Journal of Statistical Physics, Vol. 59, Nos. 3/4, 1990

Cluster Variation Method and Möbius Inversion Formula

T. Morita[1]

Received July 31, 1989; revision received November 22, 1989

The cluster variation method is reformulated with the aid of the Möbius inversion formula.

KEY WORDS: Cluster variation method; Möbius inversion formula; Möbius function.

1. INTRODUCTION

The cluster variation method is a method giving a series of approximations in the statistical-mechanical theory of lattice systems. It was invented by Kikuchi,[1] and a reformulation was presented by Morita.[2,3] Schlijper[4] noted that the formulation due to Morita[2,3] is related to the Möbius inversion formula. Considering this fact, An[5] showed how the Möbius function plays an important role in the formulation. In the present article, we rewrite the formulation given in ref. 3 by using the Möbius function. The terminology on the mathematics follows Rota.[6]

In Section 2, mathematical terminologies and the Möbius inversion formula are given. In Section 3, the free energy in an approximation of the cluster variation method is expressed in terms of the reduced density matrices or distribution functions for clusters of lattice sites, with the aid of the Möbius inversion formula. In expressing the free energy in an approximation in Section 3, we consider two alternative choices for the set of clusters of lattice sites. In Section 4, we see that the expression obtained in Section 3 does not depend on the two choices.

[1] Department of Engineering Science, Faculty of Engineering, Tohoku University, Sendai 980, Japan.

0022-4715/90/0500-0819$06.00/0

2. A FORMULA DUE TO MÖBIUS INVERSION FORMULA

In the present paper, we are concerned with partially ordered sets, and, in particular, with *mathematical* finite lattices. A *partially ordered set* is such a set of elements that any two elements α and β are either comparable, namely, related by one of the order relations $\alpha \leqslant \beta$ and $\beta \leqslant \alpha$, or not comparable. A *finite lattice* L is such a partially ordered set of a finite number of elements that there exist elements $\bigvee S$ and $\bigwedge S$ of L for every subset S of L. Here $\bigvee S$ and $\bigwedge S$ are called a least upper bound and a greatest lower bound, respectively, and satisfy the conditions that $\alpha \leqslant \bigvee S$ for all the elements α of S and, if β is an element of L and $\alpha \leqslant \beta$ for all the elements α of S, then $\bigvee S \leqslant \beta$, and that $\bigwedge S \leqslant \alpha$ for all the elements of α of S and, if γ is an element of L and $\gamma \leqslant \alpha$ for all the elements α of S, then $\gamma \leqslant \bigwedge S$. If S is a set of two elements α and β, $\bigvee S$ and $\bigwedge S$ are denoted by $\alpha \vee \beta$ and $\alpha \wedge \beta$, respectively. $\bigvee L$ and $\bigwedge L$ are denoted by $\mathbf{1}$ and $\mathbf{0}$, respectively, so that all the elements α of L satisfy $\mathbf{0} \leqslant \alpha \leqslant \mathbf{1}$.

The notation $\beta \geqslant \alpha$ is also used to represent $\alpha \leqslant \beta$. The notation $\alpha = \beta$ denotes that α and β are identical, so that both $\alpha \leqslant \beta$ and $\beta \leqslant \alpha$ are valid, and $\alpha < \beta$ and $\beta > \alpha$ denote that $\alpha \leqslant \beta$ and $\alpha \neq \beta$. If $\alpha < \beta$ and there is no element γ of L such that $\alpha < \gamma < \beta$, we say that β *covers* α. An element covering $\mathbf{0}$ is called an *atom*, and an element covered by $\mathbf{1}$ a *dual atom*.

We consider a finite lattice L. We assume that a quantity Q_α is associated with each element α of L. We then define $\bar{Q}_\alpha$ such that

$$Q_\beta = \sum_{\alpha \leqslant \beta} \bar{Q}_\alpha \tag{1}$$

This is possible since we have $\bar{Q}_{\mathbf{0}} = Q_{\mathbf{0}}$ if we put $\beta = \mathbf{0}$ in (1), and if $\bar{Q}_\alpha$ are determined for $\alpha < \beta$, then $\bar{Q}_\beta$ is equal to $Q_\beta - \sum_{\alpha < \beta} \bar{Q}_\alpha$. The Möbius inversion formula applied to (1) says that

$$\bar{Q}_\beta = \sum_{\alpha \leqslant \beta} Q_\alpha \mu(\alpha, \beta) \tag{2}$$

where $\mu(\alpha, \beta)$ is the Möbius function, which is defined for $\alpha \leqslant \beta$ by

$$\sum_{(\alpha \leqslant \gamma \leqslant \beta)}^{\gamma} \mu(\alpha, \gamma) = \delta_{\alpha\beta} \tag{3}$$

If we put $\beta = \alpha$ in (3), we see that $\mu(\alpha, \alpha) = 1$, and then $\mu(\alpha, \beta)$ for $\beta > \alpha$ are determined recursively by $\mu(\alpha, \beta) = -\sum_{\gamma(\alpha \leqslant \gamma < \beta)} \mu(\alpha, \gamma)$. By substituting (2) into the right-hand side of (1) and using (3), we confirm that (2) gives the solution of (1).

If we choose a fixed α and then Q_β such that $Q_\beta = 1$ when $\beta \geqslant \alpha$ and 0 otherwise, then $\bar{Q}_\gamma = \delta_{\gamma\alpha}$ by (1), and (2) reads

$$\sum_{(\alpha \leqslant \gamma \leqslant \beta)} \mu(\gamma, \beta) = \delta_{\alpha\beta} \tag{4}$$

By putting $\beta = \mathbf{1}$ in (2) and using the fact that $\mu(\mathbf{1}, \mathbf{1}) = 1$, we obtain

$$Q_{\mathbf{1}} - \bar{Q}_{\mathbf{1}} = -\sum_{\alpha < \mathbf{1}} Q_\alpha \mu(\alpha, \mathbf{1}) \tag{5}$$

Substituting (1) for $\beta = \mathbf{1}$ on the left-hand side, we have

$$\sum_{\alpha < \mathbf{1}} \bar{Q}_\alpha = -\sum_{\alpha < \mathbf{1}} Q_\alpha \mu(\alpha, \mathbf{1}) \tag{6}$$

(5) and (6) are the basic relations used in the formulation of the cluster variation method.

3. CLUSTER VARIATION METHOD

We consider a *physical* finite lattice and clusters of lattice sites on the lattice, which are denoted by α, β, γ, etc. If a cluster α is a subcluster of β, we write $\alpha \leqslant \beta$ or $\beta \geqslant \alpha$, and if α is a proper subcluster of β, $\alpha < \beta$ or $\beta > \alpha$. When $\alpha \leqslant \beta$ or $\beta \leqslant \alpha$, we say that α and β are comparable. We denote the clusters of all the lattice sites and of no lattice site by $\mathbf{1}$ and $\mathbf{0}$, respectively. We now see that $\leqslant$ in this sense is an order relation and a set of clusters is a partially ordered set. We shall use the term *finite lattice* only to represent a mathematical one in the following. For a set S of a number of clusters, we shall use notations $\bigcup S$ and $\bigcap S$ to denote the union and the intersection, respectively, of the clusters belonging to S, so that

$$\bigcup S = \bigcup_{\alpha \in S} \alpha, \qquad \bigcap S = \bigcap_{\alpha \in S} \alpha$$

An approximation in the cluster variation method is specified by a set B of basic clusters $\gamma_1, \gamma_2, \ldots$, and γ_k, which are chosen such that any two are not comparable and the union of all is the cluster of all the lattice sites, so that $\bigvee B = \bigcup B = \mathbf{1}$. We consider the set P of clusters $\mathbf{1}, \gamma_1, \gamma_2, \ldots, \gamma_k$ and all the subclusters of $\gamma_1, \gamma_2, \ldots, \gamma_k$, including the empty cluster denoted by $\mathbf{0}$. This set is a finite lattice, where for a subset S of P, $\bigwedge S = \bigcap S$, and $\bigvee S$ is equal to $\bigcup S$ or $\mathbf{1}$ according as $\bigcup S$ is a subcluster of a cluster belonging to B or not. An[5] discussed the formulation by using the partially ordered set P' which is obtained from this finite lattice P by discarding $\mathbf{0}$ and $\mathbf{1}$. We

consider also the set Q defined as follows: a cluster α is an element of Q if $\alpha = \mathbf{1}$, or $\alpha = \mathbf{0}$, or α is an element of B, or $\alpha = \bigcap S$ for some subset S of B. This set Q is also a finite lattice; $\bigvee S$ and $\bigwedge S$ for a subset S of Q are given in the Appendix. The set $Q \backslash \{\mathbf{0}, \mathbf{1}\}$ was suggested by Hijmans and de Boer[7] and adopted in ref. 3. $\gamma_1, \gamma_2, \ldots, \gamma_k$ are dual atoms both in P and Q, and a cluster of only one lattice site is an atom in P and is an atom in Q if it is in Q. In the following formulation, we choose one of the finite lattices P and Q, and clusters α, β, γ are assumed to belong to it, if not otherwise stated.

We are interested in the free energy of the system on the cluster $\mathbf{1}$. The set of variables necessary to describe all the states of a system on a cluster α is denoted by s_α. The Hamiltonian of the system on the cluster $\mathbf{1}$ is expressed as

$$H_{\mathbf{1}}(s_{\mathbf{1}}) = \sum_{\mathbf{0} < \alpha \leqslant \mathbf{1}} \bar{H}_\alpha(s_\alpha) \tag{7}$$

discarding an uninteresting constant. The summation with respect to all the variables in s_α is denoted by tr_α. When the summation is taken with respect to all the variables in s_β excluding those in s_α, it is denoted by $\mathrm{tr}_{\beta \backslash \alpha}$. The basic variational principle giving the free energy $F_{\mathbf{1}}$ of the system is

$$F_{\mathbf{1}} = \operatorname*{Min}_{\rho_{\mathbf{1}}(s_{\mathbf{1}})} \mathscr{F}_{\mathbf{1}}(\rho_{\mathbf{1}}(s_{\mathbf{1}})) \tag{8}$$

where

$$\mathscr{F}_{\mathbf{1}}(\rho_{\mathbf{1}}(s_{\mathbf{1}})) = E_{\mathbf{1}} - TS_{\mathbf{1}} \tag{9}$$

$$E_{\mathbf{1}} = \mathrm{tr}_{\mathbf{1}}\, \rho_{\mathbf{1}}(s_{\mathbf{1}})\, H_{\mathbf{1}}(s_{\mathbf{1}}), \tag{10}$$

$$S_{\mathbf{1}} = -k_{\mathrm{B}}\, \mathrm{tr}_{\mathbf{1}}\, \rho_{\mathbf{1}}(s_{\mathbf{1}}) \ln \rho_{\mathbf{1}}(s_{\mathbf{1}}) \tag{11}$$

Here T is the temperature and k_{B} is the Boltzmann constant. The minimum is taken with respect to the distribution function $\rho_{\mathbf{1}}(s_{\mathbf{1}})$ under the subsidiary condition

$$\mathrm{tr}_{\mathbf{1}}\, \rho_{\mathbf{1}}(s_{\mathbf{1}}) = 1 \tag{12}$$

We introduce distribution functions $\rho_\alpha(s_\alpha)$ for $\mathbf{0} < \alpha < \mathbf{1}$ by

$$\rho_\alpha(s_\alpha) = \mathrm{tr}_{\mathbf{1} \backslash \alpha}\, \rho_{\mathbf{1}}(s_{\mathbf{1}}) \tag{13}$$

We then obtain the consistency conditions among them:

$$\rho_\alpha(s_\alpha) = \mathrm{tr}_{\beta \backslash \alpha}\, \rho_\beta(s_\beta), \qquad \mathbf{0} < \alpha < \beta < \mathbf{1} \tag{14}$$

By (12) and (13), we have the normalization conditions

$$\mathrm{tr}_\alpha \rho_\alpha(s_\alpha) = 1, \qquad \mathbf{0} < \alpha < \mathbf{1} \tag{15}$$

By substituting (7) into (10) and then using (13), we have

$$E_{\mathbf{1}} = \sum_{\mathbf{0} < \alpha \leqslant \mathbf{1}} \mathrm{tr}_\alpha \rho_\alpha(s_\alpha) \bar{H}_\alpha(s_\alpha) \tag{16}$$

We introduce S_α for $\alpha < \mathbf{1}$ by

$$S_\alpha = -k_{\mathrm{B}} \mathrm{tr}_\alpha \rho_\alpha(s_\alpha) \ln \rho_\alpha(s_\alpha) \qquad 0 < \alpha < 1, \tag{17}$$

and $S_{\mathbf{0}} = 0$. We define $\bar{S}_\alpha$ from S_α by (1). By using (5) and $S_{\mathbf{0}} = 0$, we have

$$S_{\mathbf{1}} = -\sum_{\mathbf{0} < \alpha < \mathbf{1}} S_\alpha \mu(\alpha, \mathbf{1}) + \bar{S}_{\mathbf{1}} \tag{18}$$

We now consider a system for which the Hamiltonian is expressed as (7) with $\bar{H}_{\mathbf{1}}(s_{\mathbf{1}}) = 0$. Then (16) does not involve $\rho_{\mathbf{1}}(s_{\mathbf{1}})$. In the approximation of the cluster variation method, we put $\bar{S}_{\mathbf{1}} = 0$ and ignore the reducibility (13) of $\rho_\alpha(s_\alpha)$ from $\rho_{\mathbf{1}}(s_{\mathbf{1}})$, retaining only the consistency relations (14) and the normalization conditions (15). The variational principle in the approximation is now expressed as

$$F_{\mathbf{1}} = \operatorname*{Min}_{\{\rho_\alpha(s_\alpha)\}} \mathscr{F}_{\mathbf{1}}\{\rho_\alpha(s_\alpha)\} \tag{19}$$

$$\mathscr{F}_{\mathbf{1}}\{\rho_\alpha(s_\alpha)\} = \sum_{\mathbf{0} < \alpha < \mathbf{1}} \mathrm{tr}_\alpha \rho_\alpha(s_\alpha) \bar{H}_\alpha(s_\alpha) + T \sum_{\mathbf{0} < \alpha < \mathbf{1}} S_\alpha \mu(\alpha, \mathbf{1}) \tag{20}$$

where the S_α are given by (17). The variations are to be taken with respect to $\rho_\alpha(s_\alpha)$ for $\mathbf{0} < \alpha < \mathbf{1}$, under the subsidiary conditions (14) and (15).

The first sum on the right-hand side of (18) is compared with (15) in An's paper,[5] where a_α appears in place of $-\mu(\alpha, \mathbf{1})$. Since An did not consider the element $\mathbf{1}$, he only obtained his relation (16′), but if it is compared with (4) for $\beta = \mathbf{1}$, we see that $a_\alpha = -\mu(\alpha, \mathbf{1})$ in the lattice P.

4. EQUIVALENCE OF THE CHOICES *P* AND *Q*

In the preceding section, we could choose P or Q for the finite lattice of clusters of lattice sites. We now introduce the set R of clusters α for which $\mathbf{0} < \alpha < \mathbf{1}$ and $\mu(\alpha, \mathbf{1}) \neq 0$. We show that the set R and nonzero value of $\mu(\alpha, \mathbf{1})$ for $\alpha \in R$ do not depend on whether we choose P or Q at the outset. R is a subset of Q, and we have the same expression of the variational function $\mathscr{F}_\alpha\{\rho_\alpha(s_\alpha)\}$ in terms of $\rho_\alpha(s_\alpha)$ and $\mu(\alpha, \mathbf{1})$ with α belonging

to R, for both choices, provided that $\bar{H}_\alpha(s_\alpha)$ in (7) are chosen such that $\bar{H}_\alpha(s_\alpha)=0$ if $\alpha \notin R$.

For a pair of elements α and β of a finite lattice L such that $\alpha \leqslant \beta$, the subset of all the elements γ satisfying $\alpha \leqslant \gamma \leqslant \beta$ is called a *segment* $[\alpha, \beta]$, that is, a finite lattice in which $\mathbf{0}=\alpha$ and $\mathbf{1}=\beta$. The $\mu(\alpha, \beta)$ is defined by (3), that is,

$$\sum_{(\alpha \leqslant \gamma \leqslant \beta)} \mu(\alpha, \gamma) = \delta_{\alpha\beta} \tag{21}$$

Calculating $\mu(\alpha, \beta)$ in the lattice L is equivalent to calculating $\mu(\mathbf{0}, \mathbf{1})$ in the lattice $[\alpha, \beta]$. Hence we consider the segment $[\alpha, \mathbf{1}]$ in calculating $\mu(\alpha, \mathbf{1})$.

Methods of calculating the Möbius function $\mu(\mathbf{0}, \mathbf{1})$ have been given.[6] Here we use one of them.

A subset of elements $\alpha_0, \alpha_1, \ldots, \alpha_n$ of L is called a *maximal chain* stretched between α_0 and α_n, when α_i covers α_{i-1} for every i satisfying $1 \leqslant i \leqslant n$.

A subset S of L is said to be a *spanning subset* if $\bigvee S = \mathbf{1}$ and $\bigwedge S = \mathbf{0}$.

A *cross-cut* C of L is a subset of L such that (1) C does not contain $\mathbf{0}$ nor $\mathbf{1}$; (2) no two elements of C are comparable; (3) any maximal chain stretched between $\mathbf{0}$ and $\mathbf{1}$ contains an element of C.

Cross-cut Theorem. Let C be a cross-cut of a lattice L. For every integer $k \geqslant 2$, let q_k denote the number of spanning subsets of C containing k elements. Then the Möbius function $\mu(\mathbf{0}, \mathbf{1})$ is given by

$$\mu(\mathbf{0}, \mathbf{1}) = q_2 - q_3 + q_4 - + \cdots \tag{22}$$

We now choose P and assume that α is a cluster of P but not of Q. We consider the segment $[\alpha, \mathbf{1}]$. We assume that $\gamma_1, \gamma_2, \ldots, \gamma_m$ are the dual atoms belonging to it. We see that $\gamma_1, \gamma_2, \ldots, \gamma_m$ form a cross-cut, and apply the cross-cut theorem by using this cross-cut. For every subset S of the cross-cut, $\bigwedge S$ is equal to $\bigcap S$, which is an element of Q and hence is not equal to α. This means that there is no spanning set, and hence $\mu(\alpha, \mathbf{1}) = 0$. Using this fact in (21) with $\beta = \mathbf{1}$, we can easily see that $\mu(\alpha, \mathbf{1})$ for a cluster α belonging also to Q takes the same value as in the finite lattice Q.

APPENDIX. *Q* IS A FINITE LATTICE

We give $\bigwedge S$ and $\bigvee S$ for each subset S of Q. We first consider the case that S consists of n elements which are expressed as $\bigcap B_i = \bigcap_{\gamma \in B_i} \gamma$, respectively, for $i = 1, 2, \ldots, n$, by subsets B_i of B. If we denote the union and the intersection of the sets $B_1, B_2, \ldots, B_n$, by B' and B'', respectively,

then $B' = \bigcup_{j=1}^{n} B_j \supset B_i \supset B'' = \bigcap_{j=1}^{n} B_j$, so that $\bigcap B' \leqslant \bigcap B_i \leqslant \bigcap B''$ for $i = 1, 2, \ldots, n$, and

$$\bigwedge S = \bigcap B', \qquad \bigvee S = \bigcap B''$$

We note that $\bigwedge S = \bigcap S$. If S involves **0** in addition, $\bigwedge S = \mathbf{0}$ and $\bigvee S$ does not change by the presence of **0**. If S involves **1**, then $\bigvee S = \mathbf{1}$ and $\bigwedge S = \bigcap S$.

ACKNOWLEDGMENTS

I am grateful to Prof. T. Horiguchi for valuable advice on the manuscript.

REFERENCES

1. R. Kikuchi, *Phys. Rev.* **81**:988 (1951).
2. T. Morita, *J. Phys. Soc. Japan* **12**:753 (1957).
3. T. Morita, *J. Math. Phys.* **13**:115 (1972); *J. Stat. Phys.* **34**:319 (1984).
4. A. G. Schlijper, *Phys. Rev. B* **27**:6841 (1983).
5. G. An, *J. Stat. Phys.* **52**:727 (1988).
6. G.-C. Rota, *Z. Wahrsch.* **2**:340 (1964).
7. J. Hijmans and J. de Boer, *Physica* **21**:471 (1955).