SYSTEMS WITH SEPARABLE INTERACTIONS

BIBLIOTHEEK
werk theoretische natuurkunde
Nieuwsteeg 18-Leiden Nederland

P. A. J. TINDEMANS

- 9 APR. 1975

SYSTEMS WITH SEPARABLE INTERACTIONS

Proefschrift

ter verkrijging van de graad van doctor in de wiskunde en natuurwetenschappen aan de Rijksuniversiteit te Leiden, op gezag van de Rector Magnificus Dr. A.E. Cohen, hoogleraar in de faculteit der letteren, volgens besluit van het college van dekanen te verdedigen op woensdag 23 april 1975 te klokke 16.15 uur

door

PETER ANTOON JOZEF TINDEMANS
geboren te Nederweert in 1947

INSTITUUT-LORENTZ
voor theoretische natuurkunde
Nieuwsteeg 18-Leiden-Nederland

kast dissertaties

Krips Repro - Meppel

Promotoren: Dr. H.W. Capel

Prof.dr. P.W. Kasteleyn

Het in dit proefschrift beschreven werk werd gedaan als deel van het onderzoekprogramma van de "Stichting voor Fundamenteel Onderzoek der Materie" (F.O.M.) met financiële steun van de "Nederlandse Organisatie voor Zuiver-Wetenschappelijk Onderzoek" (Z.W.O.).

STELLINGEN.

I. Het bewijs dat Ky Fan geeft voor de matrixongelijkheid (II.1) is minder geschikt dan het hier gegeven bewijs als het gaat om het afleiden van de noodzakelijke voorwaarden waaronder de ongelijkheid een gelijkheid wordt.

Ky Fan, Proc.N.A.S. (USA) 37 (1951) 760.
Dit proefschrift, hoofdstuk II.

II. De berekening van de vrije energie voor systemen met separabele interakties, zoals gegeven in dit proefschrift, kan uitgebreid worden tot systemen waarin naast de separabele interakties ook andere interakties (bv. van korte dracht) aanwezig zijn. Wel zullen de separabele interakties dan doorgaans uitsluitend van het ferromagnetische type moeten zijn.

P.A.J.Tindemans en H.W.Capel, wordt gepubliceerd in Physica.

III. Hallers beschouwt in de molekulaire-veldbenadering enkele modellen voor het optreden van metaal-halfgeleiderovergangen. Een daarvan is dan ekwivalent met een spin 1-model met nulveldsplitsing. In het biezonder kan het optreden van eerste-ordeovergangen in het metaal-halfgeleidersysteem gekoppeld worden aan het optreden van eerste-ordeovergangen in het tweede systeem.

J.J.Hallers, proefschrift Groningen 1972,hoofdstuk IV.
H.W.Capel, Physica 32 (1966) 966.

IV. Ten onrechte denkt Danielian door het invoeren van een absolute "interval"-temperatuurschaal in plaats van de gebruikelijke absolute temperatuurschaal, de onbereikbaarheid van het absolute nulpunt te kunnen ontkoppelen van de derde hoofdwet van de thermodynamika.

A.Danielian, Phys.Lett. 51A (1975) 61.

V. Een 1-dimensionaal spinmodel met symmetrische interakties tussen de x-en y-komponenten van de spins van naaste buren is slechts in schijn algemener dan het XY-model. Dat is niet het geval voor een systeem met antisymmetrische interakties.

S.A.Pikin en V.M.Tsukernik, Sov.Phys. JETP 23 (1966) 914.

VI. Het diagonaliseren van een hamiltoniaan die bilineair is in fermion-kreatie- en annihilatieoperatoren kan in een aantal gevallen vereenvoudigd worden door gebruik te maken van een"deeltje-gat"transformatie in plaats van de Bogoliubov-Valatintransformatie.

VII. De Landauontwikkeling van de vrije energie, die bv. gebruikt wordt om de eigenschappen ven een systeem in de buurt van een trikritisch punt te onderzoeken, is vaak in eerste instantie een ontwikkeling in twee of meer ordeparameters. Het is van belang te onderzoeken of uiteindelijk toch niet volstaan kan worden met één ordeparameter. De klasse van in dit proefschrift behandelde modellen biedt daartoe een goed aangrijpingspunt.

VIII. De beschouwing die McWeeny houdt bij het invoeren van één- en tweedeeltjesdichtheidsfunkties voor een systeem van N elektronen, is gebaseerd op een onjuiste interpretatie van de golffunktie die zo'n systeem beschrijft.

R.McWeeny, On the origin of electronic properties of molecules, in: Orbital theories of molecules and solids (ed. by N.H.March), Oxford 1974.

IX. Het feit dat in het experimenteel gevonden gedrag van de magnetisatie van kwasi-tweedimensionale antiferromagneten geen aanwijzigingen worden gevonden voor een overgang naar driedimensionaal gedrag, kan op eenvoudige wijze worden verklaard met behulp van de theorie van Liu en Stanley.

L.L.Liu en H.E.Stanley, Phys.Rev.B8 (1973) 2279.
zie ook: C.A.W.Citteur. proefschrift Leiden 1973.

X. Het "Demokratisch Tegenvoorstel" voor de organisatie van de wetenschapsbeoefening in Nederland verdient op een aantal essentiële punten de voorkeur boven de voorstellen gedaan in de "Nota Wetenschapsbeleid". Tot die punten behoort de positie van het universitaire onderzoek.

BWA/VWO-Werkgroep Wetenschapsbeleid,W&S juni 1974.
Nota Wetenschapsbeleid,Staatsuitgeverij januari 1975.

XI. Wetenschapstheoretische diskussies over de grenzen van de autonome ontwikkeling van de wetenschap verdienen meer aandacht, ook binnen de natuurwetenschappen. Met name zullen de gevolgen van dergelijke ideeën voor het werk in onderzoekgroepen (bv. universitaire vakgroepen) onderzocht moeten worden.

G.Böhme, W.van den Daele en W.Krohn, Z.f.Soziologie,1 (1972)302, 2 (1973) 128.

Peter Tindemans, 23 april 1975.

ter herinnering aan mijn moeder

aan mijn vader

aan José.

C O N T E N T S

I. INTRODUCTION AND SUMMARY 7

References 13

II. AN INEQUALITY FOR THE TRACE OF A PRODUCT OF MATRICES 15

1. Introduction 15
2. The theorem 15
3. Proof of the theorem 17
4. Remarks 23
5. The Hölder inequality 25

References 30

III. AN EXACT CALCULATION OF THE FREE ENERGY IN SYSTEMS WITH SEPARABLE INTERACTIONS 31

1. Introduction 31
2. The Hamiltonian 33
3. Commuting operators 36
4. An integral representation 37
5. Investigation of the absolute minimum of G_1 41
6. An upper bound for the free energy 44
7. A lower bound for the free energy 45
8. Discussion 50

Appendix A 54

Appendix B 57

Appendix C 60

References 61

IV. AN EXACT CALCULATION OF THE FREE ENERGY IN SYSTEMS WITH SEPARABLE INTERACTIONS II 63

1. Introduction 63
2. An upper bound for the free energy 66
3. An integral representation for the partition function 70
4. The absolute minimum 75
5. The second derivatives of G_1 82
6. The correction due to the second derivatives 88

7. Parameters independent of $\vec{\xi}$. 91
Appendix A 94
Appendix B 98
Appendix C 101
Appendix D 105
Appendix E 106
References 108

V. APPLICATIONS OF THE GENERAL RESULTS 110
1. Introduction 110
2. Extension to general T 111
3. Generalized interactions 114
4. Bragg-Williams formulation 117
5. Short range and long range interactions 120
6. Dipolar and quadrupolar ordering 123
7. A model for antiferromagnetism 126
8. Ferromagnetism and superconductivity 129
References 131

SAMENVATTING 133

GEGEVENS 135

Chapters III and IV have appeared before in Physica as
Tindemans, P.A.J. and Capel, H.W., Physica 72 (1974) 433.
Tindemans, P.A.J. and Capel, H.W., Physica 75 (1974) 407.
(Physica is published by The North-Holland Publishing Co., Amsterdam).

I. INTRODUCTION AND SUMMARY

One of the most important problems in classical or quantumstatistical mechanics is the calculation of the thermodynamic properties of a many-body system consisting of a large number of particles on the basis of the microscopic interactions between the particles. For this purpose one has to evaluate the logarithm of the partition function of the system (or the free energy) from which other thermodynamic quantities can be obtained. However, usually the interactions between the particles are such that an exact calculation is impossible. Therefore, in many cases one has to use approximate theories. One of the most important, and from a qualitative point of view most successful, approximation schemes is based on the construction of a one-particle Hamiltonian. For this we note that the influence that is exerted on a given particle as a result of its interactions with all other particles, can be viewed at as an instantaneous field acting on the particle. In the one-particle approximation this instantaneous field is replaced by an average, effective field which is commonly denoted by "molecular" or "internal" field. One then constructs an effective Hamiltonian which contains the molecular field (or fields) as parameters. The best approximation in the framework of the one-particle description is obtained by choosing the parameters in such a way that the free energy calculated from the effective one-particle Hamiltonian assumes its minimal value. All these one-particle theories originate from the molecular-field theory introduced by Weiss [1] in order to describe the phenomenon of ferromagnetism. The first statistical formulations of these theories have been given by Hartree [2] and Fock [3], and by Bragg and Williams [4], and are denoted by the HF and BW approximations.

The basic idea behind these approximations is to neglect completely the fluctuations in the instantaneous field on a certain particle due to its interaction with the other particles. Therefore it is to be expected that in situations where these fluctuations cannot be neglected, the molecular-field approximation breaks down. In fact, all one-particle theories lead to definite predictions for the so-called critical exponents, which are usually in disagreement with experimental data. However, according to recent

developments [5] in the theory of critical phenomena the critical exponents originating from a one-particle theory seem to be correct in the neighbourhood of a tricritical point. It can be argued that the width of the critical region, where fluctuations are important, decreases if the effective number of particles interacting with a given particle, grows. This corroborates the general idea that molecular-field approximations in which fluctuations in instantaneous fields are discarded, become better and better the larger the number of particles in interaction with any particle is.

Given a microscopic interaction, its range determines how many particles are interacting with a chosen particle. Therefore one expects there to be a relationship between the reliability of molecular-field type of approximations and the range of the interactions in a system.

In investigating such a relation one can distinguish between two situations. Firstly, one may study systems with interactions of a long, but finite range. In the process of taking the thermodynamic limit the range becomes smaller and smaller compared to the size of the system and afterwards a second limiting procedure is applied in which the range tends to infinity. Usually this situation leads to a great deal of mathematical complications. There are not many systems for which one can carry out the calculations in the case of interactions of finite range to such an extent that the second limit can be taken after the thermodynamic limit.

The second possibility is to consider interactions which have a range proportional to the size of the system. In the thermodynamic limit therefore both the size of the system and the range of the interaction tend to infinity. An important simplification occurs if the interactions are assumed to be separable. By a separable interaction one means an interaction between pairs of particles, $\Sigma_{k,\ell} V(k,\ell)$, which is such that the operator $V(k,\ell)$ describing the interaction between particles k and ℓ, can be written as a product $V(k)\cdot V(\ell)$. One can distinguish between negative-definite (attractive), and positive-definite (repulsive) separable interactions. Clearly, if the operator $V(k)$ is independent of k, the interaction between any pair of particles is the same, irrespective of their distance. As a consequence these interactions can be called equivalent-neighbour interactions and can be viewed upon as interactions of an extremely long range. Though, of course, such interactions represent a rather idealized situation, the advantage of considering them is that one may be able to investigate in a

more exact way a rather large class of physical situations. In particular, one may expect effective-field approximations to be exact for systems with this kind of interactions.

In this thesis systems with separable interactions will be investigated and in particular their free energy per particle will be calculated. In this calculation one has to face the difference between classical and quantum-mechanical systems. For classical systems the operators in the Hamiltonian commute among each other. Quantummechanical systems are much more difficult to deal with since the Hamiltonian contains non-commuting operators. In the partition function one finds the operator $e^{-\beta\mathcal{H}}$, where $\mathcal{H}$ is in many cases a sum of a number of operators which on themselves can be diagonalized in a trivial way. However, in general not much can be said about the diagonalization of their sum $\mathcal{H}$.

Most investigations on systems with an interaction of a long, but finite range have been carried out for classical systems. In connection with this we can mention the work by Kac, Uhlenbeck and Hemmer [6] on the van der Waals gas, where the interaction between particles at a distance r is chosen to be $-\gamma^{-1}e^{-\gamma r}$, and also the work by Siegert and Vezzetti [7] on Ising systems. Lebowitz and Penrose [8], however, have given a quantummechanical extension of the theory of the van der Waals gas. As a general conclusion the molecular-field approximation turns out to be allright except in a small interval around the critical point which decreases with the inverse of the range of the interaction.

Systems with separable interactions have been considered for a larger class of physical situations. The simple cases of an Ising model with equivalent-neighbour interactions and the corresponding lattice gas have been treated before by Mühlschlegel and Zittartz [9] and Husimi and Temperley [10]. As a first example of a quantummechanical system with separable interactions we mention the so-called reduced BCS-Hamiltonian [11] in the theory of superconductivity. Here the interaction between Cooper pairs (i.e. pairs of electrons with opposite momenta $\vec{k}$ and $-\vec{k}$, and spins σ and $-\sigma$) is assumed to be constant in an energy interval around the Fermi energy, and zero, otherwise. In $\vec{k}$-space this interaction can be considered as an equivalent-neighbour interaction. Mühlschlegel's treatment [12] of the BCS-Hamiltonian is one of the basic ingredients for the present investigation.

So far, two different approaches to the study of quantummechanical

systems with separable interactions have been developed. First of all, we can mention the extensive investigations by Bogoliubov jr. [13], primarily on systems with negative-definite separable interactions. Perhaps due to Bogoliubov's emphasis on systems with BCS-type of interactions, the generality of his approach seems not to have been appreciated sufficiently in the literature. A number of special models has been investigated using this approach, cf. several recent Dubna preprints [14]. For a more detailed discussion of Bogoliubov's line of reasoning, we refer to section 2 of chapter V. In the second place a number of specific systems has been treated in the framework of a C^*-algebra approach [15]. Starting from a condition which ensures the validity of the Bogoliubov-Haag procedure [16] (a procedure in which certain sum operators are replaced by c-numbers), the free energy can be obtained for extremal homogeneous states satisfying the KMS condition, for each of the systems separately.

Returning now to the present investigations on systems with separable interactions, we shall give explicitly the class of Hamiltonians to be considered. They describe systems of N particles, labeled by k and ℓ.

$$\mathcal{H} = \sum_{k=1}^{N} T(k) - (2N)^{-1} \sum_{f=1}^{p} \sum_{k,\ell=1}^{N} V_f(k)V_f(\ell) + (2N)^{-1} \sum_{a=1}^{q} \sum_{k,\ell=1}^{N} W_a(k)W_a(\ell).$$

Here the first term is a kinetic energy term or a term representing the influence of external fields. The second and third term describe negative-definite and positive-definite separable interactions. They will often be referred to as "ferromagnetic" and "antiferromagnetic" interactions resp., a terminology which is borrowed from the situation that one investigates magnetic ordering. It is important to note that the labels k and ℓ need not refer to particles at a certain position, they can as well be used to denote one-particle states corresponding to another property such as the momentum. An example is provided by the BCS-reduced Hamiltonian. In such a situation one deals with a given number of one-particle states, rather than with a fixed number of particles. All thermodynamic averages are then grand-canonical averages. The operators $T(k)$, $V_f(k)$, $f = 1, \ldots, p$, and $W_a(k)$, $a = 1, \ldots, q$, are completely general, bounded, hermitean one-particle operators and in particular do not satisfy any prescribed commutation

relations. *)

We shall give a rigorous calculation of the free energy for the class of systems defined above. The expression we obtain turns out to display a molecular-field character: i.e. it can be found from an effective one-particle Hamiltonian containing a number of parameters each of which corresponds to one of the separable interactions. These parameters can be considered as order parameters and characterize the different phases in which a system can be. So in the BCS-model they are related to the energy-gap, while in magnetic systems they represent the total or a sublattice magnetization. Their values have to be determined from molecular-field equations with the additionnal requirement that the expression for the free energy in terms of these parameters be minimal. For the systems under investigation an unequivocal meaning can now be given to statements such as "The effect of a separable interaction between the particles is equivalent to an additional field acting on the particles". For systems with more general interactions which are not of the separable type, the investigations reported in this thesis give a general and unambiguous way to obtain the results which would have been found in the molecular-field approximation. For that purpose one replaces the interactions between the particles by equivalent-neighbour interactions. The general scheme given here can also be useful in establishing relationships between models which, in spite of the variety of physical situations for which they can be used, have a similar underlying mathematical structure.

Finally we shall give an outline of the method used in this thesis. The free energy per particle will be obtained from an upper and a lower bound which will be shown to be equal in the thermodynamic limit. The upper bound on the free energy is derived by means of a variational type of argument based on Bogoliubov's inequality [17]. In order to obtain a lower bound we follow a standard approach in statistical mechanics and proceed to derive an integral representation for the partition function $Z = \int e^{-NG}$, where G

---

*) In chapter V we point out that in the case that the "antiferromagnetic" operators are absent, the results are valid also if $\Sigma_k T(k)$ and $\Sigma_k V_f(k)$ are replaced by more general operators T and V_f. In that case the name "quadratic" interactions is, perhaps, to be preferred to "separable" interactions.

is in general a complex function depending on the integration variables used in this representation. Here the quamtummechanical case presents additional complications. Due to the fact that the operators in the Hamiltonian do not commute, the operator $e^{-\beta\mathcal{H}}$ cannot be factorized into a product of exponential operators, each containing one separable interaction. Since such a separation seems necessary in order to obtain a useful integral representation by applying a familiar trick due to Stratonovich [18], a special procedure is required. Mühlschlegel uses the ordering procedure due to Feynman [19] for that purpose. We prefer, however, the Trotter product formula [20] for reasons to be explained in chapter III. As a consequence of the use of Trotter's formula (or of any similar procedure, for that matter) the number of integration variables in the integral representation must go to infinity <u>before</u> the thermodynamic limit is taken. We want to obtain a lower bound on the free energy by estimating $Z = \int e^{-NG}$ to be $\int e^{-NG_1}$, where G_1 is the real part of G, and by calculating the resulting integral by means of Laplace's method. Usually, one obtains as a result of such a calculation an asymptotic series in powers of N^{-1}. The leading order term is obtained immediately from the absolute minimum of G_1. The first order correction term can be expressed in terms of the determinant of the matrix of second derivatives of G_1, evaluated at the absolute minimum. In both aspects the quantummechanical case is much more difficult than the classical one. The function G_1 contains the trace of an arbitrary number of operators and the determination of its absolute minimum depends on the possibility of finding both an estimate on such traces of products of operators, and the precise conditions under which the estimate becomes an identity. Such an estimate is provided by a generalization of Hölder's inequality to operators, which can be derived from a sharper inequality for the trace of a product of operators. In chapter II a simple and straightforward proof of the latter inequality will be given and the conditions under which it turns into an equality will be established.

As to the first correction term in the asymptotic series, a term which is related to the second derivatives of G_1, we remark that in classical cases this term can easily be shown to lead to a contribution to the free energy which is independent of the volume, and hence can be neglected in the thermodynamic limit. In the quantummechanical case, however, the essentially infinite number of integration variables makes a careful investigation necessary. In particular a crude estimate on the integral $\int e^{-NG}$ could lead

to unphysical divergencies. It will be shown that our calculation leads to a correction term which is finite except possibly on a discrete set of temperatures. For Hamiltonians with separable interactions of the "ferromagnetic" type only, it will be shown, moreover, that the temperatures where a second or higher order phase transition occurs, belong to this set.

In chapter II we deal with the inequality for the trace of a product of matrices, which leads to the Hölder inequality for operators. Also a number of other applications is discussed. In chapter III we give the calculation of the free energy for systems described by a Hamiltonian containing a number of "ferromagnetic" separable interactions. In chapter IV the calculation will be extended to Hamiltonians containing an arbitrary, finite number of "antiferromagnetic" separable interactions as well. In chapter V, first some remarks are made about an extension to Hamiltonians of a somewhat more general nature, the equivalence to a Bragg-Williams formulation is established, and another formulation avoiding the explicit introduction of "ferromagnetic" and "antiferromagnetic" operators is given. Finally we treat from our general point of view a number of specific models which have been investigated in the literature.

REFERENCES

1) Weiss, P., J.Phys.Rad. 4 (1907) 469.
2) see Hartree, D.R., Rep. Progr.Phys. 11 (1946/7) 113 for a review.
3) Fock, V., Z.Phys. 61 (1930) 126.
4) Bragg, W.L. and Williams, E.J., Proc.Roy.Soc. A145 (1934) 699.
5) Riedel, E.K. and Wegner, F.J., Phys.Rev.Lett. 29 (1972) 349.
Wegner, F.J. and Riedel, E.K., Phys.Rev. B7 (1973) 248.
6) Kac, M., Uhlenbeck, G.E. and Hemmer, P.C., J.Math.Phys. 4 (1963) 216,229.
7) Siegert, A.J.F. and Vezzetti, D.J., J.Math.Phys. 9 (1968) 2173.
8. Lebowitz, J.L. and Penrose, O., J.Math.Phys. 7 (1966) 98.
9) Mühlschlegel, B. and Zittartz, H., Z.Phys. 175 (1963) 553.
10) Temperley, H.N.V., Proc.Phys.Soc. 67 (1954) 233.
11) Bardeen, J., Cooper, L.N. and Schrieffer, J.R., Phys.Rev. 108 (1957)1175.
12) Mühlschlegel, B., J.Math.Phys. 3 (1962) 522.
13) For a review, see Bogoliubov Jr., N.N., A method for studying model Hamiltonians, Pergamon Press (Oxford, 1972).

14) Brankov, J.G., Zagrebnov, V.A. and Tonchev, N.S., Preprint JINR R4-7917, Dubna 1974.
Brankov, J.G., Zagrebnov, V.A. and Tonchev, N.S., Preprint JINR R4-7735, Dubna 1974.
Brankov, J.G., Shumovsky, A.S. and Zagrebnov, V.A., Physica 78 (1974)183.
15) We refer to work by Vertogen, Hallers, Smit, de Vries and Kraak. For references see chapters III and IV.
16) Haag, R., Nuovo Cimento 25 (1962) 287.
17) e.g. Falk, H., Am.J.Phys. 38 (1970) 858.
18) Stratonovich, R.L., Soviet Physics-Doklady 2 (1958) 416.
19) Feynman, R.P., Phys.Rev. 84 (1951) 108.
20) Trotter, H., Proc.Am.Math.Soc. 10 (1959) 545.

II. AN INEQUALITY FOR THE TRACE OF A PRODUCT OF MATRICES

1. Introduction

Estimates on the trace of a matrix play an important role in calculations in quantum statistical mechanics. This is easily understood by noting the occurrence of traces in the basic expressions for the free energy, the entropy (cf. e.g. ref. 1) or correlation functions of quantummechanical systems.

In the course of the evaluation of the free energy for the systems that will be investigated in this thesis, one naturally encounters the problem of finding an upper bound on $|\mathrm{Tr}\, A_1 \ldots A_n|$. Here $A_1, \ldots, A_n$ are arbitrary finite-dimensional matrices and n is an arbitrary integer. Secondly, and not less important, there is the problem of establishing necessary and sufficient conditions under which the inequality to be derived will turn into an equality. The right-hand side of the inequality is expressed in terms of the eigenvalues of the matrices $\sqrt{A_k^+ A_k}$, and after applying the Hölder inequality to this expression we obtain another inequality which may be called the Hölder inequality for operators.

In section 5 of this chapter a number of applications of this inequality will be discussed. These applications deal with e.g. the convexity of the free energy and with convergence properties of Dyson expansions.

2. The theorem

We state the theorem in the following way (alternative formulations will be given in section 4).

Let $U_1, \ldots, U_n$ be unitary operators on the m-dimensional complex Euclidean space $\mathbb{C}^m$. *) Let A_k for $k = 1, \ldots, n$ be semipositive definite (hermitean) operators on $\mathbb{C}^m$ with eigenvalues $\alpha_1^{(k)} \geq \alpha_2^{(k)} \geq \ldots \geq \alpha_m^{(k)} \geq 0$. Then

$$|\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n| \leq \sum_{i=1}^{m} \prod_{k=1}^{n} \alpha_i^{(k)} . \tag{1}$$

---

*) We shall not distinguish between operators on $\mathbb{C}^m$ and the corresponding $m \times m$ matrices.

Let $p \leq m$ be the largest integer such that

$$\prod_{k=1}^{n} \alpha_p^{(k)} > 0. \tag{2}$$

The equality sign in (1) holds iff there exists an orthonormal basis $\{\psi_i\}$ of $\mathbb{C}^m$, such that for the first p vectors ψ_i $(i = 1, \ldots, p)$:

$$\mathcal{A}_k \psi_i = \alpha_i^{(k)} \psi_i \quad , \quad \text{for all } k \tag{3}$$

$$U\psi_i = \lambda \psi_i \quad . \tag{4}$$

Here

$$\mathcal{A}_k \equiv U_1 \ldots U_k \, A_k (U_1 \ldots U_k)^{-1} \quad , \tag{5}$$

$$U \equiv U_1 \ldots U_n \quad . \tag{6}$$

The case $n = 1$ is trivial. For $n = 2$, the theorem has been treated first by von Neumann [2], while a simplified proof has been given by Mirsky [3].

The general case is, as far as the inequality is concerned, due to Ky Fan [4]. He derived the theorem in the form of a maximum principle (cf. the second formulation in section 4). He used as his main tool an elegant lemma by Horn [5], based on an inequality for a determinant and a lemma by Polya [6]. An extension to compound matrices *) has been given by Marcus and Moyls [7], and by Mirsky [8]. There it is shown that a similar inequality holds for these compound matrices if one replaces the simple summation in the right-hand side of (1) by the elementary symmetric functions **).

A review of related properties of matrices derived on the basis of properties of determinants, can be found in a paper by de Bruijn [9].

The proof that will be given here is rather straightforward. Use will be made of the spectral resolution of a hermitean operator A (i.e. it can be written as a linear combination of projection operators), and of a number of

---

*) A compound matrix associated with a given $m \times m$ matrix A is a $\binom{m}{r} \times \binom{m}{r}$ matrix which has as elements the subdeterminants of order r of A.

**) The rth elementary symmetric function of the numbers $a_1, \ldots, a_m$ is defined as

$$E_r(a_1, \ldots, a_m) \equiv \sum_{1 \leq i_1 < i_2 \ldots < i_r \leq m} a_{i_1} \ldots a_{i_r} \quad .$$

relatively simple properties of projection operators. In this way also the necessary conditions for the equality sign to hold can be obtained.

3. Proof of the theorem

As a first step we consider four lemma's which will be used in the proof.

Lemma 1: Let $P_1,\ldots,P_n$ be projection operators on $\mathbb{C}^m$, i.e. $P^2 = P$ and $P^\dagger = P$. ($P^\dagger$ is the hermitean conjugate of P). Then for all $\psi\in\mathbb{C}^m$:

$$(\psi,\psi) \geq (\psi,\ P_1\ldots P_n P_n\ldots P_1\psi) \quad , \tag{7}$$

with equality sign iff

$$P_k\psi = \psi \quad , \qquad k = 1,\ \ldots,\ n. \tag{8}$$

Proof: For n = 1 the lemma is trivial, since $(\psi,\ (1-P^2)\psi) = (\psi,\ (1-P)^2\psi) \geq 0$ with equality sign iff $P\psi = \psi$. Suppose it is correct for n-1. Consider $\chi \equiv P_{n-1}\ \ldots\ P_1\psi$. Then

$$(\psi,P_1\ \ldots\ P_n P_n\ \ldots\ P_1\psi) = (\chi,\ P_n^2\chi) \leq (\chi,\chi) \leq (\psi,\psi) \quad ,$$

with equality signs iff $P_n\chi = \chi$ and $P_k\psi = \psi$, $k = 1,\ \ldots,\ n-1$, respectively. Then eq. (8) immediately follows.

Lemma 2: Let $P_1,\ldots,P_n$ be projection operators, and U be a unitary operator. Then for all $\psi\in\mathbb{C}^m$:

$$(\psi,\psi) \geq |(\psi,\ P_1\ \ldots\ P_n U\psi)| \quad , \tag{9}$$

with equality sign iff

$$P_k\psi = \psi \qquad k = 1,\ \ldots,\ n \ , \tag{10}$$

$$U\psi = \lambda\psi \qquad \text{for some } \lambda \text{ with } |\lambda| = 1 \ . \tag{11}$$

Proof: Apply Schwartz's inequality and lemma 1.

$$|(\psi,\ P_1\ \ldots\ P_n U\psi)| \leq (\psi,\ P_1\ \ldots\ P_n UU^\dagger P_n\ \ldots\ P_1\psi)^{\frac{1}{2}}(\psi,\psi)^{\frac{1}{2}} \leq (\psi,\psi) \ . \tag{12}$$

The equality sign in the first inequality of (12) applies iff

$$U^\dagger P_n\ \ldots\ P_1\psi = \mu\psi \ , \tag{13}$$

for some μ, whereas the second inequality turns into an equality iff

$$P_k\psi = \psi \qquad k = 1, \ldots, n \; . \tag{14}$$

Eqs. (13) and (14) are equivalent to (10) and (11).

<u>Lemma 3</u>: Let $P_1, \ldots, P_n$ be projection operators, U a unitary operator and d_k the dimension of the subspace $E^{(k)}$ of $\mathbb{C}^m$ on which P_k projects (i.e. the number of eigenvalues equal to 1) and let q be such that $d_q = \min\{d_1,\ldots,d_n\}$. Then

$$|\mathrm{Tr}\, P_1 \ldots P_n U| \leq \min\{d_1, \ldots, d_n\} \; . \tag{15}$$

The equality sign holds iff for all eigenvectors ψ_i corresponding to eigenvalue 1 of P_q:

$$P_k\psi_i = \psi_i \quad \text{or } E^{(k)} = E^{(q)} \oplus R^{(k)} \; , \qquad k = 1, \ldots, n \; , \tag{16a}$$

(i.e. the d_q-dimensional space $E^{(q)}$ is a subspace of all $E^{(k)}$), and

$$U\psi_i = \lambda\psi_i \; . \tag{16b}$$

<u>Proof</u>: Suppose d_q is the smallest number among $d_1, \ldots, d_n$. $\mathrm{Tr}\, P_1 \ldots P_n U = \mathrm{Tr}\, P_q B$ where $B = P_{q+1} \ldots P_n U P_1 U^{-1} \ldots U P_{q-1} U^{-1} U$. Consider an orthonormal basis $\{\psi_i\}$ with respect to which P_q is diagonal. Then

$$|\mathrm{Tr}\, P_1 \ldots P_n U| \leq \sum_{i=1}^{m} (\psi_i, P_q\psi_i)|(\psi_i, B\psi_i)| \leq \sum_{i=1}^{m} (\psi_i, P_q\psi_i) = d_q \; . \tag{17}$$

In the second step use has been made of lemma 2. (B is a product consisting of projection operators and a unitary operator). The equality sign in the second step of (17) holds iff

$$P_k\psi_i = \psi_i \quad \text{for } k \geq q+1; \quad UP_kU^{-1}\psi_i = \psi_i \quad \text{for } k \leq q-1 \; ,$$

$$U\psi_i = \lambda\psi_i \; ,$$

for all eigenvectors ψ_i corresponding to eigenvalue 1 of P_q, or equivalently:

$$P_k\psi_i = \psi_i \quad \text{for all } k, \quad \text{and } U\psi_i = \lambda_i\psi_i \; . \tag{18}$$

The equality sign in the first step of (17) holds iff $\arg(\psi_i, B\psi_i)$ is independent of i. Hence $\lambda_i = \lambda$ and eqs. (16a) and (16b) have been proved.

<u>Lemma 4</u>: Let $z_i^{(k)}$ be arbitrary complex numbers, defined for k = 1, ..., n

and $i = 1, \ldots, m$. Then

$$\sum_{i_1,\ldots,i_n=1}^{m} \left(\prod_{k=1}^{n} z_{i_k}^{(k)} \right) \min\{i_1,\ldots,i_n\} = \sum_{i=1}^{m} \prod_{k=1}^{n} (z_i^{(k)} + z_{i+1}^{(k)} + \ldots + z_m^{(k)}). \quad (19)$$

<u>Proof</u>: For arbitrary values $i_1, i_2, \ldots, i_n$ both sides of (19) contain $\Pi_k z_{i_k}^{(k)}$. We show that the coefficients in front of these terms are equal. The right-hand side of eq. (19) can be written as:

$$(\text{r.h.s.}) = \prod_k (z_1^{(k)} + z_2^{(k)} + \ldots + z_m^{(k)}) + \prod_k (z_2^{(k)} + \ldots + z_m^{(k)}) + \ldots + \prod_k z_m^{(k)} \equiv$$

$$\equiv T_1 + T_2 + \ldots + T_m \quad . \quad (20)$$

Consider an arbitrary term $\Pi_k z_{i_k}^{(k)}$. Let q be the smallest value of the subscripts $i_1, \ldots, i_n$. This particular term is found in the (r.h.s.) once in $T_1, T_2, \ldots, T_{q-1}, T_q$, but not in $T_{q+1}, T_{q+2}, \ldots, T_m$. So the coefficient in front of this term is q, as it is in the (l.h.s.).

This completes the proof of the lemma's. We now proceed to the theorem (1) itself.

<u>Proof</u>: We write

$$|\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n| = |\mathrm{Tr}\, \mathcal{A}_1 \ldots \mathcal{A}_n U| \ , \quad (21)$$

where

$\mathcal{A}_k \equiv U_1 \ldots U_k A (U_1 \ldots U_k)^{-1}$ and $U \equiv U_1 \ldots U_n$ as in (5) and (6). $\mathcal{A}_k$ is hermitean and has the same eigenvalues as A_k. We use the spectral resolution of a hermitean operator A with eigenvalues $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_m$:

$$A = (\alpha_1 - \alpha_2) P_1 + (\alpha_2 - \alpha_3) P_2 + \ldots + (\alpha_{m-1} - \alpha_m) P_{m-1} + \alpha_m P_m \quad . \quad (22)$$

Here

$$P_j = \sum_{i \leq j} P_{\alpha_i} \ , \quad (23)$$

where P_{α_i} is the projection operator on the i-dimensional subspace of the ith eigenvector, which belongs to eigenvalue α_i. On account of

$$P_{\alpha_i} P_{\alpha_j} = P_{\alpha_i} \delta_{ij} \ , \quad (24)$$

P_j is the projection operator on the j-dimensional subspace of the largest j eigenvalues. Define:

$$z_i^{(k)} \equiv \alpha_i^{(k)} - \alpha_{i+1}^{(k)} \quad \text{if } i = 1, \ldots, m-1 \text{ and } z_m^{(k)} \equiv \alpha_m^{(k)}, \text{ for all } k. \tag{25}$$

On substituting

$$\mathcal{A}_k = \sum_{i_k=1}^{m} z_{i_k}^{(k)} P_{i_k}^{(k)}, \quad \text{for } k = 1, \ldots, m \quad , \tag{26}$$

we have from (21)

$$|\mathrm{Tr} \prod_{k=1}^{n} U_k A_k| \leq \sum_{i_1,\ldots,i_n=1}^{m} \left(\prod_{k=1}^{n} z_{i_k}^{(k)}\right) |\mathrm{Tr} \left\{\left(\prod_{k=1}^{n} P_{i_k}^{(k)}\right)U\right\}|, \tag{27}$$

with equality sign iff all terms $\mathrm{Tr}\left(\prod_{k=1}^{n} P_{i_k}^{(k)}\right)U$ for which $\Pi_k z_{i_k}^{(k)}$ does not vanish, have the same argument.

Using lemma's 3 and 4 we find

$$|\mathrm{Tr} \prod_{k=1}^{n} U_k A_k| \leq \sum_{i_1,\ldots,i_n=1}^{m} \left(\prod_{k=1}^{n} z_{i_k}^{(k)}\right) \min\{i_1,\ldots,i_n\} = \sum_{i=1}^{m} \prod_{k=1}^{n} \alpha_i^{(k)} . \tag{28}$$

The proof of inequality (1) is hereby completed.

We now investigate the necessary and sufficient condition under which the inequality turns into an equality.

Eq. (2) implies that there exists a number k_0 such that

$$\alpha_i^{(k_0)} = 0 \quad \text{for } i = p+1, \ldots, m \qquad (p < m) \quad . \tag{29}$$

If (3) and (4) are satisfied, we have

$$|\mathrm{Tr}\, \mathcal{A}_1 \ldots \mathcal{A}_n U| = |\mathrm{Tr}\, \mathcal{A}_{k_0+1} \ldots U \mathcal{A}_1 \ldots \mathcal{A}_{k_0}| =$$

$$= |\lambda \sum_{i=1}^{p} \prod_{k=1}^{n} \alpha_i^{(k)} + \sum_{i=p+1}^{m} (\psi_i, \mathcal{A}_{k_0+1} \ldots U \ldots \mathcal{A}_{k_0} \psi_i)| \; .$$

Since

$$\mathcal{A}_{k_0}\psi_i = 0 \quad \text{for } i = p+1, \ldots, m \quad , \tag{30}$$

we find that

$$|\mathrm{Tr}\, \mathcal{A}_1 \ldots \mathcal{A}_n U| = \sum_{i=1}^{p} \prod_{k=1}^{n} \alpha_i^{(k)} \quad , \tag{31}$$

thereby proving the sufficiency of conditions (3) and (4).

Suppose now the equality sign holds. Then we have in particular from (27) and (28):

$$|\mathrm{Tr}\ \{(\prod_{k=1}^{n} P_{i_k}^{(k)})U\}\ | = \min\{i_1, \ldots, i_n\}\ , \tag{32}$$

if

$$\prod_{k=1}^{n} z_{i_k}^{(k)} > 0\ . \tag{33}$$

In order to prove (3) and (4) we shall consider suitable sets of integers $\{i_1,\ldots,i_n\}$ such that (33) is satisfied.

Let us suppose that $z_i^{(k)} > 0$ and let $E_i^{(k)}$ be the (i-dimensional) space on which the operator $P_i^{(k)}$ projects. We show that $E_i^{(k)} \equiv E_i$, i.e. $E_i^{(k)}$ is independent of k; furthermore we shall prove that

$$E_i \subset E_j\ , \quad \text{for } i < j\ , \quad i,j \leq p \tag{34}$$

(p being the largest integer such that $\Pi_{k=1}^{n}\alpha_p^{(k)} > 0$, as defined by (2)).

We may restrict ourselves to those values of i such that $z_i^{(k)} > 0$ for at least one value of k, otherwise $P_i^{(k)}$ would not appear in any $\mathcal{A}_k$.

Let us select a pair $i,j \leq p$, $i < j$, and consider first the following sets of integers.

If $z_i^{(k)} > 0$, we take $i_k = i$. If $z_i^{(k)} = 0$, i_k may be chosen to be any integer $q_k > i$ such that $z_{q_k}^{(k)} > 0$. (There is always such a value q_k. If $\alpha_m^{(k)} > 0$, we can choose $q_k = m$. If $\alpha_m^{(k)} = 0$, however, there must be a q_k with $p \leq q_k \leq m$ since $\alpha_p^{(k)} > 0$).

So for such a set we have $\min\{i\ ,\ldots,i_n\} = i$. Applying the necessary conditions for the equality sign in lemma 3, we find that the subspaces $E_i^{(k)}$, defined for $z_i^{(k)} > 0$, are independent of k. The resulting i-dimensional space is denoted by E_i. Since this line of reasoning can be applied to any number $\leq p$, we find in particular for a fixed value $j > i$ that all operators $P_j^{(k)}$ (for all k such that $z_j^{(k)} > 0$) project on the same space E_j.

Now the relation $E_i \subset E_j$ is trivial if there exists a value k such that both $P_i^{(k)}$ and $P_j^{(k)}$ occur in $\mathcal{A}_k$. If such a k does not exist, we consider a special set satisfying the requirements mentioned above with $i_\ell = i$ and $i_m = j$ for some $\ell \neq m$. Frmm the conditions for the equality sign in lemma 3 we then have

$$E_j^{(m)} = E_i^{(\ell)} \oplus R^{(m)}\ ,$$

where $R^{(m)}$ is a space with dimension $j-i > 0$. Since $E_j^{(m)} = E_j$ and $E_i^{(\ell)} = E_i$, we have $E_i \subset E_j$. So far we have investigated all projection operators $P_i^{(k)}$, $i \leq p$, for k such that $z_i^{(k)} > 0$.

We now must consider the projection operators $P_q^{(k)}$ with $q > p$ and $z_q^{(k)} > 0$. If for given k, $z_p^{(k)} = 0$, we select for $i=p$ some special sets of the type considered above, viz. those where q subsequently has all values $> p$ such that $z_q^{(k)} > 0$. The conditions for the equality sign in lemma 3 tell us then that

$$E_q^{(k)} = E_p \oplus R_q^{(k)} \quad . \tag{35a}$$

For k-values such that $z_p^{(k)} > 0$, it is easily seen from the spectral resolution that

$$E_q^{(k)} = E_p \oplus R_q^{(k)} \quad . \qquad (z_q^{(k)} > 0) \quad . \tag{35b}$$

From (34) and (35) it is obvious that there exists a set of p orthonormal vectors $\psi_1, \ldots, \psi_p$, such that

$$P_i \psi_\ell = \begin{cases} \psi_\ell & \text{if } \ell \leq i \qquad (i \leq p) \\ 0 & \text{if } \ell > i \quad , \end{cases} \tag{36}$$

and

$$P_i^{(k)} \psi_\ell = \psi_\ell \quad \text{if } i > p \text{ and } z_i^{(k)} > 0 \; . \tag{37}$$

Starting from the spectral resolution

$$\mathcal{A}_k = (\alpha_1^{(k)} - \alpha_2^{(k)})P_1 + \ldots + (\alpha_p^{(k)} - \alpha_{p+1}^{(k)})P_p + (\alpha_{p+1}^{(k)} - \alpha_{p+2}^{(k)})P_{p+1}^{(k)} + \\ + \ldots + \alpha_m P_m^{(k)} \quad , \tag{38}$$

where terms with $\alpha_i^{(k)} - \alpha_{i+1}^{(k)} = 0$ do not contribute, we find then immediately that condition (3) holds, i.e.

$$\mathcal{A}_k \psi_i = \alpha_i^{(k)} \psi_i \quad , \qquad i = 1, \ldots, p \; .$$

As to equation (4) we note that the conditions for the equality sign in lemma 3 applied in the case $i = p$ of a set of the above-mentioned type, also imply that $U\psi = \lambda\psi$ for any $\psi \in E_p$. Then (4) is obvious.

4. Remarks

In this section we give two alternative formulations and a slight generalization of the inequality (1). Here $A_1, \ldots, A_n$ will denote arbitrary complex $m \times m$ matrices with singular values $\alpha_1^{(k)} \geq \alpha_2^{(k)} \geq \ldots \geq \alpha_m^{(k)}$, $k = 1, \ldots, n$.

a) Consider the so-called polar decomposition: for every complex $m \times m$ matrix A there exists a unitary matrix S such that $A = S\sqrt{A^\dagger A}$ [10].

Then it is obvious that the inequality (1) can be replaced by

$$|\mathrm{Tr}\, A_1 \ldots A_n| \leq \sum_{i=1}^{m} \prod_{k=1}^{n} \alpha_i^{(k)} . \tag{39}$$

If one specifies the matrices $A_1, \ldots, A_n$ to be positive definite hermitean, then it immediately follows from (3) that the equality sign in (39) holds iff $A_1, \ldots, A_n$ are such that there exists an (orthonormal) basis ψ_i, $i=1,\ldots,m$, satisfying

$$A_k \psi_i = \alpha_i^{(k)} \psi_i , \quad \text{for all } k . \tag{40}$$

b) Consider the expression $|\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n|$ for all unitary $m \times m$ matrices $U_1, \ldots, U_n$ and fixed matrices $A_1, \ldots, A_n$. From (39) it is clear that

$$|\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n| \leq \sum_{i=1}^{m} \prod_{k=1}^{n} \alpha_i^{(k)} . \tag{41}$$

Using the relations $A = S\sqrt{A^\dagger A}$, $\sqrt{A^\dagger A} = V \Lambda V^{-1}$, $\Lambda = W \mathring{\Lambda} W^{-1}$, where S, V and W are unitary matrices, Λ and $\mathring{\Lambda}$ diagonal and where $\mathring{\Lambda}$ has the same entries as Λ but ordered in decreasing magnitude, it is easy to see that the equality sign in (41) can be realized. Hence we arrive at

$$\max_{U_1, \ldots, U_n} |\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n| = \sum_{i=1}^{m} \prod_{k=1}^{n} \alpha_1^{(k)} . \tag{42}$$

Since in this formulation one considers $|\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n|$ for given matrices $A_1, \ldots, A_n$ as a function of $U_1, \ldots, U_n$, it is desirable to give the necessary and sufficient conditions on the matrices $U_1, \ldots, U_n$ such that the maximum is attained.

It is not difficult to show that the conditions (3) and (4) are

---

*) The singular values of a matrix A are the non-negative square roots of the eigenvalues of $A^\dagger A$, where $A^\dagger$ is the hermitean conjugate of A.

equivalent to the following statement.

Let $\phi_1^{(k)},\ldots,\phi_p^{(k)}$ be any set of p orthonormal eigenvectors belonging to the p largest eigenvalues $\alpha_1^{(k)}, \ldots, \alpha_p^{(k)}$ of $\sqrt{A_k^\dagger A_k}$ ordered in decreasing magnitude.

Let Z_k be any unitary operator on $\mathbb{C}^m$ that satisfies the relations

$$Z_k\phi_i^{(k)} = \phi_i^{(k-1)} \quad , \qquad i = 1, \ldots, p \; ; \qquad k = 2, \ldots, n.$$
$$Z_1\phi_i^{(1)} = \phi_i^{(n)} \quad . \tag{43}$$

Note that Z_k depends on, but is in general not uniquely determined by the eigenvectors $\phi_1^{(k)}, \ldots, \phi_p^{(k)}$. Then the unitary matrices $U_1,\ldots, U_n$ should satisfy the relation

$$U_k = Z_k S_k^\dagger \qquad k = 1, \ldots, n \; , \tag{44}$$

where S_k is defined by $A_k = S_k\sqrt{A_k^\dagger A_k}$ and Z_k satisfies (43) for a certain choice of eigenvectors $\phi_1^{(k)}, \ldots, \phi_p^{(k)}$, $k = 1, \ldots, n$.

c) Let $\psi_1, \ldots, \psi_m$ be an arbitrary orthonormal basis of $\mathbb{C}^m$, and let P_ℓ be the projection operator on the space spanned by the first ℓ of these vectors. Since P_ℓ has ℓ eigenvalues 1 while all other eigenvalues are zero, one finds by applying (39) to the product $P_\ell A_1 \ldots A_n$

$$\left|\sum_{i=1}^{\ell} (\psi_i, A_1\ldots A_n\psi_i)\right| = |\mathrm{Tr}\, P_\ell A_1\ldots A_n| \le \sum_{i=1}^{\ell}\prod_{k=1}^{n} \alpha_i^{(k)} \quad . \tag{45}$$

If we now replace A_1 by DA_1 where the unitary diagonal matrix D has elements

$$(\psi_i, D\psi_j) = \delta_{ij}\, \exp[-i\, \arg(\psi_i, A_1\ldots A_n\psi_i)\,] \, ,$$

then eq. (45) reduces to

$$\left|\sum_{i=1}^{\ell} (\psi_i, DA_1A_2\ldots A_n\psi_i)\right| = \sum_{i=1}^{\ell} |(\psi_i, A_1\ldots A_n\psi_i)| \le \sum_{i=1}^{\ell}\prod_{k=1}^{n} \alpha_i^{(k)} \quad . \tag{46}$$

Define

$$b_i = |(\psi_i, A_1\ldots A_n\psi_i)| \text{ and } a_i = \prod_k \alpha_i^{(k)} \quad . \tag{47}$$

We label the basis vectors in such a way that $b_1 \ge b_2 \ge \ldots \ge b_m$. Eq. (46) now reads

$$b_1 + b_2 + \ldots + b_\ell \le a_1 + \ldots + a_\ell \quad . \qquad \ell \le m \quad . \tag{48}$$

Applying a lemma of Polya [6], we find for an arbitrary convex increasing function $\omega(x)$

$$\sum_{i=1}^{\ell} \omega(b_i) \leq \sum_{i=1}^{\ell} \omega(a_i) \quad . \tag{49}$$

5. The Hölder inequality

A useful inequality can be derived from (1) by applying the Hölder inequality to its right-hand side. If $0 < \theta_k^{-1} < 1$, $\Sigma_{k=1}^{n} \theta_k^{-1} = 1$ and $\alpha_i^{(k)} > 0$, the Hölder inequality reads

$$\sum_{i=1}^{m} \prod_{k=1}^{n} \alpha_i^{(k)} \leq \prod_{k=1}^{n} \left[\sum_{i=1}^{m} (\alpha_i^{(k)})^{\theta_k} \right]^{\theta_k^{-1}} , \tag{50}$$

with equality sign iff

$$(\alpha_i^{(k)})^{\theta_k} = \lambda_{k\ell} (\alpha_i^{(\ell)})^{\theta_\ell} , \tag{51}$$

for every pair (k,ℓ) with $\lambda_{k\ell}$ independent of i.

So we obtain from (1) the Hölder inequality for operators

$$|\mathrm{Tr}\, U_1 A_1 \ldots U_n A_n| \leq \prod_{k=1}^{n} \left[\mathrm{Tr}\, A_k^{\theta_k} \right]^{\theta_k^{-1}} , \tag{52}$$

with equality sign iff

$$\mathcal{A}_k^{\theta_k} = \lambda_{k\ell} \mathcal{A}_\ell^{\theta_\ell} \quad \text{for all } k, \ell , \tag{53}$$

$$U = \lambda I \quad . \tag{54}$$

Here $\mathcal{A}_k \equiv U_1 \ldots U_k A_k (U_1 \ldots U_k)^{-1}$ and $U \equiv U_1 \ldots U_n$, as in (5) and (6).

A different treatment of the Hölder inequality for n = 2, 3, which is also valid for compact operators in an infinite dimensional Hilbert space, has been given by Dunford and Schwartz [11], based on a rather deep convexity theorem by M. Riesz [12]. The inequality for arbitrary n can then be derived easily. In ref. 11 no necessary conditions for the equality sign to hold have been given.

Eqs. (52), (53) and (54) will be important tools for the calculation of the free energy for systems with separable interactions, as is shown in

chapters III and IV of this thesis.

We mention a number of other applications of the inequality (52).

a) On many occasions it may be sufficient to approximate some of the operators A in a given product by their norms $\|A\|$. Such a norm can be arrived at by choosing the θ corresponding to the particular operator A to be ∞, since we have the formula

$$\lim_{\theta\to\infty} [\mathrm{Tr}\,(A^{\dagger}A)^{\frac{1}{2}\theta}]^{\theta^{-1}} = \|A\| . \tag{55}$$

Divide now the operators $A_1,\ldots,A_n$ (which need not be positive definite hermitean) into two sets $A_{i_1}\ldots A_{i_r}$, and $A_{j_1}\ldots A_{j_{n-r}}$. Choosing $\theta_{i_1} = \ldots = \theta_{i_r} = \infty$ and $\theta_{j_1},\ldots,\theta_{j_{n-r}}$ in such a way that $\Sigma_{\sigma=1}^{n-r}\theta_{j_\sigma}^{-1} = 1$, one finds from (52) and (55) the special case

$$|\mathrm{Tr}\,A_1\ldots A_n| \le \|A_{i_1}\|\ldots\|A_{i_r}\| \prod_{\sigma=1}^{n-r}\left[\mathrm{Tr}(A_{j_\sigma}^{\dagger}A_{j_\sigma})^{\frac{1}{2}\theta_\sigma}\right]^{\theta_\sigma^{-1}} \text{*)} . \tag{56}$$

By identifying the operators in an appropriate way one obtains the specialization to finite demensional spaces of the following formula given by Ginibre and Gruber [14].

Let H be self-adjoint on some Hilbert space, with $\mathrm{Tr}\,e^{-\alpha H} < \infty$ for each α. Let $A_1,\ldots,A_n$ be bounded and $\alpha_k \ge 0$ with $\Sigma_k\alpha_k = 1$. Then

$$|\mathrm{Tr}\,(\prod_{k=1}^{n} e^{-\alpha_k H} A_k)| \le \prod_{k=1}^{n}\|A_k\| \cdot \mathrm{Tr}\,e^{-H} . \tag{57}$$

As an example of the use of the "norm-version" of the Hölder inequality we consider the Dyson expansion for a Hamiltonian $\mathcal{H} = \mathcal{H}_0+h$, in which the perturbation is assumed to have a finite operator norm $\|h\| < \infty$. The Dyson expansion can be written

$$e^{-\beta\mathcal{H}} = e^{-\beta\mathcal{H}_0}\sum_{n=0}^{\infty}(-)^n\int_0^{\beta}d\tau_1\int_0^{\tau_1}d\tau_2\ldots\int_0^{\tau_{n-1}}d\tau_n\, h(\tau_1)\ldots h(\tau_n) , \tag{58}$$

where

$$h(\tau) \equiv e^{\tau\mathcal{H}_0}\, h\, e^{-\tau\mathcal{H}_0} . \tag{59}$$

---

*) An estimate involving operator norms can, of course, also be given for the r.h.s. of (1), since for a given k all $\alpha_i^{(k)}$, i = 1, ..., n satisfy $\alpha_i^{(k)} \le \alpha_1^{(k)} = \|A_k\|$, cf. theorem 4 of ref. 13.

Let now the average value of an operator A with respect to the Hamiltonian $\mathcal{H}_0$ be defined by

$$\langle A \rangle_{\mathcal{H}_0} \equiv \frac{\mathrm{Tr}\, A\, e^{-\beta\mathcal{H}_0}}{\mathrm{Tr}\, e^{-\beta\mathcal{H}_0}} \quad . \tag{60}$$

Then

$$\langle e^{\beta\mathcal{H}_0} e^{-\beta(\mathcal{H}_0+h)} \rangle_{\mathcal{H}_0} = \frac{\mathrm{Tr}\, e^{-\beta(\mathcal{H}_0+h)}}{\mathrm{Tr}\, e^{-\beta\mathcal{H}_0}} = 1 + \sum_{n=1}^{\infty} a_n \quad , \tag{61}$$

where

$$a_n = (\mathrm{Tr}\, e^{-\beta\mathcal{H}_0})^{-1} (-)^n \int_0^{\beta} d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{n-1}} d\tau_n \, \cdot$$

$$\cdot \mathrm{Tr} \left\{ e^{-(\beta-\tau_1)\mathcal{H}_0} h\, e^{-(\tau_1-\tau_2)\mathcal{H}_0} h \ldots h\, e^{-\tau_n\mathcal{H}_0} \right\} ; . \tag{62}$$

Applying (57) we have immediately

$$|a_n| \leq \frac{\beta^n}{n!} \|h\|^n \quad , \tag{63}$$

so that the series appearing in the r.h.s. of (61) is convergent. Similar considerations have been applied in the evaluation of the high-temperature expansion of xx-spin correlation functions in the spin-cyclic one-dimensional XY model, cf. ref. 15, app. A.

b) The Hölder inequality for operators can also be useful in connection with convex functions.

Consider an arbitrary product of nonsingular operators. This can always be written in the form

$$\prod_k e^{-\lambda_k H_k} \quad ,$$

where Re $\lambda_k \geq 0$ and H_k is hermitean, $k = 1, \ldots, n$. Therefore we consider the function

$$G(\underline{\lambda}) = -\beta^{-1} \ln \left| \mathrm{Tr} \prod_{k=1}^{n} e^{-\beta\lambda_k H_k} \right| \quad , \tag{64}$$

where $\underline{\lambda}$ is the n-tuple of complex numbers $\lambda_1, \ldots, \lambda_n$, Re $\lambda_k \geq 0$, and H_k is hermitean, $k = 1, \ldots, n$.

Suppose that $\Sigma_{k=1}^{n} \operatorname{Re} \lambda_k = c \quad (0 < c < \infty)$; then one finds using (52) with the substitution $\theta_k = c(\operatorname{Re} \lambda_k)^{-1}$,

$$G(\underline{\lambda}) \geq -\beta^{-1} \sum_{k=1}^{n} c^{-1} \operatorname{Re} \lambda_k \ln \operatorname{Tr} e^{-\beta c H_k} = \sum_{k=1}^{n} c^{-1} \operatorname{Re} \lambda_k \, G(c \underline{1}_k) , \tag{65}$$

where $\underline{1}_k$ is the n-tuple with $\lambda_k = 1$, and $\lambda_i = 0$, $i \neq k$. Eq. (65) may be considered as a kind of convexity property. Consider now the special case $n = 2$. If we restrict ourselves to real values $\lambda_1 = \lambda$ and $\lambda_2 = 1-\lambda$, the function $g(\lambda)$ defined by

$$g(\lambda) \equiv G(\lambda, 1-\lambda) \tag{66}$$

satisfies the inequality

$$g(\lambda) \geq \lambda g(1) + (1-\lambda) g(0) , \qquad \text{for } 0 \leq \lambda \leq 1 . \tag{67}$$

Eq. (67) implies in particular

$$g(1) \leq g(0) + g'(0) . \tag{68}$$

Noting that $g'(0)$ is given by $\langle \mathcal{H}_1 - \mathcal{H}_2 \rangle_{\mathcal{H}_2}$, (68) can be written

$$F[\mathcal{H}_1] \leq F[\mathcal{H}_2] + \langle \mathcal{H}_1 - \mathcal{H}_2 \rangle_{\mathcal{H}_2} , \tag{69}$$

where the free energy F corresponding to an operator $\mathcal{H}$ is defined by

$$F[\mathcal{H}] = -\beta^{-1} \ln \operatorname{Tr} e^{-\beta \mathcal{H}} . \tag{70}$$

Eq. (69) is the so-called Bogoliubov inequality which is well-known in statistical mechanics and which will be used in the next two chapters.

The Bogoliubov inequality in turn implies the so-called Peierls inequality and the convexity of the free energy. This can be shown very easily, following the line of reasoning of ref. 1.

We first consider the Peierls inequality. Let $\{|i\rangle\}$ be an arbitrary orthonormal basis and let $\mathcal{H}$ be an arbitrary hermitean operator. The diagonal part $\mathcal{H}^D$ of $\mathcal{H}$ for the given basis is defined by

$$\langle i|\mathcal{H}^D|j\rangle = \delta_{ij} \langle i|\mathcal{H}|j\rangle . \tag{71}$$

Now obviously $\langle \mathcal{H} - \mathcal{H}^D \rangle_{\mathcal{H}^D} = 0$, so that

$$F[\mathcal{H}^D] \geq F[\mathcal{H}] . \tag{72}$$

This is the well-known Peierls inequality.

In order to show that the free energy is concave we must prove the inequality

$$F[(1-\lambda)\mathcal{H}_0 + \lambda\mathcal{H}_1] \geq (1-\lambda)F[\mathcal{H}_0] + \lambda F[\mathcal{H}_1] . \tag{73}$$

This can be done easily by choosing a basis $\{|i\rangle\}$ such that $(1-\lambda)\mathcal{H}_0 + \lambda\mathcal{H}_1$ is diagonal. Then from (65) and (72)

$$F[(1-\lambda)\mathcal{H}_0 + \lambda\mathcal{H}_1] = -\ln \mathrm{Tr}\, e^{-\beta(1-\lambda)\mathcal{H}_0^D} e^{-\beta\lambda\mathcal{H}_1^D} \geq (1-\lambda)F[\mathcal{H}_0^D] + \lambda F[\mathcal{H}_1^D] \geq$$

$$\geq (1-\lambda)F[\mathcal{H}_0] + \lambda F[\mathcal{H}_1] .$$

c) We now consider an inequality which can be applied to imaginary rotation operators. An imaginary rotation operator over an angle α around an axis defined by the unit vector $\vec{e}$, has the form

$$e^{-\alpha\vec{e}\cdot\vec{J}} ,$$

where $\vec{J}$ is an angular momentum operator. The trace of such an operator is given by

$$\mathrm{Tr}\, e^{-\alpha\vec{e}\cdot\vec{J}} = \frac{\sinh \alpha(J+\frac{1}{2})}{\sinh \frac{\alpha}{2}} , \tag{74}$$

where J is the norm of any of the components of $\vec{J}$.

Let us now consider a number of successive imaginary rotations. Then we have the inequality

$$\left|\mathrm{Tr} \prod_{i=1}^{n} e^{-\alpha_i \vec{e}_i\cdot\vec{J}}\right| \leq \frac{\sinh \alpha(J+\frac{1}{2})}{\sinh \frac{\alpha}{2}} , \tag{75}$$

where $\alpha = \Sigma_{i=1}^{n} \alpha_i$. Eq. (75) can be found from (52) by substituting $U_k = I$ and $\theta_k = \alpha/\alpha_k$. The equality sign holds iff $\vec{e}_i = \vec{e}$, independent of i. The version for $J = \frac{1}{2}$ of eq. (75) has been used by Mühlschlegel [16] in his proof that the calculation by Bardeen, Cooper and Schrieffer [17] of the free energy starting from the so-called reduced Hamiltonian is exact in the thermodynamic limit.

d) In Chapter III use will be made of the following inequality

$$\left|\ln \mathrm{Tr} \prod_{i=1}^{n} e^{-\frac{\beta}{n}(A - \sum_{f=1}^{p} x_{fi}B_f)}\right| \leq \frac{1}{n}\sum_{i=1}^{n} \ln \mathrm{Tr}\, e^{-\beta(A-\sum_{f=1}^{p} x_{fi}B_f)} , \tag{76}$$

where $A_1, B_1, \ldots, B_p$ are hermitean operators and x_{fi} real numbers. The equality sign holds iff

$$\sum_{f=1}^{p} (x_{fi} - x_{fj})B_f = c_{ij} \quad , \tag{77}$$

where c_{ij} is a c-number for all values $i, j = 1, \ldots, n$.

REFERENCES

1) Falk, H., Am.J.Phys. 38 (1970) 858.
2) Neumann, J. von, Tomsk Univ.Rev. 1 (1937) 286; Coll.Works Vol. IV, p. 205. Pergamon Press (New York 1962).
3) Mirsky, L., Math.Nachr. 20 (1959) 171.
4) Ky Fan, Proc. N.A.S. (U.S.A.) 37 (1951) 760.
5) Horn, A., Proc. N.A.S. (U.S.A.) 36 (1950) 374.
6) Polya, G., Proc. N.A.S. (U.S.A.) 36 (1950) 49.
7) Marcus, M. and Moyls, B.N., Canad.J.Math. 9 (1957) 313.
8) Mirsky, L., Proc.Glasgow Math.Assoc. 4 (1959) 1215.
9) Bruijn, N.G. de, Nieuw Archief v. Wiskunde (Amsterdam) 4 (1956) 18.
10) Gantmacher, F., Matrizenrechnung IX.12, VEB Deutscher Verlag der Wissenschaften (Berlin, 1965).
11) Dunford, N. and Schwartz, J., Linear Operators, XI.9.14 and XI.9.20, Wiley Interscience (New York, 1963).
12) Riesz, M., Acta Math. 49 (1927) 465.
13) Visser, C. and Zaanen, A.C., Proc. KNAW (Amsterdam) A55 (1952) 71.
14) Ginibre, J. and Gruber, C., Comm.Math.Phys. 11 (1969) 198.
15) Capel, H.W., van Dongen, E.J. and Siskens, Th.J., Physica 76 (1974) 445.
16) Mühlschlegel, B., J.Math.Phys. 3 (1962) 522.
17) Bardeen, J., Cooper, L.N. and Schrieffer, J.R., Phys.Rev. 108 (1957) 1175.

III. AN EXACT CALCULATION OF THE FREE ENERGY IN SYSTEMS WITH SEPARABLE INTERACTIONS

Synopsis

By deriving an upper and a lower bound we give a rigorous calculation of the free energy, in the thermodynamic limit, for a general class of model systems, characterized by a hamiltonian that contains a one-particle part and separable two-particle operators. The result is an expression for the free energy which is of the molecular-field type. The upper bound is obtained by a variational type of argument. Using the Trotter product formula and a well-known integral representation the partition function can be expressed as a multidimensional integral of a function e^{-NG}. In the derivation of the lower bound we have employed Laplace's method. The absolute minimum of the function G can be obtained using Hölder's inequality for operators.

In addition the second derivatives of G at the minimum are investigated in detail.

1. *Introduction.* It is well known in the literature that molecular-field type of approximations can give rigorous results in the thermodynamic limit if the range of the interactions in the model under consideration tends to infinity. The simplest example is the Ising model with equivalent-neighbour interactions, *i.e.*, a model in which all the spins have equal interactions, independent of the positions.

This model has been treated by several authors, *e.g.* Mühlschlegel and Zittartz[1]) and Niemeyer[2]).

A lattice-gas model which is equivalent to this particular Ising model was proposed by Husimi and Temperley[3]), and has also been studied by Katsura[4]).

For results on Ising models (with $S = \frac{1}{2}$) with more general interactions that are also essentially of the long-range type, the work of Kac[5]), Baker[6]), and Siegert and Vezzetti[7]) may be mentioned.

The one-dimensional Van der Waals gas with long-range interactions has been considered by Kac, Uhlenbeck and Hemmer[8]), by Van Kampen[9]), and also by Lebowitz and Penrose[10]).

Another way to obtain the molecular-field approximation is to use an expansion of the partition function in inverse powers of the number of nearest neighbours z; *cf.*, the expansions by Brout[11]) and Horwitz and Callen[12]). In ref. 6 a similar

expansion is employed. The leading term in these expansions corresponds to the molecular-field approximation. In addition Fisher and Gaunt[13]) have shown by numerical extrapolation on nearest-neighbour Ising models on d-dimensional hypercubical lattices, that if d, the number of dimensions, tends to infinity, important molecular-field results are obtained, *e.g.* the nature of the divergence of the susceptibility.

Several rigorously soluble models investigated by Hallers and Vertogen[14]) as special cases of the Hubbard hamiltonian[15]) in the theory of metal–insulator transitions, are equivalent to Ising models with equivalent-neighbour interactions.

The case of $S = 1$ ions with zero-field splitting and equivalent-neighbour interactions, where some interesting first-order phase transitions can occur (*cf.* refs. 16, 17) has been treated by Blume and Watson[18]).

So far we have only mentioned models in which all operators occurring in the hamiltonian commute with one another. The case of noncommuting operators is in general much more difficult to handle.

However, the isotropic Heisenberg model with equivalent-neighbour interactions presents not more difficulties than the corresponding Ising model. It has been considered by Niemeyer[2]), while also the computer calculations by Kittel and Shore[19]) may be mentioned.

In relation with the general noncommuting case it is important to refer to a paper by Mühlschlegel[20]) who proved, starting from the so-called reduced hamiltonian in the BCS theory of superconductivity[21]), that the calculation of the free energy by BCS is exact in the thermodynamic limit. (Quite a different and more complicated proof of this had been given previously by Bogoliubov, Zubarev and Tserkovnikov[22]). In addition various models with equivalent-neighbour interactions have been studied by a C*-algebra type of approach using the Bogoliubov–Haag procedure (*cf.*, *e.g.* refs. 23–26).

In the present paper we give a calculation of the free energy in the thermodynamic limit for a fairly general class of hamiltonians, *viz.* those which contain a one-particle part and separable two-particle operators (the interaction being essentially of the equivalent-neighbour type).

The method of calculation is essentially a generalization of Mühlschlegel's treatment of the BCS theory; there are, however, some differences. While Mühlschlegel follows Feynman[27]) in introducing dummy ordering parameters in order to deal with the problem of noncommuting operators occurring in the exponents, we prefer to use the so-called Trotter product formula[28,29]) instead, for reasons to be explained in section 4.

Although we also start to bring the partition function into a form that is suitable for a saddle-point integration, where the large parameter is provided by the number of particles (or the volume) of the system, we will use this saddle-point integration only to derive a lower bound for the free energy. It will be seen that an upper bound can be found rather easily by a variational type of argument on the basis

of a general thermodynamic inequality of Bogoliubov[30]). This derivation of the upper bound for the free energy establishes a direct connection with a molecular-field treatment of the hamiltonian. Since there are in fact no restrictions on the separable two-particle operators that may occur in the hamiltonian, a generalization of Hölder's inequality to operators[31]) is needed to get the appropriate information on the absolute maximum of the integrand in the expression for the partition function.

The paper is arranged as follows. In section 2 we define the class of hamiltonians that we are going to investigate. Also some specific examples will be given. In section 3 we deal with the case of commuting operators in order to introduce some notations and to describe the method that will be used in handling the general problem of noncommuting operators. In section 4 an integral representation for the partition function Z is derived using Trotter's product formula and a well-known integral trick. In section 5 we investigate the absolute maximum of the integrand in the expression for Z which plays a crucial role in the saddle-point integration. In section 6 an upper bound on the free energy is derived using a variational type of argument. The derivation of the lower bound by means of a saddle-point integration can be found in section 7. Much attention is paid to the determinant of the matrix of second derivatives. It turns out that in general the influence of this determinant is negligible in the thermodynamic limit. However, there are some exceptions and in section 8 one finds a discussion on the relation between such a situation and the occurrence of a phase transition.

2. *The hamiltonian.* In this paper we are going to consider a set of model systems which we define by their hamiltonians. Let H_k $(k = 1, \ldots, N)$ be finite-dimensional Hilbert spaces. Let $T(k)$ be a hermitean operator defined on H_k for each k. Let $V_f(k, l)$ be an operator defined on the direct-product spaces $H_k \otimes H_l$ for arbitrary k and l. We can define the corresponding operators in the direct-product space $H = \otimes H_k$ by taking direct products of $T(k)$ and $V_f(k, l)$ with unit operators in the remaining spaces. The resulting operators will again be denoted by $T(k)$ and $V_f(k, l)$. Obviously operators with different indices k (acting on different "particles") will commute.

The hamiltonian that we consider has the following form

$$\mathscr{H} = \sum_{k=1}^{N} T(k) - (1/2N) \sum_{f=1}^{p} \sum_{k,l=1}^{N} V_f(k, l), \tag{1}$$

where we assume that the operators $V_f(k, l)$ are separable, *i.e.* they can be written as a product of two hermitean operators

$$V_f(k, l) = V_f(k)\, V_f(l). \tag{2}$$

The hamiltonian is defined on H. It can be written as

$$\mathscr{H} = \sum_{k=1}^{N} T(k) - (1/2N) \sum_{f=1}^{p} \left(\sum_{k=1}^{N} V_f(k) \right)^2 . \tag{3}$$

Each Hilbert space H_k can be interpreted to be the space of states of a particle k. In this terminology the product space H is the space of N-particle states, $T(k)$ is a one-particle operator and $V_f(k, l)$ is a separable interaction between k and l. Here it is assumed that the unit operator $1(k)$ in the space H_k and the operators $V_f(k)$, $f = 1, \ldots, p$, are linearly independent. This is not a serious restriction since, if there is a linear relation between these operators, a suitable linear transformation will lead again to a hamiltonian of form (3), however, with $p - 1$ instead of p quadratic interaction terms.

The partition function of the system is given by

$$Z_N = \mathrm{Tr}\, \mathrm{e}^{-\beta \mathscr{H}}, \tag{4}$$

where the trace is to be taken over a basis in the product space H. The free energy per particle is given by

$$f_N = -(1/N)\, k_B T \ln Z_N . \tag{5}$$

In the present paper we shall give a rigorous calculation of the free energy per particle in the thermodynamic limit:

$$f = \lim_{N \to \infty} f_N . \tag{6}$$

It will turn out that f is equal to the free energy per particle f_{mol} which would be obtained after applying the molecular-field approximation to the hamiltonian $\mathscr{H}$. In addition grand-canonical ensembles can be treated as we shall see shortly.

(i) A simple example of a hamiltonian like (1) is a system consisting of N spins S with equivalent-neighbour interactions, independent of the spins k and l. This interaction can be an anisotropic ferromagnetic Heisenberg interaction [for antiferromagnetic interactions we do not get a hamiltonian of form (3) with hermitean operators $V_f(k)$]

$$-(1/2N) \sum_{k,l} (J_1 S_k^x S_l^x + J_2 S_k^y S_l^y + J_3 S_k^z S_l^z),$$

or a biquadratic exchange interaction[32])

$$-(K/2N) \sum_{k,l} (S_k^z)^2 (S_l^z)^2 ,$$

or even a more complicated expression. The one-particle operator $T(k)$ can be any spin operator, for instance a Zeeman interaction with a magnetic field which

may depend on the site k, $\sum_k H_k S_k^z$ or a zero-field splitting term $\sum_k D_k (S_k^z)^2$, with coefficients that may depend on k as well. In this case Z_N is the usual canonical partition function.

(ii) Other examples can be found among fermion systems. Let the possible one-fermion states be divided into groups k, where each group k contains a finite number of one-fermion states $a_{kg}^\dagger |0_k\rangle$, $g = 1, \ldots, \mathcal{N}_k$, where $|0_k\rangle$ is the vacuum state of the Hilbert space H_k. A basis for the $2^{\mathcal{N}_k}$ dimensional space H_k is given by

$$\prod_g (a_{kg}^\dagger)^{n_{kg}} |0_k\rangle, \tag{7}$$

where all numbers n_{kg} can have the values 0 or 1. This is consistent with the terminology as described just after eq. (3), if we consider the states (7) as the one-particle states of a particle k.

The set of operators $\mathscr{H}$ satisfying the condition stated in the discussion following (1), contains as hamiltonians

$$\mathscr{H} = \sum_k T(k) - (1/2N) \sum_f \sum_{k,l} V_f(k)\, V_f(l),$$

where $T(k)$ and $V_f(k)$ are assumed to be polynomials in the operators a_{kg} and $a_{kg}^\dagger$ such that each term contains an even number of creation and annihilation operators, as is usually the case. This condition is sufficient to ensure that operators $T(k)$ and $V_f(k)$ acting on different particles commute. As an example we consider

$$\mathscr{H} = \sum_k \sum_g (\varepsilon_k - \mu)\, a_{kg}^\dagger a_{kg} - (1/2N) \sum_f \Big(\sum_k V_f(k)\Big)^2.$$

Here each group k has the same number of one-electron states and the one-electron energy ε_k within each group k is taken to be constant; μ is the chemical potential. V_f is assumed to be nonzero and independent of k, if ε_k belongs to an interval

$$|\varepsilon_k - \mu| < \hbar\omega. \tag{8}$$

For other values of k, V_f is assumed to be zero. This implies that the interaction between k and l is independent of k and l and nonzero if $|\varepsilon_k - \mu| < \hbar\omega$, and $|\varepsilon_l - \mu| < \hbar\omega$. In this case $\mathrm{Tr}\, e^{-\beta\mathscr{H}}$ is the grand-canonical partition function and the thermodynamic limit is obtained by letting the volume Ω (which is proportional to the number N of groups k satisfying eq. (8)) tend to ∞.

A special example is the reduced hamiltonian in the BCS theory.

$$\mathscr{H} = \sum_k (\varepsilon_k - \mu)\, (a_{k\uparrow}^\dagger a_{k\uparrow} + a_{-k\downarrow}^\dagger a_{-k\downarrow})$$
$$-(V/4N) \Big(\sum_{|\varepsilon_k - \mu| < \hbar\omega} (a_{k\uparrow}^\dagger a_{-k\downarrow}^\dagger + a_{-k\downarrow} a_{k\uparrow}) \Big)^2$$

$$-(V/4N)\Big(\sum_{|\varepsilon_k-\mu|<\hbar\omega} \mathrm{i}\,(a^\dagger_{k\uparrow}a^\dagger_{-k\downarrow} - a_{-k\downarrow}a_{k\uparrow})\Big)^2$$
$$= \sum_k (\varepsilon_k - \mu)\,(a^\dagger_{k\uparrow}a_{k\uparrow} + a^\dagger_{-k\downarrow}a_{-k\downarrow}) - (V/N) \sum_{\substack{\mu-\hbar\omega<\varepsilon_k,\\ \varepsilon_l<\mu+\hbar\omega}} a^\dagger_{k\uparrow}a^\dagger_{-k\downarrow}a_{-l\downarrow}a_{l\uparrow}.$$

Here $\mathcal{N}_k = 2$ for each k. The Hilbert space H_k is constructed from the basis states

$$|0_k\rangle, \quad a^\dagger_{k\uparrow}|0_k\rangle, \quad a^\dagger_{-k\downarrow}|0_k\rangle \quad \text{and} \quad a^\dagger_{k\uparrow}a^\dagger_{-k\downarrow}|0_k\rangle.$$

After these examples we shall give in the following sections an explicit calculation of the free energy per particle for the general class of systems defined by the hamiltonian (3). Here it may be important to note that nothing has been assumed on the commutation properties of the operators $T(k)$ and $V_f(k)$. In the special case that these operators commute, the calculation is rather easy. This case will be treated in section 3. The more complicated situation of noncommuting operators constitutes the main part of the present paper.

3. *Commuting operators.* In this section we calculate the free energy per particle under the assumption that all the operators $T(k)$ and $V_f(k)$ commute. In that case the result is simple. Applying a well-known integral representation (*cf.*, *e.g.*, ref. 33)

$$\mathrm{e}^{A^2} = \pi^{-\frac{1}{2}} \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2}\, \mathrm{e}^{2xA}, \tag{9}$$

to any of the operators $A = \sum_k V_f(k)$ $(f = 1, \ldots, p)$, and after changing to integration variables $\xi_f = (\frac{1}{2}\beta N)^{-\frac{1}{2}} x_f$, the operator $\mathrm{e}^{-\beta\mathscr{H}}$ can be written as

$$\mathrm{e}^{-\beta\mathscr{H}} = (\beta N/2\pi)^{\frac{1}{2}p} \int_{-\infty}^{\infty} \mathrm{d}\xi \exp(-\tfrac{1}{2}\beta N\, \xi\cdot\xi) \exp\Big(-\beta \sum_k \mathscr{H}(k,\xi)\Big), \tag{10}$$

where

$$\mathscr{H}(k,\xi) = T(k) - \xi\cdot V(k). \tag{11}$$

Here ξ and V are p-dimensional vectors with components $\xi_1, \ldots, \xi_p$ and $V_1, \ldots, V_p$, respectively, and

$$\xi\cdot V = \xi_1 V_1 + \cdots + \xi_p V_p, \tag{12}$$

is the scalar product; $\mathrm{d}\xi$ denotes the p-dimensional volume element $\mathrm{d}\xi_1 \cdots \mathrm{d}\xi_p$. From the fact that operators acting on different particles commute, *i.e.*, $[\mathscr{H}(k,\xi), \mathscr{H}(l,\xi)] = 0$ if $k \neq l$ it follows that

$$\mathrm{Tr}\exp\Big(-\beta\sum_k \mathscr{H}(k,\xi)\Big) = \prod_k \mathrm{tr}_k\,\{\exp[-\beta\mathscr{H}(k,\xi)]\}, \tag{13}$$

where Tr is the trace over the product space and tr_k the trace over the one-particle space H_k. Using (13) we can write for the partition function

$$Z = (\beta N/2\pi)^{\frac{1}{2}p} \int_{-\infty}^{\infty} \mathrm{d}\xi \, \mathrm{e}^{-NG(\xi)}, \tag{14}$$

where

$$G(\xi) = \tfrac{1}{2}\beta\, \xi \cdot \xi - (1/N) \sum_k \ln \mathrm{tr}_k \exp\left[-\beta \mathscr{H}(k, \xi)\right]. \tag{15}$$

The integral can now be treated by Laplace's method. Let $G(\xi)$ have an absolute minimum for $\xi = \xi^0$, then

$$Z = (\beta N/2\pi)^{\frac{1}{2}p} \int_{-\infty}^{\infty} \mathrm{d}\xi \, \mathrm{e}^{-NG(\xi^0)} \exp\left(-\tfrac{1}{2}N(\xi - \xi^0)(\xi - \xi^0) : \frac{\partial^2 G}{\partial\xi\, \partial\xi}\right). \tag{16}$$

Noting that the second derivatives $\partial^2 G/\partial\xi\, \partial\xi$ are finite, and transforming back to the integration variables $x_f = (\frac{1}{2}\beta N)^{\frac{1}{2}}\, \xi_f$ it can be shown that

$$Z = \mathrm{e}^{-NG(\xi^0)}\, C, \tag{17}$$

where C is a finite constant independent of N. In the thermodynamic limit the free energy per particle is then given by

$$f = \beta^{-1} G(\xi^0), \tag{18}$$

and ξ^0 is the value of ξ such that $G(\xi)$ has an absolute minimum.

4. *An integral representation.* We now consider the situation of noncommuting operators. This case is certainly more difficult to handle, although the final result will be the same. In the first place we have in the expression for $\mathrm{e}^{-\beta\mathscr{H}}$, an exponential operator containing in the exponent a sum of squares of one-particle operators. In order to be able to apply the integral trick, the exponent of the sum should be replaced by a product of exponential operators.

In the case of noncommuting operators this can be done using a straightforward generalization of the Lie–Trotter formula

$$\exp(A_1 + A_2 + \cdots + A_p) = \lim_{n\to\infty} \left[\exp\left(\frac{A_1}{n}\right) \exp\left(\frac{A_2}{n}\right) \cdots \exp\left(\frac{A_p}{n}\right)\right]^n, \tag{19}$$

cf. for instance refs. 28, 29 for the special case of two operators*. Here for the moment the number of particles, N, is assumed to be finite so that all operators $A_1, \ldots, A_p$ act on a finite-dimensional Hilbert space H. The thermodynamic limit will be taken at the end of the calculation (after the limit $n \to \infty$).

An alternative way of dealing with the case of noncommuting operators is to introduce ordering parameters as was done first by Feynman[27]). Then one treats all operators as if they commute, keeping track of the correct order by using an ordering operator. Finally the so-called "disentangling process" removes the parameters again. This method has been used by Mühlschlegel and also by other authors in different problems (*e.g.*, Schrieffer[35]).

In the present paper we prefer to use Trotter's formula, one of the reasons being that one sees clearly how certain commutators between operators lead in the free energy to a term which remains finite for large values of the number of particles. This term can only be neglected in the thermodynamic limit. A more detailed discussion will be given later in this section.

Using eq. (19) the operator $e^{-\beta\mathscr{H}}$ can be written

$$e^{-\beta\mathscr{H}} = \lim_{n\to\infty} \left\{ \exp\left[-\frac{\beta}{n} \sum_k T(k) \right] \prod_f \exp\left[+ \frac{\beta}{2nN} \left(\sum_k V_f(k) \right)^2 \right] \right\}^n .$$

Applying the integral trick (9) this can be changed into

$$e^{-\beta\mathscr{H}} = \lim_{n\to\infty} \left(\frac{\beta N}{2\pi n} \right)^{\frac{1}{2}pn} \left[\int_{-\infty}^{\infty} d\xi \exp\left(-\frac{\beta N}{2n} \xi \cdot \xi \right) \right.$$

$$\left. \times \exp\left(-\frac{\beta}{n} \sum_k T(k) \right) \prod_f \exp\left(\frac{\beta}{n} \xi_f \sum_k V_f(k) \right) \right]^n . \tag{20}$$

In order to handle the integral it is convenient to replace the exponential factors

$$\exp\left(-\frac{\beta}{n} \sum_k T(k) \right) \prod_f \exp\left(\frac{\beta}{n} \xi_f \sum_k V_f(k) \right),$$

by one exponential factor

$$\exp\left[-\frac{\beta}{n} \sum_k \left(T(k) - \sum_f \xi_f V_f(k) \right) \right]. \tag{21}$$

* Dr. J. Vlieger kindly pointed out to us that the Trotter formula had already been applied in 1955, *cf.* ref. 34.

At first sight the validity of this procedure seems to raise no doubts since the commutators between the operators $(\beta/n)\sum_k T(k)$ and $(\beta/n)\sum_k V_f(k)$ are of order β^2/n^2 and therefore seem to be negligible in the limit $n \to \infty$. In appendix A it will be shown that the contribution of such commutators to the integral amounts to a constant factor which can be neglected in the calculation of the free energy per particle in the thermodynamic limit. So in this limit we can make the replacement (21). In the remaining formulae for $e^{-\beta\mathscr{H}}$ and the partition function finite values of n and N occur. In the notation it will henceforth be implied that two limits have to be taken: first the limit $n \to \infty$, *cf.* eq. (20), and afterwards the thermodynamic limit. (This is necessary since in view of the replacement (21) the results would be incorrect for finite N.)

As a result we write

$$e^{-\beta\mathscr{H}} = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn} \left[\int_{-\infty}^{\infty} d\xi \exp\left(-\frac{\beta N}{2n}\xi\cdot\xi\right) \exp\left(-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi)\right)\right]^n, \tag{22}$$

where $\mathscr{H}(k,\xi) = T(k) - \xi\cdot V(k)$ as in (11), and where the vector notation, *cf.* (11) and (12), has been used.

The partition function can now be expressed in terms of a pn-fold integral

$$Z = \mathrm{Tr}\, e^{-\beta\mathscr{H}} = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn} \int_{-\infty}^{\infty}\cdots\int \prod_{i=1}^{n} d\xi_i$$
$$\times \exp\left(-\frac{\beta N}{2n}\sum_{i=1}^{n}\xi_i\cdot\xi_i\right) \mathrm{Tr}\prod_{i=1}^{n}\prod_{k=1}^{N}\exp\left(-\frac{\beta}{n}\mathscr{H}(k,\xi_i)\right). \tag{23}$$

Here we used that operators acting on different particles commute, so

$$[\mathscr{H}(k,\xi), \mathscr{H}(l,\xi)] = 0 \quad \text{if} \quad k \neq l. \tag{24}$$

In fact, $\mathscr{H}(k,\xi_i)$ and $\mathscr{H}(l,\xi_j)$, for $k \neq l$, commute irrespective of ξ_i and ξ_j. From this one easily checks that, *cf.* eq. (13) in section 3,

$$\mathrm{Tr}\prod_{i=1}^{n}\prod_k \exp\left(-\frac{\beta}{n}\mathscr{H}(k,\xi_i)\right) = \prod_k \mathrm{tr}_k \prod_{i=1}^{n} \exp\left(-\frac{\beta}{n}\mathscr{H}(k,\xi_i)\right). \tag{25}$$

Here Tr is the trace over the product space H and tr_k the trace over the one-particle space H_k.

After inserting (25) into (23) we obtain the following integral representation for the partition function:

$$Z = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn} \int_{-\infty}^{\infty}\cdots\int \prod_i d\xi_i\, e^{-NG(\{\xi_i\})}, \tag{26}$$

where

$$G(\{\xi_i\}) \equiv G(\xi_1, \ldots, \xi_n)$$
$$= \frac{\beta}{2n} \sum_{i=1}^{n} \xi_i^2 - \frac{1}{N} \sum_{k=1}^{N} \ln \operatorname{tr}_k \left(\prod_{i=1}^{n} \exp \left(-\frac{\beta}{n} \mathscr{H}(k, \xi_i) \right) \right). \tag{27}$$

Note that in general $G(\{\xi_i\})$ will not be real though Z, of course, is. So we can write

$$G = G_1 + \mathrm{i} G_2, \tag{28}$$

where

$$G_1 = \frac{\beta}{2n} \sum_i \xi_i^2 - \frac{1}{N} \sum_k \ln \left| \operatorname{tr}_k \prod_i \exp \left(-\frac{\beta}{n} \mathscr{H}(k, \xi_i) \right) \right|$$

and

$$G_2 = -\frac{1}{N} \sum_k \arg \left(\operatorname{tr}_k \prod_i \exp \left(-\frac{\beta}{n} \mathscr{H}(k, \xi_i) \right) \right).$$

In the following sections we consider the integral (26) in more detail. If all operators commute the situation is simple: in section 3 it was shown that the free energy per particle is given by

$$f = \beta^{-1} G(\xi^0),$$

where $G(\xi)$ has been defined by (15), and ξ^0 is the value of ξ such that $G(\xi)$ has an absolute minimum for $\xi = \xi^0$.

In the present case the situation is more complicated. We have an integration over pn variables and the limit $n \to \infty$ has to be taken. In principle this could be done by using the saddle-point method, which includes in particular the calculation of the determinant of a $pn \times pn$ matrix consisting of second derivatives.

In the present paper we proceed by deriving an upper and a lower bound for the free energy per particle, *viz.*

$$f \leq \beta^{-1} G_0, \tag{29a}$$

$$f \geq \beta^{-1} G_0, \tag{29b}$$

where G_0 is the value of the real part G_1 of the function G at the absolute minimum of G_1.

The proof of (29a) is simple using a variational type of argument based on the Bogoliubov inequality. This will be discussed in section 6.

The derivation of the lower bound, *i.e.* eq. (29b), is much harder. We use the inequality

$$\int \cdots \int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-N(G_1 + \mathrm{i}G_2)}$$

$$= \int \cdots \int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-NG_1} \cos NG_2 \leqslant \int \cdots \int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-NG_1}. \tag{30}$$

and apply Laplace's method to the right-hand side. In order to be able to do so we need some properties of the minimum G_0 of the function G_1. These will be derived in section 5.

As a result of Laplace's method we find that the partition function is equal to e^{-NG_0} multiplied essentially by the inverse square root of the determinant of second derivatives of G_1 in the absolute minimum. In section 7 it will be shown that this determinant does not give a contribution to the free energy per particle in the thermodynamic limit.

5. *Investigation of the absolute minimum of* G_1. In this section we shall investigate the minimum of G_1. First it will be shown that in the absolute minimum

$$\xi_i = \xi^0 \qquad \text{(independent of } i\text{)}. \tag{31}$$

Eq. (31) implies that in the minimum we also have

$$G_2 = 0, \tag{32a}$$

$$\left(\frac{\partial G_2}{\partial \xi_i}\right)_{\min} = 0 \tag{32b}$$

(this will be shown at the end of this section).

Eq. (32) ensures that the phase of the integrand is stationary in the absolute minimum of G_1. If this property would not be satisfied, the absolute minimum of G_1 would give no contribution to the integral because of rapid phase variations.

Eqs. (32a) and (32b), however, will not be used explicitly in the proof that $f = \beta^{-1} G_0$, *cf.* (29).

Eq. (31) would be trivial if $G_1(\{\xi_i\})$ would be of the form $G_1 = (1/n) \sum_i g(\xi_i)$. This, however, is not the case; instead we shall prove that there exists a function

$$\psi(\{\xi_i\}) = (1/n) \sum_{i=1}^{n} \phi(\xi_i), \tag{33}$$

satisfying the following properties:

a) $G_1 \geq \psi$, (34a)

b) $G_1 = \psi$ iff* $\xi_i = \xi$, independent of i, for arbitrary ξ. (34b)

* If, and only if.

From (33) we immediately see that there is at least one point $\xi_i = \xi^0$, $i = 1, \ldots, n$, where ψ has an absolute minimum. Eq. (31) then follows quite easily as we shall show after proving (34a) and (34b).

The function ψ is obtained by using the following upper bound for the trace of a matrix product which is essentially a generalization of Hölder's inequality to operators:

Let $A_1, \ldots, A_n$ be arbitrary $m \times m$ positive-definite hermitean matrices, and $\theta_1, \ldots, \theta_n$ arbitrary positive numbers satisfying $\sum_{i=1}^n \theta_i^{-1} = 1$. Then

$$|\mathrm{tr}\, A_1 \cdots A_n| \leq \prod_{i=1}^n (\mathrm{tr}\, A_i^{\theta_i})^{1/\theta_i}. \tag{35}$$

The equality sign holds iff $A_i^{\theta_i} = \lambda_{ij} A_j^{\theta_j}$ for all i, j, where λ_{ij} are arbitrary constants. A simple proof of eqs. (35) and (36) has been given in ref. 31. (36)

Using (35) in the special case that $\theta_i = n$, $i = 1, \ldots, n$, we find

$$\ln \left| \mathrm{tr} \prod_{i=1}^n \exp\left(-\frac{\beta}{n} \mathscr{H}(k, \xi_i)\right) \right| \leq \frac{1}{n} \ln \left(\prod_{i=1}^n \mathrm{tr}_k \exp\left[-\beta \mathscr{H}(k, \xi_i)\right] \right). \tag{37}$$

We now define the function ψ by

$$\psi(\{\xi_i\}) = \frac{\beta}{2n} \sum_{i=1}^n \xi_i^2 - \frac{1}{Nn} \sum_{k=1}^N \sum_{i=1}^n \ln \mathrm{tr}_k \exp\left[-\beta \mathscr{H}(k, \xi_i)\right]. \tag{38}$$

It is easy to show that ψ has the desired properties: a) from (28), (37) and (38) we have $G_1 \geq \psi$; b) (36) shows that $G_1 = \psi$ iff

$$\exp\left[-\beta \mathscr{H}(k, \xi_i)\right] = \lambda_{ij} \exp\left[-\beta \mathscr{H}(k, \xi_j)\right],$$

or

$$\mathscr{H}(k, \xi_i) = \mathscr{H}(k, \xi_j) + c_{ij}(k), \tag{39}$$

where $c_{ij}(k)$ is a c-number. Using (10), eq. (39) reduces to

$$(\xi_i - \xi_j) \cdot V(k) = c_{ij}(k).$$

Since, in view of our discussion in section 2, the operators $\mathbf{1}(k)$ and $V_f(k)$ can be assumed to be linearly independent, we have $\xi_i = \xi$, $i = 1, \ldots, n$, iff the equality sign holds. In order to derive (31), we note that ψ indeed can be written as

$$\psi(\{\xi_i\}) = (1/n) \sum_{i=1}^n \phi(\xi_i),$$

if we define

$$\phi(\xi) = \tfrac{1}{2}\beta\,\xi\cdot\xi - (1/N)\sum_{k=1}^{N} \ln \operatorname{tr}_k \exp\left[-\beta\mathscr{H}(k,\xi)\right]. \tag{40}$$

Here we have to consider two different cases:

1) There is only one point ξ^0 where ϕ assumes its absolute minimum G_0. From (33) it follows that $\psi(\{\xi_i\})$ has an absolute minimum if $\xi_i = \xi^0$ $(i = 1, \ldots, n)$. But then on behalf of (34a) and (34b), also the absolute minimum of $G_1(\{\xi_i\})$ is reached at $\xi_i = \xi^0$ $(i = 1, \ldots, n)$ and this minimum is equal to G_0.

2) If there are more points, say $\xi^1, \ldots, \xi^\nu$, where $\phi(\xi)$ is absolutely minimal, then $\psi(\{\xi_i\})$ reaches its absolute minimum in ν^n points, *viz.* those points where each ξ_i $(i = 1, \ldots, n)$ can be arbitrarily chosen from the set $\xi^1, \ldots, \xi^\nu$. But from (34b) we see that only in ν points [if all ξ_i are equal to the same ξ^r $(r = 1, \ldots, \nu)$], $G_1 = \psi$. So we find that $G_1(\{\xi_i\})$ is absolutely minimal if all ξ_i are equal to the same ξ^r $(r = 1, \ldots, \nu)$. This proves (31).

In the present paper it is assumed that the function $\phi(\xi)$ can only have a finite set of isolated absolute minima. In principle, however, it is possible that $\phi(\xi)$ has a continuous curve of absolute minima in ξ space. Such a situation may be due to a symmetry property of the hamiltonian. In addition there may be difficulties using the Laplace method in order to derive a lower bound on the free energy.

In order to avoid these difficulties one could introduce an extra "anisotropy" term in the hamiltonian, which destroys the symmetry in the hamiltonian and which removes the ambiguity in the ξ^0. (The introduction of such an "anisotropy" term is not unusual in a molecular-field treatment of the Heisenberg hamiltonian in the absence of a magnetic field. The assumption of a preferred direction for the spontaneous magnetization is equivalent to the introduction of an infinitesimal magnetic field.)

Then the free energy in the presence of the "anisotropy" term can be calculated exactly and it is assumed that the free energy in the symmetric case is obtained by considering the limiting case that the anisotropy term tends to zero, after the thermodynamic limit has been taken.

The extrema of the analytic function $\phi(\xi)$ are given by $\partial\phi/\partial\xi = 0$. By using the identity

$$\frac{\partial}{\partial\lambda_j}\,\mathrm{e}^{-\beta\mathscr{H}} = \mathrm{e}^{-\beta\mathscr{H}}\int_0^\beta \mathrm{d}\tau\,\mathrm{e}^{\tau\mathscr{H}}\mathscr{H}_j\,\mathrm{e}^{-\tau\mathscr{H}}, \tag{41}$$

for an arbitrary operator $\mathscr{H}$ of the form $\mathscr{H} = -\sum_j \lambda_j\mathscr{H}_j$, it follows that

$$\frac{\partial\phi}{\partial\xi} = \beta\xi - \frac{1}{N}\sum_k \Bigg[\{\operatorname{tr}_k \exp\left[-\beta\mathscr{H}(k,\xi)\right]\}^{-1} \times \operatorname{tr}_k\left(\mathrm{e}^{-\beta\mathscr{H}(k,\xi)}\int_0^\beta \mathrm{d}\tau\,\mathrm{e}^{\tau\mathscr{H}(k,\xi)}\,V(k)\,\mathrm{e}^{-\tau\mathscr{H}(k,\xi)}\right)\Bigg]. \tag{42}$$

Hence in the extrema we have, using the cyclic invariance of the trace:

$$\xi^0 = (1/N) \sum_k \langle V(k) \rangle_k, \tag{43}$$

where

$$\langle A \rangle_k \equiv \mathrm{tr}_k (A\, \mathrm{e}^{-\beta \mathscr{H}(k, \xi^0)}) / \mathrm{tr}_k\, \mathrm{e}^{-\beta \mathscr{H}(k, \xi^0)}. \tag{44}$$

From the preceding discussion we conclude that $G_1(\{\xi_i\})$ reaches its absolute minimum in the points $\xi_i = \xi^0$ $(i = 1, \ldots, n)$ where ξ^0 is such that $\phi(\xi)$ has an absolute minimum for $\xi = \xi^0$.

We finally show that $G_2 = 0$ and $\partial G_2 / \partial \xi_i = 0$ in the absolute minimum [eqs. (32a) and (32b)]. If $\xi_i = \xi^0$ then

$$G = \tfrac{1}{2}\beta (\xi^0)^2 - (1/N) \sum_k \ln \mathrm{tr}_k \exp [-\beta \mathscr{H}(k, \xi^0)],$$

and this is a real quantity so that $G_2 = 0$ if $\xi_i = \xi^0$. In addition, using again (41) and the cyclic invariance of the trace

$$\left(\frac{\partial G}{\partial \xi_i} \right)_{\min} = \frac{\beta}{n} \xi^0 - \frac{\beta}{nN} \sum_k \langle V(k) \rangle_k.$$

This again is real (it is even zero), so in particular $\partial G_2 / \partial \xi_i = 0$ at the absolute minimum.

6. *An upper bound for the free energy.* We derive an upper bound for the free energy using a variational type of argument based on Bogoliubov's inequality. We write the hamiltonian $\mathscr{H}$ as

$$\mathscr{H} = \mathscr{H}_0(\xi) + \mathscr{H}_1(\xi), \tag{45}$$

where

$$\mathscr{H}_0(\xi) = \sum_k T(k) - \xi \cdot \sum_k V(k) + \tfrac{1}{2} N \xi \cdot \xi, \tag{46}$$

and

$$\mathscr{H}_1(\xi) = -(1/2N) \left(\sum_k V(k) - N\xi \right)^2. \tag{47}$$

Here a set $(\xi_1, \ldots, \xi_p)$ of arbitrary real parameters has been introduced. From (47) it follows that

$$\langle \mathscr{H}_1 \rangle_{\mathscr{H}_0} \leq 0. \tag{48}$$

Here use has been made of the notation

$$\langle A \rangle_B = \mathrm{Tr}\, A\, \mathrm{e}^{-\beta B} / \mathrm{Tr}\, \mathrm{e}^{-\beta B}. \tag{49}$$

Now Bogoliubov's inequality[30]) states that for any two hermitean operators $\mathscr{H}_0$ and $\mathscr{H}_1$ we have

$$F[\mathscr{H}_0 + \mathscr{H}_1] \leq F[\mathscr{H}_0] + \langle \mathscr{H}_1 \rangle_{\mathscr{H}_0}, \tag{50}$$

where

$$F[A] = -\beta^{-1} \ln \operatorname{tr} e^{-\beta A}. \tag{51}$$

From (48) and (50) we see that $F[\mathscr{H}] \leq F[\mathscr{H}_0(\xi)]$ for all ξ, hence

$$F[\mathscr{H}] \leq F[\mathscr{H}_0(\xi^0)], \tag{52}$$

where ξ^0 is such that $F[\mathscr{H}_0(\xi)]$ has an absolute minimum for $\xi = \xi^0$. Using (51), (46) and (11) we find

$$F[\mathscr{H}_0(\xi)] = \tfrac{1}{2}N\xi \cdot \xi - \beta^{-1} \sum_k \ln \operatorname{tr}_k \exp[-\beta \mathscr{H}(k, \xi)] = N\beta^{-1}\phi(\xi), \tag{53}$$

where $\phi(\xi)$ has been defined in (40). Hence from (52) and (53) we have $F[\mathscr{H}] \leq N\beta^{-1}G_0$, where G_0 is the absolute minimum value of $\phi(\xi)$, reached at $\xi = \xi^0$. This implies in particular eq. (43) and it follows that the upper bound $N\beta^{-1}G_0$ is equal to the free energy which can be calculated from the "molecular-field hamiltonian"

$$\mathscr{H}_{\mathrm{mf}} = \sum_k T(k) - \frac{1}{N}\left\langle \sum_k V(k) \right\rangle \cdot \sum_k V(k) + \frac{1}{2N}\left\langle \sum_k V(k) \right\rangle^2, \tag{54}$$

where $\langle \sum_k V(k) \rangle$ can be found from the implicit equation

$$\left\langle \sum_k V(k) \right\rangle = \sum_k \left\{ \operatorname{tr}_k \exp\left[-\frac{\beta}{n}\left(T(k) - \frac{1}{N}\left\langle \sum_k V(k) \right\rangle \cdot V(k) \right) \right] \right\}^{-1}$$
$$\times \operatorname{tr}_k V(k) \exp\left[-\frac{\beta}{n}\left(T(k) - \frac{1}{N}\left\langle \sum_k V(k) \right\rangle \cdot V(k) \right) \right]. \tag{55}$$

The free energy obtained from $\mathscr{H}_{\mathrm{mf}}$ is the free energy that can be obtained from the hamiltonian $\mathscr{H}$ in a variational formulation of the molecular-field approximation, *cf.* ref. 36 for some specific examples.

7. *A lower bound for the free energy.* In order to derive a lower bound for the free energy we apply Laplace's method to the integral

$$\int \cdots \int \prod_l d\xi_l \, e^{-NG_1(\{\xi_l\})} \qquad \text{[see (30)]}.$$

We expand G_1 in the neighbourhood of the absolute minimum of G_1 where $\xi_i = \xi^0$ $(i = 1, \ldots, n)$. Then

$$G_1 = G_0 + \tfrac{1}{2} \sum_{i,j=1}^{n} (\xi_i - \xi^0) \cdot \left(\frac{\partial^2 G_1}{\partial \xi_i \, \partial \xi_j} \right)_{\min} \cdot (\xi_j - \xi^0). \tag{56}$$

Consider

$$\begin{aligned} F &= -\beta^{-1} \ln Z \\ &= N\beta^{-1} G_0 - \beta^{-1} \ln \left(\left(\frac{\beta N}{2\pi n} \right)^{\frac{1}{2}pn} \int_{-\infty}^{\infty} \cdots \int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-N(G - G_0)} \right) \\ &\geqslant N\beta^{-1} G_0 - \beta^{-1} \ln \left(\left(\frac{\beta N}{2\pi n} \right)^{\frac{1}{2}pn} \int_{-\infty}^{\infty} \cdots \int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-N(G_1 - G_0)} \right). \end{aligned} \tag{57}$$

Substituting (56) in (57) we have

$$\begin{aligned} F &\geqslant N\beta^{-1} G_0 - \beta^{-1} \ln \left[\left(\frac{\beta N}{2\pi n} \right)^{\frac{1}{2}pn} \int_{-\infty}^{\infty} \cdots \int \prod_i \mathrm{d}\xi_i \right. \\ &\quad \left. \times \exp \left(-\tfrac{1}{2} N \sum_{ij} (\xi_i - \xi^0) \cdot \left(\frac{\partial^2 G_1}{\partial \xi_i \, \partial \xi_j} \right)_{\min} \cdot (\xi_j - \xi^0) \right) \right] \\ &= N\beta^{-1} G_0 - \beta^{-1} \ln \left[\left(\frac{\beta N}{2\pi n} \right)^{\frac{1}{2}pn} \left(\frac{2\pi}{N} \right)^{\frac{1}{2}pn} \left(\frac{n}{\beta} \right)^{\frac{1}{2}pn} (\det H)^{-\frac{1}{2}} \right] \\ &= N\beta^{-1} G_0 + (1/2\beta) \ln \det H, \end{aligned} \tag{58}$$

where $(\beta/n)\, H$ denotes the $pn \times pn$ matrix of second derivatives of G_1. Use has been made of the relation

$$\int_{-\infty}^{\infty} \cdots \int \mathrm{d}x_1 \cdots \mathrm{d}x_n \exp \left(-\tfrac{1}{2} \sum_{i,j} M_{ij} x_i x_j \right) = (2\pi)^{\frac{1}{2}n} (\det M)^{-\frac{1}{2}}. \tag{59}$$

At this stage it might be useful to point out that a direct calculation of the integral $I \equiv \int \mathrm{e}^{-(G - G_0)}$ could be possible by expanding the complex function G up to the second derivatives at the absolute minimum. However, it should be noted that the validity of eq. (59) in the case that the matrix M is related to the second derivatives of G, requires that the matrix of the second derivatives of the real part, G_1, of G be positive definite, *cf.* ref. 37. In addition it is necessary that in the limit $n \to \infty$ the determinant of H remains finite. Hence in order to justify a direct calculation of I we have to go through all the steps of the calculation of $I_1 \equiv \int \mathrm{e}^{-(G_1 - G_0)}$. The latter calculation provides, as we shall see, the appropriate lower bound to the free energy per particle in the thermodynamic limit. So a

detailed calculation of I can be useful if one is interested in the corrections to the free energy due to a finite value of N. We now proceed with the calculation of I_1.

An explicit calculation of the second derivatives of G_1 (that is the elements of matrix H) is given in appendix B. The result can be written as

$$H = 1_{pn} + (A/n) + (B/n^2), \tag{60}$$

where 1_{pn} is the $pn \times pn$ unit matrix with elements $\delta_{fg}\delta_{ij}$ ($f, g = 1, \ldots, p$ and $i, j = 1, \ldots, n$); A and B are $pn \times pn$ matrices with elements which depend on n and are at most of order unity. The elements A_{figj} of the matrix A [which are of $\mathcal{O}(1)$] are given by

$$\begin{aligned} A_{figj} = &-(\beta/N) \sum_k \{\mathrm{tr} \exp[-\beta \mathscr{H}(k, \xi^0)]\}^{-1} \sum_{p,q} V_{f_{pq}} V_{g_{qp}} \\ &\times \exp[-\tfrac{1}{2}\beta(h_p + h_q)] \cosh[\beta(\tfrac{1}{2} - |j - i|/n)(h_p - h_q)] \\ &+ (\beta/N) \sum_k \langle V_f \rangle \langle V_g \rangle. \end{aligned} \tag{61}$$

The subscripts p* and q indicate the eigenstates of the hamiltonian $\mathscr{H}(k, \xi^0)$, h_p and h_q the corresponding eigenvalues; further $V_{f_{pq}}$ is the matrix element of $V_f(k)$ between the eigenstates p and q. [In eq. (61) the k dependence of V_f and V_g, h_p and h_q is not shown explicitly.] The elements of the matrix B are bounded (in fact for $i \neq j$ these elements are of order $1/n$, while for $i = j$, they are of order unity). The precise expressions are irrelevant in the limit $n \to \infty$ in view of the following lemma given by Lenard[38]):

Let P_m and Q_m be $m \times m$ matrices such that the elements $p_{rs}(m)$ of P_m and $q_{rs}(m)$ of Q_m obey the relations $\lim_{m\to\infty} p_{rs}(m)\, m$ = constant (depending on r and s) and $\lim_{m\to\infty} q_{rs}(m)\, m^{3/2} = 0$, uniformly in the limit $n \to \infty$; then

$$\lim_{m\to\infty} [\det(1_m + P_m + Q_m) - \det(1_m + P_m)] = 0. \tag{62}$$

In our case we have a $pn \times pn$ matrix H [see (60)] and the lemma ensures that

$$\lim_{n\to\infty} \det H = \lim_{n\to\infty} \det(1_{pn} + A/n). \tag{63}$$

The matrix elements A_{figj} have been defined for $1 \leq i, j \leq n$. We note that they have the following properties:

(i) $A_{figj} = A_{gifj}$;

(ii) $A_{figj} = A_{fjgi}$;

* The label p which denotes here the eigenstates of the hamiltonian $\mathscr{H}(k, \xi^0)$ should not be confused with the number p of independent separable interactions in the hamiltonian $\mathscr{H}$.

(iii) A_{figj} depends only on $j - i$. (64)

So $A_{figj} = A_{fg,j-i}$,

where $A_{fg,i}$ is defined for $-n + 1 \leq i \leq n - 1$;

(iv) $A_{fg,n-i} = A_{fg,i} = A_{fg,-i} = A_{fg,-n+i}$ for $1 \leq i \leq n - 1$;

where in the second step use has been made of (ii). These properties can be verified directly from (61); they also follow from the corresponding properties of the matrix elements $T^{(1)}_{figj}$ in appendix B. Properties (iii) and (iv) show that $\boldsymbol{A}$ is a symmetric cyclic matrix with respect to the indices i and j. They imply that the matrix $\boldsymbol{A}$ which is a $pn \times pn$ matrix, can be reduced by a unitary transformation to n blocks of $p \times p$ matrices $\boldsymbol{D}(\varkappa)$ where $\varkappa$ runs from 0 to $n - 1$. This reduction proceeds in complete analogy to the derivation of the dynamical matrix in the theory of lattice vibrations.

The eigenvalues of the matrix $-\boldsymbol{A}/n$ are the eigenvalues $\lambda_f(\varkappa)$ for $\varkappa = 0, 1, \ldots, n - 1$ and $f = 1, \ldots, p$, of the matrices $\boldsymbol{D}(\varkappa)$ with elements

$$D_{fg}(\varkappa) = -(1/n) \sum_{j=0}^{n-1} A_{fg,j} \exp(2\pi i j\varkappa/n). \tag{65}$$

(This will be derived in more detail in appendix C.)

The determinant is now given by

$$\det(\boldsymbol{1}_{pn} + \boldsymbol{A}/n) = \prod_{\varkappa=0}^{n-1} \prod_{f} [1 - \lambda_f(\varkappa)] = \prod_{\varkappa=0}^{n-1} \det[\boldsymbol{1}_p - \boldsymbol{D}(\varkappa)] \tag{66}$$

and for the free energy we have the inequality

$$F \geqslant N\beta^{-1}G_0 + (1/2\beta) \lim_{n\to\infty} \ln \det H$$

$$= N\beta^{-1}G_0 + (1/2\beta) \lim_{n\to\infty} \sum_{\varkappa=0}^{n-1} \ln \det[\boldsymbol{1}_p - \boldsymbol{D}(\varkappa)]. \tag{67}$$

Substituting (61) into (65), we find

$$D_{fg}(\varkappa, n) = -(\beta/N) \sum_k \langle V_f(k)\rangle \langle V_g(k)\rangle \delta_{\varkappa,0}$$

$$+ (1/N) \sum_k \Big((\mathrm{tr}_k\, \mathrm{e}^{-\beta\mathscr{H}(k,\xi^0)})^{-1} \sum_{p,q} V_{fpq} V_{gqp}\, \mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)}$$

$$\times (\beta/n) \sum_{j=0}^{n-1} \mathrm{e}^{2\pi i j\varkappa/n} \cosh[\beta(\tfrac{1}{2} - j/n)(h_p - h_q)] \Big). \tag{68}$$

Performing the summation over j, one finds

$$\frac{\beta}{n}\sum_{j=0}^{n-1} e^{2\pi i j\varkappa/n}\cosh\left[\beta\left(\tfrac{1}{2}-j/n\right)(h_p-h_q)\right]$$

$$=\frac{\beta}{2n}\,\frac{\sinh\left[\tfrac{1}{2}\beta\,(h_p-h_q)\right]\sinh\left[(\beta/n)\,(h_p-h_q)\right]}{\sin^2(\pi\varkappa/n)+\sinh^2(\beta/2n)\,(h_p-h_q)}, \tag{69}$$

which reduces to

$$2\beta^2\,\frac{\sinh\left[\tfrac{1}{2}\beta\,(h_p-h_q)\right](h_p-h_q)}{4n^2\sin^2(\pi\varkappa/n)+\beta^2\,(h_p-h_q)^2}, \tag{70}$$

in the limit $n\to\infty$. Since $D(\varkappa)=D\,(n-\varkappa)$, the terms with $\varkappa$ and $n-\varkappa$ will give equal contributions to the r.h.s. of (67). Hence in studying the convergence of the sum we do not have to take into account terms with $\varkappa>\frac{1}{2}n$. Using the inequality

$$\sin x/x \geqslant 2/\pi \qquad \text{on} \qquad [0,\tfrac{1}{2}\pi], \tag{71}$$

we see on substituting (70) into (68), that the terms with $\varkappa\leq\frac{1}{2}n$ of $D\,(\varkappa,n)$ are of order $1/\varkappa^2$. This implies in particular that for large $\varkappa$ ($\leq\frac{1}{2}n$ where $n\to\infty$) we have

$$\det\left[\mathbf{1}_p-D\,(\varkappa,n)\right]=1+\mathcal{O}\,(1/\varkappa^2). \tag{72}$$

If now $D(\varkappa)$ is such that all eigenvalues are smaller than 1, the second term in the r.h.s. of (67) reduces to a value independent of N since a series of the form $\sum_\varkappa \ln(1+1/\varkappa^2)$ is convergent. In the thermodynamic limit this leads to the following lower bound for the free energy per particle.

$$f\geq\beta^{-1}G_0. \tag{73}$$

Combining this result with the upper bound derived in section 6, we find that the free energy per particle is given by

$$f=\beta^{-1}G_0. \tag{74}$$

However, for small $\varkappa$, especially $\varkappa=0$, one or more of the eigenvalues of $D(\varkappa)$ may become 1 (depending on the parameters of the problem, such as temperature, interaction strength). The divergence in (67) that results from it, is closely related to the onset of a phase transition. This relation with phase transitions will be discussed in somewhat greater detail in the next section.

Remark. As we have seen, the corrections to the free energy due to the second term on the r.h.s. of (67) are essentially independent of N. They are, however, in general not the only corrections of this magnitude. Referring to the discussion following eq. (21) we note that nonzero commutators between the different operators $V_f(k), f = 1, \ldots, p$, lead to contributions of the same order of magnitude. On the other hand, if there is only one operator $V(k)$ (so $p = 1$), the corrections to the free energy that are finite if N tends to infinity, can indeed be calculated by evaluating the second term of the r.h.s. of (67). Note that in this case all derivatives $\partial^2 G/\partial\xi_i\,\partial\xi_j$ are real, which can be seen from eqs. (B.3) and (B.5) by omitting the real-part symbol Re, so that automatically the contribution of the second derivatives of G at the absolute minimum is equal to the contribution of the second derivatives of G_1.

8. Discussion. We investigate in this section somewhat more closely the relationship between the solutions of the molecular-field equations (43) and transitions between two of them on one hand, and the properties of the eigenvalues of the matrix formed by the second derivatives of G_1 on the other hand.

We assume that for each β there is a finite number of solutions of the molecular-field equations

$$\xi = (1/N) \sum_k \langle V(k) \rangle. \tag{75}$$

These solutions are denoted by $\xi^{(r)}(\beta)$, $r = 1, 2, \ldots$ For each of them we define the function

$$F^{(r)}(\beta) \equiv N\beta^{-1} G_1 (\xi^{(r)}(\beta), \beta). \tag{76}$$

In this notation it is understood that $\xi_i = \xi^{(r)}$, independent of i. In the neighbourhood of each solution $\xi^{(r)}$ we expand G_1.

$$G_1 = G_0^{(r)} + (\beta/2n) \sum_{fgij} x_{fi}^{(r)} x_{gj}^{(r)} (\delta_{fg}\delta_{ij} + A_{figj}^{(r)}/n), \tag{77}$$

where $x_{fi}^{(r)} = \xi_{fi} - \xi_f^{(r)}$ [*cf.* eqs. (60) and (61)]. After diagonalizing $\boldsymbol{A}^{(r)}$, we can write this as

$$G_1 = G_0^{(r)} + (\beta/2n) \sum_{f\varkappa} [\eta_f^{(r)}(\varkappa)]^2 \, [1 - \lambda_f^{(r)}(\varkappa)], \tag{78}$$

where $\lambda_f^{(r)}(\varkappa)$ are the eigenvalues of $-\boldsymbol{A}^{(r)}/n$ (*cf.* appendix C), and the $\eta_f^{(r)}(\varkappa)$ are suitable linear combinations of the components of $x^{(r)}$, which can be considered as normal coordinates.

The eigenvalues $\lambda_f(\varkappa, \xi)$ can of course be defined for arbitrary points $\xi_i = \xi$, independent of i. In this notation $\lambda_f(\varkappa, \xi^{(r)}) = \lambda_f^{(r)}(\varkappa)$. We now consider for each r the maximum eigenvalue $\lambda_{\max}^{(r)} \equiv \max_{f,\varkappa} \lambda_f^{(r)}(\varkappa) = (\lambda_{\max}(\xi^{(r)})$.

We have the following possibilities: (i) if $\lambda_{\max}^{(r)} < 1$, G_1 has a minimum at $\xi^{(r)}$; (ii) if $\lambda_{\max}^{(r)} > 1$, G_1 has not a minimum at $\xi^{(r)}$; while (iii) if $\lambda_{\max}^{(r)} = 1$, no definite statement can be made.

The maximum eigenvalue $\lambda_{\max}^{(r)}(\beta)$ will be a continuous function of β and it will be assumed that it has the value 1 only in a discrete set of points. The existence of an absolute minimum of G_1 for each β implies that for each β there exists at least one r such that

$$\lambda_{\max}^{(r)} \leq 1. \tag{79}$$

We can now divide the set of all points β into three sets. (a) Those β for which $\lambda_{\max}^{(r)}(\beta) = 1$ for at least one value of r; these points form a discrete set B. (b) Those β for which $\lambda_{\max}^{(r)}(\beta) < 1$ for only one r, and $\lambda_{\max}^{(s)}(\beta) \neq 1$ for all s. Then in $\xi^{(r)}$ we have the only and therefore absolute minimum of G_1, so that the free energy is given by $F^{(r)}(\beta)$. From the continuity of $\lambda_{\max}^{(r)}(\beta)$ and $F^{(r)}$ it follows that also in a neighbourhood of such a temperature β the free energy is given by $F^{(r)}$. (c) Those β where $\lambda_{\max}^{(r)}(\beta) < 1$ for several values of r, while at the same time $\lambda_{\max}^{(s)}(\beta) \neq 1$ for all s. Then the free energy is equal to the lowest value of $F^{(r)}(\beta)$, $r = 1, 2, \ldots$.

A phase transition will occur at a point β_c if for $\beta < \beta_c$ and $\beta > \beta_c$ the free energy F will be given by two different branches which we denote by $F^{(1)}$ and $F^{(2)}$, respectively. Of course we have $F^{(1)}(\beta_c) = F^{(2)}(\beta_c)$. Two situations may occur at a transition point

$$\text{i)}\quad \xi^{(1)}(\beta_c) \neq \xi^{(2)}(\beta_c), \tag{80a}$$

$$\text{(ii)}\quad \xi^{(1)}(\beta_c) = \xi^{(2)}(\beta_c). \tag{80b}$$

In order to discuss these two possibilities we note that $\xi^{(r)}(\beta)$ can be expressed as a derivative of the function $F^{(r)}$. For this purpose we replace the one-particle operator $T(k)$ in the original hamiltonian $\mathscr{H}$ by

$$T(k, \varepsilon) = T(k) + \varepsilon \cdot V(k), \tag{81}$$

where the operators $V_f(k)$ have been introduced in (2) and $T(k)$ is independent of the real parameters ε. Then the solutions $\xi^{(r)}$ of the molecular-field equations and the corresponding functions $F^{(r)}$ depend on ε.

Now clearly

$$\frac{1}{N}\left(\frac{\partial F^{(r)}}{\partial \varepsilon}\right)_{\varepsilon=0} = \frac{1}{N}\sum_k \left\langle \frac{\partial T(k, \varepsilon)}{\partial \varepsilon}\right\rangle_{\varepsilon=0}^{(r)} = \frac{1}{N}\sum_k \langle V(k)\rangle_{\varepsilon=0}^{(r)} = \xi_{\varepsilon=0}^{(r)}, \tag{82}$$

where

$$\langle A(k)\rangle_{\varepsilon=0}^{(r)} \equiv \frac{\mathrm{tr}_k\, A_k \exp\left[-\beta\mathscr{H}(k, \xi^{(r)})\right]}{\mathrm{tr}_k \exp\left[-\beta\mathscr{H}(k, \xi^{(r)})\right]}.$$

Hence the parameters $\xi^{(r)}$ can be considered as derivatives of $F^{(r)}$ after introducing an infinitesimal one-particle operator $\varepsilon \cdot V(k)$.

Such a one-particle operator can be interpreted as a coupling of particle k with external fields $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_f$.

A special example is given by the Zeeman term in an external field $\varepsilon V(k) = -HS_k^z$. In this case ξ is the magnetization per particle $\xi = -(1/N)(\partial F/\partial H)$.

We now consider the two cases (80a) and (80b) in more detail.

(i) Since not all the components of ξ are continuous at the transition, at least one of the derivatives of the free energy F with respect to the parameters $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_f$ will have a discontinuity, so that there is a first-order phase transition.

(ii) $\xi^{(1)}(\beta_c) = \xi^{(2)}(\beta_c) = \xi_c$. (83)

Since all the derivatives of F are now continuous we have a 2nd-order or higher-order phase transition. We prove that in such a case $\beta_c \in B$, that is at least one of the maximum eigenvalues $\lambda_{\max}^{(r)}$ is equal to 1. Suppose $\lambda_{\max}^{(1)}(\beta_c) < 1$. Then there exists a number $\delta > 0$ such that $\lambda_{\max}^{(1)}(\beta) < 1$, for $\beta_c \leq \beta < \beta_c + \delta$, *i.e.* $G_1(\xi)$ has a minimum at $\xi = \xi^{(1)}(\beta)$.

For β in this interval F is given by $F^{(2)}$ so there is also a minimum at $\xi^{(2)}(\beta)$. After parametrizing the straight line that connects $\xi^{(1)}(\beta)$ and $\xi^{(2)}(\beta)$ by a parameter t, we see that G_1, considered as a function of t, has a maximum at some point $0 < t_0 < 1$. Then it can be seen that in the point $\zeta = \xi^{(1)}(\beta) + t_0[\xi^{(2)}(\beta) - \xi^{(1)}(\beta)]$ we have: $\lambda_{\max}(\zeta) \geq 1$, where $\lambda_{\max}(\zeta) \equiv \max_{\varkappa,f} \lambda_f(\varkappa, \zeta)$. Since such a ζ can be found for all β with $\beta_c \leq \beta \leq \beta_c + \delta$, we can construct a path in ξ space on which we have $\lambda_{\max}(\zeta) \geq 1$.

In particular after taking the limit $\beta \to \beta_c$, we have

$$\lambda_{\max}(\xi_c) \geq 1.$$

In addition we know that $\lambda_{\max}(\xi_c) \leq 1$. Hence

$$\lambda_{\max}(\xi_c) = 1. \tag{84}$$

As an example of the relation between second-order phase transitions and the occurrence of an eigenvalue 1, we consider the special situation that

$$\mathrm{tr}_k\, V(k)\, \mathrm{e}^{-\beta T(k)} = 0, \qquad \text{for all } \beta. \tag{85}$$

This implies that $\xi = 0$ is always a solution (to be denoted by a superscript 1) of the molecular-field equations (75). We investigate the stability of this solution. First of all we prove that for $\xi = 0$ we have

$$\lambda_{\max}^{(1)} \equiv \max_{\varkappa} \max_{f} \lambda_f^{(1)}(\varkappa) = \max_{f} \lambda_f^{(1)}(0). \tag{86}$$

Proof. From (68) and (85) we have

$$\sum_{fg} D^{(1)}_{fg}(\varkappa)\, z^*_f z_g = \frac{1}{N}\sum_k (\mathrm{tr}_k\, \mathrm{e}^{-\beta T(k)})^{-1} \sum_{pq} \left|\sum_f z^*_f V_{fpq}\right|^2 \Phi(p, q, \varkappa),$$

where z_f are arbitrary complex numbers.

The function $\Phi(p, q, \varkappa)$ is nonnegative, while also $\Phi(p, q, \varkappa) < \Phi(p, q, 0)$. Using the notation $\|z\|^2 = \sum_f |z_f|^2$, we have

$$\max_f \lambda^{(1)}_f(\varkappa) = \max_{\|z\|=1} \sum_{fg} D^{(1)}_{fg}(\varkappa)\, z^*_f z_g = \sum_{fg} D^{(1)}_{fg}(\varkappa)\, \tilde{z}^*_f \tilde{z}_g$$

$$< \sum_{fg} D^{(1)}_{fg}(0)\, \tilde{z}^*_f \tilde{z}_g \leq \max_{\|z\|=1} \sum_{fg} D^{(1)}_{fg}(0)\, z^*_f z_g = \max_f \lambda^{(1)}_f(0),$$

whence (86) follows.

Here $\tilde{z}_f$ are the components of the eigenvector of $D^{(1)}(\varkappa)$ that corresponds to the eigenvalue $\max \lambda^{(1)}_f(\varkappa)$. In the limit $\beta \to 0$, all matrix elements $D_{fg}(\varkappa)$ tend to 0, so that $\lambda^{(1)}_{\max} < 1$, *i.e.*, $\xi^{(1)} = 0$ is a stable solution.

Define now β_c as the highest temperature such that $D^{(1)}(0, \beta)$ has an eigenvalue 1. Then on account of (86) β_c is also the highest temperature for which $\lambda^{(1)}_{\max} = 1$. We assume that $\lambda^{(1)}_{\max}$ is neither degenerate, nor has a maximum at $\beta = \beta_c$; and also that there are no solutions with $\xi \neq 0$ for $\beta < \beta_c$.

Consider on the other hand the solution of the molecular-field equations for small values of ξ different from 0. Expanding the r.h.s. of (75) and writing $\xi = \lambda \boldsymbol{a}$ with $\lambda \to 0$, one finds as a condition for the temperature below which there exists a solution $\xi \neq 0$:

$$\boldsymbol{a} = \boldsymbol{a} \cdot (1/N) \sum_k (\mathrm{tr}_k\, \mathrm{e}^{-\beta T(k)})^{-1}$$

$$\times\, \mathrm{tr}_k \left(V(k)\, \mathrm{e}^{-\beta T(k)} \int_0^\beta \mathrm{d}x\, \mathrm{e}^{xT(k)}\, V(k)\, \mathrm{e}^{-xT(k)} \right) \equiv \boldsymbol{a} \cdot \boldsymbol{A}(\beta), \tag{87}$$

i.e. the matrix $\boldsymbol{A}(\beta)$ should have an eigenvalue 1.

A straightforward evaluation of the integral in $\boldsymbol{A}(\beta)$, using a representation in which $T(k) = \mathscr{H}(k, 0)$ is diagonal, shows that [*cf.* (68)] $\boldsymbol{A}(\beta) = \boldsymbol{D}^{(1)}(0, \beta)$. Now clearly β_c fulfils (87). Hence it follows that at this temperature a second-order transition will take place such that the ξ values at the absolute minimum of the function G_1 change from 0 for $\beta < \beta_c$ to values $\neq 0$ for $\beta > \beta_c$. A very simple and well-known illustration of this behaviour is provided by the Ising model with equivalent-neighbour interactions given by the hamiltonian $\mathscr{H} = -(J/2N)(\sum_k \sigma_k)^2$. The molecular-field equations are $\sigma = \tanh \beta J \sigma$ ($\sigma = J^{-\frac{1}{2}}\xi$); and the eigenvalue $\lambda(\varkappa)$ of the 1×1 matrix $D(\varkappa)$: $\lambda_\sigma(\varkappa) = D(\varkappa) = \beta J \delta_{\varkappa, 0}(1 - \sigma^2)$. One sees that above the critical temperature, given by $\beta_c J = 1$, the solution $\sigma = 0$ will be stable whereas below this temperature $\lambda_0(0) > 1$ so that this solution becomes unstable.

However, below T_c there is a stable solution with $\sigma \neq 0$. The correction term in (67) is given by $(1/2\beta) \ln [1 - \lambda_0(0)]$ for $T > T_c$; and by $(1/2\beta) \ln [1 - \lambda_\sigma(0)]$, $\sigma \neq 0$, for $T < T_c$, which shows that at T_c, where both λ's are 1, this correction term diverges (*cf.* ref. 2).

The point that the expansion we used leads to difficulties at the critical temperature, could have been anticipated: we may refer for instance to the work of Siegert and Vezzetti[7]).

In this paper we have considered the "ferromagnetic" case: separable operators with negative coefficients [*cf.* (31)]. In a following paper[39]), we deal with the "antiferromagnetic" case: separable operators with positive coefficients. Also some more explicit examples will be investigated.

Acknowledgements. The authors are much indebted to Professor P.W. Kasteleyn for many useful discussions and comments on the manuscript.

This investigation is part of the research programme of the "Stichting voor Fundamenteel Onderzoek der Materie (F.O.M.)", which is financially supported by the "Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek Z.W.O.)".

APPENDIX A

Consider eq. (20) for the operator $e^{-\beta\mathscr{H}}$ for finite values of N; in agreement with the discussion on the notation below (21) also n is taken to be large but finite. The operator can be written

$$e^{-\beta\mathscr{H}} = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn} \left[\int_{-\infty}^{\infty} d\xi \exp\left(-\frac{\beta N}{2n}\xi\cdot\xi\right) \exp\left(-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi)\right) A(\xi)\right]^n, \tag{A.1}$$

where

$$\mathscr{H}(k,\xi) = T(k) - \xi\cdot V(k)$$

and

$$A(\xi) = \exp\left(\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi)\right) \exp\left(-\frac{\beta}{n}\sum_k T(k)\right) \prod_f \exp\left(\frac{\beta}{n}\xi_f \sum_k V_f(k)\right). \tag{A.2}$$

In order to get an operator which unlike $A(\xi)$ no longer depends on ξ, we write (A.1) in the following manner:

$$e^{-\beta\mathscr{H}} = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn} \left\{\left[\int d\xi \exp\left(-\frac{\beta N}{2n}\xi\cdot\xi\right) \exp\left(-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi)\right)\right] \bar{A}\right\}^n, \tag{A.3}$$

where we have defined:

$$\tilde{A} = \left[\int d\xi \exp\left(-\frac{\beta N}{2n}\xi\cdot\xi\right)\exp\left(-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi)\right)\right]^{-1}$$

$$\times \int d\xi \exp\left(-\frac{\beta N}{2n}\xi\cdot\xi\right)\exp\left(-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi)\right)A(\xi). \tag{A.4}$$

Remembering its definition in (A.2) we can expand the operator $A(\xi)$ in a power series in $1/n$ and ξ. Obviously the zeroth-order term is 1, while there is no term proportional to $1/n$.

It is also clear from the way in which ξ occurs, that a term in the expansion with r factors ξ, contains at least a factor n^{-r}. So we have

$$A(\xi) = 1 + \sum_{s\geq 2}\sum_{r\leq s} A_{rs}(\xi), \tag{A.5}$$

where A_{rs} is the contribution of the terms that contain s factors $1/n$ and r factors ξ. More explicitly

$$A_{rs}(\xi) = \sum_{f_1,\ldots,f_r}(1/n^s)\,\xi_{f_1}\cdots\xi_{f_r}B^{(s)}_{f_1\cdots f_r},$$

where the B's are polynomials in products of the operators $\sum_k T(k)$ and $\sum_k V_f(k)$.

Inserting the expansion (A.5) into the definition of $\tilde{A}$, we have

$$\tilde{A} = 1 + \sum_{s\geq 2}\sum_{r\leq s}\tilde{A}_{rs}, \tag{A.6}$$

where $\tilde{A}_{rs}$ is the contribution due to A_{rs} in (A.4). It is not difficult to estimate the order of magnitude in $1/n$ of a particular term $\tilde{A}_{rs}$. For that purpose we note that the appropriate integration variables in eq. (A.1) are given by $\eta = (\beta N/2n)^{\frac{1}{2}}\xi$, so that each factor ξ contributes a factor $n^{\frac{1}{2}}$. Hence the order of magnitude of a particular term $\tilde{A}_{rs}$ is given by $n^{-s+r/2}$.

Since we must take the nth power of the expansion in order to get $e^{-\beta\mathscr{H}}$, the only terms of $\tilde{A}_{rs}$ that can give a contribution in the limit $n\to\infty$, are those for which $s - \frac{1}{2}r \leq 1$, which leaves only one possibility, *viz.*, $s = r = 2$, to investigate. So we write:

$$\tilde{A} = 1 + \tilde{A}_{22} + \mathcal{O}(n^{-3/2}), \tag{A.7}$$

where $\mathcal{O}(n^{-3/2})$ does not contribute to $e^{-\beta\mathscr{H}}$ if $n\to\infty$. It may be remarked that this holds whatever the N dependence of the terms in $\mathcal{O}(n^{-3/2})$ may be, since the limit $n\to\infty$ has to be taken first. In general, however, $\tilde{A}_{22}$ may give a non-vanishing contribution but we shall show that it leads to an additional term in the

free energy, which remains finite for large N and can be neglected in the thermodynamic limit. We expand $A(\xi)$, as defined in (A.2), up to order $1/n^2$, defining

$$B_0 = -\beta \sum_k T(k) \quad \text{and} \quad B_f = \beta \xi_f \sum_k V_f(k). \tag{A.8}$$

Then

$$A = \exp\left(-\sum_{i=0}^{p} \frac{B_i}{n}\right) \prod_{i=0}^{p} \exp\left(\frac{B_i}{n}\right) = 1 + \frac{1}{2n^2} \sum_{i<j} [B_i, B_j] + \mathcal{O}(n^{-3}). \tag{A.9}$$

Noting that the commutators $[B_0, B_f]$ give rise to terms which are linear in ξ_f, we have

$$\begin{aligned} A_{22}(\xi) &= +\tfrac{1}{2} \sum_{f<g} \xi_f \xi_g (\beta/n)^2 \sum_k [V_f(k), V_g(k)] \\ &= +\tfrac{1}{2} \sum_{f<g} \xi_f \xi_g (\beta/n)^2 N C_{fg}, \end{aligned} \tag{A.10}$$

where $C_{fg} = (1/N) \sum_k [V_f(k), V_g(k)]$, is a bounded operator even in the thermodynamic limit. In order to find $\tilde{A}_{22}$ we have to perform the integration over ξ. Since the appropriate integration variables in (A.4) are again $(\beta N/2n)^{\frac{1}{2}} \xi$, each factor ξ contributes after integration a factor $(n/N)^{\frac{1}{2}}$. Consequently

$$\tilde{A}_{22} = (\beta^2/n) \sum_{f<g} Q_{fg}, \tag{A.11}$$

where the Q_{fg} are bounded also if $N \to \infty$.

Substituting (A.7) and (A.11) in the expression for $e^{-\beta \mathcal{H}}$ we find

$$Z = \operatorname{Tr} e^{-\beta \mathcal{H}} = (\beta N/2\pi n)^{\frac{1}{2}pn} \operatorname{Tr} (O_1 O_2)^n, \tag{A.12}$$

where

$$\begin{aligned} O_1 &\equiv \int d\xi \exp\left(-\frac{\beta N}{2n} \xi \cdot \xi\right) \exp\left(-\frac{\beta}{n} \sum_k \mathcal{H}(k, \xi)\right), \\ O_2 &\equiv 1 + (\beta^2/n) \sum_{f<g} Q_{fg}. \end{aligned} \tag{A.13}$$

In general, for finite values of N, the commutators, *i.e.*, the operator O_2 will give a non-negligible contribution to the free energy.

However, in the thermodynamic limit this contribution can be neglected. This can be shown by using a special case of the Hölder inequality for operators (*cf.* ref. 31).

$$\|O_2^{-1}\|^{-n} \operatorname{Tr} O_1^n \leq |\operatorname{Tr} (O_1 O_2)^n| \leq \|O_2\|^n \operatorname{Tr} O_1^n. \tag{A.14}$$

Here the operator O_1 is positive-definite hermitean and O_2 is a nonsingular operator for sufficiently large values of n. In the limit $n \to \infty$, the operator norms $\|O_2\|^n$ and $\|O_2^{-1}\|^n$ reduce to finite constants which can be neglected in the thermodynamic limit.

In this limit the free energy per particle is given by

$$f = -\lim_{N\to\infty} (\beta N)^{-1} \lim_{n\to\infty} \ln [(\beta N/2\pi n)^{\frac{1}{2}pn} \operatorname{Tr} O_1^n], \tag{A.15}$$

which corresponds to eq. (22).

APPENDIX B

In order to calculate the $(pn \times pn)$ matrix H which up to a factor β/n consists of the second derivatives of the function G_1, we introduce the short-hand notations

$$\varrho(i) = \exp [-(\beta/n)\, \mathscr{H}(k, \xi_i)] \quad \text{and} \quad T(\{\xi_i\}) = \operatorname{tr} \prod_{i=1}^{n} \varrho(i). \tag{B.1}$$

For the sake of convenience the k dependence has not been written down explicitly. From (28) and (B.1) we have

$$\frac{\partial G_1}{\partial \xi_j} = \frac{\beta}{n} \xi_j - \frac{1}{N} \operatorname{Re} \sum_k \frac{1}{T(\{\xi_i\})} \frac{\partial T(\{\xi_i\})}{\partial \xi_j}, \tag{B.2}$$

where again the k dependence is not shown explicitly. Hence

$$\frac{\partial^2 G_1}{\partial \xi_{fi}\, \partial \xi_{gj}} = \frac{\beta}{n} \delta_{fg} \delta_{ij} - \frac{1}{N} \operatorname{Re} \sum_k \frac{1}{T} \frac{\partial^2 T}{\partial \xi_{fi}\, \partial \xi_{gj}} + \frac{1}{N} \operatorname{Re} \sum_k \frac{1}{T^2} \left(\frac{\partial T}{\partial \xi_{fi}}\right) \left(\frac{\partial T}{\partial \xi_{gj}}\right). \tag{B.3}$$

First we take $j > i$. Considering the second term in the r.h.s. of (B.3) we see from (B.1) that

$$\frac{\partial^2 T}{\partial \xi_{fi}\, \partial \xi_{gj}} = \operatorname{tr} \left[\left(\prod_{l=1}^{i-1} \varrho(l) \right) \frac{\partial \varrho(i)}{\partial \xi_{fi}} \left(\prod_{l=i+1}^{j-1} \varrho(l) \right) \frac{\partial \varrho(j)}{\partial \xi_{gj}} \left(\prod_{l=j+1}^{n} \varrho(l) \right) \right].$$

Eq. (41) enables us to calculate the derivatives and (30) shows that at the minimum we have $\varrho(i) = \varrho = \exp [-(\beta/n)\, \mathscr{H}(k, \xi^0)]$, independent of i. Then, using also

the cyclic invariance of the trace

$$\left(\frac{\partial^2 T}{\partial \xi_{fi}\,\partial \xi_{gj}}\right)_{\min} = \int_0^{\beta/n} \mathrm{d}\tau_1 \int_0^{\beta/n} \mathrm{d}\tau_2 \,\mathrm{tr}\,[\mathrm{e}^{\tau_1 \mathscr{H}(k,\xi^0)}\, V_f(k)$$

$$\times\, \mathrm{e}^{-\tau_1 \mathscr{H}(k,\xi^0)}\, \varrho^{j-i}\, \mathrm{e}^{\tau_2 \mathscr{H}(k,\xi^0)}\, V_g(k)\, \mathrm{e}^{-\tau_2 \mathscr{H}(k,\xi^0)}\, \varrho^{n-j+i}]$$

$$= \int_0^{\beta/n} \mathrm{d}\tau_1 \int_{-\tau_1}^{(\beta/n)-\tau_1} \mathrm{d}\tau\, \mathrm{tr}\,[V_f(k)\, \varrho^{j-i}\, \mathrm{e}^{\tau \mathscr{H}(k,\xi^0)}\, V_g(k)$$

$$\times\, \mathrm{e}^{-\tau \mathscr{H}(k,\xi^0)}\, \varrho^{n-j+i}]. \tag{B.4}$$

We now consider a representation in which $\mathscr{H}(k, \xi^0)$ is diagonal. The eigenvalues are denoted by h_p; the k dependence is not indicated explicitly. Then:

$$T_{figj} \equiv \sum_k \mathrm{Re}\, \frac{n^2}{\beta^2} \left(\frac{\partial^2 T}{\partial \xi_{fi}\,\partial \xi_{gj}}\right)_{\min} = \sum_k \sum_{p,q} \mathrm{Re}\,(V_{f_{pq}} V_{g_{qp}}) \exp\left(-\frac{\beta}{n}(j-i)\, h_q\right)$$

$$\times \left[\exp\left(-\frac{\beta}{n}(n-j+i)\, h_p\right)\right] \frac{n^2}{\beta^2} \int_0^{\beta/n} \mathrm{d}t \int_{-t}^{(\beta/n)-t} \mathrm{d}\tau\, \mathrm{e}^{\tau(h_q - h_p)}. \tag{B.5}$$

Using $(V_{pq})^* = V_{qp}$, and evaluating the integral one finds (we still have $j > i$)

$$T_{figj} = \sum_k \sum_{p,q} V_{f_{pq}} V_{g_{qp}} \exp\left[-\tfrac{1}{2}\beta\,(h_p + h_q)\right] \cosh \gamma$$

$$\times\, \frac{2n^2}{\beta^2} \left(\frac{\cosh\left[(\beta/n)\,(h_q - h_p)\right] - 1}{(h_p - h_q)^2}\right), \tag{B.6}$$

where $\gamma = \beta\,[\tfrac{1}{2} - (j-i)/n]\,(h_q - h_p)$. (B.7)

We note four properties of the matrix elements T_{figj}

1) $T_{figj} = T_{gjfi}$

which is a direct consequence of the definition of the matrix elements as second derivatives,

2) $T_{figj} = T_{gifj}$, (B.8)

3) T_{figj} is a function of $j - i$ only;
it will be denoted by $T_{fg,\,j-i}$ and it is defined for $1 \le j - i \le n - 1$,

4) T_{figj} is a function of $|\tfrac{1}{2} - (j-i)/n|$
which shows that $T_{fg,\,j-i} = T_{fg,\,n-j+i}$.

Property 1 enables us to find T_{figj} for $i > j$, and together with property 2 it ensures that γ, see (B.7), which contains the whole i and j dependence of T_{figj} may be replaced by $\beta\,(\frac{1}{2} - |j - i|/n)\,(h_p - h_q)$, which is now valid for both $i < j$ and $j < i$.

In the calculation of T_{figi} one should note that the operators $e^{\tau\mathscr{H}}$ and $e^{-\tau\mathscr{H}}$, such as in eq. (41), depend on ξ_i. A straightforward but tedious calculation leads to the result

$$T_{figi} = \sum_{k}\sum_{p,q} \mathrm{Re}\,(V_{fpq}V_{gqp})\,\frac{n^2}{\beta^2}\,e^{-\beta h_p}\,\frac{2}{h_p - h_q}\left(\frac{e^{(\beta/n)(h_p - h_q)} - 1}{h_p - h_q} - \frac{\beta}{n}\right). \tag{B.9}$$

A comparison with (B.5) shows that (B.9) cannot be obtained from (B.6) by inserting $i = j$. Eqs. (B.6) and (B.9), however, are identical as far as the lowest-order terms in $1/n$ are concerned. Hence we can write

$$T_{figj} = T^{(1)}_{figj} + B_{figj}/n, \tag{B.10}$$

where

$$T^{(1)}_{figj} = \sum_{k}\sum_{p,q} V_{fpq}V_{gqp}\,e^{-\frac{1}{2}\beta(h_p + h_q)}\cosh\left[\beta\left(\tfrac{1}{2} - |j - i|/n\right)(h_p - h_q)\right], \tag{B.11}$$

for all values of j and i, including $j = i$ $(i, j = 1, \ldots, n)$, and where the matrix elements B_{figj} are bounded (for $i \neq j$, the elements are even of order $1/n$). The precise expressions for B_{figj} are irrelevant since in view of the application in section 7 of a lemma given by Lenard[38]), the B_{figj} will not give a contribution to the determinant of second derivatives in the limit $n \to \infty$. It is easy to show that the first three properties of (B.8), which are valid for the matrix elements T_{figj}, hold as well for the elements $T^{(1)}_{figj}$. In addition we have, using properties 2) and 3) of (B.8):

$$T^{(1)}_{fg,n-i} = T^{(1)}_{fg,i} = T^{(1)}_{fg,-i} = T^{(1)}_{fg,-n+i} \qquad \text{for} \qquad 1 \le i \le n-1, \tag{B.8'}$$

showing the periodicity in the label i.

Passing on now to the last term in the r.h.s. of (B.3) we see that we have from (41) and (44)

$$\left(\frac{1}{T}\frac{\partial T}{\partial \xi_j}\right)_{\min} = \frac{\beta}{n}\,\langle V(k)\rangle. \tag{B.12}$$

Combining eqs. (B.3) and (B.10)–(B.12) we finally arrive at

$$\left(\frac{\partial^2 G_1}{\partial \xi_{fi}\,\partial \xi_{gj}}\right)_{\min} = \frac{\beta}{n}\left(\delta_{ij}\delta_{fg} + \frac{A_{figj}}{n} + \frac{B_{figj}}{n^2}\right), \tag{B.13}$$

with

$$A_{figj} = \frac{\beta}{N} \sum_k \langle V_f(k) \rangle \langle V_g(k) \rangle - \frac{\beta}{N} \sum_k (\mathrm{tr}_k \, \mathrm{e}^{-\beta \mathscr{H}(k, \xi^0)})^{-1}$$

$$\times \sum_{pq} V_{f_{pq}} V_{g_{qp}} \, \mathrm{e}^{-\frac{1}{2}\beta (h_p + h_q)} \cosh [\beta (\tfrac{1}{2} - |j - i|/n) (h_p - h_q)]), \quad \text{(B.14)}$$

where $\boldsymbol{A}$ satisfies the same properties as $T^{(1)}$.

APPENDIX C

In this appendix we show in some more detail how the $pn \times pn$ matrix $\boldsymbol{A}$ is reduced to a sum of n blocks of $p \times p$ matrices.

Denoting the eigenvalues of $-\boldsymbol{A}/n$ by λ we have the following eigenvalue problem:

$$(1/n) \sum_{gj} A_{figj} \phi_{gj} = -\lambda \phi_{fi}. \quad \text{(C.1)}$$

Here ϕ_{gj} ($g = 1, \ldots, p$ and $j = 1, \ldots, n$) is the gj component of the eigenvector corresponding to the eigenvalue $-\lambda$. Up to now $A_{fg,i}$ was defined for $i = -n + 1, \ldots, 0, \ldots, n - 1$ only. Using (64 iv) we can give a periodic extension of the definitions of $A_{fg,i}$ and ϕ_{fi}, *viz.* for integers i' satisfying $i' = i \pmod n$ we define

$$A_{fg,i'} \equiv A_{fg,i} \quad \text{and} \quad \phi_{fi'} \equiv \phi_{fi}. \quad \text{(C.2)}$$

We arrive at the eigenvalue problem

$$\sum_{gj} A_{fg,j-i} \phi_{gj} = -\lambda \phi_{fi}, \quad \text{(C.3)}$$

where now i and j run through the set of all integers. The conclusion is that the eigenvectors can be chosen to be basis functions for the irreducible representations of the abelian group of translations $\{T_i\}$, hence satisfy

$$\phi_{fj} = \phi_f(\varkappa) \exp (2\pi i j \varkappa / n). \quad \text{(C.4)}$$

From the boundary condition (C.2) it follows that $\varkappa$ is an integer, so that we may restrict ourselves to $\varkappa = 0, \ldots, n - 1$.

The eigenvalue problem reduces to

$$\sum_{g=1}^{p} D_{fg}(\varkappa) \, \phi_g(\varkappa) = \lambda(\varkappa) \, \phi_f(\varkappa), \qquad \varkappa = 0, \ldots, n - 1, \quad \text{(C.5)}$$

where

$$D_{fg}(\varkappa) = -(1/n)\sum_{j=0}^{n-1} A_{fg,j} \exp(2\pi i j\varkappa/n), \tag{C.6}$$

and the eigenvalues of $-A/n$ are given by the eigenvalues $\lambda_f(\varkappa)$, $f = 1, \ldots, p$, of the matrices $D(\varkappa)$ for $\varkappa = 0, \ldots, n-1$.

REFERENCES

1) Mühlschlegel, B. and Zittartz, H., Z. Phys. **175** (1963) 553.
2) Niemeijer, Th., Physica **48** (1970) 467.
3) Temperley, H.N.V., Proc. Phys. Soc. **67** (1954) 233.
4) Katsura, S., Progr. theor. Phys. **13** (1955) 571.
5) Kac, M., Phys. Fluids **2** (1959) 8.
6) Baker Jr., G.A., Phys. Rev. **122** (1961) 1477.
7) Siegert, A.F.J. and Vezzetti, D.J., J. math. Phys. **9** (1968) 2173.
8) Kac, M., Uhlenbeck, G.E. and Hemmer, P.C., J. math. Phys. **4** (1963) 216, 229.
9) Van Kampen, N.G., Phys. Rev. **135A** (1964) 362; Physica **48** (1970) 313.
10) Lebowitz, J.L. and Penrose, O., J. math. Phys. **7** (1966) 98.
11) Brout, R., Phys. Rev. **118** (1960) 1009.
12) Horwitz, G. and Callen, H.B., Phys. Rev. **124** (1961) 757.
13) Fisher, M.E. and Gaunt, D.S., Phys. Rev. **133** (1964) 224.
14) Hallers, J.J. and Vertogen, G., Phys. Rev. **B4** (1971) 2351; Phys. Rev. Letters **27** (1971) 404.
Vertogen, G. and Hallers, J.J., Phys. Rev. Letters **28** (1972) 563.
15) Hubbard, J., Proc. Roy. Soc. **A276** (1963) 238.
16) Capel, H.W., Physica **32** (1966) 966, **33** (1967) 295.
17) Blume, M., Phys. Rev. **141** (1966) 517.
18) Blume, M. and Watson, R.E., J. appl. Phys. **38** (1967) 991A.
19) Kittel, C. and Shore, H., Phys. Rev. **138A** (1965) 1165.
20) Mühlschlegel, B., J. math. Phys. **3** (1962) 522.
21) Bardeen, J., Cooper, L.N. and Schrieffer, J.R., Phys. Rev. **108** (1957) 1175.
22) Bogoliubov, N.N., Zubarev, D.N. and Tserkovnikov, I.A., Soviet Physics–Doklady **2** (1958) 535.
23) Haag, R., Nuovo Cimento **25** (1962) 287.
24) Thirring, W. and Wehrl, A., Commun. math. Phys. **4** (1967) 303.
Thirring, W., Commun. math. Phys. **7** (1968) 181.
25) Emch, G.C., J. math. Phys. **8** (1967) 13.
Emch, G.C. and Knops, H.J.F., J. math. Phys. **11** (1970) 3008.
26) Vertogen, G. and de Vries, A.S., Physica **59** (1972) 634.
27) Feynman, R.P., Phys. Rev. **84** (1951) 108.
28) Trotter, H., Proc. Amer. Math. Soc. **10** (1959) 545.
29) Reed, M. and Simon, B., Methods of Mathematical Physics, part 1: Functional Analysis, p. 295. Academic Press (New York, 1972).
30) *E.g.*, Falk, H., Amer. J. Phys. **38** (1970) 858.
31) Capel, H.W. and Tindemans, P.A.J., Rep. math. Phys., to be published.
32) *E.g.*, Elliott, R.J. and Thorpe, M.F., J. appl. Phys. **39** (1968) 802.
33) Stratonovich, R.L., Soviet Physics–Doklady **2** (1958) 416.

34) Butler, S.T. and Friedman, M.H., Phys. Rev. **98** (1955) 287.
35) Schrieffer, J.R., Lecture Notes C.A.P. Summer School, Banff, 1969.
36) Sivardière, J. and Blume, M., Phys. Rev. **B5** (1972) 1126.
Blume, M., Emery, V.J. and Griffiths, R.B., Phys. Rev. **A4** (1971) 2071.
37) Cramer, H., Mathematical Methods of Statistics, Princeton University Press (Princeton 1946), p. 120.
38) Lenard, A., J. math. Phys. **5** (1964) 930.
39) Tindemans, P.A.J. and Capel, H.W., to be published.

IV. AN EXACT CALCULATION OF THE FREE ENERGY IN SYSTEMS WITH SEPARABLE INTERACTIONS. II

Synopsis

In this paper we extend the calculation of the free energy in systems with separable interactions given in a previous paper, to a more general class of systems characterized by a hamiltonian which contains a number of separable two-particle operators of the antiferromagnetic type in addition to separable ferromagnetic interactions and one-particle operators.

By deriving an upper bound and a lower bound we establish an expression for the free energy which is of the molecular-field type. In the derivation of the lower bound we have used Laplace's method in order to evaluate a multidimensional integral of a function e^{-NG}. The proof that the second derivatives at the absolute minimum do not give a contribution in the thermodynamic limit is more complicated than in the ferromagnetic case and is given in detail.

1. *Introduction.* In a previous paper[1]) we considered a general class of systems described by a one-particle operator and some separable two-particle operators. The hamiltonian can be written

$$\mathscr{H} = \sum_{k=1}^{N} T(k) - (1/2N) \sum_{f=1}^{p} \left\{ \sum_{k=1}^{N} V_f(k) \right\}^2 . \tag{1}$$

Here $T(k)$ and $V_f(k)$, for $f = 1, \ldots, p$, are operators which are defined on a finite dimensional Hilbert space H_k, for $k = 1, 2, \ldots, N$, respectively. Various examples of such a hamiltonian have been mentioned in ref. 1. Here we note that the Hilbert space H_k may be interpreted as the space of states belonging to a "particle k" (although grand canonical ensembles of many fermion systems are by no means excluded). In this language $T(k)$ and $V_f(k)$ are one-particle operators.

The interaction between different particles k and $l \neq k$ is given by

$$-(1/N) \sum_f V_f(k, l) = -(1/N) \sum_f V_f(k)\, V_f(l), \tag{2}$$

i.e. each two-particle operator $V_f(k, l)$ can be written as a product of one-particle operators $V_f(k)$ and $V_f(l)$. In the particular case of operators $\sum_k V_f(k)$ which are

invariant for permutations of the particles, the interaction between k and l is independent of the choice of k and l and can be called an *equivalent neighbour interaction*. It can also be considered as an extreme case of a long-range interaction.

In ref. 1 we have given an exact calculation of the free energy per particle in the thermodynamic limit for systems described by the hamiltonian (1). In this proof no assumptions have been made on the commutation properties of the operators $T(k)$ and $V_f(k)$. As a result we obtained an expression for the free energy of the molecular-field type, *i.e.* the expression can be obtained from a suitable one-particle hamiltonian in terms of p parameters corresponding to the different interactions in eq. (1). These parameters satisfy equations of the molecular-field type. In addition the parameters must be chosen in such a way that this function assumes its absolute minimum. This result is in agreement with the general idea that the molecular-field approximation can give rigorous results in the thermodynamic limit, if the range of the interaction tends to infinity, *cf.*, *e.g.* Mühlschlegel and Zittartz[2]), Kac[3]), Baker[4]), Siegert and Vezetti[5]) for the Ising model, Kittel and Shore[6]) and Niemeyer[7]) for the Heisenberg model and Mühlschlegel[8]) in the case of the so-called reduced hamiltonian in the BCS theory of superconductivity[9]). Additional references, for instance on the C^* algebra type of approach can be found in ref. 1.

Note that the coefficients of all separable interactions in eq. (1) are *negative*. For that reason these interactions may be called "*ferromagnetic*", *cf.* the simple case that $V_f(k) = S_k^z$, *i.e.* the z component of spin k. In the present paper we extend our considerations to a larger class of systems including also a finite number of separable interactions with positive coefficients. These interactions may be called "antiferromagnetic" and the hamiltonian can be written

$$\mathscr{H} = \sum_{k=1}^{N} T(k) - (1/2N) \sum_{f=1}^{p} \left\{ \sum_{k=1}^{N} V_f(k) \right\}^2 + (1/2N) \sum_{a=1}^{q} \left\{ \sum_{k=1}^{N} W_a(k) \right\}^2 . \quad (3)$$

Here $T(k)$, $V_f(k)$, $f = 1, \ldots, p$, and $W_a(k)$, for $a = 1, \ldots, q$ are bounded hermitean operators which are defined on the Hilbert space H_k and which may be interpreted as one-particle operators. The antiferromagnetic and ferromagnetic interactions have been labeled by a and f, respectively.

A particular system belonging to the general class, as defined by eq. (3), has been treated by De Vries, Vertogen and Kraak[10]). The hamiltonian of this system is given by

$$\mathscr{H} = (1/N) \left\{ J_1 \sum_{i,j=1}^{N} \boldsymbol{S}_i \cdot \boldsymbol{S}_j + J_2 \sum_{i,j=1}^{N} \boldsymbol{T}_i \cdot \boldsymbol{T}_j + J_3 \sum_{i,j=1}^{N} \boldsymbol{S}_i \cdot \boldsymbol{T}_j \right\} . \quad (4)$$

Eq. (4) describes an equivalent-neighbour coupling between spins $\boldsymbol{S}$ and $\boldsymbol{T}$ belonging to different sublattices A and B, respectively. J_1 and J_2 are the interaction

strengths between spins within the same sublattice A and B, respectively. J_3 gives the interaction between spins belonging to different sublattices. Starting from a condition which ensures the validity of the so-called Bogoliubov–Haag procedure[11]), they use a C^* algebra approach in order to calculate the free energy in extremal homogeneous states satisfying the KMS condition.

In this paper we shall show that the free energy for a system defined by eq. (3) can be expressed in terms of p parameters $\xi_1, \xi_2, \ldots, \xi_p$ corresponding to the interactions $V_1, V_2, \ldots, V_p$ und q parameters $\eta_1, \eta_2, \ldots, \eta_q$ corresponding to the interactions $W_1, W_2, \ldots, W_q$.

The result is

$$f = \lim_{N\to\infty} (-N^{-1}k_{\mathrm{B}}T \ln Z_N) = \lim_{N\to\infty} \left(\tfrac{1}{2} \sum_{f=1}^{p} \xi_f^2 - \tfrac{1}{2} \sum_{a=1}^{q} \eta_a^2 \right.$$
$$\left. - N^{-1}k_{\mathrm{B}}T \sum_{k=1}^{N} \ln \mathrm{tr}_k \exp\left[-\beta \left\{ T(k) + \sum_{a=1}^{q} \eta_a W_a(k) - \sum_{f=1}^{p} \xi_f V_f(k) \right\} \right] \right). \tag{5}$$

Here tr_k is the trace over the Hilbert space H_k and ξ_f and η_a are determined by the implicit equations

$$\xi_f = N^{-1} \sum_{k=1}^{N} T_k^{-1} \,\mathrm{tr}_k\, V_f(k) \exp\left[-\beta \left\{ T(k) + \sum_{a=1}^{q} \eta_a W_a(k) - \sum_{f=1}^{p} \xi_f V_f(k) \right\} \right],$$
$$\eta_a = N^{-1} \sum_{k=1}^{N} T_k^{-1} \,\mathrm{tr}_k\, W_a(k) \exp\left[-\beta \left\{ T(k) + \sum_{a=1}^{q} \eta_a W_a(k) - \sum_{f=1}^{p} \xi_f V_f(k) \right\} \right], \tag{6}$$

where

$$T_k^{-1} \equiv \mathrm{tr}_k \exp\left[-\beta \left\{ T(k) + \sum_{a=1}^{q} \eta_a W_a(k) - \sum_{f=1}^{p} \xi_f V_f(k) \right\} \right]. \tag{7}$$

If eq. (6) has more than one solution we should chose that particular solution which leads to the lowest value of f. This lowest value of f is the free energy per particle in the thermodynamic limit.

In the absence of antiferromagnetic interactions W, eq. (5) reduces to the free energy given in ref. 1, *cf.* eqs. (I.29) and (I.15). In ref. 1 the free energy was obtained by deriving an upper bound on the basis of a variational type of argument and a lower bound using Laplace's method. In the present paper we shall also prove that the right-hand side of eq. (5) provides an upper bound as well as a lower bound to the free energy per particle.

The upper bound of the free energy will be derived in section 2. The derivation of the lower bound is much harder and will be treated in the remaining sections of this paper. This derivation involves in particular the introduction of a set of

parameters corresponding to the antiferromagnetic interactions. This point will be discussed in more detail in section 3, where we derive an integral representation $\int e^{-NG}$ for the partition function.

In addition we prove that the real part of the function G cannot be smaller than βf where f is given by eq. (5). A careful analysis is necessary in order to show that the integral $\int e^{-N(G-\beta f)}$ does not give a contribution to the free energy per particle in the thermodynamic limit. For that purpose we derive in section 4 another integral representation which will enable us to study both the absolute minimum of the real part of G and the contribution due to its matrix of second derivatives. This will be done in sections 5 and 6. Finally in section 7 a simple example is discussed.

2. *An upper bound for the free energy.* The upper bound for the free energy will be derived by using a variational type of argument based on Bogoliubov's inequality.

We write the hamiltonian $\mathscr{H}$ as

$$\mathscr{H} = \tfrac{1}{2}N(\xi^2 - \eta^2) + \mathscr{H}_0 + \mathscr{H}_1, \tag{8}$$

where

$$\mathscr{H}_0(\xi, \eta) = \sum_{k=1}^{N} \{T(k) - \xi \cdot V(k) + \eta \cdot W\} \tag{9}$$

and

$$\mathscr{H}_1(\xi, \eta) = \mathscr{H}'_1(\xi, \eta) + \mathscr{H}''_1(\xi, \eta) \tag{10a}$$

with

$$\mathscr{H}'_1 = -(1/2N)\Big(\sum_k V(k) - N\xi\Big)^2$$

$$\text{and} \quad \mathscr{H}''_1 = (1/2N)\Big(\sum_k W(k) - N\eta\Big)^2. \tag{10b}$$

Here we have introduced the following vector notations: ξ and $V(k)$ are p-dimensional vectors with components $\xi_1, \xi_2, \ldots, \xi_p$ and $V_1(k), V_2(k), \ldots, V_p(k)$, respectively, whereas η and $W(k)$ denote q-dimensional vectors $\eta_1, \ldots, \eta_q$ and $W_1(k), \ldots, W_q(k)$, respectively. No confusion should arise from this notation: it is implied that the symbols ξ and V always refer to p-dimensional vectors, whereas η and W refer to q-dimensional vectors exclusively.

The sets $(\xi_1, \ldots, \xi_p)$ and $(\eta_1, \ldots, \eta_q)$ consist of arbitrary real parameters.

On account of the Peierls–Bogoliubov inequality[12]), we have

$$F[\mathscr{H}_0 + \mathscr{H}_1] \leq F[\mathscr{H}_0] + \langle \mathscr{H}_1 \rangle_{\mathscr{H}_0}, \tag{11}$$

for any two hermitean operators $\mathscr{H}_0$ and $\mathscr{H}_1$, where

$$F[A] \equiv -\beta^{-1} \ln \operatorname{Tr} e^{-\beta A}, \tag{12}$$

$$\langle B \rangle_A \equiv \operatorname{Tr} B\, e^{-\beta A} / \operatorname{Tr} e^{-\beta A}. \tag{13}$$

In contrast with the ferromagnetic case we do not have the simple inequality $\langle \mathscr{H}_1 \rangle_{\mathscr{H}_0} \leq 0$, since the second term $\langle \mathscr{H}_1'' \rangle_{\mathscr{H}_0}$ gives always a positive contribution.

On the other hand, we have the parameters $\boldsymbol{\eta}$ at our disposal and if we can choose these parameters in such a way that $\langle \mathscr{H}_1'' \rangle_{\mathscr{H}_0}$ is of order 1 rather than of order N, the second term can be neglected in the thermodynamic limit.

From the definition (10) we have,

$$\langle \mathscr{H}_1'' \rangle_{\mathscr{H}_0} = (1/2N) \sum_{k,l} \langle \boldsymbol{W}(k) \cdot \boldsymbol{W}(l) \rangle_{\mathscr{H}_0} - \boldsymbol{\eta} \cdot \sum_k \langle \boldsymbol{W}(k) \rangle_{\mathscr{H}_0} + \tfrac{1}{2} N \boldsymbol{\eta}^2. \tag{14}$$

Since $\mathscr{H}_0(\boldsymbol{\xi}, \boldsymbol{\eta})$ is a sum of one-particle operators $\mathscr{H}(k, \boldsymbol{\xi}, \boldsymbol{\eta})$, where

$$\mathscr{H}(k, \boldsymbol{\xi}, \boldsymbol{\eta}) = T(k) + \boldsymbol{\eta} \cdot \boldsymbol{W}(k) - \boldsymbol{\xi} \cdot \boldsymbol{V}(k), \tag{15}$$

we have for $k \neq l$,

$$\langle \boldsymbol{W}(k) \cdot \boldsymbol{W}(l) \rangle_{\mathscr{H}_0} = \langle \boldsymbol{W}(k) \rangle_k \cdot \langle \boldsymbol{W}(l) \rangle_l, \tag{16}$$

where

$$\langle \boldsymbol{W}(k) \rangle_k \equiv \operatorname{tr}_k \boldsymbol{W}(k) \exp[-\beta \mathscr{H}(k, \boldsymbol{\xi}, \boldsymbol{\eta})] / \operatorname{tr}_k \exp[-\beta \mathscr{H}(k, \boldsymbol{\xi}, \boldsymbol{\eta})], \tag{17}$$

and the trace tr_k is taken over the Hilbert space of particles k. If we choose

$$\boldsymbol{\eta} = (1/N) \sum_k \langle \boldsymbol{W}(k) \rangle_k, \tag{18}$$

then

$$\langle \mathscr{H}_1'' \rangle_{\mathscr{H}_0} = (1/2N) \sum_k \{ \langle \boldsymbol{W}^2(k) \rangle_{\mathscr{H}(k, \boldsymbol{\xi}, \boldsymbol{\eta})} - \langle \boldsymbol{W}(k) \rangle^2_{\mathscr{H}(k, \boldsymbol{\xi}, \boldsymbol{\eta})} \} = \mathscr{O}(1), \tag{19}$$

where $\mathscr{O}(1)$ is a shorthand notation for a term which is of order 1 rather than of order N.

Using eqs. (8), (11), (19), and the relation $\langle \mathscr{H}_1' \rangle_{\mathscr{H}_0} \leq 0$, we have

$$F[\mathscr{H}] \leq F[\mathscr{H}_0(\boldsymbol{\xi}, \boldsymbol{\eta})] + (N/2)(\boldsymbol{\xi}^2 - \boldsymbol{\eta}^2) + \mathscr{O}(1).$$

For the free energy per particle it follows that

$$f \leq \beta^{-1} \phi(\boldsymbol{\xi}, \boldsymbol{\eta}), \tag{20}$$

where

$$\phi(\xi, \eta) \equiv \tfrac{1}{2}\beta(\xi^2 - \eta^2) - (1/N)\sum_k \ln \mathrm{tr}_k\, \mathrm{e}^{-\beta\mathscr{H}(k,\xi,\eta)}. \tag{21}$$

Eq. (20) holds for all ξ and η, provided that eq. (18) is satisfied. Consider now the identity

$$\frac{\partial}{\partial\lambda_j}\mathrm{e}^{-\beta A} = \mathrm{e}^{-\beta A}\int_0^\beta \mathrm{d}\tau\, \mathrm{e}^{\tau A} A_j\, \mathrm{e}^{-\tau A} \tag{22}$$

for an arbitrary operator $A = -\sum_i \lambda_i A_i$. From (21) and (22) it follows that

$$\partial\phi/\partial\eta = -\beta\left\{\eta - (1/N)\sum_k \langle W(k)\rangle_k\right\}, \tag{23}$$

so that (18) is equivalent to the condition $\partial\phi/\partial\eta = 0$. (24)

The upper bound can now be written

$$f \leq \beta^{-1}\min_\xi \phi(\xi, \eta_{mf}(\xi)), \tag{25}$$

where $\eta_{mf}(\xi)$ is a solution of $\partial\phi/\partial\eta = 0$.

The upper bound [*i.e.* essentially the function $\phi(\xi, \eta)$] clearly has a molecular-field like nature. Therefore, and also for later reference, it is useful to discuss some properties of $\phi(\xi, \eta)$.

I. Eq. (24) defines a unique function $\eta = \eta_{mf}(\xi)$, which is such that for fixed ξ, $\phi(\xi, \eta)$ has a maximum for $\eta = \eta_{mf}(\xi)$. In the proof use will be made of Bogoliubov's inequality (11) with

$$\mathscr{H} \equiv \mathscr{H}_0(\xi, \eta), \qquad \mathscr{H}_0 \equiv \mathscr{H}_0(\xi, \eta_0), \qquad \mathscr{H}_1 \equiv \sum_{k=1}^{N}(\eta - \eta_0)\cdot W(k), \tag{26}$$

where η_0 is an arbitrary point. Then

$$F[\mathscr{H}_0(\xi, \eta)] \leq F[\mathscr{H}_0(\xi, \eta_0)] + (\eta - \eta_0)\cdot\left\langle\sum_k W(k)\right\rangle_{\mathscr{H}_0(\xi,\eta_0)}. \tag{27}$$

We now expand the left-hand side into a Taylor-series around the point $\eta = \eta_0$. Since the first derivative of F with respect to η is just the average of $\sum_k W(k)$ with respect to $\mathscr{H}_0(\xi, \eta_0)$, it follows that in any point η_0, the matrix of second derivatives $\partial^2 F/\partial\eta\,\partial\eta$ is negative semi-definite. From eqs. (21), (15) and (26) we now have for $\eta \neq \eta_0$:

$$(\eta - \eta_0)\cdot\left(\frac{\partial^2\phi}{\partial\eta\,\partial\eta}\right)_{\eta=\eta_0}\cdot(\eta - \eta_0) = -\tfrac{1}{2}\beta(\eta - \eta_0)^2$$
$$+ \beta N^{-1}(\eta - \eta_0)\cdot\left(\frac{\partial^2 F}{\partial\eta\,\partial\eta}\right)_{\eta=\eta_0}\cdot(\eta - \eta_0) < 0. \tag{28}$$

Assume now that eq. (24) for fixed ξ has two different solutions

$$\eta_1 \quad \text{and} \quad \eta_2 = \eta_1 + e,$$

then there exists a point

$$\eta^* = \eta_1 + te \quad \text{with} \quad 0 < t < 1,$$

such that

$$e \cdot \frac{\partial^2 \phi}{\partial \eta \, \partial \eta} \cdot e = 0 \quad \text{for} \quad \eta = \eta^*. \tag{29}$$

Eq. (29) is in contradiction with (28) so that (24) can have at most one solution. On the other hand, there must be a solution since $\phi(\xi, \eta)$ for fixed ξ tends to $-\infty$ in the limit $|\eta| \to \infty$ so that $\phi(\xi, \eta)$ as a function of ξ must have a maximum.

II. Since eq. (24) has only one solution the upper bound (25) can be written

$$f \leq \beta^{-1} \min_{\xi} \Phi(\xi), \tag{30}$$

where $\Phi(\xi) = \phi(\xi, \eta_{mf}(\xi))$ is a unique function of ξ. The minimum of $\Phi(\xi)$ satisfies $\mathrm{d}\Phi/\mathrm{d}\xi = 0$ and from (24) it follows that also $\partial\phi/\partial\xi = 0$, so that

$$f \leq \beta^{-1} \min \phi(\xi, \eta) \tag{31a}$$

under the condition that

$$\partial\phi/\partial\xi = 0, \qquad \partial\phi/\partial\eta = 0. \tag{31b}$$

From (23) and a similar calculation of $\partial\phi/\partial\xi$ it follows that eqs. (31b) are equivalent to the molecular-field equations (6).

III. From eq. (18) it follows that $|\eta_{mf}(\xi)| \leq (1/N) \sum_k \|W(k)\|$ so that $\eta_{mf}(\xi)$ is a bounded function of ξ. An implicit equation for its derivative can be found from (18) using again (22):

$$\mathrm{d}\eta_a/\mathrm{d}\xi = (1/N) \sum_{k=1}^{N} T_k^{-1} \, \mathrm{tr}_k \left\{ \int_0^\beta \mathrm{d}\tau \, \mathrm{e}^{(\tau-\beta)\mathscr{H}(k,\xi,\eta)} \, \tilde{V}_f(k) \, \mathrm{e}^{-\tau\mathscr{H}(k,\xi,\eta)} \, \tilde{W}_a(k) \right\}. \tag{32}$$

Here, as well as in the rest of this section, η is meant to be η_{mf}. Furthermore T_k is defined in (7), and

$$\tilde{V}_f(k) = V_f(k) - \langle V_f(k) \rangle_k - (W(k) - \langle W(k) \rangle_k) \cdot \mathrm{d}\eta/\mathrm{d}\xi_f \tag{33a}$$

and

$$\tilde{W}_a(k) = W_a(k) - \langle W_a(k)\rangle_k, \tag{33b}$$

where the averages are with respect to $\mathscr{H}(k, \xi, \eta)$.

IV. The molecular-field equations (6) have already been seen to be equivalent to $\mathrm{d}\Phi/\mathrm{d}\xi = 0$. Only those solutions for which the matrix $\mathrm{d}^2\Phi/\mathrm{d}\xi\,\mathrm{d}\xi$ is positive definite, are stable (*i.e.* Φ has a minimum). The matrix $\mathrm{d}^2\Phi/\mathrm{d}\xi\,\mathrm{d}\xi$ can be evaluated using (22). The result is

$$\beta^{-1}\,\mathrm{d}^2\Phi/\mathrm{d}\xi\,\mathrm{d}\xi = 1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - (1/N)\sum_{k=1}^{N} T_k^{-1}$$
$$\times\ \mathrm{tr}_k\left\{\tilde{V}(k)\int_0^\beta \mathrm{d}\tau\, \mathrm{e}^{(\tau-\beta)\mathscr{H}(k,\xi,\eta)}\,\tilde{V}(k)\,\mathrm{e}^{-\tau\mathscr{H}(k,\xi,\eta)}\right\}, \tag{34}$$

where 1_p is the $p \times p$ unit matrix. The symbol $(\mathrm{d}\eta/\mathrm{d}\xi)\cdot(\mathrm{d}\eta/\mathrm{d}\xi)$ will be used in this paper to denote the $p \times p$ matrix obtained from the dyadic product $(\mathrm{d}\eta/\mathrm{d}\xi) \times (\mathrm{d}\eta/\mathrm{d}\xi)$ by taking the inner product with respect to the η's. So $(\mathrm{d}\eta/\mathrm{d}\xi)\cdot(\mathrm{d}\eta/\mathrm{d}\xi) = \sum_{a=1}^{q}(\mathrm{d}\eta_a/\mathrm{d}\xi)(\mathrm{d}\eta_a/\mathrm{d}\xi)$.

3. *An integral representation for the partition function.* In the derivation of a lower bound for the free energy complications can arise from the presence of antiferromagnetic operators in the hamiltonian (3). At this stage it may be worthwile to point out that there is *a priori* no reason to expect that the antiferromagnetic interactions W should be treated in a similar way as the ferromagnetic interactions V. For this purpose we review some aspects of the derivation of the upper bound in section 2. There it was seen that the average $\langle\mathscr{H}'_1\rangle_{\mathscr{H}_0}$ of the negative definite operator $\mathscr{H}'_1$ which is due to the ferromagnetic operators, could be neglected in obtaining the upper bound. This, however, was not true for the average of $\mathscr{H}''_1$ which contains the antiferromagnetic interaction operators. Instead, we had to make a very definite choice for the parameters η which were introduced in (9), thereby minimizing the effect of the quadratic antiferromagnetic operators in $\mathscr{H}''_1$. As a result the contribution of the average of $\mathscr{H}''_1$ could be neglected in the thermodynamic limit. From this discussion it seems reasonable to treat the ferromagnetic and the antiferromagnetic operators in a different way: the ferromagnetic operators will be dealt with in a manner similar to ref. 1, whereas the effect of the antiferromagnetic ones will be treated by introducing parameters comparable to those that were used to obtain an upper bound for the free energy.

First we apply the generalized Lie–Totter formula[13]) in order to replace $\mathrm{e}^{-\beta\mathscr{H}}$ by a product of exponential operators

$$\exp(A_0 + A_1 + \cdots + A_{p+q}) = \lim_{n\to\infty}\left(\exp\frac{A_0}{n}\exp\frac{A_1}{n}\cdots\exp\frac{A_{p+q}}{n}\right)^n, \tag{35}$$

where

$$A_0 = -\beta \sum_k T(k),$$

$$A_f = (\beta/2N) \left[\sum_k V_f(k)\right]^2, \quad f = 1,\ldots,p, \tag{36}$$

$$A_{p+a} = -(\beta/2N) \left[\sum_k W_a(k)\right]^2, \quad a = 1,\ldots,q.$$

In ref. 1 we linearized the squares occurring in the exponential operators by means of a simple integral trick, *cf.*, *e.g.* Stratonovich[14])

$$e^{A^2} = \pi^{-\frac{1}{2}} \int_{-\infty}^{\infty} dx\, e^{-x^2} e^{2xA}. \tag{37}$$

In view of the remarks made above on the antiferromagnetic operators we apply this integral transformation for the moment only to each of the ferromagnetic operators $\exp\,[(\beta/2nN)\,(\sum_k V_f(k))^2]$.

The result is

$$e^{-\beta\mathscr{H}} = \lim_{n\to\infty} \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn} \left\{ \int_{-\infty}^{\infty} d\xi \exp\left(-\frac{\beta N}{2n}\,\xi\cdot\xi\right) \right.$$

$$\times \exp\left[-\frac{\beta}{n}\sum_k T(k)\right] \prod_{f=1}^{p} \exp\left[\frac{\beta}{n}\,\xi_f \sum_k V_f(k)\right]$$

$$\left. \times \prod_{a=1}^{q} \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k)\right)^2\right]\right\}^n. \tag{38}$$

As for the antiferromagnetic operators, we hope to "tame" then to a large extent by introducing an arbitrary set of real functions $\eta_a(\xi)$, $a = 1, \ldots, q$, corresponding to the antiferromagnetic operators W_a.

This can be done by using the trivial identity

$$\exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k)^2\right)\right] = \exp\left[\frac{\beta N}{2n}\,\eta_a^2(\xi)\right] \exp\left[-\frac{\beta}{n}\sum_k W_a(k)\,\eta_a(\xi)\right]$$

$$\times \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a(\xi)\right)^2\right]. \tag{39}$$

The right-hand side of eq. (39) contains three factors, first a constant factor $\exp\,[\beta N\,(2n)^{-1}\,\eta_a^2(\xi)]$, then a Boltzmann factor corresponding to the sum of one-particle operators and a Boltzmann factor which contains interactions between all pairs of particles k and l. Such an interaction term is in general difficult to deal

with, but here we have the parameters $\eta_a(\xi)$ which are arbitrary real functions of the parameters $\xi = (\xi_1, \xi_2, \ldots, \xi_p)$ and which can be chosen in such a way that the interaction term becomes small. We may hope that this will prove out to be sufficient for the derivation of a lower bound. (As we shall see later on, parameters η which are independent of ξ, will not always lead to the correct result.) In the calculation of the integral over ξ which can be obtained from (38) by substituting the identity (39), it is convenient to make the replacement

$$\exp\left[-\frac{\beta}{n}\sum_k T(k)\right]\prod_{f=1}^{p}\exp\left[\frac{\beta}{n}\xi_f\sum_k V_f(k)\right]\prod_{a=1}^{q}\exp\left[-\frac{\beta}{n}\sum_k W_a(k)\,\eta_a(\xi)\right]$$
$$\times\quad \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a(\xi)\right)^2\right]$$
$$\to \exp\left[-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi,\eta(\xi))\right]\prod_a \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a(\xi)\right)^2\right], \tag{40}$$

where $\mathscr{H}(k, \xi, \eta)$ has been defined in eq. (15).

This replacement is similar to the one used in ref. 1, *cf.* eq. (I.21); the proof of its validity, however, is somewhat more complicated due to the exponential operators involving $W_a(k)$. In appendix A it is shown that the contribution to the integral due to the commutators which are neglected, amounts to an additional factor. This factor remains finite in the thermodynamic limit and can be neglected in the calculation of the free energy per particle. In the remaining formulae for $e^{-\beta\mathscr{H}}$ and the partition function finite values of n and N occur, just as in ref.1. In the notation it will henceforth be implied that two limits have to be taken, first the limit $n \to \infty$ and afterwards the thermodynamic limit $N \to \infty$. Inserting (40) and (39) into (38), we arrive at the following integral representation for the partition function

$$Z_N = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}pn}\int_{-\infty}^{\infty}\prod_{i=1}^{n} d\xi_i\, e^{-NG(\{\xi_i\})}, \tag{41}$$

where $G(\{\xi_i\})$ is a function of all the np variables ξ_{if}. The explicit formula is given by

$$G(\{\xi_i\}) = (\beta/2n)\sum_{i=1}^{n}(\xi_i^2 - \eta_i^2(\xi))$$
$$- N^{-1}\ln \mathrm{Tr}\prod_{i=1}^{n}\left\{\exp\left[-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi_i,\eta(\xi_i))\right]\right.$$
$$\left.\times\prod_a \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a(\xi_i)\right)^2\right]\right\}. \tag{42}$$

In the absence of antiferromagnetic operators W and with the choice $\eta_a(\xi_i) = 0$, eq. (41) reduces to the integral representation (I.26) derived for the ferromagnetic case. In ref. 1 the absolute minimum of $G_1 \equiv \mathrm{Re}\, G$ was determined by using the Hölder inequality for operators together with the condition under which the equality sign holds. The lower bound for the free energy was obtained by using the inequality

$$\left|\int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-N(G-G_0)}\right| \leq \int \prod_i \mathrm{d}\xi_i \, \mathrm{e}^{-N(G_1-G_0)}. \tag{43}$$

The integral in the right-hand side of (43) was evaluated in ref. 1 by using Laplace's method. The contribution to the free energy from the determinant of second derivatives of G_1 at its absolute minimum was shown to be finite in the limits $n \to \infty$, $N \to \infty$.

In the present case, where antiferromagnetic operators are included in the hamiltonian, we cannot determine the absolute minimum of G_1 exactly. However, we can easily give a lower bound for G_1 by means of the following form of the Hölder inequality for operators[15]).

Let $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ be arbitrary matrices, then

$$|\mathrm{Tr}\, A_1 B_1 \cdots A_n B_n| \leq \prod_{i=1}^{n} \{\mathrm{Tr}\,(A_i^\dagger A_i)^{\frac{1}{2}\theta_i}\}^{1/\theta_i} \prod_{i=1}^{n} \{\mathrm{Tr}\,(B_i^\dagger B_i)^{\frac{1}{2}\lambda_i}\}^{1/\lambda_i} \tag{44}$$

for arbitrary positive numbers λ_i and θ_i satisfying $\sum_{i=1}^{n} (\lambda_i^{-1} + \theta_i^{-1}) = 1$. We apply this inequality with

$$A_i = \exp\left[-\frac{\beta}{n} \sum_k \mathscr{H}(k, \xi_i, \eta(\xi_i))\right],$$
$$B_i = \prod_{a=1}^{q} \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a(\xi_i)\right)^2\right], \tag{45}$$

in the case that $\lambda_i = \infty$ and $\theta_i = n$, and use the fact that

$$\lim_{\lambda_i \to \infty} \{\mathrm{Tr}\,(B_i^\dagger B_i)^{\frac{1}{2}\lambda_i}\}^{1/\lambda_i} = \|(B_i^\dagger B_i)^{\frac{1}{2}}\| \leq 1. \tag{46}$$

Then

$$\begin{aligned} G_1(\{\xi_i\}) \equiv \mathrm{Re}\, G(\{\xi_i\}) &\geq (\beta/2n) \sum_i (\xi_i^2 - \eta^2(\xi_i)) \\ &\quad - (nN)^{-1} \sum_{i=1}^{n} \sum_{k=1}^{N} \ln \mathrm{tr}_k \exp\left[-\beta \mathscr{H}(k, \xi_i, \eta(\xi_i))\right] \\ &= (1/n) \sum_{i=1}^{n} \phi(\xi_i, \eta(\xi_i)), \end{aligned} \tag{47}$$

where the function $\phi(\xi, \eta)$ has been defined in (21) and where tr_k is the trace over the Hilbert space H_k of particle k. From (47) we have

$$\beta^{-1} G_1(\{\xi_l\}) \geq \beta^{-1} \min_{\xi} \phi(\xi, \eta(\xi)) = \beta^{-1} \phi(\xi^0, \eta(\xi^0)), \tag{48}$$

where $\phi(\xi, \eta(\xi))$ assumes its absolute minimum at the point $\xi_l = \xi^0$, which obviously depends on η. So far, we have considered an arbitrary set of functions $\eta(\xi)$ and the lower bound to $\beta^{-1} G_1(\{\xi_l\})$ is valid for any choice of these functions. The best lower bound can be obtained by choosing $\eta(\xi)$ in such a way that $\phi(\xi, \eta)$ is maximal for $\eta = \eta(\xi)$. As we have seen in section 2 this choice is realized, if η is the unique solution of the molecular-field equation (25), *i.e.* $\eta \equiv \eta_{mf}(\xi)$.

For the free energy per particle we now get the inequality

$$f \geq \beta^{-1} \min_{\xi} \phi(\xi, \eta_{mf}(\xi)) - \beta^{-1} \lim_{N \to \infty} \lim_{n \to \infty} N^{-1} \ln \left[\left(\frac{\beta N}{2\pi n} \right)^{\frac{1}{2}pn} \int_{-\infty}^{\infty} \prod_l \mathrm{d}\xi_l \, \mathrm{e}^{-N(G_1(\{\xi_l\}) - G_0)} \right], \tag{49}$$

where

$$G_0 = \min_{\xi} \phi(\xi, \eta_{mf}(\xi)), \tag{50}$$

and $G_1(\{\xi_l\})$ is the real part of the function $G(\{\xi_l\})$ given by (42) in the special case that $\eta = \eta_{mf}(\xi)$. Then, assuming that the second term of eq. (49) can be neglected in the thermodynamic limit, we would have a lower bound for the free energy, which is equal to the upper bound (25), and the free energy per particle would be given by (5). However, the validity of this assumption is by no means obvious, just as in the ferromagnetic case which we have treated in ref. 1. We may anticipate that the convergence of the correction term is ensured by the fact that the function $\Phi(\xi)$, as defined in (30), assumes a minimum. An appropriate estimate of the correction term, however, must be accurate and a Hölder inequality like (44) when applied to an integral such as occurring in the second term of (47) will in general not lead to a convergent result. In fact, if we replace $G_1(\{\xi_l\})$ by the right-hand side of (47) which does not involve the antiferromagnetic interactions except through the parameters η, the integral can be estimated to be a product of n integrals and the result would be $\lim \gamma^n$ where γ is a finite constant which is, certainly for almost all temperatures, different from 1, so that the expression diverges if $n \to \infty$. In addition, we cannot proceed in the same way as in ref. 1, since the function G, *cf.* (42) contains the interaction terms in eq. (39) and starting from (42) one cannot determine in a direct way the absolute minimum of G_1 and the second derivatives.

Remark. Finally it may be noted that it is really necessary to introduce parameters $\boldsymbol{\eta}$, which are functions of $\boldsymbol{\xi}$, after applying the integral relation (37). In order to see this, let us consider a fixed value of $\boldsymbol{\eta}$. In that case we can start with the hamiltonian (3)

$$\mathscr{H} = -\tfrac{1}{2}N\boldsymbol{\eta}^2 + \sum_k T(k) + \sum_k \boldsymbol{\eta}\cdot \boldsymbol{W}(k)$$

$$- (1/2N) \sum_{f=1}^{p} \Big(\sum_k V_f(k)\Big)^2 + (1/2N) \sum_{a=1}^{q} \Big(\sum_k W_a(k) - N\eta_a\Big)^2 . \tag{51}$$

Using the same line of reasoning as above we would obtain the inequality

$$f \geq \beta^{-1} \min_{\xi} \phi\,(\boldsymbol{\xi}, \boldsymbol{\eta}) + \text{correction term}, \tag{52}$$

where the correction term can be obtained from the second term in the right-hand side of (49) by replacing $\boldsymbol{\eta}_{mf}(\boldsymbol{\xi})$ by the constant value $\boldsymbol{\eta}$. If we assume that the correction term can be neglected, we have the following inequality

$$\beta^{-1} \max_{\eta} \min_{\xi} \phi\,(\boldsymbol{\xi}, \boldsymbol{\eta}) \leq f \leq \beta^{-1} \min_{\xi} \max_{\eta} \phi\,(\boldsymbol{\xi}, \boldsymbol{\eta}). \tag{53}$$

The lower bound in eq. (53) is the best one we can obtain from (52) by choosing a constant value of $\boldsymbol{\eta}$, the upper bound is equivalent to (25) since $\max\ \phi\,(\boldsymbol{\xi}, \boldsymbol{\eta}) = \phi\,(\boldsymbol{\xi}, \boldsymbol{\eta}_{mf}(\boldsymbol{\xi}))$. Although in some simple cases such as *e.g.* a spin-$\frac{1}{2}$ Ising antiferromagnet, the upper bound and lower bound in eq. (53) can be shown to be equal, this is by no means true in general. A relatively simple example of such a situation will be treated in section 7.

4. *The absolute minimum*. In order to prove that the correction term in (49) is convergent, we must take into account the interaction terms

$$\exp\,[-\tfrac{1}{2}\beta\,(Nn)^{-1}\,(\textstyle\sum_k W_a(k) - N\eta_a\,(\boldsymbol{\xi}))^2]$$

in eq. (39). In this section we shall derive another integral representation $\int \mathrm{e}^{-Ng}$ for the partition function which is such that we can determine the absolute minimum of the real part $g_1 \equiv \mathrm{Re}\, g$ exactly. We apply the relation

$$\mathrm{e}^{-A^2} = \pi^{-\frac{1}{2}} \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2}\, \mathrm{e}^{2xiA} \tag{54}$$

to each of the operators $\exp\,[-\tfrac{1}{2}\beta\,(Nn)^{-1}\,(\sum_k W_a(k) - N\eta_a\,(\boldsymbol{\xi}))^2]$ occurring in (39). Then Z_N can be expressed in terms of the trace of the nth power of a $(p+q)$-dimensional integral $\int \mathrm{d}\boldsymbol{\xi}\, \mathrm{d}\boldsymbol{\chi} \cdots$ and by writing this nth power as a $(p+q)\,n$-

dimensional integral $\int \prod_{i=1}^{n} \mathrm{d}\xi_i \, \mathrm{d}\chi_i \ldots$, we find

$$Z_N = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}n(p+q)} \int_{\infty}^{\infty} \prod_i \mathrm{d}\xi_i \, \mathrm{d}\chi_i \exp\left[-\frac{\beta N}{2n} \sum_i (\xi_i^2 + \chi_i^2 - \eta^2(\xi_i))\right]$$

$$\times \operatorname{Tr} \prod_{i=1}^{n} \exp\left[-\frac{\beta}{n} \sum_k \mathscr{H}(k, \xi_i, \eta(\xi_i))\right] V(\chi_i, \eta(\xi_i)), \tag{55}$$

where

$$V(\chi_i, \eta(\xi_i)) = \prod_{a=1}^{q} \exp\left[(\mathrm{i}\beta\chi_{ia}/n)\left\{\sum_k W_a(k) - N\eta_a(\xi_i)\right\}\right] \tag{56}$$

and $\mathscr{H}(k, \xi_i, \eta(\xi_i))$ is given by (15).

We now use the fact that operators acting on different particles commute. This implies in particular that the trace of a product $\operatorname{Tr} \prod_k$ can be written as a product of traces tr_k over the one-particle Hilbert spaces H_k. For the partition function we obtain the integral representation

$$Z_N = \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}n(p+q)} \int_{-\infty}^{\infty} \prod_i \mathrm{d}\xi_i \, \mathrm{d}\chi_i \, \mathrm{e}^{-Ng(\{\xi_i\}, \{\chi_i\})}, \tag{57}$$

where

$$g(\{\xi_i\}, \{\chi_i\}) = \frac{\beta}{2n} \sum_{i=1}^{n} \{\xi_i^2 + \chi_i^2 - \eta^2(\xi_i)\}$$

$$- \frac{1}{N} \sum_{k=1}^{N} \ln \operatorname{tr}_k \left[\prod_{i=1}^{n} \exp\left[-\frac{\beta}{n} \mathscr{H}(k, \xi_i, \eta(\xi_i))\right] U(k, \chi_i, \eta(\xi_i))\right] \tag{58}$$

and

$$U(k, \chi_i, \eta(\xi_i)) = \prod_{a=1}^{q} \exp\left[(\mathrm{i}\beta\chi_{ia}/n)\{W_a(k) - \eta_a(\xi_i)\}\right]. \tag{59}$$

One of the advantages of the present integral representation is that we can now determine in a rigorous way the location of the absolute minimum g_0 of the function $g_1 = \operatorname{Re} g$. [This was not possible in section 3, where we derived a lower bound of $\operatorname{Re} G$, *cf.* (47).]

First it will be shown that g_1, for a fixed, but arbitrary function $\eta(\xi)$, can only assume an absolute minimum under the condition

$$\xi_i = \xi^0 \quad \text{(independent of } i\text{)},$$

$$\chi_i = 0. \tag{60}$$

The derivation of (60) will be based on the properties

a) $g_1(\{\xi_i\}, \{\chi_i\}) \geq (1/n) \sum_{i=1}^{n} \phi(\xi_i, \eta(\xi_i))$,

b) $g_1(\{\xi_i\}, \{\chi_i\}) = (1/n) \sum_{i=1}^{n} \phi(\xi_i, \eta(\xi_i))$, (61)

if and only if $\xi_i = \xi =$ independent of i, $\chi_i = 0$.

In the proof use will be made of the Hölder inequality for operators and the condition under which the equality sign holds[15]). In addition it will be assumed that the unit operator $1(k)$ in the Hilbert space H_k and the operators $V_1(k)$, $V_2(k), \ldots, V_p(k)$, $W_1(k)$, $W_2(k), \ldots, W_q(k)$ are linearly independent.

Let $A_1, A_2, \ldots, A_n$ be positive definite hermitean $m \times m$ matrices and S_1, $S_2, \ldots, S_n$ be unitary $m \times m$ matrices, then

$$|\mathrm{tr}\, A_1 S_1 A_2 S_2 \cdots A_n S_n| \leq \prod_{i=1}^{n} (\mathrm{tr}\, A_i^{\theta_i})^{1/\theta_i}, \tag{62}$$

for arbitrary positive numbers θ_i satisfying $\sum_{i=1}^{n} \theta_i^{-1} = 1$. In addition

$$|\mathrm{tr}\, A_1 A_2 \cdots A_n| = \prod_{i=1}^{n} (\mathrm{tr}\, A_i^{\theta_i})^{1/\theta_i}, \tag{63}$$

if and only if

$$A_i^{\theta_i} = \lambda_{ij} A_j^{\theta_j}, \tag{64}$$

where the λ_{ij} are constants.

From the explicit expressions (58), (59) we have, after application of (63) in the special case that $\theta_i = n$

$$g_1(\{\xi_i\}, \{\chi_i\}) \geq \frac{\beta}{2n} \sum_{i=1}^{n} \{\xi_i^2 + \chi_i^2 - \eta^2(\xi_i^2)\}$$

$$- (1/Nn) \sum_{i=1}^{n} \sum_{k=1}^{N} \ln \mathrm{tr}_k \exp[-\beta \mathscr{H}(k, \xi_i, \eta(\xi_i))]$$

$$\geq (1/n) \sum_{i=1}^{n} \phi(\xi_i, \eta(\xi_i)). \tag{65}$$

Hence eq. (61a) has been proved.

Obviously the equality sign in (65) can only apply for $\chi_i = 0$ and using (64), we have the additional condition

$$\exp[-(\beta/n) \mathscr{H}(k, \xi_i, \eta(\xi_i))] = \lambda_{ij}(k) \exp[-(\beta/n) \mathscr{H}(k, \xi_j, \eta(\xi_j))]. \tag{66}$$

Eq. (66) implies that for all k

$$(\xi_i - \xi_j) \cdot V(k) - \{\eta(\xi_i) - \eta(\xi_j)\} \cdot W(k) = c_{ij}(k), \tag{67}$$

where $c_{ij}(k)$ is a c-number for all i and j.

Since the operators $1(k), V_1(k), \ldots, V_p(k), W_1(k), \ldots, W_q(k)$ have been assumed to be linearly independent, we can conclude from (67) that eq. (61b) is satisfied.

Using the same line of reasoning as in ref. 1, *cf.* the discussion below eq. (I.40), we can conclude from (61b) that g_1 assumes an absolute minimum $g_{0\eta}$ at the point $\xi_i = \xi^0$, for $i = 1, \ldots, n$, $\chi_i = 0$, where ξ^0 is such that $\phi(\xi, \eta(\xi))$ has an absolute minimum, *i.e.*

$$g_{0\eta} = \phi(\xi^0, \eta(\xi^0)) = \min_{\xi} \phi(\xi, \eta(\xi)). \tag{68}$$

We now consider the derivatives of $g \equiv g_1 + \mathrm{i}g_2$ at the absolute minimum of g_1.

Using (22) it follows that

$$\left(\frac{\partial g}{\partial \xi_i}\right)_0 = \frac{\beta}{n}\left[\xi - \frac{1}{N}\sum_k \left\langle V(k) - W(k) \cdot \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right\rangle_{\mathscr{H}(k,\xi,\eta(\xi))} - \eta \cdot \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right],$$

$$\left(\frac{\partial g}{\partial \chi_i}\right)_0 = \frac{\mathrm{i}\beta}{n}\left[\eta(\xi) - \frac{1}{N}\sum_k \langle W(k)\rangle_{\mathscr{H}(k,\xi,\eta(\xi))}\right]. \tag{69}$$

Comparing this with (23) and a similar formula for $\partial\phi/\partial\xi$, we have

$$\left(\frac{\partial g_1}{\partial \xi_i}\right)_0 = \frac{\beta}{n}\frac{\mathrm{d}\Phi}{\mathrm{d}\xi} = \frac{\beta}{n}\left\{\frac{\partial\phi}{\partial\xi} + \frac{\partial\phi}{\partial\eta}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right\}, \qquad \left(\frac{\partial g_2}{\partial \xi_i}\right)_0 = 0, \tag{70}$$

$$\left(\frac{\partial g_1}{\partial \chi_i}\right)_0 = 0, \qquad \left(\frac{\partial g_2}{\partial \chi_i}\right)_0 = \frac{\beta}{n}\frac{\partial\phi}{\partial\eta}. \tag{71}$$

So far, we have considered an arbitrary set of real functions $\eta(\xi)$. In order to obtain the best lower bound for the free energy, apart from corrections due to second derivatives, the function $\eta(\xi)$ should be chosen in such a way that $g_{0\eta}$ has its maximal value. Using property I of the function $\phi(\xi, \eta)$, discussed in section 2, it follows that the best lower bound is obtained by taking $\eta(\xi)$ to be $\eta_{mf}(\xi)$, *i.e.* the unique solution of eq. (24). Eq. (71) shows that this choice also guarantees that $\partial g_2/\partial\chi_i = 0$ and provides together with $\partial g_2/\partial\xi_i = 0$ a stationary phase of the integrand $\mathrm{e}^{-N\phi}$ in (57) at the absolute minimum of g_1.

If the function $\phi(\xi, \eta_{mf}(\xi))$ has more than one absolute minimum, we shall assume that the number of such minima is *finite*. A continuous degeneracy can be dealt with in the same way as in ref. 1 by introducing an additional term in the hamiltonian which removes the degeneracy.

Remark. The assumption that the operators $1(k), V_1(k), \ldots, V_p(k), W_1(k), \ldots, W_q(k)$ are linearly independent is not necessary. This is obvious, if the function Φ assumes its absolute minimum at only one point ξ_0. On the other hand, if there are more points at which $\phi(\xi, \eta_{mf}(\xi))$ attains its absolute minimum, it can be shown from eqs. (67) and (31b), that for all pairs $i, j = 1, \ldots, n$, we have

$$\xi_i^2 - \xi_j^2 = \eta_{mf}^2(\xi_i) - \eta_{mf}^2(\xi_j). \tag{72}$$

From eq. (72) we may conclude that $\xi_i = \xi_j$ for all temperatures, apart from some isolated values under rather pathological conditions.

From now on the $\eta(\xi_i)$ will be fixed to be $\eta_{mf}(\xi_i)$ and we shall omit the subscript *mf* which has been used up till now to denote these particular functions.

Since it has been shown to be possible to determine exactly the absolute minimum of the function g_1, while in section 3 we could not do the same thing for G_1, we might try to evaluate the correction to the lower bound on the free energy by using the estimate

$$Z_N \leq \mathrm{e}^{-NG_0} \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}(p+q)n} \int \prod_i \mathrm{d}\xi_i \, \mathrm{d}\chi_i \, \mathrm{e}^{-N(g_1 - g_0)}, \tag{73}$$

where $g_0 \equiv G_0 = g_{0\eta}$ in the case that $\eta = \eta_{mf}(\xi)$, *cf.* (68). If we could show that the correction due to the integral in the right-hand side of (73) is negligible, we would obtain the correct lower bound on the free energy. To see to what extent this is true we can apply Laplace's method to the right-hand side of (73).

Using a similar line of reasoning as in the derivation of eq. (I.58), it can be shown that

$$F \geq N\beta^{-1} g_0 + (1/2\beta) \ln \det H, \tag{74}$$

where

$$H \equiv \begin{pmatrix} H^{(1)} & H^{(3)} \\ \tilde{H}^{(3)} & H^{(2)} \end{pmatrix}. \tag{75}$$

Here H is a $n(p+q) \times n(p+q)$ matrix consisting of matrices $H^{(1)}$, $H^{(2)}$, $H^{(3)}$, $\tilde{H}^{(3)}$, which are $np \times np$, $nq \times nq$, $np \times nq$, $nq \times np$ matrices, respectively, and which can be expressed in terms of the second derivatives of g_1 at the absolute

minimum, *viz.*

$$H^{(1)}_{ifjg} = \frac{n}{\beta}\left(\frac{\partial^2 g_1}{\partial\xi_{if}\,\partial\xi_{jg}}\right)_0, \qquad H^{(2)}_{iajb} = \frac{n}{\beta}\left(\frac{\partial^2 g_1}{\partial\chi_{ia}\,\partial\chi_{jb}}\right)_0,$$
$$H^{(3)}_{ifja} = \frac{n}{\beta}\left(\frac{\partial^2 g_1}{\partial\xi_{if}\,\partial\chi_{ja}}\right)_0, \qquad \tilde{H}^{(3)}_{iajf} = \frac{n}{\beta}\left(\frac{\partial^2 g_1}{\partial\chi_{ia}\,\partial\xi_{jf}}\right)_0. \tag{76}$$

We can now write the matrices H as expansions in inverse powers of n. The result is (note that the matrices H are the real parts of the matrices L which are calculated in appendix B)

$$H^{(1)}_{ifjg} = \delta_{ij}\left\{\delta_{fg} - \frac{\mathrm{d}\eta}{\mathrm{d}\xi_f}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi_g}\right\} + \frac{\mathrm{Re}\,A^{(1)}_{ifjg}}{n} + \frac{\mathrm{Re}\,B^{(1)}_{ifjg}}{n^2},$$
$$H^{(2)}_{iajb} = \delta_{ij}\,\delta_{ab} + \frac{\mathrm{Re}\,A^{(2)}_{iajb}}{n^2}, \tag{77}$$
$$H^{(3)}_{ifja} = \frac{\mathrm{Re}\,A^{(3)}_{ifja}}{n} + \frac{\mathrm{Re}\,B^{(3)}_{ifja}}{n^2},$$

where the matrices A and B are bounded with respect to n. Explicit expressions for the matrices Re A can be found in eqs. (94a, b), and the precise form of the matrices B is irrelevant in view of a lemma discussed by Lenard[16]), which ensures that in the limit $n \to \infty$ the matrices B do not give a contribution to det H. Then, noting that the matrices A have a cyclic structure, we can introduce much in the same way as we did in ref. 1 dynamical matrices $D^{(1)}(\varkappa)$, $D^{(2)}(\varkappa)$, $D^{(3)}(\varkappa)$, which are essentially Fourier transforms of Re $A^{(1)}$, Re $A^{(2)}$ and Re $A^{(3)}$.

Then the correction term in the right-hand side of eq. (74) can be written

$$\frac{1}{2\beta}\ln\det H = \lim_{n\to\infty}\sum_{\varkappa=0}^{n-1}\ln\det\begin{pmatrix} 1_p - (\mathrm{d}\eta/\mathrm{d}\xi)\cdot(\mathrm{d}\eta/\mathrm{d}\xi) - D^{(1)}(\varkappa) & D^{(3)}(\varkappa) \\ D^{(3)'}(\varkappa) & 1_q - D^{(2)}(\varkappa)\end{pmatrix}, \tag{78}$$

where 1_p and 1_q are $p \times p$ and $q \times q$ unit matrices and the detailed expressions of $D^{(1)}(\varkappa)$, $D^{(2)}(\varkappa)$ and $D^{(3)}(\varkappa)$ are given in eqs. (101) and (102). Since the matrices $D^{(1)}(\varkappa)$, $D^{(2)}(\varkappa)$ are of order $\mathcal{O}\,(1/\varkappa^2)$ and $D^{(3)}(\varkappa)$ is of order $\mathcal{O}\,(1/\varkappa)$ for large values of $\varkappa$, and $(\mathrm{d}\eta/\mathrm{d}\xi)\cdot(\mathrm{d}\eta/\mathrm{d}\xi)$ is independent of $\varkappa$, the right-hand side of (78) diverges as soon as η depends on ξ.

Hence we must conclude that the inequality (73) is too crude to provide us with an appropriate estimate of the correction term. It is obvious that the *imaginary*

part of the function g must be taken into account to cancel the singularities in the right-hand side of (78). In order to establish a well-behaved lower bound to the free energy, f, we must return to eq. (49) and prove that the second term in the right-hand side vanishes in the thermodynamic limit. To this end we need a detailed analysis of the function $G_1(\{\xi_i\})$, as defined in eq. (42), and the integral representation (57) will turn out to be useful. From the two representations (41) and (57) it is obvious that the function G_1 can be expressed in terms of an integral over the variables χ_i, *viz.*

$$G_1 = -N^{-1} \ln \left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}qn} \left| \int_{-\infty}^{\infty} \prod_i \mathrm{d}\chi_i \, \mathrm{e}^{-Ng(\{\xi_i\},\{\chi_i\})} \right|. \tag{79}$$

In order to apply Laplace's method to the function G_1, we should calculate the integral in the right-hand side of (79), determine the absolute minimum of G_1 and in addition calculate the matrix of second derivatives at the absolute minimum. Now there is a difficulty since it is in general not true that the function G_1 will assume its absolute minimum at the same point $\xi_i = \xi^0$ as the function g_1, although as we shall argue later on, the difference turns out to be irrelevant in the thermodynamic limit. In addition the calculation of the second derivatives would involve a qn fold integration over the variables χ_i and this integration cannot be done analytically.

In order to avoid these difficulties we use a slightly different line of reasoning. The Laplace method involves the calculation of the matrix of second derivatives of a real function F occurring in an integral like $\int \mathrm{e}^{-NF}$ at its absolute minimum F_0. Once it has been established that the second derivatives do not give rise to singularities, we know that in the free energy of the system there is a well-behaved correction which is of order 1. In addition one could also consider the contributions due to higher-order derivatives of the function F. The exponentials of these terms can be expanded and it can be argued that these higher-order derivatives will lead to corrections of order N^{-1} to the free energy.

Let us now return to the integral appearing in the right-hand side of eq. (79). The function g may be expanded around the absolute minimum of g_1, where $\xi_i = \xi^0$ and $\chi_i = 0$ and we may estimate the correction due to the different derivatives. The free energy of the system can be thought of as an asymptotic series containing a term proportional to N, a term, which is of order 1, a term which is of order N^{-1} and so on. If, as is the usual procedure in the Laplace method, we want to prove that the term, which is of order 1, is well behaved, then we can restrict ourselves to the second derivatives of the function g at the absolute minimum of g_1, since, if this is the case, we may expect that the higher-order derivatives will lead to terms of order N^{-1}. Hence, if we do not want to go into the tedious details of discussing the behaviour of terms involving N^{-1} and higher inverse powers of N, the func-

tion g in the right-hand side of eq. (79) can be replaced by

$$g = g_0 + \frac{\beta}{2n} \sum_{i,j=1}^{n} x_i \cdot L_{ij}^{(1)} \cdot x_j + \frac{\beta}{2n} \sum_{i,j=1}^{n} \chi_i \cdot L_{ij}^{(2)} \cdot \chi_j$$

$$+ \frac{\beta}{n} \sum_{i,j=1}^{n} x_i \cdot L_{ij}^{(3)} \cdot \chi_j, \tag{80}$$

where $x_i \equiv \xi_i - \xi^0$ and the matrices L are related to the second derivatives of the function g at the absolute minimum of g_1

$$L_{ifjg}^{(1)} = \frac{n}{\beta} \left(\frac{\partial^2 g}{\partial \xi_{if} \, \partial \xi_{jg}} \right)_0, \qquad L_{iajb}^{(2)} = \frac{n}{\beta} \left(\frac{\partial^2 g}{\partial \chi_{ia} \, \partial \chi_{jb}} \right)_0,$$

$$L_{ifja}^{(3)} = \frac{n}{\beta} \left(\frac{\partial^2 g}{\partial \xi_{if} \, \partial \chi_{ja}} \right)_0. \tag{81}$$

After substituting eq. (80) into the right-hand side of eq. (79) it can be shown that G_1 has its absolute minimum at the same point $\xi_i = \xi^0$ as the function g_1. In addition the matrix of second derivatives of G_1 at this point can now be calculated exactly. This will be done in section 5.

Finally, in section 6, it will be shown that the contribution to the free energy due to the second derivatives is *finite*, provided that the function $\Phi(\xi) \equiv \phi(\xi, \eta(\xi))$ has a strict minimum for $\xi = \xi^0$, *i.e.* the matrix of second derivatives of Φ is *strictly positive definite*.

5. *The second derivatives of* G_1. In this section we calculate the second derivatives of G_1 at the point $\xi_i = \xi^0$, where $g_1 = \mathrm{Re}\, g$ has its absolute minimum. As has been explained in the previous section, the second derivatives can be calculated from (79) by substituting the expansion (80). In the first place it may be noted that in view of (80) the function G_1 has an extremum at the point $\xi_i = \xi^0$, *i.e.*

$$\left(\frac{\mathrm{d}G_1}{\mathrm{d}\xi_i} \right)_0 = 0, \tag{82}$$

where the suffix 0 indicates that the derivatives are taken at the point $\xi_i = \xi^0$. In order to see this, we expand the function

$$\exp\left[-N(g - g_0)\right] \exp\left[\beta (2n)^{-1} \sum_i \chi_i^2\right]$$

as a power series in $x_i = \xi_i - \xi^0$ and χ_i. From (80), (76) and (77) it follows that the terms that are linear in x also contain a factor χ, so that we are left with an

integral of the form

$$\int e^{-\chi^2}(1 + ax\chi + bx^2 + c\chi^2)\,d\chi,$$

from which it is obvious that no terms linear in x_i will occur in the expansion of G_1 around the point $\xi_i = \xi^0$. From (79)–(81) we see that the expansion of G_1 can be written

$$G_1(\{\xi_i\}) = g_0 - N^{-1}\ln\left(\frac{\beta N}{2\pi n}\right)^{\frac{1}{2}qn}\left|\int_{-\infty}^{\infty}\prod_i d\chi_i \times \exp\left[-\frac{\beta N}{2n}\sum_{i,j}\chi_i\cdot L_{ij}^{(2)}\cdot\chi_j\right]\right| + \tfrac{1}{2}\sum_{i,j}x_i\cdot\left(\frac{d^2G_1}{d\xi_i\,d\xi_j}\right)_0\cdot x_j. \quad (83)$$

At the point $\xi_i = \xi^0$, the function $G_1(\{\xi_i\})$ is given by the first two terms of (83). Apart from the term g_0, there is a correction term containing an integration over all χ_i. This correction term, however, vanishes in the thermodynamic limit; the convergence of the integrals is implied by the fact that all eigenvalues of the $qn \times qn$ matrix $\mathrm{Re}\,L^{(2)}$ are larger than 1, *cf.* the remark below eq. (110).

The second derivatives are given by

$$\left(\frac{d^2G_1}{d\xi_i\,d\xi_j}\right)_0 = -N^{-1}\,\mathrm{Re}\left\{\frac{\dfrac{d^2}{d\xi_i\,d\xi_j}\displaystyle\int\prod_i d\chi_i\,e^{-N\eta(\{\xi_i\},\{\chi_i\})}}{\displaystyle\int\prod_i d\chi_i\,e^{-N\eta}}\right\} \quad (84)$$

and

$$\frac{\partial^2 e^{-N\eta}}{\partial\xi_i\,\partial\xi_j} = -N\,e^{-N\eta}\left(\frac{\partial^2 g}{\partial\xi_i\,\partial\xi_j} - N\frac{\partial g}{\partial\xi_i}\frac{\partial g}{\partial\xi_j}\right), \quad (85)$$

where the derivatives ∂ should be taken at constant values of χ_i, but the function $\eta(\xi)$ must be taken into account. From (80), (81), (84) and (85) it follows that

$$\left(\frac{d^2G_1}{d\xi_i\,d\xi_j}\right)_0 = \beta n^{-1}\,\mathrm{Re}\,L_{ij}^{(1)} - \beta n^{-1}\,\mathrm{Re}\sum_{k,l=1}^{n}L_{ik}^{(3)}\cdot\langle\langle\chi_k\chi_l\rangle\rangle\cdot\bar{L}_{jl}^{(3)}, \quad (86)$$

where

$$\langle\langle\chi_k\chi_l\rangle\rangle \equiv \beta N n^{-1}\,\frac{\displaystyle\int\prod_i d\chi_i\,e^{-N\gamma}\,\chi_k\chi_l}{\displaystyle\int\prod_i d\chi_i\,e^{-N\gamma}}, \quad (87)$$

and γ is a shorthand notation for

$$\gamma(\{\chi_i\}) = g(\{\xi_i = \xi^0\},\{\chi_i\}) - g_0 = \tfrac{1}{2}\beta n^{-1}\sum_{i,j}\chi_i\cdot L_{ij}^{(2)}\cdot\chi_j. \quad (88)$$

$\tilde{L}^{(3)}_{jl}$ is a $q \times p$ matrix with elements $(\tilde{L}^{(3)}_{jl})_{af} \equiv (L^{(3)}_{jl})_{fa} = L^{(3)}_{jfla}$. In order to evaluate the right-hand side of eq. (86) we need expressions for the matrices $L^{(1)}$, $L^{(2)}$ and $L^{(3)}$. These matrices have a cyclic structure (a property which originates from the cyclic invariance of the trace) in the indices i and j:

$$L^{(\alpha)}_{ij} = L^{(\alpha)}_{j-i}, \qquad L^{(\alpha)}_{-j} = L^{(\alpha)}_{n-j}, \qquad \alpha = 1, 2, 3. \tag{89}$$

The explicit expressions for the matrices $L^{(\alpha)}$ are derived in appendix B, *cf.* eqs. (B.10), (B.11) and (B.13). These formulae are rather complicated and there is no need to give them here. Since the limit $n \to \infty$ must be taken before the thermodynamic limit, we can restrict ourselves to the expansions

$$L^{(1)}_{ij} = \delta_{ij} \left\{ \mathbf{1}_p - \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}\xi} \cdot \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}\xi} \right\} + \frac{A^{(1)}_{ij}}{n} + \frac{B^{(1)}_{ij}}{n^2}, \tag{90a}$$

$$L^{(2)}_{ij} = \delta_{ij} \mathbf{1}_q + \frac{A^{(2)}_{ij}}{n}, \tag{90b}$$

$$L^{(3)}_{ij} = \frac{A^{(3)}_{ij}}{n} + \frac{B^{(3)}_{ij}}{n^2} + \mathrm{i}\delta_{ij} \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}\xi}. \tag{90c}$$

Here $\mathbf{1}_p$ and $\mathbf{1}_q$ denote the p and q dimensional unit matrices and all the elements of the matrices $A^{(\alpha)}$ and $B^{(\alpha)}$ are bounded with respect to n. In the same way as in ref. 1 the application of a lemma given by Lenard[16]) ensures that the matrices $B^{(\alpha)}$ do not give a contribution to the determinant of the matrix of the second derivatives of G_1 at the point $\xi_l = \xi^0$ in the limit $n \to \infty$.

The explicit expressions for the matrices $A^{(\alpha)}$ can be found from (B.10), (B.11) and (B.13). By decomposing these equations in a real and an imaginary part it follows that Re $A^{(1)}_{ij}$, Re $A^{(2)}_{ij}$ and Im $A^{(3)}_{ij}$ are given by

$$-\beta N^{-1} \sum_k T_k^{-1} \sum_{p,q} X^{(1,2,3)}(p, q)\, \mathrm{e}^{-\frac{1}{2}\beta(h_p + h_q)} \cosh \gamma_{j-i}, \tag{91a}$$

and Im $A^{(1)}_{ij}$, Im $A^{(2)}_{ij}$ and Re $A^{(3)}_{ij}$ are given by

$$\mathrm{i}\beta N^{-1} \sum_k T_k^{-1} \sum_{p,q} X^{(1,2,3)}(p, q)\, \mathrm{e}^{-\frac{1}{2}\beta(h_p + h_q)} \sinh \gamma_{j-i}. \tag{91b}$$

Here the summation over k is a summation over the particles $k = 1, \ldots, N$ and T_k has been defined in eq. (7); p and q denote the eigenstates of the one-particle hamiltonian $\mathscr{H}(k, \xi^0, \boldsymbol{\eta}(\xi^0))$, as given in eq. (15), h_p and h_q are the corresponding eigenvalues and O_{pq} is the matrix element of an operator O between eigenstates p and q. Of course, p, q, h_p, h_q and O_{pq} depend on k, but for the sake of notation the explicit k-dependence has been omitted in eq. (91) and the following formulae.

Furthermore, *cf.* also (33),

$$X^{(1)}(p, q) = \tilde{V}_{pq}\tilde{V}_{qp}, \tag{92a}$$

$$X^{(2)}(p, q) = -\tilde{W}_{pq}\tilde{W}_{qp}, \tag{92b}$$

$$X^{(3)}(p, q) = \tilde{V}_{pq}\tilde{W}_{qp}. \tag{92c}$$

Finally γ_j is given by

$$\gamma_j = \beta\left(\frac{1}{2} - \frac{|j|}{n}\right)(h_p - h_q)\frac{j}{|j|}, \quad \text{if} \quad j \neq 0,$$
$$\gamma_0 = \tfrac{1}{2}\beta(h_p - h_q). \tag{93}$$

From eq. (91) and the trivial relation $\gamma_{-j} = \gamma_{n-j}$ it is obvious that the matrices $A_{ij}^{(\alpha)}$ are cyclic with respect to the indices i and j; *i.e.* for every pair of indices (f, g), (a, b), (f, a), we have

$$A^{(1)}_{ifjg} \equiv A^{(1)}_{fg,\,j-i}; \qquad A^{(2)}_{iajb} \equiv A^{(2)}_{ab,\,j-i}; \qquad A^{(3)}_{ifja} \equiv A^{(3)}_{fa,\,j-i}$$
$$A^{(1)}_{fg,\,-j} = A^{(1)}_{fg,\,n-j}; \qquad A^{(2)}_{ab,\,-j} = A^{(2)}_{ab,\,n-j}; \qquad A^{(3)}_{fa,\,-j} = A^{(3)}_{fa,\,n-j}. \tag{94}$$

In addition, it is shown in appendix C that the matrix $\langle\langle\chi_i\chi_j\rangle\rangle$ which occurs in eq. (86) has also a cyclic *structure*, *i.e.* for every pair of indices a, b, we have

$$\langle\langle\chi_{ia}\chi_{jb}\rangle\rangle \equiv \chi_{ab,\,j-i}; \qquad \chi_{ab,\,-j} = \chi_{ab,\,n-j}. \tag{95}$$

Now all $n \times n$ matrices O_{ij} occurring in eqs. (94) and (95) for fixed values of f, g, a and b can be diagonalized by the same unitary transformation, *cf.* appendix A of ref. 17, *viz.*

$$(U^{\dagger}OU)_{lm} = \lambda_l\delta_{lm}, \qquad \text{for} \qquad l, m = 1, \ldots, n, \tag{96}$$

where

$$U_{lm} = n^{-\frac{1}{2}}\,\mathrm{e}^{2\pi\mathrm{i}lm/n}, \tag{97}$$

and

$$\lambda_l = \sum_{j=0}^{n-1} O_j\,\mathrm{e}^{2\pi\mathrm{i}jl/n}.$$

Using eq. (86), the expansions (90) and the cyclic properties (94) and (95), it is obvious that the matrix

$$H_{i\varrho j\varrho} \equiv \frac{n}{\beta} \frac{\mathrm{d}^2 G_1}{\mathrm{d}\xi_{i\varrho}\, \mathrm{d}\xi_{j\varrho}} \tag{98}$$

has a cyclic structure with respect to i and j, since the product of a number of cyclic matrices is again a cyclic matrix, *i.e.*

$$H_{i\varrho j\varrho} \equiv H_{\varrho, j-i}; \qquad H_{\varrho, -j} = H_{\varrho, n-j}. \tag{99}$$

Then the unitary transformation determined by the $pn \times pn$ matrix

$$V_{i\varrho j\varrho} = \delta_{\varrho\varrho} U_{ij}, \tag{100}$$

where U_{ij} is given by (97) transforms H into a direct sum consisting of n $p \times p$ matrices; we write

$$V^{\dagger} H V = \bigoplus_{\varkappa=0}^{n-1} \left\{ 1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi} \cdot \frac{\mathrm{d}\eta}{\mathrm{d}\xi} - D(\varkappa) \right\}, \tag{101}$$

where

$$D(\varkappa) = \sum_{j=0}^{n-1} \left\{ \delta_{j,0} \left(1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi} \cdot \frac{\mathrm{d}\eta}{\mathrm{d}\xi} \right) - H_j \right\} \mathrm{e}^{2\pi \mathrm{i}\, j\varkappa/n}. \tag{102}$$

We shall show that the matrix $D(\varkappa)$ can be expressed in terms of the matrices $D^{(\alpha)}(\varkappa)$ and $E^{(\alpha)}(\varkappa)$ defined by

$$D^{(\alpha)}(\varkappa) = -n^{-1} \sum_{j=0}^{n-1} (\mathrm{Re}\, A_j^{(\alpha)})\, \mathrm{e}^{2\pi \mathrm{i}\, j\varkappa/n}, \tag{103}$$

$$E^{(\alpha)}(\varkappa) = -\mathrm{i} n^{-1} \sum_{j=0}^{n-1} (\mathrm{Im}\, A_j^{(\alpha)})\, \mathrm{e}^{2\pi \mathrm{i}\, j\varkappa/n}, \tag{104}$$

for $\alpha = 1, 2, 3$ and $\varkappa = 0, 1, \ldots, n-1$. From (86), (90a), (98) and (102) we have

$$D(\varkappa) = D^{(1)}(\varkappa) + D^{(33)}(\varkappa), \tag{105}$$

where

$$D^{(33)}(\varkappa) = \sum_{j,k,l=0}^{n-1} \mathrm{Re}\, (L_k^{(3)} \cdot \langle\langle \chi_k \chi_{l+j} \rangle\rangle \cdot \tilde{L}_l^{(3)})\, \mathrm{e}^{2\pi \mathrm{i}\, j\varkappa/n}. \tag{106}$$

In (106) use has been made of the cyclic properties of the matrices $L^{(3)}$ and $\langle\langle\chi_k\chi_l\rangle\rangle$. The matrix $\langle\langle\chi_k\chi_{l+j}\rangle\rangle$ is given by

$$\langle\langle\chi_k\chi_{l+j}\rangle\rangle = n^{-1}\sum_{\varkappa=0}^{n-1} N(\varkappa)\exp\left[2\pi\mathrm{i}\,(l+j-k)\,\varkappa/n\right], \tag{107}$$

cf. eqs. (C.2) and (C.3) in appendix C.

The matrix $N(\varkappa)$ can be expressed in terms of the matrices $D^{(2)}(\varkappa)$ and $E^{(2)}(\varkappa)$. The calculation is tedious and is given in appendix C. We find

$$N(\varkappa) = [1 - D^{(2)}(\varkappa) + E^{(2)}(\varkappa)\,\{1 - D^{(2)}(\varkappa)\}^{-1}\,E^{(2)}(\varkappa)^{\dagger}]^{-1}$$
$$\times\,[1 + E^{(2)}(\varkappa)\,\{1 - D^{(2)}(\varkappa)\}^{-1}]. \tag{108}$$

By using the expansion (90c) and the inverse relations to (103) and (104), one calculates in a straightforward way the matrix $D^{(33)}(\varkappa)$. The details will be given in appendix D, and one arrives at

$$\begin{aligned} D^{(33)}(\varkappa) = {} & \tfrac{1}{2}D^{(3)}(\varkappa)\,\{N(\varkappa) + N^{\dagger}(\varkappa)\}\,D^{(3)\dagger}(\varkappa) \\ & - \tfrac{1}{2}\left(E^{(3)}(\varkappa) - \mathrm{i}\,\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right)(N(\varkappa) + N^{\dagger}(\varkappa))\left(E^{(3)}(\varkappa) - \mathrm{i}\,\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right)^{\dagger} \\ & + \tfrac{1}{2}D^{(3)}(\varkappa)\,(N(\varkappa) - N^{\dagger}(\varkappa))\left(E^{(3)}(\varkappa) - \mathrm{i}\,\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right)^{\dagger} \\ & - \tfrac{1}{2}\left(E^{(3)}(\varkappa) - \mathrm{i}\,\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right)(N(\varkappa) - N^{\dagger}(\varkappa))\,D^{(3)\dagger}(\varkappa). \end{aligned} \tag{109}$$

So far we have expressed all second derivatives (98) in terms of the matrices $D^{(\alpha)}$ and $E^{(\alpha)}$, as defined in (103) and (104), *cf.* (101), (105), (109) and (108).

The explicit expressions for $D^{(\alpha)}$ and $E^{(\alpha)}$ can be evaluated easily if one uses eqs. (91). It follows that in the limit $n \to \infty$, $D^{(1)}(\varkappa)$, $D^{(2)}(\varkappa)$ and $-\mathrm{i}E^{(3)}(\varkappa)$ are given by

$$\tfrac{1}{2}\beta N^{-1}\sum_{k} T_k^{-1}\sum_{p,q} X^{(1,2,3)}(p,q)\,\mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)}\sinh\tfrac{1}{2}\beta\,(h_p - h_q)$$
$$\times\,\frac{4\beta\,(h_p - h_q)}{4n^2\sin^2(\pi\varkappa/n) + \beta^2\,(h_p - h_q)^2}, \tag{110}$$

where X has been defined in (92). Note that because of the minus sign in the definition of $X^{(2)}$, the matrix $D^{(2)}(\varkappa)$ is negative definite so that all the eigenvalues of $\mathrm{Re}\,L^{(2)}$ are larger than 1. Furthermore $-\mathrm{i}E^{(1)}(\varkappa)$, $-\mathrm{i}E^{(2)}(\varkappa)$ and $D^{(3)}(\varkappa)$ are

given by

$$\tfrac{1}{2}\beta N^{-1} \sum_k T_k^{-1} \sum_{p,q} X^{(1,2,3)}(p,q)\, \mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)} \sinh \tfrac{1}{2}\beta(h_p - h_q)$$

$$\times \frac{4n \sin 2\pi\varkappa n^{-1}}{4n^2 \sin^2(\pi\varkappa/n) + \beta^2 (h_p - h_q)^2}. \tag{111}$$

From (110) and (111) one obtains a number of symmetry relations for the matrices $D^{(\alpha)}$ and $E^{(\alpha)}$, *viz.*

$$D^{(1)}(\varkappa) = D^{(1)}(-\varkappa) = D^{(1)}(\varkappa)^*, \qquad D^{(1)}(\varkappa) = \tilde{D}^{(1)}(\varkappa); \tag{112a}$$

$$D^{(2)}(\varkappa) = D^{(2)}(-\varkappa) = D^{(2)}(\varkappa)^*, \quad E^{(2)}(\varkappa) = -E^{(2)}(-\varkappa) = E^{(2)*}(\varkappa), \qquad D^{(2)}(\varkappa) = \tilde{D}^{(2)}(\varkappa), \quad E^{(2)}(\varkappa) = -\tilde{E}^{(2)}(\varkappa); \tag{112b}$$

$$D^{(3)}(\varkappa) = -D^{(3)}(-\varkappa) = -D^{(3)}(\varkappa)^*; \qquad E^{(3)}(\varkappa) = E^{(3)}(-\varkappa) = -E^{(3)*}(\varkappa). \tag{112c}$$

Here $-\varkappa$ is meant to be $n - \varkappa$ if $\varkappa$ is restricted to the values $\varkappa = 0, \ldots, n-1$. For the matrix $N(\varkappa)$ we have, *cf.* eqs. (C.1) and (C.3) of appendix C:

$$N(\varkappa) = \tilde{N}(-\varkappa) = N^*(\varkappa). \tag{113}$$

6. *The correction due to the second derivatives.* From (98) and (101) it follows in complete analogy to ref. 1, that eq. (49) can be replaced by

$$f \geq \beta^{-1} \min_{\xi} \phi(\xi, \eta_{mf}(\xi)) + N^{-1} \lim_{n\to\infty} \tfrac{1}{2}\beta^{-1} \sum_{\varkappa=0}^{n-1} \ln\det\left(1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - D(\varkappa)\right), \tag{114}$$

where the matrix $D(\varkappa)$ is given by (102). In order that the correction term is negligible in the thermodynamic limit, the following two properties should hold:
(i) all matrices $1_p - (\mathrm{d}\eta/\mathrm{d}\xi)\cdot(\mathrm{d}\eta/\mathrm{d}\xi) - D(\varkappa)$ are positive definite;
(ii) the behaviour of $D(\varkappa)$ for large $\varkappa$ is such that the sum converges.
The first property is of course equivalent to saying that the function G_1 really has a minimum at the point $\xi_i = \xi^0$. In this section we shall first show that this is the case, using the fact that the function $\Phi(\xi) \equiv \phi(\xi, \eta_{mf}(\xi))$ has an absolute minimum at the point $\xi = \xi^0$.

In order to prove property (i), we consider the matrices

$$M_1 \equiv D^{(3)}(\varkappa)\,(1 - D^{(2)}(\varkappa))^{-1}\,D^{(3)\dagger}(\varkappa) - D^{(33)}(\varkappa) \tag{115a}$$

and

$$M_2 \equiv D^{(1)}(0) - D^{(1)}(\varkappa) - D^{(3)}(\varkappa)\,(1 - D^{(2)}(\varkappa))^{-1}\,D^{(3)\dagger}(\varkappa). \tag{115b}$$

The proof rests on the fact that these two matrices are positive definite, *i.e.*

$$M_1 \geq 0, \qquad M_2 \geq 0. \tag{116}$$

This will be established in appendix E.

From (115) and (116) it follows immediately that

$$D^{(1)}(\varkappa) - D^{(1)}(0) + D^{(33)}(\varkappa) \leq 0, \tag{117}$$

so that, by using (105), we have

$$\begin{aligned} 1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - D(\varkappa) &= 1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - D^{(1)}(\varkappa) - D^{(33)}(\varkappa) \\ &\geq 1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - D^{(1)}(0). \end{aligned} \tag{118}$$

On the other hand we can evaluate eq. (34) for the second derivatives $\mathrm{d}^2\Phi/\mathrm{d}\xi\,\mathrm{d}\xi$ in a representation in which $\mathscr{H}\,(k, \xi^0, \eta(\xi^0))$ is diagonal. Comparing the result with the explicit expression (110) for $D^{(1)}(\varkappa)$ if $\varkappa = 0$, we find

$$\beta^{-1}\,\mathrm{d}^2\Phi/\mathrm{d}\xi\,\mathrm{d}\xi = 1_p - D^{(1)}(0) - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi}. \tag{119}$$

Hence the inequality (118) can be written as

$$1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - D(\varkappa) \geq \beta^{-1}\,\frac{\mathrm{d}^2\Phi}{\mathrm{d}\xi\,\mathrm{d}\xi}. \tag{120}$$

So, indeed, if Φ has a strict minimum, all determinants occurring in the right-hand side of (114) are positive and cannot give rise to singularities if we take the logarithm.

We finally must investigate the behaviour of the matrices $1_p - (\mathrm{d}\eta/\mathrm{d}\xi)\cdot(\mathrm{d}\eta/\mathrm{d}\xi) - D(\varkappa)$ for large values of $\varkappa$. First of all we note that

$$D(\varkappa) = D\,(n - \varkappa), \tag{121}$$

so that in the summation over $\varkappa$ we can restrict ourselves to terms with $\varkappa \leq \frac{1}{2}n$. Eq. (121) is obvious if one considers the expression for $D(\varkappa)$ in terms of the matrices $D^{(\alpha)}(\varkappa)$ and $E^{(\alpha)}(\varkappa)$, *cf.* (105), (109), (108), and uses the relations

$$D^{(1)}(\varkappa) = D^{(1)}(n-\varkappa), \qquad D^{(2)}(\varkappa) = D^{(2)}(n-\varkappa),$$

$$E^{(3)}(\varkappa) = E^{(3)}(n-\varkappa), \qquad E^{(1)}(\varkappa) = -E^{(1)}(n-\varkappa), \tag{122}$$

$$E^{(2)}(\varkappa) = -E^{(2)}(n-\varkappa), \qquad D^{(3)}(\varkappa) = -D^{(3)}(n-\varkappa).$$

These relations in turn are a special case of the relations (112). Using the inequality

$$\frac{2}{\pi} \leq \frac{\sin x}{x} \leq 1 \quad \text{on} \quad \left[0, \frac{\pi}{2}\right], \tag{123}$$

we see that in the limit $\varkappa \to \infty$, (for $\varkappa \leq \frac{1}{2}n$)

$$D^{(1)}(\varkappa),\ D^{(2)}(\varkappa) \text{ and } E^{(3)}(\varkappa) \text{ are of order } \mathcal{O}(1/\varkappa^2) \tag{124a}$$

whereas

$$E^{(1)}(\varkappa),\ E^{(2)}(\varkappa) \text{ and } D^{(3)}(\varkappa) \text{ are of order } \mathcal{O}(1/\varkappa). \tag{124b}$$

This implies that, *cf.* (113)

$$N(\varkappa) = 1 + E^{(2)}(\varkappa) + \mathcal{O}(1/\varkappa^2), \tag{125}$$

and since $E^{(2)}(\varkappa)$ is antihermitean, we have

$$\begin{aligned} N(\varkappa) + N^{\dagger}(\varkappa) &= 1 + \mathcal{O}(1/\varkappa^2), \\ N(\varkappa) - N^{\dagger}(\varkappa) &= \mathcal{O}(1/\varkappa). \end{aligned} \tag{126}$$

Noting that $E^{(3)}(\varkappa)$ is hermitean we then find from (109)

$$D^{(33)}(\varkappa) = -\frac{d\eta}{d\xi} \cdot \frac{d\eta}{d\xi} + \mathcal{O}\left(\frac{1}{\varkappa^2}\right), \tag{127}$$

so that

$$1_p - \frac{d\eta}{d\xi} \cdot \frac{d\eta}{d\xi} - D(\varkappa) = 1_p - \mathcal{O}\left(\frac{1}{\varkappa^2}\right). \tag{128}$$

The term $-(d\eta/d\xi) \cdot (d\eta/d\xi)$ in the left-hand side is cancelled by the same term in $D^{(33)}(\varkappa)$ and this last term arises from the *imaginary* part of $L^{(3)}$, hence from

the second derivatives of the imaginary part of the function g at the point $\xi_l = \xi^0$, $\chi_l = 0$.

Now the convergence of a series like $\sum_\varkappa \ln(1 + 1/\varkappa^2)$ ensures that the sum over $\varkappa$ in the second term of (114) is finite, so that the correction term can be neglected in the thermodynamic limit.

Hence it has been proved that the free energy per particle cannot be smaller than $\beta^{-1} \min_\xi \Phi(\xi)$, which value is equal to the upper bound (30).

7. *Parameters independent of* ξ. We have seen already in section 4 that the difficulties in connection with the correction due to the second derivatives stem from the fact that the parameters η were chosen to be ξ-dependent, *cf.* the discussion below (78).

In this section we shall give an example that we do not get the correct lower bound if the parameters are taken to be independent of ξ. In that case, of course, all the properties of the absolute minimum of g_1, as derived in section 4, remain valid. For fixed parameters η the value of the minimum of g_1 is $g_{0\eta} = \min_\xi \phi(\xi, \eta)$, *cf.* (68). Since the matrices $D^{(1)}(\varkappa)$, $D^{(2)}(\varkappa)$, $D^{(3)}(\varkappa)$ are of order $\mathcal{O}(1/\varkappa^2)$, $\mathcal{O}(1/\varkappa^2)$ and $\mathcal{O}(1/\varkappa)$, respectively *cf.* (124), the correction (78) due to the second derivatives is finite and we have the following lower bound for the free energy

$$f \geq \beta^{-1} \min_\xi \phi(\xi, \eta) \qquad \text{for all } \eta. \tag{129}$$

Combined with the upper bound derived in section 2, this results in

$$\max_\eta \min_\xi \phi(\xi, \eta) \leq \beta f \leq \min_\xi \max_\eta \phi(\xi, \eta). \tag{130}$$

Note that for the lower bound we have

$$\max_\eta \min_\xi \phi(\xi, \eta) \geq \min_\xi \phi(\xi, 0) = \beta f_{fm}, \tag{131}$$

where f_{fm} is the free energy derived in ref. 1 for the corresponding pure ferromagnetic case without operators $W(k)$. Hence we see that the introduction of antiferromagnetic operators never can lower the free energy; *i.e.* we always have

$$f \geq f_{fm}.$$

In order to see that the lower bound in (130) can actually be lower than the upper bound, we shall discuss the following model.

Consider a spin system consisting of two sublattices of N spins each. The spin variables on one lattice are given by $S^z(k)$, $k = 1, \ldots, N$, $S^z(k) = \pm 1$, and those on the other by $T^z(k)$, $k = 1, \ldots, N$, $T^z(k) = 1, 0, -1$. The hamiltonian is given by

$$\mathscr{H} = \frac{J}{2N} \sum_{k,l=1}^{N} S^z(k)\, T^z(l). \tag{132}$$

We assume that the coupling between the two lattices is antiferromagnetic, *i.e.* $J > 0$. The hamiltonian can be written as:

$$\mathscr{H} = \frac{J}{8N}\left[\sum_k S^z(k) + \sum_k T^z(k)\right]^2 - \frac{J}{8N}\left[\sum_k S^z(k) - \sum_k T^z(k)\right]^2, \tag{133}$$

so we have one "ferromagnetic" and one "antiferromagnetic" operator. This hamiltonian has the form (3) with $T(k) = 0$ and

$$\begin{aligned} W(k) &= \tfrac{1}{2}J^{\frac{1}{2}}\,(S^z(k) + T^z(k)), \\ V(k) &= \tfrac{1}{2}J^{\frac{1}{2}}\,(S^z(k) - T^z(k)). \end{aligned} \tag{134}$$

The operators $V(k)$ and $W(k)$ act on the six-dimensional product space of a particle with spin $\frac{1}{2}$ and one with spin 1. The function $\phi\,(\xi, \eta)$, *cf.* (21), is given by

$$\phi\,(\xi, \eta) = \tfrac{1}{2}\beta\,(\xi^2 - \eta^2) - \ln \operatorname{tr} \exp \beta\,(\xi V - \eta W), \tag{135}$$

where the trace is over the six-dimensional space. Using the notation

$$\langle A \rangle = \frac{\operatorname{tr}\{A \exp[\beta\,(\xi V - \eta W)]\}}{\operatorname{tr} \exp[\beta\,(\xi V - \eta W)]}, \tag{136}$$

the molecular-field equations are

$$\xi = \langle V \rangle \qquad \text{and} \qquad \eta = \langle W \rangle. \tag{137}$$

Similarly one has

$$\begin{aligned} \partial^2\phi/\partial\xi^2 &= \beta - \beta^2\,\{\langle V^2 \rangle - \langle V \rangle^2\}, \\ \partial^2\phi/\partial\eta^2 &= -\beta - \beta^2\,\{\langle W^2 \rangle - \langle W \rangle^2\}, \\ \partial^2\phi/\partial\xi\,\partial\eta &= \beta^2\,\{\langle VW \rangle - \langle V \rangle\,\langle W \rangle\}. \end{aligned} \tag{138}$$

We introduce new variables, v, w, x and y by

$$V = \tfrac{1}{2}J^{\frac{1}{2}}v, \qquad W = \tfrac{1}{2}J^{\frac{1}{2}}w, \qquad x = \tfrac{1}{2}\beta J^{\frac{1}{2}}\xi, \qquad y = \tfrac{1}{2}\beta J^{\frac{1}{2}}\eta. \tag{139}$$

Also $b \equiv (J/4)\,\beta$. The function ϕ, expressed in these variables, will again be denoted by $\phi\,(x, y)$ and reads

$$\phi\,(x, y) = (1/2b)\,(x^2 - y^2) - \ln \operatorname{tr} e^{xv - yw}. \tag{140}$$

Eqs. (137) and (140) can be written as

$$x + y = 2b \tanh (x - y),$$
$$x - y = 4b \frac{\sinh (x + y)}{1 + 2 \cosh (x + y)} \tag{141}$$

and

$$\phi (x, y) = \tfrac{1}{2} b^{-1} (x^2 - y^2) - \ln 2 \{\cosh 2x + \cosh 2y + \cosh (x - y)\}. \tag{142}$$

From (141) one sees that apart from the solution $x = y = 0$ which, of course, exists for all b (*i.e.* for all temperatures), there exists a solution $(x_1, y_1) \neq (0, 0)$ for $b > \frac{1}{4}\sqrt{6}$. In addition $(-x_1, -y_1)$ is a solution.

If one calculates the second derivatives for the solution $x = y = 0$ (or $\xi = \eta = 0$) one finds that $\partial^2\phi/\partial x^2$ (or $\partial^2\phi/\partial\xi^2$), which is positive for small b (high temperatures) becomes zero at $b = \frac{3}{5}$.

Furthermore $\mathrm{d}^2\phi/\mathrm{d}x^2$ which is larger than $\partial^2\phi/\partial x^2$ becomes zero at $b = \frac{1}{4}\sqrt{6}$. Here $\mathrm{d}^2\phi/\mathrm{d}x^2$ is the derivative which takes into account the variation of the function $y = y_{mf}(x)$ determined by $\partial\phi/\partial y = 0$. In this discussion we shall need in particular the function

$$\phi (x, 0) = \tfrac{1}{2} b^{-1} x^2 - \ln 2 (\cosh 2x + \cosh x + 1). \tag{143}$$

In order to determine the minimum of this function we write down

$$0 = \frac{\partial\phi (x, 0)}{\partial x} = b^{-1} x - \frac{\sinh 2x + 2 \sinh 2x}{1 + \cosh x + \cosh 2x}. \tag{144}$$

From eq. (144) it can be seen that for $b < \frac{3}{5}$ only the solution $x = 0$ exists, whereas for $b > \frac{3}{5}$ apart from $x = 0$, other solutions, which we denote by $\pm x_0 \neq 0$, exist. Since $\phi (x, 0) \to \infty$ if $x \to \pm\infty$, for $b < \frac{3}{5}$, $\phi (x, 0)$ has a minimum in $x = 0$, whereas for $b > \frac{3}{5}$ $\phi (x, 0)$ has a maximum in $x = 0$ and minima in $\pm x_0$.

Consider now first temperatures larger than $5J/12k_B$ (*i.e.* $b < \frac{3}{5}$). There is only one solution, $x = y = 0$, of the molecular-field equations. The upper bound is $\phi (0, 0)$ whereas the lower bound is $\max_y \min_x \phi (x, y)$. By taking $x = 0$ we see from our discussion above that $\min_x \phi (x, 0) = \phi (0, 0)$. Hence the bounds are equal, so that the free energy is given by $\beta^{-1}\phi (0, 0)$.

If we now take $\frac{3}{5} < b < \frac{1}{4}\sqrt{6}$, the only solution of (141) is still $x = y = 0$, so that the upper bound is $\phi (0, 0)$. In order to calculate the lower bound we note the following three relations for $\phi (x, y)$:

$$\phi (x, y) = \phi (-x, -y), \tag{145a}$$

$$\phi (|x|, y) > \phi (-|x|, y) \qquad \text{if } y \geq 0 \text{ and } x \neq 0, \tag{145b}$$

$$\phi (-|x|, y) \text{ is a monotonically decreasing function of } y \text{ for } y \geq 0. \tag{145c}$$

Eqs. (145a, b) are obvious from (142). Eq. (145c) follows directly from the fact that

$$\frac{\partial \phi(-|x|, y)}{\partial y} = -\frac{y}{b} - \frac{2 \sinh 2y + \sinh(|x| + y)}{\cosh 2|x| + \cosh(|x| - y) + \cosh 2y} < 0.$$

Because of (145a) we may restrict ourselves to $y \geq 0$. Since $\phi(x, y) \to \infty$ if $x \to \pm\infty$ for fixed y, $\phi(x, y)$ has at least one minimum which we denote by $x_1(y)$. On behalf of (145b) the absolute minimum must occur for $x_1(y) \leq 0$. Then

$$\phi(x_1(y), y) \equiv \min_x \phi(x, y) \leq \phi(-x_0, y) \leq \phi(-x_0, 0)$$

$$= \min_x \phi(x, 0) < \phi(0, 0), \tag{146}$$

where (145c) has been used and where in $-x_0$ the function $\phi(x, 0)$, defined by (143) assumes its absolute minimum.

Eq. (146) shows that the lower bound to βf is equal to $\phi(-x_0, 0)$ and is lower than the upper bound which is $\phi(0, 0)$.

To give a numerical example for this difference we consider $b = 0.605$. Then the lower bound to βf is -1.79187, whereas the upper bound is -1.79176. (Note that the general proof of this paper shows that the free energy in this region is given by $\beta^{-1}\phi(0, 0)$.)

Acknowledgements. The authors are much indebted to Professor P.W. Kasteleyn for many interesting discussions. This investigation is part of the research programme of the "Stichting voor Fundamenteel Onderzoek der Materie (F.O.M.)", which is financially supported by the "Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek (Z.W.O.)".

APPENDIX A

In this appendix we give a justification for the replacement (40), *i.e.* in the calculation of the free energy per particle, we may use the expression

$$\varrho_1 \equiv \mathrm{Tr}\left\{\int_{-\infty}^{\infty} d\xi \exp\left[-\frac{\beta N}{2n}(\xi^2 - \eta^2)\right] \exp\left[-\frac{\beta}{n}\sum_k \mathscr{H}(k, \xi, \eta)\right]\right.$$

$$\left.\times \prod_{a=1}^{q} \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a\right)^2\right]\right\}^n, \tag{A.1}$$

instead of

$$\varrho_2 = \mathrm{Tr}\left\{\int_{-\infty}^{\infty} \mathrm{d}\xi \exp\left[-\frac{\beta N}{2n}(\xi^2 - \eta^2)\right] \exp\left[-\frac{\beta}{n}\sum_k T(k)\right]\right.$$

$$\times \prod_f \exp\left[\frac{\beta}{n}\xi_f \sum_k V_f(k)\right] \prod_a \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a(k) - N\eta_a\right)^2\right]$$

$$\left.\times \exp\left[-\frac{\beta}{n}\sum_k W_a(k)\,\eta_a\right]\right\}^n. \tag{A.2}$$

Note that the ξ-dependence of η has not been written down explicitly. Just as in ref. 1 use will be made of a special case of the Hölder inequality for operators, *viz.*

$$\|O_2^{-1}\|^{-n}\,\mathrm{Tr}\,O_1^n \le |\mathrm{Tr}\,(O_1 O_2)^n| \le \|O_2\|^n\,\mathrm{Tr}\,O_1^n, \tag{A.3}$$

where the operator O_1 is assumed to be positive definite hermitean. In the absence of antiferromagnetic interactions W, the expression for ϱ_1 can be expressed as $\mathrm{Tr}\,O_1^n$, where O_1 is positive definite hermitean. This is not true in the present case. Therefore we shall compare ϱ_1 and ϱ_2 with the expression

$$\varrho_3 \equiv \mathrm{Tr}\left\{\int_{-\infty}^{\infty} \mathrm{d}\xi\, O_1(\xi)\right\}^n, \tag{A.4}$$

where

$$O_1(\xi) \equiv \exp\left[-\frac{\beta N}{2n}(\xi^2 - \eta^2)\right]$$

$$\times \exp\left[-\frac{\beta}{n}\sum_k \mathscr{H}(k, \xi, \eta) - \sum_a \frac{\beta N}{2n}\left(\sum_k W_a(k) - N\eta_a\right)^2\right]. \tag{A.5}$$

We shall prove that in the calculation of the free energy per particle ϱ_1 and ϱ_2 can be replaced by ϱ_3, so that ϱ_1 and ϱ_2 can be interchanged. In order to prove this, we write

$$\varrho_1 = \mathrm{Tr}\left[\int_{-\infty}^{\infty} \mathrm{d}\xi\, O_1(\xi)\, A^{(1)}(\xi)\right]^n,$$

$$\varrho_2 = \mathrm{Tr}\left[\int_{-\infty}^{\infty} \mathrm{d}\xi\, O_1(\xi)\, A^{(2)}(\xi)\right]^n, \tag{A.6}$$

where

$$A^{(1)}(\xi) = \exp\left[\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi,\eta) + \sum_a \frac{\beta}{2nN}\left(\sum_k W_a - N\eta_a\right)^2\right]$$

$$\times \exp\left[-\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi,\eta)\right]\prod_{a=1}^{q}\exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a - N\eta_a\right)^2\right], \tag{A.7}$$

$$A^{(2)}(\xi) = \exp\left[\frac{\beta}{n}\sum_k \mathscr{H}(k,\xi,\eta) + \sum_a \frac{\beta}{2nN}\left(\sum_k W_a - N\eta_a\right)^2\right]\exp\left[-\frac{\beta}{n}\sum_k T(k)\right]$$

$$\times \prod_f \exp\left[\frac{\beta}{n}\xi_f \sum_k V_f(k)\right]\prod_a \exp\left[-\frac{\beta}{2nN}\left(\sum_k W_a - N\eta_a\right)^2\right]$$

$$\times \exp\left[-\frac{\beta}{n}\sum_k W_a\eta_a\right].$$

As in ref. 1, we define operators $\tilde{A}^{(1)}$ and $\tilde{A}^{(2)}$ which no longer depend on ξ, *viz.*

$$\tilde{A}^{(1)} = \{\textstyle\int d\xi\, O_1(\xi)\}^{-1}\int d\xi\, O_1(\xi)\, A^{(1)}(\xi),$$
$$\tilde{A}^{(2)} = \{\textstyle\int d\xi\, O_1(\xi)\}^{-1}\int d\xi\, O_1(\xi)\, A^{(2)}(\xi), \tag{A.8}$$

so that

$$\varrho_1 = \mathrm{Tr}\,(O_1\tilde{A}^{(1)})^n, \qquad \varrho_2 = \mathrm{Tr}\,(O_1\tilde{A}^{(2)})^n,$$

where

$$O_1 \equiv \int d\xi\, O_1(\xi). \tag{A.9}$$

We now expand the operators $A(\xi)$ in (A.7) as a power series in $1/n$ and ξ. Note that η depends on ξ so that such an expansion in principle would contain all derivatives of η, which turns out to lead to divergences if the orders in $1/n$ are considered separately. However, it is sufficient to restrict ourselves to functions η which are bounded with respect to ξ, since ultimately we are only interested in the choice $\eta = \eta_{mf}$, and the function $\eta_{mf}(\xi)$ has been shown to be bounded in section 2. This shows that a term $f_{rs}(\eta)/(\xi^r/n^s)$, which is obtained by expanding $A(\xi)$ only as far as its explicit ξ-dependence is concerned, and which therefore contains an η-dependent (hence ξ-dependent) coefficient $f_{rs}(\eta)$, is bounded by $M_{rs}(\xi^r/n^5)$ where the finite constant M_{rs} no longer depends on ξ. Hence in such an expansion which is meant to see for which values of r and s terms of the form ξ^r/n^s contribute after integration ($\xi \sim n^{\frac{1}{4}}$) and after taking the limit $n \to \infty$, the ξ-dependence of η need not be taken into account.

We write the expansion as

$$A(\xi) = 1 + \sum_{s \geq 2} \sum_{r \leq s} A_{rs}(\xi), \tag{A.10}$$

where $A_{rs}(\xi)$ is the contribution of the terms that contain s factors $1/n$ and r factors ξ. In accordance with the discussion given above only the explicit ξ-dependence is taken into account. For the ξ-independent operators $\tilde{A}$ we have a similar expansion, where $\tilde{A}_{rs}$ is the contribution due to $A_{rs}(\xi)$ in (A.8). Using the same argument as in appendix A of ref. 1, we see that, apart from the constant 1, only $\tilde{A}_{22}$ may give a nonvanishing contribution to the free energy.

We now consider the expression (A.7) for $A^{(1)}(\xi)$ and $A^{(2)}(\xi)$ in more detail. Clearly

$$A_{22}^{(1)}(\xi) = 0, \tag{A.11}$$

since the terms that are proportional to $1/n^2$ contain at most one factor ξ. Hence

$$\lim_{n\to\infty} \mathrm{Tr} \left\{ \int_{-\infty}^{\infty} \mathrm{d}\xi\, O_1(\xi)\, A^{(1)}(\xi) \right\}^n = \lim_{n\to\infty} \mathrm{Tr} \left\{ \int_{-\infty}^{\infty} \mathrm{d}\xi\, O_1(\xi) \right\}^n. \tag{A.12}$$

Eq. (A.12) shows that the replacement $\varrho_1 \to \varrho_3$ is correct, even for finite systems. On the other hand,

$$\begin{aligned} A^{(2)}(\xi) &= \exp\left[-\frac{1}{n} \sum_{i=0}^{p+2q} B_i \right] \prod_{i=0}^{p+2q} \exp\left(\frac{B_i}{n} \right) \\ &= 1 + \frac{1}{2n^2} \sum_{i<j} [B_i, B_j] + \mathcal{O}(n^{-3}), \end{aligned} \tag{A.13}$$

where

$$\begin{aligned} B_0 &= -\beta \sum_k T(k), \\ B_f &= \beta \xi_f \sum_k V_f(k), \qquad f = 1, \ldots, p, \\ B_{p+2a-1} &= -\beta\, (2N)^{-1} \Big(\sum_k W_a(k) - N\eta_a \Big)^2, \\ B_{p+2a} &= -\beta \sum_k W_a \eta_a, \qquad a = 1, \ldots, q. \end{aligned} \tag{A.14}$$

For $A_{22}^{(2)}(\xi)$ we obtain the same expression as in ref. 1, *i.e.*

$$A_{22}^{(2)}(\xi) = \frac{\beta^2}{2n^2} \sum_{f<g} \xi_f \xi_g \sum_k [V_f(k), V_g(k)], \tag{A.15}$$

which leads to

$$\tilde{A}_{22}^{(2)} = \frac{\beta^2}{n} \sum_{f<g} Q_{fg}, \tag{A.16}$$

where the Q_{fg} are bounded in the limit $N \to \infty$.

Now $\varrho_2 = \mathrm{Tr}\,\{O_1\,(1 + \tilde{A}_{22}^{(2)})\}^n$ and using (A.3), we have

$$\|O_2^{-1}\|^{-n} \varrho_3 \leq \varrho_2 \leq \|O_2\|^n \varrho_3, \tag{A.17}$$

where $O_2 = 1 + \tilde{A}_{22}^{(2)}$.

In the limit $n \to \infty$, the operator norms $\|O_2^{-1}\|^{-n}$ and $\|O_2\|^n$ reduce to finite constants. Hence in the thermodynamic limit the replacement $\varrho_2 \to \varrho_3$ is correct.

APPENDIX B

In this appendix we calculate the elements of the matrices $L^{(1)}$, $L^{(2)}$ and $L^{(3)}$ as defined in (81).

Analogously to appendix B of ref. 1 we express the matrix elements in terms of the derivatives of $T(k, \{\xi_i\}, \{\chi_i\})$ defined by*

$$T(k, \{\xi_i\}, \{\chi_i\}) = \mathrm{tr}_k \prod_{j=1}^{n} \varrho_j(k)\, \sigma_j(k), \tag{B.1}$$

where

$$\varrho_j(k) = \exp\left[-(\beta/n)\, \mathscr{H}\,(k, \xi_j, \eta(\xi_j))\right] \tag{B.2}$$

and

$$\sigma_j(k) = \prod_{a=1}^{q} \exp\left[(\mathrm{i}\beta/n)\, \chi_{ja} W_a\,(k)\right]. \tag{B.3}$$

We have *e.g.*

$$L_{ij}^{(1)} = \delta_{ij} \left(1_p - \frac{1}{2} \frac{\mathrm{d}^2\eta^2}{\mathrm{d}\xi\, \mathrm{d}\xi}\right) + \frac{n}{N\beta} \sum_k T_k^{-2} \left(\frac{\partial T(k, \{\xi_i\}, \{\chi_i\})}{\partial \xi_i}\right)_0$$
$$\times \left(\frac{\partial T(k, \{\xi_i\}, \{\chi_i\})}{\partial \xi_j}\right)_0 - \frac{n}{\beta N} \sum_k T_k^{-1} \left(\frac{\partial^2 T(k, \{\xi_i\}, \{\chi_i\})}{\partial \xi_i\, \partial \xi_j}\right)_0. \tag{B.4}$$

* This function should not be confused with the operator $T(k)$ in the hamiltonian (3).

Note from (58) that instead of $T(k, \{\xi_l\}, \{\chi_l\})$ we should have taken $T(k, \{\xi_l\}, \{\chi_l\}) \cdot \exp[-(\mathrm{i}\beta/n)\,\eta(\xi_l)\cdot\chi_l]$. The last factor, however, only gives a contribution $\mathrm{i}\delta_{lj}(\mathrm{d}\eta/\mathrm{d}\xi)$ to $L_{lj}^{(3)}$, and does not change $L^{(1)}$ and $L^{(2)}$. For $L_{lj}^{(2)}$ and $L_{lj}^{(3)} - \mathrm{i}\delta_{lj}(\mathrm{d}\eta/\mathrm{d}\xi)$ we have formulae similar to (B.4). From the expression

$$\left(\frac{\partial^2 T}{\partial\xi_{if}\,\partial\xi_{jg}}\right)_0 = \mathrm{tr}_k\left\{\varrho^i\left(\varrho^{-1}\frac{\partial\varrho}{\partial\xi_f}\right)\varrho^{j-i}\left(\varrho^{-1}\frac{\partial\varrho}{\partial\xi_g}\right)\varrho^{n-j}\right\}, \qquad \text{for } i<j, \tag{B.5}$$

where

$$\varrho \equiv \varrho(k) = \exp[-(\beta/n)\,\mathscr{H}(k, \xi^0, \eta(\xi^0))], \tag{B.6}$$

and the analogous expressions for $j < i$ two general conclusions may be drawn, using the invariance of the trace for cyclic permutations.

1) the derivatives have a cyclic structure, *i.e.* $L_{lj}^{(\alpha)} \equiv L_{j-i}^{(\alpha)}$ and $L_{-j}^{(\alpha)} = L_{n-j}^{(\alpha)}$, for $\alpha = 1, 2, 3$. This property enables one to write down the expression for $i > j$ directly from the case $i < j$.

2) the second term in (B.4) may be combined with the last one by considering instead of $\varrho^{-1}(\partial\varrho/\partial\xi)$ in (B.5) the expression $(\varrho^{-1}(\partial\varrho/\partial\xi) - T^{-1}(\partial T/\partial\xi))$. This corresponds to replacing

$$V - W\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} \quad \text{by} \quad V - W\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - \left\langle V - W\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right\rangle_k \equiv \tilde{V} \tag{B.7a}$$

and

$$W \quad \text{by} \quad W - \langle W\rangle_k \equiv \tilde{W} \tag{B.7b}$$

in the other derivatives.

In the calculations use is made of the expressions

$$\frac{\partial\varrho_i(k)}{\partial\xi_i} = \int_0^{\beta/n} \mathrm{d}\tau \exp\left[\left(\tau - \frac{\beta}{n}\right)\mathscr{H}(k, \xi_i, \eta(\xi_i))\right]\left(V(k) - W(k)\cdot\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right)$$
$$\times \exp[-\tau\mathscr{H}(k, \xi_i, \eta(\xi_i))], \tag{B.8a}$$

and

$$\frac{\partial\sigma_i(k)}{\partial\chi_{ia}} = \frac{\mathrm{i}\beta}{n}\prod_{b\leq a}\exp\left[\frac{\mathrm{i}\beta}{n}\chi_{ib}W_b\right]W_a\prod_{b>a}\exp\left[\frac{\mathrm{i}\beta}{n}\chi_{ib}W_b\right], \tag{B.8b}$$

and all derivatives eventually must be evaluated at $\xi_i = \xi^0$, $\chi_i = 0$. From (B.4)–(B.8) and the cyclic properties of $L^{(1)}$ one finds, *cf.* also (I.B.6), for $i \neq j$

$$L_{ij}^{(1)} = \delta_{ij}\left(1_p - \eta \cdot \frac{\mathrm{d}^2\eta}{\mathrm{d}\xi\,\mathrm{d}\xi} - \frac{\mathrm{d}\eta}{\mathrm{d}\xi} \cdot \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right) .$$

$$- \frac{\beta}{nN} \sum_k T_k^{-1} \sum_{p,q} \tilde{V}_{pq}\tilde{V}_{qp}\, \mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)} \frac{2n^2\{\cosh(\beta/n)(h_p - h_q) - 1\}}{\beta^2(h_p - h_q)^2}$$

$$\times (\cosh\gamma_{j-i} - \sinh\gamma_{j-i}). \tag{B.9}$$

Here h_p and h_q are eigenvalues of the hamiltonian $\mathscr{H}(k, \xi^0, \eta(\xi^0))$ and O_{pq} are the matrix elements of an operator O in a representation in which $\mathscr{H}(k, \xi^0, \eta(\xi^0))$ is diagonal; γ_j has been defined in (96). For $i = j$ one should be careful. First of all, as in ref. 1, it turns out that (B.9) may only be used to first order in $1/n$ to evaluate $L_{jj}^{(1)}$. There is, however, also an extra term which arises from $\mathrm{d}^2\mathscr{H}(k, \xi_j, \eta(\xi_j))/\mathrm{d}\xi_j\,\mathrm{d}\xi_j$. This term can be seen to be

$$-\frac{n}{\beta N} \sum_k T_k^{-1}\, \mathrm{tr}_k \left\{ \varrho^j \int_0^{\beta/n} \mathrm{d}\tau\, \mathrm{e}^{\tau\mathscr{H}(k,\xi^0,\eta(\xi^0))} \right.$$

$$\left. \times \left(\frac{\mathrm{d}^2\mathscr{H}(k, \xi_j, \eta(\xi_j))}{\mathrm{d}\xi_j \cdot \mathrm{d}\xi_j}\right)_0 \mathrm{e}^{-\tau\mathscr{H}(k,\xi^0,\eta(\xi^0))}\, \varrho^{n-j} \right\},$$

which at the minimum, is equal to

$$\frac{1}{N} \sum_k \langle W(k)\rangle \cdot \frac{\mathrm{d}^2\eta}{\mathrm{d}\xi\,\mathrm{d}\xi},$$

thereby cancelling the term $-\delta_{ij}\eta \cdot (\mathrm{d}^2\eta/\mathrm{d}\xi\,\mathrm{d}\xi)$ in $L_{ij}^{(1)}$ in view of (18). So the result is

$$L_{ij}^{(1)} = \delta_{ij}\left(1_p - \frac{\mathrm{d}\eta}{\mathrm{d}\xi} \cdot \frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right) - \frac{\beta}{nN} \sum_k T_k^{-1} \sum_{p,q} \{\tilde{V}_{pq} \cdot \tilde{V}_{qp}\, \mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)}$$

$$\times (\cosh\gamma_{j-i} - \sinh\gamma_{j-i})\} + \mathcal{O}(1/n^2). \tag{B.10}$$

Using a similar line of reasoning one finds for $L_{ij}^{(2)}$: $(i \leq j)$

$$L_{ij}^{(2)} = \delta_{ij}1_q + \frac{\beta}{nN} \sum_k T_k^{-1}\, \mathrm{tr}_k \{\varrho^{n+i-j}(k)\, \tilde{W}(k)\, \varrho^{j-i}(k)\, \tilde{W}(k)\}$$

$$= \delta_{ij}1_q + \frac{\beta}{nN} \sum_k T_k^{-1} \sum_{p,q} \tilde{W}_{pq}\tilde{W}_{qp}\, \mathrm{e}^{-\frac{1}{2}\beta(h_p+h_p)} (\cosh\gamma_{j-i} - \sinh\gamma_{j-i}), \tag{B.11}$$

which is also valid for $i > j$ in view of the relations

$$L_{ij}^{(2)} \equiv L_{j-i}^{(2)}; \qquad L_{-j}^{(2)} = L_{n-j}^{(2)}.$$

For $L_{ij}^{(3)}$ we have for $i \leq j$

$$L_{ij}^{(3)} = \mathrm{i}\,\delta_{ij}\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - \frac{\mathrm{i}}{N}\sum_k T_k^{-1}\,\mathrm{tr}_k\left\{\int_0^{\beta/n}\mathrm{d}\tau\,\mathrm{e}^{\tau\mathscr{H}(k,\,\xi^0,\,\eta(\xi^0))}\right.$$

$$\left.\times\,\tilde{V}(k)\,\mathrm{e}^{-\tau\mathscr{H}(k,\,\xi^0,\,\eta(\xi^0))}\,\varrho^{j-i}(k)\,\tilde{W}(k)\,\varrho^{n-j+i}(k)\right\}. \tag{B.12}$$

The expression for $j < i$ can again be derived using the cyclic properties. The result is

$$L_{ij}^{(3)} = \mathrm{i}\,\delta_{ij}\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - \frac{\mathrm{i}}{N}\sum_k T_k^{-1}\sum_{p,\,q}\tilde{V}_{pq}\tilde{W}_{qp}\,\frac{\mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)}}{h_p - h_q}$$

$$\times\left\{\cosh\gamma_{j-i}\left(\cosh\frac{\beta}{n}(h_p - h_q) - 1\right) - \sinh\gamma_{j-i}\sinh\frac{\beta}{n}(h_p - h_q)\right.$$

$$\left.-\,\sinh\gamma_{j-i}\left(\cosh\frac{\beta}{n}(h_p - h_q) - 1\right) + \cosh\gamma_{j-i}\sinh\frac{\beta}{n}(h_p - h_q)\right\}$$

$$= \mathrm{i}\delta_{ij}\frac{\mathrm{d}\eta}{\mathrm{d}\xi} - \frac{\mathrm{i}\beta}{nN}\sum_k T_k^{-1}\sum_{p,\,q}\tilde{V}_{pq}\tilde{W}_{qp}\,\mathrm{e}^{-\frac{1}{2}\beta(h_p+h_q)}$$

$$\times\,\{\cosh\gamma_{j-i} - \sinh\gamma_{j-i}\} + \mathcal{O}(1/n^2). \tag{B.13}$$

APPENDIX C

In this appendix we prove that the matrix $\langle\langle\chi_k\chi_l\rangle\rangle$ has a cyclic structure and in addition we calculate the matrix $N(\varkappa)$, *cf.* eq. (107).

In order to do this, we introduce the Fourier components

$$\chi(\varkappa) = n^{-\frac{1}{2}}\sum_{j=1}^{n}\chi_j\,\mathrm{e}^{-2\pi\mathrm{i}\,j\varkappa/n}, \tag{C.1a}$$

and we have the inverse relation

$$\chi_j = n^{-\frac{1}{2}}\sum_{\varkappa=0}^{n-1}\chi(\varkappa)\,\mathrm{e}^{2\pi\mathrm{i}\,j\varkappa/n} \tag{C.1b}$$

Then

$$\langle\langle \chi_k \chi_l \rangle\rangle = n^{-1} \sum_{\varkappa, \lambda = 0}^{n-1} \langle\langle \chi(\varkappa^*)\, \chi(\lambda) \rangle\rangle \exp\left[(2\pi i/n)\,(l\lambda - k\varkappa)\right]. \tag{C.2}$$

If we want to prove that $\langle\langle \chi_k \chi_l \rangle\rangle$ is cyclic, we have to show that

$$\langle\langle \chi^*(\varkappa)\, \chi(\lambda) \rangle\rangle = N(\varkappa)\, \delta_{\varkappa, \lambda}, \tag{C.3}$$

i.e. the left-hand side is diagonal in $\varkappa$.

Using eqs. (C.1b), (90b), (103) and (104) we find that the quantity γ, *cf.* (88), is diagonal in $\varkappa$ and can be written

$$\begin{aligned} \gamma &= \tfrac{1}{2}\beta n^{-1} \sum_{k, l} \chi_k \cdot L_{kl}^{(2)} \cdot \chi_l \\ &= \tfrac{1}{2}\beta n^{-1} \sum_{\varkappa} \chi^*(\varkappa) \cdot (1_q - D^{(2)}(\varkappa) - E^{(2)}(\varkappa)) \cdot \chi(\varkappa). \end{aligned} \tag{C.4}$$

Since χ_j is real, we have

$$\chi^*(\varkappa) = \chi\,(n - \varkappa). \tag{C.5}$$

In order to perform the integrations in $\langle\langle \chi^*(\varkappa)\, \chi(\lambda) \rangle\rangle$, *cf.* (87), it is convenient to use real variables. For these we can choose

if n is even:

$$\begin{aligned} &\operatorname{Re} \chi(0) = \chi(0); \quad \operatorname{Re} \chi(\varkappa); \quad \operatorname{Im} \chi(\varkappa) \quad \text{for} \quad \varkappa = 1, \ldots, \tfrac{1}{2}n - 1, \\ &\text{and} \quad \operatorname{Re} \chi\,(\tfrac{1}{2}n) = \chi\,(\tfrac{1}{2}n); \end{aligned} \tag{C.6}$$

if n is odd:

$$\operatorname{Re} \chi(0) = \chi(0); \quad \operatorname{Re} \chi(\varkappa); \quad \operatorname{Im} \chi(\varkappa); \quad \varkappa = 1, \ldots, \tfrac{1}{2}\,(n - 1).$$

The jacobian associated with the transformation of χ_j, $j = 1, \ldots, n$ to these new variables does not depend on these variables; hence it disappears from $\langle\langle \chi^*(\varkappa)\chi(\lambda) \rangle\rangle$. From (C.4) it is obvious that the integrals $\int e^{-N\gamma}\, \chi^*(\varkappa)\, \chi(\lambda)$ can only be non-zero if $\lambda = \varkappa$ or $\lambda = n - \varkappa$.

First we prove that

$$\langle\langle \chi^*(\varkappa)\, \chi\,(n - \varkappa) \rangle\rangle = 0 \qquad (\varkappa \neq 0, \tfrac{1}{2}n). \tag{C.7}$$

Define

$$y \equiv y(\varkappa) = \operatorname{Re} \chi(\varkappa), \qquad z \equiv z(\varkappa) = \operatorname{Im} \chi(\varkappa). \tag{C.8}$$

If we write $\gamma = \sum_{\varkappa=0}^{n-1} \gamma(\varkappa)$, *cf.* (C.4), and use eqs. (122), we have (C.9)

$$\gamma(\varkappa) + \gamma(n-\varkappa) = (\beta/n)\{y \cdot (1 - D^{(2)}) \cdot y + z \cdot (1 - D^{(2)}) \cdot z$$
$$+ \mathrm{i}z \cdot E^{(2)} \cdot y - \mathrm{i}y \cdot E^{(2)} \cdot z\}. \tag{C.10}$$

The quantities in the right-hand side of (C.10) depend on $\varkappa$, but this dependence has not been shown explicitly for the sake of convenience. Using (C.8) one arrives at

$$\langle\langle \chi^*(\varkappa)\chi(n-\varkappa)\rangle\rangle = \langle\langle yy\rangle\rangle - \langle\langle zz\rangle\rangle - \mathrm{i}\{\langle\langle yz\rangle\rangle + \langle\langle zy\rangle\rangle\}, \tag{C.11}$$

where now *e.g.*

$$\langle\langle yy\rangle\rangle = (\beta N/n)\int \mathrm{d}y\,\mathrm{d}z\,\mathrm{e}^{-N(\gamma(\varkappa)+\gamma(n-\varkappa))}\,yy/\int \mathrm{d}y\,\mathrm{d}z\,\mathrm{e}^{-N(\gamma(\varkappa)+\gamma(n-\varkappa))}. \tag{C.12}$$

In view of (C.10) it is easy to see from simple transformations of variables like $(y, z) \to (z, y)$ or $(y, z) \to (y, -z)$, that

$$\begin{aligned} \langle\langle yy\rangle\rangle &= \langle\langle yy\rangle\rangle^* = \langle\langle zz\rangle\rangle, \\ \langle\langle yz\rangle\rangle &= -\langle\langle yz\rangle\rangle^* = \langle\langle zy\rangle\rangle^*. \end{aligned} \tag{C.13}$$

From (C.11) and (C.13) it follows that $\langle\langle \chi^*(\varkappa)\chi(n-\varkappa)\rangle\rangle = 0$, and eq. (C.3) has been proved.

In order to calculate the matrix $N(\varkappa) \equiv \langle\langle \chi^*(\varkappa)\chi(\varkappa)\rangle\rangle$, for $\varkappa \neq 0, \frac{1}{2}n$ we introduce a $2q$-dimensional vector $t = (y, z)$ and the $2q$-dimensional matrix T defined by

$$T(\varkappa) = \begin{pmatrix} 1 - D^{(2)}(\varkappa) & -\mathrm{i}E^{(2)}(\varkappa) \\ \mathrm{i}E^{(2)}(\varkappa) & 1 - D^{(2)}(\varkappa) \end{pmatrix}. \tag{C.14}$$

Then

$$\gamma(\varkappa) + \gamma(n-\varkappa) = \beta n^{-1} t \cdot T \cdot t, \tag{C.15}$$

and

$$N(\varkappa) = 2\beta N n^{-1} \int \mathrm{d}t\,\mathrm{e}^{-\beta N n^{-1} t \cdot T \cdot t}\,(yy + \mathrm{i}yz)/\int \mathrm{d}t\,\mathrm{e}^{-\beta N n^{-1} t \cdot T \cdot t}. \tag{C.16}$$

From (112) it follows that T is a symmetrical matrix; also, in view of eq. (110) which shows that $1 - D^{(2)}(\varkappa)$ is positive definite, Re T is positive definite.

For an arbitrary symmetrical $m \times m$ matrix M with positive definite real part,

we have the relations, *cf.* ref. 18

$$\int_{-\infty}^{\infty} \mathrm{d}x_1 \cdots \mathrm{d}x_m \exp\left(-\tfrac{1}{2}\sum_{i,j=1}^{m} x_i M_{ij} x_j\right) = (2\pi)^{\frac{1}{2}m} (\det M)^{-\frac{1}{2}}, \tag{C.17}$$

$$\int_{-\infty}^{\infty} \mathrm{d}x_1 \cdots \mathrm{d}x_m \, x_i x_j \exp\left(-\tfrac{1}{2}\sum_{i,j} x_i M_{ij} x_j\right) = (2\pi)^{\frac{1}{2}m} (M^{-1})_{ji} (\det M)^{-\frac{1}{2}}, \tag{C.18}$$

where M^{-1} is the inverse matrix of M. After applying eqs. (C.17) and (C.18) to (C.16) we find that the elements of the matrix N can be written

$$N_{ab}(\varkappa) = (T^{-1})_{ba} + \mathrm{i}\,(T^{-1})_{q+b,a}, \tag{C.19}$$

where $a, b = 1, \ldots, q$.

Eq. (C.19) is also valid if $\varkappa = 0, \frac{1}{2}n$. In fact a straightforward calculation shows that $N_{ab}(\varkappa)$ in these cases is given by the matrix $(1 - D^{(2)}(\varkappa))^{-1}_{ba}$, which is equivalent to the right-hand side of (C.19) since we have $E^{(2)}(\varkappa) = 0$ if $\varkappa = 0, \frac{1}{2}n$, *cf.* (112).

It is now easy to derive an explicit formula for the matrix $N(\varkappa)$ in terms of the matrices $D^{(2)}(\varkappa)$ and $E^{(2)}(\varkappa)$. In order to do this, we start from (C.14). We write the inverse matrix T^{-1} as

$$T^{-1} \equiv \begin{pmatrix} A & B \\ \Gamma & \Delta \end{pmatrix}, \tag{C.20}$$

where A, B, Γ and Δ are $q \times q$ matrices satisfying the relations

$$\Lambda A - \mathrm{i}E\Gamma = 1, \qquad \mathrm{i}EA + \Lambda\Gamma = 0. \tag{C.21}$$

Here Λ and E are shorthand notations for

$$\Lambda \equiv 1 - D^{(2)}(\varkappa), \qquad E \equiv E^{(2)}(\varkappa). \tag{C.22}$$

From (C.21) the matrices A and Γ can be solved and we have

$$A = \Lambda_E^{-1}, \qquad \Gamma = -\mathrm{i}\Lambda^{-1}E\Lambda_E^{-1}, \tag{C.23}$$

where

$$\Lambda_E = \Lambda + E\Lambda^{-1}E^{\dagger}. \tag{C.24}$$

The inverse matrices Λ^{-1} and Λ_E^{-1} exist since $D^{(2)}(\varkappa)$ is negative definite, so that $\Lambda = 1 - D^{(2)}(\varkappa)$ and $E\Lambda^{-1}E^{\dagger}$ are positive definite.

From (C.19), (C.23) and (112) it follows that the matrix N is given by

$$N = A + \mathrm{i}\tilde{\Gamma} = \Lambda_E^{-1}(1 + E\Lambda^{-1})$$

$$= [1 - D^{(2)}(\varkappa) + E^{(2)}(\varkappa)(1 - D^{(2)}(\varkappa))^{-1} E^{(2)\dagger}(\varkappa)]^{-1}$$
$$\times [1 + E^{(2)}(\varkappa)(1 - D^{(2)}(\varkappa))^{-1}]. \tag{C.25}$$

APPENDIX D

In order to derive eq. (109), we define the matrix

$$\mathscr{D}^{(3)}(\varkappa) \equiv D^{(3)}(\varkappa) + E^{(3)}(\varkappa) - \mathrm{i}\,\frac{\mathrm{d}\eta}{\mathrm{d}\xi}. \tag{D.1}$$

Then, in view of (90c), we have

$$L_{ij}^{(3)} = -n^{-1} \sum_{\varkappa=0}^{n-1} \mathscr{D}^{(3)}(\varkappa)\, \mathrm{e}^{2\pi \mathrm{i} j\varkappa/n}. \tag{D.2}$$

From appendix C, *cf.* eqs. (C.2) and (C.3) we conclude that $\langle\langle \chi_k \chi_{l+j} \rangle\rangle$ can be written

$$\langle\langle \chi_k \chi_{l+j} \rangle\rangle = \sum_{\varkappa=0}^{n-1} N(\varkappa) \exp[(2\pi\mathrm{i}/n)(l + j - k)\varkappa]. \tag{D.3}$$

Substituting (D.2) and (D.3) into (106) we get

$$D_{fg}^{(33)}(\varkappa) = \tfrac{1}{2} n^{-3} \sum_{j,k,l=0}^{n-1} \sum_{\lambda,\mu,\nu=0}^{n-1} \sum_{a,b=1}^{q} \mathscr{D}_{fa}^{(3)}(\lambda)\, \mathscr{D}_{gb}^{(3)}(\mu)\, N_{ab}(\nu)$$
$$\times \exp[(2\pi\mathrm{i}/n)(j\varkappa - k\lambda - l\mu + (l + j - k)\nu)]$$
$$+ \tfrac{1}{2} n^{-3} \sum_{j,k,l} \sum_{\lambda,\mu,\nu} \sum_{a,b} \mathscr{D}_{fa}^{*(3)}(\lambda)\, \mathscr{D}_{gb}^{*(3)}(\mu)\, N_{ab}^{*}(\nu)$$
$$\times \exp[(2\pi\mathrm{i}/n)(j\varkappa + k\lambda + l\mu - (l + j - k)\nu)]. \tag{D.4}$$

The summations over k, l and j give a number of Kronecker delta's, *viz.* $\delta_{\lambda,-\nu}$, $\delta_{\mu,\nu}$ and $\delta_{\nu,-\varkappa}$ in the first term and $\delta_{\lambda,-\nu}$, $\delta_{\mu,\nu}$ and $\delta_{\nu,\varkappa}$ in the second term. As a result we have

$$D_{fg}^{(33)}(\varkappa) = \tfrac{1}{2} \sum_{a,b} \mathscr{D}_{fa}^{(3)}(\varkappa)\, \mathscr{D}_{gb}^{(3)}(-\varkappa)\, N_{ab}(-\varkappa)$$
$$+ \tfrac{1}{2} \sum_{a,b} \mathscr{D}_{fa}^{(3)*}(-\varkappa)\, \mathscr{D}_{gb}^{(3)*}(\varkappa)\, N_{ab}^{*}(\varkappa). \tag{D.5}$$

Using (113), we can write (D.5) in matrix notation as

$$D^{(33)}(\varkappa) = \tfrac{1}{2}\mathscr{D}^{(3)}(\varkappa)\, N(-\varkappa)\, \tilde{\mathscr{D}}^{(3)}(-\varkappa) + \text{h.c.}, \tag{D.6}$$

where h.c. as usual denotes the hermitean conjugated operator. Substituting (D.1) into (D.6) and applying some of the properties (112) one readily obtains eq. (109).

APPENDIX E

We first prove that M_1, *cf.* (115a), is positive definite, *i.e.*

$$z^* \cdot D^{(33)} \cdot z \leq z^* \cdot D^{(3)} (1 - D^{(2)})^{-1} D^{(3)\dagger} \cdot z, \tag{E.1}$$

for arbitrary complex z. Eq. (E.1) will be proved as a special case of the more general inequality for the matrix N

$$\begin{aligned} \psi(\zeta, \chi) \equiv\ & \zeta^* \cdot (N + N^\dagger) \cdot \zeta - \chi \cdot (N + N^\dagger) \cdot \chi + \zeta^* \cdot (N - N^\dagger) \cdot \chi \\ & - \chi^* \cdot (N - N^\dagger) \cdot \zeta - 2\zeta^* (1 - D^{(2)})^{-1} \cdot \zeta \leq 0, \end{aligned} \tag{E.2}$$

for arbitrary q-dimensional vectors ζ, χ. In order to prove (E.2) we note that the matrix N can be written as

$$N = \Lambda_E^{-1} (1 + E\Lambda^{-1}), \tag{E.3}$$

where $\Lambda_E = 1 - D^{(2)}$, $E = E^{(2)}$ and $\Lambda_E = \Lambda + E\Lambda^{-1}E^+$, *cf.* (C.22), (C.24) and (C.25). We use the relations

$$\begin{aligned} & \Lambda_E (N + N^\dagger) \Lambda_E = 2\Lambda_E, \\ & \Lambda_E (N - N^\dagger) \Lambda_E = 2E\Lambda^{-1}\Lambda_E = 2\Lambda_E\Lambda^{-1}E, \\ & \Lambda_E\Lambda^{-1}\Lambda_E = \Lambda_E - E\Lambda^{-1}\Lambda_E\Lambda^{-1}E. \end{aligned} \tag{E.4}$$

Then by using the notations

$$x = -\Lambda^{-1}E\Lambda_E^{-1} \cdot \zeta \quad \text{and} \quad y = \Lambda_E^{-1} \cdot \chi, \tag{E.5}$$

so that

$$x^* = \zeta^* \cdot \Lambda_E^{-1}E\Lambda^{-1} \quad \text{and} \quad y^* = \chi^* \cdot \Lambda_E^{-1},$$

we have from (E.4)

$$\tfrac{1}{2}\psi(\zeta, \chi) = -(x - y)^* \cdot \Lambda_E \cdot (x - y) \leq 0, \tag{E.6}$$

since the matrix Λ_E is positive definite. Hence (E.2) has been proved and (E.2) reduces to (E.1) by substituting

$$\zeta = D^{(3)\dagger} \cdot z, \qquad \chi = \left(E^{(3)} - \mathrm{i}\,\frac{\mathrm{d}\eta}{\mathrm{d}\xi}\right)^{\dagger} \cdot z. \tag{E.7}$$

Next we consider M_2. $M_2 \geq 0$ is equivalent to

$$\xi^* \cdot [D^{(1)}(\varkappa) - D^{(1)}(0) + D^{(3)}(\varkappa)\,(1 - D^{(2)}(\varkappa))^{-1}\,D^{(3)\dagger}(\varkappa)] \cdot \xi \leq 0, \tag{E.8}$$

for arbitrary ξ. Eq. (E.8) will be proved as a special case of the more general inequality

$$\phi\,(\xi, \chi) \equiv \xi^* \cdot D^{(1)}(\varkappa) \cdot \xi + \chi^* \cdot D^{(2)}(\varkappa) \cdot \chi + \xi^* \cdot D^{(3)}(\varkappa) \cdot \chi$$
$$+ \chi^* \cdot D^{(3)\dagger}(\varkappa) \cdot \xi \leq \xi^* \cdot D^{(1)}(0) \cdot \xi, \tag{E.9}$$

for arbitrary p-dimensional ξ and q-dimensional χ. From the explicit expressions (110), (111) and (92), we have

$$\phi = \beta\,(2nN)^{-1} \sum_k T_k^{-1} \sum_{p,q} \mathrm{e}^{-\frac{1}{2}\beta\,(h_p + h_q)}\,\frac{4n \sinh \frac{1}{2}\beta\,(h_p - h_q)}{\beta\,(h_p - h_q)\,(1 + c_{pq}^2)}$$
$$\times \{|x_{pq}|^2 - |y_{pq}|^2 + (x_{pq}^* y_{pq} + y_{pq}^* x_{pq})\,c_{pq} \cos(\pi\varkappa/n)\}, \tag{E.10}$$

where x_{pq}, y_{pq} and c_{pq} are shorthand notations for

$$x_{pq} = \tilde{V}_{qp} \cdot \xi; \qquad y_{pq} = \tilde{W}_{qp} \cdot \chi; \qquad c_{pq} = n \sin(\pi\varkappa/n)/\beta\,(h_p - h_q). \tag{E.11}$$

It is clear that the term between brackets in (E.10) is smaller than

$$|x_{pq}|^2 - |y_{pq}|^2 + 2\,|c_{pq}|\,|x_{pq}|\,|y_{pq}| \leq (1 + c_{pq}^2)\,|x_{pq}|^2. \tag{E.12}$$

Using (E.12), we then have

$$\phi \leq \beta\,(2nN)^{-1} \sum_k T_k^{-1} \sum_{p,q} |x_{pq}|^2\,\mathrm{e}^{-\frac{1}{2}\beta\,(h_p + h_q)}\,\frac{4n \sinh \frac{1}{2}\beta\,(h_p - h_q)}{\beta\,(h_p - h_q)}$$
$$= \xi^* \cdot D^{(1)}(0) \cdot \xi, \tag{E.13}$$

which proves (E.9).

Eq. (E.8) can be obtained as a special case of (E.9). By substituting

$$\chi = -D^{(3)-1}(\varkappa)\,D^{(3)\dagger}(\varkappa) \cdot \xi$$

into (E.9), one finds that

$$\xi^* \cdot [D^{(1)}(\varkappa) - D^{(1)}(0) - D^{(3)}(\varkappa)\, D^{(2)-1}(\varkappa)\, D^{(3)\dagger}(\varkappa)] \cdot \xi \leq 0. \tag{E.14}$$

Since the matrix $D^{(2)}(\varkappa)$ is negative definite, we have

$$D^{(2)-1}(\varkappa) + (1 - D^{(2)}(\varkappa))^{-1} = D^{(2)-1}(\varkappa)\,(1 - D^{(2)}(\varkappa))^{-1} < 0. \tag{E.15}$$

Then also

$$\xi^* \cdot D^{(3)}(\varkappa)\, D^{(2)-1}(\varkappa)\, D^{(3)\dagger}(\varkappa) \cdot \xi + \xi^* \cdot D^{(3)}(\varkappa)\,(1 - D^{(2)}(\varkappa))^{-1}\, D^{(3)\dagger}(\varkappa) \cdot \xi \leq 0, \tag{E.16}$$

and (E.8) is now obvious from (E.16) and (E.14).

Note added in proof. A very interesting investigation of models with separable interactions has been carried out by Bogoliubov Jr.[19]). In ref. 19, the treatment of hamiltonians with mixed ferromagnetic and antiferromagnetic interactions is restricted to a subclass of the general hamiltonian (3) in the present paper, *viz.* essentially a generalized BCS-type of interaction. In the ferromagnetic case, the approach by Bogoliubov appears to have a much more general validity. Recent applications can be found in several Dubna preprints, *cf.* also *e.g.* ref. 20.

One of the authors (P.T.) wishes to thank Dr. J. G. Brankov for an interesting discussion.

REFERENCES

1) Tindemans, P.A.J. and Capel, H.W., Physica **72** (1974) 433.
2) Mühlschlegel, B. and Zittartz, H., Z. Phys. **175** (1963) 553.
3) Kac, M., Phys. Fluids **2** (1959) 8.
4) Baker Jr., G.A., Phys. Rev. **122** (1961) 1477.
5) Siegert, A.J.F. and Vezzetti, D.J., J. math. Phys. **9** (1968) 2173.
6) Kittel, C. and Shore, H., Phys. Rev. **138A** (1965) 1165.
7) Niemeijer, Th., Physica **48** (1970) 467.
8) Mühlschlegel, B., J. math. Phys. **3** (1962) 522.
9) Bardeen, J., Cooper, L.N. and Schrieffer, J.R., Phys. Rev. **108** (1957) 1175.
10) De Vries, A.S., Vertogen, G. and Kraak, J., Physica **63** (1973) 619.
11) Haag, R., Nuovo Cimento **25** (1962) 287.
12) See *e.g.* Falk, H., Am. J. Phys. **38** (1970) 858.
13) Reed, M. and Simon, B., Methods of Mathematical Physics, part 1: Functional Analysis, Academic Press (New York, 1972) p. 295.
14) Stratonovich, R.L., Soviet Physics – Doklady **2** (1958) 416.
15) Capel, H.W. and Tindemans, P.A.J., Rep. math. Phys., to be published.
16) Lenard, A., J. math. Phys. **5** (1964) 930.
17) Mazur, P. and Siskens, Th.J., Physica **69** (1973) 259.

18) Cramer, H., Mathematical Methods of Statistics, p. 120. Princeton University Press (Princeton, 1946).
19) Bogoliubov Jr., N.N., A Method for Studying Model Hamiltonians, Pergamon Press (Oxford, 1972).
20) Brankov, J.G., Shumovsky, A.S. and Zagrebnov, V.A., Physica, to be published.

V. APPLICATIONS OF THE GENERAL RESULTS

1. Introduction

In the introduction to chapters III and IV we have already mentioned a number of models that have been considered in the literature. These models cover a fairly wide range of physical situations. They include e.g. the phenomenon of superconductivity which can be described in terms of a Hamiltonian containing long-range interactions. Furthermore one can also consider problems concerning ferromagnetism and antiferromagnetism both of the localized and the itinerant type in situations in which the interactions are approximated by interactions of the extreme long-range type.

All such models can be described by Hamiltonians which belong to the class that has been studied in chapters III and IV. One is interested in the thermodynamic behaviour of these systems and it was seen in the foregoing chapters that due to the extremely long-range character of the interactions between the constituents ("particles") of the systems, their free energy could be calculated exactly. It turned out to be of the molecular-field type and contained a number of order parameters (analogous to the total or sub-lattice magnetizations in magnetic systems), whose values must be determined from so-called molecular-field equations. The solutions of these equations and especially the stability properties of the free energy as a function of the order parameters (it should be a minimum) have to be investigated in order to find the correct free energy. All this is a necessary prerequisite for studying different aspects of thermodynamic behaviour of these systems, e.g. the occurrence of phase transitions and of critical or tricritical points. In connection with this it may be noted that the classical critical exponents seem to be exact in the case of tricritical points in a three-dimensional system, in contrast to the exponents for a normal critical point [1]. Therefore we present in this chapter a number of general results valid for the whole class of systems under investigation. Though to some extent they can be considered as different formulations of results stated previously, they can be useful in dealing with the equations corresponding to some particular system . First, however, it is noted that a generalization can be made to a much larger class of systems: i.e. the one-particle operator T may be replaced by a more general operator containing

interactions, at least for systems with purely ferromagnetic quadratic interactions (cf. also the work of Bogoliubov jr. [2), 3)]). This allows one to study the simultaneous occurrence of short-range and long-range interactions, at least in principle. After these more general considerations we turn to a more detailed discussion of several models which have been proposed for the investigation of a great variety of physical situations. These models are either members of the class treated in this thesis, or have been studied in a molecular-field approximation. In both cases the present formulation presents a general and unambiguous way for obtaining the equations relevant for the discussion of the physical behaviour.

2. Extension to general T

The Hamiltonian that was considered in chapter III, consisted of a one-particle part and of a quadratic ferromagnetic two-particle part

$$\mathcal{H} = \sum_k T(k) - (2N)^{-1} \sum_{f=1}^{p} (\sum_k V_f(k))^2 . \tag{1}$$

Here the (bounded) operators $T(k)$ and $V_f(k)$, $f = 1, \ldots, p$ can be interpreted as one-particle operators.

Instead we consider the more general class of Hamiltonians, describing systems of N **particles**

$$\mathcal{H} = T - (2N)^{-1} \sum_{f=1}^{p} V_f^2 . \tag{2}$$

Here T and V_f, $f = 1, \ldots, p$, are hermitean operators, defined on the Hilbertspace of the N-particle system. It is supposed that the following operator norms remain finite in the thermodynamic limit $N \to \infty$, viz.

$$\| N^{-1}T\| , \| N^{-1}V_f \| , \| N^{-1}[T,V_f] \| , \| N^{-1}[V_f,V_g] \| , \quad f,g = 1,\ldots,p. \tag{3}$$

The generalization, as compared to chapter III, lies in the fact that T need no longer be a one-particle operator, such as a kinetic energy term or a term representing the influence of external fields. In fact, T may contain any interactions between the constituents ("particles") of the system. In the literature one finds discussions of one-dimensional spin systems, where T contains a nearest neighbour Ising or an XY-interaction [4),5),6)]. Also the operators V_f may be more general. To our knowledge, however, no specific examples of such a situation have been considered.

Folowing the line of arguments of chapter III, the free energy per

particle in the thermodynamic limit, f, for a system described by Hamiltonian (2), is given by

$$f = \lim_{N\to\infty} \min_{\vec{\xi}} F[\tfrac{1}{2}N\vec{\xi}^2 + T - \vec{\xi}\cdot V] \quad . \tag{4}$$

Here F[A] is, as usual, defined by

$$F[A] = -\beta^{-1} \ln \mathrm{Tr}\, e^{-\beta A} \quad . \tag{5}$$

The result shows that the influence of the long-range, separable, ferromagnetic interactions can be accounted for by replacing them by $-\vec{\xi}\cdot V$ with suitably chosen parameters $\vec{\xi}$. Of course the practical applicability of this result is restricted to cases where the free energy of the Hamiltonian $\mathcal{H}_0(\vec{\xi}) = \frac{1}{2}N\vec{\xi}^2 + T - \vec{\xi}\cdot\vec{V}$ can be calculated for arbitrary $\vec{\xi}$.

The proof of eq. (5) is completely analogous to the proof presented in chapter III. In fact, during all different steps in the derivation we have not used in an essential way the assumption that the operators T or V_f should be one-particle operators. The details of this proof will not be given here, but can be found in a forthcoming publication, cf. ref. 7.

Just as in chapter III we derive an upper bound and a lower bound to the free energy per particle. As a result we have the inequality

$$\beta^{-1}G_0 \geq f \geq \beta^{-1}G_0 + (2\beta N)^{-1} \lim_{n\to\infty} \sum_{\kappa=0}^{n-1} \ln \det(\vec{\vec{I}}_p - \vec{\vec{D}}(\kappa)) \ , \tag{6}$$

where $\beta^{-1}G_0$ is given by the right-hand side of (4) and where the $p\times p$ matrix $\vec{\vec{D}}(\kappa)$ is of order $1/\kappa^2$ for large $\kappa \leq \frac{1}{2}n$ and $\vec{\vec{I}}_p - \vec{\vec{D}}(\kappa)$ is positive definite if F has an absolute minimum as a function of $\vec{\xi}$.

In connection with this it may be noted that an extensive research, though along different lines, on systems of this general type with ferromagnetic separable interactions has been carried out by N.N. Bogoliubov jr. (see ref. 2 and especially ref. 3 for a detailed review). It seems that the general nature of his approach has not been appreciated sufficiently in the literature, a possible reason being Bogoliubov's emphasis on systems with a BCS-type of Hamiltonian, cf. chapter III for references.

Recently applications to several other physical systems have been given in a number of Dubna preprints [8),9)], cf. also 10). In the framework of our approach systems with a BCS-type of Hamiltonian can be dealt with in a simpler way following the line of reasoning used by Mühlschlegel [33)]. In ref.33 use is made of the simple inequality (II. 75) in terms of imaginary

rotation operators with spin $\frac{1}{2}$. This inequality can be proved by elementary methods. For the more general case, however, use must be made of the Hölder inequality together with the necessary and sufficient conditions for the equality sign to hold.

In Bogoliubov's treatment of the ferromagnetic case the construction of an upper bound to the free energy per particle in the thermodynamic limit proceeds in the same way as in our approach using the Bogoliubov-Peierls inequality. For the lower bound, on the other hand, Bogoliubov jr. uses a number of subtle inequalities and an ingenious integration over complex variables in order to show that in the thermodynamic limit the free energy differs from the free energy of the model Hamiltonian $\frac{1}{2}N\vec{\xi}^2 + T - \vec{\xi}\cdot\vec{V}$ by a power of N which is smaller than 1. As a consequence, however, Bogoliubov's method does not seem to be very well suited, if one is interested in the evaluation of e.g. the volume-independent corrections to the total free energy of large systems.

These corrections can be investigated by a direct calculation. We have followed in chapters III and IV an approach which is familiar in statistical mechanics by using an integral representation for the partition function and Laplace's method. In addition, in section 8 of chapter III we have discussed a link between the presence of divergencies due to second derivatives and the occurrence of phase transitions.

In the case that also antiferromagnetic quadratic operators are present, Bogoliubov's treatment is restricted to a subclass of the systems described by the Hamiltonian (IV. 1), viz. essentially those with generalized BCS-type of interactions. At this stage it may be mentioned that the generalization to an arbitrary operator T in the case of ferromagnetic interactions can be used as an important tool in the construction of a simplified derivation of the free energy per particle corresponding to the Hamiltonian (IV. 1) containing a one-particle operator $\Sigma_k T(k)$ and separable interactions both of the ferromagnetic and antiferromagnetic type. In this way some of the rather complicated considerations on the convergence of the matrix of second derivatives in chapter IV, although of interest for the purpose of investigating the Laplace method and the contributions to the free energy due to terms independent of the volume, can be bypassed. However, in the presence of antiferromagnetic interactions the generalization to an arbitrary operator T is in general not correct; the validity of the derivation is restricted to the cases, in which T is a one-particle operator. These recent

developments can be found in a forthcoming paper [7].

3. Generalized interactions

In this section we shall consider a slightly generalized form of the quadratic-interaction part of the Hamiltonian. We shall show that the free energy may be calculated in an obvious way by introducing parameters corresponding to the operators that occur in this more general form.

Consider the Hamiltonian

$$\mathcal{H} = T - \tfrac{1}{2}N^{-1} \sum_{i,j=1}^{n} A_{ij} X_i X_j \quad , \tag{7}$$

where T, X_i, $i = 1, \ldots, n$ are hermitean one-particle operators, i.e. $T = \Sigma_{k=1}^{N} T(k)$, $X_i = \Sigma_k X_i(k)$, and the matrix $\vec{A}$ with elements A_{ij} is real. Without loss of generality $\vec{A}$ can be taken to be symmetric and we suppose for the moment that its inverse exists.

We denote by $\vec{O}$ the orthogonal matrix that diagonalizes $\vec{A}$, so $\vec{O}^{-1}\vec{A}\vec{O}$ is diagonal, and by λ_i the eigenvalues of $\vec{A}$. Now the eigenvalues $\lambda_1, \ldots, \lambda_n$ can be positive or negative. We assume that the first p eigenvalues are positive and the remaining $n-p = q$ eigenvalues negative, i.e.

$$\lambda_f > 0, \quad \text{for } f = 1, \ldots, p, \quad \text{and } \lambda_{p+a} < 0, \quad \text{for } a = 1, \ldots, q. \tag{8}$$

We then can define new operators

$$V_f \equiv \lambda_f^{\frac{1}{2}} \sum_{j=1}^{n} O_{fj}^{-1} X_j \ , \quad f = 1, \ldots, p \ , \tag{9}$$

$$W_a \equiv |\lambda_{p+a}|^{\frac{1}{2}} \sum_{j=1}^{n} O_{p+a,j}^{-1} X_j \ , \quad a = 1, \ldots, q \ . \tag{10}$$

The Hamiltonian can now be written as

$$\mathcal{H} = T - \tfrac{1}{2}N^{-1} \sum_{f=1}^{p} V_f^2 + \tfrac{1}{2}N^{-1} \sum_{a=1}^{q} W_a^2 \ , \tag{11}$$

which has the form that we have treated before.

The free energy is then given by

$$f = \min_{\vec{\xi}} \max_{\vec{\eta}} f(\xi,\eta) \tag{12}$$

where

$$f(\vec{\xi},\vec{\eta}) = \lim_{N\to\infty} N^{-1} F[\mathcal{H}_0(\vec{\xi},\vec{\eta})] \tag{13}$$

and

$$\mathcal{H}_0(\vec{\xi},\vec{\eta}) = \tfrac{1}{2}N\left(\sum_{f=1}^{p} \xi_f^2 - \sum_{a=1}^{q} \eta_a^2\right) + T - \sum_{f=1}^{p} \xi_f V_f + \sum_{a=1}^{q} \eta_a W_a \ . \tag{14}$$

In eq. (12) the parameters $\eta_1, \ldots, \eta_q$ are chosen for each value of $\vec{\xi}$ in such a way that $N^{-1}F[\mathcal{H}_0(\vec{\xi},\vec{\eta})] = f(\vec{\xi},\vec{\eta})$ is maximal. Next the parameters $\xi_1, \ldots, \xi_p$ are determined so that $N^{-1}F[\mathcal{H}_0(\vec{\xi},\vec{\eta}_{mf}(\vec{\xi}))]$ is minimal. In view of e.g. section 2 of IV, the parameters $\xi_1, \ldots, \xi_p$, $\eta_1, \ldots, \eta_q$ are those solutions of the molecular-field equations

$$\partial f(\vec{\xi},\vec{\eta})/\partial\vec{\xi} = 0 \quad , \quad \partial f(\vec{\xi},\vec{\eta})/\partial\vec{\eta} = 0 \tag{15}$$

that lead to the lowest value of $f(\vec{\xi},\vec{\eta}_{mf}(\vec{\xi}))$.

The molecular-field equations (15) can be written as

$$\xi_f = N^{-1} \langle V_f \rangle_{\mathcal{H}_0(\vec{\xi},\vec{\eta})} \quad , \quad f = 1, \ldots, p \tag{16}$$

$$\eta_a = N^{-1} \langle W_a \rangle_{\mathcal{H}_0(\vec{\xi},\vec{\eta})} \quad , \quad a = 1, \ldots, q \ . \tag{17}$$

Define now parameters ζ_i, $i = 1, \ldots, n = p+q$ by

$$\xi_f = \lambda_f^{\frac{1}{2}} \sum_{j=1}^{n} O_{fj}^{-1} \zeta_j \tag{18}$$

$$\eta_a = |\lambda_{p+a}|^{\frac{1}{2}} \sum_{j=1}^{n} O_{p+a,j}^{-1} \zeta_j \ . \tag{19}$$

Since the inverse of the matrix $\vec{A}$ exists (i.e. none of the eigenvalues λ_i is zero), this transformation is one-to-one. Obviously, we have the identity

$$\mathcal{H}_0(\vec{\xi},\vec{\eta}) = \mathcal{H}_0(\{\zeta_i\}) \quad , \tag{20}$$

where

$$\mathcal{H}_0(\{\zeta_i\}) \equiv \tfrac{1}{2}N \sum_{i,j=1}^{n} \zeta_i A_{ij} \zeta_j + T - \sum_{i,j=1}^{n} \zeta_i A_{ij} X_j \ . \tag{21}$$

Construct the free energy per particle corresponding to this Hamiltonian $\mathcal{H}_0(\{\zeta_i\})$ and consider the equations

$$\partial f[\mathcal{H}_0(\{\zeta_i\})] \ / \ \partial\zeta_i = 0 \quad , \tag{22}$$

or equivalently

$$\zeta_j = N^{-1} \langle X_j \rangle_{\mathcal{H}_0(\{\zeta_i\})} \tag{23}$$

Now it is clear that if ζ_i^0, $i = 1, \ldots, n$, is a solution of (23), the corresponding parameters ξ_f^0, $f = 1, \ldots, p$, and η_a^0, $a = 1, \ldots, q$, are solutions of (16) and (17). We have the identity $\mathcal{H}_0(\vec{\xi}^0, \vec{\eta}^0) = \mathcal{H}_0(\{\zeta_i^0\})$ and $f(\vec{\xi}^0, \vec{\eta}^0) = f(\{\zeta_i^0\})$, where $f(\{\zeta_i^0\}) = \lim_{N\to\infty} N^{-1} F[\mathcal{H}_0(\{\zeta_i^0\})]$. Due to the one-to-one correspondence between ξ_f, η_a on the one hand and ζ_i on the other, also the inverse statement is true: if ξ_f^0, $f = 1, \ldots, p$, and η_a^0, $a = 1, \ldots, q$, satisfy (16) and (17), then ζ_i^0, $i = 1, \ldots, n$, satisfy (23). Hence the free energies calculated with either of the two sets of parameters are equal. As a result we can conclude that the free energy f of the system described by the Hamiltonian (1) is equal to

$$f = \lim_{N\to\infty} N^{-1} F[\mathcal{H}_0(\{\zeta_i\})] \ , \tag{24}$$

where

$$\mathcal{H}_0(\{\zeta_i\}) = T - \sum_{i,j=1}^{n} \zeta_i A_{ij} X_j + \tfrac{1}{2} N \sum_{i,j=1}^{n} \zeta_i A_{ij} \zeta_j \ . \tag{25}$$

Here the parameters ζ_i should be chosen as those solutions of the molecular-field equations

$$\zeta_j = N^{-1} \langle X_j \rangle_{\mathcal{H}_0(\{\zeta_i\})} \tag{26}$$

that lead to the lowest value of the free energy.

Finally we remark that the restriction to matrices $\vec{\vec{A}}$ with non-zero eigenvalues is not necessary .

Suppose that the inverse of the matrix of coefficients $\vec{\vec{A}}$ does not exist. Then, for sufficiently small positive values of ε, $\vec{\vec{A}}^{(\varepsilon)} \equiv \vec{\vec{A}} - \varepsilon \vec{\vec{I}}$ has no eigenvalues zero. We write the Hamiltonian as

$$\mathcal{H} = T - \tfrac{1}{2} N^{-1} \sum_{i,j=1}^{n} A_{ij}^{(\varepsilon)} X_i X_j - \tfrac{1}{2} N^{-1} \varepsilon \sum_{i=1}^{n} X_i^2 \ . \tag{27}$$

Considering first only the last term with ferromagnetic quadratic interactions, we know from eq. (4) that the free energy is given by

$$\lim_{N\to\infty} N^{-1} F[\mathcal{H}_\varepsilon] \ , \tag{28}$$

where

$$\mathcal{H}_\varepsilon \equiv \tfrac{1}{2}N \sum_i \rho_i^2 + T - \tfrac{1}{2}N^{-1} \sum_{i,j} A_{ij}^{(\varepsilon)} X_i X_j - \varepsilon^{\frac{1}{2}} \sum_i \rho_i X_i \quad . \tag{29}$$

Here the parameters ρ_i, $i = 1, \ldots, n$ are those solutions of

$$\rho_i = \varepsilon^{\frac{1}{2}} N^{-1} \langle X_i \rangle_{\mathcal{H}_\varepsilon} \tag{30}$$

that lead to the lowest value of the free energy given in (28). The averages in (29) are taken with respect to $\mathcal{H}_\varepsilon$. Applying now the result obtained earlier in this section, as well as eq. (28), we find that the free energy per particle corresponding to the Hamiltonian $\mathcal{H}$, is given by

$$\lim_{N\to\infty} F[\mathcal{H}_0^\varepsilon] \tag{31}$$

where

$$\mathcal{H}_0^\varepsilon \equiv \tfrac{1}{2}N \sum_i \rho_i^2 + T - \sum_{i,j} \zeta_i A_{ij}^{(\varepsilon)} X_j + \tfrac{1}{2}N \sum_{i,j} \zeta_i A_{ij} \zeta_j - \varepsilon^{\frac{1}{2}} \sum_i \rho_i X_i \quad . \tag{32}$$

The parameters ζ_i, $i = 1, \ldots, n$ are the solutions of

$$\zeta_i = N^{-1} \langle X_i \rangle_{\mathcal{H}_0^\varepsilon} \tag{33}$$

that result in the lowest value for the free energy (31). Note that the averages in (33) are taken with respect to $\mathcal{H}_0^\varepsilon$, as defined in (32), and not with respect to the Hamiltonian $\mathcal{H}_\varepsilon$ (cf. (30)). The parameters ρ_i of course still satisfy (30).

If we now let ε go to zero, we see from (30) that the parameters ρ_i become zero, and the result given by (31), (32) and (33) reduces to the previous result for the case that the matrix $\overset{\rightarrow}{\vec{A}}$ has no zero eigenvalues. Obviously, if we had applied a molecular-field approximation to the Hamiltonian (7) we would have arrived at the same result. That is to say, for the class of systems described by the Hamiltonian (7) the molecular-field approximation leads to a rigorous expression for the free energy per particle in the thermodynamic limit.

4. Bragg-Williams formulation

For the purpose of describing molecular-field type approximations often a different formulation is used: the so-called Bragg-Williams formulation.[11,12]

Here too an effective one-particle Hamiltonian is considered which contains a number of parameters that have to be chosen in such a way that the free energy calculated with the one-particle Hamiltonian is minimal. The set of parameters chosen in the B.W.-formulation is, however, different: viz. it consists of the occupation numbers of the eigenstates of the one-particle Hamiltonian. Such a formulation may be of use in dealing with the equations governing the physical behaviour of the systems under consideration.

Therefore we show that the free energy per particle can easily be transformed into an expression of the Bragg-Williams type.

Consider a system of N identical "particles" described by the Hamiltonian (7)

$$\mathcal{H} = T - \tfrac{1}{2}N^{-1} \sum_{i,j=1}^{n} A_{ij}X_iX_j \quad . \tag{34}$$

T, X_i, $i = 1, \ldots, n$ are hermitean one-particle operators, i.e. $T = \Sigma_{k=1}^{N} T(k)$, $X_i = \Sigma_{k=1}^{N} X_i(k)$.

Since the particles are identical we see from (24)-(26) that the free energy per particle in the thermodynamic limit is given by

$$f = f[h(\{\zeta_i\})] + \tfrac{1}{2} \sum_{i,j=1}^{n} \zeta_i A_{ij} \zeta_j \quad , \tag{35}$$

where

$$h(\{\zeta_i\}) = t - \sum_{i,j=1}^{n} \zeta_i A_{ij} x_j \tag{36}$$

Here we have denoted operators for one particle by h, t and x_i, $i = 1,\ldots,n$; so we have, e.g. $T(k) = t$ for all k. The parameters ζ_i should be chosen as those solutions of the molecular-field equations

$$\partial f/\partial \zeta_i = 0 \quad \text{or} \quad \zeta_i = \langle x_i \rangle_{h(\{\zeta_j\})} \tag{37}$$

that lead to the lowest free energy.

Denote the eigenstates of the one-particle Hamiltonian (36) by $|\sigma\rangle$, $\sigma = 1, \ldots, m$, and its corresponding eigenvalues by e_σ, then we can define occupation numbers ρ_σ by

$$\rho_\sigma \equiv e^{-\beta e_\sigma} / \sum_\sigma e^{-\beta e_\sigma} \quad . \tag{38}$$

From (37) we have

$$\zeta_i = \sum_\sigma x_i(\sigma)\rho_\sigma \quad , \tag{39}$$

where $x_i(\sigma)$ is the diagonal element $\langle \sigma | x_i | \sigma \rangle$. Using (38) it is easy to show that

$$-\beta^{-1} \ln \operatorname{tr} e^{-\beta h(\{\zeta_i\})} = \sum_\sigma e_\sigma \rho_\sigma + \beta^{-1} \sum_\sigma \rho_\sigma \ln \rho_\sigma \ . \tag{40}$$

From (39) and (40) it then follows that

$$f = \tfrac{1}{2} \sum_{\sigma,\sigma'} \left\{ \sum_{i,j=1}^{n} A_{ij} x_i(\sigma) x_j(\sigma') \right\} \rho_\sigma \rho_{\sigma'} + \sum_\sigma e_\sigma \rho_\sigma + \beta^{-1} \sum_\sigma \rho_\sigma \ln \rho_\sigma \ . \tag{41}$$

Using the relation

$$e_\sigma = t(\sigma) - \sum_{\sigma'} \sum_{i,j=1}^{n} A_{ij} x_i(\sigma') x_j(\sigma) \quad , \tag{42}$$

we can write (41) also as

$$f = -\tfrac{1}{2} \sum_{\sigma,\sigma'} \left\{ \sum_{i,j=1}^{n} A_{ij} x_i(\sigma) x_j(\sigma') \right\} \rho_\sigma \rho_{\sigma'} + \sum_\sigma t(\sigma)\rho_\sigma + \beta^{-1} \sum_\sigma \rho_\sigma \ln \rho_\sigma \ . \tag{43}$$

If one uses the condition $\Sigma_\sigma \rho_\sigma = 1$ to express, say, ρ_m in terms of the independent variables $\rho_1, \ldots, \rho_{m-1}$, it is not difficult to show from (43) and (38) that the ρ_σ are such that

$$\partial f / \partial \rho_\sigma = 0, \quad \sigma = 1, \ldots, m-1. \tag{44}$$

From (43) and (44) one easily recovers the formulation of section 3. Define parameters ζ_i, $i = 1, \ldots, n$ by

$$\zeta_i = \sum_\sigma x_i(\sigma)\rho_\sigma \quad . \tag{45}$$

From the condition $\partial f/\partial \rho_\sigma = 0$ for the parameters ρ_σ, $\sigma = 1, \ldots, m-1$, in which f is expressed, one finds

$$\rho_\sigma = e^{-\beta e_\sigma} / \sum_{\sigma'} e^{-\beta e_{\sigma'}} \ . \tag{46}$$

Then, of course, the free energy as written down in (43), is identical to $f(\{\zeta_i\})$, cf. (35), and the condition $\zeta_i = \Sigma_\sigma x_i(\sigma)\rho_\sigma = \langle x_i \rangle$ is equivalent to

$$\partial f(\{\zeta_i\}) / \partial \zeta_i = 0 \ . \tag{47}$$

Eq. (41) in conjunction with (42), or eq. (43), expresses the free energy in

terms of the occupation numbers ρ_σ, $\sigma = 1, \ldots, m$. Together with (44) this constitutes what may be called the Bragg-Williams formulation for the molecular-field approximation. Given a set of one-particle states one uses in the description of the free energy as parameters the occupation numbers of the states. In particular one can calculate the free energy at fixed values of $\rho_1, \ldots, \rho_m$. One is then led to a free energy function of the type (43). The actual values of the parameters ρ_σ, $\sigma = 1, \ldots, m$, are now determined to be such that this expression for the free energy is minimal; in particular the parameters should satisfy (44).

The relation to the Hartree-Fock approximation is obvious: there one expresses the free energy in terms of the occupation numbers of a basis of one-particle states. First these states are determined by the condition that the free energy is minimal for fixed occupation numbers. Then the matrix-elements of the effective one-particle Hamiltonian obey Hartree-Fock equations. *) Afterwards the occupation numbers of the one-particle levels are found by requiring the free energy to be minimal.

5. Short-range and long-range interactions

In the case of ferromagnetic quadratic interactions, the generalization given in section 2 shows that it is possible to study the simultaneous occurrence of short-range interactions and of long-range interactions of the separable type. Of course, the practical applicability of this result is restricted to cases where the free energy of the model Hamiltonian $\mathcal{H}_0(\{\vec{\xi}\}) \equiv \frac{N}{2}\vec{\xi}^2 + \mathcal{H}_s - \vec{\xi}\cdot\vec{V}$ can be found for arbitrary $\vec{\xi}$. Since the quadratic interactions can be replaced by an effective field, this means that the free energy of the system with short-range interactions only, must be known in an arbitrary field.

We consider two such systems. First the linear Ising chain, and secondly the linear XY chain, both with nearest neighbour interactions of either ferromagnetic or antiferromagnetic type. To both Hamiltonians we add a long-range quadratic ferromagnetic interaction between the spins.

Consider a system of N spins $\frac{1}{2}$ with Hamiltonian

---

*) These equations are equivalent to the condition that the effective one-particle Hamiltonian should be diagonal with respect to the basis of one-particle states.

$$\mathcal{H} = \mathcal{H}_{Is}(H) - \frac{J_1}{2N} \sum_{i,j=1}^{N} S_i^z S_j^z \quad , \qquad J_1 > 0 \ , \tag{48}$$

where

$$\mathcal{H}_{Is}(H) = - J \sum_{i=1}^{N} S_i^z S_{i+1}^z - H \sum_{i=1}^{N} S_i^z \ , \tag{49}$$

$(S_{N+1}^z \equiv S_1^z)$; J may be positive or negative.

Using the formulation of section 3, we obtain the free energy from the following effective Hamiltonian

$$\mathcal{H}_0(\zeta) = \frac{N}{2} J_1 \zeta^2 + \mathcal{H}_{Is}(H) - J_1 \zeta \sum_{i=1}^{N} S_i^z = \frac{N}{2} J_1 \zeta^2 + \mathcal{H}_{Is}(H + J_1 \zeta) \quad . \tag{50}$$

Now the free energy per particle for the Ising chain in field H is given by [13]

$$f_{Is}(H) = - \frac{J}{4} - \beta^{-1} \ln \left[\cosh \frac{\beta H}{2} + \left\{ \sinh^2 \frac{\beta H}{2} + e^{-\beta J} \right\}^{\frac{1}{2}} \right] \quad . \tag{51}$$

Then the free energy per particle for the system described by the Hamiltonian (48) is

$$f = \min_{\zeta} \left\{ \tfrac{1}{2} J_1 \zeta^2 - \frac{J}{4} - \beta^{-1} \ln \left[\cosh \frac{\beta}{2} (H + J_1 \zeta) + \right.\right.$$
$$\left.\left. + \left\{ \sinh^2 \frac{\beta}{2} (H + J_1 \zeta) - e^{-\beta J} \right\}^{\frac{1}{2}} \right] \right\} \quad . \tag{52}$$

The parameter ζ can be found from the implicit equation

$$\zeta = \tfrac{1}{2} \frac{\sinh \frac{\beta}{2} (H + J_1 \zeta)}{\left\{ \sinh^2 \frac{\beta}{2} (H + J_1 \zeta) + e^{-\beta J} \right\}^{\frac{1}{2}}} \ , \tag{53}$$

and if this equation has more than one solution we must select the solution that leads to the lowest free energy per particle. This model has been considered before by Suzuki [2] and by Nagle [3]. Suzuki argues that the effect of a quadratic ferromagnetic interaction is an extra field in addition to the external field H. In ref. 3 the free energy is obtained from the maximum term in the partition function, taking into account the appropriate combinatorial factors.

For a discussion of the various phase transitions and the critical behaviour we refer to ref. 3.

The second model is described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_{XY}(H) - \frac{J_1}{2N} \sum_{i,j=1}^{N} S_i^z S_j^z \quad , \quad J_1 > 0 \quad , \tag{54}$$

where

$$\mathcal{H}_{XY}(H) = - J \sum_{i=1}^{N} \{(1+\gamma)S_i^x S_{i+1}^x + (1-\gamma)S_i^y S_{i+1}^y\} - H \sum_{i=1}^{N} S_i^z \quad , \tag{55}$$

$(\vec{S}_{N+1} = \vec{S}_1)$; γ, $-1 \leq \gamma \leq 1$, is the anisotropy parameter.
The effective Hamiltonian is $\frac{N}{2} J_1 \zeta^2 + \mathcal{H}_{XY}(H + J_1\zeta)$, and we use the well-known expression for the free energy per spin in the thermodynamic limit for the XY-chain, viz. [14]

$$f_{XY}(H) = - \frac{\beta^{-1}}{2\pi} \int_0^{2\pi} d\phi \ln 2 \cosh \tfrac{1}{2}\beta\Lambda(\phi) \ , \tag{56}$$

with

$$\Lambda(\phi) = \{(H + J \cosh \phi)^2 + J^2\gamma^2 \sin^2\phi\}^{\frac{1}{2}} \quad . \tag{57}$$

For the free energy corresponding to (54) we find

$$f = \min_{\zeta} \{\tfrac{1}{2}J_1\zeta^2 - \beta^{-1} \frac{1}{2\pi} \int_0^{2\pi} d\phi \ln 2 \cosh \tfrac{1}{2}\beta\tilde{\Lambda}(\phi)\} \quad , \tag{58}$$

where

$$\tilde{\Lambda}(\phi) = \{ (H + J_1\zeta + J \cos \phi)^2 + J^2\gamma^2 \sin^2 \phi\}^{\frac{1}{2}} \quad ,. \tag{59}$$

The parameter ζ satisfies

$$\zeta = \frac{1}{4\pi} \int_0^{2\pi} d\phi \ \frac{H + J_1\zeta + J \cos \phi}{\tilde{\Lambda}(\phi)} \ \tanh \tfrac{1}{2}\beta\tilde{\Lambda}(\phi) \quad . \tag{60}$$

The expressions for the isotropic case (γ=0) are given by Suzuki [2]. However, there are some inaccuracies in his paper. Gibberd [4] obtained the eqs. (58)-(60) by noting that the familiar transformations to fermion creation and annihilation operators (i.e. a Jordan-Wigner and a running wave transformation [14],[15]) bring the long-range interaction into a form analogous to the interaction part in the so-called reduced Hamiltonian in the BCS-theory of superconductivity [16]. Since the XY-part is transformed into an expression bilinear in the fermion operators, the resulting Hamiltonian can be diagonalized by techniques used in the theory of superconductivity.

6. Dipolar and quadrupolar ordering

Consider a system of N spins of magnitude S. We suppose that the Hamiltonian is given by

$$\mathcal{H} = -\sum_{i=1}^{N} S_i^z - L\sum_{i=1}^{N} (S_i^z)^2 - \frac{J}{2N}\sum_{i,j=1}^{N} S_i^z S_j^z - \frac{K}{2N}\sum_{i,j=1}^{N} (S_i^z)^2 (S_j^z)^2 . \quad (61)$$

The sign of the constants J, K and L is arbitrary.

The first term is a Zeeman-term; H is the magnetic field. The second term represents a zero-field splitting: the symmetry of the electrostatic crystal field can remove the degeneracy that a given level would have in the case of zero field. In the case of axial symmetry such an effect can be described by a term $-L\Sigma_i (S_i^z)^2$. The third term is an exchange interaction of the usual type, viz. bilinear in the spin operators. The last term is called a quadrupolar *) interaction. There may be different physical reasons for taking into account such a term.

The two interaction parts of the Hamiltonian have both been taken to be of the separable type. One then obtains the same results that one would get by applying a molecular-field approximation to interactions of the type $-\frac{1}{2}\Sigma_{ij} J_{ij} S_i^z S_j^z$ and $-\frac{1}{2}\Sigma_{ij} K_{ij} (S_i^z)^2 (S_j^z)^2$ respectively. The interaction constants J and K are $J(0) \equiv \Sigma_j J_{ij}$ and $K(0) \equiv \Sigma_j K_{ij}$ respectively.

Applying our general formalism we are able to write down immediately the free energy for systems described by Hamiltonians of the class (61). After doing so, we will identify a number of special cases that have been discussed in the literature. Note that we include only the z-components of the spin operators; i.e. we restrict ourselves to Ising-like Hamiltonians. For a discussion of the different types of ordering and transitions in more general systems with dipolar and quadrupolar interactions, within the framework of a molecular-field approximation, see ref. 17.

Introducing two order parameters ζ_1 and ζ_2 we obtain from (24) and (25) the following expression for the free energy per particle in the thermodynamic limit

$$f = \tfrac{1}{2}(J\zeta_1^2 - K\zeta_2^2) - \beta^{-1} \ln \operatorname{tr} \exp\left[\beta(H + J\zeta_1)S^z + \beta(L + K\zeta_2)(S^z)^2\right] . \quad (62)$$

The parameters ζ_1 and ζ_2 are those solutions of

---

*) The terms "dipolar" and "quadrupolar" should not be interpreted in terms of magnetic dipoles or electric quadrupoles.

$$\zeta_1 = \langle S^z \rangle \quad , \quad \zeta_2 = \langle (S^z)^2 \rangle \tag{63}$$

for which the free energy is minimal.

Here the averages are taken with respect to the Hamiltonian

$$h_0(\zeta_1,\zeta_2) = -(H + J\zeta_1)S^z - (L + K\zeta_2)(S^z)^2 \quad .$$

We shall give the explicit results for $S = 1$ and $S = 3/2$. For higher spinvalues $S \geq 2$ the expressions become more complicated. In the case $S = \frac{1}{2}$ the terms in (61) involving K and L reduce to c-numbers. This particular Ising model has been treated by Mühlschlegel and Zittartz [18], and Niemeijer [19]. A lattice gas model equivalent to this has been proposed by Husimi and Temperley [20] and has also been studied by Katsura [21]. Define

$$x = \beta(H + J\zeta_1) \quad \text{and} \quad y = \beta(L + K\zeta_2) \quad . \tag{64}$$

<u>$S = 1$.</u>

The free energy per spin is now

$$f = \tfrac{1}{2}J\zeta_1^2 + \tfrac{1}{2}K\zeta_2^2 - \beta^{-1}\ln(1 + 2e^y \cosh x) \quad , \tag{65}$$

where ζ_1, ζ_2 are the solutions of the equations

$$\zeta_1 = \frac{2e^y \sinh x}{1 + 2e^y \cosh x} \quad , \tag{66}$$

$$\zeta_2 = \frac{2e^y \cosh x}{1 + 2e^y \cosh x} \tag{67}$$

that lead to the lowest free energy.

<u>$S = \frac{3}{2}$.</u>

The free energy per spin is

$$f = \tfrac{1}{2}J\zeta_1^2 + \tfrac{1}{2}K\zeta_2^2 - \tfrac{1}{4}(L + K\zeta_2) - \beta^{-1}\ln(2\cosh\tfrac{x}{2} + 2e^{2y}\cosh\tfrac{3x}{2}) \quad . \tag{68}$$

Again ζ_1, ζ_2 are those solutions of

$$\zeta_1 = \frac{\sinh\frac{x}{2} + 3e^{2y}\sinh\frac{3x}{2}}{2(\cosh\frac{x}{2} + e^{2y}\cosh\frac{3x}{2})} \quad , \tag{69}$$

$$\zeta_2 = \frac{\cosh\frac{x}{2} + 9e^{2y}\cosh\frac{3x}{2}}{2(\cosh\frac{x}{2} + e^{2y}\cosh\frac{3x}{2})} \quad , \tag{70}$$

for which the free energy is minimal.

In connection with the class defined by (61) we mention a spin 1 model with nearest neighbour exchange coupling and zero-field splitting that has been introduced by Capel [22]. The Hamiltonian is

$$\mathcal{H} = -D + D\sum_i (S_i^z)^2 - J\sum_{\langle i,j\rangle} S_i^z S_j^z - \mu H\sum_i S_i^z \quad , \tag{71}$$

where $\langle i,j\rangle$ denotes a pair of neighbours i and j.

In ref. 22 use has been made of the Bragg-Williams approximation and as a result implicit equations are obtained which reduce to (66) and (67) in the special case that $K = 0$. One merely needs to make the following replacement (we shall always give first our parameters) H by μH, L by $-D$, J by zJ (z is the number of nearest neighbours), ζ_1 by σ; ζ_2 does not occur, since $K = 0$. (Also a constant $-D$ must be added to the free energy). For details on the occurrence of first-order and second-order phase transitions we refer to ref. 22. A very similar model has been treated by Blume [23]. Furthermore, Blume and Watson [24] have shown that in the case of equivalent-neighbour interaction, which can be considered as an example of (6.1), the molecular-field approximation is exact.

As a second example we mention the spin $\frac{3}{2}$ version of the same Hamiltonian, discussed by Taggart and Tahir-Kheli [25]. They use a Green functions technique and apply a decoupling procedure equivalent to the Hartree-Fock approximation. Their expressions for the free energy and the molecular-field equations can be found from (68) and (69) by replacing H by $-\mu H$, L by D, J by $2J(0)$ and ζ_1 by σ . Again $K = 0$, so ζ_2 does not occur (In their expression for the free energy a factor 2 in the logarithm is missing).

A model with both order parameters ζ_1 and ζ_2 different from 0 has been investigated by Sivardière and Blume [26] in order to explain the phase transitions observed in $DyVO_4$. Their Hamiltonian reads ($S = \frac{3}{2}$)

$$\mathcal{H} = -h\sum_i S_i^z - \sum_{i,j} J_{ij} S_i^z S_j^z - \sum_{i,j} K_{ij}\left[(S_i^z)^2 - \tfrac{1}{3}S(S+1)\right]\cdot\left[(S^z)^2 - \tfrac{1}{3}S(S+1)\right]. \tag{72}$$

Their results, obtained by a molecular-field approximation, can be found from (68), (69) and (70) by replacing H by h, J by $2J(0)$, K by $2K(0)$,

L by $-\frac{5}{2}K(0)$, ζ_1 by M and ζ_2 by $Q + \frac{5}{4}$. Besides there is a constant $-\frac{25}{16}K(0)$ added to the free energy.

As a final example we mention a spin 1 Ising model that has been constructed by Blume, Emery and Griffiths [27] to discuss the λ transition and phase separation in He^3-He^4 mixtures. The Hamiltonian is, apart from a constant

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i^z S_j^z - K \sum_{\langle i,j \rangle} (S_i^z)^2 (S_j^z)^2 + \Delta \sum_i (S_i^z)^2 . \tag{73}$$

Their results (again obtained using a molecular-field approximation) follow from (65), (66) and (67) by replacing L by $-\Delta$, J by zJ, K by zK, ζ_1 by M and ζ_2 by 1-x. Here the parameter M plays the role of the superfluid order parameter, and x is a parameter reflecting the possibility of phase separation.

7. A model for antiferromagnetism

In 1956 Gorter and van Peski-Tinbergen [28] considered the problem of obtaining a qualitative explanation for the phase diagram of ordinary antiferromagnets. Using a molecular-field approximation on a two-sublattice model they discussed various solutions of the resulting molecular-field equations in the case of orthorombic symmetry and identified the different phases, viz. the antiferromagnetic, the paramagnetic and the so-called spin flop phase. Several special cases of their model have been reconsidered later and have been analyzed in some detail, cf. e.g. Vertogen, de Vries, Kraak [29] and Bidaux, Carrara, Vivet [30]. In order to discuss a model belonging to our general class of systems treated in chapter IV, which gives the same results as the model in ref. 28 in the molecular-field approximation, we consider two magnetic lattices, each containing N spins $\vec{S}_i$, and $\vec{T}_i$ respectively, i = 1, ..., N, of magnitude S and T. The Hamiltonian contains interactions between spins in each lattice and between spins of different lattices. It is given by

$$\mathcal{H} = -\frac{1}{2N} \sum_{i,j} \vec{S}_i \cdot \vec{\mathcal{J}}_1 \cdot \vec{S}_j - \frac{1}{2N} \sum_{i,j} \vec{T}_i \cdot \vec{\mathcal{J}}_2 \cdot \vec{T}_j + \frac{1}{2N} \sum_{i,j} (\vec{S}_i \cdot \vec{\mathcal{J}}_3 \cdot \vec{T}_j + \vec{T}_i \cdot \vec{\mathcal{J}}_3 \cdot \vec{S}_j) -$$

$$- H \sum_i (S_i^z + T_i^z) \quad . \tag{74}$$

The magnetic field has only a z component.

Here we have introduced for the sake of notation interaction matrices $\vec{\vec{J}}_1$, $\vec{\vec{J}}_2$ and $\vec{\vec{J}}_3$ which are diagonal. The diagonal elements can be taken to be different. Following the formulation of section 3, we introduce parameters $\vec{m}_1 \equiv (m_1^x, m_1^y, m_1^z)$ and $\vec{m}_2$. The effective one-particle Hamiltonian is then given by

$$\mathcal{H}_0(\vec{m}_1,\vec{m}_2) = \frac{N}{2}\,(\vec{m}_1\cdot\vec{\vec{J}}_1\cdot\vec{m}_2 + \vec{m}_2\cdot\vec{\vec{J}}_2\cdot\vec{m}_2 - 2\vec{m}_1\cdot\vec{\vec{J}}_3\cdot\vec{m}_2) - \vec{D}\cdot\sum_i \vec{S}_i - \vec{E}\cdot\sum_i \vec{T}_i, \quad (75)$$

where we introduced the effective fields $\vec{D}$ and $\vec{E}$ acting on the spins of the sublattices 1 and 2, respectively

$$\vec{D} = H + \vec{m}_1\cdot\vec{\vec{J}}_1 - \vec{m}_2\cdot\vec{\vec{J}}_3 \quad (76)$$

$$\vec{E} = H + \vec{m}_2\cdot\vec{J}_2 - \vec{m}_1\cdot\vec{\vec{J}}_3 \quad . \quad (77)$$

The free energy is then given by

$$f = \tfrac{1}{2}\vec{m}_1\cdot\vec{\vec{J}}_1\cdot\vec{m}_1 + \tfrac{1}{2}\vec{m}_2\cdot\vec{\vec{J}}_2\cdot\vec{m}_2 - \vec{m}_1\cdot\vec{\vec{J}}_3\cdot\vec{m}_2 - \beta^{-1}\ln\left(\frac{\sinh\beta(S+\frac{1}{2})D}{\sinh\frac{1}{2}\beta D}\right) -$$

$$- \beta^{-1}\ln\left(\frac{\sinh\beta(T+\frac{1}{2})E}{\sinh\frac{1}{2}\beta E}\right) , \quad (78)$$

and the parameters $\vec{m}_1$ and $\vec{m}_2$ are those solutions of

$$\vec{m}_1 = \frac{1}{N}\langle\sum_i \vec{S}_i\rangle_{\mathcal{H}_0} = SB_S(\beta SD)\vec{d} \quad (79)$$

$$\vec{m}_2 = \frac{1}{N}\langle\sum_i \vec{T}_i\rangle_{\mathcal{H}_0} = TB_T(\beta TE)\vec{e} \quad (80)$$

that lead to the lowest value of the free energy. Here $\vec{d}$ and $\vec{e}$ are unit vectors such that

$$\vec{d} = \frac{\vec{D}}{D} \quad \text{and} \quad \vec{e} = \frac{\vec{E}}{E} \quad , \quad (81)$$

and the Brillouin function $B_S(x)$ is defined by

$$B_S(x) = \frac{2S+1}{2S}\coth\frac{2S+1}{2S}x - \frac{1}{2S}\coth\frac{x}{2S} \quad . \quad (82)$$

Defining

$$\lambda = \frac{SB_S(\beta SD)}{D} \quad \text{and} \quad \mu = \frac{TB_T(\beta TE)}{E} \quad , \quad (83)$$

we can rewrite the molecular-field equations as

$$m_1^x = \lambda(m_1^x J_1^x - m_2^x J_3^x)$$

$$m_1^y = \lambda(m_1^y J_1^y - m_2^y J_3^y) \tag{84}$$

$$m_1^z = \lambda(m_1^z J_1^z - m_2^z J_3^z + H) \quad ,$$

and

$$m_2^x = \mu(m_2^x J_2^x - m_1^x J_3^x)$$

$$m_2^y = \mu(m_2^y J_2^y - m_1^y J_3^y) \tag{85}$$

$$m_2^z = \mu(m_2^z J_2^z - m_1^z J_3^z + H) \ .$$

A large variety of ordered phases, transitions and critical points can be obtained from the solutions of these equations and a more complete investigation may be of interest. Such an investigation, however, falls outside the scope of this thesis. Using a C^*-algebra approach Vertogen, de Vries and Kraak [29] have evaluated the free energy corresponding to (74) in the isotropic case, i.e. $\vec{J}_1 = J_1\vec{I}$, $\vec{J}_2 = J_1\vec{I}$, and $\vec{J}_3 = J_3\vec{I}$, where $\vec{I}$ is the unit matrix, for $H = 0$. From (84) and (85) it is seen that the molecular-field equations in that case reduce to

$$\vec{m}_1 = \lambda(J_1\vec{m}_1 - J_3\vec{m}_2)$$

$$\vec{m}_2 = \mu(J_2\vec{m}_2 - J_3\vec{m}_1) \quad , \tag{86}$$

and it follows that the two magnetizations, $\vec{m}_1$ and $\vec{m}_2$, are parallel or antiparallel. Hence we need only two scalar order parameters m_1 and m_2 satisfying

$$m_1 = (J_1m_1 - J_3m_2)\cdot S\cdot B_S[\beta S(J_1m_1 - J_3m_2)](J_1m_1 - J_3m_2)^{-1}$$

$$m_2 = (J_2m_2 - J_3m_1)\cdot T\cdot B_T[\beta T(J_2m_2 - J_3m_2)](J_2m_2 - J_3m_2)^{-1} \ . \tag{87}$$

The free energy per particle is

$$f = \tfrac{1}{2}J_1m_1^2 + \tfrac{1}{2}J_2m_2^2 - J_3m_1m_2 - \beta^{-1}\ln\left(\frac{\sinh[\beta(S+\frac{1}{2})(J_1m_1 - J_3m_2)]}{\sinh\frac{1}{2}\beta(J_1m_1 - J_3m_2)}\right) -$$

$$- \beta^{-1}\ln\left(\frac{\sinh[\beta(T+\frac{1}{2})(J_2m_2 - J_3m_3)]}{\sinh\frac{1}{2}\beta(J_2m_2 - J_3m_3)}\right) .$$

Another special case has been investigated by Bidaux, Carrara and Vivet [30]. They considered the completely anisotropic case (i.e. the Ising version): $J_1^x = J_1^y = J_2^x = J_2^y = J_3^x = J_3^y = 0$, $J_1^z = J_2^z$ for spin $S = T = \frac{1}{2}$. Of course in such a case the possibility of a flop phase is absent. However, in spite of its simplicity the model contains a number of interesting phase transitions. Their implicit equations follow immediately from (87).

8. Ferromagnetism and superconductivity

A simple model for the coexistence of superconductivity and ferromagnetism has been given by Smit, Vertogen and Kraak [31]. Their model consists of a system of conduction electrons with spin $\frac{1}{2}$ and a set of magnetic impurities of spin S. Besides a kinetic energy there are two interactions present, both of the separable type. First an interaction of the well-known BCS-type [16] between the conduction electrons, and secondly an interaction of the Zener s-d type [32] between the itinerant conduction electron spins and the localized impurity spins. From considerations on this kind of models one can get a first impression on the dependence of the occurrence of ferromagnetism and superconductivity on the parameters (e.g. temperature, interaction strenghts) of the model.

The Hamiltonian is

$$\mathcal{H} = \sum_{\vec{k}} (\varepsilon_{\vec{k}} - \mu)(n_{\vec{k}\uparrow} + n_{\vec{k}\downarrow}) - \frac{V}{N} \sum_{\vec{k},\vec{\ell}} a^{\dagger}_{\vec{k}\uparrow} a^{\dagger}_{-\vec{k}\downarrow} a_{-\vec{\ell}\downarrow} a_{\vec{\ell}\uparrow} - \frac{2J}{N} \sum_{\vec{k}} \sum_{m=1}^{cN} S^z_{\vec{k}} S^z_m \,. \qquad (88)$$

Here $\vec{k}$ labels the N electron states (e.g. **Bloch states**); for each $\vec{k}$ there are two spin directions. The second term is the BCS interaction with strength V, and the third term represents a Zener s-d interaction of strength J; c is the concentration of magnetic impurities. Only a zz interaction is taken into account. The BCS-part can be rewritten in a manifestly separable form, cf. e.g. chapter III, section 2; some diagonal terms of order 1 are neglected.

We introduce parameters ρ_1, ρ_2, ζ and ξ corresponding to the operators

$$N^{-1} \sum_{\vec{k}} (a^{\dagger}_{\vec{k}\uparrow} a^{\dagger}_{-\vec{k}\downarrow} + a_{-\vec{k}\downarrow} a_{\vec{k}\uparrow}), \quad iN^{-1} \sum_{\vec{k}} (a^{\dagger}_{\vec{k}\uparrow} a^{\dagger}_{-\vec{k}\downarrow} - a_{-\vec{k}\downarrow} a_{\vec{k}\uparrow}), \quad (cN)^{-1} \sum_m S^z_m$$

and $N^{-1} \sum_{\vec{k}} S^z_{\vec{k}}$, respectively .

The linearized Hamiltonian from which the free energy has to be calculated, reads

$$\mathcal{H}_0 = 2NJ\zeta\xi + \frac{NV}{4}|\rho|^2 - 2J\xi\sum_m S_m^z + \mathcal{H}_0^{e\ell}, \tag{89}$$

where

$$\mathcal{H}_0^{e\ell} \equiv \sum_{\vec{k}} h(\vec{k}) = \sum_{\vec{k}} \Big[(\varepsilon_{\vec{k}}-\mu)(n_{\vec{k}\uparrow}+n_{-\vec{k}\downarrow}) - J\zeta(n_{\vec{k}\uparrow}-n_{-\vec{k}\downarrow}) -$$

$$- \frac{\rho V}{2} a^{\dagger}_{\vec{k}\uparrow} a^{\dagger}_{-\vec{k}\downarrow} - \frac{\rho^* V}{2} a_{-\vec{k}\downarrow} a_{\vec{k}\uparrow} \Big] . \tag{90}$$

Here the parameter ρ is defined as $\rho_1 + i\rho_2$.

In the free energy per particle corresponding to $\mathcal{H}_0$, the term $-2J\xi\Sigma_m S_m^z$ gives a contribution

$$-\beta^{-1} c \ln \left(\frac{\sinh 2\beta J\xi(S+\frac{1}{2})}{\sinh \beta J\xi} \right) . \tag{91}$$

The Hamiltonian $h(\vec{k})$, being bilinear in the operators $a^{\dagger}_{\vec{k}\uparrow}$, $a_{\vec{k}\uparrow}$, $a^{\dagger}_{-\vec{k}\downarrow}$, $a_{-\vec{k}\downarrow}$, can easily be diagonalized, leading to eigenvalues

$$(\varepsilon_{\vec{k}}-\mu) \pm \left\{ (\varepsilon_{\vec{k}}-\mu)^2 + \frac{V^2|\rho|^2}{4} \right\}^{\frac{1}{2}} \quad \text{and} \quad \pm J\zeta + (\varepsilon_{\vec{k}}-\mu) . \tag{92}$$

Then one finds for the free energy per particle corresponding to the Hamiltonian (88)

$$f = 2J\zeta\xi + \frac{V}{4}|\rho|^2 - \beta^{-1} c \ln \left(\frac{\sinh 2\beta J\xi(S+\frac{1}{2})}{\sinh \beta J\xi} \right) +$$

$$+ N^{-1} \sum_{\vec{k}} \varepsilon_{\vec{k}} - \beta^{-1} N^{-1} \sum_{\vec{k}} \ln \left(2 \cosh \beta \left[(\varepsilon_{\vec{k}}-\mu)^2 + \frac{V^2}{4}|\rho|^2 \right]^{\frac{1}{2}} + 2 \cosh \beta J\xi \right). \tag{93}$$

Calculating the derivatives of the right-hand side of (93) we find that the parameters $|\rho|$, ζ, ξ satisfy the equations

$$\zeta = cSB_S(2\beta J\xi S)$$

$$\xi = \frac{1}{2N}\sum_{\vec{k}} \sinh \beta J\zeta \cdot \left\{\cosh \beta J\zeta + \cosh \beta\left[(\varepsilon_{\vec{k}}-\mu)^2 + \frac{V^2}{4}|\rho|^2\right]^{\frac{1}{2}}\right\}^{-1} \tag{94}$$

$$|\rho| = \frac{V}{2N}|\rho| \sum_{\vec{k}} \left[(\varepsilon_{\vec{k}}-\mu)^2 + \frac{V^2|\rho|^2}{4}\right]^{-\frac{1}{2}} \sinh \beta J\zeta \cdot$$

$$\cdot \left\{\cosh \beta J\zeta + \cosh \beta\left[(\varepsilon_{\vec{k}}-\mu)^2 + \frac{V^2}{4}|\rho|^2\right]^{\frac{1}{2}}\right\}^{-1} .$$

In the case that these equations allow more solutions, that solution leading to the lowest free energy must be chosen.

By replacing $|\rho|^2$ by $\frac{4A^2}{V^2}$, ξ by M and ζ by $cS\frac{|J|}{J}$, we see that the equations (93) and (94) reduce to the expressions obtained by Smit, Vertogen and Kraak. (Note that S here is not the magnitude of the impurity spins). These authors considered these equations in two cases: the so-called limit of strong coupling, or of narrow bands ($\varepsilon_{\vec{k}} = \varepsilon$) and the case of a constant density of states in the conduction band. For details about this numerical work we refer to ref. 31.

REFERENCES

1) Riedel, E.K. and Wegner, F.J., Phys.Rev.Lett. 29 (1972) 349.
2) Bogoliubov jr., N.N., Physica 32 (1966) 933.
3) Bogoliubov jr., N.N., A method for studying model Hamiltonians, Pergamon Press (Oxford, 1972).
4) Suzuki, M., J.Phys.Soc. Japan 21 (1966) 2140.
5) Nagle, J.F., Phys.Rev. A2 (1970) 2124.
6) Gibberd, R.W., Aust.J.Phys. 23 (1970) 927.
7) Tindemans, P.A.J. and Capel, H.W., to be published in Physica.
8) Brankov, J.G., Zagrebnov, V.A. and Tonchev, N.S., Preprint JINR R4-7917 Dubna (1974).
9) Brankov, J.G., Zagrebnov, V.A. and Tonchev, N.S., Preprint JINR R4-7735 Dubna (1974).
10) Brankov, J.G., Shumovsky, A.S. and Zagrebnov, V.A., Physica 78 (1974)183.

11) Domb, C., Adv.Phys. 9 (1960) 149.
12) Bragg, W.L. and Williams, E.J., Proc.Roy.Soc. A145 (1934) 699.
13) cf. e.g. Huang, K., Statistical Mechanics, Wiley (New York 1963).
14) Katsura, S., Phys.Rev. 127 (1962) 1508.
15) Lieb, E., Schultz, T. and Mattis, D., Ann.Phys. 16 (1961) 407.
16) Bardeen, J., Cooper, L.N. and Schrieffer, J.R., Phys.Rev. 108 (1957) 1175.
17) Chen, H.H. and Levy, P.M., Phys.Rev. B7 (1973) 4267.
18) Mühlschlegel, B. and Zittartz, H., Z.Phys. 175 (1963) 553.
19) Niemeijer, Th., Physica 48 (1970) 467.
20) Temperley, H.N.V., Proc.Phys.Soc. 67 (1954) 233.
21) Katsura, S., Progr.Theor.Phys. 13 (1955) 571.
22) Capel, H.W., Physica 32 (1966) 966, 33 (1967) 295.
23) Blume, M., Phys.Rev. 141 (1966) 517.
24) Blume, M. and Watson, R.E., J.Appl.Phys. 38 (1967) 991A.
25) Taggert, G.B. and Tahir-Kheli, R.A., Physica 44 (1969) 321.
26) Sivardière, J. and Blume, M., Phys.Rev. B5 (1972) 1126.
27) Blume, M. Emery, V.J. and Griffith, R.B., Phys.Rev. A4 (1971) 1071.
28) Gorter, C.J. and van Peski-Tinbergen, Tineke, Physica 22 (1956) 273.
29) Vertogen, G., de Vries, A.S. and Kraak, J., Physica 63 (1973) 619.
30) Bidaux, R., Carrara, P. and Vivet, B., J.Phys.Chem.Solids 28 (1967)2453.
31) Smit, W.A., Vertogen, G. and Kraak, J., Physica 74 (1974) 97.
32) Zener, C., Phys.Rev. 81 (1951) 440.
33) Mühlschlegel, B., J.Math.Phys. 3 (1962) 522.

SAMENVATTING

Een centraal probleem in de statistische mechanica is de berekening van thermodynamische grootheden, in het bijzonder de vrije energie, uitgaande van de microskopische interacties. Bijna altijd zijn de interacties zo gecompliceerd dat een exacte berekening onmogelijk is. Men neemt dan zijn toevlucht tot benaderingen, en een veel gebruikte en in kwalitatief opzicht succesvolle benadering is de moleculaire-veldbenadering. Daarbij worden de fluctuaties in het instantane veld dat op een gegeven deeltje werkt ten gevolge van zijn wisselwerking met de andere deeltjes, geheel verwaarloosd. De mate waarin deze fluctuaties van belang zijn, zal bepalend zijn voor de betrouwbaarheid van de moleculaire-veldbenadering. In het bijzonder zal zo'n benadering betere resultaten geven naarmate meer deeltjes effectief in wisselwerking zijn met een gegeven deeltje, d.w.z. naarmate de dracht van de wisselwerking langer is.

Men kan nu twee wegen inslaan om deze relatie nader te bestuderen. Enerzijnds kan men interacties van lange, maar eindige dracht beschouwen, waarbij men pas ná het nemen van de thermodynamische limiet de dracht naar oneindig laat gaan. Anderzijds kan men de situatie bekijken waarin de dracht gekoppeld is aan de afmetingen van het systeem. Een voorbeeld dáárvan vormen de separabele interacties die in dit proefschrift bestudeerd zullen worden. Bij een separabele interactie kan men voor alle paren van deeltjes de operator die de interactie tussen twee deeltjes voorstelt, schrijven als het produkt van twee operatoren, voor elk deeltje één. Er kunnen twee soorten separabele interacties onderscheiden worden, die men naar analogie met de situatie van magnetische ordening kan aanduiden als "ferromagnetische" en "antiferromagnetische" interacties. In dit proefschrift zal een exacte berekening van de vrije energie gegeven worden voor een algemene Hamiltoniaan die naast een één-deeltjesstuk (corresponderend met een kinetische-energieterm of een veldterm), een willekeurig, eindig, aantal van beide soorten separabele interacties bevat. Het resultaat is dat de eigenschappen van een dergelijk systeem beschreven worden met een effectieve één-deeltjes Hamiltoniaan, die een zeker aantal zogenaamde ordeparameters bevat. Voor systemen met interacties die niet van het separabele type zijn, bieden de hier weergegeven onderzoekingen de mogelijkheid tot een éénduidige formulering van de moleculaire-veldbenadering. Daarnaast kunnen modellen voor geheel verschillende fysische situaties, maar met eenzelfde mathematische structuur, met elkaar in verband gebracht worden.

De methode die bij het bewijs gebruikt zal worden, sluit aan bij een aanpak die in de statistische mechanica gebruikelijk is: men probeert de toestandssom te schrijven als een integraal van het type $\int e^{-NG}$, die men met behulp van de methode van Laplace berekent als een asymptotische reeks in N^{-1}. De vrije energie per deeltje wordt dan in essentie bepaald door het absolute minimum van G, terwijl de hogere-ordetermen correcties kunnen geven, die van belang zijn voor eindige systemen.

In het quantummechanische geval waarin men te maken heeft met niet-commuterende operatoren, treden complicaties op bij het gebruik van de methode van Laplace. Men ziet dan namelijk dat in feite het aantal integratievariabelen oneindig groot wordt, hetgeen zijn weerslag heeft zowel op het bepalen van het absolute minimum van de functie G (of zijn reële gedeelte), als op het aantonen dat de correcties ten gevolge van de hogere-ordetermen te verwaarlozen zijn. Voor het eerste probleem blijkt een uitbreiding van de Hölder-ongelijkheid tot operatoren noodzakelijk, terwijl het tweede zeer zorgvuldige afschattingen vereist, in het bijzonder als "antiferromagnetische" separabele operatoren aanwezig zijn.

Na een inleidend hoofdstuk vindt men in hoofdstuk II een algemene ongelijkheid voor het spoor van een produkt van matrices. Als één van de toepassingen van deze ongelijkheid wordt de Hölder-ongelijkheid voor operatoren afgeleid die in hoofdstuk III en IV gebruikt wordt bij het berekenen van de vrije energie van systemen met in hoofdstuk III alleen "ferromagnetische", separabele interacties, terwijl in hoofdstuk IV ook "antiferromagnetische" interacties worden toegelaten. In hoofdstuk V tenslotte worden na enkele algemenere beschouwingen, diverse modellen die in de literatuur te vinden zijn, vanuit het hier gegeven algemene standpunt bekeken.

Enige studiegegevens: Na in juni 1965 het eindexamen gymnasium β te hebben afgelegd aan het Bisschoppelijk College "St. Jozef" in Weert, begon ik in september van dat jaar aan de natuurkundestudie te Leiden. Het kandidaatsexamen natuurkunde en wiskunde met sterrekunde legde ik af in december 1967, en het doctoraalexamen theoretische natuurkunde met bijvakken wiskunde en mechanica in juni 1971. Het experimentele werk verrichtte ik in de groep van Prof.dr. R. de Bruyn Ouboter; dit lag op het terrein van de tunnelverschijnselen in supergeleidende metalen. Ik deed mee aan een werkgroep die door enkele studenten en medewerkers van het Kamerlingh Onnes Laboratorium was opgezet om de theoretische aspecten van supergeleiding te bestuderen. De bijzondere aandacht die ik in dat verband aan de BCS-theorie besteedde, resulteerde uiteindelijk in het hier beschreven onderzoek aan algemene systemen met separabele wisselwerkingen. Ik was daartoe sedert oktober 1971 als wetenschappelijk medewerker in dienst van de Stichting F.O.M. en werkzaam in de werkgroep Vaste Stof Theorie Leiden onder leiding van Prof.dr. P.W. Kasteleyn en Dr. H.W. Capel. De Stichting F.O.M. stelde mij in staat enige conferenties en een zomerschool te bezoeken, waaronder de EPS-Conference on "Selected Problems in Magnetism" te Bochum in 1972, de Van der Waals Centennial Conference on Statistical Mechanics te Amsterdam in 1973, en de NUFFIC-zomerschool over "Fundamental Problems in Statistical Mechanics" te Wageningen in 1974.
Daarnaast vervulde ik enige organisatorische functies en was (en ben) ik op een aantal manieren betrokken bij pogingen om meer aandacht te krijgen voor de maatschappelijke aspecten van wetenschap en wetenschapsbeoefening.

Typewerk: Mevrouw S. Hélant Muller-Soegies.