

# Random copolymers and the Morita approximation: polymer adsorption and polymer localization

E Orlandini<sup>1</sup>, A Rechnitzer<sup>2</sup> and S G Whittington<sup>3</sup>

<sup>1</sup> INFN, Dipartimento di Fisica, Università di Padova, Padova, Italy

<sup>2</sup> Department of Mathematics and Statistics, York University, Toronto, Canada M3J 1P3

<sup>3</sup> Department of Chemistry, University of Toronto, Toronto, Canada M5S 3H6

Received 25 June 2002

Published 28 August 2002

Online at [stacks.iop.org/JPhysA/35/7729](http://stacks.iop.org/JPhysA/35/7729)

## Abstract

We analyse directed walk models of random copolymer adsorption and localization. Ideally we would like to solve the quenched problem, but it appears to be intractable even for simple directed models. The annealed approximation is solvable, but is inadequate in the strong interaction regime for the adsorption problem and gives a qualitatively incorrect phase diagram for the localization problem.

In this paper, we treat these directed models using an approximation suggested by Morita (1964 *J. Math. Phys.* **5** 1401–5) in which the proportion of each comonomer is fixed. We find that the Morita approximation leads to behaviour that is closer to that of the quenched average model and this is particularly interesting in the localization problem where the phase diagram is (at least qualitatively) very similar to that of the quenched average problem. We also show that the phase boundaries in the Morita approximation are bounds on the locations of the phase boundaries of the quenched model.

PACS numbers: 05.40.Fb, 05.50.+q, 82.35.Jk

## 1. Introduction

There has been considerable recent interest in the statistical mechanics of random copolymers, especially for systems which exhibit a phase transition such as collapse (Sfatos and Shakhovich 1997, Garel *et al* 1998, Monari *et al* 1999, Chuang *et al* 2001) or adsorption (Garel *et al* 1989, Gutman and Chakraborty 1994, Orlandini *et al* 1999). Suppose the copolymer has two types of monomers, which we shall denote as  $A$  and  $B$ . The sequence of monomers in a particular polymer molecule is determined by some random process, but is then fixed, so that random copolymers are an example of a system with quenched randomness (Brout 1959). If  $\chi$  represents a particular sequence of monomers (i.e. a sequence of  $A$ s and  $B$ s) then the partition function, and other properties, will depend on  $\chi$ . If we write  $Z_n(\chi)$  for the partition function

for a polymer with  $n$  monomers, and monomer sequence  $\chi$ , then the appropriate free energy of the system is the *quenched average free energy*  $\langle \kappa_n(\chi) \rangle = \langle n^{-1} \log Z_n(\chi) \rangle$ , where the angular brackets denote an average over all monomer sequences (Brout 1959, Mazo 1963). In order to investigate phase transitions in such systems, one is interested in the thermodynamic limit ( $n \rightarrow \infty$ ) and the existence of the limit  $\bar{\kappa} = \lim_{n \rightarrow \infty} \langle \kappa_n(\chi) \rangle$  has been proved for several models (see, for instance, Orlandini *et al* (1999) for a typical approach to this type of proof).

In this paper, we shall be concerned with two different physical situations. The first is adsorption of a random copolymer at an impenetrable surface and the second is localization of a random copolymer at an interface between two immiscible solvents.

One of the standard models of the configurational properties of polymer molecules in dilute solution is a self-avoiding walk on a lattice. The vertices of the walk are coloured (according to some random process) to represent the sequence of monomers in the polymer and, once the colouring is determined, it is fixed. Self-avoiding walk models of both the physical situations have been investigated and, although some qualitative properties of the quenched average free energy (such as convexity, continuity, the presence of a singularity corresponding to a phase transition) can be established, the evaluation of the quenched average free energy poses formidable problems (Orlandini *et al* 1999, Maritan *et al* 1999, Martin *et al* 2000). Even for simplified models such as directed walks, no one has succeeded in computing the quenched average free energy completely for these systems (Bolthausen and den Hollander 1997).

One way to treat such problems is to approximate the quenched average with an annealed average. This corresponds to calculating the logarithm of the average of the partition function giving

$$\kappa_n^a = n^{-1} \log \langle Z_n(\chi) \rangle \quad (1.1)$$

and the corresponding annealed limiting free energy

$$\kappa^a = \lim_{n \rightarrow \infty} \kappa_n^a. \quad (1.2)$$

By the arithmetic mean-geometric mean inequality  $\kappa^a \geq \bar{\kappa}$ , but the two quantities can differ substantially in some cases.

The annealed approximation does not guarantee that even the lower moments of the distribution of colours are correct. In an important paper, Morita (1964) proposed a set of successive approximations in which the quenched average is approximated by partially annealed averages in which successive moments of the colour distribution are guaranteed to be correct. Kühn (1996) showed that successive approximations give successively better bounds on the quenched average free energy. The idea behind Kühn's proof is as follows. Mazo (1963) showed that the quenched average system could be described using a variational principle in which the entropy is maximized subject to the fixed probability of occurrence of the quenched random variables. If we relax this constraint by fixing only a set of moments of the distribution of the quenched variables, the space over which the optimization of the entropy is being carried out is larger, so the maximum of the entropy cannot decrease. Fixing more moments of the distribution decreases the space and therefore gives successively better bounds on the quenched system. The annealed approximation fixes none of the moments so is the least restrictive condition, and gives the weakest bound.

In some circumstances, insisting that the first moment is correct gives a substantial improvement over the simple annealed approximation (Trovato *et al* 1998). In this paper, we investigate the application of this approximation to models of copolymer adsorption and localization. Even within this approximation, we are unable to treat the self-avoiding walk model and we confine our attention to randomly coloured Dyck paths (see for instance Janse van Rensburg (2000)) which are directed walks on a lattice. For random copolymer

adsorption we find that, at large values of the interaction parameter (or low temperature), the free energy is changed when we condition on the first moment of the colour distribution being correct, but the location of the adsorption transition is unchanged. For localization of a random copolymer, we find that this additional condition gives rise to a substantial change in the nature of the phase diagram, and brings it much closer to that of the quenched average version of the problem (Bolthausen and den Hollander 1997, Maritan *et al* 1999, Martin *et al* 2000).

In section 2 we discuss the application of Morita’s idea to the random copolymer adsorption problem. We begin in the more general setting of a self-avoiding walk model but, in order to make real progress, we specialize to a Dyck path model of adsorption in section 3. In section 4 we investigate a Dyck path model of localization. We end with a brief discussion of our results in section 5.

**2. Adsorption of random copolymers at a surface**

We first consider the self-avoiding walk model of homopolymer adsorption. We consider the  $d$ -dimensional hypercubic lattice  $\mathbb{Z}^d$  whose vertices are the integer points in  $\mathbb{R}^d$  and we write  $(x, y, \dots, z)$  for the coordinates of a vertex in the lattice. We consider  $n$ -edge self-avoiding walks on this lattice, starting at the origin and having no vertex with negative  $z$ -coordinate. Suppose that  $c_n(v)$  is the number of such walks with  $v + 1$  vertices in the hyperplane  $z = 0$ , and define the partition function

$$Z_n(\alpha) = \sum_v c_n(v) e^{\alpha v}. \tag{2.1}$$

The corresponding free energy is  $\kappa_n(\alpha) = n^{-1} \log Z_n(\alpha)$  and it is known (Hammersley *et al* 1982) that the limit  $\kappa(\alpha) = \lim_{n \rightarrow \infty} \kappa_n(\alpha)$  exists for all  $\alpha < \infty$ . Moreover,  $\kappa(\alpha)$  is independent of  $\alpha$  for  $\alpha \leq 0$  and there is a finite and positive number  $\alpha_c$  defined by

$$\alpha_c = \sup[\alpha | \kappa(\alpha) = \kappa(0)]. \tag{2.2}$$

This singular point corresponds to the adsorption transition. For  $\alpha < \alpha_c$  the walk has a zero limiting fraction of its vertices in  $z = 0$  while for  $\alpha > \alpha_c$  the walk has a non-zero limiting fraction of its vertices in  $z = 0$ .

Next we turn to the corresponding problem in which the zeroth vertex is uncoloured and the vertices  $i = 1, 2, \dots, n$  of the walk are randomly assigned one of the two colours,  $A$  and  $B$ , independently, so that the probability that a vertex is coloured  $A$  is  $p$ . We write  $\chi_i = 1$  if the  $i$ th vertex is coloured  $A$  and  $\chi_i = 0$  if it is coloured  $B$ . We write  $\chi$  as a shorthand for the sequence  $\chi_1, \chi_2, \dots, \chi_n$ . Let  $c_n(v_A | \chi)$  be the number of  $n$ -edge walks with colouring  $\chi$ , having  $v_A$  vertices coloured  $A$  in  $z = 0$ . Define the corresponding partition function

$$Z_n(\alpha | \chi) = \sum_{v_A} c_n(v_A | \chi) e^{\alpha v_A}. \tag{2.3}$$

In the quenched version of this problem, we are interested in the behaviour of the quenched average free energy

$$\bar{\kappa}(\alpha) = \lim_{n \rightarrow \infty} \langle n^{-1} \log Z_n(\alpha | \chi) \rangle \tag{2.4}$$

where the angular brackets denote an average over all possible colourings. This problem has been investigated by Orlandini *et al* (1999) who proved the existence of the limit in (2.4) and that  $\bar{\kappa}(\alpha) = \kappa(0)$  for all  $\alpha \leq 0$ . Moreover, there exists a finite number  $\alpha_q$  defined by

$$\alpha_q = \sup[\alpha | \bar{\kappa}(\alpha) = \kappa(0)] \tag{2.5}$$

and  $\alpha_q \geq \alpha_c$ .

The problem can be approximated by the annealed version in which the order of the logarithm and average is reversed. We are then interested in the average of the partition function

$$\begin{aligned} \langle Z_n(\alpha|\chi) \rangle &= \sum_{v, v_A} c_n(v) \binom{v}{v_A} p^{v_A} (1-p)^{v-v_A} e^{\alpha v_A} \\ &= \sum_v c_n(v) \sum_{v_A} \binom{v}{v_A} (p e^\alpha)^{v_A} (1-p)^{v-v_A} \\ &= \sum_v c_n(v) e^{\gamma v} \end{aligned} \quad (2.6)$$

where  $\gamma = \log(p e^\alpha + (1-p))$ . This implies that the annealed random copolymer adsorption problem could be solved if the homopolymer version were solved. If  $p > 0$  the annealed problem has a critical point  $\alpha_a$ , defined in an analogous way to the definition of  $\alpha_q$ , and  $\alpha_a > \alpha_c$  if  $p < 1$ . The annealed and quenched average free energies are related through the arithmetic mean-geometric mean inequality and this implies that  $\alpha_a \leq \alpha_q$ . (Incidentally, this implies that  $\alpha_q > \alpha_c$ ,  $0 < p < 1$ .) Even for the simple directed models we describe below, the question of whether or not  $\alpha_a$  is equal to  $\alpha_q$  is open.

The annealed approximation is exact for  $\alpha \leq \alpha_a$  (Orlandini *et al* 1999) but becomes very poor for large values of  $\alpha$ . Morita (1964) suggested a partial annealing procedure in which moments of the distribution of colours were constrained to have their correct values. At the lowest level, for this problem, this means that we apply the constraint

$$n^{-1} \sum_{i=1}^n \chi_i = p. \quad (2.7)$$

In this paper, we shall refer to this first moment constraint as *the Morita approximation*.

This may be accomplished by building a Lagrange multiplier,  $\lambda$ , into the average of the partition function,

$$\begin{aligned} \langle Z_n(\alpha|\chi) \rangle_M &= \int d\chi_1 \dots \int d\chi_n \prod_{i=1}^n [p \delta(\chi_i - 1) \\ &\quad + (1-p) \delta(\chi_i)] \sum_{\omega_n} \exp \left\{ \alpha \sum_i \chi_i \Delta_i - \lambda \left( \sum_i \chi_i - np \right) \right\} \end{aligned} \quad (2.8)$$

where the sum over  $\omega_n$  is a sum over all  $n$ -edge walks, and  $\Delta_i = 1$  if the  $i$ th vertex is in the hyperplane  $z = 0$ , and 0 otherwise. Performing the integrations (which may be decoupled) and re-arranging gives

$$\begin{aligned} \langle Z_n(\alpha|\chi) \rangle_M &= e^{\lambda np} \sum_v c_n(v) (p e^{\alpha-\lambda} + 1 - p)^v (p e^{-\lambda} + 1 - p)^{n-v} \\ &= e^{\lambda np} (p e^{-\lambda} + 1 - p)^n Z_n(\gamma) \end{aligned} \quad (2.9)$$

where

$$\gamma = \log \left[ \frac{p e^{\alpha-\lambda} + 1 - p}{p e^{-\lambda} + 1 - p} \right]. \quad (2.10)$$

If we specialize to the case  $p = \frac{1}{2}$  (as we shall do later in the paper) this reduces to

$$\langle Z_n(\alpha|\chi) \rangle_M = e^{\lambda n/2} 2^{-n} (e^{-\lambda} + 1)^n Z_n(\gamma) \quad (2.11)$$

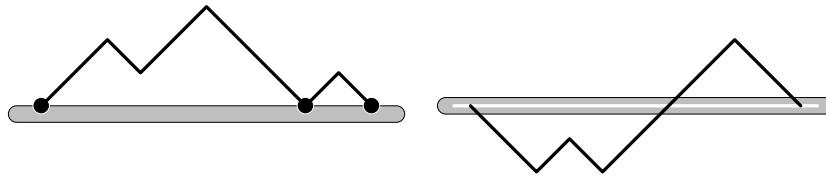


Figure 1. An example of a Dyck path and a bilateral Dyck path.

with

$$\gamma = \log \left[ \frac{e^{\alpha-\lambda} + 1}{e^{-\lambda} + 1} \right]. \quad (2.12)$$

Now we need to determine the value of  $\lambda$ . We can calculate  $\langle \chi_i \rangle$  by differentiating with respect to  $\lambda$  since

$$\frac{\partial \log \langle Z_n(\alpha|\chi) \rangle_M}{\partial \lambda} = - \left\langle \sum_i \chi_i \right\rangle + np. \quad (2.13)$$

Since we want to impose the condition that  $\langle \sum_i \chi_i \rangle = np$  this means that

$$\frac{\partial \log \langle Z_n(\alpha|\chi) \rangle_M}{\partial \lambda} = 0. \quad (2.14)$$

Setting  $p = \frac{1}{2}$ , this condition is equivalent to

$$\frac{1}{n} \frac{\partial \log Z_n(\gamma)}{\partial \gamma} = \frac{(e^\lambda - 1)(e^\alpha + e^\lambda)}{2e^\lambda(e^\alpha - 1)}. \quad (2.15)$$

In the infinite  $n$  limit, the left-hand side is zero for  $\gamma \leq \gamma_c$  and so  $\lambda = 0$  when  $\gamma \leq \gamma_c$ . Hence the Morita treatment agrees with the annealed approximation for  $\alpha \leq \alpha_a$ , where the annealed approximation is exact. To compute  $\lambda$  for super-critical values of  $\gamma$  we would need to know the free energy of the homopolymer adsorption model. For self-avoiding walks, this is unknown. Consequently, we must turn to a simpler model.

### 2.1. Directed paths

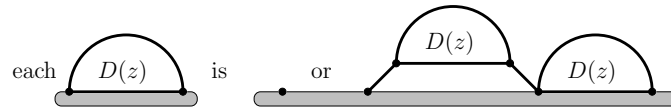
In the remainder of this paper, we shall concentrate on directed path models in two dimensions. A directed path is a walk whose vertices lie in the vertex set  $\mathbb{Z}^2$ , and whose edges are the vectors  $(1, \pm 1)$ . It simplifies the combinatorics of the situation to consider only those directed paths which start and end on the line  $y = 0$  and it is easy, using the methods of Hammersley *et al* (1982), to show that the free energy of this subset is identical to that of the full set.

For the adsorption problem, the polymer is confined to be in or on one side of the adsorbing surface, and hence we consider the subset in which the vertices of the path have non-negative  $y$ -coordinate. In the combinatorial literature, such walks are called *Dyck paths* (see e.g. Janse van Rensburg (2000)). For the localization problem, the polymer can cross the interfacial plane from one phase to the other, so the appropriate directed model is a *bilateral Dyck path*, which is a directed path that is constrained to start and end in the line  $y = 0$  and can cross this line (see figure 1).

### 3. Adsorbing Dyck paths

#### 3.1. Generating functions

Dyck paths, with generating function  $D(z)$ , can be canonically factored according to the following scheme:



The Dyck path is either a single vertex or can be factored into two Dyck paths, at the point at which the Dyck path first returns to the line  $y = 0$ .

Suppose that  $d_n$  is the number of Dyck paths with  $2n$  edges, i.e. with half-length  $n$ . The above factorization implies the following recurrence:

$$\begin{aligned} d_0 &= 1 \\ d_n &= \sum_{k=0}^{n-1} d_k d_{n-k-1}. \end{aligned} \quad (3.1)$$

Converting this to an equation for the generating function,  $D(z) = \sum_n d_n z^n$ , yields

$$D(z) = 1 + zD(z)^2 \quad (3.2)$$

where  $z$  is conjugate to the half-length of the paths. Hence

$$D(z) = \frac{1 - \sqrt{1 - 4z}}{2z}. \quad (3.3)$$

A path of half-length  $n$  contains  $2n + 1$  vertices, and so the generating function of paths counted by two variables  $z$  and  $v$ , which are conjugate to the half-length and number of vertices respectively, is given by

$$D(z, v) = vD(zv^2) = \frac{1 - \sqrt{1 - 4zv^2}}{2zv}.$$

Using the above factorization, we can also keep track of the number of vertices in the line  $y = 0$ . Suppose that  $w$  is conjugate to the number of vertices in the line  $y = 0$  (which we refer to as *visits*) and  $u$  is conjugate to the number of vertices not in  $y = 0$ . The corresponding generating function,  $F(z, u, w)$ , satisfies the following equation,

$$F(z, u, w) = 1 + zwD(z, u)F(z, u, w) \quad (3.4)$$

where we use the convention that the zeroth vertex (which always lies in the line  $y = 0$ ) does not contribute a factor of  $w$  to the generating function. Hence

$$F(z, u, w) = \frac{1}{1 - zwD(z, u)} = \frac{2u}{2u - w + w\sqrt{1 - 4zu^2}}. \quad (3.5)$$

#### 3.2. Dyck path model of the homopolymer

In the homopolymer model, only those vertices lying in the line  $y = 0$  contribute to the energy so we wish to ignore the vertices counted by  $u$ ; therefore we set  $u = 1$ . This gives the homopolymer generating function

$$H(z, w) = \frac{2}{2 - w + w\sqrt{1 - 4z}}. \quad (3.6)$$

The behaviour of the system in the thermodynamic limit ( $n \rightarrow \infty$ ) is determined by the behaviour of the limiting free energy which, in turn, is determined by the behaviour of the generating function close to its boundary of convergence (with respect to  $z$ ). The above function has two singularities: a square-root singularity when  $z = \frac{1}{4}$  and a simple pole when the denominator is equal to zero. The simple pole is determined by the solution of the following equation:

$$2 - w + w\sqrt{1 - 4z} = 0. \quad (3.7)$$

Hence there is a simple pole along the curve  $z = \frac{w-1}{w^2}$  (when  $w > 2$ ) and this singularity determines the boundary of convergence when  $w > 2$ . When  $w = 2$ , the simple pole and square root singularity coalesce and, when  $w < 2$ , the boundary of convergence is determined by the square root singularity at  $z = \frac{1}{4}$ .

The density of visits  $\rho_n(w)$  is given by the following weighted average over walks of half-length  $n$ ,

$$\begin{aligned} \rho_n(w) &= \frac{1}{2n} \frac{\sum_{\omega \in \text{walks}} V(\omega) w^{V(\omega)}}{\sum_{\omega \in \text{walks}} w^{V(\omega)}} \\ &= \frac{w}{2n} \frac{\partial}{\partial w} \log Z_n(w) \end{aligned}$$

where the number of visits in a walk,  $\omega$ , is denoted by  $V(\omega)$ , and  $Z_n(w)$  is the partition function defined by

$$Z_n(w) = \sum_{\omega \in \text{walks}} w^{V(\omega)}. \quad (3.8)$$

For large  $n$  the dominant asymptotic behaviour of  $Z_n(w)$  is determined by the dominant singularity of the generating function (see for instance Wilf (1990), especially chapter 5). If we denote the dominant singularity by  $z_c(w)$ , then

$$\lim_{n \rightarrow \infty} n^{-1} \log Z_n(w) = -\log z_c(w). \quad (3.9)$$

We are able to use this to determine the limiting density of visits:

$$\rho(w) = \lim_{n \rightarrow \infty} \rho_n(w) = \frac{-w}{2} \frac{\partial}{\partial w} \log z_c(w). \quad (3.10)$$

We note that the factor of  $\frac{1}{2}$  takes into account the fact that  $z$  is conjugate to the half-length rather than the length. The density of visits for this system is given by

$$\rho(w) = \begin{cases} 0 & \text{for } w \leq 2 \\ \frac{1}{2} \frac{w-2}{w-1} & \text{for } w > 2 \end{cases} \quad (3.11)$$

and so we interpret  $w < 2$  as a desorbed phase (since the fraction of vertices which are visits is zero) and  $w > 2$  as an adsorbed phase (since there is a positive density of visits). The point  $w = 2$  is the adsorption critical point, and it is not hard to show that the number of visits grows with the square-root of the length. We note that in the large  $w$  limit the density becomes  $\frac{1}{2}$ , which reflects the fact that only every second vertex may lie in the interface.

### 3.3. The annealed model

As an approximation to the quenched system, in the annealed model we interchange the order of the logarithm and the average over colourings, in computing the free energy. This is equivalent to fixing the conformation and then tossing a coin to decide on the colour of each vertex. This model is substantially easier to treat than the quenched problem in that it can be related directly to the homopolymer model.

The only contribution to the energy is when an  $A$ -vertex is in the line  $y = 0$ . The other vertices do not contribute to the energy. This means that the colours of the vertices which are not in  $y = 0$  are irrelevant. The colours of visit vertices, on the other hand, are important, and contribute a term  $a = e^\alpha$  to the partition function if they are coloured  $A$ , and 1 otherwise. We use the convention that the zeroth vertex is uncoloured, and we only consider the case where the probability,  $p$ , that a vertex is coloured  $A$  is  $\frac{1}{2}$ .

Consequently, the annealed partition function,  $\langle Z_n(a) \rangle$ , of walks of half-length  $n$  is given by

$$\langle Z_n(a) \rangle = \frac{\sum_{\chi} \sum_{\omega \in \text{walks}} a^{V(\omega|\chi)}}{\sum_{\chi} 1} \quad (3.12)$$

where  $V(\omega|\chi)$  is the number of  $A$ -vertices in the interface for the walk  $\omega$  which is coloured by the sequence of colours  $\chi$ . Massaging this expression gives

$$\begin{aligned} \langle Z_n(a) \rangle &= \frac{1}{2^{2n}} \sum_{\omega \in \text{walks}} \sum_{\chi} a^{V(\omega|\chi)} \\ &= \frac{1}{2^{2n}} \sum_{\omega \in \text{walks}} \sum_{\chi} \prod_{i=1}^{2n} a^{\Delta_i(\omega|\chi_i)} \\ &= \sum_{\omega \in \text{walks}} \prod_{i=1}^{2n} \sum_{\chi_i=0,1} \frac{1}{2} a^{\Delta_i(\omega|\chi_i)} \\ &= \sum_{\omega \in \text{walks}} \left( \frac{1+a}{2} \right)^{V(\omega)} \end{aligned}$$

where  $V(\omega)$  is the number of vertices in the walk  $\omega$  that lie in the line  $y = 0$ , and  $\Delta_i(\omega|\chi_i)$  is 1 if the  $i$ th vertex of  $\omega$  lies in  $y = 0$  and is coloured  $A$ , and is zero otherwise.

Hence we make the following substitutions in the generating function  $F(z, u, w)$ ,

$$\begin{aligned} u &\longrightarrow 1 \\ w &\longrightarrow (a+1)/2 \end{aligned}$$

so that the generating function of the annealed model is

$$N(z, a) = \frac{4}{3 - a + (1+a)\sqrt{1-4z}} \quad (3.13)$$

where  $a$  is conjugate to the number of  $A$ -vertices in  $y = 0$ . That is,  $a$  counts the vertices which contribute to the energy of the system.

Again we determine the thermodynamic properties of the model by studying the singularities of the generating function. There are two singularities, the square root singularity at  $z = \frac{1}{4}$  and a simple pole. By an argument similar to that used for the homopolymer case, there is a simple pole at

$$z = \frac{2(a-1)}{(a+1)^2} \quad (3.14)$$

which is dominant when  $a > 3$ . At  $a = 3$  the simple pole and the square root singularity coalesce and the square root singularity determines the boundary of convergence when  $a < 3$ .

This singularity structure implies that the density of visits is given by

$$\rho(a) = \begin{cases} 0 & \text{for } a \leq 3 \\ \frac{1}{2} \frac{a(a-3)}{(a^2-1)} & \text{for } a > 3. \end{cases} \quad (3.15)$$



We note that for large  $a$  the density of  $A$ -vertices in  $y = 0$  becomes  $\frac{1}{2}$ , which is the same as in the homopolymer case. In the quenched case, we expect the average number of  $A$ -vertices in the interface to be  $\frac{1}{4}$ , since only half of the vertices can be in the interface and, of these, only half will be  $A$ . This can be established rigorously using the methods of Orlandini *et al* (1999).

### 3.4. The Morita approximation

The Morita condition alters the annealed system by requiring that the average number of  $A$ -vertices be half of the total number of vertices. We could, in principle, require that *exactly* half of the vertices are coloured  $A$  (i.e. consider the problem in a micro-canonical ensemble), but this is significantly more difficult and we do not explore this further.

We define a new statistic on a coloured walk, called the *colour distance*, which we define to be the number of  $A$ -vertices *minus the half-length*. Hence we may rephrase the Morita condition as the condition that the average colour distance is zero. This is equivalent to introducing a Lagrange multiplier as described in section 2.

We introduce a new variable,  $L$ , conjugate to the colour distance, into the annealed partition function (recall that  $n$  is the half-length of the walk),

$$\langle Z_n(a; L) \rangle = \frac{\sum_{\chi} \sum_{\omega \in \text{walks}} a^{V(\omega|\chi)} L^{\sum_i \chi_i - n}}{\sum_{\chi} 1} \tag{3.16}$$

where  $\chi_i$  is defined to be 1 if the  $i$ th vertex is coloured  $A$  and 0 otherwise. Manipulating this equation gives

$$\begin{aligned} \langle Z_n(a; L) \rangle &= \frac{1}{L^n 2^{2n}} \sum_{\chi} \sum_{\omega \in \text{walks}} \prod_{i=1}^{2n} a^{\Delta_i(\omega|\chi_i)} L^{\chi_i} \\ &= \frac{1}{L^n} \sum_{\omega \in \text{walks}} \prod_{i=1}^{2n} \sum_{\chi_i=0,1} \frac{1}{2} a^{\Delta_i(\omega|\chi_i)} L^{\chi_i} \\ &= \frac{1}{L^n} \sum_{\omega \in \text{walks}} \left(\frac{aL+1}{2}\right)^{V(\omega)} \left(\frac{L+1}{2}\right)^{2n-V(\omega)} \end{aligned} \tag{3.17}$$

where we have again used  $V(\omega)$  to denote the number of vertices in the walk  $\omega$  that lie in  $y = 0$ , and  $\Delta_i(\omega|\chi_i)$  is 1 if the  $i$ th vertex of  $\omega$  lies in  $y = 0$  and is coloured  $A$ , and otherwise is 0, and  $\chi_i = 1$  if the  $i$ th vertex is coloured  $A$  and 0 otherwise.

We note that the generating function  $F(z, u, w)$ , discussed in section 3.1, may be rewritten as

$$F(z, u, w) = \sum_{n \geq 0} z^n \sum_{\omega \in \text{walks}} w^{V(\omega)} u^{2n-V(\omega)} \tag{3.18}$$

and so by making the following substitutions,

$$\begin{aligned} u &\longrightarrow (L+1)/2 \\ w &\longrightarrow (aL+1)/2 \\ z &\longrightarrow z/L \end{aligned}$$

we obtain the generating function,  $M(z, a; L)$ , corresponding to the partition function defined in equation (3.17):

$$M(z, a; L) = \frac{2(L+1)}{(2-a)L+1+(aL+1)(1-\sqrt{1-z(L+1)^2/L})}. \tag{3.19}$$

Again there are two singularities in this generating function. The square root singularity now occurs at  $z = \frac{L}{(1+L)^2}$ . Let us now examine carefully the zero of the denominator, determined by the solution of the equation

$$(2-a)L + 1 + (aL+1)(1 - \sqrt{1 - z(L+1)^2/L}) = 0$$

which is equivalent to

$$\sqrt{1 - z(L+1)^2/L} = \frac{aL - 2L - 1}{aL + 1}$$

where the right-hand side must be non-negative. Then the zero of the denominator is

$$z = \frac{4(a-1)L^2}{(L+1)(aL+1)^2} \quad (3.20)$$

which only exists when  $\frac{aL-2L-1}{aL+1} \geq 0$ . This is the case provided  $L \geq 1/(a-2)$ .

The Morita condition is equivalent to the condition that the average colour distance is zero, or that the average colour distance *per vertex* is zero. We may obtain the average colour distance per vertex (as  $n \rightarrow \infty$ ) in the same way as we obtained the limiting density of visits; this gives the condition

$$\frac{L}{2} \frac{\partial}{\partial L} (-\log z_c(a, L)) = 0 \quad (3.21)$$

where  $z_c(a, L)$  is the radius of convergence of the generating function.

Since the dominant singularity depends on the value of  $a$ , the value of  $L$  which satisfies the above condition will also depend on  $a$ . In particular we expect that, since there are two singularities, there will be different values of  $L$  depending on which singularity is dominant.

### 3.5. The desorbed phase

For small values of the interaction parameter  $a$ , the square-root singularity will dominate and  $z_c = \frac{L}{(1+L)^2}$ . The Morita condition implies that

$$\frac{L}{2} \frac{\partial}{\partial L} \left( -\log \left( \frac{L}{(1+L)^2} \right) \right) = \frac{L-1}{2(L+1)} = 0 \quad (3.22)$$

and so we require that  $L = 1$ . If we set  $L = 1$  then the radius of convergence is  $\frac{1}{4}$ .

We note that when  $L = 1$ , the condition for the existence of the simple pole becomes  $\frac{1}{a-2} \leq 1$  or  $a \geq 3$ . Hence, for this value of  $L$ , the simple pole does not exist until  $a \geq 3$  (i.e. at or above the annealed critical point).

Since  $z_c = \frac{1}{4}$  is independent of  $a$ , the density of  $A$ -vertices in the line  $y = 0$  is zero—confirming that this is indeed the desorbed phase.

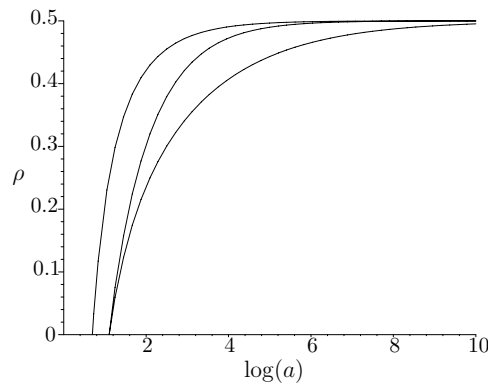
### 3.6. The adsorbed phase

For large values of  $a$ , the dominant singularity is the simple pole at  $z = \frac{4(a-1)L^2}{(L+1)(aL+1)^2}$ . The Morita condition is then

$$\frac{aL^2 - L - 2}{(aL+1)(L+1)} = 0. \quad (3.23)$$

This has two solutions,  $L = \frac{1 \pm \sqrt{1+8a}}{2a}$ , of which only  $L = \frac{1 + \sqrt{1+8a}}{2a}$  is positive (and hence physical). We need to verify that the simple pole actually exists for this value of  $L$  and indeed one can verify that

$$L = \frac{1 + \sqrt{1+8a}}{2a} \geq \frac{1}{a-2} \quad (3.24)$$



**Figure 2.** The mean density of  $A$ -visits in the line  $y = 0$  for the case where all vertices are coloured. The curves are (from top to bottom) the homopolymer, annealed and Morita models.

and, further, that the equality is only reached when  $a = 3$ . Hence this singularity does exist for  $a \geq 3$ , and  $L = 1$  at  $a = 3$ . This means that the density of visits is continuous at  $a = 3$  and so is equal to zero.

In the adsorbed phase, the radius of convergence is given by

$$z_c(a, L) = \frac{(8a^2 + 20a - 1) - (1 + 8a)^{3/2}}{2a(a - 1)^2}. \tag{3.25}$$

One can verify that the square root singularity at  $z = \frac{1}{4}$  does not dominate this singularity for this value of  $L$  (though they coalesce at  $a = 3$ ).

One can also verify that this gives a density of  $A$ -visits equal to  $\frac{1}{2}$  in the large  $a$  limit, as is the case for the homopolymer and annealed models (though all three density curves are different—see figure 2). This means that the Morita condition does not bring us closer to the quenched case at large  $a$ . The large  $a$  behaviour of the quenched case can be analysed using the methods developed for self-avoiding walks by Orlandini *et al* (1999) and shows that the limiting density of  $A$ -visits is  $\frac{1}{4}$  for the Dyck path model. This inadequacy is caused by the restriction that only alternate vertices can lie in the line  $y = 0$ , for the Dyck path model, and so the energy of the model can be optimized and still satisfy the Morita condition.

This situation may be improved by considering a related model in which, rather than colouring all vertices, only alternate vertices are coloured. Since the zeroth vertex lies on the line  $y = 0$ , and only even vertices can lie on this line, we colour only the even vertices. We define the generating function

$$F_e(z, u, w) = \sum_{n \geq 0} \sum_{\omega \in \text{walks}} w^{V_e(\omega)} u^{n - V_e(\omega)} \tag{3.26}$$

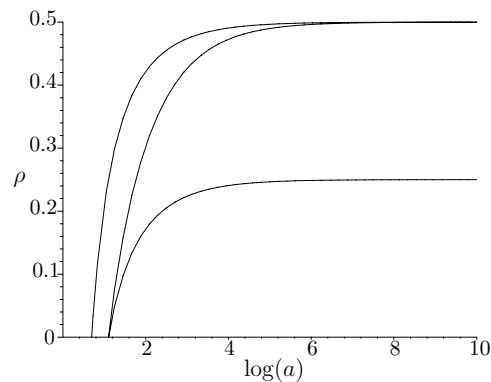
where  $V_e(\omega)$  is the number of even vertices of the walk  $\omega$  which are in  $y = 0$  and  $n - V_e(\omega)$  is the number of even vertices of  $\omega$  which are not in  $y = 0$ . Then, by the standard factorization (section 3.1),

$$F_e(z, u, w) = 1 + zwD_o(z, u)F_e(z, u, w) \tag{3.27}$$

where

$$D_o(z, v) = \frac{1 - \sqrt{1 - 4zv}}{2zv} \tag{3.28}$$

is the generating function of Dyck paths in which  $v$  counts odd vertices. We note that we need to count the number of odd vertices in the Dyck path, since these vertices then become even



**Figure 3.** The mean density of  $A$ -visits in the line  $y = 0$  for the case where only even vertices are coloured. The curves are (from top to bottom) the homopolymer, annealed and Morita models. We note that the homopolymer and annealed models are unchanged from the previous figure, while the Morita model has the same horizontal asymptote as the quenched model.

vertices in the factorization (by adding one bond at each end). Hence

$$F_e(z, u, w) = \frac{2u}{2u - w(1 - \sqrt{1 - 4zu})}. \quad (3.29)$$

The expectation of the partition function is

$$\begin{aligned} \langle Z_n(a; L) \rangle &= \frac{\sum_{\chi} \sum_{\omega} a^{V_e(\omega|\chi)} L^{\sum_{i=1}^n \chi_{2i} - n/2}}{\sum_{\chi} 1} \\ &= \frac{1}{L^{n/2}} \sum_{\omega} \left( \frac{aL+1}{2} \right)^{V_e(\omega)} \left( \frac{L+1}{2} \right)^{n-V_e(\omega)} \end{aligned} \quad (3.30)$$

where  $V_e(\omega|\chi)$  is the number of even vertices of the walk  $\omega$  with colouring  $\chi$  which are in the line  $y = 0$ . Then the generating function

$$\begin{aligned} M_e(z, a; L) &= \sum_{n \geq 0} z^n \langle Z_n(a; L) \rangle \\ &= F_e(z/\sqrt{L}, (aL+1)/2, (L+1)/2) \\ &= \frac{2(L+1)}{2(L+1) - (aL+1) + (aL+1)\sqrt{1 - 2z(L+1)/\sqrt{L}}}. \end{aligned} \quad (3.31)$$

The annealed and homopolymer models are unchanged.

A similar analysis yields the following expression for the radius of convergence in the adsorbed phase:

$$z_c = \frac{3\sqrt{3}(a-1)}{8a^{3/2}}. \quad (3.32)$$

The location of the critical point for the Morita approximation does not change when we only colour even vertices. However, the density of  $A$ -vertices in  $y = 0$  does change radically and in particular has a horizontal asymptote of  $\frac{1}{4}$  which is same as for the quenched model (see figure 3). This also implies that the limiting free energies of the two models have the same limiting slope (as a function of  $\log a$  at large  $a$ ). For this model, we find that the Morita approximation is a considerable improvement over the annealed approximation and leads to a free energy which is much closer to the quenched average free energy.

### 4. Localization of random copolymers at an interface

In this section, we investigate a bilateral Dyck path model of the localization of random copolymers at an interface. We work in two dimensions where the plane is divided (by the line  $y = 0$ ) into two semi-infinite half spaces that represent two phases which we call *oil* and *water*. The vertices of the Dyck path are independently coloured  $A$  and  $B$ , each with probability  $\frac{1}{2}$ . Vertices coloured  $A$  contribute an energy term when they are in the oil phase (i.e. have positive  $y$ -coordinate) and vertices coloured  $B$  contribute a (possibly different) energy when they are in the water phase (i.e. have negative  $y$ -coordinate). Vertices in the interface  $y = 0$  contribute no energy term.

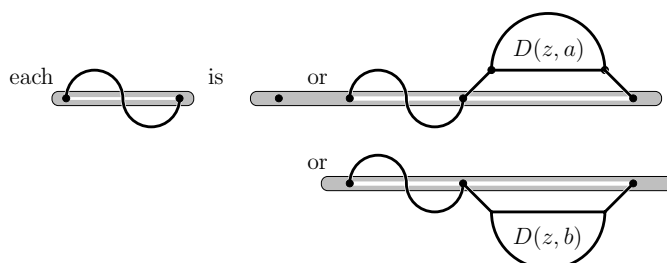
We treat an annealed version of this model and also a version in which we use the Morita approximation where the average proportion of vertices coloured  $A$  is fixed at  $\frac{1}{2}$ . In order to handle these problems, we first introduce a particular homopolymer version of the model, in which all vertices are identical but have a different contribution to the energy of the system depending on whether they are in the oil or water phases.

#### 4.1. The generating function of bilateral Dyck paths

We use  $p_n(v_o, v_w, v_i)$  to denote the number of bilateral Dyck paths of length  $2n$  with  $v_o, v_w$  vertices in the oil and water phases (respectively) and  $v_i$  vertices in the interface between these two phases. We then define the generating function of bilateral Dyck paths to be

$$B(z, a, b, c) = \sum_{n \geq 0} z^n \sum_{v_o, v_w, v_i} p_n(v_o, v_w, v_i) a^{v_o} b^{v_w} c^{v_i} \tag{4.1}$$

where the generating variables  $a, b$  and  $c$  are conjugate to  $v_o, v_w$  and  $v_i$  respectively. This generating function may be obtained from the Dyck path generating function,  $D(z, v)$ , via the following factorization:



Hence we have

$$\begin{aligned} B(z, a, b, c) &= 1 + zcB(z, a, b, c)D(z, a) + zcB(z, a, b, c)D(z, b) \\ &= \frac{1}{1 - zc(D(z, a) + D(z, b))} \\ &= \frac{1}{1 - c\left(\frac{1 - \sqrt{1 - 4za^2}}{2a} + \frac{1 - \sqrt{1 - 4zb^2}}{2b}\right)} \end{aligned} \tag{4.2}$$

where we have used the convention that the zeroth vertex contributes a factor of 1 rather than  $c$ .

#### 4.2. The homopolymer model

In this subsection, we consider a model of localization of a homopolymer in which each vertex is identical and vertices in the oil and water phases contribute energies  $\log(a)$  and

$\log(b)$  respectively. Vertices in the interface  $y = 0$  make no contribution to the energy. The generating function of this model is then

$$H(z, a, b) = B(z, a, b, 1). \quad (4.3)$$

The thermodynamic behaviour of this homopolymer system is determined by the singularities of the generating function. An examination of the generating function shows that it contains three singularities:

$$z_1 = 1/4a^2 \quad (4.4)$$

$$z_2 = 1/4b^2 \quad (4.5)$$

$$z_3 = \frac{(b-1)(a-1)(a+b-ab)}{(2ab-b-a)^2}. \quad (4.6)$$

The singularities at  $z_1$  and  $z_2$  are square root singularities which exist for all physical (i.e. real and positive) values of  $a$  and  $b$ . When the walk resides in the oil phase (respectively water phase) the singularity at  $z_1$  (respectively  $z_2$ ) is dominant. When the singularity at  $z_3$  is dominant the walk is localized at the interface (i.e. the walk crosses the interface a positive density of times).

Unlike the singularities at  $z_1$  and  $z_2$ , the singularity at  $z_3$  does not exist for all  $a$  and  $b$ ; the position of the singularity  $z_3$  is given by the zero of the denominator of  $H(z, a, b)$ :

$$b\sqrt{1-4a^2z} + a\sqrt{1-4b^2z} = a + b - 2ab. \quad (4.7)$$

This equation only has a solution (for  $z$ ) if  $a + b - 2ab \geq 0$  or  $b < \frac{a}{2a-1}$ . Hence the singularity at  $z_3$  does not exist anywhere in the first quadrant of the  $(\log a, \log b)$ -plane ( $a, b \geq 1$ ).

The dominant singularity of the generating function determines the density of vertices in each of three phases (oil, water and the interface). If a singularity  $z_c(a, b)$  is dominant, then the density of vertices in each of the three phases is given by

$$\rho_{\text{oil}} = -\frac{a}{2} \frac{\partial}{\partial a} \log z_c(a, b) \quad (4.8)$$

$$\rho_{\text{water}} = -\frac{b}{2} \frac{\partial}{\partial b} \log z_c(a, b) \quad (4.9)$$

$$\rho_{\text{interface}} = 1 - \rho_{\text{oil}} - \rho_{\text{water}} \quad (4.10)$$

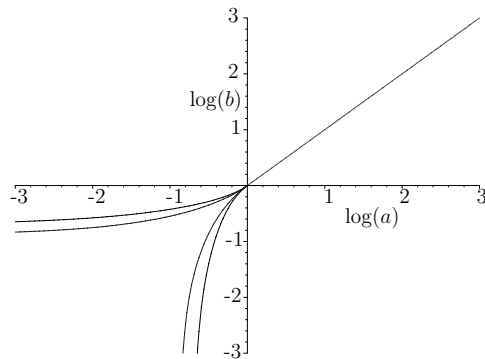
where the factor of  $\frac{1}{2}$  takes into account the fact that  $z$  is conjugate to the half-length of the walk, rather than its length.

When  $z_1$  is dominant  $\rho_{\text{oil}} = 1$ , while the density of vertices in the other two phases is zero. By  $a \leftrightarrow b$  symmetry, when  $z_2$  is dominant  $\rho_{\text{water}} = 1$ , while the density of vertices in the other two phases is zero. When  $z_3$  is dominant (which may happen in the third quadrant of the  $(\log a, \log b)$ -plane,  $(a, b) \leq (1, 1)$ ) there is a positive density of vertices in each of the three phases (and these densities are continuous functions of  $a$  and  $b$ ).

Since we have two energy contributions ( $\log a$  and  $\log b$ ) we now have phase boundaries in the  $\log a - \log b$  plane rather than isolated phase transition points. We determine these phase boundaries by examining the curves along which pairs of the singularities  $z_1, z_2$  and  $z_3$  are equal.

The intersection of  $z_1$  and  $z_2$  occurs along the line  $a = b$ . The singularity  $z_1$  dominates  $z_2$  below this line, and  $z_2$  dominates  $z_1$  above it. The intersection of  $z_1$  and  $z_3$  occurs along the curve

$$b = \frac{a(2a-1)}{1-2a+2a^2} \quad (4.11)$$



**Figure 4.** The phase boundaries of the homopolymer and annealed systems. The boundaries are the same in the first quadrant, while in the third quadrant the boundaries of the annealed system are the inner curves.

and  $z_1$  dominates below this curve, and  $z_3$  above it. By  $a \leftrightarrow b$  symmetry, there is also a phase boundary between  $z_2$  and  $z_3$  along the curve  $a = \frac{b(2b-1)}{1-2b+2b^2}$ .

Since  $z_3$  does not exist in the first quadrant of the  $(\log a, \log b)$ -plane ( $a, b > 1$ ), the only transition in this region is a first-order transition along the line  $a = b$  dividing two thermodynamic phases in which the walk resides either in the oil phase ( $a > b$ ) or the water phase ( $a < b$ ). There are no transitions in the second ( $a < 1, b > 1$ ) and fourth ( $a > 1, b < 1$ ) quadrants of the  $(\log a, \log b)$ -plane and so the walk resides in the water and oil phases (respectively).

In the third quadrant there are three thermodynamic phases. When  $b < \frac{a(2a-1)}{1-2a+2a^2}$  the walk resides in the oil phase and when  $a < \frac{b(2b-1)}{1-2b+2b^2}$  the walk resides in the water phase. Otherwise the walk is localized at the interface and this constitutes a third thermodynamic phase. The phase boundaries in the third quadrant correspond to second-order transitions. The phase boundaries show that the walk is not in the oil phase for any value of  $b$  when  $a < \frac{1}{2}$ . Similarly, the walk is not in the water phase for any value of  $a$  when  $b < \frac{1}{2}$ . The phase boundaries are shown in figure 4.

Along the line  $a = b > 1$  the density of vertices in the interface is zero (i.e. there is no interfacial phase) while in the third quadrant where  $z_3$  is the dominant singularity there is a positive density of vertices in the oil and water phases and in the interface. In other words, when  $z_3$  dominates the walk is ‘pushed’ out of the oil and water phases and is localized at the interface, while when  $a = b > 1$  the walk is attracted to both the oil and water phases and spends (essentially) no time at the interface.

#### 4.3. The annealed model

The partition function  $\langle Z_n(a, b) \rangle$  of walks of half-length  $n$  in the annealed model is given by

$$\langle Z_n(a, b) \rangle = \frac{\sum_{\chi} \sum_{\omega \in \text{walks}} a^{V_A(\omega|\chi)} b^{V_B(\omega|\chi)}}{\sum_{\chi} 1} \quad (4.12)$$

where  $V_A(\omega|\chi)$  is the number of vertices in the oil phase coloured  $A$  in the walk  $\omega$  which is coloured by the sequence of colours  $\chi$  (with a similar definition for  $V_B(\omega|\chi)$ ). Similarly, we use  $V_o(\omega)$  and  $V_w(\omega)$  to denote the number of vertices (regardless of colour) in the oil and

water phases. We recall that the zeroth vertex is uncoloured. Re-arranging (4.12) gives

$$\begin{aligned} \langle Z_n(a, b) \rangle &= \frac{1}{2^{2n}} \sum_{\omega \in \text{walks}} \sum_{\chi} a^{V_A(\omega|\chi)} b^{V_B(\omega|\chi)} \\ &= \frac{1}{2^{2n}} \sum_{\omega \in \text{walks}} (1+a)^{V_o(\omega)} (1+b)^{V_w(\omega)} \\ &= \sum_{\omega \in \text{walks}} \left( \frac{1+a}{2} \right)^{V_o(\omega)} \left( \frac{1+b}{2} \right)^{V_w(\omega)}. \end{aligned} \quad (4.13)$$

Thus the generating function of the annealed system may be obtained from  $B(z, a, b, c)$  by the substitutions

$$a \longrightarrow \frac{a+1}{2} \quad b \longrightarrow \frac{b+1}{2} \quad c \longrightarrow 1.$$

Hence the generating function,  $N(z, a, b)$ , of the annealed system is given by

$$N(z, a, b) = B\left(z, \frac{a+1}{2}, \frac{b+1}{2}, 1\right). \quad (4.14)$$

As for the homopolymer case, we determine the behaviour of this system by analysing the singularities of the generating function, and we find that there are three singularities,

$$z_1 = \frac{1}{(a+1)^2} \quad (4.15)$$

$$z_2 = \frac{1}{(b+1)^2} \quad (4.16)$$

$$z_3 = \frac{(1-b)(1-a)(3+a+b-ab)}{4(1-ab)^2} \quad (4.17)$$

where  $z_1$  and  $z_2$  correspond to the walk residing in the oil and water phases (respectively), and  $z_3$  to the system being localized at the interface. Again we find that the singularity  $z_3$  does not exist in the first quadrant of the  $\log a - \log b$  plane ( $a, b > 1$ ). The location of this singularity is given by the zero of the denominator of the generating function,

$$(1+b)\sqrt{1-(1+a)^2z} + (1+a)\sqrt{1-(1+b)^2z} = 1-ab \quad (4.18)$$

and this may be solved for  $z$  only when  $ab < 1$ , i.e.  $b < 1/a$ . Hence  $z_3$  does not exist anywhere in the first quadrant.

*4.3.1. Phases and phase boundaries.* Again we find the density of vertices in the oil and water phases and in the interface by examining the singularities of the generating function. However we now concentrate on the density of  $A$ -vertices in each phase, which we denote by  $\rho_{\text{oil}}^A$ ,  $\rho_{\text{water}}^A$  and  $\rho_{\text{interface}}^A$ .

When  $z_1$  is dominant  $\rho_{\text{oil}}^A = \frac{a}{1+a}$  and  $\rho_{\text{oil}}^B = \frac{1}{1+a}$ , while the density of vertices (of either colour) in the water phase and in the interface is zero. By  $a \leftrightarrow b$  symmetry, the density of  $B$ -vertices in the water phase when  $z_2$  is dominant is  $\frac{b}{1+b}$ , and hence  $\rho_{\text{water}}^A = \frac{1}{1+b}$ . Consequently, for large values of  $a$  we expect all vertices to become  $A$ -vertices in the oil phase, and for large values of  $b$  we expect all vertices to become  $B$ -vertices in the water phase. When  $z_3$  is dominant we find that there is a positive density of  $A$ - and  $B$ -vertices in each of the three phases and that these densities vary continuously with  $a$  and  $b$ .

The singularities  $z_1$  and  $z_2$  are equal along the line  $a = b$ , with  $z_1$  dominant below this line and  $z_2$  dominant above it. This transition is first order.



When  $ab < 1$ , the singularity at  $z_3$  exists and is equal to  $z_1$  along the curve  $b = \frac{a^2+2a-1}{a^2+1}$  corresponding to a second-order transition. We note that this value of  $b$  becomes negative when  $a < \sqrt{2} - 1$ . By symmetry, when  $ab < 1$ , the singularities  $z_2$  and  $z_3$  meet along the curve  $a = \frac{b^2+2b-1}{b^2+1}$  corresponding to a second-order transition. Again we note that this implies that when  $a, b < \sqrt{2} - 1$ , the walk is always localized.

Consequently, we find that the behaviour of the annealed model is quite similar to that of the homopolymer model. The phase boundaries for the two models are compared in figure 4.

#### 4.4. The Morita approximation

We apply the Morita approximation to the localization problem in much the same way as it was applied to the adsorption problem; we introduce a new parameter,  $L$ , to ensure that on average the fraction of vertices in the walk that are coloured  $A$  is equal to  $\frac{1}{2}$ . We define the colour-distance of a walk of length  $2n$  to be the number of  $A$ -vertices in the walk minus  $n$ . Let  $L$  be conjugate to the colour distance of the walk.

Following similar reasoning to that used to derive the annealed generating function, we make the following substitutions in the bilateral Dyck path generating function:

$$a \rightarrow \frac{La + 1}{2} \quad b \rightarrow \frac{L + b}{2} \quad c \rightarrow \frac{L + 1}{2} \quad z \rightarrow \frac{z}{L}.$$

Hence the generating function,  $M(z, a, b; L)$ , of the Morita system is given by

$$M(z, a, b; L) = B\left(\frac{z}{L}, \frac{La + 1}{2}, \frac{L + b}{2}, \frac{L + 1}{2}\right) \tag{4.19}$$

where  $a$  is conjugate to the number of  $A$ -vertices in the oil,  $b$  is conjugate to the number of  $B$ -vertices in the water phase, and  $L$  is conjugate to the colour distance of the walk. We note that this generating function possesses the following symmetry:

$$M(z, a, b; L) = M(z, b, a; 1/L). \tag{4.20}$$

We impose the requirement that the average colour distance be zero by choosing an appropriate value of  $L$ . If  $z(a, b, L)$  is the dominant singularity, then the average colour distance (per vertex) is given by

$$\langle \text{colour distance per vertex} \rangle = -\frac{L}{2} \frac{\partial}{\partial L} \log z(a, b, L) \tag{4.21}$$

where the factor of  $\frac{1}{2}$  arises since  $z$  is conjugate to the half-length. The appropriate value of  $L$  depends on the dominant singularity, and so varies with  $a$  and  $b$ .

4.4.1. Singularities of the generating function. As in the homopolymer and annealed cases, we find that this generating function has three singularities:

$$z_1 = \frac{L}{(La + 1)^2} \tag{4.22}$$

$$z_2 = \frac{L}{(L + b)^2} \tag{4.23}$$

$$z_3 = \frac{4L^2(1 - a)(1 - b)(1 + L + L^2 + La + Lb - Lab)}{(1 + L)^2(1 + La + Lb + L^2 - b - 2Lab - L^2a)^2}. \tag{4.24}$$

Again we find that  $z_1$  and  $z_2$  exist everywhere in the  $(\log a, \log b)$ -plane, but that  $z_3$  exists only in certain regions. By examining the function that defines the zeros of the denominator of  $M$ ,

we see that it only has physical solutions for  $z$  when  $L(2-ab) - ((a-1)L+1)((b-1)+L) \geq 0$ . Hence  $z_3$  only exists when

$$b \leq 1 - \frac{L(a-1)(L+1)}{L(2a-1)+1} = b_{\text{crit}}. \quad (4.25)$$

We note that for all  $L \geq 0$ , when  $a \geq 1$ ,  $b_{\text{crit}} < 1$ . Hence  $z_3$  does not exist anywhere in the first quadrant ( $a, b > 1$ ).

4.4.2. *Determining the value of  $L$ .* When  $z_1$  is dominant

$$\langle \text{colour distance per vertex} \rangle = \frac{La-1}{La+1}. \quad (4.26)$$

Consequently, if we set  $L = 1/a$  then the mean colour distance will be zero. We shall write this value of  $L$  as  $L_1 = 1/a$ . This choice gives  $z_1 = z_1(a, b, 1/a) = 1/4a$ . Similarly, when  $z_2$  is dominant

$$\langle \text{colour distance per vertex} \rangle = \frac{L-b}{L+b} \quad (4.27)$$

and we choose  $L = L_2 = b$ , which gives  $z_2 = z_2(a, b, b) = 1/4b$  (which is consistent with the symmetries of the generating function).

When  $z_3$  is dominant the resulting expression for the average colour distance is substantially more complicated, and the value of  $L = L_3$  that makes the average colour distance zero is a zero of the following quintic polynomial:

$$\begin{aligned} 2(a-1)L^5 - 3(1-a)(1+a+b-ab)L^4 - (3+2a^2b^2-5a+2b-3ab^2+2ab \\ - 2a^2b+b^2)L^3 + (2a^2b^2+2ab-5b+2a+a^2-3a^2b-2ab^2+3)L^2 \\ + 3(1-b)(1+a+b-ab)L + 2(1-b). \end{aligned} \quad (4.28)$$

Again we see that this polynomial possesses the same symmetry as the generating function under the interchange  $(a, b, L) \leftrightarrow (b, a, 1/L)$ , and hence so does  $L_3$ .

4.4.3. *The phase boundaries in the Morita approximation.* The analysis of the phase boundaries in the Morita approximation is more complicated than in the homopolymer and annealed models. This is because we must choose  $L$  according to the dominant singularity and take care that, in choosing this value of  $L$ , we do not change which singularity is dominant. Hence we use the following strategy.

We begin by assuming that we are in an area of the  $a$ - $b$  plane in which the singularity  $z_1$  is dominant (such as when  $a \gg 1$  and  $b \ll 1$ ). In this region  $L = 1/a$  and  $z_1 = 1/4a$ . Hence the density of  $A$ -vertices in the oil phase is  $\frac{1}{2}$ , and there is a zero density of vertices in the water phase and in the interface. This is consistent with the Morita condition that on average half the vertices are coloured  $A$ , and that this phase represents the walk residing entirely in the oil phase.

Similarly, when we are in a region of the  $a$ - $b$  plane in which the singularity  $z_2$  is dominant, we choose  $L = b$  making  $z_2 = 1/4b$ . This implies that there is a zero density of vertices in the oil phase and in the interface, and that the density of  $B$ -vertices in the water phase is  $\frac{1}{2}$ . Hence the density of  $A$ -vertices in the water phase is also  $\frac{1}{2}$ . This is consistent with the  $a$ - $b$  symmetry of the model.

The complicated form of  $L_3$  makes it difficult to obtain similar exact expressions for the densities of  $A$ - and  $B$ -vertices when  $z_3$  is the dominant singularity. We would expect that there is a positive density of  $A$ - and  $B$ -vertices in the oil and water phases and in the interface.

We start by finding the boundary between the  $z_1$ -dominated and  $z_3$ -dominated regions in the  $a$ - $b$  plane. We do this by starting in the region of the plane dominated by  $z_1$  and then

reduce the value of  $a$  until there is a change in the dominant singularity—this defines the boundary. We then confirm that along this boundary there is continuity in the value of  $L$ .

When  $z_1$  is dominant,  $L = L_1 = 1/a$ , and

$$z_1 = \frac{1}{4a} \quad (4.29)$$

$$z_3 = \frac{4a^2(a-1)(b-1)(1+a+ab+2a^2-a^2b)}{(a+1)^2(1-a+ab+2a^2-3a^2b)^2}. \quad (4.30)$$

A little algebra shows that  $z_1 = z_3$  (for this value of  $L$ ) along the curve

$$b = \frac{6a^3 - a^2 - 1}{(1 - 2a + 5a^2)a} \quad 0.611\,858 \dots \leq a < 1 \quad (4.31)$$

where the bounds on  $a$  are obtained by noting that  $b$  may not be negative, and that  $z_3$  does not exist in the first quadrant. This phase boundary represents a second-order transition.

Performing a similar analysis but starting in the region of the plane dominated by  $z_3$  is difficult because  $L_3$  is the solution of a quintic. However, we note that along the curve  $b = \frac{6a^3 - a^2 - 1}{(1 - 2a + 5a^2)a}$ , the polynomial which defines  $L_3$  simplifies so that  $(1 - La)$  is a factor. This indicates that the value of  $L$  is continuous across this phase boundary.

A similar analysis of the boundary between the regions of the  $a$ - $b$  plane dominated by  $z_2$  and  $z_3$  shows that there is a second-order transition along the curve

$$a = \frac{6b^3 - b^2 - 1}{(1 - 2b + 5b^2)b} \quad 0.611\,858 \dots \leq b < 1 \quad (4.32)$$

as expected on symmetry grounds.

This shows that the behaviour of the Morita model in the third quadrant of the  $a$ - $b$  plane is quite similar to that of the annealed and homopolymer models. We find that the behaviour in the first quadrant, where there is a boundary between the  $z_1$  and  $z_2$  singularities, is distinctly different.

*4.4.4. The first quadrant.* We recall that the singularity at  $z_3$  does not exist anywhere in the first quadrant ( $a, b > 1$ ) for any positive value of  $L$ .

Let us again assume that we are in a region of the  $a$ - $b$  plane in which  $z_1$  dominates and hence  $L = L_1 = 1/a$ . This means that

$$z_1 = \frac{1}{4a} \quad (4.33)$$

$$z_2 = \frac{a}{(1+ab)^2}. \quad (4.34)$$

Starting at some fixed value of  $a$  and  $b$ , and then increasing  $b$ , we find that these two singularities meet when

$$b = 2 - \frac{1}{a} \quad (4.35)$$

and  $z_1$  is dominant for  $b < 2 - \frac{1}{a}$  and  $z_2$  otherwise (for fixed  $L = L_1$ ).

Similarly, if we start from a region of the plane in which  $z_2$  is dominant and  $L = L_2 = b$ , we see that

$$z_1 = \frac{b}{(1+ab)^2} \quad (4.36)$$

$$z_2 = \frac{1}{4b}. \quad (4.37)$$

Starting from some fixed values of  $a$ , and  $b$  and increasing the value of  $a$ , we find that these two singularities meet when

$$a = 2 - \frac{1}{b} \quad (4.38)$$

where  $z_2$  dominates for  $a < 2 - \frac{1}{b}$  and  $z_1$  dominates otherwise (for this fixed value of  $L$ ).

This analysis shows that there is a special region of the  $a$ - $b$  plane,

$$a > 2 - \frac{1}{b} \quad \text{and} \quad b > 2 - \frac{1}{a} \quad (4.39)$$

that does not exist in the annealed or homopolymer models. In this region, if we decide that  $z_1$  is dominant, then our choice of  $L$  (determined by the Morita condition) makes  $z_2$  dominant. Similarly, if we decide that  $z_2$  is dominant, then the choice of  $L$  (determined by the Morita condition) makes  $z_1$  dominant. Since neither singularity can dominate the other without giving rise to a contradiction, the two singularities must be equal throughout this region.

The condition that  $z_1 = z_2$  implies that

$$L = L_{\text{crit}} = \frac{b-1}{a-1} \quad (4.40)$$

which implies that

$$z_1 = z_2 = \frac{(a-1)(b-1)}{(ab-1)^2}. \quad (4.41)$$

Since the two singularities are equal, i.e. we have two equal free energies, we interpret this region as the coexistence of two thermodynamic phases. We shall examine this from two points of view. The generating function in the Morita approximation can be written as

$$M(z, a, b; L) = f_0(z, a, b; L) + f_1(z, a, b; L)\sqrt{1-z/z_1} + f_2(z, a, b; L)\sqrt{1-z/z_2} \quad (4.42)$$

where the  $f_i$  are rational functions slowly varying close to  $z_1$  and  $z_2$ . Consequently, the partition function in the Morita approximation can be written as

$$\langle Z_n(a, b; L) \rangle = \left( X(a, b; L) \frac{z_1^{-n}}{n^{3/2}} + Y(a, b; L) \frac{z_2^{-n}}{n^{3/2}} \right) (1 + o(1)) \quad (4.43)$$

because the contribution of  $f_0$  to the partition function is exponentially smaller compared to that of  $f_1$  and  $f_2$ . We now compute the limiting average colour distance per vertex,  $\langle l \rangle$ , by taking the limit of the logarithmic derivative of the partition function with respect to  $L$ . Some algebra leads to

$$\begin{aligned} \langle l \rangle &= - \left( \frac{X}{X+Y} \frac{L}{2} \frac{\partial}{\partial L} \log z_1(a, b, L) + \frac{Y}{X+Y} \frac{L}{2} \frac{\partial}{\partial L} \log z_2(a, b, L) \right) \Big|_{L=L_{\text{crit}}} \\ &= \rho_o \langle \text{colour distance in oil} \rangle + \rho_w \langle \text{colour distance in water} \rangle \end{aligned} \quad (4.44)$$

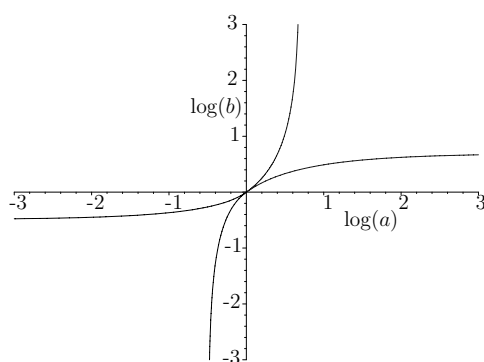
since  $\rho_o + \rho_w = 1$ , we interpret  $\rho_o$  and  $\rho_w$  as the proportions of the walk in the oil and water phases. This also shows that the density of vertices in the interface is zero.

In order to satisfy the overall Morita condition the average colour distance must be zero, which implies that

$$\rho_o \langle \text{colour distance in oil} \rangle + \rho_w \langle \text{colour distance in water} \rangle = 0 \quad (4.45)$$

which together with the fact that  $\rho_o + \rho_w = 1$  gives

$$\rho_o = \frac{ab+1-2b}{2(ab+1-a-b)} \quad \rho_w = \frac{ab+1-2a}{2(ab+1-a-b)}. \quad (4.46)$$



**Figure 5.** The phase boundaries in the Morita approximation. The asymptotes in the first quadrant are at  $\log a, \log b = \log 2 = 0.693\dots$ , while in the third quadrant they are at  $\log a, \log b = -0.491\dots$ .

One may verify that  $\rho_o = 1$  along the curve  $b = 2 - 1/a$ , and  $\rho_o = 0$  along the curve  $a = 2 - 1/b$  (as required).

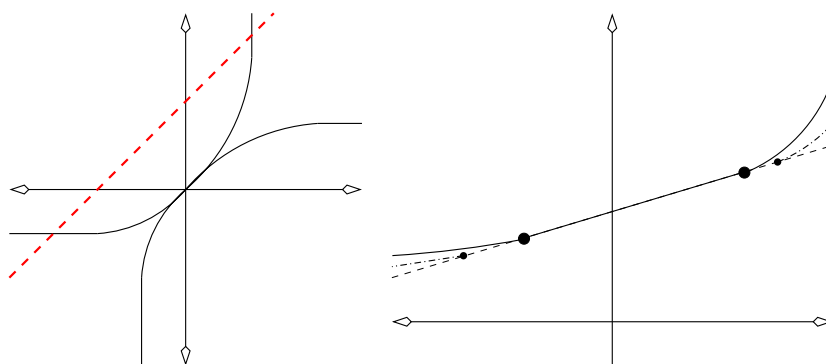
An alternative and equivalent point of view is to write the free energy of the system as the sum of two terms, each corresponding to the free energy of a macroscopic portion of the walk lying in each of the two phases. This leads to an expression equivalent to equation (4.44).

We note that the behaviour in the Morita approximation in this phase differs from that of the quenched problem. It can be shown, using the methods of Martin *et al* (2000), that there is a true localization phase in the quenched problem, in that the walk has a non-zero fraction of vertices in the interface. The Morita approximation leads to coexistence of two phases so that a macroscopic portion of the walk is in each of the oil and water phases, but the walk does not necessarily cross the interface frequently (see figure 5).

**4.4.5. Bounds on boundaries.** In order to find bounds for the locations of the quenched phase boundaries, we make use of two facts. First, the free energy in the Morita approximation is an upper bound on the quenched average free energy (Kühn 1996). Second, the free energy in the Morita approximation is equal to the quenched average free energy in the second and fourth quadrants. Using the methods of Martin *et al* (2000) we can compute the quenched average free energy in these two quadrants explicitly, and find that they are equal to those given by the Morita approximation ( $\kappa_2 = \log(4b)$  and  $\kappa_4 = \log(4a)$ ) as calculated in section 4.4. The result for the quenched average free energy essentially follows because the walk is almost entirely in the oil phase in the fourth quadrant of the  $\log a$ – $\log b$  plane and almost entirely in the water phase in the second quadrant (Martin *et al* 2000).

In figure 6 we show a sketch of the phase diagram in the Morita approximation and a distinguished line. In the same figure, we also sketch the free energy in the Morita approximation along this distinguished line. This shows that the phase boundaries in the Morita approximation are bounds on the phase boundaries for the quenched model. The figure implies that the phase boundaries are distinct, although we are not able to determine whether or not they are, in fact, coincident.

While the phase boundaries in the Morita approximation may be a faithful reflection of those of the quenched problem, the nature of this new phase in the first quadrant is different in the two models, as noted above. In the Morita approximation, the walk spends roughly half its time in each of the oil and water phases and rarely crosses the interface. In the quenched



**Figure 6.** Left: a sketch of the phase boundaries of the Morita approximation. Right: the solid line is the free energy in the Morita approximation along the dashed line shown in the phase diagram (left). The dashed line is a continuation of the linear region, and the dot-dashed line is a sketch of the quenched average free energy. This shows that the quenched phase boundaries lie to the north-east (first quadrant) and south-west (third quadrant) of those of the Morita approximation.

model, on the other hand, in the localized phase the walk spends a positive fraction of its time in each of the oil and water phases but crosses the interface frequently so that a positive fraction of its vertices lies in the interface.

## 5. Discussion

We have investigated a Morita approximation for directed walk models of random copolymer adsorption and localization, based on the combinatorial objects called Dyck paths. This approximation, unlike the annealed approximation, requires the additional condition that on average half the vertices be coloured  $A$ .

For the adsorption problem, we investigated two different Morita approximations in which either all vertices are coloured or only alternate vertices are coloured. In the first case, the differences from the annealed model are minor, and in particular the two models have the same critical point (corresponding to adsorption) and the limiting free energies have the same slope. The second approximation, on the other hand, leads to a significant difference, in that the slope of the limiting free energy is identical to that of the quenched problem, and hence has the same limiting density of  $A$ -visits. In this model, the critical point is also the same as that of the annealed problem, and we suspect that the quenched model may also have the same critical point (although we have no proof).

The localization problem is more interesting and we find that the Morita approximation displays a markedly different phase diagram from that of the annealed problem, and is much closer to that of the quenched model. In particular, the annealed model displays only two thermodynamic phases in the first quadrant, while in the Morita approximation and quenched model there are three. However, the nature of this new phase in the Morita approximation is different from that of the quenched model. In the Morita approximation, the colourings may be re-arranged so as to optimize the energy, provided that half the vertices are labelled  $A$ . This means that the Morita approximation has no control over correlations in the sequence of colours and so we expect that a typical walk in this new phase will spend roughly half of its time in the oil phase with most vertices coloured  $A$ , and half in the water phase with most vertices coloured  $B$ , and with few vertices in the interface. For most colourings in the quenched problem, the runs of  $A$ s and  $B$ s will be short and the walk will therefore cross the

interface frequently in order to optimize its energy. We also show that the phase boundaries of the Morita approximation bound the locations of the phase boundaries of the quenched model.

### Acknowledgments

The authors wish to thank Aleks Owczarek for helpful comments on the manuscript, and NSERC for financial support.

### References

- Bolthausen E and den Hollander F 1997 *Ann. Probab.* **25** 1334–66
- Brout R 1959 *Phys. Rev.* **115** 824–35
- Chuang J, Grosberg A and Kardar M 2001 *Phys. Rev. Lett.* **87** 078104
- Garel T, Huse D A, Leibler S and Orland H 1989 *Europhys. Lett.* **8** 9–13
- 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
- Gutman L and Chakraborty A K 1994 *J. Chem. Phys.* **101** 10074–91
- Hammersley J M, Torrie G and Whittington S G 1982 *J. Phys. A: Math. Gen.* **15** 539–71
- Janse van Rensburg E J 2000 *The Statistical Mechanics of Interacting Walks, Polygons, Animals and Vesicles* (Oxford: Oxford University Press)
- Kühn R 1996 *Z. Phys. B* **100** 231–42
- Maritan A, Riva M P and Trovato A 1999 *J. Phys. A: Math. Gen.* **32** L275–80
- Martin R, Causo M S and Whittington S G 2000 *J. Phys. A: Math. Gen.* **33** 7903–18
- Mazo R M 1963 *J. Chem. Phys.* **39** 1224–5
- Monari P, Stella A L, Vanderzande C and Orlandini E 1999 *Phys. Rev. Lett.* **83** 112–6
- Morita T 1964 *J. Math. Phys.* **5** 1401–5
- Orlandini E, Tesi M C and Whittington S G 1999 *J. Phys. A: Math. Gen.* **32** 469–77
- Sfatos C D and Shaknovich E I 1997 *Phys. Rep.* **288** 77–108
- Trovato A, van Mourik J and Maritan A 1998 *Eur. Phys. J. B* **6** 63–73
- Wilf H S 1990 *Generating Functionology* (London: Academic)