Progress of Theoretical Physics, Vol. 126, No. 5, November 2011

First-Order Transition in Potts Models with "Invisible" States

— *Rigorous Proofs* —

Aernout C. D. van Enter,[*] Giulio Iacobelli[**] and Siamak Taati[***],[†]

Johann Bernoulli Institute, Rijksuniversiteit Groningen, Nijenborgh 9, 9747 AG Groningen, The Netherlands

(Received June 23, 2011; Revised September 12, 2011)

In some recent papers by Tamura, Tanaka and Kawashima [R. Tamura, S. Tanaka and N. Kawashima, Prog. Theor. Phys. **124** (2010), 381; S. Tanaka, R. Tamura and N. Kawashima, J. Phys.: Conf. Ser. **297** (2011), 012022; S. Tanaka and R. Tamura, arXiv:1012.4254] a class of Potts models with "invisible" states was introduced, for which the authors argued, by numerical arguments and by a mean-field analysis, that a first-order transition occurs. Here we show that the existence of this first-order transition can be proven rigorously, by relatively minor adaptations of existing proofs for ordinary Potts models. In our argument, we present a random-cluster representation for the model, which might also be of interest for general parameter values of the model.

Subject Index: 012, 040, 042, 370

§1. Introduction

In Refs. 1) and 2), the authors introduce a variant of the Potts model, in which along with the q ordinary "visible" colours (the ordinary Potts states), r "invisible" colours are possible. A ferromagnetic nearest-neighbour interaction acts exclusively among the visible colours. The invisible colours, on the other hand, are neutral, and have no interaction with their neighbours, regardless of the state of the neighbours.

Although the number of ground states and low-temperature states equals q, and there is, at low temperatures, spontaneous breaking of the q-fold permutation symmetry just as in the standard q-state Potts model, the transition for low q ($q = 2, 3, 4$) and high r is different from the second-order transition of the ordinary two-dimensional q-state Potts model. In fact a first-order transition in the temperature-parameter appears.

The occurrence of such a first-order transition contradicts a simple form of universality which would predict that all systems with the same broken symmetry in the same dimension with short-range interactions have the same type of transition. However, such a universality property is known to be too strong to be true. The question of first-order versus second-order is *not* a universal question. Some counterexamples illustrating this point are the two-dimensional 3-state Kac-Potts model,[3] in which a

[*] E-mail: A.C.D.van.Enter@rug.nl

[**] E-mail: dajegiu@gmail.com

[***] E-mail: siamak.taati@gmail.com

[†] Present address: Department of Mathematics, Utrecht University, P.O. Box 80.010, 3508 TA Utrecht, The Netherlands.

first-order transition occurs in presence of a broken 3-fold rotation (= permutation) symmetry, or the three-dimensional versions of the nonlinear $O(n)$-models treated in Refs. 4)–8), in which a first-order transition occurs in presence of a broken continuous rotation symmetry. In both cases, the same type of symmetry breaking is also known to be possible with a second-order transition. This occurs for the standard nearest-neighbour 3-state two-dimensional Potts model, or for the standard three-dimensional classical Heisenberg or XY models respectively. For another closely related system, which includes the model with invisible colours, and in which both first and second-order transitions occur but without a change in the broken symmetry, see Ref. 9).

We shall argue that the model with many invisible states has a first-order transition for the same reason the high-q Potts model has a first-order transition. In fact, as far as the order of the transition is concerned, the Potts model with q visible states and r invisible states, when q or r or both are large, should be compared with the standard $(q + r)$-state Potts model rather than the q-state Potts model, as might be suggested by the nature of the broken symmetry. This is precisely the feature that leads to the distinctive behaviour of the (q, r)-model; while the broken symmetry is determined by the number of "visible" colours (states), the order of the transition depends on the total number of colours (states). In particular, for every q for which the original model has a second-order transition, one can make the transition first-order by increasing r, without changing the nature of the symmetry breaking.

At the transition temperature, low-energy (*ordered*) phases coexist along with a high-entropy (*disordered*) phase, and the different phases are separated by "free-energy barriers". In each of the q low-energy phases, most sites have the same (visible) colour. In the high-entropy phase, the colours of different sites are almost independent of one another. The coexistence of the low-energy and high-entropy phases can be proven by a form of a Peierls-type free-energy-contour argument. For the standard Potts model, by now there exists a variety of such proofs,[10)–13)] whether by Reflection Positivity and Chessboard Estimates, or by a Pirogov-Sinai argument, either within a spin description or in a random-cluster version. Those proofs can be adapted to include the model described above.

The ordered and disordered regions in the standard Potts model may roughly be identified locally by looking at bonds connecting neighbouring sites on the lattice. In an *ordered* bond, the two sites have the same colour, which globally identifies the phase. In a *disordered* bond, the two sites take their colours freely and independently of each other; for a large number of colours, this essentially means that the two sites of the bond have different colours. The "free-energy barriers", which will be in the form of *contours*, consist of sites that are attached to both ordered and disordered bonds, and have most of the time neither all neighbouring sites different, nor all neighbours equal. Equivalently, one can choose the contours to consist of bonds whose sites take their colours independently from each other but not freely: the colour of at least one of their sites is determined "globally" (i.e., belongs to an ordered region).

At the transition temperature, the free energy of an ordered bond (which is purely energetic) approximately equals the free energy of a disordered bond (which

is almost purely entropic). If a bond is neither ordered nor disordered (which, as described above, depends on the sites of the bond as well as their neighbourhood), its free energy is higher by an amount proportional to the logarithm of the number of colours. When the number of colours is large, this means that such bonds become unlikely, and tend to occur in far-apart small groups. This ensures the stability of each of the ordered and disordered phases.

In the model with invisible colours, the sites of an ordered bond are required to take the same *visible* colour. However, the colour of the sites in an ordered bond is global within the encompassing ordered phase, and hence does not contribute to the free energy of the bond. On the other hand, the excess free energy of a bond that is neither ordered nor disordered remains of the order of the logarithm of the total number of colours (visible + invisible). This enables us to increase the probability cost of a contour (by increasing r) without changing the degeneracy of the ordered phase (given by q).

The heuristic argument above suggests why the Potts model with q visible colours and r invisible colours, as long as the order of the phase transition is concerned, should be compared with the standard $(q+r)$-Potts model. In fact, in the model with invisible colours, the probability of a contour of length l is proportional to $(q+r)^{-l}$, exactly as it is for the $(q+r)$-Potts model. The new aspect of the model introduced by Tamura, Tanaka and Kawashima, as compared to the standard Potts model, is that for any q, the probability of a contour can be made small, by increasing the number of invisible colours r. We shall indicate how this reasoning can be made precise, following the approach of Refs. 12), 14), and 15).

It is enough to consider only the two-dimensional version, but this is not essential, and the arguments directly generalize to higher dimensions. As the presence of a first-order transition in higher-dimensional Potts models is less surprising, the main interest seems to be in two dimensions.

As a further comment we mention that the term "invisible" is actually a bit of a misnomer, as at high temperatures the density of "invisible" colours is higher than those of the "visible" ones when $r \geq q$. Thus most of the colours which appear would be the "invisible" ones.

§2. Main result

At each site of the d-dimensional cubic lattice, there is a discrete-valued spin variable which can take one out of $q+r$ colours, q of which "visible" and r "invisible". The (q,r)-model then is defined by the following (formal) Hamiltonian:

$$H(\sigma) = -\sum_{\langle i,j\rangle} \delta(\sigma_i,\sigma_j) \sum_{\alpha=1}^{q} \delta(\sigma_i,\alpha)\delta(\sigma_j,\alpha),$$

where δ denotes Kronecker's delta function. The first sum is over all pairs of nearest-neighbour sites in the lattice. Now we have the following result.

Theorem 1. *For $q+r$ large enough the above model undergoes a first-order transition in temperature. At the transition temperature q ordered extremal Gibbs states coexist*

with a disordered extremal Gibbs state.

Proof sketch. There are various ways in which one may adapt existing proofs. For example, the proof originally due to Kotecký and Shlosman,[11] later also treated in Refs. 16) and 17), could be adapted by observing that

- Our model has a C-potential (in Georgii's terms), so reflection positivity holds.
- An ordered bond now will be a bond whose two sites have the same *visible* colour.
- The "restricted ensemble" for the disordered phase is formed by all configurations having disordered bonds only, which has an approximate entropy density $\ln(q+r)$, when $q+r$ is large.

Then the arguments used in section 19.3 of Ref. 16) or in Ref. 17), using chessboard estimates to provide a contour estimate, apply.

In this paper, we will sketch, in some more detail, how to adapt an alternative proof of the first-order phase transition in the ordinary q-state Potts model to the model with invisible states. We elaborate further the heuristic arguments given in the introduction, and discuss one way it can be made into a rigorous proof, which is based on the Fortuin-Kasteleyn random-cluster representation of the model and the Pirogov-Sinai machinery. Such a proof for the standard Potts model was first derived in Ref. 12), and later treated, for example, in Ref. 14).

The distinction between order and disorder is not yet fully clear, especially along the boundaries between two regions. However, let us assume, for the moment, that we have a clear-cut and unambiguous way of identifying the ordered and disordered regions and the boundaries separating them.

Each bond in an ordered region has energy -1. The entropy per bond of an ordered region, however, is negligible (at least when the region is large), for a single choice of colour is shared among all the sites in the region. Therefore, the free energy per bond of an ordered region is approximately -1.

In a disordered region, the entropy per bond is $\frac{1}{d}\log(q+r)$, while the average energy per bond is $-\frac{q}{(q+r)^2}$. When $q+r$ is large, the term $-\frac{q}{(q+r)^2}$ becomes negligible compared to temperature times $\frac{1}{d}\log(q+r)$, and hence the free energy per bond of a disordered region can be approximated by $-\frac{1}{\beta d}\log(q+r)$, where β is the inverse temperature.

Let us now consider a bond that is neither in an ordered region, nor in a disordered region. Such a bond is either on the boundary between two ordered regions, or on the boundary between an ordered and a disordered region. In the first case, the free energy of such a bond is $-\frac{1}{q}$, in the second case $-\frac{1}{2\beta d}\log(q+r)$.

In the high-temperature regime $\beta \ll \frac{1}{d}\log(q+r)$, the disordered bonds have the lowest free energy, and hence are the most likely. In this case, with high probability, the sites form a large "sea" of disorder with small "islands" of order. Therefore, the system has a unique stable phase, which is disordered. In the low-temperature regime $\beta \gg \frac{1}{d}\log(q+r)$, the minimal free energy is achieved by the ordered bonds. Hence, with high probability, the configuration of the model consists of a large "sea"

of order with small "islands" of disorder. This large "sea" of order could have either of the q visible colours, leading to q stable ordered phases.

The transition occurs when $\beta \approx \frac{1}{d}\log(q+r)$. In this case, the ordered bonds and the disordered bonds have the same free energy, whereas a bond which is neither ordered nor disordered (a bond on the boundary between two different regions) has an excess free energy that is at least $\frac{1}{2\beta d}\log(q+r)$. If $q+r$ is large, the bonds separating two different regions become very unlikely. Therefore, probable configurations are made of a large "sea" of either order or disorder, with small islands of disturbance. This leads to $q+1$ stable phases — q ordered ones and one disordered one.

In order to make this argument precise, we first need to identify, in an unambiguous fashion, what we mean by the "ordered" and "disordered" regions and the "boundary" separating them. Note that it is not clear whether a pair of neighbouring sites having the same colour should be considered as an "ordered" or "disordered". Indeed, in a disordered bond, by virtue of their independence, the two sites could also take the same visible colour with probability $\frac{q}{(q+r)^2}$.

In the standard Potts model, an elegant way to formalize order/disorder is provided by the random-cluster representation 18)–20). A similar representation can be constructed for the model with invisible colours. The idea is to resolve the ambiguous situations by flipping a coin:

- A bond whose endpoints either have different colours or at least one of them has an invisible colour, is considered as *disordered*.
- For a bond whose endpoints have the same visible colour, we flip a (temperature-biased) coin with probability $p_\beta = 1 - e^{-\beta}$ of having a head. If we get a head, we consider the bond as *ordered*; otherwise as *disordered*.

The particular choice of p_β ensures that, conditioned on which bonds are designated as ordered and which as disordered, the partitioning of the lattice into ordered and disordered regions has the desired properties we were after:

a) The sites within an ordered cluster (i.e., a maximal connected subgraph induced by ordered bonds) all have the same visible colour.
b) The colours of the sites that are completely inside a disordered region (i.e., are not incident to any ordered bond) are independent of one another, each chosen uniformly from the $q+r$ possible colours.

The probability distribution induced on the bond configurations (i.e., the configurations of ordered/disordered bonds) is a perturbation of the percolation model, in which to each non-singleton cluster induced by ordered bonds is given an extra weight q and to each isolated site (a site not incident to any ordered bond) is given an extra weight $q+r$. We call it the *r-biased random-cluster model*. For $r=0$, this model reduces to the standard random-cluster model. Many of the properties of the standard random-cluster model[14),21)] extend to this generalized model.

We shall now describe, in a more detailed manner, how this model is related to the (q,r)-Potts model. Afterwards we will indicate how, for large values of $q+r$, a first-order transition for the biased random-cluster representation can be proven. For a more detailed treatment we refer to Refs. 22) and 23).

Let $\mathbb{G} = (S, B)$ be a finite graph, where S denotes the set of sites, and B the set of bonds in the graph. The *r-biased random-cluster* model on $\mathbb{G}$ is given by a probability distribution on the sets $X \subseteq B$. The distribution has three parameters $0 \leq p \leq 1$, $q > 0$ and $r \geq 0$ and is defined by

$$\phi_{p,q,r}(X) = \frac{1}{Z^{\mathsf{RC}}_{p,q,r}(\mathbb{G})} \left[\prod_{b\in B} p^{\delta(b\in X)}(1-p)^{\delta(b\notin X)}\right] (q+r)^{\kappa_0(S,X)} q^{\kappa_1(S,X)} ,$$

in which $\kappa_0(S, X)$ denotes the number of isolated vertices of the graph (S, X), $\kappa_1(S, X)$ the number of non-singleton connected components of (S, X) and $Z^{RC}_{p,q,r}(\mathbb{G})$ the partition function. Notice that for $r = 0$, the model reduces to the standard random-cluster model, in which both singleton and non-singleton connected components have weight q. For $r > 0$, the above model induces a bias towards singleton connected components (isolated sites). Namely, the singleton connected components have weight $(q+r)$, whereas the non-singleton connected components have weight q. We emphasize that, while the standard random-cluster model does not distinguish between isolated sites and non-singleton connected components, the r-biased version distinguishes between them.

Let us now see how the (q, r)-Potts model, for positive integer q and r, is related to the r-biased random-cluster model. Let Ω be the set of (q, r)-Potts configurations on $\mathbb{G}$. The partition function of this model can be rewritten as

$$\begin{aligned}
Z_\beta(\mathbb{G}) &= \sum_{\sigma\in\Omega} e^{\beta\sum_{\{i,j\}\in B}\delta(\sigma_i=\sigma_j\leq q)} \\
&= \sum_{\sigma\in\Omega}\prod_{\{i,j\}\in B} e^{\beta\delta(\sigma_i=\sigma_j\leq q)} \\
&= \sum_{\sigma\in\Omega}\prod_{\{i,j\}\in B}\Big[1+\delta(\sigma_i=\sigma_j\leq q)(e^\beta-1)\Big] \\
&= \sum_{\sigma\in\Omega}\sum_{X\subseteq B}\prod_{\{i,j\}\in X}\delta(\sigma_i=\sigma_j\leq q)(e^\beta-1)^{|X|} \\
&= \sum_{\sigma\in\Omega}\sum_{X\subseteq B}\pi(\sigma,X) , \qquad (1)
\end{aligned}$$

where

$$\pi(\sigma,X) = e^{\beta|B|}\prod_{\{i,j\}\in B}\Big[\delta(\{i,j\}\in X)\delta(\sigma_i=\sigma_j\leq q)(1-e^{-\beta}) + \delta(\{i,j\}\notin X)e^{-\beta}\Big] .$$

The expression above describes a coupling of the (q, r)-Potts distribution on $\Omega = \{1, ...q+r\}^S$ and a probability distribution on the space $\{0, 1\}^B$ (compare Ref. 20)). The marginal of this coupling on the space $\{0, 1\}^B$ is simply the r-biased random-cluster distribution $\phi_{p_\beta,q,r}$ with $p_\beta = 1 - e^{-\beta}$. In particular, the weight $\pi(\sigma, X)$ can also be expressed as

$$\pi(\sigma,X) = e^{\beta|B|}\cdot 1_{F_r}(\sigma,X)\cdot\prod_{\{i,j\}\in B}\left[p_\beta\,\delta(\{i,j\}\in X) + (1-p_\beta)\,\delta(\{i,j\}\notin X)\right] ,$$

where

$$F_r \triangleq \{(\sigma, X) : \sigma_i = \sigma_j \leq q \text{ for all } \{i,j\} \in X\} \ .$$

Summing over σ we obtain

$$\sum_{\sigma \in \Omega} 1_{F_r}(\sigma, X) = q^{\kappa_1(S,X)}(q+r)^{\kappa_0(S,X)} \ . \tag{2}$$

The latter, together with Eq. (1), gives us

$$Z_\beta(\mathbb{G}) = e^{\beta|B|} Z^{\mathsf{RC}}_{p_\beta,q,r}(\mathbb{G}) \ . \tag{3}$$

The second step of the proof, which we can apply once we have set up our random cluster representation, is based on Pirogov-Sinai theory. Pirogov-Sinai theory is a powerful extension of the famous Peierls contour argument to situations in which there is no symmetry relation between the different phases. The technical problem which Pirogov-Sinai solves is that in such situations there is no independence between contours containing other contours. In its original form,[24),25)] the theory shows that under rather weak conditions a model in dimension at least 2 with a certain finite number of translation-invariant ground-state configurations has the same number of low-temperature phases on some coexistence curve. The high energy cost of large contours at low temperatures makes them sufficiently improbable, so that a typical phase consists of an infinite connected "sea" of sites on which the spins take a particular ground state value, with finite clusters of sites (droplet excitations) in which the spins take mostly different values.

In the high-q Potts model, the usual high-energy contours are replaced by high-free-energy contours, in which the role of the large inverse temperature is taken by a large value of $\log(q+r)$. In our random-cluster version, if we define energies of bond configurations (rather than spin configurations) as the logarithms of their probabilities, we get a model which is more similar to the original Pirogov-Sinai theory. The two "ground states" now become the all-occupied or the all-empty bond configurations. Due to the lack of symmetry in the model between "occupied" and "empty", we need to introduce two different categories of contours: one consisting of contours which separate order from disorder (i.e. which have order in its exterior and disorder in its interior) and one of contours which separate disorder from order. We call the first type *order* contours and the second type *disorder* contours. Amongst both types of contours, a special role (to infer the stability of each of the ordered and disordered phases) is played by those contours separating a unique infinite "sea" of one phase (which could be either ordered or disordered) from a finite "island" of disturbance: we refer to them as *external* contours. When $q+r$ is large enough, the probability of a large contour is very small, which ensures (almost surely with respect to any Gibbs measure) that no infinite cascade of contours occurs. Then we will have an infinite "sea" of one of the phases, with "islands" bounded by external contours. In the high-temperature regime, due to the fact that the lowest free energy is that of the disordered bonds, the external contours are of the disorder type, implying the stability of the disordered phase. In the low-temperature regime, due to the lowest

free energy now belonging to the ordered bonds, the external contours are of the order type. As in the previous case, if $q+r$ is large, they again become very unlikely, which now leads to the stability of the ordered phase. At the transition temperature, for $q+r$ large, both types of external contours might occur, due to the ordered and disordered bond having the same free energy, which makes it impossible to impose a "preference" between order and disorder. As for $q+r$ large, both type of external contours become unlikely, the configurations can be cast in the "sea/islands" picture with overwhelming probability. However, at the transition point, the "sea" could be either ordered or disordered. This leads to the coexistence of the ordered an the disordered phase.

For the original Potts model, this approach was developed in Ref. 12) and it is reviewed in Ref. 14). The main point where the proof for the (q,r)-Potts model differs from that of the standard Potts model is that the leading term in the probability cost of the external contours depends on $q+r$ instead of q. For a more detailed and self-contained analysis we refer the reader to Ref. 22).

□

§3. Comments and conclusions

In this note, we have shown how the Potts model with many invisible states, introduced by Tamura, Tanaka and Kawashima, can be proven to have a first-order phase transition, similarly as occurs for the standard high-q Potts models. The transition temperature is asymptotically given by $\beta \approx \frac{1}{d}\log(q+r)$. The proofs, as usual, apply for quite high values of q and/or r. The numerical approach of Refs. 1) and 2) might give a better indication of the values at which the first-order transition first occurs.

Our proof employed a random-cluster representation of the model. This has the advantage that it extends to values of q and r which need not correspond to a spin model interpretation, e.g. $q=1$, or q and/or r non-integer.

We conjecture that the dynamical properties of the Potts model with r invisible states, which were considered in Ref. 26), have a similar corresponding behaviour as occurs in the ordinary Potts model. These were rigorously analysed in Ref. 27).

Acknowledgements

G. I. and S. T. thank the Nederlandse organisatie voor Wetenschappelijk Onderzoek NWO for support.

References

1) R. Tamura, S. Tanaka and N. Kawashima, Prog. Theor. Phys. **124** (2010), 381.
2) S. Tanaka, R. Tamura and N. Kawashima, J. Phys.: Conf. Ser. **297** (2011), 012022.
3) T. Gobron and I. Merola, J. Stat. Phys. **126** (2007), 507.
4) H. W. J. Blöte, W. Guo and H. J. Hilhorst, Phys. Rev. Lett. **88** (2002), 047203.
5) E. Domany, M. Schick and R. H. Swendsen, Phys. Rev. Lett. **52** (1984), 1535.
6) A. C. D. van Enter and S. B. Shlosman, Phys. Rev. Lett. **89** (2002), 285702.
7) A. C. D. van Enter and S. B. Shlosman, Commun. Math. Phys. **255** (2005), 21.

8) A. C. D. van Enter and S. B. Shlosman, Markov Proc. Rel. Fields **13** (2007), 239.
9) A. N. Berker, S. Ostlund and F. A. Putnam, Phys. Rev. B **17** (1978), 3650.
10) J. Bricmont, K. Kuroda and J. L. Lebowitz, Commun. Math. Phys. **101** (1985), 501.
11) R. Kotecký and S. B. Shlosman, Commun. Math. Phys. **83** (1982), 495.
12) L. Laanait, A. Messager, S. Miracle-Sole, J. Ruiz and S. B. Shlosman, Commun. Math. Phys. **140** (1991), 81.
13) D. H. Martirosian, Commun. Math. Phys. **105** (1986), 281.
14) G. Grimmett, *The random-cluster model* (Springer, Berlin, 2006; corrected reprint, 2009).
15) R. Kotecký, "Geometric representation of lattice models and large volume asymptotics", in *Probability and Phase Transition* (Kluwer, Dordrecht, 1994).
16) H. O. Georgii, *Gibbs Measures and Phase Transitions* (De Gruyter, Berli, New York, 1988; 2nd Edition, 2011).
17) S. B. Shlosman, Russ. Math. Surveys **41** (1986), 83.
18) C. M. Fortuin and P. W. Kasteleyn, Physica **57** (1972), 536.
19) P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Jpn. **26** (1969) Suppl., p. 11.
20) R. G. Edwards and A. D. Sokal, Phys. Rev. D **38** (1988), 2009.
21) H. O. Georgii, O. Häggström and C. Maes, *Phase Transitions and Critical Phenomena*, vol. 18 (Academic Press, 2000), p. 1.
22) A. C. D. van Enter, G. Iacobelli and S. Taati, arXiv:1109.0189.
23) G. Iacobelli, PhD Thesis.
24) S. A. Pirogov and Ya. G. Sinai, Theor. Math. Phys. **25** (1976), 1185.
25) Ya. G. Sinai, *Theory of phase transitions: Rigorous results* (Pergamon Press, Oxford, 1982).
26) S. Tanaka and R. Tamura, arXiv:1012.4254.
27) C. Borgs, J. T. Chayes, J. H. Kim, A. Frieze, P. Tetali, E. Vigoda and V. H. Vu, *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science, New York, 1999*, p. 218.