

# DNA Physics Near Melting Point

*By*

**Tanmoy Pal**

**PHYS07200904004**

Institute of Physics, Bhubaneswar

*A thesis submitted to the*

*Board of Studies in Physical Sciences*

*In partial fulfillment of requirements*

*For the Degree of*

**DOCTOR OF PHILOSOPHY**

*of*

**HOMI BHABHA NATIONAL INSTITUTE**



**January, 2016**

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI. Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Date:- 5<sup>th</sup> January 2016

(Tanmoy Pal)

## CERTIFICATE

This is to certify that the thesis entitled “**DNA Physics Near Melting Point**”, which is being submitted by **Shri Tanmoy Pal**, in partial fulfillment of the degree of **Doctor of Philosophy in Physics** of **Homi Bhabha National Institute** is a record of his own research work carried by him. He has carried out his investigations for the last six years on the subject matter of the thesis under my supervision at **Institute of Physics, Bhubaneswar**. To the best of our knowledge, the matter embodied in this thesis has not been submitted for the award of any other degree.

Signature of the Candidate

Signature of the Supervisor

Tanmoy Pal  
Institute of Physics  
Bhubaneswar

Prof. Somendra Mohan Bhattacharjee  
Professor  
Institute of Physics  
Bhubaneswar

Date:- 5<sup>th</sup> January 2016.

## DECLARATION

I, Tanmoy Pal, hereby declare that the investigations presented in the thesis have been carried out by me. The matter embodied in the thesis is original and has not been submitted earlier as a whole or in part for a degree/diploma at this or any other Institution/University.

Date: 5<sup>th</sup> January 2016

(Tanmoy Pal)

*To My Maa*

---

# Contents

<b>List of Figures</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Duplex DNA . . . . .	2
1.2 Triplex DNA . . . . .	3
1.3 DNA and Thermodynamics . . . . .	5
1.4 Denaturation of dsDNA . . . . .	6
1.4.1 Chemical Denaturation . . . . .	6
1.4.2 Thermal Melting . . . . .	6
1.4.3 Forced Unzipping . . . . .	8
1.5 Classical Models of DNA . . . . .	9
1.5.1 Freely Jointed Chain (FJC) Model . . . . .	9
1.5.2 Semiflexible Chain Model . . . . .	10
1.6 Directed Polymers . . . . .	10
1.7 A Simple Model For Thermal Melting . . . . .	11
1.8 Triplex DNA Melting And The Efimov-DNA . . . . .	13
1.8.1 A New State of DNA . . . . .	13
1.8.2 The Quantum Efimov Effect . . . . .	14
1.8.3 Efimov Effect in DNA . . . . .	14
1.8.4 Issues With The Polymer RG . . . . .	16
1.9 DNA Elasticity . . . . .	17
1.9.1 Puzzles in the DNA Looping . . . . .	17
1.9.2 Rigidity of a Melting DNA . . . . .	18
1.10 Organization of The Thesis . . . . .	19

<b>Appendix A</b>	<b>An Exactly Solvable Model For Forced Unzipping</b>	<b>21</b>
<b>Appendix B</b>	<b>Necklace Model</b>	<b>25</b>
<b>Appendix C</b>	<b>Renormalization Group</b>	<b>27</b>
<b>2</b>	<b>Temperature Dependence of Entropic Elasticity of DNA</b>	<b>30</b>
2.1	The Model . . . . .	31
2.1.1	Recursion Relation And Observables . . . . .	32
2.1.2	Elastic Response Under a Stretching Force . . . . .	33
2.1.3	Role of The Bubbles . . . . .	40
2.1.4	Elastic Response in Presence of Unzipping Force . . . . .	41
2.2	Discussion . . . . .	45
2.3	Conclusion . . . . .	45
<b>3</b>	<b>Bubble Mediated Softening of Rigid DNA</b>	<b>47</b>
3.1	Bias Induced Melting . . . . .	48
3.1.1	Elastic Response . . . . .	51
3.2	Rigid Model . . . . .	54
3.2.1	Thermal Melting : $g_s = 0$ . . . . .	55
3.2.2	Elastic Response : $g_s \neq 0$ . . . . .	57
3.2.3	Role of The Bubbles . . . . .	60
3.3	Discussion . . . . .	62
3.4	Conclusion . . . . .	63
<b>4</b>	<b>Absence of Efimov-DNA Without Denaturation Bubbles</b>	<b>64</b>
4.1	The Model . . . . .	65
4.1.1	Single chain . . . . .	65
4.1.2	Bound state, Y fork, duplex, $g_2$ and $g_3$ . . . . .	66
4.2	Diagrammatic definitions and rules . . . . .	67
4.2.1	Free chain . . . . .	68
4.2.2	Bound state . . . . .	68
4.2.3	$\mathbf{k}$ and $s$ Conservation . . . . .	69
4.3	Two Chains . . . . .	69

4.4	Three Chains . . . . .	71
4.5	Conclusion . . . . .	74
<b>Appendix D Bound State</b>		<b>75</b>
<b>Appendix E Rules Of Diagrammatic Calculations</b>		<b>76</b>
<b>5</b>	<b>Efimov-DNA and Renormalization Group Limit Cycle</b>	<b>80</b>
5.1	Model and result . . . . .	81
5.1.1	Strand exchange and $g_3$ . . . . .	81
5.1.2	Qualitative description . . . . .	82
5.2	Three Chain Problem With Bubbles: $g_2 \neq 0$ . . . . .	84
5.2.1	Critical Case: $g_2 = g_{2c}$ . . . . .	86
5.2.2	Scale-free Limit . . . . .	87
5.2.3	For $\Lambda < \infty$ . . . . .	88
5.2.4	Complex Fixed Points and Periodicity . . . . .	89
5.2.5	Discrete Scale Invariance . . . . .	92
5.3	Off-critical: $g_2 \neq g_{2c}$ . . . . .	93
5.4	Discussion . . . . .	94
5.5	Conclusion . . . . .	97
<b>Appendix F A simple <math>\beta</math> function for <math>\hat{g}_3</math></b>		<b>98</b>
<b>6</b>	<b>Summary</b>	<b>100</b>

## Acknowledgement

I take this opportunity to express my sincere gratitude to my advisor Prof. Somendra Mohan Bhattacharjee for the continuous support for my PhD study and research. Special thanks for being patient with me. His quest to find something new every other time motivates me a lot. I really enjoyed discussing Physics with him which rekindled my interest in Biology. And, as a bonus, I also gained ideologically.

Many of my colleagues helped me with my work on different occasions. Some of them are Dr. Trilochan Bagarti, Dr. Poulomi Sadhukhan, Dr. Jaya Maji and Dr. Sourabh Lahiri.

I also acknowledge the support I got from Dr. Goutam Tripathy, Dr. Biju Sekhar, Prof. Kalyan Kundu, Prof. A. M. Jayannavar and Dr. Sumedha.

Special thanks to my football and tennis mates.

## Synopsis

As a most important molecule for life DNA has attracted the attention of the physicists for decades. Applications of known physical principles in DNA have already revealed a great deal of results relevant in biology as well as physics. Apart from being a challenging testing ground of applicability of the known physics this field has the potential of breeding even new physics.

When suspended in a solvent a number of hydrogen-bonded base pairs of DNA are always broken depending on the solvent condition and the temperature. These single stranded unbonded segments are called denaturation bubbles. By varying a control parameter, say temperature, behavior of these bubbles can be manipulated. The increase in thermal fluctuation energy due to increase in temperature results in increasing fluctuations in the average number of bubbles as well as the average bubble length. At a critical temperature ( $T_c$ ) these fluctuations become so severe that the two strands of the double stranded DNA (dsDNA) separate out of each other giving two independent single strands. This second order phase transition is called the thermal denaturation of DNA or DNA melting. Here we have investigated two aspects of the physical properties of DNA near this thermal melting point. They are elastic response of the dsDNA and the Efimov-DNA phase of the three, or more, stranded DNA.

Inside a cell DNA is often subjected to stretching and bending to facilitate fundamental processes such as DNA replication, gene regulation etc. For this it is very important to study how the DNA responds under mechanical forces. Around kilo base pairs level elastic properties of DNAs have been studied extensively and classical DNA-elasticity models such as the worm-like chain model provides satisfactory answers. For DNAs as short as 100 base pairs there are a few studies of it's elastic properties. These studies currently do not give clear cut answers whether the worm-like chain model works here or not. For reasonably long DNA chains it is possible to see signatures of DNA melting which ideally requires infinite chain length. Here we have attempted to explore this area by introducing very simple models. It is a well established fact that the single stranded DNA is far more flexible than the dsDNA. The existence of denaturation bubbles, even of shorter lengths, in a dsDNA may thus significantly increase the flexibility of a dsDNA. As the number as well as the length of the bubbles fluctuate severely around the melting point we expect dramatic effects

of these bubbles in elastic response of the DNA. Our aim is to study the effects of the bubbles.

We have studied here three related but different models of DNA. We call them Model-I, Model-II and Model-III. In all the three models we consider the free single strands as 2 dimensional directed random walk on a square lattice. The base pair interaction is mimicked by a contact interaction between the two polymers if the monomers are at the same contour length. In the case of directed polymers, this native base pair interaction is satisfied in a natural way. We ignore all other structural details of the DNA. The Binding-unbinding transition in this system of two polymers represents DNA melting. In fact the melting behavior of the low dimensional model resembles many features of higher dimensional models. This allows us to use the two dimensional model to probe the question of elasticity by using exact analytical and numerical techniques. To measure the elastic response we apply a co-ordinate independent stretching force of same strength at the rear end of the two polymers. Polymers are attached to the origin at the near end.

Model-I consists of two Gaussian chains of fixed lengths with contact interactions. This model is exactly solvable even in presence of an unzipping force. We get melting by imposing the constraint that the polymers can not cross each other. Above the critical temperature, when the DNA is in unbound state, an application of the stretching force induces an unbound-bound phase transition with a critical stretching force. We have shown that this is a second order phase transition. The macroscopic elastic modulus shows a finite jump at the critical stretching force signifying the unbinding-binding transition. We have also studied this model under both the stretching force and the unzipping force all applied at the rear ends of the two strands. The presence of even a small unzipping force changes the nature of the binding-unbinding transition under stretching from second order to first order. Extensibility curve shows a finite jump resulting in a delta function peak in the elastic modulus at the critical stretching force. We have solved this model exactly by using generating function technique and verified the results numerically by employing the transfer matrix method. We are thus able to draw the complete three dimensional phase diagram in the temperature-stretching force-unzipping force space. The zero force thermal melting point is now a tricritical point.

In Model-II we impose a Boltzmann weight to those two-chain configurations where the two chains bend in some particular direction while staying in the bound state. We do not apply the non-crossing constraint here. The restriction resulting from the Boltzmann weight alone induces a thermal melting transition. We get a second order binding-unbinding transition under only the stretching force and a first order transition when both the forces are present. Unlike Model-I, there is a spatial anisotropy which manifests in melting of a bound state by a stretching force if pulled in the unfavorable direction. Above the critical stretching force at a given temperature there is no bound state for reasonably high strengths of the stretching force. This is a force induced melting transition. We have analytically solved the purely thermal zero force part of this model. For the analysis with both the forces we rely on the exact numerical data.

In Model-III we introduce a bubble opening and closure weight to track the bubble related quantities. To make the model simpler we have completely excluded the possibility of the two strands going in a particular direction while in the bound state. To differentiate between a normal contact and a bubble opening or closure we impose the restriction that if a contact occurs immediate next step is also a contact. To do these we had to give memory in the transfer matrix formulation unlike the previous two models which were Markovian. Like Model-II here also we get melting transition without imposing the non-crossing constraint. Under a stretching force the system again undergoes a binding-unbinding transition of second order. The elastic modulus shows a finite jump at the critical stretching force. But now the jump is significant in that it jumps well above the unbound state elastic modulus indicating an anomalous behavior. We study the connections between the anomalous elastic behavior and the bubble fluctuations, by using the exact numerical methods.

Now we come to the Efimov-DNA part of our work. Although DNA exists commonly as the classical B-DNA many other alternative conformations, both intramolecular and inter-molecular, are possible. Relevant to our work is the case of the three-stranded DNA. In one case a third strand sits in the major groove of a Watson-Crick base paired DNA and attaches with one of the two strands by forming hydrogen bonds through Hoogsteen base pairing forming a triple-helix. In another case, relevant for our work, if one of the strands in an unbound segment of a DNA finds

a third strand of suitable complementarity then also triple-stranded DNA complex forms. This second process is called strand exchange and we use this mechanism to formulate the Efimov physics problem in a three-chain system. The origin of the Efimov physics is the three-body quantum problem. In such a system with short-range pair potential only a series of three-body bound states appear when the interaction strength is not enough to hold even a pair in the bound state. Occurrence of these strange trimers is attributed to the quantum fluctuation generated long range attractive effective potential. It was reported that this pure quantum effect has its classical equivalence by mapping the three-body quantum problem to a Gaussian-three-chain problem. The classical counterpart of the quantum fluctuation, here, is the fluctuation in the bubble lengths due to thermal fluctuations. Severe bubble length fluctuations around the dsDNA melting point increase the probability of strand exchange with a third strand and in the process creates an effective attraction which can accommodate a three-stranded DNA bound state when no two are bound. This state of the three-stranded DNA is called an Efimov-DNA. Our aim is to show that this Efimov-DNA is represented by a characteristic renormalization group limit cycle.

We study this problem from renormalization group (RG) perspective. In RG one deals with the length scale dependence of the system parameters. These parameters goes to fixed values with increasing length scale. We take a system of three directed Gaussian polymers in  $(3 + 1)$  dimension. To avoid the technical difficulties of solving the full three chain problem we assume that any two of the three chains are in the bound state. Due to thermal fluctuations at any finite temperature there will be bubbles in the dsDNA bound state. We call this dsDNA with bubbles as duplex DNA. The bubbles are loops of Gaussian chains which are very flexible and the bound segments are not allowed to bend in our model. The bubbles are created by successive dissociation of a bound state into two single strands and corresponding closure of these two strands into another bound segment. We give these dissociation and closure same statistical weight. This weight is our two-body parameter. We do a necklace model type analysis of this model of duplex DNA. The results show a second order melting transition at a critical value of the two-body parameter. We evaluate the partition function of the duplex around this critical point.

Along with the duplex we add a third single strand. This third strand can interact

with the duplex giving rise to our three-body parameter. If we do not allow any bubble to form then the third strand has to interact with the rigid bound state. A RG analysis of this model shows that this problem is equivalent to the problem of absorbance of single polymer in a surface. The resulting states are described by real fixed points of the three-body parameter and no special three-body effects emerge here. We then allow the bubbles to form in the duplex and perform a momentum-shell like analysis of this problem at the duplex melting point after adding a third strand. We basically want to find the third virial coefficient of this system which is related to the force between the three chains. Instead of going to real fixed points three-body parameter now runs into complex fixed points and becomes log periodic. By going into the complex plane of the three-body parameter we explicitly show how the closed trajectories of the limit cycle oscillations appear. By redefining our parameters we were able to reproduce the square root behavior of the famous Efimov plot for our three-stranded DNA system. Right at the melting point total number of Efimov-trimers is infinite. By using specific limits of our model we have extended, through interpolation, this analysis of our system for non-critical values of the two-body parameter. As we go far from the critical point the total number of trimers becomes finite and decreases.

The quantum Efimov trimers are very non-trivial to detect experimentally. This is due to the difficulty of tuning the interaction parameters as desired. As the polymer related parameters are easily tunable we expect our work may help detecting these elusive trimers. We have discussed a number of scenarios where one may find the experimental signatures of the existence of the Efimov states. We specifically suggest to measure the third virial coefficient. Small angle neutron scattering (SANS) technique may be of great help to detect the oscillations and the zero wavelength divergence suggested in our work.

## Publications:

(a) Published :

- \* Renormalization Group Limit Cycle for Three-Stranded DNA, Tanmoy Pal, Poulomi Sadhukhan, and Somendra M. Bhattacharjee, Phys. Rev. Lett. 110, 028105 (2013).
- \* Efimov-like phase of a three-stranded DNA and the renormalization group limit cycle, Tanmoy Pal, Poulomi Sadhukhan, and Somendra M. Bhattacharjee, Phys. Rev. E 91, 042105 (2015).

(b) To Be Communicated :

- \* Rigidity of a Melting DNA, Tanmoy pal, Somendra M. Bhattacharjee, arXiv.1508.04122.

\* Papers included in the thesis

# List of Figures

1.1	Single Strand DNA Backbone : Nucleotide units made of sugar and phosphate form the backbone of a single strand through covalent bonding. The chemically different P and S ends are called 5' and 3' ends. . . . .	2
1.2	Schematic diagrams of (a) a DNA double-helix, (b) a DNA triple-helix. The double-helix is shown as an ideal B-DNA structure. The B-DNA has enough space in its major groove to accommodate an extra single strand. The third strand can utilize this space and extra H-bonding sites of the Purines (here A) to form a triple-helix. . . . .	3
1.3	Watson-Crick and Hoogsteen mechanism of base pairing. . . . .	4
1.4	Strand exchange mechanism between a single strand (solid black line) and a duplex DNA. The single strand invades the unboned opening and binds through H-bonding (vertical green dotted lines). . . . .	5
1.5	Schematic diagram of DNA denaturation. (a) Thermal denaturation through bubble nucleation. Here $T_1 < T_2 < T_3$ . (b) Forced unzipping by applying an unzipping force $g_u$ . The arrows show the directions in which $g_u$ acts. In these type of experiments generally the near end of the DNA is anchored to a substrate shown as a brown rectangle. . . . .	7
1.6	A DNA configuration with three bubbles (U) and four bound states (B). . . . .	12
1.7	Schematic diagram of strand exchange and the equivalent coarse-grained three chain interaction. A single strand (solid black line) pairs with the strands (dash-dot lines (left)) of a bubble on a duplex (thick line (right)). Short green dotted lines indicate base pairings. . . . .	16

A.1	Two strands of a DNA is schematically represented as two directed non-crossing random walks in $D = (1 + 1)$ dimensional square lattice. $y$ is the contact interaction weight. $g_u$ is the unzipping force. The arrows show directions of $g_u$ . . . . .	21
A.2	Fraction of bound base pairs discontinuously goes to zero at $y = 1.43$ for a fixed $g_u = 0.1$ indicating a first order transition. . . . .	23
A.3	Re-entrant phase diagram. . . . .	24
B.1	Graphical solution of the duplex state singularity for $1 < \Psi < 2$ . The solid blue line is the bubble partition function and the red straight lines are $1/G_B(s)$ for different temperatures. The parameter values are taken as $\epsilon = f_u = -g_2 = -1$ . . . . .	26
C.1	Renormalization group flows for different $\epsilon$ values. The arrows show the flow directions with increasing length scale. The red and the green discs are the unstable and the stable fixed points respectively. . . . .	28
2.1	Schematic representation of the dsDNA as two directed walks in $(1 + 1)$ dimension. The polymers can not cross each other. The bound segments can bend left or right freely. The direction in which the forces act are indicated by the arrows. (a) DNA under stretching force, (b) DNA under unzipping force. . . . .	32
2.2	Flexible Model : The phase diagram of the system under a stretching force. The phase boundary separates the bound phase from the unbound phase. In the region $4/3 > y > 1$ there exists a critical $g_{sc}$ for every value of $y$ . Above $y = 4/3$ the system is by default in the bound state. Red squares are the numerically obtained critical points and solid blue line is the analytical curve. . . . .	34
2.3	Flexible Model : The variation of $n_c$ with $g_s$ for different values of $g_u$ and $y$ . A non-zero $n_c$ indicates a bound state. For $g_u = 0$ , there is no unbound state for $y \geq 4/3$ . The cyan line with triangles is for the melting point $y = y_c$ , and shows the quadratic dependence on $g_s$ . The black curve with filled diamonds is for $y = 1.4 > y_c$ . . . . .	36

2.4	Flexible Model : Plot of the average extension as a function of the stretching force with a fixed $y = 1.20$ . The average extension varies continuously around the critical point. . . . .	37
2.5	Flexible Model : Elastic modulus curve for $y = 1.20$ . The solid brown line is the analytically obtained curve. The dashed cyan line is for $\kappa_b = 4 \operatorname{sech}^2(2g_s)$ Eq. (2.18a) for the case with no bubbles. Other curves are the plots of $\kappa$ for different system lengths. . . . .	38
2.6	Flexible Model : Magnified version of 2.5 around the same crossing point of all the curves. The chain length for each curve is given in the legend. . . . .	38
2.7	Flexible Model : Elastic modulus as a function of $y$ for zero stretching force. . . . .	39
2.8	Flexible Model : Elastic modulus as a function of $g_s$ at $y = 4/3 = y_c$ and $y = 1.40 > y_c$ . . . . .	40
2.9	Flexible Model : Three dimensional phase diagram of the system in the the presence of both stretching and unzipping forces. All the point on the surface except the curve for $g_{uc} = 0$ represents first order phase transition points. The unzipping line for $g_s = 0$ is shown by the thick black line. . . . .	42
2.10	Flexible Model : $x$ vs $g_s$ plot with $g_u = 0.4$ and $y = 1.20$ . The inset figure magnify the critical region. The solid magenta line is the analytical curve. All the other curves are for finite system sizes shown in the legend box. The discontinuity at a $g_{sc}$ indicates a first order transition. . . . .	43
2.11	Flexible Model : $\kappa$ vs $g_s$ plot with $g_u = 0.4$ and $y = 1.20$ . The inset figure magnifies the critical region. The solid magenta lines are the analytical curves. All the other curves are for finite system sizes shown in the legend box. The peak height increases proportionally with $N$ signaling a $\delta$ -function peak which is not shown in the analytical curve. . . . .	44

3.1	Schematic representation of the dsDNA as two directed walks in $(1+1)$ dimension. The polymers can cross each other. (a) The options for the bound segments can be restricted by introducing a statistical weight $b$ . In this figure $b$ is associated to the right degree of freedom. (b) Rigid Model : The bound segments can not bend to the right and they are at least 2 bonds long. $v$ is the weight associated to the bubble opening or closure. . . . .	48
3.2	$n_c$ vs $y$ plot for $b = 0.5$ . $n_c$ becomes finite at $y = 1.142$ . Numerical data is also consistent with the analytical result. . . . .	50
3.3	$C_c$ vs $y$ plot for $b = 0.5$ . Plot shows a finite discontinuity at $y = 1.142$ . Numerical data also shows very similar behavior. . . . .	50
3.4	Numerical $x$ vs $g_s$ plot for $b = 0.5$ for different system sizes. $y = 1.20$ . System size dependence of $x$ is very weak. . . . .	51
3.5	Numerical $\kappa$ vs $g_s$ plot for $b = 0.5$ and $y = 1.20$ for different system sizes. There are very small but identifiable gradually increasing peaks. All the curves has a common crossing point at $g_s = 0.18$ . . . . .	52
3.6	Numerical $n_c$ vs $g_s$ plot for $b = 0.5$ and $y = 1.20$ for different system sizes. $n_c$ becomes non-zero to zero continuously. . . . .	53
3.7	Numerical $C_c$ vs $g_s$ plot for $b = 0.5$ and $y = 1.20$ for different system sizes. All the curves has a common crossing point at $g_s = 0.18$ . . . . .	53
3.8	Numerical phase diagram in the $y_c$ - $g_{sc}$ plane for $b = 0.5$ . . . . .	54
3.9	Rigid Model : $n_c$ increases from zero to non-zero values continuously at a $y > 1$ before approaching the saturation value 1. This indicates a binding-unbinding transition. . . . .	56
3.10	Rigid Model : $C_c$ vs $y$ plots for different $N$ . Peak Height increases with increasing $N$ but eventually saturates creating a finite discontinuity. The discontinuity occurs at $y = 1.18$ where all the curves meet. . . . .	56
3.11	Rigid Model : Continuous stretching of the DNA to its full extent by both positive and negative forces. $y$ value is fixed at 1.20. . . . .	57

3.12	Rigid Model : $\kappa$ shows an increasing peak around $g_s = 0.02$ with increasing system length $N$ with a fixed $y = 1.20$ . Maximum values of $\kappa$ , $\kappa_{\max}$ , are plotted vs $1/N$ in the inset. A linear fit with the first four points (solid blue line) gives the estimate for $N \rightarrow \infty$ , $\kappa_{\max} = 4.95$ . This indicates there is a finite discontinuity in $\kappa$ . The solid black line is the plot of the function $2 \operatorname{sech}^2(g_s)$ , the unbound state elastic constant. For $g_s > 0.02$ , $\kappa$ matches with the unbound state modulus. . . . .	58
3.13	Rigid Model : $y$ value is fixed at 1.20. The inset image shows $n_c$ decreases continuously from finite values to zero indicating a continuous phase transition. Peak heights in $C_c$ vs $g_s$ curves saturate indicating a finite discontinuity. Around $g_s = 0.02$ all curves meet which is the critical point. . . . .	59
3.14	Rigid Model : Numerical phase diagram for the binding-unbinding transition. . . . .	59
3.15	Rigid Model : $n_b$ vs $g_s$ and $l_b$ vs $g_s$ plot for $y = 1.20$ . As the critical point approaches $n_b$ becomes very small but $l_b$ becomes as large as $N$ . . . . .	60
3.16	Rigid Model : $n_b$ vs $g_s$ and $C_b$ vs $g_s$ plot for $y = 1.20$ . Around the critical point $n_b$ is very small but its fluctuation $C_b$ is very large. . . . .	60
3.17	Rigid Model : $\frac{1}{x}\sqrt{n_b}$ vs $y$ curves for different system lengths collapse to a single master curve in the bound region. $g_u = g_s = 0$ . $y_c = 1.18$ . . . . .	61
3.18	Rigid Model : $\frac{1}{\kappa}\sqrt{C_b}$ vs $y$ curves for different system lengths collapse to a single master curve in the bound region. $g_u = g_s = 0$ . $y_c = 1.18$ . . . . .	61
4.1	Basic building blocks. Panel (a) represents $Z(\mathbf{k}, s)$ for a Gaussian chain, (b) $Z_b(\mathbf{k}, s)$ for a two chain bound state, (c) a Y fork representing the interface between a bound pair and two open strands. It has a weight $g_2$ . . . . .	68
4.2	(a) The duplex partition function as an infinite series of bound pairs and bubbles, (b) Y fork for a duplex. (c) A three-chain interaction, $g_3$ , involving a free chain and a duplex. . . . .	69
4.3	Interaction of free chain with a bound state in absence of bubbles. . . . .	72

4.4	In the limit $a \rightarrow 0$ , $V_0 \rightarrow \infty$ for a finite $V_0 a^d$ . In this limit a particle with energy $-E$ looks close to the threshold and remains outside of the well most of the time. The solid brown line represents the wave function. . . . .	74
E.1	A bound state with one bubble showing variables in the Fourier-Laplace space variable obeying the $\mathbf{k}$ and the $s$ conservations. . . . .	76
5.1	(Color online) Schematic diagram of a strand exchange and the equivalent coarse-grained three-chain interaction $g_3$ . (a) A single strand [blue line marked (1)] pairs with one strand of a bubble on a duplex [brown lines marked (2)]. The short vertical (green) lines indicate base pairings, with the energy per unit length $\epsilon$ . The junction weight $g_2$ is associated with each fork or the interface on a duplex. In (a) there are four interfaces as indicated by the arrows. (b) A coarse-grained version of (a) where the duplex is represented by a thick line interacting with the single line. The filled circle represents the three-chain interaction $g_3$ . . . . .	81
5.2	Diagrammatic representation of the three-chain partition function. The hatched circle is the effective interaction $W$ . This figure translates into an integral equation involving interactions to all order. . . . .	86
5.3	Closed elliptical trajectories in the complex $H$ plane as $\Lambda$ is varied. These are drawn for different starting values of $(H_1, H_2)$ . All loops in the upper half plane have the fixed point $(1 + is_0)/(1 - is_0)$ as a focus while $(1 - is_0)/(1 + is_0)$ as a focus for the lower half plane. . . . .	91
5.4	Plot of $H$ as a function of $\Lambda$ showing zeros and divergences with $\Lambda_* = 1$ .	92
5.5	(Color online) Schematic flow diagram in the $\hat{g}_2$ - $\hat{g}_3$ plane. $\hat{g}_2 = 0$ corresponds to the bound state (no bubbles) while $\hat{g}_2 = \hat{g}_2^*$ (box) is the duplex melting point. The flow along the thick vertical line through $\hat{g}_2 = \hat{g}_2^*$ is periodic. Along the $\hat{g}_2 = 0$ line, there is a stable fixed point $\hat{g}_3 = \hat{g}_{3c}$ (filled disk) which represents the peeling of one polymer from the rigid bound state. A typical flow from a point away from the melting point is shown. . . . .	95

F.1 (Color online) A schematic diagram of the flow of the RG equation, Eq. (F.2), showing the approach to the limit cycle. The limit cycle is an ellipse in the  $\hat{g}_2 = 0.5$  plane [blue line marked (a)]. A point with off-critical  $\hat{g}_2$  is shown to approach the planar limit cycle as  $\Lambda \rightarrow \infty$  [red line marked (b)]. . . . . 99

# Chapter 1

## Introduction

DNA is a bio-molecule found, without any known exception, in every cellular life-form. In biology the fundamental phase of reproduction is the cell-division. During cell division the DNA of the parent cell is first copied and then transmitted to the offspring. This DNA transmission is absolutely necessary because it contains all the hereditary informations. Without the cellular environment DNA can not be replicated. For example, the virus can replicate their DNA only by hijacking the replication machinery of a host cell which they enter through infection. As the life is mortal, without reproduction or rather DNA replication, its sustenance is not possible. The DNA has been called ‘the book of life’ and rightly so [1–3].

Biology has attracted physicists for quite a long time. It’s back in the 1944 when Schrödinger posed and attempted to answer the question *What is Life?* [4]. Then came the discovery of the molecular structure of DNA by Franklin, Watson and Crick in 1953 [5]. This discovery has boosted the growth of this branch of physics many-fold. Since then DNA has been a testing ground for known physics. Physical theories like statistical mechanics, information theory, topology, network theory etc. has been employed to study it with varying success. Active research has produced many fascinating results and at the same time posed more challenging questions. From the current scenario one may be tempted to speculate that DNA might produce ‘new physics’ one day.

Millions of years of evolution has resulted in a huge varieties of living organisms.

Though different in many aspects the way in which they store their hereditary information in DNA is the same. To know this information storage procedure one needs to have a knowledge of the structure of a DNA.

## 1.1 Duplex DNA

DNA is most commonly found in the B-DNA form [1–3]. It is a very long double stranded polymer. Each strand is made of many monomers which are called nucleotides. Nucleotides are made off one sugar-phosphate group and a base. They form covalent bonds between each other to create a linear chain. At the one end of this chain sits a sugar group and a phosphate group rests on the other. This gives a directionality to the single strands which plays very important roles in the physical mechanisms of the DNA. The phosphate and the sugar ends are called 5' and 3' ends respectively, see Figure 1.1. The sugar-phosphates form the backbone of the each

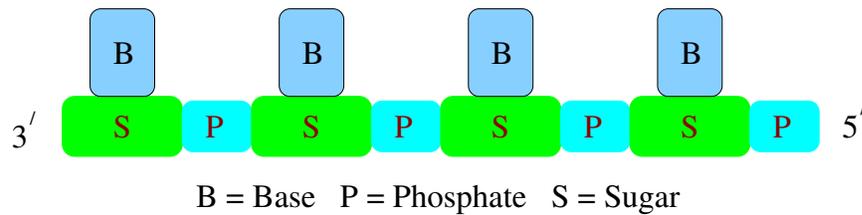


Figure 1.1: Single Strand DNA Backbone : Nucleotide units made of sugar and phosphate form the backbone of a single strand through covalent bonding. The chemically different P and S ends are called 5' and 3' ends.

strand and the bases can vary monomer to monomer. There exist only four kinds of bases namely A, T, G and C where A and G are the Purines (Pu) and C and T are the Pyrimidines (Py). Any hereditary information is stored through these four letter alphabet. Different sequences of the bases are expressed through different proteins which is the reason of the amazing variety of species prevalent in the nature. One base can pair with another base by forming Hydrogen-bonds (H-bonds) with the restriction that one A can pair with a T only and one G with a C. This is the Watson-Crick complementarity. Two strands of a DNA are complementary to each other. They wound around each other to form a double-helix structure while the bases are paired

through the H-bonds along the contour. Ideally two nucleotides of different strands can be H-bonded together to form a base-pair (bp) if they are at the same contour length along the single strands. The structure of a double stranded DNA (dsDNA)

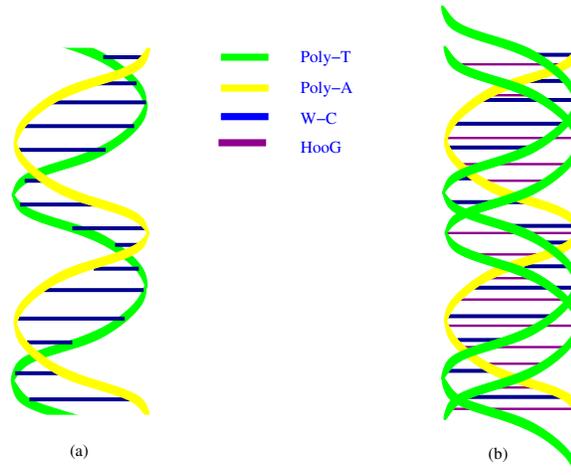


Figure 1.2: Schematic diagrams of (a) a DNA double-helix, (b) a DNA triple-helix. The double-helix is shown as an ideal B-DNA structure. The B-DNA has enough space in its major groove to accommodate an extra single strand. The third strand can utilize this space and extra H-bonding sites of the Purines (here A) to form a triple-helix.

is shown schematically in Figure 1.2(a) as a right handed double-helix. Two strands of the double-helix are anti-parallel to each other in the sense that if one strand has a direction  $5' \rightarrow 3'$ , the other strand must have the direction  $3' \rightarrow 5'$ .

## 1.2 Triplex DNA

The structural features of the DNA stated up to now is a bit idealized. In *in vivo* situations, the structure of DNA is found to be rather dynamic [3]. DNA can exist in many other forms. One of them, most relevant to this thesis, is the triplex DNA (tsDNA). Occurrence of the triple helical structures were reported as early as 1957 [6]. The Poly(Pu) strand of a normal Watson-Crick DNA of Poly(Pu).Poly(Py) can H-bond with a third Poly(Py)/Poly(Pu) strand to form a tsDNA. The third strand binds with the dsDNA through Hoogsteen pairing [7–9]. A dot (.) is used to represent

the Watson-Crick pairing, an asterisk (\*) will be used to represent the Hoogsteen H-bonding. For example, a single strand of Poly(T) can bind with the Poly(A) strand of a Poly(A).Poly(T) dsDNA to form a Poly(T)\*Poly(A).Poly(T) triple helix. Beside the Hoogsteen pairing an alternative reverse Hoogsteen mechanism also can take place. If the third strand in the tsDNA is a Poly(Py), it binds with the Poly(Pu) central strand of the Watson-Crick DNA in parallel through Hoogsteen pairing and if the third strand is a Poly(Pu) it binds through reverse Hoogsteen pairing in the anti-parallel fashion. The corresponding chemical bondings are shown in Figure 1.3 schematically. The major groove of a Watson-Crick DNA has enough space to accommodate a third

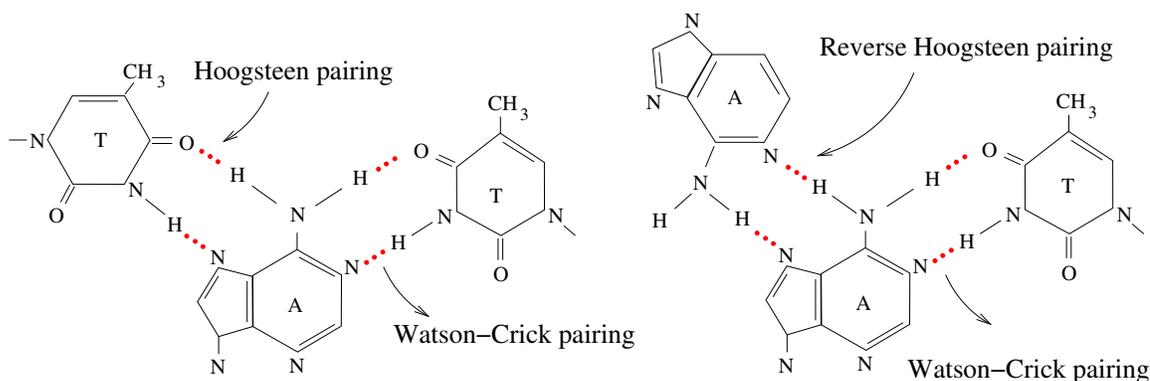


Figure 1.3: Watson-Crick and Hoogsteen mechanism of base pairing.

strand. Only the Purines has the additional H-bonding sites that can be utilized through the Hoogsteen mechanism. For that the central strand of this kind of a triplex must be a Poly(Pu) sequence. A triple-helical form is shown schematically in Figure 1.2(b). Even the tsDNA comes in many varieties. When the third single strand is supplied by another molecule the resulting triplex is called an inter-molecular triplex. DNA is a very long molecule and it can locally melt for various reasons to form single strand pairs. One single strand of the unbound DNA segment can bind with its parent DNA through Hoogsteen pairing forming an intra-molecular triplex. H-DNA is a famous example of the last kind [10–12]. Triplex structures has been found both in *in vitro* and *in vivo* conditions [13]. They are expected to play major roles in cell-mechanisms like gene regulation, DNA replication, transcription etc. [14,15]. Another major potential field of application for this unconventional structures is the anti-gene strategy. Synthetically made oligonucleotides can recognize specific pre-determined

sequences in a duplex DNA and form local triplex structures and thus modulate its normal physical activities. These are called triplex forming oligonucleotides (TFO). Their special ability to recognize DNA sequences without rupturing the double helix makes them a potential candidate for universal anti-gene drugs [7]. The activities and the utilities of the triplex structures have drawn serious attentions quite recently and it is a very active field of research now.

Triple stranded DNA can exist sans triple helical structure. Such structures are formed during genetic recombination through *strand exchange* mechanism [16]. In this process a dsDNA opens up locally exposing its constituent single strands. Then a single strand, from a second molecule or from itself, invades the exposed segment and binds with one of its strands provided that it finds its exact complementary sequence. This is schematically shown in Figure 1.4. For a long dsDNA this can happen simultaneously at many places. In this way the genetic information is exchanged between two DNAs. It also helps in DNA repair and replication.

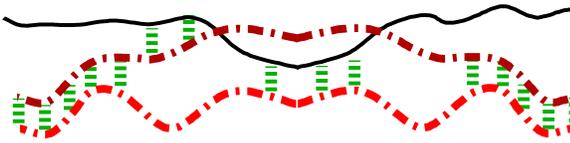


Figure 1.4: Strand exchange mechanism between a single strand (solid black line) and a duplex DNA. The single strand invades the unbonded opening and binds through H-bonding (vertical green dotted lines).

### 1.3 DNA and Thermodynamics

Be in the *in vitro* experiments in salt solutions or in the *in vivo* cellular conditions DNA has to tackle thermally noisy environments. The estimated energy of a base pair is 0.13 - 0.26 eV which is fairly comparable to the thermal fluctuation energy characterized by  $k_B T$  where  $k_B$  is the Boltzmann constant and  $T$  is the absolute temperature [17]. For example, at room temperature (300 K)  $k_B T = 0.026$  eV. So, at normal temperatures thermal fluctuation can disrupt some base-pairing randomly to form locally melted segments, known as denaturation *bubbles*, in the otherwise bound

DNA. The bubbles are made of unpaired single strands which play many important roles in physical processes. When all the base pairs are considered bound on the average even then the energy required for a transition from one configuration to another configuration, e.g. through bending, is  $\sim 10 k_B T$  [18]. These make the DNA a topic of interest for statistical physicists. In thermodynamics  $(T, S)$  constitutes a conjugate pair where  $S$  is the entropy. For a DNA the entropy originates from the polymeric correlations. When a mechanical force ( $g$ ) is applied the response of DNA,  $x$ , acts as a conjugate variable. Depending on the nature of the force  $x$  could be the measure of either extension or distance between the two strands. Here  $g$  acts as the intensive variable [19].

## 1.4 Denaturation of dsDNA

In this process all the bound base-pairs get broken and as a result of that two completely separated single strands appear as an end product. The dsDNA is a remarkably stable molecule. The factors which contribute the most to stabilize it is the H-bonding and base-pair stacking interactions [1–3]. To achieve denaturation one must overcome these stabilizing forces. The denaturation can happen in three different ways: chemically, through thermal melting and through forced unzipping.

### 1.4.1 Chemical Denaturation

If the pH of a DNA solution is increased beyond 12 or decreased below 2 the constituent bases can no longer form H-bonds as they get ionized [3, 20]. The reagents urea and formamide, in appropriate concentrations, form H-bonds with the bases blocking their complementary bases. In the absence of the H-bonds, two strands separate. While using different chemicals one should be careful not to create any cleavage in the strands [21].

### 1.4.2 Thermal Melting

If the temperature of a DNA in a solution is increased at a critical temperature ( $T_c$ ) its two constituent strands separate. This is called the thermal melting which is a

genuine phase transition [22–25]. It can be described through bubble nucleation as shown in Figure 1.5(a). As the temperature is raised, more and more base pairs

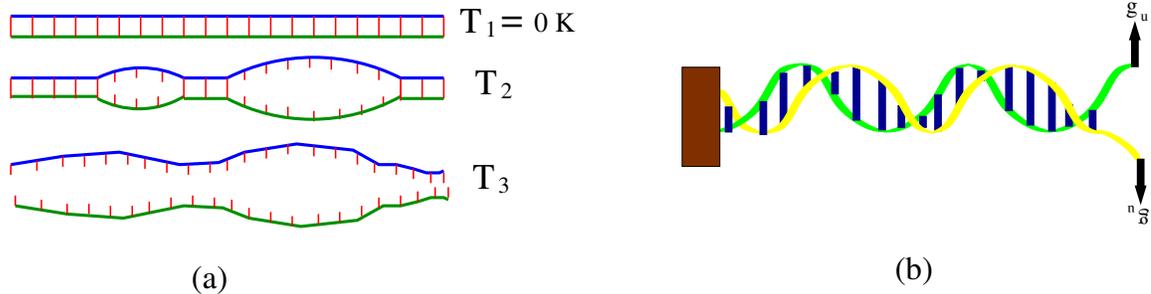


Figure 1.5: Schematic diagram of DNA denaturation. (a) Thermal denaturation through bubble nucleation. Here  $T_1 < T_2 < T_3$ . (b) Forced unzipping by applying an unzipping force  $g_u$ . The arrows show the directions in which  $g_u$  acts. In these type of experiments generally the near end of the DNA is anchored to a substrate shown as a brown rectangle.

open and in the process smaller bubbles coalesce together to form larger bubbles. Around  $T_c$  the bubble size becomes equal to the DNA length separating two strands completely. Depending on the salt conditions of the solution and DNA sequence,  $T_c$  can vary  $80 - 100 \text{ } ^\circ\text{C}$ . The base A binds with the base T forming two H-bonds while G pairs with C through three H-bonds. This makes the G.C pairs stronger than the A.T pairs.  $T_c$  is influenced by this asymmetry. A DNA with high G.C content will melt at higher temperatures than the one with higher A.T content. The melting is a reversible process. After denaturation if the temperature is lowered slowly then the DNA regains its previous double-helix form but a fast decrease results in a globule formation. Though thermal melting is an old problem the order of the transition is not completely settled yet. Depending on the specific interactions one uses in a theoretical model the order can vary. There are proposals of it being a first order [26–30], second order [26, 31, 32], even infinite order [33] phase transition.

### 1.4.3 Forced Unzipping

DNA hides its base pairs at the center of the helix exposing only the sugar-phosphate backbone to the environment. This helps it keeping safe the invaluable genetic information to be passed to an offspring. On the other hand, it needs to open up at different places along its contour to facilitate the essential physiological processes e.g. DNA replication, RNA transcription, gene regulation, genetic recombination [1–3,21]. As the cells of the most of the mammals are pH neutral with temperature 36-37 °C the chemical and the thermal denaturation can not help to induce a denaturation due to their unphysical requirements discussed above. The cell employs force instead. Different helicases like dnaB, PcrA etc. exert mechanical force to unzip the DNA partly or totally [34–39]. This forced unzipping is a cooperative phenomena when done in a solution.

With the advancements of the experimental methods it is now possible to manipulate a single isolated molecule of DNA through the optical tweezer or an AFM tip [40–42]. A typical set up used in the single molecule forced unzipping experiments is shown schematically in Figure 1.5(b). By anchoring one end to a substrate forces are applied at the other ends of the single strands in the opposite direction and perpendicular to the helix-axis. This is done below the melting temperature. As the strength of the force is increased, two strands are peeled apart abruptly at a critical force for a fixed temperature resulting in two separate single strands. The unzipping force required to unzip a DNA at room temperature can be measured very accurately [43–45]. Although experimental verification of the reported temperature dependence of the critical unzipping force has still not been materialized.

The unzipping as a cooperative phenomena was first demonstrated by Bhattacharjee by analyzing a non-Hermitian Hamiltonian for a directed two polymer system with contact interactions and unzipping force [46]. Recognizing its importance in biology many other works have come up using both equilibrium and non-equilibrium set ups [47–53]. In the non-equilibrium case the application of a periodic force results in a hysteresis loop [54–57]. For the equilibrium case all the studies agree that unzipping is a first order transition and the critical force is a function of the temperature or vice-versa. It is now also an established fact that the unzipping force can not penetrate the bound state [19] (also see Appendix A).

## 1.5 Classical Models of DNA

A long chain of DNA easily bends and when suspended in a solution it takes the random coil form. Shorter chains are less easily bendable. If one goes on fragmenting, below a length chains are very hard to bend. This is the *persistence length* which is normally  $\sim 150$  bp. DNA has its own elasticity and it resists deformations in its structure. As the A.T pairs are softer than the G.C pairs A.T rich segments are deformed relatively easily. Although statistical models of DNA elasticity often ignore the microscopic details. Below we describe two basic models of this kind which are used extensively.

### 1.5.1 Freely Jointed Chain (FJC) Model

In this simplest model a polymer is considered as a random walk of  $N$  steps with step length  $a$  such that the total length of the polymer is  $L = Na$  [58]. The monomers do not interact with each other and can rotate freely at any angle in the continuum. Let us consider this problem in one dimension with the starting end of the polymer being fixed at  $x = 0$ , where  $x$  is the position coordinate of a monomer. The probability of finding the other end at a distance  $X$  is given by

$$P(x) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{x^2}{2N}}, \quad (1.1)$$

where we have set  $a = 1$ . This gives the most probable value of  $X = 0$ . Treating  $a$  as a constant means the polymer is inextensible. Though changing from one configuration to another configuration does not cost any energy the model has its own free energy which originates from entropy. The free energy of the system is

$$F \approx -k_B T \frac{X^2}{2N}, \quad (1.2)$$

such that the force required to hold the polymer at a small  $X \ll N$  is given by

$$g = k_B T \frac{X}{N}. \quad (1.3)$$

This is the familiar Hookean elasticity with a temperature dependent entropic elastic modulus.

## 1.5.2 Semiflexible Chain Model

A FJC is rather flexible. Stiffness can be introduced in a polymer by introducing a bending energy cost. In a semiflexible chain, a tangent at a monomer interacts with its next neighbor tangent through angular correlation [58, 59]. The Hamiltonian for a polymer of chain length  $N$  subjected to a constant mechanical force  $\mathbf{g}$  is given by

$$H = -K \sum_{i=1}^{N-1} \mathbf{t}_i \cdot \mathbf{t}_{i+1} - \mathbf{g} \cdot \sum_{i=1}^{N-1} \mathbf{t}_i, \quad K > 0 \quad (1.4)$$

where  $K$  is the bending elastic constant,  $\mathbf{t}_i$  is the tangent vector at the  $i$ th monomer and the smallest bond length is again taken to be unity. The two-point correlation function of this model is given by

$$\langle \mathbf{t}_i \cdot \mathbf{t}_{i+n} \rangle = e^{-n/l_p}, \quad (1.5)$$

where the persistence length  $l_p \approx K/k_B T$ . As the correlation drops exponentially any two monomers separated by a distance  $b \approx 2l_p$  are virtually independent of each other and the polymer can be treated as a FJC of smallest bond length  $b$  provided  $K \gg k_B T$ . This discretized model is known as the Kratky-Porod model and its continuum version is known as the worm-like chain model. The force extension relation for this model can be given by the following interpolation formula [60]

$$g \approx \frac{k_B T}{l_p} \left[ \frac{X}{L} + \frac{1}{4} \frac{1}{(1 - X/L)^2} - \frac{1}{4} \right]. \quad (1.6)$$

## 1.6 Directed Polymers

Consider a polymer in  $d$  dimension. Now embed it in  $D = d + 1$  dimension with its contour length as the extra direction, say the  $z$ -axis. As the length is a variable like time which can only grow the polymer can not trace back in that direction. Although its coordinates can fluctuate freely up to their maximum extent (the polymer length  $N$ ) in all the other transverse directions. This is a description of a directed polymer (DP). A freely jointed chain in this  $D$  dimensional continuum can be described by the Hamiltonian

$$H = \frac{1}{2} \int_0^N \left( \frac{\partial \mathbf{r}(z)}{\partial z} \right)^2 dz, \quad (1.7)$$

where  $\mathbf{r}(z)$  is the position coordinate of a monomer of contour length  $z$  and the chain starts from the origin with  $\mathbf{r}(z = 0) = \mathbf{0}$ . The corresponding partition function is given by

$$Z(\mathbf{r}, z) = \int_{\mathbf{0},0}^{\mathbf{r},z} \mathcal{D}[\mathbf{r}(z)] e^{-\beta H}, \quad (1.8)$$

where  $\beta = 1/k_B T$  and  $\mathcal{D}[\mathbf{r}(z)]$  represents integration over all possible configurations. This partition function obeys the following diffusion equation

$$\frac{\partial Z(\mathbf{r}, z)}{\partial z} = D \nabla^2 Z(\mathbf{r}, z), \quad (1.9)$$

where the diffusion constant  $D = 1/(2\beta)$ . Its solution is given by the Gaussian

$$Z(\mathbf{r}, z) = \frac{1}{(2\pi\beta z)^{d/2}} e^{-\frac{\mathbf{r}^2}{2\beta z}}, \quad (1.10)$$

which is nothing but the probability of finding the far end point of the polymer at  $(\mathbf{r}, z)$ .

## 1.7 A Simple Model For Thermal Melting

The zero force thermal denaturation of a DNA is described by the Poland-Scheraga model simply and elegantly [28, 29]. Taking advantage of the finite temperature bubbles a DNA can be considered as a sequence of double stranded (bound state) segments and single stranded loops (bubbles). For this bead-chain structure it is called the necklace model. With the assumption that the bound states (B) and the bubbles (U) do not interact with each other the contribution of a sequence in the partition function is obtained simply by multiplying the individual partition functions of the each segments. The system partition function is the sum of all the contributions from the all possible sequences. For simplicity, we take the starting and the ending segments as bound states. A configuration with three bubbles and four bound segments is shown schematically in Figure 1.6. Let  $Z_B(z)$  and  $Z_U(z)$  be the partition functions of a bound state and a bubble of length  $z$  respectively. Other than that an additional Boltzmann weight  $g_2$  is associated for each bound-unbound state junction. With the constraint that the total length of the DNA,  $N$ , is fixed the canonical partition function is given by

$$Z_D(N) = Z_B(N) + \int_0^N dz_2 \int_0^{z_2} dz_1 Z_B(N - z_2) g_2 Z_U(z_2 - z_1) g_2 Z_B(z_1) + \dots, \quad (1.11)$$

where the first term is for no bubble, the second term is for one bubble and so on. The length constraint can be relaxed by going to the grand canonical ensemble



Figure 1.6: A DNA configuration with three bubbles (U) and four bound states (B).

where length of the each segment varies arbitrarily. This is done through the Laplace transformation with respect to the length coordinate. It can be easily recognized that the terms in Eq. 1.11 has the convolution form in length. So a Laplace transformation of that equation with the definitions

$$G_X(s) = \int_0^\infty Z_X(N)e^{-sN}dN, \quad (1.12)$$

where  $X = (D, B, U)$  will result in a geometric series given by

$$G_D(s) = G_B(s) + g_2^2 G_B(s)^2 G_U(s) + \dots \quad (1.13)$$

which is easily summed to have

$$G_D(s) = \frac{G_B(s)}{1 - g_2^2 G_B(s) G_U(s)}. \quad (1.14)$$

The free energies of the completely bound and the denatured states are given by the singularities of  $G_B(s)$  and  $G_U(s)$  in  $s$ -space respectively. For a non-zero  $g_2$  there can be a third state, the duplex, where the bubbles and the bound segments co-exist. The singularity obtained by equaling the denominator of Eq. 1.14 to zero which is same as solving the equation

$$\frac{1}{G_B(s)} = g_2^2 G_U(s) \quad (1.15)$$

for a  $s$  corresponds to the duplex state. If the single strands of the bubbles are modeled as directed FJCs and a bound state as a rigid rod in  $D = d + 1$  dimension then the individual partition functions can be taken as

$$Z_B(N) = e^{-\beta\epsilon N}, \quad \text{and} \quad Z_U(N) \approx \frac{e^{-Nf_u}}{N^\Psi}, \quad (1.16)$$

where  $\Psi = \frac{d}{2}$  is the reunion exponent [61] and  $\epsilon$  and  $f_u$  are the free energies of the bound and the unbound states respectively. The largest singularity changes as one

varies the temperature and there can be a melting transition as a result.  $\Psi$  plays a major part in this transition. For  $\Psi < 1$  there can not be any transition, for  $1 < \Psi < 2$  the transition is second order (see Appendix B) and for  $\Psi > 2$  the transition happens discontinuously.

## 1.8 Triplex DNA Melting And The Efimov-DNA

The necklace model describes the melting of a dsDNA. What happens when a triplex DNA is heat treated ? It also melts. Let us consider the inter-molecular triplex melting case here. The dsDNA is a more stable structure than the tsDNA as the Hoogsteen pairing is weaker than the Watson-Crick pairing. For this the third strand will separate first at temperature  $T_t$  resulting in a free single strand and a duplex DNA. If the temperature is increased further, the duplex will also melt when the temperature reaches  $T_c$  ( $> T_t$ ). For a continuous melting the bound segments and the bubbles coexist at  $T_c$  and it is dominated by the bubbles of large lengths which increases the possibility of strand exchange. This scenario can be modeled as a system consisting of a duplex DNA at its melting point when a third strand is added to it.

### 1.8.1 A New State of DNA

Consider a three-particle quantum system with pairwise short-range potential. Apart from the occurrence of the usual three-body bound state, a very special phenomenon occurs at the critical two-body zero-energy state. An infinite number of three-body bound states appear though the corresponding potential is not appropriate to bind any two of them; the removal of any one of them destroys the bound state. This phenomenon, valid for any short-range interaction, is known as the Efimov effect. The size of the three-body bound states, or Efimov trimers, is large compared to the potential range, and so it is a purely quantum effect [62–66].

It was argued in Refs. [67–69] that near the thermal melting of a dsDNA, the formation of large bubbles enhances the possibility of a strand exchange involving each strand of the bubble. This leads to an effective long range attraction of the original pair, mediated by the third strand. As a result, a three-strand bound state is formed where no two are bound. Such a novel state of DNA, produced by fluctuations, has

been called an Efimov-DNA in analogy with the Efimov effect in three-body quantum mechanics.

### 1.8.2 The Quantum Efimov Effect

The Efimov effect was studied originally as a three-body quantum mechanics problem by solving the Schrödinger equation in the Fadeev approach [62–66]. Several approximate methods were also used, most notable among which is the use of the Born-Oppenheimer approximation for a separable short range potential [70]. Such calculations show the emergence of the scale-free  $1/r^2$  attraction at the quantum critical point of unbinding, as does the polymer scaling of Ref. [67]. Field theoretic methods were initiated much later on as an effective field theory for few-body quantum mechanics. Some successes were met in the diatom approach where the effect was studied in the scattering of a single particle from a diatom [66, 71]. The major effort in these attempts was to see the Efimov effect as a universal effect, emerging from a diverging length scale, here the scattering length. It led to the idea of a limit cycle renormalization group (RG) flow [72], which was introduced in a different context in Ref. [73]. The characteristic signature of the effect is the geometric sequence ( $E_n = a^n E_0$ ) of energy eigenvalues, the Efimov tower, and it is believed to emanate from this limit cycle behavior. It occurs only at the point where the length scale diverges. For nearby points one still gets three-body bound states but with a finite number of states. Although proposed in nuclear physics, experimental signatures for the Efimov effect started pouring in only after the technological developments in handling cold atoms [74–76].

### 1.8.3 Efimov Effect in DNA

The possibility of an Efimov-DNA was first pointed out by using a scaling argument and by a real space renormalization group approach to three polymers that can be implemented exactly for hierarchical lattices [67, 68]. That the melting transition with a large or diverging length scale is crucial (equivalent to infinite scattering length in the quantum version) was clearly brought out by the polymer scaling, that reproduces the  $1/r^2$  interaction, where  $r$  is the distance between two polymers. This interaction

owes its origin to the long range polymer correlations in a big bubble. The importance of the transition was also made clear in studies of the DNA problem in lower dimensional fractal lattices. In fact, a mixed phase different from the Efimov DNA was predicted, for which a quantum analog is not known [69]. On the polymer front, the strangeness of the long-range interaction is evident from the renormalization-group analysis of two polymers with  $-g/r^2$  interaction in the presence of a short range attraction [77–79]. The unbinding transition is described, as usual, by a fixed point, but that is not all. First, the fixed point is  $g$ -dependent, and, second, the order of the transition is determined by the reunion exponent of the bubbles at the  $g$ -dependent fixed point describing the unbound phase [61]. More unusual is the possibility of complex fixed points. This happens for  $g > 1/4$ . The complex fixed points are responsible for certain periodicity of various thermodynamic quantities and is similar to the origin of the Efimov tower.

The similarity between the zero temperature quantum problem and the classical thermal system of polymers actually follows from an imaginary time transformation of the quantum problem in the path integral approach [67]. For example, a path integral computation of the quantum problem could identify polymer like phases [80]. The fluctuations in the size of the polymer bubbles near the melting point of dsDNA play a similar role as that of quantum fluctuations near the unbinding transition of a pair of particles. The DNA bubbles correspond to the paths in the classically forbidden region of the short range potential [81]. There are many other similarities of the results for a DNA with the diatom trick used in the quantum version and it reinforces the idea that the features of the quantum Efimov effect could be observed in a classical setting of DNA in a solution [67]. The effect of fluctuations is not just restricted to the melting point itself. Even above the melting point, the bound state persists, eventually melting at a temperature higher than the duplex melting temperature. The limit cycle that occurs at the melting point is actually unstable as we move away from this special point. The number of turns in a way determines the number of bound states. Beyond a certain point all such states vanish. This point or temperature will be the melting point of the Efimov DNA.

### 1.8.4 Issues With The Polymer RG

The idea of RG is to look at a problem based on length scales and not on the numerical values of the parameters *per se*. How the various parameters change with the length scale then tells us the behavior of the macroscopic system in the large size limit (see Appendix C). The procedure is to (i) integrate out the small length scale fluctuations, especially bubbles and (ii) then by rescaling generate a similar system but with renormalized parameters. The flows of the parameters as the scale is changed give us all the crucial large length-scale results. In general, an RG approach is expected to lead to fixed points and separatrices, at most lines of fixed points. The fixed points represent states of the system which show scale invariance under a continuous rescaling of lengths. As pointed out above, it is rather rare to see a limit-cycle-like behavior because its periodicity would produce a discrete scale invariance only.

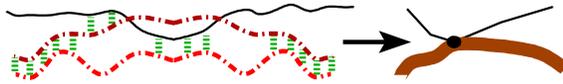


Figure 1.7: Schematic diagram of strand exchange and the equivalent coarse-grained three chain interaction. A single strand (solid black line) pairs with the strands (dash-dot lines (left)) of a bubble on a duplex (thick line (right)). Short green dotted lines indicate base pairings.

Polymer problems traditionally start with a random walk or a Gaussian polymer as the primary representation of a polymer. Most of the DNA melting theories are of this type where the free model represents the unbound states. In the Gaussian polymer model (no self-interaction), the melting transition is continuous [82, 83] so that the closer one is to the melting temperature, the larger is the size of the thermally generated bubbles. In this approach the two-chain and the three-chain problems have critical dimensionality  $d = 2$  and  $d = 1$  respectively [84]. Evidently a strand exchange for a small bubble in a three-chain system would, on scales larger than the bubble size, look like a three-strand interaction (Figure 1.7). Therefore, in three dimensions, one needs to consider irrelevant variables and, so, a straightforward perturbative RG fails here. This makes a complete analysis of the three-chain problem formidable. To circumvent this, a different approach has to be adopted.

## 1.9 DNA Elasticity

With its H-bonding, base stacking interactions and other factors DNA is a fairly rigid molecule. Yet DNAs are found in the bent state very frequently. Long chains of DNAs are wrapped around the histone or histone like proteins for DNA packaging. For example, a human cell of length  $\sim 1 \mu m$  contains DNA of total length  $\sim 1$  meter which is packaged into the microscopic space of the cell nucleus. Other than that different DNA binding proteins induce appreciable bending. A smooth bending is associated to the A-tracts which are continuous runs of A bases in a sequence. Two different models were introduced to explain the relation between the bending and A-tracts. They are the junction model and the wedge model. In the junction model it is suggested that whenever a part of a B-DNA segment admits non-B-DNA structures it bends [85]. On the other hand, the wedge model associates a specific bend angle to each AA dinucleotide. Thus depending on how many AA dinucleotides the A-tract has, it imparts a smooth bend of appropriate angle [86]. The bends play crucial role in the gene regulation. For example, a sequence can be made unavailable to the RNA-polymerase through bending [87]. This blocks the expression of that specific sequence. Bends also become handy for the repair enzymes [88]. Wherever mismatch (violation of complementarity) of base pairs occurs DNA bends and the repair enzymes feel the bend to recognize and correct the mismatch. It was also found that the replication origins contain natural or induced bends [89].

### 1.9.1 Puzzles in the DNA Looping

To facilitate different fundamental biological processes like replication, gene expression, assembly of functional nucleoprotein structures, packaging of viral DNA and others, DNA has to go through a lot of twisting and stretching along with bending [1, 89–94]. Generally different proteins induce these conformational changes in DNA but not without facing any resistance. This is because when subjected to an external mechanical force DNA responds elastically. A single stranded DNA may be flexible, easy to bend, but a double stranded DNA is known to be more rigid with its rigidity determined by the angular interactions between its nearest neighbor tangent vectors. This interaction results in a characteristic length scale, the persistence length

( $\sim 150$  base pairs), within which a dsDNA acts more or less like a rigid rod. Thus, it seems, a larger force is required to bend a DNA of length smaller than its persistence length than the one of a longer length. However the flexibility of short DNA fragments are important for different *in vivo* mechanisms, like those already mentioned, where loops or bends of length as small as 100 base pairs are involved [95, 96], and also in *in vitro* experiments where fragments are used in open or hairpin geometries. It is therefore important to probe the elastic response of a dsDNA not only in the thermodynamic limit of large length — dsDNA is long — but also for finite size systems.

If a dsDNA is treated as a free Gaussian polymer where the different monomers do not interact with each other, even then it shows an entropic elasticity [58, 59]. However, it is often treated as a semiflexible chain with an intrinsic rigidity to accommodate a persistence length. Recent debates [18, 97–107] on the behavior of short segments brought into focus the importance of broken base pairs on its eventual or effective rigidity. As the base pair energy is comparable to the thermal energy at physiological temperatures ( $\sim 2\text{-}3$  kcal/mole), bubbles form spontaneously or are produced by external forces (see, e.g. [108] for an earlier study). The issue of elastic response of a dsDNA cannot therefore be studied in isolation as its intrinsic property but rather needs to be coupled to the inner degrees of freedom, namely base pairings responsible for the bound state.

### 1.9.2 Rigidity of a Melting DNA

The thermal melting of DNA is by itself an interesting problem and important for different *in vivo* or *in vitro* processes. A notable example is the polymer chain reaction (PCR) which is used extensively in DNA amplification. Above the melting temperature the dsDNA separates into two independent single strands. In the temperature region below the melting point there can be local melting at different positions creating denaturation bubbles which are nothing but the single stranded loops preceded and succeeded by double stranded segments. As a single stranded DNA is far more flexible than the paired ones these thermally generated bubbles can provide flexible hinges which can make a dsDNA significantly flexible [18, 99]. Generally the average length of these bubbles increases as one approaches the critical point and equals the

system length at the melting temperature. Different from thermal melting is the unzipping transition where the two strands of a dsDNA are pulled apart at temperatures below the melting point. The unzipping transition is generally first order [19,46]. Since the unzipping force does not penetrate the bound state [19,43], the nature of the bubble distribution does not change in presence of an unzipping force. A change in the elastic behavior is therefore expected near the transition. The DNA melting is a genuine phase transition for which the DNA length has to satisfy the thermodynamic limit. But still, the existence of a transition point is sufficient to affect a finite size system even when it is away from the critical point.

Different statistical mechanical models have been applied with varying success to study the DNA elasticity problem. The classical semiflexible chain model with no denaturation bubbles has been employed by a number of investigators [109,110]. Segments made of single strands can be introduced in the discretized semiflexible DNA, by considering models comprising of two-state internal coordinates, and also, by coupling these internal coordinates to the external rotational degrees of freedom of its tangent vectors [111–113]. Their primary concern was to explain the recently reported higher looping probability of short ( $\sim 100$  bp) DNA fragments than what predicted by the worm-like chain model. The length of the bubbles considered in this model are often a few base pairs long. It is quite instructive to consider a model which studies the effects of the large bubbles in the flexibility of a bound DNA and compares it with the unbound state flexibility in a single set up.

## 1.10 Organization of The Thesis

We have organized the thesis in following way.

In Chap. 2, we study the temperature dependence of DNA elasticity by using a model where rigidity comes only through entropic effects. Using the stretching force as a probe, we obtain the elastic modulus for different temperatures including the thermal melting temperature. Above the melting temperature the system becomes bound after a critical stretching force continuously. We obtain the phase diagram of this transition in the critical-stretching-force–temperature plane. The elastic properties are also studied in the presence of an unzipping force which changes the nature

of the transition making it first order. Corresponding phase diagram is presented in the three dimensional temperature–stretching-force–unzipping-force space.

In Chap. 3, we introduce a model of DNA with intrinsic rigidity. After finding its zero force thermal melting point, we study its elastic behavior at different temperatures. The bubble related quantities like the bubble number, the bubble length and the bubble number fluctuation are also calculated as a function of stretching force and temperature. For the zero force case, we present quantitative relation between the extension and the square root of the bubble number. We also show the elastic modulus is related proportionally to the square root of the bubble number fluctuation for zero force.

In Chap. 4, a Poland-Scheraga type model of DNA melting is introduced. Here, we model the double stranded segments of the DNA at a finite temperature as rigid rods which can not bend but can rotate freely. The locally melted segments are modeled as Gaussian chain loops. We show this model goes through a second order melting transition. By considering the interaction between a rigid bound state and a single chain we show that no three-body features appear as the system shows similar results like a two-chain system.

In Chap. 5, we solve the full three-chain problem by considering a finite temperature duplex DNA as a bound state of the Poland-Scheraga model introduced in Chap. 4. By allowing strand exchange between the duplex and a single strand, added separately, we calculate the third virial coefficient. We define a dimensionless three-body parameter and obtain its flow equation using a non-perturbative RG technique. The Efimov-DNA appears at the duplex melting point and it corresponds to a RG limit cycle. We investigate the existence of the Efimov states above the duplex melting point by introducing a toy RG flow equation which shows that the number of Efimov-trimers decreases as the system moves away from the critical point. The possible experimental scenarios where an Efimov-DNA may be detected are also suggested at the discussion.

We summarize our results in Chap. 6. Throughout the thesis, necessary supplementary materials are provided at the chapter ending appendices.

# Appendix A

## An Exactly Solvable Model For Forced Unzipping

Here we introduce a transfer matrix based exactly solvable model of DNA unzipping [114]. Each strand of a DNA is represented by a  $D = (1+1)$  dimensional random walk in a square lattice shown schematically in Figure A.1. The chains can grow only in

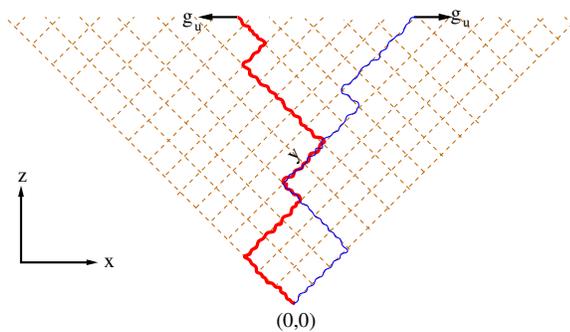


Figure A.1: Two strands of a DNA is schematically represented as two directed non-crossing random walks in  $D = (1 + 1)$  dimensional square lattice.  $y$  is the contact interaction weight.  $g_u$  is the unzipping force. The arrows show directions of  $g_u$ .

the positive  $z$ -direction and can not trace back. This leaves the chains only with two options to move, either left or right. The coordinates of any monomer can be found simply by taking projections on the  $x$  and the  $z$  axes. The native base pairing which only occurs at the same contour length of the both strands is naturally achieved here and represented by the attractive contact interaction weight  $y = e^{-\beta\epsilon}$  where  $\epsilon = -|\epsilon|$

is binding energy. First let us consider the  $g_u = 0$  case where this model shows a thermal melting transition when the chains are not allowed to cross each other. The corresponding recursion relations are given by

$$Z_{n+1}(x_1, x_2) = \sum_{(i,j)=\pm 1} Z_n(x_1 + i, x_2 + j)[1 - (1 - y)\delta_{x_1, x_2}], \quad (\text{A.1})$$

where  $Z_n(x_1, x_2)$  is the canonical partition function of the system of two polymers, each of length  $n$  and the spatial positions of the  $n$ th monomers of polymer 1 and polymer 2 are  $x_1$  and  $x_2$  respectively. For a given monomer number if  $x_1$  becomes equal to  $x_2$  then there is a contact. We set the initial condition as  $Z_0(0, 0) = y$  such that two strands start from the origin. The non-crossing constraint is implemented by not letting  $x_1$  becoming greater than  $x_2$  ( $x_1 \leq x_2$ ). We solve Eq. A.1 through generating function technique. In this technique a fugacity like variable  $z$  is introduced through which one can define the generating function as the following

$$G(z, x_1, x_2) = \sum_{n=0}^{\infty} z^n Z_n(x_1, x_2). \quad (\text{A.2})$$

We make the ansatz

$$G(z, x_1, x_2) = A(z)\lambda(z)^{|x_1 - x_2|}. \quad (\text{A.3})$$

By multiplying the both sides of Eq. A.1 with  $z^n$  and summing over  $n$   $Z$ s can be replaced by  $G$ s with the help of Eq. A.2. Then using the ansatz Eq. A.3 in the resulting equations  $A(z)$  and  $\lambda(z)$  are easily solved to have

$$A(z) = \frac{2y}{2 + y[-1 - 2z + \sqrt{1 - 4z}]}, \quad (\text{A.4a})$$

$$\lambda(z) = \frac{1 - 2z + \sqrt{1 - 4z}}{2z}. \quad (\text{A.4b})$$

The singularities in of  $A(z)$  and  $\lambda(z)$  are a simple pole and a branch point singularity which correspond to the bound state and the unbound state singularities respectively given by

$$z_b = \frac{1 - y + \sqrt{y^2 - y}}{y}, \quad \text{and} \quad z_u = \frac{1}{4}, \quad (\text{A.5})$$

where  $z_b$  is the bound state singularity and  $z_u$  is the denatured state singularity. The corresponding free energies obtained as

$$f_b = \log z_b, \quad \text{and} \quad f_u = \log z_f, \quad (\text{A.6})$$

where we set  $\beta = 1$ . The melting here is continuous similar to the necklace model. The critical point is obtained by equating  $z_b$  to  $z_u$  and solving for a  $y$  which is given by  $y_c = \frac{4}{3}$ . Let us now apply a constant unzipping force  $g_u$  in the manner shown in Figure A.1. The corresponding generating function is given by

$$\mathbf{G}(z, g_u) = \sum_{x_1, x_2} e^{g_u(x_1 - x_2)} G(z, x_1, x_2) = \frac{A(z)}{1 - \lambda(z)e^{g_u}}. \quad (\text{A.7})$$

While the unzipping force modifies  $z_u$  it leaves  $z_b$  unchanged indicating that it can not penetrate the bound state. The new unbound state singularity becomes

$$z'_u = \frac{1}{4 \cosh^2 g_u}. \quad (\text{A.8})$$

The phase diagram of unzipping can now be obtained by equating  $z'_u$  to  $z_b$  and solving for a critical

$$g_{uc} = \cosh^{-1} \left( \frac{\sqrt{y}}{2\sqrt{-y + \sqrt{(y-1)y + 1}}} \right). \quad (\text{A.9})$$

For a fixed  $g_u$  the number of bound base pairs becomes abruptly zero at the corresponding critical  $y$  which indicates it's a first order transition (see Figure A.2). The

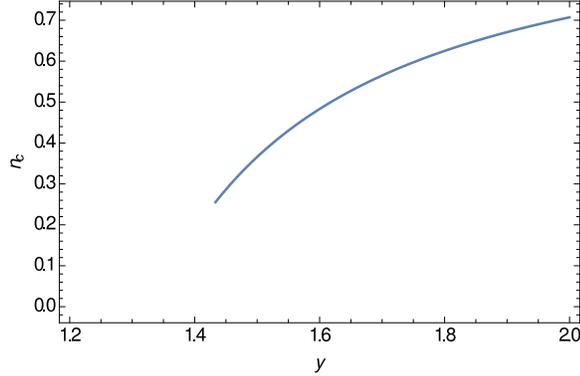


Figure A.2: Fraction of bound base pairs discontinuously goes to zero at  $y = 1.43$  for a fixed  $g_u = 0.1$  indicating a first order transition.

phase diagram of actual force  $tg_{uc}$  vs temperature  $t = 1/\log y$  with  $k_B = 1$  is shown in Figure A.3.

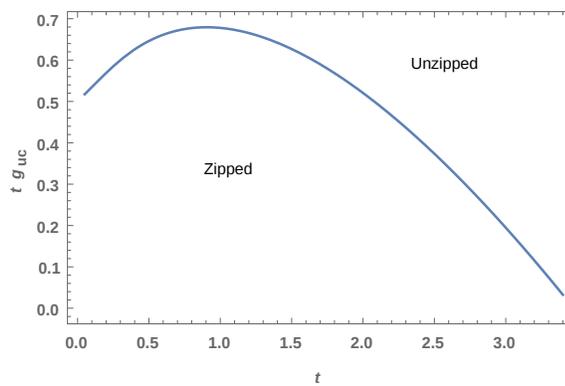


Figure A.3: Re-entrant phase diagram.

# Appendix B

## Necklace Model

For example, let us consider the following model partition functions

$$Z_B(N) = e^{-\beta\epsilon N}, \quad \text{and} \quad (\text{B.1a})$$

$$Z_U(N) \approx \frac{e^{-Nf_u}}{N^\Psi}, \quad (\text{B.1b})$$

where  $\Psi$  is the reunion exponent [115, 116] such that for  $1 < \Psi < 2$

$$G_B(s) = \frac{1}{s + \beta\epsilon}, \quad \text{and} \quad (\text{B.2a})$$

$$G_U(s) \approx \frac{\Gamma(1 - \Psi)}{(s + f_u)^{1 - \Psi}}. \quad (\text{B.2b})$$

Eq. 1.15 can be solved graphically as shown in Figure B.1. If  $1/G_b(s)$  and  $g_2^2 G_U(s)$  are plotted at the same graph as functions of  $s$  the intersection points, if there is any, for different values of  $\beta$  give the duplex state singularities. The free energy of the system is given by the largest singularity. For  $\beta < \beta_c = \frac{|f_u|}{|\epsilon|}$ ,  $s = f_u$  remains the largest singularity and the DNA stays denatured. But for  $\beta > \beta_c$  the new largest singularity is given by the solution of Eq. 1.15 and the DNA becomes bound. As shown in Figure B.1 the transition happens continuously with  $\beta_c = 1/(k_B T_c)$  where  $T_c$  is the critical temperature of the transition.

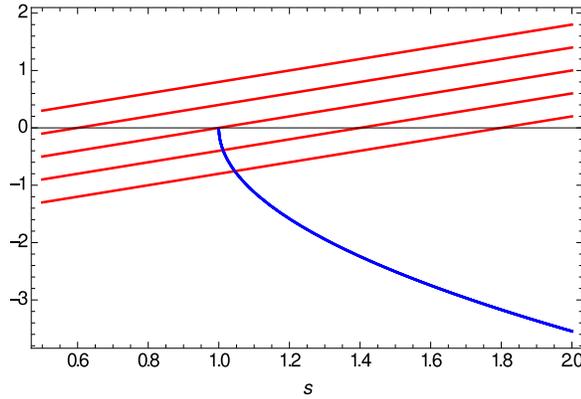


Figure B.1: Graphical solution of the duplex state singularity for  $1 < \Psi < 2$ . The solid blue line is the bubble partition function and the red straight lines are  $1/G_B(s)$  for different temperatures. The parameter values are taken as  $\epsilon = f_u = -g_2 = -1$ .

# Appendix C

## Renormalization Group

Around the critical point of a continuous phase transition a system goes through violent large wavelength fluctuations. The mean field description of a phase transition ignores these fluctuations. This simplification however often costs it vital informations. An efficient theory which can handle these fluctuations is the renormalization group (RG) method [117, 118]. It considers the parameters of a system length scale dependent and aims to find this dependence quantitatively at the large length scale limit. As the length scale is increased some parameters will perish. They are called irrelevant. Some will remain influential which are called relevant parameters. The flow equations of these relevant parameters determine the behavior of the system around the critical point. Here we will consider an example of the application of RG in a two chain polymer system [84]. This simple model captures the basic ideas of the RG method.

Consider a two-chain directed polymer system in  $D = d + 1$  dimension with contact interactions only. The interaction potential is given by a Dirac delta function with strength  $v_0$  which ensures contacts happen only at the same space and length coordinate.  $v_0$  has a canonical dimensionality  $[v_0] \sim L^{d-2}$  where  $L$  is an arbitrary length scale. This problem is exactly solvable through dimensional regularization technique. The flow of the only relevant parameter, also known as the  $\beta$ -function, is given by the following differential equation

$$\beta(u) = L \left( \frac{\partial u}{\partial L} \right) \Big|_{v_0} = \epsilon u - u^2, \quad (\text{C.1})$$

where  $\epsilon = 2 - d$  and  $u$  is the dimensionless renormalized coupling constant related to

$v_0$  through the equation

$$v_0 L^d = u(1 + D_1 u + D_2 u^2 + \dots), \quad (\text{C.2})$$

with  $D_n = (2\pi\epsilon)^{-n}$ . The coefficient of  $u$  in Eq. C.1 comes from the canonical dimension which a mean field theory would have predicted. The  $u^2$  term is solely due to RG. Eq. C.1 has two fixed points, one stable and one unstable, where  $u$  does not change when  $L$  is varied. It is illuminating to solve the flow equation around a fixed point. Let us take the initial value  $u_0 = u^* + \Delta_0$  for a length scale  $L_0$  where  $\Delta_0$  is arbitrarily small. For small  $u$  the linearization of Eq. C.1 gives

$$L \frac{\partial \Delta u}{\partial L} = \epsilon \Delta u, \quad (\text{C.3})$$

which has the solution  $\Delta u = \Delta_0 (L/L_0)^\epsilon$  where  $\Delta u = u - u^*$ . When  $\Delta u$  is finite we can define a length scale  $\xi$  by setting  $\xi \sim L$  which diverges as the the system approaches the fixed point in the limit  $\Delta_0 \rightarrow 0$ . The manner in which  $\xi$  diverges is given by

$$\xi \sim |\Delta_0|^{-\zeta}, \quad \text{with} \quad \zeta = \frac{1}{\epsilon}. \quad (\text{C.4})$$

In Figure C.1 RG flows of  $u$  for different  $\epsilon$  values are shown. Flow to the negative

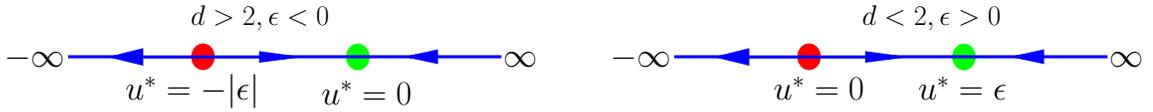


Figure C.1: Renormalization group flows for different  $\epsilon$  values. The arrows show the flow directions with increasing length scale. The red and the green discs are the unstable and the stable fixed points respectively.

infinity indicates a bound state formation. For  $d < 2, \epsilon > 0$  the interaction is relevant for small negative  $u$  values as they flow to the negative infinity. For  $d > 2, \epsilon < 0$  the interaction becomes irrelevant for small  $u$  as there is a negative energy cut off above which there is no bound state as all the points flow to the stable fixed point  $u^* = 0$  which represents the high temperature unbound phase. The unstable fixed point acts as the critical point of the transition. In this way  $\epsilon$  controls the critical behavior of the system.

For this simple system the stable-unstable fixed point description serves us very well. But for more complicated systems, for example, a two-chain directed polymer system with long range attraction [78], the  $\beta$ -function can take a form like

$$\beta(u) = \epsilon u + au^2 + b, \tag{C.5}$$

where  $a$  and  $b$  are constants. For a range of these constant values the  $\beta$ -function does not have any real roots. We reserve the discussion of this interesting problem for Chapter 5.

## Chapter 2

# Temperature Dependence of Entropic Elasticity of DNA

In this chapter we study the temperature dependence of the elastic properties of a model of DNA. There is no intrinsic rigidity in this model and the elasticity comes totally from the entropic effects. We use the transfer matrix method through recursion relations to find the partition function of the system. For analytical solution we use the generating function technique. For numerical calculations we iterate the recursion relations for finite lengths and find the exact partition function. The effect of the unzipping force in the elasticity is also explored. The general case of two unequal forces can always be transformed into a case of unzipping and stretching forces. Since unzipping and stretching of a dsDNA are independent of each other, we are able to generate a general phase diagram in three variables the temperature, the stretching force and the unzipping force.

The organization of this chapter is given below. In Sec. 2.1 first we describe the model qualitatively. In Sec. 2.1.1 we introduce the corresponding recursion relations and define the observables necessary for analysis of the model. The elastic properties of the flexible model are explored in Sec. 2.1.2 where we show a finite discontinuity in the elastic modulus at the melting point. In Sec. 2.1.3 we compare our results with the bubble-less Y-model. The phase diagram in the presence of an unzipping force and a stretching force is obtained in Sec. 2.1.4. The transition is now first order and the elastic modulus shows a  $\delta$ -function peak at the transition point. After a brief

discussion in Sec. 2.2 we conclude in Sec. 2.3.

## 2.1 The Model

We consider each single strand as a directed polymer in a two dimensional square lattice, see Figure 2.1. The base pairing is represented by contact interactions between the monomers which can only occur at the same space and length coordinates. The directedness of the polymers ensures correct base pair bonding. The chains are inextensible [122], of same lengths and attached to each other at the origin. We mimic the DNA melting by the binding-unbinding phase transition in the system. The statistical weight for a contact interaction is  $y$  which is the Boltzmann factor  $\exp(\beta\epsilon)$ ,  $-\epsilon$  being the gain in the energy per contact and  $\beta$  is inverse temperature with the Boltzmann constant set to  $k_B = 1$ . Two chains have a hardcore repulsion that forbids them to cross. The perfectly bound DNA remains as flexible as the single strand so that at any non-zero temperature there is only the emergent entropic elasticity.

To probe the elastic modulus we apply an external space independent mechanical stretching force independently at the end of each strand. If the two forces are in the same direction, the dsDNA is said to be under a stretching force  $g_s$ . On the other hand it will be the unzipping force  $g_u$  if the forces are in the opposite directions. The average extension and the elastic modulus can be obtained from the free energy simply by taking derivatives with respect to the stretching force once and twice respectively. Though the elastic response we are interested in is important for DNA activity, from a phase transition point of view, it is not the primary response function that diverges at a critical point, like the magnetic susceptibility for a ferro-para magnetic transition or the elastic modulus in a liquid-gas transition. One generally associates an exponent  $\gamma$  to such a primary response function, but no such general results can be used here. This justifies the necessity of a detailed study of the rigidity of a melting DNA.

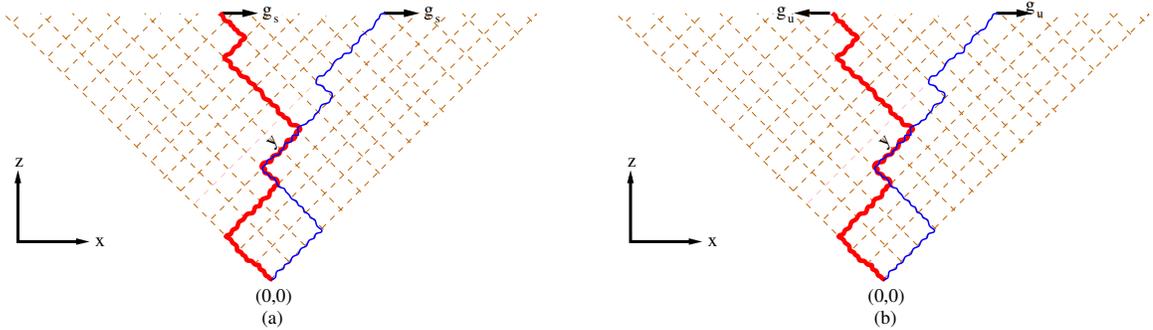


Figure 2.1: Schematic representation of the dsDNA as two directed walks in  $(1 + 1)$  dimension. The polymers can not cross each other. The bound segments can bend left or right freely. The direction in which the forces act are indicated by the arrows. (a) DNA under stretching force, (b) DNA under unzipping force.

### 2.1.1 Recursion Relation And Observables

In the absence of any force the recursion relation followed by this system is given by [123]

$$Z_{n+1}(x_1, x_2) = \sum_{(i,j)=\pm 1} Z_n(x_1 + i, x_2 + j)[1 - (1 - y)\delta_{x_1, x_2}], \quad (2.1)$$

where  $Z_n(x_1, x_2)$  is the canonical partition function of the system of two polymers, each of length  $n$  and the spatial positions of the  $n$ th monomers of polymer 1 and polymer 2 are  $x_1$  and  $x_2$  respectively. For a given monomer number if  $x_1$  becomes equal to  $x_2$  then there is a contact. Each contact has its Boltzmann weight  $y = e^{\beta\epsilon}$ . We set the initial condition as  $Z_0(0, 0) = y$  such that two strands start from the origin. The non-crossing constraint is implemented by not letting  $x_1$  becoming greater than  $x_2$  ( $x_1 \leq x_2$ ).

We apply constant stretching force  $g_s$  at the each end point of the two strands. The partition function of an  $n$  length DNA in presence of this stretching force is given by

$$Z(g_s) = \sum_{x_1, x_2} Z_n(x_1, x_2)e^{g_s(x_1 + x_2)}, \quad (2.2)$$

where the sum is over all the allowed values of  $x_1$  and  $x_2$ .

The elastic response of the system under a stretching force can be quantified through the average extension ( $x$ ) and the elastic modulus ( $\kappa$ ). We define them in

the following way

$$x = \frac{\partial f}{\partial g_s} = \frac{1}{N} \frac{\partial \ln Z(g_s)}{\partial g_s}, \quad \text{and} \quad \kappa = \frac{\partial x}{\partial g_s}, \quad (2.3)$$

where  $f = \beta F$  is the free energy of the system scaled by  $\beta$ , and  $N$  is the length of the strands ( $N \rightarrow \infty$ ). Using Eq. (2.2) we can rewrite them as the following

$$x = \frac{\sum_{x_1, x_2} Z_N(x_1, x_2) e^{g_s(x_1+x_2)} (x_1 + x_2)}{N \sum_{x_1, x_2} Z_N(x_1, x_2) e^{g_s(x_1+x_2)}}, \quad (2.4a)$$

$$\kappa = \frac{\sum_{x_1, x_2} Z_N(x_1, x_2) e^{g_s(x_1+x_2)} (x_1 + x_2)^2}{N^2 \sum_{x_1, x_2} Z_N(x_1, x_2) e^{g_s(x_1+x_2)}} - x^2. \quad (2.4b)$$

An inspection of Eq. (2.4a),(2.4b) shows that  $x$  is related to the average vectorial position of the center of mass of the end points of the two strands of length  $n$  under a stretching force  $g_s$  and, as expected,  $\kappa$  is related to the fluctuation of  $x$ . If  $x_1, x_2$  is uncorrelated, then  $\kappa$  is the sum of the individual elastic constants. This will be the case in the unbound phase of dsDNA. According to this definition for a given force the bigger the value of  $\kappa$  the larger is the flexibility of the dsDNA. Two other important quantities are the average number of contacts between two strands ( $n_c$ ) and it's fluctuation ( $C_c$ ). Two extreme values of  $n_c$ , zero and one, represent the unbound and the bound states respectively. As  $y$  is a temperature like variable one can derive the specific heat of the system from  $C_c$ . We call it specific heat for brevity. These are defined as

$$n_c = \frac{y}{N} \frac{\partial F}{\partial y} \quad \text{and} \quad C_c = y \frac{\partial n_c}{\partial y}. \quad (2.5)$$

We follow these definitions in the rest of this paper.

## 2.1.2 Elastic Response Under a Stretching Force

### Generating Function and The Free Energy

By employing the generating function technique the recursion relation Eq. (2.1) can be exactly solved. We define

$$G(z, x_1, x_2) = \sum_{n=0}^{\infty} z^n Z_n(x_1, x_2). \quad (2.6)$$

By doing this we are going to the grand-canonical ensemble from the canonical ensemble. By multiplying both sides of Eq. (2.1) by  $z^n$  and then summing over  $n$  we

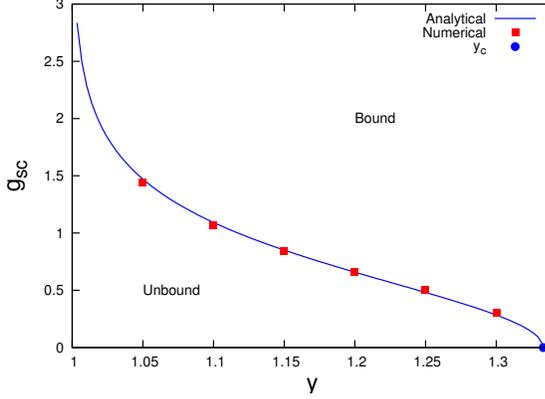


Figure 2.2: Flexible Model : The phase diagram of the system under a stretching force. The phase boundary separates the bound phase from the unbound phase. In the region  $4/3 > y > 1$  there exists a critical  $g_{sc}$  for every value of  $y$ . Above  $y = 4/3$  the system is by default in the bound state. Red squares are the numerically obtained critical points and solid blue line is the analytical curve.

get two independent equations. One for nonzero unequal values of  $x_1$  and  $x_2$ , and another for  $x_1 = x_2 = 0$ . These are given by

$$\frac{1}{z}G(z, x_1, x_2) = \sum_{(i,j)=\pm 1} G(z, x_1 + i, x_2 + j), \quad (2.7a)$$

$$\frac{1}{yz}G(z, 0, 0) = \frac{1}{z} + \sum_{(i,j)=\pm 1} G(z, i, j). \quad (2.7b)$$

The free energy per unit length of the DNA is determined by the singularity of  $G(z, x_1, x_2)$  closest to the origin in the complex  $z$ -plane.

Assuming a power law form for  $G(z, x_1, x_2)$  with respect to relative position coordinate we make the ansatz

$$G(z, x_1, x_2) = A(g_s, z)\lambda(g_s, z)^{(x_1-x_2)/2}e^{g_s(x_1+x_2)}, \quad (2.8)$$

where  $A$  and  $\lambda$  are functions of  $z$  and independent of position coordinates. Eq. (2.8) generalizes the ansatz of Ref. [123]. Using the ansatz Eq. (2.8) into Eq. (2.7a) and Eq. (2.7b) two unknowns  $A(g_s, z)$  and  $\lambda(g_s, z)$  can easily be solved. Their forms are

given by

$$A(g_s, z) = \frac{-1/(2z)}{\cosh(2g_s) - \sqrt{\left(\cosh(2g_s) - \frac{1}{2z}\right)^2 - 1 + \frac{(y-2)}{2yz}}}, \quad (2.9a)$$

$$\lambda(g_s, z) = \sqrt{\left(\cosh(2g_s) - \frac{1}{2z}\right)^2 - 1 + \frac{1}{2z} - \cosh(2g_s)}. \quad (2.9b)$$

The free energies of the different phases of the system are obtained from the singularities of  $G(g_s, z)$ . The singularity  $z_b$  of  $A(g_s, z)$  corresponds to the bound state free energy and the branch point singularity  $z_f$  of  $\lambda(g_s, z)$  gives the free energy of the unbound state. They are calculated as

$$z_b(y, g_s) = \frac{y-1}{y \operatorname{sech}(2g_s)} \left[ \sqrt{\frac{\operatorname{sech}^2(2g_s)}{y-1} + 1} - 1 \right], \quad (2.10a)$$

$$z_f(y, g_s) = \frac{\operatorname{sech}^2(g_s)}{4}. \quad (2.10b)$$

The difference in the force term can be understood by looking at the low energy excitations. In the case of the free chains a force  $g_s$  flips a bond interchanging the energies  $\pm g_s$ . This gives the  $\operatorname{sech}^2(g_s)$  term. In the bound state, with coincident end points, a bound bond gets flipped under a force  $2g_s$ , yielding the  $\operatorname{sech}^2(2g_s)$  term. From here onwards we suppress  $y$  and  $g_s$  as arguments for notational simplification and show them whenever necessary. The corresponding free energies per unit length are given by

$$f_b = \ln z_b, \quad (2.11a)$$

$$f_f = \ln z_f. \quad (2.11b)$$

There are two parameters in this formulation,  $y$  and  $g_s$ . The singularities move when these two parameters are changed. Consider a situation where the system is in the bound state with free energy given by  $f_b$ . Now, we can vary the parameters in such a way that the unbound state singularity  $z_f$  crosses  $z_b$  and becomes closest to the origin. In this situation the free energy of the system becomes  $f_f$ . The crossing of the singularities defines the transition point from bound to unbound by the force at

$$g_{sc} = \frac{1}{2} \cosh^{-1} \left( \frac{2-y}{2(y-1)} \right). \quad (2.12)$$

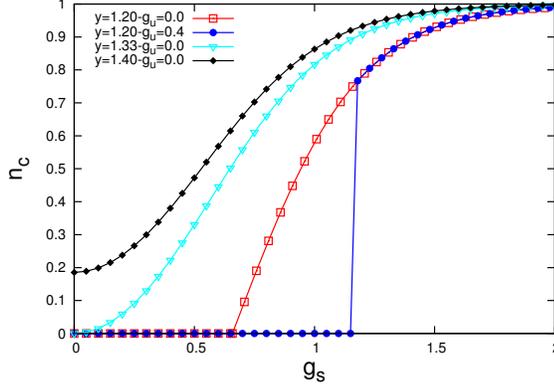


Figure 2.3: Flexible Model : The variation of  $n_c$  with  $g_s$  for different values of  $g_u$  and  $y$ . A non-zero  $n_c$  indicates a bound state. For  $g_u = 0$ , there is no unbound state for  $y \geq 4/3$ . The cyan line with triangles is for the melting point  $y = y_c$ , and shows the quadratic dependence on  $g_s$ . The black curve with filled diamonds is for  $y = 1.4 > y_c$ .

The phase diagram in the  $y$ - $g_{sc}$  plane is shown in Figure 2.2. The phase boundary has the following limiting forms

$$g_{sc} \sim \sqrt{y_c - y} \quad \text{for } y \rightarrow y_c^-, \quad \text{and} \quad (2.13a)$$

$$g_{sc} \sim -\frac{\ln(y-1)}{\ln y} \quad \text{for } y \rightarrow 1^+. \quad (2.13b)$$

## Results and Discussions

**Long Chain Limit :** When the two strands are in the unbound state, they come closer and form contacts in the influence of the stretching force. In this way an unbound state becomes a bound state above the critical stretching force. Figure 2.3 shows how  $n_c$  becomes non-zero before saturating at one as the stretching force crosses a critical value for a  $y < y_c$ . This shows that the transition is continuous. It is already known that at the point  $(4/3, 0)$  in the phase diagram the system goes through a second order phase transition. Beyond this critical point the system always remains in the bound state, thus excluding any other possibilities of phase transition. The only effect of the stretching force there is to influence the bubble statistics. The

asymptotic behavior of  $n_c$  is given by

$$n_c \approx \begin{cases} \frac{9}{2}(g - g_{sc})\sqrt{y - y_c}, & \text{for } y \rightarrow y_c+, g_s \rightarrow g_{sc}+, \\ \frac{3}{2} g_s^2, & \text{for } y = y_c, g_s \rightarrow 0, \\ \frac{y-2+\sqrt{y(y-1)}}{2(y-1)} + \frac{g_s^2}{\sqrt{y(y-1)}}, & \text{for } y > y_c, g_s \rightarrow 0, \\ \frac{27}{8} (y - y_c), & \text{for } y \rightarrow y_c+, g_s = 0. \end{cases} \quad (2.14)$$

That this is a second order phase transition can be corroborated by examining the average extension of the center of mass due to the application of the stretching force and the elastic modulus. They are calculated using the definitions of Eq. (2.4a),(2.4b). Figure 2.4 shows how the average extension changes continuously as the system crosses

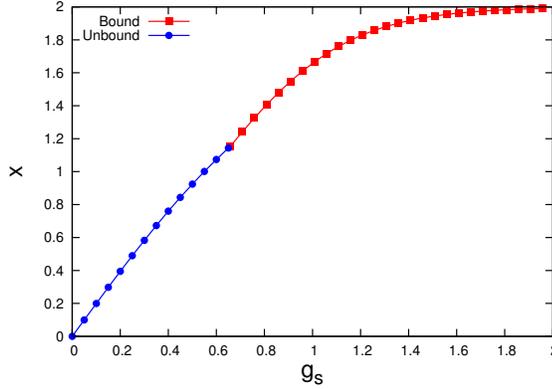


Figure 2.4: Flexible Model : Plot of the average extension as a function of the stretching force with a fixed  $y = 1.20$ . The average extension varies continuously around the critical point.

the critical stretching force. For a zero force  $x$  is zero, consistent with the Gaussian chain behavior while the fully stretched state under a large force has  $x = 2$ . The slope discontinuity at the transition point  $g_{sc}$  of Eq. (2.12) gives rise to a jump in the elastic constant as shown in Figure 2.5. To be noted here is that there is no pre-transitional signature on either side of the transition. However for a finite size system the scenario is different. All the other curves in Figure 2.5 except the analytical curve are for different finite sizes of the system. In the unbound phase each strand has the equation of state  $x = \tanh(g_s)$  so that total  $x = 2 \tanh(g_s)$ . This is the  $g_s < g_{sc}$  branch. The corresponding entropic elastic constant is  $\kappa = 2 \operatorname{sech}^2(g_s)$ . The

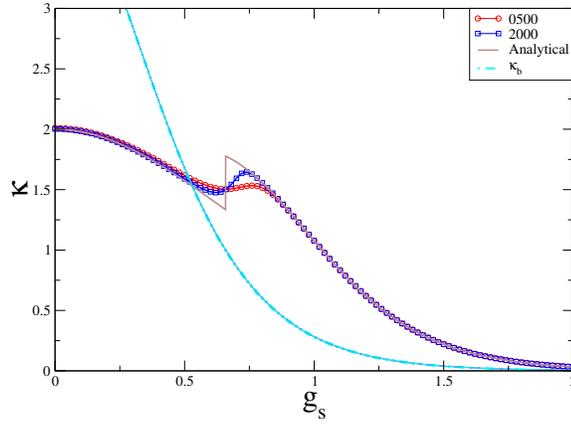


Figure 2.5: Flexible Model : Elastic modulus curve for  $y = 1.20$ . The solid brown line is the analytically obtained curve. The dashed cyan line is for  $\kappa_b = 4 \operatorname{sech}^2(2g_s)$  Eq. (2.18a) for the case with no bubbles. Other curves are the plots of  $\kappa$  for different system lengths.

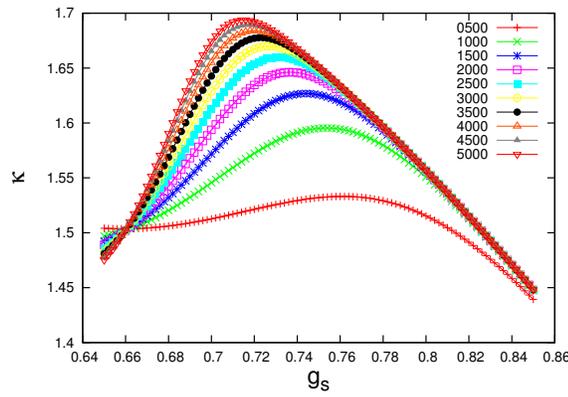


Figure 2.6: Flexible Model : Magnified version of 2.5 around the same crossing point of all the curves. The chain length for each curve is given in the legend.

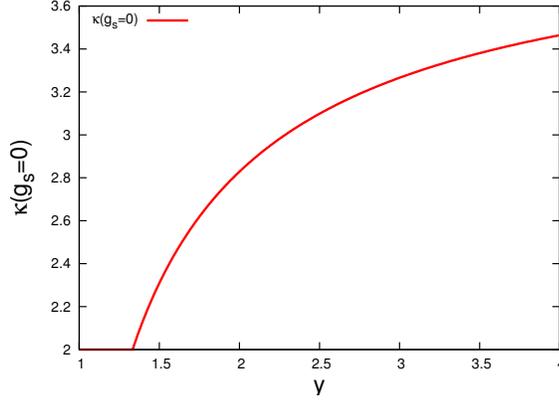


Figure 2.7: Flexible Model : Elastic modulus as a function of  $y$  for zero stretching force.

completely bound state, in absence of any bubbles, should have a similar equation of state with an elastic constant of purely entropic origin given by  $\kappa = 4 \operatorname{sech}^2(2g_s)$ . But the bubbles give an extra contribution. The exact form of the elastic constant can be determined for a few special cases. The  $y$  dependence of the zero force  $\kappa$  is given by (see Figure 2.7)

$$\kappa(g_s = 0) = \begin{cases} 2, & \text{for } y < y_c, \\ 4\sqrt{\frac{y-1}{y}}, & \text{for } y > y_c. \end{cases} \quad (2.15)$$

The elastic constant as a function of force at the melting point  $y = y_c$  is

$$\kappa(y = y_c) = \frac{64w(w+1)}{(w^2 + 14w + 1)^{3/2}}, \quad \text{with } w = e^{4g_s}. \quad (2.16)$$

The behavior of  $\kappa$  for  $y = y_c$  and  $y > y_c$  is shown in Figure 2.8.

**Finite Length DNA :** The contribution of the bubbles become significant in finite length DNA as shown in Figure 2.6. The finite size effects become significant when the length is comparable or smaller than the length of bubble fluctuations. The elastic constant for a finite length DNA is necessarily continuous, devoid of any singularity, but it should evolve into a discontinuous function as the length is made larger. This indicates that shorter chains will show larger deviation from the thermodynamic limit over a range of forces. A finite size scaling from is

$$\kappa = f((g_s - g_{sc})N^{1/\nu}), \quad (2.17)$$

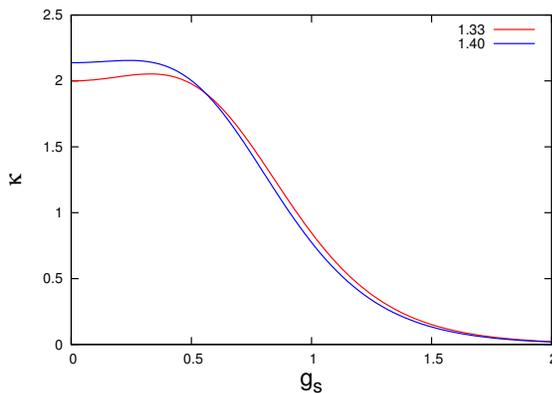


Figure 2.8: Flexible Model : Elastic modulus as a function of  $g_s$  at  $y = 4/3 = y_c$  and  $y = 1.40 > y_c$ .

so that  $\kappa = f(0)$  at  $g_s = g_{sc}$  for all finite  $N$ . Therefore all the finite size curves pass through a common point as shown in the Figure 2.6 which is the critical point. By identifying the common points for other  $y$  values we can now find out the phase diagram numerically. This is shown in the Figure 2.2. All the points in this phase boundary including the thermal melting point ( $y = 4/3, g_{sc} = 0$ ) are second order critical points. The behavior of  $\kappa$  shows that the system is most flexible when it is in the unbound state and under no external force as it has the highest value of  $\kappa$ .

### 2.1.3 Role of The Bubbles

To highlight the importance of the bubbles we compare our results with the Y-model which is similar to the flexible model except the bubble formation is not allowed there [124]. The bound state of this model is the same as the completely bound state of the flexible model and it has a zero force melting point (a first order transition) at  $y_c = 2$ . In the presence of a  $g_s$  the corresponding singularities and elastic contrasts are given by

$$z_b = \frac{1}{2y \cosh 2g_s}, \quad \kappa_b = 4 \operatorname{sech}^2(2g_s), \text{ and} \quad (2.18a)$$

$$z_f = \frac{1}{2 \cosh^2 g_s}, \quad \kappa_f = 2 \operatorname{sech}^2(g_s), \quad (2.18b)$$

where  $\kappa_b$  and  $\kappa_f$  are the bound state and unbound state elastic constants respectively. We obtain the phase boundary by equating  $z_b$  with  $z_f$  and solving for a

$$g_{sc} = \frac{1}{2} \cosh^{-1} \left( \frac{1}{y-1} \right). \quad (2.19)$$

The phase boundary has the similar asymptotics for  $y \rightarrow y_c (= 2)$  and  $y \rightarrow 1$  as in Eqs. (2.13a),(2.13b). In Figure 2.5 we compare  $\kappa_b$  with the flexible model results. It shows that the flexibility of the bound state of the flexible model is mostly due to the bubbles.

### 2.1.4 Elastic Response in Presence of Unzipping Force

It is well known that this model undergoes an unzipping transition under the influence of unzipping force in the absence of stretching force. This unzipping transition is known to be a first order phase transition. The unzipped state consists of two completely separated independent single strands. When the DNA is in the double stranded form the unzipping force tries to unzip it into two single strands. On the other hand, when the DNA is in the unzipped state the stretching force tries to make them bound. Now, if we apply both the forces simultaneously we expect a competition between the opposing effects. In this section we study this problem, again both analytically for infinite system and numerically for finite systems. We use the same definitions of quantities as in Eqs. (2.4a), (2.4b) and Eq. (2.5).

Let us apply a spatially independent unzipping force  $g_u$  at the rear end of the DNA, i.e., the forces act exactly in opposite directions. In the presence of the stretching force  $g_s$  the generating function is given by  $G(g_s, z) = A(g_s, z)\lambda(g_s, z)^{(x_1-x_2)/2}e^{g_s(x_1+x_2)}$  where  $A(g_s, z)$  and  $\lambda(g_s, z)$  are given by Eq. (2.9a) and Eq. (2.9b). The generating function in the presence of both the forces is given by

$$\mathcal{G}(g_s, g_u, z) = \sum_{x_1, x_2} G(g_s, z) e^{g_u(x_1-x_2)}. \quad (2.20)$$

So the bound state singularity remains the same as  $z_b$ , consistent with the hypothesis of non-penetration of forces in the bound state [19], but the unbound state singularity is now given by the solution of the equation  $e^{2g_u} = \lambda(g_s, z)$ . This equation is easily obtained by performing the summation over  $x_1$  and  $x_2$  in Eq. (2.20). Solving this

equation for  $z$  we find the unbound state singularity  $z_{fu}$  is given by

$$z_{fu} = \frac{1}{2 [\cosh(2g_s) + \cosh(2g_u)]}, \quad (2.21)$$

which corresponds to the partition function of the two chains under forces  $g_s + g_u$  and  $g_s - g_u$ , namely  $4 \cosh(g_s + g_u) \cosh(g_s - g_u)$ . For  $g_u = 0$ , the corresponding singularity matches with Eq. (2.10b). The transition as before is given by the crossing of the singularities at

$$g_{uc} = \frac{1}{2} \cosh^{-1} \left( \frac{1}{2z_b} - \cosh(2g_s) \right). \quad (2.22)$$

This expression reduces to the known unzipping line [124] for  $g_s = 0$  and Eq. (2.12) for  $g_u = 0$ .

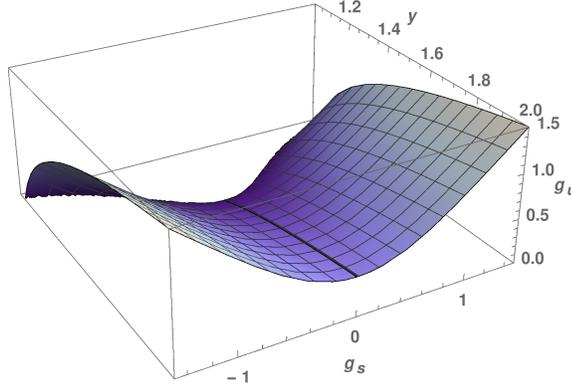


Figure 2.9: Flexible Model : Three dimensional phase diagram of the system in the presence of both stretching and unzipping forces. All the point on the surface except the curve for  $g_{uc} = 0$  represents first order phase transition points. The unzipping line for  $g_s = 0$  is shown by the thick black line.

### Complete Phase Diagram With Unzipping Force

After the introduction of the unzipping force we now have three control parameters. Changing any one of them keeping the other two fixed one can induce a phase transition in the system. The transition points of the system will naturally be distributed on a surface in the  $y$ - $g_s$ - $g_u$  space given by Eq. (2.22). In Figure 2.9 we plot this function. The critical curve for  $g_{uc} = 0$  in the surface is a second order line. Around  $g_s = 0$ ,  $g_{uc}(g_s) = g_{uc}(0) + ag_s^2 + \dots$ , so that the unzipping line for  $g_s = 0$  lies along the

locus of the local minima on the surface. The first order surface ends on the  $g_u = 0$  plane in a critical line that contains the usual melting point at  $y_c(g_s = g_u = 0)$ . Except the critical line all the other possible lines in the surface are first order lines. To show that there is indeed a first order transition we plot  $n_c$  as a function of  $g_s$  in Figure 2.3 with  $g_u$  and  $y$  are being kept fixed.

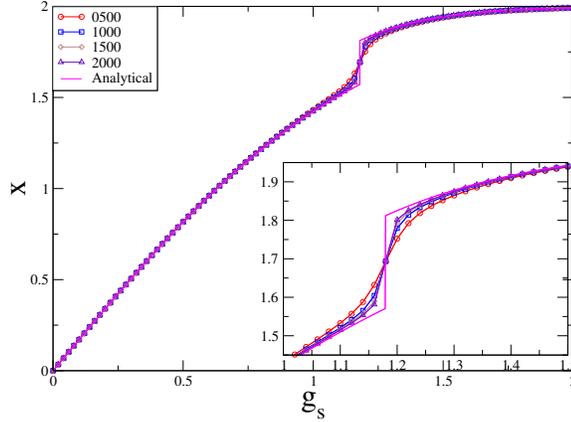


Figure 2.10: Flexible Model :  $x$  vs  $g_s$  plot with  $g_u = 0.4$  and  $y = 1.20$ . The inset figure magnify the critical region. The solid magenta line is the analytical curve. All the other curves are for finite system sizes shown in the legend box. The discontinuity at a  $g_{sc}$  indicates a first order transition.

### Elastic Constant

We plot  $x$  vs  $g_s$  in Figure 2.10 and  $\kappa$  vs  $g_s$  in Figure 2.11 keeping  $g_u$  and  $y$  at fixed values. The magenta curves are the analytical ones for infinite system length while all other curves are for finite system sizes which gradually matches with the analytical curve as  $N$  becomes larger.  $x$  shows a finite discontinuity at a critical  $g_{sc} = 1.18$ . The analytical curve for  $\kappa$  has a  $\delta$ -function peak at  $g_{sc}$  which is not shown in Figure 2.11. The uniform increase of the peak height with increasing system size in  $\kappa$  at the critical point is the signature of the delta peak. Below we list various useful limiting values of  $x$  and  $\kappa$ .

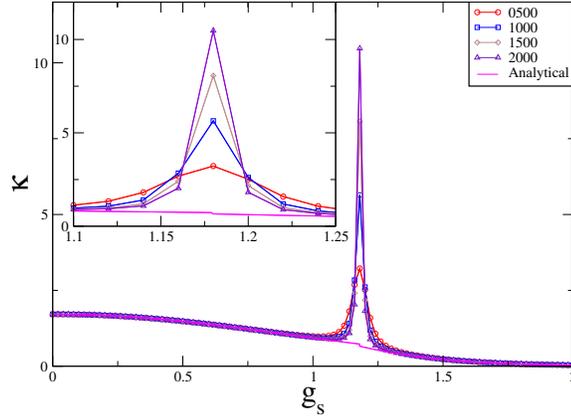


Figure 2.11: Flexible Model :  $\kappa$  vs  $g_s$  plot with  $g_u = 0.4$  and  $y = 1.20$ . The inset figure magnifies the critical region. The solid magenta lines are the analytical curves. All the other curves are for finite system sizes shown in the legend box. The peak height increases proportionally with  $N$  signaling a  $\delta$ -function peak which is not shown in the analytical curve.

1. For  $g_s \rightarrow 0, g_u > g_{uc}(y)$ ,

$$x \approx 2 \operatorname{sech}^2(g_u) g_s, \quad (2.23a)$$

$$\kappa \approx 2 \operatorname{sech}^2(g_u) + 2 \frac{\cosh(g_u) - 2}{\cosh^4(g_u)} g_s^2. \quad (2.23b)$$

2. For  $g_s \rightarrow 0, g_u < g_{uc}(y)$ ,

$$x \approx 4 \sqrt{\frac{y-1}{y}} g_s, \quad (2.24a)$$

$$\kappa \approx 4 \sqrt{\frac{y-1}{y}} + \frac{8(3-2y)\sqrt{y-1}}{y^{3/2}} g_s^2. \quad (2.24b)$$

3. For  $g_s \rightarrow g_{sc}^-, y = 1.2, g_u = 0.4$ ,

$$x \approx 1.57107 + 0.73049(g_s - g_{sc}), \quad (2.25a)$$

$$\kappa \approx 0.73049 - 1.03646(g_s - g_{sc}). \quad (2.25b)$$

4. For  $g_s \rightarrow g_{sc}^+, y = 1.2, g_u = 0.4$ ,

$$x \approx 1.81211 + 0.66067(g_s - g_{sc}), \quad (2.26a)$$

$$\kappa \approx 0.66067 - 2.01486(g_s - g_{sc}). \quad (2.26b)$$

For an infinite system the transition occurs suddenly at a single point. On the other hand, for a finite size system the effect of the transition remains relevant for a domain of  $g_s$  values containing  $g_{sc}$  beyond the scaling regime.

## 2.2 Discussion

There are a few points which we feel need to be clarified in some more detail. (a) As the free ends of the two strands are stretched more and more with increasing forces in the same direction they are bound to come closer due to their equal lengths and the starting ends being attached to each other. This coming closer together increases the possibility of forming a bound base pair by gaining energy. (b) The transition in the presence of the unzipping force is definitely an unzipping transition. This is true even if the  $g_u$  is very small. (c)  $\kappa$  of the flexible model is sensitive to the changes in  $g_u$ ,  $g_s$  and  $y$ . Elasticity is entropic in nature which emerges from collective behavior. (d) The single molecule DNA experimental set-up in which stretching force is achieved by placing the DNA in a directional flow can be a testing platform of our models. (e) In the nanopore sequencing technique, a dsDNA is unzipped and a single strand is passed through a nanopore [125]. Other than that during bacterial conjugation or infection of a cell by a virus the DNA comes under similar geometry. Our study may be relevant in these cases. (f) The single molecule DNA unzipping experiments are normally done at the room temperature. In Eq. (2.22) we provided a phase boundary which not only depends on the temperature and the unzipping force, but also on the stretching force. It remains a challenge to generate this phase boundary experimentally with temperature as a variable in single molecule experiments.

## 2.3 Conclusion

We studied the effect of melting of DNA in its elasticity using a  $(1 + 1)$  dimensional model by employing exact numerical and analytical methods. Under the stretching force DNA goes through a second order binding-unbinding phase transition. The dependence of DNA flexibility on the stretching force, the unzipping force and the temperature has also been discussed. In the presence of both the forces the system

goes through a first order unzipping transition. The complete phase diagram in the  $y-g_s-g_u$  plane is obtained. Though the binding-unbinding transition is very sharp for infinite length system the transition point can influence the elastic behavior of DNA for a broad region of parameter values when the system length is finite. Consequently the elastic response of small length DNA, as used extensively in experiments, has to be widely different from that of long chain DNA. Furthermore, though DNA is a very long molecule it can locally melt depending on the environment it is in. Thus our study will help understanding the importance of these locally melted regions of smaller lengths in determining the elastic properties of the DNA as a whole.

## Chapter 3

# Bubble Mediated Softening of Rigid DNA

The model introduced in the previous chapter uses hard core repulsion to induce melting. In this chapter we first show that the melting can be achieved by making the bound segments semi-rigid against bending and without introducing the hard core repulsion. Then we consider a model, the rigid model, where the bound states are completely rigid and flexibility comes only through the flexible patches provided by the bubbles. For this model we quantitatively relate the extension and the elastic modulus to the bubble number and the bubble number fluctuations respectively.

The organization of this chapter is the following. In Sec. 3.1 we demonstrate a semi-rigidity induced melting transition. The rigid model is introduced in Sec. 3.2. After listing its governing recursion relations and defining required observables specific to this model we obtain its thermal melting, a continuous transition, point in Sec. 3.2.1. In Sec. 3.2.2 we show that the corresponding elastic modulus becomes anomalous around the melting point as it surpasses the unbound state elastic modulus. The roles played by the bubbles are shown quantitatively in Sec. 3.2.3. Only the stretching force is considered for the rigid model. After a brief discussion about the relevance of our results in Sec. 3.3, we summarize and conclude in Sec. 3.4.

### 3.1 Bias Induced Melting

(1+1) dimensional directed polymer models of DNA introduce the non-crossing constraint to implement the mutual hard core repulsion between the two strands. From the perspective of flexibility the bound state of those models are not distinguishable from the unbound state. A semi-rigidity can be introduced by imposing an energy cost against the bending of a bound segment. For that we have to include another microscopic parameter,  $b$ , other than the contact interaction  $y$ , see Figure 3.1(a). But the advantage of this approach is that we can now easily modulate the flexibility of the bound state and relax the non-crossing constraint in a controlled manner. By

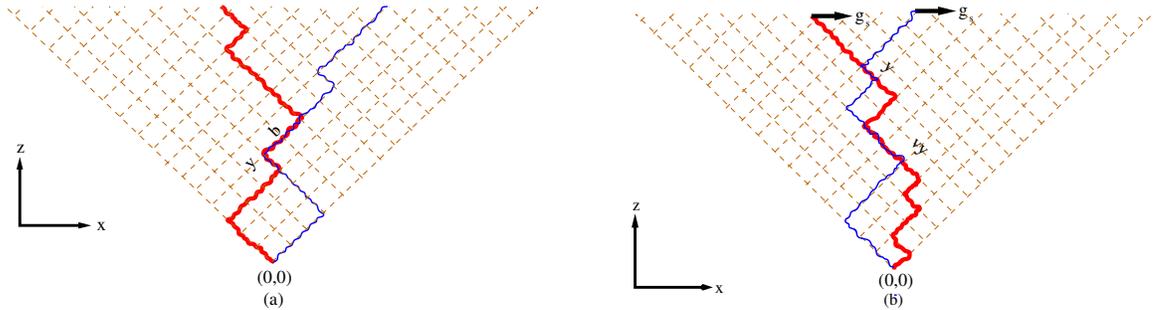


Figure 3.1: Schematic representation of the dsDNA as two directed walks in (1 + 1) dimension. The polymers can cross each other. (a) The options for the bound segments can be restricted by introducing a statistical weight  $b$ . In this figure  $b$  is associated to the right degree of freedom. (b) Rigid Model : The bound segments can not bend to the right and they are at least 2 bonds long.  $v$  is the weight associated to the bubble opening or closure.

fulfilling the requirement that for  $b = 1$  we must get back the good old free Gaussian chain the recursion relation of the system is given by

$$Z_{n+1}(x_1, x_1) = y[Z_n(x_1 + 1, x_1 + 1) + Z_n(x_1 + 1, x_1 - 1) + Z_n(x_1 - 1, x_1 + 1) + bZ_n(x_1 - 1, x_1 - 1)], \quad (3.1a)$$

$$Z_{n+1}(x_1, x_2) = [Z_n(x_1 + 1, x_2 + 1) + Z_n(x_1 + 1, x_2 - 1) + Z_n(x_1 - 1, x_2 + 1) + Z_n(x_1 - 1, x_2 - 1)]$$

for  $x_1 \neq x_2$ , (3.1b)

where  $Z_n(x_1, x_2)$  is the partition function for the system length  $n$ . We set the initial condition as  $Z_0(0, 0) = y$  such that two strands starts from the origin. The microscopic parameter  $b$  is a Boltzmann weight which controls the possibility of the two polymers both going to the right hand side while being in the bound state. All the notations and definitions of the observables which are used in this model are same as the previous chapter.  $z$ -transforming Eqs. (3.1a),(3.1b) using the initial condition we get two independent equations for  $x_1 \neq x_2 \neq 0$  and  $x_1 = x_2 = 0$ . They are given by

$$\begin{aligned} \frac{1}{z}G(z, x_1, x_2) &= G(z, x_1 + 1, x_2 + 1) + G(z, x_1 - 1, x_2 + 1) \\ &\quad + G(z, x_1 + 1, x_2 - 1) + G(z, x_1 - 1, x_2 - 1), \end{aligned} \quad (3.2a)$$

$$\begin{aligned} \frac{1}{yz}G(z, 0, 0) &= \frac{1}{z} + G(z, 1, 1) + G(z, -1, 1) \\ &\quad + G(z, 1, -1) + bG(z, -1, -1). \end{aligned} \quad (3.2b)$$

To solve these independent equations we make the ansatz for the generating function as  $G(z, x_1, x_2) = A(z)\lambda(z)^{\text{Abs}[(x_1-x_2)/2]}$ . Substituting this ansatz in Eq. (3.2a) and Eq. (3.2b) and solving for  $A$  and  $\lambda$  we get

$$A = \frac{1}{-bz + \frac{1}{y} + z + \sqrt{1 - 4z} - 1}, \quad (3.3a)$$

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

The bound state and the unbound state singularities are given by

$$z_b = \frac{b - 1 - y(b + 1) + \sqrt{y}\sqrt{(b + 1)^2y - 4b + 4}}{(b - 1)^2y}, \quad (3.4a)$$

$$z_f = \frac{1}{4}. \quad (3.4b)$$

From these singularities all the other relevant quantities can be derived. Figure 3.2 shows how  $n_c$  varies with  $y$ . At high enough  $y$ ,  $n_c$  saturates to its maximum value, 1. In Figure 3.3 we plot  $C_c$  against  $y$ . In both Figure 3.2 and Figure 3.3 we have also shown the corresponding numerical data for different finite system sizes.  $C_c$  obeys a finite size scaling form

$$C_c = g((y - y_c)N^{1/\delta}) \quad (3.5)$$

such that at  $y = y_c = 1.142$ ,  $C_c(y_c) = g(0)$ . The numerical data is consistent with the analytical results. From these two figures we conclude that the system is going

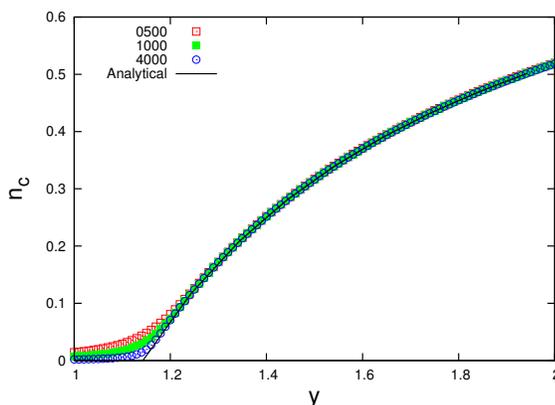


Figure 3.2:  $n_c$  vs  $y$  plot for  $b = 0.5$ .  $n_c$  becomes finite at  $y = 1.142$ . Numerical data is also consistent with the analytical result.

through an usual second order binding-unbinding phase transition. We obtain the critical point of this transition analytically by matching  $z_b$  with  $z_f$ . The critical value

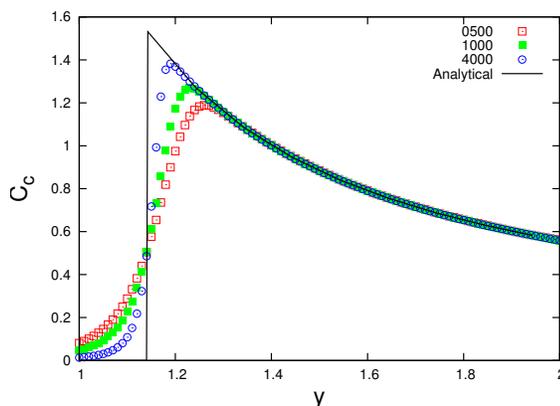


Figure 3.3:  $C_c$  vs  $y$  plot for  $b = 0.5$ . Plot shows a finite discontinuity at  $y = 1.142$ . Numerical data also shows very similar behavior.

of  $y$  is now  $b$  dependent and varies with  $b$  as  $y_c = 4/(3 + b)$ . For  $b = 1$  we have the same recursion relation as that of a system with two Gaussian chains which can freely cross each other and there is no phase transition. In our case also we get  $y_c = 1$ , which means there is no phase transition at finite temperatures. Another interesting limit is when  $b = 0$ ,  $y_c = 4/3$ , the critical point for two Gaussian chains with non-crossing constraint although the recursion relations for these two cases are not the

same. Moreover the  $b$  dependence of  $y_c$  in our model adds extra flexibility in that we can now tune the critical point by tuning  $b$  for a wide range of values.

From the recursion relation it is clear that the parameter  $b$  just modulates one of the four possible contributions in the partition function of the  $n$ th generation from the  $(n-1)$ th generation.  $b$  may be associated with any of the four possible contributions. The case when  $b$  is attached with the  $z(x_1 - 1, x_2 + 1)$  term is of special interest because the  $b \rightarrow 0$  limit is exactly the flexible non-crossing case discussed in the previous chapter. This current case also is exactly solvable through the generating function technique and gives exactly the same  $b$  dependent critical melting point  $y_c = 4/(3+b)$ . So we can now actually modulate the non-crossing constraint through  $b$  and the corresponding critical point as well.

### 3.1.1 Elastic Response

After establishing the melting transition let us now explore the elastic properties for this case. Here, like the flexible case, we again apply a constant stretching force,  $g_s$  at the end of the two chains and use the earlier definitions of  $x$  and  $\kappa$ . We use exact numerical transfer matrix method for the analysis of finite size systems. In Figure 3.4

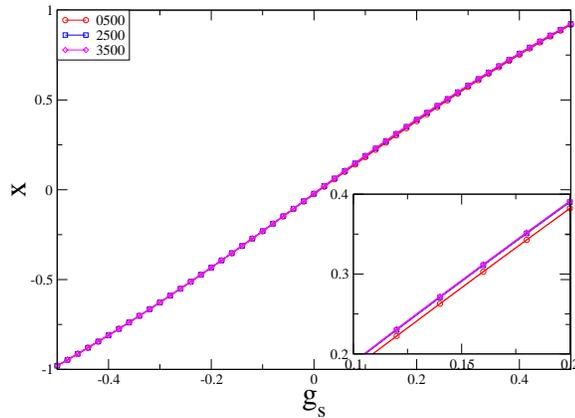


Figure 3.4: Numerical  $x$  vs  $g_s$  plot for  $b = 0.5$  for different system sizes.  $y = 1.20$ . System size dependence of  $x$  is very weak.

and Figure 3.5 we plot the average extension  $x$  and elastic modulus  $\kappa$  respectively for different chain lengths.  $x$  has almost unrecognizable chain length dependence.  $\kappa$  also

shows a small smearing for a range of  $g_s$  values. From these plots we can not conclude whether there is any force induced binding-unbinding transition in this model. To see if there is any transition we study the behaviors of  $n_c$  and  $C_c$ . Figure 3.6 and Figure

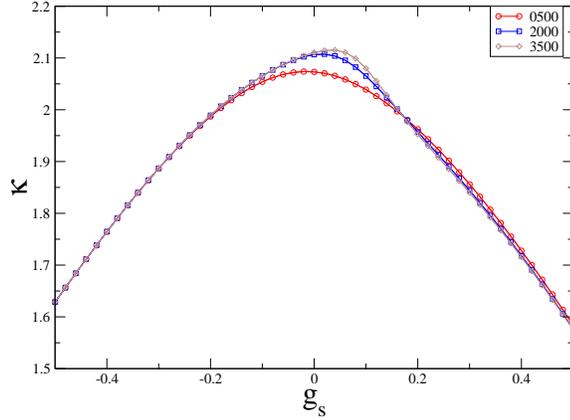


Figure 3.5: Numerical  $\kappa$  vs  $g_s$  plot for  $b = 0.5$  and  $y = 1.20$  for different system sizes. There are very small but identifiable gradually increasing peaks. All the curves has a common crossing point at  $g_s = 0.18$ .

3.7 show how  $n_c$  and  $C_c$  varies with increasing  $g_s$  respectively for fixed fixed  $y = 1.20$  and  $b = 0.5$ . These figures show  $n_c$  and  $C_c$  have considerable system size dependence. As we vary  $g_s$  from  $-2$  to  $2$   $n_c$  tends to zero. Plots of  $C_c$  on the other hand, tend to show a finite discontinuity as the peak heights gradually saturate. We thus conclude that the little smearing in the  $\kappa$  is actually signifies a second order binding-unbinding phase transition. Now the task in hand is to obtain the phase diagram of the system by identifying the critical points. Figure 3.7 shows at  $g_{sc} = 0.18$  all the curves meet. For the other  $y$  values we similarly get one common crossing point each. As discussed earlier these crossing points are nothing but the critical points of the transition. By listing the critical points in this way we plot the  $y$ - $g_{sc}$  phase diagram in Figure 3.8. From the phase diagram we see that the system becomes unbound for  $g_s > g_{sc}$  for a fixed  $y$ . If the force is increased further the bound state will appear again. Let us fix the force at a  $g$  where  $g > g_{sc}$  for fixed  $y$  and  $b$ . The energy per unit length of a completely bound state on the unfavorable direction is given by  $E_1 \approx 2g - 1 - \frac{\ln b}{\ln y}$  and corresponding energy of the two completely stretched but unbound single strands is  $E_2 \approx 2g$ . By comparing  $E_1$  and  $E_2$  it can be concluded that to make the bound state

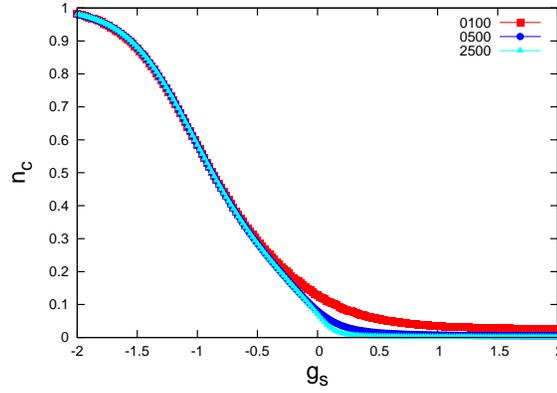


Figure 3.6: Numerical  $n_c$  vs  $g_s$  plot for  $b = 0.5$  and  $y = 1.20$  for different system sizes.  $n_c$  becomes non-zero to zero continuously.

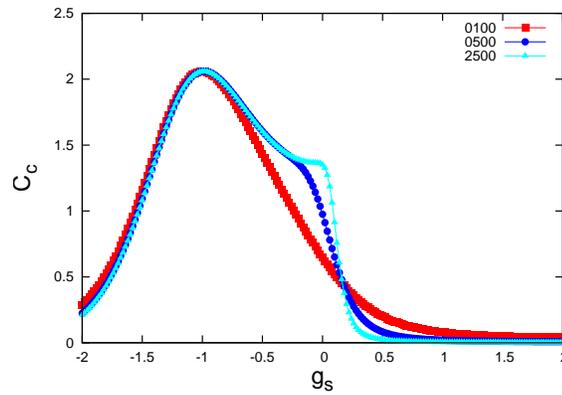


Figure 3.7: Numerical  $C_c$  vs  $g_s$  plot for  $b = 0.5$  and  $y = 1.20$  for different system sizes. All the curves has a common crossing point at  $g_s = 0.18$ .

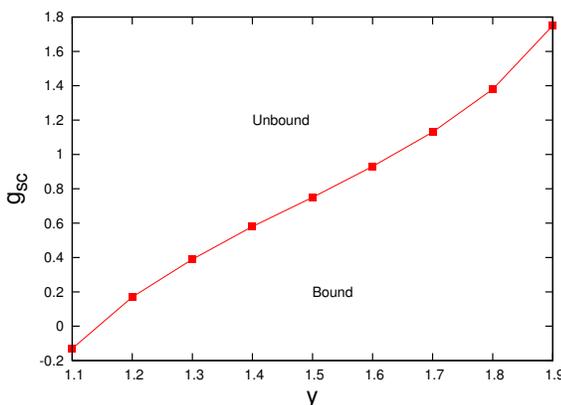


Figure 3.8: Numerical phase diagram in the  $y_c$ - $g_{sc}$  plane for  $b = 0.5$ .

favorable we must increase  $y > 1/b$ . The infinite temperature behavior of the system will be same as the flexible case discussed in the previous chapter because  $b$  does not affect the unbound state.

## 3.2 Rigid Model

Here we customize the previous model to incorporate explicit weights for bubble formation. Doing that in the transfer matrix format is a bit involved. To identify a bubble we need to ensure that an unbound region is attached in between two bound segments. A bound segment is defined as a DNA patch where every base pair is in the bound state and the minimum length it can have is 2. We implement this by putting the constraint that a bound base-pair can form only if another bound base-pair precedes it. So, for every step in the generation of the polymers we need to keep track of its previous step. We introduce an inbuilt rigidity to the dsDNA by putting a bias against the bending towards right for the bound segments. For computational simplicity we here completely switch off the right option. By doing this we are introducing a bias in the propagation of the DNA in favor of one direction. Other than the usual contact weight  $y$  we introduce another Boltzmann weight  $v$  if a bound segment opens to form two single strands or two single strands recombine to form a bound segment, see Figure 3.1(b). The recursion relations which obey these

rules are given by

$$z_n(x_1, x_2) = \begin{cases} y[vz_m(i, l) + vz_m(j, k) + z_m(j, l)], & \text{if } x_1 = x_2 \& n > 0 \\ vz_m(i, l) + z_m(i, k) + z_m(j, k) + z_m(j, l), & \text{if } x_1 - x_2 = 2 \& n = 1 \\ vz_m(j, k) + z_m(i, k) + z_m(i, l) + z_m(j, l), & \text{if } x_1 - x_2 = -2 \& n = 1 \\ vy z_{m-1}(i+1, l+1) + z_m(i, k) + z_m(j, k) + z_m(j, l), & \text{if } x_1 - x_2 = 2 \& n \geq 2 \\ vy z_{m-1}(j+1, k+1) + z_m(i, l) + z_m(i, k) + z_m(j, l), & \text{if } x_1 - x_2 = -2 \& n \geq 2 \\ z_m(i, k) + z_m(i, l) + z_m(j, k) + z_m(j, l), & \text{if } |x_1 - x_2| > 2 \& n > 0 \end{cases}, \quad (3.6)$$

with  $i = x_1 - 1$ ,  $j = x_1 + 1$ ,  $k = x_2 - 1$ ,  $l = x_2 + 1$  and  $m = n - 1$ . The first two steps fix the initial configurations. To fix the configuration at the  $n$ th step we need to keep the informations of not only the  $(n - 1)$ th step but also of the  $(n - 2)$ th step. Using the bubble weight  $v$  we can now count the average number of bubbles ( $n_b$ ) and calculate the average bubble length ( $l_b$ ). They are given by the following formulas

$$n_b = \frac{v^2}{N} \frac{\partial F(y)}{\partial v^2}, \quad C_b = v^2 \frac{\partial n_b}{\partial v^2}, \quad \text{and} \quad l_b = \frac{(N - n_c)}{N n_b}, \quad (3.7)$$

where  $C_b$  describes the fluctuations in  $n_b$ . Once  $Z(x_1, x_2)$  is known, the force dependent partition function can be obtained with the help of Eq. (2.2) and the corresponding elastic constant from Eq. (2.3). Only stretching force  $g_s$  is considered here.

### 3.2.1 Thermal Melting : $g_s = 0$

First let us show that this model goes through a binding-unbinding transition as  $y$  is varied in the absence of any external forces. For the analysis of this section we set  $v = 1$ . Here we use the exact numerical transfer matrix method for finite size systems.

In Figure 3.9 we show how the average number of contacts vary with  $y$ . As the system size is increased one part of the curve gradually touches the  $y$ -axis. And it is also evident that  $n_c$  will saturate at  $n_c = 1$  for appropriately high  $y$ -values. These indicate a binding-unbinding transition. To find out the order of the transition and the corresponding critical value of  $y$ ,  $y_c$ , we plot  $C_c$  vs  $y$  in Figure 3.10.  $C_c$  obeys a

finite size scaling relation similar as Eq. (3.5) which indicates a finite discontinuity. At  $y = 1.18$  all the curves pass through a common point indicating  $y_c = 1.18$ . The finite discontinuity at  $y_c$  establishes that this is a second order phase transition. Note

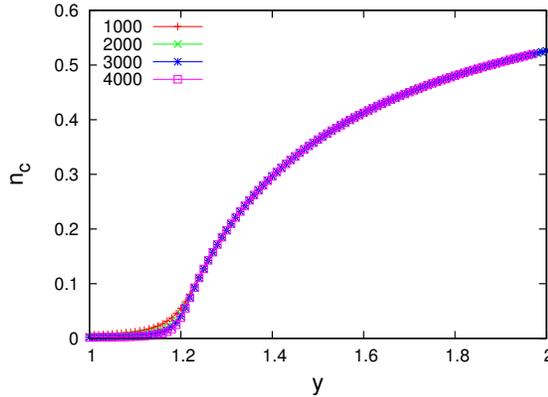


Figure 3.9: Rigid Model :  $n_c$  increases from zero to non-zero values continuously at a  $y > 1$  before approaching the saturation value 1. This indicates a binding-unbinding transition.

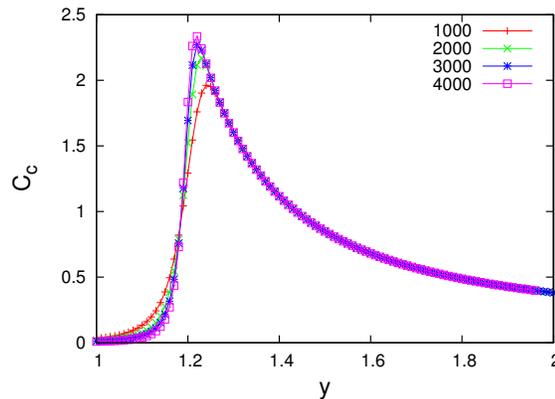


Figure 3.10: Rigid Model :  $C_c$  vs  $y$  plots for different  $N$ . Peak Height increases with increasing  $N$  but eventually saturates creating a finite discontinuity. The discontinuity occurs at  $y = 1.18$  where all the curves meet.

that we have not imposed the non-crossing constraint here. The restriction imposed on the bound state is sufficient to induce a bound-unbound transition. In one spatial dimension the entropy of a system dominates over binding energy which implies that

no ordering here. Imposition of special restrictions to limit the entropy may result in a energy dominated ordered state. The non-crossing constraint in the previous model was doing exactly that by decreasing the total number of configurations. The restriction imposed here in the degrees of freedom of the bound segment is playing the similar role, decreasing the total number of configurations of the DNA.

### 3.2.2 Elastic Response : $g_s \neq 0$

Let us now discuss the elastic properties of this system. As discussed earlier the inherent asymmetry in this model favors extension of DNA in one direction and opposes on the other direction. So under the influence of a spatially independent stretching force the DNA is more flexible in one direction compared to the other direction. As the system is no longer symmetric under  $g_s \leftrightarrow -g_s$  we need to give

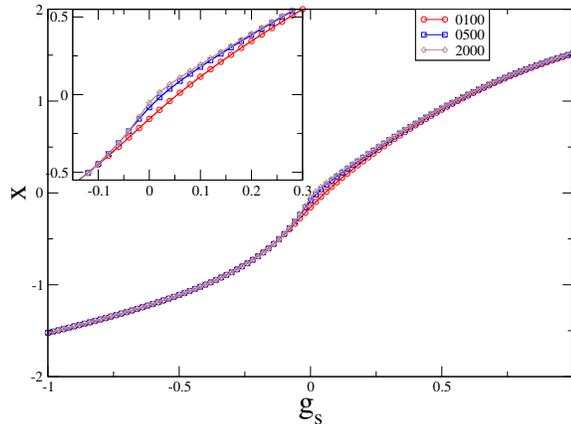


Figure 3.11: Rigid Model : Continuous stretching of the DNA to its full extent by both positive and negative forces.  $y$  value is fixed at 1.20.

attention to the negative values of  $g_s$  too. Figure 3.11 shows  $x$  varies continuously with  $g_s$ , reaching  $\pm 2$  for large positive/negative  $g_s$ . In Figure 3.12 we plot  $\kappa$  vs  $g_s$  keeping  $y$  fixed. The curves shows a peak around  $g_s = 0.02$  which increases in height as  $N$  increases. Maximum of  $\kappa$ ,  $\kappa_{\max}$ , goes to finite value in the  $N \rightarrow \infty$  limit as shown in the inset of Figure 3.12. Inset image in Figure 3.13 shows how  $n_c$  changes as we increase  $g_s$  for  $y = 1.20$ . This indicates a continuous binding-unbinding phase transition. Figure 3.13 shows  $C_c$  has a finite jump at a critical  $g_{sc} = 0.02$  which can

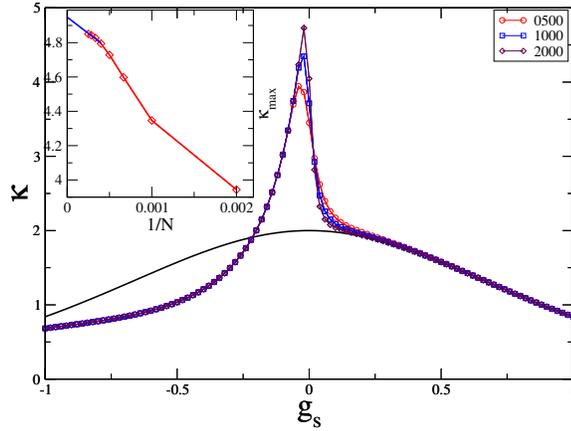


Figure 3.12: Rigid Model :  $\kappa$  shows an increasing peak around  $g_s = 0.02$  with increasing system length  $N$  with a fixed  $y = 1.20$ . Maximum values of  $\kappa$ ,  $\kappa_{\max}$ , are plotted vs  $1/N$  in the inset. A linear fit with the first four points (solid blue line) gives the estimate for  $N \rightarrow \infty$ ,  $\kappa_{\max} = 4.95$ . This indicates there is a finite discontinuity in  $\kappa$ . The solid black line is the plot of the function  $2 \operatorname{sech}^2(g_s)$ , the unbound state elastic constant. For  $g_s > 0.02$ ,  $\kappa$  matches with the unbound state modulus.

be identified as the common point in the peak region through which every finite size curve passes. For  $g_s < g_{sc}$ ,  $\kappa$  shows anomalous behavior as close to the transition point it can reach values which are much greater than the entropic elastic modulus of the unbound state given by  $2 \operatorname{sech}^2(g_s)$  shown in Figure 3.12 as a solid black line. By collecting similar common points for different  $y$  values we draw the numerical phase diagram of the system which is shown in the Figure 3.14. Noticeable here is that the stretching force actually works like an unzipping force because the bound state becomes unbound on application of  $g_s > g_{sc}$ . The reason for this is the following. Due to the bias the bound state formation on the positive  $x$ -axis is unfavorable and the DNA prefers to go to the negative  $x$ -direction. As  $g_s$  is increased it wins over the bias eventually and pulls the DNA towards the positive  $x$ -direction. Because the bound state is forbidden in that direction the bound DNA unzips as a result.

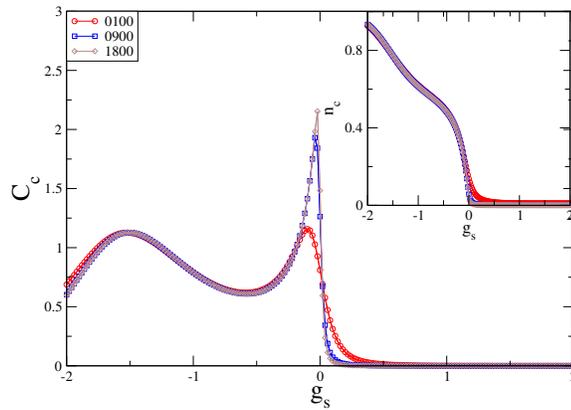


Figure 3.13: Rigid Model :  $y$  value is fixed at 1.20. The inset image shows  $n_c$  decreases continuously from finite values to zero indicating a continuous phase transition. Peak heights in  $C_c$  vs  $g_s$  curves saturate indicating a finite discontinuity. Around  $g_s = 0.02$  all curves meet which is the critical point.

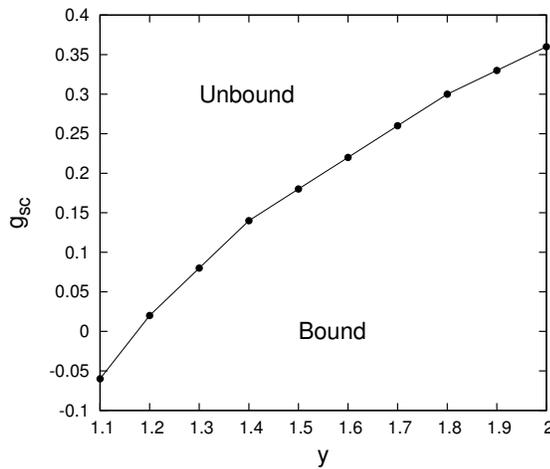


Figure 3.14: Rigid Model : Numerical phase diagram for the binding-unbinding transition.

### 3.2.3 Role of The Bubbles

The flexibility of the bound state comes solely due to the bubbles as the bound segments are absolutely rigid in this model. Figure 3.15 shows that  $n_b$  becomes

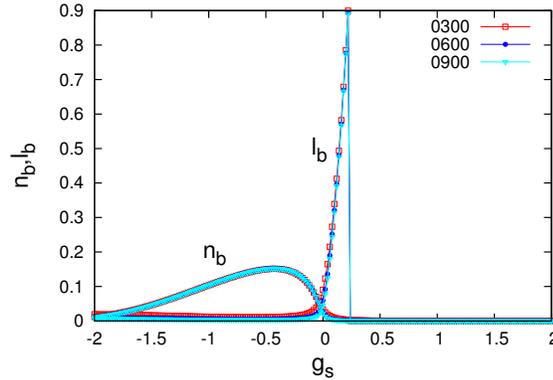


Figure 3.15: Rigid Model :  $n_b$  vs  $g_s$  and  $l_b$  vs  $g_s$  plot for  $y = 1.20$ . As the critical point approaches  $n_b$  becomes very small but  $l_b$  becomes as large as  $N$ .

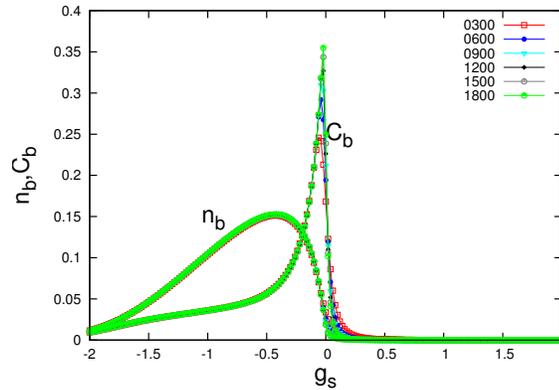


Figure 3.16: Rigid Model :  $n_b$  vs  $g_s$  and  $C_b$  vs  $g_s$  plot for  $y = 1.20$ . Around the critical point  $n_b$  is very small but it's fluctuation  $C_b$  is very large.

very small while  $l_b$  increases to almost equal to  $N$  as we approach the transition point. Here  $l_b \approx (N \text{ or } 0)$  means the DNA is in the unzipped state. The peak in the  $n_b$  curves indicates large number of bubbles but at the same time of very small average length. From these two observations we can now say that as we approach the transition point many small bubbles coalesce together to form large bubbles with decreasing numbers

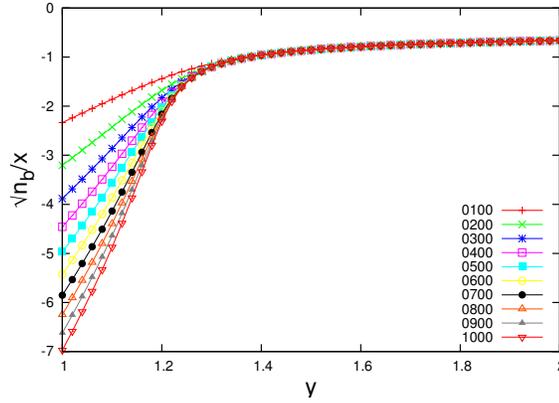


Figure 3.17: Rigid Model :  $\frac{1}{x}\sqrt{n_b}$  vs  $y$  curves for different system lengths collapse to a single master curve in the bound region.  $g_u = g_s = 0$ .  $y_c = 1.18$ .

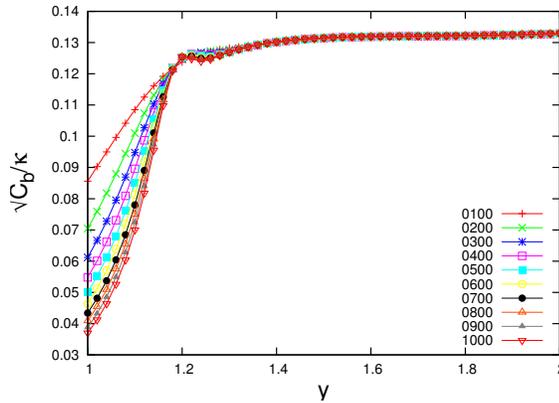


Figure 3.18: Rigid Model :  $\frac{1}{\kappa}\sqrt{C_b}$  vs  $y$  curves for different system lengths collapse to a single master curve in the bound region.  $g_u = g_s = 0$ .  $y_c = 1.18$ .

and eventually  $n_b$  becomes zero when the two strands get completely separated. The fluctuation in  $n_b$ ,  $C_b$ , also becomes large around the critical point which is shown in Figure 3.16. Earlier we have shown that  $\kappa$  of this model behave anomalously and the anomalous behavior occurs at the same region where  $l_b$  and  $C_b$  are the largest.  $\kappa$  in our model is the fluctuation of extension  $x$  by definition and it depends on  $n_b$ . For example near the transition point  $n_b$  is very small and  $x$  is also very small although  $l_b$  is large. This is because in the (1 + 1) dimension  $x$  for a single strand is zero due to its Gaussian nature. But for a bound DNA,  $x \propto \sqrt{n_b}$  as shown in Figure 3.17. It is then expected that the  $\kappa$  will be determined by the fluctuations in  $n_b$ ,  $C_b$ . In Figure 3.18 we plot  $\frac{1}{\kappa}\sqrt{C_b}$  vs  $y$  for different system sizes in the absence of any external force. For the bound region ( $y > 1.18$ ) the curves collapse into a single master curve which is almost  $y$  independent inferring

$$\kappa \approx 7.7\sqrt{C_b}. \quad (3.8)$$

We therefore conclude that  $C_b$  is the important factor in determining the elastic behavior of the system.

### 3.3 Discussion

For the rigid model the DNA unzips for the positive stretching forces. After unzipping if the stretching force is increased further one may expect the strands to become bound again at a large force as they are anchored at the origin. This does not happen here as we have completely forbidden the bound state formation along the positive  $x$  direction when there is no bubble. When the chains are fully stretched there is no room for fluctuation and hence no bubbles. So, even though the elasticity goes to zero the chains are not bound. On the other hand, for small forces there can be bubbles in the bound state. As a result depending on the attraction strength  $y$  we get bound state up to some positive stretching force. As discussed earlier role of small bubbles in DNA flexibility has already been addressed by a number of studies. By considering the role of big bubbles which appear around the melting point we get something puzzling. Our results show that the bound DNA is way more flexible than the unbound DNA around the critical point.

### 3.4 Conclusion

We have shown that the introduction of semi-rigidity in a flexible Gaussian DNA leads to melting transition. By introducing a model with intrinsic rigidity we have studied the role of the bubbles in DNA flexibility for different temperatures. The average bubble length and the average bubble number for our model for different parameter values are also studied. We have shown that the DNA flexibility is related to the bubble number fluctuations. For zero external forces, the extension of the DNA is temperature independent and varies with square root of bubble numbers proportionally while the elastic modulus is also proportional to the square root of the bubble number fluctuation.

## Chapter 4

# Absence of Efimov-DNA Without Denaturation Bubbles

In this chapter we introduce a necklace type model of DNA melting. The two-chain bound segments are considered as rigid rods without any bending. Although they can freely rotate around their starting point. The single chains are modeled as flexible Gaussian chains. This model will be used to represent a duplex DNA in the next chapter. Before that, here we consider the interaction between a rigid bound state and a free single chain. We show for this case Efimov-DNA does not appear.

The outline of the chapter is given below. The model is defined in Sec. 4.1 in real space. It involves the bound state of dsDNA as a rigid rod and the third strand as a flexible chain. The interfacial term that helps in bubble formation through forking, and the three chain interaction are defined here too. The calculations are done in the Fourier-Laplace space. The necessary rules for diagrams and the form of the partition functions are given in Sec. 4.2. We approach from the bound to the unbound side. For the two-chain bound state at finite temperatures, forking is allowed with some energy cost. Two successive forkings result in the formation of a bubble which can be infinite in number. The corresponding duplex partition function and the melting of the rigid DNA are discussed in Sec. 4.3. The three-chain case where we consider the interaction between a dsDNA with no bubbles and a single chain can be found in Sec. 4.4. We conclude this chapter in Sec. 4.5.

## 4.1 The Model

Our model of three polymer chains is defined, *à la* Poland-Scheraga, through their partition functions. Every monomer lives in a  $d$ -dimensional space. All chains are of equal length  $N$ . They are tied at one end at origin in space while the other end may also be tied together at a point  $\mathbf{r}$ , though the latter constraint may be relaxed. One end needs to be kept together to prevent any chain from flying away, thereby facilitating the bookkeeping for entropy. The monomers on different chains interact only when they are at the same space ( $\mathbf{r}$ ) and length ( $z$ ) coordinates. Such an attractive interaction corresponds to the native base pairing of DNA.

The basic constituents of the model are the partition functions for a single chain, for a bound pair, the weight  $g_2$  for dissociation of a pair or joining of two chains (Y forks), and the interaction  $g_3$  between a single chain and a bound pair. The two important parameters in this problem are  $g_2$  and  $g_3$ .

We first define the basic partition functions, viz., (i)  $Z(\mathbf{r}, N)$  for a single chain, and (ii)  $Z_b(\mathbf{r}, N)$  for a pure bound state of two strands.

### 4.1.1 Single chain

A single strand is a flexible Gaussian chain with

$$Z(\mathbf{r}, N) = \mu^{N\Lambda^2} \frac{1}{(2\pi N)^{d/2}} e^{-\frac{\mathbf{r}^2}{2N}}, \quad (4.1)$$

where  $\mu^{N\Lambda^2}$  is the total number of configurations, and  $\Lambda^{-1}$  is a short distance or microscopic cutoff. For a Gaussian chain the overall size  $R$  of a polymer scales as

$$R^2 \sim N, \quad (4.2)$$

which allows us to set the dimension of  $N$  as

$$[N] = L^2, \quad \text{when } [r] = [\Lambda^{-1}] = L, \quad (4.3)$$

the square bracket [...] indicating the dimensionality of the enclosed entity. In Eq. (4.1),  $\Lambda$  is used to make  $N$  dimensionless in the  $\mu$ -dependent factor. The unconstrained entropy of a free chain is taken as  $\propto \ln \mu$  per unit length to avoid the problem of an infinite entropy of a continuous chain. The Gaussian factor in Eq. (4.1) is the

probability density of finding the end of the polymer at  $\mathbf{r}$ , so that the total partition function after integration over all space is dimensionless.

### 4.1.2 Bound state, Y fork, duplex, $g_2$ and $g_3$

The second partition function needed is for the two-chain bound state which is taken as a rigid rod with  $\epsilon\Lambda^2$  as the binding energy per unit length. The bound DNA with one end fixed can rotate in space as a whole but can not bend. The partition function of the bound state of length  $N$  is given by

$$Z_b(\mathbf{r}, N) = \frac{1}{4\pi} e^{-\epsilon N\Lambda^2} \delta(\mathbf{r} - N\Lambda\hat{\mathbf{n}}), \quad (4.4)$$

where  $\hat{\mathbf{n}}$  is a unit vector giving the direction of the rigid rod.

Now, there are finite temperature fluctuations in the form of pair breaking. The bound state then locally dissociates into two single strands to form a Y fork. The two free strands may rejoin to produce a bubble. We assign an interface weight  $g_2$  in the partition function for every Y fork. This is like an extra interfacial contribution and is referred to as the ‘‘co-operativity factor’’ [69]. It is shown below Eq. (4.15), that dimensionwise

$$[g_2^2] = [\Lambda]^{4-d}, \quad (4.5a)$$

$$= [\Lambda]^1, \quad (d = 3). \quad (4.5b)$$

The importance of bubbles is well-recognized both in the biological functions and in physical properties of DNA. There have been many studies, in recent times, both theoretical and experimental, on the nature and functions of the bubbles under various situations, like DNA under a force or topological constraints [126–130], in breathing dynamics [131–134], in hysteresis [55, 57], with semiflexibility [135], etc.

The introduction of finite temperature bubbles makes the bound state flexible and as a result it can bend. We name this bound state with bubbles a *duplex*. The  $g_2$ -dependent partition function  $Z_d$  is discussed in Sec 4.3. We see there that the formation of large bubbles, by thermal fluctuations leads to the melting of a duplex into two free chains, at a critical value of  $g_2 = g_{2c}$ . The role of  $g_2$  may be represented

by the sequence of partition functions

$$\underbrace{Z_b}_{\{g_2=0\}} \xrightarrow{\text{crossover}} \underbrace{Z_d}_{\{g_2 < g_{2c}\}} \xrightarrow{\text{melting}} \underbrace{Z^2}_{\{g_2 > g_{2c}\}}, \quad (4.6)$$

where the change from a rigid to an elastic one by  $g_2$  is a crossover (not a transition). An interaction between a rigid bound state and a single chain is our three-body parameter  $g_3$ . We show below Eq. (4.24) that it has the dimensionality

$$[g_3] = [\Lambda]^{2-d}, \quad (4.7a)$$

$$= [\Lambda]^{-1} \quad (d = 3). \quad (4.7b)$$

## 4.2 Diagrammatic definitions and rules

Since we shall be using diagrammatic representations for many equations, it is prudent to define our model in terms of diagrams and the rules for computations. To take advantage of the convolution property of the Fourier and the Laplace transforms, we work in the Fourier-Laplace space  $(\mathbf{k}, s)$  instead of  $(\mathbf{r}, N)$ . The conventions for these transforms are

$$\hat{Z}(\mathbf{k}, N) = \int Z(\mathbf{r}, N) e^{-i\mathbf{k}\cdot\mathbf{r}} d^d r, \quad (4.8)$$

$$Z(\mathbf{k}, s) = \int_0^\infty \hat{Z}(\mathbf{k}, N) e^{-Ns} dN, \quad (4.9)$$

with the inverse transforms defined as

$$Z(\mathbf{r}, N) = \frac{1}{(2\pi)^d} \int \hat{Z}(\mathbf{k}, N) e^{i\mathbf{k}\cdot\mathbf{r}} d^d k, \quad (4.10)$$

$$Z(\mathbf{k}, N) = \frac{1}{2\pi i} \oint Z(\mathbf{k}, s) e^{Ns} ds. \quad (4.11)$$

The contour for the integration in Eq. (4.11) is the usual Mellin's contour.

The dimensionalities of the partition functions, as per our conventions, are as follows

$$[Z(\mathbf{r}, N)] = L^{-d}, [\hat{Z}(\mathbf{k}, N)] = L^0, \quad (4.12a)$$

$$[Z(\mathbf{k}, s)] = L^2, \text{ with } [s] = L^{-2}. \quad (4.12b)$$

These  $L$  dependencies are used to identify the dimensionalities of the remaining parameters.

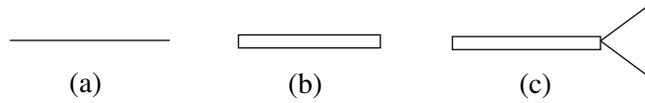


Figure 4.1: Basic building blocks. Panel (a) represents  $Z(\mathbf{k}, s)$  for a Gaussian chain, (b)  $Z_b(\mathbf{k}, s)$  for a two chain bound state, (c) a Y fork representing the interface between a bound pair and two open strands. It has a weight  $g_2$ .

### 4.2.1 Free chain

In the Fourier-Laplace space the single chain partition function, Eq. (4.1), becomes

$$Z(\mathbf{k}, s) = \frac{1}{s - \Lambda^2 \ln \mu + \frac{k^2}{2}}. \quad (4.13)$$

To be noted here is that the free energy per unit length comes from the pole of Eq. (4.13) in the complex- $s$  plane. This partition function, Eq. (4.13), to be called a propagator, is represented by a solid line in Figure 4.1(a).

### 4.2.2 Bound state

The rigid bound state of Eq. (4.4) is direction dependent. For simplicity we take an average over all directions by integrating over the solid angle subtended by  $\hat{\mathbf{n}}$  (see Appendix D). The corresponding Fourier-Laplace transformed partition function is given by

$$Z_b(\mathbf{k}, s) = \frac{1}{k\Lambda} \arctan \frac{k\Lambda}{s + \epsilon\Lambda^2} \stackrel{k \rightarrow 0}{=} \frac{1}{s + \epsilon\Lambda^2}. \quad (4.14)$$

The  $k \rightarrow 0$  form of  $Z_b(\mathbf{k}, s)$  can be used close to the melting defined shortly. Here also the pole in the complex  $s$  plane gives the free energy of the rigid bound state. The bound-state partition function of Eq. (4.14) is represented by an unfilled rectangular box in Figure 4.1(b).

At finite temperatures, the inclusion of the Y forks gives the duplex partition function  $Z_d(\mathbf{k}, s)$  represented by a filled black box in Figure 4.2(a).

The Y-fork junction,  $g_2$ , is represented in Figure 4.1(c) by a vertex where a rectangular box (a bound pair) and two solid lines (free chains) meet. The diagram in Figure 4.2(c) represents an interaction between a bound state and a free chain. The interaction is the three-body coupling constant given by  $g_3$ .

### 4.2.3 $\mathbf{k}$ and $s$ Conservation

As all the partition functions have translation invariance we have  $\mathbf{k}$  conservation at each vertex. Following the standard nomenclature, the  $\mathbf{k}$  vectors are called “momentum.” So momentum is conserved at each vertex. Also cutoff  $\Lambda$  is called a momentum cutoff.

The Y fork and the three-chain interaction can take place anywhere along the length of the polymers which are infinitely long. This invariance (translational invariance along the contour) leads to an  $s$ -conservation at any point. In other words, at a junction  $s$ -values get distributed, but the total remains the same. More details are discussed in Appendix E.

## 4.3 Two Chains

At first let us find the duplex partition function. In the grand canonical ensemble (Laplace space) the singularity of the partition function closest to the origin gives the free energy of the system; contributions from others are suppressed in the thermodynamic limit. Whenever there is a switching of the nearest singularity due to a change in some parameter of the system, we have a phase transition. Henceforth all the calculations are done in  $d = 3$ .

Considering an arbitrary number of bubbles we can write the finite temperature bound state as a sum of an infinite number of diagrams shown in Figure 4.2(a). In terms of the Laplace variable  $s$ , the duplex partition function, denoted by  $Z_d(\mathbf{k}, s)$ ,

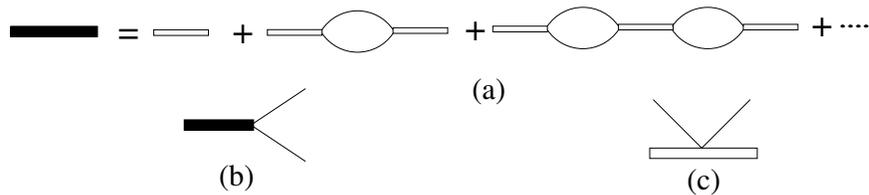


Figure 4.2: (a) The duplex partition function as an infinite series of bound pairs and bubbles, (b) Y fork for a duplex. (c) A three-chain interaction,  $g_3$ , involving a free chain and a duplex.

can be written as a geometric series

$$\begin{aligned} Z_d(\mathbf{k}, s) &= Z_b(\mathbf{k}, s) + g_2^2 Z_b(\mathbf{k}, s) I_0 Z_b(\mathbf{k}, s) + \dots \\ &= \frac{1}{\frac{1}{Z_b(\mathbf{k}, s)} - g_2^2 I_0}, \end{aligned} \quad (4.15)$$

where  $I_0$ , the single bubble contribution, is given by

$$\begin{aligned} I_0 &= \int \frac{d\mathbf{q}}{(2\pi)^3} \frac{d\bar{s}}{2\pi i} Z\left(\frac{\mathbf{k}}{2} - \mathbf{q}, \bar{s}\right) Z\left(\frac{\mathbf{k}}{2} + \mathbf{q}, s - \bar{s}\right) \\ &= \frac{1}{2\pi^2} \left[ \Lambda - \sqrt{s' + k^2/4} \arctan \frac{\Lambda}{\sqrt{s' + k^2/4}} \right], \end{aligned} \quad (4.16)$$

and  $s' = s - 2\Lambda^2 \ln \mu$ . Equation (4.15), with Eqs. (4.12a) and (4.12b), sets the dimension of  $g_2$  as quoted in Eq. (4.5a).

To evaluate  $I_0$  we do the  $\bar{s}$  integral by the method of residues. See Appendix E for details. The only contribution comes from the simple pole at  $\bar{s} = \Lambda^2 \ln \mu - (\mathbf{k}/2 - \mathbf{q})^2/2$ . In the limit  $(s' + k^2/4) \rightarrow 0$  and  $\Lambda$  finite, which is the relevant limit near the transition point, we have

$$I_0 = \frac{1}{2\pi^2} \left[ \Lambda - \frac{\pi}{2} \sqrt{s' + k^2/4} \right]. \quad (4.17)$$

So the duplex partition function becomes

$$Z_d(\mathbf{k}, s) = \left\{ s + \epsilon\Lambda^2 - \frac{1}{2\pi^2} g_2^2 \left[ \Lambda - \frac{\pi}{2} \sqrt{s' + k^2/4} \right] \right\}^{-1}. \quad (4.18)$$

We identify here three different singularities in the partition function  $Z_d$ , which correspond to three distinct states. The branch point singularity of Eq. (4.18) at  $s = 2\Lambda^2 \ln \mu$ , owing its origin to  $Z$ , gives the completely unbound (denatured) state. This is the high temperature phase. The singularity at  $s = -\epsilon\Lambda^2$  corresponds to the completely bound state when  $g_2 = 0$ . This, being the singularity of  $Z_b$ , is the zero temperature phase and does not survive when  $g_2 \neq 0$ . The third singularity  $s'_*$  comes from the zero of the denominator of  $Z_d(0, s)$ . As  $s'_*$  continuously evolves with  $g_2$  from  $s = -\epsilon\Lambda^2$ , it corresponds to a bound state with bubbles. In the absence of  $g_2$ , i.e., in absence of any interface or junction point, there can be two states only; the system stays either in the completely bound state or in the completely unbound state. There could be a denaturation transition, necessarily first order, by changing  $\epsilon$

or  $\mu$ . This case is of no interest to us. The presence of the interface alters the nature of the bound state because of the bubbles and also makes the transition critical [see Eq. 4.6].

Our main aim is to concentrate on the behavior of the system near duplex melting where the contributions from the bubbles (loops) of large sizes dominate the duplex partition function. In the small  $s'$  limit with  $k = 0$ ,  $s'_*$  is given by

$$\sqrt{s'_*} = -\frac{\Delta t}{2\pi^2 g_2^2 \Lambda^{-2}}, \quad (4.19)$$

where  $\Delta t \equiv (2\pi)^3(2 \ln \mu + \epsilon) - 4\pi g_2^2 \Lambda^{-1}$ .  $s'_*$  can be identified as the difference of free energies between duplex and two free chain states. So,  $\Delta t$  is a measure of deviation from the duplex melting point. The equation  $\Delta t = 0$  [136] gives the critical point as

$$g_{2c} = \sqrt{2\pi^2(2 \ln \mu + \epsilon)\Lambda}. \quad (4.20)$$

The thermal melting of a dsDNA can also be illustrated from this model [137]. Using Eq. (4.19), we define a diverging length scale  $\xi$  in the following way

$$s'_* \sim \xi^{-2}, \text{ with } \xi \sim |\Delta t|^{-1}. \quad (4.21)$$

If we now make a scale change such that, for arbitrary  $b$ ,  $\mathbf{k} \rightarrow b^{-1}\mathbf{k}$ , the length scale changes as  $\xi \rightarrow b\xi$ . And as  $s'_*$  is the free energy difference, the free energy scales as  $f \rightarrow b^{-2}f$ . This is the continuous scale invariance satisfied at the thermal dsDNA melting.

By tuning  $g_2$  the length scale  $\xi$  can be made divergent for some critical value of  $g_2$ ,  $g_{2c}$ . Beyond it, the system goes to a stable high temperature phase with two free chains. The full duplex partition function can be written for small  $s'$  as

$$Z_d(\mathbf{k}, s) = \frac{(2\pi)^3}{2\pi^2 g_2^2 \left[ -\xi^{-1} + \sqrt{s' + k^2/4} \right]}, \quad (4.22)$$

which explicitly shows the  $\xi$  dependence.

## 4.4 Three Chains

Now consider the three-chain problem. As discussed in the introduction of this thesis the effect of thermal fluctuations (bubbles) are very important in getting an Efimov-DNA. To make this more clear, let us first consider the case of a free chain interacting

with a bound state. We will consider the case with bubbles in the next chapter. Here we set  $g_2 = 0$  such that there are no bubbles in the bound state. This problem with interaction up to all orders can be solved by solving the diagrammatic integral equation shown in Figure 4.3 (see Appendix E).

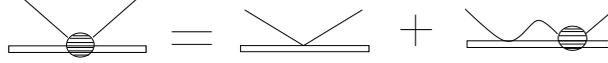


Figure 4.3: Interaction of free chain with a bound state in absence of bubbles.

The bare single chain bound state contact interaction is  $g_3$  and we denote the corresponding renormalized interaction by  $V$ . Evaluations of Figure 4.3 in  $d$ -dimensions give us

$$V = -g_3 - g_3 V \int \frac{d\mathbf{q}}{(2\pi)^d} \frac{d\bar{s}}{2\pi i} Z(\mathbf{q}, \bar{s}) Z_b(-\mathbf{q}, s - \bar{s}). \quad (4.23)$$

Here  $V$  is taken as a function of  $\Lambda$  but not  $\mathbf{q}, s$ . The above equation can be used to obtain the dimension of  $g_3$  as quoted in Eq. (5.1a). The  $\bar{s}$  integral is evaluated by the method of residues as

$$V = -g_3 - g_3 V \int_0^\Lambda \frac{\Omega_d}{(2\pi)^d} \frac{q^{d-1} dq}{s - \Lambda^2 \ln \mu + \Lambda^2 \epsilon + q^2/2}, \quad (4.24)$$

where  $\Omega_d$  is the surface area of the  $d$ -dimensional unit hyper sphere. The unbound phase consists of two independent members, a rigid bound pair and a free polymer, with the system free energy determined by  $s = \Lambda^2(\epsilon - \ln \mu)$ . Considering the system to be in this unbound state we get

$$\hat{V} = -\hat{g}_3 - \hat{g}_3 \frac{\hat{V}}{d-2}, \quad (4.25)$$

where

$$\hat{g}_3 = \frac{2\Omega_d \Lambda^{d-2}}{(2\pi)^d} g_3, \quad \hat{V} = \frac{2\Omega_d \Lambda^{d-2}}{(2\pi)^d} V, \quad (4.26)$$

are dimensionless quantities. Equation (4.25) can be rewritten as

$$\hat{g}_3 = -\frac{\hat{V}}{1 + \frac{\hat{V}}{d-2}}. \quad (4.27)$$

The RG flow equation of  $\hat{g}_3$  is obtained simply by differentiating with respect to  $\Lambda$ , keeping  $V$  constant. The result is

$$\Lambda \frac{\partial \hat{g}_3}{\partial \Lambda} = (d-2)\hat{g}_3 + \hat{g}_3^2, \quad (4.28)$$

where the linear term on the right hand side can be linked to dimensional analysis, Eq. (4.26). The quadratic term is the loop contribution. A small loop, quadratic in  $g_3$ , on a bigger scale would look like an effective interaction, modifying the coupling constant.

Writing down the RG flow equation in terms of  $\Lambda$  for the bare values is similar to the use in quantum problems. Here one studies the flow of the bare values for a fixed renormalized coupling, while the converse is done in the usual polymer RG. Consequently the stability of the fixed points are of opposite nature compared to the polymer RG fixed points of say Ref. [77, 78, 84].

To illustrate this further we map our problem with contact interaction to the quantum problem of a particle in a finite rectangular potential well with  $V_0$  and  $a$  as the depth and the width of the well respectively. The contact interaction can be represented with a  $\delta$ -function potential  $V\delta(x)$  such that the coefficient  $V$  represents the interaction strength. As the potential energy have the same dimension as the kinetic energy ( $\nabla^2 \sim L^{-2}$ ) the dimensionality of  $V = L^{d-2}$ . We need to ensure  $V_0 a^d$  remains finite in the limit  $V_0 \rightarrow \infty, a \rightarrow 0$  while we get a potential  $V_0 a^d \delta(x)$ . A comparison between the coefficients of the both  $\delta$ -functions give  $\hat{V} = \frac{V_0}{\Lambda^2}$ . If we now take the limit  $\Lambda \rightarrow \infty$  to get a constant  $\hat{V} \lesssim 0$ ,  $V_0$  must go to infinity. This means  $\hat{V}/V_0 \rightarrow 0$  for  $\Lambda \rightarrow \infty$  and the particle looks at the edge of the well. The wave function outside of the well decays as  $\text{Exp}(-\alpha x)$  where  $\alpha \sim \sqrt{|E|}$  and  $E$  is the energy of the particle. But as we are keeping the energy fixed the width of the wavefunction remains fixed. In this situation the shrinking of the width of the well means the particle stays outside of the well most of the time. This is schematically shown in Figure 4.4. In terms of the bubbles, the length-scale for the bubbles, be it above or below the transition ( $V \lesssim 0$ ), looks much larger compared to the range of the interaction and, therefore, closer to the critical point which has a diverging length scale. This explains why the critical point in this scheme corresponds to a stable fixed point.

The flow equation is similar to the RG flow of the two-chain coupling [84] with a stable and an unstable fixed point. For  $d = 3$ , the stable fixed point  $\hat{g}_3^* = -1$  corresponds to the critical point of unbinding and  $\hat{g}_3^* = 0$  represents the unbound state. This is expected because the bound DNA acts as a single rigid polymer with

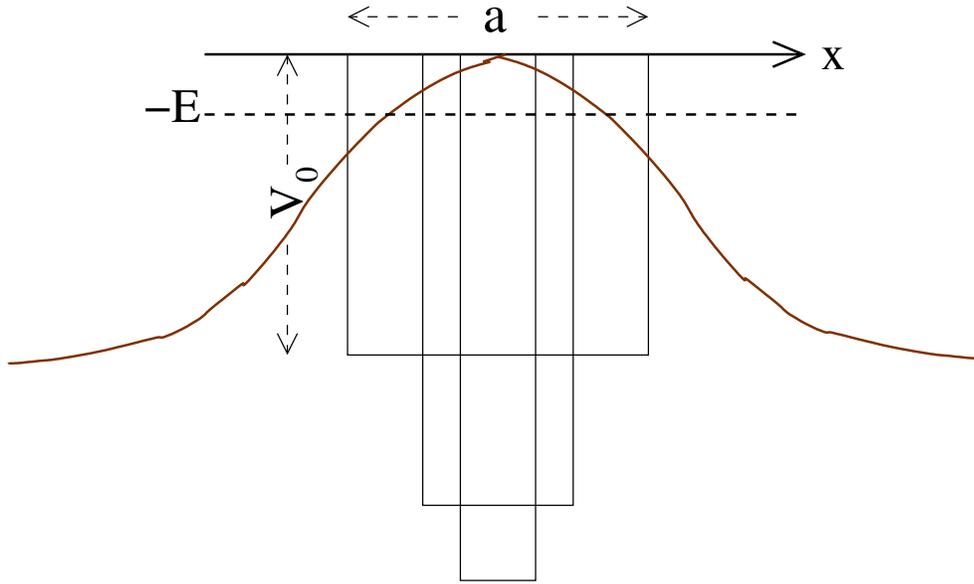


Figure 4.4: In the limit  $a \rightarrow 0$ ,  $V_0 \rightarrow \infty$  for a finite  $V_0 a^d$ . In this limit a particle with energy  $-E$  looks close to the threshold and remains outside of the well most of the time. The solid brown line represents the wave function.

no internal structure so that the problem is effectively like the unbinding of two dissimilar DNA strands.

## 4.5 Conclusion

In this chapter we presented a Poland-Scheraga type model which shows a continuous melting at three dimension. This model is generic in nature and may be used for other purposes, for example, to study the DNA elasticity problem. Here we also introduced the rules to calculate the diagrammatic expressions in the Fourier-Laplace space which will be used again in the next chapter.

# Appendix D

## Bound State

The Fourier-Laplace transformed bound state partition function is given by

$$\begin{aligned} Z_b(\mathbf{k}, s) &= \int d\hat{\mathbf{n}} \int_0^\infty dN e^{-Ns} \\ &\quad \times \int d^3\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} \frac{e^{-\epsilon N\Lambda^2}}{(4\pi)} \delta(\mathbf{r} - N\Lambda\hat{\mathbf{n}}) \\ &= \int \frac{d\hat{\mathbf{n}}}{4\pi} \int_0^\infty dN e^{-Ns} e^{-\epsilon N\Lambda^2} e^{iN\Lambda\mathbf{k}\cdot\hat{\mathbf{n}}} \\ &= \frac{1}{2} \int_0^\pi \frac{\sin\theta d\theta}{s + \epsilon\Lambda^2 - ik\Lambda \cos\theta} \\ &= \frac{1}{k\Lambda} \arctan \frac{k\Lambda}{s + \epsilon\Lambda^2} \end{aligned} \tag{D.1}$$

# Appendix E

## Rules Of Diagrammatic Calculations

In this appendix we list all the rules to evaluate diagrams we have used.

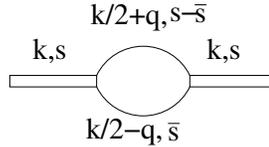


Figure E.1: A bound state with one bubble showing variables in the Fourier-Laplace space variable obeying the  $\mathbf{k}$  and the  $s$  conservations.

Every partition function has two arguments: one space and one length respectively. With translational invariance, the arguments would be the difference of the corresponding quantities at the two ends of each piece. The dissociation of a bubble or a duplex is our two-body vertex  $g_2$  [Figure 4.1(c) and Figure 4.2(b)]. The interaction between one single chain and a bound state [Figure 4.2(c)] is the three-chain vertex  $g_3$ . The algebraic expression for any diagram is obtained by sequentially multiplying the partition functions and the vertexes as arranged, with integrations over the intermediate variables. A renormalized vertex is called a vertex function.

To see how the  $k$  conservation appears, consider a bound state with one bubble of

the type in Figure E.1. Applying the above stated rules, this diagram is evaluated as

$$I(\mathbf{r}, N) = \int Z_b(\mathbf{r}_1|z_1)Z^2(\mathbf{r}_1 - \mathbf{r}_2|z_2 - z_1) \times Z_b(\mathbf{r} - \mathbf{r}_2|N - z_2)d\mathbf{r}_1d\mathbf{r}_2dz_1dz_2. \quad (\text{E.1})$$

The convolution form in the real space leads to a product form in the Fourier space. We suppress the  $z$  integrals for the time being. Fourier transforming both sides from variable  $\mathbf{r}$  to  $\mathbf{k}$  and rewriting right hand partition functions in terms of their Fourier transformed functions we get

$$\begin{aligned} \hat{I}(\mathbf{k}, N) &= \int I(\mathbf{r}, N)e^{-i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r} \\ &= \frac{1}{(2\pi)^{4d}} \int Z_b(\mathbf{k}_1|z_1)Z(\mathbf{k}_2|z_2 - z_1)Z(\mathbf{k}_3|z_2 - z_1)Z_b(\mathbf{k}_4|N - z_2) \\ &\quad \times e^{i\mathbf{r}_1\cdot(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3)}e^{i\mathbf{r}_2\cdot(\mathbf{k}_2 + \mathbf{k}_3 - \mathbf{k}_4)}e^{i\mathbf{r}\cdot(\mathbf{k}_4 - \mathbf{k})}d\mathbf{r} \prod_{j=1,2} \{d\mathbf{r}_jdz_j\} \prod_{l=1}^4 d\mathbf{k}_l, \\ &= \frac{1}{(2\pi)^{4d}} \int Z_b(\mathbf{k}_1|z_1)Z(\mathbf{k}_2|z_2 - z_1)Z(\mathbf{k}_3|z_2 - z_1)Z_b(\mathbf{k}_4|N - z_2) \times \\ &\quad \delta(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3)\delta(\mathbf{k}_2 + \mathbf{k}_3 - \mathbf{k}_4)\delta(\mathbf{k}_4 - \mathbf{k}) dz_1dz_2 \prod_{j=1}^4 d\mathbf{k}_j. \quad (\text{E.2}) \end{aligned}$$

Performing three  $\delta$  function integrals we get the following relation between different  $\mathbf{k}$ 's:

$$\mathbf{k}_1 = \mathbf{k}_2 + \mathbf{k}_3, \quad \mathbf{k} = \mathbf{k}_4, \quad \mathbf{k}_4 = \mathbf{k}_2 + \mathbf{k}_3 \quad \text{and} \quad \mathbf{k} = \mathbf{k}_1. \quad (\text{E.3})$$

From these relations it is clear that overall there is one single  $\mathbf{k}$  and it is conserved at every junction point. Now as there are four unknown  $\mathbf{k}$ s and we have only three constraints, there is one undetermined  $\mathbf{k}$  left. This is the characteristic of the loop in the diagram. Whenever there is a loop there is an undetermined  $\mathbf{k}$  over which we have to integrate ( $\frac{1}{(2\pi)^d} \int d\mathbf{k}$ ). The integration over  $\mathbf{k}$  corresponds to a bubble in real space with two ends fixed, as, e.g., in  $Z^2(\mathbf{r}_1 - \mathbf{r}_2|z_2 - z_1)$  in Eq. (E.1).

Now we show how the  $s$  conservation appears by evaluating Eq. (E.1). Laplace transforming both sides in  $N$  and rewriting every term in the right hand side through

their inverse Laplace transformation we have

$$\begin{aligned}
A(\mathbf{r}, s) &= \int_0^\infty e^{-sN} I(\mathbf{r}, N) dN \\
&= \frac{1}{(2\pi i)^4} \int_0^\infty e^{-sN} dN \int \prod_{j=1,2} \{d\mathbf{r}_j dz_j\} \prod_{l=1}^3 ds_l Z_b(\mathbf{r}_1, s_1) e^{s_1 z_1} Z(\mathbf{r}_2 - \mathbf{r}_1, s_2) e^{s_2(z_2 - z_1)} \\
&\quad \times Z(\mathbf{r}_2 - \mathbf{r}_1, s_2) e^{s_3(z_2 - z_1)} Z_b(\mathbf{r} - \mathbf{r}_2, s_4) e^{s_4(N - z_2)}, \quad (\text{E.4})
\end{aligned}$$

where the  $s_i$  integrals are the usual Mellin integrals. We evaluate  $z$  integrations with limit  $z_1 = 0$  to  $z_2$ , and  $z_2 = 0$  to  $N$  to get

$$A(\mathbf{r}, s) = \int_0^\infty dN \int_{r_i, s_i} (\dots) \left[ \frac{e^{N(s_1 - s_4)} - 1}{(s_1 - s_4)(s_1 - s_2 - s_3)} - \frac{e^{N(s_2 + s_3 - s_4)} - 1}{(s_2 + s_3 - s_4)(s_1 - s_2 - s_3)} \right] e^{(s_4 - s)N}. \quad (\text{E.5})$$

If we now do the  $s_i$  integrations using the method of residues, contributions come only from the poles at  $s_1 = s_4$ ,  $s_1 = s_2 + s_3$  and from the  $N$  integration  $s = s_4$ . As the first bound segment is labeled by  $s_1$ , two free chains are labeled by  $s_2$  and  $s_3$  and the end bound state is labeled by  $s_4$ , we see the conservation at every point with an overall  $s$ . And similar to the case of  $\mathbf{k}$  conservation, every loop in Laplace space also possesses one undetermined  $s$  as there are three relations and four  $s$ 's to be determined. So whenever a loop comes, we have to integrate over that undetermined  $s$  ( $\frac{1}{2\pi i} \int ds$ ).

We can now label the diagram in the Fourier-Laplace space using the above conservation rules as shown in Figure E.1(c). When evaluated algebraically it gives  $Z_b^2(\mathbf{k}, s) I_0$  with the loop integral

$$\begin{aligned}
I_0 &= \int \frac{d\bar{s}}{2\pi i} \frac{d\mathbf{q}}{(2\pi)^3} Z\left(\frac{\mathbf{k}}{2} - \mathbf{q}, \bar{s}\right) Z\left(\frac{\mathbf{k}}{2} + \mathbf{q}, s - \bar{s}\right) \\
&= \int \frac{d\bar{s}}{2\pi i} \frac{d\mathbf{q}}{(2\pi)^3} \frac{1}{\bar{s} - \Lambda^2 \ln \mu + \frac{(\mathbf{k}/2 - \mathbf{q})^2}{2}} \frac{1}{s - \bar{s} - \Lambda^2 \ln \mu + \frac{(\mathbf{k}/2 + \mathbf{q})^2}{2}}. \quad (\text{E.6})
\end{aligned}$$

We evaluate the  $\bar{s}$  integral by employing the method of residues. There is a simple pole at  $\bar{s} = \Lambda^2 \ln \mu - \frac{(\mathbf{k}/2 - \mathbf{q})^2}{2}$ . All the contribution to the integral comes only from this simple pole. So replace the rest of the  $\bar{s}$  by its value at the pole and the prefactor  $\frac{1}{2\pi i}$  cancels out yielding

$$\begin{aligned}
I_0 &= \int \frac{d\mathbf{q}}{(2\pi)^3} \frac{1}{s' + \frac{k^2}{4} + q^2} \\
&= \frac{4\pi}{(2\pi)^3} \left[ \Lambda - \sqrt{s' + k^2/4} \arctan \frac{\Lambda}{\sqrt{s' + k^2/4}} \right]
\end{aligned}$$

where  $s' = s - 2\Lambda^2 \ln \mu$ . A similar kind of integrals also appear while evaluating three-chain diagrams. We employ this same procedure to evaluate them.

All the diagrams of this chapter and the next chapter are evaluated by using the rules and procedures discussed in this appendix.

## Chapter 5

# Efimov-DNA and Renormalization Group Limit Cycle

A three-stranded DNA with short range base pairings only is known to exhibit a classical analog of the quantum Efimov effect, viz., a three-chain bound state at the two chain melting point where no two are bound. By using a non-perturbative renormalization-group method for a rigid duplex DNA and a flexible third strand, with base pairings and strand exchange, in this chapter we show that the Efimov-DNA is associated with a limit cycle type behavior of the flow of an effective three-chain interaction. The analysis also shows that thermally generated bubbles play an essential role in producing the effect. A toy model for the flow equations shows the limit cycle in an extended three-dimensional parameter space of the two-chain coupling and a complex three-chain interaction.

The outline of the chapter is given below. How the strand exchange mechanism is used in our model is demonstrated in Sec. 5.1.1. A brief overview of what is expected in a limit cycle RG is given in Sec. 5.1.2. The calculations are done in the Fourier-Laplace space. The three-chain case where we consider the duplex-free-chain interaction can be found in Sec. 5.2. The whole three-chain part is analyzed at the two body critical point. A few details can be found in the appendixes. In particular, Appendix F is about a toy example that extrapolates between the flow equations for the three chain interaction with no bubbles and the same at the critical melting point. We end with a discussion of the results and their experimental consequences in Sec.

5.4 and a short summary in Sec. 5.5.

## 5.1 Model and result

Here we allow bubble formation in the bound state by taking  $g_2 \neq 0$ . All the definition of the partition functions and the rules for Diagrammatic calculations remain the same as described in the previous chapter. But the two-body parameter  $g_2$  is now defined by dissociation of a duplex and the three-body parameter  $g_3$  is defined by the interaction between the duplex and a single chain. As the duplex has open single strands it can now participate in strand exchange with a third strand.

### 5.1.1 Strand exchange and $g_3$

For the three-polymer system, we consider the situation where a pair is in the bound or duplex state and the other one is free. The third chain is allowed to interact with a free strand of a bubble allowing it to form a duplex locally. This is called strand exchange. Figure 5.1 illustrates this process schematically. The pair interaction

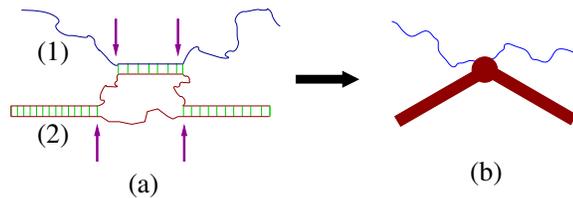


Figure 5.1: (Color online) Schematic diagram of a strand exchange and the equivalent coarse-grained three-chain interaction  $g_3$ . (a) A single strand [blue line marked (1)] pairs with one strand of a bubble on a duplex [brown lines marked (2)]. The short vertical (green) lines indicate base pairings, with the energy per unit length  $\epsilon$ . The junction weight  $g_2$  is associated with each fork or the interface on a duplex. In (a) there are four interfaces as indicated by the arrows. (b) A coarse-grained version of (a) where the duplex is represented by a thick line interacting with the single line. The filled circle represents the three-chain interaction  $g_3$ .

together with the strand exchange would, in principle, be sufficient to formulate the three-chain DNA problem, but in a renormalization group approach, the three-chain

bound state is described by a three-chain interaction. Therefore, in anticipation of its generation, we allow a three-chain interaction (between a single chain and a duplex)  $g_3$  (Figure 5.2(b)). This interaction parameter is our three-body coupling. We see below Eq. (4.23) that it has the dimensionality

$$[g_3] = [\Lambda]^{2-d}, \quad (5.1a)$$

$$= [\Lambda]^{-1} \quad (d = 3). \quad (5.1b)$$

The dimensionless three-chain parameter may now be constructed as [139]

$$H(\Lambda) = -\frac{g_3}{4g_2^2}\Lambda^2, \quad (5.2)$$

with a factor of 4 for convenience.

The partition function for three chains depends on both  $g_2$  and  $g_3$  but for a duplex it depends only on  $g_2$ .

### 5.1.2 Qualitative description

Our aim is to see the effect of strand exchange on the three-chain system near the duplex melting point where large sized bubbles are expected. Right at the melting point, we may concentrate on how  $g_3$  or  $H$  evolves as  $\Lambda \rightarrow \infty$  for large  $N$ . A flow of  $H$  from zero to  $+\infty$  is an indication of a three-chain bound state [note the negative sign in Eq. (5.2)], the Efimov-DNA case of a bound three-chain system where no two are bound.

In the RG approach the flows of parameters are obtained in a few steps. With a reciprocal space upper cutoff  $\Lambda$  (a short distance scale  $\sim \Lambda^{-1}$ ), the effects over a range  $\Lambda$  to  $\Lambda - d\Lambda$  are taken into account by redefining the problem for scales up to  $\Lambda - d\Lambda$ . A subsequent rescaling brings back the problem to the original scale with renormalized parameters. The changes in the parameters, as continuous variables, gives us the flow equations or  $\beta$  functions

$$\Lambda \frac{\partial H}{\partial \Lambda} = \beta(H). \quad (5.3)$$

The behavior of a system is then characterized by the flows which generally terminate at stable fixed points (or at infinity) separated by unstable fixed points. The fixed

points represent the phases and the phase transitions in the system. This is the generic picture of RG and this is where the three-chain problem stands out.

In our approach,  $g_2$  is the control parameter for the duplex melting. We first determine the critical point of melting by locating the critical value  $g_{2c}$  at which a suitably defined length scale  $\xi$  diverges. At this particular point we determine the RG flow or the  $\beta$  function for  $H(\Lambda)$ . Quantitatively this is implemented by calculating the third virial coefficients, obtained from the connected three-chain partition function (i.e., for polymers connected by the interactions).

A naive use of the definition, Eq. (5.2), suggests the form

$$\beta(H) = 2H, \quad (\text{naive}). \quad (5.4a)$$

This however gets modified by the effects of strand exchange and other fluctuations to a form

$$\beta(H) = 2H + \mathcal{F}(H), \quad (\text{with renormalization}), \quad (5.4b)$$

and all the nontriviality comes from these additional terms. In general, it has a form

$$\beta(H) = -AH^2 + BH + C, \quad (5.5)$$

$A, B, C$  being all real. In conventional cases,  $B$  is mainly determined by the naive dimensional analysis, while a nonzero  $A$  is the extra addition of length rescaling and renormalization. A constant term  $C$  is unusual and appears if there are marginal parameters, which do not change with length scales. An example of a marginal parameter is  $g$  of the inverse square interaction mentioned in the Introduction. This identification may be turned around to argue that a constant term in the  $\beta$  function signals the presence of some, may be hidden, marginal parameter in the problem.

If  $C$  in Eq. (5.5) is such that there are two real roots of  $\beta(H) = 0$ , the standard picture remains valid with one stable and one unstable fixed points, but not so if the roots are complex conjugate pairs [for  $C < -B^2/(4A)$ ]. Such is the case for the problem in hand. We find

$$\Lambda \frac{\partial H}{\partial \Lambda} = -A(H - H_0)(H - H_0^*), \quad (5.6)$$

where  $H_0, H_0^*$  form a complex-conjugate pair.

The procedure we adopt to derive Eq. (5.6) is different from the conventional RG way. The traditional approach is to take into account the effects at the short distance level to redefine the parameters on a larger length scale. Instead of such an approach we determine the effective parameter as an integral equation and then use a thin-shell integration method to get the  $\Lambda$  dependence of  $H$  by demanding the existence of a cutoff independent limit. From this we reconstruct the  $\beta$  function. Then we argue that at the duplex melting point the  $\beta$  function of the dimensionless scaled three-body interaction parameter  $H$  has the same form as that of Eq. (5.6) and the limit cycle describes the three-body bound Efimov states.

The nonconformity with the standard picture of fixed points has a far reaching consequence of converting the continuous scaling symmetry at the unstable fixed point to a discrete symmetry. The continuous scaling symmetry at a real fixed point leads to power law behaviors of physical quantities. Contrary to that, complex fixed points invoke a limit-cycle-type behavior in the RG flow trajectories. An outcome of the generated periodicity is a discrete scaling symmetry and the relevant parameter, here  $H$ , repeats itself in a log periodic manner. In the quantum language, this discrete symmetry leads to the Efimov tower of the energies.

## 5.2 Three Chain Problem With Bubbles: $g_2 \neq 0$

Here we consider the full Efimov-DNA problem allowing thermal-fluctuation generated bubbles in the bound state. Here we always consider situations where any two of the three chains have formed a duplex and the other free chain is interacting with that duplex. This consideration simplifies the problem immensely. We formulate our analysis at the two-chain melting point  $g_2 = g_{2c}$  to find the three-chain partition function. From this partition function the effective three-body coupling at the duplex melting point can be determined. There are no small parameters in the problem and therefore we need to sum terms up to infinite order or equivalently solve the integral equation shown diagrammatically in Fig 5.2.

Let us generalize the effective interaction  $V$  of Sec. 4.4 to a three-chain vertex function as  $W$  (see Appendix E), which in general depends on the input and the output momenta and the  $s$  values [Figure 5.2(a)]. Two successive Y forks producing

a strand exchange at a small separation would look like a three-chain interaction (see Figure 5.2(c)). This is an  $O(g_2^2)$  term. One may also couple this strand-exchanged configuration to the rest of the three-chain interactions, Figure 5.2(e), generating a term of  $O(g_2^2 W)$ . The  $g_3$ -dependent terms of Figure 4.3 also occur but with the replacement of the bound propagator (unfilled rectangles) by that of the duplex (filled rectangles), Figure 5.2(b) and 5.2(d). By combining all these, we have,

$$\begin{aligned}
W(\mathbf{k}, \mathbf{k}', s_1, s'_1, s) &= 2g_2^2 Z(\mathbf{k} + \mathbf{k}', s - s_1 - s'_1) - g_3 \\
&+ 2g_2^2 \int \frac{d\mathbf{q}}{(2\pi)^3} \frac{d\bar{s}}{2\pi i} Z(\mathbf{q}, \bar{s}) Z(\mathbf{k} + \mathbf{q}, s - s_1 - \bar{s}) \\
&\times Z_d(-\mathbf{q}, s - \bar{s}) W(\mathbf{q}, \mathbf{k}', \bar{s}, s'_1, s) \\
&- g_3 \int \frac{d\mathbf{q}}{(2\pi)^3} \frac{d\bar{s}}{2\pi i} Z(\mathbf{q}, \bar{s}) Z_d(-\mathbf{q}, s - \bar{s}) W(\mathbf{q}, \mathbf{k}', \bar{s}, s'_1, s). \quad (5.7)
\end{aligned}$$

Notice the factor of 2 in the diagrams with strand exchange because the chains are distinguishable [140].

If we do the  $\bar{s}$  integration by residues, the only contribution is from the pole of  $Z(\mathbf{q}, \bar{s})$  at  $\bar{s} = \Lambda^2 \ln \mu - q^2/2$ . This relation between  $\bar{s}$  and  $\mathbf{q}$  is analogous to the real space relation for size [Eq. (4.3)], which means the free chain is in a relaxed state [141]. Small distortions around the average size of a free polymer in equilibrium can be described by the Gaussian distribution around its average. Therefore this residue guarantees that no special large stretching takes place in a strand-exchange and the free chain remains more or less like an average chain. So we have

$$\begin{aligned}
W(\mathbf{k}, \mathbf{k}', s_1, s'_1, s) &= 2g_2^2 Z(\mathbf{k} + \mathbf{k}', s - s_1 - s'_1) - g_3 \\
&+ \int \frac{d^3 q}{(2\pi)^3} \left( \frac{2g_2^2}{s - s_1 - 2\Lambda^2 \ln \mu + q^2/2 + (k^2 + q^2)/2 + \mathbf{k} \cdot \mathbf{q}} - g_3 \right) \\
&\times W(\mathbf{q}, \mathbf{k}', s'_1, s) Z_d(-\mathbf{q}, s - \Lambda^2 \ln \mu + \mathbf{q}^2/2). \quad (5.8)
\end{aligned}$$

To simplify let us do the angle averaging, i.e. replacing  $Z_d$  by Eq. (4.22), so that  $W, Z_d$  are functions of the magnitudes of the wave vectors. The remaining angular integral from  $\mathbf{k} \cdot \mathbf{q}$  can be done. Assuming the external single chains are in their relaxed states such that  $s_1 = \Lambda^2 \ln \mu - k^2/2$  and  $s'_1 = \Lambda^2 \ln \mu - k'^2/2$  we have the angle averaged partition function

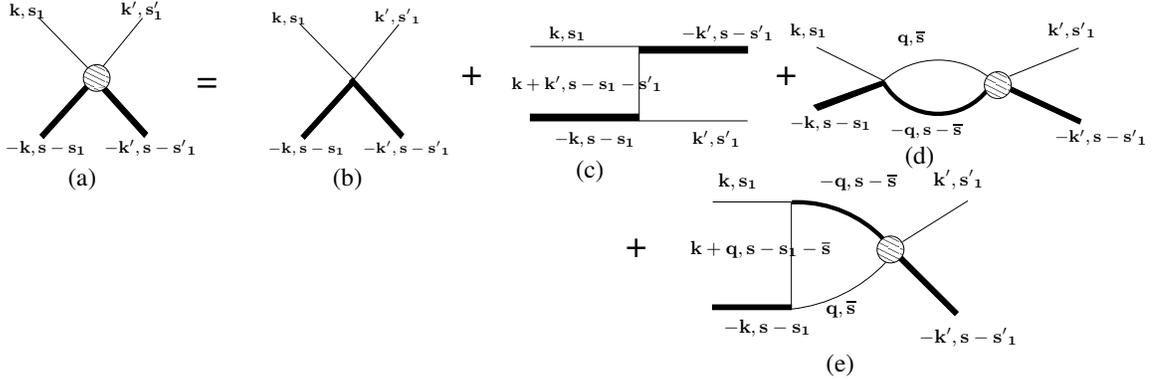


Figure 5.2: Diagrammatic representation of the three-chain partition function. The hatched circle is the effective interaction  $W$ . This figure translates into an integral equation involving interactions to all order.

$$\begin{aligned}
W(k, k') &= \frac{g_2^2}{kk'} \ln \frac{s'' + k^2 + k'^2 + kk'}{s'' + k^2 + k'^2 - kk'} - g_3 \\
&+ 4\pi \int_0^\Lambda \frac{dq}{(2\pi)^3} q^2 \left( \frac{g_2^2}{qk} \ln \frac{s'' + q^2 + k^2 + qk}{s'' + q^2 + k^2 - qk} - g_3 \right) \\
&\quad \times W(q, k') Z_d(-q, s + q^2/2),
\end{aligned} \tag{5.9}$$

where  $s'' = s - 3\Lambda^2 \ln \mu$ .

### 5.2.1 Critical Case: $g_2 = g_{2c}$

The limit required here is  $s'' \rightarrow 0$  which asserts that all chains are critical simultaneously as we have three chains now. To make sure that we are around the two-body critical point we take the limit  $\xi \rightarrow \infty$  in  $Z_d(-q, s + q^2/2)$ , Eq. (4.21). In this limit we expect to have significant contributions from the loop diagrams, and so we neglect the tree diagrams. By defining dimensionless quantities, like  $H$  in Eq. (5.2),

$$\bar{W}(q, k') = qW(q, k'), \tag{5.10}$$

and using Eq. (4.22), at the melting point, we have

$$\bar{W}(k, k') = \frac{8}{\sqrt{3}\pi} \int_0^\Lambda \frac{dq}{q} \left[ \ln \frac{q^2 + kq + k^2}{q^2 - kq + k^2} + 2kq \frac{H(\Lambda)}{\Lambda^2} \right] \bar{W}(q, k'). \tag{5.11}$$

The main reason behind the difference in the form of Eqs. (5.11) and (4.23) lies in the criticality of the duplex partition function used here. Since it is possible to consider

the  $k' \rightarrow 0$  limit of Eq. (5.11), the functional dependence of  $W$  on  $k'$  is not important for our calculations. We therefore suppress  $k'$  hereafter.

### 5.2.2 Scale-free Limit

In the limit  $H \rightarrow 0$  and  $\Lambda \rightarrow \infty$  there is no scale left in the problem because  $g_2$  has already been tuned to its critical value where  $\xi \rightarrow \infty$ . In this scale-free limit, the eigen-function type equation for  $\overline{W}$  is

$$\overline{W}(k) = \mathcal{I}_{k,q} \overline{W}(q) \equiv \frac{8}{\sqrt{3}\pi} \int_0^\infty \frac{dq}{q} \left[ \ln \frac{q^2 + kq + k^2}{q^2 - kq + k^2} \right] \overline{W}(q). \quad (5.12)$$

Since  $\overline{W}$  is dimensionless, a manifestly dimensionless form of Eq. (5.12) is obtained by replacing  $k, q$  by  $\hat{k} = k/\Lambda_*, \hat{q} = q/\Lambda_*$ , for some arbitrary  $\Lambda_*$ . Furthermore there is a large- $k$  – small- $k$  duality of integral operator  $\mathcal{I}_{\hat{k},\hat{q}}$  which suggests two degenerate solutions for Eq. (5.12). This is a consequence of the invariance of the integral operator under a transformation

$$\hat{Q} = \frac{\Lambda_*}{q}, \hat{K} = \frac{\Lambda_*}{k}, \quad \text{with} \quad \mathcal{I}_{\hat{k},\hat{q}} \equiv \mathcal{I}_{\hat{K},\hat{Q}}, \quad (5.13)$$

Thus if  $f(k/\Lambda_*)$  is an eigenfunction of  $\mathcal{I}_{k,q}$ , then so is  $f(\Lambda_*/k)$ . The general solution of  $\overline{W}(k)$  can then be written as a sum of the two degenerate solutions.

Taking note of the scale-free form, we can have a power law ansatz

$$\overline{W}(k) \approx \left( \frac{k}{\Lambda_*} \right)^s, \quad (5.14)$$

which on substitution in Eq. (5.12) yields

$$s = \frac{16}{\sqrt{3}} \frac{\sin(\pi s/6)}{\cos(\pi s/2)}. \quad (5.15)$$

This equation has solutions for pure imaginary values,  $s = \pm i s_0$ , with

$$s_0 = 1.5036, \quad \text{or} \quad \exp(\pi/s_0) = 8.0713\dots, \quad (5.16)$$

which is different from 22.7 obtained by Efimov.

The solution for  $\overline{W}$  is a linear combination of  $\exp[\pm i s_0 \ln(k/\Lambda_*)]$ , which can be recast in a trigonometric form

$$\overline{W}(k) = C \cos \left( s_0 \ln \frac{k}{\Lambda_*} \right), \quad (5.17)$$

with  $C, \Lambda_*$  as two arbitrary constants.

### 5.2.3 For $\Lambda < \infty$

In the general case, ( $\Lambda < \infty$ ) we may still proceed to find the  $\Lambda$  dependence of  $H(\Lambda)$  by assuming that  $\overline{W}$  approximately retains its form ( as in Eq. (5.17)) by changing only its constants [71].

Defining the function  $f(\Lambda) = \frac{H(\Lambda)}{\Lambda^2}$  we can rewrite Eq. (5.11) as

$$\overline{W}(k) = \frac{8}{\sqrt{3\pi}} \int_0^\Lambda \frac{dq}{q} \left[ \ln \frac{q^2 + kq + k^2}{q^2 - kq + k^2} + 2kqf(\Lambda) \right] \overline{W}(q). \quad (5.18)$$

$\overline{W}$  is related to the third virial coefficient of the system. For this  $\overline{W}$  must be independent of  $\Lambda$  which is introduced arbitrarily. We take advantage of this fact to compare the value of  $\overline{W}$  for two infinitesimally different  $\Lambda$ s. The cutoff independence is preserved by equating the residual pieces to zero. By integrating over a small shell of radius  $\Lambda dl$  we have

$$\overline{W}(k) = \frac{8}{\sqrt{3\pi}} \int_0^{\Lambda e^{-dl}} \frac{dq}{q} \left[ \ln \frac{q^2 + kq + k^2}{q^2 - kq + k^2} + 2kqf(\Lambda) \right] \overline{W}(q) \quad (5.19)$$

$$+ \frac{8}{\sqrt{3\pi}} \left[ \ln \frac{\Lambda^2 + k\Lambda + k^2}{\Lambda^2 - k\Lambda + k^2} + 2k\Lambda f(\Lambda) \right] \overline{W}(\Lambda) dl. \quad (5.20)$$

Rescaling back  $\Lambda \rightarrow \Lambda e^{dl}$  and retaining terms up to order  $dl$  we have

$$\begin{aligned} \overline{W}(k) &= \frac{8}{\sqrt{3\pi}} \int_0^\Lambda \frac{dq}{q} \left[ \ln \frac{q^2 + kq + k^2}{q^2 - kq + k^2} + 2kqf(\Lambda) \right] \overline{W}(q) \\ &+ dl \frac{8}{\sqrt{3\pi}} \left( 2k \frac{\partial f(\Lambda)}{\partial \Lambda} \int_0^\Lambda dq \overline{W}(q) + \left( \frac{2k}{\Lambda} + 2f(\Lambda)k\Lambda \right) \overline{W}(\Lambda) \right). \end{aligned} \quad (5.21)$$

In the previous step we used the approximation  $k \ll \Lambda$  such that

$$\ln \frac{\Lambda^2 + k\Lambda + k^2}{\Lambda^2 - k\Lambda + k^2} \approx \frac{2k}{\Lambda}, \quad (\text{for } k \ll \Lambda). \quad (5.22)$$

Now using Eq. (5.18) in Eq. (5.21) we can easily arrive at the differential equation

$$\frac{1}{\Lambda} \left[ \Lambda \frac{\partial H}{\partial \Lambda} - 2H \right] \int_0^\Lambda dq \overline{W}(q) + [1 + H] \overline{W}(\Lambda) = 0, \quad (5.23)$$

where  $d\Lambda = \Lambda dl$ . This equation for  $H$  is already of the form of Eq. (5.4b), except that the terms in addition to the naive  $2H$  term is dependent on  $\Lambda$ .

In principle, a renormalization-group  $\beta$  function is not expected to have any explicit cutoff dependence. The  $\Lambda$ -independent flow equation is derived below from the

full form for  $H(\Lambda)$ . To do so, by inserting  $\overline{W}(x) = C \cos(s \ln x)$ , where  $x = \frac{\Lambda}{\Lambda_*}$ , in that equation we obtain,

$$\frac{\partial H(x)}{\partial x} \frac{\sin(A+B)}{x} + \frac{H(x)}{x^2} (s_0 - 1) \sin(A+B) + \frac{s_0 \cos(A-B)}{x^2} - \frac{\sin(A-B)}{x^2} = 0, \quad (5.24)$$

where  $A = s_0 \ln x$  and  $B = \arctan(\frac{1}{s_0})$ . We can express the left hand side of the above equation as an exact differential,

$$\frac{\partial}{\partial x} \left[ \frac{H(x)}{x} \sin \left( s_0 \ln x + \arctan \left( \frac{1}{s_0} \right) \right) + \frac{1}{x} \sin \left( s_0 \ln x - \arctan \left( \frac{1}{s_0} \right) \right) \right] = 0. \quad (5.25)$$

Setting the boundary condition such that the integration constant vanishes, we get

$$H(\Lambda) = - \frac{\sin \left( s_0 \ln \frac{\Lambda}{\Lambda_*} - \arctan \left( \frac{1}{s_0} \right) \right)}{\sin \left( s_0 \ln \frac{\Lambda}{\Lambda_*} + \arctan \left( \frac{1}{s_0} \right) \right)}. \quad (5.26)$$

Having found out the  $\Lambda$  dependence of  $H$  we can now derive its RG flow equation by simply taking a derivative of Eq. (5.26). The  $\beta$  function of  $H$  is given by

$$\Lambda \frac{\partial H}{\partial \Lambda} = \beta(H) \equiv 2H - \frac{1}{2} (1 + s_0^2) (H + 1)^2, \quad (5.27)$$

which is of the form of Eq. (5.5), with

$$A = -C = \frac{1}{2} (1 + s_0^2), B = 1 - s_0^2. \quad (5.28)$$

The specialty of the flow equation, Eq. (5.27), is the emergence of complex conjugate fixed points,  $H_0, H_0^*$ , with

$$H_0 = \frac{(1 + is_0)}{(1 - is_0)}, \quad (5.29)$$

so that one recovers the form of Eq. (5.6). Further consequences are discussed below.

## 5.2.4 Complex Fixed Points and Periodicity

Because of the complex fixed points the flow of  $H$  consists of closed trajectories in the complex  $H$ -plane. This is at the duplex melting point, a fixed point for  $g_2$  in the renormalization group sense, and, therefore, the flows remain planar in the complex

$H$  plane. We define a new variable  $\zeta = (H - H_0)/(H - H_0^*)$  which is nothing but a conformal mapping of  $H$ . The flow equation of  $\zeta$  is the equation of a unit circle

$$\frac{\Lambda}{\zeta} \frac{\partial \zeta}{\partial \Lambda} = 2is_0. \quad (5.30)$$

In critical phenomena normally one expects the occurrence of real fixed points in the flow equation of relevant parameters. In this particular RG scheme, a stable fixed point serves as the critical point for the corresponding parameter. In the vicinity of the critical point the system is scale-free because the system length scale diverges with exponent  $\nu$  which is related to the difference of the fixed points. Unlike this, when we have a limit cycle, the continuous scaling symmetry breaks down and the relevant parameter is log periodic. A power law  $f(x) \sim x^{-\nu}$  for real  $\nu$ , is converted to an oscillatory form  $f(x) \sim e^{-i|\nu| \ln|x|}$  where  $\nu$  is imaginary ( $\nu = i|\nu|$ ). This occurs because the difference between the complex fixed points is a purely imaginary quantity now. A consequence of this is the log periodicity of  $H$  with respect to  $\Lambda$  in Eq. (5.26).

There is another convenient way to visualize the closed trajectories. Decompose  $H$  into its real and complex parts by writing  $H = H_1 + iH_2$  to get two interdependent differential equations:

$$\Lambda \frac{\partial H_1}{\partial \Lambda} = -\frac{1}{2} [s_0^2 \{(1 + H_1)^2 - H_2^2\} + (1 - H_1^2) - H_2^2], \quad (5.31a)$$

$$\Lambda \frac{\partial H_2}{\partial \Lambda} = (1 - H_1)H_2 - s_0^2(1 + H_1)H_2. \quad (5.31b)$$

By solving these two equations simultaneously for different initial values we get closed elliptical trajectories. All ellipses in the upper-half plane have one common focus at one complex fixed point  $H_0$  and the other foci are at different places. A similar thing happens in the lower-half plane too, with a common focus at  $H_0^*$ . The trajectories which start on the real line ( $H_2 = 0$ ) always stay on the real line. These are shown in Figure 5.3.

These closed loops change over to the real form of Eq. (4.28) as  $g_2$  is detuned from the critical point. A toy model for this smooth crossover is discussed in Appendix F that shows the specialty of the closed loop in the three-dimensional space of  $g_2$  and complex  $g_3$  (in dimensionless forms).

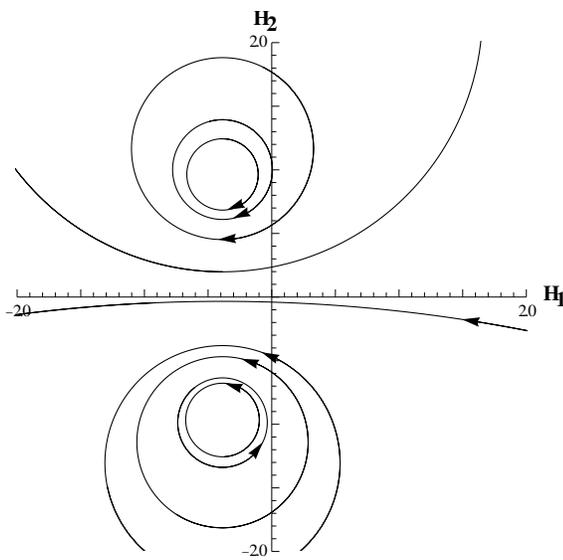


Figure 5.3: Closed elliptical trajectories in the complex  $H$  plane as  $\Lambda$  is varied. These are drawn for different starting values of  $(H_1, H_2)$ . All loops in the upper half plane have the fixed point  $(1 + is_0)/(1 - is_0)$  as a focus while  $(1 - is_0)/(1 + is_0)$  as a focus for the lower half plane.

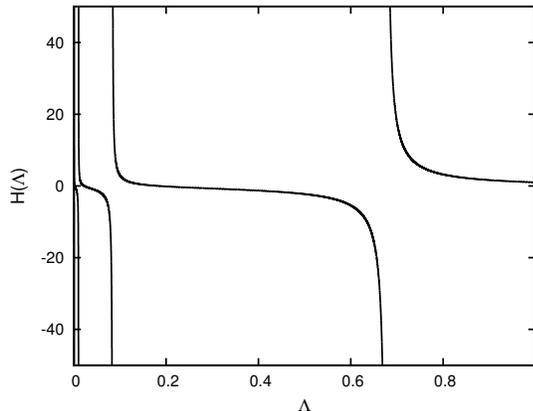


Figure 5.4: Plot of  $H$  as a function of  $\Lambda$  showing zeros and divergences with  $\Lambda_* = 1$ .

### 5.2.5 Discrete Scale Invariance

A simple inspection of Eq. (5.26) shows us that  $H$  is log periodic. As already mentioned, at the duplex melting point  $g_2$  has its fixed point value and so, from the definition of  $H$ , the dimensionless three-body interaction energy,  $\hat{g}_3 \sim g_3 \Lambda$  obeys a flow equation very similar to that of  $H$ . If we start from  $\hat{g}_3 = 0$  we arrive at the negative infinity as  $\Lambda$  is increased. At this point  $\hat{g}_3$  jumps to positive infinity and decreases to negative infinity again as  $\Lambda$  is increased further. This is shown in Figure 5.4. This behavior goes on and  $\hat{g}_3$  runs into negative infinity whenever the denominator of Eq. (5.26) becomes zero. This occurs at the points

$$\Lambda_n = \Lambda_* \left( e^{\frac{\pi}{s_0}} \right)^n \exp \left[ \frac{\arctan(s_0) - \frac{\pi}{2}}{s_0} \right], \quad (5.32)$$

where  $n$ 's are integers. We therefore see the emergence of a discrete scale invariance in this three-chain problem even though the melting itself or the three-chain interaction *per se* has no indication of this sort.

As  $\hat{g}_3$  can also be interpreted as the three-body binding energy, at those values of  $\Lambda$  we get the three-body Efimov bound states in the quantum case. The corresponding energy spectrum of the quantum three-particle system would follow a geometric relation given by

$$E_{n+1}/E_n = e^{-2\pi/s_0}, \quad (5.33)$$

where  $E_n$  is the  $n$ th energy state. So the energies of the Efimov states are related by a factor of  $e^{2\pi/s_0}$ . We observe from Eq. (5.30) that two successive windings around

the unit circle are also related through the factor  $e^{\pi/s_0}$ . As  $H$  has nontrivial values at these points we conclude that every jump from one energy level to the next one corresponds to one winding of  $\zeta$  around the closed trajectory. It can also be observed that  $\hat{g}_3$  goes to zero at the points

$$\Lambda_n = \Lambda_* \left( e^{\frac{\pi}{s_0}} \right)^n \exp \left[ \frac{\frac{\pi}{2} - \arctan(s_0)}{s_0} \right], \quad (5.34)$$

where the numerator of Eq. (5.26) becomes zero. At these special points we do not have to introduce three-body coupling. So, in this picture every jump corresponds to a switching of one Efimov state to another one and it is associated with a complete winding around a limit cycle. These states are crowded more and more as one goes in the direction of zero energy. They are infinite in number.

To summarize, we see that each of  $g_2$  and  $g_3$ , acting alone on its own, allows critical points in the form of melting or dissociation, well described by the conventional renormalization group fixed points. These points show a continuous scale invariance; under a rescaling of the system by any factor,  $L \rightarrow bL$  for any  $b$ , a critical system remains critical, statistically identical. In contrast, at such a fixed point for  $g_2$ ,  $g_3$  shows a cyclic behavior, better described as a “limit cycle” behavior, in the complex plane, because of the emergence of a periodicity. The log periodicity induces a discrete scale invariance,  $L \rightarrow b_n L$  for a discrete set  $b_n$ , breaking the continuous symmetry expected at the critical value of  $g_2$  for two chains. In quantum mechanics, this leads to an infinite set of energy eigenstates in a three-particle system at the point where the energy of any pair should have been zero. In the context of DNA, at the melting point of a double stranded DNA where the strands are not bound to each other, a third strand induces a binding of a size much larger than the hydrogen bond length. In other words the infinite correlation length scale of a duplex DNA gets transmuted to a finite value in the presence of a third one when each pair is supposed to be critical.

### 5.3 Off-critical: $g_2 \neq g_{2c}$

So far we have considered the case of the critical two-chain case. For the general situation,  $g_2 \neq g_{2c}$ , we need to go back to Eq. (5.9), and we also need the flow equation

for  $g_2$ . Instead, we may take a heuristic approach. In terms of the dimensionless two- and three-body constants  $\hat{g}_3, \hat{g}_2 \sim g_2 \Lambda^{-1/2}$ , with  $H \sim \hat{g}_3 / \hat{g}_2^2$ , the expected form of the RG flow equation for  $g_2$  is

$$\Lambda \frac{\partial \hat{g}_2}{\partial \Lambda} = \beta_2(\hat{g}_2), \quad (5.35)$$

with an unstable fixed point  $\hat{g}_2 = 0$  for the bound phase and a stable one at  $\hat{g}_2 = \hat{g}_c$  for the duplex melting. At these points  $\beta_2(\hat{g}_2) = 0$ . With these, we may formally write

$$\Lambda \frac{\partial \hat{g}_3}{\partial \Lambda} = \frac{2\hat{g}_3}{\hat{g}_2} \beta_2(\hat{g}_2) - 4 \hat{g}_2^2 \beta(H, \hat{g}_2), \quad (5.36)$$

relating the  $\beta$  function for  $\hat{g}_3$  with the others. As expected, for the duplex melting point with  $\beta_2(\hat{g}_2^*) = 0$ , the flow of  $\hat{g}_3$  is the same as that of  $H$  and therefore  $\hat{g}_3$  is to be described by the pair of complex fixed points. However, for the  $\hat{g}_2 = 0$  fixed point, to get back Eq. (4.28),  $\beta(H)$  of Eq. (5.27) is not sufficient because it does not yield a  $\hat{g}_2$ -independent limit. This indicates that the contributions of the off-critical terms in  $W$ , Eq. (5.9), are important. This can be taken as a signal that the periodicity that develops at the duplex melting point for  $g_3$  or  $\hat{g}_3$  do not survive in the off-critical limit. A flow diagram is shown in Figure 5.5 which depicts a few cycles before merging with the flow at  $g_2 = g_{2c}$ .

## 5.4 Discussion

At this point we would like to place the results of this paper in a broader context. We do so in three different contexts, namely (i) as a DNA problem, (ii) as an Efimov effect, and (iii) more formally as a renormalization-group problem. Let us first discuss the last two issues, as elaborated upon in the Introduction. The results, in conjunction with the previous works, provide a possible testing ground for the quantum Efimov physics in a classical environment, namely the melting of DNA which occurs at temperatures in the range 60–100C. Here thermal fluctuations play the role of quantum fluctuations [142]. It was shown earlier that the fluctuation induced long range inverse-square attraction has a natural basis in the polymer scaling. Here we showed how the polymer phase transitions in the various limiting situations allowed us to construct an extrapolation formula for the renormalization group  $\beta$  function that shows the development of the limit cycle behavior. This is also important in

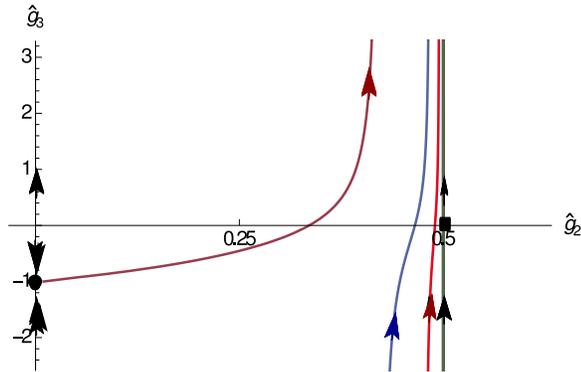


Figure 5.5: (Color online) Schematic flow diagram in the  $\hat{g}_2$ - $\hat{g}_3$  plane.  $\hat{g}_2 = 0$  corresponds to the bound state (no bubbles) while  $\hat{g}_2 = \hat{g}_2^*$  (box) is the duplex melting point. The flow along the thick vertical line through  $\hat{g}_2 = \hat{g}_2^*$  is periodic. Along the  $\hat{g}_2 = 0$  line, there is a stable fixed point  $\hat{g}_3 = \hat{g}_{3c}$  (filled disk) which represents the peeling of one polymer from the rigid bound state. A typical flow from a point away from the melting point is shown.

the general theory of renormalization group where examples of limit cycles are rather few.

Let us now look at it as a DNA problem. The paper builds on the model of a stiff duplex [138] and extends the study of the third virial coefficient of the three-polymer system to the region around the duplex melting point (at temperature  $T_c$ ). This extension, which goes beyond Ref. [138], shows that the fluctuation-induced Efimov DNA is not just a specialty of the melting point but it also exists over a region where the duplex should have been unbound. Although the fractal-like lattices of Ref. [67–69] showed the possibility of the three-chain thermodynamic phase, the lack of a metric or distance forbade any analysis of the inverse square law attraction responsible for the Efimov effect. This gap is now partly filled by the analysis of this paper. Short of a direct proof of the inverse-square attraction, the limit cycle behavior is similar to the complex fixed points known in systems with such long range interaction.

There are several issues which might be amenable to experimental verification. The third strand could be made of alternating short sequences of both the strands. (i) A direct test of the Efimov-effect in DNA would be a measurement of the melting

temperature  $T_t$  of the triple-chain system to see if  $T_t > T_c$ . The melting is expected to be first-order in nature. The difficulty of course lies in separating the melting of the Efimov DNA from that of a Watson-Crick and Hoogsteen paired triple-stranded DNA.

(ii) A different thermodynamic study would be the virial coefficients via the osmotic pressure of dilute solutions [143]. The third virial coefficient will give  $H(\Lambda)$ , Eq. (5.2). Near the melting of duplex, the cutoff parameter may be chosen as the bubble length scale  $\xi$ , Eq. (4.21). Therefore, for  $\Lambda \sim \xi^{-1} \sim \sqrt{\Delta t}$ , the virial coefficient is expected to show an oscillatory behavior whose periodicity is determined by the (non-universal) Efimov number  $s_0$ .

(iii) More detailed information might be obtained from small angle neutron scattering (SANS) or light scattering. One of the signatures of the fluctuation induced long range interaction is the  $1/k$  divergence of  $W(k)$  in Eq. (5.10), since  $\overline{W}$  is bounded. Such a divergence in the interaction is observable as a zero- $k$  peak in the scattered intensity in SANS or light scattering [144].

(iv) It is possible to design tailor-made environments where one might test some of the details that have gone in the theory. For example, one may use the geometry of two similar strands maintained at a distance larger than the hydrogen bond length to prevent direct pairing and then allow a complementary strand to form bonds via strand exchange (Figure 5.1) with both the strands. The force needed to unzip one of the original strands can then be measured to verify the inverse square nature. Such experiments will be similar in spirit to the measurement of the fluctuation induced Casimir force or the entropic interaction in DNA solution [145, 146].

(v) An assumption that has gone in the theory, actually in most field theoretic calculations (see Ref. [141]), is that the third chain between two contacts with the other two strands is in a relaxed state. This condition may be verified by using a carefully constructed bubble and then placing a third chain. A labeled chain (with say heavier isotopes) will help in separating out the scattering from this chain and provide information on its configuration.

(vi) A different experiment would be to study the DNA in a narrow pore. The constraint prevents the large loop formation, thereby cutting off the long range interaction. In this situation, the Efimov-DNA formation is not likely to happen but a novel finite size effect would be expected [147]. The effects would be observable in the configuration of the labeled chain and even in thermodynamic quantities. Such experiments of DNA in a pore has been attempted but not at the level that may explore the large loops

near DNA melting [148]. (vii) One may go beyond the three-chain problem to the possibility of a four or more chain bound state, or even a gel formation in a many strand solution through the Efimov interaction. In such a gel, just above the duplex melting point, the mesh size of the network would be similar to the bubble size. A theoretical study of this gelation phenomenon and the elastic properties of the gel remain an open and interesting topic.

## 5.5 Conclusion

This paper presents a model of a three-stranded DNA as a three-chain polymer system, which shows mathematically analogous results as that of Efimov physics, namely, the possibility of a three-chain bound state when no two are bound. The existence of a three-stranded DNA bound state (Efimov DNA) at the duplex DNA melting point is shown here analytically by a renormalization group approach. To achieve this, a nonperturbative momentum-shell type RG procedure is employed. We studied the duplex-DNA melting by introducing a rigid chain model, where the melting is induced by an interfacial term. A completely bound two-stranded state at zero temperature is the zero temperature configuration. At finite temperatures thermal fluctuations locally denature the bound state to form bubbles made of two free-chain pairs. A third similar strand, when added, can again form a duplex with one or both of the free chains of a bubble. Due to renormalization of short range interactions close to the duplex melting point an effective long range three-chain interaction is generated. The Efimov DNA is a result of this. Just as in the quantum Efimov problem, we show that the Efimov-DNA is associated with a limit cycle behavior of the RG flow of the generated three-chain interaction. Since the interaction parameters for a DNA are easily tunable, by choosing solvent quality, we hope our results would motivate experiments in detecting the Efimov effect in polymeric systems.

# Appendix F

## A simple $\beta$ function for $\hat{g}_3$

We propose an extrapolation formula that connects smoothly the flows for the  $\hat{g}_2 = 0$  case to the  $\hat{g}_2 = \hat{g}_2^*$  flows. This is a toy example to amplify the limit cycle behavior in a three-dimensional parameter space, namely,  $\hat{g}_2$  and the real and imaginary parts of  $\hat{g}_3$ . For simplicity we choose,  $s_0 = 1$ .

Take

$$\beta(\hat{g}_2) = \frac{4-d}{2}\hat{g}_2 - \hat{g}_2^2, \quad (\text{F.1})$$

with an unstable fixed point at  $\hat{g}_2 = 0$  and a stable fixed point at  $\hat{g}_2 = 1/2$ .

Use this to write for  $d = 3$  [Eq.(5.36)]

$$\beta(\hat{g}_3, \hat{g}_2) = \hat{g}_3 - 2\hat{g}_3\hat{g}_2 - 4\hat{g}_2^2 (2H - F_2(\hat{g}_2)H^2 - 2H - 1), \quad (\text{F.2})$$

where, with  $H = -\hat{g}_3/(4\hat{g}_2^2)$  as before, the  $2H$  term is made explicit for comparison with Eq. (5.4b), and we defined  $F_2(\hat{g}_2) = 4\hat{g}_2^2$ . With these choices, we recover both the limit of  $\hat{g}_2 = 0$  and  $\hat{g}_2 = 1/2$  the critical melting point.

The flow diagram in the three-dimensional space of  $\hat{g}_2, \text{Re}(\hat{g}_3), \text{Im}(\hat{g}_3)$  shows the approach to the planar cycle at the melting point. This is shown in Figure F.1. So long as the flow is controlled by the real fixed points, we see a monotonic flow. As  $\hat{g}_2$  changes, the complex fixed points take over and we get the loops.

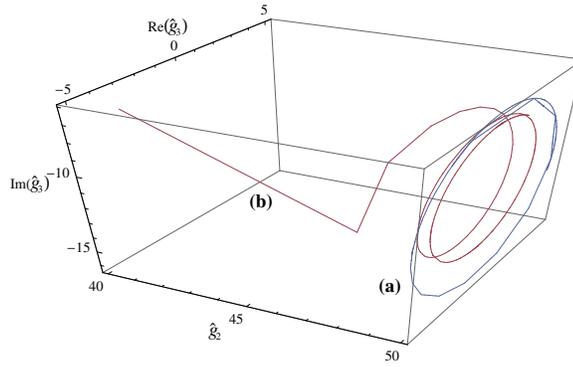


Figure F.1: (Color online) A schematic diagram of the flow of the RG equation, Eq. (F.2), showing the approach to the limit cycle. The limit cycle is an ellipse in the  $\hat{g}_2 = 0.5$  plane [blue line marked (a)]. A point with off-critical  $\hat{g}_2$  is shown to approach the planar limit cycle as  $\Lambda \rightarrow \infty$  [red line marked (b)].

# Chapter 6

## Summary

In this thesis, we studied two aspects of the physics of DNA, the temperature dependence of the DNA elastic modulus and the connection between the recently proposed Efimov-DNA and the renormalization group limit cycle. Techniques employed in this work are the generating function method, the numerical transfer matrix method and the renormalization group (RG) method.

Using a DNA model with entropic elasticity only, we studied the temperature dependence of the corresponding elastic properties for different stretching forces. After obtaining the results analytically we checked their consistency with the corresponding numerical results. For a DNA, above its melting point, a second order binding-unbinding phase transition takes place under the stretching force. When the system length is infinite, there is no pre-transition effect as the elastic modulus shows a sharp finite discontinuity at the transition point. On the contrary, in case of finite systems, the transition influences the elastic behavior away from the transition point. When the elastic modulus is obtained for this system in the presence of an unzipping force it shows a  $\delta$ -function singularity and the transition becomes first order. We also obtained the three dimensional temperature-stretching-unzipping force phase diagram.

We introduced a DNA model with intrinsic rigidity and studied its elastic properties for different temperatures employing a stretching force. For the analysis of the model we used an exact numerical transfer matrix method. The binding-unbinding transition is again second order but the elastic modulus shows an anomalous behavior around the transition point in the sense that the bound state elastic modulus

becomes much larger than the unbound state modulus. The influence of the thermal fluctuations is also explored by studying different bubble related quantities like bubble number, bubble length etc. Our analysis shows that for zero force, the extension and the elastic modulus vary with the square root of bubble number and the square root of bubble number fluctuation respectively.

To study the three-chain Efimov-DNA problem, we introduced a Poland-Scheraga type model of DNA melting by considering the bound segments as completely rigid rods without bending and the bubbles as flexible Gaussian polymer loops. Using the generating function method we showed that this model goes through a continuous thermal melting transition. The finite temperature bound state of this model is used to represent a finite temperature duplex DNA. By considering the interaction between a completely rigid bound state and a single chain we have shown through RG method that Efimov-DNA can not appear in the absence of thermal bubbles in the duplex DNA.

We solved the Efimov-DNA problem from the RG perspective. Using the strand exchange mechanism between a duplex DNA and a single strand, the third virial coefficient is calculated by employing a non-perturbative momentum shell RG method. We have shown that the Efimov-DNA corresponds to the RG limit cycle trajectories of a three-body parameter. Each winding of the limit cycle corresponds to an Efimov-trimer which repeats itself log periodically and for infinitely many times at the critical point. We also proposed a toy RG flow equation for the off critical temperatures which shows the number of Efimov-trimers decreases as one goes away from the critical point. The temperature at which the number becomes zero is the Efimov-DNA melting temperature which is greater than the duplex melting point. The possible experimental scenarios where the Efimov-DNA may be detected are also discussed in detail.

# Bibliography

- [1] J. D. Watson, *Molecular Biology of the Gene*, 7th ed. (Pearson, Cold Spring Harbor, NY, 2013).
- [2] Bruce Alberts *et al*, *Molecular Biology of The cell*, 6th ed. (Garland Science, NY, 2014).
- [3] Richard R. Sinden, *DNA Structure and Function*, 1st ed. (Academic Press Inc., San Diego, California, 1994).
- [4] E. Schrödinger, *What is Life?*, Cambridge University Press (1944).
- [5] J. D. Watson and F. H. C. Crick, *Nature (London)* **171**, 737 (1953).
- [6] G. Felsenfeld, D. R. Davies and A. Reich, *J. Am. Chem. Soc.* **79**, 2023 (1957).
- [7] M. D. Frank-Kamenetskii and S. M. Mirkin, *Annu. Rev. Biochem.* **64**, 65 (1995).
- [8] I. Radhakrishnan and D. J. Patel, *Biochemistry* **33**, 11405 (1994).
- [9] E. N. Nikolova, E. Kim, A.A. Wise, P.J. O'Brien, I. Andricioaei and H.M. Al-Hashimi, *Nature (London)* **470** 498 (2011).
- [10] V. I. Lymichev, S. M. Mirkin and M. D. Frank-Kamenetskii, *J. Biomol. Struct. Dynam.* **3**, 327 (1985).
- [11] V. I. Lymichev, S. M. Mirkin and M. D. Frank-Kamenetskii, *J. Biomol. Struct. Dynam.* **3**, 667 (1985).
- [12] S. M. Mirkin *et al.*, *Nature (London)* **330**, 495 (1987).
- [13] P. Parniewski *et al.*, *Nucleic Acids Res.* **18**, 605 (1990).

- [14] M. J. Ulrich *et al.*, J. Biol. Chem. **267**, 18649 (1992).
- [15] B. T. Brinton *et al.*, J. Biol. Chem. **266**, 5153 (1991).
- [16] See, e.g., M. J. Neale and S. Keeney, Nature (London)**442**, 153 (2006).
- [17] L. V. Yakusheich, *Nonlinear Physics of DNA*, Wiley, Chichester (1998).
- [18] J. Yan and J. F. Marko, Phys. Rev. Lett. **93**, 108108 (2004).
- [19] P. Sadhukhan and S. M. Bhattacharjee, Ind. J. Physics, **88**, 895 (2014).
- [20] M. Ageno, E. Dore and C. Frontali, Biophys. J. **9**, 1281 (1969).
- [21] A. Kornberg and T. A. Baker, *DNA Replication*, W. H. Freeman and Company, New York (1992).
- [22] J. Marmur and P. Doty, J. Mol. biol. **5**, 109 (1962).
- [23] R. M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985).
- [24] O. Gotoh, Adv. Biophys **16**, 1 (1983).
- [25] M. Daune, *Molecular Biophysics : Structures in Motion* (Oxford University Press, 1999)
- [26] D. Marenduzzo *et al.*, Phys. Rev. Lett. **88**, 028102 (2001).
- [27] B. H. Zimm, J. Chem. Phys **33**, 1349 (1960).
- [28] D. Poland and H. Scheraga, J. Chem. Phys. **45**, 1456 (1966).
- [29] M. E. Fisher, J. Chem. Phys. **45**, 1469 (1966).
- [30] M. Peyrard and A. R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989).
- [31] Y. Kafri, D. Mukamel and L. Peliti, Phys. Rev. Lett. **85**, 4988 (2000).
- [32] Y. Kafri, D. Mukamel and L. Peliti, eur. Phys. J. B **27**, 132 (2002).
- [33] M. Y. Azbel, Phys. Rev. A **20**, 1671 (1979).
- [34] S. S. Patel and K. M. Picha, Annu. Rev. Biochem. **69**, 651 (2000).

- [35] D. Fass, C. E. Bogden and J. M. Berger, *Structure* **7**, 691 (1999).
- [36] S. S. Velankar *et al.*, *Cell* **97**, 75 (1999).
- [37] P. Soultanas *et al.*, *EMBO J.* **19**, 3799 (2000).
- [38] M. R. Singleton *et al.*, *Cell* **107**, 79 (2001).
- [39] S. M. Bhattacharjee, *Europhys. Lett.* **65**, 574 (2004).
- [40] C. Bustamante, S. Smith, J. Liphardt and D. Smit, *Current Opinion in Str. Biol.* **10**, 279 (2000).
- [41] C. Bustamante, Z. Bryant and S. Smith, *Nature* **421**, 423 (2003).
- [42] F. Ritort, *J. Phys. Condens. Matter* **18**, R531 (2006).
- [43] S Kumar and M S Li, *Phys. Rept.* **486**, 1 (2010).
- [44] S. B. Smith, Y. Cui and C. Bustamante, *Science* **271**, 795 (1996).
- [45] J. M. Huguet *et al.*, *Proc. Natl. Acad. Sci.* **107**, 15431 (2010).
- [46] S. M. Bhattacharjee, *J. Phys. A* **33**, L423 (2000); **33**, 9003(E) (2000).
- [47] K. L. Sebastian, *Phys. Rev. E* **62**, 1128 (2000).
- [48] D. Marenduzzo *et al.*, *Phys. Rev. E* **64**, 031901 (2001).
- [49] D. Marenduzzo *et al.*, *Phys. Rev. Lett* **88**, 028102 (2002).
- [50] J. Z. Y. Chen, *Phys. Rev. E* **66**, 031912 (2002).
- [51] D. K. Lubensky and D. R. Nelson, *Phys. Rev. E* **65**, 031917 (2002).
- [52] R. Kapri, S. M. Bhattacharjee and F. Seno, *Phys. Rev. Lett.* **93**, 248102 (2004).
- [53] S. Kumar, D. Giri and S. M. Bhattacharjee, *Phys. Rev. E* **71**, 051804 (2005).
- [54] R. Kapri, *Phys. Rev. E* **86**, 041906 (2012).
- [55] S. Kumar and G. Mishra, *Phys. Rev. Lett.* **110**, 258102 (2013).

- [56] G. Mishra *et al.*, Phys. Rev. E **87**, 022718 (2013).
- [57] R. Kapri, Phys. Rev. E **90**, 062719 (2014).
- [58] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).
- [59] S. M. Bhattacharjee, A. Giacometti and A. Maritan, J. Phys.: C ondens. Matter **25**, 503101 (2013).
- [60] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995).
- [61] If two random walkers or directed polymers starting at one spatial point rejoin at another point (including the starting point) then it is called a reunion. The corresponding partition function of the bubble decays as a power law in the large length limit,  $Z_R \sim N^{-\psi_R}$ . This exponent  $\psi_R$  is called the reunion exponent, which plays an important role in determining the order of the melting transitions.
- [62] V. Efimov, Phys. Lett. **B33**, 563 (1970).
- [63] V. Efimov, Sov. J. Nucl. Phys. **12**, 589 (1971).
- [64] V. Efimov, Sov. J. Nucl. Phys. **29**, 546 (1979).
- [65] E. Nielsen, D.V. Fedorov, A.S. Jensen, E. Garrido, Phys. Rept. **347**, 373 (2001).
- [66] E. Braaten, H. W. Hammer, Phys. Rep. **428**, 259 (2006).
- [67] J. Maji, S. M. Bhattacharjee, F. Seno, A. Trovato; New J. Phys. **12**, 083057 (2010).
- [68] J. Maji, S. M. Bhattacharjee, Phys. Rev. E **86**, 041147 (2012).
- [69] J. Maji, S. M. Bhattacharjee, F. Seno and A. Trovato, Phys Rev E **89**, 012121 (2014).
- [70] A. Fonseca, E. Redish, and P. E. Shanley, Nucl. Phys. **A320**, 273 (1979).
- [71] P. F. Bedaque, H. -W. Hammer, U. van Kolck, Nucl. Phys **A646**, 444 (1999).
- [72] Y. Horinouchi and M. Ueda, Phys. Rev. Lett. **114**, 025301 (2015).

- [73] S. D. Glazek, K. G. Wilson, Phys. Rev. D **48**, 5863 (1993); Phys. Rev. Lett. **89**, 230401 (2002).
- [74] N T Zinner and A S Jensen, J. Phys. G: Nucl. Part. Phys. **40**, 053101 (2013).
- [75] M. Zaccanti, B. Deissler, C. D’Errico, M. Fattori, M. Jona-Lasinio, S. Müller, G. Roati, M. Inguscio and G. Modugno, Nat. Phys. **5**, 586 (2009).
- [76] S. Knoop, F. Ferlaino, M. Mark, M. Berninger, H. Schöbel, H. C. Nägerl, R. Grimm, Nat. Phys. **5**, 227 (2009).
- [77] S. M. Bhattacharjee and S. Mukherji, Phys. Rev. Lett. **83**, 2374 (1999).
- [78] S. Mukherji and S. M. Bhattacharjee, Phys. Rev. E **63**, 051103 (2001).
- [79] E. B. Kolomeisky and J. P. Straley, Phys. Rev. B **46**, 12664 (1992).
- [80] S. Piatecki and W. Krauth, Nature Communications **5**, 3503 (2014)
- [81] P. Sadhukhan and S. M. Bhattacharjee, J. Phys. A: Math. Theor. **43** 245001 (2010); Europhys. Lett. **98**, 10008 (2012).
- [82] Although it seems that DNA is a one-dimensional object, actually it is not. The dimensionality  $d$  of the embedding space in which DNA, and its monomers, live is important in determining physical properties of DNA. This dimension specific to the model of concern can also be changed, e.g., for a DNA on a surface ( $d = 2$ ) or a DNA in a solution ( $d = 3$ ). So the theorem that there cannot be any phase transition in one dimension is not applicable.
- [83] For a debate on what should be the form of a coarse-grained Hamiltonian for DNA, see, e.g., M. D. Frank-Kamenetskii and S Prakash, Physics of life reviews **11**, 153 (2014).
- [84] J. J. Rajasekaran and S. M. Bhattacharjee, J. Phys. **A24**, L371 (1991) and Phys. Rev. A **44**, 6202 (1991).
- [85] H. M. Wu and D. M. Crothers, Nature (London) **308**, 509 (1984).
- [86] E. N. Trifonov and J. L. Sussman, Proc. Natl. Acad. Sci. **77**, 3816 (1980).

- [87] R. Schleif, *Annu. Rev. Biochem.* **61**, 119 (1992).
- [88] J. A. Rise and D. M. Crothers, *Biochemistry* **28**, 4512 (1989).
- [89] K. Zahn and F.R. Blattner, *Science* **236**, 416(1987).
- [90] J. Kim, C. Zwieb, C. Wu and S. Adhya, *Gene* **85**, 15(1989).
- [91] K. Giese, J. Cox and R. Grosschedl, *Cell* **69**, 185 (1992).
- [92] S. C. Schultz, G. C. Shields and T. A. Steitz, *Science* **253**, 1001 (1991).
- [93] A. K. Nagaich, E. Appella and E. R. Huttington, *J. Biol. Chem.* **272**, 14842 (1997).
- [94] M. E. Ortega and C. E. Catalano, *Biochemistry* **45**, 5180 (2006).
- [95] S. Oehler, M. Amouyal, P. Kolkhof, B. Wilcken-Bergmann, and B. Muller-Hill, *EMBO J.* **13**, 3348 (1994).
- [96] T. J. Richmond and C. A. Davey, *Nature (London)* **423**, 145 (2003).
- [97] T. E. Cloutier and J. Widom, *PNAS* **102**, 3645 (2005).
- [98] P. A. Wiggins, T.V. D. Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker, and P. C. Nelson, *Nat. Nanotechnol.* **1**, 137 (2006).
- [99] P. Ranjith, P.B. Sunil Kumar and G. Menon, *Phys. Rev. Lett.* **94**, 138102 (2005)
- [100] Q. Du, C. Smith, N. Shiffeldrim, M. Vologodskaja and A. Vologodskii, *PNAS* **102**, 5397 (2005).
- [101] A. K. Mazur, *Phys. Rev. Lett.* **98**, 218102 (2007).
- [102] A. Noy, R. Golestanian, *Phys. Rev. Lett.* **109**, 228101 (2012).
- [103] R. P. Linna, K. Kaski, *Phys. Rev. Lett.* **100**, 168104 (2008).
- [104] H. Shroff, D. Sivak, J. J. Siegel, A. L. McEvoy, M. Siu, A. Spakowitz, P. L. Geissler and J. Liphardt, *Biophys. J.* **94**, 2179 (2008).

- [105] R. A. Forties, R. Bundschuh, and M. G. Poirier, *Nucleic Acids Res.* **37**, 4580 (2009).
- [106] T.T. Le and H. D. Kim, *Nucleic Acids Res.* **42**, 10786 (2014).
- [107] A. Vologodskii and M. D. Frank-kramenetskii, *Nucleic Acids Res.* **41**, 6785 (2013).
- [108] J. Ramstein and R. Lavery, *PNAS* **85**, 7231 (1988).
- [109] M. D. Barkley and B. H. Zimm, *J. Chem. Phys.* **70**, 2991 (1979).
- [110] M. Perard, *Nonlinearity* **17**, **R1** (2004).
- [111] N. Theodorakopoulos and M. Peyrard, *Phys. Rev. Lett.* **108**, 078104 (2012).
- [112] P. A. Wiggins, R. Phillips, and P. C. Nelson, *Phys. Rev. E* **71**, 021909 (2005)
- [113] J. Palmeri, M. Manghi and N. Destainville, *Phys. Rev. E* **77**, 011913 (2008).
- [114] D. Marenduzzo, A. Trovato, and A. Maritan, *Phys. Rev. E* **64**, 031901 (2001).
- [115] S. Mukherji and S. M. Bhattachaejee, *J. Phys. A : Math. Theor.* **26**, L1139 (1993).
- [116] S. Mukherji and S. M. Bhattachaejee, *Phys. Rev. E* **48**, 3483 (1993).
- [117] D. J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena*, (McGraw-Hill Inc., New York, 1978).
- [118] P. G. deGennes, *Scaling Concepts in Polymer Physics*, (Cornell University Press, Ithaca, 1979).
- [119] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron, *Science* **271**, 792 (1996).
- [120] D. Marenduzzo, E. Orlandini, F. Seno and A. Trovato, *Phys. Rev. E* **81**, 051926 (2010).
- [121] A. Ahsan, J. Rudinick and R. Bruinsma, *Biophys. J.* **74**, 132 (1998)

- [122] For the inextensibility of the chains there can not be a S-DNA phase in our model as described in [119–121].
- [123] R. Kapri, S. M. Bhattacharjee and F. Seno, *Phys. Rev. Lett.* **93**, 248102 (2004).
- [124] D. Marenduzzo, S. M. Bhattacharjee, A. Maritan and E. Orlandini, *Phy. Rev. Lett.* **88**, 028102 (2002).
- [125] J. Kasianowicz, E. Brandin, D. Branton, and D. Deamer, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13770 (1996).
- [126] A. Son, A.-Y. Kwon, A. Johner, S.-C. Hong and N. Lee, *Europhys. Lett.* **105**, 48002 (2014).
- [127] G. A. King, P. Grossa, U. Bockelmann, M. Modesti, G. J. L. Wuitea, and E. J. G. Peterman, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3859 (2013).
- [128] J.-H. Jeon and W. Sung, *Biophys J.* **95**, 3600 (2008).
- [129] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10579 (1998).
- [130] J. H. Jeon, J. Adamcik, G. Dietler and R. Metzler, *Phys. Rev. Let.* **105**, 208101 (2010).
- [131] G. Altan-Bonnet, A. Libchaber and O. Krichevsky, *Phys. Rev. Lett.* **90**, 138101 (2003).
- [132] T. Ambjörnsson, S. K. Banik, O. Krichevsky, R. Metzler, *Phys. Rev. Let.* **97**, 128105 (2006).
- [133] B. S. Alexandrov, Y. Fukuyo, M. Lange, N. Hirikoshi, V. Gelev, K.Ø. Rasmussen, A.R. Bishop and A. Usheva, *Nucl. Acids Res.* **40**, 10116 (2012).
- [134] F. Sicard, N. Destainville and M. Manghi, *J. Chem. Phys.* **142**, 034903 (2015).
- [135] J. H. Jeon, W. Sung, F. H. Ree, *J. Chem. Phys.* **124**, 164905 (2006).

[136] The expression for  $\Delta t$  can be written in a general form as

$$\Delta t = [Z_b^{-1}(s' \rightarrow 0, \mathbf{k} \rightarrow 0) - 4\pi g_2^2 \Lambda] \Lambda^{-2}.$$

[137] M. E. Fisher, J. Stat. Phys **34**, 667 (1984).

[138] T. Pal, P. Sadhukhan and S. M. Bhattacharjee, Phys. Rev. Lett. **110**, 028105 (2013).

[139] The three-chain coupling  $g_3$  here is the negative of the same parameter in Ref. [138]. This is done for easy comparison with polymer formulations of Refs. [77, 78, 84].

[140] For three different chains, if we label the interface weight as  $g_2^{ij}$  for pair  $i, j$ , then the strand exchange would be like  $g_2^{12}g_2^{23} + g_2^{13}g_2^{23}$ . As we are considering all the interface weights as  $g_2$ , the factor of 2 comes in.

[141] This condition is similar to the “on-shell” condition used in field theory. In field theoretical calculations on-shell and off-shell conditions referred to the center-of-mass frame or non-center-of-mass frame respectively. Our notion of a relaxed DNA segment is similar to the on-shell condition where only the Gaussian part ( $\sim q^2$ ) is important. For off-shell considerations one needs to consider non-Gaussian terms (higher order terms in  $q$ ) which we ignore in this paper.

[142] Other known examples of this type are nonrelativistic solid state physics problems providing examples of relativistic quantum mechanics, liquid crystals as testing ground of cosmological theories; see Ref. [138].

[143] T. Oohashi, K. Inoue and Y. Nakamura, Polymer J. **46**, 699 (2014).

[144] The rise in the scattered intensity at very small wave vectors in small angle neutron scattering has been used to identify effective long range attraction between protein molecules in solutions. See, e.g., Y. Liu, E. Fratini, P. Baglioni, W. R. Chen, and S. H. Chen, Phys. Rev. Lett. **95**, 118102 (2005).

[145] U. Mohideen and A. Roy, Phys. Rev. Lett. **81**, 4549 (1998).

- [146] R. Verma, J. C. Crocker, T. C. Lubensky, and A. G. Yodh, *Phys. Rev. Lett.* **81**, 4004 (1998).
- [147] A. Grosberg, *Physics of Life Reviews* **11**, 178 (2014).
- [148] W. Reisner, N.B. Larsen, A. Silahtaroglu, A. Kristensen, N. Tommerup, J.O. Tegenfeldt and H. Flyvbjerg, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13294 (2010).