SLAC – PUB – 3485
October 1984
T/E

NUMERICAL SIMULATION OF QUANTUM SPIN MODELS*

D. KUNG AND R. BLANKENBECLER

Stanford Linear Accelerator Center

Stanford University, Stanford, California, 94305

and

R. L. SUGAR†

Institute for Theoretical Physics

University of California, Santa Barbara, California 93106

Submitted to *Physical Review B*

* Work supported in part by the Department of Energy Contract DE-AC03-76SF00515.
† Work supported in part by the Department of Energy Contract DE-AT03-83ER45008, National Science Foundation Grants PHY83-13324 and PHY77-27084, and grants from E. I. DuPont de Nemours and Company, and Xerox Corporation.

ABSTRACT

The general application of the projector Monte Carlo method to spin systems is discussed. The purpose of the paper is to present several variants of the method that improve its accuracy and convergence in specific problems and to give numerical examples of the improvements. More detailed application to specific spin models is left for subsequent publications.

1. Introduction

In this paper we study the numerical simulation of quantum spin models. We present a general formalism for carrying out such simulations[1] and discuss methods for reducing systematic and statistical errors. In subsequent publications the methods presented here will be used to make detailed studies of specific spin models.[2,3]

Our general approach is to use statistical methods to project out the ground state and first excited state wave functions of the system of interest. Once these wave functions have been obtained, it is straightforward to calculate energies, gaps, correlation functions and thermodynamic quantities. This general approach was pioneered by Kalos and his collaborators,[4] Cepereley and Adler,[5] and has been applied to the study of a wide variety of physical systems.[6–9] In particular Kuti and his collaborators have been studying quantum spin models alone lines similar to ours.[10]

In Section 2 of this paper we describe our basic formalism.[9] In Sections 3 and 4 we discuss methods of minimizing the systematic and statistical errors that arise in the calculation of energies and correlation functions. In Section 5 we turn to the problem of computing the energy gap between the ground state and the first excited state. The straightforward approach of calculating the individual energies of these states and then subtracting is untenable for large systems since the energies grow linearly with the volume, while the gap approaches a volume independent limit. We present alternative approaches which avoid this difficulty. Finally in Section 6 we briefly summarize our results.

2. Formalism

In this section we present our approach to the numerical study of quantum spin models. We shall use the Ising model and Heisenberg antiferromagnet as illustrative examples.

As is well known, a classical spin model in $d+1$ dimensions can always be described by a transfer matrix in one lower dimension. For example, the Ising model with partition function

$$Z = \sum_{S_i=\pm 1} e^{g\sum_{\langle ij\rangle} S_i S_j} \tag{1}$$

has the transfer matrix

$$T = \prod_i (e^g + e^{-g}\sigma_z^i) \prod_{\langle ij\rangle} e^{g\sigma_x^i \sigma_x^j} \ . \tag{2}$$

In Eq. (1) the sum is over all nearest neighbor lattice points in a $d+1$ dimensional space. In Eq. (2) the first product is over all lattice points in a d dimensional space and the second product is over all nearest neighbor points in this space. σ_x^i and σ_z^i are the usual Pauli spin matrices defined at the ith lattice point.

Alternatively, one can start with a quantum spin Hamiltonian in d dimensions. For the Ising model in a transverse field we have

$$\begin{aligned} H &= -g\sum_{\langle ij\rangle} \sigma_x^i \sigma_x^j - h\sum_i \sigma_z^i \\ &\equiv H_1 + H_2 \ . \end{aligned} \tag{3}$$

A quantity of particular interest is the evolution operator for the propagation of

the system through an imaginary time interval $\Delta\tau$,

$$T(\Delta\tau) = e^{-\Delta\tau H} \ . \tag{4}$$

For small values of $\Delta\tau$ we can write

$$\begin{aligned} T(\Delta\tau) &\simeq e^{-\Delta\tau H_2} e^{-\Delta\tau H_1} \\ &= \prod_i (\cosh \Delta\tau h - \sigma_z^i \sinh \Delta\tau h) \prod_{\langle ij \rangle} e^{\Delta\tau g \sigma_x^i \sigma_x^j} \ , \end{aligned} \tag{5}$$

which has the same structure as Eq. (2). The important feature of Eqs. (2) and (5) is that the transfer matrix (from now on we will use the term transfer matrix to denote either T or $T(\Delta\tau)$) can be written as a product of operators each term of which involves only a few neighboring lattice sites. We shall shortly consider breakups of H other than Eq. (3) which lead to the same property.

Most quantities of physical interest can be obtained from a knowledge of the two largest eigenvalues of the transfer matrix, λ_0, λ_1 ($\lambda_0 > \lambda_1$), and their corresponding eigenvectors, $|\psi_0\rangle$, $|\psi_1\rangle$. Our strategy is to project out $|\psi_0\rangle$ and $|\psi_1\rangle$ by repeatedly applying the transfer matrix to suitable trial states. Ordinarily one is aided by the fact that $|\psi_0\rangle$ and $|\psi_1\rangle$ are eigenvectors of a symmetry operator with different eigenvalues. For the Ising model the appropriate symmetry operator (henceforth referred to as parity) is

$$P = \prod_j \sigma_z^j \tag{6}$$

and

$$P\,|\psi_i\rangle = (-)^i\,|\psi_i\rangle \ , \quad i = 0, 1 \ . \tag{7}$$

If we then introduced trial states, $|\phi_i\rangle$, such that

$$P\,|\phi_i\rangle = (-)^i\,|\phi_i\rangle \quad , \quad i = 0,1 \; , \tag{8}$$

we see that

$$T^N\,|\phi_i\rangle \xrightarrow[N\to\infty]{} c_i(\lambda_i)^N\,|\psi_i\rangle \quad , \quad i = 0,1 \tag{9}$$

where c_i are constants. It then follows that

$$\frac{\langle\chi_i|\; T^{N+M}\;|\phi_i\rangle}{\langle\chi_i|\; T^N\;|\phi_i\rangle} \xrightarrow[N\to\infty]{} (\lambda_i)^M \tag{10}$$

and

$$\frac{\langle\chi_i|\; T^N Q T^N\;|\phi_i\rangle}{\langle\chi_i|\; T^N\;|\phi_i\rangle} \xrightarrow[N\to\infty]{} \langle\psi_i|\; Q\;|\psi_i\rangle \; , \tag{11}$$

where $|\chi_i\rangle$ are trial states which also satisfy Eq. (8) and Q is an arbitrary operator. In Eqs. (10) and (11) T stands either for the transfer matrix of Eq. (2) or $T(\Delta\tau)$ of Eq. (4).

We will evaluate the matrix elements of Eqs. (10) and (11) using stochastic techniques due to the von Neuman and Ulam[11] and Kuti.[10] The first step is to introduce complete sets of states, $|j_\ell\rangle$, between each factor of T. We then write

$$\begin{aligned} Z_i(N) &= \langle\chi_i|\; T^N\;|\phi_i\rangle \\ &= \sum_{j_1\cdots j_{N+1}} \langle\chi_i|\, j_{N+1}\rangle\;\langle j_{N+1}|\; T\;|j_N\rangle\;\langle j_N|\; T\;|j_{N-1}\rangle \\ &\quad \ldots\langle j_2|\; T\;|j_1\rangle\;\langle j_1|\,\phi_i\rangle \; . \end{aligned} \tag{12}$$

It is the sum over intermediate states in Eq. (12) that we wish to perform

stochastically. To this end we write the matrix elements in the form

$$\langle j_1 | \phi_i \rangle = p^i_{j_1} s^i_{j_1} , \qquad \langle j_{\ell+1} | T | j_\ell \rangle = P_{j_{\ell+1} j_\ell} S_{j_{\ell+1} j_\ell} . \tag{13}$$

$p^i_{j_1}$ will be the probability that the first intermediate state is $|j_1\rangle$. $P_{j_{\ell+1}\, j_\ell}$ will be the probability that the $\ell + 1^{th}$ intermediate state is $|j_{\ell+1}\rangle$ given that the ℓth intermediate state is $|j_\ell\rangle$. We shall refer to $s^i_{j_1}$ and $S_{j_{\ell+1}\, j_\ell}$ as the corresponding scores. It is clearly necessary to impose the constraints

$$p^i_{j_1} , \; P_{j_{\ell+1} j_\ell} \geq 0 , \qquad \sum_{j_1} p^i_{j_1} = 1 , \qquad \sum_{j_{\ell+1}} P_{j_{\ell+1} j_\ell} = 1 . \tag{14}$$

One approach, which we have previously called the projector Monte Carlo method,[9] is to choose the state $|j_1\rangle$ with probability $p^i_{j_1}$, the state $|j_2\rangle$ with probability $P_{j_2 j_1}$, etc. We thereby generate a particular set of intermediate states $\alpha(j_1 \ldots j_{N+1})$ with a probability

$$P_{j_{N+1} j_N} \quad P_{j_n j_{N-1}} \cdots P_{j_2 j_1} \quad p^i_{j_1} .$$

Repeating this process allows us to sample many different sets of intermediate states. Clearly

$$Z_i(N) = \sum_\alpha \langle \chi^i | j_{N+1} \rangle \; S_{j_{N+1} j_N} \; S_{j_N j_{N-1}} \cdots S_{j_2 j_1} s^i_{j_1} ,$$

and

$$\begin{aligned} (\lambda_i)^M &= \frac{\sum_{\alpha'} \langle \chi^i | j_{M+N+1} \rangle \; S_{j_{N+M+1} j_{8N+M}} \cdots S_{j_2 j_1} s^i_{j_1}}{\sum_\alpha \langle \chi^i | j_{N+1} \rangle \; S_{j_{N+1} j_N} \cdots S_{j_2 j_1} s^i_{j_1}} \\ \langle \psi_i | \; Q \; | \psi_i \rangle &= \frac{\sum_{\alpha'} \langle \chi^i | j_{N+2} \rangle \; S_{j_{2N+1} j_{2N+1}} \cdots \langle j_{N+2} | \; Q \; | j_{N+1} \rangle \cdots s^i_{j_1}}{\sum_\alpha \langle \chi_i | \, j_{2N+1} \rangle S_{j_{2N_1} j_{2N}} \cdots s^i_{j_1}} . \end{aligned} \tag{15}$$

The division of the matrix elements of the transfer matrix into a transition probability and a score leaves a great deal of freedom which can be exploited

to improve the convergence of the method. If one has some idea of the underlying physics, one can use it to construct transition probabilities that favor the most important intermediate states. The calculation is kept exact by adjusting the scores in accordance with Eq. (13). Kalos[4] has emphasized that with a knowledge of the exact ground state wave function we could completely eliminate statistical fluctuations, and with a good variational trial function they could be significantly reduced.

3. Systematic Errors

Finite β

There are several sources of errors in this stochastic approach. One of the fundamental systematic errors arises because of the simple fact that in a practical calculation we can only apply the transfer matrix to the trial states a finite number of times. Because the quantity N in Eqs. (10)-(12) is finite, non-leading eigenvalues of T enter. If we denote the second leading eigenvalue with parity i by λ_i' and its corresponding eigenvector by $|\psi_i'\rangle$ then, for example,

$$Z_i(N+M)/Z_i(N) = \lambda_i^M \left\{ 1 + A\left(\lambda_i'/\lambda_i\right)^N \left[\left(\lambda_i'/\lambda_i\right)^M - 1 \right] + \dots \right\} \qquad (16)$$

where

$$A = \frac{\langle \chi_i | \psi_i' \rangle \ \langle \psi_i' | \phi_i \rangle}{\langle \chi_i | \psi_i \rangle \langle \psi_i | \phi_i \rangle} \ .$$

The problem of course is to minimize the non-leading term in the bracket on the right-hand side of Eq. (16). The most straightforward approach is to increase N until the results are independent of it. Here one relies on the fact that if the

(perhaps infinite) system has a finite gap, the quantity λ_i'/λ_i approaches a finite limit—less than one—as the spatial size of the system is increased. However, if the gap is zero in the infinite volume limit, then $\ell n(\lambda_i'/\lambda_i) \sim 1/L$ where L is some linear dimension of the system. As a result, if the gap is zero, or just small, one might have to make N impractically large, if one was relying on the $(\lambda_i'/\lambda_i)^N$ term alone for convergence.

An alternative is to make A small by minimizing the overlap of $|\phi_i\rangle$ and $|\chi_i\rangle$ with $|\psi_i'\rangle$ and other eigenvectors associated with non-leading eigenvalues. This can be accomplished if one has a good trial state for $|\psi_i\rangle$, by choosing $|\phi_i\rangle$ and $|\chi_i\rangle$ to be such a state.

We will illustrate the improvement due to the choice of the initial state by considering a nearest neighbor 4-site Heisenberg antiferromagnet, whose Hamiltonian is given by

$$H = \sum_{\langle i,j \rangle} \vec{s}_i \cdot \vec{s}_j \ . \tag{17}$$

We input a trial state of the form

$$|\phi\rangle = |\uparrow\downarrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\downarrow\rangle + b\Big(|\downarrow\uparrow\uparrow\downarrow\rangle + |\uparrow\uparrow\downarrow\downarrow\rangle + |\uparrow\downarrow\downarrow\uparrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle\Big)$$

which becomes exact when $b = 0.5$.

From Table I, we note that the error due to finite β is reduced significantly if $|\phi\rangle$ is a good approximation to the exact ground state.

The corresponding improvement due to the choice of the final state is illustrated by considering an 8-site 1-d Ising model whose Hamiltonian is

$$-\sum_{i=1}^{8} \left[\sigma_i^x + \sigma_i^z \sigma_{i+1}^z\right] \ . \tag{18}$$

Periodic boundary condition is imposed and our basis $|i\rangle$ is σ^z diagonal. This system is in the ferromagnetic regime so we choose the trial form to be

$$|\chi\rangle = \sum_i e^{-c_t \sum_{j=1}^{8} \sigma_j^z \sigma_{j+1}^z} |i\rangle \ . \tag{19}$$

We expect the optimal c_t to be negative because ferromagnetic alignment is then favored. In Fig. 1 we plot the measured energy as a function of β and the approach to the $\beta = \infty$ limit is much faster for negative c_t than for the positive ones, thus confirming our expectation.

If a suitable trial state cannot be found, it can be generated stochastically as follows. Let us imagine expanding $|\psi_i\rangle$ in terms of the basis states

$$|\psi_i\rangle = \sum_j a_i(j) |j\rangle \ . \tag{20}$$

The ideal choice for $|\phi_i\rangle$ is $|\psi_i\rangle$, so, for example,

$$Z_i(N+M)/Z_i(N) = \frac{\sum_j Y_i^j(N+M) a_i(j)}{\sum_j Y_i^j(N) a_i(j)} \ , \tag{21}$$

where we have written

$$Y_i^j(N) = \langle \chi_i | \ T^N \ | j \rangle \ . \tag{22}$$

Now for large N

$$Y_i^j(N) \simeq \lambda_i^N \langle \chi_i | \psi_1 \rangle a_i(j) \ , \tag{23}$$

so

$$Z_i(N+M)/Z_i(N) = \left[\sum_j Y_i^j (N+M)^2 / \sum_j Y_i^j (N)^2\right]^{1/2} \qquad (24a)$$

$$= \left[\langle\chi_i|\, T^{2(N+M)}\, |\chi_i\rangle \,/\, \langle\chi_i|\, T^{2N}\, |\chi_i\rangle\right]^{1/2} . \qquad (24b)$$

Thus by using Eq. (21) we can double the effective values of N and M without increasing the number of time steps in the Monte Carlo calculation. Using again the example of the Heisenberg antiferromagnet, Eq. (17), we have found that the energy calculated form Eq. (24a) is only a few parts in a thousand away from the energy calculated by inputting the exact ground state (see Table I and Ref. 1 for details). The price paid for this improvement is that one must calculate the Y_i^j for a variety of states j. However, we have found that even truncating the sum in Eq. (21) to a small subset of j can significantly reduce the systematic error due to finite N.

Breakup Error

A second important source of systematic error arises in most Hamiltonian systems. As we noted in Section 2, in order to evaluate the time evolution operator, $T(\Delta\tau)$, of Eq. (4), it is convenient to break the Hamiltonian up into two pieces each of which is a sum of commuting terms as in Eq. (3). We then write for small $\Delta\tau$

$$T(\Delta\tau) = T_2(\Delta\tau)\, T_1(\Delta\tau) \left\{1 - \frac{1}{2}(\Delta\tau)^2 [H_2, H_1] + \ldots\right\} ,$$

where

$$T_i(\Delta\tau) = e^{-\Delta\tau H_i} \qquad i = 1, 2 .$$

Ordinarily this breakup leads to errors in the energy of order $(\Delta\tau)^2$ and errors in correlation functions can be of order $\Delta\tau$. However, the magnitudes of these errors are very sensitive to the specific choice of the breakup, the parameters of the model, and the model itself. This is clearly illustrated in Fig. 2 where we present results for the Hamiltonian of Eq. (3). The points denoted by the ○'s represent the breakup

$$H_1' = \sum_{i\,\mathrm{odd}} \left[-g\sigma_z^i\sigma_z^{i+1} - h\,\sigma_x^i \right]$$
$$H_2' = \sum_{i\,\mathrm{even}} \left[-g\sigma_z^i\sigma_z^{i+1} - h\,\sigma_x^i \right] \tag{25}$$

while the •'s represent the breakup of Eq. (3) into separate g and h terms. It is remarkable that the error to the first breakup is essentially negligible while that due to the second is clearly of $O(\Delta\tau^2)$ as predicted. $H_1 + H_2$ and $H_1' + H_2'$ have the same lowest energy eigenvalue and we see that the extrapolation to the $\Delta\tau = 0$ limit is within a few parts of a thousand of the exact result.

4. Statistical Fluctuations

In Section 2 we set up a general formalism for carrying out the numerical calculation, and outlined a specific implementation, the projector Monte Carlo method. For large systems the statistical fluctuations encountered in evaluating the sums in Eq. (15) can become large since the eigenvalues of the transfer matrix, and therefore the scores $S_{i,j}$, grow exponentially with the volume of the system. However, the freedom in factoring the matrix elements of T into a probability and a score allows us to attempt to minimize the fluctuations by an appropriate choice of the probabilities.

It is interesting to ask for the optimal choice of the probabilities and scores. We define this choice to be the one for which the mean square deviation of $Z_i(N)$ is a minimum. That is we minimize

$$\Delta Z_i(N)^2 = \sum_{j_1 j_{N+1}} P_{j_{N+1},j_N}(N) \dots P_{j_2,j_1}(1) p_{j_1} \left[\chi^i_{j_{N+1}} \, S_{j_{N+1},j_n}(N) \dots S_{j_2,j_1}(1) s_{j_1}\right]^2 \tag{26}$$

with respect to the choice of the P's for fixed values of the matrix elements, $\langle j_{\ell+1}| \, T \, |j_\ell\rangle$, and subject to the constraints of Eq. (14). In writing Eq. (25) we have anticipated that the optimal probabilities and scores will depend on the time slice. We have also introduced the notation $\chi^i_{j_{N+1}} = \langle \chi_i \,|\, j_{N+1}\rangle$.

A straightforward calculation yields

$$P^{\rm opt}_{j_{\ell+1},j_\ell}(\ell) = F^i_{\ell+1}\,|j_{\ell+1}\rangle \,|\, \langle j_{\ell+1}|\, T \,|j_\ell\rangle \,|/F^i_\ell(j_\ell) \;, \tag{27}$$

where

$$F^i_{N+1}(j_{N+1}) = \chi^i_{j_{N+1}}$$

and

$$F^i_\ell(j_\ell) = \sum_{j_{\ell+1}} F^i_{\ell+1}(j_{\ell+1})|\, \langle j_{\ell+1}|\, T \,|j_\ell\rangle \,|\;.$$

For $\ell \ll N$ these equations simplify considerably. If we denote by $\Psi_i(j_{\ell+1})$, $i = 0, 1$, the left eigenvector of the matrix $|\langle j_{\ell+1} \,|\, T \,|\, j_\ell\rangle|$ with the two largest

eigenvalues, Λ_i, then

$$F_\ell^i(j_\ell) \xrightarrow[N-\ell\to\infty]{} c_i \Psi_i(j_\ell) \tag{28}$$

so

$$P^{\rm opt}_{j_{\ell+1}j_\ell}(\ell) \xrightarrow[N-\ell\to\infty]{} \frac{\Psi_i(j_{\ell+1})\,|\,\langle j_{\ell+1}|\,T\,|j_\ell\rangle\,|}{\Lambda_i \Psi_i(j_\ell)} \,. \tag{29}$$

If the matrix elements of T are all positive definite, then Eq. (29) becomes the results of Kalos, Levesque and Varlet.[4] Note that Eq. (29) is exact for all ℓ if we are farsighted enough to choose $\chi^i_{jN+1} = \Psi_i(j_{N+1})$.

Similarly we find that

$$p^{i\,\rm opt}_{j_1} = \frac{F_1^i(j_1)|\,\langle j_1|\,\phi_i\rangle|}{\sum_{j_1} F_1^i(j_1')|\,\langle j_1'|\,\phi_i\rangle|} \,. \tag{30}$$

If we choose $\langle j_1\,|\,\phi_i\rangle = \Psi_i(j_1)$ then for large N, Eq. (30) reduces to

$$p^i_{j_1} = |\Psi_i(j_1)|^2 \,, \tag{31}$$

again in agreement with Kalos, Levesque and Verlet[4] when the matrix elements of T are positive definite. The optimal choice of probabilities leads to a total score

$$S = \prod_{\ell=1}^{N} \frac{\langle j_{\ell+1}|\,T\,|j_\ell\rangle}{|\,\langle j_{\ell+1}|\,T\,|j_\ell\rangle\,|} \cdot \frac{\langle j_1|\,\phi_i\rangle}{|\,\langle j_1|\,\phi_i\rangle|} \cdot \sum_j F_1^i(j)\;\langle j|\,\phi_i\rangle \,. \tag{32}$$

Clearly if all the matrix elements of T and $\langle j_1\,|\,\phi_i\rangle$ are positive, then S is completely independent of intermediate states and there are no statistical fluctuations. In general if we could choose $\langle j\,|\,\phi_i\rangle = \langle j\,|\,\chi_i\rangle = \Psi_i(j)$ then the magnitude of S would simply be Λ_i.

Of course, obtaining the wave functions Ψ_i analytically is as difficult as solving the original problem. But as has been emphasized by Kalos, one can make very considerable progress by using variational trial functions in determining the transition probabilities. In the spin models we are presently considering, it is important to keep the Monte Carlo algorithm local in order for the calculation to be tractable. This means that we must make some compromises in our choice of transition probabilities. Let us consider the Ising model with the transfer matrix of Eq. (2) as an example. It is convenient to take the intermediate states to be eigenstates of σ_z^i. Then they will also be eigenstates of the symmetry operator, P, with eigenvalue $+1$ if an even number of spins are down and eigenvalue -1 if an odd number of spins are down. In this basis the terms $e^g + e^{-g}\sigma_z^i$ become c-numbers and the operator

$$e^{g\sigma_z^i\sigma_z^j} = \cosh g + \sinh g\, \sigma_x^i\sigma_x^j \tag{33}$$

either flips a pair of neighboring spins or leaves them unchanged. Thus we can sweep through the lattice testing whether to flip each pair of nearest neighbor spins in turn. At the simplest level we could take the spin flip probability equal to $e^{-g}\sinh g$; however, it is clearly preferable to build into the spin flip probability the effect of the σ_z^i factors. By choosing an appropriate O_T (see Ref. 9) we can take into account the configurations of the neighboring spins. For example, in the ferromagnetic domain we would want to increase the probability of a flip that tended to align the spins under consideration with their neighbors. The full transition probability between a pair of intermediate states is, of course, the product over all the elementary two spin probabilities, and the corresponding score the product of the elementary scores.

In addition to guiding the choice of intermediate states by an appropriate selection of the local transition probabilities, we can also provide guidance based on the total score over the entire spatial lattice for each imaginary time interval (application of H_1 and H_2).

We have implemented such global guidance in two ways. The first is a simple modification of the original projector method and the second is the population method of Kalos.

A. Modified Projector

We assume that N applications of the transfer matrix are sufficient to project out $|\psi_i\rangle$ from $|\phi_i\rangle$ as in Eqs. (10) and (11). (Of course, N depends on the accuracy one desires. In practice it must be determined experimentally.) We first make a few (L) passes though the lattice as described in connection with Eq. (15) to obtain a rough approximation to the average score for $T^N \mid \phi_i\rangle$.

$$\overline{S}_i(N) = \frac{1}{L} \sum_{\alpha} S_{j_{N+1} j_N} S_{j_N j_{N-1}} \cdots S_{j_2 j_1} s^i_{j_1} \tag{34}$$

We are then ready to begin the calculation in earnest. At each pass through the lattice we make N steps as in the original projector method obtaining an overall score

$$S_i(N) = S_{j_{N+1} j_N} \cdots S_{j_2 j_1} s^i_{j_1} \ . \tag{35}$$

We then form the ratio

$$S_i(N)/\overline{S}_i(N) = I + r \tag{36}$$

where I is an integer and r is a number between zero and one. In proceeding to calculate the eigenvalues λ_i or the matrix elements $\langle\psi_i|\ Q\ |\psi_i\rangle$ we use the configuration $|j_{N+1}\rangle$ not once as in the original projector method but I times (with

probability $1-r$) or $I+1$ times (with probability r). Thus the total probability for using a configuration $|j_{N+1}\rangle$ is proportional to the product of elementary probabilities for reaching the state times the product of the corresponding elementary scores. The overall score for reaching the state $|j_{N+1}\rangle$ is therefore independent of the intermediate states $|j_i\rangle, \ldots |j_N\rangle$. Equation (15) now simplifies considerably. For example,

$$(\lambda_i)^M = \frac{\sum_\alpha \langle \chi_i | j_{N+M+1}\rangle \, S_{j_{N+M+1},j_{N+M}} \cdots S_{j_{N+1},j_N}}{\sum_\alpha \langle \chi_i | j_{N+1}\rangle} \ . \tag{37}$$

B. Population Tracking

The second approach to global guidance that we have investigated is the population method of Kalos.[4] In this approach one provides guidance at each application of the transfer matrix or in other words at each time interval.

We begin by computing a rough average score for the first time interval, $\overline{SC}_{oi}$. The main calculation is then carried out not with one lattice configuration at a time, but with a population of configurations. Suppose we begin with L_0 configurations. For each of these we calculate the score for the first time interval. Comparing the individual scores to $\overline{SC}_{oi}$, we multiply or delete configurations just as in the modified projector method. At the end of the first time interval we are left with L_1 configurations each having a score $\overline{SC}_{oi}$.

In order to stabilize the population we take the trial score for the second time interval to be

$$\overline{SC}_{1,i} = \overline{SC}_{0,i}(L_1/L_0) \ , \tag{38}$$

since decreasing $\overline{SC}_{1,i}$ will tend to increase the population and vice versa.

After N applications of the transfer matrix or equivalently after N time intervals we will have a population of L_N configurations, each carrying the score

$$\begin{aligned} S_i(N) &= \overline{SC}_{N-1,i}\,\overline{SC}_{N-2,i}\cdots\overline{SC}_{0,i} \\ &= L_{N-1}\,L_{N-2}\cdots L_1(\overline{SC}_{0,i}/L_0)^{N-1}\ . \end{aligned} \tag{39}$$

If N is large enough to have projected out $|\psi_i\rangle$ from $|\phi_i\rangle$, then, we can calculate λ_i by simply going an additional step and obtaining the average score for the entire population. Of course, one can increase statistics by going a number of steps beyond N, multiplying or deleting configurations as just described, but keeping a running average of the scores. The important point is that at every step beyond N the distribution of states within the population gives a snapshot of $|\psi_i\rangle$. The particular basis state $|j\rangle$ will occur with frequency proportional to $\langle j\,|\,\psi_i\rangle$.

The calculation of $\langle\psi_i|\,Q\,|\psi_i\rangle$ is straightforward in the population method if Q is diagonal in the basis states $|j_\ell\rangle$. In this case one merely measures the value of Q for each member of the population after the N^{th} application of the transfer matrix. Then according to Eq. (11) one must proceed an additional N steps carrying this value of Q with each population member. The average value of Q for the population that survives after the $2N^{th}$ step is a measure of $\langle\psi_i|\,Q\,|\psi_i\rangle$.

If Q is not diagonal in the basis states, then its application will modify each member of the population. We can imagine applying Q stochastically so there is both a transition probability and a corresponding score, S_Q. We must propagate both the original and the modified populations forward for an additional N steps in accordance with Eq. (11). The expectation value of Q is then given by

$$\langle\psi_i|\ Q\ |\psi_i\rangle = \frac{\prod_{j=N+1}^{2N-1}(L'_j/L'_N)\sum\langle\chi|\,j'_{2N+1}\rangle\,S_Q}{\prod_{j=N+1}^{2N-1}(L_j/L_N)\sum\langle\chi|\,j_{2N+1}\rangle}\ . \tag{40}$$

Here the primed quantities refer to the population to which Q has been applied and the unprimed quantities to the one to which it has not been applied. The sums in Eq. (40) are over all members of the two populations that exists after $2N$ steps.

The three approaches just described are compared in Table II, where we give Monte Carlo results for the two largest eigenvalues of the two-dimensional Ising model. We present results for the intensive quantities

$$\epsilon_i = -\ell n\, \lambda_i / N_x \Delta\tau \ , \tag{41}$$

where N_x is the number of spatial lattice points. When we are dealing with the eigenvalues of the transfer matrix of a classical system $\Delta\tau \equiv 1$, while for a quantum Hamiltonian ϵ_i is the energy per site of the state in question.

For the sake of comparison we use the same number of sweeps through the lattice for each method. The computer time is roughly equal in each case, but the population method requires substantially more computer memory. There is little to choose between the methods for small systems, but for larger systems the population method is clearly superior. This superiority is even more pronounced for calculations of ground state expectation values where a detailed knowledge of the wave function $|\psi_i\rangle$ is more important than for the calculation of the λ_i.[3]

5. The Energy Gap

A quantity of particular interest is the energy gap

$$\Delta = N_x(\epsilon_1 - \epsilon_0) = E_1 - E_0 \tag{42}$$

where the ϵ_i are the densities defined in Eq. (41) for both the transfer matrix and the Hamiltonian systems. For small systems one can simply calculated E_1 and E_0 directly using the techniques discussed in Sections 2-4. However, since Δ will in general approach a finite limit as the volume of the system grows, a direct calculation will become untenable for large systems because of the large volume cancellation in the difference $E_1 - E_0$.

Fortunately the gap can be calculated directly by a measurement of intensive quantities. Let us begin by considering the transfer matrix for the Ising model given in Eq. (2). It is convenient to introduce an interpolating transfer matrix

$$T(\eta) = \frac{(e^g + \eta\, e^{-g}\sigma_z^j)}{(e^g + e^{-g}\sigma_z^j)}\, T \ . \tag{43}$$

Now using the fact that

$$\sigma_x^j T(\eta)\, \sigma_x^j = T(-\eta) \ , \tag{44}$$

the two largest eigenvalues are connected by

$$\lambda_0(\eta) = \lambda_1(-\eta) \ . \tag{45}$$

Thus $\lambda_0(\eta)$ interpolates smoothly between $\lambda_0(1)$ and $\lambda_1(1)$ as η goes from 1 to -1.

One strategy is to use the population method, with a set of probabilities $P_{ij}(\eta_0)$ and scores $S_{ij}(\eta_0)$ chosen to minimize fluctuations for $\eta = \eta_0$. We then define scores at $\eta_0 \pm \Delta\eta$ by

$$S_{ij}(\eta_0 \pm \Delta\eta) = \langle i| T(\eta_0 \pm \Delta\eta\, |j\rangle \,/ P_{ij}(\eta_0) \ . \tag{46}$$

After generating configurations with $P_{ij}(\eta_0)$ and $S_{ij}(\eta_0)$ a measurement of $S_{ij}(\eta_0 \pm \Delta\eta)/S_{ij}(\eta_0)$ at each time slice, which is an intensive quantity, yields $\lambda_0(\eta_0 \pm \Delta\eta)/\lambda_0(\eta_0)$. By choosing a few η_0's so as to completely overlap the region $-1 \leq \eta \leq 1$, we obtain $\lambda_1(1)/\lambda_0(1)$ and therefore Δ without subtracting two extensive quantities. Some sample results are shown in Table III.

An alternative approach which is particularly useful for quantum spin Hamiltonians is to make use of the Feynman-Hellmann theorem. In this case we introduce an interpolating Hamiltonian

$$H(\eta, j) = H + (1-\eta)\, h\, \sigma_z^j \ , \tag{47}$$

where H is given by Eq. (3). By our previous argument

$$E_0(\eta, j) = E_1(-\eta, j) \tag{48}$$

where $E_0(\eta, j)$ and $E_1(\eta, j)$ are the ground state and first excited state energies of $H(\eta, j)$. As a result

$$\Delta = E_1(1, j) - E_0(1, j) = \int_{-1}^{1} d\eta\, \frac{dE_0(\eta, j)}{d\eta} \ , \tag{49}$$

which is independent of the particular choice for the spin j. This formula can be used to improve the calculation of the gap Δ. The Feynman-Hellmann theorem

states that

$$\frac{dE_0(\eta,j)}{d\eta} = \langle\psi_0(\eta)|\,\frac{dH(\eta,j)}{d\eta}\,|\psi_0(\eta)\rangle \;/\; \langle\psi_0(\eta)|\,\psi_0(\eta)\rangle$$

which is

$$= -\frac{h\,\langle\chi_0|\,e^{-\beta H(\eta,j)}\sigma_z^j e^{-\beta H(\eta,j)}\,|\phi_0\rangle}{\langle\chi_0|\,e^{-2\beta H(\eta,j)}\,|\phi_0\rangle}\,. \tag{50}$$

It is convenient to work in the representation where the σ_z's are diagonal. the probabilities can then be chosen to be independent of η and hence j. Then with a *single* set of configurations one can calculate $\frac{dE_0(\eta,j)}{d\eta}$ for *all* values of η and j. One can now increase statistics by averaging over all values of j in Eq. (50). Since the result for different values of η is obtained from just one set of random walks, the fluctuations increase at values of η where the walks are not guided optimally. This point is clearly illustrated by the solid curve in Fig. 3 where the fluctuations for negative η's are much bigger than those for positive η's. However, by the symmetry expressed in Eq. (44), the curve $\frac{dE_0(\eta,j)}{d\eta}$ versus η obtained by inputting a parity even $|\phi_0\rangle$ is just the mirror reflection (about the η-axis) of the curve obtained by inputting a parity odd $|\phi_0\rangle$. The two curves are the solid and dash curves in Fig. 3 respectively. Hence, we can combine just the good statistics sections of the two curves to form a new curve (the dotted one) which yields an even more accurate result. The numerical results for moderate length runs are as follows:

$\Delta = 0.182 \pm 0.047$	solid curve
$\Delta = 0.191 \pm 0.017$	dotted curve
$\Delta = 0.191 \pm 0.026$	direct extensive subtraction
$\Delta = 0.197$	exact result

This technique is especially appealing because it is already comparable to direct extensive subtraction for an 8-site chain. For a 16-site chain, the derivative technique is decisively superior to direct subtraction. Note that these numbers and those of Fig.3 involve much shorter runs than those in Table II and III. Thus the accuracy as reflected in the fluctuations cannot be directly compared.

6. Conclusions

The main point of this paper is to emphasize that there are a large number of possible stochastic algorithms which can be used to advantage in extracting numerical results from Hamiltonian or transfer matrix formulations of quantum spin models. The most efficient algorithm clearly depends on the particulars of the model and the question asked.

In this paper we have studied the error induced in the eigenvalues by the breakup of the Hamiltonian into factorizable parts and how to minimize it. We have studied the error do to finite β and have suggested several methods for reducing this error by improving the initial and/or the final state. The improvment of the final state is most convenient in most applications because only the variational coefficients are needed. If suitable trial values of these weights cannot be found, one can use the stochastic method itself, at the expense of having to generate several sets of random walks.

The fluctuations can be reduced in a number of ways. We have shown in Section 4 that by using a good trial wave function in computing the transition probabilities could bring about a considerable improvement in convergence. In practice, we use an appropriate O_T to smoth out the breakup of H into H_1 and H_2. The other approach discussed here was the population method in which the

configurations are replicated in proportion to their score.Thus each configuration carries the same weight and hence they are all important in determinimg the final averages. The population scheme becomes superior for for large systems.

Finally, we have presented a method which extracts energy gaps without direct subtraction of two extensive quantities. The key ingredients here was the Feynman-Hellman theorem and the use of parallel scoring[12] to obtain several measurements from just one set of random walks.

REFERENCES

1. Some of the methods presented in this paper can be found in R. L. Sugar, Monte Carlo Methods in Quantum Problems, M. H. Kalos ed., D. Readel Publishing Co., Dordrecht, 1984.

2. D. Kung, to be published.

3. R. Scalletar, D. Scalapino, and R. L. Sugar, to be published.

4. M. Metropolis and S. Ulam, J. Am. Stat. Assoc. 44, 335 (1949); J. B. Anderson, J. Chem. Phys. 63, 1499 (1975); 65, 4121 (1976); 73, 3897 (1980); 74, 6307 (1981);J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos, J. Chem. Phys. 77, 349 (1982); M. H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A9, 2178 (1974); D. M. Cepereley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1979).

5. D. M. Ceperely and B. J. Alder, Phys. Rev. Lett. 45, 560 (1980); P. J. Reynolds, D. M. Cepereley, B. J. Adler, and W. A. Lester, J. Chem. Phys. (to be published).

6. S. E. Koonin in *Nuclear Theory 1981*, G. Bartsch ed., World Scientific Published Company, Singapore, 1981 and references contained therein.

7. D. W. Heys and D. R. Stump, Phys. Rev. D15, 2067 (1983); T. A. DeGrand and J. Potrin, Phys. Rev. D (to be published).

8. D. Campbell, T. DeGrand and S. Mazumdar, Phys. Rev. Lett. 52, 1717 (1984).

9. R. Blankenbecler and R. Sugar, Phys. Rev. D27, 1304 (1983).

10. J. Polonyi and J. Kuti, *Proceedings of the 6th Johns Hopkins Workshop on Current Problems in Particle Theory*, Florence, Italy; F. Alcavaz, M. Barber, A. Kennedy, J. Kuti, B. Pendelton, J. Polonyi and P. Rujan (to be published).

11. See G. E. Forsythe and R. A. Leibler, Math. Tabs. Aids Comput. 4, 127 (1950) and Chapter 7 of J. M. Hammersley and D. C. Handscomb, Monte Carlo Methods (Methuen & Co., Ltd., London, 1964).

12. D. Dahl and R. Blankenbecler, SLAC-PUB-3484, to be published.

Table I

Comparison of the ground state energy of a 4-site Heisenberg antiferromagnet obtained by using different initial states $|\phi\rangle$ and using Eq. (24a).

4 site Heisenberg antiferromagnet

Exact E_0 = -3.

Using $|\phi\rangle$ as input state.

E_0	b
−3.13	0.0
−3.00	0.5
−2.94	1.0
−2.90	1.5
−2.88	2.0
−2.86	2.5

Using (24a) $E_0 = -2.99$

Table II

$g = 0.40$

Table of the ground and first excited state energies for different projector methods.

$N_x = 8$	$E_0(Exact) = 0.8816$	$E_1(Exact) = -0.8558$
Method	$E_0(Monte\ Carlo)$	$E_1(Monte\ Carlo)$
Projector	-0.8819 ± 0.0002	-0.8552 ± 0.0003
Modified Projector	-0.8813 ± 0.0002	-0.8558 ± 0.0002
Population	-0.8816 ± 0.0001	-0.8559 ± 0.0001
$N_x = 16$	$E_0(Exact) = -0.8796$	$E_1(Exact) = -0.8687$
Projector	-0.8798 ± 0.0011	-0.8681 ± 0.0009
Modified Projector	-0.8794 ± 0.0004	-0.8682 ± 0.0006
Population	-0.8797 ± 0.0001	-0.8689 ± 0.0002
$N_x = 24$	$E_0(Exact) = -0.8794$	$E_1(Exact) = -0.8723$
Projector	-0.8783 ± 0.0014	-0.8655 ± 0.0014
Modified Projector	-0.8794 ± 0.0003	-0.8724 ± 0.0009
Population	-0.8790 ± 0.0001	-0.8724 ± 0.0001

Table III

GAP Results

$g = 0.40$

Table of Δ obtained by the overlap method versus the exact result for a 2-d Ising model. N_x is the number of sites in the x-direction.

N_x	*Monte Carlo*	*Exact*
2	0.525 ± 0.001	0.525
4	0.300 ± 0.001	0.298
6	0.233 ± 0.002	0.234
8	0.208 ± 0.002	0.206
10	0.189 ± 0.002	0.191
12	0.180 ± 0.002	0.183

FIGURE CAPTIONS

1. Ground state energy versus β for different final state $|\chi\rangle$ for an 8-site 1-d Ising Hamiltonian in a transverse field.
2. Ground state energy versus $\Delta\tau^2$ for two different breakups of the Hamiltonian.
3. The ground state expectation value of σ_z versus η in an 8-site Ising Hamiltonian described by Eq. (47).

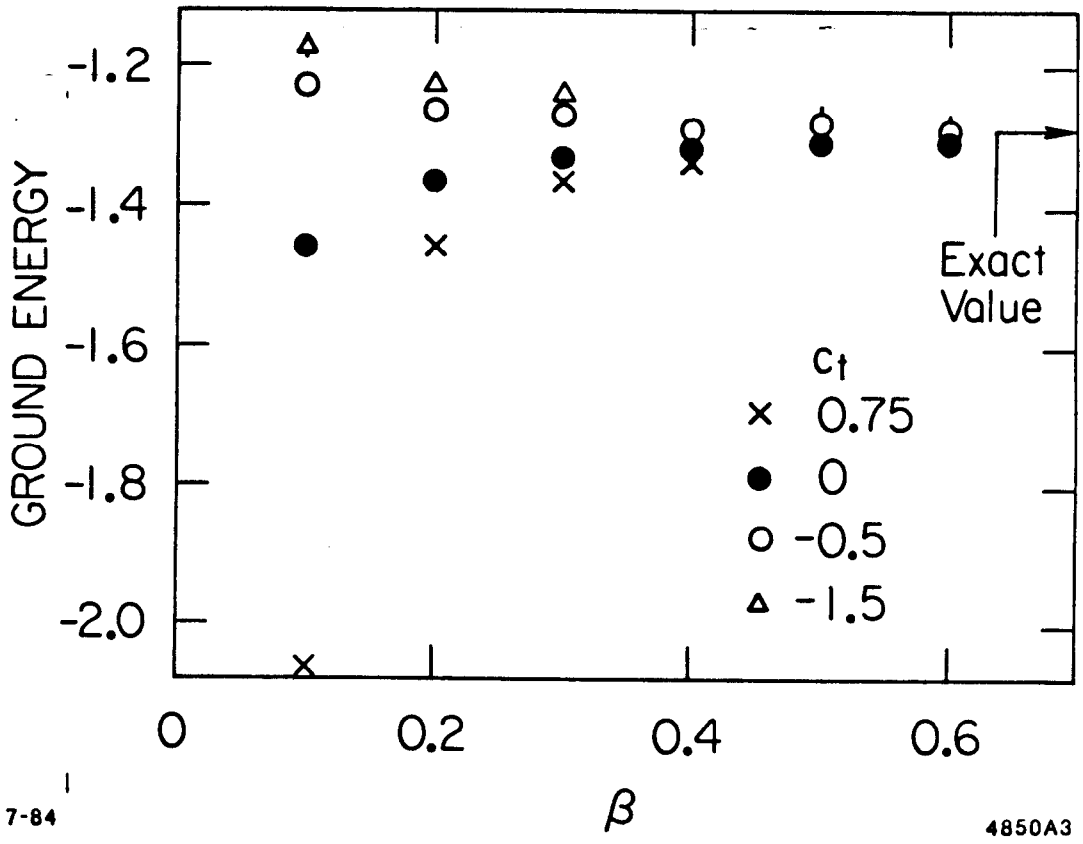
GROUND ENERGY
-1.2
-1.4
-1.6
-1.8
-2.0
0
0.2
0.4
0.6
β
Exact Value
c_t
× 0.75
● 0
○ -0.5
△ -1.5
7-84
4850A3

Fig. 1

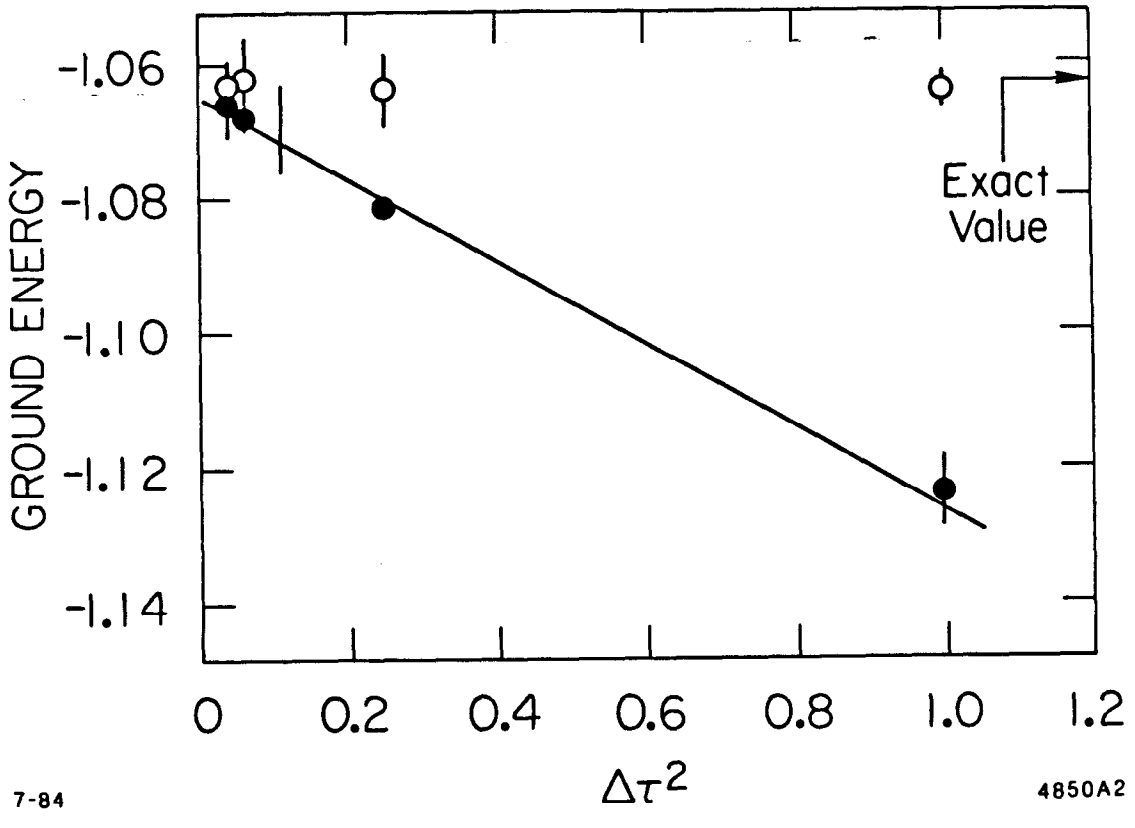
-1.06
-1.08
-1.10
-1.12
-1.14
GROUND ENERGY
0
0.2
0.4
0.6
0.8
1.0
1.2
$\Delta\tau^2$
Exact
Value
7-84
4850A2

Fig. 2

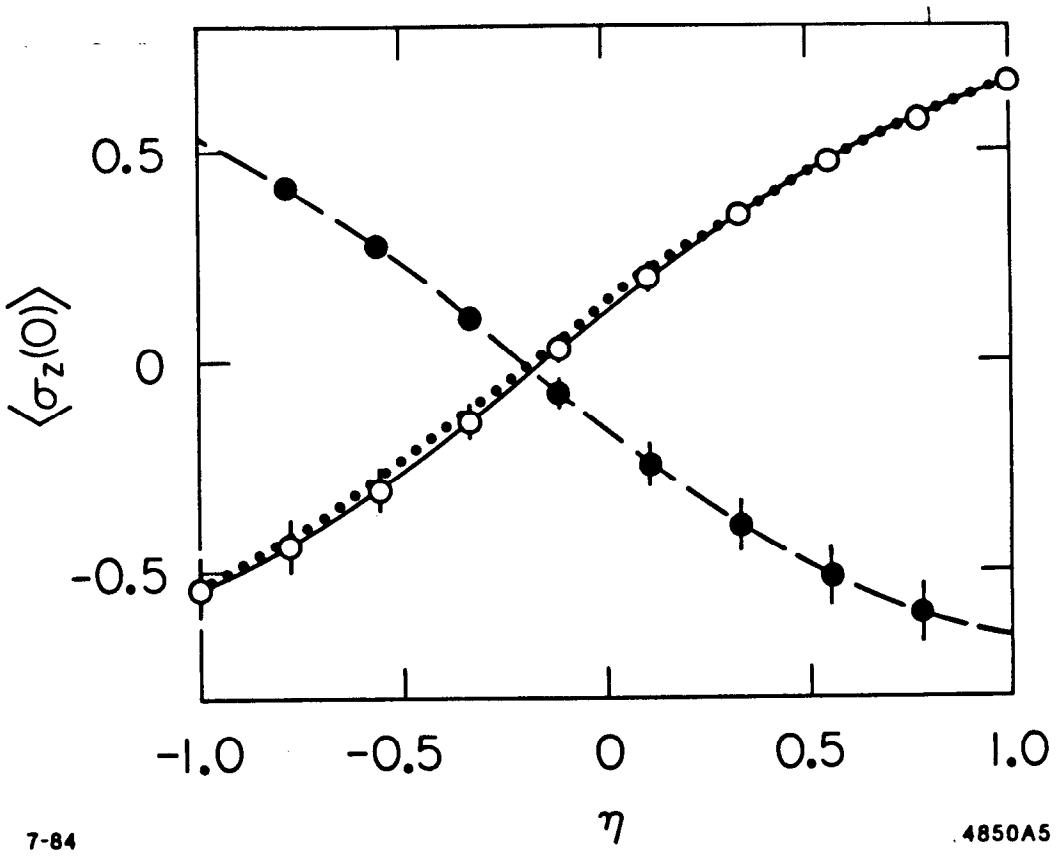

Fig. 3