Variational approximation in classical lattice systems.

Schlijper, Antoine Gerard

INTRODUCTION AND SUMMARY

Statistical mechanics provides the mathematical techniques that enable, in principle, the prediction of the thermodynamic behaviour of a system on the basis of a model for its microscopic (molecular) structure[1-3]. For any realistic interacting system of macroscopic size, however, the overwhelming complexity of the calculation makes the introduction of some kind of approximation unavoidable. In the past decades, as a consequence of the increasing availability of ever larger and faster computers, there has been a growing interest in the development and application of ever more sophisticated approximation schemes.

For lattice systems the so-called cluster approximations are widely used. Here the behaviour of the whole system is extrapolated from that of one or a few relatively small subsystems (clusters), approximations being made for the influence of the surrounding system. Despite at least thirty years of widespread application the validity of such an extrapolation has never been investigated in general. In this thesis it is shown, among other things, that to guarantee the validity of this approach the chosen cluster has to satisfy an infinite number of conditions. This makes the selection of the proper cluster into what is called an undecidable problem that cannot be solved by computer[4].

It has been known for a long time that the laws of phenomenological thermodynamics hold exactly only in the thermodynamic limit, where the system is taken infinitely large with a finite density. The success of phenomenological thermodynamics then shows that a system of macroscopic size can be approximated very well by one of infinite size. This observation has led to the development of a statistical-mechanical formalism for infinite systems, sometimes called "rigorous statistical mechanics", which reflects the fact that in dealing with the subtleties of infinite systems mathematical rigor is a necessity. Particularly for infinite lattice systems a fair number of exact results have been obtained[5-7].

The development and application of approximation techniques seems at first glance to be in a way just the opposite of the rigorous approach. Yet

there is a two-sided relation.

On the one hand, approximation results may have a rigorous interpretation. Griffiths[8] first showed that for spin - $1/2$ ferromagnets the critical temperature as predicted by the mean-field approximation was more than just an approximation: it is a rigorous upper bound. Recently more results along these lines have been obtained, most of them dealing with the mean-field approximation[9-12] (which is a one-site cluster approximation).

On the other hand, comparison of results from rigorous theory with approximation techniques can lead to a better understanding of the validity of the approximations and may also point the way to better approximation schemes[13].

In this thesis certain aspects of cluster approximations for classical lattice systems are analysed in the light of recent exact results. It turns out that a number of physically reasonable and generally accepted assumptions are far from trivial or even dubious from a mathematical point of view.

All cluster approximations, including the mean-field[14], quasi-chemical[15] and Kramers-Wannier[16] approximations, can be described within the framework of one general formalism, which is called the Cluster Variation Method (CVM)[17]. The CVM has received much attention and has been widely used. Evidence for an ongoing interest is the fact that in no year in the period 1979-1984 less than 25 papers have appeared that deal with the CVM or its application. Surprisingly enough, however, with the exception of the special case of the mean-field approximation, which has been studied thoroughly, little effort has been spent on clarifying the exact nature of the approximations involved in the CVM and their influence on the results of the calculations.

In order to summarise the contents of this thesis we shall briefly describe the CVM.

To predict the thermodynamic behaviour of a system one must find the free energy function of the system in equilibrium. The Gibbs prescription to obtain this function is equivalent to a variational principle: write down the free energy for each thermodynamic state; the equilibrium state is

that state that gives minimal free energy. The free energy F is given by

$$F = U - TS \tag{1}$$

where U is the (average) energy of the system and S is the entropy.

Consider a translation-invariant lattice $\bar{L}$. With each lattice site a there is associated an occupation index or spin σ_a. For instance, if the model describes a binary alloy of atoms A and B, σ_a may take the values A or B indicating which type of atom occupies site a. If Λ is a finite set of lattice points (a cluster) of $\bar{L}$, then $\sigma_\Lambda = \{\sigma_a \mid a \in \Lambda\}$ gives the occupation of Λ, or the configuration on Λ. The probability of finding the configuration σ_Λ on Λ is written $p_\Lambda(\sigma_\Lambda)$. The probabilities p_Λ are called the occupation variables for the cluster Λ.

Now consider a large but finite system L within $\bar{L}$. L then is a finite system for which the Gibbs prescription is valid. The thermodynamic state of this system is described by the occupation variables for L. To find the equilibrium state we must write U and S as a function of these occupation variables and minimise the difference U - TS. Since the number of variables is extremely large, this is an insurmountable task.

For many models, however, the (average) energy density u = U/N (where N is the number of sites of L) can be written as a function of only the occupation variables for one or a few small clusters:

$$u = u(p_{\Lambda_1}, p_{\Lambda_2}, \ldots, p_{\Lambda_n}) \tag{2}$$

The entropy density, s = S/N, however, always depends on all the variables $p_L(\sigma_L)$:

$$s = -\frac{k}{N} \sum_{\sigma_L} p_L(\sigma_L) \log p_L(\sigma_L) \tag{3}$$

The essence of the CVM is to approximate the entropy density s by an expression that also involves only a small number of occupation variables:

$$s \tilde{=} \sum_{i=1}^{m} \alpha_i \, S[\Lambda_i] \qquad (4)$$

where the α_i are (real) coefficients and

$$S[\Lambda] = - k \sum_{\sigma_\Lambda} p_\Lambda(\sigma_\Lambda) \log p_\Lambda(\sigma_\Lambda) \qquad (5)$$

Each particular CVM approximation is determined by specification of the clusters Λ_i and the coefficients α_i.

As a consequence of the approximation in Eq. (4) an expression for the free energy per lattice site, $f = u - Ts$, is obtained that depends only on a relatively small number of occupation variables. This expression is minimised by variation of these variables. Thus, approximations are obtained to the equilibrium free energy as well as to the equilibrium values for these occupation variables. Since no values are obtained for all the other occupation variables, the CVM thus only yields a <u>partial</u> specification of an approximation to the equilibrium state.

In Chapter 1 of this thesis, after some general remarks, the relationship is established between the CVM and the rigorous version of the variational principle for the free energy density as valid in the thermodynamic limit of an infinitely extended lattice.

Chapter 2 is concerned with the entropy approximation (Eq. (4)) that is the heart of the CVM. It is shown that this approximation is based on the truncation of a series expansion that is not necessarily absolutely convergent. We show, however, that there is a particular sequence of partial sums that converges monotonically to the correct value for the entropy density, for any translationally invariant thermodynamic state. Identification of such sequences is important: since the series expansion does not converge absolutely the mere addition of an extra term in the summation may actually deteriorate the approximation.

Chapter 3 is concerned with what we call the compatibility assumption. We saw that a CVM approximation yields only partial information on the equilibrium state. It gives values for the occupation variables for those clusters that have been used in the calculation. A thermodynamic state, however, is specified by the occupation variables for <u>all</u> clusters in the

lattice. The CVM harbours the hidden assumption that the produced partial information is compatible with such a complete specification. The reason that this "compatibility assumption" is not trivially satisfied lies in the requirement of invariance under lattice translations. The very existence of this assumption is seldom recognised; when it is not justified, nonsensical results may be obtained, as has been reported on some occasions[18,19]. We give a set of conditions that is both necessary and sufficient to guarantee compatibility. In their stated form the conditions entail an infinite number of inequalities; thus, in general they are not amenable to verification.

Some more insight is obtained by studying the CVM in the limit of zero temperature. To ensure compatibility it is then necessary to know whether a given lattice model exhibits frustration. To determine this in general is a so-called undecidable problem. This means that it is not possible to construct an algorithm (i.e. write a computer program) that will solve the problem in a finite amount of time for any arbitrary lattice model. For the CVM this means in particular that the question which clusters should be taken into account in setting up the approximation does not admit a general answer: the set of criteria that have to be met is infinitely large.

In a number of special situations the compatibility assumption can be justified by a more direct method: the explicit construction of the missing part of the specification. This is the subject of Chapter 4. The method of construction uses the theory of Markov chains.

In Chapter 5 it is shown that the CVM can yield rigorous upper and lower bounds on the free energy density. In particular we show that the Bethe (or quasi-chemical) approximation on the two-dimensional square lattice gives an upper bound on the free energy per site.

Background material on the statistical-mechanical theory of infinite lattice systems may be found in Ref. 7. A review of the cluster-variation method from the traditional, application-oriented point of view may be found in Ref. 20.

The original formulation of the CVM[17,21,22] was based on the idea of constructing an approximate expression for the combinatorial factor (i.e. the number of configurations with specified energy) that occurs in the

partition function and involved rather complicated counting procedures. Later, Morita[23] reformulated the CVM using a Möbius transformation formalism; the Möbius transformation serves to automate the counting procedures of the original formulation. Morita's formulation underlies the one we present in Chapter 1; Morita considered only finite lattices, however. Finally a paper by Woodbury should be mentioned[24]; Woodbury showed that a number of cluster approximations may be derived from considerations based on general properties of the entropy set function S[Λ]. Notions such as conditional entropy, Markov process and strong subadditivity are already implicitly present in his approach.

We want to end this introduction with a few remarks on the phenomenon of phase coexistence as it manifests itself in the CVM. It has been known for a long time that phase transitions in the sense of non-analytic behaviour of thermodynamic functions or of non-uniqueness of equilibrium states (coexistence of phases) can occur only in the thermodynamic limit of an infinitely large system[25]. In the lattice systems we are considering here, this follows immediately from the fact that for any finite system the free energy is a strictly convex functional of the thermodynamic state. Only in the thermodynamic limit the strictness is lost and there may be more than one equilibrium state. The CVM is, in a way, a finite-system calculation; however, since it employs an approximate free energy functional that is not necessarily convex, there may be more than one minimising set of occupation variables. By means of the familiar Maxwell construction this is then interpreted as indicating separation into pure phases.

In this thesis we show in Chapter 5 that the use of increasingly accurate entropy approximations in the CVM results in increasingly accurate information on some equilibrium state or states. We expect these to be the extremal equilibrium states (which represent the pure phases). Were this not to be true, it would mean that the CVM would miss part of the region of phase coexistence at any level of accuracy: a state that is interpreted as a pure phase on the edge of the coexistence region may then be in fact non-extremal, thus a mixture of pure phases and actually inside the coexistence region. A proof to substantiate our expectation would thus be of interest, but is as yet lacking.

References for the

1. T.L. Hill, "I Wesley, Readi
2. K. Huang, "St 1963).
3. R.K. Pathria, reprinted 197
4. J.E. Hopcroft
5. D. Ruelle, "S Reading, Mass
6. D. Ruelle, "T Massachusetts
7. R.B. Israel, University Pre
8. R.B. Griffiths
9. B. Simon, J. &
10. P.A. Pearce,
11. J. Slawny, J.
12. B. Siu, J. Sta
13. F. Peruggi, J.
14. W.L. Braggs & (1934).
15. E.A. Guggenhei
16. H.A. Kramers &
17. R. Kikuchi, Ph
18. C.E. Campbell
19. M. Kaburagi, T (1982).
20. D.M. Burley, i C. Domb & M.S.
21. R. Kikuchi, J.
22. M. Kurata, R.
23. T. Morita, J.
24. G.W. Woodbury,
25. C.N. Yang & T.

References for the introduction and summary

1. T.L. Hill, "Introduction to Statistical Thermodynamics", (Addison-Wesley, Reading, Massachusetts, 1960).
2. K. Huang, "Statistical Mechanics", (John Wiley & Sons, New York, 1963).
3. R.K. Pathria, "Statistical Mechanics", (Pergamon Press, Oxford, 1972; reprinted 1978).
4. J.E. Hopcroft, Sci. Am. 251, 70 (1984).
5. D. Ruelle, "Statistical Mechanics: Rigorous Results", (W.A. Benjamin, Reading, Massachusetts, 1969; reprinted 1977).
6. D. Ruelle, "Thermodynamic Formalism", (Addison-Wesley, Reading, Massachusetts, 1978).
7. R.B. Israel, "Convexity in the Theory of Lattice Gases", (Princeton University Press, Princeton, New Jersey, 1979).
8. R.B. Griffiths, Commun. Math. Phys. 6, 121 (1967).
9. B. Simon, J. Stat. Phys. 22, 491 (1980).
10. P.A. Pearce, J. Stat. Phys. 25, 309 (1981).
11. J. Slawny, J. Stat. Phys. 32, 375 (1983).
12. B. Siu, J. Stat. Phys. 38, 519 (1985).
13. F. Peruggi, J. Math. Phys. 25, 3303, 3316 (1984).
14. W.L. Braggs & E.J. Williams, Proc. Roy. Soc. (London) A 145, 699 (1934).
15. E.A. Guggenheim, Proc. Roy. Soc. (London) A 148, 304 (1935).
16. H.A. Kramers & G.H. Wannier, Phys. Rev. 60, 252, 263 (1941).
17. R. Kikuchi, Phys. Rev. 81, 988 (1951).
18. C.E. Campbell & M. Schick, Phys. Rev. A5, 1919 (1972).
19. M. Kaburagi, T. Tonegawa & J. Kanamori, J. Phys. Soc. Jpn. 51, 3857 (1982).
20. D.M. Burley, in "Phase Transitions and Critical Phenomena", Vol. 2, C. Domb & M.S. Green, Eds. (Academic Press, New York, 1972), Chapter 9.
21. R. Kikuchi, J. Chem. Phys. 19, 1230 (1951).
22. M. Kurata, R. Kikuchi & T. Watari, J. Chem. Phys. 21, 434 (1953).
23. T. Morita, J. Phys. Soc. Jpn. 12, 753, 1060 (1957).
24. G.W. Woodbury, Jr., J. Chem. Phys. 47, 270 (1967).
25. C.N. Yang & T.D. Lee, Phys. Rev. 87, 404 (1952).