Self-averaging in models of random copolymer collapse

2000 J. Phys. A: Math. Gen. 33 259

(http://iopscience.iop.org/0305-4470/33/2/303)

J. Phys. A: Math. Gen. **33** (2000) 259–266. Printed in the UK PII: S0305-4470(00)05770-X

Self-averaging in models of random copolymer collapse

E Orlandini†, M C Tesi‡ and S G Whittington§

† INFM, Dipartimento di Fisica, Università di Padova, Padova, Italy
‡ Dipartimento di Matematica, Università di Bologna, Bologna, Italy
‡ Department of Chemistry, University of Toronto, Toronto, Canada M5S 3H6

Received 7 July 1999, in final form 20 October 1999

Abstract. We give a set of conditions under which a system is thermodynamically self-averaging and show that several lattice models of interacting copolymers satisfy these conditions. We prove this result for a general potential which is linear in the numbers of various types of contacts, and show that this includes two potentials which have previously been used in models of random interacting linear copolymers.

1. Introduction

Random copolymers are an interesting example in the statistical mechanics of random systems and several papers have appeared on the adsorption (Garel *et al* 1989, Gutman and Chakraborty 1994, 1995, Bolthausen and den Hollander 1997, Biskup and den Hollander 1999, Orlandini *et al* 1999) and collapse (Sfatos and Shakhnovich 1997, Garel *et al* 1998, Monari *et al* 1999) of random copolymers. There are interesting questions concerning the existence of the thermodynamic limit of the quenched average free energy and thermodynamic self-averaging in these systems. Random systems are characterized by a Hamiltonian which depends on a set of random parameters, and in the case of random copolymers these parameters are the identities of the individual monomers making up the polymer. For a given set of values of these parameters (e.g. the sequence of comonomers along the polymer chain) one can ask for the value of the free energy of the system. If this free energy is equal (almost surely) to the quenched average free energy (in the thermodynamic limit of infinitely large systems) the system is said to be thermodynamically self-averaging. Orlandini *et al* (1999) proved that a self-avoiding walk model of random copolymer adsorption is thermodynamically self-averaging but there are rather few similar results in this area. Thermodynamic self-averaging has been proved for random spin models with both short-range (van Hemmen and Palmer 1982) and long-range interactions (van Enter and van Hemmen 1983). However, it is known that correlation functions are not self-averaging in some random spin problems (Derrida and Hilhorst 1981, Sourlas 1987).

A key ingredient in the proof of the existence of the thermodynamic limit in lattice problems is the derivation of a sub- or super-additive inequality for the logarithm of the partition function. In random systems this means that we are dealing with a sub-additive or super-additive stochastic process, and there is an extensive mathematical literature on these processes. They were first introduced by Hammersley and Welsh (1965) in the context of first passage percolation and Kingman (1968, 1973) proved an important ergodic theorem for stationary subadditive processes, in the one-parameter case. Akcoglu and Krengel (1981) generalized

0305-4470/00/020259+08$30.00

this result to multi-parameter processes and replaced Kingman's stationarity condition by an endomorphism on the measure space. The proof of self-averaging for long-range random spin systems by van Enter and van Hemmen (1983) makes use of Akcoglu and Krengel's ergodic theorem.

The theorem of Akcoglu and Krengel (1981) is a very powerful tool and, in section 2, we use it to derive sufficient conditions under which certain lattice models of copolymers are thermodynamically self-averaging.

There is overwhelming numerical evidence for a collapse transition both in models of linear (Mazur and McCrackin 1968, Webman *et al* 1981, Tesi *et al* 1996a) and branched polymers (Derrida and Herrmann 1983, Gaunt 1991, Gaunt and Flesia 1991) in the homopolymer case, though the existence of a collapse transition has not been demonstrated rigorously. In section 3 we consider several specific models of random copolymers which are expected to show a collapse transition. We show that these models each satisfy the conditions of section 2, and hence show that they are thermodynamically self-averaging.

2. A sufficient condition for self-averaging

In this section we consider a class of lattice models of polymers, such as walks, trees, animals, etc. In each case the vertices will be coloured and we shall be interested in interacting versions of these objects, so that the vertices interact with a surface, or with each other, etc. We shall consider the d-dimensional hypercubic lattice Z^d though extensions to other lattices are possible. To formalize the situation we consider a class $\mathcal{C}$ of combinatorial objects which are connected subgraphs of Z^d, viewed as embeddings of some abstract object. These subgraphs are stratified by the number (n) of vertices in the embedding and we label the vertices $i = 1, 2, \ldots, n$ in some canonical way. The vertices are each coloured randomly and independently and we represent the colour of vertex i by the random variable χ_i belonging to the probability space Y. We write χ for the sequence $\chi_1, \chi_2, \ldots, \chi_n$ where χ belongs to the probability space $X = Y \times Y \ldots$. The embeddings have an associated energy determined by an integer-valued vector s which depends on χ. We call s the *dimensionless energy* of the embedding. Let $f_n(s|\chi)$ be the number of embeddings of an object in class $\mathcal{C}$ with n vertices with colouring χ, and dimensionless energy s. Define the partition function

$$Z_n(\omega|\chi) = \sum_s f_n(s|\chi)\, \mathrm{e}^{\omega \cdot s} \tag{2.1}$$

and the associated free energy

$$F_n(\omega|\chi) = n^{-1} \log Z_n(\omega|\chi) \tag{2.2}$$

where the vector ω is conjugate to s.

Suppose that $Z_n(\omega|\chi)$ satisfies the following conditions:

1. $n^{-1} \log Z_n(\omega|\chi) \leqslant M(\omega) < \infty$ for all finite n, when all components of ω are finite and $\chi \in X$.
2. $Z_{n_1}(\omega|\chi_1) Z_{n_2}(\omega|\chi_2) \leqslant Z_{n_1+n_2}(\omega|\chi)$ where χ is the concatenation of the colourings χ_1 and χ_2.

It follows that, for $\chi \in X$ and for any fixed real ω the function $\log Z_n(\omega|\chi)$ is integrable over the probability space X. Under these conditions $\log Z_n(\omega|\chi)$ is a discrete super-additive process and in the remainder of this section we shall assume that these conditions hold.

Theorem 2.1. *The limiting quenched average free energy*

$$\lim_{n\to\infty} \langle F_n(\omega|\chi)\rangle \equiv \bar{\mathcal{F}}(\omega) \tag{2.3}$$

exists, where $\langle\cdots\rangle$ *is an average over colourings.*

Proof. Taking logarithms in condition 2 and averaging over the random colourings we obtain

$$\langle \log Z_{n_1}(\omega|\chi_1)\rangle + \langle \log Z_{n_2}(\omega|\chi_2)\rangle \leqslant \langle \log Z_{n_1+n_2}(\omega|\chi)\rangle. \tag{2.4}$$

This equation, together with condition 1 and the super-additive function theorem (Hille 1948) immediately implies (2.3). □

Lemma 2.1. *Conditions 1 and 2 imply that*

$$\liminf_{n\to\infty} F_n(\omega|\chi) \geqslant \bar{\mathcal{F}}(\omega) \tag{2.5}$$

for almost all $\chi \in X$.

Proof. For fixed m write $n = mp + q$ with $0 \leqslant q < m$. Repeated application of condition 2 gives

$$Z_n(\omega|\chi) \geqslant \left[\prod_{i=1}^{p} Z_m(\omega|\chi_i)\right] \times Z_q(\omega|\chi_0) \tag{2.6}$$

where the colourings χ_i and χ_0 are inherited from χ. Taking logarithms and dividing by n we obtain

$$n^{-1}\log Z_n(\omega|\chi) \geqslant \left[\frac{1}{m(p+q/m)}\sum_{i=1}^{p}\log Z_m(\omega|\chi_i)\right] + n^{-1}\log Z_q(\omega|\chi_0). \tag{2.7}$$

Letting $p \to \infty$ with m fixed we obtain

$$\liminf_{n\to\infty} n^{-1}\log Z_n(\omega|\chi) \geqslant \langle m^{-1}\log Z_m(\omega|\chi)\rangle \tag{2.8}$$

almost surely, where we have made use of the strong law of large numbers. Letting $m \to \infty$ and using theorem 2.1 we obtain (2.5). □

We next state a version of the local super-additive ergodic theorem of Akcoglu and Krengel (1981), appropriate to our problem.

Theorem 2.2 (Akcoglu and Krengel). *Since* $\log Z_n(\omega|\chi)$ *is a discrete super-additive process then* $\lim_{n\to\infty} F_n(\omega|\chi) \equiv \eta(\omega|\chi)$ *exists for almost all* χ.

The next theorem allows us to identify the value of this limit. Akcoglu and Krengel remarked that under certain circumstances $\eta(\omega|\chi)$ is almost surely equal to the quenched average free energy and we give a proof appropriate to our conditions.

Theorem 2.3. *Lemma 2.1 implies that*

$$\eta(\omega|\chi) = \bar{\mathcal{F}}(\omega) \tag{2.9}$$

for almost all $\chi \in X$.

Proof. By theorem 2.2 $\lim_{n\to\infty} F_n(\omega|\chi) \equiv \eta(\omega|\chi)$ exists for almost all $\chi \in X$ so lemma 2.1 implies that $\eta(\omega|\chi) \geqslant \bar{\mathcal{F}}(\omega)$ for almost all $\chi \in X$. Suppose that the set for which this is true is X^*, then X^* has measure $\mu(X^*) = 1$. Suppose, to obtain a contradiction, that there exists a positive measure set S^+ such that, for all $\chi \in S^+$, $\eta(\omega|\chi) > \bar{\mathcal{F}}(\omega)$. Let S^- be the set such that $\chi \in S^-$ implies that $\eta(\omega|\chi) < \bar{\mathcal{F}}(\omega)$. Clearly (2.5) implies that $\mu(S^-) = 0$. From condition 1 and the fact that X is a probability space, the dominated convergence theorem implies that $\eta(\omega|\chi)$ is integrable and

$$\bar{\mathcal{F}}(\omega) = \lim_{n\to\infty} \int_{X^*} n^{-1} \log Z_n(\omega|\chi)\, \mathrm{d}\chi = \int_{X^*} \eta(\omega|\chi)\, \mathrm{d}\chi. \tag{2.10}$$

Using the fact that

$$\int_{X^*-S^--S^+} \eta(\omega|\chi)\, \mathrm{d}\chi = (1 - \mu(S^-) - \mu(S^+))\bar{\mathcal{F}}(\omega) \tag{2.11}$$

this implies that

$$\int_{S^-} \eta(\omega|\chi)\, \mathrm{d}\chi + \int_{S^+} \eta(\omega|\chi)\, \mathrm{d}\chi = \bar{\mathcal{F}}(\omega)[\mu(S^+) + \mu(S^-)]. \tag{2.12}$$

However, $\mu(S^-) = 0$ so that this equation implies that

$$\langle \eta(\omega|\chi)\rangle_{\chi\in S^+} = \int_{S^+} \eta(\omega|\chi)\, \mathrm{d}\chi/\mu(S^+) = \bar{\mathcal{F}}(\omega). \tag{2.13}$$

However, this is impossible since $\eta(\omega|\chi) > \bar{\mathcal{F}}(\omega)$ for all $\chi \in S^+$. □

3. Some examples

In this section we use the results of section 2 to prove thermodynamic self-averaging for several interesting problems in polymer physics. Each of these problems could be used as a model of polymer collapse in random copolymers. We treat them in a fairly general way so that our treatment includes several of the Hamiltonians which have been suggested in the literature (Garel *et al* 1998) for modelling collapse in random copolymers. For each of these problems we need to check that the model satisfies conditions 1 and 2 above.

3.1. Self-interacting random unfolded self-avoiding walks

The proof of the existence of a collapse transition in interacting self-avoiding walks, where all vertices are treated in the same way and there is no random colouring, is still open. This problem was studied (for homopolymers) by Tesi *et al* (1996b), where they proved that the limiting free energy exists for repulsive interactions between vertices, and that it is equal to the corresponding free energy of self-interacting polygons. For attractive interactions there is not even a proof of the existence of the limiting free energy.

We are not able to say anything about self-averaging in the random interacting self-avoiding walk problem but we can prove self-averaging in a simpler model. We consider self-avoiding walks on Z^d. We shall be interested in walks with n vertices and number these vertices $i = 1, \ldots, n$ in a sequential way. We write $(x_i, y_i, \ldots, z_i)$ for the coordinates of the ith vertex and say that a walk is *unfolded* if $x_1 < x_i \leqslant x_n$ for all $i = 2, \ldots, n$. The vertices of the walk are independently coloured +1 with probability p and -1 with probability $1 - p$. Let $b_n(k_{++}, k_{--}, k_{+-}|\chi)$ be the number of unfolded walks with n vertices, k_{++} contacts between pairs of vertices coloured +1 and +1, k_{--} contacts between pairs of vertices coloured -1 and

-1 and k_{+-} contacts between pairs of vertices coloured +1 and -1, given a colouring sequence χ. Define the partition function

$$B_n(\beta|\chi) = \sum_k b_n(k|\chi)\, \mathrm{e}^{\beta g(k)} \tag{3.1}$$

where k is the vector (k_{++}, k_{--}, k_{+-}) and $g(k)$ is a linear function of k. We take $\beta < \infty$.

Since the number of unfolded walks is bounded above by $(2d)^n$ and the maximum number of contacts is $O(n)$, condition 1 is satisfied for any $\beta < \infty$.

Two unfolded walks can be concatenated as follows. We shall write $(x_i^A, y_i^A, \ldots, z_i^A)$, $i = 1, 2, \ldots, n_A$, for the coordinates of vertex i in walk A and $(x_i^B, y_i^B, \ldots, z_i^B)$, $i = 1, 2, \ldots, n_B$, for the coordinates of vertex i in walk B. Translate walk B so that $x_1^B = x_{n_A}^A + 1$, $y_1^B = y_{n_A}^A, \ldots, z_1^B = z_{n_A}^A$, and add an edge joining the n_Ath vertex of walk A to the first vertex of walk B. This produces an unfolded walk with $n_A + n_B$ vertices, $k_{++}^A + k_{++}^B$ contacts between pairs of vertices labelled +1 and +1, etc, so that the k-vectors add, and with a colouring of vertices which is exactly the concatenation of the colourings in the two subwalks. This immediately gives the inequality

$$B_{n_A}(\beta|\chi_A) B_{n_B}(\beta|\chi_B) \leqslant B_{n_A+n_B}(\beta|\chi) \tag{3.2}$$

where χ is the concatenation of the colourings χ_A and χ_B. This establishes condition 2. Since both conditions hold the model is thermodynamically self-averaging.

3.2. Self-interacting random lattice trees

The general problem of collapse in trees as a model of branched homopolymers (i.e. without random colouring of the vertices) has been studied by Madras *et al* (1990). They showed that the thermodynamic limit exists and derived bounds on the behaviour of the free energy.

For the quenched copolymer case one should randomly label each vertex of a tree (with two colours, say) and this labelling should then be fixed. The labelling should be a property of the *graph* and not of the *embedding*. Unlike the unfolded walk model, we are not able to handle this case but we can prove results for a different labelling scheme which we call *lexicographically quenched*. For each *embedding* of the graph, i.e. for each lattice tree, we label the vertices lexicographically so that the labelling is additive under the concatenation operation which we use.

We consider the d-dimensional hypercubic lattice with vertices having integer coordinates $(x, y, \ldots, z)$. A *contact* is an edge of the lattice which is not an edge of the tree but which is incident on two vertices of the tree. The tree is labelled lexicographically and each vertex of the tree is coloured +1 with probability p and -1 with probability $1 - p$, independently. With the vertices ordered lexicographically we write χ for the corresponding *colouring*, i.e. the sequence of +1 and -1. Let $t_n(k_{++}, k_{--}, k_{+-}|\chi)$ be the number of trees, modulo translation, with n vertices, with k_{++} contacts between pairs of vertices coloured +1 and +1, k_{--} contacts between pairs of vertices coloured -1 and -1 and k_{+-} contacts between pairs of vertices coloured +1 and -1, given a colouring χ of the vertices. Note that k_{++}, etc are all functions of χ and hence are random variables. Define k to be the vector (k_{++}, k_{--}, k_{+-}), and write $t_n(k|\chi)$ for $t_n(k_{++}, k_{--}, k_{+-}|\chi)$.

Define the partition function

$$T_n(\beta|\chi) = \sum_k t_n(k|\chi)\, \mathrm{e}^{\beta g(k)} \tag{3.3}$$

where $g(k)$ is a linear function of the vector k. We take $\beta < \infty$.

Since the number of trees is exponentially bounded (Klein 1981, Klarner 1967), the maximum number of contacts is $O(n)$ and $\beta < \infty$, condition 1 is automatically satisfied, so we only have to check condition 2.

Each tree has a *leftmost plane* (*rightmost plane*) such that vertices in this plane have a minimum (maximum) x-coordinate. Consider a particular tree τ. Suppose the leftmost (rightmost) plane of τ is $x = x_L(\tau)$ ($x = x_R(\tau)$). Call the lexicographically first vertex in the leftmost plane the *left vertex* of τ, and the lexicographically last vertex in the rightmost plane the *right vertex* of τ. Consider two trees τ_1 and τ_2, such that τ_1 has n_1 vertices and k-vector k_1, and τ_2 has n_2 vertices and k-vector k_2. With τ_1 fixed, having its right vertex at $(x_R^1, y_R^1, \ldots, z_R^1)$, translate τ_2 such its left vertex has coordinates $x_L^2 = x_R^1 + 1$, $y_L^2 = y_R^1, \ldots, z_L^2 = z_R^1$. Add an edge incident on these two vertices to form a single tree with $n_1 + n_2$ vertices and k-vector equal to $k_1 + k_2$ since the concatenation does not create or destroy any contacts. This implies that

$$\sum_{k_1} t_{n_1}(k_1|\chi_1)\, t_{n_2}(k - k_1|\chi_2) \leqslant t_{n_1+n_2}(k|\chi) \tag{3.4}$$

where we have written χ for the concatenation of the two colourings χ_1 and χ_2. Multiplying both sides of (3.4) by $e^{\beta g(k)}$, and summing over k, we obtain

$$T_{n_1}(\beta|\chi_1)T_{n_2}(\beta|\chi_2) \leqslant T_{n_1+n_2}(\beta|\chi) \tag{3.5}$$

where we have made use of the linearity of $g(k)$. This establishes condition 2 for this model and we have therefore established that the model is thermodynamically self-averaging.

3.3. Self-interacting random lattice animals

Madras *et al* (1990) considered six different models of interacting lattice animals which are all expected to show a collapse transition. The animal can be a sub-graph (i.e. a weak embedding) or a section graph (i.e. a strong embedding). In each case the animal can be counted by vertices or by edges. For section graphs the collapse can be induced by cycles, while for subgraphs it can also be induced by contacts. Provided that the animal is labelled lexicographically, the methods used in section 3.1 above can be used without essential modification to prove self-averaging for the three cases in which the animal is counted by vertices (a^0, sub-graphs with collapse induced by cycles, A^0, section graphs with collapse induced by cycles, and a, sub-graphs with collapse induced by contacts, in the notation of Madras *et al*). The exponential bound in condition 1 is almost immediate, using the exponential upper bound on the number of animals due to Klarner (1967), since the maximum number of cycles or contacts is in each case bounded above by $(d-1)n$. For $\beta < \infty$ we immediately have condition 1. The functional inequality in condition 2 is proved by a concatenation argument analogous to that used to prove (3.4) and hence (3.5).

4. Conclusions

We have derived a set of sufficient conditions for a random statistical mechanical model to be thermodynamically self-averaging. We have shown that several lattice models of self-interacting copolymers obey these conditions and hence are self-averaging. For the cases of linearly labelled unfolded walks and lexicograpically labelled lattice trees, where the potential depends on a linear function of the numbers of the various types of contacts, we have given an explicit proof of self-averaging and noted that these methods can be easily extended to handle several models of lexicographically labelled self-interacting lattice animals. In addition we

note that the argument given for linearly labelled unfolded walks applies *mutatis mutandis* to the model of linearly labelled directed walks, which is the random version of a model for which a transition is known to occur in the homopolymer case (Brak *et al* 1992).

Our potential is quite general and includes several models which have previously been introduced in studies of collapse in models of linear heteropolymers (Garel *et al* 1998). In particular, the binary quenched random hydrophilic–hydrophobic model, introduced by Garel *et al* (1994),

$$\sum_{i<j} \Delta_{ij}(\lambda_i + \lambda_j) = g(k_{++}, k_{--}, k_{+-}) = 2(k_{++} - k_{--}) \tag{4.1}$$

where Δ_{ij} is 1 when vertices i and j are near neighbours on the lattice and form a contact, and zero otherwise. λ_i is the colouring of vertex i and can be ± 1. For the random short-range interaction model introduced by Kantor and Kardar (1994)

$$\sum_{i<j} \Delta_{ij}\lambda_i\lambda_j = g(k_{++}, k_{--}, k_{+-}) = k_{++} + k_{--} - k_{+-}. \tag{4.2}$$

Our results show that lexicographically labelled random trees and animals and linearly labelled unfolded walks, with either of the above potentials, are self-averaging, but the corresponding question for interacting self-avoiding walks is still open.

Acknowledgments

The authors wish to thank Buks van Rensburg and Chris Soteros for helpful conversations, and NSERC of Canada for financial support.

References

Akcoglu M A and Krengel U 1981 *J. Reine Angew. Math.* **323** 53–67
Biskup M and den Hollander F 1999 A heterpolymer near a linear interface *Ann. Appl. Prob.* at press
Bolthausen E and den Hollander F 1997 *Ann. Prob.* **25** 1334–66
Brak R, Guttmann A J and Whittington S G 1992 *J. Phys. A: Math. Gen.* **25** 2437–46
Derrida B and Herrmann H J 1983 *J. Physique* **44** 1365–76
Derrida B and Hilhorst H 1981 *J. Phys. C: Solid State Phys.* **14** L539–44
Garel T, Huse D A, Leibler S and Orland H 1989 *Europhys. Lett.* **8** 9–13
Garel T, Leibler S and Orland H 1994 *J. Physique II* **4** 2139–48
Garel T, Orland H and Pitard E 1998 Protein folding and heteropolymers *Spin Glasses and Random Fields* ed A P Young (Singapore: World Scientific) pp 387–443
Gaunt D S 1991 *Physica* A **177** 146–52
Gaunt D S and Flesia S 1991 *J. Phys. A: Math. Gen.* **24** 3655–70
Gutman L and Chakraborty A K 1994 *J. Chem. Phys.* **101** 10074–91
——1995 *J. Chem. Phys.* **103** 10 733–50
Hammersley J M and Welsh D J A 1965 *Bernoulli–Bayes–Laplace Anniversary Volume* (Berlin: Springer) pp 65–110
Hille E 1948 *Functional Analysis and Semi-Groups (Am. Math. Soc. Colloq. Publ. 31)* (New York: American Mathematical Society)
Kantor Y and Kardar M 1994 *Europhys. Lett.* **28** 169–74
Kingman J F C 1968 *J. R. Statist. Soc.* B **30** 499–510
——1973 *Ann. Prob.* **1** 883–909
Klarner D A 1967 *Can. J. Math.* **19** 851–63
Klein D J 1981 *J. Chem. Phys.* **75** 5186–9
Madras N, Soteros C E, Whittington S G, Martin J L, Sykes M F, Flesia S and Gaunt D S 1990 *J. Phys. A: Math. Gen.* **23** 5327–50
Mazur J and McCrackin F L 1968 *J. Chem. Phys.* **49** 648–65
Monari P, Stella A L, Vanderzande C and Orlandini E 1999 *Phys. Rev. Lett.* **83** 112–15

Orlandini E, Tesi M C and Whittington S G 1999 *J. Phys. A: Math. Gen.* **32** 469–77
Sfatos C D and Shakhnovich E I 1997 *Phys. Rep.* **288** 77–108
Sourlas N 1987 *Europhys. Lett.* **3** 1007–11
Tesi M C, Janse van Rensburg E J, Orlandini E and Whittington S G 1996a *J. Stat. Phys.* **82** 155–81
——1996b *J. Phys. A: Math. Gen.* **29** 2451–63
van Enter A C D and van Hemmen J L 1983 *J. Stat. Phys.* 141–52
van Hemmen J L and Palmer R G 1982 *J. Phys. A: Math. Gen.* **15** 3881–90
Webman I, Lebowitz J L and Kalos M H 1981 *Macromolecules* **14** 1495–501