Statistical Mechanics of Protein Complexed and Condensed DNA

Dissertation
zur Erlangung des Grades

"Doktor der Naturwissenschaften"

am Fachbereich Physik
der Johannes Gutenberg-Univerität
in Mainz

Igor Kulić
geb. in Belgrad

Mainz, den 24. Mai 2004

Datum der Prüfung : 11.11.2004

Zusammenfassung

In der vorliegenden Doktorarbeit beschäftigen wir uns mit verschiedenen Themen, die mit der Verpackung und Komplexierung des DNS Moleküls durch Proteine sowie durch eingeschränkte Geometrien zusammenhängen. In diesem Kontext werden Fragen der statistischen Mechanik und der Dynamik der DNS mit verschiedene Methoden behandelt.
In der ersten Hälfte (Kapitel 2) der Arbeit wird zum ersten mal die Einzelmolekül-Zustandsgleichung (Kraft-Dehnungsrelation) der DNS mit Selbstkontakt (durch Zuhilfenahme von Pfadintegralmethoden) hergeleitet. Wir zeigen in diesem Zusammenhang, dass elastische Teilstrukturen wie Schleifen sowie Winkelrandbedingungen bei der Molekülverankerung (z.B. bei AFM Experimenten) eine starke Renormalisierung der scheinbaren Persistenzlänge induzieren und erklären damit rätselhafte Befunde bei Einzelmolekülexperimenten.
In Kapitel 3 wird das thermisch induzierte Wandern des Nukleosoms entlang der DNS untersucht. Nach eingehender Betrachtung der Experimente und theoretischer Modellierung der möglichen Mechanismen schliessen wir, dass der "Korkenziehermechanismus" die wahrscheinlichste Ursache für diesen biologisch wichtigen Prozess sein muss.
Das Kapitel 4 zeigt, dass "DNS-Spulen" - Strukturen, die aus zylindrisch oder toroidal gewundener DNS bestehen - eine bemerkenswerte kinetische Trägheit gegenüber kraftinduzierter Destabilisierung aufweisen. Wir schlagen damit eine Brücke zwischen verschiedenen Streckungsexperimenten an Nukleosomen und DNS-Toroiden sowie Simulationen und zeigen, dass die auftretenden "Kraftpeaks" in Kraft-Dehnungsexperimenten vom gleichen Ursprung sind.
Wir zeigen schliesslich in Kapitel 5, dass eine toroidal verpackte DNS (wie z.B. in Viren, DNS-Kondensaten oder in Spermienchromatin) einen bemerkenswerten Verdrillungsübergang aufweist, sobald der zugrundeliegende Torus ein kritisches Verhältnis zwischen Querschnitt und (äusserem) Radius überschreitet. Der vorgestellte Mechanismus rationalisiert und verbindet verschiedene experimentelle Funde aus den letzten 25 Jahren und eröffnet die Möglichkeit einer "topologischen Einkapselung" der DNS mit potentiellen Anwendungen in der Gentechnologie.

Summary

In this thesis I treat various biophysical questions arising in the context of complexed / "protein-packed" DNA and DNA in confined geometries (like in viruses or toroidal DNA condensates). Using diverse theoretical methods I consider the statistical mechanics as well as the dynamics of DNA under these conditions.
In the first part of the thesis (**chapter 2**) I derive for the first time the single molecule "equation of state", i.e. the force-extension relation of a *looped DNA* (Eq. 2.94) by using the path integral formalism. Generalizing these results I show that the presence of elastic substructures like loops or deflections caused by anchoring boundary conditions (e.g. at the AFM tip or the mica substrate) gives rise to a significant *renormalization of the apparent persistence length* as extracted from single molecule experiments (Eqs. 2.39 and 2.98). As I show the experimentally observed apparent persistence length reduction by a factor of 10 or more is naturally explained by this theory.
In **chapter 3** I theoretically consider the thermal motion of nucleosomes along a DNA template. After an extensive analysis of available experimental data and theoretical modelling of two possible mechanisms I conclude that the *"corkscrew-motion"* mechanism most consistently explains this biologically important process.
In **chapter 4** I demonstrate that DNA-spools (architectures in which DNA circumferentially winds on a cylindrical surface, or onto itself) show a remarkable *"kinetic inertness"* that protects them from tension-induced disruption on experimentally and biologically relevant timescales (cf. Fig. 4.1 and Eq. 4.18). I show that the underlying model establishes a connection between the seemingly unrelated and previously unexplained force peaks in single molecule nucleosome and DNA-toroid stretching experiments.
Finally in **chapter 5** I show that toroidally confined DNA (found in viruses, DNA-condensates or sperm chromatin) undergoes a transition to a twisted, highly entangled state provided that the aspect ratio of the underlying torus crosses a certain critical value (cf. Eq. 5.6 and the phase diagram in Fig. 5.4). The presented mechanism could rationalize several experimental mysteries, ranging from entangled and supercoiled toroids released from virus capsids to the unexpectedly short cholesteric pitch in the (toroidaly wound) sperm chromatin. I propose that the "topological encapsulation" resulting from our model may have some practical implications for the gene-therapeutic DNA delivery process.

List of Related Work

Chapter 2

"Equation of State of Looped Semiflexible Chains". I.M. Kulić, R. Thaokar and H. Schiessel, *in preparation*

Chapter 3

1. I.M. Kulić and H. Schiessel. "Nucleosome Repositioning via Loop-formation", *Biophysical Journal* 84: 3197 (2003)
2. I.M. Kulić and H. Schiessel. "Chromatin dynamics: Nucleosomes go mobile through twist defects", *Physical Review Letters* 91: 148103 (2003)
3. F. Mohammad-Rafiee, I.M. Kulić and H. Schiessel. "Theory of Nucleosome Corkscrew Sliding in the Presence of Synthetic DNA Ligands", submitted to *Journal of Molecular Biology (2004)*

Chapter 4

I.M. Kulić and H. Schiessel. "DNA spools under tension". To appear in *Physical Review Letters (2004).*

Chapter 5

"Twist-bend instability for toroidal DNA condensates". I.M. Kulić, D. Andrienko and M. Deserno. To appear in *Europhysics Letters (2004).*

Contents

Chapter 1

Introduction of the Main Characters

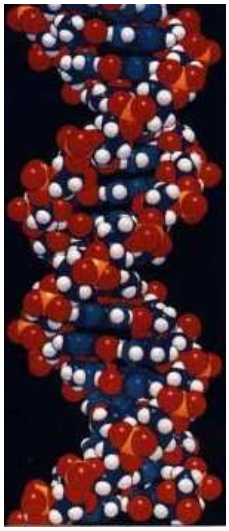
Molecular biophysics is a comparably young branch of science dealing quantitatively with living matter on the microscopic up to mesoscopic scale. By the very nature of the studied objects (that are often soft and fuzzy in their behavior) it defies all attempts of formalization. Although the concept of "life" can indeed be defined in terms of few basic criteria, the complexity of "living" phenomena can not be grasped by a limited number of axioms (as we are used from classical physical theories like classical and quantum mechanics, electrodynamics). But one should not be mislead to assume that there are no general laws of living matter at all. The first law that we meet already in elementary school is the principle of evolution. This extremely powerful principle can be brought to the simple (slightly caricaturized) form. In classical (non-living) physics a physicist is allowed to ask the question "How?" but a biologist / biophysicist may (by virtue of the evolution principle for living matter) in addition ask the question "What is it good for?" (without going into metaphysical issues).
What about other general laws in biology/ biophysics, are there any besides evolution? The answer to this question is surprisingly simple: Yes there is a myriad of general laws (most of them still awaiting discovery)- yet they are all consequences and specializations of the mighty and ubiquitous principle of evolution! But if there are myriads of laws and they are all implications of only one master principle how can they be general? The answer is that life consist of a myriad of systems, sub systems and sub-sub systems - we might call them universes without exaggeration[1]- each of them obeying similar yet in detail different general laws (general within the "class of universality" of this system).
Let us illustrate that. Take a typical representative of life, for instance a house cat lying purringly in our arms. Let us zoom in its fur. There we observe a flea. A whole colony of them happily populates their "biosphere" cat. Let us zoom further, into the flea. At a magnification of 1:100.000 we observe that the flea itself is the host for ten-thousands of micronsized bacteria which peacefully populate their universe called

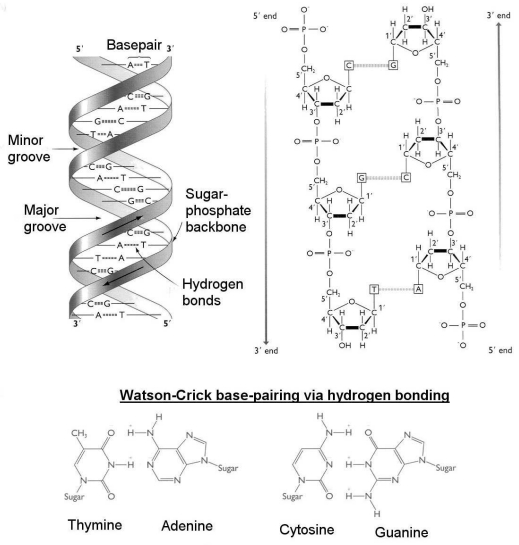
[1]The simplest cell has a much larger structural and dynamical complexity than any star in our galaxy. An organism consists of several hundred billions of cells (roughly the number of stars in our Milky Way) interacting with each other in a complex manner.

flea. If we zoom even further (by electron microscope now), into one of the bacteria we find it invaded by hundreds of viruses (so called bacteriophages) uptaking control of the bacterial metabolism right at this moment. Startled by the dramatics of this event we still continue zooming, this time inside the virus. After penetrating its protein hull we are suddenly surrounded by something very familiar. In a cloud of water and ions we recognize the blurred picture of **DNA**! Now wherever we zoom, into the cat, the flea, the bacterium or the virus all we find at the highest magnification is DNA!
This is at least how a Hollywood movie would transport the simple message: DNA is all around us, it codes the genetic information of every living being! Well to be honest this is only 99% true (true enough for a movie). There are indeed some viruses that posses RNA instead of DNA. But as viruses do not meet one criterion of life (namely they have no metabolism by themselves) they are sometimes counted only as "associate members" of the society of living creatures. All in all, DNA/RNA is the information carrier and foundation of all (known) life. And this amazing universality goes even further. Not only that the information storage medium DNA/RNA is universal also the way it codes for proteins (the underlying "compression codec") is everywhere the same: three base-pairs (three monomers) of DNA/RNA code for one amino-acid, i.e., one monomer of a protein. The famous mapping table of the DNA/RNA trimers to amino-acids[2] belongs to the deepest and most far-reaching achievements of mankind, and it plays (for terrestrial life) in the same "league of generality" as the periodic table of elements, Maxwell-equations, Pauli principle.
Now beyond this remarkable universality of information storage, the information content in the genes of different species decomposes into different universality classes each obeying individual yet very related general laws. E.g. all organisms have to replicate their genetic (DNA/RNA) content and inherit it to their descendants but some (typically lower organisms) do it in a 1:1 manner via the process of binary fission while others (typically higher organisms) shuffle their genes with other members of the same species (via sexual reproduction). In both cases biophysics can (and did) explore the underlying biology and the physics of the process and find general laws in each universality class of organisms. Another difference between higher organisms, the so called **eucaryotes** (the cat and the flea belong to them) and the simpler **procaryotes** (like bacteria) is the degree of complexity in the organization and packing of their genetic material. In procaryotes the DNA is merely loosely packed and weakly condensed by superspiralization (like a telephone cable). This simple packing mechanism works well for the comparably small procaryotic genome sizes (up to a few millions of basepairs). In contrast to that for the huge eucaryotic genomes consisting of tens of billions of basepairs, several centimeters to meters in length (per single cell!), the nature had to invent a more powerful packing mechanism. The evolution finally came up with a brilliant solution: it invented the **nucleosome** - a next to perfect complex between the DNA and a protein core able of hierarchically packing the DNA from the lowest basepair level up to full chromosomes. Coming back to the question of general principles in biology we might remark that the nucleosome is a characteristic of only one subclass of organisms - the eucaryotes. On the other hand it is an extremely general

[2] A function that maps 4^3 triplets of DNA/RNA onto 20 different monomers of amino acids and additionally some special characters like the "stop" sign.

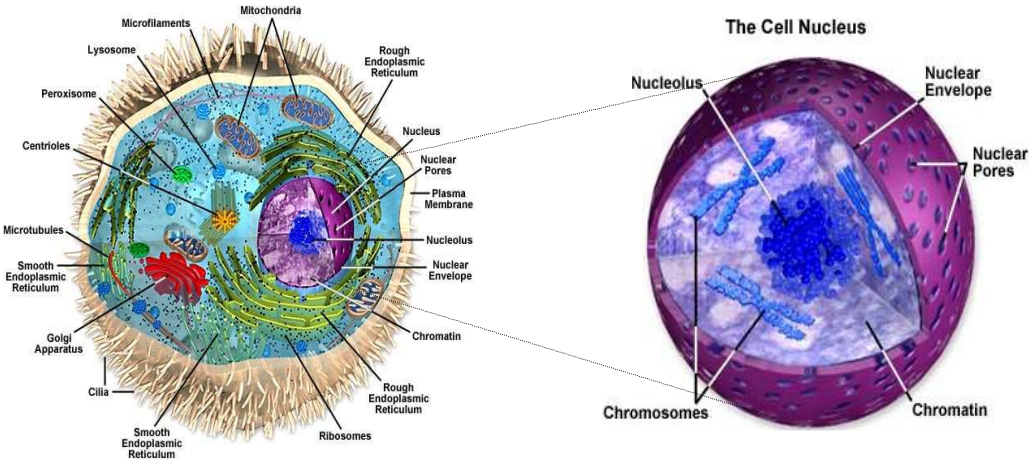
object: humans, cats, fleas and even yeast (brewing our beer) and millions other species around us have (up to small variations) the same nucleosome structure. Within the (huge) universality class of eucaryotes it is the nucleosome and not the bare "naked" DNA that rules basic processes like gene regulation, DNA duplication and DNA condensation. This is an example of a general law referring to a particular (yet huge) universality class of organisms.
In the following we will present a short yet pictorial description of the main biological "actors" which will occur in some places later in this thesis. It is not a comprehensive introduction to those structures (which can be found in later chapters and in the references in more depth) but should rather serve as an informal conceptual glossary.





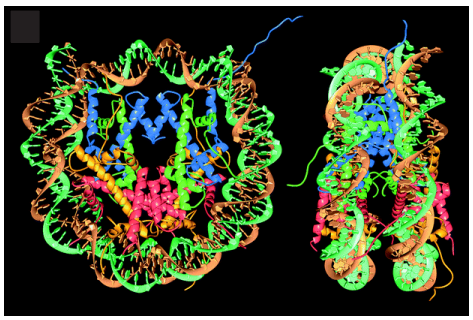
DNA: The fundamental information storage medium

Constitutes the genetic information of most organisms. Contained in every cell and reaches lengths of several meters per cell. Consists of two anti-parallel complementary strands made of a long sequence of 4 different nucleotides (adenine, guanine, cytosine and thymine) attached to a phosphate-sugar backbone. The complementarity of the strands is guaranteed by hydrogen bonds (Watson-Crick base-pairing). The typical helical structure is induced by the stacking interaction between subsequent base-pairs.

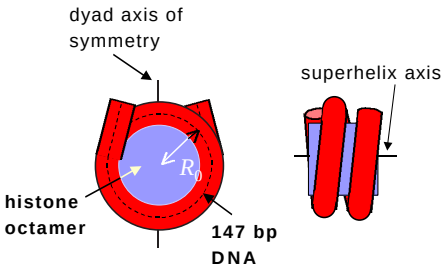


The (animal) cell

An animal cell contains several types of membrane-bound organs, or **organelles**. The **nucleus** – the brain or CPU of the cell directs cell activities and carries genetic information (in form of DNA) from generation to generation. It contains the genes packed inside of **chromatin**. Condensed chromatin appears during cell division (mitosis) in form of **chromosomes**. Other important organelles are: the **mitochondria** generating energy for the cell. Proteins are manufactured by **ribosomes.** Ribosomes are assambled in the **nucleolus** - a suborganelle of the nucleus. The **Golgi apparatus** modifies, packages, and distributes proteins while **lysosomes** store enzymes for digesting food. The entire cell is wrapped in a **lipid membrane** that selectively permits materials to pass in and out of the cytoplasm. (Pictures adapted from micro.magnet.fsu.edu/cells/).

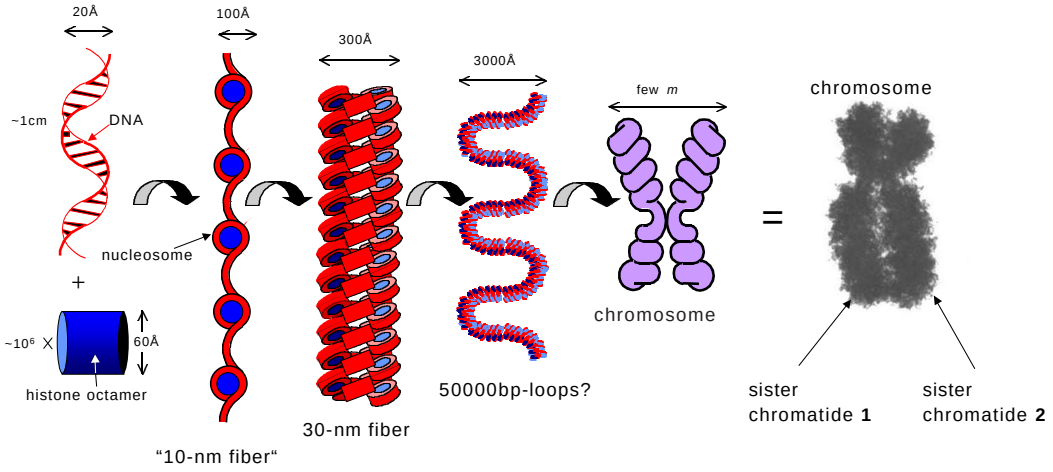


The crystal structure of the nucleosome by Luger et al. Nature 389 (1997) 251



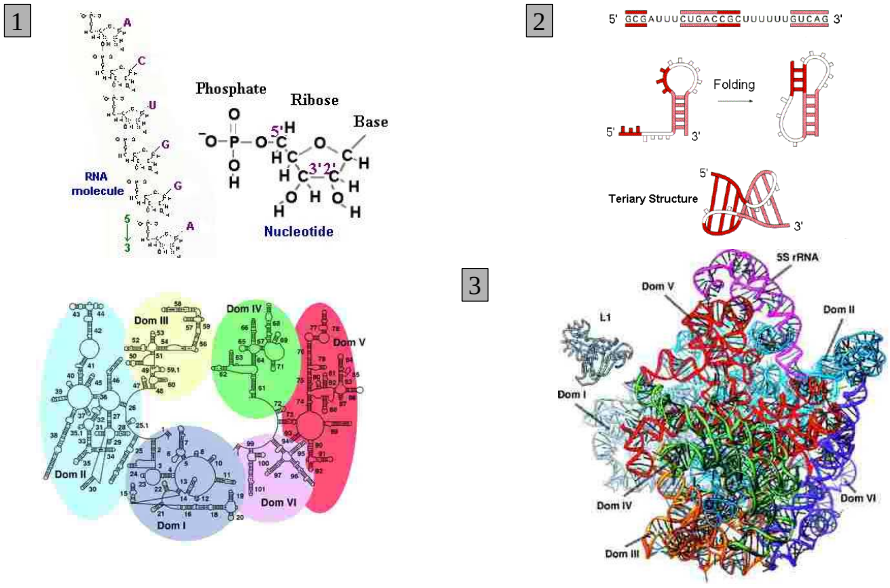
Nucleosome: The DNA packaging unit

Unlike in bacteria most of the DNA (70-90%) in all higher organisms does not appear free but packed by a protein complex called the histone octamer. The protein-DNA association between the **histone octamer** and **147 basepairs of DNA** (that wrap in 1 and 3/4 superhelical turns around it) is called the **nucleosome** (or sometimes the core-particle). The 14 discrete binding sites between the octamer surface and the DNA minor groove guarantee a high stability of the complex despite enormous DNA deformation. It compacts the bound DNA by a factor of 6 and (due to its self assembly abilities) helps the DNA to pack into higher-order structures up to chromosomes. Despite its high stability the nucleosome is surprisingly able to "slide" along DNA allowing the latter to be accessible even in the highly packed state.



Chromatin: DNA complexed with proteins

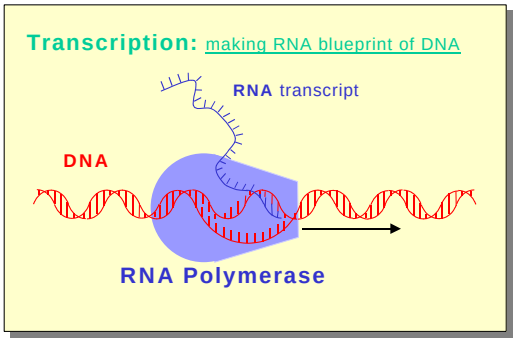
On lowest scale of condensation DNA builds a complex with **histone octamers** (in blue) forming the **nucleosomes**. The string of nucleosomes (**10-nm fiber**), already compacted by a factor of six, is then coiled into an even denser helical structure (**the 30-nm fiber**) compacting the DNA by a factor of 40. The next stage of organisation is unknown and possibly consists of large loops attached to a protein skeleton. The final product is the fully packed **chromosome** consisting of two identical copies of the same DNA molecule (**sister chromatides**). The final linear compaction factor is $\sim 10^4$!!



Large ribosomal RNA subunit: Ban et al. Science. 289 : 878 (2000)

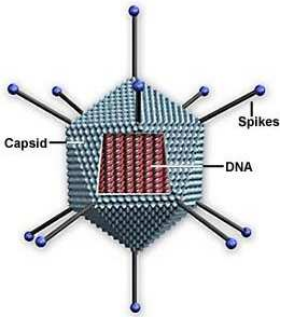
RNA: Information carrier and nanomachine: an all-rounder among biomolecules.

RNA differs chemically from DNA in only two minor points (1): its sugar molecule contains an oxygen atom not found in DNA, and RNA contains the base uracil in the place of the base thymine in DNA. Despite possessing very similar chemical structure to DNA, the conformational folding (tertiary structure) of RNA is much more variable and complex than for DNA (2). There are three types of RNA. The highly complex **ribosomal RNA** (rRNA, cf. 3) is found in the cell's ribosomes (protein assembly units). **Transfer RNA** (tRNA) carries amino acids to the ribosomes for incorporation into a protein. **Messenger RNA** (mRNA) carries the genetic blueprint copied from the sequence of bases in a cell's DNA. There are even RNAs that act as enzymes (ribozymes).

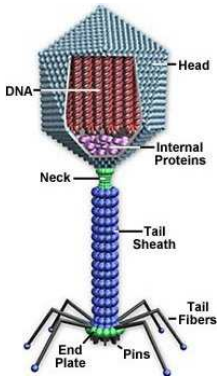


RNA Polymerase: The nanomachine performing **transcription**

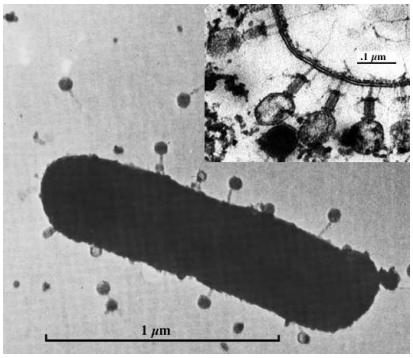
The fundamental enzyme that transcribes DNA into a 1:1 RNA offprint. It separates locally the two DNA strands and uses one of them as a complementary template for the RNA synthesis. The underlying process is called **transcription**. RNA polymerase requires energy in form of **ATP** (the basic fuel for most processes in the cell) and works far from termal equilibrium. It can generate significant **forces** of up to 30 pN, strong enough to move and remove bound DNA proteins and other obstacles on its way. Because it follows one of the two strands (that rotates around the other one once every 10 bp) it can create rotational **torques** as well. The case when the polymerase meets its most abundant obstacle - the nucleosome will become important later in this text.



Animal virus



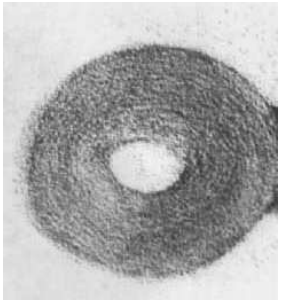
Bacteriophage



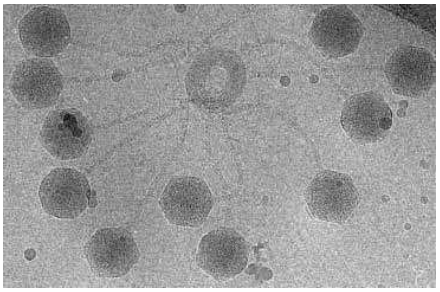
Bacteriophages in action: conquering a bacterium

from micro.magnet.fsu.edu/cells/

Virus/Phage: Creatures at the borderline of life, no metabolism – pure genetic "software". Animal/plant **viruses** and **bacteriophages** (bacterial viruses) consist essentially of pure genetic information in form of DNA or RNA and a protective protein envelope (capsid). The virus docks on its host (bacteria or animal cells) injecting its genetic information inside them. The latter integrates into the host genome and redirects its metabolism to form copies of same virus. A virus has **no own metabolism** yet highly successfully relies on that of its host, demonstrating the power of pure genetic information. The DNA containing viruses are the world record holders in DNA packing density (6 times denser than in chromatin). Typically a 15 micrometer DNA thread is fitted inside of a 50 nm diameter capsid! The DNA order inside a virus is discussed in the last chapter.



DNA Toroid



Phages eject DNA that form a Toroid

from Lambert et. al. PNAS 2000

DNA Toroids: Typical shape of a condensed single DNA

Despite its high negative charge DNA can collapse in many **poor solvents** (alcohol, small polymers like PEG, multivalent counterions) and forms bundles. In very dilute DNA solutions **single DNA** molecules undergo a **condensation** from coiled to the **toroidal bundle** state. The emerging donut shape (typical diameter around 100 nm for DNA lengths from 400 to 100000 basepairs) is a compromise between the high DNA bending rigidity (causing the hole in the centre) and the solvent induced effective DNA-DNA attraction. Together with viruses DNA toroids are the most promising candidates for gene transfer to target cells and are therefore of large biotechnological and theoretical interest. DNA inside of many viruses as well as sperm-heads is believed to have a toroidal organisation.

Chapter 2

Stretching of Looped DNA

2.1 DNA as a Wormlike Chain

The most appealing physical continuum description of the DNA molecule is the wormlike chain (WLC) model. Originating back to the first half of the last century [1] it gained renewed interest after the semiflexible nature of DNA and other (bio)polymers became clear. In the last decade it received tremendous attention as its deeper understanding became indispensable for theoretical the explanation of single molecule experiments (cf. the review [2]) that became technically feasible and extremely popular. It is not exaggerated to state that the biophysical revolution that started at that time and was feeded by the hand in hand development of theory and pioneering experiments still holds on today. To "see" single DNAs, RNAs and proteins wiggling, coiling and jumping between different states under a myriad of different conditions has not only fired our imagination but it has also deepened our physical understanding. There are a lot of surprises with general impact on general (especially non-equilibrium) statistical mechanics awaiting discovery in the single molecule world [1].
In the first part of this chapter we will review the Euler-Kirchhoff elastic description of the (constrained) ground states of DNA under tension. We explain there the remarkable analogy between an elastic rod (the DNA) and the spinning top that was discovered by Kirchhoff more than 100 years ago [5]. It is extremely useful for understanding the behavior of constrained "cold DNA". By "cold DNA" we metaphorically mean DNA in situations where the importance of its configurational entropy is negligible as compared to its elastic energy. This is typically the case for short DNA lengths (below its persistence length l_P) and large energy densities (larger than tens of k_BT's per l_P). In the second part we will switch on the temperature and see how the thermal DNA wiggling affects its behavior when we "set it on fire". We will learn that such "hot DNA" at room temperature responds purely entropically to moderate pulling forces by reviewing the well known derivation of its mechanical "equation of state", i.e. the force extension behavior of stretched DNA.

[1] A good example is the remarkable (yet widely unknown) Jarzynski equation [3] which connects measurements on single molecules far from equilibrium to the equilibrium data in a surprisingly simple and general manner. The reader is warmly recommended to discover for himself this very recent "gem" of modern physics [3] still awaiting many applications.

Having reviewed those well known basic concepts of single DNA physics we finally push forward and develop the statistical mechanics of looped DNA under tension for which we derive the equation of state (cf. Eq. 2.94). In this context we will learn how stretched DNA behaves when its new "ground state" is far from the straight configuration. The analytical machinery that is applied and developed further in this chapter has its roots in classical problems of physics like the quantum mechanical tunnelling, the statistical mechanics of dislocations in solids and nucleation of critical bubbles in overheated liquids (cf. Appendix E and references therein). The unifying concept behind all this phenomena is that of path integration in the semiclassical limit.
Despite the partly involved techniques that we apply in this chapter the main results will be stated in terms of simple analytical expressions with intuitive interpretations. Besides the force extension behavior of a DNA loop in 2-D (Eq. 2.36) and 3-D (Eq. 2.94) another interesting and potentially important result is the "renormalization" of the apparent persistence length found from loops stretching (Eq. 2.39). We show that the same results apply in the case of AFM stretching of short semiflexible polymers (shorter than 20 persistence lengths). In this case the boundary anchoring conditions become of prominent importance for the outcome of the stretching experiment and again lead to a strong reduction of the apparent persistence length (cf. Eq. 2.98). Interestingly this behavior which we might call the "ghost persistence length" effect was also recently experimentally observed in single molecule AFM stretching experiments by a group at the University of Mainz [8].

2.1.1 The Euler-Kirchhoff Elastica: the Physics of "Cold" DNA

The basic assumption of a purely elastic description of DNA (and other semiflexible polymers as well) is that the energy density of a given DNA state is given as a quadratic function of the underlying distortions from the straight state. Let us consider the simplest case and neglect the DNA twist degree of freedom at first. This can be done in cases when the DNA twist is not constrained from outside, i.e., when no torsional torques are acting on it. Then we can describe the path of the DNA of a given length L and bending constant A in 3D-space by a vector $\underline{r}(s)$ having a tangent $\underline{t}(s) = \frac{d}{ds}\underline{r}(s)$. It is convenient to choose the parameter $0 < s < L$ as the contour length and to normalize the tangent to unity $|\underline{t}(s)| = 1$. The elastic energy will then become [4]

$$E_{elastic} = \frac{A}{2}\int_0^L \left(\frac{d\underline{t}}{ds}\right)^2 ds \tag{2.1}$$

The curvature $\kappa = \frac{d\underline{t}(s)}{ds}$ has a dimension of 1/length which implies that the bending constant A has dimension energy×length. The typical value of that important material constant is around $A \approx 4\times 10^{-30} Jm$ (under physiological salt concentration $\sim 100mM$ at room-temperature). Although this is a microscopically tiny value, when expressed in units of thermal energy (at room temperature) it becomes $A = l_P k_B T$ with a lengthscale $l_P \approx 50nm$. The latter length $l_P = A/k_B T$ which is called the **persistence length** sets the upper limit to the purely elastic description of DNA as we will see in

Figure 2.1: A protein forming a loop on a short stretch of DNA

the next chapter. Here we will take a look at "cold" DNA at first, i.e., at a molecule shorter than l_P (which is true for sufficiently short DNA or low temperature) where we can neglect the entropic contributions to its behavior.
Now in a typical application of molecular biophysics the DNA is subjected to forces and geometrical constrains which are induced by the action of proteins binding to it. In the simplest case a protein loops DNA as in Figure 2.1. To describe such a situation we have to introduce an external force F acting locally on the DNA and the total stress+strain energy now writes

$$E = \int_0^L \frac{A}{2} \left(\frac{d\underline{t}}{ds} \right)^2 - \underline{F} \cdot \underline{t} ds \tag{2.2}$$

If this force is constant (e.g. of mechanical origin and acting between the two DNA ends) the problem of finding the DNA conformation reduces to the classical problem of inextensible elastic beam theory [4] of finding the energy minimizing state which satisfies given constrains and $\delta E / \delta t = 0$. In a concrete computation one would parametrize the unit tangent vector $\underline{t}$ in spherical coordinates

$$\underline{t} = \begin{pmatrix} \cos\phi \sin\theta \\ \sin\phi \sin\theta \\ \cos\theta \end{pmatrix}$$

and put the force along the $z-$axis so that the energy now writes

$$E = \int_0^L \left[\frac{A}{2} \left(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2 \right) - F \cos\theta \right] ds \tag{2.3}$$

This kind of linear elastic ansatz can be readily extended to the description of twisted DNA states by the introduction of another degree of freedom the twisting angle $\psi(s)$ in addition to $\phi(s)$ and $\theta(s)$ (cf. Fig. 2.2) and the twist-rigidity constant C:

$$E = \int_0^L \frac{A}{2} \left(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2 \right) + \frac{C}{2} \left(\phi \cos\theta + \dot{\psi} \right)^2 - F \cos\theta ds \tag{2.4}$$

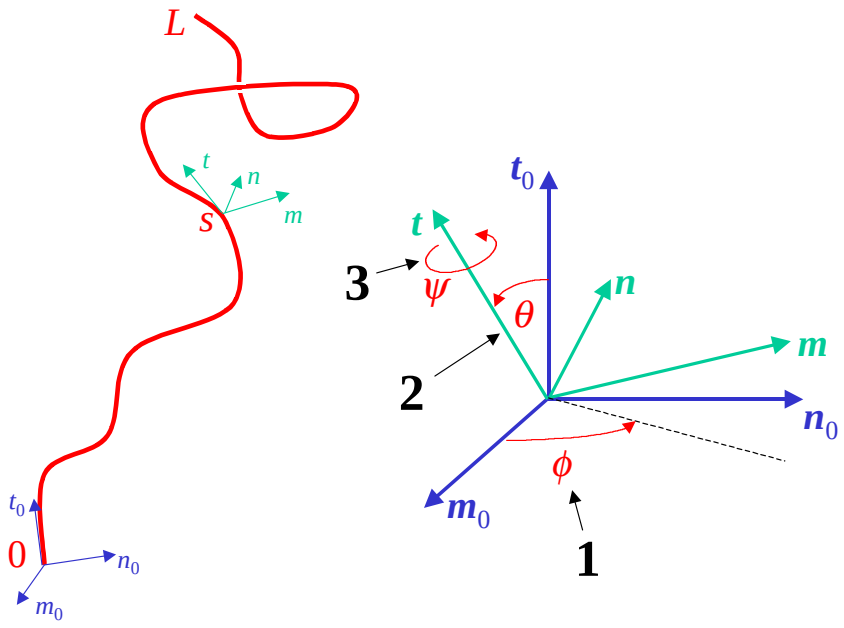


Figure 2.2: The Euler angle description of the DNA. The internal coordinate frame of the DNA is given by a material coordinate triad consisting of the DNA tangent $\underline{t}$, the normal vector $\underline{n}$ perpendicular to the helical DNA minor groove and the corresponding binormal $\underline{m}$. (This frame should not to be confused with the usual Frenet triad). The orientation of the internal DNA frame with respect to the laboratory coordinate system is given by three Euler angles θ, ϕ and ψ.

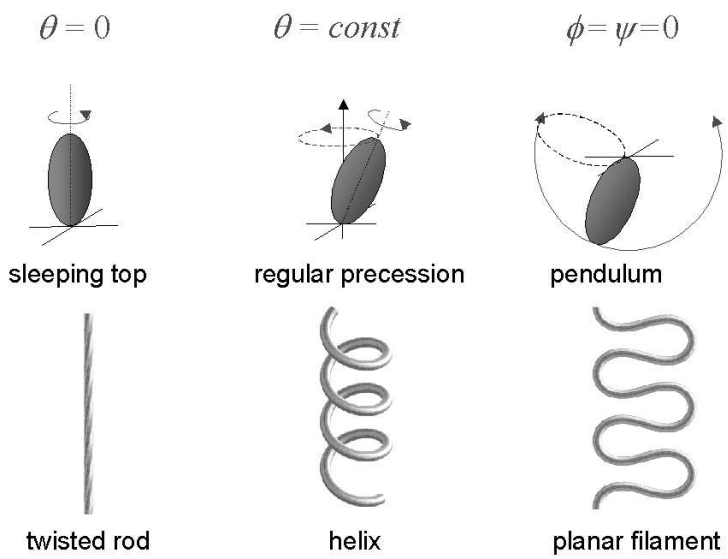


Figure 2.3: The Kirchhoff analogy between the shape of a twisted/bent rod (DNA) and the time dynamics of the spinning top

Experimentally the twist rigidity C is of the same magnitude as the bending constant ($C \approx 70 k_B T nm$) so it is not a negligible quantity. The reason why we can neglect it in some (but by far not all!) problems is that if the twist angle ψ is not explicitly constrained (no rotational torque or torsional constraining of DNA) it can always adapt so that the C multiplying term in the integral vanishes (without affecting $\phi(s)$ and $\theta(s)$).

The real beauty of the elastic energy expression in Eq. 2.4 is a powerful and very visual analogy with a well understood classical mechanical system: the spinning top! More than a century ago Kirchhoff has pointed out [5] that the **elastic energy** expression of an elastic rod can be mapped onto the **Lagrangian action** of a spinning top. The angles then $\theta(s), \phi(s)$ and $\psi(s)$ describing the local deformations of the rod along the **contour length** s become the Euler-angles $\theta(t), \phi(t)$ and $\psi(t)$ of the spinning top describing the rotation of the internal coordinates system (with respect to the space fixed frame) as functions of **time**. All the quantities appearing in Eq. 2.4 have their counterparts in the case of the spinning top. For instance the tension F becomes equivalent to (minus) gravity force (times distance to the center of mass) in the top case. The rigidity constants C and A become equivalent to the moments of inertia around the spinning and perpendicular axis of the top etc. A nice compilation of this analogy is found in ref. [6]. The Euler-Lagrange equations for Eq. 2.4 look of course the same as those for the spinning top. The resulting rod shapes are usually called **Euler-Kirchhoff filaments** (in 3D) or **Euler-elastica** (in the 2D case). Some of them are depicted in Fig. 2.3.

There is only one point of caution about this otherwise very useful analogy. The space in which the spinning top "lives" is that of the Euler-angles describing the tops position in space. In the case of a rod (DNA) the relevant space is the **integrated tangent**

space i.e. $\underline{x}(s) = \int^s \underline{t}(\theta(s'), \phi(s'), \psi(s'))\, ds'$. This seemingly harmless difference turns out to be crucial in cases when there are interactions along the DNA chain which obviously happen in the real space (rather than the tangential space). For instance in cases of DNA self-contacts (which usually appear after elastic instabilities of the DNA) the naive rod-top analogy breaks down and one is forced to use other (by far less elegant) methods to handle such problems.

2.1.2 Introduction to Statistical Mechanics of "Hot" DNA

In the previous section we have treated the case of either very short DNA (shorter than its persistence length l_P which will be explained below) or of DNA at sufficiently low temperature. Since all living beings exist at room temperatures and have DNA molecules of typically centimeter to meter length (which is even for a polymer an extraordinary dimension!) it is of course necessary to extend the methods from the previous chapter and to introduce a heat bath. We will see in the following how the conformational properties of DNA get modified and indeed change substantially at non-zero temperature. In the first introductory part we will mainly present some "state of the art" of DNA statistical mechanics. We will follow the derivations of Bresler and Frenkel [1] , Landau and Lifshits [4] , Marko and Siggia [10], and Odijk [11] dealing mainly with stretched (or short) DNA as an elementary yet instructive warm-up.
Let us start simple and forget about twisting for a moment and consider the total DNA energy as in Eqs. 3.2 and 2.3. To make our life even easier let us also switch of the force and consider a DNA in a plane. In this case we may set $\phi = 0$ in Eq. 3.2 and we are left only with

$$E = \int_0^L \frac{A}{2} \left(\frac{d\underline{t}}{ds} \right)^2 = \int_0^L \frac{A}{2} \dot{\theta}^2 ds \tag{2.5}$$

The angle $\theta(s)$ is measured with respect to an arbitrary coordinate system so we might set $\theta(0) = 0$ without restriction.
Now we can ask the following simple question: What is the mean end to end distance of a (planar) DNA molecule of length L and stiffness A at a given temperature T?
To answer this let us consider a small subsegment of the DNA with length $l \ll L$. If the latter is very short we may assume that the DNA behaves mainly elastically on that scale yet with the constraint that the boundary conditions (in this case $\theta(L) = \Delta\theta_L$) of each realization of an elastic configuration will be affected by the temperature. The minimization of Eq. 2.5, i.e., setting $\delta E = 0$ leads to the simple Euler-Lagrange equation $\ddot{\theta}(s) = 0$. By imposing the boundary conditions $\theta(0) = 0$ and $\theta(l) = \Delta\theta_l$ one obtains $\theta(s) = (s/l)\,\Delta\theta_l$ with the bending energy

$$E[\theta] = \int_0^l \frac{A}{2} (\Delta\theta_l / l)^2 \, ds = A\Delta\theta_l^2 / 2l \tag{2.6}$$

The former solution is a circular arc of length l with an opening angle $\Delta\theta_l$ and radius $R = l/|\Delta\theta_l|$. Now the energy in Eq. 2.6 is a quadratic function of the variable $\Delta\theta_l$ so by the equipartition theorem we obtain $\langle A\Delta\theta_l^2/2l \rangle = \frac{1}{2} k_B T$ or

$$\left\langle \Delta\theta_l^2 \right\rangle \approx \frac{l k_B T}{A} \text{ (for short } l\text{)}$$

The latter result means that the tangential correlation function $\langle \underline{t}(s), \underline{t}(s+l) \rangle$ of two DNA tangents at positions separated by a small distance l behaves like

$$\langle \underline{t}(s)\, \underline{t}(s+l) \rangle = \langle \cos \Delta\theta_l \rangle \approx 1 - \frac{1}{2} \left\langle \Delta\theta_l^2 \right\rangle \approx 1 - l\frac{k_B T}{2A}$$

At a twice as large distance of $2l$ we have

$$\begin{aligned}
\langle \underline{t}(s)\, \underline{t}(s+2l) \rangle &= \langle \cos(\Delta\theta_{l,1} + \Delta\theta_{l,2}) \rangle \\
&= \langle \cos \Delta\theta_{l,1} \rangle \langle \cos \Delta\theta_{l,2} \rangle - \underbrace{\langle \sin \Delta\theta_{l,1} \rangle \langle \sin \Delta\theta_{l,2} \rangle}_{=0} \\
&= \langle \underline{t}(s)\, \underline{t}(s+l) \rangle \langle \underline{t}(s+l)\, \underline{t}(s+2l) \rangle \\
\langle \underline{t}(s)\, \underline{t}(s+2l) \rangle &= \left(1 - l\frac{k_B T}{2A}\right)^2
\end{aligned}$$

In the first step we exploited the independence of the two succeeding bending angles $\theta_{l,1}$ and $\theta_{l,2}$. If we now subdivide a long DNA of length L in n short subsegments of length l we can iterate the upper procedure and get (by straightforward induction)

$$\langle \underline{t}(s)\, \underline{t}(s+nl) \rangle = \left(1 - l\frac{k_B T}{2A}\right)^n$$

So in the limit of very short subsegments l (or $n \to \infty$) we get

$$\langle \underline{t}(s)\, \underline{t}(s+L) \rangle = \lim_{n\to\infty} \left(1 - \frac{L}{n}\frac{k_B T}{2A}\right)^n = e^{-\frac{L}{2l_P}} \quad \text{with} \tag{2.7}$$

$$l_P = A/k_B T \tag{2.8}$$

For the full 3-D DNA case it is easy to see that the result gets only modified by a factor of 2 in the exponent

$$\langle \underline{t}(s)\, \underline{t}(s+L) \rangle = e^{-\frac{L}{l_P}} \quad \text{(3D case)}$$

The latter is an important result by Bresler and Frenkel [1]. It states that the tangential correlations of DNA fall off exponentially on a lengthscale l_P. The latter is in fact the (bending) **persistence length** introduced above. Looking at Eq. 2.7 we see that for $L < l_P$ one has strongly correlated tangents whereas on longer lengthscales they loose memory by thermal bombardment. Unlike the bending stiffness A which is a material constant the persistence length has an inverse temperature dependence. Visually at larger temperature the chain wiggles more so l_P decreases yet without changing stiffness A (at least in our simple macroscopic model). For DNA at room temperature l_P is around 50nm.

Having derived $\langle \underline{t}(s)\, \underline{t}(s+L) \rangle$ it is now easy to compute the mean end to end distance

$$\begin{aligned}\left\langle R^2\right\rangle &= \left\langle \left(\int_0^L \underline{t}\,(s)\,ds\right)^2 \right\rangle = \int_0^L \int_0^L \left\langle \underline{t}\,(s)\,\underline{t}\,(s')\right\rangle dsds' \\ &= \int_0^L \int_0^L e^{-\frac{|s-s'|}{2l_P}} dsds' \\ &= 4l_P^2\left(L/l_P + 2e^{-L/2l_P} - 2\right) \end{aligned} \tag{2.9}$$

From the latter equation we see that the end-to-end distance $\left\langle R^2\right\rangle^{1/2}$ behaves as L for $L/l_P \ll 1$ and $\left\langle R^2\right\rangle^{1/2} \approx 2\left(Ll_P\right)^{1/2}$ for $L/l_P \gg 1$ so that the DNA behaves as a random walker on large scales as expected. For the 3D case an analogous formula holds, namely Eq. 2.9 with l_P replaced with $2l_P$.

2.1.3 The Force-Extension Relation for a Straight DNA

An other interesting question that many people asked theoretically and experimentally [10, 11, 2] is: What is the mean end to end distance of a DNA subjected to a stretching[2] force F at a temperature T? Even without computation we can make some phenomenological observations. The applied force F and the temperature T will behave antagonistically in the stretching process: The force tries to stretch the polymer while the temperature performing constant bombardment and deflection of the DNA axis tries to contract the polymer. A (rather biased) "mediator" between these two hard opponents will be the bending stiffness A. Obviously if the latter is large the chain will be more extended so A and F share the same "desire" to stretch DNA.
Typical observables that one might be interested in is the mean and the mean squared end-to end distance, $\langle \Delta z\rangle$ and $\left\langle (\Delta z - \langle \Delta z\rangle)^2\right\rangle$, cf. Fig. 2.4. Experimentally these quantities can nowadays be measured with amazing accuracies. A typical setup for such an experiment consists of two tiny (micron sized) magnetic or polystyrene beads which are tethered to the DNA and in addition a device to excert forces on them and measure their positions (typically a low intensity laser or a strong inhomogeneous field magnet and an optical microscope able to resolve their positions). This is nowadays a standard experiment performed in dozens of labs worldwide but at the time it appeared for the first time it was a real revolution (reviewed in [2]) which subsequently accelerated the theoretical understanding of DNA.
Let us briefly rederive the well known results for the force extension behavior [10, 11]. Computationally Eq. 2.3 with the two fields θ and ϕ entering the energy in a nonlinear manner makes the problem way too difficult to be treated analytically. In order to make it more feasible we restrict ourselves to the case of small deflections of the DNA tangent with the respect to the $z-$axis (i.e., with respect to the force direction). This will be fulfilled if the molecule is short enough, the temperature low enough and/or the force large enough. In this case the energy can be "expanded around the straight

[2]The case of compressive forces is usually less relevant for DNA. It appears in the context of DNA buckling. For literature cf. Ref [12]

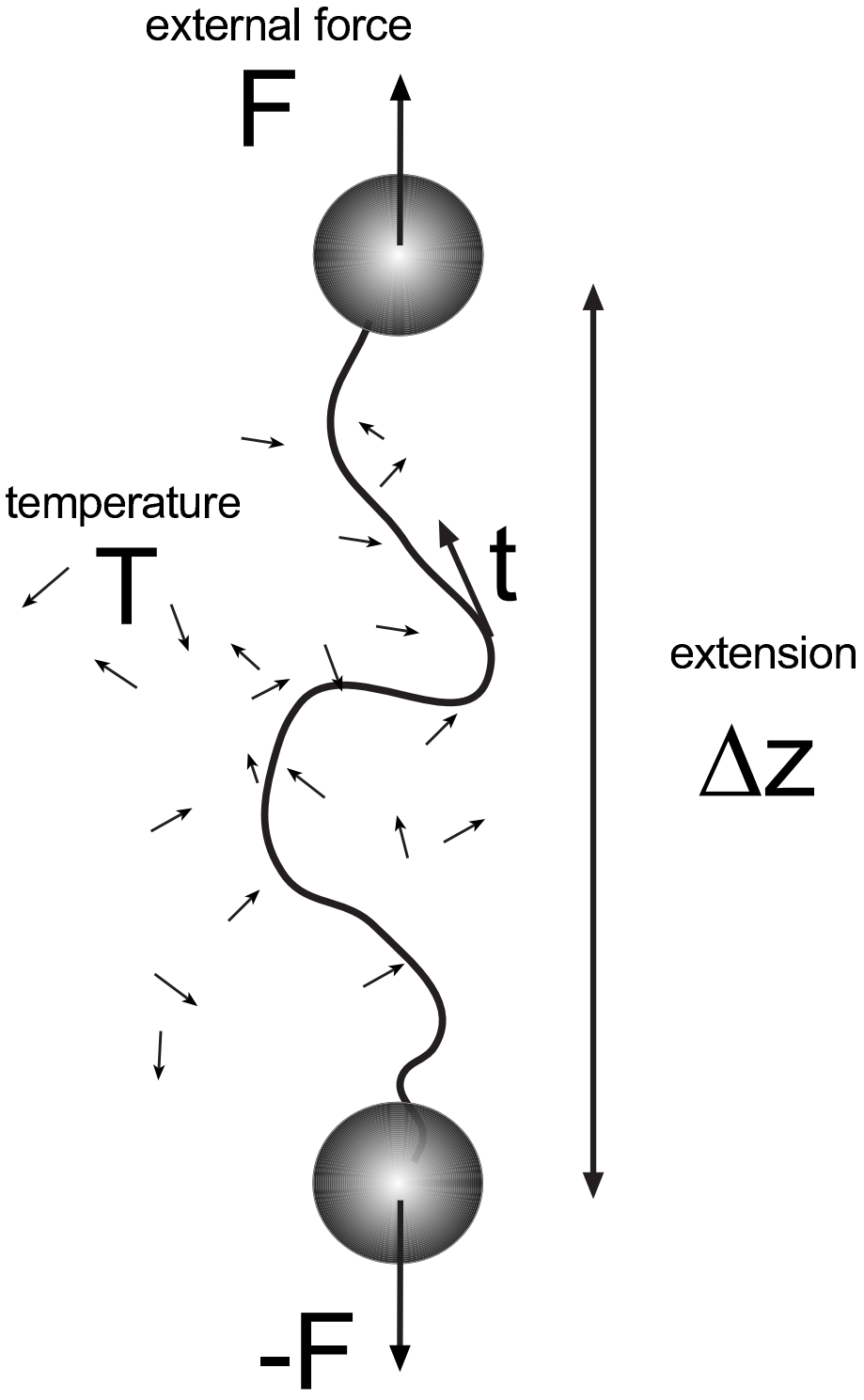


Figure 2.4: The schematic experimental setup of a single molecule DNA stretching experiment. DNA is tethered with its two ends to a pair of polystyrene (or magnetic) beads. Forces can be excerted on the latter by laser beam field gradients or (by magnetic fields) and their positions can be measured with high accuracy.

configuration". Now the somehow tricky thing about expanding Eq. 2.3 around the $z-$direction is that the coordinate system (i.e. the parametrization in θ and ϕ) is singular around the most interesting point namely $\theta = 0$! The fact that ϕ enters the expression only through its derivative $\dot{\phi}$ does well for "cold DNA" problems from the previous chapter but for doing statistical mechanics it causes serious trouble as we will see in later chapters[3].
The simplest way around this is to introduce two new angles (between the tangent projected into the two F containing perpendicular planes and to the z-axis, cf. Fig. 2.5) ϑ_x and ϑ_y and to substitute

$$\begin{aligned} \vartheta_x &= \theta \cos\phi \\ \vartheta_y &= \theta \sin\phi \end{aligned} \tag{2.10}$$

which leads to

$$\begin{aligned} \dot{\vartheta}_x^2 + \dot{\vartheta}_y^2 &= \dot{\theta}^2 + \dot{\phi}^2\theta^2 \approx \dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta \\ \vartheta_x^2 + \vartheta_y^2 &= \theta^2 \approx 2(1-\cos\theta) \end{aligned}$$

The the latter approximations hold for small $\theta \ll 1$. Using this our energy Eq. 2.3 reads

$$E\left[\vartheta_x, \vartheta_y\right] \approx \int_0^L \left[\frac{A}{2}\left(\dot{\vartheta}_x^2 + \dot{\vartheta}_y^2\right) + \frac{F}{2}(\vartheta_x^2 + \vartheta_y^2) - F\right] ds \tag{2.11}$$

$$: = H\left[\vartheta_x\right] + H\left[\vartheta_y\right] - FL \tag{2.12}$$

with a Hamiltonian H

$$H\left[\vartheta\right] = \frac{1}{2}\int_0^L \left[A\dot{\vartheta}^2 + F\vartheta^2\right] ds \tag{2.13}$$

Having the nice decoupling of the two variables ϑ_x and ϑ_y and the quadratic Hamiltonian structure of $H\left[\vartheta\right]$ the statistical mechanics problem of finding the partition function Q as an integral of $\exp\left(-E\left[\vartheta_x, \vartheta_y\right]/k_BT\right)$ over the two independent functions ϑ_x and ϑ_y is a standard exercise of path integration [13, 14]. Yet it is more illuminating to write H in Eq. 2.13 in Fourier modes of ϑ by putting[4] $\vartheta_{x/y}(s) = \sum_n \widehat{\vartheta}_{x/y,n} e^{-2\pi i n s/L}$. Then the Hamiltonian readily decouples into a sum over independent modes

$$H\left[\vartheta_{x/y}\right] = \sum_n \left(\frac{2A\pi^2}{L}n^2 + \frac{1}{2}FL\right)\left|\widehat{\vartheta}_{x/y,n}\right|^2.$$

[3]Even in the zero temperature case the singular coordinate system has caused a lot of confusion and wrong results (in highly ranked journals) which predict the stability of certain (un-) stable rod structures. This problem is discussed and resolved later on in this chapter.

[4]Here we assumed periodic boundary conditions $\vartheta_{x/y}(L) = \vartheta_{x/y}(0)$ but this of course does not change the physics for long enough L.

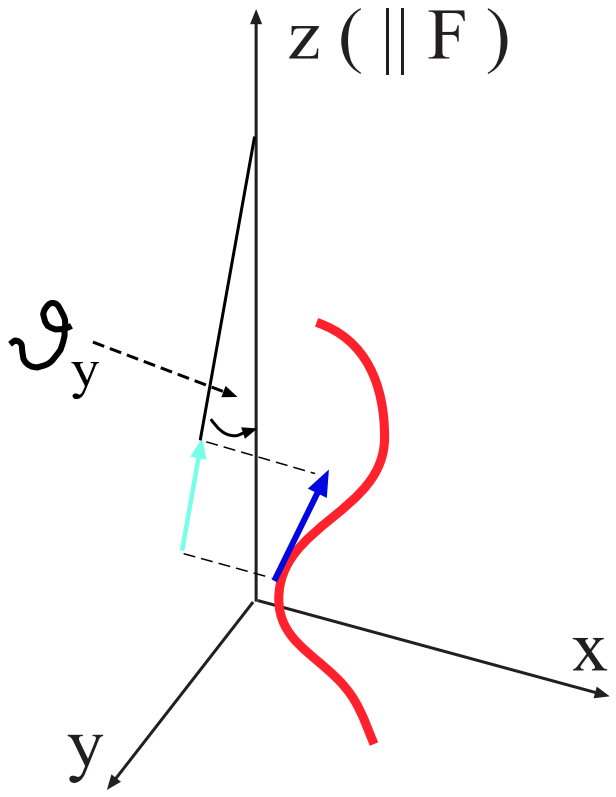


Figure 2.5: The definition of the projected tangent angles ϑ_x and ϑ_y. The projection of the DNA tangent into the y-z plane encloses the angle ϑ_y with the z-axis. ϑ_x is defined in analogous manner via a projection in the x-z plane (not shown here).

Now one can use once again (cf. previous section) the equipartition theorem for quadratic Hamiltonians stating that each mode "absorbs" $\frac{1}{2}k_BT$ energy on average. The latter implies the elementary yet important result

$$\left\langle \left| \widehat{\vartheta}_{x/y,n} \right|^2 \right\rangle = \frac{k_BT}{(4A\pi^2/L)\, n^2 + FL}$$

In order to compute the mean end-to-end distance $\langle \Delta z \rangle$ - useful formulas are

$$\begin{aligned} \Delta z &= \int_0^L \cos\theta ds \approx \int_0^L \left(1 - \frac{\vartheta_x^2(s) + \vartheta_y^2(s)}{2} \right) ds \\ \langle \Delta z \rangle &\approx L \left(1 - \frac{1}{2L} \int_0^L \left(\left\langle \vartheta_x^2(s) \right\rangle + \left\langle \vartheta_y^2(s) \right\rangle \right) ds \right) \end{aligned}$$

Now by plugging in the Fourier representation of the integrals above they simplify (by virtue of the Parceval theorem) leading to the important formula [10, 11]

$$\langle \Delta z \rangle \approx L \left(1 - \frac{k_BT}{2\sqrt{AF}} \right) \tag{2.14}$$

This can be solved for the force giving

$$F \approx \frac{(k_BT)^2}{4A} \frac{1}{\left(1 - \langle \Delta z \rangle / L\right)^2} = \frac{k_BT}{4l_P} \frac{1}{\left(1 - \langle \Delta z \rangle / L\right)^2} \tag{2.15}$$

The latter force has an obvious interpretation. It increases with temperature so it is of entropic origin. The inverse dependence on the DNA stiffness A means that the softer the chain (A smaller) the more entropy it has to loose by stretching. Consequently it is more "unwilling" to stretch and its force response is larger. Experimentally the equation Eq. 2.15 turned out to be a very powerful tool for directly and accurately determining the persistence length of DNA molecules subjected to a multitude of different solvent conditions [2].

A few words on the range of validity of Eq. 2.15 are appropriate here. This elegant force expression is valid in the limit of large forces ($F \gg \frac{k_BT}{4l_P} = 20fN$) and large relative extensions $\langle \Delta z \rangle / L \approx O(1)$. Looking at its simplicity it is somehow surprising that it is experimentally accurate for piconewton forces almost up to the point where DNA starts to melt and the wormlike chain description breaks down (around $60pN$). On the other side for very low forces (on femtonewton scale) the equation Eq. 2.15 has to be modified and is usually fitted by [21]

$$F = \frac{(k_BT)^2}{A} \left[\frac{1}{4\left(1 - \frac{\langle \Delta z \rangle}{L}\right)^2} - \frac{1}{4} + \frac{\langle \Delta z \rangle}{L} \right]$$

In the limit of small extensions $\langle \Delta z \rangle / L \ll 1$ one recovers $F = \frac{3}{2} \frac{k_BT \langle \Delta z \rangle}{l_P L}$ which is the force one expects for a Gaussian random coil perturbed by a weak force [22]. For large forces one asymptotically recovers Eq. 2.15.

2.1.4 The Partition Function for Straight DNA

The above reviewed derivation of the force extension relation Eq. 2.15 was fairly straightforward and elegant. By exploiting the equipartition theorem in Fourier space an explicit evaluation of the partition function Q and the free energy $G(F,L,T)$ was avoided there. Of course by virtue of the relation

$$\langle \Delta z \rangle = -\frac{\partial G(F,L,T)}{\partial F}$$

we can obtain $G(F,L,T)$ up to an F independent part. Nevertheless it is instructive to compute the free energy by direct evaluation of the partition function in terms of a path integral over quadratic fluctuations. This well known approach (which of course gives nothing new in the case of straight DNA stretching) is applied in the subsequent sections to a less trivial case - the stretching of looped DNA.
The partition function of the DNA chain of length L at inverse temperature β under tension F writes

$$Q(F,L,T) = \int \delta\left(|\underline{t}|-1\right) \mathcal{D}\left[\underline{t}\right] \exp\left(-\beta \int_0^L \frac{A}{2}\left(\frac{d\underline{t}}{ds}\right)^2 - \underline{F}\cdot\underline{t} ds\right)$$

The latter is a pretty nontrivial quantity to evaluate exactly even for vanishing tension F (cf. ref. [25] and refs therein). One seeming simplification is to perform the path-integral by parametrizing $\underline{t}$ by two spherical angles ϕ and θ so that the constraint $|\underline{t}| = 1$ is automatically fulfilled. By doing this another serious problem appears: in the spherical representation the integration measure has to be corrected in a highly non-trivial manner [14]. A possible way around this problem is to exploit and generalize ("quantize") the Kirchhoff kinetic analogy mentioned above and map the thermalized DNA to a quantum mechanical spinning top (or a spherical pendulum if one neglects the twist degree of freedom) as done in refs. [26]. More exactly the quantum mechanical transition amplitudes $\langle \underline{t}, 0 | \underline{t}, L\rangle$ of the QM spinning top can be mapped to the partition function $Q(F,L,T)$ of DNA under tension. The solutions of the quantum mechanical problem which are readily known can then be mapped by a transition from imaginary time $i\tau$ to the rod length L.
In our almost straight DNA case the situation is somewhat simpler. Because the deflections in the θ angle are small we can reparametrize the energy Eq. 2.10 by introducing two new angles ϑ_x, ϑ_y (cf. above). The latter has two virtues. First, the metrics of the parametrization becomes locally "flat" (for $\vartheta_x, \vartheta_y \ll 1$) and a correction of the path integral measure is not needed. Second, the partition function decouples in two one dimensional path-integrals:

$$\begin{aligned} Q(F,L,T) &= \int\int \mathcal{D}\left[\vartheta_x\right] \mathcal{D}\left[\vartheta_y\right] e^{-\beta \int_0^L \left[\frac{A}{2}\left(\dot{\vartheta}_x^2+\dot{\vartheta}_y^2\right)+\frac{F}{2}\left(\vartheta_x^2+\vartheta_y^2\right)-F\right] ds} \\ &= e^{\beta F L} Q_1^2(F,L,T) \\ Q_1(F,L,T) &= \int \mathcal{D}\left[\vartheta\right] e^{-\frac{\beta}{2}\int_0^L \left[A\dot{\vartheta}^2+F\vartheta^2\right] ds} \end{aligned} \tag{2.16}$$

The latter function Q_1 looks familiar. Indeed it is analogous to the standard path integral for the transition amplitudes of a 1-D harmonic oscillator [27, 12]. More precisely in the case of a QM harmonic oscillator the transition amplitude $\langle x_1, \tau_1 | x_0, \tau_0 \rangle$ writes [13, 14]

$$\begin{aligned} \langle x_1, \tau_1 | x_0, \tau_0 \rangle &= \int_{(x_0,\tau_0)}^{(x_1,\tau_1)} \mathcal{D}[x]\, e^{\frac{i}{\hbar}\int_{\tau_0}^{\tau_1} \frac{1}{2}M\left(\dot{x}^2 - \omega^2 x^2\right) d\tau} \\ &= \frac{1}{\sqrt{2\pi i\hbar/M}} \sqrt{\frac{\omega}{\sin\omega\left(\tau_1 - \tau_0\right)}} \\ &\quad \times \exp\left(\frac{iM\omega}{2\hbar} \frac{\left(x_0^2 + x_1^2\right)\cos\omega\left(\tau_1 - \tau_0\right) - 2x_0 x_1}{\sin\omega\left(\tau_1 - \tau_0\right)} \right) \end{aligned} \tag{2.17}$$

The path integral above differs from the one used in Eq. 2.16 by the constrained integration, i.e., the fixed boundary conditions (denoted by the upper and lower integral boundaries (x_0, τ_0) and (x_1, τ_1)) used in Eq. 2.17. By performing the substitution $\tau \to s$, $\hbar \to -ik_BT$, $M \to A$, $\omega \to i\sqrt{F/A}$ and keeping the intuitive bra-ket notation we obtain the conditional partition function

$$\begin{aligned} \langle \vartheta_L, L | \vartheta_0, 0 \rangle &= \sqrt{\frac{\sqrt{FA}}{4\pi}\beta} \sqrt{\frac{2}{\sinh\left(\frac{L}{\lambda}\right)}} \\ &\quad \times e^{-\frac{\beta\sqrt{FA}}{2} \frac{\left(\vartheta_0^2 + \vartheta_L^2\right)\cosh\left(\frac{L}{\lambda}\right) - 2\vartheta_0\vartheta_L}{\sinh\left(\frac{L}{\lambda}\right)}} \\ \text{with } \lambda &= \sqrt{A/F} \end{aligned} \tag{2.18}$$

The latter expression is proportional to the angular correlation function of the angles ϑ_0 and ϑ_L at the first and last position of the DNA molecule (up to a normalization function). The quantity $\lambda = \sqrt{A/F}$ is usually called the **deflection length** or **tension-length**[5]. The tension length λ having a dimension of length, becomes now the relevant lengthscale in the case of DNA under tension replacing the usual (tension-free!) persistence length $l_P = A/k_BT$. To clarify ourselves the meaning of λ let us consider the limiting case $L/\lambda \gg 1$ which holds for even moderately long DNA and piconewton forces[6]. In this case Eq. 2.18 writes

$$\langle \vartheta_L, L | \vartheta_0, 0 \rangle \approx (FA)^{1/4} \left(2\beta/\pi\right)^{1/2} e^{-\frac{L}{2\lambda}} e^{-\frac{\beta\sqrt{FA}}{2}\left(\vartheta_0^2 + \vartheta_L^2\right)} \tag{2.19}$$

From the latter expression (which is valid for $\vartheta_0, \vartheta_L \ll 1$) we indeed see that the chain looses orientational memory exponentially on the scale given by λ.
Now in order to obtain the partition function Q of our system (Eq. 2.16) we need to

[5]It was probably first introduced by Odijk in [27] in the context od DNA in liquid crystalline environment. We adopt here the expression "tension length" as used by Bruinsma and Rudnick in [28].

[6]For fairly moderate applied forces of say $F = 1pN$ we have $\lambda \approx 14nm$ so $L/\lambda \gg 1$ is indeed fulfilled for all single molecule experiments ($L = 1 - 15\mu m$).

integrate Eq. 2.19 over[7] ϑ_0,ϑ_L to obtain

$$\begin{aligned} Q_1(F,L,T) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \langle \vartheta_L, L|\vartheta_0, 0\rangle \, d\vartheta_L d\vartheta_0 \\ &\approx (FA)^{1/4} (2\beta/\pi)^{1/2} e^{-L/2\lambda} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{\beta\sqrt{FA}}{2}\left(\vartheta_0^2+\vartheta_L^2\right)} d\vartheta_L d\vartheta_0 \\ &= 2\sqrt{2\pi}\left(\frac{k_BT}{Fl_P}\right)^{1/4} e^{-\frac{L}{2\lambda}} \end{aligned} \tag{2.20}$$

The full partition function $Q(F,L,T) = e^{\beta FL} Q_1^2(F,L,T)$ by virtue of Eq. 2.16 finally writes

$$Q(F,L,T) \approx \frac{4\pi}{\beta\sqrt{FA}} e^{\left(\beta F - \sqrt{F/A}\right)L} \tag{2.21}$$

The free energy then reads

$$\begin{aligned} G(F,L,T) &= -\frac{1}{\beta}\ln Q(F,L,T) \\ &\approx -FL + \frac{\sqrt{F/A}}{2\beta}L + \frac{1}{\beta}\ln\left(\frac{\beta\sqrt{FA}}{4\pi}\right) \end{aligned}$$

To compute the mean extension $\langle \Delta z\rangle = -\partial G/\partial F$

$$\langle \Delta z\rangle \approx L\left(1 - \frac{k_BT}{2\sqrt{FA}} + \frac{k_BT}{2FL}\right)$$

If we now neglect the $k_BT/FL \ll 1$ term which is very small for pN forces and relevant DNA lengths ($> 1\mu m$) we recover Eq. 2.14. Having $Q(F,L,T)$ from Eq. 2.21 it is an easy exercise to derive all kind of correlation functions and moments from its partial derivatives.

Up to now we merely reviewed some well known basic facts about the wormlike chain. We dispense with their deeper elaboration and move to a somehow less trivial case in the next section.

[7] Depending on the boundary conditions we impose, ϑ_0 and ϑ_L can be linked to each other via $\vartheta_0 = \vartheta_L$ (for periodic boundary conditions). Here this unnecessary restriction (which is usually very convenient in the Fourier space representation) is dropped. Generally the boundary condition will contribute negligibly to the statistics for $L/\lambda \gg 1$.

2.2 Equation of State for Looped Semiflexible Polymers

In the last chapter we looked at the statistical mechanics and in particular the force-extension curve (the "equation of state") of a DNA molecule close to its straight ground state. The analogy with the QM harmonic oscillator reviewed above turned out to be computationally quite useful in obtaining the DNA partition function. The reason for this mapping to work was the simple quadratic shape of the Hamiltonian in Eq. 2.16. But how should one deal with non-trivial configurations (far from straight) which appear in many experimental (in vitro) and natural (in vivo) contexts? A more concrete question would be: What is the force extension curve of a DNA chain that is locally folded onto itself like in Fig. 2.6?

Such a backfolding can be caused by a ligand which brings the two distant DNA parts together but still allowing them to slide with respect to each other. Besides this somehow obvious realization, there are a multitude of different mechanisms all able of stabilizing the loop, cf. Fig. 2.6. To list just a few: a) supercoiling in twisted DNA (the same phenomenon like in a looping telephone cable), b) DNA adsorption on a surface (e.g. a membrane), c) DNA in a dense liquid crystalline environment kinetically prohibiting the loop unfolding d) DNA in a strong magnetic field trying to align it perpendicularly, e) DNA condensed by multivalent counterions and other ligands etc. Looking at the variety of experimental situations where the loop might be of relevance motivates our desire to obtain theoretically the corresponding force-extension curve in order to interpret the available experimental data. Although we will restrict ourselves to the treatment of the cases d) (for its feasibility) and e) (for its importance in DNA condensation) the other cases are treatable in the same spirit (though with additional effort).

In this section we first review and rederive the shape of the homoclinic[8] loop at zero temperature from the corresponding Euler-Lagrange equations. This simplest extended non-trivial filament shape which was already considered by Euler is essentially two-dimensional. For a given tension F the homoclinic loop turns out to be stable for arbitrary large in plane perturbations. Indeed the 2-D loop turns out to be a (static) topological soliton appearing in various reparametrizations in many contexts of contemporary physics ranging from Josephson-junctions (Sine-Gordon equations) to QM tunneling problems (cf. Appendix E). In the subsequent section we go to the third dimension by considering the out-of-plane fluctuations and how they contribute to the force response. On our way we will see that the homoclinic DNA loop is intrinsically unstable (in contrast to some false claims in literature) in the third dimension and we will learn how to introduce and deal with potentials or constraints necessary for its stabilization. We will derive the force extension relations for the DNA loop in 2-D and the 3-D case under various constraints and potentials. Finally we apply the developed results to resolve a problem that recently appeared in single molecule (AFM) stretch-

[8]The term homoclinic stems from the Kirchhoff analogy between the loop that we consider here and the homoclinic orbit of a (mathematical) pendulum which obtained just enough energy to make one full 2π rotation in an infinite time interval

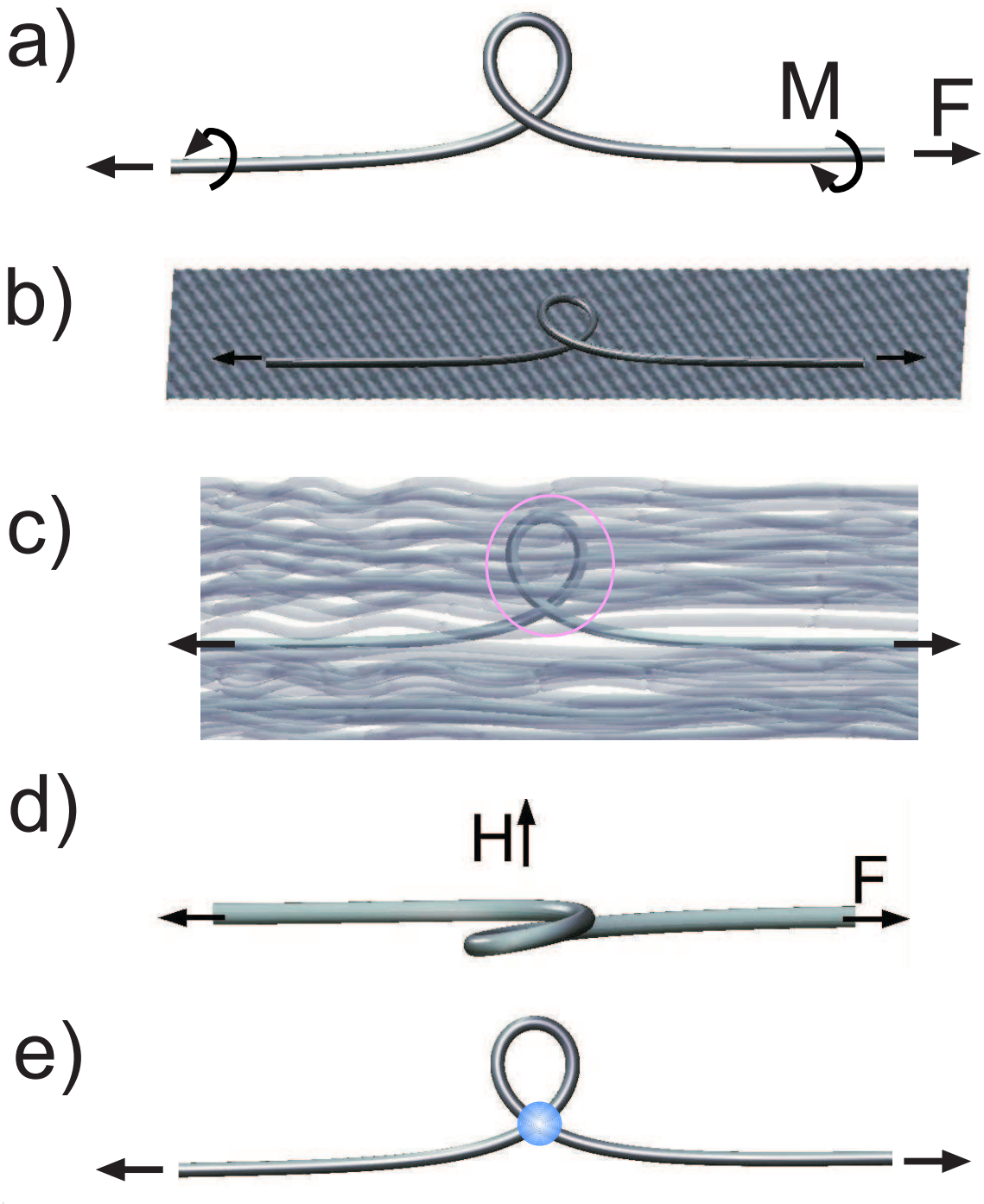


Figure 2.6: Various examples of stable loops in DNA under tension: a) Applied torque M at the ends. b) DNA adsorbed on a surface. c) DNA surrounded by a dense solution of infinitely long DNAs. Unfolding of the loop goes hand in hand with an energetically costly transient "cavity" creation (in the pink region) d) DNA in a strong magnetic field H perpendicular to the applied force e) DNA looped by a freely sliding linker ligand ("weakly condensed" DNA).

ing experiments: why is the measured persistence length of AFM stretched polymers often (much) smaller than that obtained by other methods. This "ghost" persistence length effect is a very general phenomenon that lies at the heart of stretching elastic substructures (like loops and deflections due to boundary anchoring) in wormlike chains.

2.2.1 The Planar Homoclinic Loop

Let us start with the simplest case one can imagine: a looped DNA under tension F along the z-axis. The DNA will at first be allowed to fluctuate only in plane (say a DNA on the flat surface of a membrane as in Fig. 2.6b). We neglect the DNA twist degree of freedom which if not explicitly constrained immediately decouples from DNA bending energy [9]. The latter writes in the general 3D case (cf. previous chapter)

$$\begin{aligned} E &= \int_{-L/2}^{L/2} \left(\frac{A}{2} \left(\frac{d\underline{t}}{ds} \right)^2 - F t_z \right) ds \\ &= \int_{-L/2}^{L/2} \left(\frac{A}{2} \left(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2 \right) - F\cos\theta \right) ds \end{aligned} \tag{2.22}$$

Here t_z is the z-projection of the DNA tangent and the dots again represent the differentiation with respect to the arc-length parameter s. The two variables θ, ϕ are the spherical coordinates of the tangent vector $\underline{t}$ (with symmetry axis z). In the spirit of the Kirchhoff kinetic analogy from the previous chapter the bending energy in Eq. 2.22 corresponds to the Lagrangian of a spherical pendulum in the gravitational field. The Euler-Lagrange equations of Eq. 2.22 become

$$A\frac{d}{ds}\left(\dot{\phi}\sin^2\theta \right) = 0 \tag{2.23}$$

$$A\left(\ddot{\theta} - \dot{\phi}^2 \sin\theta\cos\theta \right) - F\sin\theta = 0 \tag{2.24}$$

By virtue of the fact that ϕ becomes a cyclic variable in Eq. 2.22 the equation Eq. 2.23 is integrable and one obtains

$$\dot{\phi}\sin^2\theta = M_z/A$$

with the integration constant M_z being the overall conserved bending moment around the z-axis[10] (cf. [6]). For the planar case we have $\dot{\phi} = 0$ and consequently Eq. 2.24 simplifies to

$$\ddot{\theta} = \lambda^{-2}\sin\theta \text{ with } \lambda = \sqrt{A/F} \tag{2.25}$$

[9] The latter assumption is of course not justified in the pretty involved case depicted Fig. 2.6a) where one introduces explicit torque.

[10] The system is indeed fully integrable as it possesses a second integral which is the total stress + strain energy (or the Hamiltonian corresponding to the Lagrangian in Eq. 2.22) $\frac{A}{2}\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2 \right) + F\cos\theta = const.$

Here once more we meet the tension length λ introduced in the last chapter as the relevant lengthscale in our problem. The latter equation is the time independent Sine-Gordon equation well known and studied in many systems especially in the context of solitons (and their applications like Josephson junctions, cf. Davydov's book [29]). Therefore it is appropriate to call the solutions of Eq. 2.25 solitons or kinks. Now the equation Eq. 2.25 can be integrated twice to obtain

$$\begin{aligned} \frac{d\theta}{ds} &= \lambda^{-1}\sqrt{2\left(c-\cos\theta\right)} \\ \left(s-s_0\right)/\lambda &= \int_{\theta(s_0)}^{\theta(s)} \frac{d\theta'}{\sqrt{2\left(c-\cos\theta'\right)}} \end{aligned} \tag{2.26}$$

with a trivial integration constant s_0 reflecting the reparametrization invariance of our system and a less trivial constant c (related to the total stress+ strain energy cf. footnote on previous page). The general solution of Eq. 2.25 with arbitrary c leads to elliptic functions but in our case the solution is even simpler as the DNA curvature $d\theta/ds$ is assumed to vanish asymptotically for $s/\lambda \to \pm\infty$. This implies $c = 1$ and the solution reads[11] $\theta_{loop} = 4\arctan e^{s/\lambda}$ or

$$\cos\theta_{loop}\left(s\right) = 1 - \frac{2}{\cosh^2\left(s/\lambda\right)} \tag{2.27}$$

cf. also Fig. 2.7.
To obtain the force-extension behavior of the homoclinic loop / kink in 2-D we need to evaluate the contribution of the rod elasticity as well as the fluctuation contribution to the partition function Q_{loop}. The latter writes

$$Q_{loop} = \int_{\theta\in K} \mathcal{D}\left[\theta\right] e^{-\beta E[\theta]} \tag{2.28}$$

with

$$E\left[\theta\right] = \int_{-L/2}^{L/2} \left(\frac{1}{2}A\dot{\theta}^2 - F\cos\theta\right) ds \tag{2.29}$$

Here the path integral spans over some (functional) neighborhood K of the kink solution θ_{loop}. For large enough tensions only the quadratic fluctuations will contribute to Q_{loop} so that we can expand[12] $E\left[\theta\right]$ up to quadratic order around θ_{loop}. The linear term δE in this expansion vanishes because θ_{loop} is an extremal point of E. We have

$$E\left[\theta_{loop}+\delta\theta\right] = E_{loop} + E_{fluct}\left[\delta\theta\right]$$

with the "classical" $(T = 0)$ bending energy of the kink

$$\begin{aligned} E_{loop} &= E\left[\theta_{loop}\right] = F\int_{-L/2}^{L/2}\left(\frac{4}{\cosh^2\frac{s}{\lambda}} - 1\right) ds \\ &= -FL + 8\sqrt{AF} + O\left(e^{-L/\lambda}\right) \end{aligned} \tag{2.30}$$

[11] Here the length of the molecule L is assumed to be very large compared to the tension length λ, i.e., $L/\lambda \to \pm\infty$. The solution θ_{kink} provided here is only valid in this asymptotic limit. For finite L/λ case there are exponentially small corrections $O\left(e^{-L/\lambda}\right)$ treated in the next section.

[12] This common approximation is usually called the "saddle point approximation", cf Appendix A.

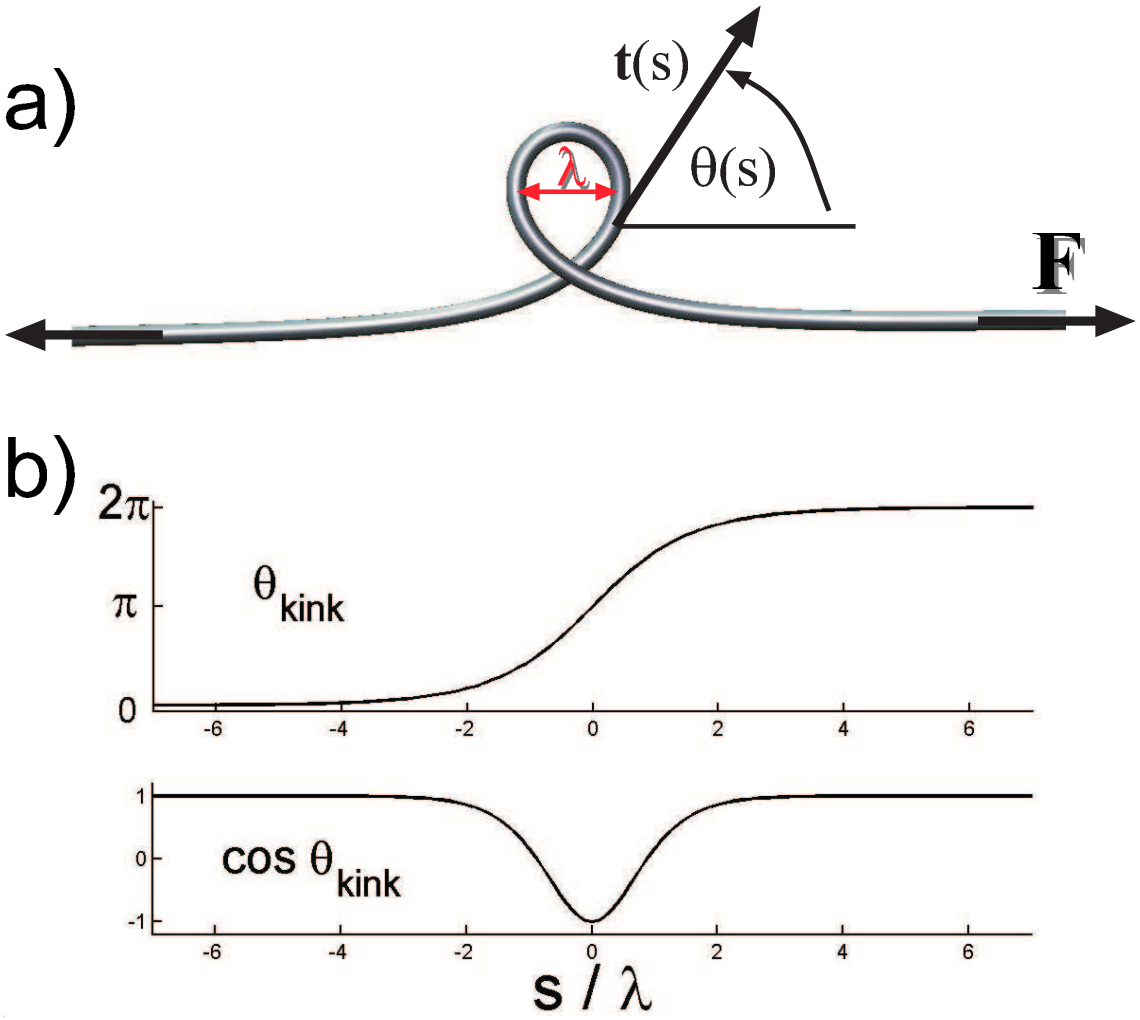


Figure 2.7: a) The definition of the Euler angle θ and the scale of the loop. The loop head diameter (red) is approximately given by $\lambda = \sqrt{A/F}$ b) The loop solution $\theta_{kink}(s)$ as given by Eq. 2.27.

and the quadratic fluctuation contribution

$$\begin{aligned} E_{fluct}\left[\delta\theta\right] &= \int_{-L/2}^{L/2}\left(\frac{A}{2}\delta\dot{\theta}^2+\frac{F}{2}\cos\left(\theta_{loop}\right)\delta\theta^2\right)ds \\ &= \frac{A}{2}\int_{-L/2}^{L/2}\left[\delta\dot{\theta}^2+\lambda^{-2}\left(1-\frac{2}{\cosh^2\left(s/\lambda\right)}\right)\delta\theta^2\right]ds \end{aligned}$$

After performing partial integration and introducing the dimensionless parameter

$$t=\frac{s}{\lambda}$$

the latter can be recast into

$$\beta E_{fluct}\left[\delta\theta\right]=\frac{1}{2}\int_{-L/2\lambda}^{L/2\lambda}\delta\theta\left(t\right)\hat{\mathbf{T}}\left(t\right)\delta\theta\left(t\right)dt$$

with the (position dependent) Schrödinger-like **fluctuation operator**

$$\hat{\mathbf{T}}=\beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2}+\left(1-\frac{2}{\cosh^2\left(t\right)}\right)\right) \tag{2.31}$$

The partition function Eq. 2.28 can now be written as a quadratic path integral over fluctuations $\delta\theta$

$$Q_{loop} = e^{-\beta E_{loop}}Q_{loop}^{fluct} \tag{2.32}$$

$$Q_{loop}^{fluct} = \int_{(0,-\frac{L}{2\lambda})}^{(0,\frac{L}{2\lambda})}\mathcal{D}\left[\delta\theta\right]e^{-\frac{1}{2}\int_{-L/2\lambda}^{L/2\lambda}\delta\theta\hat{\mathbf{T}}\delta\theta dt} \tag{2.33}$$

At first glance the evaluation of this path integral appears very tricky because of the explicit s (or t) dependence of the fluctuation operator $\hat{\mathbf{T}}$. Fortunately the "potential energy part" $V\left(t\right)\propto\left(1-2/\cosh^2\left(t\right)\right)$ is just simple enough[13] to allow an exact diagonalization of $\hat{\mathbf{T}}$. The spectrum of $\hat{\mathbf{T}}$ consists of a discrete spectrum with the only discrete eigenvalue $\mu_0=0$ and the continuum spectrum $\mu_k=\beta\sqrt{AF}\left(k^2+1\right)$, $k>0$ (cf. [13, 14, 30]). The existence of a vanishing eigenvalue (Goldstone mode) which is a consequence of translational invariance $t\rightarrow t+t_0$ of the kink solution formally causes a divergence of Eq. 2.33 [13, 14]. Fortunately it can be shown that this problem which is a consequence of infinite DNA limit ($L/\lambda=\infty$) can be cured by taking the limiting process $L/\lambda\rightarrow\infty$ properly into account. This is done in the next section where we perform the full derivation of the fluctuation partition function from Eq. 2.33. The result is

$$Q_{loop}=\frac{\sqrt{2}}{\pi}\beta LFe^{-\frac{L}{2}\sqrt{\frac{F}{A}}}e^{-\beta\left(8\sqrt{FA}-LF\right)}$$

[13]Such operators appear in the (time independent) 1D Schrödinger equation with the Morse-Rosen potential $V\left(x\right)=2/\cosh^2 x$. The latter happens to be one of the few exactly soluble non-trivial cases [30] where the spectrum and the eigenfunctions are given analytically. It is a member of the supersymmetric potential family generally admitting analytic solutions.

and the corresponding free energy writes

$$G_{loop} = \frac{k_B T}{2} L \sqrt{\frac{F}{A}} + \left(8\sqrt{AF} - LF\right) + k_B T \ln\left(\pi^{-3/4}\beta LF\right) \tag{2.34}$$

Comparing the latter free energy expression with the free energy G_0 of the stretched straight DNA (without a loop, cf. last chapter) we get

$$\begin{aligned} \Delta G_{loop-0} &= G_{loop} - G_0 \\ &= 8\sqrt{AF} + O\left(k_B T \ln\left(\frac{LF}{k_B T}\right)\right) \end{aligned} \tag{2.35}$$

From that we see that the free energy difference ΔG_{loop-0} is dominated by the elastic energy part $8\sqrt{AF}$ which is the second term in E_{loop}. (The first one $-FL$ is already present in the straight DNA case and cancels in the difference.) Besides that typically very large term there is a mere logarithmic correction term. The observation of weak coupling of this entropic part with the (circular) shape was already made by Odijk for circular rings [31], so we are not much surprised by the logarithmic form of this contribution[14]. The force-extension curve of a 2-D loop is then given from the mean end-to-end distance $\langle \Delta z \rangle = -\partial G / \partial F$:

$$\frac{\langle \Delta z \rangle}{L} = 1 - \left(\frac{1}{4} + 4\frac{l_P}{L}\right)\sqrt{\frac{k_B T}{F l_P}} + O\left(\frac{k_B T}{FL}\right) \tag{2.36}$$

We may drop the $O\left(k_B T/\left(FL\right)\right)$ term which is for all practical purposes negligible. We can now compare the "equation of state" of the looped DNA Eq. 2.36 with the previously obtained one for the straight configuration given by $\Delta z_0 = -\partial G_0 / \partial F$ that is

$$\frac{\langle \Delta z_0 \rangle}{L} = 1 - \frac{1}{4}\sqrt{\frac{k_B T}{F l_P}} + O\left(\frac{k_B T}{FL}\right) \tag{2.37}$$

Comparing Eq. 2.36 and Eq. 2.37 we see that they both have a leading term $F^{-1/2}$ only the prefactor gets renormalized in Eq. 2.36 by a contribution stemming from the elastic part of the loop free energy! This implies a fairly simple prediction which is useful for experimental interpretation. Suppose one performs a single molecule stretching experiment with a DNA that contains a loop. Imagine the experimentalist is "naive" and does not know anything about the presence of the loop. He/she merely fits the data by the usual straight DNA expression Eq. 2.37 (as done for instance in ref. [131]) and is of course happy that it works well (at least up to the leading term $F^{-1/2}$). From that fit he/she recovers correctly the total length of the DNA (from the asymptotic line on the Δz axis) but he/she observes that something strange happened to the "persistence length"- it is smaller then expected!

[14] One might be slightly disappointed as one could have also (wrongly) guessed stronger deviations in entropy between the two states. A naive argument that the entropy should strongly "feel" the size (and shape) of the loop around which it acts indeed fails. The reason for this is that the amplitude of the undulations goes down with increasing forces in roughly the same manner as the loop squeezes.

The explanation is simple: the **apparent persistence length** becomes $l_P^{app} = l_P/\left(1+16l_P/L\right)^2$ instead of the real persistence length l_P. For the case of N_{loop} contained loops the latter easily generalizes (as we see in next sections) to

$$l_P^{app} = \frac{l_P}{\left(1+16N_{loop}\frac{l_P}{L}\right)^2} \text{ in 2D} \tag{2.38}$$

$$l_P^{app} = \frac{l_P}{\left(1+8N_{loop}\frac{l_P}{L}\right)^2} \text{ in 3D} \tag{2.39}$$

So we see that with a growing number of loops the apparent persistence length rapidly goes down, i.e., the effective stretching resistance shoots up (as it should).

2.2.2 The Partition Function of a Planar DNA Loop: The Formal Derivation

In the previous section we have stated basic results like the partition function, free energy and force-extension relation of a looped DNA in 2-D without proof. Here we provide the full derivation. Although quite technical at some points it will reveal the detailed behavior of the 2-D loop under tension. In the (semi-classical) limit of large tensions i.e. $\beta\sqrt{AF} \gg 1$ it will provide us with exact expressions even for "big" loops which are comparable to the DNA length ($L/\lambda \sim 1$).
A standard method for computing path integrals like Eq. 2.33 with 2.31 was developed in the context of quantum mechanical tunnelling problems, and nucleation of bubbles in overheated liquids by Langer [15] (cf. also refs. [14] and [13]). In its basic formulation it requires the knowledge of the whole spectrum of the fluctuation operator $\hat{\mathbf{T}}$. We will essentially follow (at least in principle) these standard methods but at one point we will take the shortcut invented by Gelfand and Yaglom [17] to avoid the technicalities of dealing explicitly with the full spectrum of $\hat{\mathbf{T}}$.
Instead of imposing the "asymptotically vanishing" boundary conditions $d\theta\left(s\right)/ds \to 0$ for $s/\lambda \to \pm\infty$ (which is an approximation valid for $L/\lambda \gg 1$) as we did above we do it more correctly here on the finite interval $[-L/2, L/2]$. Of course the particular form of boundary conditions will not be crucial for the underlying physics in the limit $L/\lambda \to \infty$. As a matter of convenience we make the simplest choice for boundary conditions: $\theta\left(-L/2\lambda\right) = 0$ and $\theta\left(+L/2\lambda\right) = 2\pi$, i.e. we clamp the ends of the DNA in an orientation parallel to the force direction.
In this case the Euler-Lagrange equation Eq. 2.25 gives Eq. 2.26 with $s_0 = 0$ and $\theta\left(0\right) = \pi$. In terms of the dimensionless variable $t = s/\lambda$ the solution reads

$$\cos\theta_{loop}\left(t\right) = 2\mathrm{sn}^2\left(\frac{t}{\sqrt{m}}|m\right) - 1 \tag{2.40}$$

$$\theta_{loop}\left(s\right) = \pi + 2\mathrm{am}\left(\frac{t}{\sqrt{m}}|m\right) \tag{2.41}$$

with sn and am being the Jacobi elliptic function with parameter m. The latter results from the clamped boundary conditions and is implicitly given by

$$\sqrt{m}K\left(m\right) = \frac{L}{2\lambda} = \frac{L}{2}\sqrt{F/A} \tag{2.42}$$

In the same manner as above in Eq. 2.31 we can introduce the fluctuation operator

$$\hat{\mathbf{T}} = \beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2} + 2\mathrm{sn}^2\left(\frac{t}{\sqrt{m}}|m\right) - 1\right) \tag{2.43}$$

Note that in the limiting case $L/\lambda \to \infty$ we have $m \to 1$ and the operator 2.43 coincides with Eq.2.31, the expression Eq. 2.40 reduces to Eq. 2.27 - the asymptotic loop solution we derived above. In the Kirchhoff analogy the solution Eq. 2.40 describes a revolving pendulum which makes one full turn (from $\theta = 0$ to 2π) during the finite "time period" L/λ.
In the following we also need the derivative

$$\theta'(t) = \frac{d}{dt}\theta(t) = \lambda\frac{d}{ds}\theta(s) = \lambda\dot{\theta}(s)$$

with respect to t which writes for the loop solution

$$\theta'_{loop}(t) = \frac{2}{\sqrt{m}}\mathrm{dn}\left(\frac{t}{\sqrt{m}}|m\right) \tag{2.44}$$

We want now to evaluate the partition function around the loop solution θ_{loop}. We perform this computation by using standard path-integral methods developed in the context of the "saddle point approximation". These basic methods are reviewed in the Appendix A. Using these results the loop partition function

$$Q_{loop} = \int_{\left(0,-\frac{L}{2\lambda}\right)}^{\left(2\pi,\frac{L}{2\lambda}\right)} \mathcal{D}[\theta]\, e^{-\beta E[\theta]}$$

can be written as

$$Q_{loop} = \underbrace{e^{-\beta E_{loop}[\theta_{loop}]}}_{\text{classical contribution}} \quad \underbrace{\sqrt{\frac{\beta\sqrt{AF}}{2\pi D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)}}}_{\text{fluctuations around class. solution}} \tag{2.45}$$

The first factor here is the energetic contribution of the "classical" solution θ_{loop} and the second term $\sqrt{...}$ is the entropic contribution of **quadratic** fluctuations around θ_{loop} (cf. Eq.2.33). The expression $D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)$ is defined by the ratio of eigenvalues of the fluctuation operator $\hat{\mathbf{T}}$ and the corresponding free particle (cf. Appendix A Eq. 2.105 for details)

$$\frac{\lambda}{L}D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right) = \frac{\det\left(\hat{\mathbf{T}}/\left(\beta\sqrt{AF}\right)\right)}{\det\left(-\frac{\partial^2}{\partial t^2}\right)} = \prod_{k=0}^{\infty}\frac{\mu_k/\left(\beta\sqrt{AF}\right)}{\pi^2(k+1)^2\lambda^2/L^2} \tag{2.46}$$

$D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)$ can be computed very elegantly via the celebrated method of Gelfand and Yaglom [17] which consists of solving an initial value problem on the interval $[-L/2\lambda, L/2\lambda]$ (Appendix A Eqs. 2.107,2.106). Remarkably the explicit solution for

$D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)$ can be stated in terms of the "classical" solution $\theta_{loop}(s)$ (cf. Eq. 2.108). By virtue of Eq. 2.44 the latter writes

$$\begin{aligned} D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right) &= \theta'_{loop}\left(\frac{L}{2\lambda}\right)\theta'_{loop}\left(-\frac{L}{2\lambda}\right)\int_{-L/2\lambda}^{L/2\lambda}\frac{dt}{\left(\theta'_{loop}(t)\right)^2} \\ &= \frac{4}{m}\underbrace{\mathrm{dn}^2\left(K(m)|m\right)}_{=(1-m)}\frac{m^{3/2}}{4}\underbrace{\int_{-K(m)}^{K(m)}\frac{d\tau}{\mathrm{dn}^2(\tau|m)}}_{=2E(m)/(1-m)} \\ D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right) &= 2\sqrt{m}E(m) \end{aligned} \tag{2.47}$$

In the second line we have substituted $\tau = t/\sqrt{m}$ and used standard properties of the dn function (cf. [85]). $E(m)$ is the complete elliptic integral of the second kind. The loop energy E_{loop} (not to be confused with elliptic integral $E(m)$) writes

$$\begin{aligned} E_{loop}[\theta_{cl}] &= \int_{-L/2}^{L/2}\left(\frac{1}{2}A\dot{\theta}^2_{loop}(s) - F\cos\theta_{loop}(s)\right)ds \\ &= F\int_{-L/2}^{L/2}\left(\frac{2}{m}\mathrm{dn}^2\left(\frac{s}{\sqrt{m}\lambda}|m\right) - 2\mathrm{sn}^2\left(\frac{s}{\sqrt{m}\lambda}|m\right)+1\right)ds \\ &= F\frac{\lambda}{\sqrt{m}}\left[(m-2)\underbrace{\int_{-K(m)}^{K(m)}dt}_{2K(m)} + 4\underbrace{\int_{-K(m)}^{K(m)}\mathrm{dn}^2(\tau|m)\,d\tau}_{2E(m)}\right] \\ E_{loop}[\theta_{cl}] &= 2\frac{\sqrt{AF}}{\sqrt{m}}\left[K(m)(m-2)+4E(m)\right] \end{aligned} \tag{2.48}$$

In the third line we exploited the relation $m\mathrm{sn}^2 + \mathrm{dn}^2 = 1$.
Inserting Eqs. 2.47 and 2.48 in Eq. 2.45 we arrive at the elegant but deceptive expression

$$Q^{wrong!}_{loop} = \sqrt{\frac{\beta\sqrt{AF}}{4\pi\sqrt{m}E(m)}}e^{-2\frac{\beta\sqrt{AF}}{\sqrt{m}}[K(m)(m-2)+4E(m)]}\ \text{(wrong!)} \tag{2.49}$$

The upper expression for $Q^{wrong!}_{loop}$ "looks too good to be true" in all limits. Taking for instance the limit $L/\lambda \to \infty$ in Eq. 2.42 we have $m \to 1$, $E(m) \to 1$, $K(m) \to L/2\lambda$. The exponent tends to $-\beta\left(8\sqrt{AF} - FL\right)$ i.e. we indeed recover the loop energy Eq. 2.30 in this asymptotic case. But in the fluctuation prefactor we miss (at least) another factor $\sim e^{const.\times L\sqrt{F/A}}$ that would account for the fluctuations around the straight configuration[15].

[15]The latter we expect to be present for physical reasons as the largest part of a very long DNA is in a roughly straight configuration despite the loop in the middle position. Note that the loop is spacially confined to a region of length $\sim \lambda \ll L$ (in the long DNA limit).

The resolution of this discrepancy can cause some headache but it is indeed physically simple to resolve[16]. The main problem is that the expression Eq. 2.46 for $D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)$ implicitly assumes that **all** the eigenmodes of the operator $\hat{\mathbf{T}}$ contribute in a quadratic Gaussian manner i.e. that $\left(\frac{\lambda}{L}D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)\right)^{-1/2}$ can be thought as a ratio of infinitely many **Gaussian** integrals:

$$\frac{1}{\sqrt{\frac{\lambda}{L}D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)}}=\lim_{N\to\infty}\prod_{k=0}^{N}\frac{\int_{-\infty}^{\infty}e^{-\frac{1}{2}\mu_k a_k^2}da_k}{\int_{-\infty}^{\infty}e^{-\frac{1}{2}\frac{\lambda^2\pi^2(k+1)^2}{L^2}a_k^2}da_k} \tag{2.50}$$

After performing the Gaussian integrals this coincides with Eq. 2.46. The boundaries in the integrals above are taken to be $\pm\infty$ for convenience. But it is exactly this "convenience" that causes trouble in our case. Especially in the case when one of the eigenvalues, say μ_0, approaches zero, the expression Eq. 2.50 makes a serious flaw as the potential acting on that mode becomes so "flat" that the corresponding Gaussian integral would become unbounded. But thinking physically: what really matters in this limiting case is rather **the finiteness of the (state) space** that the almost-zero-eigenvalue mode can populate (rather than the vanishing potential acting on it). For the positional translational shifting mode of the loop along the DNA (mentioned in the previous section) which we expect to occur here in the asymptotic limit $L/\lambda\to\infty$ the **size of the state space becomes the (dimensionless) DNA contour length** L/λ. Therefore in the limit $\mu_0\left(L/\lambda\right)^2\ll 1$ one should correct Eq. 2.50 and rather write

$$\frac{1}{\sqrt{\frac{\lambda}{L}D_{corr}\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)}}=\underbrace{\left(\int_{-L/2\lambda}^{L/2\lambda}e^{-\frac{1}{2}\mu_0 a_0^2}da_0\right)}_{=\sqrt{\frac{2\pi}{\mu_0}}\mathrm{erf}\left(\frac{\sqrt{\mu_0}L}{2\sqrt{2}\lambda}\right)}\lim_{N\to\infty}\frac{\prod_{k=1}^{N}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\mu_k a_k^2}da_k}{\prod_{k=0}^{N}\int_{-\infty}^{\infty}e^{-\frac{\lambda^2\pi^2(k+1)^2}{L^2}a_k^2}da_k}$$

Here erf (...) is the Gaussian error function. Rewriting in terms of the (wrong) expression $D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)$ we obtain

$$\begin{aligned}D_{corr}\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right) &= \left(\frac{\int_{-\infty}^{\infty}e^{-\frac{1}{2}\mu_0 a_0^2}da_0}{\int_{-L/2\lambda}^{L/2\lambda}e^{-\frac{1}{2}\mu_0 a_0^2}da_0}\right)^2 D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)\\ &= \frac{1}{\mathrm{erf}^2\left(\frac{\sqrt{\mu_0}L}{2\sqrt{2}\lambda}\right)}D\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)\end{aligned} \tag{2.51}$$

In the two limiting cases the latter writes

$$D_{corr}\left(\frac{L}{2\lambda},-\frac{L}{2\lambda}\right)\approx\left\{\begin{array}{c}\frac{2\pi\lambda^2}{\mu_0 L^2}\ \text{for}\ \sqrt{\mu_0}\frac{L}{\lambda}\ll 1\\ 1-\frac{4\sqrt{2}\lambda}{\sqrt{\pi\mu_0}L}e^{-\frac{\mu_0 L^2}{8\lambda^2}}\ \text{for}\ \sqrt{\mu_0}\frac{L}{\lambda}\gtrsim 1\end{array}\right.\times D\left(\frac{L}{2},-\frac{L}{2}\right) \tag{2.52}$$

In the last line we used the Taylor and the asymptotic expansion of erf around 0 and ∞ respectively. From that we see that the naive expression Eq. 2.49 is valid only for large

[16]Similar problems occur in the context of the semiclassical treatment of QM tunneling (cf. [13, 14]) and result from time invariance of the tunneling event.

enough $\sqrt{\mu_0}L/\lambda$. In order to compute the corrected fluctuation factor D_{corr} explicitly we still need the lowest eigenvalue μ_0 of $\hat{\mathbf{T}}$ (Eq. 2.43). In the Appendix B we show that

$$\mu_0 = \beta\sqrt{AF}\left(\frac{1}{m}-1\right) \tag{2.53}$$

$$= \frac{2l_P}{L}\frac{K(m)(1-m)}{\sqrt{m}} \tag{2.54}$$

In the last line we replaced F via the constraint Eq. 2.42 which provides a unique mapping $m \longleftrightarrow F$. Now putting Eqs. 2.45, 2.47, 2.51 and 2.54 we obtain finally the correct partition function[17]

$$Q_{loop} = \sqrt{\frac{\beta\sqrt{AF}}{4\pi\sqrt{m}E(m)}}\operatorname{erf}\left(\frac{\sqrt{\mu_0}L}{2\sqrt{2}\lambda}\right)e^{-2\frac{\beta\sqrt{AF}}{\sqrt{m}}[K(m)(m-2)+4E(m)]} \tag{2.55}$$

Using the relations 2.54 and 2.42 we can rewrite this general expression fully in terms of m

$$Q_{loop}(m) = \left(\tfrac{l_P}{2\pi L}\tfrac{K(m)}{E(m)}\right)^{\frac{1}{2}}\operatorname{erf}\left(\left(\frac{l_P}{L}\right)^{\frac{1}{2}}K^{\frac{3}{2}}(m)(1-m)^{\frac{1}{2}}m^{\frac{1}{4}}\right)e^{-4\frac{l_P}{L}K(m)[K(m)(m-2)+4E(m)]}$$

Based on this form a treatment of the exact force extension relation for a 2-D loop for arbitrary lengths and shapes is possible[18].
Here we dispense with that and we focus on the most relevant limiting case $\frac{L}{2\lambda} \gg 1$ (i.e. the contour length L of the DNA is much larger than the loop-head size λ). In this case we have $m \to 1$ and for the two elliptic integrals the following asymptotic expansions hold:

$$K(m) = \ln 4 - \frac{1}{2}\ln(1-m) + O((1-m)\ln(1-m))$$
$$E(m) = 1 + O((1-m)\ln(1-m))$$

The involved expression $O((1-m)\ln(1-m))$ which goes sublinearly to zero as m approaches unity can be neglected. The constraint Eq. 2.42 can be solved for $m \approx 1$ and we obtain

$$m \approx 1 - 16e^{-\frac{L}{\lambda}}$$

[17]It is somehow ironic that the error we have tapped into (Eq. 2.49) at first is corrected by an *error* function. It is also remarkable and a strange coincidence (?) that the initially wrong (ghost) entropy of the μ_0 -mode in Eq. 2.49 has exactly "eaten up" the whole entropy of the wormlike chain behavior of the largest part of the chain! Identical problems with the (almost) zero mode appear in many other applications (cf. the Appendix E for references) but there their visual meaning is harder to grasp than in our example here. We have also seen that the explicit exclusion of the zero eigenvalue by proper boundary conditions does not save us from renormalizing the contribution of the μ_0 mode "by hand"!

[18]Taking Eq. 2.42 and differentiating both sides by m we obtain $\frac{dm(F)}{dF} = 1/\left(\frac{dF(m)}{dm}\right)$. Thus the derivative required for computing $\langle z\rangle$ can be performed, i.e. $\frac{d\ln Q}{dF} = \frac{dm(F)}{dF}\frac{d\ln Q}{dm}$. $\langle z\rangle = \langle z\rangle(m)$ can be written in terms of m and a parametric $\langle z\rangle(m) - F(m)$ curve can be obtained.

For the smallest eigenvalue μ_0 from Eq. 2.54 we obtain

$$\mu_0 = \frac{16 l_P}{\lambda} e^{-\frac{L}{\lambda}} \tag{2.56}$$

So indeed for $L/\lambda \to \infty$ this eigenvalue becomes asymptotically zero and the corresponding mode (the translational shifting mode) becomes a Goldstone mode as we would intuitively expect from the approximate translational invariance of the loop energy (the loop can be found at any position along the DNA as the boundary conditions do not play a role in this limit). We obtain

$$\begin{aligned} Q_{loop} &= \frac{\sqrt{2} L l_P}{\pi \lambda^2} e^{-\frac{L}{2\lambda}} e^{\frac{l_P}{\lambda}\left(\frac{L}{\lambda}-8\right)} \\ &= \frac{\sqrt{2}}{\pi} \beta L F e^{-\frac{L}{2}\sqrt{\frac{F}{A}}} e^{-\beta\left(8\sqrt{FA}-LF\right)} \end{aligned} \tag{2.57}$$

The free energy as function of F reads

$$G_{loop} = \frac{L}{2\beta}\sqrt{\frac{F}{A}} + \left(8\sqrt{AF} - LF\right) - \frac{1}{\beta}\ln\left(\frac{\sqrt{2}}{\pi}\beta L F\right) \tag{2.58}$$

Implying the force extension

$$\begin{aligned} \frac{\langle \Delta z \rangle}{L} &= -\frac{1}{L}\frac{dG}{dF} \\ &= 1 - \frac{1}{4\beta\sqrt{AF}} - \frac{4}{L}\sqrt{\frac{A}{F}} + O\left(\frac{1}{\beta F L}\right) \end{aligned}$$

This is exactly Eq. 2.36 presented in the previous section.
After this excursion (or rather "tour de force") in elliptic functions / integrals in the following sections we will merely focus on the asymptotic limit $L/\lambda \gg 1$ which is most relevant and which will make the physics behind most transparent.

2.2.3 The Homoclinic Loop in 3-D

After having understood the behavior of the homoclinic loop (the kink) in 2-D it seems that a generalization to the third dimension should be straightforward. But as we will see there are several traps and some interesting physics on the way. The first and main problem that one encounters is that the homoclinic loop is (unlike in the 2-D case) elastically unstable! The simplest way to see or better to feel this is to take an elastic cable make a loop in it and (without torsionally constraining the ends) to pull on it. Obviously if we force the loop to stay in a plane (its own weight can do it if the cable is lying on a table provided that we do not pull to strongly) it cannot escape its fate and it stays what it is: a **topological excitation** which cannot leave the rod (besides at one of its two ends)! But in contrast to this situation when redoing the same experiment in 3-space we feel slightly disappointed observing that the loop immediately unfolds

at any force[19]. So if there is any interesting physics of 3-D homoclinic loops it will have to come through constraints or loop stabilizing potentials. Before we start this program let us (as warm up) check if the 3-D loop is indeed mechanically unstable. Therefore we have to consider the 3-D elastic energy, Eq. 2.4, from the last chapter[20]. If we dispense with applying torques or torsionally constraining the DNA we can set $C = 0$ formally (as explained in the previous section) and forget about the angle ψ. Then we obtain the energy

$$E\left[\theta,\phi\right] = \int_{-L/2}^{L/2} \frac{A}{2}\left(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2\right) - F\cos\theta ds \tag{2.59}$$

By solving the corresponding Euler-Lagrange equations we see that θ_{loop} from Eq. 2.27 and $\phi_{loop} = const.$ constitute all possible (single) kink solutions with approximate boundary conditions of vanishing curvatures at infinity ($L/\lambda \to \infty$). To check the stability of the loop we need to expand Eq. 2.59 up to second order in terms of small perturbations $\delta\theta$ and $\delta\phi$ around the loop state $(\theta_{loop}, \phi_{loop})$. As the first variation $\delta E = 0$ vanishes at $(\theta_{loop}, \phi_{loop})$ we obtain

$$E\left[\theta_{loop} + \delta\theta, \phi_{loop} + \delta\phi\right] = E_{loop} + \frac{1}{2}\delta^2 E \tag{2.60}$$

$$\delta^2 E = \int_{-L/2}^{L/2} \left[\left(A\delta\dot{\theta}^2 + F\cos\left(\theta_{loop}\right)\delta\theta^2\right) + A\sin^2\left(\theta_{loop}\right)\delta\dot{\phi}^2\right] ds \tag{2.61}$$

Looking closely at the expansion above $\delta^2 E$ we see that it consists of two contributions. The first is a functional of $\delta\theta$ and $\delta\dot{\theta}$ only and appeared already in the planar problem before. The second term depends only on $\delta\dot{\phi}$ but not on $\delta\phi$. It is **positive definite**! We are surprised, what went wrong? Did we misinterpret our experiment with the telephone cable, is it a deception of senses? Is the 3-D kink indeed stable? The answer is **no**[21]! The subtlety lies in the singularity of the (θ, ϕ) coordinate system: around $\theta = 0$ the angle ϕ becomes arbitrary. This leads to peculiar effects. For instance the energy stays invariant under a transformation $J : \phi(s) \to \phi(s) + const.H(s - s_0)$ where H is the Heaviside step function ($H(s) = 0$ for $s \leq 0$ and $H(s) = 1$ otherwise) and s_0 is a point where $\sin\theta(s)$ crosses the zero line. This interesting invariance under the transformation J (call it the "jump-gauge") says that admissible perturbations $\delta\phi$ need not even to be continuous ($\delta\dot{\phi}$ can contain delta peaks) at points where $\sin\theta = 0$. This indicates that the naive expansion like done in Eq. 2.60 might be inappropriate. Instead of going into mathematical subtleties of this peculiar coordinate system we decide to "throw it overboard" and chose a new one by rotating it.
Instead of explicitly rotating the coordinate system it is equivalent but more convenient to rotate the force direction F and put it along the x-axis so that the potential energy

[19]Here for simplicity we neglect the possibility of applying torques at the ends. If the ratio of the torque and tension is large enough the loop can indeed be mechanically stabilized as one can again see by playing with the rod.

[20]For simplicity we go to the limit $L/\lambda \gg 1$ where the rather sophisticated elliptic functions from previous section degenerate to hyperbolic functions.

[21]At this point some smart researchers (not to be named here) in literature stepped into the trap answering this question with yes!

part writes now $-F\cos\phi\sin\theta$. Further we introduce for convenience a new angle $\vartheta = \theta - \pi/2$ instead of θ, i.e., we measure the angle ϑ with respect to the equatorial plane now (as on a globe). The energy $E[\vartheta,\phi]$ then writes

$$E[\vartheta,\phi] = \int_{-L/2}^{L/2} \frac{A}{2}\left(\dot{\phi}^2\cos^2\vartheta + \dot{\vartheta}^2\right) - F\cos\phi\cos\vartheta ds \tag{2.62}$$

Though the elastic energy looks slightly different from before the physics governed by it does not change. The Euler-Lagrange equations in this system write

$$\begin{aligned} \ddot{\vartheta} &= \frac{F}{A}\cos\phi\sin\vartheta - \dot{\phi}^2\cos\vartheta\sin\vartheta \\ \ddot{\phi}\cos^2\vartheta - 2\dot{\phi}\dot{\vartheta}\cos\vartheta\sin\vartheta &= \frac{F}{A}\sin\phi\cos\vartheta \end{aligned}$$

Here we are merely interested in the $\vartheta = 0$ solution, i.e., we put the loop into the x-y-plane. This imposes no restriction as we can always rotate the coordinate system around the $x-$axis to achieve $\vartheta = 0$. In this case we have $\ddot{\phi} = F/A\sin\phi$ which is the same as Eq. 2.25, only that ϕ and θ interchanged their roles. Of course the kink solution in this case again writes

$$\cos\phi_{loop} = 1 - \frac{2}{\cosh^2(s/\lambda)} \tag{2.63}$$

$$\vartheta_{loop} = 0 \tag{2.64}$$

Now we can again expand Eq. 2.62 up to second order and obtain this time

$$\begin{aligned} E[\delta\vartheta, \phi_{loop}+\delta\phi] &= E_{loop} + \frac{1}{2}A\int_{-L/2}^{L/2}\left(\delta\dot{\vartheta}^2 + \frac{F}{A}\delta\vartheta^2\left(\cos\phi_{loop} - \frac{A}{F}\dot{\phi}_{loop}^2\right)\right)ds \\ &\quad + \frac{1}{2}A\int_{-L/2}^{L/2}\left(\delta\dot{\phi}^2 + \frac{F}{A}\cos\phi_{loop}\delta\phi^2\right)ds \end{aligned} \tag{2.65}$$

With $t = s/\lambda$ the latter can again be recast in a more illuminating dimensionless quadratic form

$$\begin{aligned} \beta E[\delta\vartheta, \phi_{loop}+\delta\phi] &= \beta E_{loop} + \frac{1}{2}\int_{-L/2\lambda}^{L/2\lambda}\delta\phi(t)\,\hat{\mathbf{T}}_{\parallel}(t)\,\delta\phi(t)\,dt \\ &\quad + \frac{1}{2}\int_{-L/2\lambda}^{L/2\lambda}\delta\vartheta(t)\,\hat{\mathbf{T}}_{\perp}(t)\,\delta\vartheta(t)\,dt \end{aligned} \tag{2.66}$$

with the **in** and **out** of plane fluctuation operators $\hat{\mathbf{T}}_{\parallel}$ and $\hat{\mathbf{T}}_{\perp}$ given by

$$\hat{\mathbf{T}}_{\parallel} = \beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2} + \left(1 - 2/\cosh^2 t\right)\right) \text{ and} \tag{2.67}$$

$$\hat{\mathbf{T}}_{\perp} = \beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2} + \left(1 - 6/\cosh^2 t\right)\right) \tag{2.68}$$

The operator $\hat{\mathbf{T}}_{\parallel}$ governing the fluctuations of the angle ϕ in the $\vartheta = 0$ plane is of course the same as in 2-D (cf. Eq. 2.31). It is indeed a special case of Eq. 2.43 for $m \to 1$. Although very similar to $\hat{\mathbf{T}}_{\parallel}$ the behavior of the out of plane operator $\hat{\mathbf{T}}_{\perp}$ is fundamentally different. The discrete spectrum of $\hat{\mathbf{T}}_{\perp}$ consists of two eigenvalues (cf. for instance [30]) $\mu_{-1}^{\perp} = -3\beta\sqrt{AF}$ and $\mu_0^{\perp} = 0$ the first of which is negative ! That of course means that the loop conformation has one unstable out of plane direction corresponding to $\mu_{-1}^{\perp}$. The zero eigenvalue mode of $\hat{\mathbf{T}}_{\perp}$ comes from the rotational symmetry around the x-axis in a similar manner as the translational invariance of the loop causes a vanishing eigenvalue[22] in $\hat{\mathbf{T}}_{\parallel}$. It is illuminating to check the latter statement by looking at infinitesimal rotational transformations of the loop in 3-D. Up to quadratic order a rotation of a kink with $\vartheta_{loop} = 0$ around the x-axis by a small angle ε looks like

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-\frac{1}{2}\varepsilon^2 & \varepsilon \\ 0 & -\varepsilon & 1-\frac{1}{2}\varepsilon^2 \end{pmatrix} \begin{pmatrix} \cos\phi_{loop} \\ \sin\phi_{loop} \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\phi_{loop} \\ \sin\phi_{loop} - \frac{1}{2}\varepsilon^2 \sin\phi_{loop} \\ -\varepsilon\sin\phi_{loop} \end{pmatrix} \tag{2.69}$$

On the other hand the perturbed solution $(\delta\vartheta, \phi_{loop} + \delta\phi)$ implies a representation of the tangent

$$\begin{pmatrix} \cos\delta\vartheta\cos\left(\phi_{loop}+\delta\phi\right) \\ \cos\delta\vartheta\sin\left(\phi_{loop}+\delta\phi\right) \\ -\sin\delta\vartheta \end{pmatrix} \approx \begin{pmatrix} \cos\phi_{loop} - \delta\phi\sin\phi_{loop} - \frac{1}{2}\left(\delta\phi^2+\delta\vartheta^2\right)\cos\phi_{loop} \\ \sin\phi_{loop} + \delta\phi\cos\phi_{loop} - \frac{1}{2}\left(\delta\phi^2+\delta\vartheta^2\right)\sin\phi_{loop} \\ -\delta\vartheta \end{pmatrix} \tag{2.70}$$

Here we expanded the left side up to second order in perturbations $\delta\vartheta$ and $\delta\phi$. Matching the two right-hand sides of Eqs. 2.69 and 2.70 we obtain

$$\begin{aligned} \delta\vartheta &\approx \varepsilon\sin\phi_{loop} \\ \delta\phi &\approx -\frac{1}{2}\varepsilon^2\sin\phi_{loop}\cos\phi_{loop} = O\left(\varepsilon^2\right) \end{aligned}$$

By taking the partial derivative $\partial/\partial\varepsilon...|_{\varepsilon=0}$ (at $\varepsilon = 0$) we finally obtain the (unnormalized) mode corresponding to the rotational symmetry of the system. After a convenient normalization ($\int_{-\infty}^{\infty}\vartheta_0^2 dt = 1$) the mode writes

$$\begin{aligned} \vartheta_0(t) &= \left(\tfrac{3}{2}\right)^{1/2}\sin\phi_{loop} = \left(\tfrac{3}{2}\right)^{1/2}\frac{\sinh(t)}{\cosh^2(t)} \\ \phi_0(t) &= 0 \end{aligned} \tag{2.71}$$

Direct insertion of Eq. 2.71 into Eq. 2.68 confirms that ϑ_0 is annihilated by the perpendicular fluctuation operator $\hat{\mathbf{T}}_{\perp}$ (ϕ_0 is annihilated by $\hat{\mathbf{T}}_{\parallel}$ trivially) convincing us that we indeed obtained the rotational zero-eigenvalue Goldstone mode. Note that in lowest order this mode leaves $\phi_0 = 0$ unaffected and that ϑ_0 and ϕ_0 formally decouple.

[22] This appears also formally when the exponentially small eigenvalue μ_0 i.e. $m = 1$ (Eq. 2.56) from last section is set to zero. The (normalized) translational shifting mode becomes then $\phi_0 = 1/\left(\sqrt{2}\cosh t\right)$.

2.2.4 The Unstable Mode

Now it is time to treat the most interesting mode- the unstable mode ϑ_{-1} corresponding to the negative eigenvalue $\mu_{-1}^{\perp} = -3\beta\sqrt{AF}$. One can check by insertion into Eq. 2.68 that

$$\vartheta_{-1}(t) = \left(\tfrac{3}{4}\right)^{1/2} \frac{1}{\cosh^2 t} \tag{2.72}$$

with eigenvalue

$$\mu_{-1}^{\perp} = -3\beta\sqrt{AF} \tag{2.73}$$

is the searched (normalized) unstable mode. The fact that the value $\left|\mu_{-1}^{\perp}\right|$ (which acts as a negative "spring constant" of the mode ϑ_{-}) increases with the force is intuitive as the loop gains energy during the unfolding (and larger forces give larger gains).
Let us now consider cases in which an external potential or DNA self-interaction can compensate for the unstable mode. The simplest, yet formal way to stabilize the loop would be to go to the eigenmode representation of $\hat{\mathbf{T}}_{\perp}$ and introduce an additional potential acting on the ϑ_{-} mode. But it is hard to imagine an interaction or potential living in real space that would act on one mode alone. One more physical situation in which the DNA loop is stabilized is if we switch on (a very strong) magnetic field along the z-axis perpendicular to the force direction (x-axis), cf. Fig. 2.6d). The DNA nucleotides (having π-electrons) are known to prefer alignment perpendicular to the field, i.e., DNA exhibits a negative diamagnetic anisotropy [32]. The application of a magnetic field H along the z-axis drives the DNA molecule into a plane parallel to the x-y plane (perpendicular to the field along the z-axis). The total energy of the DNA writes in this case

$$E[\vartheta,\phi] = \int_{-L/2}^{L/2} \frac{A}{2}\left(\dot{\phi}^2\cos^2\vartheta + \dot{\vartheta}^2\right) - F\cos\phi\cos\vartheta + \frac{\kappa}{2}\sin^2\vartheta ds \tag{2.74}$$

The last term gives the coupling between the DNA tangent and the magnetic field H where $\kappa = -\chi_a H^2/h$ characterizes the coupling strength. Here χ_a is the (experimentally accessible) diamagnetic anisotropy of a single DNA basepair [32] and $h = 0.34nm$ the distance between the subsequent DNA basepairs. Note that χ_a is negative here, i.e., $\kappa > 0$ so $\vartheta = 0$ is the preferred rod orientation for large κ. Expanding $E[\vartheta,\phi]$ again as in Eq. 2.65 we obtain the same expansion[23] as in Eq. 2.66 except that $\hat{\mathbf{T}}_{\perp}$ is replaced by a new out of plane fluctuation operator $\hat{\mathbf{T}}_{\perp}^{\kappa}$

$$\hat{\mathbf{T}}_{\perp}^{\kappa} = \beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2} + \left(1 - \frac{6}{\cosh^2(t)}\right) + \frac{\kappa}{F}\right) \tag{2.75}$$

The spectrum of Eq. 2.75 is of course trivially given by shifting the spectrum of $\hat{\mathbf{T}}_{\perp}$ by the constant $\beta\sqrt{AF}\kappa/F$, i.e., the new eigenvalues now write $\mu_s^{\kappa} = \mu_s^{\perp} + \beta\lambda\kappa$. The rotational Goldstone mode is immediately destroyed for any non-zero coupling constant $\kappa > 0$. More importantly the previously unstable mode ϑ_{-1} now becomes stable provided that $\kappa/F > 3$ i.e. for $\kappa > \kappa_{crit} = 3F$. As the variables ϑ and ϕ decouple

[23] Note that the Euler-Lagrange equation write now slightly different but the homoclinic loop is *still a solution* to the problem.

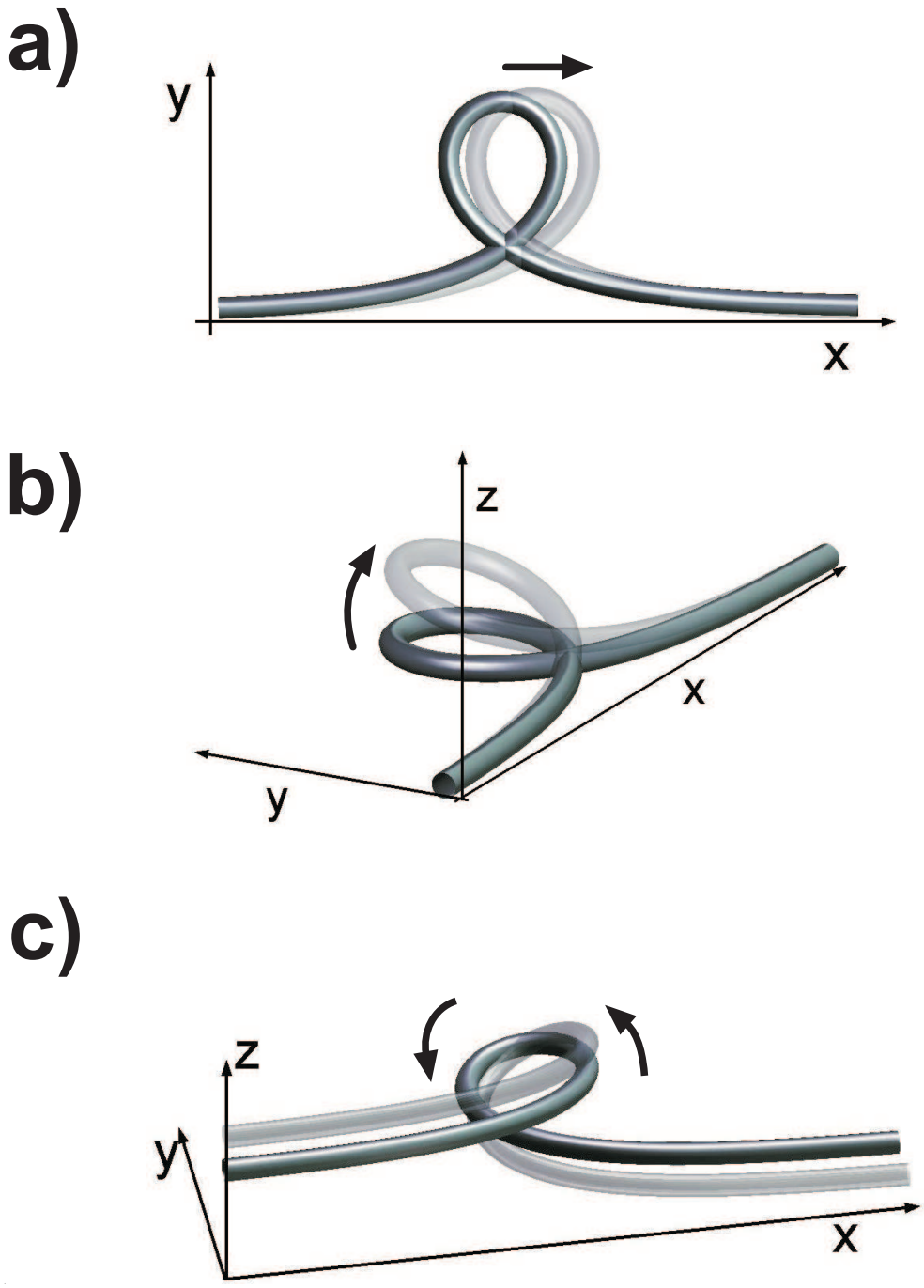


Figure 2.8: The three discrete eigenmodes in action: a) The translational mode $\phi_0 = 1/\left(\sqrt{2}\cosh t\right)$, b) The rotational mode $\vartheta_0 = \sqrt{3/2}\sinh t\cosh^{-2} t$ (Eq. 2.71) and c) The unstable (out of plane tilting) mode $\vartheta_{-1} = \sqrt{3/4}\cosh^{-2} t$ (Eq. 2.72).

(in quadratic expansion) here the full fluctuation partition function $Q^{fluct} = Q_{\parallel} Q_{\perp}^{\kappa}$ factorizes into a product of the in plane ($Q_{\parallel}$) and out of plane ($Q_{\perp}^{\kappa}$) partition functions. The partition function $Q_{\parallel}$ is nothing else but the fluctuation part of the planar partition function from Eq. 2.57 from above. The out of plane partition function $Q_{\perp}^{\kappa}$ is derived via the Gelfand-Yaglom [17] method in the Appendix C. The two expressions write respectively

$$\begin{aligned} Q_{\parallel} &= \frac{\sqrt{2} L l_P}{\pi \lambda^2} e^{-\frac{L}{2\lambda}} && (2.76) \\ Q_{\perp}^{\kappa} &= \sqrt{\frac{l_P}{\lambda} \frac{c\,(c+2)\,(c+1)}{\pi (c-2)(c-1)}} e^{-c\frac{L}{2\lambda}} && (2.77) \\ \text{with } c &= \sqrt{1+\frac{\kappa}{F}} \end{aligned}$$

The free energy then writes

$$\begin{aligned} \beta G &= -\ln\left(e^{-\beta E_{loop}} Q_{\parallel} Q_{\perp}\right) \\ &= -\beta F L + 8\frac{l_P}{\lambda} + \frac{(1+c)\,L}{2\lambda} - \ln\left(\sqrt{\tfrac{2c(c+2)(c+1)}{(c-2)(c-1)}} \tfrac{L l_P^{3/2}}{\pi^{3/2}\lambda^{5/2}}\right) && (2.78) \\ &= -\beta F L + 8\beta\sqrt{FA} + \frac{\left(1+\sqrt{1+\frac{\kappa}{F}}\right) L\sqrt{F/A}}{2} - \ln\left(\sqrt{\tfrac{2c(c+2)(c+1)}{(c-2)(c-1)}} \tfrac{L A^{1/4}}{\pi^{3/2}\beta^{3/2}} F^{5/4}\right) \end{aligned}$$

By differentiating the upper expression by F we immediately obtain the force-extension relation for all forces $F < \kappa/3$. This turns out to be a lengthy expression that we omit here. Here we merely state the result for the limiting case[24] $\kappa \gg F \gg A/L^2$

$$\begin{aligned} \frac{\Delta z}{L} &= -\frac{1}{L}\frac{\partial G}{\partial F} \\ &= \underbrace{1 - \frac{1}{4\beta\sqrt{A\kappa}}}_{I} - \underbrace{\frac{1}{4\beta\sqrt{AF}}\left(1 - \frac{1}{2}\left(\frac{F}{\kappa}\right)^{3/2} + O\left(\frac{F}{\kappa}\right)^{5/2}\right)}_{II} - \underbrace{\frac{4}{L}\sqrt{\frac{A}{F}}}_{III} + O\left(\frac{1}{\beta F L}\right) \end{aligned}$$

It is interesting to have a look at the particular terms appearing in the force-extension relation above. The term II states that (up to corrections in powers of the small parameter F/κ) the chain behaves essentially as a 2-D WLC, cf. Eq. 2.36. The $F-$independent term I says that the effective contour length will appear renormalized by a small relative amount ($\propto \kappa^{-1/2}$). This is valid for $\kappa \gg F$ and disappears if $\kappa \sim F$ (but then the stability of the loop is destroyed and the expansion above becomes invalid). The last term III that can also be written as $4\lambda/L$ corresponds to the length required for the loop formation itself.
Finally, could we experimentally observe the force extension curve derived above? Unfortunately the coupling parameter κ turns out to be too small for reasonable magnetic

[24] The second relation $F \gg A/L^2$ (that usually holds for all DNA stretching experiments) makes the contributions of the $\ln(...)$ term (which scale as $O(\beta F L)^{-1}$) negligible.

fields[25] to be of physical relevance in practice ($\kappa > 3F$). Nevertheless, the formal diamagnetic term $\kappa \sin^2 \vartheta$ introduced in Eq.2.74 is conceptually quite useful to understand the (otherwise unstable) behavior of the DNA loop in 3-D. This case provides a simpler warm-up for the more relevant case of a linker stabilized loop treated in the next section.

2.2.5 The DNA Self-Attraction and the Homoclinic Loop

In the last subchapter we merely "calibrated" our computational machinery and learned how to deal with discrete eigenmodes which turned out to rule the physics of the loop. Here we treat an experimentally much more relevant case in which the 3-D loop turns out to be stable: the DNA self attraction. DNA is known to effectively attract itself in many solvents despite its strong negative bare charge. Typical situations inducing DNA self-attraction are poor solvents (like alcohol, small neutral polymers like PEG) the presence of multivalent counterions (like CoHex and Spermidine) or small cationic proteins acting as linkers between two DNA surfaces. Indeed it was a single molecule stretching experiment on DNA condensed with multivalent counterions(Baumann et al. [131]) that made us think about the force response of loops.
How should we deal with the DNA self-interaction? A formal treatment that first comes to our mind is to introduce a potential $V\left(\|\underline{x}(s_1) - \underline{x}(s_2)\|\right)$ acting between any pair of points s_1 and s_2 on the DNA molecule and to write the total interaction energy as a double integral (over s_1 and s_2) as an additional term in our Hamiltonian. The hope that we can gain any insight by performing this formally correct procedure (in the spirit of ref. Doi and Edwards [23]) quickly turns out to be illusive in our case. Many serious obstacles block our way, like the fact that we (most elegantly) describe the DNA here by the two spherical angles (ϑ and ϕ) of its *tangent* vector. In contrast to that the self-interaction acts in *real space* ("integrated tangent space") which for our purposes makes the Hamiltonian virtually intractable. This stresses the need for a tractable simplification of the DNA self-attraction.
To this end, we make here two simplifying assumptions:

1. There is only *a single discrete* DNA self contact point, given by the crossing point of the homoclinic loop solution.

2. The interaction potential $V\left(\|\underline{x}(s_1) - \underline{x}(s_2)\|\right)$ is *short ranged* enough so that the interaction energy at the crossing becomes independent of the crossing angle i.e. other parts of the DNA (apart from the crossing point) do not interact with each other.

These fairly reasonable assumptions imply that the loop ground state solution will not be significantly modified by the self-attraction and only the fluctuations around it will

[25] The diamagnetic anisotropy per length χ_a/h of a single basepair is around $3 \times 10^{-6} pN/T^2$ [32] i.e. for even very large magnetic fields of say $H = 15T$ we have $\kappa \approx 7 \times 10^{-4} pN$. So for DNA this effect is rather small.

be affected. This means that we can write down the (linearized) loop energy around the solution $\vartheta = 0, \phi = \phi_{loop}$ in a way similar to the last section, namely

$$\begin{aligned}\beta E\left[\delta\vartheta, \phi_{loop} + \delta\phi\right] &= \beta E_{loop} + \tfrac{1}{2}\int_{-L/2\lambda}^{L/2\lambda} \delta\phi \hat{\mathbf{T}}_{\parallel} \delta\phi dt + \tfrac{1}{2}\int_{-L/2\lambda}^{L/2\lambda} \delta\vartheta \hat{\mathbf{T}}_{\perp} \delta\vartheta dt \\ &\quad + V\left(D_c\left(\delta\vartheta\right)\right) \end{aligned} \tag{2.79}$$

The last term $V\left(D_c\right)$ that we introduced here (in accordance with our assumptions 1. and 2.) is the interaction potential of two overcrossing parts of DNA which have a (closest approach) distance D_c. To keep the problem tractable we approximate here the distance D_c by the *perpendicular distance* of the two crossing DNA parts at the *equilibrium (mean) crossing point* t_c (note: $t = s/\lambda$ as above) of the homoclinic loop

$$D_c\left(\delta\vartheta\right) \approx \lambda \int_{-t_c}^{t_c} \sin\delta\vartheta\left(t\right) dt \approx \lambda \int_{-t_c}^{t_c} \delta\vartheta\left(t\right) dt \tag{2.80}$$

The crossing point t_c will be given by the (in plane) projected loop self-crossing. This implies the condition that the integral (over the interval $[-t_c, t_c]$) of the x-component of the loop tangent vanishes i.e.

$$\int_{-t_c}^{t_c} \cos\phi_{loop}\left(t\right) dt = 0 \text{ implying}$$

$$t_c = 2\tanh t_c$$

A numeric solution of the upper implicit condition gives[26] $t_c \approx 1.915$. Before we compute further it is interesting to have a short look at D_c from Eq. 2.80. Because D_c depends only on the out-of-plane perturbations $\delta\vartheta$ the in-plane ($\delta\phi$) problem stays unaffected. It is also intuitive that the out-of-plane rotational Goldstone mode ϑ_0 (the generator of an infinitesimal rotation) leaves the distance D_c unaffected: formally because $\vartheta_0\left(t\right)$ is an odd function, physically because rotations leave distances fixed.
Now it is time to compute the partition function resulting from Eq. 2.79 for any given V. We can write it as

$$Q^V = e^{-\beta E_{loop}} Q_{\parallel} Q_{\perp}^V \tag{2.81}$$

where the in plane partition function $Q_{\parallel}$ is again given by Eq. 2.76. The out of plane partition function $Q_{\perp}^V$ *is* modified by the presence of the contact potential and can be written as

$$Q_{\perp}^V = \frac{1}{\lambda}\int_{-\infty}^{\infty} e^{-\beta V(D_c)} Q_{\perp}\left(D_c\right) dD_c \tag{2.82}$$

where the expression $Q_{\perp}\left(D_c\right)$ denotes the constrained partition function

$$Q_{\perp}\left(D_c\right) = \int_{(0,-\frac{L}{2\lambda})}^{(0,\frac{L}{2\lambda})} \delta\left(\frac{D_c}{\lambda} - \int_{-t_c}^{t_c} \delta\vartheta dt\right) e^{-\frac{1}{2}\int_{-L/2\lambda}^{L/2\lambda} \delta\vartheta \hat{\mathbf{T}}_{\perp} \delta\vartheta dt} \mathcal{D}\left[\delta\vartheta\right] \tag{2.83}$$

Physically the δ function introduced here enforces that the perpendicular distance at the crossing point has the exact value D_c. This path integral is nothing else but

[26] This corresponds to the actual **loop circumference** which is $2 \times 1.915\lambda$.

a linearly constrained Gaussian integral. This can be computed by replacing the δ function by its Fourier representation

$$\delta\left(\frac{D_c}{\lambda}-\int_{-t_c}^{t_c}\delta\vartheta dt\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ip\left(\frac{D_c}{\lambda}-\int_{-t_c}^{t_c}\delta\vartheta dt\right)}dp \tag{2.84}$$

and by rewriting the integral in the exponent more elegantly as a scalar product of $\delta\vartheta$ with a "boxcar" function (called so in literature for its characteristic boxcar like shape)

$$\begin{aligned}\int_{-t_c}^{t_c}\delta\vartheta dt &= \int_{-L/2\lambda}^{L/2\lambda}\Pi\left(t\right)\delta\vartheta\left(t\right)dt=\langle\Pi|\delta\vartheta\rangle \text{ with} \\ \Pi\left(t\right) &= H\left(t+t_c\right)-H\left(t-t_c\right)\end{aligned} \tag{2.85}$$

Here the boxcar function Π is most elegantly represented as a difference of two Heaviside step functions H. The latter is given by $H\left(x\right)=1$ for $x>0$, $H\left(x\right)=1/2$ for $x=0$ and $H\left(x\right)=0$ otherwise. Above, for convenience we introduced the scalar product on the interval $[-L/2\lambda, L/2\lambda]$

$$\langle f|g\rangle=\int_{-L/2\lambda}^{L/2\lambda}f\left(t\right)g\left(t\right)dt$$

In the this notation and by virtue of Eqs.2.84 and 2.85 the partition function $Q_\perp\left(D_c\right)$ (Eq. 2.83) writes more transparently

$$Q_\perp\left(D_c\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ip\frac{D_c}{\lambda}}\int_{(0,-\frac{L}{2\lambda})}^{(0,\frac{L}{2\lambda})}e^{-\frac{1}{2}\langle\delta\vartheta|\hat{\mathbf{T}}_\perp|\delta\vartheta\rangle-ip\langle\Pi|\delta\vartheta\rangle}\mathcal{D}\left[\delta\vartheta\right]dp \tag{2.86}$$

Equivalently, its Fourier transform $\widehat{Q}_\perp\left(p\right)$ writes more compact

$$\widehat{Q}_\perp\left(p\right)=\int_{(0,-\frac{L}{2\lambda})}^{(0,\frac{L}{2\lambda})}e^{-\frac{1}{2}\langle\delta\vartheta|\hat{\mathbf{T}}_\perp|\delta\vartheta\rangle-ip\langle\Pi|\delta\vartheta\rangle}\mathcal{D}\left[\delta\vartheta\right] \text{ (inaccurate)} \tag{2.87}$$

Written in terms of the scalar product $\langle.|.\rangle$ the upper expression looks quite familiar: a Gaussian integral with an additional linear term $ip\langle\Pi|\delta\vartheta\rangle$. So there is a good hope of solving this.

But before we try to evaluate Eq. 2.87 we have to resolve a technical problem first (the reason for Eq. 2.87 being "inaccurate"). The rotational mode ϑ_0 corresponds to a zero eigenvalue and causes divergence of the partition function $Q_\perp$ (and $\widehat{Q}_\perp$ as well) in its naive form like in Eq. 2.87. The problem results from the fact that a rotation of the kink around the x-axis costs no energy and consequently the entropic contribution of this state space direction seemingly diverges (at least in the Gaussian approximation, implied by the saddle point approximation used here). To avoid this problem we make the following trick: instead of $\hat{\mathbf{T}}_\perp$ we use $\hat{\mathbf{T}}_\perp^\kappa$ from Eq. 2.75 and after performing all other calculations we let $\kappa\to 0$ (Note that $\hat{\mathbf{T}}_\perp^\kappa|_{\kappa=0}\equiv\hat{\mathbf{T}}_\perp$!). Physically this procedure corresponds to infinitesimally breaking the rotational symmetry (around the force direction) and restoring it afterwards in a controlled manner in the limit

$\kappa \to 0$. Hand in hand with this limit taking we have to properly renormalize the rotational mode contribution in a similar way we did it in Eq. 2.51 for the translational mode of $\hat{\mathbf{T}}_{\|}$:

$$\widehat{Q}_{\perp}^{ren}(p) = \lim_{\kappa\to 0}^{ren} \int_{(0,-\frac{L}{2\lambda})}^{(0,\frac{L}{2\lambda})} e^{-\frac{1}{2}\langle\delta\vartheta|\hat{\mathbf{T}}_{\perp}^{\kappa}|\delta\vartheta\rangle - ip\langle\Pi|\delta\vartheta\rangle}\mathcal{D}[\delta\vartheta] \tag{2.88}$$

The limiting process together with the "proper renormalization" we denote with a "ren" over the limit sign. In the Appendix D (cf. Eq. 2.143) we evaluate Eq. 2.88. Here we merely state the final result that writes

$$\widehat{Q}_{\perp}^{ren}(p) = i2\sqrt{6}\left(\frac{l_P}{\lambda}\right) e^{-\frac{L}{2\lambda}} e^{\frac{3\lambda}{8l_P}\left(3t_c^2-10\right)t_c p^2} \tag{2.89}$$

This is an easy Gaussian in the Fourier variable p. We now "Fourier-back" to obtain

$$\begin{aligned} Q_{\perp}(D_c) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ip\frac{D_c}{\lambda}}\widehat{Q}_{\perp}^{ren}(p)\,dp \\ &= i\left(\frac{\sqrt{6}l_P}{\pi\lambda}\right)e^{-\frac{L}{2\lambda}} \underbrace{\int_{-\infty}^{\infty} e^{ip\frac{D_c}{\lambda}} e^{\frac{3}{8}\frac{\lambda}{l_P}\left(3t_c^2-10\right)t_c p^2} dp}_{=-i\sqrt{\frac{8\pi l_P}{3\left(3t_c^2-10\right)t_c\lambda}}\exp\left(\frac{2l_P}{3\left(3t_c^2-10\right)t_c\lambda}\left(\frac{D_c}{\lambda}\right)^2\right)} \end{aligned}$$

We can finally write down the important result for the constrained partition function

$$Q_{\perp}(D_c) = 2\sqrt{\frac{6\Gamma}{\pi}}\left(\frac{l_P}{\lambda}\right)^{3/2} e^{-\frac{L}{2\lambda}} \exp\left(\Gamma\frac{l_P}{\lambda}\left(\frac{D_c}{\lambda}\right)^2\right) \tag{2.90}$$

with the scale-independent (negative) elasticity constant for out of plane tilting

$$\Gamma = \frac{2}{3\left(3t_c^2-10\right)t_c} \approx 0.35 \tag{2.91}$$

The result Eq. 2.90 is quite physical: the larger the perpendicular distance D_c (on the λ lengthscale!) the larger the partition function, i.e., the "happier" the system. This is clear as the system without the constraint "D_c=fixed" is intrinsically unstable tending to increase the contact distance D_c!
Having the central formula Eq. 2.90 the rest is an easy exercise: We can compute the partition functions $Q_{\perp}^{V}$ and Q^{V} (Eq. 2.82 and 2.81) for any given (reasonable) potential $V(x)$. Combining Eqs. 2.90, 2.82, 2.81, 2.76 together with $\beta E_{loop} = 8l_P/\lambda - \beta FL$ (Eq. 2.30) we obtain the nice result

$$Q^{V} = \frac{4\sqrt{3\Gamma}Ll_P^{5/2}}{\pi^{3/2}\lambda^{9/2}} e^{\beta FL - 8\frac{l_P}{\lambda} - \frac{L}{\lambda}} \int_{-\infty}^{\infty} e^{\Gamma\frac{l_P}{\lambda}\left(\frac{x}{\lambda}\right)^2 - \beta V(x)} dx \tag{2.92}$$

This expression has to be taken seriously only for sufficiently fast growing interaction potentials $V(x)$ for which the integral above stays finite. Otherwise the system is

metastable and the integral diverges. But even in the case when the bound state, say $x = x_0$, is just a local metastable state the integral above Q^V still makes some sense if the $V(x)$ is very deep. In this case the system can be considered as being in quasi-equilibrium (on some experimentally relevant timescale). Then we can approximate $V(x)$ locally by a quadratic potential

$$V(x) = \frac{1}{2}K(x - x_0)^2$$

and we obtain

$$\begin{aligned} Q^V &= \frac{4\sqrt{6\Gamma}Ll_P^{5/2}}{\pi\lambda^{7/2}\sqrt{\beta\lambda^2 K - 2\Gamma l_P/\lambda}} e^{\frac{\Gamma\beta K l_P/\lambda}{\lambda^2\beta K - 2\Gamma l_P/\lambda}x_0^2} e^{\beta FL - 8\frac{l_P}{\lambda} - \frac{L}{\lambda}} \\ &\approx \frac{4\sqrt{6\Gamma}Ll_P^{5/2}}{\pi\lambda^{9/2}\sqrt{\beta K}} e^{\frac{\Gamma l_P}{\lambda^3}x_0^2} e^{\beta FL - 8\frac{l_P}{\lambda} - \frac{L}{\lambda}} \end{aligned}$$

The last expression is valid in the limit of strong localization i.e. $K\lambda^2 \gg \sqrt{AF}$.
Let us finally write down the force-extension relation resulting from the general expression 2.92. Therefore we conveniently express the free energy $G^V = -\beta^{-1}\ln Q^V$ in terms of F and A (instead of $\lambda = \sqrt{A/F}$)

$$\begin{aligned} \beta G^V &= -\beta FL + (L + 8l_P)\sqrt{\frac{F}{A}} - \ln\left(L^{-1}\int_{-\infty}^{\infty} e^{\beta\Gamma A^{-1/2}F^{3/2}x^2 - \beta V(x)}dx\right) \\ &\quad - \ln\left(\frac{4\sqrt{3\Gamma}L^2\beta^{5/2}A^{1/4}F^{9/4}}{\pi^{3/2}}\right) \end{aligned}$$

The force extension relation resulting from that is

$$\begin{aligned} \frac{\langle x\rangle}{L} &= -\frac{1}{L}\frac{\partial G}{\partial F} \\ &= 1 - \frac{1}{2}\left(1 + 8\frac{l_P}{L}\right)\frac{1}{\beta\sqrt{AF}} + \frac{3\Gamma}{2}\underbrace{\frac{\sqrt{F}}{L\sqrt{A}}}_{(L\lambda)^{-1}}\underbrace{\frac{\int_{-\infty}^{\infty} x^2 e^{\frac{F}{A}\Gamma x^2 - \beta V(x)}dx}{\int_{-\infty}^{\infty} e^{\frac{F}{A}\Gamma x^2 - \beta V(x)}dx}}_{=\langle D_c^2\rangle \sim \text{DNA cross-section}} + O\left(\frac{1}{\beta FL}\right) \end{aligned} \quad (2.93)$$

The last term $O\left((\beta FL)^{-1}\right)$ is always negligible for large forces. The second last term $\sim \langle D_c^2\rangle / (L\lambda)$ is also negligibly small because $\langle D_c^2\rangle$ typically scales as the squared polymer cross-section (for a short ranged surface contact interaction). In the most extreme case the contact distance D_c could become comparable to λ (the loop head size) i.e. $\langle D_c^2\rangle \lesssim \lambda^2$, but the latter is still $\ll L\lambda$. That means that (for reasonable parameters of F and L) the force extension relation of a DNA loop with attractive contact interaction will essentially be independent of the concrete realization of the self-interaction potential $V(x)$. It will consist only of the first terms in Eq. 2.93 and write

$$\frac{\langle x\rangle}{L} = \underbrace{1 - \frac{1}{2\beta\sqrt{AF}}}_{\text{"straight WLC"}} - \underbrace{4\frac{l_P}{L}\frac{1}{\beta\sqrt{AF}}}_{\text{"The loop signature"}} \quad (2.94)$$

The first two terms are the usual "straight WLC" contribution, the last term is the force extension signature of the DNA loop.

2.2.6 Summary: A Simple View of Loop Stretching

We invested quite a lot of work to convince ourselves of something that finally looks almost trivial. The result is fairly independent of many details like how we stabilize the loop in 3-D. In terms of the two relevant lengthscales the persistence length $l_P = \beta A$ and the tension length $\lambda = \sqrt{A/F}$ the mean extension (in very good approximation) looks like:

$$\langle x \rangle = \underbrace{L}_{\text{full length}} - \underbrace{\frac{\lambda}{2l_P} L}_{\text{length "eaten" by fluctuations}} - \underbrace{4\lambda}_{\text{length "eaten" by the loop}} \tag{2.95}$$

The physical reason for this simple decomposition has its roots in the fact that the WLC fluctuations (leading to the 2. term) and its state of deformation (3. term) couple only negligibly in the large force regime (giving merely rise to weak logarithmic corrections in the free energy and negligible $O\left((\beta F L)^{-1}\right)$ corrections in the extension $\langle x \rangle$).

2.3 The "Ghost - l_P" Effect in Polymer Stretching by AFM

Here we answer an experimentally puzzling question [8]: When stretching short semi-flexible polymers by an AFM-tip the persistence length seems to be strongly reduced as compared to that obtained with other methods (e.g. light scattering). We suspect that one of the main reasons for the observed "renormalization" of the persistence length is the same effect that we found in loop stretching. Therefore we shortly recapitulate the basic observations concerning loop stretching:

1. For large forces F the state of the DNA (i.e. looped or not) negligibly modifies the entropy of the chain under tension.

2. The bending energy + potential energy resulting from a nontrivial DNA shape scales as $c\sqrt{AF}$ with some **shape dependent constant** c (e.g. for a complete loop $c = 8$). It contributes significantly to the force extension behavior and modifies the force extension relation by a term $\frac{1}{2}c\lambda$ with $\lambda = \sqrt{A/F}$ the tension length.

3. The length l_{elast} elastically "eaten up" by the conformation

$$l_{elast}(F) = \frac{c}{2}\lambda = \frac{c}{2}\sqrt{A/F} \tag{2.96}$$

   scales in the same manner with F as the length l_{entr} consumed by the entropic undulations ($\langle l_{entr} \rangle = L\left(2\sqrt{l_P F/k_B T}\right)^{-1}$ for a DNA in 3-D)!

4. Instead of the real persistence length in a stretching experiment one would observe the **apparent persistence length** l_P^{app} which we define as

$$\langle x \rangle = L\left(1 - \frac{1}{2}\left(\frac{k_B T}{l_P^{app}}\right)^{1/2} F^{-1/2}\right)$$

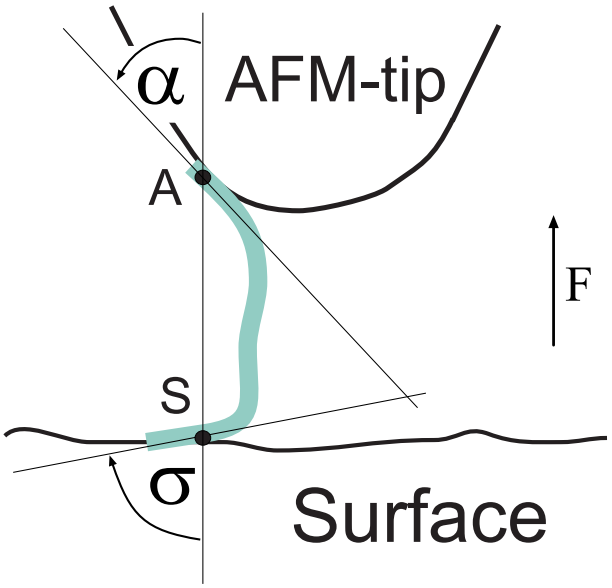


Figure 2.9: AFM stretching of a semiflexible polymer from a suface.

It is given by the elasticity factor c (shape constant) via

$$l_P^{app} = \frac{l_P}{\left(1 + c\frac{l_P}{L}\right)^2} \tag{2.97}$$

For short DNA molecules with $L \sim l_P$ one can measure a l_P^{app} many times smaller than l_P!

Now interestingly the same effect can play a role in stretching DNA (or a general semiflexible polymer) attached to a surface (mica) and the AFM tip on the other side. Consider Fig. 2.9 where a typical setup of a single molecule AFM pulling experiment is depicted. Here an AFM-tip catches one end of the polymer at a certain point A and at a certain angle α with respect to the line $\overline{AS}$ connecting A with the surface attachment point S (here for simplicity taken parallel to the pulling direction). The angle between the polymer tangent and $\overline{AS}$ at the surface is denoted by σ. The importance of the two orientational angles α and σ for the outcome of the force-extension measurement has to our knowledge been completely overseen in the "force - spectroscopy" literature. Let us see how crucial α and σ really are by computing the constant $c = c\,(\alpha, \sigma)$ giving the "elastically lost" length l_{elast} in Eq. 2.96 and l_P^{app} in Eq. 2.97.
We assume here that the force is large enough so that $\lambda/L \ll 1$, i.e., $F \gg A/L^2$ (which is given for typical polymer lengths and the large forces in AFM experiments, $F \sim 0.1 - 1nN$). The two polymer ends in Fig. 2.9 decay to the straight configuration[27] in the same manner as the tangent deflections around the loop-head decayed to the force direction, cf. Fig. 2.7. In fact the two ends are asymptotically ($\lambda/L \ll 1$) nothing else

[27] If we refrain from thermal undulations which decouple from the shape as we have seen before.

but parts of the homoclinic loop in Fig. 2.7. The "lost length"at the point A is given by the expression

$$\begin{aligned} l^A_{elast}(F,\alpha) &= \int_{s_0(\alpha)}^{\infty} (1-\cos\theta_{loop}(s))\,ds = 2\int_{s_0}^{\infty} \frac{ds}{\cosh^2(s/\lambda)} \\ &= 2\lambda\,(1-\tanh(s_0/\lambda)) \\ &= 2\lambda\left(1-\cos\frac{\alpha}{2}\right) \end{aligned}$$

with θ_{loop} from Eq. 2.27. In the last line we used the identity $\theta_{loop}(s_0)=\alpha$ so that via Eq. 2.27 we can eliminate the hyperbolic functions of s_0 in favor of trigonometric ones of α, in particular $\cos\frac{\alpha}{2}=\tanh\left(\frac{s_0}{\lambda}\right)$. In the same manner we also obtain the "lost length"at the point S

$$l^S_{elast}(F,\sigma) = 2\lambda\left(1-\cos\frac{\sigma}{2}\right)$$

And for the total "lost length" $l_{elast}=l^A_{elast}+l^S_{elast}$ we obtain

$$l_{elast}(F,\alpha,\sigma) = 4\lambda\left(1-\frac{1}{2}\left(\cos\frac{\alpha}{2}+\cos\frac{\sigma}{2}\right)\right)$$

We see that the shape constant $c=2l_{elast}/\lambda$ can be very large. Consequently the apparent persistence length can be much smaller than expected, namely

$$l_P^{app}(\alpha,\sigma) = \frac{l_P}{\left(1+8\left(1-\frac{\cos(\alpha/2)+\cos(\sigma/2)}{2}\right)\frac{l_P}{L}\right)^2} \tag{2.98}$$

This is the central formula derived in this section. For a typical experiment with short polymers, $L/l_P=5$ with anchoring angles $\alpha=\sigma=\pi/2$ (i.e. perpendicular to the force direction) we have $l_P^{app}=0.46l_P$! For longer polymers the effect becomes smaller and asymptotically vanishes but that very slowly: for $L/l_P=20$ we still have $l_P^{app}=0.8l_P$ and for $L/l_P=40$ we have $0.9l_P$. That means the effect we discovered here can indeed be experimentally easily resolved for even quite large L/l_P! To turn it around and formulate sharply: **if one does not account for the boundary anchoring effects the experimentally fitted (apparent) "persistence length" is a useless information for polymers of short to intermediate length ($L/l_P \leq 20$)!**
How should an experimentalist resolve this ambiguity? The assumption $\alpha=\sigma=\pi/2$ (tangents perpendicular to the force $\overrightarrow{F}$ and tangential to the two surfaces) will probably be a good initial guess for adsorbed polymers on flat surfaces on both ends. But the real unknown will be the shape of the AFM-tip and the exact anchoring point A. By looking at the Fig. 2.9 one can state: The sharper the tip is the more unpredictable the outcome will be (due to large differences in possible angles α). Can one somehow measure directly the angles α and σ by a stretching experiment?
A good idea for doing that is to use an additional degree of freedom that the AFM has: the lateral shifting of the tip. Such a situation is depicted in Fig. 2.10. After a shift by a distance $\Delta x=\left|\overline{AB}\right|$ parallel to the surface the connection line $\overline{AS}$ is no more parallel to the force. The *effective angles* $\alpha_{eff}=\alpha+\tau$ and $\sigma_{eff}=\sigma+\tau$ giving rise to

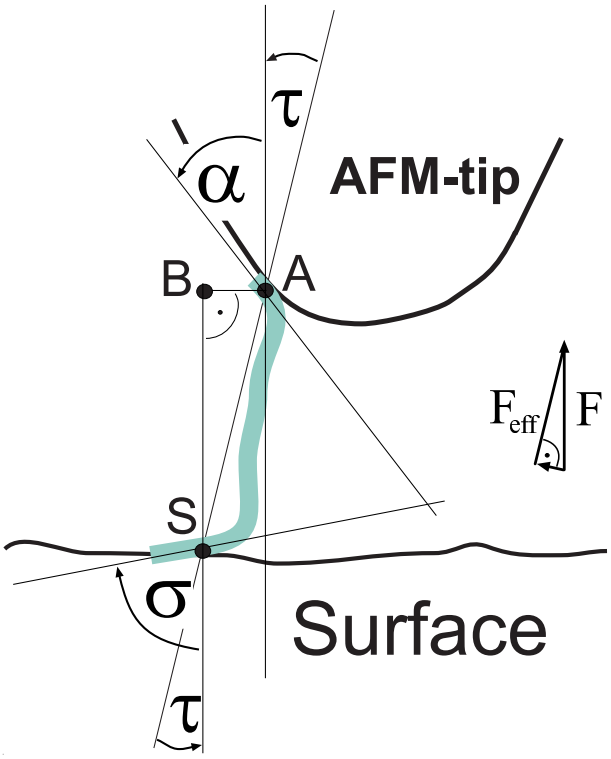


Figure 2.10: AFM polymer stretching with a lateral shift.

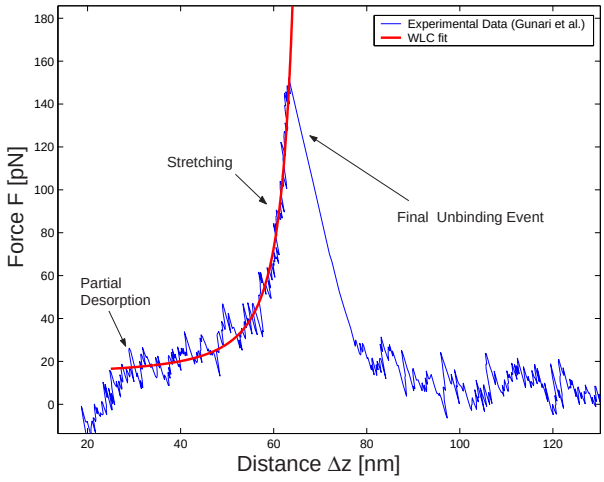


Figure 2.11: A typical AFM force-extension curve of a single PNI-PAM brush polymer (data kindly provided by Gunari et. al. [8]). The measured data (blue thin line) are well fitted by the wormlike chain expression $F = \frac{k_B T}{4 l_P^{app}}(1 - \Delta z/L)^{-2} + F_{ads}$ (red thick line) with an additional offset force F_{ads} resulting from a partial desorption of the polymer. The optimal fitting parameters are $F_{ads} = 13.7pN$, $L = 69.8nm$ and $l_P^{app} = 0.86nm$

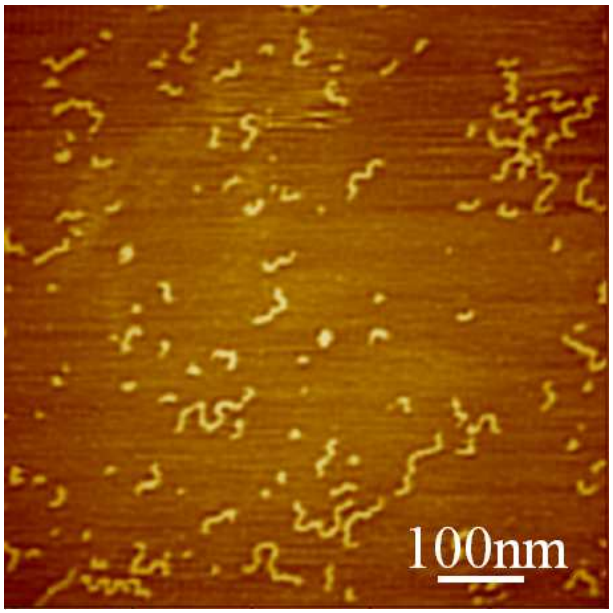


Figure 2.12: AFM picture of the cylindrical PNIPAM brush polymers adsorbed on mica substrate (in water at T=20C). Picture kindly provided by Gunari et al. [8]

an elastic response (along the effective direction $\overline{AS}$) get modified by the tilting angle $\tau = \arctan(\Delta x/\Delta z)$. By modifying Δx properly while measuring the force-extension behavior we can recover the anchoring angles α and σ! The details of this procedure would lead us too far here and will be discussed elsewhere in future.

Let us finally take a look at a typical experimental force-extension curve [8] obtained for the cylindrical-brush polymers with PNIPAM side chains [7] by Gunari et. al., cf. Fig. 2.11. These PNIPAM brushes are known to behave as semiflexible chains with a persistence length close to 30nm, cf. Fig. 2.12. A closer inspection of the experimental data reveals that our model of a tangentially end-grafted polymer is slightly idealized: The polymer indeed shows a typical continuous desorption behavior in the small force region. This effect is neglected in our theory that assumes point anchoring and neglects desorption for simplicity. The typical (apparent) adsorption force $F_{ads} = 14pN$ indicates an adsorption energy line density of roughly $4k_BT/nm$. Despite this complication, in the large force regime ($F > 40pN$) the curve shows the typical wormlike chain behavior. The apparent persistence length obtained from that region is tiny (typically $0.5-2nm$) and indeed much smaller than the "real" persistence length $l_P \sim 30nm$. The latter can be estimated from scattering experiments and from AFM pictures of surface adsorbed PNIPAM brushes (cf. Fig.2.12).

The reduction of the apparent persistence length is in qualitative agreement with our theory but quantitatively it turns out to be significantly stronger than predicted by Eq. 2.98. Inserting the experimental parameters $L = 69.8nm$ and $l_P = 30nm$ into Eq. 2.98 and assuming again tangential anchoring ($\alpha = \sigma = \pi/2$) we obtain the apparent persistence length $l_P^{app} = 7.4nm$ which is still an order of magnitude larger than the one obtained from the fit ($l_P^{app} = 0.86nm$). A closer theoretical inspection of

the experimental setup reveals that the simplified picture from figure 2.9 with the line AS of anchoring points being parallel to the force direction is generically violated in the experiments of Gunari et al. In fact due to continuous chain desorption the angle $\widehat{ASB}$ (cf. Fig. 2.10) is changing in the course of the experiment in a way that reflects the 2-D (quenched) persistent random walk behavior of the adsorbed part of the chain (on the substrate surface). The measured force F and the actual tension acting in the chain F_{eff} are no longer identical. It is easy to see that this effect reduces the apparent persistence length by an other factor of $\cos\left(\widehat{ASB}\right)$.

The details of these effects are still under intense joint theoretical and experimental [8] investigations.

2.4 Appendices

2.4.1 Appendix A: The Saddle Point Approximation for Path Integrals

Here we shortly review the most basic method for lowest order approximation of many path integrals. In literature it is usually called the saddle point, semiclassical or WKB approximation.
Consider a path integral for the transition amplitude of a quantum mechanical particle in a most general potential $V(x)$

$$(x_1,t_1|x_0,t_0)_V = \int_{(x_0,t_0)}^{(x_1,t_1)} \mathcal{D}[x]\, e^{\frac{i}{\hbar}\mathcal{A}[x]} \tag{2.99}$$

with the Lagrangian L action $\mathcal{A}$ given by

$$\mathcal{A}[x] = \int_{t_0}^{t_1} L[x]\, dt \tag{2.100}$$

$$L[x] = \frac{1}{2}m\dot{x}^2 - V(x) \tag{2.101}$$

The integral in Eq. 2.99 is in general not analytically accessible but there are simple approximation schemes. The simplest of them is the saddle point approximation which consists of expanding the action around a "classical" solution x_{cl} given by

$$\delta\mathcal{A}[x]\,|_{x=x_{cl}} = 0$$

i.e. the Euler-Lagrange equation

$$m\ddot{x}_{cl}(t) = -V'(x_{cl}(t)) \tag{2.102}$$

with the two boundary conditions $x\left(t_{0/1}\right) = x_{0/1}$. Here $(...)'$ is the derivative with respect to the variable x.
Taking now $x = x_{cl} + \delta x$ the action $\mathcal{A}[x]$ (Eq. 2.100) can be expanded up to second order in the small perturbation δx around the classical path:

$$\begin{aligned}\mathcal{A}[x_{cl}+\delta x] &= \mathcal{A}[x_{cl}] + \int_{t_0}^{t_1}\left(\frac{\partial L}{\partial \dot{x}}\delta\dot{x} + \frac{\partial L}{\partial x}\delta x\right)dt \\ &\quad + \frac{1}{2}\int_{t_0}^{t_1}\left(\frac{\partial^2 L}{\partial \dot{x}^2}\delta\dot{x}^2 + 2\frac{\partial^2 L}{\partial x \partial \dot{x}}\delta x\delta\dot{x} + \frac{\partial^2 L}{\partial x^2}\delta x\right)dt + ...\end{aligned}$$

The second term above (the first variation) vanishes and one obtains

$$\mathcal{A}[x_{cl}+\delta x] = \mathcal{A}[x_{cl}] + \frac{1}{2}\int_{t_0}^{t_1}\left[m(\delta\dot{x})^2 - V''(x_{cl}(t))(\delta x)^2\right]dt + ...$$

Inserting this in Eq. 2.99 gives

$$(x_1,t_1|x_0,t_0)_V = e^{\frac{i}{\hbar}\mathcal{A}[x_{cl}]}\int_{(0,t_0)}^{(0,t_1)}\mathcal{D}[\delta x]\, e^{\frac{i}{2\hbar}\int_{t_0}^{t_1}\left[m(\delta\dot{x})^2 - V''(x_{cl}(t))(\delta x)^2\right]dt} + ... \tag{2.103}$$

The first factor here comes from the classical action. The second gives the contribution of quadratic fluctuations around x_{cl} and it is formally the path integral of a harmonic oscillator with a time dependent frequency $\omega^2(t) = \frac{1}{m}V''(x_{cl}(t))$ (but with Dirichilet boundary conditions $\delta x(t_0) = \delta x(t_1) = 0$). There are (at least) three equivalent methods to write Eq. 2.103 explicitly and depending on the physical problem one might consider the one or the other is more elegant, didactic or intuitive.

1. The first method rewrites the exponent in Eq. 2.103 (after partial integration) as

$$\frac{i}{2\hbar}\int_{t_0}^{t_1}\left[m(\delta\dot{x})^2 - V''(x_{cl}(t))(\delta x)^2\right]dt = \frac{im}{2\hbar}\int_{t_0}^{t_1}(\delta x)\hat{T}(\delta x)\,dt$$

with the so called fluctuation operator

$$\hat{T} = -\left(\frac{\partial^2}{\partial t^2} + \frac{V''(x_{cl}(t))}{m}\right)$$

The latter operator being Hermitian has real eigenvalues μ_n and a complete orthonormal set of eigenfunctions v_n. Writing Eq. 2.103 in this base $\hat{T}$ diagonalizes and one obtains

$$(x_1,t_1|x_0,t_0)_V = \sqrt{\frac{m}{2\pi i\hbar D(t_1,t_2)}}e^{\frac{i}{\hbar}\mathcal{A}[x_{cl}]} \tag{2.104}$$

with the fluctuation determinant

$$\frac{D(t_1,t_0)}{t_1-t_0} = \frac{\det\left(\hat{T}\right)}{\det\left(-\frac{\partial^2}{\partial t^2}\right)} = \prod_{k=1}^{\infty}\frac{\mu_k}{\pi^2k^2/(t_1-t_0)^2} \quad \text{(Method 1)} \tag{2.105}$$

The last expression (i.e. the product of eigenvalue ratios of the two operators with Dirichilet boundary conditions) can be taken as a definition for $D(t_1,t_0)$. This method requires the knowledge of the whole spectrum μ_k of $\hat{T}$ and can be quite technical. However it can be quite instructive in cases when some μ_k approach zero (or turn negative) and the physics behind formal divergence of $(x_1,t_1|x_0,t_0)_V$ can be understood better.

2. A second method by Gelfand and Yaglom [17] uses Eq. 2.103 but computes the eigenvalue ratio $D(t_1,t_0)$ without explicit knowledge of the whole spectrum. It can be shown that $D(t,t_0)$ satisfies the following initial value problem

$$\frac{d^2D(t,t_0)}{dt^2} = -\omega^2(t)D(t,t_0) \text{ with initial conditions} \tag{2.106}$$

$$D(t_0,t_0) = 0 \text{ and } \dot{D}(t_0,t_0) = 1 \text{ (Method 2)} \tag{2.107}$$

Here we set $\omega^2(t) = \frac{1}{m}V''(x_{cl}(t))$. The eigenvalue ratio is then given by

$$D(t_1,t_0) = D(t,t_0)|_{t=t_1}$$

A practical way of solving of Eqs. 2.106 and 2.107 comes from the observation that Eqs. 2.106 without the condition Eq. 2.107 (by virtue of Eq. 2.102) admits a particular solution, namely $\dot{x}_{cl}(t)$. Finally by d'Alembert's construction a solution that satisfies the boundary conditions Eq. 2.107 can be written down as

$$D(t,t_0) = \dot{x}_{cl}(t_0)\,\dot{x}_{cl}(t)\int_{t_0}^{t}\frac{d\tau}{\dot{x}_{cl}^2(\tau)} \text{ (improved Method 2)} \tag{2.108}$$

Evaluation of that at $t = t_1$ gives the fluctuation factor $D(t_1, t_0)$.

3. The third method is useful if we explicitly know the action $\mathcal{A}[x_{cl}]$ from Eq. 2.100 as function of the two boundary values x_0, x_1, i.e. we if can write $\mathcal{A}[x_{cl}] = \mathcal{A}[x_{cl}; x_0, x_1]$. Then the integral above can be computed giving the following beautiful and simple formula (cf. [13, 14])

$$(x_1,t_1|x_0,t_0)_V = \sqrt{\frac{i}{2\pi\hbar}\frac{\partial^2\mathcal{A}[x_{cl};x_0,x_1]}{\partial x_0\partial x_1}}\,e^{\frac{i}{\hbar}\mathcal{A}[x_{cl};x_0,x_1]} \text{ (Method 3)} \tag{2.109}$$

This remarkable expression by Van Vleck, Pauli and Morette says that the fluctuation contribution (up to quadratic order) is given by partial derivatives of the classical action $\mathcal{A}[x_{cl}]$! A nice thing about "Method 3" is that it possesses a straightforward generalization to n-dimensional systems (with $\underline{x}(t) = \left(x^k(t)\right)$ now a n-dimensional vector function of t)

$$(\underline{x}_1,t_1|\underline{x}_0,t_0)_V = \sqrt{\left(\frac{i}{2\pi\hbar}\right)^n \det\left(\frac{\partial^2\mathcal{A}[\underline{x}_{cl};\underline{x}_0,\underline{x}_1]}{\partial x_0^k\partial x_1^l}\right)}\,e^{\frac{i}{\hbar}\mathcal{A}[\underline{x}_{cl};\underline{x}_0,\underline{x}_1]} \text{ (Method 3 in n-D)}$$

The involved expression $\det(...)$ is called the Van Vleck-Pauli-Morette determinant.

An illustration: a (1-D) harmonic oscillator has the classical action

$$\mathcal{A}_h = \frac{m\left(\omega\left(x_0^2+x_1^2\right)\cos\omega\left(\tau_1-\tau_0\right)-2x_0x_1\right)}{2\sin\omega\left(t_1-t_0\right)}$$

which by virtue of Eq. 2.109 gives the well known propagator

$$(\underline{x}_1,t_1|\underline{x}_0,t_0)_h = \sqrt{\frac{\omega m}{2\pi i\hbar\sin\omega\left(\tau_1-\tau_0\right)}}\,e^{\frac{im\omega}{2\hbar}\frac{\left(x_1^2+x_2^2\right)\cos\omega\left(\tau_1-\tau_0\right)-2x_1x_2}{\sin\omega\left(\tau_1-\tau_0\right)}}$$

Here the saddle point approximation gives the exact result as the Lagrangian of the harmonic oscillator is a quadratic function.

If we want to apply these results to the WLC we just have to rename the involved constants and the time variable τ via the formal mapping: $\tau \to -it$ ("Wick rotation") with $t = s/\lambda$ and the replacement $1/\hbar \to \beta$, $m \to \sqrt{AF}$, $\omega^2(t) \to 1$ for the straight DNA case or generally $\omega^2(t) \to$"Fluctuation potential" (e.g. "$1 - 2/\cosh^2(t)$" in the planar loop case).

2.4.2 Appendix B: Lamé - Type Eigenvalue Problem

Consider an operator of the form

$$\hat{\mathbf{L}}_N = -\frac{\partial^2}{\partial t^2} + \left(N\left(N+1\right)\overline{m}\mathrm{sn}^2\left(t|\overline{m}\right) - 1\right) \tag{2.110}$$

and the corresponding eigenvalue problem

$$\hat{\mathbf{L}}_N f_k\left(t\right) = \nu_k f_k\left(t\right) \tag{2.111}$$

The problem of finding these eigenvalues falls into a class of "quasi exactly solvable" problems and typically appears in quantum mechanical problems (like for electrons in a periodic potential). The corresponding differential equation is called the Lamé equation [18]. It can admit simple solutions in terms of polynomials of elliptic functions sn, cn and dn provided that N is an integer.
In the simplest case $N = 1$ the discrete spectrum and the corresponding eigenfunctions of $\hat{\mathbf{L}}_2$ are given by

$$\begin{aligned}
\nu_{-1} &= \overline{m} - 1 \text{ and } f_{-1}\left(t\right) = \mathrm{dn}\left(t|\overline{m}\right) \\
\nu_0 &= 0 \text{ and } f_0\left(t\right) = \mathrm{cn}\left(t|\overline{m}\right) \\
\nu_1 &= \overline{m} \text{ and } f_1\left(t\right) = \mathrm{sn}\left(t|\overline{m}\right)
\end{aligned}$$

For a given eigenvalue ν_k in addition to $f_k\left(t\right)$ there is always another linearly independent solution $\widetilde{f}_k\left(t\right)$ with just the opposite even / odd symmetry given by the d'Alembert construction for linear second order differential equations

$$\widetilde{f}_k\left(t\right) = f_k\left(t\right)\int_0^t \left(\frac{1}{f_k\left(t'\right)}\right)^2 dt' \tag{2.112}$$

In addition $\hat{\mathbf{L}}_N$ has a continuous spectrum starting after a gap above ν_1.
Now the operator from Eq.2.110 is not exactly the one we need in Eq. 2.43. To obtain the latter we have to analytically continue the operator $\hat{\mathbf{L}}_N$ to values of $\overline{m}$ larger than 1 and introduce the parameter $m = 1/\overline{m} < 1$ then by Jacobi's real transformation [85] we obtain

$$\hat{\mathbf{T}}_N = -\frac{\partial^2}{\partial t^2} + \left(N\left(N+1\right)\mathrm{sn}^2\left(\frac{t}{\sqrt{m}}|m\right) - 1\right) \tag{2.113}$$

whose discrete spectrum and eigenvalues can be obtained from the corresponding ones of $\hat{\mathbf{L}}_N$. For $N = 1$ they write

$$\begin{aligned}
\nu_0 &= 0 \text{ and } f_0\left(t\right) = \mathrm{dn}\left(\frac{t}{\sqrt{m}}|m\right) \\
\nu_{-1} &= \frac{1}{m} - 1 \text{ and } f_{-1}\left(t\right) = \mathrm{cn}\left(\frac{t}{\sqrt{m}}|m\right) \\
\nu_1 &= \frac{1}{m} \text{ and } f_1\left(t\right) = \mathrm{sn}\left(\frac{t}{\sqrt{m}}|m\right)
\end{aligned}$$

We can now impose boundary conditions as in text above

$$f\left(\pm\sqrt{m}K\left(m\right)\right)=0 \tag{2.114}$$

We immediately observe that for these boundary conditions $\nu_0 = 0$ is not an eigenvalue anymore because **no** linear combination of $f_{-1}(t)$ and the corresponding $\widetilde{f}_0(t)$ (given by Eq. 2.112) can be constructed to satisfy Eq. 2.114. But the mode f_{-1} **does** satisfy Eq. 2.114 so that $\nu_{-1} = \frac{1}{m} - 1$ becomes **the smallest eigenvalue** of $\hat{\mathbf{T}}_N$ (Eq. 2.112) with these boundary conditions.
Note 1: The eigenvalue $\nu_{-1} = \frac{1}{m} - 1$ is denoted with μ_0 in the main text.
Note 2: In the limiting case $\overline{m}, m \to 1$ the elliptic functions reduce to hyperbolic functions and we recover the fluctuation operator Eq. 2.31 up to a trivial prefactor.

2.4.3 Appendix C: The Fluctuation Determinant for $\hat{\mathbf{T}}_{\perp}^{\kappa}$

Here we compute the fluctuation partition function Q^{fluct} for the DNA loop in magnetic field. $Q^{fluct} = Q_{\parallel} Q_{\perp}^{\kappa}$ is the product of two independent partition functions $Q_{\parallel}$ and $Q_{\perp}^{\kappa}$ (Eqs. 2.76 and 2.77) for the in plane and out of plane fluctuations, the first of which was already computed. Here we derive $Q_{\perp}^{\kappa}$ leading to Eq. 2.77 in the main text. It is defined in terms of the corresponding fluctuation operator

$$\hat{\mathbf{T}}_{\perp}^{\kappa} = \beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2} - \frac{N\left(N+1\right)}{\cosh^2\left(t\right)} + c^2\right) \tag{2.115}$$

with $c = \sqrt{1+\frac{\kappa}{F}}$ and $N = 2$ via

$$Q_{\perp}^{\kappa} = \int_{\left(0,-\frac{L}{2\lambda}\right)}^{\left(0,\frac{L}{2\lambda}\right)} e^{-\frac{1}{2}\int_{-L/2\lambda}^{L/2\lambda}\delta\vartheta\hat{\mathbf{T}}_{\perp}^{\kappa}\delta\vartheta dt}\mathcal{D}\left[\delta\vartheta\right] = \sqrt{\frac{\beta\sqrt{AF}}{2\pi D\left(-\frac{L}{2\lambda},\frac{L}{2\lambda}\right)}} \tag{2.116}$$

The fluctuation determinant $D\left(-\frac{L}{2\lambda},\frac{L}{2\lambda}\right)$ is most elegantly computed via the Gelfand-Yaglom initial value problem described above (Appendix A) which states that $D\left(-\frac{L}{2\lambda},\frac{L}{2\lambda}\right) = f\left(t\right)|_{t=L/2\lambda}$ where $f(t)$ is the solution to

$$\hat{\mathbf{T}}_{\perp}^{\kappa} f\left(t\right) = 0 \text{ with initial values} \tag{2.117}$$

$$f\left(-\frac{L}{2\lambda}\right) = 0 \text{ and } \dot{f}\left(-\frac{L}{2\lambda}\right) = 1 \tag{2.118}$$

The solution of Eq. 2.117 with 2.115 can be written in terms of the two linearly independent solutions (cf. [19])

$$f_{\pm}\left(x\right) = \cosh^{N+1}\left(t\right)\left(\frac{1}{\cosh t}\frac{d}{dt}\right)^{N+1} e^{\pm ct} \tag{2.119}$$

For $N = 2$ the general solution writes

$$\begin{aligned} f\left(t\right) = {} & C_1 e^{ct}\left((c-2\tanh t)(c-\tanh t) - \cosh^{-2}\left(t\right)\right) \\ & + C_2 e^{-ct}\left((c+2\tanh t)(c+\tanh t) - \cosh^{-2}\left(t\right)\right) \end{aligned} \tag{2.120}$$

Enforcing the Gelfand-Yaglom initial conditions in the limit $\frac{L}{2\lambda} \gg 1$ (where we may safely set $\tanh\left(\pm\frac{L}{2\lambda}\right) \approx \pm 1$, $\cosh^{-2}\left(\frac{L}{2\lambda}\right) \approx 0$) we obtain the conditions

$$\begin{aligned} f\left(-\frac{L}{2\lambda}\right) &= C_1(c+2)(c+1)e^{-c\frac{L}{2\lambda}} + C_2(c-2)(c-1)e^{c\frac{L}{2\lambda}} = 0 \\ \dot{f}\left(-\frac{L}{2\lambda}\right) &= C_1 c(c+2)(c+1)e^{-c\frac{L}{2\lambda}} - C_2 c(c-2)(c-1)e^{c\frac{L}{2\lambda}} = 1 \end{aligned}$$

which imply

$$\begin{aligned} C_1 &= \frac{1}{2c\,(c+2)\,(c+1)}e^{c\frac{L}{2\lambda}} \\ C_2 &= -\frac{1}{2c\,(c-2)\,(c-1)}e^{-c\frac{L}{2\lambda}} \end{aligned}$$

Evaluating $f\,(t)$ at the right boundary $t = \frac{L}{2\lambda}$ we obtain

$$\begin{aligned} D\left(-\frac{L}{2\lambda}, \frac{L}{2\lambda}\right) &= \frac{(c-2)(c-1)}{2c\,(c+2)\,(c+1)}e^{c\frac{L}{\lambda}} \times \left(1 + O\left(e^{-\frac{L}{2\lambda}}\right)\right) \\ &\approx \frac{(c-2)(c-1)}{2c\,(c+2)\,(c+1)}e^{c\frac{L}{\lambda}} \end{aligned} \tag{2.121}$$

The asymptotically unimportant term $O\left(e^{-\frac{L}{2\lambda}}\right)$ in the first line can be safely neglected for $L/2\lambda \gg 1$.
Inserting finally Eq.2.121 into Eq.2.116 we obtain

$$\begin{aligned} Q_\perp^\kappa &= \sqrt{\frac{l_P}{\lambda}\frac{c\,(c+2)\,(c+1)}{\pi(c-2)(c-1)}}e^{-c\frac{L}{2\lambda}} \\ \text{with } c &= \sqrt{1+\frac{\kappa}{F}} \end{aligned} \tag{2.122}$$

Note that unlike for the in plane operator $Q_\parallel$ case (where a close to 0 eigenmode appears and creates artifacts) here we need not to renormalize $D\left(-\frac{L}{2\lambda}, \frac{L}{2\lambda}\right)$ and $Q_\perp^\kappa$ as long as c is sufficiently larger[28] than 2.

2.4.4 Appendix D: Computation of the Perpendicularly Constrained Partition Function

Here we evaluate the path-integral from Eq. 2.88. It is a special realization of the general path integral of a QM harmonic oscillator with a time dependent frequency

[28] If on the other hand c is sufficiently close to 2 we will meet the problem of a formal divergence of $Q_\perp^\kappa$. This again is an artifact as for $\kappa \approx 3F$ the system becomes unstable (the loop unfolds). If one is really interested in the limit $\kappa \to 3F$ one has to check again how much of the state-space does the almost zero eigenvalue mode really *physically* occupy (beyond the Gaussian treatment). The vanishing mode in this case represents the loop unfolding direction and its state space volume will scale with λ (loop-head-size) rather than with L as it is the case for the translational mode contribution to the in plane fluctuations.

$\omega(\tau)$ and a driving force term $j(\tau)$. The latter writes

$$I[\omega, j] = \int_{(x_0,\tau_0)}^{(x_1,\tau_1)} \mathcal{D}[x] \underbrace{e^{\frac{i}{\hbar}\left(\int_{\tau_0}^{\tau_1} \frac{m}{2}\left(\dot{x}^2(\tau)-\omega^2(\tau)x^2(\tau)\right)d\tau + \int_{\tau_0}^{\tau_1} j(\tau)x(\tau)d\tau\right)}}_{=e^{\frac{i}{\hbar}S[j,x]}} \tag{2.123}$$

The latter can be computed exactly (cf. [13, 14])

$$I[\omega, j] = \underbrace{\sqrt{\frac{m}{2\pi i\hbar D(\tau_1,\tau_0)}}}_{\text{fluctuation contribution}} \underbrace{e^{\frac{i}{\hbar}S[j,x_{cl}]}}_{\text{classical contribution}} \tag{2.124}$$

The first term $D(\tau_1,\tau_0)$ here is the functional determinant of the (j-independent!) operator $\hat{\mathbf{T}} = d^2/d\tau^2 + \omega^2(\tau)$ normalized by the free-particle operator $d^2/d\tau^2$ i.e.

$$\frac{D(\tau_1,\tau_0)}{\tau_1-\tau_0} = \det\left(\frac{d^2/d\tau^2+\omega^2(\tau)}{d^2/d\tau^2}\right) \tag{2.125}$$

The Gelfand-Yaglom method described above can be used to immediately obtain $D(\tau_1,\tau_0)$ (exactly as in the previous case without the driving force j).
The second term in Eq. 2.124 involves the classical action $S[j,x_{cl}]$ which after short computation (partial integration and exploiting the equation of motion Eq. 2.127 below) can be recast into

$$S[j,x_{cl}] = \frac{m}{2}x_{cl}(\tau)\dot{x}_{cl}(\tau)|_{\tau_0}^{\tau_1} + \frac{1}{2}\int_{\tau_0}^{\tau_1} j(\tau)x_{cl}(\tau)d\tau \tag{2.126}$$

The latter involves the classical path $x_{cl}(\tau)$ depends now on j and is the solution of the corresponding Euler-Lagrange equation (classical equation of motion)

$$m\ddot{x}_{cl}(\tau) + m\omega^2(\tau)x_{cl}(\tau) = j(\tau) \tag{2.127}$$

with boundary conditions $x_{cl}\left(\tau_{0/1}\right) = x_{0/1}$.
Now in our concrete case we have to evaluate

$$\begin{aligned}\widehat{Q}_\perp(p) &= \int_{(0,t_0)}^{(0,t_1)} e^{-\frac{1}{2}\int_{t_0}^{t_1}\delta\vartheta\hat{\mathbf{T}}_\perp^\kappa\delta\vartheta dt + \int_{t_0}^{t_1} j(t)\delta\vartheta dt}\mathcal{D}[\delta\vartheta] \\ &= \int_{(0,t_0)}^{(0,t_1)} e^{-\frac{\beta\sqrt{AF}}{2}\int_{t_0}^{t_1}\left(\delta\dot{\vartheta}^2+\left(1-\frac{6}{\cosh^2(t)}+\frac{\kappa}{F}\right)\delta\vartheta^2\right)dt+\int_{t_0}^{t_1} j(t)\delta\vartheta dt}\mathcal{D}[\delta\vartheta]\end{aligned} \tag{2.128}$$

with the (asymptotic) boundary conditions

$$t_{1/0} = \pm\frac{L}{2\lambda} \to \pm\infty \tag{2.129}$$

The operator $\hat{\mathbf{T}}_\perp^\kappa$ is again given by Eq. 2.115.

$$\hat{\mathbf{T}}_\perp^\kappa = \beta\sqrt{AF}\left(-\frac{\partial^2}{\partial t^2} + 1 - \frac{6}{\cosh^2(t)} + \frac{\kappa}{F}\right) \tag{2.130}$$

The source term is here given by

$$j(t) = -ip\left(H(t+t_c) - H(t-t_c)\right) \tag{2.131}$$

with $H(t)$ the Heaviside step function, p and t_c (≈ 1.915) constants. Note that the path integral in Eq. 2.128 is just the same as that in Eq. 2.123 after the usual "Wick rotation" $\tau \to it$ and the replacement $1/\hbar \to \beta$, $m \to \sqrt{AF}$, $\omega^2(t) \to \left(1 - 6/\cosh^2(t) + \kappa/F\right)$. The Euler-Lagrange equation giving the functional extremum of the exponent in Eq. 2.128 corresponding to the Eq. 2.127 reads

$$-\beta\sqrt{AF}\delta\ddot{\vartheta}_{cl} + \beta\sqrt{AF}\left(1 - \frac{6}{\cosh^2(t)} + \frac{\kappa}{F}\right)\delta\vartheta_{cl} - j(t) = 0$$

or written in terms of $\hat{\mathbf{T}}_{\perp}^{\kappa}$

$$\underbrace{\left(-\frac{\partial^2}{\partial t^2} + 1 - \frac{6}{\cosh^2(t)} + \frac{\kappa}{F}\right)}_{=\left(\beta\sqrt{AF}\right)^{-1}\hat{\mathbf{T}}_{\perp}^{\kappa}}\delta\vartheta_{cl}(t) = \left(\beta\sqrt{AF}\right)^{-1} j(t) \tag{2.132}$$

With boundary conditions $\delta\vartheta_{cl}(\pm\infty) = 0$. To solve this inhomogeneous Eq. 2.132 we construct the Green's function[29] $G(t,t')$ which is the solution to

$$\left(\beta\sqrt{AF}\right)^{-1}\hat{\mathbf{T}}_{\perp}^{\kappa} G(t,t') = \delta(t-t') \tag{2.133}$$

with $G(t,t') = G(t',t)$ and proper boundary conditions $G(\pm\infty,t') = 0$. The latter is very useful as it immediately gives the solution to Eq. 2.132 via the simple convolution

$$\delta\vartheta_{cl}(t) = \left(\beta\sqrt{AF}\right)^{-1}\int_{-\infty}^{\infty} G(t,t')\, j(t')\, dt \tag{2.134}$$

For our Dirichilet boundary conditions the Green's function generally writes[30] :

$$G(t,t') = -\frac{H(t'-t)\, f_2(t')\, f_1(t) + H(t-t')\, f_1(t')\, f_2(t)}{W} \tag{2.135}$$

with f_1 two f_2 being two (arbitrary) linearly independent solutions to the homogeneous equation $\left(\beta\sqrt{AF}\right)^{-1}\hat{\mathbf{T}}_{\perp}^{\kappa} f = 0$ satisfying the (one sided!) boundary conditions $f_1(-\infty) = 0$ and $f_2(\infty) = 0$ respectively. The constant W is the Wronski determinant of the two solutions i.e.

$$W = f_1(t)\,\dot{f}_2(t) - \dot{f}_1(t)\, f_2(t) = const. \tag{2.136}$$

[29]This will work without any problems if the homogeneous equation $\hat{\mathbf{T}}_{\perp}^{\kappa}\delta\vartheta = 0$ (with $\delta\vartheta$ vanishing at **both** boundaries) possesses **no nontrivial solutions**. Such zero eigenmodes of $\hat{\mathbf{T}}_{\perp}^{\kappa}$ are excluded if κ is slightly larger than 0. This is the main idea behind taking $\kappa > 0$ instead of the deceptively harmless $\kappa = 0$.

[30]This solution is valid if the prefactor of the leading order operator (i.e. of "$-d^2/dt^2$") is normalized to 1. This is the case for $\left(\beta\sqrt{AF}\right)^{-1}\hat{\mathbf{T}}_{\perp}^{\kappa} = -d^2/dt^2 + ...$ For a general exposition of Green's functions cf. for instance the nice book by Barton [20].

Now, we are already in possession of the two solutions (set $C_2 = 0$ and $C_1 = 0$ respectively in Eq. 2.120). They write

$$\begin{aligned} f_1(t) &= e^{ct}\left((c-2\tanh t)(c-\tanh t)-\cosh^{-2}(t)\right) \\ f_2(t) &= e^{-ct}\left((c+2\tanh t)(c+\tanh t)-\cosh^{-2}(t)\right) \end{aligned} \tag{2.137}$$

with

$$c=\sqrt{1+\frac{\kappa}{F}}$$

Their Wronskian 2.136 is given after short computation by

$$W=-2c\left(c^2-2\right)\left(c^2-1\right)$$

Inserting that and Eq. 2.137 into Eq. 2.135 gives a lengthy expression for $G(t,t')$. Fortunately there is no need for writing out explicitly neither $G(t,t')$ nor $\delta\vartheta_{cl}(t)$ as we are only interested in $S[j,x_{cl}]$ from Eq. 2.126 (with $x_{cl}=\delta\vartheta_{cl}$). In our case the classical action $S[j,\delta\vartheta_{cl}]$ writes

$$\beta S[j,\delta\vartheta_{cl}] = \underbrace{\frac{\beta\sqrt{AF}}{2}\delta\vartheta_{cl}(t)\,\delta\dot{\vartheta}_{cl}(t)\,|_{-\infty}^{\infty}}_{=0} - \frac{1}{2}\int_{-\infty}^{\infty} j(t)\,\delta\vartheta_{cl}(t)\,dt = \tag{2.138}$$

$$= -\frac{1}{2\beta\sqrt{AF}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} j(t)\,G(t,t')\,j(t')\,dtdt' \tag{2.139}$$

$$= \frac{p^2}{2\beta\sqrt{AF}}\int_{-t_c}^{t_c}\int_{-t_c}^{t_c} G(t,t')\,dtdt' \tag{2.140}$$

In the second line we used Eq. 2.134 and in the third line we exploited that $j(t)$ is proportional to the characteristic function of the interval $[-t_c,t_c]$, cf. Eq. 2.131. Now we can exploit that $f_2(t)=f_1(-t)$ and simplify the involved integral

$$\begin{aligned} \beta S[j,\delta\vartheta_{cl}] &= \frac{\left(\beta\sqrt{AF}\right)^{-1}p^2}{4c\left(c^2-2\right)\left(c^2-1\right)}\int_{-t_c}^{t_c}\int_{-t_c}^{t_c}\underbrace{\left(H(t'-t)\,f_2(t')\,f_1(t)+H(t-t')\,f_1(t')\,f_2(t)\right)}_{=(H(t'-t)f_1(-t')f_1(t)+H(t-t')f_1(t')f_1(-t))}dtdt' \\ &= \frac{\left(\beta\sqrt{AF}\right)^{-1}p^2}{4c\left(c^2-2\right)\left(c^2-1\right)}\int_{-t_c}^{t_c}\int_{-t_c}^{t_c} 2H(t-t')\,f_1(t')\,f_1(-t)\,dtdt' \\ &= \frac{\left(\beta\sqrt{AF}\right)^{-1}p^2}{2c\left(c^2-2\right)\left(c^2-1\right)}\underbrace{\int_{-t_c}^{t_c} f_1(t')\left(\int_{t'}^{t_c} f_1(-t)\,dt\right)dt'}_{I(c,t_c)} \end{aligned}$$

The involved double integral $I(c,t_c)$ in the last line will depend on the variable $c=\sqrt{1+\frac{\kappa}{F}}$ and the numerical constant t_c in a complicated manner. But luckily we are

only interested in the case $\kappa \to 0$ i.e. $c \to 1$. The expansion of the integrand around $c = 1$ (to lowest order) followed by the double integration gives

$$\begin{aligned} I(c, t_c) &= -6(c-1)\left(2t_c + 3\frac{t_c}{\cosh^2 t_c} - 5\tanh t_c\right) + O\left((c-1)^2\right) \\ &= \underbrace{\frac{3}{2}t_c\left(3t_c^2 - 10\right)}_{\approx 2.88}(c-1) + O\left((c-1)^2\right) \end{aligned}$$

In the second line we merely exploited the definition of t_c namely $t_c = 2\tanh t_c$. The limit $c \to 1$ i.e. $\kappa \to 0$ can now be performed safely and the action $S[j, \delta\vartheta_{cl}]$ writes

$$\begin{aligned} \lim_{\kappa\to 0} \beta S[j, \delta\vartheta_{cl}] &= \lim_{c\to 1}\left\{\frac{\left(\beta\sqrt{AF}\right)^{-1} p^2}{2c\left(c^2-2\right)\left(c^2-1\right)}\frac{3}{2}t_c\left(3t_c^2-10\right)(c-1)\right\} \\ &= -\frac{3}{8}\frac{\left(3t_c^2-10\right)t_c}{\beta\sqrt{AF}}p^2 \\ &\approx -\frac{0.72}{\beta\sqrt{AF}}p^2 \end{aligned} \tag{2.141}$$

Now the Eq. 2.124 in our case (after "Wick rotation" $\tau \to -it$ and the replacement $1/\hbar \to \beta$, $m \to \sqrt{AF}$, $\omega^2(t) \to \left(1 - 6/\cosh^2(t) + \kappa/F\right)$ etc.) writes

$$\widehat{Q}_\perp(p) = \underbrace{\sqrt{\frac{\beta\sqrt{AF}}{2\pi D(t_1, t_0)}}}_{\text{fluctuation contribution}} \underbrace{e^{-\beta S[j, x_{cl}]}}_{\text{classical contribution}} \tag{2.142}$$

The fluctuation determinant and the "fluctuation contribution" (which both do **not** depend on $j(t)$ and the parameter p) we already computed before in Eqs. 2.121 and 2.122. In the case of the "fluctuation contribution" (coinciding with $Q_\perp^\kappa$ in previous appendix) we cannot carelessly go to the limit[31] $\kappa \to 0$ and $t_{1/0} = \pm\frac{L}{2\lambda} \to \pm\infty$. The reason is again the diverging entropic contribution of a zero eigenmode. In the case $\kappa = 0$ (i.e. $c = 1$) this mode is just the rotational mode of the DNA configuration (i.e. the ϑ_0 in Eq. 2.71). To cure the problem we have to divide out the wrong Gaussian contribution and multiply the real "state-space volume" corresponding to this mode, similarly as we did this in Eq. 2.51 for the translational mode. Compiling Eqs.2.142, 2.141 and 2.121 together with $\lambda = \sqrt{A/F}$, $l_P = \beta A$ we obtain

$$\widehat{Q}_\perp(p) = \sqrt{\frac{l_P}{\lambda}\frac{c(c+2)(c+1)}{\pi(c-2)(c-1)}}e^{-c\frac{L}{2\lambda}}e^{\frac{3}{8}\frac{\lambda}{l_P}\left(3t_c^2-10\right)t_c p^2}$$

We are almost done besides that we have to perform the "proper renormalization" of $\widehat{Q}_\perp(p)$ mentioned in the main text. More concretely we have to correct for the (unphysically) diverging contribution of the rotational mode $\vartheta_0(t) = \sqrt{3/2}\sinh(t)\cosh^{-2}(t)$

[31] In contrast to the case of the more tame "classical contribution" where both limits cause no problems. Therefore for the "classical contribution" we go to the limit $\kappa \to 0$ ($c \to 1$) immediately.

from Eq. 2.71. The eigenvalue $\mu_0(\kappa)$ of $\hat{\mathbf{T}}_\perp^\kappa$ corresponding to ϑ_0 is easily seen to be

$$\mu_0(\kappa) = \beta\lambda\kappa = \frac{l_P}{\lambda}(c-1)(c+1)$$

This eigenvalue (artificially) enters the partition function $\widehat{Q}_\perp(p)$ through a Gaussian integral which we need to divide out and replace with a 2π - the state space volume of the rotational mode. After doing that we are on safe shores and we can switch off the formally introduced "magnetic field" (i.e. we can take the limit $\kappa \to 0$, or $c \to 1$)

$$\begin{aligned}\widehat{Q}_\perp^{ren}(p) &= \lim_{c\to 1}\left(\frac{2\pi}{\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\mu_0 x^2}dx}\widehat{Q}_\perp(p)\right) = \lim_{c\to 1}\left(\sqrt{2\pi\mu_0}\widehat{Q}_\perp(p)\right) \\ &= \lim_{c\to 1}\sqrt{\left(\frac{l_P}{\lambda}\right)^2\frac{2c(c+2)(c+1)^2}{(c-2)}}e^{-c\frac{L}{2\lambda}}e^{\frac{3}{8}\frac{\lambda}{l_P}\left(3t_c^2-10\right)t_c p^2}\end{aligned}$$

The limit now is well behaved and we obtain the final result

$$\widehat{Q}_\perp^{ren}(p) = i2\sqrt{6}\left(\frac{l_P}{\lambda}\right)e^{-\frac{L}{2\lambda}}e^{\frac{3}{8}\frac{\lambda}{l_P}\left(3t_c^2-10\right)t_c p^2} \tag{2.143}$$

This is a simple "inverted" (positive exponent) Gaussian!

2.4.5 Appendix E: Analogies with other Systems, Kinks, Instantons, Bubbles

Whenever one obtains some equations that look complicated but admit simple solutions one can make an almost sure bet that (in the best case) some or (in worst case) a whole community of people uses and solves the same in some other context. This is not much different in our case. Our problem possesses strong analogies in many branches of physics. This of course not the place to work them out here - we merely bring them to our attention:

1. **Nucleation of critical droplets ("bubbles")** in overheated liquids and similar problems. The classical reference on this topic is Langer [15] (cf. also [13, 14]) who treats for the first time the general Kramers (classical) activated barrier crossing in an infinitely dimensional system.

2. **Kinks in (overdamped) soliton bearing systems** like solids with dislocations (or Josephson junctions, cf. [29]). A classical reference on that is the work of Büttiker and Landauer [16] where they treat dilute soliton gases. A "kink" in their work essentially corresponds to a 2-D loop here and the corresponding partition functions can be mapped onto each other.

3. **Quantum mechanical tunneling** in single and double well (cubic / quartic) potentials of all kinds. The main issue here is the evaluation of a transition

amplitude (the propagator) $(x_1, t_1|x_0, t_0)$ of the kind

$$(x_1, t_1|x_0, t_0)_V = \int_{(x_0,t_0)}^{(x_1,t_1)} \mathcal{D}[x]\, e^{\frac{i}{\hbar}\mathcal{A}[x]} \text{ with the action}$$

$$\mathcal{A}[x] = \int_{t_0}^{t_1} \left(\frac{1}{2}M\dot{x}^2 - V(x)\right) dt$$

for a potential of the form $V(x) = a_4x^4 + ... + a_1x$ (a_4 can be zero) and the particle mass M. If the involved action integral $\mathcal{A}$ happens to be $\gg \hbar$ than most of the particle paths cancel (by destructive interference) and only the classical path together with paths close to this classical trajectory remain. The classical trajectory will be given by the Euler-Lagrange equation $\delta\mathcal{A}[x_{cl}] = 0$ which leads to

$$\ddot{x}_{cl}(t) = -\frac{1}{M}\left(4a_4x_{cl}^3(t) + ... + a_1\right) \tag{2.144}$$

Such an equation is generally solved by elliptic functions and in most interesting limiting cases one usually recovers some reparametrization of a kink (loop) solutions obtained in the text. The contribution of quantum fluctuations around a "tunnelling path"[32] x_{cl} between the two valleys of the double well potential are evaluated in an analogous manner (within the saddle point approximation) as we did for thermal undulations of the DNA chain around the loop solution. This is not surprising as quantum fluctuations are formally equivalent to thermal fluctuations in imaginary time. For an introduction to this topic cf. refs. [13, 14] and refs. therein.

4. **Flexible $\leftrightarrow$ semiflexible-polymer analogy**. This analogy is particularly beautiful and very visual. To my best knowledge it has not been worked out properly in literature. Take a flexible Gaussian polymer chain and put it an external potential $V(x)$. This classical problem was treated in exquisite detail in the past century [23, 22]. For a typical application of that system in the context of flexible polymer activated barrier crossing (polymer through a pore problem) cf. for instance ref. [24]. The free - energy functional of a chain configuration $\underline{x}(n)$ parametrized by the bead number n is given by

$$F[\underline{x}] = \int_0^N \left(\frac{a}{2}\dot{\underline{x}}^2(n) + V(\underline{x}(n))\right) dn$$

with the entropic spring constant a proportional to temperature T. Take now $\underline{x}(n)$ to be on a sphere of radius 1 (say a flexible polymer adsorbed on a sphere). Then the squared derivative $\dot{x}^2 = (d\underline{x}/dn)^2$ writes in the spherical coordinates

$$\dot{\underline{x}}^2 = \dot{\phi}^2(n)\sin^2\theta(n) + \dot{\theta}^2(n)$$

where $\theta(n)$ and $\phi(n)$ are the corresponding spherical coordinates of the polymer path. If we now subject the chain to a potential

$$V(n) = -F\cos\theta(n) \tag{2.145}$$

[32] The tunnelling path x_{cl} is sometimes also called the "instanton" because it jumps within a very short time period (almost instantaneously) from one potential valley to the other.

we obtain a mapping to the wormlike chain bending and potential energy (cf. Eq. 2.74). The constant a becomes now the stiffness A. So we can state that the unit (!) tangent of a semiflexible chain can be mapped onto the real - space position vector $\underline{x}(n)$ of completely flexible chain on a sphere subjected to the potential Eq. 2.145. In Fig. 2.13 we represent the magnitude of different acting ("tangent space") potentials as distance from the origin so that the potential energy can be represented as a spherical landscape. Note that by introducing these hills and valleys we do not really mean that the metrics (line elements, areas) on the sphere is modified (it is assumed to be constant throughout the sphere) - it is done only for the sake of visualization that otherwise would be difficult.

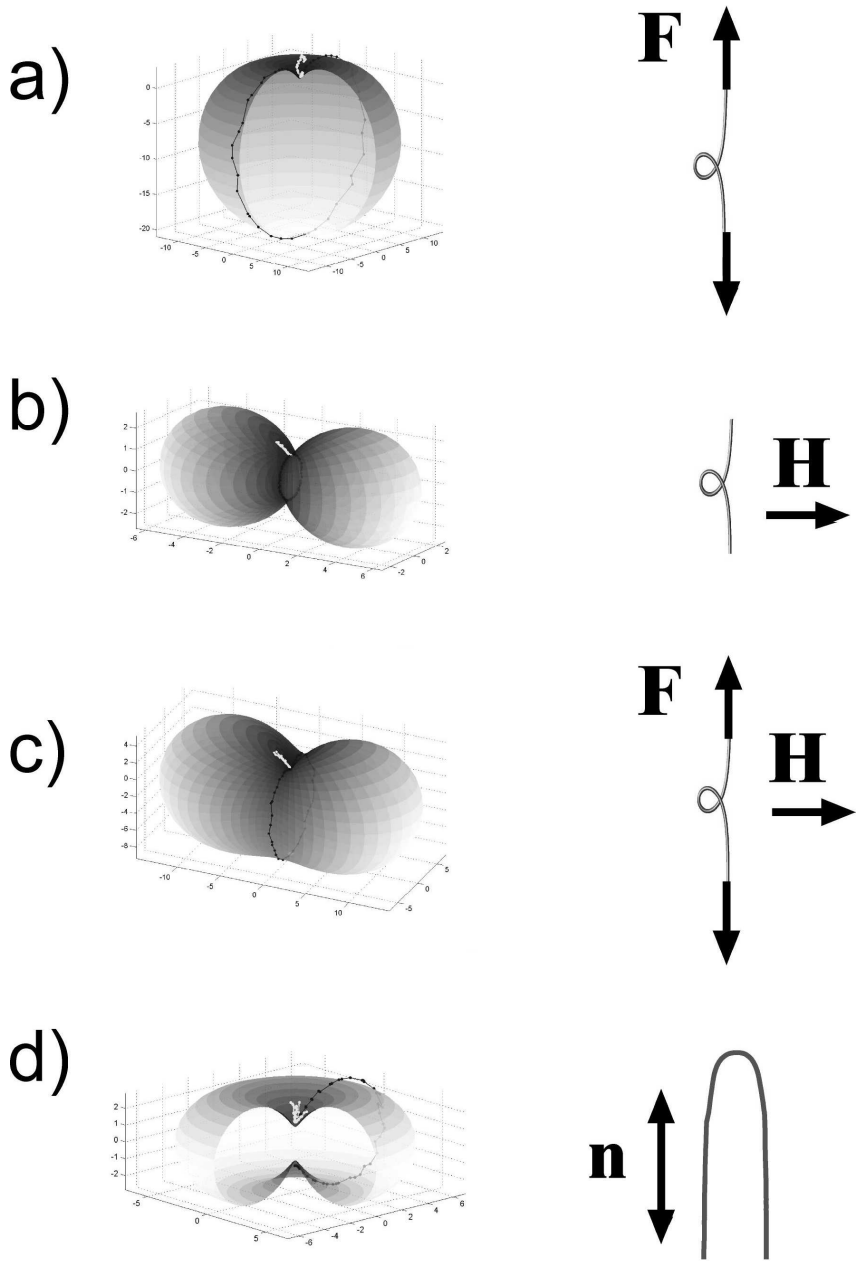


Figure 2.13: The analogy between a flexible polymer on sphere and a semiflexible polymer. The potential energy landscape acting on the flexible polymer (left) is represented by the distance from the coordinate origin. The depicted potentials correspond in the semiflexible case (right) to the action of a) tension F (the "apple"), b) magnetic field H (the "dumb-bell"), c) both F and H (the "peanut") and d) a nematic field n (the "red-blood-cell"). The black coils (left) correspond to the semiflexible excitations (loops,hairpins on right), the white coils are small deviations from straight "rods".

Chapter 3

The Thermal Motion of Nucleosomes

The genetic information of all higher organisms is organized in huge beads-on-a-chain arrays consisting of centimeters to meters of DNA wrapped around globular aggregates of so-called histone proteins (cf. the introductory chapter). The basic unit of chromatin, the nucleosome, is a tiny 10×6 nm sized spool composed of 147 base pairs (bps) DNA tightly wrapped around an octamer made from 8 histone proteins. Each nucleosome is connected via a stretch of "linker" DNA to the next such protein spool. The wrapped DNA, being coiled in $\sim 1\frac{3}{4}$ turns of a left handed helix with radius ~ 4.2 nm, is strongly distorted from its preferred straight ground state due to strong interactions with the histone octamer, namely short range electrostatics (between the negatively charged DNA sugar-phosphate backbone and the positively charged octamer surface) and through extensive hydrogen-bonding – both localized at 14 discrete interaction patches helically arranged along the octamer surface [126] (cf. the picture gallery in chapter 1).

Higher order structures, from the 30nm-chromatin fiber up to the highest level of DNA condensation, the fully folded chromosome, are designed to achieve a huge DNA volume fraction. They all rely on the significant stability of the nucleosome complex. On the other hand, fundamental life processes like transcription (making RNA offprints from the underlying DNA) and DNA replication seem to be in conflict with the picture of a stable nucleosome, as they are all performed by protein machines that track the DNA helix. The latter inevitably implies that every DNA bound obstacle (protein) has to be penetrated or even completely removed from its DNA target. In fact, the numbers are quite dramatic: A typical gene extends over hundreds of nucleosomes, each contributing $30 - 40\ k_BT$ net adsorption energy [140, 136]. Also other mechanisms like the gene activation rely on regulatory protein binding to specific DNA sequences that are often covered by nucleosomes making them inaccessible.

A key to the understanding of these seemingly contradictory features might be the *physical* phenomenon of thermally driven nucleosome "sliding" along DNA (also called nucleosome repositioning) which has repeatedly observed in well-defined *in vitro* experiments [36, 103, 38]. Spontaneous repositioning is strongly temperature dependent; at room temperature nucleosomes move a few tens of bps within an hour. Despite clear

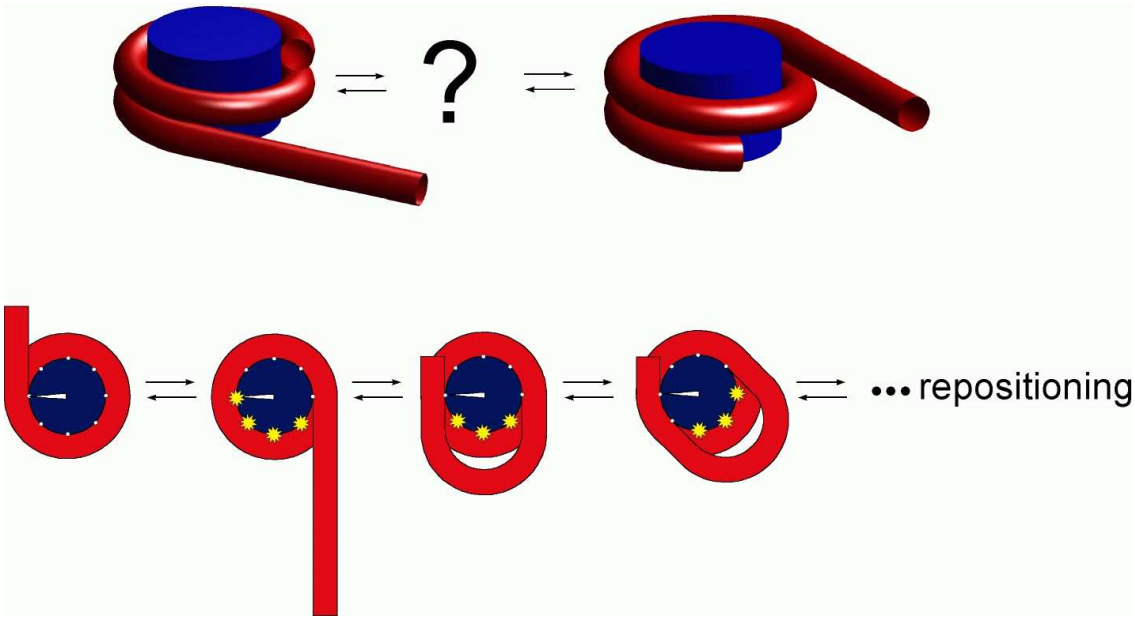


Figure 3.1: The basic problem setting: how does the histone-octamer move along the DNA template? Below: the DNA loop mechanism as proposed in in Ref. [69]

evidence for repositioning the underlying mechanism has been the matter of longstanding controversy, especially due to the lack of any quantitative theoretical treatment of nucleosome statics and dynamics that has to rely on the detailed knowledge of the molecular structure and its underlying parameters.
In this chapter we theoretically consider two distinct models that both could account for many experimental observations and draw conclusions and theoretical predictions that are intrinsically linked to both of them respectively. Finally in the last section of this chapter we will discuss and theoretically model some long awaited and very recent experimental findings. The result of this modelling will finally shed a surprisingly clear light on the underlying mechanism.

3.1 Repositioning via Loop Formation

The first explanation which appears to be consistent with the discrete nucleosome "jumps" (10 bp quantized) and the large barriers observed by Pennings et. al. has been recently proposed in Ref. [69]. In this model the basic step in the repositioning process is a partial unwrapping of DNA from the very ends of the nucleosome [70, 71] followed by a backfolding of DNA with a small 10 bp mismatch (cf. Fig. 3.1). The result of this process is the formation of a small DNA bulge or loop on the octamer surface. Once trapped on the nucleosome surface this small defect carrying some discrete quantum of DNA extra length (a multiple of 10 bp - the DNA helical repeat length) can propagate by diffusion in both directions. If the loop happens to surround the nucleosome and comes out at the opposite side (in respect to where it was created) the nucleosome is eventually repositioned by a distance given by the "pulled in" extra

length. The energetic barrier and rates of repositioning were computed [69, 111] and were shown to be consistent with the Pennings et al. experiment [67, 68]. Moreover, the 10 bp discrete step repositioning observed in the experiment (discrete bands, no 1 bp spaced intermediates) came out as a natural consequence of the loop length quantization. The latter is enforced by the strongly preferred DNA minor groove - octamer interaction and the discrete binding sites at the nucleosome surface as deduced from the crystallographic structures [73].
In Ref. [69] small loops with short excess length of typically $\sim 1 - 2{\times}10$ bp were considered and it was shown that the looping energies involved increase rapidly with the excess length implying that only the shortest (10 bp) loop contributes significantly to the repositioning mechanism. Consequently the model predicts a classical discrete random walk with a jump-size of 10 bp – instead of a 1 bp motion that would be implied by a corkscrew motion considered later in this chapter (cf. section 3.2, Fig. 3.11). Apart from the discrepancy in the elementary step size, both models predict very similar behavior: a local one-dimensional diffusive motion along the DNA chain.
In this section we will carefully reanalyze the idea of loop-mediated repositioning by applying the classical tool of the Kirchhoff kinetic analogy (cf. chapter 2) which provides us with analytic solutions of the loop problem and enables us to look at loops of virtually any given excess length. The main outcome of our study will be a different picture of repositioning which physically results from the looping mechanism: on short and moderately long segments of up to 2-3$\times$ l_P (l_P: DNA persistence length) the repositioning is a jumpy process with largest possible loops being the most dominant ones in contrast to short 10 bp steps as conjectured before. For longer and very long (infinite) DNA segments there is an optimal jump size of order $\sim O(l_p)$ and the behavior is superdiffusive in contrast to the previously predicted diffusive mechanism. As we will see below, these predictions allow us to clearly distinguish between different repositioning mechanisms in experiments expected to be performed in near future.

3.1.1 Energetics of Loops

Let us now consider the energetics of an intranucleosomal DNA loop. We will describe it within the framework of the Euler-Kirchhoff theory for the static equilibrium of rods that we already used in chapter 2, cf. Fig. 3.2. For simplicity and because of the approximate planarity of the problem we can in first approximation assume the nucleosome and the loop-forming DNA to be in one plane and the DNA to be free of any twisting deformation. It this case the internal energy of our system is simply divided into two components, the planar elastic DNA-bending and a histone-octamer DNA interaction:

$$U_{tot} = U_{bend} + U_{ads} \tag{3.1}$$

The bending energy (within the linear elasticity approximation) can be written in terms of the local DNA curvature κ

$$U_{bend} = \frac{A}{2} \int_{-L/2}^{L/2} \kappa^2 (s)\, ds \tag{3.2}$$

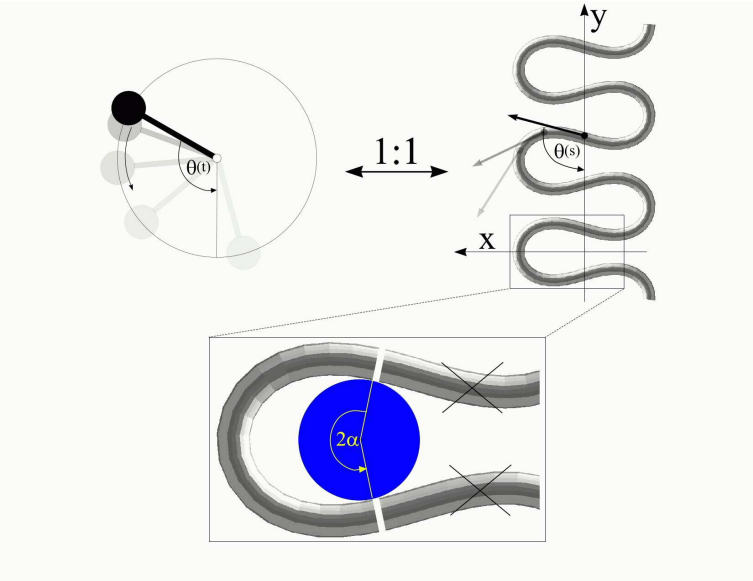


Figure 3.2: The Kirchhoff kinetic analogy between the spinning top and the bent/twisted rod depicted for a special case: the plane pendulum - planar rod equivalence. The inset shows how an intranucleosomal loop can be constructed by inscribing the octamer (gray disk) into the bent rod. The nucleosome opening angle 2α accounts for the adsorption energy cost (see text for details).

with $A \approx 50\ nm \cdot k_B T$ being the bending rigidity of DNA at room temperature and physiological salt concentrations [74]. As in chapter 2 the DNA is assumed to be parametrized by its contour length parameter s ranging from $-L/2$ to $L/2$ with L being the total length of the loop. The latter can be expressed in terms of two independent quantities: the excess length ΔL and the nucleosome opening angle α (Fig. 3.2)

$$L(\alpha, \Delta L) = 2\alpha R + \Delta L \tag{3.3}$$

where $R \approx 4$ nm is the effective nucleosome radius, or more precisely the distance from the center of the nucleosome to the central DNA axis. Because the DNA can enter the nucleosome only in quantized orientations (with its minor groove phosphates) and bind only to discrete positions on the protein surface [73] (cf. the gallery of pictures at the beginning of the thesis), the excess length $\Delta L = n \times h_{DNA}$ is to a good approximation an integer multiple of the DNA repeat length $h_{DNA} = 3.4nm$.
The second part in the total energy Eq. 3.1 U_{ads} comes from the (predominantly electrostatic) interaction between the positively charged protein surface and the negatively charged DNA. It can be roughly measured from experiments probing the competitive protein binding to nucleosomal DNA [70, 71]. Neglecting the discreteness of charges (binding sites) on the histone octamer surface it can in first approximation be assumed to be proportional to the opening angle α and the *effective*[1] adsorption energy density

[1]By "effective" we mean here the adsorption energy density *renormalized* by the bending energy density on the octamer surface (with a radius of curvature $R = 4nm$). That means: ε_{ads}= (interaction energy per length) - (bending energy per length).

ε_{ads}

$$U_{ads} = 2\alpha R \varepsilon_{ads} \tag{3.4}$$

with $\varepsilon_{ads} \approx 0.5 - 1.0\ k_B T/nm$ as roughly extracted from [70] [2]. Here and in the following we assume an intermediate value of $\varepsilon_{ads} = 0.7\ k_B T/nm$.

3.1.2 Ground States of Trapped Loops

In order to compute the ground state for a trapped intranucleosomal loop we have to consider shapes that minimize the total energy 3.1 under two constraints:

1. The excess length ΔL is prescribed. Therefore we have the relation Eq. 3.3 between the opening angle and the total loop length L

$$\Delta L = L - 2\alpha R = const. \tag{3.5}$$

2. At the two ends $s_\pm = \pm L/2$ the rod has to be tangential on an inscribed circle of given radius (representing the nucleosome) [3]:

$$R = \left| \frac{y\left(\frac{L}{2}\right)}{-x'\left(\frac{L}{2}\right)} \right| = const. \tag{3.6}$$

Here $x(s)$ and $y(s)$ are the Cartesian coordinates of the rod axis as a function of the arc-length parameter s (cf. Fig. 3.2). The absolute value in the second constraint needs to be introduced formally for dealing with crossed rod solutions (which we consider later on) and can be omitted for simple uncrossed loops.
For an analytical description it is convenient to use the angle $\theta = \theta(s)$ between the DNA tangent and the y axis as a variable describing the DNA centerline. In this case the integrated sine (cosine) of θ over the arc-length parameter s gives the x (y) Cartesian coordinate of any point along the rod, and the squared derivative $(\theta')^2$ gives the rod curvature κ. Furthermore the nucleosome opening angle α is simply related to θ at the boundary

$$\alpha = \begin{cases} \theta(L/2) & \text{for simple loops} \\ \pi - \theta(L/2) & \text{for crossed loops} \end{cases}$$

The two constraints Eq. 3.5 and Eq. 3.6 can be rewritten in terms of θ and then be introduced into the minimization by two Lagrange multipliers $\mu_{1/2}$. We then arrive at the following functional

$$\begin{aligned} \widehat{U}_{tot} \;=\; & A \int_0^{L/2} (\theta')^2 \, ds \; + 2\alpha R \varepsilon_{ads} \\ & + \mu_1 \left[L - (\Delta L + 2\alpha R) \right] \\ & + \mu_2 \left[\int_0^{L/2} \cos\theta ds - R \sin\alpha \right] \end{aligned} \tag{3.7}$$

[2] In Eq. 3.4 we assume that the interaction is only short ranged (contact interaction) which is justified by the very short Debye screening length of $\approx 1nm$ under physiological salt conditions.

[3] Because of the symmetry we have to impose the conditions only on one side.

Here the first line is the bending + adsorption energy contribution, the second and third line are the imposed length and tangency constraint. Eq.3.7 can be rearranged in a more familiar form

$$\int_0^{L/2} \left(A\left(\theta'\right)^2 + \mu_2 \cos\theta \right) ds + \text{b.t.} \tag{3.8}$$

Here b.t. denotes the boundary terms (depending on $\theta\left(L/2\right)$ only) that obviously do not contribute to the first variation inside the relevant s interval. The integral in Eq.3.8 is evidently analogous to the action integral of the plane pendulum with $A\left(\theta'\right)^2$ corresponding to the kinetic and $-\mu_2 \cos\theta$ to the potential energy of the pendulum. This was explained in the chapter 2 in the context of the Kirchhoff kinetic analogy.
The nice thing about Kirchhoff's analogy apart from its esthetic content is that it provides us with explicit expressions for DNA shapes subjected to twist, bending and various geometric topological constraints. In our simple planar and twistless case, the "spinning top" simply reduces to the plane pendulum. The corresponding planar and twistless rods, also called the *Euler elastica*, are most generally given by

$$\cos\theta\left(s\right) = 1 - 2m\mathrm{sn}^2\left(\frac{s}{\lambda} \mid m\right) \tag{3.9}$$

which can be integrated to obtain the general planar rod shape in Cartesian coordinates:

$$x\left(s\right) = 2\sqrt{m}\lambda\,\mathrm{cn}\left(\frac{s}{\lambda} \mid m\right) \tag{3.10}$$

$$y\left(s\right) = 2\lambda E\left(\frac{s}{\lambda} \mid m\right) - s \tag{3.11}$$

with sn, dn, cn$(. \mid m)$ being the Jacobi elliptic functions with the parameter m and

$$E\left(u \mid m\right) := \int_0^u \mathrm{dn}^2\left(v \mid m\right) dv \tag{3.12}$$

denoting the incomplete elliptic integral of the second kind in its "practical" form[4]. The two parameters $m > 0$ and λ (the tension length, cf. chapter 2) in Eqs. 3.10 and 3.11 characterize the shape and the scale of the solution, respectively. These solutions are up to trivial plane rotations, translations, reflections and shifting of the contour parameter $s \to s + s_0$ the most general planar Euler elastica corresponding to the plane pendulum. For different parameters m one obtains different rod shapes corresponding to different solutions of the spinning top (plane pendulum) motion [84]. The case $m = 0$ describes a pendulum at rest corresponding to a straight rod, for $0 < m < 1$ one has strictly oscillating pendulums corresponding to point symmetric rod shapes which for $m < 0.92$ are free of self intersections like the one depicted in Fig. 3.2. For m higher than 0.92 the rods show varying complexity with a multitude of self-intersections and for $m = 1$ one has the so-called homoclinic pendulum orbit corresponding to a rod solution having only one self intersection and becoming asymptotically straight for $s \to \pm\infty$ (for details see Ref. [84]). For even higher values [5] of m, i.e., for $m \geq 1$ we

[4]Some useful formulas and relations for the elliptic functions and integrals are briefly sketched in [84] and found in [85] in full depth.

[5]Usually the parameter m is artificially assumed to be confined to $0 \leq m \leq 1$ but by the Jacobi's real transform for elliptic functions [85] they stay well-defined even for $m > 1$.

have revolving pendulum orbits corresponding to rods with self-intersections lacking point symmetry. Finally, the limiting case $m \to \infty$ corresponds to the circular rod shape.

In order to describe a trapped loop we need to use Eqs. 3.10 and 3.11 imposing the constraints Eqs. 3.5 and 3.6. It turns out to be more convenient to replace the parameter set (λ, m, L) with the new (but equivalent) set $(\lambda, m, \sigma := \frac{L}{2\lambda})$ where we introduced the new dimensionless parameter σ which we call the "contact parameter"[6]. From Eq. 3.6 together with 3.10 and 3.11 we can immediately extract the scaling parameter λ and the opening angle in terms of the contact parameter σ and the shape parameter m

$$\lambda(\sigma, m) = R\left|\frac{\mathrm{sn}(\sigma \mid m)\,\mathrm{dn}(\sigma \mid m)}{2E(\sigma \mid m) - \sigma}\right| \tag{3.13}$$

$$\alpha(\sigma, m) = \arccos\left[\pm\left(2\mathrm{dn}^2(\sigma \mid m) - 1\right)\right] \tag{3.14}$$

$$\pm := sign\left(2E(\sigma \mid m) - \sigma\right) \tag{3.15}$$

Plugging this into Eq. 3.5 we obtain the final form of the implicit constraint

$$\frac{\Delta L}{2R} = \sigma\left|\frac{\mathrm{sn}(\sigma \mid m)\,\mathrm{dn}(\sigma \mid m)}{2E(\sigma \mid m) - \sigma}\right| - \arccos\left[\pm\left(2\mathrm{dn}^2(\sigma \mid m) - 1\right)\right] \tag{3.16}$$

The curvature $\kappa(s)$ and the bending energy Eq. 3.2 follow from the explicit solution Eq. 3.9 to be

$$\kappa(s) = \frac{2\sqrt{m}}{\lambda}\mathrm{cn}\left(\frac{s}{\lambda} \mid m\right) \tag{3.17}$$

$$U_{bend} = \frac{4mA}{\lambda}\int_0^\sigma \mathrm{cn}^2(t \mid m)\,dt \tag{3.18}$$

$$= \frac{4A}{\lambda}\left[(m-1)\,\sigma + E(\sigma \mid m)\right] \tag{3.19}$$

The latter expression together with Eqs. 3.1, 3.4 - 3.15 gives a lengthy expression for the total energy with the sign chosen $\pm$ as in Eq. 3.15.

$$U_{tot}(\sigma, m) = \frac{4A}{R}\left|\frac{\left[2E(\sigma \mid m) - \sigma\right]\left[E(\sigma \mid m) + (m-1)\,\sigma\right]}{\mathrm{sn}(\sigma \mid m)\,\mathrm{dn}(\sigma \mid m)}\right| + 2R\varepsilon_{ads}\arccos\left[\pm\left(2\mathrm{dn}^2(\sigma \mid m) - 1\right)\right] \tag{3.20}$$

Now our problem of finding the ground state loop for given excess length ΔL reduces to a two variable (σ, m) minimization of Eq. 3.20 under the constraint Eq. 3.16. This final step is easily performed numerically.

3.1.3 Loop Zoology: Simple and Crossed Loops

We can scan now through the $\sigma - m$ parameter plane and look at the shapes of the solutions and their energies. In Fig. 3.3 we see a small (but most important) part of the

[6] A more visual parameter set (α, m, λ) using the opening angle $\alpha = \alpha(\sigma, m)$ produces technical problems with non-uniqueness of loop representation.

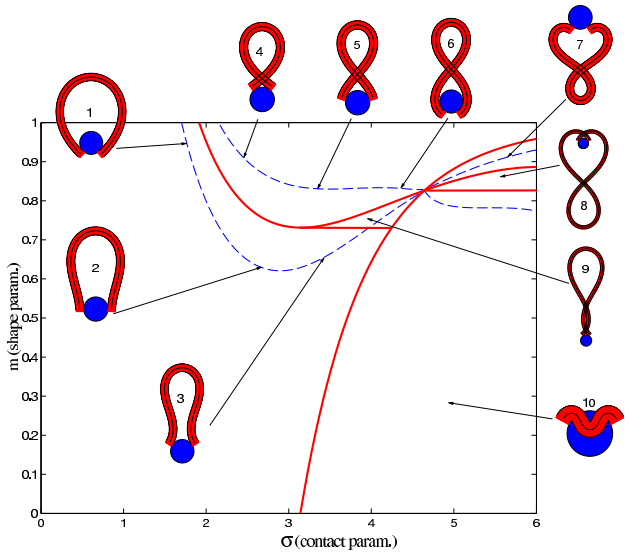


Figure 3.3: The set of possible ground-state solutions is characterized by two parameters, the contact point parameter σ and the loop shape parameter m. Solutions with constant excess length ΔL (here $10 \times 3.4nm$) are located along the dashed lines (e.g. loops 1-7). The solid lines separate loops with different geometric characteristics: simple (1,2,3), crossed (4,5,6) and "exotic" (7,8,9,10) loop shapes.

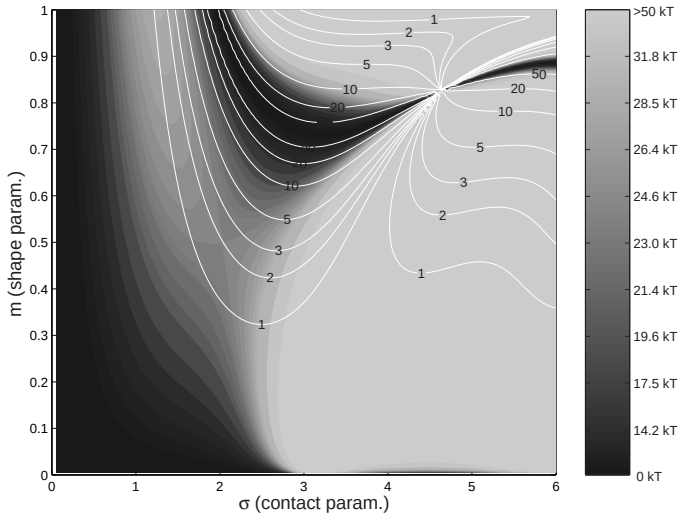


Figure 3.4: Density plot of the total loop energy Eq. 3.20 (grayscale levelsets) as a function of σ and m (same parameter range as in Fig. 3.3). The white contours denote lines of constant excess length $\Delta L = 1, 2, 3, 5, 10, 20, 50 \times 3.4$ nm. For given excess length the ground state is the point on the corresponding white line with the darkest background (note the different branches for given ΔL). The parameters are $\varepsilon_{ads} = 0.7 k_B T/nm$ and $A = 50nm \times k_B T$ and $R = 4nm$.

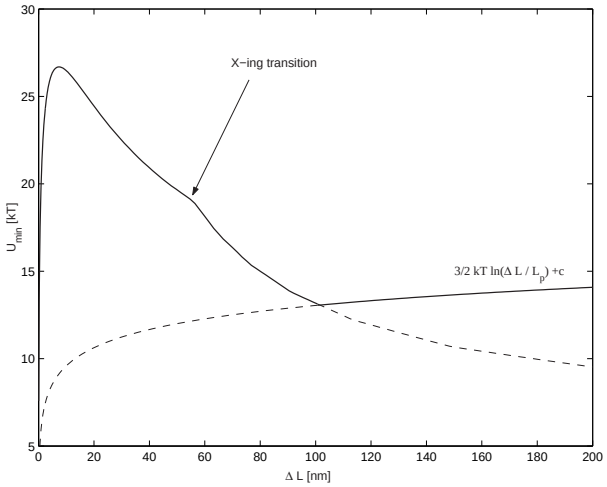


Figure 3.5: The ground state loop energy plotted vs the excess length ΔL. Note the energy maximum occurring for shorter loops. For much longer loops (around $\Delta L = 60nm$) a transition from simple uncrossed to crossed loop shapes occurs leading to a kink in $U_{\min}(\Delta L)$. In the regime of low $\Delta L \lesssim l_P$ the elastic energy prevails strongly over entropy whereas for large loops the entropy starts to dominate the behavior producing a shallow energy minimum in the cross-over regime which roughly defines the predominant loop size.

whole parameter space and the corresponding different loop geometries. The dashed lines indicate parameter values which lead to constant excess length $\Delta L = 10 \times 3.4nm$ (corresponding to 100 bps) in accordance with the constraint Eq. 3.16. The shapes 1-7 are examples of 100bp-loops with different geometries. The whole parameter plane is subdivided by separation lines (solid) into regions of structurally different solutions. The large region starting at $\sigma = 0$ contains exclusively simple loops (like 1,2 and 3) without self-intersections and nucleosome penetration. Above that simple-loop-region we find loops with a single self-intersection (4,5,6) and to the right the loops penetrate the nucleosome, like loop 10. There are also three other regions with single and double crossing points (7,8,9) where the loop can also be on the "wrong" side of the nucleosome like in 7 and 8.

We are interested in the energy minimizing loops and the underlying minimal energies as functions of the excess length ΔL. A density plot of these energies as function of the parameters σ and m together with the corresponding lines of constant ΔL (with $\Delta L = 1, 2, .., 50 \times 3.4nm$) is given in Fig. 3.4. As can be seen from Fig. 3.3 there are, for a given ΔL, different branches of (σ, m) values corresponding to uncrossed, simply crossed and other exotic structures. Of all these structures for short excess lengths, $\Delta L \lesssim 20 \times 3.4nm$, the energetically dominant ones are simple (uncrossed) loops which we study first. Loops with larger excess length form crossed structures and are studied later below.

3.1.4 Simple Loops

For simple uncrossed loops it is a straightforward numerical task to minimize Eq. 3.20 under the constraint of constant excess length, Eq. 3.16. For $\varepsilon_{ads} = 0.7 k_B T/nm$ and all the other parameters as above ($A = 50nm \times k_B T$, $R = 4nm$) the ground state energy U_{min} as a function[7] of the excess length ΔL is shown in Fig. 3.5 (for $\Delta L \lesssim 60nm$; for longer ΔL-values crossed loops are more favorable as discussed in the next section). Remarkably we find that the loop energy is non-monotonous: For small ΔL U_{min} increases with ΔL as $(\Delta L)^{1/3}$ (in accordance with Ref. [69] where only small loops were studied). At some critical excess length $\Delta L = \Delta L_{crit}$ (which is approximately $\Delta L_{crit} \approx 2.2 \times 3.4nm$ for $\varepsilon_{ads} = 0.7\ k_B T/nm$) the loop energy reaches a maximum (here $U_{min}(\Delta L_{crit}) \approx 26 k_B T$). Beyond that the energy decreases with increasing ΔL.
In the following we show how this behavior can be explained on the basis of the loop geometry. Naively one might argue as follows: For excess lengths shorter than the persistence length of DNA it is increasingly difficult to store additional length into the loop because it requires increasing DNA deformation. On the other hand, for loops longer than l_P the bending energy contribution becomes very small and hence one expects such ground state loops relaxing with increasing ΔL. However the reason for occurrence of a maximum of U_{min} around ≈ 2 excess DNA lengths, a value which is considerably smaller than the persistence length, is not obvious. In order to understand this finding one has to go beyond the simple handwaving heuristics and needs to take a close look at the details of the loop geometry.
To this end we introduce here a simple approximation technique which leads to explicit expressions which can be more easily handled than the exact yet complicated expressions given above. We call this method the *circle-line approximation* and give a detailed exposition in the Appendix. As we will see this method is quite accurate and at the same time very intuitive.
Looking at the geometrical shapes of the loops in Fig. 3.3 we notice that each of them is subdivided into several sections of very high and very low curvature (cf. also Eq.3.17). In first approximation we replace the high curvature regions by sections of circles, the low curvature regions by straight lines (cf. Fig. 3.6). Furthermore, to keep the smoothness we assume that the lines are tangents to the circles. Generally in order to have reasonable approximations of all possible loop shapes we would need to consider compositions of several circles and lines (cf. for instance loops 3, 6, 7). However, if the adsorption energies are not to high, i.e., if the opening angle α is "soft enough" and does not impose such a severe bending like in loop 3, such multiply bent loops will not be relevant as ground state solutions. As it turns out for our problem we already obtain a quite good approximation by assuming that the loop consists of a *single* circular arc and two lines only. It is characterized by two quantities : 1) the arc radius r and 2) the nucleosome opening angle α (cf. Fig. 3.6 and Appendix). With these assumptions and after some elementary geometry the constraint Eq.3.5 becomes simply

$$\Delta L = 2\,(R - r)\,(\tan\alpha - \alpha) = const. \tag{3.21}$$

[7] Formally the quantisation condition $\Delta L = 1, 2, ... \times 3.4nm$ holds as mentioned above. Nevertheless for clarity we consider the values in between as well.

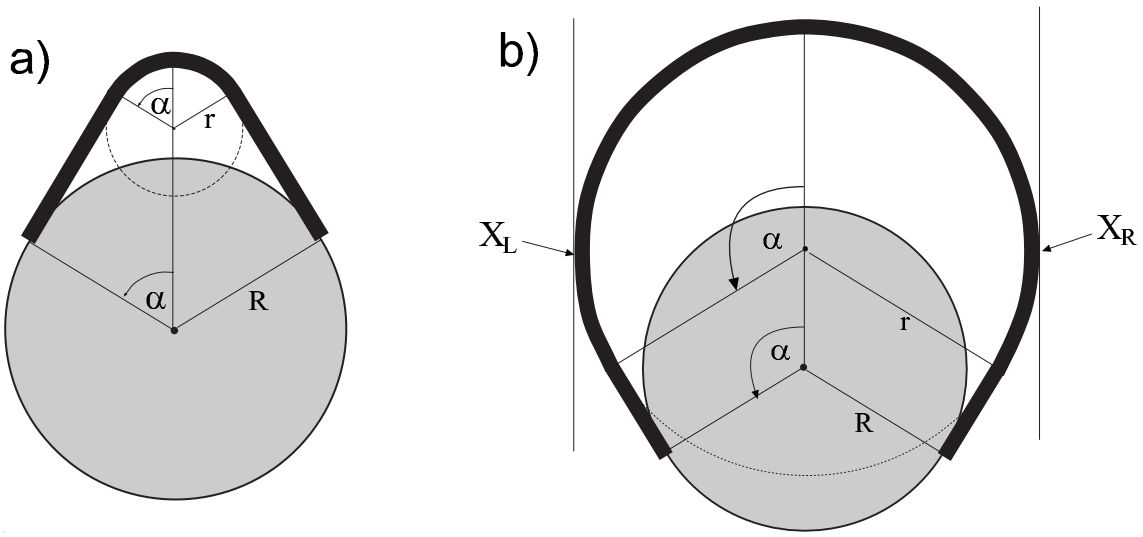


Figure 3.6: Two generic types of simple loop geometries (in the circle-line approximation): a) the subcritical loop with opening angle $\alpha < \pi/2$ and b) the supercritical loop with $\alpha > \pi/2$. In the former case the introduction of further excess length leads to an energy increase but in the latter case to a relaxation of stress: The introduction of additional length at points X_L and X_R followed by a relaxation of the structure obviously decreases the total energy.

Note that the (more complex) second constraint Eq. 3.6 is eliminated through the "ansatz" *per se*. The total loop energy is given in terms of the loop radius r and the opening angle α

$$U_{tot}(\alpha, r) = A\frac{\alpha}{r} + 2\alpha R\varepsilon_{ads}$$

and by applying the constraint Eq. 3.21 (which this time can be solved explicitly!) we obtain U_{tot} in terms of α and given ΔL

$$U_{tot}(\alpha) = 2\alpha\left(A\frac{\tan\alpha - \alpha}{2R(\tan\alpha - \alpha) - \Delta L} + R\varepsilon_{ads}\right) \tag{3.22}$$

which is explicit in α. We note that this approximation for U_{tot} is only reasonable for $2R(\tan\alpha - \alpha) > \Delta L$, i.e., for not too small α (vs. ΔL), otherwise the bending contribution diverges or becomes even negative (the latter is obviously absurd). The reason for this is that for very small angles α (compared to ΔL) uncrossed[8] circle-line loops cannot exist for geometrical reasons. There this most basic approximation breaks down and we would have to approximate the loop by more than one circular segment. But as mentioned above, such loops (α small compared to ΔL) are not candidates for the ground state for moderate $\varepsilon_{ads} \sim O(1)$, and we therefore dispense with giving a discussion of this case.
The nice thing about Eq. 3.22 is that despite its simplicity and approximate nature it reproduces the position of the maximum in Fig. 3.5 quite well. We find the condition for the critical excess length ΔL_{crit} from a simple geometric distinction between two loop

[8]In contrast to crossed loops there still are solutions for small α (cf. next subsection).

shapes: the subcritical loop (Fig. 3.6a) with its tangents not being parallel to the Y axis ($\alpha = 0$) and the supercritical loop (Fig. 3.6b) having two or more tangents parallel to the line $\alpha = 0$. Suppose now we add excess length to a subcritical loop by keeping the angle $\alpha = const.$ Obviously the loop-energy increases because the loop radius r becomes smaller. On the other hand in the supercritical case we have the opposite situation: the loop energy decreases with increasing ΔL. This is simply because we could cut the loop at two points (X_L and X_R in Fig. 3.6), introduce there the additional length (without changing the energy) and then relax the shape by letting it evolve to the new equilibrium while keeping $\alpha = const.$ Thus we can obtain the condition for the critical excess length ΔL_{crit} by assuming that the corresponding minimum $\alpha^{\min}$ of U_{tot} just crosses the critical line $\pi/2$ line, i.e., $\alpha^{\min}(\Delta L^{crit}) \stackrel{!}{=} \pi/2$ for the searched ΔL_{crit}.

$$\left.\frac{d}{d\alpha}\right|_{\alpha=\pi/2} U_{tot}(\alpha) \stackrel{!}{=} 0 \tag{3.23}$$

which can be solved for ΔL^{crit}

$$\Delta L^{crit} = \frac{4R}{\pi} + \frac{8R^3}{\pi A}\varepsilon_{ads} \tag{3.24}$$

The latter can now be inserted in Eq. 3.22 leading to

$$U_{tot}^{crit} = \frac{\pi A}{2R} + \pi R\varepsilon_{ads} \tag{3.25}$$

For the given values of R, A, ε_{ads} ($R = 4nm$, $A = 50\ nmk_BT$, $\varepsilon_{ads} = 0.7\ k_BT/nm$) we obtain $\Delta L^{crit} = 7.37$nm and $U_{tot}^{crit} = 28.4k_BT$ which is in satisfactory agreement with the exact numeric results ($\Delta L^{crit} = 7.19nm$, $U_{tot}^{crit} = 26.7k_BT$). More generally, for not to high adsorption energies ($\varepsilon_{ads} = 0.5 - 2.0\ k_BT/nm$) the circle-line approximation works well and Eqs. 3.24 and 3.25 reproduce the exact positions of the critical point typically with a 5-15% accuracy.
For an explicit parametric representation of the minimal energy curve within the circle-line approximation, which in particular implies the upper results, the reader is referred to the appendix where the usefulness of this approach is also demonstrated for some other examples.

3.1.5 Crossed and Entropic Loops

A closer inspection of Fig. 3.4 shows that the ground state of loops switches from simple uncrossed loops to crossed loops when one reaches an excess length around 50 nm. However, as can be seen for the crossed structures 4, 5 and 6 in Fig. 3.3 these loops have a self-penetration at the crossing point. Therefore, a planar theory is in principle not sufficient to describe such structures. One possible formal cure for this problem would be to leave the plane and to consider the rod's self-contacts with the corresponding point-forces etc. in 3D as done by Coleman et al. in a general theory of rod self-contacts [86]. However such a procedure leads to a significant loss of transparence, not only because of the third dimension entering the scene but also

due to the necessity to subdivide the rod into different regions with different forces acting in each of them. Instead of following Coleman at al. [86] we decided to treat the self-interaction in a perturbational manner as follows. If the self-contact point is not too close to the nucleosome the rod is not severely deflected out of the plane by its self-interaction. Thus it remains roughly planar with some out of plane bending in Z-direction of the rod sections between the nucleosome and the crossing point. This will cost some additional bending energy U_{def} that is roughly given by (cf. Appendix)

$$U_{def}(\sigma, m) = \frac{2A}{R} \frac{\rho \arctan\left(\frac{\rho \tan\alpha(\sigma,m)}{\tan^2\alpha(\sigma,m)-\rho^2}\right)}{\tan^2\alpha(\sigma,m) - \rho^2} \tag{3.26}$$

Here $\rho := d/R$ with $d \approx 1nm$ is the DNA radius. We neglect the slight twisting of the rod induced by the non-planarity of the DNA and consider the bending only. The deflection energy Eq. 3.26 can be phenomenologically incorporated into the model by simply adding it to Eq. 3.20 as a correction term to obtain the final form of the total energy U^*_{tot}

$$U^*_{tot}(\sigma, m) = \begin{cases} U_{tot}(\sigma, m) \text{ for uncrossed (simple) loops} \\ U_{tot}(\sigma, m) + U_{def}(\sigma, m) \text{ for crossed loops} \end{cases}$$

With this additional modification of U_{tot} we computed numerically the minimal energy (ground state) solution for any given excess length ΔL. The graph of the ground state energy versus ΔL is shown if Fig. 3.5. We find that even with the inclusion of the out-of-plane deflection there is still a critical length ΔL_{cross} (here $\approx 60nm$) where the crossed loops become energetically more favorable than the simple uncrossed. This behavior that we call the "crossing transition" can be rationalized by noting that for long enough loops the adsorption energy (proportional to α) starts to dominate over the bending energy so that loops with smaller α become increasingly favorable. From the critical length ΔL_{cross} on, the gain in adsorption energy (by diminishing α) is more than sufficient to outweigh the (slight) increase in bending energy together with the additional self-interaction term, Eq. 3.26.

Increasing the length even further we leave the elastic energy dominated regime in which the entropic effects can be neglected due to short loop length ($\lesssim$ persistence length). For larger lengths entropic effects become more and more important and we ultimately enter the entropic loop regime. The crossover between these two regimes is hard to handle analytically [87]; for the case of closed loops a perturbative description has been given in Ref. [88]. For our purpose it is sufficient only to consider the asymptotic behavior. In the large loop limit where the loop is longer than several l_P the chain looses its "orientational memory" exponentially and behaves roughly as a random walk which starts from and returns to the same point. The entropic cost for gluing the ends of a random walk (long loop) together is then given by

$$U = 3/2 k_B T \ln(\Delta L / l_P) + E_0 + S_0 \tag{3.27}$$

The first constant, $E_0 \approx 6.5\ k_B T$ is the bending + adsorption energy contribution of the overcrossing DNA segments leaving / entering the nucleosome which can be determined

by minimizing the crossed loop energy (cf. Appendix Eq. 3.46) for $\Delta L \to \infty$. The second additive constant $S_0 \sim O(k_B T)$ accounts for the entropic contribution of DNA-histone octamer interaction volume (the proximity necessary for the histone octamer and DNA to see each other). Although the latter constant is not easy to estimate the following prediction is not sensitive to any additive constant. We expect a free energy minimum to occur at the overlap between the elastic ($\Delta L \lesssim l_P$) and entropic ($\Delta L \gg l_P$) region where the decreasing elastic energy is overtaken by the increasing entropic contribution.
The free energy, Eq. 3.27, leads to an algebraically decaying probability $w(\Delta L)$ for the jump lengths scaling as $w \sim (\Delta L)^{-3/2}$. In general, power law distributions of the form $w \sim (\Delta L)^{-\gamma}$ with $\gamma > 1$ lead to superdiffusive behavior of the random walker (here the nucleosome). According to Levy's limit theorem the probability distribution of the random walker (more precisely, the distribution of the sums of independent random variable drawn out from the same probability distribution $w \sim (\Delta L)^{-\gamma}$) converges to a stable Levy distribution of index $\gamma - 1$ [89, 90, 91]. This so-called Levy-flight differs in many respects from the usual diffusion process as for short time intervals big jumps are still available with significant probability. Moreover, all moments (besides possibly the first few ones) diverge. For our case $\gamma = 3/2$ even the first moment does not exist. We note that the value 3/2 is based on the assumption of an ideal chain (no excluded volume); in general the excluded volume leads to self-avoiding-walk statistics with a slightly larger value of γ around 2.2 [91] (cf. also Ref. [92]). In that case one has a finite value of the first moment, i.e., of the average jump length.

3.1.6 The Dynamics of Nucleosome Repositioning

In the preceding sections we have computed the typical energies involved in the formation of arbitrary sized loops. Assuming that a slow creation followed by a fast thermal migration of loops around the nucleosome is the governing mechanism for nucleosome repositioning we start now considering the repositioning dynamics. In order to describe the time-dependent evolution of the nucleosome position we consider its probability distribution along a DNA segment of a length $N \times 10bp$ and write the master equation governing the jump process

$$\frac{d}{dt} p_i = \sum_{j=1, j\neq i}^{N} w_{ji} p_j - p_i \sum_{j=1, j\neq i}^{N} w_{ij} \tag{3.28}$$

where p_i is the probability for the nucleosome being at the admissible[9] position i on the DNA segment. The transition rate matrix $\underline{\underline{W}} = (w_{ij})$ is given by

$$w_{ij} = \begin{cases} C_A \exp\left(-\frac{1}{k_B T} U_{\min}\left(h_D\,|i-j|\right)\right) & \text{for } i \neq j \\ -\sum_{k=1, k\neq i}^{N} w_{ij} & \text{for } i = j \end{cases} \tag{3.29}$$

where $h_D = 3.4nm$ (DNA helical pitch). C_A denotes the Arrhenius constant involved in the loop formation process that has in principle to be determined experimentally.

[9] Spaced by a multiple of 10 bp from the initial position

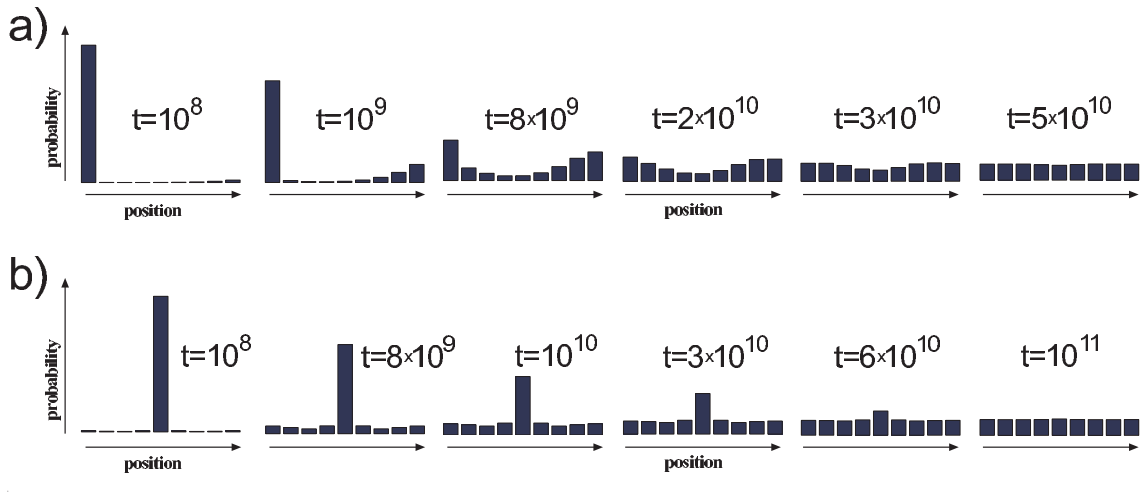


Figure 3.7: Relaxation dynamics of two initial states of nucleosome positions on a short DNA segment (147 + 90 bp): a) the nucleosome starting from an end and b) the nucleosome starting from the middle position. The time unit is the inverse Arrhenius activation factor C_A^{-1} (compare text).

The rough estimate of $C_A^{-1} = 10^{-6}s$ is provided in Ref. [69] where it was shown that C_A is essentially given by the inverse lifetime of the loop (denoted by A in that paper). This means that typical repositioning times range from seconds to hours.
The (formal) explicit solution of Eqs. 3.28, 3.29 together with the previously obtained minimal energy $U_{\min}$ is given by

$$\underline{p}\,(t) = \exp(\underline{\underline{W}}t)\underline{p}\,(0)$$

The latter solution can now be considered for different cases: for short or long DNA chains and for the nucleosome placed in the middle or at the end of the chain.
For short DNA segments we expect a slow repositioning rate due to high energies involved in small loop formation. In Fig. 3.7 we depict the repositioning of a nucleosome on a DNA piece of a length 147+90 bp. Starting from an end positioned nucleosome (Fig. 3.7a) we observe a behavior that is completely unlike a local diffusion mechanism: the jumps bigger than $\approx 2 \times 3.4nm$ start to dominate over the smaller local ones, which follows from the loop formation energy cf. Fig. 3.5. Consequently, in the initial phase of repositioning (of such an end-positioned population) the nucleosomes will predominantly jump between the two end positions. Later, on a much larger timescale they gradually start to explore the positions towards the middle of the DNA segment. Could we extract such a behavior from an experiment using gel-electrophoretic separation (as in [67], [68])? The basis of such separations is the fact that the gel-electrophoretic mobility of nucleosomes on DNA pieces (longer than 147bp) increases roughly linearly with its distance from the middle position, i.e., DNA pieces with the nucleosome sitting close to the end run much faster in gels than equivalent middle positioned nucleosomes do. We can exploit this (empiric) fact to mimic the outcome of a gel-electrophoresis experiment (cf. Figs. 3.8 and 3.10). In Fig. 3.8a we depict such a simulated gel pattern for the middle positioned nucleosome. Since symmetric species are not distinguished by

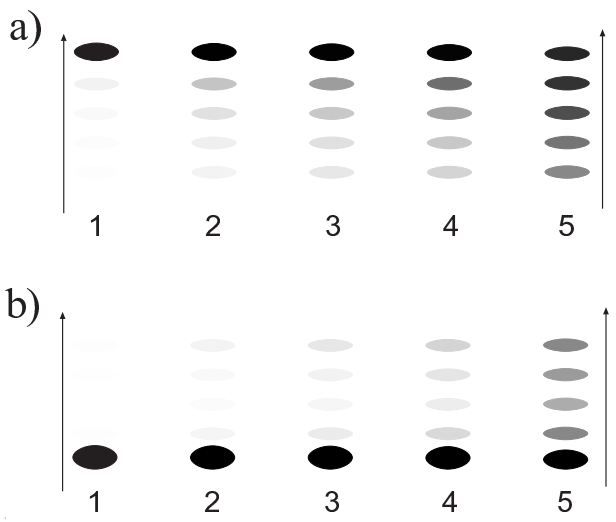


Figure 3.8: Typical (1-D) gel electrophoresis signatures expected for the relaxation dynamics of the two species from Fig. 3.7: a) nucleosome starts from an end and b) from the middle position. The lanes 1-5 correspond to incubation times $(1,5,10,20,100)\times 10^8$ C_A^{-1} respectively. Note: the population of distant bands in b) lanes 2-4 occurs first, in sharp contrast to what we expect from a simple (local) diffusive behavior.

this experimental method and are projected onto the same bands (symmetric left/right positions lead to the same mobility), the expected non-locality of motion cannot be extracted from the structure of the bands.

For the same short segment, but with the nucleosome starting from the middle position (Fig. 3.7b) the situation is slightly different: the neighboring positions are populated more homogeneously, although there is a small initial underpopulation of the $2\times 3.4nm$ distant position as expected from the energy maximum occurring there. In this case, a slight initial "population gap" can be observed in gel electrophoresis (Fig. 3.8b) which in this case would be sufficient to distinguish between a jumpy and a diffusive behavior, since the latter would obviously lack the "population gap".

In the case of longer DNA (but still not entropic segments) like the 147+300 bp segment in Figs. 3.8 and 3.9, similar effects as for the short segments are expected but with significantly faster relaxation times by typically 2-3 orders of magnitude as compared to the corresponding short segment populations. The corresponding (simulated) electrophoretic gels are shown in Fig. 3.10 where for the centrally positioned case (Fig. 3.10b) the "population gap" effect is even more pronounced than in the short segment case.

For even longer DNA segments we expect the gap effect to persist (data not shown) and the optimal jump size to be around 2-3$\times l_P$ corresponding to the free energy minimum in Fig. 3.5. For very long DNA segments, the nucleosome repositioning behavior implied by the big-loop-mechanism becomes strongly non-local which contrasts a local diffusive motion as expected from cork-screwing motion (cf. Refs. [65, 66, 68, 67]) or small loop repositioning as considered by Ref. [69]. As mentioned above, this superdif-

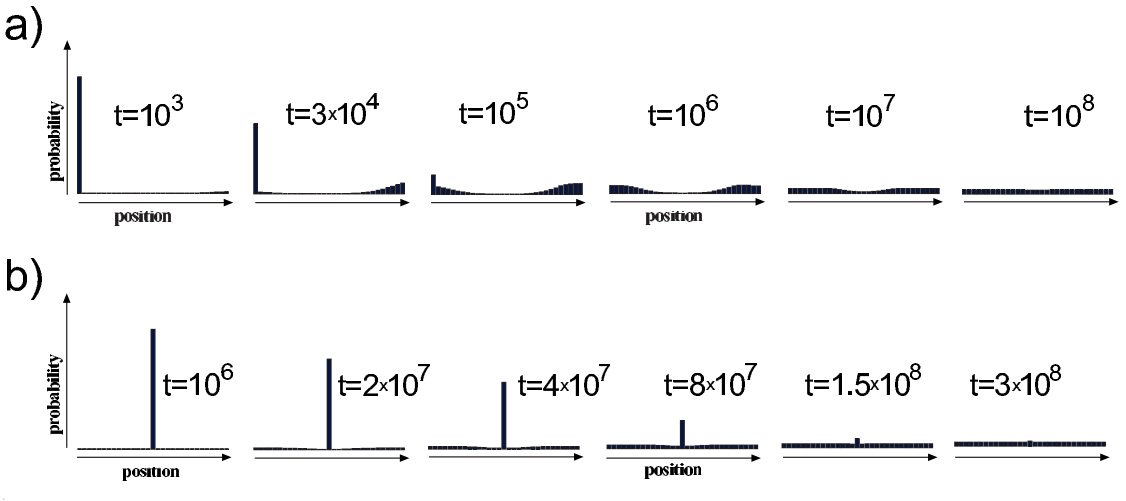


Figure 3.9: Relaxation dynamics of two initial states of nucleosome positions on a longer DNA segment (147 + 300 bp): a) end positioned and b) centrally positioned initial species. Note the initial difference in relaxation timescales for a) and b) (which are due to different loop energies involved).

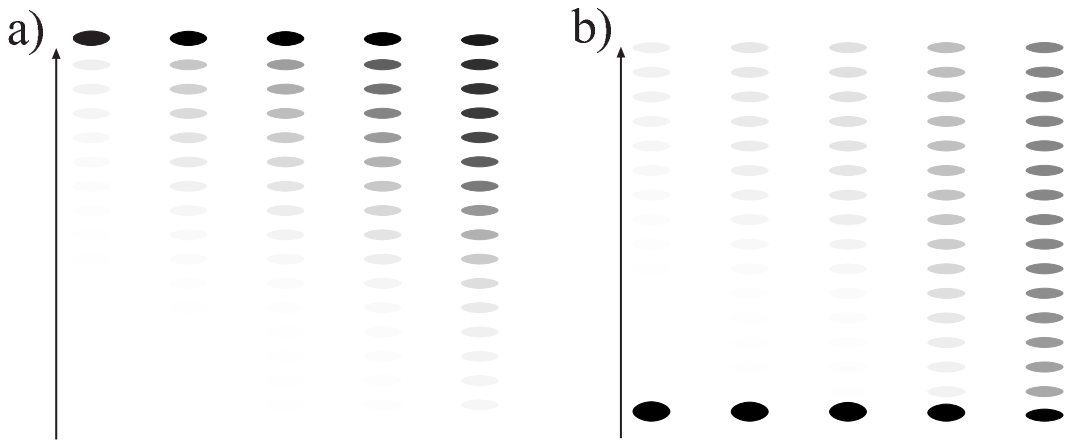


Figure 3.10: The (1-D) gel electrophoresis signatures simulated for the relaxation dynamics of the two initial species from Fig. 3.9. a) End positioned (lanes 1-5 corresponding to incubation times (1,2,3,10,50)$\times 10^4$ C_A^{-1}) and b) centrally positioned (incubation times (1,2,3,10,50)$\times 10^6$ C_A^{-1}).

fusive behavior has diverging moments which implies strongly enhanced nucleosome transport along very long DNA pieces. However such an ideal superdiffusion of nucleosomes could hardly occur *in vivo* because free DNA segments between subsequent nucleosomes (DNA linkers) are never longer than $\sim O(l_P)$. Furthermore the neighboring nucleosomes might be a significant barriers (if not for loop formation then) for loop migration around the nucleosome which is an indispensable event for loop-mediated repositioning.

3.1.7 Conclusions on the Loop Mechanism

In this section we examined the first possible mechanism for the repositioning of nucleosomes along DNA which is based on the formation and diffusion of intranucleosomal loops. The most important outcome of this section is the prediction of two classes of loops that might occur: (1) small 10bp-loops and (2) large loops with a wide distribution of stored lengths with a weak peak at roughly two times the DNA persistence length.

The small loops were already discussed in Ref. [69] and led to the prediction of repositioning steps of 10bps. Furthermore, the repositioning time should be of the order of an hour, a consequence of the large activation energy required to form a loop. This might explain the strong temperature dependence of the typical repositioning time [67]. In fact, by lowering the temperature from 37° to $4°C$ no redistribution within one hour was detected in that experiments. Assuming a loop formation energy of $23k_BT$ one finds indeed a slowing down of this process by factor of 13.

On the other hand, the large loop repositioning considered here turns out to be energetically much more favorable. Loops with an extra length of $2l_P$ have an energy that is roughly 12-13k_BT smaller than that of a 10bp-loop. To a certain extent this is because such loops can have a very small nucleosome opening angle by forming crossed loops but the main contribution stems from the significantly decreased DNA bending energy. One therefore expects that repositioning via large loops should be the dominant process on sufficiently large DNA pieces and that the typical times are much shorter than the one for small loop repositioning (say, of the order of minutes).

So far, however, the experiments did not report such events. Meerseman et al. [68, 67], for instance, found on short DNA pieces of 207bps length results that are consistent with 10bp repositioning – as we would expect for such short DNA fragments. However, when they redid the experiment with a 414bp long piece, a tandem repeat of the 207bp DNA, their analysis of the complicated band patterns observed in 2D gel electrophoresis did not show any indication that the nucleosome was able to move from one half to the other.

Hence, the question arises if the repositioning observed in these experiments was mediated via the loop mechanism or if it occurred via a different process. An analysis of the results is made especially difficult by two complications: (a) the nucleosomes seem to prefer to sit on the ends of the DNA fragments and (b) most of the experiments use strong positioning sequences (like the 5S rDNA sequence). This means that, independent of what the repositioning process might be, the nucleosomes have certain preferred positions and these might obscure the underlying repositioning process.

Despite those possible artifacts in the experiments performed up to now and the blurred and fuzzy picture they might draw it cannot be denied that the basic theoretical prediction of the loop model (the non-locality of motion) is *not observed experimentally*. We seem to be in a dead end with the loop model - not because of its logic inconsistency but rather because nature seemingly "refuses" to act that way. At this point one is reminded of the fact that dealing theoretically with such a highly complex and dynamic mesoscopic object like the nucleosome (consisting of more than 10000 atoms, without solvent and ions!) usually requires additional assumptions (a simplified model). The validity of the latter is just a working hypothesis always requiring experimental confirmation.
In the light of that we need to consider other repositioning mechanisms that one could imagine. The first and indeed most trivial mechanism is that the nucleosome detaches *completely* from the DNA and attaches at some other position (or even a different DNA molecule). This process, however, seems to be excluded by two facts (among others). First that no repositioning from one half to the other of the 414bp DNA or to competitor DNA fragments was observed [68, 67]. Secondly, once completely detached from the DNA template the histone octamer becomes unstable and disintegrates into a tetrameric and two dimeric subunits which makes an effective nucleosome reconstitution difficult (if not impossible). So one has to think of a completely different mechanism. In the next section we will describe a more promising scenario, and finally resolve the problem on the basis of new experimental findings.

3.2 Repositioning via Twist Diffusion

In this section we consider a different mechanism for thermal nucleosome repositioning: the twist diffusion or (more pictorially) the corkscrew motion, cf. Fig. 3.11a. The model is in some respect similar to the loop model presented above as it assumes that the nucleosome motion is carried out by the formation and migration of **defects** in the adsorbed DNA. But in contrast to the loop model the carrier of motion in this model is a **twist defect** that contains one missing or one extra bp (in contrast to a 10 bp containing loop defects). Experimentally, twist defects have been observed in the high resolution crystal structure of the nucleosome [126]. In that study the nucleosomes were reconstituted from histones and DNA of 146 bp assuming that this would be its optimal length in the crystal. However, the latter turned out to be 1bp longer, i.e. 147 bp. It was found that the missing bp of the 146 bp DNA was not localized at its terminus but instead at a 10 bp stretch close to the dyad axis (cf. Fig. 4d in Ref. [126]). That twist defect allows the DNA termini of adjacent particles in the crystal to come close in order to mimic a bp step. Obviously this gives an upper bound for the energetic cost of a single twist defect, namely the stacking energy of the blunt ends $\sim 10 - 20 k_B T$ [39]. From that we already see that twist defects might be energetically less costly than loop defects (from the previous section).
In order to model the twist diffusion mechanism we map the nucleosomal DNA on a Frenkel-Kontorova (FK) chain of particles connected by harmonic springs in a spatially periodic potential (cf. Fig. 3.11). The original FK model was introduced more than

sixty years ago to describe the motion of dislocations in crystals [43]. In the meantime variants of this model were applied to many different problems including charge density waves [44], sliding friction [45, 46], ionic conductors [47, 48], chains of coupled Josephson junctions [49] and adsorbed atomic monolayers [50, 109]. Here, in the context of DNA adsorbed on the octamer, the beads represent the base pairs. The springs in between have an equilibrium distance $b = 0.34$ nm and a stiffnes C that reflects the coupled DNA twist-stretch elasticity. Specifically

$$E_{elastic}\left(\{x_n\}\right) = \sum_k C\left(\frac{x_{k+1} - x_k}{b} - 1\right)^2 \tag{3.30}$$

Here the conformation of the wrapped DNA is given by the set $\{x_n\}$ where x_n is the position of the nth bp measured along the helical backbone; $C \simeq 70 - 100 k_B T$ is the combined twist and stretch spring constant including the (here unfavorable) twist-stretch coupling [52] and the summation goes over all bp associated with the wrapped DNA. In addition there is the external potential of the 14 contact points to the octamer with neighboring points being 10 bp apart [126] that we model as follows

$$\begin{aligned} E_{ads}\left(\{x_n\}\right) &= -U_0 \sum_k \sum_{l=1}^{14} \left(\left(\frac{x_k - 10bl}{a}\right)^2 - 1\right)^2 \\ &\quad \times\theta\left(a - |x_k - 10bl|\right) \end{aligned} \tag{3.31}$$

with θ being the Heaviside step function. The two parameters of the external potential, its depth U_0 and its width a, can be estimated as follows. U_0 represents the pure adsorption energy per point contact which follows from competitive protein binding [140] to be of order $6k_BT$. The other parameter, a, can be estimated from the fluctuations of the DNA in the crystal measured by the B-factor (cf. Fig. 1b in [126]) at different nucleosome positions. The ratio of DNA helix fluctuations $R_{fluct} = \langle x^2_{middle}\rangle / \langle x^2_{bond}\rangle \approx 3$ at positions between the binding sites and at the bound sites is a measure of DNA localization. Using a quadratic expansion of Eq. 3.31 one finds from a straightforward normal mode analysis that $a = \left(5U_0/[(R_{fluct} - 1)\, C]\right)^{1/2} b \sim b/2$, i.e., the adsorption regions lead to a strong localization of the DNA. Knowing all involved parameters the total energy of the DNA chain confined in the nucleosome can be written down

$$E_{tot} = E_{elastic} + E_{ads} + E_{sd} \tag{3.32}$$

The last term E_{sd} is the sequence dependent part of the total energy which we will neglect first. In the following we study the mechanism for thermal motion of DNA governed by E_{tot}. Generally two scenarios are possible: (*i*) the injection of a kink (1 bp missing) or antikink (1 additional bp) at either nucleosome end and (*ii*) the generation of kink-antikink pairs inside the nucleosome. Since the second mechanism is energetically roughly twice as costly than the first one, we will focus here on the (anti)kink injection mechanism only.

How and how fast does the kink step around the nucleosome? Due to the strong DNA localization at the binding sites ($a/b < 1$) for a realistic range of parameters U_0 and

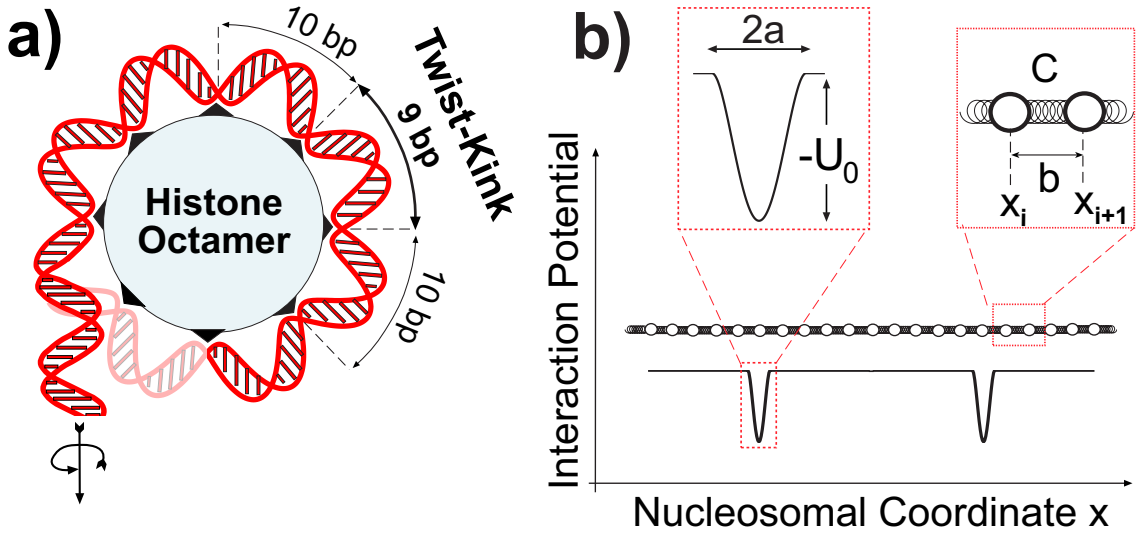


Figure 3.11: The twist-diffusion mechanism for nucleosome repositioning. a) A concerted translational and rotational motion of DNA leads to injection of twist-defects (kinks) which migrate between the octamer adsorption sites (black triangles) leading to a "creep" motion of DNA. b) The corresponding Frenkel-Kontorova model for twist diffusion and its characteristic parameters (cf. text for details).

C a given kink is localized either between two adsorption positions, i.e., smeared out over 10 bp (denoted by the K_{10} state), or between three of them, i.e., smeared out over 20 bp (the K_{20} state). It is obvious that the motion of a (anti)kink will consist of an alternation between K_{10} and K_{20} states similarly to an earthworm creep motion. To model this process we introduce the effective kink coordinate x_K describing the coordinate of that bp that goes from being pinned to being depinned during a single kink step, so that $x_K \approx 0$ and $x_K \approx b/2$ correspond to K_{10} and K_{20}, respectively, whereas $x_K \approx b$ means that the kink moved by one bp step. The Peierls-Nabarro potential experienced by the kink is then given by

$$U_{PN}(x_K) = \begin{cases} C_{eff}\left(x_K/b - 1/2\right)^2 - U_0\left(x_K^2/a^2 - 1\right)^2 \theta(a - x) & \text{for } 0 < x_K < b/2 \\ U_{PN}(b - x_K) & \text{for } b/2 \leq x_K < b \end{cases}$$

Here $C_{eff} = \frac{2}{10 \pm 1} C$ with "$-$" referring to a kink and "$+$" to an antikink. Depending on the ratio of parameters U_0 and C, the state K_{20} corresponds to a local minimum or maximum of U_{PN} whereas K_{10} is always stable for the relevant parameter range. The rate for the kink step process is then given by the expression

$$f_{step} = \frac{k_B T j_0}{b^2 \zeta_{eff}}$$

with

$$j_0^{-1} = \left(\int_0^1 e^{-U_{PN}(sb)/k_B T} ds\right)\left(\int_0^1 e^{+U_{PN}(sb)/k_B T} ds\right)$$

and

$$\zeta_{eff} = \frac{4\pi^2}{10b}\mu_{spin}$$

the effective kink friction constant. Here $\mu_{spin} = 1.3\times10^{-20}Ns$ is roughly the rotational friction for a single basestep [53]. To determine the rate at which twist defects are formed at the entry/exit points of the DNA one can now use an argument similar to the one presented in Ref. [41]: The ratio of the life time t_{life} of a kink to the time interval t_{inj} between two kink injection events at the end of the wrapped DNA portion equals the probability to find a defect on the nucleosome, i.e.

$$t_{life}/t_{inj} \simeq N_{site}e^{-U_{Kink}/k_BT}$$

Here $N_{site} = 13$ denotes the number of possible positions of the defect between the 14 binding sites and $U_{Kink} \simeq C/10$ is the energetic cost for a single kink (cf. above).
How is the average life time t_{life} of a defect related to t_{step}, the typical time needed for one step? This can be determined from the mean first passage times τ_{left} and τ_{right} for a defect that forms, say, at the left end to leave the nucleosome at the same or at the other end, respectively. From Ref. [54] one finds $\tau_{left} = (25/6)\, t_{step}$ and $\tau_{right} = 28t_{step}$. Furthermore, the probability to leave at the left end is $p_{left} = 12/13$ and at the right end $p_{right} = 1/13$ [54] which gives the life time as the weighted average $t_{life} = 6t_{step}$. Only a fraction p_{right} of the defects reaches the other end and will lead to a repositioning step, i.e., the time of a 1bp diffusion step of the nucleosome along the DNA is given by $T = t_{inj}/p_{right}$. Putting all this together we arrive at $T \simeq 6b^2\zeta_{eff}j_0^{-1}/k_BT\exp(C/10k_BT)$. For realistic parameter values $C = 100k_BT$, $U_0 = 6k_BT$ and $R_{fluct} = 3$ we find $T \simeq 10^{-3}$ s implying a nucleosome diffusion constant $D = 580$ bp^2/s $= 6.6\times10^{-17}$ m^2/s. Note that $U_{Kink} \simeq 9k_BT$ for K_{10} and $\simeq 11k_BT$ for K_{20}.
Hence we find repositioning rates that are orders of magnitude faster than the ones observed in experiments [103]. Even worse, the experimental observation of an apparent 10 bp jump length [103] seems to be inconsistent with our predictions. We show now how these facts can be explained by the existence of additional barriers with a 10 bp periodicity. To do so we have to extend our simple model to deal with the quenched disorder stored in the DNA bp sequence. The sequence dependent anisotropic bendability, i.e., the propensity of DNA to bend in different directions with different elastic constants turns out to be essential. It has been known for long [55, 56] that (A/T) rich dinucleotide steps (dns) prefer to face the octamer in the minor groove (i.e., at the octamer contact points) whereas (G/C) rich dns prefer to face the octamer in the major groove (i.e., between contact points). This reflects different propensities of the dinucleotides to widen or compress towards the DNA minor groove. To incorporate these anisotropic effects into our model we first note that the bending state of the DNA molecule is fully constrained by its helical path on the octamer surface. Moving a DNA sequence via twist diffusion by a few bp ($<$ 10 bp) along that path changes the relative rotational setting of the bent DNA with respect to its preferred bending direction causing an energetic penalty, whereas a motion by 10 bp restores the initial rotational setting. We address this by introducing a 10bp periodic "bending field" $F_{bend}(x) = -\cos[2\pi x/(10b)]$ attached to the octamer surface. We assume the DNA

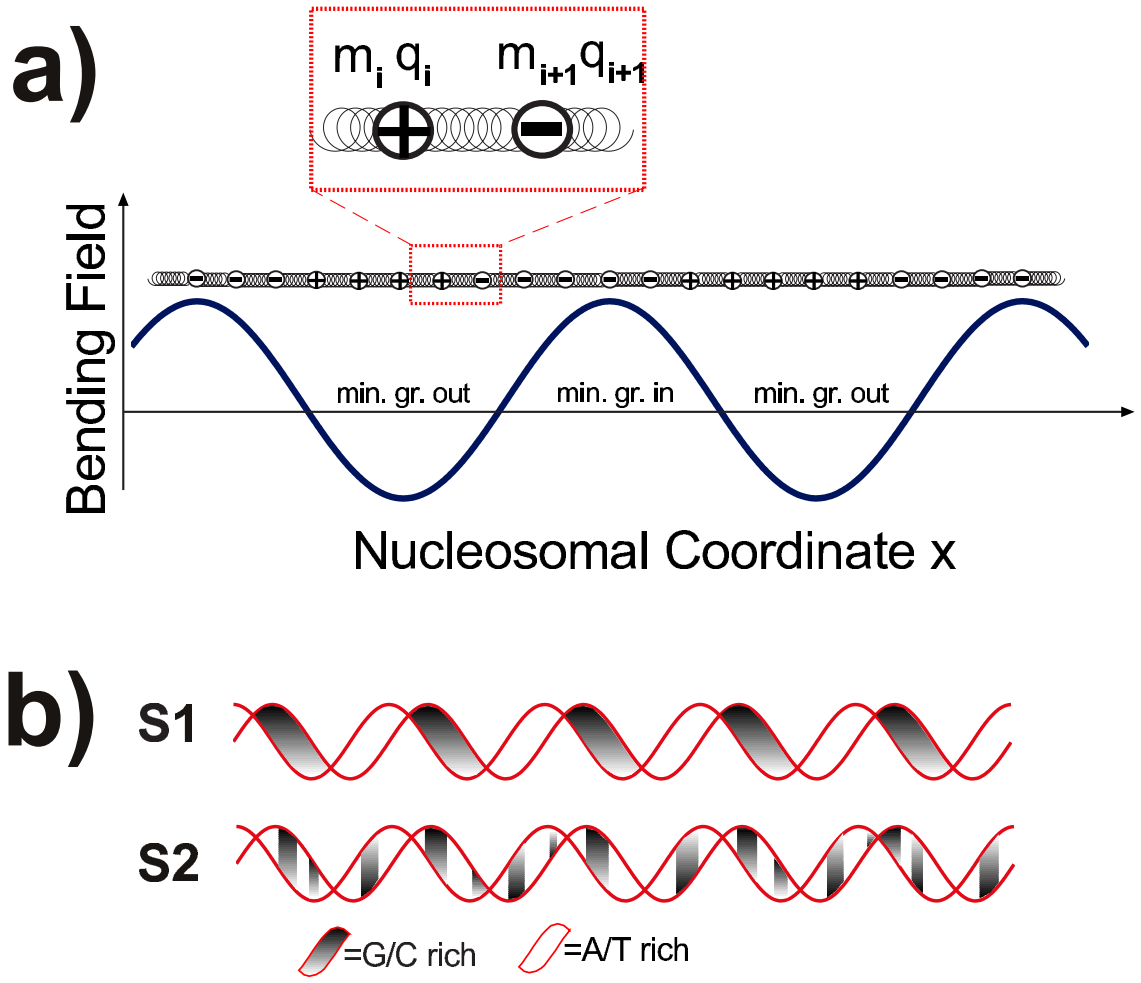


Figure 3.12: The extended Frenkel-Kontorova model includes effects from anisotropic bp sequences. a) DNA sequences couple additionally to an octamer-fixed "bending field" through the anisotropic bending parameters q_i ("bending charge"). b) Two sequences with extremely different mobilities. S1: highly anisotropic, 10 bp phased ("TG"-like) sequence with $D_{sd} \approx 10^{-4} - 10^{-5}$ bp^2/s. S2: random sequence corresponding to $> 95\%$ of the genome with $D_{sd} \approx 10^2$ bp^2/s.

sequence to couple linearly to that field through "bending charges" q_k attached to each of the dns. This gives us finally the third term in Eq. 3.32:

$$E_{sd} = \sum\nolimits_k q_k F_{bend}\left(x_k\right) + m_k \tag{3.33}$$

In addition to the anisotropic term we also introduced here the isotropic bending parameters m_k to include isotropic flexibility effects (which become important when the q_k's vanish or average out). The summation involved is again over all base pairs incorporated in the nucleosome. q_k and m_k both have units of energy and can be extracted from competitive protein binding experiments [56] for each of the 10 dns (AA, AT, GC...). To obtain a rough estimate we distribute the dns into three classes: 1) (G/C) containing dns, 2) (A/T) containing dns and 3) mixed dns (like AG, CT etc.) and treat the dns in each class as identical. Using the available experimental data [56, 57] we then arrive at $q_{G/C} \approx\ 95$, $q_{A/T} \approx -85$, $q_{mixed} \approx 0$ and $m_{G/C} \approx 20$, $m_{A/T} \approx -3$, $m_{mixed} \approx 7$, where all energies are in cal/mol per dns.
It turns out that the nucleosome mobility depends strongly on the underlying bp sequence. When shifting the position of all beads by l bp steps, $x_k \rightarrow x_k + lb$, we find $E_{sd}\left(l\right) = \left(A/2\right)\cos\left(2\pi l/10 - \phi\right)$ to vary as a cosine function of l with phase ϕ and amplitude A determined by the DNA sequence, which is assumed to be appropriately periodic here. Arranging G/C and A/T tracts properly and taking the sequence dependent q and m values given above we can easily reach amplitudes A (i.e. barriers to repositioning) that exceed $10 - 12$ kcal/mol! A very effective sequence arrangement called the "TG"-sequence which leads to a strong nucleosome stability and localization was experimentally constructed in Ref. [56] by putting G/C tracts around positions $k = 0, 10, 20...$ and A/T tracts around $k = 5, 15, 25....$ In our picture this means to put the "bending charges" q along the DNA such that they couple favorably to the bending field F_{bend} for a distinct rotational setting whereas a 5 bp shift is extremely costly. The 5S-RNA sequence which was used in most nucleosome mobility experiments shows also the effect of an optimal rotational setting. It is less pronounced than in the "TG" case, yet it is still detectable. More involved theoretical computations relying on molecular sequence dependent deformability parameters [58] reveal barriers $A \approx 5 - 6$ kcal/mol for this particular sequence. The sequence dependent barrier height A exponentially suppresses the bare (sequence independent) diffusion constant D obtained above leading to the sequence dependent diffusion constant D_{sd}:

$$D_{sd} = DI_0^{-2}\left(A/2k_BT\right) \approx \frac{\pi j_0 A}{12\zeta_{eff}} e^{-(A+C/10)/k_BT} \tag{3.34}$$

with I_0 being the modified Bessel function.
Equation 3.34 predicts that mobility experiments with highly anisotropic sequences like "TG" (instead of the standard "5S-RNA") would find hardly any appreciable repositioning on the one hour timescale *if* it would be solely mediated via twist defects ($D_{sd} = 10^{-6} - 10^{-7} \times D = 10^{-4} - 10^{-5}$ bp^2/s). The typical path for a nucleosome to escape from such a rotational trap goes very likely via the previously considered loop formation mechanism that allows "tunneling" over sequence barriers, thus dominating over twist-diffusion for extremely anisotropic sequences. An experimental test for this

prediction would be to increase the free DNA segment length which in this regime should strongly enhance the loop mediated mobility [111] whereas it would leave the twist diffusion unaffected. Going to the other extreme, in the most relevant case of random isotropically bendable sequences which make up more than 95% of the eucaryotic genome one should observe that the twist diffusion mechanism is strongly enhanced by 2-3 orders of magnitude as compared to the *in vitro* measurements on "5S-RNA".
In conclusion the following picture is implied by the twist diffusion mechanism: On physiological timescales the majority of genomic nucleosomes seems to be intrinsically highly mobile. However, only a small fraction ($< 5\%$) of all nucleosomes has strongly reduced mobility due to anisotropic DNA sequences which they populate. We may speculate that only the latter require the action of active (ATP consuming) remodelling mechanisms [59] making them hotspots and switching elements for global chromatin rearrangements.
Despite the plausibility of the twist-diffusion model (and its much better consistency with the experimental data than the loop model gave before) a decisive experiment confirming this mechanism was still missing at the time we studied it first [112]. In the next section we will close this gap by carefully analyzing the newest experimental data that became available only very recently.

3.3 Nucleosome Corkscrew Dynamics in the Presence of DNA Ligands

3.3.1 The Experiment

In the previous sections we have considered two different possible scenarios for thermally driven nucleosome motion along a DNA template. Both mechanisms are equally plausible and we can only rely on experimental observations to decide which one of them (if not both) is valid for the nucleosome. Although the present experiments that show a locality of nucleosome motion (no large jumps observed) seem slightly to favor the twist diffusion over the loop formation mechanism the available data are not straight forward to interpret (and still may suffer from artifacts).
A new experimental approach that could clarify the situation was taken recently by Gottesfeld *et al.* [115]. The authors considered the usual experimental setup i.e. a 216 bp DNA fragment that again contained the sea urchin 5S rDNA nucleosome positioning sequence. They followed the heat induced nucleosome repositioning (as done in all previous experiments) but this time in the presence of pyrrole-imidazole polyamides (PIP s), synthetic minor-groove binding DNA ligands that are designed to bind to specific target sequences with high (nanomolar) affinities. Experiments have been performed in the presence of one of 4 different ligands, each having one binding site on the nucleosomal DNA. The general outcome of this study was as follows: (1) A one-hour incubation at 37° in the absence of any ligand leads to a redistribution of the nucleosomes. (2) In the presence of 100 nM ligands **no** repositioning of nucleosomes is detected after such an incubation if (and only if) the target sequence of this specific

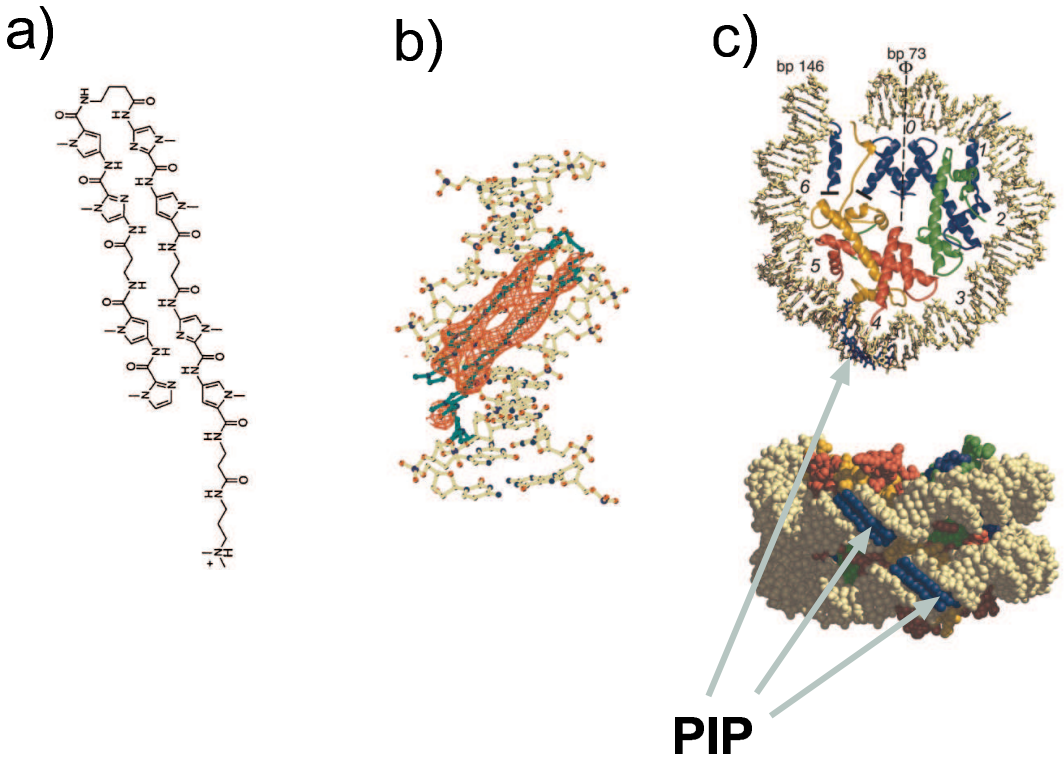


Figure 3.13: Sequence specific DNA ligands pyrrole-imidazole polyamides (PIPs) adopted form Ref. [115]: a) The chemical structure of a typical PIP. b) A PIP bound to the minor groove of DNA. c) PIP binding to specific DNA sequences on the nucleosome. A PIP can bind to the nucleosome only if its DNA binding sequence is faced with its minor groove outwards (away from the histone octamer surface)

ligand faces to the solution when the DNA is bent in its preferred direction. (3) If a ligand has been added whose binding site faces the octamer in its preferred rotational frame, the ligand has no detectable effect on the reposition dynamics.
This raises the question whether the described experiment is capable to distinguish between loop- and twist-defect induced nucleosome mobility. Since the ligands bind to the minor groove (cf. the co-crystal complexes between nucleosomes and such ligands [116] and Fig. 3.13) it is quite likely that a bound ligand will block the overall corkscrew motion of the DNA: The DNA can only rotate on the nucleosome up to a point where the bound ligand comes close to one of the 14 binding sites. A further rotation of the DNA is not possible because of steric hindrance and twist defects that would induce that rotation will be deflected once they encounter the ligand site. In other words: The observed suppression of mobility through ligand binding agrees qualitatively well with the twist defect picture. What about a loop defect encountering a bound ligand? In this case the answer is at first not obvious. One should expect that a bound ligand does not hinder bulge diffusion – at least sterically (for their small size). Of course, the ligand might in principle locally stiffen or soften the DNA and might influence the loop

formation mechanism (either enhancing or suppressing loop formation). But this seems unlikely as the ligand affinity for DNA is insensitive to the actual DNA curvature as seen from the almost unaltered ligand binding constants to free DNA as compared to the strongly curved (and distorted) DNA inside of the nucleosome. Consistently with that the co-crystal of the PIP ligands with DNA shows only minor distortions of the DNA. This again indicates that the PIP binding induces a fairly small (if not negligible) conformational perturbation of the underlying DNA molecule which could not account for any significant modulation of the loop diffusion mechanism. It therefore seems that the influence of ligand binding on nucleosome mobility supports much more the idea of twist diffusion as the underlying mechanism.
In the following we provide a theoretical model for nucleosome repositioning in the presence of DNA ligands. We make use of the upper results on repositioning via twist defects in the absence of ligands that essentially provides us with the nucleosomal diffusion constant as a function of temperature and underlying DNA sequence. Assuming thermodynamic equilibrium we will then calculate the diffusion constant in the presence of ligands. We find below – in agreement with the experiments – that in the presence of 100nM ligands the repositioning on the 5S positioning sequence is essentially completely blocked if the ligand binding site prefers to face the solution. On the other hand, when the binding site faces the octamer surface, the ligands have a negligible influence on the nucleosome mobility again in complete agreement with the experiment.

3.3.2 Nucleosome-Ligand Complex: Equilibrium Properties

In this subsection we determine the equilibrium properties of a nucleosome in the presence of a finite concentration $[L]$ of one synthetic ligand targeting one specific site on the nucleosomal DNA. In Fig. 3.14 we represent the different possible states by nodes and the possible pathways from one state to the next state by connecting lines. Fig. 3.14(a) shows the case of a DNA template with an isotropic bendability. The open and filled circles correspond to nucleosomes at different positions without a bound ligand. The filled circles (state "1") correspond to states where ligands can bind, i.e., to states where the ligand binding site (assumed to be located on the wrapped DNA portion) faces away from the octamer surface and is accessible. In this case a ligand can bind; the nucleosome with bound ligand is represented by an open square (state "0"). We assume that in this case the nucleosome looses its mobility, i.e., we have no line connecting this state to a neighboring state. Before the nucleosome can "slide" to a neighboring position the ligand has to unbind, i.e., one has to go back to state "1". If the nucleosome is in a position where the ligand binding site faces the octamer (open circle, state "2") the site is blocked. At these positions the nucleosome mobility is not affected by the ligands. For simplicity, we will assume here that always 5 consecutive bp positions (corresponding to one half turn of the corkscrew motion) have the ligand binding site exposed to the solvent and represent these 5 positions by the filled circle. Likewise the other 5 positions are collapsed into the open circle.
Fig. 3.14(b) and 3.14(c) show the case of a rotational positioning sequence as used in the experiment [115]. In case (b) the situation is such that the ligand can bind when the

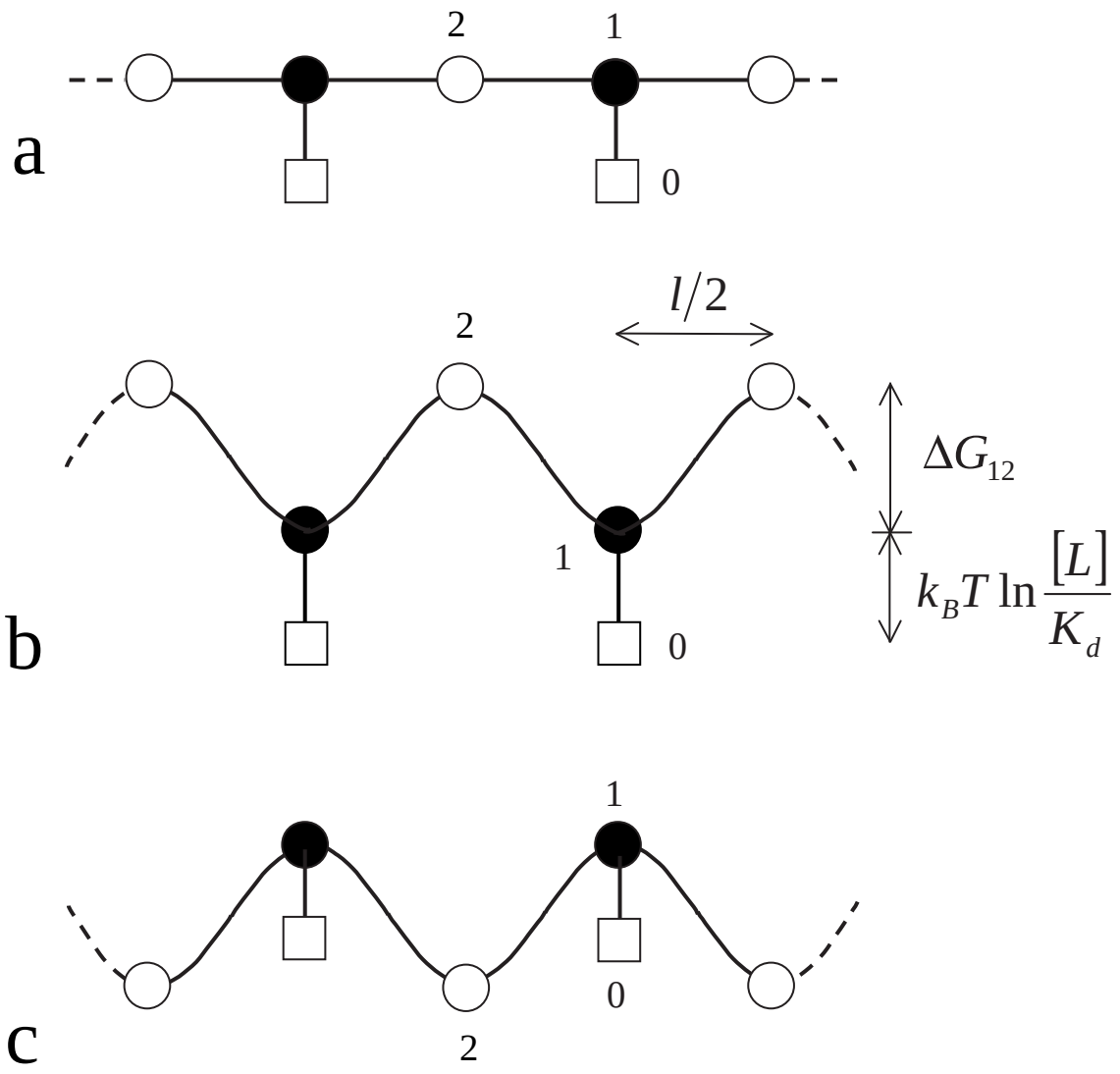


Figure 3.14: Nucleosome repositioning in the presence of ligands. (a) In the case of a homogenously bendable DNA template the states with exposed binding sites ("1") and occluded ones ("2") have the same elastic energy, $\Delta G_{12} = 0$. State "0" represents the immobile state with a bound ligand. (b) and (c) For templates with a rotational positioning one state has a higher DNA bending energy than the other. For case (b) the preferred rotational frame with respect to the octamer corresponds to an open binding site, in case (c) to a closed one. Each node in this scheme represents 5 consecutive bp positions so that the periodicity of the bending potential corresponds to one helical pitch.

DNA sits on the nucleosome in its preferred bending direction. In this case the states "1" sit in the potential wells of the elastic energy landscape. We call the difference in the DNA bending energy between the top and the bottom $\Delta G_{12} = G_2 - G_1$. The case $\Delta G_{12} > 0$ correspond to the situation where a ligand can most effectively bind to the nucleosome and block the repositioning (in the experiment this corresponds to the ligands 1 and 4 [115]). Fig. 3.14(c) depicts the other extreme where the binding site faces the octamer in the preferred rotational frame (this corresponds to ligands 2 and 3 in the experiment [115]).
We denote by p_i the probability for the nucleosome to be in state i. Detailed balance relates these probabilities as follows:

$$\frac{p_2}{p_1} = \frac{\omega_{12}}{\omega_{21}} = e^{-\Delta G_{12}/k_B T} = f \tag{3.35}$$

and

$$\frac{p_0}{p_1} = \frac{\omega_{10}}{\omega_{01}} = K = \frac{[L]}{K_d} \tag{3.36}$$

Here ω_{ij} denotes the transition rate from state i to state j, K is the equilibrium constant for the ligand and K_d its dissociation constant. Eqs. 3.35 and 3.36 together with $p_1+p_2+p_3 = 1$ yield immediately the occupation probabilities for the three states:

$$p_0 = \frac{K}{1 + K + f} \tag{3.37}$$

$$p_1 = \frac{1}{1 + K + f} \tag{3.38}$$

and

$$p_2 = \frac{f}{1 + K + f} \tag{3.39}$$

Let us consider some special cases. In the absence of ligands one has $[L] = 0$ and $K = 0$. For a homogeneous template one finds in addition f (defined in Eq. 3.35) to equal unity. Then $p_1 = p_2 = 1/2$ and, of course, $p_0 = 0$. Using a rotational positioning sequence the nucleosome prefers to be in its optimal rotational frame. Then for $K = 0$ one finds $p_2/p_1 = f$. For the 5S positioning sequence one has $f \approx e^{-10}$ so that state "1" is populated with a roughly 20000 times higher probability then state "2". This explains the electrophoretic band structure with a 10 bp periodicity as observed in most repositioning experiments [103, 106]. The presence of ligands changes the relative weight of the different states. We mention here the most intriguing situation: A rotational positioning sequence with the ligand binding facing inwards in the preferred rotational frame where $\Delta G_{12} < 0$ and hence $f > 1$ (cf. Fig. 3.14(c)). The probability to find the nucleosome in its mechanically unfavorable states "0" and "1" will be higher than to find it in state "2", the usually preferred state, if $p_0 + p_1 > p_2$, i.e., for $K + 1 > f$. For sufficiently high concentration and affinity the presence of the ligand overrules the positioning sequence! The ligands used in the experiment [115] have dissociation constants ranging from 0.7 to $6.0nM$. For a strong positioning sequence f is too large for the above inequality to hold for reasonable ligand concentration, say $100nM$. However, for less strong sequences this might play a role. Also in the case

of a 146 bp template corresponding to the total wrapping length as consider by Suto *et al.* [116] it might well be that ligands shift the preferred centered DNA position to an off-centered position. A 1-9 bp shift would cost the opening of one binding site but might allow ligands to bind more effectively, especially if the binding site(s) at the centered DNA positions are (partially) occluded. In one case [116] (polyamid 2) such an effect might have been indeed observed, cf. Fig. 5 in that paper.

3.3.3 Nucleosome Mobility in the Presence of Ligands

We are now in the position to determine the diffusion constant of a nucleosome along DNA in the various cases. The diffusion constant can be determined from the average of the diffusion constant for the nucleosome to jump from state "1" to one of the two neighboring states "1" and that of going from "2" to neighboring "2's". Let us denote by ω_1 the rate to go from a given state "1" to the next position "1" to the right and by ω_2 the rate of jumps to the right from "2" to "2". Then the diffusion constant is given by $D = (p_1\omega_1 + p_2\omega_2)\, l^2$ where l is the jump length, here $l = 10bp$. Now $\omega_1 = \omega_2$ follows from Kramers' rate theory [121] to be $\nu_0 e^{-|\Delta G_{12}|/k_BT}$ with the attempt frequency $\nu_0 \approx D_0/l^2$ for $|\Delta G_{12}| \lesssim k_BT$ and $\nu_0 \approx \pi\, |\Delta G_{12}|\, D_0/\left(k_BTl^2\right)$ for $|\Delta G_{12}| \gg k_BT$.
Using Eqs. 3.38 to 3.39 we arrive at the final formula for the diffusion constant for the case $\Delta G_{12} \geq 0$ (i.e. $f \leq 1$):

$$D_> = \frac{\nu_0}{1+K+f}\left(f+f^2\right) l^2 \tag{3.40}$$

In the opposite case, $\Delta G_{12} \leq 0$ (i.e. $f \geq 1$), we find

$$D_< = \frac{\nu_0}{1+K+f}\left(f^{-1}+1\right) l^2 \tag{3.41}$$

Let us now consider special cases:
(*i*) homogeneous DNA bendability, no ligands ($\Delta G_{12} = 0$, $[L] = 0$): In that case $f = 1$, $K = 0$. Both formulas, Eqs. 3.40 and 3.41 reduce to $D_> = D_< = D_0$.
(*ii*) homogeneous DNA bendability but ligands present ($\Delta G_{12} = 0$, $[L] > 0$), cf. Fig. 3.14(a): $f = 1$ leads to

$$D_> = D_< = D = \frac{2D_0}{2+K} \tag{3.42}$$

(*iii*) rotational positioning sequence, no ligands present ($|\Delta G_{12}| \gg k_BT$, $[L] = 0$): Equations 3.40 and 3.41 reduce to Eq. 3.34 with $|\Delta G_{12}| = A \gg k_BT$, the case that has been already discussed in the previous section [112].
(*iv*) rotational positioning sequence, ligands present with binding site exposed for the preferred orientational frame ($\Delta G_{12} \gg k_BT$, $[L] > 0$), cf. Fig. 3.14(b): Using $f \ll 1$ we obtain from Eq. 3.40

$$D_> = \frac{\pi\, |\Delta G_{12}|\, f}{k_BT}\frac{D_0}{1+K} \tag{3.43}$$

(*v*) rotational positioning sequence, ligands present with exposed binding site occluded for the preferred orientational frame ($\Delta G_{12} \ll k_BT$, $[L] > 0$), cf. Fig. 3.14(c): Here

$f \gg 1$ and from Eq. 3.41 we find

$$D_< = \frac{\pi \left|\Delta G_{12}\right|}{k_B T} \frac{D_0}{f + K} \tag{3.44}$$

We are now in the position to check how effectively the ligands can reduce the repositioning in the various cases. We estimate in the following the typical equilibration time on a 216 bp long template (as it has been used in Ref. [115]) to be $T_{70bp} = (216 - 146)^2 \, bp^2/(2D)$. Let us start with case (i) where $D = D_0 \approx 580 bp^2/s$. This leads to the typical time $T_{70bp} = 4s$. Adding now a ligand with $[L] = 100nM$ and $K_d = 1nM$ (case (ii)) this leads to a 50 fold reduction of the diffusion constant, $D \approx 12 bp^2/s$, cf. Eq. 3.42, and to an equilibration time $T_{70bp} \approx 3.5min$. If one uses a positioning sequence instead with $|\Delta G_{12}| = 9 k_B T$ one finds in the absence of ligands (case (iii)) from Eq. 3.34 $D \approx 2 bp^2/s$ and $T_{70bp} \approx 20min$. Repositioning experiments on such sequences are thus typically performed on time scale of an hour to ensure equilibration [103, 115]. Adding now again a ligand with $[L] = 100nM$ and $K_d = 1nM$ and having its binding site facing the solution in the preferred rotational frame (case (iv), Fig. 3.14(b)) we predict from Eq. 3.43 an additional dramatic reduction of the diffusion constant by a factor of 100: $D_> \approx 2 \times 10^{-2} bp^2/s$ and $T_{70bp} \approx 34h$. In other words, in this situation one does not observe any repositioning of the nucleosomes on the time scale of an hour. This is in accordance with the experimental observations, cf. Fig. 5, lane 1 and 4 in the study by Gottesfeld *et al.* [115]. On the other hand, for the case of a ligand with same affinity and concentration but with the binding site in the unfavorable orientation (case (v), Fig. 3.14(c)) one finds hardly any effect; in fact the diffusion constant as compared to the ligand free case, case(iii), is reduced by approximately 1 percent, cf. Eq. 3.44. As a matter of fact in the experiment [115] these two cases were indeed indistinguishable as seen in Fig. 5, lane 0, 2 and 3 in that paper.
In conclusion the upper analysis explains the whole set of experimental data provided in [115] on the basis of the twist diffusion model. Strikingly this analysis required no adjustable parameters. Taken together this provides a strong theoretical support for the twist diffusion model.

3.4 Appendix: The Circle-Line Approximation

Although the Kirchhoff's analogy provides us with essentially analytic solutions for the rod deformed in plane, the occurrence of boundary conditions (like Eqs. 3.5 and 3.6) prevents us in most cases from obtaining analytical expressions of all the parameters characterizing the solution (like σ and m above). To overcome this problem, we suggest here a simple geometric approximation scheme which will prove to be useful in obtaining analytic results for loops within a reasonable accuracy (usually with a deviation of 5-15% from the exact numeric results).
The main idea is the following. The curvature and the energy (Eqs. 3.17 and 3.18) of the loop contains the $\mathrm{cn}(\sigma|m)$ function which for $0 < m < 1$ has the typical oscillatory behavior depicted in Fig. 3.15 (left). This suggests to approximate the curvature

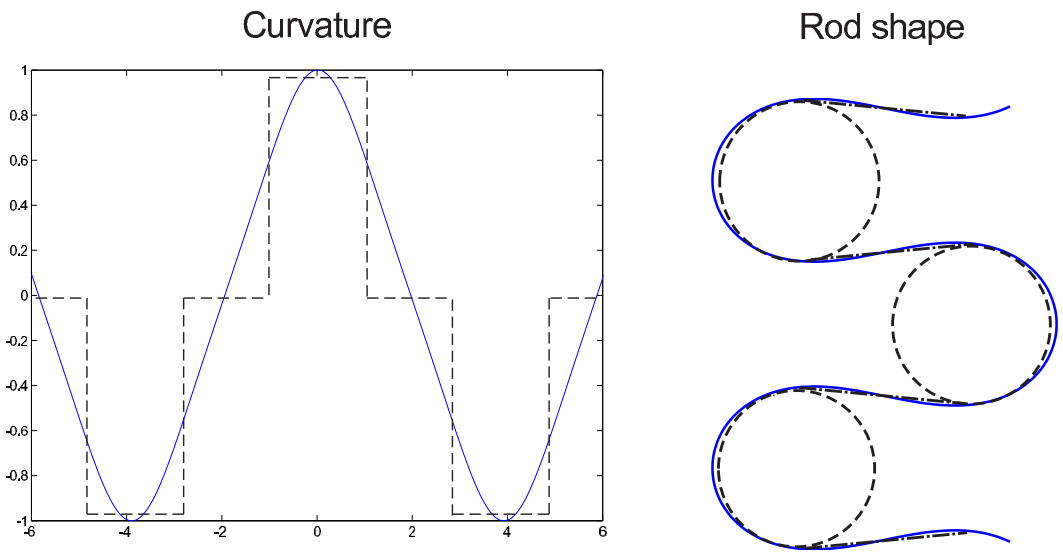


Figure 3.15: The *circle-line approximation* for planar rods. The curvature of an equilibrium rod shape (cn-function, cf. Eq. 3.17) is approximated by a periodic sequence of step-functions. The latter corresponds to an approximation of the rod shape by a sequence of straight lines ($\kappa = 0$) and circles (κ =const.) glued together in a smooth manner (continuous tangents).

function simply by a step function consisting of an alternating sequence of negative, zero and positive piecewise constant curvatures. Consequently the corresponding rod shape (Fig. 3.15 right) is approximated by a sequence of circles (positive / negative constant curvature) and lines (zero curvature). An analogous approximation procedure can also be performed in the case $m > 1$ where the cn function has a natural analytical continuation into a dn function with a modified second argument (cf. Ref. [85]).
Using this approximation ansatz several problems concerning planar rods reduce to elementary geometry as seen from the following simple but illustrative examples.
1) The Yamakawa-Stockmayer angle [88]: Two points on the rod are glued together without restricting the orientation of the tangents, e.g., a protein connects two distant points on DNA (cf. Fig. 3.16a). What is the preferred angle χ between the tangents in the ground state of the rod? By imposing a fixed total rod length L we have the simple constraint $L = \left(2\cot\frac{\chi}{2} + \chi + \pi\right) r$ from which we can eliminate r and write the elastic energy of the configuration as $U_{DNA}^{bend} = \frac{A}{L}(\chi + \pi)\left(2\cot\frac{\chi}{2} + \chi + \pi\right)$. Its minimization leads to the transcendent condition $\chi_{\min} + \pi = \tan\chi_{\min}$ with the only relevant solution $\chi_{\min} \approx 77.5°$. The latter angle differs by 5% from the exact result $\chi_{\min} \approx 81.6°$ (by Yamakawa and Stockmayer in [88]) which is satisfactory regarding the simplicity of the computation.
2) Simple and crossed loops (Fig. 3.16 b,c)**:** We can easily derive an approximate energy expression for simple / crossed loops as a function of the excess length ΔL and the opening angle α. By applying simple geometry the excess length constraint can be easily eliminated (the tangency constraint is trivially fulfilled by the ansatz) and we

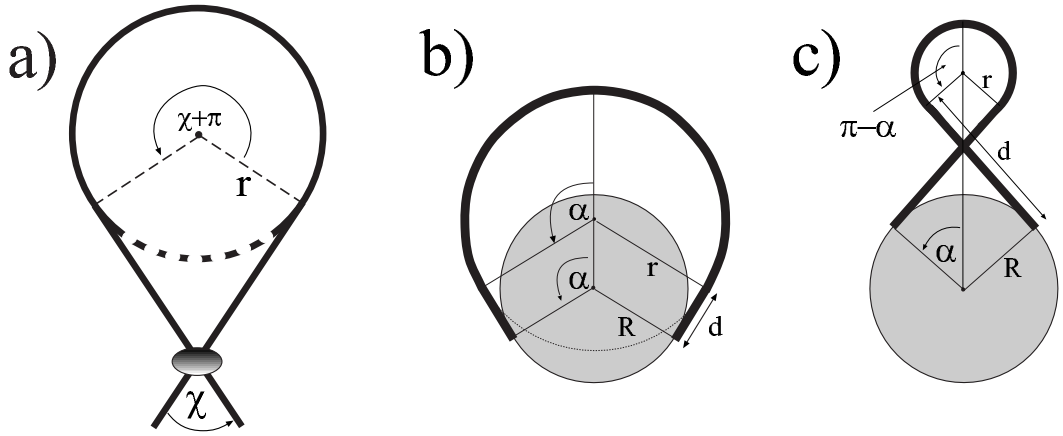


Figure 3.16: Three applications of the circle-line approximation: Problems with complex constraints reduce to simple geometries leading to good approximations: a) the Yamakawa-Stockmayer angle b) simple loops and c) crossed loops (see the Appendix text for details).

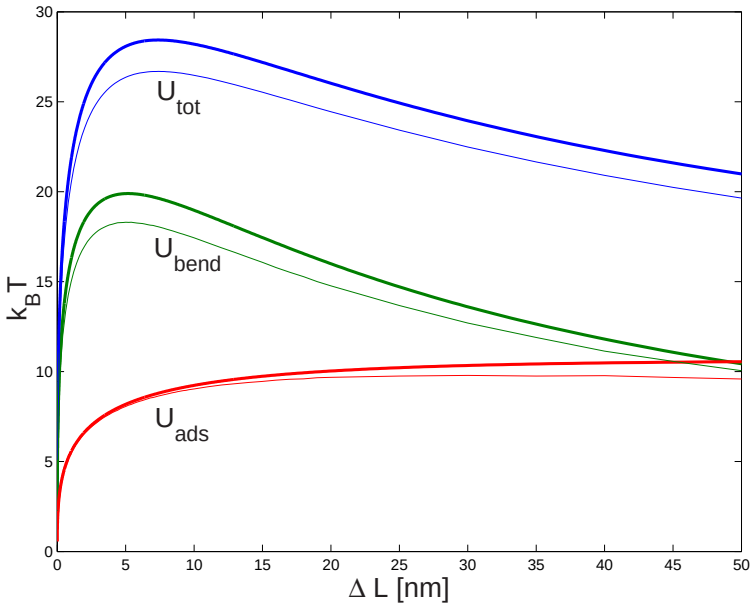


Figure 3.17: Comparison of the adsorption and bending energy contributions (U_{ads} and U_{bend}) as well as the total ground state energy U_{tot} of the simple loop. The fat lines represent the circle-line approximation (cf. Eq. 3.45) whereas the thin lines show the corresponding exact expressions, Eqs. 3.1 and 3.20 (thin line). The parameters are $\varepsilon_{ads} = 0.7 k_B T/nm$ and $A = 50nm \times k_B T$ and $R = 4nm$.

arrive at

$$U_{simp}(\alpha) = 2\alpha\left(A\frac{\tan\alpha-\alpha}{2R(\tan\alpha-\alpha)-\Delta L}+R\varepsilon_{ads}\right) \tag{3.45}$$

for simple loops and

$$\begin{aligned} U_{cross}(\alpha) &= 2\alpha\left(A\frac{\pi+\tan\alpha-\alpha}{\Delta L-2R(\tan\alpha-\alpha)}+R\varepsilon_{ads}\right) \\ &\quad +U_{def}(\alpha) \end{aligned} \tag{3.46}$$

for crossed loops where A, R and ε_{ads} defined as above and U_{def} being the excluded volume interaction at the crossing point, which is considered below (and applied in the main text as Eq. 3.26). We remark that the above expressions for U_{simp} and U_{cross} are valid within certain α intervals which are given by the restriction $0 < \alpha < \pi$ and by the condition that the first terms in the brackets of Eqs. 3.45 and 3.46 are positive (these are the necessarily positive bending energy contributions in the two cases.)
These fairly simple expressions can now be used in the two cases to obtain explicitly the ground state energies by minimizing Eq. 3.45 and Eq. 3.46 with respect to α. For instance, setting $U'_{simp}(\alpha) = 0$ we obtain a transcendental equation for α. We can now use the fact that this condition is algebraic in ΔL so that we can solve it for $\Delta L = \Delta L(\alpha)$. Thus instead of finding $\alpha = \alpha(L)$ (which cannot be given in an explicit form) we obtain explicitly its inverse:

$$\begin{aligned} \frac{\Delta L(\alpha)}{R} &= \frac{(2-c)G(\alpha)+cH(\alpha)}{1-c} \\ &\quad +\chi\frac{\sqrt{[(2-c)G(\alpha)+cH(\alpha)]^2-4(1-c)G^2(\alpha)}}{1-c} \end{aligned} \tag{3.47}$$

with the abbreviations

$$\begin{aligned} G(\alpha) &= \tan\alpha - a \\ H(\alpha) &= \alpha\tan^2(\alpha) \end{aligned}$$

In Eq. 3.47 the introduced dimensionless constant is $c = (1+2R^2\varepsilon_{ads}/A)^{-1}$ ($0 < c < 1$, and $c = 0.69$ here) and χ is the sign accounting for different branches of the α parametrized solution

$$\chi = \begin{cases} -1 \text{ for } 0 \le \alpha \le \pi/2 \\ \pm 1 \text{ for } \pi/2 \le \alpha \le \alpha_{\max}(c) \end{cases} \tag{3.48}$$

Note that for $\alpha \le \pi/2$ there is only one branch but for $\alpha > \pi/2$ we have two branches[10] (± 1) for $\Delta L(\alpha)$. The maximal opening angle $\alpha_{\max}(c)$ is obtained by setting the discriminant (expression below the square root) in Eq. 3.47 equal to 0.

[10]The latter means that for $\pi/2 \le \alpha \le \alpha_{\max}$ there are two different excess loop lengths leading to the same (equilibrium) angle α ,i.e., with increasing ΔL the nucleosome angle α opens but after passing some critical point on the ΔL axis, it starts closing again.

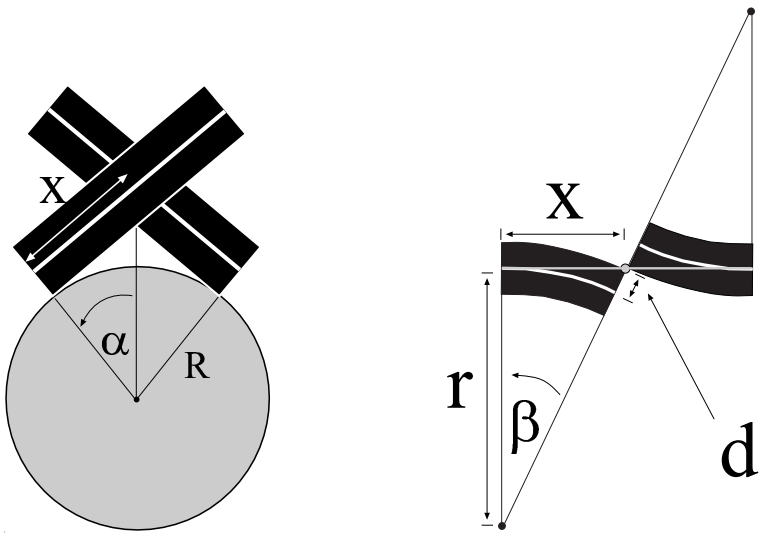


Figure 3.18: The out of plane deflection of the incoming/outgoing DNA due to excluded volume in the top projection (left) and seen from the side (right). In the latter case (for the sake of visual clarity) the two rods are depicted in a single plane, i.e., rotated around their contact point (grey dot).

From Eq. 3.47 together with Eq. 3.45 we obtain an explicit parametric representation of the minimal energy curve for simple loops. A comparison of the approximate minimal energies (Eq. 3.47 and Eq. 3.45) with the exact minimal energy (cf. also Fig. 3.5 for $\Delta L \lesssim 60nm$) is shown in Fig. 3.17. We find that the quantitative agreement is quite satisfactory taking the simplicity of our ansatz into account. We note here that analogous computations as we have shown for simple loops can be performed for crossed loops as well.

For $\Delta L \to 0$ we find after an appropriate expansion of U_{simp} around $\alpha = 0$ that the ground state energy scales as $U_{simp} \sim (\Delta L/R)^{1/3}$ in agreement with Ref. [69]. Further we obtain the excess length at which the loop ground state energy is maximal by setting $\partial U_{simp}(\alpha)/\partial\alpha|_{\alpha=\pi/2} = 0$. From this follows the critical length ΔL_{crit} as discussed in the main text (cf. Eq. 3.24). This simple approximate expression for ΔL_{crit} agrees within 2-15% with the exact numerical result for a wide range of adsorption energies with deviations becoming larger for adsorption energies above $\varepsilon_{ads} = 2.0\ k_BT/nm$ (data not shown).

3) The overcrossing potential for crossed loops (Fig. 3.18): The outgoing DNA path is perturbed out of the plane due to the interaction with the ingoing DNA (and vice versa in a symmetrical manner). Because of that our simple planar and phantom model (no self interaction) needs modifications. Instead of solving this (nonplanar) problem within the general theory of self-interacting deformed rods as in Ref. [86] (which is a feasible but rather technical numerical task) we can treat the out-of-plane deformation perturbationally. The first assumption we make here is that the overall shape of the crossed loop does not deviate much from a planar configuration though the orientation of its (effective) plane might be slightly deflected from the nucleosomal

plane. Consequently the small perturbation out of the plane and the deformation in plane essentially decouple into a sum of two energy contributions as in Eq.3.46. Again by simple geometry (cf. Fig. 3.17), the second (out of the plane) term in 3.46 can in first approximation be written as

$$U_{def}(\alpha) = \begin{cases} 2A\dfrac{d\arctan\left(\frac{2dx(\alpha)}{x^2(\alpha)-d^2}\right)}{x^2(\alpha)-d^2} \text{ for } x(\alpha) > d \\ \infty \text{ otherwise} \end{cases} \tag{3.49}$$

where $d \approx 1nm$ is the thickness of DNA and $x(\alpha) := R\tan\alpha$ the length of the crossed segment. In our simple approximation the self-interaction energy diverges for $x \longrightarrow d+0$ as $\frac{\pi}{2}A(x-d)^{-1}$ (extreme deformation) and approaches zero for $x \longrightarrow \infty$ as $4Ad^2x^{-3}$ (weak deformation).
We finally note that besides the above given examples it is possible to apply the circle-line approximation to several other standard problems of rod theory like the first and especially the higher order Euler buckling instabilities to obtain qualitatively the known results from buckling theory with very little effort. Thus the circle-line approximation when applied appropriately turns out to be very useful and generally allows computationally inexpensive qualitative and quantitative insights into the behavior of deformed rods.

Chapter 4

Kinetic Behavior of DNA Spools under Tension

Wrapped DNA-protein complexes are ubiquitous in nature [122] and play key roles in many fundamental life processes. Prominent examples of DNA wrapping proteins are: the Lac1 repressor [123] participating in the bacterial gene regulation, the DNA-gyrase [124] directing changes in DNA topology, RNA polymerase [125] copying DNA to RNA, and of course the histone octamer [126]. Besides the natural wrapped architectures there are attempts to design nanoparticles imitating that motive [127] as a means to efficiently pack and transport DNA into cells. In most of these ligand-DNA complexes the geometry and chemistry of the ligand surface enforces the DNA to follow a superhelical wrapping path with one or more tight turns. Remarkably, upon addition of multivalent condensing agents (like in sperm cells) or under high crowding conditions (like in virus capsids or during ψ-condensation) DNA also shows an intrinsic ability to self-organize into large toroidal spools [144].

In the past decade single molecule experiments have become available allowing to apply tension to individual polymers in order to probe their mechanical properties [129] as well as their interaction with ligands [130, 131, 132] and molecular motors [133]. Static and dynamic force spectroscopy [134, 135] developed into a powerful tool for measuring equilibrium as well as kinetic characteristics of single molecules, going far beyond the information accessible by classical bulk experiments. Application of these methods to DNA-spool geometries has been awaited for long and was reported only recently for single nucleosomes [136] and single DNA toroidal condensates [131, 132]. These experiments – at first glance completely unrelated – reside on different length and energy scales and ground on different mechanisms of wrapping. Despite that, they both reveal the same surprising result apparently contradicting all the available bulk data: the unfolding of wrapped DNA from the spools is a catastrophic event, i.e., it is sudden and quantized and happens one DNA turn at a time. The aim of this chapter is to theoretically explain this unusual non-equilibrium effect and to demonstrate the universality behind it. Our theory is then applied to nucleosomes and DNA toroids allowing to extract from experiments the relevant energetic parameters and to resolve apparent "oddities" in the dynamics of these systems.

4.1 The Model

In this section we introduce the geometry and energetics of a DNA spool under tension, cf. Fig. 4.1. We again model the DNA as a Worm-Like Chain(WLC) as in all previous chapters. The WLC can be considered a a semi-flexible tube characterized by two moduli, the bending and the torsional stiffnesses. The torsional stiffness will be neglected in the following since we consider the case of freely rotationg ends as in the experiments [136, 131, 132] to which we will compare our model in the following sections. Then the elastic energy of a WLC of length L and curvature κ (cf. Eq. 2.1) writes

$$E_{bend} = \frac{A}{2}\int_0^L ds\kappa^2(s) \tag{4.1}$$

The DNA is assumed to be adsorbed on the protein spool surface along a predefined helical path with radius R and pitch height H. This path accounts for the typical chemical structure of such a protein spool surface (e.g. distribution of charges, hydrogen donors/acceptors etc.). The DNA is adsorbed with a net adsorption energy density ε_{ads} given by the difference of the total DNA-protein binding energy and the stored DNA bending energy per length along the helical path.
The degree of DNA adsorption is described by the desorption angle α which is defined to be zero for one full turn wrapped (cf. Fig. 4.1). After short inspection it becomes clear that the unwrapping problem is non-planar and that the spool needs to rotate transiently out of the plane while performing a full turn – an effect already pointed out by Cui and Bustamante [137]. Therefore a second angle, β, is introduced to describe the out-of-plane tilting of the spool. When a tension F (along the Y–axis here) acts on the two outgoing DNA "arms" the system (i.e., the wrapped spool together with the free DNA ends) will simultaneously respond with (*i*) DNA deformation, with (*ii*) spool tilting and with (*iii*) DNA desorption from the spool.
The total energy of the system as a function of α and β has three contributions:

$$E_{tot} = 2R\varepsilon_{ads}\alpha + 2E_{bend} - 2F\Delta y \tag{4.2}$$

The first term in Eq. 4.2 is the adsorption energy, the second the bending deformation energy of the *two* free DNA portions, and the third term describes the gain in potential energy by pulling out the DNA ends, each by a distance Δy.
In order to proceed further we need to find the optimal shape of the DNA arms. We will not account for entropic shape fluctuations that decouple from the deformation state in our case (large forces F) as shown in chapter 2. For given boundary conditions (i.e. given values of the angles α and β) it is possible to find the optimal shape that minimizes the bending energy, Eq. 4.1, by applying the Kirchhoff analogy (introduced in the chapter 2) which relates stationary points of that WLC energy, to the well-studied classical mechanics problem of the trajectory of a symmetric spinning top in a gravity field [6]. For the twistless case under consideration this analogy reduces to that between planar untwisted rods, the Euler elastica, and the plane pendulum. One of the boundary conditions is that the DNA arms should be asymptotically straight. This leads us directly to the conformation that corresponds to the homoclinic orbit within the pendulum analogy ([6], cf. chapter 2). The natural parametric representation of a

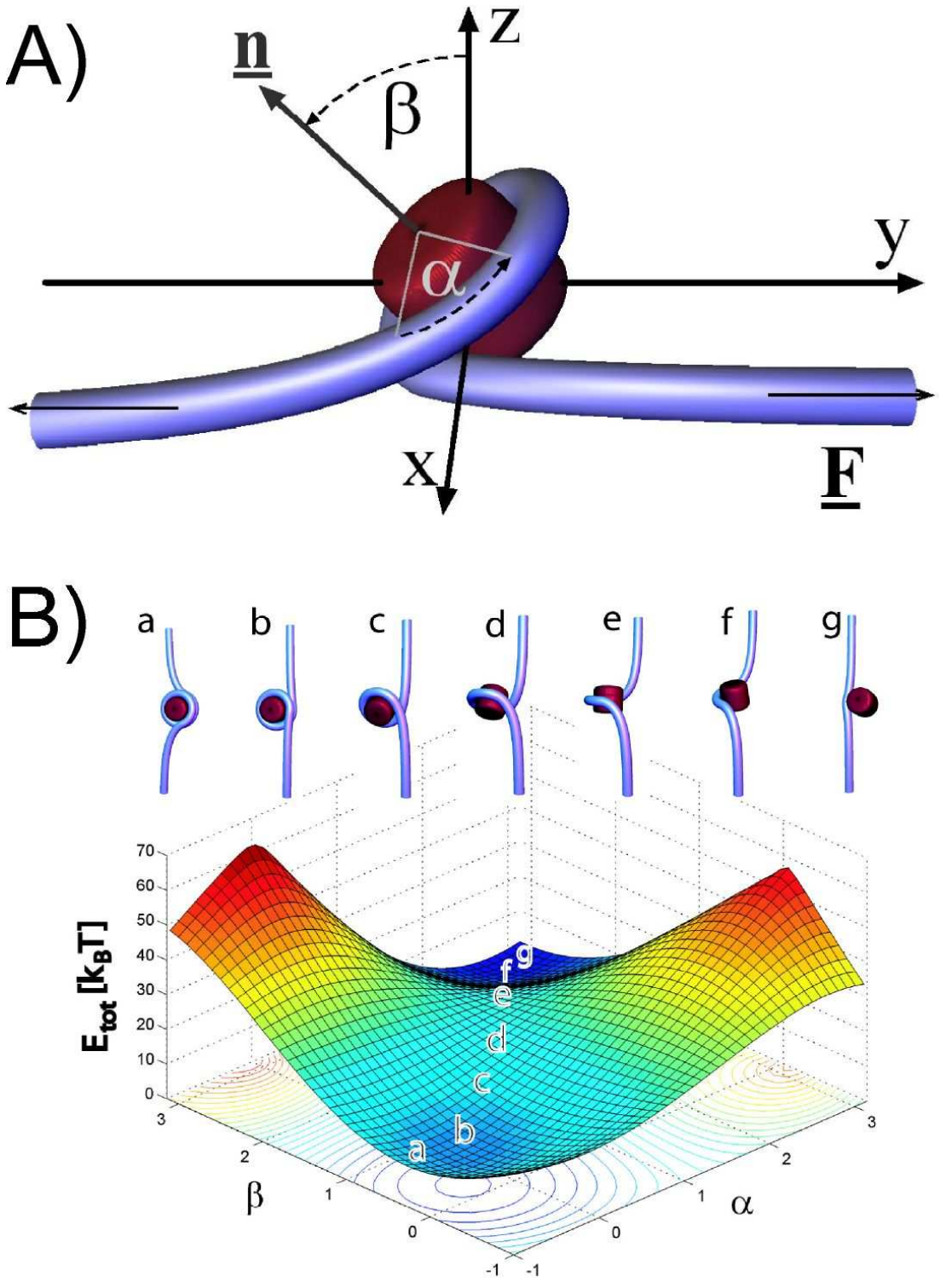


Figure 4.1: A) An unfolding DNA-protein spool under tension is characterized by two angles: the desorption angle α and the tilt angle β. B) The energy landscape of a DNA-spool (as given by Eq. 4.18 for $A = 50nm$, $F = 4pN$, $\varepsilon_{ads} = 0.7k_BT/nm$, $R = 4.2nm$, $H = 2.5nm$) and its unfolding pathway.

DNA arm within its plane (the $x-y$-plane which is in general tilted with respect to the X, Y, Z coordinates shown in Fig.4.1) is then given by

$$\begin{aligned} y(s) &= s - 2\lambda \tanh(s/\lambda) \\ x(s) - x(s_0) &= \frac{2\lambda}{\cosh(s/\lambda)} \end{aligned} \tag{4.3}$$

leading to the local curvature

$$\kappa(s) = \frac{1}{\lambda} \frac{2}{\cosh(s/\lambda)} \tag{4.4}$$

Here again (as in chapter 2) we introduce the tension length $\lambda = \sqrt{A/F}$. Note that the upper expressions describe nothing else but the homoclinic loop solution described in chapter 2 in depth [1]. The equations Eq. 4.3 and Eq. 4.3 are respectively the real-space (integrated tangent) coordinates and the curvature resulting from Eq. 2.27.
The bending energy of one arm is then given by

$$E_{bend} = \frac{A}{2} \int_{s_0}^{\infty} \kappa^2(s)\, ds = \frac{2A}{\lambda} (1 - \tanh(s_0/\lambda)) \tag{4.5}$$

The important offset parameter s_0 is related to the angle θ, the angle between the Y-axis and the tangent of the DNA at the point where it leaves the nucleosome. This angle is a function of α and β (see below). The boundary condition $(x'(s_0), y'(s_0)) = (\sin\theta, \cos\theta)$ leads to the following relation between s_0 and θ:

$$\tanh \frac{s_0}{\lambda} = \cos \frac{\theta}{2} \tag{4.6}$$

for $\theta < \pi$. This allows to rewrite the bending energy as

$$E_{bend}(\theta) = \frac{2A}{\lambda} \left(1 - \cos\frac{\theta}{2}\right) = 2\sqrt{AF} \left(1 - \cos\frac{\theta}{2}\right) \tag{4.7}$$

In the non-planar case we have a tilting of the spool normal $\underline{n}$ with respect to z-axis by an angle β as depicted in Fig. 4.1. Due to the symmetry in the problem the nucleosome dyad axis of the spool always coincides with the x-axis. Consequently $\underline{n}$ is always in the $Z-Y$-plane. The helical wrapping path of DNA on the spool and its tangent for an untilted geometry, $\beta = 0$, is as follows:

$$\underline{h}(t) = \begin{pmatrix} R\cos t \\ R\sin t \\ \frac{H}{2\pi}(\pi - t) \end{pmatrix} \quad \text{and} \quad \underline{h}_0'(t) = \frac{1}{\sqrt{R^2 + \left(\frac{H}{2\pi}\right)^2}} \begin{pmatrix} -R\sin t \\ R\cos t \\ -\frac{H}{4\pi} \end{pmatrix} \tag{4.8}$$

with $\alpha < t < 2\pi - \alpha$. To obtain the path for a non-vanishing value of β one has to tilt the spool by applying a rotation using the rotation matrix

[1] Note that the roles of the x and y axes are interchanged here with respect to the notation in chapter 2

$$\underline{\underline{\mathcal{R}}}(\beta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix} \tag{4.9}$$

This leads to the following helical path

$$\underline{h}(t,\beta) = \underline{\underline{\mathcal{R}}}(\beta)\,\underline{h}(t) = \begin{pmatrix} R\cos t \\ R\cos\beta\sin t - \frac{H}{2\pi}(\pi - t)\sin\beta \\ R\sin\beta\sin t + \frac{H}{2\pi}(\pi - t)\cos\beta \end{pmatrix} \tag{4.10}$$

and its tangent

$$\underline{h}_0^{'}(t,\beta) = \frac{\underline{\underline{\mathcal{R}}}(\beta)\,\underline{h}^{'}(t)}{\left\|\underline{\underline{\mathcal{R}}}(\beta)\,\underline{h}^{'}(t)\right\|} = \frac{1}{\overline{R}} \begin{pmatrix} -R\sin t \\ R\cos\beta\cos t + \frac{H}{2\pi}\sin\beta \\ R\sin\beta\cos t - \frac{H}{2\pi}\cos\beta \end{pmatrix} \tag{4.11}$$

where $\overline{R} = \sqrt{R^2 + \left(\frac{H}{2\pi}\right)^2}$.
We have now the helical wrapping path of the tilted spool in the $X-Y-Z$ coordinates, Eq. 4.10, and the shape of the DNA arms, Eq. 4.3, in its natural representation. In the following we have to "glue" the two DNA arms to the ends of the wrapped portions, i.e. at $t = \alpha$ and $t = 2\pi - \alpha$. This procedure has to fulfill the following three conditions:

1. The free arm attached at $t = \alpha$ is in the plane spanned by the helix tangent $\underline{h}_0^{'}(\alpha,\beta)$ (which coincides with the tangent vector of the arm at $s = s_0$) and the unit vector in Y-direction, $\underline{e}_y$. The angle θ from above is then given by

$$\cos\theta = \underline{h}_0^{'}(\alpha,\beta)\cdot\underline{e}_y = \frac{R}{\overline{R}}\cos\beta\cos a + \frac{H}{2\pi\overline{R}}\sin\beta \tag{4.12}$$

2. The offset of the rod at $t = \alpha$ is given by $\underline{h}(\alpha,\beta)$, i.e., the rod conformation can be written as

$$\begin{aligned} \underline{r}(s) &= \underline{h}(\alpha,\beta) + \left(\frac{2\lambda}{\cosh(s/\lambda)} - \frac{2\lambda}{\cosh(s_0/\lambda)}\right)\underline{h}_{\perp}^{'}(\alpha,\beta) \\ &\quad - \left((s - s_0) - 2\lambda\left(\tanh(s/\lambda) - \tanh(s_0/\lambda)\right)\right)\underline{e}_y \end{aligned} \tag{4.13}$$

   with $\underline{h}_{\perp}^{'}(\alpha,\beta)$ denoting the normalized orthogonal component of $\underline{h}_0^{'}(\alpha,\beta)$ with respect to $\underline{e}_y$. The other side of the DNA rod (starting at $t = 2\pi - \alpha$) is given by the upper solution reflected at the X-axis.

3. The torque must vanish for $s \to \infty$ on both sides. Therefore the Z-component of $\underline{r}(s)$ must vanish in that limit, i.e.

$$\underline{e}_z\cdot\underline{r}(\infty) = \underline{e}_z\cdot\left[\underline{h}(\alpha,\beta) - \left(\frac{2\lambda}{\cosh(s_0/\lambda)}\right)\underline{h}_{\perp}^{'}(\alpha,\beta)\right] = 0 \tag{4.14}$$

   Using Eqs. 4.6, 4.12 and 4.13 this condition leads to relation between the applied tension F and the angles α and β

$$\frac{\sqrt{2}\lambda}{R}\sqrt{\frac{1-\left(\frac{R}{\overline{R}}\cos\beta\cos\alpha+\frac{H}{2\pi\overline{R}}\sin\beta\right)}{\sin^2\alpha+\left(\sin\beta\cos\alpha-\frac{H}{2\pi R}\cos\beta\right)^2}}\left(\sin\beta\cos\alpha-\frac{H}{2\pi R}\cos\beta\right)$$
$$= \sin\beta\sin\alpha+\frac{H}{2\pi R}(\pi-\alpha)\cos\beta \tag{4.15}$$

If the wrapped portion would be fixed, i.e. for a given value of α, then Eq. 4.15 is an implicit equation giving the spool tilting angle β as a function of F.
From Eqs. 4.7 and 4.12 we obtain the bending energy as a function of α and β:

$$E_{bend}(\alpha,\beta)=2\sqrt{AF}\left(1-\sqrt{\frac{1+\left(\frac{R}{\overline{R}}\cos\beta\cos\alpha+\frac{H}{2\pi\overline{R}}\sin\beta\right)}{2}}\right) \tag{4.16}$$

Using the explicit shape of the DNA arms, Eq. 4.13, we are able to calculate the pullout length Δy which is the last term missing in Eq. 4.2. We obtain

$$\Delta y = R\alpha-\underline{h}(\alpha,\beta)\cdot\underline{e}_y+2\lambda\left(\cos\frac{\theta(\alpha,\beta)}{2}-1\right) \tag{4.17}$$

Here the first term is the amount of DNA released by unpeeling of a DNA arm, the second accounts for the Y-position of the point where the DNA leaves the spool and, finally, the third term describes the wasted length by bending of the free DNA. That term is zero for vanishing chain stiffness, $A=0$, when the conformation of the DNA arm is straight and points parallel to the Y-axis.
We are now in the position to present the total energy, Eq. 4.2. From Eqs. 4.16 and 4.17 follows

$$\begin{aligned}E_{tot}(\alpha,\beta) &= 2R\varepsilon_{ads}\alpha+2FR\left(\cos\beta\sin\alpha-\frac{H}{2\pi R}(\pi-\alpha)\sin\beta-\alpha\right) \\ &\quad +8\sqrt{AF}\left(1-\sqrt{\left(1+\frac{R}{\overline{R}}\cos\beta\cos\alpha+\frac{H}{2\pi\overline{R}}\sin\beta\right)/2}\right)\end{aligned} \tag{4.18}$$

This is the central formula derived here. The first term in Eq. 4.18 describes the cost of desorption due to unpeeling of the wrapped chain portion. The second term describes the gain/loss of potential energy by spool opening (change of α) and rotation (change of β). Finally, the last and most remarkable term accounts for the stiffness of the non-adsorbed DNA portions. Two effects contribute equally to this term: (*i*) the bending energy of the deformed DNA arms, Eq. 4.16, and (*ii*) the loss of potential energy by "wasting" length due to DNA deformation, third term of Eq. 4.17. Eq. 4.18 gives the total energy of the spool under a given tension F as a function of α and β. In general this does not lead to torque-free geometries that fullfill constraint 4.14. However, it can be shown that this constraint leads to the energy minimizing value of the tilting angle β for given values of α and F, i.e. $\partial E_{tot}(\alpha,\beta)/\partial\beta=0$.
To understand the implications of Eq. 4.18 on the kinetics of unwrapping we consider two limiting cases. First let us look at the case of a large thin spool, i.e., $R\gg A/k_BT$

(or, equivalently, an infinitely flexible polymer) and $R \gg H$, where we may neglect all the contributions that stem from bending of the two arms. Then Eq. 4.18 simplifies to

$$E_{tot}(\alpha, \beta) = 2R\varepsilon_{ads}\alpha + 2FR\cos\beta\sin\alpha \tag{4.19}$$

For $F > \varepsilon_{ads}$ the spool moves from the (thermodynamically) metastable state M_1 with $\alpha = \alpha_0 = -\arccos(1 - \varepsilon_{ads}/F)$ and $\beta = 0$ via a saddle point S at $\alpha = 0$ and $\beta = -\alpha_0$ into a more favorable minimum M_2 at $\alpha = \pi + \alpha_0$ and $\beta = \pi$. Remarkably S constitutes a significant energetic barrier between M_1 and M_2, namely

$$\Delta E_{tot} = 2FR(\alpha_0\cos\alpha_0 - \sin\alpha_0) \tag{4.20}$$

This barrier has a tremendous effect on the unwrapping kinetics: For reasonable parameter values, say $R = 50nm$, $\varepsilon_{ads} = 1k_BT/nm$ and $F = 2\varepsilon_{ads}$, one finds a huge barrier of $\Delta E_{tot} \approx 70k_BT$.
A second interesting limit of Eq. 4.18 is given by a flat spool and high polymer stiffness, i.e., $A \gg Rk_BT$ and $R \gg H$. For not too large forces ($F \lesssim A/R^2$) and $\varepsilon_{ads} \lesssim F$ the kinetic behavior is roughly dominated by the term E_{stiff}. In this case we find a transition path from $(\alpha, \beta) = (0, 0)$ over the saddle point $(\pi/2, \pi/2)$ to the state (π, π) with a barrier height

$$\Delta E_{tot} = 8\sqrt{AF}\left(1 - 1/\sqrt{2}\right) \tag{4.21}$$

Note that in this limit the DNA actively participates in the suppression of unwrapping ($\Delta E_{tot} \sim A^{1/2}F^{1/2}$) which can even give rise to negative resistance effects [139] for small forces. In Appendix A we present a simpler geometry where the cylindrical spool is replaced by a short rod. In that case the effect of DNA stiffness can be understood more clearly.
In preliminary conclusion, in both limiting cases the unwrapping meets significant kinetic barriers but for different reasons: because of unfavorable projection of the force in terms of the (α, β) configurational space in the first limit and due to significant transient bending of the DNA arms during the transition in the second limit. For realistic DNA-spools we are somewhere in between these two cases.

4.2 Applications

4.2.1 Nucleosomes under Tension

The most abundant DNA spool in nature is the nucleosome. The question about the equilibrium and kinetic stability of nucleosomes is one of the important experimentally unsettled questions in present molecular biology. How can nucleosomes be highly stable with its wrapped DNA being highly accessible at the same time [140]? A recently performed experiment [136] (cf. Fig. 4.2) measuring the critical force required to unwrap single nucleosomes reveals an interesting and unexpected behavior[2]. When small

[2]The experiments were performed on DNA chains with up to 17 nucleosomes complexed at well-defined positions. In the force range of interest their coupling can be safely neglected since the inter-nucleosomal distance $d \sim 40nm$ exceeds the DNA-linker induced interaction length $\sim \lambda = (A/F)^{1/2}$.

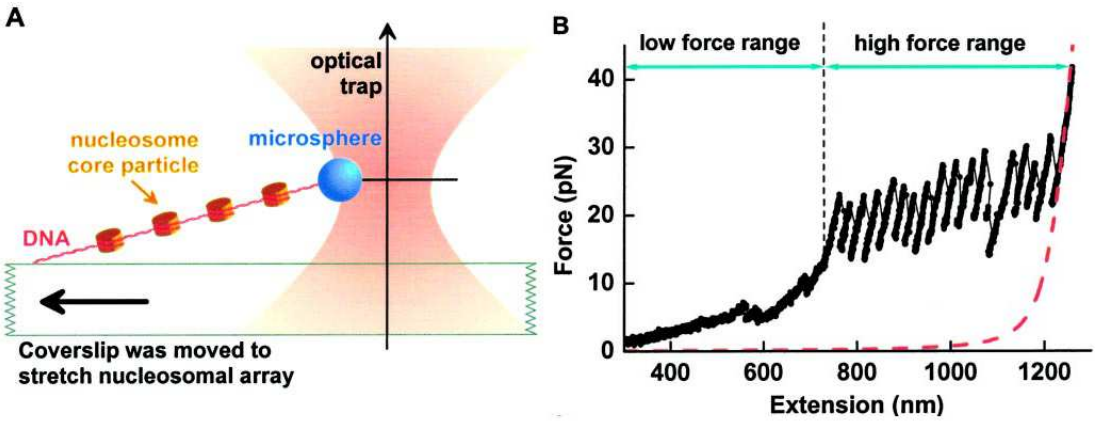


Figure 4.2: Unwrapping nucleosoms: A) The schematic experimental setup: In the actual experiment 17 nucleosoms were bound to the DNA. B) A typical experimental force-extension curve: the end-to-end distance is increased at constant speed (20 nm/s) and force is measured. The 17 peaks correspond exactly to the 17 nucleosomes placed on the DNA sample. The dashed curve corresponds to the same DNA without nucleosomes. Picture adapted from Brower-Toland et. al. [136].

forces ($F < 10pN$) are applied for short times ($\sim 1 - 10\ s$) the nucleosome unwraps only partially by releasing the outer 60-70 bp of wrapped DNA (moving from state a to b in Fig. 4.1) in a gradual and equilibrium fashion. For higher forces ($F \gtrapprox 20pN$) nucleosomes show a pronounced sudden non-equilibrium release behavior of the remaining 80 bp (cf. c-g in Fig. 4.1B) – the latter force being much larger than expected from equilibrium arguments [141]. In fact, experiments [140] measuring spontaneous partial unwrapping of nucleosomal DNA suggest $30k_BT$ per 147bp leading to an unpeeling force of $\sim 2.5pN$. To explain this peculiar finding Brower-Toland et al. [136] conjectured that there must be a barrier in the adsorption energy located after the first 70-80 bp which reflects some biochemical specificity of the nucleosome structure at that position. Their analysis of the dynamical force spectroscopy measurements revealed an apparent barrier of $\sim 38k_BT$ smeared out over not more than 10 bp. However, there is no experimental indication of such a huge specific barrier – neither from the crystal structure [126] nor from the equilibrium accessibility to nucleosomal DNA [140]. Consequently the question arises if the barrier is really caused by biochemistry of the nucleosome or, as we show below, by its underlying geometry and physics.
To see that the effect is mainly physical we apply Eq. 4.18 to compute the barrier. For this purpose we model the nucleosomal adsorption energy density as

$$\varepsilon_{ads}(\alpha) = \varepsilon^0_{ads} + H(\alpha)\,\varepsilon_{es}$$

where $\varepsilon^0_{ads} \approx 0.7k_BT/nm$ is taken from the reversible part (for the first 60-70bp) of the measurement in Ref. [136]. The introduction of the step function ($H = 0$ for $\alpha < 0$ and $H = 1$ for $\alpha \geq 0$) together with a new parameter ε_{es}, the electrostatic interaction energy density, accounts for the DNA-DNA repulsion of the two adjacent helical gyres which acts only for $\alpha < 0$ reducing the net ε_{ads}. Using ε^0_{ads} from above

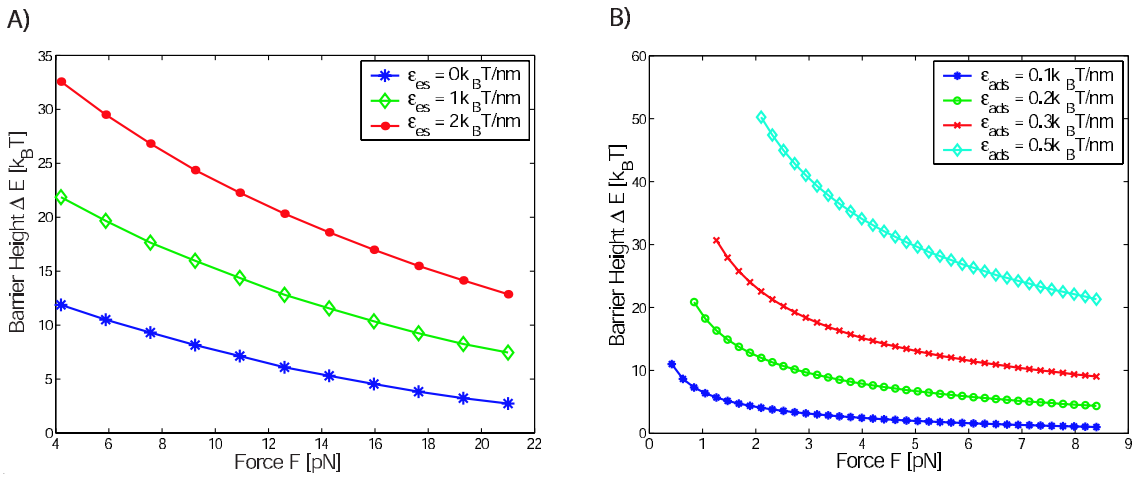


Figure 4.3: Kinetic barriers opposing DNA-spool unfolding as function of applied tension computed for: A) The nucleosome ($R = 4.2nm$, $H = 2.4nm$, $\varepsilon^0_{ads} = 0.7k_BT/nm$, $A = 50k_BTnm$,cf. text) for various interstrand repulsion energy densities ε_{es} and B) the DNA toroid ($R = 50nm$, $H = 2.4nm$, $A = 40k_BTnm$) for various adsorption energy densities ε_{ads}.

we can compute the barrier height for nucleosome unfolding for various values of ε_{es} as done in Fig. 4.3A. To relate the barrier heights from Fig. 4.3A to the dynamical force spectroscopy (DFS) measurements in Ref. [136] we generalize the classical relation between the loading rate r_F and the most probable rupture force F^* [134, 135] to the case of nonlinear force-barrier dependence and obtain (cf. Appendix)

$$\ln\left(r_F/r_0\right) = \ln\left[-\nu_{att}k_BT/\left(r_0\Delta E'\right)\right] - \Delta E/k_BT \tag{4.22}$$

Here r_F and ΔE are functions of F^* and $\Delta E' = \partial\left(\Delta E\right)/\partial F^*$. $r_0 = 1pNs^{-1}$ is an arbitrary scale on the r_F axis and ν_{att} is the typical attempt frequency of the nucleosome. Assuming ν_{att} in the range 10^5 to $10^8 s^{-1}$ we can fit the experimental data from [136] to obtain the corresponding values of ε_{es}, cf. Fig. 4.4. The attempt frequency ν_{att} is dominated by the slowest process involved in the unfolding event. The rotational attempt frequency of a nucleosome-sized sphere is of the order $2\frac{k_BT}{8\pi R^3\eta_s} \approx 10^5 - 10^6 s^{-1}$ with η_s being the water viscosity (a centipoise). The typical frequency that characterizes the relaxation of the DNA arms is comparable to that – even if one accounts for additional complexed nuclesomes as it is the case in the experiment [136]. There the first unfolding nucleosome is surrounded by 16 other nucleosomes that have to move via a distance $\Delta s \approx 25nm$ (unfolding length) under a force of $F \approx 10 - 20pN$ leading to an lower bound $10^5 - 10^6 s^{-1}$ of that frequency. Hence we estimate $\nu_{att} \lessapprox 10^6 s^{-1}$. The latter implies $\varepsilon_{es} \approx 1.4 - 1.7k_BT/nm$. So indeed, at the line $\alpha = 0$ there is clear jump in adsorption energy density as we would naively expect from repulsive DNA-DNA electrostatics under these conditions[3]. This explains why under equilibrium conditions (at $F = 0$) the DNA deeply inside the nucleosomes (almost the whole bound

[3]Note that the latter alone would not explain the sudden catastrophic behavior and the slow kinetics of unfolding if the effects described by Eq. 4.18 were not included explicitly.

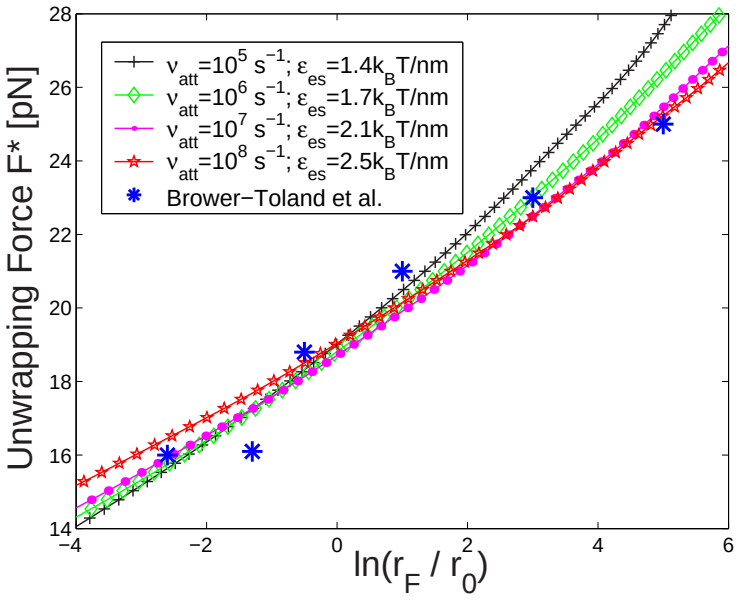


Figure 4.4: Optimal fits of the DFS data from Ref. [136] for various attempt frequencies give the corresponding electrostatic DNA-DNA-repulsion ε_{es}.

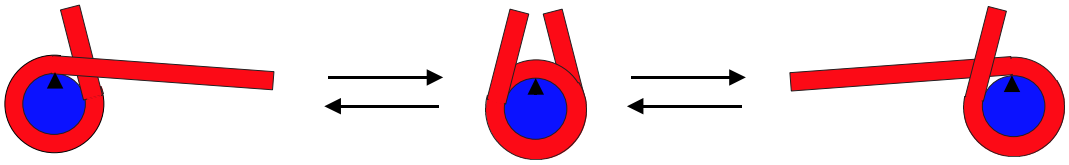


Figure 4.5: The reason for the nucleosome being highly stable yet highly accessible: One sided DNA unpeeling is assisted by repulsive DNA-DNA interaction, leading to high DNA accessibility on both sides. Further opening beyond that point (e.g. simultaneous opening) becomes energetically much more costly, leading to high nucleosome stability.

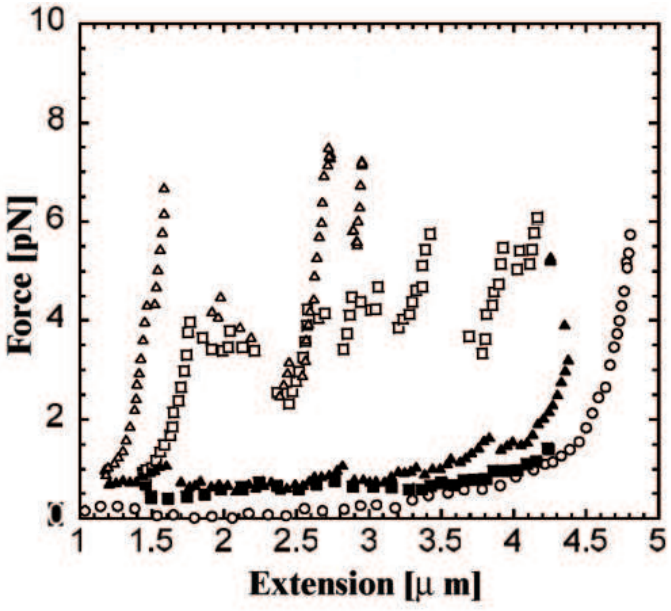


Figure 4.6: Typical force-extension behavior of spermidine condensed DNA from Ref. [132]. The typical distance between peaks is a multiple of 300 nm indicating a regular toroidal organization of the condensate.

DNA!) can be rather easily accessed by proteins [140] but the nucleosome is still highly stable: The line $\alpha = 0$ can be moved to each position inside the nucleosome if the left and right DNA arms are adsorbed/desorbed in a consistent manner, cf. Fig. 4.5. Beyond that ($\alpha > 0$) the assisting electrostatics switches off and the nucleosome is suddenly strongly stabilized (by $60 - 70 k_B T$ in total!).

4.2.2 DNA Toroids under Tension

When long DNA molecules condensed with multivalent counterions were stretched in a single molecule experiment [131] Baumann et al. found a surprising behavior. When a critical force (typically $F \approx 4 - 12 pN$) is reached large portions of DNA are released in packets in a discontinuous manner ("stick release pattern"). When the same experiment was redone recently by another group [132] a pronounced quantization in the DNA release length of $\approx 300nm$ was clearly demonstrated. It was noted in Refs. [131, 132] that the latter correlates exactly with the typical size ($R \approx 50nm$) of toroidal condensates formed in solution and led those authors to the conclusion that a single turn of DNA unwraps from the toroid spool at a time. Despite that interesting finding the mechanism behind this non-equilibrium effect remained unexplained. However in the light of our theory the explanation might be again straightforward as a DNA-toroid exhibits a spool geometry with $R \approx 50nm$ and $H \ll R$. The "first limit" in Eq. 4.20 considered above gives here a good approximation. The barrier heights for different values of ε_{ads} as a function of force are computed in Fig. 4.3B. Similar as in the case of the nucleosome the attempt frequency ν_{att} is dominated by the rotational friction –

here of the $50nm$ sized toroid object – leading to $\nu_{att} \approx 3 \times (10^2 - 10^3)\, s^{-1}$. For high concentrations of the condensing agent spermidine one finds $\varepsilon_{ads} \approx 0.2 - 0.3 k_B T/nm$ (cf. [131, 132] and the references therein). In case of equilibrium this means a very small peeling-off force of $F \approx 1 - 1.5pN$. Our model together with Fig. 4.3 allows us now to predict the activated non-equilibrium behavior to have very low unfolding frequencies $\nu_{unf} = v_{att} \exp\left(-\Delta E\left(F\right)/k_B T\right)$, for instance 10^{-6} to $10^{-2} s^{-1}$ for $F = 2pN$, 10^{-3} to $1 s^{-1}$ for $F = 4pN$ and 0.3 to 50 s^{-1} for $F = 8pN$ [4] – consistent with experimental findings [131, 132].

4.3 Conclusion

We have shown that DNA-spools ranging from protein-DNA complexes to DNA toroids share a universal feature inherited by their geometry: They are strongly kinetically protected from mechanical disruption upon applied tension. In the case of chromatin fibers consisting of large arrays of nucleosomes and other DNA spooling proteins this effect provides a great biological advantage. Strong molecular motors like RNA polymerase and helicase or microtubuli during cell division are known to act on the fiber with significant transient tensions of the order of $20pN$ or even more. While a hypothetical "fiber A" consisting of DNA and non-spooling proteins (say only DNA bending proteins) would immediately lose most of its protein content a "fiber B" constituted of DNA-spools would survive long time periods (up to $10^6 - 10^8$ times longer than "fiber A"). We can speculate that this obvious advantage was not overlooked by nature and has flown into the chromatin fiber design and the nucleosome-spool shape. The remarkable universality of the "kinetic protection" also shows up in the case of DNA-toroids which are roughly ~ 10 times larger while the DNA is ~ 10 times weaker adsorbed than for typical DNA-protein spools. While the biological implications of this finding still have to be fully figured out it seems that this might play a role in the injection / ejection process of DNA from viral capsids inducing similar quantization effects as found here. Looking at the wealth of peculiar effects revealed by the single molecule experiments [131, 132] we feel that the present understanding of DNA condensation kinetics is still incomplete, yet one partial mystery seems resolved.

4.4 Appendix

4.4.1 Simplified geometry

We discuss here a related yet much simpler system where the subtle interplay of bending and potential energy shows up more transparently and causes an elastic instability. Consider the two DNA arms attached to a stiff rod of length L as depicted in Fig. 4.7. Each arm is attached to one arm at an end of the rod with the constraint that the tangent of the rod and the attached DNA are always parallel. This leads to the condition

[4] The upper/lower estimate correspond to $\varepsilon_{ads} = 0.2$ / $0.3 k_B T/nm$; note the strong dependence of ν_{unf} on ε_{ads}.

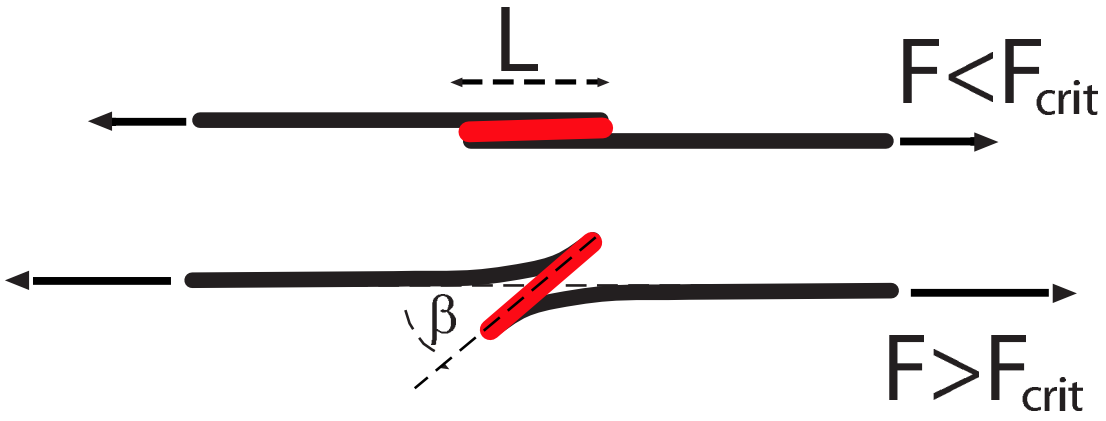


Figure 4.7: Two semiflexible rods are glued tangentially onto a short stiff rod (in red) of length L. For small forces the straight configuration $\beta = 0$ is stable. After crossing a critical force F_{crit} the system goes through a bifurcation and configuration starts to tilt developing a nonzero angle β.

$$\tanh \frac{s_0}{\lambda} = \cos \frac{\beta}{2} \tag{4.23}$$

This relation corresponds to Eq. 4.6 in the spool case. Eq. 4.23 can be rewritten as follow $\cosh^{-2}(s_0/\lambda) = \sin^2(\beta/2)$. Again we require vanishing torque on the structure leading to the condition

$$\frac{1}{\cosh^2(s_0/\lambda)} = \left(\frac{L}{4\lambda}\right)^2 \sin^2 \beta \tag{4.24}$$

Combining Eqs. 4.23 and 4.24 we arrive at the equilibrium condition

$$(1 - \cos\beta)\left(2\left(\frac{L}{4\lambda}\right)^2 (1 + \cos\beta) - 1\right) = 0 \tag{4.25}$$

There are three solutions to Eq. 4.25: $\beta = 0$, $\beta = \pi$ and

$$\cos\beta = \frac{8A}{FL^2} - 1 \tag{4.26}$$

For small tensions Eq. 4.26 has no solution and the rod remains untilted, $\beta = 0$ even though a tension is applied. When a critical force

$$F_{crit} = 4A/L^2 \tag{4.27}$$

is reached the rod begins to rotate in either direction with β given by Eq. 4.26.
This indicates that there is an instability which can also be seen by inspecting the energy of the system. It has two contributions: $E_{tot} = 2E_{bend} - 2F\Delta y$. E_{bend} is again the bending energy per DNA arm, Eq. 4.7 with $\theta = \beta$. The length Δy in second term,

the potential energy, is the sum of two effects: the gain by rod rotation and the cost by "wasting" length due to chain bending

$$\Delta y = L\left(1-\cos\beta\right) + 4\lambda\left(\cos\frac{\beta}{2}-1\right) \tag{4.28}$$

The total energy writes then

$$E_{tot} = 8\sqrt{AF}\left(1-\cos\frac{\beta}{2}\right) - FL\left(1-\cos\beta\right) \tag{4.29}$$

The pitchfork bifurcation at F_{crit} can indeed be recovered for the small β-expansion of Eq. 4.29:

$$E_{tot}\left(\beta\right) \overset{\beta\to 0}{\to} \left(\sqrt{AF}-\frac{1}{2}FL\right)\beta^2 + \frac{1}{48}\left(2FL-\sqrt{AF}\right)\beta^4 \tag{4.30}$$

Let us study how this tilting instability is affected by a finite height H between the two DNA strands (i.e. we assign a thickness to the stiff red segment). For that case the total energy can be calculated in a similar fashion as Eq. 4.29

$$E_{tot} = 8\sqrt{AF}\left(1-\cos\frac{\beta}{2}\right) - FL\left(1-\cos\beta\right) - FH\sin\beta \tag{4.31}$$

For small deformations $\beta << 1$ we find

$$E_{tot}\left(\beta\right) \approx \left(-FH\right)\beta + \left(\sqrt{AF}-\frac{1}{2}FL\right)\beta^2 + \frac{1}{6}FH\beta^3 + \frac{1}{48}\left(2FL-\sqrt{AF}\right)\beta^4 \tag{4.32}$$

For $H << L$ we can use Eq. 4.32 to investigate what happend to the pitchfork bifurcation in the case of finite H. In fact, for small values of F we have only one minimum at $\beta \approx FH/\left(2\sqrt{AF}-FL\right)$. At $F_{crit} \approx 4AL^2/\left(L^2+2H^2/3\right)^2$ we encounter a saddle point bifurcation with the middle branch (closest to the $\beta \approx 0$ line) being unstable, the other two (the upper and the lowest) stable.

4.4.2 Dynamical Force Spectroscopy

We give here a short discussion of dynamical force spectroscopy [134, 135]. This method allows to estimate the barrier which has be overcome during the unbinding of non-covalent bonds under the application of an external force. In the present study the bound state corresponds to that of a wrapped spool with N turns, the unbound state to that of a spool with one turn less. The rupture rate of the bound state under a force F with the barrier height E_b and its distance (along the force direction) from the bound state x_b is given by the Kramers rate

$$\nu_{fail}\left(F\right) \approx \nu_0\left(F\right) e^{-\left(E_b - F x_b(F)\right)/k_B T} \tag{4.33}$$

If we assume a very steep potential, the force F changes neither the attempt frequency ν_0 nor the position of the maximum x_b. In that case we may write $\nu_0\left(F\right) \approx \nu_0\left(0\right) = \nu_0$ and $x_b\left(F\right) \approx x_b\left(0\right) = x_b$ and

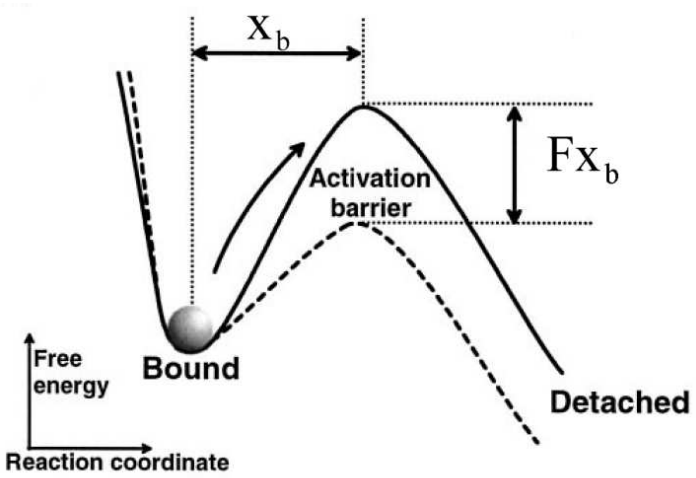


Figure 4.8: The principle of a dynamical force spectroscopy (DFS) measurement: The barrier hight (at a distance x_b from the bound state) and the Kramers' escape rate is modulated by an applied force F. By applying a time varying force the hight and the distance x_b of the barrier can be recovered from the typical unbinding times/forces. Picture adopted from [136]

$$\nu_{fail}\left(F\right) \approx \nu_0 e^{-(E_b - F x_b)/k_B T} \tag{4.34}$$

Assume that the force varies with time $F = F\left(t\right)$. The probability for bond failure in a small time interval $[t, t + dt]$ is then given by

$$w_{fail}\left(t\right) dt = \nu_{fail}\left(F\left(t\right)\right) P_{surv}\left(t\right) dt \tag{4.35}$$

with

$$P_{surv}\left(t\right) = \exp\left(-\int_0^t \nu_{fail}\left(F\left(t'\right)\right) dt'\right) \tag{4.36}$$

The first factor in Eq. 4.35 is the dissociation rate at time t that has to be multiplied by the probability that the bond survives until the time t, i.e. by $P_{surv}\left(t\right)$.
Now assume that the applied force increases linearly in time $F\left(t\right) = r_F t$ with a loading rate r_F. Then the typical quantity to determine is the maximum value of $w_{fail}\left(t\right) = \frac{d}{dt}\left(1 - P_{surv}\left(t\right)\right)$ which gives the peak in the distribution of unbinding forces. This leads to the condition

$$\frac{d^2 \exp\left(-\nu_b \rho_F\left(t\right)\right)}{dt^2} = 0 \tag{4.37}$$

with $\nu_b = \nu_0 e^{-\beta E_b}$ and $\rho_F\left(t\right) = \int_0^t e^{\beta x_b F(t')} dt'$. This can be rewritten as follows

$$\frac{\rho_F''\left(t\right)}{\rho_F'^2\left(t\right)} = \nu_b \tag{4.38}$$

In the case of a linear force ramp $\rho_F(t)$ can be computed $\rho_F(t) = \int_0^t \exp(\beta x_b r_F \tau)\, d\tau = \frac{1}{\beta x_b r_F}\left(e^{\beta x_b r_F t} - 1\right)$ ($\beta = 1/k_B T$). From Eq. 4.38 (and using $F = r_F t$) follows then the peak in the rupture force distribution:

$$F^* = \frac{1}{\beta x_b} \ln \frac{\beta x_b r_F}{\nu_0} + \frac{E_b}{x_b} \tag{4.39}$$

The latter is the known Evans-Ritchie result used in virtually all DFS experiments. It is usually rewritten as

$$F^*(r_F) = \frac{1}{\beta x_b} \ln \frac{r_F}{r_0} + \frac{1}{\beta x_b} \ln \frac{r_0 \beta x_b}{\nu_0} + \frac{E_b}{x_b} \tag{4.40}$$

with $r_0 = 1\ p\mathrm{N}/\mathrm{s}$ being some arbitrary loading scale. Note that only the first term in Eq. 4.40 depends on the loading rate r_F whereas the other terms are constant.
Having measured $F^*(r_F)$ over a range of loading rates, the recipe for computation of x_b and E_b is obvious: (1) Obtain x_b from the slope and (2) compute E_b from the offset constant. For step 3 molecular attempt frequency ν_0 (usually $\sim 10^9 - 10^{10}\,\mathrm{s}^{-1}$) is needed that cannot be obtained from $F^*(r_F)$.
When comparing our theoretical results to the experimental force spectroscopy plot extracted from nucleosome unwrapping in Ref. [136] we will make use of a generalized version of Eq. 4.40 that we derive in the following. This takes into account that $\nu_{fail}(F)$ shows in general a more complex dependence on F than assumed in Eq. 4.34. In this case the condition for the peak F^* in the force distribution takes now the form $\nu'_{fail}(r_F t_m) = \nu^2_{fail}(r_F t_m)$

$$\nu^{-2}_{fail}(F^*) \frac{d\nu_{fail}(F^*)}{dF} = \frac{1}{r_F} \tag{4.41}$$

For the special case $\nu_{fail}(F) = \nu_0 e^{-(E_b - F x_b)/k_B T}$ we recover Eq. 4.39.
On the other hand, for some arbitrary dependence of the barrier height U on F the failure rate is of the form $\nu_{fail}(F) = \nu_0 \exp(-U(F)/k_B T)$. Using Eq. 4.41 leads to a generalized relation between r_F and F^*:

$$\ln \frac{r_F(F^*)}{r_0} = \ln\left(-\frac{\nu_0 k_B T}{r_0 U'(F^*)}\right) - \frac{U(F^*)}{k_B T} \tag{4.42}$$

In this case the force spectroscopy plot F^* vs. $\ln r_F$ in general does not lead to a linear relation between these quantities. This can indeed be seen in Fig. 4.4 where we fit the experimental plot with nonlinear curves $F^*(r_F)$.

Chapter 5

The Internal Structure of Toroidal DNA Condensates

5.1 The Experimental Facts

Single polymers collapse from a random coil conformation to a dense state once the solvent gets sufficiently poor [142]. For a flexible chain the condition of minimal surface energy yields an approximately spherical globule, but for semiflexible polymers the situation is more complex [143]: The local structure of the dense phase then consists of essentially straight chains with a basically parallel alignment, in order to minimize bending energy and maximize density, respectively. Such a state can be characterized by a smooth field of tangent vectors, but in the spherical case this field must have at least two energetically unfavorable defects on the surface (one can't comb a sphere). However, for a *torus* many defect-free fields are possible. Indeed, DNA, the probably best studied semiflexible polymer, readily forms beautiful nanotori after adding any one of a variety of possible condensing agents (like polyethylenglycol (PEG), multivalent counterions, or bundling proteins) to a dilute solution of DNA chains [144]. These tori are surprisingly monodisperse, having a radius comparable to the persistence length of DNA ($\ell_\mathrm{p} \approx 50\,\mathrm{nm}$) basically independent of the condensation method [144].
Consider such a condensate, in which the chain is wound up like a garden hose to form a torus with axial and tubular radii r_1 and r_2, respectively. Since r_1 is the average radius of curvature of the chain, a simple scaling analysis balancing a bending energy A/r_1^2 per unit length of polymer, where $A = k_\mathrm{B}T\,\ell_\mathrm{p}$ is the bending modulus of a semiflexible chain, and a surface energy σ per unit area of the torus yields [143] $r_1 \sim (\sigma/A)^{-2/5}V^{1/5}$ and $r_2 \sim (\sigma/A)^{1/5}V^{2/5}$, where the chain volume $V \sim r_1 r_2^2$ as well as the packing density are assumed constant. Hence, the aspect ratio $\xi = r_1/r_2 \sim (\sigma/A)^{-3/5}V^{-1/5}$ shrinks if σ or V increase (*i. e.*, if the solvent gets poorer or the chain longer), and thus the torus "fattens". In this case it is no longer justified to calculate the bending energy with some average radius of curvature $\langle\rho\rangle = r_1$. In fact, since $\langle\rho^{-2}\rangle \geq \langle\rho\rangle^{-2}$ (by virtue of Jensen's inequality [145]), the actual curvature energy should be larger. However, the same argument indicates that the condensate can *lower* its bending energy by redistributing curvature more evenly. In this chapter we demonstrate that indeed below a critical aspect ratio ξ_c (or above a critical surface tension σ_c) the system spontaneously relaxes

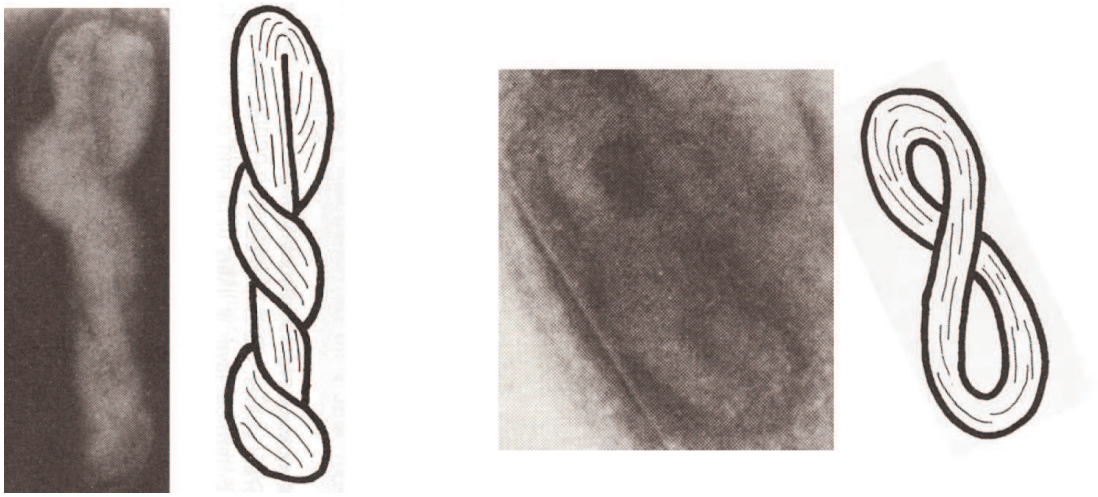


Figure 5.1: Supercoiled toroids formed after DNA release from mutated page T4 heads. Different toroids show either left- or right-handed supercoiling, indicating a spontaneous symmetry breaking mechanism. Pictures adopted from ref. [150]

bending energy by *twisting* the bundle of polymer strands. In other words, given that the overall *shape* of the condensate is a torus, its internal *structure* still depends on further elastic details. However, the underlying mechanism is very generic and thus independent of specific microscopic details of the polymer under study.
Indirect indications of such a twisted state can be found in computer simulations [146]. Analyzing Cryo-EM experiments on DNA toroids [147] Hud et al. proposed a rosette-like DNA winding (slowly precessing off-centered loops which yield a "spirograph" motif) [148]. Furthermore, such a non-trivial *local* organization of DNA in toroids suggests natural explanations for various remarkable *large scale* findings from *in vivo* studies: Certain bacteriophages, whose DNA is (due to a genetic modification) no longer end-attached to their nucleocapsid, display unusually strong knotting of the genome [149]. DNA toroids released from giant T4 phages can undergo supercoiling [150], cf. Fig. 5.1. And the chirality of the highly confined sperm-chromatin is surprisingly pronounced, with a pitch 10 times shorter than in vitro [151]. These facts make us wonder whether there is a connection to the topological ripening we will now discuss.

5.2 Nematic Elastic Energy Functional

Let us begin our quantitative analysis of the situation by neglecting the connectivity of the chain. More specifically, we will first formulate a *local* theory which is based on the above mentioned nematic field $\boldsymbol{n}$ of unit tangent vectors [152]. The path of the actual polymer will later be recovered as an integral curve of this flow field, and its global topological properties can then be studied. The elastic energy e per unit volume, describing the deviation from perfectly parallel alignment, is the Frank-Oseen

free energy density of a uniaxial nematic liquid crystal [153]

$$e \;=\; \frac{1}{2}K_1\big[\nabla\cdot\boldsymbol{n}\big]^2+\frac{1}{2}K_2\big[\boldsymbol{n}\cdot(\nabla\times\boldsymbol{n})\big]^2+\frac{1}{2}K_3\big[\boldsymbol{n}\times(\nabla\times\boldsymbol{n})\big]^2\;, \tag{5.1}$$

where the three terms correspond to splay, twist, and bend deformations, respectively[1]. Assuming the condensate to behave like an incompressible liquid (thereby accounting for excluded volume effects in an effective way) the first term (splay) must vanish identically in order to maintain a constant polymer density throughout the condensate[2], while the other two terms divide the elastic energy between themselves. The total energy is of course the integral of Eqn. (5.1) over the torus volume.
It is convenient to treat this situation in suitable toroidal coordinates $\{r,\vartheta,\varphi\}$, defined by

$$x=(r_1-r\cos\vartheta)\cos\varphi\quad,\quad y=(r_1-r\cos\vartheta)\sin\varphi\quad,\quad z=r\sin\vartheta\;. \tag{5.2}$$

The nematic field is now represented as $\boldsymbol{n}=n_r\boldsymbol{e}_r+n_\vartheta\boldsymbol{e}_\vartheta+n_\varphi\boldsymbol{e}_\varphi$ (where the $\boldsymbol{e}_i=\partial_i\boldsymbol{r}/|\partial_i\boldsymbol{r}|$ are the toroidal unit tangent vectors). When inserting this into Eqn. (5.1) and integrating over the volume, we obtain the elastic energy as a functional of $n_r(r,\vartheta,\varphi)$, $n_\vartheta(r,\vartheta,\varphi)$, and $n_\varphi(r,\vartheta,\varphi)$. At the surface of the torus the boundary condition $n_r(r=r_2)=0$ must hold, and owing to rotational symmetry we will henceforth make the (nontrivial but very reasonable) assumption that none of the coordinate functions depends on φ. These steps reduce the task of finding the optimal polymer winding to a two-dimensional variational problem.

5.3 Numerical Minimization

Now that the mathematical problem is formulated, let us first have a look at the full solution, which we obtained numerically via a conjugate gradient minimization [154]. The results confirm the suspicions made above: For large enough aspect ratio $\xi=r_1/r_2$ the equilibrium nematic field is $\boldsymbol{n}=\boldsymbol{e}_\varphi$, corresponding to simple circumferential winding of the polymer. But as the torus grows sufficiently fat, a continuous transition occurs to a state in which (simultaneously throughout the entire torus) $\boldsymbol{n}$ acquires components in $\boldsymbol{e}_\vartheta$- and $\boldsymbol{e}_r$-direction, *i. e.*, the polymer additionally winds around the tubular circle, see Fig. 5.2. This winding relaxes bending energy, but only at the expense of the additional twist, which is zero when $\boldsymbol{n}=\boldsymbol{e}_\varphi$. Consequently, this twist instability occurs more readily if the ratio $\eta=K_2/K_3$ of twist and bend modulus is small. All this is confirmed in Fig. 5.3, where we show the maximum twist angle $\tau\equiv\arccos(n_{\varphi,\min})$, which is a suitable order parameter, as a function of the aspect ratio ξ of the torus. Note that the calculation is reliable down to the value $\xi=1$, where the hole in the torus degenerates to a point; here the only possible tangent vector is $\boldsymbol{e}_\vartheta$ and thus $\tau=\pi/2$.

[1] Here we neglect the usually weak cholesteric interaction [151] stemming from DNA chirality.

[2] In infinitely long polymer liquid crystals without hairpins the splay strongly couples to density variations [152]

Figure 5.2: Illustration of the flow field on a toroidal condensate which features additional twist. The aspect ratio is $\xi = 1.5$.

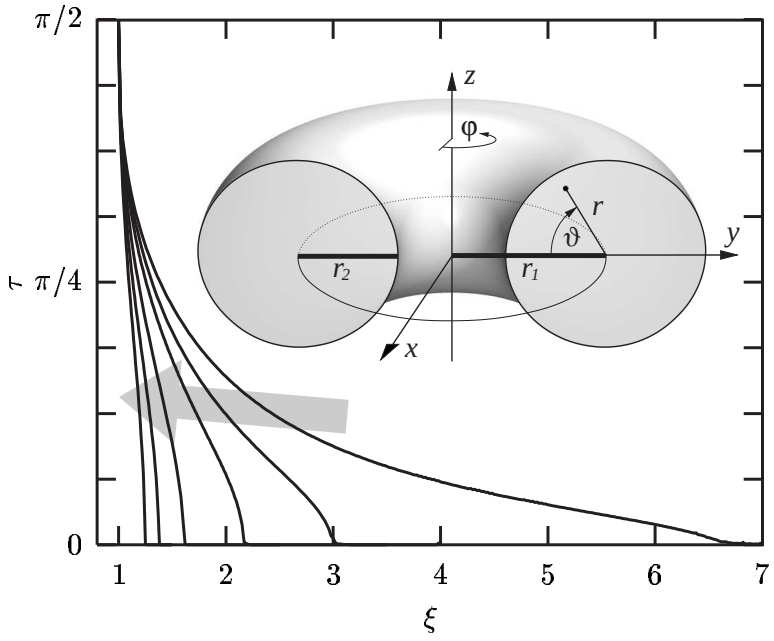


Figure 5.3: Twist order parameter $\tau \equiv \arccos(n_{\varphi,\min})$ as a function of the aspect ratio $\xi = r_1/r_2$. The curves correspond to different ratios of elastic moduli, $\eta = K_2/K_3 \in \{0.01, 0.05, 0.1, 0.2, 0.3, 0.4\}$, the gray arrow pointing toward increasing values. The inset illustrates the toroidal coordinate system used in the text.

5.4 Variational Ansatz

The mathematical task of functional minimization can often be accurately approximated by devising a variational ansatz which is analytically tractable. The following choice turns out to be remarkably good: We will first assume that the nematic field does not have a component in $\boldsymbol{e}_r$-direction. The tangential boundary condition $n_r(r = r_2) = 0$ is then automatically taken care of. The remaining two components must satisfy the normalization condition $n_\vartheta^2 + n_\varphi^2 = 1$, and it thus suffices to specify one of them, say n_ϑ. It is easy to check that any ansatz of the form $n_\vartheta = f(r)/[1 - (r/r_1)\cos\vartheta]$ with an arbitrary function $f(r)$ yields a divergence-free nematic field. We choose a linear $f(r) = \omega r/r_2$, *i. e.*

$$n_\vartheta(r, \vartheta; \omega) \;=\; \omega \, \frac{r/r_2}{1 - (r/r_1)\cos\vartheta} \;, \tag{5.3}$$

where ω, which we may call the "twisting strength", is the only free parameter of the ansatz. With this choice we go back into the Frank-Oseen free energy (5.1), calculate the derivatives, and integrate over the volume of the torus. Since the sign of ω only determines the *handedness* of the twisted structure, it cannot influence the free energy E, which thus must be an even function of ω. An expansion in powers of ω^2 then yields

$$\frac{E}{K_3\, r_2} \;=\; g_0(\xi) + g_2(\xi, \eta)\, \omega^2 + g_4(\xi, \eta)\, \omega^4 + \mathcal{O}(\omega^6) \;, \tag{5.4}$$

where the expansion coefficients g_i are functions of the system parameters[3]. In particular,

$$g_0(\xi) \;=\; 2\pi^2 \Big(\xi - \sqrt{\xi^2 - 1} \Big) \;\stackrel{\xi\to\infty}{\longrightarrow}\; \pi^2/\xi \;. \tag{5.5}$$

This term contributes even if the twist ω vanishes. In fact, it coincides with the bending energy of a chain of length L which has a curvature energy A/ρ^2 per unit length (ρ being the local radius of curvature), and which is wound without additional looping within a torus of volume V—provided the (intuitively clear) relation $K_3 V = AL$ holds. This relates the nematic bending modulus K_3 to the more usual polymer bending stiffness A.

While g_0 helped us to map our parameters, g_2 will localize the phase transition. The reason is that Eqn. (5.4) has the form of a Landau free energy as it occurs for phase transitions with a scalar order parameter, ω, and it predicts a continuous transition (which breaks chiral symmetry) at the point where the coefficient of the quadratic term vanishes, *i. e.*, $g_2(\xi, \eta) = 0$. This results in the phase boundary

$$\eta \;=\; \frac{1}{2} + \frac{\xi^2 - 1}{4\,\xi^2} \Big[1 + 6\,\xi \Big(\sqrt{\xi^2 - 1} - \xi \Big) \Big] \;. \tag{5.6}$$

This boundary is shown in Fig. 5.4, together with exact points originating from the full functional minimization. It quantifies the conclusion that toroidal condensates will

[3]For the particular ansatz (5.3) the resulting coefficient g_4 does in fact not depend on η; but generally it does.

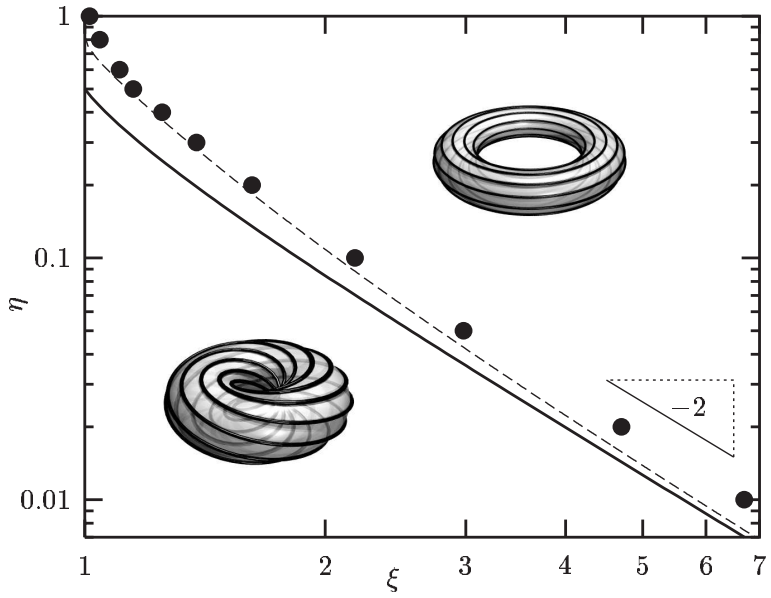


Figure 5.4: Structural phase diagram of the toroidally wound complex on a log-log scale. For small $\xi = r_1/r_2$ and $\eta = K_2/K_3$ the polymer is wound in a twisted way, for large ξ and η it prefers to wind straight. The dots are results from the full numerical minimization, the solid line is Eqn. (5.6), and the dashed line stems from the "improved ansatz" (see text).

spontaneously twist if the aspect ratio $\xi = r_1/r_2$ is small (and the torus thus fat), and if the ratio of elastic moduli $\eta = K_2/K_3$ is small, *i. e.* if twisting a bundle is easy compared to bending it. For large ξ Eqn. (5.6) asymptotically behaves like $\eta \sim \frac{5}{16}\xi^{-2}$, thereby explaining the exponent -2, which is also indicated in Fig. 5.4 and which actually describes the whole phase boundary quite well. The agreement between the simple one-parameter variational ansatz and the full calculation is remarkably good. It can even be improved by including an additional prefactor $1 - \frac{r}{r_1}$ into Eqn. (5.3). Particularly for small ξ this improved ansatz agrees better with the exact answer, basically since the new prefactor cancels the unphysical divergence of the denominator for $r \to r_1$ at $\vartheta = 0$. The analytical expression for the phase boundary is quite involved and will not be shown here, but it has the same large ξ asymptotics (see Fig. 5.4).

Both the ansatz as well as the full numerical solution point to an upper critical ratio of elastic moduli, η_c, beyond which the system will no longer spontaneously twist. Even at the lowest possible aspect ratio $\xi = 1$ the energetic cost for twisting has become so large that it no longer admits a bending relaxation. The ansatz (5.3) gives $\eta_c = \frac{1}{2}$, the improved ansatz gives $\eta_c \approx 0.829$, while the full numerical solution suggests the deceptively simple result $\eta_c = 1$. We have no analytical support for the latter, but we also note that the limit $\xi \to 1$ is somewhat academic, because our tacit assumption that the condensate shape is strictly toroidal will most likely break down in this case [155].

We have thus seen that the ratio η and the torus geometry ξ uniquely specify the twist-

state of the condensate. However, while one can easily measure ξ in an experiment (just by visual inspection), it is hard to specify in advance. In contrast to that, the surface tension σ can be readily changed (for instance via the concentration of condensing agents), but its actual value is hard to measure. In our simple model it is of course not difficult to add a tension term σ times the torus surface S to the condensate energy. Using $S \propto \xi^{1/3} V^{2/3}$, one can re-express the twisting transition as being driven by increasing σ, and it remains continuous. However, practical considerations suggest that one tunes σ only for the purpose of modifying ξ, but subsequently use ξ as the independent variable. This way one needs no longer (neither theoretically nor practically) worry about how a particular concentration of condensing agents gives rise to a particular torus geometry.

5.5 Global Aspects

After these local considerations it is time to study global aspects of the polymer structure, *i. e.*, consequences of the chain *connectivity* which we have neglected so far. Let us start with the flow itself. It can be shown that incompressibility, $\nabla \cdot \boldsymbol{n} = 0$, together with axial symmetry causes the flow to be Hamiltonian—hence one more conservation law exists[4]. In our case it forces the flow lines to stay on two-dimensional slightly deformed toroidal surfaces, such that the total flow consists of a nested structure of invariant tori. In fact, our ansatz (5.3) follows readily from the quadratic Hamiltonian $H = \frac{1}{2}\xi\omega r^2$, which is constant on circular tubular layers. Of course, the actual polymer has to switch between these layers, reminding us that irrespective of twist none of the above structures can be realized without localized defects [156], even if the underlying nematic flow field is everywhere smooth.

There is one global aspect of the polymer structure in which it differs fundamentally from a plainly wound torus: As is visible in Fig. 5.2, the path of the polymer threads it repeatedly through the middle hole. Moreover, the amount of this looping (as measured *e. g.* by the average change in ϑ per turn) depends on the layer. This effect implies that the entire polymeric strand is heavily entangled with itself. A rough estimate for $\xi = 1.5$ and polymer length $L = 15\mu$m gives about 30 threadings through the hole. In fact, were it not for the two free ends, these knotted states would be topologically inaccessible. In other words, unlike the initial collapse, the second stage, the structural ripening, relies on the motion of the free chain ends and is thus a much slower process. On the other hand this structural ripening meets no kinetic barriers during the relaxation to its twisted ground state, as it proceeds downhill on the free energy landscape. The motion of the two free ends is then energetically directed and their local rearrangement does not involve the highly improbable threading through the toroid hole in 3D space. In addition, the weak chiral interaction of DNA molecules [151] neglected above, gives rise to a (small) chiral term in the elastic free energy, which might contribute to

[4]The following transformation of the field and coordinates given in a cylindrical coordinate system, $\{\rho, z\} \to \{\rho^2/2, z\}$, $\{n_\rho, n_z\} \to \{\rho n_\rho, n_z\}$, together with the axial symmetry, reduces the 3D divergence-free nematic field to the (new) 2D divergence-free field. Then the director field can be written as $n_\rho = \rho^{-1}\partial H/\partial z$, $n_z = -\rho^{-1}\partial H/\partial \rho$, where $H(\rho, z)$ is the Hamiltonian.

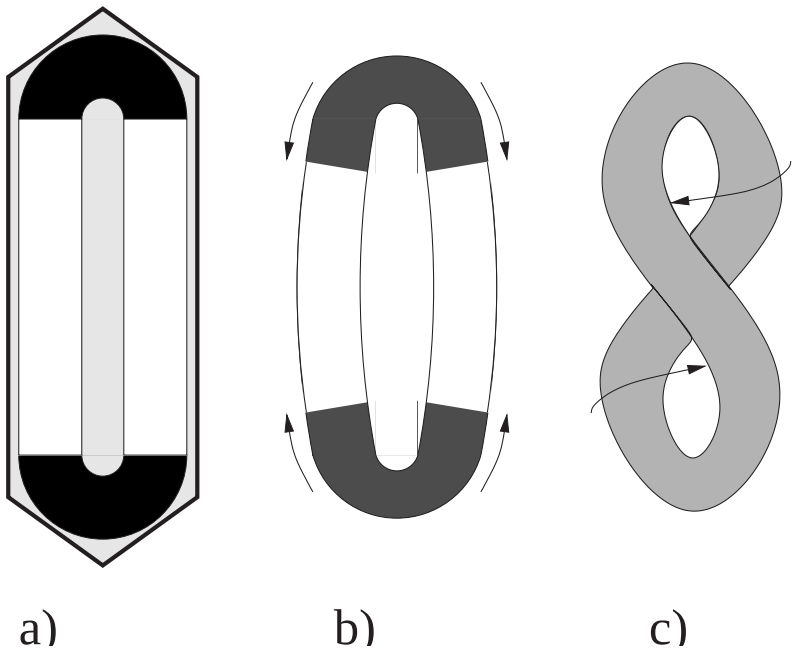


Figure 5.5: Possible mechanism for plectonemic supercoiling [150] of the genome of giant T4 phages: (a) Initially the toroidal genome is only twisted (dark shading) at the two poles. (b) After removing the capsid, twist propagates into the remaining DNA, but since the twist-creating confinement is removed, the remaining bundle suddenly is overtwisted. (c) This then induces a global supercoiling.

the symmetry breaking and "guide" the twisting in a preferred direction.

5.6 Discussion

Is the predicted effect strong enough to be of some relevance for DNA condensation? For typical experimental parameters of DNA length $L = 15 - 30\mu$m, $\xi \sim 1.5 - 2.5$, bending stiffness $A \sim 50 k_{\rm B}T{\cdot}$nm, and inter-helical distance $d \sim 3$nm we obtain first the elastic constant $K_3 \sim A/d^2 \sim 20$pN, which is dominated by the bending stiffness [152]. The twist constant K_2 can be estimated by the decondensation force ~ 2pN obtained in single molecule experiments with condenser spermidine [157]. Then the difference in elastic energy between the twisted and untwisted states, as bounded below by the the variational ansatz, lies in the range $15 - 30 k_{\rm B}T$. This indicates that topological ripening stabilizes the condensate. In addition, if the solvent quality abruptly improves, the twisted toroid will unfold more inertly than its untwisted counterpart, due to heavy entanglement with itself. If this stabilization occurs on the typical time scales relevant for gene therapeutical applications, it might prevent a premature digestion of the genetic material by the host organism and influence (positively) the efficiency of the gene delivery process.

Finally, it is tempting to speculate that the twist-bend instability is responsible for

several puzzling phenomena observed in biology. For instance, sperm chromatin shows an unusually short DNA cholesteric pitch (10 times shorter than *in vitro*) [151]. Since the genome is highly confined in sperm cells, basically at the same density as the DNA tori discussed above, it is potentially prone to a twist-bend instability (recall that all one really needs is an externally imposed bending of a bundle of DNA). While the weak intrinsic DNA chirality might thus determine the *handedness* of the resulting twist, its *pitch* would then be given by the twisted state of polymer strands after the topological ripening took place. As a second example, if the capsid of bacterial phages is broken (for instance by osmotic shock), their DNA spills onto the underlying substrate. The amount of "knotting" displayed by the freed genome is significantly larger for genetically modified phages, in which the end of the viral genome is not first attached to the capsid at the time of loading it [149]. Since the ends are free to move, the torus- or spool-like genome inside such mutant phages can again undergo topological ripening, explaining the enhanced entanglement of DNA after breaking the capsid. Finally, Earnshaw et. al [150] have shown that when mutant T4 bacteriophages (displaying a greatly enhanced capsid aspect ratio) are opened under conditions in which the toroidally wound-up genome remains condensed, structures resembling "twisted skeins of yarn" are sometimes observed (see Figs. 5 and 7 in Ref. [150]), cf. Fig. 5.1. Within the framework of the twist-bend instability discussed in the present chapter, the following explanation is tempting (see Fig. 5.5): Inside the elongated phage head the genome is present as a torus which is squeezed flat, and only at the two "poles" curvature exists which promotes the creation of twist. After opening the capsid, the torus can resume a properly round shape. Its rather large aspect ratio no longer necessitates a twist, but for kinetic reasons it cannot instantaneously get rid of it. However, it can rather quickly exchange twist for *writhe* [158] and thus undergo *plectonemic supercoiling*, a state which indeed characterizes the observed structures rather accurately. Since supercoiling inherits its handedness from the underlying twist, which itself emerged via spontaneous breaking of chiral symmetry, both right- and left-handed supercoiled states should exist, which indeed appears to be borne out by observation [150].

Bibliography

[1] Frenkel-Bresler theory is reviewed in L.D. Landau and E.M. Lifshitz Statistical Mechanics,Pergamon Press, Oxford (1996). O. Kratky and G. Porod, Recl. Trav. Chim. Pays-Bas 68, 1106 (1949)

[2] T. R. Strick et al. Rep. Prog. Phys. **66**, 1 (2003)

[3] C. Jarzynski, Phys Rev. Lett. 78, 2690 (1997)

[4] L.D. Landau and E.M. Lifshitz Theory of Elasticity ,Pergamon Press, Oxford (1996).

[5] G. Kirchhoff, J. Reine Angew. Math. (Crelle) 56, 285-313 (1859).

[6] Nizette M. and A. Goriely . Towards a classification of Euler-Kirchhoff filaments. J. Math. Phys. 40,2830 (1999).

[7] C. Li, N. Gunari, K. Fischer, A. Janshoff and M. Schmidt, Angew. Chem. Int. Ed. 2004, 43.

[8] N. Gunari, A. Janshoff and M. Schmidt private communication.

[9] L.D. Landau and E.M. Lifshitz Statistical Physics, ,Pergamon Press, Oxford (1996).

[10] J. F. Marko and E. D. Siggia Stretching DNA, Macromolecules 28, 8759 (1995).

[11] T. Odijk, Stiff Chains and Filaments under Tension, Macromolecules 28, 7016 (1995).

[12] T. Odijk, Microfibral Buckling within Fibers under Compression, J. Chem. Phys. 108, 6923 (1998).

[13] L.S. Schulman, Techniques and Applications of Path Integration, Wiley Classics Library (1996).

[14] H. Kleinert, Path Integrals, World Scientific Publishing Co., Singapore (2002)

[15] J.S. Langer, Ann. Phys. 41, 108 (1967)

[16] M. Büttiger and R. Landauer, Phys. Rev. A, 23, 1397 (1981).

[17] I.M. Gelfand and A.M. Yaglom, J. Math. Phys. 1, 48 (1960).

[18] F.M. Arscott, Periodic Differential Equations, Pergamon Press (1964).

[19] E. Kamke, Differential Equations, Methods and Solutions, Leipzig (1959).

[20] G. Barton, Elements of Green's Functions and Propagation, Oxford Science Publications (1989)

[21] C. J. Bustamante et al. Entropic Elasticity of λ-Phage DNA. Science 265, 1599 (1994).

[22] P. G. de Gennes, Scaling Concepts in Polymer Physics, Ithaca, Cornell University Press (1979).

[23] M. Doi and S.F. Edwards, Theory of Polymer Dynamics, Oxford University Press, (1988).

[24] P. J. Park and W. Sung, Dynamics of a polymer surmounting a potential barrier: The Kramers problem for polymers, J. Chem. Phys. 111, 5259 (1999). K. L. Sebastian and A. K. R. Paul, Kramers problem for a polymer in a double well, Phys. Rev. E 62, 927-939 (2000).

[25] H. Yamakawa, Helical Wormlike Chains in Polymer Solutions. Springer Berlin, (1997).

[26] C. Bouchiat and M Mezard, Phys. Rev. Lett. 80,1556 (1998); Eur.Phys. J. E 2,377 (2000); J.D. Moroz and P. Nelson Macromol. 31, 6333 (1998).

[27] T. Odijk, Macromolecules 19, 2313 (1986).

[28] R. Bruinsma and J. Rudnick, DNA-Protein Cooperative Binding through Variable-Range Elastic Coupling, Biophys. J. 76, 1725, (1999)

[29] A.S. Davydov, Solitons in Molecular Systems. Dordrecht, Netherlands: Kluwer (1985)

[30] L.D. Landau and E.M. Lifshitz,Quantum Mechanics (1),Pergamon Press, Oxford (1996).

[31] T. Odijk, DNA in Liquid-Crystalline Environment: Tight Bends, Rings and Supercoils, J. Chem. Phys. 105, 1270 (1996)

[32] G. Maret et al., Orientation of DNA in High Magnetic Fields, Phys. Rev. Lett. 35, 397 (1975)

[33] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, Nature (London) **389**, 251 (1997).

[34] K. J. Polach and J. Widom, J. Mol. Biol. **254**, 130 (1995); **258**, 800 (1996).

[35] B. D. Brower-Toland, C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, and M. D. Wang, Proc. Natl. Acad. Sci. USA **99**, 1960 (2002).

[36] P. Beard, Cell **15**, 955 (1978); C. Spadafora, P. Oudet, and P. Chambon Eur. J. Biochem. **100**, 225 (1979).

[37] S. Pennings, G. Meersseman, and E. M. Bradbury, J. Mol. Biol. **220**, 101 (1991); Proc. Natl. Acad. Sci. USA **91**, 10275 (1994); G. Meersseman, S. Pennings, and E. M. Bradbury, EMBO J. **11**, 2951 (1992).

[38] A. Flaus and T. J. Richmond, J. Mol. Biol. **275**, 427 (1998); J. M. Gottesfeld, J. M. Belitsky, C. Melander, P. B. Dervan, and K. Luger, J. Mol. Biol. **321**, 249 (2002).

[39] R. L. Ornstein, R. Rein, D. L. Breen and R. D. Macelroy, Biopolymers **17**, 2341 (1978)

[40] H. Schiessel, J. Phys.: Cond. Mat. **15**, R699 (2003)

[41] H. Schiessel, J. Widom, R. F. Bruinsma, and W. M. Gelbart, Phys. Rev. Lett. **86**, 4414 (2001); **88**, 129902 (2002).

[42] I. M. Kulić and H. Schiessel, Biophys. J. **84** 3197 (2003).

[43] J. Frenkel and T. Kontorova, Zh. Eksp. Teor. Fiz. **8**, 1340 (1938); 1939 *J. Phys. (Moscow)* **1**, 137 (1939).

[44] L. M. Floria and J. J. Mazo, Adv. Phys. **45**, 505 (1996).

[45] O. M. Braun, T. Dauxois, M. V. Paliy, and M. Peyrard, Phys. Rev. Lett. **78**, 1295 (1997).

[46] T. Strunz and F.-J. Elmer, *Phys. Rev. E* **58**, 1601 (1998).

[47] L. Pietronero, W. R. Schneider, and S. Strässler, Phys. Rev. B **24**, 2187 (1981).

[48] S. Aubry, J. Phys. (France) **44**, 147 (1983).

[49] S. Watanabe, H. S. J. van der Zant, S. H. Strogatz, and T. P. Orlando, Physica D **97**, 429 (1996).

[50] W. Uhler and R. Schilling, Phys. Rev. B **37**, 5787 (1988).

[51] H. Schiessel, G. Oshanin, A. M. Cazabat and, M. Moreau, Phys. Rev. E **66**, 056130 (2002).

[52] R. D. Kamien, T. V. Lubensky, P. Nelson, and C. S. O'Hern, Europhys. Lett. **38**, 237 (1997); J. F. Marko, Europhys. Lett. **38**, 183 (1997).

[53] C. Levinthal and H. Crane, Proc. Natl. Acad. Sci. USA **42**, 436 (1956);P. Nelson, Proc. Natl. Acad. Sci. USA **96**, 14342 (1999).

[54] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[55] S. C. Satchwell, H. R. Drew, and A. A. Travers, J. Mol. Biol. **191**, 659 (1986).

[56] T. E. Shrader and D. M. Crothers, Proc. Natl. Acad. Sci. USA **86**, 7418 (1989); T. E. Shrader and D. M. Crothers, J. Mol. Biol. **216**, 69 (1990).

[57] S. Cacchione, M. A. Cerone, and M. Savino, FEBS Lett. **400**, 37 (1997).

[58] S. Mattei, B. Sampaolese, P. De Santis, and M. Savino, Biophys. Chem. **97**, 173 (2002); C. Anselmi et al., Biophys. J. **79**, 601 (2000).

[59] Y. Lorch, M. Zhang, and R. D. Kornberg, Cell **96**, 389 (1999).

[60] Widom J. 1998. Structure, dynamics, and function of chromatin in vitro. Annu Rev Biophys Biomol Struct. 27:285-327

[61] Kornberg R.D. and Y. Lorch. 1999. Twenty-five years of the nucleosome, fundamental particle of the eucaryote chromosome. Cell 98:285

[62] Wolffe A. Chromatin: Structure and Dynamics (3.ed). Academic Press, ISBN 0127619151 (1999)

[63] Peterson C.L. 2000. ATP-dependent chromatin remodeling: going mobile. FEBS Lett. 476 : 68

[64] Vignali M., A.H. Hassan , K.E. Neely and J.L. Workman. 2000. ATP-dependent chromatin-remodeling complexes. Mol. Cell. Biol. 20: 1899

[65] Beard P. 1978. Mobility of histones on chromosome of simian virus-40. Cell 15:955-967

[66] Spadafora C., P. Oudet and P. Chambon. 1979. Rearrangement of chromatin structure induced by increasing ionic-strength and temperature. Eur. J. Bioch. 100: 225-235

[67] Meersseman G., S. Pennings and E.M. Bradbury. 1992. Mobile nucleosomes–a general behavior. EMBO J. 11:2951-9.

[68] Pennings S., G. Meersseman and E.M. Bradbury. 1991. Mobility of positioned nucleosomes on 5 S rDNA. J. Mol. Biol. 220:101-10.

[69] Schiessel H., J. Widom, R.F. Bruinsma and W.M. Gelbart. 2001. Polymer reptation and nucleosome repositioning. Phys. Rev. Lett. 86: 4414-4417; Phys. Rev. Lett. 88 129902

[70] Polach K.J. and J. Widom. 1995. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. J. Mol. Biol. 254:130-49.

[71] Anderson J.D. and J. Widom. 2000. Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. J. Mol. Biol. 296:979-87.

[72] S. Mangenot, private communication.

[73] Luger K., A.W. Mader ,R.K. Richmond, D.F. Sargent and T.J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Åresolution. Nature 389:251-60.

[74] Hagerman P.J. 1988. Flexibility of DNA. Ann. Rev. Biophys. Biophys. Chem. 17:265

[75] Benham C.J. 1977 . Elastic model of supercoiling. Proc. Natl. Acad. Sci. U S A. 74:2397-401.

[76] Benham C.J. 1979. An elastic model of the large-scale structure of duplex DNA. Biopolymers 18:609-23.

[77] Le Bret M. 1979. Catastrophic variation of twist and writhing of circular DNAs with constraint? Biopolymers 18:1709-25.

[78] Le Bret M. 1984. Twist and writhing in short circular DNAs according to first-order elasticity. Biopolymers 23:1835-67.

[79] Swigon D., B.D. Coleman B.D. and I. Tobias. 1998. The elastic rod model for DNA and its application to the tertiary structure of DNA minicircles in mononucleosomes. Biophys. J. 74:2515-30.

[80] Shi Y.M. and J.E. Hearst. 1994. The Kirchhoff elastic rod, the nonlinear Schrödinger-equation, and DNA supercoiling. J. Chem. Phys. 101:5186-5200

[81] Fain B. and J. Rudnick. 1997. Conformations of linear DNA. Phys. Rev. E 55:7364

[82] Fain B. and J. Rudnick J. 1999. Conformations of closed DNA. Phys. Rev. E 60:7239

[83] Schiessel H. , J. Rudnick, R. Bruinsma and W.M. Gelbart. 2000. Organized condensation of worm-like chains. Europhys. Lett. 51:237-243

[84] Nizette M. and A. Goriely 1999. Towards a classification of Euler-Kirchhoff filaments. J. Math. Phys. 40:2830

[85] Abramowitz M. and I. Stegun. Handbook of Mathematical Functions with Formulas, Graphs and Tables. US Gov. Pr. Off. (1972)

[86] Coleman B.D., D. Swigon and I. Tobias. 2000. Elastic stability of DNA configurations. Phys. Rev. E 61:747-70.

[87] Yamakawa H. Helical Wormlike Chains in Polymer Solutions. Springer New York, ISBN 3540629602 (1997)

[88] Yamakawa H. and W.H. Stockmayer. 1972. Statistical Mechanics of Wormlike Chains II. J. Chem. Phys. 57, 2843

[89] Bouchaud J.P. and A. Georges. 1990. Anomolous diffusion in disordered media - statistical mechanics, models and physical applications. Phys. Rep. 195:127

[90] Klafter J., M.F. Shlesinger and G.Zumofen. 1993. Non-Brownian transport in complex systems. Chem. Phys. 177:821

[91] Sokolov I.M., J.Mai and A.Blumen. 1997. Paradoxical diffusion in chemical space for nearest-neighbor walks over polymer chains. Phys. Rev. Lett. 79:857

[92] de Cloizeaux J. and G. Jannick. Polymers in Solution. Clarendon Press, Oxford.1990.

[93] Längst G. and P.B. Becker. 2001. ISWI induces nucleosome sliding on nicked DNA. Mol. Cell 8:1085-1092

[94] Bazett-Jones D.P., J. Cote, C.C. Landel, C.L. Peterson and J.L. Workman. 1999. The SWI/SNF complex creates loop domains in DNA and polynucleosome arrayes and can disrupt DNA-histone contacts within these domains. Mol. Cell. Biol. 19: 1470

[95] Felsenfeld G. , D. Clark and V. Studitsky. 2000. Transcription through nucleosomes. Biophys. Chem. 86:231

[96] Rees A.R., R.W. Keller, J.P. Vesenka, G. Yang and C. Bustamante. 1993. Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. Science 260:1646

[97] Rivetti C., M. Guthold and C. Bustamante. 1999. Wrapping of DNA around the E.coli RNA polymerase open promoter complex. EMBO J. 18:4464

[98] Schulz A., N. Mucke, J. Langowski and K. Rippe. 1998. Scanning force microscopy of E.coli RNA Polymerase σ^{45} holoenzyme complexes with DNA in buffer and air. J. Mol. Biol. 283:821

[99] Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal Structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251-260.

[100] Workman, J. L. & Kingston. R. E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* **67**, 545-579.

[101] Polach, K. J. & Widom, J. (1995). Mechanism of protein access to specific DNA sequences in chromatin: A dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254**, 130-149.

[102] Anderson, J. D. & Widom, J. (2000). Sequence and position dependence of the equilibrium accessibility of nucleosomal DNA target sites. *J. Mol. Biol.* **296**, 279-287.

[103] Pennings, S., Meersseman, G. & Bradbury, E. M. (1991). Mobility of positioned nucleosomes on 5S rDNA. *J. Mol. Biol.* **220**, 101-110.

[104] Meersseman, G., Pennings, S. & Bradbury, E. M. (1992). Mobile nucleosomes-a general behavior. *EMBO J.* **11**, 2951-2959.

[105] Pennings, S., Meersseman, G. & Bradbury, E. M. (1994). Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proc. Natl. Acad. Sci.* **91**, 10275-10279.

[106] Flaus, A. & Richmond, T. J. (1998). *J. Mol. Biol.* **275**, 427.

[107] Kornberg, R. D. & Lorch, Y. (1999). Twenty-five years of the nuclesome, fundamental particle of the eukaryote chromosome. *Cell*, **98**, 285-294.

[108] Becker, P. B. (2002). Nucleosome sliding: facts and fiction. *EMBO J.* **21**, 4749-4753.

[109] Schiessel, H. (2003). The physics of chromatin. *J. Phys.: Condens. Matter*, **15**, R699-R774.

[110] Flaus, A. & Owen-Hughes, T. (2003). Mechanisms for nucleosome mobilization. *Biopolymers*, **68**, 563-578.

[111] Kulić, I. & Schiessel, H. (2003). Nucleosome repositioning via loop formation. *Biophys. J.* **84**, 3197-3211.

[112] Kulić, I. & Schiessel, H. (2003). Chromatin dynamics: nucleosomes go mobile through twist defects. *Phys. Rev. Lett.* **91**, 148103.

[113] Anselmi, C., Bocchinfuso, G., De Santis, P., Savino, M. & Scipioni, A. (2000). A theoretical model for the prediction of sequence-dependent nucleosome thermodynamic stability. *Biophys. J.* **79**, 601-613.

[114] Mattei, S., Sampaolese, B., De Santis, P. & Savino, M. (2002). *Biophys. Chem.* **97**, 173.

[115] Gottesfeld, J. M., Belitsky, J. M., Melander, C., Dervan, P. B. & Luger, K. (2002). Blocking transcription through a nucleosome with synthetic DNA ligands. *J. Mol. Biol.* **321**, 249-263.

[116] Suto, R. K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B. & Luger, K. (2003). Crystal structures of the nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* **326**, 271-280.

[117] Studitsky, V. M., Clark, D. J. & Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template *Cell*, **76**, 371-382.

[118] Studitsky, V. M., Clark, D. J. & Felsenfeld, G. (1995). Overcoming a nucleosomal barrier to transcription. *Cell*, **83**, 19-27.

[119] Studitsky, V. M., Kassavetis, G. A., Geiduschek, E. P. & Felsenfeld, G. (1997). Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. *Science*, **278**, 1960-1963.

[120] Bednar, J., Studitsky, V. M., Gregoryev, S. A., Felsenfeld, G. & Woodcock, C. L. (1999). The nature of the nucleosomal barrier to transcription: Direct observation of paused intermediates by electron cryomicroscopy. *Molecular Cell*, **4**, 377-386.

[121] Hänggi, P., Talkner, P. & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.*, **62**, 251-341.

[122] Reviewed in R.M Saecker and M.T. Record, Curr. Op. Struct. Biol. **12**, 311 (2002)

[123] O.V. Tsodikov et al., J. Mol. Biol. **294**, 639 (1999)

[124] S.C. Kampranis, A.d. Bates and A. Maxwell, Proc. Natl. Acad . Sci. USA **96**, 8414 (1999)

[125] C. Rivetti, M. Guthold and C. Bustamante, EMBO J. **18**, 4464 (1999)

[126] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, Nature (London) **389**, 251 (1997).

[127] K.B. Thurmond, E.E. Remsen, T. Kowalewski and K.L. Wooley, Nucl. Acids Res. **27**, 2966 (1999)

[128] Reviewed in V.A. Bloomfield, Curr. Op. Struct. Biol. **6**, 334 (1996); Biopolymers. **44**: 269 (1997).

[129] T. R. Strick et al. Rep. Prog. Phys. **66**, 1 (2003)

[130] M.C. Wiliams et al. Proc. Natl. Acad. Sci. USA **98**, 6121 (2001); M. Hegner, S. B. Smith and C. Bustamante, Proc. Natl. Acad. Sci. USA **96**, 10109 (1999);

[131] C. G. Baumann et al. Biophys. J. **78**, 1965 (2000);

[132] Y. Murayama, Y. Sakamaki and M. Sano. Phys. Rev. Lett. **90**, 018102 (2003)

[133] K. Svoboda and S.M. Block Cell **77**, 773 (1994); H. Yin et al. Science **270**, 1653 (1995)

[134] E. Evans and K. Ritchie, Biophys J. **72**, 1541 (1997)

[135] E. Evans, Biophys. Chem. **82,** 83 (1999)

[136] B. D. Brower-Toland et al. Proc. Natl. Acad. Sci. USA **99**, 1960 (2002).

[137] Y. Cui and C. Bustamante Proc. Natl. Acad. Sci. USA **97**, 127 (2002).

[138] Love A E, *A treatise of the Mathematical Theory of Elasticity* 4th edn (Dover, New York, 1944).

[139] G.C. Cecchi and M.O. Magnasco, Phys. Rev. Lett. **76**, 1968 (1996); D. Bartolo, I. Derényi and A. Ajdari, Phys. Rev. E, **65**, 051910 (2002)

[140] K. J. Polach and J. Widom, J. Mol. Biol. **254**, 130 (1995); **258**, 800 (1996).

[141] J. F. Marko and E. D. Siggia, Biophys. J. **73**, 2173 (1997); K.-K. Kunze and R. R. Netz, Phys. Rev. E **66**, 011918 (2002)

[142] A. Y. Grosberg and A. R. Khokhlov, *Statistical physics of macromolecules, AIP Press* (1994); I. M. Lifshitz, A. Y. Grosberg, and A. R. Khokhlov, *Rev. Mod. Phys.* **50**, 683 (1978).

[143] A. Y. Grosberg, Biofizika, **24**, 32 (1979); A. Yu. Grosberg and A. R. Khokhlov, *Adv. Polym. Sci.*, **41**, 53, (1981).

[144] V. A. Bloomfield, *Curr. Op. Struct. Biol.* **6**, 334 (1996); *Biopolymers* **44**, 269 (1997).

[145] A convex function $f(x)$ lies above any of its tangents. Constructing the tangent at $x = \langle x \rangle$ shows that $f(x) \geq f(\langle x \rangle) + (x - \langle x \rangle) f'(\langle x \rangle)$. Averaging yields Jensen's inequality $\langle f(x) \rangle \geq f(\langle x \rangle)$. For concave f the inequality sign reverses. J. L. W. V. Jensen, Acta Math. **30**, 175 (1906).

[146] M. J. Stevens, *Biophysical J.* **80**, 130 (2001); M. R. Stukan et al., *J. Chem. Phys.* **118**, 3392 (2003)

[147] K. A. Marx and G. C. Ruben, *Nucl. Acids Res.* **11**, 1839 (1983).

[148] N. V. Hud, K. H. Downing, and R. Balhorn, *Proc. Natl. Acad. Sci. USA* **92**, 3581 (1995);

[149] L. F. Liu et al. *Proc. Natl. Acad. Sci. USA* **78**, 5498 (1981); J. Arsuaga et al. ibid. **99**, 5373 (2002).

[150] W. C. Earnshaw, J. King, S. C. Harrison, and F. A. Eiserling, Cell **14**, 559 (1978).

[151] F. Livolant, *Physica A* **176**, 117 (1991); F. Livolant and A. Leforestier, *Prog. Polym. Sci.* **21**, 1115 (1996)

[152] P. G. de Gennes, *Mol. Cryst. Liq. Cryst. Lett.* **34**, 177 (1977).; R. Meyer in *Polymer Liquid Crystals*, A. Cifferi et. al. (eds.), Academic NY (1982); H. H. Strey, V. A. Parsegian, and R. Podgornik, *Phys. Rev. E* **59**, 999 (1999).

[153] P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, 2^{nd} ed., Clarendon Press, Oxford (1995).

[154] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Fortran*, 2^{nd} ed., Cambridge University Press, Cambridge (1992).

[155] J. Ubbink and T. Odijk, *Europhys. Lett.* **33**, 353 (1996).

[156] For hexagonal-crystalline toroids this effect was pointed out by S. Y. Park, D. Harries, and W. M. Gelbart, *Biophys. J.* **75**, 714 (1998).

[157] C. G. Baumann et al. *Biophys. J.* **78**, 1965 (2000); Y. Murayama, Y. Sakamaki, and M. Sano, *Phys. Rev. Lett.* **90**, 018102 (2003)

[158] F. B. Fuller, Proc. Nat. Acad. Sci. USA **68**, 815 (1971).