# Complexity of Ising Spin Glasses

T. Aspelmeier, A. J. Bray, and M. A. Moore

*Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom*

(Received 8 September 2003; published 26 February 2004)

We compute the complexity [logarithm of the number of Thouless-Anderson-Palmer (TAP) states] associated with minima and index-one saddle points of the TAP free energy. Higher-index saddles have smaller complexities. The two leading complexities are equal, consistent with the Morse theorem on the total number of turning points, and have the value given by Bray and Moore [J. Phys. C **13**, L469 (1980)]. In the thermodynamic limit, TAP states of all free energies become marginally stable.

DOI: 10.1103/PhysRevLett.92.087203 PACS numbers: 75.10.Nr, 75.50.Lk

The computation of the complexity of spin glasses has recently attracted an avalanche of renewed interest. More than two decades ago, two of us [1] [Bray and Moore (BM)] computed the average number of solutions, $\langle N_s \rangle_J$, of the Thouless-Anderson-Palmer (TAP) equations [2] for the Sherrington-Kirkpatrick Ising spin-glass model. We found $\langle N_s \rangle_J \sim \exp[N\Sigma(T)]$, where $\Sigma(T)$ is the complexity (per spin) at temperature $T$ and $N$ is the number of spins. The function $\Sigma(T)$ vanishes at the spin-glass critical point $T_c$, while for $T \to 0$ it approaches the value $0.1992\ldots$ obtained from independent calculations of the number of one-spin-flip-stable states at $T = 0$ [1,3].

Various aspects of this calculation have recently been (implicitly or explicitly) criticized [4–9]. The essence of this criticism is as follows. The TAP equations take the form $G_i = 0$, $i = 1, \ldots, N$, where each $G_i$ is a function of the $N$ variables $m_i$ which denote the local magnetizations, $m_i = \langle S_i \rangle$, at each site, where the brackets are thermal averages and $S_i = \pm 1$ is an Ising spin variable. The number of solutions of the $N$ equations $G_i = 0$ is given by $N_s = \int_{-1}^{1} \prod_i dm_i \prod_i \delta(G_i) |\det \partial G_i/\partial m_j|$, where the modulus sign on the determinant ensures that each solution is counted with weight unity. In BM (and many subsequent calculations) the modulus sign was dropped. This is valid if the matrix $\partial G_i/\partial m_j$ is positive definite, which requires all TAP solutions to be local minima of the TAP free energy. We argue that, with one modification, this idea is essentially correct. This seems, at first sight, paradoxical because, if the modulus sign is dropped, TAP solution $s$ is weighted by $(-1)^{n_s}$, where $n_s$ is the index (number of negative eigenvalues) of the saddle point, giving $N_s = \sum_s (-1)^{n_s}$. The Morse theorem states that this sum is a topological invariant equal (in this case) to unity [10], implying that not all TAP states can be minima. We resolve this apparent contradiction by showing that, in the thermodynamic limit and within the BM solution, the matrix $\partial G_i/\partial m_j$ is positive *semidefinite*, with exactly one null eigenvalue. For $N$ large but finite, we argue that the solutions appear as pairs consisting of a minimum and an index-one saddle point, both types of solution having the same (extensive) complexity thereby rescuing the Morse theorem.

Recent work on spin-glass complexity [4–8] exploits a symmetry of the action, the so-called Becchi-Rouet-Stora-Tyutin (BRST) supersymmetry [11], that enters the calculation of $N_s$. If this symmetry is unbroken, the Morse theorem follows. However, the solutions that preserve this symmetry have been shown [8] to violate the convexity inequality $\langle \ln N_s \rangle_J \leq \ln \langle N_s \rangle_J$, where the brackets represent disorder averages, leading Crisanti *et al.* to conjecture that there might be no extensive complexity [8]. The BM solution breaks the BRST symmetry and satisfies all physical requirements, provided the apparent difficulties with the Morse theorem can be resolved. It is the purpose of this Letter to present such a resolution. As a spin-off from our calculation, we note that the marginal stability of TAP states provides a possible explanation for why they are so difficult to find numerically.

The free- energy (multiplied by $\beta = 1/k_B T$) of a TAP state is given by [2]

$$F = -\frac{\beta}{2}\sum_{i,j} J_{ij} m_i m_j - \frac{N}{4}\beta^2(1-q)^2 - N\ln 2 + \sum_i \left(\frac{1}{2}\ln(1-m_i^2) + m_i \tanh^{-1} m_i\right), \quad (1)$$

where $q = (1/N)\sum_i m_i^2$, and the bonds $J_{ij}$ are drawn from a Gaussian distribution of mean zero and variance $1/N$.

The TAP equations are given by $G_i \equiv \partial F/\partial m_i = 0$, for all $i = 1, \ldots, N$, where

$$G_i = \tanh^{-1} m_i + \beta^2(1-q)m_i - \beta \sum_j J_{ij} m_j. \quad (2)$$

Using $G_i = 0$, the free energy $F$ can be rewritten as a sum of single-site terms, $F = \sum_i f_1(m_i)$, where

$$f_1(m) = -\ln 2 - (1/4)\beta^2(1-q^2) + (m/2)\tanh^{-1} m + (1/2)\ln(1-m^2). \quad (3)$$

The number of solutions (per unit free-energy range) with free energy per spin $f = F/N$ is given by

  

$$N_s(f) = N^2 \int_0^1 dq \int_{-1}^1 \prod_i (dm_i)\delta\Big(Nq - \sum_i m_i^2\Big) \times \delta\Big(Nf - \sum_i f_1(m_i)\Big)\prod_i \delta(G_i)|\det\mathbf{A}|, \quad (4)$$

where $\mathbf{A}$ is the Hessian matrix,

$$A_{ij} = \partial G_i/\partial m_j = \partial^2 f/\partial m_i \partial m_j = \Big[\frac{1}{1-m_i^2} + \beta^2(1-q)\Big]\delta_{ij} - \beta J_{ij} - \frac{2\beta^2}{N} m_i m_j. \quad (5)$$

The final term in Eq. (5) is $O(1/N)$ and was omitted in BM. It does not contribute to the extensive part of the complexity, only to the prefactor of the exponential in the relation $\langle N_s(f)\rangle_J \sim \exp[N\Sigma(f)]$. It has, however, the form of a projector and may play an important role in the eigenvalue spectrum of the matrix $\mathbf{A}$, as emphasized by Plefka [12]. In particular, it may determine the sign of $\det\mathbf{A}$. We show that, for $N \to \infty$, the projector term splits off a single isolated eigenvalue from the continuous spectrum of $\mathbf{A}$. Furthermore, the continuous part contains only positive eigenvalues, while the isolated eigenvalue is a null eigenvalue outside the continuum.

Equation (4) is the common starting point for all calculations of $N_s(f)$. Here we focus on the configuration average (sometimes called the "annealed average" or "white average"), $\langle N_s(f)\rangle_J$, over realizations of the disorder. However, it is not straightforward to do this while retaining the modulus on the determinant. Dropping the modulus, the calculation can be completed and the result takes the form [1]

$$\frac{1}{N}\ln\langle N_s(f)\rangle_J = -\lambda q - uf - (B+\Delta)(1-q) + (B^2-\Delta^2)/2\beta^2 + \ln I, \quad (6)$$

where $I$ is a function of the parameters $\lambda$, $q$, $u$, $B$, and $\Delta$ and is defined by the integral

$$I = \int_{-1}^1 \frac{dm}{\sqrt{2\pi P}}\Big(\frac{1}{1-m^2} + B\Big) \times \exp\Big[\lambda m^2 + uf_1(m) - \frac{(\tanh^{-1}m - \Delta m)^2}{2P}\Big], \quad (7)$$

where $P = \beta^2 q$. The left-hand side of (6) is the complexity or, more properly, the "annealed complexity." The parameters $\lambda, \ldots, \Delta$ originally entered the calculation as integration variables: $\lambda$ and $u$ appear in auxiliary integrations that relax the delta function constraints on $q$ and $f$, respectively, while the other parameters were introduced via Hubbard-Stratonovich transformations that reduce the problem to a single-site problem. Full details can be found in [1,4]. The resulting five-dimensional integral can be evaluated (for $N \to \infty$) by the method of steepest descents, so the five parameters take values corresponding to the appropriate saddle point in the five-dimensional space. Note that a calculation of the *total* number of solutions, independent of their free energy, requires setting $u = 0$.

It is straightforward to derive the five saddle-point equations from Eqs. (6) and (7). The same five equations appear in [1,4]. Equivalent equations have also been derived by De Dominicis *et al.* [13]. The equations admit the solution $B = 0$, and this is the solution adopted in all three papers. The differences between the various subsequent treatments [4–8] arise in the solution of the remaining four equations.

Consider first the case $u = 0$, corresponding to a calculation of the total number of solutions. Cavagna, Giardina, Parisi, and Mézard (CGPM) [4] note that the BM solution apparently violates the Morse theorem and propose a new BRST-symmetric solution that gives vanishing complexity for $u = 0$. As $u$ is decreased, states of lower free energy are selected and, within their solution, CGPM find that there exists a threshold $f$ below which $\ln\langle N_s(f)\rangle_J$ is nonzero. Unfortunately, however, an important inequality, $x_p \equiv 1 - (\beta^2/N)\sum_i (1-m_i^2)^2 \geq 0$, is violated in the CGPM solution, rendering it unphysical [6]. The condition $x_p \geq 0$ is necessary for the internal consistency of the TAP equations [12]. This inequality is satisfied by the BM solution [6]. The BM solution is internally consistent, therefore, provided one can demonstrate that the matrix $\mathbf{A}$ is positive definite, guaranteeing the positivity of the determinant and justifying the replacement of $|\det\mathbf{A}|$ by $\det\mathbf{A}$ in the calculation, and provided one can understand the apparent violation of the Morse theorem that ensues. The remainder of the Letter is devoted to these points.

We first rewrite Eq. (5) in the form $A_{ij} = (X^{-1})_{ij} - (2\beta^2/N)m_i m_j$, in which the projector term has been separated off. The matrix $\mathbf{A}^{-1}$ is the susceptibility matrix, $(A^{-1})_{ij} = \partial m_i/\partial h_j$, giving the response to a site-dependent external field, and $X_{ij}$ is the $O(1)$ contribution to it. The eigenvalue spectrum of $\mathbf{X}^{-1}$ can be obtained using either Pastur's theorem [14] or the "locator expansion" of Ref. [15]. In the limit $N \to \infty$, the spectrum consists of a continuous band of positive eigenvalues for both $x_p > 0$ and $x_p < 0$ (though, as noted, the TAP equations themselves become unphysical for $x_p < 0$), and the left edge of the band reaches zero only for $x_p = 0$. For the BM solution $x_p > 0$, so all eigenvalues of $\mathbf{X}^{-1}$ are positive. When the projector term is included, an isolated eigenvalue, outside the main band, is produced. Using the eigenvectors of $\mathbf{X}^{-1}$ as a basis, it is easy to show [16] that this eigenvalue has a non-negative value provided the inequality

$$1 \geq \frac{2\beta^2}{N}\sum_{ij} m_i X_{ij} m_j = 2\beta^2 H \quad (8)$$

is satisfied, where the final equality defines $H$. The same result can be obtained using the variational trial function $v_i = \sum_j X_{ij} m_j$ for the eigenvector of $\mathbf{A}$ with smallest eigenvalue, i.e., $\lambda_{\min} \leq \sum_{i,j} v_i A_{ij} v_j / \sum_i v_i^2 \propto (1 - 2\beta^2 H)$.

If the inequality (8) becomes an equality, the isolated eigenvalue $\lambda_{\min}$ has the value zero, and the variational eigenfunction becomes exact. A variant of the inequality (8) was derived earlier [12], with only the diagonal terms, $i = j$, appearing on the right. The off-diagonal terms were missing due to the use of Pastur's theorem outside its range of validity [17].

The quantity $H$ in Eq. (8) can be computed as follows. We introduce an additional factor of unity, expressed as

$$1 = \frac{1}{\sqrt{\det \mathbf{X}}} \int_{-\infty}^{\infty} \prod_i \left( \frac{d\phi_i}{\sqrt{2\pi}} \right) \exp\left( -\frac{1}{2} \sum_{i,j} \phi_i (X^{-1})_{ij} \phi_j \right), \quad (9)$$

in the integrand of Eq. (4) and obtain $H$ from $H = (1/N)\sum_{ij} \langle m_i m_j \langle \phi_i \phi_j \rangle_\phi \rangle_m$, where the averages $\langle \ldots \rangle_\phi$ and $\langle \ldots \rangle_m$ are over the variables $\{\phi_i\}$ and $\{m_i\}$, respectively. The weight function for the $\phi_i$ integrals is given the integrand in Eq. (9), while for the $m_i$ integrals it is given by the integrand in Eq. (4).

After a straightforward but lengthy calculation one finds [16]

$$H = \frac{A_3 q^2}{(q - A_1)^2 + A_3[\beta^2 q(1-q) - A_2]}, \quad (10)$$

where

$$A_1 = \langle (1 - m^2) m (\tanh^{-1} m - \Delta m) \rangle, \quad (11)$$

$$A_2 = \langle (1 - m^2)(\tanh^{-1} m - \Delta m)^2 \rangle, \quad (12)$$

$$A_3 = \langle m^2 (1 - m^2) \rangle, \quad (13)$$

and the averages are now over the weight function given by the integrand in Eq. (7). Carrying out the required integrals numerically (with $B = 0$ as usual) one obtains a remarkable result: the quantity $2\beta^2 H$ is unity for all temperatures $T < T_c$ and all values of the free energy per spin, $f$. It follows that the inequality (8) is satisfied as an equality and that the Hessian has one null eigenvalue [18].

The result that, in the thermodynamic limit, there is always one exactly zero eigenvalue, but no negative eigenvalue, is the key to resolving all the puzzles surrounding this problem. First, $\det \mathbf{A}$ vanishes, so the prefactor of the exponentially large number of TAP states is, for $N \to \infty$, exactly zero, in accordance with the result of Kurchan [10] and its extension to general values of $u$ (the variable conjugate to $f$) [8]. However, the exponential itself diverges for $N \to \infty$, so the product of exponential and prefactor is not defined in this limit. To make sense of it, one has to keep $N$ large but finite. The result, confirmed by numerical studies, is that the zero eigenvalue is shifted, for finite $N$, to a small positive or negative value, corresponding to a TAP minimum or to a saddle of index one, respectively. The shift would be expected to be of order $1/\sqrt{N}$ [16]. No examples of more than one negative eigenvalue were found. Furthermore, for a given sample the two types of solution typically occur together as a closely related pair, in a sense we clarify below. The extrema of the finite-$N$ TAP free energy are therefore dominated by minima and index-one saddles.

This picture can be further clarified by constructing a fictive free-energy function

$$F_q = \tilde{F} + \frac{\beta^2}{2}(1 - q)\left( \sum_i m_i^2 - Nq \right), \quad (14)$$

where $\tilde{F}$ is a function of the $m_i$ and $q$. It is given by Eq. (1), but with $q$ regarded as an independent variable, unrelated to the $m_i$; i.e., $F_q$ is a function of the $N + 1$ variables $m_1, \ldots, m_N, q$, whereas the original TAP free energy $F$ depends only on the $N$ variables $m_1, \ldots, m_N$ (with $q$ *defined* as $q = (1/N)\sum_i m_i^2$). One readily verifies that the stationarity equations for $F_q$ reproduce the TAP equations: $\partial F_q / \partial m_i = G_i = 0$. However, for these new equations, the quantity $Q \equiv (1/N)\sum_i m_i^2$ is, in general, not equal to the parameter $q$ appearing in the equations. The additional stationarity equation, $0 = \partial F_q / \partial q = (\beta^2/2)(Nq - \sum_i m_i^2)$ forces $Q = q$ at stationary points in the full $(N + 1)$-dimensional space. The free-energy functions $F$ and $F_q$ have, therefore, the same stationary points and the same values at these points. By formally eliminating the variables $m_i$, one can obtain the function $F_q(q)$ as a function of the single variable $q$. Its first derivative is $dF_q/dq = (\beta^2/2)(Nq - \sum_i m_i^2)$, where the $m_i$ are implicit functions of $q$ through the TAP equations. In practice, of course, there will be exponentially many TAP solutions, $\langle N_s \rangle_J \sim \exp[N\Sigma_q]$, for each fixed value of $q$. Their number can be calculated from the same equations, (6) and (7) as before, but with $\lambda = 0$, since $q$ is no longer constrained to equal $(1/N)\sum_i m_i^2$, and $u = 0$ since $f$ is not fixed. The functions $F_q(q)$ and $Q(q)$, however, are *self-averaging* and therefore well-defined, being determined by averages of the appropriate functions of $m$, e.g., $Q = \langle m^2 \rangle$, where the weight function for the averages is the integrand of Eq. (7), with $\lambda = 0 = u$.

We have solved these TAP-like equations numerically for a range of $q$ and determined the corresponding values of $Q$ and $f_q = F_q/N$. An example is shown in Fig. 1. First

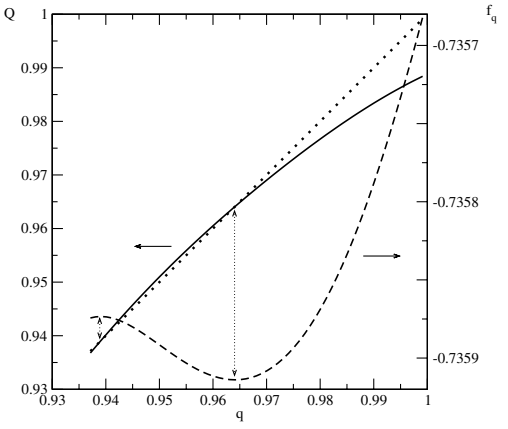


FIG. 1. The functions $Q(q)$ (continuous line) and $f_q(q)$ (dashed line) defined in the text. Physical states occur where $Q(q)$ crosses the dotted line $Q = q$. The data were obtained using $N = 200$ spins at a temperature $T = 0.2$.

a solution of the standard TAP equations (i.e., with $q = (1/N)\sum_i m_i^2$) was found, and solutions for other $q$ values were generated iteratively from the previous value, starting from the TAP solution. The iterative procedure typically fails to converge when $q$ becomes too small. The physical solutions in Fig. 1 are the two points where the function $Q(q)$ intersects the line $Q = q$. They correspond to turning points of the function $f_q(q)$. The solution with the larger $q$ always corresponds to a minimum of $F$, the other solution to a saddle point of index one. The difference vector, $\delta m_i$, between the solutions typically has a large overlap with the eigenvector, $e_i$, of $\mathbf{A}$ with the smallest eigenvalue: $\sum_i e_i \delta m_i / [\sum_i e_i^2 \sum_j (\delta m_j)^2]^{1/2} \approx 0.1$–1, with a typical value around 0.5, for system size $N = 1000$. This shows that one moves from the minimum to the saddle point by moving roughly in the direction of the isolated "soft mode." This agrees with our expectation based on the relation $dm_i/dq = \beta^2 \sum_j X_{ij} m_j$, which follows from the TAP equation. Recall that $v_i = \sum_j X_{ij} m_j$ becomes, for $N \to \infty$, the null eigenfunction of $\mathbf{A}$, so $dm_i/dq \propto v_i$ in this limit. The minimum and the saddle point will coalesce as the small eigenvalue tends to zero with increasing $N$, and the two turning points of $f_q(q)$ will merge to form an inflection point. One can see this formally by differentiating the relation $df_q/dq = (\beta^2/2)[q - (1/N)\sum_i m_i^2]$ to obtain

$$\frac{d^2 f_q}{dq^2} = \frac{\beta^2}{2}\left(1 - \frac{2\beta^2}{N}\sum_{ij} m_i X_{ij} m_j\right) = 0. \qquad (15)$$

It is important to recall that the isolated eigenvalue of order $1/\sqrt{N}$ does not enter the result for the extensive part of the complexity, because the projector term in Eq. (5) responsible for it is $O(1/N)$ and drops out of the complexity at leading order. The upshot is that the BM calculation, in which the projector term is neglected, counts minima and index-one saddles, both with positive sign, since without the projector term the Hessian matrix is, for $N \to \infty$, positive definite. We have shown that including the projector produces one null eigenvalue in the thermodynamic limit, i.e., the prefactor in the calculation of $\langle N_s \rangle_J$ vanishes as required by exact analysis [8,10]. For finite $N$, however, the marginally stable states become pairs of minima and index-one saddles.

This suggests a scenario in which the complexities associated with minima and index-one saddles are extensive and equal, and no other solutions are possible in the limit $N \to \infty$ except the trivial solution, $m_i = 0$ for all $i$. The Morse theorem would then be identically satisfied. Recent numerical studies [19] lend strong support to this scenario: The only solutions found are minima and index-one saddles, which occur in equal numbers and consist of close pairs as predicted here, plus the trivial solution. The distribution of states over free energy is also broadly consistent, for the system sizes studied, with the result for $\langle N_s(f) \rangle$ obtained by BM [1].

We conclude that the BM theory remains the only viable theory of Ising spin-glass complexity. The striking result that TAP solutions become inflection points of the TAP free energy in the thermodynamic limit is a direct consequence (via the appearance of a null Hessian eigenvalue) of the broken supersymmetry in the BM solution and should have important consequences for the dynamics. By contrast, in the much-studied spherical $p$-spin model the supersymmetry remains unbroken [9], the free-energy function has well-defined minima and saddle points, and the dynamics is more straightforward. For Ising spin glasses the thermodynamic limit is rather subtle. For finite $N$, the inflection points are replaced by pairs of minima and index-one saddles. The $N$ dependence of the free-energy barrier separating a minimum from its nearby saddle will play an important role in the finite-$N$ dynamics and merits further study.

We thank Irene Giardina, Andrea Cavagna, Giorgio Parisi, and Tommaso Rizzo for making their unpublished results available. T. A. acknowledges financial support from the European Community's Human Potential Programme under Contract No. HPRN-CT-2002-00307, DYGLAGEMEM.

[1] A. J. Bray and M. A. Moore, J. Phys. C **13**, L469 (1980).
[2] D. J. Thouless, P.W. Anderson, and R. G. Palmer, Philos. Mag. **35**, 593 (1977).
[3] F. Tanaka and S. F. Edwards, J. Phys. F **10**, 2471 (1980).
[4] A. Cavagna, I. Giardina, G. Parisi, and M. Mézard, J. Phys. A **36**, 1175 (2003).
[5] A. Annibale, A. Cavagna, I. Giardina, and G. Parisi, Phys. Rev. E **68**, 061103 (2003).
[6] A. Crisanti, L. Leuzzi, G. Parisi, and T. Rizzo, Phys. Rev. B **68**, 174401 (2003).
[7] A. Annibale, A. Cavagna, I. Giardina, G. Parisi, and E. Trevigne, cond-mat/0307465.
[8] A. Crisanti, L. Leuzzi, G. Parisi, and T. Rizzo, cond-mat/0307543.
[9] A. Crisanti, L. Leuzzi, and T. Rizzo, Eur. Phys. J. B **36**, 129 (2003).
[10] J. Kurchan, J. Phys. A **24**, 4969 (1991).
[11] C. Becchi, A. Stora, and A. Rouet, Commun. Math. Phys. **42**, 127 (1975); I. V. Tyutin, Lebdev Report No. FIAN 39, 1975 (unpublished).
[12] T. Plefka, J. Phys. A **15**, 1971 (1982).
[13] C. De Dominicis, M. Gabay, T. Garel, and H. Orland, J. Phys. (Paris) **41**, 923 (1980).
[14] L. A. Pastur, Russ. Math. Surv. **28**, 1 (1974).
[15] A. J. Bray and M. A. Moore, J. Phys. C **12**, L441 (1979).
[16] Details will be presented elsewhere.
[17] J. C. Owen, J. Phys. C **15**, L1071 (1982).
[18] G. Parisi and T. Rizzo, e-print cond-mat/0401509, have shown that this is a general property of the supersymmetry-breaking solution.
[19] A. Cavagna, I. Giardina, and G. Parisi, e-print cond-mat/0312534.