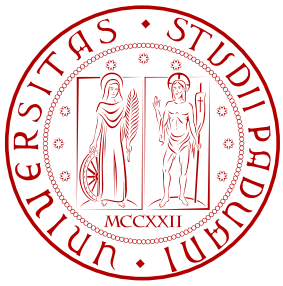


UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica ed Astronomia "Galileo Galilei"

Corso di Laurea Magistrale in Fisica

Efimov bound states in polymer physics

Relatore:
Dott. Antonio Trovato

Candidata:
Federica Mura

Correlatore:
Prof. Flavio Seno

Anno Accademico 2013/2014

Abstract

The aim of this thesis is to investigate, using simplified lattice models, whether the Efimov effect, well known in quantum mechanics, can arise in polymer physics too, particularly in a triple stranded DNA system. Such effect consists in the formation of stable trimer bound states when dimer bound states are not stable.

We base strand interaction rules on a Poland-Scheraga model for a directed DNA-like polymer, double-stranded first and triple-stranded later. Within the Poland-Scheraga scheme, only interactions between base pairs with the same monomer indexes along the strands are permitted. Thus, the identification of the monomer index with imaginary time allows the formal mapping to the quantum problem in which particles interact at the same time. According to the transfer matrix method, we are able to extract information about the free-energy of the system from matrix eigenvalues. Studying them at the critical point of the unzipped-zipped phase transition we look for the similarities between the bound states in the polymer model and those predicted by Efimov theory for a three particle quantum system. In particular, in Efimov theory, an infinite series of trimer energy levels, with a constant ratio between consecutive levels, is found at the critical point of the 2-body problem.

In chapter 1 we briefly describe the main characteristics of double stranded DNA physical and chemical structure, triple stranded DNA and the denaturation process which will be the main arguments of this work.

In chapter 2 we give an elementary explanation of the Efimov effect in quantum physics concentrating on the conditions for which the effect arises. We also explain why to expect that an analogue of the Efimov effect could take place in a polymer system.

In the third chapter we give an analytic procedure to solve exactly the two strands DNA system. After that, we define the Poland-Scheraga model on a 1+1 directed lattice and we introduce the transfer matrix method for the two chain system, starting from a recursive equation for the partition function computed on the lattice. We see how the matrix eigenvalues are related to the free energy of the system and then to the phase transition in the denaturation process. The results achieved in this chapter also act as a check of such method on a problem that is analytically solvable.

In chapter 4 we face the task to build up the transfer matrix for the more complex three stranded system, then in chapter 5 we table the achieved results in comparison with those of the double stranded system. In particular we observe a shift between the transition point of the three and two chain system, accompanied by a free energy variation, in agreement with the analogy suggested by the quantum case. On the other hand, we do not find any evidence for the existence of the analogy with the infinite series of Efimov trimers.

Contents

Acknowledgements

I would like to sincerely thanks all those who helped me in this work, with opinions, corrections, critics and encouragements, that were fundamental for me from the beginning to the end.
Firstly my supervisor Dott. Antonio Trovato of having involved me in this project and led constantly, giving most of his time and meeting me every time I need. I am also grateful to my co-supervisor Prof. Flavio Seno, even if there were few occasions to work with him directly, I know that he has always been interested in shaping this thesis and willing to give us his co-operation and opinion. My thanks also go to Giovanni who helped me, instead of being in England, and to Hicham who adjusted the terrible English that I used to write these acknowledgements. Last but not least, I wish to give thanks to Stefano, not only for the unique practical and moral support he has given me in this thesis period, but also for everything that he has taught me during this academic path, with his sympathy also during his most challenging period, making always available his aptitude for giving me answers to my too many and infinite questions.

Chapter 1

Introduction

Deoxyribonucleic acid, more commonly known as DNA, is a complex molecule that contains all the information necessary to build and maintain living organisms. However, DNA does more than specifying the structure and function of living systems things, it also serves as the primary unit of heredity in organisms of all types. In other words, whenever organisms reproduce, a portion of their DNA is passed along to their offspring. This transmission of all or part of an organism's DNA helps ensure a certain level of continuity from one generation to the next, while still allowing for slight changes that contribute to the diversity of life.

The discovery of DNA double-helical structure by Watson, Crick, Wilkins and Franklin in 1953 [11] was an important milestone of modern biology. They observed that DNA is formed by two complementary strands where, through hydrogen bonds, an adenine pairs with thymine, and guanine with cytosine forming A-T and G-C base pairs. The succession of base pairs defines the genetic information and gives the information to the cell to accomplish its vital functions.

Four years later, Felsenfeld et al [9] found that in DNA there are acceptor and donor groups that can form hydrogen bond interactions with a third strand. This was the starting point for a number of studies that have had great impact in medical progress, for example with the development of the Antigen Therapy [5], based on gene silencing approach. This technique delivers short, single stranded pieces of DNA, called oligonucleotides, that bind specifically between a gene's two DNA strands. This binding makes a triple-helix structure that blocks the DNA from being transcribed into mRNA.

Of course, a good model is a fine starting point to understand the implications and the potential of the triple helix structure. Since DNA is a polymer formed by a great number of monomers it is reasonable to describe it with a mechanical statistical approach, and this is what we are going to do in this

work with particular interest in the three chains structure.
Firstly, we need a clear understanding of the biological structure of double and triple helical, in order to decide how to build our statistical model.

1.1 DNA structure

DNA in the form proposed for the first time by Watson and Crick is a long polymer, made from repeated units called nucleotides. The basic structure comprises two helical chains of nucleotides each coiled round the same axis, and each with a pitch of 34 ångströms (Å, 1 Å= 0.1 nm) and a radius of 10 Å. When measured in a particular solution, the DNA chain measured 22 to 26 Å in width, and one nucleotide unit measured 3.3 Å long [18]. Although each individual repeating unit is very small, DNA polymers can be very large molecules containing millions of nucleotides. For instance, the largest human chromosome, chromosome number 1, consists of approximately 220 million base pairs and would be 85 mm long, if completely stretched.

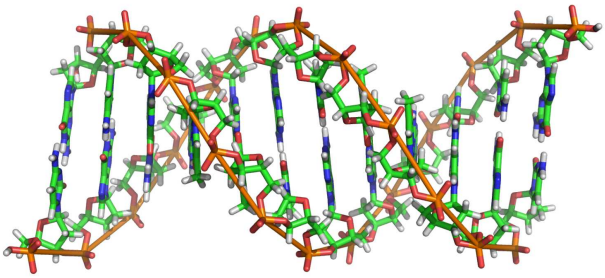


Figure 1.1: A section of DNA.

Each nucleotide consists of a 5-carbon sugar (deoxyribose), a nitrogen containing base attached to the sugar, and a phosphate group. There are four different types of nucleotides found in DNA, differing only in the nitrogenous base. The four nucleotides are given one letter abbreviations as shorthand for the four bases: adenine (A), guanine (G), cytosine (C), thymine (T). Adenine and guanine are purines (5 carbon, 4 nitrogen) that are the largest of the two types of bases found in DNA, while thymine and cytosine are pyrimidines (4 carbon, 2 nitrogen).
The deoxyribose sugar of the DNA backbone has 5 carbons and 3 oxygens. The carbon atoms are numbered 1', 2', 3', 4', and 5' to distinguish from the numbering of the atoms of the purine and pyrmidine rings. Along the single

strand each phosphate group binds the 3'-carbon of a deoxyribose molecule with the 5'-carbon of the sequent sugar, forming phosphodiester bonds.
In a double stranded DNA (dsDNA), each type of nucleobase on one strand pairs with just one type of nucleobase on the other strand. This is called complementary base pairing. Here, purines form hydrogen bonds with pyrimidines, with adenine bonding only to thymine with two hydrogen bonds, and cytosine bonding only to guanine with three hydrogen bonds.

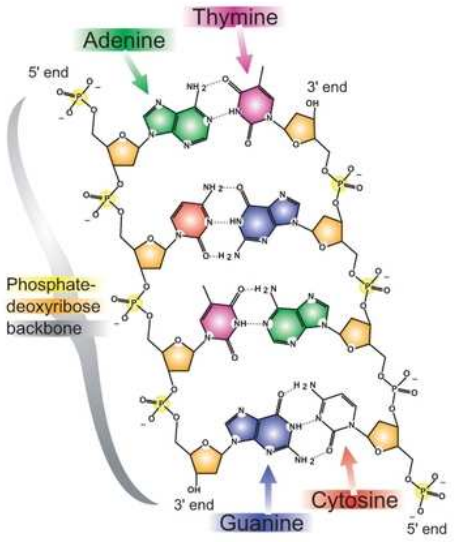


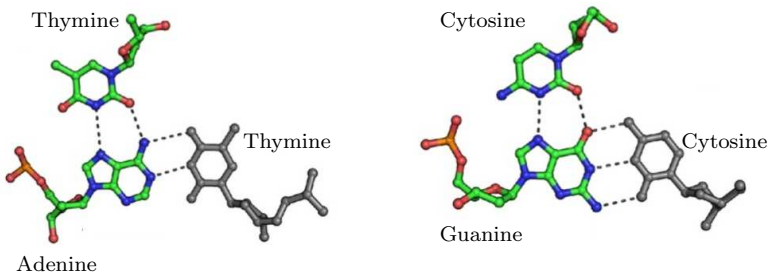
Figure 1.2: DNA chemical structure.

Moreover the double helical structure is stabilised not only by hydrogen bonds, but also by the hydrophobic effect and especially by van der Waals interactions between parallel bases, also known as base stacking interactions, characterized by cooperativity.

1.2 Triple stranded DNA

Triple-stranded DNA (tsDNA) is a structure of DNA in which three oligonucleotides wind around each other and form a triple helix. Triplex formation is governed by sequence-specific binding rules that are conceptually similar

to the familiar Watson-Crick base-pairing rules. The third nucleotide strand binds to dsDNA through Hoogsteen or reversed Hoogsteen hydrogen bonds [10].
In the Hoogsteen pairing the adenine base is flipped as compared to Watson-Crick pairing and, since one nucleotide side is connected to one helix and the other is connected to the other helix, that's going to change the shape of that portion of the DNA. In Fig.(1.3) are shown the two possible Hoogsteen pairs for the TAT couple or CGC.



$T * A - T$ $C * G - C$

Figure 1.3: TAT and CGC pairing in tsDNA. The symbol $*$ indicate the Hoogsteen base pairs and the $-$ the Watson-Crick base pairs

Moreover the Hoogsteen pairing is stable only in determinate contexts, as pH 5 or lower, and the G-C Watson-Crick pairing consists of one more hydrogen bond than Hoogsteen pairing. Indeed normally, the Watson-Crick structure is favoured over the Hoogsteen scheme.
Besides the TAT and CGC pairing in the tsDNA also the GGC and AAT are permitted. In these last cases the Hoogsteen pairing G*G and A*A is said to be *reverse* because the purines sugars are located in the opposite side compared to the first case.
The central strand in tsDNA have to be formed by purines only, because the pyrimidines cannot have multiple hydrogen bonds on both sides simultaneously.
Two particular types of tsDNA have been identified:

- the *intermolecular* tsDNA, discovered in 1957, formed between triplex forming oligonucleotides (TFO) consisting in a single DNA strand that binds to a target sequences on duplex DNA.

- the *intramolecular* tsDNA (Fig.(1.4)), discovered in 1987 [23], that does

not involve an additional strand, but formed thanks to the double helix opening, so that one of the two free strands bind to the dsDNA itself.

Figure 1.4: Scheme of intramolecular tsDNA .

Artificially synthesised TFOs are promising gene-drugs, which can be used in an anti-gene strategy, that attempts to modulate gene activity in vivo. Transcriptional processes can be stopped by different strategies based on the binding of the TFO to a target site and the subsequent creation of a physical block to a normal cellular process. These applications can be relevant in creating cancer therapies that inhibit gene expression at the DNA level. Since aberrant gene expression is a hallmark of cancer, modulating these endogenous genes' expression levels could potentially act as a therapy for multiple cancer types. However, despite much in vitro success, there has been limited achievement in cellular applications potentially due to target accessibility.
Limitations also include concerns regarding binding affinity and specificity, in vivo stability, and uptake into cells. Researchers are attempting to overcome these limitations by improving TFO characteristics through chemical modifications, such as modifying the TFO backbone to reduce electrostatic repulsions between the TFO and the DNA duplex [7].
The intramolecular tsDNA could be also related to degenerative deseas, since it is prone to the structural alterations of DNA. Friedreich's ataxia (FRDA),

the autosomal recessive degenerative disorder of nervous and muscles tissue caused by the transcriptional silencing of the FXN gene, is the only disease known so far to be associated with DNA triplex [25].

1.3 DNA denaturation

The hydrogen bonds between bases of different strands are not covalent so that they can be broken and reformed relatively easily. The two strands of DNA in a double helix can therefore be pulled apart like a zipper, either by a mechanical force or temperature increase. When DNA is destabilised by temperature, the process is called melting, and occurs at high temperature, low salt and high pH. However low pH also melts DNA, but since DNA is unstable in this context, low pH is rarely used.
Since the base pair A-T forms two hydrogen bonds and G-C three, the G-C base pair is stronger so that the dsDNA stability depends on the GC-content. Other indicators for stability are the sequence, because the stacking is sequence specific, and the length, because long molecules are more stable. A common way to measure the stability is the *melting temperature* T_{mc}, which is the temperature at which 50% of the nucleotides are converted from the close (*folded*) to the open (*unfolded*) state.
The primary experimental tool for studying the denaturation process is the measurement of UV light absorption at a wavelength of about 270 nm. Light at this wavelength is preferentially absorbed by the single strands and it thus provides a measure for the fraction of the double-stranded pairs, $\theta(T)$, at any given temperature T. This is known as the optical melting curve of the DNA. We expects θ to decrease with temperature and to vanish at very high temperature[21]. This kind of analysis required μg amounts of DNA, and often took hours to complete. Contemporary DNA melting analysis uses fluorescence because it is a more sensitive method. It needs only ng amounts of DNA, shortening the melting time to a few minutes. (Fig.(1.5))
The denaturation process is energetically disadvantaged ($E_{ss} - E_{ds} = \Delta E > 0$) because of the hydrogen bonds, while is entropically favoured ($S_{ss} - S_{ds=} \Delta S > 0$) because the single strands are more flexible over the double strands so that the spatial configurations are more numerous for the denatured DNA. Indeed the transition occurs when the system reaches a particular temperature above which the free energy variation $F_{ss} - F_{ds} = \Delta F = \Delta E - T\Delta S$ became negative.

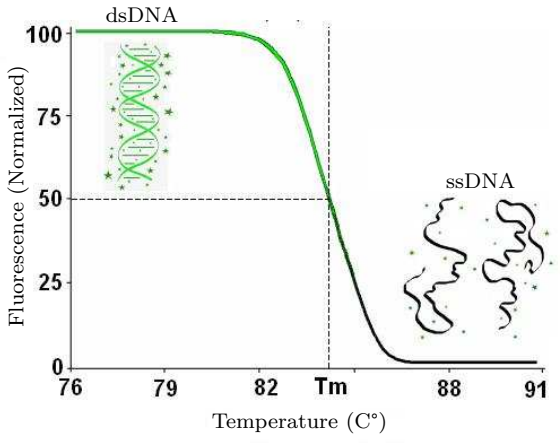


Figure 1.5: Fluorescence melting curve

1.4 DNA and random walks

There are several methods to convert the DNA sequences into digital sequences, each one of these characterized by different accuracy levels.
A detailed study of sequence dependent effects in DNA can be faced through molecular dynamics methods which afford detailed structural and dynamical insights at atomic level. Thanks to molecular dynamics, noteworthy advances have been made in the study of DNA-ligand interactions, but also in simulation of large conformational transitions, including folding and unfolding of short DNA duplexes.
Obviously the great computational cost of such methods does not make it possible to simulate entire systems on a realistic timescale, even if the spectacular improvements in software and hardware of the last years has enabled the tackling of systems of increasing complexity and simulation of reasonably large system dynamics over a timescale of microseconds have now been achieved.
Moreover large system as DNA can be addressed in statistical terms, considering the thermodynamics limit. For example the thermal denaturation can be seen as a phase transition in a polymeric system of infinite dimension.
In order to describe such process, we will consider a simplified model in which DNA is intended as a homopolymer neglecting all the details about the presence of different base pairs in the sequence and the interaction strength between them. In these models DNA is described by means of two interacting random walks. The steps of such random walks cannot be interpreted as related to the real DNA monomer, in fact the helicoidal structure of DNA is

stiff because subsequent monomers are constrained to maintain a fixed angle between them for chemical reasons. We can interpret the random walk model as an equivalent polymer in which each monomer is formed by a set of different nucleotides. This model is known as *Equivalent freely jointed chain* and consist in a coarse graining of the real model of DNA strands.
Another fundamental aspect to consider in DNA modellisation is the *excluded volume* effect. Experimental studies on DNA melting are always carried on in solution and a property of a good solvent is that interaction of the polymer with solvent molecules is energetically favoured over interaction with other monomers. Indeed each monomer is surrounded by solvent molecules, forming an inaccessible region for the other chain elements. The most common way to describe such effect is the *self-avoiding walk* (SAW), that is a model in which self-intersection in the chain paths are banned. The Poland-Scheraga model that we are going to introduce in the Ch.(3) involve two different SAW, representing dsDNA, of which only elements with the same distance from the origin, measured along the strands, can interact. This option is useful to approximate the DNA base pairing and will enable us to develop an analogy with the quantum case of two interacting particles. In fact is possible to interpret the step index in the random walks as the time parameter in the quantum problem in which particles can interact only at the same time.

Chapter 2

Efimov states

The Efimov effect is an effect in quantum mechanics first predicted by V. Efimov in 1970 [28] for identical bosons that occupy a spatially symmetric s-state and interact with a short-range pair-wise potential.
In these circumstances, their spectrum obeys a geometrical scaling law, such that the ratio of the successive energy eigenvalues of the system is a constant and accumulation of states near zero energy takes place. When the dimer binding is zero and the two-body state is exactly at the dissociation threshold, the number of three-body bound states is infinite. Indeed, three body bound states are stable when two-body ones not. An issue to emphasise is the fact that Efimov effect is independent of the details of the interaction, for example, whether it is between atoms with a pair-wise van der Waals interaction, or between nuclei with a nuclear force.
In 2005, for the first time the research group of Rudolf Grimm and Hanns-Christoph Nägerl from the Institute for Experimental Physics (University of Innsbruck, Austria) experimentally confirmed such a state in an ultracold gas of caesium atoms [16]. At such low temperatures (10 nK) thermal motion does not mask quantum effects.
In this chapter an elementary exposition of the Efimov effect, will be given using basic concepts of quantum mechanics.

2.1 Two body problem

2.1.1 Two-body scattering at low energy

We need to recall the most important results about the quantum mechanical problem of a beam of particles incident at low energy upon a target. In principle, if we assume that all the in-going particles are represented by wavepackets of the same shape and size, our challenge is to solve the full time-dependent

Schrödinger equation for such a wavepacket

$$i\hbar\partial_t\Psi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\Psi(\mathbf{r},t) \tag{2.1}$$

and find the probability amplitudes for out-going waves in different directions at some later time after scattering has taken place. However, if the incident beam of particles is switched on for times very long as compared with the time a particle would take to cross the interaction region, steady-state conditions apply. Moreover, if we assume that the incident wavepacket has a well-defined energy (and hence momentum), so it is many wavelengths long, we may consider it a plane wave. Setting $\Psi(\mathbf{r},t) = \psi(\mathbf{r})e^{-iEt/\hbar}$ we may therefore look for solutions $\psi(\mathbf{r})$ of the time-independent Schrödinger equation

$$E\psi(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) \tag{2.2}$$

subject to the boundary condition that the incoming component of the wavefunction is a plane wave, $e^{i\mathbf{k}\cdot\mathbf{x}}$. Here $E = \mathbf{p^2}/2m = \hbar^2\mathbf{k^2}/2m$ denotes the energy of the incoming particles.
In the three-dimensional system the wavefunction well outside the localized target region will involve a superposition of the incident plane wave and the scattered spherical wave

$$\psi(\mathbf{r}) \simeq e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta)\frac{e^{ikr}}{r} \tag{2.3}$$

where we assume that the potential perturbation $V(r)$ depends only on the radial coordinate and the function $f(\theta)$ records the relative amplitude and phase of the scattered components along the direction θ relative to the incident beam.
Using the partial wave method the wavefunction can be expanded in a series of Legendre polynomials and $f(\theta)$ results

$$f(\theta) = \sum_{l=0}^{\infty}(2l+1)f_l(k)P_l(\cos\theta) \tag{2.4}$$

where $P_l(\cos\theta)$ denote the Legendre Polynomials and $f_l(k)$ is the *partial wave scattering amplitude*. At low energies is possible to show that the total scattering amplitude is dominated by the s-wave ($l = 0$) channel and results

$$f(\theta) \simeq f_0(k) = (k\cot\delta_0(k) - ik)^{-1} \tag{2.5}$$

where $\delta_0(k)$ is the s-wave phase shift caused by potential. The scattering between the two particles is well described [12] at low energies by two shape-independent parameters, the scattering length a and the effective range r_0, where

$$k \cot \delta_0(k) = -\frac{1}{a} + \frac{1}{2} r_0 k^2 + \dots \tag{2.6}$$

The scattering length a is linked to potential strength and for a square well potential, considering $\hbar = 2m = 1$, $a = \left(1 - \frac{tan\gamma}{\gamma}\right)$ with $\gamma = U_0^{1/2} R$, U_0 the depth and R the width of well potential. At low energies, $k \to 0$, the scattering cross-section, $\sigma = 4\pi a^2$ is fixed by the scattering length alone. If $\gamma \ll 1$, a is negative. As γ is increased, when $\gamma = \pi/2$, both a and σ diverge - there is said to be a zero energy resonance. This condition corresponds to a potential well that is just able to support an s-wave bound state. We'll see that the Efimov effect arise in this context.

2.1.2 Separable two-body potential

Consider the scattering of a heavy particle of mass M interacting with a light one of mass m, with the mass ratio $\rho = \frac{M}{m}$. The relative energy is given by $E_2 = k^2/\nu'$ where we have set again $\hbar = 2m = 1$ and $\nu' = \rho/(\rho + 1)$. For simplicity we assume $\rho \gg 1$ and $\nu' = 1$. The eigenvalues equation is

$$\left(k^2 - H_0\right) |\psi_k\rangle = V |\psi_k\rangle \tag{2.7}$$

where H_0 is the kinetic energy operator. The Eq.(2.7) takes the form

$$\left(k^2 + \nabla^2\right) \psi_k(\mathbf{r}) = \int \langle \mathbf{r}| V |\mathbf{r}'\rangle \, \psi_k(\mathbf{r}') d^3 r' \tag{2.8}$$

in the **r**-space with **r** the relative coordinate between the two particles, and

$$\left(k^2 - p^2\right) \psi_k(\mathbf{p}) = \int \langle \mathbf{p}| V |\mathbf{p}'\rangle \, \psi_k(\mathbf{p}') d^3 p' \tag{2.9}$$

in the **p**-space, where **p** is the relative momentum between the two particles. A separable potential in Hilbert space may be written as $V = -\lambda |g\rangle \langle g|$, where the negative sign is taken for attraction, and $\lambda > 0$ determines the strength of the potential. In the coordinate representation the separable potential in **r**-space is given by $\langle \mathbf{r}| V |\mathbf{r}\rangle = -\lambda g(r) g(r')$ where g(r) is taken to be real. The Schrödinger equation in **p**-space for a bound state $k^2 = -k_0^2$ is given by

$$(k_0^2 + p^2)\psi_{k_0}(\mathbf{p}) = \lambda g(p) \int g(p')\psi_{k_0}(p') d^3 p' \tag{2.10}$$

where

$$g(p) = \int \langle \mathbf{p}|\mathbf{r}\rangle g(r) d^3r = \frac{1}{(2\pi)^{3/2}} \int exp(-i\mathbf{p}\cdot\mathbf{r}) g(r) d^3r \tag{2.11}$$

Rewriting Eq.(2.10) as

$$\psi_{k_0}(p) = \lambda C_{k_0} \frac{g(p)}{(k_0^2 + p^2)} \tag{2.12}$$

with $C_{k_0} = \int g(p')\psi_{k_0}(p') d^3p'$ and multiplying both sides by $g(p)$ and integrating over d^3p we obtain the equation that determines the binding energy k_0^2 for a given potential

$$\lambda \int \frac{g(p)^2}{(k_0^2 + p^2)} d^3p = 1 \tag{2.13}$$

For explicit calculations we choose the popular Yamaguchi form $g(p) = (p^2 + \beta^2)^{-1}$ giving

$$g(r) = \frac{1}{(2\pi)^{3/2}} \int exp(i\mathbf{p}\cdot\mathbf{r}) g(p) d^3p = \sqrt{\frac{\pi}{2}} \frac{exp(-\beta r)}{r} \tag{2.14}$$

and

$$\psi_{k_0}(r) = \lambda C_{k_0} \sqrt{\frac{\pi}{2}} \left(\frac{exp(-k_0 r)}{r} - \frac{exp(-\beta r)}{r} \right) \tag{2.15}$$

2.2 Efimov spectrum and the inverse square potential

Let us now focus upon the problem of a single particle of mass m in an inverse square potential $V(r) = (\hbar^2/2m)\lambda'/r^2$ where λ' is a dimensionless coupling constant.
It's possible to show that for $\lambda' > -1/4$ there is no bound states, while infinite and continuous number of bound states arise for $\lambda' < -1/4$. In order to solve this problem we need to take a short distance cut-off in some point $r = r_c$ and impose that eigenfunctions vanish on it [27][6].
We are interested in the situation where the potential is inverse square only for $r > r_c$ so we can write the radial ($u(r) = r\psi(r)$) Schrödinger equation for this case:

$$\left[-\frac{d^2}{dr^2} - \frac{s_0^2 + 1/4}{r^2} \right] u(r) = \frac{2m}{\hbar^2} E u(r) \tag{2.16}$$

where $s_0^2 \geq 0$ is just a way of parametrizing the strength of an inverse square potential whose coupling constant is smaller than $-1/4$. For bound states, we

set $(2m/\hbar^2)E = -k^2$, and require that wave functions vanish at infinity. This equation give us an energy spectrum for shallow bound states that follows a geometric scaling law given by [13] :

$$\frac{E_{n+1}}{E_n} = exp(-2\pi/s_0) \qquad\qquad n = 0, 1, 2, ...\infty \tag{2.17}$$

Note that as n becomes larger, the states become shallower, with an infinite number of states accumulating at zero energy. This results will be useful, as we'll see in due course, to solve in the three body problem the Schrödinger equation for radial coordinate.

2.3 The three body problem

Our aim is now to obtain an inverse square interaction, like that of the previous section, in the three body problem where the particles interact with a short range pair-wise potential. This objective is best served by taking two identical heavy particles 1 and 2, each of mass M, and particle 3 of mass $m \ll M$. The analysis is simplified if the relative coordinates $\mathbf{R} = (\mathbf{r}_1 - \mathbf{r}_2)$ and $\mathbf{r} = \mathbf{r}_3 - (\mathbf{r}_1 + \mathbf{r}_2)/2$ are introduced, where $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_3$ are the coordinates of particles 1, 2 and 3 measured from an arbitrary origin. Following the common convention, we denote by V_1 the interparticle potential between particles 2 and 3, and likewise for V_2 and V_3 in cyclic order. The Schrödinger equation results

$$H\Psi(\mathbf{r}, \mathbf{R}) = E\Psi(\mathbf{r}, \mathbf{R}) \tag{2.18}$$

with

$$H = -\frac{1}{\mu}\nabla_R^2 - \frac{1}{\nu}\nabla_r^2 + V_1(\mathbf{r} - \mathbf{R}/2) + V_2(\mathbf{r} + \mathbf{R}/2) + V_3(R) \tag{2.19}$$

$$\mu = \rho/2 \qquad \nu = 2\rho/(2\rho + 1) \tag{2.20}$$

where E is the total energy of the system and $\rho = M/m$, recalling that $\hbar = 2m = 1$. For $\rho \gg 1$, $\nu \to 1$, and the motion of the heavy particles is very slow compared to that of the light particle of mass m, so that we can apply the adiabatic Born-Oppenheimer approximation to solve the three-body Schrödinger equation in two stages. First, the wave function is decomposed according to

$$\Psi(\mathbf{r}, \mathbf{R}) = \psi(\mathbf{r}, \mathbf{R})\phi(\mathbf{R}) \tag{2.21}$$

where $\psi(\mathbf{r}, \mathbf{R})$ with eigenenergy $\epsilon(R)$, is first solved for the relative motion of the light-heavy system, keeping $\mathbf{R}$ as parameter. For fixed $\mathbf{R}$ the relative kinetic energy of the heavy particle is zero, and the potential $V_3(R)$ is a constant

shift in energy and can be neglected at this stage. Eq.(2.18) becomes

$$\left[-\nabla_r^2 + V_1(\mathbf{r} - \mathbf{R}/2) + V_2(\mathbf{r} + \mathbf{R}/2)\right] \psi(\mathbf{r}, \mathbf{R}) = \epsilon(R)\psi(\mathbf{r}, \mathbf{R}) \tag{2.22}$$

Next, we can solve within the adiabatic approximation the two-body Schrödinger equation with $\epsilon(R)$ as the adiabatic potential between the two heavy particles

$$[-\nabla_R^2/\mu + V_3(R) + \epsilon(R)]\phi(R) = E\phi(R) \tag{2.23}$$

For V_1 and V_2 in Eq.(2.22), we take the short-range separable potential $-\lambda \left|g\right\rangle \left\langle g\right|$. Thus for a fixed parameter $\mathbf{R}$ we have

$$\langle \mathbf{r}' - \mathbf{R}/2 | \, V_1 \, | \mathbf{r} - \mathbf{R}/2 \rangle = -\lambda \langle \mathbf{r}' - \mathbf{R}/2 | g \rangle \langle g | \mathbf{r} - \mathbf{R}/2 \rangle \tag{2.24}$$

$$\langle \mathbf{r}' + \mathbf{R}/2 | \, V_2 \, | \mathbf{r} + \mathbf{R}/2 \rangle = -\lambda \langle \mathbf{r}' + \mathbf{R}/2 | g \rangle \langle g | \mathbf{r} + \mathbf{R}/2 \rangle \tag{2.25}$$

In the $\mathbf{p}$-space after some algebra and letting $\epsilon(R) = -k^2(R)$ the Eq.(2.22) takes the form

$$\lambda \int \frac{g^2(p_r)}{p_r^2 + k^2} d^3p_r + \lambda \int \frac{g^2(p_r) exp(i\mathbf{p}_r \cdot \mathbf{R})}{p_r^2 + k^2} d^3p_r = 1 \tag{2.26}$$

The coupling constant λ may be eliminated by relating it to the binding k_0^2 of the two body problem from Eq.(2.13). For the Yamaguchi form $g(p) = (p^2 + \beta^2)^{-1}$ the integrals in Eq(2.26) may be performed analytically, giving

$$1 - \left(\frac{\beta + k_0}{\beta + k}\right)^2 = \left(\frac{\beta + k_0}{\beta + k}\right)^2 \left[\frac{2\beta}{(\beta - k)^2} \frac{e^{-kR} - e^{-\beta R}}{R} - \frac{\beta + k}{\beta - k} e^{-\beta R}\right] \tag{2.27}$$

As the distance R between the two heavy atoms is increased, the light atom tend to attach to one of them, and $k^2(R) \to k_0^2$. We are particularly interested in this large R behaviour of $\epsilon(R)$ as the two-body binding k_0^2 goes to zero. It's convenient to define $\zeta \equiv k - k_0$ and substitute it into Eq.(2.27). In the resonant limit, consider $a \to +\infty$ so that $k_0 = 1/a \to 0$ [13], and moreover $k_0 \ll \beta$, $\beta R \gg 1$, $R/a \to 0$, $\zeta \ll \beta$ so that $exp(-\beta R) \to 0$ and $exp(-k_0 R) \to 1$. The Eq.(2.27) in this limit results

$$\frac{e^{-\zeta R}}{\zeta R} = 1 \tag{2.28}$$

which has the solution $\zeta R = A$ where $A = 0.5671\ldots$ Hence $k = k_0 + \zeta = k_0 + A/R$. In the resonant limit we see that $\epsilon(R) = -k^2 = -A^2/R^2$, giving the desired inverse square potential.

In our analysis, we set $\nu = 1$, which only approximately holds for $\rho \gg 1$. However, setting $\nu = 1$ is not necessary, but it's possible to show [1] that $\epsilon(R) = -\mu A^2/(\nu R^2) = -(1/4A^2)(1 + 2\rho)/R^2$. Substituting this expression in

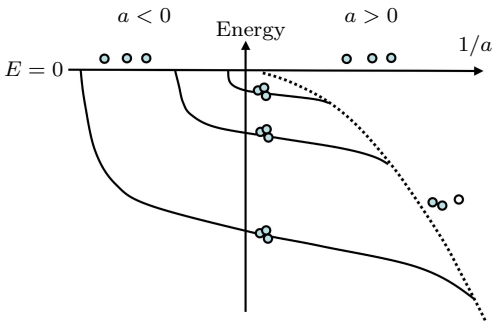


Figure 2.1: Bound and unbound states for three and two particles system. The bound Efimov states are shown schematically by solid lines, with a scaling factor set artificially at 2 rather than 22.7. We see that for $a < 0$, even though there is no two body bound state (dimer), more three-body bound states (trimers) are formed. For $a > 0$ the trimers break up at the atom-dimer continuum, indicated by the dotted curve.

Eq.(2.23) and neglecting the short-range potential V_3 we obtain an equation similar in form to that achieved in the previous section for single particle in an inverse square potential.
Indeed Efimov spectrum arise in three body problem too, and the scaling parameter s_0 depends on the mass ratio ρ. For three identical bosons we get $exp(\pi/s_0) \simeq 22.694\ldots$. The number of shallow bound states is given approximately by [1]

$$N \simeq \frac{s_0}{\pi} \ln \frac{|a|}{r_0} \tag{2.29}$$

and we note that it diverges in the resonant limit $a \to \infty$ with finite r_0 for a short range potential. Fig.(2.1) depicts the Efimov scenario for the three boson states plotted as function of $1/a$.

2.4 Efimov effect in polymer physics

After seeing the underlying causes of Efimov states in quantum mechanics, we're going to investigate whether this effect can arise in polymer physics too. To look into that, consider three Gaussian polymers interacting through a DNA base pairing type short range potential. The Hamiltonian for such system results [3]

$$\beta H = \int_0^N ds \left[\sum_{j=1}^{3} \frac{K_j}{2} \left(\frac{\partial \mathbf{r}_j(s)}{\partial s} \right)^2 + \sum_{k<l} V_{kl}(\mathbf{r}_k(s), \mathbf{r}_l(s)) \right] \tag{2.30}$$

where s is the length variable measured along the contour of the chain and identify the monomer, j is the index of the chain ($j = 1, 2, 3$), $\mathbf{r}_j(s)$ is the monomer position, V_{kl} is the short-range attractive interaction between chains k and l, K_j is a bending rigidity controlling the flexibility of chain j.
We can obtain the partition function of the system by summing over ($\int \mathcal{D}R$) all configurations of the three chains:

$$Z = \int \mathcal{D}R \exp(-\beta H) \tag{2.31}$$

The DNA-quantum correspondence relates the quantum critical threshold to the thermal melting of duplex DNA, a continuous transition in this model of Gaussian chains, with a diverging length scale. Suppose that we have two non-interacting chains 1 and 3, both interacting with another one, chain 2. For simplicity, although not essential, in the spirit of Born-Oppenheimer approximation we may take 1 and 3 as relatively stiffer compared to 2, noting that in Eq.(2.30) the stiffness parameter K_j in the kinetic term, takes the place of mass of the quantum case. The system composed by chain 1 and chain 2 or chain 2 and chain 3 is characterized generally by the presence of many crossings which define a sequence of bubbles (Fig.2.2). These bubbles are described by two scales: $\xi_\perp$ for the spatial extent and $\xi_\parallel$ for the length of the bubbles along the chain. Since $\xi_\perp$ is linked to the correlation length of the system and $\xi_\parallel$ to the relaxation time, these are related by $\xi_\parallel \sim \xi_\perp^z$, where z is a size exponent for polymers. Moreover, we note that the melting transition of double chain system takes place in the limit $\xi_\perp \to \infty$. The diffusive or Gaussian nature of the free chains for our case implies $z = 2$ [20]. If $\xi_\perp > R$, where R is the distance between chain 1 and 3, then in between two contacts with chain 1, chain 2 is expected to meet chain 3, thus mediating an effective interaction between chains 1 and 3.
By expliciting in the Eq(2.31) the sum over the configurations of each chain we achieve:

$$\begin{aligned} Z = \int DR_1 \int \mathcal{D}R_2 \int \mathcal{D}R_3 \exp \Bigg[-\beta \int_0^N ds \frac{K_1}{2} \left(\frac{\partial \mathbf{r}_1}{\partial s} \right)^2 + \\ + \frac{K_2}{2} \left(\frac{\partial \mathbf{r}_2}{\partial s} \right)^2 + \frac{K_3}{2} \left(\frac{\partial \mathbf{r}_3}{\partial s} \right)^2 + V(\mathbf{r}_1, \mathbf{r}_2) + V(\mathbf{r}_2, \mathbf{r}_3) + V(\mathbf{r}_1, \mathbf{r}_3) \Bigg] \end{aligned} \tag{2.32}$$

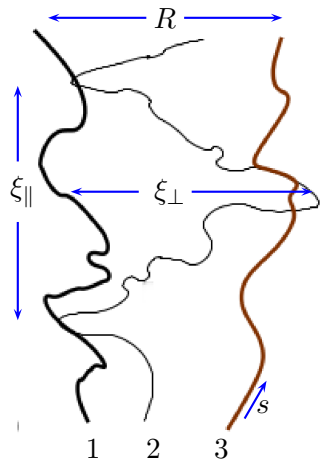


Figure 2.2: Two non interacting Gaussian chains 1 and 3 separated by a distance R. Each of these can pairs with a flexible chain 2, denoted by the thin line. The extent along the bubble contour is $\xi_\parallel$ and the spatial extent is $\xi_\perp$,while s is the length variable measured along the contour of the chain.

Born-Oppenheimer approximation allows to perform the integration over the configurations of flexible chain 2, fixing that of the other two chains. The analogy with the quantum case suggests that the outcome of such operation could be the emergence of three chains bound states at two chains critical point, resulting in a reduction of the free energy of the system. As in the quantum case the effect of the light particle induces an attractive interaction, at long length scales, between the other two particles, so in the polymer system we expect the arising of similar interaction between the stiffer chains.
We can assume the change in free energy as:

$$\Delta F \sim -\frac{N}{\xi_\parallel} f(R/\xi_\perp) \tag{2.33}$$

where the first factor is the numbers of bubbles and $f(x)$ is a scaling function. In the limiting behaviour suggested by the quantum analogy, for $\xi_\perp \to \infty$, Eq.(2.33) should be finite, requiring $f(x) \sim x^{-z}$ as $x \to 0$. Indeed, in this limit and for $z = 2$ we have

$$\epsilon(R) \equiv \frac{\Delta F}{N} \simeq -\frac{A}{R^z} = -\frac{A}{R^2} \tag{2.34}$$

where A is constant. We see the emergence of an inverse square interaction like in the quantum case suggesting the origin of three chain Efimov-like bound

states. Different polymer models based on hierarchical diamond lattice or Euclidean lattice in 1+1 dimensions have been developed by Jaya Maji *et al.*[20] and have given us the numerical evidence that the melting point in DNA unzipped-zipped transition is different in the case of three or two chains. Thus three chains bound states would appear to arise even if there are no two chain bound states yet. In the next chapters we will try to find a direct evidence of Efimov states in the well known Poland-Scheraga polymer model using the transfer matrix technique that allows us to directly access the free energy of a system of infinite chains confined in a finite lattice strip, eventually highlighting the presence of bound states.

Chapter 3

Double chain DNA models

From a statistical mechanics point of view the two chain system is obviously simpler than that with three chains. In the latter case, many difficulties are encountered in the calculation of the partition function, so that we need an alternative technique to obtain information about the system. In order to do that, we will use the well known transfer matrix method.
On the other hand different theoretical models have been developed to study the double chains DNA configurations by a statistical mechanics approach. In this chapter we will discuss one of these models and we will see how to obtain the partition function exactly and other analytical results. This will be helpful because it allows to compare these analytical outcomes with those achieved by the transfer matrix technique applied to the two chain system, testing therefore such method in this context.

3.1 Generating function

A simple but accurate way to describe DNA structure was introduced by Poland and Scheraga [4] and consists of two infinite random walks that can interact when two point of different walks, whit the same distance from the origin, touch each other. This is useful to simulate complementary DNA base pairing, where, to a good approximation, only bases with the same chemical distance from the origin can join together. Configurations of partially melted DNA are represented in this model as alternating sequences of double stranded segments (rods) and single stranded loops (bubble), as depicted in Fig.(3.1). To make analytical computations feasible, we shall ignore the variations in binding energy for different nucleotides, and assign an average energy $\epsilon < 0$ per double-stranded base pair. The weight of a rod segment formed by l nucleotides is thus $exp(-\beta\epsilon l) = w^l$. We assign also a weight σ for each bubble opening to take into account the difficulty to pass from a rod segment to the

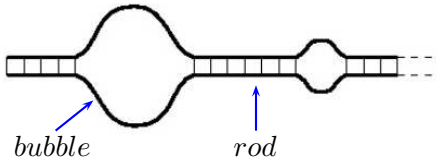


Figure 3.1: Schematic representation of the Poland-Scheraga model.

bubble and the corresponding increase in free energy when two bases break up. The parameter σ reflect the breaking of cooperative stacking interaction between consecutive base pairs. In real DNA $\sigma \simeq 10^{-4} \div 10^{-5}$ and denaturating bubbles are strongly suppressed.
For the sake of simplicity we choose to start and end the DNA chain with rod segments. We consider indeed a chain of N steps, formed by s bubbles and $s+1$ rods with lengths, in order of position along the chain, $i_0, j_1, i_1, \ldots, j_s, i_s$, where j are the lengths of the bubbles and i those of the rods. Since all internal segments must measure one step at least, it is required that

$$i_0 + \sum_{\tau=1}^{s} (i_\tau + j_\tau) = N, \qquad 0 \leq s \leq \left[\frac{N-1}{2} \right] \tag{3.1}$$

The canonical partition function results

$$Z_N(w) = \sum_{0}^{[(N-1)/2]} \sum_{\{i_\tau, j_\tau\}} \left(\prod_{\tau=1}^{s} v_{i_\tau} w^{i_\tau+1} g\sigma u_{j_\tau} \right) v_{i_0} (wz)^{i_0} \tag{3.2}$$

where $\{i_\tau, j_\tau\}$ denotes the sum over all possible lengths of segments such that $i_0 + \sum_\tau (i_\tau + j_\tau) = N$; v_{i_τ} is the number of different configurations of rods with length i_τ and u_{j_τ} is the number of different configurations of bubbles with length j_τ. The g factor is 2 when the crossing of two strands is allowed and 1 otherwise. In other words, we are defining bubbles for two chains that are not allowed to cross each other in a one-dimensional system.

The grand canonical partition function, with fugacity z, is

$$\begin{aligned} G(w,z) &= \sum_{N=0}^{\infty} z^N Z_N(w) \\ &= \sum_{N=0}^{\infty} \sum_{s=0}^{[(N-1)/2]} \sum_{\{i_\tau, j_\tau\}} \left(\prod_{\tau=1}^{s} v_{i_\tau} w^{i_\tau+1} g\sigma u_{j_\tau} z^{i_\tau+j_\tau} \right) v_{i_0} (wz)^{i_0} \\ &= \sum_{s=0}^{\infty} \sum_{N=2s}^{\infty} \sum_{\{i_\tau, j_\tau\}} \left(\prod_{\tau=1}^{s} v_{i_\tau} w (wz)^{i_\tau} g\sigma u_{j_\tau} z^{j_\tau} \right) v_{i_0} (wz)^{i_0} \\ &= \sum_{s=0}^{\infty} \prod_{\tau=1}^{s} \sum_{i_\tau=0}^{\infty} \sum_{j_\tau=0}^{\infty} v_{i_\tau} w (wz)^{i_\tau} g\sigma u_{j_\tau} z^{j_\tau} \sum_{i_0=0}^{\infty} v_{i_0} (wz)^{i_0} \\ &= \sum_{s=0}^{\infty} \left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_i w (wz)^i g\sigma u_j z^j \right)^s \cdot \sum_{i_0=0}^{\infty} v_{i_0} (wz)^{i_0} \end{aligned} \tag{3.3}$$

We define now the generating function $V(x)$ for rods and $U(x)$ for bubbles

$$V(x) = \sum_{i=0}^{\infty} v_i x^i, \qquad\qquad U(x) = \sum_{j=0}^{\infty} u_j x^j \tag{3.4}$$

Indeed we obtain

$$G(w,z) = V(wz) \sum_{s=0}^{\infty} (wg\sigma U(z) V(wz))^s = \frac{V(wz)}{1 - wg\sigma U(z) V(wz)} \tag{3.5}$$

3.1.1 Directed lattice model

To compute explicitly the generating functions in Eq.(3.4) we introduce the Poland-Scheraga model [4] for a directed self-avoiding polymer in a 1+1 square lattice. Particularly, consider two strands moving in oblique-lattice as showed in Fig.(3.2); at each step we enforce them to increase their coordinate along the $(1,1)$ direction so that the possibles moves need to be along $(1,0)$ or $(0,1)$ directions.

In this way self-intersections are banned and only interactions between bases with the same distance from the origin are permitted, as it should happen for two complementary DNA strands.
The last request is important regarding the problem of the quantum mapping in which particles interact at the same time. We can map, in fact, the N-th strand step to the N-th instant of discretized time in the formalism for quantum particles.

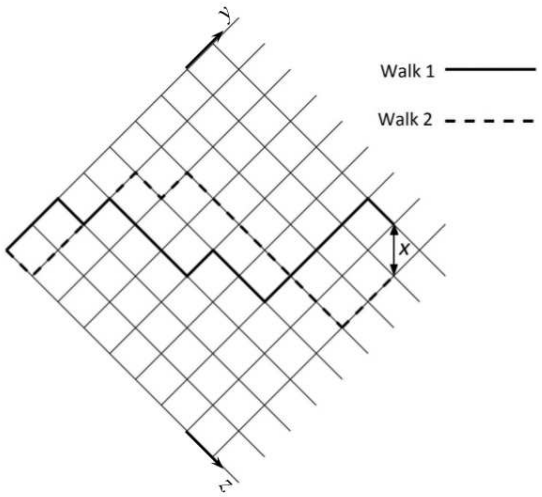


Figure 3.2: Double stranded polymer in oblique lattice.

These rules demand two different possibilities at each step and therefore the number of different paths for a rod with length i is $v_i = 2^i$ and V results:

$$V(x) = \sum_{i=0}^{\infty} 2^i x^i = \frac{1}{1-2x} \tag{3.6}$$

To find an explicit expression for the bubble generating function is more complicated and it results [2]

$$U(z) = \frac{1-2z-\sqrt{1-4z}}{2} \tag{3.7}$$

Finally, the grand canonical partition function results

$$\begin{aligned} G(z,w) =& \frac{\frac{1}{1-2wz}}{1-wg\sigma\frac{1}{1-2wz}\frac{1-2z-\sqrt{1-4z}}{2}} \\ =& \frac{2}{2-4wz-wg\sigma(1-2z-\sqrt{1-4z})} \end{aligned} \tag{3.8}$$

There are two points of non analyticity for G:

- $z_U = \frac{1}{4}$ that is the radius of convergence of $U(z)$
- z_c such that $2-4wz = wg\sigma(1-2z-\sqrt{1-4z})$, namely

$$z_c(w) = \frac{1}{(2-g\sigma)^2}\left[\frac{2}{w} - g\sigma\left(\frac{1}{w}+1-\sqrt{1-\frac{2}{w}+\frac{g\sigma}{w}}\right)\right] \tag{3.9}$$

Since the free energy density $f = F/N$, where N is the steps number and F the total free energy, is linked to the radius of convergence r_c of the grand canonical partition function [14] by

$$f = \lim_{N\to\infty} K_B T \ln r_c(w) \tag{3.10}$$

we obtain that the denaturation transition take place at w_t such that the two singularities are equal

$$w_t = \frac{4}{2 + g\sigma} \tag{3.11}$$

We can mark that in the case of allowed crossing ($g = 2$) and $\sigma = 1/2$ it results $w_t = 4/3$.
Henceforward we'll be interested only in the case with crossing, so we set $g = 2$ and this will the value used in the following.
In this discussion we enforce the rearmost segments to join together, but in general we can find configurations with open strands ending in a "fork configuration". Anyhow, the generating function of fork segments don't insert other singularities in the partition function of the system, so that the value of w at the transition point remains the same [24].

3.2 Transfer matrix method

An alternative form to express the canonical partition function is

$$Z_N(w, \sigma) = \sum_{x=0}^{N} d_N(x, w, \sigma) \tag{3.12}$$

where x is the unsigned distance between the strands at the end of the chain and $d_N(x, w, \sigma)$ is the partition function for all configurations with fixed x.
So let's see now how a recursive equation for d_N can be achieved. The two strands in the square lattice can move in four different ways, so that each configuration with a fixed ending distance x at the N-th step can be obtained in four different ways starting from the $(N-1)$-th step: two starting from distance x, one from $x+1$ and another one from $x-1$ (Fig.3.13).
In general we should apply the relation $d_N(x) = \sum_{x'} d_{N-1}(x')\mu(x, x')p(x, x')$ where $\mu(x, x')$ is the number of ways of passing from distance x' to distance x and $p(x, x')$ is the weight due to bubble opening and base pair interaction in double-stranded segments. An alternative formulation could be given if bubble closure is weighted; the two formulations (weighting either opening or closure) yield different results for finite size, but the same results in the

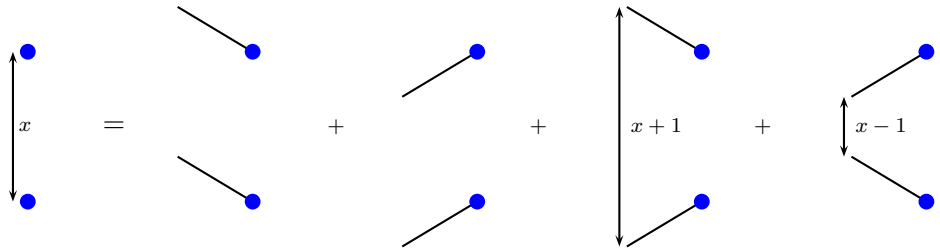


Figure 3.3: Scheme of recurrence relation for the coefficients $d_N(x)$.

thermodynamic limit.
Therefore, the recurrence relation for the partition function with fixed x results

$$d_N(x) = [2d_{N-1}(x) + d_{N-1}(x-1)(2\sigma)^{\delta_{x,1}} + d_{N-1}(x+1)]w^{\delta_{x,0}} \qquad (3.13)$$

Obviously for a system in an infinite lattice the $d_N(x)$ are vectors of infinite dimension $[d_N(0), d_N(1), ..., d_N(\infty)]^T$ because all ending distances x between strands are permitted. But, with appropriate periodic boundary conditions, we can confine paths in a finite strip of transverse size L and infinite length. In this way the $d_N(x)$ are vectors of finite dimension: $[d_N(0), d_N(1), ..., d_N(L/2)]^T$. In fact, if we impose the top and bottom side of the strip to coincide, the paths are enforced to move in a cylinder of base circumference L and the largest distance which can be achieved is $L/2$ [Fig.(3.4)].

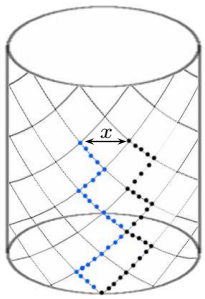


Figure 3.4: Paths in a strip with periodic boundary conditions.

The boundary conditions now in place are (for the case of L even)

$$\begin{aligned} d_N(L/2) =& 2d_{N-1}(L/2) + d_{N-1}(L/2-1) \\ d_N(L/2-1) =& 2d_{N-1}(L/2-1) + d_{N-1}(L/2-2) + 2d_{N-1}(L/2) \\ d_N(0) =& 2d_{N-1}(0)w + d_{N-1}(1)w \\ d_N(1) =& 2d_{N-1}(1) + 2d_{N-1}(0)\sigma + d_{N-1}(2) \end{aligned} \tag{3.14}$$

In particular when the ending distance is $x = L/2$, two of the four possible moves don't change the distance and the other two lower it to $L/2 - 1$.
In other words we enforce reflecting boundary conditions at $x = L/2$ to exploit the mirror symmetry $x \to L - x$ within the cylindrical strip. Similarly, reflecting boundary conditions are enforced at $x = 0$ because the two chains are allowed to cross each other.
Thanks to this confining it is possible to write the Eq.(3.13) in matrix form

$$d_N(x) = \sum_{y=0}^{L/2} T_{xy} d_{N-1}(y) \tag{3.15}$$

where we have introduced the transfer matrix T, which is tridiagonal and results

$$T = \begin{pmatrix} 2w & w & & & & 0 \\ 2\sigma & 2 & 1 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & \ddots & \ddots & 2 \\ 0 & & & & 1 & 2 \end{pmatrix} \tag{3.16}$$

Considering the Eq.(3.12) and the Eq.(3.15) we obtain immediately

$$\begin{aligned} Z_N =& \sum_x T_{xy} d_{N-1}(y) \\ =& \sum_x T_{xy} T_{ym} d_{N-2}(m) \\ =& \sum_x T^2_{xm} T_{mn} d_{N-3}(n) = \cdots \\ =& \sum_x T^N_{xl} d_0(l) \end{aligned} \tag{3.17}$$

where we adopt the Einstein's convention for the sum over the repeated indexes.

If in general we expand the initial state d_0 as a superposition of the transfer matrix eigenvectors $[\mathbf{v}_1, \cdots, \mathbf{v}_{L/2}]$, from Eq.(3.17) we obtain

$$\begin{aligned} Z_N =& \sum_x T_{xl}^N d_0(l) \\ =& \sum_x T_{xl}^N (a_1 v_1(l) + a_2 v_2(l) + \cdots + a_{L/2} v_{L/2}(l)) \\ =& \sum_x (a_1 \lambda_1^N v_1(x) + a_2 \lambda_2^N v_2(x) + \cdots + a_{L/2} \lambda_{L/2}^N v_{L/2}(x)) \\ =& \sum_x \lambda_1^N \left(a_1 v_1(x) + a_2 \left(\tfrac{\lambda_2}{\lambda_1}\right)^N v_2(x) + \cdots + a_{L/2} \left(\tfrac{\lambda_{L/2}}{\lambda_1}\right)^N v_{L/2}(x) \right) \end{aligned} \tag{3.18}$$

3.2.1 Free energy from the transfer matrix method

A general property of the interacting system which could be described using the transfer matrix formalism is that thermodynamic quantities will depend on its largest eigenvalue in the thermodynamic limit.
Supposing that eigenvalues are arranged in descending order, so that λ_1 is the largest one, considering the limit $N \to \infty$, in the expression that we obtained for the partition function only the first term survives. The free energy density in natural units is therefore given by:

$$\begin{aligned} f =& \lim_{N\to\infty} \frac{1}{N}(-\ln Z_N) \\ \simeq& \frac{1}{N}\left(-N \ln \lambda_1 - \ln\left(\sum_x a_1 v_1(x)\right)\right) \\ \simeq& -\ln \lambda_1 \end{aligned} \tag{3.19}$$

Importantly, the free energy density depends only on the largest eigenvalue. Hence, the thermodynamic quantities which are different derivatives of the free energy density will also depend on λ_1.
The ratio between the second and the first eigenvalues influence the time employed by the system to reach the equilibrium $\xi_\parallel$ according to

$$\frac{1}{\xi_\parallel} = \ln\left(\frac{\lambda_1}{\lambda_2}\right) \tag{3.20}$$

in turn related to the correlation length of the system by $\xi_\parallel = \xi_\perp^z$, with $z = 2$. We expect both quantities to diverge for a phase transition, with $\xi_\perp$ linear with the system size.

3.2.2 Numerical results

In this section we present the main results of the transfer matrix method applied to the two chain model. Note that all energies and free energies values given in this thesis are in unit of $\mathrm{K}_B\mathrm{T}$, whereas the temperature scale is set by $w = exp(-\epsilon/\mathrm{K}_B\mathrm{T})$ via the choice of the pairing energy parameter $\epsilon \sim -8.9$ Kcal/mol [22].
The analytical achievement discussed in section 3.1.1 tells us that the phase transition point is located at $w = 4/3$. From Eq.(3.10) we also expect the free energy density to be constant for $w \leq 4/3$, with $r_c = 1/4$, and start to decrease for $w > 4/3$ with r_c given by Eq.(3.9).
In Fig.(3.5) is shown, as a function of w, the free energy obtained from the largest eigenvalue of the transfer matrix, for different values of matrix dimension $L/2$ and compared with the expected analytical function $f(w)$. We can see, already for $L/2 = 500$, the numerical data essentially agreeing with the theoretical prediction. The vertical dashed lines correspond to the w values for which we have observed the free energy trend at varying L, as depicted in Fig.(3.6). Particularly, the free energy is shown as a function of $1/(L/2)$, so that the limit $L \to \infty$ can be seen in $1/(L/2) \to 0$. We note that in Fig.(3.6a) and Fig.(3.6b) the limit for $1/(L/2) \to 0$ is $-\ln 4$, while in Fig.(3.6c) is $\ln(z_c)$ with z_c given by Eq.(3.9) computed for $w = 1.34$, $\sigma = 1/2$ (and $g = 2$).
The results for unbound states below the transition at $w = 4/3$ can also be justified with an entropic argument: at each step the strands have four different possible moves, so that after N steps the entropy results $S = k_B \ln 4^N$. Starting from $F = U - TS$ and considering two non interacting strands ($U = 0$, $\sigma = w = 1$), the free energy density in natural units $1/k_BT$, becomes exactly $f = -\ln 4$. Since we know that for $w < w_t$ and $L \to \infty$ the free energy density must be constant, so it results $f = -\ln 4$ in the entire interval $[1, w_t]$.
At $w = w_t = 4/3$, it can be shown explicitly that $\lambda_1 = 4$ is an eigenvalue of the transfer matrix of Eq.(3.16) for all L values, as indeed shown in Fig.(3.6b).
In Fig.(3.7) we plot the relaxation time $\xi_{\|}$, rescaled to the square system size, given from Eq.(3.20) and computed for different values of $L/2$. Recalling that we expect $\xi_{\|}$ to diverge at transition point, we note that its peak raises at growing L, suggesting that the phase transition takes place in the limit $L \to \infty$.

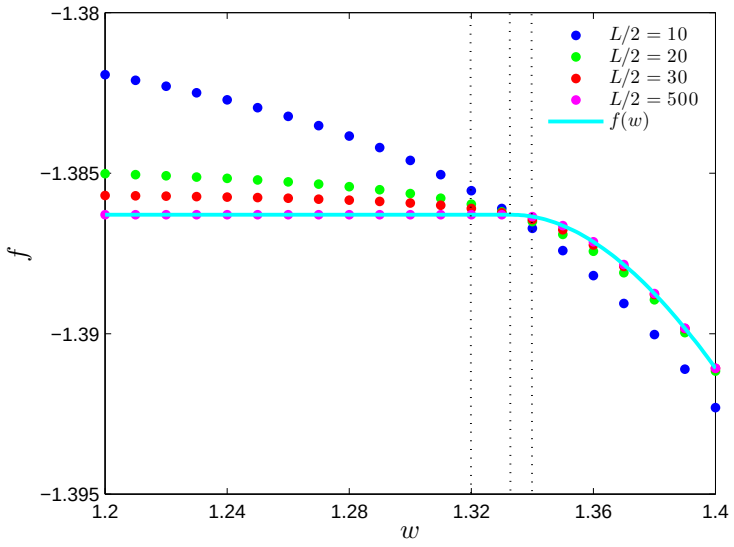
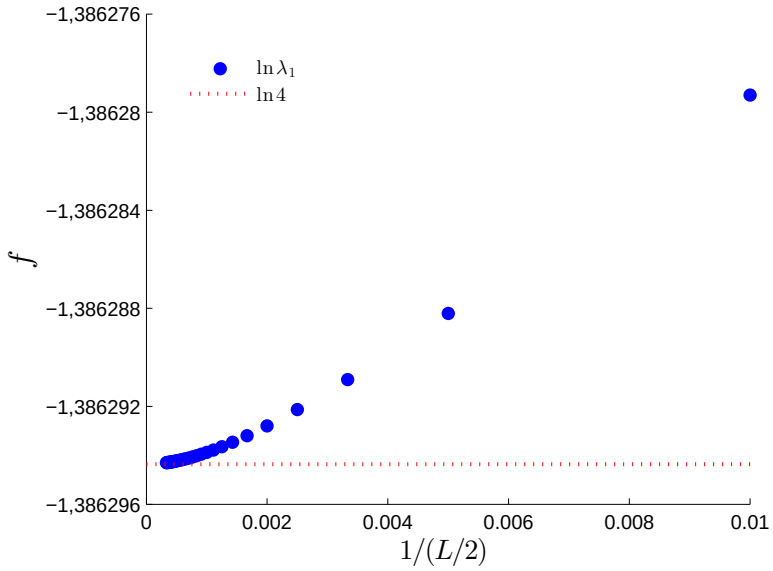
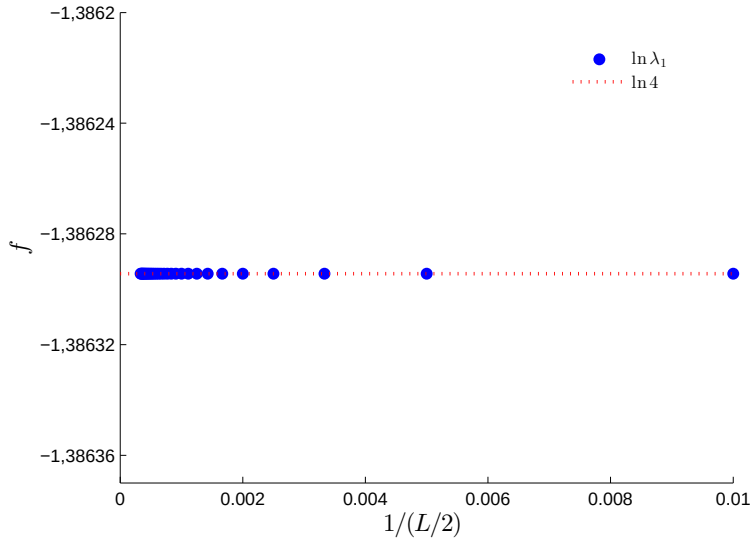


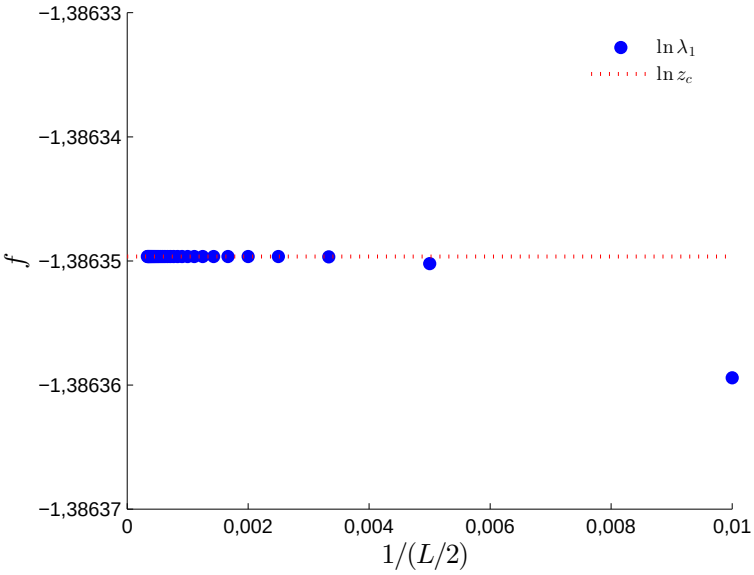
Figure 3.5: Free energy, $\sigma = 1/2$



(a) $w = 1.33$, $\sigma = 1/2$



(b) $w = 4/3$, $\sigma = 1/2$



(c) $w = 1.34$, $\sigma = 1/2$

Figure 3.6: $L \to \infty$ limit for different values of w.

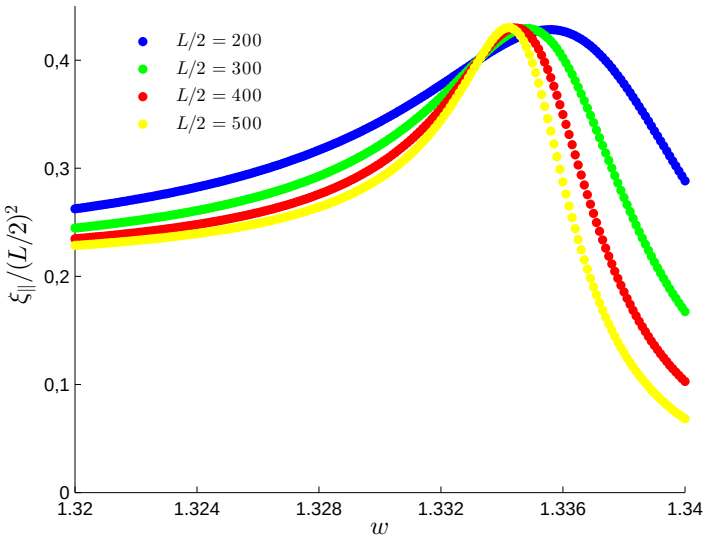


Figure 3.7: $\xi_{\|}$ rescaled to the square system size, for different values of $L/2$, $\sigma = 1/2$.

It is clear that the peaks of different curves in Fig.(3.7) move towards left when $L/2$ increases, approaching the transition point of the infinite system at $4/3$. Therefore, we could reasonably estimate w_t looking for the positions of the peaks at every L. To improve the precision of peak localization we have computed the relaxation time, at fixed L, for different w at steps of $\Delta w = 10^{-4}$ around the theoretical transition point. Successively, we have selected a subset of ten points located around the maximum and interpolated with a 4th order polynomial. The peak of the relaxation time profile was then computed by analytically finding the local maximum of the interpolated curve. This process was repeated varying $L/2$ between 200 to 2000 and the resulting w_t estimates were plotted against $1/(L/2)$ in Fig.(3.8) in order to point out the $L \to \infty$ limit at $1/(L/2) = 0$.

3.2.3 Transfer matrix eigenvectors

It is interesting to deal with the meaning of the transfer matrix eigenvectors in our specific problem. From Eq.(3.18) it is evident that in the $N \to \infty$ limit the only contribute to the partition function is given from the first eigenvector $\mathbf{v_1}$ linked to the largest eigenvalue λ_1. In particular the probability to have a

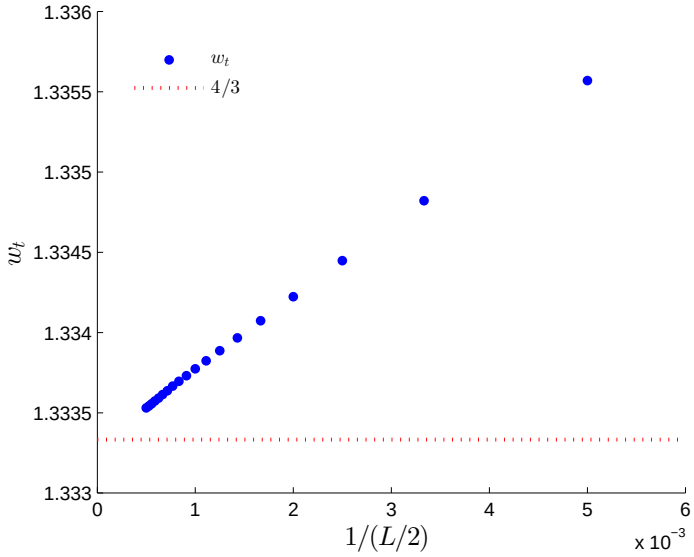


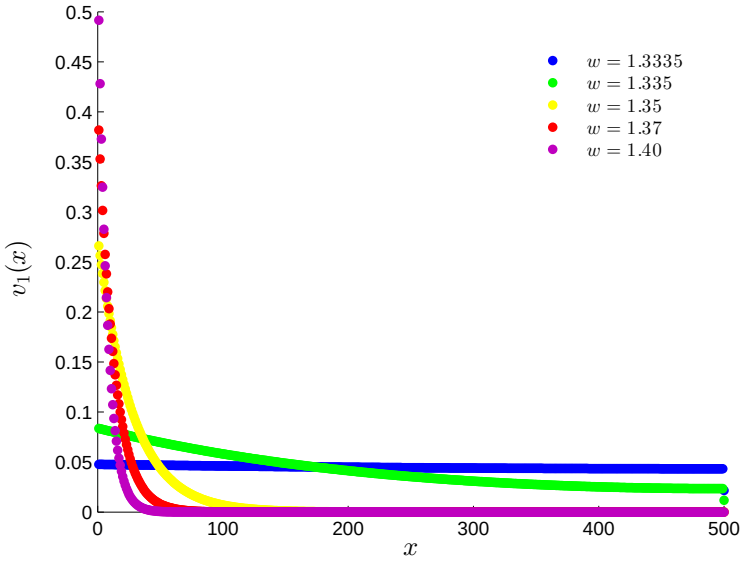
Figure 3.8: Transition point w_t against $1/(L/2)$ varying $L/2$ from 200 to 2000 at steps of 100.

certain ending distance x results

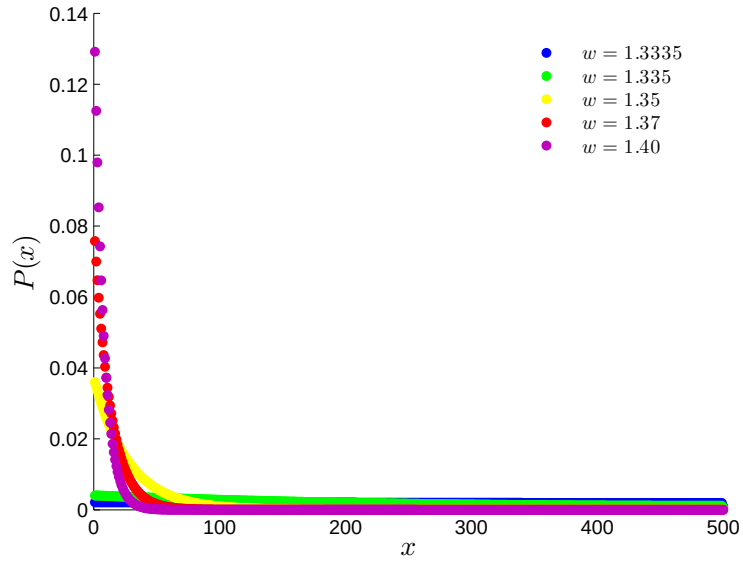
$$\begin{aligned} P(x) =& \frac{d_N(x)}{Z_N} = \frac{a_1 \lambda_1^N v_1(x)}{\sum_x a_1 \lambda_1^N v_1(x)} \\ =& \frac{v_1(x)}{\sum v_1(x)} \end{aligned} \tag{3.21}$$

Therefore, when $N \to \infty$ the probability distribution of ending distance x depends on the first eigenvector components and this vector represents the stationary state of the system. This means that whatever the initial condition may be, the system will go to position itself in the same state. If we plot the first eigenvector components against x (Fig.(3.9a)) or the correspondent probability distribution $P(x)$ (Fig.(3.9b)) for different values of w in the zipped phase ($w > 4/3$), we can see that, over the transition, the states of the system with small final distances x between strands are favoured with an exponential decay, reminiscent of the ground state wavefunction in a δ-function potential well that would be obtained in the corresponding quantum problem.

Ultimately we want to highlight the meaning of the other eigenvectors of the transfer matrix. In fact from Eq.(3.18) it is clear that these play a significant role at fixed N. Their components are indeed linked to the time spent by the system with a certain distance x before the steady state is reached after infinite



(a) First eigenvector components



(b) Probability distribution of distance x

Figure 3.9: Stationary distribution of the system.

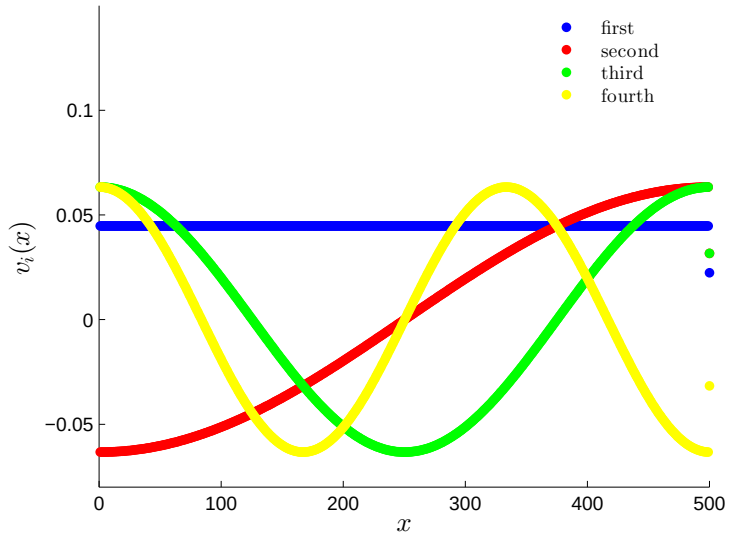


Figure 3.10: First four eigenvectors at $w = 4/3$,$\sigma = 1/2$.

steps. If we perturb the system, the time spent at some distance x could be related to the corresponding components of the eigenvectors that follows the first one. In Fig.(3.10) are showed the eigenvector components against the distance x, for the vectors linked to the first four largest eigenvalues, computed at the phase transition point, $w = 4/3$. These are reminiscent of the states found in the corresponding quantum problem of the particle in a box of size $L/2$ with reflecting boundaries.

3.2.4 Mathematical notes

We need to stress that in our discussion we have implicitly assumed the first eigenvalue of the transfer matrix to be non degenerate, so that in the Eq.(3.18) only the first term survives in the limit $N \to \infty$.
This hypothesis is satisfied thanks to the following theorem [8].

Theorem (Perron-Frobenius for irreducibile matrices)**.** *Let A be an irreducible non-negative $n \times n$ matrix with period h and spectral radius $r = max(|\lambda_i|)$, then the following statements hold*

1. *The number r is a positive real number and it is an eigenvalue of the*

matrix A, called the ***Perron-Frobenius eigenvalue***

2. *The Perron-Frobenius eigenvalue r is simple. Both right and left eigenspaces associated with r are one-dimensional*

3. *A has a left eigenvector* $\mathbf{v}_l$ *with eigenvalue r whose components are all positive*

4. *Likewise, A has a right eigenvector* $\mathbf{v}_r$ *with eigenvalue r whose components are all positive*

5. *The only eigenvectors whose components are all positive are those associated with eigenvalue r*

6. *Matrix A has exactly h eigenvalues with absolute value r*

For the sake of completeness we need to give the following definitions

Definition (Irreducibility)**.** *A matrix A is irreducible if it is not similar via a permutation to a block upper triangular matrix*

Definition (Periodicity)**.** *A square matrix A such that the matrix power* $A^{k+1} = A$ *for k a positive integer is called a periodic matrix. If k is the least such integer, then the matrix is said to have period k.*

In order to apply the theorem we need to verify that the transfer matrix T satisfies the condition of irreducibility and aperiodicity ($h = 1$). In order to do this check, it is convenient to move on the correspondent directed graph. The graph is defined by a set of nodes connected by edges, where the edges have a direction associated with them. It has exactly n vertices, where n is size of A, and there is an edge from vertex i to vertex j precisely when $A_{ij} > 0$. Then the matrix A is *irreducible* if and only if its associated graph G_A is strongly connected, that is, for every couple of states i and j exists an oriented path that connects them.
When A is irreducible, the period can be defined as the greatest common divisor of all the lengths of the closed directed paths in G_A [15].

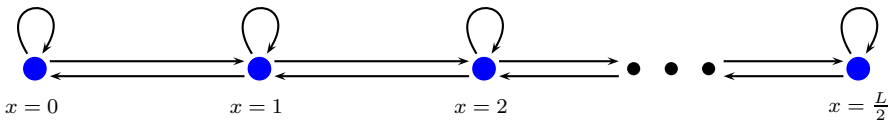


Figure 3.11: Directed graph G_T of the transfer matrix.

In Fig.(3.11) is depicted the correspondent graph of the transfer matrix T and it is clear that the conditions for irreversibility and aperiodicity are respected. In fact, the graph is strongly connected and moreover the smallest closed directed path, obtained remaining at the same state, has length 1, so that 1 is also the common divisor of all the possible paths lengths.
Another assumption of our discussion consists in supposing the coefficient a_1 in the first term of Eq.(3.18) to be non-zero. This hypothesis is equivalent to consider that the initial condition has a non zero component along the first eigenvector of the transfer matrix. Anyway, in order to obtain a partition function with a physical meaning, the initial state $d_0(x)$ of the system could not be a generic vector, but it needs to be non negative. Moreover, since we could express a_1 as a scalar product between $d_0(x)$ and $v_1(x)$, that is strictly positive from Perron-Frobenius theorem, it is straightforward to show that also a_1 must be strictly positive.

Chapter 4

Triple chain DNA models

The key theme of this work is to study the denaturation process with a basic model of triple chain DNA and verify if an analogue of the Efimov effect takes place in that system. In fact, the polymeric system that we described in Chap.2, under scaling hypothesis, is characterized by the presence of an inverse square interaction between chains, very similar to that arising in the quantum case when the Efimov states occur. Indeed, as in the quantum case the Efimov effect predicts the appearance of three particles bound states at the two particles resonance point, so we can suppose that an infinite number of triple chain DNA bound states could appear at the transition point of the two chain system.
In order to verify this fact, we will build a system analogue to the previously described two chain system, but with three chains interacting together. Obviously such system is more complicated than the previous, both from the analytical and the computational point of view. In fact, a generic configuration of three chains cannot be schematised by a sequence of bubbles and rods, as we did in the analytic model for the two chains DNA, because other different configurations are possible. On the other hand, the great numbers of achievable configurations increase a lot the computation times needed to address the problem with the transfer matrix method. Since we are interested to see the $L \to \infty$ limit, it will be appropriate to take advantage of some system symmetries, as we will see in the next section, in order to be able to compute the eigenvalues of the transfer matrix for a system of dimension L as large as possible.
Our aim is to verify that variations, predicted in Eq(2.33) for the free energy of the system, that is related to the transfer matrix eigenvalues, effectively take place.

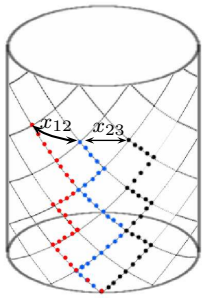


Figure 4.1: Three chains paths with periodic boundary conditions.

4.1 Three chains in a cylinder

Consider three interacting strands able to move in a cylindrical directed lattice with the rules imposed by the Poland-Scheraga model, described in Sec.(3.1.1), that is banning self intersection and permitting interactions only between points with the same distance from the origin. (Fig.(4.1)). We associate a weight w when two chains join together and a weight σ when there is a bubble opening between two strands, where w and σ are analogue parameters to those already defined for the two chain model. Instead, when three chains are involved, or in joint or in breaking a triple-stranded bond, we associate respectively a weight w^2 or $\sigma^2\rho$, where ρ is a parameter useful to introduce a different stacking interaction in the three body case. We set $\rho = 1/\sigma$ since we choose chains to interact like two pairs within the triple-stranded arrangement (Fig.(1.3)), based on the chemical nature of the Hoogsteen base pairs.

As it is clear from Fig.(4.1), this time there are three possible ending distances to describe the system states, x_{12}, x_{23}, x_{13}, which are respectively the distance between chain 1 and chain 2, between chain 2 and chain 3 and between chain 3 and chain 1. Obviously only two of these are independent, so that the partition function can depend only from them. Basing ourselves on the natural definition of distance, we choose to utilize the two smallest unsigned distances of the three. Define indeed x and y the smallest distances and z the largest one. Consider moreover that the largest distance which can be reached in a cylinder of base circumference L is $L/2$. Then we notice that the definition of z as a function of x and y must change if the condition $(x + y) \leq L/2$ or $(x + y) > L/2$ occurs.

The resulting requests are summarized in Eq.(4.1) and Eq.(4.2).

1. $x + y \leq L/2$

$$\begin{cases} x + y = z \\ x \leq z \\ y \leq z \end{cases} \implies \begin{array}{l} x \geq 0 \\ y \geq 0 \end{array} \tag{4.1}$$

2. $x + y > L/2$

$$\begin{cases} x + y + z = L \\ x \leq z \\ y \leq z \end{cases} \implies \begin{array}{l} x \geq (L - y)/2 \\ y \geq (L - x)/2 \end{array} \tag{4.2}$$

In Fig.(4.2) are depicted the domains resulting from Eq.(4.1) end Eq.(4.2), and their union that takes in all the possible configurations for the smallest distances x and y. From now on we will assume L to be a multiple of six.
To write the partition function at fixed distances $d_N(x, y)$ we have to consider eight different possible moves and their respective weights, as shown in Fig.(4.3).

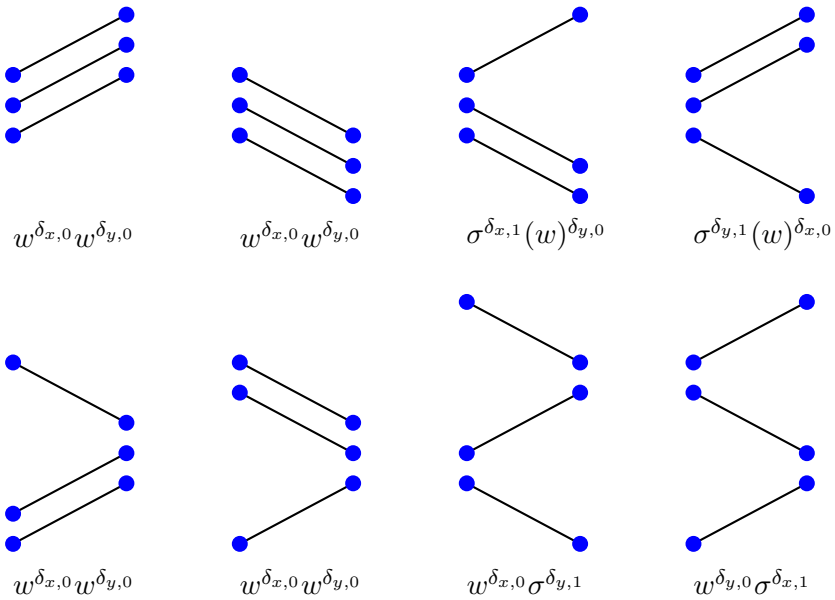
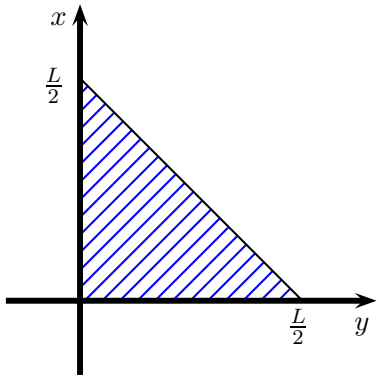
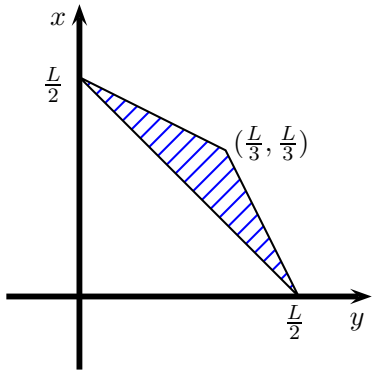


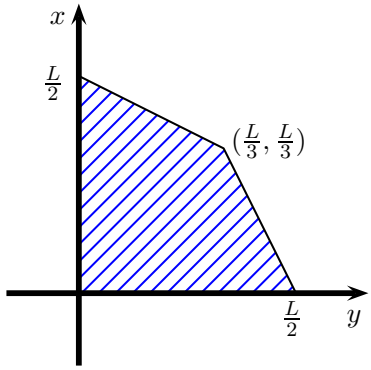
Figure 4.3: Possible moves for three strands in an oblique square lattice.



(a) Domain obtained when the condition $x + y \leq L/2$ occurs



(b) Domain obtained when the condition $x + y > L/2$ occurs



(c) Total domain

Figure 4.2: Domains of relative distances x and y.

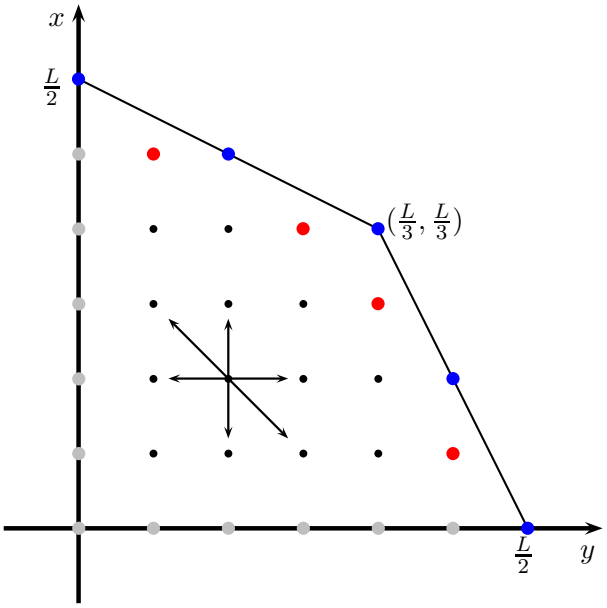


Figure 4.4: Possible pairs of distances x and y. The arrows outline the possible moves for a chain starting from a particular point, and coloured points are those for which not all moves are allowed, so that they need special boundary conditions.

In order to build the recursive equations for $d_N(x, y)$ we need to take care imposing the boundary conditions at the polygon edges in Fig.(4.2c). Obviously our system is discrete because the strands move in a lattice, so that the possible couples of distances x and y are given by the points that are located into the polygon and separated by the steps shown in Fig.(4.4). The coloured points in the picture need a special attention because they cannot allow all permitted moves; in fact they are affected by edges presence. The *blue* points are those in which $x = z$ (upper side) or $y = z$ (right lateral side), and starting from them the moves $(\uparrow, \rightarrow, \nwarrow)$ (upper side) and $(\uparrow, \rightarrow, \searrow)$(right lateral side) are denied, but there are two different ways to take the steps $(\leftarrow, \downarrow, \searrow)$ and $(\leftarrow, \downarrow, \nwarrow)$. To take this fact into account we have to add a factor 2 in the recursive equation for the $d_N(x, y)$ every time there is a contribute from such edges. In order to better figure out this arguments it is useful to see Fig.(4.5) where a graphic depiction is given.
The *red* points, besides being reached in two ways from each nearby boundary point, have also another peculiarity. In fact, starting from this points the moves $\uparrow$(under upper side) and $\rightarrow$ (under right lateral side) are banned, but there are three ways to remain in the same point instead of two. For example, if we start from the red point with coordinates $(L/2-1, 1)$ to take the step $\rightarrow$,

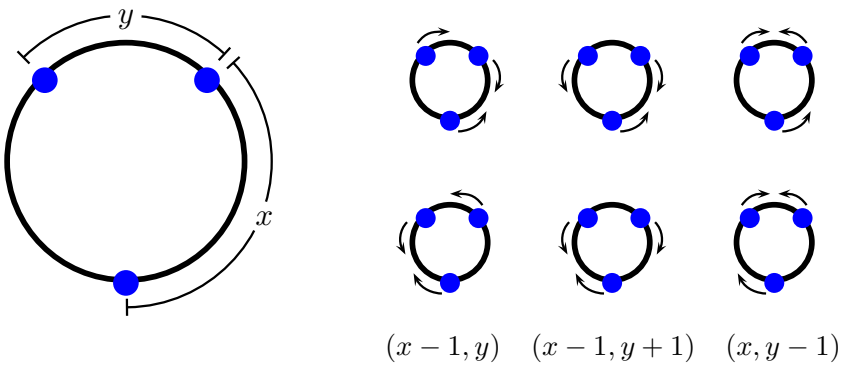


Figure 4.5: Possible moves starting from a blue point with $x = z$. The circle should depicts the cylindrical lattice seen from above, and the points are the strands extremities. It is clear that there are two possibilities to take the steps $(\downarrow, \searrow, \leftarrow)$

we see that as the distance y grow of a unity the distance z decrease to $L/2-1$ becoming the smallest together with x, so that the couple of distances which define the partition function d_N remains unchanged. To take that into account we need to introduce a factor which multiplies by 3 the terms $d_{N-1}(x, y-1)$ that contribute to $d_N(x, y-1)$ when x and y belong to the upper side, and the terms $d_{N-1}(x-1, y)$ that contribute to $d_N(x-1, y)$ when x and y belong this time to the right lateral side.

The gray points in the left lateral side and in the down side are those in which at least one distance x or y is zero. Starting from one of this points it is not possible to take the steps $(\leftarrow, \nwarrow)$ (lateral side) or $(\downarrow, \searrow)$ (down side), but there are two ways to reach the points respective to the moves $(\rightarrow, \searrow)$(lateral side) and $(\uparrow, \nwarrow)$ (down side). We need indeed to add in the recursive equation a factor 2 which multiplies every term with a factor σ related to the bubble opening and a factor 3 which multiplies the terms related to the double bubble opening. This last is to take account of the point at the origin, because the moves $(\leftarrow, \downarrow, \rightarrow, \nwarrow)$ cannot be performed starting from it, whereas $(\uparrow, \rightarrow)$ can be performed each three times.
In the coloured point that we have listed are included also the polygon vertices as special points, as we will see later.
We could further simplify the problem considering the symmetry of the system under the exchange of distances x and y. In fact, states identified by the distances (x, y) or (y, x) results clearly equivalent. Therefore, we could identify opposite states with respect to the line $y = x$ and reduce the computation of the partition function to those states for which $y \leq x$, introducing new boundary conditions (Fig.(4.6)).

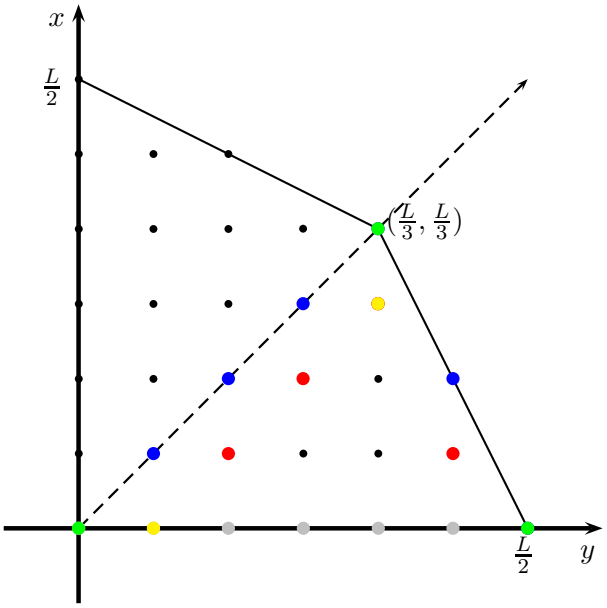


Figure 4.6: The final domain in the bottom triangle, obtained taking advantage of the system simmetry under the exchange of distances x and y.

If we consider only the bottom triangle, starting from the points on the straight line $x = y$ the moves $(\leftarrow, \nwarrow, \uparrow)$ cannot be allowed, but there are two ways to do the moves $(\rightarrow, \searrow, \downarrow)$. To take that into account we introduce a factor 2 in the terms coming from this edge which give a contribute in the equation for d_N. Moreover, we note that also starting from the new red points under the line $y = x$ there are three ways to remain in the same state.
The yellow points have the peculiarity that starting from them there are four ways to remains in the same points, and particularly for those in the bottom side there are two ways that required bubble opening and the other two not.
Finally, let's see the vertices, identified by green dots. Starting from the vertex on top, or from that on the lower left there is only one points that can be reached in six different ways. While from the vertex on the lower right there are two points accessible, the red in four ways and the gray in two.

Now we are able to write the recursive equation for the partition function at fixed distances $d_N(x, y)$. It is useful to define the following functions

$$f(y,x) = \begin{cases} 3 & \text{if } (y,x) \in \text{red dots} \\ 4 & \text{if } (y,x) = (L/3, L/3-1) \\ 2+2\sigma & \text{if } (y,x) = (1,0) \\ 2 & \text{otherwise} \end{cases} \tag{4.3}$$

$$b(y_i,x_i,y,x)=\begin{cases}2 & \text{if } (y_i,x_i)\in \text{edge } \wedge\ (y,x)\notin \text{same edge}\\ 6 & \text{if } (y_i,x_i)=(0,0)\ |\ |(y_i,x_i)=(L/3,L/3)\\ 2 & \text{if } (y_i,x_i)=(L/2,0)\ \wedge\ (y,x)=(y_i-1,x_i)\\ 4 & \text{if } (y_i,x_i)=(L/2,0)\ \wedge\ (y,x)=(y_i-1,x_i+1)\\ 1 & \text{otherwise}\end{cases} \tag{4.4}$$

Then the recursive equation for $d_N(x,y)$ can be written as

$$\begin{aligned}
&d_N(x,y)=f(y,x)d_{N-1}(x,y)w^{\delta_{x,0}}w^{\delta_{y,0}}+\\
&+d_{N-1}(x-1,y)\sigma^{\delta_{x,1}}w^{\delta_{y,0}}b(y,x-1,y,x)+\\
&+d_{N-1}(x,y-1)(\sigma)^{\delta_{y,1}}w^{\delta_{x,0}}b(y-1,x,y,x)+\\
&+d_{N-1}(x-1,y+1)w^{\delta_{y,0}}\sigma_{\delta_{x,1}}b(y+1,x-1,y,x)+\\
&+d_{N-1}(x+1,y)w^{\delta_{x,0}}w^{\delta_{y,0}}b(y,x+1,y,x)+\\
&+d_{N-1}(x,y+1)w^{\delta_{x,0}}w^{\delta_{y,0}}b(y+1,x,y,x)+\\
&+d_{N-1}(x+1,y-1)w^{\delta_{x,0}}\sigma^{\delta_{y,1}}b(y-1,x+1y,x)
\end{aligned} \tag{4.5}$$

where the points (x,y) are contained in the triangle with vertexes in $(0,0)$, $(L/3,L/3)$ and $(L/2,0)$.
We can now define the transfer matrix for the three chains system, so that

$$\begin{aligned}
d_N(x,y)&=\sum_{(x',y')}T_{[(x,y),(x',y')]}d_{N-1}(x',y')\\
d_N(i)&=T_{i,j}d_{N-1}(j)=T^N_{i,j}d_0(j)
\end{aligned} \tag{4.6}$$

where we have indexed all the states of the system with a new index i that runs over every row inside the bottom triangle of Fig.(4.6) in ascending order. An example of the transfer matrix for $L=12$ is reported in Eq.(4.8). We note that the matrix dimension M is equal to the number of points inside the triangular domain, and can be obtained by the Pick' theorem [26], resulting

$$M=A+L/2+1=(L^2/12)+L/2+1 \tag{4.7}$$

where A is the area of the triangular domain.

$$T = \begin{pmatrix}
2w^2 & w^2 & & & & & & & & & & & & & & & \\
6w\sigma & (2+2\sigma)w & w & & & & & 2w & & & & & & & & & \\
 & w & 2w & w & & & & 2w & w & & & & & & & & \\
 & & w & 2w & w & & & & w & w & & & & & & & \\
 & & & w & 2w & w & & & & w & w & & & & & & \\
 & & & & w & 2w & 2w & & & & w & w & & & & & \\
 & & & & & w & 2w & & & & & w & & & & & \\
 & 2\sigma & 2\sigma & & & & & 2 & 1 & & & & & & & & \\
 & & 2\sigma & 2\sigma & & & & 2 & 3 & 1 & & & 2 & & & & & \\
 & & & 2\sigma & 2\sigma & & & & 1 & 2 & 1 & & 2 & 1 & & & & \\
 & & & & 2\sigma & 2\sigma & & & & 1 & 2 & 1 & & 1 & 1 & & & \\
 & & & & & 2\sigma & 4\sigma & & & & 1 & 3 & & & 1 & 2 & & \\
 & & & & & & & 1 & 1 & & & & 2 & 1 & & & & \\
 & & & & & & & & 1 & 1 & & & 2 & 3 & 1 & & 2 & \\
 & & & & & & & & & 1 & 1 & & 1 & 2 & 2 & 2 & 1 \\
 & & & & & & & & & & & 1 & & 1 & 2 & & 1 \\
 & & & & & & & & & & & & 1 & 1 & & 2 & 1 & \\
 & & & & & & & & & & & & & 1 & 2 & 2 & 4 & 6 \\
 & & & & & & & & & & & & & & & & 1 & 2
\end{pmatrix} \tag{4.8}$$

Since we are in interested in studying the analogy between the Efimov quantum effect and the behaviour of tsDNA bound states, we need to underline the analogy between the Eq.(4.6) and an equation of time evolution in the quantum problem

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(0)\rangle \tag{4.9}$$

Such formal analogy, obtained by the identification of the monomer index with the imaginary time, suggests us to expect a common behaviour of the hamiltonian eigenvalues and the logarithm of the eigenvalues of the transfer matrix. Obviously this analogy could be affected by the fact that Eq.(4.9) holds for a continuous time and the other one is defined on a discrete lattice.

Chapter 5

Results and discussion

In this chapter we report the results we have been able to obtain by increasing the system dimension as much as possible. In order to optimize the numerical computing we took advantage of sparse matrices computation modules under Matlab platform (R2013a), and particularly of the *eigs* function, useful to compute the largest eigenvalues of the matrix [17]. In this way we have reached a limit system dimension of $L = 2940$ that corresponds to a transfer matrix dimension of 721.771.
In Fig.(5.1) it is shown an illustrative graph for the free energy density with varying w. The different curves correspond to different small values of L, just to highlight the behaviour of the system near to the transition point. This picture is enough to figure that in the limit $L \to \infty$ the curve before the phase transition ($w < w_t$) tends to become constant with limit value $-\ln 8 = -2,0794..$, similarly to the two chains system with $-\ln 4$. This result it is also justifiable with entropic arguments, in fact now we have eight possible moves for the strands in the lattice, instead of four. Thus we expect to find, in natural units and in the limit $L \to \infty$, $f = -\ln 8$ for $w = \sigma = 1$ and therefore for the whole stretch ($w < w_t$) before the phase transition, in the denatured phase.
In the spirit of Efimov effect we expect to find, at the point of the two chain transition $w = 4/3$, bound states for the triple chains DNA, resulting in a reduction of the free energy. This would correspond to find the largest eigenvalue of the transfer matrix bigger than eight, in the limit $L \to \infty$. In Fig.(5.2) is plotted the variation of the free energy density with respect to the denatured phase ($\Delta f = -\ln(\lambda_1/8)$) against $1/L$. Since we observe it to converge to a negative value, Fig.(5.2) is in agreement with our prediction of the analogy with the Efimov effect.

As in the case of double chains we plot in Fig.(5.3) the relaxation time $\xi_{\|}$, rescaled to the squared system size, against w for different value of L, to

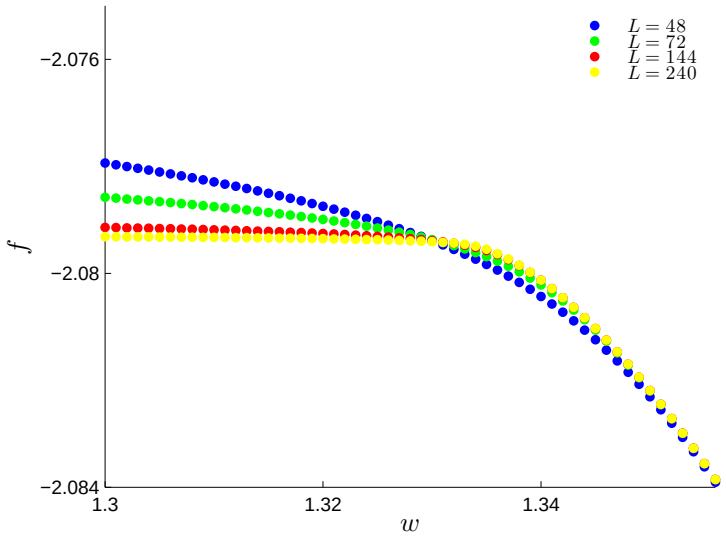


Figure 5.1: Free energy density, $\sigma = 1/2$, for varying w.

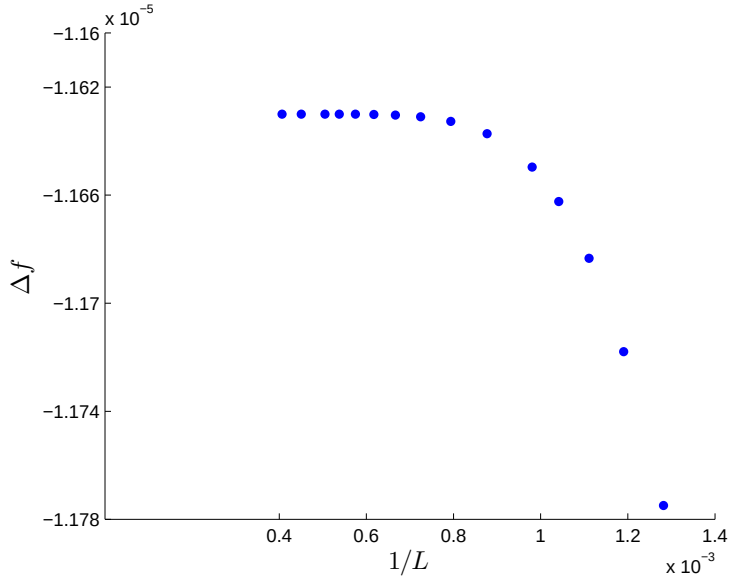


Figure 5.2: Variation of the free energy density against L, at fixed $\sigma = 1/2$, $w = 4/3$; the last value is achieved for $L = 2940$.

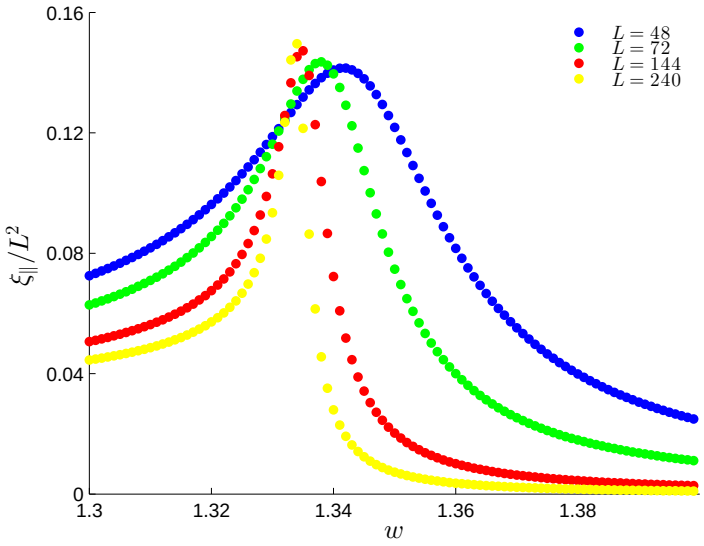


Figure 5.3: $\xi_{\|}$ rescaled to L^2, $\sigma = 1/2$, varying w.

emphasize that we expect $\xi_{\|}$ to diverge, at the transition point, when $L \to \infty$.

To estimate w_t we fit, with a fourth degree polynomial, a set of ten points located around the maximum with variable w increments and we look for the maximum $\xi_{\|}$. In Fig.(5.4) we compare the achieved results with those obtained from the two chain system with the same L; it is clear that the two sets of data converge at different values. Particularly the limit value of w_t of the data set for the three chain system is clearly smaller than that of the two chain case (4/3), under the reasonable assumption that the sequence be monotonic. Such shift in the transition point too gives us an evidence of the presence of at least onethree stranded DNA bound state when the double stranded transition has not yet taken place.

The problem of assessing the existence of excited trimer bound states is much more difficult and we will address it in the next section.

5.1 Extrapolation

The problem to extrapolate the limit of an sequence of data depending on a parameter L, when the latter tends to infinity, is a difficult one. Any possible choice of a particular functional form, for data dependence on L, could bias the reliability of the final outcome.

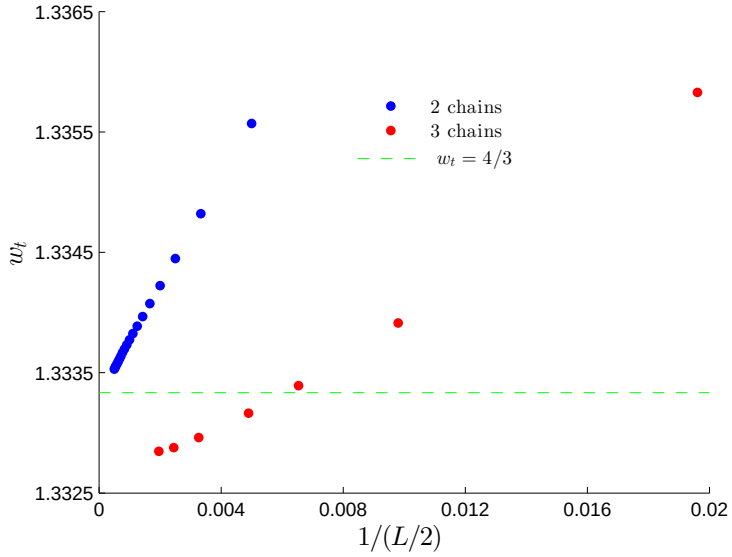


Figure 5.4: Comparison between the two and the three chains w_t estimates.

One possible choice to face this task is to assume a particular functional form for sequence dependence on L. The particular function opted for will depends on a certain number of free parameters that we need to set in order to interpolate the data points computed until a finite value L_{max}. As has already been said, it is useful to observe the sequence as a function of $1/L$ so that we will consider as our final estimate of the data sequence limit for $L \to \infty$ the function value extrapolated at $1/L = 0$.
It is clear that such method introduces a systematic error due to the choice of the function form. Our data can be considered free from error except from that introduced by numerical precision, that we estimate to be of the order of 10^{-12}. Indeed, obtaining the function from a least square interpolation, the statistical error obtained from the fit cannot be interpreted as a standard error.
To put into practice what has just been said, we fit our data with different functions, obtaining different estimates, and observe how such estimates change by varying the function form. The variability of final extrapolated values, under changing function, gives us an idea of the systematic error introduced by function choice.
In Fig.(5.5) are shown the interpolations resulting from the comparison of two and three stranded system transition point. In the two chain case a linear fit of data returns a value of w_t entirely in agreement with the analytical re-

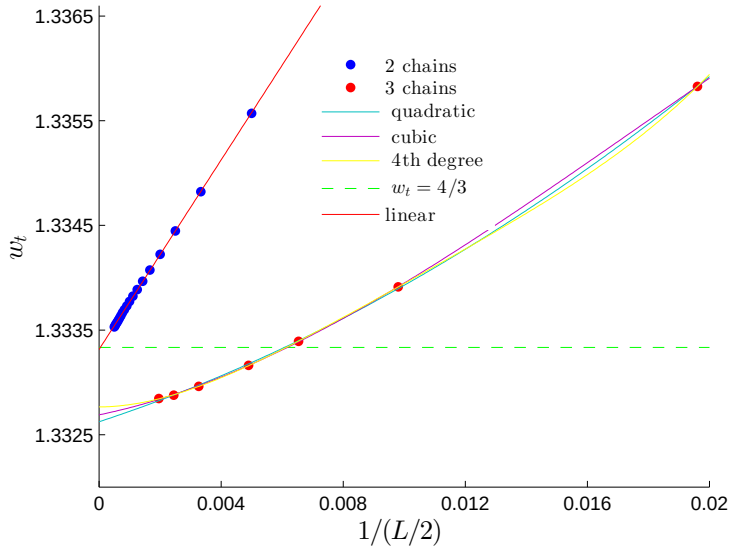


Figure 5.5: Comparison between the two and the three chains w_t limit, and resulting fit functions.

sult within the fourth digit ($w_t = 1.3333$). While for the three chain system we try to fit data with a quadratic ($w_t = 1.3326$), cubic($w_t = 1.3327$) and quartic ($w_t = 1.3328$) polynomial function . We can indeed assume a systematic error of about 10^{-4}, and give an estimate of the transition point of $w_{t3} = 1.3327 \pm 0.0001$, in agreement with the value reported in literature, $w_t = 1.3326 \pm 0.0001$ [20].

Having quantitatively checked the presence of a shift of the three chain transition point in relation to that of the two chain case, we are now interested to understand if this behaviour is a consequence of an Efimov-like effect. In order to do that we are going to search in our polymeric system, as main features of the quantum Efimov effect, the presence of an infinite number of three particles bound states which obey a geometrical scaling law (Eq.(2.17)). In the transfer matrix theory at the thermodynamic limit we have seen only the largest eigenvalue to be meaningful in the free energy computation, but based on the formal analogy between Eq.(4.6) and Eq.(4.9) we expect that also the logarithm of the transfer matrix eigenvalues, following the first, could behave in the classical polymer system in the same way of the correspondent quantum energies. If that were the case, we also expect, recalling Eq.(2.33), that the quantities $\Delta f_i = -\ln(\lambda_i/8)$ share the same properties of the hamiltonian eigenvalues of the quantum problem, so that they are negative and follow

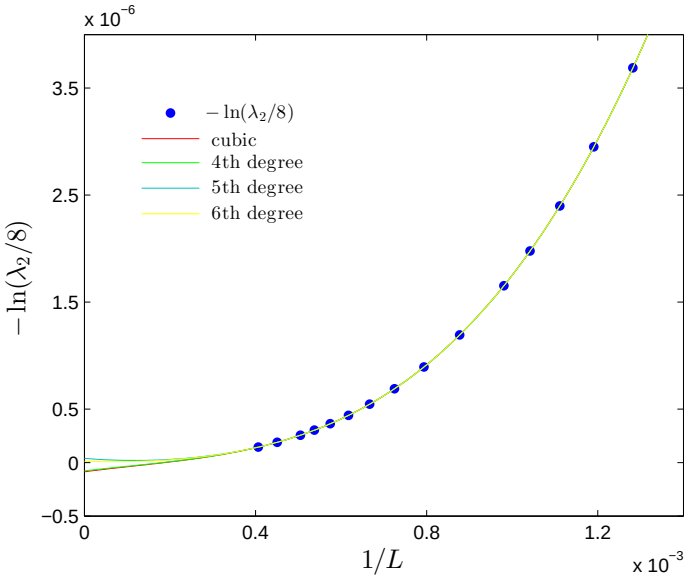


Figure 5.6: Δf_2 and different interpolating functions, $w = 4/3$, $\sigma = 1/2$, three chain system.

the analogue of the quantum mechanics scaling law

$$\frac{\Delta f_{i+1}}{\Delta f_i} = c \tag{5.1}$$

where c is a constant value dependent on system features.
The transfer matrix eigenvalues, corresponding to the excited states, result smaller than eight even at the largest value of L that we have been able to reach. In order to verify if the quantities Δf_i may turn out to be negative in the $L \to \infty$, we will use the extrapolation method already described. In Fig.(5.6) and Fig.(5.7) we plotted respectively Δf_2 and Δf_3 against $1/L$, with different choices of function forms for the interpolation. In Tab.(5.1) and in Tab.(5.2) are listed the extrapolated values, respectively of Δf_2 and Δf_3, by varying the polynomial function, and compared with those of the three chain non interacting case ($w = \sigma = 1$) for which we don't expect Efimov effect, since any chain behaves independently from the other chain. Similar results are obtained for the other eigenvalues.

We can note that changing polynomial order the limit values ranges around zero, so that we are not able to reach a conclusion about his sign; even if it is interesting to observe that the results obtained for the interacting system differ by four orders of magnitude with respect to that obtained for the non

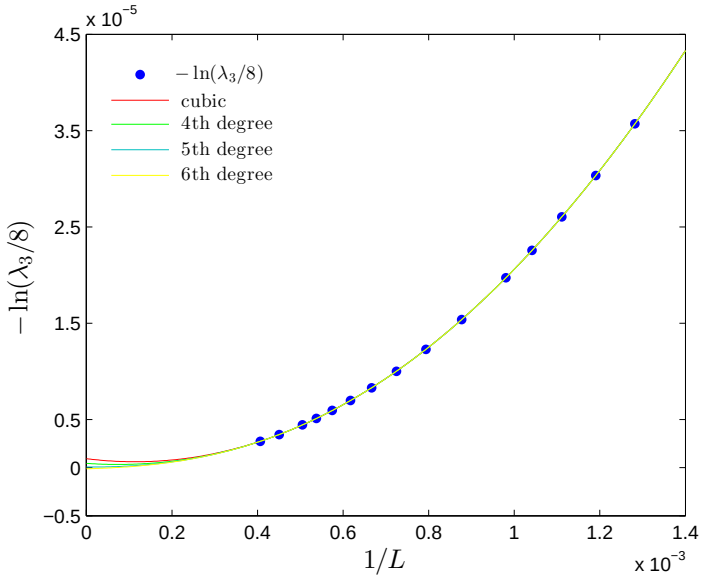


Figure 5.7: Δf_3 and different interpolating functions, $w = 4/3$, $\sigma = 1/2$, three chain system.

Table 5.1: Limit value of Δf_2 for $L \to \infty$, achieved with different interpolating functions.

Polynomial type	$w = 4/3$, $\sigma = 1/2$	$w = 1$, $\sigma = 1$
Cubic	$-8.6 \cdot 10^{-8}$	$-7.2 \cdot 10^{-12}$
4th degree	$-7.7 \cdot 10^{-8}$	$-2 \cdot 10^{-12}$
5th degree	$4 \cdot 10^{-8}$	$-6 \cdot 10^{-12}$
6th degree	$1.2 \cdot 10^{-8}$	$-1 \cdot 10^{-12}$

Table 5.2: Limit value of Δf_3 for $L \to \infty$, achieved with different interpolating functions.

Polynomial type	$w = 4/3$, $\sigma = 1/2$	$w = 1$, $\sigma = 1$
Cubic	$9.3 \cdot 10^{-7}$	$-5.5 \cdot 10^{-11}$
4th degree	$4.4 \cdot 10^{-7}$	$-7.1 \cdot 10^{-12}$
5th degree	$6.1 \cdot 10^{-8}$	$-1.7 \cdot 10^{-11}$
6th degree	$-1.2 \cdot 10^{-7}$	$-1.1 \cdot 10^{-11}$

interacting one.
To summarize, our data do not provide any evidence of the infinite series of Efimov trimers. If present, Efimov excited bound states have an energy higher than $-10^{-7}\mathrm{K}_B\mathrm{T}$.

5.2 Eigenvectors

The fact that in the triple chain system the partition function depends on two variable x and y does not change obviously the transfer matrix eigenvectors role in the problem. Indeed, the components of the first eigenvector are linked to the probability distribution of the two ending distances x and y. We remind that x, y are two smallest distances between the three possible chain pairs.
Recalling that we have indexed the system states running over every row inside the triangular domain in ascending order, in Fig.(5.8) we plot the components of the first eigenvector normalized in such triangle for different w values and $L = 480$. Before the transition point the probability to find the system in the three-stranded state $x = y = 0$ is not significant yet, and the system prefers the states characterised by large distances between strands. After transition to the three-stranded phase ($w > w_t$) we observe that the states close to the origin $x = y = 0$ are more probable, also at the critical point of the two chain system, $w = 4/3$. Noticeably, our system has a direct transition from unbound to three-stranded states, without an intermediate phase in which two chains bound states are favoured.
In Fig.(5.9) are reported the eigenvectors related to the first six largest eigenvalues of the transfer matrix at the critical point $w = 4/3$ of the two chain system. In order to enabling us to compare them, we multiply each vector for the same normalizing constant of the first one.
The statistical meaning of those eigenvectors in this problem remains the same

of previous case, that is their components give us information about the time employed by the system to approach the equilibrium state in the thermodynamic limit.

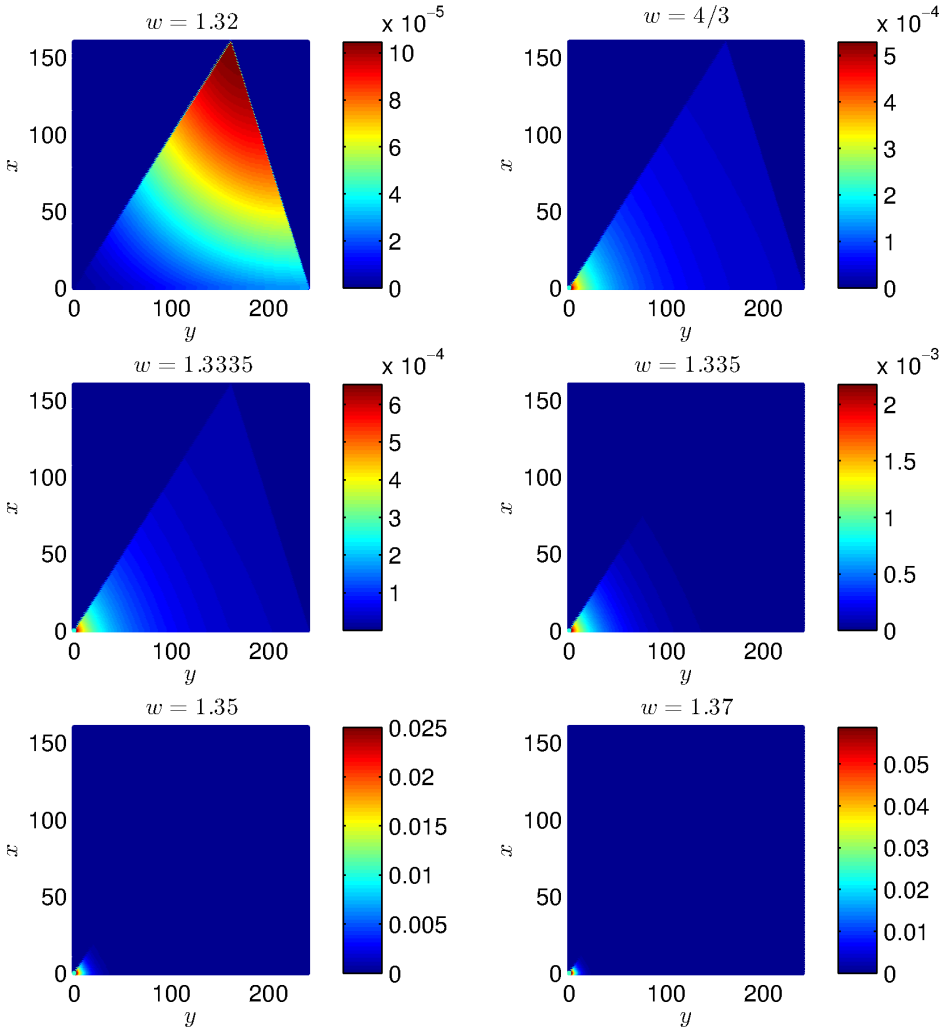


Figure 5.8: Components of the first normalized transfer matrix eigenvector, computed at $L = 480$, $\sigma = 1/2$ for different w values. Note the different scales in the colour bars.

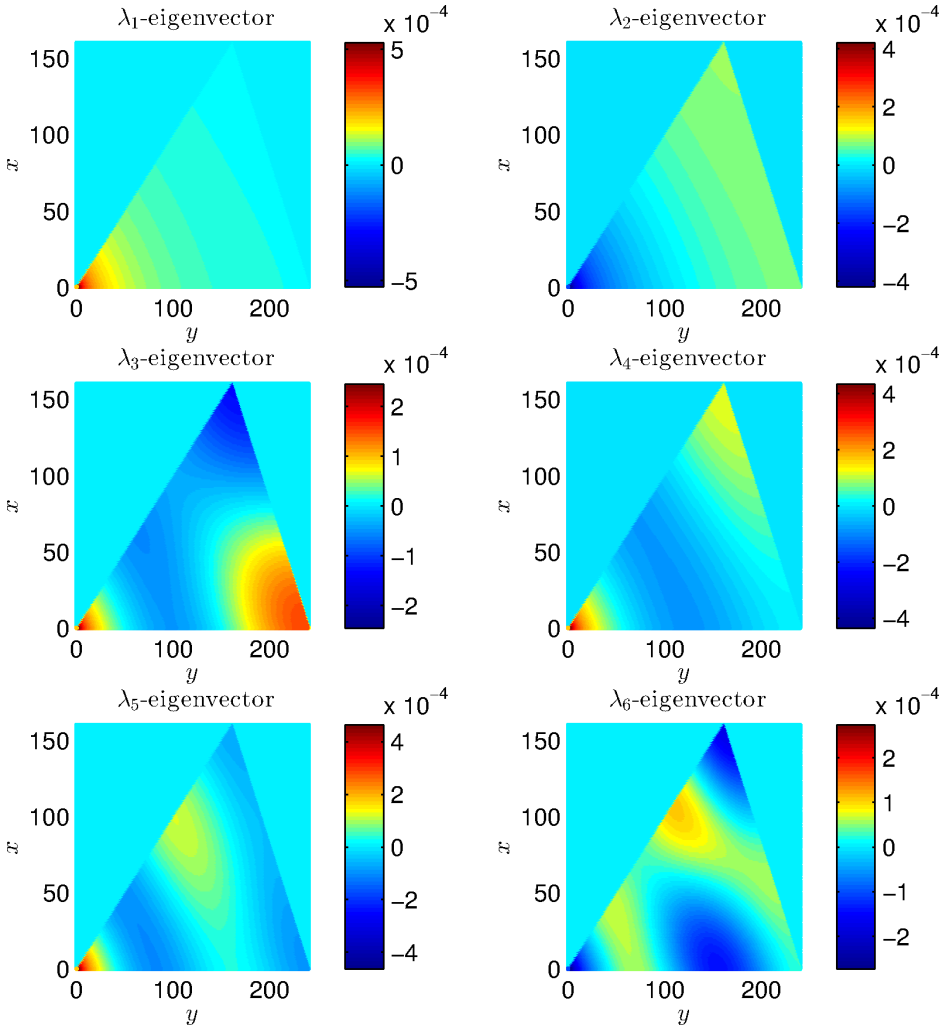


Figure 5.9: Components of the first six transfer matrix eigenvectors, computed at $L = 480$, $\sigma = 1/2$, $w = 4/3$.

We observe that only for the λ_1-eigenvector all components are positive, in agreement with the Perron-Frobenius theorem.
The λ_1-eigenvector describes a three-stranded bound state, as expected at $w > w_t$. Other eigenvectors describe oscillatory states, that are similar to those obtained for the non-interacting three-chain ($w = \sigma = 1$). This is consistent with the eigenvalues analysis performed in the previous section.

Chapter 6

Conclusions and future perspectives

In this work we studied the denaturation transition of double-stranded and triple-stranded DNA-like polymers setting up the transfer matrix technique within simplified 1+1 directed lattice models.
We obtained results that give us a new evidence supporting the notion that a shift in the transition point takes place moving from the two chain to the three chain system. The achieved value of the transition point for the triple stranded DNA is in agreement with that previously reported in literature of $w_t = 1.3326 \pm 0.0001$ [20], to be compared with the exact result $w_t = 4/3$ for the double-stranded system. $w = \exp(-\epsilon/K_B T)$ is an adimensional parameter related to the base-pair interaction energy ϵ. In order to have a finite denaturation temperature (that is $w_t > 1$), we set $\sigma = 1/2$, where σ is the adimensional cooperativity parameter weighting the opening of denaturation bubbles.
The transfer matrix method allows us to reach a novel result, estimating the variation in the free energy, upon addition of the third chain, in correspondence of the critical point of the two chain system, that results of the order of $-10^{-5}\mathrm{K}_B\mathrm{T}$.
These results are in line with the prediction by the polymer analogue of the quantum mechanics Efimov effect, that is the presence of three chain bound states when there is not a correspondent two chain bound state yet.
On the other hand, we have not been able to verify the analogy with the other features present in the Efimov effect, as the appearance of infinite trimer bound states and the geometrical scaling law for the energy spectrum.
The reasons for this could be due to the restrictions imposed by the model which has been used. In fact, the analogue of the quantum effect is predicted in a continuum system, but the transfer matrix method was used to study a

model defined on a simplified directed 1+1 lattice.
In order to understand the reasons of the incomplete analogy it would be interesting to study a three-dimensional model for discrete polymer chains. Such choice would enable us to discover if the discrepancies can be due to the discretization or to the dimensionality of the system. Moreover, the transfer matrix method cannot be used to study a 3-D Poland-Scheraga model and numerical simulations would be probably the best option, as done for example in [22].
Another interesting perspective would be to further explore the model used in this thesis by varying the value of the parameter ρ, controlling the cooperativity of the three chain bubble opening in the system. Indeed we can foresee that, by setting it to a small value, a mixed phase of a two chain bound state and the other one free, takes place in the DNA denaturation process, as already predicted by studying DNA model on a Sierpinsky Gasket [19]. In fact, even though very often the results obtained on the Sierpinsky gasket turn out to be true in the real world, nevertheless it would be interesting to obtain the same result in a more realistic 1+1 lattice.

Bibliography

[1] A.Fonseca and P.E.Shanley. "Efimov effect in an analitically solvable model". In: *Nucl.Phys* (1979).

[2] A.J.Guttmann and T.Prellberg. "Staircase polygons, elliptic integrals, Heun functions, and lattice Green functions". In: *Phys. Rev.* 47 (1993).

[3] Somendra M Bhattacharjee. "Unzipping DNAs: towards the first step of replication". In: *Journal of Physics A: Mathematical and General* 33 (2000).

[4] D.Poland and A.Scheraga. "Phase transitions in one dimension and the helix-coil transition in polyamino acids". In: *J. Chem. Phys* 45 (1996).

[5] Maria Duca et al. "The triple helix: 50 years later, the outcome". In: *Nucleic Acids Research* (2008).

[6] Andrew M. Essin and David J. Griffiths. "Quantum mechanics of the $1/x^2$ potential". In: *American Journal of Physics* (2005).

[7] Keith R. Fox and Tom Brown. "Formation of stable DNA triplexes". In: *Biochemical Society Transactions* (2011), 629–634.

[8] Georg Frobenius. "Ueber Matrizen aus nicht negativen Elementen". In: *Sitzungsber. Königl. Preuss. Akad. Wiss.* (1912), pp. 456–477.

[9] Felsenfeld G, Davies DR, and Rich A. "Formation of a three-stranded polynucleotide molecule". In: *J. Am. Chem. Soc* (1957).

[10] K. Hoogsteen. "The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine". In: *Acta Crystallographica* 16.9 (1963), pp. 907–916.

[11] Watson JD and Crick FH. "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid". In: *Nature* (1953), 737–738.

[12] J.M.Blatt and V.F.Weisskopf. *Theoretical Nuclear Physics*. 1952.

[13] Rajat K.Bhaduri. "An elementary exposition of the Efimov effect". In: *American Journal of Physics* (2011).

[14] Gerasim K.Iliev. "Phase transitions in polymeric systems: a directed walk study". In: *PhD thesis, University of Toronto* (2008).

[15] Bruce Kitchens. *Symbolic dynamics: one-sided, two-sided and countable state markov shifts*. Ed. by Springer. 1998.

[16] T. Kraemer et al. "Evidence for Efimov quantum states in an ultracold gas of caesium atoms". In: *Nature* (2006).

[17] R.B. Lehoucq and D.C. Sorensen. "Deflation Techniques for an Implicitly Re-Started Arnoldi Iteration". In: *SIAM J. Matrix Analysis and Applications* (1996), 789–821.

[18] Mandelkern M et al. "The dimensions of DNA in solution". In: *J Mol Biol* (1981).

[19] Jaya Maji et al. "Melting behavior and different bound states in three-stranded DNA models". In: *Physical review* 89 (2014).

[20] Jaya Maji et al. "When a DNA triple helix melts: an analogue of the Efimov state". In: *New Journal of Physics* (2010).

[21] M. Mandel and J. Marmur. "Use of Ultraviolet Absorbance-Temperature Profile for Determining the Guanine plus Cytosine Content of DNA". In: *Methods in Enzymology* (1968), pp. 198–206.

[22] D. Marenduzzo et al. "Different pulling modes in DNA overstretching: A theoretical analysis". In: *Physical review* 81 (2010).

[23] Mirkin et al. "DNA H form requires a homopurine-homopyrimidine mirror repeat". In: *Nature* (1987).

[24] M.Masetto. "Denaturazione termica del DNA a tre eliche:un approccio meccanico statistico". In: *Bachelor's degree thesis, University of Padua* (2011).

[25] Rajeswari MR. "DNA triplex structures in neurodegenerative disorder, Friedreich's ataxia". In: *journal of biosciences* (2012), pp. 519–32.

[26] Georg Pick. "Geometrisches zur Zahlenlehre". In: *Sitzungsberichte des deutschen naturwissenschaftlich-medicinischen Vereines für Böhmen "Lotos"* (1899), 311–319.

[27] S.A.Coon and B.R.Holstein. "Anomalies in quantum mechanics: The $1/r^2$ potential". In: *Am.J.Phys.* (2002).

[28] V.Efimov. "Energy levels arising from resonant two-body forces in a three-body system". In: *Physics Letters B* 33 (1970), pp. 563–564.