Statistical-mechanical formalism for spin-glasses

A. C. D. van Enter and J. L. van Hemmen
Sonderforschungsbereich 123, Universität Heidelberg, D-6900 Heidelberg 1, Federal Republic of Germany
(Received 13 December 1982)

A consistent description of the spin-glass transition and the spin-glass phase is possible within the context of *equilibrium* statistical mechanics. This conclusion is based on two arguments. First, every equilibrium state may be decomposed uniquely into its ergodic components, which represent the pure thermodynamic phases. Second, all the ergodic components have the same free energy. A detailed analysis of the ensuing conceptual structure is given and its agreement with experiments on field-cooled spin-glasses is pointed out.

I. INTRODUCTION

A spin-glass is a *disordered,* magnetic system with a well-defined freezing temperature T_f such that for $T<T_f$ the magnetic moments are frozen in random orientations without a conventional long-range order. The spin-glass problem is the question of whether the freezing of the spins is a dynamical, nonequilibrium phenomenon[1] or a new thermodynamic phase with an equilibrium phase transition at $T=T_f$.[2,3] In this paper we address the spin-glass problem, with particular emphasis on metallic spin-glasses, such as *Au*Fe and *Cu*Mn.

Since there has been some debate[4–6] about the interpretation of equilibrium statistical mechanics, we will first give an overview (Sec. II) of the thermodynamic formalism as it evolves from the notation of broken ergodicity: Below T_f the canonical equilibrium state is not ergodic with respect to the dynamics but may be decomposed into several ergodic components. We present, in Sec. III, a detailed analysis of the ensuing conceptual structure and apply the equilibrium formalism to the spin-glass case. One needs to bear in mind constantly the real object of our study: a spin-glass in thermodynamic *equilibrium,* which may be obtained, for instance, through field-cooling. The interpretation of a spin-glass as a *metastable* state and the Monte Carlo one-spin dynamics are considered in Sec. IV, where we also outline some new experiments to check the dynamical picture. We give our conclusion in Sec. V. In the Appendix we prove the surprising result that, notwithstanding the randomness, all ergodic components have the same free energy.

II. THERMODYNAMIC FORMALISM

Ergodicity is a central notion in equilibrium statistical mechanics. Since the number of particles, spins, etc., is very large we will take the thermodynamic limit $N\to\infty$ explicitly[7] and discuss the conditions which are to be satisfied by an equilibrium state of an infinite system. Such an equilibrium state can be decomposed into its extremal (ergodic) components. A critical temperature T_c is such that below T_c the extremal components are ergodic whereas the equilibrium state is not. The phenomenon of broken ergodicity is closely related to the occurrence of a phase transition at T_c and the relevance of an order parameter below T_c.

A. Ergodicity

Suppose, for the time being, that we have a finite *classical* system, say, N particles in a box, which are described by a phase space Ω and a dynamical evolution T_t. That is, if ω is a point in Ω at time $t=0$, $T_t\omega$ is the position of ω at time t. An observable f is a function on the phase space Ω. If we integrate f with respect to a probability measure μ, we get the phase-space average of f,

$$\langle f\rangle=\int_\Omega d\mu(\omega)f(\omega)\ . \tag{2.1}$$

Let A denote an arbitrary "event" in Ω; $\mu(A)$ is its probability. Throughout what follows, we suppose that μ is time invariant, i.e., $\mu(T_tA)=\mu(A)$ for any A. We then have the necessary ingredients for ergodicity: a phase space Ω, a dynamics T_t which transforms Ω into itself, and a T_t-invariant probability distribution μ. We say μ is *ergodic* with respect to T_t if the empty set $\varnothing$ and Ω itself are the only sets in Ω which are invariant under T_t. More precisely,

$$T_tA\subseteq A\to\mu(A)=0\ \text{ or }\ 1. \tag{2.2}$$

As a consequence,[8] time average and phase-space average agree with probability 1,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T dt\, f(T_t\omega)=\langle f\rangle\ . \tag{2.3}$$

The set of all ω where (2.3) does not hold has probability 0 and, hence, will be discarded.

Alternatively, μ is *non*ergodic with respect to T_t if we can split up Ω into disjoint, invariant subsets Ω_i of positive measure, so that $\mu(\Omega_i)>0$ and

$$\Omega=\bigcup_i\Omega_i,\ \ \Omega_i\cap\Omega_j=\varnothing,\ \ T_t\Omega_i\subseteq\Omega_i\ . \tag{2.4}$$

Call μ_i the normalized restriction of μ to Ω_i $[\mu_i(\Omega_i)=1]$ and $\alpha_i=\mu(\Omega_i)$. Each μ_i is T_t invariant, and

$$\mu=\sum_i \alpha_i \mu_i, \quad \sum_i \alpha_i = 1 \,. \tag{2.5}$$

In other words, μ is *non*ergodic if μ may be written as a *convex* combination of other invariant probability measures. The converse is also true: One can show[9] that μ is ergodic if and only if μ can*not* be written as a convex combination of other invariant measures. In the limit $N \to \infty$ the notion of energy surface loses its meaning, but the above notion of ergodicity survives without difficulty—together with (2.3). One says that μ is ergodic or extremal invariant.

The μ_i in (2.5) live on disjoint subsets of Ω. Suppose now that μ can be written as a convex combination of two ergodic measures μ_1 and μ_2. Is it true that for disjoint Ω_1 and Ω_2

$$\Omega=\Omega_1 \cup \Omega_2, \quad \mu_1(\Omega_2)=\mu_2(\Omega_1)=0 \; ? \tag{2.6}$$

Indeed it is.[9] If μ_1 and μ_2 are both ergodic with respect to T_t, then either $\mu_1=\mu_2$ or (2.6) holds. So the extremal components of an invariant measure μ are ergodic and live on *disjoint* subsets Ω_i of the phase space.

In the thermodynamic limit the microcanonical and the canonical ensemble are equivalent. It is then convenient to think of the components Ω_i as being separated by *infinitely* high free energy barriers. Of course, infinitely high is also a matter of time scales. The point is that for practical purposes a free energy barrier may become "infinitely" high if it cannot be passed within the time available for the experiment. Here we need not go into this problem since it has already been discussed very carefully by Münster.[10]

B. Classical lattice systems

The real objects of our study are not continuous systems, but lattice spin systems which, for the sake of simplicity, are assumed to be classical and of the Ising type. Now, Ising spins do not have a dynamics by themselves, so they must be given one. Usually a Glauber dynamics[11,12] is taken, which is a *one*-spin dynamics: Only one spin can change during one step. The procedure is widely used in Monte Carlo simulations.

Throughout what follows, the interaction between the spins is taken to be a pair potential. The Hamiltonian for N spins in a magnetic field is given by

$$H_N = -\sum_{i,j} J_{ij} S_i S_j - h \sum_i S_i \,, \tag{2.7}$$

where $S_i = \pm 1$, $1 \le i \le N$, and $J_{ii}=0$. Each pair $\{i,j\}$ is counted only once. For reasons of convenience we assume the potential to be of finite range r, i.e., $J_{ij}=0$ if $|i-j|>r$. Unless stated otherwise we take free boundary conditions and $h=0$.

For a finite system consisting of N spins thermodynamic equilibrium at inverse temperature β is described by the *Gibbs* state $\mu_{\beta,N}$ with density $C\exp(-\beta H_N)$ where H_N is the Hamilton function. The Glauber dynamics is constructed in such a way that $\mu_{\beta,N}$ is invariant and even ergodic with respect to this dynamics.[13] We now turn to the problem how to describe the system as $N \to \infty$.

C. The thermodynamic limit

In a macroscopic system the order of magnitude of N is 10^{23}. Working with energy hypersurfaces in such a high-dimensional space is quite deceptive[7] and we, therefore, take the practical point of view that an *infinite* system is a reasonable approximation of a macroscopic system. Moreover, taking the system infinite is unavoidable if one wants to study equilibrium phase transitions.[14] One only may wonder what to do with $\exp(-\beta H_N)$ as $N \to \infty$.

Let μ_β or simply μ be an equilibrium state of the infinite system. It is a probability measure on the system's phase space Ω, the set of all possible spin configurations. If A and B are two events in Ω with $\mu(B)>0$ (say spin up in i and spin up in j), then the conditional probability

$$\mu(A \mid B)=\mu(A \cap B)/\mu(B) \tag{2.8}$$

is the probability that A happens *given* B. Pick an arbitrary lattice point i, and let $\Lambda(i)$ be the sphere around i with radius r but with i deleted. The J_{ij} with j outside $\Lambda(i)$ vanish. Let Λ' be a domain which contains $\Lambda(i)$; see Fig. 1. Finally we denote by S_Λ an arbitrary set of spin values S_i, $i \in \Lambda$, and by $\mu(S_\Lambda)$ the probability of getting S_Λ. Then μ is said to be an equilibrium state or Gibbs state if, whatever i, μ satisfies the Dobrushin-Lanford-Ruelle (DLR) condition[15]

$$\mu(S_i \cap S_{\Lambda' \setminus \{i\}})=\mu(S_i \mid S_{\Lambda(i)})\mu(S_{\Lambda' \setminus \{i\}}) \,, \tag{2.9}$$

where $\mu(S_i \mid S_{\Lambda(i)})$ is given by

$$\mu(S_i \mid S_{\Lambda(i)})=C\exp\left[\beta \sum_{j \in \Lambda(i)} J_{ij} S_i S_j\right] , \tag{2.10}$$

the constant C in (2.10) being chosen in such a way that $\mu(S_i=+1 \mid S_{\Lambda(i)})+\mu(S_i=-1 \mid S_{\Lambda(i)})=1$. Stated in words, each point is in thermal equilibrium with its surroundings. This is just what one expects from an equilibrium state.

To derive (2.9) and (2.10) we imagine Λ' to be contained

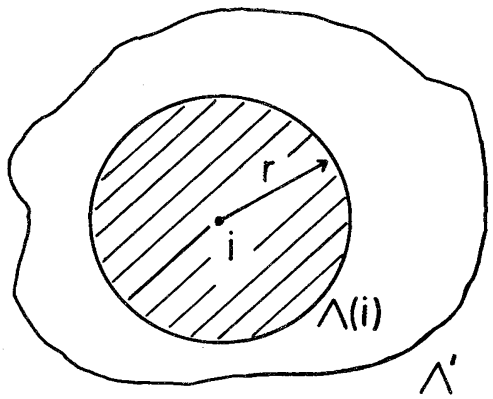

FIG. 1. Spin at i interacts only with spin inside the sphere $\Lambda(i)$.

in a large volume, with $\mu_{\beta,N}$ determined by the Boltzmann factor $\exp(-\beta H_N)$. We then apply (2.8) and note that, by the finite range of the interaction, the conditional probability $\mu_{\beta,N}(S_i \mid S_{\Lambda'\setminus\{i\}})$ reduces to $\mu_{\beta,N}(S_i \mid S_{\Lambda(i)})$. Having fixed $S_{\Lambda(i)}$ we see that the spins S_j with j in $\Lambda(i)$ produce an external field $\sum_{j\in\Lambda(i)} J_{ij}S_j$ at the site i and, hence, the Boltzmann factor (2.10). We now send $N\to\infty$ and obtain (2.9). One can show[16] that, given (2.9) and (2.10) for each i, the same holds true with i replaced by an arbitrary finite region. The above arguments may be extended straightforwardly to so-called short-range potentials which satisfy condition (3.8) below.

Alternatively one could argue that an equilibrium state should be obtained as a limit of finite-volume Gibbs states,

$$\langle A\rangle = \lim_{N\to\infty} \mathrm{Tr}[A\exp(-\beta H_N)]/\mathrm{Tr}\exp(-\beta H_N)\ . \tag{2.11}$$

Then $\langle A\rangle = \int d\mu(\omega)A(\omega)$ and μ satisfies the DLR equations. For ferromagnetic interactions the existence of the limit in (2.11) is guaranteed by correlation inequalities, but this need not be the case in spin-glasses below T_f. Instead of $\{\mu_{\beta,N}\}$ one could pick a (weakly) converging subsequence $\{\mu_{\beta,N}\}$ (this is possible by a compactness argument), but a better way out is presented by the DLR equations themselves, which do not involve a limit any more. Some properties of solutions to the DLR equations are listed below.

Any solution to the DLR equations is called a Gibbs state (Gibbs measure). Gibbs states have some remarkable properties. First, for a *finite* system there is a unique solution to the DLR equations: the Gibbs state $\mu_{\beta,N}$ we started with. This is another way of saying that a finite system does not exhibit a phase transition. Second, any convex combination of Gibbs states gives another solution to the DLR equations, as is most easily seen by noticing that the right-hand side of (2.10) does not depend on μ. Conversely, every Gibbs measure μ may be decomposed uniquely[15,17] into extremal or *ergodic* components. This is the ergodic decomposition of μ with respect to time translation. It is basic to all that follows. Finally, ergodic components are space-*clustering*, i.e., if $\mu^{(\lambda)}$ is ergodic, then

$$\lim_{|x|\to\infty} |\mu^{(\lambda)}(A\tau_x B) - \mu^{(\lambda)}(A)\mu^{(\lambda)}(\tau_x B)| = 0\ , \tag{2.12}$$

where τ_x is a space translation by x.

D. Phase transitions and order parameters

If the temperature T is high enough, there is a *unique* solution μ_β to the DLR equations (Dobrushin's uniqueness theorem). By the uniqueness μ_B is ergodic and the limit of finite-volume Gibbs states. We say the system exhibits a phase transition at T_c if below T_c we have *several* solutions to the DLR equations. So the following picture emerges. For $T > T_c$ there is a unique, ergodic equilibrium state. Below T_c an equilibrium state as produced by (2.11) with free boundary conditions is never ergodic, but it may be decomposed into ergodic components, which represent the *pure* thermodynamic phases. That is, below T_c *the ergodicity is broken.*[18]

The free energy per spin $f(\beta)$ is given by

$$-\beta f(\beta) = \lim_{N\to\infty} \frac{1}{N} \ln \mathrm{Tr}\exp(-\beta H_N)\ . \tag{2.13}$$

The limit exists and does not depend on the boundary condition. In the case of short-range random systems, $f(\beta)$ does not depend on the random configuration of the J_{ij} either.[19] As a first and surprising consequence of this result we note that if a short-range random system has a phase transition, all samples have the same T_c.

We can also assign a free energy to each of the ergodic components of a Gibbs state μ_β. Thermal equilibrium is characterized by a (local) variational principle[20]: The free energy has to be minimal. Accordingly, the free energy of the ergodic components should not exceed $f(\beta)$ as given by (2.13). If, on the other hand, their free energy *agrees with* $f(\beta)$, the interpretation of the ergodic components as pure thermodynamic phases makes sense and we get a conventional phase transition, i.e., below T_c the system "picks" one of the ergodic components. We mean the following. Though probabilistically μ_β is a convex combination of ergodic components $\mu_\beta^{(\lambda)}$, each with its own weight (probability) α_λ, the system certainly picks *one* component λ.[21] We just do not know which one. The probabilities α_λ are consistent with our information (ignorance) about the system: It is in thermal equilibrium. In the Appendix we show that all the ergodic components have the same free energy, which agrees with $f(\beta)$ as given by (2.13). Thus also for random systems the notion of conventional phase transition is firmly rooted in the formalism of equilibrium statistical mechanics.

Suppose now that we have determined all the ergodic components $\mu_\beta^{(\lambda)}$. How do we discriminate between them? The answer is simple: by finding an order parameter. More specifically, an *order parameter* is an observable A, or a group of observables A_i, $1\le i\le n$, such that the numbers $\mu_\beta^{(\lambda)}(A)$ or $\mu_\beta^{(\lambda)}(A_i)$, $1\le i\le n$, determine λ and, hence, $\mu_\beta^{(\lambda)}$ *uniquely.* At the moment we do not specify n. In Sec. III we will see that the formerly somewhat mysterious notion of an order parameter has been turned into a sensible, well-defined concept.

III. BROKEN ERGODICITY

The purpose of this section is to analyze the phenomenon of broken ergodicity. This will be accomplished by studying several examples, including the two-dimensional Ising model and the three-dimensional Heisenberg model (both translationally invariant) as well as disordered spin systems. We will also discuss the Edwards-Anderson order parameter, the breakdown of linear response, and the possible number of ergodic components.

A. The two-dimensional Ising model

The two-dimensional Ising model with translationally invariant and ferromagnetic nearest-neighbor interactions is defined by (2.7) with $J_{ij}=J>0$ if i and j are neighbors,

and $J_{ij}=0$ otherwise. Onsager showed in his classic paper[22] that the model has a phase transition at a positive temperature T_c. Below T_c the model has two equilibrium states, a $(+)$ state with positive magnetization and a $(-)$ state with negative magnetization. More precisely, if we take $(+)$ boundary conditions and let $N\to\infty$, we find the state $\mu_\beta^{(+)}$, with $(-)$ boundary conditions we arrive at the state $\mu_\beta^{(-)}$. If $T>T_c$, then $\mu_\beta^{(+)}=\mu_\beta^{(-)}$ and the spontaneous magnetization m vanishes, whereas for $T<T_c$ we have $\mu_\beta^{(+)}\neq\mu_\beta^{(-)}$ with $m\neq 0$, and thus at least two different solutions to the DLR equations. Both $\mu_\beta^{(+)}$ and $\mu_\beta^{(-)}$ are ergodic, and the model has no other ergodic equilibrium states.[23]

The natural symmetries of the model are translational invariance, and spin-flip symmetry $S_i\to -S_i$ for all i. By flipping the spins we map $\mu_\beta^{(+)}$ onto $\mu_\beta^{(-)}$, and conversely. If, however, μ_β is the state we get by taking the limit (2.11) with, either free or periodic boundary conditions, then μ_β is invariant under the spin-flip symmetry operation. In fact, for these states

$$\mu_\beta=\tfrac{1}{2}\mu_\beta^{(+)}+\tfrac{1}{2}\mu_\beta^{(-)}\,. \tag{3.1}$$

Below T_c the ergodic components $\mu_\beta^{(+)}$ and $\mu_\beta^{(-)}$ are not invariant under the full symmetry group of the Hamiltonian and the ergodicity of μ_β has been broken. Both the $(+)$ state and the $(-)$ state have the same energy, entropy, and thus the same free energy. Hence the phase transition is conventional (Sec. III D). A nice illustration of (3.1) has been given by Aizenman.[23]

An order parameter is easily found. We take S_0, the spin at 0. If $\mu_\beta^{(\lambda)}(S_0)>0$ we have $\lambda=+$, and if $\mu_\beta^{(\lambda)}(S_0)<0$ we deduce $\lambda=-$. So the sign of $\mu_\beta^{(\lambda)}(S_0)$ already determines λ completely.

B. The three-dimensional Heisenberg model

The Hamiltonian of the classical three-dimensional isotropic Heisenberg model with nearest-neighbor interactions is given by

$$H_N=-J\sum_{\langle i,j\rangle}\vec{S}_i\cdot\vec{S}_j\,, \tag{3.2}$$

where the sum is over nearest neighbors only. The natural symmetries of the model are given by translational invariance and the symmetry group SO(3): $\vec{S}_1\to g\vec{S}_i$ with g in SO(3). As in the previous model the interaction is ferromagnetic and at $T=0$ all spins are parallel. In such a ground state the remaining symmetry group (the isotropy subgroup) is SO(2), so the coset space SO(3)/SO(2) labels the ground states.[24] This coset space is homeomorphic to S_2, the surface of the three-dimensional unit sphere.

Indeed, the model has a phase transition at a positive temperature T_c,[25] and below T_c we can label the uncountably many ergodic states by the elements of SO(3)/SO(2) or, equivalently, by S_2. Above T_c we have a unique and, thus, ergodic equilibrium state, while the ergodicity is broken below T_c.

What is the gist of what we have done? If the Hamiltonian has a symmetry group G and an ergodic state μ has a symmetry group H, which is a subgroup of G (the isotropy subgroup), then we can produce as many ergodic states as there are elements in the coset space G/H. To see this we first note that if μ is ergodic and g is in G, then $\mu'=\mu\circ g$ is also ergodic. [μ' is defined by $\mu'(A)=\mu(gA)$ for any observable A.] Suppose, namely, that μ' were not ergodic, then $\mu'=\alpha_1\mu_1+\alpha_2\mu_2$ with $\alpha_i\geq 0$ and $\alpha_1+\alpha_2=1$. Hence $\mu=\alpha_1\ (\mu_1\circ g^{-1})+\alpha_2(\mu_2\circ g^{-1})$, which is a contradiction unless $\mu_1=\mu_2=\mu'$. Here we used the fact that also $\mu_i\circ g^{-1}$ is an equilibrium state; simply notice (2.9) and (2.10). Factoring out the symmetry group H of μ itself we end up with G/H, as was to be shown. We now return to the Heisenberg model.

If μ_β is the state we get by taking the limit (2.11) with either periodic or free boundary conditions, then μ_β is invariant under whatever element from SO(3), and its ergodic decomposition is given by

$$\mu_\beta=\int_{S_2}dP(\lambda)\mu_\beta^{(\lambda)}\,, \tag{3.3}$$

where $dP(\lambda)$ is the normalized uniform distribution on the surface of the three-dimensional unit sphere, S_2. We note again that μ_β is invariant under the full symmetry group SO(3) whereas the ergodic, components $\mu_\beta^{(\lambda)}$ are not. We must take, so to speak, a convex combination of ergodic states to restore the full symmetry of the Hamiltonian.

Since the ergodic components are transformed into each other by the elements of the symmetry group, they all have the same energy, the same entropy, and the same free energy. Hence we have a conventional phase transition at T_c (Sec. II D).

Here also the order parameters are easily found. We take x, y, and z components of the vector S_0. Below T_c the vector $\mu_\beta^{(\lambda)}(S_0)$ uniquely determines a direction and, hence, a point on the surface of the unit sphere in $\mathbb{R}^3$, i.e., λ itself.

In the above example the elements of a symmetry group *exhaustively* label the ergodic components $\mu_\beta^{(\lambda)}$ of the canonical Gibbs measure μ_β; cf. Eq. (3.3). More precisely, by factoring out the isotropy subgroup H of the full symmetry group G so as to form the coset space G/H we find all the ergodic components. This phenomenon is called broken symmetry. One, therefore, might wonder[5] whether this is characteristic for a translationally invariant, nonrandom system. It is not. There are examples[26] where the elements of the symmetry group do *not* exhaustively label the ergodic components. So we must conclude that, in general, broken ergodicity and not broken symmetry is the important phenomenon in translationally invariant systems also.

C. Quantum system

Thermal equilibrium states of *quantum* spin systems may also be decomposed into ergodic components. We will make a small detour to sketch the necessary modifications. Full details may be found in Refs. 15 and 27.

Quantum spin systems have the pleasant property that they possess a natural dynamics which, for a finite system,

is given by ($\hbar=1$)

$$\alpha_t^N(A)=e^{iH_Nt}Ae^{-iH_Nt} . \tag{3.4}$$

We now take a local observable, for example, the product of finitely many spin operators, and send $N\to\infty$ so as to get the dynamics α_t of the infinite system.

For one reason or another quantum equilibrium states are denoted by ω or ρ. Specifically, ρ is said to be an equilibrium state at inverse temperature β if ρ satisfies the Kubo-Martin-Schwinger (KMS) condition,

$$\rho((\alpha_tA)B)=\rho(B(\alpha_{t+i\beta}A)) , \tag{3.5}$$

for all local observables A and B. Note that α_t is the dynamics of the infinite system. This condition implies, among other things, that ρ is stationary in time. In the case of a finite system with Hamiltonian H_N, it *uniquely* specifies the canonical Gibbs state $\rho_{\beta,N}$ given by

$$\rho_{\beta,N}(A)=\mathrm{Tr}[A\exp(-\beta H_N)]/\mathrm{Tr}\exp(-\beta H_N) . \tag{3.6}$$

In the case of an infinite system, the KMS condition is satisfied by the infinite volume limit of $\rho_{\beta,N}$. With the use of (3.6) it is straightforward to verify (3.5).

A KMS state ρ may be decomposed uniquely into a convex combination of *extremal* KMS states $\rho^{(\lambda)}$, which are ergodic with respect to the dynamics α_t, i.e., for each pair of observables A and B, we have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T dt\,\rho^{(\lambda)}((\alpha_tA)B)=\rho^{(\lambda)}(A)\rho^{(\lambda)}(B) . \tag{3.7}$$

In addition, relation (2.12) holds with $\mu^{(\lambda)}$ replaced by $\rho^{(\lambda)}$.

Summarizing, here also the ergodic decomposition is well defined, and the notion of phase transition and order parameter may be defined as in Sec. II D. Nevertheless, we will assume the spins to be classical throughout what follows (except in the Appendix).

D. Spin-glasses

A spin-glass is a disordered magnetic system where the bonds between the spins are either ferromagnetic or antiferromagnetic, and cannot be satisfied simultaneously (frustration). One may assume the spins are on a regular lattice and take the interaction as random.[2] According to Edwards and Anderson a field-cooled spin-glass is a new thermodynamic phase with a well-defined phase transition at a critical temperature $T_c=T_f$. We now probe some of the consequences of this suggestion within the context of equilibrium statistical mechanics. We study the notion of T_f, treat the breakdown of linear response, discuss the Edwards-Anderson order parameter, and analyze the assumption[5] that the number of components grows with N as $\exp(aN)$.

A phase transition shows up as a nonanalyticity in β of the free energy $f(\beta)$. We have seen in Sec. II D that, whatever the random configuration of the J_{ij}, we always get the same free energy and, hence, the same T_c. The finite range of the potential is not really essential in the sense that we only have to require

$$||J||=\sup_i\left\{\sum_j|J_{ij}|\right\}<\infty , \tag{3.8}$$

i.e., the tails of the interaction have to be uniformly small. This condition and a standard argument[28] suffice. If the electron mean-free path is finite, the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction is exponentially damped[29] and, hence, satisfies (3.8). So all the previous arguments also apply to metallic spin-glasses. Moreover, the free energy of the ergodic components agrees with $f(\beta)$ as given by Eq. (2.13); cf. the Appendix. So a conventional phase transition is possible and, up to now, the equilibrium picture is fully consistent.

Experimentally it is well-known that T_f and the free energy are *reproducible* quantities, which do not depend on the microscopic random configuration of the J_{ij}: Two alloys whose microscopic structures differ but whose macroscopic constitution and preparation are identical (same concentration) give the same experimental outcomes. One can take a *specific* sample and need not average over an ensemble of them. The above equilibrium formulation of the free energy and the phase transition at T_f satisfy the reproducibility criterion. Any theory of spin-glasses which finds no reproducibility of thermodynamic quantities is, in our opinion, particularly suspect.

Suppose now that the J_{ij} are independent random variables whose distribution only depends on $(i-j)$. For a finite system of Ising spins the zero-field susceptibility $\chi_0(T)$ is given by

$$\chi_{0,N}(T)=\frac{\beta}{N}\sum_{i,j}[\mu_{\beta,N}(S_iS_j)-\mu_{\beta,N}(S_i)\mu_{\beta,N}(S_j)] . \tag{3.9}$$

As to $\mu_{\beta,N}$ we take free boundary conditions. It is tempting to rewrite (3.9) in the form

$$\chi_{0,N}(T)=\frac{\beta}{N}\sum_i\left[\sum_j\mu_{\beta,N}(S_iS_j)-\mu_{\beta,N}(S_i)\mu_{\beta,N}(S_j)\right] \tag{3.10}$$

and to take the limit $N\to\infty$ so as to conclude that *with probability 1*

$$\chi_0(T)=\beta\sum_j\left[\langle\mu_\beta(S_0S_j)\rangle-\langle\mu_\beta(S_0)\mu_\beta(S_j)\rangle\right] \tag{3.11}$$

by the ergodic theorem. The angular brackets denote an average over the randomness. Equation (3.11) is a so-called fluctuation-dissipation relation. The simplicity of its "derivation" is a bit specious, however. One has to spell out the arguments more carefully.[30,31] The case $h\neq0$ only requires a small modification and need not be given here.

Let first $T>T_c=T_f$. Then the equilibrium state μ_β exists as the thermodynamic limit ($N\to\infty$) of $\mu_{\beta,N}$. (Note that *below* T_c the existence of the limit is not guaranteed.) By spin-flip symmetry $\mu_{\beta,N}(S_i)=\mu_\beta(S_i)=0$. If the temperature is high enough, $\mu_\beta(S_iS_j)$ decreases exponentially fast as $|i-j|\to\infty$, and we may rewrite the sum (3.9) in

the form (3.10), knowing for sure that, if $T > T_c$ and given i, the sum over j will converge as $N \to \infty$. Hence (3.9) holds with probability 1 and

$$\chi_0(T) = \beta \sum_j \langle \mu_\beta(S_0 S_j) \rangle \ . \tag{3.12}$$

Usually not that much of the right-hand side of (3.12) remains, since in many spin-glass models one takes an *even* distribution for J_{ij}, which allows a local gauge transformation $S_0 \to -S_0$, $J_{0j} \to -J_{0j}$, within each average. Thus all averaged two-point correlation functions are zero[32] and $\chi_0(T) = \beta$ because $j = 0$ is the only term in (3.12) which survives; the susceptibility has a *pure Curie behavior.*

Three aspects of this result are worth noticing. First, it is exact and holds with probability 1, i.e., it is reproducible and does not depend on the specific sample. Second, the result also holds for classical Heisenberg spins provided the spin-spin interaction is isotropic. Third, the assumption that the interaction is of finite range and the distribution of the J_{ij} is even *must* lead to a pure Curie behavior above T_c.

Numerically exact results[33] indicate that $T_c = 0$ for nearest-neighbor interactions in two dimensions. The reason behind this behavior is the occurrence of a network of zero-energy loops[34] (which require an explanation themselves). Let us assume that such a network of zero-energy loops could be generalized to a network of zero-energy surfaces in three dimensions. Then equilibrium statistical mechanics could not account for a positive T_f in spin-glasses with allegedly short-range interactions as in $Eu_xSr_{1-x}S$ unless long-range dipole-dipole interactions, which are not absolutely summable [cf. Eq. (3.8)], play an important role. The assumption that zero-energy loops are capable of generalization to three dimensions might be wrong, however.[35]

What happens to (3.11) below T_c? It is generally believed that (3.11) does not hold below T_c ("breakdown of linear response"). This belief is certainly correct. As it stands Eq. (3.11) is not even true for the translationally invariant, ferromagnetic two-dimensional Ising model (Sec. III A). Before turning to the random case we will study this model in some detail.

With free or periodic boundary conditions.

$$\mu_\beta(S_0 S_j) = \tfrac{1}{2}\mu_\beta^{(+)}(S_0 S_j) + \tfrac{1}{2}\mu_\beta^{(-)}(S_0 S_j) \to m^2 \neq 0 \tag{3.13}$$

as $|j| \to \infty$, where m is the spontaneous magnetization.[36] Accordingly, the series in (3.12) diverges. Moreover, we know (Sec. II D) that below T_c the thermal equilibrium behavior of the system is described by an *ergodic* component. Let us, therefore, return to Eq. (3.10) and take either $(+)$ or $(-)$ boundary conditions. We then find, as $N \to \infty$,

$$\chi_0(T) = \sum_j \left[\mu_\beta^{(\lambda)}(S_0 S_j) - \mu_\beta^{(\lambda)}(S_0)\mu_\beta^{(\lambda)}(S_j) \right] , \tag{3.14}$$

where λ denotes one of the ergodic components $(+)$ or $(-)$. The convergence of the infinite series (3.14) is well known[30] and, in fact, consistent with (2.12). So, to make sense out of the linear-response formula (3.11), we must at least take an ergodic component. Note, however, that we were fairly liberal in interchanging the limits $h \to 0$ and $N \to \infty$. A more careful treatment could be based on, e.g., Ref. 31.

However, what should we do in a disordered system where we do not have translational invariance? As in the above examples we can take suitable (generalized) boundary conditions[37] so as to single out an ergodic component λ as $N \to \infty$. Then for large N the following formula correctly represents $\chi_0(T)$ of the ergodic component λ:

$$\chi_0(T) = \frac{1}{N} \sum_{i,j} [\mu_\beta^{(\lambda)}(S_i S_j) - \mu_\beta^{(\lambda)}(S_i)\mu_\beta^{(\lambda)}(S_j)] \ . \tag{3.15}$$

In passing we note that Eq. (3.15), *mutatis mutandis,* also holds for $h \neq 0$. We now rewrite this equation in a form which reminds us of the central limit theorem,

$$\chi_0(T) = \mu_\beta^{(\lambda)} \left[\left[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} S_i - \mu_\beta^{(\lambda)}(S_i) \right]^2 \right] . \tag{3.16}$$

For ferromagnetic systems it is known[38] that the finiteness of the susceptibility *as* $N \to \infty$ implies and is implied by a central limit theorem for the magnetization. In this case the susceptibility is finite if $T \neq T_c$. We speculate that the theorem also holds for disordered systems and that the $\chi_0(T)$ thus obtained does not depend on the specific component λ. The second half of our speculation can be substantiated considerably. We know (cf. the Appendix) that all ergodic components have the same free energy, which agrees with the usual $f(\beta,h)$ given by (2.13); here we have also included the dependence upon h.

The function $-\beta f(\beta,h)$ is convex in β and h. For given β, $-\beta f(\beta,h)$ is differentiable almost everywhere in h and its derivative (proportional to the magnetization) is monotonic.[39] Monotonic functions are themselves differentiable almost everywhere. Thus $f(\beta,h)$ is almost everywhere in h twice differentiable and its second derivative equals $\chi(T,h)$. Nondifferentiability is expected only there where we have a phase transition. Since the χ's for different components are derived from the same f, they have to agree (almost everywhere)—as they should, since they are reproducible.

Except for the spin-flip symmetry a disordered Ising spin system does not have evident symmetries. Nevertheless, the notion of an order parameter is well defined and should satisfy the following criteria (Sec. II D): (a) for a *fixed* random configuration of the J_{ij} it should single out (b) *precisely one* ergodic component. The Edwards-Anderson order parameter,

$$q_{EA} = \langle [\mu_\beta(S_i)]^2 \rangle \ , \tag{3.17}$$

does not satisfy either of the above criteria. The inner average $\mu_\beta(S_i)$ is the canonical Gibbs average (hence zero) and the outer average is taken with respect to the randomness, whereas we consider one fixed sample. There is a very important fundamental truth about random systems we must always keep in mind[40]: "No real atom is an average atom, nor is an experiment ever done on an ensemble

of samples." Another candidate,

$$q'_{\mathrm{EA}}=\langle\mu_\beta(S_iS_j)\rangle,\quad i\neq j \tag{3.18}$$

is hardly an improvement. Both do not single out an ergodic component. Even worse, taken literally, both are identically zero. In our opinion, any averaging over the J_{ij} is to be avoided.

Let us pause briefly to see what we have gained. First, and most importantly, there is conservation of ergodicity: Above T_c we have a unique, canonical equilibrium state μ_β while below T_c, though the ergodicity of μ_β is broken and there are several ergodic components $\mu_\beta^{(\lambda)}$, the system still picks one specific $\mu_\beta^{(\lambda)}$, which is *ergodic*. The ergodic components all have the same free energy $f(\beta,h)$. Since singularities of $f(\beta,h)$ in β and h occur at the same β and h for all components, there is no "bifurcation" of different components at different temperatures. All this needs no modification in quantum mechanics. In addition, the following consistent picture arises.

(a) In each component the thermodynamic relations, like $df=-s\,dT-m\,dh$, are valid—as follows by formal differentiation of f with respect to β and h.

(b) The linear-response formula (3.15) applies ($N\to\infty$).

(c) The Maxwell relations hold.[41]

The thermodynamic (equilibrium) quantities derived from (a)—(c) do not vary as we go from one ergodic component to another, nor do they depend on the specific random configuration. Hence they are reproducible—as they should be[42]—provided, of course, there is no phase transition.

The main problem one has to face when describing the spin-glass phase within the context of equilibrium statistical mechanics is to characterize and label the ergodic components. Once (and not before) one knows, or can guess, the ergodic decomposition one can choose the order parameters. Because of the complicated random structure of spin-glasses this seems a formidable task.

Though we cannot label the ergodic components yet ($T<T_f$) we can try to estimate their number. Estimating this number is a sensible problem because the ergodic components are disjoint (Sec. II A). According to a recent suggestion by Palmer[5] there are $\exp(aN)$ disjoint ergodic components as $N\to\infty$. Nevertheless the spin-glass phase transition is to be conventional in the sense that, at sufficiently low temperature, the system is caught in a specific component. We will show that *in the context of equilibrium statistical mechanics* this is not possible. To facilitate the discussion we take N finite but very large and estimate the free energy $f(\beta)$ of the canonical Gibbs state μ_β; cf. Eqs. (2.11) and (2.13).

According to the above suggestion the phase space Ω can be decomposed into $M=\exp(aN)$ disjoint components Ω_λ, and

$$\mu_\beta=\sum_{\lambda=1}^{M}\alpha_\lambda\mu_\beta^{(\lambda)},\quad \mu_\beta^{(\lambda)}(\Omega_\nu)=\delta_{\lambda\nu}\,. \tag{3.19}$$

Then the entropy of μ_β is given by

$$S_N(\mu_\beta)=\sum_{\lambda=1}^{M}\alpha_\lambda S_N(\mu_\beta^{(\lambda)})-\sum_{\lambda=1}^{M}\alpha_\lambda\ln\alpha_\lambda\,, \tag{3.20}$$

where $S_N(\mu_\beta^{(\lambda)})$ is the entropy of the λth component. To estimate the second term in (3.20), which is called the complexity,[5]

$$I_N=-\sum_{\lambda=1}^{M}\alpha_\lambda\ln\alpha_\lambda\,, \tag{3.21}$$

we note that $\alpha_\lambda\simeq\exp(-aN)$ with $a>0$, and find $I_N\simeq Na$. Hence the entropy per spin may be written

$$s(\mu_\beta)=\sum_{\lambda=1}^{M}\alpha_\lambda s(\mu_\beta^{(\lambda)})+a\,. \tag{3.22}$$

The energy per spin is easily found,

$$u(\mu_\beta)=\frac{1}{N}\mu_\beta(H_N)=\sum_{\lambda=1}^{M}\alpha_\lambda u(\mu_\beta^{(\lambda)})\,, \tag{3.23}$$

and thus, since all the ergodic components have the same free energy and $f(\beta)=u(\mu_\beta)-Ts(\mu_\beta)$,

$$f(\beta)=\sum_{\lambda=1}^{M}\alpha_\lambda f(\beta)-aT=f(\beta)-aT\Longrightarrow a=0\,, \tag{3.24}$$

which contradicts our assumption that $a>0$. If the disjointness of the components were only approximately true, the same conclusion would hold.[43] In fact, we have shown that $N^{-1}I_N=o(N)$: The canonical free energy cannot be lower than the free energy of each of its ergodic components. Phrased differently, if the complexity would have a physical meaning, this result would exclude a conventional phase transition.

Finally, we turn to a closely related sampling problem which occurs when quantities are not reproducible. Below T_c the system is in a specific ergodic component λ. Measurement of a physical quantity A gives rise to the expectation value $A^{(\lambda)}$. As λ varies, $A^{(\lambda)}$ may also vary. We only know that this is not the case for reproducible quantities like the susceptibility or the specific heat. Whatever A, one may define[4,5]

$$\bar{A}=\sum_\lambda\alpha_\lambda A^{(\lambda)} \tag{3.25}$$

when

$$\mu_\beta=\sum_\lambda\alpha_\lambda\mu_\beta^{(\lambda)}\,. \tag{3.26}$$

If the variance $v(A)=(\overline{A^2}-\bar{A}^2)>0$, one has to face the problem of how to interpret $\bar{A}$. The point is that $\bar{A}$ cannot have a direct physical significance: How shall we sample it in view of many components? The few values we get from repeating an experiment *certainly* do not suffice in the case of $\exp(aN)$ components. Accordingly it is hard to interpret the complexity[4,5] as a physical entropy simply because parts of the phase space which the system does not visit cannot contribute to the entropy.

IV. NONEQUILIBRIUM AND SPIN DYNAMICS

In this section spin freezing and other spin-glass properties are treated as dynamical phenomena of a nonergodic nature which are characterized by metastability and decay with "infinite" relaxation times as $t \to \infty$. We estimate the number of potential valleys, study the sampling and entropy problems, which are inherent to this description, and suggest two new experiments for checking the internal consistency of the nonequilibrium approach. We first try to grasp what the classical energy landscape looks like.

One imagines[5,6,44–46] that the phase space of a spin-glass contains many potential valleys separated by large but *finite* energy barriers; the depths of the valleys and the heights of the barriers are randomly distributed. There may also be some passes connecting two valleys. Because the passes are narrow, they do not change the picture essentially. The barriers must be finite since otherwise we would have some kind of equilibrium phase transition. Inside the valleys one finds only very small hills separated by wide plains, each of which is to be associated with a large ground-state degeneracy and, hence, nonzero entropy. When a spin-glass is cooled through T_f it is frozen into a particular valley. It does not sample microscopic states in other valleys that have the same energy but are inaccessible because of the intervening barriers. Cooling down further, and doing it carefully, one finally reaches a ground state, which may be used to label the valley.

One of the main questions which now arise is how to generate so many valleys that the barrier heights between them remain finite. In the case of the two-dimensional *Ising* model an answer has been suggested by Morgenstern and Horner.[45] According to their suggestion there exists a network of zero-energy loops. Inside a (nontrivial) loop we have about 150 spins.[47] We can flip them all at the same time for no cost in energy, by the very definition of zero-energy loop. However, since an Ising model has no natural dynamics we imagine it to be endowed with a *one-spin* dynamics of the Glauber type (Sec. II B) so that we have to flip the spins one after the other. It, therefore, takes quite a long time to travel from one valley to a neighboring one over a high energy barrier. If the temperature is low, the system has no chance to finish the trip within the finite amount of time available for a Monte Carlo experiment: It has been frozen into a certain valley. Note that the hills and the valleys *are determined by the dynamical process one assumes.* Moreover, it remains to be shown that the Monte Carlo dynamics has a direct bearing on the system's real dynamics, it being understood that the system is in contact with a heat bath.

The number of nontrivial zero-energy loops is proportional to N, the number of spins, and flipping all the spins inside one loop gives another valley, so that the total number of valleys is proportional to 2^{bN} for a certain $b > 0$, i.e., $\exp(aN)$ with $a > 0$. The picture does not change essentially when we add more spins: We simply get some more zero-energy loops. This explains why thermodynamic quantities are extensive. However, they are also reproducible and precisely here we get into serious problems. Before proceeding we remind the reader that reproducible means that two samples (say, alloys) whose microscopic structures differ but whose macroscopic constitution and preparation are identical (same concentration) give the same experimental outcomes. Examples include thermodynamic observables such as the freezing temperature T_f, the susceptibility χ, and (the magnetic part of) the specific heat.

As we already noted in Sec. III D, there is the sampling problem itself. It is, for instance, reasonable[6] to assume that cooling in different fixed magnetic fields leaves a sample trapped in different regions of phase space. Because the number of potential valleys is extremely high we will visit another valley each time we repeat an experiment. Since it is not to be expected that all valleys are equal, the ensuing experimental results are bound to be *irreproducible,* that is, they show a statistical scatter, hence a nonergodic behavior, and hence a contradiction, thermodynamic observables being reproducible.[41]

In view of the well-defined and reproducible freezing temperature T_f (Refs. 3, 48–50) the difficulties associated with the sampling problem become even more pronounced. The point is this. In a real glass[4] we do not have a sharp transition from the undercooled melt into the glassy phase but instead we find a transition *regime* Φ_g which indicates that the barrier height distribution has a finite width. On the other hand, in the spin-glass case the transition is fairly sharp and thus the width of the barrier height distribution has to vanish, at least as $N \to \infty$, which is somewhat surprising. Even if it did, it is hard to see how the transition could be sharp in a *canonical* description, the system being in contact with a heat bath. In this case, even if we take the large-volume limit, the canonical and the microcanonical descriptions are not equivalent.

Suppose there were a uniform barrier height and a sharp transition at T_f (by mechanisms which we do not understand yet). If $T > T_f$ the system may wander through the whole phase space. T_f is characterized by the system being caught in a specific valley, which it *de facto* will not leave for $T < T_f$. Then the system has to lose nearly all its entropy at T_f and, hence, a considerable amount of latent heat, which is amenable to experimental verification. However, as far as we know, the entropy is smooth.

Finally we note another interesting phenomenon which is also a consequence of the occurrence of zero-energy surfaces and also amenable to experimental verification: *spurious magnetization.* Each domain of size M, which is surrounded by a zero-energy surface, has at low T, a net magnetic moment proportional to $\sqrt{M}$ and, thus, a magnetization per spin $\sim 1/\sqrt{M}$. All the domains have about the same size.[47] If one cools the system in a very weak external field, one should be able to direct the magnetic moment of each *domain* parallel to the external field and, in so doing, generate a net or spurious magnetization, whose order of magnitude is $M^{-1/2}$ times the saturation magnetization. This effect is clearly borne out by Fig. 1 of Ref. 46 (dotted lines). However, evaluating carefully the experimental results of Ref. 51 we cannot but con-

clude that there is *no* spurious magnetization. It would be interesting to check whether this is also true in $Eu_xSr_{1-x}S$ or another spin-glass with nearest-neighbor-type interactions.

Though the dynamical picture may be attractive from a Monte Carlo point of view and quite suggestive for explaining metastability and irreversibility, some fundamental questions about its consistency remain open: The absence of a statistical scatter of the barrier heights and the sharpness of the spin-glass transition at T_f are still unexplained, the predicted nearly complete loss of entropy at T_f has not been observed yet, and the nonexistence of spurious magnetization in archetypical spin-glasses like *Au*Fe and *Cu*Mn seems firmly established. One also has to bear in mind that the interaction between the spins is usually long range[52] instead of nearest-neighbor, which seems unfavorable for the occurrence of zero-energy surfaces, and that the very existence of zero-energy surfaces in *three* dimensions still has to be shown.[35] Finally, the nonergodic point of view contradicts the validity of the Maxwell relations.[41,6]

V. CONCLUSION

In this paper we have treated the bare essentials of the spin-glass problem with a special emphasis on certain statistical-mechanical aspects. For the sake of definiteness we have neglected some other approaches like Néel's superparamagnetism[53,54] and percolation.[42,52] Most of the time the interactions were taken to be short-range so as to avoid technical complications and present the arguments as simply as possible. With these restrictions in mind we now turn to the results.

We have unambiguously shown (Sec. III) that it is possible to give a consistent description of the spin-glass transition and the spin-glass phase within the context of *equilibrium* statistical mechanics, if the spin-glass is field-cooled.

Above T_f we have a unique equilibrium state μ_β, which is the canonical Gibbs state for the infinite system ($N\to\infty$). The state μ_β has a free energy $f(\beta,h)$ which is obtained through the thermodynamic limit—as usual—with β as the inverse temperature and h as the external magnetic field; $f(\beta,h)$ does *not* depend on the randomness. Morever, μ_β is ergodic with respect to the system's dynamics.

At T_f we find a conventional phase transition. This means that below T_f the ergodicity of μ_β has been broken and μ_β may be decomposed into *several* ergodic components $\mu_\beta^{(\lambda)}$. The system picks one of the components as its equilibrium state. So, in a sense, we have "conservation of ergodicity." The ergodic components all have the same free energy $f(\beta,h)$, which agrees with the one obtained via the thermodynamic limit, i.e., the free energy of the canonical Gibbs state μ_β itself (see the Appendix). Our prescription is *un*changed in quantum mechanics. Thermodynamic quantities which are obtained by formal differentiation of $f(\beta,h)$ with respect to β and h neither depend on the randomness nor on the specific component. In particular, the Maxwell relations are always obeyed. This agrees with experiment. The main problem is to find and characterize the ergodic components. The choice of the order parameters, which uniquely determine the components, depends crucially on our knowledge of the decomposition.

Phase transitions are characterized by a singularity of $f(\beta,h)$ in either β or h. Thus, by its very definition, T_f does not depend on the randomness. If one ergodic component has a phase transition, so do all the other components. This does not mean, however, that they bifurcate. In a typical bifurcation the outer prongs of the fork branch off and are stable whereas the original solution remains "in the middle" and becomes unstable. It is tempting to imagine that below T_f the original μ_β may be *continued* in some sense so as to be a metastable (or unstable) state. This is not the case.[55,56] In fact, though there has been some progress,[56,57] a satisfying theory of metastability still has to be found. Nevertheless it is possible to give a consistent statistical-mechanical description of the equilibrium properties of a spin-glass.

ACKNOWLEDGMENTS

We thank G. Toulouse for useful correspondence and gratefully acknowledge stimulating discussions with J. A. Mydosh, I. Morgenstern, and R. G. Palmer. We also thank Dr. Palmer for sending us his recent work (Ref. 5) prior to publication. This work has been supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 123).

APPENDIX

Let ρ be a state which satisfies either the DLR equations (classical) or the KMS conditions (quantum mechanical). We show that the free energy per site of ρ is given by

$$-\beta f(\beta)=\lim_{\Lambda\to\infty}\frac{1}{|\Lambda|}\ln \mathrm{Tr}\exp(-\beta H_\Lambda)\,, \tag{A1}$$

provided the limit exists (see below). Special emphasis will be given to the quantum case as it is more difficult. Throughout what follows Λ denotes a finite region with $|\Lambda|$ sites, and $\Lambda\to\infty$ means that Λ goes to infinity in the sense of van Hove.[15,28] We write

$$\tilde{H}_\Lambda=H_\Lambda+W_\Lambda\,, \tag{A2}$$

where H_Λ and W_Λ, respectively, represent the interaction between the spins in Λ (free boundary conditions) and the interaction between the spins in Λ and those in (its complement) Λ^c, i.e., those which are not in Λ. It is assumed that the interactions have finite range and that their random distribution allows the application of the ergodic theorem. Then the limit in (A1) exists with probability 1 and does not depend on the specific random configuration.[19]

Following Sewell[57] we define a conditional free energy $\tilde{F}_\Lambda(\rho)$ which is designed to represent the free energy of the

open system consisting of the spins in Λ, interacting with one another and with the spins outside that region. $\tilde{F}_\Lambda(\rho)$ is minimal for variations in the state which leave it unchanged *outside* Λ. We will first show that

$$\lim_{\Lambda\to\infty}\frac{1}{|\Lambda|}\tilde{F}_\Lambda(\rho)=f(\beta)\ . \tag{A3}$$

To this end we need some definitions.

Given a state ρ, its restriction ρ_Λ to a finite region Λ is defined by

$$\rho(A)=\mathrm{Tr}\rho_\Lambda A,\quad \forall A\in\mathcal{O}_\Lambda\ , \tag{A4}$$

where $\mathcal{O}_\Lambda$ denotes the collection of all observables inside Λ. One may interpret the density matrix ρ_Λ as the thermal average of ρ with respect to all degrees of freedom outside Λ (partial trace).

Given ρ and ρ_Λ, the entropy of ρ in Λ is given by

$$S_\Lambda(\rho)=-\mathrm{Tr}\rho_\Lambda\ln\rho_\Lambda\ . \tag{A5}$$

The entropy S_Λ is strongly subadditive[15,58]

$$S_{\Lambda_1\cup\Lambda_2}+S_{\Lambda_1\cap\Lambda_2}\le S_{\Lambda_1}+S_{\Lambda_2}\ , \tag{A6}$$

and thus, if $\Lambda\subseteq\Lambda'$,

$$S_{\Lambda'}\le S_\Lambda+S_{\Lambda'\setminus\Lambda}\ , \tag{A7}$$

which we rewrite in the form

$$S_{\Lambda'}-S_{\Lambda'\setminus\Lambda}\le S_\Lambda\ . \tag{A8}$$

By the strong subadditivity (A6) the left side of (A8) is monotonically decreasing in Λ'; moreover it is bounded from below. Hence we find, as $\Lambda'\to\infty$, the *conditional entropy* $\tilde{S}_\Lambda(\rho)$:

$$\tilde{S}_\Lambda(\rho)=\lim_{\Lambda'\to\infty}[S_{\Lambda'}(\rho)-S_{\Lambda'\setminus\Lambda}(\rho)]\le S_\Lambda(\rho)\ . \tag{A9}$$

$\tilde{S}_\Lambda(\rho)$ is used to define the *conditional free energy* $\tilde{F}_\Lambda(\rho)$,

$$\tilde{F}_\Lambda(\rho)=\rho(H_\Lambda)+\rho(W_\Lambda)-\beta^{-1}\tilde{S}_\Lambda(\rho)\ . \tag{A10}$$

Suppose now that ρ is an arbitrary KMS state (the DLR case may be handled analogously). We have to show that the free energy of ρ,

$$\begin{aligned}f(\beta;\rho)&=\lim_{\Lambda\to\infty}\frac{1}{|\Lambda|}F_\Lambda(\rho)\\&=\lim_{\Lambda\to\infty}\frac{1}{|\Lambda|}[\rho(H_\Lambda)-\beta^{-1}S_\Lambda(\rho)]\ ,\end{aligned} \tag{A11}$$

exists and equals $f(\beta)$ as given by (A1). Instead of this we will first prove (A3); the result will follow easily. The point is that $\tilde{F}_\Lambda(\rho)$ satisfies a variational principle[20]

$$\tilde{F}_\Lambda(\rho)=\inf_{\substack{\rho'\\ \rho'_{\Lambda^c}=\rho_{\Lambda^c}}}[\rho'(\tilde{H}_\Lambda)-\beta^{-1}\tilde{S}_\Lambda(\rho')]\ , \tag{A12}$$

where the infimum is taken over all ρ' which agree with ρ outside Λ. Furthermore we have[59]

$$S_\Lambda(\rho)-2\beta||W||\le\tilde{S}_\Lambda(\rho)\le S_\Lambda(\rho)\ , \tag{A13}$$

which we accept for now. The right inequality is just (A9).

Let $\rho_{\beta,\Lambda}=\exp(-\beta H_\Lambda)/\mathrm{Tr}\exp(-\beta H_\Lambda)$ be the canonical Gibbs state of the region Λ. Its free energy F_Λ also satisfies a variational principle, which is well known,[60]

$$\begin{aligned}F_\Lambda&=\rho_{\beta,\Lambda}(H_\Lambda)-\beta^{-1}S_\Lambda(\rho_{\beta,\Lambda})\\&=\inf_{\rho'}[\rho'(H_\Lambda)-\beta^{-1}S_\Lambda(\rho')]\ .\end{aligned} \tag{A14}$$

where ρ' ranges through all the density matrices defined on Λ. We know, by (A1), that $|\Lambda|^{-1}F_\Lambda$ converges to $f(\beta)$ as $\Lambda\to\infty$.

To use (A12) we choose $\rho'=\rho_{\beta,\Lambda}\otimes\rho_{\Lambda^c}$. Then, since ρ' is a product state,

$$\tilde{S}_\Lambda(\rho')=S_\Lambda(\rho')=S_\Lambda(\rho_{\beta,\Lambda}),\quad \rho'(H_\Lambda)=\rho_{\beta,\Lambda}(H_\Lambda)\ . \tag{A15}$$

Thus, by (A12) and (A14),

$$\begin{aligned}\tilde{F}_\Lambda(\rho)&\le\rho_{\beta,\Lambda}(H_\Lambda)+\rho'(W_\Lambda)-\beta^{-1}S_\Lambda(\rho_{\beta,\Lambda})\\&\le F_\Lambda+||W_\Lambda||\ .\end{aligned} \tag{A16}$$

On the other hand, by (A10), (A13), and (A14),

$$\begin{aligned}\tilde{F}_\Lambda(\rho)&=\rho(H_\Lambda)+\rho(W_\Lambda)-\beta^{-1}\tilde{S}_\Lambda(\rho)\\&\ge\rho(H_\Lambda)+\rho(W_\Lambda)-\beta^{-1}S_\Lambda(\rho)\\&\ge\rho_{(H_\Lambda)}-\beta^{-1}S_\Lambda(\rho)-||W_\Lambda||\\&\ge F_\Lambda-||W_\Lambda||\ .\end{aligned} \tag{A17}$$

Combining (A16) and (A17) we find

$$-||W_\Lambda||+F_\Lambda\le\tilde{F}_\Lambda(\rho)\le F_\Lambda+||W_\Lambda||\ , \tag{A18}$$

which implies (A3), since $|\Lambda|^{-1}F_\Lambda\to f(\beta)$ and $|\Lambda|^{-1}||W||\to 0$ as $\Lambda\to\infty$. Furthermore, (A11) and $f(\beta;\rho)=f(\beta)$ also follow directly since, by (A13),

$$\tilde{F}_\Lambda(\rho)\le F_\Lambda(\rho)\le\tilde{F}_\Lambda(\rho)+2||W_\Lambda||\ . \tag{A19}$$

Finally, we prove (A13). Use $\rho'=\rho_\Lambda\otimes\rho_{\Lambda^c}$ in (A12). Then $\tilde{S}_\Lambda(\rho')=S_\Lambda(\rho)$ and

$$\begin{aligned}\tilde{F}_\Lambda(\rho)&=\rho(H_\Lambda)+\rho(W_\Lambda)-\beta^{-1}\tilde{S}_\Lambda(\rho)\\&\le\rho(H_\Lambda)+\rho'(W_\Lambda)-\beta^{-1}S_\Lambda(\rho)\end{aligned} \tag{A20}$$

so that

$$\tilde{S}_\Lambda(\rho)\ge S_\Lambda(\rho)-2\beta||W_\Lambda||\ . \tag{A21}$$

This completes the proof.

[1]K. Binder, in *Fundamental Problems in Statistical Mechanics V,* edited by E. G. D. Cohen (North-Holland, Amsterdam, 1980), pp. 21—51.
[2]S. F. Edwards and P. W. Anderson, J. Phys. F 5, 965 (1975).
[3]A. P. Malozemoff and Y. Imry, Phys. Rev. B 24, 489 (1981).
[4]J. Jäckle, Philos. Mag. B 44, 533 (1981).
[5]R. G. Palmer, Adv. Phys. 31, 669 (1982).
[6]W. E. Fogle, J. D. Boyer, N. E. Phillips, and J. Van Curen, Phys. Rev. Lett. 47, 352 (1981).
[7]J. L. van Hemmen, Phys. Rep. 65, 43 (1980), in particular Sec. 1.
[8]A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949), pp. 29—30.
[9]Reference 7, lemmas 7.1 and 7.2
[10]A. Münster, *Statistische Thermodynamik* (Springer, Berlin, 1956), Sec. 5.12.
[11]R. J. Glauber, J. Math. Phys. 4, 294 (1963).
[12]*Monte Carlo Methods in Statistical Physics,* edited by K. Binder (Springer, Berlin, Heidelberg, New York, 1979).
[13]R. Kindermann and J. L. Snell, *Markov Random Fields and Their Applications* (American Mathematical Society, Providence, R.I., 1980), Sec. 3. Here one finds a nice discussion of the finite-N behavior and its interpretation.
[14]C. N. Yang and T. D. Lee, Phys. Rev. 87, 404 (1952). A second argument will be given shortly.
[15]R. B. Israel, *Convexity in the Theory of Lattice Gases* (Princeton University Press, Princeton, N.J., 1979).
[16]J. Fernando Perez and R. H. Schonmann, J. Stat. Phys. 28, 479 (1982).
[17]D. Ruelle, Commun. Math. Phys. 3, 133 (1966).
[18]As far as we know, the name "broken ergodicity" was coined by Palmer in Ref. 5; see also P. W. Anderson, in *Ill-condensed Matter, Les Houches Lectures, 1978,* edited by R. Balian, Roger Maynard, and Gerard Toulouse (North-Holland, Amsterdam, 1979), p. 163. However, we argue that broken ergodicity is a classical notion, which dates back at least to the midsixties; cf. Refs. 15, 17, and 28 (Ch. 6).
[19]J. L. van Hemmen and R. G. Palmer, J. Phys. A 15, 3881 (1982).
[20]G. L. Sewell, Phys. Rep. 57, 307 (1980), in particular see Sec. 3.2.
[21]Our point of view is corroborated by a result of Georgii: Taking the limit (2.11) we find an ergodic state for "almost every" boundary condition. So, in a sense, free and periodic boundary conditions are quite special. For a discussion and references see, J. Bricmont, J. L. Lebowitz, and C. E. Pfister, J. Stat. Phys. 21, 573 (1979).
[22]L. Onsager, Phys. Rev. 65, 117 (1944).
[23]M. Aizenman, Phys. Rev. Lett. 43, 407 (1979); Commun. Math. Phys. 73, 83 (1980); Y. Higuchi, Colloquia Mathematica Societatis János Bolyai, 27, 517 (1982). A nice illustration of Eq. (3.1) has been given by M. Aizenman in *Mathematical Problems in Theoretical Physics,* Vol. 116 of Springer Lecture Notes in Physics, edited by K. Osterwalder (Springer, Berlin, 1980), p. 143, Fig. 1.
[24]N. D. Mermin, Rev. Mod. Phys. 51, 591 (1979), Sec. IV.
50, 79 (1979).
[26]S. A. Pirogov and Ya. G. Sinai, Theor. Math. Phys. 25, 1185 (1975); 26, 39 (1976); R. Kotecký and S. B. Shlosman, Commun. Math. Phys. 84, 87 (1982).
[27]N. M. Hugenholtz, in *Mathematics of Contemporary Physics,* edited by R. F. Streater (Academic, New York, 1972), pp. 145—182.
[28]D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), Theorem 2.3.3.
[29]D. C. Mattis, *The Theory of Magnetism I* (Springer, Berlin, Heidelberg, New York, 1981), Sec. 6.6.
[30]C. J. Thompson, *Mathematical Statistical Mechanics* (Macmillan, New York, 1972; Princeton University Press, Princeton, N.J. 1979), Sec. 6.4.
[31]A. D. Sokal, J. Stat. Phys. 25, 25 (1981); in particular see the Appendix.
[32]J. Chalupa, Solid State Commun. 22, 315 (1977).
[33]I. Morgenstern and K. Binder, Phys. Rev. Lett. 43, 165 (1979); Phys. Rev. B 22, 288 (1980).
[34]I. Morgenstern and H. Horner, Phys. Rev. B 25, 504 (1982); J. L. van Hemmen and I. Morgenstern, J. Phys. C 15, 4353 (1982).
[35]S. T. Chui, G. Forgacs, and D. M. Hatch, Phys. Rev. B 25, 6952 (1982). The results of this paper are *not* in favor of zero-energy surfaces in three dimensions.
[36]C. N. Yang, Phys. Rev. 85, 808 (1952).
[37]See Bricmont *et al.* (Ref. 21).
[38]C. M. Newman, Commun. Math. Phys. 74, 119 (1980).
[39]A. W. Roberts and D. E. Varberg, *Convex Functions* (Academic, New York, 1973), Sec. II and remark 12.F.
[40]P. W. Anderson, Rev. Mod. Phys. 50, 199 (1978).
[41]This has been confirmed succinctly by a recent discussion in Phys. Rev. Lett.: E. M. Gullikson and S. Schultz, Phys. Rev. Lett. 49, 238 (1982); L. E. Wenger and J. A. Mydosh, *ibid.* 49, 239 (1982); W. E. Fogle, J. D. Boyer, N. E. Phillips, and J. Van Curen, *ibid.* 49, 241 (1982).
[42]J. A. Mydosh in *Disordered Systems and Localization,* Vol. 149 of Springer Lecture Notes in Physics, edited by C. Castellani *et al.* (Springer, Berlin, 1980), pp. 87—106.
[43]B. Simon and A. D. Sokal, J. Stat. Phys. 25, 679 (1981), in particular Secs. 1 and 2.
[44]R. E. Walstedt and L. R. Walker, Phys. Rev. Lett. 47, 1624 (1981).
[45]I. Morgenstern and H. Horner, Phys. Rev. B 25, 504 (1982).
[46]I. Morgenstern, Phys. Rev. B 25, 6067 (1982).
[47]I. Morgenstern (private communication). For the two-dimensional $\pm J$ model, $M \approx 13^2$.
[48]B. Barbara, A. P. Malozemoff, and Y. Imry, Physica 108B, 1289 (1981).
[49]R. V. Chamberlin, M. Hardiman, and R. Orbach, J. Appl. Phys. 52, 1771 (1981).
[50]C. A. M. Mulder, A. J. van Duyneveldt, and J. A. Mydosh, Phys. Rev. B 23, 1384 (1981).
[51]P. Monod and H. Bouchiat, J. Phys. (Paris) Lett. 43, L45 (1982).
[52]J. A. Mydosh, J. Mag. Magn. Mat. 7, 237 (1978).
[53]L. Néel, Ann. Géoph. 5, 99 (1949); and in *Cours de Physique Théorique, Les Houches Lectures, 1961* (Gordon and Breach, New York, 1962), pp. 413—440.
[54]J. L. Tholence and R. Tournier, J. Phys. (Paris) Colloq. 35, C4-229 (1974).
[55]O. E. Lanford and D. Ruelle, Commun. Math. Phys. 13, 194 (1969), Sec. 4.
[56]O. Penrose and J. L. Lebowitz, in *Fluctuation Phenomena,* edited by E. W. Montroll and J. L. Lebowitz (North-Holland, Amsterdam, 1979), Chap. 5.
[57]G. L. Sewell, Phys. Rep. 57, 307 (1980).
[58]E. H. Lieb and M. B. Ruskai, J. Math. Phys. 14, 1938 (1973).
[59]A. C. D. van Enter, Lett. Math. Phys. 6, 289 (1982).
[60]Reference 15, pp. 46—47.