Relative entropy in the Sherrington–Kirkpatrick spin glass model

2002 J. Phys. A: Math. Gen. 35 7773

(http://iopscience.iop.org/0305-4470/35/36/305)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 7773–7778 PII: S0305-4470(02)36697-6

Relative entropy in the Sherrington–Kirkpatrick spin glass model

Peter Varga

Institute of Informatics and Mathematics, University of Debrecen, Debrecen, Hungary

E-mail: varga@math.klte.hu

Received 6 May 2002
Published 28 August 2002
Online at stacks.iop.org/JPhysA/35/7773

Abstract

It has been observed that in the Sherrington–Kirkpatrick spin glass model the relative entropy of the probability distribution of the product of two independent replicas with respect to the uniform distribution exhibits a sort of phase transition above the critical point of the model. In this paper, we show that this is not an isolated phenomenon as, for example, the product of three replicas shows analogous behaviour. We also study the probability distribution of the product of a spin glass and a ferromagnetic system.

PACS numbers: 05.70.Fh, 02.50.Cw, 75.10.Nr

1. Introduction

The Sherrington–Kirkpatrick (SK) [3] model of the mean field theory of spin glasses exhibits interesting behaviour even in the high-temperature regime. Catoni [1] and Comets [2] observed that at inverse temperature $\beta = 1/\sqrt{2}$ (which is above the critical point $\beta_c = 1$) the model goes through a sort of secondary phase transition. They calculated the relative entropy of the distribution of the product of two independent replicas with respect to the uniform probability measure. They found that this relative entropy is very small if $\beta < 1/\sqrt{2}$, while its per site value is positive in the thermodynamic limit if $\beta > 1/\sqrt{2}$. Since one expects some sort of uniformity of behaviour in the high-temperature regime, it is reasonable to look for similar phenomenon also at other values of β. In this paper, we show that around $\beta \approx 0.799$ the product of three independent replicas shows similar behaviour.

The relative entropy measures how evenly a probability measure (PM) is distributed among the spin configurations. The PM on the product of two replicas corresponds to the convolution of the PMs of the individual replicas. The convolution of PMs usually gives a more evenly distributed PM. So these sorts of results measure how far we need to spread out or smooth the spin glass PM in order to reach the uniform distribution. However, this information is quite implicit, since the spin glass PM is not a very simple object. So it is reasonable to try to compute the convolution of the spin glass PM with something simpler, such as the PM of a ferromagnetic Curie–Weiss (CW) model. The PM of the CW model is basically the uniform distribution in the high-temperature phase, so in that case the result of the convolution should also be the uniform distribution. This observation gives a simple proof of the fact that the free energy of the SK model is not changed by a sufficiently small ferromagnetic perturbation of the random couplings. This phenomenon was already observed by Sherrington and Kirkpatrick in [3].

We describe the computation of the relative entropy of the product of three replicas in section 2, while section 3 contains the study of the convolution of the PMs of a spin glass and a ferromagnetic system.

2. The product of three spin glass replicas

Let us recall that the SK model is defined by the Hamiltonian

$$H_J(s) = -\sum_{0<i<j\leqslant N}^{N} J_{ij} s_i s_j \tag{1}$$

where J_{ij} are independent Gaussian random variables with zero mean and variance $1/N$, and $s_i = \pm 1$. $H_J(s)$ defines the probability distribution on the s spin configurations:

$$\rho_{J,\beta}(s) = Z_J(\beta)^{-1} \exp\{-\beta H_J(s)\} \qquad \text{where} \qquad Z_J(\beta) = \sum_s \exp\{-\beta H_J(s)\}. \tag{2}$$

The s spin configurations might be taken as members of the Abelian group $\mathbb{Z}_2^N$ with sitewise product,

$$(s \cdot s')_i = s_i s_i'. \tag{3}$$

The convolution of two discrete probability measures is defined by

$$(p_1 * p_2)(s) = \sum_s p_1(s') p_2(s'^{-1} s). \tag{4}$$

The relative entropy of two distributions is

$$h(p_1, p_2) = \sum_s p_1(s) \log \frac{p_1(s)}{p_2(s)}. \tag{5}$$

Catoni and Comets showed that

$$\frac{1}{N} \mathbb{E}_J h(u, \rho_{J,\beta} * \rho_{J,\beta}) \tag{6}$$

tends to zero as $N \to \infty$ if $\beta < 1/\sqrt{2}$ while, if $\beta > 1/\sqrt{2}$, it approaches a positive value. (Here u denotes the uniform probability measure and $\mathbb{E}_J$ denotes the expectation value with respect to the distribution of the J_{ij} couplings.) With the help of the replica method, we determine the point where

$$\frac{1}{N} \mathbb{E}_J h(u, \rho_{J,\beta} * \rho_{J,\beta} * \rho_{J,\beta}) = \frac{1}{N} \mathbb{E}_J h\left(u, \rho_{J,\beta}^3\right) \tag{7}$$

shows similar behaviour.

So let us compute $h\big(u, \rho^3_{J,\beta}\big)$:

$$h\left(u, \rho^3_{J,\beta}\right) = \sum_{\psi} \frac{1}{2^N} \log\left[\frac{1}{2^N}\left(\frac{\sum_{s^1,s^2} \exp\{-\beta(H_J(s^1)+H_J(s^2)+H_J(s^1\cdot s^2\cdot\psi))\}}{\sum_{s^1,s^2,s^3} \exp\{-\beta(H_J(s^1)+H_J(s^2)+H_J(s^3))\}}\right)^{-1}\right]$$
$$= -N\log 2 + \log[Z_J(\beta)]^3 - \frac{1}{2^N}\sum_{\psi}\log\sum_{s^1,s^2}\exp\{-\beta[H_J(s^1) + H_J(s^2) + H_J(s^1\cdot s^2\cdot\psi)]\}. \tag{8}$$

We need to compute the expectation value of $h\big(u, \rho^3_{J,\beta}\big)$ with respect to the distribution of the J_{ij} couplings. It is known [3] that $\mathbb{E}_J \log Z_J(\beta) = N(\beta^2/4 + \log 2)$, while the last term can be computed by the standard techniques of the replica method. Note that

$$\mathbb{E}_J \frac{1}{2^N}\sum_{\psi}\log\sum_{s^1,s^2}\exp\{-\beta(H_J(s^1)+H_J(s^2)+H_J(s^1\cdot s^2\cdot\psi))\}$$
$$= \mathbb{E}_J \log\sum_{s^1,s^2}\exp\{-\beta H_{J,\mathbf{1}}(s^1,s^2)\} \tag{9}$$

where $H_{J,\mathbf{1}}(s^1, s^2) = H_J(s^1) + H_J(s^2) + H_J(s^1\cdot s^2)$, since the distribution of the couplings is invariant under the $J_{ij} \to J^{\psi}_{ij} = J_{ij}\psi_i\psi_j$ gauge transformation. For the computation of the expected value of the logarithm, we employ the replica trick and compute

$$\mathbb{E}_J Z_J(\beta)^n = \mathbb{E}_J\left(\sum_{s^1,s^2}\exp\{-\beta H_{J,\mathbf{1}}(s^1,s^2)\}\right)^n$$
$$= \int \mathrm{d}\mu(J_{ij}) \sum_{s^{1,a},s^{2,a}} \exp\left\{-\beta\sum_{a}^{n}\sum_{i<j} J_{ij}\left(s_i^{1,a}s_j^{1,a} + s_i^{2,a}s_j^{2,a} + s_i^{1,a}s_i^{2,a}s_j^{1,a}s_j^{2,a}\right)\right\}$$
$$= \sum_{s^{1,a},s^{2,a}} \exp\left\{\frac{\beta^2}{2N}\sum_{i<j}\left(\sum_a s_i^{1,a}s_j^{1,a} + s_i^{2,a}s_j^{2,a} + s_i^{1,a}s_i^{2,a}s_j^{1,a}s_j^{2,a}\right)^2\right\}. \tag{10}$$

The expectation value of $\log Z_J$ can be recovered from the derivative of Z^n_J at $n = 0$. After some straightforward [4] manipulations, we obtain that

$$\mathbb{E}_J Z^n_J = \int \prod_{A;B} \mathrm{d}Q_{A;B}\left(\frac{N\beta^2}{2\pi}\right)^{1/2} \exp\{-NA[Q]\} \tag{11}$$

$$A[Q] = -\frac{3n\beta^2}{4} + \frac{\beta^2}{2}\sum_{A;B} Q^2_{AB} - \log Z[Q] \tag{12}$$

$$Z[Q] = \sum_{S^1,S^2}\exp\{-\beta H[Q, S^1, S^2]\} \tag{13}$$

$$H[Q, S^1, S^2] = -\beta\left(\sum_{a<b} Q^{1\cdot 1}_{ab} S^1_a S^1_b + Q^{2\cdot 2}_{ab} S^2_a S^2_b + Q^{12\cdot 12}_{ab} S^1_a S^2_a S^1_b S^2_b + \sum_{a,b} Q^{1\cdot 12}_{ab} S^1_a S^1_b S^2_b + Q^{2\cdot 12}_{ab} S^2_a S^1_b S^2_b + Q^{1\cdot 2}_{ab} S^1_a S^2_b\right). \tag{14}$$

At this point, we make the replica symmetric approximation for Q:

$$Q_{ab}^{1\cdot 1} = Q_{ab}^{2\cdot 2} = q \qquad Q_{ab}^{1\cdot 2} = r \qquad Q_{ab}^{12\cdot 12} = s \qquad Q_{ab}^{1\cdot 12} = Q_{ab}^{2\cdot 12} = t. \tag{15}$$

Although the replica symmetric ansatz for Q does not predict correctly the exact value of $\mathbb{E}_J Z_J^n$, it does provide the correct location of the critical point. With these choices, we obtain that

$$\frac{1}{N}\mathbb{E}_J h\left(u, \rho_{J,\beta}^3\right) = -\log 2 + 3(\beta^2/4 + \log 2) + \mathrm{crit}(f) \tag{16}$$

where

$$f = -\frac{3}{4}\beta^2 + \frac{\beta^2}{2}(-q^2 - s^2) + \beta^2(q + s/2) - \int \prod_{i=1}^{9} \frac{\mathrm{d}x_i}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{9} x_i^2\right\}$$
$$\times \log(\mathrm{e}^{A+B+C} + \mathrm{e}^{-A-B+C} + \mathrm{e}^{A-B-C} + \mathrm{e}^{-A+B-C}) \tag{17}$$

$$A = \sqrt{q}x_1 + \sqrt{t/2}(x_4 + \mathrm{i}x_5) + \sqrt{r/2}(x_8 + \mathrm{i}x_9) \tag{18}$$

$$B = \sqrt{q}x_2 + \sqrt{t/2}(x_6 + \mathrm{i}x_7) + \sqrt{r/2}(x_8 - \mathrm{i}x_9) \tag{19}$$

$$C = \sqrt{s}x_3 + \sqrt{t/2}(x_4 - \mathrm{i}x_5) + \sqrt{t/2}(x_6 + \mathrm{i}x_7) \tag{20}$$

and $\mathrm{crit}(f)$ denotes the critical value of $f(q, r, s, t)$. In the standard SK model, one has to maximize the free energy with respect to the overlap parameter q. Since q occurs $n(n-1)$ times in the overlap matrix $Q_{a,b}$, which is negative at small positive values of n, the maximization of the free energy with respect to negative numbers of parameters is regarded as the correct analogue of the usual minimization of the free energy. Similar considerations should be applied in our case to the q and s parameters, but we should minimize f with respect to the r, t parameters since they occur $n^2 > 0$ times in the $Q_{ab}^{1\cdot 2}$, $Q_{ab}^{1\cdot 12}$ or $Q_{ab}^{2\cdot 12}$ matrices. The f can be computed after some (computer) algebra up to second order:

$$f(q, r, s, t) \approx -\log 4 + \beta^2\left(-\frac{3}{4} + \frac{qs\beta^2}{4} - \frac{q^2}{2} + \frac{3q^2\beta^2}{4} - \frac{s^2}{2} + \frac{r^2\beta^2}{4} + \frac{5s^2\beta^2}{16} + \frac{t^2\beta^2}{2}\right). \tag{21}$$

The $q = r = s = t = 0$ critical point loses its stability when the quadratic part of f in q and s becomes semidefinite. This happens when

$$\frac{11}{64}\beta^4 - \frac{17}{32}\beta^2 + \frac{1}{4} = 0 \quad \Rightarrow \quad \beta_3 = \frac{\sqrt{17 - \sqrt{113}}}{\sqrt{11}} \approx 0.761. \tag{22}$$

After this value $\frac{1}{N}\mathbb{E}_J h\left(u, \rho_{J,\beta}^3\right)$ attains a positive value. Note that this 'critical point' corresponds to a somewhat lower temperature than $\beta_2 = 1/\sqrt{2}$, where $\frac{1}{N}\mathbb{E}_J h\left(u, \rho_{J,\beta}^2\right)$ becomes positive. This result is quite reasonable, since the convolution of probability measures usually smooths them, so $\rho_{J,\beta}^3$ should be closer to the uniform distribution than $\rho_{J,\beta}^2$. It is clear that similar phenomenon occurs for the higher powers of $\rho_{J,\beta}$ above the critical inverse temperature $\beta = 1$.

3. The product of a spin glass and a ferromagnetic system

Let us consider besides the SK model the ferromagnetic Curie–Weiss mean field model defined by the Hamiltonian

$$H_{\mathrm{CW}} = -\frac{J_0}{N}\sum_{0<i<j\leqslant N} \sigma_i\sigma_j. \tag{23}$$

For $J_0\beta < 1$ the model is in the disordered phase and the PM of the σ spins is very close to the uniform distribution. Now let us compute the relative entropy of the PM of the product $\psi_i = \sigma_i s_i$ with respect to the u uniform distribution. Here s is the same as in section 2.

The joint Hamiltonian of s and σ is

$$\begin{aligned} H(s,\sigma) &= -\sum_{0<i<j\leqslant N} J_{ij}s_i s_j + \frac{J_0}{N}\sigma_i\sigma_j \\ &= -\sum_{0<i<j\leqslant N}\left(J_{ij}\psi_i\psi_j + \frac{J_0}{N}\right)\sigma_i\sigma_j = H_\psi(\sigma). \end{aligned} \tag{24}$$

The probability distribution of ψ is given by

$$p_J(\psi) = \frac{\sum_\sigma \exp\left\{\beta\sum_{0<i<j\leqslant N}\left(J_{ij}\psi_i\psi_j + \frac{J_0}{N}\right)\sigma_i\sigma_j\right\}}{\sum_{\sigma,s}\exp\left\{\beta\sum_{0<i<j\leqslant N} J_{ij}s_i s_j + \frac{J_0}{N}\sigma_i\sigma_j\right\}}. \tag{25}$$

The relative entropy of p_J is

$$h(u,p_J) = \sum_\psi \frac{1}{2^N}\log\frac{\frac{1}{2^N}}{p_J(\psi)}. \tag{26}$$

So

$$\begin{aligned} \frac{1}{N}\mathbb{E}_J h(u,p_J) &= \frac{1}{N}\mathbb{E}_J\left(\sum_\psi \frac{1}{2^N}\left[\log\frac{1}{2^N} + \log\sum_{\sigma,s}\mathrm{e}^{-\beta H(s,\sigma)} - \log\sum_\sigma \mathrm{e}^{-\beta H_\psi(\sigma)}\right]\right) \\ &= \frac{1}{N}\mathbb{E}_J\left(-N\log 2 + \log\sum_s \mathrm{e}^{-\beta H_{\mathrm{SK}}(s)} + \log\sum_\sigma \mathrm{e}^{-\beta H_{\mathrm{CW}}(\sigma)}\right. \\ &\quad \left. - \log\sum_\sigma \exp\left\{-\beta\sum_{i<j}\left(J_{ij} + \frac{J_0}{N}\right)\sigma_i\sigma_j\right\}\right). \end{aligned} \tag{27}$$

Here we considered that the distribution of the J_{ij} couplings is invariant under the $J_{ij} \to J_{ij}\psi_i\psi_j$ transformation.

If $J_0\beta < 1$ then this expression should be zero, since the p_J distribution is a convolution where one factor is already the almost uniform distribution of the σ spins. As $-N\log 2$ cancels $\log\sum_\sigma \exp\{-\beta H_{\mathrm{CW}}(\sigma)\}$ in the high-temperature phase, the following equation must hold:

$$\mathbb{E}_J \log\sum_s \exp\left\{\beta\sum_{i<j} J_{ij}s_i s_j\right\} = \mathbb{E}_J\log\sum_s\exp\left\{\beta\sum_{i<j}\left(J_{ij} + \frac{J_0}{N}\right)s_i s_j\right\}. \tag{28}$$

From this, we conclude that if $J_0 < 1/\beta$ than the ferromagnetic perturbation $J_{ij} \to J_{ij} + J_0/N$ does not change the free energy of the spin glass system. This phenomenon was already observed by Sherrington and Kirkpatrick [3]. Let us note that the observation of Catoni [1] on the relative entropy of the PM of the product of two spin glass replicas was a by-product of his derivation of improved bounds of the free energy of the SK model.

References

[1] Catoni O 1996 The Legrende transform of two replicas of the S-K spin glass model *Prob. Theor. Rel. Fields* **105** 369–92

[2] Comets F 1998 The martingale method for mean-field disordered systems at high temperature *Mathematical Aspects of Spin Glasses and Neural Networks* ed A Bovier and P Picco *(Progress in Probability vol 41)* (Basle: Birkhauser) pp 90–113

[3] Sherrington D and Kirkpatrick S 1978 Infinite-ranged models of spin glasses *Phys. Rev. B* **17** 4384–403

[4] Mezard M, Parisi G and Virarsoro M A 1988 *Spin Glass Theory and Beyond* (Singapore: World Scientific)