Two-Dimensional Ising Model as a Soluble Problem of Many Fermions

T. D. SCHULTZ AND D. C. MATTIS

International Business Machines Corporation, Thomas J. Watson Research Center, Yorktown Heights, New York

E. H. LIEB

Belfer Graduate School of Science, Yeshiva University, New York, New York

I. INTRODUCTION

The two-dimensional Ising model for a system of interacting spins (or for the ordering of an AB alloy) on a square lattice is one of the very few nontrivial many-body problems that is exactly soluble and shows a phase transition. Although the exact solution in the absence of an external magnetic field was first given almost twenty years ago in a famous paper by Onsager[1] using the theory of Lie algebras, the flow of papers on both approximate and exact methods has remained strong to this day.[2] One reason for this has been the interest in testing approximate methods on an exactly soluble problem. A second reason, no doubt, has been the considerable formidability of the Onsager method. The simplification achieved by Bruria Kaufman[3] using the theory of spinor representations has diminished, but not removed, the reputation of the Onsager approach for incomprehensibility, while the subsequent application of this method by Yang[4] to the calculation of the spontaneous magnetization has, if anything, helped to restore this reputation.

The principal alternative line to these algebraic methods has been to reduce the problem to one of counting polygons on a lattice—the so-called combinatorial method.[5] The first important steps in this direction were achieved by Kac and Ward[6] and by Potts and Ward,[7] although a rigorous combinatorial solution has only recently been given by Hurst and Green[8] and Kasteleyn,[9] the spontaneous magnetization being calculated by Montroll, Potts, and Ward.[10] The combinatorial method, despite its reliance on certain ingenious topological tricks and on relatively unfamiliar entities call Pfaffians, certainly seems simple in contrast to the earlier algebraic methods and may have seemed destined to displace the former as the classical method of solution of this famous problem.

The present paper tries to restore the balance by presenting the algebraic approach in a way that is both very simple and intimately connected with the problem of a soluble many-fermion system. Except for one or two crucial steps, the approach is straightforward and requires no more than a knowledge of the elementary properties of spin $\frac{1}{2}$ and the second quantization formalism for fermions. Certain similarities to the method of linear canonical transformations employed by Bogolubov[11] and Valatin[12] in the theory of superconductivity is obvious but familiarity with these methods is not necessary to understand the present approach. It becomes apparent that the two-dimensional Ising model, rather than being entirely different from the trivially soluble many-body problems, reduces in some ways to one of them, being just the diagonization of a quadratic form.

In Sec. II the transfer matrix formalism, which is the heart of the algebraic approach, is rederived in the conventional way and also in a way that to us seems more natural. In Sec. III the transfer matrix is diagonalized for a lattice wrapped on a torus, using the formalism of second quantization for fermions. In Sec. IV, the two-spin correlation function in an infinite lattice is investigated and the different approaches of previous authors are reconciled. In Sec. V, various definitions of the spontaneous magnetization are contrasted and one of these is used to derive previous results. The rigorous steps heretofore omitted in the

* Present address: Belfer Graduate School of Science, Yeshiva University, New York, New York.

[1] L. Onsager, Phys. Rev. **65**, 117 (1944).

[2] For reviews, see G. F. Newell and E. W. Montroll, Rev. Mod. Phys. **25**, 353 (1953); and C. Domb, Phil. Mag. Suppl. **9**, 151 (1960).

[3] B. Kaufman, Phys. Rev. **76**, 1232 (1949); B. Kaufman and L. Onsager, Phys. Rev. **76**, 1244 (1949). See also Y. Nambu, Progr. Theoret. Phys. (Kyoto) **5**, 1 (1950); K. Husimi and I. Syôzi, Progr. Theoret. Phys. (Kyoto) **5**, 177 (1950); and I. Syôzi Progr. Theoret. Phys. (Kyoto) **5**, 341 (1950).

[4] C. N. Yang, Phys. Rev. **85**, 808 (1952).

[5] B. L. van der Waerden, Z. Physik **118**, 473 (1941).

[6] M. Kac and J. C. Ward, Phys. Rev. **88**, 1332 (1952). See also S. Sherman, J. Math, Phys. **1**, 202 (1960); and P. N. Burgoyne, J. Math. Phys. **4**, 1320 (1963) who have supplied proofs necessary to make the approach of Kac and Ward rigorous.

[7] R. B. Potts and J. C. Ward, Progr. Theoret. Phys. (Kyoto) **13**, 38 (1955).

[8] C. A. Hurst and H. S. Green, J. Chem. Phys. **33**, 1059 (1960). See also A. M. Dykhne and Yu. B. Rumer, Usp. Fiz. Nauk **75**, 101 (1961) [English transl.: Soviet Phys.—Usp. **4**, 698 (1962)].

[9] P. W. Kasteleyn, J. Math. Phys. **4**, 287 (1963).

[10] E. W. Montroll, R. B. Potts, and J. C. Ward, J. Math, Phys. **4**, 308 (1963).

[11] N. N. Bogolubov, Nuovo Cimento **7**, 794 (1958).

[12] J. G. Valatin, Nuovo Cimento **7**, 843 (1958).

calculation of the spontaneous magnetization are all supplied.

The paper ends with a skeleton outline of the various steps required in the demonstration. The reader may find it advantageous at this time to examine this outline, given in Sec. VI, and also to return to it from time to time while following the detailed development of Secs. II–V. He should also avail himself of at least one of the Refs. 1 and 2 for collateral reading, discussion of the specific heat curve, etc.

II. FORMULATION OF PROBLEM

A. Introduction

We consider a set of spin $\frac{1}{2}$'s arranged on a square lattice of M columns and N rows, interacting only with nearest neighbors and with a magnetic field $\mathfrak{H}$. (We, ultimately, let M and N tend to infinity together, i.e., with M/N a fixed ratio, and denote this limit by $\lim_{M,N\to\infty}$.) Let the Hamiltonian be

$$\mathcal{H}(\sigma_{11},\cdots,\sigma_{NM})=-\mathfrak{H}\sum\sigma_{nm}-J_1\sum\sigma_{nm}\sigma_{n+1,m}-J_2\sum\sigma_{nm}\sigma_{n,m+1}. \quad (2.1)$$

(nm) refers to the site in the nth row and mth column, J_1 and J_2 are, respectively, the bond strengths within the columns and within the rows, and we have set the Bohr magneton equal to unity. Each σ_{nm} is a classical variable taking on the values ±1; or equivalently, the σ_{nm} are operators having eigenvalues ±1 (e.g., $\sigma^x{}_{nm}=2S^x{}_{nm}$, where $S^x{}_{nm}$ is the x component of the nmth spin operator). Because the interaction strengths of the horizontal and vertical bonds in the lattice are not assumed the same, one might call $\mathcal{H}$ the Hamiltonian for a "rectangular" Ising lattice. For boundary conditions, we can assume either that the lattice is wrapped on a torus, so that $\sigma_{n,M+1}=\sigma_{n1}$, $\sigma_{N+1,m}=\sigma_{1m}$, and the sums extend to N and M, or we can assume that the lattice has free ends, in which case the sums extend only to $N-1$ and $M-1$. Each boundary condition has its advantages and disadvantages, as we shall see.

The free energy F per spin of the system is obtained from the partition function either by

$$\exp(-\beta FNM)=Z\equiv\sum_{\sigma_{11}=\pm1}\cdots\sum_{\sigma_{NM}=\pm1}\exp[-\beta\mathcal{H}(\sigma_{11},\cdots,\sigma_{NM})] \quad (2.2)$$

if the σ's are considered as classical variables, or by

$$\exp(-\beta FNM)=Z\equiv \mathrm{tr}_{\sigma^x{}_{11}}\cdots\mathrm{tr}_{\sigma^x{}_{NM}}\times\exp[-\beta\mathcal{H}(\sigma^x{}_{11},\cdots,\sigma^x{}_{NM})] \quad (2.2')$$

if the σ's are considered as twice the x components of spin operators.

Also of interest, in calculating the spontaneous magnetization, is a two-spin correlation function defined by

$$\langle\sigma^x{}_{n'm'}\sigma^x{}_{nm}\rangle=Z^{-1}\,\mathrm{tr}_{\sigma^x{}_{11}\cdots\sigma^x{}_{NM}}\sigma^x{}_{n'm'}\sigma^x{}_{nm}\times\exp[-\beta\mathcal{H}(\sigma^x{}_{11}\cdots\sigma^x{}_{NM})]. \quad (2.3)$$

Averages of other functions of the spins can be defined in a similar way.

The problem under consideration is a special case of the more general Ising problem in any number of dimensions defined by the Hamiltonian

$$\mathcal{H}=-\tfrac{1}{2}\sum_{ij}J(\mathbf{R}_i-\mathbf{R}_j)\sigma_{\mathbf{R}_i}\sigma_{\mathbf{R}_j}-\mathfrak{H}\sum_i\sigma_{\mathbf{R}_i}.$$

If the $\sigma_{\mathbf{R}}$ variables were continuous, rather than confined to the two values ±1, then even this Hamiltonian and the corresponding free energy would be trivially soluble by simply introducing running wave variables

$$\sigma_{\mathbf{q}}=(NM)^{-\frac{1}{2}}\sum_i\exp(i\mathbf{q}\cdot\mathbf{R}_i)\sigma_{\mathbf{R}_i}$$

in terms of which the Hamiltonian is immediately diagonalized to

$$\mathcal{H}=-\sum\epsilon_{\mathbf{q}}\sigma^*{}_{\mathbf{q}}\sigma_{\mathbf{q}}-\mathfrak{H}\sigma_0(NM)^{\frac{1}{2}}$$

with

$$\epsilon_q=(2NM)^{-1}\sum J(\mathbf{R}_i-\mathbf{R}_j)\exp[i\mathbf{q}\cdot(\mathbf{R}_i-\mathbf{R}_j)].$$

Unfortunately this running-wave transformation is not available for the actual Ising problem because the constraint $\sigma_{\mathbf{R}_i}=\pm1$ leads to extremely complex conditions on the transformed variables, $\sigma_{\mathbf{q}}$. The spherical model, in which this constraint is replaced by the much weaker constraint $\sum\sigma_{\mathbf{R}}{}^2=NM$, is solvable by introducing running waves, but it is neither an exact solution to the Ising problem nor even a reliable guide. As we see, running waves *can* be introduced into the Ising problem to take advantage of the translational degeneracy, but only after some subtle algebraic preparations which are only possible with nearest neighbor interactions in one or two dimensions. This, then, is the reason for restricting the Hamiltonian to the special form of (2.1).

B. Conventional Derivation of the Transfer Matrix

The fundamental observation on which the algebraic approach is based is that the partition function is the trace of the Nth power of a certain matrix, which we call the *transfer matrix*. In this section we recall the conventional derivation of this result,[13] which has the advantage of algebraic simplicity. In the next section

[13] See the review of Newell and Montroll, Ref. 2. This and related methods were originally discovered by E. W. Montroll, J. Chem. Phys. **9**, 706 (1941); H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252 (1941); E. N. Lassettre and J. P. Howe, J. Chem. Phys. **9**, 747 (1941); and R. Kubo, Busseiron Kenkyu **1**, 1943.

we give an alternate derivation which, although not so simple in its details, is perhaps more natural and straightforward (in the sense that it is formulated in the language of density matrices and reduced density matrices familiar from the many-body problem).

Consider first a one-dimensional cyclic lattice, so that

$$Z=\sum_{\sigma_1,\cdots,\sigma_N}\exp(K_1\Sigma\sigma_n\sigma_{n+1})\exp(H\Sigma\sigma_n), \quad (2.4)$$

where

$$K_1=\beta J_1 \quad \text{and} \quad H=\beta\mathfrak{H}. \quad (2.5)$$

To bring out the fact that the multiple sum is a matrix product, we rewrite (2.4):

$$Z=\sum_{\sigma_1\sigma_1'\cdots\sigma_N\sigma_N'}[\exp(K_1\sigma_1'\sigma_2)][\exp(H\sigma_2)\delta_{\sigma_2\sigma_2'}]\cdots \times[\exp(K_1\sigma_N'\sigma_1)][\exp(H\sigma_1)\delta_{\sigma_1\sigma_1'}]. \quad (2.6)$$

If one defines the two 2×2, matrices

$$(\mathbf{V}_1)_{\sigma_i\sigma_j}=\exp(K_1\sigma_i\sigma_j)$$

or

$$\mathbf{V}_1=\begin{array}{c|cc|c}&\sigma_j=+1&\sigma_j=-1&\\\hline\sigma_i=+1&e^{K_1}&e^{-K_1}&\\\sigma_i=-1&e^{-K_1}&e^{K_1}&\end{array} \quad (2.7a)$$

and

$$(\mathbf{V}_2)_{\sigma_i\sigma_j}=\exp(H\sigma_i)\delta_{\sigma_i\sigma_j}$$

or

$$\mathbf{V}_2=\begin{array}{c|cc|c}&\sigma_j=+1&\sigma_j=-1&\\\hline\sigma_i=+1&e^{H}&0&\\\sigma_i=-1&0&e^{-H}&\end{array}, \quad (2.7b)$$

then Z is obviously the trace of a matrix product:

$$Z=\operatorname{tr}\mathbf{V}_1\mathbf{V}_2\cdots\mathbf{V}_1\mathbf{V}_2=\operatorname{tr}(\mathbf{V}_1\mathbf{V}_2)^N. \quad (2.8)$$

From the invariance of the trace under cyclic permutations of factors, Z can also be written as

$$Z=\operatorname{tr}(\mathbf{V}_2^{\frac{1}{2}}\mathbf{V}_1\mathbf{V}_2^{\frac{1}{2}})^N\equiv\operatorname{tr}\mathbf{V}^N \quad (2.9)$$

or as

$$Z=\operatorname{tr}(\mathbf{V}_1^{\frac{1}{2}}\mathbf{V}_2\mathbf{V}_1^{\frac{1}{2}})\equiv\operatorname{tr}\mathbf{V}'^N, \quad (2.9')$$

which have the advantage that V and V′ are symmetric. We call $\mathbf{V}$ (or $\mathbf{V}'$) the *transfer matrix*.[14]

[14] The transfer matrix, formulated first in general terms by Montroll[13] and also at the same time by Lasettre and Howe,[13] by Kramers and Wannier,[13] and perhaps others, is a powerful operational technique for describing the propagation down an arbitrary line of physical elements. It can be useful even when there is no translational invariance, e.g., the one-dimensional chain of vibrating atoms with random masses analyzed by Schmidt.[15] The one-dimensional Ising model with random bond strengths can also be solved, although the two-dimensional Ising model with some kind of randomness in the bonds has not yet been solved.

[15] H. Schmidt, Phys. Rev. 105, 425 (1957).

If Λ_1 and Λ_2 are the larger and smaller eigenvalues of $\mathbf{V}$, respectively, then

$$Z=\Lambda_1^N+\Lambda_2^N=\Lambda_1^N[1+(\Lambda_2/\Lambda_1)^N]. \quad (2.10)$$

As $N\to\infty$, the free energy per spin is

$$F=-kT\log\Lambda_1, \quad (2.11)$$

the contribution of Λ_2^N being negligible. Thus *finding the free energy reduces to determining the largest eigenvalue of the transfer matrix.*

The matrices $\mathbf{V}_1$, $\mathbf{V}_2$, and $\mathbf{V}$, being 2×2 matrices, can be written in terms of the four Pauli matrices, which we call $\mathbf{1}$, τ^x, τ^y and τ^z (we use τ's rather than σ's to avoid confusion with the σ variables already introduced):

$$\mathbf{V}_1=e^{K_1}\mathbf{1}+e^{-K_1}\tau^x=e^{K_1}(\mathbf{1}+e^{-2K_1}\tau^x), \quad (2.12a)$$

$$\mathbf{V}_2=\mathbf{1}\cosh H+\tau^z\sinh H. \quad (2.12b)$$

It is convenient for the generalization to two dimensions to express $\mathbf{V}_1$ and $\mathbf{V}_2$ as exponentials of the Pauli matrices. For this purpose recall the property of the τ^i, $(i=x, y, z)$, that because $(\tau^i)^2=1$,

$$\exp(a\tau^i)=\mathbf{1}\cosh a+\tau^i\sinh a=(\cosh a)(\mathbf{1}+\tau^i\tanh a) \quad (2.13)$$

for any number a. Then $\mathbf{V}_2$ becomes immediately

$$\mathbf{V}_2=\exp(H\tau^z). \quad (2.14)$$

To simplify $\mathbf{V}_1$ with (2.13), we define K_1^* by[16]

$$\tanh K_1^*\equiv e^{-2K_1}, \quad (2.15a)$$

whereupon

$$\mathbf{V}_1=(2\sinh 2K_1)^{\frac{1}{2}}\exp(K_1^*\tau^x). \quad (2.16)$$

We have used the simple identities

$$\tanh K_1\equiv e^{-2K_1^*} \quad \text{and} \quad \sinh 2K_1\sinh 2K_1^*\equiv1 \quad (2.15b)$$

implied by (2.15a).

The generalization to the two-dimensional problem follows immediately. Instead of summing over the two orientations of each spin, we must sum over the 2^M configurations of each row. The matrix $\mathbf{V}_2$ is still diagonal and can be written

$$\mathbf{V}_2=\exp(K_2\textstyle\sum\tau^z_m\tau^z_{m+1}+H\sum\tau^z_m).$$

Because the two sums in the exponent commute and represent different physical mechanisms, it is customary to replace $\mathbf{V}_2$ in (2.8) or (2.9) by $\mathbf{V}_2\mathbf{V}_3$, where

$$\mathbf{V}_2=\exp(K_2\textstyle\sum\tau^z_m\tau^z_{m+1}) \quad \text{and} \quad \mathbf{V}_3=\exp(H\sum\tau^z_m). \quad (2.16)$$

[16] K_1^* should not be confused with "complex conjugate of K_1", which is real of course.

The matrix $\mathbf{V}_1$ is similarly generalized from the one-dimensional case:

$$\mathbf{V}_1=(2\sinh 2K_1)^{M/2}\exp\,(K_1^*\textstyle\sum\tau^x_m). \quad (2.17)$$

Now the matrices τ^x_m and τ^z_m are $2^M\times 2^M$ matrices defined as direct products[17]

$$\tau^z_m=\mathbf{1}\times\cdots\times\mathbf{1}\times\tau^z\times\mathbf{1}\times\cdots\times\mathbf{1} \quad (2.18a)$$

$$\tau^x_m=\mathbf{1}\times\cdots\times\mathbf{1}\times\tau^x\times\mathbf{1}\times\cdots\times\mathbf{1} \quad (2.18b)$$

and

$$Z=\operatorname{tr}[(\mathbf{V}_2\mathbf{V}_3)^{\frac{1}{2}}\mathbf{V}_1(\mathbf{V}_2\mathbf{V}_3)^{\frac{1}{2}}]^N=\operatorname{tr}\mathbf{V}^N. \quad (2.19)$$

This is the conventional formulation in terms of the transfer matrix. Although very direct, it introduces the matrices τ^z_m and τ^x_m as an artificial device, distinct from the operators σ^z_{nm}, σ^x_{nm}.

C. Alternate Derivation of the Transfer Matrix

Because the τ matrices are isomorphic to the σ operators of a single row, it is possible to derive the transfer matrix formalism without introducing the τ's at all. In this section we present such a derivation, which is in some ways more natural than that of the preceding section, although the results are the same.

Consider a lattice of N rows and M columns, *not* on a torus (although perhaps on a cylinder, so that each row is cyclic). If N is large, it is physically reasonable to expect that the probability of a given configuration of the last row, when one averages over all configurations of the other $N-1$ rows, should be asymptotically independent of N. This suggests trying to find a recursion relation between the probability distribution of configurations of the final row in lattices of N and $N+1$ rows.

To derive such a recursion relation we consider the *density operator* for $N+1$ rows,

$$\rho_N(\Sigma_0,\Sigma_1,\cdots,\Sigma_N)=\exp\Big[\sum_1^N(\mathcal{JC}_n+\mathcal{K}_n^{(1)}+\mathcal{K}_n^{(2)})\Big]\rho_0(\Sigma_0), \quad (2.20)$$

where

$$\mathcal{JC}_n=H\sum_m\sigma^x_{nm},$$

$$\mathcal{K}_n^{(1)}=K_1\sum_m\sigma^x_{n-1,m}\sigma^x_{nm},$$

$$\mathcal{K}_n^{(2)}=K_2\sum_m\sigma^x_{nm}\sigma^x_{n,m+1}, \quad (2.21)$$

[17] If one has two vector spaces with sets of basis vectors denoted, respectively, by $|\alpha\rangle$ and $|a\rangle$ and operators $\mathfrak{A}$ and $\mathbf{A}$ defined, respectively, in these vector spaces, then the direct product $\mathfrak{A}\times\mathbf{A}$ is defined in the vector space with basis vectors denoted by $|\alpha a\rangle$ by the matrix elements

$$\langle\beta b\,|\,\mathfrak{A}\times\mathbf{A}\,|\,\alpha a\rangle=\langle\beta\,|\,\mathfrak{A}\,|\,\alpha\rangle\langle b\,|\,\mathbf{A}\,|\,a\rangle.$$

Thus the operator $\mathfrak{A}$ can be considered in the larger space as the direct product $\mathfrak{A}\times\mathbf{1}$, etc., where $\mathbf{1}$ is the unit operator in the space of the $|a\rangle$'s. The generalization to direct products of more than two operators is obvious.

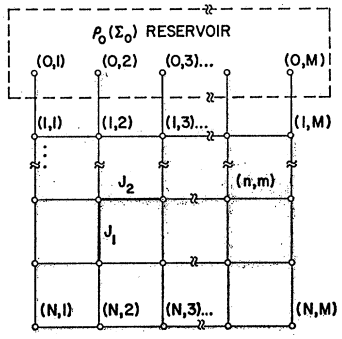

FIG. 1. Two-dimensional Ising lattice of $N+1$ rows and M columns, in which zeroth row is connected to a special reservoir.

and Σ_n is a shorthand symbol for $(\sigma^x_{n1},\cdots,\sigma^x_{nM})$. We assume that the x components, rather than the more conventional z components, of the spins interact with each other and with the magnetic field, to simplify some of the algebra. The density operator for the zeroth row, $\rho_0(\Sigma_0)$, is as yet arbitrary and allows us to see the effect of the conditions at one edge of the lattice on the bulk properties. The model described by (2.20) is shown pictorially in Fig. 1.

The *reduced density operator* for the last row is

$$\rho_N(\Sigma_N)\equiv\operatorname{tr}_{\Sigma_0,\Sigma_1,\ldots,\Sigma_{N-1}}\rho_N(\Sigma_0,\Sigma_1,\cdots,\Sigma_N). \quad (2.22)$$

This operator can be expressed in terms of the reduced-density operator for the last row in an N row system by the obvious recursion relation

$$\rho_N(\Sigma_N)=\exp\,(\mathcal{JC}_N+\mathcal{K}_N^{(2)})\operatorname{tr}_{\Sigma_{N-1}}\times\exp\,(\mathcal{K}_N^{(1)})\rho_{N-1}(\Sigma_{N-1}). \quad (2.23)$$

The trace on the right can be considerably simplified if we observe that, because $(\sigma^x_m)^{2r}=1$ and $(\sigma^x_m)^{2r+1}=\sigma^x_m$, one can think of $\rho_{N-1}(\Sigma_{N-1})$ as being expanded in a canonical form:

$$\rho_{N-1}(\Sigma_{N-1})=a^{(0)}+\sum_{m_1}a^{(1)}{}_{m_1}\sigma^x_{m_1}+\sum_{m_1\neq m_2}a^{(2)}{}_{m_1m_2}\sigma^x_{m_1}\sigma^x_{m_2}+\cdots, \quad (2.24)$$

where we have designated $\sigma^x_{N-1,m}$ by σ^x_m, etc. There are

$$\binom{M}{0}+\binom{M}{1}+\cdots+\binom{M}{M-1}=2^M$$

coefficients in the series (2.24), corresponding to the 2^M configurations of the $(N-1)$st row. Consider the contribution to the trace in (2.23) of a typical term in $\rho_{N-1}(\Sigma_{N-1})$:

$$\operatorname{tr}_\Sigma\exp\,(\mathcal{K}_N^{(1)})\sigma^x_{m_1}\cdots\sigma^x_{m_r}=\operatorname{tr}_\Sigma\prod_m\exp\,(K_1\sigma^x_m{}'\sigma^x_m)\sigma^x_{m_1}\cdots\sigma^x_{m_r} \quad (2.25)$$

where we also designate σ^x_{Nm} by $\sigma^x{}_m'$. Now

$$\mathrm{tr}_{\sigma^x{}_m} \exp\,(K_1\sigma^x{}_m'\sigma^x{}_m) = 2\cosh\,(K_1\sigma^x{}_m') = 2\cosh K_1 \tag{2.26a}$$

and

$$\mathrm{tr}_{\sigma^x{}_m} \exp(K_1\sigma^x{}_m'\sigma^x{}_m)\sigma^x{}_m = 2\sinh\,(K_1\sigma^x{}_m') \tag{2.26b}$$
$$= 2\sigma^x{}_m'\sinh K_1.$$

Thus

$$\mathrm{tr}_\Sigma \exp(\mathcal{K}_N{}^{(1)}\sigma^x{}_{m_1}\cdots\sigma^x{}_{m_r})$$
$$= (2\cosh K_1)^M(\tanh K_1)^r\sigma^x{}_{m_1}'\cdots\sigma^x{}_{m_r}' \tag{2.27}$$

and, if $\rho_{N-1}(\sigma^x{}_1, \cdots, \sigma^x{}_M)$ is considered to be written in its canonical form, then

$$\mathrm{tr}_{\Sigma_{N-1}} \exp(\mathcal{K}_N{}^{(1)})\rho_{N-1}(\sigma^x{}_{N-1,1}, \cdots, \sigma^x{}_{N-1,M})$$
$$= (2\cosh K_1)^M\rho_{N-1}(\sigma^x{}_{N1}\tanh K_1, \cdots, \sigma^x{}_{NM}\tanh K_1). \tag{2.28}$$

We now try to find an operator $\mathcal{O}$ having the property that

$$(2\cosh K_1)^M\rho_{N-1}(\sigma^x{}_{N1}\tanh K_1, \cdots, \sigma^x{}_{NM}\tanh K_1)$$
$$= \mathcal{O}\rho_{N-1}(\sigma^x{}_{N1}, \cdots, \sigma^x{}_{NM}), \tag{2.29}$$

because we would then have the simple recursion relation

$$\rho_N(\Sigma_N) = \exp\,(\mathcal{K}_N + \mathcal{K}^{(2)}{}_N)\,\mathcal{O}\rho_{N-1}(\Sigma_N). \tag{2.30}$$

If $\mathcal{O}$ exists it must clearly have the form

$$(2\cosh K_1)^M(\tanh K_1)^{\mathfrak{N}},$$

where $\mathfrak{N}$ is an operator that counts the number of σ^x's in any term of the canonical expansion. In particular, for the zeroth-order term, we would require that $\mathfrak{N}\cdot 1 = 0$ or $\mathfrak{N} = 0$, which does not meet the requirements for other terms of the expansion. Thus there exists no operator $\mathcal{O}$ in the algebra of the σ^x's that makes (2.29) an operator identity. It is possible, however, to satisfy (2.29) when both sides are applied to a particular state. It is especially convenient (and sufficient to obtain $\mathrm{tr}\,\rho_N(\Sigma_N)$, as we shall see) to take as that state the state with all spins down, which we call the "vacuum" and denote by $|\,0\rangle$:

$$\sigma^-{}_{Nm}\,|\,0\rangle = 0, \quad \text{all } m. \tag{2.31}$$

Then

$$\sigma^x{}_{Nm}\,|\,0\rangle = \sigma^+{}_{Nm}\,|\,0\rangle, \tag{2.32}$$

so that the operator $\mathfrak{N}$, which must count the number of distinct σ^x's present in a product, is just the operator that counts the number of up spins,

$$\mathfrak{N} = \sum_m \sigma^+{}_{Nm}\sigma^-{}_{Nm}. \tag{2.33}$$

In this way, from (2.28) we obtain the recursion relation

$$\rho_N(\Sigma_N)\,|\,0\rangle = \exp\,(\mathcal{K}_N + \mathcal{K}_N{}^{(2)})\,(2\cosh K_1)^M$$
$$\times(\tanh K_1)^{\Sigma\sigma^+{}_m\sigma^-{}_m}\rho_{N-1}(\Sigma_N)\,|\,0\rangle, \tag{2.34}$$

where here, and henceforth, we denote $\sigma^x{}_{Nm}$ simply by $\sigma^x{}_m$, etc. If we define $K_1{}^*$ as before by

$$\tanh K_1 = e^{-2K_1{}^*} \tag{2.15b}$$

we have

$$\rho_N(\Sigma_N)\,|\,0\rangle = (2\cosh K_1)^M \exp\,(\mathcal{K}_N + \mathcal{K}_N{}^{(2)})$$
$$\times \exp\,(-2K_1{}^*\Sigma\sigma^+{}_m\sigma^-{}_m)\rho_{N-1}(\Sigma_N)\,|\,0\rangle. \tag{2.35}$$

Let us now consider $\mathrm{tr}\,\rho_N(\Sigma_N)$ in the representation in which all spins are in the $\pm z$ directions. We notice that any product of one or more distinct σ^x's has no diagonal matrix elements in this representation. Thus only the zeroth-order term in the canonical expansion of $\rho_N(\Sigma_N)$ contributes to the trace, and its contribution is the same for all 2^M states; i.e.,

$$\mathrm{tr}\,\rho_N(\Sigma_N) = 2^M\langle 0\,|\,\rho_N(\Sigma_N)\,|\,0\rangle, \tag{2.36}$$

justifying our remark that knowledge of $\rho_N(\Sigma_N)\,|\,0\rangle$ would be sufficient.

Finally we consider possible choices for ρ_0, each one of which defines a slightly different physical problem. The simplest choice, in which the zeroth row is in a thermal reservoir at infinite temperature is (if we normalize ρ_0) the constant

$$\rho_0 = (\tfrac{1}{2})^M. \tag{2.37}$$

In this case

$$Z = \langle 0\,|\,(V_2V_3V_1)^N\,|\,0\rangle \tag{2.38}$$

with

$$V_1 = (2\cosh K_1)^M \exp\,(-2K_1{}^*\Sigma\sigma^+{}_m\sigma^-{}_m)$$
$$= (2\sinh 2K_1)^{M/2}\exp\,[-2K_1{}^*\Sigma(\sigma^+{}_m\sigma^-{}_m - \tfrac{1}{2})],$$
$$V_2 = \exp\,(K_2\Sigma\sigma^x{}_m\sigma^x{}_{m+1}),$$
$$V_3 = \exp\,(H\Sigma\sigma^x{}_m). \tag{2.39}$$

Another possible choice is

$$\rho_0 = V_2V_3. \tag{2.40}$$

If we let $N\to N-1$, this choice gives an $N\times M$ lattice with no connections to an external reservoir—the classical Ising problem with free edges—which, therefore, has the partition function

$$Z = \langle 0\,|\,(V_2V_3V_1)^{N-1}(V_2V_3)\,|\,0\rangle$$
$$= \langle 0\,|\,(V_2V_3V_1)^NV^{-1}{}_1\,|\,0\rangle \tag{2.41}$$
$$= (2\cosh K_1)^{-M}\langle 0\,|\,(V_2V_3V_1)^N\,|\,0\rangle.$$

In either case the σ's are the operators associated with the spins of the last row, in contrast to the τ matrices of the previous section. To obtain the V's of the last section, we must make the canonical transformation

$$\sigma^x{}_m \to \sigma^z{}_m$$
$$\sigma^z{}_m \to -\sigma^x{}_m$$

and then identify the σ operators with the corresponding τ matrices.

It is again convenient to deal with a symmetric

operator,[18] so we write (2.38) as

$$Z=\langle 0\mid (V_2V_3)^{-\frac{1}{2}}V^N(V_2V_3)^{\frac{1}{2}}\mid 0\rangle, \tag{2.42}$$

where

$$V=(V_2V_3)^{\frac{1}{2}}V_1(V_2V_3)^{\frac{1}{2}}. \tag{2.43}$$

As $N\to\infty$, only the largest eigenvalue of V survives, and factors like $\langle 0\mid (V_2V_3)^{-\frac{1}{2}}\mid \Psi_{\max}\rangle$ have a negligible effect on the free energy per spin, i.e., the slight differences between (2.42) and (2.19) due to the different boundary conditions at the first and last rows are negligible. For convenience we henceforth work with the cyclic analog to (2.42):

$$Z=\operatorname{tr} V^N. \tag{2.44}$$

It should be remarked that spin correlation functions such as (2.4) can be expressed in the transfer matrix formalism quite simply. For example, by similar methods, one easily obtains for the two-spin correlation function

$$\langle\sigma^x{}_{n'm'}\sigma^x{}_{nm}\rangle=\operatorname{tr}(\sigma^x{}_{m'}V^{n'-n}\sigma^x{}_mV^{N-(n'-n)})/\operatorname{tr} V^N. \tag{2.45}$$

III. EIGENVALUES AND EIGENVECTORS OF THE TRANSFER MATRIX

A. Introduction of Fermion Operators

The operators V_1, V_2, and V_3 involve linear and quadratic forms of the spin raising and lowering operators, which obey the mixed set of commutation-anticommutation rules

$$\begin{aligned}&[\sigma^{\pm}{}_m,\sigma^{\pm}{}_n]=0, && m\neq n\\ &\{\sigma^{+}{}_m,\sigma^{-}{}_m\}=1, && (\sigma^{+}{}_m)^2=(\sigma^{-}{}_m)^2=0.\end{aligned} \tag{3.1}$$

The part-boson, part-fermion nature of these rules is a principal source of difficulty, because no simple linear transformation among the σ^+'s and σ^-'s, such as would be required to diagonalize a quadratic form, leaves these rules invariant. It is nevertheless well-known how to change such operators to ones obeying a complete set of anticommutation rules. We introduce annihilation and creation operators by the transformation[19]

$$\begin{aligned}C_m&=[\exp(\pi i\sum_1^{m-1}\sigma^{+}{}_j\sigma^{-}{}_j)]\sigma^{-}{}_m;\\ C_m{}^{\dagger}&=[\exp(\pi i\sum_1^{m-1}\sigma^{+}{}_j\sigma^{-}{}_j)]\sigma^{+}{}_m.\end{aligned} \tag{3.2}$$

The C's and $C^{\dagger}$'s are readily seen to be fermion operators. Then

$$C_m{}^{\dagger}C_m=\sigma^{+}{}_m\sigma^{-}{}_m, \tag{3.3}$$

so that the inverse transformation is simply

$$\begin{aligned}\sigma^{-}{}_m&=[\exp(\pi i\sum_1^{m-1}C_j{}^{\dagger}C_j)]C_m;\\ \sigma^{+}{}_m&=[\exp(\pi i\sum_1^{m-1}C_j{}^{\dagger}C_j)]C_m{}^{\dagger}.\end{aligned} \tag{3.4}$$

Despite its apparently complicated nonlinear structure, this transformation is useful because quadratic forms involving products of operators on the same site, such as occur in V_1, or of operators on adjacent sites, such as occur in V_2, remain quadratic in the new operators. Thus, because of (3.3), V_1 is just

$$V_1=(2\sinh 2K_1)^{M/2}\exp[-2K_1{}^*\textstyle\sum(C_m{}^{\dagger}C_m-\frac{1}{2})]. \tag{3.5}$$

Also, for $m<M$, one finds

$$\begin{aligned}\sigma^{+}{}_m\sigma^{-}{}_{m+1}&=C_m{}^{\dagger}C_{m+1},\\ \sigma^{+}{}_m\sigma^{+}{}_{m+1}&=C_m{}^{\dagger}C_{m+1}{}^{\dagger},\\ \sigma^{-}{}_m\sigma^{-}{}_{m+1}&=-C_mC_{m+1},\end{aligned} \tag{3.6}$$

so that if each row is assumed to have free ends,

$$V_2=\exp[K_2\sum_1^{m-1}(C_m{}^{\dagger}-C_m)(C_{m+1}{}^{\dagger}+C_{m+1})]. \tag{3.7}$$

The operator V_3 is not left in any simple form by the transformation to fermion operators, and it is this feature which has prevented the solution of the Ising problem in a magnetic field. Also, if interactions are between other than nearest neighbors, V_2 involves higher-order terms than quadratic. *For these reasons, we are confined to nearest neighbor interactions and, for the rest of this seztion, to zero external magnetic field.*

If cyclic boundary conditions are imposed, the bond between the last and first spins introduces complications in V_2, because instead of (3.6) we have

$$\begin{aligned}\sigma^{+}{}_M\sigma^{-}{}_1&=-(-)^{\mathfrak{N}}C_M{}^{\dagger}C_1\neq C_M{}^{\dagger}C_1\\ \sigma^{+}{}_M\sigma^{+}{}_1&=-(-)^{\mathfrak{N}}C_M{}^{\dagger}C_1{}^{\dagger}\neq C_M{}^{\dagger}C_1{}^{\dagger}\\ \sigma^{-}{}_M\sigma^{-}{}_1&=(-)^{\mathfrak{N}}C_MC_1\neq -C_MC_1,\end{aligned} \tag{3.8}$$

so that

$$V_2=\exp K_2[\sum_1^{M-1}(C_m{}^{\dagger}-C_m)(C_{m+1}{}^{\dagger}+C_{m+1})-(-)^{\mathfrak{N}}(C_M{}^{\dagger}-C_M)(C_1{}^{\dagger}+C_1)]. \tag{3.9}$$

Here $\mathfrak{N}$ is again the total number operator,

$$\mathfrak{N}=\sum_1^M\sigma^{+}{}_m\sigma^{-}{}_m=\sum_1^MC_m{}^{\dagger}C_m. \tag{3.10}$$

The difficulties created by this more complicated form for V_2 are not insuperable. To see this we observe

[18] Recall that the operator V, which is identical with Montroll and Newell's P, is only one possible symmetrization of the transfer matrix. Another, $V'=V_1{}^{\frac{1}{2}}(V_2V_3)V_1{}^{\frac{1}{2}}$, was used by Yang to derive the spontaneous magnetization. The operator V' is diagonalizable with less algebraic manipulation, but the operator V simplifies expressions for the spontaneous magnetization and the two-spin correlation function. Hence we use the latter. This is discussed further at the end of Sec. IV.

[19] This method was known at least as early as 1928 [See P. Jordan and E. Wigner, Z. Physik **47**, 631 (1928)] and has been rediscovered many times since. It has been used to change spin operators into fermions [E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. **16**, 407 (1961)], to change electrons into bosons with a "hard core" [E. Lieb and D. Mattis, Phys. Rev. **125**, 164 (1962); Appendix], and to change hard core bosons into fermions [T. Schultz, J. Math. Phys. **4**, 666 (1963)].

that because all terms in V_1 and V_2 involve bilinear products of fermion operators, the evenness or oddness of $\mathfrak{N}$ is conserved:

$$[(-)^{\mathfrak{N}}, V_1]=[(-)^{\mathfrak{N}}, V_2]=0. \quad (3.11)$$

We can consider separately eigenstates of V in which only even numbers of fermions are present, and other states in which only odd numbers are present. Acting on the even states, V_2 is equivalent to the operator

$$V_2^+= \exp\left[K_2\sum_1^M (C_m^\dagger - C_m)(C_{m+1}^\dagger + C_{m+1})\right] \quad (3.11a)$$

with

$$C_{M+1}=-C_1 \quad \text{and} \quad C_{M+1}^\dagger=-C_1^\dagger; \quad (3.12a)$$

acting on the odd states, V_2 is equivalent to the operator

$$V_2^-= \exp\left[K_2\sum_1^M (C_m^\dagger - C_m)(C_{m+1}^\dagger + C_{m+1})\right] \quad (3.11b)$$

with

$$C_{M+1}=C_1 \quad \text{and} \quad C_{M+1}^\dagger=C_1^\dagger. \quad (3.12b)$$

Thus we are led to seek the eigenvalues and eigenvectors of the operators

$$V^\pm= (2\sinh 2K_1)^{M/2} \exp\left[\tfrac{1}{2}K_2\sum_1^M (C_m^\dagger - C_m)\right.$$
$$\left.\times (C_{m+1}^\dagger + C_{m+1})\right] \exp\left[-2K_1^*\sum_1^M (C_m^\dagger C_m - \tfrac{1}{2})\right]$$
$$\times \exp\left[\tfrac{1}{2}K_2\sum_1^M (C_m^\dagger - C_m)(C_{m+1}^\dagger + C_{m+1})\right], \quad (3.13)$$

the $\pm$ referring to the anticyclic and cyclic definitions of C_{M+1}, $C_{M+1}^\dagger$, respectively. It is understood that the eigenvectors (and eigenvalues) of V consist of those eigenvectors of V^+ having even numbers of fermions and those eigenvectors of V^- having odd numbers of fermions. *The "odd" eigenvectors of V^+ and the "even" eigenvectors of V^- are irrelevant and must be discarded.*

It should be emphasized that this apparent complication arises solely because of the cyclic boundary conditions; with free ends, there is only one operator V to consider. This implies several advantages for the free-end boundary condition, but the price paid through the loss of translational invariance tends to offset these. In this section, we consider only the case that the rows are cyclic.

Our problem then is to diagonalize the product of three exponentials of quadratic forms in fermion operators. This is somewhat, but not impossibly, more complicated than the diagonalization of a single quadratic form. As one might expect, there exists a general linear transformation to a new set of fermion operators of the type

$$\xi_q= \sum_m g_{qm}C_m + h_{qm}C_m^\dagger, \quad (3.14)$$

in terms of which V appears in the simplest form one could hope to obtain

$$V= \exp\left(-\Sigma \epsilon_q \xi_q^\dagger \xi_q + \text{constant}\right). \quad (3.15)$$

The problem is simply to determine the transformation matrices g_{qm} and h_{qm}, the one-fermion "energies" ϵ_q, and the unknown constant. There are several ways to proceed; we follow one that in our view consists of just a few simple and natural steps.

Because of the translational symmetry in the exponents of (3.5) and (3.13), the first step is obviously to go to running wave operators by the linear canonical transformation

$$C_m= M^{-\frac{1}{2}} e^{-i\pi/4} \sum_q e^{iqm}\eta_q. \quad (3.16)$$

The factor $\exp(-i\pi/4)$ is introduced for future convenience to ensure real coefficients in all terms of $V^\pm$ when expressed in terms of the η_q's. The anticyclic condition (3.12a) requires that

$$q\equiv l=\pm\pi/M, \ \pm 3\pi/M, \ \cdots, \ \pm(M-1)\pi/M, \quad (3.17a)$$

while the cyclic condition (3.12b) requires that

$$q\equiv k=0, \pm 2\pi/M, \pm 4\pi/M, \cdots, \pm(M-2)\pi/M, \pi. \quad (3.17b)$$

We have assumed for convenience that M is even.

Direct substitution of (3.16) into (3.13) yields $V^\pm$ in the form

$$V^\pm= (2\sinh 2K_1)^{M/2} \prod_{0\le q\le \pi} V_q, \quad (3.18)$$

where the q's are either all l's or all k's. For $q\neq 0$ or π,

$$V_q= (V_{2q})^{\frac{1}{2}} V_{1q} (V_{2q})^{\frac{1}{2}} \quad (3.19a)$$

with

$$V_{1q}= \exp\left[-2K_1^*(\eta_q^\dagger\eta_q + \eta_{-q}^\dagger\eta_{-q} - 1)\right], \quad (3.19b)$$

and

$$V_{2q}= \exp\{2K_2[(\cos q)(\eta_q^\dagger\eta_q + \eta_{-q}^\dagger\eta_{-q}) + (\sin q)(\eta_q\eta_{-q} + \eta_{-q}^\dagger\eta_q^\dagger)]\}. \quad (3.19c)$$

For $q=0$,

$$V_0= \exp\left[-2(K_1^*-K_2)(\eta_0^\dagger\eta_0 - \tfrac{1}{2})\right], \quad (3.20)$$

and for $q=\pi$,

$$V_\pi= \exp\left[-2(K_1^*+K_2)(\eta_\pi^\dagger\eta_\pi - \tfrac{1}{2})\right]. \quad (3.21)$$

Because the operators V_q and $V_{q'}$ involve bilinear products of operators that anticommute, they themselves commute and can be diagonalized simultaneously and independently, an enormous simplification. The operators V_0 and V_π are obviously diagonal in the occupation number representation. For more general values of q, the problem reduces to the diagonalization of V_q between the four vectors Φ_0, Φ_q, Φ_{-q}, Φ_{-qq}. The notation is self-explanatory, and the sign of Φ_{-qq} is determined by $\Phi_{-qq}=\eta_{-q}^\dagger\eta_q^\dagger\Phi_0$.

The second step is to observe that the vectors Φ_q and Φ_{-q} are eigenvectors of V_q, being simultaneously eigenvectors of the operators $\eta_q\eta_{-q}$ and $\eta_{-q}^\dagger\eta_q^\dagger$ with eigenvalue zero, and eigenvectors of the operator $\eta_q^\dagger\eta_q+\eta_{-q}^\dagger\eta_{-q}$ with eigenvalue unity. Thus for $q\neq 0$ or π,

$$V_q\Phi_q=\exp(2K_2\cos q)\Phi_q \tag{3.22}$$

and similarly for Φ_{-q}.

The third step is to diagonalize V between the vectors Φ_{-qq} and Φ_0. It is easy to compute the matrix elements of V_{1q} and V_{2q} between these vectors. V_{1q} is already diagonal:

$$V_{1q}=\begin{pmatrix}\exp(-2K_1^*) & 0\\ 0 & \exp(2K_1^*)\end{pmatrix}. \tag{3.23}$$

To compute the matrix elements of $(V_{2q})^{\frac{1}{2}}$ we introduce the pair annihilation and creation operators

$$b^-{}_q=\eta_q\eta_{-q} \quad\text{and}\quad b^+{}_q=\eta_{-q}^\dagger\eta_q^\dagger \tag{3.24}$$

which, in the space of Φ_{-qq} and Φ_0, obey the simple rules of spin-lowering and raising operators. Then, *in this space*, b^- and b^+ are represented by Pauli spin matrices so that

$$\eta_q^\dagger\eta_q+\eta_{-q}^\dagger\eta_{-q}=2b^+{}_qb^-{}_q\equiv b^z{}_q+1 \tag{3.25a}$$

and

$$\eta_q\eta_{-q}+\eta_{-q}^\dagger\eta_q^\dagger=b^x{}_q, \tag{3.25b}$$

so that

$$\begin{aligned}(V_{2q})^{\frac{1}{2}}&=\exp\{K_2[(b^z{}_q+1)\cos q+b^x{}_q\sin q]\}\\&=\exp(K_2\cos q)\exp(K_2b^z{}_q')\\&=\exp(K_2\cos q)(\cosh K_2+b^z{}_q'\sinh K_2)\\&=\exp(K_2\cos q)\\&\times\begin{pmatrix}\cosh K_2+\sin K_2\cos q & \sinh K_2\sin q\\ \sinh K_2\sin q & \cosh K_2-\sinh K_2\cos q\end{pmatrix},\end{aligned} \tag{3.26}$$

where $b^z{}_q'=b^z{}_q\cos q+b^x{}_q\sin q$ and has the important property that $(b^z{}_q')^2=1$. Combining (3.23) and (3.26), we obtain the matrix of V_q between the vectors Φ_{-qq} and Φ_0 in the form

$$V_q=\exp(2K_2\cos q)\begin{pmatrix}A_q & C_q\\ C_q & B_q\end{pmatrix}, \tag{3.27}$$

where

$$\begin{aligned}A_q&=\exp(-2K_1^*)(\cosh K_2+\sinh K_2\cos q)^2\\&\quad+\exp(2K_1^*)(\sinh K_2\sin q)^2\\B_q&=\exp(-2K_1^*)(\sinh K_2\sin q)^2\\&\quad+\exp(2K_1^*)(\cosh K_2-\sinh K_2\cos q)^2\\C_q&=(2\sinh K_2\sin q)\\&\quad\times(\cosh 2K_1^*\cosh K_2-\sinh 2K_1^*\sinh K_2\cos q).\end{aligned}$$

The two eigenvalues are obtained by elementary algebra:

$$\begin{aligned}&\exp(2K_2\cos q)\{\tfrac{1}{2}(A_q+B_q)\pm[(\tfrac{1}{2}(A_q-B_q))^2+C_q^2]^{\frac{1}{2}}\}\\&\qquad=\exp(2K_2\cos q)e^{\pm\epsilon_q},\end{aligned} \tag{3.28}$$

where ϵ_q is the positive root of

$$\begin{aligned}\cosh\epsilon_q&=\cosh 2K_2\cosh 2K_1^*\\&\quad-\sinh 2K_2\sinh 2K_1^*\cos q.\end{aligned} \tag{3.29}$$

The eigenvectors belonging, respectively, to the upper and lower signs in (3.28) are also obtained by elementary algebra:

$$\begin{aligned}\Psi_0&=\cos\phi_q\Phi_0+\sin\phi_q\Phi_{-qq},\\\Psi_{-qq}&=-\sin\phi_q\Phi_0+\cos\phi_q\Phi_{-qq},\end{aligned} \tag{3.30}$$

where ϕ_q is defined modulo π by

$$\tan\phi_q=C_q/(e^{\epsilon_q}-A_q). \tag{3.31}$$

Simple manipulation gives a more convenient definition

$$\tan 2\phi_q=2C_q/(B_q-A_q), \tag{3.32a}$$

which defines ϕ_q modulo $\pi/2$. The solutions of (3.32) that also solve (3.31) satisfy the additional requirement

$$\operatorname{sgn}2\phi_q=\operatorname{sgn}q. \tag{3.32b}$$

The paired nature of the quadratic forms in (3.19) and the structure of the four eigenvectors are extremely reminiscent of the pair-Hamiltonian and the energy states introduced by Bardeen, Cooper, and Schrieffer[20] in the theory of superconductivity, particularly in Anderson's formulation[21] in which electron-pair operators were in fact formally replaced by spin operators [analogous to our $b^\pm{}_q$ of Eq. (3.24)], in the subspace where pairs were not "broken up." We can simplify the results for V_q in precisely the same way that Bogolubov[11] and Valatin[12] simplified the BCS theory, if we introduce the transformation

$$\begin{aligned}\xi_q&=\cos\phi_q\eta_q+\sin\phi_q\eta_{-q}^\dagger,\\\xi_{-q}&=\cos\phi_q\eta_{-q}-\sin\phi_q\eta_q^\dagger.\end{aligned} \tag{3.33}$$

Then it is readily verified that the four eigenvectors are defined by

$$\begin{aligned}&\xi_q\Psi_0=\xi_{-q}\Psi_0=0, \qquad \Phi_q\equiv\Psi_q=\xi_q^\dagger\Psi_0,\\&\Phi_{-q}\equiv\Psi_{-q}=\xi_{-q}^\dagger\Psi_0, \qquad \Phi_{-qq}=\xi_{-q}^\dagger\xi_q^\dagger\Psi_0.\end{aligned} \tag{3.34}$$

In terms of the ξ-operators, V_q becomes

$$V_q=\exp(2K_2\cos q)\exp[-\epsilon_q(\xi_q^\dagger\xi_q+\xi_{-q}^\dagger\xi_{-q}-1)]. \tag{3.35}$$

[20] J. Bardeen, L. Cooper, and J. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[21] P. W. Anderson, Phys. Rev. **112**, 1900 (1958).

The special cases $q=0$ and $q=\pi$ can also be written in this notation if we define

$$\phi_0=0 \quad \text{and} \quad \epsilon_0=2(K_1^*-K_2) \tag{3.36a}$$

$$\phi_\pi=0 \quad \text{and} \quad \epsilon_\pi=2(K_1^*+K_2), \tag{3.36b}$$

so that

$$V_0V_\pi=\exp\left[2K_2(\cos 0+\cos\pi)\right] \times\exp\left[-\epsilon_0(\xi_0^\dagger\xi_0-\tfrac{1}{2})-\epsilon_\pi(\xi_\pi^\dagger\xi_\pi-\tfrac{1}{2})\right]. \tag{3.37}$$

We thus obtain the extremely simple form for $V^\pm$:

$$V^\pm=(2\sinh 2K_1)^{M/2}\exp\left[-\sum_{\text{all } q}\epsilon_q(\xi_q^\dagger\xi_q-\tfrac{1}{2})\right], \tag{3.38}$$

where we have used the fact that $2K_2\sum\cos q=0$.

It is to be noted that a state with an even (odd) number of ξ-particles involves only even (odd) numbers of η-particles. *Thus the allowable eigenstates of V^+ must have even numbers of ξ-particles; of V^-, odd numbers of ξ-particles.* We denote the eigenstates of V^+ and V^- by Ψ^+ and Ψ^- whenever it is useful to distinguish them.

It is also important to observe that for $T\neq T_c$,

$$\epsilon_q\geq\epsilon_{\min}>0, \qquad q\neq 0$$

and

$$\epsilon_0=\pm\lim_{q\to 0}\epsilon_q\gtrless 0 \quad \text{for} \quad T\gtrless T_c. \tag{3.39}$$

Here T_c is defined by $K_1^*=K_2$, or equivalently, by Eq. (2.15b),

$$\sinh(2J_1/kT_c)\sinh(2J_2/kT_c)=1, \tag{3.40}$$

the famous relation for the transition temperature, first derived for the square lattice by Kramers and Wannier[13] prior to Onsager's exact solution.

The partition function, from which one immediately obtains all the well-known thermodynamic properties of the two-dimensional Ising model, is just the Nth power of the largest eigenvalue of V, in analogy with the one-dimensional case.[22] The fact that the other eigenvalues have negligible effect when N, $M\to\infty$ is shown in the argument leading to Eq. (5.21) of Sec. VB. Let us then inquire into the maximum eigenvalue and the corresponding "maximum eigenvector."

The maximum eigenvalue of V^+ is for the ξ_l-particle vacuum, Ψ^+_0, and is

$$\Lambda^+_0=(2\sinh 2K_1)^{M/2}\exp\left[\tfrac{1}{2}\sum_l\epsilon_l\right]. \tag{3.41}$$

The eigenvector Ψ^+_0, having an even number of ξ_l-particles, is allowable and is therefore an eigenvector of V. All other eigenvalues of V^+ are less than Λ^+_0 by factors like $\exp(-\epsilon_{\min})$ or smaller.

The maximum eigenvalue of V^- is either for the ξ_k-particle vacuum, Ψ^-_0, or for the eigenvector with the $k=0$ particle present, $\Psi^-_{k=0}=\xi_{k=0}^\dagger\Psi^-_0$, depending on whether $\epsilon_{k=0}$ is, respectively, positive or negative (i.e., whether $T>T_c$ or $T<T_c$). The largest eigenvalue of V^- in both cases is

$$\Lambda^-_0 \text{ (or, respectively, } \Lambda^-_{k=0})=(2\sinh 2K_1)^{M/2} \times\exp\left[\tfrac{1}{2}\sum_k|\epsilon_k|\right], \tag{3.42}$$

and again all other eigenvalues of V^- are less by factors of $\exp(-\epsilon_{\min})$ or smaller.

Above T_c the maximum eigenvector Ψ^-_0, containing an even number of ξ_k-particles, is not allowable, so the maximum eigenvalue of V is just Λ^+_0 and is nondegenerate. Below T_c, the maximum eigenvector $\Psi^-_{k=0}$, containing an odd number of ξ_k-particles, is also allowable. The two maximum eigenvalues of V, Λ^+_0 and $\Lambda^-_{k=0}$ contain sums that are both approximations to

$$\frac{M}{2\pi}\int_{-\pi}^{\pi}\epsilon_q dq,$$

the integral of a periodic analytic function. The errors in the approximations are both exponentially small[23] for large M, as Onsager originally remarked, because all correction terms in the Euler–MacLaurin expansion vanish. Thus the largest eigenvalues of V^+ and V^- are exponentially degenerate in the limit $M\to\infty$. We call them simply Λ_0:

$$\Lambda_0=(2\sinh 2K_1)^{M/2}\exp\left(\frac{M}{4\pi}\int_{-\pi}^{\pi}\epsilon_q dq\right). \tag{3.43}$$

The effect on the free energy per spin of this degeneracy is $N^{-1}\ln 2$, i.e., negligible, so that one obtains

$$F=-kT\left[\ln(2\sinh 2K_1)^{\frac{1}{2}}+\frac{1}{4\pi}\int_{-\pi}^{\pi}\epsilon_q\,dq\right]. \tag{3.44}$$

The degeneracy nevertheless plays an important role in the long-range order, as was first shown by Ashkin and Lamb,[24] and as we see in Sec. V.

We need not repeat how (3.44) can be written in a form that brings out the similarity between K_1 and K_2, how it leads to a logarithmic singularity in the specific heat, etc., subjects which are discussed by Onsager, Domb, Newell and Montroll, and Huang.[25]

IV. SPIN–SPIN CORRELATION FUNCTION

We consider now the evaluation of the correlation function between two spins in the same row:

$$\langle\sigma^x_{nm}\sigma^x_{nm'}\rangle= \lim_{M,N\to\infty}\operatorname{tr}\sigma^x_{nm}\sigma^x_{nm'}\exp(-\beta\mathcal{H}_0)/\operatorname{tr}\exp(-\beta\mathcal{H}_0). \tag{4.1}$$

Expressed in the transfer matrix formalism, this is

$$\langle\sigma^x_{nm}\sigma^x_{nm'}\rangle=\lim_{M,N\to\infty}(\operatorname{tr}\sigma^x_m\sigma^x_{m'}\cdot V^N/\operatorname{tr}V^N). \tag{4.2}$$

[22] If the largest eigenvalue is twofold degenerate, as it is for $T<T_c$, the partition function is increased by a factor of two, making a negligible change in the free-energy per spin.

[23] See remark of Domb, Ref. 2, p. 194.

[24] J. Ashkin and W. E. Lamb, Phys. Rev. **64**, 159 (1943).

[25] K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, London, 1963), pp. 369 ff.

If we expand in the eigenvectors of V belonging to the eigenvalues Λ_α, we have

$$\langle \sigma^x_{nm}\sigma^x_{nm'}\rangle = \lim_{M,N\to\infty}[(\sum_\alpha \langle\Psi_\alpha \mid \sigma^x_m\sigma^x_{m'} \mid \Psi_\alpha\rangle\Lambda_\alpha{}^N)/\sum_\alpha \Lambda_\alpha{}^N]. \quad (4.3)$$

It is shown in Sec. VB that only the eigenvector(s) belonging to the largest eigenvalue of V contribute in the numerator and denominator when we take the limit M, $N\to\infty$. Thus, for $T<T_c$, when $M\to\infty$,

$$\langle \sigma^x_{nm}\sigma^x_{nm'}\rangle = \lim_{M\to\infty} \tfrac{1}{2}[\langle\Psi^+_0 \mid \sigma^x_m\sigma^x_{m'} \mid \Psi^+_0\rangle_M + \langle\Psi^-_{k=0} \mid \sigma^x_m\sigma^x_{m'} \mid \Psi^-_{k=0}\rangle_M]. \quad (4.4)$$

The two expectation values depend on M through the eigenvectors. We now consider the first of these in detail, and then show that the two are equal.

The evaluation is made easy if we write the product $\sigma^x_m\sigma^x_{m'}$ in terms of fermion operators. From the original transformation (3.4) we have

$$\sigma^x_m\sigma^x_{m'} = C^x_m[\exp(\pi i\sum_m^{m'-1} C_j{}^\dagger C_j)]C^x_{m'} = (iC^y_m)[\exp(\pi i\sum_{m+1}^{m'-1} C_j{}^\dagger C_j)]C^x_{m'}, \quad (4.5a)$$

where

$$C^x_j = C_j{}^\dagger + C_j \quad \text{and} \quad iC^y_j = C_j{}^\dagger - C_j. \quad (4.5b)$$

Observing that

$$(-)^{c_j{}^\dagger c_j} = (C_j{}^\dagger + C_j)(C_j{}^\dagger - C_j) = C^x_j(iC^y_j), \quad (4.6)$$

we obtain the product of $2(m'-m)$ fermion operators

$$\sigma^x_m\sigma^x_{m'} = (iC^y_m)(C^x_{m+1})(iC^y_{m+1})\cdots(C^x_{m'-1}) \times (iC^y_{m'-1})(C^x_{m'}). \quad (4.7)$$

To evaluate the "vacuum" expectation value of this product, we use Wick's theorem, which states that we must associate the operators in pairs (with the operators of each pair written in the original order), replace each pair by its contraction (in this case, vacuum expectation value), multiply the product of these contractions by $(-)^p$ [where $(-)^p$ is the signature of the permutation necessary to bring paired operators next to each other from the original ordering], and sum these products over all pairings. The three kinds of contractions that occur are readily evaluated if we combine the transformation (3.16) and the inverse of (3.33):

$$C_m = M^{-\frac{1}{2}} e^{-i\pi/4} \sum_q e^{iqm}(\cos\phi_q\xi_q - \sin\phi_q\xi_{-q}{}^\dagger) \quad (4.8)$$

with

$$\phi_{-q} \equiv -\phi_q. \quad (4.9)$$

The q's range over either the set of l's or the set of k's, whichever is more convenient. Then[26]

$$\langle\Psi^+_0 \mid C^x_iC^x_j \mid \Psi^+_0\rangle = \langle\Psi^+_0 \mid C^y_iC^y_j \mid \Psi^+_0\rangle = 0, \quad i\neq j \quad (4.10)$$

and

$$\langle\Psi^+_0 \mid iC^y_iC^x_j \mid \Psi^+_0\rangle = -M^{-1}\sum_l \exp[il(i-j)]\exp(-i2\phi_l). \quad (4.11)$$

If one defines

$$a_{ij} = \langle\Psi^+_0 \mid iC^y_iC^x_{j+1} \mid \Psi^+_0\rangle, \quad (4.12)$$

then the sum over all pairings is

$$\langle\Psi^+_0 \mid \sigma^x_m\sigma^x_{m'} \mid \Psi^+_0\rangle = \sum_P (-)^P a_{m,Pm}\cdots a_{m'-1,P(m'-1)} = \begin{vmatrix} a_{m,m} & a_{m,m+1} & \cdots & a_{m,m'-1} \\ a_{m+1,m} & & & \vdots \\ \vdots & & & \\ a_{m'-1,m} & & \cdots & a_{m'-1,m'-1} \end{vmatrix}, \quad (4.13)$$

i.e., the correlation function between two spins $m'-m$ sites apart in the same row is a determinant of order $m'-m$. For the cyclic problem we are considering, we see that

$$a_{ij} = a_{j-i} = -M^{-1}\sum_l \exp[-il(j-i)]\exp[-i(2\phi_l+l)], \quad (4.14)$$

so that the determinant in (4.13) is a Toeplitz determinant (the ijth element depends only on $i-j$) and its asymptotic behavior can be calculated by a theorem of Szegö[27] as generalized and applied by Montroll, Potts, and Ward (MPW).[10] In fact, the determinant (4.13) with the definition (4.14) is identical with theirs, so we do not need to repeat its evaluation.

To see the identity of a_{j-i} defined by (4.14) and MPW's a_r, we need an explicit expression for $\exp(-i2\phi_q)$. From (3.32a) one obtains

$$\exp(-i4\phi_l) = \frac{e^{2il} - 2\operatorname{ctnh} 2K_1{}^*\operatorname{ctnh} K_2e^{il} + \operatorname{ctnh}^2 K_2}{e^{-2il} - 2\operatorname{ctnh} 2K_1{}^*\operatorname{ctnh} K_2e^{-il} + \operatorname{ctnh}^2 K_2},$$

[26] This situation is identical to what we found for the "XY model" of quantum mechanical spins in one dimension, where it is discussed in more detail. See E. Lieb, D. Mattis, and T. Schultz, Ann. Phys. 16, 407 (1961), especially pp. 417 ff.

[27] See V. Grenander and G. Szegö, *Toeplitz Forms and their Applications* (University of California Press, Berkeley, California, 1958). For a proof that is extendable to non-Hermitian kernels, as is the case here, see M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, Inc., New York, 1959).

which can be written in factored form as

$$\exp(-i4\phi_l)=\frac{(e^{il}-x_1)(e^{il}-x_2)}{(e^{-il}-x_1)(e^{-il}-x_2)}, \quad (4.15)$$

with

$$x_1=\text{ctnh } K_1^* \text{ ctnh } K_2>1$$

and

$$x_2=(\text{ctnh } K_2/\text{ctnh } K^*_1)\gtrless 1 \quad \text{for} \quad T\gtrless T_c. \quad (4.16)$$

The appropriate square root of (4.15) consistent with the requirement (3.32b) is

$$\exp(-i2\phi_l)$$

$$=e^{il}\left[\frac{(1-x_1^{-1}e^{il})(1-x_2e^{-il})}{(1-x_1^{-1}e^{-il})(1-x_2e^{il})}\right]^{\frac{1}{2}} \quad \text{for} \quad T<T_c, \quad (4.17a)$$

$$=-\left[\frac{(1-x_1^{-1}e^{il})(1-x_2^{-1}e^{il})}{(1-x_1^{-1}e^{-il})(1-x_2^{-1}e^{-il})}\right]^{\frac{1}{2}} \quad \text{for} \quad T>T_c, \quad (4.17b)$$

where the phase angles of all four factors on the right are taken to be in $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$.

If we identify

$$2\phi_l+l=\delta^*(l), \quad (4.18)$$

then in the limit $M\to\infty$, the elements a_{j-i} and the determinant (4.18) agree exactly with those obtained by Montroll, Potts, and Ward.[28]

The analysis of $\langle\Psi^-_{k=0}\,|\,\sigma^x_m\sigma^x_{m'}\,|\,\Psi^-_{k=0}\rangle$ proceeds in exactly the same way, with all contractions being defined as expectation values in the state $\Psi^-_{k=0}$, rather than in Ψ^+_0. The $k=0$ term in the sum corresponding to (4.14) changes sign, but in the limit $M\to\infty$, this has no effect on a_{j-i}. Therefore

$$\langle\Psi^-_{k=0}\,|\,\sigma^x_m\sigma^x_{m'}\,|\,\Psi^-_{k=0}\rangle=\langle\Psi^+_0\,|\,\sigma^x_m\sigma^x_{m'}\,|\,\Psi^+_0\rangle, \quad (4.19)$$

as asserted earlier.

Below the critical temperature, the asymptotic value of the determinant (4.13) is evaluated by MPW using Szegö's theorem. It turns out to be

$$\lim_{m'-m\to\infty}\lim_{M,N\to\infty}\langle\sigma^x_{nm}\sigma^x_{nm'}\rangle_{MN}$$

$$=\left[1-\frac{(1-\tanh^2 K_1)^2(1-\tanh^2 K_2)^2}{16\tanh^2 K_1\tanh^2 K_2}\right]^{\frac{1}{4}}. \quad (4.20)$$

A remarkable property of this correlation function is that it is symmetric in J_1 and J_2; *the two-spin correlation function along a row and along a column are asymptotically equal even in an anisotropic lattice.*

Above the critical temperature, the asymptotic value can be shown to be zero directly by similar methods,[29] or indirectly by using the results proved in Sec. VA.

It is interesting to recall that Kaufman and Onsager, using similar algebraic methods, nevertheless derived the two-spin correlation function as the sum of two determinants rather than as a single determinant, and that Montroll, Potts, and Ward proved the equivalence of the two results only with considerable effort. The reason for the two possible forms, and their identity, can be seen immediately within the algebraic formalism. The Kaufman and Onsager result, rather than the single determinant, is obtained if one works with the transfer matrix V',

$$V'=V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}}, \quad (4.21)$$

which differs from V only in the way it is symmetrized. Then a similar diagonalization procedure can be followed, leading to a canonical transformation analogous to (4.8) but with an angle ϕ_q' rather than ϕ_q, defined by

$$\tan 2\phi_q'=\frac{\sinh 2K_2\sin q}{\sinh 2K_1^*\cosh 2K_2-\cosh 2K_1^*\sinh 2K_2\cos q}, \quad (4.22a)$$

with the requirement

$$\text{sgn } 2\phi_q'=\text{sgn } q. \quad (4.22b)$$

Instead of (4.2) the correlation function takes the form

$$\langle\sigma^x_{nm}\sigma^x_{nm'}\rangle=\lim_{M,N\to\infty}\frac{\text{tr } V_1^{-\frac{1}{2}}\sigma^x_m\sigma^x_{m'}V_1^{\frac{1}{2}}V'^N}{\text{tr } V'^N}, \quad (4.23)$$

so that one must consider the expectation value

$$\langle\Psi^+_0{}'\,|\,(V_1^{-\frac{1}{2}}\sigma^x_mV_1^{\frac{1}{2}})(V_1^{-\frac{1}{2}}\sigma^x_{m'}V_1^{\frac{1}{2}})\,|\,\Psi^+_0{}'\rangle_M$$

$$=\langle\Psi^+_0{}'\,|\,(\sigma^x_m\cosh K_1^*+i\sigma^y_m\sinh K_1^*)$$

$$\times(\sigma^x_{m'}\cosh K_1^*+i\sigma^y_{m'}\sinh K_1^*)\,|\,\Psi^+_0{}'\rangle_M, \quad (4.24)$$

and a corresponding one for $\Psi^-_{k=0}{}'$, which can be shown to be the same when $M\to\infty$. $\Psi^+_0{}'$ is obviously the eigenvector of V' corresponding to Ψ^+_0 for V, etc.

Of the four terms of the product in (4.24), only the terms involving $\sigma^x_m\sigma^x_{m'}$ and $\sigma^y_m\sigma^y_{m'}$ have nonzero expectation values [the other two terms, when simplified using Wick's theorem, always involve a factor of the form $\langle\Psi^+_0{}'\,|\,C^x_iC^x_j\,|\,\Psi^+_0{}'\rangle$ or $\langle\Psi^+_0{}'\,|\,C^y_iC^y_j\,|\,\Psi^+_0{}'\rangle$, $i\neq j$, which vanish as in Eq. (4.10)], so we obtain

$$\langle\sigma^x_{nm}\sigma^x_{nm'}\rangle=\cosh^2 K_1^*\langle\Psi^+_0{}'\,|\,\sigma^x_m\sigma^x_{m'}\,|\,\Psi^+_0{}'\rangle_{M\to\infty}$$

$$+\sinh^2 K_1^*\langle\Psi^+_0{}'\,|\,(i\sigma^y_m)(i\sigma^y_{m'})\,|\,\Psi^+_0{}'\rangle_{M\to\infty}. \quad (4.25)$$

The two expectation values on the right lead to the two determinants of Kaufman and Onsager.

[28] In $\delta^*(l)$ for $T<T_c$, a term equal to π has been omitted by MPW, but this does not affect the conclusion. Also, to make the agreement manifest, we must interchange K_1 and K_2, K_1^* and K_2^*, etc.

[29] Dr. Montroll and Dr. Potts have kindly informed us that their use of the Szegö–Kac theorem in this case is not correct as it appears in MPW, but that it can be corrected.

V. SPONTANEOUS MAGNETIZATION

A. General Remarks

In contrast to the free energy, the spontaneous magnetization of the Ising model on a square lattice, correctly defined, has never been solved with complete mathematical rigor. Starting from the only sensible definition of the spontaneous magnetization, the methods of Yang, and of Montroll, Potts, and Ward are each forced to make an assumption that has not been rigorously justified. The assumptions appear to be quite different; however, from the similarities between the difficulties encountered in trying to justify them, and the identity of the results obtained, one might conclude that they are closely related. In this section we discuss these assumptions; in the following two sections we discuss the MPW and Yang approaches, respectively, in more detail.

Ordinarily in statistical mechanics, the kind of mathematical questions we wish to examine *are completely ignored.* But in the Ising model, which is exactly soluble in certain respects, it is rewarding to be more careful, and to see precisely what one must assume even here.

For a finite lattice in a magnetic field $\mathfrak{H}$ (in the x direction), the free energy per site and the magnetization (magnetic moment per site) are defined by

$$F_{MN}(\mathfrak{H}) = -(1/\beta MN)\ \ln \operatorname{tr} \exp(-\beta \mathcal{K}_0)\ \exp(\beta \mathfrak{H} \textstyle\sum \sigma^x_{mn}) \tag{5.1}$$

and

$$\mathfrak{M}_{MN}(\mathfrak{H}) = \frac{1}{MN}\ \frac{\operatorname{tr} \sum \sigma^x_{mn} \exp(-\beta \mathcal{K}_0)\ \exp(\beta \mathfrak{H} \Sigma \sigma^x_{ij})}{\operatorname{tr} \exp(-\beta \mathcal{K}_0)\ \exp(\beta \mathfrak{H} \Sigma \sigma^x_{ij})} = -\partial F_{MN}/\partial \mathfrak{H}, \tag{5.2}$$

where $\mathcal{K}_0$ is the Hamiltonian in the absence of a field, and again we have set the Bohr magneton equal to unity. The spontaneous magnetization is in some sense the limit of the magnetization when $\mathfrak{H} \to 0+$, but considerable care must be taken with the limit. Thus, it is clear that the spontaneous magnetization is *not* obtained from

$$\lim_{\mathfrak{H} \to 0+} \mathfrak{M}_{MN}(\mathfrak{H}),$$

which vanishes for all finite M and N, as can be seen by expanding each term of (5.2) in powers of $\mathfrak{H}$. For finite M and N, $F_{MN}(\mathfrak{H})$ is analytic for all real $\mathfrak{H}$ and is quadratic in $\mathfrak{H}$ for sufficiently small $\mathfrak{H}$.

The correct definition of the spontaneous magnetization requires that $\lim_{M,N\to\infty}$ be taken *before* $\lim_{\mathfrak{H}\to 0+}$. That is, *the spontaneous magnetization is defined by*

$$\mathfrak{M}_s = \lim_{\mathfrak{H}\to 0+} \lim_{M,N\to\infty} \mathfrak{M}_{MN}(\mathfrak{H}), \tag{5.3}$$

which definition we take as fundamental. Physically, this definition corresponds to the experimental situation, in which $\mathfrak{H}$ is allowed to become small on a macroscopic scale and $\mathfrak{M}_s$ is the extrapolation of the $\mathfrak{M}(\mathfrak{H})$ curve. The values of $\mathfrak{H}$ determining the extrapolated value are all much greater than M^{-1} or N^{-1} (in units of kT/μ_{Bohr}), so that one must pass to the limit $M, N \to \infty$ *before* letting $\mathfrak{H}$ tend to zero. Mathematically, although the sequence of functions $F_{MN}(\mathfrak{H})$ are all quadratic in $\mathfrak{H}$, for small $\mathfrak{H}$, the limit function

$$F(\mathfrak{H}) = \lim_{M,N\to\infty} F_{MN}(\mathfrak{H}) \tag{5.4}$$

may vary as $|\mathfrak{H}|$ for small $\mathfrak{H}$, and just such a behavior would lead to spontaneous magnetization. Unfortunately, no one has succeeded in calculating the function $F(\mathfrak{H})$ for $\mathfrak{H} \neq 0$.

As we shall see in Sec. VC, Yang calculated a quantity $\mathfrak{M}_{\text{Yang}}$ defined by a different limiting process:

$$\mathfrak{M}_{\text{Yang}} = \Big|\lim_{\alpha\to 0+} \lim_{\substack{M,N\to\infty \\ \mathfrak{H}=\alpha/M}} \mathfrak{M}_{MN}(\mathfrak{H})\Big|. \tag{5.5}$$

It has not been possible to show the equivalence of Eq. (5.5) with the spontaneous magnetization as given by the fundamental definition. We have, however, been able to show that if one takes the limit $\alpha \to \infty$ instead of $\alpha \to 0$, in Eq. (5.5), then the resulting expression is the spontaneous magnetization. Unfortunately, the expression $\lim_{M,N\to\infty} \mathfrak{M}_{MN}(\alpha/M)$ has only been computed for α very small.

The important reason why Yang's definition appears to agree with the fundamental one is that the quadratic dependence of $F_{MN}(\mathfrak{H})$ persists as $\mathfrak{H}$ increases only so long as $\mathfrak{H} = O(1/MN)$; for larger $\mathfrak{H}$, $F_{MN}(\mathfrak{H})$ appears to be linear in $\mathfrak{H}$, and presumably the slope is the same for $\mathfrak{H} = O(1/M)$ and for $\mathfrak{H} = O(1)$.

In order to verify this quadratic behavior for $\mathfrak{H} = O(1/MN)$ and to obtain a third possible definition for spontaneous magnetization on which both the calculations of Yang and of MPW can be based, it is useful to consider the probability density that the magnetization have the value $\mathfrak{M}$ in the absence of a magnetic field:

$$p_{MN}(\mathfrak{M})\, d\mathfrak{M} = \sum_{S(\mathfrak{M})} \exp(-\beta \mathcal{K}_0) \Big/ \sum_{\text{all states}} \exp(-\beta \mathcal{K}_0), \tag{5.6}$$

where $S(\mathfrak{M})$ is the set of all states for which

$$\mathfrak{M} < (MN)^{-1} \sum_{nm} \sigma^x_{nm} < \mathfrak{M} + d\mathfrak{M}.$$

The limiting density is defined as

$$p(\mathfrak{M}) = \lim_{M,N\to\infty} p_{MN}(\mathfrak{M}). \tag{5.7}$$

$p_{MN}(\mathfrak{M})$ is even in $\mathfrak{M}$ by symmetry. Also, because it is the probability density of a macroscopic quantity, one expects it to peak sharply at points we shall call $\pm\mathfrak{M}_0$ and which we subsequently identify with $\pm\mathfrak{M}_s$.[30]

[30] Above the critical temperature, of course, there is only one sharp peak around $\mathfrak{M}_0 = 0$.

The magnetization in a field $\mathfrak{H}$ is then

$$\mathfrak{M}_{MN}(\mathfrak{H})=\frac{\int_{-\infty}^{\infty}\mathfrak{M}p_{MN}(\mathfrak{M})\exp(\beta\mathfrak{H}\mathfrak{M}MN)\,d\mathfrak{M}}{\int_{-\infty}^{\infty}p_{MN}(\mathfrak{M})\exp(\beta\mathfrak{H}\mathfrak{M}MN)\,d\mathfrak{M}}$$

$$=\frac{\int_{0}^{\infty}\mathfrak{M}p_{MN}(\mathfrak{M})\sinh(\beta\mathfrak{H}\mathfrak{M}MN)\,d\mathfrak{M}}{\int_{0}^{\infty}p_{MN}(\mathfrak{M})\cosh(\beta\mathfrak{H}\mathfrak{M}MN)\,d\mathfrak{M}}. \quad (5.8)$$

The quadratic behavior for extremely small $\mathfrak{H}$ follows if $p_{MN}(\mathfrak{M})$ satisfies the very weak condition that it be negligible for $\mathfrak{M}>\mathfrak{M}_0\times$(constant of order unity). Then, if $0<\mathfrak{H}\ll(\beta\mathfrak{M}_0MN)^{-1}=O(1/MN)$, we can expand the hyperbolic functions in (5.8) obtaining

$$\mathfrak{M}(\mathfrak{H})=(\beta\mathfrak{H}MN)\int_0^\infty\mathfrak{M}^2p(\mathfrak{M})\,d\mathfrak{M}\Big/\int_0^\infty\mathfrak{M}p(\mathfrak{M})\,d\mathfrak{M}$$

$$=\beta\mathfrak{H}MN\langle\mathfrak{M}^2\rangle/\langle|\,\mathfrak{M}\,|\rangle; \quad (5.9)$$

corresponding to the quadratic dependence of $F_{MN}(\mathfrak{H})$ on $\mathfrak{H}$. We note that in this linear region of the magnetization curve, the susceptibility is proportional to the size of the lattice.

For $\mathfrak{H}$ of macroscopic size, the hyperbolic functions in (5.8) differ negligibly from $2\exp(\beta\mathfrak{H}\mathfrak{M}MN)$ so that we can write

$$\mathfrak{M}_{MN}(\mathfrak{H})$$

$$=\frac{\int_0^\infty\mathfrak{M}p_{MN}(\mathfrak{M})\exp[\beta\mathfrak{H}(\mathfrak{M}-\mathfrak{M}_0)MN]\,d\mathfrak{M}}{\int_0^\infty p_{MN}(\mathfrak{M})\exp[\beta\mathfrak{H}(\mathfrak{M}-\mathfrak{M}_0)MN]\,d\mathfrak{M}}. \quad (5.10)$$

Now, only for an *extremely* sharply peaked $p_{MN}(\mathfrak{M})$, an example of which might be the form

$$p_{MN}(\mathfrak{M})\sim\exp[-A(\mathfrak{M}-\mathfrak{M}_0)^2]$$
$$+\exp[-A(\mathfrak{M}+\mathfrak{M}_0)^2],\quad\text{with } A=O(MN), \quad (5.11)$$

will the effect of the exponentials in Eq. (5.10) for small but macroscopic $\mathfrak{H}$ be small.[31] Then, we would obtain

$$\mathfrak{M}_s=\lim_{\mathfrak{H}\to0+}\lim_{M,N\to\infty}\mathfrak{M}_{MN}(\mathfrak{H})=\langle|\,\mathfrak{M}\,|\rangle$$

$$=\lim_{M,N\to\infty}2\int_0^\infty\mathfrak{M}p_{MN}(\mathfrak{M})\,d\mathfrak{M}. \quad (5.12)$$

[31] It should be remarked that an ansatz like (5.11) is *much* stronger than the more usual requirement $\langle(\mathfrak{M}-\langle\mathfrak{M}\rangle)^2\rangle=O(1/MN)$ or even analogous requirements on any higher moments up to moments of order MN.

If $p_{MN}(\mathfrak{M})$ is so strongly peaked, it is more convenient to calculate $\mathfrak{M}_s$ from

$$\mathfrak{M}_s^2=\langle\mathfrak{M}^2\rangle=\lim_{M,N\to\infty}2\int_0^\infty\mathfrak{M}^2p_{MN}(\mathfrak{M})\,d\mathfrak{M}, \quad (5.13)$$

which can be considered as a third definition of the spontaneous magnetization. Although the MPW calculation has been, and the Yang calculation can be, based on this definition, it has not been possible to show rigorously that it is equivalent to (5.3), i.e., that $p_{MN}(\mathfrak{M})$ *is* sufficiently peaked.

Let us turn to a more detailed consideration of (5.13), deriving both the MPW method and the Yang method.

B. Spontaneous Magnetization and Long-Range Order

Definition (5.13) for $\mathfrak{M}_s$, when written explicitly, is

$$\mathfrak{M}_s^2=\lim_{M,N\to\infty}\langle[(MN)^{-1}\textstyle\sum\sigma^x_{nm}]^2\rangle_{MN}$$

$$=\lim_{M,N\to\infty}(MN)^{-2}\sum_{nmn'm'}\langle\sigma^x_{n'm'}\sigma^x_{nm}\rangle_{MN}, \quad (5.14)$$

where $\langle\cdots\rangle_{MN}$ denotes a thermal average in zero field for an $M\times N$ lattice. For the cyclic problem, the thermal average of (5.14) depends only on $n'-n$ and $m'-m$, so we can write

$$\mathfrak{M}_s^2=\lim_{M,N\to\infty}(MN)^{-1}\sum_{m'n'}\langle\sigma^x_{n'm'}\sigma^x_{nm}\rangle_{MN}. \quad (5.15)$$

This is the starting point of Montroll, Potts, and Ward. Their use of (5.15) has rested on the observation that the terms in the sum $\sum_{m'n'}$ for $m'-m$ and $n'-n$ less than some large value do not survive the multiplication by $(MN)^{-1}$ and the limit $M, N\to\infty$. One has therefore to consider the long-range order $\langle\sigma^x_{n'm'}\sigma^x_{nm}\rangle_{MN}$ for $M, N\to\infty$ and $n'-n$, $m'-m$ large.

Actually, there are at least two kinds of long-range order: the "*short* long-range order" defined by

$$f(\alpha)=\lim_{\substack{r,s\to\infty\\ s/r=\tan\alpha}}\lim_{M,N\to\infty}\langle\sigma^x_{n+s,m+r}\sigma^x_{nm}\rangle_{MN} \quad (5.16)$$

and calculated for $\alpha=0$ in the previous section, and the "*long* long-range order" defined by

$$g(\mu,\nu)=\lim_{M,N\to\infty}\langle\sigma^x_{n+\nu N,m+\mu M}\sigma^x_{nm}\rangle_{MN},\qquad 0<\mu,\nu<1. \quad (5.17)$$

Previously it has first been assumed that (5.15) can be written in terms of the short long-range order, when in fact it is the long long-range order that ultimately enters in (5.15). It has second been found necessary to assume that $f(\alpha)$ is independent of α, because only for $\alpha=0$ (or $\alpha=\frac{1}{2}\pi$) has $f(\alpha)$ been easy to calculate. The first assumption has been made without any justification. The second is suggested (but hardly proved) by the fact that the explicit calculation (as we have observed) gives $f(0)=f(\frac{1}{2}\pi)$, even for the ani-

sotropic case $J_1 \neq J_2$. We now prove these two assumptions by showing that

$$g(\mu, \nu) \equiv g, \qquad \text{all } \mu > 0, \nu > 0, \tag{5.18a}$$

and

$$f(\alpha) \equiv g, \qquad \text{all } \alpha. \tag{5.18b}$$

Thus[32]

$$\mathfrak{M}_s^2 = \int_0^1 \int_0^1 d\mu \, d\nu \, g(\mu, \nu) = f(0), \tag{5.19}$$

the basic formula usually assumed.

For fixed M and N, we consider $g_{MN}(\mu, \nu)$, which in the transfer matrix formalism becomes

$$g_{MN}(\mu, \nu) = \text{tr } (\sigma^x_{m+\mu M} V^{\nu N} \sigma^x_m V^{(1-\nu)N}) / \text{tr } V^N. \tag{5.20}$$

We assume $T < T_c$, so that the two largest eigenvalues of V are degenerate (more precisely, they differ by $O(e^{-M})$ for large M; such differences are consistently neglected). For $T > T_c$ there is no degeneracy so all proofs are similar but simpler.

The denominator of (5.20) can be accurately bounded because $\epsilon_q \geq \epsilon_{\min} > 0$ for all $q \neq 0$:

$$2\Lambda^N_0 < \text{tr } V^N < 2\Lambda^N_0 [1 + \exp(-N\epsilon_{\min})]^{M-1} < 2\Lambda^N_0 \times \exp[(M-1)\exp(-N\epsilon_{\min})].$$

For large M and N, then, with negligible error,

$$\text{tr } V^N = 2\Lambda^N_0, \tag{5.21}$$

i.e., the sole contribution when N and M are large comes from the two maximum eigenvalues of V.

If we expand the numerator of (5.20) in the eigenvectors of V we have

$$g_{MN}(\mu, \nu) = \tfrac{1}{2} \sum_{\alpha,\beta} \langle \Psi_\alpha \mid \sigma^x_{m+\mu M} \mid \Psi_\beta \rangle \langle \Psi_\beta \mid \sigma^x_m \mid \Psi_\alpha \rangle \left(\frac{\Lambda_\alpha}{\Lambda_0}\right)^{(1-\nu)N} \left(\frac{\Lambda_\beta}{\Lambda_0}\right)^{\nu N}. \tag{5.22}$$

We have suppressed the subscript M on all matrix elements and eigenvalues. Let us separate explicitly terms in which one or both eigenvectors belong to the maximum eigenvalue:

$$\begin{aligned} g_{MN}(\mu, \nu) = {} & \tfrac{1}{2}[\langle \Psi^+_0 \mid \sigma^x_{m+\mu M} \mid \Psi^-_{k=0} \rangle \langle \Psi^-_{k=0} \mid \sigma^x_m \mid \Psi^+_0 \rangle + \text{c.c.}] \\ & + \tfrac{1}{2}\Bigg[\sum_{\beta \neq (k=0)} \langle \Psi^+_0 \mid \sigma^x_{m+\mu M} \mid \Psi^-_\beta \rangle \langle \Psi^-_\beta \mid \sigma^x_m \mid \Psi^+_0 \rangle \left(\frac{\Lambda_\beta}{\Lambda_0}\right)^{\nu N} + \sum_{\beta \neq 0} \langle \Psi^-_{k=0} \mid \sigma^x_{m+\mu M} \mid \Psi^+_\beta \rangle \langle \Psi^+_\beta \mid \sigma^x_m \mid \Psi^-_{k=0} \rangle \left(\frac{\Lambda_\beta}{\Lambda_0}\right)^{\nu N} \Bigg] \\ & + \tfrac{1}{2}\Bigg[\sum_{\alpha \neq (k=0)} \langle \Psi^-_\alpha \mid \sigma^x_{m+\mu M} \mid \Psi^+_0 \rangle \langle \Psi^+_0 \mid \sigma^x_m \mid \Psi^-_\alpha \rangle \left(\frac{\Lambda_\alpha}{\Lambda_0}\right)^{(1-\nu)N} + \sum_{\alpha \neq 0} \langle \Psi^+_\alpha \mid \sigma^x_{m+\mu M} \mid \Psi^-_{k=0} \rangle \langle \Psi^-_{k=0} \mid \sigma^x_m \mid \Psi^+_\alpha \rangle \left(\frac{\Lambda_\alpha}{\Lambda_0}\right)^{(1-\nu)N} \Bigg] \\ & + \tfrac{1}{2}\Bigg[\sum_{\substack{\alpha \neq 0, (k=0) \\ \beta \neq 0, (k=0)}} \langle \Psi_\alpha \mid \sigma^x_{m+\mu M} \mid \Psi_\beta \rangle \langle \Psi_\beta \mid \sigma^x_m \mid \Psi_\alpha \rangle \left(\frac{\Lambda_\alpha}{\Lambda_0}\right)^{(1-\nu)N} \left(\frac{\Lambda_\beta}{\Lambda_0}\right)^{\nu N} \Bigg]. \end{aligned} \tag{5.23}$$

We have explicitly recognized the fact that operators like $\sigma^x_{m+\mu M}$ and σ^x_m only connect states of odd numbers of fermions with states of even numbers.

We now wish to show that, when $M, N \to \infty$, $g_{MN}(\mu, \nu)$ is approximated with arbitrary accuracy by the first bracket on the right. It may be remarked that this first bracket can be simplified because

$$\langle \Psi^-_{k=0} \mid \sigma^x_m \mid \Psi^+_0 \rangle = \langle \Psi^-_{k=0} \mid \sigma^x_{m+\mu M} \mid \Psi^+_0 \rangle$$

by translational invariance, so the two terms within the bracket are the same.

Using Schwarz's inequality, the first term in the second bracket is bounded in absolute value by

$$\exp(-\nu N \epsilon_{\min}) \Big[\sum_{\beta \neq (k=0)} |\langle \Psi^+_0 \mid \sigma^x_{m+\mu M} \mid \Psi^-_\beta \rangle|^2 \Big]^{\frac{1}{2}} \times \Big[\sum_{\beta \neq (k=0)} |\langle \Psi^-_\beta \mid \sigma^x_m \mid \Psi^+_0 \rangle|^2 \Big]^{\frac{1}{2}}. \tag{5.24}$$

If we include the term with $\beta = (k=0)$ in the sums and use the fact that $(\sigma^x_m)^2 = 1$, etc. this is in turn bounded by $\exp(-\nu N \epsilon_{\min})$. Similar arguments hold for the other term in the second bracket and both terms in the third bracket. The fourth bracket is similarly bounded, in absolute value, by $2^M \exp(-N\epsilon_{\min})$. Thus

$$|g_{MN}(\mu, \nu) - |\langle \Psi^+_0 \mid \sigma^x_m \mid \Psi^-_{k=0} \rangle_M|^2| \leq \exp(-\nu N \epsilon_{\min}) + \exp[-(1-\nu)N\epsilon_{\min}] + 2^M \exp(-N\epsilon_{\min}). \tag{5.25}$$

And so, taking the limit $N, M \to \infty$,

$$g(\mu, \nu) \equiv \lim_{M,N \to \infty} |\langle \Psi^+_0 \mid \sigma^x_m \mid \Psi^-_{k=0} \rangle_M|^2 \equiv g, \qquad \nu > 0. \tag{5.26}$$

A similar argument with the transfer matrix going from column to column shows that $g(\mu, \nu)$ is a constant, but restricted only by the requirement $\mu > 0$. Thus (5.26) holds on both the μ and ν axes as well as in the quadrant $\mu > 0$, $\nu > 0$ and

$$|\mathfrak{M}_s| = \lim_{M \to \infty} |\langle \Psi^+_0 \mid \sigma^x_m \mid \Psi^-_{k=0} \rangle_M|. \tag{5.27}$$

[32] Strictly speaking (5.15) is $\mathfrak{M}_s^2 = \lim_{M,N\to\infty} (MN)^{-1} \Sigma_{\mu\nu} g_{MN}(\mu,\nu)$ with x and y ranging over a quasicontinuous mesh. The replacement of the Riemann sum of the function $g_{MN}(\mu,\nu)$ by the Riemann integral of the limit function $g(\mu,\nu)$ can be justified because of the exponentially rapid convergence of the $g_{MN}(\mu,\nu)$ to $g(\mu,\nu)$.

It is this form of (5.15), essentially, from which Yang computed the spontaneous magnetization.

Having proved that the long long-range order is a constant, independent of direction of separation of the two sites, we now observe that this is also true for the short long-range order and in fact the two are the same. We have only to point out that in deriving (5.26) we required only that $\nu N\to\infty$, i.e., that $n'-n\to\infty$, and not that ν remain fixed. Thus $f(\alpha)\equiv g$ for all α, and we have a second equivalent form for the spontaneous magnetization obtained from the short long-range order in a row, the form used by MPW:

$$\mathfrak{M}_s=[\lim_{m'-m\to\infty}\lim_{M,N\to\infty}\langle\sigma^x_{m'}\sigma^x_m\rangle_{MN}]^{\frac{1}{2}}$$

$$=\left[1-\frac{(1-\tanh^2 K_1)^2(1-\tanh^2 K_2)^2}{16\tanh^2 K_1\tanh^2 K_2}\right]^{1/8}. \quad (5.28)$$

We remark that for $T>T_c$ there is only one eigenvector of maximum eigenvalue. There is therefore no term corresponding to the first bracket in (5.23); and so, $g(x,y)\equiv f(\alpha)\equiv 0$, and $\mathfrak{M}_s=0$.

C. Yang's Approach to Spontaneous Magnetization

We now consider Yang's derivation of the spontaneous magnetization to make explicit the implicit assumption that $\mathfrak{H}=O(1/M)$. Yang observed that if the spontaneous magnetization is to be nonvanishing, the partition function and therefore the largest eigenvalue of the transfer matrix must have a first-order contribution from $\mathfrak{H}$ or H. Let us then first derive the expansion of the transfer matrix to this order. Recalling that V_2 and V_3 commute, we can write

$$V(H)=V_3^{\frac{1}{2}}VV_3^{\frac{1}{2}}, \quad (5.29)$$

where by V we mean the transfer matrix in the absence of a magnetic field. Expanding $V_3^{\frac{1}{2}}$,

$$V_3^{\frac{1}{2}}=1+\tfrac{1}{2}H\sum\sigma^x_m+\cdots \quad (5.30)$$

we obtain a formal expansion for $V(H)$:

$$V(H)=V+\tfrac{1}{2}H\{\sum\sigma^x_m, V\}+\cdots. \quad (5.31)$$

We can already see that because $\sum\sigma^x_m$ is an operator whose eigenvalues are of order M, the second term is of order HM times the first term. For H larger than $O(1/M)$ it would be essential to include terms of all orders and hope that the correction to the eigenvalue of V could in some sense be exponentiated (as is almost always the case in many-body theory), i.e., we would expect $\Lambda_0(H)$ to be of the form

$$\Lambda_0(H)=\Lambda_0\exp(b_1|H|+b_2H^2+\cdots)$$

where b_1, b_2, etc. are constants of order M. Since we must content ourselves with the expansion (5.31), we assume that $H=\alpha/M$ and that α ultimately tends to zero. At best only b_1 is then obtained.

We now observe that the diagonal matrix elements of the term in (5.31) that is linear in H, for the eigenvectors of V, all vanish. Thus for $T>T_c$, when the largest eigenvector is nondegenerate, there is no modification of Λ_0 or the free energy to first order in H and so $\mathfrak{M}_s=0$. Below the transition temperature, when the largest eigenvectors are exponentially degenerate, in the limit $M\to\infty$, there is a first-order effect obtained by diagonalizing $V(H)$ between the two maximum eigenvectors. The matrix of $V(H)$, in this limit is

$$V(H)\simeq\begin{pmatrix}\Lambda_0 & \Lambda_0H\langle\Psi^+_0|\sum\sigma^x_m|\Psi^-_{k=0}\rangle\\ \Lambda_0H\langle\Psi^+_0|\sum\sigma^x_m|\Psi^-_{k=0}\rangle^* & \Lambda_0\end{pmatrix} \quad (5.32)$$

and the maximum eigenvalue [belonging to the vector $2^{-\frac{1}{2}}(\Psi^+_0+\Psi^-_{k=0})$] is

$$\Lambda_0(H)\simeq\Lambda_0(1+|H||\langle\Psi^+_0|\sum\sigma^x_m|\Psi^-_{k=0}\rangle|)$$

$$\simeq\Lambda_0\exp(|H||\langle\Psi^+_0|\sum\sigma^x_m|\Psi^-_{k=0}\rangle|), \quad (5.33)$$

from which one derives essentially Yang's expression[33] for the spontaneous magnetization,

$$\mathfrak{M}_s=\lim_{M\to\infty}|\langle\Psi^+_0|M^{-1}\sum_m\sigma^x_m|\Psi^-_{k=0}\rangle_M|. \quad (5.34)$$

[33] Yang's expression actually reduces to

$$\Lambda_0\exp(|H||\langle\Psi_0'|\Sigma\sigma^x_m|\Psi_{k=0}'\rangle|\cosh 2K_1^*),$$

where Ψ_0' and $\Psi_{k=0}'$ are vectors defined for the transfer matrix $V'=V_1^{\frac{1}{2}}V_2V_1^{\frac{1}{2}}$ rather than for $V=V_2^{\frac{1}{2}}V_1V_2^{\frac{1}{2}}$ as presented here.

For cyclic boundary conditions in the rows, the matrix element $\langle\Psi^+_0 \mid \sigma^x_m \mid \Psi^-_{k=0}\rangle$ is strictly independent of m so that (5.34) reduces to

$$\mathfrak{M}_s = \lim_{M\to\infty} \mid \langle\Psi^+_0 \mid \sigma^x_1 \mid \Psi^-_{k=0}\rangle_M \mid \qquad (5.34')$$

agreeing with (5.27).[34]

It is an easy matter to express $\sigma^x_1 = C_1^\dagger + C_1$ in terms of the ξ_l's and $\xi_l^\dagger$'s or in terms of the ξ_k's and $\xi_k^\dagger$'s. The difficulty in evaluating $\langle\Psi^+_0 \mid \sigma^x_1 \mid \Psi^-_{k=0}\rangle$ is that Ψ^+_0 is the Fermi vacuum for the ξ_l particles, while $\Psi^-_{k=0}$ is obtained from the Fermi vacuum for the ξ_k particles. The two vacuum states Ψ^+_0 and Ψ^-_0 are not simply related. This difficulty can be circumvented if one introduces free-end boundary conditions. There is then only one Fermi vacuum, and the second-degenerate state is a simple single-excitation state from that vacuum. Now, however, the translational degeneracy is lost and one must consider the matrix element of σ^x_m, $m \gg 1$, not just of σ^x_1.

---

[34] The condition that $H = \alpha/M$, $\alpha \to 0$, which is required for the exponentiation of (5.33a) to obtain (5.33b), seems unnecessarily restrictive. One wonders if higher terms in the series (5.33a) confirm the exponential form (5.33b). For example, to calculate the H^2 term in $\Lambda_0(H)$, we observe that the H^2 term in the formal expansion (5.31) is

$$\tfrac{1}{4}H^2 \sum_{mm'} (\sigma^x_m V \sigma^x_{m'} + \tfrac{1}{2}\{\sigma^x_m \sigma^x_{m'}, V\}).$$

The H^2 term in the expansion of $\Lambda_0(H)$ has two contributions, the diagonal contribution of this second-order term from the eigenvector $2^{-\frac{1}{2}}(\Psi^+_0 + \psi^-_{k=0})$, and the second-order perturbation theoretic contribution from $\frac{1}{2}H\{\Sigma\sigma^x_m, V\}$. Straightforward calculation gives, to order H^2

$$\Lambda_0(H) = \Lambda_0[1 + \mid H \mid \langle\Psi^+_0\mid \sum_m \sigma^x_m \mid \Psi^-_{k=0}\rangle + \tfrac{1}{2}H^2 \langle\Psi^+_0\mid \sum_{mm'} \sigma^x_m \sigma^x_{m'} \mid \Psi^+_0\rangle + H^2 \sum_\alpha{}' \Lambda_\alpha (\Lambda_0 - \Lambda_\alpha)^{-1} \mid \langle 2^{-\frac{1}{2}}(\Psi^+_0 + \Psi^-_{k=0}) \mid \Sigma\sigma^x_m \mid \Psi_\alpha\rangle\mid^2],$$

where we have used the facts that $\langle\Psi^+_0\mid \sigma^x_m \mid \Psi^-_{k=0}\rangle$ is real and that $\langle\Psi^+_0\mid \Sigma\sigma^x_m\sigma^x_{m'} \mid \Psi^+_0\rangle = \langle\psi^-_{k=0}\mid \Sigma\sigma^x_m\sigma^x_{m'} \mid \Psi^-_{k=0}\rangle$. In the last term, the set of $\{\Psi_\alpha\}$ includes all allowable eigenvectors of V but the two maximum ones.

We observe first that, because of (5.26), the first three terms are, to within corrections terms of order H^2M, precisely what one would obtain if $\Lambda_0(H)$ had the form $\exp(\mid H \mid \langle\Psi^+_0\mid \Sigma\sigma^x_m \mid \psi^-_{k=0}\rangle)$. For the last term, because $\Lambda_\alpha < \Lambda_0$, one can obtain the upper bound

$$(\Lambda_0/\Lambda_1 - 1)^{-1} H^2 \{\langle\Psi^+_0\mid (\Sigma\sigma^x_m)^2 \mid \Psi^+_0\rangle - \mid \langle\Psi^+_0 \mid \sigma^x_m \mid \Psi^-_{k=0}\rangle \mid^2\} = O(H^2M).$$

Thus the asymptotic equality (5.26) is sufficient to verify the exponential form of $\Lambda_0(H)$ up to second order in H when $M \to \infty$. More detailed analysis shows that this holds for terms of higher order in H until terms of order $H^{O(M)}$. This is precisely analogous to the problem of proving the sufficient sharpness of $p_{MN}(\mathfrak{M})$ through analysis of its very high moments.

VI. SUMMARY

In this paper we have tried to reduce the algebraic solution of the Ising model on a square lattice to a sequence of simple steps, so that the whole solution, originally considered so complex and obscure, would be transparent and not require knowledge of special algebraic techniques. It might be useful in conclusion to summarize these steps.

First we introduced the transfer matrix, which relates the density operator of the last row in an N-row lattice to that for an $(N-1)$-row lattice and expressed this matrix in terms of the spin operators for the last row. This step is not limited to the two-dimensional problem (for three dimensions, the "last row" becomes the last plane).

Second, we transformed the spin operators in that row to fermion annihilation and creation operators. Such a transformation is always possible when the spins can be numbered in some order, although the transformed operators are convenient when only nearest neighbors in the ordering interact (thus limiting the method to two dimensions) and when the operators of interest are all bilinear in spin raising and lower operators (thus limiting the method to vanishing magnetic field).

Third, we introduced running-wave fermion operators, made possible by cyclic boundary conditions. The transfer operator is then a product of commuting operators each of which involves simple quadratic forms of the paired type, similar to what occurs in the BCS theory of superconductivity.

Fourth, we diagonalized the transfer matrix by a Bogolubov–Valatin transformation. The complications introduced by using periodic boundary conditions were not fundamental and only required more careful bookkeeping.

Fifth, we used Wick's theorem to show that the correlation function between two spins in the same row is a determinant, and we showed the relation to the results previously obtained in the literature.

Finally, we have analyzed the problem of spontaneous magnetization, pointing out the discrepancies between the correct definition of spontaneous magnetization and those previously used. In the course of this analysis we have shown rigorously that what we have called the *long* and *short* long-range orders are equal and independent of direction.

ACKNOWLEDGMENTS

We should like to thank Dr. D. Jepsen and Dr. J. Slonczewski for several interesting conversations.