

Localization of random copolymers and the Morita approximation

G. Iliev[†], A. Rechnitzer[‡] and S.G. Whittington^{†‡}

[†]Department of Chemistry, University of Toronto,
Toronto, Canada M5S 3H6

[‡]Department of Mathematics and Statistics,
The University of Melbourne,
Parkville, Victoria 3010, Australia

Abstract. We discuss and analyse the Morita approximation for a number of different models of quenched random copolymer localization at the interface between two immiscible liquids. We focus on two directed models, bilateral Dyck paths and bilateral Motzkin paths, for which this approximation can be carried through analytically.

We study the form of the phase diagram and find that the Morita approximation gives phase boundaries which are qualitatively correct. This is also true when a monomer-interface interaction is included in the model. When this interaction is attractive it can lead to separation of the phase boundaries, which is also a feature of the quenched problem. We note the existence of non-analytic points on the phase boundaries which may reflect tricritical points on the phase boundaries of the full quenched average problem. In certain regions of the phase plane this approximation can be extended to the self-avoiding walk model.

1. Introduction

A random copolymer is a polymer with at least two types of monomers distributed at random along the polymer chain. We shall be concerned here with random copolymers with exactly two kinds of monomers, which we call A and B . We write $\chi_i = A$ if the i th monomer is of type A and $\chi_i = B$ if it is of type B . The χ_i are independent random variables and the probability that a monomer is A is p . Suppose that we have two immiscible liquids α and β , which we refer to as *oil* and *water* respectively for convenience. Monomers of type A prefer to be in the oil phase and monomers of type B prefer to be in the water phase. At high temperatures the polymer will delocalize into one of the two bulk phases to optimise the entropy of the system while at low temperatures it will cross the interface frequently to optimise the energy, in which case we say that the polymer is localized at the interface. For convenience we fix one end of the polymer in the interface.

One needs an underlying model for the conformational properties of the polymer chain. Various models have been considered including random walk (Maritan *et al* 1999), directed walk (Bolthausen and den Hollander 1997, Biskup and den Hollander 1999, Orlandini *et al* 2002) and self-avoiding walk models (Maritan *et al* 1999, Martin *et al* 2000, Madras and Whittington 2003, James *et al* 2003). Regardless of the chosen model we have a lattice and a hyperplane dividing it into two half-spaces. We call the hyperplane *the interface* and the two half-spaces *oil* and *water*. Label the vertices of the n -edge walk $i = 0, 1, \dots, n$ and colour each of them (except the zeroth vertex) A with probability p and B with probability $1 - p$. Given a colouring $\chi = \{\chi_1, \chi_2, \dots, \chi_n\}$ write $c_n(v_o, v_w, v_i | \chi)$ for the number of n -edge walks with colouring χ having v_o vertices coloured A in the oil, v_w vertices coloured B in the water, and $v_i + 1$ vertices in the interface. We define $A(\chi)$ (and similarly $B(\chi)$) to be the number of vertices coloured A (and similarly B) by χ . Define the partition function

$$Z_n(a, b, c | \chi) = \sum_{v_o, v_w, v_i} c_n(v_o, v_w, v_i | \chi) a^{v_o} b^{v_w} c^{v_i}, \quad (1.1)$$

where a , b , and c are parameters associated with the interaction of various types of vertices with oil, water and the interface[‡]. We are interested in the behaviour of the quenched average free energy

$$\bar{\kappa}(a, b, c) = \lim_{n \rightarrow \infty} \langle n^{-1} \log Z_n(a, b, c | \chi) \rangle, \quad (1.2)$$

where the angular brackets denote the average over all possible colourings.

When the paths are self-avoiding walks on \mathbb{Z}^d a number of results are available about the qualitative nature of the phase diagram. For $c = 1$ there are two delocalized phases D_{oil} and D_{water} in which the walk is delocalized into oil and water respectively. The free energy is then given by (Martin *et al* 2000, Madras and Whittington 2003)

$$\bar{\kappa}(a, b, 1) = \kappa_d + p \log a \quad (1.3)$$

in D_{oil} and

$$\bar{\kappa}(a, b, 1) = \kappa_d + (1 - p) \log b \quad (1.4)$$

[‡] We note that the variables a , b and c are Boltzmann factors associated with interaction parameters $\alpha = \log a$, $\beta = \log b$ and $\gamma = \log c$.

in D_{water} , where κ_d is the connective constant[§] of \mathbb{Z}^d . In the localized phase (where there is a positive density of vertices in the interface)

$$\bar{\kappa}(a, b, 1) > \kappa_d + \max[p \log a, (1 - p) \log b]. \quad (1.5)$$

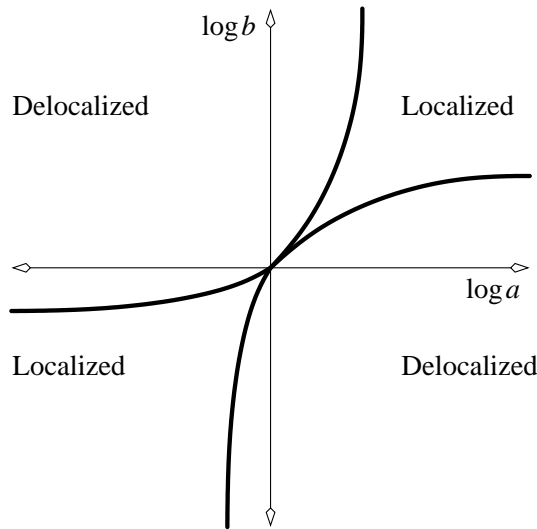


Figure 1. The qualitative form of the phase boundaries of the quenched average localization model with $p = 1/2$ and $c = 1$.

It is known that there are phase boundaries (where the free energy is singular) between the localized phase and D_{oil} and between the localized phase and D_{water} , and that these phase boundaries meet (only) at the origin. The order of the phase transition is not known rigorously (except between D_{oil} and D_{water} at $(a, b, c) = (1, 1, 1)$ where it is first order) but numerical evidence (Causo and Whittington 2003) suggests that the localization phase transition is second order in the third quadrant of the $(\log a, \log b)$ plane and higher than second order in the first quadrant. It is not known where the change from second to higher order occurs but the origin is a reasonable guess. In this paper we find a remnant of this behaviour in the Morita approximation which agrees with the location of the change of the order of the transition being at the origin.

When $c > 1$ the situation is somewhat different. It is known that there is a number $c_1 \geq 1$ such that, for all $c > c_1$, the two phase boundaries no longer share a common point and a weak upper bound is known for c_1 (Madras and Whittington 2003). Nothing is known about the order of the localization transition when $c > 1$.

Handling the quenched average free energy (1.2) is extremely difficult, even for simple path models such as directed walks. It is known (Orlandini *et al* 2002) that the annealed approximation (where the average of the logarithm of the partition function is replaced by the logarithm of the average) gives qualitatively wrong results for the phase diagram when $c = 1$. A partial annealing treatment, using an idea due to Morita (1964) (see also Mazo (1963) and Kühn (1996)), in which the average fraction

[§] We follow the definition of the connective constant given originally by Hammersley (1957) which means that κ is a reduced free energy. Often it is confused with the growth constant, $\mu = e^\kappa$.

of vertices coloured A is fixed at p , gives a phase diagram in which the phase boundaries have the correct qualitative shape. However, the behaviour in the localized phase is not faithful when $a, b > 1$. This treatment predicts a difference in behaviour between the first and third quadrants (in the $(\log a, \log b)$ plane) of the localized phase which is consistent with a change in the nature of the transition at the origin in the quenched average system.

In this paper we extend this treatment to $c \neq 1$ to derive bounds on the locations of the phase boundaries and we also attempt to find information about where the orders of the phase transitions may change. We introduce bilateral Dyck paths in Section 2 and bilateral Motzkin paths in Section 3. Examples of these objects are given in Figure 2. In Section 4 we discuss the homopolymer phase diagrams of these models.

In Section 5 we introduce the annealed and Morita approximations for a general lattice model of a copolymer and in Section 6 we apply these techniques to the bilateral Dyck and Motzkin path models. These techniques can be extended (formally) to the full self-avoiding walk model, but since we do not have sufficiently detailed information about the homopolymer model we cannot derive precise phase boundaries in the Morita approximation. This is discussed in Section 7.

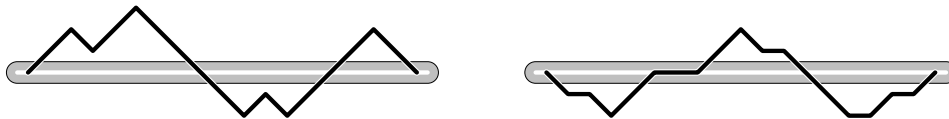


Figure 2. An example of a bilateral Dyck path (left) and a bilateral Motzkin path (right).

2. A bilateral Dyck path model

In this section we consider a Dyck path model for the underlying configurational properties of the polymer. We define a *directed path* to be a walk on the square lattice \mathbb{Z}^2 whose edges are the vectors $(1, \pm 1)$. A *Dyck path* is a directed path which starts at the origin, has no vertices with negative y -coordinate and whose last vertex is in the line $y = 0$. A *bilateral Dyck path* is a directed path starting at the origin whose last vertex is in the line $y = 0$. That is, a bilateral Dyck path can cross the line $y = 0$. The restriction that the last vertex lies in the axis simplifies the combinatorics but does not change the thermodynamics of the model; the free energies are the same.

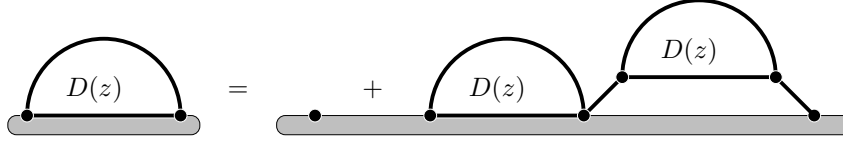
Let d_n be the number of Dyck paths with n edges, and define $d_0 = 1$. Define the generating function of Dyck paths as

$$D(z) = \sum_{n \geq 0} d_n z^n. \quad (2.1)$$

This generating function satisfies the relation

$$D(z) = 1 + z^2 D(z)^2 \quad (2.2)$$

which follows from the factorization of Dyck paths shown below.



Hence

$$D(z) = \frac{1 - \sqrt{1 - 4z^2}}{2z^2}. \quad (2.3)$$

We note that these objects always contain an even number of edges and so frequently they are enumerated according to their half-length ($z^2 \mapsto z$ in the above generating function).

To keep track of the number of vertices we can write the generating function $D(z, v)$ where z is conjugate to the length and v is conjugate to the number of vertices, and we have $D(z, v) = vD(zv)$ since each Dyck path with n edges has $n + 1$ vertices. In a similar way we can count bilateral Dyck paths keeping track of vertices with $y > 0$, $y < 0$ and $y = 0$, by the generating function $B_D(z, a, b, c)$ where z is conjugate to the length, a is conjugate to the number of vertices with $y > 0$, b is conjugate to the number of vertices with $y < 0$ and c is conjugate to the number of vertices with $y = 0$. By a similar factorization (see, for example, Orlandini *et al* (2002)):

$$B_D(z, a, b, c) = 1 + z^2 c [D(z, a) + D(z, b)] B_D(z, a, b, c) \quad (2.4)$$

so that

$$B_D = \frac{2ab}{2ab - ac - bc + ac\sqrt{1 - 4z^2b^2} + bc\sqrt{1 - 4z^2a^2}}. \quad (2.5)$$

B_D is singular when $z = z_1 = 1/2a$, when $z = z_2 = 1/2b$ and when $z = z_3$, corresponding to the situation when the denominator of B_D is zero. When z_1 is dominant the system is in D_{oil} , when z_2 is dominant the system is in D_{water} while when z_3 is dominant the system is in the localized phase. These three singularities determine the phase behaviour for the homopolymer model.

The singularity z_3 only occurs in a certain region of the (a, b, c) -space. This is discussed in Orlandini *et al* (2002). As z_3 is the solution of

$$2ab - ac - bc + ac\sqrt{1 - 4z^2b^2} + bc\sqrt{1 - 4z^2a^2} = 0 \quad (2.6)$$

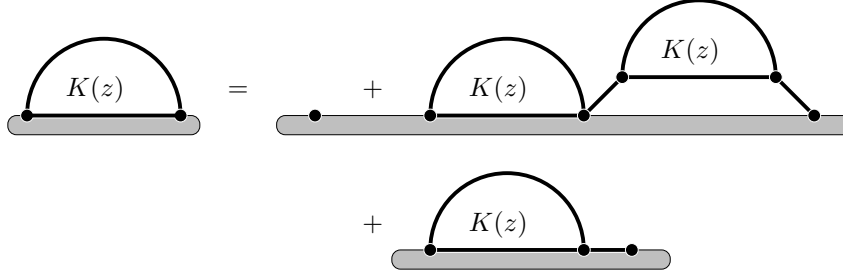
and, since the square root terms are non-negative, z_3 only occurs when

$$b \leq \frac{ac}{2a - c}. \quad (2.7)$$

We note that the above generating function also has singularities at negative z values, however one can also show that the negative singularities are never dominant over the positive singularities. Also only real positive singularities can give real free-energies and so all other singularities are considered *non-physical*.

3. A bilateral Motzkin path model

One disadvantage of the above polymer model is that the walk cannot lie in the interfacial line $y = 0$. To overcome this problem, one can allow a third kind of edge in the direction $(1, 0)$. This gives rise to a family of directed paths which are known as Motzkin paths. Like Dyck paths, they can be simply factored as follows:



where $K(z)$ is the generating function of Motzkin paths counted by the number of edges. Hence $K(z)$ satisfies

$$K(z) = 1 + z^2 K(z)^2 + zK(z) \quad (3.1)$$

and so is given by

$$K(z) = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z^2}. \quad (3.2)$$

Again to keep track of the number of vertices we form the generating function $K(z, v) = vK(zv)$, where v is conjugate to the number of vertices in the path.

Similarly if one considers the bilateral version of Motzkin paths one obtains the factorization of the generating function B_K :

$$B_K(z, a, b, c) = 1 + z^2 c \left[K(z, a) + K(z, b) \right] B_K(z, a, b, c) + zc B_K(z, a, b, c). \quad (3.3)$$

This can then be solved to give

$$B_K(z, a, b, c) = \frac{2ab}{2ab - ac - bc + bc\sqrt{(1+za)(1-3za)} + ac\sqrt{(1+zb)(1-3zb)}}. \quad (3.4)$$

This generating function has six singularities with respect to z , only three of which are real and positive and therefore physically relevant. Two of these singularities are branch cuts from the square roots at $z = z_1 = 1/3a$ and $z = z_2 = 1/3b$. The other, $z = z_3$, is a simple pole where the denominator of the above expression is zero and so is the solution of

$$2ab - ac - bc + bc\sqrt{(1+za)(1-3za)} + ac\sqrt{(1+zb)(1-3zb)} = 0. \quad (3.5)$$

Since the square root terms cannot be negative, a solution to this equation, and hence the corresponding pole, can only exist when

$$b \leq \frac{ac}{2a - c}. \quad (3.6)$$

We note that this inequality is the same for the bilateral Dyck path model.

4. The homopolymer phase diagram

The behaviour of the generating functions of the above models is determined by their dominant singularities. A change in the dominant singularity expresses itself as a change in the stable phase of the model. The phase boundaries are the loci of points at which the radius of convergence is determined by two or more dominant singularities. For example, at $(a, b, c) = (1, 1, 1)$ the three physical singularities of the Dyck and Motzkin path generating functions coalesce and are codominant. Since we know the positions of the singularities for the above two models we can readily find the phase boundaries (see, for example, Janse van Rensburg 2000 and Orlandini *et al* 2002).

Below we plot the phase diagram of the homopolymer Dyck path model at $p = 1/2$, $c = 2/3$ (left) and $c = 3/2$ (right). Note that these diagrams are not translates of each other.

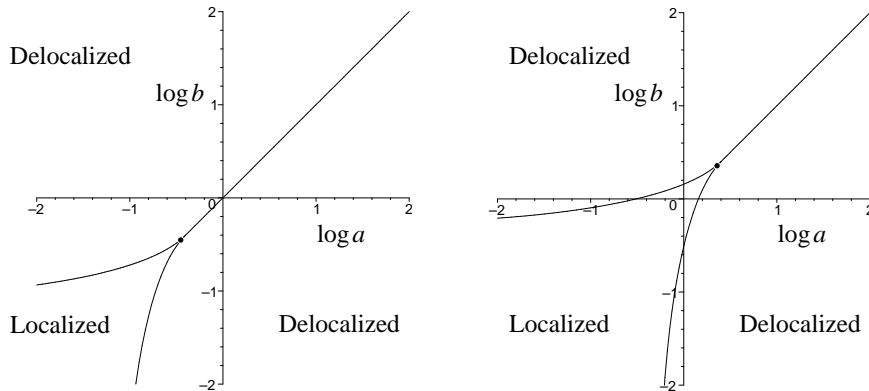


Figure 3. The phase diagram for the Dyck path model at $p = 1/2$ and $c = 2/3$ (left) and $c = 3/2$ (right).

The corresponding phase diagrams for the Motzkin path and self-avoiding walk models are qualitatively the same, though the boundaries that delineate the localized phase *are* different in the three models (see Section 7 for a discussion of the self-avoiding walk model). The point at which the three phase boundaries meet is identical for both Motzkin and Dyck paths models and is given by $a = b = c$. We believe that this is also true for self-avoiding walks on the hypercubic lattice. The best we can prove is that it happens at $(a, b, c) = (a, a, c^\dagger)$ where

$$a \leq c^\dagger \leq a \exp(2\kappa_d - \kappa_{d-1} - \sinh^{-1} \cosh \kappa_d). \quad (4.1)$$

In three dimensions this bound is approximately $a \leq c^\dagger \leq 1.63a$ and in two dimensions it is approximately $a \leq c^\dagger \leq 2.1a$. These bounds follow from results of Hammersley *et al* (1982) on the location of the adsorption transition for self-avoiding walks at a penetrable surface.

5. The annealed and Morita approximations

The overall forms of the homopolymer and quenched average copolymer phase diagrams (in Figures 1 and 3) are quite different. Consequently the homopolymer system does not adequately model the behaviour of the quenched average copolymer system. Unfortunately, even for directed models, such as those considered above, the quenched average free energy appears to be intractable (Bolthausen and den Hollander 1997, Biskup and den Hollander 1999, and Orlandini *et al* 2002). Consequently we must consider approximations to the quenched system.

Perhaps the simplest approximation that one may consider is the annealed system, in which the order of application of the expectation and the logarithm is reversed in the definition of the average free energy. That is, the limiting annealed free energy is given by

$$\kappa^a(a, b, c) = \lim_{n \rightarrow \infty} n^{-1} \log \langle Z_n(a, b, c | \chi) \rangle. \quad (5.1)$$

The geometric-arithmetic mean inequality shows that this is an upper bound on the quenched average free energy.

The following argument shows that the expectation of the partition function can be simply related back to the homopolymer partition function. The annealed partition function can be written as

$$\langle Z_n(a, b, c | \chi) \rangle = \sum_{\chi} \Pr(\chi) \sum_{\omega \in \Omega_n} a^{U_{oil}(\omega | \chi)} b^{V_{water}(\omega | \chi)} c^{W_i(\omega)}, \quad (5.2)$$

where for a given walk ω in the set, Ω_n , of walks of length n and colouring χ , $U_{oil}(\omega | \chi)$ is the number of vertices of type A in the oil, $V_{water}(\omega | \chi)$ is the number of vertices of type B in the water and $1 + W_i(\omega)$ is the number of vertices of either type lying in the interface. (The extra term comes from the fact that the zeroth vertex is always fixed in the interface.) Since the colouring is *iid* (identically and independently distributed) we only need to know the number of vertices coloured A in each of the phases, and not their positions. Therefore we can rewrite the above equation in terms of the numbers of walks with given numbers of vertices in each of the oil, water and interface. Let $c_n(u, v, w)$ be the number of walks of length n with u vertices in the oil, v vertices in the water and w vertices in the interface (excluding the zeroth vertex). Of the u vertices, u_A may be coloured A in $\binom{u}{u_A}$ ways, and similarly for the vertices in the other phases. This leads to

$$\begin{aligned} \langle Z_n(a, b, c | \chi) \rangle &= \sum_{u, v, w} \sum_{u_A, v_B} c_n(u, v, w) \binom{u}{u_A} p^{u_A} (1-p)^{u-u_A} \\ &\quad \times \binom{v}{v_B} p^{v-v_B} (1-p)^{v_B} a^{u_A} b^{v_B} c^w. \end{aligned} \quad (5.3)$$

Summing over u_A and v_B gives

$$\begin{aligned} \langle Z_n(a, b, c | \chi) \rangle &= \sum_{u, v, w} c_n(u, v, w) (pa + (1-p))^u (p + (1-p)b)^v c^w \\ &= Z_n(pa + (1-p), p + (1-p)b, c). \end{aligned} \quad (5.4)$$

This implies that the phase diagram of the annealed system is a simple linear transform of the homopolymer phase diagram, and so does not adequately approximate the quenched average phase diagram.

To understand why the annealed approximation is inadequate, one needs to examine the proportion of vertices of types A and B . In the quenched case, the expected fraction of A vertices is fixed at p , while this is not assured in the annealed model. In fact, one finds that in the D_{oil} region of the phase diagram, the fraction of A vertices goes to 1 as $a \rightarrow \infty$. The Morita approximation (Morita 1964, Orlandini *et al* 2002) avoids this problem by restricting the annealing process so that the mean fraction of vertices coloured A is fixed at p .

Following the above reasoning, we can write down an expression for the expectation of the partition function with an additional variable, L , acting as a Lagrange multiplier to enforce the Morita condition:

$$\begin{aligned} \langle Z_n(a, b, c; L | \chi) \rangle &= \sum_{\chi} \Pr(\chi) \sum_{\omega \in \Omega_n} a^{U_{oil}(\omega | \chi)} b^{V_{water}(\omega | \chi)} c^{W_i(\omega)} L^{A(\chi) - np} \\ &= L^{-np} Z_n(paL + (1-p), pL + (1-p)b, c(pL + (1-p))). \end{aligned} \quad (5.5)$$

The variable L counts the number of vertices of type A , regardless of which phase they lie in, minus np . This means that walks with the expected number of vertices of

type A are weighted by L^0 . We are then able to choose L to force the mean fraction of vertices coloured A to be p , and this value of L depends on a , b and c .

The mean fraction of vertices coloured A is then

$$\begin{aligned} \frac{1}{n} \langle A(\chi) \rangle_n &= \frac{\sum_{\chi} \Pr(\chi) \sum_{\omega \in \Omega_n} A(\chi) a^{U_{oil}(\omega|\chi)} b^{V_{water}(\omega|\chi)} c^{W_i(\omega)} L^{A(\chi)}}{n \sum_{\chi} \Pr(\chi) \sum_{\omega \in \Omega_n} a^{U_{oil}(\omega|\chi)} b^{V_{water}(\omega|\chi)} c^{W_i(\omega)} L^{A(\chi)}} \\ &= \frac{1}{n} L \frac{\partial}{\partial L} \log \langle Z_n(a, b, c; L|\chi) \rangle + p, \end{aligned} \quad (5.6)$$

where the additional term p comes from the fact that L does not count the number of A vertices, rather it counts the number of A vertices minus np . Hence in the limit $n \rightarrow \infty$ the fraction of vertices coloured A is determined by the radius of convergence, $z_c^M(a, b, c; L)$, of the generating function

$$\begin{aligned} M(z, a, b, c; L) &= \sum_{n \geq 0} \langle Z_n(a, b, c; L|\chi) \rangle z^n L^{-np} \\ &= B(zL^{-p}, paL + (1-p), pL + (1-p)b, c(pL + (1-p))), \end{aligned} \quad (5.7)$$

where $B(z, a, b, c)$ is the generating function of the homopolymer localization model. This expression is true for many models including Dyck paths, Motzkin paths and self-avoiding walks.

In the limit $n \rightarrow \infty$ we require that the mean fraction of vertices coloured A is p , and so we choose the value of L that satisfies

$$0 = -L \frac{\partial}{\partial L} \log z_c^M(a, b, c; L). \quad (5.8)$$

Since the Morita generating function is simply related to the homopolymer generating function by a change of variables, we can similarly relate z_c^M to the radius of convergence, z_c^H , of the generating function B :

$$\begin{aligned} a_H &\mapsto paL + (1-p) \\ b_H &\mapsto pL + b(1-p) \\ c_H &\mapsto c(pL + (1-p)), \end{aligned} \quad (5.9)$$

so that

$$z_c^M(a, b, c; L) = L^p z_c^H(paL + (1-p), pL + (1-p)b, c(pL + (1-p))), \quad (5.10)$$

where we have used the subscript H to denote the variables of the homopolymer system.

In the next section we give the details of the calculation of the phase diagram for the Motzkin path model in the Morita approximation for general c . The corresponding results for Dyck paths for $c = 1$ can be found in Orlandini *et al* (2002). The extension to general c is similar to that of Motzkin paths.

6. The Morita approximation for Motzkin paths

In order to apply the Morita approximation to the Motzkin path model, we need to determine as a function of a , b and c the value of L that satisfies equation (5.8). For a given value of (a, b, c) , the value of L is determined by the dominant singularity of M . The dominant singularity also determines the ‘‘physical’’ behaviour of the model.

The analysis of the homopolymer model showed the existence of three physical singularities — corresponding to three distinct phases:

$$\begin{aligned} z_1^H &= 1/3a \rightarrow \text{the } D_{oil} \text{ phase} \\ z_2^H &= 1/3b \rightarrow \text{the } D_{water} \text{ phase} \\ z_3^H &= \text{the solution of equation (3.5)} \rightarrow \text{the localized phase.} \end{aligned}$$

Equation (5.7) shows that the generating function of the Morita model is really that of the homopolymer model with transformed variables. This allows us to analyse the Morita approximation by mapping it back to the homopolymer model. Equation (5.10) shows that the above singularities become the following singularities in the Morita approximation:

$$\begin{aligned} z_1^M &= L^p/3(paL + (1-p)) \\ z_2^M &= L^p/3(pL + (1-p)b) \\ z_3^M &= L^p \times (\text{solution of (3.5) after the substitutions of (5.9)}). \end{aligned}$$

When z_1^M is dominant, we satisfy the Morita condition, equation (5.8), by setting $L = 1/a$. This implies that $z_1^M = 1/3a^p$ and $\kappa = \log 3 + p \log a$ and hence the system is in the delocalized phase D_{oil} .

Similarly, when z_2^M is dominant, equation (5.8) implies that $L = b$, and so $z_2 = 1/3b^{1-p}$ and $\kappa = \log 3 + (1-p) \log b$. The system is in the delocalized phase D_{water} .

If z_3^M is dominant, equation (5.8) can be satisfied, but the resulting value of L is a complicated algebraic function of a , b and c . This shows that there is a positive density of A -vertices in the oil and water as well as in the interface, and so constitutes a localized phase.

6.1. The delocalized phase

Let us explore the phases of the Morita approximation by determining the points (a, b, c) where z_1^M is the dominant singularity. In order to satisfy the Morita condition we must set $L = 1/a$. Some care must now be taken — while the dominance of z_1^M does imply that $L = 1/a$, it is *not* the case that $L = 1/a$ forces z_1^M to be dominant over the other two singularities. We determine the phase boundary of D_{oil} by finding the values of (a, b, c) which keep z_1^M dominant when $L = 1/a$. In Orlandini *et al* (2002), this was done by direct examination of the singularities of the Morita generating function. However here we do so by mapping the Morita model back to the homopolymer model. This approach affords us greater generality and will later enable us to say something about the corresponding self-avoiding walk model.

When $L = 1/a$, the Morita generating function is given by

$$\begin{aligned} M(z, a, b, c; 1/a) &= B(za^p, 1, p/a + (1-p)b, c(p/a + (1-p))) \\ &= B(za^p, 1, b_H, c_H), \end{aligned} \tag{6.1}$$

where we use the subscript H to denote variables of the homopolymer model. Equation (5.10) implies that we can show that z_1^M is the radius of convergence of M by examining the radius of convergence of $B(za^p, 1, b_H, c_H)$. We now need to confirm that the homopolymer generating function at the point $(1, b_H, c_H)$ is dominated by the singularity z_1^H which implies that the point $(1, b_H, c_H)$ lies in the D_{oil} phase.

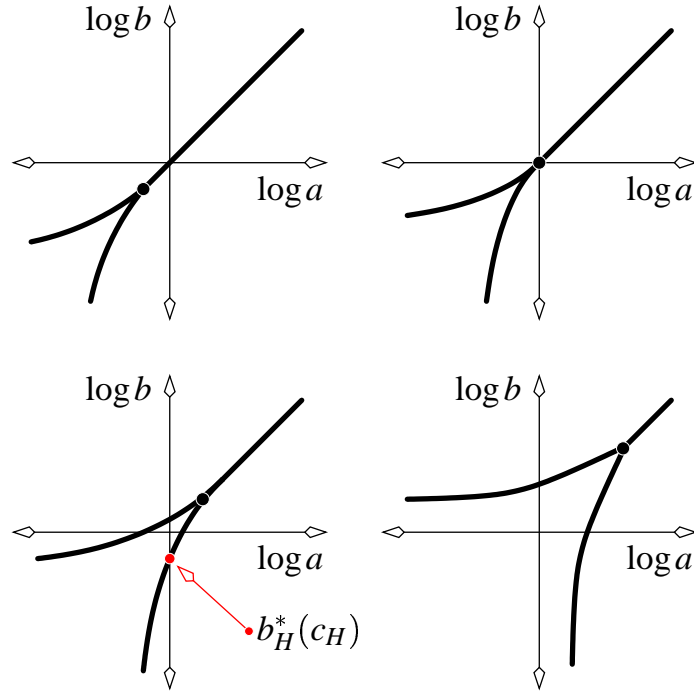


Figure 4. Homopolymer phase diagrams for $c_H < 1$ (top-left), $c_H = 1$ (top-right), $1 < c_H < c_H^*$ (bottom-left) and $c_H \geq c_H^*$ (bottom-right). When $c_H = c_H^*$ the asymptotes of the phase boundaries are the axes.

From this point on we give the details of the calculations only for the case $p = 1/2$ in order to simplify the discussion. All the calculations can be carried out for general p and we show some results for $p \neq 1/2$.

Consider the phase diagrams of the homopolymer model sketched in Figure 4. Whether or not a point $(1, b_H, c_H)$ lies inside the D_{oil} phase depends on the value of c_H . There are three cases we need to consider: $c_H \leq 1$, $1 < c_H < c_H^*$ and $c_H \geq c_H^*$, where c_H^* is the minimum value of c in the homopolymer model for which the third quadrant of the $(\log a, \log b)$ plane is free of any phase boundaries.

When $c_H \leq 1$ all points $(1, b_H)$ for $b_H < 1$ lie within D_{oil} . Using the mappings in equation (5.10) we may map these inequalities in the homopolymer variables, b_H and c_H , to inequalities in the Morita variables a , b and c :

$$\begin{aligned} 0 < c_H \leq 1 &\Rightarrow \frac{c}{2-c} \leq a < \infty, \\ 0 < b_H < 1 &\Rightarrow 0 < b < 2 - \frac{1}{a}. \end{aligned} \quad (6.2)$$

In Figure 5 we plot the region in the $(\log a, \log b)$ plane bounded by these inequalities for two values of c . The part of the curve $b = 2 - 1/a$ to the right of the line $a = c/(2-c)$ is part of the phase boundary of D_{oil} . In order to determine the remainder of the boundary, we have to consider $c_H > 1$.

Consider now the case $1 < c_H < c_H^*$ and the homopolymer phase boundary intersects the negative $\log b$ axis at the point $0 < b_H^*(c_H) < 1$ (see Figure 4, bottom-left). Similar inequalities to those above still hold, but they are now dependent on the

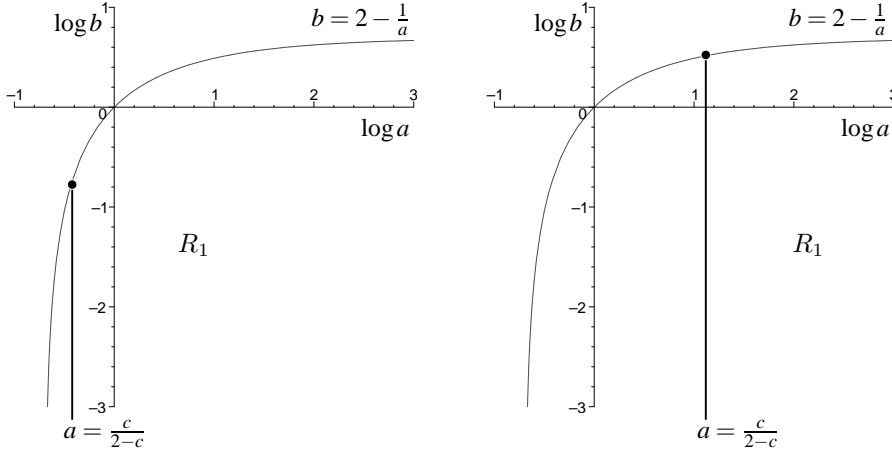


Figure 5. In the case $p = 1/2$ and $c_H \leq 1$ we show the region, R_1 , bounded by the inequalities (6.2) for $c = 4/5$ (left) and $c = 3/2$ (right). The curve is $b = 2 - 1/a$ and the vertical lines are $a = 2/3$ (left) and $a = 3$ (right).

value of $b_H^*(c_H)$.

$$1 < c_H < c_H^* \Rightarrow \frac{c}{2c_H^* - c} < a < \frac{c}{2 - c}$$

$$0 < b_H < b_H^* \Rightarrow 0 < b < 2b_H^* - \frac{1}{a}. \quad (6.3)$$

In Figure 6 we plot the region, R_2 , of the $(\log a, \log b)$ plane bounded by the inequalities (6.3) for Motzkin paths when $c = 3/2$. Note that the value of b_H^* is a function of $c_H = c(a + 1)/2a$ and depends on the underlying model. For Motzkin paths one can evaluate this function:

$$b_H^* = \frac{c_H(3 - 2c_H)}{c_H^2 - 3c_H + 3}. \quad (6.4)$$

From this we see that $c_H^* = 3/2$. Performing a similar analysis for Dyck paths one finds that $c_H^* = 2$ and the following expression for b_H^* :

$$b_H^* = \frac{c_H(2 - c_H)}{c_H^2 - 2c_H + 2}. \quad (6.5)$$

Finally, when $c_H \geq c_H^*$ there is no value of b_H such that the homopolymer model is dominated by the singularity z_1^H at $a_H = 1$. Hence the Morita approximation is not dominated by z_1^M and our choice of $L = 1/a$ is invalid.

The phase D_{oil} is the union of the two regions defined by the inequalities (6.2) and (6.3). That is, $D_{oil} = R_1 \cup R_2$. The phase D_{water} can be found by $(a, b, L, p) \leftrightarrow (b, a, 1/L, 1 - p)$ symmetry. When both R_1 and R_2 exist they meet along the line $a = c/(2 - c)$. The intersection of this line with the boundary of D_{oil} defines a point where the phase boundary is non-analytic:

$$(a, b) = \left(\frac{c}{2 - c}, 3 - \frac{2}{c} \right). \quad (6.6)$$

Hence if $c > 2$ then R_1 does not exist, while if $c < 2/3$ then R_2 does not exist. For general p the non-analytic point on the phase boundary is given by

$$(a, b) = \left(\frac{cp}{1 - c(1 - p)}, \frac{c(2 - p) - 1}{c(1 - p)} \right). \quad (6.7)$$

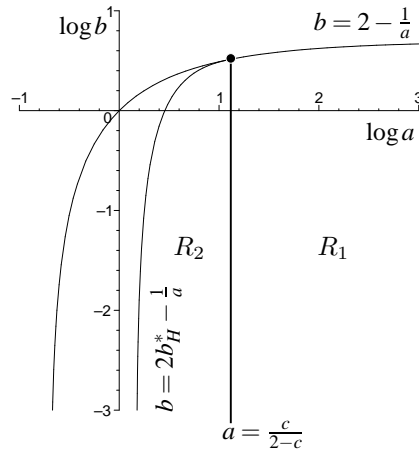


Figure 6. In the case $p = 1/2$ and $1 < c_H < c_H^*$ we show the region, R_2 , bounded by the inequalities (6.3) for $c = 3/2$. In the same figure we also show the region, R_1 , corresponding to the inequalities (6.2) for $c_H \leq 1$. We note that the diagram is similar for $c < 1$, but now R_1 and R_2 meet at $a < 1$.

And similarly if $c > \frac{1}{1-p}$ then R_1 does not exist, and if $c < \frac{1}{2-p}$ then R_2 does not exist.

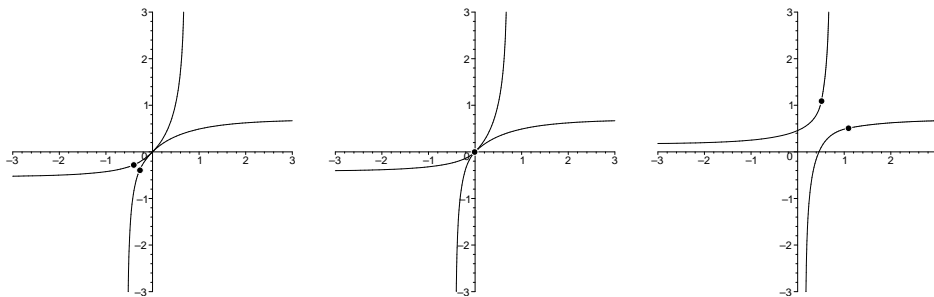


Figure 7. The boundaries of the delocalized phases for $p = 1/2$ and $c = 2/3$ (left), $c = 1$ (centre) and $c = 3/2$ (right). The non-analytic points in the boundaries are marked.

In Figure 7 we show the phase boundaries of the delocalized phases at various values of c when $p = 1/2$. In Figure 8 we show the corresponding phase boundaries for $p = 3/4$.

6.2. The localized phase

The case $c = 1$ for Dyck paths was studied previously by Orlandini *et al* 2002. They found that in the third quadrant of the $(\log a, \log b)$ plane there was a localized phase with a positive density of vertices in the interface. This agrees qualitatively with the results of Biskup and den Hollander (1999), who derived rigorous results for the path properties of a quenched average model of localization closely related to the Dyck path model we study here.

In the first quadrant of the $(\log a, \log b)$ plane the Morita approximation does not

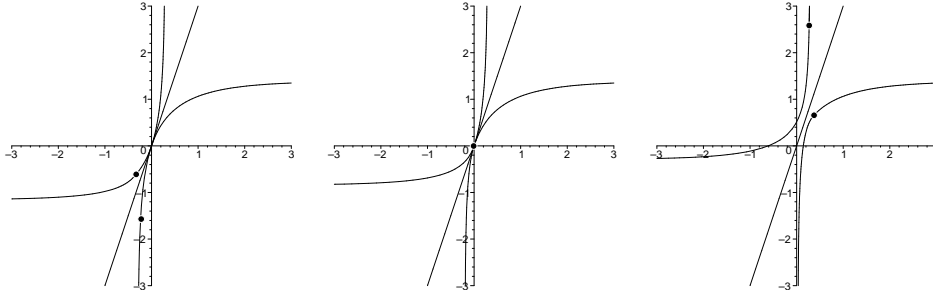


Figure 8. The boundaries of the delocalized phases for $p = 3/4$ and $c = 5/6$ (left), $c = 1$ (centre) and $c = 13/10$ (right). The non-analytic points in the boundaries are marked, and the line $\log b = 3 \log a$ is drawn as a guide to the eye.

give a faithful representation of the localized phase of the quenched average problem. While it does predict a phase in which there is a positive density of A vertices in the oil and a positive density of B vertices in the water, it does not predict a positive density of vertices in the interface. This is attributed (Orlandini *et al* 2002) to the fact that the Morita condition can be satisfied by placing half the vertices in the oil and colouring them A , and placing the other half of the vertices in the water and colouring them B . This means that the walk need not cross the interface often in order to satisfy the condition that mean fraction of A vertices is p . In contrast, in the case of a quenched colouring we find many short runs of A 's and many short runs of B 's. In order for the walk to maximise its energy by placing many of the A 's in the oil and many of the B 's in the water, it must cross the interface a positive density of times (Biskup and den Hollander 1999).

We have analysed the case $c \neq 1$ and find similar results, but with a significant new feature. We find a curve (that does not pass through $(0, 0, \log c)$) which joins the D_{oil} and D_{water} boundaries at the non-analytic points and separates the localized phase into two regions. We refer to the south-west region as L_1 and the north-east region as L_2 (see Figure 9)

We find that L_1 is a true localized region, in that there is a positive density of vertices in the interface. On the other hand in L_2 the density of vertices in the interface is 0; this is an artifact of the Morita approximation and corresponds to the results of Orlandini *et al* (2002). The boundary between these two regions is a phase boundary, in the sense that along it the free energy is non-analytic. This curve is determined by the locus of existence of z_3^M . This is also the set of points at which $z_1^M = z_2^M = z_3^M$, ie for $p = 1/2$

$$b(a, c) = 1 + \frac{(a-1)(c-1)}{a-c}. \quad (6.8)$$

The numerical results of Causo and Whittington (2003) suggest the existence of tricritical points on the phase boundaries of the quenched problem and the non-analytic points predicted by the Morita approximation may reflect these tricritical points. We do not know if the phase boundary predicted by the Morita approximation is reflected by some change in the localized phase of the quenched problem, but we do not expect a phase boundary in the thermodynamic limit. Further study is required to verify these observations.

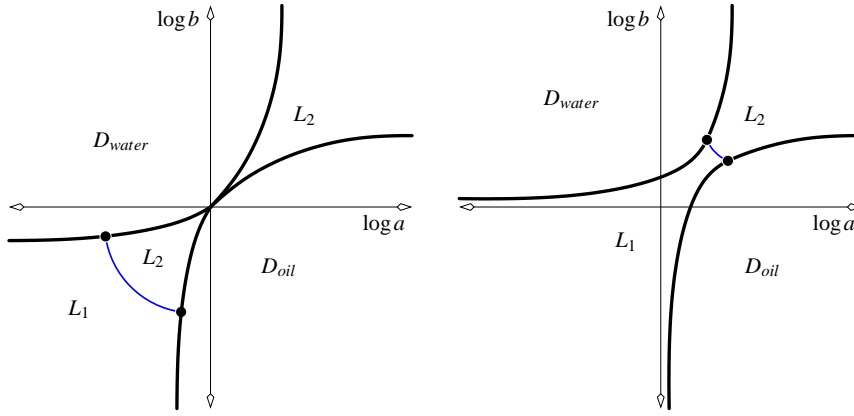


Figure 9. The localized phase for the Morita approximation at $c < 1$ (left) and $c > 1$ (right). We note that the localized phase is separated into two regions which we refer to as L_1 and L_2 . The Morita approximation predicts a new phase boundary separating L_1 and L_2 .

7. The Morita approximation for self-avoiding walks

As was the case for the Motzkin path and Dyck path models, we need to discuss the phase diagram of the localization of the self-avoiding walk model of homopolymers before we can discuss the application of the Morita approximation.

The self-avoiding walk model we consider is similar to the directed homopolymer models considered above, except that the paths are now self-avoiding walks (and so undirected) and the final vertex is not required to lie in the interface. Using the methods of Madras and Whittington (2003) one can show that the homopolymer is delocalized in the oil phase when $a \geq \max(c, b)$ and is delocalized in the water phase when $b \geq \max(c, a)$. This also implies that for a fixed value of c there is a triple point in the (a, b) plane at $(a, b) = (a^\dagger, a^\dagger)$ where $a^\dagger \leq c$. Note that this is consistent with equation (4.1).

One can also show that the phase boundary between the D_{oil} and the localized phase is concave in the $(\log a, \log b)$ plane and monotone increasing with increasing a . The boundary between D_{water} and the localized phase can be obtained by reflection in the line $\log b = \log a$. However we cannot rule out the possibility that these phase boundaries are horizontal and vertical lines (respectively). In spite of this we expect that the phase diagram is qualitatively the same as those depicted in Figure 4.

Irrespective of the details of the phase boundary we still obtain the boundary $b = 2 - 1/a$ for $a \geq c/(2 - c)$ which is the same as that found for Dyck and Motzkin paths. This is also a bound on the location of the phase boundary for the quenched problem (see Orlandini *et al* (2002) for a discussion of this point).

If we did know the precise functional form of the homopolymer phase boundary we could repeat the arguments used in Section 6 to find the complete phase boundaries for the Morita approximation to the quenched average self-avoiding walk model. Without these details we can only use the horizontal and vertical line bounds on homopolymer phase boundaries (discussed above). This gives precisely the results of Theorem 8 of Madras and Whittington (2003) which are derived in a different manner.

8. Discussion

In this paper we have extended the Morita approximation for directed walk models of copolymer localization to the case where there is an interaction (characterised by a parameter c) of the monomers with the interface. In this approximation the phase boundaries have a common point at the origin for $c \leq 1$ and are disjoint for $c > 1$. Such behaviour is expected for the full quenched problem though the best that has been proved is that for some $c_1 \geq 1$ the phase boundaries are disjoint for $c > c_1$ (Madras and Whittington 2003).

When $c = 1$ the single common point of the two phase boundaries is a non-analytic point (along the boundary). When $c \neq 1$ the point splits into two non-analytic points, one on each phase boundary. These points move along the boundaries as c changes. We believe that these points may be remnants of tricritical points where the order of the transition changes in the full quenched problem. In the Morita approximation these points are endpoints of an additional phase boundary which we do not believe to be present in the quenched problem. This boundary separates two localized phases. In one of these phases the Morita approximation gives a good representation of the localized behaviour. In the other, the Morita approximation does not give a faithful representation of the path properties of the quenched copolymer; although the copolymer has a positive density of vertices in each of the bulk phases, it has zero density of vertices in the interface.

In principle these results extend to the self-avoiding walk model of copolymer localization. We find a mapping between the homopolymer phase boundaries and the bounds predicted by the Morita approximation. Unfortunately we know very little about the precise location of the homopolymer boundaries in the repulsive regime and so we are unable to give good bounds on the copolymer phase boundaries in this regime. On the other hand, in the attractive regime we can derive bounds on the location of the phase boundaries which coincide with those derived by Madras and Whittington (2003). Their results are also based on a Morita approximation, but are derived by a somewhat different technique.

Acknowledgements

SGW wishes to thank the members of the Department of Mathematics and Statistics at the University of Melbourne, and AMSI for their kind hospitality, and similarly ADR wishes to thank the Department of Chemistry at University of Toronto. We would also like to thank E Orlandini, A L Owczarek, C E Sotoros and M Zabrocki for their helpful comments on the manuscript.

This research was partially funded by NSERC, ARC, the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems and the Australian Mathematical Sciences Institute.

References

- Biskup M and den Hollander F 1999 *Ann. Appl. Probab.* **9** 668-687
- Bolthausen E and den Hollander F 1997 *Ann. Probab.* **25** 1334-1366
- Causo M S and Whittington S G 2003 *J. Phys. A: Math. Gen.* **36** L189-L195
- Hammersley J M 1957 *Proc. Camb. Phil. Soc.* **53** 642-645
- Hammersley J M, Torrie G M and Whittington S G 1982 *J. Phys. A: Math. Gen.* **15** 539-571
- James E W, Soteris C E and Whittington S G 2003 *J. Phys. A: Math. Gen.* **36** 11187-11200
- 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** 231-242
- Madras N and Whittington S G 2003 *J. Phys. A: Math. Gen.* **36** 923-938
- Maritan A, Riva M P and Trovato A 1999 *J. Phys. A: Math. Gen.* **32** L275-L280
- Martin R, Causo M S and Whittington S G 2000 *J. Phys. A: Math. Gen.* **33** 7903-7918
- Mazo R M 1963 *J. Chem. Phys.* **39** 1224-1225
- Morita T 1964 *J. Math. Phys.* **5** 1401-1405
- Orlandini E, Rechnitzer A and Whittington S G 2002 *J. Phys. A: Math. Gen.* **35** 7729-7751