Journal of Statistical Physics, Vol. 50, Nos. 3/4, 1988

Tiling Problems and Undecidability in the Cluster Variation Method

A. G. Schlijper[1]

Received July 11, 1987; revision received October 19, 1987

In cluster approximations for lattice systems the thermodynamic behavior of the infinite system is inferred from that of a relatively small, finite subsystem (cluster), approximations being made for the influence of the surrounding system. In this context we study, for translation-invariant classical lattice systems, the conditions under which a state for a cluster admits an extension to a global translation-invariant state. This extension problem is related to undecidable tiling problems. The implication is that restrictions of global translation-invariant states cannot be characterized purely locally in general. This means that there is an unavoidable element of uncertainty in the application of a cluster approximation.

KEY WORDS: Classical lattice system; variational principle; cluster variation method; tiling problem; undecidability.

1. INTRODUCTION

This paper is the fourth in a series[1-3] that reports on an investigation into mathematical aspects of the cluster variation method (CVM).[4-6] The CVM is a systematic approach toward the generation of approximate expressions for the free energy density of translation-invariant classical lattice systems. As the approximate expression depends on only a finite number of variables, a straightforward minimization procedure then yields information about the equilibrium thermodynamics of the system.

Despite its limitations (e.g., it is known that critical exponents cannot be reproduced correctly by the CVM[7]), the CVM seems to be the best

[1] Koninklijke/Shell Exploratie en Produktie Laboratorium, P.O. Box 60, 2280 AB Rijswijk, The Netherlands.

0022-4715/88/0200-0689$06.00/0

available approximation technique for calculating phase diagrams of realistic systems from molecular interaction.[8] Despite extensive use during the past decades, little effort has been spent on clarifying the mathematical structure of the CVM.

The CVM can conveniently be related to the variational principle for the free energy as valid for an infinite lattice system, which involves variation over all translation-invariant thermodynamic states.[9,10] It then turns out that the CVM involves two distinct steps of approximation: the first one is a truncation of a cluster expansion for the mean entropy, and the second one is allowing variation over a certain set of local states, defined for a finite part of the lattice, often referred to as the basic cluster of the approximation.[1] This second step is justified if such a local state allows extension to a translation-invariant state of the infinite system.[3,11,12] This extension problem is the subject of the present paper.

In order to make the problem definition precise, we first give a description of the lattice system that we consider and introduce some notation. We follow largely Ref. 10.

With each latice point or site a of the υ-dimensional lattice Z^υ we associate a variable ("spin") σ_a, which can have values in a finite set Ω_0, the one-site configuration space. The configuration space for a subset X of Z^υ is then $\Omega_X = (\Omega_0)^X$. By putting on Ω_0 the discrete metric, we can make Ω_0 into a metric and a topological space. The topology of Ω_X will be the product topology. In this topology Ω_X is compact, even if X is infinite.

The configuration space Ω_{Z^υ} for the thermodynamic (infinite) system on the whole lattice will be denoted simply by Ω. If $X, Y \subset Z^\upsilon$ and $X \cap Y = \varnothing$ and $\omega_X \in \Omega_X$, $\omega_Y \in \Omega_Y$, we denote by $\omega_X \times \omega_Y$ the configuration on $X \cup Y$ that coincides with ω_X on X and with ω_Y on Y. If ω is any configuration on X and $Y \subset X$, we denote the restriction of ω to Y by ω_Y.

We let $C(\Omega_X)$ denote the real-valued continuous functions on Ω_X. If $Y \subset X$, then $C(\Omega_Y)$ has a natural embedding in $C(\Omega_X)$, which will never be made explicit, i.e., any $f \in C(\Omega_Y)$ will be regarded as an element of $C(\Omega_X)$ if it is convenient to do so.

On Ω_0 we take as *a priori* measure the (unnormalized) counting measure μ_0. The product measure on Ω_X will be denoted by μ_0^X. Integration with respect to μ_0^X will be denoted by the symbol $\langle \cdot \rangle_0^X$. A state of the system is a positive linear functional ρ on $C(\Omega)$, with normalization $\rho(1) = 1$. By restriction to $C(\Omega_X)$, X finite, it defines a density function $\rho[X] \in C(\Omega_X)$ such that for all $f \in C(\Omega_X)$

$$\rho(f) = \langle f \cdot \rho[X] \rangle_0^X \tag{1}$$

For convenience we define $\rho[\varnothing]$ to be 1.

Translation over $a \in Z^\nu$ will be denoted by τ_a; the translations on Z^ν induce translations on Ω, on $C(\Omega)$, and on the states; the symbol τ_a will be used indiscriminately for all of these. I is the set of translation-invariant states. It is convex and compact in the weak* (w^*)-topology.

We also define a notion of translational invariance for states defined on a subset Λ of Z^ν: a state ρ_Λ on $C(\Omega_\Lambda)$ is locally translational invariant[2,3,12] if

$$\rho_\Lambda(f) = \rho_\Lambda(\tau_a f) \tag{2}$$

for all $f \in C(\Omega_\Lambda)$ and all $a \in Z^\nu$ such that $\tau_a f \in C(\Omega_\Lambda)$. (Recall that embeddings are not made explicit.) The set of such states is denoted by I_Λ.

The interaction Φ: $X \subset Z^\nu$, X finite $\to \Phi[X] \in C(\Omega_X)$ is required to be translation-invariant; the range $R(\Phi)$

$$R(\Phi) = \bigcup \{X \subset Z^\nu \mid |X| < \infty, 0 \in X, \Phi[X] \neq 0\} \tag{3}$$

is required to be finite: $|R(\Phi)| < \infty$. Here $|X|$ denotes the number of sites of X. An observable that represents the mean energy in any translation-invariant state is then

$$A_\Phi = \sum_{X \ni 0} \frac{\Phi[X]}{|X|} \tag{4}$$

The mean entropy $s(\rho)$ of a $\rho \in I$ is given by

$$s(\rho) = \lim_{\Lambda \to Z^\nu} \frac{S_\rho[\Lambda]}{|\Lambda|} \tag{5}$$

with the entropy for the cluster (= finite set of lattice sites) Λ given by

$$S_\rho[\Lambda] = -\rho(\log \rho[\Lambda]) = -\langle \rho[\Lambda] \log \rho[\Lambda] \rangle_0^\Lambda \tag{6}$$

The limit in Eq. (5) is to be taken in the sense of van Hove and exists for all $\rho \in I$.

Translation-invariant equilibrium states are characterized by the fact that they yield a minimum for the free energy per lattice point or free energy density $f_\Phi(\rho)$:

$$f_\Phi(\rho) = \rho(A_\Phi) - s(\rho) \tag{7}$$

Here $\beta = (kT)^{-1}$ has been absorbed into the interaction.

Practical application of this characterization thus involves treating the following minimization problem:

$$\bar{f}_\Phi = \min\{\rho(A_\Phi) - s(\rho) \colon \rho \in I\} \tag{8}$$

The energy term $\rho(A_\Phi)$ depends only on the restriction of ρ to $C(\Omega_{R(\Phi)})$ and thus is a function of a finite number of variables. This is not the case for the entropy term $s(\rho)$, however. The essence of the CVM is to replace $s(\rho)$ in the minization problem of Eq. (8) by a function of a finite number of variables. Specifically, $s(\rho)$ is replaced by a linear combination of cluster entropies

$$T_\Lambda(\rho) = \sum_{X \subset \Lambda} \alpha_{\Lambda X} S_\rho[X] \tag{9}$$

The $\alpha_{\Lambda X}$ are real coefficients that are determined according to some scheme[4,5] from the choice of the basic cluster Λ, which is chosen such that $R(\Phi) \subset \Lambda$.

As a result, the object function for minimization now depends only on the restriction of $\rho \in I$ to $C(\Omega_\Lambda)$ and is thus a function of a finite number of variables. The final step in the construction of the CVM approximation is to allow variation over $\rho \in I_\Lambda$ instead of $\rho \in I$. Thus, the CVM solves

$$\bar{f}^\Lambda_\Phi = \min\{\rho(A_\Phi) - T_\Lambda(\rho)\colon \rho \in I_\Lambda\} \tag{10}$$

instead of the minization problem of Eq. (8).

Not all elements of I_Λ need correspond to restrictions of elements of I. This fact can lead to erroneous and misleading predictions from the CVM, as has been observed in a number of cases.[13,14] To prevent this problem it would be necessary to have some suitable characterization of the set I^e_Λ of local states on $C(\Omega_\Lambda)$ that correspond to restrictions of elements of I. It is the problem of characterizing I^e_Λ that will be discussed in the following sections.

A probably more familiar setting for essentially the same characterization problem results in the limit $T \downarrow 0$. Then we are dealing with the problem of calculating ground states and ground-state energies for finite-ranged interactions. As is well known, this problem can be highly nontrivial for models in which frustration occurs, and frustration is just the phenomenon that the local energy-minimizing states are not extendable to global translation-invariant states. In fact, similar problems can arise in conjunction with most applications of variational approximation techniques. An account of such matters, including illustrations of the use and usefulness of having even partial characterizations of the correct search set in variational problems, may be found in Ref. 25. Another area of current interest where specifically the characterization of I^e_Λ is encountered is the local structure theory of cellular automata.[26]

To end this introduction, we remark that the description of the CVM given above describes a generic situation. In practical application, details may differ, the main difference being that in general it is not A_Φ that is used to calculate the mean energy, but some other, equivalent observable. Also, sometimes more than one basic cluster is used.

2. EXTENDABLE STATES

In the following Λ will be a finite set of lattice points (a cluster) of Z^ν. Theorem 1 will give necessary and sufficient conditions for a linear functional λ on $C(\Omega_\Lambda)$ to be the restriction of a translation-invariant state on $C(\Omega)$.

Definition 1. For any $f \in C(\Omega)$

$$p(f) = \max\{\rho(f) \colon \rho \in I\} \tag{11}$$

Lemma 1. (i) p is sublinear, i.e., p is subadditive

$$p(f+g) \leqslant p(f) + p(g)$$

for all f, $g \in C(\Omega)$; and positive-homogeneous,

$$p(\alpha f) = \alpha p(f)$$

for all $\alpha \geqslant 0$ in R and $f \in C(\Omega)$.

(ii) p is convex.

Proof. Trivial.

Definition 2:

$$N = \{f \in C(\Omega) \mid \rho(f) = 0 \text{ for all } \rho \in I\} \tag{12}$$

N is a closed linear subspace of $C(\Omega)$. The quotient space $C(\Omega)/N$ consists of the equivalence classes of observables that have the same expectation value for all states in I; we denote this quotient space by M. The associated quotient map is denoted by q:

$$q \colon \quad f \in C(\Omega) \to q(f) \in M \tag{13}$$

If $q(f) = q(g)$, then we will say that f and g are equivalent with respect to translations and we write $f \cong g$.

Theorem 1. Let L be a linear subspace of $C(\Omega)$ and let λ be a real-valued linear functional on L. Then there exists a state $\rho \in I$ such that

$\rho(f) = \lambda(f)$ for all $f \in L$ if and only if λ satisfies the following two conditions:

(C.1) If $f \in L$, $g \in L$, $f \cong g$, then $\lambda(f) = \lambda(g)$.

(C.2) $\lambda(f) \leqslant p(f)$ for all $f \in L$.

Proof. Let $Q = q(L)$ be the image of L under the quotient map q. Then Q is a linear subspace of M. If λ satisfies C.1, we may define a linear functional γ on Q by

$$\gamma(q(f)) = \lambda(f), \qquad f \in L \tag{14}$$

Since

$$f \cong g \quad \text{implies} \quad p(f) = p(g) \tag{15}$$

we may define a functional π on M by

$$\pi(q(f)) = p(f), \qquad f \in C(\Omega) \tag{16}$$

Then π is a sublinear (and thus convex) functional on M and if λ satisfies C.2, then

$$\gamma(x) \leqslant \pi(x) \qquad \text{for all} \quad x \in Q \tag{17}$$

By the Hahn–Banach theorem there exists a linear extension $\tilde{\gamma}$ of γ to the whole of M with

$$\tilde{\gamma}(x) \leqslant \pi(x) \qquad \text{for all} \quad x \in M \tag{18}$$

Define ρ: $C(\Omega) \to R$ by

$$\rho(f) = \tilde{\gamma}(q(f)), \qquad f \in C(\Omega) \tag{19}$$

Then for $f \in L$ we have $\rho(f) = \tilde{\gamma}(q(f)) = \gamma(q(f)) = \lambda(f)$. That ρ is a translation-invariant linear functional is obvious. It remains to show that ρ is a state. To that end, note that for all $f \in C(\Omega)$

$$\rho(f) = \tilde{\gamma}(q(f)) \leqslant \pi(q(f)) = p(f) \tag{20}$$

Consequently,

$$\rho(1) \leqslant p(1) = 1$$

and

$$\rho(1) = -\rho(-1) \geqslant -p(-1) = 1$$

so $\rho(1)=1$; and similarly, if $f \geqslant 0$, then

$$\rho(f) = -\rho(-f) \geqslant -p(-f) \geqslant 0$$

Thus, Eq. (19) defines an extension of λ to the whole of $C(\Omega)$ that is an element of I.

This proves sufficiency of C.1 and C.2. The necessity is obvious, since both conditions are satisfied by any element of I.

Theorem 2. Let λ be a state on $C(\Omega_\Lambda)$. Then the following statements are equivalent:

(i) $\lambda \in I_\Lambda$, i.e., $\lambda(f) = \lambda(\tau_a f)$ for all $f \in C(\Omega_\Lambda)$ and all $a \in Z^\nu$ such that $\tau_a f \in C(\Omega_\Lambda)$.

(ii) λ obeys condition C.1, i.e., if $f, g \in C(\Omega_\Lambda)$ and $f \cong g$, then $\lambda(f) = \lambda(g)$.

Proof. (ii) implies (i), since $\tau_a f \cong f$. For the reverse implication, let $\lambda \in I_\Lambda$. We must show that $f \in C(\Omega_\Lambda) \cap N$ implies $\lambda(f) = 0$. To that end, recall that any $f \in C(\Omega)$ has a unique decomposition (see, e.g., Ref. 15, Section 2.2)

$$f = \sum_{\substack{X \subset Z^\nu \\ |X| < \infty}} f[X] \tag{21}$$

with

$$\text{(i)} \quad f[X] \in C(\Omega_X) \tag{22}$$

$$\text{(ii)} \quad \sum_{\omega_Y \in \Omega_Y} f[X](\omega_Y \times \omega_{X \setminus Y}) = 0 \tag{23}$$

for all $Y \subset X$ and all $\omega_{X \setminus Y}$, provided $X \neq \varnothing$.

Take arbitrarily $f \in C(\Omega_\Lambda) \cap N$ and decompose f as above. Let υ be the product measure on Ω_{Λ^c} of the normalized counting measure on Ω_0, where $\Lambda^c = Z^\nu \setminus \Lambda$. Then

$$f = \int f\, d\upsilon = \sum_{X \subset \Lambda} \int f[X]\, d\upsilon + \sum_{X \not\subset \Lambda} \int f[X]\, d\upsilon = \sum_{X \subset \Lambda} f[X]$$

on account of Eqs. (22) and (23). Since the decomposition of f is unique,

$$f[X] = 0 \qquad \text{if} \quad X \not\subset \Lambda \tag{24}$$

Actually, for $f \in C(\Omega_\Lambda)$ with Λ and Ω_0 finite, the decomposition (21) can easily be made explicit: with the use of a partial trace notation

$$(\mathrm{tr}_X f)(\omega_{\Lambda \setminus X}) = \frac{1}{|\Omega_X|} \sum_{\omega_X \in \Omega_X} f(\omega_X \times \omega_{\Lambda \setminus X}) \qquad \text{if} \quad X \subset \Lambda, X \neq \varnothing$$

$$\mathrm{tr}_X f = f \qquad \text{if} \quad X = \varnothing$$

we can *define* $f[X]$ by

$$f[X] = \mathrm{tr}_{\Lambda\setminus X}\left[\sum_{Y\subset X}(-1)^{|Y|}\,\mathrm{tr}_Y f\right] \qquad \text{if} \quad X\subset\Lambda$$

$$f[X]=0 \qquad \text{if} \quad X\not\subset\Lambda$$

and verification of Eqs. (21)–(24) is straightforward.

Now define

$$\tilde{f} = f[\varnothing] + \sum_{X\ni 0}\frac{\tilde{f}[X]}{|X|} \tag{25}$$

with, for $X\neq\varnothing$,

$$\tilde{f}[X] = \sum_{a\in Z^\nu}\tau_{-a}f[\tau_a X] \tag{26}$$

[this is a finite sum due to Eq. (24)]. Then (cf. Ref. 15, Section 7.1):

$$\text{(i)} \quad \tilde{f}[X]\in C(\Omega_X) \tag{27}$$

$$\text{(ii)} \quad \tau_a\tilde{f}[X] = \tilde{f}[\tau_a X] \tag{28}$$

$$\text{(iii)} \quad \tilde{f}\cong f \tag{29}$$

Furthermore, with $\|\cdot\|$ denoting the supremum norm,

$$\inf_{g\cong f}\sum_X\|g[X]\| = |f[\varnothing]| + \sum_{X\ni 0}\frac{\|\tilde{f}[X]\|}{|X|} \tag{30}$$

(cf. Ref. 15, Theorem 7.2).

Since $f\in N$, we have $f\cong 0$, so

$$f[\varnothing]=0 \tag{31}$$

and

$$\tilde{f}[X]=0 \qquad \text{for all} \quad X \tag{32}$$

Combining Eqs. (24), (26), and (32) gives

$$\sum_{\substack{a\in Z^\nu:\\ \tau_a X\subset\Lambda}}\tau_{-a}f[\tau_a X]=0 \qquad \text{for all} \quad X \tag{33}$$

For $X\subset\Lambda$ we can apply λ to this expression:

$$0=\sum_{\substack{a\in Z^\nu\\ \tau_a X\subset\Lambda}}\lambda(\tau_{-a}f[\tau_a X]) = \sum_{\substack{a\in Z^\nu\\ \tau_a X\subset\Lambda}}\lambda(f[\tau_a X]), \qquad X\subset\Lambda \tag{34}$$

Hence

$$\lambda(f)=\lambda\left(\sum_{X\subset \Lambda} f[X]\right)=\sum_{X\subset \Lambda} \lambda(f[X])=\sum_{X}{}^{*} \sum_{\substack{a\in Z^{v}\\ \tau_a X\subset \Lambda}} \lambda(f[\tau_a X])=0$$

where $\sum^*$ denotes summation over equivalence classes with respect to translations.

Definition 3. If $H\subset \Omega_\Lambda$, then its characteristic function $\chi_H^\Lambda \in C(\Omega_\Lambda)$ is given by

$$\begin{aligned}\chi_H^\Lambda(\omega)&=1 \qquad \text{if} \quad \omega\in H\\ &=0 \qquad \text{if} \quad \omega\notin H\end{aligned}$$

where $\omega\in\Omega_\Lambda$. If H consists of a single configuration σ, then we write χ_σ^Λ for $\chi_{\{\sigma\}}^\Lambda$.

Since any $f\in C(\Omega_\Lambda)$ has the representation

$$f=\sum_{\sigma\in\Omega_\Lambda} f(\sigma)\cdot\chi_\sigma^\Lambda \tag{35}$$

any state λ on $C(\Omega_\Lambda)$ is completely specified by the values $\lambda(\chi_\sigma^\Lambda)$; moreover, to guarantee that $\lambda\in I_\Lambda$, it suffices to have

$$\lambda(\chi_\sigma^X)=\lambda(\tau_a\chi_\sigma^X) \tag{36}$$

for any $X\subset\Lambda$ and $a\in Z^v$ such that $\tau_a X\subset\Lambda$. With

$$\lambda(\chi_\sigma^X)=\sum_{\omega\in\Omega_{\Lambda\setminus X}} \lambda(\chi_{\sigma\times\omega}^\Lambda) \tag{37}$$

this means that condition C.1 of Theorem 1 is equivalent to a finite number of linear equality constraints on the parameters $p_\Lambda(\sigma)=\lambda(\chi_\sigma^\Lambda)$. Indeed, it is precisely this set of constraints that is taken into account in practical application of the CVM.

With respect to the boundedness condition C.2, the situation is different. As pointed out in the Introduction, in the CVM the variation is performed over I_Λ and no condition whatsoever is imposed to guarantee C.2. This negligence cannot be justified *a priori*, since in general not all elements of I_Λ satisfy C.2; in other words, the set of extendable states I_Λ^e is in general a proper subset of I_Λ. This may be illustrated by the following example on Z^2. Take $\Omega_0=\{-, 0, +\}$ and let Λ consist of three points of Z^2:

$$\Lambda=\{a, b, c\}=\{(0, 1), (0, 0), (1, 0)\}$$

Define a state λ on $C(\Omega_\Lambda)$ by specification of the probabilities $p_\lambda(\sigma_a, \sigma_b, \sigma_c)$ assigned by λ to configurations on Λ:

$$\begin{aligned} &p_\lambda(0, 0, +) = p_\lambda(+, -, -) = p_\lambda(-, +, 0) = 1/3 \\ &\text{all others} = 0 \end{aligned} \tag{38}$$

One easily verifies, using Eq. (36), that $\lambda \in I_\Lambda$. Yet λ cannot be extended to an element λ^+ of I_{Λ^+} with $\Lambda^+ = \Lambda \cup \{(1, 1)\}$, let alone to an element of I. Indeed, there is no configuration in Ω_{Λ^+} to which λ^+ can assign a positive probability without immediately contradicting either translational invariance or the specifications of Eq. (38).

Another, very familiar, example arises in the application of the CVM with the nearest neighbor pairs as basic clusters to the Ising antiferromagnet on the triangular lattice. Then, the approximation predicts that the probability of finding opposite spins on each nearest neighbor pair of sites tends to one as the temperature tends to zero.[11] Such a state, however, has no extension either, because of the frustration effect.

Theorem 3. I_Λ is compact and convex and has a finite number of extremal points. I_Λ^e is a compact and convex (in general proper) subset of I_Λ.

Proof. As each state λ on $C(\Omega_\Lambda)$ is defined by the values $\lambda(\chi_\sigma^\Lambda) = p_\lambda(\sigma)$ (i.e., the probability of finding the configuration σ on Λ in the state λ), I is isomorphic with a subset of R^N, with $N = |\Omega_0|^{|\Lambda|}$ (recall $|\Lambda| < \infty$); this subset is defined by a finite number of equalities (from normalization and translational invariance) and inequalities (from positivity), from which the first statement follows.

The compactness and convexity of I_Λ^e are consequences of the w^*-compactness and of the convexity of I.

At this point one might speculate whether in general I_Λ^e is the convex hull of the extendable extremal elements of I_Λ. Such speculation arises from consideration of a practice that is sometimes empoyed in applications of the CVM.[27] If an extremal element of I_Λ is obviously nonextendable (it then corresponds to what de Fontaine refers to as a "nonconstructable structure"), then the search set is changed to the convex hull of the remaining extremal points, in an *ad hoc* attempt to prevent erroneous results. Similar procedures are employed in the determination of the $T = 0$ phase diagram (cf. Section 4).

To provide a theoretical foundation for this practice, the above speculation must be proven to be true. If then also a general method can be found to decide whether or not any given extremal element of I_Λ is extendable, then it would generally be possible to perform the variation in the CVM over I_Λ^e instead of over I_Λ and thus to increase the reliability of this approximation method.

At the start of this investigation, our goal was to carry through this program. As it turns out, this goal is unattainable. In the next section we show this by establishing a connection with tiling problems. One last observation must be made to set the scene: the extremal points of I_Λ are such that the defining probabilities $p_\Lambda(\sigma)$ are all rational numbers. This follows since the constraints that define I_Λ involve only integers. In other words, if I_Λ is viewed as a subset of R^N, as in the proof of Theorem 3, then its extremal points are all in Q^N. It is thus, at least in principle, possible to calculate all the extremal elements of I_Λ with absolute precision, i.e., without any numerical inaccuracy (an algorithm to do so can be found in Ref. 16, Section 53).

3. TILING AND UNDECIDABILITY

In this section we show that the problem of characterizing the set I_Λ^e of extendable, locally translation-invariant states on $C(\Omega_\Lambda)$, Λ finite, is intimately connected with tiling problems and the so-called domino problem, which is known to be formally undecidable.

Definition 4. The support $\mathrm{supp}(\lambda)$ of a state λ on $C(\Omega_\Lambda)$ is the subset of Ω_Λ with the following property:

$$\omega \in \mathrm{supp}(\lambda) \Leftrightarrow \lambda(\chi_\omega^\Lambda) > 0 \tag{39}$$

Definition 5. A subset H of Ω_Λ is said to be tiling if there exists a configuration $\omega \in \Omega$ on Z^ν such that

$$\tau_{-a}\omega_{\tau_a\Lambda} \in H \tag{40}$$

for all $a \in Z^\nu$. One says that H is strictly tiling if H is tiling and every proper subset of H is not tiling.

Theorem 4. If $\lambda \in I_\Lambda^e$, then $\mathrm{supp}(\lambda)$ is tiling.

Proof. $\lambda \in I_\Lambda^e$, so λ has an extension $\rho \in I$. Suppose $\mathrm{supp}(\lambda)$ is not tiling. This implies that there is some cube $\tilde{C}_n$ with sides n in Z^ν with the following property: for every configuration $\sigma \in \Omega_{\tilde{C}_n}$ there is an $a \in Z^\nu$ such that $\tau_a\Lambda \subset \tilde{C}_n$ and

$$\tau_{-a}\sigma_{\tau_a\Lambda} \in K_\Lambda = \Omega_\Lambda \backslash \mathrm{supp}(\lambda)$$

Indeed, if this were not the case, then we would be able to take a sequence of cubes C_n with $C_n \to Z^\upsilon$ and a corresponding sequence of configurations $\sigma_n \in \Omega_{C_n}$ such that we would have

$$\tau_{-a}\sigma_{n_{\tau_a A}} \in \operatorname{supp}(\lambda)$$

for all $a \in Z^\upsilon$ with $\tau_a A \subset C_n$. Taking arbitrary extensions $\omega_n \in \Omega$ of these local configurations σ_n, the compactness of Ω would then provide us with the existence of a limit configuration $\omega \in \Omega$ that satisfies the tiling condition: $\tau_{-a}\omega_{\tau_a A} \in \operatorname{supp}(\lambda)$ for all $a \in Z^\upsilon$.

Now consider the restriction ρ_n of the state $\rho \in I$ to $C(\Omega_{\tilde{C}_n})$. Since $\Omega_{\tilde{C}_n}$ is finite, there is at least one local configuration $\tilde{\sigma} \in \Omega_{\tilde{C}_n}$ such that

$$\rho_n(\chi_{\tilde{\sigma}}^{\tilde{C}_n}) > 0$$

However, for some a we have $\tau_{-a}\tilde{\sigma}_{\tau_a A} = \bar{\sigma} \in K_A$. Denote $\tau_a A$ by A'; using translational invariance, we find

$$\begin{aligned}
\lambda(\chi_{\bar{\sigma}}^{A}) &= \rho(\chi_{\bar{\sigma}}^{A}) \\
&= \rho(\chi_{\tilde{\sigma}_{A'}}^{A'}) \\
&= \rho_n(\chi_{\tilde{\sigma}_{A'}}^{A'}) \\
&= \sum_{\sigma_c \in \Omega_{\tilde{C}_n \setminus A'}} \rho_n(\chi_{\tilde{\sigma}_{A'} \times \sigma_c}^{\tilde{C}_n}) \\
&\geqslant \rho_n(\chi_{\tilde{\sigma}}^{\tilde{C}_n}) > 0
\end{aligned}$$

and this contradicts that $\bar{\sigma} \in K_A = \Omega_A \setminus \operatorname{supp}(\lambda)$.

Thus, $\operatorname{supp}(\lambda)$ must be tiling.

Theorem 5. Let $H \subset \Omega_A$.

(i) If H is tiling, then there exists a $\lambda \in I_A^e$ with $\operatorname{supp}(\lambda) \subset H$.

(ii) If H is strictly tiling, then there exists a $\lambda \in I_A^e$ with $\operatorname{supp}(\lambda) = H$.

Proof. (i) Since H is tiling, there exists an $\omega \in \Omega$ such that

$$\tau_{-a}\omega_{\tau_a A} \in H$$

for all $a \in Z^\upsilon$. Let C_n be the cube in Z^υ with the origin as its first point (in lexicographic order) and sides n. Let $\bar{\omega}_n \in \Omega$ be the periodic continuation of the restriction of ω to C_n. Define $\rho_n \in I$ by

$$\rho_n(f) = \frac{1}{n^\upsilon} \sum_{a \in C_n} \tau_a f(\bar{\omega}_n) \tag{41}$$

Let d be the side of the smallest cube that contains Λ. Then, for arbitrary $\sigma \in H^c = \Omega_\Lambda \backslash H$,

$$\begin{aligned} \rho_n(\chi_\sigma^\Lambda) &= \frac{1}{n^\upsilon} \sum_{a \in C_n} \tau_a \chi_\sigma^\Lambda(\bar{\omega}_n) \\ &\leqslant \frac{1}{n^\upsilon} \{n^\upsilon - (n-d)^\upsilon\} \\ &= 1 - \left(1 - \frac{d}{n}\right)^\upsilon \end{aligned} \tag{42}$$

Since I is w^*-compact, there exists a subsequence (ρ_n) that converges w^* to some $\rho \in I$. Let λ be the restriction of ρ to $C(\Omega_\Lambda)$. Then $\lambda \in I_\Lambda^e$ and by Eq. (42)

$$\lambda(\chi_\sigma^\Lambda) = \rho(\chi_\sigma^\Lambda) = 0 \tag{43}$$

for all $\sigma \in H^c$, which means that $\mathrm{supp}(\lambda) \subset H$.

(ii) Now suppose we have in addition that H is strictly tiling. Take an arbitrary $\sigma \in H$. Then $H^- = H \backslash \{\sigma\}$ is not tiling. This implies that there is an N such that the restriction of the tiling configuration ω to any translate of the cube C_N has the property that

$$\tau_{-a} \omega_{\tau_a \Lambda} = \sigma \tag{44}$$

for some $a \in Z^\upsilon$ with $\tau_a \Lambda$ contained in this translate of C_N (cf. the argument used in the proof of Theorem 4). Now cover the lattice with disjunct translates of C_N. Loosely speaking, Eq. (44) then states that in the configuration ω we will find the local configuration σ at least once within every such translate. Consequently,

$$\begin{aligned} \rho_n(\chi_\sigma^\Lambda) &= \frac{1}{n^\upsilon} \sum_{a \in C_n} \tau_a \chi_\sigma^\Lambda(\bar{\omega}_n) \\ &\geqslant \frac{1}{n^\upsilon} \left[\frac{n}{N}\right]^\upsilon \\ &> \left(\frac{1}{2N}\right)^\upsilon \end{aligned} \tag{45}$$

for all $n \geqslant N$ and thus

$$\lambda(\chi_\sigma^\Lambda) = \rho(\chi_\sigma^\Lambda) \geqslant \left(\frac{1}{2N}\right)^\upsilon > 0 \tag{46}$$

Since this holds for arbitrary $\sigma \in H$, we have now that $\mathrm{supp}(\lambda) = H$.

Next, we shall relate the concept of tiling to the functional p (cf. Definition 1).

Lemma 1. For any $f \in C(\Omega)$

$$p(f) = \inf_n \sup_{\omega \in \Omega} \frac{1}{n^\upsilon} \sum_{a \in C_n} \tau_a f(\omega) \tag{47}$$

where C_n is the cube in Z^υ with sides n and the origin as its first point in lexicographic order.

Proof. Define an auxiliary function F_f, defined on finite subsets X of Z^υ, for each $f \in C(\Omega)$:

$$F_f[X] = \sup_{\omega \in \Omega} \sum_{a \in X} \tau_a f(\omega) \tag{48}$$

One verifies without difficulty that F_f has the following properties:

(i) F_f is subadditive: if $X_1 \cap X_2 = \varnothing$, then

$$F_f[X_1 \cup X_2] \leqslant F_f[X_1] + F_f[X_2] \tag{49}$$

(ii) F_f is translation-invariant:

$$F_f[\tau_a X] = F_f[X] \tag{50}$$

(iii) The following condition holds:

$$\inf_{\omega \in \Omega} f(\omega) \leqslant F_f[X]/|X| \leqslant \sup_{\omega \in \Omega} f(\omega) \tag{51}$$

By a standard argument (see, for instance, Ref. 15, Theorem 4.10), the above properties imply that the limit

$$p_2(f) = \lim_{n \to \infty} \frac{F_f[C_n]}{|C_n|} \tag{52}$$

exists and that

$$p_2(f) = \inf_n \frac{F_f[C_n]}{|C_n|} \tag{53}$$

Now write

$$R_f^n = \frac{1}{n^\upsilon} \sum_{a \in C_n} \tau_a f \tag{54}$$

Then for any $\rho \in I$ and any n

$$\rho(f) = \rho(R_f^n) \leqslant \sup_{\omega \in \Omega} R_f^n(\omega)$$

Thus,

$$p(f) = \max_{\rho \in I} \rho(f) \leqslant \inf_n \sup_{\omega \in \Omega} R_f^n(\omega) = p_2(f) \tag{55}$$

To prove the reverse inequality, we note that the functional p_2 is sublinear and convex, as can be shown by a straightforward calculation. Moreover, if $f \cong g$, then $p_2(f) = p_2(g)$, by the following argument: $f \cong g$ implies $g - f = h \cong 0$. Then, by Theorem 7.1 of Ref. 15,

$$\lim_{n \to \infty} \frac{1}{n^{\upsilon}} \left\| \sum_{a \in C_n} \tau_a h \right\| = 0$$

Thus

$$\lim_{n \to \infty} \sup_{\omega \in \Omega} R_h^n(\omega) = p_2(h) \leqslant 0$$

By the sublinearity of p_2 we then have

$$p_2(g) = p_2(f + h) \leqslant p_2(f) + p_2(h) \leqslant p_2(f)$$

Reversing the roles of f and g yields $p_2(f) \leqslant p_2(g)$, and thus $p_2(f) = p_2(g)$.

We now define a sublinear functional π_2 on $M = C(\Omega)/N$ by

$$\pi_2(q(f)) = p_2(f), \qquad f \in C(\Omega) \tag{56}$$

Take $f \in C(\Omega)$ fixed. We shall show the existence of a $\rho_f \in I$ with $\rho_f(f) = p_2(f)$, which implies $p(f) \geqslant p_2(f)$.

Define a linear functional β on the subspace of M that is spanned by the element $q(f) \in M$ by

$$\beta(\alpha q(f)) = \alpha \beta(q(f)) = \alpha p_2(f), \qquad \alpha \in R \tag{57}$$

For $\alpha \geqslant 0$, we have, since p_2 is positive-homogeneous,

$$\beta(\alpha q(f)) = p_2(\alpha f) = \pi_2(\alpha q(f)) \tag{58}$$

For $\alpha < 0$, since p_2 is subadditive and thus

$$0 = p_2(0) \leqslant p_2(f) + p_2(-f) \tag{59}$$

we have

$$\beta(\alpha q(f)) = \alpha p_2(f) = (-\alpha)\cdot(-p_2(f))$$
$$\leqslant (-\alpha)\cdot p_2(-f) = p_2(\alpha f) = \pi_2(\alpha q(f)) \qquad (60)$$

Equations (58) and (60) show that β is majorised by π_2 and the Hahn–Banach theorem ensures the existence of a linear extension $\tilde{\beta}$ of β to the whole of M with

$$\tilde{\beta}(q(g)) \leqslant \pi_2(q(g)) \qquad \text{for all} \quad g \in C(\Omega) \qquad (61)$$

Now define ρ_f by

$$\rho_f(g) = \tilde{\beta}(q(g)), \qquad g \in C(\Omega) \qquad (62)$$

As in the proof of Theorem 1, it can be shown that $\rho_f \in I$. Since $\rho_f(f) = p_2(f)$ by construction, we have now shown that

$$p(f) = \max\{\rho(f) \colon \rho \in I\} \geqslant p_2(f) \qquad (63)$$

Equations (55) and (63) yield the lemma.

Lemma 2. For $f \in C(\Omega_\Lambda)$

$$p(f) = \max\{\lambda(f) \colon \lambda \in I_\Lambda^e\} \qquad (64)$$

Proof. Trivial.

Theorem 6. $H \subset \Omega_\Lambda$ is tiling if and only if $p(\chi_H^\Lambda) = 1$.

Proof. Suppose H is tiling. Then there is an $\bar{\omega} \in \Omega$ such that, for all $a \in Z^\nu$, $\tau_{-a}\bar{\omega}_{\tau_a \Lambda} \in H$, or $[\tau_{-a}\bar{\omega}]_\Lambda \in H$; thus,

$$\tau_a \chi_H^\Lambda(\bar{\omega}) = \chi_H^\Lambda(\tau_{-a}\bar{\omega}) = 1$$

Hence, for all n,

$$\frac{1}{n^\nu} \sum_{a \in C_n} \tau_a \chi_H^\Lambda(\bar{\omega}) = 1$$

Thus,

$$\sup_{\omega \in \Omega} \frac{1}{n^\nu} \sum_{a \in C_n} \tau_a \chi_H^\Lambda(\omega) \geqslant 1$$

and thus, by Lemma 1, $p(\chi_H^\Lambda) \geqslant 1$.

Obviously $p(\chi_H^\Lambda) \leqslant 1$, so it follows that $p(\chi_H^\Lambda) = 1$, which proves the first part of the theorem.

As to the reverse implication, suppose $p(\chi_H^A)=1$. By Lemma 1 this implies

$$\sup_{\omega\in\Omega}\frac{1}{n^v}\sum_{a\in C_n}\tau_a\chi_H^A(\omega)=1 \tag{65}$$

for all n. This means that for all n there exists a configuration $\bar{\omega}_n$ on $\Lambda_n=\bigcup\{\tau_a\Lambda: a\in C_n\}$ such that for each $a\in C_n$

$$\tau_a\chi_H^A(\bar{\omega}_n)=1 \qquad \text{or} \qquad \tau_{-a}\bar{\omega}_{n_{\tau_a\Lambda}}\in H$$

The compactness of Ω then yields the existence of a tiling configuration $\bar{\omega}\in\Omega$.

Remark. Theorem 6 in combination with Lemma 2 immediately yields alternative proofs of Theorems 4 and 5. In fact, we now have the following result:

Corollary 1. Let $E(I_\Lambda^e)$ denote the set of extremal points of I_Λ^e.

(i) $H\subset\Omega_\Lambda$ is tiling if and only if there exists a $\lambda\in E(I_\Lambda^e)$ with $\mathrm{supp}(\lambda)\subset H$.

(ii) If $H\subset\Omega_\Lambda$ is strictly tiling, then there exists a $\lambda\in E(I_\Lambda^e)$ with $\mathrm{supp}(\lambda)=H$.

Proof. Suppose H is tiling. By Theorem 6 this implies $p(\chi_H^A)=1$. By Lemma 2 this implies $\max\{\lambda(\chi_H^A): \lambda\in I_\Lambda^e\}=1$. Since this involves the maximum of a linear functional of λ over the compact and convex set I_Λ^e (Theorem 3), the maximum is assumed in an extremal point $\lambda_0\in E(I_\Lambda^e)$. Then $\lambda_0(\chi_H^A)=1$ implies $\mathrm{supp}(\lambda_0)\subset H$. Theorem 4 implies that $\mathrm{supp}(\lambda_0)$ is tiling; so, first, if H is strictly tiling, then $\mathrm{supp}(\lambda_0)=H$, and second, this implies again that H is tiling.

Corollary 2. $H\subset\Omega_\Lambda$ is tiling if and only if

$$\max\{\lambda(\chi_H^A): \lambda\in E(I_\Lambda^e)\}=1$$

Proof. A trivial consequence of Theorem 6.

As pointed out at the end of Section 2, the simple structure of the set I_Λ of locally translation-invariant states on $C(\Omega_\Lambda)$ (over which set the variation in the CVM is performed) makes it possible to identify with absolute precision all the extremal points of I_Λ. In this sense a complete and exact characterization of I_Λ is possible. Corollary 2 implies that the set I_Λ^e of extendable, locally translation-invariant states (over which set one would like to perform the variation in the CVM) does not permit such a characterization. To show this, we consider the problem of determining in general whether any given subset H of $\Omega_\Lambda = (\Omega_0)^\Lambda$, for any given Ω_0 and any given $\Lambda \subset Z^\nu$, is tiling or not. Let us call this problem the tiling problem. The tiling problem is a so-called undecidable problem[17–20] if $\nu \neq 1$. Roughly speaking, this means that there does not exist a general criterion that, given Ω_0, Λ, and H, enables one to verify, in a finite amount of time, whether H is tiling or not. The word "general" here means that the criterion must be applicable to, and yield a definite answer for, any triple $\{\Omega_0, \Lambda, H\}$. Another formulation of this undecidability problem is the following: it is not possible to write a computer program, no matter how complicated, that is guaranteed to come up with the answer to the question: "Is H tiling?" in a finite amount of time, for arbitrary $\{\Omega_0, \Lambda, H\}$. This does not preclude the possibility that such a program exists for subclasses of the general problem, however; the number of "decidable subclasses" will then be infinite, or else they will not cover the entire general problem.

The tiling problem is undecidable because the so-called domino problem, which is known to be undecidable,[21] may be identified with a subclass of the general tiling problem.

The domino problem is the following.

Suppose that we are given a finite set of unit squares, the dominoes, whose edges are marked with symbols (say, colors), each domino in a different manner. Assume that we have an unlimited number of copies of each type of domino. We seek to assemble the dominoes on the infinite plane, ruled into unit squares, according to the following rules:

1. No domino may be rotated or reflected.
2. A domino must be placed exactly over a ruled square.
3. The symbols (colors) on adjacent domino edges must match.
4. Every square must be covered with a domino.

The domino set is called solvable if and only if the dominoes can be so assembled. A solvabe domino set is also said to "admit a tiling of the plane."

The domino problem deals with the class of all domino sets. It consists of deciding, for each domino set, whether or not it is solvable. The domino problem is said to be decidable or undecidable according to whether there exists or does not exist an algorithm which, given the specifications of an arbitrary domino set, will decide whether or not the set is solvable. Implicit in the notion of an algorithm is that the answer will be found in a finite amount of time. Such an algorithm is also called a decision procedure.

Some background on the domino problem may be found in Ref. 22.

Theorem 7 (Berger, 1966).[21] The domino problem is undecidable.

The obvious connection between the domino problem and the tiling problem is made explicit in the following theorem.

Theorem 8. The tiling problem is undecidale if $\upsilon \neq 1$.

Proof. The domino problem deals with the class of all domino sets D. The tiling problem in υ dimensions deals with the class of all triples $\{\Omega_0, \Lambda, H\}$, where Λ is a finite subset of Z^υ and $H \subset \Omega_\Lambda = (\Omega_0)^\Lambda$, with Ω_0 a finite set.

To each domino set D we assign such a triple $\{\Omega_{0D}, \Lambda_D, H_D\}$ in the following way.

Since D is a finite set, we may number its elements, the dominoes, from 1 to N. The set $\{1,..., N\}$ will be the one-site configuration space Ω_{0D}. For Λ_D we take a square with sides 2 in Z^υ. Now construct all 2×2 squares of four dominoes that are allowed under the constraint of matching edges. This set of squares then defines the subset H_D of $(\Omega_{0D})^{\Lambda_D}$. Note that H_D may be empty.

It is now obvious that the domino set D is solvable if and only if H_D is tiling.

Suppose the tiling problem for $\upsilon = 2$ is decidable. Then there exists a decision procedure (an algorithm) to determine whether H_D is tiling or not. This algorithm thus also constitutes a decision procedure for the domino problem. Since this contradicts Berger's theorem, the tiling problem for $\upsilon = 2$ must be undecidable.

The undecidability of the tiling problem for $\upsilon > 2$ follows by induction, by considering the υ-dimensional lattice as a stack of $(\upsilon - 1)$-dimensional lattices.

Now suppose we had succeeded in carrying out the program outlined at the end of Section 2, i.e., suppose we had been able to prove that all extremal elements of I^e_Λ are extremal in I_Λ and that we had found some finite algorithm to decide whether or not any given extremal element of I_Λ is extendable. We would then also be able to construct a decision preocedure for the tiling problem: as noted before, finding all the extremal elements of I_Λ is a finite task; since there is only a finite number of them, verifying the extendability of each of them in turn would then also be a finite task; finally, since each of these extremal elements is defined by rational configuration probabilities, application of Corollary 2 would decide the tiling problem, because an explicit and finite calculation would suffice to establish whether or not the maximum figuring in that Corollary equals one. Since this contradicts Theorem 8, however, at least one of the following statements must be true:

1. I_A^e can have extremal elements that are not extremal in I_A.

2. The problem of finding whether or not any given extremal element of I_A is extendable is undecidable (i.e., there is no general finite algorithm guaranteed to accomplish this task).

4. CVM AT $T=0$ AND FRUSTRATION

In the previous section we saw that a general discussion of the cluster variation method leads in a natural way to the problem of characterizing the set of extendable, locally translation-invariant states and hence to a consideration of tiling problems. These tiling problems also occur in the construction of ground-state configurations from local energy-minimizing configurations, in which situation the impossibility of tiling is usually called "frustration," and also in this respect there is an obvious connection with the CVM.

We recall the minimization problem of the CVM, Eq. (10), and we now make the dependence on $\beta=(kT)^{-1}$ explicit:

$$\beta \bar{f}_{\Phi}^{A}=\min\{\beta\rho(A_{\Phi})-T_{A}(\rho)\colon \rho\in I_{A}\} \tag{66}$$

with

$$T_{A}(\rho)=\sum_{X\subset A}\alpha_{AX}S_{\rho}[X] \tag{67}$$

Definition 6. The local ground-state energy e_0^A is

$$e_0^A=\min\{\rho(A_{\Phi})\colon \rho\in I_A\} \tag{68}$$

States $\rho_0\in I_A$ with $\rho_0(A_{\Phi})=e_0^A$ will be called local ground states. The set of local ground states will be denoted by I_A^0.

Theorem 9. Let $\bar{\rho}_T\in I_A$ be a CVM approximation to the restriction of an equilibrium state at inverse temperature $\beta=(kT)^{-1}$, i.e., let

$$\beta \bar{f}_{\Phi}^{A}=\beta\bar{\rho}_{T}(A_{\Phi})-T_{A}(\bar{\rho}_{T}) \tag{69}$$

Let $\bar{\rho}_0$ be a limit of such states as $T\downarrow 0$. Let

$$T_0=\max\{T_A(\rho)\colon \rho\in I_A^0\} \tag{70}$$

Then

$$\bar{\rho}_0(A_\Phi) = e_0^A$$

and

$$T_A(\bar{\rho}_0) = T_0$$

Proof. $T_A(\rho)$ is a continuous function of $\rho \in I_A$ and it is bounded:

$$|T_A(\rho)| \leqslant B = \sum_{X \subset A} |\alpha_{AX}| \cdot |X| \cdot \log |\Omega_0| \tag{71}$$

Let ρ_0 be any local ground state. Then

$$\begin{aligned} \beta e_0^A &\leqslant \beta \bar{\rho}_T(A_\Phi) = \beta \bar{f}_\Phi^A + T_A(\bar{\rho}_T) \\ &\leqslant \beta \rho_0(A_\Phi) - T_A(\rho_0) + T_A(\bar{\rho}_T) \\ &\leqslant \beta e_0^A + 2B \end{aligned}$$

Hence

$$\bar{\rho}_0(A_\Phi) = e_0^A \tag{72}$$

Similarly, let $\tilde{\rho} \in I_A^0$ be such that $T_A(\tilde{\rho}) = T_0$ (its existence is guaranteed, since I_A^0 is compact). Then

$$\begin{aligned} T_A(\bar{\rho}_T) &= \beta \bar{\rho}_T(A_\Phi) - \beta \bar{f}_\Phi^A \\ &\geqslant \beta \bar{\rho}_T(A_\Phi) - \beta \tilde{\rho}(A_\Phi) + T_A(\tilde{\rho}) \\ &= \beta \bar{\rho}_T(A_\Phi) - \beta e_0^A + T_0 \\ &\geqslant T_0 \end{aligned}$$

Hence, $T_A(\bar{\rho}_0) \geqslant T_0$, and thus, because of Eq. (71),

$$T_A(\bar{\rho}_0) = T_0$$

Remark. Theorem 9 is just the CVM version of the exact variational principle valid at $T=0$ as stated by Aizenman and Lieb.[23]

In applications of the CVM it is assumed that the phase diagram at $T>0$ is a continuous deformation of the phase diagram at $T=0$. The CVM free energy minimization procedure is then used to establish the regions of stability of phases that are known *a priori*: the phases that are present at $T=0$. All $T=0$ limits of CVM equilibrium states $\bar{\rho}_T$ are found in the (compact) set I_A^0. A necessary condition for a physically acceptable

CVM approximation would then be that $I_\Lambda^0 \cap I_\Lambda^e \neq \varnothing$; otherwise all considered phases would, at low enough temperatures, correspond to nonextendable, i.e., physically unacceptable, local states.

The standard example where this happens is in the application of the CVM pair approximation to the Ising antiferromagnet on the triangular lattice: restricting attention to nearest neighbor pairs of lattice points means that the frustration phenomenon is not taken into account at all.

Investigation of the condition mentioned above is usually a part of the procedure that is followed in practical applications of the CVM, in the following form: the extremal elements of I_Λ^0 are identified (note that these are also extremal in I_Λ) and one tries to establish extendability or non-extendability of each of them "by inspection".[27] For simple situations a verdict can usually be reached, and $I_\Lambda^0 \cap I_\Lambda^e$ is taken to be the convex hull of the extendable extremal elements of I_Λ^0. This set of local states is then used in a determination of the $T=0$ phase diagram. This *ad hoc* procedure is obviously rather unsatisfactory: the "inspection" is difficult for more complicated lattices, such as the face-centered cubic lattice, and large basic clusters Λ and it is not guaranteed to reach any verdict at all. Not surprisingly, also here attempts to devise automated procedures come up against the undecidability properties of tiling problems.

Definition 7. The set of local ground-state configurations $G_\Lambda \subset \Omega_\Lambda$ is

$$G_\Lambda = \bigcup \{\operatorname{supp}(\rho) \colon \rho \in I_\Lambda^0\}$$

Theorem 10. A necessary and sufficient condition for I_Λ^0 and I_Λ^e to have a nonempty intersection is that G_Λ is tiling.

Proof. Let $\rho \in I_\Lambda^0 \cap I_\Lambda^e$. Since $\rho \in I_\Lambda^0$, $\operatorname{supp}(\rho) \subset G_\Lambda$. Since $\rho \in I_\Lambda^e$, $\operatorname{supp}(\rho)$ is tiling (Theorem 4). Hence G_Λ is tiling. This proves necessity. Now suppose G_Λ is tiling. By Theorem 5 there exists a $\lambda \in I_\Lambda^e \subset I_\Lambda$ with $\operatorname{supp}(\lambda) \subset G_\Lambda$. Then $\lambda \in I_\Lambda^0$, which yields sufficiency, by the following reasoning.

If $|G_\Lambda| = 1$, i.e., G_Λ consists of one configuration σ_0 on Λ, the situation is trivial: then I_Λ^0 consists of only one element λ_0, defined by $\lambda_0(\chi_{\sigma_0}^\Lambda) = 1$ and $\operatorname{supp}(\lambda) \subset G_\Lambda$ implies $\lambda = \lambda_0$.

If $|G_\Lambda| > 1$, we argue as follows. For any $\sigma \in G_\Lambda$ there is a $\rho_\sigma \in I_\Lambda^0$ with $\rho_\sigma(\chi_\sigma^\Lambda) > 0$. Define

$$\bar{\rho} = \frac{1}{|G_\Lambda|} \sum_{\sigma \in G_\Lambda} \rho_\sigma$$

Then $\bar{\rho} \in I_A^0$ and $\text{supp}(\bar{\rho}) = G_A$; with

$$m = \min\{\bar{\rho}(\chi_\sigma^A): \sigma \in G_A\}$$
$$M = \max\{\bar{\rho}(\chi_\sigma^A): \sigma \in G_A\}$$

we have for all $\sigma \in G_A$

$$0 < m \leqslant \bar{\rho}(\chi_\sigma^A) \leqslant M < 1 \tag{73}$$

Now suppose $\lambda(A_\Phi) > e_0^A$. Choose α such that $0 < \alpha < m$ and write

$$\lambda_1 = (\bar{\rho} - \alpha\lambda)/(1 - \alpha) \tag{74}$$

Then $\lambda_1 \in I_A$, $\text{supp}(\lambda_1) = G_A$, but $\lambda_1(A_\Phi) < e_0^A$, which contradicts the definition of e_0^A. Thus, it must be that $\lambda(A_\Phi) = e_0^A$, or $\lambda \in I_A^0$.

If G_A is not tiling, then we encounter frustration: local energy-minimizing configurations cannot be extended to global ones. The only relevant freedom at this point in setting up the CVM approximation is the choice of the basic cluster A, and it should thus be chosen large enough to prevent this frustration phenomenon. While in many situations that arise in practice it will be obvious how large a frustration-preventing basic cluster should be, a general finite criterion for such a choice does not exist. For this would imply the existence of an algorithm to decide whether G_A is tiling or not, and such an algorithm cannot exist. The fact that sets of local ground-state configurations are not arbitrary subsets of Ω_A does not detract from the validity of the argument: this is a consequence of the observations that precisely those subsets of Ω_A that support some element of I_A are the sets of local ground-state configurations for some interaction Φ with $A_\Phi \in C(\Omega_A)$, and that this includes all the strictly tiling subsets of Ω_A. A detailed discussion of this point may be found in Ref. 12.

5. DISCUSSION

In the CVM an approximate expression for the free energy density of a classical lattice system is minimized to obtain information about the equilibrium situation. This minimum is sought by variation over the set of local states I_A. The set I_A can easily be characterized by a finite set of local conditions. However, only those elements of I_A that are the restriction of an element of I, i.e., of a global translation-invariant state, can have a physical interpretation. Neglecting this aspect basically means that essential frustration effects are not taken into account by the approximation, and this can result in completely erroneous predictions. To prevent this, one would want to solve the minimization problem in the set I_A^e of extendable locally translation-invariant states. For that, a suitable characterization of I_Λ^e would be needed.

Theorem 1 identifies necessary and sufficient conditions for a linear functional λ defined on a subspace of $C(\Omega)$ to be the restriction of an element of I. Both conditions C.1 and C.2 as stated in this theorem are not directly amenable to verification in practice. Theorem 2 shows that the infinite set of equality constraints that make up C.1 can be replaced by a finite number of constraints that can be used in actual calculations. The reduction of C.2 to a form that can be used in practical application of the CVM is not as readily accomplished. Guided by existing *ad hoc* procedures, we formulated an attempt at such a reduction. This led to an investigation of the relationship between the problem of characterizing I_Λ^e and so-called tiling problems. From the formal undecidability of the tiling problem, we concluded that our attempt was doomed to fail, and hence the *ad hoc* practice cannot be given a rigorous justification in general. Corollary 2 actually implies the nonexistence of any characterization of I_Λ^e that would allow an exact (i.e., without numerical inaccuracy) solution of the maximization problem stated there.

This still leaves the practical problem of characterizing I_Λ^e in some way that can be employed in the actual CVM variational calculation, since we have only succeeded in showing that the existing *ad hoc* procedure will not do in the general case. Perhaps the most interesting question in this context is if (or when) I_Λ^e has a finite number of extreme points. If such is the case, then only a finite number of the inequalities of condition C.2 are effective constraints and the problem is reduced to identifying these effective ones. (Note that this possibility is not *a priori* excluded by the relation to the tiling problem: if nonrational configuration probabilities are involved, it may no longer be possible to determine in a finite amount of time whether the maximum in Corollary 2 equals 1 or not.) Otherwise, practical advantages might still be obtained if it is possible to characterize some suitable sequence of approximations A_n to I_Λ^e that converges to I_Λ^e in some sense. Replacing I_Λ^e by I_Λ would then correspond to something like a zero-order approximation. Improvement of a CVM approximation could then take two directions: a better approximation of the mean entropy than is given by T_Λ or a better approximation to I_Λ^e than is provided by I_Λ. Suggestions for improvement in the first direction are contained in Ref. 12. Improvement in the second direction only results if I_Λ^e is actually a proper subset of I_Λ and if $\rho(A_\Phi) - T_\Lambda(\rho)$ takes its minimum value in $I_\Lambda \backslash I_\Lambda^e$. In current practice the only way to improve on any given approximation is to choose a larger basic cluster Λ; this, however, increases the dimensionality of the minimization problem and thus the required computational effort, and is therefore not always feasible.

With respect to the construction of ground-state configurations and the choice of a "frustration-preventing" basic cluster, it may be interesting to remark that this problem not only occurs in the CVM but, e.g., also in the work of Holsztynski and Slawny on verification of the Peierls condition[24]: they remark on the necessity of choosing a potential for a given Hamiltonian in such a way that local ground-state configurations can be patched together; their selection problem is identical to the one in the CVM, and the nonexistence of a systematic selection procedure thus also pertains to their verification criterion.

Finally, we stress that all the undecidability results depend crucially on the generality of the problem that is considered. Decision procedures for certain subclasses of the general problem will often exist. Thus, for certain clusters Λ or certain configuration spaces Ω_0 a complete, finite, and exact characterization of I^e_Λ may be obtainable. The identification of such cases can be of practical relevance. Some results in this direction may be found in Refs. 3 and 12.

ACKNOWLEDGMENTS

I am indebted to the Management of Shell Research B.V. for permission to publish this paper. Stimulating and enlightening discussions with Prof. M. Winnink, R. Kuik, and T. Dorlas are gratefully acknowledged. The proof of Theorem 2 was communicated by R. Kuik.

REFERENCES

1. A. G. Schlijper, *Phys. Rev. B* **27**:6841 (1983).
2. A. G. Schlijper, *J. Stat. Phys.* **35**:285 (1984).
3. A. G. Schlijper, *J. Stat. Phys.* **40**:1 (1985).
4. R. Kikuchi, *Phys. Rev.* **81**:988 (1951).
5. T. Morita, *J. Phys. Soc. Jpn.* **12**:753, 1060 (1957).
6. D. M. Burley, in *Phase Transitions and Critical Phenomena*, Vol. 2, C. Domb and M. S. Green, eds. (Academic Press, New York, 1972), Chapter 9.
7. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, London, 1982).
8. J. M. Sanchez, D. de Fontaine, and W. Teitler, *Phys. Rev. B* **26**:1465 (1982).
9. D. Ruelle, *Thermodynamic Formalism* (Addison–Wesley, Reading, Massachusetts, 1978).
10. R. B. Israel, *Convexity in the Theory of Lattice Gases* (Princeton University Press, Princeton, New Jersey, 1979).
11. A. G. Schlijper, *Phys. Rev. B* **31**:609 (1985).
12. A. G. Schlijper, Variational approximation in classical lattice systems, Ph. D. Thesis, University of Groningen, The Netherlands (1985).
13. C. E. Campbell and M. Schick, *Phys. Rev. A* **5**:1919 (1972).
14. M. Kaburagi, T. Tonegawa, and J. Kanamori, *J. Phys. Soc. Ppn.* **51**:3857 (1982).
15. N. M. Hugenholtz, *Proc. Symp. Pure Math.* **38** (Part 2):407 (1982).

16. A. W. Roberts and D. E. Varberg, *Convex Functions* (Academic Press, New York, 1973).
17. A. Church, *J. Symbolic Logic* **1**:40, 101 (1936).
18. A. M. Turing, *Proc. Lond. Math. Soc. Ser. 2* **42**:230 (1936); **43**:544 (1937).
19. J. R. Büchi, *Math. Annalen* **148**:201 (1962).
20. M. L. Minsky, *Computation: Finite and Infinite Machines* (Prentice-Hall, Englewood Cliffs, New Jersey, 1967).
21. R. Berger, *Mem. Am. Math. Soc.* **66** (1966).
22. R. M. Robinson, *Inv. Math.* **12**:177 (1971).
23. M. Aizenman and E. M. Lieb, *J. Stat. Phys.* **24**:279 (1981).
24. W. Holsztynski and J. Slawny, *Commun. Math. Phys.* **61**:177 (1978).
25. J. K. Percus, in *Statistical Mechanics: New Concepts, New Problems, New Applications*, S. A. Rice, K. F. Freed, and J. C. Light, eds. (University of Chicago Press, Chicago, 1972), p. 337.
26. H. A. Gutowitz and J. D. Victor, *Complex Systems* **1**:57 (1987).
27. D. de Fontaine, private communication.