Journal of Statistical Physics, Vol. 115, Nos. 3/4, May 2004 (© 2004)

Poland–Scheraga Models and the DNA Denaturation Transition

C. Richard[1] **and A. J. Guttmann**[2]

Received April 17, 2003; accepted November 20, 2003

Poland–Scheraga models were introduced to describe the DNA denaturation transition. We give a rigorous and refined discussion of a family of these models. We derive possible scaling functions in the neighborhood of the phase transition point and review common examples. We introduce a self-avoiding Poland–Scheraga model displaying a first order phase transition in two and three dimensions. We also discuss exactly solvable directed examples. This complements recent suggestions as to how the Poland–Scheraga class might be extended in order to display a first order transition, which is observed experimentally.

KEY WORDS: DNA denaturation; exactly solvable models; self-avoiding walks; phase transitions.

1. INTRODUCTION

When a solution of DNA is heated, the double stranded molecules denature into single strands. In this process, looping out of AT rich regions of the DNA segments first occurs, followed eventually by separation of the two strands as the paired segments denature. This *denaturation* process corresponds to a phase transition.[35]

A simple model of the DNA denaturation transition was introduced in 1966 by Poland and Scheraga[30, 31] (hereinafter referred to as PS) and refined by Fisher.[10, 11] The model consists of an alternating sequence (chain) of straight paths and loops, which idealize denaturing DNA, consisting of a sequence of double stranded and single stranded molecules. An

[1] Fakultät für Mathematik, Universität Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany; e-mail: richard@mathematik.uni-bielefeld.de

[2] Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia.

0022-4715/04/0500-0925/0 © 2004 Plenum Publishing Corporation

attractive energy is associated with paths. Interactions between different parts of a chain and, more generally, all details regarding real DNA such as chemical composition, stiffness or torsion, are ignored. It was found that the phase transition is determined by the critical exponent c of the underlying loop class. Due to the tractability of the problem of random loops, that version of the problem was initially studied by PS.[31] The model displays a continuous phase transition in both two and three dimensions. It was argued[10] that replacing random loops by self-avoiding loops, suggested as a more realistic representation accounting for excluded volume effects within each loop, sharpens the transition, but does not change its order.

However, the sharp jumps observed in the UV absorption rate in DNA melting[3] experiments,[35] which correspond to a sudden breaking of large numbers of base pairs, indicated that a first order phase transition would be the appropriate description. The question whether such an asymptotic description, which implies very long chains, is valid for relatively short DNA sequences, has been discussed recently.[16, 22] Nevertheless, a directed extension of the PS model, being essentially a one-dimensional Ising model with statistical weighting factors for internal loops, is widely used today and yields good coincidence of simulated melting curves with experimental curves for known DNA sequences.[4, 5, 35] Another recent application of PS models analyses the role of mismatches in DNA denaturation.[13] A numerical approach to DNA denaturation, which we will not discuss further, uses variants of the Peyrard–Bishop model,[29, 34] a Hamiltonian model of two harmonic chains coupled by a Morse potential.

With the advent of efficient computers, it has more recently been possible to simulate analytically intractable models extending the PS class, which are assumed to be more realistic representations of the biological problem. One of these is a model of two self-avoiding and mutually avoiding walks, with an attractive interaction between different walks at corresponding positions in each walk.[1, 3, 8, 9] The model exhibits a first order phase transition in $d = 2$ and $d = 3$. The critical properties of the model are described by an exponent c' related to the loop length distribution,[1, 3, 8, 20, 21, 23] see also Fisher's review article.[11] This exponent is called c again. Indeed, for PS models, it coincides with the loop class exponent c if $1 < c < 2$, see below. Within a refined model, where different binding energies for base pairs and stiffness are taken into account, the exponent c'

[3] In recent years, a number of other properties of DNA molecules have been studied by refined techniques such as optical tweezers and atomic force microscopy, and theoretical descriptions of underlying effects such as unzipping[3, 21, 23, 26, 27] have been proposed. These will not be discussed here.

seems to be largely independent of the specific DNA sequence and of the stiffness of paired walk segments corresponding to double stranded DNA parts.[8] There are, however, no simulations of melting curves for known DNA sequences which are compared to experimental curves for this model.

An approximate analytic derivation of the exponent related to this new model was given by Kafri *et al.*[20, 21, 23] using the theory of polymer networks. They estimated the excluded volume effect arising from the interaction between a single loop and two attached walks. This approach (refined recently[3]) yields an approximation to the loop length distribution exponent c', which agrees well with simulation results of interacting self-avoiding walk pairs.[1, 3, 8] There is a recent debate about the relevance of this approximation to real DNA.[16, 22] The polymer network approach, as initiated previously,[20] led to a number of related applications.[2, 3, 17, 21, 23]

In this article, we reconsider PS models for three principal reasons. Firstly, the older articles are short in motivating the use of particular loop classes, which may be misleading in drawing conclusions about the thermodynamic effects of different loop classes. In fact, the loop classes discussed in the early approaches[10, 31] are classes of *rooted* loops and lead to chains which are not self-avoiding. This seems unsatisfactory from a biological point of view, since real DNA is self-avoiding. Secondly, the common view holds that PS models with self-avoiding loops cannot display a first order transition in two or three dimensions. In fact, this view led to extending the PS class[9, 20, 21, 23] in order to find a model with a first order transition. However this view is incorrect, as we demonstrate by a self-avoiding PS model with self-avoiding loops. Thirdly, the two exponents c and c', extracted from different expressions as described above, are used in the literature without distinction, although there are subtle differences, which we will point out.

This paper is organized as follows. In the next section, we give a rigorous discussion of PS models and their phase transitions. We derive the scaling functions which describe the behavior below the critical temperature, as the critical point is approached. We then derive the loop statistics, thereby analyzing the occurrence of the loop class exponent in the loop length distribution. The third section reviews the prevailing PS models, with emphasis on motivation for the underlying loop classes. We then introduce a self-avoiding PS model displaying a first order phase transition in two and three dimensions. In Section 5, exactly solvable directed PS models are discussed. We will give explicit expressions for generating functions. The critical behavior of some of these models has been analyzed previously by different methods.[24–27] This is followed by a discussion of models extending the PS class. We conclude with a discussion of some open questions.

2. POLAND–SCHERAGA MODELS: GENERAL FORMALISM

We define PS models and consider analytical properties of their free energies, thereby refining previous expositions[10, 11, 30, 31] in a rigorous manner. We analyze phase transitions, extract the asymptotic behavior of chains at temperatures near the phase transition by means of scaling functions, and consider the loop length distribution of chains.

2.1. Definition of the Model

We consider a discrete model, defined on the hypercubic lattice $\mathbb{Z}^d$. Double stranded DNA segments are modeled by paths, and single stranded DNA segments are modeled by loops on the edges of the lattice. Each loop is assumed to have two marked vertices to indicate where paths are attached. (Like PS, we do not *a priori* assume the "DNA condition" that the marked vertices divide the loop into parts of equal lengths. There is a recent discussion about the effect of mismatches.[13]) Any alternating sequence of paths and (marked) loops is called a *chain*. A *PS model* consists of all chains obtained by concatenation of paths and loops from a given path class and a given loop class, where the initial segment and the final segment of a chain are both paths. (The case of an open end affects the behavior above the critical temperature[22] and will not be considered here.[3]) Note that, in general, such chains are not self-avoiding, in contrast to real DNA. Self-avoidance may be violated by paths, by loops, or by the way segments are concatenated. A chain is called *segment-avoiding* if there are no overlaps, i.e., every two non-neighboring segments have no vertex in common, and every two neighboring segments have exactly the marked vertex in common. We call a PS model *self-avoiding* if paths and loops are self-avoiding and if all chains of the model are segment-avoiding. We will discuss below several examples of self-avoiding PS models, which arise from self-avoiding walks.

The requirement of self-avoidance restricts the admissible path classes and loop classes. A simple subclass of self-avoiding PS models are directed PS models: We call a chain *directed* if there is a preferred direction such that the order of the chain segments induces the same order on the vertex coordinates (w.r.t. the preferred direction), for each pair of vertices taken from two different chain segments. Such chains are then segment-avoiding. We call a PS model *directed* if paths, loops and chains are directed. We will discuss below several examples of directed PS models which arise from classes of directed walks.

For a given PS model, let $z_{m,n}$ denote the number of chain configurations with m contacts and length n. The generating function is defined by

$$Z(x, w) = \sum_{m,n} z_{m,n} w^m x^n, \tag{2.1}$$

where the activity x is conjugate to the chain length n. The Boltzmann factor $w = e^{-E/kT}$ takes into account the attractive interaction (achieved by setting the energy $E < 0$) between bonds. T is the temperature, and k is Boltzmann's constant. Over the relevant temperature range $0 < T < \infty$, we have $\infty > w > 1$. Note that there is no interaction between different segments in a chain.

The generating function $Z(x, w)$ can be expressed in terms of the generating functions for paths $V(x)$ and loops $U(x)$. These are

$$V(x) = \sum_{n=0}^{\infty} b_n x^n, \qquad U(x) = \sum_{n=1}^{\infty} p_{2n} x^n, \tag{2.2}$$

where b_n is the number of paths of length n, and p_{2n} is the number of loops[4] of length $2n$. Due to the chain structure, we get a geometric series in $V(wx)\,U(x)$,

$$Z(x, w) = \frac{V(wx)}{1 - U(x)\,V(wx)} = \sum_{n=1}^{\infty} Z_n(w)\,x^n. \tag{2.3}$$

Since we want to analyze phase transitions of the model, which can only occur in the infinite system, we define the free energy of the model as

$$f(w) = \lim_{n \to \infty} \frac{1}{n} \log Z_n(w) = -\log x_c(w), \tag{2.4}$$

where, for fixed w, $x_c(w)$ is the radius of convergence of $Z(x, w)$. Concatenation arguments and supermultiplicative inequalities can be used to show that the free energy exists.[18] We will alternatively investigate properties of the free energy in terms of the generating functions for paths and loops.[5] Throughout the paper, we employ the following assumption.

Assumption 1. Assume that the generating functions $U(x)$ and $V(x)$, defined in (2.2), have radius of convergence x_U and x_V, respectively, where $0 < x_U < x_V \leqslant 1$. At the critical point $x = x_V$, assume that $V(x_V^-) := \lim_{x \to x_V^-} V(x) = \infty$.

Remark. $U(x)$ is a generating function (i.e., a series with nonnegative coefficients), such that $U(x)$ and the derivative $U'(x)$ are strictly

[4] In simulations of melting curves like MELTSIM,[4,5] a cooperative parameter $\sigma \approx 10^{-5}$ is used. In our context, this amounts to replacing the loop generating function $U(x)$ by $\sigma U(x)$.

[5] The reader may find it illuminating in following this very general discussion that now follows to refer to a specific model, discussed in Section 5.1.

positive for $0 < x < x_U$. A corresponding statement holds for $V(x)$. The assumption that $0 < x_U < 1$ reflects the requirement that the number of configurations p_{2n} grows exponentially in length. The models commonly discussed[10, 11, 21, 30, 31] have $x_V = 1$ and $V(0) = 1$. For loops and paths of the same type, we have $x_U = x_V^2 < x_V < 1$, see also the examples discussed below. The assumption $V(x_V^-) = \infty$ is satisfied for typical classes of paths. Note that typically $U(x_U^-) := \lim_{x \to x_U^-} U(x) < \infty$, but there are loop classes where $U(x_U^-) = \infty$ such as convex polygons.[6] We have $U(0) = 0$ by definition.

The radius of convergence $x_c(w)$ of $Z(x, w)$ is the minimum of the radius of convergence of $U(x)\,V(wx)$ and the point $x_1(w)$ where the denominator in (2.3) vanishes. Define $F(x, w) := U(x)\,V(wx)$. As a function of argument x, $F(x, w)$ is continuous and monotonically increasing for $0 < x < \min(x_U, x_V/w)$. Note that $F(0, w) = 0$. If $w \geqslant x_V/x_U$, then $\lim_{x \to x_V/w} F(x, w) = \infty$, such that there exists a unique solution $x_1(w) < x_U$ with $F(x_1(w), w) = 1$. Assume that $1 \leqslant w < x_V/x_U$. If $U(x_U^-) = \infty$, there exists a unique solution $x_1(w) < x_U$. If $U(x_U^-) < \infty$ and $U(x_U^-)\,V(x_U) \geqslant 1$, we have $F(x_U, w) \geqslant 1$ for all $w \geqslant 1$, i.e., there exists a unique solution $x_1(w) < x_U$. If $U(x_U^-) < \infty$ and $U(x_U^-)\,V(x_U) < 1$, we define $w_c > 1$ by the condition $U(x_U^-)\,V(w_c\,x_U) = 1$. Then, for $w \geqslant w_c$, we have a unique solution $x_1(w) < x_U$. If $w < w_c$, there is no such solution, and the dominant singularity of $Z(x, w)$ occurs at $x = x_U$. We thus proved the following theorem.

Theorem 1. If Assumption 1 is satisfied, the PS model (2.3) has free energy

$$f(w) = \begin{cases} -\log x_c(w) & (w_c \leqslant w < \infty) \\ -\log x_U & (1 < w < w_c), \end{cases} \tag{2.5}$$

where the radius of convergence $x_c(w)$ of $Z(x, w)$ is, for $w_c \leqslant w < \infty$, the unique positive solution of

$$U(x_c(w))\,V(w\,x_c(w)) = 1. \tag{2.6}$$

If $U(x_U^-) := \lim_{x \to x_U^-} U(x) = \infty$ or $U(x_U^-)\,V(x_U) \geqslant 1$, we have $w_c = 1$. Otherwise, $w_c > 1$ is implicitly given by $U(x_U^-)\,V(w_c\,x_U) = 1$.

Remark. If $w_c > 1$, this point is a critical point, see the following section.

2.2. Phase Transitions

For $w > w_c$, there is no phase transition, since the radius of convergence $x_c(w) > 0$ is analytic for $w_c < w < \infty$, as can be inferred from (2.6). A necessary condition for a phase transition at a finite temperature $w_c > 1$ is $U(x_U^-)\,V(x_U) < 1$. The nature of the transition at $w = w_c > 1$ is exhibited by the fraction of shared bonds $\theta(w) = w\frac{df(w)}{dw}$, which is defined for $w \neq w_c$. Note that $\theta(w) = 0$ for $w < w_c$. For $w > w_c$, we have

$$\theta(w) = -w\frac{x_c'(w)}{x_c(w)} = w\left(\frac{U'}{U}\frac{V}{V'} + w\right)^{-1} > 0. \tag{2.7}$$

Consider the limit $w \to w_c^+$. We have $0 < V < \infty$, since $UV = 1$ and U is finite. Since $V(x)$ is a generating function, this implies $0 < V' < \infty$. If $U'(x_U^-) = \infty$, it follows that $\theta(w) \to 0$, such that the phase transition is continuous. If $U'(x_U^-) < \infty$, $\theta(w) \to \theta_c > 0$, such that the phase transition is of first order. This leads to the following statement.

Theorem 2. Let Assumption 1 be satisfied. If $U(x_U^-) = \infty$ or $U(x_U^-)\,V(x_U) \geqslant 1$, the PS model (2.3) has no phase transition at finite temperature. Otherwise, if $U(x_U^-) < \infty$ and $U'(x_U^-) < \infty$, a continuous phase transition will occur at $w = w_c > 1$ defined in Theorem 1. If $U(x_U^-) < \infty$ and $U'(x_U^-) = \infty$, a first order phase transition will occur at $w = w_c > 1$ defined in Theorem 1.

Remark. The phase transition condition $U(x_U^-)\,V(x_U) < 1$ is typically satisfied in more realistic models, where $U(x)$ is multiplied by the cooperativity parameter $\sigma \approx 10^{-5}$.

We conclude that the nature of the transition is determined by the singularity of the loop generating function $U(x)$ at $x = x_U$. It can be more directly related to the asymptotic properties of loops, which are typically of the form[6]

$$p_{2n} \sim A x_U^{-n} n^{-c} \qquad (n \to \infty), \tag{2.8}$$

for some constants $A > 0$ and $c \in \mathbb{R}$. The exponent c determines the singularity at $x = x_U$, which is, to leading order and for $c \in \mathbb{Q} \setminus \mathbb{N}$, algebraic. This specializes Theorem 2.

[6] $A_n \sim B_n$ for $n \to \infty$ means that $\lim_{n\to\infty} A_n/B_n = 1$. Similarly, $f(x) \sim g(x)$ for $x \to x_c$ means that $\lim_{x\to x_c} f(x)/g(x) = 1$.

Proposition 1. Let Assumption 1 be satisfied and assume that, for $c \in \mathbb{R} \setminus \mathbb{N}$, the singular part of $U(x)$ is, as $x \to x_U^-$, asymptotically given by

$$U^{(\text{sing})}(x) \sim U_0 (x_U - x)^{c-1} \qquad (x \to x_U^-) \tag{2.9}$$

for some constant $U_0 \neq 0$. Then, the PS model (2.3) has no phase transition if $0 < c < 1$ or if $U(x_U^-) V(x_U) \geqslant 1$. Otherwise, a continuous phase transition occurs for $1 < c < 2$, and a first order transition occurs if $c > 2$, at $w = w_c > 1$ defined in Theorem 1.

Remark. If $c = 1$ in (2.8), we get a logarithmic singularity in $U(x)$, hence no phase transition. If $c = 2$ in (2.8), we get a logarithmic singularity in the derivative $U'(x)$, resulting in a continuous phase transition. See also Fisher.[11]

2.3. Scaling Functions

In the vicinity of a phase transition point, critical behavior of the form

$$Z(x, w) \sim (x_c - x)^{-\theta} F((w - w_c)/(x_c - x)^{\phi}) \qquad (x, w) \to (x_c^-, w_c^+), \tag{2.10}$$

uniformly in w, is expected with critical exponents θ and ϕ. Here $F(s)$ is a scaling function which only depends on the combined argument $s = (w - w_c)/(x_c - x)^{\phi}$. The scaling function is extracted from the generating function $Z(x, w)$ by replacing w using the variable s of combined argument and expanding to leading order in $x_c - x$. To this end, let us assume as in (2.9) that $U(x)$ behaves as

$$U(x) = U(x_U^-) + U_0 (x_U - x)^{c-1} + o((x_U - x)^{c-1}) \qquad (x \to x_U^-), \tag{2.11}$$

and $1 < c \in \mathbb{R} \setminus \mathbb{N}$. In this case, the constant U_0 is negative. (If $c < 1$, no phase transition occurs at positive temperature according to the considerations above.) We have the following result.

Theorem 3. Let Assumption 1 be satisfied. If the leading singularity of the loop generating function $U(x)$ is of the form (2.11) with an exponent $c > 1$ and if $U(x_U^-) V(x_U) \leqslant 1$, the PS model (2.3) has critical exponents and scaling function (2.10)

$$\theta = \phi = c-1, \qquad F(s) = \frac{1}{|U_0| - U^2(x_U^-)\, V'(w_c\, x_U)\, x_U s} \qquad (1 < c < 2),$$
$$\theta = \phi = 1, \qquad F(s) = \frac{1}{U^2(x_U^-)\, V'(w_c\, x_U)[w_c - x_U s]} \qquad (c > 2), \tag{2.12}$$

where $w_c \geqslant 1$ is implicitly given by $U(x_U^-)\, V(w_c\, x_U) = 1$.

Remark. If $c > 2$, the critical exponents are independent of c. In both cases, the scaling function is proportional to $F_a(s) = 1/(1-as)$ with a positive constant a.

It can be shown[18] that under mild assumptions such a scaling function implies a certain asymptotic behavior of the function $Z_n(w)$ for large n. That is to say,

$$Z_n(w) \sim x_c^{-n-\theta} n^{\theta-1} h\left(\left(\frac{n}{x_c}\right)^{\phi} (w - w_c)\right) \qquad (n \to \infty,\ w \to w_c^+), \tag{2.13}$$

uniformly in w, where $h(x)$ is the finite-size scaling function $h(x) = \sum_{k=0}^{\infty} f_k x^k / \Gamma(k\phi + \theta)$, where $\Gamma(z)$ denotes the Gamma function, and the coefficients f_k appear in the Taylor expansion of the scaling function $F(s) = \sum_{k=0}^{\infty} f_k s^k$. This can be used to derive the critical behavior of the fraction of shared bonds above the phase transition[10] for $1 < c < 2$,

$$\theta(w) = A(w - w_c)^{\frac{2-c}{c-1}} \qquad (w \to w_c^+), \tag{2.14}$$

for some constant $A > 0$. In our case $F_a(s) = 1/(1-as)$, we have $h_a(x) = e^{ax}$, if $\theta = 1$. For $0 < \theta = c - 1 < 1$ and θ rational, the finite size scaling function can be expressed in terms of hypergeometric functions. For the case $c = 3/2 = \theta + 1$, which will be relevant below, we get

$$h_a(x) = \frac{1}{\sqrt{\pi}} + axe^{(ax)^2}(1 + \operatorname{erf}(ax)), \tag{2.15}$$

where $\operatorname{erf}(x) = 2/\sqrt{\pi} \int_0^x e^{-t^2}\, dt$ denotes the error function.

2.4. Loop Statistics

Within a given PS model, consider the set of all chains of length n with m contacts. Denote the number of loops of length $2l$ in this set by $g_{m,n,l}$. Define the generating function

$$L(x, w, y) = \sum_{m,n,l} g_{m,n,l} w^m x^n y^l. \tag{2.16}$$

Again, due to the directedness of the model, this generating function can be expressed in terms of the loop generating function $U = U(x)$, $\bar{U} = U(xy)$ and the path generating function $V = V(xw)$. An argument as in Section 2.1 yields

$$\begin{aligned} L(x, w, y) &= V\bar{U}V + (V\bar{U}VUV + VUV\bar{U}V) + \cdots \\ &= V\bar{U}V(1 + 2UV + 3(UV)^2 + \cdots) \\ &= V\bar{U}V \frac{1}{(1-UV)^2} = U(yx)\, Z^2(x, w). \end{aligned} \tag{2.17}$$

The generating function for the total number of loops is then given by setting $y = 1$,

$$T(x, w) = L(x, w, 1) = U(x)\, Z^2(x, w), \tag{2.18}$$

and the generating function for the sum of all loop lengths is given by

$$S(x, w) = y \frac{d}{dy} L(x, w, y)\bigg|_{y=1} = xU'(x)\, Z^2(x, w). \tag{2.19}$$

The finite size behavior of these quantities about the phase transition can be computed from the scaling function results of Section 2.3. For the total number of loops, we get

$$\lim_{w \to w_c^+} [x^n]\, T(x, w) \sim A x_U^{-n-2\theta} n^{2\theta-1} \qquad (n \to \infty), \tag{2.20}$$

where $A > 0$ is some amplitude,[7] and for the number of loops of length $2l$, we get

$$\begin{aligned} \lim_{w \to w_c^+} [x^n y^l]\, L(x, w, y) &\sim p_{2l} x_U^{l-n-2\theta} n^{2\theta-1} \qquad (n \to \infty) \\ &\approx A l^{-c} x_U^{-n-2\theta} n^{2\theta-1} \qquad (1 \ll l \ll n, n \to \infty). \end{aligned} \tag{2.21}$$

Note that we obtained (2.21) under the assumption that l is asymptotically large but that $l \ll n$. This is due to the estimate $[y^l]\, U(yx) = p_{2l} x^l \sim p_{2l} x_U^l$ for $x \to x_U^-$ and l finite, but $p_{2l} \sim A x_U^{-l} l^{-c}$ for $l \to \infty$, where we assumed (2.8) to be valid.

[7] In order to simplify notation, we will denote all following amplitudes by the letter A, with the convention that their values may be different.

For the sum of the loop lengths, we distinguish between the cases $1 < c < 2$ and $c > 2$, where the derivative of the loop generating function $U'(x_U^-)$ is finite or infinite at the critical point. We find

$$\lim_{w \to w_c^+} [x^n]\, S(x, w) \sim \begin{cases} A x_U^{-n-c} n^{c-1} & (1 < c < 2) \\ A x_U^{-n-2} n & (c > 2) \end{cases} \qquad (n \to \infty). \tag{2.22}$$

Let us compute the probability of a loop of length $2l$ within chains of length n. We get, in the limit $w \to w_c^+$, for the loop length distribution

$$P(l, n) = \lim_{w \to w_c^+} \frac{[x^n y^l]\, L(x, w, y)}{[x^n]\, T(x, w)} \approx A l^{-c} \qquad (1 \ll l \ll n, n \to \infty). \tag{2.23}$$

Furthermore, for the mean loop length $\langle l \rangle_n$ we get, in the limit $w \to w_c^+$,

$$\langle l \rangle_n = \sum_{l=0}^{n} l\, P(l, n) = \lim_{w \to w_c^+} \frac{[x^n]\, S(x, w)}{[x^n]\, T(x, w)} \sim \begin{cases} A n^{2-c} & (1 < c < 2) \\ A n^0 & (c > 2) \end{cases} \qquad (n \to \infty). \tag{2.24}$$

The behavior of the mean loop length reflects the nature of the phase transition: If $1 < c < 2$, the mean loop length diverges, such that the transition is continuous. If $c > 2$, the mean loop length stays finite, indicating a first order transition.

3. TWO PROMINENT LOOP CLASSES

Since the critical behavior of PS models is essentially determined by the properties of loops, PS,[30, 31] and later Fisher,[10] were led to consider various loop classes (together with straight paths for the double stranded segments). Whereas PS analyzed loop classes derived from random walks, Fisher considered loop classes derived from self-avoiding walks.

3.1. Loops and Walks

An oriented, rooted loop of length n is a walk of length $n-1$, whose starting point and end point are lattice nearest neighbors. As usual (ref. 28, Section 3.2), we identify such loops if they have the same shape, i.e., if they are equal up to a translation, possibly followed by a change of orientation. These objects we call unrooted, unoriented loops, or simply loops. Each loop of length n has at most $2n$ corresponding walks. If the walks are self-avoiding, each loop has exactly $2n$ corresponding walks. The number of loops of length n is denoted by p_n. For example, for self-avoiding loops

on $\mathbb{Z}^2$ we have $p_4 = 1$ and $p_6 = 2$. For a given class of walks, the above description defines the corresponding (unmarked) loop class. Within a chain structure, two paths are attached to each loop. Different choices for attachment positions increase the number of (marked) loop configurations to $\tilde{p}_n \geqslant p_n$. If we assume the DNA condition that the two paths attached to a loop bisect it into pieces of equal length, then the number of possible attachments of two paths to a loop of length n is less than or equal to $2n$. (We distinguish the two strands of a marked loop.) PS and Fisher consider classes of oriented rooted loops. Self-avoiding oriented rooted loops can be interpreted as loops with $2n$ possible attachment positions of paths to a loop of length n. A similar interpretation for oriented rooted random loops is not obvious. We stress that these loop classes result in chains which are not segment-avoiding, as paths will intersect the loops. Both models cannot therefore represent real (self-avoiding) DNA.

3.2. Oriented Rooted Random Loops

The first simple example of loops, which are discussed by PS,[31] are oriented rooted random loops derived from random walks. Random walks on $\mathbb{Z}^d$ have the generating function $V_d(x) = 1/(1-2dx)$. The asymptotic behavior of the number of oriented rooted random loops of length $2n$ is given [ref. 28, Appendix A] by

$$\tilde{p}_{2n} \sim A(2d)^{2n}\,(2n)^{-d/2} \qquad (n \to \infty). \tag{3.1}$$

This implies no phase transition in $d = 2$ and a continuous phase transition in $d = 3$, since the phase transition condition $U(x_U^-)\,V(x_U) < 1$ is satisfied in $d = 3$: We have

$$\tilde{p}_{2n} = \sum_{k+l+m=n} \frac{(2n)!}{(k!)^2\,(l!)^2\,(m!)^2} = \frac{((2n)!)^2}{(n!)^4}\,{}_3F_2(-n, -n, -n; 1, -n+1/2; 1/4), \tag{3.2}$$

where ${}_3F_2(a_1, a_2, a_3; b_1, b_2; z)$ is a hypergeometric function. With $x_U = 1/36$, we extracted the amplitude $U(x_U^-)$ numerically[14] using the values $\tilde{p}_{2n}$ where $n \leqslant 30$, getting $U(x_U^-) = 0.51638461326(7)$. The result follows for straight paths since $1/(1-x_U) = 36/35$ and also for random walks since $V_3(x_U) = 6/5$.

3.3. Oriented Rooted Self-Avoiding Loops

The PS results led to the question[10] whether accounting for excluded volume effects within a loop increases the loop class exponent c, which

might change the order of the phase transition. This led to considering *self-avoiding loops*, which are loops derived from self-avoiding walks. (A self-avoiding walk [ref. 28, Section 1.1] is a random walk which never visits a vertex twice.) By definition, the self-avoiding loop class fully accounts for excluded volume interactions within a loop. In $d = 3$, self-avoiding loops of length $n \geqslant 24$ may be knotted. Self-avoiding walks and loops are well studied objects.[18, 28] Fisher considered oriented rooted loops $\tilde{p}_{2n} = 4np_{2n}$. Their loop class exponent c, $\tilde{p}_{2n} \sim B\mu_d^n n^{-c}$, is related to the mean square displacement exponent ν of self-avoiding walks by the hyperscaling relation [ref. 28, Section 2.1] $c = d\nu$, where $\nu = 1/2$ for $d \geqslant 4$, $\nu = 0.5877(6)$ for $d = 3$ and $\nu = 3/4$ for $d = 2$. At present, there is no proof for the values of ν in dimensions $d \leqslant 4$. Explicitly, we have

$$c = \begin{cases} d/2 & d \geqslant 4 \\ 1.7631(18) & d = 3 \\ 3/2 & d = 2. \end{cases} \tag{3.3}$$

For unknotted self-avoiding loops, which is the preferrable model from a biological point of view, it has been proved[33] that the exponential growth constant is strictly less than that of all self-avoiding loops, while the exponent (if it exists) is expected to coincide with that of all self-avoiding loops.

Fisher concluded that the above values of the loop class exponent c imply a continuous phase transition in $d = 2$ and $d = 3$. We have to check the phase transition condition: In $d = 2$, a numerical analysis of the oriented rooted SAP series data[19] gives $x_U = 0.1436806285(8)$ and $U(x_U^-) = 0.6523866(2)$. The value of the generating function for straight paths is $1/(1 - x_U) < 1.678$. In $d = 3$, a corresponding analysis with unknotted oriented rooted SAP data and SAW data [ref. 28, Appendix C] gives $x_U = 0.045578(3)$ and $U(x_U^-) = 0.10(1)$. The path generating function value is bounded by that of random walks $1/(1 - 6x_U) < 1.38$. Thus, the phase transition condition is satisfied in both $d = 2$ and in $d = 3$ for straight paths. Note that in $d = 2$, self-avoiding walks as paths will result in no phase transition, since $V(x_U) > 2$ for such paths, as follows from SAW data analysis.

4. SELF-AVOIDING POLAND–SCHERAGA MODELS

As discussed above, the previous PS models are not self-avoiding. Given a particular class of walks, for example, SAW, a PS model with segment-avoiding chains may be defined as follows. As paths, take only those walks with extremal first and last vertex: If $v(0), v(1), \ldots, v(n)$ are the

vertices of an n-step walk v, this walk is taken as a path iff $v_x(0) < v_x(i) < v_x(n)$ for all $1 < i < n$. Such walks are bridges [ref. 28, Section 1.2], whose last step is in x-direction. For (unmarked) loops, take loops derived from the walks, as explained in Section 3.1. In our example these will be SAW loops. Marking of the loops may be achieved in different ways. To this end, consider for a given loop the sets A_l (A_r) of vertices of smallest (largest) x-coordinate. We distinguish four different types of marking: complete marking (with DNA constraint), where we mark a loop at all vertex pairs from A_l and A_r (whenever the DNA condition is satisfied), and unique marking (with DNA constraint), where we only mark a loop at a single vertex pair, for example the bottom vertex and the top vertex in a lexicographic ordering (if they satisfy the DNA condition).

If the walk class is self-avoiding walks, this will result in self-avoiding PS models. Unique marking would then imply $\tilde{p}_n = 2p_n$ (we distinguish the two strands of a marked loop) and hence increase the previous exponents by one. Hence, such a PS model displays a first order transition with $c = 5/2$ in $d = 2$ and with $c = 2.7631(18)$ in $d = 3$. (The phase transition condition is satisfied, as follows from the values $U(x_U^-)$ given in Section 3.3, and from the estimate $V(x_U) < 1/(1-dx_U) - 1$ for bridges with last step in x-direction.) Unique marking with DNA constraint results in $\tilde{p}_n \leqslant 2p_n$. If we assume that the exponential growth constant for marked self-avoiding loops is still given by μ_d (which we expect to be true, compare Section 5.2), this implies a critical exponent c greater or equal to the model with unique marking, i.e., a first order phase transition in $d = 2$ and $d = 3$. For complete marking, we have $2p_n \leqslant \tilde{p}_n \leqslant 2np_n$, which rules out a decrease of c by more than one. We however expect that the number of possible markings is of order 1 as $n \to \infty$, such that c remains unchanged, and the model again displays a first order phase transition. Similar considerations apply for the case of complete marking with DNA constraint.

5. DIRECTED POLAND–SCHERAGA MODELS

We present two classes of directed PS models which are exactly solvable and, by definition, take into account excluded volume interactions between all parts of the structure. The first class is derived from fully directed walks and is solvable in arbitrary dimension.[15] We will give explicit expressions for the generating function in $d = 2$ and in $d = 3$. This extends,[24–27] where the critical behavior has been derived. The model displays a first order phase transition only for $d > 5$. The second class, derived from partially directed walks, is considered in $d = 2$ only. Different variants of the second model display both a first order phase transition and a continuous phase transition in $d = 2$.

5.1. Fully Directed Walks and Loops

This directed PS model consists of fully directed walks for the paths in the chain. These only take steps in positive directions. The corresponding loops are staircase polygons, which consist of two fully directed walks, which do not intersect or touch, but have a common starting point and end point. Paths are attached to these points. We distinguish the two strands of a loop. This model satisfies all the assumptions discussed in Section 2.1, and also satisfies the DNA condition that the two segments of the loops are the same length.

In $d = 2$, the generating functions for paths and marked loops are[18]

$$V(x) = \frac{1}{1-2x}, \qquad U(x) = 1-2x-\sqrt{1-4x}. \tag{5.1}$$

$U(x)$ is twice the generating function of staircase polygons. The loop class exponent is $c = 3/2$. We have $x_V = 1/2$, $x_U = 1/4$, and $w_c = 1$, i.e., a phase transition at $T = \infty$. (If the empty path would not be allowed, a phase transition occurred at a finite temperature.) The free energy $f(w)$ is given by

$$f(w) = \log\left(\frac{2(w-1)^2}{\sqrt{1+(w-1)^2}-1}\right) \qquad (1 \leqslant w < \infty). \tag{5.2}$$

The fraction of shared bonds follows as

$$\theta(w) = \frac{2w}{w-1}\,\frac{\sqrt{1+(w-1)^2}-1-(w-1)^2/2}{1+(w-1)^2-\sqrt{1+(w-1)^2}} \qquad (1 \leqslant w < \infty), \tag{5.3}$$

which approaches zero linearly in $w-1$. The asymptotic behavior of $Z_n(w)$ about $w = 1$ is given by

$$Z_n(w) \sim \frac{4^{n-1}}{n^{1/2}}\, h_{\sqrt{n}/8}(w-1) \qquad (n \to \infty,\ w \to 1^+), \tag{5.4}$$

uniformly in w, where $h_a(x)$ is given by (2.15).

The PS model of fully directed walks and loops is exactly solvable in arbitrary dimension.[15] Consider fully directed walks on $\mathbb{Z}^d$. If there are k_i steps in direction i, the number of distinct walks, starting from the origin, is given by the multinomial coefficient $\binom{k_1+k_2+\cdots+k_d}{k_1, k_2, \ldots, k_d}$. This results in $V_d(x) = 1/(1-dx)$. The generating function of a pair of such walks is given by

$$\tilde{Z}_d(x) = \sum_{k_1, k_2, \ldots, k_d = 0}^{\infty} \binom{k_1+k_2+\cdots+k_d}{k_1, k_2, \ldots, k_d}^2 x^{k_1+\cdots+k_d}. \tag{5.5}$$

The generating function $\tilde{Z}_d(x)$ can be interpreted as a chain, where each link consists of a staircase polygon or a double bond. Thus, it is related to the generating function $U_d(x)$, where staircase polygons are counted twice, by

$$\tilde{Z}_d(x) = \frac{1}{1-(dx+U_d(x))}. \tag{5.6}$$

The functions $\tilde{Z}_d(x)$ satisfy Fuchsian differential equations of order $d-1$, from which the singular behavior of $\tilde{Z}_d(x)$, and hence of $U_d(x)$ can be derived.[15] In dimensions $2 < d < 5$, $U_d(x)$ can be expressed in terms of Heun functions. We get in $d = 3$

$$\tilde{Z}_3^2(x) = \left(\frac{2}{\pi}\right)^2 (1-9x)^{-1} (1-x)^{-1} K(k_+) K(k_-), \tag{5.7}$$

where $K(k)$ is the complete elliptic integral of the first kind, and

$$\begin{aligned} k_\pm^2 &= \frac{1}{2} \pm \frac{x_1}{4} (4-x_1)^{\frac{1}{2}} - \frac{1}{4} (2-x_1)(1-x_1)^{\frac{1}{2}} \\ x_1 &= -\frac{16x}{9x^2-10x+1}. \end{aligned} \tag{5.8}$$

We have $x_U = 1/9$. In arbitrary dimension $d \geqslant 3$, it has been shown[15] that the models have a critical point $x_U = 1/d^2$ with exponent $c = (d-1)/2$, with logarithmic corrections in $d = 3$. Noting that $U_d(x_U^-) = 1 - 1/d - 1/\tilde{Z}_d(x_U^-)$ and $V_d(x_U) = d/(d-1)$, we conclude that the phase transition condition $U_d(x_U^-)\, V_d(x_U) < 1$ is satisfied iff $\tilde{Z}_d(x_U^-) < \infty$. This is the case[15] in $d \geqslant 4$. For the corresponding PS model, the phase transition is thus first order for $d \geqslant 6$, and we have a continuous phase transition in dimensions $2 \leqslant d \leqslant 5$, being at finite temperature in $d \geqslant 4$ only.

5.2. Partially Directed Walks and Loops

We consider a directed PS model, where paths are partially directed walks on the square lattice $\mathbb{Z}^2$. These walks are self-avoiding and not allowed to take steps in the negative x direction. The number of paths of length n can be obtained[18] by considering the ways a walk of length n can be obtained from a walk of length $n-1$. This yields the path generating function

$$V(x) = \frac{1+x}{1-2x-x^2}, \tag{5.9}$$

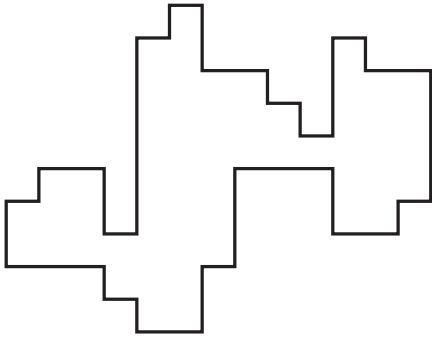

Fig. 1. A column-convex polygon.

whose dominant singularity is a simple pole at $x_V = 1-\sqrt{2}$. The associated loop class is column-convex polygons, which has been analyzed in ref. 7, see Fig. 1 for an example. We assume unique marking as explained in Section 4. The (unmarked) loop generating function $U(x)$ satisfies the algebraic equation of order four [ref. 7, Eq. (5.2)]

$$\begin{aligned} 0 = &-(x-1)^4 x^2 + (x^3-7x^2+3x-1)(x-1)^3 U(x) \\ &+ 2(2x^3-11x^2+10x-4)(x-1)^2 U(x)^2 \\ &+ (5x^3-35x^2+47x-21)(x-1) U(x)^3 \\ &+ (2x^3-23x^2+38x-18) U(x)^4. \end{aligned} \tag{5.10}$$

The dominant singularity at $x_U = x_V^2 = 3-2\sqrt{2}$ is a square-root singularity with an exponent $c = 3/2$. The critical amplitude is given by $U(x_U) = 0.086983(1)$. Thus $2U(x_U) V(x_U) < 1$. This implies a continuous phase transition for the PS model at $w_c = 1.922817(1)$.

However, the loop class defined above does not obey the DNA constraint, which might increase the exponent c. Using standard techniques,[7] it is possible to derive a functional equation for the generating function of column-convex polygons counted by perimeter and, in addition, by their upper and lower walk length. The corresponding expression is, however, much more difficult to handle than (5.10), and there seems to be no practical way to extract the loop class exponent for the model with DNA constraint. On the other hand, it is possible to bound the loop class exponent from above by considering a refined (artificial) model. As we will argue below, this refined model has a loop class exponent $c = 5/2$. For the former model, we thus expect a value of $3/2 \leqslant c \leqslant 5/2$, with a value of

$c = 2$ seeming plausible. This PS model would then be at the border between a first order and a continuous phase transition.

For the refined model, let us further assume that the upper and lower walk in a loop are not allowed to touch the x-axis. The loop class of the refined model can be described in terms of bar-graph polygons.[32] These are column-convex polygons, one of whose walks is a straight line. Loops are composed of two bar-graph polygons, one of them in the half-plane $y \geqslant 0$, the other in the half-plane $y \leqslant 0$. Both polygons have the same horizontal coordinates and are constrained to have equal walk lengths. Paths of the model are partially directed walks, which are attached to the two loop vertices with vertical coordinate zero.

Let $g_{m,n}$ denote the number of bar-graph polygons with $2m$ horizontal and $2n$ vertical steps. The anisotropic perimeter generating function $G(x, y) = \sum g_{m,n} x^m y^n$ satisfies an algebraic equation of degree two,[32]

$$G(x, y) = xy + (y + x(1+y))\, G(x, y) + xG^2(x, y). \tag{5.11}$$

A generalization of the Lagrange inversion formula[12] can be used to obtain a closed formula for its series coefficients. It is

$$[x^m y^n]\, G(x, y) = g_{m,n} = \sum_{k=0}^{n} \frac{1}{m+k} \binom{m+k}{m} \binom{m}{n-1-k} \binom{m}{n-k}. \tag{5.12}$$

The number p_{2n} of all combinations of bar-graph polygons subject to equal walk lengths is then

$$p_{2n} = \sum_{k=1}^{(n-1)/2} g_{n-2k,k}^2. \tag{5.13}$$

The first few numbers p_{2n} are 1, 1, 2, 10, 38, 126, 483, 126, 483,... for $n = 3, 4, 5, \ldots$. We used standard methods of numerical series analysis[14] to estimate the critical point and critical exponent of $U(x)$. An analysis with first order differential approximants, using the coefficients p_{2n} for $70 \leqslant n \leqslant 80$, yields the estimates

$$x_U = 0.17157287(1), \qquad c = 2.5000(1). \tag{5.14}$$

This is, to numerical accuracy, indistinguishable from $x_U = 3 - 2\sqrt{2} = 0.171572875253\ldots$, as expected, and $c = 5/2$. Note that the phase transition condition $2U(x_U^-)\, V(x_U) < 1$ is satisfied, since the loop generating function is bounded by that of column-convex polygons, which is known to satisfy the above inequality.

6. MODELS EXTENDING THE POLAND–SCHERAGA CLASS

The PS model with oriented rooted self-avoiding loops,[10] reviewed in Section 3.3, displays a continuous phase transition in $d = 2$ and $d = 3$. This led to the question whether introducing excluded volume effects between different segments of the chain can change the nature of the transition.[1, 3, 8, 9, 20, 21, 23] In Section 4, we introduced self-avoiding PS models in order to answer that question in the affirmative, staying within the PS class.

Another approach, which extends the PS class, consisted in simulations of self-avoiding walk pairs.[1, 3, 8, 9] We emphasize that these models are not PS models as defined in this paper, and that loop class exponents as defined in (2.8) bear no meaning for these models. The appropriate generalization is the exponent c' of the loop length distribution $P(l, n)$ of the chain, which is assumed to behave in the limit $w \to w_c^+$ like[2, 3, 8, 20, 21, 23]

$$P(l, n) \approx l^{-c'} g(l/n) \qquad (1 \ll l \leqslant n, n \to \infty), \tag{6.1}$$

where $g(x)$ is a scaling function, assumed to be constant in some analyses.[2, 20, 21, 23] Then, the same conclusions as in Section 2.2 about the nature of the phase transition determined by the loop length distribution exponent c' hold. This follows from the behavior of the mean loop length $\langle l \rangle_n$ in the limit $w \to w_c^+$, which is obtained from (6.1) by integration,

$$\langle l \rangle_n = \sum_{l=0}^{n} l\, P(l, n) \approx A n^{2-c'} \qquad (n \to \infty). \tag{6.2}$$

If $1 < c' < 2$, the mean loop length diverges, indicating a continuous transition. If $c' > 2$, the mean loop length stays finite, indicating a first order transition.[2, 8] Note that assumption (6.1) is not fully consistent with PS models: The behavior (2.23) is consistent with (6.1), which justifies the use of the same name for the two exponents. We found (6.2) to be satisfied for PS models if $1 < c < 2$ in Section 2.4. If $c > 2$, however, the mean loop length exponent is independent of c for PS models—see (2.24). This implies that a scaling form like (6.1) cannot hold for PS models with loop class exponent $c > 2$.

The simulation results are $c' = 2.44(6)$ in $d = 2$[3] and $c' = 2.14(4)$ in $d = 3$,[1] indicating a first order phase transition in $d = 2$ and in $d = 3$. This led to the conclusion that introducing excluded volume effects between different segments to the PS model of oriented rooted self-avoiding loops drives the transition from continuous to first order.[20, 21, 23] The perturbative analytic arguments using the theory of polymer networks[20, 21, 23] yield a very good approximation to the value of the loop length distribution

exponent found in simulations. This suggests that the main mechanism responsible for the change of critical behavior in the PS model of oriented rooted self-avoiding loops is due to "local" excluded volume effects arising from forbidden attachments of paths to loops.

7. CONCLUSION

We discussed PS models with emphasis on the order of their phase transitions. We re-analyzed the old model of oriented rooted self-avoiding loops[10] and found that the resulting PS model is not self-avoiding. Hence the conclusions about the order of the phase transition, which are most relevant to the recent discussions,[1–3, 8, 9, 16, 20–23] rely on a model where most chain configurations are not self-avoiding and thus cannot represent real DNA. Our self-avoiding PS model (unique marking, with self-avoiding bridges and (unrooted) self-avoiding loops as defined in Section 4) yields a first order phase transition in both $d = 2$ and $d = 3$. It is expected that the other variants considered in Section 4 also yield a first order phase transition, but a detailed (numerical) analysis is needed to answer this question, as in the discussion in Section 4.

For the presumably more realistic model of pairs of interacting self-avoiding walks,[1, 3, 8, 9, 20, 21, 23] our results suggest an interpretation of excluded volume effects, which complements the common one.[20, 21, 23] The self-avoiding PS models of Section 4 correctly account for excluded volume effects within a loop, but overestimate excluded volume effects between different segments of a chain, due to their directed chain structure. Since this leads to a first order phase transition in $d = 2$ and $d = 3$, we conclude that the relaxation of excluded volume effects between different segments of the chain does not change the nature of the transition.

We also discussed several directed examples. One of these models displays a first order phase transition in $d = 2$. Despite being exactly solvable, these models seem to be of limited relevance to the biological problem due to their directed structure and other limitations.

The key question, as to which effects are responsible for the observed behavior in real DNA, is only partially answered by these results. With respect to directed models, which also account for different base pair sequences, it seems surprising however that simulations of melting curves[4, 35] agree so well with experimental curves. In these simulations, a heuristic critical exponent accounting for the statistical weight of internal loops has to be inserted. It has recently been suggested[5] that a value more realistic than the commonly used Fisher loop class exponent[10] $c \approx 1.763$ might be the loop length distribution exponent of the polymer network approximation[20] $c' \approx 2.115$. This suggestion implicitly assumes that the loop length

distribution exponent of a non-directed model can be interpreted as an "effective" loop class exponent within a directed model. Although appearing plausible, it seems difficult to give this assumption a quantitative meaning. On the other hand, the MELTSIM simulation approach[4] is robust against a change of parameters. Indeed, melting curves can be simulated with satisfactory coincidence for both exponents, if the cooperativity parameter for the loops is adjusted accordingly.[5] It would certainly be illuminating to obtain such simulated melting curves from a self-avoiding walk pair modeling.[1, 3, 8, 9] We also mention the recent debate about the relevance of the self-avoiding walk pair model to real DNA.[16, 22]

The stiffness of the double stranded segments does not seem to qualitatively alter the critical properties of chains,[8] as can be seen by comparing the PS model result of Section 2.1, showing that critical properties are largely independent of the path generating function. The corresponding question of introducing energy costs for bending and torsion to single stranded segments has, to our knowledge, not been discussed. Even within the PS class, it is not clear how this affects the loop class exponents, since the changes just discussed might lead to loop classes with different exponential growth constants, but there is no argument as to how the exponents might change.

In conclusion, the question of the mechanisms applying in real DNA which are responsible for the denaturation process and which explain multistep behavior as observed in melting curves, are still far from being satisfactorily answered in our opinion.

ACKNOWLEDGMENTS

We would like to acknowledge useful discussions with Stu Whittington. C.R. would like to acknowledge financial support by the German Research Council (DFG) and like to thank the Erwin Schrödinger International Institute for Mathematical Physics for support during a stay in December 2002, where part of this work was done. A.J.G. would like to acknowledge financial support from the Australian Research Council, and helpful discussions with Marianne Frommer. We acknowledge clarifying comments on an earlier version of the manuscript by Marco Baiesi, Enrico Carlon, Andreas Hanke, and Ralf Metzler.

REFERENCES

1. M. Baiesi, E. Carlon, Y. Kafri, D. Mukamel, E. Orlandini, and A. L. Stella, Inter-strand distance distribution of DNA near melting, *Phys. Rev. E* **67**:21911–21917 (2002); cond-mat/0211236.

2. M. Baiesi, E. Carlon, E. Orlandini, and A. L. Stella, A simple model of DNA denaturation and mutually avoiding walk statistics, *Eur. Phys. J. B* **29**:129–134 (2002); cond-mat/0207122.
3. M. Baiesi, E. Carlon, and A. L. Stella, Scaling in DNA unzipping models: denaturated loops and end-segments as branches of a block copolymer network, *Phys. Rev. E* **66**:21804–21812 (2002); cond-mat/0205125.
4. R. D. Blake *et al.*, Statistical mechanical simulation of polymeric DNA melting with MELTSIM, *Bioinformatics* **15**:370–375 (1999).
5. R. Blossey and C. Carlon, Reparametrizing loop entropy weights: Effect on DNA melting curves, preprint (2002); cond-mat/0212457.
6. M. Bousquet-Mélou and A. J. Guttmann, Enumeration of three-dimensional convex polygons, *Ann. Comb.* **1**:27–53 (1997).
7. R. Brak, A. L. Owczarek, and T. Prellberg, Exact scaling behavior of partially convex vesicles, *J. Stat. Phys.* **76**:1101–1128 (1994).
8. E. Carlon, E. Orlandini, and A. L. Stella, The roles of stiffness and excluded volume in DNA denaturation, *Phys. Rev. Lett.* **88**:198101–198104 (2002); cond-mat/0108308.
9. M. S. Causo, B. Coluzzi, and P. Grassberger, Simple model for the DNA denaturation transition, *Phys. Rev. E* **62**:3958–3973 (2000).
10. M. E. Fisher, Effect of excluded volume on phase transitions in biopolymers, *J. Chem. Phys.* **45**:1469–1473 (1966).
11. M. E. Fisher, Walks, walls, wetting, and melting, *J. Stat. Phys.* **34**:667–729 (1984).
12. P. Flajolet and R. Sedgewick, Analytic combinatorics: Functional equations, rational and algebraic functions, INRIA preprint 4103 (2001).
13. T. Garel and H. Orland, On the role of mismatches in DNA denaturation, Saclay preprint T03/045 (2003); cond-mat/0304080.
14. A. J. Guttmann, Asymptotic analysis of power-series expansions, in *Phase Transitions and Critical Phenomena*, C. Domb and J. L. Lebowitz, eds., Vol. 13 (Academic, New York, 1989), pp. 1–234.
15. A. J. Guttmann and T. Prellberg, Staircase polygons, elliptic integrals, Heun functions, and lattice Green functions, *Phys. Rev. E* **47**:R2233–R2236 (1993).
16. A. Hanke and R. Metzler, Comment on "Why is the DNA denaturation transition first order?" *Phys. Rev. Lett.* **90**:159801 (2003); cond-mat/0110164.
17. A. Hanke and R. Metzler, Entropy loss in long-distance DNA looping, preprint (2002); cond-mat/0211468.
18. E. J. Janse van Rensburg, *The Statistical Mechanics of Interacting Walks, Polygons, Animals, and Vesicles* (Oxford University Press, New York, 2000).
19. I. Jensen, A parallel algorithm for the enumeration of self-avoiding polygons on the square lattice, *J. Phys. A: Math. Gen.* **36**:5731–5745 (2003); cond-mat/0301468.
20. Y. Kafri, D. Mukamel, and L. Peliti, Why is the DNA denaturation transition first order?, *Phys. Rev. Lett.* **85**:4988–4992 (2000); cond-mat/0007141.
21. Y. Kafri, D. Mukamel, and L. Peliti, Melting and unzipping of DNA, *Eur. Phys. J. B* **27**:135–146 (2001); cond-mat/0108323.
22. Y. Kafri, D. Mukamel, and L. Peliti, Reply to comment by Hanke and Metzler, preprint (2001); cond-mat/0112179, cond-mat/0302589.
23. Y. Kafri, D. Mukamel, and L. Peliti, Denaturation and unzipping of DNA: Statistical mechanics of interacting loops, *Physica A* **306**:39–50 (2002).
24. E. B. Kolomeisky and J. P. Straley, Universality classes for line-depinning transitions, *Phys. Rev. B* **46**:12664–12674 (1992).
25. R. Lipowsky, Typical and exceptional shape fluctuations of interacting strings, *Europhys. Lett.* **15**:703–708 (1991).

26. D. Marenduzzo, A. Trovato, and A. Maritan, Phase diagram of force-induced DNA unzipping in exactly solvable models, *Phys. Rev. E* **64**:031901–031912 (2001); cond-mat/0101207.
27. D. Marenduzzo, S. M. Bhattacharjee, A. Maritan, E. Orlandini, and F. Seno, Dynamical scaling of the DNA unzipping transition, *Phys. Rev. Lett.* **88**:28102–28105 (2002); cond-mat/0103142.
28. N. Madras and G. Slade, *The Self-Avoiding Walk* (Birkhäuser, Boston, 1993).
29. M. Peyrard and A. R. Bishop, Statistical mechanics of a nonlinear model for DNA denaturation, *Phys. Rev. Lett.* **62**:2755–2758 (1989).
30. D. Poland and H. A. Scheraga, Phase transitions in one dimension and the helix-coil transition in polyamino acids, *J. Chem. Phys.* **45**:1456–1463 (1966).
31. D. Poland and H. A. Scheraga, Occurrence of a phase transition in nucleic acid models, *J. Chem. Phys.* **45**:1464–1469 (1966).
32. T. Prellberg and R. Brak, Critical exponents from nonlinear functional equations for partially directed cluster models, *J. Stat. Phys.* **78**:701–730 (1995).
33. D. W. Sumners and S. G. Whittington, Knots in self-avoiding walks, *J. Phys. A* **21**:1689–1694 (1988).
34. N. Theodorakopoulos, T. Dauxois, and M. Peyrard, Order of the phase transition in models of DNA thermal denaturation, *Phys. Rev. Lett.* **85**:6–9 (2000); cond-mat/0004487.
35. R. M. Wartell and A. S. Benight, Thermal denaturation of DNA molecules: A comparison of theory with experiment, *Phys. Rep.* **126**:67–107 (1985).