Some Motzkin path models of random and periodic copolymers

N Habibzadah, G K Iliev, A Saguia and S G Whittington
Department of Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 3H6
E-mail: swhittin@chem.utoronto.ca

Abstract. Motzkin paths are one of several directed walk models of polymers which have been used to investigate configurational properties. They have the advantage, like Dyck paths, that the models can often be solved exactly, and they have a further advantage over Dyck paths when considering physical systems involving an interface or surface. We illustrate their utility by considering several physical situations where Motzkin path models can be used. We consider the localization of periodic and random copolymers at an interface, under the influence of an applied force, and the unzipping (by an applied force) of a duplex polymer with a random monomer sequence. For the random cases the quenched system cannot be solved exactly and we use the Morita approximation but show that this gives a bound on the force required in the quenched system.

1. Introduction
Motzkin paths are a class of directed paths in two dimensions [1] which are proving useful as a model of the configurational and thermodynamic properties of polymers, in a variety of contexts [2, 3] including situations where the polymer is being micromanipulated [4, 5] by application of a force.

Consider the square lattice, $Z^2$, i.e. the integer lattice in Euclidean 2-space. We attach a coordinate system $(x_1, x_2)$ where $x_1$ and $x_2$ are integers. Motzkin paths are directed walks on $Z^2$ which obey the following constraints:

(i) The walk starts at the origin and all vertices have non-negative $x_2$-coordinate.
(ii) The possible steps are $(1, \pm 1)$ and $(1, 0)$.
(iii) The last vertex of the walk has $x_2$-coordinate equal to zero.

They are closely related to Dyck paths which are defined in a similar way but can only have steps in the $(1, \pm 1)$ directions.

If we write $c_n$ for the number of distinct Motzkin paths with $n$ edges then $c_1 = 1$, $c_2 = 2$, $c_3 = 4$, etc. If we define the generating function

$$M_0(z) = 1 + \sum_{n \geq 1} c_n z^n,$$  \hspace{1cm} (1.1)

then $M_0(z)$ satisfies the relation

$$M_0(z) = \frac{1}{1 - z} \left[1 + z^2 M_0(z)^2\right].$$  \hspace{1cm} (1.2)
One can see this by noting that the path begins as a set of edges in \( x_2 = 0 \) (including a single vertex and no edges), then leaves the line \( x_2 = 0 \), behaves as translated Motzkin path, returns to \( x_2 = 0 \) and can then be completed by adding another Motzkin path. From this we have

\[
M_0(z) = \frac{1 - z - (1 - 2z - 3z^2)^{1/2}}{2z^2}.
\]

(1.3)

\( M_0(z) \) has a singularity on the positive real axis at \( z = 1/3 \) which corresponds to the fact that \( \lim_{n \to \infty} c_n / n = 3 \). In there are about \( 3^n \) Motzkin paths with \( n \) edges, for large \( n \).

Motzkin paths can be modified by adding a weight to produce a simple model of polymer adsorption at an impenetrable surface. Suppose we regard the line \( x_2 = 0 \) as the surface at which adsorption can occur and weight Motzkin paths according to the number of vertices in \( x_2 = 0 \). Let \( c_n(v) \) be the number of Motzkin paths with \( n \) edges having \( v + 1 \) vertices in the line \( x_2 = 0 \). We say that the walk visits the line \( v \) times, or that it has \( v \) visits. Let

\[
M(x, z) = \sum_n \sum_v c_n(v)x^vz^n.
\]

(1.4)

A factorization argument shows that the generating function \( M(x, z) \) satisfies the relation

\[
(1 - xz)M(x, z) = 1 + xz^2M_0(z)M(x, z)
\]

(1.5)

so that

\[
M(x, z) = \frac{1}{1 - xz - xz^2M_0(z)}.
\]

(1.6)

\( M(x, z) \) has a square root singularity at \( z = z_1 = 1/3 \) and a singularity at \( z = z_2(x) \) corresponding to a zero of the denominator. These two singularities meet at \( x = x_c \). For \( x < x_c \) the walk has a zero fraction of its vertices in \( x_2 = 0 \) and is desorbed. For \( x > x_c \) the fraction of vertices in \( x_2 = 0 \) is positive and the walk is adsorbed.

In this paper we shall be concerned with several related problems in which Motzkin paths are used to model phase transitions in some polymer systems. The polymers will be copolymers, ie polymers with more than one type of monomer. To model this situation we shall colour the vertices of the path \( A, B \), etc, to represent the different types of monomers. In general the Hamiltonian \( H(\omega|\chi) \) will depend on both the colouring sequence \( \chi \) and on the walk \( \omega \).

2. Localization of random copolymers at an interface

If a copolymer has two types of monomers, one hydrophilic and the other hydrophobic, it can localize at the interface between bulk phases of oil and water. At low temperatures the polymer will localize at and close to the interface to optimise the numbers of monomers in their preferred phases, while at high temperatures it will delocalize into one of the bulk phases to optimise the entropy of the system [6]. In this section we investigate the localization of a random copolymer at an interface between two immiscible solvents, and consider the force required to pull the polymer from the interface into one of the bulk phases. Since the polymer can cross the interface it is natural to consider a bilateral Motzkin path. This is a directed walk with possible steps \((1, \pm 1)\) and \((1, 0)\), which starts at the origin and ends on the line \( x_2 = 0 \).

We consider a sequence of colours \( \chi = \chi_1, \chi_2, \ldots \) where \( \chi_i \in \{A, B\} \) and the colours are chosen independently, with \( A \) being chosen with probability \( p \) and \( B \) with probability \( 1 - p \). The vertices of a bilateral Motzkin path with \( n \) edges are labelled \( i = 0, 1, 2, \ldots, n \) and vertex \( i \) is associated with colour \( \chi_i \), \( i = 1, 2, \ldots, n \). We call the half-space \( x_2 > 0 \) the \( \alpha \)-phase and the half-space \( x_2 < 0 \) the \( \beta \)-phase. Suppose that \( b_n(v_A, v_B, w|\chi) \) is the number of bilateral Motzkin paths with \( n \) edges, with vertex colouring \( \chi \), having \( v_A \) vertices coloured \( A \) in the \( \alpha \)-phase, \( v_B \) vertices coloured \( B \) in the \( \beta \)-phase, and
$w + 1$ vertices (of either colour) in the interfacial line $x_2 = 0$. The partition function, at fixed colouring $\chi$, is given by

$$Z_n(a, b, c|\chi) = \sum_{v_A,v_B,w} b_n(v_A,v_B,w|\chi) a^{v_A} b^{v_B} c^w$$

where $a$, $b$ and $c$ are the Boltzmann factors $a = \exp[-\epsilon_1/kT]$, $b = \exp[-\epsilon_2/kT]$ and $c = \exp[-\epsilon_3/kT]$, $\epsilon_1$, $\epsilon_2$ and $\epsilon_3$ are suitable energies per vertex, $k$ is Boltzmann’s constant and $T$ is the absolute temperature.

Since the sequence of monomers, once chosen, is then fixed, one would like to compute the quenched average free energy

$$\bar{\epsilon}(a, b, c) = \lim_{n \to \infty} n^{-1} (\log Z_n(a, b, c|\chi))$$

where the angular brackets denote an average over the colouring sequences $\chi$. Even for simple models like Motzkin paths deriving precise information about the behaviour of the quenched average free energy is a formidable task [7, 8]. It is relatively easy to calculate the annealed free energy, where the order of the expectation and the logarithm is reversed, but this approximation is known to give a phase diagram which is qualitatively incorrect [9]. Instead we use an idea due to Morita [10] which is a partial annealing with a Lagrange multiplier incorporated to ensure that the proportion of monomers coloured $A$ is fixed at $p$. This does not ensure that higher order correlations are correct but it is known to give a qualitatively correct phase diagram and the phase boundaries in this approximation are bounds on the phase boundaries for the quenched system [9, 11]. We want to calculate the average partition function

$$\langle Z_n(a, b, c|\chi) \rangle_M = \sum_{\chi} p^{\{|A(\chi)|\}} (1-p)^{n-|A(\chi)|} \sum_{v_A,v_B,w} b_n(v_A,v_B,w|\chi) a^{v_A} b^{v_B} c^w \langle A(\chi) \rangle_{|np|}$$

where $|A(\chi)|$ is the number of vertices labelled $A$ by the first $n$ elements of $\chi$. $L$ is a Lagrange multiplier to be chosen such that $\langle |A(\chi)| - np \rangle_M = 0$.

We first compute the generating function for bilateral Motzkin paths. Let $B(a, b, c, z)$ be the generating function of bilateral Motzkin paths, where $a$ is conjugate to the number of vertices in the $\alpha$-phase, $b$ is conjugate to the number of vertices in the $\beta$-phase, $c$ is conjugate to $w$ and $z$ is conjugate to the number of edges, $n$. Then $B$ satisfies the relation

$$B(a, b, c, z) = \frac{1}{1-cz} \left[1 + cz^2 B(a, b, c, z) (aM_0(az) + bM_0(bz)) \right].$$

Hence

$$B = \frac{2ab}{2ab - ac - bc + ac\sqrt{1 - 2b^2 - 3b^2z^2} + bc\sqrt{1 - 2az - 3a^2z^2}}.$$  

We then obtain the Morita approximation by making the substitutions $a \to paL + (1-p)$, $b \to pL + (1-p)b$, $c \to pcL + c(1-p)$ and $z \to z/L^p$. The generating function $B(paL + (1-p), pL + (1-p)b, pcL + c(1-p), z/L^p)$ has three physically relevant singularities. The value of $L$ depends on the values of $a$, $b$ and $c$ and can be determined by the condition that

$$\lim_{n \to \infty} \langle |A|/n \rangle_M = p.$$  

This procedure yields three singularities $z_1 = 1/(3a^p)$, $z_2 = 1/(3b^{1-p})$ and $z_3$ which has a complicated dependence on $a$, $b$ and $c$. The set of values of $a$, $b$ and $c$ where $z_1$ ($z_2$) is dominant corresponds to delocalization into the $\alpha$-phase ($\beta$-phase). The free energy is then $\log 3 + p \log a (\log 3 + (1-p) \log b)$ and is equal to the quenched average free energy. So the Morita approximation is exact in these regimes. The regime where $z_3$ is dominant corresponds to the localized phase, and the free energy in the Morita approximation is then a bound on the quenched free energy [9]. The phase boundaries are determined by the solutions of the equations $z_i = z_j$, $i, j = 1, 2, 3, i \neq j$. It is known [9, 12] that the Morita approximation gives qualitatively correct results, not only for the locations of the phase boundaries but
Figure 1. Force-temperature curves for the interface to bulk transition for a random copolymer when \( p = 1/2 \). The parameters for the four curves (top to bottom) are (i) \( \epsilon_1 = \epsilon_2 = 1, \epsilon_3 = -0.1 \), (ii) \( \epsilon_1 = \epsilon_2 = 1, \epsilon_3 = 0 \), (iii) \( \epsilon_1 = 0.8, \epsilon_2 = 1, \epsilon_3 = -0.1 \), and (iv) \( \epsilon_1 = 0.8, \epsilon_2 = 1, \epsilon_3 = 0 \).

also for the behaviour in the interior of the localized phase in a certain region of the \((a, b, c)\)-space. For \( c = 1 \) this includes the localization phase for \( 0 < a, b \leq 1 \).

The force required to pull the polymer, localized around the interface, into one of the bulk phases has already been investigated for this model for the special case \( c = 1 \) [11]. Here we extend this work to \( c \neq 1 \). One should think of the boundary between the localized and delocalized phases as a surface in \((a, b, c)\)-space. The idea is to calculate the critical force required as a function of the temperature. Since the point \((1, 1, 1)\) corresponds to infinite temperature, the force must be calculated along a ray through the origin in the \((\alpha, \beta, \gamma)\)-space, where \( \alpha = \log a, \beta = \log b \) and \( \gamma = \log c \).

In order to incorporate a force we need to extend the model to directed walks which are the analogue of bilateral Motzkin paths but without the restriction that the last vertex has \( x_2 = 0 \). In fact we need to calculate the number of such paths which have their last vertex in \( x_2 = h \). We only need to consider \( h \geq 0 \) [13]. Let \( Q(a, b, c, y, z) \) be the generating function for these walks where \( y \) is conjugate to \( h \) and the other symbols retain their previous meaning. By a factorization argument it is easy to show that

\[
Q(a, b, c, y, z) = B(a, b, c, z)[1 + ayzM_1(y, az)]
\]  

(2.7)

where

\[
M_1(y, z) = M_0(z)[1 + yzM_1(y, z)].
\]  

(2.8)

After making the substitutions into \( Q \) for the Morita approximation we find four physically relevant singularities. Again \( L \) depends on the values of \( a, b \) and \( c \) and must be determined to satisfy (2.6). This yields the three singularities found earlier, \( z_1, z_2 \) and \( z_3 \) together with a fourth singularity which depends on \( y \),

\[
z_4 = \frac{y}{(1 + y + y^2)a^p},
\]  

(2.9)

which corresponds to the situation in which the walk is pulled into the \( \alpha \)-phase under the influence of the applied force. The critical value of the force \( f \) is found from the solution of the equation \( z_3 = z_4 \) with the substitution \( y = \exp[f/kT] \). The values of \( \epsilon_1, \epsilon_2 \) and \( \epsilon_3 \) are fixed and these values determine
the ray which is followed as the temperature is changed. We focus on $\epsilon_1, \epsilon_2 > 0$ and in fact take $\epsilon_2 = 1$ without essential loss of generality. We work in units such that $k = 1$ and results for the temperature dependence of the force when $p = 1/2$ are given in Figure 1. When $\epsilon_1 < \epsilon_2$ there is a maximum in the force-temperature curve and the force goes to zero at a finite critical temperature corresponding to the localization-delocalization transition in the absence of a force. When $\epsilon_1 = \epsilon_2$ there is no delocalization in the absence of a force. Note that the force required for delocalization is greater when $\epsilon_3 < 0$ than when $\epsilon_3 = 0$.

One can show, using the methodology of [11], that $-\log z_4$ is exactly the quenched average free energy when the polymer is pulled into the $\alpha$-phase by the applied force. By the general arguments of Kühn [14], $-\log z_3$ is a bound on the quenched free energy in the localized phase. From this it follows that the critical force calculated in the Morita approximation is an upper bound on the force required in the quenched problem.

3. Localization of periodic copolymers at an interface

In this section we address a similar situation but where the sequence of monomers is periodic. In fact we shall only treat the case where the sequence is strictly alternating, though other periodic sequences (of short period) could be handled by the same approach. The model is identical to that discussed in Section 2 except that $\chi$ is either $ABABAB\ldots$ or $BABABA\ldots$.

Let $G_{AA}(a, b, c, z)$ be the generating function for bilateral Motzkin paths starting and ending with an $A$ vertex and with the colouring being strictly alternating. Define $G_{AB}, G_{BA}$ and $G_{BB}$ in a similar way. The variables $a$, $b$ and $c$ are conjugate to $v_A$, the number of vertices coloured $A$ in the $\alpha$-phase, to $v_B$, the number of vertices coloured $B$ in the $\beta$-phase and to $w$, respectively. Let $M_i(z)$ count Motzkin paths with an odd number of edges and let $M_e(z)$ count Motzkin paths with an even number of edges. A factorization argument gives the following simultaneous equations:

$$
G_{AA} = \frac{1}{1-cz^2 z^2} [1 + cz^2 \left( M_{BA} G_{BA} + M_{BB} G_{AA} + bM_{BA} G_{BA} + bM_{BB} G_{AA} \right) + c^2 z^3 \left( aM_{AA} G_{BA} + aM_{AB} G_{AA} + M_{AA} G_{BA} + M_{AB} G_{AA} \right)] \quad (3.1)
$$
Figure 3. Comparison of the force-temperature curves for the alternating and random ($p = 1/2$) cases. The graphs on the left are all for $\epsilon_1 = \epsilon_2 = 1$ and the graphs on the right are for $\epsilon_1 = 0.8, \epsilon_2 = 1$. In each case the top pair of curves is for $\epsilon_3 = -0.2$ and the bottom pair of curves is for $\epsilon_3 = 0$. The full curve is for the alternating case and the dashed curve is for the random case.

and

$$G_{BA} = \frac{1}{1 - c^2z^2}[cz\left(1 + cz^2(M_{BA}^{(a)}G_{BA} + M_{BB}^{(a)}G_{AA} + bM_{BA}^{(b)}G_{BA} + bM_{BB}^{(b)}G_{AA})\right) + c^2z\left(aM_{AA}^{(a)}G_{BA} + aM_{AB}^{(a)}G_{AA} + M_{AB}^{(b)}G_{BA} + M_{AB}^{(b)}G_{AA}\right)],$$  

(3.2)

where

$$M_{AA}^{(a)} = M_{BB}^{(a)} = M_e(z\sqrt{a}), \quad M_{AA}^{(b)} = M_{BB}^{(b)} = M_e(z\sqrt{b}),$$

$$M_{AB}^{(a)} = M_o(z\sqrt{a}/\sqrt{a}), \quad M_{AB}^{(b)} = \sqrt{b}M_o(z\sqrt{b}),$$

$$M_{BA}^{(a)} = \sqrt{a}M_o(z\sqrt{a}), \quad M_{BA}^{(b)} = M_o(z\sqrt{b}/\sqrt{b}).$$  

(3.3)

There is a similar pair of simultaneous equations relating $G_{BB}$ and $G_{AB}$.

All four of the generating functions $G_{AA}, G_{BB}, G_{AB}$ and $G_{BA}$ have the same three physically relevant singularities $z_1 = 1/3\sqrt{a}$, $z_2 = 1/3\sqrt{b}$ and a third singularity, $z_3$ which is a complicated function of $a, b$ and $c$. When $z_1$ is dominant the system is delocalized into the $\alpha$-phase and the free energy is $\log 3 + \alpha/2$. When $z_2$ is dominant the system is delocalized into the $\beta$-phase and the free energy is $\log 3 + \beta/2$, while when $z_3$ is dominant the system is localized and the free energy is greater than $\max[\log 3 + \alpha/2, \log 3 + \beta/2]$. The phase boundaries between the localized and delocalized phases are the solutions of the equations $z_1 = z_3$ and $z_2 = z_3$. The phase boundaries in the third quadrant of the $(\alpha, \beta)$-space are shown in Figure 2, for $\log c = 0$ and $\log c = 0.2$. When $\log c = 0$ the phase boundaries for the alternating and random cases meet at the origin and the alternating phase boundaries are inside the random phase boundaries, i.e. it is easier to localize in the random case. For $\log c = 0.2$ the alternating phase boundary is also inside the random phase boundary and, in both cases, the origin in the $(\alpha, \beta)$-plane is in the interior of the localized phase. The phase boundaries for the alternating case meet in the first quadrant (for $\log c > 0$) and there is no localized phase beyond this point. For the random case the phase boundaries do not intersect when $\log c > 0$ [12].
When we include a force we need to extend the treatment to a generalization of Motzkin paths without the last vertex being required to be in \( x_2 = 0 \), as in Section 2. We write \( Q_{AA}(a, b, c, y, z) \) for the partition function corresponding to that defined in (2.7) where the walk starts and ends with an \( A \)-vertex, and we define \( Q_{AB}, Q_{BA} \) and \( Q_{BB} \) in an analogous way. Then factorization arguments give

\[
Q_{AA}(a, b, c, y, z) = G_{AA}[1 + yz\sqrt{a}M_1^+(y, z\sqrt{a})] + ayzG_{AB}M_1^+(y, z\sqrt{a})
\]

(3.4)

where

\[
M_1^+(y, z) = [M_1(y, z) + M_1(y, -z)]/2, \quad M_1^-(y, z) = M_1(y, z) - M_1^+(y, z)
\]

(3.5)

and \( M_1(y, z) \) is given by (2.8). Similar equations can be derived for \( Q_{AB} \) etc. The four generating functions \( Q_{AA} \) etc all have the same four physically relevant singularities. One is \( z_4 \) given by (2.9) with \( p = 1/2 \) and one is a complicated function of \( a, b \) and \( c \), corresponding to localization. Equating these two singularities and substituting as in Section 2 gives the temperature dependence of the force. We compare the force-temperature behaviour for the random and alternating cases in Figure 3.

4. Unzipping a random duplex copolymer

Simple models of the unzipping of duplex molecules such as DNA continue to attract attention [15]. In this section we describe how a variant of Motzkin paths, together with the Morita approximation, can be used to investigate the average force required to unzip a duplex copolymer with a random monomer sequence. (By average force we mean the average over monomer sequences, with some given constraints.)

The model that we consider is a pair of directed walks, \( \omega_1 \) and \( \omega_2 \), on the square lattice. Both walks have steps \((1, 0)\) and \((0, 1)\). Both walks start at the origin and each walk has \( n \) edges and so \( n + 1 \) vertices. We number the vertices of each walk with a randomly chosen colour \( \chi_i \), where \( \chi_i \in \{A, B, C\} \), so that the \( i \)th vertex on each walk has the same colour. We colour the \( i \)th vertex of each walk with a randomly chosen colour \( \chi_i \) where \( \chi_i \in \{A, B, C\} \), so that the \( i \)th vertex on each walk has the same colour. We write \( \chi \) as a shorthand for \( \chi_1, \chi_2, \ldots \). If \( \omega_1(i) = \omega_2(i) \) then we call the \( i \)th vertex a contact and each contact contributes an energy \( \epsilon_1, \epsilon_2 \) or \( \epsilon_3 \) according as the vertex is coloured \( A, B \) or \( C \). We write \( \omega = \{\omega_1, \omega_2\} \) for the walk-pair. The partition function for a fixed colouring sequence \( \chi \) is then given by

\[
Z_n(a, b, c|\chi) = \sum_{\omega} a^{k_A(\omega|\chi)}b^{k_B(\omega|\chi)}c^{k_C(\omega|\chi)}
\]

(4.1)

where the first sum runs over walk-pairs and \( k_A(\omega|\chi) \) is the number of \( A \)-contacts in walk-pair \( \omega \) with colouring sequence \( \chi \), etc, and \( c_n(k_A, k_B, k_C|\chi) \) is the number of walk pairs of length \( n \) with \( k_A \) \( A \)-contacts, etc, given the colouring sequence \( \chi \). \( a, b \) and \( c \) are Boltzmann factors, \( \alpha = \log a = -\epsilon_1/kT \), \( \log b = -\epsilon_2/kT \) and \( \log c = -\epsilon_3/kT \).

We call a step \((1, 0)\) an East or \( E \) step and a step \((0, 1)\) a North or \( N \) step. If we think of the possible pairs of steps which can be added to the walk-pair they are

(i) \( O \equiv (E, N) \)
(ii) \( C \equiv (N, E) \)
(iii) \( S_1 \equiv (N, N) \)
(iv) \( S_2 \equiv (E, E) \).
Figure 4. The phase boundaries for the duplex copolymer for \( p = 0 \) (inner curve), \( 1/4 \), \( 1/2 \) and \( 3/4 \) (outer curve).

Figure 5. Force-temperature curves for unzipping a duplex copolymer, for various values of \( p \): \( p = 1/8 \) (top curve), \( 1/4 \), \( 3/8 \) and \( 1/2 \) (bottom curve). In each case \( \epsilon_1/\epsilon_2 = 2/3 \).

The first element is the step being added to \( \omega_1 \) and the second element is the step being added to \( \omega_2 \).

If \( \omega_1(n) = \omega_2(n) \) then the two walks have a final common vertex. In this case the walk-pair can be mapped to a bicoloured Motzkin path [16] by considering \( O \) as a step in the North-East direction, \( C \) as a step in the South-East direction and \( S_1 \) and \( S_2 \) as two different steps in the East direction. Each contact corresponds to a visit in the mapped problem so there are three kinds of visits. In order to introduce a force we need to incorporate additional pairs of steps after the last contact and keep track of the distance apart of the \( n \)th vertices in \( \omega_1 \) and \( \omega_2 \). This corresponds to adding additional edges to the bicoloured Motzkin path and keeping track of the height above the surface line \( x_2 = 0 \). For the
The relevant singularities of $P$ are $z_1 = 1/4$, 

$$z_2 = \frac{u - w + \sqrt{w^2 - uw}}{uw}$$

and

$$z_3 = \frac{y}{u(1 + y)^2}.$$  

Suppose that the vertices $i = 1, 2, \ldots, n$ are coloured independently so that $\chi_i \in \{A, B, C\}$ and a vertex is coloured $A$ with probability $p_1$, $B$ with probability $p_2$ and $C$ with probability $p_3 = 1 - p_1 - p_2$. To treat the random case in the Morita approximation we need to incorporate two Lagrange multipliers $L_1$ and $L_2$ which will allow us to fix the first moments

$$\lim_{n \to \infty} \langle |A|/n \rangle_M = p_1 \quad \lim_{n \to \infty} \langle |B|/n \rangle_M = p_2.$$  

The average of the partition function is then

$$\langle Z_n(a, b, c|\chi) \rangle_M = \sum_{\chi} p_1^{[A(\chi)]} p_2^{[B(\chi)]} p_3^{[C(\chi)]} \sum_{k_A, k_B, k_C} c_n(k_A, k_B, k_C|\chi) a^{k_A} b^{k_B} c^{k_C} L_1^{[A] - np_1} L_2^{[B] - np_2},$$  

where $L_1$ and $L_2$ are chosen to satisfy (4.6). This implies the substitutions $u \to p_1 L_1 + p_2 L_2 + p_3$, $w \to p_1 a L_1 + p_2 b L_2 + p_3 c$ and $z \to z/(L_1^{p_1} L_2^{p_2})$ into (4.2).

We shall consider the case where $a$ and $b$ measure the strengths of two different monomer-monomer interactions and where $c$ is associated with a mismatch between monomer pairs (so we set $c = 1$). We write $p_3 = p$ and $p_1 = p_2 = (1 - p)/2$, so we take the $A$ and $B$ monomers (or monomer pairs) to occur with equal frequency. In Figure 4 we show the phase boundary between the zipped and unzipped phases (at zero force) in the $(\alpha, \beta)$-plane where $\alpha = -\epsilon_1/kT$ and $\beta = -\epsilon_2/kT$, for various values of $p$. As expected, as $p$ increases unzipping becomes easier, i.e. occurs at lower temperatures. In Figure 5 we show the force-temperature diagrams for several values of $p$ for the case $\epsilon_1/\epsilon_2 = 2/3$. (This corresponds roughly to the case for duplex DNA where adenine-thymine base pairs contribute two hydrogen bonds and cytosine-guanine base pairs contribute three hydrogen bonds.) At fixed $T$ the critical force is a monotone decreasing function of $p$.

5. Low temperature behaviour

One can try to understand the low temperature behaviour of the various force-temperature curves by constructing an approximate theory which is either exact in the $T \to 0$ limit or gives a good approximation in this regime. First consider the case where the polymer is being pulled out of the localized phase into a bulk phase (Sections 2 and 3). We consider the random copolymer under the
influence of a force $f$ with $n - m$ monomers in or near the interface and $m$ monomers which have been pulled out into the bulk phase. Since the part of the polymer in the bulk phase is under tension we ignore any entropy associated with this part of the polymer. There are several contributions to the free energy, including:

(i) an elastic energy term from the part of the polymer extending into the bulk,
(ii) an non-elastic energy term for this part of the polymer,
(iii) an energy term for the part of the polymer in and near the interface, and
(iv) an entropy term for the part of the polymer near the interface.

First take $\epsilon_1, \epsilon_2 > 0$ and $\epsilon_3 < 0$ so that the ground state is non-degenerate and there is no ground state entropy. We can write the free energy, $F_n$, approximately as

$$F_n = -mf + mp\epsilon_1 + (n - m)\epsilon_3.$$  \hspace{1cm} (5.1)

Differentiating with respect to $m$ and setting the derivative equal to zero gives a critical force $f_0$ of

$$f_0 = p\epsilon_1 - \epsilon_3.$$  \hspace{1cm} (5.2)

If $f > f_0$ then the free energy is dominated by the case $m = n$ while if $f < f_0$ the free energy is dominated by the case $m = 0$, and the polymer is completely adsorbed. (Recall that this is essentially a zero temperature argument.) The above argument suggests that the limiting derivative $\partial f_0 / \partial T$ is zero in the $T \to 0$ limit.

To extend the argument to non-zero temperature we need to consider states close to the ground state. Suppose that $|\epsilon_3| < \min|\epsilon_1, \epsilon_2|$ so that lowest lying states are those in which an $A$-vertex is in the $\beta$-phase or a $B$-vertex is in the $\alpha$-phase. In each case the energy penalty is $|\epsilon_3|$. This gives the following
approximation for the free energy

\[ F_n = -mf + mpe_1 - T \log \left( \sum_{l=0}^{n-m} \left( \frac{n-m}{l} \right) e^{-le_3/T} \right) \]

which gives the following expression for the critical force

\[ f_0 = pe_1 - \epsilon_3 + T \log(1 + e^{\epsilon_3/T}). \] (5.4)

This argument works equally well for the random and alternating cases (if we put \( p = 1/2 \)). In Figure 6 we compare the low temperature behaviour of the two cases with the prediction of (5.4). The agreement at low temperatures is remarkably good.

If we look at the case \( \epsilon_1, \epsilon_2 > 0 \) and \( \epsilon_3 = 0 \) the situation is rather different. First think of the alternating case. The ground state is degenerate since every \( A \) vertex (\( B \) vertex) can be either in the \( \beta \)-phase (\( \alpha \)-phase) or in the line \( x_2 = 0 \). In the ground state every vertex is in one of the lines \( x_2 = 0, \pm 1 \). If the \( m \)th vertex is out of the line \( x_2 = 0 \) then the \((m + 1)\)th vertex must be in this line. If the \( m \)th vertex is in this line the \((m + 1)\)th vertex can be in or out of the line. This gives a ground state entropy of \( \log[(1 + \sqrt{5})/2] \) and the low temperature behaviour of the free energy should be given by

\[ F_n = -mf + (n - m)\left( \frac{1-p}{2} \right) (\epsilon_1 + \epsilon_2) - T(n - m)[(1 - p) \log 2 + p \log 3] \] (5.5)

so the critical force is

\[ f_0 = p\epsilon_1 - \epsilon_3 + T \log(1 + e^{\epsilon_3/T}). \] (5.6)

Hence, at low temperatures the critical force should approach \( p\epsilon_1 \) linearly as \( T \to 0 \). This is precisely the behaviour of the alternating model, as can be seen in Figure 6 where the force-temperature curve for the alternating case approaches the ordinate with the slope predicted by (5.6) as \( T \to 0 \). For the random case the above argument underestimates the ground state entropy. This is easy to see since the colouring will contain strings of \( A \)s (\( B \)s) which can form a loop in the \( \beta \)-phase (\( \alpha \)-phase) at no energy cost, but giving a contribution to the entropy. The limiting slope for the random case (see Figure 6) is higher than predicted by (5.6).

We can treat the low temperature behaviour of the duplex copolymer similarly. We restrict our attention here to the case \( p_3 = p, p_1 = p_2 = (1 - p)/2, \epsilon_1, \epsilon_2 < 0, \epsilon_3 = 0 \). At low temperature the free energy is approximately

\[ F_n = -mf + (n - m)\left( \frac{1-p}{2} \right) (\epsilon_1 + \epsilon_2) - T(n - m)[(1 - p) \log 2 + p \log 3] \] (5.7)

for small \( p \), so the critical force is given by

\[ f_0 = -\left( \frac{1-p}{2} \right) (\epsilon_1 + \epsilon_2) + T[(1 - p) \log 2 + p \log 3]. \] (5.8)

This gives the zero temperature force (by substituting \( T = 0 \)) which agrees with the results of Section 4, and shows that the force-temperature curve has positive slope at \( T = 0 \). The prediction is compared with the results of the full calculation in Figure 7. The limiting slope is in good agreement with the prediction from (5.8) for small values of \( p \).
Figure 7. Force-temperature curves for unzipping a random duplex copolymer at low temperatures. The three pairs of curves are for $p = 0$ (top), $p = 0.2$ and $p = 0.4$ (bottom). In each case $\epsilon_1 = \epsilon_2 = -1$ and $\epsilon_3 = 0$. The full curves are from the treatment in Section 4 and the dashed lines are from (5.8).

6. Discussion
There is considerable current interest in the use of directed walk models to study certain phase transitions in homopolymers and copolymers, such as the adsorption and localization transitions [6]. Motzkin paths are a simple class of directed walks with many advantages. They are almost as easy to handle as Dyck paths and, for adsorption and localization problems, they have the advantage that they can lie entirely in the line which is the obvious candidate to play the role of the surface or interface. By colouring the vertices, either randomly or periodically, they can be adapted to model copolymers. In Sections 2 and 3 we have shown how they can be used to study localization in random and periodic copolymers. For the random case one would like to compute the quenched average free energy but this is currently impossible even for such simple models as Motzkin paths [6, 7]. Instead we have treated this problem in the Morita approximation [10] which is a partial annealing procedure where the first moment is constrained to have its appropriate value. We have focused on the prediction of the temperature dependence of the critical force required to pull the polymer out of the interface (in the localized phase) and into one of the bulk solvents. We have compared the results for the random and periodic cases.

In Section 4 we considered a simple model for unzipping a duplex polymer by an applied force. We focused on the case of a random copolymer, as a very crude model of duplex DNA. This model can be mapped to an adsorption problem (bicoloured Motzkin paths [16]) and we treated this model using the Morita approximation.

The force-temperature curves can show remarkable richness. The critical force can go through a maximum as the temperature increases, and this is seen both in the localization and unzipping models. The initial rise is associated with a loss of entropy when the polymer is subject to a force. At very low temperatures the behaviour can be understood by using simple arguments (see Section 5) where the entropy in the ground state plays an important role.

Acknowledgments
The authors would like to thank NSERC of Canada for financial support.
References

[1] Lando S K 2003 Lectures on Generating Functions (Providence: American Mathematical Society)
[2] Vanderzande C 1998 Lattice Models of Polymers (Cambridge: Cambridge University Press)
[3] Janse van Rensburg E J 2003 J. Phys. A: Math. Gen. 36 R11–R61
[4] Ortiz C and Hadziioannou G 1999 Macromolecules 32 780–787
[5] Bemis J E, Akhremitev B B and Walker G C 1999 Langmuir 15 2799–2805
[6] Soteros C E and Whittington S G 2004 J. Phys. A: Math. Gen. 37 R279-R325
[7] Bolthausen E and den Hollander F 1999 Ann. Prob. 25 1334–1366
[8] Bodineau T and Giacomin G 2004 J. Stat. Phys. 117 801–818
[9] Orlandini E, Rechnitzer A and Whittington S G 2002 J. Phys. A: Math. Gen. 35 7729–7751
[10] Morita T 1964 J. Math. Phys. 5 1401–1405
[11] Iliev G, Orlandini E and Whittington S G 2004 Eur. Phys. J. B 40 63–71
[12] Iliev G, Rechnitzer A and Whittington S G 2005 J. Phys. A: Math. Gen. 38 1209-23
[13] Orlandini E and Whittington S G 2004 J. Phys. A: Math. Gen. 37 5305-5314
[14] Kühn R 1996 Z. Phys. B 100 231–42
[15] Richard C and Guttmann A J 2004 J. Stat. Phys. 115 925–947
[16] Janse van Rensburg E J 2000 The statistical mechanics of interacting walks, polygons, animals and vesicles (Oxford: Oxford University Press)