Mean-field theory of random-site q-state Potts models

1988 J. Phys. A: Math. Gen. 21 791

(http://iopscience.iop.org/0305-4470/21/3/035)

J. Phys. A: Math. Gen. 21 (1988) 791-801. Printed in the UK

Mean-field theory of random-site q-state Potts models

A C D van Enter, J L van Hemmen† and C Pospiech†

Department of Physics, Technion-Israel Institute of Technology, Haifa 32000, Israel

Received 21 July 1987, in final form 22 September 1987

Abstract. A class of random-site mean-field Potts models is introduced and solved exactly. The bifurcation properties of the resulting mean-field equations are analysed in detail. Particular emphasis is put on the relation between the solutions and the underlying symmetries of the model. It turns out that, in contrast to the Ising case, the introduction of randomness in the Mattis-Potts model can change the order of the transition. For $q \leq 6$ the transition becomes second order.

1. Introduction

Recently there has been considerable interest in random-site mean-field models, both for modelling spin glasses and for describing pattern recognition in neural networks [1-15]. In these papers Ising spins are considered and, to avoid the fact that by a Mattis (or gauge) transformation the model becomes equivalent to a ferromagnet, one chooses a number of random variables per site which is larger than one. Here we consider the same kind of model for q-state Potts spins.

It turns out that Mattis transformations do not apply in this case, and even the Mattis-Luttinger choice [16, 17] for the site random variables gives new effects. In contrast to the first-order transition of the ferromagnet, which occurs for $q > 2$, the random model transition becomes first order only for $q > 6$. The method of solution we employ has been developed previously for Ising models with a finite number of random-site variables [1-3, 15], and has also been applied to the random-field problem [18, 19] and anisotropic spin glasses [20].

We describe the model and the main results in the case of the Mattis-Luttinger distribution in § 2. In § 3, we carefully study the interplay of bifurcation and symmetry in a more abstract setting. Exploiting the symmetries of the Potts model under consideration, we then present a comprehensive bifurcation analysis of the mean-field equations. Application of our formalism to neural network models (cf [6, 10, 11]) is then straightforward and will not be given here.

2. The model and its solution

We consider a q-state Potts model with Hamiltonian

$$H_N = -\sum_{i,j=1}^{N} J(i,j)\delta(\sigma_i, \sigma_j). \tag{2.1}$$

† Present address: Sonderforschungbereich 123, Im Neuenheimer Feld 294, D-6900 Heidelberg, Federal Republic of Germany.

Here the σ_i and σ_j denote the q possible states at site i and j. If $J(i,j)>0$, the model is ferromagnetic; if $J(i,j)<0$, it is antiferromagnetic. For spin-glass models, the $J(i,j)$ are random variables.

One can write the Potts Hamiltonian in different ways, using different representations of the Potts spins.

(i) Following Wu [21], we can write

$$\delta(\sigma,\sigma')=\frac{1}{q}[1+(q-1)\boldsymbol{e}^{\sigma}\cdot\boldsymbol{e}^{\sigma'}] \tag{2.2}$$

where the $\boldsymbol{e}^{\sigma}$, $\sigma=0,\ldots,q-1$, are q unit vectors, pointing in q directions, which span a hypertetrahedron in $\mathbb{R}^{q-1}$. This is the representation we will mostly use later in this paper.

(ii) Following Mittag and Stephen [22], we can take for the σ points on the unit circle in the complex plane: $\sigma=1$, $\mathrm{e}^{\mathrm{i}\omega}$, $\mathrm{e}^{\mathrm{i}2\omega},\ldots,\mathrm{e}^{\mathrm{i}(q-1)\omega}$ with $\omega=2\pi/q$ and $\mathrm{e}^{\mathrm{i}\omega}$ the qth root of unity. Then

$$\delta(\sigma,\sigma')=\frac{1}{q}\sum_{k=0}^{q-1}\sigma^k\overline{\sigma'^k}=\frac{1}{q}\left(1+\sum_{k=1}^{q-1}\sigma^k\overline{\sigma'^k}\right). \tag{2.3}$$

The simplest example of a model with site disorder is obtained by the Mattis-Luttinger choice $J(i,j)=N^{-1}\xi_i\xi_j$ where the ξ_i are independent random variables which are ±1 with equal probability. The Hamiltonian then becomes

$$H_N=-\varepsilon N\frac{q-1}{2q}\left(\frac{1}{N}\sum_{i=1}^{N}\xi_i\boldsymbol{e}^{\sigma_i}\right)^2\equiv-\varepsilon N\frac{q-1}{2q}\boldsymbol{m}_N^2 \tag{2.4}$$

if we use (2.2), or becomes

$$\begin{aligned}H_N&=-\frac{\varepsilon N}{2q}\sum_{k=1}^{q-1}\left(\frac{1}{N}\sum_{i=1}^{N}\xi_i\sigma_i^k\right)\left(\frac{1}{N}\sum_{j=1}^{N}\xi_j\overline{\sigma_j^k}\right)\\&=-\frac{\varepsilon N}{2q}\sum_{k=1}^{q-1}\left|\frac{1}{N}\sum_{i=1}^{N}\xi_i\sigma_i^k\right|^2\\&=-\frac{\varepsilon N}{2q}\sum_{k=1}^{q-1}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\xi_i\cos k\omega_i\right)^2+\left(\frac{1}{N}\sum_{i=1}^{N}\xi_i\sin k\omega_i\right)^2\right]\\&=-\frac{\varepsilon N}{2q}(\boldsymbol{m}_{1,N}^2+\boldsymbol{m}_{2,N}^2)\end{aligned} \tag{2.5}$$

if we use (2.3). For $q>2$, the ξ_i cannot be removed through a gauge (Mattis) transformation.

We have dropped constant terms from the Hamiltonian. The quantities $\boldsymbol{m}_N$, $\boldsymbol{m}_{1,N}$ and $\boldsymbol{m}_{2,N}$ are vectors in $\mathbb{R}^{q-1}$. They represent the relevant order parameters. In (2.4) the spin configurations are described by the $\{\sigma_i\}_{i=l,\ldots,N}$ and in (2.5) by the $\{\omega_i\}_{i=1,\ldots,N}$. In both (2.4) and (2.5) ε has the dimension of energy.

To solve the model (in general), we can apply the same techniques which were developed before to treat the case of Ising spins. This means that we have to calculate the appropriate c functions [3]. The different representations coming from (2.2) and (2.3) give rise to correspondingly different c functions. In the Wu representation (2.2) we obtain, for $t \in \mathbb{R}^{q-1}$ and $m_N = N^{-1}\sum_{i=1}^{N} \xi_i e^{\sigma_i} \in \mathbb{R}^{q-1}$,

$$\begin{aligned} c(t) &= \lim_{N\to\infty} \frac{1}{N} \ln \mathbb{E}_\sigma\left[\exp\left(t \cdot \sum_{i=1}^{N} \xi_i e^{\sigma_i}\right)\right] \\ &= \lim_{N\to\infty} \frac{1}{N} \ln \mathbb{E}_\sigma[\exp(t \cdot N m_N)] \\ &= \langle \ln \mathrm{Tr}_\sigma[\exp(\xi t \cdot e^\sigma)]\rangle_\xi. \end{aligned} \tag{2.6}$$

Here Tr_σ is the normalised trace at one site, $\mathbb{E}_\sigma$ is the normalised trace over N sites and $\langle\ldots\rangle_\xi$ denotes an average with respect to ξ. In the Mittag–Stephen representation we obtain, for x and y in $\mathbb{R}^{q-1}$, that

$$\begin{aligned} c(x, y) &= \lim_{N\to\infty} \frac{1}{N} \ln \mathbb{E}_\omega\left\{\exp \sum_{i=1}^{N} \xi_i \left(\sum_{k=1}^{q-1} (x_k \cos k\omega_i + y_k \sin k\omega_i)\right)\right\} \\ &= \left\langle \ln \mathrm{Tr}_\omega\left[\exp\left(\xi \sum_{k=1}^{q-1} (x_k \cos k\omega_i + y_k \sin k\omega_i)\right)\right]\right\rangle. \end{aligned} \tag{2.7}$$

Taking advantage of some general considerations [1–3], one easily derives an expression for the free energy

$$-\beta f(\beta) = \sup_m (F(m) - c^*(m)) = \sup_t (c(t) - F^*(t)). \tag{2.8}$$

Here $c^*(m)$ and $F^*(t)$ are the Legendre transforms of $c(t)$ and $F(m)$, where $F(m)$ is defined below. Both $c(t)$ and $c^*(m)$ are convex. The mean-field (fixed-point) equation which determines the maximum in (2.8) is

$$m = \nabla c(\nabla F(m)). \tag{2.9}$$

In the Wu representation (2.4) we then find

$$F(m) = \beta\varepsilon\left(\frac{q-1}{2q}\right) m^2 \equiv \tfrac{1}{2} k m^2 \tag{2.10}$$

with

$$k = \beta\varepsilon\left(\frac{q-1}{q}\right).$$

Alternatively, in the special case of (2.4) and (2.5), we could also linearise the squares (as was done in [4, 7]) so as to compute $f(\beta)$. The present method is far more general, however.

To obtain the well known ferromagnetic solution, we put $\xi \equiv 1$ in (2.6). Then the mean-field equation is

$$m = \nabla c(km) = \frac{\mathrm{Tr}_\sigma e^\sigma \exp(km \cdot e^\sigma)}{\mathrm{Tr}_\sigma \exp(km \cdot e^\sigma)}. \tag{2.11}$$

The vector $m = 0$ is always a solution, and for small k it is the only one. Suppose now that we have spontaneous magnetisation in the direction $\boldsymbol{e}^0$; that is, $\boldsymbol{m}=\lambda\boldsymbol{e}^0$. Then

$$\lambda\boldsymbol{e}^0=\frac{\boldsymbol{e}^0\exp(k\lambda)+(\Sigma_{j=1}^{q-1}\boldsymbol{e}^j)\exp[-k\lambda/(q-1)]}{\exp(k\lambda)+(q-1)\exp[-k\lambda/(q-1)]}. \tag{2.12}$$

Because $\Sigma_{\sigma=0}^{q-1}\boldsymbol{e}^\sigma=0$, this is equivalent to

$$\lambda\boldsymbol{e}^0=\boldsymbol{e}^0\frac{\exp[qk\lambda/(q-1)]-1}{\exp[qk\lambda/(q-1)]+q-1}. \tag{2.13}$$

Equation (2.13) corresponds to equations (7) and (11) of [22].

We now turn to the Mattis–Luttinger case. Using a trick described in [11], we obtain

$$-\beta f(\beta)=F(\boldsymbol{m})-\boldsymbol{m}\cdot\nabla F(\boldsymbol{m})+c(\nabla F(\boldsymbol{m})) \tag{2.14}$$

where $\boldsymbol{m}$ satisfies the mean-field equation (2.9) and maximises (2.14). From (2.6) and (2.9) we obtain the mean-field equation

$$\boldsymbol{m}=\left\langle\frac{\mathrm{Tr}_\sigma\,\xi\boldsymbol{e}^\sigma\exp(\xi k\boldsymbol{m}\cdot\boldsymbol{e}^\sigma)}{\mathrm{Tr}_\sigma\exp(\xi k\boldsymbol{m}\cdot\boldsymbol{e}^\sigma)}\right\rangle_\xi. \tag{2.15}$$

To consider in what direction we have to take $\boldsymbol{m}$ so as to get a stable mean-field solution, we start by studying the ground states. We call points i for which $\xi_i=1$, blue and points for which $\xi_i=-1$, red. As

$$H_N=-\frac{\varepsilon}{2N}\sum_{i,j=1}^{N}\xi_i\xi_j\delta(\sigma_i,\sigma_j)$$

we see that points of the *same* colour interact ferromagnetically and points with a *different* colour interact antiferromagnetically. Hence we obtain ground states by putting all the blue points in one direction, say $\boldsymbol{e}^{\sigma_1}$, and all the red points in another direction, say $\boldsymbol{e}^{\sigma_2}$. The system is not frustrated and has order parameter $\boldsymbol{m}=\frac{1}{2}(\boldsymbol{e}^{\sigma_1}-\boldsymbol{e}^{\sigma_2})$. The number of ground states equals the number of ordered pairs (σ_1,σ_2), i.e. $\frac{1}{2}q(q-1)$. The mean-field equation which we obtain by assuming $\boldsymbol{m}=\lambda(\boldsymbol{e}^0-\boldsymbol{e}^1)$ is

$$\begin{aligned}\lambda(\boldsymbol{e}^0-\boldsymbol{e}^1)&=\frac{1}{2}\frac{\begin{array}{l}\boldsymbol{e}^0\exp(k\lambda)+\boldsymbol{e}^1\exp(-k\lambda)+(\Sigma_{\sigma=2}^{q-1}\boldsymbol{e}^\sigma-\boldsymbol{e}^0)\\ \qquad\times\exp(-k\lambda)-\boldsymbol{e}^1\exp(k\lambda)-(\Sigma_{\sigma=2}^{q-1}\boldsymbol{e}^\sigma)\end{array}}{\exp(k\lambda)+\exp(-k\lambda)+q-2}\\ &=\frac{1}{2}(\boldsymbol{e}^0-\boldsymbol{e}^1)\frac{2\sinh(k\lambda)}{2\cosh(k\lambda)+q-2}.\end{aligned} \tag{2.16}$$

If we define

$$f(y)=\frac{\sinh y}{2\cosh y+q-2} \tag{2.17}$$

it is straightforward to calculate

$$f'(y)=\frac{(q-2)\cosh y+2}{(2\cosh y+q-2)^2} \tag{2.18}$$

and

$$f''(y)=\sinh y\frac{-2(q-2)\cosh y+q^2-4q-4}{(2\cosh y+q-2)^3}. \tag{2.19}$$

Hence $f''(y)\leqslant 0$ and f is *concave* for all $y\geqslant 0$, if $q\leqslant 6$, but $f''(0^+)>0$ for $q>6$.

Accordingly the transition becomes first order for $q>6$. Note that, due to the evenness of the ξ distribution, the map $\boldsymbol{m} \rightarrow -\boldsymbol{m}$ is a symmetry of the system, in contrast to the ferromagnet. As a curiosity we remark that a similar calculation for $\boldsymbol{m} = \lambda \boldsymbol{e}^0$ (which is an unstable direction) shows that the first-order behaviour in this direction already occurs if $q>4$.

Combining (2.4) and (2.10) we obtain the energy $u(\beta) = -\frac{1}{2}k\boldsymbol{m}^2$ where $\boldsymbol{m}$ satisfies (2.9) and maximises (2.8). Differentiating $u(\beta)$ with respect to the temperature we find the specific heat. This is most easily done by implicitly differentiating $\boldsymbol{m}$ in (2.9) with respect to β. In a similar vein one obtains the susceptibility. For details, see § 9 of [3].

With the methods of [1-3, 15] it is also straightforward to consider more general distributions of the ξ or cases where there are two or more random variables per site.

3. Bifurcation analysis

The conclusion of the previous section was obtained through the ansatz $\boldsymbol{m} = \lambda(\boldsymbol{e}^0 - \boldsymbol{e}^1)$. To finish the proof we have to classify *all* solutions which bifurcate from zero, determine their stability, and show that none of them contradicts the above conclusion. This is done in the present section. Here we extend the Wu representation and study the properties of a somewhat more general class of c functions, which satisfy both the q-fold Potts permutation symmetry and the reflection symmetry of the ξ distribution. We start by developing a general theory. At the end of the section we specialise to Potts models.

3.1. General theory

Let $c: \mathbb{R}^q \mapsto \mathbb{R}$ be a twice continuously differentiable and even† function, which is symmetric under permutation of the arguments‡. We denote vectors in $\mathbb{R}^q$ by $\vec{a}$, instead of $\boldsymbol{a}$ for vectors in $\mathbb{R}^{q-1}$.

Formally,

$$c(g\vec{x}) = c(\vec{x}) \qquad \text{for all } \vec{x} \in \mathbb{R}^q,\ g \in \mathrm{G} := C_2 \times S_q. \tag{3.1}$$

The group G is the direct product $C_2 \times S_q$ of the permutation group S_q and the reflections $C_2 = \{-1, +1\}$. Then the derivative ∇c of c is *equivariant* with respect to the group G, that is

$$\nabla c(g\vec{x}) = g\nabla c(\vec{x}) \qquad \text{for all } \vec{x} \in \mathbb{R}^q,\ g \in \mathrm{G} = C_2 \times S_q. \tag{3.2}$$

One can see this by taking the derivative of (3.1). We obtain

$$\nabla c(\vec{x}) \cdot y = \nabla c(g\vec{x}) \cdot gy = g^{-1}\nabla c(g\vec{x}) \cdot y \qquad \text{for all } y \in \mathbb{R}^q,\ g \in \mathrm{G}. \tag{3.3}$$

Multiplying (3.3) with $g \in \mathrm{G}$ from the left yields (3.2). The property (3.2) allows us to restrict ourselves to certain subspaces, which we will now introduce.

Definition. Let $\mathrm{H} \subseteq \mathrm{G}$ be a subgroup of G. The *fixed space* $(\mathbb{R}^q)^{\mathrm{H}}$ of H is defined to be the subspace of vectors in $\mathbb{R}^q$ which are invariant under $\mathrm{H} \subseteq \mathrm{G}$, i.e.

$$(\mathbb{R}^q)^{\mathrm{H}} := \{\vec{x} \in \mathbb{R}^q \mid hx = x \text{ for all } h \in \mathrm{H}\}. \tag{3.4}$$

† This holds because of the evenness of the distribution of the ξ.
‡ This follows from the Potts symmetry.

For all subgroups $\mathrm{H} \subseteq \mathrm{G}$ we have

$$\nabla c((\mathbb{R}^q)^{\mathrm{H}}) \subseteq (\mathbb{R}^q)^{\mathrm{H}}. \tag{3.5}$$

Namely, let $h \in \mathrm{H}$ and $\vec{x} \in (\mathbb{R}^q)^{\mathrm{H}}$. By (3.2) we then have

$$\nabla c(\vec{x}) = \nabla c(h\vec{x}) = h\nabla c(\vec{x}). \tag{3.6}$$

We will denote the restriction of ∇c to $(\mathbb{R}^q)^{\mathrm{H}}$ by ∇c^{H}.

If $\vec{m}$ solves the fixed-point equation (cf (2.9)),

$$\vec{m} = \nabla c(k\vec{m}) \tag{3.7}$$

then $g\vec{m}$ also solves (3.7). If $\vec{m} \in (\mathbb{R}^q)^{\mathrm{H}}$, $g\vec{m}$ stays fixed under the conjugate subgroup $g\mathrm{H}g^{-1}$.

We can also draw some conclusions on the second derivative of c by the equivariance property (3.2). In future applications, a bifurcating branch is stable if all the eigenvalues of the second derivative of c are less than one [11].

Definition. Let G be a group operating on a real vectorspace V by orthogonal matrices. Then V is called *irreducible*, if there is no proper subspace $W \subseteq V$ satisfying $gW \subseteq W$ for all $g \in \mathrm{G}$.

Every real vectorspace V with a group G acting on it as above has a decomposition

$$V = \bigoplus_{i=1}^{n} W_i \tag{3.8}$$

in irreducible subspaces W_i.

Indeed, if V is not irreducible, we can find a proper subspace W with orthogonal complement $W^\perp$, such that

$$gW \subseteq W \qquad gW^\perp \subseteq W^\perp \qquad \text{for all } g \in \mathrm{G}.$$

Since W and $W^\perp$ have lower dimension than V, we can proceed by induction.

Theorem 3.1. Let $\vec{x}_0 \in \mathbb{R}^q$, and let $\mathrm{H} \subseteq \mathrm{G}$ be a subgroup of G, such that $\vec{x}_0 \in (\mathbb{R}^q)^{\mathrm{H}}$. Suppose now that the W_i in the decomposition of $\mathbb{R}^q$ into irreducible subspaces (with respect to H)

$$\mathbb{R}^q = \bigoplus_{i=1}^{n} W_i \tag{3.9}$$

are pairwise non-equivalent†. Then each of the W_i is contained in some eigenspace of the second derivative $L := \mathrm{D}^2 c(\vec{x}_0)$ of c at $\vec{x}_0$ (L is viewed as a matrix). In particular, the number of real *distinct* eigenvalues of L is less than or equal to n; see (3.9).

Proof. The second derivative $L = \mathrm{D}^2 c$ of c at $\vec{x}_0$ is symmetric and equivariant with respect to H. This can be seen by differentiating (3.2). Now let Π_i be the projection onto W_i parallel to the other W_j‡ and look at the composite map

$$\Phi_{i,j}: \quad W_i \overset{\subseteq}{\hookrightarrow} \mathbb{R}^q \overset{L}{\mapsto} \mathbb{R}^q \overset{\Pi_j}{\mapsto} W_j. \tag{3.10}$$

† That is, there is no H-equivariant isomorphism from one such space to another.

‡ This projection is automatically equivariant.

The kernel and image of this equivariant map are H invariant. So they are either zero or the whole (irreducible) space W_i or W_j, respectively. Thus $\Phi_{i,j}$ is either zero or an isomorphism; in the latter case $i = j$ since the W_i are pairwise non-equivalent. As $\Phi_{i,i}$ is symmetric, it must have a real eigenvalue μ the eigenspace of which again is H invariant. Thus, this eigenspace must be W_i by irreducibility. Therefore we have for all $w \in W_i$:

$$\mu\omega = \Phi_{i,i}\omega = L\omega. \tag{3.11}$$

This proves the first assertion; the second is a direct consequence of the first.

Remark. The theorem can be generalised [23], but it is sufficient for our purposes in this form.

3.2. Dimensions of fixed spaces and the corresponding groups

Given some subgroup $\mathrm{H} \subseteq \mathrm{G}$, we consider the action of H on the set

$$M := \{\pm\vec{e}_\sigma \,|\, 0 \leqslant \sigma \leqslant q-1\} \tag{3.12}$$

where the $\vec{e}_\sigma$ denote the standard basis of $\mathbb{R}^q$. $H(\pm\vec{e}_\sigma)$ is the *orbit* of $\vec{e}_\sigma$ under H. If an orbit is such that $H(\vec{e}_\sigma) = H(-\vec{e}_\sigma)$, it is called *irrelevant.* This is motivated by the following proposition.

Proposition. The dimension of $(\mathbb{R}^q)^{\mathrm{H}}$ is just half the number of relevant orbits. A basis of $(\mathbb{R}^q)^{\mathrm{H}}$ is given by

$$\left\{ \vec{b}(\vec{e}_\sigma) := \frac{1}{|H|} \sum_{h \in \mathrm{H}} h\vec{e}_\sigma \,\middle|\, \vec{e}_\sigma \in I \right\}$$

where $I \subseteq M$ will be defined in the proof.

Proof. First note that $\vec{b}(\vec{e}_\sigma) = 0$ if the orbit $H(\vec{e}_\sigma)$ is irrelevant; for then $H(\vec{e}_\sigma) = H(-\vec{e}_\sigma)$ and all terms in the sum are cancelled by their opposites. For relevant orbits, $H(\vec{e}_\sigma)$ and $H(-\vec{e}_\sigma)$ are disjoint, so relevant orbits always come in pairs. To construct I, we choose a representative from each orbit pair $H(\vec{e}_\sigma) \cup H(-\vec{e}_\sigma)$ of relevant orbits. By construction, any two $\vec{b}(\vec{e}_\sigma)$ are orthogonal and their number equals half the number of relevant orbits.

If the $\vec{b}(\vec{e}_\sigma)$ did not span $(\mathbb{R}^q)^{\mathrm{H}}$, then we could find some $0 \neq \vec{x} = (x_0, \ldots, x_{q-1}) \in (\mathbb{R}^q)^{\mathrm{H}}$ orthogonal to all $\vec{b}(\vec{e}_\sigma)$. So we get for all $0 \leqslant \sigma \leqslant q-1$ that

$$0 = \vec{x} \cdot \vec{b}(\vec{e}_\sigma) = \frac{1}{|H|} \sum_{h \in \mathrm{H}} \vec{x} \cdot h\vec{e}_\sigma = \frac{1}{|H|} \sum_{h \in \mathrm{H}} h^{-1}\vec{x} \cdot \vec{e}_\sigma = \frac{1}{|H|} \sum_{h \in \mathrm{H}} \vec{x} \cdot \vec{e}_\sigma = x_\sigma .$$

Hence $\vec{x} = 0$, which contradicts our assumption.

Next we are interested in the number of different fixed spaces $(\mathbb{R}^q)^{\mathrm{H}}$. As shown in the following two theorems it turns out that we can find them without writing down all the subgroups H explicitly (this is possible but does not give much insight).

Theorem 3.2. (a) If H contains only permutations, then all orbits are subsets of either

$$M_+ := \{\vec{e}_\sigma \,|\, 0 \leqslant \sigma \leqslant q-1\} \qquad \text{or} \qquad M_- := \{-\vec{e}_\sigma \,|\, 0 \leqslant \sigma \leqslant q-1\}.$$

(b) If H contains some element $(-1, \pi) \in C_2 \times S_q$, then in each orbit half of its elements are in M_+ and half are in M_-.

Proof. Part (a) is obvious. Part (b) is reduced to (a). In case (b) we have for every $0 \leq \sigma \leq q-1$ that

$$H(\vec{e}_\sigma) = \tilde{H}(\vec{e}_\sigma) \cup \tilde{H}(-\vec{e}_{\pi\sigma})$$

where $\tilde{H} := H \cap [\{+1\} \times S_q]$ is of the type described in (a) and both $\tilde{H}(\vec{e}_\sigma) \subseteq M_+$ and $\tilde{H}(-\vec{e}_{\pi i}) \subseteq M_-$ are disjoint orbits, which are of equal size since $(-1, \pi)$ maps the one isomorphically onto the other.

Theorem 3.3. For every decomposition

$$M = \bigcup_{\alpha=1}^{\kappa} M_\alpha$$

of M into disjoint subsets of either type (a) or type (b) there is a group H, namely

$$\mathrm{H} := \{g \in \mathrm{G} \mid gM_\alpha \subseteq M_\alpha \text{ for all } 1 \leq \alpha \leq \kappa\}$$

that has precisely the orbits M_α.

Proof. By construction every orbit $H(\vec{e}_\sigma)$ is already contained in some M_α. It remains to construct, for any two $\pm\vec{e}_{\sigma_1}, \pm\vec{e}_{\sigma_2} \in M_\alpha$, some $h \in \mathrm{H}$ that maps the one onto the other. If both elements have the same sign, the permutation of σ_1 and σ_2 will do. If not, all M_β are of type (b). Then for each β arrange the indices of the $\pm\vec{e}_\sigma \in M_\beta$ into a cyclic permutation γ_β, such that the corresponding elements in M_β have alternating signs and such that σ_1 is taken to σ_2 by γ_β. The desired element then is given by $h = (-1, \Pi_\beta \gamma_\beta) \in C_2 \times S_q$.

3.3. Applications to the Potts glass

To link the results in this section to those of § 2 we note that the $\boldsymbol{e}^\sigma$ are connected to the $\vec{e}_\sigma$ by

$$\boldsymbol{e}^\sigma = \text{constant}(\vec{e}_\sigma - \vec{b}) \qquad \sigma = 0, \ldots, q-1 \qquad \vec{b} = \frac{1}{q} \sum_{\sigma=0}^{q-1} \vec{e}_\sigma. \qquad (3.13)$$

The hypertetrahedron which is spanned by the endpoints of the q unit vectors $\vec{e}_\sigma$ lies in a hyperplane of dimension $(q-1)$ orthogonal to the vector $\vec{b}$ (see figure 1).

The results we present here have been obtained mainly by straightforward calculations, which do not provide much insight. Therefore we decided to state some results without proof. The mean-field equation (3.7) with c function given by

$$c(\vec{t}) := \langle \ln[\mathrm{Tr}_\sigma \exp(\xi \vec{t} \cdot \vec{e}_\sigma)] \rangle_\xi = \left\langle \ln\left(\frac{1}{q} \sum_{\sigma=0}^{q-1} \exp(\xi \vec{t} \cdot \vec{e}_\sigma) \right) \right\rangle_\xi$$

satisfies the following lemma.

Lemma. The image of ∇c, and therefore every solution of the mean-field equation (3.7), lies in the orthogonal complement $(\vec{b})^\perp$ of

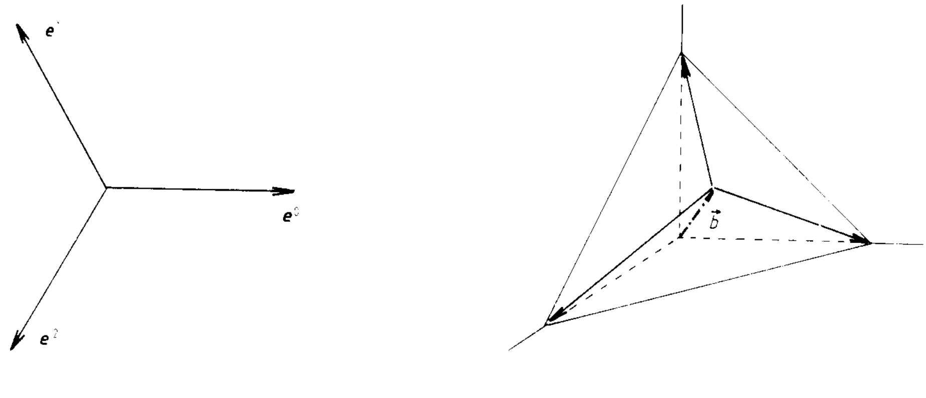
$$\vec{b} = \mathrm{Tr}_\sigma \, \vec{e}_\sigma \left(= \frac{1}{q} \sum_{\sigma=0}^{q-1} \vec{e}_\sigma \right).$$

The linearisation† $L = L(k, \vec{m})$ of ∇c at $(k, \vec{m})$ satisfies $L(\vec{b}) = 0$ and

$$L(\langle \vec{b} \rangle^\perp) \subseteq \langle \vec{b} \rangle^\perp.$$

This shows that $\vec{b}$ is a dummy direction, and that we indeed have reduced everything to the Wu representation.

† Up to a constant this is the second derivative of c.



(a) (b)

Figure 1. (*a*) The Wu representation (2.2) in the case $q=3$. The sum of the unit vectors e^{σ} is zero. (*b*) The very same representation is obtained through the prescription (3.13). The q endpoints of the standard basis vectors $\vec{e}_{\sigma}$, $0 \leq \sigma \leq q-1$, span a simplex, whose centre of mass is the endpoint of $\vec{b}$. Up to normalisation, e^{σ} equals $\vec{e}_{\sigma} - \vec{b}$, as is easily seen in the figure for $q=3$.

The vector $\vec{m} = \vec{0}$ is always a solution of (3.7). The linearisation $L(k, \vec{0})$ has the following eigenvalues:

Eigenvalue	Eigenspace	Multiplicity
0	$\langle \vec{b} \rangle$	1
k/q	$\langle \vec{b} \rangle^{\perp}$	$q-1$

Certain subgroups H determine a direction for some bifurcating branch. They will be examined in the sequel.

If H contains only pure permutations (see theorem 3.2(a)), we consider the case where M_+ contains two (relevant) orbits M_1 and M_2 of length q_1 and q_2, respectively. The corresponding fixed space $(\mathbb{R}^q)^{\mathrm{H}}$ has a basis $\{\vec{b}_1, \vec{b}_2\}$, where $\vec{b}_i = q_i^{-1} \Sigma_{\sigma \in M_i} \vec{e}_{\sigma}$. This fixed space contains the vector $\vec{b}$, and hence we can restrict ourselves to the subspace spanned by the vector $\vec{b}_1 - \vec{b}_2$, which is orthogonal to $\vec{b}$.

In the case of theorem 3.2(b) things are different. We get a one-dimensional fixed space if M has two† relevant orbits of length $2q_1$, and one irrelevant orbit containing the rest of M. Considering the orbits under the 'pure permutations' of H, we see that M_+ contains three of those, called M_0, M_1 and M_2, of length q_0, q_1 and $q_2 = q_1$, respectively. M_0 is contained in the irrelevant orbit while the other two are part of the relevant orbits. Having defined $\vec{b}_i$, $i = 0, 1, 2$, in analogy to the previous case a spanning vector for the fixed space is given by $\vec{b}_1 - \vec{b}_2$.

Theorem 3.4. For the $\vec{b}_i$ defined above we have

$$\nabla c^{\mathrm{H}}\left(x \frac{\vec{b}_1 - \vec{b}_2}{\|\vec{b}_1 - \vec{b}_2\|} \right) = f(x) \frac{\vec{b}_1 - \vec{b}_2}{\|\vec{b}_1 - \vec{b}_2\|} \tag{3.14}$$

where f is given by

$$f(x) = \left\langle \xi \frac{p_1 \gamma \bar{v} \sinh(x \xi \bar{v})}{p_1 \cosh(x \xi \bar{v}) + p_2} \right\rangle_{\xi}. \tag{3.15}$$

† Remember that relevant orbits always come in pairs!

The values of the constants are given by

Case	p_1	p_2	γ	$\vec{v}$
Theorem 3.2(a)	$2q_1q_2$	$q_1^2+q_2^2$	$\frac{1}{2}$	$\|\vec{b}_1-\vec{b}_2\|$
Theorem 3.2(b)	$2q_1$	q_0	1	$\frac{1}{2}\|\vec{b}_1-\vec{b}_2\|$

Corollary. f is strictly monotone and thus *all* non-trivial solutions of (3.7) lying in $(\mathbb{R}^q)^{\mathrm{H}}$ form smooth curves, which approach the ground-state solutions (of (3.14)) $x=\pm\frac{1}{2}\|\vec{b}_1-\vec{b}_2\|\langle|\xi|\rangle_\xi$ for large k†. The bifurcation is supercritical‡ if and only if

$$\begin{aligned} &2-\sqrt{3}\leq q_1/q_2\leq 2+\sqrt{3} && \text{case of theorem 3.2(a)} \\ &q\leq 6q_1 && \text{case of theorem 3.2(b).} \end{aligned} \tag{3.16}$$

The following tables provide the necessary information to determine the stability of the different branches.

In the case of theorem 3.2(a) the matrix $L(k,x)$ has the following spectrum:

Eigenvalue	Eigenspace	Multiplicity
$\lambda_1(k,x)=0$	$\langle\vec{b}\rangle$	1
$\lambda_2(k,x)=kf'(kx)=kq\left\langle\xi^2\dfrac{2q_1q_2+(q_1^2+q_2^2)\cosh(\xi xk\beta)}{[2q_1q_2\cosh(\xi xk\beta)+q_1^2+q_2^2]^2}\right\rangle_\xi$	$\langle\vec{b}_1-\vec{b}_2\rangle$	1
$\lambda_3(k,x)=k\left\langle\xi^2\dfrac{q_1+q_2\cosh(\xi xk\beta)}{2q_1q_2\cosh(\xi xk\beta)+q_1^2+q_2^2}\right\rangle_\xi$	$\langle M_1-\vec{b}_1\rangle$	q_1-1
$\lambda_4(k,x)=k\left\langle\xi^2\dfrac{q_2-q_1\cosh(\xi xk\beta)}{2q_1q_2\cosh(\xi xk\beta)+q_1^2+q_2^2}\right\rangle_\xi$	$\langle M_2-\vec{b}_2\rangle$	q_2-1

Remark. Except for λ_1, all eigenvalues are large for k large. Hence all these branches are unstable at *low* temperatures.

In the case of theorem 3.2(b) $L=L(k,x)$ has the following spectrum:

Eigenvalue	Eigenspace	Multiplicity
$\lambda_1(k,x)=0$	$\langle\vec{b}\rangle$	1
$\lambda_2(k,x)=kf'(kx)=k\left\langle\xi^2\dfrac{2q_1+q_0\cosh(\xi xk\beta)}{[2q_1\cosh(\xi xk\beta)+q_0]^2}\right\rangle_\xi$	$\langle\vec{b}_1-\vec{b}_2\rangle$	1
$\lambda_3(k,x)=k\left\langle\xi^2\dfrac{q\cosh(\xi xk\beta)}{[2q_1\cosh(\xi xk\beta)+q_0]^2}\right\rangle_\xi$	$\langle\vec{b}_0-\vec{b}\rangle$	$\begin{cases}0 & \text{if } 2q_1=q\\ 1 & \text{if } 2q_1\neq q\end{cases}$
$\lambda_4(k,x)=k\left\langle\xi^2\dfrac{1}{2q_1\cosh(\xi xk\beta)+q_0}\right\rangle_\xi$	$\langle M_0-\vec{b}_0\rangle$	$\max(0,q_0-1)$
$\lambda_5(k,x)=k\left\langle\xi^2\dfrac{\cosh(\xi xk\beta)}{2q_1\cosh(\xi xk\beta)+q_0}\right\rangle_\xi$	$\langle(M_1-\vec{b}_1)\cup(M_2-\vec{b}_2)\rangle$	$2(q_1-1)$

† See § 2 for a special case.

‡ This means that the bifurcating branches 'bend to the right'.

Remark. All eigenvalues except for λ_5 tend to zero for k large (low temperature). In particular all bifurcating branches of the type described in theorem 3.2(b) are unstable for k large, unless $q_1 = 1$, in which case λ_5 does not exist.

With (3.16) we recover all the results of § 2. Note that in the present section we have found the bifurcation properties not only of the *stable* branches, but also of the *unstable* ones.

4. Conclusion

We have investigated a class of random-site mean-field Potts models. The solutions are described by mean-field equations which contain both the Potts (permutation) symmetry and a reflection symmetry. We have investigated how this symmetry influences the number and the stability of the different solutions. In particular, it turns out that the value of the Potts parameter above which there occurs a first-order transition is different from the ferromagnetic case ($q > 6$ instead of $q > 2$).

Acknowledgments

This work has been supported by SFB 123 of the Deutsche Forschungsgemeinschaft. ACDvE is supported by the Lady Davis Foundation and also thanks the University of Groningen for its hospitality during a visit when an earlier draft of this paper was written.

References

[1] van Hemmen J L 1982 *Phys. Rev. Lett.* **49** 409
[2] van Hemmen J L, van Enter A C D and Canisius J 1983 *Z. Phys.* B **50** 311
[3] van Hemmen J L 1983 *Heidelberg Colloq. on Spin Glasses* ed J L van Hemmen and I Morgenstern (Berlin: Springer) pp 203–33, appendix
[4] Provost J P and Vallée G 1983 *Phys. Rev. Lett.* **50** 598
[5] Benamira A, Provost J P and Vallée G 1985 *J. Physique* **46** 1269
[6] Amit D, Gutfreund H and Sompolinsky H 1985 *Phys. Rev.* A **32** 1007
[7] Choy T and Sherrington D 1984 *J. Phys. C: Solid State Phys.* **17** 739
[8] van Hemmen J L 1986 *J. Phys. C: Solid State Phys.* **19** L379
[9] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554; 1984 *Proc. Natl Acad. Sci. USA* **81** 3088
[10] van Hemmen J L and Kühn R 1986 *Phys. Rev. Lett.* **57** 913
[11] van Hemmen J L 1986 *Phys. Rev.* A **34** 3435
[12] Grensing D and Kühn R 1986 *J. Phys. A: Math. Gen.* **19** L1153
[13] Kühn R, Grensing D and Huber A 1986 *Z. Phys.* B **63** 445
[14] Jedrzejewski J and Komoda A 1986 *Z. Phys.* B **63** 247
[15] van Hemmen J L, Grensing D, Huber A and Kühn R 1986 *Z. Phys.* B **65** 53
[16] Mattis D C 1976 *Phys. Lett.* **56A** 421
[17] Luttinger J 1976 *Phys. Rev. Lett.* **37** 778
[18] Salinas S R and Wreszinski W F 1985 *J. Stat. Phys.* **41** 299
[19] Fernando Perez J, Pontin L F and Baéta Segundo J A 1986 *Phys. Lett.* **116A** 287
[20] de Almeida J R L and Brady Moreira F G 1986 *Z. Phys.* B **63** 365
[21] Wu F Y 1982 *Rev. Mod. Phys.* **54** 235
[22] Mittag S and Stephen L 1974 *J. Phys. A: Math. Gen.* **7** L109
[23] Pospiech C 1987 unpublished