Ising-Chain Statistics*

JAMES S. MARSH
The Johns Hopkins University, Baltimore, Maryland
(Received 13 December 1965)

The general n-spin correlation function for the Ising chain is calculated exactly. The perpendicular magnetic susceptibility is calculated for arbitrary values of the parallel magnetic field.

I. INTRODUCTION

THE statistics of the Ising chain are quite simple to calculate exactly, yet do not seem to be widely known. By statistics, we mean the probability of any particular arrangement of spins or, equivalently, the value of the correlation function involving any set of spins in the chain.

The knowledge of these is useful in several cases. There seems to be an application to at least one type of physical system, namely, rare-earth chloride crystals,[1,2] and presumably results from the Ising-chain problem could be applied with similar good effect to other axial systems. It is useful to have a nontrivial, solvable, system in which one can check approximations that must be made for real systems. For instance, Lee[3] has checked the approximations made by McMillan and Opechowski[4] in a line-shift and shape calculation by repeating the derivation exactly for the Ising-chain case.

In this paper we expose the Ising-chain statistics and apply them to two cases: We calculate the perpendicular magnetic susceptibility of the Ising chain for arbitrary values of the parallel magnetic field, and we show how Lee's calculation may be done in a few lines. This is made possible by an apparently novel point of view regarding the necessary statistical sums which allows us to evaluate them by inspection in terms of the joint probability of the various arrangements of three spins in a row. The same sort of considerations are applicable to more general Ising models, but in those cases, the necessary joint probabilities have not been calculated.

II. STATISTICS

We write the Hamiltonian of the Ising chain in the form

$$E=-\tfrac{1}{2}J\sum_{\langle i,j\rangle}\sigma_i\sigma_j-H\sum_i\sigma_i,$$

where σ_i is the variable describing particle i, and σ_i takes on the values ±1. H is the external magnetic field times the various factors needed to make the energy come out right. The summation index, $\langle i,j\rangle$, means that the sum is taken over near neighbors in the chain and each "bond" is counted once. Then the partition function, Z, is given by

$$Z=\mathrm{tr}(e^{-\beta E}),$$

where

$$\beta=1/kT.$$

Suppose now that one wants to know the probability that any particular spin, spin i, is up or down, $P\{\sigma_i=\pm1\}$. We call this $P(K_i)$, where $K_i=\pm1$ for $P\{\sigma_i=\pm1\}$. This will be given by the sum over all the Boltzmann factors involving $\sigma_i=K_i$, divided by the partition function. We can extend this sum over the whole phase space by multiplying the Boltzmann factor by a projection operator which is one when $\sigma_i=K_i$ and zero when $\sigma_i=-K_i$. Such an operator is $\hat{P}(K_i)=\frac{1}{2}(1+K_i\sigma_i)$, so that

$$P(K_i)=Z^{-1}\,\mathrm{tr}(\hat{P}(K_i)e^{-\beta E})\equiv\langle\hat{P}(K_i)\rangle=\tfrac{1}{2}(1+K_i\langle\sigma\rangle),$$

where the bracket means the thermodynamic average and $\langle\sigma_i\rangle=\langle\sigma\rangle$ because of the translational invariance of the chain.

Obviously the joint probability of any particular arrangement of any n spins is given by the following expression, where we make an obvious extension of notation:

$$\begin{aligned}&P(K_i,K_j,\cdots,K_r)\\&\quad=\langle\hat{P}(K_i)\hat{P}(K_j)\cdots\hat{P}(K_r)\rangle\\&\quad=2^{-n}(1+\langle\sigma\rangle\sum_i K_i+\sum_{i,j}K_iK_j\langle\sigma_i\sigma_j\rangle+\cdots\\&\qquad\qquad+K_iK_j\cdots K_r\langle\sigma_i\sigma_j\cdots\sigma_r\rangle),\end{aligned}$$

where the summations are over the spins involved. We must therefore calculate the general spin correlation function, $\langle\sigma_i\cdots\sigma_k\rangle$.

This is true, of course, for any Ising model, but for the Ising chain, these correlation functions can be evaluated exactly using, for example, matrix methods.[5,6]

Consider the two-by-two matrix

$$H(\sigma_i\sigma_{i+1})=\exp(K\sigma_i\sigma_{i+1}+\tfrac{1}{2}C(\sigma_i+\sigma_{i+1}))$$

* This work was supported under the auspices of the U. S. Atomic Energy Commission.

[1] J. S. Marsh, dissertation, The Johns Hopkins University, 1966 (unpublished).

[2] G. Prinz, dissertation, The Johns Hopkins University, 1966 (unpublished).

[3] Y. Y. Lee, Can. J. Phys. **39**, 1733 (1961).

[4] M. McMillan and W. Opechowski, Can. J. Phys. **38**, 1168 (1960).

[5] H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252, 263 (1941).

[6] B. Kaufman and L. Onsager, Phys. Rev. **76**, 1244 (1949).

or

$$H=\begin{pmatrix} e^{K+C} & e^{-K} \\ e^{-K} & e^{K-C} \end{pmatrix},$$

where $K=J/2kT$, and $C=H/kT$.

As is well known,[7] the partition function for the Ising chain is given by

$$Z=\mathrm{tr}(H(\sigma_1\sigma_2)H(\sigma_2\sigma_3)\cdots H(\sigma_N\sigma_1)), \qquad (1)$$

where there are N spins in the chain and the ends of the chain are joined.

If we transform H into diagonal form,

$$\lambda=SHS^{-1},$$

$$\lambda=\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

where

$$\lambda_{1,2}=e^K\cosh(C)(1\pm\gamma),$$

and

$$\gamma=(1+(e^{-4K}-1)\,\mathrm{sech}^2C)^{1/2},$$

then

$$Z=\lambda_1{}^N+\lambda_2{}^N.$$

Since γ is always positive, $\lambda_1>\lambda_2$, and as N gets very large, $Z\sim\lambda_1{}^N$, which is the usual result.

Suppose now that one wants to calculate $\langle\sigma_i\rangle$. This can be done if, in Eq. (1), the part that involves $\sigma_i=+1$ is taken with a plus sign and the part that involves $\sigma_i=-1$ is taken with a minus sign, and the sum divided by Z. This will obviously happen if

$$\hat{\sigma}=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is inserted between the two matrices in Eq. (1) that involve σ_i before the trace is taken.

$$\langle\sigma_i\rangle=Z(\sigma_i)\equiv Z^{-1}\,\mathrm{tr}(H(\sigma_1\sigma_2)\cdots \times H(\sigma_{i-1}\sigma_i)\hat{\sigma}H(\sigma_i\sigma_{i+1})\cdots H(\sigma_N\sigma_1)).$$

Similarly, any n spin correlation function can be written, with the obvious extension of notation,

$$\langle\sigma_i\sigma_{i+r_1}\cdots\sigma_{i+r_1+\cdots+r_n}\rangle=Z(\sigma_i\sigma_{i+r_1}\cdots\sigma_{i+r_1+\cdots+r_n}).$$

If we now let $\Sigma=S\hat{\sigma}S^{-1}$, where S is the same matrix that diagonalizes H, then

$$\langle\sigma_i\sigma_{i+r_1}\cdots\sigma_{i+r_1+\cdots+r_n}\rangle=\lambda_1{}^{-N}\,\mathrm{tr}(\lambda^{N-r_1-\cdots-r_n}\Sigma\lambda^{r_1}\Sigma\cdots\lambda^{r_n}\Sigma).$$

Since $H(\sigma_i\sigma_{i+1})$ is real and Hermitian, S is real and orthogonal, and we have $\Sigma_{11}=-\Sigma_{22}$ and $\Sigma_{12}=\Sigma_{21}$.

For instance, we have

$$\langle\sigma_i\rangle=\lambda_1{}^{-N}\,\mathrm{tr}(\lambda^N\Sigma)=\Sigma_{11}+\Sigma_{22}(\lambda_2/\lambda_1)^N=\Sigma_{11}=\langle\sigma\rangle$$

as N gets large. This evaluates Σ_{11} and Σ_{22} since

$$\langle\sigma\rangle=\partial\ln\lambda_1/\partial C=\tanh C/\gamma.$$

$$\langle\sigma_i\sigma_{i+r}\rangle=\lambda_1{}^{-N}\,\mathrm{tr}(\lambda^{N-r}\Sigma\lambda^r\Sigma)=\Sigma_{11}{}^2+\Sigma_{12}{}^2\epsilon^r,$$

where we have set $\epsilon=\lambda_2/\lambda_1$. When $r=0$, we have

$$\langle\sigma_i\sigma_i\rangle=1=\langle\sigma\rangle^2+\Sigma_{12}{}^2 \quad\text{or}\quad \Sigma_{12}{}^2=1-\langle\sigma\rangle^2.$$

When $r=1$,

$$\langle\sigma_i\sigma_{i+1}\rangle\equiv\langle\sigma\sigma\rangle=\langle\sigma\rangle^2+(1-\langle\sigma\rangle^2)\epsilon.$$

Now $\langle\sigma\sigma\rangle$ is given by $\langle\sigma\sigma\rangle=\partial\ln\lambda_1/\partial K$, so that

$$\epsilon=(\langle\sigma\sigma\rangle-\langle\sigma\rangle^2)/(1-\langle\sigma\rangle^2)$$

evaluates λ_2/λ_1, which is also given by $\epsilon=(1-\gamma)/(1+\gamma)$. The latter form is convenient for calculations while the former form is more expressive of the physical content.

$$\langle\sigma_i\sigma_{i+r}\sigma_{i+r+s}\rangle=\lambda_1{}^{-N}\,\mathrm{tr}(\lambda^{N-r-s}\Sigma\lambda^r\Sigma\lambda^s\Sigma)$$
$$=\Sigma_{11}{}^3+\Sigma_{12}{}^2\Sigma_{11}(\epsilon^r+\epsilon^s)+\Sigma_{12}{}^2\Sigma_{22}\epsilon^{r+s}$$
$$=\langle\sigma\rangle^3+\langle\sigma\rangle(1-\langle\sigma\rangle^2)(\epsilon^r+\epsilon^s-\epsilon^{r+s}).$$

In general we have

$$\langle\sigma_i\sigma_{i+r_1}\cdots\sigma_{i+r_1+\cdots+r_n}\rangle$$
$$=\lambda_1{}^{-N}\sum_{j,k,\cdots,p}(\lambda_j{}^{N-r_1-\cdots-r_n}\delta_{jk}\Sigma_{kr}\lambda_r{}^{r_1}\delta_{rs}\cdots\lambda_q{}^{r_n}\delta_{qp}\Sigma_{pj})$$

$$=\Sigma_{11}{}^{n+1}+\Sigma_{11}{}^{n-1}\Sigma_{12}{}^2(\epsilon^{r_1}+\epsilon^{r_2}+\cdots+\epsilon^{r_n})+\Sigma_{11}{}^{n-3}\Sigma_{12}{}^4(\epsilon^{r_1+r_3}+\epsilon^{r_1+r_4}+\cdots+\epsilon^{r_1+r_n}+\epsilon^{r_2+r_4}+\cdots+\epsilon^{r_{n-2}+r_n})$$
$$+\Sigma_{11}{}^{n-2}\Sigma_{22}\Sigma_{12}{}^2(\epsilon^{r_1+r_2}+\epsilon^{r_2+r_3}+\cdots+\epsilon^{r_{n-1}+r_n})+\Sigma_{11}{}^{n-5}\Sigma_{12}{}^6(\epsilon^{r_1+r_3+r_5}+\epsilon^{r_1+r_3+r_6}+\cdots+\epsilon^{r_{n-4}+r_{n-2}+r_n})$$
$$+\cdots+\Sigma_{22}{}^{n-1}\Sigma_{12}{}^2\epsilon^{r_1+r_2+\cdots+r_n},$$

where we have neglected terms proportional to ϵ^N. This is written down by inspection as follows: The trace involves λ_1 and λ_2 to the various powers $r_1\cdots r_n$; we start by writing down first the part coming from the λ_1 parts of all the λ matrices, i.e., $\Sigma_{11}{}^{n+1}$; this gives the first term. Then we have the part that comes from the λ_2 part of one of the λ matrices; this introduces a Σ_{12} factor from each side of the λ matrix in question and Σ_{11} occurs only $n-1$ times. This gives the second term. Next we have the terms from the λ_2 parts of two

[7] D. ter Haar, *Elements of Statistical Mechanics* (Rinehart and Company, Inc., New York, 1956).

of the λ matrices. When the two λ matrices in question are not adjacent, we get Σ_{11} occuring $n-3$ times and Σ_{12} 4 times, two for each λ matrix; this gives the third term. When the λ matrices are adjacent we get Σ_{11} $n-2$ times, Σ_{22} once from the Σ between the two λ matrices, and Σ_{12} twice to get the fourth term; and so on. Thus we can calculate the general n-spin correlation function.

III. APPLICATIONS

For a broad class of problems one is interested in the relative number of transitions between two sets of states induced by some one-particle operator. By one-particle operator, we mean an operator of the form $\hat{O}=\sum_i \hat{O}_i$, where $\hat{O}_i$ operates only on the coordinates of the ith spin. These transitions are classified by the energy difference between the initial and final states, and since each spin is coupled only to its nearest neighbors, this means classification according to different arrangements of three spins in a row.

We will illustrate this first by calculating the perpendicular magnetic susceptibility for the Ising chain for any value of the parallel magnetic field. This has already been done by Fisher[8] and Katsura[9] for the case when the parallel field is zero. We do this by calculating the correction to the free energy of the Ising chain in a perpendicular magnetic field H_x to second order in H_x. The susceptibility $\chi_\perp$ is then found by taking the appropriate derivatives.

If the Hamiltonian of the crystal is given by $\mathcal{H}=E+V$, and the states of the unperturbed Hamiltonian are labeled with the indices n, i.e., $E|n\rangle=E_n|n\rangle$, and the matrix elements of the perturbing potential, V, with the eigenstates of E are called V_{nm}, then the free energy of the system described by $\mathcal{H}$ can be expanded in the form[10]

$$F=F_0+F_1+F_2+\cdots,$$

where $F_0=-\beta^{-1}\ln \mathrm{tr}(e^{-\beta E})\equiv-\beta^{-1}\ln Z_0$ is the "unperturbed free energy,"

$$F_1=\langle V_{nn}\rangle$$

and

$$F_2=-\tfrac{1}{2}\sum_{n,m}[(\omega_m-\omega_n)|V_{nm}|^2/(E_n-E_m)]$$
$$-\tfrac{1}{2}\beta\langle(V_{nn}-\langle V_{nn}\rangle)^2\rangle.$$

TABLE I. In the "Neighbors" column, the signs on either side of the comma indicate the orientation of the neighboring spins on either side of spin i; E_n-E_m is the corresponding energy denominator; P_n and P_m are explained in the text.

Neighbors	E_n-E_m	P_n	P_m
+, +	$-2(H_z+J)$	$P(UUU)$	$P(UDU)$
+, −; −, +	$-2H_z$	$P(UUD)$	$P(UDD)$
−, −	$-2(H_z-J)$	$P(DUD)$	$P(DDD)$

Here, the angular brackets mean the thermodynamic average over the "unperturbed distribution," $e^{-\beta E_n}$, and ω_n is the probability of occurrence of the unperturbed state, $\omega_n=Z_0^{-1}e^{-\beta E_n}$.

E will be the Hamiltonian for the Ising chain, so that $Z_0=\lambda_1{}^N$, and $V=-H_x\sum_i(\sigma_i{}^+ +\sigma_i{}^-)\equiv\sum_i V_i$, where again the factors that make the energy come out right have been absorbed into H_x, and

$$\sigma^+=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\quad \sigma^-=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}.$$

In this case, $V_{nn}=0$, so that the correction to second order comes from the first term of F_2, which may be written

$$F_2=\tfrac{1}{2}\sum_{n,m,i}(\omega_n-\omega_m)|V_{i,nm}|^2/(E_n-E_m).$$

Since the V_i cause only one spin at a time to flip. $|V_{i,nm}|^2=H_x{}^2$ when it is nonzero, and it is nonzero only between states that differ only in the orientation of spin i. Thus the energy denominators fall into three classes determined by the orientation of the nearest neighbors to spin i. If we let the index n refer to those states where spin i is up ($\sigma_i=+1$) and m refer to the states where spin i is down, the neighborhoods and corresponding energy denominators are as given in Table I. It is apparent that when $\omega_n-\omega_m$ is performed over the states which lead to a particular energy denominator, we will get just the joint probabilities of the various arrangements of three spins in a row. These are given in the table in the columns headed P_n and P_m where, for instance, $P(UUD)\equiv P(K_{i-1}K_iK_{i+1})$, when $K_{i-1}=K_i=+1$ and $K_{i+1}=-1$, i.e., spins i and $i-1$ are up and spin $i+1$ is down. The subsequent summation over i gives the factor N. Thus

$$F_2=-NH_x{}^2\left[\frac{P(UUU)-P(UDU)}{2(H_z+J)}+\frac{P(UUD)-P(UDD)}{H_z}+\frac{P(DUD)-P(DDD)}{2(H_z-J)}\right]$$

$$=-H_x{}^2NK/2J\left[\frac{\langle\sigma\rangle+2\langle\sigma\sigma\rangle+\langle\sigma\sigma\sigma\rangle}{2(C+2K)}+\frac{\langle\sigma\rangle-\langle\sigma\sigma\sigma\rangle}{C}+\frac{\langle\sigma\rangle-2\langle\sigma\sigma\rangle+\langle\sigma\sigma\sigma\rangle}{2(C-2K)}\right]\equiv-H_x{}^2NG/2J=-H_x{}^2\chi_\perp/2,$$

[8] M. E. Fisher, Physica **26**, 618 (1960).
[9] S. Katsura, Phys. Rev. **127**, 1508 (1962).
[10] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Permagon Press Ltd., London, 1958).

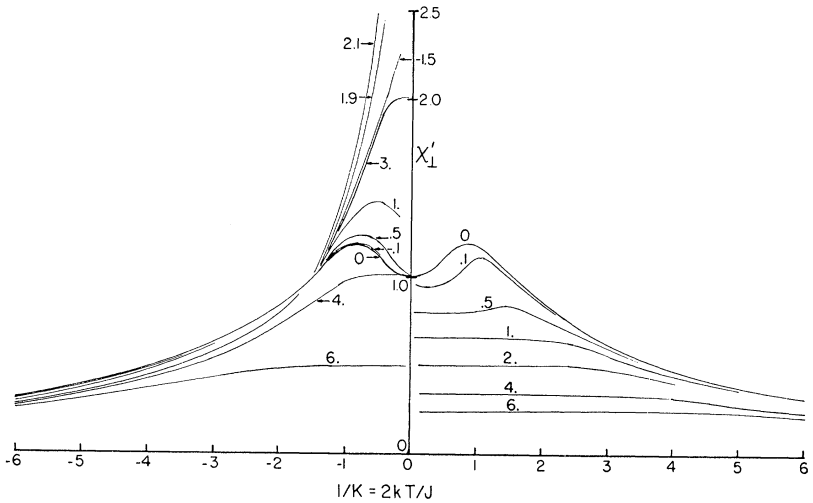

FIG. 1. Perpendicular magnetic susceptibility of the Ising chain as a function of temperature for different values of the parallel magnetic field. $\chi_\perp' = J\chi_\perp/N$. The numbers on each curve give the ratio $2H_z/J$. Ferromagnetic coupling is on the right-hand side of the graph and antiferromagnetic coupling is on the left.

since $\chi_\perp = -\partial^2 F/\partial H_x^2$. Here, $\langle\sigma\sigma\sigma\rangle = \langle\sigma_{i-1}\sigma_i\sigma_{i+1}\rangle$, and we have used the fact that $P(UUD) = P(DUU)$, etc.

One might be alarmed that something nasty will happen to F_2 when one of the denominators approaches zero. However, when this happens, F_2 approaches a finite limit and in fact it can be shown, by taking suitable linear combinations of the offending states so that V is diagonal among these states and calculating the contributions of these matrix elements to F_1 and the second term of F_2, that G approaches the correct limit.

Thus, when $H_z \to 0$, $J\chi_\perp/N \sim K\,\mathrm{sech}^2 K + \tanh K$, which is the result of Fisher and Katsura.

In Fig. 1 we show $J\chi_\perp/N$ as a function of $1/K$ for various values of $C/K = 2H_z/J$. When J is positive (ferromagnetic coupling) the chain becomes more "rigid," i.e., the spins are locked in the z direction, as H_z is increased and it becomes harder for H_x to induce a moment σ_x in the x direction. For constant H_z the fluctuations in σ_x are controlled by the magnetic energy until the thermal energy overcomes the magnetic energy. In fact, $kT\chi_\perp \sim \langle\sigma_x^2\rangle$, so we see $\langle\sigma_x^2\rangle$ rising linearly from $T=0$ and becoming constant for temperatures higher than $kT \sim H_z$.

When J is negative (antiferromagnetic coupling) at low temperatures, we see the susceptibility rising as H_z increases toward J, for, as this happens, the spins become more and more "free" as the magnetic energy tends to cancel the coupling energy and so the fluctuations $\langle\sigma_x^2\rangle$ become large. As H_z increases beyond this point and dominates the orientation energy, the chain resembles more and more the ferromagnetically coupled chain, as is shown by the behavior of the susceptibility for high fields.

Thus the perpendicular susceptibility behaves more or less the way it is supposed to.

We now discuss Lee's calculation.[3] Lee is calculating the relative intensities of lines seen in an EPR-type experiment on an Ising chain. The lines come from a flip of one spin at a time and the corresponding energies are just the energy denominators in Table I and they arise from the corresponding neighborhoods. We see by inspection that if Lee's Eq. (4) is divided on the left by the partition function of the Ising chain, the statistical sums will be just the differences of the joint probabilities of the neighborhoods as found in the last two columns of Table I. Thus, we have immediately, in unnormalized form, Lee's Eq. (7).

$$I(2H_z+2J) = P(UUU) - P(UDU),$$
$$I(2H_z) = 2(P(UUD) - P(UDD)),$$
$$I(2H_z-2J) = P(DUD) - P(DDD),$$

except for differences from the different way in which the Ising chain Hamiltonian is written.