Magnetic Properties of Some Itinerant-Electron Systems at $T > 0$

Michael Aizenman[(a)]
Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

Elliott H. Lieb
Departments of Physics and Mathematics, Princeton University, P. O. Box 708, Princeton, New Jersey 08544
(Received 6 June 1990)

The Lieb-Mattis theorem on the absence of one-dimensional ferromagnetism is extended here from ground states to $T>0$ by proving, *inter alia*, that $M(\beta,h)$, the magnetization of a quantum system in a field $h>0$, is always *less* than the pure paramagnetic value $M_0(\beta,h)=\tanh(\beta h)$, with $\beta\equiv 1/kT$. Our proof rests on a new formulation in terms of path integrals that holds in any dimension; another of its applications is that the Nagaoka-Thouless theorem on the Hubbard model also extends to $T>0$ in the sense that $M(\beta,h)$ *exceeds* $M_0(\beta,h)$.

PACS numbers: 75.10.Lp, 05.30.Fk

To many physicists the study of magnetism begins with a Heisenberg model, or something similar, in which localized spins interact with each other. Unfortunately, nature does not present us with such a starting point. Instead, at a reasonably (though not absolutely) fundamental level, we are presented with the Schrödinger equation for electrons whose interaction is, to a very good approximation, spin independent. It is left to the Pauli principle to induce a spin dependence, and for this it is essential that the electrons are *itinerant*, i.e., *not* localized *a priori*.

Given the subtlety of the phenomenon, it is crucial to elucidate the conditions which encourage or discourage the emergence of ferromagnetism. Some years ago Lieb and Mattis[1] showed that one *never* gets ferromagnetism in the ground states of itinerant electrons on a line. The results presented here include an extension of that earlier theorem (by a different method) to positive temperatures, and an opposite statement for systems in any dimension with a certain *parity constraint* (requiring *hard-core* repulsions)—for which we show that parallel spin alignments *are* favored at all temperatures. Examples of the second kind are one-dimensional systems of odd numbers of electrons on a ring (i.e., *periodic* boundary conditions), for which we extend the $T=0$ result of Herring,[2] and a particular case of the Hubbard model whose ground state was shown to be ferromagnetic by Nagaoka[3] and Thouless.[4,5]

We deal with the following general (spin-independent) Hamiltonian for N identical, itinerant particles of mass m ($\hbar=1$):

$$H=-(2m)^{-1}\sum_{i=1}^{N}\nabla_i^2+V(x_1,x_2,\ldots,x_N)\,, \tag{1}$$

where V is some potential-energy function[6] which is symmetric in its N variables. Boundary conditions are stated below. We consider simultaneously (a) the continuum model (as just stated) and (b) the lattice model in which the x_i's have integer components and ∇^2 is the second difference operator.

If the particles are spin-$\frac{1}{2}$ fermions (e.g., electrons), the eigenstates of H can be classified according to the eigenvalues (j,s) of the total spin angular momentum operator J and of S—the z component ($j=N/2$, $N/2-1,\ldots,0$ or $\frac{1}{2}$, and $|s|\le j$). The eigenvalues of H depend on j but (for j specified) not on s. The ground-state energies are denoted by $E_0(j)$. Two simple benchmarks are the following: (i) Immobile, noninteracting, particles—with $m\to\infty$ and $V=0$. In this limit $E_0(j)$ is *independent* of j. The system is a *pure paramagnet*, with spins responding to a magnetic field independently of each other. (Recall that $S=\sum\sigma_i^z/2$.) In a uniform field, $\mathbf{h}=(0,0,h)$, the partition function $Z(\beta,h)=\mathrm{Tr}\exp[-\beta(H-\sum\mathbf{h}\cdot\boldsymbol{\sigma}_i)]$ satisfies

$$Z(\beta,h)/Z(\beta,h=0)\equiv\langle e^{2\beta hS}\rangle_{h=0}=[\cosh(\beta h)]^N$$

in this limit, and the magnetization is

$$M(\beta,h)=d[\ln Z(\beta,h)]/d(\beta h)=N\tanh(\beta h)\,. \tag{2}$$

(ii) Noninteracting itinerant system, with $V=0$ and $0<m<\infty$. The spectrum can be analyzed in terms of one-particle states, and one finds

$$E_0(j+1)>E_0(j) \tag{3}$$

for all $j<N/2$. Thus, the "noninteracting" itinerant system shows a tendency towards *anti*ferromagnetism (resulting from the Pauli principle, which allows two electrons to occupy the same one-particle eigenstate of ∇^2 only if their spins are antialigned). At $T=0$ and weak field the magnetization is 0 or 1.

The theorem of Lieb and Mattis[1] states that (3) always holds in 1D, even when $V\neq 0$. Although the theorem has some practical value for one-dimensional systems, its main conceptual value lies in the observation that theories of ferromagnetism based naively on Heisenberg's "exchange integral" can be misleading since two-body potentials with either ferromagnetic or antiferromagnetic exchange integrals can easily be constructed in one dimension.[7]

Equation (3) states that increasing the spin costs ground-state energy. The first aim of this paper is to provide similarly precise information about the free energy when the inverse temperature β is finite. We denote the partition function with the trace restricted to subspaces with specified values of j, s, or both by $Z_{JS}(\beta, h;j,s)$, $Z_S(\beta,h;s)$, and $Z_J(\beta,h;j)$. When h is omitted, it is understood to take the value $h=0$. We also denote by $Y_{JS}^{(N)}(j,s)$, $Y_J^{(N)}(j)$, and $Y_S^{(N)}(s)$, the dimensions of these spaces for N noninteracting spins. The Y's [see (9) *et seq.*] serve as natural normalization constants for the Z's [because Eqs. (4) and (5) become equalities in the pure paramagnetic case].

Theorem 1.—Let N spin-$\frac{1}{2}$ fermions have the Hamiltonian (1) (with $m<\infty$) in a bounded region $[-L,L]^N$, with Dirichlet ($\psi=0$), Neumann ($\psi'=0$), or sticky $[\psi'/\psi(\pm L)=\lambda_\pm]$ boundary conditions (either in the continuum or on a lattice). Then, for $\beta<\infty$,

$$Z_S(\beta;s+1)/Y_S^{(N)}(s+1) < Z_S(\beta;s)/Y_S^{(N)}(s)\,, \qquad (4)$$

$$Z_J(\beta;j+1)/Y_J^{(N)}(j+1) < Z_J(\beta;j)/Y_J^{(N)}(j)\,, \qquad (5)$$

for all j and $s\ge 0$. In a field $h>0$ the magnetization satisfies

$$0 < M(\beta,h) < N\tanh(\beta h)\,. \qquad (6)$$

Furthermore, there are strictly positive functions C_N, $C_{N-2}, C_{N-4},\ldots,C_0$ or C_1 (which depend on β but not on s) such that, for $|s|\le N/2$,

$$Z_S(\beta;s)=\sum_{k=0}^{N} C_k(\beta)Y_S^{(k)}(s)\,, \qquad (7)$$

and

$$Z(\beta,h)=\sum_{k=0}^{N} C_k(\beta)[\cosh(\beta h)]^k\,, \qquad (8)$$

where it is understood that $C_k\equiv 0$ if $N-k$ is odd, and $Y_S^{(k)}(s)\equiv 0$ if $k<2|s|$.

Remarks.—(1) Equations (7) and (8) are equivalent. They assert that, as far as the total spin is concerned, the thermal equilibrium state is effectively a (positive) superposition of states corresponding to systems of fewer (k) spins with no interaction.

(2) It is false that the susceptibility $\chi(h)\equiv\beta^{-1}\partial M(\beta, h)/\partial h$ does not exceed the pure paramagnetic value $N\cosh^{-2}(\beta h)$. By an elementary calculation using (8), the opposite is *always* the case for *very large h* (with the threshold possibly dependent on N).

(3) Equation (8) is derived below from (13); before turning to it let us note that (4) and (5), which are *not* derivable from each other, follow from (7) by elementary algebra. To do that, let us first recall that by the theory of angular momentum, for $|s|\le j$,

$$Z_{JS}(\beta;j,s)=Z_{JS}(\beta;j,j)=Z_S(\beta;j)-Z_S(\beta;j+1)$$
$$=Z_J(\beta;j)/(2j+1)\,. \qquad (9)$$

Similar identities hold for $Y_{JS}^{(N)}(j,s)$, $Y_J^{(N)}(j)$, and $Y_S^{(N)}(s)$, with $Y_S^{(N)}(s)=N!/(N/2+s)!(N/2-s)!$.

To derive (4) from (7), one needs to establish the monotonicity in $|s|$ of $Y_S^{(k)}(s)/Y_S^{(N)}(s)$, for $N-k>0$ even. To do that, it is convenient to write this ratio as a telescopic product, whose factors are the monotone functions

$$Y_S^{(k)}(s)/Y_S^{(k+2)}(s)=[(k/2+1)^2-s^2]/(k+2)\,.$$

Equation (5) is derived by a similar tactic. Equation (6) is an elementary consequence of (8).

In Theorem 1 it is assumed that the interaction has no hard-core repulsion, or other singularities which are strong enough to induce nodal surfaces in the configuration space. The only role of this assumption is to allow (4)-(6) to be stated as *strict* inequalities. For our second result we let H include such repulsive interactions, which are the *sine qua non* for the *parity restriction* which is essential for Theorem 2. The term "dynamically allowed" refers there to those permutations that can be achieved by a motion of the particles that does not pass through a nodal surface induced by the positive divergence of the potential. For example, in 1D permutations are restricted if the interaction includes a hard-core repulsion which disallows particle encounters.

Theorem 2.—Suppose that in a system of N spin-$\frac{1}{2}$ fermions with Hamiltonian (1), in arbitrary dimension and in the continuum or on a lattice, the dynamically allowed permutations are all *even* and not restricted to the identity. (This parity requirement can be satisfied only if H includes a *hard-core repulsion.*) Then: (i) Among the ground states there is one with $j=N/2$. (ii) In a field $h>0$ the magnetization satisfies

$$M(\beta,h) > N\tanh(\beta h)\,. \qquad (10)$$

(iii)

$$Z(\beta,h)=\sum_{\{n\}:\sum n_i=N} D_{\{n\}}(\beta)\prod_i\cosh(\beta h n_i) \qquad (11)$$

for some $D_{\{n\}}(\beta)\ge 0$, with $\{n\}=\{n_1,n_2,\ldots\}$ varying over partitions of N. $D_{\{n\}}(\beta)>0$ if and only if there is an allowed permutation whose cycle lengths are $n_1,n_2,\ldots$.

Remarks.—(1) The content of (11) is that as far as the z-component S is concerned, at thermal equilibrium the system is in a superposition of states in which the particles form independent "cliques" of sizes n_i whose contribution to S is $\pm n_i/2$. In such a state the system shows enhanced response to magnetic fields. In particular, Eq. (10) [which should be contrasted with (6)] follows from (11) by writing $M(\beta,h)=E(\sum_j n_j \times\tanh(\beta h n_j))$, where $E(\cdot)$ is a normalized average over partitions with weights proportional to $D_{\{n_i\}} \times\prod_i\cosh(\beta h n_i)$, and applying the strict monotonicity of $\tanh(\cdot)$. By a similar argument, all the moments $\langle S^k\rangle$ exceed their paramagnetic values. (A similar clique structure can be inferred for the ferromagnetic Ising model from the Fortuin-Kasteleyn[8] representation.) Assertion (i) follows from the uniform bound $Z_S(\beta,N/2)/Z(\beta) \geq 2^{-N}$, implied by (11).

(2) Two examples to which Theorem 2 applies are the following: (a) The Hubbard model in the strong-coupling limit, with one hole on a bipartite lattice. In this case part (i) [in the stronger sense that $E_0(j)$ is strictly lowest at $j = N/2$] was established by Nagaoka and Thouless.[3-5] Our contribution is to extend their result to positive temperatures, in the sense of (10). (b) Systems of *odd* numbers N of spin-$\frac{1}{2}$ fermions in a ring, i.e., in $[0,L]^N$ with periodic boundary conditions, and a *hard-core repulsion.* The dynamically allowed permutations are powers of the cyclic shift $P \equiv (1,2,\ldots,N)$, whose parity equals $(-1)^{N-1} = 1$. Part (i) was proved in that case by Herring.[2] Herring's and our method also yield[9] $E_0(N/2-1) > E_0(N/2)$.

(3) Theorem 2 has a very restrictive hypothesis which is not satisfied in cases of direct physical interest. Its value lies in two points: (i) the light it sheds on the mechanism of ferromagnetism; (ii) the construction used in the proof, which introduces ideas not ordinarily encountered in the study of electron correlations, e.g., the notion of cliques and the representation (13) which is always valid regardless of the hypothesis.

We turn now to the proof of (8) and (11), by first analyzing the Feynman-Kac path-integral representation[10] of $\mathrm{Tr}\, e^{-\beta H}$:

$$Z(\beta,h) = \int_{\Omega_\beta} \rho(d\omega) \exp\left[-\int_0^\beta V(x(t))dt\right](-1)^{\pi(\omega)} \times \exp\left[\beta h \sum_{i=1}^{N} \sigma_i\right]. \quad (12)$$

Here, $\omega = \{(x(t),\sigma(t)) | 0 \leq t \leq \beta\}$, which we refer to as a "path," represents a trajectory of the N particles, with $x = (x_1,\ldots,x_N)$, $\sigma = (\sigma_1,\ldots,\sigma_N)$. Ω_β is the collection of paths such that (i) the spins $\sigma_i(t) = \pm 1$ are constant in time, i.e., independent of t; (ii) the particle configuration at time β is a permutation, $\pi(\omega)$, of the configuration at time zero, with $(-1)^{\pi(\omega)}$ the parity of that permutation; and (iii) there are no encounters of particles of equal spin. $\rho(d\omega)$ is a (*positive*-) probability measure on Ω_β, determined by just the kinetic component of H, which we now describe in more detail for the two cases considered here.

In the discrete case, with x_i taking values on some lattice, the particle trajectories are piecewise constant functions of time, and hop to neighboring sites independently and at random times. The jumps form a Poisson process, with the probability ρ that a specified particle does not jump during the time interval $t \in [a,b]$ given by $\exp(-|a-b|m)$.

In the continuum case, $\rho(d\omega)$ is a product of Wiener measures constrained to satisfy the periodicity requirement (ii); i.e., $x(t)$ is a "Brownian bridge" with $x_i(\cdot)$ continuous in time. For the Dirichlet Laplacian the paths are killed at the boundary; for Neumann they are "reflected" (in a sense thoroughly discussed in textbooks).

The third condition is absent *a priori*, but it is a standard observation that its inclusion does not affect the value of the integral. (The proof of this version of the Pauli exclusion principle can be based on the symmetry transformation that exchanges the trajectories leading to the earliest equal-spin encounter.) A useful implication for us is that paths in Ω_β have no triple encounters; this follows from the other properties of Ω_β and the fact that the spins take only two values.

The measure in (12) possesses a hidden symmetry which we shall uncover. We are going first to define for each path ω a collection $\Gamma(\omega)$ of loops γ. Each loop will have a winding number $w(\gamma)$. There will be a symmetry group $\mathcal{G}$ having 2^N elements each of which maps Ω_β in a one-to-one way onto Ω_β preserving $\Gamma(\omega)$, $\rho(d\omega)$, $\int V(x(t))dt$, and $(-1)^{\pi(\omega)}$. However, $\mathcal{G}$ does not preserve $\pi(\omega)$ and the spin values $s = \sum \sigma_i$. The values attained by s within the ensemble of paths related by the action of $\mathcal{G}$ are of the form $\sum_{\gamma \in \Gamma(\omega)} \pm w(\gamma)/2$, with all the configurations of $\pm$ signs occurring with equal weights. Symmetrization of the integrand in (12) then yields the following useful representation:

$$Z(\beta,h) = \int_{\Omega_\beta} \rho(d\omega) \exp\left[-\int_0^\beta dt\, V(x(t))\right](-1)^{\pi(\omega)} \times \prod_{\gamma \in \Gamma(\omega)} \cosh[\beta h w(\gamma)]. \quad (13)$$

The loops are associated with the space-time picture of the collection of particle trajectories of the path ω (see Fig. 1). They are drawn by starting from the $t=0$ location of any of the particles, and tracing the particle's location forward in space-time until its first encounter with another particle; at which point the tracing line switches to the world line of the other particle in the reversed orientation in time. Such an orientation switch is repeated whenever a particle encounter is reached. When the trace line reaches the time $t=0$ or β, it reemerges at the

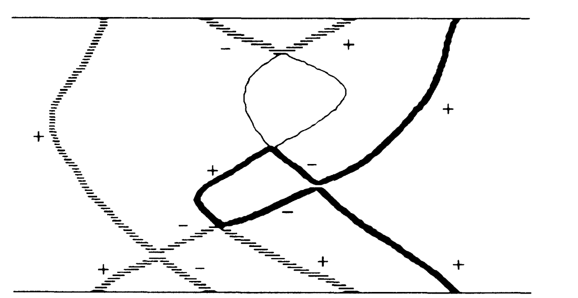

FIG. 1. This collection of particle world lines has two loops with winding numbers $\{1,1\}$. The statistical distribution of s under the symmetry group is that of two noninteracting spin-$\frac{1}{2}$ particles. In dimensions $d > 1$, and for 1D periodic boundary conditions, permutations can occur without there being particle encounters, and then the winding numbers can be larger.

same location with time treated as periodic. This is continued until a loop is closed. For each $\omega \in \Omega_\beta$ the loops will all have only finite numbers of orientation reversals, since that number can diverge only if there is a triple encounter. We denote the collection of the distinct loops (γ) by $\Gamma(\omega)$, and define $w(\gamma)$ to be the absolute value of the loop's winding number, i.e., $w(\gamma)=|\int_\gamma dt/\beta|$.

A path $\omega \in \Omega_\beta$ is completely characterized by three things: the initial configuration $x(0)$, the union of the world lines, and the spin assignments along the world lines. It is evident that consistency with the constraints (i)-(iii) is achieved if and only if the spin assignment alternates along each loop, as shown in Fig. 1. By this condition, reading along the $t=0$ line, the admissible values of $2s$ are of the form $\sum_i \sigma_i = \sum_{\gamma \in \Gamma(\omega)} \pm w(\gamma)$.

We now define the group $\mathcal{G}$ to have N commuting generators g_i. In the space-time picture the map g_i flips the spin values $(+\frac{1}{2} \leftrightarrow -\frac{1}{2})$ along the loop γ containing the $t=0$ position of the ith particle. The spin flip is accompanied by the unique rearrangement of particle trajectories needed to maintain the constraints (i)-(iii). The paths ω and $g_i\omega$ differ in the $\pm$sign attached to $w(\gamma)$.

To show that g_i preserves the parity of $\pi(\omega)$, we note that $\pi(g_i\omega)$ is obtained from $\pi(\omega)$ by a product of transpositions associated with the orientation reversals in the loop γ; their number is always even. The invariance of $\rho(d\omega)$ is evident in the lattice case with discrete time steps. The invariance in the continuum case can be deduced by verifying that g_i are continuous maps in the uniform topology of Ω_β and invoking Donsker's theorem[11] on the weak convergence of the random-walk measures to the Wiener measure.

Equation (13)$\rightarrow$(8): Evidently, on a line $\pi(\omega)$ can only be the identity, and $w(\gamma)$ can only take the values 0 and 1. Then (13) directly implies (8) with $C_k(\beta)$ given by the integral in (13) restricted to ω's for which the number of loops with $w(\gamma)=1$ is exactly k. The strict positivity is assured by the nonvanishing contributions from paths ω with collisions occurring only within the first $(N-k)/2$ pairs of adjacent particles. Q.E.D.

Equation (13)$\rightarrow$(11): The hard-core repulsion precludes encounters, so the loops are dynamically simple but have nontrivial winding numbers, satisfying $\sum w(\gamma) = N$. Q.E.D.

Finally, we remark that Theorem 1 admits a *parastatistical* extension—as did the ground-state theorem of Ref. 1.

This research was supported in part by NSF Grants No. PHY-892067 (M.A.) and No. PHY-8515288-A04 (E.H.L.).

[a]Also in the Physics Department. Present address: Departments of Physics and Mathematics, Princeton University, Princeton, NJ 08544.

[1]E. H. Lieb and D. C. Mattis, Phys. Rev. **125**, 164–172 (1962).

[2]C. Herring, Phys. Rev. B **11**, 2056 (1975).

[3]Y. Nagaoka, Phys. Rev. **147**, 392 (1967).

[4]D. Thouless, Proc. Phys. Soc. London **86**, 893 (1965).

[5]E. H. Lieb, in *Phase Transitions, Proceedings of the Fourteenth Chemistry Solvay Conference* (Wiley, New York, 1971), pp. 45–63.

[6]The restrictions on V are fairly lax:

$$\sup_x \int_{|y|\le a} d^{Nd}y \|y\|^{-(Nd-2)} |V(y-x)| \rightarrow 0$$

as $a \rightarrow 0$. If V consists of k-body interaction, that condition translates to a lower-dimensional condition of similar form, with the weight becoming $\ln\|y\|$ for $kd=2$, and 1 for $kd=1$.

[7]R. E. Peierls, *Quantum Theory of Solids* (Oxford Univ. Press, New York, 1955).

[8]C. M. Fortuin and P. W. Kasteleyn, Physica (Utrecht) **57**, 536 (1972).

[9]At $T=0$ all the powers of P occur with equal weights in Eq. (13). By Eq. (9), $Z_J(\beta,N/2-1)/Z_J(\beta,N/2) \rightarrow 0$ as $\beta \rightarrow \infty$.

[10]B. Simon, *Functional Integration and Quantum Physics* (Academic, New York, 1979).

[11]M. Donsker, in *Four Papers on Probability*, Memoirs of the American Mathematical Society Vol. 6 (American Mathematical Society, Providence, RI, 1951).