



# ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

SISSA  ISAS

Scuola Internazionale Superiore di Studi Avanzati  
International School for Advanced Studies

## Thermal and Mechanical Denaturation of DNA

Thesis submitted for the degree of  
*Magister Philosophiæ*

**Candidate:**  
Davide Marenduzzo

**Supervisor:**  
Prof. Amos Maritan

October 2000





Scuola Internazionale Superiore di Studi Avanzati  
International School for Advanced Studies

# Thermal and Mechanical Denaturation of DNA

Thesis submitted for the degree of  
*Magister Philosophiæ*

**Candidate:**  
Davide Marenduzzo

**Supervisor:**  
Prof. Amos Maritan

October 2000



# Contents

|   |           |
|---|-----------|
| <b>Preface</b>  | <b>3</b>  |
| <b>1 The biological problem</b>   | <b>7</b>  |
| 1.1 Introduction . . . . .  | 7         |
| 1.2 Micromanipulation experiment . . . . .  | 8         |
| 1.3 DNA replication . . . . .   | 9         |
| <b>2 Model and solution</b>   | <b>11</b> |
| 2.1 Model and approximations . . . . .  | 11        |
| 2.2 Exact solution of the model . . . . .   | 13        |
| 2.2.1 Behaviour of the models at non zero force . . . . .   | 14        |
| 2.3 Scaling laws for thermal melting and unzipping . . . . .  | 20        |
| <b>3 Discussion and role of disorder</b>  | <b>23</b> |
| 3.1 Physical explanation of reentrance . . . . .  | 23        |
| 3.1.1 Re-entrance in the continuum . . . . .  | 24        |
| 3.2 Quenched disorder and analytical treatment . . . . .  | 28        |
| 3.2.1 Integer moments of the canonical partition function: the<br>Cayley tree (thermal melting) . . . . . | 28        |
| 3.2.2 Y configurations . . . . .  | 32        |
| 3.2.3 Full model with bubbles . . . . .   | 33        |
| <b>Conclusion</b>   | <b>35</b> |
| <b>Appendix A</b>   | <b>36</b> |
| <b>Appendix B</b>   | <b>39</b> |
| <b>Appendix C</b>   | <b>44</b> |
| <b>Bibliography</b>   | <b>48</b> |



# Preface

The goal of this thesis is to give a theoretical discussion of a DNA mechanical unzipping experiment, from the statistical mechanics point of view. The structure of this thesis is as follows.

In the first chapter, we give the biological and physical (experimental) motivations for our study.

In the second chapter, we introduce a simple model for the mechanical denaturation of DNA and give the details leading to the exact solution in the case of homogeneous DNA, i.e. all base pairs are AT or CG. We discuss the  $D$ -dimensional case in this chapter with the generating function scheme.

In the third chapter, we discuss our results and give a physical explanation for the phase diagram that we observe. We also discuss in some details a simplified case when heterogeneity can be considered exactly with the replica method.

Finally, we highlight the results and discuss some possible extension of this work.





# Chapter 1

## The biological problem

DNA is a biomolecule basic for all life functions and responsible for transmitting genetic codes. It consists of two strands paired together. Each strand is made up of a succession of bases chosen among the four basic constituents, which are adenine, thymine, cytosine and guanine (A,T,C,G). The strands are kept together by base pairing energies (hydrogen bonds) between the bases, which are specific and can be formed either between C and G or between A and T.

### 1.1 Introduction

In recent years, micro and nanomanipulation of single biological macromolecules has become feasible due to a dramatic improvement of experimental techniques. By using devices such as optical tweezers [1, 2], soft microneedles [3] and atomic force microscopes [4], it is now possible to access physical and mechanical properties of fundamental biological objects, namely proteins, nucleic acids, and molecular motors, on the scale of individual molecules. Special effort has been devoted to the measurement of force-elongation characteristics of double stranded DNA molecules (dsDNA), determining its response to external forces and torques in the absence of enzymes. The *mechanical unzipping* of dsDNA structure by a force pulling the end of one of the two strands, the end of the other strand being anchored to some physical support, has been studied by Bockelmann *et al.*[5, 6], who have measured the average force along the opening of the two strands. Mechanical forces are in fact exerted on the DNA molecule by different enzymes during the process of DNA replication or transcription [7, 8]. On the other hand, the double helical structure of dsDNA may be disrupted 'in vitro' by changing pH, solvent conditions and/or temperature [9]. This transition is known as *melting denaturation*, and it has been long studied by theoretical physicists [10]. Instead, only very recently DNA mechanical denaturation has been the subject of theoretical studies.

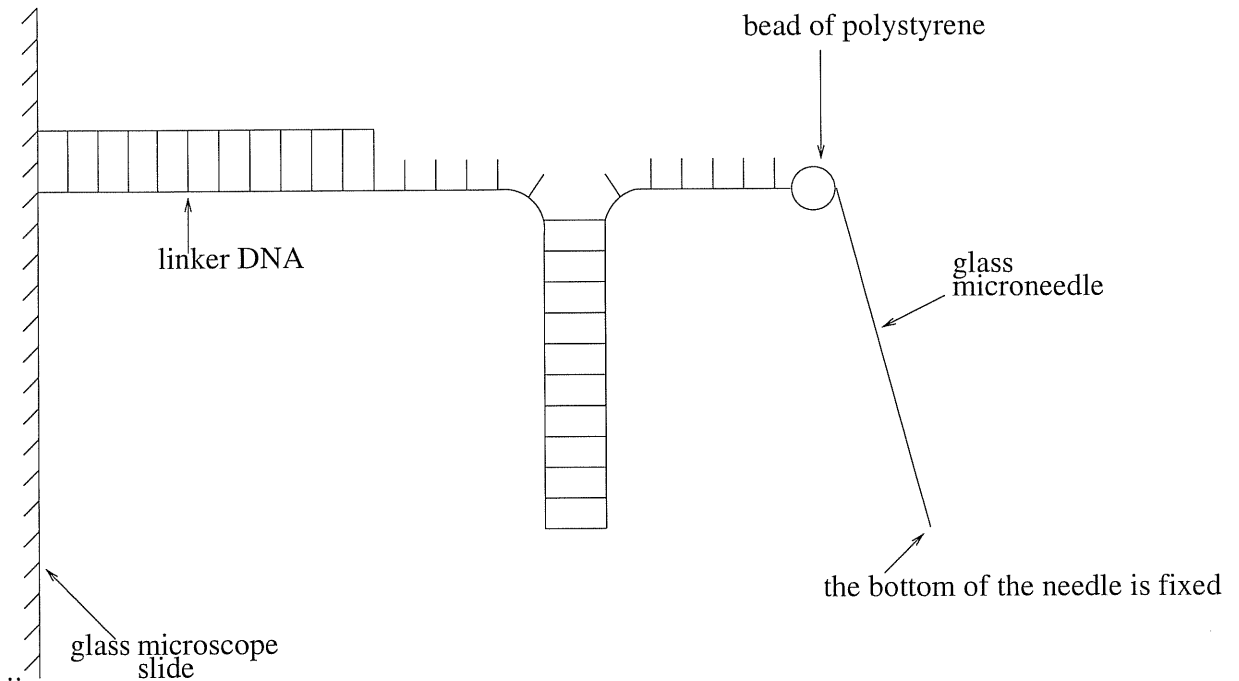


Figure 1.1: Schematical representation of a DNA micromechanical experiment, with the use of a soft microneedle.

## 1.2 Micromanipulation experiment

In Fig.1.1 we show a schematical representation of a DNA micromechanical experiment (see [5, 6] for details).

During a typical DNA unzipping experiment, the distance between the two ends of two strands of DNA is controlled, and the average force necessary to keep the two ends at that fixed distance is measured. In the schematical view presented in Fig.1.1, one end is kept fixed via a microneedle and the other end is displaced at a constant velocity (typically several nm/s). The force is measured by detecting the average deflection of the microneedle. The typical order of magnitude experimental value of the force is some pN (15 pN on average, but the force is sequence dependent). Though the natural ensemble would thus be the fixed stretch ensemble, one can consider the conjugate fixed strength ensemble where calculations can be performed analytically in any dimension. The value of the average force is sequence dependent and indeed this has been used as a way to partially sequence the DNA molecule. Experimentalists report the presence of hysteresis in the force vs. extension characteristic curves, indicating a first order transition.

## 1.3 DNA replication

While experiments on single molecule micromanipulation are the major motivation for DNA unzipping analysis in theoretical models, it is also important to point out that many speculations and theoretical hypothesis are being made by physicists and statistical mechanists on the *interpretation* of these experiments.

We deem the following speculation interesting and motivating for our work. It has recently been hypothesized by Bhattacharjee[13] that equilibrium and non-equilibrium calculations of DNA force-induced unzipping may shed some light on the still not fully understood mechanism of DNA replication.

The replication process starts with a DnaA protein attaching itself, in interaction with the membrane, at the “origin” to start the Y-fork ( or “eye”). The next step is the binding of the various enzymes/proteins like single-strand binding (SSB) protein, helicase, topoisomerase, polymerases etc. Unlike the latter proteins/enzymes, the functionality of DnaA protein is not well understood. It is apparent that the whole process requires a strong correlation in space and time, whose source or origin is not obvious. In order to explain such correlations in subsequent processes, a structural unit “replisome” has been postulated which has never been isolated.

Bhattacharjee makes the hypothesis that

- the function of DnaA is actually to exert a pulling force (let us take this time  $t = 0$ ) with a force which is close to, but slightly below, the critical force for unzipping. This force tends to form a Y-fork or an eye-type bubble (depending on the location of the origin).
- the large length and time scales for the unzipping process, leads to a space time correlation in the initial nonequilibrium dynamics of unzipping. It is postulated that this correlation controls the subsequent processes especially the dynamics of bindings of the subsequent enzymes/proteins.

This hypothesis implying that the statics of DNA unzipping and subsequently its dynamics might help one understanding the principles behind DNA replication is in itself an interesting motivation for our study.



# Chapter 2

## Model and solution

### 2.1 Model and approximations

As is customary in Statistical Mechanics, we choose a minimal model which contains only the basic details of the system. The belief is indeed that the neglected details will not have a drastic effect on the general conclusions that we can draw from such a simple model.

- We focus on a model on the  $d = 2$  lattice (see Fig.2.1) (even though the solution which we give is more general and works for the  $D = d + 1$  dimensional case), where the two strands of the DNA molecule are mimicked by two directed and self avoiding walks. Similar approximations date back to the original model of Poland and Scheraga in the field of thermal denaturation of DNA.
- The hydrogen bonds between bases labelled by the same monomer index (or projection along the lattice diagonal  $(1, 1)$ ) is modelled by a contact attractive interaction  $-\epsilon < 0$ .
- The presence of the external force,  $\vec{g}$  is inserted into the model by giving every pair of walks a Boltzmann weight proportional to  $\exp(\beta \vec{g} \cdot \vec{x})$ . To obtain a meaningful geometry, the external force is taken in the  $(1, -1)$  direction, i.e.  $\vec{g} = g(1, -1)$

In particular, we note that the double helical nature of the molecule is not treated (each strand is a directed self avoiding chain). Secondly, we do not consider stacking interaction between the bases. We also consider for the moment homogeneous DNA, in contrast to the most realistic case in which the sequence of base pairs is random (and is usually taken in the simplest model to be uncorrelated). Self-avoidance of the strands *is* taken into account, but only partially, as we restrict our analysis to a subclass of self-avoiding chains. Finally, “wrong base pairing”, i.e. hydrogen bonds forming between non corresponding base pairs are

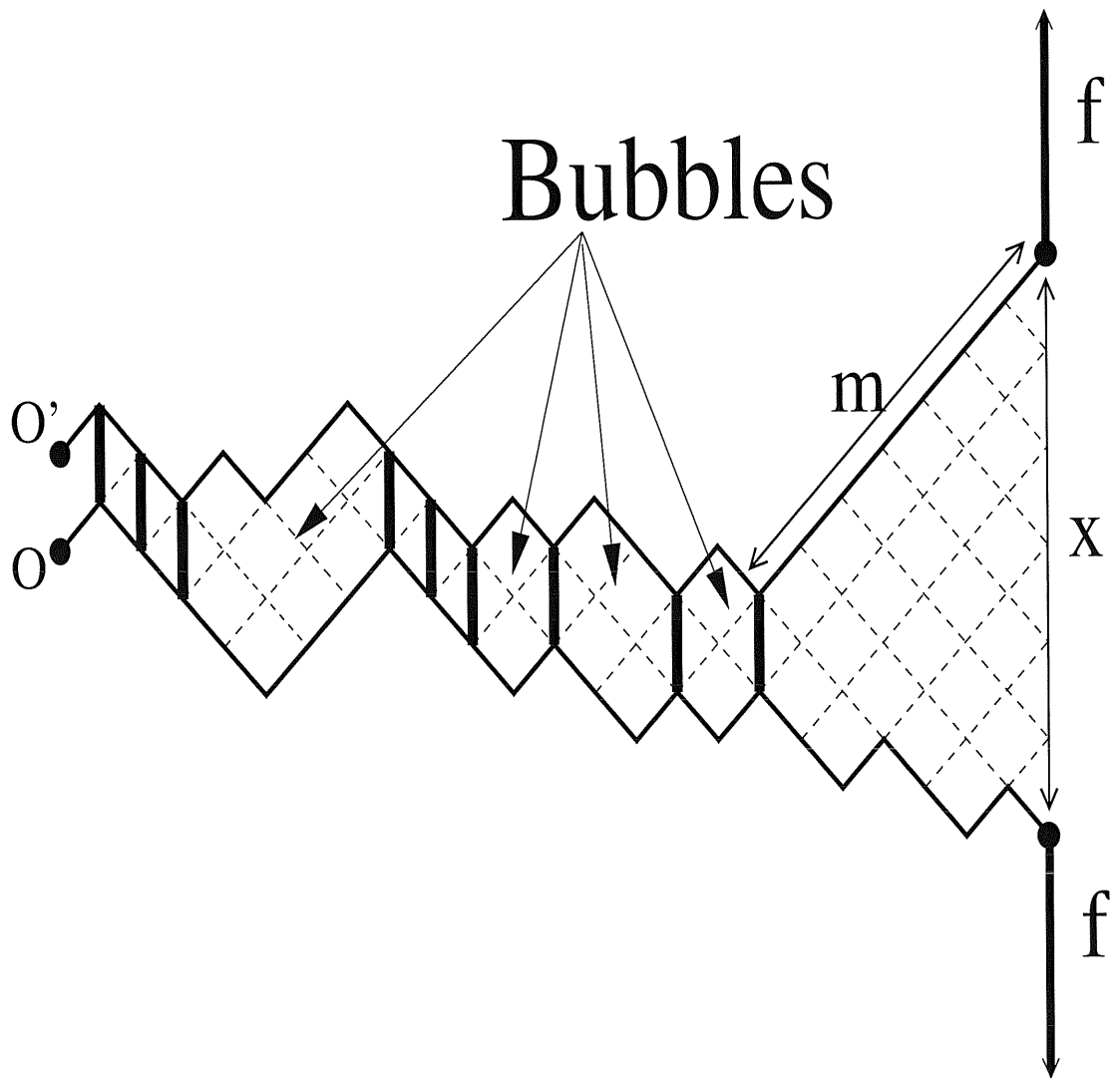


Figure 2.1: An example of the configurations that are taken into considerations. Denaturation bubbles are also present and shown.

not possible in this model. All these approximations are to be considered when we want to extract significant and experimentally relevant results from our calculations. On the other hand, such approximations are quite generally expected to capture the essence of the phase transition in similar models of DNA thermal denaturation.

## 2.2 Exact solution of the model

We are now ready to present the analytical solution of the model presented above.

For completeness we give the details for the solution in the  $D = d + 1$  dimensional hypercubic lattice  $\mathbf{Z}^D$ . The two strands of a homogeneous DNA molecule with  $N$  base pairs are mimicked by two SAWs, directed along the  $(1, \dots, 1)$  direction. The two chains have one end in common, while at the other end a force  $\vec{g}$  is pulling in the  $(1, -1, 0, \dots, 0)$  direction. As reported above, the walks gain a binding energy  $-\epsilon$  ( $\epsilon > 0$ ) every time the bases with the same monomer index (or same projection upon  $(1, \dots, 1)$ ) interact, i.e. *wrong base pairing* is forbidden.

The canonical partition function for two  $N$ -step directed self avoiding walks (DSAWs) is:

$$Z_N(\beta, \vec{g}) = \sum_{\vec{x} \in \mathbf{Z}^D} p_N(\vec{x}, \beta\epsilon) \exp(\beta\vec{g} \cdot \vec{x}), \quad (2.1)$$

where  $p_N(\vec{x}, \beta\epsilon)$  is the canonical partition function of two directed interacting strands whose last base pairs are at relative distance  $\vec{x}$ . The following recursion relation holds for  $p_N(\vec{x}, \beta\epsilon)$ :

$$p_{N+1}(\vec{x}, \beta\epsilon) = \sum_{i,j=1}^D p_N(\vec{x} - \vec{e}_i + \vec{e}_j, \beta\epsilon) (1 + (\exp(\beta\epsilon) - 1)\delta_{\vec{x}, \vec{0}}), \quad (2.2)$$

where  $\vec{e}_i$ ,  $i = 1, \dots, D$ , are the canonical euclidean versors of the  $D$ -dimensional space and  $\delta$  is the Kronecker delta.

These and similar equations have been intensively studied, at  $\vec{g} = \vec{0}$ , within simple models of DNA thermal melting [17, 19], and, in  $D = 2$ , within the context of random walk adsorption [20] and wetting problems [21]. Note that the choice of the  $(1, \dots, 1)$  direction along which the walks are directed, is crucial in allowing us to write local recursion relations. It is thus no surprise that the model belongs to the same universality class of random walks in  $d = D - 1$  dimensions [22].

It is well known (see [17] for a recent review), that, in the absence of a pulling force, within the model defined by equation (2.2), dsDNA undergoes a phase transition between a low temperature double stranded phase and a high temperature denaturated state only for  $D > 3$  ( $d > 2$ ): in  $D = 2, 3$  dsDNA remains double stranded at all temperatures.

However, we note that in (2.2) the two strands are allowed to cross each other, without any restriction. This seems rather unphysical, since every real

chain has a finite "hard core" distance. Thus, one should consider also the case in which crossing between different strands is forbidden (recent studies of both homogeneous [23] and heterogeneous [24] DNA melting transition have indeed considered the  $d = 1$  case with forbidden crossing). Let us focus for the moment on the  $D = 2$  case, with no pulling force, where the effect of forbidding crossing is most dramatic. Here, the direction  $(1/\sqrt{2}, 1/\sqrt{2})$  can be identified as "time", and its normal  $(1/\sqrt{2}, -1/\sqrt{2})$  as "space" ( $x$ ). The model with crossing (w.c.) implies no restriction on the relative distance  $x$ , whereas the one without crossing (w.o.c.) implies that  $x$  cannot change sign, *e.g.*  $x \geq 0$ . The model w.o.c. is equivalent to surface adsorption models previously considered [20]. While the thermal melting of the two strands w.c. takes place at  $T_c = \infty$ , in the model w.o.c.  $T_c = \epsilon/\log(4/3)$  as it can be deduced from ref.[20]. To elucidate this point further, we have tackled an intermediate model, where we do not forbid crossing, but we make it disadvantageous, by assigning a cost  $V > 0$  each time the strands pass through one another. With calculations similar to those reported below, we find that melting takes place at the critical temperature (as a function of  $V$ ):

$$T_c(V) = \epsilon \left[ \log \left( \frac{4e^V}{3e^V + 1} \right) \right]^{-1}. \quad (2.3)$$

As expected, as  $V \rightarrow \infty$ , equation (2.3) yields  $T_c \rightarrow \epsilon/\log(4/3)$ , whereas  $T_c \sim \epsilon/V$  as  $V \rightarrow 0$ .

To implement the non-crossing constraint in the model in generic dimension, we require that, if the DSAWs join, one coming from direction  $\vec{e}_i$  and the other from direction  $\vec{e}_j$  (with  $i \neq j$ ), when they divide again, they cannot both proceed along their previous direction; i.e we forbid that the first walk goes along direction  $\vec{e}_i$  and the second one along  $\vec{e}_j$ . This excludes one configuration out of  $D(D-1)$  at the splitting point (the remaining  $D$  possibilities would lead to the zipping of the two strands). As regards the  $D = 3$  thermal melting, again the model w.c. has  $T_c = \infty$ , while, if crossing is forbidden as described above, the critical temperature is  $T_c = \epsilon/\log(9/8)$ . This result was also found in a similar calculation by Rubin [25]. For  $D > 3$  the models with and without crossing both undergo a denaturation transition at a finite, though different, critical temperature, so that the effect of forbidding crossing here is less important (the critical temperatures are the same at the leading order  $1/D$  in the  $D \rightarrow \infty$  expansion, as expected).

### 2.2.1 Behaviour of the models at non zero force

Let us now turn to the models with the pulling force  $\vec{g}$  in generic dimension  $D$ , with  $\vec{g} = (g, -g, 0, \dots, 0)$ . We can find the asymptotic value of the canonical partition function (2.1) by locating [17, 19, 26] the singularity closest to the origin of its generating function:



$$Z(z, \beta\epsilon, \beta g) = \sum_{N=0}^{\infty} z^N \sum_{\vec{x}} p_N(\vec{x}, \beta\epsilon) \exp(\beta \vec{g} \cdot \vec{x}), \quad (2.4)$$

which we will also refer to as the grand partition function. We stress that  $p_N(\vec{x}, \beta\epsilon)$  is a generic partition function: its detailed form will depend on whether or not we allow crossing.

To proceed, it is useful to partition the DNA molecule in ds helices (with the strands attached to each other) and *bubbles*, sequences in which the two chains share just the first and the last base pairs (see Fig.2.1.); we also single out the contribution of the unzipped end of the two DSAWs, *i.e.* the part from the last contact to the end. In this way, the grand partition sum (2.4) can be expressed as:

$$Z(z, \beta\epsilon, \beta g) = \frac{1}{1 - Dz \exp(\beta\epsilon)} \frac{\exp(\beta\epsilon)}{1 - \frac{\exp(\beta\epsilon)}{1 - Dz \exp(\beta\epsilon)} B(z)} S(z, \beta g), \quad (2.5)$$

where we have defined:

$$B(z) = \sum_{N=2}^{\infty} z^N b_N, \quad S(z, \beta g) = 1 + \sum_{N=1}^{\infty} z^N \sum_{\vec{x}} \left[ c_{N-1}(\vec{x}) - c_{N-1}(\vec{0}) \delta_{\vec{x}, \vec{0}} \right] \exp(\beta \vec{g} \cdot \vec{x}) \quad (2.6)$$

In (2.6),  $b_N$  is the number of  $2N$ -step bubbles, and  $c_N(\vec{x})$  the partition function of two  $N$ -step DSAWs never touching each other and having their last sites at a mutual distance  $\vec{x}$ , and their initial sites at a relative distance  $\vec{e}_{i_0} - \vec{e}_{j_0}$  for some  $i_0 \neq j_0$ . By summing over all possibilities for initial conditions (see equation (2.8) below),  $b_N = \sum_{i \neq j; i, j=1}^D c_{N-2}(\vec{e}_i - \vec{e}_j)$ , so that we need to find an explicit expression for the  $c_N$ 's. The equations they obey are:

$$c_{N+1}(\vec{x}) = \sum_{i, j=1}^D c_N(\vec{x} - \vec{e}_i + \vec{e}_j) - c_N(\vec{0}) \sum_{i \neq j; i, j=1}^D \delta_{\vec{x}, \vec{e}_i - \vec{e}_j}. \quad (2.7)$$

Notice that in eq.(2.7)  $\vec{0}$  acts as a sink or absorbing state, *i.e.* once the two walks join, they never leave. In this way,  $c_N(\vec{x})$  with  $\vec{x} \neq \vec{0}$  counts the number of pairs of walks that never touch each other (and with the last monomers at a relative distance  $\vec{x}$ ), while  $c_N(\vec{0})$  counts the number of remaining pairs of walks that at some point come into contact and then remain stuck together. This last quantity plays the role of an arbitrary constant and has to be subtracted away from the final counting as in (2.6) (see [25] for another example in which this prescription was used in a  $D = 3$  example).

In order to simplify the presentation from now on we restrict our calculations on models w.o.c. (detailed calculations for this case and the outline of calculations for the simpler models w.c. are deferred to Appendix A). The initial conditions for such models are:

$$c_0(\vec{x}) = \left( \sum_{i \neq j; i, j=1}^D \delta_{\vec{x}, \vec{e}_i - \vec{e}_j} \right) - \delta_{\vec{x}, \vec{e}_{i_0} - \vec{e}_{j_0}}, \quad (2.8)$$

where  $\vec{e}_{i_0}$  and  $\vec{e}_{j_0}$  are the directions forbidden as described in the previous subsection.

By performing a Fourier and a discrete Laplace transform on eq.(2.7), and by using the relation between  $b_N$  and  $c_N$  given above, we can derive an explicit expression for the bubble generating function  $B(z)$  (see eq. (3.35)). The singularities of the second denominator of the grand partition function (2.5) are necessary to find the melting temperature. We find that the equations locating these two singularities are :

$$z_1 = \frac{1}{D^2}, \quad (2.9)$$

$$\exp(-\beta\epsilon) - Dz_2 = B(z_2). \quad (2.10)$$

The first singularity  $z_1$  is leading to the usual random walk behavior in the absence of any interaction. The second singularity  $z_2(\beta\epsilon)$  is a function of the strength of the attractive interaction between the two strands, determining thus the behaviour in the native zipped phase. At  $g = 0$ , the critical temperature for thermal melting is obtained when the two singularities above coincide ( $z_1 = z_2$ ). Unlike the models w.c., which have  $T_c = \infty$  in  $D \leq 3$ , those w.o.c. have a finite critical temperature in any  $D \geq 2$  (see Appendix A for further details).

As regards the force-dependent third factor in (2.5), it reads:

$$S(z, \beta g) = zS_{sing}(z, \beta g) + 1 - zc(z, \vec{0}). \quad (2.11)$$

A third new singularity, depending on the external force, and fundamental in our calculations arises when computing:

$$S_{sing}(z, \beta g) \equiv \int_{[-\pi, \pi]^D} \frac{d^D \vec{q}}{(2\pi)^D} \tilde{c}(z, \vec{q}) \sum_{\vec{x}} \exp[(\beta \vec{g} - i\vec{q}) \cdot \vec{x}], \quad (2.12)$$

where  $\tilde{c}(z, \vec{q})$  and  $c(z, \vec{0})$  are given in eqs. (3.33,3.34). Notice that in eq.(2.12) one can immediately compute the integrals on  $dq_3 \dots dq_D$ , by using the well known identity  $2\pi\delta(q) = \sum_x \exp(-iqx)$ , so that one is left with the double integral:

$$\int_{-\pi}^{\pi} \frac{dq_2}{(2\pi)} \int_{-\pi}^{\pi} \frac{dq_1}{(2\pi)} \tilde{c}(z, q_1, q_2, 0, \dots, 0) \sum_{x_1, x_2} \exp[(\beta g - iq_1)x_1 + (-\beta g - iq_2)x_2]. \quad (2.13)$$

We now give details for the evaluation of the inner integral only, the outer one being equivalent. Defining  $h(q_1, q_2) \equiv \left( \tilde{c}(z, q_1, q_2, 0, \dots, 0) \right)^{-1}$ , we aim at proving

that the complex translation  $q_1 \rightarrow q_1 - i\beta g$  can be performed in the integral  $\int_{-\pi}^{\pi} \frac{dq_1}{2\pi} \frac{\exp(-iq_1 x_1)}{h(q_1, q_2)}$ . Notice that  $h$  is a periodic function of  $q_1$ . We extend  $q_1$  to complex numbers and consider the contour  $\gamma$  as shown in Fig. 2.2. By the residue theorem, one can write:

$$\begin{aligned} & \int_{-\pi}^{\pi} \frac{dq_1}{2\pi} \frac{\exp(-iq_1 x_1)}{h(q_1, q_2)} + \int_0^{-\beta g} \frac{idy}{2\pi} \frac{\exp(-i\pi x_1 + x_1 y)}{h(\pi + iy, q_2)} - \\ & \int_{-\pi}^{\pi} \frac{dq_1}{2\pi} \frac{\exp(-iq_1 x_1 - \beta g x_1)}{h(q_1 - i\beta g, q_2)} + \int_{-\beta g}^0 \frac{idy}{2\pi} \frac{\exp(i\pi x_1 + x_1 y)}{h(-\pi + iy, q_2)} = \sum_{q_0} \text{Res}(\mathcal{H}, q_0), \end{aligned} \quad (2.14)$$

where the sum runs over all poles  $q_0$  of  $\mathcal{H} \equiv \frac{\exp(-iq_1 x_1)}{h(q_1, q_2)}$  (at fixed  $q_2$ ) inside the contour  $\gamma$  and  $\text{Res}(\mathcal{H}, q_0)$  is the residue of  $\mathcal{H}$  at  $q_0$ . Note that due to the periodicity of  $h$  the second and fourth terms in the previous equation cancel and thus the complex translation can be performed provided that no pole of  $\mathcal{H}$  can be found inside the contour  $\gamma$ . An equivalent conclusion can be reached for the outer integral on  $dq_2$ . The condition of having no poles inside the contours of integration is satisfied as long as  $z < z_3(\beta g)$ , where:

$$z_3(\beta g) = \frac{1}{(D-2)^2 + 2 + 2 \cosh(2\beta g) + 4(D-2) \cosh(\beta g)} \quad (2.15)$$

is found by solving  $h(-i\beta g, i\beta g) \propto (1 - zf(-i\beta \vec{g})) = 0$ , with  $f(\vec{q})$  defined in eq. (3.31). In this case, we find consistently that  $z_3(\beta g)$  is just the singularity in the resulting expression  $S(z, \beta g) = 1 + z [\tilde{c}(z, -i\beta \vec{g}) - c(z, \vec{0})]$ .

As  $z_3$  is always smaller than  $z_1$  (for  $g \neq 0$ ), the singularity closest to the origin has to be determined between  $z_2$ , controlled by the attractive energy  $\epsilon$ , and  $z_3$ , controlled by the pulling force  $g$ . If  $z_2 < z_3$ , the DNA molecule is zipped, otherwise it is unzipped. The free energy per monomer  $f$  and the average distance between the two ends (projected onto  $\hat{g} \equiv (1, -1, 0, \dots, 0)$ ),  $\langle x_g \rangle$ , are defined as:

$$f \equiv \lim_{N \rightarrow \infty} -T \frac{\log [\sum_{\vec{x}} p_N(\vec{x}, \beta \epsilon) \exp(\beta \vec{g} \cdot \vec{x})]}{N} \quad (2.16)$$

$$\langle x_g \rangle \equiv \lim_{N \rightarrow \infty} \langle \hat{g} \cdot \vec{x} \rangle = -\frac{\partial f}{\partial g} N \quad (2.17)$$

These quantities read in the thermodynamic limit:

$$f = T \log z_2(\beta \epsilon) \quad g < g_c(T, \epsilon) \quad (2.18)$$

$$\frac{\langle x_g \rangle}{N} = 0 \quad g < g_c(T, \epsilon) \quad (2.19)$$

$$f = T \log z_3(\beta g) \quad g > g_c(T, \epsilon) \quad (2.20)$$

$$\frac{\langle x_g \rangle}{N} = z_3(\beta g) [4 \sinh(2\beta g) + 4(D-2) \sinh(\beta g)] \quad g > g_c(T, \epsilon) \quad (2.21)$$

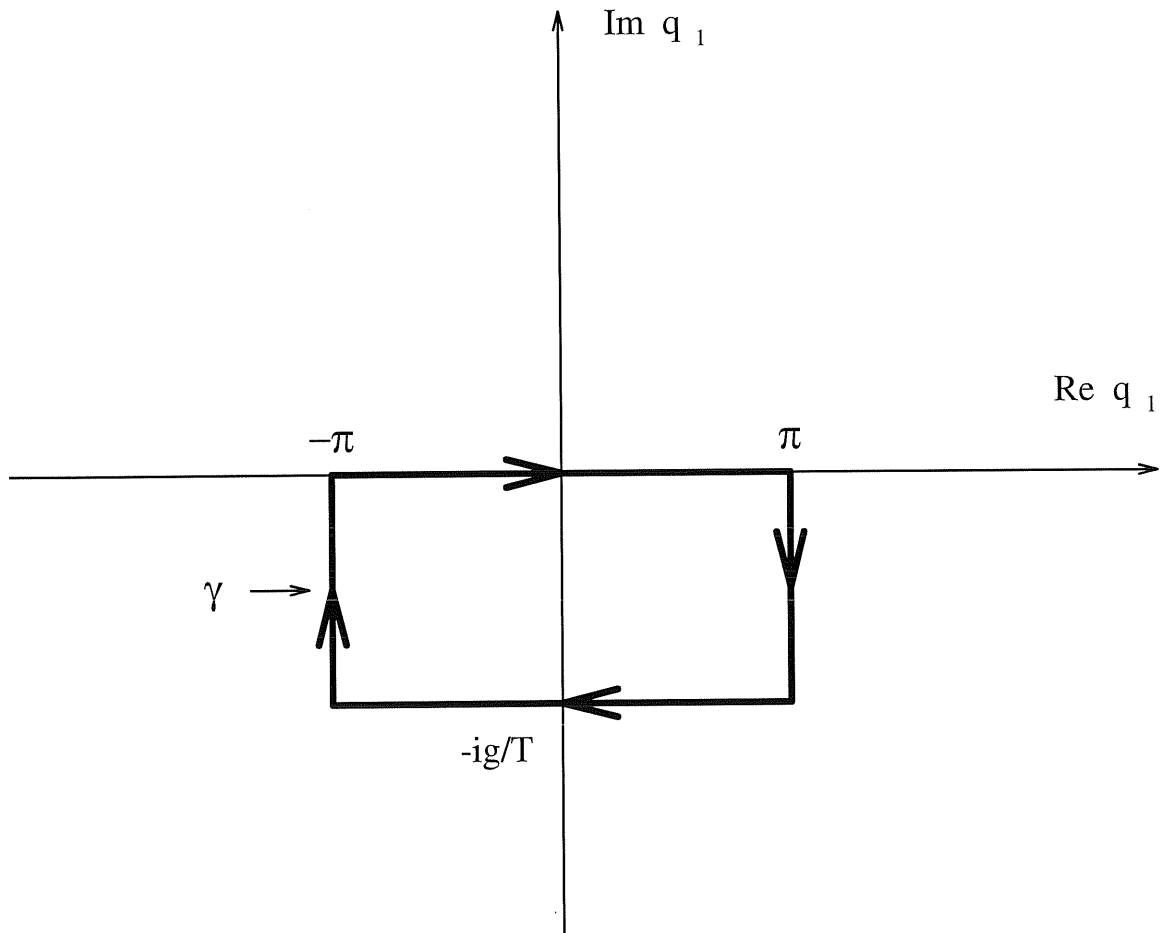


Figure 2.2: Contour of integration  $\gamma$  in the complex plane used for the evaluation of the integrals in (2.14).

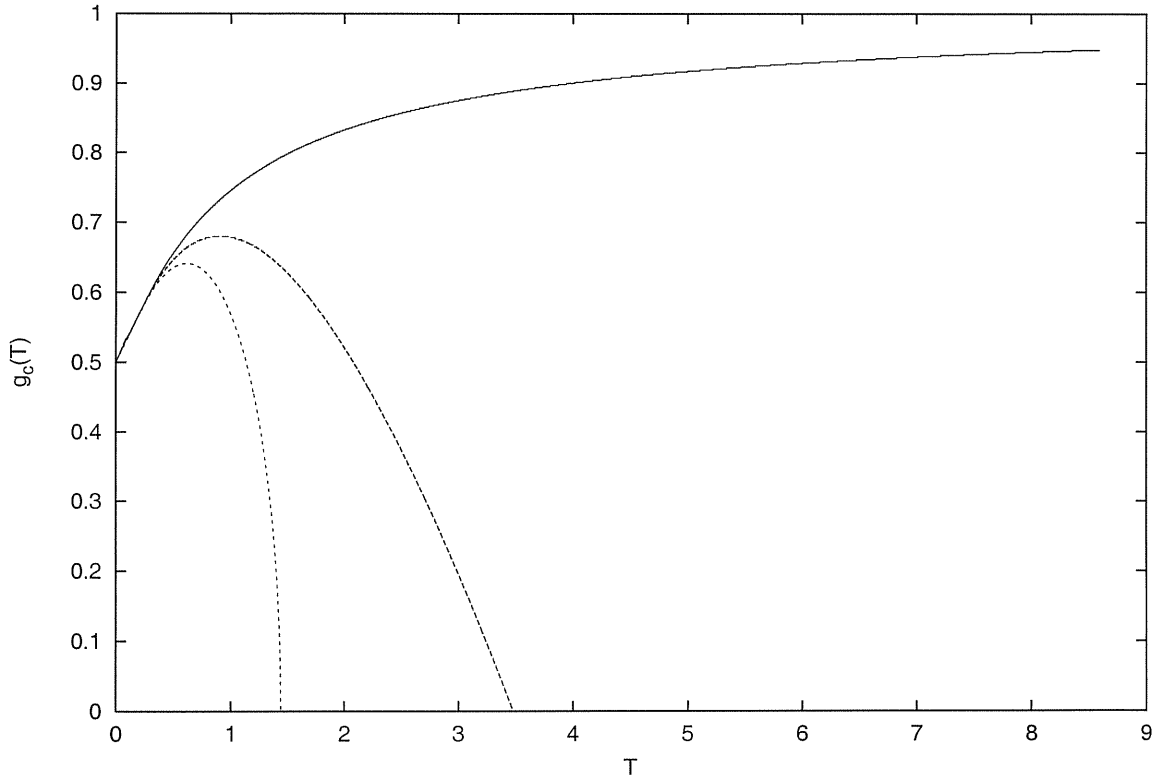


Figure 2.3: Phase diagram for two-dimensional models of DSAWs ( $\epsilon = 1$  in the figure). The solid curve refers to the model with crossing, the long-dashed one to that without crossing and the dashed one to the model which considers only Y-shaped configurations.

where the critical force  $g_c(T, \epsilon)$  is found by imposing  $z_2(\beta\epsilon) = z_3(\beta g)$ , as given in eqs. (2.10, 2.15). The above equations show the existence of a first order transition at  $g = g_c(T, \epsilon)$ , if  $g_c(T, \epsilon) > 0$ . The phase diagram for the models w.c. can be found exactly in the same way (see Appendix A for some details on this). It is interesting to notice that the singularity  $z_3(\beta g)$  does not depend on whether or not we allow crossing, whereas  $z_2$  in the w.c. case is different from the w.o.c. case. This is enough of course to make the whole phase diagram different for the two kinds of models, as will be discussed in section III.

We stress here that the main interest in using directed walks is that they are a sub-class of SAWs in the same dimension. However, qualitatively similar results can be obtained also with simple random walks (RWs), but in that case it would be not physically meaningful to forbid crossing as done above.

Experimentally, however, it should not be hard to set up an experiment closely related to the calculation we have performed, by stretching the ds molecule in one given direction before applying the external force.

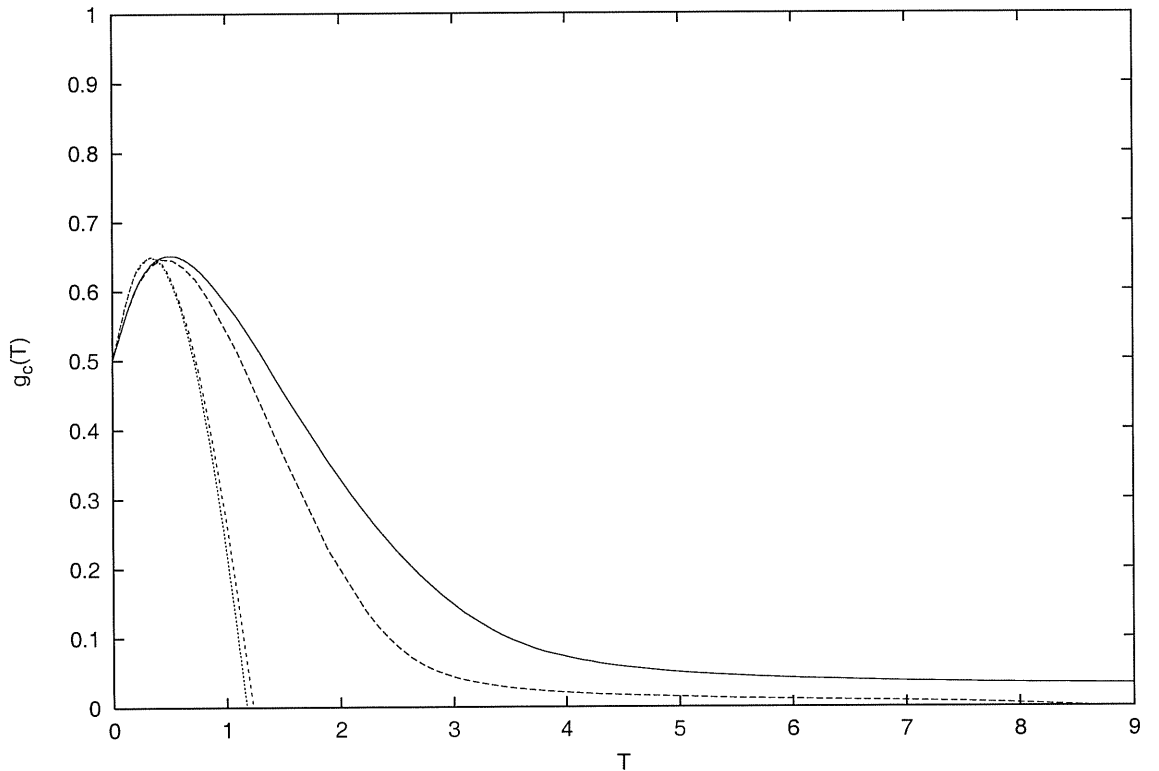


Figure 2.4: Phase diagrams in  $D=3$  and  $4$  dimensions for models with and without crossing ( $\epsilon = 1$  in the figure). It can be seen that the difference between the four dimensional models is negligible. In  $D=3$  (model w.o.c.),  $T_c \simeq 8.49\epsilon$ . Solid line: model  $D=3$  w.c.; Long-dashed line: model  $D=3$  w.o.c.; Dashed line: model  $D=4$  w.c.; Dotted line: model  $D=4$  w.o.c.

## 2.3 Scaling laws for thermal melting and unzipping

Let us focus for concreteness on the result for the model in  $D = 2$  w.o.c., in order to analyze the scaling laws of the system. The phase diagram, as found also in [19], explicitly reads (see Fig. 2.3):

$$g_c(T, \epsilon) = \frac{T}{2} \cosh^{-1} \left[ \frac{1}{2} \frac{1}{\sqrt{1 - \exp(-\beta\epsilon)} - (1 - \exp(-\beta\epsilon))} - 1 \right]. \quad (2.22)$$

The model exhibits a first-order “unzipping” transition if we move at a fixed value of  $T < \epsilon / \log(4/3)$ , as shown in (2.18). As is shown in Fig. 2.3,  $\lim_{T \rightarrow 0} g_c(T, \epsilon) = \frac{\epsilon}{2}$  and  $g_c(T, \epsilon)$  attains its maximum at  $T = T_M \simeq 0.9\epsilon$  where  $g_c(T_M) \simeq 0.68\epsilon > \frac{\epsilon}{2}$ . The transition is second order at  $g = 0$ ; as in [17] we find that close to the melting

point

$$\langle n \rangle \equiv \frac{\partial}{\partial(\beta\epsilon)} \log Z_N \sim \tau^{-1} f(\tau N^{\phi_t}), \quad (2.23)$$

where  $\langle n \rangle$  is the mean number of native contacts, with  $\tau = (T - T_c)/T_c$ , and  $\phi_t = 1/2$  for the crossover exponent at thermal melting in the  $D = 2$ -case. The scaling function  $f(x)$  behaves as usual

$$\begin{cases} f(x) \rightarrow 0 & \text{as } x \rightarrow +\infty, \\ f(x) \sim x & \text{for } |x| \ll 1, \\ f(x) \sim -|x|^{1/\phi_t} & \text{for } x \rightarrow -\infty, \end{cases} \quad (2.24)$$

such that  $\langle n \rangle \sim N^{\phi_t}$  at the transition.

The exact expressions for  $\langle n \rangle$  and for  $\langle x_g \rangle$  can be found by inverse Laplace transforms of the quantities  $Z(z, \beta g, \beta\epsilon)$ ,  $\frac{\partial}{\partial(\beta\epsilon)} Z(z, \beta g, \beta\epsilon)$ ,  $\frac{\partial}{\partial(\beta g)} Z(z, \beta g, \beta\epsilon)$ . However, if we only need scaling relations in the thermodynamic limit, we can use the discrete Tauberian theorem (see [27] and Appendix B), which relates the critical behaviour of a series to the asymptotic behaviour of its coefficients. By using this method, we find, in the vicinity of the unzipping mechanical transition:

$$\langle n \rangle \sim \frac{h_1 [N^{\phi_1} (A_1 \gamma + A_2 \tau)]}{A_1 \gamma + A_2 \tau}, \quad (2.25)$$

$$\langle x_g \rangle \sim -\frac{h_2 [-N^{\phi_2} (A_1 \gamma + A_2 \tau)]}{A_1 \gamma + A_2 \tau}, \quad (2.26)$$

where  $A_1(T_c, g_c)$ ,  $A_2(T_c, g_c)$  are determined in such a way that  $\frac{dg_c(T, \epsilon)}{dT} = -\frac{A_2 g_c}{A_1 T_c}$ ,  $\gamma = (g - g_c)/g_c$ ,  $\tau = (T - T_c)/T_c$ , and  $\phi_1 = \phi_2 = 1$  for the two crossover exponents, consistently with the fact that the unzipping transition is first-order (the explicit form of  $A_1(T_c, g_c)$ ,  $A_2(T_c, g_c)$  is worked out in Appendix B). Note that  $A_1 > 0$ , whereas  $A_2$  is negative in the re-entrant part of the transition curve. The scaling functions  $h_{1,2}(x)$  behave in a similar way to  $f(x)$ , and thus we get that  $\langle n \rangle \sim N^{\phi_1}$  and  $\langle x_g \rangle \sim N^{\phi_2}$  are both extensive at the transition point. The physical interpretation is that a macroscopic portion of the chain is still in the double stranded state, but the rest of the chain is unzipped; these are just the two phases coexisting at the first order transition. This result will be used later to justify the use of Y-shaped configurations.

It is instructive to compare this phase diagram with that of the same  $D = 2$ -case when we allow crossing. The transition line (plotted in Fig. 2.3), obtained from  $z_2 = z_3$  (see Appendix A) is:

$$g_c(T, \epsilon) = T \tanh^{-1} [1 - \exp(-\beta\epsilon)] = \frac{\epsilon}{2} + \frac{1}{2\beta} \log [2 - \exp(-\beta\epsilon)], \quad (2.27)$$

Both the models behave similarly near  $T = 0$ , namely they yield a transition to the *cold denaturated* state described in [19]. When we move at a constant

force  $g$  such that  $g_c(0) < g < \text{Max}_T [g_c(T, \epsilon)]$ , at low enough temperatures the molecule is unzipped, and it zips by increasing  $T$ . This feature of the phase diagrams seems at first sight paradoxical, since one would expect the critical force to decrease monotonically as the temperature is increased. The physical explanation of this result will be given below. In the model w.c., moreover,  $g_c$  always increases, approaching  $\epsilon$  as  $T \rightarrow \infty$ , and the two strands remain zipped for every temperature when  $g < \epsilon$ . In the model w.o.c., instead, the two strands unzip again by further increasing the temperature. As regards scaling, on the other hand, the two models are identical so that we can say that the effect of forbidding crossing is irrelevant in the renormalization group sense, but has a dramatic effect on the form of the critical line.

Let us now discuss the results that we have obtained in higher dimensions (the critical lines for three and four dimensional models are shown in Fig. 2.4). Common to all models is the cold unzipping transition found in the  $D = 2$  case. Moreover, the scaling laws at the unzipping transition do not depend on dimensionality in the crossover exponents. On the other hand, the thermal melting has dimensionality-dependent critical exponents (see [11, 17]). Consequently, the behaviour of the critical line  $g_c(T, \epsilon)$  near the end point  $T = T_c$  also depends on  $D$ . As  $T \rightarrow T_c^-$  the result that we find in generic dimension is:

$$\frac{g_c(T, \epsilon)}{T} \sim \begin{cases} \epsilon/T - \epsilon/T_c(D) & D = 2, 4 \\ \exp \left[ -\frac{a}{(\epsilon/T - \epsilon/T_c(D))^2} \right] & D = 3 \\ [\epsilon/T - \epsilon/T_c(D)]^{1/2} [\log(\epsilon/T - \epsilon/T_c(D))]^{-1/2} & D = 5 \\ [\epsilon/T - \epsilon/T_c(D)]^{1/2} & D > 5 \end{cases}, \quad (2.28)$$

where  $a$  is a constant, and  $T_c = \infty$  in  $D = 2, 3$  for models w.c.



# Chapter 3

## Discussion and role of disorder

### 3.1 Physical explanation of reentrance

To obtain some physical insight into our analytical results, we compute the free energy of a Y-shaped molecule of DNA with a force pulling at the extremities (see Fig. 3.1). In this way, we neglect all the configurations with *bubbles*. Such an approximation is valid at low  $T$  and yields the exact result in the limit  $T \rightarrow 0$ . Since the configurations of the unzipped part are weighted by  $\exp(\beta \vec{g} \cdot \vec{x})$ , in this limit only the completely stretched configuration will contribute to the free energy for the unzipped part of the Y. In the limit of low temperature, the free energy is then:

$$F(m, N) \sim -(N - m)(\epsilon + T \log \mu) - 2gm, \quad (3.1)$$

where  $m$  is the number of monomers in the unzipped part,  $N$  is the total number of bases and  $\mu$  is the connective constant of a single walk. Note that eq. (3.1) (with  $\vec{g} = g(1, 0, \dots, 0)$ ) is valid also in the case of two self-avoiding walks (SAWs), since the connective constant of SAWs constrained to avoid the fully stretched unzipped part, is the same as for ordinary SAWs [27, 28] (for a more detailed discussion on SAWs we refer to Appendix C). By minimizing with respect to  $m$ , we find the critical force  $g_c(T, \epsilon)$  such that (3.1) is minimum for  $m = 0$  when  $g < g_c(T, \epsilon)$ , and for  $m = N$  when  $g > g_c(T, \epsilon)$ :

$$g_c(T, \epsilon) \stackrel{T \rightarrow 0}{\sim} \frac{1}{2}(\epsilon + T \log \mu). \quad (3.2)$$

This is indeed what we have found in our calculations, and confirms that reentrance is a robust feature of lattice models, not depending on dimensionality or self-avoidance. In other words, (3.2) means that, at low  $T$ , it is more difficult to open a dsDNA helix as  $T$  is increased, because the energy gain obtained through the unzipping is more than compensated by the entropy loss, since there is only one possible completely stretched configuration versus  $\mu^{N-m}$  possibilities for the

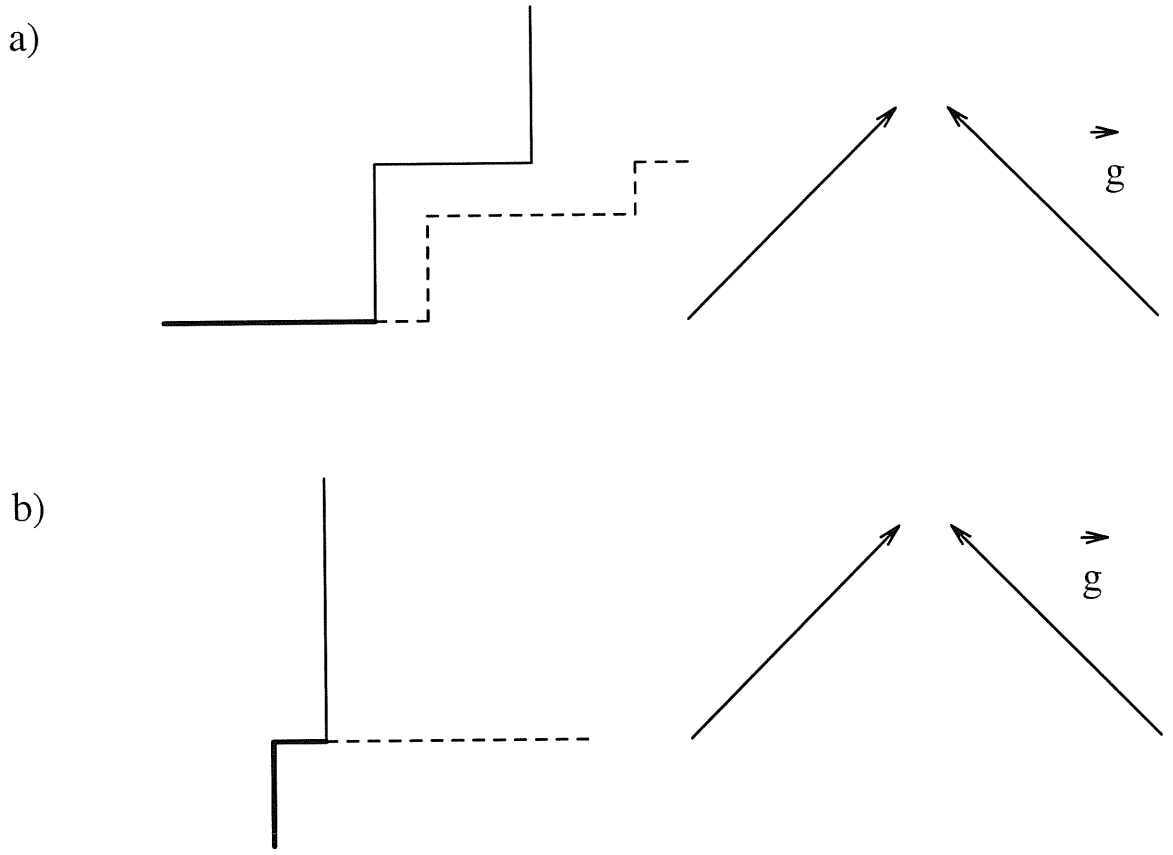


Figure 3.1: a) We show here an example of Y-shaped configuration for the DNA molecule: in this approximation bubbles are neglected. b) A completely stretched configuration for directed walks, which is dominant for  $T \rightarrow 0$ .

double stranded portion of the chain (we will see in the following section that the entropy loss in the continuum space exhibits a power law correction).

### 3.1.1 Re-entrance in the continuum

We now prove that also discrete chains in the continuum space undergo a cold denaturation for  $T \rightarrow 0$ . Let us consider two  $N$ -monomer chains in  $\mathbf{R}^d$  with no constraint on the directedness, with a force  $\vec{g}$  pulling at the extremities and with constant unitary distance between consecutive monomers. As  $\beta \rightarrow \infty$ , bubbles can be neglected and only Y-shaped configurations contribute to the partition sum. The partition function for a Y-configuration then reads:

$$Z_N(\beta\epsilon, \beta g) = \sum_{m=0}^{N-1} Z_{N-m}^z(\beta\epsilon) Z_{2m}^u(\beta g), \quad (3.3)$$

where  $Z_{N-m}^z(\beta\epsilon) = \int_{\mathbf{R}^d} d^d \vec{x}_1 \dots d^d \vec{x}_{N-m-1} \delta(|\vec{x}_1| - 1) \dots \delta(|\vec{x}_{N-m-1}| - 1) \exp((N-m)\beta\epsilon)$  is the partition function of the zipped portion of the strands, and  $Z_{2m}^u(\beta g) =$

$\int_{\mathbf{R}^d} d^d \vec{y}_1 \dots d^d \vec{y}_{2m} \delta(|\vec{y}_1| - 1) \dots \delta(|\vec{y}_{2m}| - 1) \exp(\beta \vec{g} \cdot \sum_{i=1}^{2m} \vec{y}_i)$  is that of the unzipped end. Chain discreteness is crucial in ensuring the validity of eq.(3.3) for  $\beta \rightarrow \infty$ , since when one bubble has formed, the length of the stretched portion of the chain can be at most  $\sum_{i=1}^{2m-1} \vec{y}_i = 2(m-1) \frac{\vec{g}}{|\vec{g}|}$ . The integrals in (3.3) can be performed in any dimensions, and the evaluation of  $Z_{2m}^u(\beta g)$  involves the modified Bessel function of the first kind of order  $\frac{d-2}{2}$ . In particular,  $Z_{2m}^u(\beta g) \stackrel{\beta \rightarrow \infty}{\sim} \left(\frac{2\pi}{\beta g}\right)^{m(d-1)} \exp(2m\beta g)$ , and there is a power-law correction with respect to the lattice result. We find that  $Z_N(\beta\epsilon, \beta g) \stackrel{\beta \rightarrow \infty}{\sim} \sum_{m=0}^{N-1} \Omega_d^{N-m-1} \exp((N-m)\epsilon) Z_{2m}^u(\beta g)$ , where  $\Omega_d$  is the surface area of the unit sphere in  $d$  dimensions. From this, the critical force is easily found:

$$g_c(T, \epsilon) \stackrel{T \rightarrow 0}{\sim} \frac{\epsilon}{2} - \frac{d-1}{2} T \log \frac{T}{\epsilon} - \frac{T}{2} \log \left[ 2^{2d-3} \pi^{\frac{d}{2}-1} \Gamma\left(\frac{d}{2}\right) \right], \quad (3.4)$$

where  $\Gamma$  is the Euler Gamma function. The critical line  $g_c(T, \epsilon)$  increases at low  $T$  and re-entrance is present also in the continuum, even enhanced with respect to the lattice case. The leading term  $T \log T$  in equation (3.4) (due to the power-law correction in  $Z_{2m}^u(\beta g)$ ) is indeed not present in the  $d = 1$  case, when one correctly recovers equation (3.2) for a lattice random walk with  $\mu = 2$ .

We finally wish to discuss the relation of the present treatment with previous work ([13, 14, 15, 16]) done on the unzipping of homo-DNA in the limit of a continuum chain. The ds molecule with the pulling force had been described by means of an effective hamiltonian, which, apart from an irrelevant center of mass term, reads:

$$H = \int_0^N dn \left( \frac{Td}{b^2} \left( \frac{d\vec{r}}{dn} \right)^2 + V(\vec{r}(n)) - \vec{g} \cdot \frac{d\vec{r}}{dn} \right), \quad (3.5)$$

where  $\vec{r}$  is the relative separation between the strands,  $b$  is the effective Kuhn length of single-stranded DNA and  $V$  is a realistic short range attractive potential, namely a delta function or a potential well (the two cases should be equivalent according to the standard quantum theory as long as  $\int d\vec{r} V(\vec{r})$  is the same). The system described by (3.5) is equivalently represented by a quantum system with a non-hermitian hamiltonian. One finds [14, 15] that there is a first order unzipping transition when the force reaches the critical value:

$$g_c(T, \epsilon) = \sqrt{-\frac{4\epsilon_0(T)Td}{b^2}}, \quad (3.6)$$

where  $\epsilon_0(T)$  is the ground state energy of the quantum hamiltonian obtained from (3.5) when  $\vec{g} = \vec{0}$ . In the quantum mechanical system, for  $g < g_c(T, \epsilon)$  the ground state is a bound state, while for  $g > g_c(T, \epsilon)$  the spectrum of (3.5) is continuous.

Let us focus on the  $d = 1$  case, corresponding to our two-dimensional models without crossing constraints. It is well known that at  $g = 0$  both a symmetric square well and a delta function always have at least one bound state, meaning that DNA remains ds at all  $T$ . In Fig. 3.2 we show the critical force for the two potentials (obtained simply from (3.6)). It can be seen from the figure that at high  $T$  both the potentials saturate towards the same limit, as in our calculation for the model w.c. This is expected since the delta potential case can be recovered from our equation (2.2) in the limit of both continuum space and time (see Fig. 3.2). At low  $T$ , however,  $g_c(T, \epsilon) \sim \left(\frac{T}{\Delta}\right)^{1/2}$  for a square well of width  $\Delta$ , and is constant for the delta potential, so that re-entrance is present only in the former case, but with a behaviour different from that found in the lattice. This already signals that the low temperature behaviour of the solution obtained through the quantum mapping is rather unphysical. This is due to the fact that in the limit of continuum chain the chain constraint is modelled in a ‘soft’ way, by using a harmonic potential between consecutive beads along the chain. This interaction is usually assumed to be entropic (as in eq. (3.5)) and thus effectively vanishes at  $T = 0$ . In that limit eq. (3.5) is describing a set of  $N$  ‘free’ particles moving in an effective potential determined by  $V$  and  $g$ . Consequently, the quantum mapping is expected to describe the system well except in the low temperature limit. Indeed, the re-entrance found for the square well potential is due to the unboundedness of the potential felt by the “free particles” as soon as  $g \neq 0$ . As  $T \rightarrow 0$ , fluctuations are not important, and the particles stay close to the minimum of such potential, which is  $-\infty$  for a force whatever small but finite. This is enough to unzip the strands.

As Zhou has observed in [16], one can improve this model by placing a hard core either inside the potential well or at a finite distance from the delta (this is analogous in spirit to introducing the crossing constraint in our lattice models). Once again, one finds that for  $T$  near  $T_c$ , which is now finite, both potentials predict  $g_c(T, \epsilon) \stackrel{T \rightarrow T_c}{\sim} (T_c - T)$  as in the lattice model (see eq.(2.28)), whereas for low  $T$  the delta shows no re-entrance and the potential well displays a  $T^{1/2}$  behaviour. We note once again that the low temperature behavior of such ‘quantum’ models is an artifact, when considered as polymer chain models, due to the ‘soft’ enforcing of the chain constraint. In this respect, discrete chain models in the continuum space are more realistic, either with constant bond length or with harmonic springs between consecutive beads. In the first case we have indeed proved (see eq. (3.4)) the existence of cold mechanical denaturation. In the second case it has been shown to occur by means of numerical simulations [19].

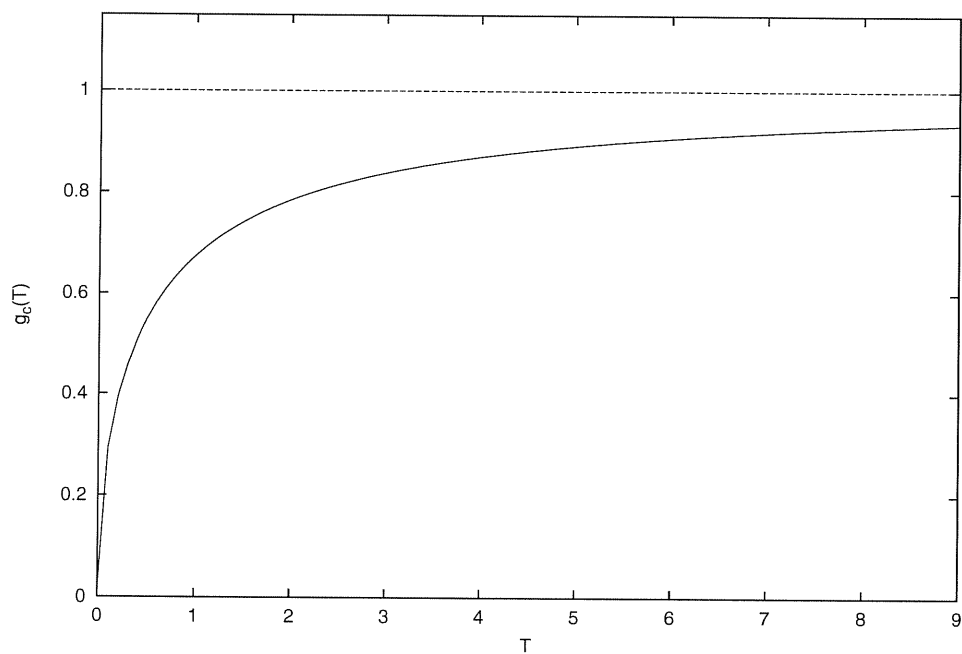


Figure 3.2: Critical force in  $d = 1$ , found with the quantum mapping. The solid line is the result with a symmetric square well, the dashed one with a delta function. The parameters are chosen so that the integral of the potentials over space is the same.

## 3.2 Quenched disorder and analytical treatment

In real DNAs, the contact energy  $\epsilon_i$  ( $i$  is the index of the monomer along the chain) depends on  $i$  and is not constant, because the CG and AT base pairs have different pairing energies. We take a bimodal distribution of disorder, meaning that for every monomer index one has  $\epsilon_i = \epsilon \pm \Delta$  with equal probability.

The partition function for two chains having their  $N$ -th monomers at a relative lattice distance  $x$  can be written in terms of a transfer matrix  $W$  as  $Z_N(x) = \langle x | W_1 \dots W_N | x \rangle$ . The ket (bra)  $|x\rangle$  ( $\langle x|$ ) denotes a column (row) vector whose only non-zero entry is the  $x$ -th, which is equal to 1, and  $W$  is defined in terms of its matrix elements  $\langle x' | W_k | x \rangle \equiv ((\exp(\beta\epsilon_k) - 1)\delta_{x,0} + 1)(2\delta_{x,x'} + \delta_{x,x'+1} + \delta_{x,x'-1})$ . This approach renders the problem unsolvable analytically and one has to resort to transfer matrix techniques.

If on the other hand we restrict the allowed configurations to the Y configurations discussed above in which bubbles, i.e. rejoining of two strands once they have cleaved from one another, are forbidden, the model remains exactly treatable even in the presence of quenched disorder. We first solve the zero force case (in this case restricting oneself to Y configurations means considering the problem on the Cayley tree).

### 3.2.1 Integer moments of the canonical partition function: the Cayley tree (thermal melting)

Let us consider two strands of DNA on the Cayley tree with coordination number  $\mu \geq 2$ . This system is easier to solve than the full system of directed walks, as when the two strands divide, they can never join again because of the tree structure of this lattice. We consider  $g = 0$  in this calculation. The result will give us a procedure to use in the generic case. Our aim is to find an expression for the quantity:

$$Z_n(z, \beta, \epsilon, \Delta) = \sum_{N=0}^{\infty} z^N \overline{Z_N^n} \quad (3.7)$$

where  $\Delta$  is defined before,  $Z_n$  is the partition function of a  $n$ -times replicated DNA,  $z$  is the step fugacity,  $Z_N$  is the usual  $N$ -monomer partition function and  $\overline{(\cdot)}$  indicates disorder averaging. To understand physically how this system behaves, it is useful to partition the  $n$ -times replicated double stranded DNA in  $n+1$  sections: in the first all the  $n$  replicas are in the ds phase, in the second one is unzipped, while the other remain zipped, and so on: in the  $i$ -th section there are  $n-i+1$  replicas ds and  $i-1$  in the denatured state. Now, each of these sections will bring in a singularity in (3.7), and therefore the  $n$  times replicated system will be characterized by the smallest of such singularities. Not surprisingly, the two competing singularities are the one pertaining to all  $n$  replicas zipped and

that coming from the section with all strands opened. This competition generates a phase transition at a temperature  $T_{c,n}$  which depends on the replica index  $n$ .

We recall that the pure version of this model has a critical temperature  $T_{c,pure} = \frac{\epsilon}{\log \mu}$ .

In the following calculation, we take for simplicity:

$$Z_N = \sum_{k=0}^N \mu^k \exp\left(\sum_{i=1}^k \beta \epsilon_i\right) \mu(\mu-1) \mu^{2(N-k)} \quad (3.8)$$

so that we can write:

$$Z_{N+1} = \mu(\mu-1) \mu^{N+1} \exp\left(\beta \sum_{i=1}^{N+1} \epsilon_i\right) + \mu^2 Z_N \quad (3.9)$$

(Notice that  $Z_0 = \mu(\mu-1)$ . We also in fact calculate  $Z_N/(\mu(\mu-1))$  to avoid unnecessary constants.

We calculate the annealed average of the partition function, then the second moment and finally we give the expression for generic  $n$ . We will then discuss the applicability of the replica trick, in order to find an expression for the free energy and the mean number of contacts in the thermodynamic limit.

- $\overline{Z_N}$ : the annealed model

$$Z_1(z, \beta, \epsilon, \Delta) = \sum_{N=0}^{\infty} z^N \overline{Z_N} \quad (3.10)$$

The recursion relation is:

$$\overline{Z_{N+1}} = \mu^{N+1} (\overline{\exp(\beta \epsilon)})^{N+1} + \mu^2 \overline{Z_N} \quad (3.11)$$

This implies

$$Z_1(z) = \frac{1}{1 - \mu (\overline{\exp(\beta \epsilon)}) z} + \mu^2 z Z_1(z) \quad (3.12)$$

and so

$$Z_1(z) = \frac{1}{(1 - \mu^2 z)(1 - \mu (\overline{\exp(\beta \epsilon)}) z)} \quad (3.13)$$

The two resulting singularities for the annealed model are:

$$z_1 = \frac{1}{\mu} \exp\left(-\beta\epsilon - \beta^2\Delta^2/2\right) \quad (3.14)$$

$$z_2 = \frac{1}{\mu^2}$$

$T_{c,1}$  (annealed) is given by:

$$T_{c,1} = \frac{\epsilon + \sqrt{\epsilon^2 + 2\Delta^2 \log \mu}}{2 \log \mu} \quad (3.15)$$

In the low temperature (zipped) phase, the annealed free energy is given by:

$$f = T \log z_1 = -T \log \mu - \epsilon - \frac{\Delta^2}{2T} \quad (3.16)$$

so that the entropy per particle  $s = -\frac{\partial f}{\partial T}$  gets negative for  $T < T_n \equiv \frac{\Delta}{\sqrt{2 \log \mu}}$ . Below  $T_n$ , the annealed approximation thus cannot work.

- n=2:  $\overline{Z_N^2}$  and  $\frac{\overline{Z_N^2}}{Z_N^2}$

We need the following recursion relations, easily found from (3.9):

$$\overline{Z_{N+1}^2} = \mu^{2(N+1)} \left( \overline{\exp(2\beta\epsilon)} \right)^{N+1} + \mu^4 \overline{Z_N^2} + \quad (3.17)$$

$$2\mu^2 \mu^{N+1} \overline{\left( \exp\left(\sum_{i=1}^{N+1} \beta\epsilon_i\right) Z_N \right)}$$

$$\overline{Z_{N+1} \exp\left(\sum_{i=1}^{N+1} \beta\epsilon\right)} = \mu^{N+1} \left( \overline{\exp(2\beta\epsilon)} \right)^{N+1} + \mu^2 \overline{Z_N \exp\left(\sum_{i=1}^{N+1} \beta\epsilon_i\right)}$$

As before, we need  $Z_2(z, \beta, \epsilon, \Delta)$ . We have:

$$Z_2(z) = \frac{1}{1 - \mu^2 z \exp(2\beta\epsilon + 2\beta^2\Delta^2)} + \mu^4 z Z_2(z) + \quad (3.18)$$

$$2\mu^3 z Z_{1,1}(z) \exp(\beta\epsilon + \beta^2\Delta^2/2)$$

where we have defined:



$$Z_{1,1}(z) = \sum_{N=0}^{\infty} z^N \overline{\mu^N \exp\left(\sum_{i=1}^N \beta \epsilon_i\right) Z_N} \quad (3.19)$$

We get for (3.19):

$$Z_{1,1}(z) = \frac{1}{(1 - \mu^3 z \exp(\beta\epsilon + \beta^2 \Delta^2/2))(1 - z \mu^2 \exp(2\beta\epsilon + 2\beta^2 \Delta^2))} \quad (3.20)$$

and therefore  $Z_2(z)$  reads:

$$Z_2(z) = \frac{1 + z \mu^3 \overline{\exp(\beta\epsilon)}}{(1 - \mu^4 z)(1 - z \mu^3 \overline{\exp(\beta\epsilon)})(1 - z \mu^2 \overline{\exp(2\beta\epsilon)})} \quad (3.21)$$

The three singularities in  $Z_2(z)$  are:

$$\begin{aligned} z_1 &= \frac{1}{\mu^2} \exp(-2\beta\epsilon - 2\beta^2 \Delta^2) \\ z_2 &= \frac{1}{\mu^3} \exp(-\beta\epsilon - \beta^2 \Delta^2/2) \\ z_3 &= \frac{1}{\mu^4} \end{aligned} \quad (3.22)$$

It can be seen that the smallest singularity is to be found between  $z_1$  and  $z_3$ , and that  $z_2$  never comes in in this evaluation. The critical temperature of the second moment is:

$$T_{c,2} = \frac{\epsilon + \sqrt{\epsilon^2 + 4 \log \mu \Delta^2}}{2 \log \mu} \quad (3.23)$$

From eqs.(3.21,3.12) we can find the asymptotical behaviours of  $\overline{Z_N^2}$  and  $\overline{Z_N}^2$  by using the discrete Tauberian theorem. For example, we write down the behaviour for large  $N$  of  $\frac{\overline{Z_N^2}}{\overline{Z_N}^2}$  in the unzipped phase of the second moment, where we know it to be exactly 1 when  $\Delta = 0$ :

$$\frac{\overline{Z_N^2}}{\overline{Z_N}^2} \underset{N \rightarrow \infty}{\sim} \frac{(1 + \frac{\exp(+\beta\epsilon + \beta^2 \Delta^2/2)}{\mu})(1 - \frac{\exp(+\beta\epsilon + \beta^2 \Delta^2/2)}{\mu})}{1 - \frac{\exp(+2\beta\epsilon + 2\beta^2 \Delta^2)}{\mu^2}} \quad (3.24)$$

- $n$ -th moment

For the  $n$ -th moment, we can proceed in a similar way. The two competing singularities of  $Z_n(z)$  are:

$$z_1 = \frac{1}{\mu^n} \exp\left(-n\beta\epsilon - n^2\beta^2\Delta^2/2\right) \quad (3.25)$$

$$z_{n+1} = \frac{1}{\mu^{2n}}$$

The critical temperature of the  $n$ -th moment is:

$$T_{c,n} = \frac{\epsilon + \sqrt{\epsilon^2 + 2n\Delta^2 \log \mu}}{2 \log \mu} \quad (3.26)$$

The behaviour of  $T_{c,n}$  for high  $n$  is that found in the REM:  $T_{c,n} \stackrel{n \rightarrow \infty}{\sim} n^{1/2}$   
For generic  $n$ , we can thus find, through the Tauberian theorem, that:

$$\overline{Z_N^n} \stackrel{N \rightarrow \infty}{\sim} c_n z_{1(n+1)}^{-N} \quad (3.27)$$

where  $c_n$  does not depend on  $N$ . Indeed we can find a recursion relation for  $c_n$ , which, in the unzipped phase of the  $n$ -th moment, reads:

$$c_n = \frac{1}{1 - \alpha^n} \left( 1 + \sum_{k=1}^{n-1} \binom{n}{k} c_{n-k} \overline{\alpha^{n-k}} \right) \quad (3.28)$$

where  $\alpha = \frac{\exp(\beta\epsilon)}{\mu}$ .

If, in eq.(3.27) we neglect  $c_n$  in the limit  $N \rightarrow \infty$ , we can formally apply the replica trick to obtain the free energy of the quenched model, which in this way is found to be equal to that of the corresponding pure model if we substitute  $\epsilon$  of the pure model with the average of the corresponding disordered model.

### 3.2.2 Y configurations

The strategy used to calculate the average of the moments of the partition function for the Cayley tree can be generalized to the case when we consider only Y configurations for the two directed walks, this time at whatever force  $g$ . In a similar way to the one shown before, a bit more cumbersome because we have to keep track of the non-crossing constraint between the chains, we can find that in the Y or fork model the critical force remains the same as that of the pure model with

$\epsilon$  being the average of the dicotomic attractive energy. Consequently, the transition remains first order in any point of the phase diagram. These conclusions are supported by a numerical calculation. As a result, we can find that in this Y or fork model the critical force remains the same as that of the pure model with  $\epsilon$  being the average of the dicotomic attractive energy. Consequently, the transition remains first order in any point of the phase diagram. These conclusions are supported by a numerical calculation as well.

### 3.2.3 Full model with bubbles

Also in this case it is possible to set up a replica approach: the second moment of the partition function corresponds for example to a system of two pairs of directed walks. Each pair is subject to a pulling force and is kept together by a contact energy, but when the two pairs are simultaneously, *i.e.* at the same site index, joined, the contact energy becomes higher than  $2\epsilon$ , as is always the case when studying moments of disordered systems. It is not possible to solve this model exactly on the lattice. We can however resort to numerical methods which make use of the transfer matrix to study directly the quenched model. At any points with  $g \neq 0$  we find numerically that the transition remains first order, even if in this case the critical line in the phase diagram is not the same as in the pure phase diagram. We can move in the phase diagram either at constant  $T$  or at a constant  $g$ . At  $g \neq 0$  the order parameters, *e.g.* the average end-to-end distance per monomer  $\frac{\langle x \rangle}{N}$ , obey simple scaling laws, different from the pure case as found in[14]:

$$\frac{x}{N} = F_g(T - T_c)N^\phi \quad (3.29)$$

if we move for example at constant  $g$ , where  $F_g$  is a suitable scaling function. One finds that  $\phi$  is  $\frac{1}{2}$ , as opposed to the pure case where it is known to be 1. The plots of the order parameter at fixed  $g$  for various  $T$  which lead to the collapse validating eq.(3.29) are shown in Fig.3.3.

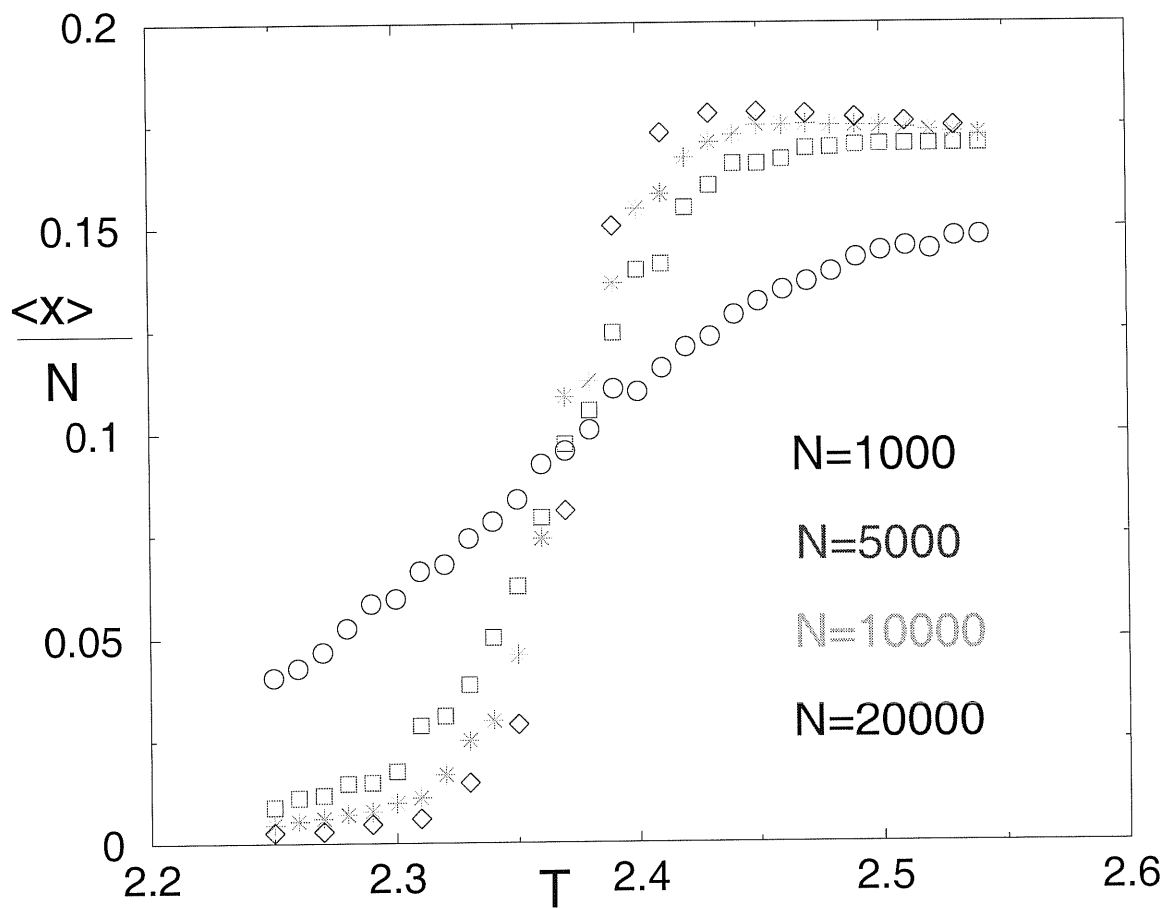


Figure 3.3: Finite size behaviour of the order parameter, quenched averaged case, for different system sizes.

# Conclusions

In this thesis we have considered a simplified model for the DNA mechanical unzipping. Given the high current interest in experiments of thermal and mechanical denaturation, this topic can be relevant to experimentalists as well.

Although the model neglects many details, it contains the important features of the system. In particular, we have mimicked the two strands of the DNA being pulled by means of two directed walks. The resulting model is analytically solvable by means of the generating function technique.

As a result, we have found that for a homogeneous DNA the unzipping induced by an external force is a phase transition from a zipped phase, where only a finite number of base pairs are not in the native bound state, to an unzipped phase, where the number of broken base pairs is extensive. This transition has been proved to be first order, in contrast to the usual melting transition which within the model treated is second order. More intriguingly, the exact solution makes it possible to find exactly the phase diagram of the system, which reveals the presence of a cold temperature mechanical unzipping which could in principle be observed experimentally. This reentrant region seems at first striking but has a simple physical meaning, being due to the entropy loss caused in fully stretching the two strands in the direction of the pulling force, which is not compensated for by the free energy gain obtained in the opening.

If we put quenched disorder in the system, in general the position of the transition line changes, as expected as it is not a universal quantity. The unzipping transition however remains first order, and the reentrant region is still present: this has been proved by means of numerical and analytical calculations. On the other hand, the scaling relations valid for the unzipping change in the quenched disorder case, in agreement with the Harris criterium which predicts that the disorder is relevant for this transition.

Other developments, such as the study of the non-equilibrium effects and of the role of helicity in this picture could help in trying to render these results more directly related to experiments. We wish to pursue these in the future.



# Appendix A: Computation of $B(z)$ in models of DSAWs without crossing and of the phase diagram for models with crossing

We here first work out the details to obtain an explicit expression for  $B(z)$ , defined in eq.(2.6). Starting from equations (2.7), we can perform a Fourier and a discrete Laplace transform to obtain:

$$\tilde{c}(z, \vec{q}) \equiv \sum_{N=0}^{\infty} z^N \sum_{\vec{x}} \exp(i\vec{q} \cdot \vec{x}) c_N(\vec{x}) = \frac{1}{1 - zf(\vec{q})} (\tilde{c}_0(\vec{q}) - zf_0(\vec{q})c(z, \vec{0})), \quad (3.30)$$

where we have made the following definitions:

$$f_0(\vec{q}) \equiv 2 \sum_{i < j; i, j=1}^D \cos(q_i - q_j); \quad f(\vec{q}) \equiv D + f_0(\vec{q}) \quad (3.31)$$

$$c(z, \vec{x}) \equiv \sum_{N=0}^{\infty} z^N c_N(\vec{x}); \quad \tilde{c}_0(\vec{q}) \equiv \sum_{\vec{x}} \exp(i\vec{q} \cdot \vec{x}) c_0(\vec{x}) \quad (3.32)$$

Using  $c(z, \vec{0}) = \int_{[-\pi, \pi]^D} \frac{d^D \vec{q}}{(2\pi)^D} \tilde{c}(z, \vec{q})$ , equation (3.30) is easily solved since, by permuting variables inside the resulting integrals, it is possible to see that they do not depend on the particular step which we forbid. We obtain:

$$\tilde{c}(z, \vec{q}) = \frac{1}{1 - zf(\vec{q})} \frac{\tilde{c}_0(\vec{q}) - zW_1(z)[A(D)f_0(\vec{q}) - \tilde{c}_0(\vec{q})]}{1 + zW_1(z)}, \quad (3.33)$$

$$c(z, \vec{0}) = A(D) \frac{W_1(z)}{1 + zW_1(z)}, \quad (3.34)$$

where  $A(D) = 1 - \frac{1}{D(D-1)}$  is the reduction factor due to the crossing constraint. As a result, using  $B(z) = z^2 \sum_{i \neq j; i, j=1}^D c(z, \vec{e}_i - \vec{e}_j)$ , we obtain

$$B(z) = A(D) \frac{z^2 W_2(z)}{1 + z_2 W_1(z)}, \quad (3.35)$$

where

$$W_1(z) = \int_{[-\pi, \pi]^D} \frac{d^D \vec{q}}{(2\pi)^D} \frac{f_0(\vec{q})}{1 - zf(\vec{q})}, \quad W_2(z) = \int_{[-\pi, \pi]^D} \frac{d^D \vec{q}}{(2\pi)^D} \frac{f_0^2(\vec{q})}{1 - zf(\vec{q})}. \quad (3.36)$$

The singularity  $z_1 = 1/D^2$ , leading to the usual random walk behavior, comes in when evaluating the above integrals; their denominator becomes negative as  $\vec{q} \rightarrow 0$  for  $z > z_1$ .

We give here also a brief outline of the calculation necessary to find out the phase diagram for the models of Section II of DSAWs when crossing is allowed. These models are simpler to solve than the corresponding models w.o.c., since it is not necessary to partition the molecule of DNA as done in the text in eq.(2.5), which allows us to write a more explicit formula for the grand partition function. Indeed we can simply start from the recursion relations (2.2) and, after Fourier and Laplace transform, find the expression:

$$\tilde{p}(z, \vec{q}, \beta\epsilon) = \frac{1}{1 - zf(\vec{q})} \frac{1}{1 - (1 - \exp(-\beta\epsilon)) W_0(z)}, \quad (3.37)$$

for the quantity  $\tilde{p}(z, \vec{q}, \beta\epsilon)$ , defined as:

$$\tilde{p}(z, \vec{q}, \beta\epsilon) = \sum_{N=0}^{\infty} z^N \sum_{\vec{x}} \exp(i\vec{q} \cdot \vec{x}) p_N(\vec{x}, \beta\epsilon) \quad (3.38)$$

with  $p_N(\vec{x}, \beta\epsilon)$  defined as in eq.(2.1) and

$$W_0 = \int_{[-\pi, \pi]^D} \frac{d^D \vec{q}}{(2\pi)^D} \frac{1}{1 - zf(\vec{q})}. \quad (3.39)$$

The final form of the grand partition function, for DSAWs w.c. in the presence of a pulling force, is:

$$Z(z, \beta\epsilon, \beta g) = \frac{1}{1 - z/z_3(\beta g)} \frac{1}{1 - (1 - \exp(-\beta\epsilon)) W_0(z)}, \quad (3.40)$$

with  $z_3(\beta g)$  defined as in eq.(2.15). Alternatively, one can proceed exactly as in Section II, but with initial conditions

$$c_0(\vec{x}) = \sum_{i \neq j; i, j=1}^D \delta_{\vec{x}, \vec{e}_i - \vec{e}_j}. \quad (3.41)$$

The only resulting difference would be to have  $A(D) = 1$  in eqs.(3.33) and (3.35); after a bit of algebra one can convince himself that the partition function found by recollecting the different factors in eq.(2.5) is the same as that found in eq.(3.40). In doing this, one needs to compute the relations between the integrals  $W_0(z)$ ,



$W_1(z)$ ,  $W_2(z)$ , which may be done as, e.g., in [25]. For example, the bubble generating function for both kind of models could also be simply expressed as

$$B(z) = A(D) (1 - Dz - 1/W_0(z)). \quad (3.42)$$

Changing the dimension dependent factor  $A(D)$  from  $A(D) = 1$ , when crossing is allowed, to  $A(D) = 1 - \frac{1}{D(D-1)}$ , when crossing is forbidden, is enough for the melting transition temperature to become finite for  $D \leq 3$ .



# Appendix B: Tauberian theorem and its applications

The discrete Tauberian theorem (see [27]) states that the relations

$$c(z) \equiv \sum_{N=0}^{\infty} c_N z^N \underset{z \rightarrow z^{*-}}{\sim} \left( \frac{1}{1 - z/z^*} \right)^\rho L \left( \frac{1}{1 - z/z^*} \right), \quad (3.43)$$

and

$$c_N \underset{N \rightarrow \infty}{\sim} \frac{N^{\rho-1}}{(z^*)^N \Gamma(\rho)} L(N), \quad (3.44)$$

are equivalent provided that  $\rho > 0$ ,  $\{c_N\}$  is a positive monotonic sequence and  $L$  is “slowly varying” in the sense specified in [27].

We apply this theorem here first to find the scaling laws of the unzipping transition to give an example of how the relations reported in section III can be recovered. When approaching the critical line at  $g = g_c$  from the unzipped phase, the smallest singularity is  $z_3(\beta g)$ , as defined in eq. (2.15). The leading behavior as  $z \rightarrow z_3^-$  is:

$$Z(z, \beta\epsilon, \beta g) \underset{z \rightarrow z_3^-}{\sim} \frac{1}{\exp(-\beta\epsilon) - Dz_3 - B(z_3)} \frac{1}{1 - z/z_3}, \quad (3.45)$$

$$\frac{\partial Z(z, \beta\epsilon, \beta g)}{\partial(\beta g)} \underset{z \rightarrow z_3^-}{\sim} \frac{1}{\exp(-\beta\epsilon) - Dz_3 - B(z_3)} \frac{1}{(1 - z/z_3)^2} \frac{\partial z_3 / \partial(\beta g)}{z_3(\beta g)}, \quad (3.46)$$

$$\frac{\partial Z(z, \beta\epsilon, \beta g)}{\partial(\beta\epsilon)} \underset{z \rightarrow z_3^-}{\sim} \frac{\exp(-\beta\epsilon)}{(\exp(-\beta\epsilon) - Dz_3 - B(z_3))^2} \frac{1}{1 - z/z_3}, \quad (3.47)$$

where we have neglected inessential factors when  $g \sim g_c$ . Consequently, the application of the Tauberian theorem yields:

$$Z_N(\beta\epsilon, \beta g) \underset{N \rightarrow \infty}{\sim} \frac{1}{\exp(-\beta\epsilon) - Dz_3 - B(z_3)} z_3^{-N}, \quad (3.48)$$

$$\frac{\partial Z_N(\beta\epsilon, \beta g)}{\partial(\beta g)} \underset{N \rightarrow \infty}{\sim} \frac{N}{\exp(-\beta\epsilon) - Dz_3 - B(z_3)} z_3^{-N} \frac{\partial z_3 / \partial(\beta g)}{z_3(\beta g)}, \quad (3.49)$$

$$\frac{\partial Z_N(\beta\epsilon, \beta g)}{\partial(\beta\epsilon)} \underset{N \rightarrow \infty}{\sim} \frac{\exp(-\beta\epsilon)}{(\exp(-\beta\epsilon) - Dz_3 - B(z_3))^2 z_3^{-N}}. \quad (3.50)$$

The canonical partition function  $Z_N$  can be expressed in terms of previously defined quantities as  $Z_N(\beta\epsilon, \beta g) = \sum_{\vec{x}} p_N(\vec{x}, \beta\epsilon) \exp(\beta\vec{g} \cdot \vec{x})$  (see eqs.(2.1) and (2.4)).

Note that at the critical line  $z_2(\beta\epsilon) = z_3(\beta g_c)$ , implying that the  $\epsilon$ -dependent denominator in the above equations vanishes, as the critical line is approached (that is  $\gamma \equiv (g - g_c)/g_c \ll 1$  and  $\tau \equiv (T - T_c)/T_c \ll 1$ ). In that limit the order parameters obey

$$\langle n \rangle \equiv \frac{\partial Z_N / \partial(\beta\epsilon)}{Z_N} \underset{N \rightarrow \infty}{\sim} \frac{\exp(-\beta\epsilon)}{A_1\gamma + A_2\tau}, \quad (3.51)$$

$$\langle x_g \rangle \equiv -\frac{\partial Z_N / \partial(\beta g)}{Z_N} \underset{N \rightarrow \infty}{\sim} \frac{\partial z_3 / \partial(\beta g)}{z_3(\beta g_c)} N. \quad (3.52)$$

The coefficients  $A_1(T_c, g_c), A_2(T_c, g_c)$  determine the normal direction in the  $(\log(T), \log(g))$  plane, and are reported below:

$$\begin{aligned} A_1(T_c, g_c) &= \beta_c g_c \left( D + \frac{\partial B(z_3(\beta_c g_c))}{\partial z_3} \right) z_3^2(\beta_c g_c) [4 \sinh(2\beta_c g_c) + 4(D-2) \sinh(\beta_c g_c)] \\ A_2(T_c, g_c) &= -\beta_c g_c \left( D + \frac{\partial B(z_3(\beta_c g_c))}{\partial z_3} \right) z_3^2(\beta_c g_c) [4 \sinh(2\beta_c g_c) + 4(D-2) \sinh(\beta_c g_c)] \\ &\quad \beta_c \epsilon \exp(-\beta_c \epsilon). \end{aligned} \quad (3.5)$$

As expected, it may be easily checked that, whatever direction we choose to approach the critical curve, the physical quantities as (3.51) only depend on the projection of this direction on the normal to the critical line. To show this, one only needs first to observe that criticality is achieved through the vanishing of the denominator in (3.51), and then to apply the implicit function theorem to this denominator.

Recently, some authors ([13, 14]) have suggested, as an alternative order parameter for the unzipping transition, the number  $m$  of monomers from the last contact to the end. By using the same tools as before, we can find some quantities of interest such as the probability distribution of having  $m$  monomers ‘‘liberated’’ (in the equations below we suppose  $1 \ll m \ll N$ ). In general:

$$P(m) = \frac{\sum_W \delta_{n_{u.z.}, m} p_W}{\sum_W p_W} \quad (3.55)$$

where  $W$  are the pairs of directed walks,  $p_W$  is their Boltzmann weight and  $n_{u.z.}$  is the number of ‘‘liberated’’ monomers of the configuration. The possible enforcing of the crossing constraint does not affect the following equations, simply shifting the critical melting temperature  $T_c$ .

At zero force and  $T < T_c$ , in the native state:

$$P(m) \propto \exp[-m \log(z_1/z_2)] m^{\alpha(d)}, \quad (3.56)$$

where  $m^{\alpha(d)} = m^{-1/2}$  in  $d=1$ ,  $(\log(m))^{-1}$  in  $d=2$  and constant for  $d > 2$ . We recall that  $z_1 = 1/D^2$ , with  $D = d + 1$ . At criticality:

$$P(m) \propto (N - m)^{\phi_t - 1} m^{\alpha(d)}, \quad (3.57)$$

where in the first factor there are logarithmic corrections in  $d = 2, 4$  to the power-law behavior controlled by the crossover thermal exponent  $\phi_t$  [17]. Above  $T_c$  we have:

$$P(m) \propto (N - m)^{\beta(d)} m^{\alpha(d)}. \quad (3.58)$$

with  $\beta(d) = -|1 - d/2| - 1$  if  $d \neq 2$ , in  $d = 2$   $m^{\beta(d)} = 1/(N \log^2(N))$ . At force different from zero, in the zipped phase, we get:

$$P(m) \propto \exp(-m \log(z_3/z_2)), \quad (3.59)$$

while in the unzipped phase:

$$P(m) \propto \exp(-(N - m) \log(z_2/z_3)), \quad (3.60)$$

if  $T < T_c$ , and

$$P(m) \propto \exp(-(N - m) \log(z_1/z_3)), \quad (3.61)$$

if  $T > T_c$ . Just at criticality, the probability distribution is flat.

For the Y-shaped configuration, we get, at force different from zero:

$$P(m) \propto \exp(-(N - m) \log(z_y/z_3)) \quad \text{u.z. phase}, \quad (3.62)$$

$$P(m) \propto \exp(-m \log(z_3/z_y)) \quad \text{z. phase}, \quad (3.63)$$

$$P(m) \propto \text{const.} \quad \text{at criticality}, \quad (3.64)$$

where  $z_y = \exp(-\beta\epsilon)/\mu$  is the singularity controlling the free energy per monomer in the Y-zipped phase, and  $\mu = D$  the connective constant of a single directed walk. At zero force we get:

$$P(m) \propto \exp(-(N - m) \log(z_y/z_1)) m^{\alpha(d)} \quad \text{denatured phase}, \quad (3.65)$$

$$P(m) \propto \exp(-m \log(z_1/z_y)) m^{\alpha(d)} \quad \text{native phase}, \quad (3.66)$$

$$P(m) \propto m^{\alpha(d)} \quad \text{at criticality}. \quad (3.67)$$

Let us give the details in a simple case, to get *e.g.* eqns.(3.62,3.63,3.64). One has:

$$P(m) = \frac{Z_{N-m}^{PS} Z_m^{free}}{Z_N} \quad (3.68)$$

where  $Z_{N-m}^{PS}$  is the partition function of two  $(N-m)$ -monomer chains with the end points in common (this is the original Poland-Sheraga model),  $Z_m^{free}$  is the partition of two  $m$ -monomer chains with no base pair in contact, and  $Z_N$  is given by eq.(2.1). We have, for the Y at non zero force,  $Z_{N-m}^{PS} = z_y^{-(N-m)}$ ,  $Z_m^{free} \sim z_3^{-m}$ , and

$$Z_N \sim z_y^{-N} \quad \text{native phase,} \quad (3.69)$$

$$Z_N \sim z_3^{-N} \quad \text{above criticality.} \quad (3.70)$$


---

# Appendix C: Y configurations with SAWs

In this Appendix we extend the analysis of section IV to a Y configuration where the two chains are generic SAWs and not directed walks. We take here  $\vec{g} = g(1, 0, \dots, 0)$ .

Let us call  $c_N(\vec{0}, \vec{x})$  the number of  $N$ -step SAWs ending in  $\vec{x}$  and starting in  $\vec{0}$ , and  $c_{N,\vec{g}} = \sum_{\vec{x}} c_N(\vec{0}, \vec{x}) \exp(\beta\vec{g} \cdot \vec{x})$  (so that  $c_N \equiv c_{N,\vec{g}=\vec{0}}$ ). As a first step, we need to prove the existence of a connective constant  $\mu_{\vec{g}}$  for a single SAW in the presence of a pulling force *i.e.*  $c_{N,\vec{g}} \sim \mu_{\vec{g}}^N$  for large  $N$ . The existence of a connective constant then follows [27] from the inequality:

$$\sum_{\vec{x}} c_N(\vec{0}, \vec{x}) \exp(\beta\vec{g} \cdot \vec{x}) \leq \sum_{\vec{x}, \vec{x}_1} c_{N_1}(\vec{0}, \vec{x}_1) c_{N-N_1}(\vec{x}_1, \vec{x}) \exp(\beta\vec{g} \cdot \vec{x}_1) \exp(\beta\vec{g} \cdot (\vec{x} - \vec{x}_1)) \quad (3.71)$$

which implies the subadditivity of  $c_N$  [27]. From the connective constants of RWs and DSAWs, we can establish the following bounds (recall that here  $\vec{g} = g(1, 0, \dots, 0)$ ) for the connective constant of a SAW in  $d$ -dimensions:

$$d - 1 + \exp(\beta|\vec{g}|) \leq \mu_{\vec{g}} \leq 2(d - 1) + 2 \cosh(\beta|\vec{g}|). \quad (3.72)$$

Let us now introduce the canonical partition function. It reads:

$$Z_N = \sum_{m=0}^{N-1} \exp(\beta(N - m)\epsilon) c_{N-m-1} \sum_{\vec{x}_1, \vec{x}_2} c'_{2m}(\vec{x}_1, \vec{x}_2) \exp(\beta\vec{g} \cdot (\vec{x}_2 - \vec{x}_1)), \quad (3.73)$$

where  $c'_{2m}(\vec{x}_1, \vec{x}_2)$  counts the SAWs which do not cross the first  $N - m$  attached monomers of the Y and  $\vec{x}_1, \vec{x}_2$  are the end-points of the strands.

An upper bound is found if we allow the opened part of the Y to cross the first  $N - m$  zipped monomers:

$$Z_N^{upper} = \sum_{m=0}^{N-1} \exp(\beta(N - m)\epsilon) c_{N-m-1} \sum_{\vec{x}_1, \vec{x}_2} c_{2m}(\vec{x}_1, \vec{x}_2) \exp(\beta\vec{g} \cdot (\vec{x}_2 - \vec{x}_1)). \quad (3.74)$$

The grand canonical partition function defined from (3.74) displays singularities in  $z \equiv z_1 = \frac{1}{\mu_{\vec{g}}^2}$  and in  $z \equiv z_2 = \frac{1}{\mu \exp(\beta\epsilon)}$ , where  $\mu$  is the connective constant of a standard SAW in  $d$ -dimensional space, *i.e.*  $\mu = \mu_{\vec{g}=\vec{0}}$ .

To find a lower bound for  $Z_N$ , we restrict the Y configurations to those with the joined part of the Y constrained to lay in the half-space  $\{\vec{x}|\vec{x}\cdot\vec{e}_1 > 0, \vec{e}_1\cdot\frac{\vec{g}}{|\vec{g}|} = 0\}$  (the origin is in the bifurcation point), and the unzipped part forced to stay in the other half-space.

$$Z_N^{lower} = \sum_{m=0}^{N-1} \exp(\beta(N-m)\epsilon) c_{N-m-1}^{left} \sum_{\vec{x}_1, \vec{x}_2} c_{2m}^{right}(\vec{x}_1, \vec{x}_2) \exp(\beta\vec{g}\cdot(\vec{x}_2 - \vec{x}_1)), \quad (3.75)$$

where we have indicated with  $c_n^{left(right)}$  the number of  $n$ -step SAWs constrained to stay in the left(right) half space. As the connective constants can be rigorously proved to be the same even for walks constrained in such a way (see [27, 28] for details, it is important here that walks generated from a reflection through the hyperplane  $\{\vec{x}|\vec{x}\cdot\vec{e}_1 > 0, \vec{e}_1\cdot\frac{\vec{g}}{|\vec{g}|} = 0\}$  leave the scalar product  $\vec{g}\cdot\vec{x}$  unchanged), the singularities of the grand partition function obtained from (3.75) are the *same* as those of  $\sum_{N=0}^{\infty} Z_N^{upper} z^N$ . Thus the grand partition function singularities associated to  $Z_N$  in eq.(3.73) are  $z = z_1, z_2$  given above.

In summary, we have proved that a mechanical unzipping transition takes place when  $\mu_{\vec{g}}^2 = \mu \exp(\beta\epsilon)$ , similarly to the directed walk case. We suggest that a standard computation of  $\mu_{\vec{g}}$  (for instance through exact enumeration) could give an accurate phase diagram for the ‘‘Y approximation’’ with SAWs.

The bounds in eq.(3.72) give immediately two bounds within which the transition line  $g_c(T, \epsilon)$  lies. In the  $T \rightarrow 0$  limit eq.(3.2) follows since both bounds give the same asymptotic transition line. This should be the same also for generic SAWs, since in the  $T \rightarrow 0$  limit it is sufficient to consider only Y-shaped configurations. This demonstrates the existence of cold unzipping at sufficiently low temperature for the generic self-avoiding case.



# Acknowledgments

I would like to thank a lot my supervisor, Prof. Amos Maritan for proposing me this topic of research, which I believe is really intriguing and very interesting, and for his strong encouragements and help, really effective when things start to be hard.

For the argument of this thesis, I have had the luck to cooperate with Antonio, and with such bigs in Physics as Flavio and Somen. They have all helped me a lot and I learned many things from all of them. I hope to have the pleasure to collaborate with them in the future also.

My deepest thanks goes to my wife Ilenia, for bearing me every single day! I also want to thank my cat Ortica, who is a great company when I come home during the weekdays. I am indebted to my family, who always enthusiastically (after thinking up a bit) approves my choices in life and manages to make me feel near them.

I also want to thank my friends from Cittadella, Cristiano and Andrea (who always complain that I do not answer to their e-mails!) for coming here in Trieste every now and then to have a good old evening together.

And last, but not least, I thank all my new friends in Trieste, the old foxes of Condensed Matter Sector, i.e. my former roommates Gianluca and Gianni; Andrea and Lorenzo, and my new roommates Sandro and Hoang. And of course my Sissa Condensed Matter friends (the Russians also for allowing me to beat them at table tennis all the times...)!



# Bibliography

- [1] K. Svoboda, S. M. Block, *Annu. Rev. Biophysics. Biomol. Structure*, **23**, 247, (1994).
- [2] A. Ashkin, *Proc. Natl. Acad. Sci. USA*, **94**, 4853, (1997).
- [3] A. Kishino, T. Yanagida, *Nature*, **34**, 74, (1988).
- [4] H. G. Hansma, *J. Vac. Sci. Technol.*, **B14**, 1390, (1995).
- [5] U. Bockelmann, B. Essevaz-Roulet, F. Heslot, *Phys. Rev. Lett.*, **79**, 4489, (1997).
- [6] B. Essevaz-Roulet, U. Bockelmann, F. Heslot, *Proc. Natl. Acad. Sci. USA*, **94**, 11935, (1997).
- [7] H. Yin, M. D. Wang, K. Svoboda, R. Landick, S. M. Block, J. Gelles, *Science*, **270**, 1653, (1995).
- [8] A. Kornberg, T. Baker, *DNA replication*, Freeman, New York, 1991.
- [9] R. M. Wartell, A. S. Benight, *Phys. Rep.*, **126**, 67, (1985).
- [10] B. H. Zimm, J. K. Bragg, *J. Chem. Phys.*, **31**, 526, (1959).
- [11] D. Poland, H. A. Sheraga, *J. Chem. Phys.*, **45**, 1456,1464, (1966).
- [12] M. Peyrard, A. R. Bishop, *Phys. Rev. Lett.*, **62**, 2755, (1989)
- [13] S. M. Bhattacharjee, *J. Phys. A*, **33**, L423 (2000); **33**, 9003(E) (2000); cond-mat/0010132, (2000).
- [14] D. K. Lubensky, D. R. Nelson, *Phys. Rev. Lett.*, **85**, 1572, (2000).
- [15] K. L. Sebastian, *Phys. Rev. E.*, **62**, 1128, (2000).
- [16] H. Zhou, cond-mat/0007015, (2000).
- [17] M. S. Causo, B. Coluzzi, P. Grassberger, , *Phys. Rev. E*, **62**, 3958, (2000).

- [18] Y. Kafri, D. Mukamel, L. Peliti, *Phys. Rev. Lett.*, **85**, 4988, (2000).
- [19] F. Seno *et al.*, submitted for publication, (2000).
- [20] R. J. Rubin, *J. Chem. Phys.*, **43**, 2392, (1965).
- [21] G. Forgacs *et al.*, in *Phase Transitions and Critical Phenomena*, Vol.14, Edited by C. Domb, J. L. Lebowitz. Vol. I, Clarendon Press, Oxford, 1995.
- [22] J. Cardy, *J. Phys. A*, **16**, L355, (1983).
- [23] N. Theodorakopoulos, T. Dauxois, M. Peyrard, *Phys. Rev. Lett.*, **85**, 6, (2000).
- [24] D. Cule, T. Hwa, *Phys. Rev. Lett.*, **79**, 2375, (1997).
- [25] R. J. Rubin, *J. Chem. Phys.*, **44**, 2130, (1966).
- [26] S. Lifson, *J. Chem. Phys.*, **40**, 3705, (1964).
- [27] B. D. Hughes, *Random Walks and Random Environments*, Vol. I, Clarendon Press, Oxford, 1995.
- [28] J. M. Hammersley *et al.*, *J. Phys. A: Math. Gen.*, **15**, 539, (1982).