Finite-volume effects in the ferromagnetic phase of the Derrida model at absolute zero in connection with coding

D. B. Saakyan
Erevan Physics Institute, 375036 Erevan, Armenia

(Submitted 18 November 1994; resubmitted 23 January 1995)
Pis'ma Zh. Éksp. Teor. Fiz. **61**, No. 8, 631–635 (25 April 1995)

The free energy is calculated. An expression is derived for finite-volume effects. This expression is analogous to the probability for the decoding error for an optimum channel with intense noise. © *1995 American Institute of Physics.*

Ideas from physics (thermodynamic entropy) have found their way over the years into information theory. A situation which has developed in recent years demonstrates again the profound relationship between these two sciences: coding theory, on the one hand, and spin glasses, on the other. A study of such effects, which lie at the border between two important fields of natural science, may prove to be of interest to each of them. It was hypothesized in Ref. 1 that the Derrida model[2,3] yields the optimum coding in the Shannon sense.[4] This hypothesis was proved for the general case in Refs. 5 and 6.

Let us clarify the physical meaning of the problem. We wish to transmit information from N numbers ξ_i, $i=1,\dots,N$. In place of the N numbers ξ_i, we send a communication of Z numbers τ_j, $j=1,\dots,Z$, where these τ's are calculated by some algorithm from the N numbers ξ_i. In the course of the transmission of the information, the initial τ_j may, with some nonzero probability $(1-m)/2$, change sign at random, while there is a probability $(1+m)/2$ that the value will remain the same. We denote the new constants by τ_j'. If, in the absence of noise, each of the constants τ_j contains an amount of information ln 2, then the τ_j' now contain just a little less information:

$$\ln 2-h(m)\equiv\ln 2+\frac{1+m}{2}\ln\frac{1+m}{2}+\frac{1-m}{2}\ln\frac{1-m}{2}. \tag{1}$$

It is necessary to develop an algorithm which can reconstruct the initial values of the N numbers ξ_i from the noisy constants τ_j'.

Let us construct τ_j in the following way. For the P indices $i_1\dots i_P$ we choose

$$\tau_{i_1\dots i_P}=\xi_{i_1}\dots\xi_{i_P}. \tag{2}$$

From the total number C_N^P of different sets of indices $(i_1\dots i_P)$ we choose, at random, Z sets $\tau_{i_1\dots i_P}$. Using Z coupling constants we construct a function of the N variables σ_i:

$$H(\sigma_i,\xi_j)=-\sum_{(i_1\dots i_P)}\tau'_{i_1\dots i_P}\sigma_{i_1}\dots\sigma_{i_P}. \tag{3}$$

In (3) there is a summation over Z different sets of indices $(i_1,\dots,i_P)$.

When m is equal to 1, we have $\tau_j' = \tau_j$, and the ground state of Hamiltonian (3) is the configuration $\sigma_i = \xi_i$. As was shown in Ref. 5, the configuration $\sigma_i = \xi_i$ is the ground state of Hamiltonian (3) for nonzero values of the noise probability $(1-m)/2$, up to the value of m determined from the equality

$$Z[\ln 2 - h(m)] = N \ln 2. \tag{4}$$

The coding scheme, according to Ref. 1, thus runs as follows: The initial information consists of the values of the N spins ξ_i. From them we construct Z numbers τ_j. In the course of the transmission of information, these are partially replaced by τ_j'. For the decoding we seek the ground state of the Hamiltonian, considering, for example, the thermodynamics of the system as $T \to 0$. We can also calculate an average magnetization,

$$M = \frac{1}{N} \sum_{i=1}^{N} \langle \sigma_i \rangle \xi_i. \tag{5}$$

In the ferromagnetic phase, $M \approx 1$, there are only exponentially small corrections,

$$M \approx 1 - C \exp[-E(m,R)N], \tag{6}$$

where $R \equiv N/Z$ is called the "rate of information transmission" in information theory.

For optimum coding we would have $E(m,R) > 0$ for $R < R_c$. In this case $E(m,R)$ depends on the particular coding method. There exists some maximum value above which $E(m,R)$ cannot be raised. In speaking of $E(m,R)$ below, we mean this maximum value. Values of the function $E(m,R)$ are known only in a certain interval of R values near the critical value R_c found from (4). Finding the function $E(m,R)$ is an important unresolved problem in information theory.

In this letter we calculate finite-volume effects for the Derrida model in the case of complete connectedness. This case corresponds to the limit $m \to 1/2$ in our terminology or the case of intense noise in information theory. In this case, $E(m,R)$ is known over the entire region $0 < R < R_c$.

In our case (coding with the help of the Derrida model) we have one more adjustable parameter, P. As the results of this letter show, when we make the choice $P = N/2$, the finite-volume corrections correspond to this extreme value $E(m,R)$.

Instead of (3) we consider the Hamiltonian

$$H = -\sum_{1 \le i_1 \ldots < i_P = N} \left(J_0 N / C_N^P + j_{i_1 \ldots i_P} \sqrt{N / C_N^P} \right) \sigma_{i_1} \ldots \sigma_{i_P}, \tag{7}$$

where $j_{i_1, \ldots, i_P}$ are random Gaussian constants with a distribution

$$P(j_{i_1 \ldots i_P}) = \frac{1}{\sqrt{\pi}} \exp[-(j_{i_1 \ldots i_P})^2]. \tag{8}$$

Finite-volume effects were calculated in Ref. 2 (with $J_0 = 0$) for spin-glass and paramagnetic phases. We will calculate finite-volume effects using that technique. We need to calculate the quantity

$$\langle \ln Z(B,j) \rangle_j, \tag{9}$$

where the average is over the constant j. A system of N spins has 2^N energy levels. Instead of (9) we can thus consider

$$\int \prod_{\alpha=1}^{2} dE_\alpha P(E_1 ... E_{2^N}) \ln\left[\sum_{\alpha=1}^{2^N} \exp(-BE_\alpha)\right]. \tag{10}$$

Accordingly, we need to somehow calculate the distribution function $P(E_1 ... E_M)$ for M spin configurations. We denote by $\{\sigma_i^1\}$ the ferromagnetic configuration with $\sigma_i^1 = 1$. For this configuration we have

$$P(E_1) = \frac{1}{\sqrt{\pi N}} \exp\left[-\frac{(E_1 + J_0 NB)^2}{N}\right]. \tag{11}$$

For other configurations with δN flipped spins we find an expression

$$P(E_\alpha) = \frac{1}{\sqrt{\pi N}} \exp\left[-\left(E_\alpha + \frac{J_0 A}{N} e^{-K\delta N}\right)^2\right]. \tag{12}$$

This suppression of the second term in the exponential function in (12) in comparison with the corresponding expression in (11) stems from our successful choice $P = N/2$.

In calculating $P(E_1 ... E_M)$ for $M \geqslant 2$ we find an expression of the type

$$P(E_1 ... E_M) = \frac{1}{\sqrt{\pi N}} \exp\left[-\frac{(E_1 + J_0 NB)^2}{N}\right]$$
$$\times \prod_{\alpha=2}^{N} \frac{1}{\sqrt{\pi N}} \exp\left(-\frac{E_\alpha^2}{N}\right) \exp\left[-J_0 \sum_{\alpha<\beta=1}^{N} C_{\alpha\beta} E_\alpha E_\beta\right], \tag{13}$$

where

$$C_{\alpha\beta} \sim \sum_{1 \leqslant i_1 < i_2 ... < i_p \leqslant N} (\sigma_{i_1}^\alpha \sigma_{i_1}^\beta) ... (\alpha_{i_p}^\alpha \sigma_{i_p}^\beta). \tag{14}$$

For (14) we again have the estimate

$$C_{\alpha\beta} \sim \exp(-K\delta N), \tag{15}$$

where K is a combinatorial constant. Let us ignore the last factor in (13). In the calculation of the free energy, this approximation leads to an exponential accuracy. Since the finite-volume effects calculated from a factorized expression for

$$P(E_1 ... E_M) = \frac{1}{\sqrt{\pi N}} \exp\left[-\frac{(E_1 + J_0 NB)^2}{N}\right] \prod_{\alpha=2}^{M} \frac{1}{\sqrt{\pi N}} \exp\left[-\frac{E_\alpha^2}{N}\right], \tag{16}$$

reaches values on the order of one as the boundary of the ferromagnetic phase is approached, although our expressions are correct at these values of J_0, close to $J_0 = \sqrt{\ln 2}$. Incidentally, the agreement of the exponential functions of the corrections with the limiting possible $E(m,R)$ is evidence that it is correct throughout the region $0 < R < R_c$.

Let us use the technique of Ref. 2 to find the average $\langle \ln Z \rangle$:

$$\langle \ln Z\rangle=\int_0^\infty\left[\frac{e^{-t}}{t}-\frac{\langle e^{-tZ}\rangle}{t}\right]dt=\Gamma'(1)-\int_0^\infty \ln t d\langle e^{-tZ}\rangle. \tag{17}$$

Here we need to use expression (16) for $P(E_1...E_M)$. If we had only $P(E_1)$ in (16), we would find the following expression from (17):

$$J_0NB=\Gamma'(1)-\int_{-\infty}^{\infty}uf'(u+J_0NB), \tag{18}$$

where

$$f(u)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}dY\exp[-Y-\exp(u+\lambda Y)],\quad \lambda=\sqrt{N}B.$$

When the other factors in (16) are taken into account, we find the expression

$$\langle \ln Z(J,B)\rangle_j=J_0NB-\int_{-\infty}^{\infty}f(u+J_0NB)[1-f(u)^{2^N-1}]. \tag{19}$$

The function $f(u)$ falls off monotonically, from a value of 1 at $-\infty$ to zero (at an exponential accuracy) at positive values of u. For $f(u)$ we have different asymptotic regimes:

$$0<u,\quad f(u)=\frac{1}{\sqrt{\pi}}\left|\Gamma\left(\frac{2u}{\lambda^2}\right)\right|\exp\left(-\frac{u^2}{\lambda^2}\right), \tag{20}$$

$$\frac{-\lambda^2}{2}<u<0,\quad f(u)=1-\frac{1}{\sqrt{\pi}}\left|\Gamma\left(\frac{2u}{\lambda^2}\right)\right|\exp\left(-\frac{u^2}{\lambda^2}\right), \tag{21}$$

$$-\lambda^2<u<-\lambda^2/2,\quad f(u)=1-\exp(u+\lambda^2/4). \tag{22}$$

The function $[1-f(u)^{2^N-1}]$ has the behavior

$$\exp[N\ln 2-u^2/\lambda^2] \tag{23}$$

in the region $u<2\sqrt{N\ln 2}$, while it behaves as one at $u>2\sqrt{N\ln 2}$.

If we ignore the pre-exponential factors, we need consider only the expression

$$\int_{-J_0NB}^{-\sqrt{N\ln 2}B}\exp\left[-\frac{(u+J_0NB)^2}{\lambda^2}+\left(\frac{N\ln 2-u^2}{\lambda^2}\right)\right]+\int_{-\sqrt{N\ln 2B}}^{\infty}du\exp\left[-\frac{(u+J_0NB)^2}{\lambda^2}\right]. \tag{24}$$

The contribution of the second term in (24) is always small in comparison with that of the first. The saddle point for the first integral is

$$u_0=-J_0NB/2. \tag{25}$$

When it lies in the integration, i.e., under the condition $J_0>2\sqrt{\ln 2}$, we have the following expression for the corrections to the free energy:

$$\exp\{N[-(J_0)^2/2+\ln 2]\}. \quad (26)$$

If, in contrast, u_0 does not lie in the integration cut, i.e., if $\sqrt{\ln 2}<J_0<2\sqrt{\ln 2}$, we have

$$\exp[-(J_0-\sqrt{\ln 2})^2 N]. \quad (27)$$

Expressions (26) and (27) are the same as the corresponding regimes for $\exp[-E(m,R)N]$ in information theory.

Determining the situation in the general case $m \neq 1/2$ is a goal for further research.

I deeply thank E. Arutyunyan for a consultation on information theory and A. Allakhverdyan and Sh. Rukhani for a useful discussion.

This study has financial support from the International Science Foundation (Grant MVM000) and the Ministry of Research and Technology of the Federal Republic of Germany (Grant 211-5291).

[1] N. Sourlas, Nature **339**, 693 (1989).
[2] B. Derrida, Phys. Rev. B **24**, 2613 (1981).
[3] D. Gross and M. Mezard, Nucl. Phys. B **240**, 43 (1984).
[4] I. Csiszar and J. Korner, *Information Theory* (Academic, New York, 1981; Moscow, Mir, 1995).
[5] D. B. Saakyan, JETP Lett. **55**, 192 (1992).
[6] D. B. Saakyan, Phys. Lett. A **180**, 163 (1993).

Translated by D. Parsons