{cases} 1 & \text{if } |D| \text{ even} \\ 0 & \text{if } |D| \text{ odd} \qquad D \in \mathcal{B}(S) \end{cases}$$

Proof: It is easily seen that $E\gamma_D \neq 0$ if $D = \{n,m\}$ for $n,m \in S$. The rest follows easily from this and the definition of H.

Lemma 5-2: For a Gibbs state P on Ω with 0-external field

$$\hat{P}(D) = E\gamma_D = 0 \quad \text{if} \quad |D| \text{ is odd}$$

Proof: For a given configuration, $\prod_D X_n = \gamma_D = 1$ iff $|D \cap \chi|$ is even and $\gamma_D = -1$ iff $|D \cap \chi|$ is odd. Ω_D can be decomposed into $\Omega_D = \Omega_0 \cup \Omega_1$, where

$$\Omega_0 = \{\underline{\bar{X}} \in \Omega_D,\ \gamma_D(\underline{\bar{X}}) = 1\}$$

$$\Omega_1 = \{\underline{\bar{X}} \in \Omega_D,\ \gamma_D(\underline{\bar{X}}) = -1\}$$

Ω_0 and Ω_1 are homeomorphic, take for example $f(\underline{\bar{X}}) = -\underline{\bar{X}}$; $f(\underline{\bar{X}}) \in \Omega_0 \Rightarrow \underline{\bar{X}} \in \Omega_1$, and $f = f^{-1}$. $f(\underline{\bar{X}})$ and $\underline{\bar{X}}$ have the same distribution. Therefore

$$E\gamma_D = P(\gamma_D = 1) - P(\gamma_D = -1)$$

$$= \sum_{\Omega_0} P(\underline{\bar{X}}) - \sum_{\Omega_1} P(\underline{\bar{X}}) = 0.$$

The Haar measure on Ω is Bernoulli $(\frac{1}{2})$ measure Q. So the Fourier transform of $f \in L_1(\Omega)$ becomes

$$\hat{f}(D) = \int (-1)^{|D \cap \chi|} f dQ = \int \prod_D X_n f(\bar{\underline{X}}) Q d\bar{\underline{X}} \ . \tag{5.6}$$

From the Riemann-Lebesgue Lemma follows that for $\varepsilon > 0$ $\exists F \in \mathcal{B}(S)$ such that

$$|\hat{f}(D)| < \varepsilon \quad \forall D \subset F^c, \ D \in \mathcal{B}(S)$$

5-3. Examples of Infinitely Divisible Gibbs States

(a) Infinite temperature model: For infinite temperature there is no interaction among the particles so the model can be described by

$$\Phi_A(\bar{\underline{X}}) = \begin{cases} -HX_n & \text{if} \quad A = \{n\}, \quad \text{s.th.} \quad |\beta H| < \infty \\ 0 & \text{otherwise} \end{cases}$$

which means the only influence is due to an external field H. Since no interaction occurs $P(A|A_{F^c}) = P(A)$, $\forall A \in A_F$, $\forall F \in \mathcal{B}(S)$ and therefore it follows from the Kolmogoroff construction that $\mathcal{G}(\Phi)$ is a singleton $\forall \Phi$, i.e. $\forall H$.

Let $F = \{n\}$

$$P(X_n=1) = \frac{\exp \beta H}{2 \cosh \beta H} =: p, \qquad P(X_n=-1) = 1-p =: q.$$

Note that

$$p \leq \frac{1}{2} \quad \text{iff} \quad H \leq 0.$$

Then for $F \in \mathcal{B}(S)$

$$\begin{aligned} P([A,F] &= P(X_n=-1, \ n \in A, \ X_n=1, \ n \in F \setminus A) \\ &= p^{|F \setminus A|} q^{|A|} \end{aligned}$$

With this $\hat{P}$ can be obtained: Let $D \in \mathcal{B}(S)$

$$\hat{P}(D) = \int (-1)^{|D\cap\chi|} dP = \sum_{i=0}^{|D|} \int_{U_i} (-1)^{|D\cap\chi|} dP$$

where

$$U_i = \{\underline{\bar{X}} \in \Omega,\ |D\cap\chi| = i\}$$

so

$$\hat{P}(D) = \sum_{i=0}^{|D|} (-1)^i P(U_i) = \sum_{i=0}^{|D|} (-1)^i \binom{|D|}{i} q^i p^{|D|-i}$$

$$= (p-q)^{|D|} = (2p-1)^{|D|} \tag{5.7}$$

If P would be infinitely divisible, there would exist a P_n for each n with $\hat{P} = (\hat{P}_n)^n$. It is possible to construct infinite temperature models satisfying this: Consider for each n the infinite temperature model with external field H_n such that

$$\beta H_n = \frac{1}{2} \ln \frac{1+\sqrt[n]{1-p}}{1-\sqrt[n]{1-p}}$$

and

$$p_n := \frac{\exp \beta H_n}{2 \cosh \beta H_n}$$

Then

$$p_n = \frac{1+\sqrt[n]{2p-1}}{2} \quad \text{or} \quad (2p_n-1)^n = 2p-1$$

and thus $\hat{P}(D) = (\hat{P}_n(D))^n$ if $P_n \in \mathcal{G}(\Phi_n)$ for $\Phi_n \sim H_n$.

This shows:

Lemma 5-3: The infinite temperature model is infinitely divisible.

(b) One dimensional Ising model: This refers to the case when $S = Z$. For this case it is known that $G(\Phi)$ is a singleton $\forall\ \Phi$ (see [6]). It turns out that in the case of 0-external field the Gibbs state associated with a n.n. potential is infinitely divisible and thus the correlation function has exponential form. Let now $P \in G(\Phi)$ with

$$\Phi_A(\underline{\bar{X}}) = \begin{cases} -J_{n,m}X_nX_m & \text{if} \quad A = \{n,m\},\ n \in \partial\{m\} \\ 0 & \text{otherwise} \end{cases}$$

where $J_{n,m} = J_{m,n} = J$.

From (4) it is known that P is a MRF.

Lemma 5-4: A Markov chain (X_n), $n \in Z$ is a MRF and vice versa.

Proof: A MRF is trivially a Markov chain. So assume now (X_n), $n \in Z$ is a Markov chain. Assume further that

$$P(X_n, \ldots, X_{n+k}) > 0 \quad \forall\ n,k$$

Then

$$P(X_n|X_{n+1},\ldots,X_{n+k}) = \frac{P(X_n,X_{n+1},\ldots,X_{n+k})}{P(X_{n+1},\ldots,X_{n+k})}$$

$$= \frac{P(X_{n+k}|X_n,\ldots,X_{n+k-1})}{P(X_{n+k}|X_{n+1},\ldots,X_{n+k-1})} \frac{P(X_n,\ldots,X_{n+k-1})}{P(X_{n+1},\ldots,X_{n+k-1})}$$

$$= \frac{P(X_{n+k}|X_{n+k-1})}{P(X_{n+k}|X_{n+k-1})} \frac{P(X_n,\ldots,X_{n+k-1})}{P(X_{n+1},\ldots,X_{n+k-1})}$$

so by induction

$$P(X_n|X_{n+1},\ldots,X_{n+k}) = P(X_n|X_{n+1}) \qquad \forall\ k$$

$$P(X_n|X_{n-1},X_{n+1},\ldots,X_{n-k},X_{n+k}) = \frac{P(X_n,X_{n-1},X_{n-2},\ldots,X_{n-k}|X_{n+1},\ldots,X_{n+k})}{P(X_{n-1},\ldots,X_{n-k}|X_{n+1},\ldots,X_{n+k})}$$

$$= \frac{P(X_n,X_{n-1},\ldots,X_{n-k}|X_{n+1})}{P(X_{n-1},\ldots,X_{n-k}|X_{n+1})}$$

$$= \frac{P(X_n,X_{n+1}|X_{n-1},\ldots,X_{n-k})}{P(X_{n+1}|X_{n-1},\ldots,X_{n-k})}$$

$$= P(X_n|X_{n-1},X_{n+1})$$

So (X_n), $n \in Z$ is a MRF.

<u>Remark 5-4</u>: For $A_k^{(n)} := \sigma(X_{n-1},X_{n+1},\ldots,X_{n-k},X_{n+k})$ then

$$P(X_n|A_k^{(n)}) \to P(X_n|A_0^{(n)})[P]$$

where

$$A_0^{(n)} := \sigma(X_m,\ m\in Z\backslash\{n\})$$

since

$$A_k^{(n)} \uparrow A_0^{(n)}$$

By (4) every Gibbs state is a Markov-chain and is completely determined by conditional probabilities of the form (4.2) where

$$p := P(X_n=1|X_{n-1}=1) = \frac{\exp(-\beta J)}{2 \cosh \beta J} = P(X_n=-1|X_{n-1}=-1)$$

and

$$q = P(X_n=1|X_{n-1}=-1) = P(X_n=-1|X_{n-1}=1) = 1-p$$

so the transition matrix of P becomes

$$M = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$$

The equilibrium distribution is $r = (r_+, r_-) = (\frac{1}{2}, \frac{1}{2})$. Therefore:

$$P(X_n=1, X_{n-1}=1) = \frac{\exp(-\beta J)}{4 \cosh \beta J}$$

<u>Lemma 5-5</u>: The convolution of two one dimensional Ising models with prescribed potential is again an Ising model with potential of that kind.

<u>Proof</u>: Let J_1, J_2 be the coupling constants for the Gibbs states and Markov chains P_1, P_2 and let

$$P = P_1 * P_2$$

$\underline{\bar{X}}$ will denote a configuration for the model P_1, $\underline{\bar{Y}}$ for P_2 and

$$Z = \underline{\bar{X}} \cdot \underline{\bar{Y}} \quad \text{for } P$$

For all in the following occurring $n,m \in Z$ it will be assumed that $|n-m| = 1$. Let $\alpha_m, \beta_m \in \{-1,1\}$

$$P(Z_n=1|X_m=\alpha_m, Y_m=\beta_m)$$

$$= \sum_{\alpha_n \in (-1,1)} P(X_n=\alpha_n, Y_n=\alpha_n \; X_m=\alpha_m, Y_m=\beta_m)$$

$$= \sum_{\alpha_n \in (-1,1)} P(X_n=\alpha_n \; X_m=\alpha_m) P(Y_n=\alpha_n \; Y_m=\beta_m)$$

$$= \left(\frac{\exp \beta J_1 \alpha_m}{2 \cosh \beta J_1}\right)\left(\frac{\exp \beta J_2 \beta_m}{2 \cosh \beta J_2}\right) + \left(\frac{\exp -\beta J_1 \alpha_m}{2 \cosh \beta J_1}\right)\left(\frac{\exp -\beta J_2 \beta_m}{2 \cosh \beta J_2}\right)$$

$$= \frac{(\exp \beta J_1 \alpha_m)(\exp \beta J_2 \beta_m) + \exp(-\beta J_1 \alpha_m) \exp(-\beta J_2 \beta_m)}{4 \cosh \beta J_1 \cosh \beta J_2}$$

Now if $\alpha_m = \beta_m \Rightarrow Z_m = 1$ and

$$P(Z_n=1|X_m=\alpha_m, Y_m=\beta_m) = P(Z_n=1|Z_m=1)$$

$$= \frac{\exp \beta(J_1+J_2) + \exp(-\beta(J_1+J_2))}{4 \cosh \beta J_1 \cosh \beta J_2} \tag{5.8}$$

If $\alpha_m = -\beta_m \Rightarrow Z_m = -1$ and

$$P(Z_n=1|Z_m=-1) = \frac{\exp \beta(J_1-J_2) + \exp \beta(J_2-J_1)}{4 \cosh \beta J_1 \cosh \beta J_2} \tag{5.9}$$

From (5.8), (5.9) follows that

$$P(Z_n=1|X_m, Y_m)$$

is a function of $X_m \cdot Y_m = Z_m$

The same calculation goes through for

$$P(Z_n=-1|X_m,Y_m)$$

where $P(Z_n=-1|Z_m=1)$ has form (5.9), $P(Z_n=-1|Z_m=-1)$ form (5.8).

To see that (Z_n), $n \in Z$ is a Markov chain let

$$A_n := \sigma\{(X_k,Y_k) \quad k < n\}$$

$$A_n^Z := \sigma\{(Z_k) \quad k < n\}$$

Since $Z_k = X_k \cdot Y_k$ is a continuous function on $\{-1,1\}^S$, $(Z_k) = (X_k)(Y_k)$, (coordinatewise), is a continuous function and it follows that

$$A_n^Z \subset A_n \ .$$

Let now $A \in \sigma(Z_n)$, $(\Rightarrow A \in \sigma(X_n, Y_n))$

$$\begin{aligned} E(A|A_n^Z) &= E(E(A|A_n)|A_n^Z) \\ &= E(E(A|X_{n-1},Y_{n-1})|A_n^Z) \\ &= E(A|Z_{n-1}) \end{aligned}$$

Since $E(A|X_{n-1},Y_{n-1})$ is a function of Z_{n-1}, independent of A_{n-1}^Z. Therefore (Z_n), $n \in Z$ is a Markov chain.

In order that P has a nearest neighborhood potential with $H = 0$ there has to exist some J such that

$$\frac{\exp \beta J}{2 \cosh \beta J} = \frac{\exp(\beta(J_1+J_2)) + \exp(-\beta(J_1+J_2))}{4 \cosh \beta J_1 \cosh \beta J_2} \tag{5.10}$$

and

$$\frac{\exp(-\beta J)}{2 \cosh \beta J} = \frac{\exp(\beta(J_1-J_2)) + \exp(\beta(J_2-J_1))}{4 \cosh \beta J_1 \cosh \beta J_2} \tag{5.11}$$

Subtracting (5.11) from (5.10) yields

$$\tanh \beta J = (\tanh \beta J_1)(\tanh \beta J_2) \tag{5.12}$$

which has a solution for J.

Note: (i) $J > 0$

(ii) Since $\tanh \beta J < 1$ it follows from (5.12) that $\tanh \beta J < \tanh \beta J_i$, $i = 1,2$ and so $\beta J < \beta J_i$, $i = 1,2$ which means convolution leads to higher temperature.

Corollary 5-2: A one dimensional Ising model with n.n. potential and zero external field is infinitely divisible if $J > 0$.

Proof: For given n choose J_n such that

$$\tanh \beta J = (\tanh \beta J_n)^n$$

Since

$$|\tanh \beta J| < 1,$$

$$\beta J_n = \text{arc} \tanh(\sqrt[n]{\tanh \beta J})$$

exists and $\hat{P} = (\hat{P}_n)^n$ where $P_n \in G(\Phi_n)$, $\Phi_n \sim J_n$.

The above results were first observed in [18].

5-4. The Levy-Khintchin Representation of the One Dimensional Model

Assume an Ising model as in Corollary 5-2; since it is infinitely divisible its correlation admits the Levy-Khintchin representation

$$\hat{P}(D) = \hat{\lambda}_H(D)\ \exp\{\int[(-1)^{|D\cap\chi|}-1]dF\} \quad \forall\ D \in \mathcal{B}(S).$$

By Lemma 5-2 and Corollary 5-1 (if H is taken as there)

$$\hat{P}(D) = \begin{cases} \exp\{\int[(-1)^{|D\cap\chi|}-1]dF\} & \text{if} \quad |D| \text{ is even} \\ 0 & \text{if} \quad |D| \text{ is odd} \end{cases}$$

In order to get the Levy-Khintchin measure, the correlation function will be derived first. Various details left to the reader in [18] are supplied here.

Lemma 5-6: If (X_n), $n \in Z$ is a ± 1 valued Markov-chain with transition matrix

$$M = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$$

$$EX_nX_m = EX_nX_{m-1}EX_{m-1}X_m \ ; \qquad n,m \in Z \tag{5.13}$$

Proof: It is easily checked that

$$E(X_k|X_{k-1}) = X_{k-1}E(X_k\cdot X_{k-1}) \tag{5.14}$$

then

$$E(X_nX_m) = E[X_nE(X_m|X_{m-1},\dots,X_n)]$$

$$= E[X_nE(X_m|X_{m-1})]$$

$$= E[X_nX_{m-1}X_{m-1}E(X_m|X_{m-1})]$$

$$= EX_nX_{m-1}EX_{m-1}X_m$$

<u>Corollary 5-3</u>: If (X_n), $n \in Z$ is as in Lemma 5-6

$$EX_nX_m = EX_nX_{n+1}EX_{n+1}X_{n+2}\dots EX_{m-1}X_m \quad \forall\ n,m \tag{5.15}$$

<u>Proof</u>: Induction, using Lemma 5-6.

<u>Corollary 5-4</u>: If (X_n), $n \in Z$ is as in Lemma 5-6 and $n_1,\dots,n_k$ are any numbers in Z

$$E(X_{n_1}\dots X_{n_k}) = E(X_{n_1}\dots X_{n_k-1})EX_{n_k-1}X_{n_k} \tag{5.16}$$

<u>Proof</u>: Like proof of Lemma 5-6.

<u>Lemma 5-7</u>: For the one dimensional Ising model with n.n. potential and zero external field

$$\hat{P}(D) = \begin{cases} 0 & \text{if } |D| \text{ is odd} \\ (p-q)^{\mu(D)} & \text{if } |D| \text{ is even} \end{cases} \tag{5.17}$$

where

$$\mu(D) = \sum_{i-1}^{k} n_{2i} - n_{2i-1}$$

for

$$D = \{n_1,\dots,n_{2k}\} \quad (n_i < n_{i+1})$$

Proof: The result for $|D|$ odd was already proved. Assume $|D|$ is even, $D = \{n_1,\ldots,n_{2k}\}$. Note that $\hat{P}(D) = EX_{n_1}\ldots X_{n_{2k}}$. The proof follows by induction over k.

$k = 1$: $\quad EX_nX_{n+1} = p - q$

Therefore by Corollary 5-3

$$EX_{n_1}X_{n_2} = (p-q)^{|n_2-n_1|} = (p-q)^{\mu(D)}$$

Assume now (5.17) for $D_0 = \{n_1,\ldots,n_{2k-2}\}$ and let

$D = D_0 \cup \{n_{2k-1},n_{2k}\}$

$$\hat{P}(D) = E\prod_D X_{n_i} = E(\prod_{D_0} X_{n_i}\cdot X_{n_{2k-1}}\cdot X_{n_{2k}-1})EX_{n_{2k}-1}X_{n_{2k}}$$

by Corollary 5-4. Successive application yields:

$$\hat{P}(D) = E(\prod_{D_0} X_{n_i}X^2_{n_{2k-1}})E(X_{n_{2k-1}}X_{n_{2k-1}+1})\ldots E(X_{n_{2k}-1}X_{n_{2k}})$$

$$= (p-q)^{\mu(D_0)}(p-q)^{(n_{2k}-n_{2k-1})} = (p-q)^{\mu(D)}$$

Notes: - For D connected: $\mu(D) = \frac{|D|}{2}$ if $|D|$ even

- $\mathrm{Cov}(X_n,X_m) = (p-q)^{\mu(D)}$ with $D = \{n,m\}$.

Now the Levy-Khintchin form will be derived. For $D \in \mathcal{B}(S)$ with $|D|$ even

$$\hat{P}(D) = \exp\{\int[(-1)^{|D\cap\chi|}-1]dF\}$$

or

$$\mu(D) \ln (p-q) = \int [(-1)^{|D\cap\chi|} - 1] dF. \tag{5.18}$$

The right hand side will only be non zero if $|D\cap\chi|$ is odd. Then

$$\mu(D) \ln (p-q) = -2F(|D\cap\chi| \text{ is odd})$$

$$= -2 \sum_{\substack{\bar{\underline{X}}\in\Omega \\ |D\cap\chi| \text{ is odd}}} F(\bar{\underline{X}})$$

Let now $\bar{\underline{X}}_{(k)} \in \Omega$ be a configuration with

$$X_{(k)_i} = \begin{cases} -1 & i \leq k \\ 1 & i > k \end{cases}$$

In $\{\bar{\underline{X}}\in\Omega, |D\cap\chi| \text{ is odd}\}$ there are $\mu(D)$ configurations of the form $\bar{\underline{X}}_{(k)}$. If F is now restricted to these configurations with equal mass distribution one gets

$$\mu(D) \ln (p-q) = -2\mu(D) F(\bar{\underline{X}}_{(k)}) \qquad \text{for some fixed } k$$

so choose

$$F(\bar{\underline{X}}) = \begin{cases} -\ln(\sqrt{p-q}) & \text{if } \bar{\underline{X}} = \bar{\underline{X}}_{(k)} \quad \forall k \\ 0 & \text{otherwise} \end{cases} \tag{5.19}$$

F satisfies the conditions of the Levy-Khintchin measure in Theorem 5-1 and from the uniqueness of the representation follows that (5.18) is the Levy-Khintchin measure for the one dimensional Ising model with n.n. potential and 0 external field.

5-5. Characterization of Tail Triviality by the Correlation Function

Let P be a Gibbs state on an arbitrary countable set S. The correlation function is then given by the Fourier-Stiltjes transform of P and is a function on $\mathcal{B}(S)$:

$$\hat{P}(D) = E\gamma_D \tag{5.20}$$

where $\gamma_D \in \Omega$, $\gamma_D(\underline{\bar{X}}) = (-1)^{|D \cap \chi|} = \prod_D X_n$.

The continuous characters on Ω can now be viewed as ± 1 valued random variables with mean

$$E\gamma_D = \hat{P}(D)$$

and covariance

$$\begin{aligned} \mathrm{Cov}\ \gamma_A\gamma_B &= E\gamma_A\gamma_B - E\gamma_A E\gamma_B \\ &= \hat{P}(A\Delta B) - \hat{P}(A)\hat{P}(B), \quad A,B,D \in \mathcal{B}(S) \end{aligned}$$

With this it is possible to characterize tail triviality of Gibbs states in terms of correlation. In (3) it was stated that (3.14) is equivalent to tail triviality. The idea is now to write (3.14) in terms of the correlation:

For $f \in C(\Omega)$

$$f(x) = \sum_{D \in \mathcal{B}(S)} \gamma_D(x)\hat{f}(D)$$

and so

$$\int fgdP - \int fdP\int gdP = \int \sum_A \sum_B \gamma_A\gamma_B\hat{f}(A)\hat{g}(B) - (\int \sum_A \gamma_A\hat{f}(A))(\int \sum_B \gamma_B\hat{g}(B))$$

$$= \sum_A \sum_B \hat{f}(A)\hat{g}(B)\int\gamma_A\gamma_B - (\sum_A \hat{f}(A)\int\gamma_A)(\sum_B \hat{g}(B)\int\gamma_B)$$

$$= \sum_{A,B} \hat{f}(A)(\int\gamma_A\gamma_B - \int\gamma_A\int\gamma_B)\hat{g}(B), \quad \text{integrals w.r. to } P$$

$$= \sum_{A,B} \hat{f}(A) \operatorname{Cov} \gamma_A\gamma_B\hat{g}(B); \quad A,B \in \mathcal{B}(S) \tag{5.21}$$

This gives rise to

<u>Lemma 5-8</u> (Waymire): A random field P on Ω is tail trivial if $\forall f \in C(\Omega)$, $\exists F \in \mathcal{B}(S)$ such that

$$\left| \sum_{A,B} \hat{f}(A) \operatorname{Cov}(\gamma_A\gamma_B)\hat{g}(B) \right| \leq \| g \|_\infty \tag{5.22}$$

$\forall g \in C(\Omega)$, $g \in \mathcal{A}_{F^c}$.

<u>Example 5-1</u>: Infinite temperature model with 0-external field. In this case (X_n), $n \in S$ is a family of i.i.d. Bernoulli $(\frac{1}{2})$ distributed random variables. Tail triviality follows then by Kolmogoroff's 0-1 law, but also by Lemma 5-8. Let $f \in C(\Omega)$; by the Riemann Legesgue Lemma there exists an $F \in \mathcal{B}(S)$ such that

$$|\hat{f}(D)| < 1 \qquad \forall D \subset F^c, D \in \mathcal{B}(S). \tag{5.23}$$

Let $g \in \mathcal{A}_{F^c}$, $g \in C(\Omega)$.

For $A, B \in \mathcal{B}(S)$

$$\mathrm{Cov}(\gamma_A, \gamma_B) = E \prod_{A\Delta B} X_n - E \prod_A X_n E \prod_B X_n$$

$$= \prod_{A\Delta B} EX_n \{1 - \prod_{A\cap B} (EX_n)^2\}$$

$$= \begin{cases} 0 & \text{for} \quad A \neq B \\ 1 & \text{otherwise} \end{cases} \tag{5.24}$$

If $B \cap F \neq \emptyset$ decompose B into $B = B_1 \cup B_2$; $B_1 \subset F$, $B_2 \subset F^c$

$$\hat{g}(B) = \int \gamma_B g dP = \int \gamma_{B_1} \gamma_{B_2} g dP = \int \gamma_{B_1} dP \int \gamma_{B_2} g dP = 0 \tag{5.25}$$

So by (5.23-5.25):

$$\Big| \sum_{A,B} \hat{f}(A) \, \mathrm{Cov}(\gamma_{A_1} \gamma_B) \hat{g}(B) \Big| = \Big| \sum_{A \neq \emptyset} \hat{f}(A) \hat{g}(A) \Big|$$

$$= \Big| \sum_{A \subset F^c} \hat{f}(A) \hat{g}(A) \Big|$$

$$< \sum_{A \subset F^c} |\hat{g}(A)|$$

$$\sum_{A \subset F^c} |\hat{g}(A)| \leq \sum_A |\gamma_A(\underline{\bar{X}_+}) \hat{g}(A)| \leq \| g \|_\infty$$

Since $\gamma_A(\underline{\bar{X}_+}) = 1 \ \forall \ A \in \mathcal{B}(S)$ and therefore P is tail trivial by Lemma 5-8.

Chapter 6. GIBBS STATES ON THE BETHE LATTICE

Gibbs states on the Bethe Lattice or tree T_N were defined in Chapter 4. Only n.n. potentials, i.e. symmetric, invariant pair potentials with the n.n. property, are considered; so Gibbs states are also MRF's. Within the class of MRF's on Ω there exists a subclass for which calculations simplify and which will be the main object of study in the following. This subclass will be called the set of Markov-chains on Ω and can be defined as follows:

Let M be a stochastic matrix

$$M = \begin{pmatrix} M(1,1) & M(1,-1) \\ M(-1,1) & M(-1,-1) \end{pmatrix} = \begin{pmatrix} s & 1-s \\ 1-t & t \end{pmatrix} \qquad 0 < t,s, < 1 \tag{6.1}$$

and $r = (r_+,r_-)$ the unique invariant distribution.

A random field μ can then be defined by cylinder set probabilities:

$$\mu([A,F]) = \pi(X_1) \prod_{i=2}^{|F|} M(X_{(i)},X_i) \tag{6.2}$$

where $A \subset F \in \mathcal{B}(T_N)$ and

$$\pi(X_1) = \begin{cases} r_+ & \text{if} \quad X_1=1 \\ r_- & \text{if} \quad X_1=-1 \end{cases}$$

This defines a projective family of finite dimensional distributions, as can be seen by an easy summation and therefore gives rise to a unique state on Ω, by the Kolmogoroff construction exhibited in Chapter 3.

Note: (6.2) is independent of the labeling, as long as the labeling introduced in Chapter 4 is used.

States defined in this way by a stochastic matrix (6.1) will be called Markov chains, since they behave like one dimensional Markov chains along paths.

Lemma 6-1: Every Markov chain is a MRF and hence a Gibbs state for a certain n.n. potential Φ on Ω.

Proof: Note that finite dimensional probabilities are positive. Let $n \in F \subset \mathcal{B}(T_N)$; $\mathcal{A}_{F\setminus\{n\}} = \sigma(X_k, k\ F\setminus\{n\})$

$$P(X_n=1|\mathcal{A}_{F\setminus\{n\}}) = P(X_n=1|X_k, k\in F\setminus\{n\})$$

$$= \frac{P(X_n=1, X_k, k\in F\setminus\{n\})}{P(X_k, k\in F\setminus\{n\})}$$

$$= \prod_{i\in\{n\cup\partial\{n\}\}} M(X_{(i)}, X_i)$$

i.e. $P(X_n=1|\mathcal{A}_{F\setminus\{n\}})$ depends only on the sites in $\partial\{n\}$ $\forall\ F \in \mathcal{B}(T_N)$ with $n \in F$. Let now $F \uparrow T_N$ in an increasing way, to get the result.

Note: Markov-chains are homogeneous, i.e. independent of the position of measurable sets on the tree. This follows from the fact that MRF's are Gibbs states with invariant potential and from

$$r_+(1-s) = r_-(1-t)$$

If M denotes the set of all Markov-chains

$$M \subset \bigcup_{\Phi} G(\Phi)$$

the set of all Gibbs states with n.n. potential, i.e. the set of all MRF's on Ω.

As it was stated in Chapter 4, to each $\mu \in M$ corresponds a unique n.n. potential such that

$$\mu \in G(\Phi)$$

which means every stochastic matrix of form (6.1) determines a potential through the probabilities (6.2) (see also [16]). The question which arises now is, does there exist an element $\mu \in M$ with $\mu \in G(\Phi)$ for a given n.n. potential Φ and is it uniquely determined by this potential. The answer in the case $N = 2$ is, that for every n.n. potential Φ

$$|M \cap G(\Phi)| = 1,2 \text{ or } 3.$$

If $|M \cap G(\Phi)| > 1$, phase transition has occurred. This will be shown below.

Two Gibbs states μ_1, μ_2 belong to the same $G(\Phi)$ if their corresponding conditional probabilities (4.2) are equal. These will be calculated next.

Let

$$M = \begin{pmatrix} s & 1-s \\ 1-t & t \end{pmatrix} \quad \text{and} \quad P \sim M$$

Then

$$\frac{1}{P(X_n=1 \mid k \text{ neighbors are } -1)} - 1$$

$$= \frac{P(k \text{ neighb. are } -1) - P(X_n=1, k \text{ neighb. are } -1)}{P(X_n=1, k \text{ neighb. are } -1)}$$

$$= \frac{P(X_n=-1, k \text{ neighb. are } -1)}{P(X_n=1, k \text{ neighb. are } -1)}$$

$$= \frac{r_- t^k (1-t)^{N-k}}{r_+ (1-s)^k s^{N-k}} = \frac{t^k (1-s)(1-t)^{N-k}}{(1-s)^k s^{N-k}} \tag{6.3}$$

and therefore:

$$\alpha_k^{(n)} = P(X_n=1 \mid k \text{ neighb. are } -1) = \frac{r_+ (1-s)^k s^{N-k}}{r_- t^k (1-t)^{N-k} + r_+ (1-s)^k s^{N-k}}$$

$$= \frac{(1-s)^k s^{N-k}}{t^k (1-t)^{N-k-1} (1-s) + (1-s)^k s^{N-k}} \qquad \forall\, n \in T_N \tag{6.4}$$

If for a Markov chain $\tilde{P}$ with

$$\tilde{M} = \begin{pmatrix} \tilde{s} & 1-\tilde{s} \\ 1-\tilde{t} & \tilde{t} \end{pmatrix}$$

$\tilde{\alpha}_k = \alpha_k \quad \forall\, k = 1 \ldots N$, then P and $\tilde{P}$ belong to the same $\mathcal{G}(\Phi)$.

Note also that

$$s = P(X_n=1 \mid X_k=1, k \in \partial\{n\}) = \frac{\exp \beta J+H}{2 \cosh \beta J+H}$$

$$t = P(X_n=-1 \mid X_k=-1, k \in \partial\{n\}) = \frac{\exp \beta J-H}{2 \cosh \beta J-H}$$

and hence

$$\frac{s}{1-s} \frac{t}{1-t} = \exp 4\beta J, \tag{6.5}$$

So if $P,\tilde{P} \in G(\Phi)$, $P,\tilde{P} \in M$

$$\frac{s}{1-s}\,\frac{t}{1-t} = \frac{\tilde{s}}{1-\tilde{s}}\,\frac{\tilde{t}}{1-\tilde{t}}$$

$$\Longleftrightarrow \qquad \frac{s(1-\tilde{s})}{(1-s)\tilde{s}} = \frac{(1-t)\tilde{t}}{t(1-\tilde{t})}\,. \tag{6.6}$$

The following calculations investigate for T_2 in which cases

$$|M \cap G(\Phi)| = 1,2 \text{ or } 3$$

Let $P,\tilde{P} \in M$ satisfy $\alpha_k = \tilde{\alpha}_k$, $k = 1,2,3$ and (6.6). From $\alpha_0 = \tilde{\alpha}_0$ and (6.3) results

$$\frac{(1-t)^{N-1}(1-s)}{s^N} = \frac{(1-\tilde{t})^{N-1}(1-s)}{\tilde{s}^N}$$

$$\Longleftrightarrow \qquad \frac{\tilde{s}}{s}\,\frac{1-s}{1-\tilde{s}} = \left[\frac{1-\tilde{t}}{1-t}\,\frac{s}{\tilde{s}}\right]^{N-1} \tag{6.7}$$

Substituting (6.6) into (6.7) yields:

$$\frac{t(1-\tilde{t})}{(1-t)\tilde{t}} = \left[\frac{s\,\dfrac{t(1-\tilde{t})}{(1-t)\tilde{t}} + 1 - s}{(1-t)\,\dfrac{t(1-\tilde{t})}{(1-t)\tilde{t}} + t}\right]^{N-1}$$

or

$$x^{N-1} = \left[\frac{sx^{N-1} + 1 - s}{(1-t)x^{N-1} + t}\right]^{N-1} \tag{6.8}$$

where

$$x = \left[\frac{t(1-\tilde{t})}{(1-t)\tilde{t}}\right]^{1/N-1} > 0$$

then

$$\frac{\tilde{t}}{1-\tilde{t}} = x^{-N+1} \frac{t}{1-t} , \qquad \frac{\tilde{s}}{1-\tilde{s}} = \frac{s}{1-s} x^{N-1} \tag{6.9}$$

If Equation (6.8) has more than one positive, real solution then

$$|M \cap G(\Phi)| > 1$$

and phase transition has occurred among the class of Markov chains.

Analysis of (6.8): $x = 1$ is always a solution to (6.8) verifying that

$$|M \cap G(\Phi)| \geq 1 \quad \forall\ \Phi, \Phi \text{ n.n.}$$

In this case $\tilde{s} = s$, $\tilde{t} = t$ as it can be seen from (6.9).

For $N = 1$ (6.8) reduces to

$$(1-t)x^2 + (t-s)x - (1-s) = 0$$

which has only $x = 1$ as a positive real solution.

(6.8) is equivalent to

$$x - \frac{sx^{N-1} + (1-s)}{(1-t)x^{N-1} + t} = 0$$

$$\Longleftrightarrow \quad (1-t)x^N + tx - sx^{N-1} - (1-s) = 0 \tag{6.10}$$

Since $x = 1$ is a solution to (6.10) divide by $(x-1)$ to get

$$(1-t)x^{N-1} + (1-t-s)x^{N-2} + \ldots + (1-t-s)x^2 + (1-t-s)x + (1-s) = 0 \tag{6.11}$$

which becomes for $N = 2$

$$(1-t)x^2 + (1-t-s)x + (1-s) = 0 \tag{6.12}$$

and has solutions

$$x = \frac{(s-1+t) \pm \sqrt{(s-1+t)^2-4(1-t)(1-s)}}{2(1-t)} \tag{6.13}$$

Therefore, for $N = 2$, (6.8) has three solutions if

$$(s-1+t)^2 - r(1-t)(1-s) > 0 \tag{6.14}$$

and none of the solutions (6.13) equals 1; in this case the solutions are real and positive. (6.14) is equivalent to

$$(s-t)^2 + 2(s+t) > 3. \tag{6.15}$$

If $(s-t)^2 + 2(s+t) = 3$, (6.8) has two real, positive solutions if the solution in (6.13) is not equal to 1. In all the other cases, $x = 1$ is the only real, positive solution to (6.8), except when (6.15) holds and one solution in (6.13) is 1. Assume that (6.15) holds with a '$\geq$', i.e. (6.13) is real and positive:

(i) If (6.12) factors into $(x-1)^2$, (6.13) will be 1 in both cases. This occurs for $s = t = \frac{3}{4}$ and in this case (6.8) has only 1 positive, real solution.

(ii) If $(1-t) + (1-t-s) + (1-s) = 0 \iff s + t = \frac{3}{2}$ then one solution to (6.12) is 1.

Summary:

(a) For $N = 1$, $|M \cap G(\Phi)|$ will always be one. Since this case refers to the one dimensional model, this result confirms again that in one dimension there occurs no phase transition for n.n. potentials since $M = \bigcup_{\Phi} G(\Phi)$, for n.n. potentials.

(b) $N = 2$:

(i) $|M \cap G(\Phi)| = 1$ for $(s-t)^2 + 2(s+t) < 3$ and $s = t = \frac{3}{4}$

(ii) $|M \cap G(\Phi)| = 2$ for $(s-t)^2 + 2(s+t) = 3$, $s,t \neq (\frac{3}{4},\frac{3}{4})$ and for $s + t = \frac{3}{2}$, $s,t \neq (\frac{3}{4},\frac{3}{4})$

(iii) $|M \cap G(\Phi)| = 3$ for $(s-t)^2 + 2(s+t) > 3$.

Note that for the case $s + t < 1$, referred to as the repulsive case, there is no phase transition among the class of Markov-chains.

The fact that phase transition occurs within the class of Markov chains will be taken as a reason to further study this class of Gibbs states.

Lemma 6-2: Isotherms for Markov chain Ising models are constants for $\frac{s}{1-s}\frac{t}{1-t}$. An attractive Markov chain P_1 and a repulsive P_2 represent the same temperature iff

$$s_1 + t_1 = 1 + s_2 + t_2$$

Proof: This follows from (6.5) and the fact that a repulsive coupling constant J is negative.

6-1. Infinite Divisibility of Markov Chains on the Tree

The use of infinite divisibility was already pointed out in Chapter 5. It was shown that the convolution of two one dimensional n.n. Ising models with no external field is again of that kind. As will be shown this is also true for the subclass of Markov chains with no external field on the tree. The way to prove this is slightly different than in one dimension (compare Lemma 5-5). The results here are new.

Lemma 6-3: Let $\underline{\bar{X}}$, $\underline{\bar{Y}}$ be two Markov chain Ising models with no external field and Z their convolution. Then

$$P(Z_n=\varepsilon_n | X_k=\alpha_k,\ Y_k=\beta_k,\ k\in\partial\{n\}), \qquad n\in T_N$$

is not a function of $X_k \cdot Y_k$, $k \in \partial\{n\}$.

Proof: This is easily checked by writing out the probabilities and choosing α_k, β_k appropriately.

Note: In one dimension it was possible to restrict to only one neighbor to prove the converse of Lemma 6-3; this is not possible on the tree.

Theorem 6-1: The convolution of two Markov chains on the tree with 0-external field is again of that kind.

Note: $r_+ = r_- = \frac{1}{2}$ for $H = 0$.

The proof is prepared by the following results:

Lemma 6-4: If Z is a MRF on $\Omega = \{1,-1\}^{T_N}$, $D \in \mathcal{B}(T_N)$ connected and

$$P(Z_n=1,\ n\in D) = r_+ s^{|D|-1}$$

$$P(Z_n=-1,\ n\in D) = r_- t^{|D|-1}$$

for some $s,t \in (0,1)$ and (r_+,r_-) with $r_+ + r_- = 1$ and $r_+(1-s) = r_-(1-t)$, then

$$P(Z_n=\varepsilon_n,\ n\in D) = \pi(Z_1) s^{e_+}(1-s)^{o_+} t^{e_-}(1-t)^{o_-}$$

where

$$\pi(Z_1) = \begin{cases} r_+ & \text{if} \quad Z_1 = 1 \\ r_- & \text{if} \quad Z_1 = -1 \end{cases}$$

e_+ = # of bonds from +1 to +1

o_+ = # of bonds from +1 to -1

e_- = # of bonds from -1 to -1

o_- = # of bonds from -1 to +1

Proof: Case 1: $s = t =: p$. Assume $D \in \mathcal{B}(T_N)$, $|D| = 2$; then the assertion holds. Assume now the assertion holds for $D \in \mathcal{B}(T_N)$ with $|D| = n - 1$, D connected. Let now $D = D_0 \cup \{m\}$, $D_0 \in \mathcal{B}(T_N)$, $|D_0| = n - 1$, D connected, where m is such that $m > k$, $\forall\, k \in D_0$

$$\begin{aligned} P(Z_k=\varepsilon_k,\ k\in D_0,\ Z_m=\varepsilon_m) &= P(Z_m=\varepsilon_m | Z_k=\ _k,\ k\in D_0) P(Z_k=\varepsilon_k,\ k\in D_0) \\ &= P(Z_m-\varepsilon_m | Z_{k_0}=\varepsilon_{k_0},\ k_0\in\partial\{m\}) P(Z_k=\varepsilon_k,\ k\in D_0) \\ &= p^{e_+ + e_-} q^{o_+ + o_-} \end{aligned}$$

Note that m has exactly one neighbor k_0 in D_0. The proof for $s \neq t$ works the same way.

<u>Lemma 6-5</u>: Let $D \in \mathcal{B}(T_N)$, connected and

$$\Gamma = \{\bar{\underline{X}} \in \Omega,\ X_1 \text{ fixed},\ 1 \in D\}$$

where 1 is the lowest label in D. Then there are $\binom{|D|-1}{k}$ configurations in Γ with k odd bonds, where a bond is odd if it connects two sites with different values.

<u>Proof</u>: For $D = 2$ obvious. Assume now the assertion holds for D_0. $|D_0| = n - 1$, D_0 connected. Let $D = D_0 \cup \{n\}$ connected, where again $n > k$, $\forall\, k \in D_0$. The bond from n to (n), where $(n) \in D_0$ is either odd or even:

(i) Even: To get k odd bonds in D, they have to be among D_0, by assumption there are $\binom{|D|-2}{k}$ configurations in Γ with this property.

(ii) Odd: To get k odd bonds in D, D_0 has to contain $k-1$. There are $\binom{|D|-2}{k-1}$ configurations in Γ with this property.

Together there are $\binom{|D|-2}{k} + \binom{|D|-2}{k-1} = \binom{|D|-1}{k}$ configurations in Γ with k odd bonds.

<u>Proof of Theorem 6-1</u>: Let P_x, P_y be two Markov chain Ising models and $P_Z = P_x * P_y$. Let $D \in \mathcal{B}(T_N)$ be connected. (Z_n), $n \in T_N$ is a MRF: This is the case when $P(Z_n=1 | Z_k=\varepsilon_k,\ k \in D\backslash\{n\})$, where $n \in D$, is a function of (Z_k), $k \in \partial\{n\}$. For $|D| = 2$ this is obvious.

Assume now $P(Z_n=1|Z_k=\varepsilon_k,\ k\in D\backslash\{n\})$ is a function of (Z_k), $k \in \partial\{n\}$ when $|D| = j - 1$. Let then $|D| = j$ and m denote the highest label in D. Note that m has exactly one neighbor in D. W.L.O.G. $m \notin \partial\{n\}$.

$$P(Z_n=1|Z_k=\varepsilon_k,\ k\in D\backslash\{n\})$$

$$= \frac{P(Z_n=1,\ Z_k=\varepsilon_k,\ k\in D\backslash\{n\})}{P(Z_k=\varepsilon_k,\ k\in D\backslash\{n\})}$$

$$= \frac{\sum\limits_{\alpha_k,k\in D} P(X_n=\alpha_n,\ X_k=\alpha_k)P(Y_n=\alpha_n,\ Y_k=\alpha_n\cdot\varepsilon_k)}{\sum\limits_{\alpha_k,k\in D\backslash\{n\}} P(X_k=\alpha_k)P(Y_k=\alpha_k\cdot\varepsilon_k)}$$

$$= \frac{\{ab+(1-a)(1-b)\}\sum\limits_{\alpha_k,k\in D\backslash\{m\}} P(X_n=\alpha_n,\ X_k=\alpha_k)P(Y_n=\alpha_n,\ Y_k=\alpha_k\cdot\varepsilon_k)}{\{ab+(1-a)(1-b)\}\sum\limits_{\alpha_k,k\in D\backslash\{m,n\}} P(X_k=\alpha_k)P(Y_k=\alpha_k\cdot\varepsilon_k)}$$

$$= P(Z_n=1|Z_k=\varepsilon_k,\ k\in D\backslash\{n,m\})$$

which is, by assumption, a function of (Z_k), $k \in \partial\{n\}$.

$$a \in \{p_x, 1-p_x\}, \quad b \in \{p_y, 1-p_y\}$$

The Markov property for (Z_n), $n \in T_N$ follows then by applying a martingale limit theorem. To see that (Z_n), $n \in T_N$, is also a Markov chain:

$$P(Z_n=1,n\epsilon D) = \sum_{(\varepsilon_n),n\epsilon D} P_x(X_n=\varepsilon_n,n\epsilon D)P_y(Y_n=\varepsilon_n,n\epsilon D)$$

$$= \sum_{(\varepsilon_n),n\epsilon D} \frac{1}{2} \prod_{i=2}^{|D|} M_x(\varepsilon_{(i)},\varepsilon_i) \frac{1}{2} \prod_{i=1}^{|D|} M_y(\varepsilon_{(i)},\varepsilon_i)$$

$$= \sum_{\varepsilon_1=1} \frac{1}{4} \prod_{i=2}^{|D|} M_x(\varepsilon_{(i)},\varepsilon_i)M_y(\varepsilon_{(i)},\varepsilon_i)$$

$$+ \sum_{\varepsilon_1=-1} \frac{1}{4} \prod_{i=2}^{|D|} M_x(\varepsilon_{(i)},\varepsilon_i)M_y(\varepsilon_{(i)},\varepsilon_i)$$

These summations are each summation over Γ as introduced in Lemma 6-5 and thus:

$$P(Z_n=1,n\epsilon D) = \frac{1}{4} \sum_{k=0}^{|D|-1} \binom{|D|-1}{k} (p_x p_y)^{|D|-k}(q_x q_y)^k$$

$$+ \frac{1}{4} \sum_{k=0}^{|D|-1} \binom{|D|-1}{k} (p_x p_y)^{|D|-k}(q_x q_y)^k$$

$$= \frac{1}{2} (p_x p_y + q_x q_y)^{|D|-1} = \frac{1}{2} p_Z^{|D|-1} \tag{6.16}$$

where

$$p_Z = p_x p_y + q_x q_y \ .$$

$$P(Z_n=-1,n\epsilon D) = \sum_{(\varepsilon_n),n\epsilon D} P(X_n=\varepsilon_n,n\epsilon D)P(Y_n=\varepsilon_n,n\epsilon D)$$

$$= \sum_{(\varepsilon_n),n\epsilon D} \frac{1}{4} \prod_{i=2}^{|D|} M_x(\varepsilon_{(i)},\varepsilon_i)M_y(-\varepsilon_{(i)},-\varepsilon_i)$$

$$= \frac{1}{2} p_Z^{|D|-1}$$

since $M_y(-\varepsilon_{(i)},-\varepsilon_i) = M_y(\varepsilon_{(i)},\varepsilon_i)$. Therefore Lemma 6-4 applies and proves the assertion.

Corollary 6-1: The attractive Markov chain Ising model with 0-external field is infinitely divisible.

Proof: If J denotes the coupling constant of the convolution of two Markov chains with 0 external field and J_1, J_2 respectively, their coupling constants then (5.10) through (5.12) hold and so the proof is the same as of Corollary 5-2.

Remarks 6-1:

(i) For a given model with $p > \frac{1}{2}$ (H = 0), there exists an attractive and a repulsive model, representing the same temperature, whose convolution yields the given model: If p_1, p_2 are chosen to be

$$p_1 = \frac{1}{2} + \frac{\sqrt{2p-1}}{2}$$

$$p_2 = \frac{1}{2} - \frac{\sqrt{2p-1}}{2}$$

p_1 represents an attractive, p_2 a repulsive Markov chain Ising model, both representing the same temperature by Lemma 6-2. The convolution yields the given model.

(ii) The convolution of two attractive or two repulsive models is attractive and the convolution of an attractive with a repulsive model is repulsive. This is easily checked by (6.16).

Corollary 6-3: A repulsive Markov-chain Ising model is not infinitely divisible with Markov-chains as factors (still $H = 0$).

Proof: From Remark 6-1 (ii) follows that the convolution of a Markov chain Ising model with itself is never repulsive.

6-2. Notes to the Correlation on T_N

The correlation function for n.n. Gibbs states, being of Markov chain type corresponding to attractive potentials admits the representation

$$\hat{P}(D) = \hat{\lambda}_H(D) \exp\{\int(\prod_D X_n - 1)dF\}, \tag{6.17}$$

$D \in \mathcal{B}(T_N)$, H a compact subgroup and λ the Haar measure on H.

This fact, resulting from the infinite divisibility of the specified states, will be used to describe the general form of the correlation. For the following only Markov chain Ising models with attractive potential are considered; these also detect phase transition whereas repulsive models do not.

The fact that Markov-chains on trees act like one dimensional Markov chains on paths, admits the application of Lemma 5-6 and Corollaries 5-3 and 5-4. Denote by $\bar{D}$ the smallest, connected subset of T_N containing $D \in \mathcal{B}(T_N)$. $\bar{D} \in \mathcal{B}(T_N)$.

Lemma 6-6: Let $D \in \mathcal{B}(T_N)$ be an arbitrary connected subset with $|D| = n$, $D = \{1,2\ldots n\}$. Then

$$E(X_1 \cdot \ldots X_n) = E(X_1 \cdot \ldots X_{n-1} \cdot X_{(n)}) E(X_{(n)} \cdot X_n) \tag{6.18}$$

If $(n) = n-1$:

$$E(X_1 \cdot \ldots X_n) = E(X_1 \cdot \ldots X_{n-2}) E(X_{n-1} \cdot X_n)$$

Proof:

$$E(X_1 \cdot \ldots X_n) = E(E(X_1 \cdot \ldots X_n | A_{\{n\}^c}) = E(X_1 \cdot \ldots X_{n-1} E(X_n | X_{(n)}))$$
$$= E(X_1 \cdot \ldots X_{n-1} \cdot X_{(n)}) E(X_{(n)} \cdot X_n)$$

using (5.12) and the fact that n has exactly 1 neighbor in D.

Remark 6-2: If D is not connected, (6.17) holds for (n) the unique neighbor of n in $\bar{D}$; $\bar{D}$ is labelled in the usual way, D has the induced labelling.

In the one dimensional case the correlation had the form

$$\hat{P}(D) = (p-q)^{\mu(D)} \quad \text{for} \quad D \in \mathcal{B}(T_2), \ |D| \text{ even}$$

and

$$\hat{P}(D) = 0 \quad \text{for} \quad |D| \text{ odd}$$

where $\mu(D)$ is given by (5.14).

In this case μ was a counting function, counting the number of bonds between succeeding sites in the subset. The idea for the tree case is to get a similar counting function, which, starting from the site with highest label, counts the number of bonds to the nearest site and continues this way. The problem here is, that this nearest neighbor may not be unique if the corresponding subset is disconnected. Therefore the connected case will be treated first.

<u>Proposition 6-1</u>: Let $\bar{X}$ be a Markov chain Ising model with 0-external field and $p > \frac{1}{2}$. Let $D \in \mathcal{B}(T_2)$ be connected. Then

$$\hat{P}(D) = \begin{cases} (p-q)^{\mu(D)} & |D| \text{ even} \\ 0 & |D| \text{ odd} \end{cases}$$

where

$$\mu(D) = \sum_{\substack{i \in D_k^* \\ k>i}} \{\# \text{ of bonds between } (i)^* \text{ and } i\}$$

and the sum starts at the highest label of D, $(i)^*$ denotes the unique nearest site of i in

$$D_i^* = D \setminus \{k,(k)^*,k>i\}$$

Notes: - $(n)^* = (n)$ and $D_n^* = D$ if n is the highest label in D,

$D_{n-1}^* = D \setminus \{n,(n)\}$

- D_i^* is in general not connected

- i is the highest label in D_i^* if $i \in D_i^*$.

<u>Proof</u>: From the way of labeling and the connectedness follows that each $i \in D$ has a unique $(i)^*$ in D_i^*.

$$\hat{P}(D) = E(\prod_D X_k) = E(E(\prod_D X_k | A_{\{n\}^c}))$$

$$= E(\prod_{D\setminus\{n\}} X_k E(X_n | X_{(n)}))$$

$$= E(\prod_{D\setminus\{n\}} X_k \cdot X_{(n)}) \cdot E(X_{(n)} \cdot X_n)$$

$$= E(\prod_{D\setminus\{n,(n)\}} X_k) \cdot E(X_{(n)} \cdot X_n) =$$

$$= E(\prod_{D^*_{n-1}} X_k) \cdot E(X_{(n)} \cdot X_n)$$

$$= \begin{cases} E(\prod_{D^*_{n-1}} X_k) \cdot E(X_{(n)} \cdot X_n) & \text{if} \quad (n) \neq n-1 \\ E(X_1 \ldots X_{n-2}) \cdot (p-q) & \text{if} \quad (n) = n-1 \end{cases}$$

In general:

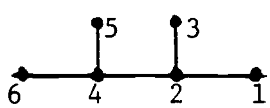
$$E(\prod_{D^*_i} X_k) = E(\prod_{D^*_{i-1}} X_k) E(X_{(i)^*} X_i)$$

from where the assertion follows. Note that if $i-1 = (i)^*$ then $D^*_{i-1} = D^*_{i-2}$ and that

$$E(X_{(i)^*} X_i) = (p-q)^{\alpha} \; ; \qquad \alpha = \{\# \text{ of bonds between } i \text{ and } (i)^*\}$$

which follows from Corollary 5.4.

<u>Example 6-1</u>: Let $D = \{1,\ldots,6\}$ have the form



Both the formal calculation and Proposition 6-1 yield

$$\hat{P}(D) = (p-q)^5$$

The calculation of the correlation for disconnected subsets works principally the same way, but is much more involved and will therefore be skipped. However some estimates will be given.

Let P represent an attractive Markov chain Ising model with 0-external field. If λ_H in (6.17) is taken to be the same as in Chapter 5, then

$$\hat{P}(D) = \exp\{\int(\gamma_D-1)dF\}$$

for $D \in \mathcal{B}(T_N)$, $|D|$ even

$$\hat{P}(D) = \exp\{\int_{\Gamma_D} (\gamma_D-1)dF\}$$

where

$$\Gamma_D = \{\underline{\bar{X}}\epsilon\Omega,\ \prod_D X_n=-1\}$$

and

$$\ln \hat{P}(D) = -2F(\Gamma_D).$$

Let now $D = D_1 \cup D_2$, $D_1 \cap D_2 = \emptyset$, $|D_1|$, $|D_2|$ even, then

$$\Gamma_D = \Gamma_{D_1} \Delta \Gamma_{D_2}$$

and

$$\ln \hat{P}(D) = -2\{F(\Gamma_{D_1})+F(\Gamma_{D_2})-F(\Gamma_{D_1}\cap\Gamma_{D_2})\} .$$

Note that

$$|\tfrac{1}{2} \hat{P}(D)| = |F(\Gamma_D)| < \infty.$$

Therefore:

$$\hat{P}(D) = \hat{P}(D_1)\hat{P}(D_2) \exp\{2\int_{\Gamma_{D_1}\cap\Gamma_{D_2}} dF\} \tag{6.19}$$

Lemma 6-7: For $D \in \mathcal{B}(T_2)$, $|D| > 2$, even, $\exists\, D_1, D_2 \subset D$ such that

$$\hat{P}(D) = \hat{P}(D_1)\hat{P}(D_2) \tag{6.20}$$

Proof: Assume $|D| = n$. For connected D the assertion follows from Proposition 6-1. Let D be disconnected and let n be the site with highest label and longest path to the lowest label in D.

Case I: n has a unique nearest site in D, $\tilde{n}$. Let $D_1 = \{n, \tilde{n}\}$, $D_2 = D \setminus D_1$ and the assertion follows from Remark 6-2.

Case II: n does not have a unique nearest site; $\exists$ sites m,k with the same number of bonds to n. Assume $m > k$ and let $D_1 = \{n,m\}$

$$\begin{aligned}
\hat{P}(D) &= E(E(\prod_D X_i | A_{\{n\}^c})) \\
&= E(E(X_j^2 \prod_D X_i | A_{\{n\}^c})) \\
&= E(X_j \prod_{D\setminus\{n\}} X_k) E(X_n \cdot X_j)
\end{aligned}$$

where $j = (n)$ in $\bar{D}$, $j \notin D$

W.l.o.g. j is the nearest site of m in $D \cup \{j\}$

$$\begin{aligned}
\hat{P}(D) &= E(X_n X_j) E(X_m X_j) E(\prod_{D\setminus\{n,m\}} X_i) \\
&= E(X_n X_m) E(\prod_{D\setminus\{n,m\}} X_i) \\
&= \hat{P}(\{n,m\}) \hat{P}(D\setminus\{n,m\})
\end{aligned}$$

by the fact that j is on the path between n and m and Lemma 5-6.

<u>Corollary 6-5</u>: If $D \in \mathcal{B}(T_2)$, $|D|$ even, $\exists\ D_i \subset D$, $i \in I$, $|I| < \infty$ such that

$$\hat{P}(D) = \prod_{I} \hat{P}(D_i)$$

and D_i is a path $\forall\ i$.

<u>Proof</u>: The result follows by applying Lemma 6-7 to D_1, D_2 and their factors. Moreover D_i can be taken to have $|D_i| = 2\ \forall\ i$.

<u>Corollary 6-6</u>: For $D \in \mathcal{B}(T_2)$, $|D|$ even

$$\hat{P}(D) = (p-q)^{\mu(D)}$$

where $\mu(D)$ is a function of the set D.

<u>Proof</u>: Follows from Corollary 6-5, taking

$$\mu(D) = \sum_{i \in I} \mu(D_i)$$

<u>Remark 6-3</u>:

$$\begin{aligned} \mu(D) &= \min\{ \sum_{i \in I} \mu(D_i);\ (D_i) \text{ finite decomposition of } D\} \\ &= \min\{ \sum_{i \in I} \mu(D_i);\ (D_i) \text{ decomposition of } D \text{ with } \\ &\qquad\qquad |D_i| = 2\ \ \forall\ i\} \end{aligned}$$

follows inductively from Lemma 6-7 and its proof.

Properties of μ: (defined for $D \in \mathcal{B}(T_2)$, $|D|$ even)

(i) $\mu(D) \geq \frac{|D|}{2}$ by Corollary 6-5.

Let $D_1, D_2 \in \mathcal{B}(T_2)$, $|D_1|, |D_2|$ even, $D_1 \cap D_2 = \emptyset$.

(ii) $\mu(D_1 \cup D_2) \leq \mu(D_1) + \mu(D_2)$ by Remark 6-3.

(iii) If $\bar{D}_1 \cap \bar{D}_2 = \emptyset$ then $\mu(D_1 \cup D_2) = \mu(D_1) + \mu(D_2)$ by Corollary 6-5 and Remark 6-3.

(iv) $\hat{P}(D_1 \cup D_2) \geq \hat{P}(D_1)\hat{P}(D_2)$ by (ii).

(v) $\exp\{2\int_{\Gamma_{D_1}\Gamma_{D_2}} dF\} \geq 1$ by (6.19) and (iv), therefore

$\int_{\Gamma_{D_1}\Gamma_{D_2}} dF \geq 0 \quad \forall\ D_1, D_2$ disjoint and thus F non negative.

Let $D_1, D_2 \in \mathcal{B}(T_2)$, $|D_1|$, $|D_2|$ even:

(vi) $\hat{P}(D_1 \Delta D_2) \geq \hat{P}(D_1)\hat{P}(D_2)$ by (iv).

(vii) $0 \leq \text{Cov}\ \gamma_{D_1}\gamma_{D_2} \leq 1$ by (vi).

(viii) $\text{Cov}\ \gamma_{D_1}\gamma_{D_2} < p-q$ since if either $(p-q)^{\mu(D_1 \Delta D_2)} < p-q$ or

$(p-q)^{\mu(D_1 \Delta D_2)} = p-q$ and then $\mu(D_1) + \mu(D_2) > 0$.

The following lemma gives a sufficient condition for tail triviality, showing that the Markov chain Ising models we have been looking at are extreme.

Lemma 6-8 (see [19]): P, representing a n.n. Gibbs state on $\Omega = \{1,-1\}^{T_N}$, is tail trivial, if for each $D \in \mathcal{B}(T_N)$, $\varepsilon > 0$ $F \in \mathcal{B}(T_N)$ such that

$$\sum_{\substack{A \in \mathcal{B}(T_N) \\ A \subset F^c}} \left| \max_{B \subset D} \operatorname{Cov}(\gamma_A \gamma_B) \hat{g}(A) \right| \leq \| g \|_\infty \quad \forall \quad g \in A_{F^c},\ g \in C(\Omega)$$

Corollary 6-7: The Markov chain Ising model with 0-external field is extreme.

Proof: For $D \in \mathcal{B}(T_2)$, choose F such that $F = \bar{F} \supset \bar{D}$, then $\bar{A} \cap \bar{B} = \emptyset \ \forall \ B \subset D$, $A \subset F^c$ and hence $\mu(A \Delta B) = \mu(A) + \mu(B)$ by property (iii) $\Rightarrow \operatorname{Cov}(\gamma_A, \gamma_B) = 0 \ \forall \ A \subset F^c$, $B \subset D$, $A,B \in \mathcal{B}(T_2)$, if $|A|$, $|B|$ are even. For $|A|$, $|B|$ odd, the result follows by a similar argument as in Example 5-1.

Chapter 7. ANOTHER APPROACH FOR THE GRAPH Z^d

The following approach will emphasize more the property of the interaction of being invariant under a certain group of graph isomorphisms and will develop the theory of Gibbs states by considering the set of invariant measures for this invariance group, where Gibbs states will then be characterized within this set.

Let $S = Z^d$ for some d. A family of graph isomorphisms is defined by the group of translations

$$a \in Z^d, \quad T_a : Z^d \to Z^d$$

$$T_a(n) = n+a = a+n, \quad n \in Z^d$$

Each T_a defines also a mapping $T_a : \Omega \to \Omega$ by

$$T_a(\underline{\bar{X}}) = \underline{\bar{Y}}, \quad \underline{\bar{X}}, \underline{\bar{Y}} \in \Omega$$

where $Y_n = X_{n+a}$.

Each T_a is continuous as a mapping on Ω and thus defines also a mapping

$$T_a : C^*(\Omega) \to C^*(\Omega)$$

by

$$T_a \mu(A) = \mu(T_a^{-1}(A)), \quad A \in \mathcal{A}, \ \mu \in C^*(\Omega)$$

<u>Remark 7-1</u>: For $f \in C(\Omega)$

$$\int f \circ T_a d\mu = \int f dT_a \mu \quad \forall \ a \in Z^d$$

which follows by the transformation lemma.

Let $M(\Omega,T_a)$ denote the set of probability measures on Ω, which are invariant under T_a, i.e.

for $\quad \mu \in M(\Omega,T_a) \quad T_a\mu = \mu$

The set

$$M(\Omega,H) = \bigcap_{a\in Z^d} M(\Omega,T_a)$$

is then the set of probability measures, which are invariant under the group H of all translations on Z^d.

Properties of $M(\Omega,T_a)$:

(i) $M(\Omega,T_a)$ is non empty: Let μ be the normalized Haar measure on Ω.

(ii) $M(\Omega,T_a)$ is convex.

(iii) $M(\Omega,T_a)$ is compact: $M(\Omega,T_a)$ is contained in the unit ball of $C^*(\Omega)$, which is compact. Assume $\mu_n \in M(\Omega,T_a)$, $\mu_n \to \mu$ weakly. Then

$$\int f \circ T_a d\mu_n = \int f d\mu_n \to \int f d\mu$$

but also

$$\int f \circ T_a d\mu_n \to \int f \circ T_a d\mu \qquad f \in C(\Omega)$$

and hence $T_a\mu = \mu$, $\mu \in M(\Omega,T_a)$ so $M(\Omega,T_a)$ is closed.

(iv) $\mu \in M(\Omega,T_a)$ is extreme iff T_a is ergodic under μ:

(a) Assume $\mu \in M(\Omega,T_a)$ and T_a not ergodic, i.e. $\exists\, A \in \mathcal{A},\ 0 < \mu(A) < 1$ and $T_aA = A$. Then

$$\mu_1(B) := \frac{\mu(AB)}{\mu(A)} \qquad \mu_2(B) := \frac{\mu(A^cB)}{\mu(A^c)}$$

define two measures in $M(\Omega,T_a)$ with

$$\mu = \alpha\mu_1 + (1-\alpha)\mu_2\ , \quad \alpha = \mu(A)$$

(b) Assume T_a is ergodic under μ and $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ then $\mu_1 << \mu$ and

$$\mu_1(E) = \int_E \frac{d\mu_1}{d\mu}\, d\mu = \int_{E\cap T_aE} \frac{d\mu_1}{d\mu}\, d\mu + \int_{E\setminus T_aE} \frac{d\mu_1}{d\mu}\, d\mu$$

Since $\mu_1(T_aE) = \mu_1(E)$:

$$\mu_1(T_aE) = \int_{E\cap T_aE} \frac{d\mu_1}{d\mu}\, d\mu + \int_{T_aE\setminus E} \frac{d\mu_1}{d\mu}\, d\mu$$

and so

$$\int_{T_aE\setminus E} \frac{d\mu_1}{d\mu}\, d\mu = \int_{E\setminus T_aE} \frac{d\mu_1}{d\mu}\, d\mu \tag{7.1}$$

If $\mu \neq \mu_1$ then for $E = \{\frac{d\mu_1}{d\mu} < 1\}$ one has that $\mu(E\setminus T_aE) = \mu(T_aE\setminus E)$ and so (7.1) can only hold if $\mu(E) = \mu_1(E) = 0$ which means μ is extreme.

From the definition of $M(\Omega,H)$ follows that this set has also the properties (i) through (iii). Furthermore:

(iv)' If $\mu \in \text{ext } M(\Omega,T_a)$ for all $a \in Z^d$ then $\mu \in \text{ext } M(\Omega,H)$.

The converse is not true.

Let $F \in \mathcal{B}(Z^d)$.

Definition 7-1: For $\mu \in M(\Omega,H)$ the entropy of μ in F is defined by

$$h_F(\mu) = -\sum_{\bar{\underline{X}}\in\Omega_F} \mu(\bar{\underline{X}}) \log \mu(\bar{\underline{X}}), \quad \mu(\bar{\underline{X}}) = \mu(\bar{\underline{X}}|\text{sites in } \Omega_{F^c} \text{ are } 1)$$

log = ln .

Let for a H-invariant potential Φ $G_0(\Phi)$ denote the set of H-invariant Gibbs states in $G(\Phi)$.

Lemma 7-1: If $\mu \in G_0(\Phi)$, where Φ is an H-invariant potential, then

$$\log Z_F^\Phi = h_F(\mu) - \sum_{\bar{\underline{X}}\in\Omega_F} \mu(\bar{\underline{X}}) U_F^\Phi(\bar{\underline{X}})$$

$$= h_F(\mu) - \int U_F^\Phi d\mu, \quad Z_F^\Phi := \sum_{\bar{\underline{X}}\in\Omega_F} \exp\{-U_F^\Phi(\bar{\underline{X}})\}$$

where U_F^Φ is the energy function defined by (3.1).

Proof:

$$h_F(\mu) = -\sum_{\bar{\underline{X}}\in\Omega_F} \mu(\bar{\underline{X}})[-U_F^\Phi(\bar{\underline{X}}) - \log Z_F^\Phi]$$

so

$$h_F(\mu) - \sum_{\underline{\bar{X}}\in\Omega_F} \mu(\underline{\bar{X}})U_F^{\Phi}(\underline{\bar{X}}) = \log Z_F^{\Phi} \sum_{\underline{\bar{X}}\in\Omega_F} \mu(\underline{\bar{X}})$$

$$= \log Z_F^{\Phi}$$

<u>Definition 7-2</u>:

$$h(\mu) := \lim_{F\to Z^d} |F|^{-1} h_F(\mu) \qquad (7.2)$$

is called entropy; $F \to Z^d$ in the sense of van Hove.

Note: $h(\mu)$ exists and is finite $\forall\ \mu \in M(\Omega,H)$ (see [14], p. 46).

For a given H-invariant potential Φ, define

$$A_{\Phi} := \lim_{F\to Z^d} |F|^{-1} U_F^{\Phi}$$

$F \in \mathcal{B}(Z^d)$, $F \to Z^d$ in the sense of van Hove.

Note: A_Φ is a continuous function.

<u>Definition 7-3</u>: The function

$$P(\Phi,H) := \sup\{h(\mu) - \int A_\Phi d\mu,\ \mu \in M(\Omega,H)\}$$

is called pressure.

Notes: - The supremum in Definition 7-3 is finite.

- If the supremum is obtained for $\mu \in \mathcal{G}_0(\Phi) \subset M(\Omega,H)$ then

$$P(\Phi,H) = \lim_{F\to Z^d} |F|^{-1}[h_F(\mu) - \int U_F^{\Phi} d\mu]$$

$$= \lim_{F\to Z^d} |F|^{-1} \log Z_F^{\Phi}$$

by Lemma 7-1, i.e. the pressure for Gibbs states is the free energy per site.

Definition 7-4: A measure $\mu \in M(\Omega,H)$ is called an equilibrium state for Φ, whenever

$$h(\mu) - \int A_{\Phi} d\mu = P(\Phi,H)$$

The set of equilibrium states for Φ is denoted by $\dot{M}_{\Phi}(\Omega,H)$.

Remark 7-2: Equilibrium states maximize the difference between entropy and energy per site, which is an expected fact since equilibrium states in physics represent states with highest entropy and lowest energy. Pressure is the maximum of the difference between entropy and energy per site. This is referred to as the variational principle.

Let for $\mu \in M(\Omega,H)$ and $F \in \mathcal{B}(Z^d)$

$$Z_F^{\Phi} := \sum_{\underline{\bar{X}}\in\Omega_F} \exp\{-U_F^{\Phi}(\underline{\bar{X}})\}$$

Lemma 7-2: For $\mu \in M(\Omega,H)$, $F \in \mathcal{B}(Z^d)$

$$h_F(\mu) - \int U_F^{\Phi} d\mu \leq \log Z_F^{\Phi} \tag{7.3}$$

Proof:

$$h_F(\mu) - \int U_F^{\Phi} d\mu = -\sum_{\bar{\underline{X}} \in \Omega_F} \mu(\bar{\underline{X}}) \{\log \mu(\bar{\underline{X}}) + U_F^{\Phi}(\bar{\underline{X}})\}$$

$$= \sum_{\bar{\underline{X}} \in \Omega_F} \mu(\bar{\underline{X}}) \log \frac{\exp\{-U_F^{\Phi}(\bar{\underline{X}})\}}{\mu(\bar{\underline{X}})}$$

and the result follows by the concavity of the logarithm.

Corollary 7-1: If $\mu \in G_0(\Phi)$ then $\mu \in M_\Phi(\Omega,H)$ and vice versa.

Proof: Equality in (7.3) holds iff $\mu(\bar{\underline{X}}) = (Z_F^{\Phi})^{-1} \exp\{-U_F^{\Phi}(\bar{\underline{X}})\}$ for $\bar{\underline{X}} \in \Omega_F$, i.e. $\mu \in G_0(\Phi)$. Taking the limit in the sense of van Hove of $|F|^{-1}[h_F(\mu) - \int U_F^{\Phi} d\mu]$ as $F \to Z^d$, yields the result. So one has $G_0(\Phi) = M_\Phi(\Omega,H)$ for any H-invariant potential ϕ and also that $M_\Phi(\Omega,H)$ is non empty, convex and compact. Extreme Gibbs states, i.e. extreme invariant equilibrium states can be described by properties of translations:

Lemma 7-3: H-invariant sets are either in A_∞ or have μ-probability 0, for μ a Gibbs state.

Proof: Assume $T_a A = A \;\forall\, a,\ A \in A_\infty$. Then $\exists\ A_1, B_1 \neq \emptyset$ s.th. $A = A_1 \cap B_1,\ A_1 \in A_F,\ B_1 \in A_{F^c}$ for some $F \in B(Z^d)$. From the invariance follows then immediately that A is a singleton.

Corollary 7-2: If μ is extreme in $M_\Phi(\Omega,H)$, all H-invariant sets are trivial: Extreme states in $M_\Phi(\Omega,H)$ are called H-ergodic.

Proof: Since non trivial H-invariant sets are contained in A_∞, the result follows by Proposition 3-4.

Lemma 7-4: If for some $a \neq 0$, T_a is ergodic under $\mu \in M_\Phi(\Omega,H)$, then μ is extreme in $M_\Phi(\Omega,H)$.

Proof: $M_\Phi(\Omega,H) \subset M(\Omega,H) \subset M(\Omega,T_a)$ $\forall$ $a \in Z^d$. If $a \in Z^d$ is such that T_a is ergodic under μ, then μ is extreme in $M(\Omega,T_a)$ and if μ is also in $M_\Phi(\Omega,H)$, μ is also extreme in $M_\Phi(\Omega,H)$. Note that for $a = 0$ no extreme state in $M(\Omega,T_a)$ is in $M_\Phi(\Omega,H)$ if Φ is non trivial.

From now on let $H = \{T_a\}$, $a \neq 0$. Then, by [17] Theorem 9.13, noting that Corollary 7-1 remains true:

Corollary 7-3: For H-invariant potential Φ, $\mu \in G_0(\Phi)$ is a pure state iff T_a is ergodic under μ, for any $a \neq 0$.

The representation of non pure Gibbs states in Corollary 3-2 is called ergodic decomposition.

Proposition 7-1: The following are equivalent for H-invariant Φ:

(i) $\mu \in G_0(\Phi)$ is extreme for $H = \{T_a\}$, $a \neq 0$.

(ii) For any $a \in Z^d\backslash\{0\}$, T_a is ergodic under μ.

(iii) If $f \in A$ and $f \circ T_a = f[\mu]$ then f is constant, for any $a \in Z^d\backslash\{0\}$

(iv) $\forall$ $A,B \in A$

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(T_a^{-i}A\cap B) \to \mu(A)\mu(B)[\mu] \quad \forall \quad a \neq 0$$

(v) μ has trivial tail.

Proof: This summarizes preceding results and different ways to describe ergodicity.

Note: (i) => (iv) is implied by Theorem 3-4.

Lemma 7-5: For $\mu \in G_0(\Phi)$, $a \neq 0$; T_a is weak mixing, T_a is mixing imply that μ has trivial tail.

Proof: Mixing implies weak mixing, which implies ergodic; ergodicity implies tail triviality by Proposition 7-1.

Lemma 7-6: If $\mu \in G_0(\Phi)$ is extreme, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T_a^i \underline{\bar{X}}) = \int f d\mu[\mu], \qquad f \in L_1(\Omega),\ a \neq 0,\ \underline{\bar{X}} \in \Omega$$

Proof: This is Birkhoff's ergodic theorem, which applies by Proposition 7-1. With Lemma 7-6 it is possible to describe correlation: Note that for $D \in B(Z^d)$, $\gamma_D \in L_1(\Omega)$.

Corollary 7-4: For $D \in B(Z^d)$, $\mu \in G_0(\Phi)$ extreme

$$\hat{\mu}(D) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \gamma_D(T_a^i(\underline{\bar{X}}))[\mu] \qquad \forall\ D. \tag{7.4}$$

Using the representation (7.4) it is possible to derive conditions for tail triviality of Gibbs states in terms of the correlation.

Lemma 7-7: If $\mu \in M_\Phi(\Omega,H)$, $D \in \mathcal{B}(Z^d)$, $a \in Z^d$ then

$$\lim_{|a|\to\infty} \hat{\mu}(D+a) = \hat{\mu}(D)$$

Proof:

$$\hat{\mu}(D+a) = \int \gamma_{D+a} d\mu = \int \gamma_D \circ T_a d\mu$$

$$= \int \gamma_D dT_a \mu = \int \gamma_D d\mu$$

which remains the same by taking the limit.

Lemma 7-8: If $\mu \in \text{ext } M_\Phi(\Omega,H)$, then

$$\lim_{|j|\to\infty} \hat{\mu}(M\Delta N+j) = \hat{\mu}(M)\hat{\mu}(N), \quad \forall\ M,N \in \mathcal{B}(Z^d) \tag{7.5}$$

Proof: $\lim_{|j|\to\infty} \hat{\mu}(M\Delta N+j)$ exists since $\mu \in \text{ext } M_\Phi(\Omega,H)$ and by (7.4)

$$\lim_{|j|\to\infty} \hat{\mu}(M\Delta N+j) = \lim_{|j|\to\infty} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \gamma_{M\Delta N+j}(T_a^i(\underline{\bar{X}}))[\mu]$$

$$= \lim_{n\to\infty} \lim_{|j|\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \gamma_M(T_a^i(\underline{\bar{X}}))\gamma_{N+j}(T_a^i(\underline{\bar{X}}))[\mu] \tag{7.6}$$

where the inner limit exists. Therefore also

$$\lim_{|j|\to\infty} \frac{1}{|j|} \sum_{k=0}^{|j|-1} \gamma_{N+k}(T_a^i(\underline{\bar{X}})) = \hat{\mu}(N+ia) \tag{7.7}$$

exists and equals $\lim_{|j|\to\infty} \gamma_{N+j}(T_a^i(\underline{\bar{X}}))$ with probability 1. Hence (7.6) becomes

$$\lim_{|j|\to\infty} \hat{\mu}(M\Delta N+j) = \lim_{n\to\infty} \frac{1}{n}\left\{\sum_{i=0}^{n-1} \gamma_M(T_a^i(\underline{\bar{X}}))\cdot\hat{\mu}(N+ia)\right\}$$

$$= \hat{\mu}(N)\cdot\hat{\mu}(M)$$

by Lemma 7-7. (7.5) means if $\mu \in \text{ext } M_\Phi(\Omega,H)$ then

$$\lim_{|j|\to\infty} \text{Cov}(\gamma_N\gamma_{M+j}) = 0, \quad N,M \in \mathcal{B}(Z^d)$$

And as a last characterization of pure states:

Proposition 7-2: For $\mu \in M_\Phi(\Omega,H)$ the following are equivalent.

(i) μ is a pure state.

(ii) For a probability measure m on Ω with $m \ll \mu$, $a \in Z^d\backslash\{0\}$

$$\frac{1}{n}\sum_{i=0}^{n-1} T_a^i m \to \mu.$$

(iii) $$\frac{1}{n}\sum_{i=0}^{n-1} \delta_{T_a^i(\underline{\bar{X}})} \to \mu[\mu]$$

Proof: (i) => (ii): Let $f \in C(\Omega)$, then

$$\int fd\ \frac{1}{n}\sum_{i=0}^{n-1} T_a^i m = \frac{1}{n}\sum_{i=0}^{n-1}\int f\circ T_a^{-i} dm$$

$$= \frac{1}{n}\sum_{i=0}^{n-1}\int f\circ T_a^{-i}\frac{dm}{d\mu}\, d\mu$$

$$\to \int fd\mu \int\frac{dm}{d\mu}\, d\mu = \int fd\mu$$

by (3.13).

(ii) = (i): For given $f \in C(\Omega)$, $g \in L_1(\Omega)$ with respect to μ. Define for $B \in A$:

$$m(B) = C \int 1_B g d\mu \quad ; \qquad C = (\int g d\mu)^{-1}$$

Then

$$\frac{1}{n} \sum_{i=0}^{n-1} \int f \circ T_a^i \cdot g d\mu \rightarrow \int f d\mu \int g d\mu [\mu]$$

which is equivalent to μ being ergodic.

(i) => (iii): Apply Birkhoff's ergodic theorem.

(iii) => (i): From the construction of measurable functions

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ T_a^i \rightarrow \int f d\mu [\mu], \qquad f \in A$$

Let $f \in C(\Omega)$, $g \in L_1(\Omega)$, then

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ T_a^i g \rightarrow g \int f \mu [\mu]$$

Using the dominated convergence theorem yields

$$\frac{1}{n} \sum_{i=0}^{n-1} \int f \circ T_a^i g d\mu \rightarrow \int g d\mu \int f d\mu [\mu]$$

If an H-invariant, extreme Gibbs state for $H = \{T_a, a \in Z^d\}$ is not extreme among all T_a-invariant Gibbs states, μ has an ergodic decomposition in terms of extreme, T_a-invariant Gibbs states and symmetry breaking is said to occur. In the other case all results hold with $H = \{T_a, a \in Z^d\}$.

BIBLIOGRAPHY

[1] Bauer, H: "Wahrscheinlichkeits Theorie und Grund-Züge der Mass Theorie," 3. Auflage, de Gruyter (1978)

[2] Bhattacharya, R.N.; Waymire, E.C.: "Stochastic Processes and Applications," Preprint.

[3] Dobrushin, R.L.: "Description of a Random Field by Means of Conditional Probabilities and the Conditions Governing its Regularity," Theor. Prob. Appl. 10, 193-213 (1968)

[4] Doob, J.L.: "Stochastic Processes," Wiley (1953)

[5] Halmos, P.R.: "Measure Theory," Springer Verlag (1976)

[6] Ising, E.: "Beitrag zur Theorie des Ferromagnetismus," Z. Phys. 31, 253-258 (1924)

[7] Israel, R.B.: "Convexity of the Theory of Lattice Gases," Princeton University Press (1979)

[8] Jacobs, K.: "Measure and Integral," Academic Press (1978)

[9] Kindermann, R.; Snell, J.L.: "Markov Random Fields and Their Applications," AMS (1980)

[10] Lanford, O.E.; Ruelle, D.: "Observables at Infinity with Short Range Correlations in Statistical Mechanics," Comm. Math. Phys. 13, 194-215 (1969)

[11] Loomis, L: "Abstract Harmonic Analysis," Van Nostrand (1953)

[12] Parthasarathy, K.R.: "Probability Measures on Metric Spaces," Academic Press (1967)

[13] Preston, C.J.: "Gibbs States on Countable Sets," Cambridge University Press (1974)

[14] Ruelle, D.: "Thermodynamic Formalism," Addison-Wesley (1978)

[15] Spitzer, F.: "Markov Random Fields and Gibbs Ensembles," MAA, 142-154 (1973)

[16] Spitzer, F.: "Markov Random Fields on an Infinite Tree," Annals of Prob. 3, 387-398 (1975)

[17] Walters, P.: "An Introduction to Ergodic Theory," Springer Verlag (1982)

[18] Waymire, E.: "Infinitely Divisible Gibbs States," Rocky Mtn. J. of Math, 14(3), 667-680, (1984)

[19] Waymire, E.: "Lattice Random Fields and Short Range Correlation," Preprint (1984)

[20] Yosida, K.: "Functional Analysis," Springer Verlag (1978)